



MODELOS NEURAIIS PARA A CORREÇÃO DE ESTIMATIVAS DE  
EVENTOS RAROS EM UM EXPERIMENTO DE ALTAS ENERGIAS

Philipp do Nascimento Gaspar

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia Elétrica, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia Elétrica.

Orientadores: José Manoel de Seixas  
Marcia Begalli

Rio de Janeiro  
Março de 2016

MODELOS NEURAIIS PARA A CORREÇÃO DE ESTIMATIVAS DE  
EVENTOS RAROS EM UM EXPERIMENTO DE ALTAS ENERGIAS

Philipp do Nascimento Gaspar

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO  
ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE  
ENGENHARIA (COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE  
JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A  
OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA  
ELÉTRICA.

Examinada por:

---

Prof. José Manoel de Seixas, D.Sc.

---

Prof. Marcia Begalli, Ph.D.

---

Prof. João Torres de Melo Neto, Ph.D.

---

Prof. Charles Casimiro Cavalcante, D.Sc.

---

Prof. Gerson Zaverucha, Ph.D.

RIO DE JANEIRO, RJ – BRASIL

MARÇO DE 2016

Gaspar, Philipp do Nascimento

Modelos Neurais para a Correção de Estimativas de Eventos Raros em um Experimento de Altas Energias/Philipp do Nascimento Gaspar. – Rio de Janeiro: UFRJ/COPPE, 2016.

XV, 89 p.: il.; 29, 7cm.

Orientadores: José Manoel de Seixas

Marcia Begalli

Dissertação (mestrado) – UFRJ/COPPE/Programa de Engenharia Elétrica, 2016.

Referências Bibliográficas: p. 83 – 89.

1. Inteligência Computacional. 2. Física de Altas Energias. 3. Redes Neurais Artificiais. I. Seixas, José Manoel de *et al.* II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Elétrica. III. Título.

*Para minha mãe, que me chama  
de "filho", e para meu pai, que  
me chama de "cara".*

# Agradecimentos

Agradeço aos meus pais, José Carlos que mesmo sem saber me ensinou a difícil tarefa de ser resiliente sem perder o bom humor; e também à Mirian, não me imagino vivendo sem esse afeto.

Aos meus orientadores José Seixas e Marcia Begalli por terem acreditado em mim e principalmente pela paciência durante a minha orientação.

À minha namorada, Gabriela Vaz, pelo incentivo e compreensão por conta dos dias em que deixei de vê-la para confeccionar este trabalho.

Aos meus companheiros de LPS, em especial ao Natanael Júnior pela ajuda que me deu durante meu primeiro ano como mestrando.

Aos amigos que a vida me deu; carrego todos vocês em mim.

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

MODELOS NEURAIIS PARA A CORREÇÃO DE ESTIMATIVAS DE  
EVENTOS RAROS EM UM EXPERIMENTO DE ALTAS ENERGIAS

Philipp do Nascimento Gaspar

Março/2016

Orientadores: José Manoel de Seixas

Marcia Begalli

Programa: Engenharia Elétrica

A medição feita pelos sensores de um sistema de instrumentação eletrônica possui um papel fundamental. Neste contexto o calorímetro do detector ATLAS utiliza milhares de canais para medir a energia de partículas sub-atômicas, produzidas no acelerador LHC, e em seguida as converte em sinais digitais de forma a armazenar estes dados. Contudo, as não-linearidades intrínsecas à resposta do detector faz com que haja uma diferença entre o valor medido e o valor real de energia destas partículas. Neste sentido, esta dissertação apresenta uma correção baseada em Redes Neurais Artificiais de forma a reduzir esta diferença e analisa seu desempenho utilizando métodos estatísticos.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

NEURAL MODELS FOR CORRECTION ESTIMATES OF RARE EVENTS IN  
A HIGH ENERGY EXPERIMENT

Philipp do Nascimento Gaspar

March/2016

Advisors: José Manoel de Seixas

Marcia Begalli

Department: Electrical Engineering

The measurement made by a system of electronic sensors has a fundamental role. In this context the ATLAS detector calorimeter uses thousands of channels to measure the energy of the sub-atomic particles produced in LHC, and then converts them into digital signals so as to store these data. However, the nonlinearities intrinsic to the detector response make a difference between the measured value and the actual value of energy of these particles. In this sense, this work presents an artificial neural network-based correction to reduce this difference and analyzes its performance using statistical methods.

# Sumário

<b>Lista de Figuras</b>	<b>x</b>
<b>Lista de Tabelas</b>	<b>xiv</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Objetivo . . . . .	2
1.2 Organização da Dissertação . . . . .	2
<b>2 A Física de Altas Energias</b>	<b>3</b>
2.1 Física de Altas Energias . . . . .	4
2.1.1 Unidades de Energia . . . . .	6
2.2 Modelo Padrão . . . . .	6
2.3 Bóson de Higgs . . . . .	8
<b>3 O Experimento ATLAS no LHC</b>	<b>11</b>
3.1 O Grande Colisor de Hádrons (LHC) . . . . .	11
3.2 O Detector ATLAS . . . . .	12
3.2.1 Sistema de Coordenadas do ATLAS . . . . .	15
3.3 Detector Interno . . . . .	16
3.4 Sistema de Calorimetria . . . . .	17
3.4.1 Calorímetro Eletromagnético . . . . .	19
3.4.2 Calorímetro Hadrônico . . . . .	20
3.5 Espectrômetro de Múons . . . . .	21
3.6 Sistema de Aquisição de Dados e Trigger . . . . .	22
<b>4 Reconstrução e Identificação de Jatos</b>	<b>24</b>
4.1 Jatos no Experimento D0 . . . . .	25
4.2 Jatos no Experimento CDF . . . . .	26
4.3 Jatos no Experimento CMS . . . . .	26
4.4 Reconstrução de Jatos no Detector ATLAS . . . . .	27
4.4.1 Cluster no Calorímetro . . . . .	27
4.4.2 Reconstrução - anti-Kt . . . . .	28



4.4.3	Correção JES - <i>Jet Energy Scale</i> . . . . .	29
4.4.4	Correção GSC - <i>Global Sequential Calibration</i> . . . . .	31
4.4.5	<i>Tagging b</i> -jatos - Identificando jatos <i>b</i> . . . . .	31
4.4.6	<i>Muon-in-jet correction</i> para jatos <i>b</i> . . . . .	33
4.4.7	Correção do Momentum Transverso dos jatos <i>b</i> . . . . .	34
<b>5</b>	<b>Método de Análise</b>	<b>36</b>
5.1	Seleção de Eventos . . . . .	36
5.2	Redes Neurais Artificiais . . . . .	36
5.3	Validação Cruzada . . . . .	41
5.4	Correções Por Jato e Por Evento . . . . .	42
5.5	Aplicação da Correção Neural . . . . .	43
5.6	Cálculo da Massa do Bóson de Higgs . . . . .	43
5.7	Divergência de Jensen-Shannon . . . . .	44
5.8	Modelo de Ajuste . . . . .	46
5.9	Conjunto de Redes Neurais ( <i>Boosting</i> ) . . . . .	47
<b>6</b>	<b>Resultados Obtidos</b>	<b>49</b>
6.1	Correções Por Jato . . . . .	49
6.1.1	Correção Utilizando o Momento Transverso dos Jatos . . . . .	49
6.1.2	Correção Utilizando um Descritor dos Jatos . . . . .	57
6.2	Correções Por Evento . . . . .	65
6.2.1	Correção Utilizando o Momento Transverso dos Jatos . . . . .	65
6.2.2	Correção Utilizando um Descritor dos Jatos . . . . .	68
6.3	Boosting das Correções Neurais . . . . .	73
6.4	Comparação dos Modelos Neurais Propostos . . . . .	75
<b>7</b>	<b>Conclusões e Trabalhos Futuros</b>	<b>81</b>
7.1	Trabalhos Futuros . . . . .	82
	<b>Referências Bibliográficas</b>	<b>83</b>

# Lista de Figuras

2.1	Branching Ratio do bóson de Higgs para diferentes canais de decaimento em função de sua massa. Fonte: [7] . . . . .	9
3.1	Visão esquemática do LHC com seus 4 experimentos principais: ALICE, ATLAS, CMS e LHCb e os aceleradores antigos que agora são usados como pré-aceleradores para o LHC. Sistemas secundários, utilizados em outros experimentos, também pode ser vistos na figura. Fonte: [20] . . . . .	13
3.2	Visão esquemática do detector ATLAS. Fonte: [17] . . . . .	14
3.3	Ilustração de diferentes partículas interagindo com os sub-detectores do ATLAS. Fonte: [18] . . . . .	14
3.4	Sistema de coordenadas do ATLAS. Fonte: [17] . . . . .	16
3.5	Visão geral do Inner Detector com os sub-detectores Pixel, SCT e TRT para as regiões do barril e end-cap. Fonte: [21] . . . . .	17
3.6	Layout do sistema de calorimetria do ATLAS. Fonte [17] . . . . .	18
3.7	Ilustração das diferentes camadas de um módulo do barril do calorímetro eletromagnético do ATLAS. Fonte: [17] . . . . .	19
3.8	Geometria do sistema magnético do ATLAS (esquerda) e uma simulação do mapa do campo magnético (direita). Fonte: [32] . . . . .	21
3.9	Visão geral do espectrômetro de múons do detector ATLAS. Fonte: [17] . . . . .	22
4.1	Etapas da análise para os dados reais e simulações de Monte Carlo junto com os <i>softwares</i> usados em cada etapa. . . . .	25
4.2	Ilustração de um topocluster com a razão sinal-ruído para cada célula. . . . .	28
4.3	Média de correção na escala de energia dos jatos obtida através de simulações de Monte Carlo. A correção só é mostrada na faixa acessível da cinemática. Fonte: [55] . . . . .	30
4.4	Resposta dos jatos, $R(E, \eta)$ , na escala de energia eletromagnética em função da pseudo-rapidez do detector $\eta_{det}$ . O inverso da Resposta é igual à média do fator de correção para cada bin. Fonte: [55] . . . . .	30

4.5	Diagrama de um jato $b$ . Fonte: [60] . . . . .	32
4.6	Eficiência do algoritmo MV1 para selecionar jatos $b$ , $c$ e jatos leves em função de $p_T$ (4.6a) e $\eta$ (4.6b). O discriminante é escolhido para ter 70% de eficiência para jatos $b$ com o $p_T > 20$ e $ \eta  < 2.5$ e avaliado numa amostra de eventos de $t\bar{t}$ . Fonte: [59] . . . . .	32
4.7	Diagramas de Feynman para o decaimento de quarks $b$ , onde um dos bósons $W$ decai semi-leptonicamente em um múon e um neutrino, enquanto o outro bóson $W$ decai hadronicamente. Este processo ocorre em aproximadamente 10.6% dos casos. . . . .	33
4.8	Fatores de correção utilizados para cada bin de $p_T$ . Fonte: [62] . . . . .	34
4.9	Massa invariante reconstruída em diferentes etapas de correção para a análise $ZH \rightarrow llb\bar{b}$ . Fonte: [38] . . . . .	35
5.1	Neurônio artificial. Fonte [63] . . . . .	37
5.2	Exemplo de rede neural com múltiplas camadas. . . . .	39
5.3	Ilustração do processo de validação cruzada por k-folds. . . . .	41
5.4	Um exemplo do formato da função densidade de probabilidade de Bukin. . . . .	46
5.5	O $\chi^2$ normalizado para diferentes graus de liberdade. Cada curva corresponde a um grau de confiabilidade do modelo ajustado. Fonte: [79] . . . . .	47
6.1	MSE médio para cada topologia das redes com todos os eventos (a), com múons (b) e sem múons (c) das correções por jato utilizando o momento transversal. . . . .	51
6.2	Distribuições de $m_{b\bar{b}}$ , simulada e corrigida pelo corretor neural <i>por jato</i> utilizando o momento transversal. . . . .	52
6.3	Ajuste da função Bukin para a distribuição de $m_{b\bar{b}}$ para eventos corrigidos com a rede neural para todos os jatos (@NN) e corrigidos com uma rede para os jatos com múon e outra para os jatos sem múon (@NN Muons/noMuons). Correções <i>por jato</i> utilizando o momento transversal. . . . .	54
6.4	Ajuste da função Bukin para a distribuição de $m_{b\bar{b}}$ para eventos gerados (@TRUTH) e reconstruídos com a correção oficial da colaboração (@GSC). . . . .	55
6.5	Distribuições de $m_{b\bar{b}}$ , simulada e corrigida pelo corretor neural <i>por jato</i> utilizando o momento transversal e aplicadas na amostra de background $WZ$ . . . . .	59

6.6	MSE médio para cada topologia das redes com todos os eventos (a), com múons (b) e sem múons (c) das correções por jato utilizando um descritor de jatos. . . . .	60
6.7	Distribuições de $m_{b\bar{b}}$ , simulada e corrigida pelo corretor neural <i>por jato</i> utilizando um descritor de jatos. . . . .	61
6.8	Ajuste da função Bukin para a distribuição de $m_{b\bar{b}}$ para eventos corrigidos com a rede neural para todos os jatos (@NN) e corrigidos com uma rede para os jatos com múon e outra para os jatos sem múon (@NN Muons/noMuons). Correções <i>por jato</i> utilizando um descritor de jatos. . . . .	62
6.9	Distribuições de $m_{b\bar{b}}$ , simulada e corrigida pelo corretor neural <i>por jato</i> utilizando um descritor de jatos e aplicadas na amostra de background $WZ$ . . . . .	63
6.10	MSE médio para cada topologia da rede neural por evento utilizando o momento transversal. . . . .	65
6.11	Distribuições de $m_{b\bar{b}}$ , simulada e corrigida pelo corretor neural <i>por evento</i> utilizando o momento transversal. . . . .	66
6.12	Ajuste da função Bukin para a distribuição de $m_{b\bar{b}}$ para eventos corrigidos com a rede neural (@NN). Correção <i>por evento</i> utilizando o momento transversal. . . . .	67
6.13	Distribuições de $m_{b\bar{b}}$ , simulada e corrigida pelo corretor neural <i>por evento</i> utilizando o momento transversal e aplicadas na amostra de background $WZ$ . . . . .	68
6.14	MSE médio para cada topologia da rede neural por evento utilizando um descritor de jatos. . . . .	69
6.15	Distribuições de $m_{b\bar{b}}$ , simulada e corrigida pelo corretor neural <i>por evento</i> utilizando um descritor de jatos. . . . .	70
6.16	Ajuste da função Bukin para a distribuição de $m_{b\bar{b}}$ para eventos corrigidos com a rede neural (@NN). Correção <i>por evento</i> utilizando um descritor de jatos. . . . .	70
6.17	Distribuições de $m_{b\bar{b}}$ , simulada e corrigida pelo corretor neural <i>por evento</i> utilizando um descritor de jatos e aplicadas na amostra de background $WZ$ . . . . .	72
6.18	Efeito na massa invariante do Higgs para cada iteração do <i>boosting</i> na correção por evento. . . . .	73
6.19	Efeito na massa invariante dos jatos $b$ na amostra de <i>background</i> ( $WZ$ ) para cada iteração do <i>boosting</i> na correção por evento. . . . .	74
6.20	Efeito na massa invariante do Higgs para cada iteração do <i>boosting</i> na correção por jato. . . . .	74

6.21	Efeito na massa invariante dos jatos $b$ na amostra de <i>background</i> ( $WZ$ ) para cada iteração do <i>boosting</i> na correção por jato. . . . .	75
6.22	Diagrama com as correções sequenciais testadas pela colaboração ATLAS. . . . .	75
6.23	Distribuições de massa reconstruída com cada uma das correções para o sinal do decaimento $H \rightarrow b\bar{b}$ . . . . .	76
6.24	Ajuste da função Bukin para a distribuição de massa invariante do bóson de Higgs para os estágios de correção @Mu e @Pt. . . . .	78

# Lista de Tabelas

2.1	Lista dos férmions do Modelo Padrão. . . . .	7
2.2	Bósons de Gauge descritos pelo Modelo Padrão. . . . .	8
2.3	Razão de Ramificação ( <i>Branching Ratios</i> ) para todos os canais de decaimento do bóson de Higgs, para $m_H = 125.4$ GeV segundo as previsões do Modelo Padrão. Fonte: [8] . . . . .	10
5.1	Parâmetros de treinamento do algoritmo LM. . . . .	40
6.1	Estimativas para o histograma de massa invariante simulada e corrigida pelo corretor neural <i>por jato</i> utilizando o momento transverso. . . . .	53
6.2	Ajuste da função Bukin para diferentes intervalos de $m_{b\bar{b}}$ . Correções <i>por jato</i> utilizando o momento transverso. . . . .	56
6.3	Comparação da divergência JS para a distribuição de massa corrigida com a correção neural <i>por jato</i> utilizando o momento transverso. . . . .	58
6.4	Estimativas para o histograma de massa invariante simulada e corrigida pelo corretor neural <i>por jato</i> utilizando um descritor de jatos. . . . .	61
6.5	Ajuste da função Bukin para diferentes intervalos de $m_{b\bar{b}}$ . Correções <i>por jato</i> utilizando um descritor de jatos. . . . .	63
6.6	Comparação da divergência JS para a distribuição de massa corrigida com a correção neural <i>por jato</i> utilizando um descritor de jatos. . . . .	64
6.7	Estimativas para o histograma de massa invariante simulada e corrigida pelo corretor neural <i>por evento</i> utilizando o momento transverso. . . . .	66
6.8	Ajuste da função Bukin para diferentes intervalos de $m_{b\bar{b}}$ . Correções <i>por evento</i> utilizando o momento transverso. . . . .	67
6.9	Comparação da divergência JS para a distribuição de massa corrigida com a correção neural <i>por evento</i> utilizando o momento transverso. . . . .	68
6.10	Estimativas para o histograma de massa invariante simulada e corrigida pelo corretor neural <i>por evento</i> utilizando um descritor de jatos. . . . .	70
6.11	Ajuste da função Bukin para diferentes intervalos de $m_{b\bar{b}}$ . Correções <i>por evento</i> utilizando um descritor de jatos. . . . .	71

6.12	Comparação da divergência JS para a distribuição de massa corrigida com a correção neural <i>por evento</i> utilizando um descritor de jatos. . .	71
6.13	Estimativas para o histograma de massa invariante gerada, reconstruída e corrigida com a correção neural. . . . .	77
6.14	Ajuste da função Bukin para as correções @Mu, @Pt, @NN (Muons/noMuons) e seu ajuste para a distribuição verdadeira @TRUTH.	79
6.15	Comparação da divergência JS para a distribuição de massa corrigida com a correção neural vencedora entre todos os modelos testados. . .	80

# Capítulo 1

## Introdução

Em um sistema de instrumentação eletrônica, a medição feita pelos seus sensores possui um papel primordial na estimação do sinal que se propõe obter. Este é um dos desafios do experimento ATLAS (capítulo 3), situado no CERN, que detecta o resultado de colisões de partículas, geradas pelo acelerador LHC, com seu sub-sistema de detectores. Uma das mais importantes características de uma partícula é a sua energia. Um dos sub-sistemas, mais precisamente o calorímetro (seção 3.4), do ATLAS utiliza centenas de milhares de canais para ler a energia depositada por essas partículas e a converte em sinais digitais para serem armazenados e analisados posteriormente pela colaboração.

As partículas detectadas pelo calorímetro podem ser de origem hadrônica ou eletromagnética (seção 2.2). Um calorímetro é dito compensante se a resposta hadrônica for idêntica à resposta eletromagnética, logo um calorímetro não-compensado terá que ser calibrado com uma função que depende da energia dessas partículas, antes de ser usado numa análise. Além da natureza não-compensada do calorímetro do ATLAS, qualquer grandeza física que foi medida pelo detector será afetada por suas imperfeições, sejam elas as não-linearidades presentes na eletrônica do detector ou perda de energia em regiões mortas. Inclusive a alta taxa de eventos do experimento pode ocasionar uma medida imprecisa por conta de um processo conhecido como *pileup* (seção 3.1).

Contudo, para minimizar essa diferença do valor medido e o valor real são necessários algoritmos de calibração e correção de energia. Neste contexto a colaboração ATLAS aplica correções de energia de forma sequencial, em diferentes estágios, utilizando tanto as informações do detector como as informações intrínsecas de cada partícula de modo a se aproximar estatisticamente do valor real de energia.



## 1.1 Objetivo

O objetivo desta dissertação consiste em derivar um novo método de correção para os jatos de partículas (capítulo 4) detectados no calorímetro, mais especificamente os jatos provenientes de um decaimento do bóson de Higgs (seção 2.3). Esta correção é feita utilizando redes neurais artificiais, uma vez que estas são consideradas mapeadores universais. Todo o trabalho foi desenvolvido em cima das simulações de Monte Carlo e a correção proposta foi comparada com as correções oficiais dadas pela colaboração do experimento. O resultado da correção é medido na massa invariante dos dois jatos que foram selecionados como sendo produtos de decaimento do Bóson de Higgs, pois a massa invariante,  $m_{b\bar{b}}$ , é o melhor discriminante conhecido para separar sinal e background.

## 1.2 Organização da Dissertação

- O Capítulo 2 faz uma introdução à Física de Altas Energias, ao Modelo Padrão e ao Bóson de Higgs, apresentando um resumo da área de estudo deste trabalho.
- O Capítulo 3 descreve o acelerador LHC e o detector ATLAS com cada característica de seus sub-sistemas, mostrando o aparato experimental onde este trabalho foi desenvolvido.
- O Capítulo 4 apresenta as técnicas de correções na energia de partículas utilizadas nos mais recentes experimentos de física de altas energias. Uma ênfase maior será dada para as correções de energia do detector ATLAS.
- O Capítulo 5 mostra o método de análise em conjunto com as técnicas empregadas.
- O Capítulo 6 contém os resultados das correções neurais propostas e sua análise estatística é discutida correção por correção.
- O Capítulo 7 termina a dissertação com uma conclusão final sobre o trabalho e também apresenta os possíveis trabalhos futuros.

## Capítulo 2

# A Física de Altas Energias

Na Grécia antiga, uma corrente de filósofos acreditavam que o Universo inteiro reduzia-se a um, dois ou alguns poucos componentes da matéria. Tales de Mileto (624 - 546 a.C) acreditava que toda a imensa diversidade da natureza podia ser representada por um único elemento: a água - "sobre qual a Terra flutua e é o começo de todas as coisas". De certa forma, ele foi o filósofo precursor desse ímpeto de explicar o mundo à sua volta. A característica marcante do trabalho de Tales de Mileto reside na procura do entendimento da natureza de modo racional, postulando que este esteja ligado a um único princípio. Suas ideias são justificadas não em termos de deuses ou forças sobrenaturais, mas em termos da lógica. Vieram outros, depois, para corrigir, modificar e ampliar essa descrição, incluindo novos elementos: para Anaxímenes de Mileto (570 - 500 a.C) o elemento primordial seria o ar; para Xenófanes da Jônia (570 - 460 a.C) a terra seria o elemento mais primitivo do Universo; para Heráclito de Éfeso (540 - 480 a.C) o fogo seria o constituinte mais elementar. Passou-se a acreditar, então, que os elementos fundamentais da natureza eram o fogo, a terra, a água e o ar [1].

Uma descrição mais elaborada deu-se por volta do quinto século antes de Cristo, quando filósofos gregos propuseram a chamada teoria atomista. A palavra *átomo* é derivada do grego *atomos*, que significa: indivisível. Para os atomistas o mundo era composto de apenas duas partes fundamentais: átomos e vazio; sendo os átomos partículas muito pequenas, indestrutíveis e imutáveis. O primeiro filósofo da antiguidade a teorizar que o Universo foi produzido pelo movimento dos átomos foi Leucipo de Mileto (480 - 420 a.C) e para ele, esse movimento dos átomos era aleatório e indeterminado. Para Leucipo, portanto, o cosmo (o mundo e todas as coisas, inclusive a alma) é formado por um turbilhão de infinitos átomos de diversos formatos que jorram ao acaso e se chocam.

O fenômeno da digestão alimentar e mesmo o cheiro dos alimentos, entre outras observações levaram Demócrito de Abdera (430 - 370 a.C), aluno de Leucipo, a desenvolver a teoria atomista criada por seu mestre. Dizia ele que a consistência

dos aglomerados de átomos que faz com que algo pareça sólido, líquido, gasoso ou anímico (alma) seria então determinada pelo formato e arranjo dos átomos envolvidos. Desse modo, os átomos de aço possuem um formato que se assemelha a ganchos, que os prendem solidamente entre si; os átomos de água são lisos e escorregadios; os átomos de sal, como demonstra seu gosto, são ásperos e pontudos; os átomos de ar são pequenos e pouco ligados, penetrando todos os outros materiais; e os átomos da alma e do fogo são esféricos e muito delicados. Evidentemente, não era possível a Demócrito valer-se de qualquer experiência que desse suporte a sua crença. Alguns de seus contemporâneos assimilaram a noção que propunha, mas não a grande maioria.

No século XVIII, o conceito de átomo foi retomado. John Dalton (1766 - 1844), um professor de inglês que se interessava pela química como passatempo, é considerado o pai da teoria atômica moderna. Para ele, a existência dos átomos era a única ideia que poderia explicar certas propriedades dos gases que observava, o que o levou a formular e postular as leis das proporções definidas, primeira tentativa de relacionar o átomo com a química [2]. A ideia principal era de que toda a matéria seria composta por átomos, todos os átomos de um dado elemento químico seriam idênticos, e átomos de diferentes elementos químicos difeririam quanto a sua massa e outras propriedades. Reviveu, assim, a antiga teoria atomista e elaborou a primeira tabela de pesos atômicos, anunciando seus resultados em 1803. Ao fim de sua vida, sua teoria atômica estava amplamente difundida entre a comunidade química e reconhecida pelo rei da Inglaterra com a Medalha Real.

Agora, no século XXI, dentro da precisão experimental, conhecemos partículas realmente elementares, a partir das quais toda a matéria observada no Universo é formada. E segundo esse entendimento, há forças fundamentais ligando essas partículas, pois a matéria não pode existir sem a ação dessas forças. A física de partículas elementares aborda a seguinte questão: "Do que a matéria é feita?" em seu nível mais fundamental, ou seja, na sua menor escala de tamanho. Podemos considerar que a matéria em nível subatômico é constituída de pequenos blocos e esses, por sua vez, aparecem em diferentes tipos (elétrons, prótons, nêutrons, neutrinos, mésons, entre outros). Apesar das partículas elementares serem objetos complexos, dadas a quantidade e variedade de suas interações, o critério que define elementar não é difícil: toda partícula que não possui estrutura interna é considerada elementar [3].

## 2.1 Física de Altas Energias

A matéria tangível do nosso cotidiano se manifesta em baixas energias, no entanto, para estudar a estrutura íntima da matéria, precisamos criar artificialmente

condições que viabilizem sua manifestação em altíssimas energias. Na década de 1950, a observação dessas partículas tornou-se por si só um ramo da física. Para realizar tal tarefa a comunidade científica passou a projetar máquinas de grande porte em que as partículas são aceleradas em altas energias, num processo parecido aos raios cósmicos, mas em condições de maior controle.

Tais experiências são realizadas nos aceleradores de partículas, a ferramenta fundamental para a pesquisa em física de altas energias. A sua importância pode ser compreendida quando consideramos sua principal característica: a energia do feixe. Não podemos utilizar luz visível para explorar estruturas atômicas e subatômicas. Entretanto, como todas as partículas também possuem propriedades ondulatórias, podemos usá-las como sondas. Então para investigar a estrutura subatômica da matéria precisamos que essas partículas possuam o menor comprimento de onda possível, ou seja, se pretendemos utilizar o acelerador como um enorme "microscópio", sua resolução deve aumentar com a energia do feixe. De acordo com a equação de *de Broglie*, a relação entre momento  $|p|$  e o comprimento de onda  $\lambda$  de um pacote de onda é dado por:

$$\lambda = \frac{h}{|p|} \quad (2.1)$$

onde  $h$  é a constante de Planck [4]. Portanto, quanto maior o momento (a energia do feixe), menor será o comprimento de onda, permitindo assim o acesso a estruturas de dimensões cada vez menores. Na procura de novas subestruturas a equação (2.1) representa uma relação fundamental, ela nos diz que a resolução aumenta quando vamos para altas energias. Por exemplo, as resoluções de  $1 \text{ GeV}/c$  e  $10^3 \text{ GeV}/c$  são:

$$|p|= 1 \frac{\text{GeV}}{c} \rightarrow 1.24 \times 10^{-15} \text{ m} \simeq \text{tamanho de um próton}$$

$$|p|= 10^3 \frac{\text{GeV}}{c} \rightarrow 1.24 \times 10^{-18} \text{ m} \simeq \text{tamanho de um quark}$$

Adicionalmente é possível usar os aceleradores para produzir novas partículas com massa maior do que as partículas que colidiram, pois durante a colisão a energia cinética das partículas originais é convertida na formação de novas partículas. Energia e massa são dois lados da mesma moeda. Massa pode ser convertida em energia e vice-versa de acordo com a famosa equação de *Einstein*

$$E = mc^2 \quad (2.2)$$

Também, por conta desta equivalência, massa e energia podem ser medidas na mesma unidade. Na escala da física de partículas essa unidade é o elétron-volt e seus múltiplos.

Logo, a Física de Altas Energias tem como objetivo a natureza da natureza, o desvendar das regras básicas que permeiam as relações entre a matéria, as leis que regem os processos naturais e os estudos dos elementos fundamentais que compõe a estrutura da matéria.

### 2.1.1 Unidades de Energia

A grandeza física *energia* possui muitas unidades que podem ser utilizadas: Joules, calorias e kilowatt hora são exemplos de medidas de energia em diferentes contextos. Somente a unidade Joule está no Sistema Internacional de Medidas, mas todas elas se relacionam através de fatores de conversão. Em física de partículas, a unidade que é utilizada com mais frequência é o eletrôn-Volt (eV) e os seus derivados keV ( $10^3$  eV), MeV ( $10^6$  eV), GeV ( $10^9$  eV) e o TeV ( $10^{12}$  eV). O eletrôn-Volt é conveniente porque as energias que os físicos de partículas trabalham são muito pequenas quando comparadas com as unidades de energia para mecânica por exemplo. No LHC a energia total de colisão é de 14 TeV, e isso faz dele o acelerador de partículas mais poderoso do mundo. Ainda assim, se convertermos essa energia para Joules teremos:

$$14 \times 10^{12} \times 1.602 \times 10^{-19} = 22.4 \times 10^{-7} \text{ Joules.}$$

A definição de eletrôn-Volt vem do fato de que um elétron acelerado por uma diferença de potencial de 1 Volt terá uma quantidade discreta de energia,  $E = qV$  Joules, onde  $q$  é a carga do elétron em Coulombs e  $V$  é a diferença de potencial em Volts. Logo  $1 \text{ eV} = (1.602 \times 10^{-19} \text{ C}) \times (1 \text{ V}) = 1.602 \times 10^{-19} \text{ J}$ .

## 2.2 Modelo Padrão

O Modelo Padrão da física de partículas elementares é uma teoria relativa às forças eletromagnética, fraca e forte, que são responsáveis por mediar a dinâmica das partículas subatômicas. Desenvolvido na segunda metade do século XX e com sua formulação atual tendo confirmações experimentais até hoje, o Modelo Padrão é tão importante para os físicos teóricos quanto para os experimentais. Para os teóricos, o Modelo Padrão é o exemplo de uma teoria quântica de campos, que apresenta uma ampla variedade de assuntos como: quebra espontânea de simetria, anomalias, comportamento não perturbativo e etc. Também serve como base para a construção de modelos mais exóticos que incorporam partículas hipotéticas, dimensões extras, e simetrias mais elaboradas (como a supersimetria) em uma tentativa de explicar os resultados experimentais em desacordo com o Modelo Padrão, tais como a existência da matéria escura e as oscilações de neutrinos. Por outro lado, os físicos experimentais têm incorporado o Modelo Padrão em simulações de Monte Carlo para ajudar

Tabela 2.1: Lista dos férmions do Modelo Padrão.

Quarks	Carga Elétrica [ $e$ ]	Léptons	Carga Elétrica [ $e$ ]
$u \ c \ t$	$2/3$	$\nu_e \ \nu_\mu \ \nu_\tau$	$0$
$d \ s \ b$	$-1/3$	$e \ \mu \ \tau$	$1$

na busca de uma nova física além do Modelo Padrão, bem como para compreender e ampliar o mesmo.

