



UNIVERSIDADE FEDERAL DE CAMPINA GRANDE
PROGRAMA DE PÓS GRADUAÇÃO EM FÍSICA
UNIDADE ACADÊMICA DE FÍSICA

Maria Gabriela Ferreira Siqueira Amaral Gomes

**Medida da Largura Invisível do Bóson Z
Utilizando os Dados Abertos do Experimento
CMS**

Campina Grande, Paraíba, Brasil

10 de junho de 2025

Maria Gabriela Ferreira Siqueira Amaral Gomes

Medida da Largura Invisível do Bóson Z Utilizando os Dados Abertos do Experimento CMS

Dissertação realizada sob orientação do Prof. Dr. Diego Alejandro Cogollo Aponte, apresentada à Unidade Acadêmica de Física em complementação aos requisitos para obtenção do título de Mestre em Física.

Orientador: Professor Dr. Diego Alejandro Cogollo Aponte

Coorientador: Professor Dr. Gilvan Augusto Alves

Campina Grande, Paraíba, Brasil

10 de junho de 2025

G633m Gomes, Maria Gabriela Ferreira Siqueira Amaral.
Medida da largura invisível do bóson Z utilizando os dados
abertos do experimento CMS / Maria Gabriela Ferreira Siqueira
Amaral Gomes. – Campina Grande, 2025.
79 f. : il. color.

Dissertação (Mestrado em Física) – Universidade Federal de
Campina Grande, Centro de Ciências e Tecnologia, 2025.
“Orientação: Prof. Dr. Diego Alejandro Cogollo Aponte, Prof.
Dr.
Gilvan Augusto Alves”.

Referências.

1. Física de Partículas. 2. Aprendizado de Máquina – Graph Neural
Networks (GNNs). 3. Bóson Z – Modelo Padrão – CMS Open
Data.
4. Improved Resistive Plate Chambers (iRPCs). I. Aponte,
Diego Alejandro Cogollo. II. Alves, Gilvan Augusto. III. Título.

UFCG/BC

CDU 539.12(043.3)



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE CAMPINA GRANDE
POS-GRADUACAO EM FISICA

Rua Aprigio Veloso, 882, - Bairro Universitario, Campina Grande/PB, CEP 58429-900

REGISTRO DE PRESENÇA E ASSINATURAS

REGISTRO DE PRESENÇA E ASSINATURAS

ATA DA DEFESA PARA CONCESSÃO DO GRAU DE MESTRE EM FÍSICA, REALIZADA
EM 06 DE JUNHO DE 2025

Aos seis dias do mês de junho do ano de dois mil e vinte e cinco, reuniram-se em caráter de solenidade pública, os membros da comissão designada para avaliar **Maria Gabriela Ferreira Siqueira Amaral Gomes** ao grau de Mestre em Física, área de concentração Física. Foram componentes da Banca Examinadora os especialistas: o professor **Diego Alejandro Cogollo Aponte** (Orientador) – Doutor em Física, **Gilvan Augusto Alves** (Coorientador) – Doutor em Física, o professor **Luiz Augusto Stuaní Pereira** – Doutor em Física, a professora **Patrícia Rebello Teles** – Doutor em Física, sendo o primeiro e o terceiro integrantes do corpo docente do PPG de Física da Universidade Federal de Campina Grande, e o segundo e a quarta, integrantes de pesquisadores do Centro Brasileiro de Pesquisas Físicas. HORA DE INÍCIO: **10h00min** – LOCAL: **Sala Virtual**. Dando início aos trabalhos, o Presidente da Banca, professor **Diego Alejandro Cogollo Aponte**, após declarar os objetivos da reunião, apresentou o(a) candidato(a) **Maria Gabriela Ferreira Siqueira Amaral Gomes**, a quem concedeu a palavra para que dissertasse oral e sucintamente sobre o tema apresentado, intitulado “**Medida da largura invisível do bóson Z utilizando dados abertos do CMS**”. Após discorrer o referido tema, o(a) candidato(a) foi arguido(a) pelos examinadores na forma regimental. Ato contínuo, passou a Comissão, em caráter secreto, a proceder a avaliação e julgamento do trabalho, concluindo por atribuir-lhe o conceito Aprovado. Face à aprovação, declarou o Presidente estar o(a) avaliado(a), legalmente habilitado(a) a receber o Grau de Mestre em Física, cabendo à Universidade Federal de Campina Grande, providências para a expedição do Diploma a que o(a) mesmo(a) faz jus. Nada mais havendo a tratar, foi lavrada a ata, que será submetida à aprovação da Comissão Examinadora. Campina Grande, 06 de junho de 2025.

Diego Alejandro Cogollo Aponte
Presidente da Comissão e Orientador

Gilvan Augusto Alves □
Coorientador

Luiz Augusto Stuaní Pereira □
Examinador Interno

Patrícia Rebello Teles
Examinadora Externa

Maria Gabriela Ferreira Siqueira Amaral Gomes
Candidato(a)

João Rafael Lucio dos Santos
Coordenador do Programa

2 - APROVAÇÃO

2.1. Segue a presente Ata de Defesa de Tese de Mestrado do(a) candidato(a) **Maria Gabriela Ferreira Siqueira Amaral Gomes**, assinada eletronicamente pela Comissão Examinadora acima identificada.

2.2. No caso de examinadores externos que não possuam credenciamento de usuário externo ativo no SEI, para igual assinatura eletrônica, os examinadores internos signatários certificam que os examinadores externos acima identificados participaram da defesa da tese e tomaram conhecimento do teor deste documento.



Documento assinado eletronicamente por **DIEGO ALEJANDRO COGOLLO APONTE, PROFESSOR(A) DO MAGISTERIO SUPERIOR**, em 06/06/2025, às 11:55, conforme horário oficial de Brasília, com fundamento no art. 8º, caput, da [Portaria SEI nº 002, de 25 de outubro de 2018](#).



Documento assinado eletronicamente por **JOAO RAFAEL LUCIO DOS SANTOS, COORDENADOR(A)**, em 06/06/2025, às 12:00, conforme horário oficial de Brasília, com fundamento no art. 8º, caput, da [Portaria SEI nº 002, de 25 de outubro de 2018](#).



Documento assinado eletronicamente por **Patricia Rebello Teles, Usuário Externo**, em 06/06/2025, às 12:03, conforme horário oficial de Brasília, com fundamento no art. 8º, caput, da [Portaria SEI nº 002, de 25 de outubro de 2018](#).



Documento assinado eletronicamente por **Maria Gabriela Ferreira Siqueira Amaral Gomes, Usuário Externo**, em 06/06/2025, às 12:16, conforme horário oficial de Brasília, com fundamento no art. 8º, caput, da [Portaria SEI nº 002, de 25 de outubro de 2018](#).



Documento assinado eletronicamente por **Gilvan Augusto Alves, Usuário Externo**, em 06/06/2025, às 12:19, conforme horário oficial de Brasília, com fundamento no art. 8º, caput, da [Portaria SEI nº 002, de 25 de outubro de 2018](#).



Documento assinado eletronicamente por **Luiz Augusto Stuaní Pereira, Usuário Externo**, em 06/06/2025, às 12:32, conforme horário oficial de Brasília, com fundamento no art. 8º, caput, da [Portaria SEI nº 002, de 25 de outubro de 2018](#).



A autenticidade deste documento pode ser conferida no site <https://sei.ufcg.edu.br/autenticidade>, informando o código verificador **5500154** e o código CRC **6B6A83FB**.

Referência: Processo nº 23096.032337/2025-26

SEI nº 5500154

Agradecimentos

Esta dissertação é fruto do trabalho que desenvolvi ao longo de dois anos. Nada disso teria sido possível sem o apoio, a orientação e o incentivo que recebi de muitas pessoas, antes e durante o mestrado, e por isso sou profundamente grata.

Me sinto privilegiada por ter contado com a orientação dos professores Diego Cogollo, Gilvan Alves, e Patricia Telles durante o mestrado. Sou sinceramente grata pelo acompanhamento, pelo apoio constante e pela confiança que depositaram em mim. As oportunidades que me proporcionaram foram determinantes para o desenvolvimento da minha trajetória acadêmica. São, para mim, grandes inspirações — como pessoas, professores e pesquisadores. Espero, sinceramente, um dia retribuir contribuindo para a formação de outros estudantes, assim como fizeram por mim.

Também agradeço ao Programa de Pós-Graduação em Física da Universidade Federal de Campina Grande (UFCG), aos excelentes professores e aos bons amigos que conquistei durante o mestrado. Agradeço especialmente ao professor Luiz Stuari por ter participado da minha banca e contribuído com sugestões valiosas. Sou igualmente grata aos professores da Universidade Estadual da Paraíba (UEPB), que foram muito importantes para mim durante a graduação e continuam sendo até hoje.

Agradeço à FAPESQ pela bolsa de mestrado e ao Programa Paraíba Sem Fronteiras, da FAPESQ em parceria com o SECTIES, pela bolsa de mestrado sanduíche.

Agradeço aos grupos Open Data CMS e RPC-CMS pela orientação, apoio e por tudo o que me ensinaram durante meu período no CERN.

Agradeço aos meus muitos e queridos amigos que, mesmo os que estão longe, estiveram sempre ao meu lado, fazendo companhia e trazendo alegria à minha rotina. Muitas vezes, foi por causa de vocês que consegui sorrir em meio à ansiedade.

À minha família — Maria, Vilanilza, Carlos, Rayanny, Jefferson, Elizabete, Cleyton, Marizabel, Zequinha e Zara — minhas pessoas preferidas no mundo, agradeço por todo o amor, incentivo e dedicação que sempre me deram. Nos momentos em que me senti incapaz, foi em vocês que pensei. Dedico este trabalho a cada um de vocês, com todo o meu carinho.

Por fim, agradeço a todas as pessoas que, de alguma forma, contribuíram para a realização desta etapa. Cada gesto, por mais simples que tenha sido, fez a diferença.

"É justo que muito custe o que muito vale"(Santa Teresa D'Ávila)

Resumo

Este trabalho apresenta a medida da largura invisível do bóson Z , um parâmetro fundamental do Modelo Padrão, utilizando dados do CMS Open Data provenientes de colisões próton-próton a $\sqrt{s} = 13$ TeV, com uma luminosidade integrada de $16,290 \text{ fb}^{-1}$. A largura invisível do bóson Z , associada ao decaimento $Z \rightarrow \nu\bar{\nu}$, é inferida a partir do momento transversal ausente em eventos $Z + \text{jatos}$. Esta abordagem permite obter informações sobre o acoplamento do bóson Z com neutrinos. Os resultados obtidos foram comparados com medidas previamente publicadas pela colaboração CMS. Adicionalmente, o trabalho inclui um estudo exploratório de seleção de eventos com aprendizado de máquina, por meio de *Graph Neural Networks* (GNNs), aplicadas como alternativa à abordagem tradicional baseada em cortes físicos. A pesquisa também contempla contribuições para a atualização do sistema de múons do CMS, com ênfase na montagem e nos testes dos Improved Resistive Plate Chambers (iRPCs). Esses detectores gasosos, projetados para operar com alta precisão temporal, serão instalados na região dos *endcaps*, ampliando a cobertura em pseudo-rapidez e aumentando significativamente a eficiência na detecção e no rastreamento de múons, especialmente em cenários de alta luminosidade previstos para o upgrade do LHC.

Palavras-chave: Bóson Z , Modelo Padrão, CMS Open Data, Largura Invisível, GNN, Aprendizado de Máquina, iRPCs, Física de Partículas.

Abstract

This work presents a measurement of the invisible width of the Z boson, a fundamental parameter of the Standard Model, using CMS Open Data from proton-proton collisions at a center-of-mass energy of $\sqrt{s} = 13$ TeV, corresponding to an integrated luminosity of $16,290 \text{ fb}^{-1}$. The invisible width of the Z boson, associated with the decay $Z \rightarrow \nu\bar{\nu}$, is inferred from the missing transverse momentum observed in $Z + \text{jets}$ events. This approach provides insight into the coupling of the Z boson to neutrinos. The result obtained is compared with previous measurements published by the CMS Collaboration. In addition, the study includes an exploratory investigation of event selection using machine learning techniques, employing Graph Neural Networks (GNNs) as an alternative to traditional cut-based selection methods. The research also involves technical contributions to the upgrade of the CMS muon system, focusing on the assembly and testing of the Improved Resistive Plate Chambers (iRPCs). These gaseous detectors, designed for high-precision timing, will be installed in the endcap regions of the detector to extend pseudorapidity coverage and significantly enhance muon detection and tracking efficiency, particularly in the high-luminosity environment expected in future LHC upgrades.

Keywords: Z Boson, Standard Model, CMS Open Data, Invisible Width, GNN, Machine Learning, Deep Learning, iRPCs, Particle Physics.

Lista de ilustrações

Figura 1	–	Esquema do complexo de aceleradores do CERN. Fonte: [1].	29
Figura 2	–	Visão esquemática do detector CMS. Fonte: [2].	31
Figura 3	–	Localização dos detectores ECAL e HCAL em uma seção transversal longitudinal de um quarto do detector. Fonte:[3]	33
Figura 4	–	Esquema dos subdetectores do CMS com destaque para os detectores de múons (DTs, CSCs, RPCs e GEMs) em uma visão longitudinal. Fonte: [4].	34
Figura 5	–	Esquema transversal do detector CMS ilustrando a reconstrução de eventos. Fonte: [5]	35
Figura 6	–	Diagrama de Feynman ilustrando um processo típico em colisões próton-próton (pp), no qual um quark d e um antiquark $\bar{d}$, originados dos prótons incidentes, aniquilam-se gerando um bóson Z . O bóson Z então decai de forma invisível em um par neutrino–antineutrino ($\nu\bar{\nu}$). Embora não representados neste diagrama, jatos hadrônicos ($Z + \text{jets}$) geralmente acompanham o processo, sendo produzidos por radiação de gluons (QCD) emitidos pelos quarks iniciais. A presença de jatos, combinada com o momento transversal ausente (p_T^{miss}) resultante dos neutrinos, caracteriza a assinatura experimental desse processo. <i>Fonte: o autor.</i>	36
Figura 7	–	Evolução da acurácia e do $F1\text{-score}$ ao longo do treinamento.	53
Figura 8	–	Evolução da função de perda (CrossEntropyLoss) durante as épocas de treinamento.	53
Figura 9	–	Matriz de confusão final após o treinamento do modelo GNN.	54
Figura 10	–	Distribuição da variável u (magnitude do vetor de recuo) para eventos selecionados por aprendizado de máquina (ML) e por cortes tradicionais.	55
Figura 11	–	Distribuição do momento transversal do jato líder (p_T) para eventos selecionados por aprendizado de máquina (ML) e por cortes tradicionais.	55
Figura 12	–	Distribuição da pseudorapidez (η) dos fótons nos eventos selecionados por ML e por cortes.	56
Figura 13	–	Distribuição do momento transversal (p_T) dos fótons nos eventos selecionados por ML e por cortes.	56
Figura 14	–	Distribuições normalizadas de <i>pileup</i> nos dados reais, na simulação antes da correção de <i>pileup</i> e na simulação corrigida por <i>reweighting</i>	59
Figura 15	–	Distribuições do vetor de recuo U	60

Figura 16 – Distribuições marginais e correlações entre os parâmetros ajustados: θ (sistemático global), r (fator de escala da largura invisível do Z), r_Z (fator de transferência para $Z + \text{jets}$) e Γ_{inv} (largura invisível).	64
Figura 17 – Comparação entre diferentes medições experimentais da largura invisível do bóson Z. A linha tracejada representa a predição do Modelo Padrão. O ponto inferior (em verde) corresponde ao resultado obtido nesta análise com dados preliminares do CMS Open Data.	66
Figura 18 – Um quadrante do experimento CMS. Em roxo, as localizações das duas estações de iRPC nos endcaps - RE3/1 e RE4/1. Fonte:[6].	67
Figura 19 – Esquema detalhado de um RPC, destacando os principais componentes: placas resistivas de HPL, tiras de leitura, gás, grafite, isolante e espaçadores. Fonte: [7].	68
Figura 20 – O gráfico mostra uma reconstrução 3D das posições de impacto de múons em uma câmara iRPC em um ponto de operação com um limiar de carga de 40 fC. Os dados foram coletados usando um <i>trigger</i> de coincidência tripla de cintiladores com área de 30x40 cm ²	70
Figura 21 – O gráfico apresenta a eficiência de detecção de múons cósmicos em função da alta tensão efetiva (HV_{eff}). A eficiência atinge um pico de 99%, com o ponto de operação definido como a tensão em que a eficiência é 95% (HV_{knee}) mais 150V.	71
Figura 22 – Eficiência para a mistura de gases padrão com fontes: OFF (verde), 10 (rosa), 3.3 (vermelho), 2.2 (preto) e 1 (azul), durante o teste de feixe de 2024.	72
Figura 23 – Eficiência para a mistura de gases 30% CO ₂ + 1.0% SF ₆ com fontes: OFF (verde), 10 (rosa), 3.3 (vermelho), 2.2 (preto) e 1 (azul), durante o teste de feixe de 2024.	73
Figura 24 – Eficiência para a mistura de gases 30% CO ₂ + 0.5% SF ₆ com fontes: OFF (verde), 10 (rosa), 3.3 (vermelho), 2.2 (preto) e 1 (azul), durante o teste de feixe de 2024.	73
Figura 25 – Eficiência para a mistura de gases 40% CO ₂ + 1.0% SF ₆ com fontes: OFF (verde), 10 (rosa), 3.3 (vermelho), 2.2 (preto) e 1 (azul), durante o teste de feixe de 2024.	74

Lista de tabelas

Tabela 1 – Conjuntos de dados primários usados nas regiões de análise	38
Tabela 2 – Lista dos conjuntos de dados primários utilizados nesta análise.	40
Tabela 3 – Amostras de eventos simulados para os processos de sinal	40

Tabela 4 – Amostras de eventos simulados para os processos de <i>background</i> dominante	41
Tabela 5 – Amostras de eventos simulados para os processos de <i>background</i> secundários	42
Tabela 6 – Conjuntos de dados primários e seus correspondentes <i>triggers</i>	43
Tabela 7 – Fluxo de cortes da região p_T^{miss} - Parte 1	45
Tabela 8 – Fluxo de cortes da região p_T^{miss} - Parte 2	46
Tabela 9 – Fluxo de cortes da região single muon - Parte 1	46
Tabela 10 – Fluxo de cortes da região single muon - Parte 2	46
Tabela 11 – Fluxo de cortes da região single electron - Parte 1	47
Tabela 12 – Fluxo de cortes da região single electron - Parte 2	47
Tabela 13 – Fluxo de cortes da região single tau - Parte 1	47
Tabela 14 – Fluxo de cortes da região single tau - Parte 2	48
Tabela 15 – Fluxo de cortes da região QCD - Parte 1	48
Tabela 16 – Fluxo de cortes da região QCD - Parte 2	48
Tabela 17 – Fluxo de cortes da região double muon	49
Tabela 18 – Fluxo de cortes da região double electron	49
Tabela 19 – Seleção "baseline" e critérios para cada região.	50
Tabela 20 – Seção de choque e razões de vários processos em relação ao DY.	58
Tabela 21 – Fontes de incerteza sistemática consideradas na análise. Fonte: [8]	63
Tabela 22 – Comparação de misturas de gases utilizadas no sistema de iRPCs do CMS e seus Potenciais de Aquecimento Global (GWP).	72

Sumário

1	INTRODUÇÃO	15
2	O MODELO PADRÃO DA FÍSICA DE PARTÍCULAS	18
2.1	Eletrodinâmica Quântica (QED) e o Princípio de Gauge	18
2.2	A Teoria Eletrofraca	19
2.3	A Teoria da Interação Forte: Cromodinâmica Quântica (QCD)	20
2.4	A Lagrangeana do Modelo Padrão	21
2.5	O Setor de Férmions	22
2.6	O Setor de Gauge	23
2.7	O Setor de Yukawa	24
2.8	A Quebra Espontânea de Simetria	25
2.9	Os Bósons de Gauge Físicos	26
2.10	O Tempo de Vida e Largura de Decaimento das Partículas	27
2.11	O Número de Neutrinos	28
3	O GRANDE COLISOR DE HÁDRONS	29
3.1	Introdução	29
3.2	O Experimento CMS	30
3.2.1	Solenóide Supercondutor	31
3.2.2	Rastreador de Silício (Silicon Tracker)	32
3.2.3	Calorímetro Eletromagnético (ECAL)	32
3.2.4	Calorímetro Hadrônico (HCAL)	33
3.2.5	Sistema de Múons	33
3.2.6	Sistema de <i>Trigger</i> e Aquisição de Dados	34
3.2.7	Sistema de Aquisição de Dados (DAQ)	34
3.2.8	Reconstrução de Eventos	35
4	RESULTADOS E DISCUSSÕES	36
4.1	Datasets	37
4.2	<i>Trigger</i> e Luminosidade	42
4.3	Seleção de Eventos	43
4.3.1	Fluxo de Cortes	45
4.3.2	Estudo de Seleção de Eventos com Machine Learning	51
4.4	Aplicação de Pesos para Normalização de Eventos	57
4.5	Correção da Contribuição de γ^* na Amostra DYJetsToLL	57
4.6	Correção de Pileup	58

4.7	Ajuste Simultâneo e Extração da Largura Invisível do Bóson Z . . .	60
4.8	Contribuição à Atualização do Sistema de Múons do CMS	67
4.8.1	Estudo de Gases para os iRPCs	71
5	CONSIDERAÇÕES FINAIS	75
	REFERÊNCIAS	77

1 Introdução

A preservação de dados é uma das prioridades do CERN (Centro Europeu de Pesquisa Nuclear). Nesse contexto, o CMS Open Data, lançado em 2014, destaca-se como uma iniciativa significativa, disponibilizando dados de colisões próton-próton e colisões de íons pesados coletados pelo experimento CMS (Compact Muon Solenoid). Esses dados têm diversas finalidades, como a reinterpretação e reanálise de resultados, a promoção da educação e da divulgação científica, além do desenvolvimento técnico e algorítmico [9].

O presente trabalho tem como objetivo realizar uma medida direta da largura invisível do bóson Z , uma importante observável prevista pelo Modelo Padrão, e comparar o resultado obtido com a medição previamente publicada pela colaboração CMS [8]. Segundo o Modelo Padrão, aproximadamente 20% dos decaimentos do bóson Z resultam na produção de um par neutrino-antineutrino [10]. Esses decaimentos, denominados invisíveis, não são diretamente observáveis no detector CMS, mas podem ser inferidos a partir do momento transversal ausente (p_T^{miss}) gerado pela presença dos neutrinos.

A medição da largura invisível do Z é um teste essencial do Modelo Padrão e fornece informações sobre o número de espécies de neutrinos que se acoplam ao bóson Z . Desvios em relação à expectativa do Modelo Padrão podem indicar a existência de novas partículas além do Modelo Padrão que contribuam para a largura invisível do Z .

A largura invisível do bóson Z é uma medição amplamente consolidada na comunidade científica, com resultados obtidos por experimentos de grande relevância, como o LEP, o CMS e o ATLAS. Apesar de ser uma observável bem conhecida, a realização desta análise no contexto do *CMS Open Data* possui grande importância. Ela será apresentada como um exemplo didático e completo de análise, abrangendo todas as etapas essenciais do fluxo de trabalho em física de partículas: desde a seleção de eventos e correções de *pileup*, passando pela visualização e exploração dos dados, até o ajuste simultâneo necessário para uma estimativa precisa da largura invisível do bóson Z .

Na estratégia de análise, são considerados eventos de sinal aqueles em que o bóson Z é produzido em associação com jatos ($Z + \text{jets}$) e decai em neutrinos, identificados pelo detector como momento transversal ausente (p_T^{miss}). Nesses casos, o p_T^{miss} equilibra o momento transversal dos jatos, evidenciando a conservação do momento na interação. A largura invisível do bóson Z é inferida por meio de um canal de referência: o decaimento do Z em léptons carregados, que compartilham características cinemáticas semelhantes com o decaimento invisível. Essa semelhança permite extrapolar as propriedades do decaimento leptônico para estimar com precisão a contribuição dos neutrinos.

$$\Gamma(Z \rightarrow \text{inv}) = \frac{\sigma(Z + \text{jets})\mathcal{B}(Z \rightarrow \text{inv})}{\sigma(Z + \text{jets})\mathcal{B}(Z \rightarrow \ell\bar{\ell})}\Gamma(Z \rightarrow \ell\bar{\ell}) \quad (1.1)$$

Nesse contexto, $\sigma(Z + \text{jets})$ representa a seção de choque do processo $Z + \text{jets}$, modificada pela razão de ramificação invisível¹ $\mathcal{B}(Z \rightarrow \text{inv})$ ou pela razão de ramificação para díleptons $\mathcal{B}(Z \rightarrow \ell\bar{\ell})$.

