Universidade de São Paulo
Instituto de Física

Medidas da produção de bósons W no canal de
elétrons em colisões próton-próton a $\sqrt{s} = 13$ TeV
no experimento ATLAS

Simão Paulo Silva

Orientador: Dr. Marco Aurelio Lisboa Leite

Tese de doutorado apresentada ao Instituto de Física
da Universidade de São Paulo, como requisito par-
cial para a obtenção do título de Doutor em Ciências.

Banca Examinadora:

Dr. Marco Aurélio Lisboa Leite - Orientador (IFUSP)

Prof. Dr. Marcelo Gameiro Munhoz (IFUSP)

Prof. Dr. Mauricio Moralles (IPEN)

Prof. Dr. Anderson Campos Fauth (UNICAMP)

Profa. Dra. Márcia Begalli (UERJ)

São Paulo

2019

FICHA CATALOGRÁFICA
Preparada pelo Serviço de Biblioteca e Informação
do Instituto de Física da Universidade de São Paulo

Silva, Simão Paulo

Medidas da produção de bósons W no canal de elétrons em colisões próton-próton a $\sqrt{s}=13$ TeV no experimento ATLAS. São Paulo, 2019.

Tese (Doutorado) – Universidade de São Paulo. Instituto de Física.
Depto. de Física Nuclear.

Orientador: Prof. Dr. Marco Aurélio Lisboa Leite
Área de Concentração: Física de Altas Energias.

Unitermos: 1. Física de alta energia; 2. Física de partículas; 3. Física experimental.

USP/IF/SBI-036/2019

University of São Paulo
Physics Institute

Measurements of W boson production in the
electron channel in proton-proton collisions at
 $\sqrt{s} = 13$ TeV in the ATLAS experiment

Simão Paulo Silva

Supervisor: Dr. Marco Aurélio Lisboa Leite

Thesis submitted to the Physics Institute of the
University of São Paulo in partial fulfillment of the
requirements for the degree of Doctor of Science.

Examining committee:

Dr. Marco Aurélio Lisboa Leite - Supervisor (IFUSP)

Prof. Dr. Marcelo Gameiro Munhoz (IFUSP)

Prof. Dr. Mauricio Moralles (IPEN)

Prof. Dr. Anderson Campos Fauth (UNICAMP)

Profa. Dra. Márcia Begalli (UERJ)

São Paulo

2019

Agradecimentos

Agradeço à minha família que sempre me apoiou, ao meu orientador Marco Leite que me aceitou como aluno e acreditou em mim quando nem eu mesmo acreditava, a todo o pessoal do LIP-IFUSP pelo apoio dado durante o desenvolvimento do trabalho, ao José e à Marisilvia pelas diversas vezes que me ajudaram, ao CNPq pelo suporte financeiro que viabilizou o desenvolvimento da presente tese, ao CERN e à colaboração ATLAS que permitira minha participação no experimento e, por meio do EPLANET, tornaram possível meu estágio presencial no CERN. Acima de tudo devo agradecer a Deus por ter me proporcionado todas as condições para a realização deste trabalho.

Resumo

Este trabalho apresenta os resultados da primeira medida da seção de choque de produção $W^\pm \rightarrow e^\pm \nu$ utilizando dados de colisões próton-próton com energia de centro de massa $\sqrt{s} = 13$ TeV pelo detector ATLAS em dois regimes diferentes de operação do *Large Hadron Collider* (LHC): um período de baixa luminosidade em que foram acumulados 81 pb^{-1} de dados e que permitiram medir a seção de choque fiducial e total da produção de bósons W^+ e W^- na região inclusiva (região de massa transversa acima de 50 GeV, abrangendo o pico de massa transversa do bóson W) e um período de alta luminosidade em que foi obtido $76,6 \text{ fb}^{-1}$ de dados tornando acessível a medida da seção de choque fiducial de produção de bósons W^+ e W^- na região de alta massa transversa (região de massa transversa entre 130 GeV e 2,5 TeV). As seções de choque medidas e as distribuições cinemáticas referentes à energia transversa faltante (E_T^{miss}), massa transversa do bóson W (m_T^W), pseudorapidez (η^e), momento transverso (p_T^e) do elétron decorrente do decaimento do W e a diferença entre o ângulo azimutal ($\Delta\phi$) do elétron e da E_T^{miss} são consistentes com previsões teóricas obtidas em NLO e normalizadas em NNLO, considerando-se as incertezas sistemáticas, estatísticas e na luminosidade.

Palavras-chave: Bósons W , LHC, detector ATLAS.

Abstract

This work presents the results of the first $W^\pm \rightarrow e^\pm \nu$ production cross section measurements in the proton-proton collisions at the center of mass energy of $\sqrt{s} = 13$ TeV obtained with the ATLAS detector in two different operation regimes: a low luminosity period when 81pb^{-1} of data was accumulated allowing the measurement of the fiducial and total production cross section of W^+ and W^- bosons in the inclusive region (region of transverse mass above 50 GeV, covering the transverse mass peak of the W boson) and a high luminosity period when $76,6\text{fb}^{-1}$ made accessible the high W boson transverse mass region (between 130 GeV and 2.5 TeV), allowing the measurement of the fiducial cross section of W^+ and W^- boson production. The measured cross sections and the kinematic distributions for the missing transverse energy, (E_T^{miss}), W boson transverse mass (m_T^W), the pseudorapidity (η^e), the transverse momentum (p_T^e) of the electron originated from the W boson decay as well as the azimuthal angle difference ($\Delta\phi$) between the electron and the E_T^{miss} are consistent with the theoretical predictions obtained at NLO and normalized to NNLO, considering the systematic, statistical and luminosity uncertainties.

Keywords: W bosons, LHC, ATLAS detector.

Lista de Figuras

1.1	Regiões de exclusão a 95% CL dos fatores de forma Y e W	21
3.1	Representação do complexo de aceleradores do CERN [32].	36
3.2	O detector do experimento ATLAS e seus subsistemas[22].	40
3.3	Cobertura em η e granularidade dos subsistemas do ATLAS [22].	41
3.4	Representação esquemática do sistema magnético do ATLAS	42
3.5	Visão em corte do detector interno (região do barril e tampas)[22].	43
3.6	Eventos de colisões de prótons a $\sqrt{s} = 13$ TeV em 2015	45
3.7	Visão em corte do sistema de calorimetria do experimento ATLAS[22].	46
3.8	Esquema simplificado do sistema de <i>trigger</i> do ATLAS[22].	48
4.1	Ilustração da topologia de um evento com energia transversa faltante [39].	52
4.2	Resolução em termos da $\sum E_T$ para os dados coletados (a) , e os dados simulados (b). Os valores do parâmetro de ajuste k estão em $\text{GeV}^{1/2}$ [40].	58
4.3	Distribuições da resolução como função de $\sum E_T$ em eventos de $Z \rightarrow \mu^+ \mu^-$ [41].	59
4.4	Distribuição da componente x da $\vec{E}_T^{\text{miss}}$ e da $\sum E_T$ para o canal $Z \rightarrow \mu^+ \mu^-$ para cada um dos algoritmos usados[51].	59
4.5	Comparação da linearidade entre os algoritmos envolvidos sem incluir jatos, esquerda, e incluindo jatos, direita[51].	60
4.6	Resolução da E_T^{miss} em termos da $\sum E_T$ para os processos $Z \rightarrow \mu^+ \mu^-$, (a,b), e $W^\pm \rightarrow \mu^\pm \nu$ (c,d). Os resultados foram separados em eventos que não incluem jatos (a,c) , e incluindo jatos (b,d)[51].	61

4.7	Eficiências na identificação de elétrons de uma amostra $Z \rightarrow e^+e^-$ (a) e de uma amostra de dijetos (b). Para a amostra de $Z \rightarrow e^+e^-$, com elétrons reconstruídos a partir de elétrons verdadeiros, a eficiência aumenta com a E_T , enquanto diminui para a amostra de dijetos, da qual os elétrons são jatos identificados incorretamente como elétrons[52].	64
5.1	Distribuição de $\langle\mu\rangle$ dos dados usados na análise na região inclusiva.	74
5.2	Distribuições do número de vértice para dados e simulações	75
5.3	Distribuição das variáveis de isolamento relativas	78
5.4	Efeito do isolamento na distribuição de m_T	79
5.5	Distribuição de m_T para $W^\pm$ após seleção com o isolamento <i>gradient</i> (a) , e com o mesmo isolamento invertido (b).	80
5.6	Efeito do isolamento do <i>trigger</i> nas variáveis de isolamento	80
5.7	Distribuições da massa transversa de $W^\pm$ obtidas calculando-se a contribuição dos multijatos por simulação (linha vermelha) e usando-se um método baseado em dados (linha azul), normalizados em relação ao número total de eventos, mostrando que a simulação não reproduz a distribuição dos dados de forma adequada.	81
5.8	Distribuição das variáveis de isolamento reais	82
5.9	Histogramas do número de eventos por variável de isolamento para o canal $W^- \rightarrow e^- \bar{\nu}$. (a,b) histogramas de $pt_{varcone20}/p_T$ e (c,d) os histogramas de $topoetcone20/p_T$, sem o corte em m_T (a,c) e sem o corte em E_T^{miss} (b,d).	83

- 5.10 Distribuições cinemáticas para produção de bósons W . O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa hachurada e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos). (a) momento transverso do elétron, (b) distribuição em pseudorrapidez dos elétrons η^e , (c) energia transversa faltante E_T^{miss} e (d) massa transversa m_T^W 85
- 5.11 Distribuições cinemáticas (p_T^e e η^e) para produção de bósons W separadas pelo sinal da carga dos elétrons reconstruído. O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa hachurada e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos). Momento transverso do elétron p_T^e ((a) e^+ , (b) e^-), distribuição em pseudorrapidez dos elétrons η^e ((c) e^+ , (d) e^-). 86

5.12	Distribuições cinemáticas (E_T^{miss} e m_T^W) para produção de bósons W separadas pelo sinal da carga dos elétrons reconstruído. O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa hachurada e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos. Energia transversa faltante E_T^{miss} ((a) e^+ , (b) e^-), massa transversa m_T^W ((c) e^+ , (d) e^-)).	87
6.1	$\langle\mu\rangle$ durante o Run-2	92
6.2	Distribuição do número de médio interações por cruzamento de feixe ($\langle\mu\rangle$) para os dados e as simulações de fundo e sinal na região inclusiva . As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa rosa e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos).	98
6.3	Distribuição de $\Delta\phi$ na região inclusiva	101
6.4	Distribuição da E_T^{miss} na região inclusiva	102
6.5	Distribuição de p_T na região inclusiva	103
6.6	Distribuição de $\Delta\phi$ na região inclusiva	105
6.7	Distribuição da E_T^{miss} na região inclusiva	106
6.8	Distribuição do momento transverso do lépton (p_T^e) na região inclusiva	107
6.9	Efeito dos diferentes isolamentos no p_T^e	109
6.10	Distribuição de $\langle\mu\rangle$ na região de alta massa para a produção de $W^\pm$.	110
6.11	Distribuição de $\Delta\phi$ na região de alta massa para produção de $W^\pm$.	111
6.12	Distribuição da E_T^{miss} na região de alta massa para produção de $W^\pm$.	112

6.13	Distribuição da p_T^e na região de alta massa para a produção de $W^\pm$.	113
6.14	Distribuição de $\Delta\phi$ na região de alta massa para W^-, W^+	115
6.15	Distribuição da E_T^{miss} na região de alta massa para W^-, W^+	116
6.16	Distribuição da $p_T^{l^-}$ na região de alta massa para W^-, W^+	117
6.17	Sistemáticas nas distribuições cinemáticas na região inclusiva	121
6.18	Sistemáticas nas distribuições cinemáticas na região de alta massa . .	122
6.19	Distribuições de discriminantes de saída calculadas pelos cinco métodos de MVA estudados, considerando apenas as variáveis cinemáticas da reconstrução do bóson W no conjunto de variáveis de entrada: (a) discriminantes de Fisher, (b) discriminates de Fisher com transformação gaussiana das variáveis de entrada, (c) <i>Boosted Decision Tree</i> adaptivo, (d) <i>Boosted Decision Tree</i> com <i>Gradient Boosting</i> e (e) <i>Boosted Decision Tree</i> adaptivo com decorrelação. Nota-se uma significativa sobreposição das distribuições de sinal e fundo presente em todos os métodos.	133
6.20	Curvas características de operação do receptor (ROC) para os cinco métodos de MVA estudados, considerando apenas as variáveis cinemáticas da reconstrução do bóson W no conjunto de variáveis de entrada.	134
6.21	Multiplicidade de jatos ($nJets$) na região de alta massa transversa para $W^\pm$	135
6.22	Distribuições de saída calculadas pelos cinco métodos de MVA estudados, incluindo a multiplicidade de jatos juntamente com as variáveis cinemáticas da reconstrução do bóson W no conjunto de variáveis de entrada. (a) discriminantes de Fisher, (b) discriminates de Fisher com transformação gaussiana das variáveis de entrada, (c) <i>Boosted Decision Tree</i> adaptivo, (d) <i>Boosted Decision Tree</i> com <i>Gradient Boosting</i> e (e) <i>Boosted Decision Tree</i> adaptivo com decorrelação. Comparado à Fig. 6.19 observa-se uma significativa melhora na capacidade de discriminação após a inclusão da multiplicidade de jatos.	137

6.23	Curvas características de operação do receptor (ROC) para os cinco métodos de MVA estudados, incluindo a multiplicidade de jatos juntamente com as variáveis cinemáticas da reconstrução do bóson W no conjunto de variáveis de entrada.	138
6.24	(a) Matriz de correlação das variáveis de entrada na amostra de sinal ($W^\pm$) e (b) na amostra de fundo ($t\bar{t}$).	139
6.25	Curvas de eficiência para a seleção do sinal, a eficiência na rejeição do fundo, a pureza do sinal, o produto entre a eficiência e pureza do sinal e a significância para cada um dos cinco métodos de MVA. . . .	141
6.26	Distribuição de $\Delta\phi$ na região de alta massa	142
6.27	Distribuição da E_T^{miss} na região de alta massa	143
6.28	Distribuição de momento transverso do elétron (p_T^e) (a,b) e multiplicidade de jatos (c,d) para produção de $W^\pm$ na região de alta massa transversa após o corte de MVA. O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa rosa e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos). (a,c) escala linear, (b,d) escala logarítmica.	144

6.29	Distribuição do número médio de colisões ($\langle\mu\rangle$) na região de alta massa transversa para produção de $W^\pm$ após o corte de MVA. O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa rosa e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos).	145
6.30	Distribuição de $\Delta\phi$ na região de alta massa	146
6.31	Distribuição da E_T^{miss} na região de alta massa	147
6.32	Distribuição do momento transverso do elétron (p_T^e) (a,b) e multiplicidade de jatos (c,d) na região de alta massa transversa após o corte de MVA. O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa rosa e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos). (a,c) W^- , (b,d) W^+	148
A.1	Razão entre os canais de elétrons e múons para as seções de choque fiduciais de produção de bósons $W^\pm$ e Z , comparados com as previsões do Modelo Padrão e as bandas médias do PDG [14].	154
A.2	Razões entre seções de choque fiduciais $\sigma_{W^+}^{\text{fid}}/\sigma_{W^-}^{\text{fid}}$ e $\sigma_{W^\pm}^{\text{fid}}/\sigma_Z^{\text{fid}}$, representadas pelas linhas vermelhas, e as razões calculadas teoricamente usando diversos conjuntos de PDF's, com as bandas de incertezas estatísticas e totais.	154

A.3	Razão entre as razões das seções de choque fiduciais teóricas e medidas para cada bóson, usando diversos conjuntos de PDFs.	155
B.1	Figura conceitual da aplicação do método dos discriminantes de Fisher mostrando a distribuição de sinal (S) e fundo (B) de duas variáveis hipotéticas e a distribuição de um discriminante de Fisher (y_{Fisher}), que pode ser entendida como a projeção dos dados num eixo cuja orientação é calculada de forma a maximizar a separação entre os valores médios das distribuições de sinal e fundo enquanto minimiza a dispersão de cada distribuição projetada [100].	157
B.2	Princípio de funcionamento do método de classificação por <i>Boosted Decision Trees</i> : A árvore é construída através de uma sequência de separações. Em cada etapa, a variável que melhor discrimina entre sinal e fundo é escolhida pelo algoritmo classificador quando um corte c_i é aplicado. No final, cada folha é classificada como sinal (S) ou fundo (B) e a árvore construída (ou treinada) dessa forma pode ser utilizada para classificar os eventos da amostra de dados.	159

Lista de Tabelas

2.1	Férmions elementares do Modelo Padrão	23
2.2	Os Bósons do Modelo Padrão	24
2.3	Férmions e seus números quânticos segundo o Modelo Eletrofraco. . .	27
3.1	Parâmetros de máquina nominais do LHC [4].	37
5.1	Eventos simulados usados na análise de bósons $W^\pm$	71
5.2	Contribuições no canal $W \rightarrow e\nu$ das amostras geradas por simulação em relação ao número total de eventossimulados sobreviventes para cada um dos canais estudados, em porcentagem.	88
5.3	Incertezas sistemáticas relativas $\delta C_W/C_W$ nos diversos canais, em porcentagem.	89
5.4	Aceitação A , fator C e seções de choque calculadas para cada canal. .	89
5.5	Seções de choque calculadas pelo gerador FEWZ utilizando PDF CT14nnlo.	89
6.1	Geradores de eventos e PDFs usados para cada uma das amostras geradas por simulação de Monte Carlo, a ordem de cálculo e as PDFs empregadas. Também indicada é a ordem da normalização da seção de choque (k -factor), calculada utilizando geradores de ordem fixa (v. texto)	94

6.2	Intervalos de massa utilizados pelo gerador de eventos Powheg na geração dos eventos de sinal ($W \rightarrow e^- \nu$, $W \rightarrow e^+ \nu$) e para cada componente do fundo ($Z \rightarrow l^+ l^-$, $W \rightarrow \mu^- \nu$, $W \rightarrow \mu^+ \nu$, $W \rightarrow \tau^- \nu$, $W \rightarrow \tau^+ \nu$). A amostra final os intervalos são combinados considerando-se as respectivas seções de choque, os k -factors e a eficiência dos filtros na seleção dos eventos de acordo com os intervalos de massa.	95
6.3	Contribuições dos fundos simulados na região inclusiva	100
6.4	Contribuições das simulações na região de alta m_T^W	109
6.5	Número de candidatos observados para os eventos no canal de elétrons do decaimento do bóson W , fundo total (correspondente a soma dos processos eletro-fracos e dos processos dos canais de produção de quarks top e di-bósons) e o número final de eventos de sinal correspondentes à diferença entre o número de candidatos iniciais e o fundo. As incertezas sistemáticas no fundo correspondem às incertezas sistemáticas experimentais	127
6.6	Fatores de correção C_W para os canais W^+ , W^- e $W^\pm$	128
6.7	Seção de choque medida para os processos $W^- \rightarrow e^- \nu$, $W^+ \rightarrow e^+ \nu$ e $W^\pm \rightarrow e^\pm \nu$	128
6.8	Seção de choque prevista pelo gerador PowhegPythia8 para os canais $W^- \rightarrow e^- \nu$, $W^+ \rightarrow e^+ \nu$ e $W^\pm \rightarrow e^\pm \nu$	129
6.9	<i>Ranking</i> das variáveis cinemáticas MET , p_T^l , m_T e $\Delta\phi$ obtidas em cada um dos métodos de MVA. Variáveis com maior poder discriminatório são classificadas com o $rank = 1$	132
6.10	<i>Ranking</i> das variáveis cinemáticas MET , p_T^l , m_T , $\Delta\phi$ e $nJets$ obtidas em cada um dos métodos de MVA. Variáveis com maior poder discriminatório são classificadas com o $rank = 1$	136
6.11	Significância máxima e respectivo ponto de corte onde a significância é máxima para cada um dos cinco métodos de MVA.	140
6.12	Seção de choque, após a aplicação do MVA, medida para os processos $W^- \rightarrow e^- \nu$, $W^+ \rightarrow e^+ \nu$ e $W^\pm \rightarrow e^\pm \nu$	149

6.13	Seção de choque, após a aplicação do MVA, prevista pelo gerador PowhegPythia8 para os canais $W^- \rightarrow e^- \nu$, $W^+ \rightarrow e^+ \nu$ e $W^\pm \rightarrow e^\pm \nu$.	149
A.1	Seções de choque calculadas combinando-se os léptons com as respec- tivas incertezas estatísticas, sistemáticas e de luminosidade.	153

Conteúdo

1	Introdução	18
2	Fundamentação Teórica	22
2.1	O Modelo Padrão	22
2.1.1	Partículas Fundamentais	22
2.1.2	Interações fundamentais	24
2.1.3	Quebra espontânea de simetria	29
2.2	Fenomenologia das colisões próton-próton	31
3	Aparato Experimental	35
3.1	LHC: O Grande Colisor de Hádrons	35
3.1.1	Operações	38
3.2	O Experimento ATLAS	39
3.2.1	Sistema de coordenadas e nomenclatura	40
3.3	O sistema magnético	41
3.4	O detector interno	43
3.5	Os calorímetros	44
3.6	O espectrômetro de múons	45
3.7	Os detectores dianteiros	46
3.8	<i>Trigger</i> e aquisição de dados	47
3.8.1	Sistema de <i>trigger</i> dos calorímetros	48
4	Reconstrução de elétrons e da energia transversa faltante no ATLAS	50
4.1	Definição da energia transversa faltante	50
4.2	Algoritmos usados no cálculo da E_T^{miss}	51
4.2.1	CST E_T^{miss}	52
4.2.2	Algoritmos baseados em traços	55

4.2.3	Método do recuo hadrônico	56
4.3	Resultados de estudos de desempenho	57
4.4	Reconstrução de elétrons	60
4.5	Identificação de elétrons	62
4.6	Isolamento de elétrons	63
4.6.1	As variáveis de isolamento	63
5	Medidas da produção $W^\pm \rightarrow e^\pm \nu$ em colisões próton-próton a $\sqrt{s} = 13$ TeV	66
5.1	Metodologia	68
5.2	Amostras de dados e amostras simuladas	69
5.3	Seleção de eventos	70
5.4	Correções e incertezas sistemáticas	72
5.4.1	Correções e sistemáticas dos objetos físicos	74
5.5	Estudos dos critérios de isolamento do elétrons na região inclusiva . .	77
5.5.1	Abordagem funcional	77
5.5.2	Isolamento no <i>trigger</i>	78
5.6	O fundo de multijatos	81
5.6.1	Metodologia	82
5.7	Resultados	84
5.8	Contribuição dos fundos determinados por simulação	88
5.8.1	O fator C_W	88
5.8.2	Cálculo da seção de choque	88
6	Medidas da produção $W^\pm \rightarrow e^\pm \nu$ na região de alta massa transversa em colisões próton-próton a $\sqrt{s} = 13$ TeV	91
6.1	Previsões de Monte Carlo para a estimativa do sinal e o fundo	92
6.2	Seleção de eventos	96
6.3	Correções e normalização	97
6.3.1	Correções nas condições de empilhamento	97
6.4	Resultados	99
6.4.1	Distribuições cinemáticas - região inclusiva	100
6.5	Estudos na região de alta massa transversa	108

6.5.1	Estudos dos critérios de isolamento do elétrons na região de alta massa transversa	108
6.6	Discussão das incertezas sistemáticas	118
6.7	Medida da seção de choque fiducial e comparação com as previsões teóricas	123
6.7.1	Cálculo da seção de choque fiducial	127
6.8	Estudo da redução do fundo de $t\bar{t}$ por análise multivariável	129
6.8.1	Implementação dos métodos de classificação multi-variável . . .	131
6.8.2	Determinação do ponto de operação da MVA e distribuições cinemáticas	140
7	Conclusões	151
A	Combinação dos resultados com os canais de múons e Z	153
B	Métodos de análise multi-variável	156
B.1	Discriminantes de Fisher	156
B.2	<i>Boosted Decision Trees</i>	157

Capítulo 1

Introdução

O Modelo Padrão [1, 2, 3] é hoje uma das teorias de maior sucesso na física, descrevendo os constituintes básicos da matéria e as forças que regem as interações fundamentais. Última peça que faltava ao modelo padrão, a descoberta do bóson de Higgs pelos experimentos ATLAS e CMS em 2012 permitiu consolidar a compreensão do mecanismo pelo qual as partículas adquirem massa.

Hoje, a fronteira da física de partículas é explorada pelos quatro maiores experimentos do *Large Hadron Collider* ou LHC [4]. A primeira fase de operações do LHC, entre 2009 e 2013, foi muito bem sucedida, tendo como principais resultados a já mencionada descoberta do bóson de Higgs pelos experimentos ATLAS e CMS [5, 6], a descoberta de pentaquarks pelo LHCb [7] e a confirmação da simetria entre núcleos e antinúcleos leves pelo ALICE [8].

Entretanto, os experimentos do LHC não realizam apenas descobertas diretas. Medidas de precisão de processos descritos pelo modelo padrão são parte importante das análises desenvolvidas e avançam para espaços de fase cada vez mais restritos à medida que os experimentos amadurecem e a quantidade de dados aumenta. Nesse cenário muitas medidas são realizadas com precisão maior que aquela atualmente presente no estado da arte dos modelos teóricos. Nessa direção, as medidas de seção de choque de produção de bósons de gauge W e Z como mediadores da força fraca oferecem um rico ambiente de teste para as previsões do Modelo Padrão uma vez que a sua assinatura clara no canal leptônico e sua grande seção de choque permite realizar medidas de alta precisão, inclusive separando o espaço de fase em várias regiões para medidas diferenciais.

Tais processos são também fundos irreduzíveis em processos que envolvem buscas além do modelo padrão ou medidas que buscam caracterizar as propriedades do bóson de Higgs, como seu acoplamento a férmions ou mesmo a bósons de gauge. Neste trabalho, as duas faces da análise envolvendo a produção de bósons de gauge (aqui restritas ao estudo da produção de bósons $W^\pm$ no canal de elétrons) são exploradas: i) medidas na região inclusiva, definida como a região ao redor do pico de massa transversa do bóson realizadas com elevada precisão auxiliada pela baixa luminosidade do LHC e alta seção de choque do processo e ii) medidas na região de alta massa ou acima do pico de massa transversa, possível apenas pela copiosa quantidade de dados produzidas nos últimos 3 anos pelo LHC na região de 13 TeV. A primeira medida têm o poder de restringir as diversas funções que descrevem a distribuição de partons no interior do próton (PDFs) e de testar propriedades fundamentais como a universalidade leptônica. A segunda têm o poder de prospectar processos que seriam acessíveis apenas a energias maiores que a atual energia do LHC, buscando possíveis diferenças na seção de choque de produção em alta massa.

O estudo da produção de bósons vetoriais Z e $W^\pm$ tem sido fundamental para o desenvolvimento de modelos que descrevem a QCD e as interações eletrofracas. Previsões teóricas hoje disponíveis em NNLO (*Next to Next Leading Order*) em QCD e em NLO (*Next to Leading Order*) em processos radiativos eletrofracos. A produção de bósons vetoriais Z e $W^\pm$ é altamente sensível às PDFs dos prótons (*Parton Distribution Function*) e, associada a uma assinatura clara no canal leptônico, é uma ferramenta importante para descrever a dinâmica da colisão na escala de TeV, além de, pela alta seção de choque, permitirem estudos de desempenho dos detectores e métodos de reconstrução. Entretanto, ainda que as medidas realizadas na região inclusiva (*on-shell*) apresentem a melhor precisão experimental, elas estão confinadas a um espaço de fase restrito pela fração x do momento dos partons e do quadrivetor da transferência do momento Q . Estender essa medida para a região de alta massa invariante no caso do bóson Z ou transversa para o caso do bóson $W^\pm$ permite cobrir uma faixa mais ampla de x e Q onde as PDFs dos quarks são conhecidas com pouca precisão. Além disso, tanto o ATLAS como o CMS tem produzido nos últimos anos buscas diretas por ressonâncias na região de alta massa (invariante ou transversa)

que possam fornecer evidências para os parceiros dos bósons Z e $W^\pm$ propostos por várias extensões do modelo padrão: os bósons Z' e W' . Ainda que uma eventual ressonância possa ser tão alta de forma a estar fora do alcance do LHC, é possível realizar medidas de precisão na região de alta massa de tal forma a evidenciar um efeito que a "cauda" da ressonância provoque no contínuo do espectro.

O potencial para a descoberta de nova física na região de alta massa tem recebido uma atenção especial nos últimos anos em virtude do desenvolvimento de modelos baseados em teorias de campo efetivas (EFT). Nesses modelos novos bósons vetoriais massivos se misturam com os bósons do modelo padrão num processo que pode ser descrito por propagadores oblíquos [9], que em LO correspondem aos propagadores $\hat{S}, \hat{T}, W$ e Y que modificam os propagadores γ, Z e $W^\pm$. W e Y são propagadores que dão origem a efeitos que aumentam com a energia de centro de massa da colisão. No modelo padrão, tais propagadores oblíquos devem ser zero. Ao implementar tais propagadores em previsões teóricas com valores diferentes de zero, é possível propor distribuições cinemáticas que sejam sensíveis a essa diferença e que podem ser testados hoje graças à grande quantidade de dados disponíveis no LHC. A Fig. 1.1 mostra como a sensibilidade do LHC se compara com a do LEP (*Large Electron-Positron collider*, colisor elétron-pósitron do CERN, antecessor do LHC [10].) para os parâmetros oblíquos W e Y . Nota-se que o LHC já possui uma sensibilidade maior que a do LEP para o parâmetro W .

Esta tese está organizada da seguinte forma: o Cap. 2 apresenta uma breve descrição de alguns aspectos-chave do Modelo Padrão e da fenomenologia das colisões. O Cap. 3 se dedica à descrição do *Large Hadron Collider* e do experimento ATLAS e seus subsistemas. O Cap. 4 discute o processo de reconstrução dos elétrons e da energia transversa faltante no ATLAS. O Cap. 5 apresenta as primeiras medidas de seção de choque na região inclusiva obtidas para a produção de bósons W a $\sqrt{s} = 13$ TeV no canal de elétrons. O Cap. 6 estende o escopo da análise apresentada no capítulo anterior para a região de alta massa transversa e finalmente o Cap. 7 apresenta as conclusões obtidas neste trabalho.

Figura 1.1: Regiões de exclusão a 95% CL dos fatores de forma Y e W [9], para diferentes processos e escalas de energia e luminosidade integrada.

Capítulo 2

Fundamentação Teórica

2.1 O Modelo Padrão

O Modelo Padrão (MP) da física de partículas fornece uma descrição das partículas elementares e suas interações. Dentre as suas previsões mais importantes e experimentalmente confirmadas, encontram-se os bósons $W^\pm$ e Z obtidos no CERN [11]; o quark top, descoberto pelo Tevatron [12]; e mais recentemente, o bóson de Higgs observado no CERN [5, 6].

O Modelo Padrão postula a existência de partículas elementares denominadas férmions, que formam a matéria conhecida. Os férmions são classificados em léptons e quarks e suas interações são descritas por campos de *gauge* mediados pela troca de bósons. Com estes ingredientes, o modelo é capaz de descrever três das quatro forças conhecidas: as forças fraca, forte e eletromagnética. Somente a força gravitacional não é incluída. A descrição teórica obedece à formulação da teoria quântica de campos, cuja premissa é a invariância sob transformações de calibre locais, também denominadas invariância de *gauge*. As bases da simetria $SU(3)_C \times SU(2)_L \times U(1)_Y$ serão detalhadas a seguir, a partir da bibliografia adotada nos cursos de formação [1, 2, 3] .

2.1.1 Partículas Fundamentais

Férmions

Experimentos têm demonstrado que os férmions, partículas elementares de spin $\frac{1}{2}$, existem em três gerações ("sabores"). As gerações dos quarks e léptons diferem entre si somente pelas massas que aumentam da primeira para a terceira geração.

