Universidade do Estado do Rio de Janeiro
Centro de Tecnologia e Ciências
Instituto Politécnico

Badilé Miranda Insali

Modelos lineares para estimação da energia em tempo real do calorímetro de telhas do ATLAS

Nova Friburgo
2025

Badilé Miranda Insali

Modelos lineares para estimação da energia em tempo real do calorímetro de telhas do ATLAS

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-Graduação em Modelagem Computacional, da Universidade do Estado do Rio de Janeiro.

Orientador: Prof. Dr. Gustavo Barbosa Libotte
Orientador: Prof. Dr. Bernardo Sotto-Maior Peralva

Nova Friburgo
2025

CATALOGAÇÃO NA FONTE
UERJ / REDE SIRIUS / BIBLIOTECA CTC/E

I59 Insali, Badilé Miranda.
Modelos lineares para estimação da energia em tempo real do calorímetro de telhas do ATLAS / Badilé Miranda Insali. - 2025.
86 f. : il.

Orientadores: Gustavo Barbosa Libotte e Bernardo Sotto-Maior Peralva.

Dissertação (mestrado) – Universidade do Estado do Rio de Janeiro, Instituto Politécnico.

1. Processamento de sinais – Métodos de simulação – Teses. 2. Estimativa de parâmetros – Teses. 3. Partículas (Física nuclear) – Teses. I. Libotte, Gustavo Barbosa. II. Peralva, Bernardo Sotto-Maior. III. Universidade do Estado do Rio de Janeiro. Instituto Politécnico. III. Título.

CDU 621.391:539.12

Bibliotecária Fernanda Souza Cruz CRB7/7361

____________________________ ____________________
Assinatura Data

Badilé Miranda Insali

Modelos lineares para estimação da energia em tempo real do calorímetro de telhas do ATLAS

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-Graduação em Modelagem Computacional, da Universidade do Estado do Rio de Janeiro.

Aprovada em 19 de fevereiro de 2025.

Banca Examinadora:

Prof. Dr. Gustavo Barbosa Libotte (Orientador)
Instituto Politécnico – UERJ

Prof. Dr. Bernardo Sotto-Maior Peralva (Orientador)
Instituto Politécnico – UERJ

Prof. Dr. Roberto Pinheiro Domingos
Instituto Politécnico – UERJ

Prof. Dr. Luciano Manhães de Andrade Filho
Universidade Federal de Juiz de Fora

Nova Friburgo

2025

DEDICATÓRIA

Dedico este trabalho aos meus amigos, colegas e professores que contribuíram para esse feito, especialmente para o nosso grupo de pesquisa e para minha família.

AGRADECIMENTOS

Gostaria de expressar minha profunda gratidão à UERJ e a todos os servidores do Instituto Politécnico, especialmente aos professores do programa de pós-graduação em Modelagem Computacional, que contribuíram de maneira significativa para minha formação acadêmica.

Meus sinceros agradecimentos vão aos meus excelentes orientadores, Prof. Dr. Gustavo Barbosa Libotte e Prof. Dr. Bernardo Sotto-Maior Peralva, por aceitarem o desafio de me orientar neste trabalho. Eles sempre se mostraram disponíveis para me auxiliar sempre que necessário.

Agradeço também aos professores que compuseram esta ilustre banca examinadora, em especial aos professores Dr. Roberto Pinheiro Domingos e Dr. Luciano Manhães de Andrade Filho, por aceitarem o nosso convite para participar deste momento tão importante.

Sou profundamente grato a todos os meus familiares, especialmente aos meus pais, Sandra da Silva Insali e José Carlos Miranda, que sempre se dedicaram incansavelmente para garantir que meus irmãos e eu tivéssemos uma educação de qualidade.

Aos meus amigos e colegas, agradeço por terem contribuído nesta jornada, especialmente àqueles que sempre acreditaram em mim e me inspiraram a seguir em frente, firme e forte, em busca dos meus objetivos.

Gostaria de agradecer à Pró-Reitoria de Políticas e Assistência Estudantil (PR4-UERJ) pelo apoio.

Por fim, agradeço à Coordenação de aperfeiçoamento de Pessoal de Nível Superior - CAPES pela bolsa de Mestrado concedida.

O conhecimento é como um jardim: se não for cultivado, não pode ser colhido!

Provérbio africano

RESUMO

INSALI, B. M. **Modelos lineares para estimação da energia em tempo real do calorímetro de telhas do ATLAS**. 2025. 86 f. Dissertação (Mestrado em Modelagem Computacional) – Instituto Politécnico, Universidade do Estado do Rio de Janeiro, Nova Friburgo, 2025.

O LHC (Large Hadron Collider) é o maior acelerador de partículas do mundo, e fornece um ambiente de inúmeros desafios para diversas áreas do conhecimento. Tendo em vista o período de atualização do LHC para alta luminosidade (HL-LHC), com término previsto para 2029, uma nova arquitetura eletrônica de leitura que permitirá a operação no modo free-running de processamento digital de sinais será implementada. Neste contexto de operação do HL-LHC, espera-se um aumento considerável no número de interações por colisão, o que traz desafios significativos para os algoritmos de processamento, especialmente nos calorímetros, onde o empilhamento de sinais pode comprometer a resolução energética. O calorímetro de telhas do experimento ATLAS no LHC, atualmente utilizado para identificar e medir partículas geradas nas colisões de alta energia entre prótons, emprega o método do Filtro Ótimo 2, que busca minimizar a variância do ruído eletrônico. No entanto, em cenários de alta luminosidade, onde os sinais de resposta do calorímetro se sobrepõem devido à alta ocupação, esse método pode introduzir erros sistemáticos, dificultando a estimativa precisa da amplitude do sinal. Diante desse cenário, este estudo busca explorar e comparar o desempenho de diferentes métodos lineares de estimativa de sinal: filtro ótimo 2, filtro ótimo 1, e mínimos quadrados com e sem restrição, considerando tanto condições com pedestal fixo quanto condições com pedestal variável. Os resultados mostram que o método dos mínimos quadrados com restrição se apresenta como melhor alternativa em condições de variação do pedestal.

Palavras-chave: modelos lineares; empilhamento de sinais; reconstrução do sinal; experimento ATLAS.

ABSTRACT

INSALI, B. M. **Linear models for real-time energy estimation of the ATLAS tile calorimeter**. 2025. 86 f. Dissertação (Mestrado em Modelagem Computacional) – Instituto Politécnico, Universidade do Estado do Rio de Janeiro, Nova Friburgo, 2025.

The LHC (Large Hadron Collider) is the largest particle accelerator in the world, providing an environment full of challenges for various fields of knowledge. Given the LHC's ongoing upgrade to high luminosity (HL-LHC), expected to be completed by 2029, a new reading electronic architecture will be implemented to enable operation in free-running mode for digital signal processing. In this HL-LHC operating context, a considerable increase in the number of interactions per collision is expected, which brings significant challenges to processing algorithms, especially in the calorimeters, where signal stacking can compromise energy resolution. The tile calorimeter of the ATLAS experiment at the LHC, currently used to identify and measure particles generated in high-energy proton collision events, employs the Optimal Filter 2 method, which aims to minimize the variance of electronic noise. However, in high luminosity scenarios, where calorimeter response signals overlap due to high occupancy, this method may introduce systematic errors, making it difficult to accurately estimate the signal amplitude. In this context, this study aims to explore and compare the performance of different linear signal estimation methods: Optimal Filter 2, Optimal Filter 1, and least squares with and without constraints, considering both fixed pedestal and variable pedestal conditions. The results show that the least squares method with constraints is the best alternative under pedestal variation conditions.

Keywords: linear models; signal pileup; signal reconstruction; ATLAS experiment.

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE ABREVIATURAS E SIGLAS

ADC	Analog-to-Digital Converter
ALICE	A Large Ion Collider Experiment
ATLAS	A Toroidal LHC ApparatuS
CERN	Conseil Européen pour la Recherche Nucléaire
CLS	Constrained Least Squares
CMS	Compact Muon Solenoid
DSP	Digital Signal Processor
LHC	Large Hadron Collider
LHCb	Large Hadron Collider beauty
LH-LHC	High-Luminosity LHC
LS	Least Squares
LSE	Least Squares Estimation
OF	Optimal Filter
TileCal	Tile Calorimeter

LISTA DE SÍMBOLOS

$\boldsymbol{A}$	Amplitude do sinal
$\widehat{\boldsymbol{A}}$	Amplitude estimada usando método de mínimos quadrados (LS e CLS)
$\widehat{\boldsymbol{A}}_{\boldsymbol{OF}}$	Estimativa da amplitude usando o Filtro Ótimo (OF)
$C[k,j]$	Elemento da matriz de covariância do ruído
$\boldsymbol{C}$	Matriz de covariância do ruído
e	Erro da estimativa
$\mathbb{E}\{\cdot\}$	Valor esperado
$\boldsymbol{H}$	Matriz das amostras
$\boldsymbol{H}$	Matriz de observação
$\boldsymbol{J(\theta)}$	Função de custo dos mínimos quadrados
$\lambda, \epsilon, \kappa$	Multiplicadores de Lagrange
$\mathcal{L}(\boldsymbol{w}, \lambda, \epsilon, \kappa)$	Função Lagrangiana associada à minimização da variância
M	Número total de amostras
N	Número de janelas de dados
$n[k]$	Ruído no tempo k
p	Pedestal
$s[k]$	Amostra no tempo k
$\boldsymbol{s}$	Vetor de sinais estimados
$x[k]$	Amostra do sinal no tempo k
$\boldsymbol{x}$	Amplitude real
$\boldsymbol{X}$	Conjunto de valores de amplitude
$g[k]$	Pulso de referência no tempo k
$g'[k]$	Derivada do pulso de referência no tempo k
$\boldsymbol{w}$	Pesos do filtro usando os métodos do filtro ótimo
$\boldsymbol{w[k]}$	Pesos do filtro no tempo k
$\boldsymbol{\theta}$	Pesos do filtro usando o método LS
$\boldsymbol{\theta_r}$	Pesos do filtro usando o método CLS

SUMÁRIO

INTRODUÇÃO

Se não posso estimular sonhos impossíveis, não devo negar o direito de sonhar com quem sonha.

Paulo Freire

O processamento digital de sinais DSP (*Digital Signal Processing*, em inglês) está relacionado aos aspectos teóricos e práticos de representar sinais portadores de informação em forma digital e ao uso de computadores ou hardware digital de propósito específico, seja para extrair essa informação ou para transformar os sinais de maneiras úteis. Áreas onde o processamento digital de sinais teve um impacto significativo incluem telecomunicações, comunicações homem-máquina, engenharia de computação, aplicações multimídia, tecnologia médica, radar e sonar, análise de dados sísmicos e sensoriamento remoto, para citar apenas algumas (Madisetti; Williams, 1999).

E um dos mais importantes problemas no processamento de sinais é o problema de estimar um sinal. Em muitas aplicações, o sinal desejado não está disponível ou não é observável diretamente. Em vez disso, o sinal observável é uma versão degradada ou distorcida do sinal original. O problema da estimação do sinal é recuperar, da melhor maneira possível, o sinal desejado a partir de sua réplica degradada (Orfanidis, 1988).

Este problema de estimação de sinal é especialmente relevante na área de física experimental de altas energias, principalmente na física de partículas, onde os sinais são frequentemente complexos e difíceis de observar diretamente devido às condições extremas e à natureza dos experimentos. Um exemplo desse tipo de problema, que estamos interessados em estudar nesta dissertação, é a reconstrução da energia da amplitude de sinais resultantes de colisões de partículas no experimento ATLAS (*A Toroidal LHC Apparatus*).

O Large Hadron Collider (LHC) é um acelerador e colisor de hádrons que utiliza tecnologia de supercondutores e é composto por dois anéis. Ele está instalado em um túnel com uma extensão total de 26,7 quilômetros. Este túnel foi originalmente construído entre 1984 e 1989 para acomodar a máquina LEP (*Large Electron-Positron Collider*) do Centro Europeu de Pesquisa Nuclear (CERN), localizado em Genebra, Suíça (Evans; Bryant, 2008).

O LHC é projetado para acelerar partículas a velocidades próximas à velocidade da luz e colidi-las, permitindo a investigação de fenômenos fundamentais da física. Na sequência, discutiremos em detalhes o funcionamento do LHC. É importante destacar que os sinais gerados pelas colisões de partículas podem ser contaminados por ruído. Portanto, será necessário implementar técnicas para filtrar e minimizar o impacto desse ruído nas medições.

O objetivo desta dissertação é analisar e comparar diferentes modelos lineares de estimação de sinal, entre os quais o método dos mínimos quadrados, tanto com quanto sem restrições, bem como os filtros ótimos 1 (OF 1) e 2 (OF 2). O foco será na reconstrução da amplitude do sinal obtido através das colisões de partículas no programa de atualização do experimento ATLAS. O intuito é minimizar o efeito de empilhamento, que pode distorcer a interpretação dos dados e prejudicar a precisão das medições.

Esperamos que este trabalho contribua para o campo da física experimental de altas energias ao abordar o complexo problema de estimação de sinal, especificamente no contexto do experimento ATLAS, uma das maiores colaborações científicas do mundo. A precisão na reconstrução de sinais de colisões de partículas é crucial para a interpretação correta dos fenômenos físicos observados no LHC. Ao focar em técnicas avançadas de estimação de sinal, como o método dos mínimos quadrados (com e sem restrições), esta pesquisa oferece uma análise detalhada das abordagens mais eficazes para minimizar o efeito de empilhamento.

Essa análise não apenas contribuirá para uma melhor compreensão dos processos subjacentes às colisões de partículas, mas também tem o potencial de influenciar futuras melhorias nos métodos de processamento de sinais em experimentos de física de partículas. Além disso, a aplicação dessas técnicas pode ser estendida a outras áreas da ciência e tecnologia, onde a estimação precisa de sinais a partir de dados ruidosos é um desafio comum.

Nesta dissertação, o primeiro capítulo, intitulado *Conceito e Contextualização*, aborda a contextualização do CERN, da física de partículas, do LHC e de seus experimentos, com ênfase no experimento ATLAS. São detalhados os calorímetros do experimento, com destaque para o Calorímetro de Telhas, seu sistema, o efeito de empilhamento de sinal e a estimação da energia.

No segundo capítulo, são discutidos o método atualmente utilizado no LHC, o OF 2, e sua versão alternativa conhecida na literatura como OF 1. Além disso, são apresentados os métodos de mínimos quadrados, com e sem restrições, propostos para a estimação da energia em consonância com as atualizações previstas para o LHC.

O terceiro capítulo descreve a metodologia adotada, incluindo a caracterização dos dados, a técnica de janelamento, a validação cruzada k-fold e as métricas empregadas para a comparação entre os métodos. Por fim, no quarto capítulo, são apresentados e analisados os resultados obtidos ao longo do estudo.

1 CONCEITOS E CONTEXTUALIZAÇÃO

Ensinar não é transferir conhecimento, mas criar as possibilidades para a sua própria produção ou a sua construção.
Paulo Freire

1.1 CERN

Em uma reunião intergovernamental da UNESCO (*United Nations Educational, Scientific and Cultural Organization*) em Paris, em dezembro de 1951, foi adotada a primeira resolução sobre a criação de um Conselho Europeu para Pesquisa Nuclear (*Conseil Européen pour la Recherche Nucléaire*, em francês) (CERN, 2024). Dois meses depois, um acordo foi assinado estabelecendo o conselho provisório – a sigla CERN nasceu (CERN, 2024). Hoje, nossa compreensão da matéria vai muito além do núcleo, e a principal área de pesquisa do CERN é a física de partículas. Por isso, o laboratório operado pelo CERN é frequentemente referido como o Laboratório Europeu de Física de Partículas (CERN, 2024).

1.1.1 Física de Partículas

A física de partículas é uma área da física dedicada ao estudo das partículas fundamentais que compõem a matéria e das forças básicas que governam suas interações (Close, 2023). Esse campo busca identificar os constituintes elementares do universo, como quarks, léptons e bósons, e compreender suas propriedades, como massa, carga e spin (Close, 2023). Além disso, investiga como essas partículas interagem entre si por meio das forças fundamentais – gravitacional, eletromagnética, forte e fraca – para explicar a formação e o comportamento de toda a matéria e energia no cosmos (Close, 2023).

De acordo com o Modelo Padrão, uma teoria que descreve as partículas fundamentais e suas interações, os léptons e quarks são considerados partículas verdadeiramente elementares, ou seja, sem estrutura interna (Moreira, 2009). Por outro lado, partículas compostas que possuem estrutura interna são chamadas de hádrons. Estas são formadas por combinações de quarks e podem ser classificadas em bárions, compostos por três quarks ou três antiquarks, e mésons, formados por um par quark-antiquark (Moreira, 2009).

Existem seis tipos de léptons: elétron, múon, tau, neutrino do elétron, neutrino do múon e neutrino do tau (Moreira, 2009). Além disso, há seis tipos de quarks: *up* (u), *down* (d), *charme* (c), *estranho* (s), *bottom* (b) e *top* (t) (Moreira, 2009).

Os quarks possuem uma propriedade única chamada carga de cor, que pode assumir três estados: vermelho, verde ou azul (Moreira, 2009). Considerando todas as combinações possíveis, existem 18 tipos distintos de quarks. Além disso, para cada partícula há uma correspondente antipartícula. Assim, no total, existem 12 léptons (6 partículas e 6 antipartículas) e 36 quarks (18 partículas e 18 antipartículas) (Moreira, 2009). Entre os léptons, o elétron é o mais conhecido, enquanto os hádrons mais familiares são o próton e o nêutron, compostos por combinações específicas de quarks (Moreira, 2009).

O CERN é a Organização Europeia para a Pesquisa Nuclear, uma das mais importantes instituições científicas do mundo, especialmente na área de física de partículas. Fundado em 1954, o CERN está localizado na fronteira entre a França e a Suíça, perto de Genebra (Evans; Bryant, 2008). A Figura 1 mostra a vista do Globo de Ciência e Inovação do CERN em um dia de inverno suíço.

O CERN é mais conhecido pelo seu Grande Colisor de Hádrons (LHC, na sigla em inglês), o maior e mais potente acelerador de partículas do mundo (Hasert, 1973). Desde o início de suas operações, o CERN tem sido responsável por diversos avanços tecnológicos, como a criação da World Wide Web (WWW) nos anos 1980, além de importantes descobertas científicas. Em 1974, o experimento Gargamelle revelou a corrente neutra das interações fracas (Hasert, 1973).

Em 1983, os experimentos UA1 e UA2 observaram as partículas W e Z (Arnison; al., 1983). Em 1989, medições mais precisas da partícula Z confirmaram a existência de apenas três famílias de partículas na natureza. Uma das descobertas mais importantes ocorreu em 2012, com a identificação do bóson de Higgs (Pimenta et al., 2013), um marco que rendeu o Prêmio Nobel de Física em 2013 a François Englert e Peter Higgs.

O bóson de Higgs é uma partícula fundamental responsável por conferir massa às demais partículas subatômicas (Pimenta et al., 2013). Com essa descoberta, o CERN consolidou seu prestígio mundial. A instituição reúne mais de 20 países membros e milhares de cientistas de todo o mundo que colaboram nos experimentos (CERN, 2024). Além da pesquisa em física fundamental, o CERN tem contribuído com avanços tecnológicos que impactam áreas como medicina e tecnologia da informação.

Como mencionado anteriormente, a criação da WWW ocorreu no CERN como uma solução para facilitar o compartilhamento de informações entre cientistas. Além disso, o CERN desempenha um papel importante na educação e divulgação científica, oferecendo programas de treinamento para estudantes, professores e jovens cientistas, além de promover visitas e eventos culturais para o público geral (CERN, 2024).

O complexo do CERN opera nove aceleradores e dois desaceleradores. Esses aceleradores fornecem partículas para experimentos ou as preparam para aceleradores maiores. Um acelerador impulsiona partículas carregadas, como prótons ou elétrons (componentes subatômicos do núcleo atômico), a velocidades próximas à da luz. Essas partículas são então colididas contra um alvo fixo ou contra outras partículas que circulam na direção oposta. O estudo dessas colisões permite que os físicos explorem o mundo do infinitamente pequeno. Quando as partículas possuem energia suficiente, ocorre um fenômeno extraordinário: a energia da colisão é convertida em matéria na forma de novas partículas, algumas das quais existiram apenas nos primórdios do universo.

Esse fenômeno é descrito pela famosa equação de Einstein, $E = mc^2$, que estabelece a equivalência entre energia e massa. A luminosidade é um indicador-chave do desempenho de um acelerador, representando o número de colisões potenciais por unidade de superfície ao longo de um determinado período de tempo. Os aceleradores do CERN são controlados 24 horas por dia pelo Centro de Controle da instituição (CERN, 2024a).

Figura 1 – Vista do Globo de Ciência e Inovação no CERN em um dia de inverno suíço.

Fonte: Adaptado de CERN (2024).

O CERN continua a liderar a pesquisa em física de partículas, com novos experimentos e atualizações do LHC planejados para os próximos anos, buscando responder a questões fundamentais sobre o universo e suas origens. Entre os experimentos conduzidos pelo CERN, o mais conhecido e também considerado o principal é o LHC.

O LHC é uma das maiores e mais complexas máquinas já construídas, sendo essencial para a exploração dos componentes básicos da matéria e das forças que atuam entre eles. Foi nesse experimento que ocorreu a confirmação do bóson de Higgs, entre outras descobertas importantes. Devido à sua relevância científica e impacto na física de partículas, abordaremos o LHC de forma mais detalhada na próxima seção, explorando suas contribuições para a ciência e os desafios tecnológicos envolvidos em seu funcionamento.

1.1.2 LHC

O Grande Colisor de Hádrons, mais conhecido como LHC, é o acelerador de partículas do CERN e o mais poderoso acelerador de partículas do mundo. Ele foi ativado pela primeira vez em 10 de setembro de 2008 e continua sendo a mais recente adição ao complexo de aceleradores do CERN. Trata-se de uma gigantesca máquina em forma de anel, com 27 quilômetros de circunferência, composta por ímãs supercondutores e várias estruturas de aceleração para aumentar a energia das partículas ao longo do caminho. O LHC está localizado a cerca de 100 metros abaixo do solo (Evans; Bryant, 2008), como ilustrado na figura 2.

Dentro do acelerador, dois feixes de prótons de alta energia viajam a uma velocidade próxima à da luz antes de colidirem. Esses feixes se deslocam em direções opostas dentro de tubos mantidos em ultra-alto vácuo. Eles são guiados ao longo do anel do acelerador por um intenso campo magnético gerado por eletroímãs supercondutores.

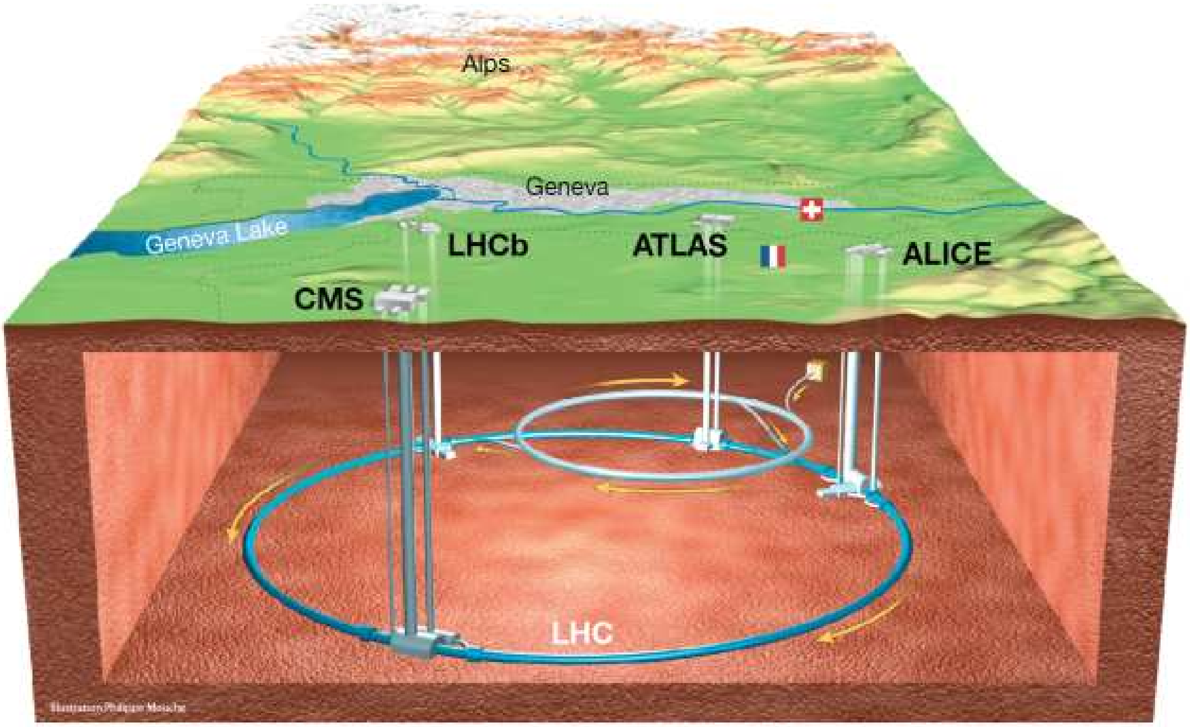
Os eletroímãs são construídos com bobinas de um cabo elétrico especial que opera em estado supercondutor, conduzindo eletricidade sem resistência ou perda de energia. Para isso, os ímãs devem ser resfriados a -271,3°C – uma temperatura mais baixa do que a do espaço exterior. Dessa forma, grande parte do acelerador está conectada a um sistema de distribuição de hélio líquido, responsável pelo resfriamento dos ímãs e pelo fornecimento de outros serviços essenciais (CERN, 2024c).

Milhares de ímãs de diferentes tipos e tamanhos são utilizados para direcionar os feixes dentro do acelerador. Entre eles, destacam-se 1232 ímãs dipolos, com 15 metros de comprimento, que curvam os feixes, e 392 ímãs quadrupolos, com comprimento entre 5 e 7 metros, que focam os feixes (CERN, 2024c). Pouco antes da colisão, outro tipo de ímã é empregado para "comprimir"as partículas, aumentando a probabilidade de colisões. As partículas são tão pequenas que fazê-las colidir é comparável a disparar duas agulhas a 10 quilômetros de distância com precisão suficiente para que se encontrem no meio do caminho (CERN, 2024c).

Todos os controles do acelerador, assim como seus serviços e infraestrutura técnica, são gerenciados a partir de um único local: o Centro de Controle do CERN. Nesse centro, os feixes são direcionados para colidirem em quatro locais ao longo do anel do acelerador, onde estão posicionados os principais detectores de partículas: CMS (Compact Muon Solenoid), LHCb (Large Hadron Collider beauty), ALICE (A Large Ion Collider

Experiment) e ATLAS, conforme ilustrado na figura 2.

Figura 2 – Visão geral do LHC, incluindo os experimentos ALICE, ATLAS, CMS e LHCb.



Fonte: Adaptado de Sánchez (2010).