Para este estudo, foram utilizados os dados disponíveis no portal CMS Open Data, referentes a colisões próton-próton a uma energia de centro de massa de $\sqrt{s} = 13 \text{ TeV}$, registrados em 2016, com uma luminosidade integrada correspondente a $16,290 \text{ fb}^{-1}$. A análise foi conduzida utilizando a linguagem Python, e o código está disponível publicamente em um repositório no GitHub, acessível em: <https://github.com/Mariaggomes>.

Este trabalho também inclui um estudo complementar de seleção de eventos utilizando técnicas de aprendizado de máquina. Especificamente, foi investigado o uso de *Graph Neural Networks* (GNNs) para identificar eventos compatíveis com a região de sinal $Z \rightarrow \nu\bar{\nu}$. Esse tipo de rede, que opera diretamente sobre representações em forma de grafos, tem se mostrado promissor em tarefas de classificação em física de partículas, especialmente quando há múltiplos objetos reconstruídos em cada evento.

No contexto desta pesquisa, também foram desenvolvidas atividades técnicas durante o período de mestrado sanduíche no CERN, no âmbito do programa *Paraíba Sem Fronteiras*, relacionadas à montagem e ao teste dos *Improved Resistive Plate Chambers* (iRPCs). Os RPCs são detectores gasosos com excelente resolução temporal, posicionados tanto na região do barril quanto nos *endcaps* do CMS². Esses detectores desempenham um papel essencial no sistema de disparo (*trigger*) para detecção de múons, fornecendo informações temporais precisas que são cruciais para distinguir eventos de colisão muito próximos.

Como parte do processo de atualização do sistema de múons do CMS, em resposta ao aumento esperado da luminosidade nos próximos anos, novas câmaras RPC, denominadas iRPCs, serão instaladas na região dos *endcaps*, onde a cobertura atual do detector é limitada para a detecção de múons. Essas iRPCs expandirão a cobertura do sistema de múons, estendendo o alcance de pseudo-rapidez de 1,9 para 2,4 [11]. Além disso, os iRPCs apresentam melhorias significativas em relação aos RPCs originais, como eletrodos mais finos e menor espaçamento nas lacunas, o que resulta em um rastreamento mais preciso de múons.

Nos capítulos seguintes, serão abordados o Modelo Padrão da Física de Partículas,

¹ A razão de ramificação invisível de um processo representa a fração do número total de decaimentos de uma partícula que ocorre em um canal específico.

² A região do *barril* corresponde à parte central e cilíndrica do detector, ao redor da linha do feixe de prótons. Enquanto os *endcaps* são as extremidades do detector, dispostas perpendicularmente ao feixe, cobrindo regiões de maior pseudorapidez (η).

o Grande Colisor de Hádrons (LHC) e o experimento CMS, bem como os resultados obtidos, incluindo o estudo com machine learning, a extração da largura invisível do bóson Z e a contribuição para a atualização do sistema de múons do CMS.

2 O Modelo Padrão da Física de Partículas

O Modelo Padrão (MP) é a teoria que descreve as partículas elementares e as interações fundamentais, com exceção da gravitação. Ele abrange três das quatro forças fundamentais: forte, fraca e eletromagnética, sendo fundamentado no princípio de gauge e no grupo de simetria [12]:

$$SU(3)_C \times SU(2)_L \times U(1)_Y. \quad (2.1)$$

As interações fundamentais no MP são descritas por meio de simetrias de gauge, que determinam a forma das interações entre os campos. Para entender a estrutura do MP, comecemos pela Eletrodinâmica Quântica (QED) como exemplo básico do princípio de gauge.

2.1 Eletrodinâmica Quântica (QED) e o Princípio de Gauge

Na QED, a simetria de gauge está associada ao grupo $U(1)$. A lagrangiana inicial para um férmion livre de massa m é:

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi, \quad (2.2)$$

onde ψ é o campo do férmion, γ^μ são as matrizes de Dirac, e ∂_μ é a derivada ordinária [13]. Essa lagrangiana é invariante sob transformações globais de fase:

$$\psi \rightarrow e^{i\alpha}\psi, \quad (2.3)$$

com α sendo uma constante.

Quando a transformação de fase passa a depender das coordenadas espaço-temporais ($\alpha = \alpha(x)$), temos uma transformação local:

$$\psi(x) \rightarrow e^{i\alpha(x)}\psi(x). \quad (2.4)$$

Nesse caso, a derivada ordinária $\partial_\mu\psi$ deixa de ser invariante:

$$\partial_\mu\psi \rightarrow e^{i\alpha(x)}(\partial_\mu\psi + i(\partial_\mu\alpha)\psi). \quad (2.5)$$

Para restaurar a invariância, introduzimos o campo de gauge A_μ , que transforma como:

$$A_\mu \rightarrow A_\mu + \partial_\mu\alpha. \quad (2.6)$$

A derivada ordinária é substituída pela derivada covariante:

$$D_\mu = \partial_\mu + ieA_\mu, \quad (2.7)$$

onde e é a carga do férmion. A lagrangiana invariante sob transformações locais é:

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \quad (2.8)$$

com o tensor de força:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \quad (2.9)$$

Essa Lagrangiana descreve a interação entre férmions carregados (como elétrons) e o fóton (A_μ), que é o bóson de gauge da QED. Portanto, a QED é uma teoria de gauge baseada no grupo $U(1)$. O primeiro termo representa a dinâmica do férmion e sua interação com o campo eletromagnético por meio da derivada covariante D_μ , enquanto o segundo termo corresponde à energia do campo eletromagnético, expressa pelo tensor de campo $F_{\mu\nu}$.

2.2 A Teoria Eletrofraca

A teoria eletrofraca unifica as interações eletromagnética e fraca e é descrita pelo grupo $SU(2)_L \times U(1)_Y$. As partículas fundamentais estão organizadas em dubletos e singletos, refletindo suas propriedades de simetria. Assim, os férmions de mão esquerda pertencem à representação fundamental, também chamada de dubleto, do grupo $SU(2)_L$ e os férmions de mão direita estão na representação de singlete [13].

$$f_L^a = \begin{pmatrix} \nu^a \\ \ell^a \end{pmatrix}_L, \quad \ell_R^a, \quad (2.10)$$

$$\begin{aligned} \nu^1 &= \nu_e, & \nu^2 &= \nu_\mu, & \nu^3 &= \nu_\tau, \\ \ell^1 &= e^-, & \ell^2 &= \mu^-, & \ell^3 &= \tau^-. \end{aligned}$$

Enquanto,

$$Q_L^a = \begin{pmatrix} u^a \\ d^a \end{pmatrix}_L, \quad u_R^a, \quad d_R^a, \quad (2.11)$$

$$\begin{aligned} u^1 &= u, & u^2 &= c, & u^3 &= t, \\ d^1 &= d, & d^2 &= s, & d^3 &= b. \end{aligned}$$

A derivada covariante para os dubletos (Q_L, L_L) é:

$$D_\mu Q_L = \partial_\mu Q_L + igW_\mu^a T^a Q_L + ig' \frac{Y}{2} B_\mu Q_L, \quad (2.12)$$

e a derivada covariante para os singletos (u_R, d_R, e_R) é:

$$D_\mu \psi_R = \partial_\mu \psi_R + ig' \frac{Y}{2} B_\mu \psi_R. \quad (2.13)$$

Nessas expressões, Q_L e L_L representam os dubletos de quarks e léptons, que transformam sob o grupo de simetria $SU(2)_L$. Os campos W_μ^a (com $a = 1, 2, 3$) e B_μ são os bósons de gauge associados aos grupos $SU(2)_L$ e $U(1)_Y$, respectivamente. Os T^a são os geradores de $SU(2)_L$, dados por $T^a = \tau^a/2$, onde τ^a são as matrizes de Pauli, e Y denota a hipercarga fraca do campo.

2.3 A Teoria da Interação Forte: Cromodinâmica Quântica (QCD)

No contexto do Modelo Padrão, o grupo $SU(2)_L \times U(1)_Y$ descreve a unificação das interações eletromagnética e fraca na teoria eletrofraca, enquanto o grupo $SU(3)_C$ está associado à interação forte [13]. A Cromodinâmica Quântica (QCD) é o componente do Modelo Padrão que trata da interação forte e descreve como os quarks e glúons interagem por meio de cargas de cor. A Lagrangeana da QCD pode ser expressa como:

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4}G_{\mu\nu}^a G^{a\mu\nu} + \sum_q \bar{\psi}_q (i\gamma^\mu D_\mu - m_q) \psi_q, \quad (2.14)$$

$$G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a + g_s f^{abc} G_\mu^b G_\nu^c \quad (2.15)$$

Sendo $G_{\mu\nu}^a$ o tensor de campo de força, G_μ^a os campos de glúons, g_s a constante de acoplamento forte, e f^{abc} os símbolos de estrutura do grupo $SU(3)_C$.

- ψ_q representa os campos de quarks de diferentes sabores.
- $D_\mu = \partial_\mu - ig_s T^a G_\mu^a$ é a derivada covariante, com T^a sendo os geradores do grupo $SU(3)_C$.
- m_q são as massas dos quarks.

Um fenômeno característico da QCD é o confinamento de cor: quando duas partículas carregadas de cor são separadas, a energia do campo entre elas aumenta até que se torne energeticamente mais favorável criar um novo par quark-antiquark, em vez de continuar a separação. Como consequência, os quarks não podem ser observados isoladamente na natureza; eles estão sempre confinados em combinações neutras de cor chamadas de hádrons [12]. Esses hádrons podem ser classificados em:

- Mésons: combinações de um quark e um antiquark ($q\bar{q}$).
- Bárions: combinações de três quarks (qqq).

Por exemplo, o próton é um bárion composto por três quarks (uud).

Outro fenômeno relevante da QCD é a liberdade assintótica. Em altas energias (ou curtas distâncias), a interação entre quarks se torna mais fraca, e os quarks podem ser tratados como partículas quase livres. Isso é descrito pela evolução da constante de acoplamento forte α_s , que decresce com o aumento da escala de energia Q^2 . A dependência de α_s é dada por:

$$\alpha_s(Q^2) = \frac{1}{\beta_0 \ln(Q^2/\Lambda_{\text{QCD}}^2)}, \quad (2.16)$$

onde:

- β_0 é o coeficiente da função beta da QCD, dependente do número de quarks ativos.
- Λ_{QCD} é a escala de energia característica da QCD.

2.4 A Lagrangeana do Modelo Padrão

Dessa forma, o Modelo Padrão é capaz de descrever a dinâmica das partículas e suas interações, incluindo a teoria eletrofraca e a teoria das interações fortes. Esse arcabouço teórico oferece uma descrição abrangente das interações fundamentais das partículas elementares, organizadas sob as simetrias de gauge $SU(3)_C \times SU(2)_L \times U(1)_Y$.

Os termos que compõem a dinâmica do Modelo Padrão são os seguintes:

- $\mathcal{L}_{\text{Léptons}}$: descreve a dinâmica dos léptons e suas interações.
- $\mathcal{L}_{\text{quarks}}$: descreve a dinâmica dos quarks e suas interações.
- $\mathcal{L}_{\text{scalar}}$: inclui a contribuição do campo escalar de Higgs, responsável pela quebra espontânea de simetria eletrofraca.
- $\mathcal{L}_{\text{Yukawa}}$: define os acoplamentos de Yukawa, que são responsáveis pelas massas das partículas fundamentais.
- $\mathcal{L}_{\text{gauge}}$: trata da dinâmica dos campos de gauge, que mediam as interações fundamentais.

Esses componentes se combinam para formar a Lagrangeana completa do Modelo Padrão, expressa como:

$$\mathcal{L}_{\text{MP}} = \mathcal{L}_{\text{leptons}} + \mathcal{L}_{\text{quarks}} + \mathcal{L}_{\text{scalar}} + \mathcal{L}_{\text{Yukawa}} + \mathcal{L}_{\text{gauge}}. \quad (2.17)$$

2.5 O Setor de Férmions

Os dois primeiros termos da Lagrangeana do Modelo Padrão representam a parte fermiônica, descrevendo as interações dos férmions com os bósons de gauge físicos, tanto neutros quanto carregados. A Lagrangeana fermiônica é expressa como:

$$\mathcal{L}_{\text{férmions}} = \sum_{a=1}^3 \left[\bar{\psi}_L^a i\gamma^\mu D_\mu^L \psi_L^a + \bar{\psi}_R^a i\gamma^\mu D_\mu^R \psi_R^a \right]. \quad (2.18)$$

Nesta expressão:

- Quando $\psi_L^a = f_L^a$ e $\psi_R^a = \ell_R^a$, estamos nos referindo à Lagrangeana leptônica.
- Quando $\psi_L^a = Q_L^a$ e $\psi_R^a = u_R^a, d_R^a$, estamos tratando da Lagrangeana de quarks.

Os campos e as derivadas covariantes transformam sob o grupo $SU(2)_L \times U(1)_Y$ da seguinte maneira:

$$\psi_L(x) \rightarrow \psi'_L(x) = e^{it^a \alpha^a(x) + i\frac{Y}{2}\beta(x)} \psi_L(x) \quad (2.19)$$

$$\bar{\psi}_L(x) \rightarrow \bar{\psi}'_L(x) = \bar{\psi}_L(x) e^{-it^a \alpha^a(x) - i\frac{Y}{2}\beta(x)} \quad (2.20)$$

$$\psi_R(x) \rightarrow \psi'_R(x) = e^{i\frac{Y}{2}\beta(x)} \psi_R(x) \quad (2.21)$$

$$\bar{\psi}_R(x) \rightarrow \bar{\psi}'_R(x) = \bar{\psi}_R(x) e^{-i\frac{Y}{2}\beta(x)} \quad (2.22)$$

As derivadas covariantes transformam conforme:

$$D_\mu^L \psi_L \rightarrow D_\mu^{L'} \psi'_L = e^{it^a \alpha^a(x) + i\frac{Y}{2}\beta(x)} D_\mu^L \psi_L, \quad (2.23)$$

$$D_\mu^R \psi_R \rightarrow D_\mu^{R'} \psi'_R = e^{i\frac{Y}{2}\beta(x)} D_\mu^R \psi_R. \quad (2.24)$$

Essas transformações garantem que a Lagrangeana dos férmions seja invariável sob as transformações do grupo $SU(2)_L \times U(1)_Y$ [12].

Os férmions do Modelo Padrão podem ser classificados em dois grupos:

- Quarks:
 - Participam de todas as interações fundamentais (forte, fraca e eletromagnética).
 - Possuem três estados de cor.
 - Existem seis sabores: up (u), down (d), charm (c), strange (s), top (t), e bottom (b).
- Léptons:
 - Não possuem carga de cor e, portanto, não participam da interação forte.

- Incluem o elétron (e), múon (μ), tau (τ) e seus respectivos neutrinos (ν_e, ν_μ, ν_τ).

Os férmions são organizados em três gerações ou famílias:

- 1ª geração (matéria ordinária):

$$\left\{ \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, e_R, \begin{pmatrix} u \\ d \end{pmatrix}_L, u_R, d_R \right\}.$$

- 2ª geração:

$$\left\{ \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \mu_R, \begin{pmatrix} c \\ s \end{pmatrix}_L, c_R, s_R \right\}.$$

- 3ª geração:

$$\left\{ \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L, \tau_R, \begin{pmatrix} t \\ b \end{pmatrix}_L, t_R, b_R \right\}.$$

Essas gerações possuem as mesmas propriedades, diferenciando-se apenas pelas massas. As partículas da segunda e terceira gerações decaem rapidamente para as partículas da primeira geração, que constituem a matéria ordinária[12].

2.6 O Setor de Gauge

Desta parte da Lagrangeana derivam-se as interações tríplexes e quárticas entre os próprios bósons de gauge, além de sua dinâmica, descrita por:

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4}G_{\mu\nu}^a G^{a\mu\nu} - \frac{1}{4}B_{\mu\nu} B^{\mu\nu}, \quad (2.25)$$

onde:

$$G_{\mu\nu}^a = \partial_\mu W_\nu^a - \partial_\nu W_\mu^a + g\epsilon^{abc}W_\mu^b W_\nu^c, \quad (2.26)$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu, \quad (2.27)$$

com $a, b, c = 1, 2, 3$.

Nesta expressão, os bósons de gauge W_μ^a e B_μ são chamados de bósons de gauge simétricos, pois pertencem a uma Lagrangeana que é invariante sob as transformações do grupo $SU(2)_L \times U(1)_Y$ [13]. Esses bósons de gauge ainda não possuem massa. Adicionar termos de massa diretamente, como $m_W^2 W_\mu^a W^{a\mu}$, quebraria explicitamente a simetria $SU(2)_L \times U(1)_Y$.

Entretanto, experimentalmente sabe-se que os bósons mediadores da interação fraca ($W^\pm$ e Z^0) são bastante massivos, enquanto o fóton (γ), associado às interações eletromagnéticas, permanece sem massa [12]. Esse fato evidencia a necessidade de um

mecanismo para gerar massa apenas para os bósons de gauge que medeiam a interação fraca, preservando a simetria para o fóton.

Para garantir a invariância de gauge, as transformações infinitesimais dos campos de gauge são dadas por:

$$W_\mu^a \rightarrow W_\mu^{a'} = W_\mu^a - \frac{1}{g} \partial_\mu \alpha^a - \epsilon^{abc} \alpha^b W_\mu^c, \quad (2.28)$$

$$B_\mu \rightarrow B'_\mu = B_\mu + \partial_\mu \beta. \quad (2.29)$$

Ao aplicar essas transformações, verifica-se que termos de massa, como $m_W^2 W_\mu^a W^{a\mu}$, não permanecem invariantes, uma vez que seriam modificados como:

$$W_\mu^a W^{a\mu} \rightarrow \left(W_\mu^a - \frac{1}{g} \partial_\mu \alpha^a - \epsilon^{abc} \alpha^b W_\mu^c \right) \left(W^{a\mu} - \frac{1}{g} \partial^\mu \alpha^a - \epsilon^{abc} \alpha^b W^{\mu c} \right). \quad (2.30)$$

Como consequência, não é possível introduzir massas diretamente para os bósons de gauge simétricos sem quebrar a simetria $SU(2)_L \times U(1)_Y$. O mecanismo responsável por gerar as massas para os bósons de gauge $W^\pm$ e Z^0 , enquanto mantém o fóton sem massa, será discutido posteriormente. Ele é fundamental para explicar o alcance extremamente curto da interação fraca e a dinâmica associada às interações eletrofracas.

2.7 O Setor de Yukawa

Um termo de massa para os léptons carregados, do tipo $M\bar{\ell}\ell = M[\bar{\ell}_L\ell_R + \bar{\ell}_R\ell_L]$, violaria explicitamente a invariância de gauge. Para contornar essa dificuldade, exige-se que, antes da quebra espontânea de simetria, os léptons carregados sejam sem massa, e que suas massas surjam apenas após a quebra de simetria. No Modelo Padrão, a inclusão de termos de massa para os léptons carregados, de forma invariante de gauge, é realizada através do acoplamento do dubleto de léptons f_L^a com um dubleto de escalares Φ e com o singlete de mão direita ℓ_R^b :

$$\mathcal{L}_{Y,\text{léptons}} = -G_{ab} \left[(\bar{f}_L^a \Phi) \ell_R^b \right] + \text{H.C.}, \quad (2.31)$$

onde G_{ab} representam as constantes de acoplamento de Yukawa, e o termo H.C. (*Hermitian Conjugate*) corresponde ao conjugado hermitiano da expressão anterior, sendo incluído para assegurar que a Lagrangiana total seja real. Após a quebra espontânea de simetria, essa Lagrangeana transforma-se em:

$$\mathcal{L}_{Y,\text{léptons}} = - \left[\bar{\ell}_L^a \left(\frac{v_\Phi G_{ab}}{\sqrt{2}} \right) \ell_R^b \right] + \text{H.C.}, \quad (2.32)$$

sendo a matriz de massa para os léptons carregados dada por:

$$M_\ell = \frac{v_\Phi G_{ab}}{\sqrt{2}}. \quad (2.33)$$

2.8 A Quebra Espontânea de Simetria

A geração de massa para os bósons de gauge no Modelo Padrão é realizada por meio de uma quebra espontânea de simetria utilizando o mecanismo de Higgs. Esse mecanismo é implementado pela inclusão de uma Lagrangeana de campos escalares, invariante sob as transformações do grupo $SU(2)_L \times U(1)_Y$. A Lagrangeana escalar é construída a partir de um dubleto de escalares complexos Φ , definido como:

$$\Phi = \begin{pmatrix} \Phi^+ \\ \Phi^0 \end{pmatrix}, \quad Y_\Phi = 1. \quad (2.34)$$

A Lagrangeana escalar do Modelo Padrão é dada por:

$$\mathcal{L}_{\text{escalar}} = (D_\mu^L \Phi)^\dagger (D_\mu^L \Phi) - V(\Phi), \quad (2.35)$$

onde o potencial escalar é escrito como:

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2. \quad (2.36)$$

A quebra espontânea de simetria ocorre quando o campo escalar complexo Φ adquire um valor esperado no vácuo diferente de zero ($\langle \Phi^0 \rangle_0 \neq 0$). Isso leva à quebra da simetria $SU(2)_L \times U(1)_Y$ para $U(1)_{\text{EM}}$, preservando a invariância associada ao eletromagnetismo [13]. A configuração de campo que minimiza $V(\Phi)$, denotada como $\langle \Phi \rangle_0$, depende do sinal de μ^2 : - Para $\mu^2 > 0$, o potencial tem um mínimo na origem:

$$\langle \Phi \rangle_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Nesse caso, a simetria $SU(2)_L \times U(1)_Y$ não é quebrada. - Para $\mu^2 < 0$, o potencial adquire um mínimo não trivial:

$$\langle \Phi \rangle_0^\dagger \langle \Phi \rangle_0 = \frac{-\mu^2}{2\lambda}.$$

Parametrizando o campo escalar como:

$$\Phi = \begin{pmatrix} \phi^+ \\ (v + H + iG^0)/\sqrt{2} \end{pmatrix},$$

a condição de mínimo é satisfeita para $\phi^+ = G^0 = H = 0$, resultando em:

$$\langle \Phi \rangle_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix},$$

onde $v = \sqrt{-\mu^2/\lambda}$ é o valor esperado no vácuo (VEV) do campo escalar ϕ^0 .

O operador carga elétrica é definido como:

$$\exp(iQ)\langle\Phi\rangle_0 = \langle\Phi\rangle_0, \quad (2.37)$$

o que implica:

$$\begin{aligned} (1 + iQ)\langle\Phi\rangle_0 &= \langle\Phi\rangle_0, \\ \langle\Phi\rangle_0 + iQ\langle\Phi\rangle_0 &= \langle\Phi\rangle_0. \end{aligned}$$

Portanto:

$$Q\langle\Phi\rangle_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ v_\Phi/\sqrt{2} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (2.38)$$

Isso demonstra que a carga elétrica do vácuo é nula, ou seja, o operador carga elétrica aniquila o vácuo [13]. Essa condição garante que o fóton permaneça sem massa após a quebra de simetria, enquanto os demais bósons de gauge ($W^\pm$ e Z^0) adquirem massa devido ao mecanismo de Higgs.

2.9 Os Bósons de Gauge Físicos

A derivada covariante do Modelo Padrão pode ser escrita de forma matricial como:

$$D_\mu^L = \partial_\mu + \frac{ig}{2} \begin{pmatrix} W_\mu^3 & \sqrt{2}W_\mu^+ \\ \sqrt{2}W_\mu^- & -W_\mu^3 \end{pmatrix} + \frac{ig'}{2} Y B_\mu, \quad (2.39)$$

onde os campos $W_\mu^\pm$ são definidos como:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2).$$

Para o campo de Higgs com $Y_\phi = 1$, temos:

$$D_\mu^L\langle\Phi\rangle_0 = \begin{pmatrix} \partial_\mu + \frac{ig}{2}W_\mu^3 + \frac{ig'}{2}B_\mu & \frac{ig}{\sqrt{2}}W_\mu^+ \\ \frac{ig}{\sqrt{2}}W_\mu^- & \partial_\mu - \frac{ig}{2}W_\mu^3 + \frac{ig'}{2}B_\mu \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v_\Phi \end{pmatrix}. \quad (2.40)$$

Eliminando os termos relacionados à cinética do bóson de Higgs, restam os termos responsáveis pelas massas dos bósons de gauge:

$$D_\mu^L\langle\Phi\rangle_0 = \begin{pmatrix} \frac{ig}{\sqrt{2}}W_\mu^+v_\Phi \\ -\frac{iv_\Phi}{\sqrt{2}}(W_\mu^3 - tB_\mu) \end{pmatrix}, \quad (2.41)$$

com $t = g'/g$.

