FEDERAL UNIVERSITY OF BAHIA

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

POSTGRADUATE PROGRAM IN ELECTRICAL AND COMPUTER ENGINEERING

MARTON SANDES DOS SANTOS

Ph.D Thesis

COMPUTATIONAL INTELLIGENCE APPLIED TO CROSSTALK MITIGATION IN A LIQUID ARGON CALORIMETER

Salvador

2024

MARTON SANDES DOS SANTOS

Computational intelligence applied to crosstalk mitigation in a liquid argon calorimeter

Ph.D Thesis

Text of the Ph.D. thesis presented to the Postgraduate Program in Electrical and Computer Engineering at the Federal University of Bahia as one of the requirements for obtaining the degree of Doctor in Electrical Engineering.

Supervisors:

Prof. Dr. Eduardo F. de Simas Filho – UFBA
Prof. Dr. Paulo César M. de A. Farias – UFBA
Prof. Dr. Bertrand Laforge – Sorbonne Université

Salvador

2024

S237 Santos, Marton Sandes
Computational intelligence applied to crosstalk mitigation in a liquid argon calorimeter/ Marton Sandes dos Santos. – Salvador, 2024-
134 p. : il. color;

Orientador: Prof. Dr. Eduardo F. de Simas Filho.
Orientador: Prof. Dr. Paulo César M. de A. Farias.
Orientador: Prof. Dr. Bertrand Laforge.

Tese (doutorado) – Programa de Pós-graduação em Engenharia Elétrica e de Computação – Universidade Federal da Bahia, Escola Politécnica, 2024.

1. ATLAS Detector. 2. Analytical Simulator. 3. Crosstalk Mitigation. 4. Statistical Signal Processing. 5. Machine Learning. 6. Lorenzetti Framework.
I. Simas Filho, Eduardo F. II. Farias, Paulo César M. de A. III. Laforge, Bertrand. IV. Universidade Federal da Bahia. V. Título.

CDD: 333.79

Marton Sandes dos Santos

Computational Intelligence applied to Crosstalk Mitigation in a Liquid Argon Calorimeter

Ph.D. Thesis

Text of the Ph.D. thesis presented to the Postgraduate Program in Electrical and Computer Engineering at the Federal University of Bahia as one of the requirements for obtaining the degree of Doctor in Electrical Engineering.

Ph.D. Thesis. Salvador, December 11 of 2024:

Prof. Dr. Eduardo F. de Simas Filho – PPGEEC/UFBA
Advisor

Prof. Dr. Paulo César M. de A. Farias –
PPGEEC/UFBA Advisor

Prof. Dr. Bertrand Laforge – LPNHE/Sorbonne
Université Advisor

Documento assinado digitalmente
gov.br ANTONIO CARLOS LOPES FERNANDES JUNIOR
Data: 11/02/2025 21:58:54-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. Antonio C. L. Fernandes Jr. –
PPGEEC/UFBA

Prof. Dr. Bernardo Sotto-Maior Peralva – UERJ

Documento assinado digitalmente
gov.br EDUARDO DA SILVA ALMEIDA
Data: 17/02/2025 09:40:49-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. Eduardo da Silva Almeida – IF/UFBA

Salvador

2024

I dedicate this work to my parents,
who are great supporters and mentors.

Acknowledgements

In everything give thanks: for this is the will
of God in Christ Jesus concerning you.
(1 Thessalonians 5:18)

Render therefore to all their dues:
tribute to whom tribute is due;
custom to whom custom;
fear to whom fear; honor to whom honor.
(Romans 13:7)

This Ph.D journey is coming to an end, and I have so many people to thank; I'm sure I'll miss mentioning some names. First, I thank God for His support and for allowing me to meet everyone who contributed to this work over these years.

Thanks to my mom and dad (*in memoriam*) for teaching me values and principles. To my wife for her unwavering support and patience along this path, and to God once again for my sons, who were born during my first year in France, and the fifth year of this Ph.D. This story began during my final term in electrical engineering when I started working with Professor Simas. Soon after, I met Edmar, who was presenting a seminar on neural networks in a course taught by Professor Paulo César. Thank you, for your countless times of support. From then on, I embarked on my challenging journey in high-energy physics, a field that demands high energy from the beginning, to persist, and to grasp its pursuit of answers to some of nature's most intriguing questions.

Thanks to my colleagues from our "bat-cave" at UFBA, for the many enriching discussions and the exchange of insights across diverse research areas. I can't count how many times a single conversation saved time or helped avoid simple errors in signal processing methods or algorithms.

In the ATLAS Brazil collaboration, I had the privilege of participating in workshops and learning from dedicated researchers from various fields. Thanks to Werner for his support and for answering my many questions with patience and expertise.

At Sorbonne, I was fortunate to work alongside some of the many scientists who are part of the ATLAS Collaboration team. I'm grateful to my colleagues at the LPNHE laboratory for the discussions about physics, where I listened intently, always learning a bit more. Thanks to Mykola for our discussions, both online and in person when he came to Salvador, and for his contributions to the EMshower simulator. Special thanks go to Bertrand for his support, patience, attentiveness, and the many rich discussions that helped guide this research forward, always pushing to find the right answers.

Today, I can echo the words of Sir Isaac Newton: *"If I have seen further, it is by standing on the shoulders of giants"*. This is precisely how I feel about my advisors, Professor Simas, Professor Paulo César, Professor Seixas (ATLAS Brazil coordinator), and Professor Bertrand Laforge.

The author would like to thank the financial support of FAPESB, FAPERJ, RENAFAE, CNPq, and CAPES, in Brazil, and COFECUB and CNRS in France. This study was financed in part by the

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

"Neither is there any creature that
is not manifest in his sight: but
all things are naked and opened
unto the eyes of him
with whom we have to do".
(Hebrews 4:13)

Resumo

O ATLAS (*A Toroidal LHC ApparatuS*) é o maior experimento do acelerador LHC (*Large Hadron Collider*) situado no CERN (*Centre Européen pour la Recherche Nucléaire*). O ATLAS está localizado em um dos pontos de colisão no túnel do acelerador. É composto de detectores especializados para caracterizar as partículas produzidas por colisões próton-próton (pp) com energia de centro de massa de 13 TeV. Um dos seus detectores especializados é o calorímetro de Argônio Líquido (LAr), com cerca de 187.000 células de sensores para registar a produção de partículas electromagnéticas. Os calorímetros são detectores largamente utilizados em experimentos de física de partículas, os quais têm a função de absorver e registrar a energia das partículas produzidas nas colisões no interior dos experimentos. O calorímetro em LAr do ATLAS têm uma granularidade fina e alta densidade de células, que em associação com as altas taxas de colisão, 40×10^6 por segundo, e a estrutura mecânica e eletrônica do sistema de leitura do detector, produzem efeitos de interferência proveniente dos canais eletrônicos vizinhos (*crosstalk* – XT). O XT aumenta a dificuldade de estimativa da energia e do tempo de propagação da partícula incidente. O Filtro Ótimo é o método padrão utilizado no ATLAS para estimar a energia e o tempo, mas no seu projeto de coeficientes, o XT não é levado em consideração, o que traz um erro na estimativa da energia e do tempo de voo das partículas. Esta tese investigou possíveis soluções para a mitigação do XT utilizando técnicas de aprendizagem de máquinas e métodos estatísticos de processamento de sinais. A solução proposta baseia-se no desenvolvimento de um estimador baseado em redes neurais artificiais para mitigar os efeitos indesejáveis do XT. Duas bases de dados simulados com base em modelos de física de partículas eletromagnéticas foram utilizados na aplicação e avaliação do modelo. Os sinais obtidos a partir dos simuladores desenvolvidos são utilizados para desenvolver e avaliar possíveis soluções em aprendizado de máquina. Utilizando os dados produzidos nos simuladores disponíveis, a abordagem utilizando métodos supervisionados mostrou-se eficiente tanto para a estimativa da energia quanto do tempo de voo das partículas utilizando a estrutura de *Multilayer Perceptron* (MLP) numa tarefa de regressão do valor da energia e do tempo das partículas sem a influência do XT. Os resultados obtidos produziram desempenho de regressão com baixo RMSE (*root mean squared error*) em relação ao método padrão utilizado para estimar a energia e o tempo de voo das partículas. Apresentando redução do erro na estimação da energia, siginifacante redução do erro na estimação do tempo de voo. Para o valor da energia a melhor rede possui uma camada oculta, já para o tempo, uma rede com três camadas ocultas produziu o melhor resultado. O método proposto reduziu o erro da estimação da energia em até 15,5% para energia de impacto 10 GeV, e reduziu a erro para a estimativa do tempo em até 2 ordens de grandeza na mesma energia de impacto.

Palavras-chave: ATLAS, Simulador Analítico, *Crosstalk*, Processamento Estatístico de Sinais, Aprendizado de Máquina, Lorenzetti Framework.

Abstract

ATLAS (A Toroidal LHC ApparatuS) is the largest experiment at the LHC (*Large Hadron Collider*) located at CERN (*Centre Européen pour la Recherche Nucléaire*). ATLAS is positioned at one of the collision points in the accelerator tunnel. It is composed of specialized detectors designed to characterize particles produced in proton-proton (pp) collisions with a center-of-mass energy of 13 TeV. One of its specialized detectors is the Liquid Argon Calorimeter (LAr), featuring approximately 187,000 sensor cells to record electromagnetic particle production. Calorimeters are widely used in particle physics experiments to measure and absorb the energy of particles produced during collisions within the experiment. The ATLAS LAr calorimeter has fine granularity and a high density of cells. Combined with high collision rates and the mechanical and electronic structure of the detector's readout system, this setup generates interference effects from neighboring electronic channels (*crosstalk* – XT). XT complicates energy and timing estimation for incident particles. The Optimal Filter (OF) is the standard method used in ATLAS for energy and time estimation. However, XT is not accounted for in the coefficient design, leading to errors in the estimated energy and time-of-flight of the particles. This thesis investigated potential solutions to mitigate XT using machine learning techniques and statistical signal processing methods. The proposed solution is based on the development of an artificial neural network (ANN) estimator to mitigate the undesirable effects of XT. Two simulated datasets, based on electromagnetic particle physics models, were used to apply and evaluate the model. The signals obtained from the developed simulators were employed to develop and assess potential machine-learning solutions. Using data produced by the available simulators, the supervised approach proved effective for estimating both the energy and time-of-flight values of particles. The *Multilayer Perceptron* (MLP) architecture was applied for regression tasks to predict the energy and time values of particles without XT influence. The results showed low RMSE (*root mean squared error*) compared to the standard method for estimating particle energy and time-of-flight. The approach achieved a significant reduction in energy estimation errors and a notable reduction in time-of-flight estimation errors. For energy estimation, the best network architecture featured a single hidden layer. For time estimation, a network with three hidden layers produced the best results. The proposed method reduced energy estimation errors by up to 15.5% for a 10 GeV impact energy and reduced time estimation errors by up to two orders of magnitude at the same impact energy.

Keywords: ATLAS, Analytical Simulator, Crosstalk, Statistical Signal Processing, Machine Learning, Lorenzetti Framework.

Résumé

ATLAS (*A Toroidal LHC ApparatuS*) est la plus grande expérience du LHC (*Large Hadron Collider*) située au CERN (Centre Européen pour la Recherche Nucléaire). ATLAS est positionné à l'un des points de collision dans le tunnel de l'accélérateur. Il est composé de détecteurs spécialisés conçus pour caractériser les particules produites lors des collisions proton-proton (pp) avec une énergie au centre de masse de 13 TeV. L'un de ses détecteurs spécialisés est le calorimètre à argon liquide (LAr), qui dispose d'environ 187 000 cellules de capteurs pour enregistrer la production de particules électromagnétiques. Les calorimètres sont largement utilisés dans les expériences de physique des particules pour mesurer et absorber l'énergie des particules produites lors des collisions. Le calorimètre LAr d'ATLAS possède une granularité fine et une densité élevée de cellules. Associée aux taux de collision élevés et à la structure mécanique et électronique du système de lecture du détecteur, cette configuration génère des effets d'interférence provenant des canaux électroniques voisins (*crosstalk* – XT). Le XT complique l'estimation de l'énergie et du temps pour les particules incidentes. Le Filtre Optimal (OF) est la méthode standard utilisée dans ATLAS pour l'estimation de l'énergie et du temps. Cependant, le XT n'est pas pris en compte dans la conception des coefficients, ce qui conduit à des erreurs dans l'estimation de l'énergie et du temps de vol des particules. Cette thèse a exploré des solutions potentielles pour atténuer le XT en utilisant des techniques d'apprentissage automatique et des méthodes de traitement statistique des signaux. La solution proposée repose sur le développement d'un estimateur basé sur un réseau neuronal artificiel (RNA) pour atténuer les effets indésirables du XT. Deux ensembles de données simulées, basées sur des modèles de physique des particules électromagnétiques, ont été utilisés pour appliquer et évaluer le modèle. Les signaux obtenus à partir des simulateurs développés ont servi à développer et évaluer des solutions potentielles basées sur l'apprentissage automatique. En utilisant les données produites par les simulateurs disponibles, l'approche supervisée s'est révélée efficace pour estimer les valeurs d'énergie et de temps de vol des particules. L'architecture *Multilayer Perceptron* (MLP) a été appliquée aux tâches de régression pour prédire les valeurs d'énergie et de temps des particules sans l'influence du XT. Les résultats ont montré un faible RMSE (*root mean squared error*) par rapport à la méthode standard d'estimation de l'énergie et du temps de vol des particules. L'approche a permis une réduction significative des erreurs d'estimation de l'énergie et une réduction notable des erreurs d'estimation du temps de vol. Pour l'estimation de l'énergie, la meilleure architecture de réseau comportait une seule couche cachée. Pour l'estimation du temps, un réseau avec trois couches cachées a produit les meilleurs résultats. La méthode proposée a réduit les erreurs d'estimation de l'énergie jusqu'à 15,5% pour une énergie d'impact de 10 GeV et a réduit les erreurs d'estimation du temps jusqu'à deux ordres de grandeur pour la même énergie d'impact.

Mots-clés : ATLAS, Simulateur Analytique, Diaphonie, Traitement Statistique du Signal, Apprentissage Automatique, Simulateur Lorenzetti Showers.

List of Figures

List of Tables

List of Abbreviations and Acronyms

ACM	Association for Computing Machinery. 25, 26, 130
ADC	Analogue to Digital Converter. 26, 29
AE	AutoEncoder. 26, 69
AI	Artificial Intelligence. 25, 26, 130
ALEPH	Apparatus for LEP Physics. 26, 132
ALICE	A Large Ion Collider Experiment. 26, 1, 12
AMS	American Mathematical Society. 25, 26, 130
ATLAS	A Toroidal LHC apparatus. 15, 16, 17, 18, 19, 21, 24, 26, 34, 1, 2, 3, 4, 5, 12, 16, 22, 23, 24, 25, 26, 27, 29, 35, 39, 40, 41, 42, 43, 45, 49, 50, 55, 67, 68, 83, 84, 86, 93, 94, 95, 97, 98, 99, 122, 124, 125, 129, 132, 133
BC	Bunch Crossing. 26, 4
BCID	Bunch Crossing Identifier. 26, 28
BDT	Boosted Decision Trees. 26, 129, 132
BE	Back end electronics. 26
BP	Backpropagation. 26
CALIB	Calibration. 26
CALICE	Calorimeter Linear Collider Experiment. 26, 4
CBTE	Centro de Ciência e Tecnologia do Bioetanol. 26, 124
CDF	Cumulative Distribution Function. 26, 46
CDM	Cold Dark Matter. 26, 11
CERN	*Centre Européen pour la Recherche Nucléaire*. 26, 1, 3, 10, 12, 13, 49, 83, 84, 133
CMS	Compact Muon Solenoid. 26, 1, 3, 9, 10, 12, 13, 28, 124, 125, 129, 132, 133
CNN	Convolutional Neural Network. 26, 129, 132
CNPEM	National Center for Energy and Materials Research. 26, 124
CONT	Controller. 26
CP	Charge Parity. 26, 11, 14
DAQ	Data acquisition. 26, 29
DELPHI	Detector with Lepton, Photon and Hadron Identification. 26, 132
DM	Dark Matter. 26, 11
DNN	Deep Neural Networks. 26, 129
DSP	Digital Signal Processor. 26
EBS	Extremely Brilliant Source. 26, 124
ELU	Exponential Linear Unit. 26
EM	Electromagnetic. 16, 18, 23, 26, 5, 7, 16, 17, 19, 21, 26, 27, 39, 40, 41, 43, 49, 50, 57, 99
EMB	Electromagnetic Barrel Calorimeter. 26
EMC	Electromagnetic Compatibility. 26, 34
EMCAL	Electromagnetic calorimeter. 19, 26, 1, 3, 5, 23, 41, 42, 43, 45, 49, 50, 68
EMEC	Electromagnetic Endcap Calorimeter. 26
EMI	Electromagnetic Interference. 26, 34
EMshower	Electromagnetic shower. 26

ESRF European Synchrotron Radiation Facility. 26, 124
eV electron volt. 26
FASER Forward Search Experiment. 26, 12
FCAL Forward Calorimeter. 26
FE Front Electronic Board. 26
FEB Front Electronic. 18, 26, 41
FEC Front end Crate. 26
FECMS Front end Crate Monitoring System. 26
FF Feedforward. 26, 45
FPGA Field Programmable Gate Array. 26, 133
GAN Generative Adversarial Networks. 26, 4, 133
GEANT4 Geometry and Tracking version 4. 26, 43, 60, 91
GeV Gigaelectron volt. 16, 23, 24, 25, 26, 1, 8, 9, 10, 11, 12, 13, 14, 20, 21, 22, 41, 42, 49, 50, 57, 58, 107, 125
GUTs Grand Unified Theories. 26, 11
H1 First sampling layer from hadronic calorimeter. 26
H2 Second sampling layer from hadronic calorimeter. 26
H3 Third sampling layer from hadronic calorimeter. 26
HCAL Hadronic Calorimeter. 26
HEC Hadronic Endcap Calorimeter. 26, 23
HEP High energy physics. 26, 33, 43, 132, 133, 134
HFEC Half Front end Crate. 26
HFFR Halogen-free, flame-retardant. 26
HGTD High Granularity Timing Detector. 26, 12
HL Hidden Layer. 19, 20, 21, 22, 23, 24, 26, 45, 53, 54, 56, 58, 67, 104, 106, 108, 129
HL-LHC High Luminosity Large Hadron Collider. 26, 4, 23
HS Hidden Sectors. 26, 14
HVPS High voltage power supply. 26
IBL Insertable B-layer. 16, 26, 25
ID Inner detector. 26, 23, 25
KeV Kiloelectron volt. 26
KNN K-Nearest Neighbor. 26, 129, 133
KS Kolmogorov-Smirnov Test. 26, 46, 51, 61, 65
L1 Level 1. 26
L1CAL Level 1 calorimeter trigger. 26
LAr Liquid Argon. 17, 18, 23, 24, 26, 1, 4, 5, 26, 28, 30, 37, 38, 39, 40, 41, 42, 43, 49, 50, 55, 67, 68, 69, 84, 94, 97, 99, 103, 124, 127
LEP Large Electron–Positron. 26, 10, 122, 132
LHC Large Hadron Collider. 15, 23, 26, 33, 1, 2, 5, 7, 9, 10, 12, 13, 14, 15, 23, 28, 29, 33, 83, 87, 122, 124, 129, 132, 133
LHCb Large Hadron Collider beauty. 26, 1, 12
LHCf Large Hadron Collider forward. 26, 12
Linac2 Linear Accelerator 2. 26, 13
LNBio Laboratório Nacional de Biociências. 26, 124
LNLS Laboratório Nacional de Luz Síncrotron. 26, 124
LNNano Laboratório Nacional de Nanotecnologia. 26, 124
LPNHE *Laboratoire de physique nucléaire et des hautes énergies*. 26, 43

LReLU Rectified Linear Unit. 26, 128
LZT Lorenzetti Showers Framework. 26, 33, 34, 39, 43, 68
MC Monte Carlo. 26, 69
MCTI Ministério de Ciência e Tecnologia. 26, 124
meV millielectronvolt. 26, 15
MeV Megaelectron volt. 23, 26, 12, 13, 14, 17, 29
ML Machine Learning. 15, 18, 26, 4, 5, 41, 44, 45, 46, 49, 67, 101, 127, 129, 130, 132, 133, 134
MLP Multilayer Percepton. 26, 5, 45, 68, 129, 132, 133
MoEDAL Monopole and Exotics Detector At the LHC. 26, 12
MSE Mean squared error. 26, 45, 54, 56, 59, 131
NLPCA Nonlinear Principal Component Analysis. 26
NN Neural Networks. 19, 20, 21, 22, 23, 24, 26, 45, 53, 54, 55, 56, 57, 58, 59, 67, 68, 104, 105, 106, 107, 108, 109, 127, 128, 129, 131, 132, 133
OF Optimal Filtering. 17, 18, 20, 21, 22, 26, 1, 29, 30, 37, 38, 41, 46, 51, 52, 53, 54, 55, 56, 57, 58, 67, 68, 99, 105, 106, 108
PCA Principal Component Analysis. 26, 129, 133
PCB Printed Circuit Board. 26, 33, 34
PDFs Probability Density Functions. 26, 20, 42
PLU Parametric Linear Unit. 26
PReLU Parametric Rectified Linear Unit. 26, 45
PS Proton Synchrotron. 26, 13
PS Pre sampler. 26, 25
PSB Proton Synchrotron Booster. 26, 13
PSpice Personal Simulation Program with Integrated Circuit Emphasis. 26, 4
QCD Quantum Chromodynamics. 26, 8, 9, 10
ReLU Leaky ReLU. 26
RMS Root mean squared. 26, 23
RMSE Root mean squared error. 23, 24, 26, 46, 54, 56, 57, 60, 104, 107
RNN Recurrent Neural Networks. 26, 129, 132
ROD Readout Driver. 26
S1 first sampling layer from electromagnetic calorimeter. 26, 25, 26, 42, 50, 99
S2 second sampling layer from electromagnetic calorimeter. 16, 19, 23, 24, 26, 25, 26, 40, 42, 43, 49, 50, 60, 99
S3 third sampling layer from electromagnetic calorimeter. 16, 26, 25, 26, 42, 50, 99
SCT Semiconductor Tracker. 16, 26, 25
SiPM Silicon Photomultiplier. 26, 4
SLAC Stanford Linear Accelerator Center. 26, 133
SM Standard Model. 15, 26, 1, 7, 9, 10, 11, 14, 18
SPAC Serial Protocol for ATLAS Calorimeters (serial control system). 26
SPS Super Proton Synchrotron. 26, 8, 13
SUSY Supersimmetry. 26, 14, 15
SVM Support Vector Machine. 26, 129, 133
T2Calo ATLAS detector discriminator. 26

TBB Tower Builder Boards. 26

TDB Tower Driver Boards. 26

TeV Teraelectron volt. 15, 16, 26, 2, 13, 14, 15, 20, 21, 122

TileCal Tile calorimeter. 26

TOTEM Total Elastic and diffractive cross section Measurement. 26, 12

TRT Transition Radiation Tracker. 16, 26, 25

TTC Trigger Timing and Control system. 17, 26, 28, 30

TTCrx TTC receiver chip. 26

UERJ Universidade do Estado do Rio de Janeiro. 26, 12

UFBA Universidade Federal da Bahia. 26, 12

UFJF Universidade Federal de Huiz de Fora. 26, 12

UFRJ Universidade Federal do Rio de Janeiro. 26, 12

UFSJ Universidade Federal de São João del Rei. 26

USP Universidade do Estado de São Paulo. 26, 12

VAE Variational AutoEncoder. 26, 4, 133

XT Crosstalk. 15, 18, 19, 20, 21, 26, 1, 3, 4, 5, 34, 38, 39, 40, 41, 42, 44, 45, 49, 50, 51, 52, 53, 54, 56, 57, 59, 60, 67, 68, 84, 94, 99, 100, 101, 103, 106, 109

List of Symbols

c	Vacuum Speed of light: 299 792 458 ms^{-1}. 26, 11, 14
E_c	Critical Energy. 26, 19
P_T	Transverse Momentum. 26, 132
Pb	Lead. 26, 21
Cu	Copper. 26, 21
Fe	Iron. 26, 21
E_0	Impact Energy. 26, 42
σ_E	Energy fluctuation. 26
R_M	Moliere Radius. 26, 17
$\sqrt{s}$	Energy on center of mass. 26
e^-	Electron. 26
R^2	Coefficient of determination. 26, 46, 59
H^0	Higgs boson. 26
σ_Z	Impact point fluctuation. 26
α	Learning rate. 26
$\mathcal{L}$	Luminosity. 26
μ	Muon. 26
ν	Neutrino. 26
η	Pseudorapidity. 26
ϕ	Azimuthal angle. 26
e^+	Positron. 26
γ	Photon. 26
$< \mu >$	Pile-up. 26, 33
$\oplus$	Quadrature summation - ex.: $m \oplus n = \sqrt{m^2 + n^2}$. 26
X_0	Radiation Length. 16, 26, 16, 20, 21, 22, 26, 50
σ	Standard Deviation. 26
$\otimes$	Tensor Product: $(\mathbf{v} \otimes \mathbf{w})_{ij} = v_i w_j$. 26
σ_t	Time fluctuation. 26
τ	Time of flight for particle. 26
E_T	Transverse energy. 26
dB	DeciBel. 26
W	Tungstene. 26, 21
U	Uranium. 26, 21
$W^{\pm}$	Weak interaction mediator boson. 26
λ_{int}	Interaction Lenght. 26, 17
$Z^{'}$	Z *prime*. 26
Z^+	Weak Interaction mediator boson. 26

Contents

1 Introduction

Human curiosity and the ability to ask questions about their world have fueled the most diverse scientific discoveries known today. One of the essential physics questions is: What is the matter made of? Thales of Miletus (624 – 547 B.C.), a Greek philosopher, among the first to raise this question and seek to explain the nature of the physical world. The discovery of subatomic structures in the early 1900s was a breakthrough that raised a series of new questions, among which the nature of dark matter (CLOWE et al., 2006) was required to explain many important aspects at various macroscopic scales in our universe (PINHEIRO; COSTA; MOREIRA, 2011).

In the last century, progress in our knowledge has brought expressive growth thanks to the complex experiments developed to unveil the nature of matter. Designing such experiments requires innovative solutions for different engineering problems in electronics, signal processing, computing, and data analysis. Currently, the world's largest particle accelerator in operation is the Large Hadron Collider (LHC) (EVANS; BRYANT, 2008), which is located at the *Centre Européen pour la Recherche Nucléaire* (CERN) (CERN Collaboration, 2016) on the border between France and Switzerland.

The Large Hadron Collider (LHC) benefits from a dense and complex structure of successive particle accelerators available at CERN. The proton-proton (pp) collisions provided by the LHC are studied thanks to several experiments, such as A Toroidal LHC apparatus (ATLAS); A Large Ion Collider Experiment (ALICE), Compact Muon Solenoid (CMS) and Large Hadron Collider beauty (LHCb). The accelerator circular tunnel is about 27 km long, and proton bunches collisions can occur at a rate of up to 40 $\times 10^6$ times per second (EVANS; BRYANT, 2008).

ATLAS is the largest experiment of LHC. It is of general purpose, was built in cylindrical format, and is formed by the following sub-detectors: tracking detector, calorimeters (electromagnetic, and hadronic), and the muon chamber. The volume of data produced in ATLAS is about 68 TB/s. Such data rate resulted in 2010 a total of 1 PB only in ATLAS (FREUND, 2011). The center of the ATLAS experiment is one of the nominal collision points of the proton beams of the LHC. ATLAS comprises different subdetectors. Among them, the calorimeters are responsible for measuring the energy of outcoming particles either deposited from an electromagnetic shower for electrons (e^-), positrons (e^+), and photons (γ) or from a hadronic shower for all the quark-composed particles. Beyond energy and direction measurements, granted by the granularity of the detector, calorimeter information is crucial for particle identification and triggering. The correct identification and characterization of electromagnetic particles is a key point for the Higgs boson search because that particle could decay into two γ or two pairs of $e^{\mp}$ (POVH et al., 2015).

The ATLAS Liquid Argon (LAr) Electromagnetic calorimeter (EMCAL) has a fine granularity and high cell density, which in association with the high collision rates and the mechanical and electronic structure of the detector readout, produces Crosstalk (XT) effects between cells which bias introduces biases the reconstruction process for the energy and time of the detected particle. When a cell records a physical or calibration signal, the electrical coupling between cells on the readout channel affects the reconstruction of shower shapes. This Crosstalk (XT) provokes distortion of the signal in each cell and induces a bias on energy and time reconstructed because the coefficients of the standard algorithm, Optimal Filtering (OF), are estimated using a signal shaped without the XT contributions in each cell.

Previous studies demonstrated that XT induces a bias around 94 ps on particle's arrival time estimation in the ATLAS Electromagnetic calorimeter (EMCAL) for an electron of a few tens of Gigaelectron volt (GeV) (COLAS et al., 1999; HUBAUT, 2000; HUBAUT et al., 2000; LABBÉ; ISHMUKHAMETOV, 2008). So the XT treatment is essential because reducing this bias in the particle time of flight increases the primary particle vertex position estimation accuracy. The XT treatment allows us to reduce the fluctuation on the particle identification (LOPEZ SOLIS, 2017).

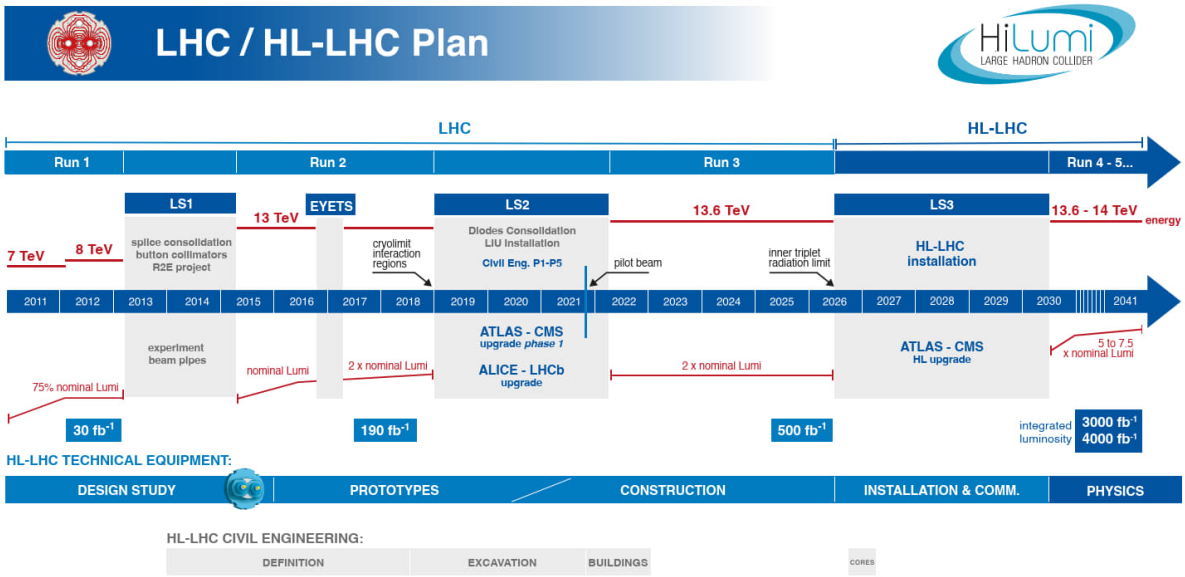
1.1 Motivation

The study of matter, its primary constituents, and their interactions resulted in the Standard Model (SM) of particle physics (CLOWE et al., 2006). The studies of Standard Model (SM) and searches for new physics beyond SM are at the knowledge frontier of humankind. Each discovery usually requires improving existing detector techniques and often the invention of new detection strategies and methodological approaches, especially in data analysis.

In the ATLAS experiment, there are approximately 100,000,000 electronic channels; approximately 3,000 km of cables; totaling 7,000 tons in a cylindrical-shaped structure with a diameter of 25 m and a length of 46 m (ATLAS Collaboration, 2010). All this complex structure is organized in concentric layers aiming to record the physics of rare occurrences, such as the Higgs boson decay (HIGGS, 1964) and other phenomena predicted by theoretical models.

The LHC operation starts in 2009 (CERN, 2015), with the first collision using an energy of 900 GeV[1] in the center of mass. Figure 1 begins in 2011 due to the word record energy of 7 TeV achieved for the first time. In Run 2 (2015–2018), collisions occurred at 13 Teraelectron volt (TeV) in the center of mass. That is, in 11 years, its operating energy has increased 15.56 times (see Figure 1). The number of particles in beam bunches, the luminosity[2] and temporal separation between each bunch follow a detailed schedule since the beginning of LHC operations. See Figure 1 (CERN, 2015)[3].

Figure 1 – LHC agenda for the updates on energy level and luminosity in the collisions for the next years. The plot begins with a world record beam intensity of 7 TeV in 2011.



Source: CERN (2024)

Run 3 will operate at 13.6 TeV, which is the target energy of LHC but with a luminosity[4] much higher than its design value by a factor of 5 to 7 (Béjar Alonso et al., 2020). A high luminosity should increase the probability of recording physics events with rare occurrences that are expected to answer physics questions. Beyond Run 3 (2022–2026), the plan is to operate the LHC in an even higher luminosity setting.

The signal processing chain in the ATLAS calorimeters should be able to handle the stringent conditions of LHC operation, such as several particle interactions per collision, very high energy, and collision rate, to provide accurate information on the energy deposits in each cell. In addition, some electronic drawbacks may contribute to the estimation errors. One important issue is the cross-interference (or crosstalk - XT) between adjacent readout channels in the LAr calorimeter. The XT effect introduces spurious information in each cell coming from other cells in the vicinity, which produces energy estimation errors.

[1] Electron-volt ($1\text{eV} = 1.602 \cdot 10^{-19}\text{J}$) energy gained by an electron accelerated by an electrical potential difference of 1 volt.

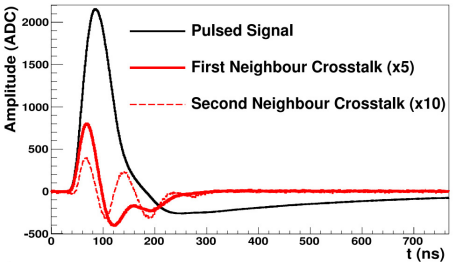
[2] Measure of the number of collisions per square centimeter produced every second [$\text{cm}^{-2}\text{s}^{-1}$].

[3] The LS marked points refer to long shutdown.

[4] Number of collisions that occur in a cross-section inside a collider.

In the Figure 2, it is possible to see the typical signals response in a cell on the front layer (S1) of the EMCAL and the XT from the first and second order neighbors.

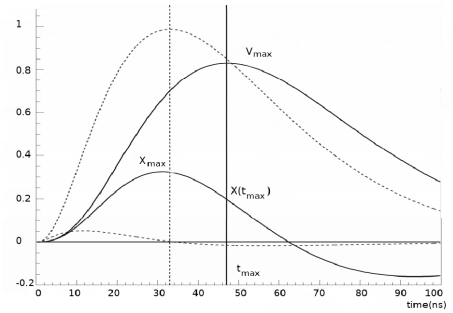
Figure 2 – Typical signal shapes on the LAr cells obtained at calibration runs at ATLAS final installation stage. In the solid black line, the pulsed signal from a cell is sensitized. In the solid red line, the crosstalk signals in the front layer from the first neighbors in η are scaled by a factor of 5, and in the dashed red line, the crosstalk from the second neighbor signal is scaled by a factor of 10.



Source: Labbé and Ishmukhametov (2008)

Figure 3 shows the main effect of the XT, i.e., the modification of the shape of the channel pulse, shifting the maximum position and the amplitude of the signal. This affects the process of signal reconstruction used to estimate the energy (V_{max}) and the time-of-flight (t_{max}) of the particle which is based on the expected pulse shape (solid line in Figure 3) that fired the channel. This effect introduces a bias in the scale measurement of the energy and time value (LABBé; ISHMUKHAMETOV, 2008).

Figure 3 – Example for distortion caused by XT in a cell of the ATLAS calorimeter in a normalized signal representation. The cell signal is the solid line with a peak around 0.8. XT signal the solid line with a peak around 0.3. The dashed line is the composed signal cell plus XT with the peak amplitude changed and time of peak shifted from ∼47 ns to ∼33 ns.



Source: Colas et al. (1999)

XT in high-energy physics experiments have been studied to achieve an accurate estimation of the channel coupling transfer functions, as well as to develop suitable mitigation strategies. Studies characterizing crosstalk in photodiodes have been ongoing since 1998 (KINDT; ZEIJL; MIDDELHOEK, 1998). In Ficorella et al. (2018), a measurement of this effect is made as a function of the distance between two different avalanche diodes. The Compact Muon Solenoid (CMS) experiment at CERN has drawn attention to crosstalk in a readout chip as they update their instrumentation systems for a new data-taking period (EBRAHIMI, 2019). In Yadav et al. (2018), Yadav and Shrimali (2019), analytical expressions of noise and crosstalk were proposed for a silicon particle detector with a charge-sensitive amplifier. Crosstalk is also an issue for X-ray spectroscopy detectors (TARTONI et al., 2020).

An accurate estimation of the time of flight is crucial in particle detection experiments. In another application, starting from the prototype stage, the XT rate has been evaluated for its contribution in dedicated circuits aimed at estimating particle time information (BUTALLA; HOHLMANN, 2020). In the context of the high-granularity calorimeter

with scintillator tiles in the Calorimeter Linear Collider Experiment (CALICE) project, the XT effects have been carefully considered and led to the replacement of several individual scintillator tile modules with a single piece (WEITZEL et al., 2020).

In Hohlmann (2020) a capacitive coupling model was developed using Personal Simulation Program with Integrated Circuit Emphasis (PSpice) to study XT in a micro-pattern gas detector, replicating experimental XT measurements. In some cases, the requirement is precision on the order of 10 ps for time measurements in dedicated circuits (QIN et al., 2022). It has also been observed that in a Liquid Argon (LAr) chamber equipped with two external tiles for detecting photons using a Silicon Photomultiplier (SiPM)-based system, XT effects, and other noise sources, thereby degrading overall performance (BOULAY et al., 2023). In the ATLAS experiment, the XT effects in LAr calorimeters, have been evaluated since 1999 (COLAS et al., 1999) before this experiment started data taking. On the sequence, posterior works added more information about this signal and its effects in the calorimeter cell information reconstruction (HUBAUT, 2000; HUBAUT et al., 2000; HUBAUT et al., 2003; LOPEZ SOLIS, 2017).

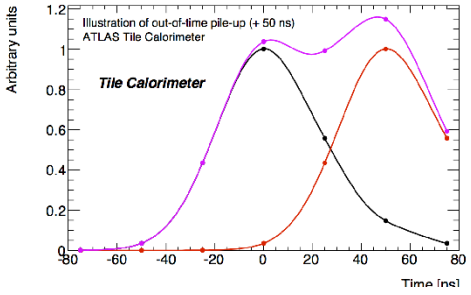
This work deals with the crosstalk mitigation problem in a segmented electromagnetic calorimeter comprising alternate volumes of liquid argon (LAr) as sensitive material and lead (Pb) as absorber material. For this, analytical modeling of capacitive and inductive crosstalk couplings between adjacent sensor cells is proposed to generate simulated events with the Lorenzetti Showers framework (ARAÚJO et al., 2023) and also with a simplified simulator developed in the context of this thesis (EMshower). A machine learning crosstalk mitigation method, based on artificial neural networks, is proposed to receive the signals distorted, as inputs to produces an estimation for energy and time without XT effects.

Machine Learning (ML) techniques are applied in particle physics experiments for different purposes such as feature extraction, as in Arpaia et al. (2021), in which ML techniques were evaluated for beam dynamics studies trying to correct possible beam losses at the LHC. In Pang (2021) Generative Adversarial Networks (GAN), and Variational AutoEncoder (VAE) are used to extract features to describe heavy-ion collisions. Two ML structures, GAN or VAE are very successful in generating models that can learn statistical information from data (HEP ML Community, 2024).

Another growing challenge in particle detectors is signal pile-up which adds challenges for the data processing system as the information from multiple events may be superimposed. In Figure 4a is presented an illustration of pile-up events in ATLAS tile calorimeter cell. In Figure 4a shows the successive detection of two events produced in two successive bunch collisions (out of time pile up), one in black and another in red. As a result, the spurious magenta signal that combines black and red events didn't carry information from black and red signals.

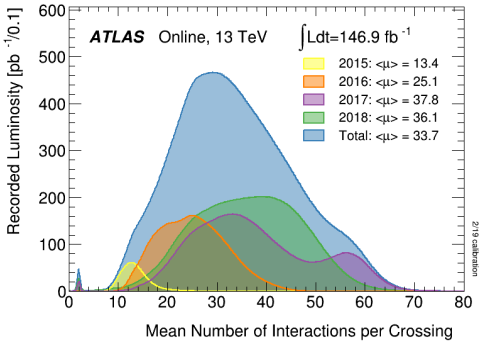
As illustrated in Figure 4b during Run 2, the average pile-up was between 30 and 40 simultaneous *pp* collisions per Bunch Crossing (BC), and it will rise to 200 during the High Luminosity Large Hadron Collider (HL-LHC) phase. The different events are produced in a time window corresponding to the bunches superposition, which lasts for about 200 ps and is called *in time* events or in-time pile-up.

Figure 4 – Pile-up events on the ATLAS tile calorimeter. In (a), one is black (curve with center on 0), and the other is red (curve with center on $\approx$ 50), two registered events in a short time interval. The result is the shapeless magenta event. Furthermore, the consequence of the pile-up event is the loss of both events, black and red. In (b), the average pile-up ($< \mu >$)from the first collision until the end of Run 2 with max equal to $\sim$70 at green and magenta curves.



(a)

Source: Peralva and Seixas (2015)



(b)

Source: Boyd (2020).

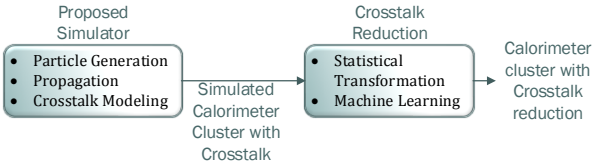
Similarly to XT, pileup also distorts the signal pulse shape. The combination of these phenomena contributes to an increase in the energy estimation error. In this context, a method capable of reducing the XT may allow for a better characterization of the pileup contributions, which is of major importance for future HEP experiments.

1.2 Objective

1.2.1 General Objective

The purpose of this work is to develop a Neural Network model for crosstalk reduction on a Liquid Argon calorimeter. Using simulated data it is possible to train a supervised regressor that maps the crosstalk-distorted information into the target (crosstalk-free) signals. As described in Figure 5.

Figure 5 – Simplified flowchart of the process of energy reconstruction of a particle after a collision. The blocks of expert knowledge and ML are the targets of this doctoral work.



1.2.2 Specific Objectives

- To study the physics of the energy deposition of electromagnetic particles;
- To study the structure of the electromagnetic calorimeter of the ATLAS experiment;
- To model the sources of fluctuation during the shower development process to feed the electromagnetic shower simulator;
- To develop an analytical simulator of electromagnetic particle shower based on the profile of energy deposition of electromagnetic particles and characteristics of EMCAL;
- To embed the XT coupling model at the first neighborhood into the Lorenzetti simulation framework (ARAÚJO et al., 2023), which allows studies on high-energy calorimeter electronics and signal processing systems;
- To develop an algorithm using ML techniques based on Multilayer Percepton (MLP) structure as a supervised method for the treatment of crosstalk.

1.3 Structure of the document

This Ph.D. thesis has 5 chapters. The Chapter 2 presents a brief historical summary of how the standard model of particle physics was built, from the first discussions and ideas about mater nature until our days. It also brings information about accelerators and particle detectors including the LHC and ATLAS experiment.

Chapter 3 presents the proposed method to treat the crosstalk problem in ATLAS LAr electromagnetic calorimeter cells. An analytical electromagnetic shower simulator based on ATLAS LAr and electromagnetic shower decay profiles was developed to provide signals to develop an ML approach to reconstruct energy and time without XT influence. A high-energy calorimeter simulation framework was also used to validate the proposed method.

Chapter 4 brings the results and analysis of the proposed method. Conclusions are derived in Chapter 5. Additionally, the Appendix A, presents the publications that originated from this work. **??** describes the standard method used to reconstruct energy and time for ATLAS LAr calorimeter cells. Appendix B presents the physical modeling for Electromagnetic (EM) energy profiles. Appendix C shows the analytical modeling developed in this work to simulate EM particle showers. Appendix D presents all tables and figures concerning the results of the method developed. Appendix E brings additional information about physics and specific parameters of experiments used.

2 Physics and LHC

Contents

Introduction

After a brief presentation of the SM of particle physics, this chapter presents how the origin of the mass of particles is a difficult question from a theoretical point of view and how the Higgs mechanism solves the problem. The sequence shows how we learned about the Higgs boson at the LHC since its discovery in 2012 and describes the ATLAS apparatus, one of the experiments used to discover the Higgs boson. Further findings are possible at the LHC and will benefit from an idea to improve the detector performance. This chapter will also focus on the EM calorimeter, which is the target of the research project addressed in this work.

2.1 High-Energy Physics (HEP) and LHC

2.1.1 The SM History - Short description

How is our universe organized? What are the fundamental constituents of the matter? Such questions have worried philosophers and physicists since antiquity. Tales of Miletus (624/623 – c. 548/545 B.C.), traditionally the most ancient Greek physicist, sought to explain the nature of the physical world. Later, Leucippus[5] postulated the idea of atoms[6] and void. This theory defines the atom as indivisible because the divisibility only exists due to void (KIRK; RAVEN; SCHOFIELD, 2010; PINHEIRO; COSTA; MOREIRA, 2011).

The first to differentiate atoms from molecules by taking up Leucippus atomist definition was Pierre Gassendi[7], in 1647. In his model, the molecules had to be small groups of atoms (PINHEIRO; COSTA; MOREIRA, 2011). Atom's definition didn't consider structure until the end of the XIX century. This picture changed with studies and the development of electricity (OSTERMANN, 2001) and the discovery of radioactivity.

[5] Leucippus from Eleia of Miletus (500-450 B.C.), philosopher.
[6] atoms - from the Greek, a-negation, tom-parts.
[7] Pierre Gassendi (1592-1655) - French philosopher and mathematician.

In the early XIX century, Dalton[8] along with Gay-Lussac[9] and Avogadro[10] took the first steps toward an atomic theory inspired by Lavoisier's work in meteorology[11] on air composition, demonstrating that it is composed of at least two gases (PINHEIRO; COSTA; MOREIRA, 2011).

A little more than a hundred years later, Thomson[12] obtained evidence of the electron presence within the atom, which led him to propose a model for the atom in 1903, now with the existence of electrons. In his model, Thomson defines the atom would be a uniformly charged positive sphere. Within such a sphere, electrons with a negative charge are uniformly arranged, which results in an electrical balance. This model presents inconsistencies because the neutrality of atoms requires opposite charges (ROCHA, 2002; PINHEIRO; COSTA; MOREIRA, 2011).

In 1904, Nagaoka[13] proposed a model in which the nucleus was a sphere, large and rigid, and the electrons would be distributed in saturnian-like rings (INAMURA, 2016). In parallel with the development of the atomic model, quantum mechanics emerge to help answer questions in which classical physics is inadequate, having as promoters Bohr[14] and Plank[15]. In 1911 Rutherford published a work about his experiments in 1909, which presents one peace inside of atom which is larger than the electron and where all positive charge does it concentered. This part he called *nucleus* and the electrons in a uniform distribution around the nucleus, which allows the atom to be neutral.

Some years later, in 1930, based on the energy conservation principle, Pauli[16] postulates the existence of an elusive particle, the neutrino, which was discovered in later experiments. In the 1930s, the following particles were observed: the electron (e^-), the proton (p), the neutron (n), the positron (e^+)[17] and the neutrino (ν) (1932) was postulated with the first observation recorded in 1970. The positron observation is thanks to the Dirac[18] studies based on the symmetry properties of his equation. Positron was the first antiparticle observed in the lab. As a consequence, this fact established the antiparticle concept. In 1935 Anderson and Neddermeyer[19] discovered the muon particle using cloud chamber measurements of cosmic rays (OSTERMANN, 2001).

Experimental facilities known as particle accelerators were designed to study the fundamental nature of matter. In the beginning, these machines accelerated electrons at high energy to hit a fixed target until colliding beams could be set up. Nowadays, accelerators use electrons and protons, antiprotons, or charged ions to study the decay of high-energy particles created in collisions. These machines allowed us to observe new subatomic particles and measure their characteristics (mass, electric charge, spin). These numerous products from strong nuclear interactions are called hadrons, and are represented by Greek letters (π, Γ, Λ, Ω, $\Sigma_c^{++}\ldots$) (OSTERMANN, 2001). Studying hadrons lead to identifying regularities that could be explained by postulating their compositeness and the existence of quarks bound by gluons in hadrons.

Advancing in history, in 1954 the proton synchronon was developed, known as the bevatron (BERKELEY LAB, 2021). This accelerator, at the time, was state of the art in particle accelerators, as it required a complex mechanical and electrical design in its development. It could work with energies above 30 GeV. With this accelerator, the University of California first produced antiproton and antimatter, which resulted in the 1958 Nobel Prize in physics. In the 1970s, a set of important discoveries in particle physics happened: in 1973, the Gargamelle experiment discovered weak neutral currents, which resulted in the unification of weak and electromagnetic forces, the first step before Z, W discovered at Super Proton Synchrotron (SPS) in 1983; in 1974, J/ψ discovery demonstrates the existence of the charm quark; in 1975, Martin Perl finds the tau lepton (3^{rd} family); in 1979, the gluon discovery at PETRA (DESY) establishes the Quantum Chromodynamics (QCD) (CHAO et al., 2023).

Some years later, in 1983, at Fermilab, a North American collider, the Tevatron came into operation and was

[8] John Dalton (1766 – 1844) - English chemist, physicist, and meteorologist.
[9] Joseph-Louis Gay-Lussac (1778 – 1850) - French chemist
[10] Amadeo Avogadro (1776-1856) - Italian chemist.
[11] Antoine Laurent de Lavoisier (1743 – 1794), French chemist.
[12] Joseph John Thomson (1856 – 1940), a British physicist with a Nobel Prize in Physics in 1906 for his investigations of the conduction of electricity in gases
[13] Hantaro Nagaoka (1865-1950), a Japanese physicist.
[14] Niels Henrick David Bohr (1885 – 1962), Danish physicist.
[15] Max Karl Ernest Plank (1858 – 1947), a German physicist.
[16] Wolfgang Pauli (1900 – 1958) Austrian physicist.
[17] Carl David Anderson (1905–1991), an American physicist using a cloud chamber experiment.
[18] Paul Adrien Maurice Dirac (1902-1984), a British theoretical physicist.
[19] Seth Neddermeyer (1907–1988), an American physicist.

registered as the first 512 GeV center of the mass-energy collision. Accelerating $\overline{pp}$ beams at 99,999954% speed of light and even recorded collisions with an energy of up to 1.8 TeV in 1986. However, the Tevatron closed its activities in September 2011, being the second-largest accelerator in operation at the time, behind only the LHC (FERMILAB, 2014) which began its activities in 2009, having its first recorded collision with the energy of 900 GeV. In Appendix E is presented a time-ordered summary of the most relevant discoveries in particle physics and its researchers until the ATLAS and CMS observations were announced at the same time Higgs boson detection at ATLAS and CMS experiments at CERN.

2.1.2 The Standard Model (SM) of Particle Physics

The Standard Model (SM) aims to establish the origin of electromagnetic, strong nuclear, and weak nuclear forces. The SM define for those three forces a particle to generate them, establishing that the photons are responsible for the electromagnetic force, for the strong force, the gluons, and the W and Z bosons for the weak force.

The SM constituents are organized into two main groups: the fermions[20] and bosons. The first one is subdivided into leptons and quarks, while the latter is the interactions' mediators. The leptons do not undergo strong interactions and have a half-integer spin $\frac{1}{2}$. There are six leptons, three with charge $Q = -1$ (electron *e*, muon μ and tau τ), and three without charge their respective neutrino partner with which each forms a pair, see Table 1In SM the basic constituents of matter are the elementary fermions: quarks, leptons, whose interactions are mediated by bosons (γ, Z^0, $W^{\pm}$) and gluons. For each fermion, there is a corresponding antiparticle with the same mass, and spin, but with opposite quantum numbers (KANE, 2003; CLOWE et al., 2006; BRAIBANT; GIACOMELLI; SPURIO, 2012; ELLWANGER, 2012).

Table 1 – Leptons and Quarks particle groups organized by family classification according to physics characteristics. Q is the electric charge, L_e is the electronic lepton number, B the baryonic number.

	Families			Characteristics		
	1^{st}	2^{nd}	3^{rd}	Q	L_e	B
Leptons	ν_e	ν_μ	ν_τ	0	1	–
	e^-	μ^-	τ^-	-1	1	–
Quarks	u	c	t	$+2/3$	–	$+1/3$
	d	s	b	$-1/3$	–	$+1/3$

Source: Braibant, Giacomelli and Spurio.

The second fermion subgroup comprises the quarks, which interact through all fundamental interactions. Quarks are the only particles that engage in a strong nuclear force. They cannot exist on their own but instead form the proton (uud), neutron (udd), and other hadrons[21], which are the basic constituents of the atomic nuclei. Similarly to leptons, quarks come in six so-called flavors. Three of them - up (u), charm (c), and top (t) - carry an electric charge of $+\frac{2}{3}$, represented as $+\frac{2}{3}e$, while the other three - down (d), strange (s), and bottom (b) - have a charge of $-\frac{1}{3}$, represented as $-\frac{1}{3}e$, as shown in Table 1.

In particle physics, the strong interactions between quarks are described by the theory of Quantum Chromodynamics (QCD) (National Research Council, 1998; MARTIN; SHAW, 2008; PIMENTA et al., 2013; De ANGELIS; PIMENTA, 2018). This theory introduces eight types of gluons, the bosons responsible for mediating the strong force that binds quarks into hadrons. There are three fundamental groups of particle physics in SM. Hadrons are particles formed by quarks. Part of these particles is called mesons[22], that are bosons, formed by one quark and one antiquark. Another part of the hadrons is called baryons[23], which are formed by three quarks (ELLWANGER, 2012; De ANGELIS; PIMENTA, 2018; BRAIBANT; GIACOMELLI; SPURIO, 2012).

[20] Particles with a half-integer spin. *spin* - intrinsic characteristic for the elementary particles; one of the quantum numbers that help to define a particle.

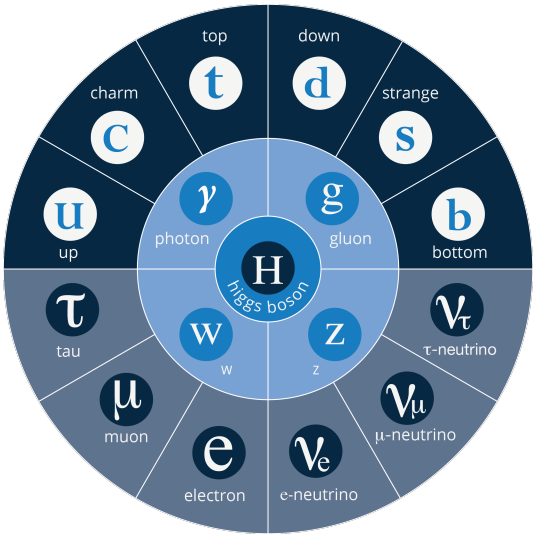
[21] From the Greek *hadrós* — strong. The quark composite particles are subject to strong nuclear force

[22] From the Greek *méson* meaning intermediate or medium, used to identify particles with medium mass.

[23] From the Greek *barýs* - heavy. This name was because, until the baryon discovery, the previously known particles had lower masses than baryons.

Bosons are the mediators of interactions. Das and Ferbel (2003) resume particle physics that way: beyond QCD all physics particles on the nature are either bosons[24], that is obey the Bose-Einstein[25] distribution, or are they fermions which obey the Fermi-Dirac distribution[26]. Those particles are responsible for generating: the quantum field, electromagnetic field, weak force field, strong force field, and gravitational field. The Higgs boson, according to current theory, is the excitation of the Higgs field, which permeates the entire universe. Thus, from the interaction of the Higgs boson with a particle, there is a mass-energy transfer from the Higgs field to the elementary particle (PIMENTA et al., 2013). This set of particles and classification, taking into account physical parameters, building the SM as shown in Figure 6,

Figure 6 – Simplified diagram of SM with the two main groups, Quarks and Leptons, and the bosons mediators of interactions. At each quadrant, a boson mediator for a set of three fundamental particles, γ and g massless bosons and W and Z mass bosons, and the Higgs, at the center, participate in all interactions.



Source: ATLAS Collaboration (2021).

2.1.3 The Higgs Boson

This spin 0 particle was theorized in 1960 by Peter Higgs (HIGGS, 1964), François Englert, and Robert Brout (ENGLERT; BROUT, 1964)[27] as responsible for the generation of the mass of all fundamental particles. However, its mass was a free parameter of the model (BRAIBANT; GIACOMELLI; SPURIO, 2012). Before LHC experiments started, the maximum mass for the Higgs had been searched up to 114 GeV, through experiments developed on Large Electron–Positron (LEP)[28] (ELLWANGER, 2012). This range was not enough for Higgs boson observation and triggered the LHC program, which allowed extensive studies.

Nath et al. (2010) presents the two main physical decay channels for the Higgs Boson at the LHC:

- $H^0 \to \gamma\gamma$ – A channel with sensitive in the range, $120\,GeV \leq M_H \leq 140\,GeV$ that could be registered due to the resolution of the electromagnetic channels at the ATLAS and CMS experiments;
- $H^0 \to ZZ$ – This channel, $H^0 \to ZZ \to 4\ell$, (where $\ell = e\,,\mu$), was called the *golden channel*, because it produces a clear peak in a 4ℓ invariant mass spectrum, with the main backgrounds, $t\bar{t}$, ZZ, $Zb\bar{b}$. In this channel the mass is up to $M_H \sim 500\,GeV$ a 30 [fb^{-1}][29];

[24] Particle with integer spin.
[25] Particles that build particle fields, *bosons*.
[26] Particle of matter, *fermions* (De ANGELIS; PIMENTA, 2018).
[27] Robert Brout (1928 – 2011) a Belgian theoretical physicist.
[28] Accelerator used for the CERN between years 1989 and 2000
[29] f – femto, equal to 10^{-15} m; barn – Cross section unit equal to $10^{-28}\,m^2$, then 1 $fb = 10^{-43}\,m^2$.

The LHC started in 2009 with a center of mass energy equal to $\sqrt{s} = 900$ GeV. In 2011 the energy level was increased to $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV when for the first time the Higgs boson was detected. For the ATLAS detection $M_H = 126 \pm 0.4\,(stat.) \pm 0.4(sist.)$ GeV, and for CMS detection $M_H = 125.3 \pm 0.4\,(stat.) \pm 0.6(sist.)$ GeV (ATLAS Collaboration, 2012a)[30].

2.2 Open Questions in SM

This section intends not to cover all open questions in S.M. or provide rich details about each open question, but keep in mind that many questions are without answers yet. The theoretical physicist hunt for a Grand Unified Theories (GUTs) that describes nature with all the particles in the universe, and their interactions makes SM the best and most close to this aim. Nevertheless, some open questions even in the S.M. or beyond it are described in Ghosh et al. (2020) as follows:

- Gravity - The SM does not describe gravitation interactions, one of the four fundamental forces, which is essential to explain interactions at large scales;
- Matter and Antimatter Asymmetry - The known universe has predominance of matter over antimatter. This question could be related to a breaking of the Charge Parity (CP) symmetry-related with matter and antimatter in strong interaction
- GUTs - This theory is expected to unify fundamental interactions described by the SM into one single model that aims to be a complete description. The strings theory is one of the most known theories in this direction;
- Dark Matter (DM) - Is expected to be an explanation for the fact that less than 5% of the observable universe is composed of ordinary matter that SM describes. Most specifically, Cold Dark Matter (CDM) is believed to constitute around five times the ordinary matter in the universe. Another thing is that SM didn't provide any particle candidate to describe DM or dark energy;
- Neutrino Mass - According to S.M., the neutrino is a massless particle, but the discoveries of neutrinos oscillations ask for an S.M. extension to explain these observations.

2.3 Particles Accelerators and Detectors

The studies on particle physics became possible only due to the development of dedicated technologies: accelerators and particle detectors. The intense discussion and grown-up knowledge about the nature of matter allowed for a tremendous technological development of the machines and equipment that resulted from the former particle physics studies and discoveries.

According to Povh et al. (2015) the experiments in high-energy physics could be classified into:

- Scattering - In this type of experiment, a beam of probe particles with known energy and momentum is directed towards a target. This target is the object to be studied. The information about the object is studied by the changes in kinematics caused by the beam interaction process. An example of this process is the photoelectric effect, in which an object, a metal, most of the case, emits electrons when exposed to electromagnetic radiation. The large energy corresponds to the lowest spatial scales being investigated;
- Colliders - In this experiment, elementary particles, are accelerated close to speed of light (c) to collide. (PERELSTEIN, 2011);
- Spectroscopy - Is used to describe the decay product of the excited states, changes in energy levels, as well as interactions between their constituents. An example is the particles produced by beam collisions to study rare and theoretical particle decay.

For detectors, the main classes are:

[30] *stat* - refers to statistical procedure to determine the mass value, and *sys* - refers to the systematic uncertainties combined described on the paper and their references.

- Scintillators - Provide fast timing information and energy measurement as an important part of some calorimeters;
- Gaseous - Cover a large area and provide a good spatial resolution, and are used in combination with a magnetic field to measure the momentum of charged particles;
- Semiconductor - Have high energy and spatial resolution. In the ATLAS experiment, the High Granularity Timing Detector (HGTD) detector was developed based on a low gain avalanche detector, with a timing resolution equal to 30 ps. This precision in time aims to mitigate the effects of the pileup (ALLAIRE, 2018);
- Cherenkov - Counters based on transition radiation that are used for particle identification;
- Bubble Chamber - Composed of a liquid superheated to a high temperature at high pressure. If a particle enters the chamber, then ionizes the atoms, and the size and number of bubbles grow long, the tracks and photographs are taken (JAYAKUMAR, 2012);
- Calorimeters - Used to measure the total energy.

