

UNIVERSIDADE FEDERAL DE SÃO JOÃO DEL REI  
DEPARTAMENTO DE CIÊNCIAS NATURAIS

Dominique Lopes Ramos

APLICAÇÃO DO APRENDIZADO DE MÁQUINA  
NA PROCURA POR MATÉRIA ESCURA NO LHC

São João del Rei - MG  
2021

São João del Rei - MG

2021

**Dominique Lopes Ramos**

APLICAÇÃO DO APRENDIZADO DE MÁQUINA NA  
PROCURA POR MATÉRIA ESCURA NO LHC

Dissertação apresentada ao Programa de Pós-Graduação em Física da Universidade Federal de São João del Rei, em associação ampla com a Universidade Federal de Alfenas, como requisito para a obtenção do título de Mestre em Física.

Área de concentração: Física de Partículas e Campos.

Orientadora: Profa. Dra. Maria Aline Barros do Vale

Co-orientador: Prof. Dr. Fernando Marroquim Leão de Almeida Júnior

São João del Rei - MG

Departamento de Ciências Naturais - UFSJ

2021

Ficha catalográfica elaborada pela Divisão de Biblioteca (DIBIB)  
e Núcleo de Tecnologia da Informação (NTINF) da UFSJ,  
com os dados fornecidos pelo(a) autor(a)

R175a Ramos, Dominique Lopes.  
Aplicação do Aprendizado de Máquina na Procura por  
Matéria Escura no LHC / Dominique Lopes Ramos ;  
orientadora Maria Aline Barros do Vale;  
coorientador Fernando Marroquim Leão de Almeida  
Junior. -- São João del-Rei, 2021.  
100 p.

Dissertação (Mestrado - Física) -- Universidade  
Federal de São João del-Rei, 2021.

1. Física de Partículas. 2. Matéria Escura. 3.  
Modelo ZP-TP-DM. 4. Aprendizado de Máquina. I.  
Barros do Vale, Maria Aline , orient. II. Marroquim  
Leão de Almeida Junior, Fernando, co-orient. III.  
Título.



MINISTÉRIO DA EDUCAÇÃO  
UNIVERSIDADE FEDERAL DE SÃO JOÃO DEL-REI  
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA

**NOMEAÇÃO DE BANCA EXAMINADORA Nº 20 / 2021 - PPGF (13.29)**

**Nº do Protocolo: 23122.027951/2021-88**

**São João del-Rei-MG, 11 de agosto de 2021.**

A dissertação de mestrado "**Aplicação do Aprendizado de Máquina na Procura por Matéria Escura no LHC**" elaborada por **Dominique Lopes Ramos** e aprovada por todos os membros da banca examinadora, foi aceita pelo Programa de Pós-graduação em Física da Universidade Federal de São João del-Rei como requisito parcial à obtenção do título de

## **MESTRE EM FÍSICA**

Assinado por concordância de acordo com ata enviada por email

Prof. Dr. Jean Carlos Coelho Felipe - UFVJM

*(Assinado digitalmente em 12/08/2021 10:09 )*

HERON CARLOS DE GODOY CALDAS  
PROFESSOR DO MAGISTERIO SUPERIOR  
DCNAT (12.12)  
Matricula: 1212928

*(Assinado digitalmente em 12/08/2021 11:25 )*

MARIA ALINE BARROS DO VALE  
PROFESSOR DO MAGISTERIO SUPERIOR  
DCNAT (12.12)  
Matricula: 364974

Para verificar a autenticidade deste documento entre em <https://sipac.ufsj.edu.br/public/documentos/index.jsp> informando seu número: **20**, ano: **2021**, tipo: **NOMEAÇÃO DE BANCA EXAMINADORA**, data de emissão: **11/08/2021** e o código de verificação: **110bfb5aff**

*Aos meus pais...*

# Agradecimentos

Primeiramente, agradeço a Deus por me sustentar e guiar em toda a minha jornada.

Agradeço aos meus pais, Lídia e Mauro, que sempre me apoiaram e incentivaram em todas as etapas da minha vida.

À minha orientadora Professora Doutora Maria Aline Barros do Vale e ao meu co-orientador Professor Doutor Fernando Marroquim Leão de Almeida Júnior, que me acompanharam ao longo do mestrado e com paciência me instruíram e ensinaram. Agradeço também a todos os professores que fizeram parte da minha formação no mestrado.

Ao meu professor da graduação, Luiz Paulo Colatto, agradeço por sempre ter me incentivado a prosseguir na carreira acadêmica.

Agradeço a todos os amigos e amigas que conquistei durante esses anos de mestrado. Sou grata aos jovens da Primeira Igreja Batista de São João del Rei que foram de extrema importância na minha vida. Em especial, agradeço à Beatriz Biasi, Felipe Vasconcellos, Thais Heringer e Thayane por todo suporte e amizade. Carregarei essas amizades no meu coração. Obrigada PIBSJ por ter me acolhido com tanto carinho. Agradeço também a minha amiga Dayana Cristine com quem pude desabafar em dias complicados e que me deu apoio em diversos momentos. À Aria, agradeço pelas várias dicas de programação que foram de grande ajuda e pelas boas conversas no laboratório que me alegravam. Muito obrigada ao meu amigo Ederson Neri com quem compartilhei pensamentos sobre os mais diversos assuntos.

Por fim, agradeço a UFSJ pelo apoio financeiro.

*“Be strong and courageous. Do not be afraid; do not be discouraged, for the Lord your God will be with you wherever you go.”*

– Joshua 1:9

*“SHINZOU WO SASAGEYO!”*

(“心臓を捧げよ!”)

– Erwin Smith

# Resumo

Apesar do Modelo Padrão (MP) ser capaz de explicar muitos fenômenos do mundo das partículas, existem muitos problemas ainda não esclarecidos por esse modelo. A natureza da matéria escura (ME), que compõe cerca de 27% de todo o Universo, é um enigma tanto do ponto de vista da Física de Partículas quanto da Cosmologia. A ME não interage com a matéria comum da forma como conhecemos, por isso não pode ser detectada nos aparatos experimentais atuais, sendo a percepção da sua existência devido apenas à sua interação gravitacional com a matéria ordinária. No Grande Colisor de Hádrons (LHC), a procura por ME está associada a sua assinatura caracterizada por energia transversa faltante ( $\cancel{E}_T$ ) em experimentos.

Diversas teorias além do MP consideram uma nova partícula candidata a ME que pode ser uma partícula escalar, vetorial, um férmion de Dirac ou Majorana. Neste trabalho, consideramos um candidato escalar  $\phi$  (com spin 0) para a ME prevista pelo modelo ZP-TP-DM que surge do decaimento de um novo bóson de calibre  $Z'$  (spin 1) em um par  $T'\bar{T}'$  (chamados parceiros top fermiônico com spin 1/2), dando origem ao estado final  $t\bar{t}\phi\phi$ .

A partir das técnicas de Aprendizado de Máquina (AM) separamos os eventos de fundo ( $pp \rightarrow t\bar{t}Z$ , com  $Z$  decaindo nos neutrinos e seus respectivos antineutrinos) dos eventos de sinal ( $pp \rightarrow T'\bar{T}' \rightarrow t\bar{t}\phi\phi$ ) gerados pelo CalcHEP.

Construímos uma rede neural profunda (RNP) para separar os eventos de fundo e de sinal e obtemos valores próximos a 1 para a área abaixo da curva Característica de Operação do Receptor (ROC - *Receiver Operating Characteristic Curve*) para diferentes massas. Isso indica que o classificador criado separou os eventos de forma eficiente. Entretanto, os resultados obtidos para a significância estatística representa uma situação ideal, já que não incluímos na análise o decaimento do par  $t\bar{t}$ .

**Palavras-chaves:** Física de Partículas, Matéria Escura, Modelo ZP-TP-DM, Aprendizado de Máquina.



# Abstract

Although the Standard Model (SM) is able to explain many phenomena in the world of particles, there are many problems that have not yet been clarified by this model. The nature of dark matter (DM), which makes up about 27 % of the entire Universe, is an enigma both from the point of view of Particle Physics and Cosmology. DM does not interact with ordinary matter as we know it, so it cannot be detected in current experimental apparatus, being the perception of its existence only due to its gravitational interaction with ordinary matter. In the Large Hadron Collider (LHC), the search for DM is associated with its signature characterized by missing transverse energy ( $\cancel{E}_T$ ) in experiments.

Several theories beyond the SM consider a new DM candidate particle that can be a scalar, vectorial, a fermion of Dirac or Majorana particle. In this work, we consider a scalar DM candidate  $\phi$  (with spin 0) predicted by ZP-TP-DM model that come from the decay of a new gauge boson  $Z'$  (spin 1) into a pair  $T'\bar{T}'$  (called top fermionic partners with spin 1/2), giving rise to the final state  $t\bar{t}\phi\phi$ .

Using Machine Learning (AM) techniques, we separate the background events ( $pp \rightarrow t\bar{t}Z$ , with  $Z$  decaying in neutrinos and their respective antineutrinos) from the signal events ( $pp \rightarrow T'\bar{T}' \rightarrow t\bar{t}\phi\phi$ ) generated by CalcHEP.

We built a deep neural network (DNN) to separate background and signal events and obtained values close to 1 for the area under the Receiver Operating Characteristic curve (ROC), indicating that the created classifier efficiently separated the events. However, the results obtained for statistical significance represent an ideal situation, since we did not include in the analysis the decay of the pair  $t\bar{t}$ .

**Keywords:** Particle Physics, Dark Matter, ZP-TP-DM Model, Machine Learning.

# Sumário

<b>Sumário</b>	<b>8</b>
<b>Lista de Figuras</b>	<b>10</b>
<b>Lista de Tabelas</b>	<b>14</b>
<b>1 Introdução</b>	<b>15</b>
<b>2 Modelo Padrão</b>	<b>18</b>
2.1 Problemas com o Modelo Padrão . . . . .	23
2.2 Extensões do Modelo Padrão . . . . .	24
<b>3 Matéria Escura (ME)</b>	<b>26</b>
3.1 Cosmologia Padrão . . . . .	26
3.2 Evidências da ME . . . . .	30
3.3 Propriedades da Partícula de ME . . . . .	34
3.4 Possíveis Candidatos para ME . . . . .	35
3.5 Fenomenologia da ME . . . . .	35
3.6 Métodos de Detecção e Processos de Interação da ME . . . . .	36
3.6.1 Detecção Direta (DD) . . . . .	37
3.6.2 Detecção Indireta (DI) . . . . .	38
3.6.3 Colisores . . . . .	40
<b>4 O Grande Colisor de Hádrons, Experimento ATLAS e Ferramentas de Simulação</b>	<b>41</b>
4.1 Centro Europeu de Pesquisas Nucleares . . . . .	41
4.2 LHC . . . . .	43

4.3	Detector ATLAS e Quantidades Transversas . . . . .	44
4.4	Procura por Matéria Escura no LHC . . . . .	48
4.5	Ferramentas de Simulação . . . . .	50
4.5.1	CalcHep . . . . .	50
4.5.2	DELPHES e PYTHIA . . . . .	51
<b>5</b>	<b>Modelo ZP-TP-DM</b>	<b>52</b>
5.1	Fenomenologia . . . . .	53
5.2	Eventos de Sinal e de Fundo . . . . .	55
5.3	Predição Teórica da Seção de Choque e Distribuições Cinemáticas . . . . .	55
<b>6</b>	<b>Aprendizado de Máquina</b>	<b>60</b>
6.1	Técnicas de AM . . . . .	62
6.1.1	Aprendizado Supervisionado . . . . .	62
6.1.2	Aprendizado Não-Supervisionado . . . . .	69
6.2	Curva ROC e AUC . . . . .	71
6.3	Linguagens de Programação . . . . .	72
<b>7</b>	<b>Análise dos Dados Simulados e Resultados</b>	<b>74</b>
7.1	Implementação do Modelo . . . . .	74
7.2	Eventos Simulados . . . . .	75
7.2.1	Principais Eventos de Fundo . . . . .	76
7.2.2	Distribuição de Variáveis para os Eventos de Sinal e de Fundo . . . . .	77
7.3	Treinamento e Teste da Rede . . . . .	79
7.4	Resultados . . . . .	81
7.5	Significância do sinal . . . . .	86
<b>8</b>	<b>Conclusões</b>	<b>91</b>
	<b>Referências Bibliográficas</b>	<b>93</b>

# Lista de Figuras

3.1	O gráfico descreve o tamanho relativo do Universo, distância média entre as Galáxias, durante os seus bilhões de anos sob diferentes cenários sobre como o espaço e a matéria evoluem com o tempo. Ao lado de cada curva está a representação geométrica do Universo: Universo aberto, plano e fechado. O tempo presente destacado em rosa equivale a aproximadamente 13,8 bilhões de anos. . . . .	28
3.2	Curvas das velocidades rotacionais como função da distância radial para sete galáxias. . . . .	32
3.3	Lentes gravitacionais. (a) Imagem obtida por telescópio do Aglomerado Abell 2218 onde podemos ver os arcos gravitacionais formados devido a geometria curva do espaço pela presença de corpos super massivos. (b) Exemplo esquemático de como é vista a imagem de um objeto por um observador na Terra . . . . .	34
3.4	Possíveis canais de detecção de matéria escura. (a)Detecção direta. (b)Detecção indireta. (c)Colisores. . . . .	36
3.5	Resultados experimentais de detecção direta e limites de exclusão para seções de choque de interações WIMP-núcleon independente do spin. Os contornos fechados são indicações do sinal de ME. As curvas delimitam os limites de exclusão do espaço de parâmetros acima. . . . .	38
3.6	Predições da seção de choque de aniquilação da ME em função da massa da matéria escura ( $M_\phi$ ). Estão sendo comparados os limites de experimentos DI, Fermi-LAT e H.E.S.S., e do experimento DD LUX para os canais $\phi\phi \rightarrow b\bar{b}$ e $\phi\phi \rightarrow Z\gamma$ , através de observações de galáxias e Centro do Halo Galáctico. . . . .	39
4.1	Complexo de aceleradores no CERN. . . . .	42
4.2	Esquema 3D do ATLAS com as indicações do momento transversal $\mathbf{p}_T$ e momento linear $\mathbf{p}$ . . . . .	45

4.3	É mostrado um evento no detector ATLAS de um experimento de busca por matéria escura. O momento transverso (barra amarela) é balanceado pela energia transversa faltante (linha vermelha tracejada).	47
4.4	Alguns exemplos dos diagramas de Feynman dos processos estudados nas pesquisas de ME.	49
5.1	Diagramas de Feynman para a produção de $t\bar{t}\phi\phi$ pelo decaimento de $T'\bar{T}'$ vindo de (a) $Z'$ e (b e c) dos glúons retirado de	53
5.2	Seções de choque teóricas do modelo para diferentes valores de $M_{Z'}$ e $M_{T'}$ .	56
5.3	Distribuições normalizadas da fração do número de eventos em função da pseudo-rapidez do quark top (a) e pseudo-rapidez do antiquark top (b).	57
5.4	Distribuições normalizadas da fração do número de eventos na escala logarítmica em função do momento transverso do quark top (a) e momento transverso do antiquark top (b).	58
5.5	Distribuições normalizadas da fração do número de eventos na escala logarítmica em função da massa transversa do par $t\bar{t}$ (a) e energia transversa faltante (b).	59
6.1	Exemplo da Classificação de spam em e-mails.	62
6.2	Regressão: dada uma <b>característica 1</b> , a partir dos dados (informação sobre os valores - eixo y), é possível prever eventos futuros estatisticamente prováveis sobre o valor que a <b>característica 1</b> assumirá.	63
6.3	O esquema da figura compreende em três árvores de decisão treinadas independentemente.	64
6.4	Arquitetura da RNA.	65
6.5	Esquema de um recorte de uma RNA, com $N$ neurônios de entrada e $N$ neurônios na primeira camada escondida, que ilustra como cada neurônio se conecta matematicamente. O neurônio ativado (em azul) recebe todas as informações dos pesos ( $W$ ) e recursos ( $X$ ) de cada neurônio da camada anterior, somado a um valor do <i>bias</i> ( $b_0$ ) e multiplicando por uma função de ativação $\varphi$ .	66
6.6	Funções de Ativação: a curva roxa é a função ReLU, a curva azul é a função de ativação $\sigma$ e a curva laranja é a função tangente hiperbólica.	67
6.7	Gráfico de perda da rede.	69
6.8	Clustering: agrupamento de características semelhantes num domínio bidimensional.	70

6.9	Redução de Dimensionalidade: a partir de uma transformação $\varphi$ adequada, os algoritmos de RD tentam representar dados com dimensão alta ( $x_1 \times x_2$ ) em dados com dimensão baixa ( $\xi_1$ ). . . . .	70
6.10	Detecção de Anomalias. . . . .	71
6.11	Curva ROC: A curva em vermelho caracteriza a eficiência do classificador criado e a linha tracejada em azul é uma curva padrão que delimita a eficiência geral de classificadores. Quanto mais afastadas as curvas ROC são da linha tracejada (área abaixo da curva $\approx 1$ ), maior a probabilidade de que você tenha criado um bom classificador. . . . .	72
7.1	Produção do par $T'\bar{T}'$ em pares $t\bar{t}$ e duas partículas de ME ( $\phi\phi$ ) no estado final vindo do decaimento do bóson $Z'$ (a) e glúons (b e c). . . . .	78
7.2	Comparação entre as distribuições normalizadas da energia transversa faltante (a) e massa transversa (b) dos eventos gerados pelo CalcHEP vindos do Modelo Padrão (linha azul) e do modelo ZP-TP-DM para $M_{Z'} = 2500GeV$ e $M_{T'} = 1250GeV$ (linha verde), e $M_{Z'} = 3000GeV$ , $M_{T'} = 1500GeV$ (linha vermelha). . . . .	79
7.3	Número de eventos na escala logarítmica em função da saída no neurônio r. Nas figuras é apresentado as classificações da rede para os eventos vindos do Modelo Padrão ( $pp \rightarrow t\bar{t}Z$ ) e do modelo ZP-TP-DM ( $pp \rightarrow T'\bar{T}' \rightarrow t\bar{t}\phi\phi$ ) para (a) $M_{Z'} = 2500GeV$ , $M_{T'} = 1250GeV$ e $M_\phi = 300GeV$ e (b) $M_{Z'} = 3000GeV$ , $M_{T'} = 1500GeV$ e $M_\phi = 300GeV$ . Os eventos de fundo correspondem à área laranja e os eventos de sinal à azul. . . . .	82
7.4	Em (a) e (b) temos a saída da rede referente à perda e à acurácia, respectivamente, para o conjunto <b>A</b> . Também é apresentado a perda (c) e acurácia (d) para o conjunto <b>B</b> do nosso modelo relacionado à Figura 7.3 com as curvas dos dados de treinamento (linha azul) da rede e de teste (linha laranja). . . . .	83
7.5	Curva ROC para o classificador criado (linha laranja) – referente à Figura 7.3 – considerando os eventos do sinal para (a) $M_{Z'} = 2500GeV$ , $M_{T'} = 1250GeV$ e $M_\phi = 300GeV$ e (b) $M_{Z'} = 3000GeV$ , $M_{T'} = 1500GeV$ e $M_\phi = 300GeV$ . O ponto denominado por "Best" é o melhor corte, isto é, indica o ponto que tem a maior taxa de verdadeiros positivos junto com a menor taxa de falsos positivos. Abaixo da linha tracejada tem AUC=0.5. . . . .	84

7.6	Número de eventos na escala logarítmica em função da saída no neurônio r. Nas figuras são apresentadas as mesmas classificações da figura 7.3 dos eventos vindos do Modelo Padrão ( $pp \rightarrow t\bar{t}Z$ ) e do modelo ZP-TP-DM ( $pp \rightarrow T'\bar{T}' \rightarrow t\bar{t}\phi\phi$ ), para as massas (a) $M_{Z'} = 2500GeV$ , $M_{T'} = 1250GeV$ e $M_\phi = 300GeV$ e (b) $M_{Z'} = 3000GeV$ , $M_{T'} = 1500GeV$ e $M_\phi = 300GeV$ . O melhor corte do número de eventos está indicado pela linha preta equivalente a 0.25 em (a) e 0.22 em (b). . . . .	86
7.7	Variação dos valores AUC de acordo com várias massas da ME para $M_{Z'} = 2500GeV$ e $M_{T'} = 1250GeV$ ; $M_{Z'} = 3000GeV$ e $M_{T'} = 1500GeV$ . O valores do número de épocas foram escolhidos de modo a evitar o <i>overfitting</i> . . . . .	86
7.8	Comparação da variação dos valores da significância do sinal de acordo com a massa $M_\phi$ para SMA1 (em verde e azul) e SMA2 (em vermelho e amarelo). . . . .	90

# Lista de Tabelas

2.1	Gerações de partículas fundamentais do MP. . . . .	19
2.2	Interações da natureza e seus respectivos mediadores. . . . .	19
3.1	Classificação de modelos cosmológicos baseados no valor da densidade total $\rho_{tot}$ , em termos da densidade crítica $\rho_c$ . . . . .	27
7.1	Canais de produção utilizados na referência . . . . .	77
7.2	Conjunto de massas utilizados para separar os eventos. . . . .	81
7.3	Valores para AUC dos conjunto de massas <b>A</b> e <b>B</b> , respectivamente. . . . .	83
7.4	Número de eventos de acordo com os parâmetros vistos no Capítulo 6, seção 6.2, para um total de 3245 eventos de teste (sinal+fundo). . . . .	84
7.5	Número de eventos calculados pela rede para FP, VP, VN e FN dos conjuntos de massas <b>A</b> e <b>B</b> . . . . .	85
7.6	Valores-p para sua correspondente significância de 1 a 6 $\sigma$ (unidade de desvio padrão) . . . . .	88



# Capítulo 1

## Introdução

A existência de matéria escura (ME) [1] no Universo, mesmo que evidenciada por diversas observações cosmológicas a partir de objetos luminosos (que emitem radiação eletromagnética como estrelas, nuvens de gases, aglomerados galácticos, etc), tem natureza que permanece desconhecida.

Apesar de não existir experimentos que detectem uma partícula de ME, existem métodos que visam a detecção da assinatura de ME (sua provável presença) devido à interação de ME com partículas do MP. Os métodos de detecção da assinatura de ME se concentram em três diferentes formas: Experimentos de Detecção Direta (DD) [2], Indireta (DI) [3] e em Colisores [4]. A DD consiste em experimentos do recuo nuclear, onde uma partícula conhecida, ao interagir com alguma partícula de ME, é espalhada. Técnicas de DI usam observações astronômicas das partículas que conhecemos para detectar os produtos de aniquilação ou decaimento de ME na nossa Galáxia e por todo o Cosmo.

O modelo cosmológico padrão (MCP) [5] descreve alguns elementos baseados na origem e evolução do Universo que nos ajudam a compreender algumas características da matéria escura. Por exemplo, o MCP considera que a ME deva ser neutra por não interagir com fótons e que deva ser estável, porque tem a idade superior à evolução do Universo.

Uma das maiores questões atuais na Física de Partículas é, portanto, a busca por uma partícula candidata à matéria escura. Ainda não existe nenhum aparato experimental que detecte ME, já que ela não interage com partículas do Modelo Padrão (MP) [6] da maneira convencional, pois não emite radiação eletromagnética. Mas, buscas em aceleradores de partículas têm realizado muitas pesquisas promissoras.

O MP é constituído por um conjunto de teorias que descrevem as interações entre

as partículas elementares, com exceção da interação gravitacional. Além disso, o MP é considerado a teoria matemática mais sofisticada sobre a natureza, é bem definido e possui inúmeras confirmações experimentais. Apesar do MP ser capaz de explicar muitos fenômenos do mundo das partículas, existem diversos problemas ainda não esclarecidos por esse modelo. A natureza da ME, que compõe cerca de 27% de todo o Universo, é um enigma tanto do ponto de vista da Física de Partículas quanto da Cosmologia.

Como a ME não interage com a matéria comum do modo conhecido, a percepção da sua existência é devido apenas à sua interação gravitacional com a matéria ordinária<sup>1</sup>. Como partículas de ME são invisíveis aos detectores, a presença de tais partículas e de todas não detectáveis deve ser inferida a partir do desequilíbrio do momento total em um evento de colisão [7]. No Grande Colisor de Hádrons (LHC), localizado no Centro Europeu de Pesquisas Nucleares (CERN), a procura por ME está associada a esse desequilíbrio do momento vetorial no plano transversal que é o plano perpendicular à direção de um feixe de partículas próton-próton ( $pp$ ). De acordo com a conservação do momento, pode-se inferir esse desequilíbrio que é chamado de momento transversal faltante ( $\cancel{p}_T$ ). Nesta dissertação utilizaremos a energia transversal faltante ( $\cancel{E}_T$ ) que é o desequilíbrio de energia no plano transversal ao feixe de partículas.

Diversas teorias além do MP consideram uma nova partícula candidata a ME que pode ser uma partícula escalar, vetorial, um férmion de Dirac ou Majorana. Nesse trabalho, consideramos um candidato escalar para a ME previsto no modelo de extensão do modelo padrão chamado ZP-TP-DM<sup>2</sup> [8]. Se a ME é produzida nos aceleradores, acreditamos que sua produção contribua para a  $\cancel{E}_T$ , pois é invisível ao detector. Para a análise da produção da assinatura  $t\bar{t} + \cancel{E}_T$  no LHC, através de simulações rápidas, de acordo com o trabalho [9] da *Colaboração ATLAS*, consideramos a luminosidade integrada igual a  $13.2fb^{-1}$ <sup>3</sup> e energia do centro de massa  $13TeV$ <sup>4</sup>. Utilizamos o Calchep [10], um software para simulações de eventos, com o objetivo de calcular as seções de choque e gerar eventos de sinal (eventos oriundos do modelo ZP-TP-DM) e de fundo (eventos que conhecemos

---

<sup>1</sup> Toda matéria a qual conhecemos.

<sup>2</sup> ZP é abreviação para *Z Prime* e representa o bóson de calibre  $Z'$ ; TP é abreviação para *T Prime*, é o chamado top fermiônico  $T'$ ; DM vem da abreviação do inglês *dark matter* e corresponde a partícula  $\phi$  de matéria escura. Todas são novas partículas previstas pelo modelo ZP-TP-DM.

<sup>3</sup> O *barn* é uma unidade métrica de área utilizada para expressar as seções de choque de qualquer processo de espalhamento em colisões de partículas. 1 *barn* equivale a  $10^{-28}m^2$ . Geralmente utiliza-se o *femtobarn* (fb) que é  $10^{-43}m^2$ . Assim, 1 *femtobarn*<sup>-1</sup> ( $fb^{-1}$ ) =  $10^{43}m^{-2}$ .

<sup>4</sup> 1 Tera elétrons-volt ( $TeV$ ) =  $10^{12}$  elétrons-volt.

vindos do MP).

Em Física de Partículas é comum o uso de cortes retangulares de variáveis para separar os eventos de sinal dos eventos de fundo. Esse é um método tradicional e tem como característica eliminar o máximo do fundo, mantendo o máximo do sinal. Porém, abordamos uma nova estratégia para a separação a partir das técnicas de Aprendizado de Máquina (AM) [11].

O AM é uma subcategoria de Inteligência Artificial (IA)<sup>5</sup> [12] que possui algoritmos cuja eficácia melhora à medida que são expostos a mais dados. Dentro do AM ainda há diversas técnicas que podem ser aplicadas de acordo com o interesse de pesquisa. Para separar os principais eventos de fundo vindos do MP escolhidos ( $pp \rightarrow t\bar{t}Z$ ) do sinal ( $pp \rightarrow T'\bar{T}' \rightarrow t\bar{t}\phi\phi$ , onde o par  $T'\bar{T}'$  e  $\phi$  são novas partículas previstas pelo modelo ZP-TP-DM), utilizamos métodos de AM: o Aprendizado Supervisionado (AS) [13] que tem como propósito fazer com que um programa de computador tenha a capacidade de aprender, comparando sua saída atual com as saídas ensinadas. Para isso, construímos uma rede neural com aprendizagem profunda. A Rede Neural Profunda (RNP) [14] possui algoritmos baseados em redes neurais (rede com várias camadas compostas por neurônios interligados, uma alusão ao cérebro humano) que aprendem a partir de enormes quantidades de dados. A RNP construída tem como finalidade separar os eventos de fundo e de sinal os classificando em categorias (0, 1) para sinal e (1, 0) para fundo.

A organização dessa dissertação foi feita da seguinte maneira: no Capítulo 2, é apresentada uma explicação resumida do MP e questões em aberto; no Capítulo 3, é fornecida uma contextualização do cenário da ME na Cosmologia e em Física de Partículas; O Capítulo 4, descreve o CERN, o LHC, o experimento ATLAS e as ferramentas de simulação; O Capítulo 5 apresenta a descrição do modelo ZP-TP-DM, que é uma extensão do MP; No Capítulo 6, são detalhadas técnicas de AM, método que utilizamos para saber se construímos uma bom classificador de eventos e as linguagem de programação que podem ser utilizadas em AM; O Capítulo 7 mostra uma análise dos dados simulados e os resultados obtidos; No Capítulo 8 são apresentadas as conclusões.

---

<sup>5</sup>Caracterizada por qualquer técnica que permita os computadores imitarem a inteligência humana.

# Capítulo 2

## Modelo Padrão

Desde a Antiguidade, quando o átomo era tratado como algo indivisível, até nos dias atuais, vivenciamos um grande avanço do conhecimento científico que levou à muitas descobertas na área de Física de Partículas. Os átomos foram considerados elementos sem estrutura interna por muito tempo (até o final do século XIX) [15], mas pela incessante busca da compreensão acerca de como o Universo físico funciona e como ele é formado, hoje sabemos que os constituintes mais elementares da matéria são muito menores do que o átomo.

Especificamente, as contribuições vindas do desenvolvimento na área da Eletricidade foram as que mais colaboraram para a conclusão de que o átomo é, na verdade, divisível. Logo, com a ideia de um núcleo atômico composto (prótons, nêutrons e elétrons), foram desenvolvidas novas tecnologias que consistem em acelerar partículas a altas energias com o objetivo de entender as forças interatômicas e descobrir novas partículas. Para isso, os físicos da área de Física de Partículas trabalharam e continuam trabalhando com o objetivo principal de identificar novas partículas e entender a interação existente entre elas.

Atualmente, temos um conjunto de teorias que descrevem as partículas elementares conhecidas e suas interações (com exceção de uma possível partícula mediadora da interação gravitacional que, até o momento, não foi identificada experimentalmente), e essa descrição se dá pelo modelo padrão da Física de Partículas (MP) [16] [17] [18]. O MP é um exemplo concreto da Teoria Quântica de Campos (TQC), que incorpora todas as partículas subatômicas que conhecemos e suas interações. O MP, considerado a teoria matemática mais sofisticada sobre a natureza, é bem definido e possui inúmeras confirmações

experimentais.

As partículas fundamentais podem ser classificadas em dois grupos: férmions e bósons. Os férmions (Tabela 2.1) são classificados em léptons e quarks que formam toda a matéria conhecida. Elétrons, múons, táons e seus respectivos neutrinos compõem a categoria dos léptons. A matéria composta por quarks é chamada matéria hadrônica, onde os hádrons são divididos em bárions (formados por três quarks de valência) e mésons (formados por um quark e um antiquark de valência).