A Tabela 2.1 lista os férmions elementares do modelo padrão com suas respectivas cargas elétricas.

Toda a matéria é basicamente constituída somente por partículas de primeira geração: o elétron e os constituintes dos prótons e nêutrons: os quarks *up* e *down* . Os outros férmions ocorrem em processos de altas energias, tanto em raios cósmicos como em colisores de partículas como o LHC.

Tabela 2.1: Classificação dos férmions elementares em três gerações de quarks e léptons e respectivas cargas elétricas. Para cada partícula carregada, existe uma anti-partícula com carga oposta.

Férmions	1ª geração		2ª geração		3ª geração		Carga [e]
Quarks	up	u	charm	c	top	t	$+\frac{2}{3}$
	down	d	strange	s	bottom	b	$-\frac{1}{3}$
Léptons	neutrino- e	ν_e	neutrino- μ	ν_μ	neutrino- τ	ν_τ	0
	elétron	e	múon	μ	tau	τ	-1

Bósons

As interações entre os férmions são mediadas por bósons, que são partículas de spin inteiro. Os bósons vetoriais (spin 1) são dados na Tabela 2.2. O fóton, γ , é o mediador da força eletromagnética, comumente denominado luz visível num estreito intervalo de energia. O fóton é estável, não apresenta massa, sendo infinito o alcance da interação com as partículas portadoras de carga elétrica.

Os glúons (g) são portadores da força forte e interagem com partículas que possuem um número quântico denominado cor, isto é, com eles mesmos e com os quarks. Os glúons existem em oito estados de cor diferentes e também não têm massa. Entretanto, a interação forte apresenta um pequeno alcance, já que o MP proíbe a existência de partículas de cor não ligadas.

Os bósons $W^\pm$ e Z são os mediadores da força fraca. Eles são massivos e portanto, a interação fraca apresenta pequeno alcance. Eles interagem com todas as partículas portadoras de carga fraca, como descreveremos na próxima seção.

As massas dos bósons (e dos férmions via acoplamento de Yukawa) são geradas pelo mecanismo de Brout-Englert-Higgs (BEH) ou mecanismo de Higgs, que prevê a existência da partícula elementar de spin 0, o bóson de Higgs, descoberta expe-

rimantalmente pelos experimentos ATLAS e CMS do LHC. O conhecimento atual deste mecanismo será discutido na Seção 2.1.3.

Todos os quarks e léptons carregados são massivos. Oscilações no setor de neutrinos têm sido observadas experimentalmente, fornecendo evidência que suas massas são extremamente pequenas, mas não são nulas [13].

Tabela 2.2: Bósons do Modelo Padrão e suas respectivas cargas elétricas e massas. Os glúons existem em oito estados de cor.

Interação	Bósons	Carga [e]	Massa [GeV]
Eletromagnética	fóton γ	0	0
Forte	glúon g	0	0
Frac	$W^\pm$	± 1	$80,379 \pm 0,012$ [14]
	Z	0	$91,1876 \pm 0,0021$ [14]

2.1.2 Interações fundamentais

O Modelo Padrão é formado por várias componentes, cada qual descrevendo interações específicas das partículas elementares. Historicamente, a primeira parte é a Eletrodinâmica Quântica (QED). A força forte é descrita pela Cromodinâmica Quântica (QCD) e a força fraca é unificada com a QED no modelo Eletrofraco (EW) [15, 16, 17], também denominado modelo GSW (Glashow, Salam e Weinberg).

Eletrodinâmica Quântica

A QED é matematicamente descrita por uma densidade de Lagrangiana denotada por $\mathcal{L}$, derivada a partir de argumentos de simetria, especialmente a invariância sob transformações de fase locais. Um férmion livre (spin $\frac{1}{2}$) é descrito pela Lagrangiana:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi, \quad (2.1)$$

onde ψ é o spinor de Dirac do campo do férmion, $\bar{\psi} = \psi^\dagger\gamma^0$ é o spinor adjunto, γ^μ são as matrizes de Dirac e $\partial_\mu = \frac{\partial}{\partial x_\mu}$ as derivadas parciais. A equação de Dirac,

$$(i\gamma^\mu\partial_\mu - m)\psi = 0 \quad (2.2)$$

é a equação de movimento correspondente.

A Lagrangiana deve ser invariante sob transformações de fase locais da forma

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

onde a fase local é dada por $\alpha(x)$, dependendo do espaço e do tempo. Estas transformações formam o grupo unitário abeliano $U(1)$, já que $e^{i\alpha(x)}$ pode ser escrita como uma matriz $U = 1 \times 1$ com $U^\dagger U = 1$. O primeiro termo de Lagrangiana (Equação 2.1) não é invariante sob esta transformação, já que

$$\partial_\mu \psi \rightarrow \partial_\mu \psi' = e^{i\alpha(x)} \partial_\mu \psi + i e^{i\alpha(x)} \psi \partial_\mu \alpha(x). \quad (2.4)$$

A invariância é estabelecida pela introdução de um campo adicional A_μ com transformação

$$A_\mu \rightarrow A'_\mu = A_\mu + \frac{1}{e} \partial_\mu \alpha(x) \quad (2.5)$$

e a substituição de ∂_μ pela derivada covariante D_μ :

$$D_\mu = \partial_\mu - i e A_\mu. \quad (2.6)$$

Pode-se demonstrar que

$$D_\mu \psi \rightarrow D'_\mu \psi' = e^{i\alpha(x)} D_\mu \psi \quad (2.7)$$

e a Lagrangiana é reescrita por

$$\begin{aligned} \mathcal{L} &= i \bar{\psi} \gamma^\mu D_\mu \psi - m \bar{\psi} \psi \\ &= \bar{\psi} (i \gamma^\mu \partial_\mu - m) \psi + e \bar{\psi} \gamma^\mu \psi A_\mu. \end{aligned} \quad (2.8)$$

Desta forma, com a introdução do campo de *gauge* A_μ , a invariância de fase local é restaurada. Este campo se acopla ao campo do férmion através da constante e , a carga elétrica elementar. Ao se introduzir o termo cinético, usando o tensor de campo $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, o campo de *gauge* A_μ é identificado como o campo do fóton e a Lagrangiana da QED fica completa:

$$\mathcal{L}_{\text{QED}} = i \bar{\psi} \gamma^\mu \partial_\mu \psi - m \bar{\psi} \psi + e \bar{\psi} \gamma^\mu \psi A_\mu - \frac{1}{4} F_{\mu\nu} F^{\mu\nu}. \quad (2.9)$$

Os termos individuais são a energia cinética e a massa do férmion, o acoplamento entre os campos do férmion e do fóton e a energia cinética do fóton. A invariância

de fase local proíbe a introdução do termo de massa da forma $\frac{1}{2}m^2 A_\mu A^\mu$ para o campo do fóton. Como consequência, este formalismo determina que o fóton não tenha massa o que concorda com os resultados experimentais.

Cromodinâmica Quântica

A QCD descreve as interações entre quarks e glúons. Os quarks são férmions como os léptons, mas com um grau de liberdade adicional: a carga de cor, que existe nos estados vermelho (*red* - r), verde (*green* - g) e azul (*blue* - b). Os spinores de Dirac são substituídos por vetores dos três spinores que denotam os quarks:

$$\psi = \begin{pmatrix} \psi_r \\ \psi_g \\ \psi_b \end{pmatrix}. \quad (2.10)$$

Acrescentando os oito campos de *gauge* G_μ^a com $a = 1, \dots, 8$ que representam os glúons, podemos extrair a Lagrangiana da QCD:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi - g_s(\bar{\psi}\gamma^\mu\frac{\lambda_a}{2}\psi)G_\mu^a - \frac{1}{4}G_{\mu\nu}^a G_a^{\mu\nu}. \quad (2.11)$$

A Lagrangiana é invariante sob transformações do grupo unitário não abeliano $SU(3)_C$, cujos geradores são $T_a = \lambda_a/2$. g_s são a notação para a constante de acoplamento forte, λ_a são as oito matrizes de Gell-Mann e $G_{\mu\nu}^a$ é o tensor de força do campo:

$$G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - g_s f_{abc} G_\mu^b G_\nu^c. \quad (2.12)$$

As constantes de estrutura f_{abc} satisfazem $[\frac{\lambda_a}{2}, \frac{\lambda_b}{2}] = if_{abc}\frac{\lambda_c}{2}$. O último termo do tensor de força se deve ao fato de $SU(3)_C$ ser não abeliano e não possuir um equivalente na QED, permitindo que os glúons interajam entre si. Assim como na QED, a invariância de fase local requer que os portadores de força, no caso os glúons, não apresentem massa.

Modelo Eletrofraco

Todos os férmions estão sujeitos à força fraca, que se manifesta, por exemplo, nos decaimentos- β nucleares. A interação acontece pela troca de bósons W^+ e W^- ,

denominada interação de corrente carregada, e do bóson Z , denominada interação de corrente neutra. Enquanto a troca de bósons $W^\pm$ permite a mudança de sabor de quarks, não há observação de mudança de sabor por correntes neutras (*flavor-changing neutral currents*-FCNC). Os experimentos têm demonstrado que a corrente fraca carregada somente se acopla a férmions de helicidade¹ esquerda. Conservando a simetria $SU(2)_L \times U(1)_Y$ no modelo eletrofraco, os férmions de helicidade esquerda são agrupados em dubletos $SU(2)_L$ com isospin de $I = \frac{1}{2}$ e campos de *gauge* W_μ^a . Os férmions de helicidade direita são descritos por singletos $U(1)_Y$ com $I = 0$ e campo de *gauge* B_μ . Aqui, a hipercarga Y é introduzida, cuja relação com a carga elétrica $Q = I_3 + \frac{Y}{2}$ será obtida a seguir.

A Tabela 2.3 mostra um resumo dos férmions e respectivos números quânticos no modelo eletrofraco. Os autoestados fracos do quarks tipo *down*, denotados por d' , s' e b' são uma mistura de autoestados de massa d , s e b . A mistura é descrita pela matriz CKM [18].

Tabela 2.3: Férmions e seus números quânticos segundo o Modelo Eletrofraco.

	1 ^a	2 ^a	3 ^a	I	I_3	Y	$Q[e]$
quarks	$\begin{pmatrix} u \\ d' \end{pmatrix}_L$	$\begin{pmatrix} c \\ s' \end{pmatrix}_L$	$\begin{pmatrix} t \\ b' \end{pmatrix}_L$	$+\frac{1}{2}$	$+\frac{1}{2}$	$+\frac{1}{3}$	$+\frac{2}{3}$
	u_R	c_R	t_R	$+\frac{1}{2}$	$-\frac{1}{2}$	$+\frac{1}{3}$	$-\frac{1}{3}$
	d_R	s_R	b_R	0	0	$+\frac{4}{3}$	$+\frac{2}{3}$
	d_R	s_R	b_R	0	0	$-\frac{2}{3}$	$-\frac{1}{3}$
léptons	$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L$	$\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L$	$\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L$	$+\frac{1}{2}$	$+\frac{1}{2}$	-1	0
	e^-_R	μ^-_R	τ^-_R	$+\frac{1}{2}$	$-\frac{1}{2}$	-1	-1
	e^-_R	μ^-_R	τ^-_R	0	0	-2	-1

Os dubletos isospin de helicidade esquerda χ_L e os singletos de helicidade direita ψ_R se comportam, sob transformações de fase locais:

¹A helicidade é definida como a projeção do spin de uma partícula na direção de seu momento. No caso de partículas sem massa, ou com energia tal que a massa é negligenciável, a helicidade pode assumir os valores de $\pm\frac{1}{2}$, sendo as partículas de helicidade positiva ditas de helicidade direita e aquelas de helicidade negativa, de helicidade esquerda.

$$\chi_L(x) \rightarrow \chi'_L(x) = e^{i\alpha_a(x)\tau_a} e^{i\beta(x)Y} \chi_L, \quad (2.13)$$

$$\psi_R \rightarrow \psi'_R(x) = e^{i\beta(x)Y} \psi_R \quad (2.14)$$

formando a simetria $SU(2)_L \times U(1)_Y$. $\alpha(x)$ e $\beta(x)$ são as fases locais, τ_a com $a = 1, 2, 3$ são os geradores de $SU(2)_L$ e Y é o operador de hipercarga fraco, gerando $U(1)_Y$. A derivada covariante é dada por

$$D_\mu = \partial_\mu + igW_\mu^a \frac{\tau_a}{2} + ig'B_\mu \frac{Y}{2}, \quad (2.15)$$

onde g é a constante de acoplamento dos campos de *gauge* $SU(2)_L$, W_μ^a , e g' é a constante de acoplamento do campo de *gauge* $U(1)_Y$, B_μ . A Lagrangiana eletrofraca resultante é escrita como:

$$\mathcal{L}_{EW} = i\overline{\chi_L^i} \gamma^\mu D_\mu \chi_L^i + i\overline{\psi_R^i} \gamma^\mu D_\mu \psi_R^i - \frac{1}{4} W_{\mu\nu}^a W_a^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}, \quad (2.16)$$

onde a soma sobre os três dubletos de isospin e os seis singletos é realizada (índice i). Os tensores de campo são dados por:

$$W_{\mu\nu}^a = \partial_\mu W_\nu^a - \partial_\nu W_\mu^a - g\epsilon_{abc} W_\mu^b W_\nu^c \quad (2.17)$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu, \quad (2.18)$$

onde o tensor assimétrico, ϵ_{abc} , representa as constantes de estrutura de $SU(2)_L$, enquanto a constante de estrutura de $U(1)_Y$ é nula. Isto permite que os campos W_μ^a interajam entre si, mas o campo B_μ não apresenta autointeração. Os campos físicos são dados por combinações lineares de W_μ^a e B_μ ,

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2) \quad (2.19)$$

$$Z_\mu = \cos \theta_W W_\mu^3 - \sin \theta_W B_\mu \quad (2.20)$$

$$A_\mu = \sin \theta_W W_\mu^3 + \cos \theta_W B_\mu, \quad (2.21)$$

onde o ângulo de mistura fina, θ_W , que relaciona as constantes de acoplamento através de: $\cos \theta_W = g/\sqrt{g^2 + g'^2}$ e $\sin \theta_W = g'/\sqrt{g^2 + g'^2}$, foi introduzido. Reescrevendo a Lagrangiana em termos de campos físicos e comparando as componentes A_μ com o campo do fóton da QED (Equação 2.9), as seguintes relações são obtidas:

$$e = g \sin \theta_W = g' \cos \theta_W, \quad (2.22)$$

$$Q = I_3 + \frac{Y}{2}. \quad (2.23)$$

Assim como na QED e na QCD, a invariância de fase local proíbe a introdução de termos de massa para os bósons. Além do mais, termos de massa da forma $-m\bar{\psi}\psi$ para os férmions são proibidos no setor eletrofraco, o que conflita com os experimentos, cujos resultados mostram que os bósons Z , $W^\pm$ e os férmions são massivos, conforme a Tabela 2.2, no caso dos bósons. Este conflito foi resolvido pelo mecanismo a ser descrito na próxima seção.

2.1.3 Quebra espontânea de simetria

A invariância de fase local no modelo eletrofraco requer que os férmions e os bósons não apresentem massa. Entretanto, os resultados experimentais têm demonstrado que os bósons $W^\pm$ e Z^0 são bastantes massivos, conforme a Tabela 2.2. Esta contradição pode ser resolvida através de uma quebra espontânea de simetria, induzida pelo mecanismo Brout-Englert-Higgs (BEH) [19, 20] ou mais comumente conhecido como "mecanismo de Higgs".

O campo escalar de Higgs Φ , um dubleto de isospin fraco, com campos escalares complexos de hipercarga $Y = 1$, é postulado como

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \quad (2.24)$$

A Lagrangiana correspondente,

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \Phi)^\dagger (D^\mu \Phi) - \mu^2 \Phi^\dagger \Phi - \lambda (\Phi^\dagger \Phi)^2, \quad (2.25)$$

é invariante sob transformações de fase $SU(2)_L \times U(1)_Y$. O potencial, parametrizado como

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2, \quad (2.26)$$

tem o seu mínimo em $\Phi \neq 0$ para $\mu^2 < 0$ e $\lambda > 0$.

Um ponto de mínimo

$$\Phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \quad (2.27)$$

é escolhido como estado fundamental, onde $v = \sqrt{-\mu^2/\lambda}$ descreve o valor esperado do vácuo. Já que $v > 0$, a simetria do grupo é espontaneamente quebrada. O campo pode ser parametrizado como

$$\Phi(x) = \frac{e^{i\tau_a\theta_a(x)/v}}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \quad (2.28)$$

onde $\theta_a(x)$ ($a = 1, 2, 3$) e $h(x)$ são campos reais. A partícula que corresponde ao campo $h(x)$ é denominada bóson de Higgs, tendo spin 0. A exponencial contendo os campos θ_a (bósons de *Goldstone*) é eliminada na Lagrangiana devido à invariância de fase local e não apresenta significado físico.

Substituindo esta parametrização na Lagrangiana da Equação 2.25 o termo

$$\begin{aligned} \left| \left(ig \frac{\tau_a}{2} W_\mu^a + ig' \frac{Y}{2} B_\mu \right) \Phi_0 \right|^2 &= \left(\frac{1}{2} v g \right)^2 W_\mu^+ W^{\mu-} \\ &+ \left(\frac{1}{2} v g \right)^2 \frac{1}{2 \cos^2 \theta_W} Z_\mu Z^\mu \\ &= 0 \cdot A_\mu A^\mu \end{aligned} \quad (2.29)$$

é encontrado, sendo $|\dots|^2 = (\dots)^\dagger(\dots)$ a notação utilizada. Pode-se identificar os termos de massa dos bósons vetoriais e consequentemente, as seguintes relações

$$m_W = \frac{1}{2} v g, \quad m_Z = \frac{m_W}{\cos \theta_W}, \quad m_\gamma = 0. \quad (2.30)$$

Estas relações representam um marco em termos de previsões do Modelo Padrão tendo sido corroboradas experimentalmente antes da era LHC [21].

O valor esperado do vácuo $v = (\sqrt{2}G_F)^{-1/2} = 246$ GeV é determinado a partir de medidas do acoplamento de Fermi G_F em decaimentos de múons. Posteriormente, o termo de massa do bóson de Higgs pode ser determinado: $m_H = \sqrt{-2\mu^2}$. Esta relação não apresenta poder de previsão, já que o parâmetro μ pode ser livremente escolhido no MP.

De maneira similar, as massas dos fermions sao geradas a partir de seu acoplamento ao boson de Higgs, tambem denominado acoplamento de *Yukawa*, sendo descrito por uma componente invariante $SU(2)_L \times U(1)_Y$ adicional Lagrangiana,

$$\mathcal{L}_{\text{Yukawa}} = -G_l^{ij} \bar{L}_L^i \Phi l_R^j - G_d^{ij} \bar{Q}_L^i \Phi d_R^j - G_u^{ij} \bar{Q}_L^i \Phi u_R^j + \text{h.c.}, \quad (2.31)$$

onde $\bar{L}_L^i$ ($\bar{Q}_L^i$) sao os dubletos de isospin do lepton (quark), l_R^j (d_R^j, u_R^j) sao os singletos do lepton (quarks tipo down e up) e $\Phi_C(x) = \sqrt{1/2}(v + h(x), 0)^\top$. As constantes de acoplamento sao dadas por G_l^{ij} , G_d^{ij} e G_u^{ij} . No caso dos quarks, os autoestados da interao fraca sao uma mistura de autoestados de massa, enquanto que nenhuma mistura esta presente para os lepons. Por exemplo, o termo de Yukawa para o eltron dado por

$$\begin{aligned} \mathcal{L}_{\text{Yukawa}}^e &= \frac{G_e}{\sqrt{2}} v (\bar{e}_L e_R) - \frac{G_e}{\sqrt{2}} (\bar{e}_L e_R + \bar{e}_R e_L) h \\ &= -m_e \bar{e} e - \frac{m_e}{v} \bar{e} e h. \end{aligned} \quad (2.32)$$

A composio da Lagrangiana do Modelo Padro se escreve

$$\mathcal{L}_{\text{MP}} = \mathcal{L}_{\text{QCD}} + \mathcal{L}_{\text{EW}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}}. \quad (2.33)$$

Esta Lagrangiana invariante sob transformaes de fase locais do grupo de simetria $SU(3)_C \times SU(2)_L \times U(1)_Y$. O boson de Higgs interage com todas as partculas de interao fraca, incluindo ele prprio. Pode-se observar da Equao 2.32 que o acoplamento do Higgs a fermions proporcional s massas dos mesmos.

2.2 Fenomenologia das colises prton-prton

As previses do Modelo Padro e de busca por nova fsica podem ser testadas atravs de experimentos como o ATLAS [22] que faz parte do Grande Colisor de Hdrons, LHC. Desde o incio de suas operaes, o LHC tem produzido colises a uma taxa crescente e sem precedentes, atingindo 13 TeV no referencial do centro de massa do sistema prton-prton, alm de colises de ons pesados atingindo uma energia de centro de massa de 5 TeV em colises PbPb. A vantagem de se usar colises entre prtons ao invs de eltrons, deve-se massa, permitindo que o colisor alcance energias de coliso bem mais elevadas, pela reduzida perda de energia por

radiação de síncrotron, proporcional a m^{-4} . Entretanto os prótons, diferentemente dos elétrons, apresentam uma subestrutura, e como o Modelo Padrão somente descreve a interação entre as partículas elementares, a subestrutura deve ser modelada para uma completa descrição do processo de espalhamento próton-próton.

A seção de choque σ^2 , está relacionada com a probabilidade do processo de espalhamento e a ocorrência de tal processo é denominado "evento". Do ponto de vista experimental, a seção de choque é definida como

$$\sigma = \frac{R}{L}, \quad (2.34)$$

onde $R = dN/dt$ é a taxa de eventos e L a luminosidade instantânea. A seção de choque pode ser reescrita como

$$\sigma = \frac{N}{\int L dt} = \frac{N}{\mathbb{L}}, \quad (2.35)$$

onde $\mathbb{L} = \int L dt$ representa a luminosidade integrada e N o número total de eventos.

A luminosidade depende dos parâmetros do acelerador de partículas. Para um acelerador circular como o LHC, com frequência de revolução f_r e dois feixes com n_b nuvens (*bunches*) de partículas em cada, a luminosidade pode ser escrita como

$$L = f_r \frac{n_b N_1 N_2}{A}, \quad (2.36)$$

onde A é a área transversal do feixe e $N_{1,2}$ o número de partículas em cada nuvem de feixe 1 e 2, respectivamente. Assumindo que $N_1 = N_2 = N_p$ e uma gaussiana para o formato do feixe com larguras σ_x e σ_y no plano transversal, pode-se escrever a luminosidade como

$$L = f_r \frac{n_b N_p^2}{4\pi\sigma_x\sigma_y}. \quad (2.37)$$

Se a seção de choque total for um processo inelástico σ_{inel} conhecido, a luminosidade pode ser determinada como

$$L = \frac{R_{\text{inel}}}{\sigma_{\text{inel}}} = \frac{\mu n_b f_r}{\sigma_{\text{inel}}}, \quad (2.38)$$

onde μ é o número médio de interações por cruzamento de feixe.

²A seção de choque é comumente expressa em "barn", 1 b = 10^{-28} m²

Da equação 2.35, o número de eventos para um processo específico pode ser previsto:

$$N = \sigma L, \quad (2.39)$$

necessitando do cálculo teórico da seção de choque. O cálculo é complexo, dada a estrutura dos prótons, formados por três quarks de valência (uud), quarks do mar de quaisquer sabor e glúons. Esta estrutura é descrita pelas funções de distribuição partônica (PDFs - *parton distribution functions*) que devem ser experimentalmente determinadas. As PDFs, parametrizadas na forma $f_{q_i}(x_i, Q^2)$, fornecem a probabilidade de se encontrar um párton q_i com fração x_i do momento total do próton numa colisão com uma transferência de momento Q^2 no referencial de Breit, cuja definição está disponível em [27].

Cálculos em NNLO [28] demonstram que para elevados valores de x os quarks de valência up e down do próton dominam, enquanto que os glúons dominam para valores baixos de x e o mar de quarks apresenta contribuição similar à dos quarks de valência. Com o aumento de Q^2 a fração dos quarks do mar e dos glúons aumenta em comparação com os quarks de valência.

A seção de choque total para o processo $pp \rightarrow X$ é calculada a partir da seção de choque partônica, $\hat{\sigma}$, multiplicada pelas PDFs, integrada sobre o momento dos pártons com soma sobre todas as combinações partônicas, levando assim, a um estado final X [29]:

$$\sigma(pp \rightarrow X) = \sum_{i,j} \int dx_i dx_j f_{q_i}(x_i, Q^2) f_{q_j}(x_j, Q^2) \hat{\sigma}(q_i q_j \rightarrow X) \quad (2.40)$$

para pártons q_i e q_j , sob a hipótese de que a interação partônica é independente das PDFs. Este teorema de fatorização requer a introdução de uma escala de fatorização, μ_F , que separa a radiação *soft*, incluída nas PDFs das interações *hard*, que são consideradas para $\hat{\sigma}$.

A seção de choque partônica é escrita como

$$\hat{\sigma} = \int \frac{|\mathcal{M}|^2}{F} dQ, \quad (2.41)$$

onde $\mathcal{M}$ é o elemento de matriz para a transição do estado inicial para o final, F é o fluxo de partícula incidente e dQ é o fator do espaço de fase para a cinemática dada. A amplitude de probabilidade é dada por $|\mathcal{M}|^2$.

O elemento de matriz $|\mathcal{M}|$ pode ser calculado usando as densidades de Lagrangiana e teoria de perturbação. A seção de choque hadrônica para interações fortes pode ser escrita como uma expansão em série de potências da constante de acoplamento α_s :

$$\hat{\sigma} = \hat{\sigma}_0(1 + c_1\alpha_s + c_2\alpha_s^2 + \dots + c_n\alpha_s^n), \quad (2.42)$$

onde $\hat{\sigma}_0$ denota a seção de choque em ordem LO (*leading order*). De maneira similar, correções adicionais provenientes de processos eletrofracos podem ser expressos em ordens da constante de acoplamento eletrofraca. O cálculo numérico de tal série pode apenas levar em consideração um número de ordens, em LO, em NLO (*next-to leading order*), em NNLO (*next-to-next-to leading order*) e assim por diante.

Os pártons remanescentes nos prótons, que não participaram do processo de espalhamento duro, não são mais componentes estáveis, sendo denominados de eventos subjacentes (*underlying events*). Eles hadronizam e as partículas formadas sensibilizam o detector. A fragmentação dos pártons é descrita por funções de fragmentação. Estas funções não podem ser calculadas por teoria perturbativa, mas podem ser modeladas baseadas nos resultados experimentais.

Outro efeito, que torna as medidas do LHC extremamente desafiadoras, é o empilhamento (*pile-up*). Cada nuvem possui um grande número de prótons e a seção de choque total para o espalhamento duro é muito grande. Desta maneira, muitas dezenas de interações próton-próton ocorrem num único cruzamento de feixe no experimento ATLAS, que é denominada *in-time pile-up*. Outro efeito, conhecido por *out-of-time pile-up* pode ocorrer do espalhamento de produtos de um cruzamento de feixe antes ou depois daquele em consideração. Técnicas de reconstrução da E_T^{miss} específicas existem para mitigar a perda de resolução proveniente do empilhamento, que serão descritos no Capítulo 4.

A fenomenologia das colisões próton-próton é bastante complexa e todos os efeitos anteriormente discutidos em adição a muitos outros como resposta do detector, resolução, devem ser modelados por simulação de Monte Carlo que é utilizada para comparar os dados com as previsões da teoria. Esta modelagem não pode ser perfeita, de tal maneira que várias correções e incertezas devem ser aplicadas, as quais serão discutidas ao longo desta tese.

Capítulo 3

Aparato Experimental

Na Organização Europeia para Pesquisa Nuclear (CERN) existe hoje o maior aparato experimental já construído para o estudo dos constituintes básicos da matéria e de suas interações fundamentais: o Grande Colisor de Hádrons ou LHC (*Large Hadron Collider*)citeLHC. Um extenso programa de pesquisa é conduzido através do estudo dos subprodutos da colisão entre prótons ou íons acelerados a energias na escala de TeV através de sistemas de detecção instalados em 4 pontos ao redor do anel de colisão, que identificam e reconstroem as partículas produzidas. Em um desses pontos está o experimento ATLAS, um experimento de propósito geral com capacidade de explorar a física do modelo padrão, a física de íons pesados e realizar buscas por indícios de uma nova física além do modelo padrão. Este capítulo introduz as propriedades e operações do LHC e a infraestrutura do detector ATLAS, com uma breve descrição do experimento, seus subsistemas e o sistema de seleção de eventos (*trigger*).

3.1 LHC: O Grande Colisor de Hádrons

O LHC foi projetado para colidir prótons a uma energia do referencial do centro de massa $\sqrt{s} = 14$ TeV com luminosidade instantânea de $L = 10^{34} \text{cm}^2 \text{s}^{-1}$, e íons pesados a $\sqrt{s_{\text{NN}}} = 2,76$ GeV por nucleon e $L = 10^{27} \text{cm}^{-2} \text{s}^{-1}$ [30, 31]. Em operação desde 2009, o LHC tem produzido uma excepcional quantidade de dados que tem permitido explorar fenômenos cada vez mais raros e realizar buscas e medidas em espaços de fase cada vez mais restritos.

O LHC está instalado num túnel de 26,7 km de extensão na fronteira entre a Suíça

e a França, construído entre 1984 e 1989 para abrigar o colisor LEP (*Large Electron Positron*). No LHC, feixes de prótons ou íons são mantidos em órbita quase-circular graças ao campo gerado por uma sucessão de ímãs supercondutores que produzem um campo de até 8,3 T. Quadrupolos magnéticos permitem manobrar os feixes e focalizá-los nos pontos de colisão. O LHC é o último conjunto acelerador que faz parte de uma cadeia de injeção composta por vários aceleradores de partículas menores, conforme mostra a Fig. 3.1.

Figura 3.1: Representação do complexo de aceleradores do CERN [32].

Átomos de hidrogênio têm os elétrons removidos em uma fonte denominada *duo-plasmatron* e são acelerados a uma energia de 50 MeV no acelerador linear (LINAC2), que alimenta o *Proton Synchrotron Booster* (PSB) onde os prótons são acelerados a 1.4 GeV. O feixe é então injetado no *Proton Synchrotron*(PS) para uma aceleração de 25 GeV, e subsequentemente para o *Super Proton Synchrotron* (SPS) onde os prótons alcançam uma energia de 450 GeV. Os feixes de prótons são final-

mente transferidos para a canalização do LHC, onde são acelerados e condensados em nuvens de prótons para que, uma vez atingida a energia nominal sejam desviados para se cruzarem nos pontos de colisão.

Um importante parâmetro de máquina do LHC é a luminosidade instantânea das colisões, L , que depende das propriedades do feixe

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

onde N_b é o número de partículas em cada uma das n_b nuvens por feixe que circulam no túnel com uma frequência de revolução f_{rev} . O símbolo γ_r denota o fator relativístico. O formato e o foco do feixe são descritos pela sua emitância transversal ϵ_n e sua função β^* , ou comprimento focal, no ponto de colisão. O fator F leva em conta a redução geométrica da luminosidade instantânea e depende do ângulo de cruzamento de feixe θ_c e do r.m.s. dos tamanhos transversal e longitudinal, respectivamente, σ_{xy} e σ_z no ponto de interação através da seguinte relação

$$F = \left(1 + \frac{\theta_c \sigma_z}{2\sigma_{xy}} \right)^{-\frac{1}{2}}. \quad (3.2)$$

A Tabela 3.1 mostra os principais parâmetros de operação do LHC.