O Modelo Padrão descreve toda a matéria como sendo composta por partículas elementares chamadas férmions, partículas com spin semi-inteiro assim chamadas porque obedecem a estatística de Fermi-Dirac. São classificados de acordo com a maneira que interagem (ou por quais cargas eles transportam) [5]. Entre os férmions existem os quarks e léptons. Quarks são partículas de carga elétrica fracionária que interagem via força forte e aparecem em seis *sabores* (*flavours*): up, down, strange, charm, top e bottom. Partículas compostas por quarks são chamadas hádrons, como por exemplo o próton que é formado por dois quarks up e um quark down (uud) ou o nêutron que é formado por dois quarks down e um quark up (udd). Léptons são partículas de carga elétrica inteira e não interagem via força forte. Os léptons com carga elétrica neutra são denominados neutrinos ( $\nu$ ). Léptons carregados são o múon ( $\mu$ ), elétron ( $e$ ) e o tau ( $\tau$ ), cada um possui o seu respectivo neutrino: neutrino do múon, neutrino do elétron e neutrino do tau.

Quarks e léptons aparecem em três gerações indicadas na tabela (2.1). A matéria ordinária que nos cerca consiste basicamente de partículas da primeira geração. Anti-partículas, com massa idêntica mas com carga contrária, existem para cada partícula mostrada. Por exemplo, a anti-partícula do elétron ( $e^-$ ) é o pósitron ( $e^+$ ) que possui a mesma massa que o elétron porém sua carga é positiva. Assim é formado o conjunto de anti-léptons e anti-quarks [6].

Além dos léptons e quarks, existe um terceiro grupo de partículas conhecidas como bósons de Gauge. Essas partículas de spin inteiro são responsáveis por mediar as quatro interações fundamentais. Partículas com carga elétrica interagem via força eletromagnética, que é mediada pelo fóton. A força fraca tem como mediadores os bósons  $W^\pm$  e  $Z^0$ , estes bósons trabalham através de distâncias pequenas, pois possuem uma massa maior comparada com a dos outros bósons. A gravidade foi teorizada para ser mediada pelo gráviton, mas até hoje em dia nenhuma teoria quântica para a gravitação foi bem desenvolvida, assim como nenhuma evidência experimental do gráviton.

A força forte age entre quarks de modo a garantir a estabilidade do núcleo atômico. Esta força é mediada pelos glúons, que não possuem massa nem carga elétrica, porém possuem um diferente número quântico chamado *cor* ou *carga de*

Tabela 2.2: Bósons de Gauge descritos pelo Modelo Padrão.

Partícula	Carga elétrica [e]	Massa [GeV]	Interação
$Z$	0	91.2	fraca
$W^+$	+1	80.4	fraca
$W^-$	-1	80.4	fraca
$\gamma$	0	0	eletromagnética
$g$	0	0	forte

*cor*. Quarks aparecem em três cores: vermelho, verde e azul, enquanto antiquarks aparecem como ciano (anti-vermelho), magenta(anti-verde) e amarelo (anti-azul). Somente partículas cuja a combinação de seus constituintes elementares resultam em branco ou sem cor (*colorless*) são permitidas de existirem livremente na natureza. Assim partículas contendo três quarks são compostas de quarks onde cada um tem uma carga de cor diferente (vermelho, verde, azul) e partículas contendo dois quarks são compostas de um par de quarks cor-anticor. A tabela (2.2) sumariza os Bósons de Gauge.

## 2.3 Bóson de Higgs

No início da década de 1960 Goldstone propôs que a força fraca e eletromagnética, a energias muito elevadas, podem ser descritas pela mesma teoria. Essa unificação implica que a eletricidade, magnetismo, luz e alguns tipos de radioatividade são todas manifestações de uma força fundamental batizada de *eletrofraca*. Mas os resultados dos cálculos dessa unificação mostravam que as partículas mediadoras não possuíam massa.

Em 1964, Peter Higgs e François Englert propuseram um novo mecanismo de quebra de simetria, com um bóson, batizado de Bóson de Higgs, que seria o responsável por gerar massa de todas as partículas, inclusive dos bósons intermediários. Vale lembrar que o fóton nesse modelo continuava a ter massa zero, conforme observamos experimentalmente.

A aquisição de massa por uma partícula poderia ser explicada da seguinte maneira: o campo de Higgs estaria permeando todo o espaço; a partícula mediadora desse campo seria o Bóson de Higgs. Uma partícula real nesse espaço interagiria e ficaria polarizada com bósons de Higgs, haveria então uma nuvem de bósons de Higgs associada à partícula lhe dando massa. Metaforicamente igual a uma pessoa que caminha com lama até a cintura, ou seja, um *campo de lama*, essa interação lhes dá sua massa, sua inércia.

Uma vez produzido, a instabilidade do bóson de Higgs faz com que ele decaia instantaneamente em outras partículas [7]. A probabilidade de decaimento para

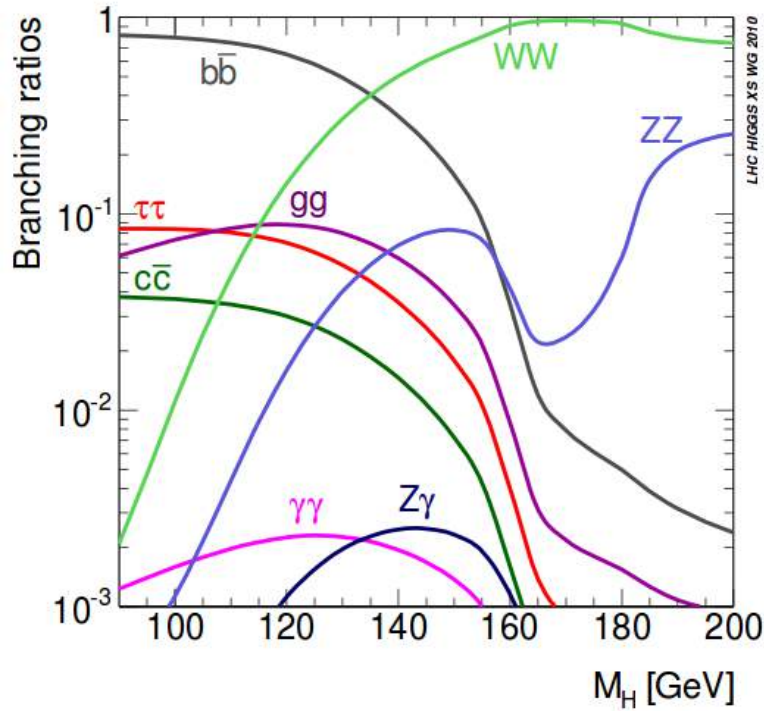


Figura 2.1: Branching Ratio do bóson de Higgs para diferentes canais de decaimento em função de sua massa. Fonte: [7]

cada um dos canais é quantificada pela sua razão de ramificação, *branching ratio* (BR), mostrada na figura 2.1 e os valores exatos mostrados na tabela 2.3. Para a região de  $m_H < 130$  GeV os canais dominante são  $H \rightarrow b\bar{b}$ ,  $H \rightarrow \tau^+\tau^-$  e  $H \rightarrow gg$ . Acima desta massa os decaimentos mais prováveis são  $WW$  e  $ZZ$ . Experimentos de física de altas energias que pretendem medir o bóson de Higgs são guiados por estas previsões do Modelo Padrão.

Apesar do decaimento  $H \rightarrow b\bar{b}$  (Higgs decaindo em dois jatos  $b$ ) ser o que possui a maior probabilidade, seu *background* é enorme, o maior entre todos os decaimentos, portanto corrigir a medida de energia destes jatos implica também em corrigir a energia do Higgs num dos canais de maior dificuldade de detecção. Os jatos do tipo  $b$  são descritos com detalhamento no capítulo 4.

No dia 4 de Julho de 2012, os experimentos ATLAS e CMS no LHC, anunciaram a descoberta de um novo bóson com massa de aproximadamente  $m_H = 125$  GeV e propriedades consistentes com as previstas pelo Modelo Padrão [9] [10]. Após essa descoberta, os experimentos do LHC entraram em uma nova fase de medidas mais precisas para as propriedades do bóson de Higgs [11] [12] [13].



Tabela 2.3: Razão de Ramificação (*Branching Ratios*) para todos os canais de decaimento do bóson de Higgs, para  $m_H = 125.4$  GeV segundo as previsões do Modelo Padrão. Fonte: [8]

Canal de Decaimento	<i>Branching Ratio</i>
$b\bar{b}$	0.571
$WW$	0.221
$gg$	$8.53 \times 10^{-2}$
$\tau\tau$	$6.25 \times 10^{-2}$
$c\bar{c}$	$2.88 \times 10^{-2}$
$ZZ$	$2.74 \times 10^{-2}$
$\gamma\gamma$	$2.28 \times 10^{-3}$
$Z\gamma$	$1.57 \times 10^{-3}$
$\mu\mu$	$2.17 \times 10^{-4}$

# Capítulo 3

## O Experimento ATLAS no LHC

O desejo contínuo da humanidade em desvendar a complexa estrutura do nosso Universo e compreender as interações fundamentais e suas partículas elementares levou ao desenvolvimento de experimentos com aceleradores de partículas para reproduzir e analisar as condições próximas ao nascimento do nosso Universo, enquanto estimula e promove constantemente a inovação em engenharia e o progresso tecnológico. Junto com muitas descobertas da física e marcos tecnológicos, este desejo e missão tem sido perseguido desde 1954 no CERN, Centro Europeu de Pesquisas Nucleares (*Conseil Européen pour la Recherche Nucléaire*) perto de Genebra, Suíça. A mais recente adição ao complexo de aceleradores no CERN é hoje o maior e mais poderoso acelerador de partículas da história, o *Large Hadron Collider* (LHC) [14].

Além dos experimentos de física de altas energias, o CERN desenvolve equipamentos e tecnologias que são utilizados em diversas áreas do conhecimento científico, tais como a invenção da *World Wide Web*, o desenvolvimento de detectores altamente segmentados para a visualização de imagens em medicina, de equipamentos de radioterapia, de tecnologias de computação em nuvem, etc...

Este capítulo descreve o acelerador LHC e seu experimento ATLAS.

### 3.1 O Grande Colisor de Hádrons (LHC)

O LHC foi construído em um túnel de 27 km de circunferência, a 100 metros abaixo da superfície, situado nos arredores de Genebra, na fronteira da França com a Suíça. O túnel foi originalmente construído para abrigar o acelerador LEP [15]. O LHC colide feixes de prótons com energia no centro de massa equivalente a 14 TeV e possui luminosidade instantânea, que representa a taxa de colisões por unidade de área,  $\mathcal{L} = 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ , bem como feixes de íons pesados (**Pb**) com energia de 2.76 TeV por nucleon e uma luminosidade de  $\mathcal{L} = 10^{27} \text{ cm}^{-2}\text{s}^{-1}$ .

Feixes de prótons ou íons pesados são colididos em quatro pontos de interação onde os principais experimentos estão localizados. São eles: ALICE (*A Large*

*Ion Collider Experiment*) [16], desenvolvido para estudar a física dos íons pesados, ATLAS (*A Toroidal LHC ApparatuS*) [17] e CMS (*Compact Muon Solenoid*) [18] que são detectores de propósito geral, ou seja, estudam toda a física envolvida nas colisões próton-próton e o LHCb (*Large Hadron Collider Beauty Experiment*) [19], especializado na física do quark  $b$ .

Como o LHC colide partículas com a mesma carga elétrica, próton-próton, dois diferentes anéis são necessários dentro do túnel para que os feixes de prótons circulem em direções opostas. Os prótons passam por diversos pré-aceleradores, como visto esquematicamente na figura 3.1, e são agrupados em pacotes (*bunches*). Com um espaço nominal por pacote de 25 ns, aproximadamente 2808 pacotes de prótons são necessários para preencher o LHC. Cada pacote possui em torno de  $10^{11}$  prótons. Dipolos e quadrupolos magnéticos, utilizam supercondutividade para gerar campos magnéticos que excedem 8 Tesla, operando em temperaturas abaixo de 2 Kelvin. São eles os responsáveis por manter os feixes na órbita circular do acelerador. Cavidades de rádio-frequência são usadas no processo de aceleração [14].

A fonte dos prótons é um simples recipiente com gás hidrogênio onde os elétrons são retirados dos átomos de hidrogênio utilizando um campo elétrico, assim os prótons restantes do átomo são levados para o acelerador linear (LINAC2) de forma a aumentar sua energia para 50 MeV antes de injetá-los no *Proton Synchrotron Booster* (PSB), atingindo 1.4 GeV de energia. Posteriormente os prótons são introduzidos no *Proton Synchrotron* (PS) onde são acelerados a 25 GeV e em seguida no *Super Proton Synchrotron* (SPS) que finalmente injeta os pacotes de prótons no LHC com uma energia de 450 GeV e lá são novamente acelerados até a energia necessária para a colisão.

## 3.2 O Detector ATLAS

O detector ATLAS (*A Toroidal LHC ApparatuS*) [17] representa umas das maiores e mais complexas máquinas científicas já construídas e foi desenvolvida para o ambiente de alta luminosidade presente no LHC (seção 3.1). Possui 44 m de comprimento, 25 m de diâmetro e pesa aproximadamente 7000 toneladas distribuídas em seu formato cilíndrico como mostrado na figura 3.2

O ATLAS consiste de um sistema de sub-detectores, com alta granularidade e resistência à radiação, para realizar medidas de alta precisão. São eles o *Inner Detector* [21], o calorímetro eletromagnético de argônio líquido (LAr) [22], o calorímetro hadrônico de telhas cintilantes (TileCal) [23] e o sistema de múons [24]. Quando as partículas atravessam cada sub-sistema elas interagem com os seus diferentes materiais e depositam sua energia de diferentes maneiras. Este princípio de identificação de partículas é mostrado na figura 3.3 e descrito a seguir.

## CERN's Accelerator Complex

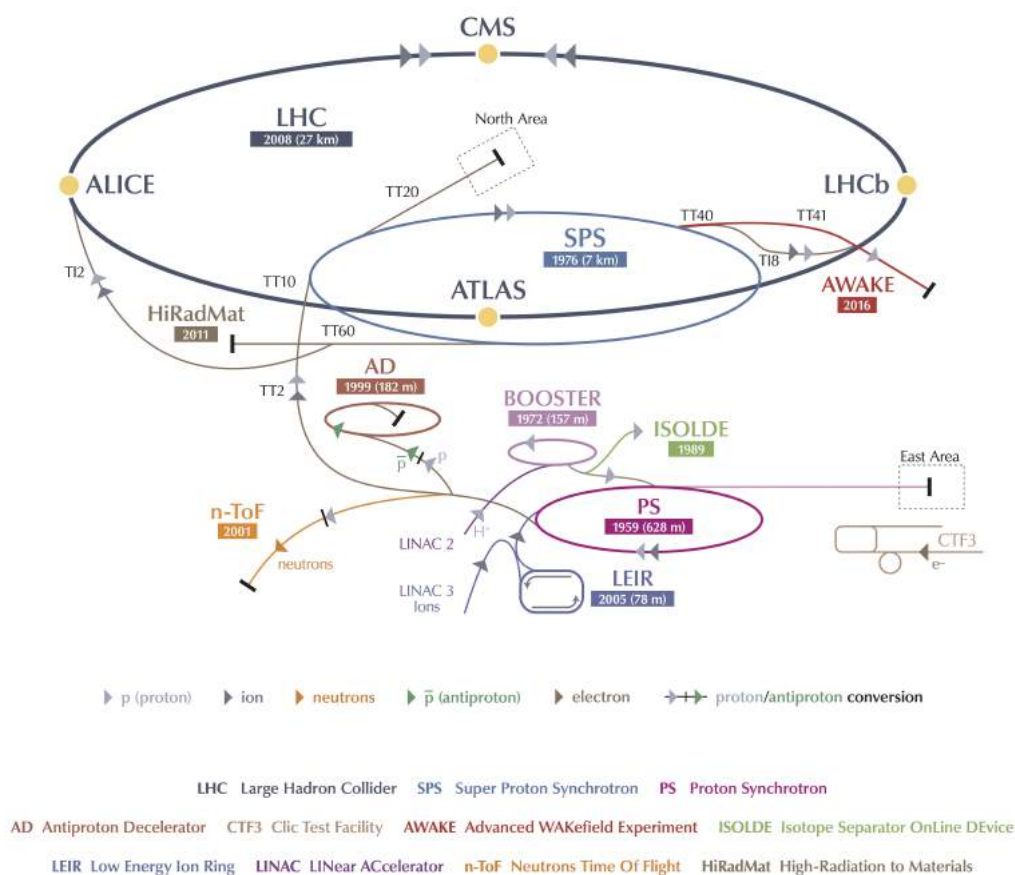


Figura 3.1: Visão esquemática do LHC com seus 4 experimentos principais: ALICE, ATLAS, CMS e LHCb e os aceleradores antigos que agora são usados como pré-aceleradores para o LHC. Sistemas secundários, utilizados em outros experimentos, também pode ser vistos na figura. Fonte: [20]

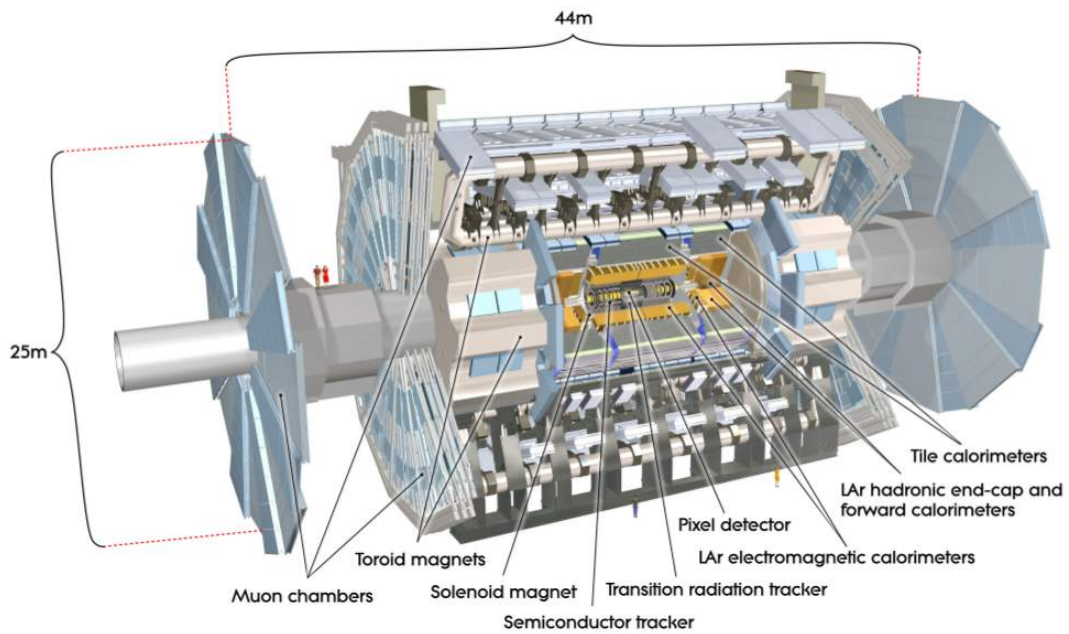


Figura 3.2: Visão esquemática do detector ATLAS. Fonte: [17]

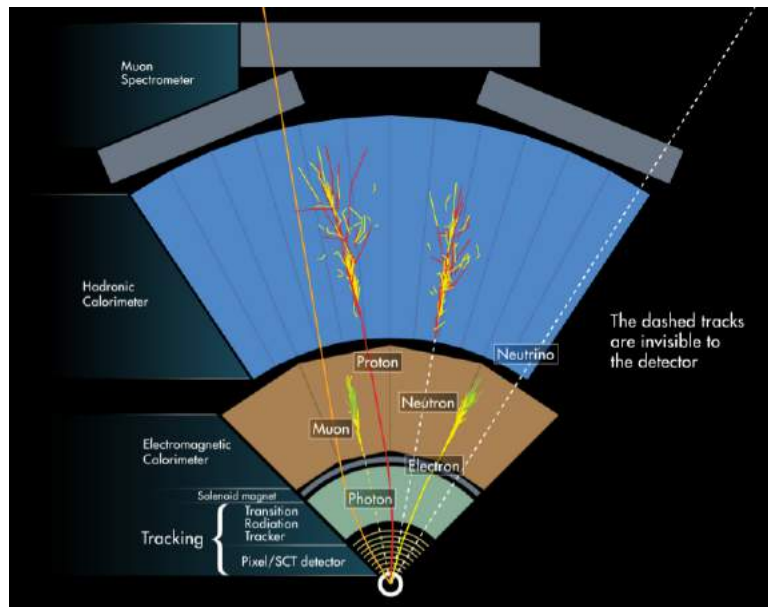


Figura 3.3: Ilustração de diferentes partículas interagindo com os sub-detectores do ATLAS. Fonte: [18]

A camada mais interna do ATLAS representa o seu sistema de traços que é operado dentro de um campo magnético solenoidal de forma a curvar a trajetória das partículas carregadas (elétrons, múons, hádrons) que passam por ali usando a força de Lorentz. Assim é possível medir o momento da partícula a partir de sua curvatura. Além disso as trajetórias identificadas pelo sistema de traços também são usadas para reconstruir o vértice de interação da colisão.

A camada intermediária é composta do sistema de calorimetria para absorver as partículas incidentes e medir sua energia. Como em experimentos de física de altas energias a taxa de eventos é elevada, esses detetores devem ter uma resposta rápida e uma boa granularidade. O calorímetro eletromagnético (EM) utiliza material cintilante para reconstruir o chuveiro de partículas. Já o calorímetro hadrônico utiliza material pesado, de alto número atômico, para reconstruir os chuveiros hadrônicos.

A camada externa do detector se trata do espectrômetro de múons que opera dentro de um campo magnético toroidal para curvar a trajetória dos múons e medir seu momento. Partículas que não interagem com o detector como os neutrinos são reconstruídas utilizando a energia perdida do evento. Uma discussão mais detalhada da reconstrução dos objetos utilizados nessa dissertação é apresentada no capítulo 4.

### 3.2.1 Sistema de Coordenadas do ATLAS

Uma descrição consistente das regiões do detector e das interações de partículas exige a definição de um sistema de coordenadas, que é introduzida aqui e usada ao longo desta dissertação.

O ponto de interação, onde ocorrem as colisões, marca o centro do sistema de coordenadas com o eixo-z ao longo do eixo do feixe e o plano x-y transverso à linha do feixe. Um z positivo é definido na direção anti-horária do LHC e o z negativo na direção horária. O x positivo é definido horizontalmente na direção do centro do anel do LHC e o y positivo é definido verticalmente para cima a partir do ponto de interação. Este sistema de coordenadas é conhecido como *right-handed coordinate system*.

Grandezas físicas transversas usadas nessa dissertação possuem o índice  $T$  como por exemplo o momento transverso  $p_T$  e são medidas no plano x-y. O ângulo azimutal  $\phi$  e o raio  $\rho$  são medidos ao redor da linha do feixe no plano x-y relativo ao eixo-x positivo e o ângulo polar  $\theta$  é medido a partir da linha do feixe e relativo ao eixo-z positivo. O ângulo polar pode ser convenientemente expressado através da pseudo-rapidez:

$$\eta = -\ln \left[ \tan \left( \frac{\theta}{2} \right) \right], \quad (3.1)$$

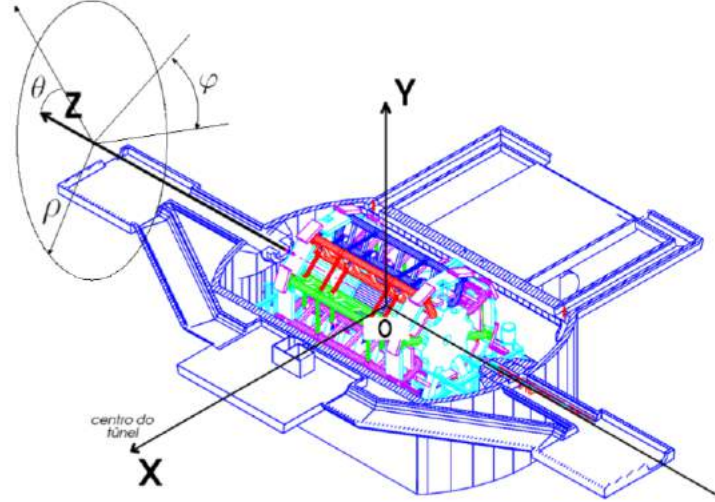


Figura 3.4: Sistema de coordenadas do ATLAS. Fonte: [17]

que é uma grandeza amplamente utilizada no estudo de colisões hadrônicas. Uma outra quantidade importante é a distância entre duas partículas no plano  $\eta - \phi$ :

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \quad (3.2)$$

A figura 3.4 mostra o sistema de coordenadas adotado para o detector ATLAS. As relações entre o momento das partículas e as grandezas  $\eta$  e  $\phi$  são dadas pelas equações 3.3 e 3.4

$$\begin{aligned} p_x &= p_T \times \cos \phi \\ p_y &= p_T \times \sin \phi \\ p_z &= p_T \times \sinh \eta \end{aligned} \quad (3.3)$$

onde  $p_T$  é o momento transverso definido como:

$$p_T = p \times \sin \theta. \quad (3.4)$$

### 3.3 Detector Interno

O Detector Interno do ATLAS [21], do inglês *Inner Detector* - ID, é um conjunto de detectores de traços de alta precisão colocado em torno do ponto de interação e projetado para a reconstrução de trajetórias de partículas carregadas com momento transverso  $p_T > 0.1$  GeV e um alcance de pseudo-rapidez  $|\eta| < 2.5$ . Devido à sua alta resolução, o ID é permite reconstruir os vértices primário e secundário do evento, que por definição são os pontos de interação e o início do decaimento das partículas, respectivamente. Os vértices secundários são cruciais para a identificação

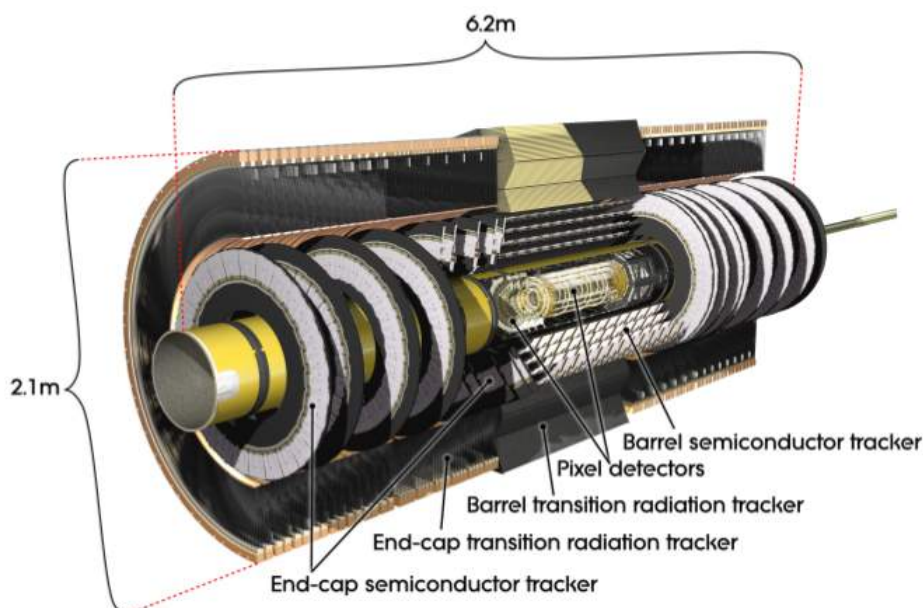


Figura 3.5: Visão geral do Inner Detector com os sub-detectores Pixel, SCT e TRT para as regiões do barril e end-cap. Fonte: [21]

de partículas compostas pelo quark  $b$ , bem como os jatos originados por um quark  $b$  (figura 4.5).

O primeiro sub-sistema do ID é o *Pixel Detector* [21] que possui resolução da ordem de  $10^{-6}$  m e utiliza uma estrutura de pixels de silício como célula de leitura. O segundo sub-sistema é o *Semi Conductor Tracker* (SCT) [25] com células de fibra de silício, semelhantes ao Pixel Detector, mas com uma resolução menor. O terceiro sub-sistema é o *Transition Radiation Tracker* (TRT) [26] que possui fibras preenchidas com um material que emite radiação de transição quando uma partícula o atravessa. A reação gerada pela partícula carregada ao passar pela fibra preenchida permite a leitura dos dados do traço nesta região. O ID é mostrado na figura 3.5.

### 3.4 Sistema de Calorimetria

A calorimetria é baseada em processos de medidas destrutivas nas quais a partícula incidente é totalmente absorvida pelo material do calorímetro produzindo um chuveiro de partículas secundárias através de interações de eletromagnéticas ou de força forte. A energia depositada pelas partículas pode ser detetada na forma de luz servindo assim como medida de energia da partícula incidente. O sistema de calorimetria do ATLAS, como mostrado na figura 3.6, consiste de calorímetros de amostragem que possuem camadas intercaladas de um material absorvedor, responsável por degradar a energia da partícula, e um meio ativo que gera o sinal de energia.

A camada interna do calorímetro consiste em calorímetros eletromagnéticos



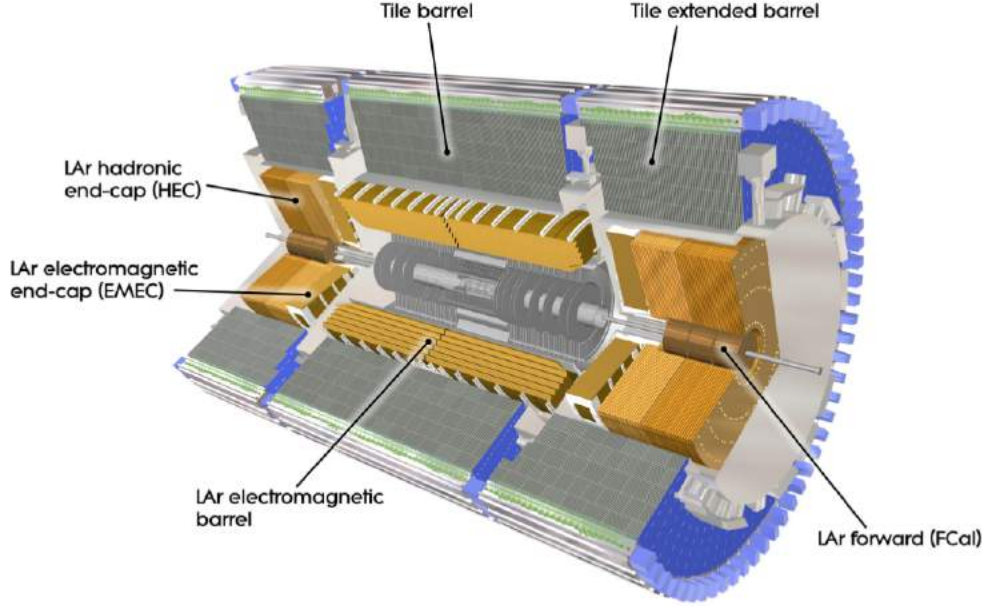


Figura 3.6: Layout do sistema de calorimetria do ATLAS. Fonte [17]

que detectam, principalmente, elétrons e fótons através das suas interações eletromagnéticas (*bremstrahlung* e produção de pares de partículas) e são complementados pela camada externa de calorímetros hadrônicos que detectam hádrons através da interação forte. O calorímetro cobre uma região de pseudo-rapidez  $|\eta| < 4.9$ . Em conjunto com as informações do ID e da câmara de múons permite reconstruir a trajetória de neutrinos utilizando a energia transversa perdida do evento.

A resolução na medida de energia do calorímetro pode ser generalizada na expressão:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c, \quad (3.5)$$

onde  $\oplus$  significa soma em quadratura.

O termo  $a$  na equação 3.5 representa a parte estocástica que existe, principalmente, devido as flutuações estatísticas intrínsecas relacionadas ao chuveiro. Geralmente a resolução melhora, ou seja, a incerteza na energia medida pelo calorímetro diminui conforme a energia aumenta [27]. Essa incerteza é um parâmetro do projeto do detector. Para o calorímetro do ATLAS tem-se:

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E [GeV]}} \oplus 0.7\% \quad (3.6)$$

ou seja, há uma incerteza de 10% na medida quando a energia depositada no calorímetro for de 1 GeV, acrescido de um erro sistemático de 0.7%.