<b>Férmions</b>	<b>1ª Geração</b>	<b>2ª Geração</b>	<b>3ª Geração</b>
quarks	<i>up</i> (u)	<i>charm</i> (c)	<i>top</i> (t)
	<i>down</i> (d)	<i>strange</i> (s)	<i>bottom</i> (b)
léptons	<i>elétron</i> (e)	<i>múon</i> ( $\mu$ )	<i>táon</i> ( $\tau$ )
	<i>neutrino do elétron</i> ( $\nu_e$ )	<i>neutrino do múon</i> ( $\nu_\mu$ )	<i>neutrino do táon</i> ( $\nu_\tau$ )

Tabela 2.1: Gerações de partículas fundamentais do MP.

Na Tabela 2.1 mostramos as divisões das partículas fundamentais do MP em gerações. Existem três gerações de quarks e três gerações de léptons: as partículas das gerações mais altas possuem maior massa e acabam decaindo rapidamente para as partículas da primeira geração, em um estado mais estável.

Os bósons de calibre (também chamadas partículas virtuais) são os transportadores de força, que mediam as interações fundamentais. Como mostra a Tabela 2.2, a interação eletromagnética pelos fótons  $\gamma$ , a interação fraca pelos bósons de calibre  $Z$ ,  $W^+$  e  $W^-$  e a interação forte é transportada pelos glúons.

Os léptons podem sofrer interações fracas ou eletromagnéticas. Partículas como os neutrinos ( $\nu_e$ ,  $\nu_\mu$  e  $\nu_\tau$ ) sofrem apenas interação fraca, pois não possuem carga elétrica. Porém, as partículas com carga ( $e$ ,  $\mu$  e  $\tau$ ) podem sofrer tanto interações fracas quanto eletromagnéticas. Os quarks, além de sofrerem interações fracas e eletromagnéticas, sofrem interações fortes.

<b>Interações</b>	<b>Mediadores (bósons de calibre)</b>
Interação Eletromagnética	$\gamma$
Interação Fraca	$W^{+-}$ , $Z$
Interação Forte	$g$ (glúons)

Tabela 2.2: Interações da natureza e seus respectivos mediadores.

Os bósons e os férmions possuem comportamentos físicos bem diferentes. Enquanto os férmions possuem spin semi-inteiro ( $1/2$ ) e satisfazem o princípio de exclusão de Pauli,

que diz que dois férmions idênticos não podem ocupar o mesmo estado quântico, os bósons de calibre não satisfazem esse princípio e possuem spin inteiro (1).

O MP é uma teoria baseada em duas classes de simetria [19]: a simetria de Poincaré em 4 dimensões e simetria interna que é a simetria de calibre do grupo  $G_{SM} = SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ , onde o grupo  $SU(3)_C$  representa as interações fortes e o subíndice C nos diz que esse é o grupo de cor da Cromodinâmica Quântica (QCD - *Quantum Chromodynamics*). O  $SU(2)_L \otimes U(1)_Y$ , portanto, se trata da interação eletrofraca (EW - *eletroweak*). A paridade<sup>1</sup> no MP não é preservada, portanto, o subíndice L representa a quiralidade<sup>2</sup> do grupo  $SU(2)_L$ . Isso quer dizer que há diferenças entre partículas de quiralidade de mão esquerda (*left-handed* - LH) e de quiralidade de mão direita (*right-handed* - RH). A Teoria da Eletrodinâmica Quântica (QED - *Quantum Eledrodynamics*) é representada pelo grupo de simetria U(1) e seu subíndice Y indica o gerador de hipercarga do grupo.

No MP os quarks e léptons LH vêm em pares ou dubletos<sup>3</sup>, como são chamados, sob a transformação do grupo  $SU(2)$  e são representados, respectivamente, em ordem crescente de massa das três gerações como:

$$\begin{pmatrix} u \\ d \end{pmatrix}_L \quad \begin{pmatrix} c \\ s \end{pmatrix}_L \quad \begin{pmatrix} t \\ b \end{pmatrix}_L \quad (2.1)$$

$$\begin{pmatrix} e \\ \nu_e \end{pmatrix}_L \quad \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix}_L \quad \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix}_L \quad (2.2)$$

Os quarks RH ( $u_R$ ,  $c_R$ ,  $t_R$ ,  $d_R$ ,  $s_R$ ,  $b_R$ ) e léptons RH ( $e_R$ ,  $\mu_R$  e  $\tau_R$ ) são singletos.

Na teoria eletrofraca a carga elétrica Q está relacionada a hipercarga Y e é dada pela relação Gell-Mann-Nishijima:

$$Q = I_3 + \frac{Y}{2}, \quad (2.3)$$

---

<sup>1</sup> Paridade: no princípio de conservação da paridade (simetria P), as leis físicas expressam-se de formas idênticas em um sistema de coordenadas levogiro ou dextrogiro. No MP essa simetria não é preservada.

<sup>2</sup> A quiralidade de uma partícula pode ser apresentada em termos de como ocorre a sua transformação em uma representação LH ou RH do grupo de Poincaré.

<sup>3</sup> As famílias de partículas semelhantes são conhecidas como multipletos isospin: famílias de duas partículas são chamadas de dubletos, famílias de três partículas são chamadas de tripletos e assim por diante.

onde  $I_3$  é um isospin<sup>4</sup>.

A simetria de calibre impede um termo de massa para os bósons  $Z$ ,  $W^+$  e  $W^-$  que deve ser de alguma forma quebrada, para que esses bósons de calibre adquiram massa. Isso pode ser feito através do mecanismo de quebra de simetria espontânea. Os férmions do MP também seriam forçados a não ter massa sob  $SU(2)_L \otimes U(1)_Y$ . A quebra espontânea da simetria de calibre resulta na geração de massa para os bósons  $Z$ ,  $W^+$  e  $W^-$ , e para um campo escalar com massa, o campo de Higgs<sup>5</sup>, que também é responsável pela atribuição de massa para os quarks e léptons devido a interação dessas partículas com esse campo escalar [20].

consistente com o MP e é a partícula de spin zero que quebra espontaneamente a simetria de  $SU(2)_L \otimes U(1)_Y$ .

Pela quebra espontânea da simetria eletrofraca,

$$SU(2)_L \otimes U(1)_Y \longrightarrow U(1)_Q, \quad (2.4)$$

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \longrightarrow SU(3)_c \otimes U(1)_Q, \quad (2.5)$$

onde  $Q$  denota o gerador de carga elétrica da equação apresentada em 2.3.

A forma como as partículas do MP interagem entre si é representada por uma Lagrangiana, descrita em termos de férmions e bósons que se relacionam sob influência das três interações citadas anteriormente: a força eletromagnética, a força forte e a força fraca. A Lagrangiana que escreve as interações entre os bósons e férmions pode ser escrita como,

$$\mathcal{L}_{MP} = \mathcal{L}_{QCD} + \mathcal{L}_{eletrofraca}. \quad (2.6)$$

De acordo com [21] o termo  $\mathcal{L}_{QCD}$  é a Lagrangiana da cromodinâmica quântica e que é dada por:

$$\mathcal{L}_{QCD} = \bar{\Psi}(i\gamma^\mu\partial_\mu - m)\Psi + g_s\bar{\Psi}\gamma^\mu\frac{\lambda_a}{2}\Psi G_\mu^a - \frac{1}{4}G_{\mu\nu}^a G^{a,\mu\nu}, \quad (2.7)$$

onde o primeiro termo  $\bar{\Psi}(i\gamma^\mu\partial_\mu - m)\Psi$  contém termos cinéticos e de massa. O segundo e terceiro são os termos de interação, onde  $\Psi$  é um tripleto de campos fermiônicos que

---

<sup>4</sup> Isospin é uma propriedade que caracteriza famílias de partículas relacionadas que diferem nos valores de sua carga elétrica. Por exemplo, os prótons e nêutrons formam um núcleon dubleto de isospin. Numa colisão de partículas o isospin é conservado.

<sup>5</sup> A descoberta do bóson de Higgs no LHC (Large Hadron Collider) permitiu a comprovação experimental da previsão teórica da existência do bóson de Higgs.

descreve os quarks (u, d, ..., t) de massa  $m$  e  $\bar{\Psi}$  descreve os antiquarks de massa  $m$ . Os campos de quarks, como também são chamados, podem ser escritos como:

$$\Psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \end{pmatrix} \quad e \quad \bar{\Psi} = \begin{pmatrix} \bar{\psi}_1 \\ \bar{\psi}_2 \\ \bar{\psi}_3 \end{pmatrix} \quad (2.8)$$

Uma representação dos geradores infinitesimais do grupo unitário  $SU(3)$  são as matrizes de Gell-Mann [22] denotadas por  $\lambda_a$  presente na equação 2.7, onde  $a = 1, 2, \dots, 8$ . Os  $G_{\mu\nu}^a$  são os tensores de campo de força para os campos de glúon  $G_\mu^a$ . As matrizes de Dirac  $4 \times 4$ ,  $\gamma^\mu$  ( $\mu = 0, 1, 2, 3$ ), são descritas em termos das matrizes de spin de Pauli:

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \gamma^k = \begin{pmatrix} 0 & \sigma_k \\ -\sigma_k & 0 \end{pmatrix}, \quad (2.9)$$

onde  $\sigma_k$ , para  $k = 1, 2$  e  $3$ , corresponde às matrizes:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (2.10)$$

A constante de acoplamento da interação forte é representada por  $g_s$ . Como o grupo  $SU(3)$  é um grupo de oito parâmetros (os glúons), ele tem oito geradores  $T_a$ . No qual,  $T_a$  é representado por:

$$T_a = \frac{\lambda_a}{2} \quad (2.11)$$

A Lagrangiana da interação eletrofraca  $\mathcal{L}_{eletrofraca}$ , ainda de acordo com [21], é dada por:

$$\mathcal{L}_{eletrofraca} = \mathcal{L}_{calibre} + \mathcal{L}_{férmions} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yukawa}. \quad (2.12)$$

O primeiro termo  $\mathcal{L}_{calibre}$  é o que descreve as partículas responsáveis pelas interações, os bósons de calibre. Podemos escrevê-la como:

$$\mathcal{L}_{calibre} = -\frac{1}{4}W_{\mu\nu}^k W^{\mu\nu,k} - \frac{1}{4}B_{\mu\nu} B^{\mu\nu}, \quad (2.13)$$

onde  $W_{\mu\nu}^k$  e  $B_{\mu\nu}$  são tensores de campos de força que possuem índices de Lorentz  $\mu$  e  $\nu$ , que variam de 0 a 3 e estão associados aos campos vetoriais  $W_\mu^k$  (onde  $k = 1, 2, 3$ ) e



$B_\mu$  que são, respectivamente, os campos de calibre dos grupos SU(2) e U(1). Os campos vetoriais descrevem partículas de spin igual a 1. Termos de massa para os campos de calibre são proibidos devido à violação da invariância de calibre.

O setor dos férmions<sup>6</sup> em 2.12 é descrito por:

$$\mathcal{L}_{f\acute{e}rmions} = \sum_j \bar{\psi}_L^j i\gamma^\mu D_\mu^L \psi_L^j + \sum_{j,\sigma} \bar{\psi}_{R\sigma}^j i\gamma^\mu D_\mu^R \psi_{R\sigma}^j. \quad (2.14)$$

Esse termo descreve as partículas de matéria<sup>7</sup>, mas não contém os termos de massa.  $D_\mu$  são derivadas covariantes e o índice  $\sigma$  que equivale a + ou - denota férmions do tipo + (up) e - (down). O índice  $j$  passa pelas três gerações de léptons e quarks.

A Lagrangiana do setor de Higgs é dada por:

$$\mathcal{L}_{Higgs} = (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi), \quad (2.15)$$

onde são acrescentados campos escalares em  $\mathcal{L}_{Higgs}$  e os termos de massa para os bósons são obtidos.  $\Phi$  são dubletos de campos escalares  $\phi$ . Campos escalares descrevem partículas que possuem spin igual a 0.

É a partir da Lagrangiana  $\mathcal{L}_{Yukawa}$  que os termos de massa para os férmions são obtidos e o acoplamento do bóson de Higgs aos férmions é permitido. Essa Lagrangiana da interação dos campos de Higgs com os férmions é dada por:

$$\mathcal{L}_{Yukawa} = - \sum_f m_f \bar{\psi}_f \psi_f - \sum_f \frac{m_f}{v} \bar{\psi}_f \psi_f H, \quad (2.16)$$

onde  $m_f$  é a massa para os férmions carregados,  $\psi_f$  e  $\bar{\psi}_f$  são campos fermiônicos para os férmions carregados,  $v$  é corresponde ao valor do vácuo esperado do campo de Higgs e  $H$  é o campo de Higgs.

## 2.1 Problemas com o Modelo Padrão

Apesar do MP ser capaz de explicar muitos fenômenos do mundo das partículas, existem muitos problemas não explicados por esse modelo que ainda precisam ser esclarecidos [19]. Dentre eles, podemos citar:

<sup>6</sup> As representações são diferentes para férmions com diferentes quiralidades.

<sup>7</sup> Que são os léptons e os quarks.

- **Gravitação:** a interação gravitacional é descrita apenas em nível clássico e não pelas teorias de calibre onde os campos são quantificados. A partícula que seria mediadora da interação gravitacional, o gráviton (ainda não observado experimentalmente), corresponde a uma partícula sem massa de spin  $s = 2$  e não há uma TQC renormalizável que descreva essas partículas a distâncias arbitrariamente pequenas.
- **Assimetria Matéria/Antimatéria:** a matéria comum no Universo corresponde cerca de 5% de sua massa total. Junto com o nascimento do Cosmo, essa quantidade de matéria foi criada e a mesma quantidade de antimatéria também deveria ter sido criada. Mas, existe uma abundância de matéria comum sobre a antimatéria que observamos no Universo atual e não existe nenhum mecanismo que explique essa assimetria no MP.
- **Matéria Escura:** a Matéria Escura (ME) é um tipo de matéria que não interage com a matéria comum (ou, no caso de interação, interage fracamente) e, talvez, nem com ela mesma. De tudo o que conhecemos em termos de massa e energia, aproximadamente 27% é matéria escura. Ela só interage gravitacionalmente podendo ser evidenciada pelos efeitos gravitacionais sobre a matéria visível. A ME não absorve nem emite nenhuma luz, não interagindo, portanto, com partículas do MP da forma como conhecemos usualmente. Desta forma, não há nenhum candidato direto para a ME até o momento.

Em meio a tantas questões para serem resolvidas, extensões do MP foram elaboradas com a finalidade de se obter uma melhor compreensão desses e de outros fenômenos existentes além da física do MP. Dois exemplos de extensões do Modelo Padrão são modelos supersimétricos e dimensões extras [23]. Esses dois modelos incluem um possível candidato para a matéria escura, como abordaremos a seguir.

## 2.2 Extensões do Modelo Padrão

Extensões do MP têm como finalidade resolver problemas que o MP não explica.

Modelos supersimétricos, por exemplo, fornecem possíveis candidatos para a matéria escura. um candidato para a ME nesses modelos seria uma superpartícula mais leve (lightest superparticle LSP) e estável. Modelos desse tipo também têm um papel im-

portante no entendimento do problema de hierarquia<sup>8</sup>, unificação e na diferença entre férmions e bósons. Extensões supersimétricas supõem um parceiro supersimétrico para cada partícula do MP.

Modelos de dimensões extras desenvolvem teorias com dimensões extras espaciais no intuito de encontrar uma teoria fundamental com uma descrição unificada de todas as interações.

A teoria de dimensões extras de Kaluza-Klein, por exemplo, unifica a interação gravitacional e a eletromagnética, supondo um Universo com cinco dimensões.

Em teorias com dimensões extras unificadas, a partícula de Kaluza-Klein é a mais leve de todos os estados correspondentes às primeiras excitações das partículas do MP e é um candidato viável a ME.

Nos últimos anos, vários modelos simplificados de ME foram identificados [24] [25] com o subsequente estudo da sua fenomenologia no LHC.

---

<sup>8</sup> Problema relativo à grande discrepância entre os aspectos da força fraca e da gravidade. Não há consenso científico sobre o porquê a força fraca é  $10^{24}$  vezes mais forte que a força da gravidade.

# Capítulo 3

## Matéria Escura (ME)

Neste capítulo, trataremos os principais conceitos associados ao Modelo Cosmológico Padrão (MCP), algumas observações que evidenciaram a existência da ME, discutiremos a relação entre ME do MCP e ME de um modelo de extensão do MP, quais possíveis candidatos para ME, sua fenomenologia e os métodos de detecção e processos de interação da ME.

### 3.1 Cosmologia Padrão

Atualmente, acreditamos que o Universo tenha cerca de 13,8 bilhões de anos, que está em expansão, que seja aproximadamente homogêneo (nenhuma posição no espaço é diferente das demais) e isotrópico (suas características são as mesmas em qualquer direção) em grandes escalas. As equações que governam a evolução do Universo podem ser derivadas da teoria da relatividade geral de Einstein [26]. A métrica mais simples resultante da equação de Einstein, que satisfaz a homogeneidade e isotropia do Universo, é a métrica proposta por Friedmann-Robertson-Walker (FRW) e é dada por:

$$ds^2 = -dt^2 + a^2(t) \left[ \frac{dr^2}{1 - kr^2} + r^2(d\theta^2 + \text{sen}^2\theta d\phi^2) \right] \quad (3.1)$$

onde  $a(t)$  é um fator de escala que fornece os tamanhos relativos das superfícies espaciais. A geometria espacial do Universo é descrita por um valor constante no tempo,  $k$ . As coordenadas  $r$ ,  $\theta$ ,  $\phi$  se expandem de acordo com a expansão do Universo.

A equação de FRW resultante que forma a base da cosmologia contemporânea (MCP),

pode ser escrita como:

$$\left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2} = \frac{8\pi G_N \rho}{3}, \quad (3.2)$$

onde  $\dot{a} = da(t)/dt$  é a taxa de variação de expansão do Universo,  $G_N$  é a constante gravitacional de Newton e  $\rho$  é a densidade dos constituintes que compõem o Universo. Uma quantidade comumente utilizada é o parâmetro Hubble ( $H$ ) que representa a taxa com que o Universo está se expandindo. A lei de Hubble foi expressa a partir de observações feitas pelo telescópio espacial Hubble que constataram o afastamento de galáxias distantes. O parâmetro  $H$  é expresso matematicamente por:

$$H = \frac{\dot{a}}{a} \quad (3.3)$$

Podemos expressar uma outra quantidade chamada de parâmetro densidade total  $\Omega$ , adimensional, como

$$\Omega \equiv \frac{\rho_{tot}}{\rho_c}, \quad (3.4)$$

onde  $\rho_c$  é chamada de densidade crítica e é equivalente a

$$\rho_c = \frac{3H_0^2}{8\pi G_N}. \quad (3.5)$$

A média da densidade de massa total do Universo é representada por  $\rho_{tot}$ .  $H_0$  é a constante de Hubble, um valor aproximado para o parâmetro Hubble, com  $H_0 \approx 67.4 \pm 0.5 \text{ km s}^{-1} \text{ Mpc}^{-1}$  [27].

Substituindo a equação 3.4 em 3.2, temos uma relação entre a curvatura espacial  $k$  e o parâmetro  $\Omega$ :

$$1 - \Omega = \frac{k}{H_0^2 a^2} \quad (3.6)$$

Na Tabela 3.1, vemos que o sinal de  $k$  é determinado se  $\Omega$  for maior, menor ou igual a 1.

$\rho_{tot} < \rho_c$	$\Omega < 1$	$k = -1$ , Universo aberto
$\rho_{tot} = \rho_c$	$\Omega = 1$	$k = 0$ , Universo plano
$\rho_{tot} > \rho_c$	$\Omega > 1$	$k = 1$ , Universo fechado

Tabela 3.1: Classificação de modelos cosmológicos baseados no valor da densidade total  $\rho_{tot}$ , em termos da densidade crítica  $\rho_c$ .

Na Figura 3.1 é apresentado um gráfico que descreve o tamanho relativo do Universo durante os seus bilhões de anos sob diferentes cenários em relação a como o espaço e a matéria evoluem com o tempo. Ao lado de cada curva estão as representações bidimensionais das geometrias dos modelos de FRW fechado, aberto e plano. O Universo plano, mostrado na figura 3.1, tem uma densidade crítica (equação 3.5) e a sua taxa de expansão está continuamente diminuindo. O Universo aberto possui baixa densidade e sua expansão também está diminuindo, mas não tanto quanto ao Universo com densidade crítica, porque a força da gravidade não é tão forte nesse caso. Por último, é apresentado um Universo que deixa de se expandir e começa novamente a se contrair. Esta contração seria acelerada e, eventualmente, produziria o *Big Crunch*<sup>1</sup>, que é o inverso do *Big Bang*. No presente momento a teoria mais aceita é a que considera a geometria do Universo sendo plana e, portanto,  $k = 0$  (ver Tabela 3.1).

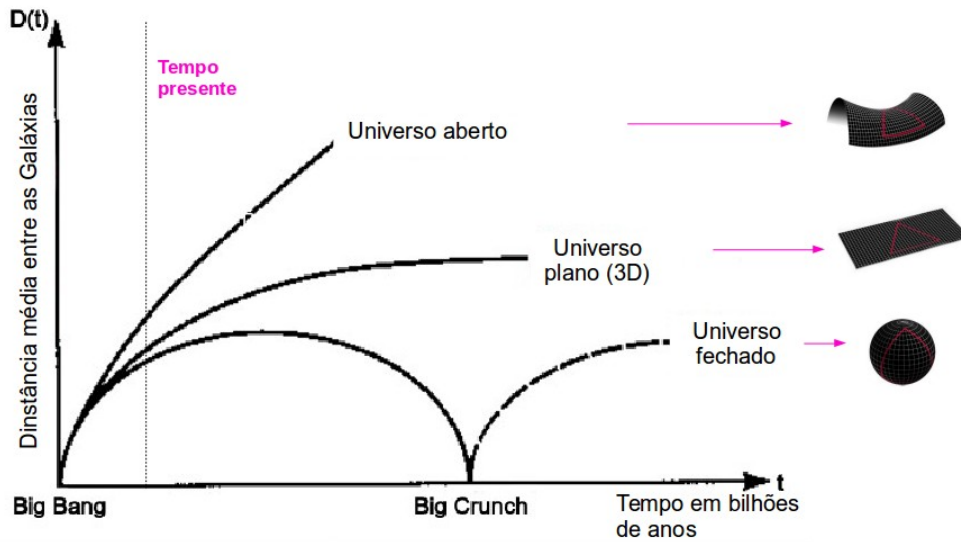


Figura 3.1: O gráfico descreve o tamanho relativo do Universo, distância média entre as Galáxias, durante os seus bilhões de anos sob diferentes cenários sobre como o espaço e a matéria evoluem com o tempo. Ao lado de cada curva está a representação geométrica do Universo: Universo aberto, plano e fechado. O tempo presente destacado em rosa equivale a aproximadamente 13,8 bilhões de anos.

A densidade de matéria total ( $\rho_{tot}$ ), pode ser escrita como uma soma de todas as contribuições das substâncias que compõem o Universo:

$$\rho_{tot} = \sum_i \rho_i , \quad (3.7)$$

onde  $i$  varia de 1 até  $n$  substâncias.

<sup>1</sup> Grande Colapso - hipótese em que o Universo começaria, no futuro, a contrair-se devido à atração gravitacional, até entrar em colapso sobre si mesmo.

Geralmente, as contribuições de maior importância consideradas para a composição aproximada de  $\rho_{tot}$  são  $\rho_M$  (o índice  $M$  refere-se a uma soma da densidade de matéria comum ( $\rho_{MC}$ ) + matéria escura ( $\rho_{ME}$ )),  $\rho_r$  (o índice  $r$  refere-se a densidade correspondente a radiação cósmica de fundo<sup>2</sup> - RCF) e  $\rho_\Lambda$  (o índice  $\Lambda$  refere-se a contribuição de energia escura).

Os resultados obtidos pelo trabalho [27] da *Colaboração Planck* para os parâmetros de densidade (equações 3.8), indicam que o Universo é, de fato, plano.

$$\begin{aligned}\Omega_M &= \rho_{MC}/\rho_c + \rho_{ME}/\rho_c = 0,3153 \pm 0,0073 & (3.8) \\ \Omega_r &\approx 5.38 \times 10^{-5} \\ \Omega_\Lambda &= 0.6847 \pm 0.0073 \\ \Omega &\approx \Omega_M + \Omega_r + \Omega_\Lambda \approx 1\end{aligned}$$

Outro parâmetro cosmológico adotado é a densidade de relíquia de alguma quantidade presente no Universo. Para evitar a incerteza relacionada à taxa de expansão do Universo, a densidade de relíquia de uma substância qualquer  $X$  é abordada como  $\Omega_X h^2$ , onde  $h$  é uma constante adimensional e é chamada constante reduzida de Hubble, sendo dada por  $h = H_0/100 km s^{-1} Mpc^{-1}$ .

Considerando a abundância de matéria escura presente no nosso Universo, sua densidade de relíquia é definida como

$$\Omega_{ME} h^2 \approx \frac{3 \times 10^{-27} cm^3 s^{-1}}{\langle \sigma v \rangle}, \quad (3.9)$$

onde  $\sigma$  é a seção de choque de aniquilação total de um par de partículas de ME em partículas do MP;  $v$  é a velocidade relativa entre duas partículas ME em relação ao seu centro de massa; os colchetes indicam uma média térmica, logo,  $\langle \sigma v \rangle$  é a média térmica da seção transversal de aniquilação de ME, comumente chamada apenas de seção de choque de aniquilação.

Medidas realizadas pelo trabalho [27] da *Colaboração Planck* determinaram que a ME é abundante ( $\Omega_{ME} h^2 = 0.120 \pm 0.001$ ) sobre a matéria bariônica<sup>3</sup> ( $\Omega_{matéria\ bariônica} h^2 =$

<sup>2</sup> Os fótons liberados durante a evolução do Universo continuam fluindo livremente com a expansão do Universo e formam a radiação cósmica de fundo. A RCF nos dá a luz (radiação) mais antiga após a última dispersão em nosso Universo [28].

<sup>3</sup> Em cosmologia, matéria bariônica é tudo aquilo que constituem os átomos. Diferentemente, em

$0.0224 \pm 0.0001$ ). Ainda assim, não sabemos a distribuição exata de ME em galáxias, mas conhecemos evidências que indicam a sua existência, como veremos a seguir na próxima sessão.

## 3.2 Evidências da ME

A primeira percepção da existência da ME ocorreu no ano de 1933 pelo astrônomo Fritz Zwicky [26] que, ao estudar aglomerados de galáxias, conseguiu inferir a razão massa-luminosidade  $M/L$  (onde  $M$  e  $L$  representam a massa total do aglomerado e a massa da matéria luminosa, respectivamente). Uma estimativa da matéria luminosa, que é a matéria que emite radiação eletromagnética, é feita a partir dos efeitos gravitacionais observados da radiação eletromagnética das estrelas ao redor do nosso Sistema Solar.

Os aglomerados são formados por galáxias ligadas gravitacionalmente e, como galáxias em aglomerados não possuem velocidades relativísticas, a física Newtoniana pode ser usada para descrever, por exemplo, as dispersões das velocidades das galáxias. Zwicky, conseguiu medir essas velocidades para o aglomerado de Coma através da aplicação do teorema do Virial. Esse teorema é derivado do problema de força central e diz que se a energia cinética total ( $E_c$ ) e energia potencial gravitacional total ( $E_g$ ) não variam consideravelmente com o tempo, em um sistema que interage gravitacionalmente, então podemos afirmar que,

$$E_c = -\frac{1}{2}E_g \quad (3.10)$$

Considerando, portanto, um sistema gravitacionalmente ligado com várias massas, para a  $i$ -ésima partícula  $i$ , com vetor posição  $\mathbf{r}_i$ , massa  $m_i$  e força  $\mathbf{F}_i$  atuante sobre essa partícula, temos que o trabalho virial  $W$  nesse sistema com  $n$  partículas é dado por:

$$W = \sum_{i=1}^n \mathbf{F}_i \cdot \mathbf{r}_i, \text{ com } \mathbf{F}_i = m_i \frac{d^2 \mathbf{r}_i}{dt^2}. \quad (3.11)$$

Utilizando a regra da cadeia nós obtemos, portanto, que

$$W = \sum_{i=1}^n m_i \left[ \frac{d}{dt} \left( \frac{d\mathbf{r}_i}{dt} \cdot \mathbf{r}_i \right) - \left( \frac{d\mathbf{r}_i}{dt} \cdot \frac{d\mathbf{r}_i}{dt} \right) \right] \quad (3.12)$$

Escrevendo a equação 3.12 em termos de momento de inércia ( $I$ ) do sistema e energia física de partículas o termo bárion remete a uma partícula formada por três quarks.



cinética total translacional, temos

$$W = \frac{1}{2} \sum_{i=1}^n \frac{d^2}{dt^2} I - 2E_c \quad (3.13)$$

O momento de inércia não muda com o tempo num sistema em equilíbrio dinâmico, assim,  $\frac{d^2 I}{dt^2} = 0$  e a equação 3.13 fica:

$$W + 2E_c = 0 \quad (3.14)$$

A força central  $F_i$  na  $i$ -ésima partícula de massa  $m_i$  desse sistema, devido a presença de todas as outras partículas pode ser escrita como:

$$F_i = \sum_{j \neq i} G_N m_i m_j \frac{(\mathbf{r}_j - \mathbf{r}_i)}{|\mathbf{r}_j - \mathbf{r}_i|^3} \quad (3.15)$$

Logo,  $W$  passa a ser:

$$W = \sum_{i=1}^n \sum_{j \neq i} G_N m_i m_j \frac{\mathbf{r}_i \cdot (\mathbf{r}_j - \mathbf{r}_i)}{|\mathbf{r}_j - \mathbf{r}_i|^3} = -\frac{1}{2} \sum_{i=1}^n \sum_{j \neq i} G_N \frac{m_i m_j}{|\mathbf{r}_i - \mathbf{r}_j|} = E_g, \quad (3.16)$$

Assim, podemos reescrever a equação 3.14 como  $E_g + 2E_c = 0$ , que é o teorema do Virial descrito em 3.10.

Portanto, dado um sistema cujo determinado aglomerado possui massa total  $M$ , velocidade de dispersão  $V$ , e raio  $R$ , podemos escrever sua energia cinética total e energia potencial gravitacional, respectivamente como:

$$E_c = \frac{1}{2} M V^2 \quad e \quad E_g = -\frac{G_N M^2}{2R}, \quad (3.17)$$

Sendo a massa

$$M = \frac{2RV^2}{G_N} \quad (3.18)$$

Em suas observações, Zwicky descobriu que as velocidades excediam o que poderia ser atribuído à matéria luminosa das galáxias, isto é, para a quantidade de matéria luminosa calculada as velocidades dos aglomerados deveriam ser muito menores. Desta forma, ele pôde verificar que a massa da equação 3.18 era muito maior do que  $L$  ( $M \gg L$ ). Então, viu-se que era necessário introduzir o termo matéria escura para indicar essa quantidade

de matéria faltante. Depois da primeira evidência da existência da matéria escura, outros astrônomos também conseguiram evidências para a sua presença no Universo através de outros métodos para a determinação da massa de aglomerados.

