



IDENTIFICAÇÃO DE ELÉTRONS BASEADA EM UM CALORÍMETRO DE
ALTAS ENERGIAS FINAMENTE SEGMENTADO

Werner Spolidoro Freund

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

Orientador: José Manoel de Seixas

Rio de Janeiro
Março de 2018

IDENTIFICAÇÃO DE ELÉTRONS BASEADA EM UM CALORÍMETRO DE
ALTAS ENERGIAS FINAMENTE SEGMENTADO

Werner Spolidoro Freund

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

Examinada por:

Prof. Alexandre Gonçalves Evsukoff, Ph.D.

Prof. Aluizio Fausto Ribeiro Araujo, Ph.D.

Prof. Ignácio De Bediaga Hickman, Ph.D.

Prof. José Manoel de Seixas, D.Sc.

Prof. Luiz Pereira Calôba, Dr.Ing.

Prof. Marley Maria Bernardes Rebuzzi Vellasco, Ph.D.

Dr. Rodrigo Coura Torres, D.Sc.

RIO DE JANEIRO, RJ – BRASIL

MARÇO DE 2018

Freund, Werner Spolidoro

Identificação de Elétrons Baseada em um Calorímetro de Altas Energias Finamente Segmentado/Werner Spolidoro Freund. – Rio de Janeiro: UFRJ/COPPE, 2018.

XXXVIII, 256 p.: il.; 29,7cm.

Orientador: José Manoel de Seixas

Tese (doutorado) – UFRJ/COPPE/Programa de Engenharia Elétrica, 2018.

Referências Bibliográficas: p. 205 – 231.

1. Redes Neurais. 2. Classificação. 3. Filtragem Online. 4. Fusão de Dados. 5. Ensemble Learning. 6. Análise de Impacto. 7. Empilhamento de Sinais. 8. Física de Altas-Energias. I. Seixas, José Manoel de. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Elétrica. III. Título.

*“Quod homo indiget potentis
sensitivis ad intelligendum.”*

Agradecimentos

Gostaria de agradecer aos meus pais pelo apoio, incentivo e compreensão. Graças a ótima criação, educação que me passaram, vocês são a razão e a luz que possibilitaram e motivaram a sede pelo desenvolvimento intelectual. Aos meus irmãos e cunhada, Stephan, Evelin e Christian, pelos bons momentos de lazer e união familiar. No caso do Christian, a realização de tarefas domésticas que garantiram o tempo vital que necessitava para possibilitar a realização desta tese. Aos meus avós, que tiveram menos tempo com o neto devido ao esforço empenhado. Ao meu tio que forneceu os *upgrades* computacionais que foram vitais para manter as engrenagens do maquinário em funcionamento.

À Alessandra por todo incentivo, compreensão, carinho e auxílio durante essa longa jornada, creio que vale aproveitar a oportunidade para me desculpar mais uma vez pelos não tão raros finais de semanas perdidos para o desenvolvimento deste trabalho.

Agradeço ao meu guia e orientador acadêmico, Seixas, que, nesse ano, compartilharemos de 10 anos de pesquisa e trabalho árduo juntos, com o sua presença e participação durante grande parte do meu desenvolvimento acadêmico. E quem diria, naquele tempo, que conseguiríamos atingir esses resultados, foi, de fato, uma longa estrada. Logo de início, quando entrei no projeto por via da iniciação científica, este ponto que chegamos era exatamente o meu objetivo, não fosse a sua supervisão, conselhos e sabedoria, nada disso seria possível. O ápice de suas contribuições ocorreram na época mais crítica do projeto, onde eu acredito ter absorvido muito conhecimento de você. Cabe aqui também dizer ao meu outro orientador acadêmico, mas também de uma diversidade de assuntos: Torres, obrigado por tudo; sou muito grato às suas orientações, e tenho certeza que como examinador você trará ainda mais contribuições para minha formação. Com certeza, a cobrança e presença de vocês foi o que me empurrou adiante de forma a atender as demandas que no início eram inusitadas para um graduando de engenharia elétrica. Ao Denis em bons momentos durante minhas estadias no CERN que foram vitais para o meu desenvolvimento academicamente, mas também como pessoa. Literalmente, foi você que abriu as portas do CERN para mim! Além disso, seu papel foi fundamental durante o comissionamento do *NeuralRinger*. Ao João, você foi um irmão

de incontáveis batalhas, linhas de código, discussões sobre *ringer* por madrugadas a dentro, correrias para completar as apresentações na virada das segundas-feiras e muitos outros perrengues que enfrentamos juntos. Não fosse você botando a mão na massa em uma ótima sinergia comigo, nunca o projeto teria chegado ao nível de excelência internacional que atingiu. A propósito, obrigado pela paciência para lidar com o chato aqui! Com certeza há muito mais que podemos alcançar juntos, as batalhas não acabaram e logo chegará a sua hora de se tornar doutor e apresentar ao mundo a outra parte do que foi feito para que este projeto se tornasse uma realidade. Mas além do trabalho, considero muito a sua honestidade, admiro sua capacidade e te respeito como um grande amigo. Aos quatro, não tenho palavras para expressar a gratidão.

Aos meus praticamente oficialmente orientados, Guilherme e Amanda, obrigado ao esforço que dedicaram ao projeto. Fico feliz em ter participado e contribuído com o melhor que podia para sua formação, mantenham o empenho em quaisquer que sejam as tarefas futuras com a qual irão lidar! Aos membros mais novos do projeto, Micael e Cadu, força, animo e trabalho em equipe, há uma grande demanda mas a recompensa vale a pena. Aos alunos de ensino médio, Hemmely, Carlos Eduardo e Guilherme, a presença de vocês no laboratório foi um prazer e fico feliz de poder ter tido vocês como alunos.

Além do Denis que teve contribuição como um orientador desde 2010, outras contribuições vieram por parte do CERN a este trabalho. Dentre eles, destaca-se o Ryan Mackenzie White, que foi uma peça chave para o avanço do projeto avançar e entrasse em operação no ATLAS. Graças a sua tradução de quais demandas eram mais urgentes por parte do ATLAS que conseguimos contribuir para o experimento com uma técnica robusta e eficiente. Outras contribuições importantes vieram de Moritz, Arantxa, Fernando e Tanya, todos *convenors* do *Trigger e/γ group*. Contribuições com outros membros do *e/γ Combined Performance group* também fazem parte do resultado final, sem contar que as análises e desenvolvimento da solução final tiveram forte influência de todo o conhecimento agregado nesse grupo. E o código de análise que também foi um dos alicerces empregado neste trabalho, tanto para a análise de impacto ou quanto na avaliação das técnicas de Aprendizado de Máquinas compondo a extensão do *ringer* para o ambiente de análise *offline*.

Ainda quanto ao desenvolvimento do trabalho, uma série de outros fatores influenciaram para que a técnica entrasse em atuação no ATLAS. Este é um projeto onde a ideia inicial tem cerca de 25 anos, logo, há um histórico muito maior que possa ser coberto em algumas linhas aqui, mas com certeza diversas contribuições de autores passados fazem parte do resultado daquilo que alcançamos. Outras pessoas que eu tenho conhecimento cujo legado contribuiu para aquilo que apenas agora se tornou realidade são Rabello, Dhiana Deva, Danilo Enoque e Torres. No caso de Ra-

bello, até hoje ainda perduram parte significativa de suas linhas de código, que foram posteriormente otimizadas por Dhiana Deva e Danilo Enoque. No caso do Danilo, foi ele quem me ensinou inicialmente como executar as minhas primeiras análises com o código do ATLAS e implementou o primeiro esboço da versão do *ringer* para análise *offline*. No caso do Torres, suas contribuições estão tanto nos códigos do algoritmo *online*, mas também o *framework* de otimização de redes neurais escrito em C++ (*FastNet*) compõe o núcleo atualmente empregado para o ajuste dos modelos neurais algumas atualizações para possibilitar o desenvolvimento de redes de fusão. Não fosse esse código, não seria possível realizar o ajuste dos milhões de modelos na LCG que, na época, não aceitava as infraestruturas de aprendizado de máquinas como o *keras*. De qualquer forma, ainda que aceitassem, não atingem nem de perto a performance de velocidade de um código implementando em C++ puro. O fato de ter isso tudo disponível nos economizou muito tempo de desenvolvimento e validação. Além disso, vocês estavam muito a frente do seu tempo, a necessidade de um método com a capacidade do *ringer* só chegou agora com o aumento do efeito de empilhamento e graças ao legado de vocês foi possível atingir o nível de excelência exigido pelo ATLAS. Outro ingrediente também se deu pelo fato de todo o legado estar justamente em um ponto crítico do HLT para a redução do estresse de processamento na fazenda e o fato da seleção original não ter sofrido intervenção para lidar com essa demanda tornou a tarefa ainda mais palpável. Graças a união de todos esses fatores, esta contribuição se tornou uma realidade, o que mostra a visão de futuro que o grupo teve. Este trabalho abre as portas e estende esse cenário para outro nível de análise, o *offline*.

Durante 1 ano morei em Genebra onde tive um momento bastante marcante de minha vida, na companhia de Victor Ferraz, Rafael Gama, João Victor, Dona Vera, Seu Eduardo e outros. Eu gostaria de agradecer muito a vocês pela companhia e pelo ótimo momento que passamos juntos. Ao Rafael, Victor e João, pelas culinárias, cervejas, filmes e passeios aleatórios e divertidos. A Dona Vera, não há palavras para expressar a gratidão pelo carinho. Ao Seu Eduardo, pelas caronas para as compras e paciência de aguentar a nossa gangue de brasileiros.

No LPS, um ótimo espírito de companheirismo tornou o desenvolvimento deste trabalho muito mais agradável. Obrigado Junior, Moura, Phil, Victor, Gama, Diego e muitos outros. No caso do Junior, devo a você os primeiros passos em *keras* para a implementação de redes especialistas e inúmeras discussões sobre Inteligência Computacional e, obviamente, diversos outros assuntos. Ah, e as aventuras para atualização e manutenção do laboratório, isso é outro ponto que não tem como esquecermos! Não tem como não olhar para trás e não ver você rindo junto comigo do perrengue que passamos. Também vale mencionar as inúmeras reuniões com o pessoal de JF, sei que em alguns casos foram complicadas, mas todos os meus

conselhos foram com a melhor intenção. À Márcia Begalli por auxiliar com uma revisão da parte da física desta tese e com ideias para a melhoria do formato.

Finalmente, aos meus amigos, sei que estive ausente, 100 % da minha desculpa para isso está contada nas páginas que seguem. . . só não sei se elas serão claras para vocês. Posso descrevê-las (ou não) de forma mais simples bebendo um chopp, mas a proposta para o chopp fica!

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

IDENTIFICAÇÃO DE ELÉTRONS BASEADA EM UM CALORÍMETRO DE ALTAS ENERGIAS FINAMENTE SEGMENTADO

Werner Spolidoro Freund

Março/2018

Orientador: José Manoel de Seixas

Programa: Engenharia Elétrica

Aplicações atuais em diferentes áreas podem se deparar com cenários envolvendo grande massas de dados, necessidade de fusão de informação, alta dimensionalidade e taxa de eventos, raros eventos de interesse e empilhamento. Todos esses aspectos estão presentes no ATLAS, o maior experimento do acelerador de partículas na vanguarda científica atual (LHC). Para isso, faz-se necessário a identificação de elétrons em um ambiente sucessivamente mais desafiador, devido ao crescimento do empilhamento de sinais, que deteriora as assinaturas dos elétrons. Empregam-se dois sistemas de detectores: calorimetria (energia) e espectrografia magnética (trajetória). A identificação ocorre em um estágio *online* (*trigger*) para depois computar-se a decisão final e de referência em *offline*. No *trigger*, utiliza-se uma sequência de estratégias híbridas, que conta com a primeira (2017) estratégia, proposta pela COPPE/UFRJ, baseada em redes neurais como método principal em experimentos similares ao ATLAS. A mesma (*NeuralRinger*) emprega anéis concêntricos de energia e é complementada, posteriormente, para a fusão de informação. Este trabalho propõe o *NeuralRinger* para a atuação *offline* e, com isso, inverte a lógica comum de desenvolvimento. Adicionou-se fusão de outras representações de informação via o ajuste por redes neurais especialistas. Desenvolveu-se uma infraestrutura completa para o complexo ambiente de análise do ATLAS. Os resultados em simulação, comparando a *NeuralRinger* com a referência (verossimilhança), mostram uma redução na taxa de falsos elétrons de 2,54 % para 1,13 % (incerteza desprezível) quando ambas operam na mesma detecção. A atuação do *NeuralRinger* no *trigger* exigiu avaliar o seu impacto no ambiente *offline*. Propôs-se um método estatístico de análise, onde se observou, em dados de colisão de 2017, que a distorção é mínima ($< 1\sigma$) e, mesmo assim, positiva, ao coletar amostras de melhor qualidade.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

ELECTRON IDENTIFICATION BASED ON A HIGH ENERGY CALORIMETER OF THIN SEGMENTATION

Werner Spolidoro Freund

March/2018

Advisor: José Manoel de Seixas

Department: Electrical Engineering

Current applications in several areas may face scenarios involving big data, need for information fusion, high dimensionality and event-rate, rare events and signal pile-up. These features are found in the ATLAS, the largest experiment of the particle accelerator in the current scientific edge (LHC). Electron identification is vital for this purpose, currently subject to ever-increasing pile-up conditions. Detector systems provide discriminant information of distinct nature: calorimetry (energy) and magnetic spectrography (tracking etc.). The identification occurs in an online stage (trigger) followed by an offline computation responsible for the ultimate and benchmark decision. In the trigger, it is employed a decision chain of hybrid methods which counts with the first strategy based on neural networks, proposed by COPPE/UFRJ, operating in experiments similar to ATLAS. This technique (NeuralRinger) is based only on concentric rings of energy and is posteriorly complemented by a decision based on likelihood for the data fusion. This work purposes the NeuralRinger for offline operation and, thereby, inverts the standard logic of development. To allow its operation, other representations of information were added through training specialist neural networks. It was developed a complete framework for the ATLAS, a complex analysis environment. The results, comparing the NeuralRinger with the reference (likelihood), show a fake electron rate reduction from 2.54 % to 1.13 % (negligible uncertainty) when operating at the same detection rate. The NeuralRinger operation in the trigger required the evaluation of its impact in the offline environment. It was purposed a statistical analysis method, where it was observed in 2017 collision data that the distortion is minimal ($< 1\sigma$) and, nevertheless, favorable, considering that better quality samples are collected.

Sumário

Sumário	xi
Lista de Figuras	xv
Lista de Tabelas	xx
Lista de Símbolos	xxiii
Lista de Abreviaturas	xxviii
Glossário	xxxv
1 Introdução	1
1.1 Motivação	3
1.2 Objetivo	5
1.3 Contribuições do Trabalho	5
1.4 Organização do Documento	6
2 Física Experimental de Altas Energias	8
2.1 Física de Partículas Elementares e o Modelo Padrão	9
2.1.1 Léptons	10
2.1.2 Quarks e hádrons	10
2.1.3 As Interações Fundamentais	11
2.2 <i>Centre Européene pour la Rechère Nucleaire</i> (CERN)	12
2.3 <i>Large Hadron Collider</i> (LHC)	13
2.3.1 Complexo de Aceleração	14
2.3.2 Colisões e Aspectos Relacionados	17
2.3.2.1 Coordenadas e Projeções	17
2.3.2.2 Categorização de Colisões de Partículas	17
2.3.2.3 Seção de Choque e Luminosidade	18
2.4 Detectores de Partículas	24
2.4.1 Interação de Partículas com a Matéria	24
2.4.1.1 Partículas Carregadas	25

2.4.1.2	Fótons	29
2.4.1.3	Comprimentos de Radiação e de Interação Nuclear	31
2.4.2	Detecção de Partículas via Calorimetria	34
2.4.2.1	Chuveiro de Partículas	36
2.4.2.1.1	Chuveiro de Partículas Eletromagnéticas	37
2.4.2.1.2	Chuveiro de Partículas Hadrônicas	39
2.4.2.2	Medição e Resolução em Energia	41
2.4.2.3	Aspectos Construtivos	43
2.4.2.4	Discussão	45
3	Levantamento Bibliográfico	48
3.1	Inteligência Computacional no Contexto de Física de Partículas de Altas Energias	48
3.2	Fusão de Informação	49
3.2.1	Técnicas Tradicionais	49
3.2.2	Das Demandas de IoT	51
3.2.3	Inteligência Computacional Aplicada a Fusão de Dados	53
3.2.3.1	Técnicas Tradicionais	53
3.2.3.2	Aprendizado Profundo	55
4	O Experimento ATLAS	59
4.1	Detector de Traços	61
4.1.1	Detector de Pixel	63
4.1.2	Detector de Rastreamento por Semicondutores (SCT)	63
4.1.3	Detector de Rastreamento por Transição de Radiação (TRT)	64
4.1.4	Solenóide Central (CS)	64
4.2	Sistema de Calorimetria	65
4.2.1	Calorímetro Eletromagnético (ECAL)	67
4.2.2	Calorímetro Pré-Amostrador (PS)	70
4.2.3	Calorímetro Hadrônico (HCAL)	71
4.2.3.1	Calorímetro Hadrônico de Telhas (<i>TileCal</i>)	71
4.2.3.2	Calorímetro Intermediário de Telhas Cintilantes (ITC)	73
4.2.3.3	Tampas do Calorímetro Hadrônico (HEC)	74
4.2.4	Calorímetro Dianteiro (FCAL)	74
4.2.5	Sistema de Criogenia	75
4.3	Sistema de Reconstrução da Física	76
4.4	Sistema de Filtragem	79
4.5	Outros Sistemas	82

5	Identificação de Elétrons no ATLAS	84
5.1	Reconstrução de Elétrons	85
5.2	As Grandezas Físicas para Discriminação de Elétrons	86
5.3	Estratégia para Discriminação e Avaliação de Performance	91
5.4	Critério de Isolamento em Elétrons	93
5.5	Resiliência ao Efeito de Empilhamento	95
5.6	Método <i>Tag and Probe</i> (T&P)	96
5.7	Cadeias para Filtragem de Elétrons	98
5.7.1	Atualizações durante a <i>Run2</i>	102
5.7.2	Discussão	106
6	Soma em Anéis para a Descrição de Chuveiros de Partículas	109
6.1	Informação Anelada de Calorimetria	109
6.2	Histórico	112
6.2.1	Proposta inicial	113
6.2.2	Contribuições Posteriores	116
6.2.3	Discussão	124
6.2.4	Emprego de Anéis Além do ATLAS	128
7	Algoritmo <i>NeuralRinger</i>	130
7.1	<i>NeuralRinger Online</i>	130
7.1.1	Método para Tomada de Decisão	132
7.1.1.1	Discussão	137
7.1.2	Método de Análise	139
7.2	Análise de Impacto na Coleta de Elétrons	140
7.3	<i>NeuralRinger Offline</i>	144
7.3.1	Ajuste de Redes Especialistas para Fusão de Informação	145
7.3.2	Método de Ajuste para a Tomada de Decisão	147
7.3.3	Método de Análise	151
7.3.4	Infraestrutura para Atuação no ATLAS	151
7.3.4.1	Otimização para Produção Oficial de Dados	153
7.4	Infraestrutura para Ajuste de Modelos	155
8	Resultados	159
8.1	Nomenclatura Empregada	159
8.2	Identificação <i>Online</i> de Elétrons	160
8.3	Análise de Impacto na Coleta de Elétrons	163
8.4	Identificação <i>Offline</i> de Elétrons	166

9	Limitações e Aperfeiçoamentos	175
9.1	Estratégia para Homogeneização dos Subconjuntos de Validação Cruzada	176
9.2	Estratégia para Redução da Demanda Computacional para o Ajuste de Modelos em Grande Massa de Dados	183
9.3	Ajuste Envolvendo Aprendizado Profundo por <i>Stacked Autoencoder</i> .	184
9.4	Pré-processamento por ICA	186
9.5	Considerações sobre a Janela	187
9.6	Tratamento de Regiões Anômalas	188
9.7	Descrição por Densidade de Energia	195
9.8	Adição de Informações de Outras Naturezas a partir do Sistema de Calorimetria	196
9.9	Região do <i>Crack</i>	197
10	Conclusões	198
10.1	Perspectivas e Trabalhos Futuros	201
	Referências Bibliográficas	205
A	Produção Acadêmica	232
A.1	Artigos Publicados em Revista	232
A.2	Artigos Submetidos em Revista	235
A.3	Artigos Publicados em Anais de Congresso	242
A.4	Resumos Publicados em Anais de Congresso	245
A.5	Autoria no Experimento ATLAS	250
A.6	Educação e Popularização de Ciência e Tecnologia	250
B	Detalhes da Infraestrutura em Atuação no SR_{off}	252
C	Pseudo-código do Algoritmo RPROP	255

Lista de Figuras

2.1	A cadeia de aceleração do LHC	15
2.2	Esboço esquemático do LHC	16
2.3	O Sistema de Coordenadas adotado para o ATLAS	18
2.4	Esboços de categorias de colisões na Física de Partículas	19
2.5	Resumo da medição pelo ATLAS de seção de choque da produção total de alguns processos físicos selecionados e apresentado em função da energia de massa (\sqrt{s}) de 7 a 13 TeV.	20
2.6	Um sumário dos resultados do ATLAS para medições de diversas seções de choque (σ) para a produção total e fiducial descrita pelo Modelo Padrão, comparada caso a caso com o esperado teórico. . . .	21
2.7	Valores medidos pelo <i>A Toroidal LHC Apparatus</i> (ATLAS) de luminosidade instantânea e integrada.	22
2.8	Número médio de interações por evento de cruzamento de pacotes para as Run 1 e 2.	23
2.9	Exemplos de descrição para a perda de energia por ionização e excitação para múons atravessando silicene.	26
2.10	Fração de perda de energia média em função da energia da partícula incidente para conhecidas interações de elétrons e pósitrons com o meio.	28
2.11	Comparação da seção de choque total em valores teóricos e as respectivas observações experimentais para fótons transpassando chumbo em função de sua energia.	31
2.12	Comprimento de interação nuclear e de radiação para elementos químicos com $Z > 20$	34
2.13	Uma simulação de um chuva desenvolvida pela incidência de um elétron em ferro.	38
2.14	Perda média de energia de múons em função de sua energia.	46
4.1	O Experimento ATLAS e seus principais sistemas de detectores	59
4.2	Esboço contendo exemplos de interação de partículas com os subsistemas do ATLAS	60
4.3	O Detector Interno (ID)	62

4.4	Os diversos subsistemas de calorimetria do ATLAS	67
4.5	Seção de corte longitudinal do ECAL	68
4.6	Esboço da granularidade das camadas do barril do ECAL e um esboço da geometria de acordeão.	69
4.7	Simulação do comprimento total em termos de X_0 e em função de η transpassado por uma partícula até o final do ECAL	71
4.8	Corte esquemático das células do Barril, Barril Estendido e Calorímetro Intermediário do Calorímetro Hadrônico de Telhas (<i>TileCal</i>)	72
4.9	Simulação do comprimento em λ_{int} e em função de η para os calorímetros do Sistema de Calorimetria do ATLAS	73
4.10	As diferentes estruturas dos subsistemas de calorimetria do ATLAS .	75
4.11	Esboço do fluxo da informação envolvendo a análise física contendo elétrons como parte dos canais avaliados no Experimento ATLAS. . .	77
4.12	Esboço da atuação da filtragem <i>online</i> realizada no ATLAS.	79
4.13	O TDAQ do ATLAS em atuação durante a <i>Run 2</i>	81
4.14	Taxas por grupos de eventos de interesse de filtragem no HLT em função do tempo em um período de coleta obtido em junho de 2017 e com pico de luminosidade de $1,53 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$	83
5.1	Esboço dos subdetectores fornecendo informação para um candidato a elétron atravessando o ATLAS.	88
5.2	Perfil em dados de simulação com condições observadas durante colisões $p - p$ de 2016 para a grandeza E_{ratio} conforme reconstruída em <i>offline</i> para elétrons e ruído físico.	89
5.3	Perfil em dados de simulação com condições observadas durante colisões $p - p$ de 2016 para a grandeza TRT_PID conforme reconstruída em <i>offline</i> para elétrons e ruído físico.	90
5.4	Esboço do comprimento simulado em unidades de X_0 do ECAL e λ_{int} do Sistema de Calorimetria contido nas regiões em η empregadas pelo ATLAS para medição de eficiência da tarefa de discriminação elétron-hádron.	94
5.5	Esboço do T&P para o bóson Z decaindo em dois elétrons.	97
5.6	Esboço dos estágios de seleção e a informação disponível para análise para a cadeia de elétrons.	99
5.7	A sequência de reconstrução e seleção empregada em cadeias de seleção de elétrons e fótons durante a <i>Run2</i>	100
5.8	A lógica da janela deslizante aplicada no L1 para seleção de elétrons e fótons.	101