O segundo termo,  $b$ , é derivado do ruído eletrônico e de sua propagação na cadeia de aquisição de dados. A redução deste termo é feita aumentando a fração

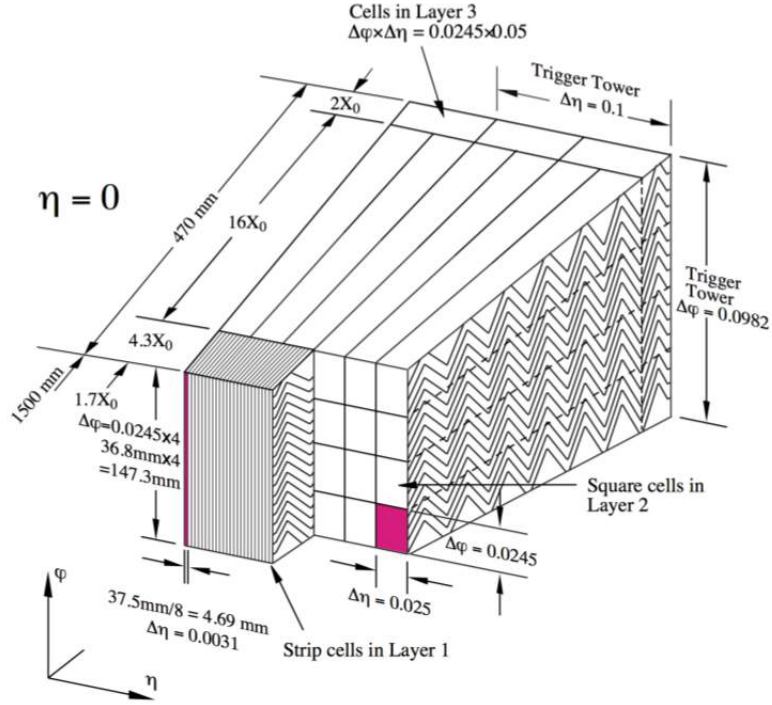


Figura 3.7: Ilustração das diferentes camadas de um módulo do barril do calorímetro eletromagnético do ATLAS. Fonte: [17]

de amostragem do calorímetro de forma a aumentar o sinal do meio ativo e adquirir uma razão sinal-ruído maior. O termo constante,  $c$ , não leva em consideração as contribuições que estão diretamente ligadas à energia da partícula, mas sim às não-homogeneidades do detector e às incertezas na sua calibração.

### 3.4.1 Calorímetro Eletromagnético

Como dito na seção 3.4 a parte mais interna do sistema de calorimetria compreende o calorímetro eletromagnético (EM), um calorímetro heterogêneo porque utiliza placas de chumbo como material passivo e absorvedor de partículas e argônio líquido como material ativo (LArg). O calorímetro EM inclui a parte central, também chamada de barril, que cobre  $|\eta| < 1.475$ , e dois *end-caps* cobrindo  $1.375 < |\eta| < 3.2$ . Um módulo do calorímetro é mostrado na figura 3.7. Para obter uma cobertura completa em  $\phi$  o material passivo é dobrado em forma de sanfona e ao longo das placas de chumbo são adicionados inúmeros finos eletrodos de cobre que, posteriormente, possibilitam uma visão em três dimensões da deposição de energia das partículas.

As placas de chumbo são submersas no argônio líquido, de forma a intercalar os materiais passivos e ativos do calorímetro, e sujeitas a um forte campo elétrico. Quando um elétron atravessa o calorímetro, ele interage com os átomos do material passivo (chumbo) liberando fótons. Estes, por sua vez, decaem em um par elétron-pósitron que também reagem com o chumbo. Assim se forma o chuveiro eletro-

magnético que só termina quando toda a energia é absorvida pelo calorímetro. Este esquema pode ser visualizado na figura 3.3. Durante este processo íons de argônio são formados e migram para o lado positivo/negativo devido ao campo elétrico, isso gera uma corrente elétrica nos eletrodos, colocados ao longo das placas de chumbo, que alimentam um circuito externo de aquisição.

A resolução de energia exigida é dada pela equação 3.6. Testes com feixes reais determinaram o termo estocástico  $a$  como 10–12%, e o termo constante  $c$  0.2–0.35% [28].

### 3.4.2 Calorímetro Hadrônico

O sistema de calorímetros hadrônicos se encontra ao redor do calorímetro EM, contornando-o completamente. A região do barril utiliza tecnologia de sensores óticos, enquanto a região das tampas *end-caps* utiliza argônio líquido, como no calorímetro EM. O calorímetro hadrônico, assim como o EM, ocupa a região de  $|\eta| < 2.5$  e todo o espaço em  $\phi$ . As duas funções básicas do calorímetro hadrônico são: a medição de energia dos jatos (capítulo 4) e a absorção das demais partículas, de origem hadrônica, não absorvidas no calorímetro EM. O sistema possui o *Tile Calorimeter* (TileCal), o *Hadronic End-cap Calorimeter* (HEC) e o *Forward Calorimeter* (FCal).

O **TileCal** consiste num barril central e mais dois barris que se estendem para cada lado cobrindo a região de  $|\eta| < 1.7$ , utiliza placas de ferro como material absorvedor intercalado com com telhas de plástico cintilante como meio ativo. Os hádrons entram no calorímetro, interagem com o núcleo do material produzindo partículas neutras e carregadas. Essas partículas carregadas ionizam o cintilador produzindo luz ultravioleta. Estes flashes de luz são então coletados e levados aos tubos fotomultiplicadores para a leitura [29].

O **HEC** é formado por dois círculos em cada região do end-cap, cobre um alcance de  $1.5 < |\eta| < 3.2$  e usa placas de cobre como material absorvedor e argônio líquido como (LAr) como meio ativo. A carga de ionização deixada pelo chuveiro de partículas no LAr é coletada através de uma alta voltagem aplicada entre os eletrodos, produzindo um pulso de corrente para o sinal de leitura [30].

O **FCal** consiste de três módulos cilíndricos em cada região do end-cap, cobrindo uma pseudo-rapidez de  $3.1 < |\eta| < 4.9$  e usa cobre e tungstênio como meio passivo e LAr como meio ativo [31].

A resolução de energia exigida é:

$$\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\% \quad (3.7)$$

para o TileCal + HEC e

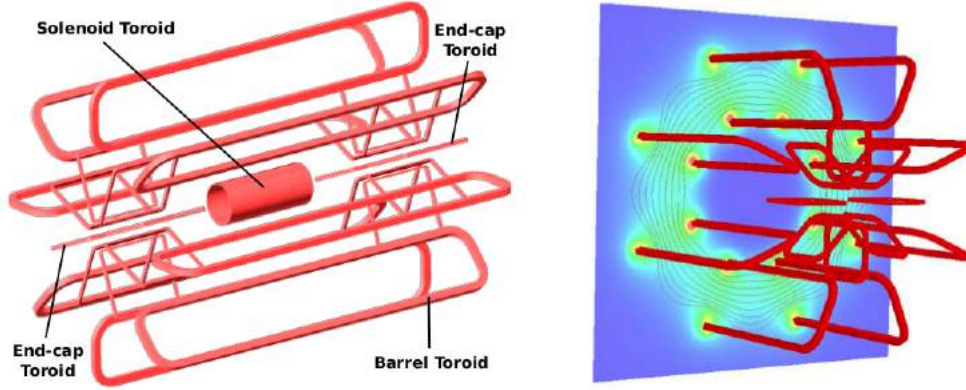


Figura 3.8: Geometria do sistema magnético do ATLAS (esquerda) e uma simulação do mapa do campo magnético (direita). Fonte: [32]

$$\frac{\sigma_E}{E} = \frac{100\%}{\sqrt{E}} \oplus 10\% \quad (3.8)$$

para o FCal. Testes com os feixes determinaram os termos estocásticos e constantes  $a$  ( $c$ ) igual a 56% (5.5%) para o TileCal, 71% (5.8%) para o HEC e 70% (3.0%) para o FCal.

### 3.5 Espectrômetro de Múons

Múons são partículas que interagem fracamente com a matéria (*minimum ionizing particles*) e por possuírem uma vida média alta são as únicas partículas, carregadas, provenientes da colisão que passam por todos os sub-detectores do ATLAS. Logo, o espectrômetro de múons (MS) faz parte da camada mais externa do ATLAS e envolve todos os calorímetros [24].

O MS utiliza um complexo campo magnético toroidal para curvar os múons no plano  $\rho - z$  (seção 3.2.1), que é ortogonal ao campo do solenoide no ID [32] como mostrado na figura 3.8, e possui diferentes câmaras na região do barril e end-cap. Estes sub-detectores são classificados em dois tipos: (i) câmaras de traços de alta precisão para reconstruir o momento dos múons e (ii) câmaras de resposta rápida para identificar de qual colisão o múon foi originado. Os detectores do tipo (i) são o *Monitored Drift Tube* (MDT) e o *Cathode Strip Chamber* (CSC) e os pertencentes ao tipo (ii) são o *Resistive Plate Chamber* (RPC) e o *Thin Gap Chamber* (TGC). Todos eles utilizam alta voltagem através dos eletrodos para a carga de ionização depositada pelos múons que a travessaram o sistema.

O MS possui diferentes câmaras na região do barril e end-cap. Estes sub-detectores são classificados em dois tipos: (i) câmaras de traços de alta precisão para reconstruir o momento dos múons e (ii) câmaras de resposta rápida para iden-

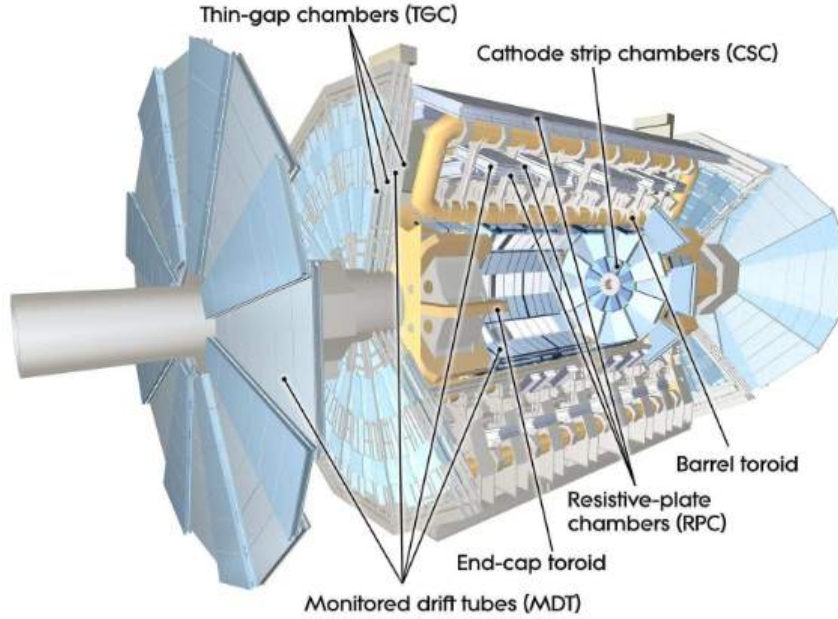


Figura 3.9: Visão geral do espectrômetro de múons do detector ATLAS. Fonte: [17]

tificar de qual colisão o múon foi originado. Os detectores do tipo (i) são o *Monitored Drift Tube* (MDT) e o *Cathode Strip Chamber* (CSC) e os pertencentes ao tipo (ii) são o *Resistive Plate Chamber* (RPC) e o *Thin Gap Chamber* (TGC). A figura 3.9 mostra uma visão geral do sistema de múons no ATLAS.

### 3.6 Sistema de Aquisição de Dados e Trigger

Com a luminosidade no LHC sendo  $L = 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ , como discutido na seção 3.1, as colisões possuem uma taxa de interação de aproximadamente  $10^9$  eventos/s que é em muitas ordens de magnitude maior do que os raros processos físicos de interesse, como por exemplo a produção do Bóson de Higgs. O ATLAS captura os dados dos eventos utilizando milhões de canais de leitura, porém há um limite tecnológico na quantidade de dados que podem ser gravadas para uma análise *offline*. Logo, deve ser implantado um sistema de seleção *online* de eventos com física de interesse para as colaborações. Esse sistema de filtragem precisa ser cuidadosamente desenvolvido, de forma a atingir uma alta taxa de rejeição de ruído de fundo, ao mesmo tempo que possui uma alta eficiência em relação à seleção dos canais físicos importantes. Esta árdua tarefa é feita pelo sistema de Aquisição de Dados e Trigger do ATLAS [33] [34], do inglês *Trigger and Data Acquisition - TDAQ*, cujo objetivo final é reduzir a taxa de armazenamento de eventos de 40 MHz para 200 Hz. Em outras palavras, um dado de evento possui tipicamente 1.5 MB de tamanho, o sistema TDAQ a taxa de armazenamento de 60 TB/s para 300 MB/s.

O sistema TDAQ é baseado em três níveis sequenciais de filtragem:

- **Nível 1:** Integra uma quantidade limitada de informação de todos os subsistemas do calorímetro e das câmaras de múons para identificar elétrons, fótons, jatos e múons de alta energia transversa,  $E_T$ . O sistema de filtragem do calorímetro procura por depósitos de energia em células com granularidade  $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$  que estão acima de uma série de limiares previamente determinados. Já o sistema de filtragem dos múons identifica padrões que são consistentes com múons originados do ponto de colisão. O primeiro nível reduz a taxa de eventos de 1 GHz para 75 kHz com cada decisão de filtragem sendo executada em menos de  $2.5 \mu s$ .
- **Nível 2:** O nível 2 é considerado um *trigger* de alto nível, de forma que toma a decisão de filtragem baseada em toda a granularidade do detector e também de informações do detector interno (ID). Esta filtragem só é operada nas regiões de interesse marcadas pelo primeiro nível. Se nenhuma das regiões de interesse forem consistentes com uma hipótese para uma determinada partícula, o evento é rejeitado. A taxa de evento é novamente reduzida, desta vez para 3.5 kHz, com um tempo de processamento de 40 ms para cada evento.
- **Nível 3:** Também chamado de *Event Filter*, utiliza a informação completa de cada evento e melhora assim a filtragem realizada no nível 2. O processamento de eventos é feito de forma *offline* numa média de quatro segundos por evento, e a taxa é reduzida para 200 Hz. Eventos que passam deste nível são transferidos para o centro de computadores do CERN para armazenamento permanente.

# Capítulo 4

## Reconstrução e Identificação de Jatos

Jatos são manifestações dos quarks e glúons e são assim denominados porque consistem em um cone de partículas colimadas, hádrons de alto  $p_T$ , resultantes da fragmentação e hadronização dos pártons. Os jatos são definidos por algoritmos de clusterização responsáveis pela sua reconstrução no detector. Combinações de jatos são usadas para identificar partículas pesadas como o top quark [35], W, Z [36] [37] e o Bóson de Higgs [38]. Portanto, o sucesso da análise consiste na habilidade de reconstruir e medir a energia dos jatos com a maior exatidão possível.

Qualquer grandeza física medida pelo detector é afetada por suas imperfeições, ou seja, as não-linearidades presentes no aparato experimental ou regiões do detector com baixa resolução. Entretanto, quando trabalhamos com amostras simuladas, usando métodos de Monte Carlo, podemos ter acesso tanto ao evento verdadeiro, *truth level*, gerado de acordo com a física presente no Modelo Padrão, quanto ao evento reconstruído pela eletrônica após a interação das partículas com a matéria que forma os diferentes detectores (*reconstructed level*). O resultado final das simulações são as amostras de Monte Carlo no mesmo formato que as amostras de dados reais com a informação adicional do evento verdadeiro previsto pelo processo físico. A figura 4.1 ilustra as etapas para a simulação de eventos de Monte Carlo, em paralelo com a aquisição de dados reais, para o detector ATLAS, junto com o *software* utilizado em cada tarefa.

Nas amostras simuladas podemos então quantificar o quão distante estatisticamente os eventos gerados, *truth level*, estão dos eventos reconstruídos. Com isso podemos ir mais adiante e derivar métodos de calibração de modo a corrigir uma grandeza física reconstruída para que seu valor fique próximo do valor verdadeiro. Essa correção deve ser função apenas de quantidades reconstruídas, isso permite aplicarmos posteriormente a mesma correção em dados reais, que não possuem quantidades geradas.



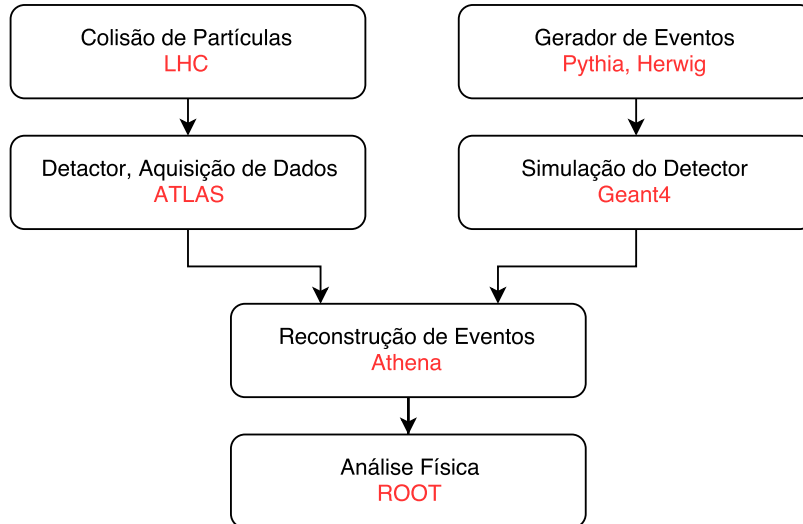


Figura 4.1: Etapas da análise para os dados reais e simulações de Monte Carlo junto com os *softwares* usados em cada etapa.

Contudo se conseguirmos corrigir os dados reconstruídos de forma que a sua distribuição de variáveis se aproxime da distribuição gerada (*truth level*), estaremos medindo melhor os nossos eventos. Para o caso do decaimento do Higgs em dois jatos, se corrigirmos os jatos estaremos corrigindo também a medida do Higgs.

Este capítulo trata dos procedimentos para a reconstrução de partículas e objetos físicos produzidos durante as colisões. Nas próximas seções vários algoritmos são discutidos começando com uma visão geral da reconstrução usada em alguns dos experimentos de física de altas energias e em seguida será dada uma ênfase maior nos objetos físicos utilizados nesta análise, jatos  $b$ , e reconstruídos pela colaboração do experimento ATLAS.

## 4.1 Jatos no Experimento D0

No experimento D0 [39] no Fermilab a reconstrução dos jatos, sejam eles de partículas estáveis ou de torres de energia detectadas pelo calorímetro, envolve o uso de um algoritmo de clusterização para assinalar essas partículas como provenientes de um jato. O algoritmo se chama *Run II Midpoint Cone Algorithm* [40]. Torres de energia no calorímetro com o raio  $\Delta R < \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < R_{jet}$  são clusterizadas [41] e a correção na energia dos jatos é então definida pela seguinte fórmula:

$$E_{ptcl} = \frac{E_{meas} - E_O}{R.S} \quad (4.1)$$

onde:

- $E_O$  representa o *offset* na energia que depende do cone do jato ( $R_{cone}$ ), da



pseudo-rapidez do detector ( $\eta_{det}$ ), do número de vértices reconstruídos ( $n_{PV}$ ) e da luminosidade instantânea ( $L$ ).

- R representa a resposta do calorímetro para as partículas que compõe o jato.
- S são as correções para chuveiros de partículas no detector.

## 4.2 Jatos no Experimento CDF

Também localizado no Fermilab o experimento CDF [42] possui correções específicas para os jatos, tais correções consistem em comparar a resposta dos jatos, entre os dados reais e as amostras de Monte Carlo, para processos físicos diferentes de forma a estimar a energia verdadeira do párton que gerou o jato detectado no calorímetro [43].

As correções são separadas em:

- **Correção dependente de  $\eta$ :** Mesmo após a calibração de energia a resposta do calorímetro do CDF não é uniforme para a pseudo-rapidez ( $\eta$ ), esta correção trata este problema.
- **Absolute Jet Energy Scale:** Esta correção visa transformar a energia do jato medido pelo calorímetro na energia do jato gerado, *particle level*. Simulações são utilizadas para derivar tais correções, portanto a acurácia desse método depende de quanto os jatos estão bem modelados nas simulações utilizadas.
- **Interações Múltiplas  $p\bar{p}$ :** Com luminosidades altas mais de uma interação  $p\bar{p}$  ocorre no mesmo cruzamento do feixe (*bunch crossing*). A interação extra pode aumentar a energia dos jatos, portanto essa correção subtrai a energia extra registrada pelo jato.

Cada uma destas correções envolve inúmeros passos e cada um deles possui uma incerteza sistemática [43]. As correções foram baseadas nos dados tomados entre 2001 e 2004 no acelerador Tevatron [43].

## 4.3 Jatos no Experimento CMS

Os jatos no detector CMS são reconstruídos [44] de acordo com o algoritmo *anti*- $k_T$  [45] onde o processo de clusterização é realizado com a soma dos quadri-momento de cada partícula. O CMS usa uma combinação dos traços de partículas carregadas e a energia depositada no calorímetro para construir o *input* do chamado fluxo

de partículas (*particle flow*) [46]. Esta técnica utiliza propriedades chave do detector CMS como o alto campo magnético gerado, sistema de traços e o calorímetro eletromagnético [47].

A calibração dos jatos no experimento CMS é similar ao ATLAS e consiste em remover a contribuição de *pile-up* [48] dentro dos jatos para em seguida aplicar uma correção multiplicativa derivada pelos eventos de Monte Carlo [49].

O fator multiplicativo é aplicado a cada componente do quadri-vetor dos jatos:

$$p_{\mu}^{cor} = C \times p_{\mu}^{raw} \quad (4.2)$$

onde  $p_{\mu}$  é uma das componentes do quadri-vetor e  $C$  é o fator de correção.

## 4.4 Reconstrução de Jatos no Detector ATLAS

Eficientes algoritmos de reconstrução de jatos se fazem necessários devido a complexidade dos estados finais produzidos nas colisões próton-próton no LHC. Na análise apresentada nesta dissertação os jatos são reconstruídos com o algoritmo anti- $k_t$  [45] que é baseado na deposição de energia nos calorímetros.

### 4.4.1 Cluster no Calorímetro

O calorímetro do ATLAS é fundamental para a reconstrução de elétrons, fótons, energia transversa perdida e, é claro, jatos. Elétrons, fótons e hádrons, ao passar pelo calorímetro, depositam sua energia nas diversas células presentes. Portanto, algoritmos de clusterização são necessários para agrupar essas células e assim calcular toda a energia depositada pelas partículas.

Clusters topológicos, também referidos como *topo-clusters*, são usados como matéria-prima para a reconstrução de jatos [50]. Topo-clusters são construídos a partir de uma célula semente e a ela são, iterativamente, adicionadas células vizinhas, resultando em clusters de tamanho variado. O processo consiste em: uma célula pode ser uma célula semente se a sua razão sinal-ruído estiver acima de um determinado limiar  $t_{seed} > 4$ , então as células vizinhas são adicionadas à célula semente, formando um *proto-cluster*, se também estiverem acima de um limiar  $t_{neighbour} > 2$ . O topocluster é finalizado quando o anel de células com razão sinal-ruído acima de 0 é adicionado como a fronteira do cluster, e assim o processo de clusterização continua até que as sementes tenham se esgotado. A figura 4.2 ilustra a formação de um topocluster.

Como as células vizinhas podem estar em camadas adjacentes do calorímetro, os topo-clusters resultantes são objetos tridimensionais que se expandem pelo calorímetro eletromagnético e hadrônico sem seguir uma forma fixa.

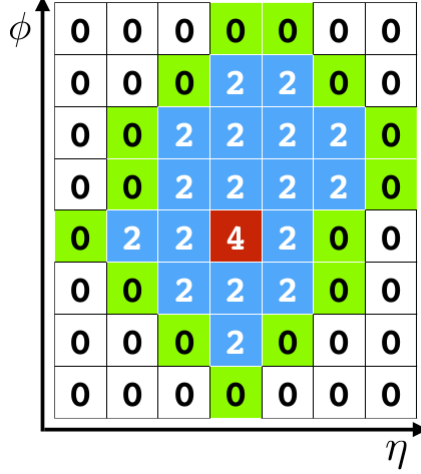


Figura 4.2: Ilustração de um topocluster com a razão sinal-ruído para cada célula.

Os topo-clusters são reconstruídos primeiramente na escala de energia eletromagnética (*EM Scale*) [51] [52] e são utilizados como *inputs* para o algoritmo de reconstrução dos jatos.

#### 4.4.2 Reconstrução - anti-Kt

Após serem selecionados os clusters no calorímetro, é preciso reconstruir os possíveis jatos a partir do depósito de energia de cada topo-cluster, para esta tarefa o algoritmo anti- $k_t$  [45] é utilizado. Definindo  $d_{ij}$  como a distância entre os objetos, partículas e clusters no calorímetro,  $i$  e  $j$  e  $d_{iB}$  entre o objeto  $i$  e o feixe do acelerador  $B$ . Para cada evento, começando com o conjunto completo de clusters, os parâmetros de distância são:

$$d_{ij} = \min(k_{ti}^{-2}, k_{tj}^{-2}) \frac{\Delta_{ij}^2}{R^2} \quad (4.3)$$

$$d_{iB} = k_{ti}^{-2} \quad (4.4)$$

Onde  $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$  e  $k_{ti}$ ,  $y_i$  e  $\phi_i$  são respectivamente o momento transversal, a rapidez e o ângulo azimutal do objeto  $i$ . A constante  $R$  é chamada de parâmetro de raio e nesta análise seu valor é  $R = 0.4$ .

Após todas as distâncias serem calculadas,  $d_{ij}$  e  $d_{iB}$ , a menor distância é selecionada. Se ela for entre os objetos  $i$  e  $j$ , eles são combinados. Se for entre o objeto  $i$  e o feixe,  $i$  é então identificado como um jato e deletado da lista de objetos. Todas as distâncias são recalculadas e o processo se repete até que não restem mais objetos, apenas jatos definidos por cada iteração do algoritmo. Em outras palavras, o algoritmo faz uma recombinação sucessiva dos clusters no calorímetro durante cada iteração de forma a aumentar o  $p_T$ . No ATLAS a implementação da reconstrução

de jatos é feita pelo FASTJET [53] [54]. O formato escolhido para combinar cada objeto leva o jato reconstruído a obter um formato de cone com o raio  $R$  para jatos de alto momento.

Com isso obtemos os jatos na escala de energia eletromagnética, também chamada de calibração EM.

### 4.4.3 Correção JES - *Jet Energy Scale*

Um dos desafios mais difíceis para a reconstrução e calibração de jatos no LHC é a presença de *pile-up*, que são interações próton-próton adicionais produzidas dentro do mesmo evento. O pileup reduz a precisão na medida da energia dos jatos e pode produzir jatos adicionais que não pertencem à interação original.

Esta correção é aplicada em duas etapas. A primeira etapa corrige as contribuições de *pile-up* [48] removendo o efeito adicional causado pela energia das partículas de pileup dentro do jato. Em seguida a direção do jato é modificada de forma a apontar para o vértice de colisão para qual o jato foi gerado e não mais para o centro de massa do detector [55].

Como foi dito no início do capítulo 4, o principal objetivo da calibração dos jatos é fazer com que a escala de energia dos jatos reconstruídos corresponda, em média, à escala de energia dos jatos gerados (*truth level jets*). A resposta (*response*) de energia dos jatos é a figura de mérito que caracteriza tal calibração e é definida como a razão, na simulação de Monte Carlo, entre a energia do jato reconstruído e do jato gerado:

$$R(E, \eta) = \left\langle \frac{E_{EM}}{E_{truth}} \right\rangle \quad (4.5)$$

onde  $E_{EM}$  é a energia do jato reconstruído na escala de energia eletromagnética, seção 4.4.2, e  $E_{truth}$  é a energia verdadeira do jato de acordo com a simulação do Modelo Padrão.

Há inúmeros efeitos que podem fazer com que a resposta da energia nos jatos possua um valor menor que 1 e que precisam ser levados em conta para a correção na JES: a não-compensação no calorímetro devido a diferença de resposta entre os chuveiros eletromagnético e hadrônico, energia perdida em regiões inativas do detector, depósitos de energia abaixo dos limiares ou fora dos topo-clusters utilizados como *input* para a reconstrução dos jatos.

Contudo, a segunda etapa consiste em derivar um fator multiplicativo, dependente da energia e da pseudo-rapidez  $\eta$ , que relacione a energia do jato reconstruído, após a subtração de pileup, ao jato gerado [55]. A figura 4.3 mostra o fator de escala para a correção de energia como função do momento transversal para três intervalos de  $\eta$  e a figura 4.4 mostra a resposta para vários bins de energia e  $\eta_{det}$ .

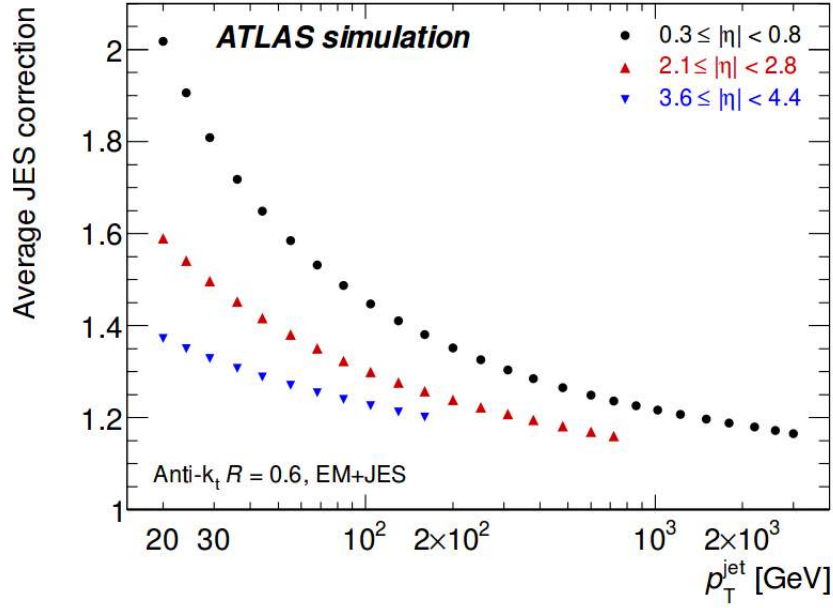


Figura 4.3: Média de correção na escala de energia dos jatos obtida através de simulações de Monte Carlo. A correção só é mostrada na faixa acessível da cinemática. Fonte: [55]

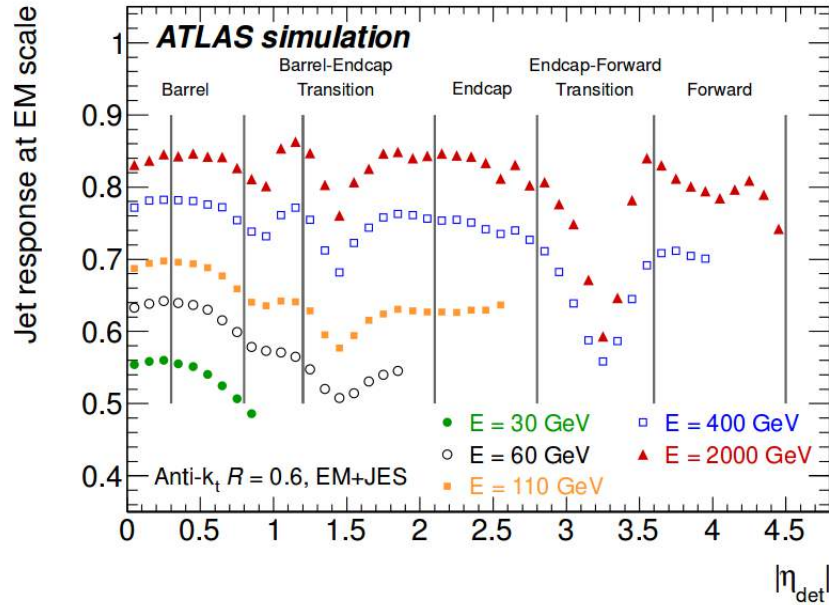


Figura 4.4: Resposta dos jatos,  $R(E, \eta)$ , na escala de energia eletromagnética em função da pseudo-rapidez do detector  $\eta_{det}$ . O inverso da Resposta é igual à média do fator de correção para cada bin. Fonte: [55]

Chamamos a coleção dessas duas correções de *jet energy scale* (JES), pois seu objetivo é melhorar a resposta dos jatos. A correção JES muda a energia, momento e  $\eta$ , mas não aplica nenhuma mudança em  $\phi$ . Apesar de tentar acertar a escala de energia dos jatos a correção EM+JES tem um efeito negativo na resolução que a próxima correção tenta corrigir.