Por volta dos anos 70, Vera Rubin [29] conseguiu obter outra evidência a favor da matéria escura através das observações dos raios-X e gases luminosos das galáxias. Com isso, foi possível analisar as curvas de rotação das galáxias a partir das velocidades como função da distância radial ao centro galáctico de estrelas e gases, como mostrado na Figura 3.2.

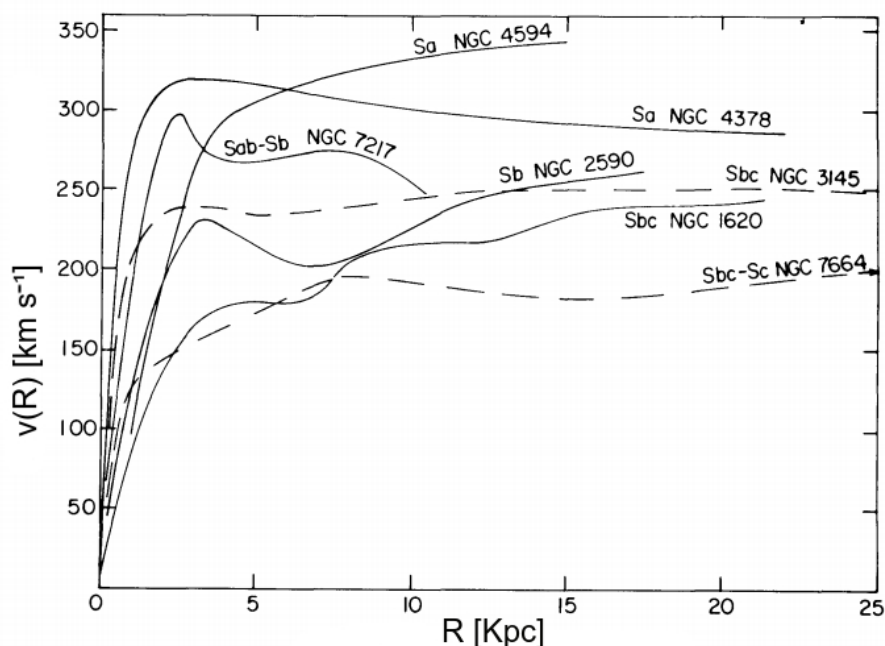


Figura 3.2: Curvas das velocidades rotacionais como função da distância radial para sete galáxias [29].

Pela dinâmica Newtoniana, de forma aproximada, a força resultante atuante em algum objeto orbitando a uma distância  $R$  do centro de rotação de alguma galáxia, é dada por:

$$F_r = F_G \longrightarrow \frac{mv^2}{R} = \frac{G_N Mm}{R^2} \quad (3.19)$$

onde  $M$  é a massa do aglomerado,  $v$  é a velocidade de rotação e  $m$  é a massa do objeto (pode ser uma estrela, por exemplo).

Da equação 3.19 temos que

$$v = \sqrt{\frac{G_N M}{R}} \quad (3.20)$$

De acordo com a predição teórica, era esperado que a velocidade de rotação de algum astro ou gás, em órbita nas galáxias da Figura 3.2, diminuísse à medida que se distanci-

assem da região central das galáxias, concordando com a interpretação da equação 3.20. Entretanto, o que foi observado, como podemos notar na Figura 3.2, era que essa velocidade se mantinha praticamente constante até a borda da galáxia. Tal observação, somente seria possível se houvesse uma quantidade de matéria maior em comparação à luminosidade da galáxia.

Um método mais atual para medir a massa de galáxias seria pelas lentes gravitacionais, cujos efeitos estão associados à curvatura da luz causada por um objeto com grande massa que muda a geometria do espaço ao seu redor. A trajetória da luz torna-se curva nas proximidades de um corpo massivo (lente), o que dá origem ao efeito de lente. Na Figura 3.3a, podemos ver uma imagem do Aglomerado Abell 2218 onde arcos gravitacionais, como indicado em vermelho, aparecem. Esses arcos são formados devido a deflexão da luz de qualquer astro que emita luz e que está mais distante do corpo chamado de lente. A Figura ao lado, 3.3b, mostra como um observador na Terra vê as imagens distorcidas de um objeto (fonte) que pode estar presente ao fundo de um aglomerado de galáxias, por exemplo, a uma distância considerável. O aglomerado funciona, portanto, como uma lente para a luz que vem do objeto que está mais ao fundo. Lentes gravitacionais em aglomerados de galáxias nos dão uma estimativa da sua massa central indicando a presença de grande matéria invisível, isto é, matéria escura.

De acordo com [28], o ângulo de deflexão de um raio de luz por conta da ação da gravidade de uma massa pontual  $M$  é dado por:

$$\alpha_D(x) = \frac{4G_N M}{xc^2}, \quad (3.21)$$

onde  $x$  é a distância<sup>4</sup> mais próxima entre o raio de luz e a massa  $M$  da lente e  $c$  é a velocidade da luz no vácuo.

A equação da lente é importante para a estimativa da massa de um aglomerado e é dada por:

$$D_s \theta = D_s \beta_s + D_{ds} \alpha_D \quad (3.22)$$

A partir da equação 3.22 pode-se medir a massa de um aglomerado  $M$  (lente) e compara-la com sua massa luminosa  $L$  através da relação  $M/L$ .

Todos os métodos citados têm como finalidade calcular a massa de galáxias ou aglome-

---

<sup>4</sup> Correspondente ao tamanho **J** na Figura 3.3b.

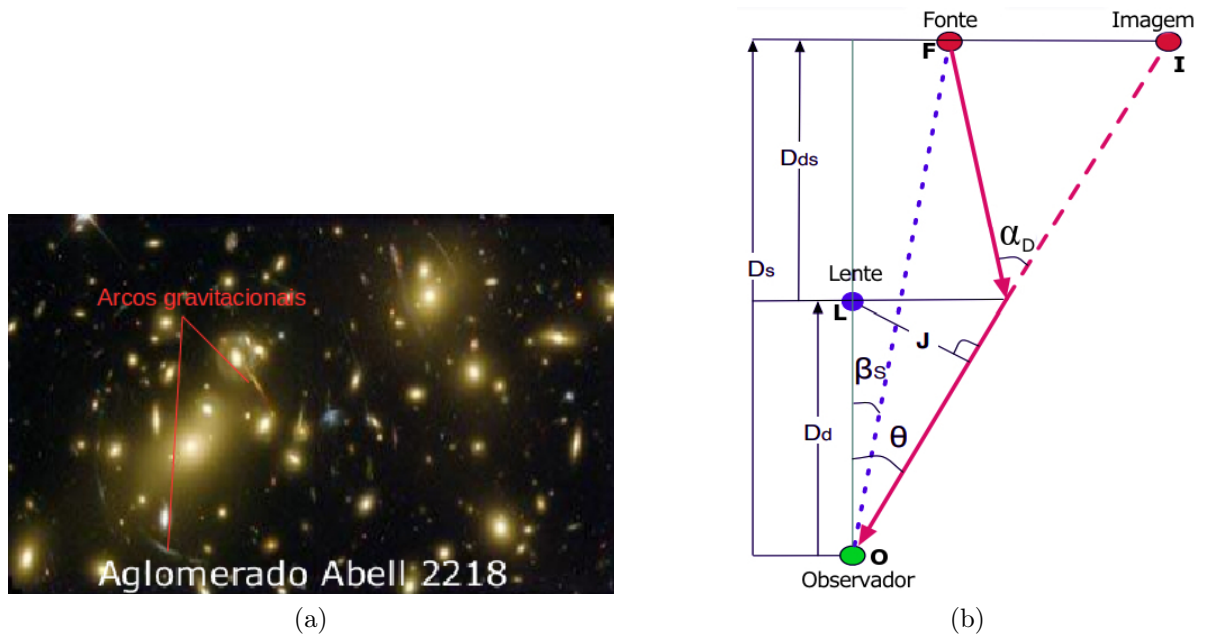


Figura 3.3: Lentes gravitacionais. (a) Imagem obtida por telescópio do Aglomerado Abell 2218 onde podemos ver os arcos gravitacionais formados devido a geometria curva do espaço pela presença de corpos super massivos. (b) Exemplo esquemático de como é vista a imagem de um objeto por um observador na Terra [30]:  $D_s$ ,  $D_{ds}$  e  $D_d$  são a distância do plano do observador até o plano da fonte, distância do plano lente-fonte e distância do observador até a lente, respectivamente; **O**, **L** e **F** são respectivamente as posições do observador, defletor (lente) e fonte (OL é o eixo de referência ótico); **J** é a distância perpendicular da reta **OI** até a lente, chamado parâmetro de impacto; os ângulos  $\alpha_D$ ,  $\beta$  e  $\theta$  são, nessa ordem, ângulo de deflexão de Einstein, o ângulo em que a fonte seria observada na ausência da lente e ângulo em que a fonte é observada; **I** é o ponto de imagem.

rados e comparar com a massa luminosa. Para mais detalhes, as referências [31] [32] [33] podem ser consultadas.

### 3.3 Propriedades da Partícula de ME

Os elementos relativos à compreensão de algumas características da ME encontram-se no Modelo Padrão da Cosmologia (MPC), o qual caracteriza essa nova física da ME associada à origem e evolução do Universo.

A relação entre ME do MPC e um possível candidato à ME do MP é dada pelas seguintes propriedades:

- um candidato à ME deve ser efetivamente neutro, já que não interage de forma detectável com fótons;
- deve ser estável devido ao tempo de vida igual à idade do Universo;

- ser pouco interagente, porque suas interações não-gravitacionais devem ser pouco intensas;
- a ME deve ser uma matéria não-bariônica, pelo fato de não se comportar como a matéria que conhecemos;
- é suficientemente abundante, pois a sua existência representa cerca de 27% do Universo.

### 3.4 Possíveis Candidatos para ME

Os candidatos para matéria escura podem ser classificados como:

- **WIMPs:** Os WIMPs [26], ou Partículas Massivas que Interagem Fracamente, são partículas não-bariônicas que interagem apenas através da gravidade e da força nuclear fraca.
- **MACHOs:** Os MACHOs [26] (Objetos Compactos que são Halo Astrofísicos Massivos), são grandes objetos astrofísicos formados de matéria bariônica com massas substancialmente maiores que as dos WIMPs.

Nem todos os candidatos de ME se encaixam nas definições de WIMPs e MACHOs. Porém, os WIMPs são os modelos mais aceitos atualmente, embora nenhum ainda tenha sido comprovado.

Como já mencionado, o MP não fornece um candidato viável para matéria escura. Sendo assim, a ME é uma motivação para buscar a física além do Modelo Padrão.

### 3.5 Fenomenologia da ME

Para efetivamente realizar uma análise fenomenológica sobre a caracterização das propriedades da ME, geralmente são usadas duas abordagens principais de modelos independentes [34]:

- **Teoria de Campo Efetivo (TCE):** A hipótese mais simples possível é que a partícula ME seja o único estado além do MP que seja acessível no LHC. Nesse caso, as interações entre estados ME e MP podem ser descritas por uma teoria de campo

efetivo contendo operadores de dimensão maior do que quatro. Nessa abordagem, as interações entre MP e ME são descritas por operadores de dimensões extras mais altas [35]. A vantagem dessa abordagem é que os únicos parâmetros livres são os coeficientes dos operadores e a massa do ME.

- **Modelos simplificados:** Além da descrição fornecida pelas teorias TCE, considerando que as partículas ME sejam as únicas cinematicamente acessíveis, podemos considerar a possibilidade de uma outra partícula leve. Essa classe de modelos é conhecida como modelos simplificados de ME [36]. A abordagem de um modelo simplificado adiciona um único mediador e uma partícula ME ao MP. Modelos simplificados são descritos por um pequeno número de parâmetros diretamente relacionados aos observáveis da física dos colisores, como massas de partículas, seções de choque e frações ramificadas.

### 3.6 Métodos de Detecção e Processos de Interação da ME

Sabemos que a matéria escura existe devido as evidências das interações gravitacionais em grande escala. Mas, para sabermos a natureza da ME é preciso utilizarmos métodos que possibilitem a sua detecção e, caso seja uma WIMP, poderemos detectá-lá no futuro. Existem três formas possíveis para detectar a assinatura de WIMPs: experimentos de detecção direta (DD) via espalhamentos de núcleos, experimentos de detecção indireta (DI) que buscam pela produção de partículas do MP a partir da aniquilação de ME e procura por sinais de ME em colisores de hádrons (ver Capítulo 4). Todos esses experimentos têm como principal objetivo encontrar limites para a massa de ME.

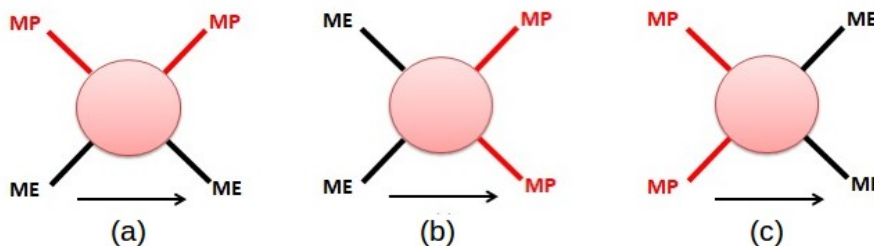


Figura 3.4: Possíveis canais de detecção de matéria escura. (a) Detecção direta. (b) Detecção indireta. (c) Colisores.

A Figura 3.4 mostra uma representação esquemática dos possíveis acoplamentos de matéria escura com partículas do modelo padrão (MP). Na Figura 3.4a o processo indicado **MEMP**  $\rightarrow$  **MEMP** caracteriza os experimentos de detecção direta, onde uma partícula de matéria escura e outra do modelo padrão interagem e produzem uma partícula **ME** e **MP**; o processo descrito em 3.4b **MEME**  $\rightarrow$  **MPMP** mostra a aniquilação de partículas de ME dando origem partículas do MP; as buscas por ME em colisores de partículas é representado por **MPMP**  $\rightarrow$  **MEME**, indicado em 3.4c.

Todos os tipos de experimentos citados requerem que a ME interaja com a matéria comum de alguma forma. Nas subseções a seguir, trataremos brevemente de cada experimento de detecção.

### 3.6.1 Detecção Direta (DD)

Considerando que a Galáxia está cheia de WIMPs, muitas dessas partículas devem passar pela Terra interagindo com a matéria bariônica. Sendo assim, essa interação pode ser observada a partir dos espalhamentos de núcleos em laboratórios pelos detectores de Detecção Direta (DD) [37]. Portanto, os experimentos de DD procuram por sinais de ME a partir da interação de ME com a matéria comum.

Há duas formas de como os WIMPs podem se acoplar com o conteúdo do spin de um núcleo: interações que dependem do spin e interações que são independentes do spin da ME.

A assinatura de ME em um experimento de DD consiste em detectar a energia perdida pelo núcleo que recuou ao interagir com um WIMP. O espectro de recuo diferencial de WIMP por unidade de massa do detector é dada por:

$$\frac{dR}{dE}(E, t) = \frac{\rho_0}{M_\phi \cdot m_A} \cdot \int v \cdot f(v, t) \cdot \frac{d\sigma}{dE}(E, v) d^3v \quad (3.23)$$

onde  $M_\phi$  é a massa da matéria escura;  $m_A$  é a massa de um núcleo;  $\frac{d\sigma}{dE}(E, v)$  é a seção de choque diferencial;  $v$  é a velocidade da ME no referencial do detector no laboratório;  $\rho_0$  e  $f(v, t)$  são parâmetros astrofísicos, densidade local da matéria escura e distribuição da velocidade do WIMP no referencial do detector, respectivamente. Esta distribuição de velocidade depende do tempo devido à revolução da Terra em torno do Sol.

Na Figura 3.5 vemos os limites de exclusão para as massas de WIMPs, cuja interação

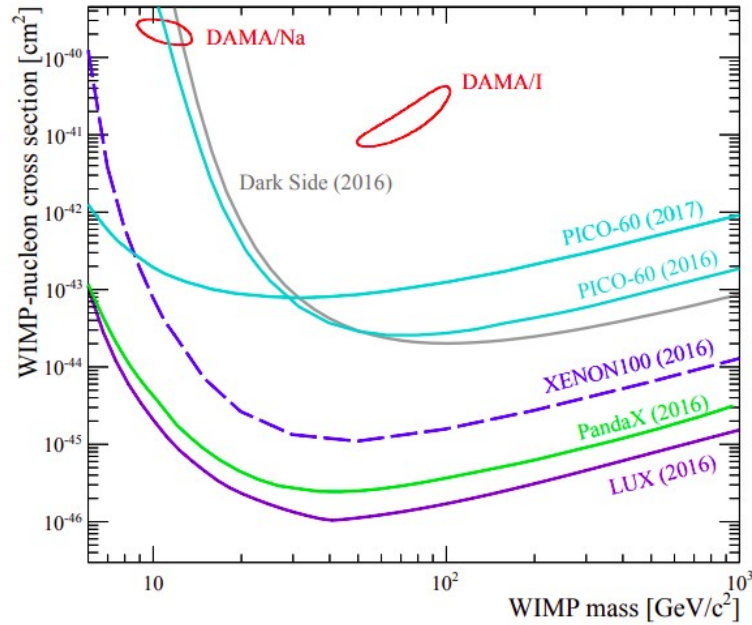


Figura 3.5: Resultados experimentais de detecção direta e limites de exclusão para seções de choque de interações WIMP-núcleon independente do spin. Os contornos fechados são indicações do sinal de ME. As curvas delimitam os limites de exclusão do espaço de parâmetros acima. Fonte: [38]

não depende do spin da partícula de ME, em experimentos de DD: Dark Side, PICO, XENON, PandaX e LUX. Vemos também as curvas fechadas que indicam sinal de ME<sup>5</sup>. Cada curva mostrada na Figura 3.5 exclui o espaço de parâmetros acima dela. O objetivo desses experimentos é diminuir cada vez mais essa região para melhorar os limites para a massa de ME.

### 3.6.2 Detecção Indireta (DI)

Experimentos de detecção indireta (DI) de ME são feitos via estudos do fluxo local de partículas ou antipartículas do MP produzidos a partir do decaimento ou aniquilação de ME [3] [37].

O fluxo  $\Phi(E, \theta)$  de partículas neutras<sup>6</sup> produzidas na aniquilação ou decaimento da ME no Centro Galáctico (CG), por exemplo, é dado por:

$$\Phi(E, \theta) = \frac{\Gamma}{4\pi M_\phi^b} \frac{dN}{dE} \int \rho^b[r(l, \theta)] dl \quad (3.24)$$

<sup>5</sup>O experimento DAMA, dependendo do átomo alvo (Na e I), tem as regiões do sinal de ME como mostradas na Figura 3.5. Porém, a origem desse sinal ainda é desconhecida e controversa, porque não está de acordo com os outros experimentos que excluem esse resultado.

<sup>6</sup>Partículas sem carga que surgem da aniquilação ou decaimento de ME que possuem termos adicionais em seu fluxo  $\Phi$  relevantes para a propagação de raios cósmicos [39].



O termo  $\Gamma$  na eq. 3.24 representa a taxa de interação das partículas de ME. Para o fluxo de aniquilação, a potência  $b$  é igual a 2 e  $\Gamma = \frac{\langle \sigma v \rangle}{2}$ ; para o fluxo de decaimento  $b = 1$  e  $\Gamma = 1/\tau$ , onde  $\tau$  é a meia vida da partícula de ME;  $M_\phi$  é a massa de ME; a derivada  $dN/dE$  é o espectro de energia de aniquilação ou decaimento da ME;  $\int \rho^b[r(l, \theta)] dl$  integra a potência <sup>7</sup> (1 ou 2) da densidade  $\rho$  da ME ao longo de uma linha de observação  $l$ ; a coordenada  $r(l, \theta)$  é escrita como  $r(l, \theta) = \sqrt{r_\odot^2 + l^2 - 2r_\odot l \cos\theta}$ , onde  $\theta$  é o ângulo entre a direção de observação e o plano do CG;  $r_\odot \approx 8,5 kpc$  é a distância do Sol ao CG.

A Figura 3.6 mostra algumas comparações dos limites de exclusão de experimentos DI e DD e previsões para a aniquilação de ME em pares  $b\bar{b}$  e  $Z\gamma$ . Os limites experimentais dados por Fermi-LAT e H.E.S.S. (experimentos de DI) são indicados para os pares  $b\bar{b}$  e  $Z\gamma$ . Um limite de exclusão de DD (LUX) é também apresentado.

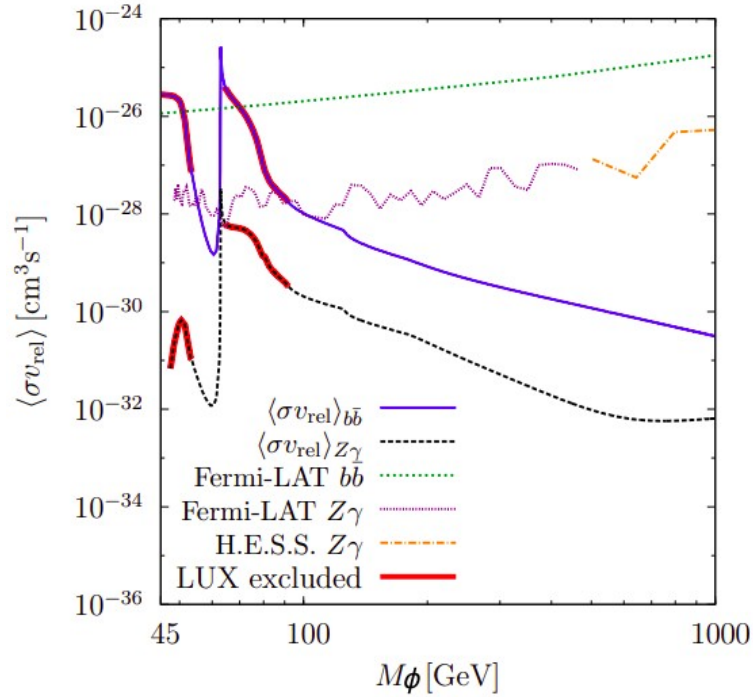


Figura 3.6: Predições da seção de choque de aniquilação da ME em função da massa da matéria escura ( $M_\phi$ ). Estão sendo comparados os limites de experimentos DI, Fermi-LAT e H.E.S.S., e do experimento DD LUX para os canais  $\phi\phi \rightarrow b\bar{b}$  e  $\phi\phi \rightarrow Z\gamma$ , através de observações de galáxias e Centro do Halo Galáctico. Fonte: [40]

<sup>7</sup>A escolha da potência se baseia em informações vindas de simulações das interações de vários corpos e diz quais os prováveis perfis a densidade de ME possui em um determinado sistema.

### 3.6.3 Colisores

Em aceleradores de partículas, esperamos que a presença de ME esteja relacionada a assinatura de energia faltante nos experimentos. Desta forma, na colisão de elétrons ou prótons em colisores, pode ser produzido pares de ME [41]. Nesse caso, a presença de ME pode apenas ser inferida pela falta de momento transversal ou energia transversal faltante, como também é chamado (cuja definição será abordada no próximo capítulo).

## Capítulo 4

# O Grande Colisor de Hádrons, Experimento ATLAS e Ferramentas de Simulação

Como visto no capítulo anterior, existem alguns métodos que buscam detectar a ME presente no nosso Universo. Em Física de Partículas, há um grande interesse em relacionar a existência de ME ou a sua produção às interações de partículas já conhecidas no Grande Colisor de Hádrons (LHC). Devido a isso, abordaremos nas próximas seções os temas relacionados ao LHC localizado no Centro Europeu de Pesquisas Nucleares, com foco na física do detector ATLAS, e algumas ferramentas de simulação que auxiliam no estudo fenomenológico de ME.

### 4.1 Centro Europeu de Pesquisas Nucleares

O Centro Europeu de Pesquisas Nucleares (CERN), fundado em 1954, é o maior laboratório de Física de Partículas que existe atualmente. Localizado a noroeste de Genebra, na fronteira entre a França e a Suíça, o CERN fornece uma variedade de instalações de aceleradores de partículas que auxiliam em pesquisas e permitem novas descobertas a cerca do que o mundo é feito e como ele funciona. Até o momento, possui 23 estados-membros e mais de 600 institutos e Universidades de todo o mundo usam as instalações do CERN.

Para experimentos em Física de Partículas, as instalações hospedadas na organização

contém 34 quilômetros de aceleradores construídos a 100 metros da superfície, detectores colocados em enormes cavernas e outras infraestruturas necessárias para pesquisa em física de altas energias.

O complexo de aceleradores do CERN consiste numa série de aceleradores lineares e circulares, construídos com o objetivo de impulsionar prótons e íons pesados em todo o caminho para o LHC (Grande Colisor de Hádrons). Como mostra a Figura 5.1, o complexo de aceleradores do CERN é composto por diferentes aceleradores que são encadeados de forma que a energia dos feixes é progressivamente aumentada. O processo de aceleração inicia-se no LINAC 4 (LINEar ACcelerator) que alimenta o Proton Synchrotron Booster (PSB). Em seguida, os prótons são injetados no Proton Synchrotron (PS) e, normalmente, usa-se o Super Proton Synchrotron (SPS) de onde o feixe segue para o LHC.

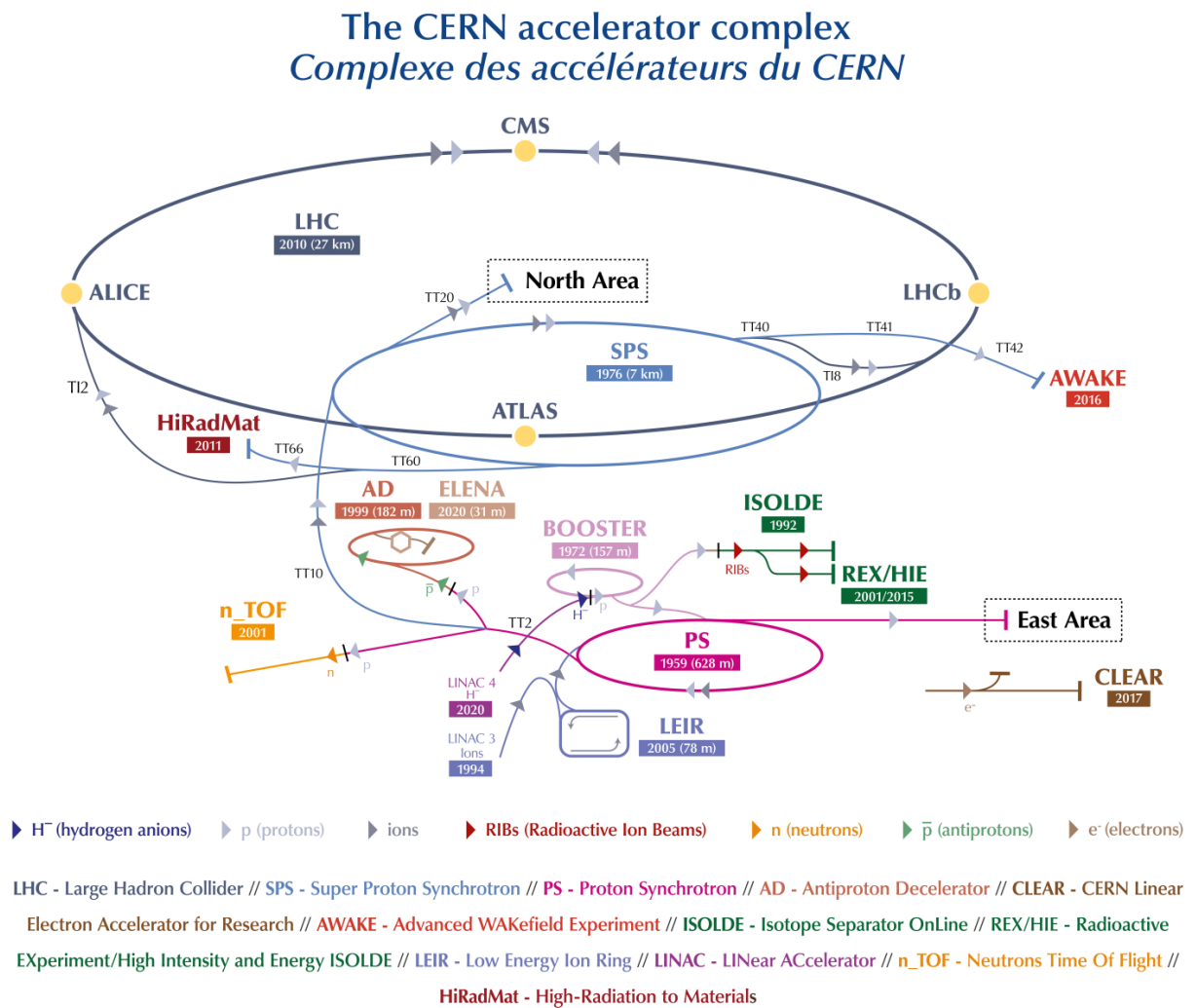


Figura 4.1: Complexo de aceleradores no CERN. Fonte: [42].

O LHC, na Figura 4.1, é o último anel (linha azul escura) em uma cadeia complexa de aceleradores de partículas. As máquinas menores são usadas em uma cadeia para ajudar a impulsionar as partículas para suas energias finais e fornecer feixes para todo um conjunto de experimentos menores.

## 4.2 LHC

Fundado em 1998, o LHC funciona desde 2008 e está situado na fronteira da França com a Suíça, é o maior e o mais poderoso acelerador de partículas que existe. Está instalado em um túnel circular subterrâneo a 175m de profundidade com 27km de circunferência. Dentro desse acelerador, dois feixes acelerados de prótons ou íons pesados viajam em direções opostas a velocidades próximas da luz e colidem à altas energias. A energia do centro de massa entre esses dois feixes é definida como  $\sqrt{s} = (p_1^\mu + p_2^\mu)$ , onde  $p_1^\mu$  e  $p_2^\mu$  são os 4-momentos dos feixes incidentes. Como no centro de massa os momentos  $\mathbf{p}_1 + \mathbf{p}_2 = 0$ , então  $\sqrt{s} = E_1 + E_2$ , em que  $E_1$  e  $E_2$  são as energias do feixe 1 e 2, respectivamente.

Em 2009, ocorreu uma colisão com energia do centro de massa  $2.36TeV$ . No ano seguinte, aconteceu a primeira colisão com feixes de energia  $3.5TeV$ , isto é, com energia total  $7.0TeV$ . No ano de 2012, realizaram as primeiras colisões a  $8.0TeV$  e em 2013 o LHC foi desligado para corrigir um problema, mas foi reiniciado em abril de 2015 para funcionar com sua energia total de  $13.0TeV$ . Novamente em 2018, o LHC foi desligado para atualizações e está programado para terminar no final de 2021 ou início de 2022 [43].