This classification shows how diverse the kinds of technologies of particle accelerators and particle detectors structures applied in particle physics across the years to increase the knowledge about matter in different countries. The rapid development of the field can be observed on the Table 5.4 in Appendix E with a shortlist of experiments developed between 1962 and 2015 according to Chao et al. (2023). Another interesting fact is that the energy level has increased from 250 Megaelectron volt (MeV) (1962) to 7 GeV (LHC in 2009) per beam, an increase of 28 times (Appendix E).

The size of the collaborations and the dataset significantly impacted the development of specific technologies that became very important for society (data grid, HTML, distributed analysis, etc.). A good example of this is the *internet*, which was developed to allow communication between research centers and CERN (FLÜCKIGER, 2013) and now it is impossible to think of countries' collaboration without the internet.

2.4 Large Hadron Collider (LHC)

LHC is the biggest particle collider ever built. It is located at the French-Swiss frontier, and has a circular tunel of $\sim$ 27 km, at 50 m to 175 m underground. The LHC has a complex structure that allows us to investigate a set of physical theories beyond the actual standard model. It's located at CERN (CERN, 2017). The project and discussions started in the early 1980s, and in December of 1994, the CERN board approved its construction that started the following year (CERN, 2015).

The LHC experimental setup is composed of eigth experiments: CMS, ATLAS, Large Hadron Collider beauty (LHCb), A Large Ion Collider Experiment (ALICE), the Total Elastic and diffractive cross section Measurement (TOTEM), Large Hadron Collider forward (LHCf), Monopole and Exotics Detector At the LHC (MoEDAL) and the Forward Search Experiment (FASER)[31], those last three, are small detectors (FREUND, 2011; CERN, 2015). Figure 7 presents where are located the main four largest experiments in the CERN tunnel structure.

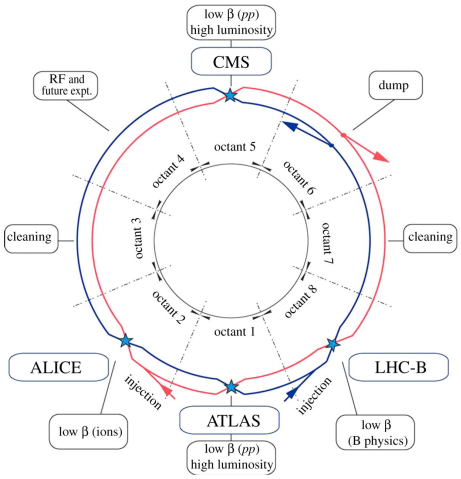
Each experiment has a specific group of researchers and institutions organized in international collaboration. The ATLAS collaboration has 3000 scientific authors coming from hundreds of institutions[32] from several countries (ATLAS Collaboration, 2020a).

Brazil is part of the ATLAS collaboration. The Brazilian cluster includes: UFRJ since 1996. USP since 1996, UFJF since 2011, UFBA since 2023, and UERJ since 2021.

[31] One of the two newest experiments situated at either side of the ATLAS cavern to detect neutrinos produced in proton collisions in ATLAS is designed to search for light and extremely weakly interacting particles.

[32] Only official institutions in 42 countries. Because there exist many other institutions that work together with official institutions during the process of recognition.

Figure 7 – Structure of the main rings of LHC where the sets of experiments are located.

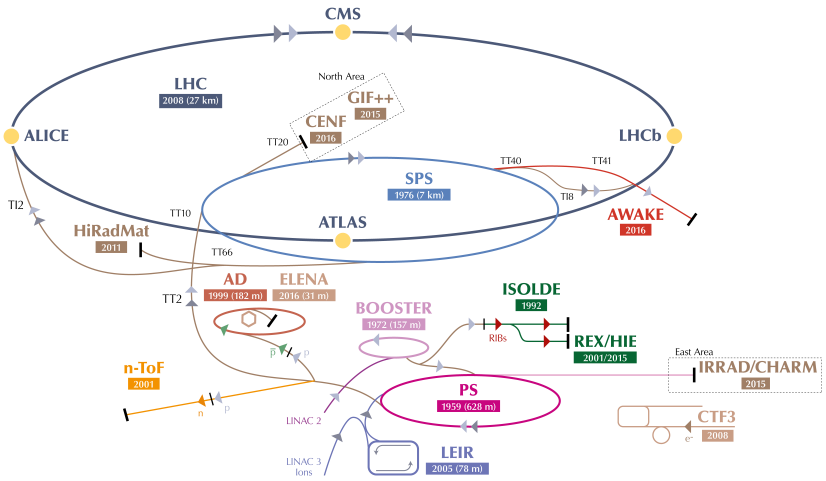


Source: (EVANS; BRYANT, 2008).

Figure 8 presents an illustration of the CERN accelerator complex, including the initial accelerating steps used to inject proton bunches at the LHC accelerator. The process starts by extracting protons from hydrogen atoms using a plasma source. In the following the acceleration process begins with the Linear Accelerator 2 (Linac2) in which protons are accelerated until 50 MeV and go to the Proton Synchrotron Booster (PSB) (BOOSTER In Figure 8) to increase energy up to 1.4 GeV.

The next step relies on the Proton Synchrotron (PS) where protons are accelerated to 25 GeV and go to SPS where they are accelerated to 450 GeV. Finally, the proton bunches go to the LHC rings into bunches circulating in a clockwise and anticlockwise direction, where bunches are accelerated for 20 minutes to their nominal energy of 7 TeV. Powerful dipoles and quadrupoles drive bunches to collide inside the different experiments, especially ATLAS (octant 1) and CMS (octant 5) (CERN, 2017).

Figure 8 – Diagram of the structure of the LHC accelerator complex. The big ring is the final step where bunches achieve the collision energy where the experiments are placed. Outside of the big ring, are the intermediate accelerating steps, and labs to monitor the process.



Source: Adapted from CERN (2017).

For reaching the energy collision in the TeV range, the proton beams should be accelerated close to the speed of light, and this is done by using five accelerating stages to increase the energy step by step. At the LHC, the final

speed of particles that collide is 99,9999991% of the speed of light at the energy of 14 TeV. Table 2 shows the different energies reached in each step of the acceleration chain (CERN, 2017, p 5). To achieve the high close to the speed of light (c) the accelerator complex is composed of a large number of electromagnetic devices used to accelerate and drive the proton bunches, see Table 3.

Table 2 – Speed and bunch energy of protons for each acceleration step until the collision. At each step accelerator, the proton-proton bunch energy is increased gradually from 50 MeV (31.4% speed of light) until the final sequential stage up to 7 TeV per bunch crossing (99.9999991% of the speed of light).

E_C proton	50 MeV	1,4 GeV	25 GeV	450 GeV	7 TeV
V (%c)	31.4	91.6	99.93	99.9998	99.9999991
Accelerator	Linac 2	PS Booster	PS	SPS	LHC
Proton mass: 0.938 GeV/c^2.					

Source: CERN (2017).

Table 3 – List of important parameters of LHC. The complex structure for the LHC accelerator requires a large number of electromagnetic devices, to control and achieve the final amount energy of 7 TeV per bunch.

Description	Quant.
Circumference	26.659 m
Magnets	9.593
RF cavities	8 / beam
Nominal energy, protons	7 TeV
Min. distance between bunches	$\sim$ 7 m
Design Luminosity	$10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$
Bunches per p beam	2.808
p per bunch (at start)	1.2×10^{11}
Num. turns/s	11.245
Num. collisions/s	1 Bi

Source: CERN (2017).

After the discovery of the Higgs boson, one of its primary initial targets, the LHC program investigated fundamental open questions about elementary particle physics and searched for new physics beyond SM (ELLWANGER, 2012, Cap. 8). Some of these are described by Nath et al. (2010) are:

- **Higgs boson characterization** - Theoretical particle predicted in 1960 as the responsibility to transfer energy to the other particles in mass form. The first time that it was recorded was in 2012 (ATLAS Collaboration, 2012b), but more studies are necessary to assess the model among which the existence of mutual interaction between the Higgs boson;
- **Supersimmetry (SUSY) search** - is an extension of the Standard Model theory to describe the basic building blocks of matter that aims to fill some of its gaps. It predicts a partner particle for each particle in the Standard Model. (CERN, 2015);
- **CP violation** - SM defines sources of CP violation which are not enough to generate the observed baryon asymmetry in the universe. LHC may allows to observe new CP violation sources in the *b* sector (KOBAYASHI; MASKAWA, 1973);
- **Dark Matter** - There exists an expectation that could explain the 23% of the invisible matter in the universe. This expected candidate could become from the SUSY proof (CLOWE et al., 2006);
- **Top-Quark physics** - It's possible to take discovery about the physics of this particle that could increase definitions on SM;
- **Exotic particles** - Some theories predict that there could be a whole new set of particles out there that physicists did not detect because they don't interact through the electromagnetic force (CERN, 2015);
- **Visible signatures in Hidden Sectors (HS)** - These models are responsible for breaking supersymmetry. These signatures are related to dark matter also and can be proved by LHC experiments;

- **Probing the Origin of Neutrino Mass on LHC** - There exists a probability to study the neutrino mass mechanism at the LHC;
- **Extra Dimensions** - In this hypothesis, the explanation for the gravity weakness concerning other fundamental forces is the fact that gravity is spread in extra dimensions (CERN, 2015). This hypothetical approach is an alternative to SUSY;

2.5 Detection using Calorimetry

In nuclear and particle physics, calorimetry means stopping the particles in a block of matter instrumented to catch the energy deposits and characterize their profiles. Calorimetry is very important and widely used for energy measurement from particles that come from collision experiments. One of the first applications of calorimetry was in the γ ray measurement that came from radioactive nucleon (LIVAN; WIGMANS, 2019).

Calorimeters form a particular detector class that is widely used in energy and particle position measurement by total energy absorbing. Calorimeters are blocks of matter in which the particle is to be measured and transforms (part of) its energy into a measurable quantity (WIGMANS, 1991). Those equipment are widely used in high-energy experiments since they allow measurements in a large dynamic range. The range starts in low-energy experiments such as astrophysics, which seeks to register and detect energy levels at the millielectronvolt (meV) level or in the experiments where they are used to measure e, γ or jets up to TeV range. They are structures designed to absorb the energy of particles as they interact with the detector material in the development of the particle shower, whether they are electromagnetic or hadronic (FABJAN; GIANOTTI, 2003; MARTIN, 2006; GRUPEN; SHWARTZ, 2008; WIGMANS, 2008).

Calorimeters are of two types. The first one is Electromagnetic, used to measure electron and photon energy from electromagnetic interactions (mainly dominated by bremsstrahlung[33] and pair production). The second kind is Hadronic, which measures hadron energies deposited from strong and electromagnetic interactions.

This detector has some interesting and relevant characteristics for the particle physics field that were summarized on (WIGMANS, 1991; FABJAN; GIANOTTI, 2003):

- Sensitive to both charged and neutral particles. This helps to identify neutrinos and possibly more exotic elusive particles by indirect detection using the energy-momentum observation that allows defining the concept of missing transverse energy;
- They are compact because their size increases logarithmically with the particle energy;
- Energy measurement can be done without a magnetic field;
- In contrast with magnetic spectrometers in which momentum resolution degrades linearly with energy, in most cases, the calorimeter's energy resolution improves with energy as $\frac{1}{\sqrt{E}}$[34] where E is the energy of the incident particle;
- They are versatile detectors. Implemented with a configurable granularity, they can be used to detect the position and direction of the particles from successive decay processes called showers. This characteristic allows for the differentiation of electrons and photons from pions (π) from muons (μ);
- Transverse Energy information can be used to select events of interest with high selectivity;
- It can be a high-speed detector in which response times smaller than 50 ns have already been reached (WIGMANS, 1991).

Although the calorimeters can detect different types of particles, these detectors are usually selectively constructed, comprising electromagnetic and hadronic segments installed in the sequence (MARTIN, 2006) to be able to identify different types of particles. Also, these detectors are classified according to their internal structure in two categories, the so-called Homogeneous and Sampling calorimeters. The first one is made of a block of matter using a

[33] radiation generated when a charged particle, like an electron, suffers deceleration and changes its trajectory and releases energy in the form of photons. From the German *Bremsen* = brake and *Strahlung* = radiation.

[34] At high energy, its resolution is dominated by a constant term.

single material to absorb and measure the particle energy. The second is made of a multilayer structure of two materials, one to absorb and another to measure the energy of particles like in the ATLAS experiment (ATLAS Collaboration, 1996).

Designing EM calorimeters requires taking into account some important physical parameters of the EM shower. The first one is the radiation length (X_0) which is described in Freund (2018) as the average distance necessary for an electron with high energy to reduce its energy by $(1 - e^{-1})$, or approximately 63,2%, by bremsstrahlung emitting photons and photons making e^+/e^- pair, by conversions. The fact that γ are uncharged makes a difference between a shower induced by a e^- or a γ. The photon shower has a $\frac{9}{7}$ ratio in its free path through a medium concerning electrons to generate its first interaction in the calorimeter as calculated by (LIVAN; WIGMANS, 2019)

$$X_0 = \frac{716.4\,\mathrm{A}}{Z(Z+1)\ln(287/\sqrt{Z})}\,[\mathrm{g\,cm^{-2}}], \tag{2.1}$$

where A is the atomic weight, and Z, is the atomic number.

When a calorimeter is composed of different materials, the effective $X_{0_{eff}}$ value is determined by (NAVAS et al., 2024):

$$\frac{1}{X_{0_{eff}}} = \sum_{i=1}^{n} \frac{w_i}{X_{0_i}}\,[\mathrm{g\,cm^{-2}}], \tag{2.2}$$

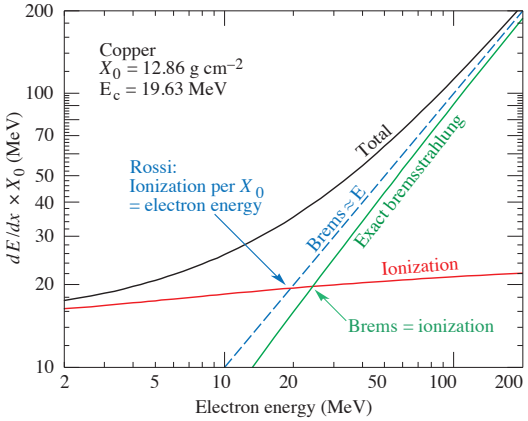
where w_i and X_{0_i} are, respectively, the weight fraction and radiation length for each material that compounds the structure.

The second important physical parameter is the Critical Energy (E_c) which has two usual definitions. The first one is defined as the energy at which the average energy losses from radiation processes equal those from ionization. However, a second definition, a little bit different, could be used too. Is the energy at which the ionization loss per radiation length (X_0) is equal to the electron energy. These definitions are equivalent in high-energy processes, see Figure 9. The Equation 2.3 shows the empirical formulae used to compute the critical energy in a solid or a liquid state, while Equation 2.4 applies to gaseous medium (WIGMANS, 2017; LIVAN; WIGMANS, 2019; NAVAS et al., 2024).

$$E_c = \frac{610}{Z + 1.24}\,[\mathrm{MeV}] \tag{2.3}$$

$$E_c = \frac{710}{Z + 0.92}\,[\mathrm{MeV}] \tag{2.4}$$

Figure 9 – Two definitions for Critical Energy. Loss of energy processes on copper. In the first one the ionization energy is equal the ionization loss per radiation length is equal to the electron energy (dashed line), and the second the loss by bremsstrahlung (solid line). In high-energy processes those definitions are equivalent.



Source: Navas et al. (2024).

Another parameter that characterizes the transverse particle shower development is the Molière Radius (R_M). Its definition in terms of the radiation length function is presented in the Equation 2.5. Equation 2.6 allows calculating the effective Moliere Radius ($R_{M_{eff}}$) of a detector composed by layers of different materials, analogue Equation 2.2, where $E_s = m_e c^2 \sqrt{4\pi/\alpha} = 21,2$ MeV.

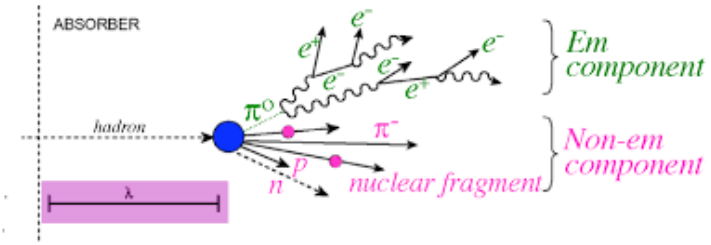
$$R_M = X_0 \frac{E_s}{E_c} \, [\mathrm{g\,cm^{-2}}] \tag{2.5}$$

$$\frac{1}{R_{M_{eff}}} = \frac{1}{E_s} \sum_{i=1}^{n} \frac{w_i \, E_i}{X_{0_i}} \, [\mathrm{g\,cm^{-2}}] \tag{2.6}$$

In a cylinder of 1 R_M radius around the shower axis contains, on average, between 85% – 90% of the energy of the incoming particle. If the cylinder has 3.5 R_M, it contains about 99% of energy. Therefore it's a statistical parameter associated with the energy loss process on particle medium interaction (WIGMANS, 2017; LIVAN; WIGMANS, 2019; NAVAS et al., 2024).

The second type of particle shower is associated with hadron interaction in the matter, which is mainly by strong interactions and EM. The hadron shower development is similar to EM shower development process, but the particle production mechanism is more complex than EM mechanism. Part of the energy is transferred to the nucleus, and this results in new hadrons (e.g., π, **K**, η, etc.) production. Some of these particles are unstable (e.g., $\pi^0 \rightarrow \gamma\gamma$, $\eta \rightarrow \gamma\gamma$, etc.) and make the hadronic shower a mix of electromagnetic showers, and in non-electromagnetic showers, with charged hadrons and nuclear fragments with a fraction of this energy that does not contribute to calorimeter signals (LIVAN; WIGMANS, 2019), like Figure 10 shows.

Figure 10 – Schematic description of a hadron shower. The energy carried by the hadron is typically deposited in the form of an electromagnetic and a non-electromagnetic component. EM components are from π^0s and ηs from nuclear interactions, the non-EM component is from charged hadrons, and one additional fraction didn't contribute to the calorimeter signals.



Source: Livan and Wigmans (2019).

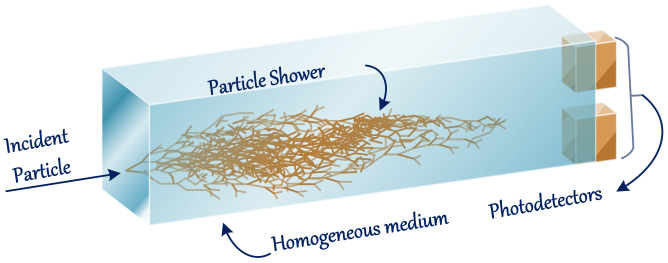
The description of hadronic showers within the material is made using a parameter defined as the interaction length (λ_{int}), which characterizes the average free distance a high-energy hadron has to travel inside a medium before a nuclear interaction occurs, which defines the shower dimension (WIGMANS, 1991)

2.5.1 Homogeneous Calorimeters

Those calorimeters are made of a single material with two functions to absorb and generate a signal that can measure the energy of the particles. They are based on scintillation light measurement, ionization, and Cherenkov radiation (GRUPEN; SHWARTZ, 2008). Figure 11 shows the structure of one of this kind of calorimeter and the particle shower that results from the interaction of an incident particle.

These detectors have good statistical precision because all charged particles are detected with good linearity due to the nonexistence of irregularities or gaps in the sensitive medium. Nevertheless, this kind of calorimeter has an expensive cost associated with its construction based on scintillating crystals, while it is challenging to build a highly segmented detector (GARUTTI, 2012) with this technology.

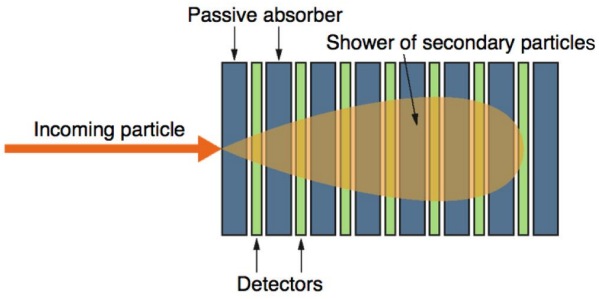
Figure 11 – Homogeneous calorimeter diagram. The indicated photodetectors could be photodiodes or PMTs – *PhotoMultiplier Tubes*.



2.5.2 Sampling Calorimeters – Non-Homogeneous

Sampling calorimeters use different materials for active and passive media. The active medium measures the charge on the light-induced by the particles passing in this volume. In contrast, a passive medium absorbs the particle energy and generates the charged particles detected in the active volume. Because of this, it's necessary to have a high-density passive medium. Figure 12 shows a diagram of a sampling calorimeter with an electromagnetic shower decay process for an electron or photon crossing the detector structure.

Figure 12 – Illustration of an electromagnetic particle shower in a sampling calorimeter. The blue layers are the passive materials to absorb particle energy and the green layers are the active medium to record particle energy.



Source: Krammer (2011).

The significant advantages of sampling calorimeters are associated with their building cost, compared with homogeneous ones, while allowing for good longitudinal and transverse segmentation, making those detectors able to measure the energy and particle direction. As a disadvantage, these calorimeters are that only part of the deposited particle energy is detected in the active layers (De ANGELIS; PIMENTA, 2018; NAVAS et al., 2024), and often require cryogenics infrastructure (LAr boiling temperature = 87 K) and high purity of the liquid (PRETZL, 2005). To a list of some important homogeneous and sampling calorimeters, see Appendix E

2.5.3 Electromagnetic Shower

Electromagnetic particles, such as electrons positrons, and photons, are important for different Higgs boson decay studies as they also major searches for new physics beyond the SM described in subsection 2.1.3: $H^0 \to \gamma\gamma$, $H^0 \to Z\,Z \to 4\ell$, $H^0 \to WW \to \ell\nu\ell\nu$, where $\ell = e\,, \mu$.

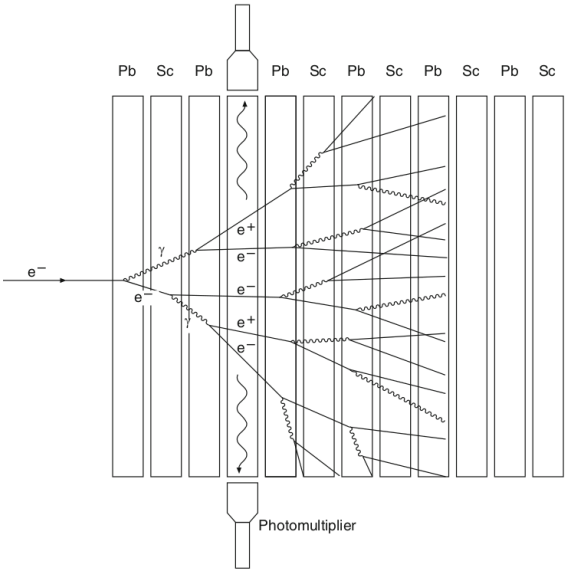
The following paragraphs will shortly describe the electromagnetic shower development process, presenting the longitudinal and radial profiles for the energy deposition. These models are part of the electromagnetic shower simulator developed to study the crosstalk effects and build a method to mitigate this interference in the present thesis (see Chapter 3).

Two main processes produce electromagnetic showers. For electrons and positrons ($e^{\mp}$) by the bremsstrahlung effect, and for photons by pair production. A simple model of EM showers can be built using the following assumptions (MARTIN; SHAW, 2008):

- Each electron with $\mathrm{E} > \mathrm{E}_c$ travels one radiation length and then transfers half its energy to a bremsstrahlung photon;
- Each photon with $\mathrm{E} > \mathrm{E}_c$ travels one radiation length and then creates an electron-positron pair with each particle having half the energy of the photon;
- Electrons with $\mathrm{E} < \mathrm{E}_c$ cease to radiate and lose the rest of their energy by collisions inside the medium;
- Ionization losses are negligible for $\mathrm{E} > \mathrm{E}_c$.

Following this model, it is possible to build the diagram in Figure 13 that shows the process generation for the electromagnetic shower. Considering an incident energy $\mathrm{E}_0 \gg \mathrm{E}_c$ the individual energy for each particle on each radiation length, see Equation 2.8, is determined by the Equation 2.7 (GRUPEN; SHWARTZ, 2008; MARTIN; SHAW, 2008).

Figure 13 – Illustration of an electromagnetic particle shower development in a sampling calorimeter. At each *bremsstrahlung* effect an electron loses energy as a photon and, each photon decays into an electron-positron pair taking into account the energy conservation principle.



Source: Povh et al. (2015).

$$E(t) = E_0\, 2^{-t}. \tag{2.7}$$

$$t = \frac{x}{X_0}. \tag{2.8}$$

The particle production processes will stop when $E(t) = E_c$, i.e. at $t = t_{max}$ then

$$t_{max} = t(E_c) = \frac{ln\left(\frac{E_0}{E_c}\right)}{ln2}. \tag{2.9}$$

The number of particles produced in the shower is determined by the Equation 2.10.

$$N_{\text{máx}} = \left(\frac{E_0}{E_c}\right). \tag{2.10}$$

In reality, the energy deposition profile is three-dimensional and can be modeled by three Probability Density Functions (PDFs), see Equation 2.11, that describes the longitudinal, radial, and azimuthal energy distribution. The ϕ distribution is uniform, then $f(\phi) = 1/2\pi$. The radial distribution measures the distance between the shower and the axis propagation on R_M units (GRINDHAMMER; RUDOWICZ; PETERS, 1990; GRINDHAMMER; PETERS, 1993). The local energy density deposited in the medium can be written as:

$$dE(\vec{r}) = Ef(t)dtf(r)drf(\phi)d\phi \tag{2.11}$$

The $f(t)$ function shown in Equation 2.13, parametrizes the average longitudinal profile of the energy deposition in an electromagnetic shower (NAVAS et al., 2024).

$$\left\langle \frac{1}{E_0}\frac{dE(t)}{dt} \right\rangle = f(t) \tag{2.12}$$

$$= \frac{(\beta t)^{\alpha-1}\beta e^{-\beta t}}{\Gamma(\alpha)}. \tag{2.13}$$

$$\text{where } \Gamma(\alpha) = \int_0^\infty \mathrm{e}^{-\alpha} x^{\alpha-1} dx. \tag{2.14}$$

In this profile, the energy deposit at the centroid is defined by Equation 2.15, and the point of maximum energy deposition by the Equation 2.16.

$$\langle t \rangle = \frac{\alpha}{\beta} \tag{2.15}$$

$$t_{max} = \frac{\alpha - 1}{\beta} = ln\left(\frac{E}{E_c}\right) + c_j; \quad j = \begin{cases} -0.5, & e^{\mp} \\ 0.5, & \gamma, \end{cases} \tag{2.16}$$

where α is the shape parameter, and β is the scaling parameter.

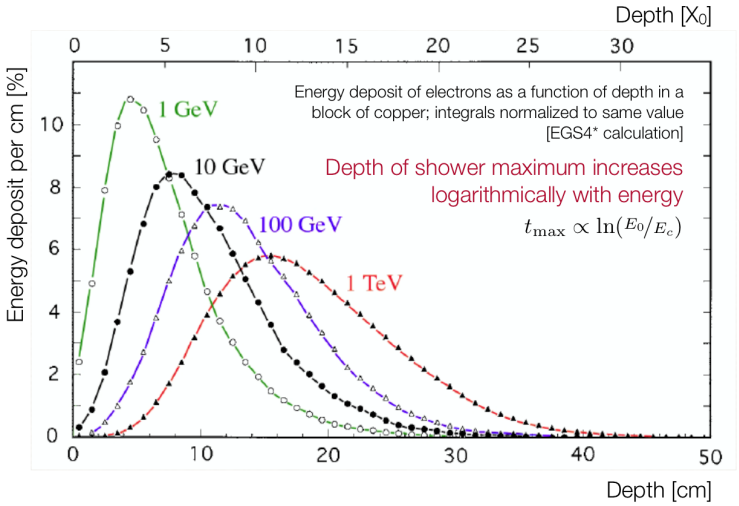
According to Grupen and Shwartz (2008) the assumptions used and the simple model for the particle cascade allow us to describe the main qualitative characteristics correctly:

- To absorb most of the energy of the incident photon, the total calorimeter thickness should be more than 10 – 15 X_0;
- The position of the maximum point for the shower increases slowly with energy. What could be visualized on Figure 14 where the energy magnitude increases by ten for each successive curve, but the peak position changes some units between consecutive curves;
- The energy leakage is caused mostly by soft photons escaping the calorimeter, mainly at the sides (lateral leakage) or the back (rear leakage) edges of a calorimeter cell.

Figure 14 shows the energy deposition along the longitudinal axis of the shower for four electromagnetic particles of increasing energy. It is interesting to observe that the difference between the lowest and highest profiles is three orders of magnitude, but the difference between peaks is ∼12 cm or ∼7 X_0 only.

Those profiles were obtained experimentally using a copper absorber. The first, 1 GeV, was totally absorbed on ∼23 cm; the second, 10 GeV, on ∼28 cm; the third, 100 GeV, on ∼33 cm, and the last, 1 TeV on ∼38 cm. This value, on average, is an increase of 3.5 cm of material for an increase of one order of magnitude in energy (LIVAN; WIGMANS, 2019).

Figure 14 – The energy deposit as a function of depth, for, 1 GeV, 10 GeV, 100 GeV e 1 TeV electron showers developing in a block of copper.



Source: Livan and Wigmans (2019).

The description for the radial profile is performed by an empirical function derived from different experimental data:

$$f(r) = \frac{1}{dE(t)} \frac{dE(t,r)}{dr} \tag{2.17}$$

This parameterization see Equation 2.18 (GRINDHAMMER; PETERS, 1993), relies on the assumption that two functions are necessary to describe that profile, one function describing the tail ($f_T(r)$) and another function describing the core ($f_C(r)$) of the shower. Those formulations were built using different materials (Fe, U, Cu, W, Pb). A parameter q is used to weight the relative intensity between tail and core,

$$\begin{aligned} f(r) &= q f_C(r) + (1-q) f_T(r). \\ &= q \frac{2 r R_C^2}{(r^2 + R_C^2)^2} + (1-q) \frac{2 r R_T^2}{(r^2 + R_T^2)^2} \end{aligned} \tag{2.18}$$

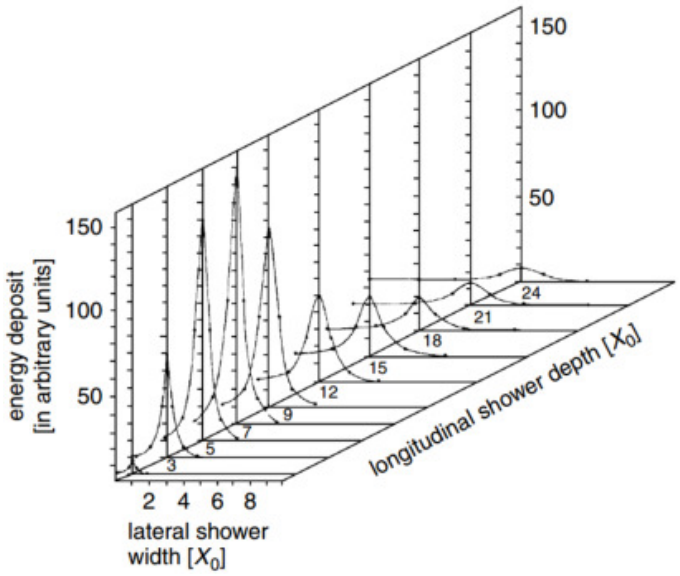
where,

$$0 \leq q \leq 1.$$

The core function $f_C(r)$ describes the vanishing shower development when the shower increases in depth. In the tail function, $f_T(r)$ ($r \gtrsim R_M$) is approximately flat in the shower at lower depths (1–2 X_0), stepper for moderate depths (5–6 X_0, 13–14 X_0) and returned to be almost flat at high depths (22–23 X_0).

In Appendix B is presented in detail the profile analysis and physical parameters that are used to describe the EM shower profiles and build the presented parameterization. Figure 15 shows both energy profiles development in function of X_0.

Figure 15 – Longitudinal and radial profiles of energy deposition in an electron shower (6 GeV) in lead. The EM shower development obeys the description used on formulation of Grindhammer and Peters (1993) with the maximum concentrated at low values for X_0.

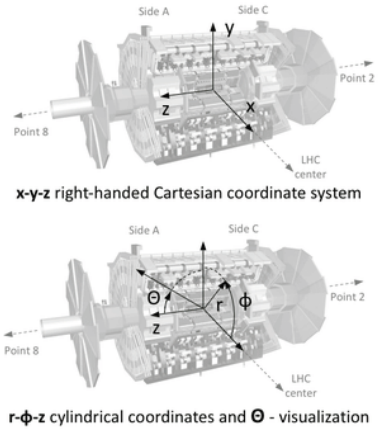


Source: Grupen and Shwartz (2008).

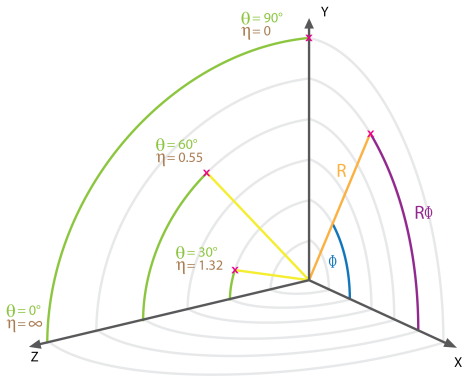
2.5.4 ATLAS Experiment

ATLAS is the largest experiment of the LHC. It has a cylindrical form, see Figure 16a, with 12.5 m of radius, a length of 44 m, and a mass of approximately 7000 tons (ATLAS Collaboration, 2016).

Figure 16 – Diagram of the detector coordinate system. The collision point is the center of the coordinating system and the origin of particles after a collision.



(a) Coordinating system.
Source: Extracted from Kuger (2017).



(b) Pseudorapidity (η) and angles θ e ϕ.
Source: Extracted from Lenzi (2013).

The coordinate system (see Figure 16a) used by the experiment defines the collision axis as the z-axis, having the nominal collision point in the center of the detector. ATLAS covers ~99% of the solid angle around the point of collision (ATLAS Collaboration, 2017).

Equation 2.19 defines the pseudorapidity η, and Equation 2.20, the azimuthal ϕ angles which indicates the particle direction after a collision. Figure 16b shows the correspondence of η for a few angles (θ). Equation 2.22 describes the transverse energy as the function of the θ angle and the energy from collision E;

$$\eta = -ln\left(\tan\left(\frac{\theta}{2}\right)\right), \tag{2.19}$$

$$\phi = \arctan\left(\frac{y}{x}\right), \tag{2.20}$$

$$\theta = \arctan\left(\frac{x}{z}\right), \tag{2.21}$$

$$E_T = Esin(\theta). \tag{2.22}$$

An essential parameter related to particle physics and accelerators is called luminosity. It is defined like in Equation 2.23 and quantifies the rate of particles that travel a cross-section at the collision point. Luminosity allows determining the number of events expected to be detected if the *pp* cross-section is known, like Equation 2.25.

$$\mathcal{L} = N f_{rev} \frac{n_1 n_2}{4\pi\sigma_x^*\sigma_y^*}\mathcal{F}\,[cm^{-2}\,s^{-1}]. \tag{2.23}$$

$$\mathcal{F} = \frac{1}{\sqrt{1+\left(\frac{\sigma_s}{\sigma_x^*}\frac{\theta_c}{2}\right)}} \tag{2.24}$$

$$N_{exp} = \sigma_{exp} \times \int \mathcal{L}(t)dt. \tag{2.25}$$

N indicates the number of the particle beam, for ATLAS, f_{rev}[35] number of turns that the particle beams develop in the rings of the LHC per second (Hz), n_1 and n_2 relative to the number of particles in the two beams, σ_x^* and σ_y^* characterize the Root mean squared (RMS) value of the transverse beam size of LHC at the collision point of the beam in both transverse directions. N_{exp} finally, corresponds to the observed number of events observed for a production process characterized by a σ_{exp} cross-section of interest (NAVAS et al., 2024). In Equation 2.24, $\mathcal{F}$ is a geometric factor related to crossing angle (θ_c), where σ_s is the RMS of the bunch length.

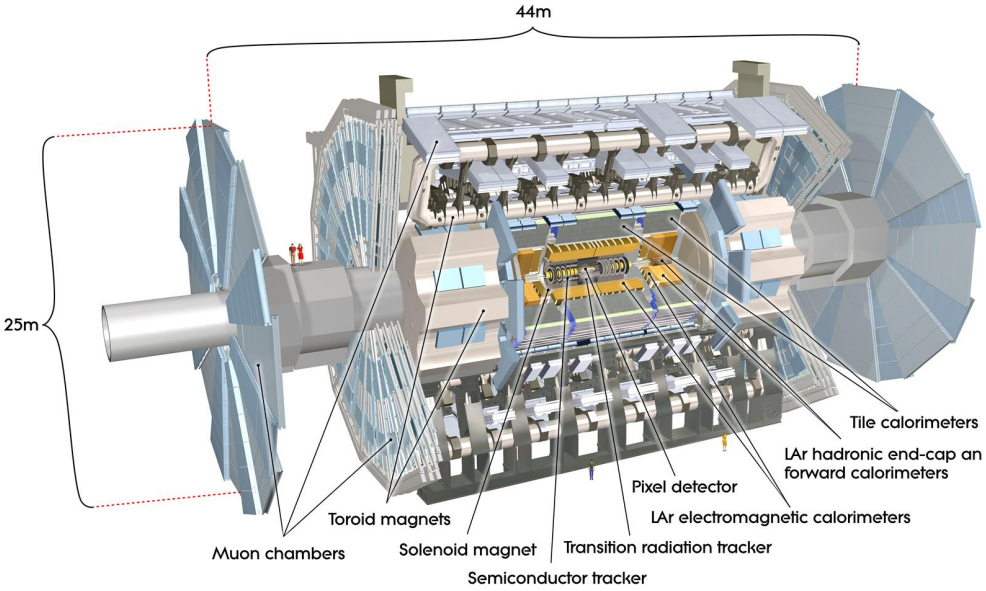
In the HL-LHC there is a forecast of updates reaching luminosity levels of up to 7.5×10^{34} cm^{-2}s^{-1}, while currently the ATLAS (LHC) reaches $2 \times 10^{34}\,cm^{-2}s^{-1}$ (APOLLINARI et al., 2017). This scenario will increase the challenges for the treatment of the information recorded in the calorimeter channels since the number of events rises, increasing the probability of more cells being sensitized and consequently increasing the influence of crosstalk incidence.

2.5.4.1 Sub-detectors

ATLAS is built from a set of sub-detectors (see Figure 17) that were designed and organized to absorb, record, and measure the properties of almost all particles produced in the *pp* collisions. The ATLAS sub-detectors are a stack of tracking systems, Inner detector (ID), the Electromagnetic Calorimeter (EMCAL), Hadronic Calorimeter (Hadronic Endcap Calorimeter (HEC)) and Muon Chamber. With this set of sub-detectors, it is possible to study a large number of particles (see Figure 18) due to the complementary information produced by each sub-detector.

[35] $f_{rev} = c/26659 = 11245.45$ Hz.

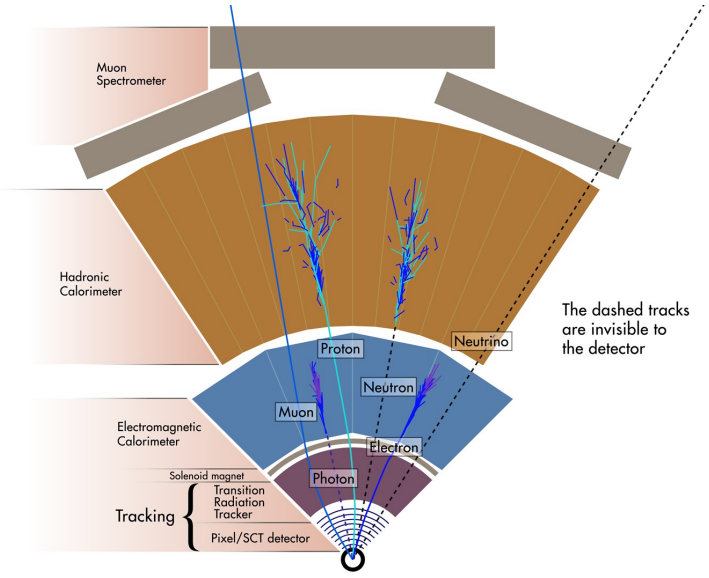
Figure 17 – ATLAS experiment diagram with emphasis to the sub-detectors EMB: LAr Electromagnetic Calorimeter; HEC: Tile Calorimeter; ID: compound by the Semiconductor tracker, Transition radiation tracker and Pixel detector, and the EMEC: LAr hadronic end-cap and forward calorimeters.



Source: Nikiforou (2013).

Figure 18 shows a sector in a transverse view of ATLAS experiment with each sub-detector and the respective specific particle interaction that occurs along their path in the ATLAS detectors after a collision. The central black circle depicts the tunnel where bunches are accelerated and collide.

Figure 18 – Cross-section view that shows particles interacting with ATLAS sub-detectors. At the bottom, is the beam tunnel, where the bunch crossing interacts. This is the detector region with the high granularity in the experiment. On the sequence the electromagnetic and hadronic calorimeters, and the most external the muon chambers.



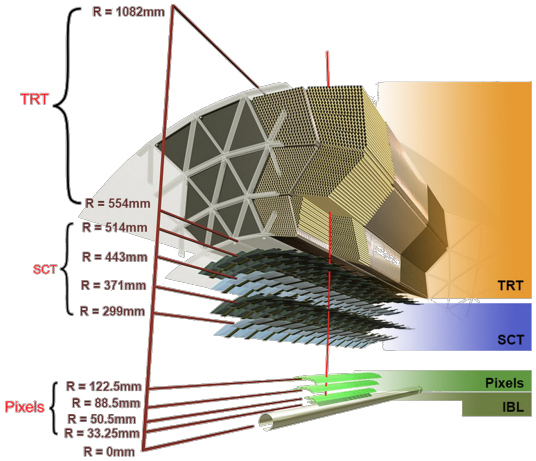
Source: Adapted from Pequenao and Schaffner (2013).

In the following ATLAS inner detector and calorimeter systems are briefly described.

Inner detector (ID)

Figure 19 presents the ID which consists of three sub-detectors utilizing three technologies: silicon pixel detectors, silicon strip detectors, and straw drift tubes. The most high-granularity of the ATLAS experiment is in this detector, around 80 million channels. The ID was designed to provide precise and efficient measurements of charged particles as they traverse the detector. The ID has full coverage in ϕ for $|\eta| < 2.5$, and is the most granular region of the ATLAS experiment, with about 86.4 million channels. (ATLAS Collaboration, 2010; FREUND, 2011; AAD; et al, 2020).

Figure 19 – Inner Detector and their sub-detectors: Insertable B-layer (IBL), Pixel detector, Semiconductor Tracker (SCT) and Transition Radiation Tracker (TRT).



Source: ATLAS Collaboration (2020b).

Calorimeters

Immediately after the ID, the calorimeters are installed. The first is the electromagnetic with four layers, Pre sampler (PS), first sampling layer from electromagnetic calorimeter (S1), second sampling layer from electromagnetic calorimeter (S2), third sampling layer from electromagnetic calorimeter (S3), while the second part is a set of hadronic layers, Tiles 0, Tiles 1, and Tiles 2. These areas are a total of seven layers, as it is displayed in Figure 20. Their respective coverage areas, and granularity, for a total of 187,652 single readout cells (AAD et al., 2017) are presented in Table 4.

Figure 20 – Illustration of the electromagnetic and hadronic calorimeter layers and their respective granularity. The red cells indicate the position of the second sampling layer from electromagnetic calorimeter (S2) hottest cell through layers for a given detected shower.

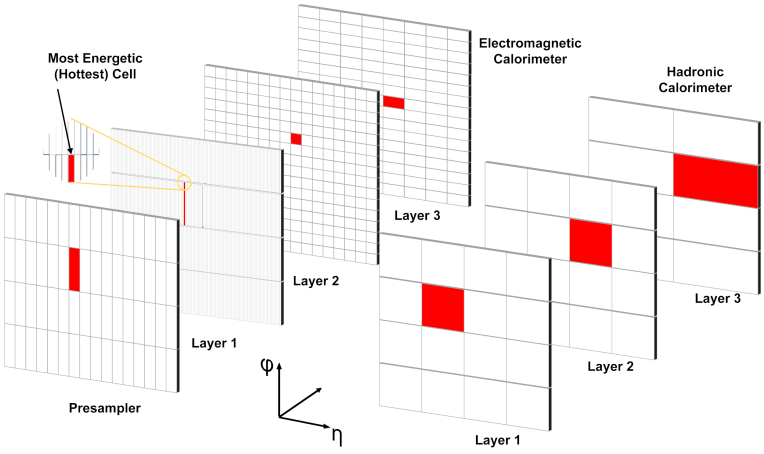
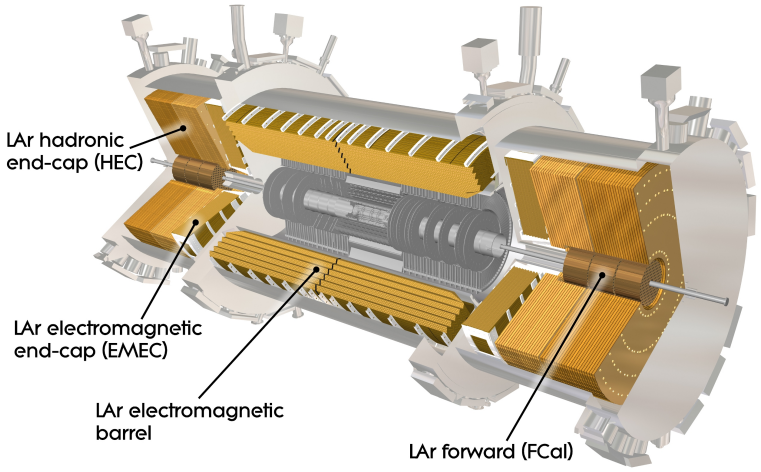


Figure 21a presents all LAr calorimeter modules of the ATLAS experiment. Two types of calorimeters are in ATLAS; electromagnetic and hadronic. As the electromagnetic particles deposit their energy at low X_0, this calorimeter is designed to be close to the tunnel where the proton beam travels to collide, on the sequence is the hadronic calorimeter. A second piece of information is that the electromagnetic LAr calorimeter has the highest granularity as is described in Table 4.

Figure 21 – Details for the set of calorimeters, in yellow, inside the ATLAS experiments.

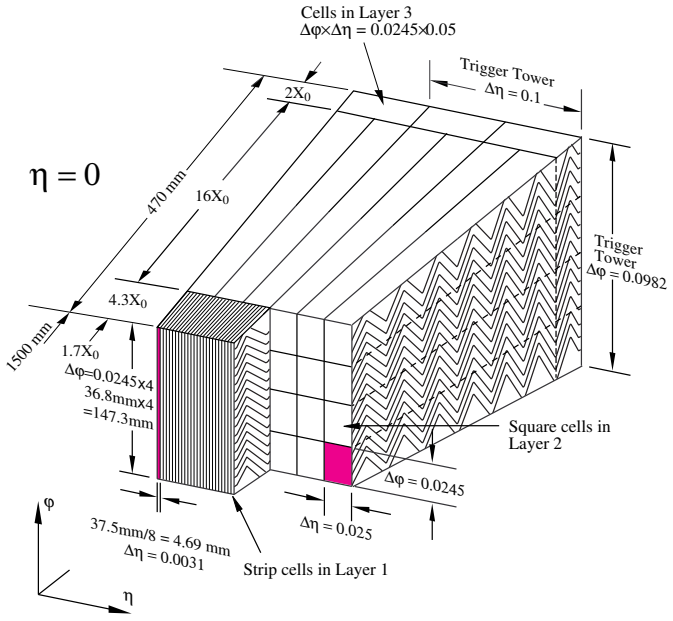


(a)

Source: Nikiforou (2013).

Figure 22a presents the accordion structure used in the electromagnetic barrel for its three EM layers, S1, S2, and S3. The total radiation length (X_0) in those layers is 24 X_0, with the high X_0 in the S2 once the EM longitudinal profile for energy deposition is concentrated with this X_0 value.

Figure 22 – Illustration of the accordion structure of the electromagnetic barrel. Provide details on the granularity among the sampling layers. The first has 4.3X_0 with its strip cells, at the sequence, S2 with squared cells in (η, ϕ), and 16X_0, and the last one, S3 with 2X_0.



(a)

Source: Nikiforou (2013).

Table 4 – Detail for ATLAS calorimeters: Coverage region in η, number of cells and granularity.

Calorimeter		Module	$N_{\text{cell.}}$	η_{region}	$\Delta\eta \times \Delta\phi$
Electromagnetic	Barrel	Presampler	7.808	$\lvert\eta\rvert < 1.520$	$0.025 \times \frac{\pi}{32}$
		EMB	109.568	$\lvert\eta\rvert < 1.520$	
		EMB1		$\lvert\eta\rvert < 1.400$	$\frac{0.025}{8} \times \frac{\pi}{32}$
				$1.400 < \lvert\eta\rvert < 1.475$	$0.025 \times \frac{\pi}{128}$
		EMB2		$\lvert\eta\rvert < 1.400$	$0.025 \times \frac{\pi}{128}$
				$1.400 < \lvert\eta\rvert < 1.475$	$0.075 \times \frac{\pi}{128}$
		EMB3		$\lvert\eta\rvert < 1.350$	$0.050 \times \frac{\pi}{128}$
	Endcap	Prsampler	1.536	$1.500 < \lvert\eta\rvert < 1.800$	$0.025 \times \frac{\pi}{32}$
		EMEC	64.744	$1.375 < \lvert\eta\rvert < 3.200$	
		EME1		$1.375 < \lvert\eta\rvert < 1.425$	$0.050 \times \frac{\pi}{32}$
				$1.425 < \lvert\eta\rvert < 1.500$	$0.025 \times \frac{\pi}{32}$
				$1.500 < \lvert\eta\rvert < 1.800$	$\frac{0.025}{8} \times \frac{\pi}{32}$
				$1.800 < \lvert\eta\rvert < 2.000$	$\frac{0.025}{6} \times \frac{\pi}{32}$
				$2.000 < \lvert\eta\rvert < 2.400$	$\frac{0.025}{4} \times \frac{\pi}{32}$
				$2.400 < \lvert\eta\rvert < 2.500$	$0.025 \times \frac{\pi}{32}$
				$2.500 < \lvert\eta\rvert < 3.200$	$0.100 \times \frac{\pi}{32}$
		EME2		$1.375 < \lvert\eta\rvert < 1.425$	$0.050 \times \frac{\pi}{128}$
				$1.425 < \lvert\eta\rvert < 2.500$	$0.025 \times \frac{\pi}{128}$
				$2.500 < \lvert\eta\rvert < 3.200$	$0.100 \times \frac{\pi}{128}$
		EME3		$1.500 < \lvert\eta\rvert < 2.500$	$0.050 \times \frac{\pi}{128}$
Hadronic	Barrel	Tiles	2.880	$\lvert\eta\rvert < 1.000$	
		Tiles 0/1			$0.1 \times \frac{\pi}{32}$
		Tiles 2			$0.2 \times \frac{\pi}{32}$
	Barrel Ext.	Tiles	2.304	$0.800 < \lvert\eta\rvert < 1.700$	
		Tiles 0/1			$0.1 \times \frac{\pi}{32}$
		Tiles 2			$0.2 \times \frac{\pi}{32}$
	Endcap	HEC	5.632	$1.500 < \lvert\eta\rvert < 3.200$	
		HEC0		$1.500 < \lvert\eta\rvert < 2.500$	$0.1 \times \frac{\pi}{32}$
				$2.500 < \lvert\eta\rvert < 3.200$	$0.2 \times \frac{\pi}{16}$
	Forward	FCAL	3.524	$3.100 < \lvert\eta\rvert < 4.900$	$\Delta x \times \Delta y$
		FCAL0		$3.100 < \lvert\eta\rvert < 3.150$	1.5 cm × 1.3 cm
				$3.150 < \lvert\eta\rvert < 4.300$	3.0 cm × 2.6 cm
				$4.300 < \lvert\eta\rvert < 4.830$	1.5 cm × 1.3 cm
		FCAL1		$3.200 < \lvert\eta\rvert < 3.240$	1.7 cm × 2.1 cm
				$3.240 < \lvert\eta\rvert < 4.500$	3.3 cm × 4.2 cm
				$4.500 < \lvert\eta\rvert < 4.810$	1.7 cm × 2.1 cm
		FCAL2		$3.290 < \lvert\eta\rvert < 3.320$	2.7 cm × 2.4 cm
				$3.320 < \lvert\eta\rvert < 4.600$	5.4 cm × 4.7 cm
				$4.600 < \lvert\eta\rvert < 4.750$	2.7 cm × 2.4 cm

Source: Aad et al. (2017).

The resolution of the different ATLAS calorimeters can be modeled using Equation 2.26[36]. The first term, *a*, is the statistical term and carries information about the intrinsic fluctuation of a shower, either hadronic or EM. This term is of lower magnitude in the homogeneous calorimeters than in sampling calorimeters. The *b* term is associated with

[36] $m \oplus n = \sqrt{m^2 + n^2}$

electronic noise from the preamplifier and shaping circuits, sampling noise, and pileup. The last term, c, is a constant that covers the non-uniformity, calibration uncertainties of the detector, and possible damages caused to the active medium due to radiation resulting from high energy experiments accumulated over time (more important for solid detectors) (ATLAS Collaboration, 2017; CAVALLARI, 2011; NAVAS et al., 2024).

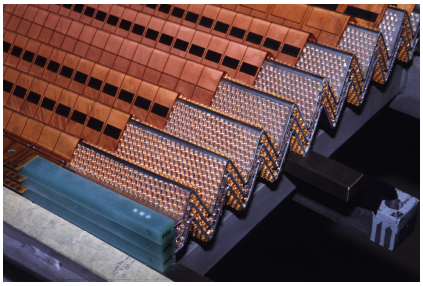
$$\frac{\sigma}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \tag{2.26}$$

Figure 23a presents a section of the accordion sampling calorimeter. This shape is to avoid dependence of the response on transverse position (BROWN; COCKERILL, 2012). Figure 23b shows a diagram of the internal stack LAr layers structure is composed by the elements (AUBERT et al., 2005; NIKIFOROU, 2013):

- Absorbers - a passive layer in the lead glued between two thin stainless sheets are 2.2 mm thick plates in an accordion shape, in gray shaded color;
- Active medium - in blue shaded is the liquid Argon;
- Electrodes - made of three copper layers insulated by two layers of polyimide;
- Intergap spacers - The inter gap spacers made of honeycomb structured NOMEX®[37] paper impregnated with phenolitic resin, keep the electrodes centered in the gap between two absorbers

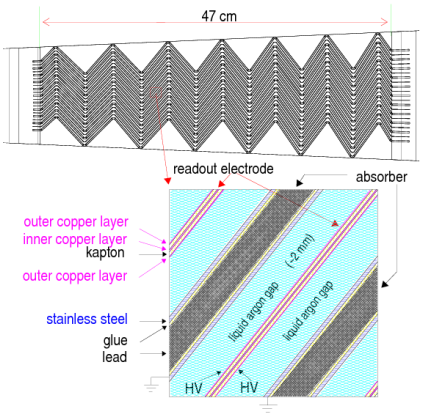
Figure 23 – The accordion sample calorimeter structure of the ATLAS experiment. (a) Image of a part of the accordion of the ATLAS; (b) detail of the layers of the accordion structure and of its charge collection gap structure.

(a)



Source: Hedberg (2009).

(b)



Source: Nikiforou (2013).

2.5.5 Timing, Synchronization and Proton Bunch Crossing Collision

A set of timing and synchronization control systems generates the reference signals using a pulse machine to guarantee that ATLAS and the other experiments are synchronized. In the LHC, this structure is called Trigger Timing and Control system (TTC). TTC is responsible for producing and distributing the reference signal for a high number of electronic circuits of the ALICE, ATLAS, CMS, and LHC-B experiments (BAILLIE et al., 1997; ASHTON et al., 2000). All LHC experiments receive a synchronism signal derived from the pulse machine, whose frequency is 40 MHz (25 ns). This signal is distributed through optical fibers to reduce losses and delays affecting the reading and recording of events with a clock jitter lower than 20 ps.

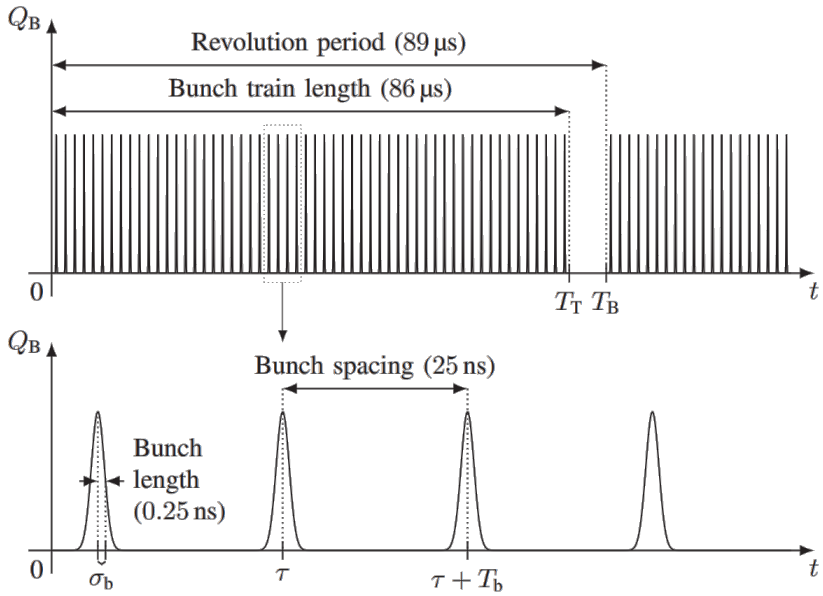
The proton beams are organized in 25 ns intervals with a spread $\sigma_b = 0.25$ ns, see Figure 24. In the context of this thesis, it means that the absolute timing of a given particle suffers from uncertainty.

The proton bunch collisions are tracked using a unique 12-bit identifier called Bunch Crossing Identifier (BCID). This identification is intended to ensure that the recorded event comes from a collision and not from external interactions,

[37] is a registered trademark of DuPont Company.

such as a cosmic ray. The structure of holes between the packets allows for the data acquisition system synchronization, and other gaps are used to adjust some necessary delays from dense electronic readout cell structure (VARELA, 2000; TAYLOR, 2002; BUCHANAN et al., 2008a; KRUPA; SOBY, 2013). On LHC, an orbit consists of 3564 periods that are called *bunches* although some of them do not contain protons. This structure synchronizes bunches and all LHC experiments. Due to this, there are gaps between proton bunches to synchronize the trigger and Data acquisition (DAQ) data. Another gap is used to adjust delays from the dense electronic readout cell structure.

Figure 24 – Diagram of a period of revolution of proton bunches within the rings of the LHC. The y axis indicates the total beam charge computed as $\int_0^\tau i_B \cdot dt$.



Source: Krupa and Soby (2013).

2.5.6 Energy Reconstruction Process

In ATLAS, the energy, and time of flight values are estimated using four [38] samples of the front end electronic signal (see the shaped signal ($g(t)$) on Figure 25a)) obtained at a sampling time equal to 25 ns. Based on these samples and a few calibration constants, see Equation 2.27 the energy deposited in a specific cell is estimated.

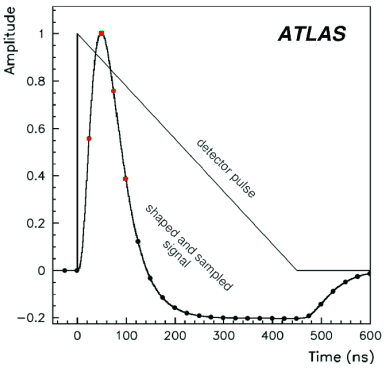
$$E_{cell} = \underbrace{F_{\mu A \to MeV}}_{\text{(i)}} \cdot \underbrace{F_{DAC \to \mu A}}_{\text{(ii)}} \cdot \underbrace{\frac{1}{\frac{M_{phys}}{M_{cal}}}}_{\text{(iii)}} .R. \underbrace{\sum_{i=1}^{4} a_i (s_i - p)}_{\text{(iv)}} \tag{2.27}$$

in which (i) refers to the sampling fraction that associates the ionization current with energy; (ii) comes from the calibration board; (iii) indicates the relationship between the ionization signal and the calibration signal (see Figure 25b). F is conversion factors to adjust units, and R is the electronic gain. (iv) describes the range of energy reconstructed from the coefficients (a_i) of the Optimal Filtering (OF)[39] in Analogue to Digital Converter (ADC) counts, based on the samples (s_i). p is the pedestal. The pedestal value is estimated by averaging over all events and all ADC samples per channel in each gain (ABDELALIM, 2008).

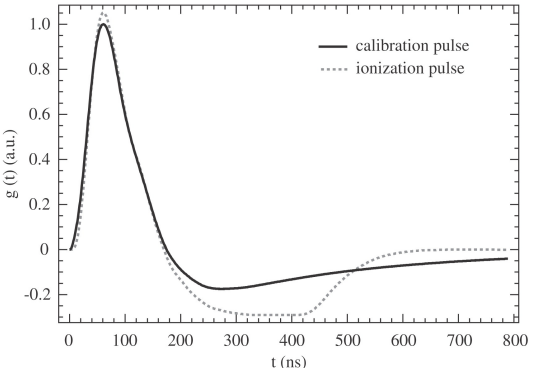
[38] In Run 1, 5 samples were used, in Run 2, were reduced to 4 to increase the level 1 (L1) frequency of the trigger from 75 kHz to 110 kHz (ATLAS Collaboration, 2017)

[39] Details on subsubsection 2.5.6.1.

Figure 25 – (a) representation of the physical (due to particle ionization) pulse and the ionization signal in a sensitized cell of the LAr calorimeter. The energy is primarily proportional to the integral of the ionization signal (triangular shape). In the sequence, the triangular signal is reshaped to a signal whose energy is proportional to the maximum of the signal peak. This new shape concentrates the energy in a short time with a peak of around 75 ns. The sampling interval equals 25 ns, synchronized by the TTC. The maximum number of samples during normal data taking is 32, but only the first 4 are stored for energy and time estimation. In (b), there are differences between the calibration and ionization signals which justify term (iii) in Equation 2.27.



(a)
Source: Nikiforou (2013).



(b)
Source: Aharrouche et al. (2007).