5.9	Taxa de saída no L1 e HLT para cadeias de seleção de elétrons em função da luminosidade instantânea não calibrada computada <i>online</i> durante o período completo de coleta de colisões p-p com centro de massa de 13 TeV em 2016.	103
5.10	Estágios de seleção avaliados pela cadeia original (<i>noringer</i>) considerando o menu de 2016 e a cadeia proposta em desenvolvimento. . . .	104
5.11	Estágios de seleção avaliados pela cadeia atuante em 2017 e sua versão duplicada com o método <i>CutBased</i> atuando no FastCalo.	105
5.12	Esboço da dinâmica dos sistemas atuando desde a tomada de dados até os resultados das análises finais do Experimento ATLAS.	107
6.1	Esquema do funcionamento do algoritmo <i>Ringer</i> para extração da base de informação anelada de calorimetria	111
6.2	Variação da probabilidade de detecção (em cima) e do falso alarme (embaixo) de diferentes normalizações avaliadas por TORRES em função da energia transversa do evento.	127
7.1	Número de amostras disponível por região do espaço de fase empregado para o ajuste dos modelos neurais a partir de dados de simulação disponibilizados para o desenvolvimento das estratégias de seleção do SF_{on} para 2017.	134
7.2	Esboço do comprimento simulado em unidades de X_0 do ECAL e λ_{int} do Sistema de Calorimetria contido nas regiões em η empregadas para análise da informação anelada de calorimetria.	135
7.3	Saída da rede neural compondo o <i>ensemble</i> neural que entrou em operação em 2017 com a função de ativação não-linear removida após o estágio de treinamento para elétrons e ruído físico.	137
7.4	Saída e P_D da rede neural compondo o <i>ensemble</i> neural que entrou em operação em 2017 com a função de ativação não-linear removida após o estágio de treinamento para elétrons em função de $\langle \mu \rangle$ $0,8 < \eta < 1,37$ e $20 < E_T < 30$	138
7.5	Esboço de um dos ajustes de redes neurais especialistas realizado neste trabalho para comporem o <i>ensemble</i> de redes neurais do <i>NeuralRinger offline</i> e suas respectivas etapas.	146
7.6	Perfil de <i>nvtx</i> para as simulações de MC empregadas no estudo para o <i>NeuralRinger offline</i> e dos dados de colisão coletados pelo ATLAS durante o ano de 2016 aplicando o método T&P para decaimentos $Z \rightarrow ee$ e exigindo, no caso de colisão, o critério <i>vloose</i> para remoção de impurezas.	148

7.7	Número de amostras disponível por região do espaço de fase empregado para o ajuste dos modelos neurais a partir de dados de simulação disponibilizados para o desenvolvimento de <i>ensemble</i> neurais para atuação no Sistema de Reconstrução (<i>Offline</i>) de Física do ATLAS.	148
7.8	Perfil de E_T e η para as simulações de MC empregadas no estudo para o <i>NeuralRinger offline</i> .	149
7.9	Módulos implementados para a geração da informação anelada e aplicação de seleção a partir de técnicas de Inteligência Computacional dentro da infraestrutura comum (<i>Athena</i>) do ATLAS.	152
8.1	Comparação da taxa de detecção na saída do HLT para cadeias primárias empregando o método <i>CutBased</i> antes da TS1 com o método <i>NeuralRinger</i> após a TS1 para dados coletados até outubro de 2017.	162
8.2	Comparação da taxa de falso alarme no estágio FastCalo para uma cadeia primária com emprego de pré-escala, utilizando o método <i>CutBased</i> antes da TS1, com o método <i>NeuralRinger</i> , após a TS1, para dados coletados até outubro de 2017	162
8.3	Análise de quadrante em η para as grandezas físicas R_η e R_{had} nas cadeias <i>ringer</i> e <i>noringer</i> atuando no critério <i>medium</i> e avaliada para dados coletados até outubro de 2017.	163
8.4	Eficiências e sua razão na saída do HLT em função de η (a) e $\langle\mu\rangle$ (b) para cadeias de seleção <i>ringer</i> (cinza) e <i>noringer</i> (azul) duplicadas para a análise de impacto durante toda a tomada de dados para colisões $p - p$ à 13 TeV.	164
8.5	Análise de impacto no <i>offline</i> ao alterar a sistemática do método T&P para as grandezas físicas R_η e R_{had} para as cadeias <i>ringer</i> e <i>noringer</i> em dados coletados para o ano de 2017.	165
8.6	Análise de impacto para a grandeza R_η reconstruída pelo <i>offline</i> considerando as amostras coletadas durante 2017 exclusivamente pela cadeia <i>ringer</i> ou <i>noringer</i>	167
8.7	Análise de impacto para a grandeza R_{had} reconstruída pelo <i>offline</i> considerando as amostras coletadas durante 2017, exclusivamente pela cadeia <i>ringer</i> ou <i>noringer</i> .	168
8.8	Eficiência em T&P e sua razão em função de E_T e η para as estratégias de seleção de elétrons desenvolvidas para atuação no SR_{off} e a atual referência atuando no critério <i>vloose</i> .	171

8.9	Falso alarme e sua razão em função de E_T e η para as estratégias de seleção de elétrons desenvolvidas para atuação no SR_{off} e a atual referência atuando no critério <i>vloose</i>	172
8.10	Eficiência em T&P e falso alarme em função de $nvtx$ para as estratégias de seleção de elétrons desenvolvidas para atuação no SR_{off} e a atual referência atuando no critério <i>vloose</i>	173
8.11	Análise de quadrante em para a grandeza física R_η de $ExNN(rings + track^+)$ e $ExNN(rings + shower + track^+)$ em relação à referência (<i>LLH</i>)	174
8.12	Análise de quadrante em ruído físico para a grandeza física R_η de $ExNN(rings + shower + track^+)$ em relação à referência (<i>LLH</i>).	174
9.1	Mapa obtido para simulações em condições compatíveis com colisões de 2016 contendo decaimentos $Z \rightarrow ee$ aplicando o método T&P na região $2,37 < \eta < 2,47$ e $30 < E_T < 40$ GeV.	181
9.2	Mapa obtido para simulações em condições compatíveis com colisões de 2016 contendo decaimentos $Z \rightarrow ee$ aplicando o método T&P na região $2,37 < \eta < 2,47$ e $30 < E_T < 40$ GeV.	182
9.3	Evolução do MSE em função das épocas de treinamento da primeira camada do SAE, com 100 neurônios na camada escondida, quando empregando a otimização por gradiente descendente (a) e ADAM (b).	185
9.4	Informação mútua para as segmentações longitudinais em camadas ou seções do Sistema de Calorimetria para a aplicação da primeira camada do SAE em função do número de neurônios empregados. No eixo da direita, o valor de SP_{max} medidos com validação cruzada por <i>k-fold</i> para um classificador MLP de uma camada escondida alimentado pelo espaço na camada interna do SAE.	186
9.5	Mapa extraído conforme indicado na Figura 9.1c, porém alterando a região em η para $1,54 < \eta < 2,37$	189
B.1	Esquema da infraestrutura implementada para o algoritmo <i>Ringer</i> no Sistema de Reconstrução.	253
B.2	Fluxo de informação de classificação para o caso de aplicação menos segmentado e mais segmentado	254

Lista de Tabelas

2.1	Os léptons, suas gerações, carga elétrica em unidades de carga absoluta do elétron e massa.	10
2.2	Os quarks, suas gerações, carga elétrica em unidades de carga absoluta do elétron e massa.	10
2.3	Resolução em energia de calorímetros eletromagnéticos em experimentos, com a respectiva tecnologia empregada e data do projeto. . .	44
4.1	Região de cobertura, granularidade e número de células de leitura dos calorímetros do ATLAS	66
5.1	Definição das grandezas empregadas pelo ATLAS para discriminação de elétrons.	87
5.2	Cobertura das regiões em η empregadas para as análises realizadas pela ATLAS.	93
5.3	Fronteira das regiões em E_T empregadas pelo ATLAS para avaliação de eficiência, aplicação de limiares de decisão e ajuste dos modelos da <i>LLH</i>	93
6.1	Descrição das amostragens longitudinais dos calorímetros compondo camadas e seções no algoritmo de anelamento.	112
6.2	Anéis ou camadas mais relevantes obtidas para as contribuições realizadas para o Algoritmo Ringer.	126
7.1	Parâmetros empregados para ajuste, escolha de topologia do modelo e escolha de modelo para operação para as redes Sistema de Filtragem do ATLAS.	133
7.2	Casos eutualmente exclusivos para a atuação das cadeias <i>ringer</i> e <i>noringer</i> em paralelo durante a tomada de dados.	142

7.3	Parâmetros empregados para ajuste, escolha de topologia do modelo e escolha de modelo para operação no <i>ensemble</i> do <i>NeuralRinger</i> ou para a extração das redes especialistas em informação anelada para atuação Sistema de Reconstrução (<i>Offline</i>) de Física do ATLAS (SR_{off}). Exceto quanto especificado, os parâmetros empregados foram os padrões empregados nas técnicas empregadas.	157
7.4	Parâmetros empregados para ajuste, escolha de topologia do modelo e escolha de modelo para operação no <i>ensemble</i> neural ou para a extração das redes especialistas em grandezas de traço e fusão especialista calorimetro-traço para atuação Sistema de Reconstrução (<i>Offline</i>) de Física do ATLAS (SR_{off}). Os demais critérios e valores foram os mesmos que o ajuste para a informação anelada.	158
7.5	Parâmetros empregados para ajuste, escolha de topologia do modelo e escolha de modelo para operação em um <i>ensemble</i> neural ou para a extração das redes especialistas em grandezas de calorimetria para atuação Sistema de Reconstrução (<i>Offline</i>) de Física do ATLAS (SR_{off}). Os demais critérios e valores foram os mesmos que o ajuste para a informação anelada.	158
7.6	Parâmetros ou critérios envolvidos especificamente para o treinamento de modelos neurais para fusão de informação para comporem o <i>ensemble</i>	158
8.1	Evolução da eficiência para os estágios de seleção do HLT para as cadeias <i>ringer</i> e <i>noringer</i> em previsões <i>offline</i> para o menu de 2016 em dados de colisão e simulação.	161
8.2	Diferenças de critérios exigidos no T&P entre as análises de impacto e de quadrante avaliados neste trabalho.	165
8.3	Valores de eficiência (P_D , SP e P_F) para a referência (LLH) medida em toda estatística disponível e do <i>ensemble</i> neural operando no ponto mais próximo ao P_D (coluna ressaltada em verde) da referência e apresentado como o valor central e 1 RMS das eficiências de validação cruzada. Nomenclatura empregada disponível na Seção 8.1.	169
8.4	Valores de eficiência (P_D , SP e P_F) para a referência (LLH) medida em toda estatística disponível e do <i>ensemble</i> neural operando no ponto mais próximo ao P_D (coluna ressaltada em verde) da referência e apresentado como o valor central e 1 RMS das eficiências de validação cruzadas integrando cada região do espaço de fase. A incerteza de validação cruzada é desprezível. Nomenclatura empregada disponível na Seção 8.1.	170

9.1	Resultados para o ajuste de SOM na informação anelada de calorimetria processada por PCA e de <i>k-means</i> para o agrupamento dos neurônios do mapa em dados de simulação compatíveis com as condições de colisão de 2016 para amostras de elétrons pré-selecionadas pelo método T&P.	179
9.2	Resultados para o ajuste de SOM na informação anelada de calorimetria processada por PCA e de <i>k-means</i> para o agrupamento dos neurônios do mapa em dados de simulação compatíveis com as condições de colisão de 2016 para amostras de ruído físico.	180
9.3	Exemplos de matrizes de correção de região anômala para a EM1 na alteração de granularidade ocorrendo em $ \eta = 1, 8$	194

Lista de Símbolos

A	peso atômico. 30
B	campo magnético. 61
d_0	ver Tabela 5.1. 61, 63, 87, 91, 102, 158
d_0/σ_{d_0}	ver Tabela 5.1. 87, 158
DB	índice de Davies-Bouldin. 177–182
$-\langle \frac{dE}{dx} \rangle$	perda média de energia por comprimento do meio percorrido. 25, 26, 29, 32
$\Delta p/p$	ver Tabela 5.1. 87, 158
$\frac{\Delta p}{x}$	perda de energia mais provavel em um material com espessura x . 26
ΔR	$\Delta R^2 = \eta^2 + \phi^2$. xxiv, xxvi, 94, 98
$\Delta \eta_1$	ver Tabela 5.1. 87, 158
D_{KL}	divergência Kullback-Leibler. 144
$\Delta \phi_2$	ver Tabela 5.1. 87, 158
$\Delta \phi_{res}$	ver Tabela 5.1. 87
E_c	energia crítica. 27, 32, 37, 39, 41, 47
E_{EM1}	energia medida na primeira camada do ECAL. 118
$E_{T,EM}$	energia transversa calculada no ECAL. 117
$E_{HAD,1}$	energia depositada no <i>cluster</i> para a primeira camada HAD. 118
EM1	primeira camada EM. 69, 85, 88, 89, 93, 110, 112, 119, 120, 122, 125, 126, 160, 170, 189, 194
EM2	segunda camada EM. 70, 85, 86, 101, 102, 110, 112, 118–120, 125–127, 153, 163, 185
EM3	terceira camada EM. 70, 85, 110, 112, 123, 125, 126, 185
E/p	ver Tabela 5.1. 87

eProbabilityHT	ver Tabela 5.1. 64, 86, 87
E_{ratio}	ver Tabela 5.1. xvi, 87–89, 101, 117, 118, 146, 151, 158, 159, 170, 200
E_T	energia transversa. Ver Tópico 2.3.2.1. xvii–xix, xxiv, xxxvii, 6, 17, 70, 71, 78, 82, 85, 87, 89, 90, 92–94, 97, 98, 100, 103, 119, 122, 125, 126, 128, 131–133, 137–139, 141, 143, 147–149, 154, 157, 160, 161, 166, 171, 172, 181, 182, 184, 187, 199–201, 250
η	pseudorrapidez. xvii–xx, xxiii, xxiv, xxxvii, 17, 61–66, 68–75, 78, 82, 83, 85, 88–94, 100–102, 110, 118, 122, 126, 132–135, 137, 138, 143, 147–149, 157, 160, 162, 164, 166, 170–172, 181–184, 187–189, 197, 199, 200
$\eta_{0,uh}$	primeira região empregada em análises da Colaboração, compreendida entre $0 < \eta < 0,6$. xxiv, 93, 94
$\eta_{1,uh}$	idem $\eta_{0,uh}$, porém $0,6 < \eta < 0,8$. 93, 94
$\eta_{2,uh}$	idem $\eta_{0,uh}$, porém $0,8 < \eta < 1,15$. 93
$\eta_{3,uh}$	idem $\eta_{0,uh}$, porém $1,15 < \eta < 1,37$. 93
$\eta_{4,uh}$	idem $\eta_{0,uh}$, porém $1,37 < \eta < 1,52$. 93
$\eta_{5,uh}$	idem $\eta_{0,uh}$, porém $1,52 < \eta < 1,81$. 93
$\eta_{6,uh}$	idem $\eta_{0,uh}$, porém $1,81 < \eta < 2,01$. 93
$\eta_{7,uh}$	idem $\eta_{0,uh}$, porém $2,01 < \eta < 2,37$. 93
$\eta_{8,uh}$	idem $\eta_{0,uh}$, porém $2,37 < \eta < 2,47$. 93, 94
$E_T^{cone,\Delta R}$	isolamento em energia. 94
$E_{T,HAD1}$	energia transversa calculada para a primeira camada do HCAL. 117
\cancel{E}_T	energia transversa faltante. xxxvi, 35, 47, 65, 73, 74, 80, 82, 83
f_1	ver Tabela 5.1. 87, 146, 158
f_3	ver Tabela 5.1. 87, 125, 146, 158
f_{em}	fração de energia eletromagnética. 39, 42
f_{inv}	fração de energia invisível. 40, 42, 127
f_{rev}	frequência de revolução. 15, 20
f_{RF}	frequência do sistema RF. 15

HAD1	primeira camada HAD. 110, 112, 125, 126, 185, 197
HAD2	segunda camada HAD. 110, 112, 122, 123, 125, 126, 197
HAD3	terceira camada HAD. 110, 112, 119, 122, 123, 125, 185, 186, 197
J/Ψ	méson J/Ψ , partícula instável que pode decair em elétrons. 202
λ_{int}	comprimento de interação nuclear. 33, 34, 39, 40, 43, 45, 67, 71, 73, 74, 94, 135
LAr	argônio líquido. 25, 44, 64, 65, 67–70, 72, 74–76, 95, 129
l	comprimento. 61, 62
\mathcal{L}	luminosidade. 19, 20, 105, 119
$\langle\mu\rangle$	número médio de interações por cruzamento de pacote. xvii, xviii, 23, 61, 92, 95, 96, 102, 105, 133, 136–138, 164
N_A	número de avogrado. 30
n_{blayer}	ver Tabela 5.1. 87
$\lceil n \rceil$	inteiro mais próximo de n no sentido de $+\infty$. Em programação, dá-se pela operação <i>ceil</i> . 110
$\lfloor n \rfloor$	inteiro mais próximo de n no sentido de $-\infty$. Em programação dá-se pela operação <i>floor</i> . 110
n_{pixel}	ver Tabela 5.1. 87
$\lceil n \rceil$	inteiro mais próximo de $n + 0,5$ no sentido de $-\infty$. Em programação, dá-se pela operação <i>round</i> . 110
n_{Si}	ver Tabela 5.1. 87
$nvtx$	número de vértices. xvii, xix, 95, 147, 148, 157, 172, 173, 181–183
\oplus	soma em quadratura: $a \oplus b = \sqrt{a^2 + b^2}$. 44, 63

P_D	probabilidade de detecção de sinal. xvii, xxi, xxxvii, 116, 121, 122, 124, 126–129, 133–136, 138, 143, 149, 155, 157, 169–171
$P_{D,sp_{max}}$	probabilidade de detecção de sinal no ponto de SP_{max} . 117–124
P_F	probabilidade de falso alarme. xxi, xxxvii, 96, 116, 121, 122, 124, 128, 129, 133–137, 149, 155, 157, 169–173
ϕ	ângulo azimutal. xxiii, 61, 63, 65, 66, 68, 70, 88, 94, 102, 110, 188
$p - p$	colisão próton-próton. xvi, xviii, 13, 17, 22, 23, 61, 83, 89, 90, 95, 99, 103, 123, 133, 147, 157, 159, 164, 165
p_T	momento transverso. xxvi, 63, 78, 85, 89, 90, 94, 95, 98, 102, 103, 148
$p_T^{\text{varcone},\Delta R}$	isolamento variável em traço. 94, 98
Q_e	erro de quantização. 177, 179, 180
R	anel. xxxvi, 110, 122, <i>veja no glossário</i> soma em anel concêntrico de energia
R_η	ver Tabela 5.1. xviii, xix, 87, 95, 101, 117, 118, 124, 146, 151, 158, 159, 162–167, 172, 174, 200
R_{had}	ver Tabela 5.1. xviii, 87, 95, 101, 158, 159, 163–166, 168
R_{had1}	ver Tabela 5.1. 87
R_M	raio de Molière. 39, 109
R_ϕ	ver Tabela 5.1. 87, 146, 151, 158, 200
σ_p	resolução de posição. 61–63
Sil	índice de silhueta. 178–182, 184
$\langle Sil \rangle_{max}$	silhueta média máxima. 178–180
SP	índice soma-produto. xxi, xxvi, 117, 118, 120, 121, 126, 136, 149, 169, 170, 185
SP_{max}	SP máximo. xix, xxvi, 117–121, 134, 155, 157, 186
T_0	comprimento de traço total. 41

T_e	erro topográfico. 177–180, 183
TRT_PID	ver Tabela 5.1. xvi, 87, 89, 90, 93, 158, 166
$w_{\eta 2}$	ver Tabela 5.1. 87, 118, 124, 146, 151, 158, 200
w_{stot}	ver Tabela 5.1. 87
X_0	comprimento de radiação. 28, 31–34, 37–39, 41, 43–45, 67, 70, 71, 74, 94, 135
Z	número atômico. 29, 33, 34
Z'	<i>Z prime</i> . 82, 84

Lista de Abreviaturas

No caso de algumas abreviaturas internacionalmente conhecidas, optou-se por mantê-las em sua língua original.

AD	Anderson-Darling. 142
ALFA	Luminosidade Absoluta para o ATLAS. 83
ALICE	<i>A Large Ion Collider Experiment</i> . 14, 15
ALP	Partículas Compatíveis com Axions. <i>Axion-like Particles</i> . 202
AOD	<i>Analysis Object Data</i> . xxxiv, 77
ATLAS	<i>A Toroidal LHC Apparatus</i> . vi, vii, ix, x, xv–xviii, xx, xxi, xxviii, xxxiii–xxxvii, 2–8, 14, 15, 17, 18, 20–24, 43, 44, 46–48, 52, 59–67, 71, 73, 75–88, 91–94, 96–98, 101, 102, 106–109, 111–114, 116–118, 121, 123–130, 132, 133, 138, 140, 145, 148, 151–155, 157, 158, 175, 176, 184, 186–188, 193–196, 198, 201–203, 250, 253
BCM	Monitoramento das Condições dos Feixes. 83
BDT	<i>Boosted Decision Tree</i> . 48
BMU	Melhor Unidade Máquina. 177–180
BOOSTER	<i>Proton Synchrotron Booster</i> . 14
BPTX	Serviços Eletrostáticos de Coleta de Feixes. 82
BT	Barril do Toroide. 64, 65
CERN	<i>Centre Européene pour la Rechère Nucleaire</i> . v, 2, 8, 12–14, 106
CMS	<i>Compact Muon Sollenoid</i> . 14, 15, 44, 46
CNN	Rede Neural Convolutiva. 55, 57
COPPE/UFRJ	Instituto Alberto Luiz Coimbra de Pós-graduação e Pesquisa de Engenharia. ix, x, xxxi, 4, 7, 59, 198, 202, 250, 251

CRNN	Rede Neural Recorrente de Valores Complexos. 55
CS	Solenóide Central. 61, 64, 65, 70, 71, 187
CTP	Processador Central de Filtragem. 79, 80
DØ	Experimento DØ. 44, 49, 128, 129
DBN	<i>Deep Belief Networks</i> . 56, 57, 147
DL	Aprendizado Profundo. 49, 55–58, 147, 153, 184
ECAL	Calorímetro Eletromagnético. xxiii, 40, 43–46, 64, 65, 67–72, 75, 76, 85, 94, 100, 118, 135, 145, 149, 196
ECT	Tampa do Toroide. 64, 65
EDM	Modelo de Dados do Evento. 110
EF	Filtro de Eventos. 112
EM	Eletromagnético. xxiii, 37, 39–46, 66, 69, 70, 74, 85, 87, 95, 99, 101, 109, 111, 114–116, 125, 127, 153, 185, 187, 195, 252
EMB	Barril do Calorímetro Eletromagnético. 66, 67, 112
EMEC	Tampa do Calorímetro Eletromagnético. 66, 67, 76, 112
FastCalo	estágio <i>online</i> de pré-seleção rápida baseada em informação de calorimetria. xvii, xviii, 100, 103–105, 107, 112, 113, 130–132, 134–136, 159, 161, 162, 196
FastElectron	estágio <i>online</i> de pré-seleção rápida de elétrons baseada em fusão da informação de calorimetria e do ID. 100, 103, 104, 135, 161
FCAL	Calorímetro Dianteiro. 65, 66, 74–76
Geant	<i>Geaometry and Tracking</i> . 25, 77, 129
GRL	Lista de <i>runs</i> em Boas Condições. <i>Good Run List</i> . 133
GSF	<i>Gaussian Sum Filter</i> . 86

HAD	Hadrônico. xxiii, xxv, 37, 39, 40, 42, 43, 45, 46, 65, 66, 73, 74, 101, 114, 119, 125, 126, 153, 185, 187, 195
HCAL	Calorímetro Hadrônico. xxiv, 40, 43–46, 65, 71, 73, 74, 95, 101, 118, 125, 145, 163
HEC	Tampa do Calorímetro Hadrônico. 65, 66, 74–76, 112
HEP	Física de Altas Energias. 76
LF	Sabores Leves. 11, 91, 94
HLT	Alto Nível de <i>Trigger</i> . vii, xviii, xxxv, xxxviii, 3, 4, 79–81, 98, 100–104, 112, 131, 133, 135, 136, 140, 145, 160–162, 164, 196, 204
HLTCalo	estágio <i>online</i> de pré-seleção precisa de elétrons baseada em informação de calorimetria. 100, 103, 104, 133–135, 161
HMM	Modelo Oculto de <i>Markov</i> . 51
<i>IBL</i>	Camada-b Interna Removível. 62, 63, 81
ICA	Análise de Componentes Independentes. 5, 120–123, 175, 186, 187, 203
ID	Detector Interno. xxix, 52, 59–64, 68, 71, 73, 77, 82, 85, 98–100, 125, 130, 145, 146, 158, 160, 166, 168, 199
IMO	Objeto Independente em Locomoção. 55, 56
IoT	<i>Internet</i> das Coisas. 2, 50–53, 57, 58
IP	Ponto de Interação. 14, 17, 20, 59, 63, 70, 105
IR	Região de Interação. 14–16
ITC	Calorímetro Intermediário de Telhas Cintilantes. xxxiv, 66, 73, 111, 112, 120, 197
JDL	<i>Joint Directors of Laboratories</i> . 50, 51
KDE	Estimação de Densidade por <i>Kernel</i> . 92, 141, 142, 165, 195
KNN	k-Vizinhos Mais Próximos. 54
KS	Kolmogorov-Smirnov. 142
L1	Primeiro Nível de Filtragem. xvi, xxxvii, xxxviii, 2, 79–81, 83, 101, 114, 117