O desempenho dos aceleradores é caracterizado pela luminosidade ( $L$ )<sup>1</sup>. A luminosidade define o número de partículas por unidade de área por segundo gerado nos feixes dos experimentos com partículas de alta energia e é geralmente expressa em  $fb^{-1}s^{-1}$ . Assim, se a luminosidade for muito grande, o número de eventos produzidos para estudo também será. Portanto, a luminosidade está relacionada com o número de eventos por segundo e a seção de choque ( $\sigma$  – que caracteriza a probabilidade de ocorrência de uma colisão) através da expressão:

$$\frac{dn}{dt} = L\sigma, \quad (4.1)$$

---

<sup>1</sup>Vimos que em Cosmologia o termo luminosidade refere-se à massa luminosa. Em Física de Partículas o seu significado é outro, como explicado no texto.

onde  $n$  é o número de colisões efetivas em uma reação.

Integrando ambos os lados da equação 4.1 num determinado intervalo de tempo  $dt$ , temos que

$$N = \sigma \mathcal{L}, \quad (4.2)$$

onde  $N$  é o número total de eventos em uma colisão e  $\mathcal{L}$  é chamada grandeza luminosidade integrada:

$$\mathcal{L} = \int L dt. \quad (4.3)$$

Atualmente o LHC possui luminosidade integrada de  $\mathcal{L} = 30 fb^{-1}$ . A pretensão é que futuramente seja alcançado valores maiores para  $\mathcal{L}$ .

Os prótons ou íons pesados sofrem colisão no LHC em quatro pontos de interação onde os principais experimentos estão localizados: ALICE (A Large Ion Collider) que estuda a física dos íons pesados; ATLAS (A Toroidal LHC ApparatuS) e CMS (Compact Muon Solenoid) que estudam toda a física envolvida nas colisões próton-próton; LHCb (Large Hadron Collider beauty experiment) que é especializado na física do quark b.

### 4.3 Detector ATLAS e Quantidades Transversas

O detector ATLAS [44] [45] [46] é um dos quatro principais experimentos de física de partículas do LHC e possui em torno de 7000 toneladas. O ATLAS foi projetado para explorar o potencial de novas descobertas na física. A colaboração do ATLAS conta com aproximadamente 5500 físicos(as), estudantes, engenheiros(as) e técnicos(as) de diversas instituições e países, incluindo o Brasil. No ATLAS é testada as previsões do MP, estudada a física além do MP e desenvolvida novas teorias que nos levam a ter um melhor entendimento acerca do Universo. A descoberta do bóson de Higgs no ano de 2012 através dos experimentos CMS e ATLAS no CERN, por exemplo, permitiu uma explicação para a origem das massas das partículas elementares mediante a interação entre partículas.

As regiões do detector ATLAS podem ser descritas por um sistema de coordenadas, como mostrado na figura 4.2. A origem do sistema de coordenadas representa o ponto de interação de dois feixes de partículas colidindo frontalmente. A direção dos feixes está ao longo do eixo  $z$ . A direção positiva do eixo  $x$  é definida como apontada para o centro do anel LHC e a direção positiva do eixo  $y$  aponta para cima do detector. O plano  $x$ - $y$  é o plano transversal ao eixo  $z$ .

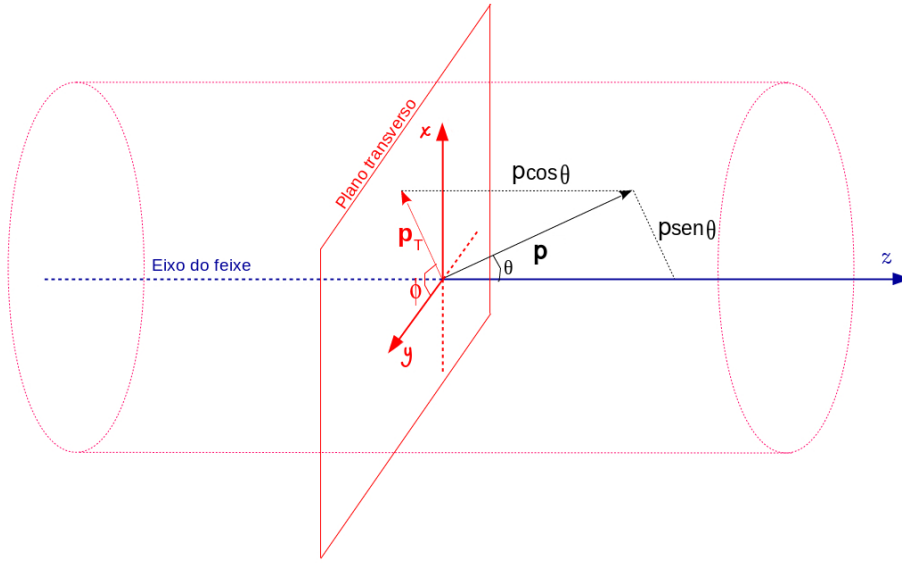


Figura 4.2: Esquema 3D do ATLAS com as indicações do momento transverso  $\mathbf{p}_T$  e momento linear  $\mathbf{p}$ .

Como mostra a figura 4.2,  $\theta$  é o ângulo polar medido a partir da linha do feixe relativo ao eixo z e  $\phi$  é o ângulo azimutal medido no plano x-y relativo ao eixo x.

Uma grandeza geralmente utilizada é a pseudorapidez que indica se uma partícula com carga é medida pelo detector ou não e é expressa em ângulo relativo ao eixo de colisão do feixe. A pseudorapidez é definida como

$$\eta = -\ln \left[ \operatorname{tg} \left( \frac{\theta}{2} \right) \right]. \quad (4.4)$$

Em um intervalo  $\eta \in [-2.5, 2.5]$  uma partícula com carga é identificada. Se  $\eta = 0$ , a partícula com carga está no plano x-y; se  $\eta = +\infty$ , ela está ao longo do eixo z; se  $\eta = -\infty$ , está ao longo da direção oposta de z. Para valores de  $|\eta| > 5$ , as partículas escapam da detecção. Mas, em um intervalo onde  $|\eta| \in [2.5, 5]$ , os momentos das partículas podem ser inferidos, entretanto, as partículas não podem ser identificadas no detector.

Outra variável muito comum no estudo do espalhamento de partículas em experimentos de colisões a altas energias é a distância de separação angular entre duas partículas ou jatos, e sua expressão matemática é

$$\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2}. \quad (4.5)$$

A colisão de hádrons é caracterizada por uma energia e momento total inicial ( $E_{in}, \mathbf{p}_{in}$ ). Pela conservação de energia e momento, após a colisão, temos que a energia e o momento

para  $n$  partículas produzidas no estado final são

$$\begin{aligned} E_{fin} &= \sum_i E_i \\ \mathbf{p}_{fin} &= \sum_i \mathbf{p}_i. \end{aligned} \tag{4.6}$$

Em muitas das vezes é medido  $E_{fin} < E_{in}$  e  $\mathbf{p}_{fin} \neq \mathbf{p}_{in}$ . Um exemplo disso, no processo  $pp \rightarrow t\bar{t}Z$ , neutrinos produzidos do decaimento do bóson  $Z$  ou quark  $t$  atravessam o detector, escapando da detecção. Isso contribui no erro da medição da soma dos momentos de cada partícula produzida no estado final. Sendo tais partículas somente inferidas a partir do desequilíbrio do momento total. Portanto, são perdidas muitas partículas na direção do feixe, assim, uma quantidade significativa da energia se perde nos eventos de colisão. Assim, projeções transversais (no plano x-y) para o momento e energia do feixe são utilizadas em quase todas as análises dos dados de aceleradores de hádrons.

A partir do momento transversal  $\mathbf{p}_T = (p_x, p_y)$  podemos escrever as componentes do momento linear  $\mathbf{p} = (p_x, p_y, p_z)$ , da figura 4.2, em função de  $\mathbf{p}_T$ , já que  $\mathbf{p} \equiv (\mathbf{p}_T, p_z)$  e também de  $\phi$  e  $\eta$ :

$$\begin{aligned} p_x &= p_T \cos\phi \\ p_y &= p_T \sin\phi \\ p_z &= p_T \sinh\eta \end{aligned} \tag{4.7}$$

Logo, a magnitude do momento transversal pode ser escrita como

$$p_T = p \sin\theta \tag{4.8}$$

A energia perdida nos eventos de colisão, mencionada anteriormente, implica na existência de um momento transversal faltante ( $\mathbf{p}_T$ ), que é uma projeção no plano x-y com mesma direção e sentido oposto de  $\mathbf{p}_T$ . Esta é uma grandeza importante quando desejamos aplicar restrições de que o momento faltante em um evento é devido às partículas de estado final não-detectadas ou invisíveis, como também podem ser chamadas.

Na literatura da área de Física de Partículas experimental e fenomenológica  $\mathbf{p}_T$  é denotado por  $\mathbf{E}_T$  (ou  $\mathbf{E}_T^{miss}$ ) e chamado de energia transversal faltante (*missing transverse energy* - MET), um vetor 2D. A energia transversal faltante (figura 4.3) é a mais ampla-



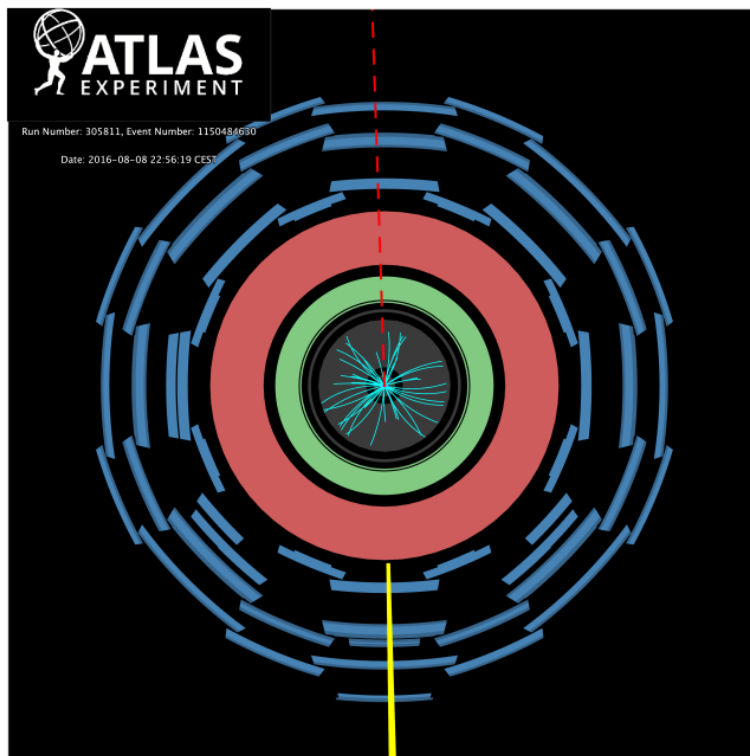


Figura 4.3: É mostrado um evento no detector ATLAS de um experimento de busca por matéria escura. O momento transversal (barra amarela) é balanceado pela energia transversal faltante (linha vermelha tracejada). Fonte: [47]

mente estudada em relação a novas pesquisas em física como, por exemplo, em modelos com candidatos para a matéria escura. Sua magnitude é dada por

$$MET = \cancel{E}_T = \sqrt{\left(\sum_n E_{x,n}\right)^2 + \left(\sum_n E_{y,n}\right)^2} \quad (4.9)$$

em que  $E_{x,n}$  e  $E_{y,n}$  são as energias da  $n$ -ésima partícula projetada ao longo dos eixos  $x$  e  $y$ , respectivamente.

A energia transversal faltante ( $\cancel{E}_T$ ) é igual a soma de todos os momentos transversais das partículas visíveis (detectáveis) do estado final, pois cada uma dessas partículas contribui para a quantidade de  $\cancel{E}_T$ :

$$\cancel{E}_T = - \sum_i \mathbf{p}_{T,i}, \quad (4.10)$$

onde  $\cancel{E}_T$  é negativa na soma vetorial de  $\mathbf{p}_T$  de todas as partículas  $i$  do estado final, por ser uma quantidade de  $\mathbf{p}_T$  que está faltando na direção de  $\mathbf{p}_T$ .

Ainda podemos definir a massa transversal do estado final em função de  $p_T$ ,  $\cancel{E}_T$ , e  $\phi$

como:

$$M_T = \sqrt{2p_T \cancel{E}_T (1 - \cos(\Delta\phi))}, \quad (4.11)$$

sendo  $p_T$  o momento transverso de um lépton com carga e  $\Delta\phi$  a variação ângular azimutal entre a direção dessa partícula visível e energia transversa faltante no plano x-y.

## 4.4 Procura por Matéria Escura no LHC

Como vimos, de fato, há uma abundância de matéria escura no Universo e apesar das observações cosmológicas serem muito convincentes para afirmarmos a sua existência, também podemos considerar uma descrição de partícula para ME pensando em modelos além do Modelo Padrão. Mesmo assim, não sabemos ao certo se a matéria escura é composta de partículas porque mesmo que a ME possua interações gravitacionais com a matéria comum, as partículas de ME são escuras, isto é, não detectáveis nos colisores de partículas atuais.

Se essa nova física da ME estiver na forma de partículas massivas que interagem fracamente, espera-se que suas assinaturas no LHC sejam caracterizadas por eventos com excesso de energia transversa faltante ( $\cancel{E}_T$ ), como já citado anteriormente. Intensos esforços experimentais estão sendo feitos no LHC para isolar essas assinaturas, embora nenhum sinal conclusivo tenha sido observado até o momento. Dessa forma, do ponto de vista experimental, podemos citar algumas estratégias relativas à procura da assinatura de ME [41]:

- Mono-jatos: a procura por partículas ME é feita através da busca nos dados coletados em colisões próton-próton no LHC por eventos nos quais um jato QCD com alto momento transverso é produzido com uma grande energia transversa faltante. A Figura 4.4a mostra um diagrama de Feynman correspondente a um canal de produção de ME em pesquisas de mono-jatos.
- Mono-V: em pesquisas mono-V ( $V = \gamma, W$  ou  $Z$ ) os bósons de calibre, irradiados no estado inicial, são produzidos junto com a assinatura de ME. A Figura 4.4b indica um processo mono-V ( $q\bar{q} \rightarrow Z\chi\bar{\chi}$ ) que produz um bóson  $Z$ , uma partícula de ME e sua antipartícula no estado final.
- mono-Higgs: caracteriza a procura por um bóson de Higgs do MP associado com

a falta de energia transversa. Na Figura 4.4c vemos um canal que tem a produção de um bóson de Higgs no estado final  $+\chi\bar{\chi}$ , cujo mediador da interação não é uma partícula do MP (um novo bóson proposto  $Z'$ ).

- ME + quarks top: a assinatura de ME está associada a produção de pares top do MP no estado final. Pesquisas desse tipo utilizam modelos que incluem novas partículas não previstas pelo MP. Como mostra a Figura 4.4d, a produção de quarks top está relacionada a assinatura de ME no estado final vindas de duas novas partículas  $\tilde{t}\tilde{t}$ .

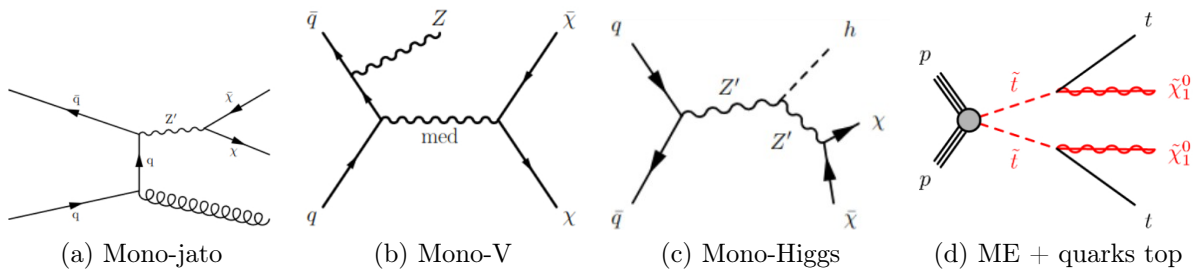


Figura 4.4: Alguns exemplos dos diagramas de Feynman dos processos estudados nas pesquisas de ME.

Partículas invisíveis, como a ME, podem ser acompanhadas por uma ou mais partículas de recuo visíveis. Se a ME é uma partícula, não parece decair. As leis de conservação podem impedir que a partícula de ME se decomponha em qualquer partícula de modelo padrão mais leve e paritária. As partículas de ME também podem ser produzidas em pares pelo decaimento de outras partículas.

## 4.5 Ferramentas de Simulação

### 4.5.1 CalcHep

O termo CalcHEP vem do inglês *Calculations in High Energy Physics* [10] [48] [49] e é um software programado capaz de fazer cálculos da produção de seção de choque transversal e largura dos decaimentos, em Física de Partículas Elementares, utilizando a teoria da perturbação<sup>2</sup> em ordem mais baixa.

O programa possui uma interface gráfica interativa (*Graphical User Interface - GUI*) orientado por menu, que guia o usuário na escolha das opções desejadas, como o modelo de interação de partículas, a inserção de um processo específico, optar entre gerar diagramas de Feynman ou diagramas quadrados de Feynman, dentre outras possibilidades.

Por meio do ambiente interativo do CalcHEP o usuário pode, a partir da Lagrangiana, obter as distribuições finais de um evento de interação. Alguns modelos teóricos já vêm implementados no programa, como o MP, mas modelos além do modelo padrão podem ser encontrados no HEPMDB (*High Energy Physics Model Database*) [51].

Ainda é possível criar ou modificar modelos teóricos usando o editor gráfico interno ou um pacote externo disponível online [52]. A implementação pode ser feita, por exemplo, pelo pacote LanHEP [53] [54] que permite a geração das regras de Feynman [55] para propagadores e vértices das interações a partir da Lagrangiana.

São em três partes que o CalcHEP processa os cálculos simbólico, numérico e em batch. O módulo simbólico é uma sessão interativa que permite ao usuário, dentre outras funções, produzir códigos em linguagem de programação C para os elementos da matriz quadrada que são exportados em código C para serem usados nos cálculos numéricos. O módulo numérico determina a seção de choque transversal ou a largura do decaimento do processo definido pelo usuário a partir da integral de todo o espaço de fase. A sessão não interativa pode ser efetuada utilizando scripts fornecidos pelo usuário, o que inclui a parte batch. Nessa interface as informações, estabelecidas anteriormente pelo usuário, são lidas

---

<sup>2</sup>A partir de Hamiltonianos simples conhecidos, é possível descrever sistemas mais complexos e reais através de uma aproximação partindo da teoria da perturbação [50]. Por exemplo, o problema do átomo de hidrogênio possui Hamiltoniano  $H_0 = \mathbf{p}^2/2m - e^2/r$ , mas ao considerarmos a sua presença em um campo elétrico ou magnético externo, o Hamiltoniano passa a ser  $H = H_0 + \lambda V$ . O potencial  $V$  é devido ao campo e  $\lambda$  é um parâmetro real contínuo, que varre valores de 0 a 1, acompanhando o número de vezes que a perturbação, devido ao campo externo, atua. O CalcHEP utiliza um método de aproximação chamado aproximação de Born que calcula a amplitude de espalhamento para diversas funções potenciais de energia.

e executadas com o cálculo dos processos de produção e decaimentos automatizados. Os resultados dos processos de produção são combinados com os decaimentos para produzir um arquivo de evento final no formato LHE (*Les Houches Event*) [56]. O arquivo final LHE também pode ser utilizado em outros softwares como, por exemplo, o DELPHES que será assunto da próxima seção.

#### 4.5.2 DELPHES e PYTHIA

O objetivo do programa DELPHES é permitir a simulação de um detector multiuso para estudos fenomenológicos [57]. O DELPHES recebe como entrada os eventos gerados mais comuns e como saída são formados os dados através de uma simulação rápida e realística de um detector geral. A partir da resposta simulada do detector, objetos físicos (como, massa de partículas, momento, energia), que podem ser usados para análise de dados, são reconstruídos. O simulador DELPHES é para colisores de hádrons, mas pode ser adaptado para experimentos de colisores de elétrons e pósitrons.

O PYTHIA8 [58] é um programa, integrado ao DELPHES, que simula o decaimento de partículas instáveis e descreve o modelo de hadronização (decaimento das partículas).

# Capítulo 5

## Modelo ZP-TP-DM

Apesar de não conhecermos muitas informações físicas a respeito da matéria escura (ME) como, por exemplo, o spin, existem modelos que descrevem possíveis candidatos a ME.

Na literatura encontramos estudos de modelos que consideram candidatos para a ME que podem ser férmions de Dirac ou de Majorana com spin  $1/2$ , vetoriais com spin  $1$  ou escalares com spin  $0$ . Muitos desses modelos incluem a produção dos candidatos a ME vindos do decaimento de duas novas partículas (partícula e antipartícula) chamadas de "*parceiros top*" que não são previstas pelo MP. Dessas novas partículas, há a produção de pares  $t\bar{t}$  (partículas previstas pelo MP) e duas partículas de ME.

De acordo com as regras de adição de momento angular, portanto, podemos ter os seguintes processos como apresentado em [59]:

- parceiros top ( $\tilde{t}\tilde{t}^*$ ) com spin  $0$  e ME com spin  $1/2$ :  $pp \rightarrow \tilde{t}\tilde{t}^* \rightarrow t\bar{t} + \tilde{\chi}^0\tilde{\chi}^0$ ;
- parceiros top ( $T'\bar{T}'$ ) com spin  $1/2$  e ME com spin  $1$ :  $pp \rightarrow T'\bar{T}' \rightarrow t\bar{t} + V^0V^0$ ;
- parceiros top ( $T'\bar{T}'$ ) com spin  $1/2$  e ME com spin  $0$ :  $pp \rightarrow T'\bar{T}' \rightarrow t\bar{t} + \phi\phi$ .

$\tilde{\chi}^0$  representa um candidato no estado fermiônico,  $V^0$  e  $\phi$  representam os candidatos vetorial e escalar para ME, respectivamente.

O modelo utilizado nesse trabalho se baseia na fenomenologia de um modelo simplificado que incorpora um novo bóson de calibre  $Z'$  (spin inteiro), uma nova partícula  $T'$  e antipartícula  $\bar{T}'$  (spin  $1/2$ ) com paridade negativa e um escalar  $\phi$  candidato a matéria escura que também possui paridade negativa. A estabilidade de  $\phi$  é protegida por uma

simetria discreta e desconhecida. No total, existem três processos de produção, um via  $Z'$  e outros dois a partir da produção QCD como mostra a figura 5.1.

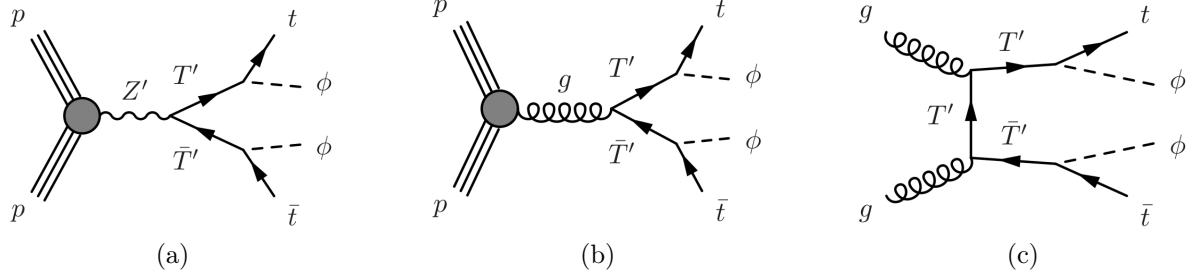


Figura 5.1: Diagramas de Feynman para a produção de  $t\bar{t}\phi\phi$  pelo decaimento de  $T'\bar{T}'$  vindo de (a)  $Z'$  e (b e c) dos glúons retirado de [8].

Estudamos, particularmente, o processo  $pp \rightarrow T'\bar{T}' \rightarrow t\bar{t} + \phi\phi$  em que o par  $T'\bar{T}'$  vem do decaimento do  $Z'$  ou da produção QCD dando a assinatura  $t\bar{t} + \cancel{E}_T$ .

## 5.1 Fenomenologia

Para estudar os produtos da colisão próton-próton, usaremos o modelo efetivo simplificado ZP-TP-DM (onde ZP chamamos de  $Z'$ , TP de  $T'$  e DM é abreviação do inglês *dark matter* para matéria escura) disponível em [60]. A Lagrangiana desse modelo [8] é dada por:

$$\mathcal{L} = \mathcal{L}_{MP} + \mathcal{L}_{kin} + \mathcal{L}_{Z'q} + \mathcal{L}_{Z'\ell} + \mathcal{L}_{Z'Q'} + \mathcal{L}_{\phi Q'} - V_\phi, \quad (5.1)$$

onde  $\mathcal{L}_{MP}$  é a Lagrangiana do Modelo Padrão apresentada anteriormente. Os outros termos, adicionais ao MP, formam a Lagrangiana do ZP-TP-DM. Para  $\mathcal{L}_{kin}$ , temos que:

$$\begin{aligned} \mathcal{L}_{kin} = & -\frac{1}{4}(\partial_\mu Z'_\nu - \partial_\nu Z'_\mu)(\partial^\mu Z'^\nu - \partial^\nu Z'^\mu) + \\ & + \frac{M_{Z'}^2}{2} Z'_\mu Z'^\mu + \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \frac{m_\phi^2}{2} \phi^2 + \\ & + \bar{T}'_s (i\not{D} - M_{T'_s}) T'_s + \bar{Q}'_d (i\not{D} - M_{T'_d}) Q'_d, \end{aligned} \quad (5.2)$$

onde  $T'_s$  é um singlete SU(2) e  $Q' = (T'_d, B'_d)^1$  é um dubleto SU(2). A massa da ME é representada por  $m_\phi$ . A Lagrangiana

$$\mathcal{L}_{Z'q} = \lambda_{Z'q\bar{q}} Z'_\mu (\bar{q} \gamma^\mu q) \quad (5.3)$$

<sup>1</sup> De acordo com [8], para o par de dubleto  $T'_d$  é esperado que tenha uma fenomenologia parecida a  $T'_s$ .

representa a interação do bóson  $Z'$  com os quarks do MP.  $\lambda_{Z'q\bar{q}}$  é a constante de acoplamento de  $Z'$  entre  $q\bar{q}$ . O quarto termo pode ser escrito como

$$\mathcal{L}_{Z'\ell} = \lambda_{Z'\ell+\ell-} Z'_\mu (\bar{\ell} \gamma^\mu \ell), \quad (5.4)$$

que vai representar a interação dos léptons do MP com  $Z'$ .  $\lambda_{Z'\ell+\ell-}$  é a constante de acoplamento de  $Z'$  com os léptons. O quinto termo representa as interações entre  $Z'$  e  $Q'$ , e é descrita por:

$$\begin{aligned} \mathcal{L}_{Z'Q'} &= \lambda_{Z'T'_s\bar{T}'_s} Z'_\mu (\bar{T}'_s \gamma^\mu q) + \\ &+ \lambda_{Z'T'_d\bar{T}'_d} Z'_\mu (\bar{T}'_d \gamma^\mu T'_d) + \\ &+ \lambda_{Z'T'_d\bar{T}'_d} Z'_\mu (\bar{B}'_d \gamma^\mu B'_d), \end{aligned} \quad (5.5)$$

onde  $\lambda_{Z'T'_s\bar{T}'_s}$  e  $\lambda_{Z'T'_d\bar{T}'_d}$  são os acoplamentos para  $Z'$  e  $T'_s\bar{T}'_s$  e para  $Z'$  e  $T'_d\bar{T}'_d$ , respectivamente.

A Lagrangiana  $\mathcal{L}_{\phi Q'}$  mostra como se dão as interações entre  $Q'$  com a matéria escura:

$$\begin{aligned} \mathcal{L}_{\phi Q'} &= \lambda_{\phi T'_s t} \phi \bar{t} T'_s + \lambda_{\phi T'_d t} \phi \bar{t} T'_d + \\ &+ \lambda_{\phi T'_d t} \phi \bar{b} B'_d + h.c., \end{aligned} \quad (5.6)$$

onde  $\lambda_{\phi T'_s t}$  e  $\lambda_{\phi T'_d t}$  são as constantes de acoplamento entre  $\phi$  e  $T'_s t$  e entre  $\phi$  e  $T'_d t$ , respectivamente.

O último termo da Lagrangiana do modelo simplificado é o potencial da partícula  $\phi$ , isto é, o potencial da matéria escura com um termo de interação com o bóson de Higgs:

$$V_\phi = \frac{\lambda_\phi}{4!} \phi^4 + \frac{\lambda_{\phi H}}{2} \phi^2 \left( |H|^2 - \frac{v^2}{2} \right), \quad (5.7)$$

onde  $\lambda_\phi$  é a constante de acoplamento da ME e  $\lambda_{\phi H}$  é a constante de acoplamento entre  $\phi$  e H (que é o bóson de Higgs), responsável pela quebra espontânea de simetria de calibre.  $v$  é um valor esperado do vácuo (VEV).

A configuração completa do espaço de parâmetros relevantes do modelo ZP-TP-DM são:

$$M_{T'_s}, m_\phi, M_{Z'}, \lambda_{Z'q\bar{q}}, \lambda_{Z'T'_s\bar{T}'_s} \quad (5.8)$$

De acordo com os valores já determinados por [8], as constantes de acoplamento têm valores:  $\lambda_{Z'q\bar{q}} = 0.25$ ,  $\lambda_{Z'T'_s\bar{T}'_s} = 2.5$ ,  $\lambda_{\phi H} = 0$ ,  $\lambda_{\phi T'_s t} = 1$  e  $\lambda_{Z'\ell+\ell-} = \lambda_{Z'q\bar{q}}$ . Além



disso, tomamos  $\lambda_\phi = 0.1$ .

## 5.2 Eventos de Sinal e de Fundo

Os eventos de sinal são os eventos importantes em uma análise de um processo. O processo  $pp \rightarrow T'\bar{T}' \rightarrow t\bar{t}\phi\phi$ , de acordo com o modelo ZP-TP-DM, será o nosso evento de sinal.

Os eventos ditos de fundo são aqueles vindos do MP que podem ser confundidos com o nosso sinal. O processo do MP escolhido foi  $pp \rightarrow t\bar{t}Z$  e por motivo de simplicidade, escolheremos para os eventos de fundo o decaimento do  $Z$  do MP em neutrinos e seus respectivos anti-neutrinos ( $Z \rightarrow \nu_e\bar{\nu}_e$ ,  $Z \rightarrow \nu_\mu\bar{\nu}_\mu$  e  $Z \rightarrow \nu_\tau\bar{\nu}_\tau$ ).

De modo geral, os eventos de fundo dominantes são [9]:  $t\bar{t}$ , único top  $Wt$ ,  $t\bar{t}+Z$  ( $\rightarrow \nu\bar{\nu}$ ) e  $W$ +jatos. Outros eventos ainda podem ser considerados, como: Dibósons,  $t\bar{t}+W$ ,  $Z$ +jatos e multijatos.