#### 4.4.4 Correção GSC - *Global Sequential Calibration*

Como a correção JES possui um efeito negativo na resolução dos jatos, uma nova correção é aplicada sobre ela. Esta nova correção utiliza propriedades internas dos jatos como por exemplo a fração de energia depositada na última camada do calorímetro eletromagnético e na primeira camada do calorímetro hadrônico, bem como a largura dos jatos, com o intuito de melhorar a resolução dos jatos e manter inalterada, em média, a escala de energia. Assim como a correção JES, a correção GSC também é aplicada em todos os jatos, independente se esses jatos foram originados de glúons ou quaisquer dos quarks, daí o seu nome *global*. Historicamente a correção GSC foi aplicada pela primeira vez na análise  $VH(b\bar{b})$  [38].

Esta correção é baseada no número de traços que possuem  $p_T > 1$  GeV associados ao jato, com a largura do jato e também conta com uma correção em sequência que é aplicada em jatos com alto  $p_T$  para levar em consideração o vazamento de energia fora do calorímetro [56]. O estágio de correção agora é EM+JES+GSC.

#### 4.4.5 *Tagging b-jatos - Identificando jatos b*

Jatos são identificados como sendo originados de um quark  $b$  através do algoritmo MV1 [57] [58] [59] que consiste numa rede neural que utiliza como entrada algumas propriedades inerentes dos jatos  $b$ , como o parâmetro de impacto (IP) e o vértice secundário, figura 4.5. A rede neural é treinada com jatos  $b$  como sinal e jatos *light flavour* como background em eventos com jatos de  $p_T > 20$  GeV em uma amostra de  $t\bar{t}$ .

Três pontos de operação são definidos para o algoritmo de *b-tagging* de acordo com a eficiência na seleção de jatos  $b$ . São eles: *loose* (L), *medium* (M) e *tight* (T) para 80%, 70% e 60% de eficiência, respectivamente. O objetivo do algoritmo é ter uma eficiência baixa na seleção de jatos leves e jatos  $c$ . É conveniente definir os fatores de rejeição como sendo o inverso destas eficiências. Por exemplo, um fator de rejeição com valor 20 significa selecionar um em cada vinte jatos, conseqüentemente uma eficiência de 5%.

As figuras 4.6 mostram o resultado do algoritmo para classificar jatos  $b$ ,  $c$  e leves como jatos  $b$  no ponto de operação a 70% (medium) em bins do  $p_T$  do jato e de  $|\eta|$ .

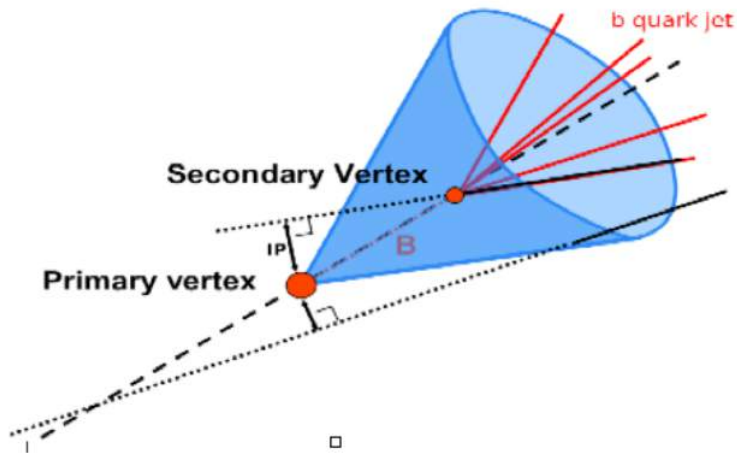
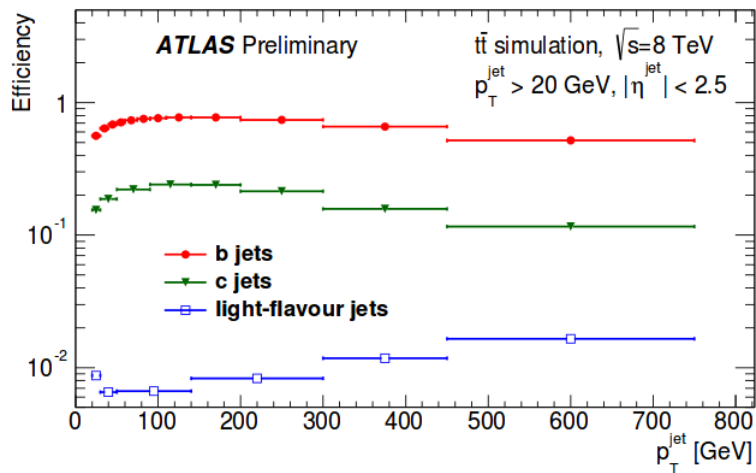
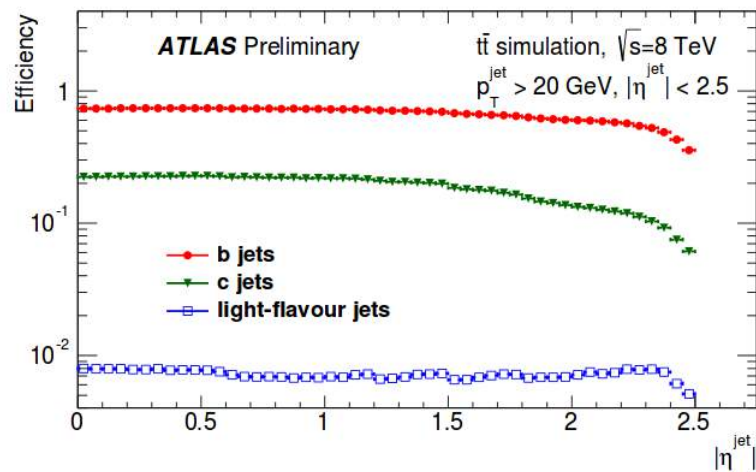


Figura 4.5: Diagrama de um jato b. Fonte: [60]



(a)



(b)

Figura 4.6: Eficiência do algoritmo MV1 para selecionar jatos  $b$ ,  $c$  e jatos leves em função de  $p_T$  (4.6a) e  $\eta$  (4.6b). O discriminante é escolhido para ter 70% de eficiência para jatos  $b$  com o  $p_T > 20$  e  $|\eta| < 2.5$  e avaliado numa amostra de eventos de  $t\bar{t}$ . Fonte: [59]

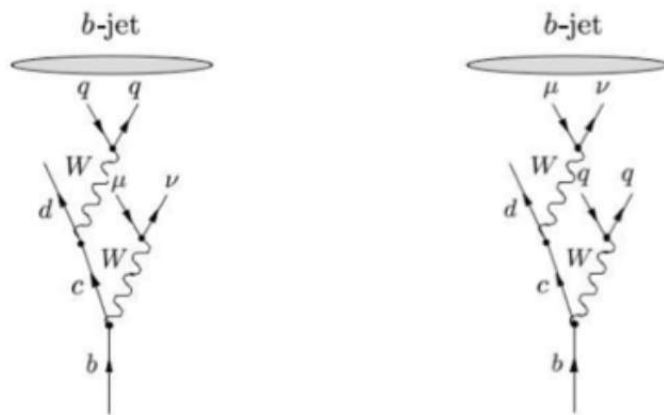


Figura 4.7: Diagramas de Feynman para o decaimento de quarks  $b$ , onde um dos bósons  $W$  decai semi-leptonicamente em um múon e um neutrino, enquanto o outro bóson  $W$  decai hadronicamente. Este processo ocorre em aproximadamente 10.6% dos casos.

Apesar dos jatos serem reconstruídos e calibrados de acordo com a sua energia depositada no calorímetro, seções 4.4.2 e 4.4.3, o *tagging* de jatos  $b$  depende somente nos traços reconstruídos no *inner detector*. Como os traços só podem ser reconstruídos para  $|\eta| < 2.5$ , esta exigência também é necessária para aplicar o  $b$  *tagging*.

#### 4.4.6 *Muon-in-jet correction* para jatos $b$

A calibração de energia dos jatos  $b$  é feita logo após o *tagging*. O quark  $b$  que se encontra dentro do hádron  $B$ , uma das partículas que constituem o jato  $b$ , decai num bóson  $W$  e em um quark  $c$  que em seguida também decai em um quark  $d$  em outro bóson  $W$ . Um destes bósons  $W$  decai semi-leptonicamente, ou seja, em um múon e um neutrino, e o outro bóson decai de forma hadrônica como mostrado na figura 4.7.

Esta característica ocorre em  $(10.6 \pm 0.2)\%$  das vezes [61]. Como os múons são partículas pouco ionizantes, sua energia registrada pelo calorímetro é mínima. Portanto, essa correção trata de corrigir a energia do jato que possui um múon em seu interior, dentro dos limites do cone de jato descrito pelo algoritmo de clusterização topológica, substituindo as informações do quadri-vetor do múon no calorímetro pelas informações provenientes da reconstrução do múon utilizando o *inner detector* e as câmaras de múons. A calibração dos múons também foi introduzida e utilizada em [38] e por definição apenas afeta os 11% dos jatos, os quais possuem um múon dentro de seu cone. Para estes jatos a resposta, equação 4.5, melhora de 20% para 10% menor que o valor ótimo. A diferença restante pode ser explicada pela presença do neutrino, figura 4.7 que escapa completamente da detecção. A nova



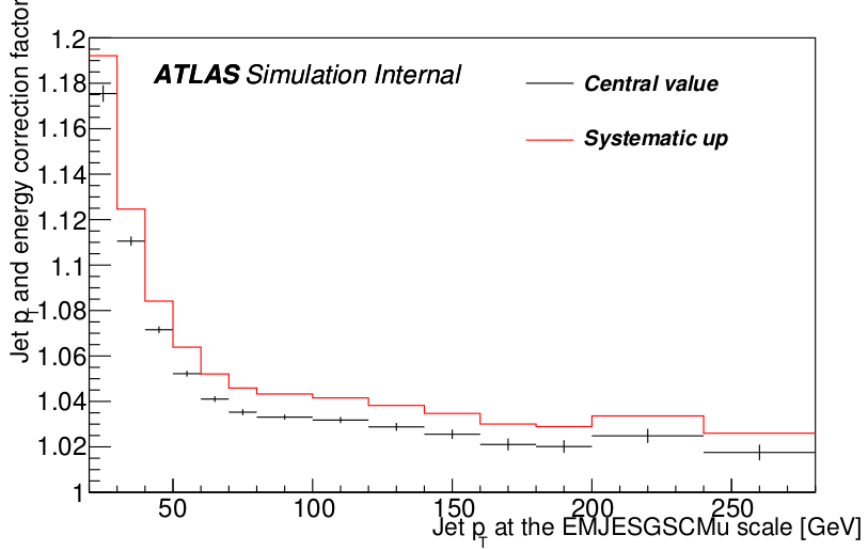


Figura 4.8: Fatores de correção utilizados para cada bin de  $p_T$ . Fonte: [62]

escala é chamada de EM+JES+GSC+Mu.

#### 4.4.7 Correção do Momentum Transverso dos jatos $b$

Mesmo após a correção de energia para os múons dentro do jato, seção 4.4.6, alguns componentes da energia dos jatos ainda não foram considerados durante a reconstrução, como o neutrino presente dentro do cone do jato  $b$ , figura 4.7, e a diferente escala de energia para cada bin de  $p_T$ . A calibração de momento transverso ( $p_T$  correction) traz a escala, definida como a razão entre o objeto reconstruído e o gerado, de cada bin da distribuição de  $p_T$  para o valor unitário. A correção de  $p_T$  também foi introduzida na análise  $VH(bb)$  de 2015 [38]. O fator de correção de cada bin é definido com a média do histograma da razão entre o  $p_T$  gerado e o reconstruído, na escala EM+JES+GSC+Mu, e é mostrado na figura 4.8 para a binagem escolhida da colaboração [62]. A figura 4.8 diz, por exemplo, que se um jato reconstruído tiver seu momento transverso entre 50 e 60 GeV este jato possui seu momento transverso sub-estimado em aproximadamente 5% e portanto precisa ter seu valor multiplicado por 1.05 de forma a se igualar com o momento transverso do jato gerado. A nova escala de energia é denominada EM+JES+GSC+Mu+Pt.

O efeito de cada correção na massa invariante dos di-jatos (mais detalhes na seção 5.6) é mostrado na figura 4.9 [38] para o canal de decaimento  $ZH \rightarrow llb\bar{b}$  com os fatores de correção derivados neste mesmo canal. A figura 4.9 mostra que a resolução da massa invariante dada pela largura do ajuste da distribuição com a função Bukin (seção 5.8) possui os valores 16.4 GeV, 14.4 GeV e 14.1 GeV, respectivamente. Assim, adicionando a correção de múons é obtida uma melhora de 12% na resolução e adicionando ainda a correção de  $p_T$  a resolução possui uma melhora de 14%.

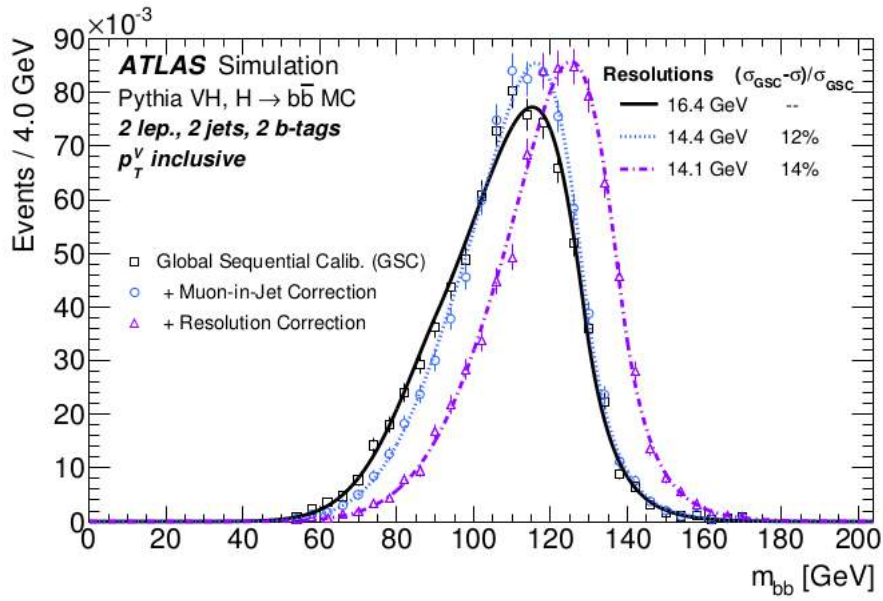


Figura 4.9: Massa invariante reconstruída em diferentes etapas de correção para a análise  $ZH \rightarrow llb\bar{b}$ . Fonte: [38]

O grupo de análise da Coppe em parceria com a Universidade de Glasgow também estudou o efeito das correções de  $p_T$  para a amostra de  $WH \rightarrow l\nu b\bar{b}$  [62], a mesma amostra de eventos utilizada nesta dissertação.

# Capítulo 5

## Método de Análise

Este capítulo discute o método de análise e cada uma de suas técnicas. A seleção de eventos é discutida na seção 5.1, os treinamentos das redes neurais artificiais e suas técnicas são mostrados nas seções 5.2 e 5.3, enquanto as correções aplicadas na energia dos jatos  $b$  são descritas nas seções 5.4 e 5.5. O cálculo da massa invariante dos dois jatos  $b$ ,  $m_{b\bar{b}}$ , é exemplificado na seção 5.6 e os métodos estatísticos usados para analisar a performance da correção neural são explicados nas seções 5.7 e 5.8. Por final é sugerido um método de correção baseado em na combinação de redes neurais 5.9.

### 5.1 Seleção de Eventos

Os dados utilizados foram fornecidos pela colaboração ATLAS e consistem no canal  $pp \rightarrow W + H + X$  com  $WH \rightarrow l\nu b\bar{b}$ , onde o candidato a Higgs (125 GeV) decai em dois jatos  $b$  e o bóson  $W$  decai em um lépton e um neutrino. Todos os dados de entrada são formados de características específicas dos jatos  $b$ . A correção neural é aplicada no topo da correção GSC de forma a aproximar os eventos reconstruídos dos gerados, onde as correções JES+GSC falharam. Além disso, a amostra utilizada para derivar a correção neural passou pela mesma seleção de eventos aprovada pela colaboração e descrita em [38]. Aproximadamente 95 mil eventos passaram pelos cortes de seleção e foram utilizados no treinamento das redes neurais.

### 5.2 Redes Neurais Artificiais

As redes neurais artificiais são modelos computacionais inspirados no sistema nervoso dos seres vivos. São definidas como um conjunto de unidades de processamento (neurônios artificiais) dispostas em uma ou mais camadas e que são interligadas por conexões (sinapses artificiais) que, geralmente, estão associadas a um determinado

tipo de peso, os quais possuem a responsabilidade de acumular o conhecimento adquirido durante o processo de aprendizagem. Os pesos sinápticos funcionam analogamente à intensidade de transmissão do sinal de um neurônio biológico para o seu vizinho mais próximo.

Logo, uma rede neural artificial pode ser vista como uma máquina adaptativa que se assemelha ao cérebro humano em dois aspectos [63]:

1. *O conhecimento é adquirido pela rede a partir de seu ambiente através de um processo de aprendizagem.*
2. *Forças de conexão entre neurônios, conhecidas como pesos sinápticos, são utilizadas para armazenar o conhecimento adquirido.*

O modelo de neurônio artificial mais simples foi proposto por McCulloch & Pitts [64] e ainda hoje é o modelo mais utilizado nas diferentes arquiteturas de redes neurais artificiais. Neste modelo os diversos sinais de entrada  $\{x_1, x_2, \dots, x_n\}$  (dendritos) são apresentados na entrada do neurônio. As ponderações de cada sinal de entrada são dadas pelo conjunto dos pesos  $\{w_1, w_2, \dots, w_n\}$  que exercem o papel análogo ao das junções sinápticas do neurônio biológico, onde a relevância de cada entrada  $\{x_i\}$  é então executada por meio de suas multiplicações com seus respectivos pesos sinápticos  $\{w_i\}$ . Assim, a saída do neurônio artificial, chamada de  $u$ , é a soma ponderada de suas entradas.

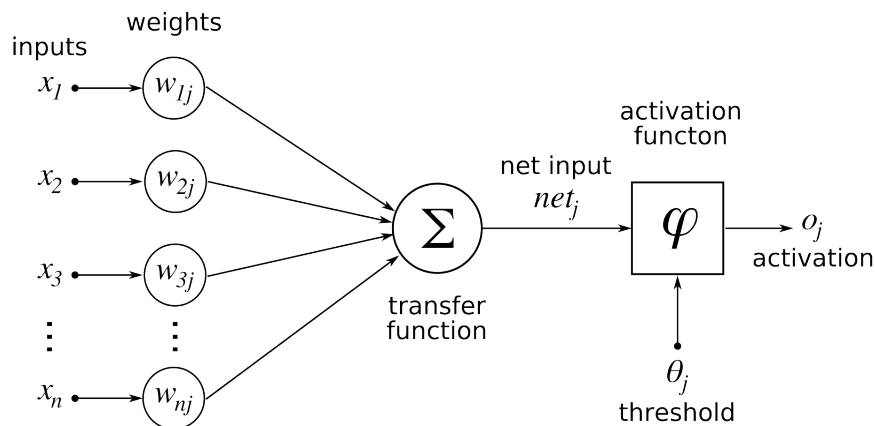


Figura 5.1: Neurônio artificial. Fonte [63]

Considerando a figura 5.1 vemos que o neurônio artificial é composto dos seguintes elementos [65]:

- Sinais de entrada  $\{x_1, x_2, \dots, x_n\}$  - São as características do meio externo e representam valores de variáveis específicas. Suas distribuições são normalizadas com o intuito de aumentar a eficiência dos algoritmos de treinamento.

- Pesos sinápticos  $\{w_1, w_2, \dots, w_n\}$  - São os ponderadores de cada variável de entrada da rede quantificando suas relevâncias em relação ao respectivo neurônio.
- Combinador linear  $\{\Sigma\}$  - Soma dos sinais de entrada com seus respectivos pesos sinápticos constituindo uma combinação linear a fim de produzir um valor de potencial de ativação.
- Limiar de ativação (*bias*)  $\{\theta\}$  - Uma variável que define o patamar apropriado para o disparo do neurônio.
- Potencial de ativação  $\{u\}$  - Diferença entre o valor produzido pelo combinador linear e o bias. Se  $u$  for positivo,  $u \geq \theta$  então o neurônio produz um potencial excitatório, caso contrário o potencial será inibitório.
- Função de ativação  $\{\varphi\}$  - São aplicadas para limitar a amplitude de saída de um neurônio. Neste trabalho usamos a função tangente hiperbólica que sempre assumirá valores reais entre  $[-1, 1]$  e é definida pela equação 5.1.

$$\tanh(v) = \frac{e^v - e^{-v}}{e^v + e^{-v}} \quad (5.1)$$

- Sinal de saída  $\{y\}$  - Valor final produzido por um neurônio em relação ao seus sinais de entrada.

As equações 5.2 e 5.3 resumizam o resultado proposto por McCulloch & Pitts para o neurônio artificial:

$$u = \sum_{i=1}^n w_i \cdot x_i - \theta \quad (5.2)$$

$$y = \varphi(u) \quad (5.3)$$

Uma rede neural do tipo *Perceptron* Multicamadas (*Multilayer Perceptron - MLP*) têm sido aplicadas na solução de diversos tipos de problemas, tais como aproximação de funções, reconhecimento de padrões, otimização e etc. Uma MLP é constituída por uma camada de entrada (*input layer*) e seus bias, uma ou mais camadas escondidas (*hidden layers*) e uma camada de saída (*output layer*) que possui dimensão igual a quantidade de sinais esperados como resposta. A figura 5.2 mostra um exemplo de rede neural com múltiplas camadas. A quantidade de camadas escondidas e de neurônios que cada rede possui, define sua complexidade e não há regra simples para determiná-los, sendo na maioria das vezes escolhidos de forma empírica [63].

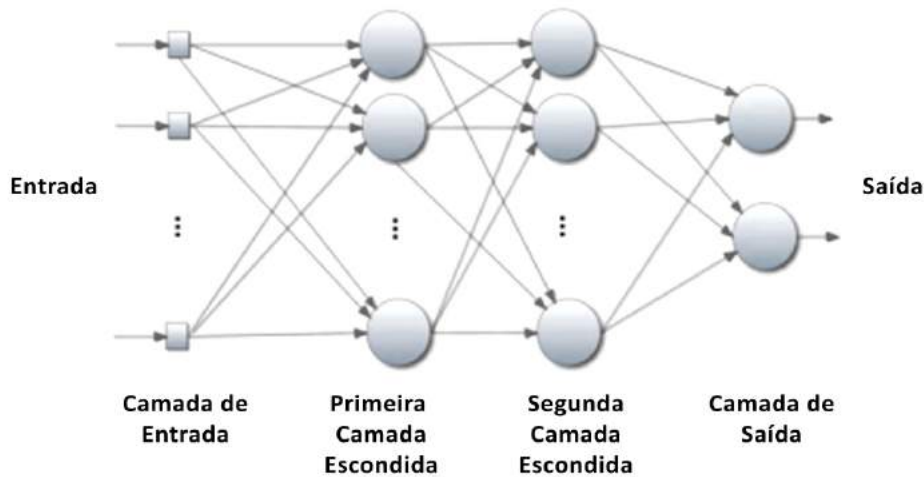


Figura 5.2: Exemplo de rede neural com múltiplas camadas.

O procedimento utilizado para realizar o ajuste dos pesos sinápticos é chamado de algoritmo de aprendizagem, ou algoritmo de treinamento, cuja a função é modificar os pesos sinápticos de forma que o conjunto de saídas da rede neural sejam próximas das saídas desejadas. Historicamente o treinamento da rede MLP é do tipo supervisionado por meio do algoritmo *backpropagation* [66]. O treinamento da rede neural é dito supervisionado quando se tem disponível para cada sinal de entrada, as respectivas saídas desejadas, ou seja, o aprendizado se comporta como se existisse um especialista ensinando para a rede qual a resposta correta para cada amostra apresentada em suas entradas.

O algoritmo de aprendizado *backpropagation* consiste de duas etapas. Na primeira etapa um padrão de entrada é apresentado à rede e seu efeito é propagado por todas as camadas até produzir um conjunto de saídas. Na segunda etapa uma função de erro calcula a diferença entre a saída da rede e o alvo (saída desejada) [65], as derivadas da função erro são calculadas com respeito aos pesos e bias da rede neural. Assim os pesos sinápticos são atualizados conforme o erro é retro-propagado de acordo com a equação 5.4, onde  $F_{erro}$  é a função erro contínua e diferenciável,  $\delta$  é uma constante chamada de passo de aprendizagem e tem seu valor entre  $[0, 1]$ ,  $w_{ji}(t)$  é o peso sináptico da entrada  $i$  relativa ao neurônio  $j$  antes da atualização e  $w_{ji}(t + 1)$  é o peso após a atualização.

$$w_{ji}(t + 1) = w_{ji}(t) + \delta \left( -\frac{\partial F_{erro}}{\partial w_{ji}} \right) \quad (5.4)$$

Como dito anteriormente a função de erro deve ser contínua e diferenciável de forma que a sua derivada em função dos pesos seja possível, usualmente a função escolhida é o Erro Médio Quadrático (*Mean Square Error - MSE*) entre a saída da rede e o seu alvo.

O algoritmo de aprendizado pode ser interrompido quando um determinado

número de épocas (iterações) é atingido ou quando o erro no conjunto de validação (seção 5.3) começar a aumentar enquanto o erro do conjunto de treino permanece o mesmo ou ainda diminui, essa técnica evita a ocorrência de *overfitting*.

Nesta dissertação o algoritmo de aprendizado utilizado é o algoritmo de Levenberg-Marquadt [67], ou simplesmente algoritmo LM, um método de otimização e aceleração da convergência do algoritmo backpropagation [68] e é considerado um método de segunda ordem pois utiliza informações da derivada segunda da função de erro. O algoritmo LM é indicado quando a rede não possui mais do que algumas centenas de pesos sinápticos a serem ajustados, isto deve-se ao fato de que o algoritmo precisa armazenar uma matriz quadrada com dimensão na ordem do número de conexões sinápticas da rede. No algoritmo LM a atualização dos pesos é realizada conforme mostrado na equação 5.5

$$W(t + 1) = W(t) - (J^T J + \mu I)^{-1} J^T F_{erro}, \quad (5.5)$$

onde  $J$  é a matriz Jacobiana com suas derivadas parciais entre a função erro e os sinais de entrada,  $I$  é a matriz identidade,  $W$  a matriz de pesos sinápticos e  $\mu$  um parâmetro que torna a matriz Hessiana definida positiva [68] [69] [70].

Quando  $\mu$  é zero a equação 5.5 se torna o Método de Newton [71] utilizando a aproximação da matriz Hessiana,  $J^T J$ . Se  $\mu$  for grande a equação 5.5 se torna o método de gradiente descendente com um passo de aprendizado pequeno. O Método de Newton é mais rápido e mais preciso quando está próximo de um mínimo da função erro, logo o objetivo é se deslocar para o método de Newton o mais rápido possível. Assim,  $\mu$  é diminuído a cada passo bem sucedido (redução da função erro) e aumentado apenas quando uma tentativa de passo de aprendizagem causa um aumento da função erro. Desta forma a função erro é sempre reduzida a cada iteração do algoritmo.

A tabela 5.1 mostra os parâmetros utilizados para o treinamento das redes neurais. O treinamento é interrompido quando o número máximo de épocas é atingido (100 épocas) ou quando o conjunto de validação possui um erro médio quadrático maior que o conjunto de treinamento durante seis épocas seguidas, também chamado de parada prematura do treinamento.

Tabela 5.1: Parâmetros de treinamento do algoritmo LM.

$\mu$ inicial	0.001
fator de diminuição em $\mu$	0.1
fator de aumento em $\mu$	10.0
$\mu$ máximo	$1.0 \times 10^{10}$
número máximo de épocas	100
falhas no conjunto de validação	6

## 5.3 Validação Cruzada

A validação cruzada é uma técnica para avaliar o modo como os resultados de uma análise estatística irão generalizar um conjunto de dados independente, ou seja, a capacidade de generalização de um modelo a partir de um conjunto de dados [72]. Tipicamente, na validação cruzada, a amostra de dados é dividida aleatoriamente em dois subconjuntos: um de treinamento e outro de teste (validação); estes conjuntos são mutuamente exclusivos. Assim, numa iteração da validação cruzada utilizamos um subconjunto, o de treinamento, para realizar a análise e o treinamento da rede neural e o subconjunto de teste para validar a análise.

Existem diferentes formas de realizar o particionamento dos subconjuntos, para esta dissertação é utilizada a técnica de *k-folds* ou *k-partições*. Nesta técnica a amostra original é particionada aleatoriamente em  $k$  partições de tamanho igual. Uma das partições é reservada para o conjunto de teste e as outras  $k - 1$  partições restantes são utilizadas no treinamento da rede neural. O processo de validação cruzada é repetido  $k$  vezes com cada uma das partições sendo usada exatamente uma única vez como o conjunto de validação. O número de partições define a porcentagem de eventos usados nos subconjuntos de treinamento e teste. Por exemplo, usando dez partições,  $k = 10$ , para cada iteração da validação cruzada temos 90% dos eventos para o treinamento e 10% dos eventos são destinados para validar o modelo. Um exemplo da validação cruzada por *k-partições* é mostrada na figura 5.3.

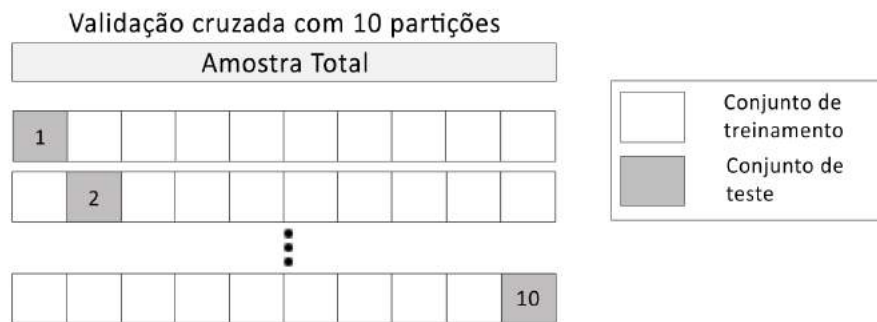


Figura 5.3: Ilustração do processo de validação cruzada por *k-folds*.

Nesta dissertação foram utilizadas exatamente 10 partições para validação cruzada e cada uma das partições foi inicializada 100 vezes. O processo de inicializar inúmeras vezes uma mesma partição é feito no intuito de escaparmos do mínimo local, que consiste num ponto onde a função que queremos minimizar, para o nosso caso o MSE, possui um valor mínimo em relação a sua vizinhança, mas possivelmente este valor é maior do que o valor de um ponto distante no resto do espaço de busca. Ao treinar uma rede neural, os seus pesos sinápticos são inicializados aleatoriamente, logo o algoritmo de treinamento possui diferentes pontos de partida para o aprendizado que dependem de cada inicialização, com isso o algoritmo pode



ficar preso em um mínimo local, pois uma inicialização ruim dos pesos sinápticos pode favorecer tal comportamento.

Contudo, para cada topologia testada foram treinadas 1000 redes neurais, 100 inicializações para cada uma das 10 partições, de modo a escolher a melhor delas. As redes escolhidas foram as que possuíam o menor valor de MSE ou o menor valor de divergência (seção 5.7) nas distribuições de probabilidade entre alvo e saída da rede neural.

## 5.4 Correções Por Jato e Por Evento

Dois tipos diferentes de correções serão derivadas. A primeira consiste em uma correção *por jato*, onde não há distinção de qual jato do evento é usado como sinal de entrada da rede neural. Por exemplo, sabemos que o canal de decaimento utilizado nesta análise é o  $H \rightarrow b\bar{b}$  (seção 2.3), onde o Higgs decai em dois jatos  $b$ , assim na correção por jato a rede neural não sabe qual destes dois jatos está sendo usado como sinal de entrada. Podendo ser o jato mais energético (*leading jet*) ou o menos energético (*sub-leading jet*), uma vez que eles alimentam a rede neural um de cada vez. Assim a rede só obtém informações de um dos jatos e não de todo o evento após o decaimento do Higgs.