O termo quadrático $|D_\mu^L\langle\Phi\rangle_0|^2$ resulta na seguinte expressão:

$$|D_\mu^L\langle\Phi\rangle_0|^2 = \frac{g^2v_\Phi^2}{4}W_\mu^-W_\mu^+ + \frac{g^2v_\Phi^2}{8}(W_\mu^3W_\mu^3 + t^2B_\mu B_\mu - 2tB_\mu W_\mu^3). \quad (2.42)$$

Dessa expressão, os bósons $W_\mu^\pm$ aparecem como autoestados de massa, conhecidos como os bósons de gauge carregados. Já os bósons W_μ^3 e B_μ , associados aos geradores diagonais, se misturam [13]. A matriz de mistura na base $\{B_\mu, W_\mu^3\}$ é dada por:

$$\frac{g^2 v_\Phi^2}{8} \begin{pmatrix} t^2 & -t \\ -t & 1 \end{pmatrix}. \quad (2.43)$$

Os autovetores dessa matriz correspondem ao fóton A_μ e ao bóson neutro Z_μ^0 :

$$\begin{pmatrix} A_\mu \\ Z_\mu^0 \end{pmatrix} = \begin{pmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix}, \quad (2.44)$$

onde o ângulo de Weinberg (θ_W) é definido como:

$$\cos \theta_W = \frac{g}{\sqrt{g^2 + g'^2}}, \quad \sin \theta_W = \frac{g'}{\sqrt{g^2 + g'^2}}. \quad (2.45)$$

Os autovalores associados são:

$$\lambda_{A_\mu} = 0, \quad \lambda_{Z_\mu^0} = \frac{g^2 v_\Phi^2}{8} (t^2 + 1). \quad (2.46)$$

O fóton A_μ é, portanto, o único bóson de gauge sem massa, enquanto Z_μ^0 e $W_\mu^\pm$ adquirem massa devido ao mecanismo de Higgs [13].

2.10 O Tempo de Vida e Largura de Decaimento das Partículas

No Modelo Padrão, cada partícula instável possui um tempo de vida associado ao seu decaimento em partículas mais leves, respeitando as leis de conservação, como energia, momento, carga elétrica e carga de cor [14]. A largura de decaimento (Γ) descreve a taxa total de decaimento de uma partícula para todos os canais possíveis e está diretamente relacionada ao tempo de vida da partícula (τ) pela relação:

$$\tau = \frac{\hbar}{\Gamma}. \quad (2.47)$$

A largura de decaimento total (Γ) é dada pela soma das larguras parciais (Γ_i) de todos os canais de decaimento:

$$\Gamma = \sum_i \Gamma_i. \quad (2.48)$$

A fração de decaimentos que ocorre em um canal específico é descrita pelo branching ($\mathcal{B}_i$):

$$\mathcal{B}_i = \frac{\Gamma_i}{\Gamma}. \quad (2.49)$$

Como citado no capítulo anterior, a largura invisível do bóson Z é determinada utilizando um sinal de referência: a largura de decaimento do Z em léptons. A similaridade

entre os decaimentos do Z em neutrinos e em léptons carregados é explorada para calcular a largura invisível. Essa relação é dada por:

$$\Gamma(Z \rightarrow \text{inv}) = \frac{\sigma(Z + \text{jets})\mathcal{B}(Z \rightarrow \text{inv})}{\sigma(Z + \text{jets})\mathcal{B}(Z \rightarrow \ell\bar{\ell})}\Gamma(Z \rightarrow \ell\bar{\ell}). \quad (2.50)$$

onde:

- $\Gamma(Z \rightarrow \text{inv})$: largura de decaimento do bóson Z em neutrinos;
- $\sigma(Z + \text{jets})$: seção de choque do processo $Z + \text{jets}$;
- $\mathcal{B}(Z \rightarrow \text{inv})$: razão de ramificação do bóson Z decaindo em neutrinos;
- $\mathcal{B}(Z \rightarrow \ell\bar{\ell})$: razão de ramificação do bóson Z decaindo em díleptons carregados ($\ell\bar{\ell}$);
- $\Gamma(Z \rightarrow \ell\bar{\ell})$: largura de decaimento do bóson Z em pares de léptons.

2.11 O Número de Neutrinos

A largura invisível do bóson Z desempenha um papel crucial na determinação do número de famílias de neutrinos que se acoplam ao Z . Este número está diretamente relacionado à largura de decaimento invisível do Z para neutrinos (Γ_{inv}). No Modelo Padrão, assume-se que cada família de neutrinos (ν_e, ν_μ, ν_τ) contribui igualmente para Γ_{inv} [15]. Assim, o número efetivo de neutrinos (N_ν) pode ser calculado pela relação:

$$N_\nu = \frac{\Gamma_{\text{inv}}}{(\Gamma_{\nu\bar{\nu}})_{\text{SM}}}. \quad (2.51)$$

onde $(\Gamma_{\nu\bar{\nu}})_{\text{SM}}$ é a largura teórica de decaimento do Z em pares neutrino-antineutrino no Modelo Padrão. Para relacionar a largura invisível do Z com a largura observável de decaimento para díleptons ($\Gamma_{\ell\bar{\ell}}$), utiliza-se a razão experimental R^{exp} , definida como:

$$R^{\text{exp}} = \frac{\Gamma_{\text{inv}}}{\Gamma_{\ell\bar{\ell}}}, \quad (2.52)$$

Combinando essas informações, o número efetivo de neutrinos pode ser reescrito como:

$$N_\nu = R^{\text{exp}} \left(\frac{\Gamma_{\ell\bar{\ell}}}{\Gamma_{\nu\bar{\nu}}} \right)_{\text{SM}}, \quad (2.53)$$

onde $\left(\frac{\Gamma_{\nu\bar{\nu}}}{\Gamma_{\ell\bar{\ell}}} \right)_{\text{SM}} = 1.991 \pm 0.001$ é a razão calculada no Modelo Padrão entre as larguras de decaimento em díleptons e neutrinos [15].

Utilizando os valores experimentais [15]:

$$R^{\text{exp}} = 5.942 \pm 0.016, \quad (2.54)$$

temos:

$$N_\nu = \frac{5.942}{1.991} = 2.984 \pm 0.008. \quad (2.55)$$

Esse valor está em concordância com a existência de três famílias de férmions no Modelo Padrão [15].

3 O Grande Colisor de Hádrons

3.1 Introdução

O Large Hadron Collider (LHC) é um acelerador de partículas localizado no CERN, na fronteira entre a Suíça e a França, projetado para colidir prótons ou íons de chumbo a altas energias [1]. Ele foi fundamental na descoberta do bóson de Higgs pelos experimentos ATLAS e CMS e desempenha um papel essencial na busca por física além do Modelo Padrão, impondo limites aos parâmetros de novos cenários teóricos. O acelerador é composto por dois anéis com 27 km de circunferência, situados a cerca de 100 metros abaixo do solo [16].

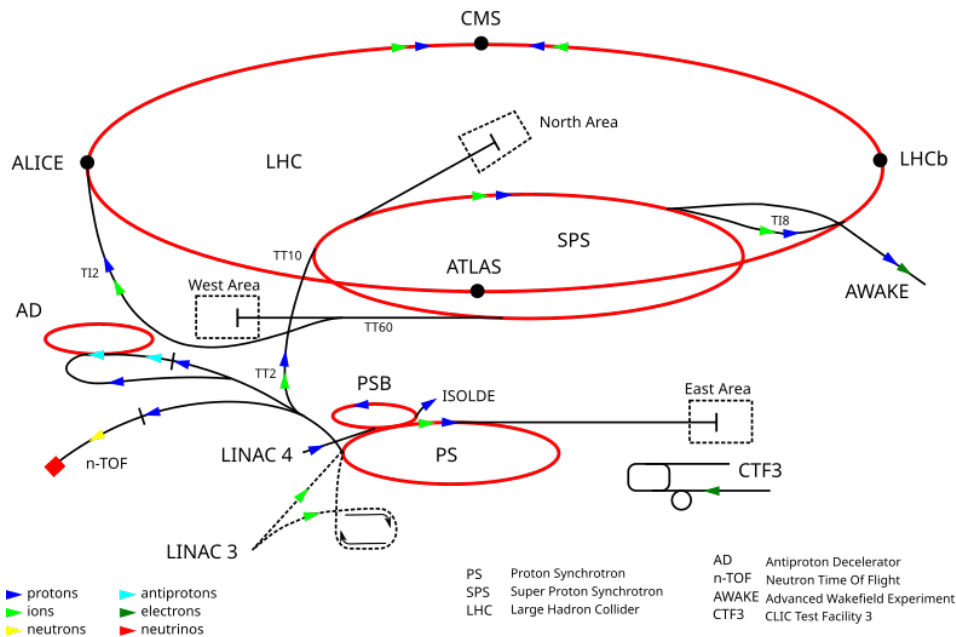


Figura 1 – Esquema do complexo de aceleradores do CERN. Fonte: [1].

O processo de aceleração começa com a ionização do gás hidrogênio, que separa os prótons de seus estados atômicos ligados [17]. Esses prótons são acelerados em estágios por meio de um complexo de aceleradores [1], composto por:

- Linac2: Um acelerador linear que eleva a energia dos prótons para 50 MeV.
- PS Booster (PSB): Um conjunto de quatro anéis síncrotrons que aumenta a energia para 1.4 GeV e intensifica o feixe de prótons.
- Proton Synchrotron (PS): Um síncrotron com 72 metros de raio que acelera os prótons para 25 GeV, organizando-os em pacotes de 4 ns espaçados por 25 ns.

- Super Proton Synchrotron (SPS): Um síncrotron de 6.9 km de circunferência que prepara os prótons com energias de até 450 GeV antes de injetá-los no LHC.

No LHC, os prótons são organizados em 2808 pacotes, separados por 25 ns, e atingem energias de até 6.5 TeV por feixe, totalizando 14 TeV no centro de massa durante as colisões [17].

O desempenho do LHC é caracterizado principalmente pela energia do feixe e pela luminosidade. A luminosidade (L) é um parâmetro fundamental para descrever o número de colisões por unidade de tempo e área. Para um feixe gaussiano, é descrita por:

$$L = \frac{N_b^2 n_b f_{\text{rev}} \gamma_r}{4\pi \epsilon_n \beta^*} F, \quad (3.1)$$

onde:

- N_b : número de prótons por pacote,
- n_b : número de pacotes por feixe,
- f_{rev} : frequência de revolução,
- γ_r : fator relativístico,
- ϵ_n : emitância transversal normalizada,
- β^* : função beta no ponto de colisão,
- F : fator de redução geométrica devido ao ângulo de cruzamento.

A taxa de eventos (N) é diretamente proporcional à luminosidade integrada e à seção de choque (σ):

$$N = L \cdot \sigma. \quad (3.2)$$

Durante o período de operação de 2016, o LHC atingiu luminosidades instantâneas de $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ [18].

3.2 O Experimento CMS

O CMS (*Compact Muon Solenoid*) é um dos dois maiores detectores multiuso do LHC, localizado no "Ponto 5" do anel do acelerador. Com dimensões de 21,6 m de comprimento, 14,6 m de diâmetro e um peso total de 12.500 toneladas, ele foi projetado para detectar partículas produzidas em colisões próton-próton com energias de até 14 TeV

[19]. Em 2016, o CMS registrou colisões com uma energia de centro de massa de 13 TeV e uma média de 25 colisões por cruzamento de pacotes, gerando cerca de 1000 partículas carregadas a cada 25 ns [18]. Esses dados foram utilizados na análise desta dissertação.

O CMS é composto por quatro subsistemas principais 2:

- Rastreador de Silício: responsável pela reconstrução de trajetórias de partículas carregadas.
- Calorímetro Eletromagnético (ECAL): mede a energia de partículas eletromagnéticas, como elétrons e fótons.
- Calorímetro Hadrônico (HCAL): responsável pela detecção e medição da energia de hádrons.
- Sistema de Múons: essencial para a identificação e medição de múons com alta precisão.

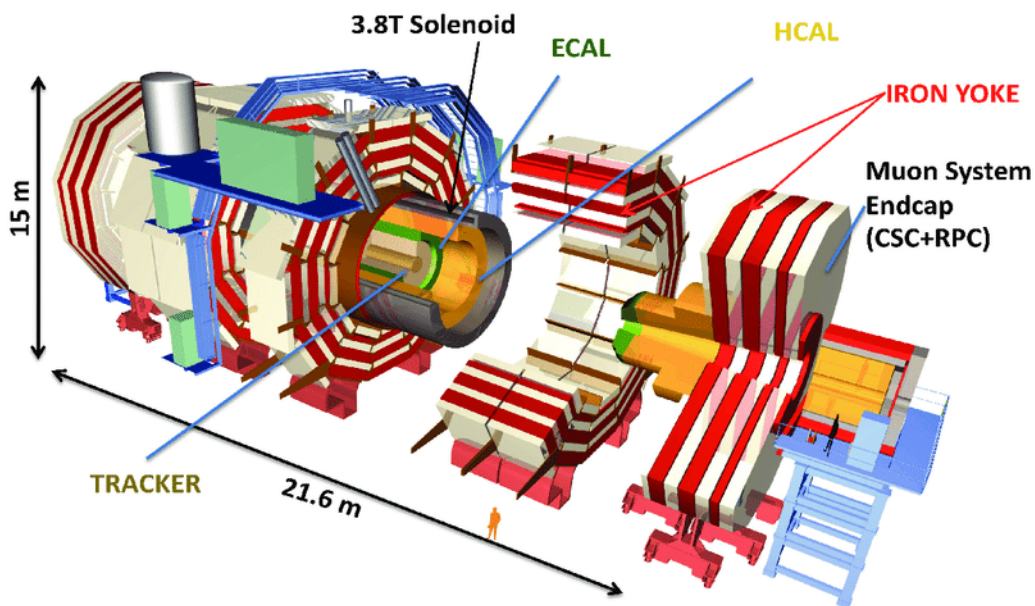


Figura 2 – Visão esquemática do detector CMS. Fonte: [2].

Além desses subsistemas, o CMS possui eletrônica avançada para leitura e aquisição de dados, permitindo o registro e a análise dos eventos das colisões.

3.2.1 Solenoide Supercondutor

O componente central do CMS é um solenoide supercondutor que gera um campo magnético de 3,8 T, cercado por um *yoke* de ferro¹ que estende o campo para a região externa, onde estão localizados os subdetectores de múons [19]. Esse campo magnético

¹ Yoke é uma estrutura de ferro usada para conter e redirecionar as linhas de campo magnético geradas pelo solenoide.

permite a curvatura dos trajetos de partículas carregadas, facilitando a determinação de sua carga elétrica e a medição precisa de seu momento.

Em conjunto com o rastreador e os calorímetros, o campo magnético desempenha um papel crucial na reconstrução do momento transversal (p_T) das partículas, contribuindo para o alto desempenho do CMS em análises de eventos complexos.

3.2.2 Rastreador de Silício (Silicon Tracker)

O rastreador de silício é o subdetector mais interno do CMS, projetado para medir a posição de partículas carregadas com alta precisão e cobre uma faixa de pseudo-rapidez de $|\eta| < 2.5$ [19]. Ele é essencial para reconstruir vértices primários e secundários, além dos rastros de partículas carregadas. O rastreador é composto por dois subsistemas principais:

- Módulos de Pixel: Localizados mais próximos ao ponto de interação, possuem três camadas cilíndricas no barril e dois discos em cada *endcap*. Com 66 milhões de pixels, oferecem alta granularidade e resolução espacial, cruciais para ambientes de alto fluxo de partículas.
- Tiras de Silício: Posicionadas ao redor do Pixel Tracker, consistem em 9,6 milhões de tiras de silício organizadas em camadas no barril e discos nos *endcaps*. Elas cobrem uma área ativa de 198 m^2 e fornecem informações complementares para a reconstrução tridimensional.

3.2.3 Calorímetro Eletromagnético (ECAL)

O ECAL 3 é um calorímetro homogêneo feito de cristais de tungstato de chumbo, projetado para medir depósitos de energia eletromagnética de elétrons e fótons [19]. Ele cobre até $|\eta| = 3.0$ e é dividido em três seções:

- Barril (EB): Cristais organizados em supermódulos, cobrindo até $|\eta| = 1.479$.
- *Endcaps* (EE): Estendem a cobertura até $|\eta| = 3.0$.
- Pré-chuveiro (ES): Instalado na frente dos *endcaps*, ajuda a distinguir fótons provenientes de decaimentos de π^0 .

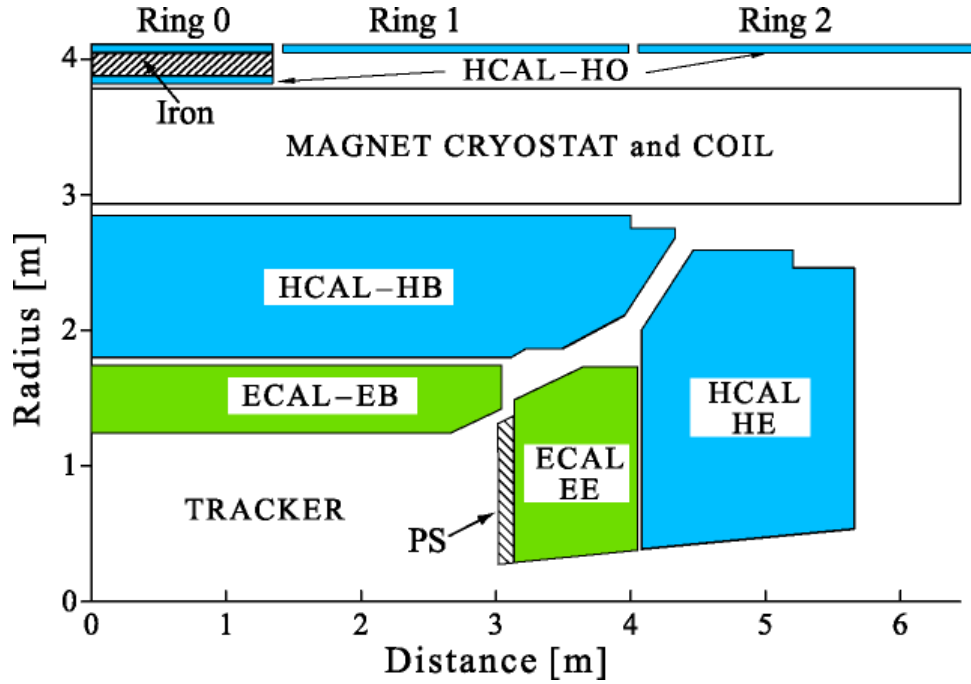


Figura 3 – Localização dos detectores ECAL e HCAL em uma seção transversal longitudinal de um quarto do detector. Fonte:[3]

3.2.4 Calorímetro Hadrônico (HCAL)

O HCAL [3](#) detecta partículas hadrônicas, cobrindo completamente até $|\eta| = 5.0$ [\[19\]](#). Ele é dividido em quatro subdetectores:

- HB (Barril): Localizado dentro do solenoide.
- HE (*Endcaps*): Complementa o barril nas regiões $|\eta| > 1.4$.
- HF (Avançado): Detecta partículas na região mais avançada do detector, $|\eta| > 3.0$.
- HO (Externo): Localizado fora do solenoide para capturar energia residual.

3.2.5 Sistema de Múons

Os múons, devido à sua alta penetração, são detectados no sistema mais externo do CMS. Ele é composto por três tipos de detectores gasosos [4](#):

- Tubo de Deriva (DT): Localizado no barril, oferece excelente resolução espacial.
- Câmaras de Tira de Cátodo (CSC): Utilizadas nos *endcaps*, fornecem reconstrução tridimensional precisa.
- Câmaras de Placas Resistivas (RPC): Com alta resolução temporal, são utilizadas para disparar eventos no sistema de *trigger*.

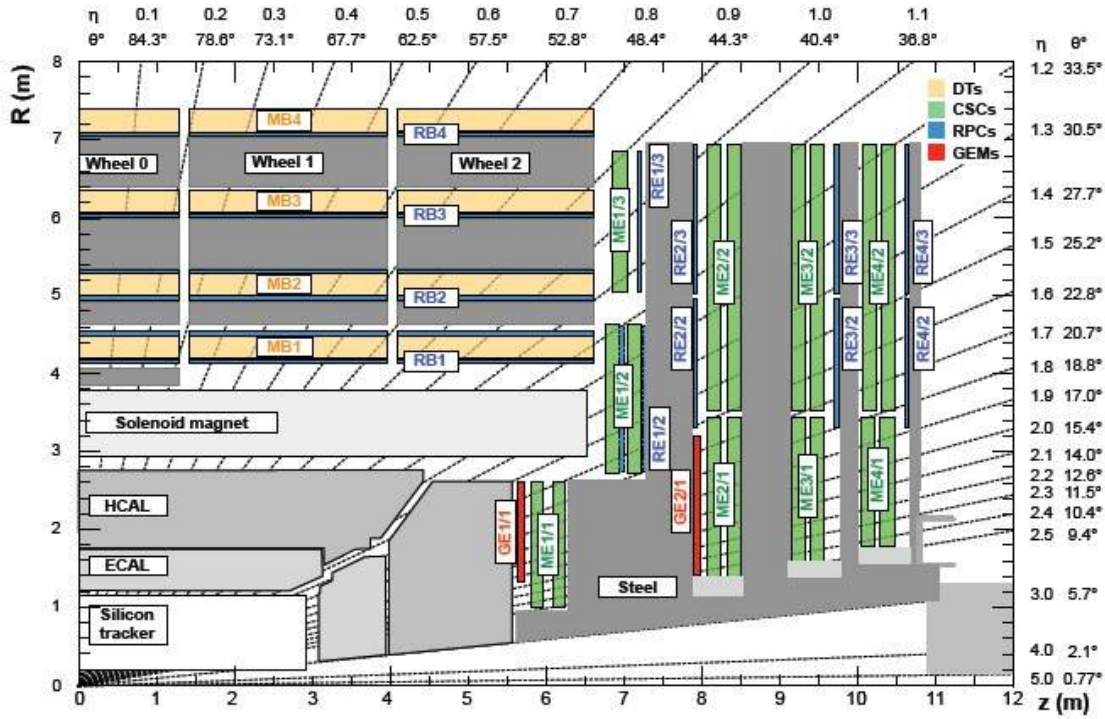


Figura 4 – Esquema dos subdetectores do CMS com destaque para os detectores de múons (DTs, CSCs, RPCs e GEMs) em uma visão longitudinal. Fonte: [4].

3.2.6 Sistema de *Trigger* e Aquisição de Dados

Devido à taxa de colisões de 40 MHz no LHC, o CMS produz até 40 TB de dados por segundo. Para lidar com esse volume, o sistema de *trigger* realiza uma seleção em dois níveis:

- *Trigger* de Nível 1 (L1): Implementado em hardware, reduz a taxa de eventos para 100 kHz.
- *Trigger* de Alto Nível (HLT): Executado em software, aplica algoritmos complexos e reduz a taxa para 100 eventos por segundo, armazenando apenas os eventos mais relevantes para análise.

3.2.7 Sistema de Aquisição de Dados (DAQ)

O sistema de aquisição de dados (DAQ) do CMS é responsável por coletar, processar e armazenar os eventos selecionados pelo sistema de *trigger*, garantindo que apenas os eventos mais relevantes sejam preservados para análise posterior. Cada subsistema do CMS possui eletrônicos de leitura que registram as informações de um evento individual. Essas informações são inicialmente armazenadas no formato bruto (RAW), passando por

várias etapas de processamento até atingir o formato *mini-AOD*², amplamente utilizado pela comunidade de física de partículas.

3.2.8 Reconstrução de Eventos

A reconstrução de eventos no experimento CMS combina informações de trilhas no rastreador, depósitos de energia nos calorímetros e dados dos detectores de múons. Utilizando o algoritmo *Particle Flow* (PF) [20], partículas individuais como múons, elétrons, fótons, e hádrons carregados e neutros são identificadas. A partir dessas partículas reconstruídas, objetos compostos, como jatos, léptons tau e a energia transversal ausente (p_T^{miss}), são calculados [5].

O p_T^{miss} , obtido como a soma vetorial negativa dos p_T de todas as partículas, é essencial para detectar partículas neutras que escapam sem interação, como neutrinos, e para identificar sinais de física além do Modelo Padrão.