## 5.3 Predição Teórica da Seção de Choque e Distribuições Cinemáticas

Utilizando os parâmetros citados na seção 5.1 geramos as seções de choque teóricas para o modelo ZP-TP-DM. A partir do CalcHEP, portanto, foram gerados os eventos do canal  $pp \rightarrow T'\bar{T}' \rightarrow t\bar{t}\phi\phi$  com energia de centro de massa igual a  $13TeV$ . Cada curva teórica da Figura 5.2 conta com eventos de QCD+ $Z'$  para valores de  $M_{Z'} = 2500GeV$ ,  $M_{Z'} = 3000GeV$ ,  $M_{Z'} = 3500GeV$ ,  $M_{Z'} = 4000GeV$ ,  $M_{Z'} = 4500GeV$  e  $M_{Z'} = 5000GeV$ . A curva que representa os eventos com apenas QCD não tem a produção de  $Z'$ . O pico das seis curvas QCD+ $Z'$  indica a produção dos pares  $T'\bar{T}'$  vindos do decaimento de  $Z'$ . Na teoria quântica de campos as partículas que satisfazem a relação energia-momento  $E^2 = \mathbf{p}^2 + m^2$ , são chamadas de on-shell e as que não satisfazem são denominadas off-shell [61]. A região on-shell no gráfico é delimitada pelos picos. Por exemplo, na curva QCD+ $Z'$  com  $M_{Z'} = 2500GeV$  com  $M_{T'}$  até  $1250GeV$  a região é on-shell depois disso a região passa a ser off-shell. No gráfico observamos também que quanto maior o valor de  $M_{Z'}$ , menor é a seção de choque de produção para os pares  $T'\bar{T}'$ . As curvas QCD+ $Z'$  com  $M_{Z'} = 4000GeV$ ,  $4500GeV$  e  $5000GeV$  os picos que representam a produção de  $Z'$

decaindo em  $T'\bar{T}'$  estão fora da escala do gráfico.

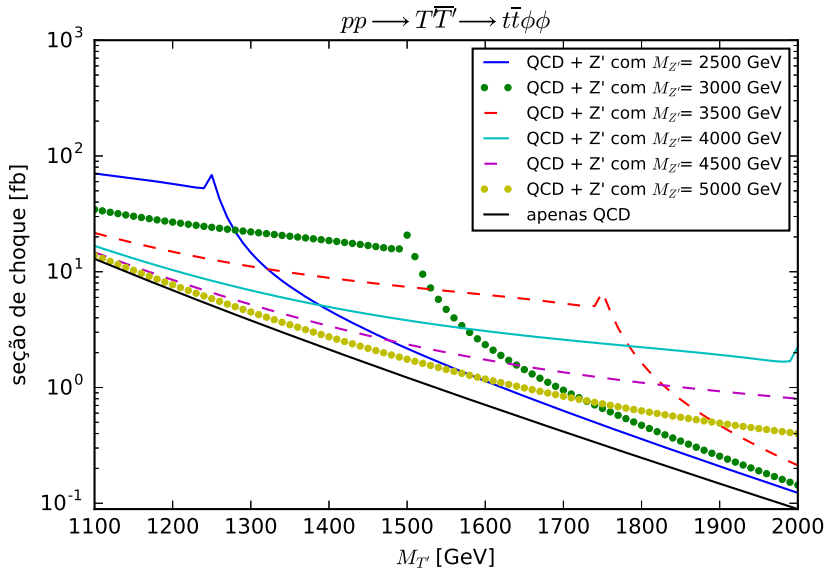


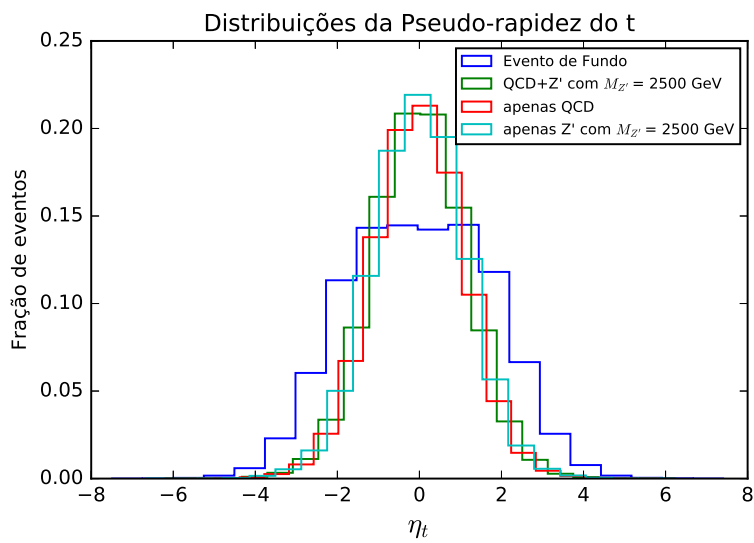
Figura 5.2: Seções de choque teóricas do modelo para diferentes valores de  $M_{Z'}$  e  $M_{T'}$ .

É impossível que todos os observáveis sejam medidos em colisores de altas energias [62]. Enquanto algumas características podem ser mensuradas pelos experimentos, outras podem apenas serem obtidas através da interpretação teórica dos resultados experimentais. A pseudo-rapidez e momento transversal, por exemplo, são obtidos através dos experimentos em colisores, pois são descritos em função de  $\theta$ , o ângulo de saída das partículas visíveis (detectáveis) após a colisão. Grandezas como a massa transversal das partículas e energia transversal faltante podem ser estimadas a partir de valores obtidos experimentalmente.

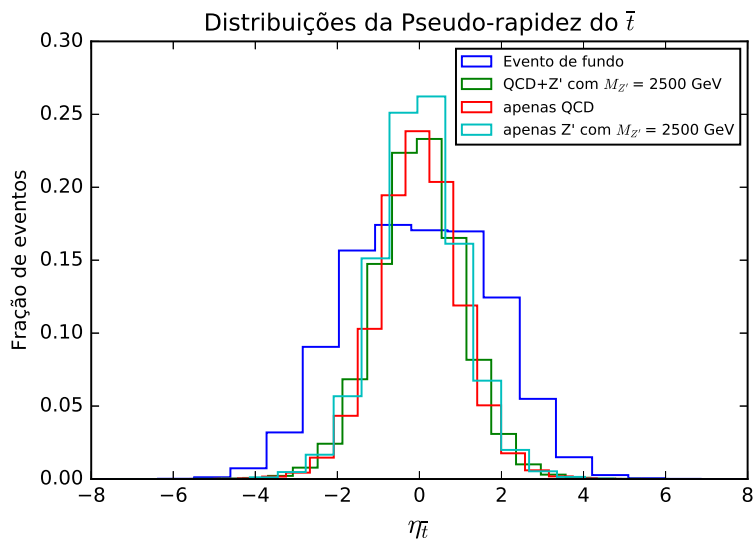
Para  $M_{Z'} = 2500\text{GeV}$  e  $M_\phi = 200\text{GeV}$  plotamos os gráficos da pseudo-rapidez ( $\eta$ ), momento transversal ( $p_T$ ), massa transversal ( $M_T$ ) do par  $t\bar{t}$  e energia transversal faltante ( $\cancel{E}_T$ ). Notamos que os eventos de fundo e sinal apresentam distribuições bem diferentes. Na figura 5.3 vemos as distribuições normalizadas da pseudo-rapidez do quark e antiquark top dos eventos de fundo  $pp \rightarrow t\bar{t}Z$  para a produção QCD+ $Z'$ , apenas QCD e produção somente de  $Z'$ . As Figuras 5.4a e 5.4b mostram o momento transversal do  $t$  e  $\bar{t}$ , respectivamente. Por fim, os gráficos da Figura 5.5 apresentam a massa transversal do par  $t\bar{t}$  e a energia transversal faltante do processo. Fazer uma análise tradicional dos gráficos obtidos significaria utilizar cortes de variáveis que têm como objetivo escolher pontos no gráfico que possuem mais eventos de sinal do que eventos de fundo. Como exemplo disso, aplicando esse método na Figura 5.5b, poderia ser escolhido um corte de  $\cancel{E}_T > 200\text{GeV}$ . Após isso, os gráficos da pseudo-rapidez, momento transversal e massa transversal faltante

deveriam ser plotados novamente com esse valor de corte em  $\cancel{E}_T$ . Passos sucessivos como esse possivelmente deverá ser feito de maneira que seja encontrado o melhor valor de corte. Devido à quantidade de passos repetidos, a utilização de cortes retangulares de variáveis com os gráficos citados seria mais trabalhoso.

De maneira a otimizar o trabalho da separação de eventos de sinal e fundo, utilizaremos nessa dissertação técnicas de Aprendizado de Máquina (AM) para tal objetivo. Portanto, no próximo capítulo será abordado técnicas de AM que tornam possível a distinção dos eventos.



(a)



(b)

Figura 5.3: Distribuições normalizadas da fração do número de eventos em função da pseudo-rapidez do quark top (a) e pseudo-rapidez do antiquark top (b).

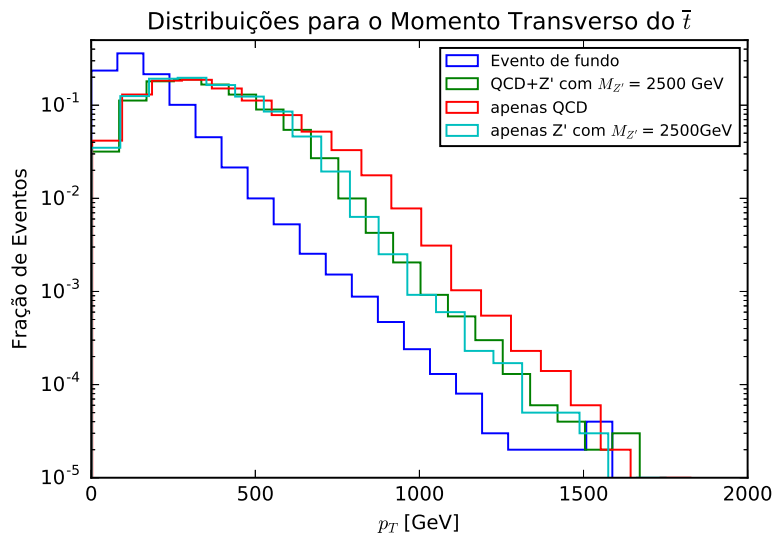
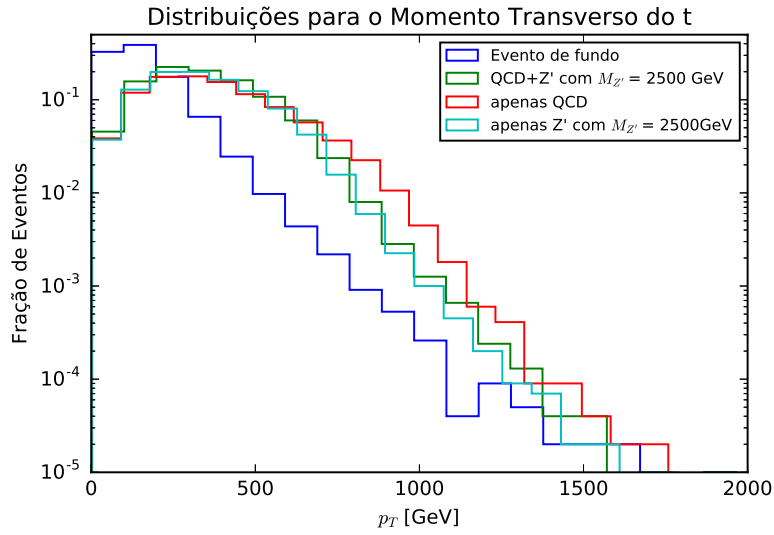
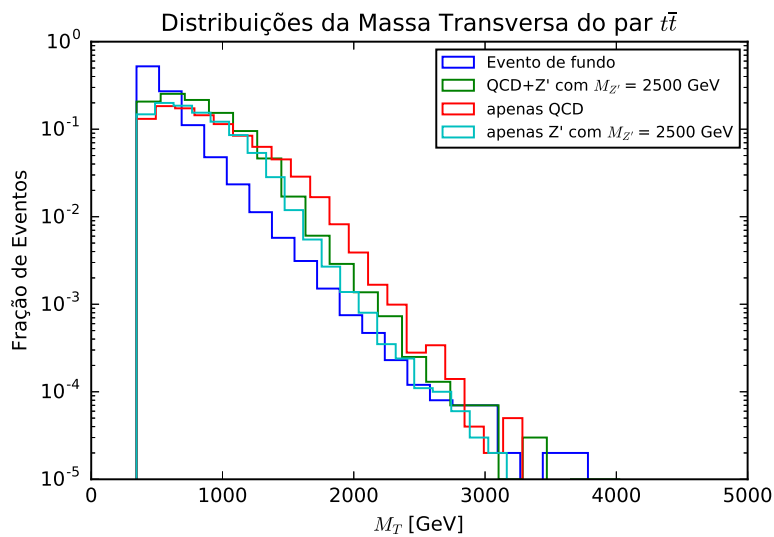
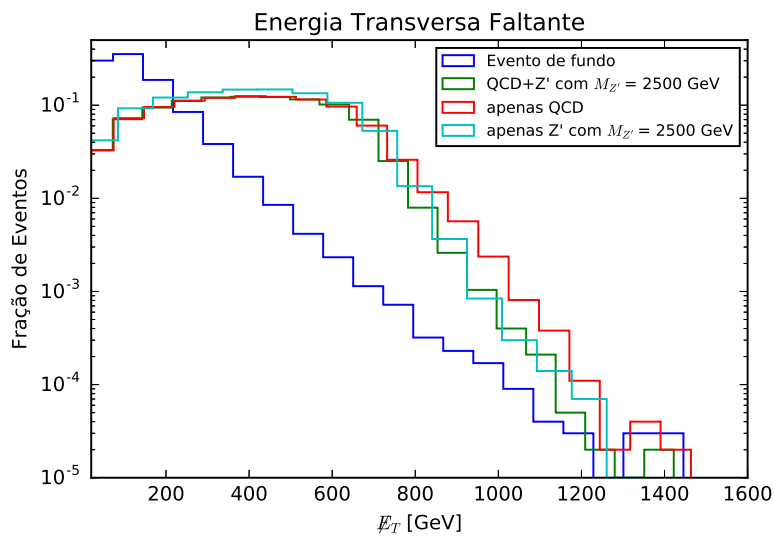


Figura 5.4: Distribuições normalizadas da fração do número de eventos na escala logarítmica em função do momento transverso do quark top (a) e momento transverso do antiquark top (b).



(a)



(b)

Figura 5.5: Distribuições normalizadas da fração do número de eventos na escala logarítmica em função da massa transversa do par  $t\bar{t}$  (a) e energia transversa faltante (b).

## Capítulo 6

# Aprendizado de Máquina

Uma das mais novas áreas de conhecimento é a Inteligência Artificial (IA) [63] [64] [12] [65] [66] que, formalmente iniciada em 1956, é caracterizada por qualquer técnica que permita os computadores imitarem as funções cognitivas dos humanos, isto é, a inteligência humana. Diversos ramos da ciência utilizam e aplicam os métodos de IA como, por exemplo, a Robótica.

O Aprendizado de Máquina (AM) [67] [68] [11] [69] é uma subcategoria da IA, cuja finalidade é fazer com que os programas de computadores possam aprender a partir de informações anteriores sem terem sido explicitamente programados. AM, portanto, é a parte estatística da IA que torna possível a resolução de problemas utilizando a experiência para resolver o mesmo problema em uma nova situação, com base em uma variedade de exemplos (dados fornecidos pelo usuário). A aplicação de AM tem crescido em todas as áreas do conhecimento, sendo as mais conhecidas, o reconhecimento de imagens faciais utilizado nos aeroportos, o reconhecimento de voz utilizando a SIRI da Apple e a recomendação de filmes na Netflix.

A área de Física de Partículas também tem se mostrado bastante apropriada para a introdução de métodos de AM. Em altas energias, especialmente, o grande volume e a grande complexidade dos dados coletados nas colisões  $pp$  (próton-próton) que ocorrem no LHC têm se mostrado como ideais para a aplicação dessas técnicas [70].

Nos experimentos do LHC a quantidade de dados coletados é demasiadamente grande. Cerca de  $1,1 \times 10^{11}$  prótons ou íons pesados sofrem colisões e cada colisão consegue produzir um grande número de partículas. Como os eventos que realmente produzem produtos significativos são raros, essa grande quantidade de dados é necessária.

Com o objetivo de selecionar da amostra de eventos registrados, informações com características semelhantes ao do sinal desejado, numa abordagem tradicional, a análise desses dados é dada em termos estatísticos devido às observações resultantes das interações das partículas serem probabilísticas. Dentro da abordagem tradicional estão incluídos:

- identificação de partículas e reconstrução dos seus momentos e energias (chamadas de tarefas de baixo nível).
- pesquisa por novas físicas e medições (chamadas de tarefas de alto nível).

Essas abordagens ditas tradicionais, em física de altas energias, são baseadas em cortes de certas grandezas físicas.

Mesmo que as abordagens tradicionais para reconstrução e seleção de eventos sejam boas, nas pesquisas que demandam grandes quantidades de dados, como nesses coletados no LHC, uma análise mais “sensível” desses dados é essencial para a busca de novas físicas. A aplicação das técnicas de AM pode ajudar nessas reduções de informações e descobrir padrões que não eram, a princípio, aparentes (eventos raros). Em AM isso é chamado de mineração de dados.

Uma das técnicas mais utilizadas em AM é a Aprendizagem Profunda (AP) que usa algoritmos<sup>1</sup> baseados em redes neurais que aprendem a partir de vastas quantidades de dados. Ferramentas como essas são muito utilizadas em Física de Partículas.

Redes neurais profundas (RNP) são versáteis e poderosas, ideais para lidar com tarefas altamente complexas, normalmente associadas à problemas de classificação. Algumas das principais aplicações em Física de Partículas das redes neurais profundas tem sido na identificação de jatos vindos de quarks pesados (c, b, t) ou quarks leves (u, d, s), na separação elétrons ou jatos e na seleção entre eventos de “sinal” e eventos de “fundo”.

Nas seções a seguir, abordaremos mais a fundo sobre as técnicas de AM e curva característica de operação do receptor (curva ROC - *Receiver Operator Characteristic curve*) e área sob a curva (AUC).

---

<sup>1</sup>Um algoritmo é uma sequência de passos lógicos (todas as tarefas que um computador executa, por exemplo) que visam obter uma solução para um determinado tipo de problema.

## 6.1 Técnicas de AM

De forma mais ampla, a AM é dividida em algumas categorias que se alicerçam em como o aprendizado é recebido ou como a resposta sobre o aprendizado é fornecido ao sistema computacional.

As técnicas mais conhecidas e utilizadas de AM são Aprendizado Supervisionado (AS) e Aprendizado Não-Supervisionado (ANS) e serão abordadas nas subseções a seguir. Por motivo de relevância, focaremos em dar mais detalhes na sessão 6.1.1. e no tópico sobre redes neurais artificiais.

### 6.1.1 Aprendizado Supervisionado

No Aprendizado Supervisionado [13] [71] é fornecido a um sistema de IA um conjunto de variáveis (a base de dados) como entrada que são rotuladas com saídas desejadas (essa é a chamada etapa de treino). Com essas variáveis de entrada é passado um valor numérico, constante ou não, e isso que é chamado de rótulo.

Esse método tem como propósito fazer com que o algoritmo tenha a capacidade de aprender, comparando sua saída atual com as saídas ensinadas, de maneira que se obtenha um melhor desempenho da tarefa a partir dos erros. Com isso, sabemos se a rede é capaz de determinar o valor do rótulo para cada variável a ser classificada. Portanto, com AS pode-se prever as variáveis de saída.

Um modelo de AM supervisionado pode ser criado utilizando técnicas de Classificação ou Regressão. A Classificação é o processo que prevê um rótulo ou uma categoria discreta de classe. Um problema de Classificação é quando a variável de resposta (saída) é uma categoria, como "sim" ou "não".



Figura 6.1: Exemplo da Classificação de spam em e-mails. Fonte: [67]

Na Figura 6.1 vemos um exemplo de como os dados são treinados na filtragem de spam. Há um conjunto de dados para treinamento, esses dados são e-mails que estão



categorizados com rótulos  $\otimes$  para spam e  $\checkmark$  para e-mails sem spam. Assim, o modelo criado aprende a classificar novos e-mails. Conseqüentemente, o algoritmo de AS será capaz de identificar posteriormente e-mails com spam não rotulados como  $\otimes$  e e-mails sem spam não tendo rótulos como  $\checkmark$ .

A técnica de Regressão envolve a previsão de um valor numérico, ou seja, a variável de saída é um valor real como, alguma característica: idade, preço de um carro, notas de alunos, etc.

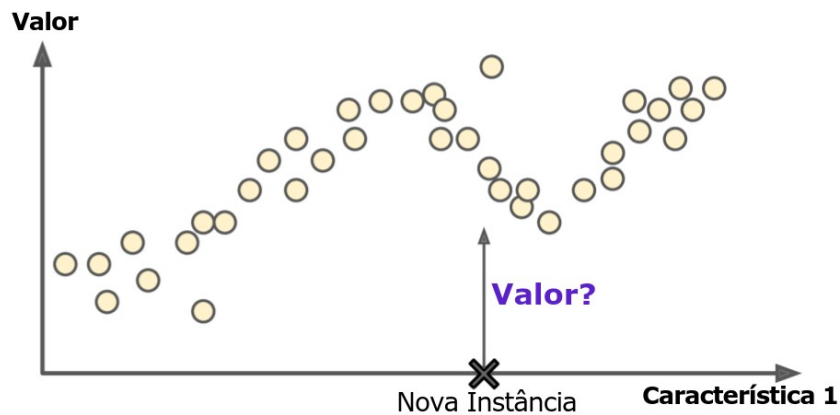


Figura 6.2: Regressão: dada uma **característica 1**, a partir dos dados (informação sobre os valores - eixo y), é possível prever eventos futuros estatisticamente prováveis sobre o valor que a **característica 1** assumirá. Fonte: [67]

A exemplo disso podemos imaginar alguém que deseja estimar valores do mercado de ações. Essa pessoa pode usar informações passadas do mercado de ações para antecipar flutuações futuras usando técnicas de Regressão. Na Figura 6.2, por exemplo, a **Característica 1** seria o mercado de ações, **Valor** representaria os dados de flutuações anteriores e a Nova Instância seria o valor num determinado momento previsto.

Existem muitos algoritmos de AS para Classificação e Regressão. Alguns deles são: Regressão Logística (RL), Árvores de Decisão (AD) e Redes Neurais Artificiais (RNA). Por motivos de interesse, nos tópicos a seguir iremos tratar RNA com mais detalhes.

## Regressão Logística

Regressão Logística prevê estatisticamente a probabilidade de um determinado evento acontecer. Por exemplo, se um cientista almeja antever a ocorrência de terremotos em determinada área (se haverá ou não), ele poderá utilizar algoritmos de RL. Se a probabilidade for maior que 50%, o modelo prevê que a instância pertence à classe positiva

(rotulada como 1). Se for menor que 50%, o modelo irá rotular como classe negativa (rotulado com 0). Essa saída é chamada de resultado binomial.

## Árvores de Decisão

Árvores de Decisão (AD) são algoritmos de AM capazes de treinar, visualizar e fazer previsões. Na mineração de dados, as AD funcionam como um modelo preditivo que esquematizam as informações sobre os dados para tirar conclusões acerca do valor final de saída com base nas variáveis de entrada.

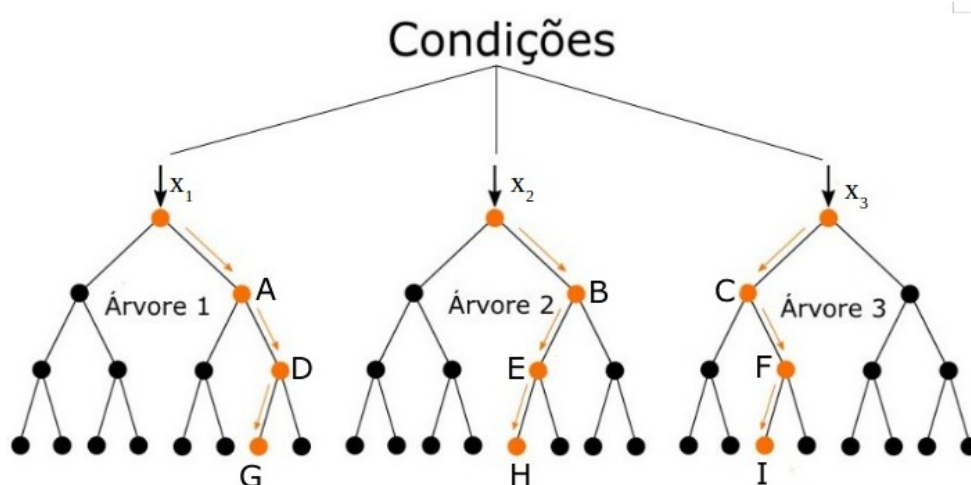


Figura 6.3: O esquema da figura compreende em três árvores de decisão treinadas independentemente.

Na Figura 6.3 vemos um exemplo de AD simplificado, onde um modelo classificador recebe dados de entrada diferentes  $x_1$ ,  $x_2$  e  $x_3$  (condições). A classe inferida de  $x_i$  é derivada das previsões individuais pela maioria de instâncias de determinada classe (A, B, C, D, E, F). A árvore classifica (prevê), portanto, qual instância é mais provável no último nó (G, H, I).

## Redes Neurais Artificiais

As Redes Neurais Artificiais [14] [72] foram criadas, inspiradas na analogia biológica dos neurônios inativos ou ativos do cérebro humano. As RNA são compostas por muitas camadas conectadas que combinam neurônios individuais como mostra a Figura 6.4.

Cada neurônio das camadas escondidas, também chamadas camadas intermediárias, recebe a saída de neurônios da camada anterior como uma soma linear ponderada e aplica a transformação não linear como a saída. A função de transformação não linear é chamada de função de ativação de um neurônio.

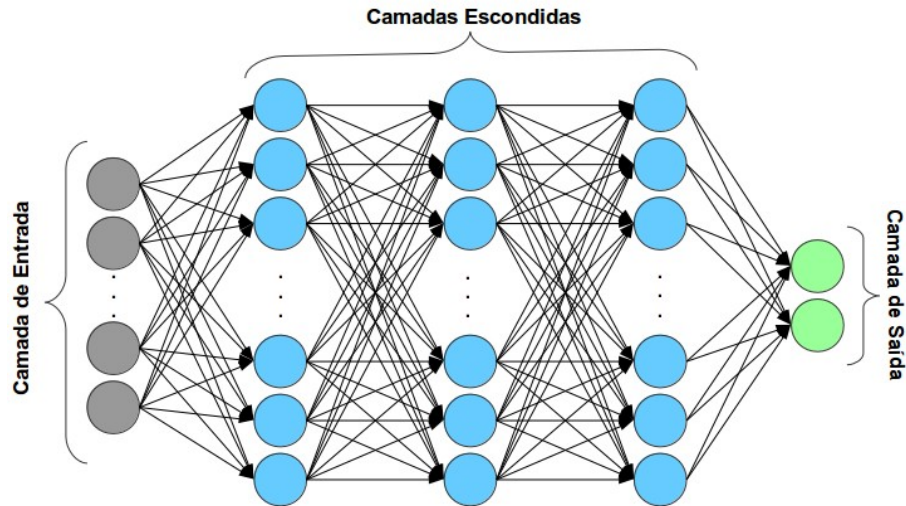


Figura 6.4: Arquitetura da RNA.

RNA com poucas camadas ocultas são referidas como Redes Neurais Artificiais Rasas (RNR), enquanto redes com muitas camadas ocultas são chamadas de Redes Neurais Artificiais Profundas (RNP).

Todo neurônio contém, basicamente, números de 0 a 1 que são chamados de ativação e o modo como a rede opera as ativações em uma camada determina as ativações da próxima camada. Assim, "padrões de ativação" (quando um neurônio é ativado) são propagados pela rede: um padrão de ativação causa muitos padrões de ativação específicos na próxima camada, na qual causa padrões na próxima até a última camada.

A transição de ativações de uma camada para outra é transmitida por meio de funções matemáticas. Portanto, as saídas das camadas de neurônios artificiais podem ser calculadas para várias instâncias, usando a equação:

$$X_{W,b} = \varphi\left(\sum_{i=1}^n W_i X_i + b\right) \quad (6.1)$$

onde  $i$  varia conforme a quantidade de camadas escondidas. A equação 6.1 computa as saídas de uma camada totalmente conectada  $X_{W,b}$  em que  $X$  representa a matriz de recursos de entrada,  $W$  é uma matriz de pesos que contém todos os pesos de conexão, exceto os pesos dos neurônios da última camada (neurônios de polarização),  $b$  é um vetor chamado *bias* que contém todos os pesos de conexão entre o neurônio de polarização e os neurônios artificiais e  $\varphi$  é chamada de função de ativação. Os pesos informam qual padrão da instância está sendo captado pelo neurônio de uma camada intermediária e, no início do treinamento, os pesos da rede neural são configurados aleatoriamente.

O vetor *bias* informa o quão alto o valor da soma ponderada  $\sum_{i=1}^n W_{0,i}X_i$  precisa ser antes que os neurônios comecem a se tornar significativamente ativos. O papel da função de ativação é garantir que a saída de cada camada escondida seja um valor de 0 a 1. Na Figura 6.5 vemos um recorte de uma RNA em que a camada  $X^{(1)}$  tem como saída  $X_0^{(1)} = \varphi(\sum_{i=1}^n W_{0,i}X_i + b_0)$ , pois é o único neurônio ativo desta camada. Essa saída será passada como uma entrada para a próxima camada e assim sucessivamente.

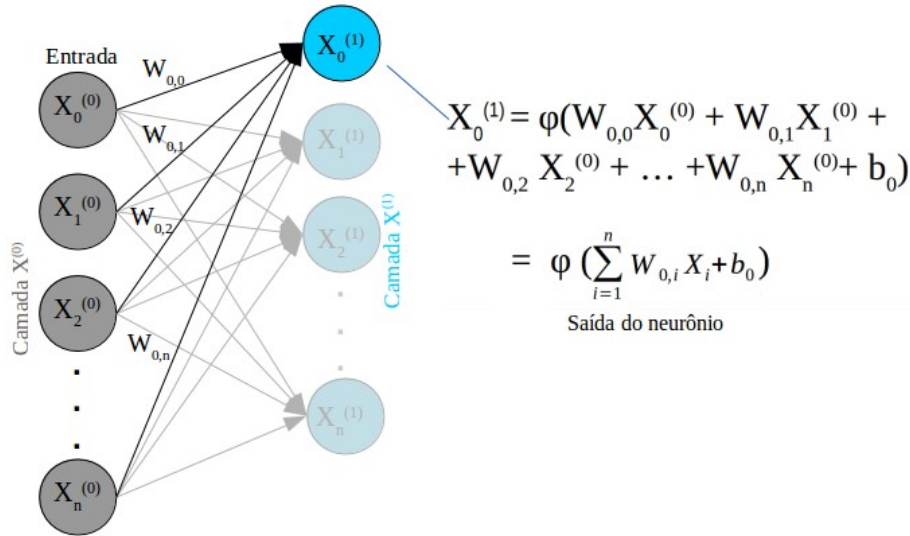


Figura 6.5: Esquema de um recorte de uma RNA, com  $N$  neurônios de entrada e  $N$  neurônios na primeira camada escondida, que ilustra como cada neurônio se conecta matematicamente. O neurônio ativado (em azul) recebe todas as informações dos pesos ( $W$ ) e recursos ( $X$ ) de cada neurônio da camada anterior, somado a um valor do *bias* ( $b_0$ ) e multiplicando por uma função de ativação  $\varphi$ .