Tabela 3.1: Parâmetros de máquina nominais do LHC [4].

energia $\sqrt{s}$	14 TeV
espaçamento entre nuvens Δt_b	25 ns
número de partículas por nuvem N_b	1.15×10^{11}
número de nuvens por feixe n_b	2808
frequência de revolução f_r	11.2 kHz
emitância transversal do feixe ϵ_n	$3.75 \mu\text{m}$
função beta β^*	0.55 m
ângulo de cruzamento no ponto de interação θ_c	$285 \mu\text{rad}$
r.m.s transversal da nuvem σ_{xy}	$16.7 \mu\text{m}$
r.m.s longitudinal da nuvem σ_z	7.55 cm

Os quatro pontos de interação do LHC são instrumentados com detectores instalados em cavernas no subsolo. Os experimentos ATLAS (*A Toroidal LHC Apparatus*) [22] e CMS (*Compact Muon Solenoid*) [33] estão instalados em posições diametralmente opostas nos Pontos 1 e 5 do LHC, onde se produz a mais elevada luminosidade instantânea de colisão. Ambos são detectores de propósito geral que

circundam o ponto de interação provendo alta hermeticidade para detecção dos produtos das colisões. O experimento LHCb (*LHC beauty*), localizado no Ponto 8 é um espectrômetro dianteiro dedicado ao estudo de violação de carga paridade em hádrons B . O experimento ALICE (*A Large Ion Collider Experiment*) está instalado no Ponto 2 e foi construído para estudo de colisões entre íons pesados.

Outros experimentos de menor tamanho também fazem parte do LHC. Os experimentos LHCf (LHC forward) [36] e TOTEM (*TOTal Elastic and diffractive cross section Measurement*) [37] localizados respectivamente, a aproximadamente 100 m de cada lado dos experimentos ATLAS e CMS são dedicados a medidas de seção de choque pp e física difrativa.

3.1.1 Operações

O LHC iniciou o comissionamento com os primeiros feixes de prótons em 2008 e as primeiras colisões ocorreram no ano de 2010 à energia de centro de massa de 0,9 TeV. Tem então início o chamado *Run I* de coleta de dados que se estendeu até 2012. O experimento ATLAS coletou $5,08 \text{ fb}^{-1}$ e $21,3 \text{ fb}^{-1}$ de dados a energia de colisão no centro de massa de $\sqrt{s}=7 \text{ TeV}$ e 8 TeV respectivamente, dados que permitiram não só a descoberta do bóson de Higgs mas um detalhado programa de medidas de precisão no Modelo Padrão.

Em 2013 o LHC sofre a primeira parada técnica para a atualização do acelerador e de alguns subsistemas dos experimentos no LS1 (*Long Shutdown 1*). As operações do LHC reiniciaram em 2015, a uma energia de colisão do referencial do centro de massa de $\sqrt{s} = 13 \text{ TeV}$, inicialmente em baixa luminosidade, com um espaçamento entre nuvens de prótons de 50 ns, numa condição favorável à medidas de precisão devido ao baixo empilhamento¹. Após duas semanas, o espaçamento foi reduzido para 25 ns (condição nominal) e a luminosidade paulatinamente aumentada. Ao final de 2018 o ATLAS acumulou mais de 149 fb^{-1} de dados do *Run II*, chegando a uma luminosidade de pico de $L = 2 \times 10^{34} \text{ cm}^2\text{s}^{-1}$ (mais do dobro da luminosidade projetada originalmente).

¹o empilhamento é devido a mais de uma colisão simultânea por cruzamento de feixe. Em um ambiente de alta luminosidade, o empilhamento reduz a capacidade de reconstrução dos objetos físicos devido à alta ocupação em certas regiões do detector

Atualmente (2019) as operações do LHC estão suspensas para a segunda parada técnica LS2, necessária para as atualizações dos injetores do complexo do acelerador, em preparo para a fase de alta luminosidade. Assim como aconteceu durante o LS1, este período será dedicado também à atualização dos sistemas de detecção e seleção de eventos do ATLAS. O LHC planeja reiniciar as operações em 2021 para o *Run III*, operando por três anos a $\sqrt{s} = 14$ TeV mantendo uma luminosidade de pico duas vezes maior que a do projeto original. Espera-se que os experimentos do LHC tenham registrado uma luminosidade integrada de 300 fb^{-1} ao final de três anos de operações.

A terceira parada técnica LS3 terá início em 2024 e concluirá a Fase I de operações do LHC, que teve início em 2008. Durante os trinta meses seguintes, o LHC passará por profundas modificações [38]. Novos quadrupolos supercondutores (Nb_3Sn), capazes de gerar um campo de até 12 T, serão instalados nos pontos de interação do ATLAS e do CMS para focalizar os feixes. Cavidades supercondutoras (*crab cavities*) serão usadas para permitir a rotação precisa das nuvens de prótons antes da colisão, reduzindo o ângulo de cruzamento e aumentando o fator F da equação 3.2. Estas atualizações aumentarão significativamente a luminosidade integrada, fazendo da Fase II do LHC ou HL-LHC, ou *High-Luminosity LHC* um instrumento para estudo de processos muito raros sem precedentes. Espera-se que a máquina forneça, durante uma década de operações, uma luminosidade integrada superior a 4000 fb^{-1} . Neste ambiente, o número de interações simultâneas por cruzamento de feixe chegará a $\langle\mu\rangle \approx 200$.

3.2 O Experimento ATLAS

O detector (ou experimento) ATLAS (*A Toroidal LHC ApparatuS*) é um dos detectores de uso geral do LHC, projetado para coletar dados em colisões de prótons ou íons pesados em altas taxas de colisão, detectando partículas numa faixa ampla de energia, a fim de possibilitar medidas precisas do Modelo Padrão e permitir a descoberta de nova Física.

A Figura 3.2 mostra a visão geral do ATLAS, com cada um de seus subsistemas indicados. Os subsistemas do ATLAS, descritos nas próximas seções, são: o detector

interno, que tem por finalidade determinar a posição dos vértices e reconstruir traços e calcular o momento de partículas carregadas eletricamente; o solenóide supercondutor, que fica ao redor do detector interno e pode gerar um campo de até 2 T; os calorímetros, que se dividem entre os calorímetros baseados em argônio líquido (LAr, *Liquid Argon*) e o calorímetro hadrônico de telhas cintilantes (TileCal); o espectrômetro de múons, no qual a medida do momento dos múons é beneficiada pela deflexão dos múons devido ao campo gerado pelos toróides eletromagnéticos; o ATLAS conta ainda com os detectores dianteiros, que são subsistemas menores e tem como principal função a medida da luminosidade [22].

Figura 3.2: O detector do experimento ATLAS e seus subsistemas[22].

3.2.1 Sistema de coordenadas e nomenclatura

Convencionalmente o eixo Z é escolhido como a direção do feixe com a origem no ponto nominal de interação, sendo que o eixo y aponta para a superfície e o eixo x aponta para o centro do anel do LHC. O lado A do detector é definido como o lado do sentido positivo do eixo z e o lado C do sentido negativo. O ângulo azimutal ϕ e o ângulo polar θ são definidos da forma usual, porém, no lugar do ângulo θ é

mais comum o uso da pseudorrapidez, definida por $\eta = -\ln \tan(\theta/2)$ ². A distância ΔR no plano $\eta \times \phi$ é definida como $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$. A granularidade de cada subsistema do detector é o tamanho mínimo de suas células, dado em coordenadas em $\eta \times \phi$. A Fig. 3.3 mostra a cobertura em pseudorrapidez e a granularidade de cada subsistema do ATLAS.

Figura 3.3: Cobertura em η e granularidade dos subsistemas do ATLAS [22].

3.3 O sistema magnético

O sistema magnético do ATLAS é formado por um solenóide central localizado em torno do detector interno e um conjunto de toróides, sendo um deles no barril, ao redor do TileCal, e um toróide em cada uma das tampas. O sistema foi projetado tendo em vista otimizar a resolução da medida do momento dos múons, que depende inversamente da intensidade do campo aplicado e do quadrado da distância percorrida pelos múons dentro do campo, e tem um papel importante em estudos envolvendo canais de decaimento leptônico no estado final. A Figura 3.4 mostra uma visão esquemática do sistema magnético do ATLAS.

²No caso de partículas de massa nula ou negligenciável, a pseudorrapidez é equivalente à rapidez $y = (1/2) \ln[(E + p_z)/(E - p_z)]$, uma variável cuja diferença entre duas partículas é invariante mediante a uma transformação de Lorentz ao longo do eixo z .

Figura 3.4: Representação esquemática do sistema magnético do ATLAS, com o solenóide na região central, e ao redor as bobinas dos toróides do barril e das tampas. O campo gerado pelo solenóide central é alinhado com a direção do feixe e os toróides geram campos toroidais de intensidade aproximada 0,5 T e 1 T, na região central e nas tampas, respectivamente[22].

O solenóide central tem como principal atributo gerar um campo magnético axial de 2 T para curvar as trajetórias das partículas eletricamente carregadas a serem rastreadas pelo detector interno. O solenóide central foi projetado com espessura reduzida, de forma a minimizar a interferência na performance do calorímetro eletromagnético que fica ao seu redor.

Os toróides são os responsáveis pela geração do campo magnético que curva a trajetória dos múons. O toróide do barril fica alojado dentro de oito criostatos e gera um campo magnético toroidal de aproximadamente 0,5 T na região central, responsável por curvar a trajetória de múons em $|\eta| < 1,4$. Cada um dos dois toróides das tampas gera um campo próximo de 1 T na região central, curvando a trajetória de múons no intervalo $1,6 < |\eta| < 2,7$. Na região de transição $1,4 < |\eta| < 1,6$, a curvatura é causada pelo campo magnético resultante da combinação dos campos dos toróides do barril e das tampas [22].

3.4 O detector interno

O detector interno é formado pelo detector de pixels, o componente mais interno do ATLAS que é constituído de três camadas de pixels de silício, pelo *tracker* semicondutor (SCT, *semiconductor tracker*), formado por camadas de microtiras de silício, e pelo detector de radiação de transição (TRT, *transition radiation tracker*) um detector a gás baseado em tubos de deriva. O detector interno foi projetado para trabalhar num ambiente de altas doses de radiação, dispondo de uma granularidade fina o suficiente para atingir a resolução imposta pela alta taxa de colisões do LHC. A Figura 3.5 mostra a disposição de cada um dos componentes do detector interno.

Figura 3.5: Visão em corte do detector interno (região do barril e tampas)[22].

O detector de pixels e o SCT são chamados de detectores de traço de precisão, cobrindo um intervalo de pseudorapidez $|\eta| < 2,5$. O TRT fornece as coordenadas no plano $R \times \phi$ de traços na região $|\eta| < 2,0$. A combinação dos detectores de precisão com o TRT melhora o reconhecimento de padrões, contribuindo para a identificação de elétrons. A distinção entre os vértices primários e secundários e a medida do parâmetro de impacto são importantes para a identificação de quarks pesados e taus.

Durante a primeira parada técnica para manutenção e atualização foi instalada uma nova camada no detector de pixels (IBL, *Insertable B-Layer*) logo abaixo da

camada mais interna, *B-Layer*, com um raio de aproximadamente 3,2 cm ao redor do tubo do feixe. A instalação dessa nova camada foi motivada pela degradação da *B-Layer* original e foi possível graças ao raio menor do novo tubo do LHC [23].

Os traços são reconstruídos a partir de uma série de interações de partículas eletricamente carregadas com o ID [24]. As colisões produzindo múltiplos traços são identificadas como vértices [25] reconstruídos através dos traços. A Figura 3.6 mostra a ilustração de eventos coletados pelo ATLAS em junho de 2015 a partir de colisões de prótons com $\sqrt{s} = 13$ TeV, com destaque para os traços que são representados pelas linhas no centro na figura.

3.5 Os calorímetros

O sistema de calorimetria do ATLAS compreende três calorímetros que utilizam Argônio líquido como meio ionizante, o calorímetro eletromagnético (EM, *electromagnetic calorimeter*), os calorímetros hadrônicos das tampas (HEC, *hadronic end-cap*) e os calorímetros dianteiros (FCal, *forward calorimeter*), além do calorímetro de telhas cintilantes, *Tilecal*, que usa a tecnologia de cintiladores plásticos. Ambos os calorímetros são divididos em camadas alternadas de material absorvedor e material detector que fazem com que as partículas deixem depósitos na forma de um chuva. A Figura 3.7 mostra a disposição dos calorímetros do ATLAS.

O calorímetro eletromagnético é dividido em um barril central que cobre a região $|\eta| < 1,475$ e dois módulos nas tampas cobrindo $1,375 < |\eta| < 3,2$. O calorímetro eletromagnético usa chumbo como material absorvedor e tem o formato de acordeon, que possibilita uma cobertura azimutal completa. Tanto no barril como nas tampas, o calorímetro eletromagnético é segmentado em três camadas, havendo ainda uma fina camada na frente do calorímetro (PS, *presampler*) para corrigir a perda de energia causada pela presença de material inativo antes da região de amostragem.

Nas regiões imediatamente mais externas aos calorímetros eletromagnéticos das tampas encontram-se os calorímetros hadrônicos das tampas, cobrindo o intervalo $1,5 < |\eta| < 3,2$. Cada peça tem a forma de um disco composto por 32 módulos em forma de cunha. Placas de cobre são usadas como metal absorvedor.

Os calorímetros dianteiros compartilham os mesmos criostatos com os HEC,

Figura 3.6: Ilustração de eventos de colisões entre prótons com $\sqrt{s} = 13$ TeV coletados pelo ATLAS em junho de 2015, com destaque para os traços reconstruídos pelo ID, no centro da figura. [26].

coabrindo o intervalo $3,1 < |\eta| < 4,9$. Cada lado do FCal consiste de um disco dividido em três módulos, sendo que o primeiro usa como metal absorvedor o cobre e é otimizado para medidas eletromagnéticas, enquanto os outros dois usam tungstênio como absorvedor e são otimizados para medidas hadrônicas.

No TileCal os cintiladores plásticos estão dispostos perpendicularmente à direção do feixe. O TileCal é dividido em um barril central ao redor do calorímetro eletromagnético, na região $|\eta| < 1,0$, e dois barris extendidos com cobertura $0,8 < |\eta| < 1,7$. O TileCal é segmentado em três camadas e os sinais são lidos por fibras óticas deslocadoras de comprimento de onda. Uma descrição detalhada do sistema de calorimetria do ATLAS pode ser encontrada em [22].

3.6 O espectrômetro de múons

O espectrômetro de múons é o componente mais externo do ATLAS, composto por câmaras de traço de alta precisão e pelos toróides eletromagnéticos, contando ainda com um sistema de *trigger* independente.

O espectrômetro de múons possui dois tipos de câmaras destinadas à medida precisa de traços. As câmaras de tubos de deriva monitoradas (MDT, *monitored*

Figura 3.7: Visão em corte do sistema de calorimetria do experimento ATLAS[22].

drift tubes) medem as coordenadas dos traços na direção principal de curvatura do campo magnético, cobrindo o intervalo $|\eta| < 2,7$. As câmaras de tiras catódicas (CSC, *cathod strip chambers*) são câmaras proporcionais multifios com os cátodos segmentados em tiras, otimizadas para suportar altas taxas de fundo de radiação e tem a função de medir traços com precisão no intervalo $2,0 < |\eta| < 2,7$.

O sistema de *trigger* é servido pelas câmaras de placas resistivas (RPC, *resistive plate chambers*) na região do barril e pelas câmaras de *gap* finas (TGC, *thin gap chambers*) nas tampas, ambas com as funções de identificar as colisões entre os feixes, fornecer os limiares de momento transverso e medir a coordenada da trajetória dos múons ortogonal àquela medida pelas câmaras de precisão [22].

3.7 Os detectores dianteiros

O LUCID (*Luminosity measurement using Cerenkov Integrating Detector*) é o principal detector responsável pelo monitoramento da luminosidade do ATLAS. O LUCID é localizado a ± 17 m do ponto de interação e detecta colisões inelásticas entre prótons em direções próximas da direção do feixe. No LS1 o LUCID foi substituído, já que o sistema original poderia não suportar o aumento da luminosidade instantânea do LHC, e a nova versão começou a operar com sucesso em 2015.

O ALFA (*Absolut Luminosity For ATLAS*) fica localizado a ± 240 m do ponto de

interação e consiste de rastreadores de fibra cintilante localizados dentro de *roman pots*, que são volumes de detector separados do vácuo do acelerador por uma fina janela, mas conectados ao tubo do feixe, projetados para se aproximarem até a 1 mm de distância do feixe.

O ZDC (*Zero-Degree calorimeter*) localiza-se a ± 140 m do ponto de interação, seus módulos são constituídos de camadas de hastes de quartzo alternadas com chapas de tungstênio e tem como principal objetivo detectar nêutrons e fótons dianteiros em colisões de íons pesados na região $|\eta| > 8,2$. Durante as colisões de prótons o ZDC é removido para evitar danos devido às elevadas doses de radiação. Informações mais detalhadas dos detectores dianteiros estão disponíveis em [22].

3.8 *Trigger* e aquisição de dados

A Figura 3.8 mostra uma visão esquemática do atual sistema de *trigger* e aquisição de dados (TDAQ) do ATLAS, que tem a função de selecionar eventos a partir de sinais eletrônicos do detector obtidos a uma frequência de colisões de 40 MHz e salvar as informações relevantes para posterior reconstrução e análise.

O sistema de *trigger* é dividido em três níveis. O primeiro nível (L1), é baseado em *hardware* e procura por assinaturas que são comparadas com a informação do *trigger* central caso sejam identificadas. Os eventos aceitos são lidos e processados pelos drivers de leitura (RODs, *Read-Out Drivers*) e enviados ao sistema de leitura (ROS, *Read-Out System*). O L1 também define pequenos pacotes de dados com informação que considera relevante, chamados regiões de interesse (ROIs, *Regions-Of-Interest*). A informação é transmitida ao L2, o segundo nível de *trigger*, a uma taxa que pode chegar a 75 kHz.

O *trigger* de alto nível (HLT, *High-Level trigger*) engloba o L2 e o terceiro nível, chamado de *Event Filter* (EF), e é implementado via *software*. O L2 analisa as ROIs definidas pelo L1, mas já usando a granularidade completa do detector para as ROIs. O L2 é configurado para aceitar aproximadamente 10% dos eventos, sendo que os eventos aceitos são completamente reconstruídos antes de serem transmitidos ao EF. O L2 entrega os eventos ao EF a uma taxa de aproximadamente 2 kHz. O EF usa algoritmos mais sofisticados para reconstruir os eventos aceitos, refinando os re-

Figura 3.8: Esquema simplificado do sistema de *trigger* do ATLAS[22].

sultados do L2. O EF entrega os dados brutos para serem gravados e posteriormente reconstruídos e analisados a uma taxa de aproximadamente 200 Hz.

3.8.1 Sistema de *trigger* dos calorímetros

O sistema de *trigger* existente é baseado no esquema de torres de *trigger*, que consiste na soma de depósitos de energia através de camadas longitudinais dos calorímetros em uma área de $\Delta\eta \times \Delta\phi = 0,1 \times 0,1$. Os sinais analógicos das torres são digitalizados pelo pré-processador que também usa um filtro para fazer a associação dos sinais às correspondentes colisões e calcula a energia transversa³ a ser usada nos algoritmos com base em uma tabela *look-up*. A informação é então transmitida paralelamente ao *Cluster Processor*, que identifica candidatos a elétrons, fótons ou taus com energia transversa acima de um limiar programado e satisfazendo critérios

³A energia transversa de uma partícula de massa m e momento transverso p_T é dada por $E_T = \sqrt{m^2 + p_T^2}$.

de isolamento, e ao *Jet/Energy-sum Processor*, que usa somas do *trigger* de jatos de área $\Delta\eta \times \Delta\phi = 0,2 \times 0,2$ para identificar jatos e calcular a soma escalar de energia transversa e a energia transversa faltante. Esses dois processadores enviam as informações ao processador de *trigger* central, onde são então tomadas as decisões de aceitar ou não os eventos. As informações dos eventos aceitos são então enviadas ao sistema de aquisição de dados, contendo dados úteis para o diagnóstico do sistema de *trigger* e do LHC, além de dados necessários para a construção das regiões de interesse a serem lidas pelo HLT.

Capítulo 4

Reconstrução de elétrons e da energia transversa faltante no ATLAS

Este trabalho se concentra na identificação das partículas e grandezas cinemáticas associadas à produção de bósons W através de seu decaimento no canal elétron-neutrino. Neste capítulo, serão discutidos os métodos empregados no experimento ATLAS para a identificação direta da produção de elétrons e indireta dos neutrinos, através da energia transversa faltante ($\vec{E}_T^{\text{miss}}$).

4.1 Definição da energia transversa faltante

Sendo desconhecidos os momentos lineares dos partons interagentes em colisões de prótons, é conveniente o uso do momento transverso, pois, pela conservação do momento e dado que as partículas iniciais tem seu movimento restrito ao longo do eixo z (direção do feixe de prótons), a soma vetorial dos momentos no plano trnasverso deve se anular. Quando o momento transverso total medido pelo detector não é zero pode ser um indício de produção de uma partícula que interage fracamente no estado final. No Modelo Padrão essas partículas são os neutrinos, mas outros modelos e extensões do Modelo Padrão possuem outras partículas fracamente interagentes, de forma que a determinação do $\vec{E}_T^{\text{miss}}$ está presente na maioria das buscas de processos envolvendo nova física. A reconstrução da energia transversa faltante é afetada por partículas mal reconstruídas, que escapam da região de aceitação do detector ou não são reconstruídas por alguma falha do detector. Dessa forma a energia transversa

faltante também desempenha um importante papel como indicador do desempenho do detector e dos processos de reconstrução de eventos.

O vetor energia transversa faltante $\vec{E}_T^{\text{miss}}$ é definido como o oposto da soma dos momentos transversos medidos pelo detector. Essa soma leva em conta tanto os objetos físicos, provenientes de espalhamento *hard*, que são elétrons, múons, taus decaindo em hádrons e jatos, como os termos *soft*, provenientes de eventos subjacentes e de radiação emitida pelos objetos produzidos no espalhamento *hard* [39].

$$\vec{E}_T^{\text{miss}} = - \sum \vec{p}_{T,\text{hard}} - \sum \vec{p}_{T,\text{soft}} \quad (4.1)$$

A $\vec{E}_T^{\text{miss}}$ tem a magnitude E_T^{miss} e a direção ϕ dados por:

$$E_T^{\text{miss}} = \sqrt{(E_x^{\text{miss}})^2 + (E_y^{\text{miss}})^2} \quad (4.2)$$

$$\phi = \arctan \frac{E_x^{\text{miss}}}{E_y^{\text{miss}}} \quad (4.3)$$

Uma grandeza útil é a soma escalar $\sum E_T$ das magnitudes dos momentos transversos medidos de todos os objetos *hard* e *soft*,

$$\sum E_T = \sum p_{T,\text{hard}} + \sum p_{T,\text{soft}}, \quad (4.4)$$

que serve como indicador da atividade de eventos em análises físicas e estudos de performance na reconstrução da E_T^{miss} . A Figura 4.1 apresenta um esquema de um evento com energia transversa faltante.

4.2 Algoritmos usados no cálculo da E_T^{miss}

Os algoritmos de reconstrução de E_T^{miss} usam tanto a informação dos traços do detector interno como a informação dos calorímetros, separadamente ou em conjunto. O método baseado nos calorímetros chamado CST (*calorimeter-based soft term*) E_T^{miss} [40] usa como termo *soft* depósitos de energia nos calorímetros não associados a nenhum objeto *hard*. Esse método tem a desvantagem de ser sensível ao empilhamento, responsável por uma contribuição extra ao termo *soft*. Um método baseado puramente nas informações do traço (*Track E_T^{miss}*), apresenta grande resalto em condições de alto empilhamento, dado que os traços podem ser associados

Figura 4.1: Ilustração da topologia de um evento com energia transversa faltante [39].

aos vértices, mas tem a aceitação limitada pela região de cobertura do detector interno e, principalmente, pelo fato de partículas neutras não deixarem traços no ID. O algoritmo TST (*track soft term*) E_T^{miss} [41], que usa o termo *soft* determinado através da informação dos traços, e que também usa as informações provenientes dos calorímetros, apresenta um bom compromisso entre os dois métodos anteriores e foi tomado como método padrão na atual fase de reconstrução e análise de dados, como será discutido nos Caps. 5 e 6. Um método alternativo é calcular a E_T^{miss} a partir do recuo hadrônico, que é definido como a soma vetorial dos momentos transversos dos quarks e glúons iniciais, que num processo Drell-Yan equivale ao momento transverso do bóson.

4.2.1 CST E_T^{miss}

A reconstrução de energia transversa faltante baseada nos calorímetros inclui a contribuição de elétrons, fótons, taus, jatos e de células que não são associadas a nenhum desses objetos. A contribuição dos múons pode vir dos calorímetros ou do espectrômetro de múons, dependendo das circunstâncias detalhadas mais adiante. Também são usados traços de baixo p_T para recuperar a informação de partículas de baixo p_T perdidas pelos calorímetros ou a informação de múons em regiões não

cobertas pelo espectrômetro de múons.

As células dos calorímetros são associadas aos objetos “pais” de alto p_T obedecendo à ordem: elétrons, fótons, taus, jatos e múons. A soma dos termos ocorre sobre todas as células associadas aos objetos específicos num intervalo de pseudorapidez de $|\eta| < 4, 5$.

Devido à alta granularidade dos calorímetros é importante suprimir o ruído e limitar a soma da energia transversa faltante às células que contêm uma contribuição significativa do sinal. Esse problema é tratado usando apenas células pertencendo a *clusters* topológicos tridimensionais, que são formados a partir de células com uma energia quatro vezes maior que o limiar de ruído daquela célula e são então expandidos nas três dimensões espaciais e, depois da expansão, uma última camada de células é adicionada ao *cluster*, sendo usados diferentes algoritmos na formação de *clusters* associados a elétrons e fótons [42] e demais objetos físicos [43].

As contribuições referentes a elétrons, fótons e taus são reconstruídas de células em *clusters* associados aos respectivos objetos. A energia faltante dos jatos é reconstruída a partir de *clusters* de jatos calibrados com $p_T > 20$ GeV, sendo incluso um termo devido a jatos originários de espalhamento *soft* que são reconstruídos a partir de células de *clusters* com p_T entre 7 GeV e 20 GeV. Os termos *soft* são calculados a partir de *clusters* não associados aos objetos reconstruídos e o termo referente aos múons é descrito a seguir. Todos os objetos que contribuem para a E_T^{miss} são calibrados de forma independente.

A contribuição dos múons

O termo referente aos múons é calculado a partir do momento de traços de múons reconstruídos num intervalo de $|\eta| < 2, 7$, sendo que na região $|\eta| < 2, 5$ é exigido que os múons reconstruídos no espectrômetro de múons tenham um traço correspondente no detector interno para rejeitar falsos múons.

Dada a distância $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ entre um múon e um jato, a energia que o múon deposita no calorímetro pode ser contabilizada ou não. Se $\Delta R > 0, 3$ o momento do múon é determinado com base apenas nas informações do detector interno e do espectrômetro de múons e a energia depositada no calorímetro não

é considerada para evitar que seja contabilizada duas vezes. Se o múon não está isolado do jato, ou seja $\Delta R < 0,3$, a energia do múon depositada no calorímetro não pode ser distinguida da energia do jato, então é somada à energia medida pelo espectrômetro de múons. Em casos nos quais há uma falta de concordância entre a informação do espectrômetro de múons e a medida combinada com informações do ID, a energia do múon deixada no calorímetro é estimada por um método paramétrico e subtraída da energia da medida combinada [44].

Os múons que atravessam a região $2,5 < |\eta| < 2,7$ estão fora da cobertura do detector interno e o seu momento é determinado exclusivamente pelo espectrômetro de múons. Múons podem ser perdidos em regiões inativas, em torno de $|\eta| = 0$ e $|\eta| \approx 1,2$, em tais casos traços do detector interno podem ser extrapolados até o espectrômetro de múons para recuperar a sua contribuição à E_T^{miss} .

Calibração

Os objetos *hard* que entram na reconstrução da energia transversa faltante são calibrados individualmente seguindo recomendações próprias com exceção dos fótons, pois estes apresentam baixa pureza, uma vez que não existe informação no detector interno de traços.

A calibração do termo *soft* é importante em análises envolvendo bósons W e Z , pois pode influenciar no p_T dos bósons em eventos com baixa multiplicidade de partículas. A contribuição de assinaturas de baixo p_T ao termo *soft* da E_T^{miss} é aprimorada por meio de um algoritmo que usa traços para recuperar a contribuição de partículas que não atingem o calorímetro ou não alimentam um *cluster* topológico. Traços com $p_T > 400$ MeV que passaram por uma seleção são extrapolados até a segunda camada do calorímetro eletromagnético e são associados a objetos ou *clusters*. Se um traço não é associado a um *cluster* ou objeto reconstruído, seu momento transversal é contabilizado no cálculo da E_T^{miss} . Se o traço é associado ao *cluster*, a energia do *cluster* é desconsiderada e a energia (ou p_T) associada ao traço é usada em seu lugar. No caso de mais de um *cluster* associado ao traço, apenas a contribuição do que tem maior energia é substituída pela energia do traço.

4.2.2 Algoritmos baseados em traços

Os algoritmos *Track* E_T^{miss} e TST E_T^{miss} usam traços que passam por critérios de seleção para assegurar a qualidade da energia transversa faltante reconstruída. Os traços devem ter um momento transverso mínimo de 0,4 GeV e estar dentro da aceitação do ID, $|\eta| < 2,5$. O vértice primário em cada evento é tomado como o que apresenta o maior valor do somatório de p_T^2 sobre os traços associados ao vértice. Também é exigido que o parâmetro de impacto transversal d_0 , que é o deslocamento do vértice na direção perpendicular ao feixe, seja menor que 1,5 cm com uma significância $d_0/\sigma(d_0) < 3$ e o parâmetro de impacto longitudinal z_0 , o deslocamento do vértice na direção do feixe, seja menor que 1,5 cm, dessa forma garantindo que o vértice reconstruído seja consistente com a geometria dos traços [41].

No *Track* E_T^{miss} a energia transversa faltante é reconstruída a partir dos traços selecionados, porém o p_T do elétron é determinado mais precisamente pelo calorímetro eletromagnético, dado que o elétron pode interagir com o ID, portanto o p_T do traço associado ao elétron é substituído pelo p_T do elétron medido no calorímetro. O termo *soft* é reconstruído a partir dos traços que não correspondem a um elétron ou múon. Usar contribuições provenientes apenas de traços originários de um vértice primário, porém ignorando o p_T das partículas reconstruídas, aumenta a estabilidade contra o empilhamento, mas leva a uma baixa resolução da E_T^{miss} devido a não inclusão do momento de partículas neutras, assim como de partículas fora do intervalo de pseudorrapidez $|\eta| < 2,5$.

O TST E_T^{miss} inclui sinais de partículas neutras usando *clusters* topológicos dos calorímetros na reconstrução do termo *soft*. O desempenho na reconstrução é inferior ao método baseado puramente nos traços por causa de uma dependência residual do empilhamento. Na determinação do termo *soft*, os traços que estão superpostos com elétrons ou fótons a uma distância $\Delta R < 0,05$ ou com taus decaindo em hádrons a uma distância $\Delta R < 0,2$ são excluídos. Traços associados a jatos que usam a técnica de associação "fantasma" [45] também são excluídos por terem incertezas no momento que podem ser maior que 40% ou não terem um depósito correspondente no calorímetro. Traços do ID associados a múons reconstruídos a partir das informações

combinadas do ID e do espectrômetro são substituídos pelos traços combinados [39, 41].

4.2.3 Método do recuo hadrônico

Um método alternativo de calcular a E_T^{miss} faz uso do chamado recuo hadrônico [46] e tem sido aplicado de forma bem sucedida em análises do primeiro período de operação do LHC, como por exemplo [47, 48, 49]. A seguir são mencionados alguns algoritmos de reconstrução do recuo hadrônico disponíveis em [50].