<i>L1Calo</i>	Primeiro Nível de Filtragem para o Sistema de Calorimetria. 79, 80, 100–103, 117, 119, 122, 124–126, 132, 162
<i>L1Muon</i>	Primeiro Nível de Filtragem para o Espectrômetro de Múons. 80
L2	Segundo Nível de Filtragem. 112
LCG	<i>LHC Computer Grid Project</i> . vii, 6, 61, 79, 117, 149, 155, 183
LDA	Análise de Discriminante Linear. 54
LHC	<i>Large Hadron Collider</i> . xxxi, xxxv, 2–4, 6, 8, 13–16, 21–24, 28, 30, 35, 36, 43, 47, 48, 59–61, 64, 69, 83, 95, 102, 105, 113, 119, 123, 124, 130, 198, 201
LHCb	<i>Large Hadron Collider beauty experiment</i> . 14, 15
LIDAR	Detecção de Luz e Alcance. 54
LINAC	Acelerador Linear. 14, 15
LPS	Laboratório de Processamento de Sinais - COPPE/UFRJ/UFRJ. vii, 183
MAE	Erro Absoluto Médio. 55
MAPE	Erro Absoluto Percentual Médio. 55
MBTS	Cintiladores de Filtragem para <i>Minbias</i> . 80, 83
MC	Monte Carlo. xvii, xviii, 77, 98, 132, 137, 138, 148, 149, 160, 176
<i>Minbias</i>	Eventos coletados com Filtragem Minimamente Restritiva. xxxi, 80
MLP	Rede Multicamadas de <i>Perceptrons</i> . xix, 5, 48, 49, 55, 103, 114, 115, 117, 122, 125–129, 133, 145, 146, 150, 155, 157, 159, 186, 199, 255
MP	Modelo Padrão de interação entre as partículas elementares. 8, 9, 11, 12, 82
MSE	Erro Quadrático Médio. xix, 55, 118, 126, 133, 157, 171, 177, 185, 195
NACAD	Núcleo Avançado de Computação da COPPE/UFRJ. 6

NARX	Rede Não-Linear Autorregressiva com Entradas Exógenas. 54
ND	Colisão Não Difrativa. 18
NLICA	Análise de Componentes Independentes Não-Linear. 121
OS	Carga Elétrica Oposta. 78, 98
P1	Ponto 1 de Inserção. 14, 76
PAW	<i>Physics Analysis Workstation Package</i> . 76
PCA	Análise de Componentes Principais. xxii, 118, 120, 121, 176, 179, 180, 183
PCD	Componentes Principais de Discriminação. 120–122, 203
<i>pdf</i>	Função de densidade de probabilidade. 21, 92, 141
PMT	Multiplicador de Fótons. 72
PNL	<i>Post Nonlinear model</i> . 121
PNN	Rede Neural Probabilística. 55
PS	Calorímetro Pré-Amostrador. 70, 71, 73, 85, 110, 112, 123, 125, 126, 185
<i>PSync</i>	<i>Proton Synchrotron</i> . 14, 15
QCD	Cromodinâmica Quântica. 11, 17, 20
QED	Eletrodinâmica Quântica. 11
RBM	<i>Restricted Boltzmann Machine</i> . 56, 57
RF	<i>Radio Frequency</i> . 15, 16, 55
RMS	Erro Médio Quadrático. xxi, 32, 133, 169, 170
RNA	Rede Neural Artificial. 53–55
ROC	<i>Receiver Operation Curve</i> . 133, 135, 149, 157
RoI	Região de Interesse. 80, 96, 100, 101, 113, 114, 132, 161, 188
<i>ROOT</i>	<i>Framework</i> de Análise ROOT. 76, 77, 252
ROS	Sistemas de Leitura. 80
RPROP	<i>Resilient Back-propagation</i> . 7, 117, 120, 133–135, 155, 157, 255, 256
RSSI	Indicador de Força de Sinal de Rádio. 55
RTMS	Sensor Remoto de Tráfego por Microonda. 55

SAE	<i>Stacked Auto-encoder</i> . xix, 5, 184–186, 203
SCT	Detector de Rastreamento por Semicondutores. 62, 63, 85–87
SF _{on}	Sistema de Filtragem. xvii, xx, xxxiii, xxxvii, 24, 61, 76, 77, 79, 80, 82, 84, 91, 96, 102, 106, 121, 123, 124, 126–128, 132–140, 144, 145, 149, 150, 153, 159, 160, 172, 175, 176, 188, 194, 196, 197, 202, <i>veja no glossário</i> Sistema de Filtragem do ATLAS
SI	Sistema Internacional de unidades e medidas. 9
SOM	Mapas Auto-Organizáveis. xxii, 54, 122, 176, 177, 179, 180, 202, 203
SPS	<i>Super Proton Synchrotron</i> . 14, 47
SR _{off}	Sistema de Reconstrução (<i>Offline</i>) de Física. xviii, xix, xxi, xxxiii, xxxvi, 61, 76, 91, 107, 118, 124, 128, 140, 141, 143, 145, 147, 148, 151–153, 157, 158, 171–173, 176, 188, 194–197, 253, <i>veja no glossário</i> Sistema de Reconstrução (<i>Offline</i>) de Física do ATLAS
SSC	<i>Superconducting Super Collider</i> . 113
SUSY	Teoria de Supersimetria. 82
SVM	Máquina de Vetores de Suporte. 54
<i>T2Calo</i>	Segundo Nível de Filtragem empregando informação do Sistema de Calorimetria. 112, 120–124, 128
T&P	Método <i>Tag and Probe</i> . xvi–xix, xxi, xxii, xxxv, 78, 93, 96, 97, 108, 123, 124, 126, 132, 137–143, 147–149, 154, 161, 164, 165, 170, 171, 173, 179, 181, 182, 202
TDAQ	Sistema de Qualidade e Filtragem de Dados. 79, 81, 98, 112, 117
<i>TileBar</i>	Calorímetro Hadrônico de telhas cintilantes (<i>TileCal</i>) cobrindo a região central do barril ATLAS. 66, 112

<i>TileCal</i>	Calorímetro Hadrônico de Telhas. xvi, xxxiii, xxxiv, 43, 47, 65, 66, 71, 72, 74–76, 80, 93, 111, 134, 135
<i>TileExt</i>	Extensão do Calorímetro Hadrônico de telhas cintilantes (<i>TileCal</i>) cobrindo a região do barril ATLAS. 66, 112
<i>TileGap</i>	Termo empregado para a divisão em amostras do ITC. 112, 197
ToT	Tempo-sobre-Limiar. 63
TRT	Detector de Rastreamento por Transição de Radiação. 62–64, 86, 87, 90, 96
TS1	Primeira Parada Técnica. <i>Technical Stop 1</i> . xviii, 131, 161, 162
TT	Torre de Filtragem. 100, 101
UFRJ	Universidade Federal do Rio de Janeiro. xxxi, 59, 250, 251
WSN	Rede de Sensores <i>Wireless</i> . 52, 53
xAOD	AOD Multi-Ambiente. 77, 78, 153, 154
ZDC	Calorímetro Contador de Zero Grau. 80, 83

Glossário

ajuste por redes neurais especialistas (ExNN) Estratégia para a fusão de informação com o ajuste de redes neurais conforme descrito na Subseção 7.3.1. 130

dados de colisão Dados experimentais obtidos a partir de colisões.

run Um único ciclo de preenchimento do LHC. xxix, 82, 105, 107, 123, 128, 133, 136–138, 160, 248

Primeira Temporada de Coleta de Dados (*Run 1*) Ciclo de preenchimentos do LHC compreendidos entre 2009 e 2012. 13, 21, 23, 46, 47, 64, 80, 110, 112, 119, 123, 124, 130, 187

Segunda Temporada de Coleta de Dados (*Run 2*) Ciclo de preenchimentos do LHC iniciados em 2015 e com previsão para se encerrarem no final de 2018. 13, 21, 23, 63, 77, 80–82, 84, 86, 100–102, 110, 201

Infraestrutura *Athena* Principal infraestrutura em *software* do ATLAS que possibilita a execução do código para os diversos fins necessários do experimento, como reconstrução física, simulação, filtragem (HLT), leitura e escrita de dados, análise etc. xviii, 77, 110, 111, 119, 145, 152

Método *Tag and Probe* Método capaz de fornecer uma pré-seleção *loose* de um determinado objeto físico a partir de conhecimento especialista em física. Esse método tem grande importância na física experimental por fazer parte de dois pontos centrais: fornecer os parâmetros necessários para as análises físicas baseadas em comparações dos resultados experimentais com os valores esperados obtidos por via de simulação; e possibilitar obter amostras de dados de colisão para o ajuste de modelos com menor dependência em modelos anteriores.

tag Objeto físico de referência, geralmente exigindo todo o conhecimento disponível a fim de garantir que o mesmo tenha sido selecionado com boa qualidade. Em determinados casos, o método T&P pode conter mais de um *tag*. xxxv, 97, 98, 161, 165

probe Objeto físico de prova que possui, junto com o *tag*, propriedades compatíveis com algum processo físico. xviii, xix, 89, 90, 97, 98, 137–141, 161, 163–165, 174

reconstrução de física Processo de tratamento da informação oriunda dos canais do experimento para obter uma estimativa de processos físicos ocorridos durante a colisão e da sequência de seu desenvolvimento para a geração dos mesmos.

objeto físico Uma partícula ou grupo de partículas observável pelo experimento. No caso do ATLAS os principais objetos disponíveis para a análise são: elétrons, fótons, táons, múons, jatos hadrônicos, energia transversa faltante (neutrinos). xxxv, xxxvi, 3, 11, 76, 81–84, 92, 98, 106, 138, 145, 152, 176, 252

representação da informação Tratamento da alta dimensionalidade através da fusão da informação para a representação em objetos capazes de sintetizar o processo de interação de objetos físicos com o material do experimento.

cluster Agrupamento de células compondo a deposição de energia de um objeto físico. xxxvi

traço Trajetória estimada de um objeto físico. xxxvi

grandezas físicas Construção de variáveis através de conhecimento especialista em Física de Altas Energias com grande poder de síntese do processo de interação com a matéria do experimento. As grandezas físicas de elétrons do ATLAS estão descritas na Seção 5.2. 200

grandezas físicas de calorimetria (*shower*) Grupo de grandezas físicas construídas unicamente a partir da informação de calorimetria. 146, 200

grandezas físicas de traço e de fusão especialista calorimetria-traço (*track⁺*) Grupo de grandezas físicas construídas a partir de informação de calorimetria e de traço, exceto aquelas oriundas unicamente a partir de informação de calorimetria. Conta com grandezas de fusão especialista entre o e o traço baseada em posição. xix, 145, 160, 166, 169, 170, 174, 200

soma em anel concêntrico de energia (*R*) Representação compactada do desenvolvimento do chuveiro via emprego de conhecimento especialista por soma de energia em anéis concêntricos. xxxvi, 5, 6, 110, 122, 145, 200

Algoritmo Anelador (*Ringer*) Método para a extração da soma em anéis concêntricos de energia. Ver Seção 6.1 para descrição do algoritmo e seções 9.5, 9.6, 9.6 e 9.8 para possíveis aprimoramentos propostos neste trabalho. 57, 110, 111, 113, 154, 161

Sistema de Reconstrução (*Offline*) de Física do ATLAS (SR_{off}) Infraestrutura em *software* responsável pela reconstrução física em *offline*. xviii, xix, xxi, 61, 76, 91, 107, 118, 124, 128, 140, 141, 143, 145, 147, 148, 151–153, 157, 158, 171–173, 176, 188, 194–197, 253

reconstrução de elétrons Processo de obtenção de candidatos de elétrons a partir de *clusters* e traços com a aplicação de critérios de seleção para o emprego nas análises físicas.

identificação de elétrons Tarefa de reconhecimento de padrões a partir

da informação discriminante de forma a disponibilizar critérios de seleção de elétrons.

critério de seleção Critérios disponibilizando diferentes qualidades de seleção de forma a possibilitar atender às necessidades diversas das análises físicas. xxxvi, xxxvii

qualidade *vloose* Critério de seleção minimamente restritivo. xvii–xix, 98, 108, 137, 138, 141, 148, 151, 161, 164–166, 171–173, 200

qualidade *loose* Critério de seleção de priorizando grande volume de estatística. xxxv, 84, 123

qualidade *medium* Critério de seleção de buscando performance equilibrada entre detecção de sinal (P_D) e contaminação por ruído físico (P_F). 78, 96, 98, 123, 137, 161

qualidade *tight* Critério de seleção priorizando pureza de estatística. xviii, 78, 96, 98, 108, 123, 137, 139–141, 161, 163, 165

métodos de seleção

Algoritmo Baseado em Cortes Lineares (*CutBased*) Estratégia tradicionalmente empregada para identificação de partículas em Física de Altas Energias com a aplicação de cortes sequenciais em grandezas físicas discriminantes. xvii, xviii, 78, 92, 103–107, 130–132, 134, 135, 149, 157, 162

Algoritmo de Verossimilhança (*LLH*) Estratégia multivariável com o emprego de verossimilhança a partir de grandezas físicas discriminantes adotada pelo ATLAS como solução para lidar com o aumento do efeito de empilhamento. A versão atual assume independência entre as grandezas físicas para obtenção do discriminante (Naïve Bayes). É a referência para a seleção de elétrons por ser o método para a seleção de elétrons nas análises físicas. xix–xxi, 78, 93, 95, 102, 103, 105, 106, 131–133, 141, 145, 147, 149–151, 157, 160, 165, 169, 170, 172, 174

Algoritmo *NeuralRinger* Estratégia baseada no emprego de redes neurais aonde a informação anelada faz parte das representações compondo o espaço de entrada. Atualmente emprega um *ensemble* de redes neurais específicas por regiões de posição de incidência ($|\eta|$) e energia (E_T) da partícula. v, ix, xvii, xviii, xxi, 4–7, 48, 84, 94, 103, 104, 107, 109, 122–128, 130–132, 135, 136, 144–146, 148, 151, 152, 154, 157, 159–162, 172, 198–204, 253

Sistema de Filtragem do ATLAS (SF_{on}) Implementa a reconstrução física *online* a fim de manter os eventos com informação relevante para os experimentos enquanto reduzindo a taxa de eventos para níveis viáveis de armazenamento e processamento *offline*. O sistema possui dois níveis de filtragem: Primeiro Nível

de Filtragem (L1) e Alto Nível de *Trigger* (HLT). xvii, xx, 2, 24, 61, 76, 77, 79, 80, 82, 84, 91, 96, 102, 106, 121, 123, 124, 126–128, 132–140, 144, 145, 149, 150, 153, 159, 160, 172, 175, 176, 188, 194, 196, 197, 202

Capítulo 1

Introdução

A proliferação de aplicações em grande massas de dados advém como resultado do avanço tecnológico durante a Era Digital. Em determinados casos, a massa de dados encobre a presença de eventos singulares que são o objetivo final das análises das aplicações. O desbalanço entre a produção de eventos de interesse e ordinários pode tornar-se um fator limitante à aplicação, se não for introduzida a atuação de um sistema de filtragem. Eventualmente, o descarte do ruído de fundo precise ser feito de forma *online*, o que adiciona requisitos de velocidade e uso de memória ao problema. Dada a própria condição restrigente, esse sistema deve operar com alta eficiência de seleção que, por sua vez, está cerceada pelas condições de recursos para a tomada de decisão *online*. Uma vez que os eventos sejam raros, deseja-se, sempre que possível, a maximização de sua taxa de produção, o que pode provocar uma alta taxa de eventos gerados na aplicação. Nesse cenário, a latência para a realização da tomada de decisão geralmente se apresenta como um fator limitante ao emprego das técnicas mais sofisticadas. A fim de possibilitar alcançar menor latência com a mesma exigência de recursos, o sistema de filtragem pode ser projetado em níveis sequencialmente mais restritivos. O escrutínio final dos eventos remanescentes ocorre em *offline*, onde se emprega plenamente a capacidade de processamento. Assim, a versão *online* de filtragem pode resultar de uma adaptação da técnica de processamento *offline* aos recursos disponíveis no ambiente *online*. Ocorre, naturalmente, uma dinâmica entre esses sistemas, uma vez que a eficiência total resulta da harmonia de ambas decisões.

Uma técnica usada para a seleção desses eventos emprega um *ensemble* de modelos. O *ensemble* pode atuar em alta dimensionalidade, possivelmente combinando a informação de natureza distinta, o que exige uma técnica para sua fusão, a fim de atingir patamares mais elevados de eficiência.

Esses aspectos estão presentes regularmente, seja isoladamente ou simultaneamente, em problemas de classificação de eventos, área de aplicação de técnicas de Aprendizado de Máquinas, Inteligência Computacional e Ciência de Dados. A proli-

feração dessas técnicas tem atingindo os mais diversos campos, exemplos envolvendo a busca de eventos de interesse são: triagem de pacientes com risco de determinadas patologias; identificação de eventos de transitório em redes elétricas; identificação de falha ou predição de anomalia em equipamentos ou na rede elétrica de uma indústria; *Internet* das Coisas (IoT); identificação de partículas em experimentos de Físicas de Altas Energias. Como no caso deste trabalho, as técnicas e análises empregadas podem ser aliadas ao conhecimento especialista do contexto da aplicação, para possibilitar uma abordagem ainda mais poderosa ao problema.

Especificamente para a área de Física Experimental de Altas Energias, os aspectos citados estão envolvidos em experimentos de grande porte, que avançam o entendimento da compreensão dos constituintes básicos do Universo e de como são regidas suas interações. A observação de eventos singulares que possibilitam obter tal compreensão é extremamente rara, sendo necessária a construção de grandes aceleradores de partículas. Atualmente, a máquina na vanguarda para a geração dessas condições é o *Large Hadron Collider* (LHC) no CERN, Genebra, Suíça. Um exemplo bem sucedido de observação deu-se na descoberta do bóson de Higgs no LHC em 2012, mecanismo previsto há cerca de meio século anteriormente, e que resultou no posterior laureamento de Peter Higgs e François Englert com o prêmio Nobel de Física. O maior dos experimentos do LHC, *A Toroidal LHC Apparatus* (ATLAS), teve um papel importante para essa descoberta.

Processos como o bóson de Higgs são extremamente raros e, para permitir o acúmulo de estatística suficiente para a caracterização de suas assinaturas em meio a uma grande massa de ruído, o LHC concentra prótons nos feixes em pacotes que são direcionados à colisão de altas energias (13 TeV no centro de massa) a cada 25 ns. No caso do ATLAS, o armazenamento por completo dos eventos na alta taxa equivalente (40 MHz em pico) exigiria uma banda de ~ 70 TB/s, enquanto que o período de coleta de dados pode se expandir por anos ou até décadas e a esmagadora parte desses eventos compõe física ordinária. Por isso, há a necessidade de atuação de um sistema de filtragem, responsável pela remoção *online* dos eventos que não agregam informação para as análises de interesse do experimento, buscando atender uma taxa alvo (1 kHz) viável de armazenamento e posterior processamento *offline*. Esse, por sua vez, realiza o processamento de grande massa de dados (atualmente superior à 5 PB/ano) envolvendo todos os sensores do ATLAS (~ 100 milhões de canais) para identificação final dos canais físicos de interesse.

A fim de atingir a baixa latência necessária para operação na taxa de eventos proveniente pelo LHC, o *trigger* (sistema de filtragem *online*) do ATLAS aplica dois níveis de processamento, em uma cadeia de seleção sequencial, tanto em profundidade de análise física, quanto em requerimento de processamento. O Primeiro Nível de Filtragem (L1), implementado em *hardware*, inicia o processo seleção e opera com

latência (tempo médio de decisão) inferior a $2,5 \mu\text{s}$, enquanto que reduz o fluxo de eventos para, no máximo, 100 kHz. O segundo nível, referido como Alto Nível de *Trigger* (HLT), é baseado em *software*, com alvo médio de latência de até 550 ms e uma média (pico) de taxa de saída de 1 kHz (1,5 kHz). No HLT, os métodos são implementados em *software* através de uma degradação do modelo *offline*, a referência empregada para a confirmação das análises físicas, de modo a satisfazer as restrições de processamento *online*. Essa estratégia permite harmonia entre as decisões dos sistemas, uma vez que ambos possuem natureza similar. O emprego de técnica insólita na filtragem *online* para a tomada de decisão exige a avaliação de seu impacto nas amostras conforme observadas pelo processamento *offline*, referência do sistema.

O processo de análise em ambos sistemas ocorre através da avaliação do sinal desenvolvido pelo resíduo das colisões empregando técnicas de reconhecimento de padrões. Elétrons são parte desses objetos físicos de interesse e participaram da descoberta do bóson de Higgs. Igualmente, elétrons fazem parte das análises em busca de outros processos ainda mais raros, que é um dos objetivos adicionais do experimento. No ATLAS, suas assinaturas provêm de detectores de duas naturezas: os calorímetros (energia) e os detectores mais internos, baseados em espectrografia magnética (trajetória da partícula, medição de momento e outros). Além disso, os padrões estão sujeitos a distorções em função da posição de sua incidência no ATLAS, da energia da partícula originando o sinal e por empilhamento de sinais oriundos de diferentes colisões.

1.1 Motivação

Com o objetivo de obter estatística suficiente para os processos ainda mais raros que o bóson de Higgs, o LHC vem elevando a taxa de colisões próton-próton. A maneira em que isso é alcançado se dá pelo estreitamento da colimação do feixe (luminosidade), o que fornece um aumento no número de colisões por cruzamento de pacotes. Apesar de extremamente positiva para alcançar, em um menor período de tempo, a estatística necessária para os estudos, essa estratégia leva a uma maior ocupação dos detectores, produzindo o empilhamento de sinais de diferentes colisões. O crescimento da taxa de colisões fica evidente ao comparar-se os períodos de dados empregados pelo ATLAS para a descoberta do bóson de Higgs, onde o LHC operou com espaçamento dos feixes de 50 ns e valores médios do número de colisões deram-se por até 20,7, com os valores atuais de 2017, contando com o espaçamento nominal de 25 ns e alcançando valores de até 60 colisões médias. Para 2018, espera-se atingir valores ainda superiores (88 colisões médias), como parte de uma atividade de contínuas melhorias. Dentre essas atividades, haverá duas grandes fases de atua-

lizações, estendendo a operação do LHC para além de 2030, a fim de atingir uma média de 200 colisões médias inelásticas.

Assim, urge a demanda por métodos eficientes de seleção, seja para o *trigger* quanto para atuação *offline*. Uma vez que as decisões finais do experimento são avaliadas empregando as técnicas mais avançadas disponíveis em *offline*, espera-se que as decisões avaliadas pelos métodos do *trigger* estejam alinhadas ao máximo com as suas referências atuantes no *offline*. Por isso, o caminho natural para o desenvolvimento, implementação e análise dá-se, primeiro, para atuação no *offline* e, depois, são estendidas para o *trigger* com as modificações necessárias (degradações) para atender às demandas de latência e memória.

Em particular, o ATLAS realça o emprego de elétrons na busca de física de interesse, com forte utilização de seus calorímetros. Duas estratégias de seleção são as referências no ATLAS para identificação de elétrons em análise *offline*. Ambas abordagens são baseadas na construção de grandezas físicas altamente descritivas, obtidas através de um conhecimento especialista agregado por mais de 50 anos de atuação na área. Originalmente, os experimentos empregavam cortes sequenciais nessas grandezas para obter a decisão final. Com o aumento do empilhamento, o ATLAS alterou o método de classificação *offline*, em 2012, para uma abordagem multivariável baseada em verossimilhança, com o objetivo de recuperar eventos em que algumas dessas grandezas obtivessem valores inferiores ao limiar. Essa estratégia foi portada para estágios finais de decisão do HLT em 2015 quando se avaliou que o emprego desta técnica não causaria a perda dos elétrons já coletados e ainda possibilitaria explorar melhor o potencial da técnica atuante em *offline*.

No *trigger*, a pressão causada pelo empilhamento afeta toda a estrutura, onde há limitação de banda para o processamento de um maior número de eventos durante a evolução da cadeia. Isso torna interessante priorizar a introdução de algoritmos mais eficientes nos estágios primordiais da cadeia quando a limitação por processamento é atingida, uma vez que há eliminação de falsos candidatos sem a necessidade de processamento de algoritmos mais complexos. No caso da filtragem de elétrons, os estágios primordiais limitam-se somente ao processamento de informação de calorimetria, por ser discriminante e de rápido processamento.

O Instituto Alberto Luiz Coimbra de Pós-graduação e Pesquisa de Engenharia (COPPE/UFRJ) propôs, no início da década de 1990, uma estratégia para filtragem de elétrons no ATLAS baseada somente em informação de calorimetria (Algoritmo *NeuralRinger*). Realiza-se a compactação da informação via conhecimento especialista, mantendo-se a interpretação do processo de interação da partícula com o detector. Assim, regiões sensoras são agrupadas em anéis concêntricos de deposição de energia de forma a explorar a geometria aproximadamente cônica da interação da partícula com o material do calorímetro. Durante o desenvolvimento dessa es-

tratégia, notou-se que redes neurais Rede Multicamadas de *Perceptrons* (MLP) se apresentavam como os melhores modelos para atuarem no processo de filtragem. A técnica possibilita uma estratégia mais eficiente para a redução de processamento e, assim, apresenta-se como uma solução para aliviar o acréscimo na demanda por conta do empilhamento. Todavia, o grande desafio para o emprego dessa estratégia dá-se pela resistência para sua adoção, devido ao temor que o treinamento do classificador introduziria tendências nos resultados do experimento e pelo fato do processamento *online* não ser proveniente da análise *offline*. A ausência de estudos do Algoritmo *NeuralRinger* atuando no ambiente *offline* e da resposta quanto a sua influência nessas análises é um obstáculo para a adoção desse método, mesmo havendo a demanda por técnicas mais eficientes para lidar com o efeito de empilhamento.

1.2 Objetivo

Este trabalho tem como objetivo desenvolver o *NeuralRinger* para a seleção *offline* de elétrons no Experimento ATLAS. Deseja-se avaliar o impacto de sua contraparte *online* na referência atualmente empregada para a análise *offline*, baseada em verossimilhança, a fim de possibilitar o seu emprego na filtragem *online*.