The energy amplitude (A) is estimated as displayed in Equation 2.28, and the time of flight (τ) is calculated by the Equation 2.29. The signal samples are presented by s_i, p is the pedestal in ADC for unipolar representation of the cell signal, and the terms a_i, b_i are the coefficients of the Optimal Filtering (OF), a linear filter, fast and efficient to implement which are described in subsubsection 2.5.6.1. OF formulation uses the knowledge of the standard pulse-shape functions to derive the filter coefficients a_i and b_i, so its accuracy is sensitive to pulse-shape variations.

2.5.6.1 Optimal Filtering

Optimal filters are digital signal processing techniques that use linear approximators. They are used to estimate the energy and time-of-flight values of the particles in the in-cell reconstruction process of the ATLAS experiment calorimeter. The process employs the use of Lagrange multipliers to obtain the coefficients associated with the amplitude of energy and time of flight, describing them as a linear combination between the coefficients and the signal samples, minimizing the influence of noise present in the samples (CLELAND; STERN, 1994).

In the Equation 2.28, it is shown how the amplitude value of the energy is obtained, and in the Equation 2.29 how the time of flight value of the particle is obtained. N is the number of signal samples, a_i are the optimal coefficients for the amplitude while b_i is the coefficients for the time, and s_i is the signal samples.

$$A = \sum_{i}^{N} a_i s_i \tag{2.28}$$

$$A\tau = \sum_{i}^{N} b_i s_i \tag{2.29}$$

We use the signal information from the cell under analysis to obtain the coefficients, $g(t_i)$. The cell signal is normalized together with the same number of samples as the signal of its first derivative, $\dot{g}(t_i)$. In the Equation 2.30 is shown the samples S_i for the reconstruction process, the $n(t)$ signal represents the noise contribution function in time. The Equation 2.31 displays the function in its approximation with its first Taylor series expansion. The Equation 2.31 describes the samples of the signal, the first derivative, and the noise in vector form.

$$s_i = Ag(t-\tau) + p \tag{2.30}$$

$$s_i \simeq A\left[g(t) - \tau\dot{g}(t)\right] + n(t) + p \tag{2.31}$$

$$\mathbf{s} = A\,\mathbf{g} - A\,\tau\,\dot{\mathbf{g}} + \mathbf{n} \tag{2.32}$$

The Equation 2.32 represents the sampling vector of the signals present in a calorimeter cell. After the Taylor expansion The Equation 2.28 and Equation 2.29 becomes the Equation 2.33 and Equation 2.34 respectively, with $< \mathbf{u} >$ and $< \mathbf{v} >$ being the average values for A and $A\tau$.

$$< \mathbf{u} >= \sum_i \mathbf{a}_i \left(Ag(t_i) - A \cdot \tau \frac{g(t_i)}{dt} + < \mathbf{n}_i > \right) \tag{2.33}$$

$$< \mathbf{v} >= \sum_i \mathbf{b}_i \left(Ag(t_i) - A\tau \frac{g(t_i)}{dt} + < \mathbf{n}_i > \right) \tag{2.34}$$

To determine the value of the coefficients a_i and b_i it is necessary to attend the constraints of the Equation 2.37 by applying Lagrange multipliers to minimize the variance of parameters $< \mathbf{u} >$ and $< \mathbf{v} >$, see Equation 2.35 and Equation 2.36, respectively. On average, the noise contribution, $< \mathbf{n}_i >$, is zero, and the term $\mathbf{R}_{ij}$ represents the autocorrelation matrix of the noise.

$$Var(\mathbf{u}) = \sum_i \mathbf{a}_i\mathbf{a}_j < \mathbf{n}_i\mathbf{n}_j >= \sum_i \mathbf{a}_i\mathbf{a}_j\,\mathbf{R}_{ij} \tag{2.35}$$

$$Var(\mathbf{v}) = \sum_i \mathbf{b}_i\mathbf{b}_j < \mathbf{n}_i\mathbf{n}_j >= \sum_i \mathbf{b}_i\mathbf{b}_j\,\mathbf{R}_{ij} \tag{2.36}$$

$$\sum_i a_i g_i = 1\,; \qquad \sum_i a_i \frac{dg(t_i)}{dt} = 0\,; \qquad \sum_i b_i g_i = 0\,; \qquad \sum_i b_i \frac{dg(t_i)}{dt} = -1 \tag{2.37}$$

The functions to be minimized are described below in Equation 2.38 and Equation 2.39, λ, κ, ι, and ρ being the associated Lagrange multipliers.

$$I_u = \sum_{ij} \mathbf{R}_{ij}\,\mathbf{a}_i\,\mathbf{a}_j - \lambda\left(\sum_i \mathbf{a}_i\,\mathbf{g}_i - 1\right) - \kappa\sum_i \mathbf{a}_i\,\dot{\mathbf{g}}_i \tag{2.38}$$

$$I_v = \sum_{ij} \mathbf{R}_{ij}\,\mathbf{b}_i\,\mathbf{b}_j - \iota\sum_i \mathbf{b}_i\,\mathbf{g}_i - \rho\left(\sum_i \mathbf{b}_i\,\dot{\mathbf{g}}_i + 1\right) \tag{2.39}$$

Calculating $\partial I_u\,|_{\mathbf{a}_i}$ e $\partial I_v\,|_{\mathbf{b}_i}$, results in the Equation 2.40 and the Equation 2.41, respectively.

$$\frac{\partial I_u}{\partial \mathbf{a}_i} = \sum_j \mathbf{R}_{ij}\,\mathbf{a}_j - \lambda\,\mathbf{g}_i - \kappa\,\dot{\mathbf{g}}_i = 0, \tag{2.40}$$

$$\frac{\partial I_v}{\partial \mathbf{b}_i} = \sum_j \mathbf{R}_{ij}\,\mathbf{b}_j - \iota\,\mathbf{g}_i - \rho\,\dot{\mathbf{g}}_i = 0, \tag{2.41}$$

and solve to find the **a**[40] and **b**[41] values results in:

$$\mathbf{a} = \lambda\,\mathbf{R}^{-1}\mathbf{g} + \kappa\,\mathbf{R}^{-1}\dot{\mathbf{g}}, \tag{2.42}$$

$$\mathbf{b} = \iota\,\mathbf{R}^{-1}\mathbf{g} + \rho\,\mathbf{R}^{-1}\dot{\mathbf{g}}. \tag{2.43}$$

[40] vector amplitude coefficients.
[41] vector time of fligth coefficients.

The values of the respective multipliers λ, κ, ι, and ρ are determined using the relationships of Equation 2.44, Equation 2.45, Equation 2.46, Equation 2.47 e Equation 2.48.

$$\lambda = \frac{Q_2}{\Delta}\ ; \quad \kappa = \frac{-Q_3}{\Delta}; \quad \iota = \frac{Q_3}{\Delta}\ ; \quad \rho = \frac{-Q_1}{\Delta} \tag{2.44}$$

$$Q_1 = \mathbf{g}^T\,\mathbf{R}^{-1}\mathbf{g} \tag{2.45}$$

$$Q_2 = \dot{\mathbf{g}}^T\,\mathbf{R}^{-1}\,\dot{\mathbf{g}} \tag{2.46}$$

$$Q_3 = \dot{\mathbf{g}}^T\,\mathbf{R}^{-1}\,\mathbf{g} \tag{2.47}$$

$$\Delta = Q_1\,Q_2 - Q_3^2 \tag{2.48}$$

After the minimization process concerning constraints, we are backing to the Equation 2.28, and Equation 2.29 to determine the amplitude (A) and time (τ) values.

2.6 Energy and Time reconstruction without noise

For the hypothetical case that we do not have noise contribution, i.e., noise equals zero and no XT influence, we can reconstruct energy and time just by considering the signal samples.

Taking into account the Equation 2.32, in terms of energy cell deposited and the cell sampling transfer function, the noise term and $\tau = 0$ is zero, so for the energy, we can compute it using Equation 2.50 the energy value from each sample.

$$s_i = E^{cell_j} g(t_i) \tag{2.49}$$

$$E^{cell_j} = \frac{s_i}{g(t_i)} \tag{2.50}$$

To evaluate the average truth energy value, the Equation 2.51.

$$E^{cell_j} = \frac{1}{4}\sum_{i=1}^{4}\frac{s_i}{g(t_i)} \tag{2.51}$$

For the time reconstruction, is used the Equation 2.32 to determine the fly time (τ). The time information is achieved by solving the linear system, see Equation 2.52, for the cell samples:

$$S = \begin{cases} s(t_1) = E^{cell_j}\cdot g(t_1) - E^{cell_j}\cdot \dot{g}(t_1)\cdot\tau \\ s(t_2) = E^{cell_j}\cdot g(t_2) - E^{cell_j}\cdot \dot{g}(t_2)\cdot\tau \\ s(t_3) = E^{cell_j}\cdot g(t_3) - E^{cell_j}\cdot \dot{g}(t_3)\cdot\tau \\ s(t_4) = E^{cell_j}\cdot g(t_4) - E^{cell_j}\cdot \dot{g}(t_4)\cdot\tau \end{cases} \tag{2.52}$$

So, to reconstruct τ, just pick up one of the equivalent solutions for the Equation 2.52 that the Equation 2.53 shows.

$$\tau = \frac{1}{E^{cell_j}} \cdot \begin{cases} \frac{g(t_2)\cdot s(t_1) - g(t_1)\cdot s(t_2)}{g(t_2)\cdot \dot{g}(t_1) - g(t_1)\cdot \dot{g}(t_2)} \\ \\ \frac{g(t_2)\cdot s(t_3) - g(t_3)\cdot s(t_2)}{g(t_2)\cdot \dot{g}(t_3) - g(t_3)\cdot \dot{g}(t_2)} \\ \\ \frac{g(t_2)\cdot s(t_4) - g(t_4)\cdot s(t_2)}{g(t_2)\cdot \dot{g}(t_4) - g(t_4)\cdot \dot{g}(t_2)} \\ \\ \frac{g(t_3)\cdot s(t_1) - g(t_1)\cdot s(t_3)}{g(t_3)\cdot \dot{g}(t_1) - g(t_1)\cdot \dot{g}(t_3)} \\ \\ \frac{g(t_4)\cdot s(t_1) - g(t_1)\cdot s(t_4)}{g(t_4)\cdot \dot{g}(t_1) - g(t_1)\cdot \dot{g}(t_4)} \\ \\ \frac{g(t_4)\cdot s(t_3) - g(t_3)\cdot s(t_4)}{g(t_4)\cdot \dot{g}(t_3) - g(t_3)\cdot \dot{g}(t_4)} \end{cases} \tag{2.53}$$

Pile-up of Event ($< \mu >$)

Event pile-up is a relevant phenomenon for High energy physics (HEP) experiments, due to high event rates that originate from particle collisions. The pile-up can occur when additional particle interactions happen in the same bunch-crossing at the vicinity of the main event (in-time pile-up) or when a given cell receives multiple particle interactions in a time interval smaller than the cell pulse duration. In this last case, the spurious information comes from different bunch-crossings generating the so-called out-of-time pile-up.

The average number of particle interactions per bunch-crossing can be calculated using Equation 2.54. Where, $\mathcal{L}_0$ is the instantaneous luminosity, $\sigma_{inelast}$ is that inelastic *pp* cross section, n_c, is the number of bunches pairs of particles colliding in the LHC and f_{rev} the revolution frequency of the bunches equal to 11.245 kHz.

Pile-up distorts the standard pulse shape in calorimeter cells and introduces correlation and non-linearities reducing the accuracy in A and τ estimation from OF. The XT mitigation could increase the pile-up treatment efficiency, as it is intended to reduce the fluctuations in particle time of flight and, consequently, increase the identification of an interesting event.

$$< \mu >= \frac{\mathcal{L}_0 \, \sigma_{\text{inelast.}}}{n_c f_{\text{rev}}} \tag{2.54}$$

2.6.1 Crosstalk

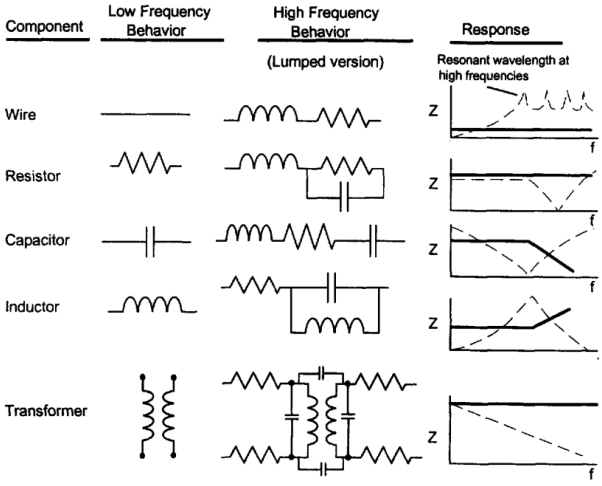
The crosstalk (XT) is the term that indicates the existence of interference in a communication channel coming from its proximity; that is, part of the current from other channels contaminates the channel of interest and distorts the observed signal. This phenomenon is also present in the ATLAS LAr detector in the three possible forms of coupling: resistive, capacitive, and inductive, of different magnitudes.

Crosstalk is a well-known and studied electrical phenomenon resulting from different physical effects. In circuits where line signals are close to each other, the magnetic field produced by one line couples to the neighbor's lines causing a parasitic inductive signal that distorts the nominal signal driven by the line. Another source arises from the effective capacitance between charge conductors separated by a short distance, like a dielectric, which forms a coupling capacitance. Those effects are present in electric circuits that operate in high voltage and current at low frequency, like power transmission lines (SILVA et al., 2011; RAUNIG et al., 2013).

In power transmission lines, the XT (mutual inductive and capacitive coupling) produces disturbances like circulating currents, reduction of the transport capacity of the circuits, which generate losses, and operation problems in high voltage systems (RAUNIG et al., 2013). In circuits at low voltage/current values, the XT occurs thanks to the effects of high frequency over passive components. To avoid this situation, the circuit project and Printed Circuit Board (PCB) layout should consider the model used for the passive components at high frequency to minimize the XT

effects. Figure 26 shows the effective lumped[42] models for the passive components at low and high frequency and their respective impedance curves.

Figure 26 – Models for wire, resistor, capacitor, inductor, and transformer at low and high frequency and their respective impedance curves according to frequency variation. To project circuits for communications, and PCB design these considerations are taking into account to minimizing the XT effects.



Source: Montrose (1998).

On communications systems and PCB projects and design, those considerations are essential to guarantee a good operation for systems avoiding interference between close channels and digital circuits that operate at high frequency like microcontrollers and processors. On those applications, the circuits are designed concerning Electromagnetic Interference (EMI) and Electromagnetic Compatibility (EMC) rules, which define criteria for electric and electronic circuits to operate reducing XT influence (MONTROSE, 1998; DĂRĂBAN; PITICĂ, 2010).

To study XT in the ATLAS calorimeter, three definitions are adopted to evaluate its level of magnitude in one cell contaminated by neighboring cells. Figure 27 shows graphically the XT observed in ATLAS LAr calorimeter (COLAS et al., 1999; HUBAUT, 2000):

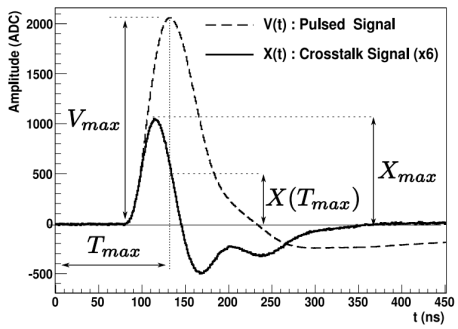
1. Peak-Peak – $\frac{X_{\max}}{V_{\max}}$ – Refers to the percentage of the ratio between the maximum XT amplitude and the maximum cell signal amplitude;

2. Under Peak – $\frac{X(T_{max})}{V_{\max}}$ - Refers to the percentage of the ratio between the XT amplitude at the time where the signal cell amplitude is the maximum;

3. Energy Contamination – In this approach (LOPEZ SOLIS, 2017) XT contributions are evaluated applying the OF coefficients and calculating $\frac{E_{XT}}{E_{cell}}$

[42] An approach for modeling and analysis considers the concentrate parameters, meaning that resistance, inductance, capacitance, and gain are concentrated into the component model.

Figure 27 – Diagram explaining the two distinct crosstalk definitions: Peak-Peak, $X_{\max}/V_{\max}$ and under the peak, i.e., the value of XT for the maximum signal instant in the sensitized cell, $X(T_{\max})/V_{\max}$.



Source: Labbé and Ishmukhametov (2008).

In Figure 28, the effect of XT on the signal of a cell is displayed. The lower magnitude signal represents the XT signal. The intermediate magnitude signal is the signal of the sensitized cell; however, with the addition of the XT, we observe the signal amplitude and peak displacement increase. According to LOPEZ SOLIS (2017), although the under-peak calculation represents a more realistic value for the XT, it is less stable due to its dependence on fluctuations in the XT signal, while the peak-peak is more stable and represents the most pessimistic value for interference. Nevertheless, the critical information for physics is how the XT impacts E and time reconstruction. So, the use of the energy contamination method implements the practical effects of XT for ϕ.

Figure 28 – Diagram explaining the two distinct crosstalk evaluation parameters in an LAr cell: Peak-Peak, and under-peak. At the left, is a representation of all signals, and at the right is a detail to explain the XT definitions. In red dot indicates (X_{max}) the maximum amplitude of the crosstalk. The blue dot indicates (V_{max}) the maximum amplitude for the cell signal. The black dot indicates the crosstalk ($X(T_{max})$) at the time when the direct signal is maximum (V_{max}). The green line indicates the new cell signal including the crosstalk contribution which has a new peak value over the cell signal which will be processed to estimate the energy and time of the particle.

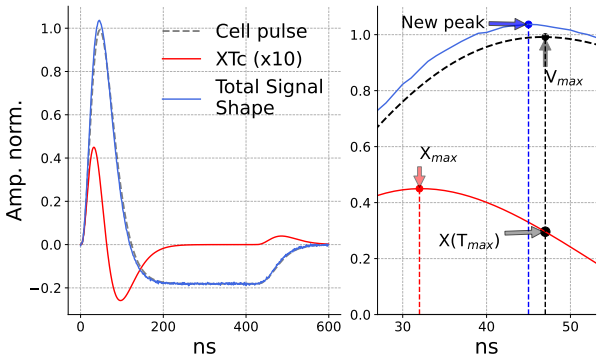
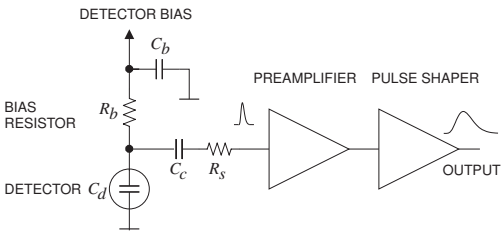


Figure 29a shows the typical circuit diagram for collecting the signal from the detector and the equivalent circuit for an LAr cell. The detector is represented by C_d, its input coupling capacitor, R_b and C_b, are the resistor and polarizing capacitor, and the R_s represents the set of input resistors present in the preamplifier input circuit. In Figure 29b, the equivalent circuit diagram for an LAr cell indicates the input of calibration, and the physics signals.

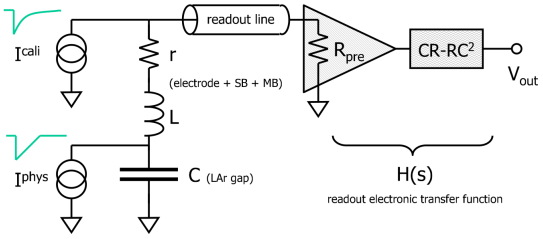
The transfer function to the signal modeling circuit is displayed in Equation 2.55, with R_f as transresistance of an ideal current amplifier followed by the $CR - RC^2$ shaping block shaper, τ_{circ} is the time constant and, in ATLAS it has a value of the order of 15 ns (COLAS et al., 1999; BANFI; DELMASTRO; FANTI, 2006). The semi-gaussian filter $CR - RC^2$, is used to produce a good representation of the input signal as a pulse whose amplitude is proportional to the detected particle energy. Another characteristic is a good signal-to-noise ratio (NOULIS; DERADONIS; SISKOS, 2007).

Figure 29 – Diagram of a typical circuit used in detector channels (a), and equivalent circuit for a LAr cell (b).



(a)

Source: Navas et al. (2024).



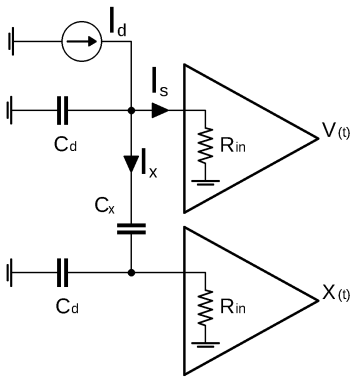
(b)

Source: Banfi, Delmastro and Fanti (2006).

$$H(s) = \frac{V_{out}(s)}{I_{in}(s)} = R_f \frac{\tau_{circ} s}{(1 + \tau_{circ} s)^3}. \tag{2.55}$$

The ATLAS electronic readout circuit has parasitic couplings between channels, which induce crosstalk signals. Figure 30 shows the diagram with the parasitic capacitive coupling C_x between two LAr cells indicating the currents of interest (I_s) and coupling current (I_x) modeled by Equation 2.56 and Equation 2.57, assuming $C_x \ll C_d$. The transfer function, for the cell of interest and the coupled cell, are given by Equation 2.58 and Equation 2.59, respectively (COLAS et al., 1999).

Figure 30 – Diagrams of the reading channel model used to estimate the capacitive coupling between two LAr cells considered in this thesis (LOPEZ SOLIS, 2017). V(t) is the cell signal for a given cell, and X(t) is the signal for the capacitive coupling.



(a)

$$I_s = \frac{I_d}{1 + sR_{in}C_t}, \tag{2.56}$$

$$I_x = \frac{sR_{in}C_x I_d}{(1 + sR_{in}C_t)^2}. \tag{2.57}$$

$$V(s) = \frac{I_o}{s} \frac{1}{1 + sR_{in}C_t} \frac{sR_f\tau_{circ}}{(1 + \tau_{circ}s)^3}, \tag{2.58}$$

$$X(s) = \frac{I_o}{s} \frac{sR_{in}C_x}{(1 + sR_{in}C_t)^2} \frac{sR_f\tau}{(1 + \tau_{circ}s)^3}, \tag{2.59}$$

$$X(s) = \frac{sR_{in}C_x}{1 + sR_{in}C_t} V(s). \tag{2.60}$$

The respective inverse Laplace transform of Equation 2.58 and Equation 2.59 are shown on Equation 2.61 e Equation 2.62. The values for the parameters used in the XT modeling in this thesis is presented in Table 16 on Appendix C. in Equation 2.62, C_t summarizes the detector and coupling capacitance between channels.

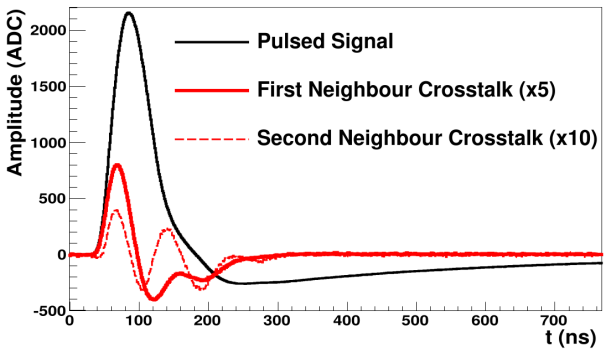
$$V(x) = R_f I_o \left[\frac{\lambda^2 e^{-x/\lambda}}{(\lambda-1)^3} - \left(\frac{x^2}{2} + \frac{\lambda x}{\lambda-1} + \frac{\lambda^2}{(\lambda-1)^2} \right) \frac{e^{-x}}{\lambda-1} \right], \tag{2.61}$$

$$X(x) = -R_f I_o \frac{\lambda C_x}{C_t} \left[\left(x - \frac{\lambda^2+2\lambda}{\lambda-1} \right) \frac{e^{-x/\lambda}}{(\lambda-1)^3} + \left(\frac{x^2}{2} + \frac{\lambda+1}{\lambda-1} x + \frac{\lambda^2+2\lambda}{(\lambda-1)^2} \right) \frac{e^{-x}}{(\lambda-1)^2} \right], \tag{2.62}$$

$$I_0 = \frac{I_d}{s}; \qquad x = \frac{t}{\tau_{circ}}; \qquad \lambda = \frac{R_{in} C_t}{\tau_{circ}}; \qquad C_t = C_x + C_d.$$

Another aspect used to study and evaluate the XT influence is associated with the source of the injected charge. This aspect differs for calibration signals injected on the board at the collecting signal connector levels and physics signals generated in the LAr gap. Figure 31 presents the intensity of the XT signal coming from a neighbor cell (in η).

Figure 31 – Typical signals from a cell and XT for the first and second neighborhood in η.



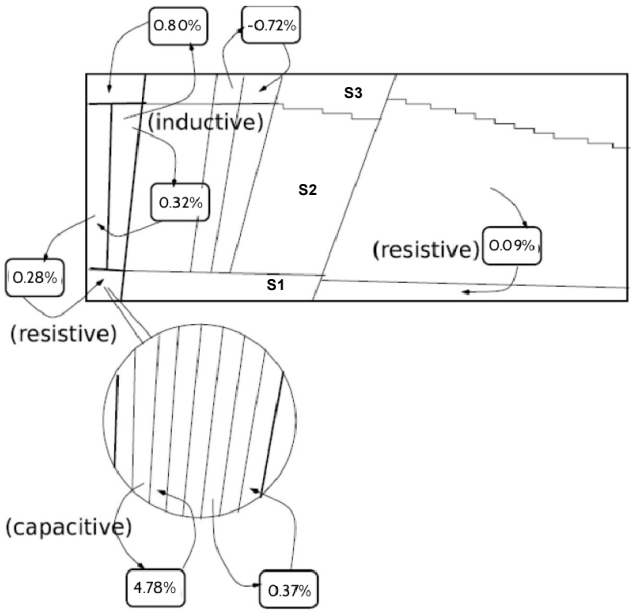
Source: Labbé and Ishmukhametov (2008).

Note that the magnitude of the XT signal from the second neighborhood cell is much smaller, around 25% of the magnitude of the signal in the first neighborhood, and has a very different shape. Labbé and Ishmukhametov (2008) comment that besides this fact, the signal of the first neighborhood approaches the signal derivative in the cell and that the XT of the second neighborhood (despite presenting distortion) maintains similarity with the second derivative of the cell signal, a characteristic of capacitive coupling.

In Figure 32 the dominant couplings in each liquid argon electromagnetic calorimeter layer are displayed for the under peak XT calculation. In the first layer (S1), the dominance is of the capacitive case, in the order of 4.78%. The predominance is inductive in layer 2 (S2), with 0.32%. In layer 3 (S3), there is also the predominance of the inductive effect, -0.72%[43]. The other measurements presented are relative to long-distance XT, that is, between layers or non-adjacent cells.

[43] The negative values are due to the fact that ϕ neighbors only present inductive XT, which tends to produce negative values when applying OF (LOPEZ SOLIS, 2017).

Figure 32 – Resistive, inductive, and capacitive crosstalk signals in the ATLAS electromagnetic calorimeter layers obtained from physical signals. XT is estimated with the OF energies.



Source: LOPEZ SOLIS (2017).

Table 5 – Summary of XT measurements on LAr layers evaluating interference level between layers from the first and second neighborhood using the energy contamination method.

Type	Physics – XT		Calibration – XT	
	Average (%)	RMS (%)	Average (%)	RMS (%)
$S_1 \to S_1$	4.78	1.20	4.31	0.65
$S_1 \to S_2$	0.197	0.177	0.080	0.017
$S_1 \to S_1$[a]	0.373	0.205	0.4(0.5)[b]	–
$S_2 \to S_1$	0.102	0.089	0.080	0.017
$S_2 \to S_2$ (η)	0.315	0.454	0.60	0.24
$S_2 \to S_2$ (ϕ)	−0.538	0.409	–	
$S_2 \to S_3$	0.804	0.699	0.48	0.17
$S_3 \to S_2$	0.846	0.465	0.48	0.17
$S_3 \to S_3$ (η)	−0.721	0.309	–	
$S_3 \to S_3$ (ϕ)	−0.389	0.122	–	

[a] Value from the second neighborhood.
[b] Contributions from measurements at $|\eta| > 1,1825$.

Source: LOPEZ SOLIS (2017).

Table 5 summarizes the physics and measured calibration XT relative magnitudes. The columns present the different sources of XTs (calorimeter layers) and the locations where the signal is coupled.

Crosstalk distorts the cell signal, altering the estimated values of amplitude and peak time, which increases the uncertainty of the process. LOPEZ SOLIS (2017) reports, based on experimental results, that XT introduces a bias of approximately 94 ps in the arrival time of particles in LAr cells.

Since particles travel nearly at the speed of light, c, a time-of-flight bias of 94 ps corresponds to a travel distance of approximately $\tilde{3}$ cm. This uncertainty is significant, as it represents nearly of the standard deviation associated with the point of collision, $\sigma_z = 3.5$ cm[44]. Reducing crosstalk could have an important impact by enabling the treatment of both in-time and out-of-time pile-up.

[44] <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/BeamSpotPublicResults>

3 Proposed Method

Contents

Introduction

This chapter presents the method used to develop a neural network estimator to treat crosstalk effects in a Liquid Argon (LAr) calorimeter. For this, two datasets were used. One comprises simulations from an analytical framework developed in this thesis to evaluate the effects of crosstalk development for single electromagnetic particles in a simplified calorimeter structure. The proposed crosstalk model was also incorporated to simulations produced with the Lorenzetti Showers framework (ARAÚJO et al., 2023), which has a calorimeter structure very similar to ATLAS Experiment and allows the manipulation of low-level information such as instrumentation pulses.

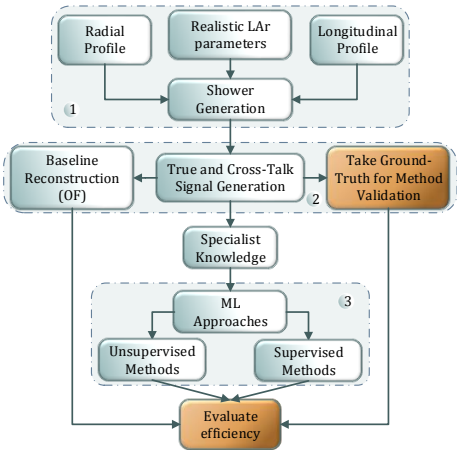
3.1 Electromagnetic Shower Analytical Simulator

3.1.1 Motivation

A particle shower analytical simulator was proposed in this work to obtain an effective method to mitigate the effects of the XT signal. Figure 33 presents the diagram with constituent blocks to build the EMshower simulator. The first block comprises the generator step, the second one shows the data structures used for this study, and the last block presents the evaluated machine-learning approaches. This simulator allows the simulation of the impact of interference on specific energy levels for a given particle of interest.

The objective of the analytical simulator of electromagnetic particle showers is to obtain a tool for rapid configuration of toy experiment to predict EM shower properties in the ATLAS detector geometry with complete flexibility and control of all parameters concerning XT contributions and cell signal. One more point is that the simulator can be executed without a sophisticated framework or infrastructure.

Figure 33 – Diagram with the steps of the algorithm design for the treatment of crosstalk in the liquid argon calorimeter cells of the ATLAS experiment.



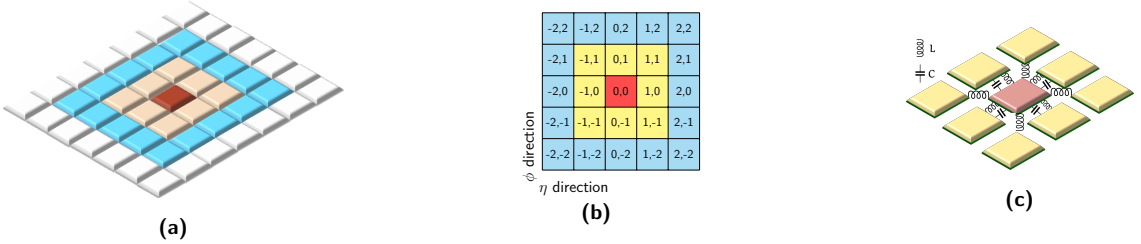
3.1.2 Simulator - General Principles

The simulator is based on the well-established description of energy profiles for an Electromagnetic (EM) shower (GRINDHAMMER; RUDOWICZ; PETERS, 1990), as detailed in subsection 2.5.3.

Figure 34 shows a diagram of the cluster in S2, which serves as the reference due to its square and homogeneous cells. This homogeneity implies that the cells have identical dimensions and physical parameters. To construct a 5×5 cluster, the impact energy and the coordinates of the hit point (η, ϕ) are selected, determining the hottest cell. Subsequently, a set of cells surrounding the impact point is taken, and the energy deposition is computed based on the energy deposition model for an EM particle, using the (η, ϕ) coordinates.

Following this, Figure 34b introduces the notation adopted for addressing cells within the 5×5 cluster: cell$_{ij}$, where i represents the row in the cluster (corresponding to the ϕ direction), and j represents the column (corresponding to the η direction).

Figure 34 – A representation of the 5×5 cluster in the ECAL L2 layer is shown in (a), the blue cells indicate the cluster limits, in yellow is a representation of the moving window (3×3) that is used to compute the crosstalk contribution from the first neighborhood cells. In red is the hottest cell of the cluster, it is the cell where the electron hits the calorimeter. In (b), the adopted cell addressing notation is shown, and (c) a representation of the main crosstalk contributions on the central cell from its neighbors at the first level is shown.



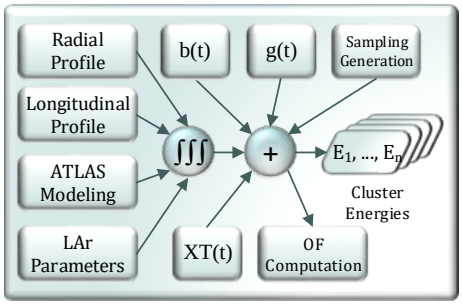
(a) (b) (c)

Figure 35 presents details of the block 1 on Figure 33. The simulator produces the information cell by cell on the LAr calorimeter based on the electromagnetic profiles to estimate the energy amplitude for the $g(t)$ cell signal. In this work, the XT coupling model is linear of the first order, which means that the electronic noise (b(t)), and XT(t) information is added on the cell energy information just considering cells around the interested cell.

The cluster production follows the strategy to achieve a good representation for the EM clusters to mitigate XT effects without adding more complexity to the model, without pile-up scenarios, and with no pulse jitter. This process stops when all cluster cells have their energy with XT information. The sampling generation block defines the sampling

time taking into account the delay among the group of cells managed by the Front Electronic (FEB)[45] (BUCHANAN et al., 2008b).

Figure 35 – Detailed block diagram for the particle electromagnetic shower simulator present in Figure 33. The integration block process produces the energy information concerning the EM energy profiles, experiment geometry, and detector parameters. As a result, the energy value per cell is determined. The sample's generation is taken in a fluctuation to reproduce the delay between cells, which FEB managing to keep synchronism in the ATLAS LAr modules, and the OF technique the standard reconstruction method, that uses two linear combinations to estimate an average value for energy and time gives a reference for ML methods of this work.



(a)

3.1.2.1 Energy Computation

The ATLAS experiment design is based on maximum energy transfer during the proton beam interaction process. Therefore, the most energetic particles are expected to go through the direction of $\eta = 0$, perpendicular to the collision point, at $z = 0$. This fact guides the simulator project, which must present convergence for calculating the integral of energy deposition for that region. To calculate the amount of energy deposited in an area of the detector, the modeling considers the energy profile functions (radial and longitudinal) in Equation 3.1 (GRINDHAMMER; RUDOWICZ; PETERS, 1990). The model convergence after the coordinates adjustment must produce the integral result equal to 1.

$$\frac{E}{E_0} = \int_{R_1}^{R_2}\int_{\eta_1}^{\eta_2}\int_{\phi_1}^{\phi_2} \underbrace{f_r(R,\,\eta,\,\phi)}_{\text{(i)}}\,\underbrace{f_t(R,\,\eta,\,\phi)}_{\text{(ii)}}\,\underbrace{J\left|\frac{f(R,\,\eta,\,\phi)}{\partial_R\,\partial_\eta\,\partial_\phi}\right|}_{\text{(iii)}} d\phi\,d\eta\,dR \approx 1, \tag{3.1}$$

in which E_0 is the impact energy, (i) is the function that describes the radial profile energy deposition, (ii) describes the longitudinal profile, and (iii) is the Jacobian associated with adjusting profile function into ATLAS EMCAL LAr coordinates.

The simulator validation depends on the numerical convergence in the region of interest inside the LAr calorimeter. Once the simulator is based on a homogeneous structure, i.e., for each layer, the cell size and parameters do not change. This structure allows the simulator to be more flexible and reduces the computational complexity associated with extensive technical design information.

The model was adjusted to generate a bunch of 748,000 electron showers with the following characteristics:

- Sampling time equal 25 ns with a time delay (δt[46] ns) to represent fluctuations between cell block on FEBs;
- Energy impact for a single electron, as a normal distribution with $\underline{\mu}$[47] = 50 GeV and $\sigma = 10$ GeV;
- Noise signal amplitude equal to 50 MeV as a normal distribution with zero mean and the sigma equal to 10 MeV (MENG, 2010);
- Transfer function modeling for capacitive XT_C coupling between two cells in the reading channel with Amplitude equal 4.2% of summation of the energy amplitude the relative neighbors (Peak-to-peak definition);

[45] These boards manage the CERN clock inside the ATLAS experiment to guarantee that LAr cells are synchronized.
[46] Generated a set of 49 delays from a normal distribution with zero mean and sigma equal 0.1 ns.
[47] This mean symbol is to differentiate the muon symbol.

- Amplitude for the XT_L contribution for a given cell equal to 2.3% (Peak-to-peak definition) of summation of the energy amplitude in its neighbors;
- At S1 the cell size is equal to $\Delta\eta = \frac{0.025}{8}$, $\Delta\phi = \frac{4\pi}{128}$. At S2 the cell size is equal to $\Delta\eta$=0.025, $\Delta\phi = \frac{\pi}{128}$. At S3 the cell size is equal to $\Delta\eta$=0.050, $\Delta\phi = \frac{\pi}{128}$.
- The fluctuations in ATLAS EMCAL cell size for large η modeled as an uniform distribution.
- Uniform cells characteristics for ATLAS EMCAL which implies that XT depends only of E_0.

A good representation of a realistic shower should present some fluctuations in the signal generated. The XT production takes into account the energy information deposited in the first neighborhood, i.e., the eight cells close to the interest cell.

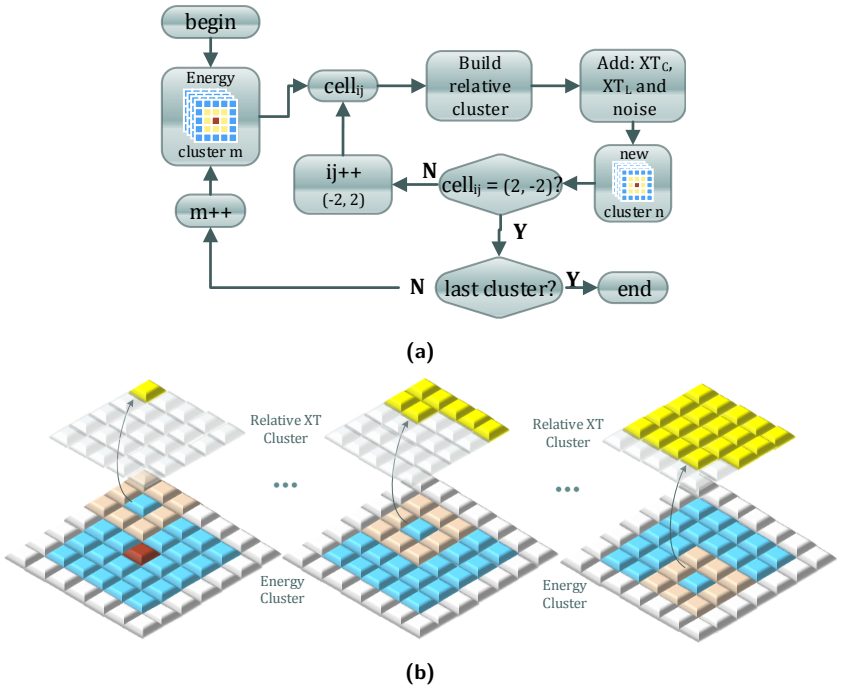
The XT contribution is a function of energy at the neighborhood for a cell and some unknown coupling function $f(E_{neighbors}, \chi)$[48]. In this work, the simulations in the XT signals are implemented by adding a source of fluctuations in the impact energy, E_0, applying a normal PDFs to represent the energy electrons decay with a mean equal to 50 GeV, and the σ_E equal to 10 GeV (ATLAS Collaboration, 2019b).

3.1.2.2 Production of final cluster information

In the EMshower simulator, a cluster contains information about the true energy deposited in the cell, the XT capacitive and inductive contributions, and the noise. In the LAr four samples are taken to perform the energy and time reconstruction, and these signals on the cluster are taken in the same sampling time.

The final 5×5 cluster considers all signals for a cell, taking into account its relative neighbors. Figure 36 shows a diagram to illustrate the process to produce a simulated cluster with all contributions of XT and noise for cells in the ATLAS LAr according to Equation 3.2, the final cell information is composed by the true energy (i) parcel, the XT_C (ii) parcel, the XT_L (iii) parcel, and the noise n_i (iv), which represents the relative XT_C, XT_L, and noise clusters respectively.

Figure 36 – (a) diagram to produce a final cluster considering all contributions for a given cell. (b), the diagram illustrates the process of building a relative XT cluster for each moving window position. Each cell of the relative cluster contains the sum of XT added in the cell at the center of the window. The moving window performs the XT contributions in a cell from its neighbors, and at the end of the process, a new cluster with full contributions is obtained.



[48] χ represent the different types of XT.

$$S_t^k = \underbrace{\hat{S}_t^k}_{\text{(i)}} + \underbrace{\sum_{i=1}^{4} XT_{C_t}^{i\to k}}_{\text{(ii)}} + \underbrace{\sum_{j=1}^{8} XT_{L_t}^{j\to k}}_{\text{(iii)}} + \underbrace{n_t^k}_{\text{(iv)}} \tag{3.2}$$

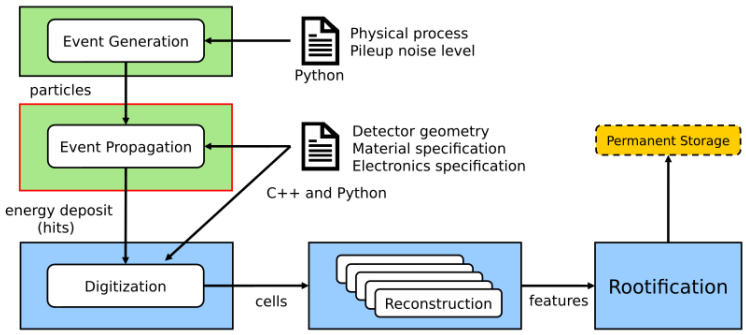
S_t^k is the sample at instant t for the interest cell k taking into account all contributions. $\hat{S}_t^k$ is the real sample value for the cell k at instant t, $XT_{C_t}^{i\to k}$ is the capacitive coupling from cell i to cell k, $XT_{L_t}^{j\to k}$, is the inductive coupling from cell j to cell k, and n_t^k is the noise sample in the cell k.

The validation of the analytical simulator is based on the numerical integration convergence for the ATLAS LAr modeling. The EM profile models used in this simulator are validated in the Geometry and Tracking version 4 (GEANT4) in a work that presents an assumption to describe the shower development for the radial and longitudinal profiles (GRINDHAMMER; PETERS, 1993).

3.2 Lorenzetti Showers Framework (LZT)

The Lorezentti Showers framework was used to study crosstalk and its effects in an electromagnetic calorimeter like ATLAS. Lorenzetti Showers (ARAÚJO et al., 2023), it is a general-purpose framework based on Geometry and Tracking version 4 (GEANT4) and PYTHIA8 to study calorimeter showers, including sampling segmented structures with electromagnetic and hadronic layers like ATLAS. In HEP, frameworks are a handle tool to allow studies in scenarios of complex structures and physical events. Figure 37 shows a summary of the general structure to generate an event in the framework. The Lorenzetti framework is a development by the ATLAS Brazil Collaboration together with the *Laboratoire de physique nucléaire et des hautes énergies* (LPNHE) group at Sorbonne Université to allow studies on HEP calorimetry.

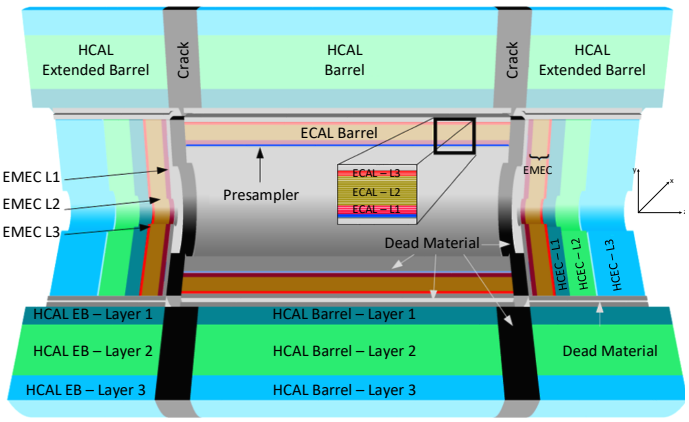
Figure 37 – Diagram of main block description of the Lorenzetti Showers framework. An event generator based on Python produces the set of particles from a collision, on the sequence, a block developed on C++ and Python has the setup for production and event generation, where those particles will interact with the detectors in the generated direction.



Source: ARAÚJO et al. (2023).

Figure 38 shows a diagram for the model of a physical experiment to studies based on calorimeter detectors used on the Lorenzetti framework. All data used in this study is on ECAL L2 (EMCAL S2 in ATLAS), which is the layer where the electromagnetic particles interact strongly (ARAÚJO et al., 2023).

Figure 38 – Diagram to detail experiment model of Lorenzetti framework. The model is composed of a cylindrical structure of different calorimeter types to characterize particle physics coming from an interacting point at the center of this cylinder. The structure is designed to absorb particles since most have very low energy levels until the high levels. This fact explains the sequence of the detector, most inner to low energies, and most external layers to absorb the most energetic particles.



Source: Adapted from ARAÚJO et al. (2023).

This study utilized a dataset generated of single electrons or positrons within the Lorenzetti framework (ARAÚJO et al., 2023) comprising 2 million events distributed across 20 impact energy levels from 10 to 200 GeV. Each energy impact value has approximately one hundred thousand events. Table 6 shows a summary of the used dataset comprising the mean and standard deviation values for the energy registered at the second electromagnetic layer for different impact energies (E_{imp}). It can be observed that as the particle loses energy also in the other calorimeter layers, the energy sampled at ECAL L2 is always smaller than the total impact energy.

Table 6 – Dataset information summary. Each impact energy comprises one hundred thousand events generated on the Lorenzetti framework for a full calorimeter detector. Estimated energy values refer to the second electromagnetic layer on the region of η range: [-0.0875, 0.0875], and ϕ range: [1.4604, 1.6322].

LZT: Average sum of energy deposited at the ECAL L2 cluster [GeV]							
E_{imp}	Energy	E_{imp}	Energy	E_{imp}	Energy	E_{imp}	Energy
10	5.44 ± 0.01	60	43.82 ± 0.01	110	85.37 ± 0.02	160	128.14 ± 0.02
20	12.50 ± 0.01	70	52.03 ± 0.01	120	93.86 ± 0.02	170	136.78 ± 0.03
30	20.01 ± 0.01	80	60.23 ± 0.02	130	102.44 ± 0.02	180	145.46 ± 0.03
40	27.82 ± 0.01	90	68.64 ± 0.01	140	110.94 ± 0.02	190	154.15 ± 0.03
50	35.78 ± 0.01	100	77.02 ± 0.02	150	119.51 ± 0.02	200	162.80 ± 0.03

In the LZT data, a set of cluster energy for single $e^{\mp}$ was produced. On the sequence, those cluster cells were sampled to generate the cluster samples four samples per cell as were made in the EMshower simulator. As a result, a 5×5 cluster cell becomes a 20×5 cluster sample.

3.3 Crosstalk mitigation approaches proposals.

This section presents the chosen approaches evaluated in this thesis to mitigate XT effects. The basis of this study is the model described in the previous section and summarized in Equation C.21. Two approaches in Machine Learning (ML), supervised and non-supervised, were, evaluated using data generated by the EMShower simulator developed for this study in this thesis proposal.

3.3.1 Supervised Methods

Our ML approach is based on Feedforward (FF) Neural Networks (NN) with Multilayer Percepton (MLP) structure. The aim is to evaluate and obtain a structure able to mitigate XT using a non-complex Neural Networks (NN) structure that could be embedded at ATLAS EMCAL reconstruction system.

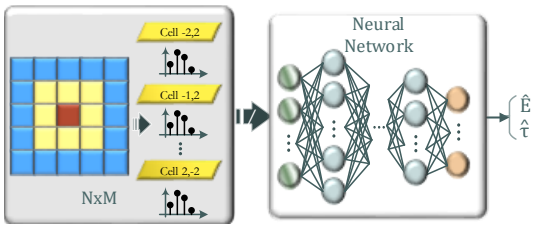
Multilayer NN

The best results presented in this work were achieved in an investigation, not exhaustive, to get a model using a MLP structure with few numbers of hidden layers and a small number of neurons. According to this strategy, a set of structures of MLP nets were trained. After that, an analysis of this set of results was made to evaluate the best structure to minimize the XT effects on the energy and time estimation. The setup used in this thesis is as follows:

- Number of Hidden Layer (HL) [1, 2, 3, 4];
- Number of features in the input layer:
  - For signal regression (E/τ): 100 (4 samples per cell in a 5×5 cluster);
  - For shower shapes analysis: 196 (4 samples per cell in a 7×7 cluster);
- Number of neurons on output layer:
  - EMshower: 25 for (E/τ);
  - LZT: 25 for an 5×5 cluster of E/τ;
  - LZT: 100 for samples regression;
  - LZT: 196 for shower shapes analysis;
- Initialization to avoid a local minimal: 15;
- Early stopping: minimal Mean squared error (MSE);
- Optimizer: Adam;
- Number of epochs: 5000 – A large number of epochs is chosen to be the last stop criteria;
- Stop criteria: minimal MSE;
- Architecture: MLP;
- Activation functions: input layer - tansig, Parametric Rectified Linear Unit (PReLU), hidden layers - PReLU with $\alpha = 0.3$, and output layer - linear;
- Normalized signal.

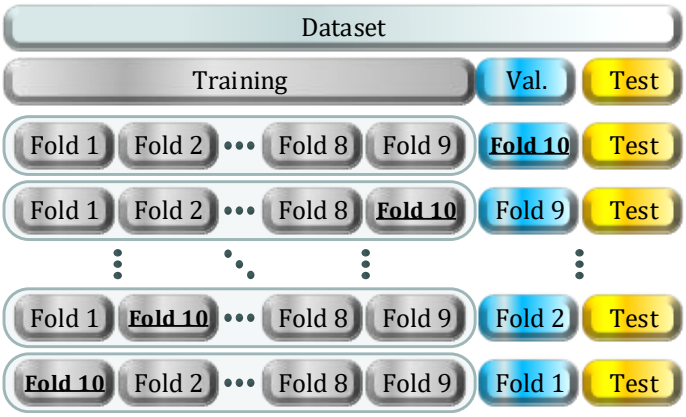
Figure 39 shows the structure and organization of each 5×5 cluster cell. The four time-sample cells were concatenated in a vector with 100 features, and it was standardized before applying to the NN. The hottest cell is in red, and their first neighborhood is yellow.

Figure 39 – Data organization used to evaluate the ML approach. A bunch of clusters with each cell and its respective set of 4 samples as input to make a regression for energy and time. In this example, 100 refers to 25 cells with 4 samples each to regression cluster 5x5.



K-fold cross-validation was adopted by dividing the datasets randomly into eleven subsets. Ten subsets were used to train and validate the model, and one was kept only to test the model, as presented in Figure 40. A common problem when training an NN model is overfitting, together with the k-fold strategy early stopping was implemented to avoid such a problem.

Figure 40 – Block diagram of the data subsets used in this work. The data set was divided into 11 equal parts. Ten parts are used for training and validation using the k-fold strategy, and one is kept only for testing.



3.3.2 Evaluation of the results

In this work, some metrics were used to evaluate the results: the root mean square error and Kolmogorov-Smirnov Test (KS), with their respective equation described below.

$$rmse = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Y[i] - \hat{Y}[i])^2}, \tag{3.3}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(\hat{Y}[i] - Y[i])^2}{\sum_{i=1}^{N}(Y[i] - \overline{Y})^2}, \tag{3.4}$$

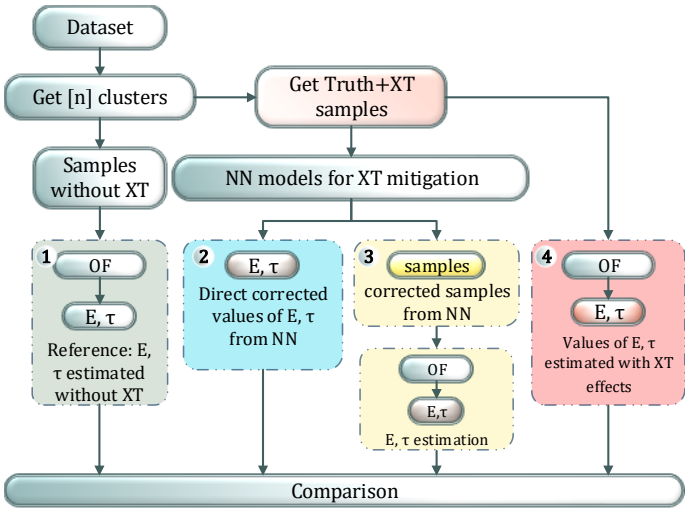
$$D_n = sup_x[|F_n(x) - F(x)|]. \tag{3.5}$$

where $Y[i]$ is the target values, $\hat{Y}$ is the predicted values, and $\overline{Y}$ is the mean values; the numerator means the squared sum of residues, and the denominator is a fixed part related to the variance of the data. $F_n(x)$ is the Cumulative Distribution Function (CDF) of the sample data, $F(x)$ is the CDF of the reference distribution, and sup_x denotes the supreme of the set of distances. The Figure 41 illustrates the comparison process between the results of the OF, ML approaches and the truth values.

The results achieved with the NN estimator will be compared with the OF results by the error value Root mean squared error (RMSE), Equation 3.3. The quality of the fit model is evaluated by the coefficient of determination ($R^2 \in [0, 1]$) presented in Equation 3.4. To evaluate the similarity of distributions achieved by the NN estimator, and the OF with respect to the Ref values the Kolmogorov-Smirnov Test (KS) (KOLMOGOROV–SMIRNOV-TEST, 2008) will be used as described in Equation 3.5.

Figure 41 presents how the results from the NN estimator and OF will be evaluated. First, the reference values are taken using the Optimal Filtering (OF) without XT. Two approaches to building an NN estimator will be evaluated: one to estimate energy and time directly from samples, and another to regress samples and calculate the energy and time using the OF. On the Chapter 4 will be presented results for block 3, which got the best results over the set of analyses developed in this work.

Figure 41 – Block diagram to describe the method in this study. Column 1 is the energy and τ reconstructed using samples with noise but without XT couplings. Column 2, an NN estimates energy and time from the samples with XT. Column 3, an NN estimates samples, and OF estimates the energy and time. The last column, 4, represents the energy and τ estimation without any mitigation method. The last block, at the bottom figure, is the comparison, to evaluate the efficiency among blocks 2, and 3 with respect to block 1.

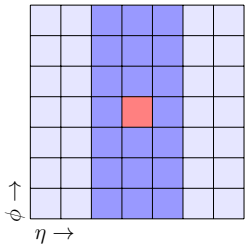


Performance evaluation considering shower shapes variables

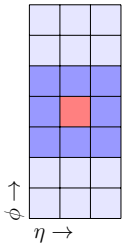
In HEP the electromagnetic particle characterization takes into account how the electromagnetic shower develops laterally (in the $\eta \times \phi$ plane) and longitudinaly. For this, a set of variables (known as shower shapes) is usually computed from the calorimeter information and used for electron and photon identification and reconstruction (ATLAS Collaboration, 2019a). These variables are important metrics for the correct identification of particles. So as the crosstalk brings a distortion on calorimeter measurements it is interesting to verify the effects on shower shapes of the proposed crosstalk mitigation method. In the ECAL L2 layer three variables are evaluated:

- $R_\eta = \frac{\sum E_{3\times7}}{\sum E_{7\times7}}$ – the ratio of the deposited energy in 3×7 cluster cells to the energy in 7×7 cluster cells centered at the position of the electron cluster.
- $R_\phi = \frac{\sum E_{3\times3}}{\sum E_{3\times7}}$ – the ratio of the energy deposited in the 3×3 cluster cells to the energy in the 3×7 cluster cells centered at the position of the electron cluster.
- $W_{\eta 2} = \sqrt{\frac{\sum E_i\ \eta_i^2}{\sum E_i} - \left(\frac{\sum E_i\ \eta_i}{\sum E_i}\right)^2}$ – Lateral shower width calculated within a window of 3×5 cells.

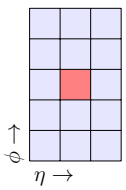
Figure 42 – Illustration to show the cluster cells around the hottest cell used in each shower-shapes variable. (a) cluster size for R_η, (b) R_ϕ, and (c) $W_{\eta 2}$



(a) 3×7 and 7×7 clusters.



(b) 3×3 and 3×7 clusters.



(c) 3×5 cluster.

4 Results

Contents

Introduction

This chapter presents the results of the proposed method which uses ML techniques to mitigate XT effects in the sample cells of a LAr calorimeter. The whole set of analyses presented uses the best fold over the 10-fold set. Two datasets were used: one based on the EMshower simulator, developed in this work to get fast simulations of single electrons. The second dataset was produced with the Lorenzetti Framework.

4.1 Analysis on EMShower Data

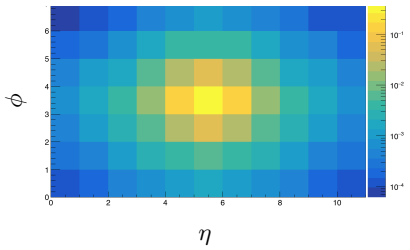
The EMShower simulation was implemented using the ROOT[49] framework based on C++ language, and a model convergence validation procedure was adopted. The validation procedure evaluates the numerical integral convergence of the whole energy deposition model across the LAr EM layers. To this end, it was used the ATLAS EMCAL physical parameters for the point $\eta = 0.3$[50], which is inside the region with fine granularity. Around this point, particles with high energy are also expected to be observed. The cell's thin granularity combined with high energy readings should provide the highest sensitivity to XT effects. The validation of the EMshower simulator means the numerical integration convergence of the modeling for ATLAS.

Once the EMshower achieves the numerical integration convergence for a given impact point (η,ϕ), a bunch of clusters is generated to evaluate the coherence of results with real data to follow with the studies to mitigate XT effects in a cell. In Table 7 it is possible to verify that on S2 was deposited 44.56 GeV or 89.13% of the impact energy. Only in the hottest cell is 85.07% of the energy in S2 layer. This result is very similar to the presented at Figure 43 which shows a window (η, ϕ) in S2 layer with size equal to 11×7 of energy deposition of electrons from $Z \rightarrow e^-e^+$ decays from ATLAS collision data. In this figure, roughly 90% of the total energy was recorded in S2 layer according to Khandoga (2020).

[49] An open-source data analysis framework used by high energy physics and other areas developed at CERN and available in: <https://root.cern/>.

[50] In fact, $\eta = 0$ is the expected point for the maximum transverse energy, however, in this region, there exists a gap between the two half-barrel modules.

Figure 43 – Energy profile of a window of 7×11 (η, ϕ) cells in the S2 ATLAS calorimeter layer for a $Z \rightarrow ee$ decay. At the center cluster, the hottest cell and the first neighbors, roughly 90% of the energy is deposited.



Source: Khandoga (2020).

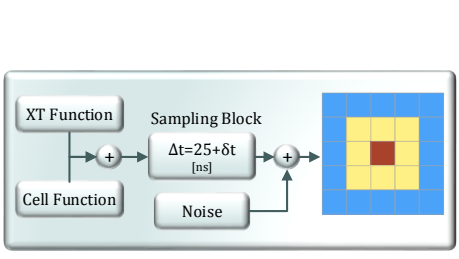
This result is consistent with the ATLAS LAr geometrical structure because this layer S2 is the thickest layer[51], which means the particle interaction with the medium produces a higher energy loss across S2 layer if compared to S1 (medium thickness) and S3 layers (smaller thickness).

Table 7 – Example of energy per cell in a standard EM cluster of size 5×5, on S2 layer produced by EMshower for an electron with energy equal to 50 GeV. In yellow is the first neighborhood around the hottest cell in red. The amount of energy deposited in S2 is equal to 44.56 GeV, 89,13% of the incoming particle energy. The cell coordinates (η, ϕ) are the first line, and first column, respectively.

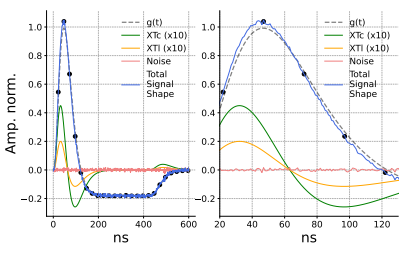
Averaged cluster energy [GeV]						
–	-2	-1	0	1	2	address
ϕ \ η	-0.025	0.000	0.025	0.050	0.075	–
$-\frac{\pi}{128}$	0.38 10^{-3}	4.19 10^{-3}	10.01 10^{-3}	4.20 10^{-3}	0.38 10^{-3}	2
0.000	4.22 10^{-3}	356.90 10^{-3}	1.27	359.00 10^{-3}	4.52 10^{-3}	1
$\frac{\pi}{128}$	10.04 10^{-3}	1.27	37.91	1.30	10.84 10^{-3}	0
$\frac{2\pi}{128}$	4.23 10^{-3}	359.90 10^{-3}	1.30	361.60 10^{-3}	4.56 10^{-3}	-1
$\frac{3\pi}{128}$	0.39 10^{-3}	4.47 10^{-3}	10.76 10^{-3}	4.52 10^{-3}	0.40 10^{-3}	-2

Figure 44b presents all signals generated with our EMshower simulator for a single cell: The typical cell pulse shape (g(t)); the electronic noise signal, together with the capacitive (XT_C) and inductive (XT_L) crosstalk signals.