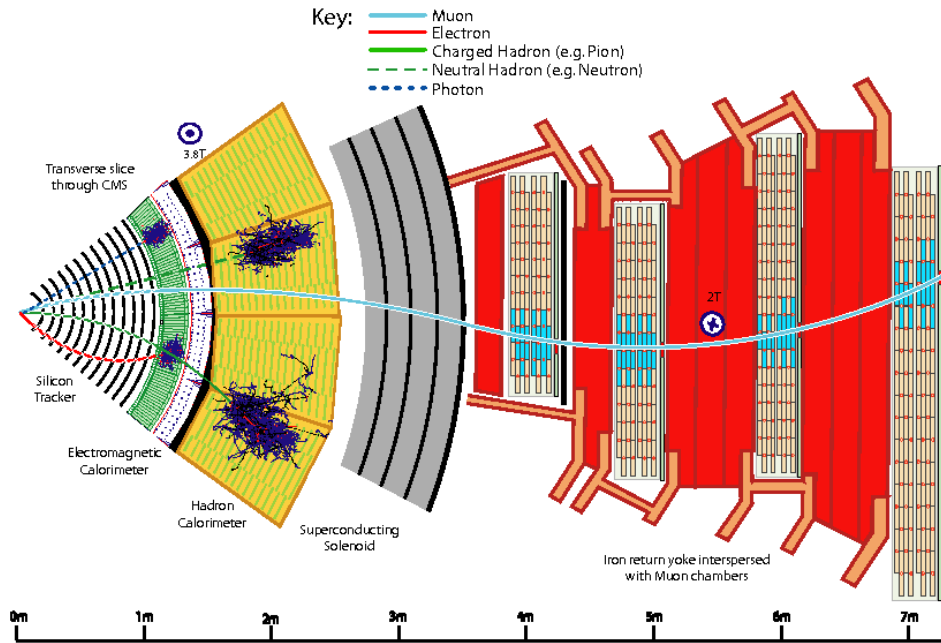


Figura 5 – Esquema transversal do detector CMS ilustrando a reconstrução de eventos.
Fonte: [5]

² O Mini Analysis Object Data é um formato de dados reduzido e otimizado desenvolvido pela colaboração CMS. Ele contém as informações essenciais para a maioria das análises físicas, como partículas reconstruídas, vértices e variáveis associadas.

4 Resultados e Discussões

Para determinar a largura invisível do Z , são analisados eventos do sinal $Z(\rightarrow \nu\nu) + \text{jets}$, caracterizados por jatos recuando contra um sistema invisível (p_T^{miss}). A Figura 6 ilustra esse tipo de processo por meio de um diagrama de Feynman simplificado. Adicionalmente, regiões de controle são definidas para estimar contribuições de processos com assinaturas semelhantes ao sinal.

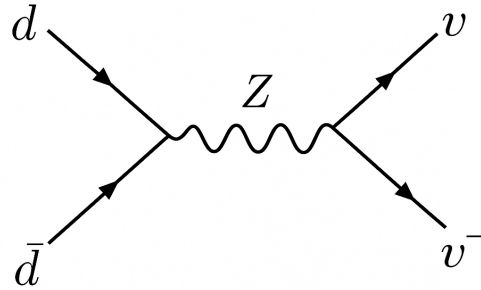


Figura 6 – **Diagrama de Feynman** ilustrando um processo típico em colisões próton-próton (pp), no qual um quark d e um antiquark $\bar{d}$, originados dos prótons incidentes, aniquilam-se gerando um bóson Z . O bóson Z então decai de forma invisível em um par neutrino-antineutrino ($\nu\bar{\nu}$). Embora não representados neste diagrama, jatos hadrônicos ($Z + \text{jets}$) geralmente acompanham o processo, sendo produzidos por radiação de glúons (QCD) emitidos pelos quarks iniciais. A presença de jatos, combinada com o momento transversal ausente (p_T^{miss}) resultante dos neutrinos, caracteriza a assinatura experimental desse processo. *Fonte: o autor.*

O decaimento do Z em léptons carregados ($Z \rightarrow \ell^+\ell^-$) é utilizado como sinal de referência para calcular a largura invisível do Z , enquanto o decaimento do W em um lépton e um neutrino ($W \rightarrow \ell\nu$) representa o principal *background*.

As regiões de controle estabelecidas são as seguintes:

- Single Muon, Single Electron e Single Tau: usadas para estimar o *background* $W \rightarrow \ell\nu$.
- Double Muon e Double Electron: utilizadas para estimar o sinal de referência $Z \rightarrow \ell^+\ell^-$.

- Região QCD: destinada a estimar o *background* de multijatos QCD.

Essas regiões são utilizadas para calcular os fatores de transferência (*transfer factors*), que estabelecem a relação entre os eventos observados nas regiões de controle e os eventos esperados na região de sinal. Esses fatores permitem extrapolar as contribuições de *backgrounds* para a região de sinal, com base nas observações realizadas nas regiões de controle.

Por fim, realiza-se um ajuste simultâneo, combinando as informações das regiões de controle e da região de sinal. Esse procedimento possibilita uma determinação precisa da largura invisível do Z .

4.1 Datasets

Os arquivos `.root` utilizados nesta análise foram obtidos por meio do portal Open Data do CERN, disponível em: <https://opendata.cern.ch>. Os dados utilizados correspondem a colisões próton-próton com energia no centro de massa de $\sqrt{s} = 13$ TeV, registradas pelo experimento CMS durante o ano de 2016.

No portal Open Data do CERN, estão disponíveis apenas os dados referentes às eras G e H (períodos específicos de coleta de dados), o que representa uma fração do total de dados coletados naquele ano. Essas eras correspondem aos *runs* no intervalo 278820–284044. Os dados estão disponibilizados no formato **NanoAOD**, uma estrutura compacta e otimizada que contém apenas os objetos físicos de alto nível (como múons, elétrons, fótons, *jets*, MET etc.). Esse formato é ideal para análises em larga escala, permitindo leitura eficiente e uso direto com ferramentas como `uproot` [21] e `awkward-array` [22].

Os conjuntos de dados de colisão próton-próton selecionados para esta análise foram: *single muon* (que garante a presença de pelo menos um múon no evento), *single electron* (que garante a presença de pelo menos um elétron no evento) e MET (*Missing Transverse Energy*) como mostra a Tabela 2. A Tabela 1 apresenta quais regiões de controle serão definidas por cada dataset.

Tabela 1 – Conjuntos de dados primários usados para a seleção das diferentes regiões de análise.

Datasets Primários	Regiões
MET	Jets + p_{miss} Single Muon Double Muon Single Tau QCD
Single Muon	Single Muon Double Muon
Single Electron	Single Electron Double Electron

Durante a tomada de dados do experimento CMS, a informação sobre a qualidade dos *runs* é registrada em um arquivo no formato JSON. Esse arquivo permite, durante a fase de análise, a seleção apenas dos eventos provenientes de *runs* considerados válidos, assegurando a integridade e a confiabilidade dos dados utilizados.

Para assegurar a certificação dos dados nesta análise, foi utilizado o arquivo JSON `Cert_278820-284044_13TeV_Legacy2016_Collisions16GH_JSON.txt`, previamente editado de forma a conter exclusivamente as informações correspondentes às eras G e H.

Além dos dados reais, esta análise também utilizou simulações de Monte Carlo (MC) para modelar os processos de sinal e de *background*, disponíveis no portal *CERN Open Data*. Foram incluídos os seguintes processos:

- Sinal: $Z \rightarrow \nu\bar{\nu}$ (Tabela 3),
- Sinal de referência: $Z \rightarrow \ell^+\ell^-$ (Tabela 3),
- *background* principal: $W + \text{jets}$ (Tabela 4),
- *backgrounds*: *Diboson*, $t\bar{t}$ (Tabela 5),
- QCD multijatos (Tabela 5),
- *backgrounds* menos relevantes: VBS, *single top* (Tabela 5).

As amostras de Monte Carlo (MC) utilizadas nessa análise foram geradas com o gerador `MadGraph5_aMC@NLO`¹, acoplado ao `Pythia8` para a simulação do processo de hadronização. Utilizou-se o conjunto de parâmetros de tunagem CP5 (`TuneCP5_13TeV`), desenvolvido com base em dados de colisões do LHC a 13 TeV e utilizando os PDFs NNPDF3.1 NNLO [23]. Enquanto isso, as amostras empregadas na análise publicada pelo

¹ *MadGraph* é um gerador de eventos usado em física de partículas para simular colisões de partículas com base em qualquer modelo que possa ser escrito na forma de uma Lagrangiana.

CMS [8] foram geradas com o tune `CUETP8M1_13TeV`, baseado nos PDFs NNPDF2.3 LO e calibrado com dados a 7 TeV [23]. O tune `CUETP8M1` foi amplamente empregado nas simulações da colaboração durante as fases iniciais do *Run 2*.

Ambos os tunes são validados e reconhecidos pela colaboração CMS. No entanto, suas diferenças — especialmente no conjunto de PDFs, no tratamento das múltiplas interações e na parametrização da radiação [23] — podem levar a variações sutis em distribuições cinemáticas e taxas de eventos. Embora geralmente permaneçam dentro das incertezas sistemáticas atribuídas à modelagem teórica.

As amostras utilizadas nesta análise foram geradas na campanha de simulação `RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v2`, que faz parte da versão *Ultra Legacy* do CMS para o ano de 2016. Essa campanha corresponde à última de reprocessamento completo dos dados simulados do Run 2, incorporando todas as correções e atualizações disponíveis para a reconstrução de objetos físicos [24]. O sufixo `asymptotic` indica que os ajustes de calibração e condições de operação do detector refletem o desempenho ideal e estável do sistema ao longo do período analisado.

Em relação a simulação Monte Carlo do sinal, o portal CERN Open Data não disponibiliza o dataset `DYJetsToNuNu`, comumente utilizado em análises do experimento CMS. No entanto, oferece a amostra `Z2JetsToNuNu`, que também descreve o processo $Z \rightarrow \nu\bar{\nu}$ acompanhado de jatos, gerado no *MadGraph* a partir do processo:

```
generate pp > nu nu~ j j [QCD]
```

A principal distinção entre os dois datasets reside na forma como os jatos adicionais são incluídos. O `DYJetsToNuNu` é enriquecido com jatos, utilizando a seguinte sequência de comandos no *MadGraph*:

```
generate pp -> nu nu~ [QCD] @0,
add process pp -> nu nu~ j [QCD] @1,
add process pp -> nu nu~ j j [QCD] @2.
```

Por sua vez, a amostra `Z2JetsToNuNu` aplica um corte na massa invariante do sistema leptônico (`lepton mass > 50 GeV`), o qual não impacta esta análise, uma vez que se restringe à região de massa do bóson Z . Dessa forma, o dataset `Z2JetsToNuNu` foi adotado como substituto adequado para a simulação do sinal $Z \rightarrow \nu\bar{\nu}$.

As informações referentes às seções de choque das amostras de Monte Carlo estão disponíveis nas páginas dos datasets no formato `MINIAOD`, acessíveis por meio do portal Open Data do CERN. Embora esta análise utilize arquivos no formato `NANOAOD`, as páginas correspondentes a esse formato não têm os valores de seção de choque para alguns datasets.

Tabela 2 – Lista dos conjuntos de dados primários utilizados nesta análise.

Dataset
/MET/Run2016G-UL2016_MiniAODv2_NanoAODv9-v1/NANOAOB
/MET/Run2016H-UL2016_MiniAODv2_NanoAODv9-v1/NANOAOB
/SingleMuon/Run2016G-UL2016_MiniAODv2_NanoAODv9-v1/NANOAOB
/SingleMuon/Run2016H-UL2016_MiniAODv2_NanoAODv9-v1/NANOAOB
/SingleElectron/Run2016G-UL2016_MiniAODv2_NanoAODv9-v1/NANOAOB
/SingleElectron/Run2016H-UL2016_MiniAODv2_NanoAODv9-v1/NANOAOB

Tabela 3 – Amostras de eventos simulados para os processos de sinal

Dataset	Cross section [fb]
/Z1JetsToNuNu_M-50_LHEFilterPtZ-50To150_[1]/[2]_v17-v1/[3]	577600
/Z1JetsToNuNu_M-50_LHEFilterPtZ-150To250_[1]/[2]_v17-v1/[3]	17410
/Z1JetsToNuNu_M-50_LHEFilterPtZ-250To400_[1]/[2]_v17-v1/[3]	1978
/Z1JetsToNuNu_M-50_LHEFilterPtZ-400ToInf_[1]/[2]_v17-v1/[3]	217.1
/Z2JetsToNuNu_M-50_LHEFilterPtZ-50To150_[1]/[2]_v17-v2/[3]	315100
/Z2JetsToNuNu_M-50_LHEFilterPtZ-150To250_[1]/[2]_v17-v2/[3]	28830
/Z2JetsToNuNu_M-50_LHEFilterPtZ-250To400_[1]/[2]_v17-v2/[3]	4976
/Z2JetsToNuNu_M-50_LHEFilterPtZ-400ToInf_[1]/[2]_v17-v2/[3]	815.1
/DYJetsToLL_LHEFilterPtZ-100To250_[1]/[2]_v17-v2/[3]	96620
/DYJetsToLL_LHEFilterPtZ-250To400_[1]/[2]_v17-v2/[3]	3738
/DYJetsToLL_LHEFilterPtZ-400To650_[1]/[2]_v17-v2/[3]	505
/DYJetsToLL_LHEFilterPtZ-650ToInf_[1]/[2]_v17-v2/[3]	47.63

[1] _TuneCP5_13TeV-amcatnloFXFX-pythia8

[2] RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic

[3] NANOAOB

Tabela 4 – Amostras de eventos simulados para os processos de *background* dominante

Dataset	Cross section [fb]
/WJetsToLNu_Pt-100To250_[1]/[2]_v17-v1/[3]	757800
/WJetsToLNu_Pt-250To400_[1]/[2]_v17-v1/[3]	27490
/WJetsToLNu_Pt-400To600_[1]/[2]_v17-v1/[3]	3501
/WJetsToLNu_Pt-600ToInf_[1]/[2]_v17-v1/[3]	542.7
/WZTo2Q2L_mllmin4p0_[1]/[2]_v17-v2/[3]	6429
/WZTo1L1Nu2Q_4f_[1]/[2]_v17-v1/[3]	9159
/WZTo1L3Nu_4f_[1]/[2]_v17-v1/[3]	3415
/WZTo3LNu_[1]/[2]_v17-v1/[3]	5218
/WWTo1L1Nu2Q_4f_[1]/[2]_v17-v1/[3]	50890
/WWTo2L2Nu_[1]/[2]_v17-v1/[3]	11090
/WWTo4Q_4f_[1]/[2]_v17-v1/[3]	51570
/ZZTo2Nu2Q_5f_[1]/[2]_v17-v1/[3]	4557
/ZZTo2L2Nu_[1]/[2]_v17-v1/[3]	973.8
/ZZTo4L_[1]/[2]_v17-v1/[3]	1325
/TTJets_[1]/[2]_v17-v1/[3]	756100

[1] _TuneCP5_13TeV-amcatnloFXFX-pythia8

[2] RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic

[3] NANOAOBSIM

Tabela 5 – Amostras de eventos simulados para os processos de *background* secundários

Dataset	Cross section [fb]
/QCD_Pt_15to30_[1]/[2]_v17-v1/[3]	1.245×10^{12}
/QCD_Pt_30to50_[1]/[2]_v17-v1/[3]	1.069×10^{11}
/QCD_Pt_50to80_[1]/[2]_v17-v1/[3]	1.571×10^{10}
/QCD_Pt_80to120_[1]/[2]_v17-v1/[3]	2.341×10^9
/QCD_Pt_120to170_[1]/[2]_v17-v1/[3]	4.074×10^8
/QCD_Pt_170to300_[1]/[2]_v17-v1/[3]	1.037×10^8
/QCD_Pt_300to470_[1]/[2]_v17-v1/[3]	6.832×10^6
/QCD_Pt_470to600_[1]/[2]_v17-v1/[3]	5.519×10^5
/QCD_Pt_600to800_[1]/[2]_v17-v1/[3]	1.565×10^5
/QCD_Pt_1000to1400_[1]/[2]_v17-v1/[3]	7.475×10^3
/QCD_Pt_1400to1800_[1]/[2]_v17-v1/[3]	6.482×10^2
/QCD_Pt_1800to2400_[1]/[2]_v17-v1/[3]	8.746×10^1
/QCD_Pt_2400to3200_[1]/[2]_v17-v1/[3]	5.235
/QCD_Pt_3200toInf_[1]/[2]_v17-v1/[3]	0.1352
/WGToLNUG_01J_5f_[1]/[2]_v17-v1/[3]	1.912×10^5
/ZGToLLG_01J_5f_[1]/[2]_v17-v1/[3]	5.144×10^4
/ZGTo2NuG_[1]/[2]_v17-v1/[3]	3.022×10^4
/ST_tW_antitop_5f_[1]/[2]_v17-v1/[3]	3.251×10^4
/ST_tW_top_5f_[1]/[2]_v17-v1/[3]	3.245×10^4
/ST_t-channel_antitop_4f_[1]/[2]_v17-v1/[3]	6.793×10^4
/ST_t-channel_top_4f_[1]/[2]_v17-v1/[3]	1.134×10^5
/EWKWMinus2Jets_[1]/[2]_v17-v1/[3]	3.210×10^4
/EWKWPlus2Jets_[1]/[2]_v17-v1/[3]	3.908×10^4
/EWKZ2Jets_ZToLL_[1]/[2]_v17-v1/[3]	6.206×10^3
/EWKZ2Jets_ZToNuNu_[1]/[2]_v17-v1/[3]	1.065×10^4

[1] `_TuneCP5_13TeV-amcatnloFXFX-pythia8`

[2] `RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic`

[3] `NANOADSIM`

4.2 Trigger e Luminosidade

Para a seleção dos datasets primários, são utilizados *triggers* sem *prescale* (unprescaled *triggers*). O *prescale* de um trigger é um fator que define a frequência com que o sistema de aquisição de dados (DAQ) grava eventos que atendem a determinados critérios de seleção. Quando o *prescale* de um trigger é igual a 1, todos os eventos que satisfazem os critérios desse trigger são armazenados. Assim, um trigger sem *prescale* é aquele em que o fator de *prescale* é igual a 1, garantindo que todos os eventos que atendem aos critérios definidos sejam gravados para análise. O *prescale* dos *triggers* foi obtido com base nas instruções fornecidas pelo CMS Open Data [25]. A Tabela 6 lista os *triggers* utilizados para cada conjunto de dados primário.

Tabela 6 – Conjuntos de dados primários e seus correspondentes *triggers*.

Datasets Primários	<i>triggers</i>
MET	HLT_PFMET170_HBHECleaned OU
	HLT_PFMET170_HBHE_BeamHaloCleaned OU
	HLT_PFMETNoMu120_PFMHTNoMu120_IDTight OU
	HLT_MET75_IsoTrk50
Single Muon	HLT_IsoMu24 OU HLT_IsoTkMu24
Single Electron	HLT_Ele27_WPTight_Gsf

A partir das instruções do CMS Open Data [26], utilizando os *triggers* mencionados acima e a ferramenta BRILCalc ², foi calculada a luminosidade integrada para os datasets primários utilizados, que resultou em 16.290 fb^{-1} .

4.3 Seleção de Eventos

Nesta análise, o objetivo é selecionar eventos compatíveis com a produção do bóson Z decaindo de forma invisível, bem como caracterizar os principais *backgrounds* associados. A assinatura típica desse sinal é composta por jatos de alta energia recuando contra o vetor de momento transverso ausente ($\vec{p}_T^{\text{miss}}$). Para descrever essa configuração cinemática, introduz-se a variável **vetor de recuo**, definida por:

$$\vec{U} = \vec{p}_T^{\text{miss}} + \sum_{\ell} \vec{p}_T^{\ell}$$

Nos eventos sem léptons carregados, $\vec{U}$ coincide com $\vec{p}_T^{\text{miss}}$, de modo que $U = |\vec{p}_T^{\text{miss}}|$.

A seleção de eventos é guiada por um conjunto de critérios básicos (*baseline*), válidos para todas as regiões de análise:

- $U > 200 \text{ GeV}$;
- O jato líder deve apresentar $p_T > 200 \text{ GeV}$;
- Eventos com jatos de $p_T > 40 \text{ GeV}$ e pseudorapidez $|\eta| > 2,4$ são vetados, visando suprimir processos de espalhamento de bósons vetoriais.

Para garantir a qualidade da reconstrução, exige-se consistência entre p_T^{miss} e U , por meio da variável:

$$\Delta = \frac{|p_{T\text{calo}}^{\text{miss}} - U|}{U}$$

Eventos com $\Delta > 0,5$ são excluídos. Adicionalmente, aplicam-se os seguintes vetos:

² O BRILCalc é uma ferramenta desenvolvida pelo CMS para calcular a luminosidade e monitorar a qualidade dos dados em experimentos de física de partículas, utilizando informações de feixes e *triggers*.

- Fótons isolados com $p_T > 25 \text{ GeV}$ e $|\eta| < 2,5$ são vetados para suprimir processos com radiação inicial e produção multijato;
- Jatos identificados como provenientes de quarks b (b -tagged) são rejeitados, utilizando o discriminador CSVv2 com ponto de operação médio ($\text{CSVv2} > 0,8$), conforme [27];
- A separação azimutal mínima entre os quatro jatos líderes e $\vec{p}_T^{\text{miss}}$, denotada $\Delta\phi_{\text{min}}$, deve satisfazer $\Delta\phi_{\text{min}} > 0,5$, a fim de reduzir o fundo QCD multijato.

Após a aplicação dos critérios *baseline*, os eventos selecionados são categorizados em diferentes regiões de análise, cada uma otimizada para estudar diferentes assinaturas do sinal e controlar contribuições de fundo específicas. A seguir, detalham-se os critérios específicos adotados para cada uma dessas regiões.

- **Região p_T^{miss} + jatos:** visa selecionar eventos sem léptons carregados, característicos de decaimentos invisíveis do bóson Z . Os seguintes vetos são aplicados:
 - Múons ou elétrons isolados com $p_T > 10 \text{ GeV}$ e $|\eta| < 2,5$;
 - Taus com $p_T > 20 \text{ GeV}$ e $|\eta| < 2,3$.
- **Região de dois léptons ($\ell^+\ell^-$):** tem como objetivo identificar eventos com dois léptons isolados no estado final, compatíveis com o decaimento do bóson $Z \rightarrow \ell^+\ell^-$, aplicando os seguintes critérios de seleção:
 - Um par de múons ou elétrons isolados com massa invariante $71 < m_{\ell\ell} < 111 \text{ GeV}$;
 - Múons com $p_T > 25 \text{ GeV}$ e elétrons com $p_T > 30 \text{ GeV}$.
- **Regiões μ +jatos e e +jatos:** destinadas à seleção de eventos com um único lépton isolado no estado final, compatíveis com o decaimento do bóson $W \rightarrow \ell\nu$.
 - **Região μ +jatos:** requer um múon isolado com $p_T > 25 \text{ GeV}$ e $|\eta| < 2,4$;
 - **Região e +jatos:**
 - * Um elétron com $p_T > 30 \text{ GeV}$;
 - * $p_T^{\text{miss}} > 100 \text{ GeV}$ para reduzir o fundo de QCD multijato.
 - Ambas as regiões exigem uma massa transversa no intervalo $30 \leq M_T(\ell, p_T^{\text{miss}}) < 125 \text{ GeV}$.
- **Região Tau + jatos:** voltada para a seleção de eventos com um único tau hadronicamente decaído e isolado, compatíveis com o processo $W \rightarrow \tau\nu$.
 - Um tau isolado com $p_T > 40 \text{ GeV}$, $|\eta| < 2,3$;

– Massa transversa no intervalo $30 < m_T < 125$ GeV.

- **Região QCD:** utilizada como região de controle para o fundo de multijatos. Define-se por:

– Inversão do corte sobre o ângulo azimutal mínimo entre os quatro jatos líderes e o vetor p_T^{miss} :

$$\min [\Delta\phi(j_1, j_2, j_3, j_4, p_T^{\text{miss}})] \leq 0,5$$

Cada uma das regiões de análise descritas inclui, além dos critérios específicos, um conjunto de cortes de veto aplicados com o objetivo de suprimir contribuições indesejadas de processos de fundo. Esses cortes visam garantir a pureza da amostra e a robustez da seleção. A descrição completa de todos os critérios de seleção, incluindo vetos e requisitos cinemáticos adotados em cada região, encontra-se detalhada na Tabela 20.

4.3.1 Fluxo de Cortes

As tabelas a seguir mostram o fluxo de cortes (cut flow) para os eventos selecionados em cada região.