1.3 Contribuições do Trabalho

A proposta de desenvolvimento do Algoritmo *NeuralRinger* para atuação em *offline* a partir de sua contrapartida no *online* inverte a lógica empregada nos experimentos. Para isso, estendeu-se a proposta *online* através da fusão da soma em anéis concêntricos de energia com as grandezas físicas por ajuste de redes neurais especialistas para cada representação. Essa estratégia mostrou o potencial da técnica em sua avaliação em dados de simulação, que possibilitou a redução da taxa de falso alarme pela metade em relação à referência (verossimilhança) para a seleção a ser empregada nas análises físicas envolvendo a busca por processos extremamente raros. O trabalho abordou outras estratégias para aprimoramento do método. Dentre elas, possíveis ajustes no algoritmo de extração de soma em anéis concêntricos de energia e a correção de suas limitações, extração de espaços latentes por *Stacked Auto-encoder* (SAE) e Análise de Componentes Independentes (ICA), adição de outras informações potencialmente discriminantes, proposta de uma nova descrição de informação de calorimetria.

Desenvolveu-se um método estatístico para avaliar o impacto da atuação do *NeuralRinger* na filtragem *online* na referência em atuação no *offline*. Essa análise possibilitou a adoção de sua contraparte *online* pelo ATLAS que se tornou a primeira aplicação baseada em redes neurais em experimentos de Física de Altas Energias

atuando como método principal para a tomada de dados. Este trabalho fornece, assim, uma base para o desenvolvimento quando desenvolvendo estratégias para filtragem *online* e posteriormente serem estendidas para atuação *offline*.

O desenvolvimento do *NeuralRinger* para a seleção *offline* de elétrons envolveu a implementação de uma infraestrutura completa para lidar com grande massas de dados no complexo ambiente de *software* do ATLAS. Essa infraestrutura disponibilizou o processamento de 11 trilhões de eventos de cruzamento pacotes (10, 5 PB de dados), correspondentes à coleta de dados ocorrida entre 2015–2017, e de conjuntos de simulações com elétrons e ruído físico em condições de observação equivalentes para a extração de soma em anéis concêntricos de energia. Esses processamentos estão disponíveis para serem empregadas por todo o experimento para elétrons com $E_T > 14$ GeV. A disponibilidade desses dados é essencial para garantir a continuidade dos desenvolvimentos do método e necessários nas análises para que a técnica atenda os requisitos necessários para a sua aplicação na análise física.

Dado a necessidade de ajuste de modelos em grande massas de dados, realizou-se o desenvolvimento em equipe de uma infraestrutura específica para o ajuste dos modelos. Tal infraestrutura foi desenvolvida considerando as necessidades conjuntas do Algoritmo *NeuralRinger offline* e *online*, o que possibilitou o desenvolvimento do método para o *offline* usufruir, e também auxiliar, de todos os desenvolvimentos realizados para o *online*. A infraestrutura encontra-se integrada com duas grandes fontes de recursos computacionais: a *LHC Computer Grid Project* (LCG), maior grade computacional do mundo disponibilizada para os experimentos do LHC; e o supercomputador Lobo Carneiro, recurso do Núcleo Avançado de Computação da COPPE/UFRJ (NACAD), cuja proposta de liberação dos recursos surgiu como parte da necessidade deste trabalho.

1.4 Organização do Documento

O trabalho encontra-se segmentado em cinco níveis com sucessivamente menor escopo de abrangência: Capítulo, Seção, Subseção, Tópico e Subtópico. Sempre que possível, o leitor será guiado para outro segmento aonde o assunto foi abordado dentro deste trabalho, de forma a facilitar uma leitura não-linear. As figuras e tabelas contêm legendas com um alto grau de detalhe a fim de possibilitar a sua compreensão sem a leitura do texto do trabalho, dado que o leitor contenha conhecimento da área. A exceção são os resultados divulgados no Capítulo 8, por sua vez, que assumem a leitura dos métodos no Capítulo 7. Uma tentativa de leitura não linear pode se dar através da leitura inicialmente do Capítulo 7 e então acompanhar as referências às outras partes do trabalho.

O leitor sem conhecimento na área de Física de Altas Energias encontrará um

levantamento dos diversos aspectos envolvidos no Capítulo 2. Todavia, esse capítulo se estende além do mínimo necessário para a compreensão do trabalho por também ter a ambição de agregar a informação para facilitar o desenvolvimento de estudantes no projeto devido à complexidade do ambiente. Essa informação, apesar de disponível, encontra-se diluída, de forma que os estudantes têm dificuldade para agregar as informações e gerar conhecimento.

O levantamento bibliográfico da aplicação de Aprendizado de Máquinas em Física de Altas Energias e das técnicas empregadas no campo de Fusão de Informação encontram-se no Capítulo 3.

A descrição do ATLAS realizada no Capítulo 4 é igualmente detalhada para catalogar as suas nuances de operação, aspecto extremamente importante para o aprimoramento dos modelos quando considerando a extensão para análise *offline*.

A reconstrução e identificação de elétrons encontra-se no Capítulo 5, onde se apresenta a atual referência para as análises *offline* e a estrutura da cadeia de elétrons. O aspecto no cerne deste trabalho é discutido nesse capítulo, na Subseção 5.7.2.

Descreve-se a informação altamente discriminante de calorimetria, a partir de descrição por soma de energia em anéis, na Seção 6.1, do Capítulo 6. Uma série de outros autores da COPPE/UFRJ abordaram o tema, sendo realizado um levantamento do conhecimento adquirido nos trabalhos anteriores.

O Capítulo 7 aborda a técnica desenvolvida para análise de impacto e o método aplicado para a proposta do *NeuralRinger* para atuação nas análises *offline* e apresenta os resultados da filtragem *online* realizada pelo *NeuralRinger*. Os resultados para os três temas abordados são apresentados no Capítulo 8.

Abordam-se outros aperfeiçoamentos para a atuação do *NeuralRinger* no Capítulo 9, onde se propõe o emprego de uma nova representação da informação a fim de explorar assimetria de desenvolvimento do chuveiro.

As conclusões são realizadas no Capítulo 10, que é prosseguido por três apêndices: a produção acadêmica (Apêndice A), os detalhes da infraestrutura implementada para possibilitar a atuação da proposta do *NeuralRinger* no ambiente *offline* (Apêndice B) e o pseudo-código do algoritmo RPROP (Apêndice C).

Capítulo 2

Física Experimental de Altas Energias

Motivada pelas proporções envolvidas em seus experimentos, a Física Experimental de Altas Energias é também conhecida por “Grande Ciência” [1]. A colaboração científica envolvida em um único experimento contém centenas ou milhares de cientistas de diversas áreas e abrange uma larga gama de nacionalidades. Os complexos experimentais alastram-se por quilômetros de distância, utilizando grandes volumes de gelo Antártico ou de lagos como um dos meios para estudar raios cósmicos, ou grandes aceleradores de partículas — como o *Large Hadron Collider* (LHC) que contém 27 km de circunferência, dentre muitas outras propriedades que o fazem único. Os experimentos são construídos em décadas e toma-se mais outra para coletar e analisar seus resultados. Os objetivos dessas análises não são menores: obter informações sobre os constituintes básicos do Universo e como se regem suas interações. Geralmente esses experimentos são realizados para possibilitar a observação (ou violação) de mecanismos ou processos previstos pelo campo teórico da Física de Partículas Elementares.

Neste Capítulo se efetua uma breve introdução dos aspectos teóricos da Física de Partículas, bem descrita pelo Modelo Padrão de interação entre as partículas elementares (Seção 2.1), necessários para a compreensão do trabalho. O Experimento *A Toroidal LHC Apparatus* (ATLAS) (Capítulo 4), do qual este trabalho faz parte da colaboração, coleta dados de colisão de partículas aceleradas pelo LHC (Seção 2.3). O mesmo está situado no complexo de aceleradores de partículas do *Centre Européene pour la Recherche Nucleaire* (CERN) (Seção 2.2), em Genebra, expandindo-se pela fronteira franco-suíça. Também se efetiva uma breve abordagem sobre os detectores construídos em Física Experimental de Altas Energias (Seção 2.4), onde se enfoca o principal tipo de detector envolvido neste trabalho — o calorímetro (Subseção 2.4.2).

2.1 Física de Partículas Elementares e o Modelo Padrão

O objetivo do estudo da Física de Partículas é a compreensão dos constituintes básicos do Universo. A noção de constituintes elementares não é recente — Demócrito, por volta de 430 a.C., já proferia esse conceito [2]. Na ciência moderna, o seu estudo teve início apenas no século 19, quando o elétron foi descoberto por Thompson [3]. O resultado de quase um século de desenvolvimento experimental e teórico do tema resultou no Modelo Padrão de interação entre as partículas elementares (MP) [4–7] — uma infraestrutura matemática que descreve partículas elementares e três forças que governam suas interações em termos de simetrias.

A base de sua formulação ocorreu na década de 1960 para explicar a fatura de partículas sendo descobertas no período e mostrou-se bem sucedida para a descrição dos fenômenos de interação com alto grau de precisão. Indo além, o MP forneceu restrições para propriedades de novas partículas que somente após décadas seriam descobertas, como o caso da partícula elementar *top* [8, 9] e do bóson de Higgs [10, 11].

As unidades normalmente utilizadas no estudo de Física de Partículas são fm para distância (equivalente a 10^{-15} m), MeV, GeV ou TeV para massa ou energia, onde 1 eV (elétron-volt) é a energia necessária para aumentar o potencial elétrico de um elétron em um volt, equivalente a $1,6 \times 10^{-19}$ J no Sistema Internacional de unidades e medidas (SI), ou em unidades de massa $1 \text{ eV}/c^2 = 1,78 \times 10^{-36}$ kg e momento $1 \text{ eV}/c = 5,3410^{-28}$ kg m/s. Em determinados casos, a normalização por c pode estar omitida para as unidades de massa e momento. A unidade para área é o *barn*, definida por $1 \text{ b} = 10^{-28} \text{ m}^2$, utilizada em termos de mb ou fb [4].

O MP contém doze férmions, partículas subatômicas. É possível separar os férmions em dois grupos: um com carga elétrica inteira em unidades da carga do elétron (léptons) e outro com carga fracionária (quarks). Além disso, existem três diferentes gerações, ou famílias, de férmions, cada com léptons e quarks com propriedades similares, porém com maior massa. Os férmions (f) também possuem uma antipartícula correspondente, geralmente representados por \bar{f} , contendo carga oposta e outros números quânticos. Em outros casos, a antipartícula pode simplesmente ser representada com o sinal inverso, como para o elétron (e^-) e sua antipartícula (e^+). Por motivos históricos, apenas a antipartícula do elétron tem nomenclatura própria [4, 5] — pósitron —, que as vezes é omitida. Quando um par de partícula e antipartícula colidem acabam se aniquilando e liberando sua energia em outras partículas. Partículas eletricamente neutras também possuem antipartículas e, em alguns casos, podem ser suas próprias antipartículas.

2.1.1 Léptons

Os léptons (Tabela 2.1) podem ser subdivididos em dois grupos, um com massa e carga elétrica, idêntica e unitária: elétron (e), múon (μ) e táon (τ); e outro neutro em carga e massa reduzida, estando relacionados com os léptons carregados: neutrino do elétron (ν_e), neutrino do múon (ν_μ) e neutrino do táon (ν_τ). Dos léptons carregados, o múon e o táon só diferem dos elétrons na sua massa e no seu tempo de vida finitos, onde o elétron é o único estável.

Tabela 2.1: Os léptons, suas gerações, carga elétrica em unidades de carga absoluta do elétron e massa [12]

Geração	Nome	Símbolo	Carga Elétrica (e)	Massa (MeV)
Primeira	elétron	e	-1	$0,5109989461 \pm 0,0000000031$
	neutrino elétron	ν_e	0	$< 2 \times 10^{-6}$
Segunda	múon	μ	-1	$105,6583745 \pm 0,0000024$
	neutrino do múon	ν_μ	0	$< 2 \times 10^{-6}$
Terceira	táon	τ	-1	$1776,86 \pm 0,12$
	neutrino do táon	ν_τ	0	$< 2 \times 10^{-6}$

2.1.2 Quarks e hádrons

As três gerações de quarks (Tabela 2.2) são formadas por pares de sabores — que descrevem a massa e carga elétrica — com carga de $+2/3$ e $-1/3$, conforme segue: *up* (u) e *down* (d); *strange* (s) e *charm* (c); *top* (t) e *bottom* (b). Além da carga elétrica, eles possuem, também, outra forma de carga (referente a força forte) fantasiosamente representada por cores: vermelho, verde e azul.

Tabela 2.2: Os quarks, suas gerações, carga elétrica em unidades de carga absoluta do elétron e massa.

Geração	Nome	Símbolo	Carga Elétrica (e)	Massa
Primeira	<i>up</i>	u	$+2/3$	$2,2_{-0,4}^{+0,6}$ MeV
	<i>down</i>	d	$-1/3$	$4,7_{-0,4}^{+0,5}$ MeV
Segunda	<i>strange</i>	s	$+2/3$	96_{-4}^{+8} MeV
	<i>charm</i>	c	$-1/3$	$1,27 \pm 0,03$ GeV
Terceira	<i>top</i>	t	$+2/3$	$173,21 \pm 0,51 \pm 0,71$ GeV
	<i>bottom</i>	b	$-1/3$	$4,18_{0,03}^{0,04}$ GeV

Os quarks não são observados diretamente devido à uma propriedade conhecida por *confinamento de quarks*. Essa observação ocorre através de sistemas compostos (hádrons), que se estendem a uma distância de até 1 fm. Bárions e mésons são os sistemas mais simples conhecidos de hádrons. Um largo espectro dos bárions podem ser explicados como uma cápsula contendo três quarks confinados por glúons (ver

Subseção 2.1.3), estando os exemplos mais conhecidos encontrados no núcleo dos átomos — os prótons e nêutrons. Os mésons são compostos essencialmente por um quark e antiquark ligados transitivamente por um glúon. Os mésons mais leves existentes são os píons: π^+ , π^- e π^0 , compostos respectivamente por pares $u\bar{d}$, $\bar{u}d$ e $u\bar{u}$ ou $d\bar{d}$, e os káons, os mais leves dos mésons estranhos (contendo quarks s), K^+ , K^0 , K_S^0 e K_L^0 , compostos pelos sistemas $u\bar{s}$, $d\bar{s}$, $\frac{d\bar{s}+s\bar{d}}{\sqrt{2}}$, onde a diferença do K_S^0 e K_L^0 está no tempo de vida médio dessas partículas. O próton (uud) é o único bárion estável, já o nêutron (udd), apesar de estável na estrutura nuclear, quando isolado tem vida média de 15 minutos [5].

A estrutura instável dos hádrons faz com que os mesmos gerem, normalmente, uma sequência de decaimentos até atingirem o material dos experimentos de Física de Altas Energias. Assim, normalmente o objeto físico envolvido para sua assinatura é mencionado por jato hadrônico, uma vez que é possível associar a interação desse grupamento de partículas com um esguicho de um jato de água no detector. Os jatos costumam ser separados nestes grupos:

- Sabores Leves (LF): oriundos de quarks de menor massa de repouso, como os quarks u , d e s , bem como originados por glúons (Subseção 2.1.3);
- Sabores Pesados (HF): refere-se normalmente aos jatos originados pelos quarks b e c .

2.1.3 As Interações Fundamentais

As interações de partículas elementares dão-se via 4 forças fundamentais: força eletromagnética: a unificação do magnetismo e eletricidade; força forte, que mantém os quarks unidos formando hádrons; força fraca: responsável por decaimentos nucleares fracos; e gravidade. No MP estão descritas as três primeiras por via de teorias de campo quânticos. Especificamente para as forças eletromagnética e forte, as teorias são denominadas, respectivamente, de Eletrodinâmica Quântica (QED) e Cromodinâmica Quântica (QCD). Em decorrência do efeito desprezível da gravidade em termos subatômicos (a massa da partículas é da ordem de 10^{-27} kg [5]), ainda não foi possível fazer observações experimentais para quaisquer propostas de teoria quântica para a mesma [13].

O MP descreve as forças através de partículas mediadoras — excitadoras de campo —, os bósons de *gauge* que tem *spin* unitário. Os bósons característicos para cada uma das forças são estes: o glúon (g): interação forte; o fóton γ : interação eletromagnética; e os bósons W e Z: interação fraca.

Os bósons W e Z, mediadores da interação fraca, possuem respectivamente massa de $80,385 \pm 0,015$ GeV e $91,1876 \pm 0,0021$ GeV [12]. Esse fato limita o alcance

da interação fraca a cerca 10^{-3} fm, uma vez que uma partícula de massa M só pode existir como parte de um estado intermediário por tempo \hbar/Mc^2 , viajando uma distância não maior que \hbar/Mc . O bóson W possui carga elétrica, enquanto o bóson Z é neutro, sendo sua própria antipartícula. O glúon e o fóton não possuem carga, assim como massa de repouso, de forma que é esperado alcance de interação infinito para os campos portados por essas partículas. Contudo, diferente do campo eletromagnético, o campo de glúons é confinado a um alcance de 1 fm. Assim, para distâncias maiores a 1 fm, a interação eletromagnética é dominante, enquanto para distâncias menores, as interações forte e fraca também ocorrem.

Além disso, o MP teoriza que em um período curto após o *Big Bang* (10^{-36} s a 10^{-12} s) a interação eletromagnética e fraca advinham de uma única. Essa união é conhecida como a interação eletrofraca, observável em energias superiores a 100 GeV. Todavia, a descrição obtida pelos teóricos na época continha uma dificuldade a ser superada. O lagrangiano do campo descrito pela combinação das peças de todos os setores do MP não continha termos de massa. Enquanto isso, os bósons W e Z contém massa, bem como os férmions. Uma das soluções possíveis violava determinadas invariâncias importantes do MP. A resolução deu-se em 1964 pelos teóricos Peter Higgs [14, 15], François Englert e Robert Brout [16] e Gerald Guralnik, Carl Hagen, e Tom Kibble [17] através do mecanismo de quebra espontânea eletrofraca através de um campo escalar mediado pelo bóson de Higgs. Somente meio século depois, em 2012, essa predição do MP foi observada experimentalmente [10, 11].

Dos férmions, os quarks sofrem influência de todas as interações descritas pelo MP. Por outro lado, nenhum dos léptons sofrem interação forte. No caso do subgrupo dos léptons neutrinos, eles também não interagem via força eletromagnética, só havendo a possibilidade de interação via força fraca e, por esse motivo, são de difícil detecção [4, 5].

Os físicos esperam adicionar a interação gravitacional através da partícula gráviton, que deverá ter *spin* equivalente a 2 unidades e massa nula.

2.2 *Centre Européene pour la Rechère Nucleaire (CERN)*

O CERN [18] é o maior laboratório de Física de Partículas do mundo. Situado na fronteira da Suíça com a França, a organização possui 22 estados-membros, e conta a colaboração de físicos e engenheiros provenientes de 70 países — dentre eles o Brasil, que possui acordos de cooperação —, com 120 nacionalidades em torno de 600 instituições ou universidades.

Desde sua fundação em 1954, tem sido uma das referências de avanços tec-

nológicos, especialmente dedicado a avanços na Física Experimental de Partículas através de aceleradores de partículas. Dentre seus feitos constam a construção do primeiro acelerador de partículas realizando colisões próton-próton ($p - p$) (1971), a descoberta da corrente neutra (1973), dos bóson Z e W (1983), do bóson de Higgs (2012) e dos pentaquarks (2015).

Porém, devido à ampla colaboração científica e interdisciplinaridade necessária para o sucesso dos experimentos, o CERN é referência de avanços em outras áreas. Por exemplo, a invenção da *web* (1990) ocorreu no CERN como uma solução para facilitar a comunicação e colaboração de físicos de seus experimentos.

2.3 *Large Hadron Collider (LHC)*

Atualmente, o LHC [19, 20] é o principal acelerador de partículas em operação do CERN. Ele constitui-se de um circuito fechado aproximadamente circular de 27 km de circunferência e situado no subsolo, dentre 50 a 175 m. O circuito é percorrido por dois feixes em sentidos opostos que são colimados para colisão em determinados pontos contendo experimentos. Além de empurrar as fronteiras de estudo da Física de Partículas para um ambiente nunca antes explorado, esse acelerador conta com algumas peculiaridades impressionantes: a temperatura de operação de seus supercondutores são inferiores (1,9 K) a média do espaço; e o vácuo desenvolvido nos seus tubos de feixes (10^{-13} atm) é menos denso que o vácuo interestelar [18], prolongando-se no maior complexo de vácuo do planeta — 9.000 m³ nos 54 km percorridos por ambos feixes.

Esse acelerador de partículas, o mais ambicioso [5, 13] já construído, realiza colisões $p - p^1$ com energia de 13 TeV no centro de massa. As altas energias das partículas são atingidas por meio de uma cadeia de aceleração (Subseção 2.3.1), sendo o LHC a ponta final da cadeia. Os objetos físicos provenientes das colisões do acelerador podem conter informação detalhada sobre o funcionamento de mecanismos físicos, que permitem físicos complementar ou adaptar os modelos atuais de Física de Partículas, validar teorias, ou até mesmo se deparar com fenômenos não previstos ou desconhecidos.

O período atual equivale à *Run 2*, que começou em final de 2015 e se estenderá até 2018 [21]. Durante a *Run 1*, realizada entre 2009 a 2013, ocorreu na primeira observação da existência do bóson de Higgs [10, 11]. Essa descoberta resultou no laureamento conjunto de Peter Higgs e François Englert do Nobel de Física de 2013 [22], físicos teóricos que haviam previsto a sua existência. Para que esses mecanismos físicos possam ser bem descritos, faz-se necessário a produção de alta

¹Outros elementos acelerados, que fogem do escopo deste trabalho, são íons de chumbo, possibilitando análises de colisões $p - Pb$ e $Pb - Pb$.

estatística. Com esse objetivo, o LHC alcança níveis sem precedentes de taxas de colisão, obtido pela alta concentração das partículas em seu feixe — mensurada através da luminosidade. Esses e outros aspectos envolvidos nas colisões estão disponíveis na Subseção 2.3.2.

2.3.1 Complexo de Aceleração

A escala de energia obtida nas colisões do LHC é consequência de uma cadeia de aceleração (Figura 2.1) incluindo parte dos aceleradores anteriores do CERN. No caso dos feixes contendo prótons, inicia-se a extração de prótons de átomos de hidrogênio que são acelerados pelo LINAC 2 (50 MeV) — um acelerador linear — que inicia sequência. Realiza-se a continuação da aceleração² por via desta sequência de síncrotrons (aceleradores circulares que orientam as partículas através de magnetos): *Proton Synchrotron Booster* (BOOSTER) (1,4 GeV); *Proton Synchrotron (PSync)* (25 GeV); *Super Proton Synchrotron* (SPS) (450 GeV); até finalmente alimentar o LHC (13 TeV). O LHC é, igualmente, um síncrotron, assim como grande parte dos aceleradores de altas energias anteriormente construídos [19, 23].

O LHC não é uma circunferência. Sua topologia (Figura 2.2) é, na realidade, composta por octantes possuindo dois arcos ($\approx 2,45$ km) e um trecho retilíneo (≈ 528 m), que se iniciam e terminam no ponto de encontro intermediário dos sucessivos arcos. Os arcos contém 23 arranjos de uma mesma estrutura com magnetos responsáveis pela manutenção dos feixes na trajetória nominal³. Os trechos retilíneos do túnel do LHC são utilizados como Regiões de Inserção (IR), contendo experimento(s) ou um determinado serviço. A parte compreendida entre dois IR é chamada de setor.

As colisões ocorrem especificamente nos Pontos de Inserção (IP) em quatro dessas regiões, nas quais estão os experimentos dedicados ao estudo da Física de Partículas. Os dois IP para os experimentos de alta luminosidade estão localizadas em seções diametrais opostas, onde estão os experimentos ATLAS — IP1, referido simplesmente por Ponto 1 de Inserção (P1) pelo ATLAS — e CMS (IP5). Nas demais duas regiões encontram-se os detectores ALICE (IP2) e LHCb (IP8), onde também estão a injeção de feixes no LHC. A injeção ocorre pelo plano vertical, incidindo por baixo do plano de LHC. Quanto aos experimentos, este trabalho está envol-

²Comum para prótons e íons de chumbo.

³Existem dois magnetos principais: os dipolos, responsáveis por gerar a força centrípeta; e os quadripolos, produzindo o campo magnético que mantém as partículas do feixe concentradas em torno da trajetória nominal. Uma vez que a capacidade dos dipolos determina a energia máxima dos feixes, sua tecnologia está no limiar da ciência atual. Empregam-se supercondutores resfriados a 1,9 K por 120 t de hélio, de forma a criar um campo magnético de até 8,3 T por via de uma corrente de 11,7 kA. Já os quadripolos têm a propriedade de focar as partículas em um plano, e desfocar no plano perpendicular, o que causa oscilação (*betatron*) das partículas em torno da trajetória nominal.