Nos eventos de produção de bósons vetoriais em colisões de prótons o momento transversal do bóson em questão não é exatamente nulo, devido à radiação de estado inicial de quarks e glúons. O recuo hadrônico é então definido como o vetor oposto ao momento transversal do bóson. Particularmente em eventos de decaimentos leptônicos de bósons $W^\pm$, a energia transversal faltante pode ser obtida a partir do momento transversal do lépton $\vec{p}_T^l$ e do recuo hadrônico $\vec{H}R$ pela simples soma vetorial

$$\vec{E}_T^{\text{miss}} = - \left(\vec{H}R + \vec{p}_T^l \right) \quad (4.5)$$

Entre os algoritmos usados na reconstrução do recuo hadrônico, um dos primeiros é baseado nas informações dos calorímetros. Este algoritmo leva em conta todos os depósitos de energia nos calorímetros. Um problema desse método é a contabilidade dupla da energia de elétrons e múons, que é minimizado pelo método através da exclusão de *clusters* dentro de um cone de raio definido ao redor do lépton, porém a remoção desses *clusters* tem o efeito de degradação do recuo hadrônico. Esses problemas podem ser amenizados por uma compensação, baseada no isolamento do elétron.

Um segundo método é baseado nas informações dos traços, e usa a soma vetorial dos momentos transversos de todos os traços na determinação do recuo hadrônico. Os principais problemas desse algoritmo são a perda da informação de partículas fora da região de aceitação do detector interno ($|\eta| < 2,5$) e das partículas de carga elétrica neutra. Em comparação com o método baseado na informação calorimétrica esse método tem a resolução menos dependente de interações de empilhamento, porém apresenta uma maior degradação na resolução da $\sum E_T$.

Existe ainda um terceiro método que combina as informações de *clusters* com a informação dos traços, cujo princípio é excluir *clusters* associados a traços ligados a vértices primários que não são consistentes com vértices de produção de bósons vetoriais. A dependência da resolução em relação ao empilhamento é reduzida em comparação com o método baseado puramente nas informações calorimétricas, assim como uma melhora global na resolução também é observada.

4.3 Resultados de estudos de desempenho

Utilizando dados de colisões entre prótons à energia de centro de massa de 7TeV, o experimento ATLAS realizou diversos estudos sobre a reconstrução da energia transversa faltante[40]. Nesse estudo, realizado pelo ATLAS em 2010, foram usadas diversas amostras, sendo usada uma amostra de eventos de jatos de alto p_T e decaimento leptônico de bósons Z e $W^\pm$ chamados de eventos de tendenciosidade mínima na referência. Para amostras de bósons Z decaindo em pares de léptons espera-se que o valor médio da E_T^{miss} seja zero, já que não há partículas fracamente interagentes no estado final, logo qualquer desvio significa que há uma tendenciosidade. A Figura 4.6 mostra a resolução em função de $\sum E_T$ para as diversas amostras usadas, tanto de dados coletados como de Monte Carlo. A resolução foi tomada como a largura da distribuição combinada de E_x^{miss} e E_y^{miss} , exceto para as simulações do decaimento leptônico de $W^\pm$, onde foi considerada a largura da distribuição combinada de $E_x^{\text{miss}} - E_x^{\text{miss, True}}$ e $E_y^{\text{miss}} - E_y^{\text{miss, True}}$. Os gráficos foram ajustados a uma função da forma $\sigma = k\sqrt{\sum E_T}$. Os resultados mostraram uma boa concordância entre os dados e a simulação.

Já usando os dados coletados em 2015 a partir de colisões de prótons a $\sqrt{s} = 13$ TeV e espaçamento entre grupos de prótons de 50 ns, o ATLAS revisitou os estudos de desempenho da reconstrução de E_T^{miss} usando o algoritmo TST E_T^{miss} , comparando os dados coletados com dados de Monte Carlo [41]. Esse trabalho se baseou nos canais $Z \rightarrow \mu^+\mu^-$ e $W^\pm \rightarrow e^\pm\nu_e$. A Figura 4.3 compara a resolução em função da $\sum E_T$ obtida para o canal $Z \rightarrow \mu^+\mu^-$ entre os dados coletados e os dados simulados. Nesse estudo a resolução foi definida como o RMS das distribuições combinadas de E_x^{miss} e E_y^{miss} . Embora os dados experimentais não contenham uma

Figura 4.2: Resolução em termos da $\sum E_T$ para os dados coletados (a) , e os dados simulados (b). Os valores do parâmetro de ajuste k estão em $\text{GeV}^{1/2}$ [40].

quantidade significativa de eventos, os resultados mostraram concordância razoável com a simulação.

Um estudo comparativo dos três algoritmos [51] foi realizado pelo grupo de desempenho de E_T^{miss} e jatos do ATLAS. Esse estudo usou amostras de simulação dos processos $Z \rightarrow l^+l^-$, $W^\pm \rightarrow l^\pm\nu$ e produção de pares $t\bar{t}$, onde l pode ser elétron ou múon, em colisões de prótons com energia de centro de massa de 13 TeV e espaçamento temporal entre as colisões de 25 ns. A Figura 4.4 mostra a comparação da distribuição da componente x da energia transversa faltante para a amostra de $Z \rightarrow \mu^+\mu^-$ e da $\sum E_T$ entre os três algoritmos. Na distribuição de E_x^{miss} o TST E_T^{miss} apresenta o pico mais bem definido, enquanto o *Track* E_T^{miss} apresenta uma cauda maior por causa da insensibilidade a partículas neutras e o CST E_T^{miss} não tem um pico tão bem definido quanto ao TST E_T^{miss} mas ambos apresentam caudas semelhantes. A distribuição de $\sum E_T$ confirma uma maior atividade no calorímetro.

A Figura 4.5 mostra a comparação da linearidade entre os três algoritmos usados para o canal $Z \rightarrow \mu^+\mu^-$, sendo a linearidade definida como:

$$\text{linearidade} = \left\langle \frac{E_T^{\text{miss}} - E_T^{\text{miss, True}}}{E_T^{\text{miss, True}}} \right\rangle, \quad (4.6)$$

onde $E_T^{\text{miss, True}}$ é a energia transversa calculada a partir das partículas antes de passarem pela simulação do detector. O valor esperado é zero, mas uma pequena tendenciosidade é esperada, dado que a E_T^{miss} é sempre positiva e a resolução é finita. O algoritmo *Track* E_T^{miss} subestima substancialmente a escala de E_T^{miss} quando os

Figura 4.3: Distribuições da resolução como função de $\sum E_T$ em eventos de $Z \rightarrow \mu^+ \mu^-$ [41].

Figura 4.4: Distribuição da componente x da $\vec{E}_T^{\text{miss}}$ e da $\sum E_T$ para o canal $Z \rightarrow \mu^+ \mu^-$ para cada um dos algoritmos usados [51].

jatos são incluídos, uma vez que a contribuição das partículas neutras dentro dos jatos são ignoradas.

A Figura 4.6 mostra a resolução (quantificada como o RMS da distribuição de $E_{x,y}^{\text{miss}}$ em termos da $\sum E_T$ para os canais de Z e W . Para ambos os bósons, o CST E_T^{miss} apresenta um RMS crescente com relação à atividade de eventos, sendo que esse efeito é menos pronunciado nos algoritmos baseados em traço. Dado que a resolução dos jatos influencia mais o TST E_T^{miss} que o *Track* E_T^{miss} , o aumento do RMS com a atividade pode ser explicado pelo efeito da resolução dos jatos. Conforme

Figura 4.5: Comparação da linearidade entre os algoritmos envolvidos sem incluir jatos, esquerda, e incluindo jatos, direita[51].

a atividade aumenta, o comportamento do TST E_T^{miss} tende ao comportamento do CST E_T^{miss} , já que a contribuição dos jatos passa a ser dominante.

4.4 Reconstrução de elétrons

A reconstrução de elétrons começa pela identificação de *clusters* topológicos no calorímetro eletromagnético formados por 3×5 células de tamanho $0,025 \times 0,025$ em unidades de $\Delta\eta \times \Delta\phi$, sendo cada um deles alimentado por uma torre de *trigger* de energia transversa superior a 2,5 GeV, identificada por um algoritmo de janela móvel.

O passo seguinte é associar um traço reconstruído no detector interno com *clusters* do calorímetro usando a distância no plano $\eta \times \phi$ entre a posição do traço e o baricentro do *cluster*. Os candidatos a traços de elétrons são reconstruídos usando um padrão de reconhecimento que supõe que a perda de energia no material é compatível com a de um pión ou, caso essa hipótese não seja satisfeita mas o traço pode ser estendido até uma região de interesse ao redor do baricentro do *cluster*, é usado um padrão de reconhecimento que faz uso da hipótese da perda de energia ser de um elétron. Após construídos, os traços são redefinidos através do ajuste de seus parâmetros aos de um traço esperado para um elétron através de um algoritmo otimizado [53].

Um elétron é reconstruído se pelo menos um traço do detector interno for asso-

(a)

(b)

(c)

(d)

Figura 4.6: Resolução da E_T^{miss} em termos da $\sum E_T$ para os processos $Z \rightarrow \mu^+ \mu^-$, (a,b), e $W^\pm \rightarrow \mu^\pm \nu$ (c,d). Os resultados foram separados em eventos que não incluem jatos (a,c), e incluindo jatos (b,d)[51].

ciado a um *cluster* do calorímetro eletromagnético. Caso mais de um traço satisfaça às condições de associação com um *cluster*, um algoritmo seleciona o melhor candidato a elétron usando como principais parâmetros a distância entre o traço e o *cluster* no plano $\eta \times \phi$, o número de depósitos de energia no detector de pixels e a presença de um depósito na primeira camada do detector de pixels [54]. Os demais *clusters* são considerados como sendo de fótons. Finalmente o *cluster* é reformulado usando 3×7 ou 5×5 torres de células no barril central ou nas tampas do calorímetro eletromagnético, respectivamente, e a energia do *cluster* é calibrada usando métodos de análise multivariáveis (MVA) [55].

4.5 Identificação de elétrons

Com o objetivo de discriminar os candidatos a elétrons entre elétrons de sinal e fundos identificados incorretamente como elétrons, tais como jatos ou fótons, são usados algoritmos que fazem o uso de informações dos traços e dos diversos componentes do detector que participam da reconstrução de elétrons.

O algoritmo recomendado pelos grupos de desempenho para as análises físicas é um algoritmo multivariáveis baseado em verossimilhança. Uma função de verossimilhança do sinal $\mathfrak{L}_S$ e outra dos fundos $\mathfrak{L}_B$ são construídas a partir das densidades de probabilidade do sinal $P_{S,i}(x_i)$ ou dos fundos $P_{B,i}(x_i)$ das diversas variáveis discriminantes x_i , e um discriminante $d_{\mathfrak{L}}$, dado por

$$d_{\mathfrak{L}} = \frac{\mathfrak{L}_S}{\mathfrak{L}_S + \mathfrak{L}_B} \quad (4.7)$$

é calculado a partir dessas funções de verossimilhança dadas por

$$\mathfrak{L}_{S(B)} = \prod_{i=1}^n P_{S(B),i}(x_i). \quad (4.8)$$

Além das variáveis usadas na construção das funções de verossimilhança, cortes simples na razão entre a energia do *cluster* e o momento do traço, uma variável dependente da largura da cascata no calorímetro eletromagnético, a diferença entre o ângulo azimutal do *cluster* na camada central do calorímetro e do traço e as variáveis que contabilizam o número de depósitos do traço no detector interno. As demais variáveis discriminantes carregam informações do calorímetro hadrônico, de

cada uma das camadas do calorímetro eletromagnético, das propriedades do traço, do TRT e da associação do traço com o *cluster* e estão detalhadas em [56].

Os algoritmos de identificação são otimizados com base em simulações, principalmente dos processos $Z \rightarrow e^+e^-$ e $J/\psi \rightarrow e^+e^-$. Contudo é preciso aplicar correções nas distribuições das variáveis discriminantes das simulações baseadas no Geant 4, porque algumas dessas distribuições possuem uma maior largura ou são mais próximas das distribuições esperadas para os fundos, efeito de imprecisões de modelagem da interação das partículas com o meio de detecção e material inativo [57]. Deve-se notar que essas correções são diferentes daquelas aplicadas nas análises físicas, que são pesos aplicados nas simulações para se reproduzir as mesmas eficiências dos dados.

Os elétrons podem ser classificados de três formas distintas, de acordo com o nível de rejeição dos fundos. A classificação com maior rejeição é chamada *tight* e a de menor rejeição *loose*, e ainda há uma classificação intermediária chamada *medium*. Todos os elétrons classificados como *tight* também passam pelos critérios de seleção dos elétrons classificados como *medium*, que por sua vez passam pela seleção *loose*. A figura 4.7 mostra a eficiência por energia transversa de uma amostra simulada de $Z \rightarrow e^+e^-$ e de uma amostra simulada de eventos de dois jatos (dijatos).

4.6 Isolamento de elétrons

Normalmente os léptons verdadeiros, ou seja, aqueles que não são outras partículas identificadas como léptons, estão isolados, sem muita atividade ao seu redor. Portanto é importante definir uma energia de isolamento em seu redor de modo a evitar falsos léptons nas análises de interesse. As energias de isolamento são definidas com base no traço, que são menos afetadas pelo empilhamento de eventos, ou com base no calorímetro.

4.6.1 As variáveis de isolamento

As variáveis do tipo *etcone* são calculadas pela soma simples da energia transversa depositada em células do calorímetro dentro de um cone no plano $\eta \times \phi$ de raio R_0 centrado na direção do lépton, onde $R_0 \in \{0, 2, 3, 4\}$. Essas variáveis não

Figura 4.7: Eficiências na identificação de elétrons de uma amostra $Z \rightarrow e^+e^-$ (a) e de uma amostra de dijetos (b). Para a amostra de $Z \rightarrow e^+e^-$, com elétrons reconstruídos a partir de elétrons verdadeiros, a eficiência aumenta com a E_T , enquanto diminui para a amostra de dijetos, da qual os elétrons são jatos identificados incorretamente como elétrons[52].

tiveram bom desempenho no Run 1 em relação à resiliência ao empilhamento e à concordância entre dados coletados e simulação.

Uma melhoria em relação às variáveis *etcone* foi alcançada com a coleta da energia transversa depositada em *clusters* topológicos. A direção do elétron ou fóton é dada pela posição do *cluster* associado, usado para reconstruir sua energia. Toda a energia positiva dos *clusters* nos quais o baricentro está dentro de um cone centrado na direção do lépton ou fóton é somada e armazenada em variáveis chamadas *topoetcone*.

As variáveis de isolamento baseadas no traço são chamadas de *ptcone* e são obtidas somando-se o momento transversal dos traços dentro de um cone de raio R_0 centrado na direção do traço do lépton ou do fóton. No isolamento baseado no calorímetro, um cone de raio inferior a 0,2 não pode ser usado, e em ambientes com objetos próximos da direção do lépton ou fóton, tal isolamento torna-se inadequado. Uma variável baseada no traço com o cone de tamanho variável pode então ser usada, sendo denominada *ptvarcone*. Para essa variável, o raio R do cone torna-se menor conforme o aumento do momento transversal:

$$R = \min \left(\frac{k_T}{p_T}, R_{max} \right), \quad (4.9)$$

onde k_T é uma constante fixa em 10 GeV e R_{max} é o tamanho máximo do cone, entre 0,2 e 0,4. Essas definições serão importantes nos estudos sobre os critérios utilizados no isolamento de elétrons, alvo de discussão das análises apresentadas nos Cap. 5 e 6.

Capítulo 5

Medidas da produção $W^\pm \rightarrow e^\pm \nu$ em colisões próton-próton a $\sqrt{s} = 13 \text{ TeV}$

No ano de 2015 o LHC inaugurou uma nova fase da Física de Partículas ao colidir feixes de prótons a energia de centro de massa de 13 TeV, no período de tomada de dados chamado *Run II*. A operação nas primeiras semanas de feixe estável em luminosidade reduzida (onde o empilhamento de sinais é pequeno) permitiu realizar as primeiras medidas de precisão da seção de choque de produção de bósons W e Z nos canais de elétrons e múons utilizando 81 pb^{-1} de dados, com os resultados publicados em 2016 [58]. Este capítulo apresenta as medidas da seção de choque de produção de bósons W no canal de elétrons¹ e a comparação com as previsões teóricas.

Para o *Run II*, o ATLAS reformulou todo seu *framework* de análise dos dados. Isso exigiu o desenvolvimento de várias ferramentas para seleção e correção dos dados reconstruídos (algo que atualmente já é integrado no *framework* de redução dos dados), exigindo diversas interações a fim de validar a consistência dos resultados entre diversos outros grupos de análise. Um outro aspecto crucial dessa análise envolve a estimativa da contribuição de multijatos. Jatos de baixo momento transversal (p_T) são produzidos de forma abundante em colisores hadrônicos e podem ser confundidos com elétrons durante o processo de reconstrução. Diferentemente dos

¹Além do canal de elétrons, onde houve a participação direta do autor, foram realizadas medidas da produção de bósons W no canal de múons e medidas da produção de bósons Z nos canais de elétrons e múons. Os resultados combinados são apresentados no Apêndice A

outros fundos, as previsões teóricas não conseguem reproduzir de forma adequada a distribuição de multijatos para a seleção no canal de elétrons, devendo essa contribuição ser estimada em regiões de controle do espaço de fase utilizando métodos derivados dos próprios dados.

Para tanto, os seguintes princípios são observados no planejamento da análise :

- A medida das seções de choque fiduciais fazem uso de um espaço de fase o mais próximo possível da observação, que são os cortes aplicados na seleção de eventos. A dependência de modelos teóricos e pressupostos do Modelo Padrão é mínima, atingindo a melhor precisão para cada lépton individualmente.
- Medidas da razão entre as seções de choque inclusivas de bósons W^+ e W^- são particularmente sensíveis à distribuição de quarks de valência u e d , enquanto que medidas da razão entre as seções de choque inclusivas de bósons $W^\pm$ e Z são importantes na restrição de quarks s . Essa informação é parte das condições usadas nas previsões teóricas, portanto servem como “seletoras” entre vários modelos.
- A combinação entre os canais de elétrons e múons leva em conta as correlações entre as incertezas sistemáticas, mas o aumento dessas incertezas é controlado pelo uso do menor espaço de fase comum entre os dois léptons. A redução das incertezas estatísticas e das incertezas sistemáticas não correlacionadas torna os resultados combinados (combinação entre sabores leptônicos) mais precisos.
- A extrapolação para o espaço de fase total deve ser feita, já que permite a comparação com os resultados de outros experimentos, porém ao custo de uma maior dependência dos modelos teóricos implementados nos geradores de eventos. Os resultados da extrapolação também podem ser necessários para algumas buscas por nova Física.
- Devido à radiação de estado final (FSR, *Final State Radiation*), correções são aplicadas aos elétrons no estado final para se chegar a resultados compatíveis com aqueles que seriam obtidos se fossem usados elétrons no seu estágio inicial

da produção. O efeito dessas correções é minimizado contabilizando-se toda a radiação eletromagnética próxima aos elétrons nas simulações.

- O controle das incertezas sistemáticas tem como objetivo mantê-las no nível mais baixo possível. Os fatores de correção usados são decompostos de forma que estimativas baseadas em dados coletados sejam maximamente exploradas, como no caso das eficiências de identificação, reconstrução e *trigger* dos elétrons. De forma semelhante, os fundos de multijatos, que não são estimados adequadamente por simulação, são estimados por métodos baseados em dados.

5.1 Metodologia

Os dados usados neste capítulo passaram por uma seleção de eventos dentro da qual são selecionados eventos que estão dentro de um espaço de fase limitado chamada *região fiducial*, correspondente à uma região de operação do detector em que os dados são coletados ou reconstruídos com maior precisão. A seção de choque nesta região é denominada seção de choque fiducial σ_W^{fid} , a qual é definida como:

$$\sigma_W^{fid} \times BR(W \rightarrow l\nu) = \frac{N - B}{C_W \cdot \mathcal{L}}. \quad (5.1)$$

onde N é o número total de eventos medidos, B corresponde à soma da contribuição estimada de todos os fundos, $\mathcal{L}$ é a luminosidade integrada e o fator de eficiência C_W é a razão entre o número de eventos reconstruídos que passaram por toda a seleção e o número total de eventos gerados dentro da região fiducial.

A seção de choque total σ é dada pela expansão da seção de choque fiducial para o espaço de fase total. Na prática σ é dada pela razão

$$\sigma_W^{tot} \times BR(W \rightarrow l\nu) = \frac{\sigma_W^{fid}}{A_W}, \quad (5.2)$$

onde o fator de aceitação A_W é definido pela fração de eventos gerados que estão dentro da região fiducial, calculada a partir de eventos que ainda não passaram pela simulação do detector (reconstrução) e cujas partículas ainda não emitiram radiação de estado final.

A região fiducial para o processo $W^\pm \rightarrow e^\pm \nu$ foi definida como a região em que o momento transversal do elétron p_T^e é maior que 25 GeV, a pseudorrapidez do elétron

η^e esteja no intervalo $|\eta^e| < 2,5$ e fora da região $1,37 < |\eta^e| < 1,52$ ², o momento transverso do neutrino representado pela E_T^{miss} seja maior que 25 GeV e a massa transversa do bóson m_T^W seja maior que 50 GeV. A massa transversa é definida como

$$m_T = \sqrt{2p_{T,\ell} p_{T,\nu} [1 - \cos(\phi_\ell - \phi_\nu)]} \quad (5.3)$$

onde ϕ_ℓ é o ângulo azimutal do elétron e ϕ_ν o ângulo azimutal do neutrino (obtido no cálculo da E_T^{miss}). Essas restrições estão descritas com mais detalhes na seção 5.3, sobre a seleção de eventos.

5.2 Amostras de dados e amostras simuladas

Os dados foram coletados pelo detector ATLAS entre junho e julho de 2015 a partir de colisões de prótons com energia de centro de massa de 13 TeV, ocorrendo colisões entre os grupos de prótons (*bunches*) cada 50 ns, com um pico de luminosidade instantânea $L = 1,7 \times 10^{33} \text{ cm}^{-1}\text{s}^{-1}$. O número médio de interações de empilhamento foi $\langle \mu \rangle = 19$ e a luminosidade integrada do período foi $L = 81 \text{ pb}^{-1}$. Os dados utilizados foram reconstruídos (e simulados no caso das previsões) utilizando a cadeia de processamento do ATLAS, o ATHENA [59].

Usando apenas os dados medidos não há como distinguir quais eventos são do sinal e quais processos contribuem para o fundo. Amostras geradas por simulação foram usadas para estimar as contribuições do sinal e dos fundos, exceto a contribuição dos multijatos que está descrita na seção 4.7. As amostras de simulação de Monte Carlo são produzidas de forma centralizada pelos grupos de simulação do ATLAS. Na produção dessas amostras, a simulação do detector foi processada pelo *software* de simulação e reconstrução do ATLAS [59], fazendo uso no pacote Geant4 [57] para simular as interações entre as partículas e o detector.

Foram geradas por simulação amostras dos canais $W^\pm \rightarrow e^\pm \nu$, $W^\pm \rightarrow \mu^\pm \nu$, $W^\pm \rightarrow \tau^\pm \nu$, $Z \rightarrow e^+ e^-$, $Z \rightarrow \mu^+ \mu^-$, $Z \rightarrow \tau^+ \tau^-$, além de amostras de pares de bósons $W^+ W^-$, ZZ e $W^\pm Z$ e amostras contendo um quark top ou um par de quarks $t\bar{t}$. A maioria das simulações usou o gerador Powheg-Box v2 [71] interfaceado com o Pythia 8.186 [72], tendo em vista que análises anteriores mostraram que

²nessa região a cobertura calorimétrica é reduzida e portanto a incerteza é maior

essa combinação descreve adequadamente as distribuições cinemáticas na produção inclusiva de bósons Z e $W^\pm$. Esses geradores produzem previsões da seção de choque em NLO (*Next to Leading Order*). A fim de normalizar a seção de choque integrada para uma previsão de ordem superior (NNLO - *Next Next Leading Order*) a seção de choque nessa ordem superior é calculada por um programa de ordem fixa (FEWZ 3 [61] ou DYNNLO 1.5 [60]) e a razão entre a previsão em NNLO e NLO são utilizadas como um peso adicional aplicada evento a evento. A Tab. 5.1 apresenta a lista de amostras simuladas, o gerador empregado e a respectiva seção de choque:

5.3 Seleção de eventos

Os elétrons são reconstruídos a partir de depósitos de energia no calorímetro eletromagnético dentro da região $|\eta| < 2,47$. Esses depósitos formam *clusters* que são organizados numa rede de torres, sendo que um *cluster* tem como semente uma torre com energia transversa de pelo menos 2,5 GeV. Os *clusters* associados a uma traço bem definido no ID são os candidatos a elétrons [42]. Os múons são reconstruídos independentemente pelo ID e pelo espectrômetro de múons. As informações de cada um dos subsistemas são então combinadas para formar traços de múons a serem usados nas análises físicas [84]. Na análise foi usada a energia transversa faltante reconstruída pelo TST E_T^{miss} , já mencionado no capítulo 4. Os objetos físicos usados nessa análise tiveram que passar pelos critérios de seleção a seguir :

- Apenas eventos coletados com o detector em completas condições operacionais foram usados na análise. Foi exigido que os eventos selecionados tivessem pelo menos um vértice com dois traços associados. Mais de 97% dos eventos atendem a esse critério.
- Foram selecionados elétrons que ativaram um *trigger* com requerimento de momento transverso $p_T^e \geq 24$ e isolamento classificado como *loose*, conforme descrito na subseção 4.1.3, ou um *trigger* sem isolamento que exige um momento transverso mínimo de 60 GeV para ser ativado. Os elétrons devem ter sido identificados por um algoritmo baseado em verossimilhança e terem sido classificados como *medium*, critério que apresenta uma eficiência de apro-

Tabela 5.1: Eventos simulados utilizados na análise da produção de bósons $W^\pm$ na região fiducial. Alguns geradores são configurados para a filtrar apenas eventos de interesse para serem reconstruídos. Esses filtros possuem uma eficiência e a seção de choque utilizada para cada processo deve levar em conta esse fator. Nas amostras de di-bóson, os termos adicionais indicam: $SF^- \rightarrow$ mesmo sabor leptônico, cargas negativas; $OF^- \rightarrow$ sabor leptônico diferente, cargas negativas; $SF^+ \rightarrow$ mesmo sabor leptônico, cargas positivas e $OF^+ \rightarrow$ sabor leptônico diferente, cargas positivas.

Canal	Gerador	$\sigma \cdot \text{BR} \times \epsilon_{\text{filter}}$ [pb]
$W^+ \rightarrow e\nu$	Powheg-Box+Pythia 8	11501
$W^- \rightarrow e\nu$		8579
$W^+ \rightarrow \mu\nu$		11501
$W^- \rightarrow \mu\nu$		8579
$W^+ \rightarrow \tau\nu$		11501
$W^- \rightarrow \tau\nu$		8579
$Z \rightarrow ee$		1892
$Z \rightarrow \mu\mu$		1892
$Z \rightarrow \tau\tau$		1892
$t\bar{t}$	Powheg-Box+Pythia 6	451
Wt		34
$W\bar{t}$		34
t (t-chan)		44
$\bar{t}$ (t-chan)		26
$ZZ \rightarrow qql\bar{l}$	SHERPA2.1.1	2.3
$WZ \rightarrow qql\bar{l}$		3.76
$WZ \rightarrow l\nu qq$		12.54
$WW \rightarrow l\nu qq$		25.99
$ZZ \rightarrow ll\nu\nu$		14.02
$ZW \rightarrow lll\nu SF^-$		1.84
$ZW \rightarrow lll\nu OF^-$		3.62
$ZW \rightarrow lll\nu SF^+$		2.56
$ZW \rightarrow lll\nu OF^+$		5.02
$ZZ \rightarrow llll$		12.58
$b\bar{b}$	Pythia 8B	187710
$c\bar{c}$		58528

ximadamente 90% para elétrons com $E_T \approx 40$ GeV [85], e devem satisfazer $p_T^e > 25$ GeV e estarem no intervalo de pseudorrapidez $|\eta^e| < 2,47$, excluindo a região de transição entre o barril e as tampas do calorímetro eletromagnético $1,37 < |\eta^e| < 1,52$. Os elétrons ainda tiveram que passar pelo critério de isolamento *gradient* (cf. Sec. 5.5).

- Na análise de $W^\pm$ foram selecionados eventos contendo exatamente 1 lépton,

com energia transversa faltante $E_T^{\text{miss}} > 25$ GeV e a massa transversa do bóson $m_T > 50$ GeV. Para reconstruir a E_T^{miss} foram selecionados jatos reconstruídos pelo algoritmo *anti- k_t* [86] com o parâmetro $R = 0,4$, num intervalo de pseudorapidez $|\eta| < 4,5$ e momento transverso $p_T > 20$ GeV, após terem suas energias calibradas. Eventos contendo jatos calibrados com qualidade *LooseBad* que passaram no limiar da variável *jet-vertex-tagger* (JVT)³ [87], e possuem momento transverso superior a 20 GeV foram rejeitados.

- A reconstrução da energia transversa faltante leva em conta a superposição entre jatos e outros objetos. Foi aplicada uma remoção de léptons e jatos sobreviventes à seleção que estavam superpostos, obedecendo aos seguintes critérios em função do raio ΔR no plano $\eta \times \phi$:
 - Se ΔR entre um elétron e um jato for menor que 0,2 o jato é removido;
 - O elétron é removido se $0,2 < \Delta R < 0,4$;
 - Se $\Delta R < 0,4$ entre um múon e um jato, o múon é removido se o jato tiver mais que dois traços com $p_T > 0,5$ GeV, senão o jato é removido.

5.4 Correções e incertezas sistemáticas

A principal fonte experimental de incertezas sistemáticas é a incerteza na luminosidade, que foi 2,1%. As demais incertezas sistemáticas contribuem para a incerteza no fator C_W da Eq. 5.1 e na estimativa dos fundos.

Com o objetivo de tornar as amostras geradas por simulação as mais próximas possíveis dos dados coletados, são aplicados diversos pesos e calibrações aos eventos e objetos reconstruídos. Cada uma dessas correções tem uma incerteza associada. A propagação das incertezas sistemáticas devidas a tais correções é obtida aplicando-se as mesmas com as respectivas incertezas acrescidas e subtraídas do valor nominal. A incerteza sistemática total no fator C_W é obtida pela combinação das incertezas

³Para suprimir os jatos provenientes de empilhamento foi usada uma variável discriminante chamada *jet-vertex-tagger* (JVT). A JVT é calculada por um método multivariáveis baseado em verossimilhança a partir de outras duas variáveis, uma dependente da soma da magnitude do momento transverso dos traços associados ao jato e dos traços associados às interações de empilhamento e a outra variável definida como a razão entre a soma da magnitude do momento transverso dos traços associados aos jatos e o momento transverso total do jato.

obtidas para cada uma das correções. As correções e pesos aplicados aos objetos físicos e eventos, assim como o cálculo das incertezas sistemáticas associadas, fizeram o uso de métodos validados pelos grupos de performance do ATLAS e são descritas na subseção a seguir.

As incertezas sistemáticas que contribuíram na incerteza do fator C não associadas à correções aplicadas em objetos são devidas ao empilhamento, aos geradores de eventos usados, às incertezas nos conjuntos de PDF's e, no canal de elétrons, à identificação errada da carga, que pode ocorrer devido ao efeito da emissão de radiação que leva à reconstrução incorreta dos traços. As incertezas atribuídas aos conjuntos de PDF's contribuem com apenas 0,1% e as incertezas associadas aos geradores contribuem principalmente na incerteza da E_T^{miss} e não foram consideradas explicitamente para não serem contabilizadas duas vezes.

As incertezas devido ao empilhamento são obtidas por uma técnica validada pela colaboração, que corrige as condições de empilhamento de amostras de Monte Carlo para que sejam equivalentes às condições dos dados coletados. Quando as amostras produzidas por Monte Carlo são geradas, não são usadas as mesmas condições de empilhamento dos dados, mas sim uma distribuição genérica do número médio de interações por colisão $\langle\mu\rangle$.