Além disso, podemos dividir a correção por jatos em mais uma categoria: se eles possuem ou não múons dentro do seu cone. Como visto na seção 4.4.6 isto ocorre por definição em aproximadamente 11% dos jatos  $b$  e esses múons se encontram acompanhados de um neutrino (figura 4.7) que, devido a sua natureza, escapa completamente da detecção. Com isso, também foram geradas correções levando em consideração se o jato possui ou não múons no seu interior, de forma a recuperar o desbalanço de energia causado pelo neutrino.

Por outro lado, nas correções *por evento* utilizamos as informações dos dois jatos  $b$  como sinais de entrada para a rede neural, ou seja, eles entram na rede neural ao mesmo tempo. Desta forma a rede neural recebe características dos dois jatos presentes num evento de Higgs e não somente de um dos jatos originados no seu decaimento.

Antes de entrarem como *inputs* da rede neural, todas as variáveis de entrada, foram normalizadas de forma que suas distribuições obtivessem média nula e desvio padrão unitário. Este processo é feito para que os sinais de entrada se encontrem dentro da faixa dinâmica da função de ativação da rede (tangente hiperbólica). Esta normalização é obtida com a equação 5.6

$$\hat{X}_i = \frac{X_i - \bar{X}}{\sigma_X} \quad (5.6)$$

onde  $\bar{X}$  é a média da distribuição original e  $\sigma_X$  é seu desvio padrão. A normalização dos alvos da rede neural é feita através do pico de sua distribuição, ou seja, cada evento é dividido pelo valor mais provável (MOP - *Most Probable Value*) da distribuição.

## 5.5 Aplicação da Correção Neural

As variáveis de entrada da rede neural serão dadas por informações intrínsecas dos jatos e suas assinaturas no detector e o alvo de treinamento será o momento transversal dado pela simulação no nível *truth*, ou seja, antes de passar pela reconstrução.

Após o treinamento o fator de correção será:

$$\alpha = \frac{Pt_{NN}}{Pt_{GSC}}, \quad (5.7)$$

onde  $Pt_{NN}$  é a saída da rede neural e  $Pt_{GSC}$  é o momento transversal reconstruído no estágio da correção GSC. Para aplicar a correção neural o fator de escala,  $\alpha$ , deve ser multiplicado pelo momento transversal e pela energia do jato reconstruído, assim obteremos o novo jato com a correção dada pela rede neural.

Assim, para observar o efeito da correção neural na massa invariante do bóson de Higgs, seção 5.6, os novos valores de momento e energia serão:

$$Pt_{corr} = \alpha \times Pt_{GSC} = Pt_{NN} \quad (5.8)$$

e

$$E_{corr} = \alpha \times E_{GSC}. \quad (5.9)$$

O fator de escala  $\alpha$  deve multiplicar tanto  $E$  quanto  $p_T$  para o cálculo da massa invariante devido à relação entre essas duas variáveis.

## 5.6 Cálculo da Massa do Bóson de Higgs

O efeito da correção neural deverá ser calculado na massa invariante do candidato a Higgs, pois como foi dito na seção 1.1 esta é a variável com o maior poder de discriminação entre sinal e background. A massa invariante de uma partícula é uma característica do momento total e energia que é a mesma em todos os referenciais relacionados por transformações de Lorentz. A massa invariante do candidato a Higgs será calculada com a soma dos dois quadri-vetores de Lorentz dos objetos de seu decaimento, no nosso caso os dois jatos do tipo  $b$ . Em outras palavras, como o Bóson de Higgs decai muito rapidamente, calculamos a sua massa utilizando o produto de seu decaimento, as suas partículas filhas, os dois jatos  $b$ . Como se tratam de partículas relativísticas não podemos simplesmente somar a massa dos

dois jatos, devemos lançar mão da cinemática relativística [73].

$$E^2 = (\mathbf{p} \cdot c)^2 + (m \cdot c^2)^2 \quad (5.10)$$

onde  $E$  é a energia total,  $\mathbf{p}$  é o vetor de momento,  $m$  é a massa da partícula mãe (Bóson de Higgs) e  $c$  é a velocidade da luz. Assim a soma das energias e momentos do produto de decaimento será:

$$m^2 = \left( \frac{E_1}{c^2} + \frac{E_2}{c^2} \right)^2 - \left( \frac{\mathbf{p}_1}{c} + \frac{\mathbf{p}_2}{c} \right)^2 \quad (5.11)$$

com  $E_1$  e  $\mathbf{p}_1$  sendo a energia e o momento do primeiro jato e  $E_2$  e  $\mathbf{p}_2$  momento e energia do segundo jato. Desenvolvendo cada parêntesis:

$$m^2 = \frac{E_1^2}{c^4} + \frac{E_2^2}{c^4} + 2 \cdot \frac{E_1}{c^2} \cdot \frac{E_2}{c^2} - \frac{\mathbf{p}_1 \cdot \mathbf{p}_1}{c^2} - \frac{\mathbf{p}_2 \cdot \mathbf{p}_2}{c^2} - 2 \cdot \frac{\mathbf{p}_1 \cdot \mathbf{p}_2}{c^2} \quad (5.12)$$

como  $m_1 = \frac{E_1^2}{c^4} - \frac{\mathbf{p}_1 \cdot \mathbf{p}_1}{c^2}$  e a mesma relação vale para a massa do segundo jato  $m_2$  temos para a expressão final:

$$m^2 = m_1^2 + m_2^2 + 2 \left( \frac{1}{c^4} \cdot E_1 \cdot E_2 - \frac{1}{c^2} \cdot \mathbf{p}_1 \cdot \mathbf{p}_2 \right) \quad (5.13)$$

Como mostrado na equação 5.13 precisamos dos valores de energia e momento dos dois jatos para podermos calcular a massa invariante do bóson de Higgs antes e depois da correção neural. Além disso precisamos de uma figura de mérito que diga se a nova distribuição de massa invariante, ou seja, a distribuição da massa corrigida com o modelo neural proposto, está ou não em melhor posição quando comparada com as correções originais fornecidas pela colaboração ATLAS. Além das divergências, seção 5.7, também utilizamos o ajuste de modelos na distribuição de massa para comparar os resultados como mostrado na seção a seguir.

## 5.7 Divergência de Jensen-Shannon

A divergência de Jensen-Shannon [74] é bastante utilizada, em estatística e teoria probabilística, como um método de medir a similaridade entre duas distribuições de probabilidade e é baseada na divergência de Kulback-Liebler [75], porém é simétrica e possui limite superior igual a um. Uma métrica entre duas distribuições  $p$  e  $q$  é definida a seguir.

Considere as variáveis aleatórias  $X$  e  $Y$  definidas no mesmo espaço amostral  $\Omega = \{\epsilon_1, \epsilon_2, \dots, \epsilon_n\}$ . Cada distribuição é caracterizada por sua distribuição de probabilidade  $p, q : \Omega \rightarrow [0, 1]$ , tal que  $p(\epsilon_i) = Pr(X = \epsilon_i)$ ,  $q(\epsilon_i) = Pr(Y = \epsilon_i)$  e  $\sum p(\epsilon) = \sum q(\epsilon) = 1$ . Para um quantificador ser considerado uma métrica ele deve

seguir as seguintes propriedades:

1.  $\mathcal{D}(p, p) = 0$ , reflexividade;
2.  $\mathcal{D}(p, q) \geq 0$ , não negatividade;
3.  $\mathcal{D}(p, q) = \mathcal{D}(q, p)$ , simetria;
4.  $\mathcal{D}(p, q) \leq \mathcal{D}(p, r) + \mathcal{D}(r, p)$ , desigualdade triangular para qualquer função de probabilidade  $r$  definida no mesmo espaço de probabilidades.

Uma divergência apenas precisa satisfazer as propriedades 1 e 2 [76]. Assim a divergência de Jensen-Shannon ( $\mathcal{D}_{JS}$ ) é definida como:

$$\mathcal{D}_{JS}(p, q) = \frac{1}{2} [\mathcal{D}_{KL}(p, m) + \mathcal{D}_{KL}(q, m)] \quad (5.14)$$

onde  $m = \frac{1}{2}(p + q)$  e  $\mathcal{D}_{KL}$  é a divergência de Kullback-Liebler

$$\mathcal{D}_{KL}(p, q) = \sum_{\epsilon \in \Omega} p(\epsilon) \log \frac{p(\epsilon)}{q(\epsilon)} \quad (5.15)$$

A divergência de Kullback-Liebler não possui limites para o seu valor máximo, ou seja,  $0 \leq \mathcal{D}_{KL} \leq \infty$ , além de não atender a propriedade de simetria. Contudo foi escolhido trabalhar com a divergência de Jensen-Shannon justamente por ela possuir a característica de ser simétrica,  $\mathcal{D}_{JS}(p, q) = \mathcal{D}_{JS}(q, p)$ , sendo assim mais natural quantificar o quão distante estatisticamente uma distribuição está da outra. Adicionalmente, a divergência JS também pode ser apresentada em função da entropia de Shannon  $\mathcal{S}(p)$  [77] :

$$\mathcal{D}_{JS}(p, q) = \mathcal{S} \left( \frac{p + q}{2} \right) - \frac{1}{2} \mathcal{S}(p) - \frac{1}{2} \mathcal{S}(q) \quad (5.16)$$

em que

$$\mathcal{S}(p) = - \sum_{\epsilon \in \Omega} p(\epsilon) \log p(\epsilon). \quad (5.17)$$

A entropia é uma medida de incerteza associada a uma variável aleatória, ou uma função que indica a perda de informação de um sistema. As definições para a divergência de Jensen-Shannon dadas pelas equações (5.14) e (5.16) são equivalentes.

Neste trabalho as divergências foram usadas para medir a similaridade entre as distribuições de massa invariante, assim podemos avaliar o quão similar a distribuição de massa corrigida está em relação a massa esperada pelo processo físico.

## 5.8 Modelo de Ajuste

De posse do histograma de massa invariante do Bóson de Higgs, devemos agora ajustar o histograma à uma distribuição de probabilidade que melhor represente o processo. Idealmente a distribuição da massa invariante seguiria uma função gaussiana, assim o primeiro parâmetro ajustado pela curva,  $\mu$ , seria ao mesmo tempo a média e o pico e o segundo parâmetro da função gaussiana,  $\sigma$ , seria a resolução da distribuição. Porém, o histograma de massa possui caudas assimétricas e essas assimetrias também possuem efeito no centro da distribuição. Logo uma nova função de probabilidade deve ser proposta para o ajuste do histograma.

A função proposta por A. Bukin em 2008 [78] foi desenvolvida justamente para o problema de modelar uma distribuição com um único máximo e caudas assimétricas. Convolvendo uma função gaussiana com funções exponenciais. A função de Bukin é dada por:

$$P(x; x_p, \sigma_p, \xi, \rho) = A_p \exp \left[ \frac{\xi \sqrt{\xi^2 + 1} (x - x_1) \sqrt{2 \ln 2}}{\sigma_p (\sqrt{\xi^2 + 1} - \xi)^2 \ln(\sqrt{\xi^2 + 1} + \xi)} + \rho \left( \frac{x - x_i}{x_p - x_i} \right)^2 - \ln 2 \right], \quad (5.18)$$

onde  $\rho = \rho_1$  e  $x_i = x_1$  para  $x < x_1$  e  $\rho = \rho_2$  e  $x_i = x_2$  para  $x \geq x_2$  e

$$x_{1,2} = x_p + \sigma_p \sqrt{2 \ln 2} \left( \frac{\xi}{\sqrt{\xi^2 + 1} \pm 1} \right). \quad (5.19)$$

O parâmetro  $x_p$  é a posição do pico,  $\sigma_p$  será a resolução, definida como a largura à meia altura dividida por 2.35 (FWHM/2.35), e  $\xi$  é o parâmetro de assimetria. A figura 5.4 mostra o formato da função de Bukin para  $\xi = -0.2$ ,  $x_p = 0.5$ ,  $\sigma_p = 0.1$ ,  $\rho_1 = 0.1$  e  $\rho_2 = 0.2$ .

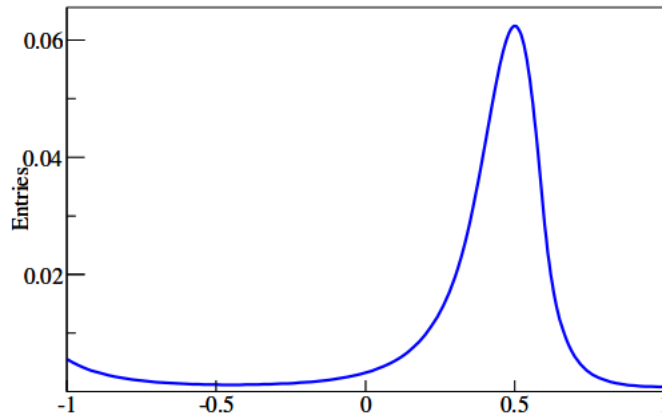


Figura 5.4: Um exemplo do formato da função densidade de probabilidade de Bukin.

Após o ajuste da distribuição de massa com a função Bukin também é necessário testar a aderência do modelo à distribuição, estes testes também são conhecidos como testes de *goodness-of-fit*. O teste do  $\chi^2$  é o escolhido, o teste trata de uma soma normalizada dos erros quadráticos entre o valor observado e o valor teórico como mostrado na equação 5.20

$$\chi^2 = \sum_{i=1}^{N_{bins}} \frac{(O_i - E_i)^2}{E_i} \quad (5.20)$$

onde  $O_i$  é o valor de cada célula do histograma, também chamadas de *bins*, e  $E_i$  é o valor da função calculada no ponto que corresponde ao centro de cada *bin* do histograma.

Geralmente o  $\chi^2/ndf$ , também chamado de  $\chi^2$  normalizado, é utilizado para os testes de aderência de modelos, onde *ndf* é o número de graus de liberdade (*number of degrees of freedom*) que é calculado como a quantidade de células, bins, presentes no intervalo da função de ajuste menos o número de parâmetros livres da função. Com o  $\chi^2/ndf$  temos acesso ao grau de confiabilidade do modelo. A figura 5.5 mostra o grau de confiabilidade para diversos valores de  $\chi^2/ndf$ .

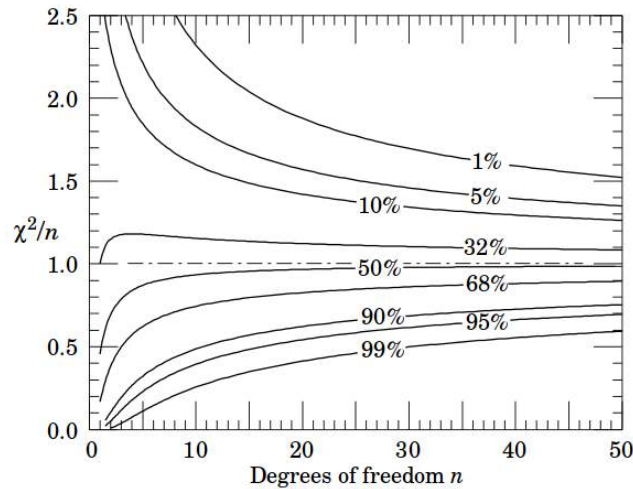


Figura 5.5: O  $\chi^2$  normalizado para diferentes graus de liberdade. Cada curva corresponde a um grau de confiabilidade do modelo ajustado. Fonte: [79]

## 5.9 Conjunto de Redes Neurais (*Boosting*)

Métodos de *Ensemble* são técnicas amplamente utilizadas em aprendizado de máquina e em inteligência artificial de forma a melhorar a acurácia de um determinado problema através da combinação de inúmeros algoritmos [80]. Isto é feito construindo-se um conjunto de algoritmos de base e combinando-os de maneiras diferentes [81]. A razão de utilizar técnicas de *ensemble* se deve ao fato de ser difícil

encontrar um único algoritmo que possua uma performance ótima para um determinado problema, então se faz uma combinação de algoritmos sub-ótimos também chamados de *weak learners* pela literatura [82].

Dentro dos métodos de *ensemble* se encontram as técnicas de *Boosting* [83]. A principal ideia dos algoritmos de *Boosting* é utilizar versões ponderadas da mesma amostra de treinamento para cada algoritmo base, onde as amostras que foram estimadas erradamente, de acordo com uma figura de mérito estabelecida, recebem um peso maior para o próximo algoritmo base. Assim o próximo algoritmo base da *ensemble* recebe mais informação sobre um determinado sinal de entrada. Os pesos para cada sinal de entrada podem ser interpretados como uma re-amostragem, ou seja, uma nova amostra de treinamento é construída de forma que as instâncias com maior peso ocorrem com mais frequência [84]. Em outras palavras, a cada iteração do *boosting* o algoritmo base recebe com mais frequência os eventos que foram estimados imprecisamente pelo algoritmo base anterior, fazendo com que em cada iteração o algoritmo base trabalhe somente com as instâncias mais complexas.

A figura de mérito escolhida para o *boosting* é mostrada na equação 5.21, onde  $y$  é a saída da rede neural,  $t$  é o valor do alvo da rede neural e  $\sigma$  é o desvio padrão da distribuição  $y - t$ . Assim os eventos que obtiverem o valor de  $Err$  acima de um determinado limiar serão classificados como uma estimacão incorreta e terão seus pesos aumentados para a próxima re-amostragem do algoritmo. A introdução de um limiar para avaliar o que é uma estimacão consistente ou uma estimacão imprecisa é feita de forma a transformar o problema de estimacão em um problema de classificacão binária.

$$Err = \frac{|y - t|}{\sigma} \quad (5.21)$$

Para aplicar a técnica de *boosting* nas redes neurais utilizamos o algoritmo Ada-Boost.MRT descrito em detalhes em [85]. O valor para o limiar foi escolhido de tal forma que os eventos estimados incorretamente pela rede neural consistem em aproximadamente 30% do total da amostra, assim ao final do treinamento todas as instâncias que estiverem acima deste limiar,  $Err$ , serão classificadas como uma estimacão imprecisa e terão um peso maior na re-amostragem da próxima iteração do algoritmo. Como não há um critério de parada específico, deixamos o algoritmo executar 5 iterações seguidas e para o resultado final é calculada a média aritmética das saídas das redes neurais de cada uma dessas iterações.

# Capítulo 6

## Resultados Obtidos

Este capítulo mostra os resultados obtidos para cada modelo de correção neural e analisa seus impactos na distribuição de massa invariante do bóson de Higgs. Como abordado na seção 5.4 as correções estudadas neste trabalho podem ser estimadas *por jato*, com uma sub-divisão a mais quando o jato contém ou não um múon dentro do seu cone, ou correções estimadas *por evento*. Além disso, também foi verificado o resultado da estimação da correção quando trabalhamos apenas com o momento transversal dos jatos como sinal de entrada da rede neural ou quando é utilizado um descritor mais completo dos jatos. Ao final do capítulo, seção 6.4, é feita uma comparação dos resultados obtidos com cada corretor neural testado.

### 6.1 Correções Por Jato

Na seção 5.5 foi discutido como um fator de correção, derivado por uma rede neural, pode ser aplicado à energia, e momento transversal, dos candidatos a jatos do decaimento do Higgs, de forma a obter uma distribuição de massa invariante que se aproxime da distribuição de massa prevista pelo Modelo Padrão. Com isso uma rede neural do tipo MLP (*Multi Layer Perceptron*), com uma única camada escondida, foi utilizada para estimar esse fator de correção. O Teorema da Aproximação Universal (*Universal Approximation Theorem*) [86], diz que qualquer função contínua pode ser aproximada por uma rede neural com uma única camada escondida e funções de ativação sigmoide e linear em suas camadas intermediária e de saída, respectivamente. Esta seção trata das correções neurais estimadas *por jato*.

#### 6.1.1 Correção Utilizando o Momento Transverso dos Jatos

Para esta correção utilizamos uma única entrada na rede neural, seu momento transversal no estágio da correção *Global Sequential Calibration* ( $p_T$  @GSC). O alvo da rede será o momento transversal previsto pelo Modelo Padrão e simulado através de



métodos de Monte Carlo ( $p_T$  @TRUTH). Assim a rede deve fazer o mapeamento não linear que leva o momento transverso reconstruído pelo detector ao momento transverso previsto pelo Modelo Padrão, desta forma podemos calcular o fator multiplicativo da correção (seção 5.5) e checar se houve alguma melhoria na distribuição de massa invariante, uma vez que esta é a variável de maior poder de discriminação entre sinal e ruído (eventos do Higgs e eventos de background).

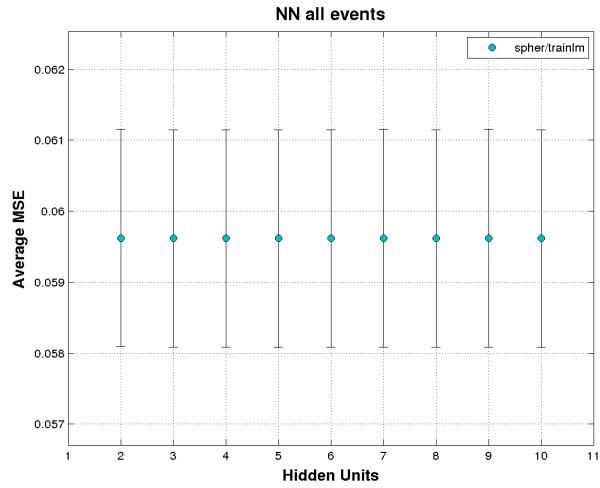
Foram testados diversos números de neurônios na camada escondida de forma a escolher uma topologia de rede que melhor generalize os resultados. Cada uma dessas topologias passou pelo mesmo processo de validação cruzada descrito na seção 5.3 e a figura de mérito para a escolha da melhor rede foi o Erro Médio Quadrático (MSE).

A figura 6.1 mostra o resultado para o estudo do número de neurônios na camada escondida de cada rede neural. A média do MSE de cada topologia é calculada entre os resultados das partições da validação cruzada ( $K$ -folds) e a barra de erro é o desvio padrão entre os valores dessas partições. Vemos que não há diferença na escolha do número de neurônios para as três redes neurais, pois todas as barras de incerteza das topologias testadas coincidem entre si. O princípio de parcimônia foi aplicado, logo para o cálculo da massa invariante escolhemos as redes com apenas dois neurônios na camada escondida, tanto para a correção de todos os jatos em conjunto, quanto para a correção dos jatos com ou sem múon separadamente.

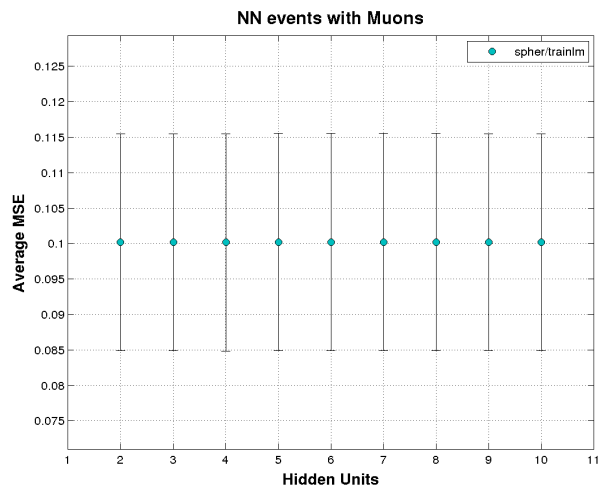
Com um modelo neural escolhido para a correção de energia podemos agora aplicar a correção multiplicativa nos jatos  $b$  e analisar seu impacto na massa invariante do bóson de Higgs conforme exposto nas seções 5.5 e 5.6. A figura 6.2 mostra o resultado da correção neural na massa do bóson de Higgs quando utilizamos uma única rede neural para todos os jatos, histograma na cor verde, a correção neural quando separamos em duas redes os jatos com ou sem múons em seu interior, histograma na cor azul, junto com a correção oficial da colaboração do experimento (*Global Sequential Calibration* - GSC), histograma vermelho, e também o *benchmark* de toda a análise que é a distribuição gerada (*Truth*) pela simulação de Monte Carlo, histograma na cor preta. A binagem escolhida para os histogramas de massa invariante é de 5 GeV para cada bin, devido à resolução finita do detector.

Para medir o impacto da correção neural utilizaremos a média do histograma de massa como medida de escala e o seu desvio padrão como a resolução. Uma correção bem sucedida deve ser aquela que aproxima a escala da distribuição de massa reconstruída em direção da escala esperada pelo Modelo Padrão, representada pela simulação de Monte Carlo no nível @TRUTH, e diminui o máximo possível a resolução do histograma.

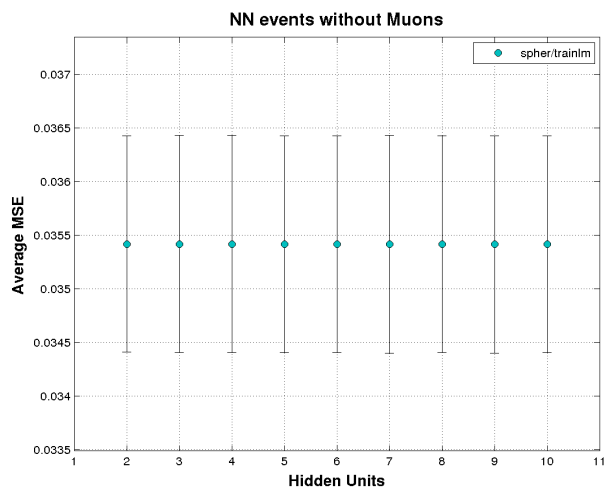
Pode ser visto que a correção neural tem performance melhor para os parâmetros de escala e resolução como mostrados nas caixas à direita na figura 6.2. Melhores no



(a)



(b)



(c)

Figura 6.1: MSE médio para cada topologia das redes com todos os eventos (a), com múons (b) e sem múons (c) das correções por jato utilizando o momento transversal.

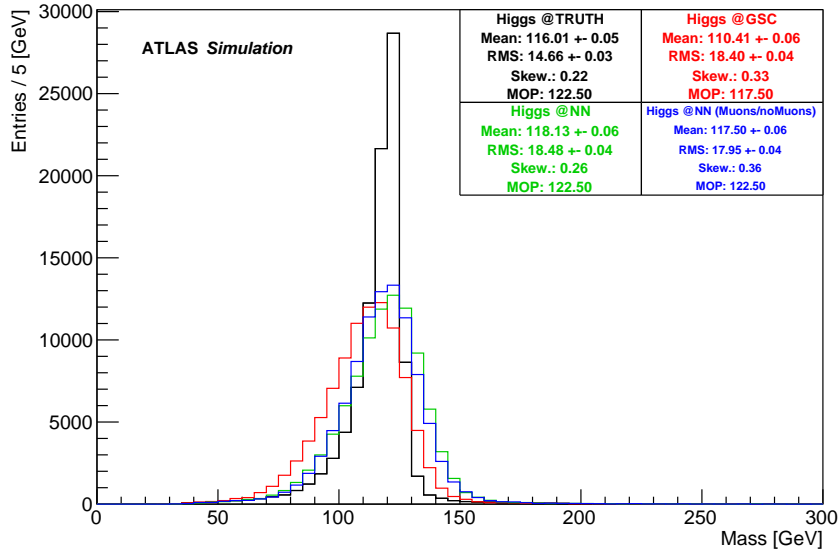


Figura 6.2: Distribuições de  $m_{b\bar{b}}$ , simulada e corrigida pelo corretor neural *por jato* utilizando o momento transverso.

sentindo de se aproximarem mais dos parâmetros da distribuição verdadeira quando comparados com os parâmetros da correção oficial da colaboração. Por exemplo, a média da distribuição @GCS possui valor  $110.41 \pm 0.06$  GeV e a média da distribuição verdadeira, de acordo com o Modelo Padrão, deve possuir uma média de  $116.01 \pm 0.05$  GeV. Isso significa que a correção @GCS subestima a média esperada em 5%, por outro lado a média da distribuição da correção neural proposta (@NN Muons/noMuons),  $117.50 \pm 0.06$  GeV, superestima a massa da distribuição verdadeira em aproximadamente 1.3% estando assim mais próxima do valor esperado. O mesmo comportamento ocorre para o valor do RMS, que diminui de  $18.40 \pm 0.04$  GeV @GSC para  $17.95 \pm 0.04$  GeV com a correção neural. Isso implica que a correção neural consegue aproximar a escala da distribuição de massa e também reduz, um pouco, o valor de sua resolução. Um outro fator importante é que a correção neural acerta perfeitamente o valor do pico (*Most Probable Value* - MOP) de massa em 122.50 GeV, uma melhoria significativa em comparação ao valor mais provável da distribuição de massa da correção @GSC, 117.50 GeV. Os valores de cada histograma são mostrados na tabela 6.1.

Uma outra forma de comparar o impacto das correções é através dos parâmetros da função densidade de probabilidade que melhor adere aos dados. A função de Bukin, seção 5.8, é a que melhor se encaixa neste papel [87]. Assim, os dois primeiros parâmetros da função Bukin representarão a escala e a resolução da distribuição de massa, pico e largura à meia altura dividida por 2.35, respectivamente. Neste trabalho o *fitting* da função Bukin aos dados é feito dentro de três intervalos distintos, são eles: 20–200 GeV, 70–150 GeV e 85–145 GeV. O primeiro intervalo pega toda

Tabela 6.1: Estimativas para o histograma de massa invariante simulada e corrigida pelo corretor neural *por jato* utilizando o momento transverso.

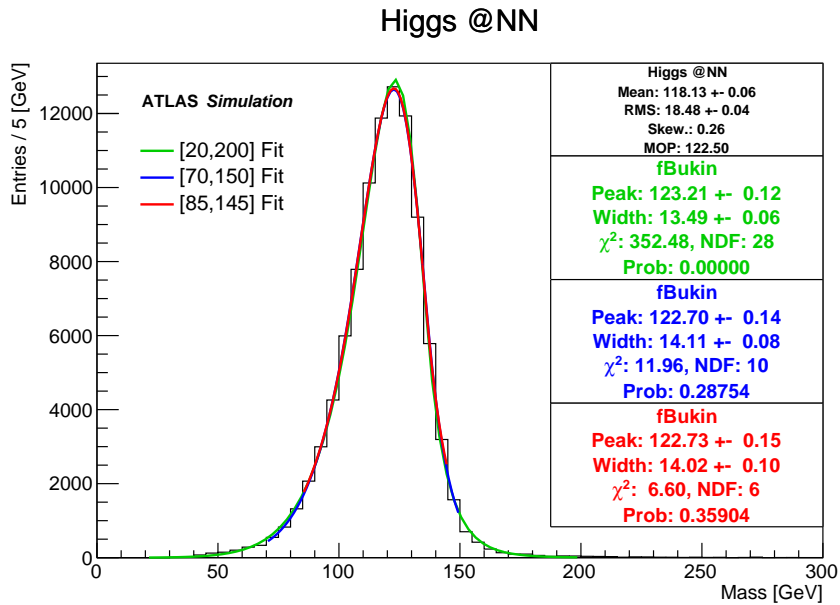
	@TRUTH	@GSC	@NN	@NN (Muons/noMuons)
Mean [GeV]	$116.01 \pm 0.05$	$110.41 \pm 0.06$	$118.13 \pm 0.06$	$117.50 \pm 0.06$
RMS [GeV]	$14.66 \pm 0.03$	$18.40 \pm 0.04$	$18.48 \pm 0.04$	$17.95 \pm 0.04$
Skew	0.22	0.33	0.26	0.36
MOP [GeV]	122.50	117.50	122.50	122.50

a faixa dinâmica da massa invariante do Higgs e os outros intervalos se concentram basicamente no centro (*core*) da distribuição de massa, e foram escolhidos desta forma porque quando há melhora na estimação de energia, este efeito é visto em maior parte no *core* dessa distribuição. Apesar dos intervalos apresentarem valores diferentes para o ajuste, os seus parâmetros de escala e resolução devem exibir valores semelhantes.