Tabela 7 – Fluxo de cortes da região p_T^{miss} - Parte 1

Cortes	Z2JetsToNuNu	Z1JetsToNuNu	WJetsToLNu	WZTo1L3Nu	WWTo2L2Nu
Todos os eventos	98.474.753	35.910.730	273.949.797	1.229.946	2.900.000
Luminosidade	98.474.753	35.910.730	273.949.797	1.229.946	2.900.000
Trigger	84.053.704	27.049.364	99.232.003	136.417	183.010
$U > 200$ GeV	52.424.077	13.140.991	62.785.171	61.743	82.492
$\frac{ \text{calo_MET}-U }{U} < 0,5$	51.495.863	12.993.015	27.800.151	41.376	28.778
$N_{\text{jets}} \geq 4$	37.037.380	4.072.441	21.917.439	30.120	23.803
Jato líder $p_T > 200$ GeV	24.765.355	3.039.123	15.334.976	17.981	12.974
Veto de fótons	17.711.990	2.586.726	8.238.960	9.660	5.179
Veto de b-jets	15.311.766	2.356.936	7.183.083	8.551	4.510
Veto de jatos $ \eta > 2,4$	2.895.832	208.024	1.557.071	1.651	1.094
Veto de múons	2.845.723	205.413	908.096	870	313
Veto de elétrons	2.322.725	189.616	600.304	565	149
Veto de taus	2.093.093	181.799	341.316	340	61
$\Delta\phi(j_1, j_2, j_3, j_4, U) > 0,5$	2.093.093	181.799	341.316	340	61

Tabela 8 – Fluxo de cortes da região p_T^{miss} - Parte 2

Cortes	ZZTo2L2Nu	ZGTo2NuG	EWKZ2Jets	Total MC	MET	DATA/MC
Todos os eventos	15.928.000	1.183.548	1.500.000	431.076.774	66.747.616	0,154
Luminosity Mask	15.928.000	1.183.548	1.500.000	431.076.774	63.250.532	0,146
Trigger selection	592.175	54.959	159.654	211.461.286	26.666.974	0,126
$U \ p_T > 200$	194.712	14.559	54.644	128.758.389	4.330.577	0,033
$(\text{calo_MET} - U /U) < 0,5$	106.156	14.074	52.694	92.532.107	2.057.038	0,022
$N_{\text{jets}} \geq 4$	80.824	8.810	34.665	63.205.482	1.456.182	0,023
Jato líder $p_T > 200 \text{ GeV}$	26.935	4.167	18.893	43.220.404	677.544	0,015
Veto de fótons	12.741	1.562	12.534	28.579.352	370.927	0,012
Veto de b-jets	10.967	1.399	10.727	24.887.939	261.535	0,010
Jet veto $ \eta > 2,4$	2.497	241	1.382	4.667.792	45.595	0,009
Veto de múons	996	238	1.351	3.963.000	33.770	0,008
Veto de elétrons	648	178	1.086	3.115.271	24.954	0,008
Veto de taus	276	167	878	2.617.930	19.598	0,007
$\Delta\phi(j_{1,2,3,4}, U) > 0,5$	276	167	878	2.617.930	19.598	0,007

Tabela 9 – Fluxo de cortes da região single muon - Parte 1

Corte	WJetsToLNu	WZTo1L1Nu2Q	WZTo1L3Nu	WZTo3LNu	WWTo1L1Nu2Q	WWTo2L2Nu	ZZTo2L2Nu	TTJets	WGToLNuG
Todos os eventos	273.949.797	3.690.271	1.229.946	10.441.724	19.976.139	2.900.000	15.928.000	89.003.534	55.939.475
Luminosidade	273.949.797	3.690.271	1.229.946	10.441.724	19.976.139	2.900.000	15.928.000	89.003.534	55.939.475
Trigger selection	119.805.807	867.035	336.468	2.738.145	4.438.327	1.050.841	3.702.528	16.594.784	6.485.523
$U > 200 \text{ GeV}$	69.560.780	191.522	67.594	414.892	816.164	106.698	390.411	4.653.907	866.112
$\frac{ \text{calo_MET} - U }{U} < 0,5$	28.305.602	63.536	43.346	88.832	245.985	31.171	208.463	1.339.763	199.339
$N_{\text{jets}} \geq 4$	22.327.618	58.657	31.227	79.868	221.142	25.455	151.944	1.332.262	162.354
Jato líder $p_T > 200 \text{ GeV}$	15.649.273	31.295	18.004	30.755	107.177	13.047	38.984	599.563	92.230
Veto de fótons	8.407.480	14.401	9.674	11.413	50.784	5.202	18.490	206.877	20.984
Veto de b-jets	7.329.936	12.049	8.565	9.742	42.215	4.528	15.933	32.970	18.324
Jet veto $ \eta > 2,4$	1.589.041	3.286	1.655	2.480	11.164	1.098	3.634	7.540	4.443
Single Muon	676.892	1.427	753	720	4.903	450	851	3.007	1.930
Massa transversa	447.179	960	390	426	3.269	246	435	1.899	1.242
Veto de elétrons	359.297	710	315	257	2.480	159	317	1.076	899
Veto de taus	104.835	219	98	53	685	21	105	307	159
$\Delta\phi(j_{1,2,3,4}, U) > 0,5$	104.835	219	98	53	685	21	105	307	159

Tabela 10 – Fluxo de cortes da região single muon - Parte 2

Corte	EWKWMinus	EWKWPlus	ST_tW_antitop	ST_tW_top	ST_t_top	ST_t_antitop	Total MC	Single Muon	MET	DATA/MC
Todos os eventos	2.202.000	2.033.000	840.000	2.491.000	25.457.000	12.557.000	518.638.886	323.952.013	66.747.616	0,753
Luminosity Mask	2.202.000	2.033.000	840.000	2.491.000	25.457.000	12.557.000	518.638.886	318.036.559	63.250.532	0,735
Trigger selection	526.954	508.895	152.809	453.515	2.122.523	1.076.464	160.860.618	211.419.030	27.259.274	1,483
$U > 200$	120.908	127.545	40.649	122.154	252.236	109.251	77.840.823	4.575.084	4.461.942	0,116
$\frac{ \text{calo_MET} - U }{U} < 0,5$	35.205	47.168	10.137	30.522	83.385	30.959	30.763.413	249.061	2.062.626	0,075
$N_{\text{jets}} \geq 4$	29.114	38.681	9.801	29.527	75.092	27.803	24.600.545	200.493	1.461.404	0,067
Jato líder $p_T > 200 \text{ GeV}$	17.323	23.900	4.538	13.837	40.012	15.200	16.695.138	129.210	678.981	0,048
Veto de fótons	7.837	11.117	1.672	5.161	19.891	7.669	8.798.652	73.705	371.649	0,050
Veto de b-jets	5.121	7.572	427	1.249	5.168	1.983	7.495.782	42.406	261.958	0,040
Jet veto $ \eta > 2,4$	912	1.178	96	328	961	364	1.628.180	8.400	45.679	0,033
Single Muon	418	509	34	145	432	175	692.646	4.725	14.195	0,027
Massa transversa	268	343	21	91	279	129	457.177	2.822	7.549	0,022
Veto de elétrons	203	269	16	58	212	98	366.366	2.318	5.871	0,022
Veto de taus	54	51	2	17	70	22	106.698	287	2.937	0,030
$\Delta\phi(j_{1,2,3,4}, U) > 0,5$	54	51	2	17	70	22	106.698	287	2.937	0,030

Tabela 11 – Fluxo de cortes da região single electron - Parte 1

Corte	WJetsToLNu	WZTo1L1Nu2Q	WZTo1L3Nu	WZTo3LNu	WWTo1L1Nu2Q	WWTo2L2Nu	ZZTo2L2Nu	TTJets	WGToLNuG
Todos os eventos	273.949.797	3.690.271	1.229.946	10.441.724	19.976.139	2.900.000	15.928.000	89.003.534	55.939.475
Luminosidade	273.949.797	3.690.271	1.229.946	10.441.724	19.976.139	2.900.000	15.928.000	89.003.534	55.939.475
Trigger selection	143.142.995	785.836	277.421	2.216.313	8.906.344	823.242	2.945.541	13.545.655	5.187.209
$U > 200$ GeV	89.518.570	247.978	73.194	486.032	1.072.498	122.364	382.927	4.936.168	1.130.876
$\frac{ \text{calo_MET}-U }{U} < 0,5$	28.617.456	64.293	42.033	79.563	248.451	30.119	160.069	1.148.117	201.915
$N_{\text{jets}} \geq 4$	22.576.163	59.378	30.579	71.966	223.494	24.900	121.979	1.141.891	164.716
Jato líder $p_T > 200$ GeV	15.755.914	31.682	18.048	30.979	108.216	13.147	39.088	515.507	93.175
Veto de fótons	8.419.151	14.350	9.662	11.301	50.654	5.181	18.421	173.612	20.938
Veto de b-jets	7.340.523	12.002	8.553	9.658	42.106	4.512	15.878	27.777	18.285
Jet veto $ \eta > 2,4$	1.590.694	3.274	1.651	2.455	11.142	1.094	3.623	6.360	4.431
Single Electron	499.546	1.142	514	812	3.734	389	1.004	2.318	1.489
$p_T^{\text{miss}} > 100$	498.161	1.135	514	808	3.703	387	1.002	2.297	1.483
Massa transversa	310.964	702	321	501	2.249	248	572	1.460	841
Veto de múons	213.518	446	195	156	1.471	86	290	687	516
Veto de taus	125.336	252	119	45	777	36	116	343	256
$\Delta\phi(j_{1,2,3,4}, U) > 0,5$	125.336	252	119	45	777	36	116	343	256

Tabela 12 – Fluxo de cortes da região single electron - Parte 2

Corte	EWKWMinus	EWKWPlus	ST_tW_antitop	ST_tW_top	ST_t_top	ST_t_antitop	Total MC	Single Electron	Data/MC
Todos os eventos	2.202.000	2.033.000	840.000	2.491.000	25.457.000	12.557.000	518.638.886	282.385.002	0,544
Luminosity Mask	2.202.000	2.033.000	840.000	2.491.000	25.457.000	12.557.000	518.638.886	279.642.412	0,539
Trigger selection	479.060	462.004	147.156	436.522	1.523.264	760.800	181.639.362	179.699.658	0,989
$U p_T > 200$	159.399	160.905	53.159	158.217	288.563	126.113	98.916.963	4.121.552	0,041
$\frac{ \text{calo_MET}-U }{U} < 0,5$	35.650	47.725	10.172	30.525	84.633	31.499	30.832.220	132.195	0,004
$N_{\text{jets}} \geq 4$	29.488	39.165	9.859	29.599	76.273	28.332	24.627.782	116.421	0,004
Jato líder $p_T > 200$	17.510	24.138	4.598	14.015	40.721	15.460	16.722.198	60.488	0,003
Veto de fótons	7.814	11.088	1.662	5.121	19.751	7.598	8.776.304	3.113	0,0003
Veto de b-jets	5.110	7.546	423	1.241	5.145	1.968	7.500.727	1.880	0,0002
Jet veto $ \eta > 2,4$	908	1.172	93	326	959	363	1.628.545	391	0,0002
Single Electron	298	373	29	122	300	117	512.187	268	0,0005
$p_T^{\text{miss}} > 100$	297	371	28	122	298	116	510.722	263	0,0005
Massa transversa	179	228	18	71	179	79	318.612	176	0,0005
Veto de múons	119	137	10	35	114	43	217.823	170	0,0007
Veto de taus	51	71	3	14	66	29	127.514	58	0,0004
$\Delta\phi(j_{1,2,3,4}, U) > 0,5$	51	71	3	14	66	29	127.514	58	0,0004

Tabela 13 – Fluxo de cortes da região single tau - Parte 1

Corte	WJetsToLNu	WZ1L1Nu2Q	WZ1L3Nu	WZ3LNu	WW1L1Nu2Q	WW2L2Nu	ZZ2L2Nu	WGToLNuG	TTJets
Todos os eventos	273.949.797	3.690.271	1.229.946	10.441.724	19.976.139	2.900.000	15.928.000	55.939.475	89.003.534
Luminosity Mask	273.949.797	3.690.271	1.229.946	10.441.724	19.976.139	2.900.000	15.928.000	55.939.475	89.003.534
Trigger selection	100.402.976	307.302	136.417	586.746	1.307.094	183.010	859.797	1.296.937	7.988.708
$U p_T > 200$	63.841.036	163.317	61.743	296.565	681.836	82.492	282.479	705.768	3.933.392
$\frac{ \text{calo_MET}-U }{U} < 0,5$	28.396.478	62.927	41.376	74.198	243.390	28.778	153.826	195.869	1.334.158
$N_{\text{jets}} \geq 4$	22.389.086	58.085	30.120	66.970	218.776	23.803	116.942	159.676	1.326.863
Jato líder $p_T > 200$	15.694.967	31.191	17.981	30.404	106.874	12.974	38.871	91.978	606.228
Veto de fótons	8.435.439	14.348	9.660	11.299	50.648	5.179	18.420	20.932	209.414
Veto de b-jets	7.354.928	12.001	8.551	9.656	42.102	4.510	15.877	18.280	33.476
Jet veto $ \eta > 2,4$	1.593.793	3.274	1.651	2.455	11.140	1.094	3.623	4.429	7.667
Seleção Tau isolado	920.921	1.690	998	892	6.072	520	1.388	2.822	3.735
Massa transversa	558.543	1.033	495	486	3.605	286	508	1.718	2.262
Veto de elétrons	421.064	711	371	324	2.563	181	353	1.184	1.352
Veto de múons	111.192	217	97	73	692	31	135	295	326
$\Delta\phi(j_{1,2,3,4}, U) > 0,5$	111.192	217	97	73	692	31	135	295	326

Tabela 17 – Fluxo de cortes da região double muon

Corte	DYJetsToLL_PtZ	WZTo2Q2L	ZZTo4L	ZGToLLG	EWKZ2Jets	Total MC	Single Muon	Double Muon	MET	DATA/MC
Todos os eventos	55.104.205	13.526.954	52.104.000	31.562.465	453.000	152.750.624	323.952.013	94.148.416	66.747.616	3,174
Luminosity Mask	55.104.205	13.526.954	52.104.000	31.562.465	453.000	152.750.624	319.828.002	92.825.116	63.250.532	3,115
Trigger selection	21.847.293	2.583.629	10.020.776	8.584.899	130.892	43.167.489	213.210.473	19.107.239	27.259.274	6,013
$U > 200$ GeV	17.408.037	608.493	505.208	465.248	27.484	19.014.470	4.575.084	894.326	4.461.942	0,522
$\frac{ \text{calo_MET}-U }{U} < 0,5$	1.649.929	32.493	35.304	17.691	1.392	1.736.809	249.061	57.209	2.062.626	1,374
$N_{\text{jets}} \geq 4$	1.430.675	30.883	31.553	15.214	1.235	1.509.560	221.796	53.991	1.461.404	1,150
Jato líder $p_T > 200$	1.085.163	17.205	6.790	7.440	675	1.117.273	129.210	37.906	678.981	0,757
Veto de fótons	491.958	7.211	2.361	1.701	323	503.554	73.715	20.648	371.649	0,925
Veto de b-jets	397.902	5.613	1.942	1.418	246	407.121	42.416	8.059	261.958	0,767
Jet veto $ \eta > 2,4$	96.877	1.485	493	356	43	99.254	8.400	1.650	45.679	0,561
Número de múons = 2	14.321	267	139	48	7	14.782	2.155	613	5.965	0,590
Double Muon	2.632	35	35	7	1	2.710	534	183	982	0,626
Massa invariante	640	14	23	1	1	679	79	36	97	0,312
Veto de elétrons	506	10	13	0	0	529	65	27	79	0,323
Veto de taus	27	0	0	0	0	27	4	3	17	0,888
$\Delta\phi(j_{1,2,3,4}, U) > 0,5$	27	0	0	0	0	27	4	3	17	0,888

Tabela 18 – Fluxo de cortes da região double electron

Corte	DYJetsToLL	WZTo2Q2L	ZZTo4L	ZGToLLG	EWKZ2Jets	Total MC	Single Electron	DATA/MC
Todos os eventos	55.104.205	13.526.954	52.104.000	31.562.465	453.000	152.750.624	282.385.002	1,848
Luminosidade	55.104.205	13.526.954	52.104.000	31.562.465	453.000	152.750.624	279.647.775	1,830
Trigger selection	32.806.331	2.770.413	7.455.388	6.674.844	130.187	49.837.163	180.991.183	3,631
$U > 200$ GeV	26.999.884	964.306	592.054	623.973	40.859	29.221.076	4.086.293	0,139
$\frac{ \text{calo_MET}-U }{U} < 0,5$	1.667.219	35.086	29.048	17.297	1.409	1.750.059	131.019	0,074
$N_{\text{jets}} \geq 4$	1.445.834	33.373	26.331	14.913	1.253	1.521.704	115.392	0,075
Jato líder $p_T > 200$	1.096.793	18.826	6.819	7.526	696	1.130.660	59.988	0,053
Veto de fótons	492.566	7.640	2.305	1.682	319	504.512	3.098	0,006
Veto de b-jets	398.507	5.945	1.898	1.403	243	407.996	1.890	0,004
Jet veto $ \eta > 2,4$	96.960	1.554	480	351	43	99.388	381	0,003
Número de elétrons = 2	8.249	181	65	34	2	8.531	82	0,009
Double Electron	53	1	0	0	0	54	0	0
Massa invariante	16	0	0	0	0	16	0	0
Veto de múons	7	0	0	0	0	7	0	0
Veto de taus	0	0	0	0	0	0	0	0
$\Delta\phi(j_{1,2,3,4}, U) > 0,5$	0	0	0	0	0	0	0	0

Após a aplicação dos critérios de seleção de eventos (cortes), observa-se uma redução expressiva na estatística de eventos nas etapas finais do fluxo de cortes, especialmente nas regiões com dois léptons no estado final. Essa limitação estatística está relacionada ao fato de que, para os dados reais de colisões próton-próton utilizados nesta análise, encontram-se disponíveis no CMS Open Data apenas os datasets correspondentes às eras **G** e **H** do ano de 2016. Dessa forma, não foi possível utilizar a totalidade dos dados coletados em 2016, mas apenas uma fração, o que restringe o volume de eventos analisados. Além disso, as amostras de Monte Carlo empregadas, embora representativas, não correspondem exatamente às utilizadas na publicação oficial do CMS [8], o que pode afetar a estimativa dos processos de fundo (*background*).

Tabela 19 – Seleção "baseline" e critérios para cada região.

Região	CrITÉrios de Seleção
Baseline	$U > 200 \text{ GeV}$ $ p_{T_{calo}}^{\text{miss}} - U /U < 0.5$ Lead jet $p_T > 200 \text{ GeV}$, $ \eta < 2.4$, $0.1 < \text{Ch. Had. EF} < 0.95$ Veto jets $p_T > 40 \text{ GeV}$, $ \eta > 2.4$ Veto photon $p_T > 25 \text{ GeV}$, $ \eta < 2.5$ Medium CSVV2 b -jet veto $p_T > 40 \text{ GeV}$, $ \eta < 2.4$
Jets+ p_T^{miss}	Baseline Veto muon $p_T > 10 \text{ GeV}$, $ \eta < 2.5$ Veto electron $p_T > 10 \text{ GeV}$, $ \eta < 2.5$ Veto tau $p_T > 20 \text{ GeV}$, $ \eta < 2.3$ $\min[\Delta\phi(j_{1,2,3,4}, U)] > 0.5$
Single Muon	Baseline 1 muon $p_T > 25 \text{ GeV}$, $ \eta < 2.4$ Veto electron $p_T > 10 \text{ GeV}$, $ \eta < 2.5$ Veto tau $p_T > 20 \text{ GeV}$, $ \eta < 2.3$ $30 \leq M_T(\mu, p_T^{\text{miss}}) < 125 \text{ GeV}$ $\min[\Delta\phi(j_{1,2,3,4}, U)] > 0.5$
Double Muon	Baseline 2 muons $p_T > 25 \text{ GeV}$, $ \eta < 2.4$ Veto electron $p_T > 10 \text{ GeV}$, $ \eta < 2.5$ Veto tau $p_T > 20 \text{ GeV}$, $ \eta < 2.3$ $71 < M_{\mu\mu} < 111 \text{ GeV}$ $\min[\Delta\phi(j_{1,2,3,4}, U)] > 0.5$
Single Electron	Baseline 1 electron $p_T > 30 \text{ GeV}$, $ \eta < 2.4$ Veto muon $p_T > 10 \text{ GeV}$, $ \eta < 2.5$ Veto tau $p_T > 20 \text{ GeV}$, $ \eta < 2.3$ $p_T^{\text{miss}} > 100 \text{ GeV}$ $30 \leq M_T(e, p_T^{\text{miss}}) < 125 \text{ GeV}$ $\min[\Delta\phi(j_{1,2,3,4}, U)] > 0.5$
Double Electron	Baseline 2 electrons $p_T > 30 \text{ GeV}$, $ \eta < 2.4$ Veto muon $p_T > 10 \text{ GeV}$, $ \eta < 2.5$ Veto tau $p_T > 20 \text{ GeV}$, $ \eta < 2.3$ $71 < M_{ee} < 111 \text{ GeV}$ $\min[\Delta\phi(j_{1,2,3,4}, U)] > 0.5$
Single Tau	Baseline 1 tau $p_T > 40 \text{ GeV}$, $ \eta < 2.3$ Veto muon $p_T > 10 \text{ GeV}$, $ \eta < 2.5$ Veto electron $p_T > 10 \text{ GeV}$, $ \eta < 2.5$ $\min[\Delta\phi(j_{1,2,3,4}, U)] > 0.5$
QCD Sideband	Baseline Veto muon $p_T > 10 \text{ GeV}$, $ \eta < 2.5$ Veto electron $p_T > 10 \text{ GeV}$, $ \eta < 2.5$ Veto tau $p_T > 20 \text{ GeV}$, $ \eta < 2.3$ $\min[\Delta\phi(j_{1,2,3,4}, U)] \leq 0.5$

4.3.2 Estudo de Seleção de Eventos com Machine Learning

Além da abordagem tradicional baseada em cortes (*cut-based*), foi conduzido um estudo complementar de seleção de eventos na região de sinal utilizando técnicas de aprendizado de máquina. Modelos supervisionados clássicos, como *Boosted Decision Trees* (BDTs), podem apresentar *overfitting*³ em cenários onde os eventos de interesse possuem características pouco diversificadas — como é o caso dos eventos da região de sinal, que se concentram em eventos com alto momento transversal ausente (p_T^{miss}), jato líder com alto p_T e vetos aplicados a léptons, o que restringe a variabilidade dos dados.

Neste caso, o problema é tratado como uma tarefa de classificação binária, cujo objetivo é treinar um modelo capaz de identificar os eventos que atendem aos critérios definidos para a região de sinal.

Nesse contexto, métodos de aprendizado profundo demonstram maior robustez e capacidade de generalização. Em particular, optou-se por utilizar **Redes Neurais em Grafos (Graph Neural Networks, GNNs)**, que têm se mostrado eficazes na classificação de eventos em física de partículas, especialmente na identificação de jatos e suas estruturas internas [28, 29]. Como os eventos utilizados nesta análise envolvem múltiplos jatos, as GNNs se apresentam como uma escolha apropriada.

As *Graph Neural Networks* (GNNs) são modelos de aprendizado profundo (*deep learning*) projetados especificamente para operar em dados representados por grafos, isto é, conjuntos de *nós* (*nodes*) e *arestas* (*edges*) que codificam relações entre entidades. Diferentemente das redes neurais convolucionais (CNNs), que assumem uma estrutura euclidiana dos dados, as GNNs são capazes de processar informações em domínios não euclidianos, o que as torna particularmente adequadas para representar a topologia de eventos de colisão, onde múltiplas partículas interagem em padrões que não seguem uma geometria regular [28, 29].

Em física de partículas, cada nó de um grafo pode representar um objeto reconstruído do evento (como um jato, um múon ou um fóton), enquanto as arestas conectam pares de partículas. O grafo resultante reflete a estrutura física subjacente ao evento, permitindo à GNN aprender padrões de interação entre partículas relevantes para a distinção entre sinal e fundo.

Para o treinamento do modelo, o conjunto de dados utilizado foi composto por **58.363 eventos de classe 1**, que satisfazem os critérios definidos para a região de sinal, e **486.097 eventos de classe 0**, que não atendem a esses critérios. Esse desbalanceamento significativo entre as classes pode comprometer o desempenho do modelo, favorecendo a classe majoritária durante o aprendizado.

³ Overfitting ocorre quando um modelo de aprendizado de máquina se ajusta excessivamente aos dados de treinamento, capturando ruídos e flutuações estatísticas específicas desses dados, o que compromete sua capacidade de generalização para novos dados.

Para mitigar esse problema, foi adotada uma estratégia mista de *under-sampling*⁴ da classe majoritária (classe 0) e *over-sampling*⁵ da classe minoritária (classe 1). O *over-sampling* foi realizado por meio do algoritmo SMOTE (*Synthetic Minority over-sampling Technique*), que cria exemplos sintéticos da classe minoritária ao interpolar eventos reais próximos entre si no espaço de características. Isso permite aumentar a diversidade dos dados da classe 1 sem simplesmente duplicar amostras já existentes.

Cada evento foi representado como um grafo com quatro nós, correspondentes às variáveis do jato líder: `lead_jet_pt`, `lead_jet_eta`, `lead_jet_phi` e `lead_jet_mass`. As conexões entre os nós foram definidas de maneira simples, ligando pares consecutivos e formando uma topologia linear. Essa abordagem fornece uma representação compacta do evento, suficiente para capturar suas características principais com custo computacional reduzido.

Embora seja possível construir grafos mais complexos — por exemplo, incorporando todas as partículas reconstruídas como nós — isso aumentaria substancialmente o custo computacional do modelo. Nesse contexto, custo computacional refere-se principalmente ao tempo necessário para treinar e executar o modelo, que cresce com o número de nós e conexões no grafo. Grafos maiores exigem mais operações de convolução e mais memória, resultando em tempos de processamento significativamente mais longos, o que dificulta aplicações práticas com infraestrutura computacional limitada, como execuções locais.