O método usado para corrigir o empilhamento das amostras simuladas faz um mapeamento que relaciona o período de coleta de dados ao número que identifica a amostra de Monte Carlo e calcula um peso comparando a distribuição de $\langle\mu\rangle$ da amostra de dados com a de Monte Carlo. A Fig. 5.1 mostra um histograma do número de interações por colisão, com a escala do eixo vertical convertida em luminosidade integrada, usado na correção das condições e empilhamento. Foi aplicado um fator adicional aos pesos de $1/1,16$ para levar em conta a fração de atividade inelástica que é diferente nas amostras de dados e de simulação, e também para equalizar o número de vértices em função de $\langle\mu\rangle$ entre os dados e as simulações. A incerteza sistemática associada ao procedimento afeta a distribuição de E_T^{miss} e as eficiências de isolamento e identificação dos léptons, logo a variação do fator C_W devido à variação dos fatores de escala das eficiências foi subtraída para evitar uma contagem dupla. A Fig. 5.2 mostra o efeito de se aplicar a correção das condições

de empilhamento, comparando duas distribuições de número de vértices primários reconstruídos geradas a partir de amostras simuladas, uma corrigida e outra não corrigida, com uma distribuição obtida a partir de dados coletados.

Figura 5.1: Distribuição do número de interações por colisão de amostras de dados usadas na análise na região inclusiva. Dados referentes a um subconjunto dos dados de colisão próton-próton a $\sqrt{s} = 13$ TeV.

5.4.1 Correções e sistemáticas dos objetos físicos

Elétrons

Os elétrons tem sua energia e momento calibrados, sendo aplicadas diferentes correções, tanto nas amostras de dados como nas amostras de simulação [88]. As correções aplicadas aos dados coletados no processo de calibração são:

- A equalização das escalas de energia das camadas do Calorímetro Eletromagnético, devido à sua segmentação longitudinal em 4 camadas.
- A correção da distribuição de material em frente ao Calorímetro Eletromagnético, que usa a razão entre as energias medidas pelas duas primeiras camadas, e leva em conta a probabilidade de conversão de fótons em frente ao PS.
- As propriedades dos *clusters* do Calorímetro Eletromagnético e informações do sistema de traços são calibradas à energia original de amostras simuladas usando um método de análise multivariáveis.

Figura 5.2: Comparação de distribuições do número de vértices geradas a partir da dados coletados (azul), de dados simulados sem a correção das condições de empilhamento (verde) e com a correção (vermelho). Dados referentes a um subconjunto dos dados de colisão próton-próton a $\sqrt{s} = 13$ TeV.

- A correção devido ao alargamento da distância entre os módulos do detector, devido ao efeito da gravidade.
- A correção devido a não homogeneidade azimutal da alta tensão aplicada nos diversos setores do Calorímetro Eletromagnético.
- Em algumas regiões do calorímetro existe uma tendenciosidade na amplificação do sinal. Essa dependência é removida aplicando-se uma correção que ajusta os dados à simulação.
- Fatores de escala são calculados para ajustar a resposta dos elétrons dos dados aos elétrons de simulações, usando uma amostra de eventos $Z \rightarrow e^+e^-$, sendo depois validados usando candidatos a elétrons de eventos $J/\psi \rightarrow e^+e^-$ de dados coletados.

Nas amostras de simulação, uma técnica validada pelo grupo de elétrons e fótons do ATLAS aplica as correções das propriedades dos *clusters* e dos traços de forma que sejam correspondentes aos valores originais de energia dos elétrons verdadeiros. Uma vez que a resolução de energia das amostras simuladas é artificialmente melhor

que a dos dados coletados, a energia dos elétrons simulados deve passar por um processo de degradação da resolução (*smearing*).

As eficiências devidas à reconstrução, à identificação, e ao isolamento devem ser corrigidas. Tais correções são dadas como fatores de escala, obtidos pela razão entre a eficiência das amostras de dados e das amostras de Monte Carlo.

Jatos

Embora os jatos não sejam explicitamente usados nesta análise, a reconstrução da energia transversa faltante E_T^{miss} e a medida de isolamento dos elétrons utilizam a informação dos jatos reconstruídos presentes no evento. O procedimento de calibração dos jatos começa ajustando-se a direção do jato de forma que ele aponte para o vértice identificado, em seguida o efeito do empilhamento é subtraído. A energia do jato é calibrada aplicando-se uma correção derivada da simulação, que é dependente de η e p_T e na qual tendenciosidades na pseudorrapidez dos jatos, devido às diferentes tecnologias adotadas em diferentes intervalos de η , são corrigidas. São aplicadas ainda correções para reduzir a dependência da energia em relação à estrutura dos jatos e reduzir os efeitos de jatos que não estão completamente contidos no calorímetro [89].

A energia dos jatos é calibrada de modo que a energia dos jatos gerados por Monte Carlo, depois de reconstruídos, seja a mesma energia antes deles passarem pela simulação do detector. Os fatores de correção são determinados a partir de amostras de dois jatos geradas por simulação, usando os jatos após terem o vértice primário e o empilhamento corrigidos. Para que a resolução em energia das amostras simuladas seja compatível com a resolução das amostras de dados, é necessário aplicar um fator de degradação nas primeiras [89, 90, 91].

Energia transversa faltante

Nas análises físicas a energia transversa faltante deve ser reconstruída a partir dos demais objetos físicos já calibrados pelos métodos validados pelos grupos de desempenho do ATLAS, conforme discutido no Cap. 4. Desse modo, a principal incerteza na E_T^{miss} é proveniente da propagação das incertezas dos jatos. As demais incertezas

associadas a E_T^{miss} são devidas à contribuição dos termos *soft*, descritos no capítulo 3, na calibração da escala e da resolução da E_T^{miss} .

5.5 Estudos dos critérios de isolamento do elétrons na região inclusiva

Uma forma tradicional de exigir que um lépton esteja isolado é calculando a razão entre uma das variáveis de isolamento *ptcone* ou *etcone* (v. Cap. 4) e o p_T do lépton, exigindo então que essa razão seja inferior a um limite. A Figura 5.3 mostra as distribuições de variáveis de isolamento divididas pelo p_T do elétron, obtidas a partir dos dados coletados no início da tomada de dados em junho de 2015, onde foi aplicada a seleção de eventos descrita na seção 5.3. A Figura 5.4 mostra o efeito da aplicação do isolamento usando as variáveis de isolamento relativas ao p_T na distribuição de m_T , sendo usados eventos que passaram pela seleção de eventos da seção 5.3, exceto pelo isolamento.

5.5.1 Abordagem funcional

A eficiência do método tradicional não apresenta uma uniformidade em η . Por isso é recomendado usar uma abordagem funcional, que apresenta uma eficiência de 90% para $p_T = 25$ GeV, chegando a 99% para $p_T = 60$ GeV, baseada tanto na informação do calorímetro (*topoetcone20/p_T*) e do traço (*ptvarcone20/p_T* para elétrons e *ptvarcone30/p_T* para múons). Esse tipo de isolamento é chamado na análise de *gradient*. A Fig. 5.5 mostra a distribuição de m_T de bósons $W^\pm$ para os dados e as simulações do sinal e dos fundos, após a seleção de eventos do sinal, descrita na seção 5.3, e após a mesma seleção, porém usando o isolamento invertido, ou seja, selecionando os eventos reprovados no critério de isolamento. Com a inversão do isolamento as simulações não reproduzem os dados por causa do excesso de multijatos que não foram simulados. Essa seleção será importante mais à frente, para definir a região de controle para a estimativa da contribuição do fundo por multijatos.

Figura 5.3: Distribuições das variáveis de isolamento relativas $ptcone20/p_T$, $ptcone30/p_T$, $etcone20/p_T$ e $topoetcone20/p_T$.

5.5.2 Isolamento no *trigger*

No primeiro nível de *trigger* do calorímetro eletromagnético (L1EM), o processador de *cluster* identifica objetos de *trigger* como um máximo local usando um algoritmo de janela móvel. A variável de isolamento é calculada na camada do EM a partir da soma das torres de *trigger* ao redor. Na camada hadrônica é usada a torre de *trigger* imediatamente atrás do máximo local. O cálculo da variável de isolamento usa quatro parâmetros e o corte é aplicado por comparação com uma tabela.

No *High Level Trigger* (HLT) o isolamento de elétrons e múons é baseado na variável $ptcone20/p_T$. Os elétrons são considerados como isolados se $ptcone20/p_T < 0,1$, nos *triggers* com isolamento classificado como *iloose* (contêm *iloose* no nome da variável que os identifica). Os múons são isolados se $ptcone20/p_T < 0,12$ nos *triggers* classificados como *iloose* e $ptcone30/p_T < 0,06$ nos *triggers* do tipo *imedium* (contêm *imedium* no nome da variável que os identifica). A Fig. 5.6 mostra o

Figura 5.4: Efeito do isolamento na distribuição de m_T usando as variáveis de isolamento relativas $ptcone20/p_T$, $ptcone30/p_T$, $etcone20/p_T$ e $topoetcone20/p_T$.

Figura 5.5: Distribuição de m_T para $W^\pm$ após seleção com o isolamento *gradient* (a) , e com o mesmo isolamento invertido (b).

efeito do isolamento no HLT na distribuição de $ptcone20/p_T$ para os primeiros dados coletados em junho de 2015, comparando a distribuição de eventos que passaram por um *trigger* com isolamento *loose* e eventos que passaram por um *trigger* sem esse isolamento.

Figura 5.6: Efeito do isolamento do HLT nas distribuições das variáveis de isolamento relativas para (a) elétrons identificados como *medium* e (b) elétrons identificados como *loose*. Foram usados dois *triggers* com limiar de $p_T^e > 24$ GeV, um sem isolamento (linha preta) e outro com isolamento *loose* (linha vermelha). Demais cortes seguiram os critérios de seleção do sinal.

5.6 O fundo de multijatos

Eventos envolvendo decaimentos semileptônicos de quarks pesados, hádrons identificados como léptons e , no caso de elétrons, elétrons provenientes de fótons, são todos chamados coletivamente de multijatos. Embora a contribuição de multijatos ao fundo seja negligenciável nas análises relacionadas a produção de bósons Z , dada a baixa probabilidade de haver dois falsos léptons isolados com massa invariante na região do polo do bóson Z num mesmo evento, a contribuição torna-se relevante nos canais de $W^\pm$. A simulação de multijatos não é satisfatória, principalmente no caso de elétrons, e um método baseado em dados é usado para estimar essa contribuição. A Fig. 5.7 deixa claro que as simulações de multijatos não são adequadas, dada a diferença entre a distribuição obtida usando-se o método baseado em dados e usando-se eventos simulados. A seleção de um lépton isolado com alta E_T^{miss} e m_T rejeita a maioria dos multijatos, mas uma contaminação ainda permanece importante nessa região do espaço de fase.

Figura 5.7: Distribuições da massa transversa de $W^\pm$ obtidas calculando-se a contribuição dos multijatos por simulação (linha vermelha) e usando-se um método baseado em dados (linha azul), normalizados em relação ao número total de eventos, mostrando que a simulação não reproduz a distribuição dos dados de forma adequada.

5.6.1 Metodologia

A metodologia de determinação dos multijatos explorou o fato dos léptons sobreviventes à seleção de eventos se encontrarem quase totalmente em regiões limitadas das variáveis de isolamento. Os elétrons sobreviventes ficaram restritos à região $ptvarcone20/p_T < 0,08$, conforme mostrado na Fig. 5.8.

Figura 5.8: Distribuição bidimensional das variáveis de isolamento relativas $ptvarcone20/p_T$ e $topoetcone20/p_T$ para elétrons e sobreviventes à seleção de eventos do sinal.

A determinação do número de eventos de multijatos foi feita por meio de ajustes de máxima verossimilhança em distribuições cinemáticas das variáveis discriminantes m_T , E_T^{miss} , p_T^l e $\Delta\phi$ entre a direção do lépton e da $\vec{E}_T^{\text{miss}}$. As distribuições foram construídas a partir de amostras enriquecidas com eventos de multijatos separadas em duas regiões de ajuste, sendo uma delas obtida removendo-se o corte em E_T^{miss} e a outra removendo-se o corte em m_T , e em ambas foi removido o corte de isolamento *gradient*. As distribuições cinemáticas foram produzidas para diferentes intervalos de uma das variáveis de isolamento, mantendo-se um corte fixo na região dos eventos sobreviventes ao isolamento *gradient* na outra. Para os elétrons a variável $ptvarcone20/p_T$ foi dividida em quatro intervalos iniciando em 0,08 e terminando em 4,0, com um corte fixo em $topoetcone20/p_T < 0,11$, e a variável $topoetcone20/p_T$ foi dividida em quatro intervalos iniciando em 0,11 e terminando em 4,0, com um

corde fixo em $ptvarcone20/p_T < 0,08$.

Para cada uma das distribuições o número de eventos de multijatos foi determinado por um ajuste de máxima verossimilhança, no qual o sinal e o fundo de multijatos estiveram livres para flutuar enquanto os fundos gerados por simulação de Monte Carlo foram mantidos fixos. O número de eventos de multijatos na região do sinal é obtido aplicando-se o requerimento de $m_T > 50$ GeV ou $E_T^{\text{miss}} > 25$ GeV para cada distribuição cinemática em cada intervalo de cada variável de isolamento. Os resultados obtidos para cada variável discriminante foram organizados em histogramas, cujos bins são os intervalos das variáveis de isolamento, e o número de eventos de multijatos foi extrapolado para as regiões correspondentes ao isolamento do sinal em cada histograma por um ajuste linear. A Fig. 5.9 ilustra o método para o canal $W^- \rightarrow e^- \bar{\nu}$.

Figura 5.9: Histogramas do número de eventos por variável de isolamento para o canal $W^- \rightarrow e^- \bar{\nu}$. (a,b) histogramas de $ptvarcone20/p_T$ e (c,d) os histogramas de $topoetcone20/p_T$, sem o corte em m_T (a,c) e sem o corte em E_T^{miss} (b,d).

O número de eventos de multijatos foi calculado para cada uma das variáveis

de isolamento e cada uma das regiões separadamente como a média ponderada dos valores obtidos para cada variável discriminante com suas incertezas. O valor nominal do número de eventos de multijatos para cada canal foi dado pela média dos resultados obtidos para cada região e cada variável de isolamento. As incertezas do ajuste linear foram propagadas como incertezas sistemáticas do método. A incerteza sistemática da escolha da variável de isolamento foi determinada calculando-se a média do número de eventos obtidos das duas regiões de ajuste para cada variável de isolamento e tomando o valor da incerteza como metade da diferença entre as duas médias. De forma semelhante a incerteza na escolha da região de ajuste foi obtida calculando-se a média do número de eventos de multijatos das duas variáveis de isolamento para cada região de ajuste e tomando-se a metade da diferença entre elas como a incerteza. Ainda contribuíram para a incerteza sistemática nos multijatos a incerteza na escala de energia dos jatos. A correlação ρ entre os fundos de multijatos das amostras de W^- e W^+ foi estimada para cada incerteza repetindo-se o procedimento para uma amostra inclusiva das duas cargas e usando-se a equação:

$$\delta(W^\pm)^2 = \delta(W^+)^2 + \delta(W^-)^2 + 2\rho\delta(W^+)\delta(W^-) \quad (5.4)$$

As frações de eventos de multijatos estimadas para os canais de elétrons foram 8% e 10% e as incertezas relativas foram 26% e 30% para W^+ e W^- respectivamente. A título de comparação, nos canais de múons as contribuições de multijatos foram 3,5% e 4% com incertezas de 22% e 30% para W^+ e W^- respectivamente.

5.7 Resultados

A boa concordância entre as simulações e os dados coletados pode ser observada nas distribuições cinemáticas. A Fig. 5.10, mostra as distribuições de momento transversal do elétron, pseudorrapidez do elétron e de energia transversa faltante para os canais de $W^\pm \rightarrow e^\pm \nu$. As mesmas distribuições, utilizando como critério adicional o sinal da carga do elétron reconstruído são apresentadas nas Figs. 5.11 e 5.12. Todas as distribuições incluem os fundos relevantes e as incertezas estão representadas na forma de bandas hachuradas.

Figura 5.10: Distribuições cinemáticas para produção de bósons W . O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa hachurada e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos). (a) momento transversal do elétron, (b) distribuição em pseudorrapidez dos elétrons η^e , (c) energia transversa faltante E_T^{miss} e (d) massa transversa m_T^W .

Figura 5.11: Distribuições cinemáticas (p_T^e e η^e) para produção de bósons W separadas pelo sinal da carga dos elétrons reconstruído. O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa hachurada e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos). Momento transverso do elétron p_T^e ((a) e^+ , (b) e^-), distribuição em pseudorapidez dos elétrons η^e ((c) e^+ , (d) e^-).

Figura 5.12: Distribuições cinemáticas (E_T^{miss} e m_T^W) para produção de bósons W separadas pelo sinal da carga dos elétrons reconstruído. O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa hachurada e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos). Energia transversa faltante E_T^{miss} ((a) e^+ , (b) e^-), massa transversa m_T^W ((c) e^+ , (d) e^-).

5.8 Contribuição dos fundos determinados por simulação

As contribuições dos fundos simulados ao canal de $W^\pm$ decaindo em elétrons foram $W^\pm \rightarrow \tau^\pm \nu$, $Z \rightarrow e^+ e^-$, $Z \rightarrow \tau^+ \tau^-$, os fundos contendo um quark top ou um par $t\bar{t}$, e os pares de bósons ZZ , WW e WZ . As contribuições de cada um, em porcentagem no número final de eventos simulados, encontram-se na Tab. 5.2, assim como as contribuições dos fundos simulados nos demais canais.

Tabela 5.2: Contribuições no canal $W \rightarrow e\nu$ das amostras geradas por simulação em relação ao número total de eventossimulados sobreviventes para cada um dos canais estudados, em porcentagem.

Processo	Contribuição
$W \rightarrow e\nu$	95,5 %
$W \rightarrow \tau\nu$	1,8 %
$Z \rightarrow \tau\tau$	0,12 %
Dibósons	0,08 %
$t\bar{t}$	1,1 %
t	0,2 %
$Z \rightarrow ee$	1,2 %

5.8.1 O fator C_W

As contribuições relativas de cada uma das incertezas sistemáticas no fator C para cada um dos canais encontra-se na Tab. 5.3.

5.8.2 Cálculo da seção de choque

As informações necessárias para o cálculo das seções de choque usando a equação 4.1, assim como as seções de choque fiducial σ^{fid} e total σ^{tot} calculadas para cada canal, incluindo as respectivas incertezas estatísticas, sistemáticas e da luminosidade são apresentadas na Tab. 5.4 :

Tabela 5.3: Incertezas sistemáticas relativas $\delta C_W/C_W$ nos diversos canais, em porcentagem.

Contribuição	$W^+ \rightarrow e^+\nu$	$W^- \rightarrow e^-\nu$
<i>Trigger</i>	0,3	0,3
Reconstrução e identificação do lépton	0,5	0,6
Isolamento do lépton	0,1	0,1
Resolução e escala na E do lépton	0,4	0,4
Identificação da carga	0,1	0,1
JES e JER	1,7	1,7
E_T^{miss}	0,1	0,1
Empilhamento	0,4	0,3
PDF	0,1	0,1
total	1,9	1,9

Tabela 5.4: Aceitação A , fator C e seções de choque calculadas para cada canal.

	W^+
Eventos	$228060 \pm 510(stat) \pm 4920(sys) \pm 200(lumi)$
C_W	$0,602 \pm 0,012(sys)$
σ^{fid} [nb]	$4,68 \pm 0,01(stat) \pm 0,14(sys) \pm 0,10(lumi)$
A	$0,383 \pm 0,007(sys)$
σ^{tot} [nb]	$12,23 \pm 0,003(stat) \pm 0,42(sys) \pm 0,27(lumi)$
	W^-
Eventos	$177890 \pm 450(stat) \pm 6110(sys) \pm 180(lumi)$
C_W	$0,614 \pm 0,012(sys)$
σ^{fid} [nb]	$3,58 \pm 0,01(stat) \pm 0,14(sys) \pm 0,08(lumi)$
A	$0,398 \pm 0,007(sys)$
σ^{tot}	$9,00 \pm 0,02(stat) \pm 0,39(sys) \pm 0,20(lumi)$

Através dos cálculos efetuados com o gerador de ordem fixa (FEWZ [61]) em NNLO e utilizando a PDF CT14nnlo [65], foram obtidos os seguintes valores para as seções de choque:

Tabela 5.5: Seções de choque calculadas pelo gerador FEWZ utilizando PDF CT14nnlo.

Canal	σ^{fid} [nb]	σ^{tot} [nb]
W^+	$4,50 \pm 0,14$	$11,55 \pm 0,32$
W^-	$3,43 \pm 0,11$	$8,54 \pm 0,24$
$W^\pm$	$7,93 \pm 0,25$	$20,10 \pm 0,54$

Com a informação das Tab. 5.4 e 5.5 é possível calcular a razão entre seções de choque, a qual possui a vantagem de cancelar a incerteza na energia do feixe e na luminosidade:

$$\frac{\sigma_{W^+}^{fid}}{\sigma_{W^-}^{fid}} = \frac{N_{sig}^{W^+}}{LC_{W^+}} \times \frac{LC_{W^-}}{N_{sig}^{W^-}} = \frac{N_{sig}^{W^+} C_{W^-}}{N_{sig}^{W^-} C_{W^+}} \quad (5.5)$$

A razão entre as seções de choque fiduciais no canal de elétrons entre W^+ e W^- , com as respectivas incertezas estatísticas e sistemáticas medidas foram

$$\sigma_{W^+}^{fid(meas)} / \sigma_{W^-}^{fid(meas)} = 1,308 \pm 0,004(stat) \pm 0,031(sys) \quad (5.6)$$

Utilizando os dados da Tab. 5.5 a razão segundo a previsão teórica é

$$\sigma_{W^+}^{fid(pred)} / \sigma_{W^-}^{fid(pred)} = 1,312 \pm 0,059(sys) \quad (5.7)$$

Esses resultados são as primeiras medidas de seções de choque de produção de bósons $W^\pm$ ⁴ a $\sqrt{s} = 13$ TeV, a maior energia de centro de massa atingida por um colisor, feitas pelo ATLAS. Os resultados acima mostram uma boa concordância com os cálculos teóricos baseados na QCD em ordem NNLO e apontam para a possibilidade de medir a razão entre as seções de choque $\sigma_{W^+}^{fid(meas)} / \sigma_{W^-}^{fid(meas)}$ com uma incerteza total de apenas 0,8%. Os resultados deste trabalho foram combinados com os resultados do canal de $W^\pm \rightarrow \mu^\pm \nu$ e também com os canais de $Z \rightarrow e^+ e^-$ e $Z \rightarrow \mu^+ \mu^-$, estudados por outros grupos da colaboração. A combinação dos resultados está disponível no Ap. A e foram publicados em [58].

⁴e também bósons Z , publicadas em conjunto em [58].

Capítulo 6

Medidas da produção $W^\pm \rightarrow e^\pm \nu$ na região de alta massa transversa em colisões próton-próton a $\sqrt{s} = 13$ TeV

Após a medida da seção de choque total na região *inclusiva* (região que compreende o pico da distribuição de massa transversa dos bósons $W^\pm$) a medida foi estendida para a região de alta massa transversa (região que compreende a faixa de massa transversa maior que 130 GeV). Medidas de precisão nessa região (conforme foi descrito na introdução desse trabalho) tem grande interesse para buscas além do modelo padrão e estratégias empregando teorias de campo efetivas para previsões. Além disso, como a produção de bósons $W^\pm$ é um fundo importante para muitas buscas na região de alta massa, a medida correta da seção de choque de produção de bósons $W^\pm$ na região de alta massa transversa é crucial para a sensibilidade dessas buscas. Deve-se mencionar que o acesso à essa região na escala de TeV só se tornou possível graças à copiosa quantidade de dados produzidas pelo LHC durante o *Run II* (aproximadamente 150 pb^{-1}) entre os anos de 2015 e 2018.

Foram selecionados dados coletados durante os anos de 2016 e 2017 pelo detector ATLAS a partir de colisões de prótons com energia de centro de massa $\sqrt{s} = 13$ TeV, com espaçamento temporal nominal entre as colisões de 25 ns. A luminosidade integrada e o número médio de interações por cruzamento de prótons ($\langle\mu\rangle$) que correspondem aos dados utilizados nesta análise foram respectivamente $\mathcal{L} = 32995 \text{ pb}^{-1}$ e $\langle\mu\rangle = 25,1$ para o ano de 2016 e $\mathcal{L} = 43584 \text{ pb}^{-1}$ e $\langle\mu\rangle = 37,8$ para 2017[94].

A Fig. 6.1 mostra a distribuição do número médio de colisões por cruzamento de feixe observada durante o Run-2 [94]. A fim de uniformizar os critérios de seleção e eficiência, os dados coletados em 2015 não foram considerados, uma vez que o *trigger* utilizou um limiar de 25 GeV no momento transversal do elétron, enquanto o limiar mínimo para os dados de 2016 e 2017 é de 27 GeV. A luminosidade integrada referente aos dados de 2015 corresponde a apenas 4% do total de dados coletados nos três anos, e portanto sua não utilização não impacta a análise. Os dados de 2018 (último ano de operação do Run-2) têm o potencial de praticamente dobrar a quantidade de dados aqui analisada. Entretanto, devido ao longo processo envolvido na preparação e análise (dados e simulações) não puderam ser aqui explorados.

Figura 6.1: Luminosidade integrada em função do número médio de colisões por cruzamento de feixe ($\langle\mu\rangle$) obtido durante o Run-2 durante os períodos de coletas de dados de 2015, 2016 e 2017 [94].

6.1 Previsões de Monte Carlo para a estimativa do sinal e o fundo

O sinal ($W^+ \rightarrow e^- \nu$, $W^+ \rightarrow e^+ \nu$) e a contribuição estimada para os fundos eletrofracos ($W^\pm \rightarrow \tau^\pm \nu$, $Z \rightarrow e^+ e^-$, $Z \rightarrow \tau^+ \tau^-$) foram produzidas pelo gerador Powheg [71] configurado para a geração de eventos contendo 1 bóson e *interfaceado* com o gerador Pythia 8.1 [72] para os processos de hadronização. A PDF utilizada é a

CT10 [105] e utiliza o *tunning* AZNLO CTEQL1[74]. O gerador EvtGen v.1.2.0 [77] foi utilizado para modelar o decaimento dos quarks b e c e o gerador Photos++ v.3.52 [78] utilizado para modelar a emissão via QED dos vértices eletrofracos e de léptons carregados. As seções de choque de produção integradas são normalizadas em NNLO por meio da previsão teórica fornecida pelo programa FEWZ [61], utilizando a PDF MSTW2008NNLO [83].

Os fundos associados à produção de quarks top ($t\bar{t}$, *single-top*, W em associação com um quark top nos canais s e t) foram produzidas pelo gerador Powheg-Box v2 [71] e o gerador Pythia v.6.4 [79] para a hadronização. A produção no canal $t\bar{t}$ é normalizada em NNLO+NNLL a partir dos resultados em ordem fixa obtidos pelo código Top++2.0 [80]. Os fundos provenientes da produção de di-bósons (WW , WZ , ZZ) foram geradas nos canais de decaimento totalmente leptônico e semi-leptônico utilizando o gerador Sherpa v.2.2.2 [81] empregando a PDF NNPDF30NNLO [66]. No ATLAS, todas as amostras (sinal ou fundo) utilizadas para análises oficiais (como neste caso) são produzidas centralmente pela colaboração e disponibilizadas em formato reconstruído para as análises físicas.

A produção central dos eventos de Monte Carlo segue a determinação e validação do ATLAS e permite harmonizar a geração entre todas as análises, sendo os geradores, as PDFs e os *tunnings* escolhidos com base na comparação com os resultados experimentais realizados em estudos prévios. As partículas produzidas pelos geradores são transportadas através do material de todos os subsistemas do ATLAS por um código baseado no Geant4 [57]. A Tab. 6.1 apresenta um sumário dos geradores e as PDFs usadas na simulação de cada um dos processos.

Nessa análise, cada uma das amostras de simulação usadas teve duas versões, cuja principal diferença entre elas é o perfil da distribuição do número médio de colisões por cruzamento de feixe $\langle\mu\rangle$. Uma das versões possui o perfil da distribuição de $\langle\mu\rangle$ adequado para ser comparado com os dados de 2015 e 2016 ($\langle\mu\rangle = 25.1$), enquanto a segunda versão é adequada para os dados de 2017 ($\langle\mu\rangle = 37.8$). Em ambos os casos (no contexto da geração das amostras) como as distribuições de $\langle\mu\rangle$ ainda não refletiam com fidelidade as mesmas condições em que os dados foram coletados, pesos devem ser aplicados evento a evento, num procedimento conhecido como *pileup*

Tabela 6.1: Geradores de eventos e PDFs usados para cada uma das amostras geradas por simulação de Monte Carlo, a ordem de cálculo e as PDFs empregadas. Também indicada é a ordem da normalização da seção de choque (k -factor), calculada utilizando geradores de ordem fixa (v. texto)

Processo	Gerador/ <i>parton shower</i>	Ordem (Norm.)	PDF
$W^\pm e^\pm \nu$	Powheg-Pythia 8	NLO (NNLO)	CT10
$W^\pm \tau^\pm \nu$	Powheg-Pythia 8	NLO (NNLO)	CT10
$Z \rightarrow e^+ e^-$	Powheg-Pythia 8	NLO (NNLO)	CT10
$Z \rightarrow \tau^+ \tau^-$	Powheg-Pythia 8	NLO (NNLO)	CT10
$t\bar{t}$	Powheg-Box v2/Pythia 8 + EvtGen 1.6	NLO (NNLO + NNLL)	CT10
Wt, Ws	Powheg-Box v1/Pythia 6 + EvtGen 1.2	NLO (NNLO + NNLL)	CT10f4
top	Powheg-Box v1/Pythia 6 + EvtGen 1.2	NLO (NNLO + NNLL)	CT10f4
dibósons (semileptônico)	Sherpa 2.2.1	NLO (NNLO)	NNPDF30NNLO
dibósons (leptônico)	Sherpa 2.2.2	NLO (NNLO)	NNPDF30NNLO

re-weighting. O observável utilizado nesse procedimento de normalização é o número de vértices primários reconstruídos, de forma que após esse processo, o número de eventos dessa distribuição em cada intervalo de $\langle \mu \rangle$ deve ser o mesmo tanto nos dados como nos eventos simulados, utilizando o mesmo procedimento descrito na Sec. 5.4.

A produção de bósons W na região de alta massa transversa possui uma seção de choque muito pequena quando comparada com a produção na região inclusiva, de forma que as amostras puramente inclusivas apresentam uma flutuação estatística intolerável no número de eventos na região de alta massa transversa, exigindo uma quantidade total de eventos extremamente alta para dominar esse problema. A solução neste caso consiste em configurar o gerador para a produção de eventos em intervalos específicos de massa do bóson para os processos $W^\pm \rightarrow e^\pm \nu$, $W^\pm \rightarrow \tau^\pm \nu$, $Z \rightarrow e^+ e^-$ e $Z \rightarrow \tau^+ \tau^-$. Essas amostras foram geradas em 19 intervalos (cobrindo a faixa de 120 GeV até 5000 GeV), tanto para o sinal como para o fundo devido à corrente neutra, conforme mostra a Tab. 6.2, com a massa dos bósons W ou Z . Na análise, essas amostras são combinadas com a produção inclusiva e devidamente normalizadas pelas suas respectivas seções de choque.