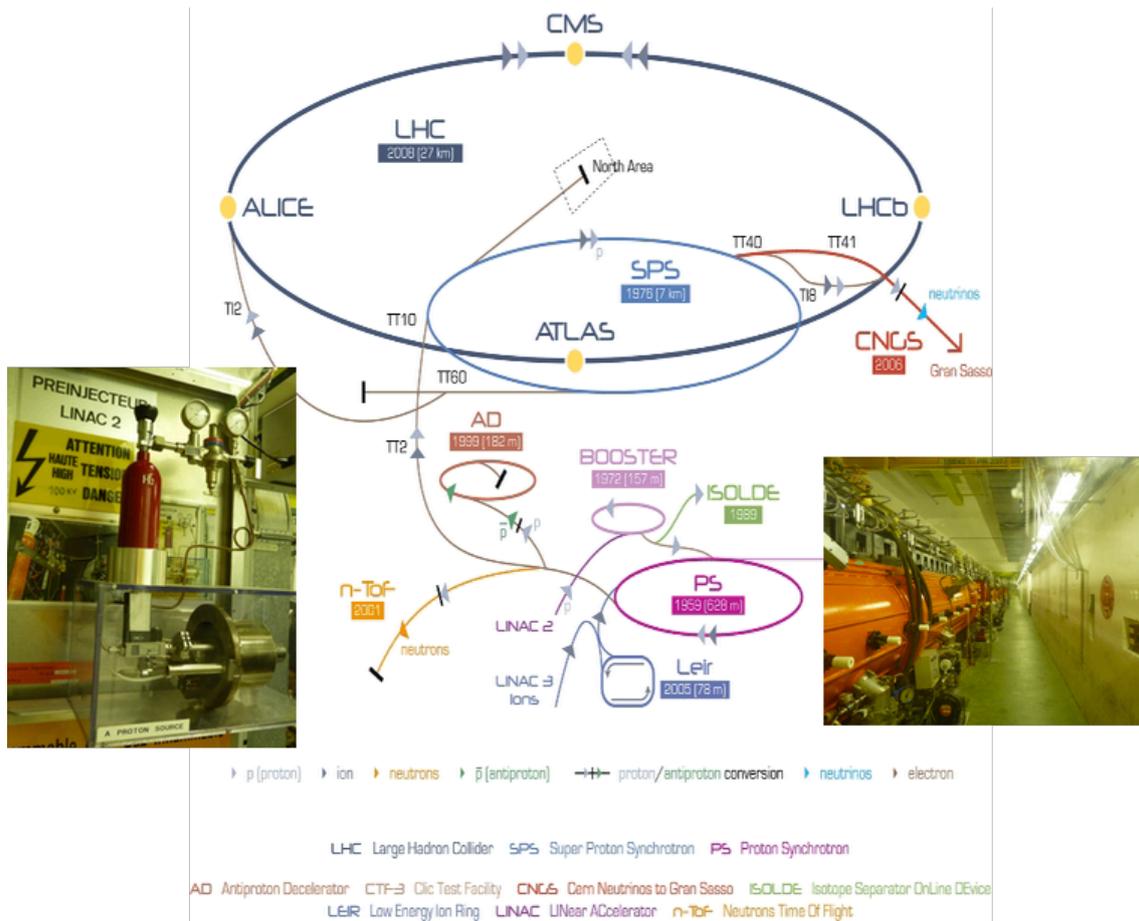


Figura 2.1: A cadeia de aceleração do LHC e seus maiores experimentos. A seta cinza claro corresponde ao sentido do deslocamento de prótons nos aceleradores. Estão disponíveis também uma foto de um exemplar da garrafa de hidrogênio (esquerda) de onde são retirados os prótons e do LINAC 2 (direita). O esquema do complexo de aceleração foi extraído de [24].

vido com o ATLAS e o Capítulo 4 dedica-se à descrição desse experimento. Outros grandes experimentos do LHC são: *Compact Muon Solenoid* (CMS) [25], *Large Hadron Collider beauty experiment* (LHCb) [26], *A Large Ion Collider Experiment* (ALICE) [27].

O acelerador possui uma série de serviços essenciais para operação instalados nos demais pontos:

- IR4 – contém sistemas independentes de *Radio Frequency* (RF), responsáveis pela aceleração dos feixes. Existem 8 células proporcionando tensão alternada com frequência do sistema RF (f_{RF}), que deve ser um múltiplo da frequência de revolução (f_{rev})⁴. A f_{RF} é de 400 MHz no eixo de propagação do feixe, e o seu potencial de aceleração é de 5 MV/m. O espaçamento dos feixes é determinado pelo acelerador *PSync*, que concentra as partículas em harmônicos a cada 25 ns. Esses valores criam um total de 3.564 harmônicos, possíveis de

⁴A f_{rev} pode ser calculada por: $c/26,659 \text{ km} = 11.253 \text{ Hz}$.

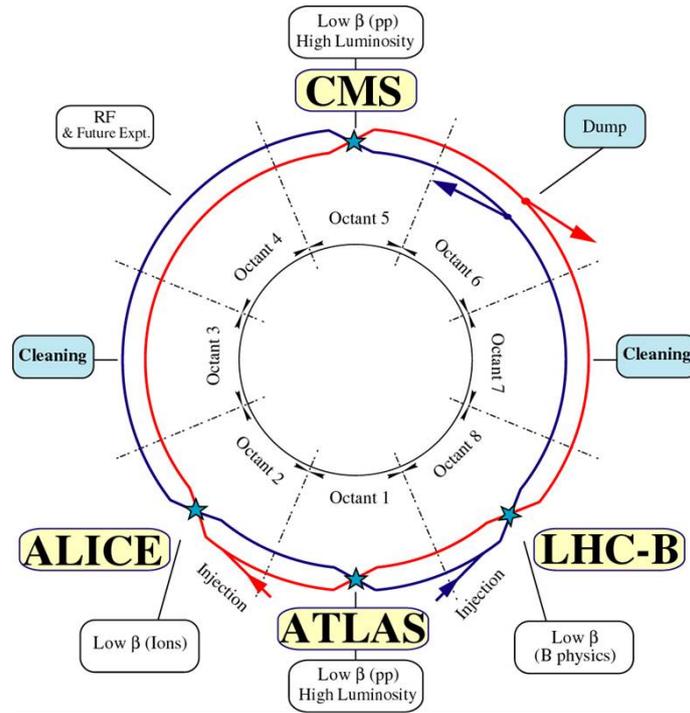


Figura 2.2: Esboço esquemático do LHC. Os octantes do LHC com suas respectivas Regiões de Inserção (IR), onde está discriminada as respectivas funcionalidades ou experimentos discriminados. O anel 1 (vermelho) percorre no sentido horário, enquanto o anel 2 (azul) no sentido anti-horário [23].

estabilidade para que uma partícula sempre encontre uma tensão acelerante. Por outro lado, nem todos os harmônicos são preenchidos, o que possibilita a remoção parcial dos feixes. Em operação nominal o LHC preenche 2.808 harmônicos em cada feixe, o que gera uma taxa média de colisão de 31,6 MHz;

- IR3 e IR7 – estão localizados dois sistemas de colimação, responsáveis pela proteção do acelerador contra as perdas inevitáveis e limpeza de irregularidades do feixes. Existe uma correção exclusiva em cada ponto: havendo sistema para a limpeza de oscilações de momento geradas pelo sistema RF (IR3); e atenuação das oscilações (*betatron*) em consequência da estratégia adotada para manter as partículas concentradas na trajetória (IR7). Essas são as áreas mais radioativas do LHC;
- IR6 – Sistema de Remoção independente para cada feixe. A sua função é extrair rapidamente os pacotes de cada um dos anéis do LHC e transportá-los até um material absorvedor. Em determinados casos, deseja-se apenas a remoção parcial do feixe, o que demanda alta velocidade de operação para evitar eventual eliminação de fração do feixe em bom estado. O Sistema de Remoção tem grande confiabilidade devido ao poder destrutivo dos feixes, o que condicionou sua concepção.

2.3.2 Colisões e Aspectos Relacionados

Ao atingirem a energia desejada, os feixes são direcionados para a colisão em determinados IP. As partículas do feixe ficam concentradas nas regiões harmônicas de aceleração (pacotes), ocorrendo colisões entre esses pacotes com alta taxa — frequência máxima (média) é de 40 (31,6) MHz. A esses fenômenos dá-se o nome de evento de colisão ou, simplesmente, evento. As colisões são muitas vezes referidas como interações, por não haver, de fato, o contato entre as partículas, mas sim a interação através dos campos. Nesta Subseção serão apresentados parâmetros importantes envolvidos nas colisões, onde se iniciará com uma breve introdução à cinemática das colisões e as coordenadas utilizadas (Tópico 2.3.2.1). Em seguida (Tópico 2.3.2.2), serão apresentados à categorização de colisões $p-p$, divididas de acordo com a intensidade da alteração da estrutura dos prótons em interação. A probabilidade de geração de um evento de interesse é proporcional à seção de choque do mecanismo físico (dependente na energia) estudado e à luminosidade dos feixes. Assim, deseja-se obter luminosidade máxima que gera efeitos colaterais. Uma apresentação da seção de choque, luminosidade e o efeito de empilhamento encontra-se no Tópico 2.3.2.3.

2.3.2.1 Coordenadas e Projeções

Como o trabalho está englobado no Experimento ATLAS, considerar-se-á o eixo de coordenadas adotado pelo mesmo. O ATLAS emprega um sistema de coordenada destro com sua origem no IP (Figura 2.3), onde a direção do eixo z segue a direção do feixe de partículas. O eixo x direciona-se do IP para o centro da circunferência do LHC e o eixo y direciona-se para cima. Coordenadas cilíndricas (r, ϕ) são utilizadas no plano transversal, onde ϕ é o ângulo azimutal em torno do eixo z e o ângulo polar θ é o ângulo do eixo z . A pseudorrapidez (η) é definida em termos do ângulo polar θ , como $\eta = -\ln \tan(\theta/2)$. O plano transversal é definido pelo corte transversal à direção de propagação do feixe (plano-xy). A projeção no eixo de propagação do feixe (z) é denominada de componente longitudinal. Simbolizar-se-á a projeção de a no plano transversal pelo subscrito a_T , enquanto se representará sua componente longitudinal por a_L . Um exemplo bastante empregado neste trabalho dá-se pela projeção da energia no eixo transversal (E_T).

2.3.2.2 Categorização de Colisões de Partículas

As colisões $p-p$ são bem descritas pela QCD, havendo uma gama de possibilidades para as interações dos prótons [29]. Dentre elas, as colisões elásticas (Figura 2.4a) não alteram as propriedades dos prótons em interação e causam apenas modificações em seus momentos. As colisões inelásticas, por sua vez, alteram a natureza dos

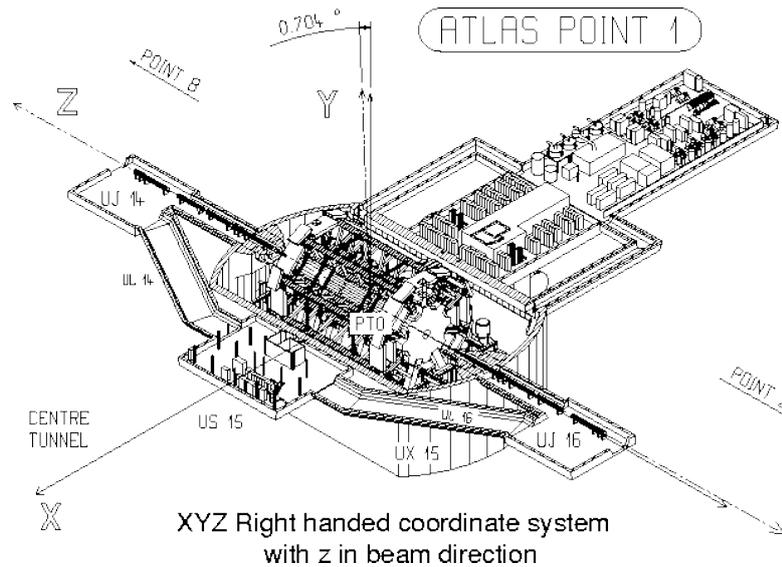


Figura 2.3: O Sistema de Coordenadas adotado para o P1, ponto de inserção do experimento ATLAS [28].

prótons, com maior ou menor intensidade. Colisões difrativas (Figura 2.4b) estão no limiar entre as colisões elásticas e inelásticas, decorrendo na produção de apenas alguns novos prótons na dissociação de um dos prótons da interação, enquanto o outro próton permanece intacto. Esse fenômeno pode ocorrer também em ambos prótons em interação (Figura 2.4c). Por outro lado, a interação dos constituintes dos prótons ocorrem quando a colisão inelástica é não-difrativa. Nelas os constituintes dos hádrons (quarks e glúons) atuam como feixes, onde é possível ocorrer múltiplas interações entre elas na mesma interação próton-próton. As Colisões Não-Difrativas (NDs) podem ser de dois tipos:

- **Colisões Inelásticas Rígidas:** decorrem em consequência de interações de curto alcance entre os constituintes do hádron. Nessas colisões (Figura 2.4e), há possibilidade de altas transferências de momento que permitem a produção de física de alta energia;
- **Colisões Inelásticas Suaves:** são o tipo de interação inelástica mais comum e devidas às interações de longo alcance entre dois hádrons cruzantes. Nelas os constituintes dos hádrons escoam uns pelos outros sem produzir colisão rígida (Figura 2.4d). Os resultados são jatos hadrônicos com baixa energia.

2.3.2.3 Seção de Choque e Luminosidade

A descoberta experimental de processos físicos na Física de Altas Energias costuma ocorrer indiretamente [31], através da reconstrução do processo físico utilizando estados finais de seus decaimentos, como no caso do bóson de Higgs [10, 11].

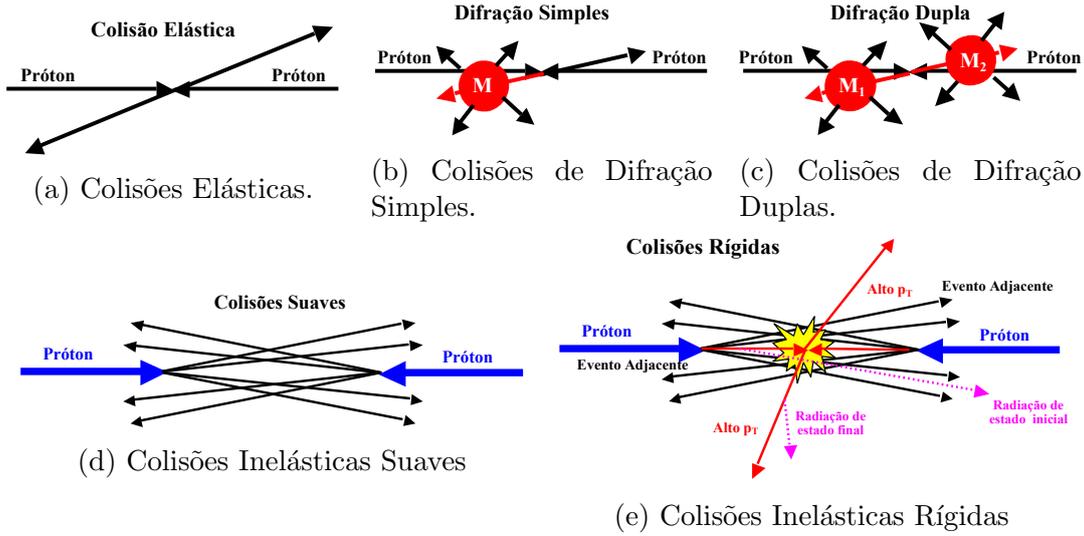


Figura 2.4: Esboços de categorias de colisões na Física de Partículas [30].

Em geral, cada um desses decaimentos tem sua própria fonte de ruído físico (mais ou menos intensa) e são analisados separadamente. Por isso, esse processo de reconstrução da física utilizando estados finais de um de seus decaimentos é assimilado por um canal de observação desse mecanismo físico. No exemplo citado, os canais fundamentais para a descoberta de tal bóson contaram com elétrons, múons e fótons isolados em seus estados finais⁵.

A análise obtém seu pleno sucesso quando há disponível estatística representativa suficiente, resultando em pequena flutuação em torno das previsões teóricas. Para que isso ocorra, as técnicas de reconstrução necessitam ser eficientes o suficiente — desde reconstrução dos estados finais e suas propriedades, como o de elétrons —, mas, também, que as colisões produzam observações necessárias para obter pequenas flutuações ao mesmo tempo que contabilize as ineficiências para a coleta de sinal e contaminação do sinal por ruído. Devido à raridade dos processos físicos a serem analisados, o problema deve ser otimizado pelas duas frentes — maximização da eficiência dos algoritmos de reconstrução e da taxa de produção dos processos físicos.

A taxa de ocorrência R de um processo físico pode ser descrito por [19]

$$R = \mathcal{L} \times \sigma, \quad (2.1)$$

onde σ é a seção de choque do processo e \mathcal{L} é a luminosidade. A seção de choque é um parâmetro do processo físico a ser determinado experimentalmente e comparado com as previsões teóricas. Ele determina a fração de partículas no feixe que reproduzem

⁵Os referidos canais foram $H \rightarrow ZZ \rightarrow 4l$, $H \rightarrow WW \rightarrow l\nu l\nu$ e $H \rightarrow \gamma\gamma$, onde o símbolo l é empregado para representar os elétrons e múons [10].

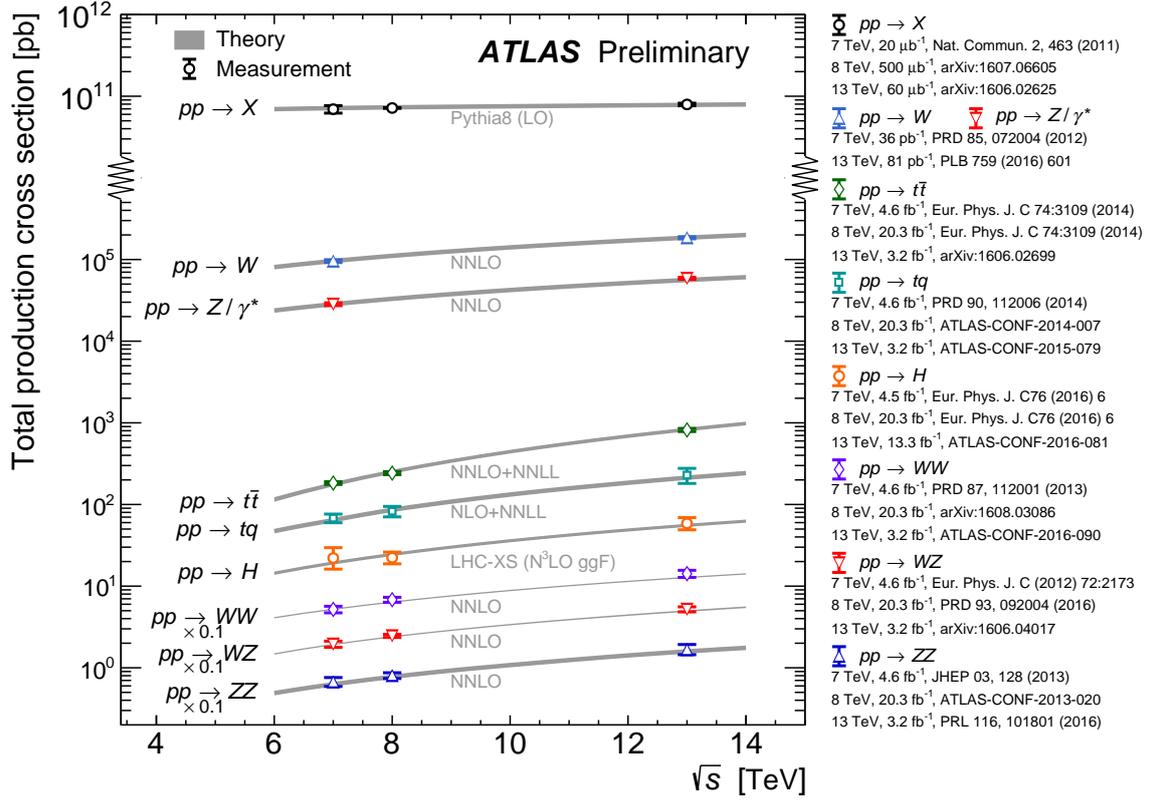


Figura 2.5: Resumo da medição pelo ATLAS de seção de choque (σ) da produção total de alguns processos físicos selecionados e apresentado em função da energia de massa (\sqrt{s}) de 7 a 13 TeV. As medições para os bósons-duplos (do inglês, *dibosons*) estão reescalados por 0,1 para permitir a apresentação sem sobreposição [32].

o processo físico em questão e tem unidades de área⁶ — usualmente em *barn* [6]. Há dependência desse parâmetro na energia das colisões (Figura 2.5), o que justifica a observação dos processos ocorrerem somente em determinadas escalas de energia. Observa-se um resumo dos valores experimentais obtidos pelo Experimento ATLAS para a produção total e para determinados canais descritos no Modelo Padrão na Figura 2.6⁷.

Logo, a seção de choque é da natureza do processo e só pode ser controlada pela energia das colisões. Por outro lado, a luminosidade [19, 35] pode ser compreendida como fluxo de partículas cruzando o plano transversal do IP por unidade de área — representada geralmente em termos de $cm^{-2}s^{-1}$ —, de forma simplificada, é descrita conforme

$$\mathcal{L} = f_{rev} n \frac{N_1 N_2}{A}, \quad (2.2)$$

onde f_{rev} é a frequência de revolução (31,6 MHz em média, 40 MHz no pico); n é o número de feixes; N_i equivale ao número de prótons em cada pacote (em valor nomi-

⁶Também se refere a seção de choque por área efetiva. [6]

⁷LO, NLO e NNLO são métodos de precisão para os cálculos empregando QCD. [34]

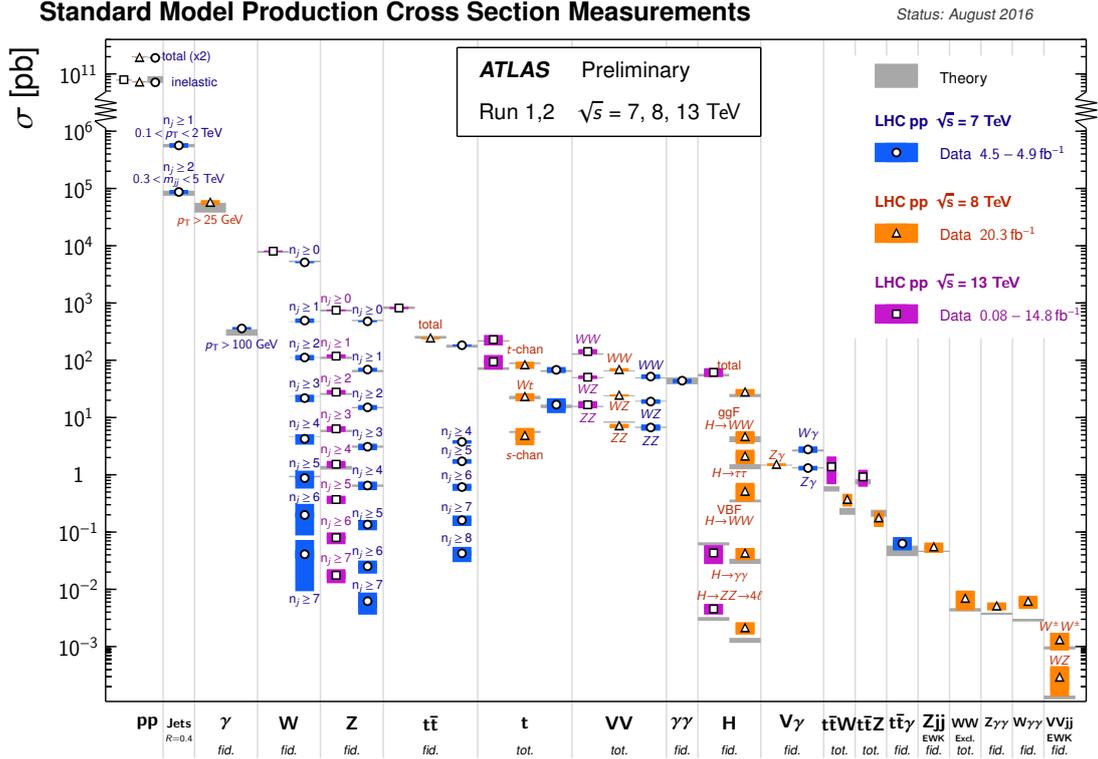


Figura 2.6: Um sumário dos resultados do ATLAS para medições de diversas seções de choque para a produção total e fiducial descrita pelo Modelo Padrão, comparada caso a caso com o esperado teórico. Todas as expectativas teóricas estão calculadas em NLO ou em ordem superior. As seções de choque inclusivas dos bósons-vetores W e Z foram calculadas em 35 pb^{-1} contidos no conjunto de dados de 2010. Todas as outras medidas foram realizadas com os conjuntos de dados de 2011, 2012 ou 2015 e 2016. A luminosidade usada em cada uma das medições é indicada próximo ao valor da medição. Incertezas para as previsões teóricas foram obtidas dos artigos originais do ATLAS. Nem sempre as mesmas usam as mesmas descrições para escalas e $pdfs$. As seções de choque teóricas para a produção de $W\gamma$ e $Z\gamma$ possuem correção não-perturbativas aplicadas para os cálculos de ordem fixa NNLO [33]. Nem todas as medições possuem significância estatística até o momento [32].

nal $N_1 = N_2 = 1,5 \times 10^{11}$); e A a área dada pela seção transversal do feixe. Os valores de luminosidade para operações do LHC medidos pelo Experimento ATLAS estão disponíveis na Figura 2.7a (Primeira Temporada de Coleta de Dados) e Figura 2.7b (Segunda Temporada de Coleta de Dados). Além disso, a luminosidade integrada para uma dada energia de colisão é uma medida de estatística disponível bastante empregada na Física de Altas energias, como pode ser observado nas Figura 2.6 e 2.5. Uma comparação para a luminosidade integrada coletada pelo ATLAS para os anos de operação do LHC até o momento está disponível na Figura 2.7c.