O modelo foi treinado por 10 épocas, utilizando métricas como acurácia, *F1-score* e matriz de confusão para avaliar o desempenho. Além disso, adotou-se um limiar de decisão ajustado, fixado em 0,6 para a probabilidade da classe 1, em vez do valor padrão de 0,5. Isso significa que um evento só seria classificado como pertencente à região de sinal se a probabilidade prevista pelo modelo para a classe 1 fosse superior a 60%. Essa escolha teve como objetivo principal reduzir a taxa de falsos positivos, ou seja, evitar que eventos de fundo fossem incorretamente classificados como sinal.

A Figura 7 apresenta a evolução da acurácia e do *F1-score* ao longo das 10 épocas de treinamento. Observa-se uma tendência geral de melhoria progressiva das métricas, culminando em um pico de desempenho na última época. Na sétima época, verifica-se uma queda abrupta nos indicadores, a qual pode ser atribuída a flutuações nos dados do conjunto de validação, um fenômeno comum quando esse conjunto possui tamanho reduzido ou inclui amostras atípicas. A rápida recuperação do desempenho nas épocas subsequentes sugere que essa oscilação não decorre de uma limitação do modelo, mas sim de uma instabilidade pontual relacionada à amostragem. Ao final do treinamento, o modelo atinge cerca de 85% tanto em acurácia quanto em *F1-score*.

⁴ Under-sampling é uma técnica que reduz a quantidade de exemplos da classe majoritária para equilibrar o número de amostras entre as classes.

⁵ *over-sampling* é uma técnica que aumenta a quantidade de exemplos da classe minoritária, gerando novas amostras ou replicando as existentes para balancear o conjunto de dados.

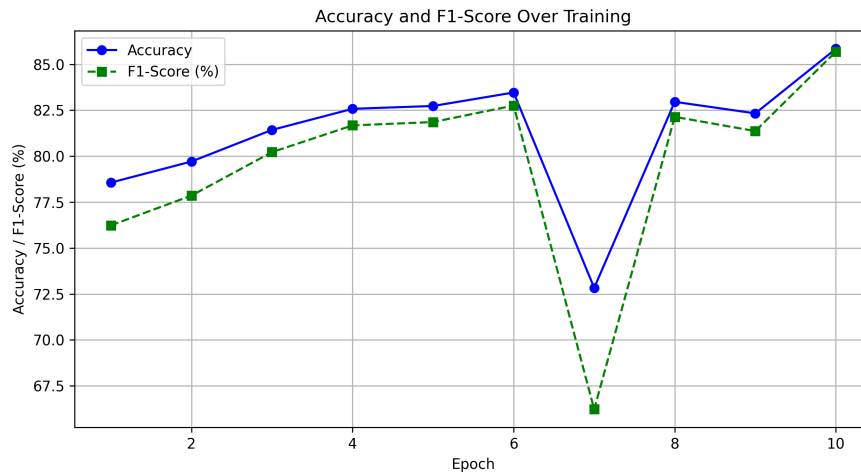


Figura 7 – Evolução da acurácia e do $F1$ -score ao longo do treinamento.

A Figura 8 complementa essa análise ao ilustrar a variação da função de perda ao longo do treinamento. Observa-se uma redução consistente da perda nas primeiras épocas, indicando que o modelo estava convergindo adequadamente. Por volta da época 7, a função de perda atinge um valor mínimo; no entanto, curiosamente, esse ponto coincide com o desempenho mais baixo em termos de acurácia e $F1$ -score. Esse comportamento sugere um possível descompasso temporário entre a minimização da perda e a capacidade do modelo de generalizar. A partir da época 9, a função de perda se estabiliza, acompanhando a recuperação e estabilização das métricas de desempenho.

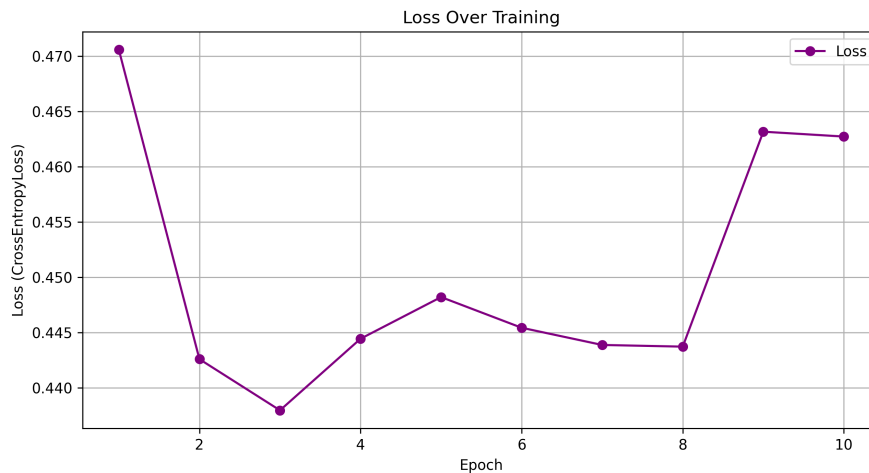


Figura 8 – Evolução da função de perda (CrossEntropyLoss) durante as épocas de treinamento.

Já a matriz de confusão final (Figura 9) mostra que o modelo foi capaz de classificar corretamente todos os eventos de classe 1 (sinal), sem gerar falsos negativos — um resultado coerente com o objetivo principal desta análise, que é identificar com eficiência os eventos

compatíveis com a região de sinal. Embora ainda existam 6.611 eventos da classe 0 (fundo) classificados incorretamente como sinal (falsos positivos), esse efeito é mitigado por uma etapa posterior: após a seleção feita pelo modelo, apenas os eventos com magnitude do vetor de recuo $u > 200$ GeV são mantidos. Esse critério adicional atua como um filtro físico, eliminando os eventos de fundo que passaram erroneamente pela GNN.

Dessa forma, o modelo cumpre seu papel com eficiência, priorizando a retenção de eventos de sinal e utilizando uma filtragem física simples para eliminar a maior parte dos falsos positivos. Essa abordagem mantém o custo computacional baixo, permitindo que todo o processo seja executado localmente, sem necessidade de recursos avançados de processamento.

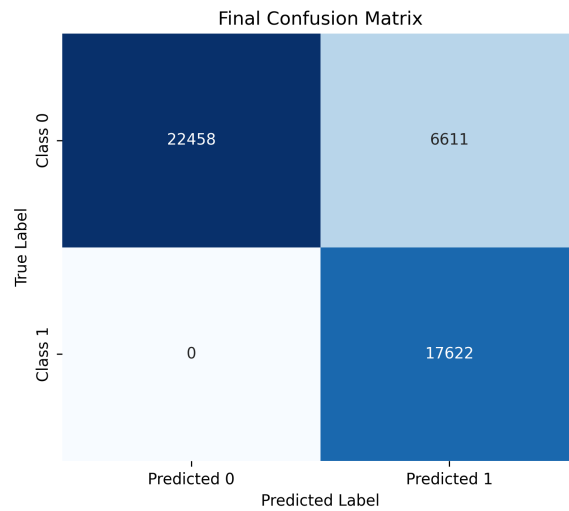


Figura 9 – Matriz de confusão final após o treinamento do modelo GNN.

Após o treinamento e validação do modelo, um novo conjunto de dados foi selecionado e submetido tanto aos cortes tradicionais da região de sinal quanto à classificação via modelo de aprendizado de máquina, com o objetivo de comparar os resultados obtidos por cada abordagem.

As Figuras 10 e 11 mostram que o modelo é capaz de selecionar eventos que passam pelos critérios de seleção — como por exemplo o corte na magnitude do vetor de recuo u e no p_T do jato líder. Além disso, observa-se que o modelo seleciona um número significativamente maior de eventos do que os cortes tradicionais.

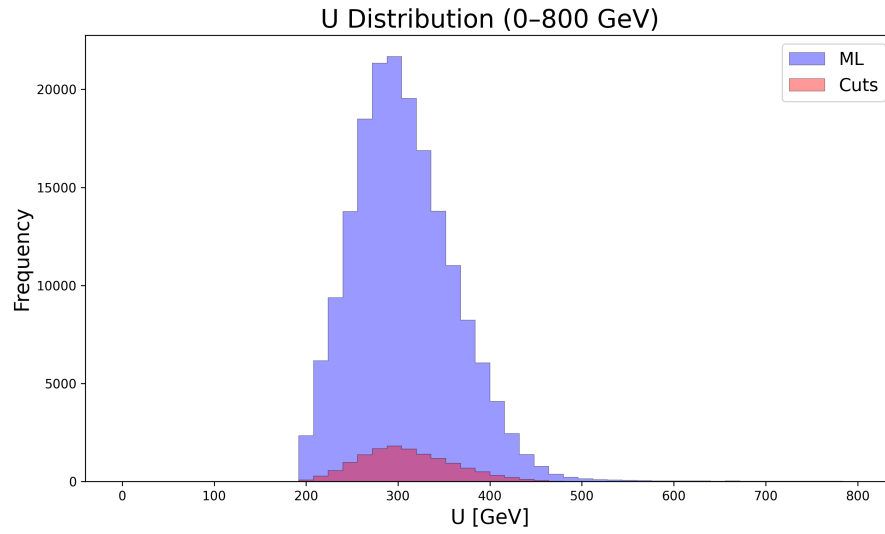


Figura 10 – Distribuição da variável u (magnitude do vetor de recuo) para eventos selecionados por aprendizado de máquina (ML) e por cortes tradicionais.

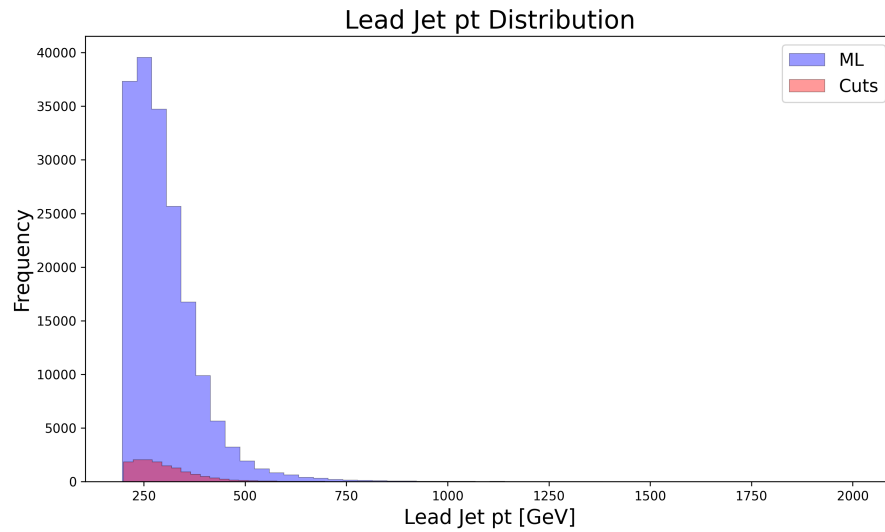


Figura 11 – Distribuição do momento transversal do jato líder (p_T) para eventos selecionados por aprendizado de máquina (ML) e por cortes tradicionais.

No entanto, ao observar as Figuras 12 e 13, percebe-se que o modelo apresenta desempenho limitado na aplicação dos vetos — como aqueles que excluem eventos com fótons em determinadas faixas de p_T e η . A presença desses eventos na seleção final indica que o modelo não aprendeu completamente os critérios de veto, o que pode explicar a quantidade significativamente maior de eventos selecionados em comparação com os cortes tradicionais.

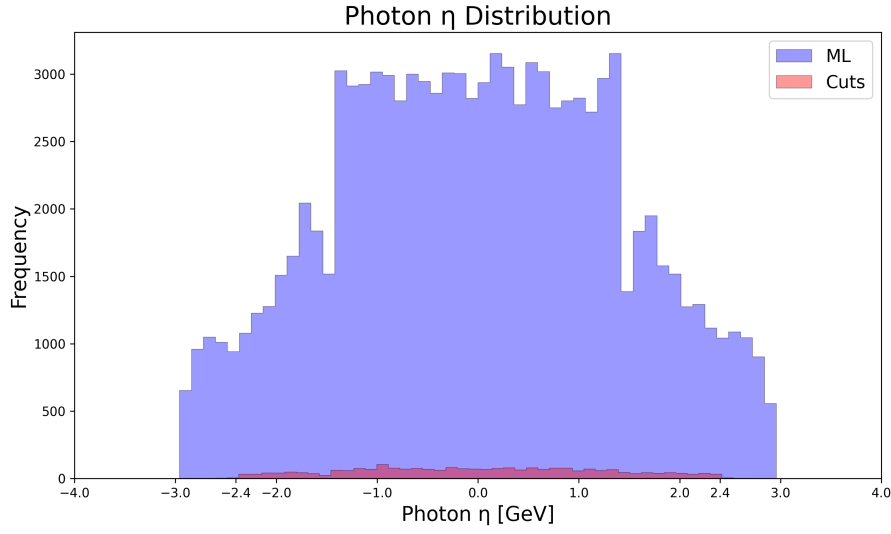


Figura 12 – Distribuição da pseudorapidez (η) dos fótons nos eventos selecionados por ML e por cortes.

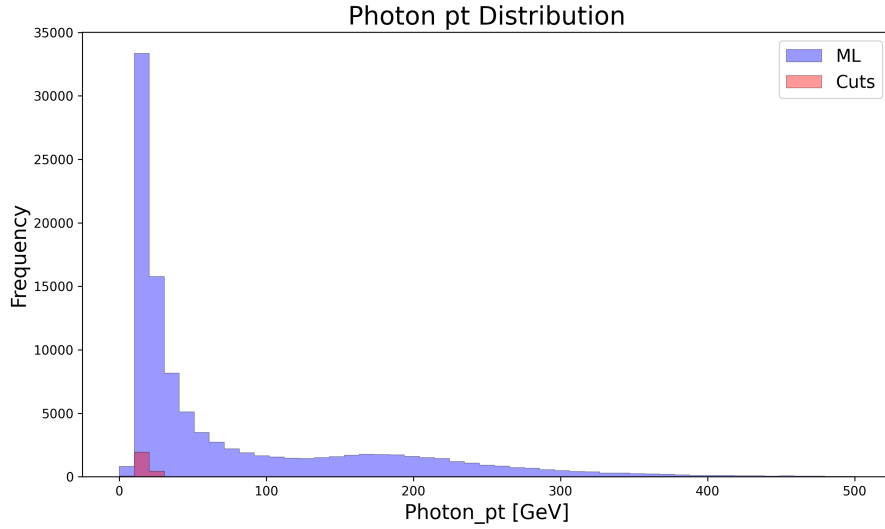


Figura 13 – Distribuição do momento transversal (p_T) dos fótons nos eventos selecionados por ML e por cortes.

Para que o modelo aprenda corretamente os vetos, duas abordagens são possíveis: (i) incluir mais exemplos no conjunto de treino que contenham tais partículas com características vetadas, e (ii) representar explicitamente essas partículas como nós em grafos, permitindo que a rede aprenda suas contribuições e aplique rejeições apropriadas.

Ainda assim, o desempenho do modelo nos cortes de seleção foi robusto, cumprindo seu objetivo central com boa eficiência e com baixo custo computacional.

4.4 Aplicação de Pesos para Normalização de Eventos

A normalização de eventos é uma etapa importante ao utilizar amostras geradas por simulação de Monte Carlo, pois assegura que as distribuições simuladas representem o comportamento observado nos dados experimentais. Essa normalização é realizada por meio da aplicação de pesos aos eventos. Conforme estabelecido nas diretrizes do CMS Open Data [30], os pesos devem ser calculados segundo a seguinte expressão:

$$w = \frac{\sigma \times \mathcal{L}}{n_{\text{gen}}}, \quad (4.1)$$

onde n_{gen} representa o número efetivo de eventos gerados, definido como $n_{\text{gen}} = \text{gw_pos} - \text{gw_neg}$, com gw_pos sendo a soma dos pesos positivos ($\sum(\text{gen_weights} > 0)$) e gw_neg a soma dos pesos negativos ($\sum(\text{gen_weights} < 0)$).

Em algumas amostras de Monte Carlo, os eventos podem ter pesos positivos ou negativos devido a técnicas de reponderação utilizadas para melhorar a precisão da simulação. Isso ocorre, por exemplo, em métodos como *unweighting* de eventos[31], que busca gerar eventos de peso unitário a partir de distribuições com pesos variáveis, aceitando ou rejeitando eventos com base em um critério probabilístico. Além disso, eventos com pesos negativos podem surgir em cálculos de ordem superior em QCD (como NLO e NNLO⁶), onde correções virtuais e termos de interferência podem resultar na necessidade de atribuir pesos negativos para garantir a consistência teórica da simulação[33].

Para os dados de colisão, o peso aplicado é igual a 1.

4.5 Correção da Contribuição de γ^* na Amostra DYJetsToLL

Nesta análise, o conjunto de dados DYJetsToLL é empregado para modelar o processo de produção do bóson Z , seguido de seu decaimento em léptons carregados. No entanto, esse dataset também inclui contribuições do fóton virtual (γ) e da interferência entre Z e γ . Portanto, torna-se necessário aplicar uma correção para isolar a contribuição do processo $Z \rightarrow \ell^+ \ell^-$.

Para esse fim, foram utilizadas as seções de choque dos processos $Z \rightarrow \ell\ell$ e $\gamma^* \rightarrow \ell\ell$, a partir das quais foram calculados os fatores de correção k_Z e k_γ , conforme apresentado na Tabela 21.

⁶ Next-to-Leading Order e Next-to-Next-to-Leading Order referem-se a cálculos de ordem superior na expansão perturbativa da QCD. NLO inclui correções de primeira ordem em α_s , como diagramas com emissão de glúons e correções virtuais. NNLO vai além, incorporando termos de segunda ordem e múltiplas contribuições de loop, aumentando a precisão dos cálculos de seções de choque e distribuições observáveis [32].

Processos	Seção de Choque (pb)	Razão em relação ao DY
DY(l)	5316	1
Z(l)	5203	0.9787 (K_z)
γ^* (l)	79	0.0149 (K_γ)

Tabela 20 – Seção de choque e razões de vários processos em relação ao DY.

Com esses fatores, as seções de choque corrigidas para Z e γ podem ser obtidas e somadas para determinar a seção de choque corrigida do DYToLL:

$$\sigma_{Z\text{corrected}} = \sigma_Z \times k_Z = 5092.2 \text{ pb}, \quad (4.2)$$

$$\sigma_{\gamma^*\text{corrected}} = \sigma_{\gamma^*} \times k_{\gamma^*} = 1.2 \text{ pb}, \quad (4.3)$$

$$\sigma_{\text{DY corrected}} = \sigma_{Z\text{ corrected}} + \sigma_{\gamma^*\text{ corrected}} = 5093.4 \text{ pb}, \quad (4.4)$$

Então é possível obter a interferência e a porcentagem de interferência no processo DYToLL:

$$\text{Interferência} = \sigma_{\text{DY}} - (\sigma_Z + \sigma_{\gamma^*}) = 34, \quad (4.5)$$

$$\frac{\text{Interferência}}{\sigma_{\text{DY}}} \times 100 = \frac{34}{5316} \times 100 = 0,64\%, \quad (4.6)$$

Por fim, é possível calcular o peso a ser aplicado no conjunto de dados DYToLL para incluir a contribuição de γ :

$$\text{Interferência} = \sigma_{\text{DY total corrigida}} \times 0,0064 = 32,6 \text{ pb}, \quad (4.7)$$

$$\sigma_{\text{DY final}} = \sigma_{\text{DY total corrigida}} - 32,6 = 5060,8 \text{ pb}, \quad (4.8)$$

$$w = \frac{\sigma_{\text{DY final}}}{\sigma_{\text{DY simulada}}} = \frac{5060,8}{5316} = 0,9519. \quad (4.9)$$

Aplicando o peso $w = 0,9519$ aos eventos do conjunto de dados DYJetsToLL, garante-se que apenas a contribuição do processo $Z \rightarrow \ell^+\ell^-$ esteja sendo considerada na análise, com as contribuições do fóton virtual (γ^*) e da interferência devidamente corrigidas.

4.6 Correção de Pileup

O termo *pileup* refere-se às múltiplas interações próton-próton que ocorrem em um mesmo cruzamento de feixes. A correção de *pileup* é aplicada para ajustar eventos gerados por simulação de Monte Carlo (MC) às condições reais de colisão observadas nos dados do experimento CMS. Isso é feito por meio de uma técnica de *reweighting*, na qual cada evento simulado recebe um peso que depende do número de interações de *pileup* associadas a ele.

A distribuição real de *pileup* é obtida conforme as instruções da colaboração CMS [34], utilizando a ferramenta `pileupCalc.py`, em conjunto com informações de luminosidade registradas e o valor recomendado da seção de choque inelástica para os dados de 2016 [34]. O resultado é um histograma no formato ROOT contendo a distribuição normalizada do número de interações de *pileup* por evento nos dados reais. Essa distribuição é então comparada à distribuição simulada de *pileup* nos eventos MC. Para cada valor de *pileup* n , calcula-se um peso de correção como a razão entre as distribuições dos dados e da simulação:

$$w(n) = \frac{P_{\text{data}}(n)}{P_{\text{MC}}(n)}, \quad (4.10)$$

onde $P_{\text{data}}(n)$ representa a probabilidade normalizada de ocorrência de n interações nos dados reais, e $P_{\text{MC}}(n)$ a mesma quantidade na simulação. Esses pesos são aplicados evento a evento, de forma que a distribuição final de *pileup* no MC reproduza a observada nos dados. Isso garante que variáveis sensíveis ao *pileup*, como o número de vértices primários e a energia depositada em diferentes regiões do detector, sejam corretamente modeladas.

A Figura 14 ilustra esse procedimento. A área sombreada em cinza representa a distribuição de *pileup* obtida a partir dos dados reais. A linha tracejada vermelha mostra a distribuição original de *pileup* da simulação antes da correção, e a linha azul representa a distribuição após o *reweighting*. Observa-se que, após a aplicação dos pesos, a distribuição simulada se ajusta à distribuição observada nos dados.

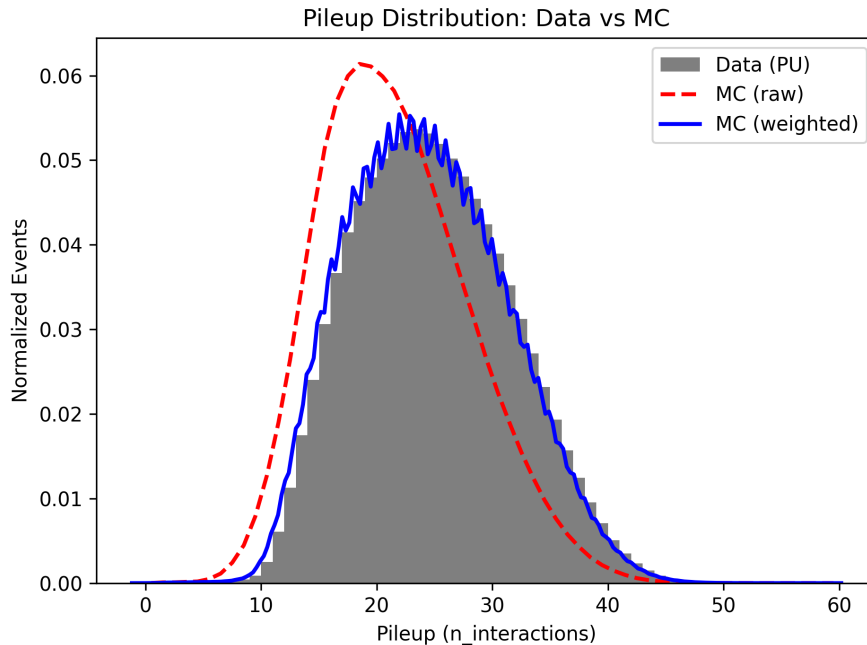


Figura 14 – Distribuições normalizadas de *pileup* nos dados reais, na simulação antes da correção de *pileup* e na simulação corrigida por *reweighting*.

4.7 Ajuste Simultâneo e Extração da Largura Invisível do Bóson Z

As etapas de pré-processamento — incluindo a normalização de eventos, a correção da contribuição de fótons virtuais (γ^*) na amostra DYJetsToLL e a correção de *pileup* — foram aplicadas aos eventos selecionados nas três regiões de análise definidas neste estudo: a **região de sinal** ($p_T^{\text{miss}} + \text{jets}$), utilizada para estimar o processo $Z \rightarrow \nu\bar{\nu}$; a **região single lepton**, composta por eventos com exatamente um múon, elétron ou tau, utilizada para modelar o principal *background* $W \rightarrow \ell\nu$; e a **região double lepton**, formada por eventos com dois múons ou dois elétrons, utilizada para estimar o canal de referência $Z \rightarrow \ell^+\ell^-$.