1.1.3 Os principais experimentos do LHC

Os principais experimentos do LHC são CMS, LHCb, ALICE e ATLAS (Evans; Bryant, 2008). O LHC conta com dois experimentos de propósito geral – que abrangem um amplo programa de física –, ATLAS (A Toroidal LHC ApparatuS) e CMS, ambos operando em alta luminosidade, com uma luminosidade máxima de $L = 10^{34}, \mathrm{cm}^{-2}, \mathrm{s}^{-1}$ para colisões de prótons (Evans; Bryant, 2008).

A luminosidade é proporcional à densidade do feixe de partículas; quanto maior a luminosidade, mais eventos são registrados nas colisões. Além desses, há dois experimentos de baixa luminosidade: LHCb, voltado para a física do quark bottom (B), com luminosidade máxima de $L = 10^{32}, \mathrm{cm}^{-2}, \mathrm{s}^{-1}$, e TOTEM, que investiga a dispersão elástica de prótons em pequenos ângulos, com luminosidade máxima de $L = 2 \times 10^{29}, \mathrm{cm}^{-2}, \mathrm{s}^{-1}$ e

156 pacotes de feixes (Evans; Bryant, 2008).

Além dos feixes de prótons, o LHC também opera com feixes de íons. Para esse fim, conta com um experimento dedicado a colisões de íons pesados, ALICE, que visa alcançar uma luminosidade máxima de $L = 10^{27}, \mathrm{cm}^{-2}, \mathrm{s}^{-1}$ em colisões de íons de chumbo (Evans; Bryant, 2008). A seguir, exploraremos com mais detalhes os experimentos CMS, LHCb, ALICE e, com maior destaque, o experimento ATLAS.

1.1.3.1 Experimento CMS

O CMS, representado na figura 3, é um detector de propósito geral instalado no LHC. Ele foi projetado para estudar diversos fenômenos físicos, incluindo aspectos do Modelo Padrão, como a investigação do bóson de Higgs, e a busca por novas descobertas além desse modelo, como evidências de dimensões extras e partículas associadas à matéria escura (CERN, 2024d). Embora compartilhe os mesmos objetivos científicos do experimento ATLAS, o CMS emprega abordagens técnicas distintas, incluindo um sistema magnético diferenciado (CERN, 2024d).

No centro do CMS, encontra-se um grande ímã solenoide, essencial para o seu funcionamento (CERN, 2024d). Esse ímã tem formato cilíndrico e é constituído por uma bobina feita de cabo supercondutor, capaz de gerar um intenso campo magnético de 4 teslas – aproximadamente 100.000 vezes mais forte que o campo magnético da Terra (CERN, 2024d). Esse campo é contido por uma estrutura de aço, que contribui significativamente para o peso total do detector, cerca de 14.000 toneladas (CERN, 2024d).

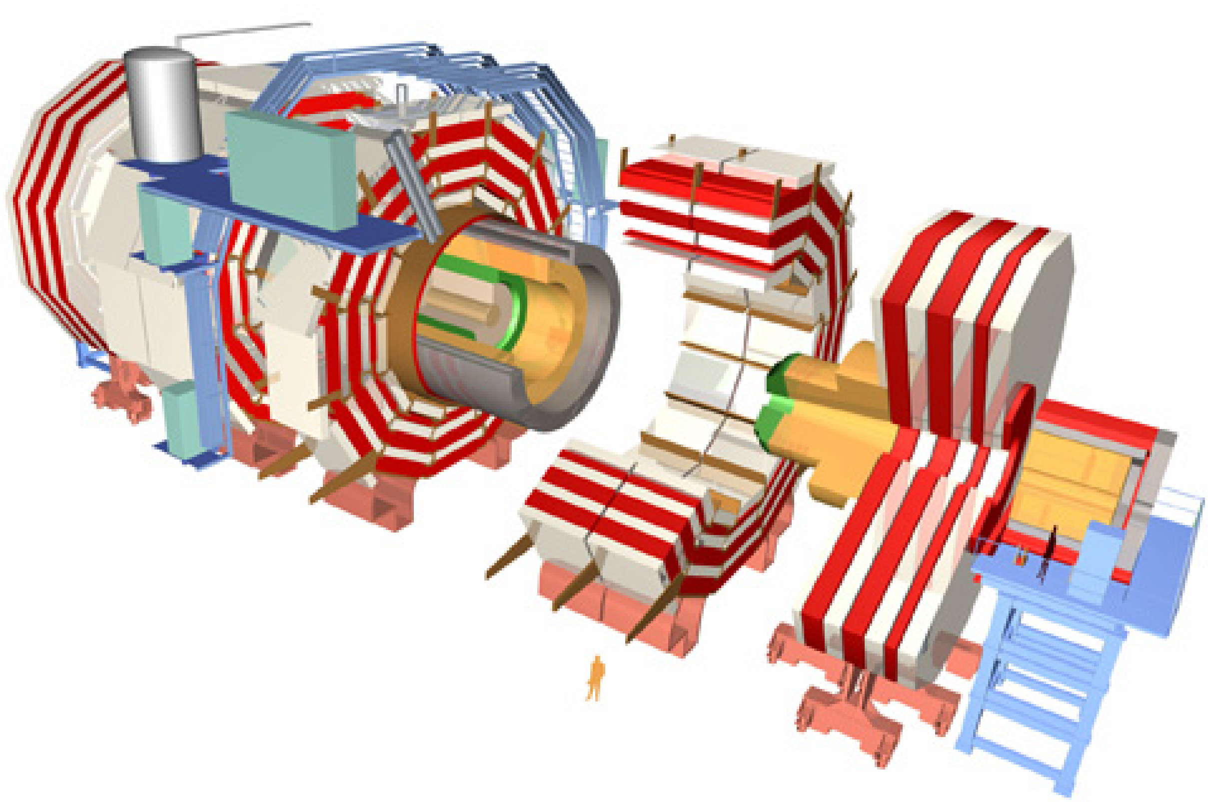
Diferente de outros grandes detectores do LHC, o CMS não foi montado diretamente no local de operação (CERN, 2024d). Suas 15 seções foram primeiramente construídas ao nível do solo e, posteriormente, baixadas até uma caverna subterrânea próxima a Cessy, na França, onde o detector foi remontado (CERN, 2024d). Quando totalmente montado, o CMS possui dimensões impressionantes: 21 metros de comprimento, 15 metros de largura e 15 metros de altura (CERN, 2024d).

Além de sua estrutura monumental, o CMS se destaca pelo grande número de colaboradores. Em maio de 2022, o experimento contava com aproximadamente 5500 membros, entre físicos de partículas, engenheiros, técnicos, estudantes e equipe de apoio, distribuídos entre 241 instituições de 54 países (CERN, 2024d).

1.1.3.2 Experimento LHCb

O experimento LHCb, ilustrado na figura 4, tem como principal objetivo investigar as pequenas diferenças entre matéria e antimatéria, concentrando-se em um tipo específico

Figura 3 – Experimento CMS.



Fonte: Adaptado de Sánchez (2010).

de partícula conhecido como "quark beleza"ou "quark b"(CERN, 2024e). Diferentemente dos detectores dos experimentos ATLAS e CMS, que cercam completamente o ponto de colisão, o LHCb adota uma abordagem distinta. Ele utiliza uma série de subdetectores dispostos em linha, projetados para detectar principalmente as partículas lançadas para frente, ou seja, aquelas que seguem em uma direção específica após a colisão (CERN, 2024e).

O primeiro subdetector está localizado próximo ao ponto de colisão, enquanto os demais são alinhados em sequência ao longo de uma extensão de 20 metros (CERN, 2024e). Durante as colisões no LHC, diversos tipos de quarks são gerados, mas rapidamente decaem em outras partículas. Para capturar os quarks b, o LHCb é equipado com detectores de rastreamento altamente sofisticados, posicionados próximos ao caminho dos feixes de partículas que circulam no acelerador (CERN, 2024e).

Com um peso de 5600 toneladas, o detector LHCb é composto por um espectrômetro frontal e detectores em formato plano. Ele possui dimensões consideráveis, medindo 21 metros de comprimento, 10 metros de altura e 13 metros de largura, e está localizado a 100 metros de profundidade, próximo à cidade de Ferney-Voltaire, na França. Em março de 2022, cerca de 1565 cientistas, engenheiros e técnicos de 20 países estavam envolvidos na colaboração LHCb (CERN, 2024e).

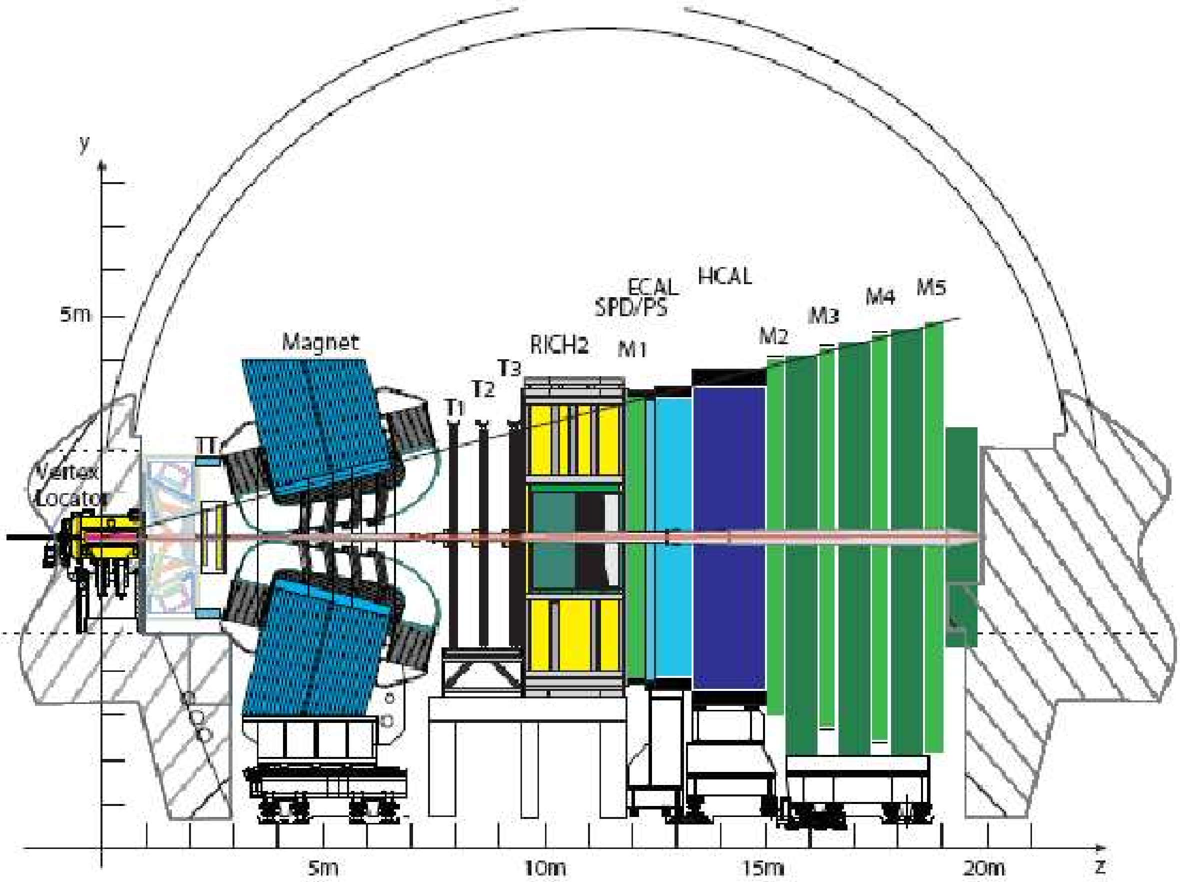
1.1.3.3 Experimento ALICE

O experimento ALICE, ilustrado na figura 5, é um detector especialmente dedicado ao estudo de íons pesados no Grande Colisor de Hádrons (LHC). Ele foi projetado para investigar as propriedades da matéria fortemente interagente em condições de densidade de energia extremamente altas, onde se forma uma fase da matéria chamada plasma de quarks e glúons (CERN, 2024b). Toda a matéria comum no universo é composta por átomos, que possuem um núcleo formado por prótons e nêutrons (exceto o hidrogênio, que não possui nêutrons), cercado por uma nuvem de elétrons (CERN, 2024b).

Tanto os prótons quanto os nêutrons são compostos por quarks, que são mantidos unidos por partículas chamadas glúons (CERN, 2024b). Até hoje, nenhum quark foi observado de forma isolada, pois tanto eles quanto os glúons estão permanentemente confinados dentro de partículas compostas, como prótons e nêutrons (CERN, 2024b). Esse fenômeno é conhecido como confinamento (CERN, 2024b). No LHC, as colisões podem gerar temperaturas mais de 100.000 vezes superiores às do centro do Sol (CERN, 2024b).

Durante parte do ano, o acelerador realiza colisões entre íons de chumbo, recriando condições semelhantes às que existiam logo após o Big Bang. Nessas condições extremas, prótons e nêutrons "derretem", liberando os quarks e glúons que estavam con-

Figura 4 – Experimento LHCb.



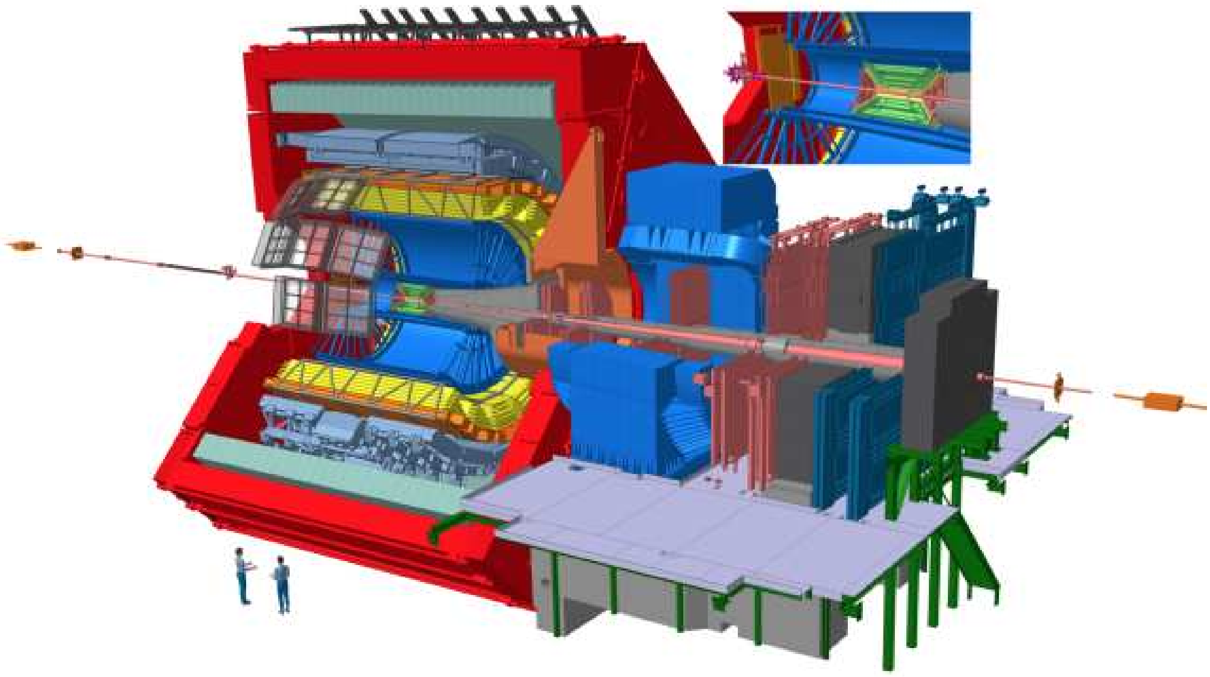
Fonte: Adaptado de Sánchez (2010).

finados (CERN, 2024b). Esse estado é o plasma de quarks e glúons, cuja existência e propriedades são questões fundamentais da teoria da cromodinâmica quântica (QCD) (CERN, 2024b).

O estudo desse plasma auxilia na compreensão do fenômeno do confinamento e da restauração da simetria quiral, um problema central na física teórica. A colaboração ALICE observa esse plasma conforme ele se expande e esfria, analisando como ele progressivamente dá origem às partículas que formam a matéria que conhecemos hoje no universo (CERN, 2024b).

O detector ALICE, que pesa cerca de 10.000 toneladas, tem 26 metros de comprimento, 16 metros de altura e 16 metros de largura, e está localizado em uma grande caverna a 56 metros abaixo do solo, perto da vila de St. Genis-Pouilly, na França (CERN, 2024b). Recebendo feixes do LHC, o ALICE reúne uma colaboração internacional composta por quase 2.000 cientistas de 174 institutos de física em 40 países (dados de abril de 2022) (CERN, 2024b).

Figura 5 – Experimento ALICE.



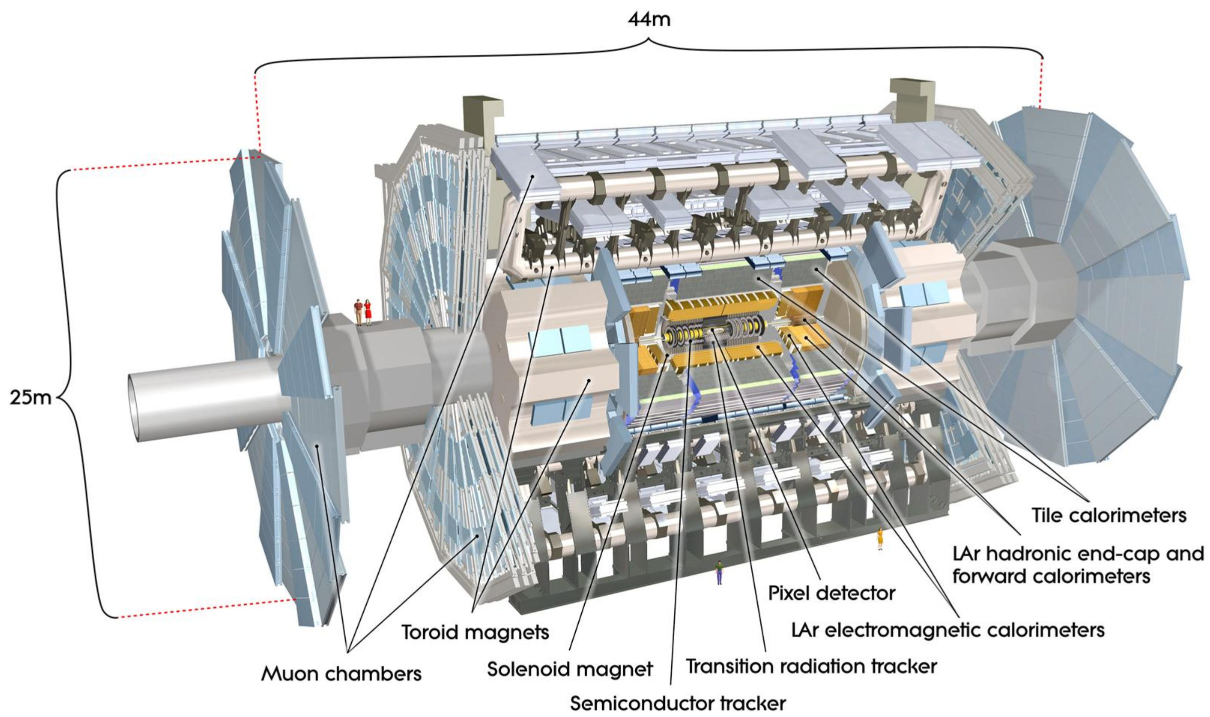
Fonte: Adaptado de Sánchez (2010).

1.1.3.4 Experimento ATLAS

O experimento ATLAS (Kordas, 2007) é um detector de propósito geral projetado para explorar todo o potencial de descoberta do LHC, sendo considerado o principal experimento do colisor. O ATLAS tem 44 metros de comprimento, 25 metros de altura e um peso total de aproximadamente 7000 toneladas. Ele é dividido em subdetectores, como ilustrado na figura 6 (Sánchez, 2010).

O Detector Interno representa a parte mais interna do ATLAS, cercado por um ímã solenoide, os Calorímetros, o sistema de Múons e um grande ímã toroidal de núcleo de ar. Ele foi projetado para operar em alta luminosidade ($10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$) com cruzamentos de pacotes a cada 25 ns (Sánchez, 2010). Para isso, emprega tecnologias altamente sofisticadas e materiais especializados.

Figura 6 – Experimento ATLAS.



Fonte: Adaptado de Sánchez (2010).

O experimento ATLAS conta com uma das maiores colaborações científicas do mundo, envolvendo aproximadamente 5900 pessoas, entre físicos, engenheiros, técnicos, estudantes e administradores (CERN, 2023). Dentro desse grupo, há 2900 autores científicos afiliados a mais de 180 instituições (ou mais de 230 institutos) ao redor do mundo, representando mais de 40 países, conforme ilustrado na figura 7. Embora o número de países oficialmente envolvidos seja superior a 40, as nacionalidades representadas na colaboração chegam a cerca de 100 (CERN, 2023).

Devido à alta frequência de colisões ocorrendo no LHC e à vasta quantidade de dados produzidos pelo experimento ATLAS, além da significativa presença de ruído de fundo resultante dessas colisões, torna-se essencial a implementação de um sistema de

Figura 7 – Mapa de colaboração do experimento ATLAS.

Fonte: Adaptado de Collaboration (2024).

filtragem online de eventos, conhecido como sistema Trigger (Sánchez, 2010).

1.1.4 Sistema Trigger

A Figura 8 apresenta uma visão geral do sistema Trigger e DAQ do ATLAS planejado para o Run-2. O sistema Trigger, conforme descrito em (Nakahama; Collaboration, 2015), é composto por dois níveis: um primeiro nível de hardware chamado L1 (Level 1) e um Trigger de alto nível (HLT, High-Level Trigger) baseado em software. Este novo sistema em duas etapas tem a função de reduzir a taxa de eventos da frequência de cruzamento de feixes de 40 MHz para 100 kHz no L1, e posteriormente para uma taxa de gravação média de 1 kHz no HLT (Nakahama; Collaboration, 2015).

No Run-1 (Nakahama; Collaboration, 2015), o sistema era composto por três níveis, com duas fases no HLT. No L1, eletrônicos rápidos e sob medida identificam regiões de interesse (RoI, Regions of Interest) a partir dos dados dos calorímetros e do sistema de múons, utilizando informações menos detalhadas, dentro de uma latência de 2,5 μs. Para o Run-2, o sistema L1 é composto pelo sistema Trigger do calorímetro (L1Calo), o sistema Trigger de múons (L1Muon), novos módulos de Trigger topológicos (L1Topo), além dos Processadores Centrais de Trigger (CTP).

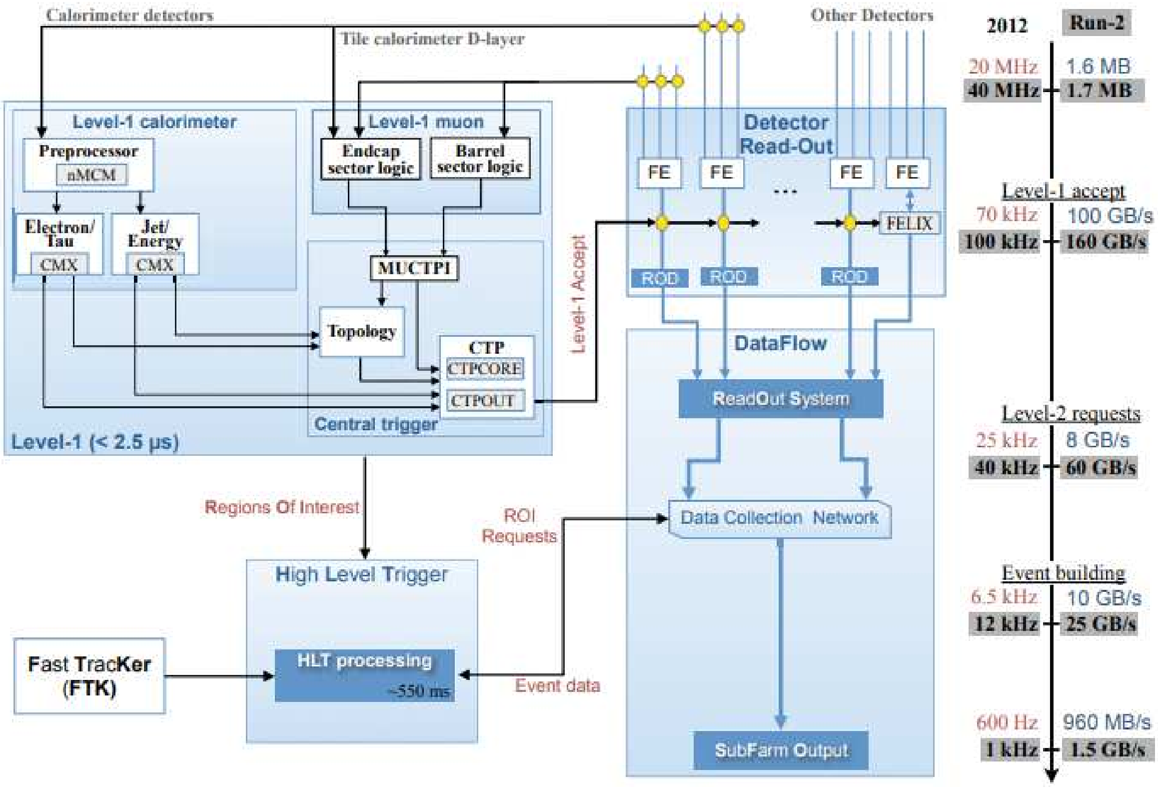
No HLT, algoritmos rápidos processam dados a partir das RoIs ou de informações completas dos eventos, utilizando um cluster único de computadores. O tempo médio de processamento é de 0,2 segundos (Nakahama; Collaboration, 2015). Até o final de 2016, está prevista a integração de um rastreador de hardware (FTK, Hardware Track Finder), que fornecerá informações de trilhas ao HLT nas taxas do L1 (Nakahama; Collaboration, 2015). Esses anais detalham as principais atualizações realizadas no L1Calo, L1Muon, L1Topo e no HLT para o Run-2, as quais proporcionam novas funcionalidades em todos os subsistemas (Nakahama; Collaboration, 2015).

Esse sistema é crucial para a operação eficiente do experimento, pois permite a seleção em tempo real dos eventos mais relevantes para análise, garantindo que os dados significativos sejam capturados enquanto se descarta o ruído e as colisões de menor interesse (Nakahama; Collaboration, 2015). A seguir, veremos como as informações selecionadas pelo sistema Trigger são recolhidas e apresentadas para serem analisadas no experimento ATLAS.

1.1.5 Calorimetria do ATLAS

A física experimental de altas energias enfrenta desafios significativos no processamento de sinais, principalmente devido à alta taxa de eventos capturada pelos sistemas

Figura 8 – Visão geral esquemática da configuração Run-2 do sistema de Trigger e DAQ.



Fonte: Adaptado de Nakahama e Collaboration (2015).

que operam em experimentos de colisões de partículas. Os sistemas de calorimetria desempenham um papel fundamental nesse contexto, sendo responsáveis por absorver, amostrar e medir a energia das partículas que interagem com o material do calorímetro.