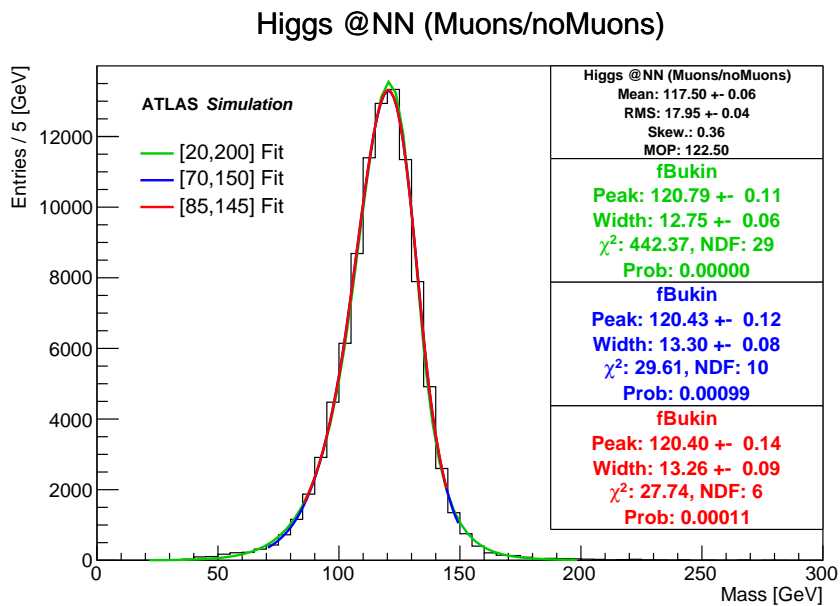
As figuras 6.3a e 6.3b mostram o resultado do ajuste da função Bukin para as duas correções neurais. Os valores dos parâmetros se mostram consistentes com o que foi discutido no parágrafo anterior, o pico e a largura da função permanecem próximos para os sub-intervalos de ajuste. Junto com os parâmetros do *fitting* também são mostrados o valor do  $\chi^2$ , o número de graus de liberdade e o grau de confiabilidade do modelo, como apresentados no final da seção 5.8. Para a comparação, a função Bukin também foi ajustada na distribuição de massa para os níveis @TRUTH e @GCS como mostrado na figuras 6.4a e 6.4b. Os resultados de cada ajuste são sumarizados na tabela 6.2.

Pode ser visto que as duas correções neurais propostas, @NN e @NN (Muons/noMuons), apresentam uma performance melhor que a correção @GSC. A rede neural treinada sem distinção de jatos aproxima a escala, *peak*, e a resolução, *width*, em direção à distribuição verdadeira em todos os intervalos de ajuste escolhidos. Já a distribuição da correção neural levando em consideração os múons dentro dos jatos tem uma resolução menor que a correção @NN para todos os intervalos, porém sua escala está mais distante da escala esperada pelo Modelo Padrão. Por exemplo, para o intervalo de ajuste de 70 – 150 GeV a escala da correção @NN (Muons/noMuons) subestima o valor esperado em 1.2% enquanto a correção @NN superestima o valor da escala em apenas 0.6%. Este comportamento pode ser explicado pelo *Bias-Variance Dillema* [63], que diz que não é possível melhorar simultaneamente a escala e a resolução da distribuição devido à natureza limitada da estatística do processo. Assim se a resolução possui uma melhora grande, o impacto na escala será em menor proporção.

Além dos parâmetros do ajuste devemos comparar a probabilidade dos dados terem sido gerados de uma função de densidade de probabilidade do tipo Bukin. A

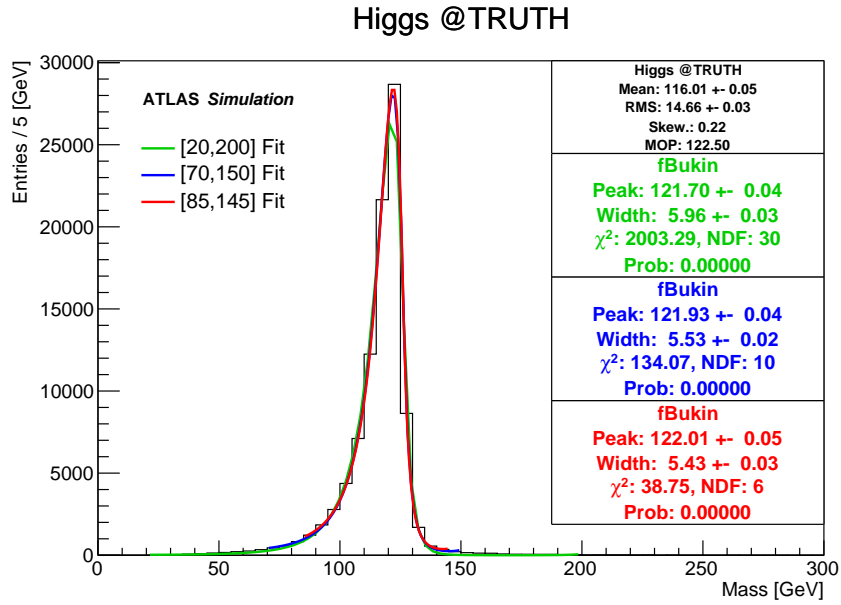


(a)

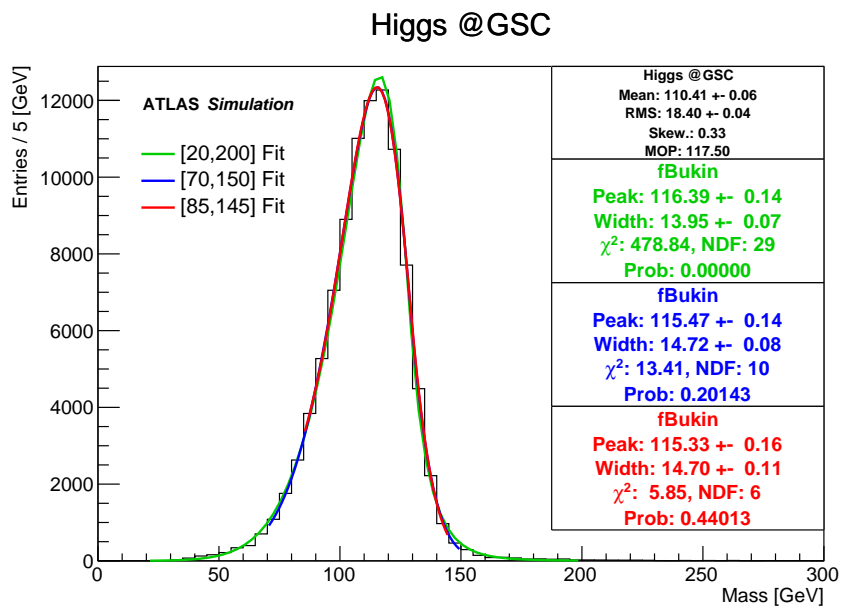


(b)

Figura 6.3: Ajuste da função Bukin para a distribuição de  $m_{b\bar{b}}$  para eventos corrigidos com a rede neural para todos os jatos (@NN) e corrigidos com uma rede para os jatos com múon e outra para os jatos sem múon (@NN Muons/noMuons). Correções *por jato* utilizando o momento transversal.



(a)



(b)

Figura 6.4: Ajuste da função Bukin para a distribuição de  $m_{b\bar{b}}$  para eventos gerados (@TRUTH) e reconstruídos com a correção oficial da colaboração (@GSC).

tabela 6.2 diz, para todos os intervalos de ajuste, que a distribuição @TRUTH tem 0% de probabilidade de ter sido originada pela função de densidade de probabilidade de Bukin, assim a correção neural também deve se afastar do modelo Bukin. Por exemplo, para o terceiro intervalo de ajuste, 85 – 145 GeV, a correção @GSC possui 44% de grau de confiabilidade no modelo, a correção neural @NN possui 36% e se afasta mais do modelo Bukin como previsto pelo ajuste dos dados @TRUTH. A correção neural @NN (Muons/noMuons) se afasta mais ainda do modelo de Bukin e atinge outra vez 0% de probabilidade, sendo assim a melhor opção de correção neural.

Tabela 6.2: Ajuste da função Bukin para diferentes intervalos de  $m_{b\bar{b}}$ . Correções *por jato* utilizando o momento transversal.

[20 – 200] GeV				
	@TRUTH	@GSC	@NN	@NN (Muons/noMuons)
Peak [GeV]	$121.70 \pm 0.04$	$116.39 \pm 0.14$	$123.21 \pm 0.12$	$120.79 \pm 0.11$
Width [GeV]	$5.96 \pm 0.03$	$13.95 \pm 0.07$	$13.49 \pm 0.06$	$12.75 \pm 0.06$
$\chi^2/NDF$	66.78	15.51	12.59	15.25
Prob.	0.00	0.00	0.00	0.00
[70 – 150] GeV				
	@TRUTH	@GSC	@NN	@NN (Muons/noMuons)
Peak [GeV]	$121.93 \pm 0.04$	$115.47 \pm 0.14$	$122.70 \pm 0.14$	$120.43 \pm 0.12$
Width [GeV]	$5.53 \pm 0.02$	$14.72 \pm 0.08$	$14.11 \pm 0.08$	$13.30 \pm 0.08$
$\chi^2/NDF$	13.41	1.34	1.20	2.96
Prob.	0.00	0.20	0.29	0.00
[85 – 145] GeV				
	@TRUTH	@GSC	@NN	@NN (Muons/noMuons)
Peak [GeV]	$122.01 \pm 0.05$	$115.33 \pm 0.16$	$122.73 \pm 0.15$	$120.40 \pm 0.14$
Width [GeV]	$5.43 \pm 0.03$	$14.70 \pm 0.11$	$14.02 \pm 0.10$	$13.26 \pm 0.09$
$\chi^2/NDF$	6.46	0.98	1.1	4.62
Prob.	0.00	0.44	0.36	0.00

Também podemos medir o impacto das correções através da similaridade entre a distribuição de probabilidade da massa gerada e das massas reconstruídas, com as correções GSC e Neural, utilizando a divergência de Jensen-Shannon descrita em detalhes na seção 5.7. A tabela 6.3 mostra os valores de divergência da correção neural em relação às funções de ajuste e aos outros estágios da distribuição de massa. Bukin 1 se refere ao primeiro intervalo de ajuste 20 – 200 GeV, Bukin 2 ao segundo intervalo 70 – 150 GeV e Bukin 3 ao intervalo 85 – 145 GeV. Pode ser visto que a divergência entre a rede @NN (Muons/noMuons) e a distribuição @TRUTH,  $\mathcal{D}(NN\ Mu|Truth) = 5.64 \times 10^{-2}$ , é menor que a divergência entre a correção @GSC

e a distribuição @TRUTH,  $\mathcal{D}(GSC|Truth) = 6.18 \times 10^{-2}$ . Assim a divergência de JS também confirma a maior similaridade entre da distribuição de probabilidade da correção neural proposta em comparação com a correção da colaboração ATLAS.

### Efeito no Background

O papel principal da correção na energia dos jatos é aproximar o máximo possível a distribuição da massa do bóson de Higgs reconstruída pelo detector em direção da massa do bóson de Higgs prevista pelo Modelo Padrão, ou seja, ambas as distribuições devem apresentar o mesmo comportamento. Após a correção ser derivada em simulações de Monte Carlo, ela deve ser aplicada aos dados reais detectados pelo ATLAS. Dados reais possuem tanto os jatos  $b$  provenientes do decaimento do bóson de Higgs quanto os jatos  $b$  de outros processos. Eventos com jatos  $b$  que não são oriundos do decaimento do Higgs são chamadas de eventos de *background*, por exemplo os jatos  $b$  resultantes do decaimento de um bóson  $Z$ . Assim o efeito da correção dos jatos também deve ser testado nos eventos de background para verificar se a correção de energia não está produzindo artificialmente eventos de Higgs nestas amostras.

A figura 6.5 mostra os efeitos de todas as correções de energia para a amostra de di-bósons  $WZ$ , onde seus jatos podem ser confundidos com jatos do decaimento do Higgs. As distribuições Higgs @Mu e Higgs @Pt são as correções da colaboração ATLAS específicas para os jatos  $b$  como discutidas nas seções 4.4.6 e 4.4.7. Vemos que há um deslocamento na distribuição de massa quando as correções são aplicadas, porém não há um pico na região de massa do Higgs, em torno de 120 GeV. Um resultado favorável, pois isto significa que as redes neurais não estão corrigindo os jatos  $b$  do bóson de Higgs e ao mesmo tempo criando artificialmente eventos de sinal nas amostras de background.

### 6.1.2 Correção Utilizando um Descritor dos Jatos

Aqui serão mostrados os resultados das correções neurais baseadas nas características internas dos jatos, dadas por um conjunto de variáveis descritas a seguir:

1.  $p_T$  - Momento transverso.
2.  $\eta$  - Pseudo-rapidez.
3.  $\phi$  - Ângulo azimutal.
4.  $E$  - Energia total.
5.  $FracEM3$  - Fração da energia depositada no terceiro layer do calorímetro eletromagnético.



Tabela 6.3: Comparação da divergência JS para a distribuição de massa corrigida com a correção neural *por jato* utilizando o momento transverso.

	Bukin 1	Bukin 2	Bukin 3	TRUTH	GSC	NN	NN (Muons/noMuons)
TRUTH	$4.24 \times 10^{-3}$	$2.55 \times 10^{-4}$	$3.20 \times 10^{-4}$	0.0	$6.18 \times 10^{-2}$	$6.90 \times 10^{-2}$	$5.64 \times 10^{-2}$
GSC	$7.13 \times 10^{-4}$	$1.55 \times 10^{-4}$	$1.19 \times 10^{-3}$	$6.18 \times 10^{-2}$	0.0	$2.90 \times 10^{-2}$	$2.36 \times 10^{-2}$
NN	$5.34 \times 10^{-4}$	$1.70 \times 10^{-4}$	$8.31 \times 10^{-4}$	$6.90 \times 10^{-2}$	$2.90 \times 10^{-2}$	0.0	$1.45 \times 10^{-3}$
NN (Muons/noMuons)	$6.85 \times 10^{-4}$	$1.95 \times 10^{-4}$	$7.36 \times 10^{-4}$	$5.64 \times 10^{-2}$	$2.36 \times 10^{-2}$	$1.45 \times 10^{-3}$	0.0

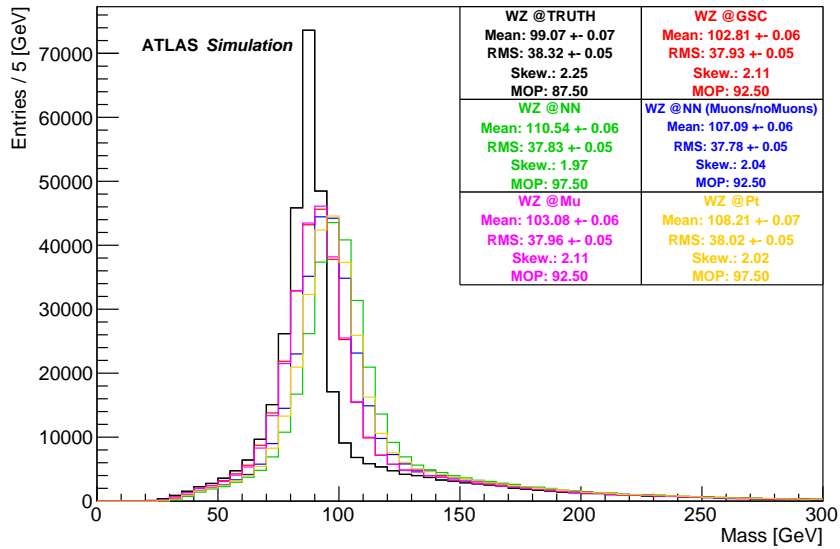
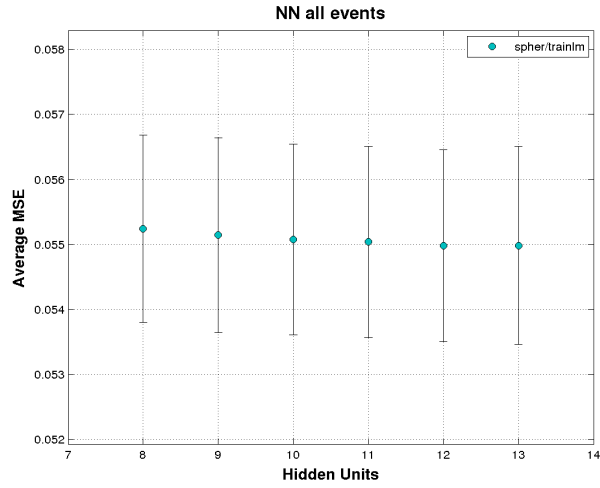


Figura 6.5: Distribuições de  $m_{b\bar{b}}$ , simulada e corrigida pelo corretor neural *por jato* utilizando o momento transverso e aplicadas na amostra de background  $WZ$ .

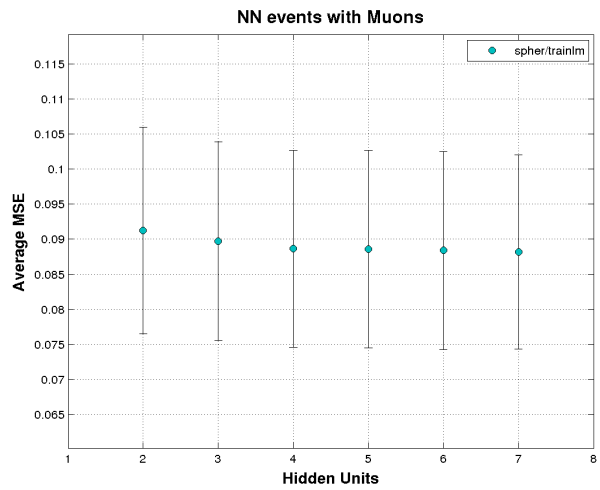
6. *FracTile0* - Fração de energia depositada no primeiro layer do *Tile Calorimeter*.
7. *TrkWidth* - Largura do jato medida de acordo com seus traços.
8. *EMF* - Fração de energia eletromagnética do jato.
9. *JVF* - Fração de vértices do jato.
10. *NTrk* - Número de traços associados ao jato.
11. *SumPtTrk* - Soma dos momentos transversos associados ao jato que foram detectados no *Inner Detector*.

Esta seção dá continuidade às correções por jato, só que desta vez utilizando características internas dos jatos como variáveis de entrada da rede neural. Assim teremos 11 sinais de entrada na rede neural, descritos anteriormente. A figura 6.6 mostra o estudo de topologia para as redes neurais treinadas com todos os eventos e as redes treinadas separadamente para os eventos com e sem múons no interior dos jatos. A topologia escolhida para a rede neural que corrige todos os jatos foi a de 10 neurônios e escolhemos 4 e 8 neurônios para as redes neurais dos jatos com e sem múons, respectivamente.

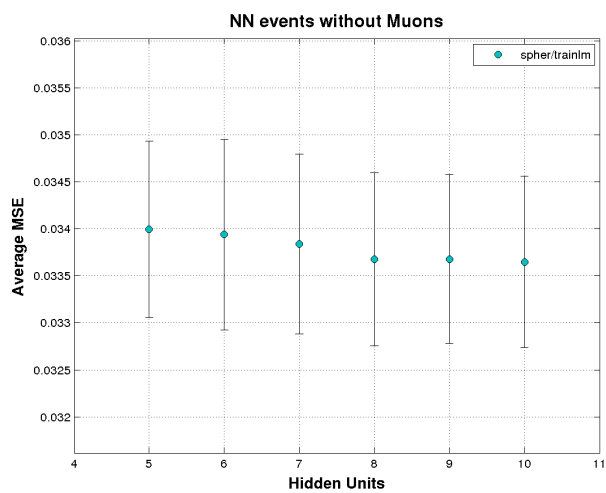
O efeito da correção neural com o uso de um descritor de jatos no histograma de massa é visto na figura 6.7 e sumarizado na tabela 6.4 onde vemos que novamente as redes neurais atingem uma performance melhor tanto nos parâmetros de escala e resolução como no valor do pico do histograma.



(a)



(b)



(c)

Figura 6.6: MSE médio para cada topologia das redes com todos os eventos (a), com múons (b) e sem múons (c) das correções por jato utilizando um descritor de jatos.

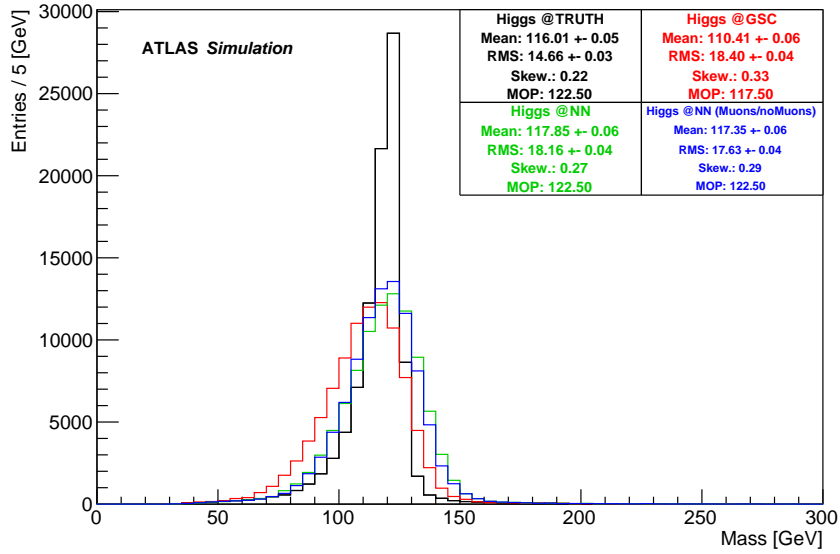


Figura 6.7: Distribuições de  $m_{b\bar{b}}$ , simulada e corrigida pelo corretor neural *por jato* utilizando um descritor de jatos.

Tabela 6.4: Estimativas para o histograma de massa invariante simulada e corrigida pelo corretor neural *por jato* utilizando um descritor de jatos.

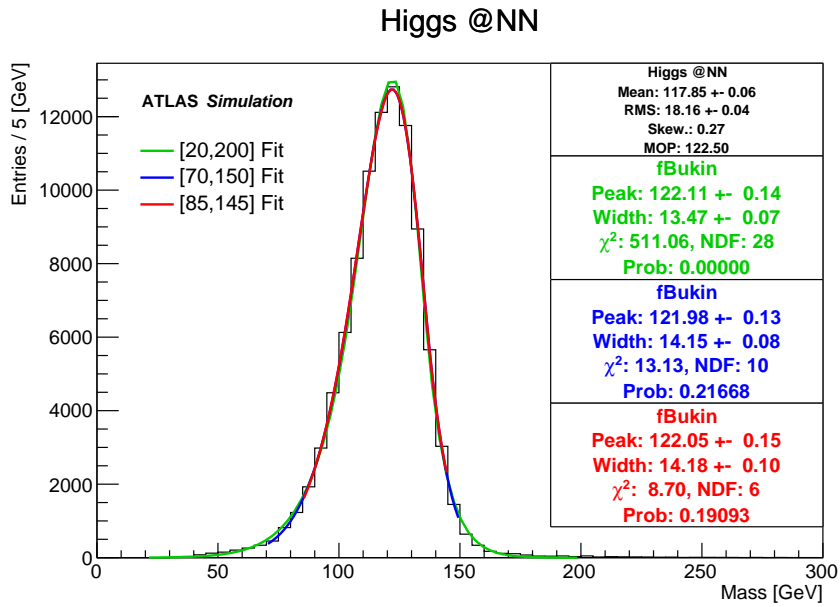
	@TRUTH	@GSC	@NN	@NN (Muons/noMuons)
Mean [GeV]	116.01 ± 0.05	110.41 ± 0.06	117.85 ± 0.06	117.35 ± 0.06
RMS [GeV]	14.66 ± 0.03	18.40 ± 0.04	18.16 ± 0.04	17.63 ± 0.04
Skew	0.22	0.33	0.27	0.29
MOP [GeV]	122.50	117.50	122.50	122.50

O ajuste da função Bukin é exibido nas figuras 6.8a e 6.8b e seus resultados são sumarizados na tabela 6.5 junto com as outras distribuições para a comparação. Vemos mais uma vez que ambas as redes neurais, @NN e @NN (Muons/noMuons), possuem performance superior à correção @GSC. Mostrando que para as correções *por jato* a rede neural com informação sobre o múon dentro do cone do jato possui a melhor performance.

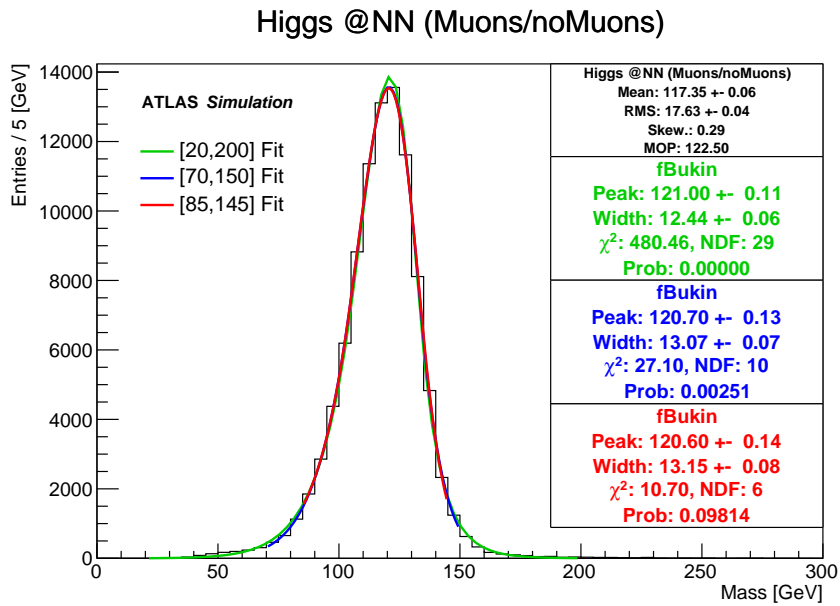
O resultado do cálculo da divergência é apresentado na tabela 6.6 mostrando novamente que a correção neural, levando em consideração os múons dentro dos jatos, é menos divergente que a correção @GSC utilizada pela colaboração ATLAS.

## Efeito no Background

O efeito da correção por jato utilizando suas características internas é mostrado na figura 6.9 onde é possível ver que as correções neurais @NN e @NN (Muons/noMuons) também não produzem um pico de massa na região do Higgs quando são aplicadas aos eventos de background.



(a)



(b)

Figura 6.8: Ajuste da função Bukin para a distribuição de  $m_{b\bar{b}}$  para eventos corrigidos com a rede neural para todos os jatos (@NN) e corrigidos com uma rede para os jatos com múon e outra para os jatos sem múon (@NN Muons/noMuons). Correções *por jato* utilizando um descritor de jatos.

Tabela 6.5: Ajuste da função Bukin para diferentes intervalos de  $m_{b\bar{b}}$ . Correções *por jato* utilizando um descritor de jatos.

[20 – 200] GeV				
	@TRUTH	@GSC	@NN	@NN (Muons/noMuons)
Peak [GeV]	$121.70 \pm 0.04$	$116.39 \pm 0.14$	$122.11 \pm 0.14$	$121.00 \pm 0.11$
Width [GeV]	$5.96 \pm 0.03$	$13.95 \pm 0.07$	$13.47 \pm 0.07$	$12.44 \pm 0.06$
$\chi^2/NDF$	66.78	15.51	18.25	16.57
Prob.	0.00	0.00	0.00	0.00
[70 – 150] GeV				
	@TRUTH	@GSC	@NN	@NN (Muons/noMuons)
Peak [GeV]	$121.93 \pm 0.04$	$115.47 \pm 0.14$	$121.98 \pm 0.13$	$120.70 \pm 0.13$
Width [GeV]	$5.53 \pm 0.02$	$14.72 \pm 0.08$	$14.15 \pm 0.08$	$13.07 \pm 0.07$
$\chi^2/NDF$	13.41	1.34	1.31	2.71
Prob.	0.00	0.20	0.21	0.00
[85 – 145] GeV				
	@TRUTH	@GSC	@NN	@NN (Muons/noMuons)
Peak [GeV]	$122.01 \pm 0.05$	$115.33 \pm 0.16$	$122.05 \pm 0.15$	$120.60 \pm 0.14$
Width [GeV]	$5.43 \pm 0.03$	$14.70 \pm 0.11$	$14.18 \pm 0.10$	$13.15 \pm 0.08$
$\chi^2/NDF$	6.46	0.98	1.45	1.78
Prob.	0.00	0.44	0.19	0.09

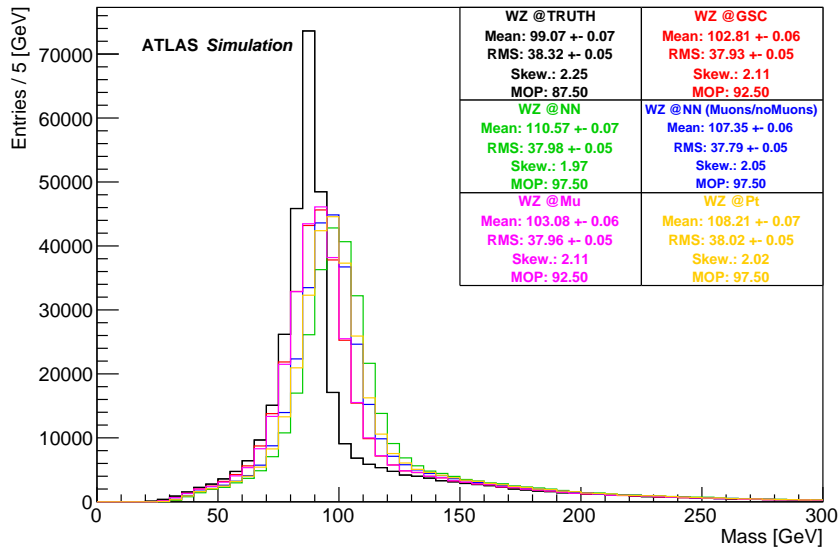


Figura 6.9: Distribuições de  $m_{b\bar{b}}$ , simulada e corrigida pelo corretor neural *por jato* utilizando um descritor de jatos e aplicadas na amostra de background  $WZ$ .

Tabela 6.6: Comparação da divergência JS para a distribuição de massa corrigida com a correção neural *por jato* utilizando um descritor de jatos.

	Bukin 1	Bukin 2	Bukin 3	TRUTH	GSC	NN	NN (Muons/noMuons)
TRUTH	$4.24 \times 10^{-3}$	$2.55 \times 10^{-4}$	$3.20 \times 10^{-4}$	0.0	$6.18 \times 10^{-2}$	$6.58 \times 10^{-2}$	$5.43 \times 10^{-2}$
GSC	$7.13 \times 10^{-4}$	$1.55 \times 10^{-4}$	$1.19 \times 10^{-3}$	$6.18 \times 10^{-2}$	0.0	$2.70 \times 10^{-2}$	$2.33 \times 10^{-2}$
NN	$9.60 \times 10^{-4}$	$1.49 \times 10^{-4}$	$7.13 \times 10^{-4}$	$6.58 \times 10^{-2}$	$2.70 \times 10^{-2}$	0.0	$1.03 \times 10^{-3}$
NN (Muons/noMuons)	$7.62 \times 10^{-4}$	$1.63 \times 10^{-4}$	$6.15 \times 10^{-4}$	$5.43 \times 10^{-2}$	$2.33 \times 10^{-2}$	$1.03 \times 10^{-3}$	0.0

## 6.2 Correções Por Evento

Neste tipo de correção, como visto na seção 5.4, utilizamos as variáveis do evento, ou seja, a cada iteração a rede recebe variáveis que em conjunto possuem a característica de terem sido originadas de um evento do bóson de Higgs, de acordo com a amostra de treinamento usada.

### 6.2.1 Correção Utilizando o Momento Transverso dos Jatos

Temos como entrada o momento transversal de cada jato  $b$  (@GSC) proveniente do decaimento do bóson de Higgs. Consequentemente também teremos dois alvos nesta rede, o momento transversal de cada jato como previsto pelo Modelo Padrão (@TRUTH) fazendo com que esta correção seja estimada através de uma rede neural com duas variáveis de entrada e duas variáveis de saída.

O MSE médio para cada topologia é mostrado na figura 6.10 onde vemos que a partir de três neurônios na camada escondida os valores de MSE estão todos iguais e dentro das barras de incerteza. Logo, para a aplicação da correção foi escolhida a rede neural com três neurônios na camada escondida, uma vez que esta é a rede neural de menor complexidade e com a performance similar às outras redes com mais neurônios na camada intermediária.