A Figura 15 apresenta a distribuição do vetor de recuo U , comparando os dados reais de colisão com as previsões do Monte Carlo. O painel da esquerda mostra a distribuição correspondente ao processo $Z \rightarrow \nu\bar{\nu}$, referenciada como $p_T^{\text{miss}} + \text{jets}$, que inclui eventos da região de sinal e das regiões single lepton. Já o painel da direita exibe a distribuição para o canal de referência $Z \rightarrow \ell^+\ell^-$, com eventos provenientes da região double lepton — denominada $\mu^+\mu^- + \text{jets}$ —, uma vez que nenhum evento passou nos critérios de seleção de dois elétrons.

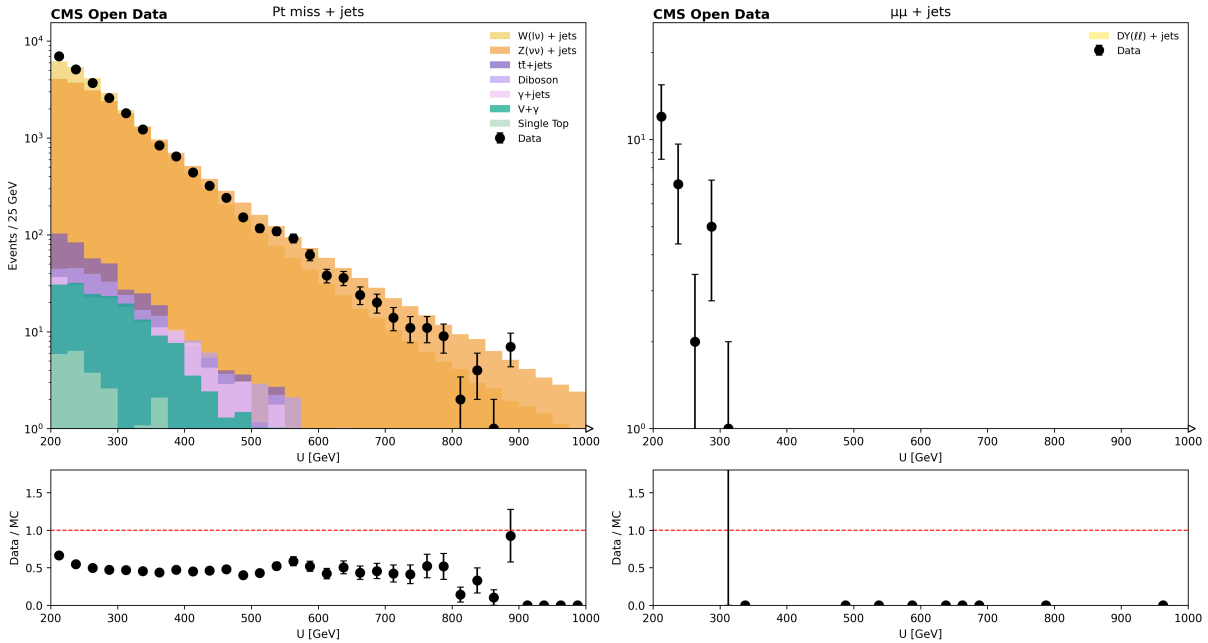


Figura 15 – Distribuições do vetor de recuo U .

Os painéis inferiores exibem a razão Data/MC avaliada bin a bin, permitindo examinar diretamente a compatibilidade entre os dados e as previsões das amostras geradas por simulação monte carlo em cada intervalo de U .

Na região $p_T^{\text{miss}} + \text{jets}$, observa-se que a razão entre dados e Monte Carlo permanece consistentemente abaixo de 1 em grande parte do espectro. Isso indica que a simulação prevê um número maior de eventos do que o efetivamente observado nos dados experimentais.

Já na região $\mu^+\mu^- + \text{jets}$, a razão apresenta flutuações significativas, atribuídas à baixa estatística de eventos. Essa limitação estatística decorre do fato de que o portal Open Data do CERN disponibiliza apenas uma fração dos dados de colisão próton-próton do ano de 2016, correspondendo exclusivamente às eras G e H do período de tomada de dados.

A distribuição do vetor de recuo U , apresentada na Figura 15, constitui a base observável sobre a qual é aplicado o ajuste simultâneo para a extração da largura invisível do bóson Z. Esse ajuste é realizado por meio da maximização de uma função de verossimilhança que leva em conta, de forma simultânea, os eventos observados nas três regiões experimentais — sinal, single lepton e double lepton — com os diferentes processos físicos que contribuem em cada uma delas.

A função de verossimilhança utilizada está expressa por:

$$\begin{aligned} \mathcal{L}(n_j, n_\ell, n_{\ell\ell} \mid r, r_Z, r_W, \theta) = & \text{Poisson}(n_j \mid r \cdot r_Z \cdot s_{Z,j}(\theta) + r_W \cdot b_{j,W}(\theta) + b_{\text{bkg},j}(\theta)) \\ & \cdot \text{Poisson}(n_\ell \mid r_W \cdot b_{\ell,W}(\theta) + b_{\text{bkg},\ell}(\theta)) \\ & \cdot \text{Poisson}(n_{\ell\ell} \mid r_Z \cdot s_{Z,\ell\ell}(\theta) + \sqrt{r_Z} \cdot s_{\text{int},\ell\ell} + s_{\gamma^*,\ell\ell}(\theta) + b_{\text{bkg},\ell\ell}(\theta)) \\ & \cdot p(\tilde{\theta}, \theta) \end{aligned} \quad (4.11)$$

onde os termos correspondem a:

- $s_{Z,j}(\theta)$: taxa esperada do processo $Z \rightarrow \nu\bar{\nu}$ na região de sinal.
- $b_{j,W}(\theta)$: taxa esperada do processo $W + \text{jets}$ na região de sinal.
- $b_{\ell,W}(\theta)$: taxa esperada do processo $W + \text{jets}$ na região single lepton.
- $s_{Z,\ell\ell}(\theta)$: taxa esperada do processo $Z \rightarrow \ell^+\ell^-$ na região double lepton.
- $s_{\text{int},\ell\ell}$: termo de interferência entre Z e γ^* na região double lepton.
- $s_{\gamma^*,\ell\ell}(\theta)$: contribuição do processo γ^* na região double lepton.
- $b_{\text{bkg},j}(\theta)$, $b_{\text{bkg},\ell}(\theta)$, $b_{\text{bkg},\ell\ell}(\theta)$: taxas esperadas de outros *backgrounds* nas regiões de sinal, single lepton e double lepton, respectivamente.
- $r_Z = \frac{N_{\text{MC}}^{\text{sinal}}(Z+\text{jets})}{N_{\text{MC}}^{\text{controle}}(Z+\text{jets})}$: fator de transferência para o processo $Z + \text{jets}$.
- $r_W = \frac{N_{\text{MC}}^{\text{sinal}}(W+\text{jets})}{N_{\text{MC}}^{\text{controle}}(W+\text{jets})}$: fator de transferência para o processo $W + \text{jets}$.
- $r = \frac{\Gamma(Z \rightarrow \nu\bar{\nu})}{\Gamma_{\text{SM}}(Z \rightarrow \nu\bar{\nu})}$: fator de escala da largura invisível do Z em relação à predição do Modelo Padrão.

A partir da razão entre os eventos simulados nas regiões de sinal e de controle, foram obtidos os seguintes fatores de transferência: $r_Z = (1,09 \pm 2,55) \times 10^5$ para o processo $Z + \text{jets}$, e $r_W = 1,4050 \pm 0,0180$ para o processo $W + \text{jets}$.

O valor obtido para o fator r_Z é anormalmente alto e apresenta uma incerteza relativa superior a 200%, isso se deve à quantidade reduzida de eventos da amostra DYJetsToLL que passam nos critérios de seleção da região de controle double lepton, comprometendo a confiabilidade da estimativa direta de r_Z .

Diante disso, é apropriado tratar r_Z como um parâmetro livre⁷ no ajuste simultâneo, permitindo que seu valor seja determinado diretamente pelos dados. Essa escolha evita a propagação de uma estimativa instável e estatisticamente pouco confiável.

O ajuste simultâneo entre as regiões da análise foi realizado por meio da técnica de amostragem por cadeias de Markov Monte Carlo (MCMC). Essa abordagem permite explorar numericamente a distribuição a posteriori dos parâmetros do modelo, a partir da função de verossimilhança construída com os dados observados e dos priors definidos para cada parâmetro.

O conjunto de parâmetros ajustados inclui o fator de escala r , os fatores de transferência r_Z e o parâmetro de incerteza sistemática global θ . O fator r_W foi mantido fixo, com valor obtido previamente.

Para incorporar incertezas experimentais e restrições teóricas no ajuste simultâneo, foram aplicadas penalizações gaussianas (priors) sobre determinados parâmetros livres do modelo:

- **Parâmetro global θ :** modela a incerteza sistemática global comum às regiões da análise. Foi aplicado um prior gaussiano com $\sigma_\theta = 0,032$, valor consistente com a incerteza sistemática combinada de 3,2%, detalhada na Tabela 21.
- **Fator de escala r :** relacionado diretamente à largura invisível do bóson Z. Um prior informativo centrado em $r = 1,0$ com $\sigma_r = 0,1$ foi utilizado, refletindo a expectativa do Modelo Padrão.
- **Fator de transferência r_Z :** aplicado ao processo $Z + \text{jets}$, com prior centrado em $r_Z = 0,7$ e $\sigma_{r_Z} = 0,2$.

⁷ Um parâmetro livre é um valor numérico cujo melhor estimador é obtido diretamente a partir do ajuste aos dados experimentais, sem ser fixado previamente nem calculado a partir de outras quantidades.

Tabela 21 – Fontes de incerteza sistemática consideradas na análise. Fonte: [8]

Fonte	Incerteza (%)
Eficiência de identificação de múons (sist.)	2,1
Escala de energia de jatos	1,8–1,9
Eficiência de identificação de elétrons (sist.)	1,6
Eficiência de identificação de elétrons (stat.)	1,0
Pileup	0,9–1,0
Eficiência de disparo (elétron)	0,7
Eficiência de veto ao tau	0,6–0,7
Eficiência de disparo em p_T^{miss} (jets + MET)	0,7
Eficiência de disparo p_T^{miss} ($Z/\gamma^* \rightarrow \mu\mu$)	0,6
Correção QCD dependente do p_T do bóson	0,5
Resolução de energia dos jatos	0,3–0,5
Eficiência de disparo (muon + jets)	0,4
Eficiência de identificação de múons (stat.)	0,3
Eficiência de reconstrução de elétrons (sist.)	0,3
Correções EWK dependentes do p_T do bóson	0,3
Funções de distribuição de partons (PDFs)	0,2
Escala de renormalização/fatorização	0,2
Eficiência de reconstrução de elétrons (stat.)	0,2
Total (combinado)	3,2

A Figura 16 apresenta o diagrama de correlação entre os parâmetros ajustados via MCMC. As distribuições marginais mostram a densidade a posteriori de cada parâmetro individualmente, enquanto os painéis 2D revelam as correlações entre pares de parâmetros.

Os parâmetros ajustados e suas respectivas incertezas, estimados a partir da distribuição a posteriori obtida via MCMC, são:

- $r = 1,272^{+0,076}_{-0,073}$
- $r_Z = 1,469^{+0,093}_{-0,088}$
- $\theta = -0,6723^{+0,0052}_{-0,0051}$

Os resultados indicam que o modelo ajustou os parâmetros de forma a compensar o valor elevado de r_Z , resultado da baixa estatística na região de controle double lepton. Para manter a consistência com os dados observados, o ajuste simultâneo aumentou os valores de $r = 1,272^{+0,076}_{-0,073}$ e $\theta = -0,6723^{+0,0052}_{-0,0051}$. Esse comportamento reflete a correlação entre os parâmetros: ao limitar o crescimento de $r_Z = 1,469^{+0,093}_{-0,088}$, o modelo redistribui o ajuste entre r e θ para melhor se adaptar às distribuições observadas.

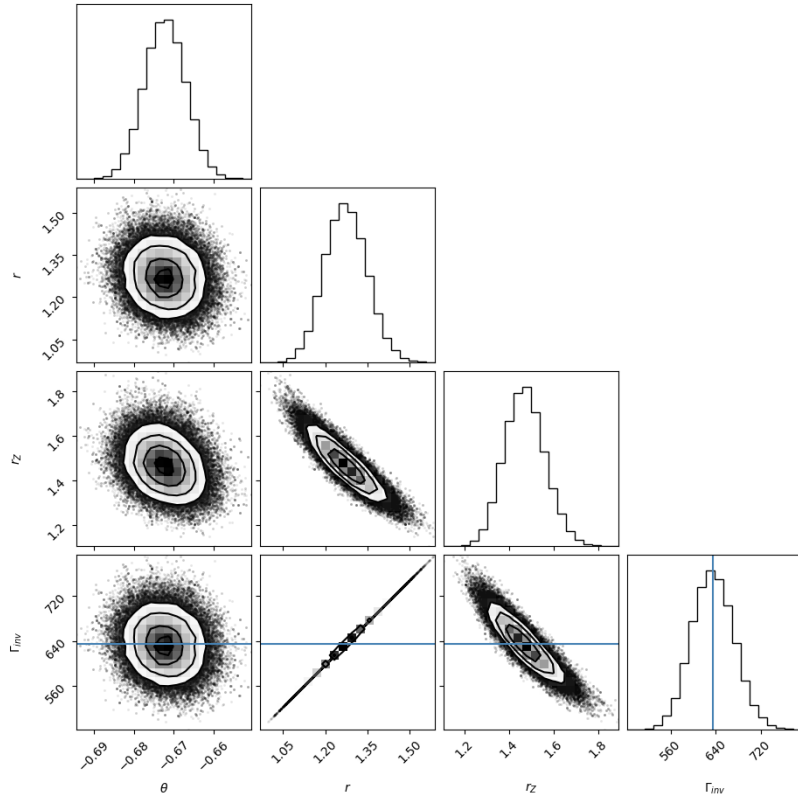


Figura 16 – Distribuições marginais e correlações entre os parâmetros ajustados: θ (sistemático global), r (fator de escala da largura invisível do Z), r_Z (fator de transferência para $Z + \text{jets}$) e Γ_{inv} (largura invisível).

Com o valor ajustado de r , é possível obter a largura invisível do bóson Z utilizando a predição teórica do Modelo Padrão para esse canal: $\Gamma_{\text{SM}}(Z \rightarrow \nu\bar{\nu}) = 499,0 \text{ MeV}$. A relação utilizada é:

$$\Gamma(Z \rightarrow \nu\bar{\nu}) = r \cdot \Gamma_{\text{SM}}(Z \rightarrow \nu\bar{\nu}). \quad (4.12)$$

Substituindo os valores, obtém-se:

$$\Gamma(Z \rightarrow \nu\bar{\nu}) = 1,272 \times 499,0 = 634,6 \text{ MeV}. \quad (4.13)$$

Com a propagação das incertezas estatísticas do parâmetro r , o resultado final da largura invisível é:

$$\Gamma(Z \rightarrow \nu\bar{\nu}) = 634,6^{+37,9}_{-36,5} \text{ MeV}. \quad (4.14)$$

Além da incerteza estatística proveniente do ajuste de r , é necessário considerar duas fontes adicionais de incerteza que impactam a estimativa final da largura invisível:

- A incerteza no fator de transferência $r_W = 1,4050 \pm 0,0180$;
- A incerteza sistemática global, representada por θ , com desvio padrão $\sigma_\theta = 0,032$.

A contribuição absoluta dessas incertezas sobre $\Gamma(Z \rightarrow \nu\bar{\nu})$ foi propagada da seguinte forma:

- Erro devido a r_W : $\Delta\Gamma_{r_W} = \Gamma_{\text{SM}} \cdot \langle r \rangle \cdot \sigma_{r_W} = 499,0 \cdot 1,272 \cdot 0,018 \approx 11,4 \text{ MeV}$;
- Erro sistemático global: $\Delta\Gamma_{\text{sist}} = \langle \Gamma \rangle \cdot \sigma_\theta = 634,6 \cdot 0,032 \approx 20,3 \text{ MeV}$.

As incertezas total superior e inferior foram então calculadas pela combinação quadrática das componentes estatísticas e sistemáticas:

$$\Delta\Gamma_{\text{total}}^+ = \sqrt{37,9^2 + 11,4^2 + 20,3^2} \approx 44,5 \text{ MeV}, \quad (4.15)$$

$$\Delta\Gamma_{\text{total}}^- = \sqrt{36,5^2 + 11,4^2 + 20,3^2} \approx 43,3 \text{ MeV}. \quad (4.16)$$

Com isso, o valor final da largura invisível do bóson Z, incorporando todas as fontes de incerteza consideradas na análise, é:

$$\Gamma(Z \rightarrow \nu\bar{\nu}) = 634,6^{+44,5}_{-43,3} \text{ MeV}. \quad (4.17)$$

De forma aproximada, esse resultado pode ser expresso como:

$$\Gamma(Z \rightarrow \nu\bar{\nu}) \approx 635 \pm 44 \text{ MeV}, \quad (4.18)$$

A Figura 17 apresenta uma comparação direta entre o resultado obtido nesta análise e outras medições experimentais da largura invisível do bóson Z. A linha vertical tracejada indica o valor previsto pelo Modelo Padrão, $\Gamma_{\text{SM}}(Z \rightarrow \nu\bar{\nu}) = 499,0 \text{ MeV}$.

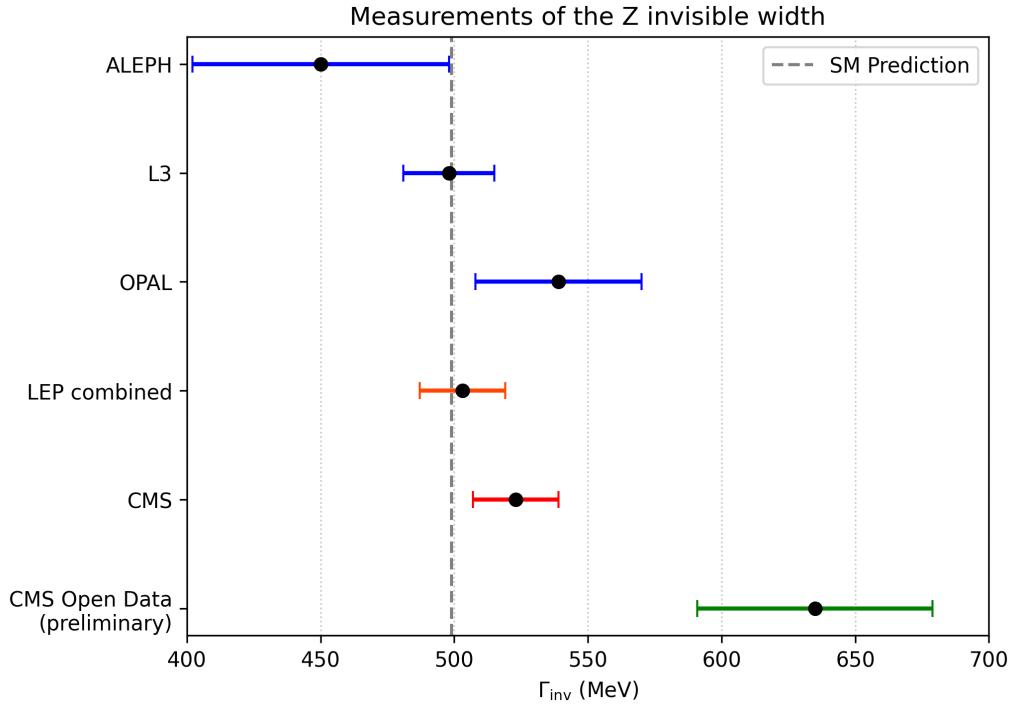


Figura 17 – Comparação entre diferentes medições experimentais da largura invisível do bóson Z. A linha tracejada representa a predição do Modelo Padrão. O ponto inferior (em verde) corresponde ao resultado obtido nesta análise com dados preliminares do CMS Open Data.

Enquanto o resultado dos experimentos do LEP (ALEPH, L3, OPAL) combinados mediram valores compatíveis com o Modelo Padrão dentro de incertezas relativamente pequenas, a análise da colaboração CMS[8] já apresenta um desvio modesto para cima. O resultado obtido nesta análise — $\Gamma(Z \rightarrow \nu\bar{\nu}) = 634,6^{+44,5}_{-43,3}$ MeV — é consideravelmente mais alto, com um valor central cerca de 27% acima da predição do SM.

Essa diferença pode ser explicada, em grande parte, pela limitação estatística das amostras disponíveis no portal open data do CERN, que abrange apenas as eras G e H do ano de 2016, além das incertezas associadas ao fator de transferência r_Z , amplificadas pela baixa estatística da simulação monte carlo na região de controle. O valor obtido nesta análise, $\Gamma(Z \rightarrow \nu\bar{\nu}) = 634,6^{+44,5}_{-43,3}$ MeV, é significativamente maior do que a estimativa mais precisa da colaboração CMS, $\Gamma = 523 \pm 16$ MeV.

Para quantificar essa diferença, calculamos a separação estatística entre os dois resultados utilizando a seguinte equação:

$$\Delta = \frac{|\Gamma_{\text{CMS Open Data}} - \Gamma_{\text{CMS}}|}{\sqrt{\sigma_{\text{Open Data}}^2 + \sigma_{\text{CMS}}^2}} = \frac{|634,6 - 523|}{\sqrt{44^2 + 16^2}} = \frac{111,6}{46,82} \approx 2,38 \sigma. \quad (4.19)$$

A diferença entre os dois resultados corresponde, portanto a aproximadamente $2,4\sigma$, embora significativo, o desvio pode ser atribuído às incertezas decorrentes da baixa estatística disponível nesta análise.

Os iRPCs introduzem avanços significativos em termos de rejeição de ruído, capacidade de detecção em altas taxas e resolução espacial aprimorada. Utilizando painéis de HPL (*High Pressure Laminate*) e lacunas de gás reduzidas a 1,4 mm, os iRPCs alcançam resoluções de 0,3 – 0,6 cm na direção ϕ e 1,5 cm na direção η [11]. Essas melhorias permitem que os iRPCs operem eficientemente no ambiente de alta luminosidade do HL-LHC, com uma resolução temporal aprimorada de até 500 ps, crucial para reconstruir trajetórias de múons com alta precisão e distinguir eventos próximos no tempo.

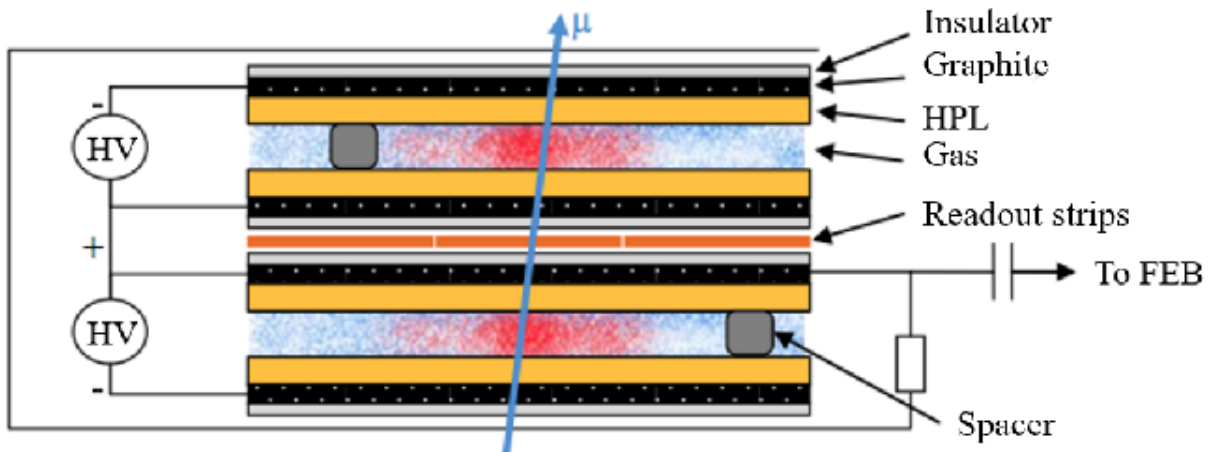


Figura 19 – Esquema detalhado de um RPC, destacando os principais componentes: placas resistivas de HPL, tiras de leitura, gás, grafite, isolante e espaçadores. Fonte: [7].