O objetivo principal desses sistemas é reconstruir a energia depositada em cada canal de leitura e, a partir disso, inferir a identidade das partículas detectadas (Wigmans, 2017), ou seja, o princípio de funcionamento dos calorímetros que compõem o sistema de calorimetria é medir a energia depositada ao longo de suas camadas pelas partículas oriundas dos processos de colisão.

Os calorímetros modernos são equipados com milhares de canais de leitura, proporcionando alta resolução espacial na amostragem da energia das partículas. Cada canal de leitura gera um sinal elétrico característico, tipicamente na forma de um pulso, que é suscetível a ruído. Esse pulso é então amplificado, conformado e digitalizado, permitindo o processamento subsequente e a estimativa de parâmetros críticos, como amplitude e fase do sinal (Rimes et al., 2020). De acordo com (Fabjan, 1985), há várias razões pelas quais os calorímetros se tornaram sistemas cruciais em muitos experimentos de física de partículas:

- **Sensibilidade a Diferentes Tipos de Partículas:** Calorímetros podem detectar tanto partículas neutras quanto carregadas, o que amplia sua capacidade de medir uma variedade de eventos de colisão.

- **Alta Eficiência na Identificação de Partículas:** Devido à maneira distinta como a energia é depositada pelas partículas, calorímetros permitem uma identificação eficiente e precisa das partículas.

- **Compactação dos Detectores:** A profundidade dos calorímetros aumenta logaritmicamente com a energia das partículas, o que permite o desenvolvimento de detectores mais compactos, capazes de conter as cascatas de partículas que se desenvolvem durante a colisão.

- **Segmentação para Medidas Precisas:** Calorímetros podem ser segmentados em diferentes regiões, possibilitando a medição tanto da energia quanto da trajetória das partículas com alta precisão.

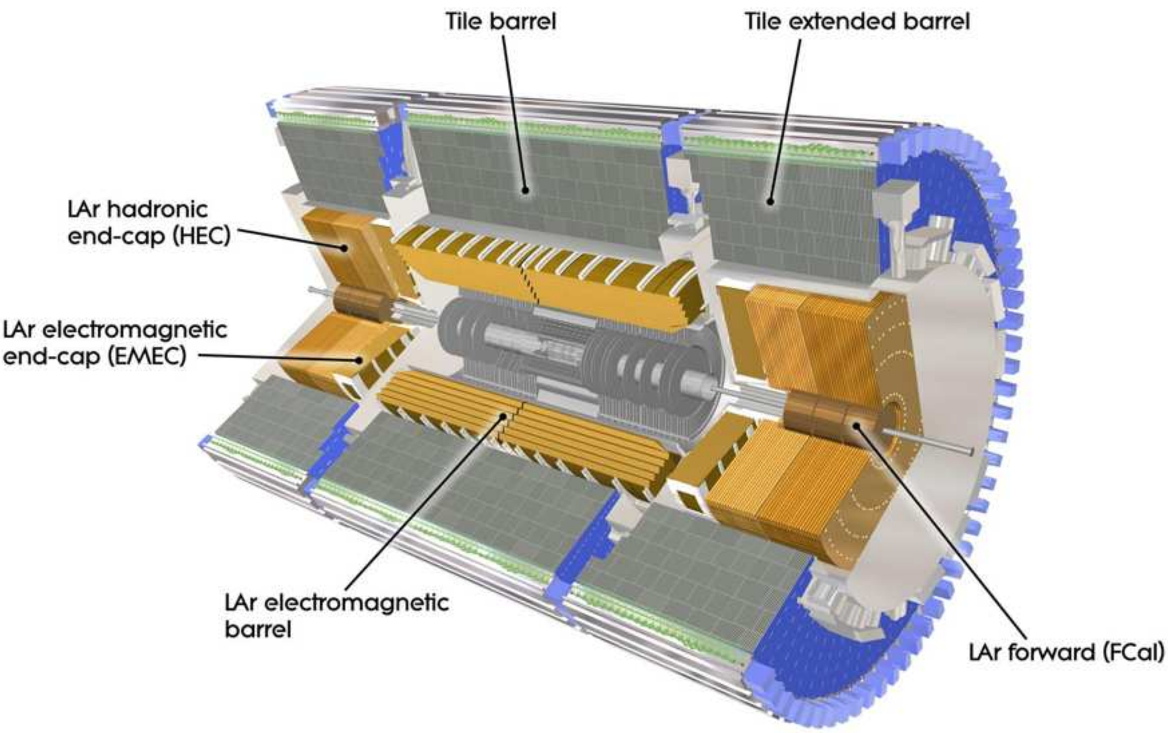
- **Resposta Rápida:** Calorímetros podem responder rapidamente, com tempos de resposta inferiores a 50 nanosegundos. Essa característica é crucial em ambientes de alta taxa de eventos, onde é necessário capturar informações rapidamente.

- **Filtragem Seletiva de Eventos:** A informação de energia obtida pelos calorímetros pode ser utilizada para filtrar eventos interessantes com alta seletividade, permitindo uma análise mais eficaz dos dados experimentais.

O sistema de calorimetria do experimento ATLAS ilustrado na 9 é dividido em duas partes principais: o calorímetro eletromagnético e o calorímetro hadrônico. O calorímetro eletromagnético foi projetado especificamente para medir a energia de partículas que interagem através da força eletromagnética, como elétrons e fótons.

Por outro lado, o calorímetro hadrônico é responsável pela medição da energia de hádrons, como prótons e nêutrons. Uma característica importante dos calorímetros é que, exceto pelos múons e neutrinos, todas as partículas que passam por eles são completamente absorvidas (Abdallah et al., 2013). Múons e neutrinos, devido à sua interação fraca ou à sua alta penetração, conseguem atravessar os calorímetros sem serem totalmente absorvidos.

Figura 9 – Calorimetria do ATLAS.

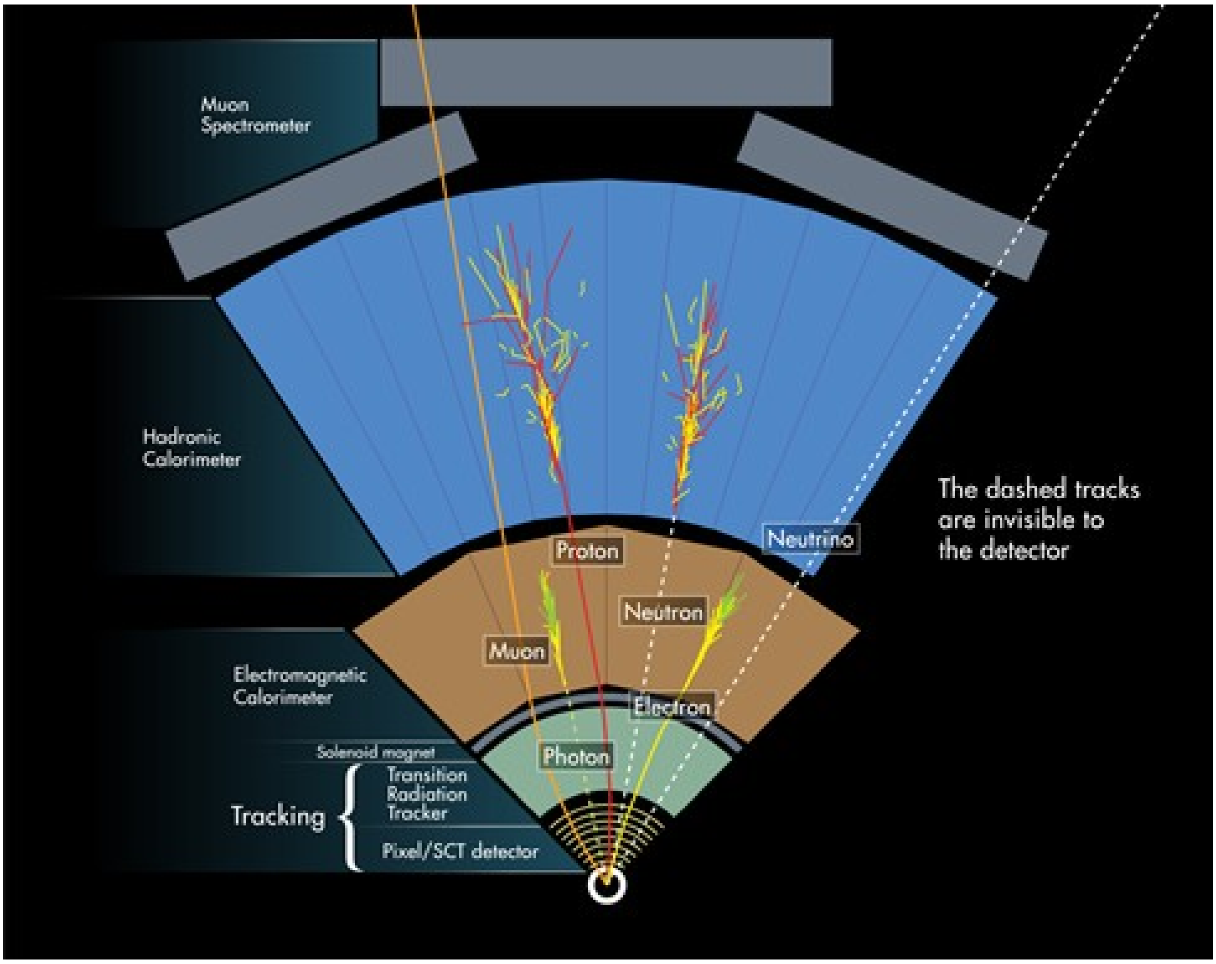


Fonte: Adaptado de ATLAS Collaboration (2005).

A figura 10 apresenta uma representação detalhada do processo pelo qual as partículas resultantes das colisões no experimento ATLAS se propagam através dos diferentes sistemas de detecção, ou melhor dizendo, o percurso das partículas desde o ponto de colisão até a sua interação com os diversos componentes do detector, destacando como cada sistema contribui para a reconstrução e análise dos dados da colisão.

O processo de propagação das partículas geradas pela colisão através dos sistemas de detecção do ATLAS é composto por várias etapas importantes, das quais podemos

Figura 10 – Propagação das partículas formadas na colisão através dos sistemas de detecção do ATLAS.



Fonte: Adaptado de ATLAS Collaboration (2005).

destacar as seguintes:

- **Passagem pelo Calorímetro:** As partículas atravessam o calorímetro, que é dividido em calorímetros eletromagnéticos e hadrônicos. O calorímetro eletromagnético mede partículas carregadas, como elétrons e pósitrons, e fornece informações sobre a energia depositada por essas partículas. Já o calorímetro hadrônico mede partículas hadrônicas, como prótons e nêutrons, que interagem fortemente com o material do detector.

- **Registro pelo Sistema de Trajetórias:** Após interagirem com o calorímetro, as partículas passam pelo sistema de trajetórias, composto por detectores de pontos de impacto (tracking detectors). Esses detectores rastreiam as trajetórias das partículas e permitem a reconstrução de suas origens e trajetórias ao longo do detector.

- **Identificação de Partículas:** Os dados coletados pelos calorímetros e pelo sistema de trajetórias são combinados para identificar o tipo de partículas e suas características. Isso envolve a análise da energia depositada, a quantidade de partículas e o caminho percorrido dentro do detector.

- **Análise de Dados:** Finalmente, os dados são analisados para reconstruir os eventos de colisão e extrair informações sobre os processos físicos que ocorreram. Esses dados são utilizados para validar modelos teóricos e realizar descobertas experimentais na física de partículas.

A seguir falaremos de forma mais detalhada sobre os dois principais tipos de calorimetria do ATLAS.

1.1.6 Calorímetro Eletromagnético

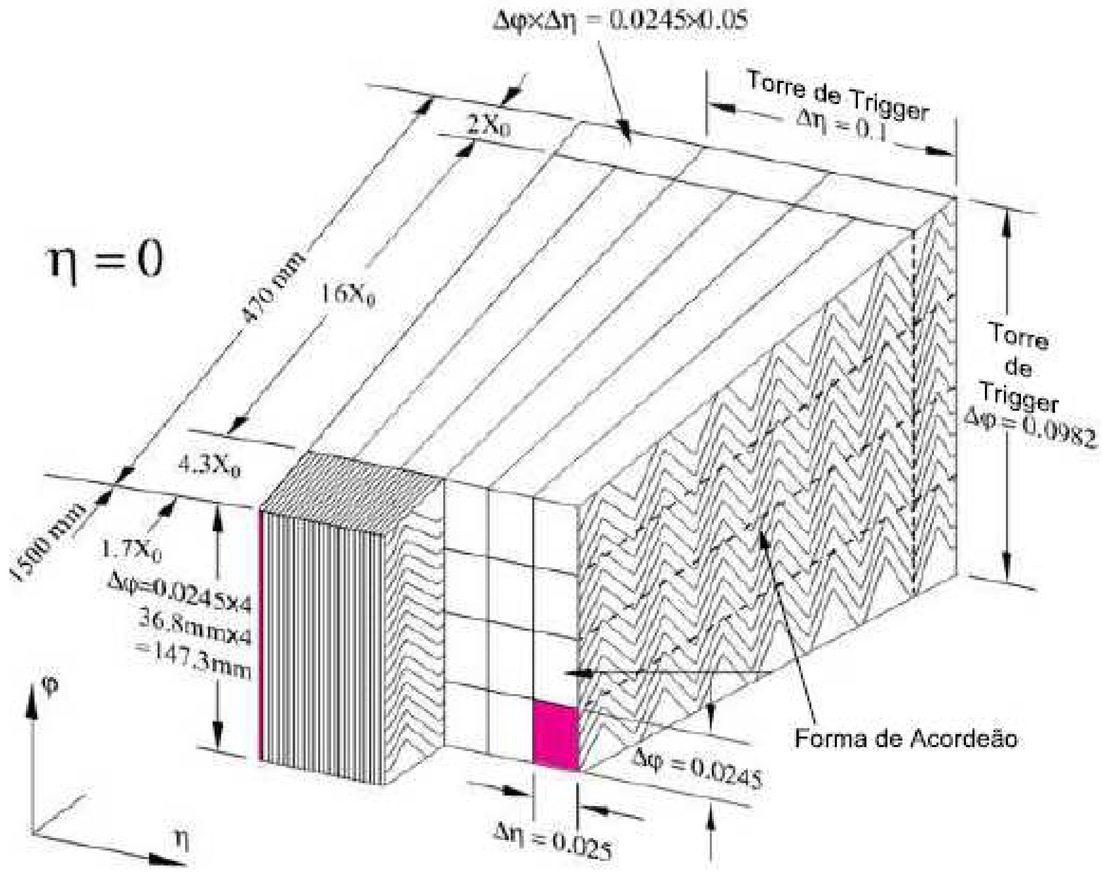
O Calorímetro de Argônio Líquido (LAr) também conhecido como o calorímetro eletromagnético (EM) do experimento ATLAS e utiliza chumbo como material absorvedor, com eletrodos em forma de acordeão imersos em argônio líquido como amostradores, responsáveis por medir a energia das partículas através da ionização das células do calorímetro (Sánchez, 2010).

Este detector cobre a pseudo-rapidez $|\eta| < 3.2$ e é dividido em duas partes principais: o barril ($|\eta| < 1.475$) e as tampas (end-caps), que cobrem até $|\eta| = 3.2$. No ponto de $|\eta| = 1.375$, o barril começa a sobrepor as tampas, que se dividem em tampas exteriores (até $|\eta| = 2.5$) e interiores (de $|\eta| = 2.5$ a $|\eta| = 3.2$) (The ATLAS Collaboration, 2008).

Este calorímetro (Perrodo, 2002) é segmentado em três camadas longitudinais, com granularidade específica em cada uma, sendo a segunda camada a mais profunda como é mostrado na figura 11. A segmentação e granularidade variam ao longo do eixo de pseudo-rapidez (η) e em torno da rotação (eixo ϕ), com maior detalhamento em regiões específicas, permitindo a identificação precisa das interações das partículas com o detector (Perrodo, 2002).

A espessura total do calorímetro eletromagnético ultrapassa 22 comprimentos de radiação para o barril e 24 para as tampas finais, garantindo alta eficiência na medição da energia das partículas, tanto no barril quanto nas tampas (Perrodo, 2002).

Figura 11 – Segmentação do calorímetro eletromagnético do ATLAS.



Fonte: Adaptado de Perrodo (2002).

1.1.7 Calorímetro Hadrônico

O experimento ATLAS (The ATLAS Collaboration, 2008) utiliza três tipos de calorímetros hadrônicos: o calorímetro hadrônico de end-cap de argônio (HEC), o calorímetro de avanço de argônio líquido e o principal calorímetro hadrônico (FCal), denominado calorímetro de telhas (TileCal).

1.1.7.1 Calorímetro Hadrônico de End-Cap de Argônio Líquido (HEC)

O Calorímetro Hadrônico de End-cap (HEC) (The ATLAS Collaboration, 2008) é composto por duas rodas independentes em cada end-cap. Essas rodas ficam localizadas diretamente atrás do calorímetro eletromagnético do end-cap e compartilham os mesmos criostatos de LAr. Para minimizar a redução da densidade do material na transição entre o end-cap e o calorímetro de avanço (em torno de $|\eta| = 3.1$), o HEC se estende até $|\eta| = 3.2$, sobrepondo-se ao calorímetro de avanço (The ATLAS Collaboration, 2008). Além disso, o intervalo de η do HEC se sobrepõe ligeiramente ao do calorímetro de tijolos ($|\eta| < 1.7$), estendendo-se até $|\eta| = 1.5$ (The ATLAS Collaboration, 2008).

Cada roda do HEC (The ATLAS Collaboration, 2008) é formada por 32 módulos idênticos em forma de cunha, montados na periferia e no centro da roda. Cada roda é dividida em dois segmentos em profundidade, totalizando quatro camadas em cada end-cap. As rodas mais próximas ao ponto de interação são feitas com placas de cobre de 25 mm de espessura, enquanto as mais distantes usam placas de cobre de 50 mm (sendo que a primeira placa de todas tem metade da espessura) (The ATLAS Collaboration, 2008).

O raio externo das placas de cobre é de 2,03 m e o raio interno é de 0,475 m (exceto na região de sobreposição com o calorímetro de avanço, onde o raio interno é reduzido para 0,372 m) (The ATLAS Collaboration, 2008). As placas de cobre são intercaladas com lacunas de LAr de 8,5 mm, que fornecem o meio ativo para este calorímetro de amostragem (The ATLAS Collaboration, 2008).

1.1.7.2 Calorímetro de Avanço de Argônio Líquido (FCal)

O Calorímetro de Avanço (FCal) (The ATLAS Collaboration, 2008) está integrado nos criostatos dos end-caps, o que oferece vantagens como uma cobertura calorimétrica mais uniforme e a redução de radiação de fundo no espectrômetro de múons. Para diminuir o albedo de nêutrons na cavidade do detector interno, a face frontal do FCal é recuada em cerca de 1,2 metros em relação ao calorímetro eletromagnético.

Esse recuo, no entanto, limita a profundidade disponível, exigindo um design de alta densidade. O FCal (The ATLAS Collaboration, 2008) tem aproximadamente 10 comprimentos de interação e é composto por três módulos em cada end-cap: o primeiro é de cobre, ideal para medições eletromagnéticas, enquanto os dois últimos, feitos de tungstênio, medem principalmente a energia das interações hadrônicas.

Cada módulo contém uma matriz metálica com canais longitudinais paralelos ao eixo do feixe, preenchidos por hastes e tubos concêntricos (The ATLAS Collaboration, 2008). O Argônio Líquido (LAr) (The ATLAS Collaboration, 2008) preenche o espaço entre as hastes e os tubos, atuando como o meio sensível. Essa configuração oferece controle preciso sobre as lacunas, que são estreitas, chegando a 0,25 mm na primeira seção, para evitar o acúmulo de íons.

1.1.7.3 Calorímetro de Telhas (TileCal)

O TileCal, também conhecido como Calorímetro de Telhas (do inglês, Tile Calorimeter), é o principal e maior calorímetro hadrônico do experimento ATLAS (Francavilla; Collaboration et al., 2012). Este componente desempenha um papel crucial na detecção e medição de partículas hadrônicas geradas nas colisões do LHC. Dada sua importância, o presente trabalho será desenvolvido no contexto deste calorímetro, com uma análise detalhada de seu funcionamento, estrutura e aplicação. A seguir, serão fornecidas descrições mais aprofundadas sobre suas características e o papel que desempenha no experimento ATLAS.

O TileCal (Davidek, 2022) é um calorímetro de amostragem utilizado no experimento ATLAS, projetado para medir a energia das partículas hadrônicas. Ele utiliza telhas cintilantes como meio ativo, que emitem luz quando atravessadas por partículas, e placas de aço como absorvedores, que capturam a energia dessas partículas (Davidek, 2022).

O calorímetro é dividido em quatro partes: dois barris principais (LB) e dois barris estendidos (EB), cobrindo um intervalo de pseudorapidez de $|\eta| < 1.7$ (Davidek, 2022). Ao longo da direção azimutal (ϕ), ele é segmentado em 64 módulos (Davidek, 2022).

A luz gerada nas telhas cintilantes é coletada por fibras de deslocamento de comprimento de onda, que a conduzem até os fotomultiplicadores (PMTs) (Francavilla; Collaboration et al., 2012). Cada telha possui duas fibras, conectadas a lados opostos em ϕ, que se direcionam para PMTs distintos, o que permite uma leitura redundante e confiável do sinal (Francavilla; Collaboration et al., 2012). Cada PMT recebe sinais de múltiplas telhas, organizadas em células de diferentes tamanhos, dependendo da pseudorapidez (η) e da profundidade do calorímetro (Francavilla; Collaboration et al., 2012).

O TileCal (Francavilla; Collaboration et al., 2012) é composto por três camadas longitudinais (A, BC, D), nas quais as células são agrupadas de forma a criar uma estrutura de torres projetivas. A granularidade é dada por $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ nas duas primeiras camadas e $\Delta\eta \times \Delta\phi = 0.2 \times 0.1$ na última camada. Uma camada especial adicional, chamada E, está conectada aos módulos do barril estendido e é lida por um único PMT para cada célula. No total, o TileCal possui 5182 células e 9852 canais de leitura, garantindo uma cobertura detalhada e precisa do detector (veja as figuras 12 e

13) (Francavilla; Collaboration et al., 2012).

Devido à leitura dupla da maioria das células, a falha de um único canal não compromete significativamente a medição de energia (Francavilla; Collaboration et al., 2012). No entanto, falhas mais sérias na eletrônica do detector, que não podem ser resolvidas remotamente, podem exigir a desativação da leitura de um módulo inteiro.

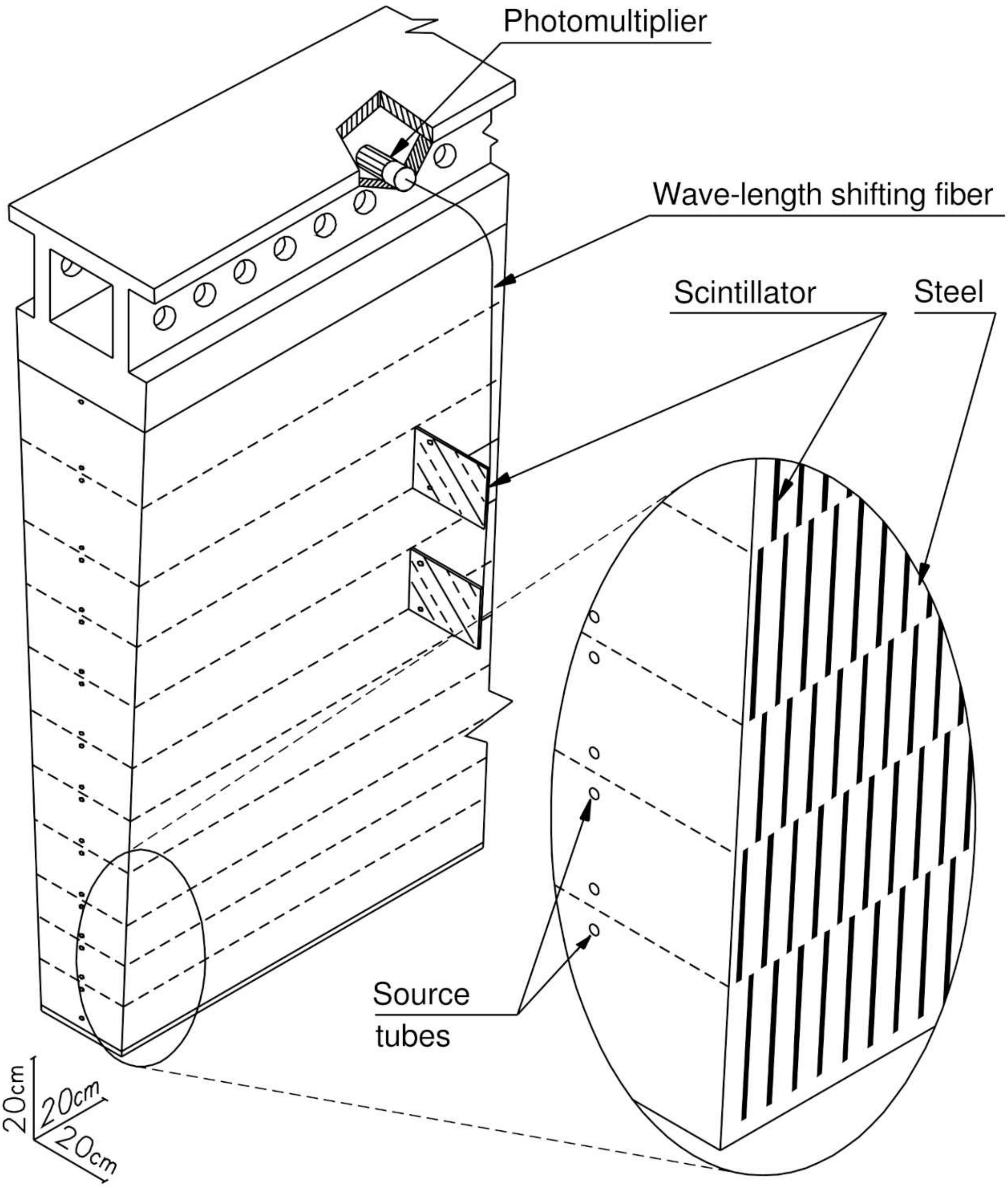
Nesse cenário, 22 células no barril e 18 células no barril estendido são mascaradas. A energia depositada nas células mascaradas é recuperada offline por meio de interpolação entre células vizinhas operacionais, o que é viabilizado pela alta granularidade espacial do calorímetro, assegurando uma boa precisão na reconstrução da energia (Francavilla; Collaboration et al., 2012).

O pulso rápido gerado na saída do fotomultiplicador (PMT) passa por um circuito de conformação (shaper), que ajusta sua forma, resultando em um pulso com amplitude proporcional à energia depositada no calorímetro (Anderson, 2005). Quando um hádron atravessa o calorímetro, ocorre uma cascata de partículas que perdem energia ao interagir com o material passivo (aço) e o material ativo (telhas cintilantes), resultando na emissão de luz. Essa luz é captada por fibras ópticas, que formam o sinal da célula. As fibras são agrupadas e direcionadas para o PMT, que converte a luz em um pulso elétrico (Gonçalves et al., 2023).