Existem muitas funções de ativação, algumas delas são:

- **Função sigmoide:** É a função de ativação mais amplamente usada, pois é uma função não linear. A função sigmoide transforma os valores no intervalo de 0 a 1 (Figura 6.6). Sua expressão matemática é definida por:  $\sigma = 1/(1 + e^{-x})$
- **Função tangente hiperbólica:** É semelhante à função sigmoide, mas é simétrica em torno da origem. Isso resulta em diferentes sinais de saídas de camadas anteriores que serão alimentadas como entrada para a próxima camada. A função tangente hiperbólica é contínua e diferenciável, os valores variam entre -1 e 1 (Figura 6.6). A equação que descreve a  $\tanh$  é:  $g(x) = \tanh(x) = 2\sigma(2x) - 1$ .
- **Função Unidade Linear Retificada (The Rectified Linear Unit Function - ReLU):** É uma função de ativação não linear amplamente utilizada em redes neurais. A vantagem de usar a função ReLU (Figura 6.6) é que todos os neurônios

não são ativados ao mesmo tempo. Isso implica que um neurônio será desativado apenas quando a saída da transformação linear for zero. Sua expressão é dada por:  $ReLU(x) = \max(0, x)$ .

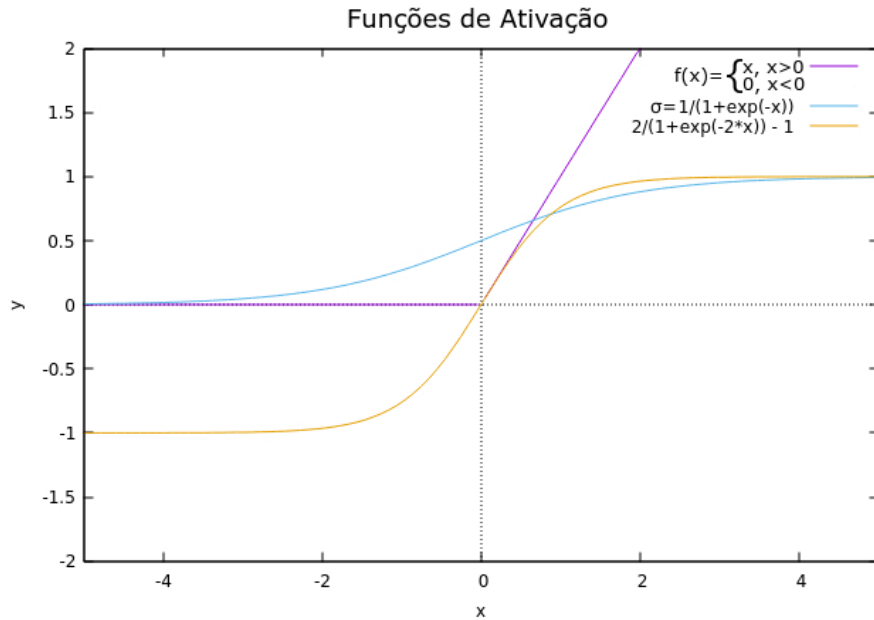


Figura 6.6: Funções de Ativação: a curva roxa é a função ReLU, a curva azul é a função de ativação  $\sigma$  e a curva laranja é a função tangente hiperbólica.

Para termos uma visão geral de como as informações chegam na última camada podemos reescrever a equação 6.1 como em 6.2: saída (nova entrada) = função de ativação([matriz peso].[matriz de recursos] + [vetor bias]).

$$X_{W,b} = \varphi \left( \begin{bmatrix} W_{0,0} & W_{0,1} & \cdots & W_{0,n} \\ W_{1,0} & W_{1,1} & \cdots & W_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ W_{k,0} & W_{k,1} & \cdots & W_{k,n} \end{bmatrix} \cdot \begin{bmatrix} X_0^{(0)} \\ X_1^{(0)} \\ \vdots \\ X_n^{(0)} \end{bmatrix} + \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_n \end{bmatrix} \right) \quad (6.2)$$

É muito comum utilizar como função de ativação para a última camada da rede neural como, por exemplo, a função chamada *softmax*. Geralmente as redes neurais com várias camadas possuem sua penúltima camada com valores de saída que flutuam em torno de números reais. Por exemplo, uma rede neural poderia ter como saída um vetor como (-0.2, 10.5), que não é uma distribuição de probabilidade válida. Por isso a função *softmax* é interessante, porque é capaz de converter as flutuações em uma distribuição de probabilidade normalizada, garantindo valores de 0 a 1 como retorno. A definição matemática da função *softmax* é dada por:

$$s(\vec{y})_i = \frac{e^{y_i}}{\sum_{j=1}^N e^{y_j}}, \quad (6.3)$$

onde  $\vec{y}$  é o vetor de entrada para a função *softmax*;  $y_i$  são os elementos do vetor de entrada para a função softmax e podem assumir qualquer valor real, positivo, zero ou negativo;  $e^{y_i}$  representa a função exponencial padrão aplicada a cada elemento do vetor de entrada; O termo  $\sum_{j=1}^N e^{y_j}$  é o termo de normalização garantindo que todos os valores de saída da função somarão um valor igual a 1 e cada valor estará no intervalo  $(0, 1)$ , constituindo assim uma distribuição de probabilidade válida;  $j$  é o número de classes do classificador.

Em um caso onde alguém queira construir um classificador que distinga imagens de roupas (classe 1) e sapatos (classe 2), por exemplo, com a função *softmax* a rede irá produzir um vetor  $[1, 0]$  quando for classe 1 e  $[0, 1]$  para classe 2.

Mesmo que treinemos bem nossos algoritmos, há perdas de informações nesse processo de aprendizagem da rede neural. Conseguimos contabilizar isso através de outras funções chamadas *funções de perda*. Essas funções são responsáveis por mensurar o desempenho de uma rede neural relativo a aprendizagem. Assim, quanto menor é a função de perda, maior é a precisão do seu classificador. Uma dessas funções é a *categorical-crossentropy* que é utilizada quando usamos a função de ativação *softmax* na última camada. Essa função perda é dada por:

$$Loss(y, \hat{y}) = - \sum_{i=0}^M \sum_{j=0}^N y_{i,j} \log(\hat{y}_{i,j}) \quad (6.4)$$

sendo  $y$  o valor observado e  $\hat{y}$  o valor predito pela rede que é a saída *softmax* da última camada. Essa função tem o papel de comparar as distribuições das predições, com a real distribuição em que a probabilidade da classe verdadeira é configurada como 1 e outras classes como 0. Assim, o erro é calculado com base nas funções de rede.

Na Figura 6.7 é mostrado como seria um gráfico da perda calculado pela rede. Vemos um valor entre 10% ~ 15% no treino e na fase de teste. Isso quer dizer que a acurácia da rede é de 85% ~ 90%. As épocas são uma das configurações das redes neurais chamadas de hiperparâmetros. O número atribuído para as épocas também é responsável por caracterizar um bom desempenho na separação de eventos. Esse valor pode ser ajustado

manualmente de modo que não aja um *underfitting*<sup>2</sup> ou um *overfitting*<sup>3</sup>.

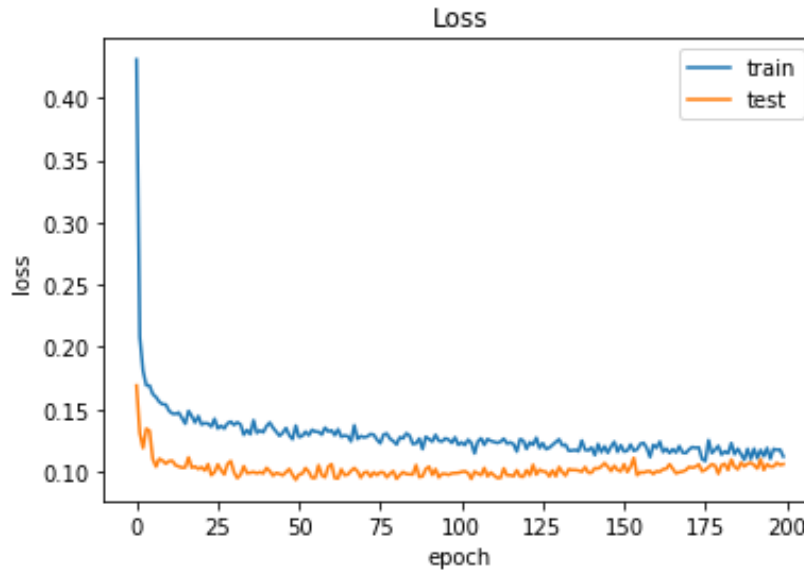


Figura 6.7: Gráfico de perda da rede.

Um método de AM que utiliza RNP é o chamado *Deep Learning* (Aprendizado Profundo) que explora redes mais profundas e podem aprender relações mais complexas, uma vez que cada camada adicional fornece uma representação mais complexa e abstrata dos dados.

O treinamento das redes neurais profundas com os parâmetros (características de entrada) é possível devido a grandes conjuntos de dados, que são necessários para treinar redes profundas. Isso acontece porque uma vez que abrangem uma enorme quantidade de parâmetros livres, tendem a se ajustar a pequenos grupos de dados.

## 6.1.2 Aprendizado Não-Supervisionado

O método de Aprendizado Não-Supervisionado (ANS) [73] [71] é definido por não rotular e não caracterizar os dados de treinamento. Isso significa que o sistema treina o conjunto de dados sem supervisão alguma e, por si só, encontra padrões ocultos entre seus dados de entrada.

Uns dos algoritmos de ANS mais importantes são *Clustering* (é a mais usada), Redução de Dimensionalidade (RD) e Detecção de Anomalias (DA).

<sup>2</sup>Caso em que a rede não aprende tudo o que poderia, tonando-se insuficiente.

<sup>3</sup>A rede, nesse caso, acaba aprendendo mais os ruídos (características que não se têm interesse) do que o sinal (uma nova instância).

*Clustering* (Agrupamento) é usado com o objetivo de agrupar pontos de dados que possuem uma similaridade. Na figura 6.8 é mostrado um gráfico com altura em função do peso que são os atributos que descrevem determinado grupo de pessoas, por exemplo.

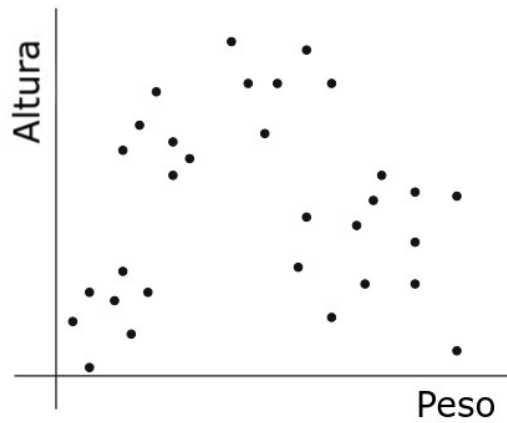


Figura 6.8: Clustering: agrupamento de características semelhantes num domínio bidimensional. Fonte: [68]

A abordagem de Redução de Dimensionalidade, também chamada seleção de característica, tem como finalidade reduzir a quantidade de atributos que se repetem para auxiliar a classificação. Assim, algoritmos desse tipo conseguem simplificar os dados sem perder muitas informações e extrair o recurso desejado. Na Figura 6.9, vemos uma representação da RD que representa os dados de baixa dimensão eficientes.

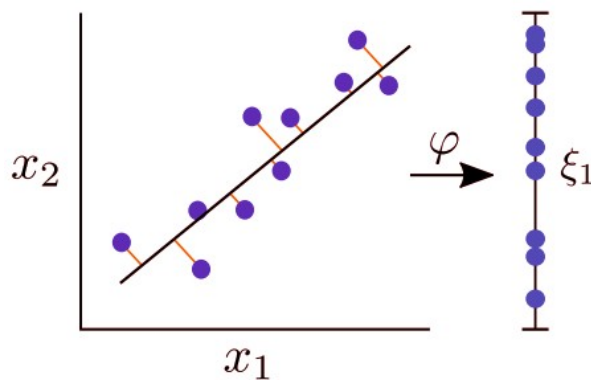


Figura 6.9: Redução de Dimensionalidade: a partir de uma transformação  $\varphi$  adequada, os algoritmos de RD tentam representar dados com dimensão alta ( $x_1 \times x_2$ ) em dados com dimensão baixa ( $\xi_1$ ). Fonte: [71]

Por fim, as técnicas de Detecções de Anomalias tem por objetivo treinar o sistema com instâncias ditas "normais" e, ao ver uma nova instância, pode dizer se ela parece normal ou se é uma provável anomalia como mostra a Figura 6.10.



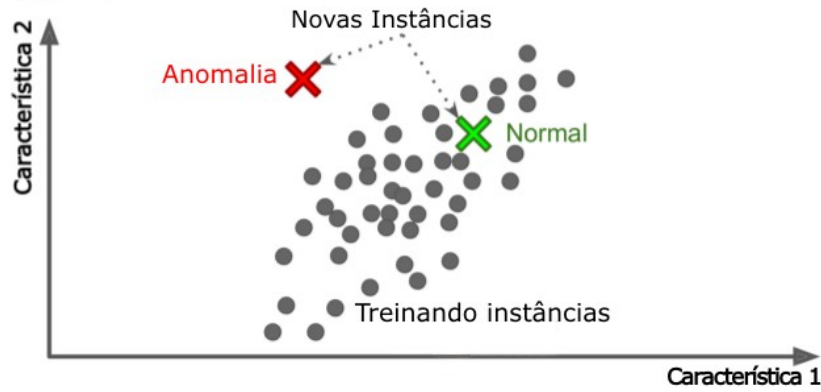


Figura 6.10: Detecção de Anomalias. Fonte: [67]

## 6.2 Curva ROC e AUC

A curva característica de operação do receptor (curva ROC - *Receiver Operating Characteristic Curve*) é uma ferramenta muito utilizada em classificadores binários e mostra o desempenho de um modelo de classificação. Esta curva traça a taxa de verdadeiro positivo (*True Positive Rate* - TPR) em relação a taxa de falso positivo (*False Positive Rate* - FPR). Esses parâmetros são dados por:

$$TPR = \frac{VP}{VP + FN} \quad (6.5)$$

onde VP = verdadeiro positivo e FN = falso negativo.

$$FPR = \frac{FP}{FP + VN} \quad (6.6)$$

onde FP = falso positivo e VN = verdadeiro negativo.

A Figura 6.11 mostra uma curva ROC típica.

Dado um determinado problema ou tarefa (como, por exemplo, a busca por uma nova assinatura de partículas em eventos de colisão), o conjunto de dados pode ser formado através de simuladores como o CalHEP, PYTHIA e DELPHES. Pois, como vimos, essas ferramentas podem ser usadas para gerar dados da mesma complexidade e formato dos dados reais da colisão que acontece em um colisor de hádrons. Então, para uma primeira amostra de dados podemos considerar os eventos simulados a partir da interação *pp*. Nesse caso, se tivermos o objetivo de separar essa nova assinatura (*signal*) dos eventos já conhecidos (*background* - eventos de fundo previstos pelo modelo padrão), podemos utilizar a curva ROC para saber o quão eficiente é o classificador criado. No exemplo

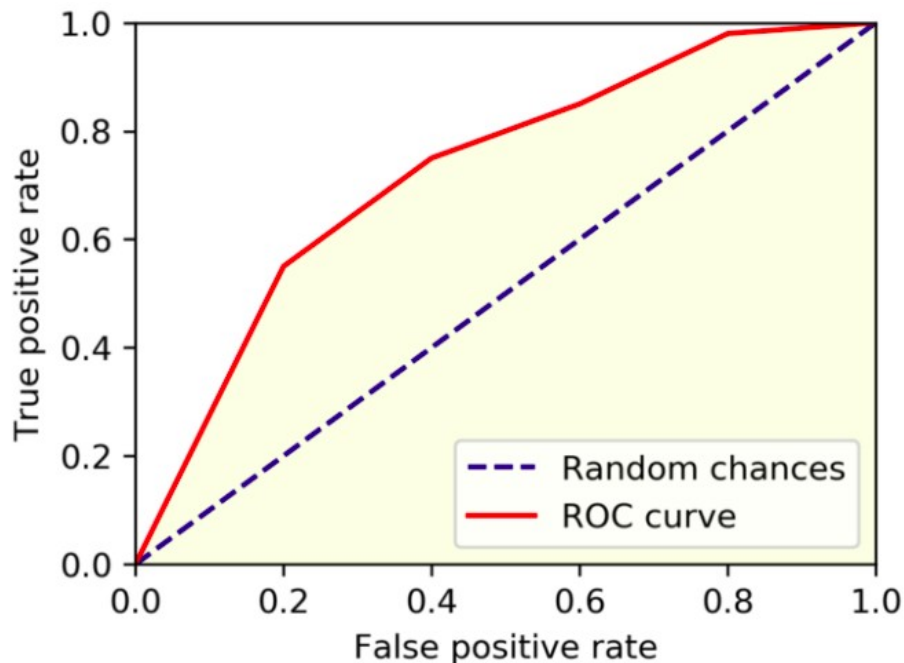


Figura 6.11: Curva ROC: A curva em vermelho caracteriza a eficiência do classificador criado e a linha tracejada em azul é uma curva padrão que delimita a eficiência geral de classificadores. Quanto mais afastadas as curvas ROC são da linha tracejada (área abaixo da curva  $\approx 1$ ), maior a probabilidade de que você tenha criado um bom classificador.

apresentado, os parâmetros citados acima seriam:

- VP = quantidade de eventos com sinal.
- VN = eventos sem sinal (eventos de fundo)
- FP = eventos que contém o sinal, mas são detectados como fundo (*background*).
- FN = eventos de fundo, mas detectados como sinal.

A área abaixo da curva ROC (AUC - *Area Under The Curve ROC*) mede toda a área bidimensional abaixo de toda a curva ROC. A AUC fornece uma medida agregada de desempenho em todos os limites de classificação possíveis. Uma maneira de interpretar AUC é como a probabilidade de que o modelo classifique um exemplo positivo aleatório mais alto do que um exemplo negativo aleatório.

### 6.3 Linguagens de Programação

A linguagem de programação Python é a mais utilizada em AM. Mas, outras linguagens como Java, R e C++ também são usadas. Sua importância nessa área dá-se pelo fato de

haver um maior desenvolvimento de estruturas de AM (algoritmos) disponibilizados em Python, como o TensorFlow e Keras, por exemplo.

TensorFlow é uma biblioteca de software que foi desenvolvida pela equipe do Google Brain e possui código aberto que pode ser usado para AM. Com essa biblioteca é possível treinar uma rede com até milhões de parâmetros. O Keras é uma ferramenta de alto nível e seus códigos estão escritos em Python e é executada na plataforma TensorFlow.

Algumas das bibliotecas de Python para AM são:

- ***numPy***: é uma biblioteca matemática para trabalhar com arrays (matrizes) N-dimensional. Ex.: matrizes, dados, funções, dicionários e imagens.
- ***SciPy***: coleção de algoritmos numéricos que inclui processamento de sinal, otimização, estatística.
- ***matplotlib***: pacote para plotar gráficos 2D e 3D.
- ***pandas(biblioteca)***: é uma biblioteca de alto nível do Python que fornece estruturas de dados de alto desempenho.
- ***Scikit-learn***: conjunto de algoritmos e ferramentas para AM, onde possui a maior parte dos algoritmos de Classificação e Regressão.

# Capítulo 7

## Análise dos Dados Simulados e Resultados

### 7.1 Implementação do Modelo

Em nosso estudo, utilizamos um modelo efetivo simplificado (ZP-TP-DM), abordado no Capítulo 5, que está disponível no HEPMDB [60]. Esse modelo considera um novo bosón de calibre  $Z'$  (com spin 1), que se acopla às partículas do MP e às novas partículas adicionais chamadas parceiros top  $T'\bar{T}'$ , que são férmions com spin 1/2. O par  $T'\bar{T}'$  pode se originar do decaimento de  $Z'$  ou da produção QCD.

Além disso, é incluído no modelo uma partícula escalar neutra  $\phi$  (spin 0) para a matéria escura (ME), que se acopla com o (anti)quark top do MP e com os parceiros top fermiônicos. A partícula  $\phi$  tem estabilidade protegida por uma simetria discreta. Para tal, é suposto uma nova simetria chamada *paridade ME* que não permite o decaimento da matéria escura. Portanto, assume-se que os parceiros top e  $\phi$  carreguem *paridade ME* negativa, enquanto o bóson  $Z'$  e as partículas do MP carreguem *paridade ME* positiva, tornando  $\phi$  uma candidata estável a matéria escura.

O modelo ZP-TP-DM foi implementado no software CalcHEP, a partir do pacote LanHEP, pelo grupo do Belyaev [8]. Usamos o CalcHEP com esse modelo para calcular as seções de choque e gerar os eventos de sinal do modelo ZP-TP-DM. O mesmo software foi utilizado para calcular as seções de choque e gerar os eventos de fundo do modelo padrão, que se confundem com o nosso sinal. Porém, antes de gerarmos os eventos, através da interface batch do CalcHEP, escolhemos todas as configurações necessárias

que serão detalhadas na próxima seção.

## 7.2 Eventos Simulados

Uma escolha muito importante para o estudo de colisões de hádrons é a definição da função de distribuição partônica<sup>1</sup> (PDF - *Parton Distribution Function*) [74] [75] [76]. As PDF's incluem informações da subestrutura dos prótons e são essenciais para determinar os parâmetros do Modelo Padrão e para fazer previsões para a produção de possíveis novas partículas pesadas em teorias além do Modelo Padrão. A PDF utilizada em nossa pesquisa foi NNPDF2.3QED [8].

Outros fatores importantes definidos foram a energia do centro de massa (ECM) da colisão e a luminosidade integrada. Para estarmos em concordância com o trabalho da *Colaboração ATLAS* citado na referência [9], utilizamos os mesmos valores adotados por eles da energia do centro de massa e luminosidade integrada. Portanto, tomamos  $\sqrt{s} = 13\text{TeV}$  e  $\mathcal{L} = 13,2\text{fb}^{-1}$  (equação 4.3). Esse valor para  $\mathcal{L}$  é a luminosidade integrada dos dados coletados pelo detector ATLAS. Seguindo a referência [8] fixamos as constantes de acoplamento do modelo (equação 5.1), como visto no Capítulo 5:

$$\lambda_{Z'q\bar{q}} = 0.25, \quad \lambda_{Z'T_s'\bar{T}'_s} = 2.5, \quad \lambda_{\phi H} = 0, \quad \lambda_{\phi T'_s t} = 1 \quad \text{e} \quad \lambda_{Z'q\bar{q}} = \lambda_{Z'\ell^+\ell^-} \quad (7.1)$$

Para a seleção do número de eventos de teste da rede utilizamos um valor proporcional à seção de choque presente no nosso arquivo de dados gerado a partir do CalcHEP. Por exemplo, se a seção de choque calculada foi  $\sigma_f = 202,25\text{fb}$  para os eventos de fundo e  $\sigma_s = 43,53\text{fb}$  para os eventos de sinal, aplicando a equação 4.2 (ver Capítulo 4), temos que  $N_f = 202,25 \times \mathcal{L}$  e  $N_s = 43,53 \times \mathcal{L}$  para o fundo e sinal, respectivamente. Então  $N_f \approx 2670$  eventos e  $N_s \approx 575$  eventos.

Os valores para as massas de  $M_{Z'}$  e  $M_{T'}$  e  $M_\phi$  foram escolhidos de acordo com [8] e serão discutidos na subseção 7.2.3.

---

<sup>1</sup>Em Física de Partículas, os hádrons são partículas compostas de pártons. Os quarks e os glúons são as partículas que chamamos de pártons.

## 7.2.1 Principais Eventos de Fundo

Os eventos de fundo escolhidos contribuem para o estado final de interesse que é a assinatura do LHC  $t\bar{t} + \cancel{E}_T$ , onde  $\cancel{E}_T$  é a energia transversa faltante. Em vista disso, como mencionado no Capítulo 5, por questão de simplicidade selecionamos para os eventos de fundo vindos do MP o canal  $pp \rightarrow t\bar{t}Z$ .

Os produtos originados do decaimento do bóson  $Z$  do MP devem possuir a soma de suas cargas elétricas iguais a zero. Isso porque o bóson  $Z$  é eletricamente neutro, então, para que haja conservação da carga elétrica, em todos os processos, o bóson  $Z$  tem que decair em um par de partícula e antipartícula. De todos os processos possíveis de  $Z$ , apenas os canais  $Z \rightarrow \nu_e \bar{\nu}_e$ ,  $Z \rightarrow \nu_\mu \bar{\nu}_\mu$  e  $Z \rightarrow \nu_\tau \bar{\nu}_\tau$  contribuem para a energia transversa faltante ( $\cancel{E}_T$ ), pelo motivo do neutrino não ser detectável pelos aparatos experimentais. Logo, o estado final do fundo se assemelha ao nosso estado final de interesse da colisão  $pp$  que é  $t\bar{t}\phi\phi$  (nosso sinal, onde  $\phi$  não é detectável, contribuindo para a assinatura do LHC).

O quark e antiquark top decaem em outras partículas que contribuem para o fundo. Os estados finais do principal processo de produção dos pares  $t\bar{t}$  podem ocorrer em três classes [77]:

1.  $t\bar{t} \rightarrow W^+ b W^- \bar{b} \rightarrow q\bar{q}' b q'' \bar{q}''' \bar{b}$ : 6 jatos no estado final (canal hadrônico).
2.  $t\bar{t} \rightarrow W^+ b W^- \bar{b} \rightarrow q\bar{q}' b \ell^- \bar{\nu}_\ell \bar{b}$  ou  $\ell^+ \nu_\ell b q'' \bar{q}''' \bar{b}$ : 4 jatos + 1 lépton + 1 neutrino (canal semi-leptônico).
3.  $t\bar{t} \rightarrow W^+ b W^- \bar{b} \rightarrow \ell^+ \nu_\ell b \ell'^- \bar{\nu}_{\ell'} \bar{b}$ : 2 jatos + 2 léptons com sinais diferentes + 1 neutrino + 1 anti-neutrino (canal leptônico).

Não fizemos o decaimento do quark top, mas consideramos o decaimento do bóson  $Z$  do MP em neutrinos e seus respectivos anti-neutrinos. Portanto, utilizamos somente os eventos de fundo  $t\bar{t}Z$  com os decaimentos de  $Z$  citados, pois são eventos de fundo que podem ser confundidos com o nosso sinal e também contribuem para a assinatura  $t\bar{t} + \cancel{E}_T$ . Contudo, como não iremos usar todos os eventos de fundo existentes, faremos uma correção de acordo com a referência [9].

Na Tabela 7.1, temos os canais de produção utilizados para o fundo como apresentado em [9]. Como já mencionado, em nosso trabalho, usamos apenas os eventos de fundo  $t\bar{t}Z$  o que corresponde ao canal de produção  $t\bar{t} + V$  da Tabela 7.1, onde o  $V$  é um bóson

de calibre. A quantidade de eventos mostrada (7.9 eventos) equivale a aproximadamente 33% dos 24 eventos de fundo totais. Portanto, o número de eventos que contribuem para o fundo na assinatura de interesse  $t\bar{t} + \cancel{E}_T$  é muito maior. Frente a isso, devemos multiplicar o número encontrado dos eventos de fundo estudados por um fator de correção. Concordando com a porcentagem de eventos que compõem o fundo conforme mostrado na Tabela 7.1, esse fator equivale a 3.03.

Canais de Produção	Número de Eventos de fundo
Single <i>top</i>	3.1
$t\bar{t}$	8.4
$t\bar{t} + V$	7.9
$W + \text{jatos}$	2.5
<i>Dibson</i>	1.2
$Z + \text{jatos}$	0.59
Total dos eventos de fundo	$\approx 24$

Tabela 7.1: Canais de produção utilizados na referência [9].

Quando calculamos as seções de choque no CalcHEP, somente os primeiros gráficos de mais baixa ordem são calculados (comumente chamados gráficos de nível árvore ou *Leading Order* - LO). Por isso, também é preciso fazer uma correção para ordens mais altas (*Next Leading Order* - NLO). Dessa forma, como indicado em [9], é aplicada uma correção à seção de choque para o processo  $t\bar{t} + Z$ . Essa correção tem o nome de fator  $k$  e é aproximadamente 1.3.

## 7.2.2 Distribuição de Variáveis para os Eventos de Sinal e de Fundo

Como já discutido anteriormente, esse trabalho se baseia na separação de eventos de fundo e sinal via aprendizagem de máquina (AM). Os processos de sinal podem ocorrer de acordo como indicado nos três diagramas de Feynman apresentados na Figura 7.1. O primeiro diagrama na Figura 7.1a mostra a produção dos parceiros top  $T'\bar{T}'$  por intermédio do bóson  $Z'$ . O par  $T'\bar{T}'$  ainda decai no estado final  $t\bar{t}\phi\phi$ . Nos diagramas 7.1b e 7.1c essa produção se dá por meio da interação QCD mediada pelos glúons. A assinatura  $t\bar{t} + \cancel{E}_T$ , portanto, pode ocorrer a partir de  $Z'$  ou pela produção QCD.

Elaboramos um programa em *Python* para selecionar as informações necessárias extraídas do nosso arquivo de dados gerado pelo CalcHEP, com a finalidade de calcularmos

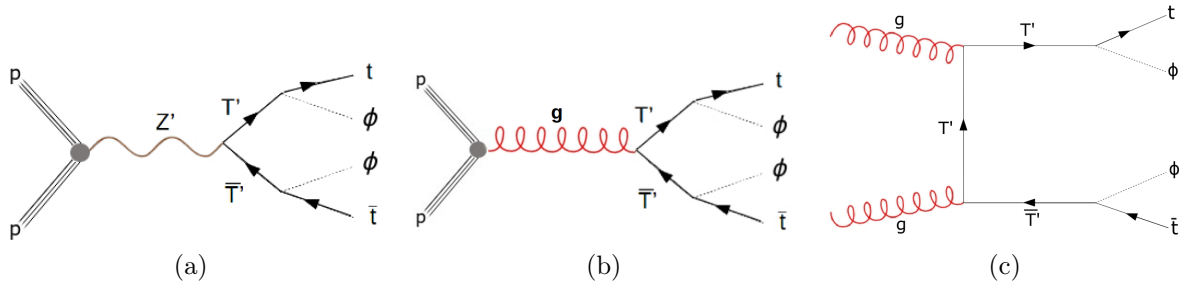


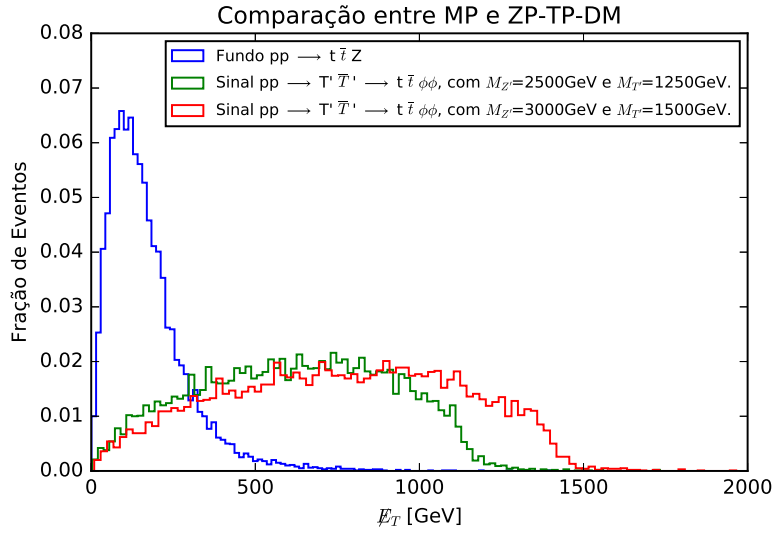
Figura 7.1: Produção do par  $T'\bar{T}'$  em pares  $t\bar{t}$  e duas partículas de ME ( $\phi\phi$ ) no estado final vindo do decaimento do bóson  $Z'$  (a) e glúons (b e c).