Tabela 6.2: Intervalos de massa utilizados pelo gerador de eventos Powheg na geração dos eventos de sinal ($W \rightarrow e^- \nu$, $W \rightarrow e^+ \nu$) e para cada componente do fundo ($Z \rightarrow l^+ l^-$, $W \rightarrow \mu^- \nu$, $W \rightarrow \mu^+ \nu$, $W \rightarrow \tau^- \nu$, $W \rightarrow \tau^+ \nu$). A amostra final os intervalos são combinados considerando-se as respectivas seções de choque, os *k-factors* e a eficiência dos filtros na seleção dos eventos de acordo com os intervalos de massa.

	Intervalo de massa [GeV]
1	[120, 180]
2	[180, 250]
3	[250, 400]
4	[400, 600]
5	[600, 800]
6	[800, 1000]
7	[1000, 1250]
8	[1250, 1500]
9	[1500, 1750]
10	[1750, 2000]
11	[2000, 2250]
12	[2250, 2500]
13	[2500, 2750]
14	[2750, 3000]
15	[3000, 3500]
16	[3500, 4000]
17	[4000, 4500]
18	[4500, 5000]
19	[5000, [

6.2 Seleção de eventos

Embora esta análise tenha como objetivo medir a seção de choque fiducial da produção de bósons W na região de alta massa ($m_T^W > 130\text{GeV}$), uma validação na região inclusiva ($m_T^W > 50\text{GeV}$) antecede essa medida, pois permite validar a metodologia e os algoritmos de reconstrução e seleção empregados.

Assim, dois critérios de seleção (**inclusivo** e **alta massa transversa**) são empregados para os observáveis cinemáticos conforme descrito a seguir. Exceto pelos limites geométricos (diretamente ligados à aceitação dos subsistemas do ATLAS) os critérios são escolhidos de forma a amostrar o máximo possível do espaço de fase e eliminar a contaminação, principalmente aquela associada ao empilhado devido ao expressivo aumento da luminosidade. Os eventos selecionados (dados e simulação) devem passar pelos seguintes critérios:

- Os eventos selecionados devem ter sido coletados em períodos nos quais o detector esteve plenamente funcional. As informações referentes à qualidade dos dados obtidas através do monitoramento do ATLAS durante a tomada de dados são utilizadas para essa seleção. Períodos em que se observa ruído excessivo em algum subsistema (principalmente nos calorímetros), grandes regiões inativas por falhas do detector ou sistema de aquisição de dados são excluídos. Devido à alta eficiência de operação, mais de 98% dos dados são aceitos por esse critério.
- Os eventos devem ter pelo menos um vértice primário reconstruído correspondente a no mínimo três traços;
- Os eventos devem ter ativado pelo menos um de três *triggers*, sendo o primeiro com limiar no momento transversal do elétron de $p_T^e > 26\text{ GeV}$ e um critério de isolamento “relaxado” (*loose*) e os outros dois com respectivos limiares de $p_T^e > 60$ e $p_T^e > 140\text{ GeV}$;
- Apenas um único elétron reconstruído deve satisfazer os critérios:
 - Pseudorrapidez no intervalo $|\eta^e| < 2,47$ e excluindo a região sem cobertura calorimétrica $1,37 < |\eta^e| < 1,52$;

- Na região inclusiva, o elétron deve ter um momento transverso mínimo de 27 GeV, limiar maior que na análise de 2015 uma vez que o *trigger* agora empregado tem um limiar de 26 GeV devido ao aumento da luminosidade. Na região de alta massa transversa o limiar é de $p_T^e > 65$ GeV;
 - Inicialmente era exigido que $ptvarcone20/p_T^e < 0,06$ (Sec.4.6), entretanto esse critério de isolamento não é suficiente para evitar a contaminação por objetos erroneamente identificados como elétrons na região de alto p_T^e . Assim, foi explorado um novo critério $topoetcone20 < 3,5$ para elétrons com $p_T^e > 60$ GeV mantendo o isolamento inicial para $p_T^e < 60$ GeV. Esse efeito do isolamento será discutido na Sec. 6.5.
- Eventos com um ou mais elétrons adicionais com $p_T^e > 20$ não isolados são excluídos;
 - Os jatos usados na reconstrução da E_T^{miss} passaram pelos mesmos critérios da medida de baixa luminosidade em 2015 e descritos na seção 5.3;
 - Os eventos devem apresentar uma energia transversa faltante $E_T^{\text{miss}} > 25$ GeV na região inclusiva e $E_T^{\text{miss}} > 65$ GeV na região de alta massa;
 - A massa transversa m_T^W mínima dos eventos selecionados deve ser maior que 50 GeV na região inclusiva e 130 GeV na região de alta massa.

6.3 Correções e normalização

6.3.1 Correções nas condições de empilhamento

Assim como foi descrito na seção 5.4, uma correção importante para essa medida são os pesos aplicados para ajustar as condições de empilhamento das simulações às condições nas quais os dados foram coletados. Porém, devido à diferença das condições de empilhamento dos dados coletados em 2016 e aqueles coletados em 2017, foi necessário empregar duas versões das amostras simuladas de cada um dos processos simulados¹. A Fig. 6.2 mostra a distribuição observada do número médio de

¹A fim de simular o efeito do empilhamento, um determinado número de eventos de *minimum bias* é sobreposto aos dados de *hard scattering* antes da reconstrução. O número exato de eventos de

colisões por cruzamento de feixe ($\langle\mu\rangle$) para o período de tomada de dados de 2016 e 2017 comprado com as simulações.

Figura 6.2: Distribuição do número de médio interações por cruzamento de feixe ($\langle\mu\rangle$) para os dados e as simulações de fundo e sinal na **região inclusiva**. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa rosa e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos).

A fim de normalizar as seções de choque dos processo simulados (calculados em NLO) para uma ordem superior (NNLO) é aplicado um fator (*k-factor*) calculado por meio de um gerador de ordem fixa como o FEWZ [61] em QCD (que gera as previsões em NNLO). Na região inclusiva, o *k-factor* é constante para todo o intervalo de massa do bóson até 120 GeV. Para a região de alta massa transversa, essa relação deixa de ser linear e exige um valor específico do *k-factor* para cada intervalo de massa. Como esse procedimento é rotineiramente empregado em análises que envolvem buscas por ressonâncias em alta massa (como W' ou Z') [98] existe um procedimento já bem estabelecido para cada gerador, que é feito durante o processo de produção central das amostras no ATLAS. Efetivamente, esse procedimento permite aplicar um peso,

minimum bias e sua distribuição é feita em função do empilhamento (e portanto da luminosidade instantânea) esperada.

evento a evento e amostra a amostra, para cada intervalo de massa, de forma que a seção de choque integrada em cada um dos intervalos descritos na Tab. 6.2 tem individualmente suas seções de choque corrigidas para a ordem superior (NNLO).

Após a reconstrução, existe um resíduo relacionado à eficiência do *trigger*, à eficiência de reconstrução dos observáveis (p_T , m_T^W , E_T^{miss} e o isolamento e identificação do elétron) e à resolução em momento e energia. Cada método de reconstrução tem associado um procedimento adicional de calibração de forma a compensar esses resíduos. Essa compensação é feita levando-se em conta a luminosidade instantânea, o momento transversal do objeto reconstruído e sua pseudo-rapidez e fazem parte do algoritmo de reconstrução que, considerando todas as variáveis descritas acima, indica um peso adicional que deve ser aplicado ao evento. Com isso, as amostras simuladas são normalizadas para a luminosidade integrada correspondente ao período de tomada de dados pela equação:

$$N_{\text{norm}} = \mathcal{L} \cdot \sigma \cdot \frac{\sum w_{\text{final}}}{\sum w_{\text{inicial}}}, \quad (6.1)$$

onde $\mathcal{L}$ é a luminosidade integrada, σ a seção de choque do processo simulado, $\sum w_{\text{final}}$ é a soma final dos pesos, sendo cada peso o produto do peso do gerador multiplicado pela correção do empilhamento e pelo *k-Factor*, e o termo no denominador $\sum w_{\text{inicial}}$ é a soma inicial dos pesos do gerador, correspondente ao número de eventos gerados de acordo com a seção de choque σ .

6.4 Resultados

A análise da região inclusiva permite verificar a consistência e validar o procedimento e ferramentas de análise, tendo em vista a excelente concordância observada na análise anterior em baixa luminosidade (Cap. 5). Para tanto, os eventos foram selecionados de acordo com os critérios descritos na Sec. 6.2 para a região inclusiva. A seguir são apresentadas as distribuições cinemáticas dos observáveis correspondentes à diferença entre o ângulo azimutal entre o elétron e o vetor energia transversal faltante ($\Delta\phi$), a distribuição de pseudo-rapidez do elétron (η^e), energia transversal faltante (E_T^{miss}), massa transversal (m_T^W) e momento transversal do elétron (p_T^e). As distribuições são apresentadas com e sem a identificação do sinal da carga do elétron

(e portanto do sinal da carga do bóson W).

A tabela 6.3 mostra o número de eventos selecionados e a contribuição de cada processo em relação aos dados. O processo mais dominante para o fundo é a produção de pares $t\bar{t}$, contribuindo com 1,93%, já que o fundo proveniente de multijatos não é considerado.

Tabela 6.3: Número de eventos após a seleção na região inclusiva dos dados e de cada um dos processos simulados e as respectivas contribuições em relação aos dados experimentais.

Processo	Eventos	Contribuição [%]
Dados	$3,510 \times 10^8$	100
$W^\pm \rightarrow e^\pm \nu$	$3,284 \times 10^8$	93,6
$t\bar{t}$	$5,887 \times 10^6$	1,68
$Z \rightarrow e^+ e^-$	$5,565 \times 10^6$	1,59
$W^\pm \rightarrow \tau^\pm \nu$	$5,428 \times 10^6$	1,55
dibósons	$5,920 \times 10^5$	0,17
top	$5,654 \times 10^5$	0,16
$Z \rightarrow \tau^+ \tau^-$	$4,868 \times 10^5$	0,14

6.4.1 Distribuições cinemáticas - região inclusiva

A Fig. 6.3 apresenta a distribuição do valor absoluto da diferença entre o ângulo azimutal $\Delta\phi$ entre o elétron e o vetor energia transversa faltante $\vec{E}_T^{\text{miss}}$ na região inclusiva. Nota-se que na maior parte dos eventos $\Delta\phi$ é próximo a π , indicando que o elétron e o neutrino provenientes do decaimento dos bósons $W^\pm$ saem preferencialmente em direções opostas.

A distribuição da pseudo-rapidez do elétron (η^e) na região inclusiva é apresentada na Fig. 6.3. Nessa distribuição é possível notar a ausência de eventos na região sem cobertura calorimétrica devido ao corte descrito na seção 6.2. A concordância entre dados e simulação é melhor na região central, e as flutuações aumentam com o aumento de η^e .

Figura 6.3: Distribuição do módulo da diferença de ângulo azimutal entre o elétron e o vetor $\vec{E}_T^{\text{miss}}$ ($\Delta\phi$) na **região inclusiva** (a,b) e distribuição em pseudorapidez dos elétrons (c,d). O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa rosa e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos). (a,c) escala linear, (b,d) escala logarítmica.

Figura 6.4: Distribuição da energia transversa faltante (E_T^{miss}) (a,b) e massa transversa (m_T^W) (c,d) na *região inclusiva*. O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa rosa e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos). (a,c) escala linear, (b,d) escala logarítmica.

Figura 6.5: Distribuição do momento transversal do elétron (p_T^e) para os eventos que passaram pela seleção da *região inclusiva*. O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa rosa e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos). (a) escala linear, (b) escala logarítmica.

Enquanto a distribuição do momento transversal dos elétrons (Fig. 6.5) apresenta uma boa concordância entre os dados e as previsões teóricas, observa-se que as distribuições de energia transversal faltante (E_T^{miss}) e, principalmente, a distribuição de massa transversal (m_T^W) possuem um excesso de dados na região de baixo E_T^{miss} e m_T^W . Esse excesso é devido à contribuição de multijatos, que deve ser estimada através de métodos baseados em dados (*data-driven*), conforme foi feito na análise do Cap. 5. Entretanto, existe uma complicação adicional nos dados desse novo período para a definição de uma região de controle, uma vez que todos os *triggers* de baixo p_T^e usados nesse período de tomada de dados têm embutidos o critério de isolamento (não é possível definir uma região em que o isolamento é relaxado). Este fator torna a estimativa dos multijatos muito mais complexa que a executada no Cap. 5. Como a análise inclusiva é apenas utilizada como verificação dos algoritmos e métodos visando a reconstrução em alta massa, a estimativa da contribuição dos multijatos para esta região não é importante, e portanto não foi objeto de estudos.

Seleção dos eventos aplicando critério de separação de cargas - região inclusiva

As distribuições que se seguem (Figs. 6.6, 6.7, e 6.8) apresentam os resultados da reconstrução obtida aplicando-se um critério adicional para selecionar os eventos de acordo com a carga reconstruída do lépton e seguem a mesma sequência das distribuições da seção anterior. Novamente, nota-se uma concordância adequada entre os dados simulados e os dados medidos dentro da incerteza sistemática, e as observações feitas na seção anterior podem igualmente ser aplicadas.

Figura 6.6: Distribuição do módulo da diferença de ângulo azimutal entre o elétron e o vetor $\vec{E}_T^{\text{miss}}$ ($\Delta\phi$) (a,b) e distribuição em pseudo-rapidez dos elétrons (c,d) na **região inclusiva**. O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa rosa e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos). (a,c) W^- , (b,d) W^+ .

Figura 6.7: Distribuição da energia transversa faltante (E_T^{miss}) (a,b) e da massa transversa (m_T^W) (c,d) na *região inclusiva*. O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa rosa e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os Dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos). (a,c) W^- , (b,d) W^+ .

Figura 6.8: Distribuição do momento transverso do elétron dos eventos na *região inclusiva*. O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa rosa e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos). (a) W^- , (b) W^+ .

6.5 Estudos na região de alta massa transversa

6.5.1 Estudos dos critérios de isolamento do elétrons na região de alta massa transversa

Conforme mencionado de antemão na Sec. 6.2, observou-se que o isolamento usando um corte simples em $ptvarcone20/p_T < 0,06$ não permite uma concordância adequada entre os dados e as previsões de Monte Carlo. A Fig. 6.9(a) apresenta essa observação na distribuição cinemática do momento transverso do elétron na região de alta massa transversa, utilizando o mesmo critério para o isolamento na região inclusiva. Observando-se a figura 6.9, fica claro que o critério de isolamento do elétron $ptvarcone20/p_T^e < 0,06$ não é adequado para altos valores de p_T^e (acima de 400 GeV, manifesto através de um excesso nos dados). Um procedimento alternativo, utilizando como critério $topoetcone20 < 3,5$ GeV foi apresentado e discutido com a colaboração e a Fig. 6.9(b) ilustra o resultado desse novo ponto de trabalho na distribuição de p_T^e , evidenciando uma melhora expressiva na concordância entre os dados e a simulação. Esse efeito deve-se, primariamente, à presença excessiva de fundos provenientes de partículas erroneamente classificadas como elétrons (e portanto não previstas nas simulações), passando pelos critérios de seleção e isolamento previamente usados. Devido ao altíssimo momento transverso, a informação do *tracker* não é eficiente na seleção do isolamento e nessa região a informação do calorímetro eletromagnético é fundamental para identificar as partículas não isoladas.

De posse de novo ponto de trabalho para o isolamento e utilizando os cortes e correções descritos nas Sec. 6.2 e 6.3 respectivamente, podem ser produzidas as distribuições cinemáticas na região de alta massa ($\Delta\phi$, η^e , E_T^{miss} , m_T^W e p_T^e), as quais são apresentadas nas Figs. 6.10, 6.11, 6.12, e 6.13. Uma observação importante para essa região diz respeito à contribuição do fundo devido à produção de pares $t\bar{t}$. Neste espaço de fase, aproximadamente 24% dos eventos é formado de pares $t\bar{t}$, conforme mostra a Tab. 6.4. Outra observação importante diz respeito ao modelo das distribuições de E_T^{miss} (m_T^W), onde existe uma deficiência na quantidade de eventos representada pelos dados entre 100 GeV e 600 GeV (200 GeV e 600 GeV). Como essa região de alta massa transversa só se tornou acessível recentemente com a alta

Figura 6.9: Distribuições de p_T^e na região de alta massa transversa aplicando o isolamento $ptvarcone20/p_T^e < 0,06$ para qualquer valor de p_T^e (a) e aplicando o isolamento $topoetcone20 < 3,5$ para $p_T^e > 60$ GeV (b). O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos). A figura apresenta apenas os dados coletados em 2017.

luminosidade proporcionada pelo LHC, é possível que os modelos que descrevem o sinal não estão adequadamente ajustados para reproduzir essas distribuições (observe, entretanto, como na região inclusiva a concordância é bastante satisfatória, cf. Fig. 6.4. A própria modelagem do fundo devido à produção de pares $t\bar{t}$ pode contribuir para esse efeito e será alvo de discussão mais adiante.

Tabela 6.4: Número de eventos após a seleção na região de alta massa transversa dos dados e cada um dos processos simulados e as respectivas contribuições em relação aos dados experimentais.

Processo	Eventos	Contribuição [%]
Dados	$1,942 \times 10^6$	100
$W^\pm \rightarrow e^\pm \nu$	$1,254 \times 10^6$	64,60
$t\bar{t}$	$4,640 \times 10^5$	23,89
$Z \rightarrow e^+ e^-$	$4,162 \times 10^4$	2,14
dibósons	$3,523 \times 10^4$	1,81
top	$3,290 \times 10^4$	1,69
$W^\pm \rightarrow \tau^\pm \nu$	$1,637 \times 10^4$	0,84
$Z \rightarrow \tau^+ \tau^-$	$1,359 \times 10^3$	0,07

Figura 6.10: Distribuição do número médio de colisões ($\langle\mu\rangle$) na *região de alta massa transversa* para a produção de $W^\pm$. O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa rosa e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos).

Figura 6.11: Distribuição do módulo da diferença de ângulo azimutal entre o elétron e o vetor $\vec{E}_T^{\text{miss}}$ (a,b) e pseudorapidez dos elétrons (c,d) na *região de alta massa transversa* para produção de $W^\pm$. O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa rosa e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos). (a,c) escala linear, (b,d) escala logarítmica.

Figura 6.12: Distribuição da energia transversa faltante E_T^{miss} (a,b) e massa transversa (m_T^W) (c,d) na *região de alta massa transversa* para produção de $W^\pm$. O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa rosa e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos). (a,c) escala linear, (b,d) escala logarítmica.

Figura 6.13: Distribuição do momento transversal do elétron (p_T^e) na *região de alta massa transversa* para produção de $W^\pm$. O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa rosa e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos). (a) escala linear, (b) escala logarítmica.

Distribuições cinemáticas separadas de acordo com a carga reconstruída do lépton

Da mesma forma que na análise realizada no Cap. 5, um dos observáveis de interesse é a razão da seção de choque de produção $\sigma_{W^+}^{fid}/\sigma_{W^-}^{fid}$ e para tal uma seleção adicional levando em conta a carga do lépton reconstruído é empregada. As distribuições a seguir apresentam os resultados da reconstrução da diferença de ângulo azimutal entre o elétron e o vetor $\vec{E}_T^{\text{miss}}$ na região de alta massa transversa e para a pseudo-rapidez do elétron ou pósitron (Fig. 6.14), energia transversa faltante E_T^{miss} e massa transversa m_T^W (Fig. 6.15) e momento transverso do elétron e pósitron (Fig. 6.16):

Figura 6.14: Distribuição do módulo da diferença de ângulo azimutal entre o elétron e o vetor $\vec{E}_T^{\text{miss}}$ (a,b) e pseudorapidez dos elétrons (c,d) na **região de alta massa transversa**. O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa rosa e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos). (a,c) W^- , (b,d) W^+ .

Figura 6.15: Distribuição da energia transversa faltante E_T^{miss} (a,b) e massa transversa (m_T^W) (c,d) na *região de alta massa transversa*. O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa rosa e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos). (a,c) W^- , (b,d) W^+ .

Figura 6.16: Distribuição do momento transverso do elétron (p_T^e) na *região de alta massa transversa*. O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa rosa e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos). (a) W^- , (b) W^+ .

6.6 Discussão das incertezas sistemáticas

As incertezas sistemáticas relativas aos observáveis são devidas às correções e calibrações aplicadas a esses objetos. Operacionalmente, as incertezas são propagadas aplicando-se as correções e calibrações, somando-se ou diminuindo-se uma variação equivalente ao um desvio padrão ao valor nominal, e repetindo-se os demais passos da análise para cada uma dessas variações. Os resultados obtidos a partir de cada variação são então somados respeitando-se as correlações entre elas. Todas as correções mencionadas, assim como suas variações, são aplicadas observando as prescrições e recomendações previamente validadas pelos grupos de desempenho da colaboração ATLAS.

Incerteza relativas aos elétrons

As incertezas sistemáticas relativas aos elétrons relevantes neste trabalho tem como fonte as correções de resolução e escala de energia aplicadas aos elétrons das amostras simuladas. As principais incertezas devem-se à correção no valor da energia, e levam em conta as propriedades dos *clusters* e traços usados em sua reconstrução, refletindo a estabilidade e a uniformidade dos ganhos e calibrações aplicados nos calorímetros e no detector interno de traços. A sistemática associada à resolução (*Electron Energy Resolution*) e a escala de energia (*Electron Energy Scale*) dos elétrons impacta diretamente os métodos de isolamento e a identificação de elétrons empregados, e portanto deve ser propagada para esses observáveis.

Incerteza associadas à energia transversa faltante

As incertezas sistemáticas associadas ao termo *soft* do cálculo do E_T^{miss} são estimadas utilizando diferentes geradores de Monte Carlo. As incertezas são dependentes da atividade devida a objetos de alto momento transversal no evento, a qual é modelada pela variável $\vec{p}_T^{\text{hard}}$, que é a soma vetorial do momento transversal de todas as partículas de alto $\vec{p}_T$ no evento, e a direção desse vetor $\vec{p}_T^{\text{hard}}$ é utilizado para discriminar entre a sistemática associada a efeitos de escala (manifesta longitudinalmente ao eixo de $\vec{p}_T^{\text{hard}}$) e a sistemática associada a efeitos de resolução (manifesta longitudinalmente e transversalmente ao eixo $\vec{p}_T^{\text{hard}}$). Essas incertezas são medidas

pelo grupo de *performance* do ATLAS utilizando eventos $Z \rightarrow \mu^+ \mu^-$ e levando em consideração possíveis diferenças na reconstrução devido à presença de material inativo no detector (o que afeta a resolução). Assim, existem 5 termos que compõem a sistemática final do E_T^{miss} : 3 termos relacionados a resolução (termo paralelo ao eixo p_T^{hard} , termo perpendicular à direção de $\vec{p}_T^{\text{hard}}$ e a correlção entre os dois) e 2 termos relacionados à variação da escala (*up* e *down*). Além disso, toda a sistemática particular aos objetos de alto momento que compõem o E_T^{miss} (elétrons, jatos, múons) são propagados no cálculo de E_T^{miss} , a qual é recalculada para cada variação.

Incertezas sistemáticas associadas aos jatos

As incertezas sistemáticas relacionadas aos jatos de maior importância são as incertezas devidas à calibração da energia (*Jet Energy Scale* - *JES*) e à degradação da resolução (*Jet Energy Resolution*- *JER*) [90]. As incertezas dos jatos dependem de 91 parâmetros (*nuisance parameters*), mas existem esquemas apropriados para a redução desses parâmetros para diferentes cenários de análise, preservando quando necessário as correlações. No ATLAS, as análises podem empregar quatro categorias de incertezas sistemáticas para os jatos:

- O conjunto **completo** de parâmetros (91) leva em consideração todas as correlações bin a bin e é o mais recomendável para análises de precisão envolvendo jatos explicitamente, sendo também recomendável que as incertezas sejam propagadas separadamente para cada parâmetro de modo que a informação completa das correlações seja mantida;
- O conjunto **globalmente reduzido** usa 21 parâmetros para a incerteza na escala da energia dos jatos, porém essa redução de parâmetros tem como custo a perda da informação das correlações, o que pode comprometer a combinação com resultados de outras análises, principalmente os sujeitos à combinação com resultados de outros experimentos. Os 21 parâmetros são agrupados em *Jet Energy Scale* e *Jet Energy Resolution* ;
- O conjunto **categoricamente reduzido** é semelhante ao globalmente reduzido, mas com um total de 29 parâmetros, o que permite que ele ainda man-

tenha informação das correlações;

- Existem ainda quatro configurações do conjunto **fortemente reduzido**, cada um com 4 parâmetros, recomendados apenas para análises não sensíveis a correlações entre jatos originadas da escala da energia, não sendo recomendadas para análises de precisão ou análises que pretendam fazer combinação de resultados.

Nesta análise foi adotado o esquema **globalmente reduzido**, uma vez que não há intenção de combinações com outras análises. Devido ao isolamento ser empregado na seleção de eventos, as sistemáticas associadas aos jatos têm um impacto em todas as distribuições.

As Figs. 6.17 e 6.18 mostram a contribuição das incertezas sistemáticas discutidas nas subseções anteriores em cada uma das distribuições cinemáticas para as regiões inclusiva e de alta massa, respectivamente. Nota-se a expressiva contribuição da sistemática associada aos jatos.

Figura 6.17: Efeito das incertezas sistemáticas referentes aos elétrons, aos jatos e à E_T^{miss} nas distribuições cinemáticas na *região inclusiva*. Em cada distribuição estão indicadas as contribuições de cada termo da sistemática.

Figura 6.18: Efeito das incertezas sistemáticas referentes aos objetos físicos nas diversas distribuições cinemáticas na *região de alta massa transversa*. Em cada distribuição estão indicadas as contribuições de cada termo da sistemática.

6.7 Medida da seção de choque fiducial e comparação com as previsões teóricas

Conforme foi discutido na introdução deste trabalho, a medida da seção de choque de produção de bósons vetoriais é sensível à dinâmica subjacente dos partons da colisão. Tais medidas constituem uma ferramenta importante para o ajuste dos modelos e das PDFs utilizadas nas previsões teóricas. Esta medida, feita pela primeira vez na região inclusiva na energia de centro de massa de 13 TeV mostrou, como foi discutido no Cap. 5 e Ap. A, o potencial discriminatório no que diz respeito às PDFs. Agora essa medida é estendida para a região de alta massa de bósons W , medida que é possível graças à copiosa quantidade de dados produzidos pelo LHC durante o RUN-2, tornando acessível uma região até agora dominada pela incerteza estatística nos dados. Da mesma forma que se procedeu para a análise dos dados da região inclusiva obtidos durante as duas primeiras semanas de operação em 13 TeV do LHC, a medida da razão da seção de choque de produção de bósons W^+ e W^- permite cancelar as incertezas associadas à determinação da luminosidade, sendo sensível à distribuição $u - d$ dos quarks de valência a valores baixos de x [107].

Como já discutido anteriormente, a seção de choque fiducial σ_W^{fid} é calculada pela Eq. 5.1:

$$\sigma_W^{fid} \times BR(W \rightarrow l\nu) = \sigma_W^{tot} \times BR(W \rightarrow l\nu) \cdot A_W = \frac{N - B}{C_W \cdot \mathcal{L}}.$$

onde

- N é o número de eventos medidos
- B é a somatória de todos os eventos estimados como fundo
- $\mathcal{L}$ é a luminosidade integrada correspondente ao período de tomada de dados
- A_W a aceitação cinemática para a identificação de bósons W
- C_W o fator de correção que leva em conta a eficiência de reconstrução, o *trigger*, a resolução (espacial, momento, energia) fatores de escala e regiões inativas do detector (como por exemplo a região do *crack* entre os calorímetros).

O fator A_W relaciona a seção de choque total com a seção de choque fiducial, e corresponde à aceitação geométrica e cinemática para a identificação de bósons W . O fator A_W é calculado contando-se o número total (N_{total}) de eventos produzidos por uma previsão teórica de determinada ordem fixa (por exemplo FEWZ [61] ou DYNNLO [60]) no espaço de fase completo e o número de eventos produzidos por esse mesmo gerador quando o espaço de fase é reduzido para a região cinemática utilizada na análise ($N_{fiducial}$):

$$A_W = \frac{N_{fiducial}}{N_{total}} \quad (6.2)$$

Como foi visto no Cap. 5, a seção de choque total é melhor determinada na região inclusiva devido à maior aceitação cinemática, incluindo a região do pico de massa transversa. A região de alta massa transversa aqui considerada está significativamente acima do pico da distribuição de massa transversa do bóson W , não havendo portanto sentido em medir a seção de choque total a partir de um espaço de fase reduzido como esse. Dessa forma, estamos interessados agora apenas na seção de choque fiducial, e o fator C_W é o único fator importante de correção. Esse fator é calculado utilizando-se a amostra de Monte Carlo de sinal (W^+ , W^- ou $W^\pm$) aplicando-se os mesmos critérios de seleção da análise para os eventos "verdadeiros" (*truth*) aplicados nos eventos reconstruídos. Critérios exclusivos aplicados na reconstrução do sinal simulado (como a qualidade da identificação do elétron e a fração de vértices reconstruídos para os Jatos (JVT, cf. Cap. 3), bem como a correção para a resolução em energia não são aplicados na seleção dos eventos verdadeiros, que se restringe portanto à uma seleção geométrica e cinemática das partículas. O fator C_W é obtido então como:

$$C_W = \frac{N_{reconstruido}}{N_{verdadeiro}} \quad (6.3)$$

Deve-se notar que, enquanto o fator A_W carrega um significado puramente físico, o fator C_W depende da qualidade e da eficiência dos métodos utilizados na reconstrução dos eventos, e pode mudar no decorrer de uma análise, dependendo por exemplo da calibração dos detectores, dos algoritmos de reconstrução e seleção de eventos etc. O empilhamento de sinais pode também influenciar o fator C_W , uma

vez que afeta a eficiência do sistema de *trigger* e principalmente a associação traço vértice.

No canal de elétrons, as seguintes incertezas sistemáticas são consideradas:

- Escala de energia e resolução: as incertezas na calibração de energia dos elétrons afetam diretamente as distribuições cinemáticas e podem causar uma pequena alteração na aceitação via a migração de eventos para baixo ou para cima do corte em p_T .
- Identificação e reconstrução dos elétrons: a eficiência da seleção de elétrons (identificação, isolamento e aplicação do ponto de trabalho *tight*, *medium* ou *loose* escolhido de forma a privilegiar a pureza da amostra (*tight*) ou a eficiência de identificação (*loose*)) são corrigidas com fatores de escala específicos e tem associados a eles incertezas sistemáticas.
- *Trigger*: Devido à variação intrínseca na eficiência do trigger existe uma variação no número de eventos selecionados. Essa variação na eficiência depende principalmente do p_T do elétron. Devido ao ponto de operação escolhido situar-se no patamar de eficiência do *trigger* (99.8%), onde a variação de eficiência é pequena ($\pm 0.2\%$), essa incerteza não é considerada nesta análise.
- Escala de energia e resolução de jatos: Embora essa análise não explore diretamente as propriedades cinemáticas dos jatos, estes são utilizados no cálculo da energia transversa faltante (E_T^{miss}) e são selecionados os eventos onde o jato mais energético tenha p_T acima de 20 GeV e que durante sua reconstrução exista um traço reconstruído associado a ele no detector interno. A incerteza na calibração da energia e resolução desses jatos afeta o cálculo de E_T^{miss} , devendo ser propagadas para esse cálculo.
- Escala de energia e resolução do E_T^{miss} : A reconstrução da energia transversa perdida é feita utilizando a soma vetorial dos elétrons, muon e jatos reconstruídos e selecionados acima de um certo limiar de p_T . Além disso, a estimativa do E_T^{miss} utiliza a contribuição de traços corretamente reconstruídos (e associados a um único vértice) mas de baixo p_T , sendo considerados

como componentes *soft*. Da mesma forma que os jatos afetam a estimativa de E_T^{miss} , a presença desses componentes *soft* possui uma sistemática associada na eficiência de reconstrução dos traços. Por ser um observável complexo formado por vários objetos reconstruídos, a estimativa do impacto da incerteza sistemática no E_T^{miss} é calculada a partir da variação sistemática de seus constituintes convoluídas pelo algoritmo de cálculo do E_T^{miss} . O cálculo de C_W é assim afetado pela migração de eventos na fronteira de corte de E_T^{miss} e m_T^W .