Esse pulso elétrico (Gonçalves et al., 2023) é padronizado pelo circuito modelador, que ajusta sua forma e garante que a amplitude seja proporcional à energia depositada. Em seguida, o pulso é convertido de analógico para digital por um Conversor Analógico-Digital (ADC), operando a uma taxa de amostragem de 40 MHz (Megahertz). O ADC captura o pulso em uma janela temporal de 150 ns (nanosegundos), gerando 7 amostras digitais que cobrem toda a duração do pulso (Anderson, 2005). A figura 14 mostra o pulso analógico e suas amostras representadas por pontos.

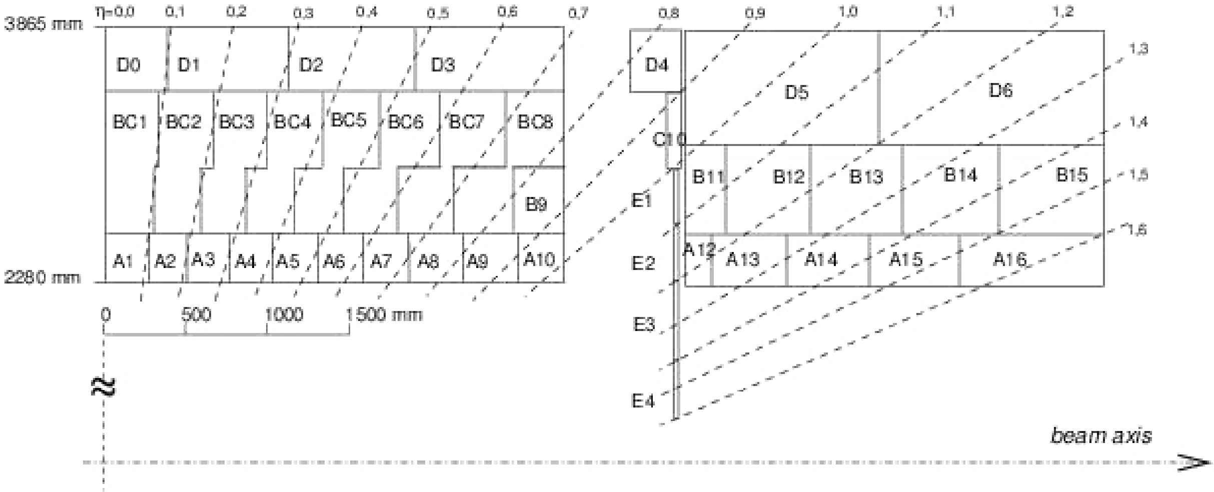
Na figura 14, o pedestal representa a linha de base do sinal, ou seja, o valor de referência a partir do qual a amplitude do sinal é medida. A amplitude corresponde à altura do sinal, medida a partir desse pedestal. A fase do sinal é determinada pela diferença de tempo entre a quarta amostra (amostra central) e o ponto de pico do pulso. Após essa coleta, as sete amostras são transmitidas via fibras ópticas para os Read Out Drivers (RODs), onde ocorre a estimativa da energia dos eventos que passaram pelo primeiro nível de filtragem, denominado nível 1 de disparo (LVL1), que seleciona os eventos mais relevantes para análise (Anderson, 2005). A seguir, será discutido o que garante a estabilidade e a confiabilidade na coleta de dados.

Figura 12 – Esquema do módulo do TileCal, mostrando 11 linhas radiais de azulejos agrupados em 3 camadas de leitura. A geometria pseudo-projetiva na pseudorapidez é obtida ao conectar fibras WLS a um único PMT.



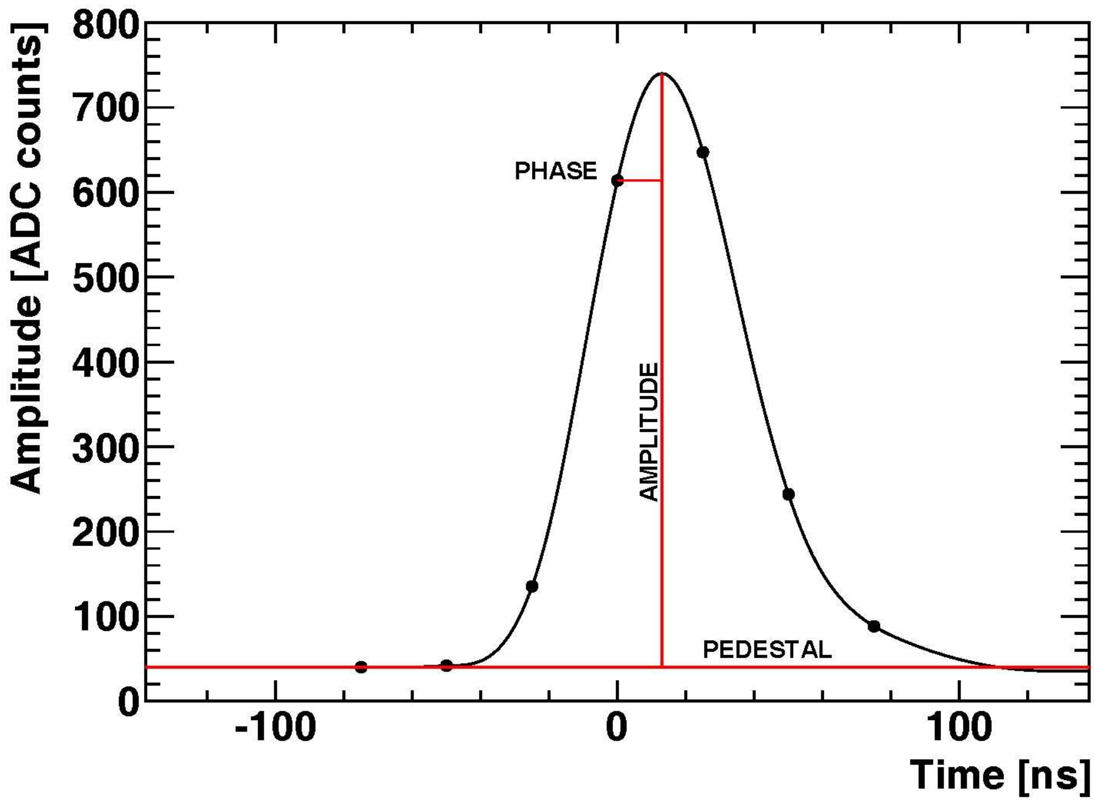
Fonte: Adaptado de Davidek (2022).

Figura 13 – Esquema de uma metade do Tile Calorimeter, com divisão em barril e barril estendido e representação da estrutura celular.



Fonte: Adaptado de Francavilla, Collaboration et al. (2012).

Figura 14 – Pulso de referência do TileCal.



Fonte: Adaptado de CERN (2023).

1.1.8 Sistema de Calibração do TileCal

O calorímetro TileCal utiliza diversos sistemas de calibração (como mostrado na figura 15) para garantir a maior estabilidade e confiabilidade na coleta dos dados provenientes das colisões de partículas. Cada sistema monitora um estágio diferente da cadeia de processamento do sinal, e são descritos brevemente abaixo.

- **Cesium:** O sistema de Cesium (Davidek, 2022) utiliza uma fonte radioativa de ^{137}Cs para medir o sinal que passa por todos os tiles do detector. O objetivo é calibrar o sistema óptico e os PMTs (tubos fotomultiplicadores), ajustando a resposta do canal. Em outras palavras, esse sistema tem como objetivo garantir a precisão das medições do detector durante o período de coleta de dados Run-2, ajustando o ganho dos PMTs para equalizar a resposta do canal (Davidek, 2022). Mudanças na resposta do ^{137}Cs são monitoradas com uma precisão de cerca de 0,3% utilizando constantes de calibração (C_{Cs}) (Davidek, 2022). O sistema também identifica e corrige a degradação dos componentes ópticos devido à radiação e as variações de ganho dos PMTs, especialmente nas áreas mais internas do detector (Davidek, 2022).

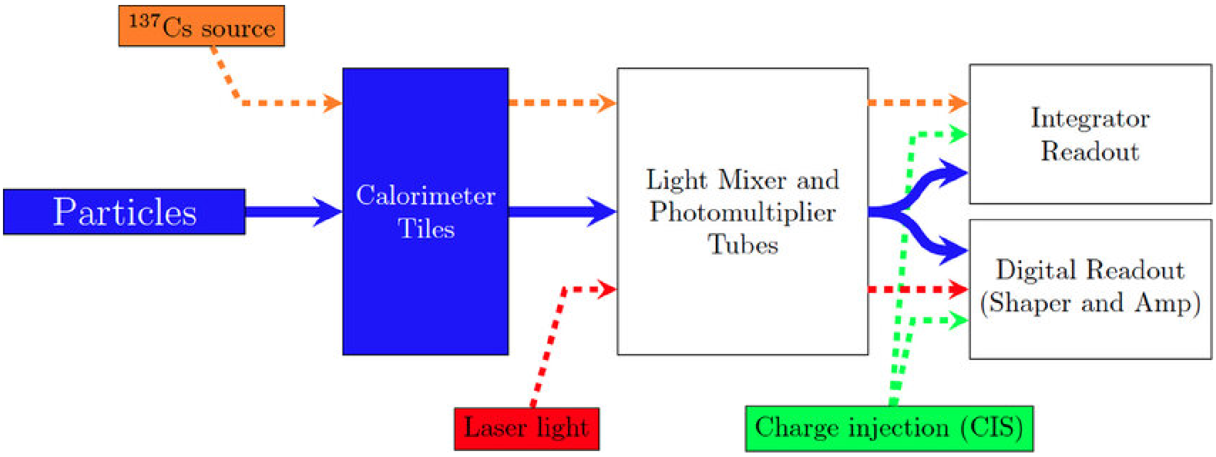
- **Laser:** Este sistema de calibração utiliza pulsos de laser curtos para monitorar e medir o ganho e a não linearidade dos PMTs (tubos fotomultiplicadores) (Davidek, 2022). Esses pulsos têm uma forma semelhante aos dados de colisão, permitindo verificar a resposta dos PMTs em condições comparáveis às de operação normal (Davidek, 2022). O objetivo é garantir a precisão do sistema de detecção, rastreando e ajustando a resposta dos PMTs. A calibração a laser fornece constantes (C_{Las}) que ajudam a avaliar a resposta dos PMTs em relação à calibração com o Cesium (^{137}Cs), com uma precisão de 0,5%. O sistema também monitora mudanças rápidas no ganho dos PMTs e a calibração temporal, ajudando a identificar e compensar variações de ganho, especialmente em regiões expostas a maiores doses de radiação (Davidek, 2022).

- **Injeção de Carga:** O sistema de injeção de carga (CIS) (Davidek, 2022) injeta pulsos de carga com especificações definidas na eletrônica de leitura rápida, cobrindo toda a faixa dinâmica de ambos os ganhos. Ele fornece um fator de conversão amplitude-carga (C_{CIS}) para cada canal e ganho, sendo utilizado para mapear não linearidades na eletrônica de leitura. O objetivo do CIS é garantir a precisão e a estabilidade do sistema de leitura, fornecendo um fator de conversão que permite a calibração da resposta dos canais em relação à carga aplicada (Davidek, 2022). Com uma precisão geral de aproximadamente 0,7% e uma estabilidade temporal muito boa (0,05% em canais individuais) durante o Run-2, o CIS ajuda a identificar e corrigir não linearidades na eletrônica de leitura, assegurando medições precisas ao longo do experimento (Davidek, 2022).

- **Mínimo Viés:** O Sistema de Mínimo Viés (Davidek, 2022) compartilha o mesmo caminho de leitura que o sistema de Cesium. Ele é usado para calibrar células especiais (E-cells) que não podem ser calibradas pela fonte de Cesium. Devido à sua natureza proporcional à luminosidade instantânea, o sistema de mínimo viés também é empregado para medir a luminosidade.

- **Calibração Temporal:** O objetivo da calibração temporal (Davidek, 2022) é garantir que os sinais das partículas sejam corretamente sincronizados em relação ao tempo, minimizando desvios temporais para uma fase $t_0 \approx 0$.

Figura 15 – Esquema do sistema de calibração do TileCal.



Fonte: Adaptado de ResearchGate ().

As descobertas empolgantes alcançadas nos primeiros anos de operação do LHC levaram a comunidade de física de altas energias a prever uma desafiadora atualização do acelerador, conhecida como LHC de alta luminosidade, ou, mais precisamente, High Luminosity LHC (HL-LHC) (Collaboration, 2010). Atualmente, esse processo de atualização do detector ATLAS está em andamento e encontra-se em sua segunda e última fase, como veremos a seguir.

1.1.9 HL-LHC

A atualização do detector ATLAS já está em andamento e encontra-se na Fase-II, conforme o cronograma do HL-LHC (Hasert, 1973). A atualização do LHC também está sendo implementada de forma gradual, com a Fase-I ocorrendo entre 2019 e 2020, e a Fase-II, que corresponderá à configuração final do HL-LHC, prevista para ser concluída entre 2023 e 2025 (Gonçalves et al., 2023).

Nesta fase, será necessária uma evolução na arquitetura eletrônica de leitura do ATLAS, permitindo que os dados sejam transmitidos diretamente para fora do detector, o que reduzirá a complexidade da eletrônica dentro do próprio detector (Gonçalves et al., 2023). A viabilização dessa mudança ocorrerá por meio do uso de links e receptores ópticos de alta qualidade e velocidade, permitindo a realocação de muitos desses componentes para fora do detector e a adoção do esquema de operação contínua (*free-running*) (Gonçalves et al., 2023).

Nesse novo cenário, os algoritmos online calcularão a energia de cada evento, e a eletrônica de leitura do TileCal será substituída, com todas as informações sendo transmitidas digitalmente para o sistema de filtragem (Trigger) do ATLAS (Gonçalves et al., 2023). Durante essa fase de alta luminosidade do LHC, espera-se que a máquina atinja luminosidades instantâneas na faixa de $5 \times 10^{34}\,\mathrm{cm^{-2}\,s^{-1}}$. A luminosidade pode ser obtida pela seguinte equação:

$$L \sim \frac{N^2}{t \cdot S} \tag{1}$$

considerando que o fator geométrico de redução da luminosidade é igual a 1. Nesta expressão, N representa o número de prótons em cada feixe, assumindo que cada partícula de um dos feixes pode colidir com qualquer partícula do outro feixe. O parâmetro t indica o tempo entre as colisões, enquanto S refere-se à seção transversal do feixe.

O objetivo da atualização é aumentar em cinco vezes a taxa de colisões instantâneas e multiplicar por dez a luminosidade integrada, em comparação com os valores nominais de design do LHC (Brüning et al., 2020). Em outras palavras, a configuração do HL-LHC baseia-se em novos modos de operação, como a operação com luminosidade nivelada e ajustes dinâmicos da ótica, além de várias tecnologias inovadoras e profundamente desafiadoras (Brüning et al., 2020).

Entre essas tecnologias, destacam-se: ímãs supercondutores de ponta, com intensidades entre 11 e 12 T (teslas); novos designs de ímãs, como os modelos *canted cosine theta* e super-férricos; cavidades supercondutoras de RF (radiofrequência) ultracompactas para rotação do feixe, com controle de fase de alta precisão; novas tecnologias e materiais para a colimação do feixe; e conexões supercondutoras de alta corrente com dissipação de energia quase nula (Brüning et al., 2020).

1.1.9.1 O Efeito de Empilhamento de Sinais

O formato do pulso gerado pelo TileCal se estende além do intervalo entre pacotes de feixes do LHC, o que resulta em uma janela de tempo maior. Em canais com alta taxa de ocupação, colisões consecutivas podem ocorrer dentro da mesma janela de leitura (Gonçalves et al., 2023). Em teoria, a informação mais relevante sobre o pulso é a sua altura máxima e o tempo em que atinge o pico, e uma única amostra nesse ponto seria suficiente para a medição de energia.

No entanto, variações no tempo do pulso introduzem imprecisões nas medições, o que afeta a amplitude proporcional do sinal. Além disso, devido ao longo tempo de decaimento do pulso, sinais de múltiplos cruzamentos consecutivos de feixes podem se sobrepor, agravando o problema (Gonçalves et al., 2023).

As colisões simultâneas de prótons, que ocorrem no mesmo cruzamento de feixe, intensificam ainda mais essa sobreposição de sinais, resultando no fenômeno conhecido como "empilhamento de sinal no tempo". Esse empilhamento distorce o sinal original, prejudicando a precisão na estimativa da amplitude.

A Figura 16 demonstra visualmente esse efeito de empilhamento. Como parte da atualização da Fase-II do LHC, a eletrônica de leitura dos calorímetros do ATLAS está sendo aprimorada para lidar com as operações em um ambiente de alta luminosidade, prevendo-se até 200 interações próton-próton simultâneas (Chiedde, 2022). Além disso, os sinais dos calorímetros são impactados pela sobreposição de até 25 colisões consecutivas, o que complica ainda mais a tarefa de reconstrução de energia (Chiedde, 2022).

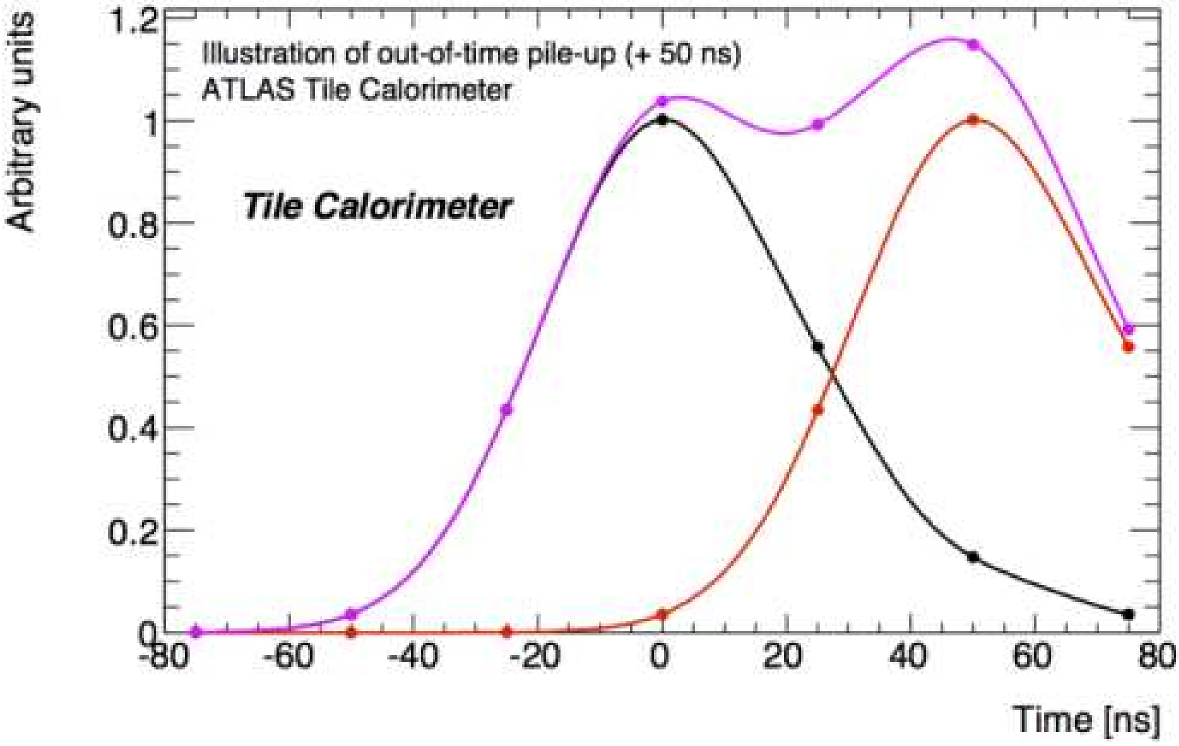
Com o aumento da luminosidade, o número de colisões também cresce devido à maior interação entre os prótons nos cruzamentos dos feixes, atribuída ao aumento no diâmetro do feixe de prótons. Embora essa alta luminosidade aumente a chance de detectar eventos físicos desejados, ela também eleva a probabilidade de múltiplos eventos ocorrerem no mesmo canal de leitura do sistema de calorimetria do ATLAS, intensificando o efeito de empilhamento de sinais (Gonçalves et al., 2023).

Nessas condições, o desempenho dos filtros digitais convencionais é prejudicado, pois a forma do pulso dos sinais se torna mais distorcida, dificultando sua análise precisa (Gonçalves et al., 2023). Além disso, é importante destacar que os sinais indesejados são compostos por duas parcelas: o ruído da eletrônica do equipamento de medição (distribuição gaussiana) e o efeito de empilhamento de sinais (distribuição lognormal) (Rimes et al., 2021).

1.2 Estimação da Energia

Em sistemas de detecção convencionais, os sinais capturados por eventos no detector passam inicialmente por um processo de preparação, onde são ajustados e modelados para facilitar seu tratamento. Em seguida, esses sinais, que inicialmente são analógicos, são convertidos para um formato digital por meio de circuitos eletrônicos localizados na etapa frontal do sistema. Após esse processamento inicial, os dados digitalizados são

Figura 16 – Efeito do empilhamento de sinal. O sinal desejado (preto) é detectado antes da chegada de um novo pulso. No entanto, simultaneamente, um sinal remanescente de uma colisão anterior (vermelho) também é registrado, causando a fusão dos dois sinais e distorcendo o resultado final (magenta).



Fonte: Adaptado de Klimek (2012).

enviados para a etapa seguinte, onde o sinal original é reconstruído para análise.

Para identificar informações importantes, como a amplitude exata do pulso gerado pelo detector, são aplicados algoritmos de filtragem digital, que removem ruídos e realçam as características desejadas. Em detectores, como os calorímetros, a amplitude obtida após essa filtragem é diretamente proporcional à energia medida no evento registrado (Gonçalves et al., 2023).

Essa sequência é essencial para sistemas que exigem alta precisão na reconstrução e análise de sinais, especialmente na física experimental, onde é crucial processar e analisar os sinais de forma eficiente para garantir a precisão das medições. Se considerarmos uma sequência de sinais digitais s no instante t, e levando em conta o efeito de desvios de fase e de linha de base no TileCal, cada amostra digital $\boldsymbol{s}$ no tempo t pode ser modelada como

$$s[k] = p + \boldsymbol{A}g\left[k + \tau\right] + n[k] \tag{2}$$

onde $\boldsymbol{s}[k]$ representa o conjunto de amostras, p corresponde ao pedestal, $\boldsymbol{A}$ refere-se à amplitude real do sinal, $\boldsymbol{g}[t]$ é o pulso de referência, k indica o instante de obtenção da amostra, $\boldsymbol{\tau}$ é a diferença de fase entre o pulso de referência e as amostras, e $\boldsymbol{n}[k]$ é a contribuição do ruído em cada amostra, para $k = 1, \ldots, M$, sendo M o número de eventos observados. Neste caso, consideramos pulsos de referência periódicos, cobrindo todos os instantes de tempo M.

Essa modelagem permite considerar tanto as variações esperadas na linha de base quanto os desvios temporais no sinal, o que é essencial para uma reconstrução precisa dos dados no calorímetro. A reconstrução dos sinais pode ser realizada por meio de modelos matemáticos, que permitem ajustar os parâmetros envolvidos para obter uma representação precisa do sinal verdadeiro.

O uso de modelos matemáticos ou computacionais permite simular e prever como um sistema se comportará sob diferentes condições ou cenários. Isso significa que podemos analisar como o sistema reage a variações em seus parâmetros. Os modelos matemáticos ou computacionais podem ser classificados em duas categorias principais: determinísticos e estocásticos, sendo os estocásticos mais complexos que os determinísticos. Além de serem classificados como determinísticos ou estocásticos, os modelos também podem ser categorizados quanto à linearidade, sendo divididos em modelos lineares e não lineares. Nesta dissertação, o nosso foco será nos modelos lineares.

2 MÉTODOS PROPOSTOS PARA ESTIMAÇÃO DE SINAL NO HL-LHC

O método atualmente empregado na reconstrução de sinais gerados pelas colisões de partículas no LHC é o Filtro Ótimo (OF, do inglês *Optimal Filtering*) na sua segunda versão, conforme descrito a seguir (Gonçalves et al., 2023). Este método busca minimizar a variância do ruído eletrônico presente nos calorímetros, aproveitando o conhecimento prévio da forma do pulso de referência (Gonçalves et al., 2023). O sinal de entrada é modelado como a soma de quatro componentes principais: o pulso de referência, o ruído, a fase do sinal e o pedestal (Gonçalves et al., 2023).

No contexto de baixa luminosidade, o método OF estima a amplitude do sinal por meio de uma soma ponderada das amostras do sinal. Os coeficientes do filtro OF são obtidos ao aproximar o sinal digital recebido por uma expansão de Taylor de primeira ordem. Essa abordagem garante que o estimador seja imparcial, ou seja, que seu valor esperado coincida com a amplitude real do sinal (Gonçalves et al., 2023), substituindo a equação 2 pela seguinte expressão:

$$\boldsymbol{A} = \widehat{\boldsymbol{A}}_{\boldsymbol{OF}} = \sum_{k=0}^{N-1} (pw[k] + \boldsymbol{A}g[k]w[k] - \boldsymbol{A}\tau g'[k]w[k] + n[k]w[k]) \tag{3}$$

Além disso, o método impõe restrições aos coeficientes para assegurar a imparcialidade em relação ao pedestal e à fase do sinal (Gonçalves et al., 2023).

2.1 OF 2

Considerando a equação 2, para que o estimador seja robusto em relação à fase e ao pedestal, são configuradas as seguintes restrições:

$$\sum_{k=0}^{N-1} g[k]w[k] = 1 \tag{4}$$

$$\sum_{k=0}^{N-1} g'[k]w[k] = 0 \tag{5}$$

$$\sum_{k=0}^{N-1} w[k] = 0 \tag{6}$$

As restrições descritas nas equações 5 e 6 garantem imunidade à fase e ao pedestal, respectivamente (Cardinot et al., 2024). Para calcular os pesos do estimador, busca-se minimizar sua variância, respeitando as restrições impostas. A variância do estimador é dada pela expressão:

$$\mathrm{Var}(\widehat{A}_{OF}) = \sum_{k=0}^{N-1} \sum_{j=0}^{N-1} w[k]w[j]C[k,j] = \boldsymbol{w}^\top \boldsymbol{C} \boldsymbol{w}, \tag{7}$$

onde $\boldsymbol{C}$ representa a matriz de covariância do ruído. Essa matriz descreve como as variações do ruído em diferentes amostras estão correlacionadas e é calculada como:

$$C[k,j] = \mathbb{E}\left\{(M[k] - \mathbb{E}\{M[k]\})(M[j] - \mathbb{E}\{M[j]\})\right\},$$

sendo $M[k]$ e $M[j]$ os valores do ruído nas amostras k e j, respectivamente, e $\mathbb{E}\{\cdot\}$ denotando o valor esperado.