$\cancel{E}_T$  e a massa transversa  $M_T$ . Para o cálculo da energia transversa faltante em x e em y ( $\cancel{E}_{Tx}$  e  $\cancel{E}_{Ty}$ ) e da massa  $M_T$ , nós implementamos as equações 4.9 e 4.11, apresentadas no Capítulo 4.

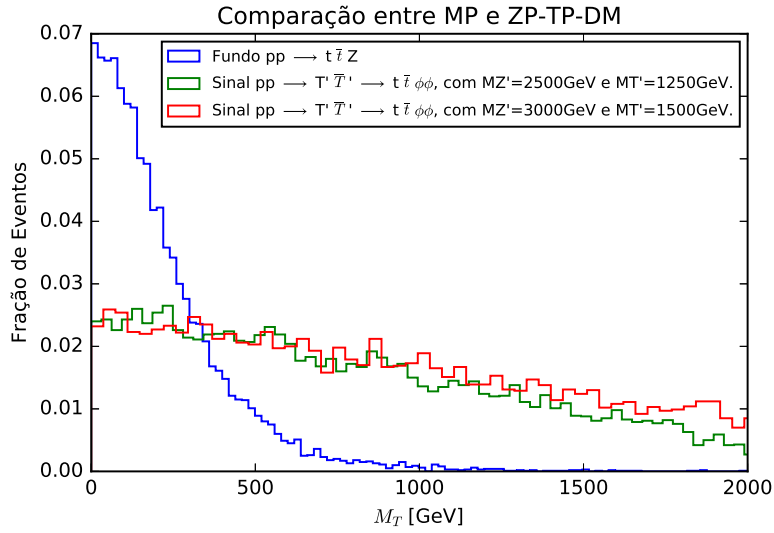
Nas Figuras 7.2a e 7.2b, mostramos os histogramas da  $\cancel{E}_T$  e  $M_T$ , respectivamente, para os eventos de fundo  $pp \rightarrow t\bar{t}Z$  e para os eventos de sinal  $pp \rightarrow T'\bar{T}' \rightarrow t\bar{t}\phi\phi$  com  $M_{Z'} = 2500\text{GeV}$ ,  $M_{T'} = 1250\text{GeV}$  e  $M_{Z'} = 3000\text{GeV}$ ,  $M_{T'} = 1500\text{GeV}$ . Ambos os gráficos estão normalizados. Em aproximadamente  $\cancel{E}_T = 300\text{GeV}$ , ponto em que há uma interseção entre as distribuições dos eventos de sinal e de fundo, vemos que para  $\cancel{E}_T > 300\text{GeV}$  grande parte dos eventos de fundo são eliminados, mas também é perdido parte dos eventos de sinal em  $\cancel{E}_T < 300\text{GeV}$ . Cortar o máximo do fundo, mantendo o máximo do sinal é característico do método tradicional baseado em cortes retangulares de variáveis. Não utilizamos esse recurso como meio de trabalho para separar os eventos do sinal e de fundo. Usamos técnicas de aprendizagem de máquina, como veremos na seção a seguir. Entretanto, observamos que as variáveis  $\cancel{E}_{Tx}$ ,  $\cancel{E}_{Ty}$  e  $M_T$  são boas para separar os eventos de sinal dos eventos de fundo, pois apresentam distribuições bem diferentes, como vemos na Figura 7.2.

Uma vez que, tanto os eventos do nosso sinal quanto os eventos de fundo resultam em estados finais  $t\bar{t} + \cancel{E}_T$ , isto é,  $t\bar{t}\phi\phi \rightarrow t\bar{t} + \cancel{E}_T$  (sinal) e  $t\bar{t}Z \rightarrow t\bar{t} + \cancel{E}_T$  (fundo), vamos utilizar essas 3 variáveis apresentadas ( $\cancel{E}_{Tx}$ ,  $\cancel{E}_{Ty}$  e  $M_T$ ), visto que carregam informações da matéria escura e/ou dos neutrinos. Outras 10 características que ajudarão nossa RNP a aprender eventos raros e separá-los do fundo são, massa, energia e momento ( $\mathbf{p}_x$ ,  $\mathbf{p}_y$ ,  $\mathbf{p}_z$ ) dos (anti)quarks  $t$  e  $\bar{t}$ , totalizando 5 variáveis para cada (anti)quark top. Comumente, todas essas grandezas são usadas quando existem partículas invisíveis ao detector no estado final, como no nosso caso. Por esse motivo, usamos as 13 variáveis citadas.





(a)



(b)

Figura 7.2: Comparação entre as distribuições normalizadas da energia transversa faltante (a) e massa transversa (b) dos eventos gerados pelo CalcHEP vindos do Modelo Padrão (linha azul) e do modelo ZP-TP-DM para  $M_{Z'} = 2500\text{GeV}$  e  $M_{T'} = 1250\text{GeV}$  (linha verde), e  $M_{Z'} = 3000\text{GeV}$ ,  $M_{T'} = 1500\text{GeV}$  (linha vermelha).

### 7.3 Treinamento e Teste da Rede

Estudamos como a utilização de redes neurais profundas podem colaborar na separação de eventos oriundos do modelo ZP-TP-DM, que prevê a existência de partículas ME, dos eventos de fundo.

Tivemos como ideia central construir portanto, um classificador a partir de técnicas de aprendizagem de máquina, baseado em uma RNP. Ao criá-la, portanto, treinamos a rede para discernir as características do sinal e de fundo diretamente dos dados simulados.

Implementamos a RNP a partir do pacote *Keras* (*tf.keras*), uma API (Application Programming Interface – Interface de Programação de Aplicativos) de alto nível para construir e treinar modelos de aprendizagem profunda. Isso foi realizado por meio da biblioteca do Tensorflow, que torna possível treinar e executar redes neurais muito grandes de forma eficiente.

A saída do nosso classificador foi usada para separar eventos semelhantes a sinais dos eventos de fundo. Assim, fornecemos como entrada para a rede um conjunto de variáveis importantes para distinguir os eventos. Em vista disso, podemos escrever a arquitetura da nossa RNP de modo simplificado, como:

$$N_{entrada}, N_{nós}, N_{nós}, N_{nós}, N_{nós}, N_{saída}. \quad (7.2)$$

Na arquitetura simplificada em 7.2,  $N_{entrada}$  é o número de características de entrada, todos os  $N_{nós}$  compõem as camadas escondidas, sendo cada  $N_{nós}$  os números de neurônios por camada escondida e  $N_{saída}$  representa o número de neurônios na saída.

Todas as 13 variáveis (massa e quadrimomentos do quark top e do antiquark top,  $\cancel{E}_{Tx}$ ,  $\cancel{E}_{Ty}$  e  $M_T$ ), citadas na seção anterior do atual Capítulo, são as nossas características de entrada ( $N_{entrada}$ ) e foram reescaladas para terem valores entre 0 e 1 antes de treinar.

A última camada da nossa RNP possui  $N_{saída} = 2$ , isto é, dois neurônios. Portanto, optamos rotular os eventos de sinal e de fundo como matrizes para combinar com as saídas da rede que classificamos como matrizes coluna, da seguinte forma:

$$sinal = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad fundo = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \quad (7.3)$$

O conjunto de dados usado para ensinar o modelo<sup>2</sup> a aprender foram os 20000 eventos gerados com o CalcHEP (10000 eventos de fundo + 10000 eventos de sinal) que contém todas as informações das características de entradas. Essa configura a etapa de treino.

Para avaliar o quão preciso nossa rede aprendeu a classificar os eventos, usamos o conjunto de teste que foi selecionado a partir do número de eventos proporcional a seção de choque, aplicando a equação 4.2.

Em todas as camadas escondidas (ver Figura 6.4 no Capítulo 6), utilizamos a função

---

<sup>2</sup>Daqui em diante, frequentemente iremos nos referir ao classificador criado para distinguir os eventos, como modelo.

de ativação ReLu (*Rectified Linear unit*), pois queremos que a saída seja sempre positiva. A RNP possui 4 camadas escondidas com 200 nós ( $N_{\text{nós}}$ ) cada. Com o objetivo de evitar o *overfitting* acrescentamos camadas extras entre as camadas secundárias. Essa é uma técnica chamada de *Dropout* que funciona reduzindo aleatoriamente o número de neurônios interconectados em uma rede neural, tornando cada neurônio suficiente independentemente. O parâmetro utilizado para o *Dropout* foi 0.2. Assim como em [70], estabelecemos a quantidade de épocas baseadas no cruzamento das curvas de acurácia, do treino e teste. Esse ponto define o valor que escolhemos para o número de épocas. Na última camada, adotamos a função de ativação *softmax*. A função de ativação *softmax* para a última camada garante que a resposta do neurônio "0"( $r_0$ ) e o neurônio "1"( $r$ ) satisfaça:

$$0 < r_0, r < 1, r_0 + r \equiv 1 \quad (7.4)$$

A distribuição  $r$  de uma rede bem treinada deve dar um pico em 1 para o sinal e 0 para o fundo, no treinamento e no teste.

Definimos a função de perda determinada como *categorical crossentropy* e escolhemos as configurações da RNP baseadas em testes da eficiência da rede, isto é, as de maior acurácia e menor perda.

Utilizamos o método de aprendizagem de máquina supervisionado para testar as camadas intermediárias no treino. Com esse método sabemos se a rede é capaz de determinar o valor do rótulo para cada variável a ser classificada.

## 7.4 Resultados

Considerando dois conjuntos diferentes de massas (**A** e **B**), como apresentado na tabela 7.2, testamos a rede com os dados gerados a partir do cálculo do número de eventos proporcional a seção de choque. Na Figura 7.3, vemos a separação dos eventos de sinal e de fundo no teste da rede.

	<b>A</b>	<b>B</b>
$M_{Z'}$ (GeV)	2500	3000
$M_{T'}$ (GeV)	1250	1500
$M_{\phi}$ (GeV)	300	300

Tabela 7.2: Conjunto de massas utilizados para separar os eventos.

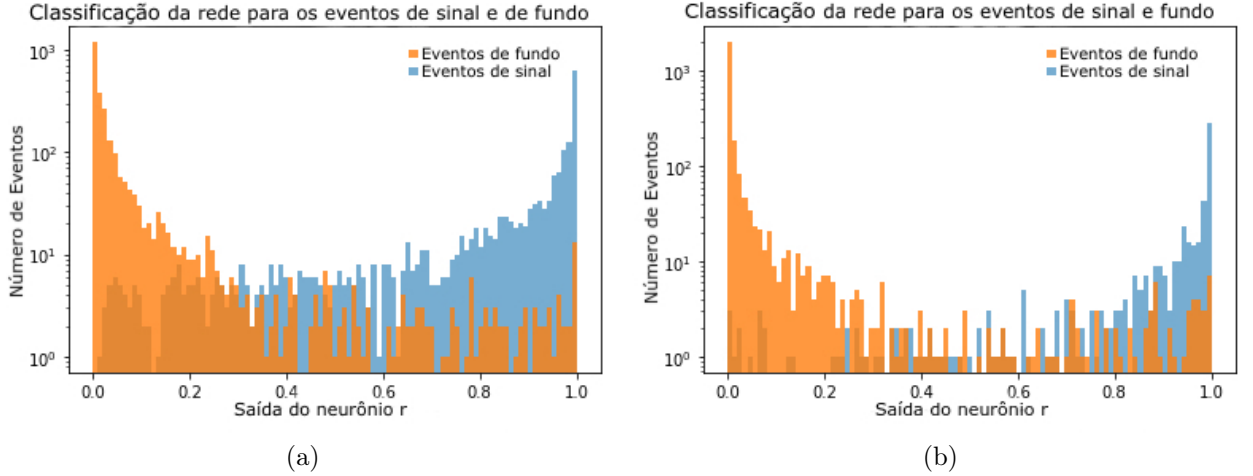


Figura 7.3: Número de eventos na escala logarítmica em função da saída no neurônio  $r$ . Nas figuras é apresentado as classificações da rede para os eventos vindos do Modelo Padrão ( $pp \rightarrow t\bar{t}Z$ ) e do modelo ZP-TP-DM ( $pp \rightarrow T'\bar{T}' \rightarrow t\bar{t}\phi\phi$ ) para (a)  $M_{Z'} = 2500\text{GeV}$ ,  $M_{T'} = 1250\text{GeV}$  e  $M_\phi = 300\text{GeV}$  e (b)  $M_{Z'} = 3000\text{GeV}$ ,  $M_{T'} = 1500\text{GeV}$  e  $M_\phi = 300\text{GeV}$ . Os eventos de fundo correspondem à área laranja e os eventos de sinal à azul.

A Figura 7.3a mostra a separação dos eventos relativos aos pontos de massa do conjunto **A**, enquanto a Figura 7.3b refere-se ao conjunto **B**. Ainda que existam eventos de sinal e de fundo sobrepostos, a utilização da RNP otimiza o tempo de trabalho comparado aos métodos convencionais baseados em cortes nas grandezas físicas, como energia e momento. Na Figura 7.4, temos a saída da rede referente à perda (Figura 7.4a) e à acurácia (Figura 7.4b) para o conjunto **A** com as curvas dos dados de treinamento da rede e de teste. Também é apresentado a perda (Figura 7.4c) e acurácia (Figura 7.4d) para o conjunto **B**. Esses gráficos indicam que o nosso classificador, que produziu as Figuras 7.3a e 7.3b para os pontos de massa citados, tem uma baixa perda (Figuras 7.4a e 7.4c, respectivamente) e uma boa acurácia (Figuras 7.4b e 7.4d, respectivamente). Notamos, portanto, que o classificador criado distingue eventos de sinal e fundo para diferentes conjuntos de massa.

Um classificador considerado ruim é aquele que apresenta uma curva ROC abaixo da linha tracejada como a apresentada na Figura 7.5. Se a  $AUC = 0$ , por exemplo, o modelo não é um classificador, pois não está classificando corretamente, retratando uma classe positiva como negativa e vice-versa. Quando a  $AUC$  é um valor muito próximo de 0.5, o modelo não tem a capacidade de distinguir entre classes positivas e negativas. Um excelente modelo é aquele com a mais alta  $AUC$ , ou seja,  $AUC = 1$  onde há uma distinção perfeita entre classes positivas e negativas. O nosso classificador assinalou os seguintes

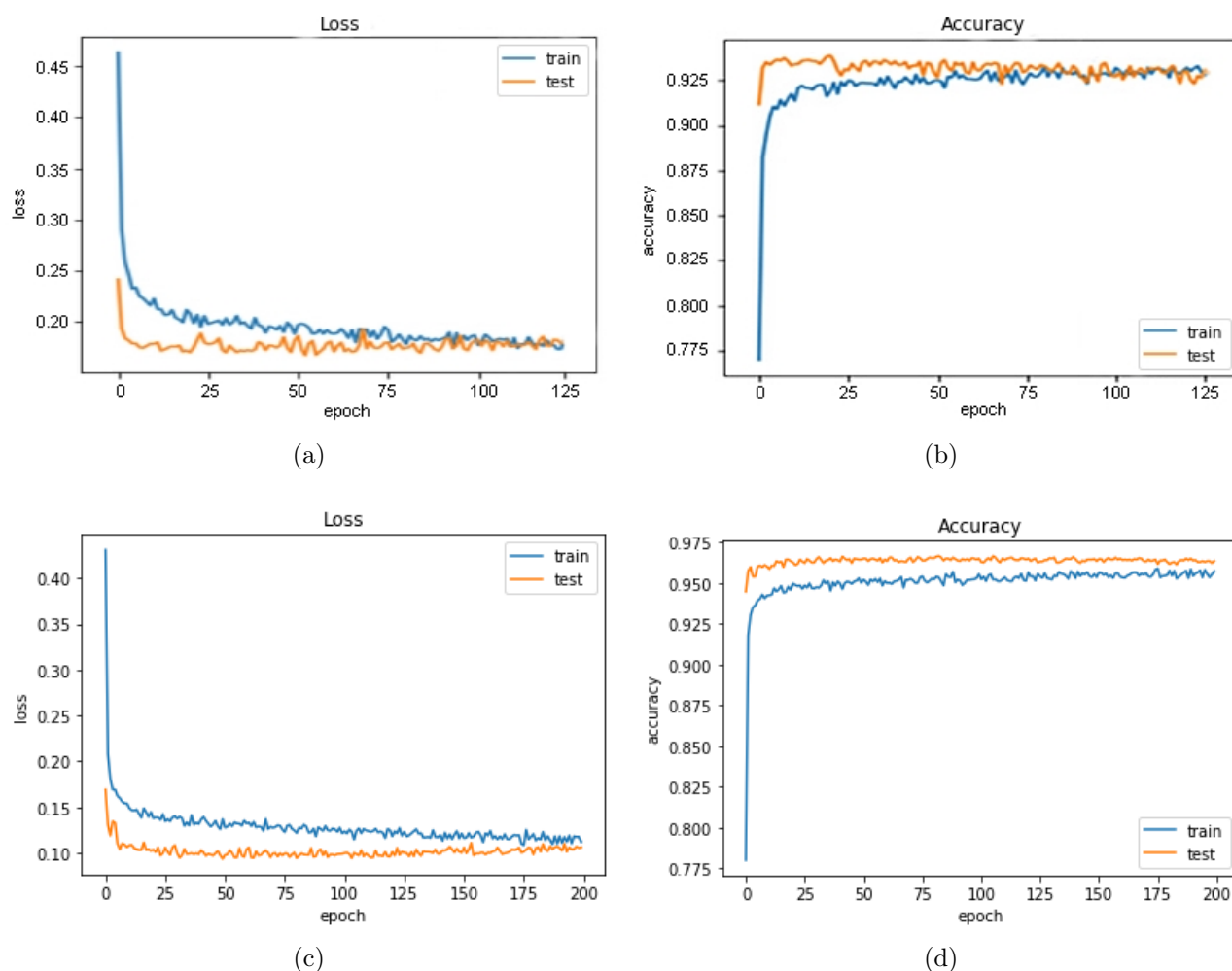


Figura 7.4: Em (a) e (b) temos a saída da rede referente à perda e à acurácia, respectivamente, para o conjunto **A**. Também é apresentado a perda (c) e acurácia (d) para o conjunto **B** do nosso modelo relacionado à Figura 7.3 com as curvas dos dados de treinamento (linha azul) da rede e de teste (linha laranja).

valores apresentados na Tabela 7.3:

AUC - <b>A</b>	AUC - <b>B</b>
0.980	0.989

Tabela 7.3: Valores para AUC dos conjunto de massas **A** e **B**, respectivamente.

Entendemos, dessa forma, que devido aos resultados apontados na Tabela 7.3, dispomos de um modelo eficiente, pois a AUC é um valor bem próximo de 1 para os conjuntos **A** e **B**. Isso quer dizer que há uma boa classificação na separação dos eventos de sinal (classe positiva) e de fundo (classe negativa).

A escolha de um valor de corte apropriado é de extrema importância para a eficácia na separação dos nossos eventos. Esse corte é calculado baseado na curva ROC. O critério

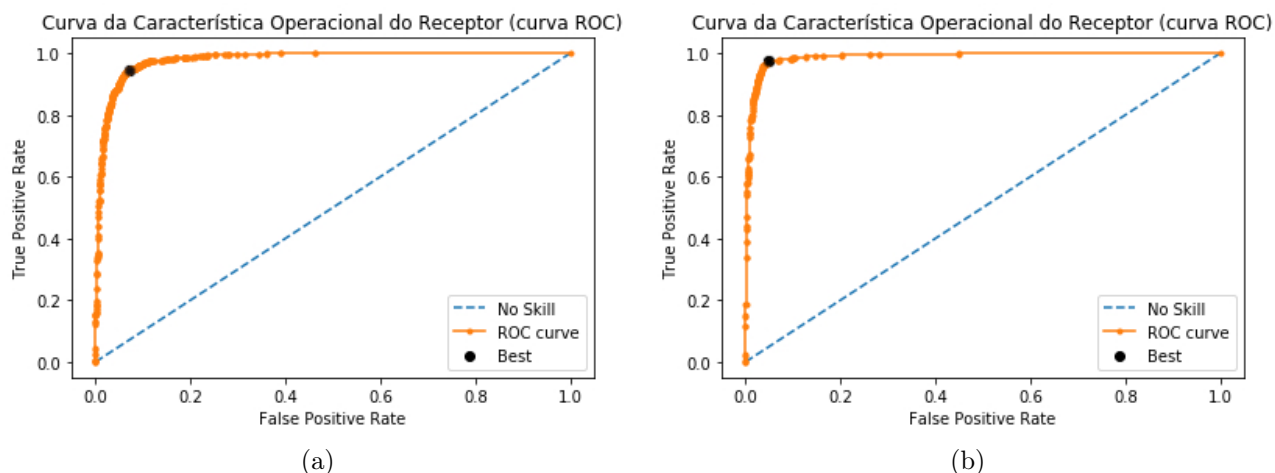


Figura 7.5: Curva ROC para o classificador criado (linha laranja) – referente à Figura 7.3 – considerando os eventos do sinal para (a)  $M_{Z'} = 2500\text{GeV}$ ,  $M_{T'} = 1250\text{GeV}$  e  $M_\phi = 300\text{GeV}$  e (b)  $M_{Z'} = 3000\text{GeV}$ ,  $M_{T'} = 1500\text{GeV}$  e  $M_\phi = 300\text{GeV}$ . O ponto denominado por "Best" é o melhor corte, isto é, indica o ponto que tem a maior taxa de verdadeiros positivos junto com a menor taxa de falsos positivos. Abaixo da linha tracejada tem  $\text{AUC}=0.5$ .

Eventos	A	B
Número total de eventos à direita do corte:	1859	687
Número total de eventos à esquerda do corte:	2575	2558
Total dos eventos de sinal:	1764	575
Total dos eventos de fundo:	2670	2670
Total dos eventos sinal+fundo:	4434	3245

Tabela 7.4: Número de eventos de acordo com os parâmetros vistos no Capítulo 6, seção 6.2, para um total de 3245 eventos de teste (sinal+fundo).

para escolher o melhor valor de corte foi selecionar o ponto que tem a maior taxa de positivos verdadeiros junto com a menor taxa de positivos falsos. Isto é, maior quantidade da taxa de eventos de sinal e menor taxa de eventos que contém o sinal que é confundido como fundo. A Figura 7.5 mostra a curva ROC da predição do nosso modelo apresentado na Figura 7.3. Cada ponto na curva ROC (Figura 7.5) corresponde a um valor de corte e está associado aos eventos de fundo e de sinal. Na Figura 7.5, o ponto preto denominado por "Best" indica o melhor corte. Na Figura 7.6, mostramos novamente as predições de separação dos eventos para os pontos de massa dos conjuntos **A** e **B**, mas com a indicação do melhor corte calculado pela curva ROC em 7.5 que nos forneceu  $\text{Best} \approx 0.25$  e  $\text{Best} \approx 0.22$  para **A** e **B**, respectivamente.

Na Tabela 7.4 mostramos os valores encontrados para a quantidade total de eventos de acordo com os parâmetros vistos no Capítulo 6, seção 6.2, para um total de 4434 (relativo à Figura 7.6a) e 3245 (relativo à Figura 7.6b) eventos de teste (sinal+fundo).

Vimos que os falsos positivos (FP) representam a quantidade de eventos que contém o sinal, porém confundidos como fundo, enquanto os verdadeiros positivos (VP) são apenas os eventos com o sinal. Os verdadeiros negativos (VN) e falsos negativos (FN), nessa ordem, são os eventos apenas de fundo e os eventos de fundo confundidos como sinal. Na tabela 7.5, apresentamos o número de eventos calculados pela rede para FP, VP, VN e FN dos conjuntos de massas **A** e **B**. Onde FP é o número de eventos de fundo à direita do corte da Figura 7.6, assim como, VP é o número de eventos à direita do corte. VN e FN são, respectivamente, o número de eventos de fundo e sinal à esquerda do corte.

$$\text{Conjunto A: } \frac{\mathbf{FP} = 190}{\mathbf{FN} = 95} \mid \frac{\mathbf{VP} = 1669}{\mathbf{VN} = 2480}$$

$$\text{Conjunto B: } \frac{\mathbf{FP} = 127}{\mathbf{FN} = 15} \mid \frac{\mathbf{VP} = 560}{\mathbf{VN} = 2543}$$

Tabela 7.5: Número de eventos calculados pela rede para FP, VP, VN e FN dos conjuntos de massas **A** e **B**.

Entendemos, dessa forma, que devido aos resultados apontados na Tabela 7.3, dispomos de um modelo adequado, pois a AUC é um valor bem próximo de 1 em ambos os conjuntos **A** e **B**. Isso quer dizer que há uma boa classificação na separação dos eventos de sinal (classe positiva) e de fundo (classe negativa).

Os valores que encontramos para esses parâmetros, no nosso teste da rede, se encontram na Tabela 7.4. De acordo com a tabela 7.4, à direita do corte da Figura 7.6a tem um total de 1859 eventos, enquanto à esquerda tem 2575 eventos. À direita do corte da Figura 7.6b tem um total de 687 eventos, enquanto à esquerda tem 2558 eventos.

Com o objetivo de analisarmos como a área abaixo da curva (AUC) varia com a massa da ME ( $M_\phi$ ) plotamos o gráfico  $AUC \times M_\phi$  para dois pontos de massa  $M_{Z'} = 2500 GeV$  e  $M_{T'} = 1250 GeV$ ;  $M_{Z'} = 3000 GeV$  e  $M_{T'} = 1500 GeV$ . Analisando a Figura 7.7, observamos que quanto maior o valor da massa da matéria escura, mais baixa é o valor da AUC. Quando  $M_{Z'} = 3000 GeV$ , temos melhores valores para AUC de  $M_\phi = 100 GeV$  até  $M_\phi = 900 GeV$ . De modo geral, para pequenos valores de massa da ME,  $M_\phi < 300 GeV$ , a AUC é a mais alta, e tem um valor acima de 0.98 aproximadamente. Apesar dessa faixa de massa ter distinção entre eventos de sinal e de fundo muito boa, como observamos pelo gráfico, precisamos da significância para determinar se, de fato, há alguma flutuação estatística sobre os dados simulados em análise.

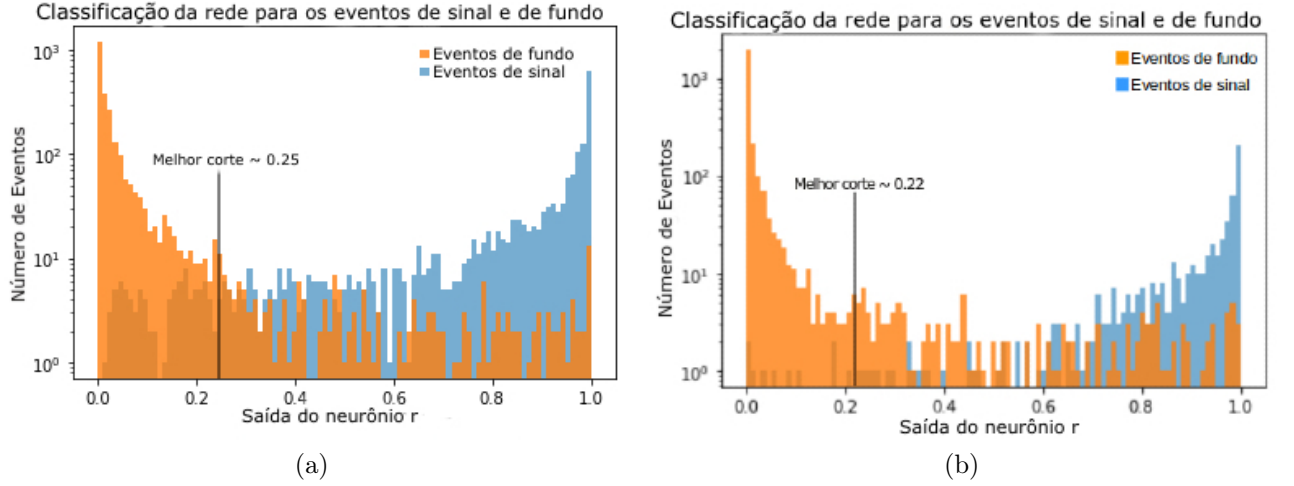


Figura 7.6: Número de eventos na escala logarítmica em função da saída no neurônio  $r$ . Nas figuras são apresentadas as mesmas classificações da figura 7.3 dos eventos vindos do Modelo Padrão ( $pp \rightarrow t\bar{t}Z$ ) e do modelo ZP-TP-DM ( $pp \rightarrow T'\bar{T}' \rightarrow t\bar{t}\phi\phi$ ), para as massas (a)  $M_{Z'} = 2500\text{GeV}$ ,  $M_{T'} = 1250\text{GeV}$  e  $M_\phi = 300\text{GeV}$  e (b)  $M_{Z'} = 3000\text{GeV}$ ,  $M_{T'} = 1500\text{GeV}$  e  $M_\phi = 300\text{GeV}$ . O melhor corte do número de eventos está indicado pela linha preta equivalente a 0.25 em (a) e 0.22 em (b).