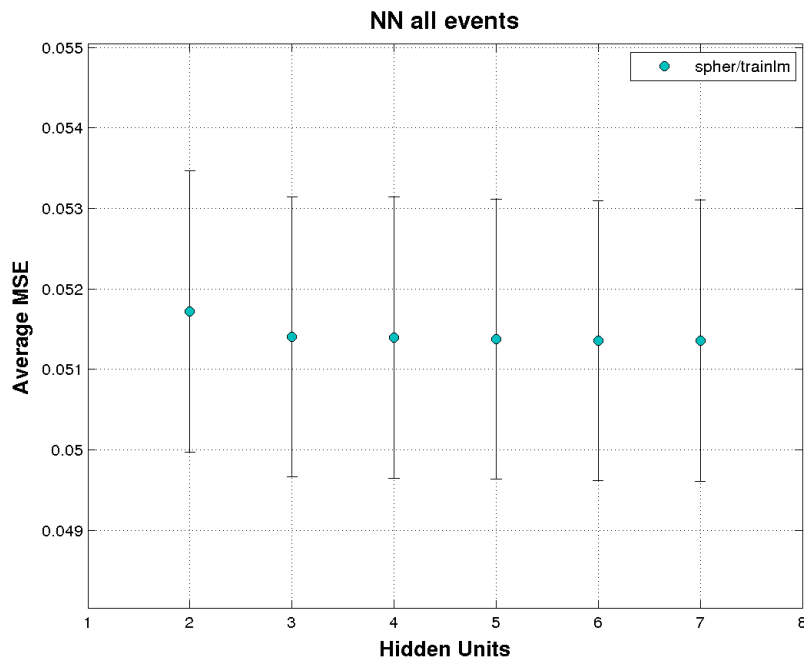


Figura 6.10: MSE médio para cada topologia da rede neural por evento utilizando o momento transversal.

O resultado desta correção para a distribuição de massa do Higgs é mostrado na figura 6.11 e resumido na tabela 6.7. A correção neural mais uma vez se mostra



melhor que a correção @GSC aproximando a média e a largura da distribuição em direção aos valores @TRUTH. Novamente a correção neural acerta o MOP da distribuição verdadeira, 122.50 GeV, e reduz em aproximadamente 4% a sua resolução em relação a correção @GSC. A rede neural também superestima a média da distribuição verdadeira em 1.6% obtendo um resultado melhor que a correção @GSC que possui média aproximadamente 5% distante do valor esperado.

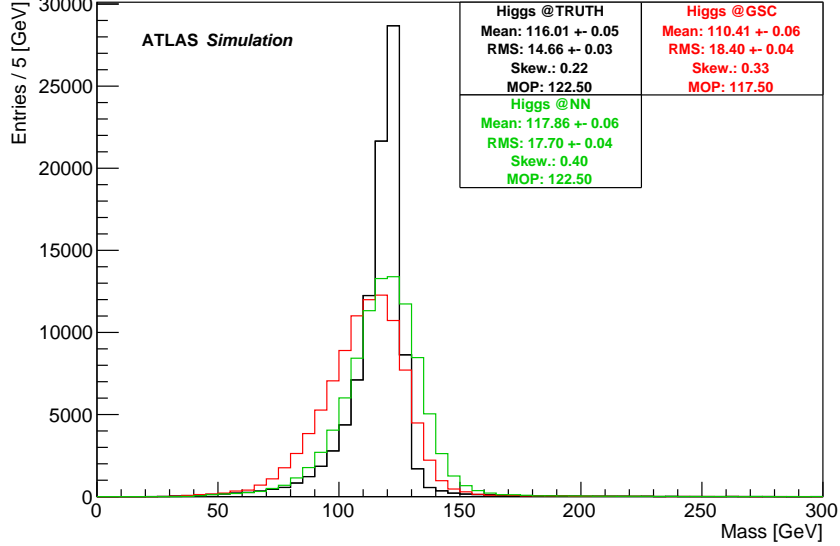


Figura 6.11: Distribuições de  $m_{b\bar{b}}$ , simulada e corrigida pelo corretor neural *por evento* utilizando o momento transversal.

Tabela 6.7: Estimativas para o histograma de massa invariante simulada e corrigida pelo corretor neural *por evento* utilizando o momento transversal.

	@TRUTH	@GSC	@NN
Mean [GeV]	116.01 ± 0.05	110.41 ± 0.06	117.86 ± 0.06
RMS [GeV]	14.66 ± 0.03	18.40 ± 0.04	17.70 ± 0.04
Skew	0.22	0.33	0.40
MOP [GeV]	122.50	117.50	122.50

O ajuste da correção neural por evento é mostrado na figura 6.12 e seus valores se encontram na tabela 6.8. Para todos os intervalos de ajuste a correção neural aproxima a escala, *peak*, e a resolução, *width*, da distribuição em direção aos parâmetros esperados pelos ajustes da distribuição @TRUTH. Porém, para os intervalos de ajuste de 70 – 150 GeV e 85 – 145 GeV a correção neural não exclui o modelo de função Bukin como visto nos valores de probabilidade 0.39 e 0.47, respectivamente. A correção neural deveria aproximar os valores dos parâmetros de escala e resolução e ao mesmo tempo se distanciar da função Bukin, uma vez que a função Bukin é excluída como modelo da distribuição @TRUTH pelo teste de  $\chi^2$ .

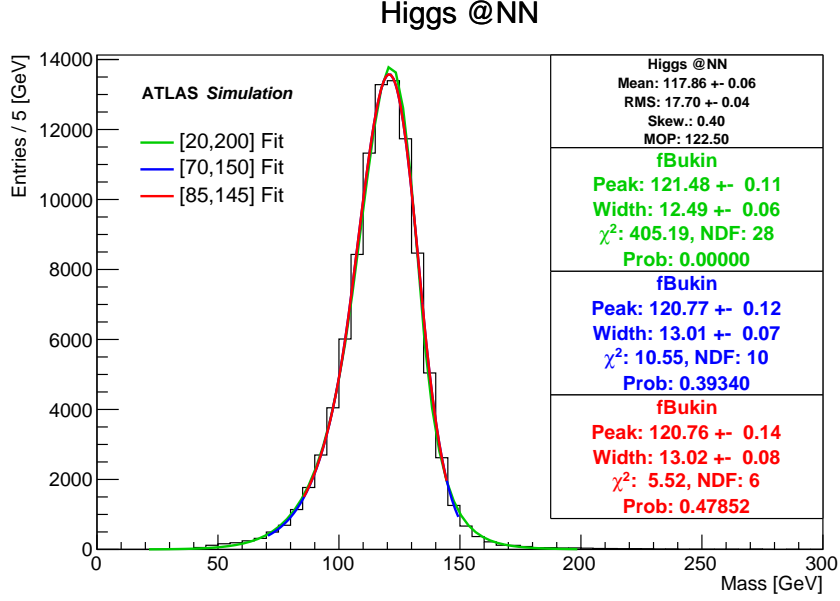


Figura 6.12: Ajuste da função Bukin para a distribuição de  $m_{b\bar{b}}$  para eventos corrigidos com a rede neural (@NN). Correção *por evento* utilizando o momento transverso.

Tabela 6.8: Ajuste da função Bukin para diferentes intervalos de  $m_{b\bar{b}}$ . Correções *por evento* utilizando o momento transverso.

[20 – 200] GeV			
	@TRUTH	@GSC	@NN
Peak [GeV]	121.70 ± 0.04	116.39 ± 0.14	121.48 ± 0.11
Width [GeV]	5.96 ± 0.03	13.95 ± 0.07	12.49 ± 0.06
$\chi^2/NDF$	66.78	15.51	14.47
Prob.	0.00	0.00	0.00
[70 – 150] GeV			
	@TRUTH	@GSC	@NN
Peak [GeV]	121.93 ± 0.04	115.47 ± 0.14	120.70 ± 0.12
Width [GeV]	5.53 ± 0.02	14.72 ± 0.08	13.01 ± 0.07
$\chi^2/NDF$	13.41	1.34	1.06
Prob.	0.00	0.20	0.39
[85 – 145] GeV			
	@TRUTH	@GSC	@NN
Peak [GeV]	122.01 ± 0.05	115.33 ± 0.16	120.76 ± 0.14
Width [GeV]	5.43 ± 0.03	14.70 ± 0.11	13.02 ± 0.08
$\chi^2/NDF$	6.46	0.98	0.92
Prob.	0.00	0.44	0.47

As divergências entre as distribuições de probabilidade são mostradas na tabela 6.9 onde, mais uma vez, vemos que a correção neural tem sua distribuição de massa menos divergente da distribuição esperada com  $\mathcal{D}(NN|Truth) = 5.66 \times 10^2$ .

Tabela 6.9: Comparação da divergência JS para a distribuição de massa corrigida com a correção neural *por evento* utilizando o momento transverso.

	Bukin 1	Bukin 2	Bukin 3	TRUTH	GSC	NN
TRUTH	$4.24 \times 10^{-3}$	$2.55 \times 10^{-4}$	$3.20 \times 10^{-4}$	0.0	$6.18 \times 10^{-2}$	$5.66 \times 10^{-2}$
GSC	$7.13 \times 10^{-4}$	$1.55 \times 10^{-4}$	$1.19 \times 10^{-3}$	$6.18 \times 10^{-2}$	0.0	$2.64 \times 10^{-2}$
NN	$6.12 \times 10^{-4}$	$1.49 \times 10^{-4}$	$6.44 \times 10^{-4}$	$5.66 \times 10^{-2}$	$2.64 \times 10^{-2}$	0.0

## Efeito no Background

O efeito das correções de energia no background  $WZ$  é exibido na figura 6.13. Embora a correção neural não apresente um pico na faixa de massa do Higgs, ela é a que possui o maior deslocamento quando comparada com todas as outras distribuições.

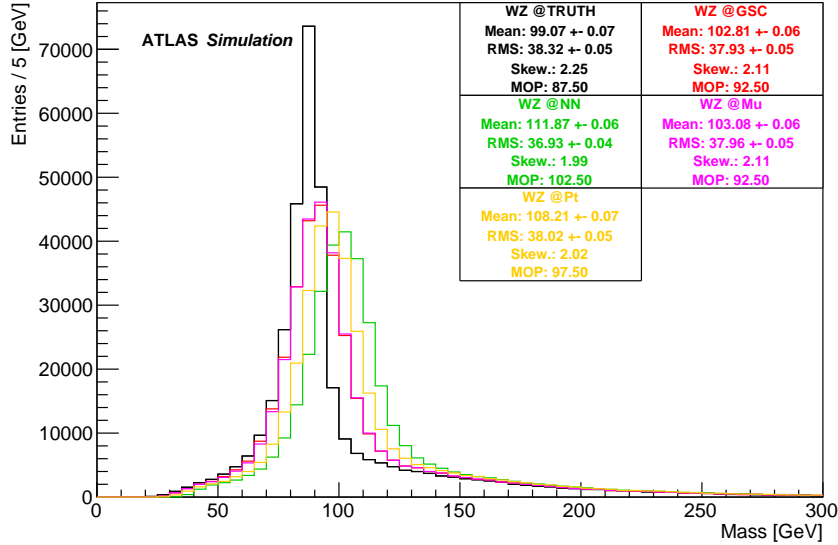


Figura 6.13: Distribuições de  $m_{b\bar{b}}$ , simulada e corrigida pelo corretor neural *por evento* utilizando o momento transverso e aplicadas na amostra de background  $WZ$ .

## 6.2.2 Correção Utilizando um Descritor dos Jatos

Aqui mostramos o resultado das correções por evento utilizando um descritor total dos jatos presentes no evento de Higgs. O resultado do treinamento pode ser visto na figura 6.14. A rede escolhida possui 22 neurônios na camada escondida, pois é a

partir desta topologia que começa a saturação no valor do erro médio quadrático. Além disto, para 22 neurônios possuímos a menor barra de incerteza.

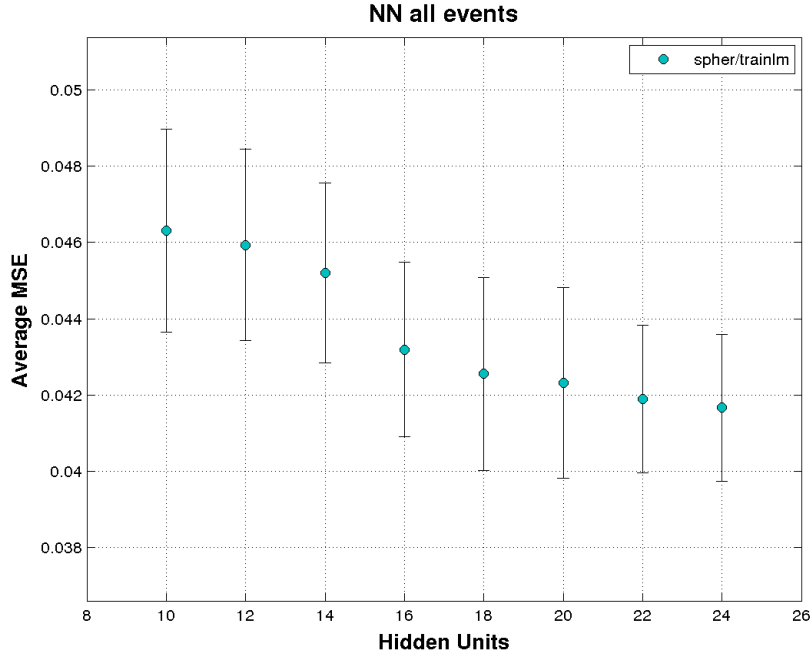


Figura 6.14: MSE médio para cada topologia da rede neural por evento utilizando um descritor de jatos.

A distribuição de massa para a correção neural por evento utilizando o descritor dos jatos é mostrada na figura 6.15 e sua função de ajuste na figura 6.16 seguidos de seus valores nas tabelas 6.10 e 6.11. Este resultado é o que mais difere das outras correções mostradas até agora. Apesar da correção não acertar o MOP da distribuição esperada, 117.50 GeV ao invés de 122.50 GeV, a escala da correção @NN superestima a escala da distribuição @TRUTH em apenas 0.4% e sua resolução diminui drasticamente,  $11.87 \pm 0.03$  GeV @NN, sendo ainda menor que a resolução da distribuição esperada pelo Modelo Padrão,  $14.66 \pm 0.03$  GeV @TRUTH.

No que tange os resultados do ajuste da função Bukin podemos dizer que a rede neural também diminui drasticamente a resolução porém seu impacto na escala da distribuição não é tão forte quanto por exemplo o da correção *por evento* utilizando apenas o momento transversal dos jatos (tabela 6.7).

Os resultados do cálculo da divergência são mostrados na tabela 6.12 onde podemos ver que a divergência entre a distribuição de massa invariante do bóson de Higgs corrigido com a correção neural proposta nesta seção, utilizando um descritor dos jatos presentes no evento, e a distribuição de massa esperada pela teoria é de  $\mathcal{D}(NN|Truth) = 2.20 \times 10^{-2}$  enquanto a divergência da correção oficial da colaboração é de  $\mathcal{D}(GSC|Truth) = 6.18 \times 10^{-2}$ , aproximadamente 2.8 vezes maior.

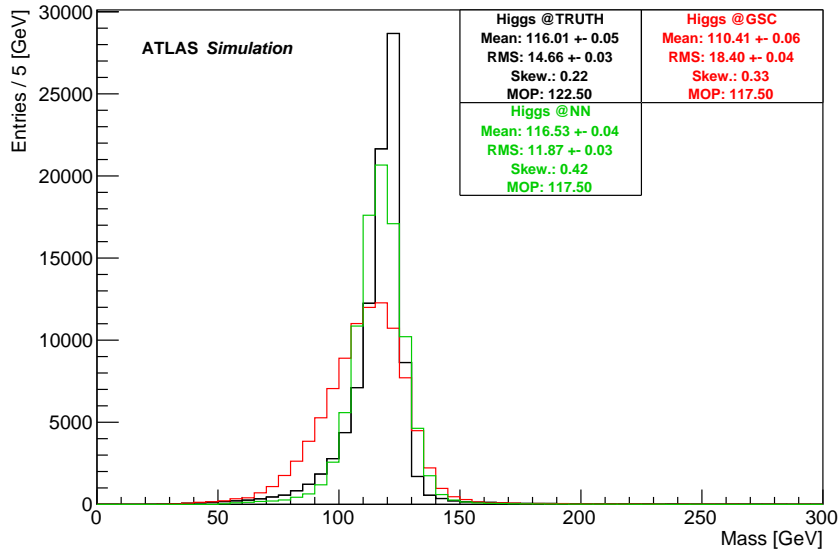


Figura 6.15: Distribuições de  $m_{b\bar{b}}$ , simulada e corrigida pelo corretor neural *por evento* utilizando um descritor de jatos.

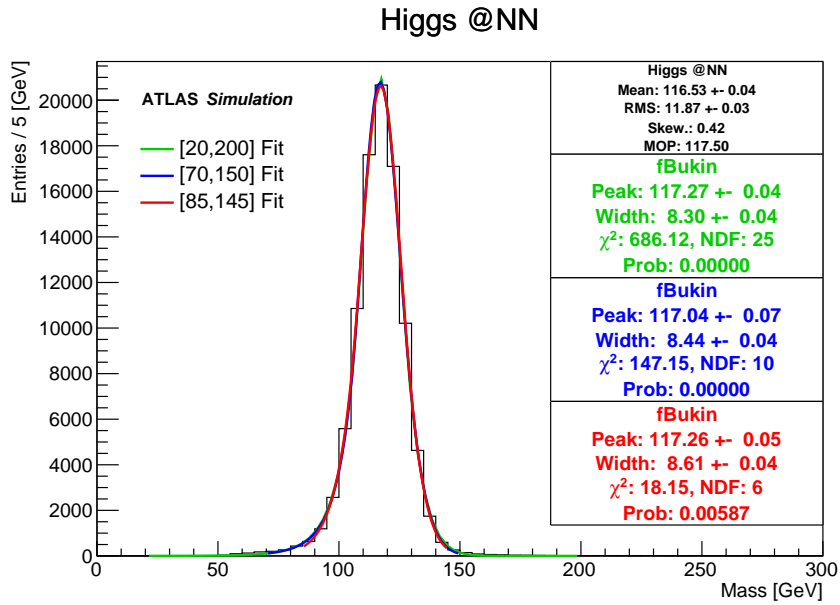


Figura 6.16: Ajuste da função Bukin para a distribuição de  $m_{b\bar{b}}$  para eventos corrigidos com a rede neural (@NN). Correção *por evento* utilizando um descritor de jatos.

Tabela 6.10: Estimativas para o histograma de massa invariante simulada e corrigida pelo corretor neural *por evento* utilizando um descritor de jatos.

	@TRUTH	@GSC	@NN
Mean [GeV]	116.01 ± 0.05	110.41 ± 0.06	116.53 ± 0.04
RMS [GeV]	14.66 ± 0.03	18.40 ± 0.04	11.87 ± 0.03
Skew	0.22	0.33	0.42
MOP [GeV]	122.50	117.50	117.50

Tabela 6.11: Ajuste da função Bukin para diferentes intervalos de  $m_{b\bar{b}}$ . Correções *por evento* utilizando um descritor de jatos.

[20 – 200] GeV			
	@TRUTH	@GSC	@NN
Peak [GeV]	$121.70 \pm 0.04$	$116.39 \pm 0.14$	$117.27 \pm 0.04$
Width [GeV]	$5.96 \pm 0.03$	$13.95 \pm 0.07$	$8.30 \pm 0.04$
$\chi^2/NDF$	66.78	15.51	27.44
Prob.	0.00	0.00	0.00
[70 – 150] GeV			
	@TRUTH	@GSC	@NN
Peak [GeV]	$121.93 \pm 0.04$	$115.47 \pm 0.14$	$117.04 \pm 0.07$
Width [GeV]	$5.53 \pm 0.02$	$14.72 \pm 0.08$	$8.44 \pm 0.04$
$\chi^2/NDF$	13.41	1.34	14.72
Prob.	0.00	0.20	0.00
[85 – 145] GeV			
	@TRUTH	@GSC	@NN
Peak [GeV]	$122.01 \pm 0.05$	$115.33 \pm 0.16$	$117.26 \pm 0.05$
Width [GeV]	$5.43 \pm 0.03$	$14.70 \pm 0.11$	$8.61 \pm 0.04$
$\chi^2/NDF$	6.46	0.98	3.03
Prob.	0.00	0.44	0.01

Tabela 6.12: Comparação da divergência JS para a distribuição de massa corrigida com a correção neural *por evento* utilizando um descritor de jatos.

	Bukin 1	Bukin 2	Bukin 3	TRUTH	GSC	NN
TRUTH	$4.24 \times 10^{-3}$	$2.55 \times 10^{-4}$	$3.20 \times 10^{-4}$	0.0	$6.18 \times 10^{-2}$	$2.20 \times 10^{-2}$
GSC	$7.13 \times 10^{-4}$	$1.55 \times 10^{-4}$	$1.19 \times 10^{-3}$	$6.18 \times 10^{-2}$	0.0	$4.80 \times 10^{-2}$
NN	$1.48 \times 10^{-4}$	$2.15 \times 10^{-4}$	$8.85 \times 10^{-5}$	$2.20 \times 10^{-2}$	$4.80 \times 10^{-2}$	0.0

## Efeito no Background

A figura 6.17 mostra o efeito da correção, descrita acima, no background  $WZ$ . Onde pode ser visto que a rede neural produz eventos de Higgs na amostra de background, comportamento apresentado pelo enorme pico da distribuição @NN na região de massa prevista pelo Modelo Padrão para o sinal do bóson de Higgs.

Este efeito pode ser explicado pela amostra de treinamento da rede neural. Em um treinamento de correção *por evento* mostramos apenas eventos de Higgs para a rede neural, de tal modo que os eventos de background que sejam semelhantes ao Higgs, em termos das características que apresentam originalmente no espaço de dimensão elevada, são mapeados similarmente no pico do Higgs. Contudo para todos os efeitos esta correção *por evento*, utilizando um descritor dos jatos presentes no evento, não deve ser utilizada por conta de seu comportamento na amostra de background.

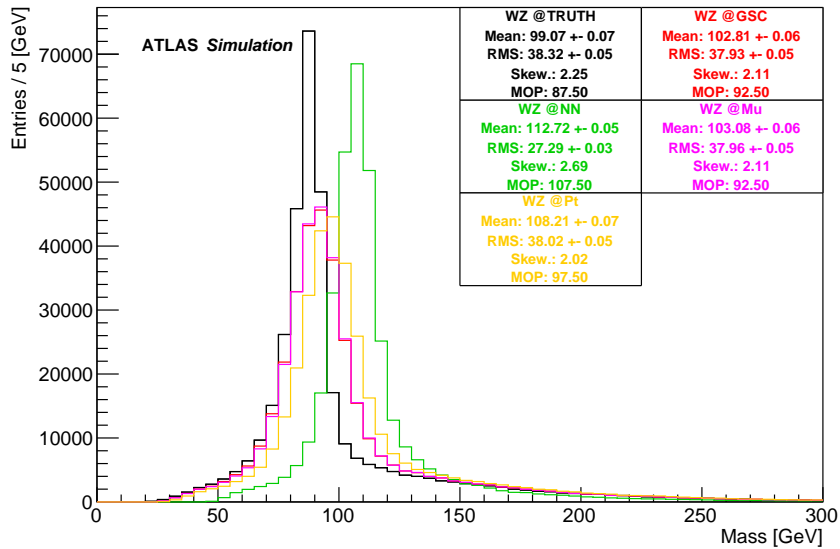


Figura 6.17: Distribuições de  $m_{b\bar{b}}$ , simulada e corrigida pelo corretor neural *por evento* utilizando um descritor de jatos e aplicadas na amostra de background  $WZ$ .

## 6.3 Boosting das Correções Neurais

Nesta seção serão mostrados os resultados do algoritmo de *ensemble* discutido na seção 5.9. Esta técnica foi aplicada nas correções que utilizaram o descritor com as características internas dos jatos de cada evento. Assim a primeira iteração do *boosting* é feita com a rede neural vencedora, por jato e por evento, mostrada anteriormente. A partir daí o algoritmo AdaBoost.MRT faz a re-amostragem dos eventos que se encontraram acima do limiar escolhido e treina uma nova rede neural com essas amostras a cada iteração.

A figura 6.18 mostra o efeito de cada iteração do algoritmo AdaBoost.MRT para a rede *por evento*. É visto que conforme as iterações avançam os valores da média vão se afastando do valor esperado. Por exemplo, na última iteração o valor de média é  $118.97 \pm 0.05$  GeV enquanto a rede original por evento @NN possui média  $116.53 \pm 0.04$  GeV estando mais próxima do valor esperado  $116.01 \pm 0.05$ . Por outro lado, além de acertar o valor do MOP, os valores de RMS vão se tornando mais plausíveis, no sentido de não estarem tão abaixo do RMS esperado pela distribuição @TRUTH, assim sabemos que a correção neural não supera a distribuição verdadeira.

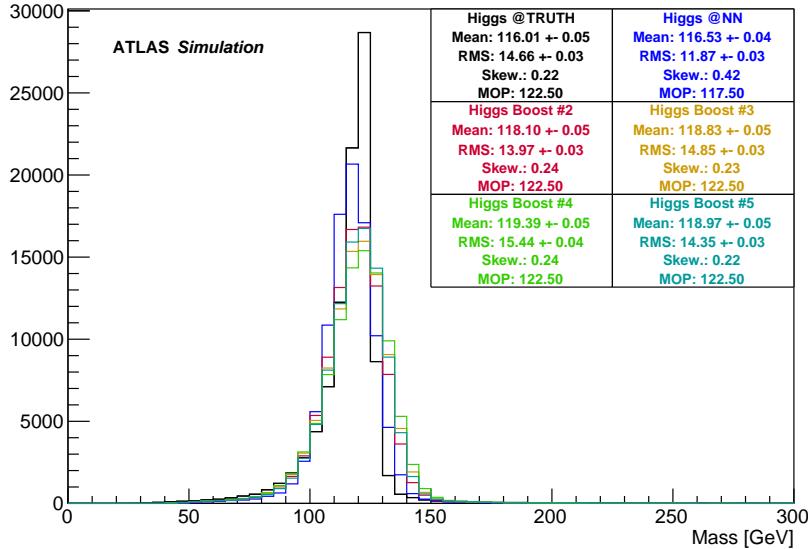


Figura 6.18: Efeito na massa invariante do Higgs para cada iteração do *boosting* na correção por evento.

Contudo, ainda é necessário checar o efeito do algoritmo de *boosting* na amostra de *background*, isto é mostrado na figura 6.19 onde é possível ver que o problema de criar eventos na faixa de massa do bóson de Higgs ainda persiste em todas as iterações testadas.

O algoritmo AdaBoost.MRT também foi testado nas correções *por jato* como é visto na figura 6.20. Pode-se ver que para este caso o algoritmo não teve um efeito



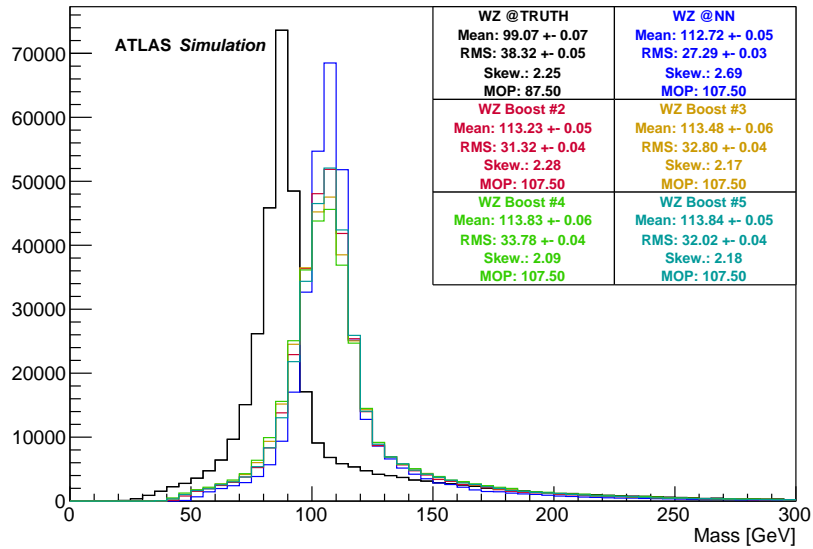


Figura 6.19: Efeito na massa invariante dos jatos  $b$  na amostra de  $background$  ( $WZ$ ) para cada iteração do  $boosting$  na correção por evento.

significativo na distribuição de massa em todas as iterações em que foi executado. Inclusive possui um efeito negativo no valor da média, se afastando mais do valor da média esperada pela distribuição @TRUTH.

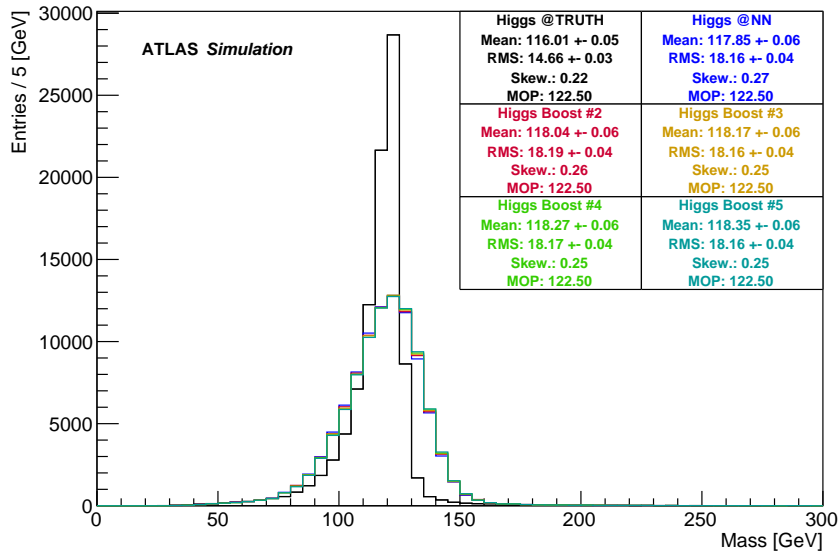


Figura 6.20: Efeito na massa invariante do Higgs para cada iteração do  $boosting$  na correção por jato.

O efeito do algoritmo no  $background$  é mostrado na figura 6.21, onde é visto que o algoritmo de  $boosting$  não causa nenhum efeito na distribuição de massa do  $background$ . A distribuição permanece praticamente a mesma desde a primeira iteração.

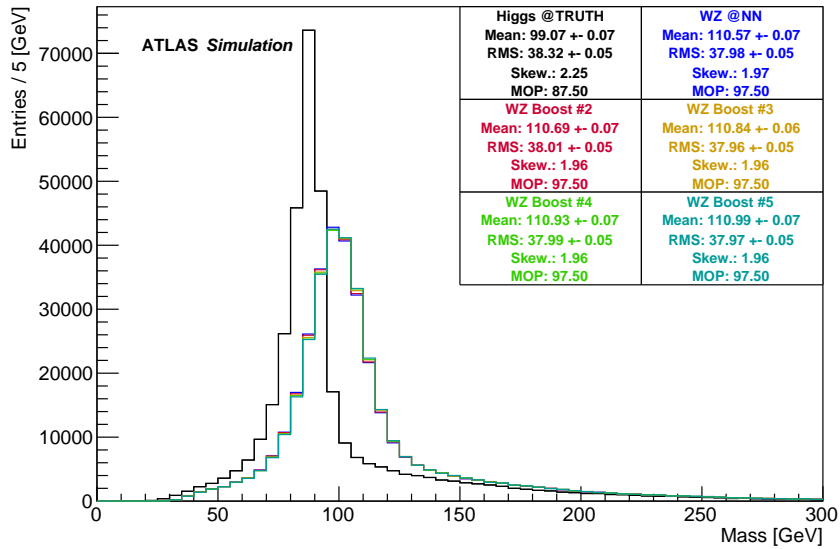


Figura 6.21: Efeito na massa invariante dos jatos  $b$  na amostra de  $background$  ( $WZ$ ) para cada iteração do  $boosting$  na correção por jato.

## 6.4 Comparação dos Modelos Neurais Propostos

Esta seção compara os modelos neurais propostos neste trabalho com as outras correções específicas para jatos  $b$  também estudadas pela Colaboração ATLAS e discutidas em detalhe nas seções 4.4.6 e 4.4.7. A figura 6.22 apresenta um diagrama com todas as correções de energia testadas pela colaboração ATLAS, inclusive as correções específicas para jatos  $b$ .

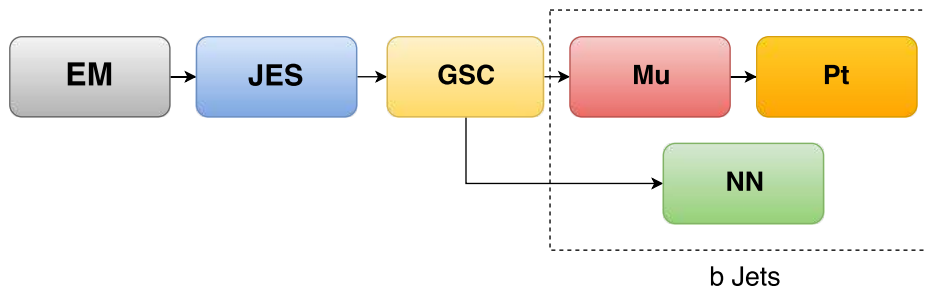


Figura 6.22: Diagrama com as correções sequenciais testadas pela colaboração ATLAS.