A minimização da variância é conduzida utilizando o método dos multiplicadores de Lagrange, que combina as restrições (equações 4, 5 e 6) à expressão da variância, resultando na função objetivo:

$$\mathcal{L}(\boldsymbol{w}, \lambda, \epsilon, \kappa) = \boldsymbol{w}^\top \boldsymbol{C} \boldsymbol{w} - \lambda \left(\sum_{k=0}^{N-1} w[k]g[k] - 1 \right) - \epsilon \left(\sum_{k=0}^{N-1} w[k]g'[k] \right) - \kappa \left(\sum_{k=0}^{N-1} w[k] \right),$$

onde λ, ϵ e κ são os multiplicadores de Lagrange associados às restrições. Derivando as condições de otimização, obtém-se o seguinte sistema de equações:

$$\begin{aligned}
&\sum_{k=0}^{N-1} w[k]g[k] = 1, \qquad \sum_{k=0}^{N-1} w[k]g'[k] = 0, \qquad \sum_{k=0}^{N-1} w[k] = 0, \\
&\sum_{j=0}^{N-1} C[k,j]w[j] - \lambda g[k] - \epsilon g'[k] - \kappa = 0, \quad \forall k.
\end{aligned}$$

Esse sistema pode ser reorganizado na forma matricial:

$$\begin{pmatrix} C[1,1] & C[1,2] & \cdots & C[1,N] & -g[1] & -g'[1] & -1 \\ C[2,1] & C[2,2] & \cdots & C[2,N] & -g[2] & -g'[2] & -1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ C[N,1] & C[N,2] & \cdots & C[N,N] & -g[N] & -g'[N] & -1 \\ g[1] & g[2] & \cdots & g[N] & 0 & 0 & 0 \\ g'[1] & g'[2] & \cdots & g'[N] & 0 & 0 & 0 \\ 1 & 1 & \cdots & 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} w[1] \\ w[2] \\ \vdots \\ w[N] \\ \lambda \\ \epsilon \\ \kappa \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

A solução desse sistema fornece os pesos $w[k]$ que minimizam a variância do estimador enquanto satisfazem as restrições impostas.

2.2 OF 1

De forma análoga, considerando a equação 2, para que o estimador seja robusto em relação à fase, são configuradas as seguintes restrições:

$$\sum_{k=0}^{N-1} g[k]w[k] = 1$$

$$\sum_{k=0}^{N-1} g'[k]w[k] = 0$$

A minimização da variância é conduzida utilizando o método dos multiplicadores de Lagrange, combinando as restrições das equações 4 e 5 à expressão da variância, resultando na função objetivo:

$$\mathcal{L}(\boldsymbol{w}, \lambda, \epsilon) = \boldsymbol{w}^\top \boldsymbol{C}\boldsymbol{w} - \lambda \left(\sum_{k=0}^{N-1} w[k]g[k] - 1 \right) - \epsilon \left(\sum_{k=0}^{N-1} w[k]g'[k] \right)$$

onde λ e ϵ são os multiplicadores de Lagrange associados às restrições. Derivando as condições de otimização, obtém-se o seguinte sistema de equações:

$$\sum_{k=0}^{N-1} w[k]g[k] = 1 \qquad \sum_{k=0}^{N-1} w[k]g'[k] = 0 \qquad \sum_{k=0}^{N-1} w[k]C[k,j] - \lambda g[k] - \epsilon g'[k] = 0, \quad \forall k$$

Na forma matricial, o sistema acima pode ser representado da seguinte maneira:

$$\begin{pmatrix} C[1,1] & C[1,2] & \cdots & C[1,N] & -g[1] & -g'[1] \\ C[2,1] & C[2,2] & \cdots & C[2,N] & -g[2] & -g'[2] \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ C[N,1] & C[N,2] & \cdots & C[N,N] & -g[N] & -g'[N] \\ g[1] & g[2] & \cdots & g[N] & 0 & 0 \\ g'[1] & g'[2] & \cdots & g'[N] & 0 & 0 \end{pmatrix} \begin{pmatrix} w[1] \\ w[2] \\ \vdots \\ w[N] \\ \lambda \\ \epsilon \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

Em ambos os métodos apresentados, o objetivo principal é determinar o vetor de pesos para estimar a amplitude $\hat{A}_{OF}$ através da seguinte equação:

$$\widehat{\boldsymbol{A}}_{\boldsymbol{OF}} = \sum_{k=0}^{N-1} x[k]w[k]$$

onde $x[k]$ representa a matriz amostra, $w[k]$ representa o vetor de pesos do estimador e N é o número total de amostras utilizadas.

Vale destacar que o desempenho dos métodos OF 1 e OF 2 depende fortemente da matriz de covariância do ruído e de sua capacidade de estimar a amplitude em condições de baixa luminosidade. Por esse motivo, propomos métodos alternativos que consideram as atualizações do LHC, especificamente para o HL-LHC (Gonçalves et al., 2023).

2.3 Métodos de Mínimos Quadrados

A técnica dos mínimos quadrados (LS, do inglês Least Squares) (Kay, 1993) é amplamente utilizada para estimar um parâmetro escalar, que denotaremos por $\boldsymbol{\theta}$. Para aplicar essa abordagem, consideramos a seguinte relação:

$$s[n] = \boldsymbol{\theta} h[n] \tag{8}$$

em que n representa um índice que indica a posição de uma amostra em uma sequência temporal e $h[n]$ é uma sequência conhecida que serve como base para a estimativa. O objetivo é minimizar a diferença entre os dados observados $x[n]$ e os valores estimados $s[n]$ (Kay, 1993). Para isso, definimos um critério de erro, conhecido como a função de custo dos mínimos quadrados, dada por:

$$\boldsymbol{J(\theta)} = \sum_{n=0}^{M-1} (x[n] - \boldsymbol{\theta} h[n])^2 \tag{9}$$

Aqui, $\boldsymbol{\theta}$ quantifica o erro quadrático entre as observações e os valores estimados, e o nosso objetivo é encontrar o valor de $\boldsymbol{\theta}$ que minimiza essa função de custo.

A minimização dessa função leva ao cálculo do estimador dos mínimos quadrados (LSE, do inglês Least Squares Estimation), que pode ser obtido pela seguinte fórmula:

$$\boldsymbol{\theta} = \frac{\sum\limits_{n=0}^{M-1} x[n]h[n]}{\sum\limits_{n=0}^{M-1} h^2[n]} \tag{10}$$

Para determinar o erro mínimo associado à estimativa de $\boldsymbol{\theta}$, substituímos a equação (10) na função de custo (9). Realizando essa substituição, obtemos:

$$\boldsymbol{J_{min}} = \boldsymbol{J(\theta)} = \sum_{n=0}^{M-1} x^2[n] - \boldsymbol{\theta} \sum_{n=0}^{M-1} x[n]h[n] \tag{11}$$

Ao substituir $\boldsymbol{\theta}$ da equação (10) na equação (11), encontramos:

$$\boldsymbol{J_{min}} = \sum_{n=0}^{M-1} x^2[n] - \frac{\left(\sum\limits_{n=0}^{M-1} x[n]h[n]\right)^2}{\sum\limits_{n=0}^{M-1} h^2[n]} \tag{12}$$

Para estender essa abordagem a um conjunto de sinais $[s[0], s[1], s[2], \ldots, s[M-1]]$, que são lineares em relação aos parâmetros desconhecidos, utilizamos a notação matricial. Nesse caso, podemos expressar a relação entre os sinais como:

$$\boldsymbol{s} = \boldsymbol{H\theta} \tag{13}$$

em que $\boldsymbol{H}$ é uma matriz de observação conhecida, de dimensão $M \times p$, com M representando o número de observações e p o número de parâmetros, sendo $N > p$. Considerando um número arbitrário p de observações contínuas do sinal, é possível reorganizar o conjunto de amostras definindo $N = \frac{M}{p}$. Essa abordagem permite a criação de "janelas"de dados sequenciais, cada uma com um tamanho p (Gonçalves et al., 2023).

Na área da física de altas energias, a energia do sinal de interesse é estimada com base na amplitude do pulso registrado pelo sistema de calorimetria (Gonçalves et al., 2023). Para isso, consideramos um conjunto de M valores de amplitude, que podemos expressar como:

$$\boldsymbol{X} = (\boldsymbol{X_1}, \boldsymbol{X_2}, \ldots, \boldsymbol{X_j}, \ldots, \boldsymbol{X_M})^\top.$$

Esses valores de amplitude estão relacionados a eventos específicos observados durante os experimentos. Para facilitar a análise dos dados, podemos extrair um subconjunto de amplitudes, representado por:

$$\boldsymbol{x} = (x_1, x_2, \ldots, x_i, \ldots, x_N)^\top$$

em que cada x_i representa o valor central de cada janela de $\boldsymbol{X}$. Para obter essa estrutura, é necessário reorganizar a matriz de maneira adequada, resultando em uma matriz de dimensões $N \times p$. Esse processo de reorganização deve seguir um procedimento semelhante ao utilizado para compor a matriz $\boldsymbol{H}$, garantindo que os dados sejam organizados de modo a facilitar a aplicação de métodos de estimativa (Gonçalves et al., 2023).

Cada uma dessas janelas pode ser representada como uma linha de uma matriz determinística, que depende de um vetor de parâmetros desconhecidos, dado por:

$$\boldsymbol{\theta} = \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_k \\ \vdots \\ \theta_p \end{pmatrix}$$

para que a relação

$$\boldsymbol{s} = \boldsymbol{H\theta}$$

seja válida. O LSE, ou seja, os pesos que buscamos, são encontrados minimizando a seguinte equação:

$$\begin{aligned} \boldsymbol{J}(\boldsymbol{\theta}) &= \sum_{n=0}^{M-1} (x[n] - s[n])^2 \\ &= (\boldsymbol{x} - \boldsymbol{H}\boldsymbol{\theta})^T (\boldsymbol{x} - \boldsymbol{H}\boldsymbol{\theta}) \end{aligned} \tag{14}$$

Para encontrar os valores ótimos de $\boldsymbol{\theta}$, realizamos a derivada parcial da função de custo $\boldsymbol{J(\theta)}$ em relação a $\boldsymbol{\theta}$:

$$\frac{\boldsymbol{\partial J(\theta)}}{\boldsymbol{\partial \theta}} = -2\boldsymbol{H}^T\boldsymbol{x} + 2\boldsymbol{H}^T\boldsymbol{H}\boldsymbol{\theta} \tag{15}$$

Igualando a equação (15) a zero e isolando $\boldsymbol{\theta}$, obtemos a fórmula:

$$\boldsymbol{\theta} = \left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1} \boldsymbol{H}^T\boldsymbol{x} \tag{16}$$

em que $\boldsymbol{x}$ representa o vetor das observações (o pulso de referência). Os estimadores dos parâmetros são então obtidos por meio da equação (16), e a amplitude estimada pode ser calculada como:

$$\boldsymbol{\hat{A}} = \boldsymbol{H}\boldsymbol{\theta} \tag{17}$$

Se assumirmos que o parâmetro $\boldsymbol{\theta}$ está sujeito a uma restrição linear, sendo $r < p$, é necessário que essas restrições sejam independentes. Para garantir essa independência, devemos evitar redundâncias, como as expressões $\theta_1 + \theta_2 = 0$ e $2\theta_1 + 2\theta_2 = 0$, onde θ_i representa elementos de $\boldsymbol{\theta}$. Portanto, (Kay, 1993) assumimos que a restrição tem a forma:

$$\boldsymbol{A}\boldsymbol{\theta} = \boldsymbol{b} \tag{18}$$

em que $\boldsymbol{A}$ é uma matriz conhecida de dimensões $r \times p$, e $\boldsymbol{b}$ é um vetor de dimensões $r \times 1$. Para ilustrar, considere o caso em que $p = 2$ e um dos parâmetros é conhecido por ser o oposto do outro. Nesse caso, a restrição será:

$$\theta_1 + \theta_2 = 0$$

o que implica que $\boldsymbol{A} = [1 \quad 1]$ e $\boldsymbol{b} = 0$. Além disso, assumimos sempre que a matriz $\boldsymbol{A}$ possui posto completo (igual a r), condição necessária para assegurar que as restrições são de fato independentes. Consequentemente, no problema de mínimos quadrados com restrições (CLE, do inglês Constrained Least Squares), existem apenas $(p-r)$ parâmetros independentes.

Para resolver o problema de mínimos quadrados com restrições lineares, aplicamos a técnica dos multiplicadores de Lagrange. Nosso objetivo é determinar $\boldsymbol{\theta_r}$ (onde o subíndice r indica a solução LSE com restrição) minimizando o Lagrangiano:

$$\boldsymbol{J_r} = (\boldsymbol{x} - \boldsymbol{H\theta})^T(\boldsymbol{x} - \boldsymbol{H\theta}) + \lambda^T(\boldsymbol{A\theta} - \boldsymbol{b}) \tag{19}$$

onde λ é um vetor de multiplicadores de Lagrange com dimensões $r \times 1$. Expandindo a expressão em (19), obtemos:

$$\boldsymbol{J_r} = \boldsymbol{x}^T\boldsymbol{x} - 2\boldsymbol{\theta}^T\boldsymbol{H}^T\boldsymbol{x} + \boldsymbol{\theta}^T\boldsymbol{H}^T\boldsymbol{H\theta} + \lambda^T\boldsymbol{A\theta} - \lambda^T\boldsymbol{b}. \tag{20}$$

Calculando a derivada parcial de (20) em relação a $\boldsymbol{\theta}$, temos:

$$\frac{\partial \boldsymbol{J_r}}{\partial \boldsymbol{\theta}} = -2\boldsymbol{H}^T\boldsymbol{x} + 2\boldsymbol{H}^T\boldsymbol{H\theta} + \boldsymbol{A}^T\lambda. \tag{21}$$

Igualando (21) a zero, obtemos:

$$\begin{aligned} \boldsymbol{\theta_r} &= (\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^T\boldsymbol{x} - \frac{1}{2}(\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{A}^T\lambda \\ &= \boldsymbol{\theta} - (\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{A}^T\frac{\lambda}{2}, \end{aligned} \tag{22}$$

onde $\boldsymbol{\theta}$ é a solução LSE sem restrição (definida na equação (16)) e λ ainda precisa ser determinada. Para encontrar λ, aplicamos a restrição definida na equação (18), resultando em:

$$\boldsymbol{A\theta_r} = \boldsymbol{A\theta} - \boldsymbol{A}(\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{A}^T\frac{\lambda}{2} = \boldsymbol{b},$$

o que leva a:

$$\frac{\lambda}{2} = \left[\boldsymbol{A}(\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{A}^T\right]^{-1}(\boldsymbol{A\theta} - \boldsymbol{b}). \tag{23}$$

Substituindo (23) em (22), obtemos a solução final para $\boldsymbol{\theta_r}$:

$$\boldsymbol{\theta}_r = \boldsymbol{\theta} - (\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{A}^T\left[\boldsymbol{A}(\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{A}^T\right]^{-1}(\boldsymbol{A}\boldsymbol{\theta} - \boldsymbol{b}). \tag{24}$$

Portanto, a solução LSE com restrição pode ser vista como uma versão ajustada da solução sem restrição. Se a restrição for satisfeita por $\boldsymbol{\theta}$, ou seja, $\boldsymbol{A}\boldsymbol{\theta} = \boldsymbol{b}$, então os estimadores obtidos serão idênticos. A estimativa da amplitude é dada por:

$$\hat{\boldsymbol{A}}_r = \boldsymbol{H}\boldsymbol{\theta}_r. \tag{25}$$

Por fim, o erro associado à estimativa tanto para OF 1, OF 2, LS e como para CLS é dado por:

$$\boldsymbol{e} = \boldsymbol{x} - \hat{\boldsymbol{A}} \tag{26}$$

Assim, ao longo deste processo, conseguimos não apenas estimar o parâmetro desejado, mas também avaliar o erro na estimativa, que é crucial para a análise de desempenho em aplicações práticas. A Tabela 1 mostra as informações necessárias para a aplicação de cada um dos métodos.

Tabela 1 – Resumo das características dos métodos analisados neste trabalho.

Método	Dados das amostras e amplitudes	Coluna de viés	Matriz da covariância	Pulso de referência
LS	✓	✓		
CLS	✓	✓		
OF 1	✓		✓	✓
OF 2	✓		✓	✓

Fonte: O autor, 2025.

3 METODOLOGIA

3.1 Descrição dos Dados

Neste estudo, foram gerados dados simulados de colisões de partículas utilizando o *Pulse Generator* (Gonçalves, 2022), uma ferramenta desenvolvida em Python. Essa ferramenta permite configurar diversos parâmetros das colisões, como o tempo de amostragem, a amplitude original, o pulso de referência e a fase (Cardinot et al., 2024).

Os dados foram estruturados no formato *free-running*, que consiste em uma sequência contínua de amostras coletadas sem interrupções. Essa abordagem, que difere das técnicas tradicionais baseadas em gatilhos ou sincronizações específicas, está alinhada com as demandas da nova fase do LHC, o HL-LHC. Esse formato exige métodos específicos para identificar e analisar os eventos registrados, refletindo as condições operacionais introduzidas pela atualização do acelerador (Cardinot et al., 2024).

Foram gerados 11 conjuntos de dados, cada um com níveis de ocupação variando de 0% a 100%, com incrementos de 10%. O nível de ocupação representa a probabilidade de ocorrência de um sinal verdadeiro em um dado instante, associado a uma colisão, ou seja, ao número médio de interações próton-próton por evento, resultando em um sinal detectável (Gonçalves et al., 2023).

Para abordar o problema utilizando os conjuntos de dados descritos, empregaremos a equação 2, assumindo inicialmente que $\boldsymbol{\tau} = 0$. Contudo, isso levanta uma questão importante: o que aconteceria se o pedestal da equação 2 não fosse fixo, mas sim variável, refletindo uma situação mais próxima da realidade? Para explorar essa possibilidade, utilizaremos os mesmos conjuntos de dados mencionados, mas com a adição de um novo vetor de dados $\boldsymbol{y}$ às amostras, a fim de modelar a variação do pedestal.

O processo de adição de um novo vetor de dados $\boldsymbol{y}$ às amostras será realizado da seguinte forma:

1. Para cada linha da matriz de amostras, é gerado um valor aleatório utilizando uma distribuição uniforme dentro de um intervalo predefinido, representado por $(100, 1000)$. Esse intervalo determina os valores mínimo e máximo que o vetor $\boldsymbol{y}$ pode assumir.

2. O vetor $\boldsymbol{y}$ gerado é somado a todas as colunas da matriz de amostras, com exceção da última coluna. Assim, cada linha da matriz modificada contém as amostras originais acrescidas do valor correspondente de $\boldsymbol{y}$, enquanto a última coluna é preservada inalterada.

Este procedimento visa introduzir variações adicionais aos dados, simulando condições experimentais mais realistas e ampliando a diversidade dos cenários avaliados. O processo assegura que a matriz original das amostras seja mantida inalterada, com as modificações aplicadas em uma cópia dos dados originais.

Em seguida, aplicaremos os métodos LS e CLS para a reconstrução do sinal utilizando os conjuntos de dados gerados, tanto para os dados originais quanto para os dados ajustados, conforme as alterações introduzidas pelo vetor $\boldsymbol{y}$. Dessa forma, investigaremos como a variação do pedestal impacta os resultados finais.

Cada conjunto contém 2.000.000 de amostras, representando colisões simuladas a cada 25 ns. A amplitude dos sinais segue uma distribuição exponencial com média de 100 ADC Counts. A fase é distribuída uniformemente no intervalo de -5 a 5 ns, enquanto o pedestal é fixado em 30 ADC Counts. Adicionalmente, é introduzida uma pequena deformação de 0,01 ADC Count para simular o desgaste dos componentes eletrônicos.

3.2 Técnica de Janelamento

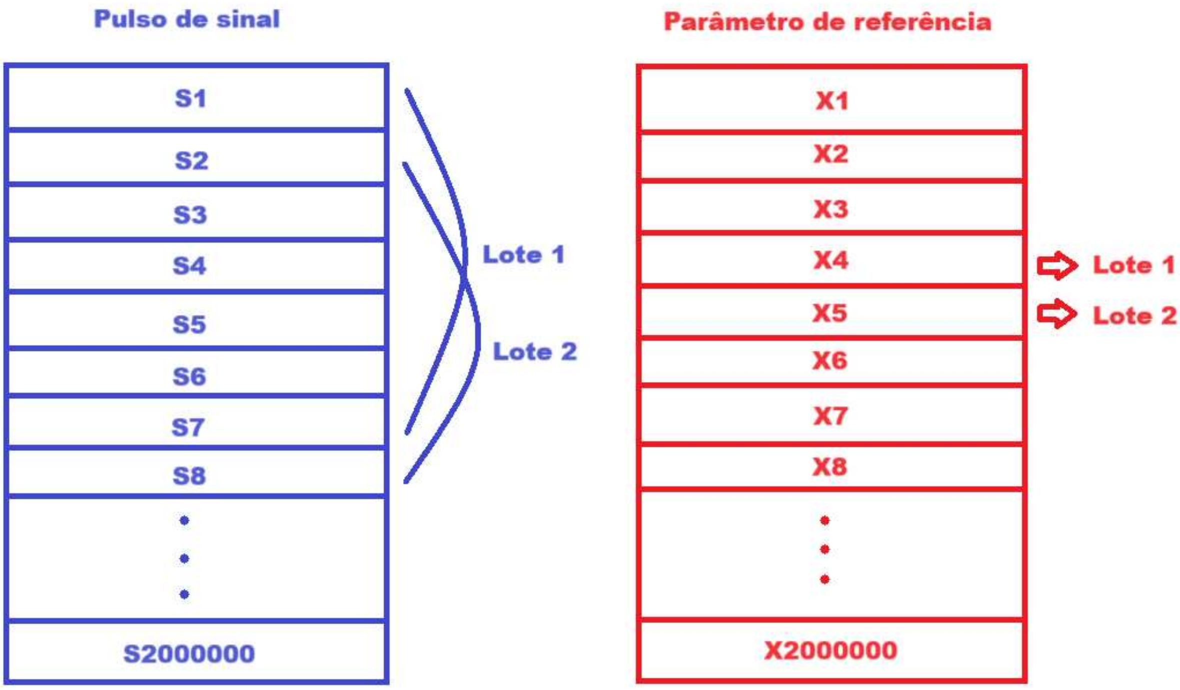
Antes da atualização do LHC, os dados eram organizados em um formato tabelar, no qual cada pulso era representado por uma linha e suas propriedades distribuídas em várias colunas (geralmente sete) (Cardinot et al., 2024). Esse formato facilitava a aplicação de métodos de estimação dos parâmetros, permitindo uma comparação direta entre a amplitude estimada e a real de cada pulso.

No entanto, após a atualização, o formato dos dados passou por uma mudança significativa: agora, todos os valores são dispostos sequencialmente em uma única coluna contínua (*free running*), ao longo de várias linhas, como representado na Figura 17. Essa nova estrutura exige uma abordagem diferente para o processamento, na qual é necessário agrupar sinais adjacentes em "janelas"de dados.

Essas “janelas” representam segmentos contínuos de dados, cujo tamanho define o número de amostras utilizadas na reconstrução do sinal. O objetivo dessa abordagem é investigar como o tamanho das janelas afeta o desempenho na reconstrução do sinal.

Para isso, adotamos janelas com tamanhos definidos por números inteiros ímpares variando de 7 a 19, com incrementos de 2, o que permite capturar de forma eficaz a amplitude central dos sinais.A escolha de janelas com um número ímpar de amostras é estratégica, pois facilita a identificação e comparação dos parâmetros centrais do sinal reconstruído, que servem como base para avaliar a precisão do método (Insali et al., 2024). Além disso, o tamanho máximo de janelamento foi fixado em 19, dentro de um limite linear imposto pelas restrições no número de pesos calculados. Essas restrições garantem que a implementação prática seja viável, respeitando as limitações do experimento.

Figura 17 – Estratégia de janelamento dos dados gerados conforme a atualização do LHC, dividindo-os em 7 janelas e representando o parâmetro de referência.



Fonte: Adaptado de Insali et al. (2024).

3.3 Técnica de validação cruzada k-Fold

A técnica de validação cruzada k-Fold é um método amplamente utilizado para avaliar a performance de modelos preditivos em problemas de aprendizado de máquina. Ela funciona dividindo os dados em k subconjuntos, ou "folds", como exemplificado na Figura 18, de tamanho aproximadamente igual. Em cada uma das k iterações, o modelo é treinado em $k-1$ dos subconjuntos e testado no subconjunto restante, que não foi utilizado no treino.

Esse processo é repetido de forma que cada subconjunto sirva como conjunto de teste exatamente uma vez, garantindo que cada dado seja usado tanto para treino quanto para teste ao longo das iterações (Wong, 2015). Valores menores de k aumentam a quantidade de dados para treinamento, mas podem reduzir a precisão da estimativa do erro de generalização, já que menos dados ficam disponíveis para a avaliação (Anguita et al., 2012).

Ao fim das k iterações, calcula-se a média das métricas de avaliação (como acurácia, erro médio, etc.), resultando em uma medida de performance mais estável e representativa. Essa abordagem reduz a variabilidade causada pela divisão dos dados, melhora a generalização do modelo e aproveita ao máximo os dados disponíveis. A validação cruzada k-Fold também facilita comparações mais precisas entre diferentes modelos ou configurações, pois permite uma avaliação robusta e confiável do desempenho (Wong, 2015).

3.4 Critérios para a avaliação dos métodos

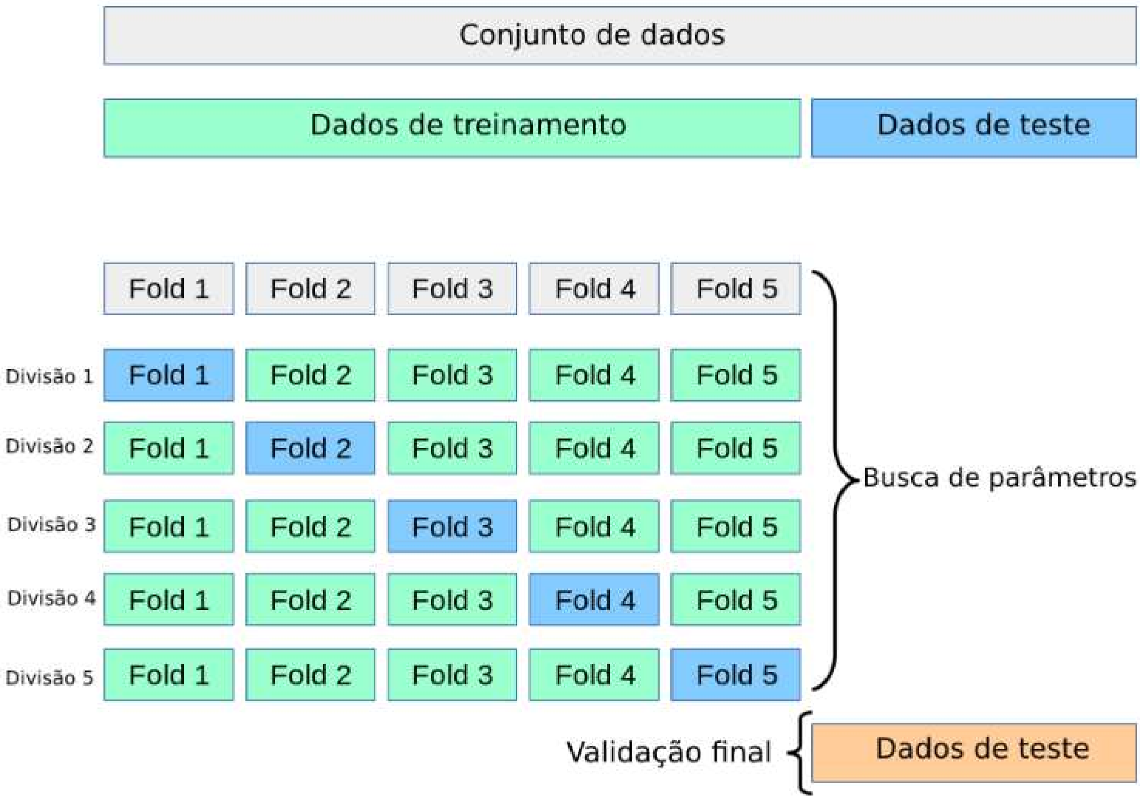
Para avaliar a eficácia dos dois métodos de estimativa de sinal, utilizaremos métricas que destacam o desempenho em cenários variados. Especificamente, observaremos o desvio padrão das amplitudes estimadas como um indicador da consistência dos métodos em diferentes contextos e a média do erro de estimação além do desvio padrão.