A montagem dos novos detectores está sendo realizada no laboratório CERN 904 em Preessin e na Universidade de Ghent. Para garantir o desempenho adequado, as câmaras iRPC passam por uma série de testes de controle de qualidade (QC, sigla em inglês para Quality Control) em cada etapa da cadeia de montagem, conforme descrito a seguir:

- **QC1: Verificação dos componentes básicos das câmaras**
 - HPL: Instituto Nacional de Física Nuclear (INFN)
 - PCB das Tiras (Printed Circuit Board das Strip Detectors): Lyon
 - Placa de Circuito Front-End (FEB): Lyon
 - Sistema de Refrigeração: Universidade da Geórgia
- **QC2: Testes de elementos da câmara**
 - Testes de lacunas no Laboratório de Detectores da Coreia:
 - * Testes de vazamento de gás
 - * Verificação de fixação dos espaçadores

- * Scans de corrente escura
- Nos locais de montagem:
 - * Repetição dos testes mencionados acima
- **QC3: Avaliação completa do desempenho da câmara após a produção**
 - QC3.1: Testes dos elementos da câmara após a montagem
 - QC3.2: Teste durante tomada de dados com raios cósmicos utilizando uma FEB (Front-End Board) portátil
- **QC4: Teste final de validação**
 - QC4.1: Testes finais da câmara
 - * Teste de vazamento do sistema de refrigeração
 - * Teste de vazamento de gás
 - * Scans de corrente escura
 - QC4.2: Estabilidade de alta tensão (HV) em longo prazo
 - QC4.3: Teste cósmico da câmara com FEBs finais

Durante a bolsa de mestrado sanduíche realizada no CERN, atuei diretamente na montagem das câmaras e na execução da etapa QC4.3, que abrange a calibração das placas Front-End Board (Febs) finais e testes de eficiência das iRPCs. Nesses testes, as câmaras são posicionadas entre dois detectores cintiladores, que funcionam como sistema de referência para identificar a passagem de múons cósmicos. Esse arranjo permite determinar com precisão se o múon foi detectado pela câmara RPC, possibilitando a avaliação da eficiência do detector em diferentes condições de operação.

O primeiro gráfico (Figura 20) exhibe uma reconstrução tridimensional das posições de impacto de múons em uma câmara iRPC, destacando a uniformidade da resposta do detector em toda a sua área ativa. Já o segundo gráfico (Figura 21) apresenta a relação entre a eficiência de detecção de múons cósmicos e a alta tensão efetiva (HV_{eff}), evidenciando que a eficiência atinge um platô de 99% em altas tensões.

A curva de eficiência é obtida a partir da razão entre o número de eventos detectados pela câmara e o número de eventos registrados pelo sistema de disparo (*trigger*), permitindo assim corrigir efeitos de ruído e falsas detecções. Para cenários em que o ruído pode impactar significativamente a medição, é adotada uma abordagem bayesiana que fornece estimativas mais robustas da eficiência real. Os dados obtidos para diferentes valores de tensão são ajustados por uma curva sigmoide, que representa o comportamento esperado de detectores gasosos operando no modo avalanche [36].

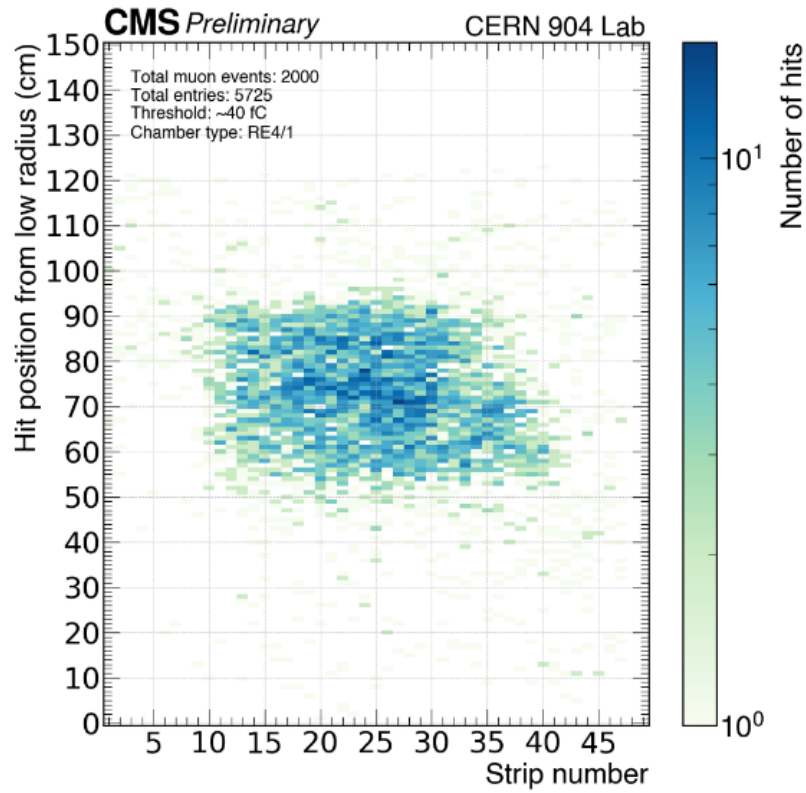


Figura 20 – O gráfico mostra uma reconstrução 3D das posições de impacto de múons em uma câmara iRPC em um ponto de operação com um limiar de carga de 40 fC. Os dados foram coletados usando um *trigger* de coincidência tripla de cintiladores com área de 30x40 cm².

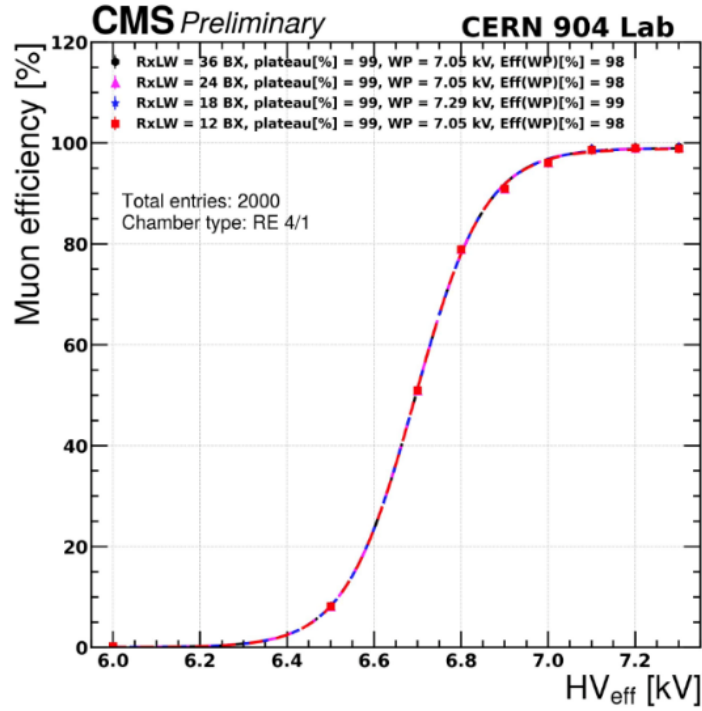


Figura 21 – O gráfico apresenta a eficiência de detecção de múons cósmicos em função da alta tensão efetiva (HV_{eff}). A eficiência atinge um pico de 99%, com o ponto de operação definido como a tensão em que a eficiência é 95% (HV_{knee}) mais 150V.

4.8.1 Estudo de Gases para os iRPCs

O estudo de misturas de gases para os iRPCs do CMS combina a necessidade de atender às demandas técnicas do HL-LHC com a responsabilidade ambiental. Gases tradicionais, como o $C_2H_2F_4$ e SF_6 , apresentam elevado potencial de aquecimento global (GWP, sigla em inglês para Global Warming Potential), contribuindo para as emissões de gases de efeito estufa. Para mitigar esse impacto, o CMS busca alternativas sustentáveis, alinhadas às metas do CERN de reduzir emissões de gases com elevado GWP em 28% até o final do Run III e alcançar zero emissões até 2050, conforme as metas da União Europeia.

Estão sendo realizados testes no *Gamma Irradiation Facility* (GIF++) do CERN com misturas alternativas de gases, incluindo composições com 30-40% de CO_2 e concentrações reduzidas de SF_6 combinadas com $C_2H_2F_4$, conforme apresentado na Tabela 1.

Tabela 22 – Comparação de misturas de gases utilizadas no sistema de iRPCs do CMS e seus Potenciais de Aquecimento Global (GWP).

Componente	Mistura Pa- drão	30% CO ₂ + 1% SF ₆	30% CO ₂ + 0.5% SF ₆	40% CO ₂ + 1% SF ₆
Freon (C ₂ H ₂ F ₄)	95.2%	64%	64.5%	54%
Isobutano (iC ₄ H ₁₀)	4.5%	5%	5%	5%
CO ₂	0%	30%	30%	40%
SF ₆	0.3%	1%	0.5%	1%
GWP	1486	1529	1337	1353

O GIF++ utiliza uma fonte de césio-137 combinada com um feixe de múons de alta energia (100-150 GeV/c) para simular as condições extremas previstas no HL-LHC. Os testes são conduzidos tanto com a fonte de radiação gama ativada quanto desativada, empregando filtros ABS com valores de 1, 2.2, 3.3, 4.6, 6.9 e 10. Cada valor de filtro corresponde a uma taxa específica de radiação, que varia conforme a distância da câmara em relação à fonte.

A relação entre o valor do filtro ABS e o nível de radiação é inversamente proporcional: quanto menor o valor do filtro ABS, maior a quantidade de radiação que atinge o detector. Durante os testes, a maior taxa de radiação de fundo observada foi de 2.2 kHz/cm² com o filtro ABS de valor 1, o que equivale a aproximadamente três vezes a taxa esperada para o HL-LHC.

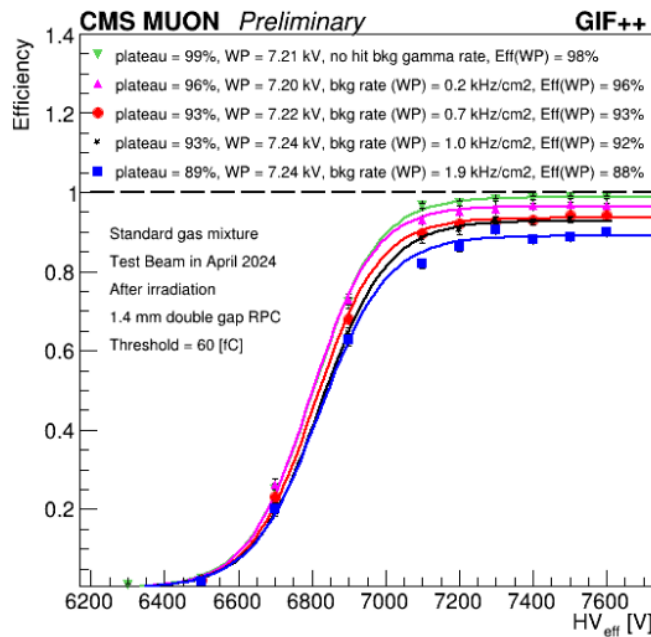


Figura 22 – Eficiência para a mistura de gases padrão com fontes: OFF (verde), 10 (rosa), 3.3 (vermelho), 2.2 (preto) e 1 (azul), durante o teste de feixe de 2024.

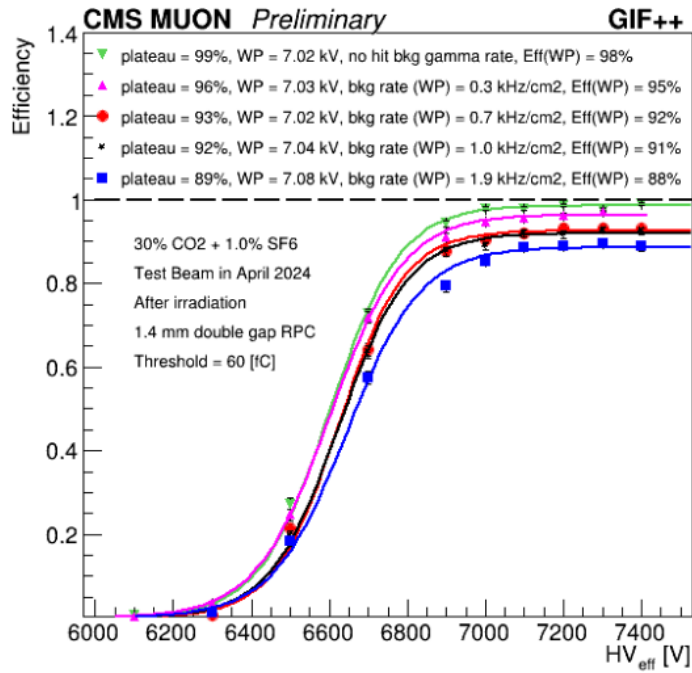


Figura 23 – Eficiência para a mistura de gases 30% CO₂ + 1.0% SF₆ com fontes: OFF (verde), 10 (rosa), 3.3 (vermelho), 2.2 (preto) e 1 (azul), durante o teste de feixe de 2024.

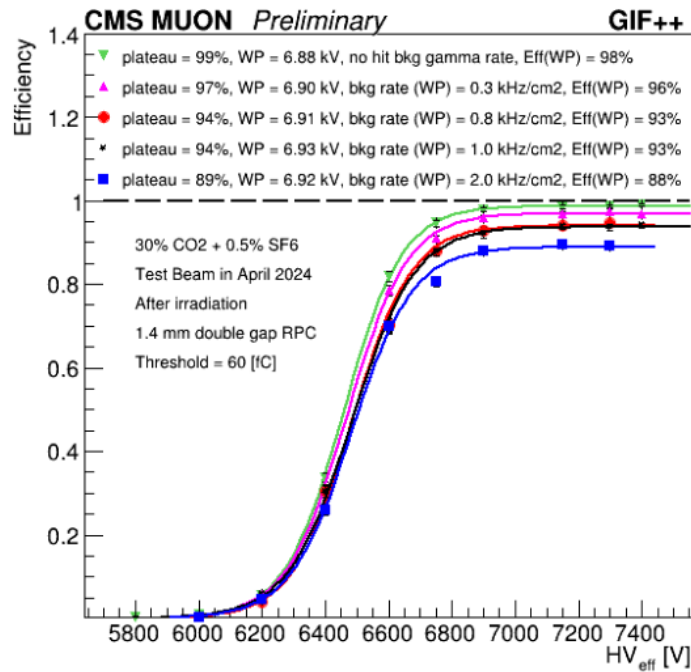


Figura 24 – Eficiência para a mistura de gases 30% CO₂ + 0.5% SF₆ com fontes: OFF (verde), 10 (rosa), 3.3 (vermelho), 2.2 (preto) e 1 (azul), durante o teste de feixe de 2024.

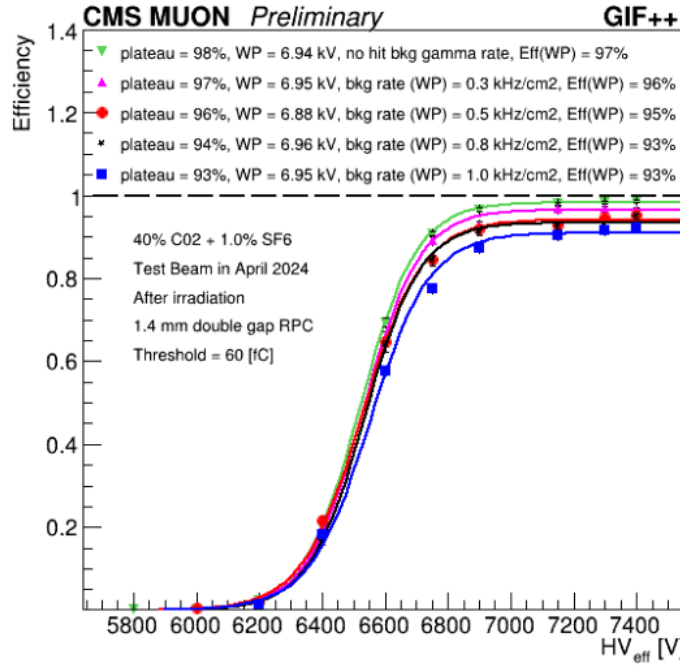


Figura 25 – Eficiência para a mistura de gases 40% CO₂ + 1.0% SF₆ com fontes: OFF (verde), 10 (rosa), 3.3 (vermelho), 2.2 (preto) e 1 (azul), durante o teste de feixe de 2024.

Durante os testes realizados no GIFF++, as câmaras RPC são instaladas na região upstream da área de testes, posicionadas entre dois detectores cintiladores que funcionam como sistema de *trigger*. A eficiência de detecção é determinada calculando-se, para cada valor de alta tensão aplicada, a razão entre o número de eventos registrados pelas câmaras RPC e o número total de eventos detectados pelo sistema de *trigger*. Os pontos obtidos são então ajustados por uma curva sigmoide, característica do comportamento de detectores gasosos operando no modo avalanche, refletindo o crescimento da eficiência até o atingimento do platô em regimes de alta tensão.

Os resultados das curvas de eficiência (Figuras 10, 11, 12 e 13) indicam que a mistura contendo 30% de CO₂ e 0,5% de SF₆ oferece o melhor equilíbrio entre eficiência e baixo GWP, garantindo desempenho adequado para as demandas do HL-LHC.

5 Considerações Finais

Neste trabalho, foi realizada uma medida direta da largura invisível do bóson Z utilizando dados públicos disponibilizados pelo *CERN Open Data portal*, referentes às eras G e H do ano de 2016, com energia de centro de massa de $\sqrt{s} = 13$ TeV. A análise considerou eventos em que o bóson Z decai em neutrinos, inferidos a partir do momento transversal ausente (p_T^{miss}), em associação com jatos. O método adotado baseou-se na extrapolação a partir de um canal de referência — o decaimento do Z em léptons carregados — e na aplicação de um ajuste simultâneo a múltiplas regiões de controle, com o uso da técnica de amostragem por cadeias de Markov Monte Carlo (MCMC).

O valor obtido para a largura invisível foi:

$$\Gamma(Z \rightarrow \nu\bar{\nu}) = 634,6_{-43,3}^{+44,5} \text{ MeV},$$

ou, de forma aproximada, 635 ± 44 MeV. Este resultado apresenta um desvio de aproximadamente $2,4\sigma$ em relação à medida realizada pela colaboração CMS, que reportou $\Gamma = 523 \pm 16$ MeV. Esse desvio pode ser atribuído, em grande parte, à limitação estatística dos dados disponíveis no CMS Open Data, bem como às incertezas no fator de transferência r_Z , impactado pela baixa estatística do Monte Carlo na região de controle double lepton.

Além da abordagem tradicional baseada em cortes físicos, foi conduzido um estudo complementar utilizando técnicas de aprendizado de máquina para a seleção de eventos da região de sinal. Especificamente, foram empregadas *Graph Neural Networks* (GNNs), que exploram representações de eventos em forma de grafos para identificar padrões complexos nas variáveis cinemáticas. O modelo treinado obteve desempenho robusto, com métricas superiores a 85% em acurácia e F1-score, demonstrando sua capacidade de generalizar para novos dados e identificar eficientemente os eventos de interesse. Apesar de certas limitações na reprodução de cortes de veto, o estudo com GNNs se mostrou promissor como alternativa à estratégia de seleção cut-based.

Apesar das limitações, esta análise demonstra a viabilidade de realizar medições sofisticadas com dados abertos. Além de validar um fluxo completo de análise — desde a seleção de eventos, correções experimentais e visualizações até o ajuste estatístico final — este estudo reforça o potencial pedagógico e científico do CMS Open Data como ferramenta para treinamento e pesquisa em física de partículas.

Adicionalmente, a dissertação apresentou uma contribuição prática relacionada ao desenvolvimento técnico no experimento CMS, com participação na caracterização e nos testes dos detectores iRPCs durante o estágio de mestrado sanduíche realizado no CERN.

Como perspectivas futuras, a análise aqui realizada pode ser refinada com a inclusão

de mais dados. A aplicação de abordagens mais avançadas, como modelos baseados em aprendizado de máquina — incluindo arquiteturas de GNNs mais complexas com múltiplos objetos reconstruídos — representa uma extensão natural e promissora para o estudo de observáveis invisíveis, para testes de precisão do Modelo Padrão, bem como para investigações de possíveis sinais de física além do modelo, utilizando dados abertos.

Referências

- 1 Wikimedia Commons contributors. *Complexo de Aceleradores do CERN*. 2023. <<https://pt.m.wikipedia.org/wiki/Ficheiro:Cern-accelerator-complex.svg>>. Acesso em: 15 maio 2025. Citado 2 vezes nas páginas 10 e 29.
- 2 CERN Courier. *CMS starts underground*. 2008. Acesso em: 15 maio 2025. Disponível em: <<https://cerncourier.com/a/cms-starts-underground/>>. Citado 2 vezes nas páginas 10 e 31.
- 3 CONNOR, P. L.; CONNOR, P. L. The large hadron collider and the compact muon solenoid. *Inclusive b Jet Production in Proton-Proton Collisions: Precision Measurement with the CMS experiment at the LHC at $s = 13$ TeV*, Springer, p. 65–93, 2019. Citado 2 vezes nas páginas 10 e 33.
- 4 ABBRESCIA, M. *The upgrade of the CMS muon system*. [S.l.], 2013. Citado 2 vezes nas páginas 10 e 34.
- 5 WALTENBERGER, W. *Development of vertex finding and vertex fitting algorithms for CMS*. [S.l.], 2004. Citado 2 vezes nas páginas 10 e 35.
- 6 PELLECCIA, A. et al. The upgrade of the cms muon system for the high luminosity lhc. *Journal of Instrumentation*, IOP Publishing, v. 19, n. 02, p. C02077, 2024. Citado 2 vezes nas páginas 11 e 67.
- 7 KUMARI, P. et al. Improved-rpc for the cms muon system upgrade for the hl-lhc. *Journal of Instrumentation*, IOP Publishing, v. 15, n. 11, p. C11012, 2020. Citado 2 vezes nas páginas 11 e 68.
- 8 CMS collaboration. Precision measurement of the Z boson invisible width in pp collisions at $\sqrt{s} = 13$ TeV. *arXiv preprint*, 2022. Citado 6 vezes nas páginas 12, 15, 39, 49, 63 e 66.
- 9 CERN Open Data Portal. *CERN Open Data Policy for LHC Experiments*. 2021. Accessed: 2025-01-20. Disponível em: <<https://opendata.cern.ch/docs/cern-open-data-policy-for-lhc-experiments>>. Citado na página 15.
- 10 HAGIWARA, K. et al. Review of particle physics: Particle data group. *Physical Review D*, American Institute of Physics, v. 66, n. 1 I, p. 100011–10001958, 2002. Citado na página 15.
- 11 SAMALAN, A. et al. Upgrade of the cms resistive plate chambers for the high luminosity lhc. *Journal of Instrumentation*, IOP Publishing, v. 17, n. 01, p. C01011, 2022. Citado 3 vezes nas páginas 16, 67 e 68.
- 12 GRIFFITHS, D. *Introduction to elementary particles*. [S.l.]: John Wiley & Sons, 2020. Citado 4 vezes nas páginas 18, 20, 22 e 23.
- 13 MANDL, F.; SHAW, G. *Quantum field theory*. [S.l.]: John Wiley & Sons, 2013. Citado 7 vezes nas páginas 18, 19, 20, 23, 25, 26 e 27.

- 27 CMS. *B-Tagging Recommendations for 80X*. 2025. Accessed: 2025-01-29. Disponível em: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation80X>. Citado na página 44.
- 28 SHLOMI, J.; BATTAGLIA, P.; VLIMANT, J.-R. Graph neural networks in particle physics. *Machine Learning: Science and Technology*, IOP Publishing, v. 2, n. 2, p. 021001, 2020. Citado na página 51.
- 29 THAIS, S. et al. Graph neural networks in particle physics: Implementations, innovations, and challenges. *arXiv preprint arXiv:2203.12852*, 2022. Citado na página 51.
- 30 CMS. *Modeling Backgrounds - CMS Open Data Workshop 2024*. 2024. <https://cms-opendata-workshop.github.io/workshop2024-lesson-background-modeling/instructor/index.html>. Accessed: 2025-01-28. Citado na página 57.
- 31 DOBBS, M.; LEFEBVRE, M. Unweighted event generation in hadronic wz production at order (α s). *Physical Review D*, APS, v. 63, n. 5, p. 053011, 2001. Citado na página 57.
- 32 FRIXIONE, S.; WEBBER, B. R. Matching nlo qcd computations and parton shower simulations. *Journal of High Energy Physics*, IOP Publishing, v. 2002, n. 06, p. 029, 2002. Citado na página 57.
- 33 DANZIGER, K.; HÖCHE, S.; SIEGERT, F. Reducing negative weights in monte carlo event generation with sherpa. *arXiv preprint arXiv:2110.15211*, 2021. Citado na página 57.
- 34 CMS. *Utilities for Accessing Pileup Information for Data*. <https://twiki.cern.ch/twiki/bin/viewauth/CMS/PileupJSONFileforData>. Accessed: 2025-01-28. Citado na página 59.
- 35 COLLABORATION, C. The phase-2 upgrade of the cms tracker. *CERN, Geneva, Tech. Rep. CERN-LHCC-2017-009. CMS-TDR-014*, 2017. Citado na página 67.
- 36 PINHEIRO, J. P. G. et al. Performance of new co2 based mixture in cms improved resistive plate chambers in hl-lhc environment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, Elsevier, v. 1065, p. 169474, 2024. Citado na página 69.