Figure 44 – (a) presents the diagram with the main simplified blocks that guide the EMShower simulator development. In (b), the set of simulated signals for a cell of the ATLAS EMCAL from the EMShwoer simulator. The cell signal is in the black dashed line. Dots indicate the samples extracted from the simulator with a period of 25 ns. In green the XT_C generated from transfer function from Colas et al. (1999). In orange, the XT_L signal is based on XT_C since this model is yet unknown. The noise contribution is in red, and in blue, the composite signal. The black dots are the samples of the composite signal used to reconstruct information about energy and time. The figure on the right presents a detailed view (zoom) within $0 < t < 150$ ns.



(a)



(b)

[51] See X_0 on section 2.5, p.15.

Figure 45 shows cluster-level details for different contributions. Figure 45a presents the true energy, in Figure 45c we can see the capacitive contribution that occurs in the perpendicular and lateral cells. The inductive crosstalk contribution is illustrated in Figure 45d. In Figure 45e the contributions XT_C and XT_L are considered, and in Figure 45b we see the full signals recorded, with the contributions of energy and XT.

Figure 45 – Average signals of 745,000 5×5 clusters (η, ϕ) produced in EMshower simulator. In (a) the true energy estimated by OF. (b) the composite cluster with energy and XT capacitive and inductive in each cell. (c) the cluster with the peak capacitive XT contribution that will be used to estimate the XT_C samples. (d) the cluster with the peak inductive XT contribution. (e) both XT contributions, capacitive and inductive.

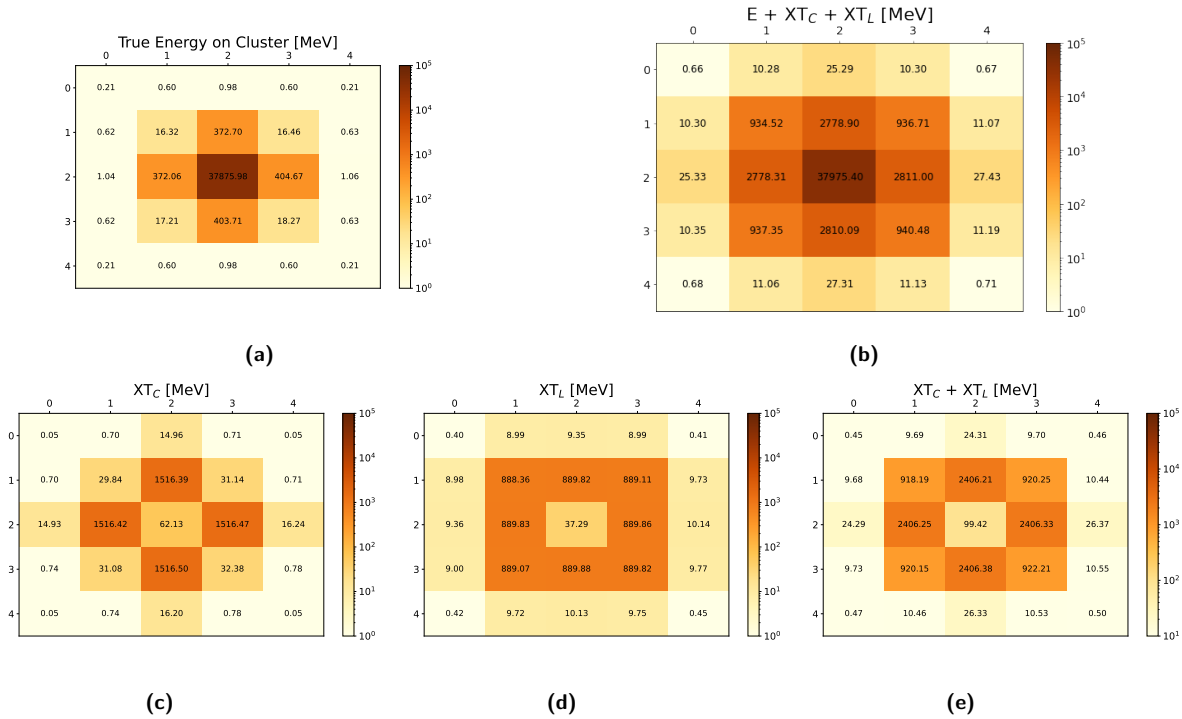


Figure 46 shows the ratio $\frac{E}{XT}$ (in dB) analysis cell by cell for the average cluster from the EMshower simulator. It is interesting to consider and evaluate this relationship once the OF coefficients are taken without XT scenario. This analysis provides important hints for explaining the next section's energy and time regression results.

Figure 46 – The $20\,log\,\frac{E}{XT}$ (in dB), cell by cell at the 5×5 cluster.

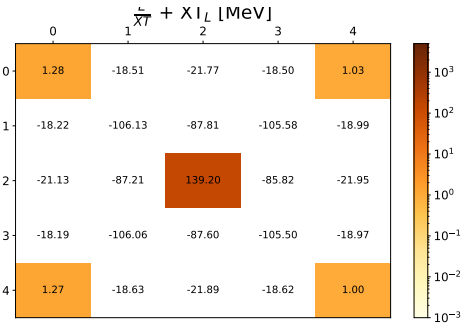
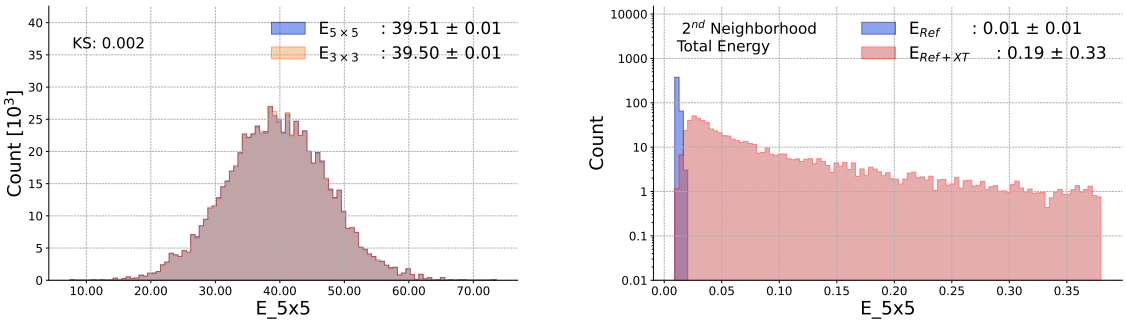


Figure 47a shows the histograms for energy to compare 3×3 and cluster 5×5 cluster. In this comparison, the 3×3 cluster preserves 99.90% of the energy in the whole 5×5 cluster, and, the distributions are very similar according to the small value for KS=0.002. Considering that, a comparison will be presented, the analysis for the center 3×3 cluster for the energy and time reconstruction. Figure 47b shows the histogram for true energy, and XT effects at the second neighbors. Due to this fact, all results presented in this chapter focus on the 3×3 cluster cells.

Figure 47 – Energy deposition comparison between the 5×5 cluster, and the 3×3 cluster. (a), in blue the true energy at 5×5 cluster, and in red the true 3×3 cluster deposition. (b), the true energy at the second neighborhood in blue, and in red the XT contribution at the second neighborhood.



(a) **(b)**

In the 3×3 cluster it is possible to organize cells in three groups:

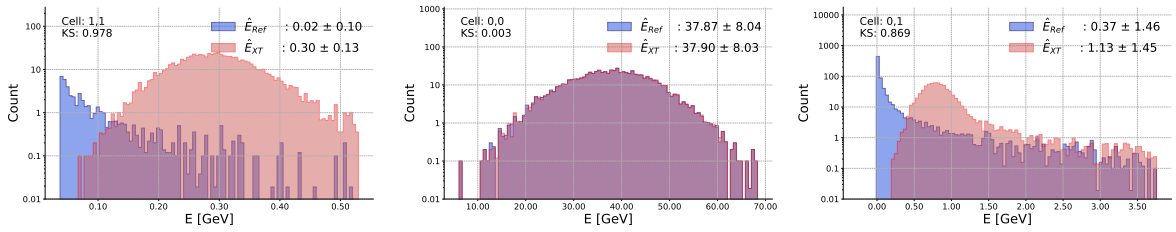
- Cells at diagonal - Under inductive XT effects;
- Cells at perpendicular - Under capacitive and inductive XT effects;
- Center cell - Hottest.

Figure 48 – Cell group organization to simplify results presentation.

-1,1	0,1	1,1
-1,0	0,0	1,0
-1,-1	0,-1	1,-1

In Figure 49 it is possible to see one of the consequences of XT (distortion) on the energy estimated for each cell in the simulated events. Instead of showing all the nine cells, for simplicity, we are considering the central cell, where most of the energy deposition occurs, a diagonal cell where inductive XT coming from the central cell is more pronunciated, and a perpendicular cell that suffers from both inductive and capacitive XT coming from the central cell. Equivalent analyses for all nine cells in the cluster are shown in Appendix D. At low energy, the histogram peak shifting is most evident due to contributions from their eight neighbor cells.

Figure 49 – Histograms for true energy and energy plus XT contributions for cell$_{ij}$ in a 3×3 cluster. (a) a diagonal cell, (b) the central cell, and (c) a perpendicular cell.



(a) Diagonal. **(b)** Center. **(c)** Perpendicular.

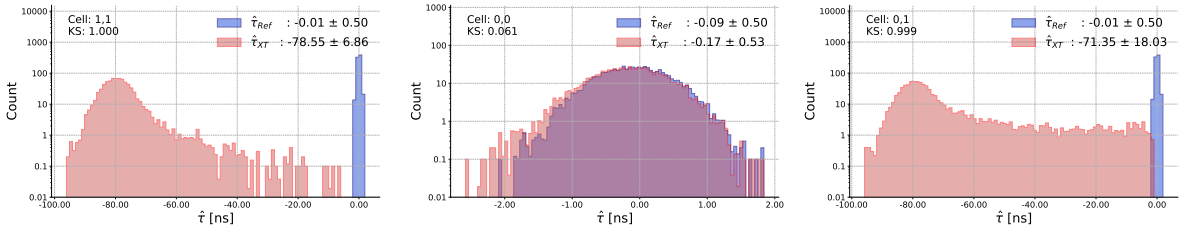
Considering the estimated time of flight ($\hat{\tau}$). The simulations considered that $\tau = 0$ without XT. Figure 50 shows the τ reconstruction using OF on samples without XT (in blue), and with XT influence (in orange). As well as in

the energy case, time reconstruction at low energies is more sensitive to XT, producing a large error in time estimation by OF.

The crosstalk effect is more severe in reconstructing the time information than for energy, which adds an extra challenge. The OF algorithm can not deal well with the crosstalk presence because the OF coefficients are estimated using the signal shape and noise contribution without the XT.

So looking at the cell of interest, see Figure 50b, the XT influence in time, $\Delta\tau \approx 60$ ps, lower than the maximum value measured by LOPEZ SOLIS (2017) on calibration runs, $\Delta\tau = 94$ ps. This result could be explained by the fact that inductive crosstalk is modeled as the capacitive transfer function.

Figure 50 – Histograms for time estimation with OF using signals without XT for cell$_{ij}$, with an arbitrary time shifting, $\tau_0 = 0.5 \pm 0.5$ ns, in blue, and energy with XT and noise, in orange, on the first neighborhood around the hottest cell (0,0) from Table 7.



(a) Diagonal. **(b)** Center. **(c)** Perpendicular.

4.1.1 Methods evaluated for Crosstalk (XT) mitigation

4.1.2 Supervised methods Crosstalk Learning

In this section are the results from the evaluation of supervised learning techniques XT mitigation. A NN receives 100 features that come from a 5×5 cluster at the input and produces an energy and time estimation at the output. For these approaches, a dataset with 745,000 clusters was used.

Multilayer NN

4.1.3 Time of Flight (τ) Estimation

This section presents the results for the time of flight estimation case. Neural networks with one, two, and three hidden layers (HL) were used and compared in Figure 58. It can be observed that the boxplots for 1 and 3 HL are similar, but in the latter case are the NN models with the smallest RMSE. Considering this, the NN with 3 Hidden Layer (HL) (25-10-5) neurons structure was used here. Table 17 on Appendix D summarizes these results.

As input, a set of 100 features was used, 4 per cell, and the output equals 25 to get the $\hat{\tau}$ of the cells cluster.

Figure 51 – Boxplots for the best NNs model for time regression using 1, 2, and 3 HL.

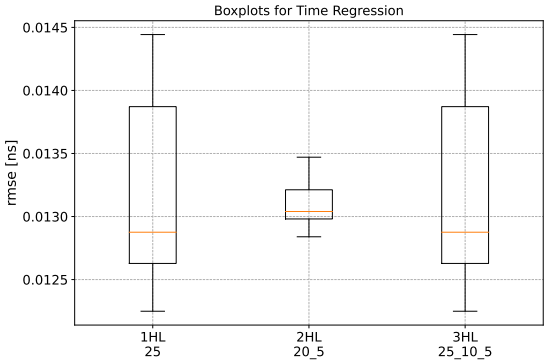


Table 8 – Summary of the results for the three best NN structures, using 3HL for time (in ns) regression task. Each cell is presented with its respectively RMSE value, and on the bottom of each structure column is the global RMSE.

OF		[100 – 10 – 10 – 5 – 25] Neur.		[100 – 25 – 10 – 5 – 25] Neur.		[100 – 30 – 10 – 5 – 25] Neur.	
RMSE	4.96e+01	RMSE	1.24e-02	RMSE	0.12e-02	RMSE	1.24e-02

Table 9 shows the output of the regression task for the particle time of flight for the best NN structures presented in Table 8 using three HL with 25-10-5 neurons. The last column shows how far the OF estimation with XT is in relation to the NN estimation.

Table 9 – Detail for the best NN structure, using 3 HL with 25-10-5 neurons for time [ns] regression task. Each cell is presented with its respectively RMSE value, and on the bottom of each structure column is the global RMSE. The last column brings an RMSE ratio between the OF and the NN.

cell	Target	OF		NN		$\frac{RMSE_{OF}}{RMSE_{NN}}$
	$\hat{\tau}$ [ns]	$\hat{\tau}$ [ns]	RMSE	$\hat{\tau}$ [ns]	RMSE	
diagonal	-0.07 ± 0.50	-80.43 ± 6.32	80.59	-0.07 ± 0.51	0.02	5068.02
perpendicular	-0.01 ± 0.50	-73.43 ± 17.07	75.36	-0.01 ± 0.51	0.02	4743.53
center	-0.09 ± 0.50	-0.18 ± 0.54	0.23	-0.09 ± 0.51	0.02	14.57

Figure 52 shows the MSE evolution through the validation epochs for the best NN structure to time estimation. Some oscillations were observed, and considering this, the training was stopped for a minimum value of the validation error.

Figure 52 – MSE for an arbitrary fold across training epochs for the best NN estimator structures at the Table 8 for time regression.

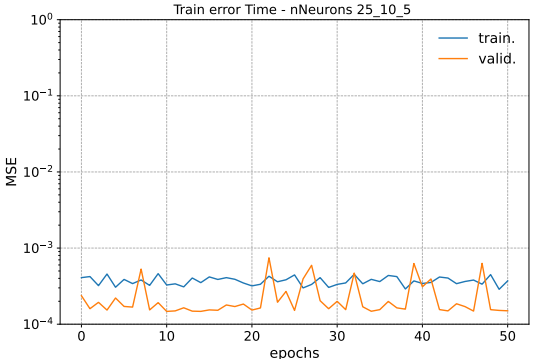
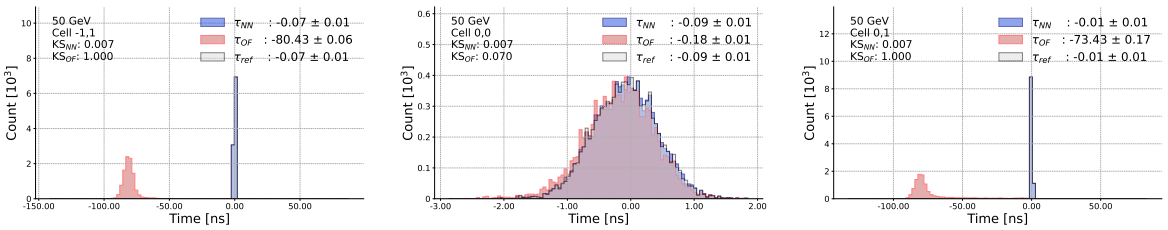


Figure 53 presents the time regression histograms for the best NN estimator structure using 3 HL with 20–10–5 neurons. NN presents a good regression with a low time spread. The auxiliary histogram shows details for the NN estimator distribution for time regression in relation to the OF technique. The time distribution for NN has good precision once the distribution achieved is the same for the target time values. Now, paying attention to the results of the OF technique, the spread is large for the cells at the diagonal and the perpendicular position, a fact that keeps the relationship with the OF project, which did not take into account the effects of XT. Look at the hottest cell (Figure 53b) the OF error is very low due to large ratio energy XT.

As electron deposition is concentrated[52], the hottest cell (0,0) records the highest energy in the cluster and the small XT contribution of its neighbor. This fact allows the OF to estimate the energy very well, with a small error. But looking for cells at diagonal or perpendicular positions, the opposite occurs. The hottest cell neighbor receives a large XT from the hottest cell. As a result, the OF error for this cell is large like Figure 53 shows.

[52] As described in section 2.5 in page 17 (WIGMANS, 2017)

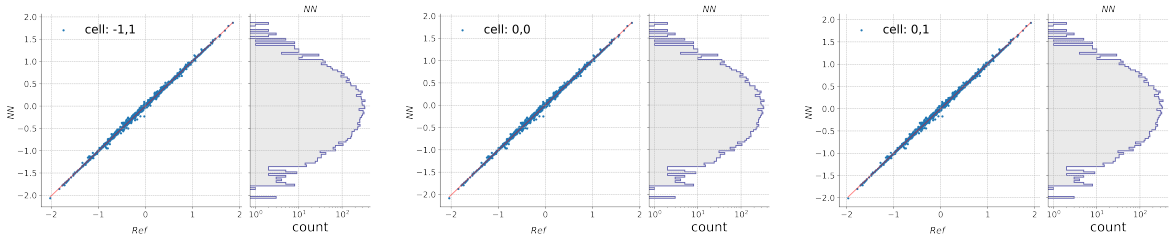
Figure 53 – Time regression histogram cell by cell comparison among OF, target and NN estimator structure ([100 – 25 – 10 – 5 – 25]). Comparison between NN structure and OF cell by cell on the center of 3×3 cluster.



(a) Diagonal. **(b)** Center. **(c)** Perpendicular.

A scatter-plot analysis is presented in Figure 54 for the proposed NN regressor. There is a good agreement comparing estimated and true time-of-flight values, reinforcing the efficiency of the proposed method.

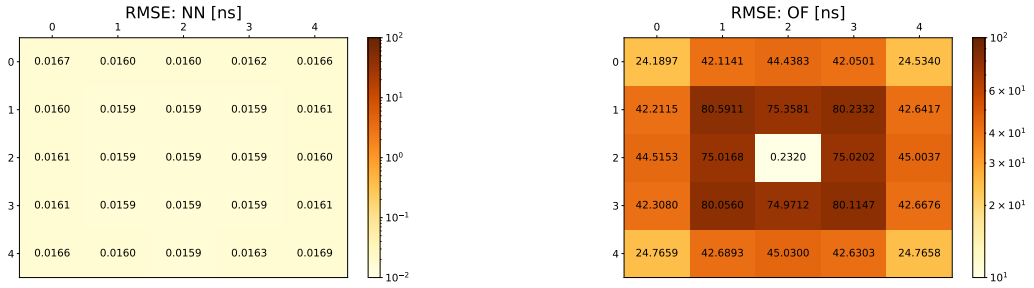
Figure 54 – Scatter plots for time reconstructed by NN with structure of 3 hidden layers ([100 – 25 – 10 – 5 – 25]). A good correlation between target values and NN output is observed for all cells in the center of 3×3 cluster around the hottest cell.



(a) Diagonal. **(b)** Center. **(c)** Perpendicular.

Figure 55 presents the errors between estimated time of flight with NN and the target, and between the OF and target. It is observed that the NN model produces small errors for the whole cluster. Differently, the OF estimation generates considerable errors for the first neighborhood of the hottest cell.

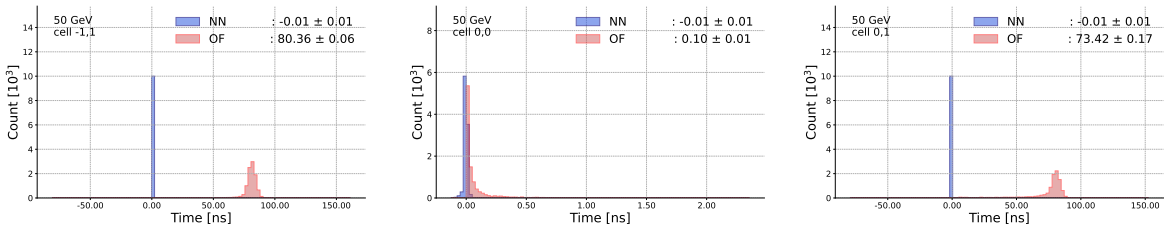
Figure 55 – RMSE over 10,000 events for the time reconstruction between NN and target in (a), and between OF and target in (b) at the cell level.



(a) **(b)**

Given the result, reconstruction error for NN is greatly improved with respect to target values, see Figure 56 compared to OF results. As is expected, the OF produces a large time error and a considerable time spread, see cell at diagonal Figure 56a, and perpendicular cell position Figure 56c. The best results for OF are on the hottest cell, see Figure 56b. However, looking at the NN error, the structure using a single layer with 20 neurons shows a low error and a narrow spread, which shows promising results to improve the real ATLAS timing reconstruction in LAr.

Figure 56 – Details for the reconstruction error for the OF and NN using 3 HL. The error for NN structure is low and with a narrow spread. For OF it is possible to see the limitation in this technique to dealing with XT influence.



(a) Perpendicular. **(b)** Center. **(c)** Perpencidular.

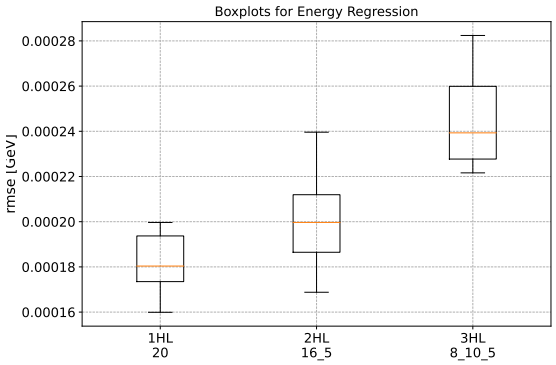
4.1.4 Energy Estimation

In this approach NN structures using 1, 2, and 3 Hidden Layer (HL) were evaluated. For each structure, eight neuron configurations were tested as follows:

- For 1 HL (Inputs – HL – Outputs): 100 – n – 25, n $\in$ {5, 8, 10, 15, 16, 20, 25, 30};
- For 2 HL (Inputs – HL_1 – HL_2 – Outputs): 100 – n – 5 – 25, n $\in$ {5, 8, 10, 15, 16, 20, 25, 30};
- For 3 HL (Inputs – HL_1 – HL_2 – HL_3 – Outputs): 100 – n – 10 – 5 – 25, n $\in$ {5, 8, 10, 15, 16, 20, 25, 30};

Figure 57 is presented the summary of training for the energy reconstruction task. The two boxplots for 1 HL and 2 HL have a narrow spread, but for the former case, the RMSE median and minimum values are smaller.

Figure 57 – Boxplots for the best NNs model for energy regression using 1, 2, and 3 HL.



In Table 10 is shown the summary for the three best NNs for the energy regression task shown in the boxplot of Figure 57 for each configuration, using 1HL, 2HL, and 3HL. The table presents the target values and the OF output to compare with the NN output. The three best NN produce a lower RMSE and a narrow spread for energy per cell compared to OF. Details are in Table 18 Appendix D

Table 10 – Summary of the best results (RMSE in GeV) using 1HL, 2HL, and 3HL on energy regression task.

Estimator	OF	NN_1	NN_2	NN_3
RMSE	903e-03	0.18e-03	1.09e-03	1.10e-03

Figure 58 presents the train and validation MSE evolution across epochs. Both training and validation errors are very small. Some fluctuations were observed for the validation case, but to avoid overtraining, the process was stopped when both training and validation errors were minimal.

Figure 58 – MSE for an arbitrary fold across epochs of training for the best NN estimator structures at the Table 10 for energy regression.

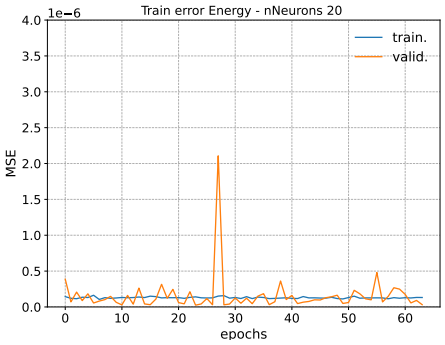


Table 11 presents the energy estimation results as presented for the time of flight estimation. This energy distribution presents a typical pattern that follows the EM deposition. Looking for the RMSE ratio column, only for the hottest cell the OF presents a low RMSE. For the diagonal and perpendicular cells, it is very high.

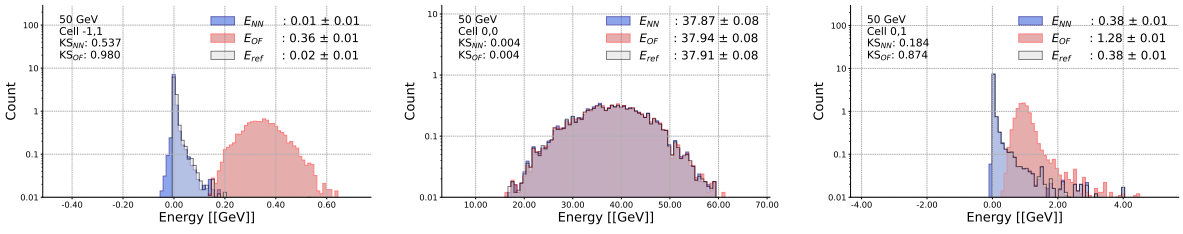
Table 11 – Summary of result for the best NN estimator, using 1 hidden layer with 20 neurons for energy (in GeV) regression task. And for each cell is presented its respective RMSE value. The last column shows an RMSE ratio between OF and the NN.

-2,2	-1,2	0,2	1,2	2,2
-2,1	-1,1	0,1	1,1	2,1
-2,0	-1,0	0,0	1,0	2,0
-2,-1	-1,-1	0,-1	1,-1	2,-1
-2,-2	-1,-2	0,-2	1,-2	2,-2

	Target	OF		NN		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	Energy	Energy	RMSE	Energy	RMSE	
diagonal	0.02 ± 0.10	0.36 ± 0.13	0.35	0.01 ± 0.09	0.01	33.00
perpendicular	0.38 ± 1.46	1.28 ± 1.46	0.92	0.38 ± 1.46	0.01	83.69
center	37.91 ± 8.01	37.94 ± 8.00	0.08	37.87 ± 8.01	0.06	1.21

Figure 59 presents the energy regression for the three cell groups inside the 3×3 cluster. The hottest cells show a Gaussian-like[53] distribution. Looking for cells at the diagonal and the perpendicular positions, which have low-energy deposition, the distribution presents a fast decay after the mean value and a tail.

Figure 59 – Histograms for energy regression comparison among target values, OF estimation using samples with XT, and an NN estimator to regress energy from 100 features from a 5×5 cluster, using one hidden layer with 20 neurons. (a) is the diagonal cell, (b) the center, and (c) the perpendicular.



(a) Diagonal. **(b)** Center. **(c)** Perpendicular.

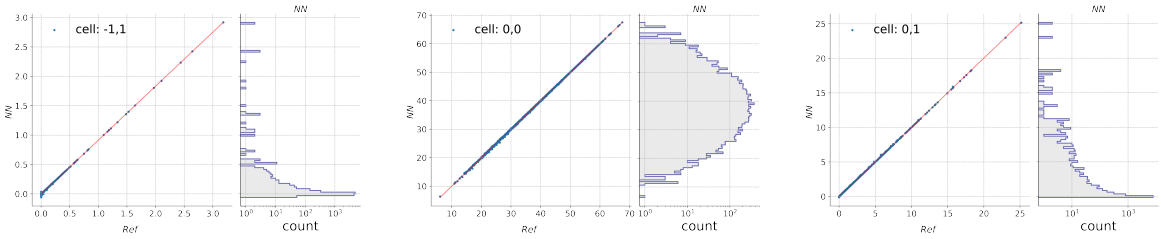
For the diagonal cells, where exists one type of XT, the inductive, the histograms show that NN and the target has a similar distribution, while OF produces a shift on the peak with the value greater at least three times the target.

Looking at perpendicular cells, where there exist two types of XT, inductive and capacitive, the reconstructed energy distributions for NNs are close to the target values again. In this group, it is interesting to observe that the energy level is greater than 1 GeV which provides a better ratio $\frac{E}{XT}$ than the previous group allowing OF to achieve a better energy estimation in comparison with cells in the first group. For the hottest cell, the NN and OF produce similar results since, in this cell, the energy is more significant than XT, which was indicated by the high ratio $\frac{E}{XT}$ in Figure 46.

[53] According to the KS=0.009 test for a Gaussian with mean and sigma for this energy distribution.

In addition, for details on the analysis, the Figure 60 presents the scatter plot between the target and NN estimation. It is observed to be a good agreement for all groups of cells, reinforcing the efficiency of the proposed method.

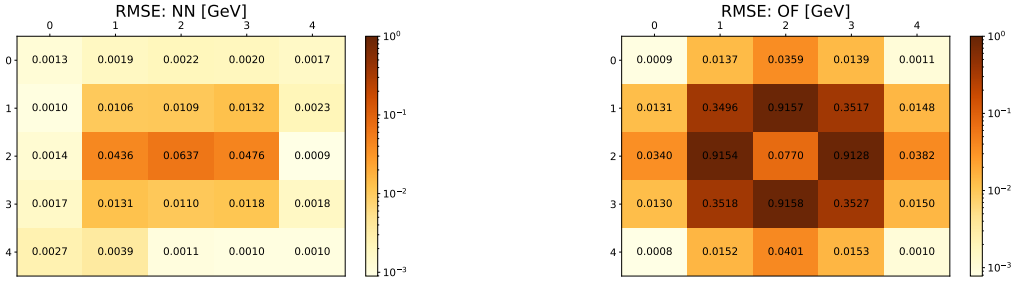
Figure 60 – Scatter plots for energy reconstructed by NN estimator, and target. Additional information is provided with a histogram for NN output.



(a) Diagonal. **(b)** Center. **(c)** Perpendicular.

Figure 61 presents the errors between estimated energy with NN and the target, and between the OF and target. It is observed that the NN model produces small errors for the whole cluster. Differently, the OF estimation generates considerable errors for the first neighborhood of the hottest cell.

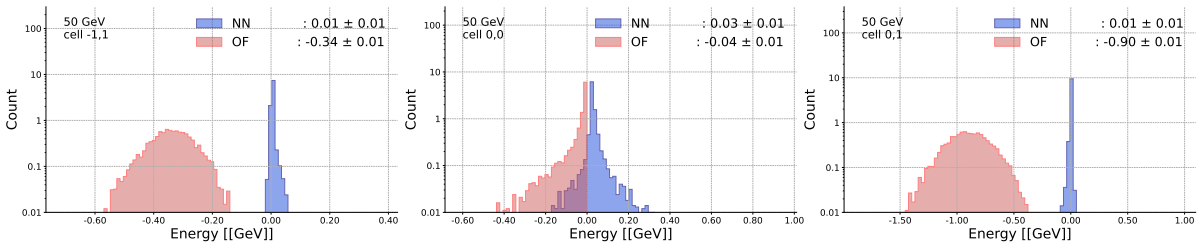
Figure 61 – RMSE over 10,000 events for the energy reconstruction between NN and target in (a), and between OF and target in (b) at the cell level.



(a) **(b)**

Figure 62 shows the energy reconstruction error histograms for NN and OF. The NN error has a lower spread than OF for all cells. Even in the cell group with low energy, the error spread is less than 1 GeV. These plots indicate that it is possible to achieve a better energy estimation using a NN estimator than with OF.

Figure 62 – Energy regression error at the diagonal, at the center, and at the perpendicular cells for an NN structure using one HL. Even in cells at the corners with low energies, the NN error spread is lower than OF.



(a) Diagonal. **(b)** Center. **(c)** Perpendicular.

4.2 Analysis on Lorenzetti data

The Lorenzetti Showers is a high-energy calorimeter simulation framework that allows to access low-level information such as the sensor's readout pulses (ARAÚJO et al., 2023). Using this simulator, the crosstalk modeling presented in Chapter 2 was implemented, and an ATLAS-like calorimeter structure was used to produce electrons and positrons from $Z -> e^{-}e^{+}$ decays. To evaluate an ML model to mitigate XT effects, simulations were generated at 20 different impact energy levels (E_{imp}), each with ∼100,000 events, totaling approximately two million events.

This section presents the results achieved using an NN estimator using as input of 196 features, 4 per cell in the 7×7 cluster as input, three HL, and 196 features as output. The models to mitigate XT effects were evaluated according to the MSE metric, and the regression description by the coefficient of determination (R^2, see Table 19. After that, the OF estimates the energy ($\hat{E}$), and time of flight ($\hat{\tau}$). A series of analyses were performed to evaluate the effectiveness of the proposed mitigation method: $\frac{\sigma_E}{E}$, to evaluate the performance evolution for the method over the whole dataset produced in Lorenzetti; σ_τ, is expected that estimation growing-up when the impact energy increases; the shower-shapes to evaluate how the shower development, at lateral point of view η, and ϕ direction, is affected by the XT effects, and how the mitigation method acting reducing they. Details on results are in section D.2.

4.2.1 Time Regression

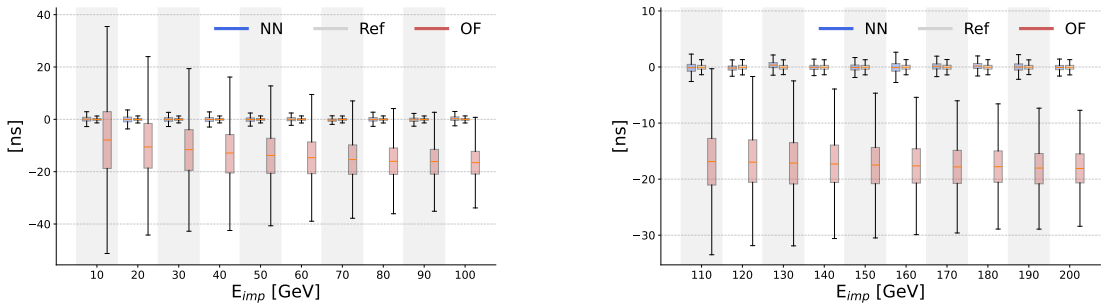
Table 12 shows the summary for the mean τ NN estimate in the center cluster for 4 energies ranges compared to the OF and Ref values. Figure 63 shows the boxplots for all impact energies in this study. The τ estimation from OF produces a higher error in comparison to the NN estimator. Detailed tables and plots for the results are shown in Appendix D.

Table 12 – Mean τ values on the cluster reconstructed per impact energy. The first column is the reference values, the second, is values applying OF on samples with XT, and the third, is the values from OF after an NN estimator to regress samples. The last column brings the RMSE ratio between OF and the NN estimator.

E_{imp} [GeV]	τ_{REF} [ns]	OF		NN		$\frac{RMSE_{OF}}{RMSE_{NN}}$
		$\hat{\tau}$	RMSE	$\hat{\tau}$	RMSE	
10	-0.04 ± 0.50	15.30 ± 13027.86	4164.83	0.18 ± 338.88	158.87	26.22
70	-0.05 ± 0.50	-16.94 ± 3746.86	1551.10	-0.22 ± 43.82	12.07	128.48
150	-0.04 ± 0.50	-22.32 ± 2861.35	1347.95	-0.09 ± 46.24	7.53	178.89
200	-0.04 ± 0.50	-21.36 ± 1771.75	623.73	-0.10 ± 2.91	1.64	379.94

Figure 63 presents a summary of the training for each E_{imp}. The OF error when estimating the time of flight is considerable in comparison with the NN estimator. A second observation is about the spread of the results.Figure 63a presents a high spread at 10 GeV, the highest, roughly 70 ns, and the smaller at 100 GeV, ∼ 30 ns. In Figure 63b the spread still large for the time of flight estimation, but a bit smaller for this set of energies, ∼ 30 ns at 110 GeV, and ∼ 28 ns for 200 GeV.

Figure 63 – Boxplot to present the τ regression after the NN estimator. For each energy three boxplots are presented: the first for the NN estimator, the second for the Ref, and the third for the OF. (a) presents the boxplots from 10 up to 100 GeV. (b) presents the boxplots from 110 up to 200 GeV.

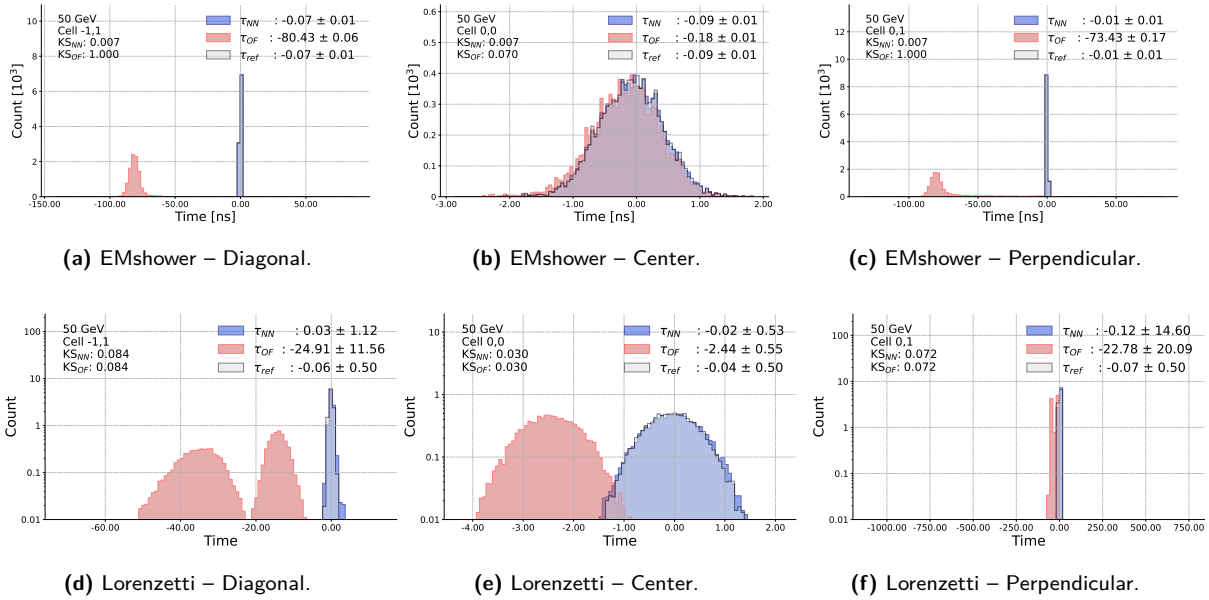


(a) (b)

EMshower × Lorenzetti Results

Figure 64 brings a method comparison between datasets EMshower and Lorenzetti for impact energy equal to 50 GeV. The method can achieve good results for both datasets. One point to mention is that the mean OF error is smaller in Lorenzetti (-24.91 ns) than with EMshower data (-80.43). One possible explanation is concerning the production of the datasets. On Lorenzetti, the physics models are rich in particle physics parameters due to GEANT4 and Pythia8 generators.

Figure 64 – Time regression histograms for EMshower, and Lorenzetti results for $E_{imp} = 50$ GeV to compare. It presents the comparison between the hottest cells, cells at the diagonal, and perpendicular positions in the cluster.



(a) EMshower – Diagonal. **(b)** EMshower – Center. **(c)** EMshower – Perpendicular.

(d) Lorenzetti – Diagonal. **(e)** Lorenzetti – Center. **(f)** Lorenzetti – Perpendicular.

4.2.2 Energy Regression

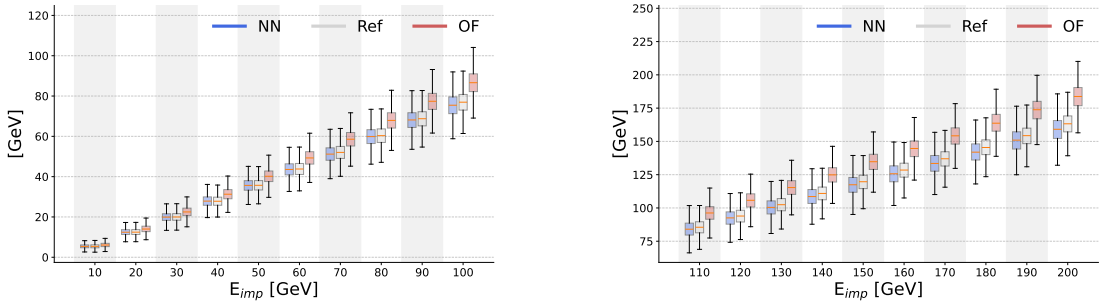
For simplicity Table 13 shows the results for 4 impact energies values. The NN estimator shows better results taking into account the RMSE values. The total energy in the cluster is presented. As the impact energy increases, the RMSE grows, once the XT is proportional to the impact energy. The NN estimator achieves an error 2.69 times lower than OF for 10 GeV, and for 200 GeV it is reduced 3,54 times. Appendix D shows all tables and plots, per impact energy.

Table 13 – Total energy cluster reconstructed by impact energy for 10 GeV, 50 GeV, 150 GeV, and 200 GeV. Column E_{REF} is the reference value sampled at S2 layer for energies without XT. Column OF refers to the values applying OF on samples with XT, and column NN refers to the values from OF after an NN estimator to regress samples. The last column is an RMSE ratio between OF and NN estimator.

E [GeV]	E_{REF} [GeV]	OF		NN		$\frac{RMSE_{OF}}{RMSE_{NN}}$
		E [GeV]	RMSE	E [GeV]	RMSE	
10	5.45 ± 0.01	6.13 ± 0.01	0.7004	5.44 ± 0.01	0.2643	2.65
70	52.00 ± 0.03	58.54 ± 0.03	6.5640	51.19 ± 0.03	1.5554	4.22
150	119.50 ± 0.04	134.53 ± 0.05	15.0639	117.21 ± 0.05	4.2093	3.58
200	162.79 ± 0.05	183.27 ± 0.06	20.5217	159.23 ± 0.06	5.6896	3.61

Figure 65 presents the boxplots for energy ($\hat{E}$) over the 20 E_{imp} energies from Lorenzetti. For the energy, the OF produces a good estimation, for all energy values evaluated.

Figure 65 – Boxplots to present the energy regression results. For each energy three boxplots are presented: the first for the NN estimator, the second for the Ref, and the third for the OF. (a) presents the boxplots from 10 up to 100 GeV. (b) presents the boxplots from 110 up to 200 GeV.



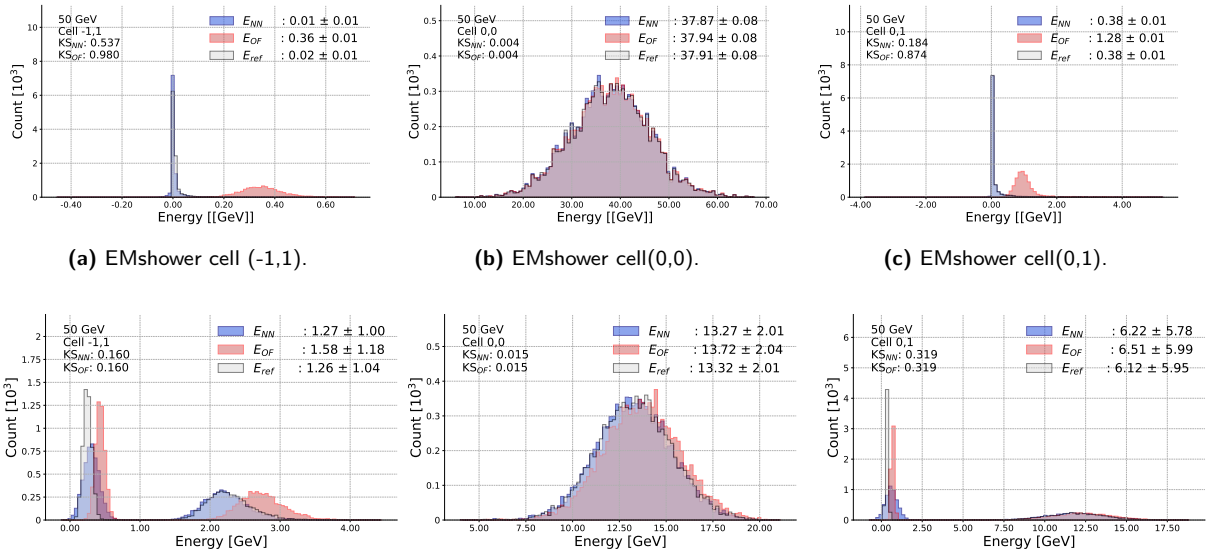
(a) **(b)**

Detailed tables for each E_{imp} are presented in Appendix D.

EMshower × Lorenzetti Results

As was made for the time of flight, a comparison between the EMshower and Lorenzetti results on energy estimation is presented in Figure 66. The results are similar in this estimation. The OF design can estimate energy well in comparison with the time of flight. The method shows a small KS, indicating similarity among distributions with the Ref values.

Figure 66 – Histograms for EMshower, and Lorenzetti results for $E_{imp} = 50$ GeV to compare .



(a) EMshower cell (-1,1). **(b)** EMshower cell(0,0). **(c)** EMshower cell(0,1).

(d) Lorenzetti cell (-1,1). **(e)** Lorenzetti cell (0,0). **(f)** Lorenzetti cell (0,1).

4.2.3 Standard Deviation Analysis

Standard deviation analysis plays a meaningful metric in particle physics as it allows us to estimate the expected fluctuations as a function of energy. To have a good particle characterization, two values were calculated for each E_{imp}: σ_τ and $\frac{\sigma_E}{E}$. The first term indicates the fluctuation in the time of flight estimation for a given particle that passes through the 5×5 cluster. The second is the energy-normalized fluctuation for the total estimated energy.

Figure 67 shows the σ_τ analysis for the set of E_{imp} generated by the Lorenzetti simulator. The standard algorithm is the Optimal Filter (OF), in red, a linear estimation based on the Lagrange multiplier, and the second

estimator is the Neural Network (NN) based on the perceptron multilayer, in blue. In a cell cluster, the time information is the same once the particle passes through the cluster and deposits energy in a nonuniform distribution. So, it is possible to estimate the time of flight considering any cell in the cluster, but, the best estimation is achieved in the cell that registers the high energy level on the cluster. To have an evaluation for this estimation, σ_τ is used, and the XT effects distort this evaluation as the Figure 67 shows the large error on time estimation for OF. The NN estimator achieves a reduction in the XT effects, showing a better time estimation compared to the standard algorithm in at least one magnitude order.

Figure 67 – Analysis for the σ_τ over a 5×5 cluster after estimation of the time of flight using the standard method, OF, in red, and the NN estimator, in blue. The RMSE is evaluated concerning the target value to OF and the NN is presented. All NN estimators trained utilize a single hidden layer to make an XT mitigation at the sample level. Every point in the plot is the result evaluated with a specific E_{imp} value, from 10 GeV to 200 GeV.

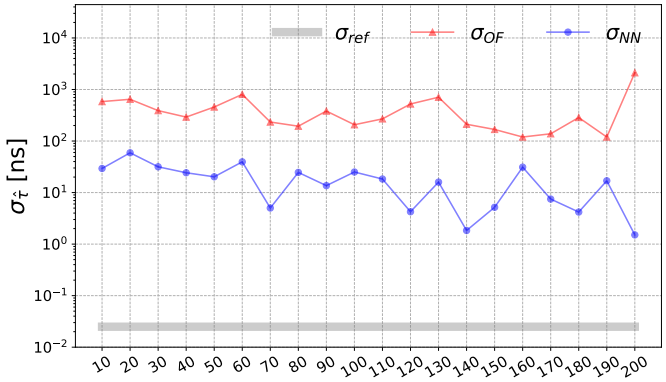
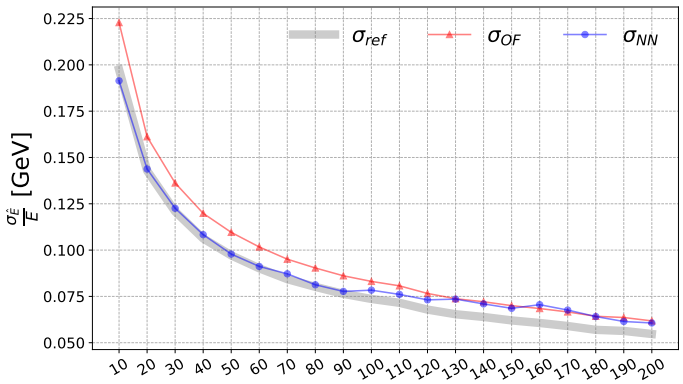


Figure 68 summarizes the evolution of the σ_E for the data produced in the Lorenzetti framework. To get a good comparison three conditions were evaluated. The reference value, which means the expectation for an event without XT contribution. The second result is the OF results to reconstruct the energy level considering the XT, and the last one is the evaluation using an NN estimator to mitigate XT. Using the NN estimator with energy until 100 GeV was possible to achieve a good estimation better than OF, for the other energy levels both algorithms achieve the same σ_E on energy estimation.

Figure 68 – Analysis for the $\frac{\sigma_E}{E}$ parameter over all 20 values for E_{imp} on the data produced on the Lorenzetti framework simulator. The large line in gray is the parameter evaluated for the energy level without XT. The thin line in gray σ_E for the energy and XT contribution. In red, σ_E after OF processing energy signal with XT. In blue, is the evaluation after the NN estimator processing the signal to mitigate XT effects.



Calorimeters are expected to present an energy resolution ($\frac{\sigma_E}{E}$) that decreases with energy (WIGMANS, 2017).

This behavior was observed in our results, but the crosstalk degrades the resolution curve achieved with the OF. By using the proposed method, it is possible to regress to a resolution profile closer to the target.

4.2.4 Shower Shapes Analysis

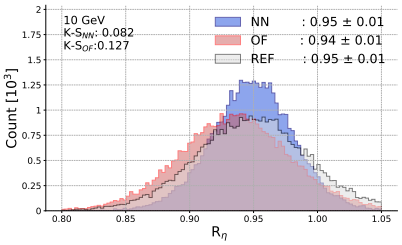
Shower shape variables, such as R_η, R_ϕ, and $W_{\eta 2}$, are essential attributes for electromagnetic and hadronic particle characterization, which are traditionally used for particle identification in online (trigger) and offline analysis. The crosstalk effect changes the profile of the shower shapes, degrading the characterization of the particles.

R_η - Shower width in η direction.

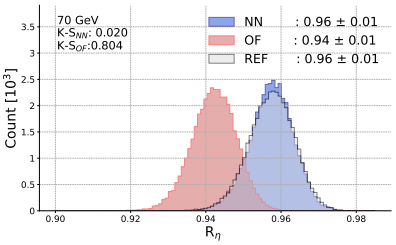
Recalling that every cell in the cluster has a coordinate (η, ϕ), this variable estimates the energy ratio evaluated between a 3×7 cluster and a 7×7 cluster. A narrow shower, such as the ones produced by electrons is expected to present R_η, close to 1.

Figure 69 shows a comparison among three situations: a reference value, the OF output, and the NN estimator. Looking at the different E_{imp} values it is possible to see a shift in R_η distribution in OF estimation due to XT effects which is smaller for low energy and increase with the energy range. On the other hand, the NN results produce a good sample regression reducing XT effects once the distribution is close to the distribution of the reference. Figure 83 presents R_ϕ in details for all 20 impact energy levels evaluated.

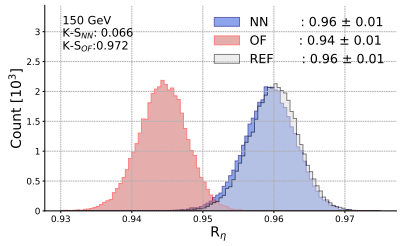
Figure 69 – R_η variable for E_{imp} energies, 10 GeV, 70 GeV, 150 GeV, and 200 GeV. In light grey, the R_η reference value without XT contribution. In light red, the OF output for the energy reconstruction considers XT contribution, and in light blue the NN estimator as a preprocessing layer to mitigate XT effects over cell samples before reconstructing energy with OF.



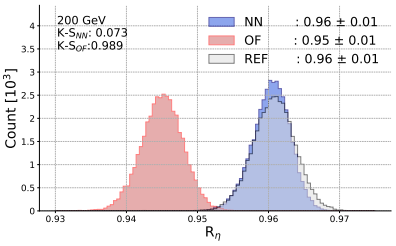
(a) E_{imp}: 10 GeV.



(b) E_{imp}: 70 GeV.



(c) E_{imp}: 150 GeV.

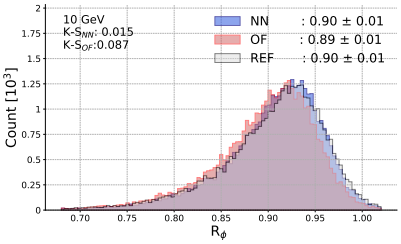


(d) E_{imp}: 200 GeV.

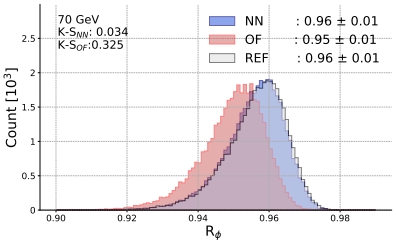
R_ϕ - Shower width in ϕ direction.

Figure 70 shows the R_ϕ evaluation for the same energy levels presented for R_η. Similarly, it is possible to observe that XT distorts the distribution adding a shift that increases proportionally to the energy level. Once again the proposed model removed the distribution shifts. Figure 84 presents R_ϕ in details for all 20 impact energy levels evaluated.

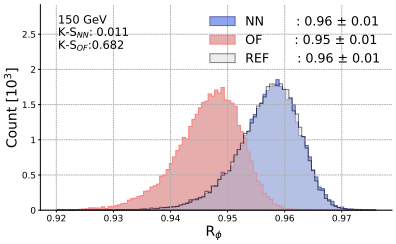
Figure 70 – R_ϕ variable. In light grey the ratio for the reference value without XT contribution. In light red the OF output for the energy reconstruction taking into account XT contribution, and in light blue the NN estimator as a preprocessing layer to mitigate XT effects over cell samples before reconstructing energy with OF.



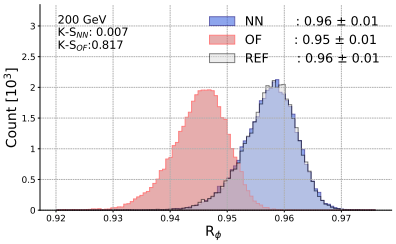
(a) E_{imp}: 10 GeV.



(b) E_{imp}: 70 GeV.



(c) E_{imp}: 150 GeV.

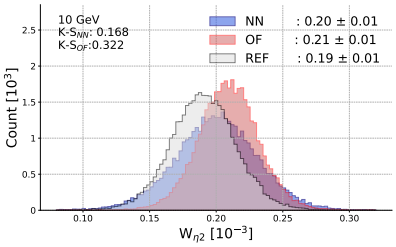


(d) E_{imp}: 200 GeV.

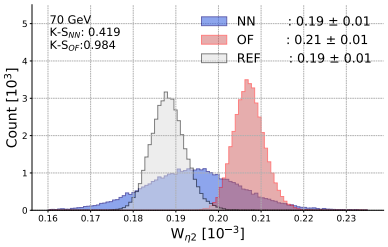
$W_{\eta 2}$ - Shower width.

Figure 71 shows how the width of the shower changes with energy levels when XT effects are present in the samples. This variable shows a large spread for the E_{imp} equal to 20 GeV in comparison with the reference and OF distribution. For this variable, the XT effect appears at a low level of the impact energy, because the displacement between OF and reference distribution starts at 10 GeV. Figure 85 presents $W_{\eta 2}$ for all 20 impact energy levels evaluated.

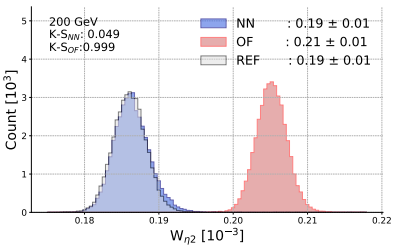
Figure 71 – $W_{\eta 2}$ indicates the width of the shower shape for an EM particle in a window 3×5. In the light gray are the reference values to show the expected value. In light red the OF processing samples with XT, and in light blue the OF processing samples after an NN estimator process samples to mitigate XT effects..



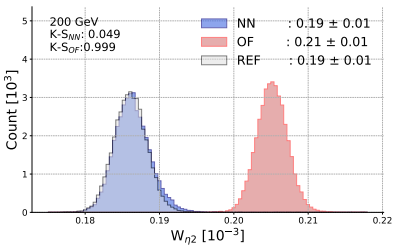
(a) E_{imp}: 10 GeV.



(b) E_{imp}: 70 GeV.



(c) E_{imp}: 150 GeV.



(d) E_{imp}: 200 GeV.

4.3 Summary of Results

Table 14 – Summary of results presented in this chapter on XT mitigation using an NN estimator.

EMshower Data	
τ Estimation	A significant reduction on RMSE (5068.02 times for diagonal cell, 4743.53, for perpendicular cell, and 14,57 times in the hottest cell) error with respect to the OF estimation looking for diagonal, perpendicular, and the hottest cell.
Energy Estimation	N estimator achieve a good energy estimation and reduces the RMSE error, from 1.21 to 83.69 times, for the hottest cell, and perpendicular cell, respectively.
An additional evaluation was made using the Kolmogorov-Smirnov Test (KS), which shows that the NN estimator produces a similar distribution for energy and time-of-flight estimations..	
Lorenzetti Data	
τ Estimation	Histograms for diagonal, perpendicular and hottest cell at 3×3 cluster shows a good estimation. A boxplot analysis shows that the NN estimator produces a good estimation, which was reinforced by the KS-test over a set of 20 E_{imp} energies. Standard deviation analysis shows that the NN can reduce the $\hat{\tau}$ estimations error between 15 dB and 50 dB in relation to OF.
Energy Estimation	Standard deviation shows that is possible to reduce the error for OF between 1.16% and 15.57% among the 20 E_{imp} energies evaluated in this work. Shower shapes (R_η, R_ϕ, $W_{\eta 2}$) analysis indicates that the NN estimator can reduce the error due to XT presence on the estimation.
Comparison between EMshower, and Lorenzetti results shows that the proposal method can mitigate XT effects.	

5 Conclusion

5.1 Introduction

This chapter summarizes the conclusions drawn from the development of an EMShower simulator designed to generate single-electron showers. The simulator supports the creation of a crosstalk (XT) mitigation method. The conclusions are based on analyses using EMShower data and the Lorenzetti (LZT) framework.

In high-energy physics, precise particle detection and characterization are crucial for validating theoretical models, discovering new physics, and achieving higher precision measurements. However, these analyses are challenging due to high energy levels, fast interaction rates, and densely packed sensor arrays. In this complex environment, unwanted effects like crosstalk can distort energy and timing measurements.

This work introduces a Machine Learning (ML)–based approach to mitigate XT distortions, estimating the energy and time of flight for single electron simulations. Using two datasets, one from the analytical simulator developed in this work with 745,000 clusters, a neural network (NN) estimator based on the Multilayer Perceptron (MLP) architecture demonstrates the feasibility of mitigating XT effects to achieve accurate energy and time estimations for particles interacting with a liquid argon (LAr) calorimeter. A second data set from the LZT framework, consisting of 20 energy levels with 100,000 electron events each, was used to develop a neural network with three hidden layers, which regressed cell samples before applying optimal filtering to estimate the energy and time measurements. The results are good compared to the standard method.

5.2 EMshower simulator

The EMshower simulator demonstrates a comparable energy distribution within a specified cluster in the detector. The results using the machine learning approach suggest that it is feasible to achieve reliable energy reconstruction by using a single hidden layer HL neural network with 20 neurons. The Neural Networks (NN) outputs indicate that this machine learning method effectively handles cell signals even in the presence of crosstalk (XT), a challenge that conventional Optimal Filtering (OF) techniques struggle to address.

The summary of the results is as follows:

- EMShower Modeling – The modeling for energy deposition in the ATLAS LAr considering real parameters has successfully converged. The simulator data shows an energy deposition pattern similar to that observed in real experimental data.
- Supervised methods – Showed good results for the estimation of energy and time of flight mitigating XT effects.
  - Energy Regression – A NN using a single HL has produced a better regression performance than Optimal Filtering (OF) technique. It is possible to organize the results of NN in three groups: The hottest cell with a good estimation for both methods, cells under inductive XT (the diagonal cells of 3×3 cluster) that energy deposition is below 1 GeV, the OF produces a large error to estimate the energy. The expectation was 0.10 GeV, and OF achieved 0.38 GeV. The last group is cells with capacitive and inductive XT (perpendicular) cells with an energy deposition rather than 1 GeV, which the NN estimator is better than OF reducing the estimation error by around 38%.
  - Time Regression – A NN with three HL provides better time reconstruction than the OF. In particular, the time regression using NN is better than OF for all cells in the first neighborhood. The RMSE achieved for the NN estimator is equal to 0.01 ns against 0.57 ns on the hottest cell, and 80.50 ns for the worst time estimation.

The summary of those results indicates that the non-specialist approach (one NN model to regress cluster information) based on MLP architecture produces a good energy and time reconstruction. The low error achieved in comparison with the OF indicates that this approach could mitigate the XT influence in ATLAS LAr EMCAL.

5.3 Lorenzetti Showers Framework (LZT)

The NN model provided stable time-of-flight estimates, reducing the estimation error achieved by the standard algorithm (OF), and achieved energy estimates comparable to those obtained with optimal filtering. Additionally, the model was evaluated using three descriptive variables for electromagnetic particle characterization, demonstrating effective XT mitigation at the cell-sample level.