O desvio padrão (Bussab; Morettin, 2010) mede a dispersão dos valores estimados em torno de sua média, sendo um indicativo da consistência das estimativas. Em Python, o desvio padrão de uma amostra é calculado usando `np.std()` com `ddof=1`, onde `ddof` (graus de liberdade) é definido como 1 para indicar que estamos calculando o desvio padrão amostral (NumPy Developers, 2024). A fórmula do desvio padrão amostral (s) é dada por:

$$s = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2} \tag{27}$$

onde N é o número de valores estimados, x_i é o valor estimado individual e $\bar{x}$ é

Figura 18 – Exemplificação da técnica de validação cruzada k-Fold com $k = 5$.



Fonte: Adaptado de Insali et al. (2024).

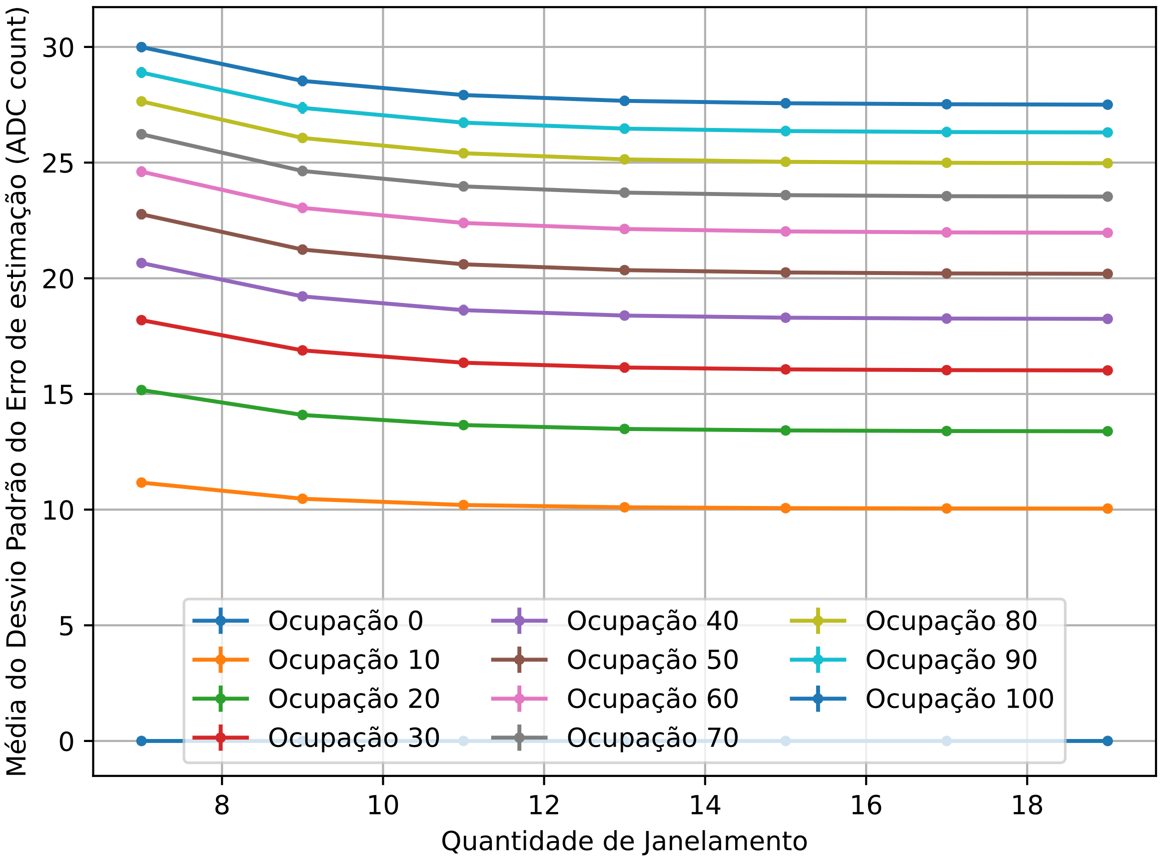
a média dos valores estimados. O erro de estimação da amplitude, que reflete a precisão dos métodos em relação aos valores esperados, será obtido de acordo com a Equação 26.

4 RESULTADOS E DISCUSSÃO

4.1 Resultados obtidos com o pedestal fixo

Os resultados apresentados a seguir foram obtidos com o pedestal fixo. O objetivo inicial é determinar o janelamento ideal para a reconstrução da energia. Esse janelamento será escolhido com base na média do desvio padrão do erro de estimação da energia.

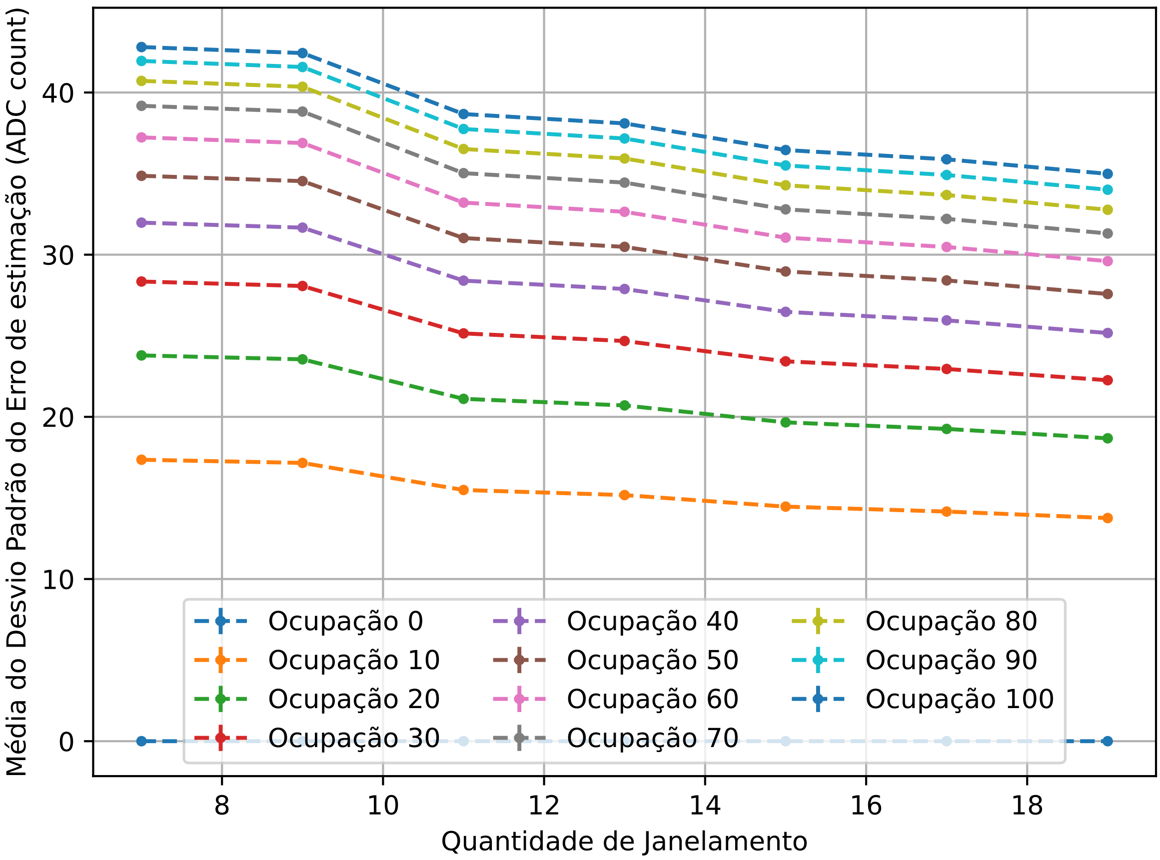
Figura 19 – Média do desvio padrão do erro de estimação em função da quantidade de janelamento LS, considerando todos os cenários de ocupação.



Fonte: O autor, 2025.

Na Figura 19, observa-se que, em todos os cenários de ocupação, a média do desvio padrão do erro de estimação se estabiliza a partir do janelamento 15, mantendo-se constante até o janelamento 19. Isso corresponde à menor média do desvio entre as janelas avaliadas, indicando que não há diferenças significativas nos resultados obtidos ao utilizar os janelamentos 15, 17 ou 19. Portanto, os janelamentos ideais para aplicação são 15, 17 e 19. No entanto, considerando o custo computacional, o janelamento 19 se apresenta como a opção mais eficiente para a reconstrução da energia com este método.

Figura 20 – Média do desvio padrão do erro de estimação em função da quantidade de janelamento CLS, considerando todos os cenários de ocupação.

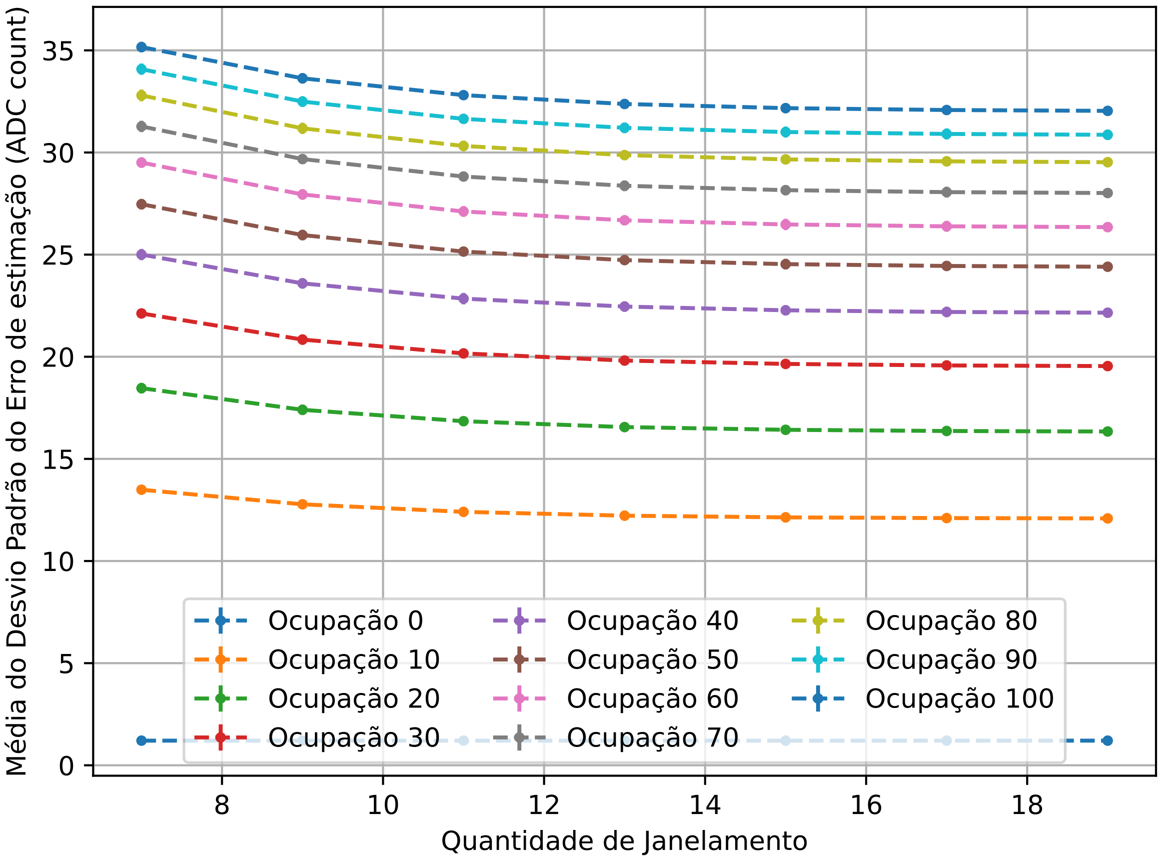


Fonte: O autor, 2025.

A Figura 20 apresenta a eficiência do método CLS na reconstrução da energia, representada pela média do desvio padrão do erro de estimação em função da quantidade de janelamento, considerando todos os cenários de ocupação. Nota-se que, quanto maior a quantidade de janelamento, menor é a média do desvio padrão do erro de estimação, resultando em maior precisão. Assim, conclui-se que o janelamento ideal para o método CLS é 19.

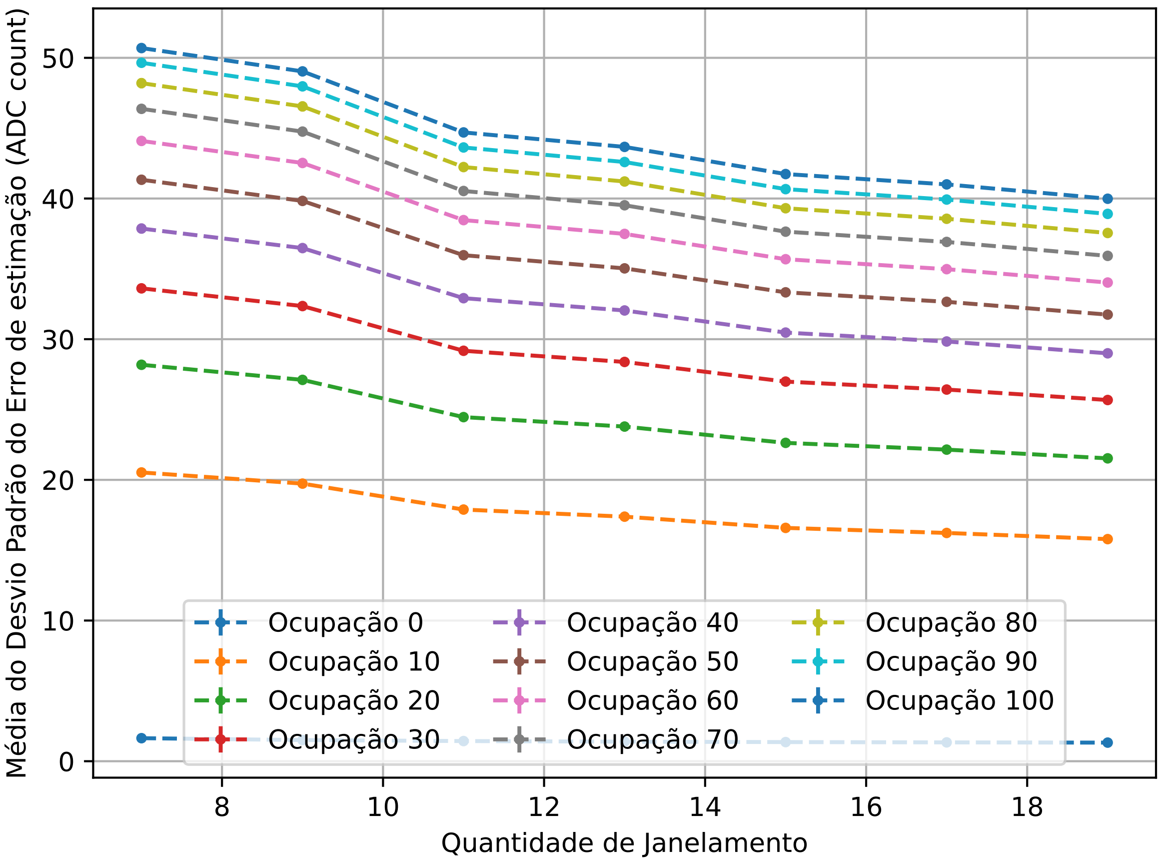
De maneira similar, a Figura 21 mostra a eficiência do método OF 1 na reconstrução da energia, representada pela média do desvio padrão do erro de estimação em função do número de janelamentos, levando em conta todos os cenários de ocupação. Verifica-se que a média do desvio padrão do erro atinge um valor mínimo e se mantém constante entre os janelamentos 17 e 19. Assim, conclui-se que os janelamentos ideais para o método OF 1 são 17 e 19. Considerando o custo computacional, o janelamento 17 se destaca como a opção mais eficiente para a reconstrução da energia com este método.

Figura 21 – Média do desvio padrão do erro de estimação em função da quantidade de janelamento OF 1, considerando todos os cenários de ocupação.



Fonte: O autor, 2025.

Figura 22 – Média do desvio padrão do erro de estimação em função da quantidade de janelamento OF 2, considerando todos os cenários de ocupação.



Fonte: O autor, 2025.

Por fim, a Figura 22 apresenta a eficiência do método OF 2 na reconstrução da energia, representada pela média do desvio padrão do erro de estimação em função da quantidade de janelamento, considerando todos os cenários de ocupação. Nota-se que, quanto maior a quantidade de janelamento, menor é a média do desvio padrão do erro de estimação, resultando em maior precisão. Assim, conclui-se que o janelamento ideal para o método OF 2 é 19.

Tabela 2 – Janelamentos ideais para diferentes métodos de reconstrução de energia.

Método	Janelamento ideal
LS	15, 17 e 19
CLS	19
OF 1	17 e 19
OF 2	19

Fonte: O autor, 2025.

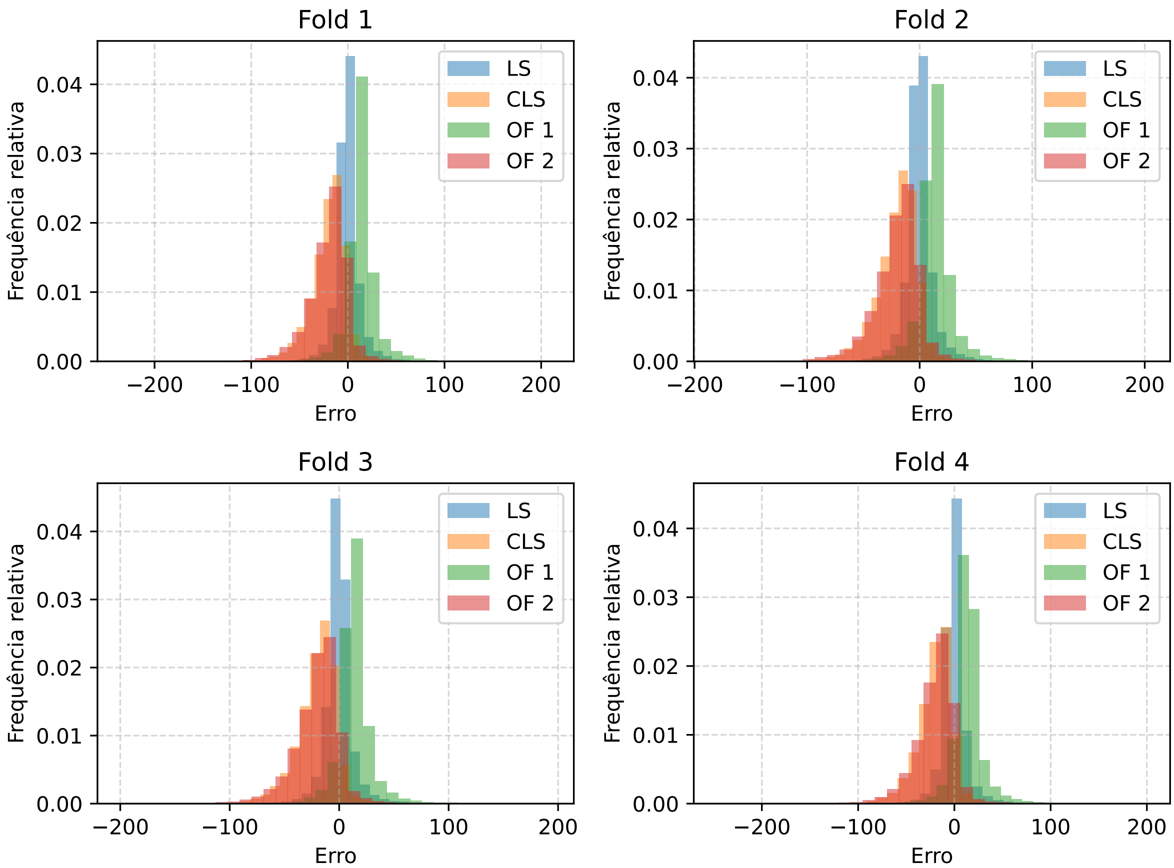
A Tabela 2 apresenta os janelamentos ideais para todos os métodos analisados. Nota-se que, em todos os casos, o janelamento 19 está entre as opções ideais. Por esse motivo, a partir deste ponto, a análise comparativa dos métodos será realizada exclusivamente com base no janelamento 19. Para isso, consideramos dois cenários distintos de níveis de ocupação (20% e 50%) para a análise dos histogramas do erro.

As Figuras 23 e 24 revelam diferenças marcantes na distribuição dos erros entre os métodos analisados nos primeiros quatro folds. O método LS apresenta o histograma mais concentrado em torno de zero, indicando maior precisão relativa em comparação com os demais métodos. O método OF 1, embora menos preciso que o LS, ainda demonstra um bom desempenho, com uma concentração considerável próxima de zero. Em contraste, os histogramas dos métodos CLS e OF 2 são os mais dispersos, sugerindo maior variabilidade nos erros e, consequentemente, menor precisão em relação aos outros métodos.

A Figura 25, referente ao último k-fold, demonstra que o método LS apresentou o melhor desempenho em termos de desvio padrão do erro para todos os níveis de ocupação analisados. O segundo melhor desempenho foi observado com o método OF 1, seguido pelo CLS e, por último, pelo OF 2, que apresentou o maior desvio padrão.

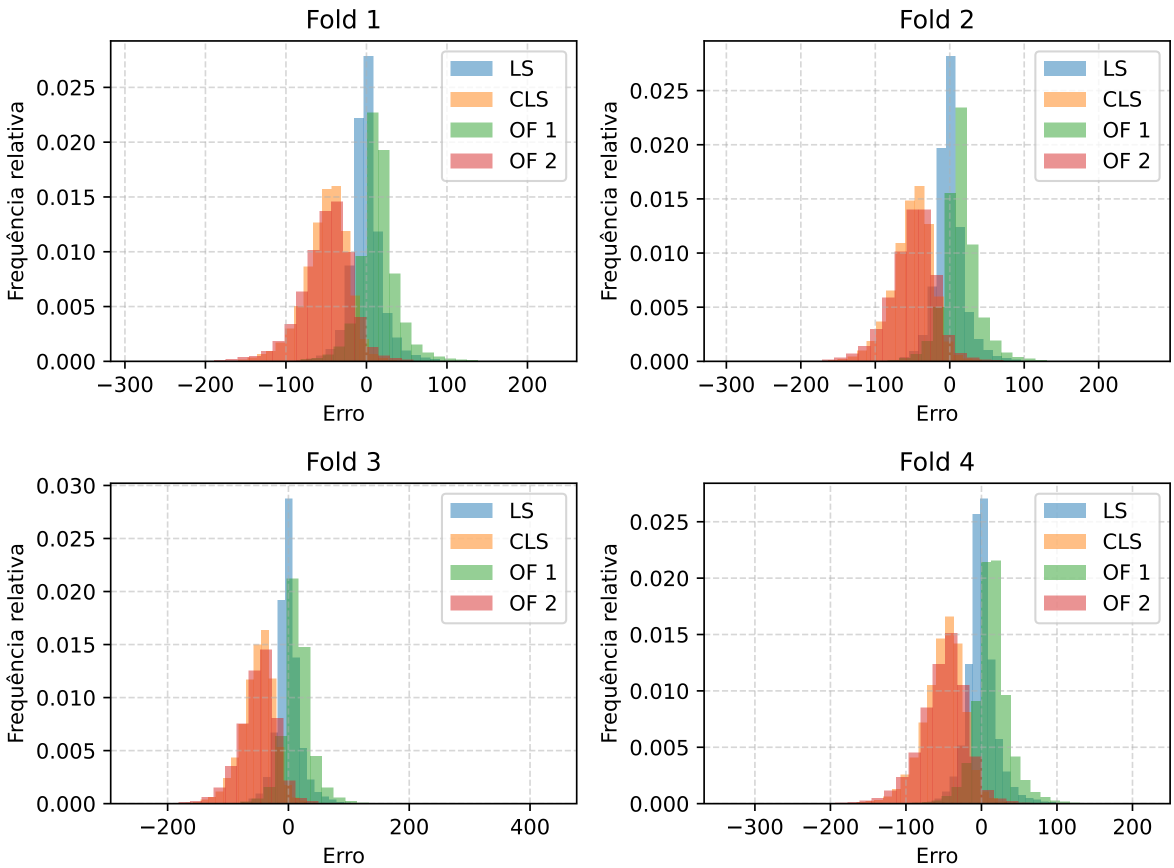
Já a Figura 26 apresenta os erros de estimativa para cada método em função da ocupação. Nota-se que os métodos CLS e OF 2 exibem comportamentos quase idênticos, com erros significativamente distantes de zero. O método OF 1, por sua vez, apresenta erro menor em comparação com o CLS e OF 2 a partir de 20% de ocupação. Mais uma vez, o método LS se destaca, com erros consistentemente próximos de zero ao longo de todas as ocupações. A seguir, veremos os resultados considerando modificações no pedestal.

Figura 23 – Erros de estimativa dos métodos LS, CLS, OF 1 e OF 2, representados pelos histogramas em azul, laranja, verde e vermelho, respectivamente, para 20% de ocupação.