Como a determinação do C_W e a estimativa dos fundos é feita através de eventos produzidos por geradores de Monte Carlo (em *Born level*)², existe intrinsecamente uma incerteza associada ao modelo teórico utilizado. Para o C_W existe a possibilidade de utilizar um conjunto de diferentes PDFs ou variar a escala de renormalização e estimar o impacto no cálculo do C_W . Ambos os casos requerem a geração de um conjunto de amostras de Monte Carlo específico para cada variação. Tais amostras não estão ainda disponíveis na região de alta massa transversa e portanto o C_W é considerado nominalmente. Entretanto, a propagação das incertezas sistemáticas conforme descritas no parágrafo anterior podem afetar a região de aceitação através da migração de eventos para dentro ou para fora dos pontos de corte, e a incerteza no C_W devido a esses fatores deve ser estimada.

A luminosidade, um dos parâmetros da Eq. 5.1, também possui uma incerteza associada. No ATLAS e em outros colisores, é possível realizar uma medida da luminosidade absoluta e calibrar os sistemas de medida de luminosidade (LUCID no ATLAS, v. Cap. 2). variando-se a posição dos feixes de prótons de modo a efetuar-se uma varredura transversal. Essa medida, associada à medida da corrente dos feixes, é chamada de varredura de Van de Meer e detalhes podem ser encontrados em [108]. Para o período de dados considerado (2016 e 2017) a incerteza no valor nominal da luminosidade integrada é de $\pm 2\%$, dessa forma tem-se:

$$L = (76580 \pm 1532) \text{ pb}^{-1}. \quad (6.4)$$

²*Born level* corresponde ao estado das partículas antes da emissão de qualquer radiação; *dressed level* corresponde ao estado das partículas após irradiação.

6.7.1 Cálculo da seção de choque fiducial

A partir dos dados da Tab.6.3 e de acordo com a Eq.5.1 é possível calcular a seção de choque fiducial na região de alta massa transversa. O fundo total B é estimado através da soma da contribuição advinda da produção de $t\bar{t}$, *single top*, dibósons e decaimentos do bóson Z e de bósons W nos canais τ e μ . A incerteza sistemática na estimativa dos eventos de fundo é somada quadraticamente para cada contribuição e leva em conta apenas as incertezas experimentais descritas na Sec. 6.6, as incertezas associadas aos modelos teóricos (escala, renormalização, PDF etc.) não são consideradas. Finalmente, são consideradas as incertezas relativas à flutuação estatística dos dados e na determinação da luminosidade, conforme foi descrito no Cap. 5. A Tab. 6.5 apresenta o número de eventos medidos na produção de W^+ , W^- e $W^\pm$ no canal de elétrons na região de alta massa transversa.

Tabela 6.5: Número de candidatos observados para os eventos no canal de elétrons do decaimento do bóson W , fundo total (correspondente a soma dos processos eletrofracos e dos processos dos canais de produção de quarks top e di-bósons) e o número final de eventos de sinal correspondentes à diferença entre o número de candidatos iniciais e o fundo. As incertezas sistemáticas no fundo correspondem às incertezas sistemáticas experimentais

Canal	Candidatos observados	Fundo Total	Sinal medido
W^+	$1000950 \pm 1000(\text{stat})$	$278320 \pm 18925(\text{sys}) \pm 407(\text{stat}) \pm 5566(\text{lumi})$	$722630 \pm 18925(\text{sys}) \pm 1079(\text{stat}) \pm 5566(\text{lumi})$
W^-	$817943 \pm 904(\text{stat})$	$272787 \pm 18549(\text{sys}) \pm 354(\text{stat}) \pm 5455(\text{lumi})$	$545156 \pm 18549(\text{sys}) \pm 971(\text{stat}) \pm 5455(\text{lumi})$
$W^\pm$	$1818890 \pm 1348(\text{stat})$	$553298 \pm 37624(\text{sys}) \pm 563(\text{stat}) \pm 11065(\text{lumi})$	$1266702 \pm 37624(\text{sys}) \pm 1460(\text{stat}) \pm 11065(\text{lumi})$

Fator de correção C_W

As distribuições cinemáticas apresentados na Sec. 6.5 mostram que o gerador de Monte Carlo utilizado permite modelar os dados de forma satisfatória. Essa observação permite afirmar que é possível utilizar essa mesma amostra de sinal para o cálculo do fator C_W . Esse cálculo é feito através da amostra simulada de sinal nos canais $W^- \rightarrow e^-\nu$, $W^+ \rightarrow e^+\nu$ e $W^\pm \rightarrow e^\pm\nu$. Os eventos "verdadeiros" ($N_{\text{verdadeiros}}$) dessa amostra são selecionados dentro de uma região fiducial (separados por carga) definida pelos seguintes cortes cinemáticos, cf. Sec. 6.2:

- $p_T^e > 65 \text{ GeV}$;

- $|\eta_\ell| < 2,47$;
- $p_T^\nu > 65 \text{ GeV}$;
- $m_T > 130 \text{ GeV}$.

A incerteza no fator C_W (δC_W) é estimada através do número de eventos em cada uma das variações sistemáticas e reflete a migração dos eventos para dentro ou para fora dos critérios de seleção. Para cada uma das sistemáticas essa variação é avaliada e a variação final corresponde à soma quadráticas de todas as variações e, calculada dessa forma, tem um impacto de $\pm 2\%$ no valor nominal de C_W . Novamente, conforme já foi observado anteriormente, a sistemática associada à reconstrução de jatos é predominante, contribuindo com mais de 70% dessa incerteza.

A Tab. 6.6 apresenta os valores obtidos de C_W para a reconstrução dos canais W^+ , W^- e $W^\pm$ e a variação.

Tabela 6.6: Fatores de correção C_W para os canais W^+ , W^- e $W^\pm$.

Canal	$C_W \pm \delta C_W$
W^+	0.737 ± 0.015
W^-	0.716 ± 0.014
$W^\pm$	0.726 ± 0.015

Dessa forma, combinando o número de eventos medidos (Tab. 6.5), a luminosidade (Eq. 6.4) e o fator C_W (Tab. 6.6) a seção de choque fiducial para a produção de bósons W na região de alta massa é:

$$\sigma_W^{fid} \times BR(W \rightarrow l\nu) = \frac{(N - B)}{C_W \cdot \mathcal{L}} \quad (6.5)$$

Tabela 6.7: Seção de choque medida para os processos $W^- \rightarrow e^-\nu$, $W^+ \rightarrow e^+\nu$ e $W^\pm \rightarrow e^\pm\nu$.

Canal	$\sigma_W^{fid} \times BR(W \rightarrow e\nu) \text{ [pb]} \text{ (medida)}$
W^+	$12.81 \pm 0.42(sys) \pm 0.017(stat) \pm 0.27(lumi)$
W^-	$9.94 \pm 0.39(sys) \pm 0.017(stat) \pm 0.22(lumi)$
$W^\pm$	$22.78 \pm 0.82(sys) \pm 0.030(stat) \pm 0.50(lumi)$

O número de eventos produzidos como sinal pelo gerador PowhegPythia8 que sobrevivem aos cortes cinemáticos após sua reconstrução podem ser usados como

uma estimativa da seção de choque fiducial por esse modelo. A Tab. 6.8 apresenta os valores previstos para a seção de choque fiducial para os eventos simulados como sinal.

Tabela 6.8: Seção de choque prevista pelo gerador PowhegPythia8 para os canais $W^- \rightarrow e^- \nu$, $W^+ \rightarrow e^+ \nu$ e $W^\pm \rightarrow e^\pm \nu$.

Canal	$\sigma_W^{fid} \times BR(W \rightarrow e\nu)$ [pb] (PowhegPythia - NNLO)
W^+	$12.12 \pm 0.86(sys) \pm 0.038(stat) \pm 0.34(lumi)$
W^-	$9.35 \pm 0.66(sys) \pm 0.030(stat) \pm 0.26(lumi)$
$W^\pm$	$21.53 \pm 1.53(sys) \pm 0.049(stat) \pm 0.61(lumi)$

Em uma comparação entre os resultados apresentados nas Tab. 6.8 e Tab. 6.7 observa-se que a incerteza dominante corresponde à sistemática, seguida da luminosidade. Outra observação importante é que a medida experimental já apresenta uma incerteza menor que a previsão teórica calculada dessa forma. A redução da incerteza da luminosidade vai muito além do escopo dessa medida, bem como da maioria das medidas de precisão do ATLAS, mas existem procedimentos, como a medida da razão das seções de choque, que possibilitam eliminar a incerteza associada com a luminosidade, conforme será discutido posteriormente. Uma forma direta de reduzir a sistemática é limitar o espaço de fase onde a incerteza sistemática é melhor controlada (como, por exemplo, limitar a pseudo-rapidez à região do barril), mas com a implicação direta da perda da informação do processo físico na região dianteira. Medidas nessa região são de suma importância para processos VBF (*Vector Boson Fusion* [96] ou estudos de assimetria em pseudorapidez na produção de léptons para medidas do ângulo de Weinberg[97]. Logo, a única alternativa prática nesse cenário é a redução do fundo estimado. Foi visto na Tab. 6.4 que o fundo nessa região de alta massa é predominantemente devido à produção de pares de quarks $t\bar{t}$. As seções seguintes propõem uma estratégia para a redução desse fundo.

6.8 Estudo da redução do fundo de $t\bar{t}$ por análise multivariável

Os resultados da seção anterior mostram que após a seleção dos eventos na região de alta massa, baseada em simples cortes cinemáticos, ainda resta uma contribuição

importante do fundo proveniente da produção de pares $t\bar{t}$. Isso é devido ao quark *top* decair principalmente em um quark b e um bóson W , que passa pelos critérios de seleção ao decair em um par elétron-neutrino. Outro problema é a incerteza associada à modelagem de $t\bar{t}$ na região de alta massa. Embora essa seja uma região normalmente explorada por análises que buscam novas partículas, como por exemplo a buscas da produção W' [98], tais análises não são tão sensíveis à modelagem incorreta do fundo e $t\bar{t}$. No caso deste trabalho, cujo objetivo é a medida de seção de choque, a rejeição desse fundo é uma etapa importante.

De forma geral, as análises que apresentam um quark b no estado final (como ocorre com a produção de quarks t) utilizam métodos de etiquetamento de jatos de sabores pesados (*b-tagging*) para suprimir eventos contendo jatos b (num processo chamado *b-veto*) [99]. Esses métodos baseiam-se no fato dos quarks b decaírem em hádrons, e devido ao tempo próprio do decaimento, apresentam um vértice deslocado com relação ao vértice primário da colisão. Tais métodos permitem identificar jatos de sabores pesados com uma eficiência de 70% com uma contaminação de 20%³. Entretanto, os métodos de *b-tagging* representam as principais fontes de incertezas sistemáticas nas análises que os empregam [99] chegando a mais de 10%. Assim, a fim de evitar propagar uma nova incerteza sistemática na região de alta massa, decidiu-se não usar o etiquetamento de jatos de sabores pesados nessa região, explorando-se em primeira instância apenas as variáveis cinemáticas já descritas na Sec. 6.5, mas agora empregando métodos de análise multi-variável para buscar discriminantes capazes de reduzir o fundo devido à produção de pares $t\bar{t}$. Deve-se mencionar que, devido à sua baixa seção de choque, a produção de quarks *top* isolados não representa um fundo significativo.

Análise multivariável

A utilização de métodos de análise multi-variável, como parte da técnica mais geral chamada de aprendizado de máquina [100], vem encontrando um domínio de aplicação cada vez maior em praticamente todas as áreas da física de altas energias: na reconstrução de objetos (jatos, elétrons, múons), na separação entre fundo e sinal,

³dependendo do ponto de trabalho, pode-se optar por uma maior eficiência com menor pureza ou maior pureza com menor eficiência [99]

na calibração e no reconhecimento de padrões topológicos em detectores segmentados, supressão de empilhamento em alta luminosidade etc. O objetivo principal destes métodos é explorar padrões embutidos nos dados a partir de um conjunto de variáveis de entrada que são utilizadas para treinar um modelo de aprendizado utilizando eventos previamente classificados como sinal ou fundo. Operacionalmente, O espaço de fase multidimensional formado por essas variáveis permite que um evento fora da amostra de treinamento possa ser classificado como sinal ou fundo.

Existem diversas implementações desses algoritmos, por exemplo o ROOT-TMVA [101], o XGBOOST [102] e o TensorFlow [103]. Dentre as opções, o ROOT-TMVA apresenta a vantagem de uma implementação que é próxima a representação e organização dos dados em física de altas energias e por esse motivo é a implementação utilizada nesta análise.

Além disso, dentro de cada implementação, existem vários métodos que podem ser empregados para a classificação de eventos em análise multi-variável. Os mais comumente empregados em física de altas energias são os discriminantes de Fisher e os *Boosted Decision Trees* (BDTs). O Ap. B apresenta de forma resumida uma descrição desses métodos.

6.8.1 Implementação dos métodos de classificação multi-variável

Conforme foi mencionado, foi utilizado o ROOT-TMVA para o implementar o treinamento e classificação dos eventos em fundo ou sinal. Cinco métodos foram explorados: discriminantes de Fisher (**Fisher**), discriminantes de Fisher com transformação gaussiana das variáveis de entrada (**FisherG**), *Boosted Decision Tree* adaptivo (**BDT**), *Boosted Decision Tree* com *Gradient Boosting* (**BDTG**) e *Boosted Decision Tree* adaptivo com decorrelação (**BTDD**).

O treinamento foi feito com 20 mil eventos das distribuições de sinal de Monte Carlo correspondentes à produção inclusiva de bósons W^+ e W^- produzidas pelo gerador de eventos PowHeg juntamente com o Pythia8 para a hadronização, conforme já discutido anteriormente. O sinal inclui tanto a produção na região inclusiva como as amostras da região de alta massa. Os pesos, conforme descritos na Sec. 6.3, são também considerados nos algoritmos de treinamento. Isso é particularmente

importante para os intervalos da região de alta massa, uma vez que a distribuição final leva em conta diferentes seções de choque de cada intervalo de massa. Da mesma forma, o fundo utilizou 20 mil eventos da amostra de $t\bar{t}$ para treinamento. Os eventos foram pré-selecionados de acordo com os cortes definidos na Sec 6.2.

As seguintes variáveis foram utilizadas como entrada para o processo de discriminação: E_T^{miss} , p_T^l , m_T , $\Delta\phi$. O *ranking* das variáveis (correspondente ao seu poder de discriminação num dado *ensemble*, conforme descrito no Apêndice B) é obtido através de cada método de MVA e é apresentada na Tabela 6.9.

Tabela 6.9: *Ranking* das variáveis cinemáticas MET , p_T^l , m_T e $\Delta\phi$ obtidas em cada um dos métodos de MVA. Variáveis com maior poder discriminatório são classificadas com o *rank* = 1

Rank	Fisher	FisherG	BDT	BDTG	BDTD
1	E_T^{miss}	MET	$\Delta\phi$	$\Delta\phi$	$\Delta\phi$
2	p_T^l	p_T^l	MET	m_T	MET
3	m_T	m_T	m_T	p_T^l	m_T
4	$\Delta\phi$	$\Delta\phi$	p_T^l	MET	p_T^l

A saída de cada um dos cinco classificadores é apresentada na Figura 6.19. A figura mostra que utilizando apenas as variáveis cinemáticas descritas no parágrafo anterior existe pouco poder de discriminação entre fundo e sinal (existe uma região grande de superposição, ou a escolha de um ponto de corte para eliminar o fundo afeta de forma significativa a eficiência (ou pureza) para a detecção do sinal, algo que também pode ser observado na curva característica de operação do receptor ou *Receiver Operating Curve* - ROC da Fig. 6.20).

Figura 6.19: Distribuições de discriminantes de saída calculadas pelos cinco métodos de MVA estudados, considerando apenas as variáveis cinemáticas da reconstrução do bóson W no conjunto de variáveis de entrada: (a) discriminantes de Fisher, (b) discriminantes de Fisher com transformação gaussiana das variáveis de entrada, (c) *Boosted Decision Tree* adaptivo, (d) *Boosted Decision Tree* com *Gradient Boosting* e (e) *Boosted Decision Tree* adaptivo com decorrelação. Nota-se uma significativa sobreposição das distribuições de sinal e fundo presente em todos os métodos.

Figura 6.20: Curvas características de operação do receptor (ROC) para os cinco métodos de MVA estudados, considerando apenas as variáveis cinemáticas da reconstrução do bóson W no conjunto de variáveis de entrada.

Inclusão da multiplicidade de jatos como variável discriminante

A utilização de etiquetamento de jatos pesados tem um grande poder de discriminação [95] mas, como já foi discutido no início dessa seção, está associado a um aumento importante da incerteza sistemática. Levando isso em consideração, foi explorada uma outra variável associada à atividade hadrônica do evento: a multiplicidade de jatos (n_{Jets}). A Fig. 6.21 mostra a multiplicidade de jatos na região de alta massa.

Figura 6.21: Multiplicidade de jatos na região de alta massa para a produção de bósons $W^\pm$. O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa rosa e somada quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos). (a) escala linear, (b) escala logarítmica

O mesmo processo de treinamento envolvendo a amostra de sinais e de fundo devido à produção de pares $t\bar{t}$ foi utilizado incluindo n_{Jets} como variável de entrada, além das variáveis E_T^{miss} , p_T^e , m_T , $\Delta\phi$ e a Tab. 6.10 mostra a classificação ou *ranking* das variáveis. Nessa tabela, exceto pelo BDTG, todos os métodos mostram que a multiplicidade de jatos possui o maior valor de discriminação, o que reforça a proposta de inclusão dessa variável. Uma análise das curvas de eficiência da Fig. 6.25 mostra que esse método (BDTG) têm um comportamento anômalo se comparado

aos outros métodos (não há um ponto de transição entre a curva de pureza do sinal e a eficiência de corte do fundo) e portanto sua informação deve ser utilizada com certo critério. A Fig. 6.22 mostra a saída dos discriminantes com a inclusão dessa variável, onde observa-se um aumento significativo da capacidade de separação entre o sinal e fundo o que também pode ser observado na curva ROC da Fig. 6.23.

Tabela 6.10: *Ranking* das variáveis cinemáticas MET , p_T^l , m_T , $\Delta\phi$ e $nJets$ obtidas em cada um dos métodos de MVA. Variáveis com maior poder discriminatório são classificadas com o $rank = 1$.

Rank	Fisher	FisherG	BDT	BDTG	BDTD
1	$nJets$	$nJets$	$nJets$	$\Delta\phi$	$nJets$
2	MET	MET	$\Delta\phi$	$nJets$	$\Delta\phi$
3	p_T^l	p_T^l	MET	p_T^l	MET
4	m_T	m_T	m_T	MET	p_T^l
5	$\Delta\phi$	$\Delta\phi$	p_T^l	m_T	m_T

Figura 6.22: Distribuições de saída calculadas pelos cinco métodos de MVA estudados, incluindo a multiplicidade de jatos juntamente com as variáveis cinemáticas da reconstrução do bóson W no conjunto de variáveis de entrada. (a) discriminantes de Fisher, (b) discriminates de Fisher com transformação gaussiana das variáveis de entrada, (c) *Boosted Decision Tree* adaptivo, (d) *Boosted Decision Tree* com *Gradient Boosting* e (e) *Boosted Decision Tree* adaptivo com decorrelação. Comparado à Fig. 6.19 observa-se uma significativa melhora na capacidade de discriminação após a inclusão da multiplicidade de jatos.

Figura 6.23: Curvas características de operação do receptor (ROC) para os cinco métodos de MVA estudados, incluindo a multiplicidade de jatos juntamente com as variáveis cinemáticas da reconstrução do bóson W no conjunto de variáveis de entrada.

A Fig. 6.24 apresenta a matriz de correlação das variáveis de entrada utilizadas pelos métodos de MVA para o sinal e o fundo. Duas importantes características são observadas na Fig. 6.24: i) a multiplicidade de jatos não apresenta correlação com as outras variáveis (o que novamente reforça a proposta de discriminação dessa nova variável) e ii) a energia faltante (E_T^{miss}), calculada no sinal apresenta uma correlação significativamente maior com o p_T reconstruído do elétron se comparada à matriz de correlação do fundo, uma vez que são restritos à massa da reconstrução do bóson W . Essa correlação é menor na amostra de $t\bar{t}$ uma vez que a (E_T^{miss}), por uma quantidade global, envolve também a E_T^{miss} do bóson W oriundo do decaimento dos quarks t e $\bar{t}$.

Figura 6.24: (a) Matriz de correlação das variáveis de entrada na amostra de sinal ($W^\pm$) e (b) na amostra de fundo ($t\bar{t}$).

6.8.2 Determinação do ponto de operação da MVA e distribuições cinemáticas

A determinação do ponto de operação ótimo para a separação entre o sinal e o fundo para cada um dos 5 métodos de MVA utilizados é feita de tal forma de que a significância s

$$s = S/\sqrt{S+B} \quad (6.6)$$

(onde S corresponde ao sinal e B o fundo) é máxima para esse ponto [100]. A Fig. 6.25 mostra a eficiência para a seleção do sinal, a eficiência na rejeição do fundo, a pureza do sinal, o produto da eficiência e da pureza do sinal e a significância para cada um dos cinco métodos avaliados. Embora todos os métodos atinjam a significância máxima (v. Tab. 6.11) os métodos de BDT permitem uma maior separação entre o fundo e o sinal (evidenciado pelas curvas de eficiência de sinal e fundo) quando comparados aos métodos de Fisher. Isso permite que eventuais *outliers* sejam adequadamente classificados mesmo na presença de uma flutuação estatística significativa, como é o caso na região de alta massa.

	Fisher	FisherG	BDT	BDTG	BDTD
$S/\sqrt{S+B}$	30.1	29.9	30.19	30.16	30.03
Ponto de corte	0.02	-0.09	-0.11	0.14	-0.03

Tabela 6.11: Significância máxima e respectivo ponto de corte onde a significância é máxima para cada um dos cinco métodos de MVA.

Observando-se as curvas de eficiência da Fig. 6.25 nota-se que o método de BDT e o BDTG permitem variar o ponto de trabalho ao redor da máxima significância de forma a selecionar uma amostra mais pura ou mais eficiente, enquanto as mesmas curvas apresentam variações mais abruptas de eficiência. No ponto de máxima significância existe uma queda importante da eficiência do sinal, tanto no BDT como no BDTD, o que pode afetar o cálculo da seção de choque. Assim, escolheu-se como ponto de trabalho efetuar o corte no ponto -0,2, o que equivale a uma amostra com 82% de pureza e 18% de contaminação pelo fundo. O método de BDTD pode então ser aplicado aos dados e aos eventos simulados, e as distribuições cinemáticas (agora incluindo a multiplicidade de jatos) utilizando esse corte são apresentadas nas

Figura 6.25: Curvas de eficiência para a seleção do sinal, a eficiência na rejeição do fundo, a pureza do sinal, o produto entre a eficiência e pureza do sinal e a significância para cada um dos cinco métodos de MVA.

Figs. 6.26, 6.27, 6.28 e 6.29 e nas Figs. 6.30, 6.31 e 6.32 considerando a separação dos eventos pelo sinal da carga do elétron. Uma observação interessante advém da análise das Figs. 6.12 e 6.27: após a redução do fundo devido à produção de pares $t\bar{t}$, existe uma perceptível melhora na concordância com as previsões, o que pode ser um importante indicador da modelagem incorreta do m_T^W na amostra de $t\bar{t}$ na região de alta massa. Após a aplicação do método BTDD no ponto de trabalho descrito no parágrafo anterior (-0.2) observou-se efetivamente uma redução de 65% no número

de eventos da amostra de $t\bar{t}$ e de 5% no número de eventos da amostra de sinal.

Figura 6.26: Distribuição do módulo da diferença de ângulo azimutal entre o elétron e o vetor $\vec{E}_T^{\text{miss}}$ ($\Delta\phi$) (a,b) e pseudorapidez dos elétrons (c,d) na **região de alta massa transversa** para produção de $W^\pm$ após o corte de MVA. O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa rosa e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos). (a) escala linear, (b) escala logarítmica.

Figura 6.27: Distribuição da energia transversa faltante (E_T^{miss}) (a,b) e e massa transversa (m_T^W) (c,d) na *região de alta massa transversa* para produção de $W^\pm$ após o corte de MVA. O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa rosa e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos). (a,c) escala linear, (b,d) escala logarítmica.

Figura 6.28: Distribuição de momento transverso do elétron (p_T^e) (a,b) e multiplicidade de jatos (c,d) para produção de $W^\pm$ na *região de alta massa transversa* após o corte de MVA. O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa rosa e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos). (a,c) escala linear, (b,d) escala logarítmica.

Figura 6.29: Distribuição do número médio de colisões ($\langle\mu\rangle$) na *região de alta massa transversa* para produção de $W^\pm$ após o corte de MVA. O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa rosa e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos).

Figura 6.30: Distribuição do módulo da diferença de ângulo azimutal entre o elétron e o vetor $\vec{E}_T^{\text{miss}}$ ($\Delta\phi$) (a,b) e pseudorapidez dos elétrons (c,d) na **região de alta massa transversa** após o corte de MVA. O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa rosa e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos). (a,c) W^- , (b,d) W^+ .

Figura 6.31: Distribuição da energia transversa faltante (E_T^{miss}) (a,b) e massa transversa (m_T^W) (c,d) na *região de alta massa transversa* após o corte de MVA. O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa rosa e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos). (a,c) W^- , (b,d) W^+ .

(a) W^-

(b) W^+

(c) W^-

(d) W^+

Figura 6.32: Distribuição do momento transversal do elétron (p_T^e) (a,b) e multiplicidade de jatos (c,d) na *região de alta massa transversa* após o corte de MVA. O sinal e as contribuições esperadas de todos os fundos são estimadas através de simulações de Monte Carlo. As incertezas sistemáticas para o sinal e o fundo estão indicadas pela região delimitada pela faixa rosa e somadas quadraticamente com a incerteza estatística. As variações estatísticas dos dados são apresentadas na forma de barras de erro associadas aos pontos. O painel inferior apresenta a razão entre os dados e a previsão de Monte Carlo (formada pela soma do sinal e dos fundos). (a,c) W^- , (b,d) W^+ .

Cálculo da seção de choque após redução do fundo devido a produção de pares $t\bar{t}$

Uma vez que o método de MVA aplicado foi capaz de reduzir significativamente o fundo devido à produção de pares $t\bar{t}$, a seção de choque fiducial pode ser novamente calculada após a seleção dos eventos pelo MVA (Tab. 6.12 e 6.13).

Tabela 6.12: Seção de choque, após a aplicação do MVA, medida para os processos $W^- \rightarrow e^- \nu$, $W^+ \rightarrow e^+ \nu$ e $W^\pm \rightarrow e^\pm \nu$.

Canal	$\sigma_W^{fid} \times BR(W \rightarrow e\nu)$ [pb] (Medida, MVA)
W^+	$13,00 \pm 0,32(sys) \pm 0,018(stat) \pm 0,26(lumi)$
W^-	$10,02 \pm 0,26(sys) \pm 0,016(stat) \pm 0,21(lumi)$
W	$23,04 \pm 0,59(sys) \pm 0,024(stat) \pm 0,47(lumi)$

Tabela 6.13: Seção de choque, após a aplicação do MVA, prevista pelo gerador PowhegPythia8 para os canais $W^- \rightarrow e^- \nu$, $W^+ \rightarrow e^+ \nu$ e $W^\pm \rightarrow e^\pm \nu$

Canal	$\sigma_W^{fid} \times BR(W \rightarrow e\nu)$ [pb] (PowhegPythia - NNLO, MVA)
W^+	$12,18 \pm 0,86(sys) \pm 0,038(stat) \pm 0,34(lumi)$
W^-	$9,29 \pm 0,66(sys) \pm 0,030(stat) \pm 0,26(lumi)$
W	$21,53 \pm 1,53(sys) \pm 0,049(stat) \pm 0,61(lumi)$

A aplicação dos cortes utilizando o MVA com o ponto de trabalho selecionado permite uma redução de 28% das incertezas sistemáticas associadas à medida das seções de choque fiduciais para a produção $W^\pm$, 24% para o canal W^+ e 34% para o canal W^- .

Razão das seções de choque $\sigma_{W^+}/\sigma_{W^-}$

A partir dos resultados obtidos após a aplicação do MVA na Tab. 6.12, a razão $\sigma_{W^+}/\sigma_{W^-}$ pode ser calculada (a partir da Eq.5.1):

$$\frac{\sigma_{W^+}^{fid}}{\sigma_{W^-}^{fid}} = \frac{N_{sig}^{W^+}}{\mathcal{L}C_{W^+}} \times \frac{\mathcal{L}C_{W^-}}{N_{sig}^{W^-}} = \frac{N_{sig}^{W^+} C_{W^-}}{N_{sig}^{W^-} C_{W^+}}. \quad (6.7)$$

Um aspecto importante é que na razão $\sigma_{W^+}/\sigma_{W^-}$, a incerteza na luminosidade é cancelada. Assim, de acordo com a Eq. 6.7 e a Tab. 6.12 a razão entre as seções de choque fiduciais é:

$$\frac{\sigma_{W^+}^{fid}}{\sigma_{W^-}^{fid}} = 1,297 \pm 0,037(sys) \pm 0,004(stat) \quad (6.8)$$

e pode ser comparada com a previsão fornecida pela combinação dos geradores Powheg e Pythia:

$$\frac{\sigma_{W^+}^{fid}}{\sigma_{W^-}^{fid}} = 1,311 \pm 0,107(sys) \pm 0,005(stat). \quad (6.9)$$

Capítulo 7

Conclusões

As medidas de seção de choque de bósons $W^\pm$ a partir dos dados coletados em 2015 mostraram a possibilidade de produzir medidas de precisão a partir de um conjunto de dados obtidos em apenas 2 semanas de operação do LHC no regime de $\sqrt{s} = 13$ TeV. A estimativa do fundo devido à produção de multijatos (predominante na região inclusiva), foi o principal desafio dessa análise, uma vez que tal fundo não é bem descrito pelos modelos teóricos e devido a sua magnitude (da ordem de 10%) contribui de forma significativa para a incerteza sistemática. A utilização de um método *data driven* para a estimativa de multijatos no canal de elétrons permitiu a obtenção da seção de choque fiducial e total na região inclusiva compatível com o previsto para o modelo padrão e com uma incerteza comparável à incerteza das previsões teóricas. A razão entre as seções de choque fiduciais de W^+ e W^- foram beneficiadas pelo cancelamento da incerteza na luminosidade, atingindo uma incerteza de apenas 0,8%, o que torna essa razão muito útil para testar os modelos teóricos implementados nos geradores assim como os diversos conjuntos de PDFs.

Os resultados dessa medida, combinados com os resultados dos canais $W^\pm \rightarrow \mu^\pm \nu$, $Z \rightarrow e^+ e^-$ e $Z \rightarrow \mu^+ \mu^-$, como mostrados no Ap. A e publicados em [58], mostram boa concordância com as previsões teóricas do Modelo Padrão. Conforme o esperado, a razão $\sigma_{W^+}^{fid}/\sigma_{W^-}^{fid}$ obtida da combinação dos canais de elétrons e múons apresenta as incertezas sistemáticas e estatísticas inferiores àquelas obtidas usando apenas elétrons. Além do mais, demonstram a capacidade do experimento na discriminação de conjuntos de PDFs, visto que as razões entre seções de choque $\sigma_{W^+}^{fid}/\sigma_{W^-}^{fid}$ e $\sigma_{W^\pm}^{fid}/\sigma_Z^{fid}$ calculadas com o uso de alguns conjuntos de PDFs se afastam dos valores

medidos nessa análise.