This method requires detailed knowledge of XT signals, and for applications like ATLAS, this represents a limitation that could be addressed with a data-driven approach. This requires the development of the specialist knowledge approach based on the model described in the section C.1. Importantly, the MLP-based neural network effectively captures XT effects; however, using MSE against true values should be replaced with a data-driven quality measure for more accurate evaluation.

Embedding the XT model in the Lorenzetti Showers Framework (LZT) allows us to proceed with additional analysis like shower shapes and sigma analysis to validate the XT mitigation approach.

Using Lorenzetti Showers Framework (LZT) data, the summary of the results is as follows:

- XT Mitigation: The ML-based approach effectively mitigated crosstalk (XT) effects, leading to improved energy and time estimations.
- Time-of-Flight: The neural network estimator produced stable time-of-flight estimates, reducing estimation error compared to standard algorithms. The RMSE of the NN estimator oscillates between 1.38 and 316.52 ns for impact energies equal to 200 GeV and 30 GeV, respectively. OF RMSE oscillates from 523.52 ns to 15280.84 ns for impact energies equal to 190 GeV and 50 GeV, respectively.
- The model produced energy estimates comparable to those obtained using optimal filtering (OF) techniques. The root mean squared error (RMSE) for the neural network (NN) estimator oscillates between 0.26 GeV at 10 GeV and 5.79 GeV at 200 GeV. In comparison, the OF RMSE oscillates between 0.70 and 20.51 GeV over the same energy range. The NN estimator achieves estimation errors between 2.15% at 40 GeV and 3.37% at 170 GeV, significantly outperforming the OF estimator, which exhibits errors between 10.62% at 190 GeV and 38.25% at 20 GeV.
- Evaluation: The results were assessed using three descriptive variables for the characterization of electromagnetic particles, which demonstrated effective XT mitigation at the cell sample level.
  - R_η – The analysis with this variable shows a sensitivity with respect to the NN estimator at low energy. In 7 values for impact energy [10, 70] GeV the NN estimator didn't get a better R_η distribution. This suggests that the NN estimator for these energies should be optimized to achieve a better NN model.
  - R_ϕ – This variable shows that the NN estimator can produce a good reconstruction.
  - $W_{\eta 2}$ – As occurs with R_ϕ variable the NN estimator presents a good regression mitigating XT effects.
- σ analysis:
  - $\frac{\sigma_E}{E}$ – The NN estimator achieves an improvement compared to the OF estimation, reducing the error from 1.1% to 15.3% relative to the OF reference value for 180 GeV and 10 GeV, respectively.
  - σ_τ – In this parameter, the NN estimator reduces the estimation error in two magnitude orders. The expectation is close to zero, and OF achieves a minimum of 10^1 at 200 GeV, and a worse value of 10^2 at 10 GeV. The NN estimator oscillates between $\sim$0.8 at 200 GeV, and $\sim 3 \cdot 10^1$ at 20 GeV.
- Neural Network Performance: The MLP-based neural network effectively captured XT effects, though the evaluation metric (MSE against true values) should be replaced by a more suitable data-driven quality measure.

This work motivates the continuation of developing an XT mitigation model based on Monte Carlo (MC) simulations and data collision, to be carried out in a Ph.D. thesis by a colleague, Mateus Hufnagel, in the ATLAS collaboration in Brazil at UFJF (Federal University of Juiz de Fora).

5.4 Future works and perspectives

The investigation concerning non-supervised methods is being developed in the UFBA (Federal University of Bahia) using AutoEncoder (AE) approach to extract sample information without XT contribution.

Some points could be explored and investigated to improve the XT coupling model to build a better mitigation approach:

- The model for XT coupling should be optimized regarding its inductive nature once this work considers the transfer function for this coupling equal to the capacitive transfer function.
- In this work the model of coupling considers the linear contribution which was an assumption to allow the development of the study.
- An XT mitigation approach based on specialist knowledge could allow a better XT characterization according to the interacting region inside the LAr.
- To investigate different ML models could get a better estimation than the MLP structure.
- To investigate an OF algorithm optimization to improve the estimation reducing the sensitivity to the XT couplings.

References

AAD, G. et al. Topological cell clustering in the ATLAS calorimeters and its performance in LHC run 1. *Eur. Phys. J.*, C77, p. 490, 2017. Available at: <https://link.springer.com/content/pdf/10.1140%2Fepjc%2Fs10052-017-5004-5.pdf>. Cited 2 times on pages 25 and 27.

AAD, G.; et al. Alignment of the ATLAS Inner Detector in Run 2. *The European Physical Journal C*, v. 80, n. 12, p. 1194, Dec 2020. ISSN 1434-6052. Available at: <https://doi.org/10.1140/epjc/s10052-020-08700-6>. Cited on page 25.

ABDELALIM, A. A. *Study on the Impact of Cross-Talk in the ATLAS Electromagnetic Calorimeter on the Signal Prediction in the Strip Layer*. Dissertation (Master of Science) — Università de Genève, 2008. Presented Sept 2008. Available at: <http://cds.cern.ch/record/1374681>. Cited on page 29.

AHARROUCHE, M. et al. Response uniformity of the atlas liquid argon electromagnetic calorimeter. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, v. 582, n. 2, p. 429 – 455, 2007. ISSN 0168-9002. Available at: <http://www.sciencedirect.com/science/article/pii/S0168900207018591>. Cited on page 30.

ALLAIRE, C. *A High-Granularity Timing Detector (HGTD) in ATLAS: Performance at the HL-LHC*. Geneva, 2018. Available at: <https://cds.cern.ch/record/2302827>. Cited on page 12.

ANJOS, A. dos et al. Neural triggering system operating on high resolution calorimetry information. *Nuclear Instruments and Methods In Physics Research A*, v. 559, p. 134–138, Abril 2006. Provided By The Sao/Nasa Astrophysics Data System. Available at: <http://adsabs.harvard.edu/abs/2006nimpa.559..134d>. Cited 2 times on pages 129 and 132.

APOLLINARI, G. et al. (Ed.). *High-Luminosity Large Hadron Collider (HL-LHC) - Technical Design Report V. 0.1*. Geneva: CERN, 2017. v. 4. (CERN Yellow Reports: Monographs, v. 4). Available at: <https://e-publishing.cern.ch/index.php/CYRM/issue/view/40>. Cited on page 23.

ARAÚJO, M. et al. Lorenzetti showers - a general-purpose framework for supporting signal reconstruction and triggering with calorimeters. *Computer Physics Communications*, p. 108671, 2023. ISSN 0010-4655. Available at: <https://www.sciencedirect.com/science/article/pii/S0010465523000164>. Cited 6 times on pages 4, 5, 39, 43, 44, and 59.

ARPAIA, P. et al. Machine learning for beam dynamics studies at the cern large hadron collider. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, v. 985, p. 164652, 2021. ISSN 0168-9002. Available at: <https://www.sciencedirect.com/science/article/pii/S0168900220310494>. Cited 3 times on pages 4, 129, and 133.

ASHTON, M. et al. *Status report on the RD12 project: timing, trigger and control systems for LHC detectors*. Geneva, 2000. Available at: <https://cds.cern.ch/record/421208>. Cited on page 28.

ATLAS Collaboration. *ATLAS calorimeter performance: Technical Design Report*. Geneva: CERN, 1996. (Technical Design Report ATLAS). Available at: <https://cds.cern.ch/record/331059>. Cited on page 16.

ATLAS Collaboration. *ATLAS detector and physics performance: Technical Design Report, 1*. Geneva: CERN, 1999. (Technical Design Report ATLAS). ISBN 9290831413, 9789290831419. Available at: <https://cds.cern.ch/record/391176>. Cited on page 94.

ATLAS Collaboration. ATLAS experiment at CERN Large Hadron Collider. *Journal of Instrumentation*, IOP Publishing, v. 3, n. S08003, p. 1–407, 2008. Available at: <https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08003>. Cited 2 times on pages 125 and 126.

ATLAS Collaboration. ATLAS fact sheet : To raise awareness of the atlas detector and collaboration on the LHC. 2010. Available at: <https://cds.cern.ch/record/1457044>. Cited 2 times on pages 2 and 25.

ATLAS Collaboration. Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc. *Physics Letters B*, v. 716, n. 1, p. 1–29, 2012. ISSN 0370-2693. Available at: <https://www.sciencedirect.com/science/article/pii/S037026931200857X>. Cited 2 times on pages 11 and 129.

ATLAS Collaboration. A particle consistent with the higgs boson observed with the ATLAS detector at the large hadron collider. *Science*, American Association for the Advancement of Science, v. 338, n. 6114, p. 1576–1582, 2012. ISSN 0036-8075. Available at: <http://science.sciencemag.org/content/338/6114/1576>. Cited 2 times on pages 14 and 122.

ATLAS Collaboration. Measurement of the photon identification efficiencies with the ATLAS detector using LHC run-1 data. measurement of the photon identification efficiencies with the ATLAS detector using LHC run-1 data. *Eur. Phys. J. C*, v. 76, n. CERN-EP-2016-110. 12, p. 666. 57 p, Jun 2016. ISSN 1434-6052. Available at: <https://cds.cern.ch/record/2158117>. Cited on page 22.

ATLAS Collaboration. *Technical Design Report for the Phase-II Upgrade of the ATLAS LAr Calorimeter*. Geneva, 2017. Available at: <https://cds.cern.ch/record/2285582>. Cited 4 times on pages 22, 28, 29, and 125.

ATLAS Collaboration. Observation of higgs boson production in association with a top quark pair at the LHC with the ATLAS detector. 2018. Cited on page 122.

ATLAS Collaboration. Electron and photon performance measurements with the atlas detector using the 2015–2017 lhc proton-proton collision data. *Journal of Instrumentation*, v. 14, n. 12, p. P12006, dec 2019. Available at: <https://dx.doi.org/10.1088/1748-0221/14/12/P12006>. Cited on page 47.

ATLAS Collaboration. Electron reconstruction and identification in the atlas experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13$ TeV. *The European Physical Journal C*, v. 79, n. 8, p. 639, Aug 2019. ISSN 1434-6052. Available at: <https://doi.org/10.1140/epjc/s10052-019-7140-6>. Cited on page 42.

ATLAS Collaboration. *The Collaboration*. 2020. Available at: <https://atlas.cern/discover/collaboration>. Accessed on: 03/09/2020. Cited on page 12.

ATLAS Collaboration. Experiment briefing: Keeping the ATLAS inner detector in perfect alignment. General Photo. 2020. Available at: <https://cds.cern.ch/record/2723878>. Cited on page 25.

ATLAS Collaboration. Performance of electron and photon triggers in ATLAS during LHC run 2. *The European Physical Journal C*, v. 80, n. 1, p. 47, 2020. ISSN 1434-6052. Available at: <https://doi.org/10.1140/epjc/s10052-019-7500-2>. Cited on page 132.

ATLAS Collaboration. ATLAS Standard Model. General Photo. 2021. Available at: <http://cds.cern.ch/record/2765709>. Cited on page 10.

AUBERT, B. et al. Development and construction of large size signal electrodes for the atlas electromagnetic calorimeter. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, v. 539, n. 3, p. 558–594, 2005. ISSN 0168-9002. Available at: <https://www.sciencedirect.com/science/article/pii/S0168900204023563>. Cited on page 28.

BAILLIE, O. V. et al. *Status report on the RD-12 project: timing, trigger and control systems for LHC detectors*. Geneva, 1997. Available at: <https://cds.cern.ch/record/338930>. Cited on page 28.

BANFI, D.; DELMASTRO, M.; FANTI, M. Cell response equalisation of the ATLAS electromagnetic calorimeter without the direct knowledge of the ionisation signals. *Journal of Instrumentation*, IOP Publishing, v. 1, n. 08, p. P08001–P08001, aug 2006. Available at: <https://doi.org/10.1088%2F1748-0221%2F1%2F08%2Fp08001>. Cited 2 times on pages 35 and 36.

BERKELEY LAB. *Bevatron Site Recognized for Historical Contributions to Physics*. 2021. Available at: <https://newscenter.lbl.gov/2021/06/25/bevatron-contributions-to-physics/>. Accessed on: 26/12/2024. Cited on page 8.

BERTACCHI, V. *DeepCore: Convolutional Neural Network for high p_T jet tracking*. Geneva, 2019. Available at: <https://cds.cern.ch/record/2797737>. Cited on page 129.

BOULAY, M. G. et al. SiPM cross-talk in liquid argon detectors. *Frontiers in Physics*, v. 11, 2023. ISSN 2296-424X. Available at: <https://www.frontiersin.org/Articles/10.3389/fphy.2023.1181400>. Cited on page 4.

BOYD, J. *LHC Run-2 and Future Prospects*. [S.l.], 2020. 9 pages, 6 figure, proceedings from lecture given at CERN-JINR ESHEP 2019 summer school in St Petersburg, Russia. Submitted for publication in a CERN Yellow Report. Available at: <https://cds.cern.ch/record/2707815>. Cited on page 4.

BRAIBANT, S.; GIACOMELLI, G.; SPURIO, M. *Particles and Fundamental Interactions: An Introduction to Particle Physics*. 2. ed. [S.l.]: Springer Netherlands, 2012. (Undergraduate Lecture Notes in Physics). ISBN 9400724632,9789400724631. Cited 2 times on pages 9 and 10.

BROWN, R.; COCKERILL, D. Electromagnetic calorimetry. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, v. 666, p. 47–79, 2012. ISSN 0168-9002. Advanced Instrumentation. Available at: <https://www.sciencedirect.com/science/article/pii/S0168900211005572>. Cited on page 28.

BUCHANAN, N. J. et al. ATLAS liquid argon calorimeter front end electronics. *Journal of Instrumentation*, IOP Publishing, v. 3, n. 09, p. P09003–P09003, sep 2008. Available at: <https://doi.org/10.1088%2F1748-0221%2F3%2F09%2Fp09003>. Cited on page 29.

BUCHANAN, N. J. et al. Design and implementation of the front end board for the readout of the ATLAS liquid argon calorimeters. *Journal of Instrumentation*, IOP Publishing, v. 3, n. 03, p. P03004–P03004, mar 2008. Available at: <https://doi.org/10.1088/1748-0221/3/03/p03004>. Cited on page 41.

BUTALLA, S. D.; HOHLMANN, M. Investigation and mitigation of crosstalk in the prototype me0 gem detector for the phase-2 muon system upgrade of the cms experiment: On behalf of the cms muon group. In: *2020 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*. [S.l.: s.n.], 2020. p. 1–6. Cited on page 3.

Béjar Alonso, I. et al. *High-Luminosity Large Hadron Collider (HL-LHC): Technical design report*. Geneva: CERN, 2020. (CERN Yellow Reports: Monographs). Available at: <https://cds.cern.ch/record/2749422>. Cited on page 2.

CAVALLARI, F. Performance of calorimeters at the LHC. *Journal of Physics: Conference Series*, IOP Publishing, v. 293, p. 012001, apr 2011. Available at: <https://doi.org/10.1088%2F1742-6596%2F293%2F1%2F012001>. Cited on page 28.

CERN. *The LEP story*. Available at: <https://home.cern/news/press-release/cern/lep-story>. Accessed on: 11/02/2020. Cited on page 122.

CERN. LHC season 2. 2015. Available at: <http://cds.cern.ch/record/2059631>. Accessed on: 15 ago. 2020. Cited 4 times on pages 2, 12, 14, and 15.

CERN. *CERN Timelines All Events*. 2016. Available at: <https://timeline.web.cern.ch/events>. Accessed on: 11/02/2021. Cited on page 122.

CERN. *Inter-Experimental LHC Machine Learning Working Group*. 2016. Available at: <https://lpcc.web.cern.ch/sites/default/files/attachments/ml_wg_mandate.pdf>. Accessed on: 07/02/2022. Cited on page 134.

CERN. LHC guide. 2017. Available at: <https://cds.cern.ch/record/2255762>. Cited 3 times on pages 12, 13, and 14.

CERN. *The HL-LHC project*. CERN, 2024. Available at: <https://hilumilhc.web.cern.ch/content/hl-lhc-project>. Accessed on: 18/11/2024. Cited on page 2.

CERN Collaboration. *CERN Homepage*. 2016. Available at: <http://www.cern.ch/>. Accessed on: 21 De Maio De 2016. Cited on page 1.

CHAO, A. W. et al. *Handbook of Accelerator Physics and Engineering*. 3. ed. [S.l.]: World Scientific Pub Co Inc, 2023. ISBN 9811269173,9789811269172. Cited 3 times on pages 8, 12, and 123.

CHOUDHARY, R.; SHUKLA, S. A clustering based ensemble of weighted kernelized extreme learning machine for class imbalance learning. *Expert Systems with Applications*, v. 164, p. 114041, 2021. ISSN 0957-4174. Available at: <https://www.sciencedirect.com/science/article/pii/S0957417420308095>. Cited on page 129.

CLELAND, W.; STERN, E. Signal processing considerations for liquid ionization calorimeters in a high rate environment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, v. 338, n. 2, p. 467 – 497, 1994. ISSN 0168-9002. Available at: <http://www.sciencedirect.com/science/article/pii/0168900294913323>. Cited on page 30.

CLOWE, D. et al. A direct empirical proof of the existence of dark matter*. *The Astrophysical Journal*, v. 648, n. 2, p. L109, aug 2006. Available at: <https://dx.doi.org/10.1086/508162>. Cited 3 times on pages 1, 9, and 14.

CNPEM. *Projeto SIRIUS: a nova fonte de luz síncrotron brasileira*. Campinas: CNPEM, 2014. Available at: <https://www.lnls.cnpem.br/sirius/livro-projeto-sirius/>. Cited on page 124.

COLAS, J. et al. *Crosstalk in the ATLAS Electromagnetic Calorimeter; preliminary version*. Geneva, 1999. Available at: <https://cds.cern.ch/record/683952>. Cited 10 times on pages 19, 1, 3, 4, 34, 35, 36, 50, 94, and 99.

DAMGOV, J.; LITOV, L. Application of neural networks for energy reconstruction. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, v. 482, n. 3, p. 776–788, 2002. ISSN 0168-9002. Available at: <https://www.sciencedirect.com/science/article/pii/S0168900201018514>. Cited 2 times on pages 129 and 132.

DĂRĂBAN, M.; PITICĂ, D. Capacitive discontinuity in cross talk effect. In: *2010 IEEE 16th International Symposium for Design and Technology in Electronic Packaging (SIITME)*. [S.l.: s.n.], 2010. p. 135–140. Cited on page 34.

DAS, A.; FERBEL, T. *Introduction to nuclear and particle physics*. 2nd. ed. [S.l.]: World Scientific Publishing Company, 2003. ISBN 9789812387448,9789812564351,9812387447. Cited on page 10.

De ANGELIS, A.; PIMENTA, M. *Introduction to Particle and Astroparticle Physics: Multimessenger Astronomy and its Particle Physics Foundations*. 2. ed. [S.l.]: Springer, 2018. ISBN 9783319781815,3319781812. Cited 4 times on pages 9, 10, 18, and 122.

DE KERRET, H. et al. The CHOOZ experiment: Proposal to search for neutrino vacuum oscillations to $\delta m^2 = 10^{-3} ev^2$ using a 1 km baseline reactor neutrino experiment. 1993. Available at: <https://inspirehep.net/files/4df82e006e3b807a61c0df26eb229c04>. Cited on page 132.

DENBY, B. Neural networks in high energy physics: A ten year perspective. *Computer Physics Communications*, v. 119, n. 2, p. 219–231, 1999. ISSN 0010-4655. Available at: <https://www.sciencedirect.com/science/article/pii/S0010465598001994>. Cited on page 132.

DRAKOULAKOS, D. G. et al. The high-precision x-ray tomograph for quality control of the atlas mdt muon spectrometer. n. CERN-OPEN-97-023, p. 25 p, Jul 1997. Available at: <https://cds.cern.ch/record/335977>. Cited on page 126.

EBRAHIMI, A. Characteristics and performance of rd53a readout chip with small-pixel silicon sensors. In: *2019 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*. [S.l.: s.n.], 2019. p. 1–5. Cited on page 3.

ELLWANGER, U. *From the Universe to the Elementary Particles: A First Introduction to Cosmology and the Fundamental Interactions*. 1. ed. [S.l.]: Springer, 2012. (Undergraduate Lecture Notes in Physics). ISBN 3642243746,9783642243745. Cited 3 times on pages 9, 10, and 14.

ENGLERT, F.; BROUT, R. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, American Physical Society, v. 13, p. 321–323, Aug 1964. Available at: <https://link.aps.org/doi/10.1103/PhysRevLett.13.321>. Cited on page 10.

EVANS, L.; BRYANT, P. LHC machine. *Journal of Instrumentation*, IOP Publishing, v. 3, n. 08, p. S08001–S08001, aug 2008. Available at: <https://doi.org/10.1088/1748-0221/3/08/s08001>. Cited 2 times on pages 1 and 13.

FABJAN, C. W.; GIANOTTI, F. Calorimetry for particle physics. *Rev. Mod. Phys.*, American Physical Society, v. 75, p. 1243–1286, Oct 2003. Available at: <https://link.aps.org/doi/10.1103/RevModPhys.75.1243>. Cited on page 15.

FERMILAB. *Tevatron*. 2014. Available at: <https://www.fnal.gov/pub/tevatron/tevatron-accelerator.html>. Accessed on: 23 ago. 2020. Cited on page 9.

FICORELLA, A. et al. Crosstalk characterization of a two-tier pixelated avalanche sensor for charged particle detection. *IEEE Journal of Selected Topics in Quantum Electronics*, v. 24, n. 2, p. 1–8, 2018. Cited on page 3.

FLÜCKIGER, F. *How the internet came to CERN*. CERN, 2013. Available at: <https://home.cern/news/opinion/computing/how-internet-came-cern>. Accessed on: 02/09/2020. Cited on page 12.

FREUND, W. S. Trabalho de Conclusão de Curso, *Algoritmo Neural Para Detecção de Partículas em um Calorímetro de Altas Energias*. 2011. Available at: <http://hdl.handle.net/11422/8148>. Cited 3 times on pages 1, 12, and 25.

FREUND, W. S. *Identificação de Elétrons Baseada em um Calorímetro de Altas Energias Finamente Segmentado*. 310 p. Ph.D. Thesis — UFRJ, 2018. Available at: <http://www.pee.ufrj.br/index.php/pt/producao-academica/teses-de-doutorado/tese-1/2016033266--108/file>. Cited on page 16.

GARUTTI, E. *Calorimeters: Energy measurement*. 2012. Available at: <https://www.desy.de/~garutti/LECTURES/ParticleDetectorSS12/L10_Calorimetry.pdf>. Accessed on: 22/09/2020. Cited on page 17.

GHOSH, S. S. et al. *General Model Independent Searches for Physics Beyond the Standard Model*. 1st ed.. ed. [S.l.]: Springer International Publishing;Springer, 2020. (SpringerBriefs in Physics). ISBN 9783030537821,9783030537838. Cited on page 11.

GIRAY, G. A software engineering perspective on engineering machine learning systems: State of the art and challenges. *Journal of Systems and Software*, v. 180, p. 111031, 2021. ISSN 0164-1212. Available at: <https://www.sciencedirect.com/science/article/pii/S016412122100128X>. Cited on page 130.

GLOROT, X.; BENGIO, Y. Understanding the difficulty of training deep feedforward neural networks. In: TEH, Y. W.; TITTERINGTON, M. (Ed.). *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*. Chia Laguna Resort, Sardinia, Italy: PMLR, 2010. (Proceedings of Machine Learning Research, v. 9), p. 249–256. Available at: <http://proceedings.mlr.press/v9/glorot10a/glorot10a.pdf>. Cited 2 times on pages 128 and 129.

GOODFELLOW, I. J.; BENGIO, Y.; COURVILLE, A. *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016. Available at: <http://www.deeplearningbook.org>. Cited 2 times on pages 130 and 131.

GRINDHAMMER, G.; PETERS, S. The parameterized simulation of electromagnetic showers in homogeneous and sampling calorimeters. In: *International Conference on Monte Carlo Simulation in High-Energy and Nuclear Physics - MC 93 Tallahassee, Florida, February 22-26, 1993*. [S.l.: s.n.], 1993. Cited 6 times on pages 16, 20, 21, 22, 43, and 91.

GRINDHAMMER, G.; RUDOWICZ, M.; PETERS, S. The fast simulation of electromagnetic and hadronic showers. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, v. 290, n. 2, p. 469 – 488, 1990. ISSN 0168-9002. Available at: <http://www.sciencedirect.com/science/article/pii/016890029090566O>. Cited 4 times on pages 20, 40, 41, and 91.

GRUPEN, C.; SHWARTZ, B. *Particle Detectors*. 2. ed. Cambridge: Cambridge University Press, 2008. (Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology). ISBN 0521840066,9780521840064,9780511388668. Available at: <https://cds.cern.ch/record/1110991>. Cited 5 times on pages 15, 17, 19, 20, and 22.

HAYKIN, S. *Neural Networks and Learning Machines*. 3rd ed. ed. Nova Jersey, Estados Unidos: Prentice Hall, 2008. Cited 2 times on pages 127 and 129.

HEDBERG, V. *The ATLAS experiment*. 2009. Available at: <http://hedberg.web.cern.ch/hedberg/home/atlas/atlas.html>. Accessed on: 07/12/2009. Cited on page 28.

HEP ML Community. *A Living Review of Machine Learning for Particle Physics*. 2024. Available at: <https://iml-wg.github.io/HEPML-LivingReview/>. Cited on page 4.

HIGGS, P. W. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, American Physical Society, v. 13, p. 508–509, Oct 1964. Available at: <https://link.aps.org/doi/10.1103/PhysRevLett.13.508>. Cited 2 times on pages 2 and 10.

HOHLMANN, M. A model for crosstalk in micropattern gas detectors. In: *2020 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*. [S.l.: s.n.], 2020. p. 1–3. Cited on page 4.

HUBAUT, F. *Crosstalk Measurements in the EM Barrel Module 0 from 99', May 00' and July 00' Beam Tests*. Geneva, 2000. Available at: <https://cds.cern.ch/record/684134>. Cited 4 times on pages 1, 4, 34, and 99.

HUBAUT, F. et al. *Test Beam Measurement of the Crosstalk in the EM Barrel Module 0*. Geneva, 2000. Available at: <https://cds.cern.ch/record/684055>. Cited 3 times on pages 1, 4, and 99.

HUBAUT, F. et al. *Crosstalk in production modules of the Electromagnetic Endcap Calorimeter*. Geneva, 2003. Available at: <https://cds.cern.ch/record/685566>. Cited 2 times on pages 4 and 99.

INAMURA, T. T. Nagaoka's atomic model and hyperfine interactions.(report). The Japan Academy, v. 92B, n. NO. 4, April 2016. ISSN 0021-4280. Cited on page 8.

JAYAKUMAR, R. *Particle Accelerators, Colliders, and the Story of High Energy Physics: Charming the Cosmic Snake*. 1. ed. [S.l.]: Springer-Verlag Berlin Heidelberg, 2012. ISBN 3642220630,9783642220630. Cited on page 12.

KANE, G. The dawn of physics beyond the standard model. *Scientific American*, v. 288, n. 6, p. 68, 2003. ISSN 00368733. Available at: <http://search-ebscohost-com.ez10.periodicos.capes.gov.br/login.aspx?direct=true&db=aph&AN=9725646&lang=pt-br&site=ehost-live>. Cited on page 9.

KELLEHER, J. D. *Deep Learning*. Paperback. [S.l.]: Mit Press, 2019. (The MIT Press Essential Knowledge). ISBN 0262537559,9780262537551. Cited on page 127.

KHANDOGA, M. *Calibration of electron shower shapes, hadronic recoil reconstruction using deep learning algorithms and the measurement of W boson transverse momentum distribution with the ATLAS detector*. PhD Thesis, 2020. Presented 22 Sep 2020. Available at: <https://cds.cern.ch/record/2752635>. Cited 2 times on pages 49 and 50.

KINDT, W.; ZEIJL, H. van; MIDDELHOEK, S. Optical cross talk in geiger mode avalanche photodiode arrays: Modeling, prevention and measurement. In: *28th European Solid-State Device Research Conference*. [S.l.: s.n.], 1998. p. 192–195. Cited on page 3.

KIRK, G.; RAVEN, J. E.; SCHOFIELD, M. *Os filósofos Pré-Socráticos: História crítica com seleção de textos*. 7. ed. [S.l.]: Fundação Calouste Gulbenkian, 2010. ISBN 9723105039, 978-9723105032. Cited on page 7.

KOBAYASHI, M.; MASKAWA, T. Cp-violation in the renormalizable theory of weak interaction. *Progress of Theoretical Physics*, v. 49, n. 2, p. 652–657, 02 1973. ISSN 0033-068X. Available at: <https://doi.org/10.1143/PTP.49.652>. Cited on page 14.

KOLMOGOROV–SMIRNOV-TEST. In: THE Concise Encyclopedia of Statistics. New York, NY: Springer New York, 2008. p. 283–287. ISBN 978-0-387-32833-1. Available at: <https://doi.org/10.1007/978-0-387-32833-1_214>. Cited on page 46.

KRAMMER, M. *Calorimeters*. 2011. Universidade de Santiago de Compostela. Available at: <http://fpmac116.usc.es/twiki/pub/Instrumentation/SiliconDetectors/Krammer_5_Calorimeters.pdf>. Accessed on: 14/09/2020. Cited on page 18.

KRUPA, M.; SOBY, L. Beam intensity measurements in the large hadron collider. In: *Proceedings of the 20th International Conference Mixed Design of Integrated Circuits and Systems - MIXDES 2013*. [S.l.: s.n.], 2013. p. 592–597. ISBN 978-83-63578-00-8. Cited on page 29.

KUGER, F. *Signal Formation Processes in Micromegas Detectors and Quality Control for large size Detector Construction for the ATLAS New Small Wheel*. 195 p. Ph.D. Theses — Julius-Maximilians-Universität Würzburg, Jul 2017. Cited on page 22.

KUMAR, V.; SAHU, M. Evaluation of nine machine learning regression algorithms for calibration of low-cost pm2.5 sensor. *Journal of Aerosol Science*, v. 157, p. 105809, 2021. ISSN 0021-8502. Available at: <https://www.sciencedirect.com/science/article/pii/S0021850221005401>. Cited on page 129.

LABBé, J.; ISHMUKHAMETOV, R. *Crosstalk Measurements in the Electromagnetic Calorimeter during ATLAS Final Installation*. Geneva, 2008. Available at: <https://cds.cern.ch/record/1143373>. Cited 6 times on pages 1, 3, 35, 37, 94, and 99.

LARKOSKI, A. J.; MOULT, I.; NACHMAN, B. Jet substructure at the large hadron collider: A review of recent advances in theory and machine learning. *Physics Reports*, v. 841, p. 1–63, 2020. ISSN 0370-1573. Jet substructure at the Large Hadron Collider: A review of recent advances in theory and machine learning. Available at: <https://www.sciencedirect.com/science/article/pii/S0370157319303643>. Cited 2 times on pages 129 and 133.

LENZI, T. *Development and Study of Different Muon Track Reconstruction Algorithms for the Level-1 Trigger for the CMS Muon Upgrade with GEM Detectors*. Dissertation (Master of Science) — U. Brussels (main), 2013. Available at: <http://inspirehep.net/record/1236817/files/arXiv:1306.0858.pdf>. Cited on page 22.

LIVAN, M.; WIGMANS, R. *Calorimetry for Collider Physics, an Introduction*. 1st ed.. ed. [S.l.]: Springer International Publishing, 2019. (UNITEXT for Physics). ISBN 978-3-030-23652-6;978-3-030-23653-3. Cited 5 times on pages 15, 16, 17, 20, and 21.

LIVINGSTON, M. S.; BLEWETT, J. P. *Particle Accelerators*. [S.l.]: McGraw-Hill Book Company, Inc., 1962. ISBN 0070381402,9780070381407. Cited on page 121.

LOPEZ SOLIS, A. *Contribution to the energy and time measurement of electrons and photons in the ATLAS experiment and search for dark matter production in association to a Higgs boson*. Ph.D. Thesis — Sorbonne, Jul 2017. CERN-THESIS-2017-214. Available at: <https://cds.cern.ch/record/2291698>. Cited 11 times on pages 17, 1, 4, 34, 35, 36, 37, 38, 53, 94, and 99.

MARTIN, B. R. *Nuclear and particle physics: [an introduction]*. 1. ed. [S.l.]: John Wiley, 2006. ISBN 9780470019993,0470019999,0470025328,9780470025321. Cited on page 15.

MARTIN, B. R.; SHAW, G. *Particle Physics*. 3. ed. [S.l.]: Wiley, 2008. (The Manchester Physics Series). ISBN 0470032936,9780470032930. Cited 2 times on pages 9 and 19.

MENG, Z. Performance of the atlas liquid argon calorimeter. In: *Physics at the LHC 2010*. [S.l.: s.n.], 2010. p. 406–408. Cited on page 41.

MONTROSE, M. I. *EMC & the Printed Circuit Board: Design, Theory, & Layout Made Simple*. 1st. ed. Wiley-IEEE Press, 1998. ISBN 078034703X,9780780347038,9780471660903. Available at: <https://onlinelibrary.wiley.com/doi/abs/10.1002/047172310X.ch1>. Cited on page 34.

NATH, P. et al. The hunt for new physics at the Large Hadron Collider. *Nuclear Physics B - Proceedings Supplements*, v. 200, p. 185 – 417, 2010. ISSN 0920-5632. Available at: <http://www.sciencedirect.com/science/article/pii/S0920563210001490>. Cited 2 times on pages 10 and 14.

National Research Council. *Elementary-Particle Physics: Revealing the Secrets of Energy and Matter*. Washington, DC: The National Academies Press, 1998. ISBN 978-0-309-06037-0. Available at: <https://www.nap.edu/catalog/6045/elementary-particle-physics-revealing-the-secrets-of-energy-and-matter>. Cited on page 9.

NAVAS, S. et al. Review of particle physics. *Phys. Rev. D*, v. 110, n. 3, p. 030001, 2024. Cited 7 times on pages 16, 17, 18, 20, 23, 28, and 36.

NIKIFOROU, N. Performance of the ATLAS liquid argon calorimeter after three years of LHC operation and plans for a future upgrade. In: *2013 3rd International Conference on Advancements in Nuclear Instrumentation, Measurement Methods and their Applications (ANIMMA)*. [S.l.: s.n.], 2013. p. 1–12. Cited 4 times on pages 24, 26, 28, and 30.

Nobel Media AB 2020. *The Nobel Prize in Physics 1956*. 2020. Available at: <https://www.nobelprize.org/prizes/physics/1956/summary/>. Accessed on: 14/09/2020. Cited on page 121.

NOULIS, T.; DERADONIS, C.; SISKOS, S. Advanced Readout System IC Current Mode Semi-Gaussian Shapers Using CCIIs and OTAs. *VLSI Design*, Hindawi Publishing Corporation, v. 2007, p. 071684, Apr 2007. ISSN 1065-514X. Available at: <https://doi.org/10.1155/2007/71684>. Cited on page 35.

OSTERMANN, F. *Partículas Elementares e Interações Fundamentais*. 2001. Textos de Apoio ao Professor de Física. Cited 2 times on pages 7 and 8.

PANG, L.-G. Machine learning for high energy heavy ion collisions. *Nuclear Physics A*, v. 1005, p. 121972, 2021. ISSN 0375-9474. The 28th International Conference on Ultra-relativistic Nucleus-Nucleus Collisions: Quark Matter 2019. Available at: <https://www.sciencedirect.com/science/article/pii/S0375947420302827>. Cited 3 times on pages 4, 129, and 133.

Particle Data Group et al. Review of particle physics. *Progress of Theoretical and Experimental Physics*, v. 2020, n. 8, 08 2020. ISSN 2050-3911. 083C01. Available at: <https://doi.org/10.1093/ptep/ptaa104>. Cited 2 times on pages 94 and 125.

PEQUENAO, J.; SCHAFFNER, P. How ATLAS detects particles: diagram of particle paths in the detector. 2013. Available at: <https://cds.cern.ch/record/1505342>. Cited on page 24.

PERALVA, B. S.-M.; SEIXAS, J. M. de. *The TileCal Energy Reconstruction for LHC Run2 and Future Perspectives*. [S.l.], 2015. Comments: 5 pages, 7 figures, LISHEP 2015, 2-9 August 2015, Manaus. Available at: <https://cds.cern.ch/record/2057907>. Cited on page 4.

PERELSTEIN, M. Introduction to collider physics. *Physics of the Large and the Small*, WORLD SCIENTIFIC, Mar 2011. Available at: <http://dx.doi.org/10.1142/9789814327183_0008>. Cited on page 11.

PIMENTA, J. J. M. et al. O bóson de higgs. *Revista Brasileira de Ensino de Física*, scielo, v. 35, p. 1 – 14, 06 2013. ISSN 1806-1117. Available at: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1806-11172013000200006&nrm=iso>. Cited 2 times on pages 9 and 10.

PINHEIRO, L. A.; COSTA, S. S. C. da; MOREIRA, M. A. *Do átomo grego ao Modelo Padrão: os indivisíveis de hoje*. 2011. Textos de Apoio ao Professor de Física. Available at: <https://www.if.ufrgs.br/public/tapf/v22_v6_pinheiro_costa_moreira.pdf>. Cited 3 times on pages 1, 7, and 8.

POVH, B. et al. *Particles and Nuclei: An Introduction to the Physical Concepts*. 7. ed. [S.l.]: Springer-Verlag Berlin Heidelberg, 2015. (Graduate Texts in Physics). ISBN 978-3-662-46320-8,978-3-662-46321-5. Cited 3 times on pages 1, 11, and 19.

PRALAVORIO, P.; SAUVAGE, D. *Review of the crosstalk in the module 0 of the Electromagnetic Endcap Calorimeter*. Geneva, 2001. Revised version number 1 submitted on 2001-02-27 17:43:41. Available at: <https://cds.cern.ch/record/684164>. Cited on page 99.

PRETZL, K. Calorimeters in astro and particle physics. *Journal of Physics G: Nuclear and Particle Physics*, IOP Publishing, v. 31, n. 7, p. R133–R149, may 2005. Available at: <https://doi.org/10.1088/0954-3899/31/7/r01>. Cited on page 18.

PU, Y. et al. Machine learning methods for rockburst prediction-state-of-the-art review. *International Journal of Mining Science and Technology*, v. 29, n. 4, p. 565–570, 2019. ISSN 2095-2686. SI: Recent Advancements in Mine Safety Science and Engineering. Available at: <https://www.sciencedirect.com/science/article/pii/S2095268619302812>. Cited on page 129.

QIN, J. et al. A waveform sampling prototype asic for picosecond-level time measurement. *IEEE Transactions on Nuclear Science*, v. 69, n. 11, p. 2271–2279, 2022. Cited on page 4.

RADOVIC, A. et al. Machine learning at the energy and intensity frontiers of particle physics. *Nature*, v. 560, n. 7716, p. 41–48, Aug 2018. ISSN 1476-4687. Available at: <https://doi.org/10.1038/s41586-018-0361-2>. Cited 3 times on pages 129, 132, and 133.

RAKITIANSKAIA, A.; ENGELBRECHT, A. Measuring saturation in neural networks. In: *2015 IEEE Symposium Series on Computational Intelligence*. [S.l.: s.n.], 2015. p. 1423–1430. Cited on page 128.

RAUNIG, C. et al. A new approach for the estimation of interference phenomena between coupled transmission lines. In: *2013 IEEE Grenoble Conference*. [S.l.: s.n.], 2013. p. 1–5. Cited on page 33.

ROCHA, J. F. M. *Origens e evolução das ideias da física*. Salvador: EDUFBA, 2002. 299 – 359 p. Organizador. Cited 2 times on pages 8 and 121.

SANTOS, M. S. dos. *Classificadores Neurais de Treinamento Rápido Aplicados na Identificação Online de Eventos no Detector ATLAS*. Dissertation (Master of Science) — UFBA, Salvador, 2018. Available at: <https://repositorio.ufba.br/ri/handle/ri/29254>. Cited on page 121.

SEIXAS, J. et al. Neural second-level trigger system based on calorimetry. *Computer Physics Communications*, v. 95, n. 2, p. 143 – 157, 1996. ISSN 0010-4655. Available at: <http://www.sciencedirect.com/science/article/pii/0010465596000124>. Cited on page 132.

SILVA, R. C. et al. Mutual coupling modeling in transmission lines directly in the phase domain. In: *2011 IEEE Electrical Power and Energy Conference*. [S.l.: s.n.], 2011. p. 374–379. Cited on page 33.

SIMON, A. Scintillating fibre detectors in particle physics. *Nucl. Instrum. Methods Phys. Res., A*, v. 355, p. 351–358. 32 p, May 1992. Available at: <https://cds.cern.ch/record/237723>. Cited on page 132.

SKANSI, S. *Introduction to Deep Learning. From Logical Calculus to Artificial Intelligence*. 1st. ed. [S.l.]: Springer, 2018. (Undergraduate Topics in Computer Science). ISBN 978-3-319-73003-5, 978-3-319-73004-2. Cited on page 130.

SMITH, W. W. et al. *The "RADIO" Handbook*. 17^{th}. ed. [S.l.]: Editors and Engineers, USA, 1940. Cited on page 121.

SZANDAŁA, T. Review and comparison of commonly used activation functions for deep neural networks. In: BHOI, A. K. et al. (Ed.). *Bio-inspired Neurocomputing*. 1st ed.. ed. [S.l.]: Springer Singapore;Springer, 2021. p. 203–224. ISBN 9789811554940,9789811554957. Cited on page 128.

TARTONI, N. et al. Hexagonal pad multichannel ge x-ray spectroscopy detector demonstrator: Comprehensive characterization. *IEEE Transactions on Nuclear Science*, v. 67, n. 8, p. 1952–1961, 2020. Cited on page 3.

TAYLOR, B. G. Timing distribution at the LHC. 2002. Available at: <https://cds.cern.ch/record/592719>. Cited on page 29.

THERRIEN, A. C. et al. Machine learning at the edge for ultra high rate detectors. In: *2019 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*. [S.l.: s.n.], 2019. p. 1–4. Cited on page 133.

VARELA, J. *Timing and synchronization in the LHC experiments*. Geneva, 2000. Available at: <https://cds.cern.ch/record/478248>. Cited on page 29.

WEITZEL, Q. et al. Development of structured scintillator tiles for high-granularity calorimeters. In: *2020 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*. [S.l.: s.n.], 2020. p. 1–7. Cited on page 4.

WIGMANS, R. Advances in hadron calorimetry. *Annu. Rev. Nucl. Part. Sci.*, v. 41, n. CERN-PPE-91-39, p. 133–185. 69 p, Feb 1991. Available at: <https://cds.cern.ch/record/217893>. Cited 2 times on pages 15 and 17.

WIGMANS, R. Calorimetry. In: . Texas: Scientifica Acta, 2008. v. 2, p. 18–55. Available at: <http://siba.unipv.it/fisica/ScientificaActa/Volume_2_1/Wigmans.pdf>. Cited on page 15.

WIGMANS, R. *Calorimetry: Energy Measurement in Particle Physics*. 2. ed. [S.l.]: Oxford Scholarship Online, 2017. (International Series of Monographs on Physics, 107). ISBN 9780198786351. Cited 6 times on pages 16, 17, 54, 62, 92, and 124.

YADAV, I.; SHRIMALI, H. Noise and crosstalk models of the particle detector with zero-pole transformation charge sensitive amplifier. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, v. 937, p. 107–116, 2019. ISSN 0168-9002. Available at: <https://www.sciencedirect.com/science/Article/pii/S0168900219306692>. Cited on page 3.

YADAV, I. et al. Analytical expressions for noise and crosstalk voltages of the high energy silicon particle detector. *Journal of Instrumentation*, v. 13, n. 01, p. C01019, jan 2018. Available at: <https://dx.doi.org/10.1088/1748-0221/13/01/C01019>. Cited on page 3.

ZORZETTO, R. Salto para um brilho maior. *Pesquisa FAPESP*, p. 18–25, Jul 2018. Available at: <https://revistapesquisa.fapesp.br/wp-content/uploads/2018/07/018-025_CAPA-Sirius_269-1.pdf>. Cited on page 124.

Appendix

APPENDIX A – Publications

A.1 Papers accepted for publication in journals.

1. M.V. ARAÚJO and M. BEGALLI and W.S. FREUND and G.I. CONÇALVES and M. KHANDOGA and B. LAFORGE and A. LEOPOLD and J.L. MARIN and B. S-M PERALVA and J.V.F. PINTO and M.S. SANTOS and J.M. SEIXAS and E.F. SIMAS FILHO and E.E.P. SOUZA – Lorenzetti Showers - A General-Purpose Framework for Supporting Signal Reconstruction and Triggering with Calorimeters. In: Computer Physics Communications, ISSN: 0010-4655, doi, 2023.
Keywords – High-energy physics simulation frameworks, calorimeters, event reconstruction, triggering systems, signal processing.

   - **Abstract**

Calorimeters play an important role in high-energy physics experiments. Their design includes electronic instrumentation, signal processing chain, computing infrastructure, and also a good understanding of their response to particle showers produced by the interaction of incoming particles. This is usually supported by full simulation frameworks developed for specific experiments so that their access is restricted to the collaboration members only. Such restrictions limit the general-purpose developments that aim to propose innovative approaches to signal processing, which may include machine learning and advanced stochastic signal processing models. This work presents the Lorenzetti Showers, a general-purpose framework that mainly targets supporting novel signal reconstruction and triggering strategies using segmented calorimeter information. This framework fully incorporates developments down to the signal processing chain level (signal shaping, energy estimation, and noise mitigation techniques) to allow advanced signal processing approaches in modern calorimetry and triggering systems. The developed framework is flexible enough to be extended in different directions. For instance, it can become a tool for the phenomenology community to go beyond the usual detector design and physics process generation approaches.

A.2 Papers accepted for presentation in Congress and Annals.

1. SANTOS, M. S.; LAFORGE, B.; SIMAS FILHO, E. F.; FARIAS, P. C. A. M; SEIXAS, J. M. – Redes Neurais Artificiais como algoritmo de reconstrução da informação depositada no Calorímetro Eletromagnético do Experimento ATLAS reduzindo os efeitos de *Crosstalk*. In: XVI Brazilian Congress on Computational Intelligence, doi, 08 to 11 of october 2023, Brazil: [s.n.], 2023.

   - **Abstract**

ATLAS is the largest experiment at the LHC accelerator complex in CERN. It is situated at one of the collision points in the accelerator tunnel. The experiment consists of a collection of specialized sub-detectors designed to characterize the particles generated by collisions in the LHC. One of these specialized detectors is the liquid argon (LAr) calorimeter, which contains approximately 187k sensor cells used to characterize electromagnetic showers. The LAr calorimeter has a high cell density, which, combined with the high collision rates and the mechanical and electronic structure of the detector readout, leads to crosstalk (XT) effects between adjacent sensor cells. Crosstalk degrades the accuracy of energy and time reconstruction for incoming particles. To address this challenge, an electromagnetic shower simulator based on the ATLAS LAr calorimeter was developed, along with a machine learning

approach to mitigate the XT effects. The results demonstrate that the energy and time-of-flight of particles can be reconstructed very closely to the desired values. The proposed approach reduces the error fluctuations in energy estimation by at least 37 times compared to the standard algorithm. Additionally, it decreases the error fluctuations in time estimation by three orders of magnitude.

2. SANTOS, M. S.; LAFORGE, B.; SIMAS FILHO, E. F.; FARIAS, P. C. A. M; SEIXAS, J. M. – Redes Neurais como um método de redução do *Crosstalk* entre sensores no calorímetro eletromagnético do experimento ATLAS . In: XV Brazilian Congress on Computational Intelligence, doi, 03 to 06 of october 2021 (virtual format), Brazil: [s.n.], 2021.

   - **Resumo**

   O ATLAS é o maior experimento do complexo do acelerador de partículas LHC no CERN. Está localizado em um dos pontos de colisão no túnel do acelerador e é composto de um conjunto de sub-detectores especializados para caracterizar as partículas produzidas nas colisões do LHC. Um desses detectores especializados é o calorímetro de Argônio Líquido (LAr), com cerca de 187k sensores para rastreamento e registro de chuveiros eletromagnéticos. O LAr tem uma granularidade fina e alta densidade de células que em associação com as altas taxas de colisão, estrutura mecânica e eletrônica e o sistema de leitura do detector produz efeitos de *crosstalk* (XT) entre células dificultando o processo de reconstrução da energia e do tempo de de voo da partícula incidente. Para tratar este desafio, foi um desenvolvido um simulador de chuveiro eletromagnético de partículas baseado no calorímetro LAr do ATLAS, junto com uma abordagem de aprendizagem de máquina para mitigar os efeitos do XT. Os resultados indicam que para altas energias a reconstrução se aproxima dos valores alvo e, por outro lado, em baixas energias é necessário um método especializado para tratar e normalizar os dados antes de aplicar uma rede neural para reconstruir os valores das energias.
   Palavras-chave – ATLAS, Calorímetro Eletromagnético, Argônio Líquido, *Crosstalk*, Redes Neurais, Simulador, Regressão.

   - **Abstract**

   ATLAS is the biggest experiment of the LHC accelerator complex at CERN. It is located at one of the collision points in the accelerator tunnel and is composed of a set of specialized sub-detectors to characterize the particles produced by LHC's collisions. One of those specialized detectors is the liquid argon (LAr) calorimeter, with around 187k sensors cells to characterize electromagnetic showers. The LAr has a high cell density, which in association with the high collision rates and the mechanical and electronic structure of the detector readout, produces crosstalk (XT) effects between sensor-cells. XT degrades the reconstruction of energy and time of incident particles.

   The solution proposed is the development of an electromagnetic shower emulator based on ATLAS LAr, together with a machine learning approach to mitigate the XT effects. The results indicate that the reconstruction is very close to the target at high energies. On the other hand, low energies need a specialized method to treat and normalize data before applying a neural network to reconstruct energies.
   Keywords – ATLAS, Electromagnetic Calorimeter, Liquid Argon (LAr), Crosstalk, Neural Networks, Simulator, Regression.

A.3 Abstracts presented in Physics Meetings.

1. E. F. SIMAS FILHO, A. LEOPOLD, B. S. PERALVA, B. LAFORGE, E. E. P. De SOUZA, G. GONÇALVES, J. V. F. PINTO, J. SEIXAS, J. L. MARIN, M. BEGALLI, M. S. SANTOS, M. V. ARAUJO, P. N. GASPAR, W. S. FREUND. Simulação e Processamento de Sinais para Futuros Desenvolvimentos em Calorimetria de Altas Energias. In: *XLII Reunião de Trabalho de Física Nuclear no Brasil - XL Encontro*

Nacional de Física de Partículas e Campos - RENAFAE - Workshop 2019 (SBF - Sociedade Brasileira de Física 2022) (SBF 2022). 25 a 28 de abril, formato remoto, Brazil: [s.n.], 2022.

- **Resumo**

Os calorímetros desempenham um papel importante na seleção (trigger) online e na análise offline em experimentos modernos e também devem trazer informações importantes para experimentos futuros, como no High-Luminosity Large Hadron Collider (HL-LHC) e no Future proton-proton Circular Collider (FCC-pp). Esses novos calorímetros enfrentarão condições extremamente rigorosas, pois espera-se energia do centro de massa de até 100 TeV, com uma luminosidade integrada também muito alta, gerando um enorme número de colisões simultâneas (podendo atingir uma média de até 1.000 colisões de partículas por cruzamento de feixes). Para projetar calorímetros de última geração, sua instrumentação eletrônica, cadeia de processamento de sinais e infraestrutura de computação, ter um ambiente de simulação aberto que permita mapear conceitos de uso geral em calorimetria e avaliar abordagens inovadoras para o processamento de sinais pode trazer benefícios para toda a cadeia de instrumentação envolvida. Este trabalho descreve diferentes contribuições para a simulação e processamento de sinais de calorímetros visando operação em alta luminosidade. A simulação completa do calorímetro geralmente requer módulos de processamento específicos, que podem ser configurados de diferentes maneiras e conectados à cadeia principal de simulação. O objetivo do simulador desenvolvido foi possibilitar a geração de efeitos como ruído eletrônico, empilhamento de sinais, erros de calibração, deformação do pulso resultante do condicionamento do sinal e erros na estimação de sinais devido ao crosstalk entre os inúmeros canais envolvidos, notadamente nas seções eletromagnéticas. De modo a contribuir para o programa de atualização do experimento ATLAS no seu sistema de calorimetria, propostas quanto à mitigação do efeito de crosstalk no calorímetro de Argônio Líquido (eletromagnético) e melhoria de granularidade em células hadrônicas foram desenvolvidas. Em ambas proposições, o uso de inteligência computacional e, em particular, do aprendizado profundo é analisado com dados simulados e em testes de calibração.

2. SANTOS, M. S.; LAFORGE, B. L. E.; SIMAS FILHO, F.; FARIAS, P. C. A. M; FREUND, W. S; SEIXAS, J. M. Redução do *Cross-Talk* no Calorímetro do Experimento ATLAS. In: *XLII Reunião de Trabalho de Física Nuclear no Brasil - XL Encontro Nacional de Física de Partículas e Campos - RENAFAE - Workshop 2019 (SBF - Sociedade Brasileira de Física 2019) (SBF 2019)*. 01 a 05 de setembro, Campos do Jordão, Brazil: [s.n.], 2019.

   - **Resumo**

O LHC apresenta um conjunto de sete experimentos distribuídos ao longo de seu perímetro de aproximadamente 27 Km, sendo o ATLAS o maior deles. A fim de registrar os eventos ocorridos durante as colisões, o ATLAS foi projetado com elevada segmenta¸cão (mais de 100 milhões de canais). Dentre os subdetectores constituintes do ATLAS estão os calorímetros eletromagnético e hadrônico, com quase 190 mil canais. Tal estrutura finamente segmentada dos calorímetros possibilita o aparecimento de cross-talk (cruzamento de informacão) entre os canais vizinhos, que produz um sinal espúrio e que passa a ser somado ao verdadeiro nível de energia. Como resultado, o nível de energia que é reconstruído pode estar incorreto, mascarando a natureza da partícula que foi registrada. Neste trabalho, propõe-se um método de processamento e tratamento da informacão para reduzir o efeito do *crosstalk* no calorímetro do ATLAS. Para isso, será utilizada uma janela de varredura sobre a região do calorímetro que foi sensibilizada por um candidato a evento de interesse. A partir do conhecimento das propriedades físicas que produzem a interacão entre os sensores adjacentes, que pode envolver fenômenos capacitivos, resistivos e indutivos, é proposto um algoritmo para recuperar a informacão verdadeira (sem influência do cross-talk) na posição central da janela de análise. Resultados com dados simulados indicam a eficiência do método proposto.

   - **Abstract**

The LHC presents a set of seven experiments distributed along its perimeter of approximately 27 km. ATLAS is the largest of them. The ATLAS was designed with high segmentation (more than 100 million channels) to record the events occurring during collisions. Among the constituent sub-detectors of ATLAS are the electromagnetic and hadronic calorimeters, with almost 190k channels. This finely segmented structure of the calorimeters enables crosstalk between adjacent channels, producing a spurious signal added to the true energy level. As a result, the reconstructed energy level may be incorrect, masking the recorded particle's nature. In this work, an information processing and handling method are proposed to reduce crosstalk's effect in the ATLAS calorimeter. To this end, a moving window will be used over the calorimeter region that a candidate event of interest has sensitized. Knowing the physical properties that produce the interaction between adjacent sensors, which may involve capacitive, resistive, and inductive phenomena, an algorithm is proposed to retrieve the true information (without the influence of crosstalk) at the central position of the analysis window. Results with simulated data indicate the efficiency of the proposed method.

3. KHANDOGA, Mykola; SANTOS, Marton; LAFORGE, L. E. B.; SIMAS FILHO, Eduardo F.; FARIAS, P. C. A. M; Crosstalk mitigation in the EM calorimeter using spatio-temporal neural networks. In: 5^{th} ATLAS Machine Learning Workshop - 13 – April 15, 2021.

   - **Abstract**

   Crosstalk is the effect of unwanted signal transmission between electronic circuits. In the context of the electromagnetic calorimeter, the crosstalk between the readout channels leads to distorted shower shapes and degrades the precision of energy and time reconstruction. The reconstruction of energy and time in the ATLAS calorimeter is performed based on a time series of signal samples from each calorimeter cluster cell. These signal samples, besides the real signal, also contain crosstalk contributions from neighboring cells. The crosstalk is supposed to have a stable pattern that can be factored out and corrected for using the spatio-temporal neural networks, which would take into account both the signal shape and the relative location of the crosstalking cells. In this ongoing study, some network architectures were tested and compared, demonstrating the effect on the reconstruction of time and energy.

4. SANTOS, M. S.; LAFORGE, B.; SIMAS FILHO, E. F.; FARIAS, P. C. A. M; SEIXAS, J. M. – Redes Neurais para redução do *Crosstalk* nas células do Calorímetro de Argônio Líquido do Experimento ATLAS. In: WORKSHOP DA RENAFAE 2021 (SBF - Sociedade Brasileira de Física) - PROJETOS PARA O FUTURO DA FÍSICA DE ALTAS ENERGIAS NO BRASI, 12 a 14 de julho 2021, Brazil: [s.n.], 2021.

   - **Resumo**

   O LHC é o maior acelerador de partículas em funcionamento e é constituído de um conjunto de sete experimentos distribuídos ao longo de seu perímetro de aproximadamente 27 Km, sendo o ATLAS o maior deles. A fim de registrar os eventos ocorridos durante as colisões, o ATLAS foi projetado com elevada segmentação de mais de 100 milhões de canais. Dentre os subdetectores constituintes do ATLAS estão os calorímetros eletromagnético e hadrônico, com quase 190 mil canais. Tal estrutura, com alta densidade de canais sensores dos calorímetros possibilita o aparecimento de *crosstalk* (XT) (cruzamento de informação) entre os canais vizinhos, que produz um sinal espúrio e que passa a ser somado ao verdadeiro nível de energia. Como resultado, o nível de energia que é reconstruído pode estar incorreto, mascarando a natureza da partícula que foi registrada e prejudicando a determinação do tempo de voo da partícula vinda da colisão. O sistema de amostragem do ATLAS utiliza quatro amostras do sinal de calibração, tomadas a cada 25ns, as quais são processadas pelo algoritmo de filtro ótimo que determina o valor da energia depositada e o tempo da partícula incidente. Porém, o filtro ótimo não consegue mitigar o efeito do XT presente nas amostras do sinal reconstruindo. Neste trabalho, foi desenvolvido um simulador de chuveiro de partículas eletromagnéticas (elétrons, pósitrons e fótons) baseado nos perfis

de deposição de energia e em parâmetros físicos do calorímetro eletromagnético, em Argônio Líquido (LAr) do ATLAS. O objetivo do desenvolvimento do simulador é a produção de eventos com diferentes condições de XT e permitir uma rápida reconfiguração e execução, sem a necessidade de utilização de uma grande infraestrutura, e de elevado conhecimento especialista do detector, como ocorre nos modelos do GEANT4. Utilizando como referência um cluster definido na segunda camada (S2) do calorímetro eletromagnético de tamanho (5×0,025) × (5×$\frac{\pi}{128}$) (η, ϕ), 745k eventos foram produzidos. Como método de processamento de sinais adotado, redes neurais, do tipo *Multilayer Perceptron*, contendo uma, duas e três camadas escondidas foram avaliadas no processo de reconstrução da energia sem a interferência do sinal XT. Como entrada, 9 células são selecionadas, a mais energética e sua primeira vizinhanças, cada uma com 4 amostras.O critério de avaliação utilizado é o menor erro médio quadrático (MSE) avaliado entre a saída da rede e os valores de energia utilizados como referência para o treinamento, obtendo um MSE, global, igual a 9,23×10^{-4}. Os ensaios têm indicado que as redes neurais do tipo MLP podem reconstruir a energia da célula de interesse, reduzindo a interferência causada pelo sinal XT.

- **Abstract**

The LHC is the largest particle accelerator in operation and consists of seven experiments distributed along its perimeter of approximately 27 km, ATLAS being the largest. To record the events that occur during collisions, ATLAS was designed with high segmentation of more than 100 million channels. Among the sub-detectors constituting ATLAS are the electromagnetic and hadronic calorimeters, with almost 190k channels. Such a structure, with a high density of sensor channels of the calorimeters, enables the appearance of (XT) (crosstalk) between neighboring channels, which produces a spurious signal that is added to the true energy level. As a result, the energy level that is reconstructed may be incorrect, masking the nature of the particle that was recorded and impairing the determination of the time of flight of the particle coming from the collision. The ATLAS sampling system uses four samples of the calibration signal, taken every 25 ns, which are processed by the optimal filter algorithm that determines the value of the deposited energy and the time of the incident particle. However, the optimal filter fails to mitigate the effect of the XT present in the reconstructing signal samples. In this work, an electromagnetic particle shower simulator (electrons, positrons, and photons) based on the energy deposition profiles, and physical parameters of the electromagnetic calorimeter, in Liquid Argon (LAr) of ATLAS was developed. The goal of the simulator development is to produce events with different XT conditions and allow for rapid reconfiguration and execution without the need for large infrastructure, and high expert knowledge of the detector, as occurs in GEANT4 models. Using as reference a cluster defined in the second layer (S2) of the electromagnetic calorimeter of size (5×0.025) × (5×$\frac{\pi}{128}$) (η, ϕ), 745k events were produced. As the adopted signal processing method, neural networks of the Multilayer Perceptron type containing one, two, and three hidden layers were evaluated in the energy reconstruction process without the interference of the XT signal. As input, nine cells are selected, the most energetic and its first neighborhood, each with four samples. The evaluation criterion used is the smallest mean squared error (MSE) evaluated between the network output and the energy values used as a reference for training, obtaining an MSE, globally, equal to 9.23×10^{-4}. Tests have indicated that MLP-type neural networks can reconstruct the energy of the cell of interest, reducing the interference caused by the XT signal.

5. SANTOS, M. S.; LAFORGE, B.; SIMAS FILHO, E. F.; FARIAS, P. C. A. M; SEIXAS, J. M. – Aprendizado de Máquina aplicado à redução de interferência entre canais sensores do Calorímetro Eletromagnético de Argônio Líquido do Experimento ATLAS. In: XLI Encontro Nacional de Física de Partículas e Campos (ENFPC), 27 to 29 of setember 2021, Brazil: [s.n.], 2021.