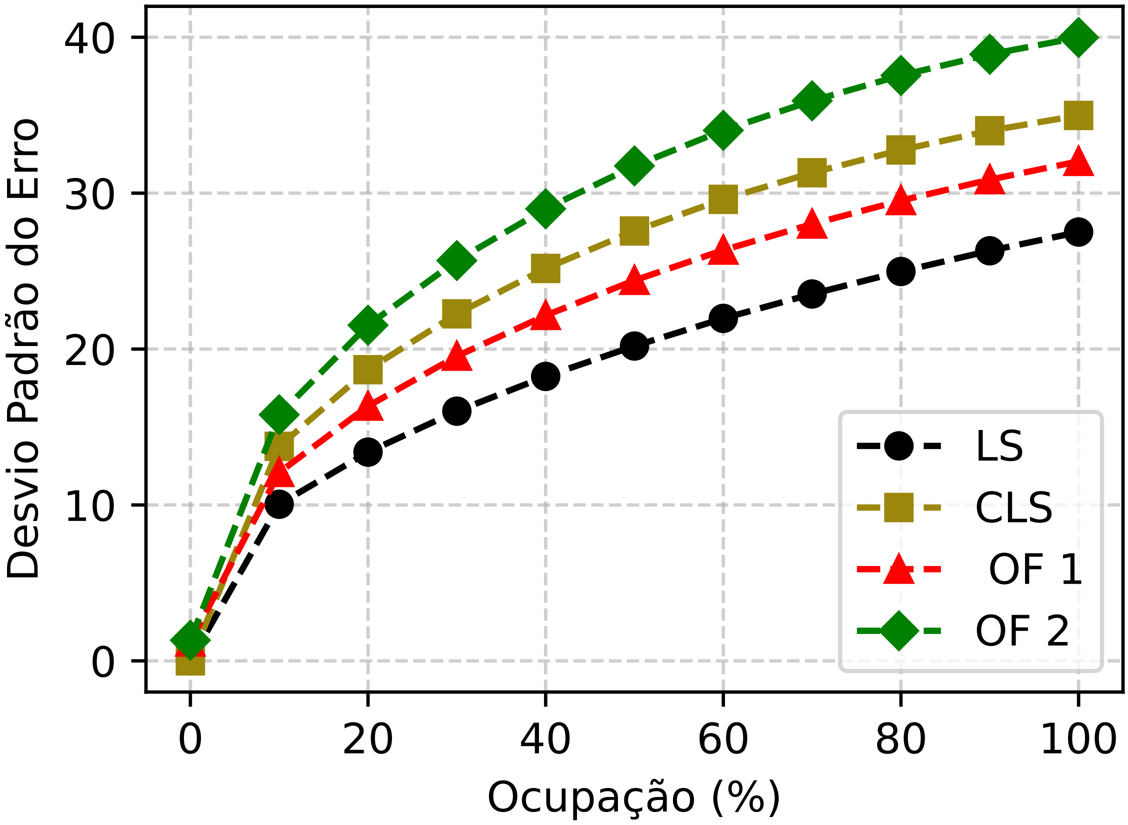
Fonte: O autor, 2025.

Figura 24 – Erros de estimativa dos métodos LS, CLS, OF 1 e OF 2, representados pelos histogramas em azul, laranja, verde e vermelho, respectivamente, para 50% de ocupação.



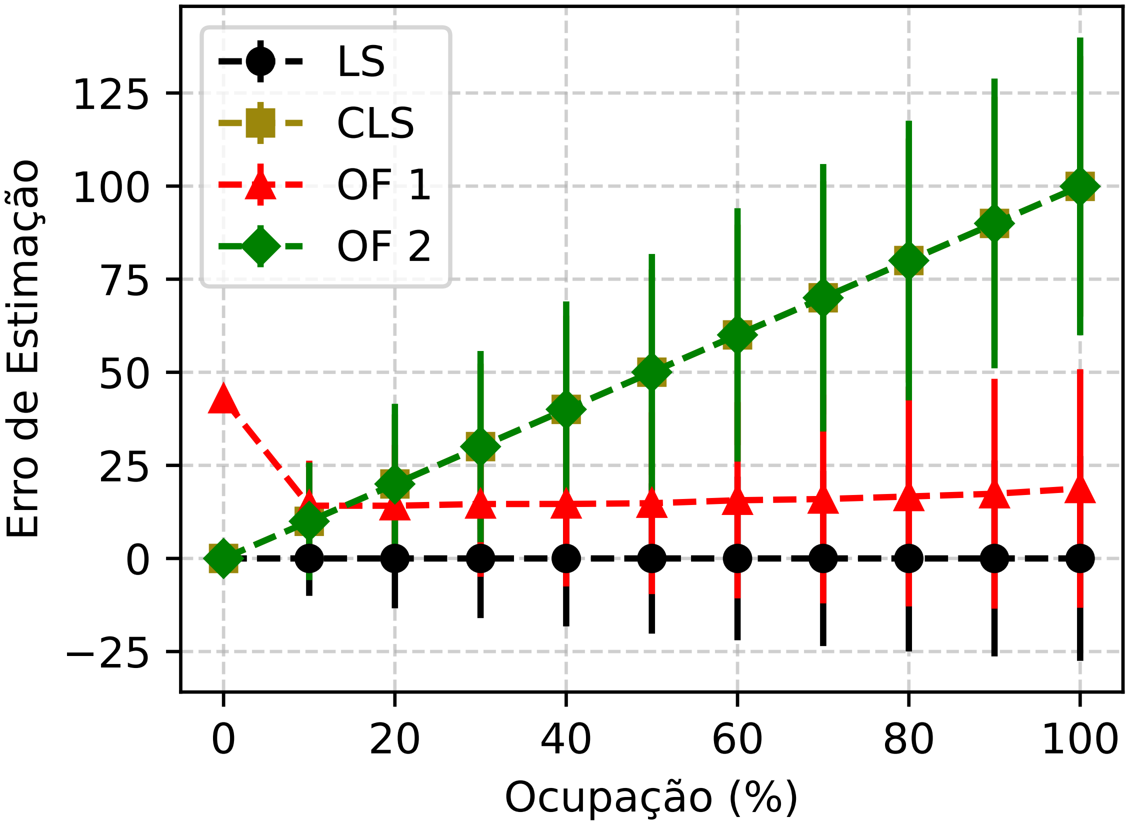
Fonte: O autor, 2025.

Figura 25 – Desvio padrão por ocupação usando os métodos LS, CLS, OF 1 e OF 2, representados pelas curvas em preto, amarelo, vermelho e verde, respectivamente.



Fonte: O autor, 2025.

Figura 26 – Erros por ocupação usando os métodos LS, CLS, OF 1 e OF 2, representados com linhas tracejadas em preto, amarelo, vermelho e verde, respectivamente.



Fonte: O autor, 2025.

4.2 Resultados obtidos com o pedestal variando

Os resultados foram obtidos pela incorporação de um novo vetor de dados, $\boldsymbol{y}$, às amostras, conforme descrito na seção 3.1. O objetivo principal é determinar o janelamento ideal para a reconstrução da energia, definido com base na análise da média do desvio padrão do erro de estimação.

A figura 27 apresenta a média do desvio padrão do erro de estimação obtido pelo método LS em diferentes cenários de ocupação. Diferentemente do comportamento observado na figura 19, onde o desvio padrão varia de forma mais pronunciada, nesta análise o desvio padrão se estabiliza a partir da janela 15, mantendo-se constante até a janela 19. No entanto, a figura 27 também evidencia que o método LS não é capaz de lidar adequadamente com as variações no pedestal, limitando sua eficácia nesses casos.

A figura 28 ilustra a eficiência do método CLS na reconstrução da energia, avaliada por meio da média do desvio padrão do erro de estimação em função do número de janelas, considerando todos os cenários de ocupação. Observa-se que, à medida que o número de janelas aumenta, a média do desvio padrão do erro de estimação diminui, indicando uma melhora na precisão. Com base nesses resultados, conclui-se que o janelamento ideal para o método CLS é a janela 19.

Além disso, ao comparar as figuras 28 e 20, verifica-se que o método CLS apresenta desempenho praticamente equivalente tanto para dados sem variação no pedestal quanto para dados com variação no pedestal, demonstrando sua robustez em ambos os casos.

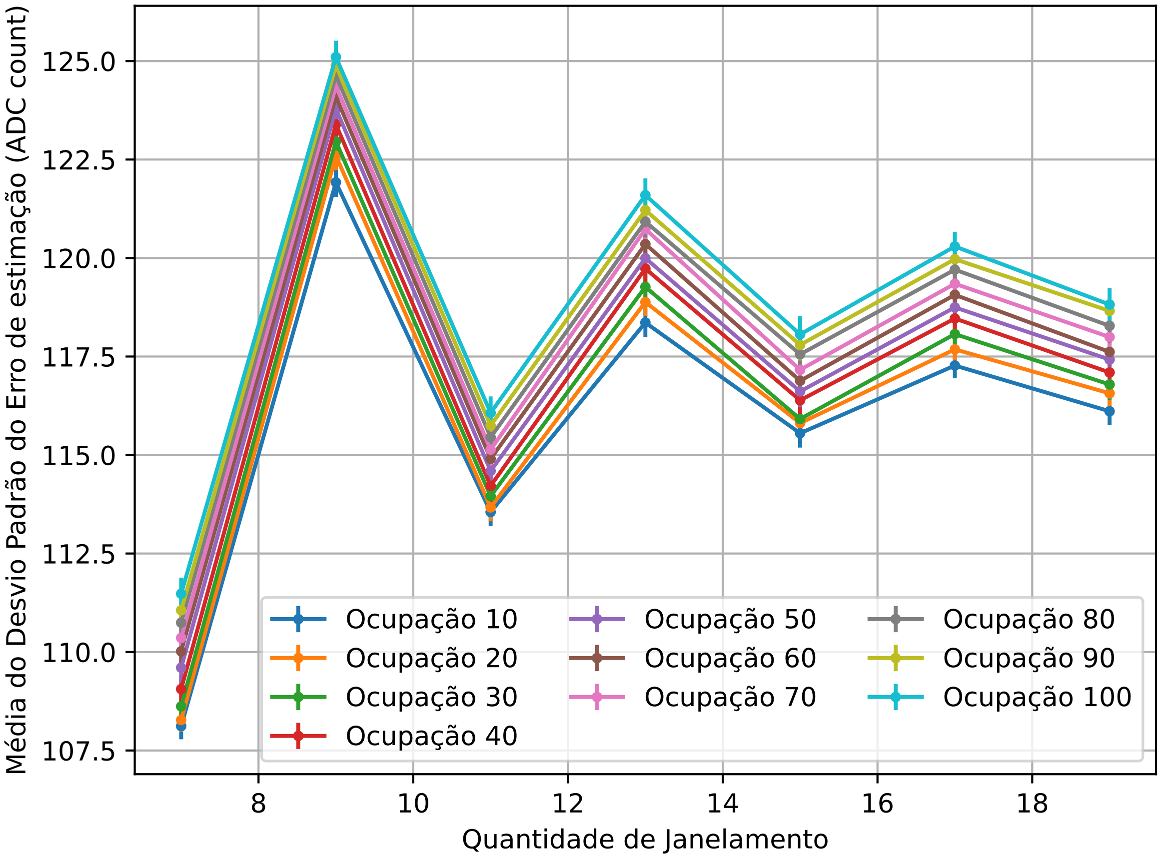
A figura 29 mostra a média do desvio padrão do erro de estimação obtido pelo método LS em diferentes cenários de ocupação. Ao contrário da figura 21, que utiliza dados sem variação no pedestal, a figura 29 destaca que o método OF 1 não consegue lidar de forma eficaz com as variações no pedestal, o que compromete seu desempenho nesses cenários.

Por fim, a figura 30 ilustra a eficiência do método OF 2 na reconstrução da energia, representada pela média do desvio padrão do erro de estimação em função do número de janelas, considerando todos os cenários de ocupação. Nota-se que, conforme o número de janelas aumenta, a média do desvio padrão do erro de estimação diminui, resultando em maior precisão. Dessa forma, conclui-se que o janelamento ideal para o método OF 2 é a janela 19.

Além disso, ao comparar as figuras 30 e 22, observa-se que o desempenho do método OF 2 é praticamente o mesmo, tanto para dados com variação no pedestal quanto para dados sem variação, o que destaca sua robustez em ambos os cenários.

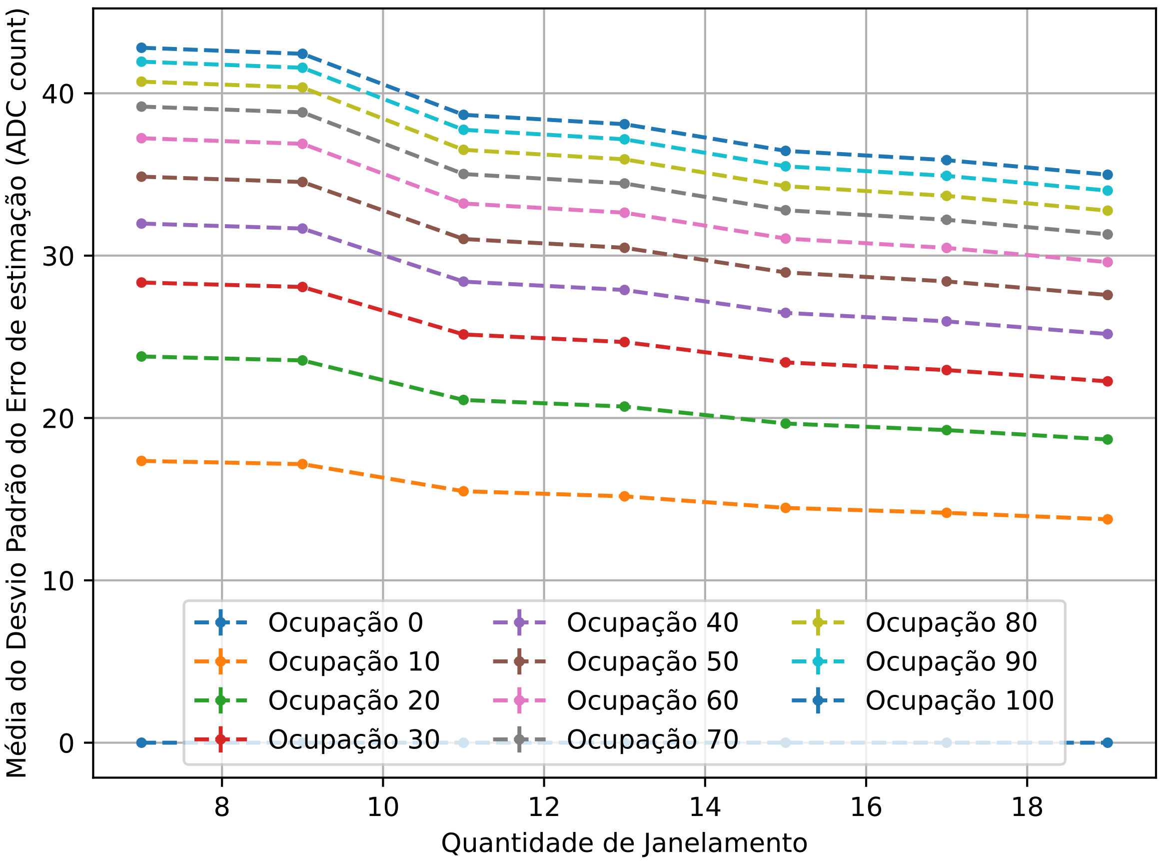
Com base nas análises realizadas usando as figuras 27, 28, 29 e 30, podemos observar que os métodos LS e OF 1 apresentaram resultados insatisfatórios no que diz respeito à reconstrução da energia com variação do pedestal, o que demonstra que esses métodos não são ideais para a reconstrução da energia quando o pedestal não é fixo.

Figura 27 – Média do desvio padrão do erro de estimação em função da quantidade de janelamento LS, considerando todos os cenários de ocupação, exceto a ocupação 0.



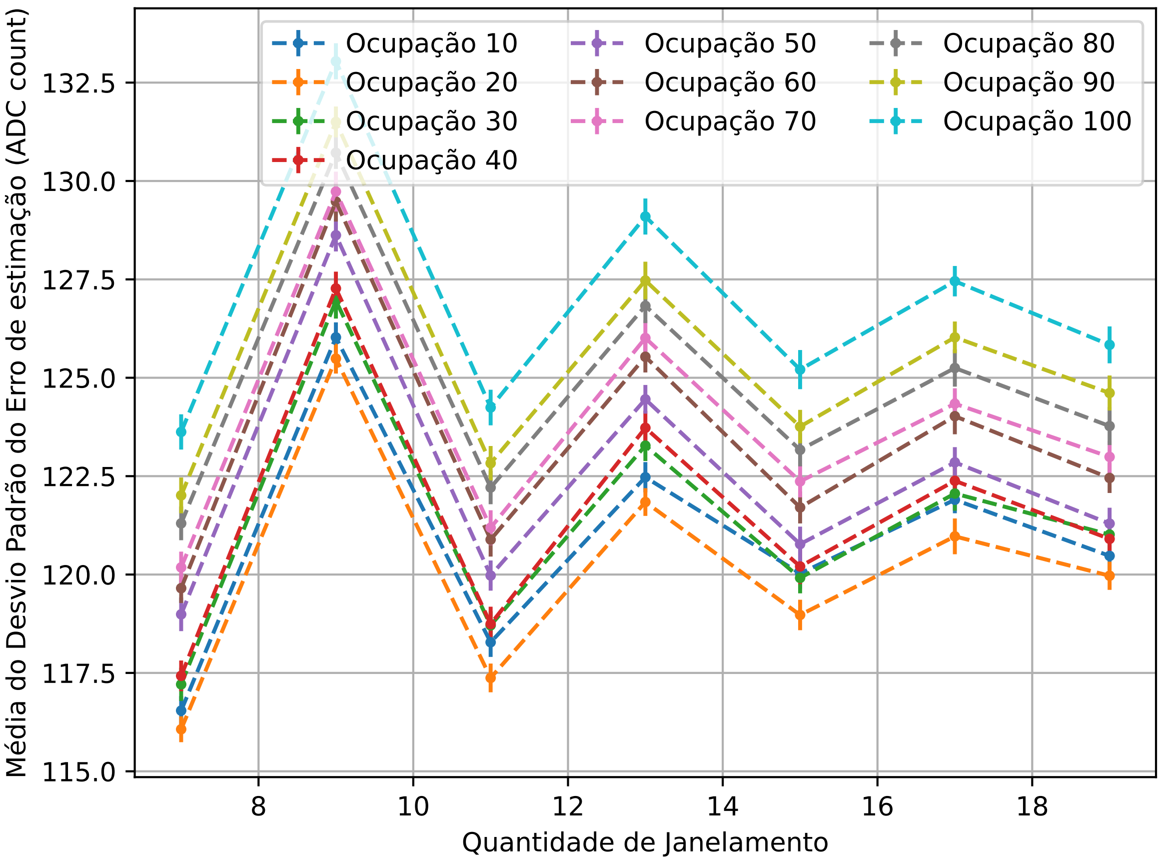
Fonte: O autor, 2025.

Figura 28 – Média do desvio padrão do erro de estimação em função da quantidade de janelamento CLS, considerando todos os cenários de ocupação.



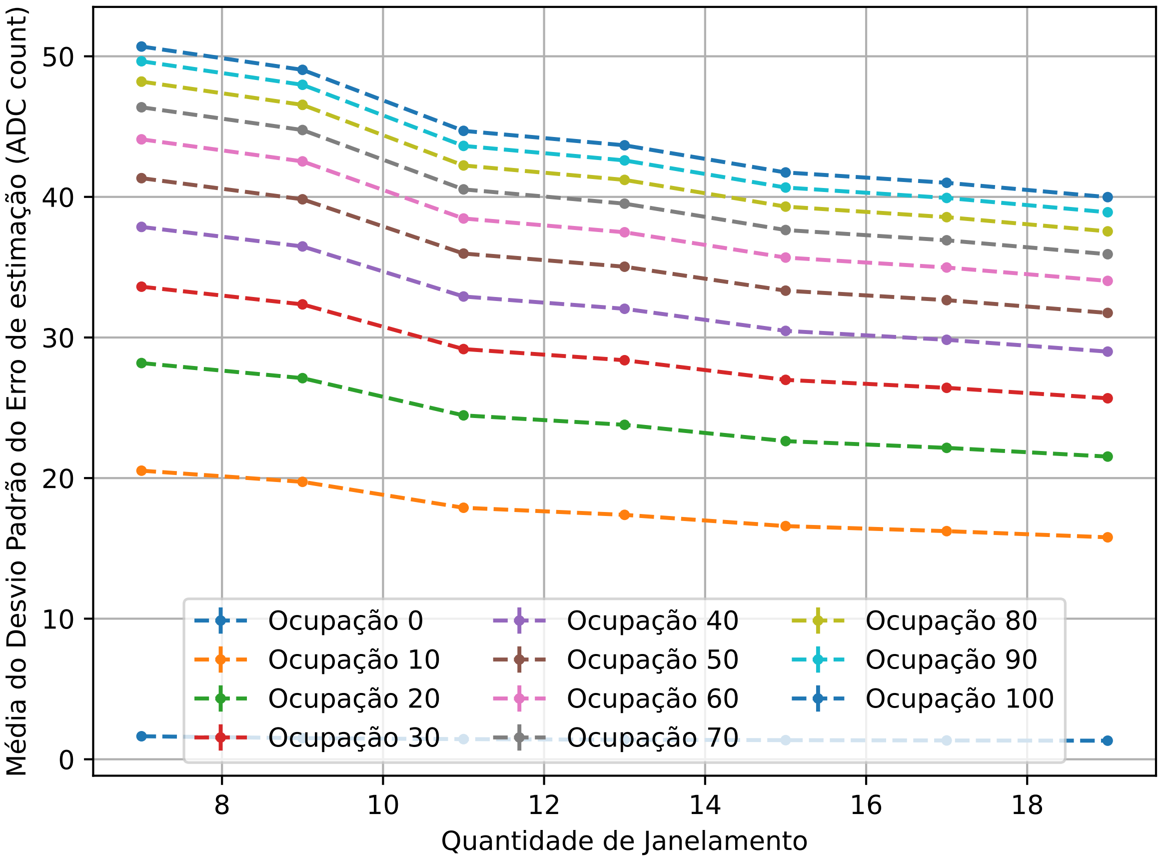
Fonte: O autor, 2025.

Figura 29 – Média do desvio padrão do erro de estimação em função da quantidade de janelamento OF 1, considerando todos os cenários de ocupação, exceto a ocupação 0.



Fonte: O autor, 2025.

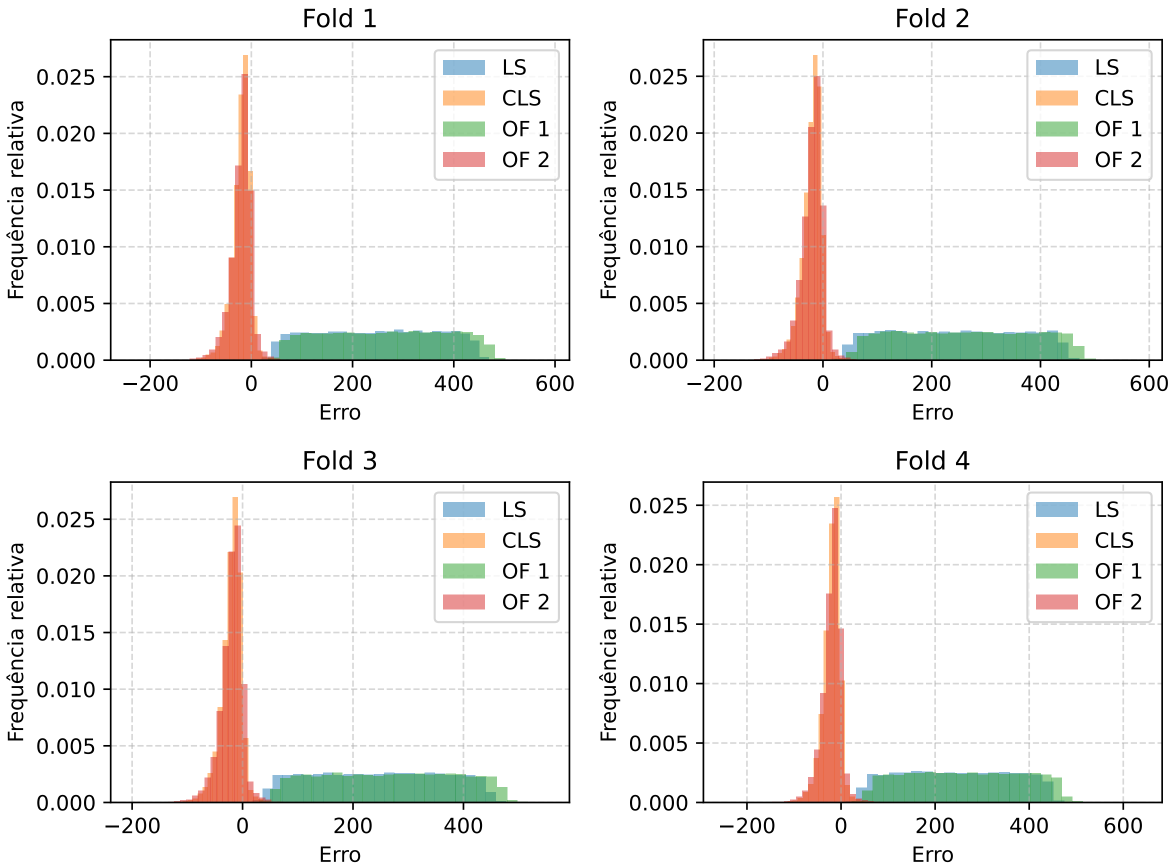
Figura 30 – Média do desvio padrão do erro de estimação em função da quantidade de janelamento OF 2, considerando todos os cenários de ocupação.



Fonte: O autor, 2025.

Em contrapartida, os métodos CLS e OF 2 demonstraram um desempenho significativamente melhor nesse contexto, com variação do pedestal, e mantiveram bom desempenho com o janelamento ideal de 19. Diante disso, concluímos que, a partir deste ponto, a análise comparativa entre os métodos será realizada exclusivamente com o janelamento 19, adotando dois cenários distintos de níveis de ocupação (20% e 50%) para a análise dos histogramas de erro.

Figura 31 – Erros de estimativa dos métodos LS, CLS, OF 1 e OF 2, representados pelos histogramas em azul, laranja, verde e vermelho, respectivamente, para 20% de ocupação.