Já na região de alta massa, que tornou-se acessível graças à grande quantidade de dados obtida a $\sqrt{s} = 13$ TeV pelo LHC nos últimos anos, o principal desafio é a redução da contribuição do fundo devida à produção de pares $t\bar{t}$. Utilizando cortes simples, esse fundo representa mais de 24% dos eventos e devido às incertezas no modelamento da produção de quarks top na região de alta massa é imprescindível buscar métodos alternativos para a identificação de eventos originados por pares $t\bar{t}$, sem sacrificar de forma importante a abrangência do espaço de fase, o que levou à investigação de métodos de análise multivariável para a redução de fundo. O método que se mostrou mais promissor para essa análise foi o *Boosted Decision Tree* com decorrelação (BDTD), e a introdução de uma variável hadrônica (multiplicidade dos jatos) como variável discriminante teve um papel importante na aplicação dos métodos de MVA, reduzindo esse fundo em aproximadamente 65% com uma perda de eficiência no sinal de apenas 5%. Uma melhora na distribuição de massa transversa (m_T^W) após a redução do fundo devido à produção de pares $t\bar{t}$ sinaliza uma possível imprecisão na modelagem desse fundo. As distribuições cinemáticas apresentam uma concordância satisfatória entre as previsões teóricas e os dados, assim como os valores para a seção de choque fiducial medidos na região de alta massa transversa mostram boa concordância com os valores calculados, conforme pode ser observado nos resultados apresentados nas Tabelas 6.12 e 6.13. A razão entre as seções de choque de produção de W^+ e W^- também foi obtida na região de alta massa e o resultado (Eq. 6.7) possui um valor central que concorda com a previsão teórica do Modelo Padrão, além de apresentar menos da metade da incerteza teórica.

Apêndice A

Combinação dos resultados com os canais de múons e Z

Os resultados da análise em baixa luminosidade, conforme já mencionado, foram combinados com o canal de múons, assumindo a universalidade dos léptons, e os resultados dos canais $Z \rightarrow e^+e^-$ e $Z \rightarrow \mu^+\mu^-$ também foram combinados. A combinação dos resultados dos dois léptons tem a vantagem de diminuir a incerteza estatística relativa devido ao maior número de eventos usados. Os valores das seções de choque medidas para os canais de W^+ , W^- , $W^\pm$ e Z estão disponíveis na Tab. A.1, com as respectivas incertezas sistemáticas, estatísticas e de luminosidade. A Fig. A.1 mostra a razão entre as seções de choque fiduciais dos canais de elétrons e múons, tanto para $W^\pm$ como Z , apresentando uma boa concordância com os valores médios do PDG [14], levando em conta as barras de erro.

Tabela A.1: Seções de choque calculadas combinando-se os léptons com as respectivas incertezas estatísticas, sistemáticas e de luminosidade.

Canal	$\sigma^{fid} \pm (sys) \pm (stat) \pm (lumi)$ [nb]	$\sigma^{tot} \pm (sys) \pm (stat) \pm (lumi)$ [nb]
W^-	$3,5 \pm 0,01 \pm 0,07 \pm 0,07$	$8,79 \pm 0,02 \pm 0,24 \pm 0,18$
W^+	$4,53 \pm 0,01 \pm 0,09 \pm 0,10$	$11,83 \pm 0,02 \pm 0,32 \pm 0,25$
$W^\pm$	$8,03 \pm 0,01 \pm 0,16 \pm 0,17$	$20,64 \pm 0,01 \pm 0,55 \pm 0,43$
Z	$0,779 \pm 0,003 \pm 0,006 \pm 0,016$	$1,981 \pm 0,007 \pm 0,038 \pm 0,042$

A razão entre as seções de choque fiduciais de W^+ e W^- foi calculada a partir dos resultados dos dois léptons combinados, e da mesma forma foi calculada a razão entre as seções de choque de $W^\pm$ e Z , também fazendo uso da combinação dos dois léptons. Conforme já mencionado, a incerteza da luminosidade não afeta a

Figura A.1: Razão entre os canais de elétrons e múons para as seções de choque fiduciais de produção de bósons $W^\pm$ e Z , comparados com as previsões do Modelo Padrão e as bandas médias do PDG [14].

razão entre as seções de choque. O valor medido para a razão $\sigma_{W^+}^{fid}/\sigma_{W^-}^{fid}$ foi $1,295 \pm 0,003(stat) \pm 0,010(sys)$ e para a razão $\sigma_{W^\pm}^{fid}/\sigma_Z^{fid}$ foi $10,31 \pm 0,04(stat) \pm 0,20(sys)$. Esses resultados também são mostrados graficamente na Fig. A.2, que também mostra os resultados previstos para os diversos conjuntos de PDFs.

Figura A.2: Razões entre seções de choque fiduciais $\sigma_{W^+}^{fid}/\sigma_{W^-}^{fid}$ e $\sigma_{W^\pm}^{fid}/\sigma_Z^{fid}$, representadas pelas linhas vermelhas, e as razões calculadas teoricamente usando diversos conjuntos de PDF's, com as bandas de incertezas estatísticas e totais.

As razões entre as seções de choque experimentais obtidas para cada um dos canais fazendo uso da combinação de léptons e entre as seções de choque teóricas calculadas com o uso de diferentes conjuntos de PDFs também foram calculadas.

Os resultados estão disponíveis na Fig. A.3.

Figura A.3: Razão entre as razões das seções de choque fiduciais teóricas e medidas para cada bóson, usando diversos conjuntos de PDFs.

Os resultados combinados finais mostram que o aparato experimental tem grande poder de discriminação entre conjuntos de PDFs e modelos teóricos. Mesmo com uma luminosidade integrada limitada, as incertezas predominantes foram as incertezas sistemáticas e as bandas de incerteza dos resultados apresentados já são da mesma ordem das incertezas teóricas. Com esses resultados é possível observar quais PDFs descrevem melhor as medidas experimentais e, portanto, são mais adequadas para serem usadas em simulações futuras.

Apêndice B

Métodos de análise multi-variável

Este apêndice descreve os conceitos envolvidos e os métodos de análise multivariável utilizados neste trabalho.

B.1 Discriminantes de Fisher

O método de classificação por discriminantes de Fisher classificam os eventos a partir dos valores médios da amostra total ($\bar{x}_k$), do sinal e do fundo ($\bar{x}_{S(B),k}$) e da matriz de covariância C , utilizando uma transformação das variáveis cinemáticas k que formam o espaço fase. Essa transformação determina um eixo separando o espaço de fase de tal forma a maximizar a separação dos grupos de eventos classificados como sinal e como fundo e minimizar nesse espaço o volume formado pelo agrupamento de eventos de mesmo tipo (sinal ou fundo).

O método do discriminante de Fisher requer que as variáveis sigam uma distribuição gaussiana. Por não ser esse o caso para a maioria das variáveis, uma modificação do método aplica uma transformação nas variáveis de entrada antes de sua de-correlação de forma que sua distribuição se torne gaussiana. Isso é feito primeiro transformando a distribuição cumulativa da variável numa distribuição uniforme e após essa etapa utilizando a função de erro inversa para transformar essa distribuição em uma distribuição gaussiana de média zero e largura unitária. Como apenas uma das classes de eventos (sinal ou fundo) pode ser transformada, a escolha foi aplicar a transformação no fundo.

Uma medida importante do poder de discriminação de cada variável de entrada utilizada pelo algoritmo é dada pela relação B_{kk}/C_{kk} , a qual é utilizada no *ranking*

das variáveis, onde

$$B_{k,l} = \frac{1}{2} \sum_{U=S,B} (x_{U,k} - \bar{x}_{U,k})(x_{U,l} - \bar{x}_{U,l}) \quad (\text{B.1})$$

e

$$C_{k,l} = \sum_{U=S,B} \langle x_{U,k} - \bar{x}_{U,k} \rangle \langle x_{U,l} - \bar{x}_{U,l} \rangle \quad (\text{B.2})$$

A Figura B.1 ilustra o conceito mostrando que o discriminante é calculado de tal forma que a projeção de uma distribuição de sinal (S) e fundo (B) de duas variáveis hipotéticas num eixo formado pela direção desse discriminante no espaço de fase formado pelas duas variáveis é tal que a separação entre as distribuições é maximizada e a dispersão minimizada. [100],[101].

Figura B.1: Figura conceitual da aplicação do método dos discriminantes de Fisher mostrando a distribuição de sinal (S) e fundo (B) de duas variáveis hipotéticas e a distribuição de um discriminante de Fisher (y_{Fisher}), que pode ser entendida como a projeção dos dados num eixo cuja orientação é calculada de forma a maximizar a separação entre os valores médios das distribuições de sinal e fundo enquanto minimiza a dispersão de cada distribuição projetada [100].

B.2 *Boosted Decision Trees*

O método de classificação por BDT constrói uma coleção de árvores de decisão individuais separando de forma binária a amostra utilizando uma única variável por

vez. O *boosting* é parte de uma técnica muito utilizada e eficiente para melhorar a estabilidade da classificação face à variação estatística dos dados de treinamento. Nesse treinamento, uma sequência de algoritmos de MVA são aplicados ao evento que é classificado e recebe um peso como sinal ou fundo de acordo com a resposta majoritária dos algoritmos. O treinamento de uma árvore decisória é definido pela forma de separação dos dados. Num primeiro passo, todos os eventos de treinamento são classificados em sinal ou fundo de acordo com um valor de corte para uma das variáveis. No próximo passo, os dois grupos de eventos são novamente classificados utilizando outra variável e outro valor de corte. Isso ocorre sucessivamente até que o ramo final contenha um número de eventos pré-determinado. Nesse ponto, o ramo é classificado como fundo ou sinal de acordo com o tipo majoritário de eventos contido nele. A cada passo, a variável e o valor de corte que proporciona a maior separação para a amostra é escolhida pelo algoritmo. Uma vantagem do BDT é sua facilidade de identificar um sobre-treinamento (ou vício) do método: nesse caso, um ramo final classificado como fundo contém apenas eventos de fundo e como sinal, apenas eventos de sinal. Nesse caso, uma "poda" pode ser feita na árvore decisória, removendo os nós com pequena significância estatística após a árvore atingir o seu tamanho máximo. Entretanto, a utilização de classificadores fracos evita o sobre-treinamento da árvore e em consequência, a necessidade da poda. A Figura B.2 mostra o princípio desse método.

No BDT, o *ranking* da variável é obtido através da sua frequência de utilização e por quão boa é a separação obtida com essa variável em cada nó.

Três tipos de algoritmos de *boosting* foram testados :

- *Adaptive Boosting*: eventos que foram classificados de forma incorreta em um determinado estágio do treinamento da árvore recebem um peso maior quando passam para o próximo estágio. Por esse processo de peso, as árvores produzidas tendem a ser menores e reduzem a chance de sobre-treinamento [104].
- *Gradiente Boosting*: Os algoritmos de *boosting* requerem a determinação de uma função de perda, que deve ser uma função diferenciável e no caso do *adaptive boosting* tem uma forma exponencial. Essa função exponencial causa

Figura B.2: Princípio de funcionamento do método de classificação por *Boosted Decision Trees*: A árvore é construída através de uma sequência de separações. Em cada etapa, a variável que melhor discrimina entre sinal e fundo é escolhida pelo algoritmo classificador quando um corte c_i é aplicado. No final, cada folha é classificada como sinal (S) ou fundo (B) e a árvore construída (ou treinada) dessa forma pode ser utilizada para classificar os eventos da amostra de dados.

perda de robustez do classificador para eventos classificados de forma incorreta (em virtude por exemplo de uma flutuação estatística alta). O método de *gradient boosting* modifica essa função de perda por uma função mais robusta [101]. Entretanto, o método de classificação torna-se mais elaborado, exigindo o cálculo do gradiente da função de perda em cada nó e ajustando cada ramo novo. Esse processo iterativo permite construir um conjunto de árvores que minimizam a função de perda.

- *Decorrelation + Adaptive Boosting*: No caso das variáveis de entrada possuírem alto grau de correlação, pode ser útil empregar uma transformação linear a uma dessas variáveis que permita sua de-correlação. Essa é uma transformação linear aplicada às variáveis que apresentam alto grau de correlação e é feita numa etapa anterior ao treinamento da árvore (pré-processamento).

Bibliografia

- [1] HALZEN, F., MARTIN, A. D. **Quarks and leptons:** an introductory course in modern particle physics. New York: John Wiley & Sons, 1984.
- [2] GRIFFITHS, D. **Introduction to Elementary Particles.** 2 ed, Weinheim: Wiley-Vch, 2008.
- [3] PESKIN, M. E., SCHOROEDER, D. V. **An Introduction to quantum field theory.** Boca raton: CRC Press, 1995.
- [4] Evans, L., Bryant, P., The CERN Large Hadron Collider: Accelerator and Experiments, **Journal of Instrumentation**, v. 3, n. 08, p. S08001, 2008.
- [5] ATLAS Collaboration, Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC, **Physics Letters B**, v. 716, n. 1, p. 1-29, 2012.
- [6] CMS Collaboration, Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC, **Physics Letters B**, v. 716, n. 1, p. 30–61, 2012.
- [7] The LHCb Collaboration, Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \rightarrow J/\psi K^- p$ Decays, **Physical Review Letters B**, v. 115, n. 7, p. 072001, 2015.
- [8] The ALICE Collaboration, Precision measurement of the mass difference between light nuclei and anti-nuclei, **Nature Physics**, v. 11, n. 10, 2015.
- [9] FARINA, M. et al., Energy helps accuracy: electroweak precision tests at hadron colliders, arXiv:1609.08157 [hep-ph].

- [10] MEYERS, S., PICASSO, E. The design, construction and commissioning of the CERN large Electron–Positron collider, **Contemporary Physics**, v. 31, n. 6, p. 387–403, 1990.
- [11] TENCHINI, R., VERZEGNASSI, C., **The Physics of the Z and W Bosons**. Singapore: World Scientific Publishing, 2008.
- [12] ABE, F. et al. Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab, **Physical Review Letters**, v. 74, n. 14, p. 2626–2631, arXiv:hep-ex/9503002 [hep-ex].
- [13] Super-Kamiokande Collaboration, FUKUDA, Y. et al., Evidence for oscillation of atmospheric neutrinos, **Physical Review Letters**, v. 81, n. 8, p. 1562–1567, 1998, arXiv:hep-ex/9807003 [hep-ex].
- [14] OLIVE, K. A. et al., Review of Particle Physics, **Chinese Physics C**, v. 38, n. 9, p. 090001, 2014.
- [15] GLASHOW, S. L., Partial-symmetries of weak interactions, **Nuclear Physics**, v. 22, n. 4, p. 579–588, 1961, <http://www.sciencedirect.com/science/article/pii/0029558261904692>.
- [16] SALAM, A., WARD, J., Electromagnetic and weak interactions, **Physics Letters**, v. 13, n. 2, p. 168–171, 1964, <http://www.sciencedirect.com/science/article/pii/0031916364907115>.
- [17] WEINBERG, S., A Model of Leptons, **Physical Review Letters**, v. 19, p. 1264–1266, 1964, <https://link.aps.org/doi/10.1103/PhysRevLett.19.1264>.
- [18] KOBAYASHI, M., MASKAWA, T., CP-Violation in the Renormalizable Theory of Weak Interaction, **Progress of Theoretical Physics**, v. 49, n. 2, p. 652–657, 1973.
- [19] HIGGS, P., Broken Symmetries and the Masses of Gauge Bosons, **Physical Review Letters**, v. 13, n. 16, p. 508–509, 1964.

- [20] ENGLERT, F., BROUT, R., Broken Symmetry and the Mass of Gauge Vector Mesons, **Physical Review Letters**, v. 13, n. 9, p. 321–323, 1964.
- [21] ALEPH Collaboration et al. Precision Electroweak Measurements and Constraints on the Standard Model, 2010, arXiv:1012.2367 [hep-ex].
- [22] The ATLAS Collaboration, The ATLAS Experiment at the CERN Large Hadron Collider, **Journal of Instrumentation**, v. 3, n. 08, p. S08003, 2008.
- [23] HÜGGING, F., The ATLAS Pixel Insertable B-layer (IBL), **Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment**, v. 650, n. 1, p. 45–49, 2011, arXiv:1012.2742 [physics.ins-det].
- [24] The ATLAS Collaboration, Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in High Pile-Up LHC Environment, ATLAS-CONF-2012-042, url: <http://cdsweb.cern.ch/record/1435196>.
- [25] The ATLAS Collaboration, Charged-particle multiplicities in pp interactions at $\sqrt{s} = 900$ GeV measured with the ATLAS detector at the LHC, **Physics Letters B**, v. 716, n. 1, p. 21, 2010.
- [26] <https://cds.cern.ch/record/2021476>, acesso em 15 de janeiro de 2019.
- [27] DE MELO, J. P. C. **Covariância na frente de luz**. Tese (Doutorado em Física) – Instituto de Física, Universidade de São Paulo. São Paulo, 1998.
- [28] MARTIN, A. D. et al. Parton distributions for the LHC, **The European Physical Journal C**, v. 63, n. 2, p. 189–285, 2009, arXiv:0901.0002 [hep-ph].
- [29] CAMPBELL, J. M., HUSTON, J. W., STIRLING, W. J., Hard Interactions of Quarks and Gluons: a Primer for LHC Physics, **Reports on Progress in Physics**, v. 70, n. 1, 2006, arXiv:hep-ph/0611148.
- [30] BACONNIER, Y. et al., LHC: the Large Hadron Collider accelerator project, CERN-AC-93-03-LHC, url: <https://cds.cern.ch/record/257706>.

- [31] LHC Study Group Collaboration, The Large Hadron Collider: conceptual design, CERN-AC-95-05-LHC, url:<https://cds.cern.ch/record/291782>.
- [32] MOBS, E., The CERN accelerator complex, Disponível em <https://cds.cern.ch/record/2197559>, acesso em 11 janeiro 2019.
- [33] The CMS Collaboration, The CMS experiment at the CERN LHC, **Journal of Instrumentation**, v. 3, n. 08, p. S08004, 2008
- [34] The ALICE Collaboration, The ALICE experiment at the CERN LHC, **Journal of Instrumentation**, v. 3, n. 08, p. S08002, 2008
- [35] The LHCb Collaboration, The LHCb experiment at the CERN LHC, **Journal of Instrumentation**, v. 3, n. 08, p. S08005, 2008
- [36] FUKUI, K. et al., The LHCf detector at the CERN Large Hadron Collider, **Journal of Instrumentation**, v. 3, n. 08, p. S08006, 2008.
- [37] The TOTEM Collaboration, The TOTEM Experiment at the CERN Large Hadron Collider, **Journal of Instrumentation**, v. 3, n. 08, p. S08007, 2008.
- [38] APOLLINARI, G. et al. **High-Luminosity Large Hadron Collider (HL-LHC): Preliminary Design Report**. Geneva: CERN Yellow Reports: Monographs, 2015, Disponível em <https://cds.cern.ch/record/2116337>.
- [39] The ATLAS Collaboration, Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV, ATL-COM-PHYS-2016-407, url: <https://cds.cern.ch/record/2149445>.
- [40] The ATLAS Collaboration, Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 7$ TeV with ATLAS, **The European Physical Journal C**, v. 72, n. 1, p. 1844.
- [41] The ATLAS Collaboration, Performance of missing transverse momentum reconstruction for the ATLAS detector in the first proton-proton collisions at $\sqrt{s} = 13$ TeV, ATL-PHYS-PUB-2015-027, url: <https://cds.cern.ch/record/2037904>.

- [42] The ATLAS Collaboration, Electron and photon reconstruction and identification in ATLAS: expected performance at haigh energy and results at 900 GerV, ATLAS-CONF-2010-005, url: <https://cds.cern.ch/record/1273197>.
- [43] LAMPL, W. et al., Calorimeter clustering algorithms: Description and performance, ATL-LARG-PUB-2008-002, url: <https://cds.cern.ch/record/1099735>.
- [44] The ATLAS Collaboration, Muon Reconstruction Performance, ATLAS-CONF-2010-064, url: <https://cds.cern.ch/record/1281339>.
- [45] CACCIARI, M., SALAM, G. P., SOYEZ, G., The Catchment Area of Jets, **Journal of High Energy Physics**, v. 0804, p. 005, 2008.
- [46] The ATLAS Collaboration, E_T^{miss} from the reconstruction and calibration of the hadronic recoil, ATL-COM-PHYS-2010-090, url: <https://cds.cern.ch/record/1241042>.
- [47] The ATLAS Collaboration, Measurement of m_W at 7 TeV: Hadronic recoil corrections, ATL-COM-PHYS-2015-344, url: <https://cds.cern.ch/record/2013274>.
- [48] The ATLAS Collaboration, Measurement of the transverse momentum distribution of W bosons in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector, **Physical Review D**, v. 85, n. 1, 2012, arXiv:1108.6308 [hep-ex].
- [49] The ATLAS Collaboration, Precision measurement and interpretation of inclusive W^+ , W^- and Z/γ^* production cross sections with the ATLAS detector, **The European Physical Journal C**, v. 77, n. 6, p. 367, 2017, arXiv:1612.03016v2 [hep-ex].
- [50] The ATLAS Collaboration, Measurement of m_W at $\sqrt{s} = 7$ TeV: Hadronic Recoil Algorithms, ATL-COM-PHYS-2014-035, url: <https://cds.cern.ch/record/1966964>.
- [51] The ATLAS Collaboration, Expected performance of missing transverse momentum reconstruction for the ATLAS detector at $\sqrt{s} = 13$ TeV, ATL-COM-PHYS-2015-347, url: <https://cds.cern.ch/record/2013489/>.

- [52] The ATLAS Collaboration, Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data, ATLAS-CONF-2016-024, url: <https://cds.cern.ch/record/2157687>.
- [53] The ATLAS Collaboration, Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung, ATLAS-CONF-2012-047, url: <https://cdsweb.cern.ch/record/1449796>.
- [54] The ATLAS Collaboration, Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data, ATLAS-CONF-2014-032, url: <http://cdsweb.cern.ch/record/1706245>.
- [55] The ATLAS Collaboration, Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data, **The European Physical Journal C**, v. 74, n. 10, p. 3071, 2014, arXiv: 1407.5063 [hep-ex].
- [56] The ATLAS Collaboration, Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data, ATLAS-CONF-2016-024, url: <https://cdsweb.cern.ch/record/2157687>.
- [57] GEANT4 Collaboration, Geant4—a simulation toolkit, **Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment**, v. 506, n. 3, p. 250–303, 2003.
- [58] ATLAS Collaboration, Measurement of $W^\pm$ and Z -boson production cross sections in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector, **Physics Letters B**, v. 759, p. 601–621, 2016, arXiv:1603.09222 [hep-ex].
- [59] The ATLAS Collaboration, ACHARYA, B., The ATLAS simulation infrastructure, **European Physical Journal C**, v.70, n. 70, p. 823–874, 2010.
- [60] CATANI, S. et al., Vector boson production at hadron colliders: a fully exclusive QCD calculation at NNLO, **Physical Review Letters**, v. 103, p. 082001, 2009.

- [61] GAVIN, R. et al., FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order, **Computer Physics Communications**, v. 182, p. 2388, 2011.
- [62] BARDIN, D. et al., SANC integrator in the process: QCD and EW contributions, **JETP Letters**, v. 96, n. 5, p. 285, 2012.
- [63] HOLLIK, W. F. L., Radiative Corrections in the Standard Model and their Role for Precision Tests of the Electroweak Theory, **Fortschritte der Physik**, v. 38, p. 165, 1990.
- [64] DITTMAYER, S., HUBER, M., Radiative corrections to the neutral-current Drell-Yan process in the Standard Model and its minimal supersymmetric extension, **Journal of High Energy Physics**, v. 1001, p. 060, 2010.
- [65] DULAT, S. et al., New parton distribution functions from a global analysis of quantum chromodynamics, **Physical Review D**, v.93, p. 033006, 2016.
- [66] BALL, R. D. et al., Parton distributions for the LHC Run II, **Journal of High Energy Physics**, v. 1504, p. 040, 2015.
- [67] HARLAND-LANG, L. A. et al., Parton distributions in the LHC era: MMHT 2014 PDFs, **European Physical Journal C**, v. 75, p. 204, 2015.
- [68] ALCKHIN, S., BLÜMLEIN, J., MOCH, S., The ABM parton distributions tuned to LHC data, **Physical Review D**, v. 89, p. 054028, 2014.
- [69] ABRAMOWICZ, H. et al., Combination of measurements of deep inelastic $e^\pm p$ scattering cross sections and QCD analysis of HERA data, **European Physical Journal C**, v. 75, p. 580, 2015.
- [70] ATLAS Collaboration, Determination of the strange quark density of the proton from ATLAS measurements of the $W \rightarrow l\nu$ and $Z \rightarrow ll$ cross sections, **Physical Review Letters**, v. 109, p. 012001, 2012.

- [71] ALIOLI, S., NASON, P., OLEARI, C., RE, E., A general framework for implementing NLO calculations in shower Monte carlo programs: the POWHEG BOX, **Journal of High Energy Physics**, v. 2010, p. 043, 2010.
- [72] SJÖSTRAND, T., MRENNNA, S., SKANDS, P., A brief introduction to PYTHIA 8.1, **Computer Physics Communications**, v. 178, n. 11, p. 852–867, 2008.
- [73] LAI, H. L. et al., New parton distributions for collider physics, **Physical Review D**, v. 82, n. 7, 2010.
- [74] The ATLAS Collaboration, Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector, **Journal of High Energy Physics**, v. 2014, n. 9, 2014, arXiv:1406.3660 [hep-ex].
- [75] The ATLAS Collaboration, Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector, **Journal of High Energy Physics**, v. 2014, n. 9, p. 1–47, 2014.
- [76] PUMPLIN, J. et al., New Generation of Parton Distributions with Uncertainties from Global QCD Analysis, **Journal of High Energy Physics**, v. 2002, n. 07, p. 012, 2002.
- [77] Lange, D. J., The EvtGen particle decay simulation package, **Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment**, v. 462, n. 1–2, p. 152, 2001.
- [78] DAVIDSON, N., PRZEDZINSKI, T., WAS, Z., PHOTOS interface in C++: Technical and physics documentation, **Computer Physics Communications**, v. 199, p. 86–101, 2016.
- [79] SJÖSTRAND, T., MRENNNA, S., SKANDS, P., PYTHIA 6.4 physics and manual, **Journal of High Energy Physics**, v. 2006, n. 05, p. 026, 2006.

- [80] CZAKON, M., MITOV, A., Top++ A program for the calculation of the top-pair cross-section at hadron colliders, **Computer Physics Communications**, v. 185, n. 11, p. 2930–2938, 2014.
- [81] GLEISBERG, T. et al., Event generation with SHERPA 1.1, **Journal of High Energy Physics**, v. 2009, n. 2, p. 007, 2009.
- [82] The ATLAS Collaboration, Summary of ATLAS Pythia 8 tunes, ATL-PHYS-PUB-2012-003, 2012, url: <https://cds.cern.ch/record/1474107>
- [83] MARTIN, A. D., STIRLING, W. J., THORNE, R. S., WATT, G., Parton distributions for the LHC, **The European Physical Journal C** v. 63, n. 2, p. 189–285, 2009.
- [84] The ATLAS Collaboration, Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13$ TeV, **The European Physical Journal C**, v. 76, n. 5, p. 1–30, 2016.
- [85] The ATLAS Collaboration, Electron identification measurements in ATLAS using $\sqrt{s} = 13$ TeV data with 50 ns bunch spacing, ATL-PHYS-PUB-2015-041, url: <https://cds.cern.ch/record/2048202>
- [86] CACCIARI, M., GAVIN, S. P., SOYEZ, G., The anti- k_t jet clustering algorithm, **Journal of High Energy Physics**, v. 20118, n. 04, p. 063, 2008.
- [87] The ATLAS Collaboration, Tagging and suppression of pileup jets with the ATLAS detector, ATLAS-CONF-2014-018, url: <https://cds.cern.ch/record/1700870>.
- [88] The ATLAS Collaboration, Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data, **The European Physical Journal C**, v. 74, n. 10, p. 1–48, 2014.
- [89] The ATLAS Collaboration, Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV, ATL-PHYS-PUB-2015-015, url: <https://cds.cern.ch/record/2037613>.

- [90] The ATLAS Collaboration, Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 13$ TeV, **The European Physical Journal C**, v. 73, n. 3, p. 1–118, 2013.
- [91] The ATLAS Collaboration, Data-driven determination of the energy scale and resolution of jets in the ATLAS calorimeters using dijet and multijet events at $\sqrt{s} = 8$ TeV, ATLAS-CONF-2015-017, url: <https://cds.cern.ch/record/2008678/>.
- [92] GLAZOV, A., Averaging of DIS cross section data, **AIP Conference Proceedings**, v. 792, p. 237, 2005.
- [93] ROJO, J. et al., The PDF4LHC report on PDFs and LHC data: Results from Run I and preparation for Run II, 2015, arXiv:1507.00556 [hep-ph].
- [94] <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2>, acesso em 20 de junho de 2018.
- [95] AABOUD, M. et al., Measurements of top-quark pair differential cross-sections in the emu channel in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector, **European Physical Journal C**, v. 77, 2017, <https://doi.org/10.1140/epjc/s10052-017-4821-x>.
- [96] AABOUD, M. et al. Measurements of electroweak Wjj production and constraints on anomalous gauge couplings with the ATLAS detector, **European Physical Journal C**, v. 77, 2017, <https://doi.org/10.1140/epjc/s10052-017-5007-2>.
- [97] The ATLAS collaboration, Aad, G., Abbott, B. et al.: Measurement of the forward-backward asymmetry of electron and muon pair-production in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector.; J. High Energ. Phys. (2015) 2015: 49. [https://doi.org/10.1007/JHEP09\(2015\)049](https://doi.org/10.1007/JHEP09(2015)049)
- [98] Search for a new heavy gauge boson resonance decaying into a lepton and missing transverse momentum in 79.8 fb^{-1} of pp collisions at

$\sqrt{s} = 13$ TeV with the ATLAS experiment, ATLAS-CONF-2018-017, url: <http://cdsweb.cern.ch/record/2621303>.

- [99] The ATLAS Collaboration, Studies of b -tagging performance and jet substructure in a high p_T $g \rightarrow b^+b^-$ rich sample of large- R jets from pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector, ATLAS-CONF-2016-002, url: <https://cds.cern.ch/record/2135187>.
- [100] BEHNKE, O. et al., **Data Analysis in High Energy Physics: A Practical Guide to Statistical Methods**, 1 ed, Wiley-VCH, 2013.
- [101] HOECKER, A., et al., Toolkit for Multivariate Data Analysis, arXiv:physics/0703039 [physics.data-an].
- [102] <https://xgboost.readthedocs.io/>, acesso em 5 de janeiro de 2019.
- [103] <https://www.tensorflow.org>, acesso em 5 de janeiro de 2019.
- [104] FREUND, Y., SCHAPIRE, R. E., A Decision-Theoretic Generalization of on-Line Learning and an Application to Boosting, **Journal of Computer and System Sciences**, v. 55, n. 1, p. 119–139, 1997.
- [105] LAI, H. L. et al., New parton distributions for collider physics, **Physical Review D**, v. 82, n. 7, 2010.
- [106] CZAKON, M., MITOV, A., Top++ A program for the calculation of the top-pair cross-section at hadron colliders, **Computer Physics Communications**, v. 185, n. 11, p. 2930–2938, 2014.
- [107] ROJO, J. et al, The PDF4LHC report on PDFs and LHC data: results from Run I and preparation for Run II, **Journal of Physics G: Nuclear and Particle Physics**, v. 42, n. 10, 2015, arXiv:1507.00556 [hep-ph].
- [108] The ATLAS Collaboration, Luminosity determination in pp collisions at $\sqrt{s}=8$ TeV using the ATLAS detector at the LHC, **The European Physical Journal C**, v. 76, n. 12, 2016, arXiv:1608.03953 [hep-ex].