Valores teóricos para a operação do LHC em luminosidade bastante inferior àquela atualmente operada ($2 \times 10^{-33} \text{ cm}^{-2} \text{ s}^{-1}$ — comparar com Figura 2.7b) reite-

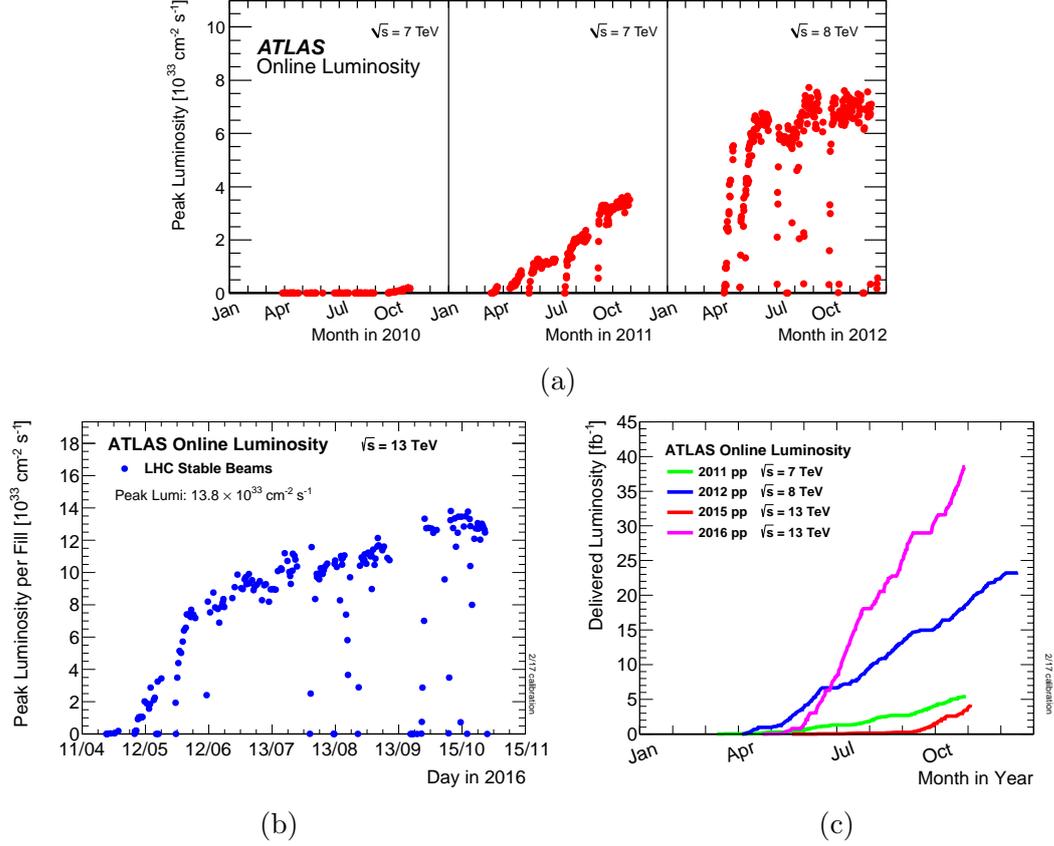


Figura 2.7: Valores medidos pelo ATLAS de luminosidade instantânea (a,b) e integrada (c). Em (a): luminosidade instantânea de pico entregue para o ATLAS para (da esquerda para a direita) os anos de 2010, 2011 e 2012. As medidas nessa Figura provém do estimador *online*. Em (b): a luminosidade instantânea de pico entregue para o ATLAS por feixes estáveis durante colisões $p - p$ com energia de 13 TeV no centro de massa em função das datas de 2016 para cada preenchimento de feixes realizados no LHC. A luminosidade é determinada empregando taxas de contagem medidas por detectores de luminosidade, baseando-se em análise preliminar de varreduras de separação de feixes de van-der-Meer durante 2016 (ver [35]). Em (c): luminosidade acumulada em função da data entregue para o ATLAS durante feixes estáveis e de colisões de alta energia de $p - p$ [36].

ram a importância para o potencial de produção de processos físicos de alta energia ao obter luminosidade, energia e taxa de eventos [6]:

- 15 minutos de colisões – o LHC produz mais quarks- b do que todo o período de operação de aceleradores anteriores $e^+ - e^-$ otimizados nessa produção (do inglês *B-factories*);
- 6 horas – mais quarks- t são produzidos que uma década de operação do Tevatron (o maior acelerador $p - p$ predecessor ao LHC);
- 1 dia – a mesma quantidade de bósons Z que o programa inteiro do acelerador LEP.

Por outro lado, existe um contraponto ao atuar com alta luminosidade. A elevação da luminosidade aumenta o número de interações próton-próton a cada evento de cruzamento dos pacotes. Isso implica em um acréscimo no número de partículas interagindo com o detector que podem gerar sobreposição do sinal caso as partículas incidam em regiões próximas — o fenômeno de empilhamento (do inglês *pile-up*). Esse fenômeno resulta em uma série de desafios para a operação, exigindo maior performance da reconstrução em diversos estágios da análise. Além disso, com a alta taxa de eventos do LHC, esse fenômeno pode transcorrer também por via de contribuições de eventos anteriores (do inglês *out-of-time pile-up*) quando o tempo de resposta do sinal for superior ao do espaçamento dos pacotes.

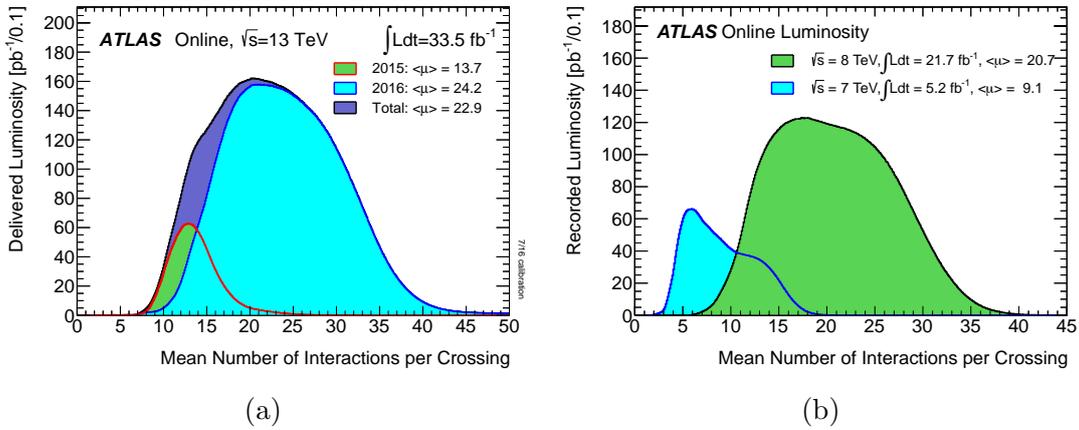


Figura 2.8: Número médio de interações por evento de cruzamento de pacotes. Em (a): distribuição das contribuições para a luminosidade da média do número de interações por cruzamento de pacotes para os conjuntos de dados de 2011 e 2012 (mais detalhes em [37]) para os períodos completos de colisões $p - p$. Em (b): a mesma distribuição, porém para os dados de 2015 e 2016. Em 2016 e em parte de 2015, as colisões foram realizadas com espaçamento entre os pacotes de 25 ns, enquanto em todos os demais períodos esse espaçamento foi de 50 ns [38].

Na Figura 2.8 estão os valores de número médio de interações por cruzamento de pacote ($\langle\mu\rangle$) para os anos da Run 1 e 2 de $p - p$ obtidos pelo ATLAS. Apesar da luminosidade de operação da Segunda Temporada de Coleta de Dados ser superior — praticamente o dobro no estágio final de 2016 em relação a 2012 (comparar Figuras 2.7a e 2.7b) —, não se observa o impacto desse aumento em $\langle\mu\rangle$. Isso é consequência do menor espaçamento entre pacotes utilizado na Segunda Temporada de Coleta de Dados, que é de 25 ns (exceto para o período inicial de coleta em 2015) ao invés dos 50 ns empregados na Primeira Temporada de Coleta de Dados. Com isso, obtém-se a mesma luminosidade com $\langle\mu\rangle$ reduzido por um fator de 2 no espaçamento de 25 ns. A contrapartida é a maior incidência de empilhamento por contribuição de eventos anteriores. Essa comutação se justifica uma vez que é possível estimar o sinal original quando o empilhamento ocorre assincronamente.

2.4 Detectores de Partículas

Nos experimentos atuais em aceleradores de partículas, as informações fundamentais obtidas por detectores não são os 4 momentos (p_0 , p_x , p_y , p_z) ou quaisquer outras propriedades das partículas a serem estudadas no programa de física coberto [6], uma vez que a detecção se dá indiretamente através de estados finais com tempo de vida longo o suficiente para interagir com o detector. Mesmo no caso dessas partículas, o processo de detecção é realizado por via da interação da partícula com a matéria (Subseção 2.4.1) do detector que gera sinais a serem tratados e manipulados para a obtenção das propriedades dessa partículas. Raramente os experimentos atuais conseguem informação suficiente para realizar tais tarefas com precisão satisfatória com apenas um tipo de detector, requerendo um conglomerado de subsistemas para o fim desejado pelo programa de física. É esse arranjo que na atualidade constitui um detector geralmente capaz de fornecer estas habilidades com precisão para as partículas envolvidas no programa de física: rastreamento; medição de momento/energia; informação para identificação; sincronização da obtenção dos sinais desses diversos subdetectores e filtragem de eventos (do inglês *trigger*). No último caso, o Sistema de Filtragem do ATLAS — responsável pela seleção *online* de eventos — é necessário para limitar a quantidade de dados armazenados, bem como a serem analisados *offline*, para ocorrências com potencial de estudo físico, que, no caso do ambiente das colisões do LHC, ocorre em altas taxas de evento.

De especial interesse para esse trabalho estão os calorímetros, um tipo de detector cuja especificidade é a absorção completa da energia da partícula incidente, decorrendo simultaneamente na geração de sinal proporcional à mesma. A aplicação de calorimetria é particularmente eficiente para a detecção de elétrons de alta energia. Porém, partículas de outras naturezas como hádrons, fótons e, menos intensamente, múons — que podem, igualmente, ser de interesse para os programas dos experimentos — também interagem com os calorímetros, o que resulta na necessidade de uma estratégia para identificação dessas partículas a fim de que essa informação possa ser empregada em análise física. Desta forma, introduzir-se-á a calorimetria na Subseção 2.4.2, focando na informação discriminante utilizada na tarefa de identificação de elétrons.

2.4.1 Interação de Partículas com a Matéria

A única maneira de detectar partículas e radiação — conseqüentemente, também, os estados finais das colisões — é através de interações com matéria. As interações das partículas com a matéria dependem da natureza da partícula incidente (ex. tipo de partícula envolvida, sua massa), de sua energia incidente e das propriedades do meio em questão (peso e número atômico, ocasionalmente também da estrutura mo-

lecular, como no caso de cristais). O entendimento dos processos de interação das partículas com a matéria são a base para a idealização de um tipo de detector [39]. A literatura disponível [12, 39–41] cobre em detalhes diversos processos, aqui se dedicando apenas em fornecer a base dos processos envolvidos na detecção por calorimetria. Existem interações específicas para partículas carregadas ou neutras, peculiarmente importante para o processo envolvido na detecção por calorimetria de elétrons. Além disso, como os mesmos são bem descritos pela literatura, é possível a aplicação de simulação de Monte Carlo, como o caso do simulador *Geaometry and Tracking* (Geant) [42], utilizado nas simulações deste trabalho.

2.4.1.1 Partículas Carregadas

Uma das principais fontes de interação de partículas carregadas (para a região de energias relativísticas intermediárias) com o meio efetiva-se via força coulombiana resultando, dependendo da proximidade com que acontecem, em ionização ou excitação dos átomos. A excitação ocorre quando a partícula carregada eleva o elétron em um átomo do meio para uma camada superior, enquanto na ionização a interação tem força suficiente para expulsar o elétron do átomo [40]. Em ambos casos, a energia envolvida é provida pela partícula atravessando o meio que interage com diversos elétrons do meio, reduzindo sucessivamente sua energia. É possível descrever o processo através da perda de energia média por comprimento de material ($-\langle \frac{dE}{dx} \rangle|_{\text{ion}}$)⁸, referida como fórmula de Bethe-Bloch. Além disso, elétrons ionizados com energia suficientemente alta podem estimular novamente ionização no meio, extirpando outros elétrons — *knock-on* ou raios- δ — do átomo.

No caso de camadas finas de material absorvedor (como em células de argônio líquido (LAr) ou cintilantes), há uma forte variação das medições de energia em relação ao valor esperado quando empregando a representação por $-\langle \frac{dE}{dx} \rangle$. Isso ocorre pois os detectores são capazes de medir apenas a energia depositada em seu material sensível (capaz de gerar sinal), e em determinados casos — por exemplo, quando elétrons atingidos saem do volume sensível do detector — apenas parcela da energia depositada no detector é observada. Assim, uma medida mais útil é dada pela perda média para o qual a energia é menor a um limiar, dada pela energia média de perda truncada ($-\langle \frac{dE}{dx} \rangle|_{\leq E_{\text{cut,ion}}}$)⁸. Porém, $-\langle \frac{dE}{dx} \rangle|_{\leq E_{\text{cut,ion}}}$ constitui-se de um corte na energia média total depositada por incidências contínuas de partículas

⁸Este trabalho realiza uma abordagem qualitativa das interações. As respectivas fórmulas e uma abordagem mais ampla para tema para informações aqui citadas podem ser encontradas em: ionização de elétrons ou pósitrons [39, p. 11]; ionização de partículas mais pesadas [39, p. 3-4]; energia média de perda truncada [12, Cap. 34, p. 10]; Teoria de Molière [12, Cap. 34, 15-17]; distribuição de Landau/Vavilov [39, p. 8]; perda mais provável de energia [12, Cap. 34, p. 32-33]; efeito de densidade de campo elétrico em velocidade relativística [12, Cap. 34, p. 7-9]; efeito de *bremstrahlung* [39, p. 19-20]; efeito de Moller e Brabha [12, Cap. 34, p. 17].

de mesma natureza e energia no meio, não o corte da energia perdida por uma única partícula no meio acima do limiar [12]. Nesses casos, uma melhor abordagem é obtida através da parametrização dessas flutuações por uma distribuição Landau/Vavilov⁸ em função da perda mais provável de energia ($\frac{\Delta_p}{x}$) para uma dada espessura x de material⁸. Esses comportamentos estão disponíveis na Figura 2.9a, enquanto a Figura 2.9b apresenta um exemplo de distribuição de Landau/Vavilov e sua versão com a correção de Bichsel para o efeito de densidade do campo elétrico para partículas em velocidade relativística⁸. Tanto $-\langle \frac{dE}{dx} \rangle|_{\leq E_{cut,ion}}$ ou quanto o valor mais provável geram um crescimento logarítmico com a energia da partícula incidente, tendendo, assim, a um limite superior em altas energias — *Fermi plateau*.

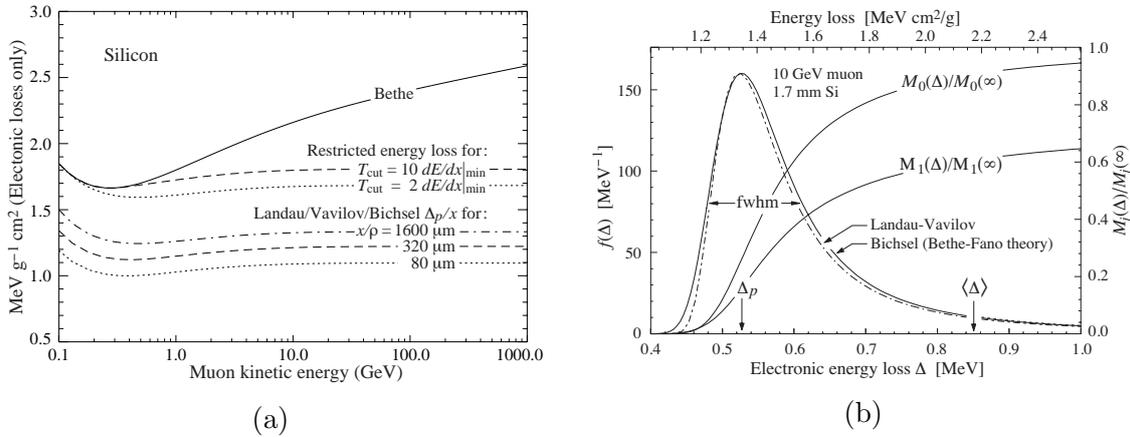


Figura 2.9: Exemplos de descrição para a perda de energia por ionização e excitação para múons atravessando silicóne. Em (a): perda média dada pela equação de Bethe e Bloch, dois exemplos de energia média de perda truncada (com corte T_{cut}), e a perda mais provável de energia por unidade de espessura. A unidade de espessura utilizada é a densidade por área do material por ser altamente independente das propriedades do material [39]. Tanto a perda média truncada ou quanto a perda mais provável de energia tendem a um limite devido ao efeito de densidade de campo elétrico para partículas em velocidade relativística (*Fermi plateau*) [12]. A mudança de $\frac{\Delta_p}{x}$ escala em função de $a \ln x + b$ conforme a espessura, diferente da perda média que é independente da mesma. Em (b): distribuição de depósito de energia de ionização e excitação para um múon de 10 GeV atravessando 1,7 mm de silicóne. A função de Landau-Vavilov (pontilhada) não aplica a correção para o efeito de densidade de campo elétrico apresentada para a função de Bichsel (contínua). Os valores acumulados normalizados para o valor médio de colisões (M_0) e perda média de energia (M_1) estão calculados para a partícula atravessando o material. Δ_p é a perda mais provável de energia e $\langle \Delta \rangle$ dividido pela espessura é o valor de perda média pela equação de Bethe [12, Cap. 34]⁹.

Além do efeito de ionização, as partículas carregadas interagem com o campo coulombiano dos núcleos e elétrons, sofrendo deflexão de sua trajetória. Esse pro-

⁹[12] provém permissão para uso das figuras desde que seja dado crédito aos seus autores conforme descrito em [43].

cesso ocorre sistematicamente enquanto a partícula transpassa o meio, fenômeno conhecido por espalhamento múltiplo. Frisa-se que nesse processo de interação não ocorre colisão com os elétrons atômicos, caso da interação por ionização. Com isso, a trajetória das partículas carregadas sofrem múltiplos desvios com alterações ditadas por pequenos ângulos em relação ao rumo original. A distribuição desses desvios é dada pela Teoria de Molière⁸ e tendem a uma distribuição gaussiana por decorrência do teorema do limite central, porém sua cauda é não-gaussiana devido à ângulos abruptos no caso de colisão das partículas com os núcleos.

Para escalas relativísticas maiores, no entanto, perdas por efeito *bremstrahlung* também devem ser consideradas. Esse efeito descreve quando a partícula sofre desaceleração no campo coulombiano do núcleo e elétrons⁸ do meio através da emissão de parcela de sua energia cinética na forma de radiação (fótons). Seu crescimento se dá diretamente proporcional à energia e inversamente proporcional à massa de repouso da partícula [39]. Assim, devido à massa diminuta de elétrons, o efeito de *bremstrahlung* desempenha maior influência sobre essas partículas que as demais. A distribuição de energia dos fótons (E_γ) emitidos segue a regra $1/E_\gamma$. Há uma tendência de emissão dos fótons na direção próxima àquela da partícula carregada ($\theta_\gamma \approx m_e c^2 / E_\gamma$) [39].

Observa-se na Figura 2.10 os efeitos de *bremstrahlung* e ionização para elétrons incidindo com energia na escala MeV em chumbo. Ao contrastar o desenvolvimento linear do efeito de *bremstrahlung* com o crescimento logarítmico em função da energia no caso da ionização, percebe-se a manifestação de um limiar — energia crítica (E_c) — no qual o efeito de *bremstrahlung* estabelece domínio sobre a ionização. No caso de elétrons, esse valor transita em torno de algumas dezenas de MeV para a maioria dos materiais [12], ressaltando-se a consonância dessa afirmação com o limiar contemplado na figura. Além disso, observa-se um desvio da proporcionalidade linear do crescimento para a faixa de energia avaliada, entre 1 a 1000 MeV, com crescimento de 60 % ($0,10 \rightarrow 0,16$) na fração de perdas por *bremstrahlung*. Porém, conforme o gráfico se aproxima da escala GeV há uma maior tendência a um crescimento linear.

Demais mecanismos também se manifestam adicionalmente ao *bremstrahlung*, especialmente em energias mais elevadas. No caso, a interação com o campo coulombiano do núcleo pode gerar pares elétron-pósitron através de fótons virtuais. Esse tipo de interação é predominante para múons de altas energias (ver Figura 2.14). Por outro lado, há predominância para altas transferências de momento para os fótons produzidos por *bremstrahlung* sobre a produção por pares [44]. Além disso, as partículas carregadas também podem interagir através de partículas de gauge virtuais com o núcleo do material, perdendo energia através de interação nuclear. As perdas fotonucleares descrevem o processo para essa interação por meio de fótons,

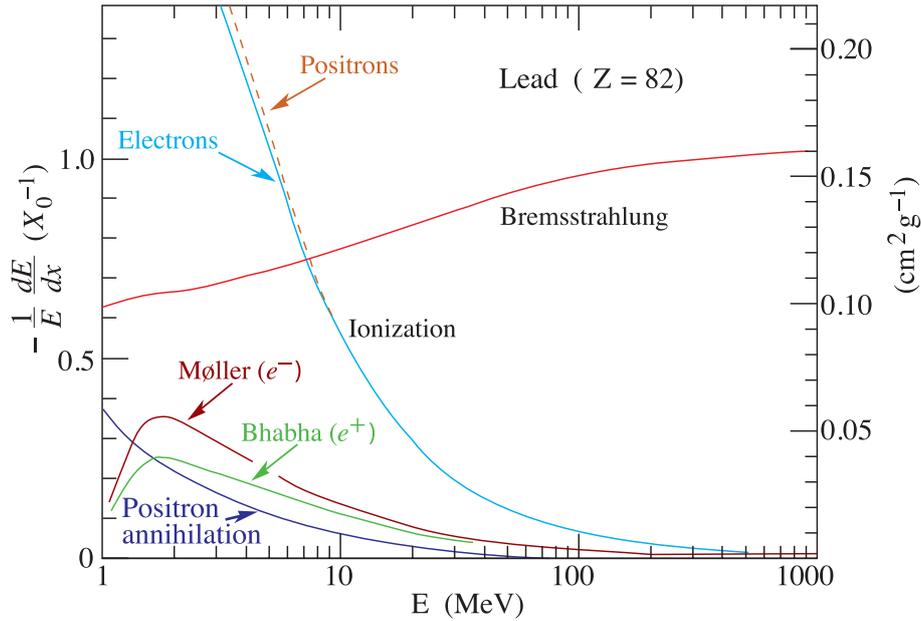


Figura 2.10: Fração de perda de energia média em função da energia da partícula incidente para conhecidas interações de elétrons e pósitrons com o meio. A fração de perda ($-\frac{1}{E} \frac{dE}{dx}$) na escala da esquerda é inversamente proporcional à comprimento de radiação (X_0) (ver 2.4.1.3). Na escala da direita, a distância é dada diretamente em unidades de densidade de área do material. Em baixas energias, elétrons e pósitrons também sofrem, respectivamente, os efeitos de Moller e Brabha⁸ [12, Cap. 34].

sendo importante para a descrição do processo para léptons de altas energias e desprezíveis para hádrons quando tendo em vista as interações diretas hádron-núcleo [39]. A interação fotonuclear gera partículas hadrônicas [45] ao invés das demais interações anteriormente mencionadas, que resultam em partículas eletromagnéticas. Para energias muito além daquelas operadas no LHC (10^{21} eV), as interações de elétrons e fótons com o núcleo superam as perdas por *bremsstrahlung*. Por fim, ambos mecanismos citados — produção de pares e fotonuclear — podem ser parametrizados através de uma função diretamente proporcional à energia da partícula incidente [39].

Diferente do efeito de ionização, contudo, os efeitos radioativos citados — *bremsstrahlung*, produção direta de pares de elétrons e interação fotonuclear — são caracterizados por altas transferências de energia e com elevadas flutuações. Com isso, geralmente decorrem grandes desvios ao descrever a perda média no meio através da soma desses efeitos [39]. Todavia, é interessante apresentar essa energia média total por ela sumarizar as possibilidades de perdas para partículas carregadas, como [39]:

$$\begin{aligned}
-\left\langle \frac{dE}{dx} \right\rangle \Big|_{\text{total}} &= -\left\langle \frac{dE}{dx} \right\rangle \Big|_{\text{ion}} - \left\langle \frac{dE}{dx} \right\rangle \Big|_{\text{brems}} - \left\langle \frac{dE}{dx} \right\rangle \Big|_{\text{par}} - \left\langle \frac{dE}{dx} \right\rangle \Big|_{\text{nucl}} \quad (2.3) \\
&= a(Z, A, E) + b(Z, A, E) \cdot E,
\end{aligned}$$

onde $a(Z, A, E)$ é a componente de interação por ionização e excitação, enquanto $b(Z, A, E)$ representa a componente de interação por radiação.