- **Resumo**

O acelerador de partículas de maior luminosidade em operação é o LHC, e está localizado no CERN. Dentre os seus experimentos, o ATLAS é o maior, com cerca de 100 milhões de canais de leitura. É formado por um conjunto de subdetectores especializados para caracterizar as partículas, produzidas nas colisões dos feixes prótons. Um desses subdetectores, é o calorímetro eletromagnético de Argônio Líquido (LAr), com cerca de 187 mil canais de leitura. Sua fina granularidade, a energia das partículas incidentes, a estrutura mecânica e o sistema de leitura dos canais possibilitam o surgimento do acoplamento (*crosstalk* - XT) entre canais vizinhos, que insere erros no processo de reconstrução da energia e do tempo de voo da partícula que interage com um canal do calorímetro. Tal acoplamento pode ser de natureza resistiva, indutiva e capacitiva. Para mitigar o XT, um simulador simplificado de chuveiros eletromagnético de partículas foi desenvolvido, com o objetivo de permitir o rápido ensaio de configurações específicas de operação deste calorímetro do ATLAS, considerando diferentes níveis de ruído e condições de acoplamento entre os canais. Desta maneira, o estudo pode se desenvolver sem a necessidade da utilização de uma grande infraestrutura e de sofisticado conhecimento especialista da estrutura do experimento ATLAS e do respectivo processo físico, como, por exemplo, ocorre ao se utilizar o GEANT4. Este simulador foi desenvolvido considerando-se os modelos de perfil de deposição, radial e longitudinal, de energia das partículas eletromagnéticas, parâmetros físicos dos canais do LAr e o modelo de acoplamento do tipo capacitivo entre canais de leitura, vizinhos, já conhecidos e validados na literatura. Neste trabalho, foram adotados os seguintes parâmetros de simulação: janela padrão ao redor da célula mais quente, na segunda do LAr, de tamanho (5×0,025) × (5×$\frac{\pi}{128}$) (η, ϕ), totalizando 25 canais. Considerou-se que o acoplamento, além da primeira vizinhança da célula de interesse, não é significativo e, no momento, o acoplamento resistivo não está sendo considerado. Dessa forma, 745 mil eventos foram produzidos. Adotou-se, como método de processamento de sinais, redes neurais, do tipo *Multilayer Perceptron* (MLP), contendo uma, duas e três camadas escondidas. Os resultados indicam que as redes MLP conseguem reconstruir o valor da energia da célula de interesse sem influência do XT, obtendo um MSE global da ordem de 10^{-4}.

A.4 Papers under preparation to submit.

1. M.S. SANTOS and E.F. SIMAS FILHO and P.C.M.A FARIAS and B. LAFORGE and J.M. SEIXAS– Lorenzetti Showers - Crosstalk Mitigation in an ATLAS-like High Energy Liquid Argon calorimeter using artificial neural networks. In: XXX.
   Keywords – Crosstalk, Calorimeter, Machine Learning, Lorenzetti Framework..

   - **Abstract**

   High-energy particle accelerators and their associated experiments contribute significantly to our understanding of the nature of matter. In this context, calorimeters are key components employed to measure the energy of the outgoing particles produced in the collisions. For this, a high-density structure of sensors organized around the beam axis and comprising several superimposed layers is used. Calorimeter information is essential for particle characterization and triggering. A high-energy sampling calorimeter is split into electromagnetic and hadronic sections and constructed using an alternate structure of sensitive and absorber materials. One possible configuration for electromagnetic calorimeters is to use liquid Argon (LAr) and lead (Pb) for sensitive and absorber materials, respectively. The high-density granular structure of sensor cells combined with high energy values of the particles to collide can produce undesired effects that distort the signal of interest in each cell. One of these effects is crosstalk (XT), a spurious signal from neighboring sensor cells. Consequently, particle characterization is affected, introducing errors in energy and particle time-of-flight estimation. This work proposes a machine-learning method for crosstalk mitigation in an ATLAS-like LAr+Pb electromagnetic calorimeter. For this, a study based on a framework to simulate particle interaction on an LAr calorimeter that accounts crosstalk effect as observed in the

ATLAS experiment was proposed. The results show that a supervised learning neural network model for time regression reduces the estimation error by 31 times for cells with the highest energy and 182 times for cells with low energy. For energy estimation, the neural estimator reduces the error by 1,09 times at high-energy cells and 12 times at low-energy cells. The results also indicate that the proposed model is capable of reducing the crosstalk distortion in shower-shape variables, which are traditionally used for particle identification in calorimetry.

APPENDIX B – Profiles for Energy Deposition - Parameterization

In the following, the parametrization for the distribution functions for the radial and longitudinal profiles of the electromagnetic particle shower for a homogeneous medium and a sampled medium will be presented. Grindhammer, Rudowicz and Peters (1990), Grindhammer and Peters (1993) provides many parameters to describe the electromagnetic particle cascade, like the profiles definition and GEANT44 constants to develop the analytical simulator.

B.1 Homogeneous Media

A homogeneous medium refers to detectors in which a single material absorbs and detects the particles during the interaction process.

B.1.1 Longitudinal Profile

The likelihood function describing the longitudinal profile is shown in the Equation B.1. Being α[54] the shape parameter and β the scale parameter; t is the depth of the profile in units of X_0; $\langle t \rangle$ indicates the center of gravity of the profile; T, the point of maximum. In Equation B.2 the gamma function is displayed.

$$\left\langle \frac{1}{E}\frac{dE(t)}{dt} \right\rangle = f(t) = \frac{(\beta t)^{\alpha-1}\beta \exp(-\beta t)}{\Gamma(\alpha)}. \tag{B.1}$$

$$\Gamma(\alpha) = \int_0^\infty \mathrm{e}^{-x} x^{\alpha-1} \mathrm{d}x. \tag{B.2}$$

$$\langle t \rangle = \frac{\alpha}{\beta}. \tag{B.3}$$

$$T = \frac{\alpha - 1}{\beta}. \tag{B.4}$$

$$T \propto \ln y = \ln \frac{E}{E_c}. \tag{B.5}$$

Following are the fits for the T and α parameters in the Equation B.6 and Equation B.7 respectively.

$$T_{hom} = \ln y - 0,858 \tag{B.6}$$

$$\alpha_{hom} = 0,21 + (0,492 + 2,38/Z) \ln y \tag{B.7}$$

B.1.2 Radial Profile

For the radial profile, two functions are used, one to describe the core ($f_C(r)$) of the profile and another for the tail ($f_T(r)$), weighted by a weight p, as shown in the Equation B.8 and r is the distance to the center of the shower in units of R_M.

[54] In the developed simulator uses $\alpha = 4.8$ and $\beta = 0.275$.

$$\begin{aligned} f(r) &= p f_C(r) + (1-p) f_T(r) \\ &= p\,\frac{2rR_C^2}{\left(r^2+R_C^2\right)^2} + (1-p)\frac{2rR_T^2}{\left(r^2+R_T^2\right)^2} \end{aligned} \tag{B.8}$$

with,

$$0 \leq p \leq 1.$$

The following are the functions used for the description of the core, Equation B.9, the tail, Equation B.10 and the weight, see Equation B.11 of the radial profile in a given material. All other constants and variables are displayed in the subsection B.2.2.

$$R_{C_{hom}}(\tau) = z_1 + z_2\,\tau \tag{B.9}$$

$$R_{T_{hom}}(\tau) = k_1\{\exp(k_3(\tau-k_2)) + \exp(k_4(\tau-k_2))\} \tag{B.10}$$

$$p_{hom}(\tau) = p_1 \exp\left\{\frac{p_2-\tau}{p_3} - \exp\left(\frac{p_2-\tau}{p_3}\right)\right\} \tag{B.11}$$

B.2 Non-homogeneous Media

It refers to detectors in which the structure is composed of layers of different materials, with one of the materials as the passive medium responsible for absorbing energy. The other material, the active one, detects the particles' energy and position.

The description of the profiles in this medium is done by adjusting the data for a homogeneous medium. The sampling frequency (F_S) parameter is defined as a function of X_0 and the thickness of the active (a) and passive (p) media, see the Equation B.12.

$$F_S = \frac{X_{0_{eff}}}{d_a + d_p} \tag{B.12}$$

Another relevant information is the average over the shower depth (e/mip[55].), information that can be roughly calculated by the Equation B.13 using the atomic number information of the constituent materials of the non-homogeneous medium.

$$\hat{e} = \frac{1}{1 + 0.007\,(Z_p - Z_a)} \approx \frac{e}{mip}. \tag{B.13}$$

In this way, it is possible to determine data equivalent to that obtained in the homogeneous medium presented in Equation B.14 and Equation B.15.

$$T_{\text{sam}} = T_{\text{hom}} + t_1 F_S^{-1} + t_2(1-\hat{e}) \tag{B.14}$$

$$\alpha_{\text{sam}} = \alpha_{\text{hom}} + a_1 F_S^{-1} \tag{B.15}$$

[55] energy independent characteristic parameter, which allows identifying the type of calorimeter. If equal to 1, it is a homogeneous calorimeter. If different non-homogeneous (WIGMANS, 2017)

B.2.1 Radial Profile

On the sequence, the respective Equation B.16, Equation B.16, and Equation B.18 for the sampled medium came from the adjustments used in a homogeneous medium in the sequence. The subscript *hom* refers to the homogeneous medium and *sam*[56] for the non-homogeneous medium.

$$R_{C_{sam}} = R_{C_{hom}} - 0.0203(1-\hat{e}) + 0.0397F_S^{-1}\exp(-\tau) \tag{B.16}$$

$$R_{T_{sam}} = R_{T_{hom}} - 0.14(1-\hat{e}) - 0.495F_S^{-1}\exp(-\tau) \tag{B.17}$$

$$p_{sam} = p_{hom} + (1-\hat{e})\left(0.348 - 0.642F_S^{-1}\exp\left(-(\tau-1)^2\right)\right) \tag{B.18}$$

The following equations refer to the effective values calculated for a material composition of a non-homogeneous medium.

$$w_i = \frac{\rho_i d_i}{\sum_j \rho_j d_j}\ (\rho = \text{ density }) \tag{B.19}$$

$$Z_{eff} = \sum_i w_i Z_i \tag{B.20}$$

$$A_{eff} = \sum_i w_i A_i \tag{B.21}$$

$$\frac{1}{X_{0,eff}} = \sum_i \frac{w_i}{X_{0,i}} \tag{B.22}$$

$$\frac{1}{R_{M_{eff}}} = \frac{1}{E_s}\sum_i \frac{w_i E_{c,i}}{X_{0,i}}\ (E_s = 21.2\text{MeV}) \tag{B.23}$$

$$E_{c,eff} = X_{0,eff}\sum_i \frac{w_i E_{c,i}}{X_{0,i}} \tag{B.24}$$

B.2.2 Variables and Constants

The constants and variables are calculable using the active medium and simulated collision data. In the Table 15 information regarding the constituent materials of the sampling calorimeter of the ATLAS experiment is displayed.

[56] sam - refers to the sampling medium.

$$t_1 = -0,59. \tag{B.25}$$

$$t_2 = -0,53. \tag{B.26}$$

$$z_1 = 0,0251 + 0.00319 * \ln E_0. \tag{B.27}$$

$$z_2 = 0,1162 - 0.000381 * Z_{LAr}. \tag{B.28}$$

$$k_1 = 0,659 - 0.00309 * Z_{LAr}. \tag{B.29}$$

$$k_2 = 0,645. \tag{B.30}$$

$$k_3 = -2,59. \tag{B.31}$$

$$k_4 = 0,3585 + 0.0421 * \ln E_0. \tag{B.32}$$

$$p_1 = 2,632 - 0,00094 * Z_{LAr}. \tag{B.33}$$

$$p_2 = 0,401 + 0,00187 * Z_{LAr}. \tag{B.34}$$

$$p_3 = 1,313 - 0,0686 * \ln E_0. \tag{B.35}$$

Table 15 – Parameters for the active medium (LAr) and passive medium (Pb) of the ATLAS sampling calorimeter.

LAr			Pb			Unit
ρ_{LAr}	=	1,396	ρ_{Pb}	=	11,350	g/cm^3
Z_{LAr}	=	18,000	Z_{Pb}	=	82,000	–
w_{LAr}	=	0,360	w_{Pb}	=	1 - w_{LAr}	–
$X0_{LAr}$	=	140,000	$X_{0_{Pb}}$	=	5,612	mm
Ec_{LAr}	=	32,840	$E_{c_{Pb}}$	=	7,430	MeV
$R_{M_{LAr}}$	=	90,430	$R_{M_{Pb}}$	=	16,020	mm
$d_a^\star$	=	4,000	$d_p^\star$	=	2,000	mm

* Layer thickness.

Source: Particle Data Group et al. (2020) and ATLAS Collaboration (1999).

B.3 XT transfer Function

In this Ph.D. thesis, the XT transfer function for capacitive coupling is presented as follows in Equation B.36. The inductive coupling was considered as the same transfer function as capacitive coupling and adjusted the magnitude according to previous works using calibration runs (COLAS et al., 1999; LABBé; ISHMUKHAMETOV, 2008; LOPEZ SOLIS, 2017).

$$XT(t) = \Bigg[2\, e^{\frac{t}{\tau_d}}\, \tau_d^2 \Big(t\,(\tau_d - \tau_{pa})(\tau_{pa} + t_d) + \tau_{pa}\,(\,3\tau_d\,\tau_{pa} + (\,2\,\tau_d + \tau_{pa}\,)\,t_d) \Big) - e^{\frac{t}{\tau_{pa}}} \Bigg(t^2\,(\tau_d - \tau_{pa})^2 *$$

$$*(\tau_d + t_d) - 2\,t\,\tau_d\,(\tau_d - \tau_{pa}) \big(2\tau_d\,\tau_{pa} + (\tau_d + \tau_{pa})\,t_d \big) + 2\,\tau_d^2\,\tau_{pa} \Big(3\tau_d\tau_{pa} + (\,2\,\tau_d + \tau_{pa}\,)\,t_d \Big) \Bigg) +$$

$$+\tau_d \Bigg(-2\, e^{\frac{t}{\tau_d} + \frac{t_d}{\tau_{pa}}}\, \tau_d\, \tau_{pa} \Big(t\,(\tau_d - \tau_{pa}) + 3\,\tau_d\,\tau_{pa} + (-\tau_d + \tau_{pa}\,)\,t_d \Big) + e^{\frac{t}{\tau_{pa}} + \frac{t_d}{\tau_d}} \Bigg(t^2\,(\,\tau_d - \tau_{pa})^2 +$$

$$+6\,\tau_d^2\,\tau_{pa}^2 + 4\,\tau_d\,(\tau_d - \tau_{pa})\,\tau_{pa}\,t_d + (\tau_d - \tau_{pa})^2\,t_d^2 - 2\,t\,(\tau_d - \tau_{pa}) \big(2\,\tau_d\,\tau_{pa} + (\,\tau_d - \tau_{pa}\,)\,t_d \big) \Bigg) \Bigg) \Bigg] *$$

$$* \frac{C_x R_f R_{in}}{2\, e^{t\left(\frac{1}{\tau_d} + \frac{1}{\tau_{pa}}\right)}\, \tau_d\,(\,\tau_d - \tau_{pa})^4\, t_d}\, \delta(t - t_d) \tag{B.36}$$

where:

- C_x - Coupling capacitance between channels in pF;
- R_f - Models the transresistance of a ideal current preamplifier;
- R_{in} - Resistance of input channel in Ω;
- τ_d - Parameter associated with detector;
- t_d - Parameter associated with detector;
- τ_{pa} - Parameter associated with preamplifier.

Table 16 shows the values used to model the XT coupling between LAr calorimeter cells at the ATLAS.

Table 16 – Values used for XT modeling in this thesis.

C_x	47.000
R_f	0.078
R_{in}	1.200
τ_d	15.820
t_d	420.000
$\tau_p a$	17.310

APPENDIX C – EMshower Modeling

For modeling the electromagnetic calorimeter of the ATLAS experiment two steps were necessary: first, to obtain the information regarding the validated physical phenomenon: Energy deposition profiles for the electromagnetic particle shower, and second, to use information regarding the ATLAS experiment, which is listed below:

- Mathematical model of the signal, g(t), from a LAr cell;
- Mathematical model of the XT(t) in a cell of the LAr;
- X_0 depth of the LAr layers;
- Radius referring to the thickness of each of the LAr layers;
- Information regarding the LAr active and passive media materials.

Next in the Figure 72 are the details of the modeling of the LAr experiment for the simulator. The XYZ coordinate system starts at the collision point, that is, at the center of the ATLAS experiment. From this point, it is defined as:

- The particle propagation axis, *t*, which indicates the penetration depth at X_0;
- The angle θ_A, formed between the propagation axis of the particle and the *z* axis;
- The radius, r_A, which describes the distance from the collision point to the point of maximum intensity of the electromagnetic shower;
- The projection of the radius r_A onto the XOY plane, ϱ;
- The φ angle;
- The ray, r_M, auxiliary, which defines the propagation limit, radial, of the shower;
- R_0 projection of the distance from the *z* axis to the S_1 layer of the LAr;
- R_A projection of the distance from the *z* axis to point A;
  - $A : \{r_A, sen(\theta_A)\cos(\phi_A),\ r_A, sen(\theta_A)sen(\phi_A),\ r_A, cos(\theta_A)\}$;

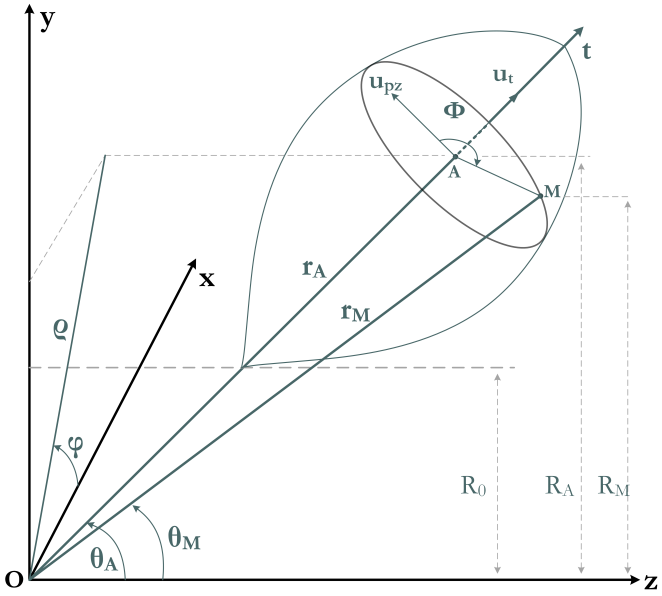
  R_M projection of the distance from the *z* axis to point M;
  - $M : \{r_M, sen(\theta_M)\cos(\phi_M),\ r_M, sen(\theta_M)sen(\phi_M),\ r_M, cos(\theta_M)\}$;
- The radius of the shower, $\|\overrightarrow{AM}\|$
- The angle of rotation, Φ, of the shower radius $\overrightarrow{AM}$;
- The unit vectors U_t, U_{pz}, referring to the direction of propagation and the orthogonal projection of the *z* axis concerning the *t* axis in the plane containing the points *A* and *M*.

These parameters are necessary for the three-dimensional description of the particle shower. The objective is to describe the variables of the mathematical model, in the physical variables of the energy deposition profile, and apply them to the taking into account the constructive characteristics of the ATLAS experiment, i.e., $f(\|\overrightarrow{AM}\|, t, \Phi) \rightarrow f(R,\ \eta,\ \phi)$ and apply the changes of variables to the longitudinal, Equation C.1, and radial, Equation C.2, respectively. There is a dependence on the incident energy, E_0, for depth, penetration into the material, and radial scattering.

$$\left\langle \frac{1}{E}\frac{dE(t)}{dt} \right\rangle = f(t) = \frac{(\beta t)^{\alpha-1}\beta \exp(-\beta t)}{\Gamma(\alpha)}. \tag{C.1}$$

$$f(r) = p(t)\,\frac{2rR_C(t)^2}{\left(r^2+R_C(t)^2\right)^2} + (1-p(t))\frac{2rR_T(t)^2}{\left(r^2+R_T(t)^2\right)^2}. \tag{C.2}$$

Figure 72 – Diagram of the electromagnetic shower modeling for the coordinates of the ATLAS experiment.



The other auxiliary terms used in the model are presented and defined by Equation C.3, Equation C.4, Equation C.5, and Equation C.6.

$$\overrightarrow{OM}(\mathrm{r_M}, \eta_\mathrm{M}, \phi_\mathrm{M}) = \{\mathrm{r_M}\sin(\theta_\mathrm{M}\cos(\phi_\mathrm{M}),\ \mathrm{r_M}\sin(\theta_\mathrm{M}\sin(\phi_\mathrm{M}), \mathrm{r_M}\cos(\theta_\mathrm{M})\}. \tag{C.3}$$

$$\overrightarrow{OA}(\mathrm{r_A}, \eta_\mathrm{A}, \phi_\mathrm{A}) = \{\mathrm{r_A}\sin(\theta_\mathrm{A}\cos(\phi_\mathrm{A}),\ \mathrm{r_A}\sin(\theta_\mathrm{A}\sin(\phi_\mathrm{A}), \mathrm{r_A}\cos(\theta_\mathrm{A})\}. \tag{C.4}$$

$$\overrightarrow{AO}(\mathrm{r_A}, \eta_\mathrm{A}, \phi_\mathrm{A}) = -\{\mathrm{r_A}\sin(\theta_\mathrm{A}\cos(\phi_\mathrm{A}),\ \mathrm{r_A}\sin(\theta_\mathrm{A}\sin(\phi_\mathrm{A}),\ \mathrm{r_A}\cos(\theta_\mathrm{A})\}. \tag{C.5}$$

$$\theta = 2\tan^{-1}\mathrm{e}^{-\eta}. \tag{C.6}$$

Each one of the variables, $f(\|\overrightarrow{AM}\|, t, \Phi)$ is described, bellow, by the variables R, η, ϕ:

$$\|\overrightarrow{AM}(\mathrm{R},\ \eta,\ \phi)\| = \mathrm{r} = \|\overrightarrow{AO}(\mathrm{R_A},\ \eta_\mathrm{A},\ \phi_\mathrm{A}) + \overrightarrow{OM}(\mathrm{R_M},\ \eta_\mathrm{M},\ \phi_\mathrm{M})\| \tag{C.7}$$

$$\mathrm{t}(\mathrm{R},\ \eta,\ \phi) = \frac{|\mathrm{R}\left(\cos(\phi_\mathrm{A}-\phi)\mathrm{sech}(\eta_\mathrm{A}) + \sinh(\eta)\tanh(\eta_\mathrm{A}) - \mathrm{R_0}\cosh(\eta_\mathrm{A})\right)|}{\mathrm{X}_{0_\mathrm{eff}}} \tag{C.8}$$

$$\Phi(\mathrm{R},\ \eta,\ \phi) = \tan^{-1}\left(\frac{\overrightarrow{AM}\cdot\vec{U}_\mathrm{pz}}{\|\mathrm{AM}\|}, \frac{\left(\overrightarrow{AM}\times\vec{U}_\mathrm{pz}\right)\cdot\vec{U}_\mathrm{t}}{\|\mathrm{AM}\|}\right) \tag{C.9}$$

After the correct description and parameters modeling, it is necessary to verify the convergence of the Equation C.10 using all information about the physics phenomena described in detail in the Appendix B for a given impact point (η, ϕ).

$$\mathrm{E} = \mathrm{E}_0 \int_{\mathrm{V}} \mathrm{dE}(\mathrm{R}, \eta, \phi) = \frac{1}{2\pi} \iiint \underbrace{f_r(R, \eta, \phi)}_{\text{(i)}} \underbrace{f_t(R, \eta, \phi)}_{\text{(ii)}} \underbrace{J \left| \frac{\partial f}{\partial_R} \frac{\partial f}{\partial_\eta} \frac{\partial f}{\partial_\phi} \right|}_{\text{(iii)}} \mathrm{dR}\, \mathrm{d}\eta\, \mathrm{d}\phi \tag{C.10}$$

The parcel (i) refers to the deposition model for the radial profile, the parcel (ii) the longitudinal profile, and the parcel (iii) is the Jacobian associated with the coordinate shifting process.

C.1 Modeling – Specialist Knowledge

In the current ATLAS reconstruction process, the timing information at a cluster level is based on the most energetic cell. This occurs because the effects of XT affect the time reconstructed using single cell OF reconstruction. Then mitigating XT could allow for a combination of cell time information at a cluster level and notably improve the e/γ time resolution. To do so, a cell neighborhood of 5×5 size is defined in the S_2 layer, see Figure 34a. For the other EM layers, the number of cells in the cluster is calculated taking into account the minimal integer number of cells that contains the area from S2, which means to apply a correspondent factor for η and ϕ dimensions. To propagate to S1, it is necessary consider the relation between layers: $8\eta_{S1} = \eta_{S2}$, and $\phi_{S1} = 4\phi_{S2}$, and to propagate to S3 the relationship is: $\eta_{S3} = 2\eta_{S2}$, and $\phi_{S3} = \phi_{S2}$. This approach results in a cluster $(\eta\,, \phi)$ 40×2 in S1, and a cluster $(\eta\,, \phi)$ 3×5 in S3.

Within the selected cluster, there is the incidence of XT, as represented in Figure 34c. In yellow is the first neighborhood, generating the highest levels of XT as shown in Figure 32 and Table 5.

The first neighborhood represents a moving window for XT treatment. In a first approach, this window will be used by moving the three-by-three cluster over the standard five-by-five cluster of interest to process cell information and reconstruct XT corrected energy and time. Equation C.11 presents the model of XT contribution for a given cell by its first neighborhood. The inductive XT contribution from all neighbors is present in the signal, while the capacitive XT is generated only by four neighbor cells, as depicted in Figure 34c

$$E^{rec} = E_{red} + \sum_{m-1}^{8} E^m XT_L^m + \sum_{\ell-1}^{4} E^\ell XT_C^\ell + n, \tag{C.11}$$

XT_C is the capacitive contribution, XT_L is the inductive contribution, *n* is the noise contribution in energy units, and E^m and E^ℓ is the respective energy deposited in the neighbors cells to produce the XT_L and XT_C signals, *m* is the index for the neighbors that add XT_L contribution, and ℓ indicates the neighbors that add XT_C.

A cluster is formed with the target cell at the center and its first eight neighbors around it, adding inductive and capacitive XT contributions. Resistive XT didn't produce any shift in the time of flight of particles, and their amplitude order contribution is around 0.1% of the energy (for peak-peak definition), and this is the reason for not considering its contribution here. This work considers the first-order XT contribution. The previous works of Colas et al. (1999), Hubaut (2000), Hubaut et al. (2000), Pralavorio and Sauvage (2001), Hubaut et al. (2003), Labbé and Ishmukhametov (2008), LOPEZ SOLIS (2017) didn't any mention about other XT contribution beyond the modeling of Equation C.11.

For each cell in the ATLAS LAr cluster, 4 samples are taken for the energy and time reconstruction process using the OF technique. Thus, taking into account the XT signal and noise ($n(t)$), it is possible to

model the composition of each of the four cell samples taking into account the contributions of noise and XT, according to Equation C.12.

$$S_i^k = \hat{S}_i^k + \sum_{j=2}^{9} XT^{j\to k} + n_i^k \tag{C.12}$$

$$S_i^1 = \hat{S}_i^1 + XT^{2\to 1} + XT^{3\to 1} + \ldots + XT^{8\to 1} + XT^{9\to 1} + n_i^1, \tag{C.13}$$

where $\hat{S}_i^k$ is the sample corresponding to the true energy in cell *k* at time *i*; n_i^k, is the noise of the cell *k*, at the sampling instant i; $XT^{j\to k}$ represents the coupling function of the neighboring cell *j* on the cell of interest *k*. Index *i*, indicates the sample, and the *j* index indicates the neighbors that add XT in the interest cell *k*.

Writing Equation C.12 for the hottest cell, cell_{00} (Figure 34b), for example, and defining this cell as $k = 1$, produces Equation C.13. So, rewriting Equation C.13 as a function of energy, $g(t_i, \tau)$ at the instant t_i and time of flight of particle (τ) in a given cell results in Equation C.14, and the last step, using the expansion of Taylor's series, for the first order, see Equation C.15, results in Equation C.16.

$$S_i^1(t_i) = E^1 g^1(t_i, \tau) + \sum_{j=2}^{9} E^j g_{XT}^{j\to 1}(t_i, \tau) + n_i^1 \tag{C.14}$$

$$g^1(t_i, \tau) \approx g^1(t_i) + \tau \dot{g}^1(t_i) \tag{C.15}$$

$$S_i^1(t_i) = E^1 \left(g^1(t_i) - \tau \dot{g}(t_i)\right) + \sum_{j=2}^{9} E^j \left(g_{XT}^{j\to 1}(t_i) + \tau \dot{g}_{XT}^{j\to k}(t_i)\right) + n_i^k \tag{C.16}$$

$$\begin{aligned} S_i^1(t_i) = {} & E^{1,truth} g^1(t_i) + E^{2,truth} g_{XT}^{2\to 1}(t_i) + \ldots + E^{8,truth} g_{XT}^{8\to 1}(t_i) + E^{9,truth} g_{XT}^{9\to 1}(t_i) + \\ & + \tau \left(E^{1,truth} \dot{g}^1(t_i) + E^{2,truth} \dot{g}_{XT}^{2\to 1}(t_i) + \ldots + E^{8,truth} \dot{g}_{XT}^{8\to 1}(t_i) + E^{9,truth} \dot{g}_{XT}^{9\to 1}(t_i)\right) \end{aligned} \tag{C.17}$$

which $E^{k,truth}$ is the true value of the energy recorded in cell *k*, $g(t_i, \tau)$ is the function that describes the ionization signal of the cell of interest with time of flight of particle (τ), $g_{XT}^{j\to k}(t_i, \tau)$ represents the coupling function which the cell of interest (k) receive from its neighbor (j).

Writing in matrix form results in Equation C.18.

$$\begin{aligned} \mathbf{S} = {} & \left[\begin{pmatrix} 0 & g_{XT}^{1\to 2}(t_i) & \ldots & g_{XT}^{1\to 8}(t_i) & g_{XT}^{1\to 9}(t_i) \\ g_{XT}^{2\to 1}(t_i) & 0 & \ldots & g_{XT}^{2\to 8}(t_i) & g_{XT}^{2\to 9}(t_i) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ g_{XT}^{8\to 1}(t_i) & g_{XT}^{8\to 2}(t_i) & \ldots & 0 & g_{XT}^{8\to 9}(t_i) \\ g_{XT}^{9\to 1}(t_i) & g_{XT}^{9\to 2}(t_i) & \ldots & g_{XT}^{9\to 8}(t_i) & 0 \end{pmatrix} + I_9 \times \begin{pmatrix} g^1(t_i) \\ g^2(t_i) \\ \vdots \\ g^8(t_i) \\ g^9(t_i) \end{pmatrix} + \right. \\ & \left. + \tau \left(\begin{pmatrix} 0 & \dot{g}_{XT}^{1\to 2}(t_i) & \ldots & \dot{g}_{XT}^{1\to 8}(t_i) & \dot{g}_{XT}^{1\to 9}(t_i) \\ \dot{g}_{XT}^{2\to 1}(t_i) & 0 & \ldots & \dot{g}_{XT}^{2\to 8}(t_i) & \dot{g}_{XT}^{2\to 9}(t_i) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \dot{g}_{XT}^{8\to 1}(t_i) & \dot{g}_{XT}^{8\to 2}(t_i) & \ldots & 0 & \dot{g}_{XT}^{8\to 9}(t_i) \\ \dot{g}_{XT}^{9\to 1}(t_i) & g_{XT}^{9\to 2}(t_i) & \ldots & g_{XT}^{9\to 8}(t_i) & 0 \end{pmatrix} + I_9 \times \begin{pmatrix} \dot{g}^1(t_i) \\ \dot{g}^2(t_i) \\ \vdots \\ \dot{g}^8(t_i) \\ \dot{g}^9(t_i) \end{pmatrix} \right) \right] \times \begin{pmatrix} E^1 \\ E^2 \\ \vdots \\ E^8 \\ E^9 \end{pmatrix} + \begin{pmatrix} n^1(t_i) \\ n^2(t_i) \\ \vdots \\ n^8(t_i) \\ n^9(t_i) \end{pmatrix} \end{aligned} \tag{C.18}$$

$$\mathbf{S} = \left(\mathbf{G_{XT}} + \mathbf{I}\,\mathbf{G} + \tau(\dot{\mathbf{G}}_{\mathbf{XT}} + \mathbf{I}\,\dot{\mathbf{G}})\right)\mathbf{E} + \mathbf{N} \tag{C.19}$$

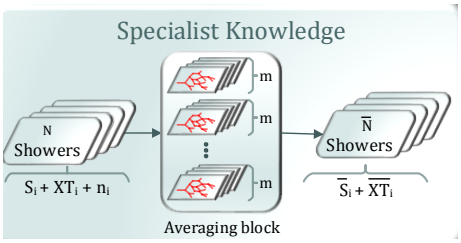
The model described in Equation C.18 refers to the signal composition for the information recorded by a cell when taking a sample at the instant of time t_i in the presence of XT and noise signals. Therefore, the

complete process of energy and time of flight reconstruction of the particle at a cell level requires calculating the information associated with the four samples, which means writing a matrix for each sample to take a solution to this matrix equation.

In Equation C.19, **G** represents the matrix of the transfer function for each cell on 3×3 cluster; $\dot{\mathbf{G}}$ refers to the first derivative of each function of the matrix **G**, obtained from the expansion of the Taylor series; $\mathbf{G_{XT}}$ represents the coupling matrix between the neighboring cells and the cell of interest on 3×3 cluster; $\dot{\mathbf{G_{XT}}}$ refers to the first derivative of each function of the coupling matrix $\mathbf{G_{XT}}$, obtained from the expansion of the Taylor series; τ is the time of flight the particle which is common to all cells inside the cluster. The **E** matrix contains the true energies of the cells, and the **N** matrix contains the noise samples. The values of **S** refer to the samples of each of the nine cells at each sampling time t_i on 3×3 cluster.

The following is a block of the specialist knowledge-based modeling to remove noise contribution in Figure 73. A subset with m clusters is chosen according to specific properties given a large set of clusters. Not only the size, a cell in the collection, for instance. The intention is to bias the average value. That procedure allows the removal of the noise contribution. It avoids some linear dependence between clusters and eliminates singularity in the solution of the linear system for a given property to estimate the XT contributions for a chosen point inside the calorimeter. After that average extraction from Equation C.19 obtains the Equation C.20 without noise.

Figure 73 – Noise mitigation step in the specialist knowledge modeling. This procedure intends to eliminate the noise contribution and organize clusters according to a specific characteristic which avoids linear dependence to solve the system.



$$\overline{\mathbf{S}} = \left(\overline{\mathbf{G_{XT}}} + \mathbf{I}\,\overline{\mathbf{G}} + \tau(\overline{\dot{\mathbf{G}}_{\mathbf{XT}}} + \mathbf{I}\,\overline{\dot{\mathbf{G}}})\right)\overline{\mathbf{E}} \tag{C.20}$$

The model in Equation C.21 describes the contribution of crosstalk types and noise in a given cell at the sample level. Precise energy information comes from the coefficients that solve this model, which allows obtaining the analytical description of the XT contribution. This modeling is the base for the ML approaches evaluated to mitigate this interference and will be described in the following sections. The upper index, *tr*, is the simplified truth index to simplify notation.

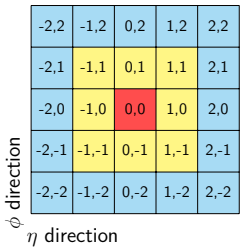
$$\begin{cases} \overline{S}_1^1 = \overline{E}^{1,tr}\overline{g}^1(t_1) + \overline{E}^{2,tr}\overline{g}_{XT}^{2\to1}(t_1) + \ldots + \overline{E}^{9,tr}\overline{g}_{XT}^{9\to1}(t_1) + \tau\left(\overline{E}^{1,tr}\overline{\dot{g}}^1(t_1) + \ldots + \overline{E}^{9,tr}\overline{\dot{g}}_{XT}^{9\to1}(t_1)\right) \\ \overline{S}_2^1 = \overline{E}^{1,tr}\overline{g}^1(t_2) + \overline{E}^{2,tr}\overline{g}_{XT}^{2\to1}(t_2) + \ldots + \overline{E}^{9,tr}\overline{g}_{XT}^{9\to1}(t_2) + \tau\left(\overline{E}^{1,tr}\overline{\dot{g}}^1(t_2) + \ldots + \overline{E}^{9,tr}\overline{\dot{g}}_{XT}^{9\to1}(t_2)\right) \\ \overline{S}_3^1 = \overline{E}^{1,tr}\overline{g}^1(t_3) + \overline{E}^{2,tr}\overline{g}_{XT}^{2\to1}(t_3) + \ldots + \overline{E}^{9,tr}\overline{g}_{XT}^{9\to1}(t_3) + \tau\left(\overline{E}^{1,tr}\overline{\dot{g}}^1(t_3) + \ldots + \overline{E}^{9,tr}\overline{\dot{g}}_{XT}^{9\to1}(t_3)\right) \\ \overline{S}_4^1 = \overline{E}^{1,tr}\overline{g}^1(t_4) + \overline{E}^{2,tr}\overline{g}_{XT}^{2\to1}(t_4) + \ldots + \overline{E}^{9,tr}\overline{g}_{XT}^{9\to1}(t_4) + \tau\left(\overline{E}^{1,tr}\overline{\dot{g}}^1(t_4) + \ldots + \overline{E}^{9,tr}\overline{\dot{g}}_{XT}^{9\to1}(t_4)\right) \end{cases} \tag{C.21}$$

APPENDIX D – Tables and Figures of Results

This appendix brings all tables and figures concerning the results of the method presented in this thesis to mitigate XT effects in an LAr calorimeter using two datasets: an analytical simulator (EMshower), and the Lorenzetti Framework shower simulator.

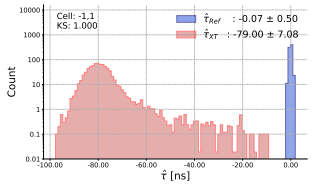
D.1 EMshower

Figure 74 – Cell addressing in the 5×5 cluster. The coordinates $(\eta, \phi) = (0,0)$ indicates the hottest cell.

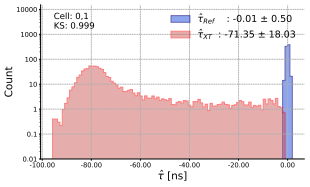


D.1.1 XT effects on Estimation

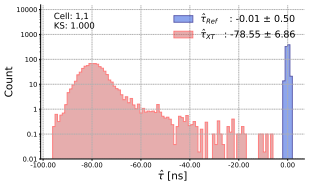
Figure 75 – Histogram for $\hat{\tau}$ with OF using signals without XT for cell$_{ij}$, with an arbitrary time distribution, with mean $= 0$ and $\sigma = 0.5$ ns, in blue, and energy with XT and noise.



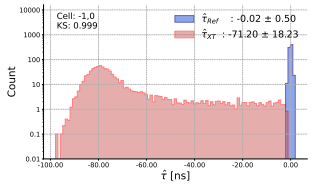
(a) (-1,1).



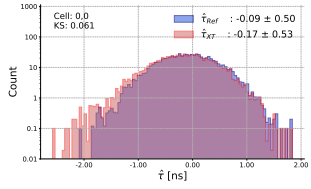
(b) (0,1).



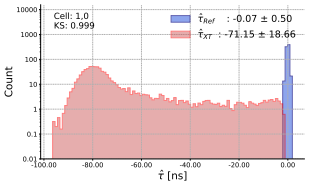
(c) 1,1.



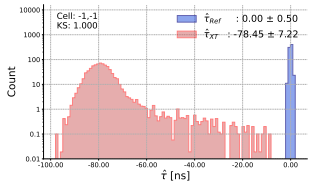
(d) (-1,0).



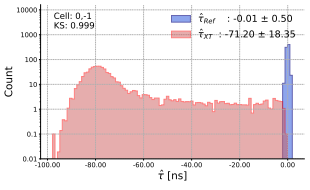
(e) (0,0).



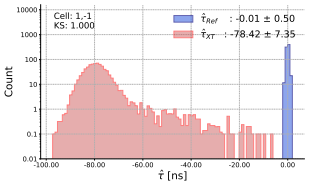
(f) (1,0).



(g) (-1,-1).

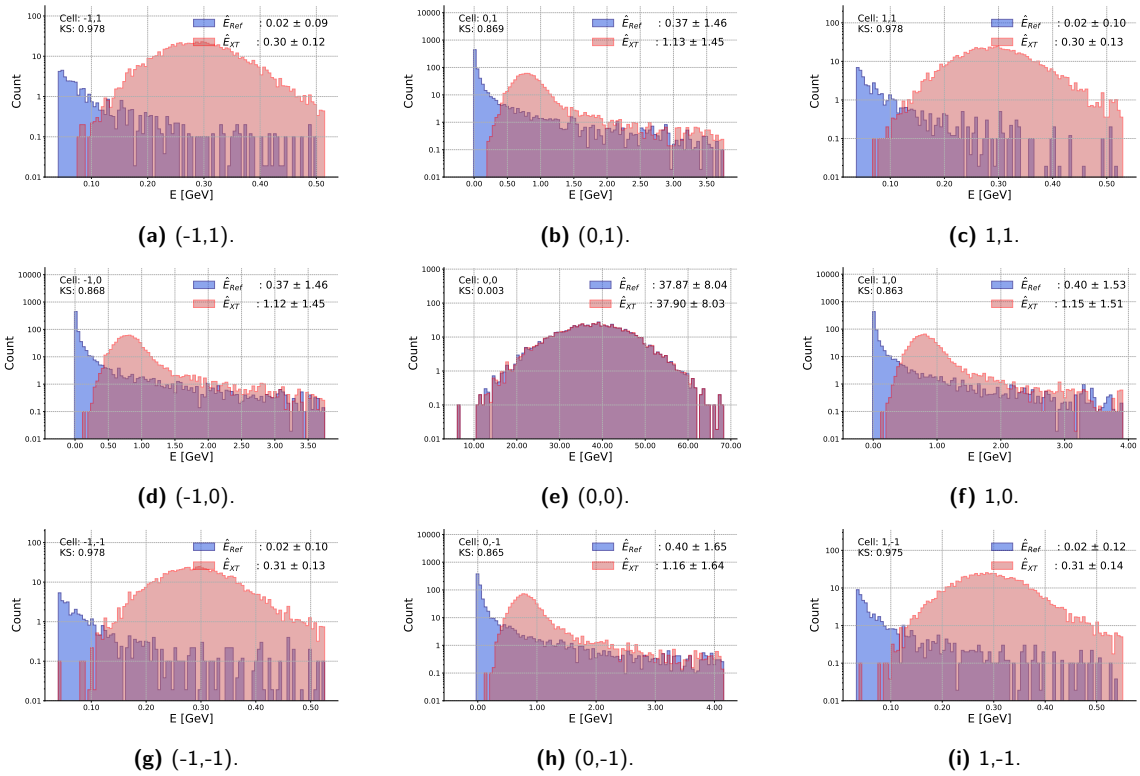


(h) (0,-1).



(i) (1,-1).

Figure 76 – XT effects on energy estimation on 3×3 cluster. In cell (0,0) it is possible to observe that the distribution with and without crosstalk did not produce high effects since this cell has a high ratio $\frac{E}{XT}$.



(a) (-1,1). **(b)** (0,1). **(c)** 1,1.

(d) (-1,0). **(e)** (0,0). **(f)** 1,0.

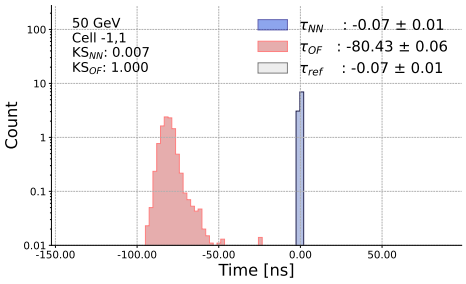
(g) (-1,-1). **(h)** (0,-1). **(i)** 1,-1.

D.1.2 Time Estimation

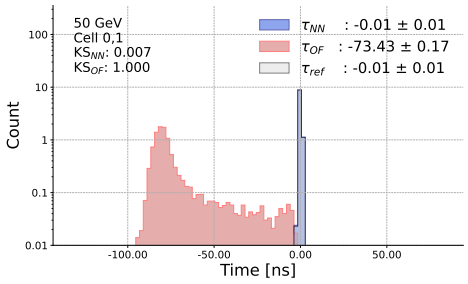
Table 17 – Detail for the best NN structure, using 3 HL with 25-10-5 neurons for τ [ns] regression task. Each cell is presented with its respectively RMSE value, and on the bottom of each structure column is the global RMSE. The last column brings an RMSE ratio between the OF and the NN.

cell	Target	OF		NN		$\frac{RMSE_{OF}}{RMSE_{NN}}$
	$\hat{\tau}$ [ns]	$\hat{\tau}$ [ns]	RMSE	$\hat{\tau}$ [ns]	RMSE	
-1,1	-0.07 ± 0.01	-80.43 ± 0.06	80.5911	-0.07 ± 0.01	0.0159	5068.02
0,1	-0.01 ± 0.01	-73.43 ± 0.17	75.3581	-0.01 ± 0.01	0.0159	4743.53
1,1	-0.01 ± 0.01	-80.03 ± 0.06	80.2332	-0.01 ± 0.01	0.0159	5053.63
-1,0	-0.02 ± 0.01	-72.96 ± 0.18	75.0168	-0.02 ± 0.01	0.0159	4719.85
0,0	-0.09 ± 0.01	-0.18 ± 0.01	0.2320	-0.09 ± 0.01	0.0159	14.57
1,0	-0.07 ± 0.01	-72.91 ± 0.18	75.0202	-0.07 ± 0.01	0.0159	4713.39
-1,-1	0.00 ± 0.01	-79.78 ± 0.07	80.0560	0.00 ± 0.01	0.0159	5047.50
0,-1	-0.01 ± 0.01	-72.83 ± 0.18	74.9712	-0.01 ± 0.01	0.0159	4720.26
1,-1	-0.01 ± 0.01	-79.85 ± 0.07	80.1147	-0.00 ± 0.01	0.0159	5036.72

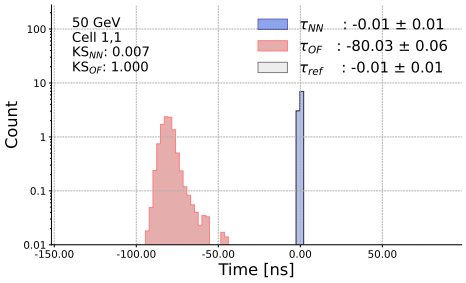
Figure 77 – Time regression histogram cell by cell comparison among OF, target and NN estimator. Comparison between NN structure and OF cell by cell on the center of 3×3 cluster.



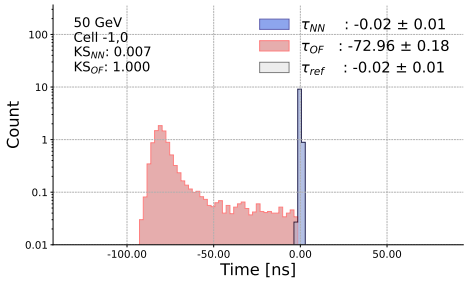
(a) (-1,1).



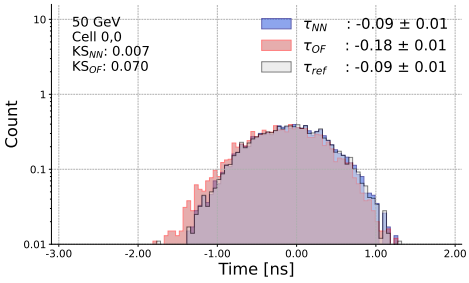
(b) (0,1).



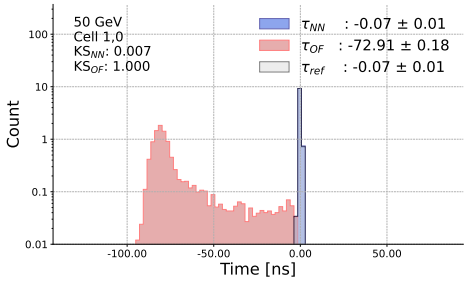
(c) (1,1).



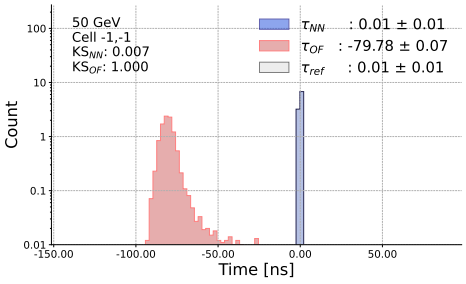
(d) (-1,0).



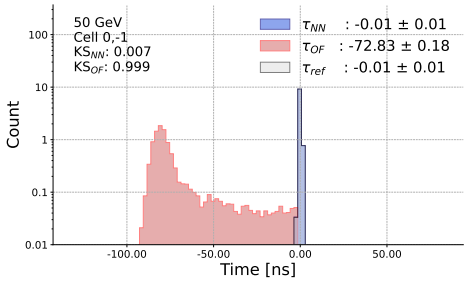
(e) (0,0).



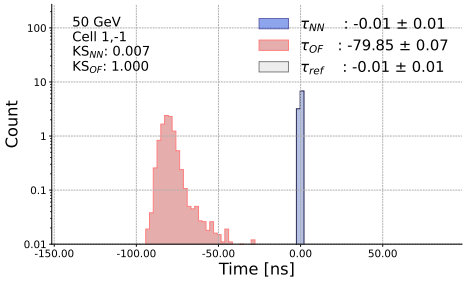
(f) (1,0).



(g) (-1,-1).

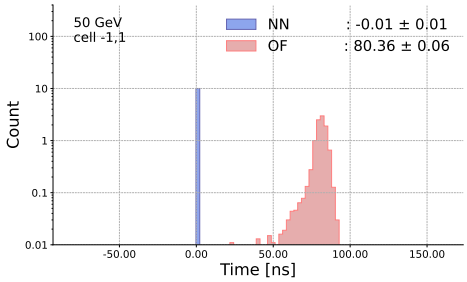


(h) (0,-1).

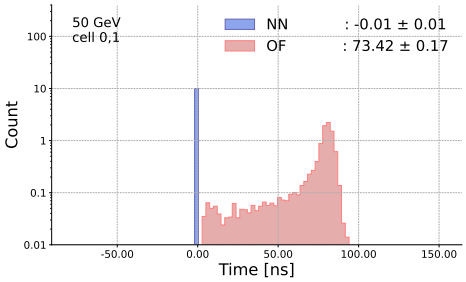


(i) (1,-1).

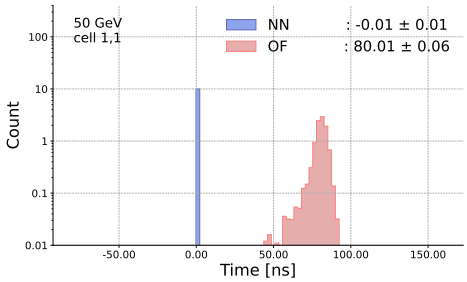
Figure 78 – Details for the reconstruction error for the OF and NN using 1 HL with 20 neurons. The error for NN structure is low and with a narrow spread. For OF it is possible to see the limitation in this technique to dealing with XT influence. The small plot on the right side is a zoom for the tiny distribution of the NN result.



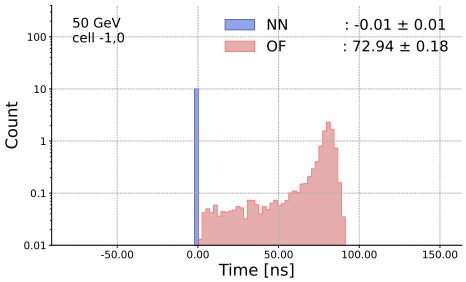
(a) (-1,1).



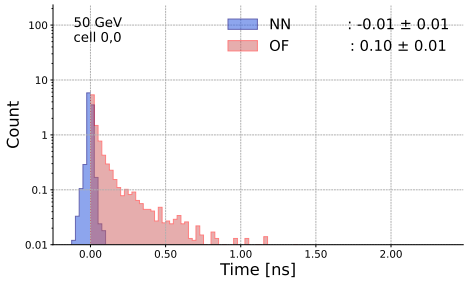
(b) (0,1).



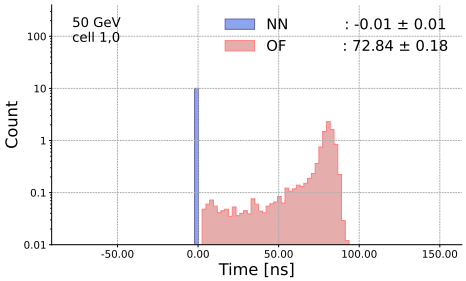
(c) (1,1).



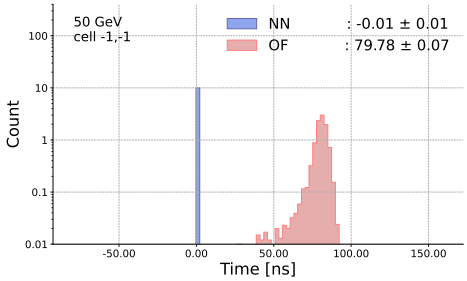
(d) (-1,0).



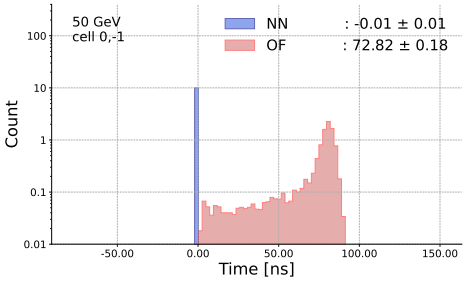
(e) (0,0).



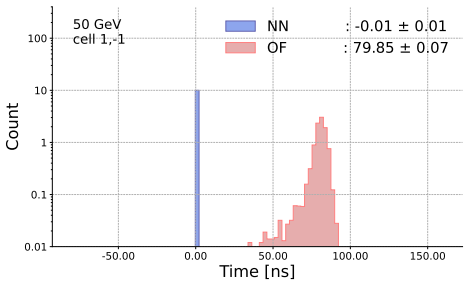
(f) (1,0).



(g) (-1,-1).

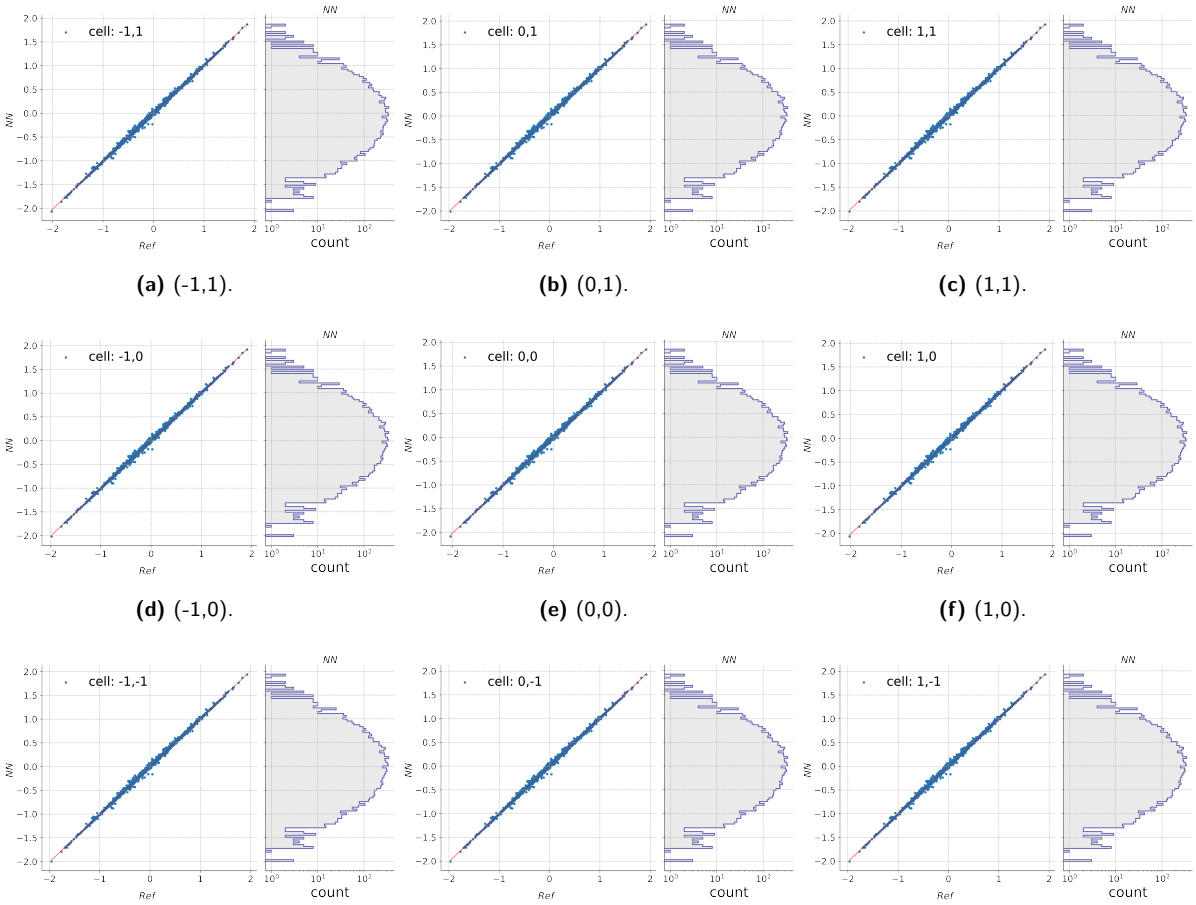


(h) (0,-1).



(i) (1,-1).

Figure 79 – Scatter plots for time reconstructed by NN with structure cell by cell. A good correlation between target values and NN output is observed for all cells in the center of the 3×3 cluster around the hottest cell.



(a) (-1,1). **(b)** (0,1). **(c)** (1,1).

(d) (-1,0). **(e)** (0,0). **(f)** (1,0).

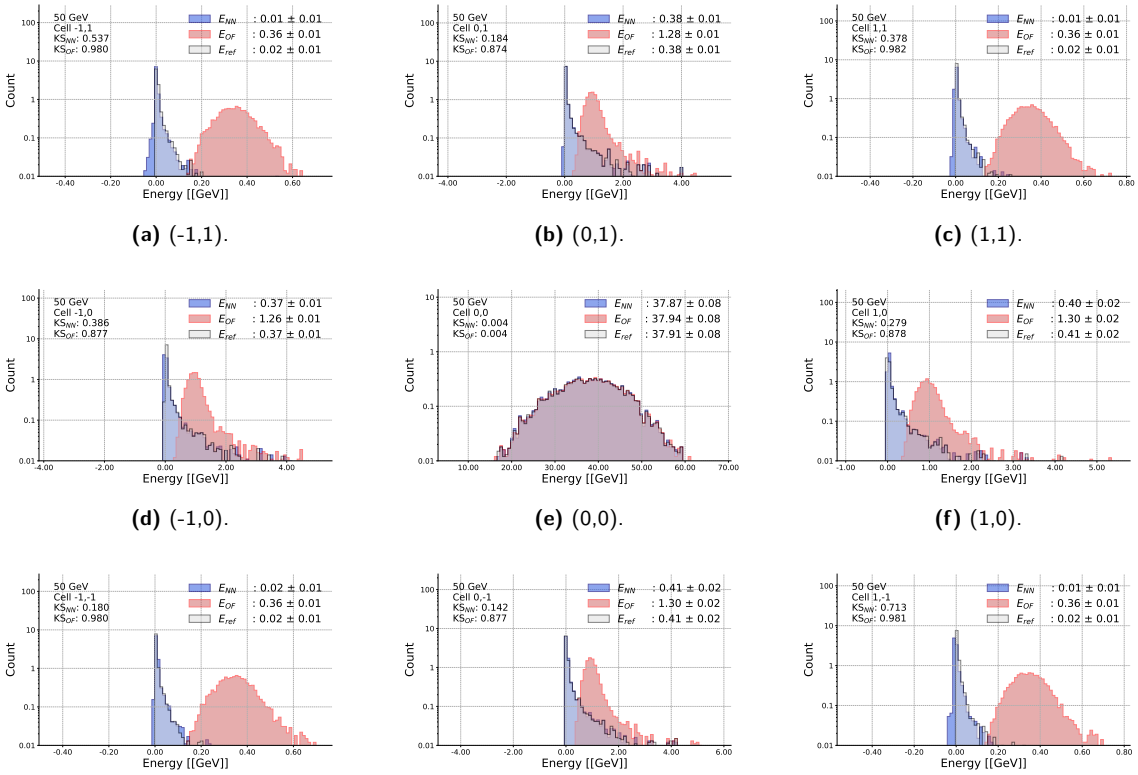
(g) (-1,-1). **(h)** (0,-1). **(i)** (1,-1).

D.1.3 Energy Estimation

Table 18 – Summary of result for the best NN estimator, using 1 hidden layer with 20 neurons for energy (in GeV) regression task for 10,000 clusters. And for each cell is presented its respective RMSE value. The last column shows an RMSE ratio between OF and the NN.

	Target	OF		NN		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	Energy	Energy	RMSE	Energy	RMSE	
-1,1	0.02 ± 0.01	0.36 ± 0.01	0.3496	0.01 ± 0.01	0.0106	33.00
0,1	0.38 ± 0.01	1.28 ± 0.01	0.9157	0.38 ± 0.01	0.0109	83.69
1,1	0.02 ± 0.01	0.36 ± 0.01	0.3517	0.01 ± 0.01	0.0132	26.59
-1,0	0.37 ± 0.01	1.26 ± 0.01	0.9154	0.37 ± 0.01	0.0436	21.01
0,0	37.91 ± 0.08	37.94 ± 0.08	0.0770	37.87 ± 0.08	0.0637	1.21
1,0	0.41 ± 0.02	1.30 ± 0.02	0.9128	0.40 ± 0.02	0.0476	19.18
-1,-1	0.02 ± 0.01	0.36 ± 0.01	0.3518	0.02 ± 0.01	0.0131	26.91
0,-1	0.41 ± 0.02	1.30 ± 0.02	0.9158	0.41 ± 0.02	0.0110	82.90
1,-1	0.02 ± 0.01	0.36 ± 0.01	0.3527	0.01 ± 0.01	0.0118	29.88

Figure 80 – Energy regression histogram cell by cell comparison among OF, target values as a normal distribution and NN structure [100 – 20 – 25]. Comparison between NN structure and OF cell by cell on the center 3×3 cluster.