Vimos que as correções *por evento* não possuem um resultado vantajoso, uma vez que a correção que utiliza apenas o momento transversal dos jatos, apesar de apresentar os parâmetros da função Bukin mais próximos dos valores da distribuição @TRUTH, não exclui o modelo Bukin pelo seu teste de  $\chi^2$  como ocorre com a distribuição esperada (tabela 6.8). Além disso, a correção *por evento* que utiliza um descritor completo de cada jato no evento possui o maior efeito no background como exibido na figura 6.17.

Nas correções *por jato* foi mostrado que é sempre mais vantajoso separar os jatos que possuem múons dos que não possuem, fazendo assim uma rede neural para cada caso. Ainda que ambas correções *por jato* não apresentem um comportamento inadequado no background, a correção utilizando um descritor mais completo dos jatos possui uma performance um pouco melhor do que a correção que utiliza somente o momento transversal como sinal de entrada. Esta melhora pode ser vista tanto nos parâmetros de seus histogramas, tabelas 6.1 e 6.4, quando nos valores de divergência em comparação com a distribuição @TRUTH,  $\mathcal{D}(NN \text{ Muons}|Truth) = 5.64 \times 10^{-2}$  para a correção com o momento transversal e  $\mathcal{D}(NN \text{ Muons}|Truth) = 5.43 \times 10^{-2}$  para a correção com o descritor dos jatos.

Dito isto, o modelo neural mais indicado para uma possível correção dos jatos  $b$  é a correção *por jato* utilizando um descritor completo de cada jato e separando em duas correções, uma para os jatos com múon e outras para os jatos sem múon, pois além de levar em consideração o múon dentro do cone do jato também corrige a energia do neutrino que escapa à detecção. A figura 6.23 mostra a comparação da correção neural vencedora com as outras correções da colaboração ATLAS, inclusive as correções específicas para jatos  $b$ . Os valores para os parâmetros dos histogramas são vistos na tabela 6.13.

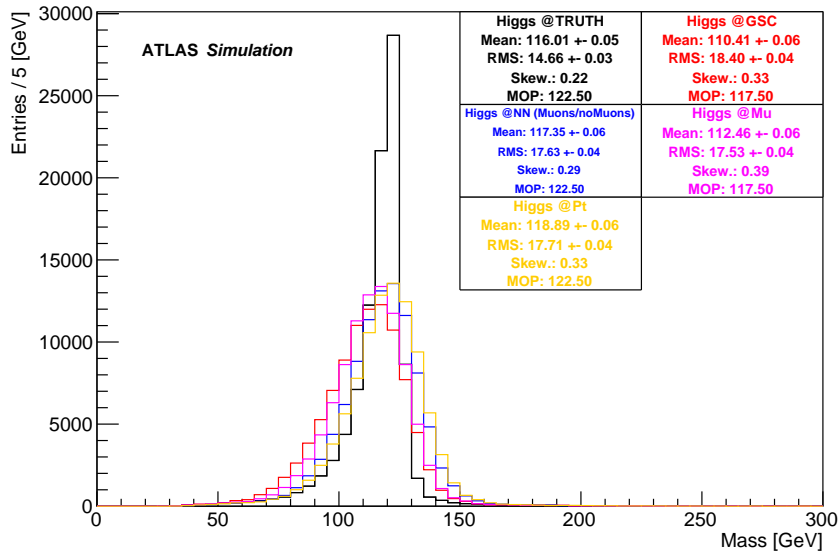


Figura 6.23: Distribuições de massa reconstruída com cada uma das correções para o sinal do decaimento  $H \rightarrow b\bar{b}$ .

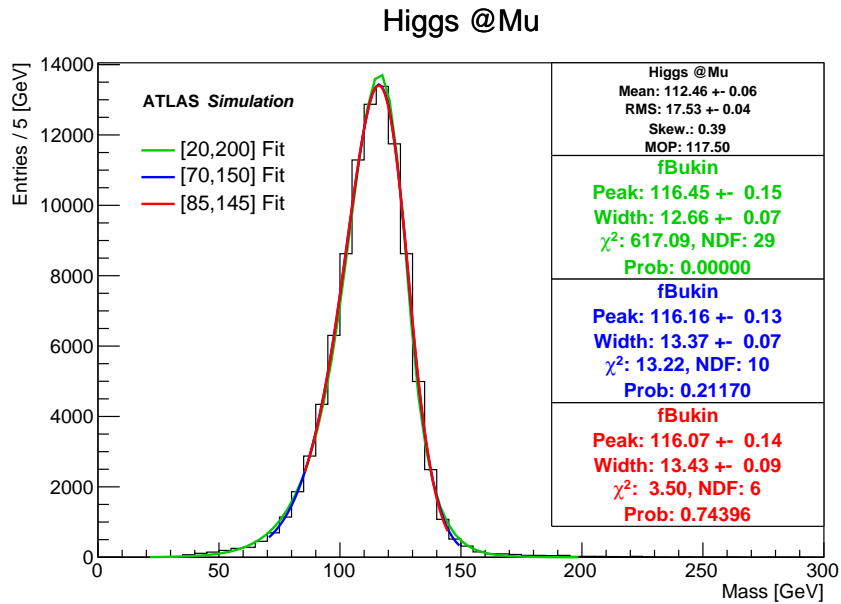
Vemos que a correção neural ainda possui o melhor parâmetro de escala quando comparada com a melhor correção para jatos  $b$ , a correção @Pt (seção 4.4.7). O corretor neural superestima a média da distribuição @TRUTH em 1.3% enquanto a correção @Pt superestima a média em 2.5%. Ambas as correções corrigem o pico da distribuição e suas resoluções, RMS, estão dentro dos intervalos de incerteza.

Tabela 6.13: Estimativas para o histograma de massa invariante gerada, reconstruída e corrigida com a correção neural.

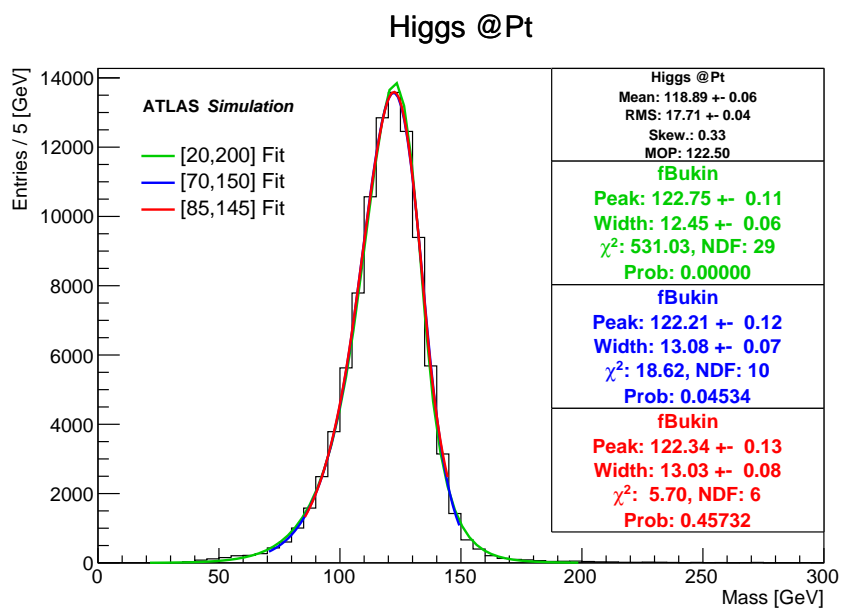
	@TRUTH	@GSC	@Mu	@Pt	@NN (Muons/noMuons)
Mean [GeV]	$116.01 \pm 0.05$	$110.41 \pm 0.06$	$112.46 \pm 0.06$	$118.89 \pm 0.06$	$117.35 \pm 0.06$
RMS [GeV]	$14.66 \pm 0.03$	$18.40 \pm 0.04$	$17.53 \pm 0.04$	$17.71 \pm 0.04$	$17.63 \pm 0.04$
Skew	0.22	0.33	0.39	0.33	0.29
MOP [GeV]	122.50	117.50	117.50	122.50	122.50

A figura 6.24 mostra os ajustes da função Bukin para as correções @Mu e @Pt. Os valores dos parâmetros se encontram na tabela 6.14. Para facilitar a comparação, junto com a escala @TRUTH só serão mostrados os valores das correções específicas para jatos  $b$ . É fácil ver que a correção neural supera a correção *Muon-in-Jet* em todos os intervalos de ajuste estudados. Para o intervalo que abrange toda a faixa dinâmica da massa do Higgs, 20 – 200 GeV, a correção neural também se mostra superior que a correção @Pt proposta e discutida na seção 4.4.7. Para o intervalo 70 – 150 GeV as correções @Pt e @NN (Muons/noMuons) possuem seus parâmetros de resolução iguais, dentro da faixa de incerteza, já a escala da correção @NN (Muons/noMuons) subestima a escala esperada em 1% enquanto a correção @Pt está distante da escala esperada em apenas 0.2%. Para o último intervalo de ajuste, 85 – 145 GeV, A correção @Pt possui seus parâmetros de escala e resolução melhores que a correção neural, porém a correção neural se afasta mais da distribuição Bukin como visto pelo teste de  $\chi^2$ , Prob = 0.09, e esperado pelo ajuste da distribuição @TRUTH.

Também foi calculada a divergência de Jensen-Shannon para as correções específicas de jatos  $b$ , seus valores são exibidos na tabela 6.15 onde vemos que a correção menos divergente, em relação à distribuição esperada de massa, é a *Muon-in-Jet* com  $\mathcal{D}(Mu|Truth) = 5.00 \times 10^{-2}$ . Apesar da correção @Mu não ter bons parâmetros de escala e resolução, tanto nos estimadores de média e RMS quanto nos parâmetros do ajuste da função Bukin, ela possui a distribuição de massa que mais se assemelha da distribuição de massa para o nível @TRUTH segundo a divergência de Jensen-Shannon. Próximo deste resultado temos a correção neural com  $\mathcal{D}(NN Mu|Truth) = 5.43 \times 10^{-2}$ . Para saber se a diferença entre estes valores é significativa, uma calibração dos valores de divergência para a distribuição  $m_{b\bar{b}}$  se faz necessária.



(a)



(b)

Figura 6.24: Ajuste da função Bukin para a distribuição de massa invariante do bóson de Higgs para os estágios de correção @Mu e @Pt.

Tabela 6.14: Ajuste da função Bukin para as correções @Mu, @Pt, @NN (Muons/noMuons) e seu ajuste para a distribuição verdadeira @TRUTH.

[20 – 200] GeV				
	@TRUTH	@Mu	@Pt	@NN (Muons/noMuons)
Peak [GeV]	$121.70 \pm 0.04$	$116.45 \pm 0.15$	$122.75 \pm 0.11$	$121.00 \pm 0.11$
Width [GeV]	$5.96 \pm 0.03$	$13.95 \pm 0.07$	$12.45 \pm 0.06$	$12.44 \pm 0.06$
$\chi^2/NDF$	66.78	21.28	18.31	16.57
Prob.	0.00	0.00	0.00	0.00
[70 – 150] GeV				
	@TRUTH	@Mu	@Pt	@NN (Muons/noMuons)
Peak [GeV]	$121.93 \pm 0.04$	$116.45 \pm 0.15$	$122.21 \pm 0.12$	$120.70 \pm 0.13$
Width [GeV]	$5.53 \pm 0.02$	$13.37 \pm 0.07$	$13.08 \pm 0.07$	$13.07 \pm 0.07$
$\chi^2/NDF$	13.41	1.32	1.86	2.71
Prob.	0.00	0.21	0.05	0.00
[85 – 145] GeV				
	@TRUTH	@Mu	@Pt	@NN (Muons/noMuons)
Peak [GeV]	$122.01 \pm 0.05$	$116.07 \pm 0.14$	$122.34 \pm 0.13$	$120.60 \pm 0.14$
Width [GeV]	$5.43 \pm 0.03$	$13.43 \pm 0.09$	$13.03 \pm 0.08$	$13.15 \pm 0.08$
$\chi^2/NDF$	6.46	0.58	0.95	1.78
Prob.	0.00	0.74	0.46	0.09

Tabela 6.15: Comparação da divergência JS para a distribuição de massa corrigida com a correção neural vencedora entre todos os modelos testados.

	Bukin 1	Bukin 2	Bukin 3	TRUTH	GSC	Mu	Pt	NN (Muons/noMuons)
TRUTH	$4.24 \times 10^{-3}$	$2.55 \times 10^{-4}$	$3.20 \times 10^{-4}$	0.0	$6.18 \times 10^{-2}$	$5.00 \times 10^{-2}$	$6.30 \times 10^{-2}$	$5.43 \times 10^{-2}$
GSC	$7.13 \times 10^{-4}$	$1.55 \times 10^{-4}$	$1.19 \times 10^{-3}$	$6.18 \times 10^{-2}$	0.0	$2.49 \times 10^{-3}$	$3.48 \times 10^{-2}$	$2.33 \times 10^{-2}$
Mu	$1.19 \times 10^{-3}$	$1.19 \times 10^{-4}$	$6.83 \times 10^{-4}$	$5.00 \times 10^{-2}$	$2.49 \times 10^{-3}$	0.0	$2.35 \times 10^{-2}$	$1.36 \times 10^{-2}$
Pt	$8.71 \times 10^{-4}$	$1.60 \times 10^{-4}$	$6.21 \times 10^{-4}$	$6.30 \times 10^{-2}$	$3.48 \times 10^{-2}$	$2.35 \times 10^{-2}$	0.0	$1.55 \times 10^{-3}$
NN (Muons/noMuons)	$7.62 \times 10^{-4}$	$1.63 \times 10^{-4}$	$6.15 \times 10^{-4}$	$5.43 \times 10^{-2}$	$2.33 \times 10^{-2}$	$1.36 \times 10^{-2}$	$1.55 \times 10^{-3}$	0.0

# Capítulo 7

## Conclusões e Trabalhos Futuros

Apesar da não-compensação do calorímetro do detector ATLAS, é possível minimizar a diferença de energia entre o valor medido e seu valor real aplicando algoritmos de calibração. Esta dissertação testou diferentes modelos de correção, todos eles baseados em redes neurais artificiais.

Nas correções *por jato* foi visto que a rede neural que separa os jatos com múons e sem múons dentro do seu interior e utiliza um descritor dos jatos  $b$ , com suas características internas e topológicas, possui uma performance melhor que a correção *Global Sequential Calibration* aprovada e utilizada pela colaboração do experimento. Esta correção neural foi melhor tanto na análise estatística quanto na análise de divergência entre suas distribuições de probabilidade.

As correções *por evento* não se mostraram favoráveis devido ao seu efeito em um dos backgrounds do  $H \rightarrow b\bar{b}$ . Principalmente a correção por evento que utiliza o descritor de cada jato do evento, esta correção neural produz eventos de Higgs na distribuição de massa do background, caracterizando assim o que é chamado de *over correction*.

O *boosting* das redes neurais não obteve benefício quando foi aplicado nas correções *por jato*, pois a distribuição de massa invariante não apresenta mudança significativa em cada re-amostragem do algoritmo. Por outro lado, quando aplicado nas correções *por evento* possui o melhor valor de resolução de todas as correções testadas, porém o problema do background permanece.

Contudo podemos concluir que, quando utilizamos informações presentes no processo físico estudado, a utilização de redes neurais se mostrou eficiente para a correção na estimativa de energia dos eventos do bóson de Higgs decaindo em dois jatos  $b$ .



## 7.1 Trabalhos Futuros

Uma vez que este trabalho utilizou os histogramas de massa, normalizados, para o cálculo das divergências, é possível apontar como um dos trabalhos futuros a utilização de métodos de *kernel* para a estimação das funções densidade de probabilidade de forma a melhorar o refinamento no cálculo das divergências entre as distribuições.

Além disso também é possível utilizar a divergência entre a distribuição de massa reconstruída e a distribuição esperada pelo Modelo Padrão como critério de treinamento da rede neural. Para isso é necessário um algoritmo de treinamento que não aplique derivadas como suas funções de atualização. Um método de atualização de pesos por enxames de partículas possui essa característica. As divergências também podem ser incorporadas como critério para a re-amostragem do algoritmo de *boosting*.

Para tentar contornar o problema da produção de eventos de sinal no background nas correções *por evento* é possível criar uma nova amostra de treinamento. Ao invés de treinar a rede neural somente com os eventos de sinal do Higgs, podemos adicionar à amostra de treinamento eventos de background, de forma que a rede neural consiga diferenciar estes eventos através de suas características no espaço de dimensão elevada.

Como trabalhos futuros também apontamos a utilização de outras técnicas de aprendizado de máquina para a estimação de correções de energia, como *Boosted Decision Trees* e métodos de *Deep Learning*.

# Referências Bibliográficas

- [1] CARUSO, F., OGURI, V. *Física Moderna - Origens Clássicas e Fundamentos Quânticos*. Elsevier/Campus - RJ, 2006.
- [2] NE'EMAN, Y., KIRSH, Y. *The Particle Hunters*. Cambridge University Press, 1996.
- [3] ALVES, G. *O Mundo das Partículas de Hoje e de Ontem*. Livraria da Física, 2012.
- [4] FEYNMAN, R. *QED the Strange Theory of Light and Matter*. Princeton University Press, 1990.
- [5] GRIFFITH, D. *Introduction to Elementary Particles*. Wiley-VCH, 2008.
- [6] HALZEN, F., MARTIN, A. D. *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. Wiley, 1984.
- [7] DITTMAYER, S., OTHERS. “Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables”, 2011. Disponível em: <<http://arxiv.org/abs/1101.0593>>.
- [8] ANDERSEN, J. R., OTHERS. “Handbook of LHC Higgs Cross Sections: 3. Higgs Properties”, 2013. Disponível em: <<http://arxiv.org/abs/1307.1347>>.
- [9] ATLAS COLLABORATION. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”, *Phys. Lett.*, v. B716, 2012. Disponível em: <<http://arxiv.org/abs/1207.7214>>.
- [10] CMS COLLABORATION. “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”, *Phys. Lett.*, v. B716, 2012. Disponível em: <<http://arxiv.org/abs/1207.7235>>.

- [11] CMS COLLABORATION. “Observation of the diphoton decay of the Higgs boson and measurement of its properties”, *Eur. Phys. J.*, v. C74, n. 10, 2014. Disponível em: <<http://arxiv.org/abs/1407.0558>>.
- [12] CMS COLLABORATION. “Measurement of Higgs boson production and properties in the WW decay channel with leptonic final states”, *JHEP*, v. 01, 2014. Disponível em: <<http://arxiv.org/abs/1312.1129>>.
- [13] ATLAS COLLABORATION. “Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector”, *JHEP*, v. 04, 2015. Disponível em: <<http://arxiv.org/abs/1501.04943>>.
- [14] EVANS, L., BRYANT, P. “LHC Machine”, *JINST*, v. 3, 2008.
- [15] *LEP design report*. Geneva, CERN, 1984. Disponível em: <<https://cds.cern.ch/record/102083>>.
- [16] ALICE COLLABORATION. “The ALICE experiment at the CERN LHC”, *JINST*, v. 3, 2008.
- [17] ATLAS COLLABORATION. “The ATLAS experiment at the CERN Large Hadron Collider”, *JINST*, v. 3, 2008.
- [18] CMS COLLABORATION. “The CMS experiment at the CERN LHC”, *JINST*, v. 3, 2008.
- [19] LHCb COLLABORATION. “The LHCb experiment at the CERN LHC”, *JINST*, v. 3, 2008.
- [20] HAFFNER, J. “The CERN accelerator complex. Complex des accélérateurs du CERN”, Disponível em: <<https://cds.cern.ch/record/1621894>>. General Photo.
- [21] WERMES, N., HALLEWEL, G. *ATLAS pixel detector: Technical Design Report*. Geneva, CERN, 1998. Disponível em: <<https://cds.cern.ch/record/381263>>.
- [22] ATLAS COLLABORATION. *ATLAS liquid-argon calorimeter: Technical Design Report*. Geneva, CERN, 1996. Disponível em: <<https://cds.cern.ch/record/331061>>.
- [23] ATLAS COLLABORATION. *ATLAS tile calorimeter: Technical Design Report*. Geneva, CERN, 1996. Disponível em: <<https://cds.cern.ch/record/331062>>.

- [24] ATLAS COLLABORATION. *ATLAS muon spectrometer: Technical Design Report*. Geneva, CERN, 1997. Disponível em: <<https://cds.cern.ch/record/331068>>.
- [25] AHMAD, A., OTHERS. *The Silicon Microstrip Sensors of the ATLAS SemiConductor Tracker*. Geneva, CERN, 2007. Disponível em: <<https://cds.cern.ch/record/1019885>>.
- [26] ABAT, E., OTHERS. “The ATLAS TRT barrel detector”, *JINST*, v. 3, 2008.
- [27] WIGMANS, R. *Calorimetry: energy measurement in particle physics*. Oxford, Clarendon Press, 2000.
- [28] FROIDEVAUX, D., SPHICAS, P. “General-purpose detectors for the Large Hadron Collider”, *Ann. Rev. Nucl. Part. Sci.*, v. 56, 2006.
- [29] ATLAS COLLABORATION. “Readiness of the ATLAS Tile Calorimeter for LHC collisions”, *Eur. Phys. J.*, v. C70, 2010. Disponível em: <<http://arxiv.org/abs/1007.5423>>.
- [30] GINGRICH, D. M., OTHERS. “Construction, assembly and testing of the ATLAS hadronic end-cap calorimeter”, *JINST*, v. 2, 2007.
- [31] ARTAMONOV, A., OTHERS. “The ATLAS Forward Calorimeter”, *JINST*, v. 9, 2014.
- [32] ATLAS COLLABORATION. “ATLAS magnet system: Technical Design Report”, 1997. Disponível em: <<https://cds.cern.ch/record/338080>>.
- [33] PANDURO VAZQUEZ, J. G. “Atlas Data Acquisition: from Run I to Run II”, 2014. Disponível em: <<https://cds.cern.ch/record/1727972>>.
- [34] KRASZNAHORKAY, A. “The evolution of the Trigger and Data Acquisition System in the ATLAS experiment”, *J. Phys. Conf. Ser.*, v. 523, 2014.
- [35] CAMPAGNARI, C., FRANKLIN, M. “The Discovery of the top quark”, *Rev. Mod. Phys.*, v. 69, 1997. Disponível em: <<http://arxiv.org/abs/hep-ex/9608003>>.
- [36] HESKETH, G. “W/Z + jet Production at the LHC”. In: *24th Rencontres de Blois on Particle Physics and Cosmology Blois, Loire Valley, France*, 2012. Disponível em: <<http://arxiv.org/abs/1211.0496>>.
- [37] CAMARDA, S. “W/Z+Jets and W/Z+HF Production at the Tevatron”. In: *Proceedings, 48th Rencontres de Moriond on QCD and High Energy Interactions*, 2013. Disponível em: <<http://arxiv.org/abs/1306.2579>>.

- [38] ATLAS COLLABORATION. “Search for the  $b\bar{b}$  decay of the Standard Model Higgs boson in associated (W/Z)H production with the ATLAS detector”, *JHEP*, v. 01, 2015. Disponível em: <<http://arxiv.org/abs/1409.6212>>.
- [39] ABACHI, S., OTHERS. “The D0 Detector”, *Nucl. Instrum. Meth.*, v. A338, 1994.
- [40] BLAZEY, G. C., OTHERS. “Run II jet physics”. In: *QCD and weak boson physics in Run II. Proceedings, Batavia, USA*, 2000. Disponível em: <<http://arxiv.org/abs/hep-ex/0005012>>.
- [41] ABAZOV, V. M., OTHERS. “Jet energy scale determination in the D0 experiment”, *Nucl. Instrum. Meth.*, v. A763, 2014. Disponível em: <<http://arxiv.org/abs/1312.6873>>.
- [42] ABE, F., OTHERS. “The CDF Detector: An Overview”, *Nucl. Instrum. Meth.*, v. A271, 1988.
- [43] BHATTI, A., OTHERS. “Determination of the Jet Energy Scale at the Collider Detector at Fermilab”, *Nucl. Instrum. Meth.*, v. A566, 2006. Disponível em: <<http://arxiv.org/abs/hep-ex/0510047>>.
- [44] THE CMS COLLABORATION. “Jet Reconstruction Performance at CMS”, *CMS Physics Analysis Summary*, 2009.
- [45] CACCIARI, M., SALAM, G. P., SOYEZ, G. “The Anti-k(t) jet clustering algorithm”, *JHEP*, v. 04, 2008. Disponível em: <<http://arxiv.org/abs/0802.1189>>.
- [46] THE CMS COLLABORATION. “Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and  $E_T^{miss}$ ”, *CMS Physics Analysis Summary*, v. CMS-PAS-PFT-09-001, 2009.
- [47] BEAUDETTE, F. “The CMS Particle Flow Algorithm”. In: *Proceedings, International Conference on Calorimetry for the High Energy Frontier (CHEF 2013)*, 2013.
- [48] CACCIARI, M., SALAM, G. P. “Pileup subtraction using jet areas”, *Phys. Lett.*, v. B659, 2008. Disponível em: <<http://arxiv.org/abs/0707.1378v2>>.
- [49] THE CMS COLLABORATION. “Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS”, *JINST*, 2011. Disponível em: <<http://arxiv.org/abs/1107.4277>>.

- [50] LAMPL, W., LAPLACE, S., LELAS, D., et al. “Calorimeter clustering algorithms: Description and performance”, 2008. Disponible em: <https://cds.cern.ch/record/1099735>.
- [51] AHARROUCHE, M., OTHERS. “Measurement of the response of the ATLAS liquid argon barrel calorimeter to electrons at the 2004 combined test-beam”, *Nucl. Instrum. Meth.*, v. A614, 2010.
- [52] ABAT, E., OTHERS. “Combined performance studies for electrons at the 2004 ATLAS combined test-beam”, *JINST*, v. 5, 2010.
- [53] CACCIARI, M., SALAM, G. P. “Dispelling the  $N^3$  myth for the  $k_t$  jet-finder”, *Phys. Lett.*, v. B641, 2006. Disponible em: <http://arxiv.org/abs/hep-ph/0512210>.
- [54] CACCIARI, M., SALAM, G. P., SOYEZ, G. “FastJet User Manual”, *Eur. Phys. J.*, v. C72, 2012. Disponible em: <http://arxiv.org/abs/1111.6097>.
- [55] ATLAS COLLABORATION. “Jet energy measurement with the ATLAS detector in proton-proton collisions at  $\sqrt{s} = 7$  TeV”, *Eur. Phys. J.*, v. C73, 2013. Disponible em: <http://arxiv.org/abs/1112.6426>.
- [56] ATLAS COLLABORATION. “Jet global sequential corrections with the ATLAS detector in proton-proton collisions at  $\sqrt{s} = 8$  TeV”, 2015.
- [57] ATLAS COLLABORATION. “Commissioning of the ATLAS high-performance  $b$ -tagging algorithms in the 7 TeV collision data”, 2011. Disponible em: <https://cds.cern.ch/record/1369219>.
- [58] ATLAS COLLABORATION. “Calibration of  $b$ -tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment”, 2014. Disponible em: <https://cds.cern.ch/record/1664335>.
- [59] ATLAS COLLABORATION. “Calibration of the performance of  $b$ -tagging for  $c$  and light-flavour jets in the 2012 ATLAS data”, 2014. Disponible em: <https://cds.cern.ch/record/1741020>.
- [60] MICONI, F. “Search for WH associated production at  $D\phi$  Tevatron”. In: *Proceedings, Symposium on Hadron Collider Physics Symposium 2011 (HCP 2011)*. Disponible em: <http://arxiv.org/abs/1201.6026>.
- [61] OLIVE, K. A., OTHERS. “Review of Particle Physics”, *Chin. Phys.*, v. C38, 2014.

- [62] BEGALLI, M., GASPAR, P., BUZATU, A., et al. “Impact of jet-energy corrections on the ATLAS Run I  $WH \rightarrow l\nu b\bar{b}$  search”. <https://cds.cern.ch/record/1982454/>, 2015.
- [63] HAYKIN, S. *Neural Networks: A Comprehensive Foundation*. Upper Saddle River, NJ, Prentice Hall, 1998.
- [64] MCCULLOCH, W., PITTS, W. “A logical calculus of the ideas immanent in nervous activity”, *Mathematical Biophysics*, v. 5, 1943.
- [65] NUNES DA SILVA, I., SPATTI, D. H., FLAUZINO, R. A. *Redes Neurais Artificiais - para engenharias e ciências aplicadas*. São Paulo - SP, Artliber, 2010.
- [66] RUMELHART, D. E., HINTON, G. E., WILLIAMS, R. J. “Neurocomputing: Foundations of Research”. MIT Press, cap. Learning Representations by Back-propagating Errors, Cambridge, MA, USA, 1988.
- [67] MARQUADT, D. “An Algorithm for Least-Squares Estimation of Nonlinear Parameters”, *Journal of the Society for Industrial and Applied Mathematics*, v. 11, 1963.
- [68] BISHOP, C. M. *Neural Networks for Pattern Recognition*. Oxford, Clarendon Press, 1995.
- [69] HAGAN, M. T., MENHAJ, M. B. “Training feedforward networks with the marquardt algorithm”, *Neural Networks, IEEE Transactions*, v. 5, 1994.
- [70] LIU, H. “On the levenberg-marquardt training method for feed-forward neural networks”, *Proceedings of the 2010 International Conference on Natural Computation*, v. 1, 2010.
- [71] SETIONO, R., HUI, L. C. K. “Use of Quasi-Newton Method in a Feedforward Neural Network Construction Algorithm”, *IEE Transactions of Neural Networks*, v. 6, 1995.
- [72] KOHAVI, R. “A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection”, *International Joint Conference on Artificial Intelligence*, 1995.
- [73] HAGEDORN, R., DAVID JACKSON, J., PINES, D. *Relativistic Kinematics: A Guide To The Kinematic Problems Of High Energy Physics*. Literary Licensing - LLC, 2012.

- [74] LIN, J. “Divergence measures based on the Shannon entropy”, *Information Theory - IEEE Transactions*, v. 37, 1991.
- [75] KULLBACK, S., LEIBLER, R. A. “On Information and Sufficiency”, *The Annals of Mathematical Statistics*, v. 22, 1951.
- [76] CHA, S., SRIHARI, S. N. “On measuring the distance between histograms”, *Pattern Recognition*, v. 35, 2002.
- [77] SHANNON, C. E. “A Mathematical Theory of Communication”, *The Bell System Technical Journal*, v. 27, 1948.
- [78] BUKIN, A. “Fitting function for asymmetric peaks”, 2007. Disponível em: <<http://arxiv.org/abs/0711.4449>>.
- [79] BERINGER, J., OTHERS. “Review of Particle Physics (RPP)”, *Phys. Rev.*, v. D86, 2012.
- [80] BROWN, G., WYATT, J. L., TIÑO, P. “Managing Diversity in Regression Ensembles”, *J. Mach. Learn. Res.*, v. 6, 2005.
- [81] ZHOU, Z.-H. *Ensemble Methods: Foundations and Algorithms*. Chapman & Hall/CRC, 2012.
- [82] ZHANG, C., MA, Y. *Ensemble Machine Learning: Methods and Applications*. Springer-Verlag, 2012.
- [83] SCHAPIRE, R., FREUND, Y. *Boosting: Foundations and Algorithms*. The MIT Press, 2012.
- [84] SCHAPIRE, R. “A Brief Introduction to Boosting”. In: *Proceedings of the 16th International Joint Conference on Artificial Intelligence - Volume 2*, 1999.
- [85] KUMMER, N., NAJJARAN, H. “Adaboost.MRT: Boosting regression for multivariate estimation”, *Artificial Intelligence Research*, v. 4, 2014.
- [86] CYBENKO, G. “Aproximation by Superpositions of a Sigmoidal Function”, *Mathematics of Control, Signals, and Systems*, v. 2, 1989.
- [87] MOURA JUNIOR, N. N. *Modelo Neural para Correção da Estimção de Energia para um Detector de Partículas de Propósito Geral*. Dissertação de Mestrado. Coppe/UFRJ, 2014. Disponível em: <<http://www.pee.ufrj.br/teses/index.php?Resumo=2014032103>>.