2.4.1.2 Fótons

As partículas neutras requerem uma interação primária com o meio (geralmente com o núcleo) de forma a alterar profundamente sua natureza a fim de serem detectáveis. Nesse processo, há transferência total ou parcial da energia para elétrons, núcleos de átomos no meio ou para partículas carregadas resultantes de reações nucleares [40]. De peculiar interesse deste trabalho estão as interações de fótons com o meio, por seus processos estarem diretamente envolvidos na detecção de elétrons por via de calorímetros (abordado mais adiante na Subseção 2.4.2).

Dentre os principais processos, dois resultam na transferência total de energia de fótons (ou seja, aniquilação do fóton) para partículas carregadas: efeito fotoelétrico e produção de pares; enquanto outro, o efeito Compton, resulta em um espalhamento (transferência variável) através de um ângulo relativamente grande. Há dependência no número atômico do material para as regiões em que cada um desses efeitos tem predominância. Sucede limites de regiões aproximados para materiais com número atômico $Z = 20$ (100) e uma breve apresentação dos efeitos [39, 40]:

- **Efeito fotoelétrico** ($\gamma + \text{átomo} \rightarrow \text{átomo}^+ + e^-$): predominância para energias inferiores a 0,1 (0,7) MeV. Nesse caso, os elétrons atômicos podem absorver por completo a energia do fóton, o que não seria possível se tais elétrons estivessem livres — consequência da conservação de momento. Em $\approx 80\%$ dos casos a energia é absorvida pela camada mais próxima do núcleo (K) devido à proximidade do núcleo do átomo, necessário para absorver a energia de recuo do elétron expelido dessa camada. Como o efeito atinge geralmente camadas internas, efeitos secundários ao fotoelétrico acontecem ao equilibrar o vácuo: emissão de raios-X com a energia da transferência de um elétron de camada superior ou através da captura de um elétron do meio. Essa energia emitida pode ser transmitida para um elétron do mesmo átomo, ocorrendo a emissão de outro elétron (efeito Auger) caso a energia incidente supere a energia de ligação da camada atingida. Todavia, os elétrons Auger contêm energia desprezível comparada ao do efeito fotoelétrico;

- **Efeito Compton** ($\gamma + e^- \rightarrow \gamma + e^-$): hegemônico entre 0,1 (0,7) < $E[MeV]$ < 10 (3). O fóton atinge um elétron do meio *quasi*-livre (ou seja, despreza-se a energia de ligação atômica do elétron), ocorrendo a deflexão do fóton com um ângulo de espalhamento θ . O elétron atingido, tido como elétron de recuo (do inglês *recoil electron*), recebe parcela da energia do fóton. O ângulo de espalhamento do fóton pode assumir qualquer valor, onde a energia máxima transferida para o elétron ocorre quando o ângulo é π . O ângulo do elétron está limitado a região $\pi/2$ devido à conservação do momento;
- **Produção de par elétron-pósitron** ($\gamma + \text{núcleo} \rightarrow e^- + e^+ + \text{núcleo}$): preponderante para a região superior de energia, $E[MeV] > 10(3)$. Esse efeito ocorre majoritariamente no campo coulombiano do núcleo, onde o fóton deve ter energia superior a, no mínimo, a massa de repouso de dois elétrons (1,02 MeV) mais a energia de recuo a ser transferida para o núcleo — essa sendo desprezível comparada a massa dos elétrons. Para fótons altamente energéticos, a seção de choque torna-se independente de energia do fóton. A divisão de energia é homogênea entre o par para baixas energias, enquanto ocorre leve assimetria para altas energias. A abertura angular entre os fótons é pequena ($\theta \approx m_e c^2 / E_\gamma$). Assim, no ambiente de altas energias do LHC, esse é o principal fenômeno envolvido na detecção de fótons, mas também de fundamental importância para a detecção por calorimetria de elétrons, por dar continuidade à absorção de energia (ver Tópico 2.4.2.1).

A intensidade I de um feixe de fótons após atravessar uma espessura x (unidade massa/unidade de área) de material com densidade ρ e coeficiente de atenuação de massa μ é dada por

$$I = I_0 e^{-x/\lambda}, \quad (2.4)$$

onde $\lambda = \left(\frac{\mu}{\rho}\right)^{-1}$. Ainda, descreve-se o coeficiente de atenuação de massa conforme

$$\mu = \frac{N_A}{A} \sum_i \sigma_i, \quad (2.5)$$

σ_i compondo a seção de choque para o i -ésimo processo, N_A é o número de Avogrado, e A é o peso atômico.

Os valores de seção de choque para as interações de fótons em chumbo estão expostos na Figura 2.11, que reforça a preponderância dos três efeitos citados. Destaca-se nessa figura uma tendência para seção de choque praticamente independente da energia do fóton partir de 1 GeV.

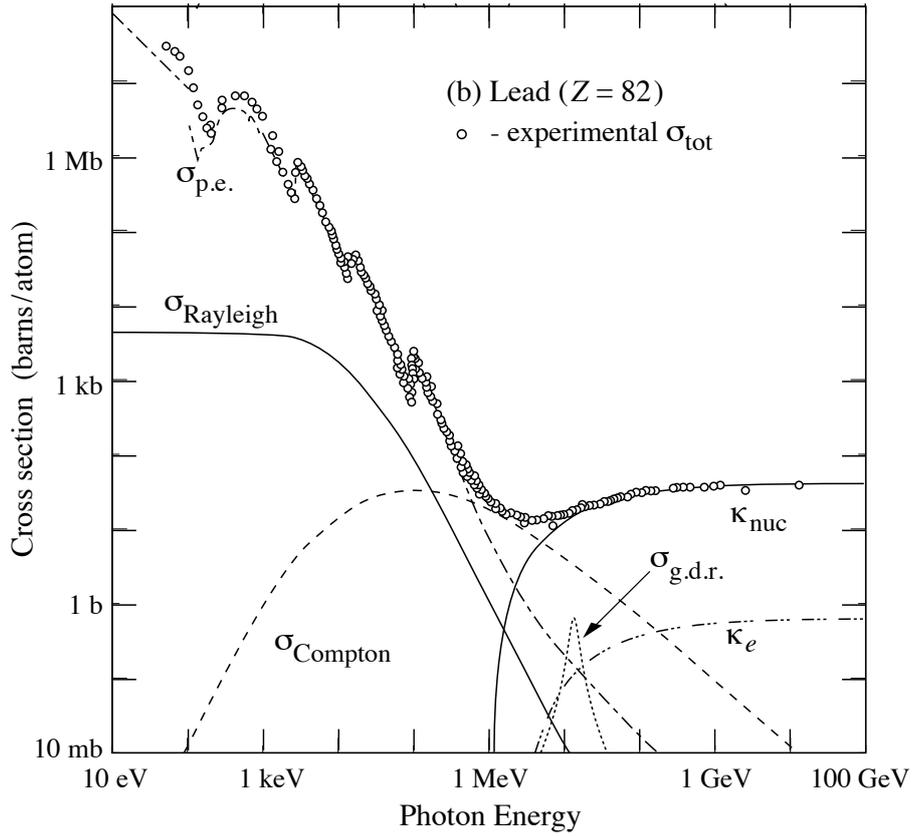


Figura 2.11: Comparação da seção de choque total dado em valores teóricos e as respectivas observações experimentais para fótons transpassando chumbo em função de sua energia. Os valores teóricos são especificados para as contribuições de diversos processos — incluindo alguns não cobertos por este trabalho. Os processos fundamentais para detecção de fótons e de maior seção de choque estão descritos por: efeito fotoelétrico ($\sigma_{p.e.}$); efeito Compton ($\sigma_{Compton}$); produção de pares (κ_{nuc}); através de interação com o núcleo [12, Cap. 34].

2.4.1.3 Comprimentos de Radiação e de Interação Nuclear

Conforme apresentado no Tópico 2.4.1.1 e 2.4.1.2, o efeito predominante para a perda de energia de elétrons de altas energias em um meio é dado via *bremmsstrahlung*. De maneira similar, a seção de choque para transferência de energia de fótons de altas energias ocorre pela produção de pares. Um parâmetro de especial importância para caracterizar esses efeitos é dado pelo X_0 , que é tanto [12]:

- a distância média na qual um elétron de alta energia tem sua energia atenuada para $1/e$ via *bremmsstrahlung*, correspondente à emissão de $\sim 63\%$ de sua energia em fótons;
- a fração de $\frac{7}{9}$ do caminho livre percorrido médio para a produção de par elétron-positron por um fóton de alta energia.

A primeira afirmação decorre da proporcionalidade linear da perda de energia conforme a energia da partícula. Ao assumir-se

$$-\left\langle \frac{dE}{dx} \right\rangle \Big|_{\text{brems}} = \frac{E}{X_0}, \quad (2.6)$$

fica evidente que X_0 é uma propriedade com unidade de distância e mede a influência entre o meio e a partícula para o efeito de *bremsstrahlung*. Na literatura, X_0 é sempre dado para a interação de elétrons com o meio. Todavia, é possível obter esse valor para demais partículas através da proporcionalidade $\tilde{X}_0 \propto m^2$. Por conseguinte, há maior pronúncia desse efeito para elétrons.

Integrando (2.6), obtém-se a energia média de elétrons após percorrer uma distância x do material conforme

$$E = E_0 e^{-x/X_0}, \quad (2.7)$$

o que conclui a primeira afirmação. Nota-se que é possível generalizar essa atenuação exponencial para a perda de todas as contribuições radioativas em decorrência da viabilidade de parametrização das mesmas por contribuições linearmente proporcionais a energia da partícula. Apesar do atenuador exponencial total não ser dado por X_0 , o mesmo continua sendo uma referência para descrever a atenuação energética devido às perdas radioativas.

Uma boa aproximação para a energia crítica (E_c) pode ser obtida ao expandir sua definição

$$-\left\langle \frac{dE}{dx} \right\rangle (E_c) \Big|_{\text{ion}} = -\left\langle \frac{dE}{dx} \right\rangle (E_c) \Big|_{\text{brems}}, \quad (2.8)$$

e, assim como X_0 , E_c escala com o quadrado da massa da partícula. No caso de múons, por exemplo, tem-se essa relação como $(m_\mu/m_e)^2 \approx 40.000$, resultando em $E_c = 890$ GeV no meio composto por ferro [39].

A segunda afirmação decorre de (2.4), onde é necessário o ajuste da seção de choque da produção de pares por um fator de escala de $\frac{7}{9}$ para obter a distância média em que ocorre a perda média da fração de $1 - \frac{1}{e}$ em energia por ocorrência do fenômeno de *bremsstrahlung* em elétrons [46]. Assim como no caso da contribuição total radioativa em partículas carregadas, o fator de escala não altera a capacidade descritiva de X_0 para a atenuação de intensidade da energia de um feixe de fótons.

Outro fenômeno abordado no Tópico 2.4.1.1, o espalhamento múltiplo de partículas carregadas, também tem sua descrição quantitativa em termos de X_0 . Nesse caso, a descrição é dada pelo RMS de desvio desse ângulo, como

$$\sigma_\theta = \sqrt{\langle \theta^2 \rangle} = \frac{13,6 \text{ MeV}/c}{p\beta} \sqrt{\frac{x}{X_0}}, \quad (2.9)$$

onde p e β são respectivamente o momento e velocidade da partícula, e x é a distância percorrida no meio pela partícula.

Ainda, cabe nesta ocasião ressaltar a distinção entre (2.7) e (2.4), onde aquela descreve a *atenuação exponencial da energia de uma partícula carregada* devido à perdas por radiação e esta dá a *atenuação de intensidade de um feixe de fótons* atravessando o material.

Além das interações descritas para partículas carregadas e fótons, existem também um grupo de interações que se fazem necessárias para compreensão deste trabalho. Os hádrons sofrem interações nucleares por intermédio da força forte. Mesmo no caso de hádrons carregados eletricamente, essas interações ainda descrevem parcelas importantes [46, 47] do processo de interação dessas partículas devido à sua maior massa em relação a de elétrons, o que debilita a intensidade de perda de energia pelo efeito de *bremstrahlung*. Devido à ampla variedade de processos com maior complexidade que as interações eletromagnéticas, as interações hadrônicas tornam os instrumentos envolvendo esses processos em sua performance e funcionamento de complicada otimização.

Apesar de mais complexas, a abordagem para as interações hadrônicas é similar àquela realizada para as interações eletromagnéticas. Considerando as contribuições das interações hadrônico-nucleares, define-se a absorção de hádrons em um meio segundo

$$N = N_0 e^{-x/\lambda_{int}}, \quad (2.10)$$

onde N_0 e N são respectivamente o número inicial de hádrons e a média de hádrons após percorrer a distância x no material, enquanto λ_{int} é o comprimento de interação nuclear [39]. Essa descrição é bastante similar àquela utilizada para fótons, onde λ_{int} determina o caminho médio livre de um hádron até que o mesmo sofra interação inelástica nuclear.

Há uma grande dependência das seções de choque hadrônicas com o núcleo em baixas energias (< 2 GeV) e, de forma rigorosa, mantém-se certa dependência mesmo no outro caso. Além disso, há certa variação na seção de choque conforme a natureza da partícula. Ambos efeitos refletem em λ_{int} , já que o mesmo é governado — e inversamente proporcional — pela seção de choque inelástica. O comprimento de interação nuclear refere-se, tipicamente, à interação de prótons com o meio. No caso de píons, por exemplo, $\tilde{\lambda}_{int,\pi} \approx 1,5 \times \lambda_{int}$, o que gera uma pequena probabilidade dessas partículas simplesmente não interagirem com o detector, ocorrência conhecida como perfuração (*punch-through*) [6].

A Figura 2.12 mostra valores de λ_{int} e X_0 para diversos materiais. Para a cobertura disponível, os elementos químicos obtêm comprimento nuclear superior ao de radiação. Em [39], cita-se que isso ocorre para os elementos químicos com $Z > 6$. Em consequência, para hipotéticos meios puros, hádrons precisam de uma maior quan-

tidade de material para que ocorra interações ao comparar com elétrons e fótons (no caso de fótons, existe o fator de escala $\frac{9}{7}X_0$). As regiões com menores λ_{int} estão compreendidos entre os elementos Tungstênio e Ouro (W-Au), seguidos por Urânio (U). Os elementos com Z superior ao do Urânio não possuem isótopos estáveis. Ferro (Fe) e Cobre (Cu) são boas escolhas quando considerando o preço. Por outro lado, materiais com alto Z possuem em geral baixo X_0 , como no caso do Chumbo (Pb) [12, Cap. 35].

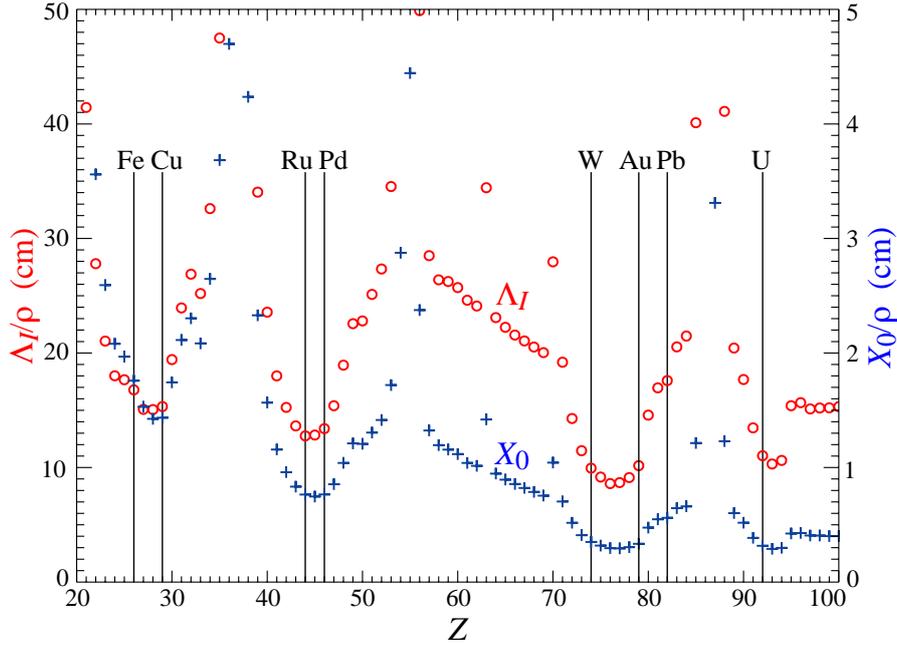


Figura 2.12: Comprimento de interação nuclear (círculos) e comprimento de radiação (cruz) para elementos químicos com $Z > 20$. Ambos comprimentos encontram-se divididos pelo ρ do material [12, Cap. 35].

2.4.2 Detecção de Partículas via Calorimetria

A reconstrução da física por meio de calorimetria [6, 12, 39, 41, 46, 47] implica na total absorção da energia da partícula incidente, ou seja, constitui-se de um processo destrutivo e, portanto, é a última informação obtida para as partículas abrangidas no processo de interação com esses detectores. Parte do desafio construtivo dos calorímetros é dado na escolha do material, sua arquitetura e quantidade de canais para que a informação retida seja representativa, ao mesmo tempo que atendendo limitações de rigidez mecânica, verba, espaço, peso etc. Apesar da grande variedade construtiva, os calorímetros para detecção de partículas são basicamente um bloco de matéria suficiente amplo para que a partícula disperse completamente sua energia através da interação com o meio e a energia dispersada nesse ínterim seja obtida de forma a representar esse processo de interação. Durante esse processo, a

energia absorvida é transformada em sinal proporcional mensurável por via de uma coletânea de células — regiões sensitivas — que detectam luz ou carga elétrica. Cabe definir duas terminologias empregadas ao se referir à segmentação das células do calorímetro. A segmentação longitudinal do calorímetro possibilita obter informação ao longo do eixo de deslocamento da partícula no material. Já a densidade de células capazes de determinar a posição de deposição de energia da partícula no plano perpendicular a esse eixo define a segmentação lateral ou densidade de células nos calorímetros.

Historicamente, as primeiras aplicações de calorímetros advieram do estudo de física de raios cósmicos. Seu emprego foi posteriormente aperfeiçoado e expandido para uso em aceleradores, hoje uma técnica bastante empregada para o estudo da física de partículas. Os detectores clássicos, como a Câmara de Bolhas, focavam na reconstrução dos 4 momentos para cada uma das partículas envolvidas no processo através de espectrografia de seus traços. O estudo da física por via de calorimetria permitiu tratar a informação em um nível mais amplo, compreendendo características que disponibilizam informação global dos eventos de colisão, como no caso de energia transversa faltante (\cancel{E}_T) e jatos hadrônicos [46], dentre outros. A calorimetria é essencial para identificação de partículas nos experimentos e, como fornece sinais de rápida aquisição, também o são para a filtragem *online*. Além disso, como mencionado no início desta seção, a combinação da informação dos mesmos com demais detectores fornecem informações suplementares que agregam mais poder para a reconstrução dos eventos de colisão. Explicitamente, dentre as propriedades dos calorímetros estão [46, 47]:

Propriedade I - Seu tamanho cresce apenas com o logaritmo da energia das partículas mais energéticas a serem detectadas, o que os torna consideravelmente compactos e eficientes em custo. Para que essa propriedade seja atingida, as partículas têm que sofrer processos com alta transferência de energia, como nos processos com perdas linearmente proporcionais a ela;

Propriedade II - Não há necessidade de campo magnético em contraste com os espectrômetros magnéticos;

Propriedade III - De uso versátil uma vez que há uma larga gama de partículas abrangidas pela Propriedade I, inclusive o caso de partículas neutras que não é coberto por espectrômetros magnéticos. Dentre os estados finais detectáveis nos calorímetros dos experimentos do LHC estão elétrons, fótons e hádrons. Possibilita a medição indireta de neutrinos através de \cancel{E}_T .

Propriedade IV - A resolução de energia dos calorímetros *melhora* conforme a energia E da partícula incidente de acordo com $1/\sqrt{E}$, diferente de es-

pectrômetros magnéticos cuja resolução deteriora linearmente. Esse fato torna sua aplicação extremamente oportuna para experimentos de altas energias;

Propriedade V - Com alta granularidade é possível obter medidas precisas da posição de interação da partícula com o detector;

Propriedade VI - Suas leituras fornecem também o tempo de voo das partículas que pode ser empregado para discriminação, porém com performance limitada a partículas com $\beta < 0.99$ [39];

Propriedade VII - Identificação de partículas ao explorar a descrição do desenvolvimento do processo de interação da partícula com o meio. Vale ressaltar que não se faz necessário absorção completa da partícula para o emprego de calorimetria para identificação de partículas. Apenas uma camada fina de material é capaz de fornecer informação discriminante. Inclusive para múons de baixa energia, onde mesmo sua assinatura com pequenos depósitos de energia pode auxiliar na sua identificação;

Propriedade VIII - Obtenção veloz de seus sinais gerando informação de calorimetria que pode, igualmente, ser tratada rapidamente. Propriedade de especial interesse para exploração na filtragem de eventos em ambientes como o do LHC, onde há altas taxas de eventos.

Propriedade IX - Para fornecerem medida de energia precisa, o calorímetro precisa ser calibrado. Normalmente se emprega partículas com identidade e momento conhecido, como o bóson Z.

No Tópico 2.4.2.1 realiza-se uma descrição do processo de interação das partículas com o calorímetro. Em seguida, descreve-se a medição e resolução de energia dos calorímetros (Tópico 2.4.2.2). Os aspectos construtivos estão em seguida (Tópico 2.4.2.3). Em vista dessas considerações, encerra-se a abordagem a calorimetria com uma discussão (Tópico 2.4.2.4).

2.4.2.1 Chuveiro de Partículas

Para elétrons, fótons e hádrons de altas energias do LHC, a interação das partículas com a matéria do calorímetro dá-se via uma cadeia contínua de eventos decorrendo na contínua multiplicação do número de partículas até sua extenuação, processo referido como *chuveiro de partículas*. Existem duas possibilidades para o desenvolvimento do chuveiro. Se a partícula em questão for um elétron ou fóton, o desenvolvimento é dado por uma cadeia de produção agregando, em sua esma-

gadora maioria¹⁰, elétrons e fótons — partículas eletromagnéticas (EM). No outro caso, partículas hadrônicas (HAD) — hádrons — desenvolvem processos que desencadeiam em tanto partículas HAD ou quanto decaimentos em partículas EM. Assume-se em média que para cada processo de interação hadrônico na cadeia, $\frac{2}{3}$ seja HAD enquanto o $\frac{1}{3}$ restante se transforme em partículas EM [39]. Assim, o desenvolvimento de chuviros HAD é caracterizado por uma componente minoritária (porém relevante) EM. Todavia, ao observar individualmente o desenvolvimento caso a caso, o seu processo de interação pode oscilar entre os extremos, contendo chuviros completamente compostos por partículas EM ou HAD.

Assim sendo, elétrons e fótons desenvolvem chuviros de partículas puramente EM, enquanto para hádrons se têm chuviros HAD com componentes variantes, em média minoritárias, EM. Compete elucidar como o processo de multiplicação de partículas ocorre em cada um dos casos e de que modo se desvanece.

2.4.2.1.1 Chuveiro de Partículas Eletromagnéticas No caso do chuveiro EM, um fóton de alta energia interage com o meio através de produção de pares (Figura 2.11), produzindo um elétron e um pósitron de menor energia e se aniquilando no processo. O elétron sofrerá o efeito de *bremstrahlung*, resultando na produção de um fóton sem que ocorra sua aniquilação. O pósitron pode ser tratado da mesma maneira que o elétron (Figura 2.10), por isso não se fará distinção no decorrer do texto adiante. Por ser dominado pelos processos de *bremstrahlung* e produção de pares, X_0 (ver Tópico 2.4.1.3) é uma boa descrição para determinar a distância característica do desenvolvimento desse chuveiro. Como visto, a produção de pares tem certo desbalanço entre a distribuição de energia entre as partículas. Particularmente para o *bremstrahlung*, há uma probabilidade maior da produção de fótons com uma pequena fração de energia do elétron em decorrência da distribuição proporcional à $1/E_\gamma$. Com isso, a distribuição de energia não é uniforme entre as partículas do chuveiro. Os elétrons com energia inferior a E_c interagem majoritariamente via ionização e, assim, não são capazes de manter a multiplicação de partículas. Logo, o chuveiro dissipa-se, na prática, quando não houverem mais partículas com energia superior a E_c , o que determina o seu desenvolvimento longitudinal. Fica evidente que chuviros com a complexidade acima descrita só podem se desenvolver acima de E_c (≈ 10 MeV). Uma boa aproximação (para $X_0 > 2$ e $E_0 > 1$ GeV) do desenvolvimento médio longitudinal (ver Figura 2.13) é dado por [48, 49]

$$\left\langle \frac{dE}{dx} \right\rangle = E_0 b \frac{(bt)^{(a-1)} e^{-bt}}{\Gamma(a)}, \quad (2.11)$$

¹⁰Como dito no Tópico 2.4.1.1, interações fotonucleares geram partículas hadrônicas, porém a seção de choque é mínima em comparação com o efeito de *bremstrahlung* para elétrons.