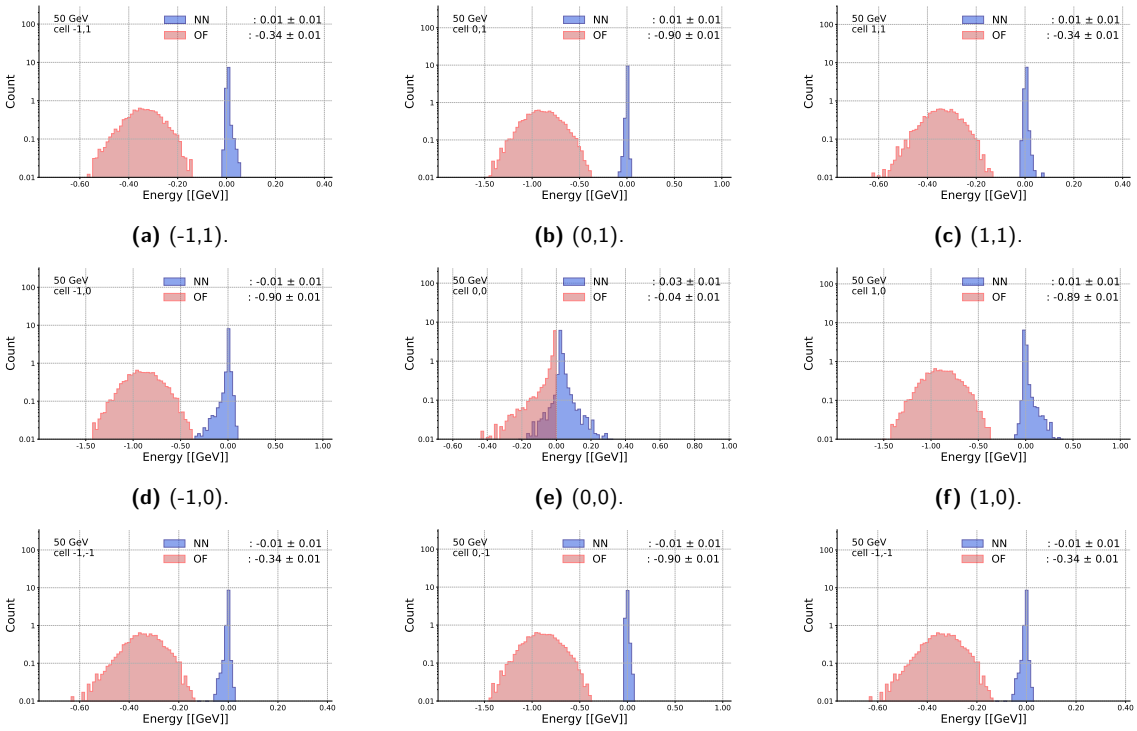


(a) (-1,1). **(b)** (0,1). **(c)** (1,1).

(d) (-1,0). **(e)** (0,0). **(f)** (1,0).

(g) (-1,-1). **(h)** (0,-1). **(i)** (1,-1).

Figure 81 – Regression error for energy for each cell in the cluster cells around the hottest cell with the NN structure using one HL. Even in cells at the corners with low energies, the NN error spread is lower than OF.

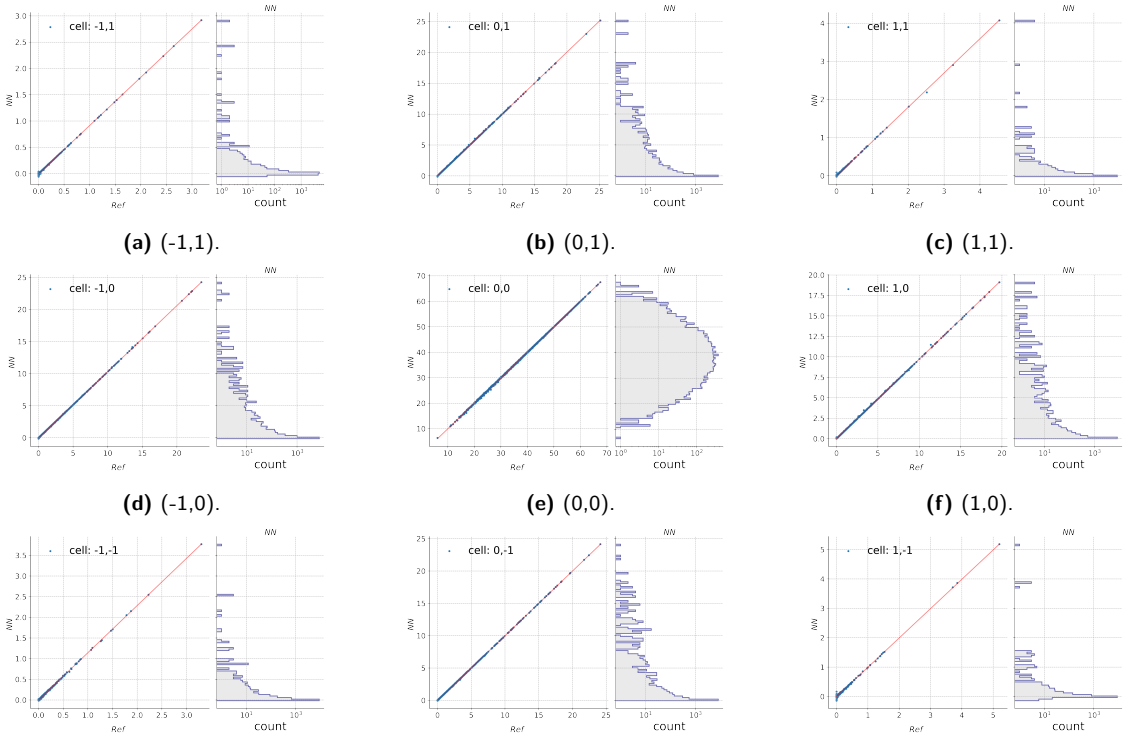


(a) (-1,1). **(b)** (0,1). **(c)** (1,1).

(d) (-1,0). **(e)** (0,0). **(f)** (1,0).

(g) (-1,-1). **(h)** (0,-1). **(i)** (1,-1).

Figure 82 – Scatter plots for energy reconstructed by NN with structure and energies target cell by cell on the center of 3×3 cluster. Additional information is provided with a histogram for NN output.



(a) (-1,1). **(b)** (0,1). **(c)** (1,1).

(d) (-1,0). **(e)** (0,0). **(f)** (1,0).

(g) (-1,-1). **(h)** (0,-1). **(i)** (1,-1).

D.2 Lorenzetti

Table 19 presents the R^2 on each model trained to mitigate XT effects on different energy level. This result is interesting, because the setup used to train the NN models was the same, and the R^2 achieves a good description of the variance on the dataset.

Table 19 – Coefficient of determination ($R^2 \in [0, 1]$) achieved during the training step for each energy impact.

E_{imp} [GeV]	R^2	E_{imp} [GeV]	R^2	E_{imp} [GeV]	R^2	E_{imp} [GeV]	R^2
10	0.9871	60	0.9826	110	0.9784	160	0.9374
20	0.9873	70	0.9726	120	0.9616	170	0.9424
30	0.9804	80	0.9747	130	0.9394	180	0.9170
40	0.9810	90	0.9704	140	0.9434	190	0.9294
50	0.9789	100	0.9826	150	0.9335	200	0.9423

D.2.1 Time Estimation.

Table 20 presents the summary for all 20 impact energy.

Table 20 – Mean $\hat{\tau}$ values on the cluster reconstructed per impact energy over 50,000 clusters. The first column is the reference values, the second, is values applying OF on sample s with XT, and the third, is the values from OF after an NN estimator to regress samples.

E_{imp} [GeV]	τ_{REF} [ns]	OF		NN		$\frac{RMSE_{OF}}{RMSE_{NN}}$
		$\hat{\tau}$	RMSE	$\hat{\tau}$	RMSE	
10	-0.04 ± 0.01	-33.51 ± 22.47	35176.5958	0.27 ± 0.71	1112.7363	31.61
20	-0.04 ± 0.01	-8.99 ± 2.23	3496.0828	0.09 ± 0.48	743.7911	4.70
30	-0.04 ± 0.01	-12.19 ± 2.63	4124.1048	-27.42 ± 22.32	34938.8843	0.12
40	-0.04 ± 0.01	-10.92 ± 1.80	2802.4103	-0.02 ± 0.59	918.1561	3.07
50	-0.04 ± 0.01	-57.85 ± 27.35	42811.1501	3.06 ± 2.45	3839.3026	11.15
60	-0.04 ± 0.01	-13.31 ± 14.05	21990.2805	-1.15 ± 0.61	963.6208	22.86
70	-0.04 ± 0.01	-18.27 ± 2.30	3607.4622	-0.17 ± 0.07	112.3198	32.10
80	-0.04 ± 0.01	-66.14 ± 31.61	49481.9078	0.12 ± 0.21	324.8949	152.23
90	-0.04 ± 0.01	-16.81 ± 2.24	3506.2846	-0.19 ± 0.24	372.4870	9.42
100	-0.04 ± 0.01	-16.88 ± 1.97	3081.3968	2.60 ± 1.27	989.7992	1.55
110	-0.04 ± 0.01	-14.91 ± 1.56	2442.0005	-1.33 ± 0.78	1213.8876	2.01
120	-0.04 ± 0.01	-23.82 ± 3.30	5162.2861	-0.21 ± 0.04	68.0527	75.87
130	-0.05 ± 0.01	-34.57 ± 56.96	89152.3899	0.24 ± 0.08	129.4762	688.42
140	-0.04 ± 0.01	-21.76 ± 6.37	9966.4853	-0.08 ± 0.01	10.5352	948.68
150	-0.05 ± 0.01	-15.36 ± 1.99	3115.7136	-0.07 ± 0.04	57.0034	54.72
160	-0.04 ± 0.01	-10.75 ± 6.99	10939.8759	-0.74 ± 0.69	1087.7008	10.06
170	-0.04 ± 0.01	-17.71 ± 2.14	3357.3886	0.19 ± 0.08	128.6695	26.07
180	-0.04 ± 0.01	-19.26 ± 2.23	3494.7691	0.06 ± 0.08	127.2972	27.36
190	-0.04 ± 0.01	-20.37 ± 1.86	2908.3917	0.42 ± 0.61	961.6941	3.03
200	-0.04 ± 0.01	-16.02 ± 2.51	3928.1845	-0.09 ± 0.01	1.8017	1993.45

Table 21 – Lorenzetti [10 GeV] - $\hat{\tau}$.

	Target	OF		$NN_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	τ [ns]	$\hat{\tau}$ [ns]	RMSE	$\hat{\tau}$ [ns]	RMSE	
-1,1	-0.05 ± 0.01	16.21 ± 44.85	10029.2457	-0.48 ± 0.64	141.9917	70.63
0,1	-0.06 ± 0.01	-23.43 ± 2.09	467.6845	1.15 ± 0.84	188.7966	2.48
1,1	-0.06 ± 0.01	18.17 ± 36.34	8125.1920	-0.61 ± 0.24	53.1146	152.97
-1,0	-0.03 ± 0.01	-18.61 ± 2.25	503.6397	0.15 ± 0.04	8.7101	57.82
0,0	-0.04 ± 0.01	-3.09 ± 0.01	3.1498	-0.02 ± 0.01	0.4970	6.34
1,0	-0.09 ± 0.01	-11.42 ± 0.16	38.3616	-0.17 ± 0.01	2.3306	16.46
-1,-1	-0.04 ± 0.01	-51.86 ± 20.21	4519.4109	-0.26 ± 0.16	36.2172	124.79
0,-1	-0.01 ± 0.01	-288.37 ± 261.33	58434.8707	-0.11 ± 0.05	10.2319	5711.06
1,-1	-0.03 ± 0.01	-19.41 ± 5.31	1188.3114	1.23 ± 0.59	132.3008	8.98

Table 22 – Lorenzetti [20 GeV] - $\hat{\tau}$.

	Target	OF		$NN_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	τ [ns]	$\hat{\tau}$ [ns]	RMSE	$\hat{\tau}$ [ns]	RMSE	
-1,1	-0.06 ± 0.01	-37.67 ± 6.98	1561.2909	0.28 ± 0.30	67.1377	23.26
0,1	-0.07 ± 0.01	-23.24 ± 0.11	33.6514	-0.59 ± 1.51	336.9053	0.10
1,1	-0.07 ± 0.01	-23.92 ± 0.58	131.4716	-2.92 ± 1.66	371.9156	0.35
-1,0	-0.03 ± 0.01	-24.87 ± 0.05	27.0282	-0.03 ± 0.02	3.7663	7.18
0,0	-0.04 ± 0.01	-3.29 ± 0.01	3.2937	-0.04 ± 0.01	0.2558	12.88
1,0	-0.10 ± 0.01	-8.45 ± 0.03	10.1869	0.93 ± 1.04	231.6371	0.04
-1,-1	-0.04 ± 0.01	-29.54 ± 2.33	520.8196	1.00 ± 0.41	91.2228	5.71
0,-1	-0.01 ± 0.01	-23.27 ± 0.11	34.5055	0.06 ± 0.68	151.6197	0.23
1,-1	-0.04 ± 0.01	-21.44 ± 1.83	409.5119	3.53 ± 1.82	405.8874	1.01

Table 23 – Lorenzetti [30 GeV] - $\hat{\tau}$.

	Target	OF		NN$_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	τ [ns]	$\hat{\tau}$ [ns]	RMSE	$\hat{\tau}$ [ns]	RMSE	
-1,1	-0.06 ± 0.01	-29.61 ± 0.08	34.4520	-0.10 ± 0.02	5.2498	6.56
0,1	-0.06 ± 0.01	-22.62 ± 0.09	30.3366	-0.01 ± 0.06	13.1966	2.30
1,1	-0.06 ± 0.01	-25.78 ± 0.06	29.2643	0.09 ± 0.03	7.4288	3.94
-1,0	-0.03 ± 0.01	-21.42 ± 0.02	21.9203	0.01 ± 0.01	0.3423	64.03
0,0	-0.04 ± 0.01	-2.41 ± 0.01	2.3887	-0.03 ± 0.01	0.2897	8.25
1,0	-0.09 ± 0.01	-15.40 ± 0.01	15.6588	-0.08 ± 0.01	0.3144	49.81
-1,-1	-0.04 ± 0.01	-29.65 ± 0.08	34.4463	0.04 ± 0.04	9.5236	3.62
0,-1	-0.01 ± 0.01	-22.64 ± 0.09	30.3628	-0.04 ± 0.06	12.9033	2.35
1,-1	-0.04 ± 0.01	-25.77 ± 0.07	29.5604	0.02 ± 0.06	13.1772	2.24

Table 24 – Lorenzetti [40 GeV] - $\hat{\tau}$.

	Target	OF		NN$_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	τ [ns]	$\hat{\tau}$ [ns]	RMSE	$\hat{\tau}$ [ns]	RMSE	
-1,1	-0.06 ± 0.01	-26.67 ± 0.05	28.9335	0.01 ± 0.09	19.4370	1.49
0,1	-0.07 ± 0.01	-23.01 ± 0.09	30.5505	0.17 ± 0.47	105.4360	0.29
1,1	-0.07 ± 0.01	-28.85 ± 0.05	30.5437	-0.16 ± 0.23	52.0816	0.59
-1,0	-0.03 ± 0.01	-17.04 ± 0.01	17.2949	-0.04 ± 0.01	0.3128	55.30
0,0	-0.04 ± 0.01	-2.38 ± 0.01	2.3511	-0.03 ± 0.01	0.2907	8.09
1,0	-0.09 ± 0.01	-21.34 ± 0.02	21.5989	-0.08 ± 0.01	0.3097	69.75
-1,-1	-0.04 ± 0.01	-26.37 ± 0.05	28.6131	-0.03 ± 0.02	3.4042	8.41
0,-1	-0.01 ± 0.01	-22.27 ± 0.09	29.9696	-0.02 ± 0.09	19.3146	1.55
1,-1	-0.04 ± 0.01	-28.64 ± 0.05	30.3772	-0.01 ± 0.03	7.6083	3.99

Table 25 – Lorenzetti [50 GeV] - $\hat{\tau}$.

	Target	OF		NN$_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	τ [ns]	$\hat{\tau}$ [ns]	RMSE	$\hat{\tau}$ [ns]	RMSE	
-1,1	-0.06 ± 0.01	-24.96 ± 0.05	27.4797	0.03 ± 0.01	0.9443	29.10
0,1	-0.06 ± 0.01	-22.81 ± 0.09	30.3374	-0.04 ± 0.04	7.9005	3.84
1,1	-0.06 ± 0.01	-30.78 ± 0.04	31.8494	0.06 ± 0.01	1.8563	17.16
-1,0	-0.03 ± 0.01	-14.17 ± 0.01	14.3191	0.01 ± 0.01	0.2185	65.54
0,0	-0.04 ± 0.01	-2.44 ± 0.01	2.4125	-0.02 ± 0.01	0.1908	12.65
1,0	-0.09 ± 0.01	-25.05 ± 0.02	25.2814	-0.09 ± 0.01	0.2229	113.40
-1,-1	-0.04 ± 0.01	-24.84 ± 0.05	27.3718	0.56 ± 0.41	92.7546	0.30
0,-1	-0.01 ± 0.01	-22.59 ± 0.09	30.2116	-0.16 ± 0.07	15.2508	1.98
1,-1	-0.04 ± 0.01	-30.75 ± 0.04	31.8457	-0.15 ± 0.02	3.3762	9.43

Table 26 – Lorenzetti [60 GeV] - $\hat{\tau}$.

	Target	OF		NN$_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	τ [ns]	$\hat{\tau}$ [ns]	RMSE	$\hat{\tau}$ [ns]	RMSE	
-1,1	-0.06 ± 0.01	-24.14 ± 0.05	27.0190	0.01 ± 0.01	1.5979	16.91
0,1	-0.06 ± 0.01	-23.01 ± 0.09	30.4515	-0.04 ± 0.02	3.5267	8.63
1,1	-0.06 ± 0.01	-32.26 ± 0.03	33.0181	-0.01 ± 0.01	0.7051	46.83
-1,0	-0.03 ± 0.01	-12.25 ± 0.01	12.3381	-0.01 ± 0.01	0.2132	57.87
0,0	-0.04 ± 0.01	-2.52 ± 0.01	2.4929	-0.01 ± 0.01	0.2005	12.43
1,0	-0.09 ± 0.01	-27.57 ± 0.02	27.7703	-0.07 ± 0.01	0.2269	122.39
-1,-1	-0.04 ± 0.01	-23.82 ± 0.05	26.7476	0.08 ± 0.02	5.2822	5.06
0,-1	-0.01 ± 0.01	-22.45 ± 0.09	30.0441	0.04 ± 0.07	15.3802	1.95
1,-1	-0.04 ± 0.01	-32.10 ± 0.03	32.8755	-0.08 ± 0.04	7.9121	4.16

Table 27 – Lorenzetti [70 GeV] - $\hat{\tau}$.

	Target	OF		NN$_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	τ [ns]	$\hat{\tau}$ [ns]	RMSE	$\hat{\tau}$ [ns]	RMSE	
-1,1	-0.06 ± 0.01	-23.62 ± 0.06	26.8390	-0.13 ± 0.01	0.3009	89.19
0,1	-0.07 ± 0.01	-23.18 ± 0.09	30.5708	-0.14 ± 0.01	0.4533	67.44
1,1	-0.07 ± 0.01	-33.16 ± 0.03	33.7291	0.07 ± 0.05	12.1354	2.78
-1,0	-0.03 ± 0.01	-10.89 ± 0.01	10.9504	-0.03 ± 0.01	0.1795	61.00
0,0	-0.04 ± 0.01	-2.61 ± 0.01	2.5789	-0.03 ± 0.01	0.1670	15.44
1,0	-0.09 ± 0.01	-29.30 ± 0.02	29.4671	-0.09 ± 0.01	0.1945	151.54
-1,-1	-0.04 ± 0.01	-23.19 ± 0.06	26.4587	0.12 ± 0.01	0.3720	71.13
0,-1	-0.01 ± 0.01	-22.45 ± 0.09	30.0212	-0.24 ± 0.01	0.8197	36.62
1,-1	-0.04 ± 0.01	-33.00 ± 0.03	33.5913	0.04 ± 0.02	5.5112	6.10

Table 28 – Lorenzetti [80 GeV] - $\hat{\tau}$.

	Target	OF		NN$_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	τ [ns]	$\hat{\tau}$ [ns]	RMSE	$\hat{\tau}$ [ns]	RMSE	
-1,1	-0.06 ± 0.01	-23.01 ± 0.06	26.5627	0.05 ± 0.02	3.3253	7.99
0,1	-0.07 ± 0.01	-22.85 ± 0.09	30.2777	-0.01 ± 0.03	5.7938	5.23
1,1	-0.07 ± 0.01	-33.68 ± 0.03	34.1056	0.08 ± 0.01	0.3425	99.57
-1,0	-0.03 ± 0.01	-9.86 ± 0.01	9.8941	-0.01 ± 0.01	0.1817	54.46
0,0	-0.04 ± 0.01	-2.71 ± 0.01	2.6713	-0.02 ± 0.01	0.1711	15.62
1,0	-0.10 ± 0.01	-30.49 ± 0.02	30.6263	-0.02 ± 0.01	0.2086	146.79
-1,-1	-0.04 ± 0.01	-23.05 ± 0.06	26.6291	-0.17 ± 0.01	2.7898	9.55
0,-1	-0.01 ± 0.01	-22.88 ± 0.09	30.3232	-0.25 ± 0.16	35.3663	0.86
1,-1	-0.04 ± 0.01	-33.71 ± 0.03	34.1698	0.13 ± 0.01	0.3359	101.71

Table 29 – Lorenzetti [90 GeV] - $\hat{\tau}$.

	Target	OF		NN$_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	τ [ns]	$\hat{\tau}$ [ns]	RMSE	$\hat{\tau}$ [ns]	RMSE	
-1,1	-0.06 ± 0.01	-22.82 ± 0.06	26.6770	-0.08 ± 0.10	21.3705	1.25
0,1	-0.07 ± 0.01	-22.91 ± 0.09	30.3112	0.01 ± 0.06	14.1675	2.14
1,1	-0.07 ± 0.01	-34.25 ± 0.02	34.6019	0.06 ± 0.01	0.2964	116.74
-1,0	-0.03 ± 0.01	-9.09 ± 0.01	9.1022	0.01 ± 0.01	0.2074	43.88
0,0	-0.04 ± 0.01	-2.79 ± 0.01	2.7573	-0.01 ± 0.01	0.2002	13.77
1,0	-0.10 ± 0.01	-31.43 ± 0.02	31.5532	-0.06 ± 0.01	0.2205	143.09
-1,-1	-0.04 ± 0.01	-22.83 ± 0.06	26.6751	-0.01 ± 0.05	10.8794	2.45
0,-1	-0.01 ± 0.01	-22.93 ± 0.09	30.3283	0.03 ± 0.01	3.1015	9.78
1,-1	-0.04 ± 0.01	-34.17 ± 0.02	34.5400	-0.02 ± 0.01	0.7657	45.11

Table 30 – Lorenzetti [100 GeV] - $\hat{\tau}$.

	Target	OF		NN$_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	τ [ns]	$\hat{\tau}$ [ns]	RMSE	$\hat{\tau}$ [ns]	RMSE	
-1,1	-0.06 ± 0.01	-22.74 ± 0.06	26.8249	0.07 ± 0.01	0.2379	112.78
0,1	-0.07 ± 0.01	-23.07 ± 0.09	30.3681	0.03 ± 0.01	0.3203	94.80
1,1	-0.07 ± 0.01	-34.53 ± 0.02	34.8181	0.04 ± 0.02	3.8773	8.98
-1,0	-0.03 ± 0.01	-8.46 ± 0.01	8.4705	-0.01 ± 0.01	0.1326	63.88
0,0	-0.04 ± 0.01	-2.88 ± 0.01	2.8472	-0.02 ± 0.01	0.1202	23.68
1,0	-0.09 ± 0.01	-32.09 ± 0.02	32.1941	-0.08 ± 0.01	0.1451	221.89
-1,-1	-0.04 ± 0.01	-22.57 ± 0.06	26.6916	-0.03 ± 0.01	2.2892	11.66
0,-1	-0.01 ± 0.01	-22.76 ± 0.09	30.1559	0.03 ± 0.01	0.4583	65.79
1,-1	-0.04 ± 0.01	-34.46 ± 0.02	34.7770	-0.10 ± 0.01	0.2316	150.17

Table 31 – Lorenzetti [110 GeV] - $\hat{\tau}$.

	Target	OF		NN$_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	τ [ns]	$\hat{\tau}$ [ns]	RMSE	$\hat{\tau}$ [ns]	RMSE	
-1,1	-0.06 ± 0.01	-22.60 ± 0.07	26.9218	-0.31 ± 0.08	17.9271	1.50
0,1	-0.07 ± 0.01	-23.01 ± 0.09	30.2866	-0.15 ± 0.01	0.4019	75.36
1,1	-0.07 ± 0.01	-34.71 ± 0.02	34.9586	0.01 ± 0.01	0.2182	160.19
-1,0	-0.03 ± 0.01	-7.94 ± 0.01	7.9446	-0.02 ± 0.01	0.1276	62.28
0,0	-0.04 ± 0.01	-2.97 ± 0.01	2.9324	-0.02 ± 0.01	0.1145	25.62
1,0	-0.09 ± 0.01	-32.54 ± 0.02	32.6402	-0.11 ± 0.01	0.1744	187.20
-1,-1	-0.04 ± 0.01	-22.53 ± 0.07	26.8537	0.14 ± 0.09	20.5082	1.31
0,-1	-0.01 ± 0.01	-22.86 ± 0.09	30.1851	0.05 ± 0.01	0.4089	73.81
1,-1	-0.04 ± 0.01	-34.65 ± 0.02	34.9227	-0.01 ± 0.01	0.1766	197.73

Table 32 – Lorenzetti [120 GeV] - $\hat{\tau}$.

	Target	OF		$NN_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	τ [ns]	$\hat{\tau}$ [ns]	RMSE	$\hat{\tau}$ [ns]	RMSE	
-1,1	-0.06 ± 0.01	-22.39 ± 0.07	26.9396	-0.02 ± 0.01	0.2937	91.73
0,1	-0.06 ± 0.01	-22.84 ± 0.09	30.1190	-0.10 ± 0.01	0.3434	87.71
1,1	-0.06 ± 0.01	-34.87 ± 0.02	35.0856	-0.05 ± 0.01	0.1477	237.57
-1,0	-0.03 ± 0.01	-7.50 ± 0.01	7.5013	0.01 ± 0.01	0.1348	55.65
0,0	-0.04 ± 0.01	-3.05 ± 0.01	3.0156	-0.03 ± 0.01	0.1113	27.10
1,0	-0.09 ± 0.01	-32.90 ± 0.02	32.9738	-0.01 ± 0.01	0.1628	202.49
-1,-1	-0.04 ± 0.01	-22.61 ± 0.07	27.1216	-0.01 ± 0.01	0.6600	41.09
0,-1	-0.01 ± 0.01	-23.07 ± 0.09	30.2978	-0.10 ± 0.01	0.5908	51.29
1,-1	-0.04 ± 0.01	-34.82 ± 0.02	35.0644	0.07 ± 0.01	2.3650	14.83

Table 33 – Lorenzetti [130 GeV] - $\hat{\tau}$.

	Target	OF		$NN_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	τ [ns]	$\hat{\tau}$ [ns]	RMSE	$\hat{\tau}$ [ns]	RMSE	
-1,1	-0.06 ± 0.01	-22.51 ± 0.07	27.2147	-0.14 ± 0.01	0.1768	153.96
0,1	-0.07 ± 0.01	-23.05 ± 0.09	30.2309	-0.03 ± 0.01	0.2132	141.76
1,1	-0.07 ± 0.01	-34.98 ± 0.02	35.1563	0.05 ± 0.01	0.2068	170.01
-1,0	-0.04 ± 0.01	-7.15 ± 0.01	7.1337	-0.02 ± 0.01	0.1030	69.25
0,0	-0.05 ± 0.01	-3.14 ± 0.01	3.0976	-0.02 ± 0.01	0.0967	32.03
1,0	-0.10 ± 0.01	-33.23 ± 0.02	33.2916	-0.02 ± 0.01	0.1511	220.26
-1,-1	-0.05 ± 0.01	-22.45 ± 0.07	27.1652	0.05 ± 0.01	0.2569	105.74
0,-1	-0.01 ± 0.01	-22.92 ± 0.09	30.1464	0.02 ± 0.01	0.3590	83.98
1,-1	-0.04 ± 0.01	-34.95 ± 0.02	35.1491	-0.08 ± 0.01	1.1003	31.94

Table 34 – Lorenzetti [140 GeV] - $\hat{\tau}$.

	Target	OF		$NN_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	τ [ns]	$\hat{\tau}$ [ns]	RMSE	$\hat{\tau}$ [ns]	RMSE	
-1,1	-0.06 ± 0.01	-22.51 ± 0.07	27.3648	0.05 ± 0.01	0.4145	66.02
0,1	-0.07 ± 0.01	-23.10 ± 0.09	30.2278	0.03 ± 0.01	0.3299	91.63
1,1	-0.07 ± 0.01	-34.98 ± 0.02	35.1325	0.05 ± 0.01	0.3400	103.33
-1,0	-0.03 ± 0.01	-6.84 ± 0.01	6.8337	-0.02 ± 0.01	0.0948	72.10
0,0	-0.04 ± 0.01	-3.22 ± 0.01	3.1785	-0.04 ± 0.01	0.0784	40.56
1,0	-0.10 ± 0.01	-33.41 ± 0.01	33.4633	-0.05 ± 0.01	0.1806	185.25
-1,-1	-0.04 ± 0.01	-22.40 ± 0.07	27.2911	-0.10 ± 0.01	0.2504	108.97
0,-1	-0.01 ± 0.01	-22.86 ± 0.09	30.0488	-0.01 ± 0.01	0.4261	70.51
1,-1	-0.04 ± 0.01	-34.95 ± 0.02	35.1353	-0.12 ± 0.01	0.1704	206.15

Table 35 – Lorenzetti [150 GeV] - $\hat{\tau}$.

	Target	OF		$NN_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	τ [ns]	$\hat{\tau}$ [ns]	RMSE	$\hat{\tau}$ [ns]	RMSE	
-1,1	-0.06 ± 0.01	-22.48 ± 0.07	27.4951	0.06 ± 0.01	0.3455	79.59
0,1	-0.07 ± 0.01	-23.04 ± 0.09	30.1145	-0.07 ± 0.01	0.2678	112.43
1,1	-0.07 ± 0.01	-35.03 ± 0.02	35.1667	-0.01 ± 0.01	0.1861	189.02
-1,0	-0.04 ± 0.01	-6.57 ± 0.01	6.5563	-0.02 ± 0.01	0.1023	64.09
0,0	-0.05 ± 0.01	-3.30 ± 0.01	3.2592	-0.02 ± 0.01	0.0904	36.05
1,0	-0.10 ± 0.01	-33.58 ± 0.01	33.6212	-0.04 ± 0.01	0.1343	250.31
-1,-1	-0.05 ± 0.01	-22.46 ± 0.07	27.4863	-0.06 ± 0.01	0.5337	51.50
0,-1	-0.02 ± 0.01	-22.94 ± 0.09	30.0644	0.05 ± 0.01	0.2507	119.93
1,-1	-0.04 ± 0.01	-34.97 ± 0.02	35.1315	0.01 ± 0.01	0.1547	227.09

Table 36 – Lorenzetti [160 GeV] - $\hat{\tau}$.

	Target	OF		$NN_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	τ [ns]	$\hat{\tau}$ [ns]	RMSE	$\hat{\tau}$ [ns]	RMSE	
-1,1	-0.06 ± 0.01	-22.61 ± 0.07	27.7485	-0.02 ± 0.01	0.4683	59.26
0,1	-0.07 ± 0.01	-23.16 ± 0.09	30.1570	-0.10 ± 0.01	0.4221	71.45
1,1	-0.07 ± 0.01	-35.01 ± 0.02	35.1277	-0.04 ± 0.01	0.1313	267.59
-1,0	-0.03 ± 0.01	-6.32 ± 0.01	6.3018	-0.03 ± 0.01	0.0892	70.64
0,0	-0.04 ± 0.01	-3.38 ± 0.01	3.3375	-0.06 ± 0.01	0.0886	37.66
1,0	-0.10 ± 0.01	-33.69 ± 0.01	33.7264	-0.16 ± 0.01	0.1200	281.16
-1,-1	-0.04 ± 0.01	-22.32 ± 0.07	27.5087	-0.05 ± 0.01	0.2952	93.18
0,-1	-0.01 ± 0.01	-22.80 ± 0.09	29.9058	-0.12 ± 0.01	0.3449	86.71
1,-1	-0.04 ± 0.01	-34.94 ± 0.02	35.0863	-0.02 ± 0.01	0.2338	150.09

Table 37 – Lorenzetti [170 GeV] - $\hat{\tau}$.

	Target	OF		$NN_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	τ [ns]	$\hat{\tau}$ [ns]	RMSE	$\hat{\tau}$ [ns]	RMSE	
-1,1	-0.06 ± 0.01	-22.48 ± 0.07	27.7732	0.08 ± 0.01	0.3158	87.95
0,1	-0.07 ± 0.01	-23.00 ± 0.09	29.9922	0.05 ± 0.01	0.2993	100.20
1,1	-0.07 ± 0.01	-34.97 ± 0.02	35.0781	0.10 ± 0.01	0.2788	125.82
-1,0	-0.03 ± 0.01	-6.10 ± 0.01	6.0888	-0.02 ± 0.01	0.0883	68.96
0,0	-0.04 ± 0.01	-3.45 ± 0.01	3.4181	-0.03 ± 0.01	0.0739	46.25
1,0	-0.09 ± 0.01	-33.74 ± 0.01	33.7697	-0.07 ± 0.01	0.1143	295.57
-1,-1	-0.04 ± 0.01	-22.48 ± 0.07	27.7710	-0.03 ± 0.01	2.0384	13.62
0,-1	-0.01 ± 0.01	-22.97 ± 0.09	29.9928	0.04 ± 0.01	0.3779	79.37
1,-1	-0.04 ± 0.01	-34.91 ± 0.02	35.0405	-0.09 ± 0.01	0.1736	201.84

Table 38 – Lorenzetti [180 GeV] - $\hat{\tau}$.

	Target	OF		$NN_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	τ [ns]	$\hat{\tau}$ [ns]	RMSE	$\hat{\tau}$ [ns]	RMSE	
-1,1	-0.06 ± 0.01	-22.38 ± 0.07	27.8224	-0.01 ± 0.01	1.1348	24.52
0,1	-0.06 ± 0.01	-22.87 ± 0.09	29.8430	-0.09 ± 0.01	0.4173	71.51
1,1	-0.06 ± 0.01	-34.85 ± 0.02	34.9469	0.02 ± 0.01	0.1656	210.98
-1,0	-0.03 ± 0.01	-5.90 ± 0.01	5.8841	-0.01 ± 0.01	0.0958	61.43
0,0	-0.04 ± 0.01	-3.53 ± 0.01	3.4979	-0.01 ± 0.01	0.0890	39.30
1,0	-0.09 ± 0.01	-33.73 ± 0.01	33.7557	-0.05 ± 0.01	0.1195	282.36
-1,-1	-0.04 ± 0.01	-22.63 ± 0.07	28.0303	0.07 ± 0.01	0.3007	93.21
0,-1	-0.01 ± 0.01	-23.06 ± 0.09	30.0007	-0.04 ± 0.01	0.3439	87.23
1,-1	-0.04 ± 0.01	-34.86 ± 0.02	34.9823	-0.20 ± 0.01	0.2648	132.09

Table 39 – Lorenzetti [190 GeV] - $\hat{\tau}$.

	Target	OF		$NN_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	τ [ns]	$\hat{\tau}$ [ns]	RMSE	$\hat{\tau}$ [ns]	RMSE	
-1,1	-0.06 ± 0.01	-22.46 ± 0.08	27.9959	0.10 ± 0.01	1.1999	23.33
0,1	-0.07 ± 0.01	-22.94 ± 0.09	29.8468	0.02 ± 0.01	0.2806	106.36
1,1	-0.07 ± 0.01	-34.82 ± 0.02	34.9114	-0.15 ± 0.01	0.1311	266.35
-1,0	-0.03 ± 0.01	-5.73 ± 0.01	5.7299	-0.03 ± 0.01	2.1527	2.66
0,0	-0.04 ± 0.01	-3.61 ± 0.01	3.5742	-0.04 ± 0.01	0.0723	49.45
1,0	-0.09 ± 0.01	-33.71 ± 0.01	33.7338	-0.10 ± 0.01	0.0997	338.38
-1,-1	-0.04 ± 0.01	-22.58 ± 0.08	28.1002	0.03 ± 0.01	1.0030	28.02
0,-1	-0.01 ± 0.01	-22.99 ± 0.09	29.9004	0.02 ± 0.01	0.2837	105.38
1,-1	-0.04 ± 0.01	-34.73 ± 0.02	34.8502	0.19 ± 0.01	0.2950	118.13

Table 40 – Lorenzetti [200 GeV] - $\hat{\tau}$.

	Target	OF		$NN_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	τ [ns]	$\hat{\tau}$ [ns]	RMSE	$\hat{\tau}$ [ns]	RMSE	
-1,1	-0.06 ± 0.01	-22.58 ± 0.08	28.2033	-0.01 ± 0.01	0.2759	102.21
0,1	-0.07 ± 0.01	-23.02 ± 0.09	29.8520	-0.01 ± 0.01	0.2224	134.22
1,1	-0.07 ± 0.01	-34.67 ± 0.01	34.7526	-0.23 ± 0.01	0.2197	158.18
-1,0	-0.03 ± 0.01	-5.57 ± 0.01	5.5764	-0.03 ± 0.01	0.1430	39.00
0,0	-0.04 ± 0.01	-3.69 ± 0.01	3.6530	-0.04 ± 0.01	0.0790	46.24
1,0	-0.10 ± 0.01	-33.66 ± 0.01	33.6739	-0.05 ± 0.01	0.1193	282.19
-1,-1	-0.04 ± 0.01	-22.56 ± 0.08	28.2071	-0.01 ± 0.01	0.6358	44.36
0,-1	-0.01 ± 0.01	-22.91 ± 0.09	29.7807	-0.02 ± 0.01	0.3295	90.38
1,-1	-0.04 ± 0.01	-34.63 ± 0.01	34.7365	-0.28 ± 0.01	0.3295	105.41

D.2.2 Energy Estimation.

Table 41 presents a summary for the total energy on cluster per impact energy comparing the NN estimator and the OF algorithm.

Table 41 – Total energy cluster reconstructed by impact energy over 50,000 clusters. Column E_{REF} is the reference value for energies without XT. Column OF is the values applying OF on samples with XT, and column NN is the values from OF after an NN estimator to regress samples.

E [GeV]	E_{REF} [GeV]	OF		NN		$\frac{RMSE_{OF}}{RMSE_{NN}}$
		E [GeV]	RMSE	E [GeV]	RMSE	
10	5.44 ± 0.01	6.12 ± 0.01	0.0292	5.44 ± 0.01	0.0408	2.65
20	12.51 ± 0.01	14.08 ± 0.01	0.0607	12.52 ± 0.01	0.0513	4.10
30	20.01 ± 0.01	22.52 ± 0.01	0.0996	19.98 ± 0.01	0.0734	4.32
40	27.82 ± 0.01	31.31 ± 0.02	0.1387	27.92 ± 0.01	0.0861	5.15
50	35.77 ± 0.02	40.26 ± 0.02	0.1776	35.67 ± 0.02	0.1027	4.72
60	43.83 ± 0.02	49.34 ± 0.02	0.2165	43.62 ± 0.02	0.1088	6.56
70	52.02 ± 0.02	58.56 ± 0.02	0.2556	51.22 ± 0.02	0.1350	4.28
80	60.18 ± 0.02	67.75 ± 0.02	0.2943	59.86 ± 0.02	0.1297	6.66
90	68.60 ± 0.02	77.22 ± 0.03	0.3342	68.05 ± 0.02	0.1483	5.92
100	76.95 ± 0.03	86.63 ± 0.03	0.3737	75.43 ± 0.03	0.2045	3.49
110	85.33 ± 0.03	96.06 ± 0.03	0.4131	84.04 ± 0.03	0.2035	4.17
120	93.80 ± 0.03	105.60 ± 0.03	0.4530	92.40 ± 0.03	0.2345	3.88
130	102.41 ± 0.03	115.29 ± 0.03	0.4933	100.36 ± 0.03	0.2708	3.37
140	110.92 ± 0.03	124.87 ± 0.04	0.5334	108.55 ± 0.04	0.2896	3.44
150	119.49 ± 0.03	134.51 ± 0.04	0.5735	117.22 ± 0.04	0.3183	3.46
160	128.19 ± 0.03	144.32 ± 0.04	0.6146	125.48 ± 0.04	0.3568	3.23
170	136.82 ± 0.04	154.03 ± 0.04	0.6552	133.41 ± 0.04	0.3651	3.33
180	145.43 ± 0.04	163.73 ± 0.04	0.6958	141.89 ± 0.04	0.3957	3.15
190	154.06 ± 0.04	173.44 ± 0.04	0.7364	150.63 ± 0.04	0.3715	3.63
200	162.73 ± 0.04	183.20 ± 0.05	0.7770	159.29 ± 0.04	0.3862	3.66

Table 42 – Lorenzetti [10 GeV] - $\hat{E}$.

	Target	OF		$\mathrm{NN}_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	Energy [GeV]	Energy [GeV]	RMSE	Energy [GeV]	RMSE	
-1,1	0.16 ± 0.01	0.20 ± 0.01	0.0433	0.16 ± 0.01	0.0269	1.61
0,1	0.60 ± 0.01	0.66 ± 0.01	0.0610	0.60 ± 0.01	0.0647	0.94
1,1	0.30 ± 0.01	0.35 ± 0.01	0.0522	0.31 ± 0.01	0.0377	1.38
-1,0	0.32 ± 0.01	0.38 ± 0.01	0.0607	0.31 ± 0.01	0.0289	2.10
0,0	1.60 ± 0.01	1.66 ± 0.01	0.0615	1.60 ± 0.01	0.0284	2.16
1,0	0.63 ± 0.01	0.70 ± 0.01	0.0693	0.64 ± 0.01	0.0347	2.00
-1,-1	0.16 ± 0.01	0.20 ± 0.01	0.0441	0.16 ± 0.01	0.0277	1.59
0,-1	0.62 ± 0.01	0.68 ± 0.01	0.0616	0.61 ± 0.01	0.0695	0.89
1,-1	0.31 ± 0.01	0.35 ± 0.01	0.0526	0.31 ± 0.01	0.0370	1.42

Table 43 – Lorenzetti 20 GeV] - $\hat{E}$.

	Target	OF		$\mathrm{NN}_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	Energy [GeV]	Energy [GeV]	RMSE	Energy [GeV]	RMSE	
-1,1	0.25 ± 0.01	0.33 ± 0.01	0.0893	0.25 ± 0.01	0.0457	1.95
0,1	1.50 ± 0.01	1.63 ± 0.01	0.1351	1.50 ± 0.01	0.1350	1.00
1,1	0.87 ± 0.01	0.99 ± 0.01	0.1225	0.87 ± 0.01	0.0984	1.25
-1,0	0.47 ± 0.01	0.59 ± 0.01	0.1265	0.47 ± 0.01	0.0536	2.36
0,0	3.58 ± 0.01	3.72 ± 0.01	0.1439	3.58 ± 0.01	0.0538	2.67
1,0	1.79 ± 0.01	1.94 ± 0.01	0.1597	1.78 ± 0.01	0.0498	3.21
-1,-1	0.25 ± 0.01	0.33 ± 0.01	0.0904	0.25 ± 0.01	0.0455	1.99
0,-1	1.52 ± 0.01	1.65 ± 0.01	0.1355	1.53 ± 0.01	0.1369	0.99
1,-1	0.87 ± 0.01	0.99 ± 0.01	0.1235	0.88 ± 0.01	0.0970	1.27

Table 44 – Lorenzetti [30 GeV] - $\hat{E}$.

	Target	OF		NN$_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	Energy [GeV]	Energy [GeV]	RMSE	Energy [GeV]	RMSE	
-1,1	0.42 ± 0.01	0.59 ± 0.01	0.1828	0.42 ± 0.01	0.0434	4.21
0,1	3.33 ± 0.02	3.54 ± 0.02	0.2217	3.36 ± 0.02	0.3106	0.71
1,1	0.64 ± 0.01	0.82 ± 0.01	0.1936	0.65 ± 0.01	0.0678	2.86
-1,0	0.75 ± 0.01	1.00 ± 0.01	0.2572	0.76 ± 0.01	0.0276	9.30
0,0	7.46 ± 0.01	7.68 ± 0.01	0.2190	7.46 ± 0.01	0.1390	1.58
1,0	1.20 ± 0.01	1.46 ± 0.01	0.2681	1.20 ± 0.01	0.0329	8.14
-1,-1	0.42 ± 0.01	0.58 ± 0.01	0.1822	0.41 ± 0.01	0.0416	4.38
0,-1	3.29 ± 0.02	3.51 ± 0.02	0.2222	3.23 ± 0.02	0.2924	0.76
1,-1	0.64 ± 0.01	0.81 ± 0.01	0.1932	0.63 ± 0.01	0.0586	3.30

Table 45 – Lorenzetti [40 GeV] - $\hat{E}$.

	Target	OF		NN$_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	Energy [GeV]	Energy [GeV]	RMSE	Energy [GeV]	RMSE	
-1,1	0.79 ± 0.01	1.03 ± 0.01	0.2677	0.78 ± 0.01	0.0840	3.19
0,1	4.74 ± 0.03	5.04 ± 0.03	0.3063	4.68 ± 0.03	0.3685	0.83
1,1	0.60 ± 0.01	0.83 ± 0.01	0.2583	0.59 ± 0.01	0.0729	3.54
-1,0	1.45 ± 0.01	1.82 ± 0.01	0.3729	1.46 ± 0.01	0.0470	7.93
0,0	10.52 ± 0.01	10.82 ± 0.01	0.3058	10.57 ± 0.01	0.1516	2.02
1,0	1.06 ± 0.01	1.42 ± 0.01	0.3617	1.06 ± 0.01	0.0304	11.88
-1,-1	0.79 ± 0.01	1.04 ± 0.01	0.2699	0.81 ± 0.01	0.0760	3.55
0,-1	4.81 ± 0.03	5.12 ± 0.03	0.3075	4.87 ± 0.03	0.3496	0.88
1,-1	0.60 ± 0.01	0.83 ± 0.01	0.2609	0.61 ± 0.01	0.0619	4.21

Table 46 – Lorenzetti [50 GeV] - $\hat{E}$.

	Target	OF		NN$_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	Energy [GeV]	Energy [GeV]	RMSE	Energy [GeV]	RMSE	
-1,1	1.27 ± 0.01	1.59 ± 0.01	0.3540	1.27 ± 0.01	0.1187	2.98
0,1	6.13 ± 0.03	6.52 ± 0.03	0.3919	6.22 ± 0.03	0.4223	0.93
1,1	0.61 ± 0.01	0.90 ± 0.01	0.3228	0.62 ± 0.01	0.0589	5.48
-1,0	2.36 ± 0.01	2.84 ± 0.01	0.4858	2.35 ± 0.01	0.0616	7.89
0,0	13.31 ± 0.01	13.70 ± 0.01	0.3973	13.27 ± 0.01	0.2865	1.39
1,0	1.05 ± 0.01	1.50 ± 0.01	0.4499	1.05 ± 0.01	0.0325	13.83
-1,-1	1.26 ± 0.01	1.58 ± 0.01	0.3542	1.25 ± 0.01	0.1119	3.17
0,-1	6.12 ± 0.03	6.51 ± 0.03	0.3929	5.98 ± 0.03	0.4133	0.95
1,-1	0.61 ± 0.01	0.90 ± 0.01	0.3231	0.60 ± 0.01	0.0530	6.09

Table 47 – Lorenzetti [60 GeV] - $\hat{E}$.

	Target	OF		NN$_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	Energy [GeV]	Energy [GeV]	RMSE	Energy [GeV]	RMSE	
-1,1	1.80 ± 0.01	2.20 ± 0.01	0.4370	1.82 ± 0.01	0.1243	3.52
0,1	7.31 ± 0.04	7.78 ± 0.04	0.4765	7.41 ± 0.04	0.4036	1.18
1,1	0.63 ± 0.01	0.97 ± 0.01	0.3810	0.64 ± 0.01	0.0587	6.49
-1,0	3.45 ± 0.01	4.04 ± 0.01	0.5969	3.42 ± 0.01	0.0589	10.14
0,0	15.90 ± 0.01	16.39 ± 0.01	0.4919	15.83 ± 0.01	0.2375	2.07
1,0	1.08 ± 0.01	1.61 ± 0.01	0.5334	1.07 ± 0.01	0.0294	18.13
-1,-1	1.84 ± 0.01	2.25 ± 0.01	0.4417	1.80 ± 0.01	0.1488	2.97
0,-1	7.47 ± 0.04	7.95 ± 0.04	0.4794	7.30 ± 0.04	0.4947	0.97
1,-1	0.64 ± 0.01	0.99 ± 0.01	0.3852	0.63 ± 0.01	0.0634	6.08

Table 48 – Lorenzetti [70 GeV] - $\hat{E}$.

	Target	OF		NN$_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	Energy [GeV]	Energy [GeV]	RMSE	Energy [GeV]	RMSE	
-1,1	2.45 ± 0.01	2.93 ± 0.01	0.5239	2.40 ± 0.01	0.1436	3.65
0,1	8.55 ± 0.05	9.11 ± 0.05	0.5634	8.37 ± 0.05	0.4667	1.21
1,1	0.68 ± 0.01	1.07 ± 0.01	0.4399	0.67 ± 0.01	0.0691	6.37
-1,0	4.68 ± 0.01	5.38 ± 0.01	0.7074	4.61 ± 0.01	0.1512	4.68
0,0	18.38 ± 0.01	18.96 ± 0.01	0.5892	18.10 ± 0.01	0.5664	1.04
1,0	1.13 ± 0.01	1.74 ± 0.01	0.6136	1.12 ± 0.01	0.0406	15.13
-1,-1	2.46 ± 0.01	2.95 ± 0.01	0.5265	2.43 ± 0.01	0.1301	4.05
0,-1	8.63 ± 0.05	9.19 ± 0.05	0.5656	8.51 ± 0.05	0.4272	1.32
1,-1	0.68 ± 0.01	1.07 ± 0.01	0.4424	0.66 ± 0.01	0.0678	6.53

Table 49 – Lorenzetti [80 GeV] - $\hat{E}$.

	Target	OF		NN$_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	Energy [GeV]	Energy [GeV]	RMSE	Energy [GeV]	RMSE	
-1,1	3.15 ± 0.02	3.72 ± 0.02	0.6108	3.14 ± 0.02	0.2004	3.05
0,1	9.76 ± 0.05	10.40 ± 0.05	0.6493	9.73 ± 0.05	0.4640	1.40
1,1	0.72 ± 0.01	1.16 ± 0.01	0.4965	0.72 ± 0.01	0.0665	7.47
-1,0	6.04 ± 0.01	6.85 ± 0.01	0.8171	6.01 ± 0.01	0.1295	6.31
0,0	20.72 ± 0.02	21.41 ± 0.02	0.6883	20.60 ± 0.02	0.4185	1.64
1,0	1.19 ± 0.01	1.88 ± 0.01	0.6897	1.19 ± 0.01	0.0399	17.29
-1,-1	3.16 ± 0.02	3.72 ± 0.02	0.6115	3.13 ± 0.02	0.2074	2.95
0,-1	9.76 ± 0.05	10.40 ± 0.05	0.6516	9.67 ± 0.05	0.4793	1.36
1,-1	0.72 ± 0.01	1.16 ± 0.01	0.4970	0.72 ± 0.01	0.0642	7.74

Table 50 – Lorenzetti [90 GeV] - $\hat{E}$.

	Target	OF		NN$_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	Energy [GeV]	Energy [GeV]	RMSE	Energy [GeV]	RMSE	
-1,1	3.95 ± 0.02	4.61 ± 0.02	0.7018	3.87 ± 0.02	0.2506	2.80
0,1	11.00 ± 0.06	11.73 ± 0.06	0.7389	10.79 ± 0.06	0.6187	1.19
1,1	0.77 ± 0.01	1.27 ± 0.01	0.5536	0.76 ± 0.01	0.0722	7.67
-1,0	7.53 ± 0.01	8.46 ± 0.01	0.9294	7.47 ± 0.01	0.1672	5.56
0,0	23.03 ± 0.02	23.81 ± 0.02	0.7914	22.82 ± 0.02	0.5103	1.55
1,0	1.26 ± 0.01	2.02 ± 0.01	0.7655	1.24 ± 0.01	0.0439	17.43
-1,-1	3.88 ± 0.02	4.53 ± 0.02	0.6982	3.88 ± 0.02	0.2361	2.96
0,-1	10.79 ± 0.06	11.52 ± 0.06	0.7397	10.81 ± 0.06	0.4953	1.49
1,-1	0.76 ± 0.01	1.25 ± 0.01	0.5505	0.76 ± 0.01	0.0614	8.97

Table 51 – Lorenzetti [100 GeV] - $\hat{E}$.

	Target	OF		NN$_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	Energy [GeV]	Energy [GeV]	RMSE	Time	RMSE	
-1,1	4.73 ± 0.02	5.46 ± 0.03	0.7867	4.66 ± 0.02	0.2422	3.25
0,1	11.97 ± 0.07	12.79 ± 0.07	0.8250	11.79 ± 0.07	0.5802	1.42
1,1	0.82 ± 0.01	1.35 ± 0.01	0.6035	0.79 ± 0.01	0.0828	7.29
-1,0	9.13 ± 0.01	10.17 ± 0.01	1.0388	8.94 ± 0.01	0.3684	2.82
0,0	25.18 ± 0.02	26.07 ± 0.02	0.8948	24.73 ± 0.02	0.9113	0.98
1,0	1.33 ± 0.01	2.16 ± 0.01	0.8365	1.30 ± 0.01	0.0903	9.26
-1,-1	4.73 ± 0.02	5.47 ± 0.03	0.7881	4.64 ± 0.02	0.2666	2.96
0,-1	11.97 ± 0.07	12.78 ± 0.07	0.8279	11.65 ± 0.06	0.6673	1.24
1,-1	0.82 ± 0.01	1.35 ± 0.01	0.6044	0.80 ± 0.01	0.0885	6.83

Table 52 – Lorenzetti [110 GeV] - $\hat{E}$.

	Target	OF		NN$_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	Energy [GeV]	Energy [GeV]	RMSE	Energy [GeV]	RMSE	
-1,1	5.55 ± 0.03	6.37 ± 0.03	0.8727	5.45 ± 0.03	0.3315	2.63
0,1	12.93 ± 0.07	13.83 ± 0.07	0.9113	12.72 ± 0.07	0.6513	1.40
1,1	0.87 ± 0.01	1.45 ± 0.01	0.6518	0.85 ± 0.01	0.0794	8.20
-1,0	10.82 ± 0.01	11.96 ± 0.01	1.1472	10.64 ± 0.01	0.3692	3.11
0,0	27.24 ± 0.02	28.24 ± 0.02	0.9986	26.86 ± 0.02	0.8547	1.17
1,0	1.40 ± 0.01	2.30 ± 0.01	0.9047	1.38 ± 0.01	0.0684	13.23
-1,-1	5.60 ± 0.03	6.42 ± 0.03	0.8765	5.57 ± 0.03	0.2932	2.99
0,-1	13.06 ± 0.07	13.96 ± 0.07	0.9156	12.90 ± 0.07	0.6264	1.46
1,-1	0.87 ± 0.01	1.46 ± 0.01	0.6556	0.86 ± 0.01	0.0691	9.49

Table 53 – Lorenzetti [120 GeV] - $\hat{E}$.

	Target	OF		NN$_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	Energy [GeV]	Energy [GeV]	RMSE	Energy [GeV]	RMSE	
-1,1	6.44 ± 0.03	7.34 ± 0.04	0.9595	6.40 ± 0.03	0.3022	3.18
0,1	13.86 ± 0.08	14.84 ± 0.08	0.9985	13.76 ± 0.08	0.6067	1.65
1,1	0.92 ± 0.01	1.54 ± 0.01	0.6989	0.91 ± 0.01	0.0705	9.91
-1,0	12.61 ± 0.01	13.86 ± 0.01	1.2566	12.39 ± 0.01	0.4702	2.67
0,0	29.24 ± 0.02	30.34 ± 0.02	1.1049	28.79 ± 0.02	1.0404	1.06
1,0	1.48 ± 0.01	2.44 ± 0.01	0.9707	1.45 ± 0.01	0.1191	8.15
-1,-1	6.55 ± 0.03	7.46 ± 0.04	0.9667	6.41 ± 0.03	0.3667	2.64
0,-1	14.13 ± 0.08	15.12 ± 0.08	1.0046	13.66 ± 0.07	0.9174	1.10
1,-1	0.93 ± 0.01	1.55 ± 0.01	0.7058	0.90 ± 0.01	0.0907	7.78

Table 54 – Lorenzetti [130 GeV] - $\hat{E}$.

	Target	OF		NN$_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	Energy [GeV]	Energy [GeV]	RMSE	Energy [GeV]	RMSE	
-1,1	7.48 ± 0.04	8.48 ± 0.04	1.0551	7.32 ± 0.04	0.4135	2.55
0,1	15.10 ± 0.08	16.17 ± 0.08	1.0909	14.78 ± 0.08	0.7962	1.37
1,1	0.98 ± 0.01	1.64 ± 0.01	0.7522	0.94 ± 0.01	0.0943	7.98
-1,0	14.50 ± 0.01	15.86 ± 0.01	1.3672	14.22 ± 0.01	0.5616	2.43
0,0	31.25 ± 0.02	32.46 ± 0.02	1.2131	30.72 ± 0.02	1.1826	1.03
1,0	1.55 ± 0.01	2.59 ± 0.01	1.0365	1.51 ± 0.01	0.0814	12.74
-1,-1	7.40 ± 0.04	8.39 ± 0.04	1.0516	7.28 ± 0.04	0.4243	2.48
0,-1	14.89 ± 0.08	15.96 ± 0.08	1.0923	14.53 ± 0.08	0.8512	1.28
1,-1	0.97 ± 0.01	1.63 ± 0.01	0.7487	0.93 ± 0.01	0.0803	9.33

Table 55 – Lorenzetti [140 GeV] - $\hat{E}$.

	Target	OF		NN$_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	Energy [GeV]	Energy [GeV]	RMSE	Energy [GeV]	RMSE	
-1,1	8.47 ± 0.04	9.55 ± 0.05	1.1439	8.25 ± 0.04	0.4883	2.34
0,1	16.04 ± 0.09	17.19 ± 0.09	1.1795	15.59 ± 0.09	0.9529	1.24
1,1	1.03 ± 0.01	1.74 ± 0.01	0.7980	0.99 ± 0.01	0.1321	6.04
-1,0	16.42 ± 0.01	17.89 ± 0.01	1.4762	16.09 ± 0.01	0.6425	2.30
0,0	33.12 ± 0.02	34.44 ± 0.02	1.3209	32.42 ± 0.02	1.2879	1.03
1,0	1.63 ± 0.01	2.73 ± 0.01	1.0991	1.59 ± 0.01	0.0985	11.16
-1,-1	8.37 ± 0.04	9.45 ± 0.05	1.1403	8.26 ± 0.04	0.3895	2.93
0,-1	15.83 ± 0.09	16.99 ± 0.09	1.1807	15.58 ± 0.09	0.8368	1.41
1,-1	1.02 ± 0.01	1.73 ± 0.01	0.7947	1.02 ± 0.01	0.0830	9.58

Table 56 – Lorenzetti [150 GeV] - $\hat{E}$.

	Target	OF		NN$_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	Energy [GeV]	Energy [GeV]	RMSE	Energy [GeV]	RMSE	
-1,1	9.42 ± 0.05	10.58 ± 0.05	1.2287	9.08 ± 0.05	0.6163	1.99
0,1	16.76 ± 0.09	18.00 ± 0.09	1.2655	16.14 ± 0.09	1.1084	1.14
1,1	1.07 ± 0.01	1.82 ± 0.01	0.8381	1.06 ± 0.01	0.0944	8.88
-1,0	18.44 ± 0.01	20.02 ± 0.01	1.5842	18.08 ± 0.01	0.7147	2.22
0,0	34.92 ± 0.02	36.34 ± 0.02	1.4302	34.27 ± 0.02	1.2793	1.12
1,0	1.71 ± 0.01	2.86 ± 0.01	1.1589	1.69 ± 0.01	0.0774	14.97
-1,-1	9.48 ± 0.05	10.65 ± 0.05	1.2338	9.43 ± 0.05	0.4070	3.03
0,-1	16.94 ± 0.09	18.19 ± 0.09	1.2710	16.85 ± 0.09	0.7440	1.71
1,-1	1.08 ± 0.01	1.83 ± 0.01	0.8430	1.07 ± 0.01	0.0730	11.55

Table 57 – Lorenzetti [160 GeV] - $\hat{E}$.

	Target	OF		NN$_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	Energy [GeV]	Energy [GeV]	RMSE	Energy [GeV]	RMSE	
-1,1	10.45 ± 0.06	11.69 ± 0.06	1.3195	10.25 ± 0.05	0.5631	2.34
0,1	17.68 ± 0.10	19.01 ± 0.10	1.3544	17.32 ± 0.10	0.9745	1.39
1,1	1.12 ± 0.01	1.91 ± 0.01	0.8823	1.11 ± 0.01	0.0797	11.07
-1,0	20.54 ± 0.01	22.23 ± 0.01	1.6932	20.13 ± 0.01	0.8213	2.06
0,0	36.69 ± 0.02	38.23 ± 0.02	1.5412	35.99 ± 0.02	1.4251	1.08
1,0	1.79 ± 0.01	3.00 ± 0.01	1.2180	1.75 ± 0.01	0.0870	14.00
-1,-1	10.56 ± 0.06	11.81 ± 0.06	1.3244	10.34 ± 0.05	0.5740	2.31
0,-1	17.81 ± 0.10	19.14 ± 0.10	1.3616	17.35 ± 0.09	1.0603	1.28
1,-1	1.13 ± 0.01	1.92 ± 0.01	0.8861	1.11 ± 0.01	0.0715	12.40

Table 58 – Lorenzetti [170 GeV] - $\hat{E}$.