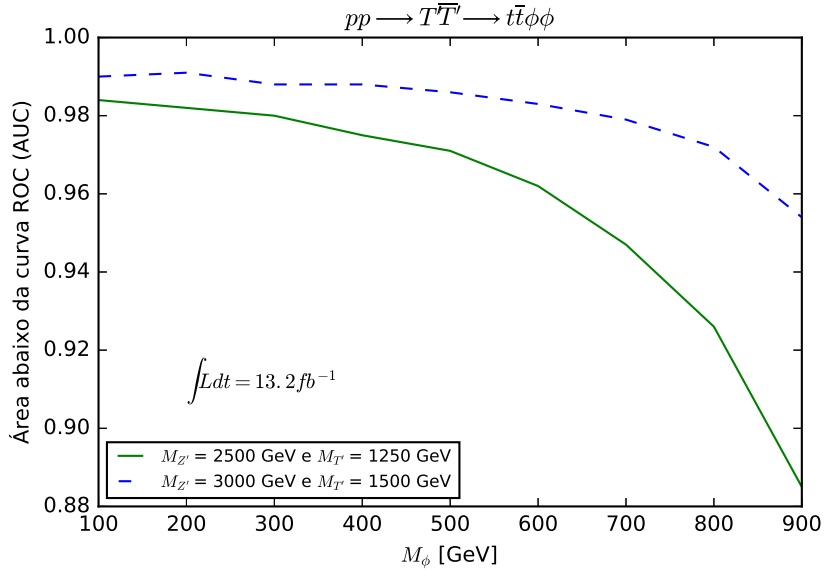


Figura 7.7: Variação dos valores AUC de acordo com várias massas da ME para  $M_{Z'} = 2500\text{GeV}$  e  $M_{T'} = 1250\text{GeV}$ ;  $M_{Z'} = 3000\text{GeV}$  e  $M_{T'} = 1500\text{GeV}$ . O valores do número de épocas foram escolhidos de modo a evitar o *overfitting*.

## 7.5 Significância do sinal

Na Física de Partículas, utilizamos este conceito de significância [78] [79] para nos ajudar a identificar um quadro estatístico consistente que possa ser quantificado. Temos como propósito utilizar um teste estatístico que possua como hipótese negar a presença do processo de sinal e testar essa hipótese contra o caso em que o sinal exista. Esse método consiste em contabilizar incertezas estatísticas na determinação de valores médios



dos eventos de sinal e de fundo. Na ocorrência de um sinal, podemos quantificar esse conceito pela sua significância estatística.

É do nosso interesse saber o número de eventos numa dada região no espaço de variáveis, região em que se encontra o sinal. Dado  $\mathcal{Y} = \{x : y(x) = s\}$ , onde  $\mathcal{Y}$  representa a região de interesse no espaço das variáveis de entrada;  $x$  é a quantidade das variáveis de entrada, com  $x \in \mathbb{R}^d$ ;  $y$  é um classificador em função das variáveis de entrada –  $y(x)$  e  $s$  é o número de eventos de sinal.

O número de eventos totais  $n$  encontrados na região  $\mathcal{Y}$  assume uma distribuição de Poisson na média  $\mu_s + \mu_f$ . No caso de variáveis com essa distribuição, a probabilidade de que determinado evento ocorra é calculado pela seguinte distribuição de Poisson:

$$P(n|\mu_s, \mu_f) = \frac{(\mu_s + \mu_f)^n}{n!} e^{-(\mu_s + \mu_f)}, \quad (7.5)$$

onde  $\mu_s$  é o número de eventos esperados de sinal,  $\mu_f$  é o número de eventos esperados de fundo e  $n$  é o número de eventos observados.

Para o teste de hipótese, vamos supor uma medida para  $\mu_s$  cujo valor difere as duas hipóteses que chamaremos de  $\mathcal{H}_0$  e  $\mathcal{H}_1$ . Seja  $\mu_s = 0$  ( $\mathcal{H}_0$ ) e  $\mu_s = \tilde{\mu}_s$  ( $\mathcal{H}_1$ ), a equação 7.5 para cada caso, respectivamente será:

$$P(n|0, \mu_f) = \frac{\mu_f^n}{n!} e^{-\mu_f}. \quad (7.6)$$

$$P(n|\tilde{\mu}_s, \mu_f) = \frac{(\tilde{\mu}_s + \mu_f)^n}{n!} e^{-(\tilde{\mu}_s + \mu_f)}, \quad (7.7)$$

onde  $\tilde{\mu}_s = n - \mu_s$  é o valor máximo estimado para  $\mu_s$  observado de  $n$  eventos na região  $\mathcal{Y}$ .

As equações 7.6 e 7.7 representam as funções de distribuição de probabilidade para  $\mu_s$  associadas às duas hipóteses. Fazendo a razão entre as equações 7.6 e 7.7 encontramos:

$$\lambda = \frac{P(n|0, \mu_f)}{P(n|\tilde{\mu}_s, \mu_f)} = \left(\frac{\mu_f}{n}\right)^n e^{(n-\mu_f)}. \quad (7.8)$$

Em concordância com o teorema de Wilks [80], o teste estatístico fica:

$$q_0 = -2 \ln \lambda, \text{ se } n > \mu_f \quad ; \quad q_0 = 0, \text{ se } n < \mu_f \quad (7.9)$$

No caso em que  $n > \mu_f$ ,  $q_0$  assume o valor:

$$q_0 = 2 \left[ n \ln \left( \frac{n}{\mu_f} \right) + \mu_f - n \right] \quad (7.10)$$

Como assumimos que a densidade de probabilidade para  $n$  é uma distribuição de Poisson, podemos realizar uma estimativa direta do chamado valor-p, uma probabilidade, de uma dada observação em  $n$  eventos na região  $\mathcal{Y}$ . De acordo com [81], a significância estatística pode ser correlacionada com o valor-p como mostra a Tabela 7.6.

<b>S</b> [ $\sigma$ ]	<b>valor-p</b>
1	0.3173
2	0.0455
3	0.0027
4	$6.3 \times 10^{-5}$
5	$5.7 \times 10^{-7}$
6	$2.0 \times 10^{-9}$

Tabela 7.6: Valores-p para sua correspondente significância de 1 a 6  $\sigma$  (unidade de desvio padrão) [81].

Para testar a descoberta do sinal, o valor-p da hipótese, em que  $\mu_s = 0$ , é calculado através da equação:

$$p = \sum_{N=n}^{\infty} \frac{(\mu_f)^n}{n!} e^{-\mu_f} \quad (7.11)$$

Usualmente converte-se o valor-p em uma significância equivalente  $S$ , definida como:

$$S = \Phi^{-1}(1 - p), \quad (7.12)$$

onde  $\Phi$  é uma distribuição cumulativa gaussiana padrão.

A quantidade  $S$ , que mede a significância estatística é dada em unidades de desvio padrão ( $\sigma$ ). Convencionalmente, temos que se  $S = 2\sigma$ , há uma fraca evidência para o sinal; para  $S = 3\sigma$ , uma forte evidência; para  $S = 5\sigma$ , podemos considerar a descoberta de uma nova física. Caso seja encontrado um resultado com um erro de  $5\sigma$ , há uma chance em 3.5 milhões de ser uma flutuação estatística sobre as experiências efetuadas. Logo,  $5\sigma$  é usado como regra geral para definir a sensibilidade necessária para uma descoberta. Por exemplo, encontrar  $S = 5\sigma$  corresponde achar um valor-p entre  $5.7 \times 10^{-7}$  e  $2.9 \times 10^{-7}$  [82]. A significância que obteríamos para um dado número de eventos  $n$  observados na região de interesse  $\mathcal{Y}$ , conhecendo a expectativa de fundo  $\mu_f$  é definida como  $S = \sqrt{q_0}$ , então a

partir da equação 7.10:

$$S = \sqrt{q_0} = \sqrt{2 \left[ n \ln \left( \frac{n}{\mu_f} \right) + \mu_f - n \right]}, \text{ para } n > \mu_f \quad (7.13)$$

Porém, não conhecemos os valores de  $n$  e  $\mu_f$ , portanto, consideramos estimativas para os eventos de fundo e de sinal. Utilizando uma Significância Mediana Aproximada (SMA) substituímos  $n$  por  $n_s + n_f$  e  $\mu_f = n_f$ , onde  $n_s$  é o número dos eventos de sinal estimado e  $n_f$  o número dos eventos de fundo estimado. A equação 7.13 fica:

$$SMA1 = \sqrt{2 \left[ (n_s + n_f) \ln \left( 1 + \frac{n_s}{n_f} \right) - n_s \right]} \quad (7.14)$$

Os valores para  $n_s$  e  $n_f$ , correspondem a taxa de verdadeiros positivos (VP) e falsos positivos (FP), respectivamente.

Para  $n_f \gg n_s$ , podemos considerar a aproximação:

$$SMA2 = \frac{n_s}{\sqrt{n_f}} \quad (7.15)$$

Na Figura 7.8, plotamos um gráfico para os valores encontrados da significância SMA utilizando a equação 7.14 e 7.15 para vários valores de  $M_\phi$  com  $M_{Z'} = 2500 GeV$  e  $3000 GeV$  e  $M_{T'} = 1250 GeV$  e  $1500 GeV$ . Como exemplo, para o cálculo da SMA em um ponto,  $M_\phi = 300 GeV$ , de acordo com a tabela 7.5, para o conjunto **A** a rede encontrou  $n_s = 1669$  eventos e  $n_f = 190$  eventos. Com as correções necessárias para  $n_f$ , temos que  $n'_f = 190 \times 3.03 \times 1.3 \approx 748$  eventos. Substituindo  $n_s$  e  $n_f$  corrigido em 7.14, obtemos  $SMA1 \approx 48.3$ . Enquanto para o conjunto **B**,  $n_s = 560$  e  $n'_f = 500$  eventos, o que nos leva a uma significância de, pela equação 7.14,  $SMA1 \approx 21.7$ . A equação 7.15 apresentou valores maiores para a significância estatística: conjunto **A**,  $SMA2 \approx 61$  e conjunto **B**,  $SMA2 \approx 21$ .

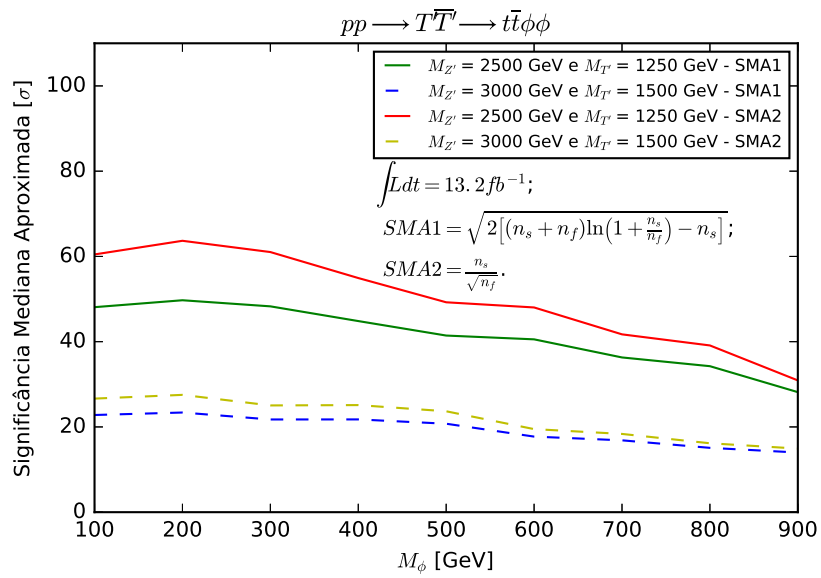


Figura 7.8: Comparação da variação dos valores da significância do sinal de acordo com a massa  $M_\phi$  para SMA1 (em verde e azul) e SMA2 (em vermelho e amarelo).

# Capítulo 8

## Conclusões

Neste trabalho, fizemos um estudo da assinatura  $t\bar{t} + \cancel{E}_T$  produzida no Grande colisor de Hádrons (LHC). Em Física de Partículas, sabemos que a Matéria Escura (ME) não interage com os detectores, mas supõe-se a sua contribuição para a  $\cancel{E}_T$ . Diante desse cenário, escolhemos trabalhar com um modelo simplificado efetivo ZP-TP-DM, uma extensão do Modelo Padrão, que considera uma partícula  $\phi$ , candidata a ME, como efetivamente neutra e estável, concordando com o Modelo Cosmológico Padrão.

Dentro deste contexto, utilizamos o software CalcHEP para calcular as seções de choque e gerar eventos de sinal e de fundo. A análise consistiu em separar os eventos de fundo,  $pp \rightarrow t\bar{t}Z$ , dos eventos de sinal,  $pp \rightarrow T\bar{T}' \rightarrow t\bar{t}\phi\phi$ , por meio da Aprendizagem de Máquina (AM). Uma característica desse estudo, foi a escolha do canal  $pp \rightarrow t\bar{t}Z$ , que possui a mesma assinatura do sinal que é  $t\bar{t} + \cancel{E}_T$ . Isso porque, o bóson  $Z$  pode decair em um par de (anti)partícula, produzindo neutrinos e seus respectivos antineutrinos que não são detectados pelos aparatos experimentais, contribuindo para a  $\cancel{E}_T$ .

Os eventos de sinal e de fundo foram gerados utilizando valores estabelecidos para energia do centro de massa,  $\sqrt{s} = 13TeV$ , e com uma luminosidade integrada  $\mathcal{L} = 13.2fb^{-1}$ . Especificamente para os eventos de sinal, consideramos os valores  $M_{Z'} = 2500GeV$  e  $3000GeV$ , com respectivos  $M_{T'}$  fixos:  $M_{T'} = 1250GeV$  e  $1500GeV$ , com diversas massas para  $M_\phi$ .

Criamos um classificador que pode ser usado para diferentes massas e, após os dados passarem pela Rede neural Profunda (RNP), obtivemos resultados significativos na separação dos eventos. Por exemplo, para os conjuntos de massas fixas **A** e **B** com  $M_\phi = 300GeV$ , apresentados na Tabela 7.2, os valores calculados para a Área Abaixo da

Curva (AUC) foram 0.980 para o conjunto **A** e 0.989 para o conjunto **B** (AUC das Figuras 7.5a e 7.5b, respectivamente). Tais resultados mostram que a rede é capaz de separar os eventos de fundo dos eventos de sinal. Sendo assim, a Aprendizagem de Máquina (AM) demonstra ser uma técnica eficaz e alternativa ao método tradicional de separação de eventos por cortes retangulares de variáveis. Entretanto, notamos uma menor eficiência do classificador conforme o aumento da massa  $M_\phi$ , como mostramos na Figura 7.7.

Percebemos que, para um valor mais alto de  $M_{T'}$ , a Significância Mediana Aproximada (SMA) diminui. Os resultados obtidos mostraram que para  $M_{T'} = 1500 GeV$ , o menor valor que SMA assume, é algo próximo a  $20\sigma$ , como mostrado na Figura 7.8. Encontramos altos valores para a significância porque não fizemos o decaimento de  $t\bar{t}$ . Portanto, a análise feita foi apenas em nível de pártons. Para um estudo mais realista, deveríamos considerar algum dos canais de produção vindos do decaimento do par  $t\bar{t}$ , canal hadrônico ( $t\bar{t} \rightarrow W^+bW^- \rightarrow q\bar{q}'bq''\bar{q}'''\bar{b}$ ), semi-leptônico ( $t\bar{t} \rightarrow W^+bW^- \rightarrow q\bar{q}'bl^-\bar{\nu}_\ell\bar{b} + \ell^+\nu_\ell bq''\bar{q}'''\bar{b}$ ) ou leptônico ( $t\bar{t} \rightarrow W^+bW^- \rightarrow \ell^+\nu_\ell bl^-\bar{\nu}_\ell\bar{b}$ ), como apresentado na seção 7.2.1. Por consequência disso, o gráfico da significância que obtemos não é realista, é idealizado, porque desconsidera o decaimento do  $t\bar{t}$ . Por isso, não pudemos comparar os nossos resultados com os resultados apresentados artigo do Belyaev [8], pois ele fez o decaimento de  $t\bar{t}$  pelo DELPHES.

Muita informação é perdida quando o decaimento do  $t\bar{t}$ , a hadronização é feita através do PYTHIA e a simulação do detector é feita com auxílio do DELPHES, o que leva a uma diminuição da significância estatística. Logo, os resultados que obtivemos para a significância iriam diminuir caso tivéssemos utilizado essas ferramentas de simulação.

Contudo, o resultado deste trabalho indica que a técnica de AM é um método bastante adequado e que pode ser aplicado à separação dos eventos de sinal e de fundo obtendo resultados semelhantes aos métodos tradicionais.

# Referências Bibliográficas

- [1] G. Bertone and D. Hooper, “History of dark matter,” *Reviews of Modern Physics*, vol. 90, no. 4, p. 045002, 2018.
- [2] R. J. Gaitskell, “Direct detection of dark matter,” *Annu. Rev. Nucl. Part. Sci.*, vol. 54, pp. 315–359, 2004.
- [3] J. Conrad, “Indirect detection of wimp dark matter: a compact review,” *arXiv preprint arXiv:1411.1925*, 2014.
- [4] P. J. Fox, R. Harnik, J. Kopp, and Y. Tsai, “Missing energy signatures of dark matter at the lhc,” *Physical Review D*, vol. 85, no. 5, p. 056011, 2012.
- [5] P. J. E. Peebles and P. J. Peebles, *Principles of physical cosmology*. Princeton university press, 1993.
- [6] G. Altarelli, “The standard model of particle physics,” *arXiv preprint hep-ph/0510281*, 2005.
- [7] C. collaboration *et al.*, “Missing transverse energy performance of the cms detector,” *Journal of Instrumentation*, vol. 6, no. 09, p. P09001, 2011.
- [8] A. S. Belyaev, T. Flacke, B. Jain, and P. B. Schaefers, “Lhc dark matter signals from vector resonances and top partners,” *Physical Review D*, vol. **98**, no. 3, p. 035019, 2018.
- [9] “Search for top squarks in final states with one isolated lepton, jets, and missing transverse momentum in  $\sqrt{s} = 13$  TeV pp collisions with the ATLAS detector,” tech. rep., CERN, Geneva, Aug 2016. ATLAS-CONF-2016-050.

- [10] A. Pukhov, “Calchep 2.3: Mssm, structure functions, event generation, batchs, and generation of matrix elements for other packages,” *arXiv preprint hep-ph/0412191*, 2004.
- [11] T. Mitchell, “Introduction to machine learning,” *Machine Learning*, vol. **7**, pp. 2–5, 1997.
- [12] J. McCarthy, “What is artificial intelligence,” *Computer Science Department, Stanford University*, p. 2, 2004.
- [13] S. B. Kotsiantis, I. Zaharakis, and P. Pintelas, “Supervised machine learning: A review of classification techniques,” *Emerging artificial intelligence applications in computer engineering*, vol. **160**, pp. 3–24, 2007.
- [14] J. J. Hopfield, “Artificial neural networks,” *IEEE Circuits and Devices Magazine*, vol. **4**, no. 5, pp. 3–10, 1988.
- [15] F. Close, M. Marten, and C. Sutton, *The Particle Odyssey: A Journey to the Heart of the Matter*. Oxford University Press, USA, 2002.
- [16] A. Romanino, “The Standard model of particle physics,” in *Physics of elementary particles and astrophysics. Proceedings, International Baikal Summer School, Bolshie Koty, Russia, July 23-30, 2009*, 2009.
- [17] Y. Nagashima and Y. Nagashima, *Elementary particle physics*, vol. **1**. Wiley Online Library, 2010.
- [18] W. N. Cottingham and D. A. Greenwood, *An introduction to the standard model of particle physics*. Cambridge university press, 2007.
- [19] S. Krippendorff, F. Quevedo, and O. Schlotterer, “Cambridge lectures on supersymmetry and extra dimensions,” *arXiv preprint arXiv:1011.1491*, 2010.
- [20] W. Buchmuller and C. Ludeling, “Field Theory and Standard Model,” in *2005 European School of High-Energy Physics*, 2006.
- [21] W. Hollik, “Quantum field theory and the standard model,” *arXiv preprint arXiv:1012.3883*, 2010.



- [22] H. Georgi, *Lie algebras in particle physics: from isospin to unified theories*. CRC Press, 2018.
- [23] G. Bertone, N. Bozorgnia, J. S. Kim, S. Liem, C. McCabe, S. Otten, and R. R. de Austri, “Identifying wimp dark matter from particle and astroparticle data,” *Journal of Cosmology and Astroparticle Physics*, vol. 2018, no. 03, p. 026, 2018.
- [24] P. De Jong, A. De Roeck, K. de Vries, D. Del Re, A. De, J. Ellis, S. Eno, E. Etzion, M. Fairbairn, B. Feldstein, *et al.*, “Simplified models for dark matter searches at the lhc,” *arXiv preprint arXiv:1506.03116*, 2015.
- [25] O. Buchmueller, M. J. Dolan, S. A. Malik, and C. McCabe, “Characterising dark matter searches at colliders and direct detection experiments: vector mediators,” *Journal of High Energy Physics*, vol. 2015, no. 1, p. 37, 2015.
- [26] E. Zackrisson, *Quasars and Low Surface Brightness Galaxies as Probes of Dark Matter*. PhD thesis, Institutionen för astronomi och rymdfysik, 2005.
- [27] N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. Banday, R. Barreiro, N. Bartolo, S. Basak, *et al.*, “Planck 2018 results. vi. cosmological parameters,” *arXiv preprint arXiv:1807.06209*, 2018.
- [28] D. Majumdar, *Dark matter: an introduction*. CRC Press, 2014.
- [29] V. C. Rubin, W. K. Ford Jr, and N. Thonnard, “Extended rotation curves of high-luminosity spiral galaxies. iv-systematic dynamical properties, sa through sc,” *The Astrophysical Journal*, vol. **225**, pp. L107–L111, 1978.
- [30] K. S. Virbhadra and G. F. Ellis, “Schwarzschild black hole lensing,” *arXiv preprint arXiv:astro-ph/9904193*, 2000.
- [31] F. Zwicky, “Republication of: The redshift of extragalactic nebulae,” *General Relativity and Gravitation*, vol. **41**, no. 1, pp. 207–224, 2009.
- [32] P. Schneider, “Detection of (dark) matter concentrations via weak gravitational lensing,” *Monthly Notices of the Royal Astronomical Society*, vol. **283**, no. 3, pp. 837–853, 1996.

- [33] V. C. Rubin and W. K. Ford Jr, “Rotation of the andromeda nebula from a spectroscopic survey of emission regions,” *The Astrophysical Journal*, vol. **159**, p. 379, 1970.
- [34] A. Belyaev, L. Panizzi, A. Pukhov, and M. Thomas, “Dark matter characterization at the lhc in the effective field theory approach,” *Journal of High Energy Physics*, vol. **2017**, no. 4, p. 110, 2017.
- [35] J. Goodman, M. Ibe, A. Rajaraman, W. Shepherd, T. M. Tait, and H.-B. Yu, “Constraints on light majorana dark matter from colliders,” *Physics Letters B*, vol. **695**, no. 1-4, pp. 185–188, 2011.
- [36] J. Abdallah, H. Araujo, A. Arbey, A. Ashkenazi, A. Belyaev, J. Berger, C. Boehm, A. Boveia, A. Brennan, J. Brooke, *et al.*, “Simplified models for dark matter searches at the lhc,” *Physics of the Dark Universe*, vol. **9**, pp. 8–23, 2015.
- [37] S. Baek, P. Ko, and P. Wu, “Top-philic scalar dark matter with a vector-like fermionic top partner,” *Journal of High Energy Physics*, vol. **2016**, no. 10, p. 117, 2016.
- [38] T. M. Undagoitia and L. Rauch, “Dark matter direct-detection experiments,” *Journal of Physics G: Nuclear and Particle Physics*, vol. **43**, no. 1, p. 013001, 2015.
- [39] E. Amato and P. Blasi, “Cosmic ray transport in the galaxy: A review,” *Advances in Space Research*, vol. **62**, no. 10, pp. 2731–2749, 2018.
- [40] M. Duerr, P. F. Pérez, and J. Smirnov, “Scalar dark matter: direct vs. indirect detection,” *Journal of High Energy Physics*, vol. **2016**, no. 6, p. 152, 2016.
- [41] F. Kahlhoefer, “Review of lhc dark matter searches,” *International Journal of Modern Physics A*, vol. **32**, no. 13, p. 1730006, 2017.
- [42] <https://cds.cern.ch/record/2684277>.
- [43] B. The Editors of Encyclopaedia, “Tevatron: Particle accelerator,” 2015.
- [44] T. A. Collaboration, *ATLAS: A 25-Year Insider Story of the LHC Experiment*. World Scientific, 2019.

- [45] G. Aad, J. Butterworth, J. Thion, U. Bratzler, P. Ratoff, R. Nickerson, J. Seixas, I. Grabowska-Bold, F. Meisel, S. Lokwitzer, *et al.*, “The atlas experiment at the cern large hadron collider,” *Jinst*, vol. **3**, p. S08003, 2008.
- [46] A. B. Poy, H. Boterenbrood, H. Burckhart, J. Cook, V. Filimonov, S. Franz, O. Gutzwiller, B. Hallgren, V. Khomutnikov, S. Schlenker, *et al.*, “The detector control system of the atlas experiment,” *Journal of Instrumentation*, vol. **3**, no. 05, p. P05006, 2008.
- [47] <https://cds.cern.ch/images/ATLAS-PHOTO-2019-012-1>.
- [48] A. Pukhov, “Calchep-calculator for high energy physics—a package for evaluation of feynman diagrams, integration over multi-particle phase space, and event generation.,” *not published*, 2005.
- [49] A. Belyaev, N. D. Christensen, and A. Pukhov, “Calchep 3.4 for collider physics within and beyond the standard model,” *Computer Physics Communications*, vol. **184**, no. 7, pp. 1729–1769, 2013.
- [50] J. J. Sakurai, J. Napolitano, *et al.*, *Modern quantum mechanics*, vol. **2nd ed.** Pearson, 2012.
- [51] <http://hepmdb.soton.ac.uk/>.
- [52] <https://theory.sinp.msu.ru/~pukhov/calchep.html>.
- [53] A. Semenov, “Lanhep—a package for the automatic generation of feynman rules in field theory. version 3.0,” *Computer Physics Communications*, vol. **180**, no. 3, pp. 431–454, 2009.
- [54] A. Semenov, “a package for automatic generation of feynman rules from the lagrangian. updated version 3.2,” *arXiv preprint arXiv:1412.5016*, 2014.
- [55] N. D. Christensen, C. Duhr, and B. Fuks, “The feynrules manual, version 1.4,” 2009.
- [56] J. Alwall, A. Ballestrero, P. Bartalini, S. Belov, E. Boos, A. Buckley, J. Butterworth, L. Dudko, S. Frixione, L. Garren, S. Gieseke, A. Gusev, I. Hinchliffe, J. Huston, B. Kersevan, *et al.*, “A standard format for les houches event files,” *arXiv preprint hep-ph/0609017*, 2007.

- [57] J. De Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaitre, A. Mertens, M. Selvaggi, D. . Collaboration, *et al.*, “Delphes 3: a modular framework for fast simulation of a generic collider experiment,” *Journal of High Energy Physics*, vol. **2014**, no. 2, p. 57, 2014.
- [58] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, “An introduction to pythia 8.2,” *Computer physics communications*, vol. **191**, pp. 159–177, 2015.
- [59] S. Kraml, U. Laa, L. Panizzi, and H. Prager, “Scalar versus fermionic top partner interpretations of  $t\bar{t}+e\bar{t}miss$  searches at the lhc,” *Journal of High Energy Physics*, vol. **2016**, no. 11, p. 107, 2016.
- [60] <https://hepmdb.soton.ac.uk/hepmdb:0717.0253>.
- [61] M. Thomson, *Modern particle physics*. Cambridge University Press, 2013.
- [62] F.-H. Liu, T. Tian, J.-X. Sun, and B.-C. Li, “What can we learn from (pseudo) rapidity distribution in high energy collisions?,” *Advances in High Energy Physics*, vol. **2014**, 2014.
- [63] S. Russell and P. Norvig, “Artificial intelligence: a modern approach,” 2002.
- [64] F. Pistono and R. V. Yampolskiy, “Unethical research: how to create a malevolent artificial intelligence,” *arXiv preprint arXiv:1605.02817v2*, 2016.
- [65] B. Goertzel, “Artificial general intelligence: concept, state of the art, and future prospects,” *Journal of Artificial General Intelligence*, vol. **5**, no. 1, pp. 1–48, 2014.
- [66] P. Wang and B. Goertzel, “Introduction: Aspects of artificial general intelligence,” in *Proceedings of the 2007 conference on Advances in Artificial General Intelligence: Concepts, Architectures and Algorithms: Proceedings of the AGI Workshop 2006*, pp. 1–16, 2007.
- [67] A. Géron, *Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems*. "O'Reilly Media, Inc.", 2017.
- [68] M. Kubat, *An introduction to machine learning*. Springer, 2017.

- [69] D. Guest, K. Cranmer, and D. Whiteson, “Deep learning and its application to lhc physics,” *Annual Review of Nuclear and Particle Science*, vol. **68**, pp. 161–181, 2018.
- [70] S. Jung, D. Lee, and K.-P. Xie, “Beyond  $m_{t\bar{t}}$ : learning to search for a broad  $t\bar{t}$  resonance at the lhc,” *The European Physical Journal C*, vol. 80, no. 2, pp. 1–13, 2020.
- [71] A. Beck and M. Kurz, “A perspective on machine learning methods in turbulence modelling,” *arXiv preprint arXiv:2010.12226*, 10 2020.
- [72] A. K. Jain, J. Mao, and K. M. Mohiuddin, “Artificial neural networks: A tutorial,” *Computer*, vol. **29**, no. 3, pp. 31–44, 1996.
- [73] M. Usama, J. Qadir, A. Raza, H. Arif, K.-L. A. Yau, Y. Elkhatib, A. Hussain, and A. Al-Fuqaha, “Unsupervised machine learning for networking: Techniques, applications and research challenges,” *arXiv preprint arXiv:1709.06599*, 2017.
- [74] S. Forte and S. Carrazza, “Parton distribution functions,” *arXiv preprint arXiv:2008.12305v1*, 2020.
- [75] R. D. Ball, V. Bertone, S. Carrazza, L. Del Debbio, S. Forte, P. Groth-Merrild, A. Guffanti, N. P. Hartland, Z. Kassabov, J. I. Latorre, *et al.*, “Parton distributions from high-precision collider data,” *The European Physical Journal C*, vol. **77**, no. 10, pp. 1–75, 2017.
- [76] R. Placakyte, “Parton distribution functions,” *arXiv preprint arXiv:1111.5452*, 2011.
- [77] K. Hagiwara, K. Hikasa, K. Nakamura, M. Tanabashi, M. Aguilar-Benitez, C. Amstler, R. Barnett, P. Burchat, C. Carone, C. Caso, *et al.*, “Review of particle physics: Particle data group,” *Physical Review D-Particles, Fields, Gravitation and Cosmology*, vol. **66**, no. 1 I, pp. 100011–10001958, 2002.
- [78] P. K. Sinervo, “Signal significance in particle physics,” *arXiv preprint hep-ex/0208005*, 2002.
- [79] C. Adam-Bourdarios, G. Cowan, C. Germain, I. Guyon, B. Kegl, and D. Rousseau, “Learning to discover: the higgs boson machine learning challenge,” *URL <http://higgsml.lal.in2p3.fr/documentation>*, p. 9, 2014.

- [80] S. S. Wilks, “The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses,” *The Annals of Mathematical Statistics*, vol. **9**, no. 1, pp. 60 – 62, 1938.
- [81] Y.-S. Zhu, “On statistical significance of signal,” *arXiv preprint arXiv:0812.2708*, 2008.
- [82] P. K. Sinervo, “Signal significance in particle physics,” *arXiv preprint hep-ex/0208005*, 2002.