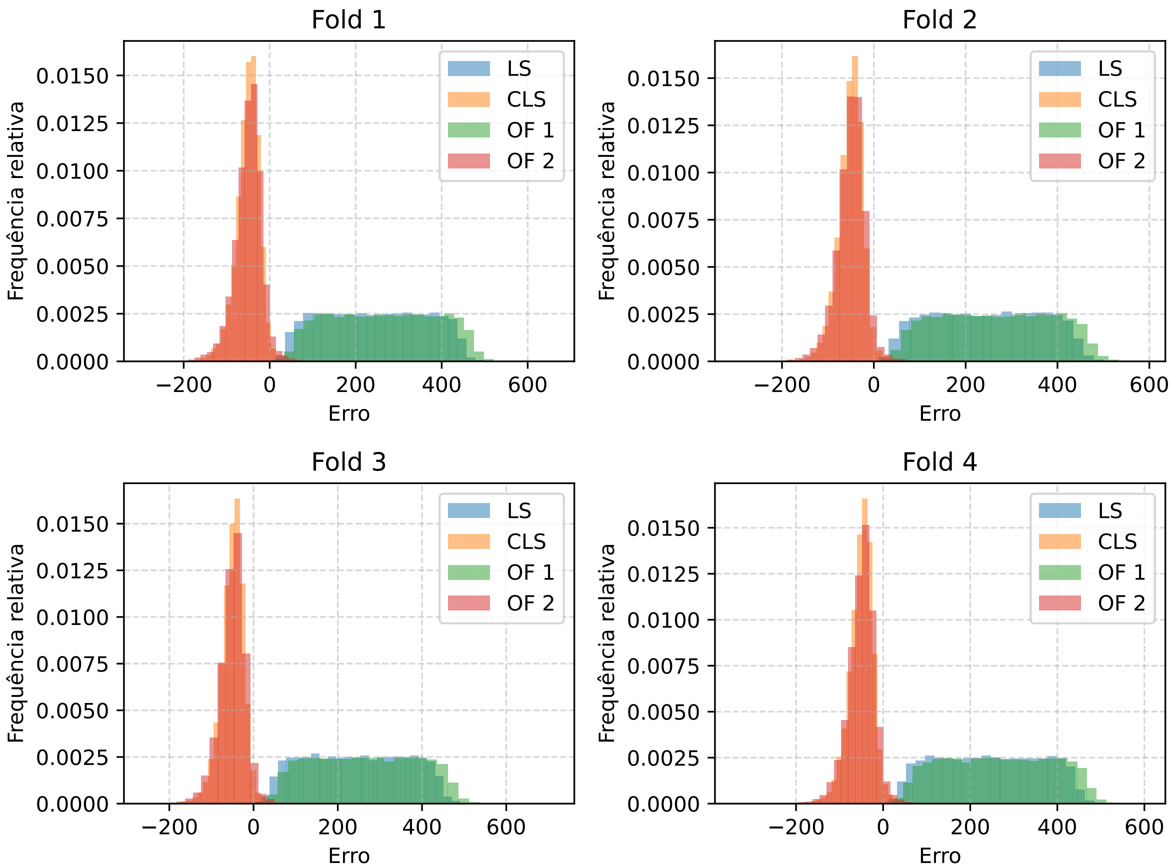


Fonte: O autor, 2025.

As figuras 31 e 32 mostram diferenças significativas na distribuição dos erros entre os métodos analisados nos primeiros 4 folds. Os métodos CLS e OF 2 apresentam histogramas com menor dispersão, estando quase sobrepostos, mas com uma leve inclinação em torno do valor zero. Por outro lado, os métodos LS e OF 1 têm histogramas mais dispersos, indicando uma maior variabilidade nos erros, o que sugere maior incerteza ou instabilidade nos resultados desses métodos.

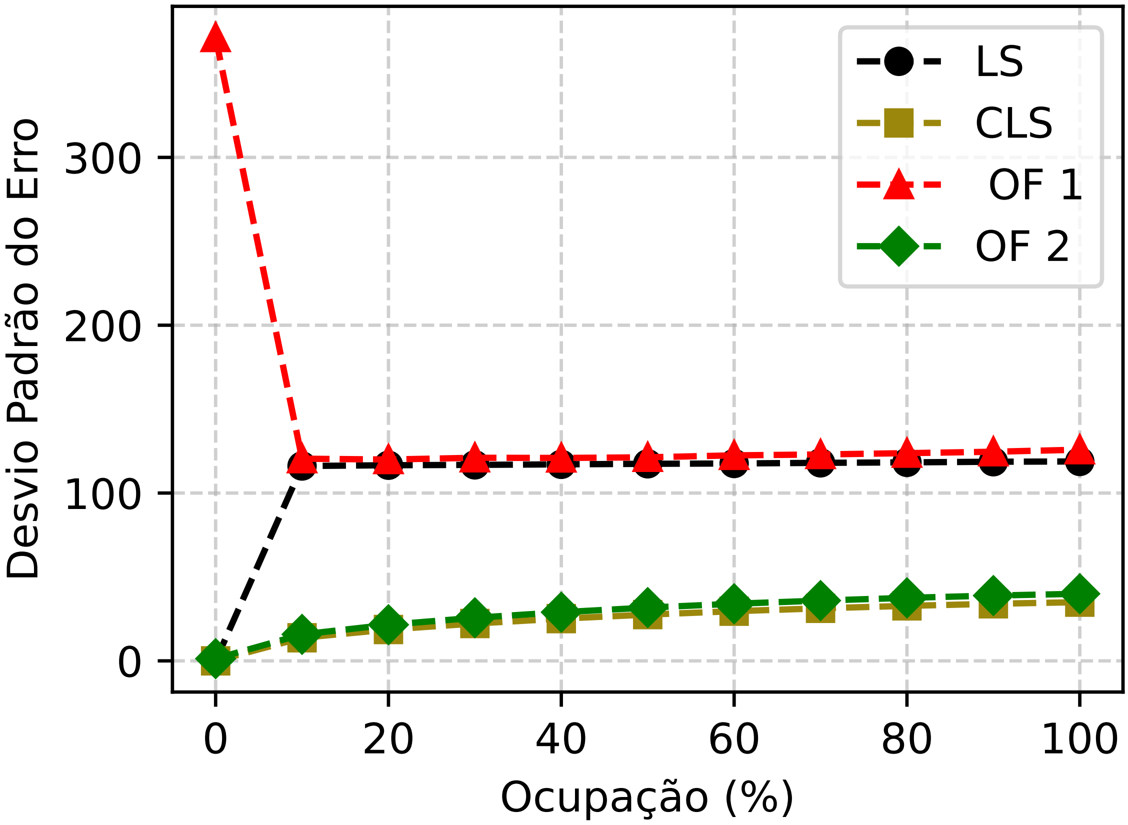
A figura 33, referente ao último *k*-fold, mostra que os métodos CLS e OF 2 apre-

Figura 32 – Erros de estimativa dos métodos LS, CLS, OF 1 e OF 2, representados pelos histogramas em azul, laranja, verde e vermelho, respectivamente, para 50% de ocupação.



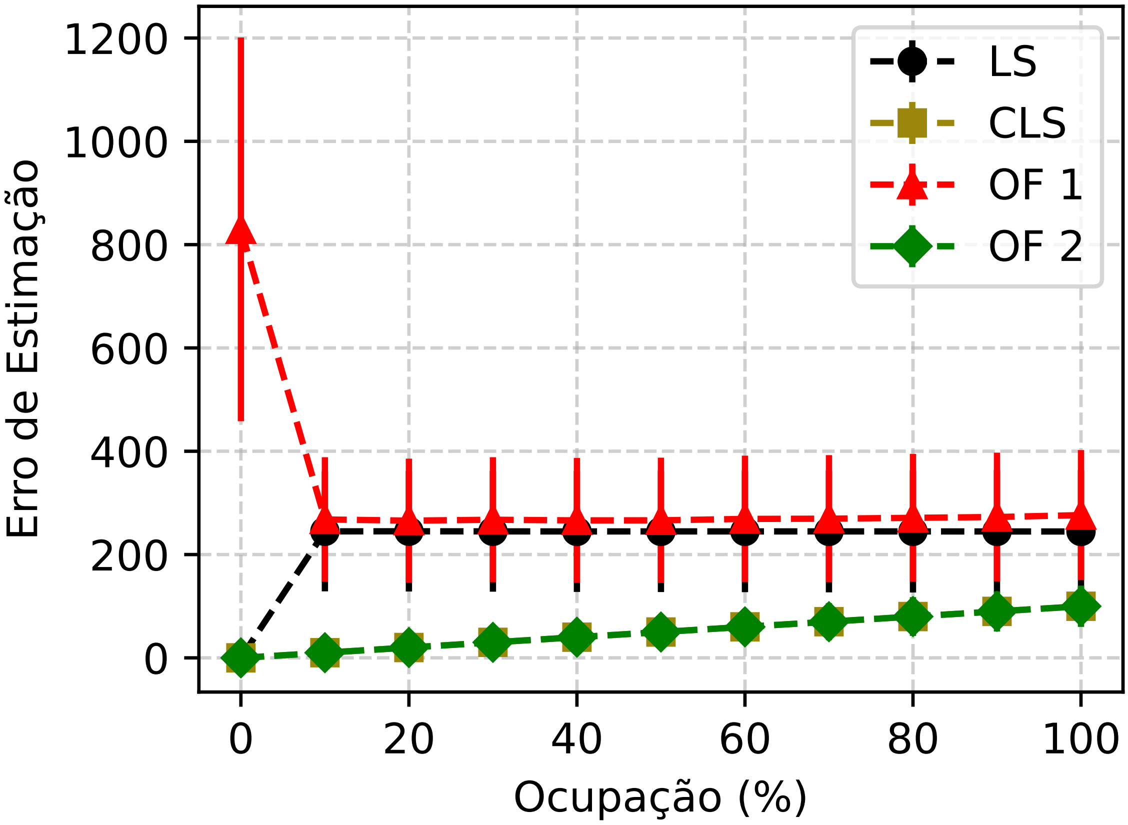
Fonte: O autor, 2025.

Figura 33 – Desvio padrão por ocupação usando os métodos LS, CLS, OF 1 e OF 2, representados pelas curvas em preto, amarelo, vermelho e verde, respectivamente.



Fonte: O autor, 2025.

Figura 34 – Erros por ocupação usando os métodos LS, CLS, OF 1 e OF 2, representados com linhas tracejadas em preto, amarelo, vermelho e verde, respectivamente.

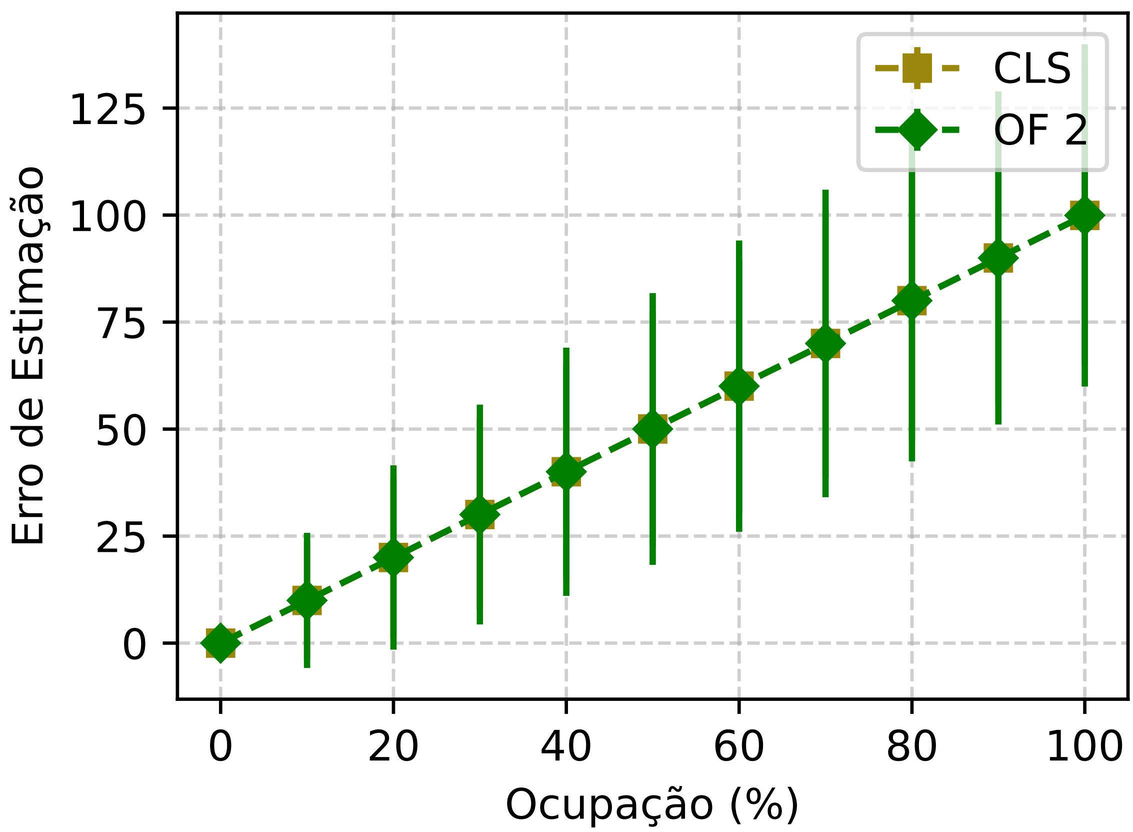


Fonte: O autor, 2025.

sentaram o melhor desempenho em termos de desvio padrão do erro para todos os níveis de ocupação analisados, com o método CLS apresentando um desempenho superior ao do OF 2. Isso pode ser visto de forma mais clara na figura 36. Por outro lado, os métodos LS e OF 1 não tiveram um bom desempenho, com resultados muito próximos entre si, exceto no caso da ocupação 0.

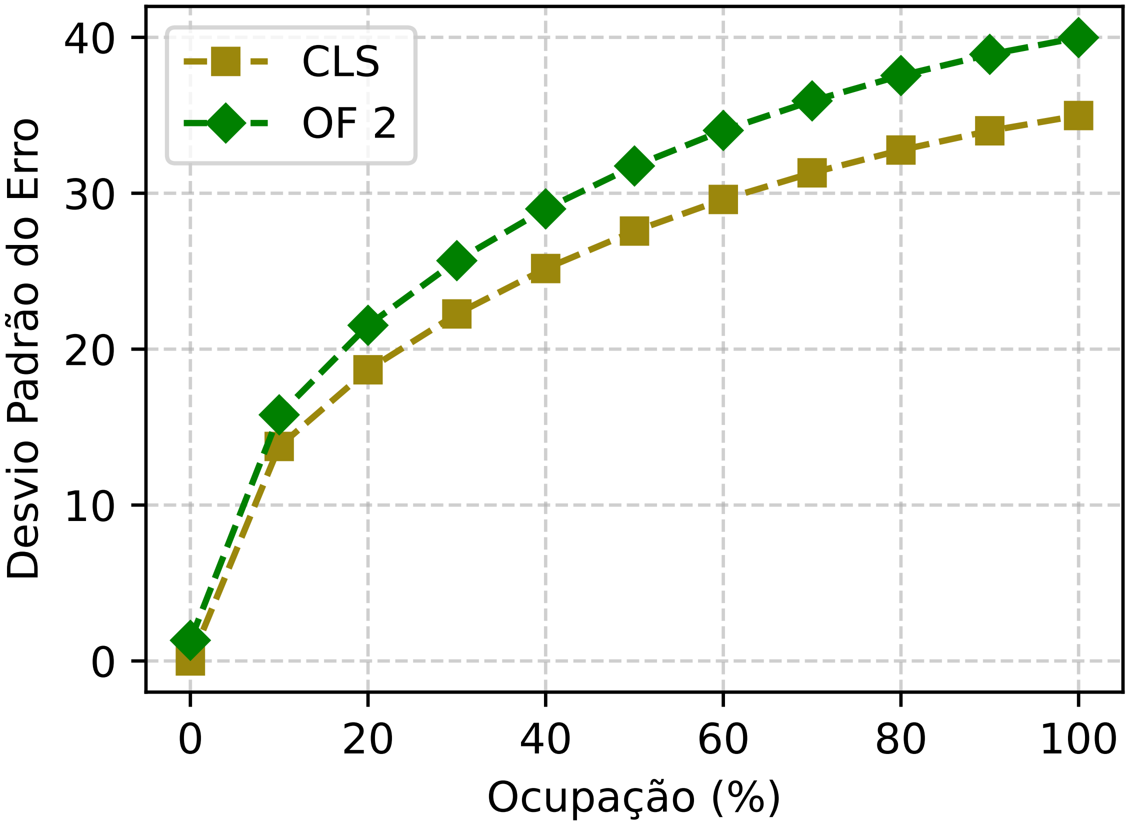
Já a figura 34 apresenta os erros de estimativa para cada método em função da ocupação. Nota-se que os métodos CLS e OF 2 exibem comportamentos quase idênticos (veja a figura 35), com erros mais próximos de zero. Já os métodos LS e OF 1 exibem erros mais distantes de zero ao longo de todas as ocupações.

Figura 35 – Erros por ocupação usando os métodos CLS e OF 2, representados com linhas tracejadas amarelo e verde, respectivamente.



Fonte: O autor, 2025.

Figura 36 – Desvio padrão por ocupação usando os métodos CLS e OF 2, representados pelas curvas em amarelo e verde, respectivamente.



Fonte: O autor, 2025.

CONCLUSÃO

No HL-LHC, espera-se um aumento no número de interações por colisão, o que causa sobreposição de sinais e dificulta a estimativa precisa da amplitude do sinal. Os métodos lineares buscam minimizar o efeito do empilhamento, melhorando a precisão das medições. Os sinais gerados pelas colisões de partículas podem ser contaminados por ruído. Os modelos lineares implementam técnicas para filtrar e minimizar o impacto desse ruído nas medições.

Os resultados apresentados indicam que, para a reconstrução da energia da amplitude do sinal proveniente das colisões de partículas no LHC, com pedestal fixo, o método mais eficiente é o LS, que apresentou os melhores resultados em comparação com CLS, OF 1 e OF 2. O segundo método mais indicado foi o OF 1, seguido pelo CLS, enquanto o OF 2 apresentou o pior desempenho nesse contexto.

No cenário de pedestal variável, que reflete uma condição mais realista, é importante destacar que o método atualmente utilizado no colorímetro de telhas do LHC é o OF 2. No entanto, os resultados obtidos mostram que o CLS apresentou o melhor desempenho, superando OF 2, LS e OF 1 na reconstrução da energia da amplitude do sinal em condições análogas às previstas para o HL-LHC. Dessa forma, para as condições do HL-LHC, o CLS é o método mais indicado, seguido por OF 1, LS e, por último, OF 2.

O estudo apresentado avaliou modelos lineares para a estimativa de energia no calorímetro de telhas TileCal do experimento ATLAS, focando no impacto do empilhamento de sinais e da variação do pedestal, e embora tenha apresentado resultados significativos, algumas limitações podem ser abordadas para aprimorar a análise e o entendimento dos fenômenos. O estudo utilizou dados simulados para as colisões de partículas, o que pode não refletir completamente as complexidades e nuances dos dados reais obtidos no experimento ATLAS. Embora a ferramenta de simulação utilizada permita configurar diversos parâmetros, a idealização dos dados pode não abranger todas as variáveis e efeitos presentes em um ambiente experimental real. A variação do pedestal foi introduzida através da adição de um vetor de dados aleatórios às amostras. Este método, apesar de simular variações no pedestal, pode ser simplificado em comparação com as flutuações complexas do pedestal observadas em experimentos reais. É necessário utilizar métodos de modelagem mais sofisticados para simular as variações do pedestal, como a incorporação de componentes de ruído mais realistas e o uso de dados do TileCal para modelar a flutuação do pedestal.

Para trabalhos futuros, também planeja-se realizar análises semelhantes utilizando dados gerados com o modelo Lorenzetti, explorando diferentes janelas de leitura resultantes das colisões.

REFERÊNCIAS

ABDALLAH, J. et al. Mechanical construction and installation of the ATLAS tile calorimeter. *Journal of Instrumentation*, v. 8, 11 2013. ISSN 17480221.

ANDERSON, K. Design of the front-end analog electronics for the ATLAS tile calorimeter. *Nuclear Instruments and Methods in Physics Research A*, v. 551, p. 469–476, 2005.

ANGUITA, D. et al. The 'K' in K-fold Cross Validation. In: *20th European Symposium on Artificial Neural Networks, ESANN 2012, Bruges, Belgium, April 25-27, 2012*. [S.l.: s.n.], 2012. Acesso em: 16 set. 2024.

ARNISON, G.; AL. et. Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at p=540 GeV. *Physics Letters B*, v. 122, n. 1, p. 103–116, 1983.

ATLAS Collaboration. *ATLAS Collaboration Technical Design Report for the ATLAS Tile Calorimeter*. Geneva, 2005. Disponível em: https://cds.cern.ch/record/1505342. Acesso em: 06 de set. 2024.

BRÜNING, O. et al. High-luminosity Large Hadron Collider (HL-LHC). *CERN Yellow Reports: Monographs*, 2020.

BUSSAB, W. d. O.; MORETTIN, P. A. Estatística básica. In: *Estatística básica*. São Paulo: Saraiva, 2010. p. xvi–540.

CARDINOT, D. A. et al. Comparação entre o Método do Filtro Ótimo com e sem Restrição para a Estimação da Amplitude de um Sinal no Contexto da Atualização do LHC. In: ENMC. *XXVII Encontro Nacional de Modelagem Computacional e XV Encontro de Ciência e Tecnologia dos Materiais*. Ilhéus, Bahia, 2024.

CERN. *A colaboração ATLAS*. Geneva, 2023. Disponível em: https://cds.cern.ch/record/2851122/files/A%20colabora%C3%A7%C3%A3o%20ATLAS.pdf. Acesso em: 3 set. 2024.

CERN. About CERN. *CERN Website*, 2024. Disponível em: https://home.cern/about. Acesso em: 25 ago. 2024.

CERN. *Accelerators*. 2024. Disponível em: https://home.cern/science/accelerators. Acesso em: 28 ago. 2024.

CERN. *ALICE Experiment*. 2024. Disponível em: https://home.cern/science/experiments/alice. Acesso em: 28 ago. 2024.

CERN. *CERN - European Organization for Nuclear Research*. 2024. Disponível em: https://home.cern/. Acesso em: 27 ago. 2024.

CERN. *CMS Experiment*. 2024. Disponível em: https://home.cern/science/experiments/cms. Acesso em: 29 ago. 2024.

CERN. *LHCb Experiment*. 2024. Disponível em: https://home.cern/science/experiments/lhcb. Acesso em: 29 ago. 2024.

CERN. Views of CERN - Image Gallery. *CERN Website*, 2024. Disponível em: https://home.cern/resources/image/cern/views-cern-images-gallery. Acesso em: 25 ago. 2024.

CHIEDDE, N. Machine learning for real-time processing of ATLAS liquid argon calorimeter signals with FPGAs. *Journal of Instrumentation*, v. 17, n. 04, p. C04010, 2022.

CLOSE, F. *Particle Physics: A Very Short Introduction*. Oxford: Oxford University Press, 2023.

COLLABORATION, A. Measurement of the $t\bar{t}$ production cross section in proton-proton collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector. *European Physical Journal C*, v. 84, n. 4, p. 223, 2024. Disponível em: https://cds.cern.ch/record/2654110/files/?ln=pt. Acesso em: 28 ago. 2024.

COLLABORATION, T. A. Readiness of the ATLAS Tile Calorimeter for LHC collisions. *European Physical Journal C*, Springer, v. 70, p. 1193–1236, 2010.

DAVIDEK, T. Performance and calibration of the ATLAS Tile Calorimeter. *Instruments*, MDPI, v. 6, n. 3, p. 25, 2022.

EVANS, L.; BRYANT, P. LHC machine. *Journal of Instrumentation*, IOP Publishing, v. 3, n. 08, p. S08001, 2008.

FABJAN, C. W. Calorimetry in high-energy physics. In: *Techniques and Concepts of High-Energy Physics III*. [S.l.]: Springer, 1985. p. 281–336.

FRANCAVILLA, P.; COLLABORATION, A. et al. The ATLAS tile hadronic calorimeter performance at the LHC. In: IOP PUBLISHING. *Journal of Physics: Conference Series*. [S.l.], 2012. v. 404, n. 1, p. 012007.

GONÇALVES, G. *Calorimetry Pulse Simulator*. 2022. Disponível em: https://github.com/ingoncalves/calorimetry-pulse-simulator. Acesso em: 28 agos. de 2024.

GONÇALVES, G. I. et al. Modelos Lineares Assistidos por Redes Neurais para Estimação Online da Energia do Calorímetro de Telhas do ATLAS. In: *XVI Brazilian Conference on Computational Intelligence (CBIC 2023)*. Salvador, Brazil: [s.n.], 2023.

HASERT, F. Search for Elastic Muon-neutrino Electron Scattering. *Physics Letters B*, v. 46, n. 1, p. 121–124, 1973.

INSALI, B. M. et al. Modelos Lineares para a Reconstrução de Sinal do Calorímetro de Telhas do ATLAS para o HL-LHC. In: SBMAC. *XLIII Congresso Nacional de Matemática Aplicada e Computacional*. Porto de Galinhas, PE, 2024.

KAY, S. M. *Fundamentals of Statistical Signal Processing: Estimation Theory*. USA: Prentice-Hall, 1993.

KLIMEK, P. ATLAS Tile Calorimeter Data Quality Assessment with Commissioning Data. In: *15th International Conference on Calorimetry in High Energy Physics*. [S.l.: s.n.], 2012. Disponível em: https://cdsweb.cern.ch/record/1473499. Acesso em: 12 dez. 2024.

KORDAS, K. The ATLAS data acquisition and Trigger: Concept, design and status. *Nuclear Physics B - Proceedings Supplements*, v. 172, p. 178–182, 2007.

MADISETTI, V. K.; WILLIAMS, D. B. (Ed.). *Digital Signal Processing Handbook*. Boca Raton: CRC Press LLC, 1999.

MOREIRA, M. A. O modelo padrão da física de partículas. *Revista Brasileira de Ensino de Física*, SciELO Brasil, v. 31, p. 1306–1, 2009.

NAKAHAMA, Y.; COLLABORATION on behalf of the A. The ATLAS Trigger System: Ready for Run-2. In: *Journal of Physics: Conference Series*. [S.l.]: IOP Publishing, 2015. v. 664, p. 082050.

NumPy Developers. *numpy.std*. [S.l.], 2024. Disponível em: https://numpy.org/doc/stable/reference/generated/numpy.std.html. Acesso em: 28 Oct. 2024.

ORFANIDIS, S. J. *Optimum Signal Processing: An Introduction*. 2nd. ed. New York: McGraw-Hill, 1988. 590 p. ISBN 0070477949.

PERRODO, P. The ATLAS liquid argon calorimetry system. In: *Proceedings of ICHEP*. Amsterdam: [s.n.], 2002. p. 909–912.

PIMENTA, J. J. M. et al. O bóson de Higgs. *Revista Brasileira de Ensino de Física*, SciELO Brasil, v. 35, p. 2306, 2013.

ResearchGate. *TileCal calibration system scheme*. Disponível em: https://www.researchgate.net/figure/TileCal-calibration-system-scheme_fig1_301658563. Acesso em: 10 set. 2024.

RIMES, S. d. M. et al. Estimação da amplitude de sinais em calorimetria de altas energias em condições de alta ocupação de eventos no detector. *XXXVIII Simpósio Brasileiro de Telecomunicações e Processamento de Sinais(SBrT). doi: https://doi.org/10.14209/SBRT*, 2020.

RIMES, S. d. M. et al. Filtragem inversa não-linear para estimação de sinais em calorímetros operando a alta taxa de eventos. Universidade do Estado do Rio de Janeiro, 2021.

SáNCHEZ, C. A. S. *Implementation of the ROD Crate DAQ Software for the ATLAS Tile Calorimeter and a Search for a MSSM Higgs Boson Decaying into Tau Pairs*. Tese (Doutorado) — Universitat de Valéncia - CSIC, Valência, Espanha, 2010. Disponível em: https://cds.cern.ch/record/1309926. Acesso em: 28 ago. 2024.

The ATLAS Collaboration. The ATLAS experiment at the CERN large hadron collider. IOP Publishing, 2008.

WIGMANS, R. *Calorimetry: Energy Measurement in Particle Physics*. 2nd. ed. Oxford: Clarendon Press, 2017. ISBN 9780198786351.

WONG, T. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recognit.*, v. 48, n. 9, p. 2839–2846, 2015. Disponível em: https://doi.org/10.1016/j.patcog.2015.03.009. Acesso em: 12 set. 2024.