	Target	OF		NN$_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	Energy [GeV]	Energy [GeV]	RMSE	Energy [GeV]	RMSE	
-1,1	11.66 ± 0.06	13.01 ± 0.06	1.4134	11.35 ± 0.06	0.6403	2.21
0,1	18.72 ± 0.10	20.14 ± 0.10	1.4460	18.20 ± 0.10	1.1201	1.29
1,1	1.18 ± 0.01	2.01 ± 0.01	0.9276	1.16 ± 0.01	0.0758	12.24
-1,0	22.69 ± 0.01	24.48 ± 0.01	1.8013	22.12 ± 0.01	0.9117	1.98
0,0	38.36 ± 0.02	40.01 ± 0.02	1.6525	37.48 ± 0.02	1.4998	1.10
1,0	1.86 ± 0.01	3.13 ± 0.01	1.2743	1.81 ± 0.01	0.1028	12.39
-1,-1	11.54 ± 0.06	12.88 ± 0.06	1.4101	11.29 ± 0.06	0.6611	2.13
0,-1	18.50 ± 0.10	19.92 ± 0.10	1.4479	18.10 ± 0.10	0.9654	1.50
1,-1	1.18 ± 0.01	2.00 ± 0.01	0.9244	1.15 ± 0.01	0.1352	6.84

Table 59 – Lorenzetti [180 GeV] - $\hat{E}$.

	Target	OF		NN$_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	Energy [GeV]	Energy [GeV]	RMSE	Energy [GeV]	RMSE	
-1,1	12.70 ± 0.07	14.12 ± 0.07	1.5013	12.41 ± 0.07	0.6747	2.23
0,1	19.39 ± 0.11	20.89 ± 0.11	1.5349	19.01 ± 0.10	0.9966	1.54
1,1	1.23 ± 0.01	2.09 ± 0.01	0.9652	1.20 ± 0.01	0.1472	6.56
-1,0	24.94 ± 0.01	26.84 ± 0.01	1.9108	24.32 ± 0.01	1.1231	1.70
0,0	40.03 ± 0.02	41.79 ± 0.02	1.7666	39.10 ± 0.02	1.6719	1.06
1,0	1.95 ± 0.01	3.28 ± 0.01	1.3307	1.88 ± 0.01	0.1117	11.91
-1,-1	12.75 ± 0.07	14.18 ± 0.07	1.5066	12.37 ± 0.07	0.7825	1.93
0,-1	19.51 ± 0.11	21.02 ± 0.11	1.5417	18.93 ± 0.10	1.2682	1.22
1,-1	1.24 ± 0.01	2.10 ± 0.01	0.9699	1.21 ± 0.01	0.1615	6.00

Table 60 – Lorenzetti $\hat{E}$ for E$_{imp}$ = 190 GeV.

	Target	OF		NN$_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	τ [ns]	τ [ns]	RMSE	τ [ns]	RMSE	
-1,1	13.88 ± 0.07	15.40 ± 0.08	1.5938	13.57 ± 0.07	0.7760	2.05
0,1	20.30 ± 0.11	21.89 ± 0.11	1.6252	19.79 ± 0.11	1.2126	1.34
1,1	1.29 ± 0.01	2.19 ± 0.01	1.0072	1.24 ± 0.01	0.2172	4.64
-1,0	27.20 ± 0.01	29.21 ± 0.01	2.0188	26.58 ± 0.01	1.0804	1.87
0,0	41.65 ± 0.02	43.52 ± 0.02	1.8795	40.77 ± 0.02	1.4865	1.26
1,0	2.04 ± 0.01	3.42 ± 0.01	1.3847	1.99 ± 0.01	0.1495	9.26
-1,-1	13.87 ± 0.07	15.39 ± 0.08	1.5954	13.51 ± 0.07	0.7534	2.12
0,-1	20.27 ± 0.11	21.87 ± 0.11	1.6306	19.77 ± 0.11	1.0737	1.52
1,-1	1.29 ± 0.01	2.19 ± 0.01	1.0086	1.23 ± 0.01	0.3940	2.56

Table 61 – Lorenzetti $\hat{E}$ for E$_{imp}$ = 200 GeV.

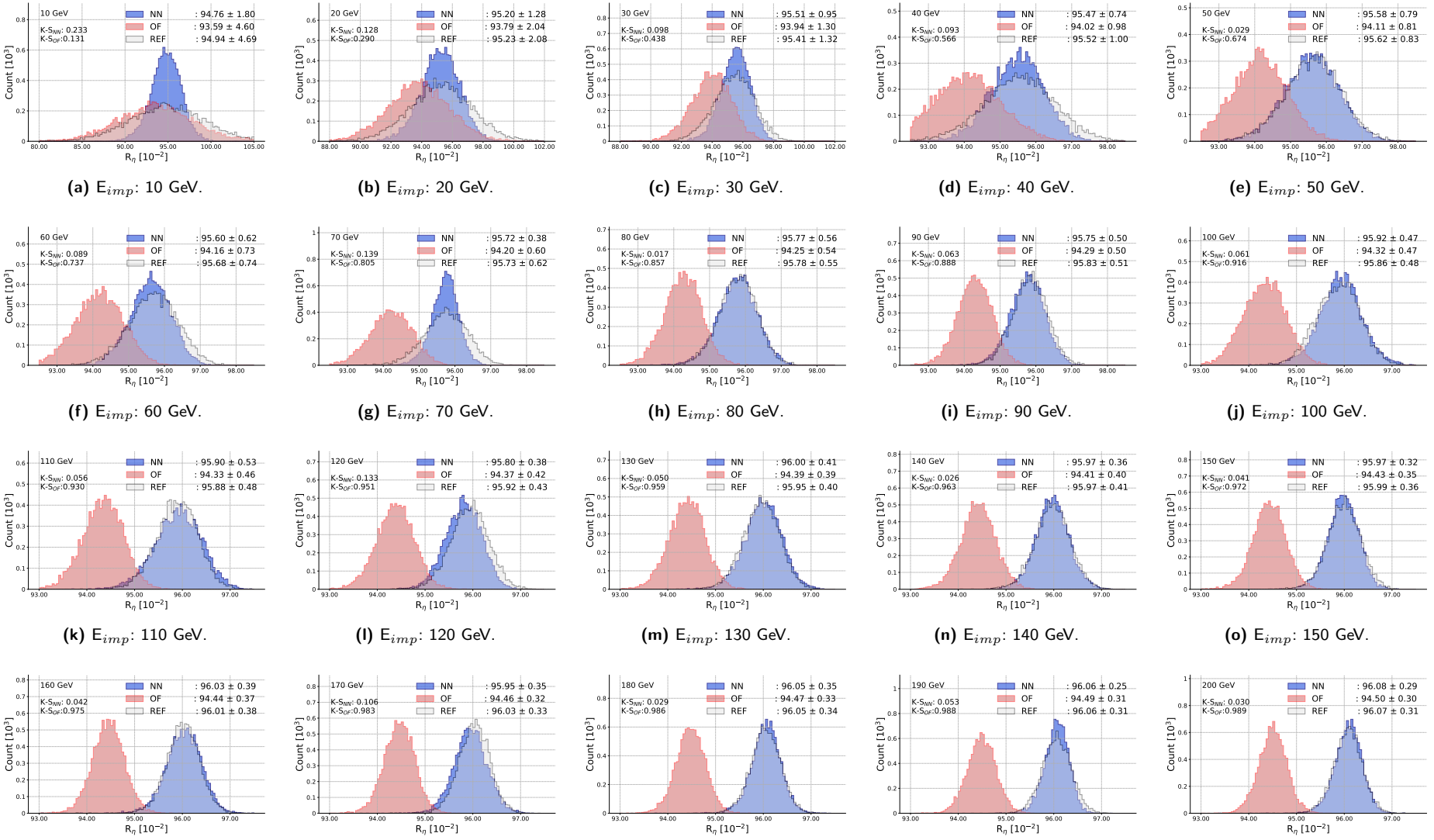
	Target	OF		NN$_{estimator}$		$\frac{RMSE_{OF}}{RMSE_{NN}}$
cell	τ [ns]	τ [ns]	RMSE	τ [ns]	RMSE	
-1,1	15.28 ± 0.08	16.89 ± 0.08	1.6905	14.73 ± 0.08	0.9639	1.75
0,1	21.33 ± 0.11	23.01 ± 0.12	1.7201	20.74 ± 0.11	1.1569	1.49
1,1	1.36 ± 0.01	2.29 ± 0.01	1.0500	1.32 ± 0.01	0.1740	6.04
-1,0	29.57 ± 0.01	31.69 ± 0.01	2.1258	28.93 ± 0.01	1.1206	1.90
0,0	43.16 ± 0.02	45.15 ± 0.02	1.9930	42.30 ± 0.02	1.5755	1.26
1,0	2.12 ± 0.01	3.55 ± 0.01	1.4359	2.11 ± 0.01	0.1336	10.75
-1,-1	14.86 ± 0.08	16.46 ± 0.08	1.6798	14.61 ± 0.08	0.7066	2.38
0,-1	20.76 ± 0.12	22.44 ± 0.12	1.7161	20.50 ± 0.11	0.8787	1.95
1,-1	1.34 ± 0.01	2.26 ± 0.01	1.0405	1.32 ± 0.01	0.1327	7.84

D.2.2.1 Summary of Shower Shapes results for 20 impact energy levels.

Table 62 – Summary for Kolmogorov-Smirnov (KS) test for all impact energies value for NN estimator and OF distributions with respect to the REF value for R_η, R_ϕ, and $W_{\eta 2}$ shower-shape. The NN estimator has a low difference, in comparison with the REF. Only in 10 GeV, the NN estimator produces a distribution worse than OF. The test result indicates that OF can not produce a distribution as the REF with XT effects for OF.

R_η										
Energy [GeV]	10	20	30	40	50	60	70	80	90	100
KS_{NN}	0.238	0.128	0.105	0.096	0.031	0.091	0.133	0.016	0.077	0.053
KS_{OF}	0.128	0.292	0.429	0.568	0.663	0.746	0.816	0.852	0.892	0.917
Energy [GeV]	110	120	130	140	150	160	170	180	190	200
KS_{NN}	0.045	0.138	0.056	0.018	0.039	0.046	0.119	0.040	0.052	0.047
KS_{OF}	0.928	0.944	0.956	0.967	0.970	0.976	0.980	0.987	0.988	0.989
R_ϕ										
Energy [GeV]	10	20	30	40	50	60	70	80	90	100
KS_{NN}	0.028	0.014	0.025	0.019	0.015	0.023	0.056	0.036	0.019	0.014
KS_{OF}	0.091	0.175	0.143	0.169	0.216	0.269	0.323	0.375	0.427	0.477
Energy [GeV]	110	120	130	140	150	160	170	180	190	200
KS_{NN}	0.017	0.044	0.027	0.039	0.044	0.026	0.030	0.016	0.025	0.011
KS_{OF}	0.516	0.565	0.612	0.647	0.680	0.714	0.750	0.774	0.796	0.819
$W_{\eta 2}$										
Energy [GeV]	10	20	30	40	50	60	70	80	90	100
KS_{NN}	0.139	0.238	0.335	0.391	0.381	0.377	0.051	0.386	0.343	0.063
KS_{OF}	0.328	0.614	0.796	0.896	0.946	0.972	0.983	0.990	0.995	0.995
Energy [GeV]	110	120	130	140	150	160	170	180	190	200
KS_{NN}	0.155	0.114	0.078	0.063	0.110	0.111	0.127	0.028	0.062	0.060
KS_{OF}	0.997	0.998	0.998	0.998	0.999	0.999	0.999	0.999	0.999	0.999

Figure 83 – R_η evolution over the 20 impact energy levels for the reference, OF reconstruction using samples with XT, and OF reconstruction after an NN estimator to regress samples reducing XT effects. Additional analysis is made using the Kolmogorov-Smirnov test to evaluate OF and NN distribution with respect to the REF value.



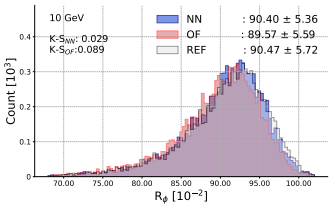
(a) E_{imp}: 10 GeV. **(b)** E_{imp}: 20 GeV. **(c)** E_{imp}: 30 GeV. **(d)** E_{imp}: 40 GeV. **(e)** E_{imp}: 50 GeV.

(f) E_{imp}: 60 GeV. **(g)** E_{imp}: 70 GeV. **(h)** E_{imp}: 80 GeV. **(i)** E_{imp}: 90 GeV. **(j)** E_{imp}: 100 GeV.

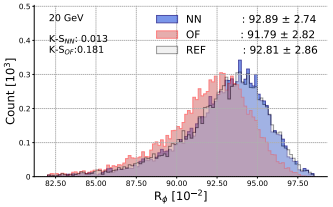
(k) E_{imp}: 110 GeV. **(l)** E_{imp}: 120 GeV. **(m)** E_{imp}: 130 GeV. **(n)** E_{imp}: 140 GeV. **(o)** E_{imp}: 150 GeV.

(p) E_{imp}: 160 GeV. **(q)** E_{imp}: 170 GeV. **(r)** E_{imp}: 180 GeV. **(s)** E_{imp}: 190 GeV. **(t)** E_{imp}: 200 GeV.

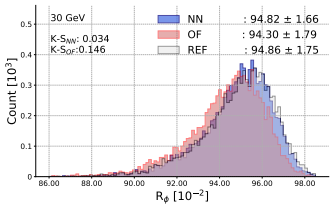
Figure 84 – R_ϕ evolution over the 20 impact energy levels for the reference, OF reconstruction using samples with XT, and OF reconstruction after an NN estimator to regress samples reducing XT effects. Additional analysis is made using the Kolmogorov-Smirnov test to evaluate OF and NN distribution with respect to the REF value.



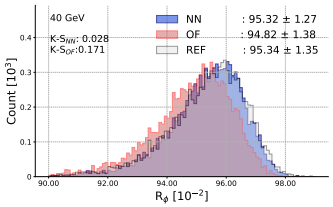
(a) E_{imp}: 10 GeV.



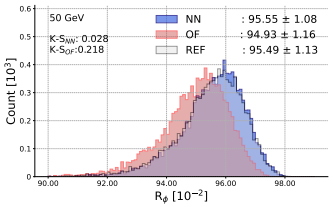
(b) E_{imp}: 20 GeV.



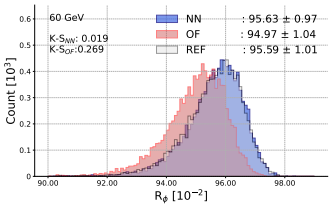
(c) E_{imp}: 30 GeV.



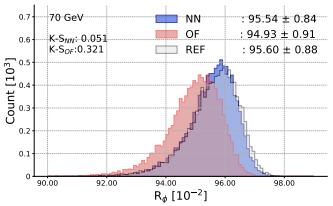
(d) E_{imp}: 40 GeV.



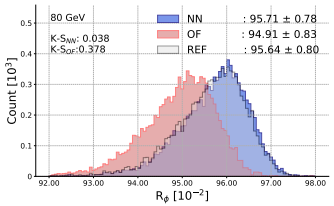
(e) E_{imp}: 50 GeV.



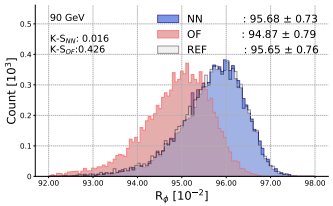
(f) E_{imp}: 60 GeV.



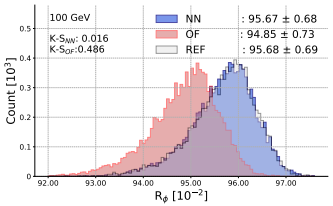
(g) E_{imp}: 70 GeV.



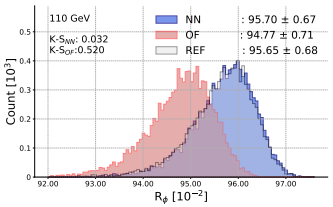
(h) E_{imp}: 80 GeV.



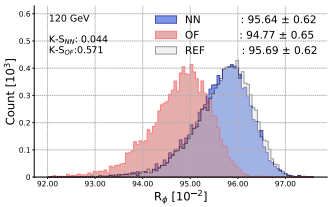
(i) E_{imp}: 90 GeV.



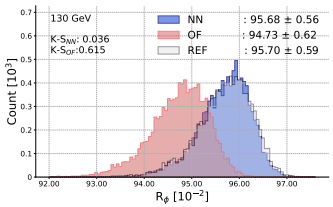
(j) E_{imp}: 100 GeV.



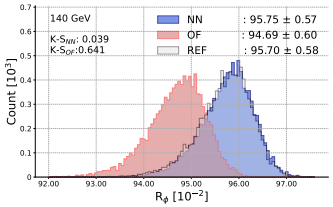
(k) E_{imp}: 110 GeV.



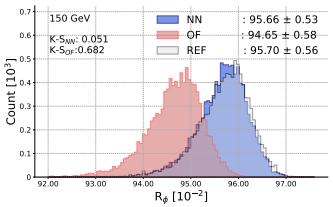
(l) E_{imp}: 120 GeV.



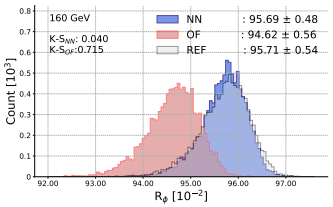
(m) E_{imp}: 130 GeV.



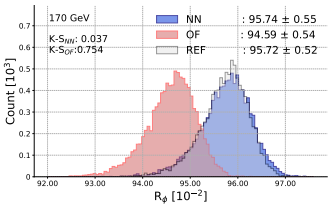
(n) E_{imp}: 140 GeV.



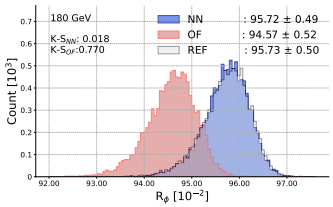
(o) E_{imp}: 150 GeV.



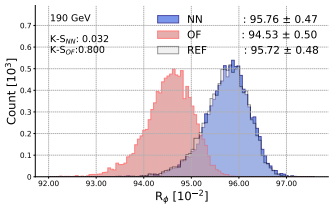
(p) E_{imp}: 160 GeV.



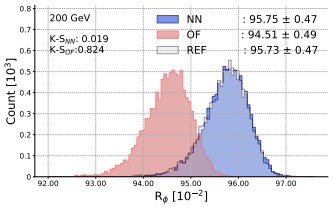
(q) E_{imp}: 170 GeV.



(r) E_{imp}: 180 GeV.

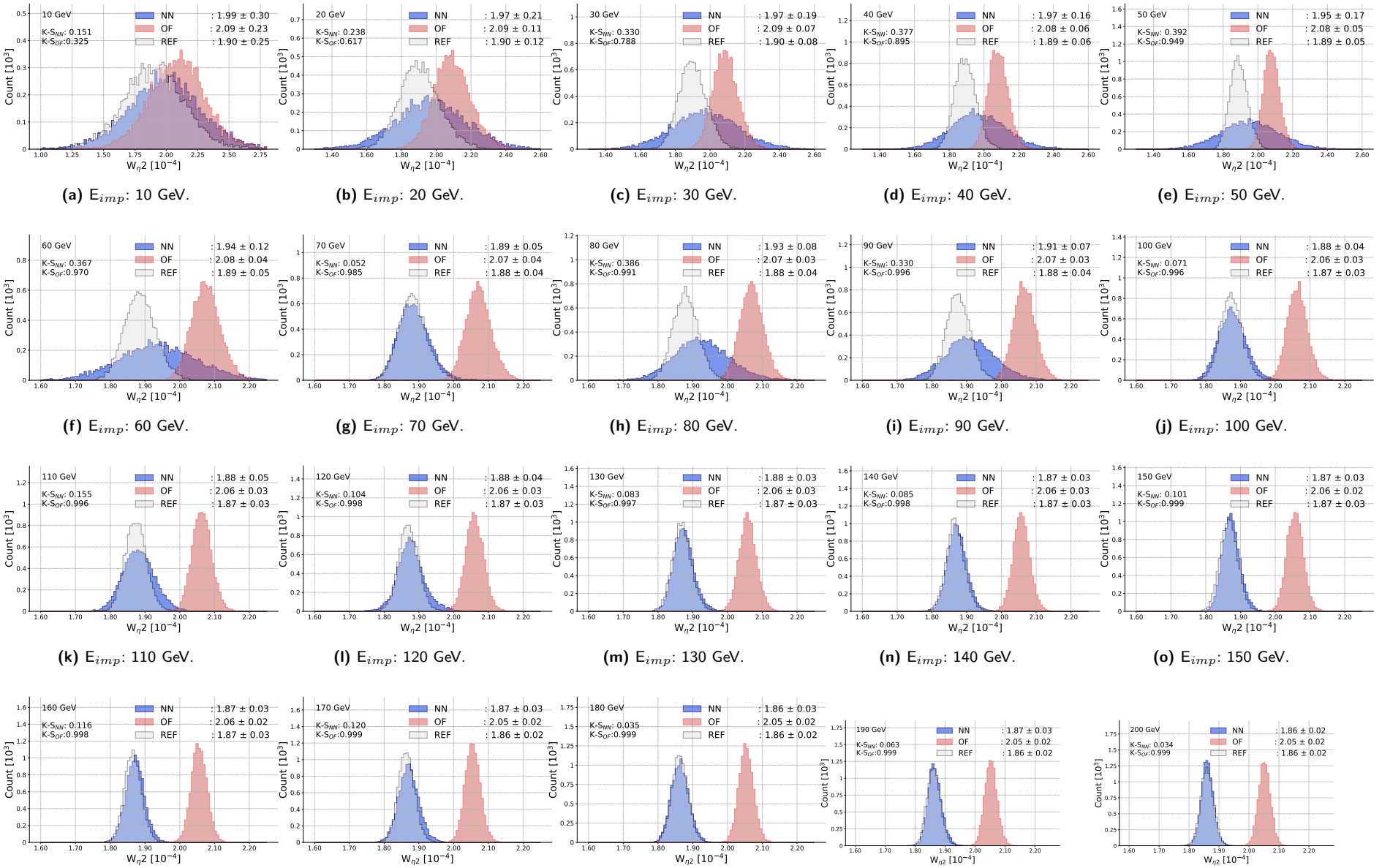


(s) E_{imp}: 190 GeV.



(t) E_{imp}: 200 GeV.

Figure 85 – $W_{\eta 2}$ evolution over the 20 impact energy levels for the reference, OF reconstruction using samples with XT, and OF reconstruction after an NN estimator to regress samples reducing XT effects. Additional analysis was made using the Kolmogorov-Smirnov test to evaluate OF and NN distribution with respect to the REF value.



(a) E_{imp}: 10 GeV. **(b)** E_{imp}: 20 GeV. **(c)** E_{imp}: 30 GeV. **(d)** E_{imp}: 40 GeV. **(e)** E_{imp}: 50 GeV.

(f) E_{imp}: 60 GeV. **(g)** E_{imp}: 70 GeV. **(h)** E_{imp}: 80 GeV. **(i)** E_{imp}: 90 GeV. **(j)** E_{imp}: 100 GeV.

(k) E_{imp}: 110 GeV. **(l)** E_{imp}: 120 GeV. **(m)** E_{imp}: 130 GeV. **(n)** E_{imp}: 140 GeV. **(o)** E_{imp}: 150 GeV.

(p) E_{imp}: 160 GeV. **(q)** E_{imp}: 170 GeV. **(r)** E_{imp}: 180 GeV. **(s)** E_{imp}: 190 GeV. **(t)** E_{imp}: 200 GeV.

APPENDIX E – Relevant Physics Information

E.1 Short historical facts in particle physics

The following are some facts concerning the discoveries in the area of relevant particle physics cited per year, gathered in Santos (2018) extracted from Rocha (2002, p. 363 – 372):

- 1897 - Discovery of the electron and its negative charge by Thomson;
- 1903 - Nobel Prize in Physics for the discovery of radioactivity: Antonie Henri Becquerel, Pierre Curie and Marie Sklodowska-Curie[57]
- 1904 - John Ambrose Fleming[58] receives the patent for the invention of the electronic valve, thermionic valve or Fleming (SMITH et al., 1940);
- 1917 - Nobel laureate in physics for the discovery of X-rays: Charles Glover Barkla[59];
- 1918 - Nobel Prize in Physics for the discovery of energy quanta: Max Plank;
- 1929 - Nobel Prize in Physics for the discovery of the wavelike nature of the electron: De Broglie[60];
- 1932 - Nobel Prize in Physics for the creation of Quantum Mechanics: Werner Heisenberg;
- 1933 - Nobel Prize in Physics for the discovery of new forms for atomic theory: Erwin Schrödinger and Paul Adrien Maurice Dirac[61];
- 1935 - Nobel Prize in Physics for the discovery of Neutron: Sir James Chadwick[62];
- 1936 - Nobel Prize in Physics was divided equally between Victor Franz Hess for his discovery of cosmic radiation and Carl David Anderson for his discovery of the positron[63]. This discovery has a connection with the introduction to anti-matter by Dirac and Schrödinger in 1993;
- 1939 - Nobel Prize in Physics to Ernest Lawrence for the creation of the Cyclotron accelerator in 1930 (LIVINGSTON; BLEWETT, 1962).
- 1949 - Nobel Prize in Physics for the discovery of Meson: Hideki Yukawa[64];
- 1954 - Nobel Prize in Physics for fundamental research on Quantum Mechanics: Max Born and Walther Bothe[65];
- 1956 - Nobel Prize in Physics for William Bradford Shockley, John Bardeen, and Walter Houser Brattain for semiconductor research and discovery of the transistor effect(Nobel Media AB 2020, 2020)[66];
- 1958 - Discovery of Antinêutron: Pavel Aleksejecic Cherenkov, Il'já Michajlovic Frank and Igor' Evgen'evic Tamm[67]

[57] Becquerel (1852 – 1908), French physicist; Pierre Curie (1859 – 1906), French physicist; Marie Curie (1867 – 1934), Polish physicist; first woman to win two Nobel Prizes; in 1911 - Nobel Prize in Chemistry for the discovery of Radio and Polonium;
[58] (1849-1945) – English electrical engineer and physicist.
[59] (1877 – 1944), an English physicist.
[60] Louis-Victor-Pierre-Raymond, 7^th Duke of Broglie (1892 – 1987), a French physicist.
[61] (1902 – 1984) a physicist the United Kingdom
[62] (1891 – 1974), an English physicist.
[63] Hess (1883 – 1964), Austrian physicist; Anderson (1905 – 1991), a physicist from the USA
[64] (1907 – 1981) a Japanese theoretical physicist.
[65] (1891 – 1957), a German physicist.
[66] Shockley (1910 – 1989) British physicist; Bardeen (1908 – 1991) e Brattain(1902 – 1987) American physicists.
[67] Cherenkov (1904 – 1990), a Russian physicist; Frank (1908 – 1990) a Russian physicist; Tamm (1895 – 1971), a Russian physicist.

- 1959 - Nobel Prize in Physics for the Discovery of the Antiproton: Emilio Gino Segrè and Owen Chamberlain[68];
- 1979 - Discovery of the gluon at the electron-positron collider PETRA of DESY (Deutsches Elektronen-Synchrotron)[69], Germany, in the late spring of 1979.
- 1983 - Tevatron, particle accelerator with collision energy up to 1.8 TeV is built at Fermilab;
- 1984 - Nobel Prize in Physics for the contributions that allowed the discovery of W and Z field particles: Carlo Rubbia and Simon van der Meer[70];
- 1989 - LEP (The Large Electron-Positron Collider) was officially inaugurated on November 13, 1989, in the presence of some 1500 guests, including Heads of State and Ministers from all of CERN's 14 Member States(CERN,);
- 1992 - Nobel Prize in Physics for the invention and development of particle detectors - the multi-wire proportional chamber: Russell A. Hulse[71];
- 1992 - HERA[72] was a particle accelerator at DESY in Hamburg. At HERA, electrons or positrons collided with protons with a center-of-mass energy of 318 GeV;
- 1995 - Nobel Prize in Physics for his contributions to the discovery of the lepton tau to Martin L. Perl and Frederick Reines for the detection of the neutrino[73];
- 2004 - Nobel Prize in Physics was awarded jointly to David J. Gross, H. David Politzer and Frank Wilczek for the discovery of asymptotic freedom in the theory of the strong interaction[74];
- 2008 - LHC goes into operation;
- 2010 - First collisions at 7 TeV on ATLAS (CERN, 2016);
- 2011 - Tevatron closes its activities on September 30;
- 2012 - July 14, ATLAS and CMS observe a particle consistent with the Higgs boson (ATLAS Collaboration, 2012b);
- 2013 - Nobel laureate in physics to Peter Higgs and François Englert[75] for the discovery of the Higgs mechanism (De ANGELIS; PIMENTA, 2018);
- 2018 - Higgs boson, second detection record. This time, associated with the quark top (ATLAS Collaboration, 2018).

E.2 List of some Particle Accelerators since 1961

The range from the first experiment until the last experiment is equal to 55 years. This shortlist presents 33 accelerators (Table 63 shows 27 LEP accelerators, and Table 64 shows 6 Hadron accelerators), a significant number, a new accelerator built and working in 1 year and seven months on average.

[68] Segrè (1905 – 1989) an Italian physicist, Chamberlain (1920 – 2006) a physicist from the USA.
[69] is a national research center in Germany that operates particle accelerators used to investigate the structure of matter.
[70] Rubbia (1930 –), an Italian physicist; Meer (1925 – 2011), Dutch physicist.
[71] (1950 –), a physicist from the USA.
[72] German: Hadron-Elektron-Ringanlage, English: Hadron-Electron Ring Accelerator
[73] Perl (1927 – 2014) a physicist of the USA, Reines (1918 – 1998) a physicist of the USA
[74] Wilckek (1951 –), Politzer (1941 –) and Gross (1941 –) physicists from the USA
[75] (1929 –) – a British physicist; Peter Englert (1932 –) – a Belgian physicist.

Table 63 – Lepton Colliders (all e^+e^- unless marked as e^-e^-) [a] DR: Double ring, SR: Single ring, LC: Linear collider. [b] Highest achieved. [c] Princeton-Stanford Colliding Beam Experiment. [d] Collisions achieved when operated in Orsay. [e] Using a detector with no solenoid field. [f] Design goal.

Location	Name (type [a])	Beam Energy [b] (GeV)	Luminosity [b] ($\mathrm{cm^{-2}\ s^{-1}}$)	Operation Period
Stanford/SLAC, USA	CBX [c] ($\mathrm{e^-e^-DR}$)	0.5	2×10^{28}	1963-1968
	SPEAR (SR)	4	1.2×10^{31}	1972-1988
	PEP (SR)	15	6×10^{31}	1980-1990
	SLC (LC)	49	2.5×10^{30}	1989-1998
	PEP-II (DR)	$9(\mathrm{e^-})3.1(\mathrm{e^+})$	1.2×10^{34}	1998-2008
Frascati, Italy	AdA	0.25	5×10^{25}[d]	1961–1964
	ADONE	1.5	6×10^{29}	1969–1993
	DAΦNE	0.51	2.4×10^{32} 4.5×10^{32}[e]	1999-present
BINP, Russia	VEP-1 ($\mathrm{e^-e^-DR}$)	0.16	5×10^{27}	1964-1968
	VEPP-2 (SR)	0.67	4×10^{25}	1966-1970
	VEPP-2M (SR)	0.7	$5 \times 10^{30}(@511)$	1974-2000
	VEPP-3 (SR)	1.55	2×10^{27}	1974-1975
	VEPP-4M (SR)	6	2×10^{31}	1984-present
	VEPP-2000 (SR)	1	5×10^{31}	2010-present
Cambridge, USA	CEA Bypass (SR)	3	8×10^{27}	1971-1973
Orsay, France	ACO (SR)	0.54	1×10^{29}	1965-1975
	DCI (SR)	1.8	1.7×10^{30}	1977-1985
DESY, Germany	DORIS (SR)	5.6	3.3×10^{31}	1973-1993
	PETRA (SR)	23.4	$2.4 \times 10^{31}(@1.75)$	1978-1986
CERN, Europe	LEP (SR)	104.5	1×10^{32}	1989-2000
Cornell, USA	CESR (SR)	5.5	1.3×10^{13}	1979-2008
KEK, Japan	TRISTAN (SR)	32	4.1×10^{31}	1986-1995
	KEKB (DR)	$8\,(\mathrm{e^-})3.5\,(\mathrm{e^+})$	2.1×10^{34}	1998-2010
	SuperKEKB (DR)	$7\,(\mathrm{e^-})4\,(\mathrm{e^+})$	8×10^{35}[f]	2016-present
IHEP, China	BEPC (SR)	2.4	$1 \times 10^{31}(@1.84)$	1988-2004
	BEPC II (DR)	2.47	$1 \times 10^{23}(@1.89)$	2009-present

Source: Chao et al. (2023, p. 14).

Table 64 – Hadron Colliders [a] DR: Double ring, SR: Single ring. [b] Highest achieved. [c] Energy per nucleon. [d] Design goal

Location	Name (type [a])	Beam Energy [b] (GeV)	Luminosity [b] ($\mathrm{cm^{-2}s^{-1}}$)	Operation Period
CERN, Europe	ISR (pp DR)	31.4	1.4×10^{32}	1971-1984
	$\mathrm{Sp\overline{p}S}$($\mathrm{p\overline{p}}$ SR)	315	6×10^{30}	1981-1991
	LHC (pp, ii, pi DR)	6500 (p) $2510/\mathrm{n(Pb)}$[c] 6500 (p) $2560/\mathrm{n(Pb)}$[c]	2.1×10^{34} 6.1×10^{27} 9×10^{29}	2009-present
Fermilab, USA	Tevatron ($\mathrm{p\overline{p}}$ SR)	980	4.3×10^{32}	1987-2011
DESY, Germany	HERA (ep DR)	27.5 (e^-) 920 (p)	5.3×10^{31}	1992-2007
Brookhaven, USA	RHIC (pp, ii DR)	250(p) $100/\mathrm{n(Au)}$[c]	2.5×10^{32} 1.6×10^{28}	2000-present
	EIC (ep, ei DR)[d]	$18(\mathrm{e^-})275(\mathrm{p})$ $18(\mathrm{e^-})110/\mathrm{n(Au)}$[c]	1×10^{34} 3×10^{30}	(under construction)

Source: Chao et al. (2023, p. 14).

Brazil participates in this development. In Brazil, SIRIUS, the light synchrotron accelerator, produces

light beams with specific wavelengths for biomedical, oil, and material studies. SIRIUS is located at the Laboratório Nacional de Luz Síncrotron (LNLS) and is part of National Center for Energy and Materials Research (CNPEM) under the Ministério de Ciência e Tecnologia (MCTI).

Four laboratories form the CNPEM: Laboratório Nacional de Biociências (LNBio); Centro de Ciência e Tecnologia do Bioetanol (CBTE); Laboratório Nacional de Nanotecnologia (LNNano); LNLS (CNPEM, 2014).

The first discussion about the importance of a synchrotron light laboratory in Brazil started in 1981, and in 1997 the laboratory was built and started its first experiments. Today SIRUS is the fourth generation of this kind of accelerator with an energy level equal to 3 GeV. The only one in Latin America this generation together another two, one in Sweden, MAX IV, the first light synchrotron source of the fourth generation, and the other in France, the Extremely Brilliant Source (EBS) located at the European Synchrotron Radiation Facility (ESRF) (ZORZETTO, 2018).

E.3 List of most common materials to build calorimeters

Following in the Table 65 it is possible to see some characteristics of materials that are used in the active and passive medium. A special detail for the X_0 difference between medium when we look at the density. For the passive medium, three has the highest density, W, Pb, and ^{238}U but the lowest values for X_0 (3,5, 5,6 e 3,2 [mm]) and E_c (8,0 7,4 e 6,8 [MeV]). These combined factors make these materials super interesting to use in calorimeters because the shower decay process starts with a low energy level and the particles travel a short distance inside the material until being absolved. (WIGMANS, 2017).

Table 65 – Some properties of typical materials used in calorimeters.

Passive Medium						Active Medium					
Material	Z	ρ (g cm^{-3})	E_c (MeV)	X_0 (mm)	R_M (mm)	Material	Z	ρ (g cm^{-3})	E_c (MeV)	X_0 (mm)	R_M (mm)
C	6	2.27	83	188.0	48.0	Si	14	2.3300	41	93.60	48
Al	13	2.70	43	89.0	44.0	Ar_{liq}	18	1.4000	37	140.00	80
Fe	26	7.87	22	17.6	16.9	Kr_{liq}	36	2.4100	18	47.00	55
Cu	29	8.96	20	14.3	15.2	Xe_{liq}	54	2.9500	12	24.00	42
Sn	50	7.31	12	12.1	21.6	Polystyrene	–	1.0320	94	424.00	96
W	74	19.30	8.0	3.5	9.3	Plexiglas	–	1.1800	86	344.00	85
Pb	82	11.30	7.4	5.6	16.0	Quartz	–	2.3200	51	117.00	49
Concrete	–	2.50	55	107.0	41.0	Air[a]		0.0012	87	304×10^3	74×10^3
Glass	–	2.23	51	127.0	53.0	Water	–	1.0000	83	361.00	92
Marble	–	2.93	56	96.0	36.0						

[a] at 20°, 1 atm.

Source: Wigmans (2017, Appendix B).

E.3.1 Shortlist of major homogeneous and sampling calorimeters

In Table 66, a shortlist of major homogeneous and sampling calorimeter characteristics is used in some large particle physics experiments. For the same X_0 value, homogeneous calorimeters have better resolution than sampling calorimeters. In LHC experiments both types are used one homogeneous for CMS and sampling calorimeters in ATLAS (LAr and Tile).

Table 66 – Typical resolution for Homogeneous and Sampling Calorimeters. E in GeV.

Homogeneous Calorimeter				
Experiment	Technology	$X_0[g\,cm^{-2}]$	Resolution	Year
Crystall Ball	NaI (Tl)	20	$2,7\%/E^{1/4}$	1983
BGO	$Bi_4Ge_3O_{12}$	22	$2\%/\sqrt{E} \oplus 0,7\%$	1993
KTeV	CsI[1]	27	$2\%/\sqrt{E} \oplus 0,45\%$	1996
BaBar	CsI	16–18	$2,3\%/E^{1/4}$	1999
BELLE	CsI	16	1,7% p. $E_\gamma > 3,5$ GeV	1998
BES III	CsI	15	2,5% p. $E_\gamma = 1$ GeV	2010
CMS	$PbWO_4$	25	$3\%/\sqrt{E} \oplus 0,5\% \oplus 0,2/E$	1997
ALICE	$PbWO_4$	19	$3,6\%/\sqrt{E} \oplus 1,2\%$	2008
OPAL	Pb crystal	20,5	$5\%/\sqrt{E}$	1990
NA48	Kr Liq	27	$3,2\%/\sqrt{E} \oplus 0,42\% \oplus 0,09/E$	1998
Sampling Calorimeter				
Experiment	Technology	$X_0[g\,cm^{-2}]$	Resolution	Year
ZEUS	Scintillator/U[b]	20–30	$18\%/\sqrt{E}$	1988
CDF	Scintillator/Pb	18	$13,5\%/\sqrt{E}$	1988
KLOE	Fiber Scint./Pb spaghetti	15	$5,7\%/\sqrt{E} \oplus 0,6\%$	1995
NA31	LAr/Pb	27	$7,5\%/\sqrt{E} \oplus 0,5\% \oplus 0,1/E$	1988
SLD	LAr/Pb	21	$8\%/\sqrt{E}$	1993
H1	LAr/Pb	20–30	$12\%/\sqrt{E} \oplus 1\%$	1998
DØ	LAr/U[b]	20,5	$16\%/\sqrt{E} \oplus 0,3\% \oplus 0,3/E$	1993
ATLAS	LAr/Pb accordion	25	$10\%/\sqrt{E} \oplus 0,7\% \oplus 0,4/E$[c]	1996

[a] Cesium Iodide.
[b] Depleted U.
[c] ATLAS Collaboration (2017).

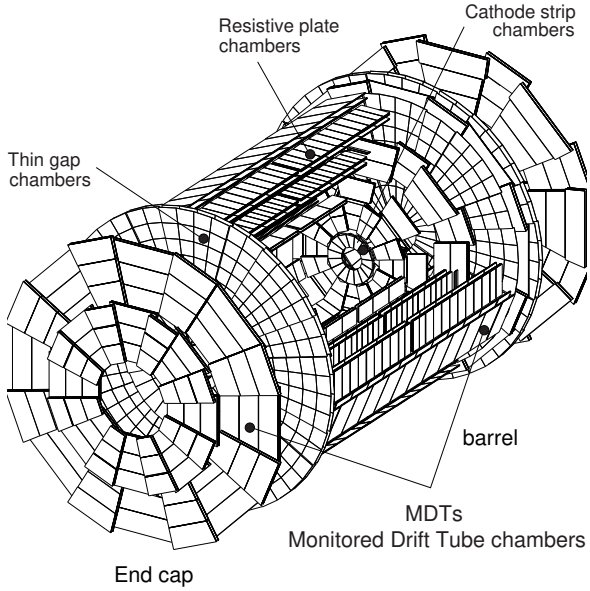
Source: Particle Data Group et al. (2020, Chap. 35.9).

E.4 ATLAS sub-detectors

Muon Spectrometer

In the outermost region, the muon spectrometer is located. There are four chambers designed for detection by magnetic deflection in toroidal superconductors with air nuclei, caused by traces of muons, see Figure 86 (ATLAS Collaboration, 2008). Table 67 shows details of the number of chambers, channels, regions in η, and the function of the chambers. Moreover, its resolution is determined by the Equation E.1.

Figure 86 – Illustration of the position of the set of components of the muon spectrometer in the outermost part of the ATLAS experiment.



Source: Drakoulakos et al. (1997).

$$\frac{\sigma_{p_T}}{p_T} \sim 3\% \quad \text{for } p_T = 100\,GeV,\ 10\% \text{ for } p_T = 1\,TeV \tag{E.1}$$

Table 67 – Muons spectrometer chambers, number of channels, and coverage region.

	Cover	Num. Chambers	Num. Channels	Function
MDT[1]	$\lvert\eta\rvert < 2,70$ $(\lvert\eta\rvert < 2,00)$	1150	354.000	Precision Tracking
CSC[2]	$2,00 < \lvert\eta\rvert < 2,70$	32	31.000	Precision Tracking
RPC[3]	$\lvert\eta\rvert < 1,05$	544	373.000	*Trigger*
TGC[4]	$1,05 < \lvert\eta\rvert < 2,70^{\star}$	3588	318.000	*Trigger*

[1] *Monitored Drift Tubes*; [2] *Cathode Strip Chambers*; [3] *Resistive Plate Chambers*; [4] *Thin Gap Chambers*. $^{\star}\lvert\eta\rvert < 2,4$ *trigger*.

Source: ATLAS Collaboration (2008).

APPENDIX F – A short review of Machine Learning ML usage in HEP

According to Haykin (2008):

> A neural network is a massively parallel distributed processor made up of simple processing units with a natural propensity for storing experiential knowledge and making it available for use.

Nowadays, ML is part of a vast number of systems and areas that are necessary to process and analyze big data for problems involving classification, regression, prediction, and data clustering. The ML area can be divided into three approaches:

- Supervised Learning[76] - In this approach, the data set is compounded by pairs of input targets. The inputs are features that characterize the signal generally in a vector format. The related outputs, for each input vector, are the target that the NN should use to evaluate the error between expected values and the NN output during the optimization step. The supervision is the fact to provide a target value for each set of input features. In this approach, NN could be designed for: classification, regression, prediction, or clusterization. This learning method was used in this research to treat the XT in the signal from LAr cells (HAYKIN, 2008).

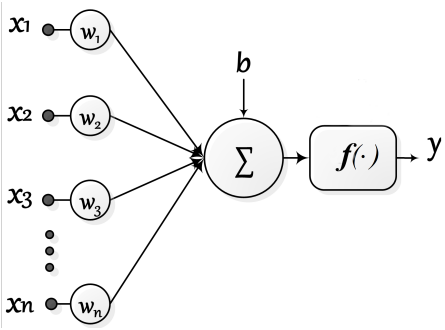
- Unsupervised Learning - For this training method, no one output related to the input features is provided to guide the optimization process to adjust weights in layers. However, it requires rich data information to provide relevant characteristics for the specific task. Most generally are clustering data and autoencoder applications (HAYKIN, 2008; KELLEHER, 2019).

- Reinforcement Learning - Here is a middle-term operation between supervised and unsupervised learning. The learning process is iterative from its current observation, using rewards or punishments at each optimization step. It is presently the most suitable method for online control tasks, such as robot control and game playing (KELLEHER, 2019).

Figure 87 and Equation F.1 show the diagram and the mathematical model used to implement a neuron used in the ML structures. These basic constituents are: the vector of input features **x**; the vector of weights **w** for each input feature x_i; b is a bias input; a block to summarize the weighted contributions plus bias; at the end, the activation function, which maps the inputs features in another space.

[76] Note: In this work, learning keeps the analogy and inspiration with the brain. Different from a brain, an artificial neural network is trained when achieving an optimal value for a specific metric according to a chosen cost function for a given and specific task, taking into account the features in its input.

Figure 87 – Mathematical model for a neuron used in ML structures, inspired on the biological neuron model where pairs x_if, w_i represents the information,**x**, from another node and level of importance respectively as weight, **w**.



$$y = f\left(\sum_{i=1}^{n} \boldsymbol{w}[i]\boldsymbol{x}[i] + b\right) \tag{F.1}$$

Another important element of ML algorithms using NN is associated with the choice of an activation function, which is a sensitive part of the optimization process for gradient descent. A good choice guarantees that the model works correctly and avoids problems with overfitting or non-trained nodes. If the choice is for non-linear functions, this allows NN to approximate any non-linearities between inputs and outputs (GLOROT; BENGIO, 2010; RAKITIANSKAIA; ENGELBRECHT, 2015). Following some activation functions are presented in Equation F.2, the *sigmoid* with $f(x) \in (0, 1)$, in Equation F.3 the *hyperbolic tangent* (tanh) with $f(x) \in (-1, 1)$, Equation F.4 is the *Rectified Linear Unit* (ReLU) with $f(x) \in [0, \infty)$ and Equation F.5 is *Parametric Rectified Linear Unit* (PReLU) a generalization of *Leaky Rectified Linear Unit* (Leaky-ReLU) where α has a fixed value equal to 0.01, with $f(x) \in (-\infty, \infty)$.

$$f(x) = \frac{1}{1 + e^{-x}} \tag{F.2}$$

$$f(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{F.3}$$

$$f(x) = \begin{cases} x, & \text{if } x > 0. \\ 0, & \text{if } x \leq 0. \end{cases} \tag{F.4}$$

$$f(x) = \begin{cases} x, & \text{if } x > 0. \\ \alpha x, & \text{if } x \leq 0 \; \alpha \in \mathbb{R}. \end{cases} \tag{F.5}$$

The Rectified Linear Unit (LReLU), in the Equation F.4 is a one-off most used in glsnn structures today because of the non-limited output, their sparsity, gradient propagation, and fast computation, until six-time than other functions. However, this function could present a known problem as *dying ReLU*. In certain circumstances, some nodes became inactive due to the node input being negative after the gradient optimization process. As a consequence, a weight or set of weights without actualization occurs in a hidden layer of NN structure (SZANDAŁA, 2021).

A variation on the ReLU function is shown in Equation F.5 with the addition of the α parameter (usually a small value $\in (0, 1)$) to adjust the function to allow negative values at node inputs avoiding the inactivation of a node if the weight optimization from gradient descent produces a negative value.

So, a NN structure is a set of this simple and basic node called *neuron*. A NN could perform the following tasks: pattern recognition, regression of models, classification, or even making a function approximation. In those tasks, special attention is given to the activation function, the core of a neuron. A correct choice for this function indicates the capability of a glsnn to work as is expected. Glorot and Bengio (2010) discuss the necessary attention to training a NN especially if NN has more than one HL.

Following is a description of each main task that a NN executes:

- Classification - Aims at separating two or more classes of interest within a data set based on the set of features provided at input. An example of this is the Neural Ringer, an event selector on the ATLAS experiment to identify electrons and reject jets from collisions (ANJOS et al., 2006; ATLAS Collaboration, 2012a). In Radovic et al. (2018) a set of examples using ML in the CMS and ATLAS experiments were used for the classification task using Boosted Decision Trees (BDT), and in LHC$_b$ a Recurrent Neural Networks (RNN) was used for jet classification.

- Clustering - According Haykin (2008): *Is a form of unsupervised learning whereby a set of observations is partitioned into natural groupings or clusters of patterns in such a way that the measure of similarity between any pair of observations assigned to each cluster minimizes a specified cost function*. Choudhary and Shukla (2021) presents a method to decompose a single problem to clustering data in a multi-sub-problem decomposing data in an ensemble to take efficient data clustering. In Arpaia et al. (2021), a set of ML techniques was evaluated for beam dynamics studies: Principal Component Analysis (PCA), K-Nearest Neighbor (KNN), MLP, and Support Vector Machine (SVM) trying to correct possible beam losses at the LHC. The goal of this work was to develop a system capable of determining the optimal parameters of the LHC setup to maximize the beam's intensity lifetime.

- Pattern recognition - In this application, an interesting pattern is a target for glsnn, and the structure should identify if the input features are equal to the expected pattern. In Bertacchi (2019) a Deep Neural Networks (DNN) based on Convolutional Neural Network (CNN) architecture was used to track jets in the CMS experiment at barrel region, using 2D-picture data from pixel detector. The aim is to produce sets of parameters to improve the jets seeds in the interested tracking region and reduce the time consumption of the standard method by iteration. As a result, the CNN shows a better performance than the standard seeding algorithm, reducing the fake rate up to 60% and seeding time by 85%.

- Prediction - For this task NN structures are used based on actual and past states to predict the following conditions, to estimate functions and process behavior. Pu et al. (2019) presents an analysis of some methods used to predict rockburst phenomena to avoid disasters in underground mines. ML structures have been used for this task using shallow, and this work shows that deep learning architectures are more effective than traditional ML methods to predict rockburst. In Pang (2021) describe an ML model to extract characteristics in heavy-ion collisions to predict parameters of heavy ions that are used in studies of properties of quark-gluon plasm.

- Regression - Statistical approach to get a representative approximation of the model under study based on the dependence and relationship among system variables. Kumar and Sahu (2021) presents an interesting analysis of nine regression ML structures to calibrate a low-cost PM$_{2.5}$ sensor to measure air pollution. The linear methods show poor results compared to ML methods for the calibration task. In Damgov and Litov (2002) a ML model to reconstruct energy deposited in CMS calorimeter, achieved an improvement in energy resolution and linearity In Larkoski, Moult and Nachman (2020) describes the jet models and ML applications on Jet identification, regression, and simulations.

Moving forward, nowadays, research is discussing the DNN - Deep Neural Networks. A dense structure based on NN, with at least two hidden layers, provides computers the power to separate and classify objects

by their characteristics or identify tendencies and estimate models to describe a specific process where the analytical description was not possible to achieve.

> ... A person's everyday life requires an immense amount of knowledge about the world. Much of this knowledge is subjective and intuitive and therefore difficult to articulate formally. Computers need to capture this same knowledge to behave intelligently. One of the critical challenges in artificial intelligence is how to get this informal knowledge into a computer (GOODFELLOW; BENGIO; COURVILLE, 2016).

So, since the first idea for a neuron, the basic processing unit from a neural network grows in capability, variety, and complexity. It is difficult to describe and present today the richness of the architectures and domains of neural networks where it is applied. Giray (2021) offers a very rich, but not definitive, engineering perspective for ML systems, discussing mainly the software engineering designed explicitly for ML, which involves the design, developing, and maintaining ML systems which include tools, libraries, and versions. This status is not definitive because they focus on the software engineering challenges of engineering ML.

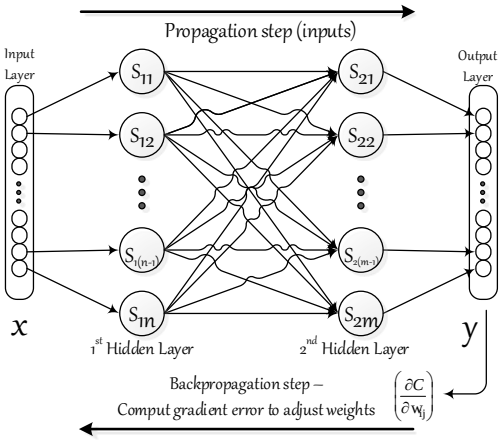
According to Skansi (2018) the American Mathematical Society (American Mathematical Society (AMS)) and the Association for Computing Machinery (Association for Computing Machinery (ACM)) are responsible for providing a formal classification of Artificial Intelligence (AI). The AMS and ACM divide AI into different subfields presented in the Table 68.

Table 68 – AI usage classification according AMS and ACM societies.

AMS	ACM
General	Natural language processing
Learning and adaptive systems	knowledge representation and reasoning
Pattern recognition and speech recognition	planning and scheduling
Theorem proving	search methodologies
Problem-solving	control methods
Logic in artificial intelligence	philosophical/theoretical foundations of AI
Knowledge representation	distributed artificial intelligence and computer vision
Languages and software systems	Machine learning is a parallel category to AI not subordinated to it
Reasoning under uncertainty	
Robotics	
Agent technology	
Machine vision and scene understanding	
Natural language processing	

The most known method used in the ML field is the multilayer perceptron using the backpropagation algorithm, see Figure 88, to adjust their weights. This concept is still alive in new approaches like graph nets.

Figure 88 – Block diagram for classical feedforward multilayer perceptron with the backpropagation algorithm to optimize their weights between layers. Here is a didactic example using two hidden layers to show a fully connected structure. In the first step, propagation, the input information goes into structure until output. The second step, backpropagation, calculates the error between the target (supervised learning approach) and predicted value to update weights from the last layer to the first until satisfying a stop criterion.



In this algorithm, there are two steps to perform the training: In the first, the input data is presented on the input layer, and the output is determined, and in the second step, the error evaluated between the predicted outcome and target is used to adjust the weights on the backpropagation step[77] until the achievement of a given stop criterion.

The optimization process uses gradient information, a first-order method to identify the direction where the function increases, to update weights in the layers during the backpropagation step according to Equation F.6. Usually, the technique focuses on the minimal value to find the best adjustment for the model. Due to this is taken a negative gradient value (GOODFELLOW; BENGIO; COURVILLE, 2016).

$$\nabla_\theta = \begin{bmatrix} \frac{\partial \theta}{\partial w_1} \\ \vdots \\ \frac{\partial \theta}{\partial w_n} \end{bmatrix} \tag{F.6}$$

Optimization algorithms are used for training models to achieve specific metrics from designed NN applications, minimize MSE for instance. To do this, let consider a set of input features x and targets y, some *loss* function $\mathrm{L}(\boldsymbol{x}^{(\mathrm{t})}, \boldsymbol{y}^{(\mathrm{t})})$, $\boldsymbol{\theta}$ represents a model parameter to update by the gradient according to Equation F.7. In equation, t is an instant where the feature and the respective target are taken, $\nabla_{\boldsymbol{\theta}}$ is the gradient of $\boldsymbol{\theta}$, and α is the *learning rate*, a hyperparameter to control size of gradient step into minimal direction (GOODFELLOW; BENGIO; COURVILLE, 2016).

$$\boldsymbol{\theta} \leftarrow \boldsymbol{\theta} + \alpha \nabla_{\boldsymbol{\theta}} \sum_t L\left(f\left(x^{(t)}; \boldsymbol{\theta}\right), \boldsymbol{y}^{(t)}; \boldsymbol{\theta}\right) \tag{F.7}$$

So gradient descent is very useful to optimize ML models. Due to this, there is some variation in this algorithm's application to avoid the minimal local problem when the ML model stops the training before achieving the minimum error between predicted values and target values. Goodfellow, Bengio and Courville (2016) discusses those variations on the gradient descent approaches to achieve the best optimization as presented in chapter 8.

[77] This step computes $\frac{\partial \theta}{\partial w_{ij}}$ and uses it to adjust weights from layers.

Considering those facts, ML has an option evaluated to solve some problems that an analytical formulation is hard to develop. In the HEP area, the ML applications are increasing due to their capabilities. Such the usage of ML in HEP is an old story, for instance. In Seixas et al. (1996) a glsnn using a preprocessing layer taking into account a parameter R_p[78], was used to increase electron/jet identification in the spacall detector (SIMON, 1992). So, this application resulted in 99% efficiency in electron identification and reduced jets error identification.

ATLAS experimental framework was also used to develop an ensemble of NN MLPs used in its high-level trigger to select events with e/γ and reduce the jet's background for posterior physical analysis. This ensemble of NN classifies electrons/jets using the calorimetry information at the cell level. The most energetic cell defines the center of one hundred concentric rings, which are distributed across the seven ATLAS calorimeter layers. These rings perform a topological description of the energy deposited by particles in cells. This approach produces a more efficient event selection and separation than the method, which uses a few physics variables associated with the previously mentioned particle shower. This approach was developed and updated to be one of the standard electron/jet discriminators in the ATLAS experiment. The previous method uses a few physics parameters to classify electrons and jets and allowed ATLAS to cope with the conditions at LHC Run2 without increasing too much the P_T threshold (ANJOS et al., 2006; ATLAS Collaboration, 2020c).

In Denby (1999) is presented a small survey of NN applications on HEP over the past 10 years since the late 1980s. At Fermilab, a glsnn for pattern recognition did sophisticated and successful exploitation of the analog neural network trigger to expand the coincidence circuit to more complex hardware to perform a very rapid device to be implemented in level 2. In the Apparatus for LEP Physics (ALEPH) experiment at LEP, a NN was used for the first time to track and reconstruct with higher efficiency concerning the conventional algorithm. Besides other applications on the trigger, there are applications in Fermilab Tevatron, Hera accelerator, CHOOZ[79] experiment, CMS experiment, Detector with Lepton, Photon and Hadron Identification (DELPHI) experiment from LEP used ML techniques in both operations mode, online and offline.

In Damgov and Litov (2002) NNs are used to recognize the type of particle that initialize the shower and to reconstruct the energy deposited in the calorimeter system of the CMS experiment made out of scintillating lead tungstate crystals. In this application, glsnn classifies hadron showers into four classes, electromagnetic interactions (mainly) ($e^{\mp}/\gamma$), hadrons from strong interactions (mostly), jets, and muons. The NN efficiently recognizes the initializing particle shower. Even if the shower is misidentified, the energy is correctly reconstructed.

An interesting point to mention is that in HEP many types of ML structures are used: BDT, MLP, CNN, RNN, graph-nets. All those structures are in a shallow or deep mode increasing particle detection, management data stored monitored by CMS and LHCb predicting which data will be more accessed to optimize the system and particle classification. The ML application already has an impressive increase in sensitivity for particle detection like Table 69 from Radovic et al. (2018) shows.

[78] This parameter is associated with the mapping the radius distance referred to the cell of maximum energy deposition.
[79] Is a reactor experiment based on neutrino oscillation technique to perform high sensitive neutrino detection with very low masses in Chooz, France (DE KERRET et al., 1993).

Table 69 – Effect of machine learning on the discovery and study of the Higgs boson. V denotes a W or Z boson, γ denotes a photon, and *b* a beauty quark. For each analysis, the sensitivity without and with machine learning is given in the P values and the equivalent number of Gaussian standard deviations σ. P ratio indicates how much the sensitivity is increased.

Analysis	Years of data collection	Sensitivity without machine learning	Sensitivity with machine learning	Ratio of P values	Additional data required
CMS $H \to \gamma\gamma$	2011-2012	2.2 σ P=0.014	2.7 σ P=0.0035	4.0	51%
ATLAS $H \to \tau^+\tau^-$	2011-2012	2.5 σ P=0.0062	3.4 σ P=0.00034	18	85%
ATLAS $VH \to bb$	2011-2012	1.9 σ P=0.029	2.5 σ, P=0.0062	4.7	73%
ATLAS $VH \to bb$	2015-2016	2.8 σ P=0.0026	3.0 σ P=0.00135	1.9	15%
CMS $VH \to bb$	2011-2012	1.4 σ P=0.081	2.1 σ P=0.018	4.5	125%

Source: Radovic et al. (2018)

What is very interesting is to observe that all parameters increased using the ML approaches and how much data information is necessary without ML. This summary indicates that ML is a useful technique for HEP applications.

HEP experiments are developed in the context of massive data production at the order of TB/s thanks to the energy level used and the higher collision rate. This context challenges the identification of interesting information among the vast available data by collisions.. So, ML has been a way of extracting knowledge and reducing offline storage structures. Therrien et al. (2019) discusses the challenge that involves HEP and signals processing where the data production sometimes is much faster than the acquisition system can collect and store the interested data at Stanford Linear Accelerator Center (SLAC) National Accelerator Laboratory. As a solution, a NN on Field Programmable Gate Array (FPGA) was evaluated as a proof of concept that a NN can select the interesting data and reduce data volumes at the source for the increasing data production at the SLAC.

In (LARKOSKI; MOULT; NACHMAN, 2020) the ML approach is used to help in the jet description. This work is a review of the state-of-the-art report of jet substructure at the LHC—such jet structure arrives from quarks and gluons, which interact via strong force dominantly. In this work, ML characteristics and capabilities are used for the event-level classification, and this review shows the improvement of the Jet substructures, which involves: identification and classification, model regression for calibration of tracking detectors, and calorimeters to measure momenta and energy, and fast simulations to a better jet comprehension.

The good ML feature extraction capabilities are explored in Arpaia et al. (2021), a set of ML techniques was evaluated for beam dynamics studies: PCA, KNN, MLP, and SVM trying to correct possible beam losses at the LHC. The goal of this work was to develop a system capable of determining the optimal parameters of the LHC setup to maximize the beam's intensity lifetime. In Pang (2021) a GAN, and VAE are used to extract features to describe heavy-ion collisions in a multistep process complex and hybrid physics model. Two ML structures, GAN or VAE are very successful in generating models that can learn statistical information from data.

Thanks to its rich set of techniques for processing data at the CERN, ML motivates the creation of the Inter-experimental Machine Learning working group. This group provides a forum for discussion in the ML community at the LHC. In this forum are regular meetings open for all interested, reviews on the latest tools and software in ML, and development of interfaces to HEP software. Other activities involve education and training

on core HEP-ML, coordination of the creation and maintenance of common benchmarks for comparison in different ML methods and tools using public information available in the forum dedicated area (CERN, 2016).