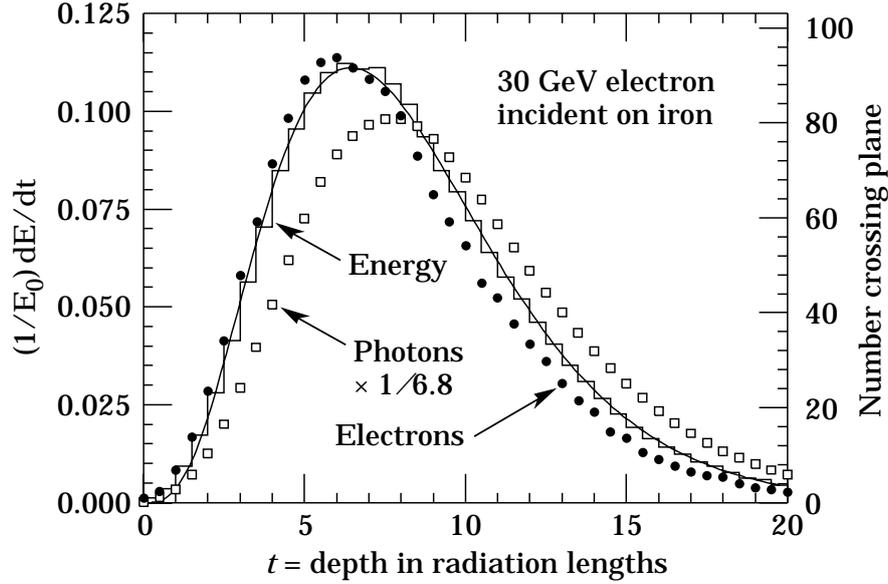


Figura 2.13: Uma simulação de um chuva desenvolvida pela incidência de um elétron em ferro. O histograma mostra a fração de energia depositada por fração de X_0 . A curva representa o ajuste através de (2.11). Os círculos indicam o número de elétrons com energia total superior a 1,5 MeV atravessando planos a cada $X_0/2$ (escala na direita), enquanto os quadrados representam o número de fótons novamente com energia superior a 1,5 MeV (normalizados para terem a área da distribuição de elétrons). O simulador empregado foi o EGS4 [49] [12, Cap. 34].

onde $t = x/X_0$; E_0 é a energia da partícula no instante de incidência com o material; a e b são parâmetros de modelo para o material; e $\Gamma(\cdot)$ é a função gama de Euler. Ao desejar obter a fração de energia E/E_0 para um determinado t , basta integrar (2.11) cancelando E_0 , o que decorre no decaimento logarítmico mencionado na Propriedade I. No caso do ponto t_{max} em que a aproximação (2.11) atinge o seu máximo, obtém-se [12, 39, 47]

$$t_{max} = \frac{a-1}{b} = \ln \frac{E_0}{E_c} + C_j, \quad j = e, \gamma, \quad (2.12)$$

no qual $C_\gamma = +0,5$ e $C_e = -0,5$. Ou seja, a variação para a posição de deposição média dos chuviscos varia aproximadamente por apenas $1 X_0$ para elétrons e fótons, elétrons atingindo o máximo ligeiramente antes que os fótons.

Ao observar a Figura 2.13, percebe-se que o chuva é dominado por fótons¹¹, inclusive havendo um decaimento mais veloz do número de elétrons compondo o chuva, o que acentua ainda mais a composição por fótons em altas profundidades.

Indo além, as partículas produzidas no chuva têm propagação bastante próxima à propagação da partícula que as originou (na ordem de $\theta = m_e c^2/E_\gamma$). As-

¹¹Atentar para a normalização do número de fótons.

sim, o desvio predominante é dado pelo efeito de espalhamento múltiplo em campo coulombiano [12, 39, 47]. Uma medida que caracteriza a distribuição lateral do chuveiro é dada pelo raio de Molière

$$R_M = \frac{21 \text{ MeV}}{E_c [\text{MeV}]} X_0 \quad \{\text{g/cm}^2\}, \quad (2.13)$$

representando a deflexão média lateral de elétrons com energia E_c após transpassarem $1 X_0$. Em média, 90 % (99 %) da energia está contida em um cilindro com $1 R_M$ ($3,5 R_M$). Porém, além de $3,5 R_M$ ocorre limitações para a descrição da energia por via dessa representação, que não leva em consideração, assim como X_0 , as peculiaridades das interações ocorrendo na região MeV [46]. O desenvolvimento do chuveiro começa por um núcleo estreito que se alarga conforme a profundidade. Essa descrição costuma ser feita pela soma de duas Gaussianas [12, Cap. 34].

2.4.2.1.2 Chuveiro de Partículas Hadrônicas Ao penetrar o meio, um hádron sofre majoritariamente interações com o núcleo, havendo uma grande multiplicidade de efeitos possíveis. A geração de partículas energéticas ocorre por meio de colisões geralmente 50 % inelásticas (ou seja, transferindo essa parcela para as partículas secundárias) [39, 47], produzindo hádrons com energias próximas à escala de energia da partícula primária (ou seja, GeV). Essas ditas rápidas componentes hadrônicas são dominadas por prótons, nêutrons, píons carregados e píons nêutrons. Dessas partículas, píons são as partículas mais relevantes. Tal fenômeno acontece, em média, a cada λ_{int} (ver Tópico 2.4.1.3) percorrido por hádrons no meio, medida característica de distância longitudinal equivalente ao X_0 para os chuveiros EM. Devido à independência de carga nas interações fortes, em média 1/3 dos píons produzidos nas colisões hadrônicas são neutros [47]. Esses píons (e outros mésons como o η^{12}) são instáveis e decaem em fótons ($\pi^0 \rightarrow \gamma\gamma$, $\eta \rightarrow \gamma\gamma$) antes de terem chance de interagirem hadronicamente. Uma vez sendo transferida a energia para uma partícula EM, ela seguirá as regras de interação do chuveiro EM descrito no Subtópico 2.4.2.1.1. Assim, essa energia será depositada através de interações EM, sendo desprezível a possibilidade de retorno para interação por vias de processos HAD. Ainda, a fração de energia EM (f_{em}) do chuveiro tende a aumentar (de forma aproximadamente logarítmica [46]) conforme a energia da partícula HAD incidente. Exemplo de valores típicos são $f_{em} = 0,5$ (0,7) para um pión incidente de 100 (1.000) GeV. Por outro lado, há uma grande dependência na quantidade dessas partículas produzidas nas primeiras interações hadrônico-nucleares, que causam uma grande flutuação da fração. Conforme se eleva a energia para escalas muito elevadas, os hádrons destinam-se a depositar sua energia através de grande emissão

¹²O méson η é composto pelo sistema $\frac{u\bar{u}+d\bar{d}-2s\bar{s}}{\sqrt{6}}$.

de fótons [47].

Uma parcela adicional de energia da colisão hadrônico-nuclear é transferida para o núcleo. Dessa energia, uma fração significativa (10 %) é transferida para a quebra de ligações nucleares, energia que não pode ser recuperada [47]. Outra parte da energia é emitida por prótons (valor típico em torno de 40 %) ou nêutrons (40 %). As partículas emitidas nesse processo têm energia na escala MeV [47]. O restante, 10 % da energia, é gasto em processos como excitação, evaporação nuclear, espalhamento etc. onde ocorre emissão de fótons de baixa energia, entre outros.

Ao avaliar os processos como um todo, o espectro de partículas no chuveiro é extremamente dominado por partículas de baixa energia, composta por elétrons, pósitrons, fótons e nêutrons — esses dois últimos de maior peso. São essas partículas compondo o espectro de baixa energia que conserva toda a informação de energia medida pelo Calorímetro Hadrônico (HCAL). Há diferença de resposta do meio para essas partículas [46, 47]. No caso de nêutrons, a única possibilidade para a perda de sua energia é através de interações fortes, o que causa uma maior demora para sua absorção (10 ps). Parte desses nêutrons — nêutrons leves — produzem fótons com energia demasiadamente pequena para serem detectáveis pelo calorímetro. Há dois grandes picos na contribuições de fótons: um em 8 MeV resultante de processos físicos nucleares típicos; e outro em 511 keV resultado da produção de fótons via aniquilação de e^+e^- . A energia carregada por esses fótons é considerável, contudo, apenas uma parcela dela é capturada pelo Calorímetro Eletromagnético (ECAL) devido ao atraso ($\lesssim 1\mu s$). Esses sinais atrasados (nêutrons e fótons) podem causar uma extensão no sinal, que, dependendo da taxa de amostragem, pode ser utilizado para discriminação entre chuveiros EM-HAD [46].

Assim, uma parcela da energia (f_{inv}) do chuveiro HAD é perdida sem que possa ser recuperada. Dentre suas contribuições — como os já citados: fótons atrasados; nêutrons leves (ou atrasados); quebra de ligações nucleares — estima-se uma perda de 20 %-40 % da energia [12, Cap. 35].

Por ser descrito em termos de λ_{int} , o chuveiro HAD tende a prolongar-se mais longitudinalmente. Por exemplo, um chuveiro desenvolvido por um pión de 300 GeV em urânio necessitaria de 80 cm para ser absorvido pelo meio, enquanto o mesmo meio seria capaz de absorver tal energia de um elétron incidente com 10 cm de material [46].

Além do desenvolvimento longitudinal mais amplo, os chuveiros HAD igualmente possuem uma largura lateral mais larga quando comparados com chuveiros EM. Como foi visto para os chuveiros EM, a sua maior contribuição para a largura lateral é dada por espalhamento múltiplo. Por outro lado, essa contribuição no caso das interações hadrônicas originam principalmente das grandes transferências de momento com as interações nucleares [39].

2.4.2.2 Medição e Resolução em Energia

Supondo o caso da absorção de elétrons em um chuveiros EM, a soma do comprimento de *traço*¹³ total percorrido por partículas carregadas do chuveiro (T_0) é proporcional a [47]

$$T_0 \propto X_0 \frac{E_0}{E_c}, \quad (2.14)$$

onde E_0/E_c representa o número de partículas no chuveiro e E_0 a energia da partícula incidente no meio. Assim, o traço total das partículas possibilita obter valores proporcionais a E_0 . Em suma, a captação de sinal dessa medida pode ser realizada por uma das duas vias: carga elétrica ou luz. Reforça-se a exclusividade dos métodos pois, apesar da abordagem ideal dar-se através da captação de ambas informações para obter uma melhor capacidade de descrição da deposição de energia, ainda não foi possível a realização de um calorímetro de larga escala capaz de captar ambas informações por decorrência das dificuldades técnicas (geometria para coleta) [47]. A informação exata captada pelo calorímetro depende da tecnologia empregada, mas, em geral, não é nenhum dos perfis disponíveis na Figura 2.13. Justamente pelo sinal medido não ser a fração de energia depositada, mas proporcional à ela, advém a necessidade de calibração (Propriedade IX). Por exemplo, no caso de contadores de gás, o sinal pode ser bem próximo ao número de elétrons. Por outro lado, para dispositivos com regiões sensíveis grossas, o mesmo pode ser próximo a deposição de energia [12, Cap. 34]. Em termos práticos, a absorção só ocorre para elétrons ou fótons acima de um limiar de energia.

A resolução intrínseca para um calorímetro ideal (tamanho infinito e sem deterioração de resposta devido à efeitos de instrumentação, inclusive por ineficiências na coleta de sinal) é dado por flutuações em T_0 . O mesmo é proporcional ao número de segmentos de traços no chuveiro que, como foi visto, é um processo estocástico descrito via seções de choque por diversos efeitos de interação das partículas com a matéria. Assim, a resolução intrínseca é dada por [47]

$$\frac{\sigma(E)}{E} \propto \frac{1}{\sqrt{T_0}} \propto \frac{1}{\sqrt{E_0}}. \quad (2.15)$$

A resolução intrínseca é um dos termos que compõe a parcela estocástica (σ_{est}) da resolução de calorímetros verídicos. Existem outras contribuições, como termo relacionado a ruído eletrônico (σ_{noise}) e outras flutuações independentes a energia (σ_{const}). Tem-se, assim, [47]

¹³Comprimento de traço no sentido de ser, realmente, todo o deslocamento realizado pelas partículas no meio, em contraposição, por exemplo, com o comprimento longitudinal.

$$\frac{\sigma(E)}{E} = \frac{\sigma_{est}}{\sqrt{E_0}} + \frac{\sigma_{noise}}{E_0} + \sigma_{const}, \quad (2.16)$$

o que mostra a dependência das contribuições de cada um dos termos para as diferentes faixas de energia. Por isso, as técnicas para o desenvolvimento de calorímetros buscando resolução ótima podem divergir de acordo com o alcance dinâmico em energia dos experimentos. Uma revisão da contribuição de cada um desses termos pode ser encontrada em [47]. Exemplos de resolução em energia de calorímetros típicos dedicados para a absorção de chuveiros EM estão disponíveis na Tabela 2.3.

Indo além, outras contribuições podem afetar a resolução, como:

- vazamento longitudinal do chuveiro, devido à flutuações longitudinais evento a evento que podem exceder a profundidade do calorímetro;
- vazamento lateral, em decorrência da limitação no tamanho da região estudada para evitar contribuições para minimizar o efeito de empilhamento e ruído eletrônico;
- perdas devido ao início prematuro do chuveiro em materiais, como em outros detectores ou material morto, anteriores ao calorímetro; e
- cobertura não hermética, quando há regiões mortas ou *crack* no calorímetro necessárias por natureza construtiva do experimento.

Todavia, os chuveiros HAD têm desenvolvimento mais complexo que resulta em resolução em energia consideravelmente pior que a dos chuveiros EM. Isso se deve a contribuições de f_{inv} que adicionam flutuações (atualmente) irreduzíveis na aquisição do sinal. Além disso, há a presença de uma componente EM com altas flutuações (altamente dependente nas primeiras interações hadrônico-nucleares) evento a evento e que cresce não linearmente (aproximadamente com o logaritmo) da energia do hádron incidente. Logo, isso tem consequências para a resposta do calorímetro caso o calorímetro tenha eficiência diferente para a medição de energia de chuveiros EM (e) ou HAD (h). Em consequência, tem-se a previsão de determinados efeitos [46]:

- Distribuição não-gaussiana para a deposição de energia de hádrons de mesma energia;
- Acréscimo de flutuação das medições devido à f_{em} , refletido na resolução;
- A resolução perde sua propriedade de melhoria conforme \sqrt{E} ;
- Há uma não linearidade do sinal em função da energia do hádron;

- A razão de sinal depositado por elétrons e píons (representando deposições hadrônicas) e/π é dependente em energia.

Conforme o crescimento da energia, esse efeito se torna o termo rapidamente dominante na contribuição. Além disso, determinados experimentos observaram flutuações devido à não linearidade na ordem de 20 % para o aumento em uma única ordem de grandeza na escala. Por isso, uma grande melhoria da resolução de energia para a medição de chuveiros HAD dá-se através da compensação da resposta dos chuveiros. Isso é realizado termos de e/h — um ajuste da resposta do calorímetro para e/π em uma faixa de energia — que fornece uma grandeza independente a energia.

O melhor HCAL já construído, para o experimento ZEUS, conseguiu obter resolução de $\sigma(E)/E = 35 \text{ \%}/\sqrt{E \text{ GeV}}$ através de compensação [39]. Porém, HCAL atuais construídos para o LHC têm conseguido resolução consideravelmente boas mesmo sem compensação. No caso do *TileCal* do ATLAS (ver Subseção 4.2.3), a resolução obtida foi de $\sigma(E)/E = 42 \text{ \%}/\sqrt{E \text{ GeV}}$ (feixe de teste) com $e/h = 1,37$ [39, 50].

2.4.2.3 Aspectos Construtivos

Emprega-se, tradicionalmente, calorímetros dedicados para a detecção de cada um dos processos de desenvolvimento de chuveiros. Não há surpresas quanto a nomenclatura: ECAL, destinado a detecção de chuveiros EM; e HCAL, responsável pela contenção da energia dos chuveiros HAD. Em ambos casos, o calorímetro deve conter profundidade suficiente, medida em termos de seus comprimentos característicos — X_0 e λ_{int} , respectivamente —, a fim de serem capazes de detectar partículas. Os valores giram em torno de 15-30 X_0 (5-8 λ_{int}) para o ECAL (HCAL) [12, Cap. 34].

Do ponto de vista construtivo, o HCAL constitui-se de calorímetro de amostragem contendo sucessivas camadas de meio absorvedor e de meio ativo. O meio absorvedor, ou passivo, é o principal responsável pela retenção da energia da partícula, composto por material denso (metálico [12, Cap. 35]). O meio ativo proporciona também a geração de sinal, geralmente constituído de um líquido nobre ionizante, uma câmara preenchida com gás, um semiconductor ou radiador de Cherenkov¹⁴ [12, Cap. 35]. Assim, apenas a parcela do sinal no meio ativo é amostrado, o que origina o nome do método construtivo. Conseqüentemente, há a adição na resolução do calorímetro devido às flutuações de amostragem, o que reduz o potencial de sua resolução. Por outro lado, a própria natureza complexa dos chuveiros HAD costuma

¹⁴A radiação de Cherenkov não foi mencionada na Subseção 2.4.1 pois o tema se distancia do cerne deste trabalho. Uma abordagem ao tema pode ser encontrada em [39, p. 142-146].

ser o termo dominante na resolução de energia, tornando o método de amostragem ideal para a construção de um HCAL até o momento.

Há, contudo, outra possibilidade com relação à construção de um ECAL, ao se desejar atingir o estado da arte em nível de resolução. Cristais cintiladores inorgânicos (BaF₂, BGO, CsI, PWO), dentre outros, possibilitam absorver e detectar o sinal em um único meio. Assim, os calorímetros são homogêneos e possuem, como propriedade, o fato de praticamente todo o seu volume ser sensível à deposição de energia. Atingir valores supremos de resolução tem sua contrapartida. Talvez, a maior delas seja a conta — cristais de cintilação têm alto custo e limitação para a produção em larga escala [39]. Além disso, esses materiais não costumam ser muito densos, aumentando o tamanho do experimento [46]. Para exemplificar a afirmação, alguns valores de X_0 para cristais cintiladores são: BGO, $X_0 = 1,12$ cm; e PbWO₄, $X_0 = 0,89$ cm; enquanto para chumbo, $X_0 = 0,56$ cm [12, 51]. Pelo fato do material ser caro, um método para redução do custo do detector é colocá-lo próximo ao ponto de colisão, o que minimiza o volume total de material envolvido. Contudo, isso significa que os calorímetros terão granularidade mais grossa por limitações mecânicas para construção das células. Com isso, sacrifica-se a capacidade de medição de posição e de discriminação do calorímetro. Deste modo, a escolha não é trivial, indo além da disponibilidade de verba. A Tabela 2.3 traz referências de resolução de alguns experimentos para comparação de calorímetros EM de amostragem e homogêneos.

Tabela 2.3: Resolução em energia de calorímetros eletromagnéticos em experimentos, com a respectiva tecnologia empregada e data do projeto. O símbolo U* representa urânio empobrecido [12, Cap. 35].

Experimento	Material e Arranjo	Profundidade	Resolução de Energia [GeV]	Data
Calorímetro Homogêneo				
Crystal Ball	NaI(Tl)	20 X_0	$2,7\%/E^{1/4}$	1983
BaBar	CsI	16-18 X_0	$2,3\%/E^{1/4} \oplus 1,4\%$	1999
CMS	PbWO ₄	25 X_0	$3\%/\sqrt{E} \oplus 0,5\% \oplus 0,2\%/E$	1997
OPAL	Cristal de Chumbo	20,5 X_0	$5\%/\sqrt{E}$	1990
NA48	Kr Líquido	27 X_0	$3,2\%/\sqrt{E} \oplus 0,42\% \oplus 0,09\%/E$	1998
Calorímetro de Amostragem				
Zeus	Cintilador/U*	20-30 X_0	$18\%/\sqrt{E}$	1988
CDF	Cintilador/Pb	18 X_0	$13,5\%/\sqrt{E}$	1988
KLOE	Fibra Cint./Pb espaguete	15 X_0	$5,7\% \oplus 0,6\%$	1995
H1	LAr/Pb	20-30 X_0	$12\% \oplus 1\%$	1998
DØ	LAr/U*	20,5 X_0	$16\% \oplus 0,3\% \oplus 0,3\%/E$	1993
ATLAS	LAr/Pb acordeão	25 X_0	$10\% \oplus 0,4\% \oplus 0,3\%/E$	1996

2.4.2.4 Discussão

Tendo em vista o que foi apresentados nesta Subseção, cabem as seguintes considerações sobre o uso de calorimetria.

Primeiro, de acordo com a Propriedade V, faz-se necessário fina granularidade para que o calorímetro possa efetuar medições de posição. Isso também é importante para maximizar a Propriedade VII, ao caracterizar fidedignamente o desenvolvimento dos chuveiros de forma a possibilitar sua discriminação entre chuveiros puramente EM ou com componentes HAD. Assim, a capacidade de discriminação de partículas na Propriedade VII é limitada a chuveiros com desenvolvimentos distintos.

Por outro lado, com o desenvolvimento do chuveiro, há um aumento de largura lateral, de forma que não há grandes ganhos com a manutenção da granularidade conforme as amostragens longitudinais se tornam mais profundas. O mesmo pode ser dito para as medições de posição, que naturalmente ocorrem nas primeiras amostragem do calorímetro, no início do desenvolvimento do chuveiro. Por isso, os calorímetros tendem a ajustar a granularidade — obtendo menor quantidade de sensores — de acordo com a profundidade das amostragens longitudinais. Em consequência, reduz-se a complexidade de construção, operação e manutenção desses detectores.

Além disso, como visto, os chuveiros HAD costumam ter maior largura lateral e serem mais irregulares lateral e longitudinalmente que os chuveiros EM. Isso significa que, ao optar pela absorção daqueles chuveiros prematuramente aos destes, perde-se o potencial discriminante entre os chuveiros em concordância com os argumentos que seguem. Do ponto de vista longitudinal: essa opção ocorreria em maior frequência de sobreposição entre os tamanhos dos chuveiros devido à irregularidade dos chuveiros HAD; e reduziria a possibilidade de explorar as ocorrências de atraso para o desenvolvimento do chuveiro HAD. Enquanto lateralmente, isso contraditaria com a natureza de desenvolvimento do chuveiro, que aumenta sua largura de acordo com a profundidade. Como os chuveiros HAD tendem a ser mais largos, é natural permiti-los propagar-se até as regiões profundas do calorímetro a fim de empregar granularidade mais grossa nessa região. Isso permite economia no número de canais sem grandes perdas na capacidade de descrição do formato do chuveiro.

Todos materiais comumente utilizados em calorímetros possuem $\lambda_{int} > X_0$ [39], havendo uma larga gama de materiais com essas propriedades — motivo para o qual um HCAL costuma ser maior que um ECAL. Exemplos para os elementos químicos puros estão disponíveis, incluindo todos os metais com isótopos estáveis, na Figura 2.12. Em posse de todas essas considerações — especialmente a distinção longitudinal dos chuveiros [46] —, é natural que os calorímetros disponham uma

configuração com o ECAL disposto anteriormente ao HCAL, o que é historicamente a configuração comumente empregada. Contudo, [46] cita a ocorrência de uma outra possibilidade para essa abordagem: como os HCAL compensados permitem a boa aquisição de ambos chuveiros, eles podem ser empregados como um calorímetro integrado, cobrindo também a detecção de elétrons, fótons, múons e, mais ambiciosamente, neutrinos (algumas considerações sobre o emprego de calorimetria para múons e neutrinos em breve). Uma das vantagens de um calorímetro integrado é não ter respostas distintas para hádrons no ECAL e HCAL, o que permitiria maior exploração da física em estados finais HAD e maior potencial de rejeição de ruído na discriminação elétron-hádron.

Todavia, o ATLAS e o CMS mantiveram a estrutura clássica, o que permite o desenvolvimento de calorímetros otimizados para a descrição de partículas EM — o CMS via excelente resolução de energia EM, enquanto o ATLAS com resolução moderada mas resposta potencialmente mais uniforme e com grande potencial de discriminação, justamente o escopo deste trabalho [47]. Não obstante, nesses casos é possível reduzir o impacto na resolução de energia para as partículas HAD através da ponderação das contribuições na segmentação longitudinal entre os dois calorímetros, como sugerido no estudo pré-Primeira Temporada de Coleta de Dados do ATLAS [52]. Com isso, evidenciam-se várias abordagens possíveis para o emprego de calorimetria, mesmo quando considerando os mesmos objetivos, que não são limitadas a abordagem tradicional o emprego conjunto de calorímetros dedicados para a contenção de chuveiros EM e HAD.

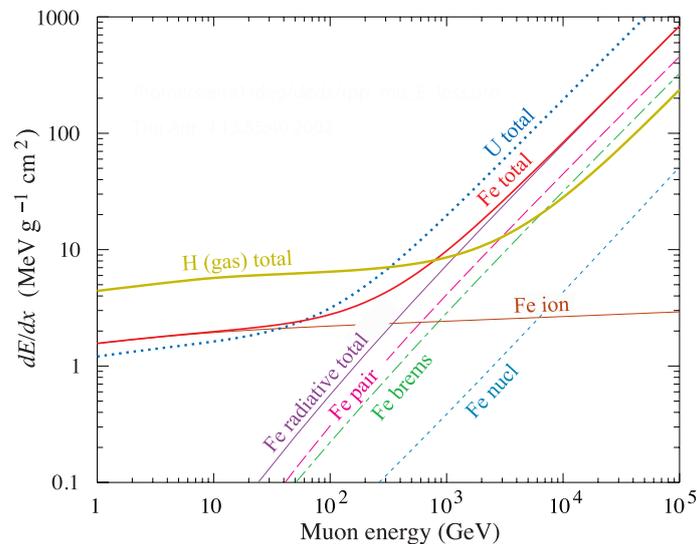


Figura 2.14: A perda média de energia de múons em hidrogênio gasoso (H total), ferro (Fe total) e urânio (U total) em função da energia do múon. No caso do ferro, as contribuições em dE/dx para as interações por ionização e radioativas — Fe pair: produção de pares; Fe brems: *bremstrahlung*; Fe nucl: fotonuclear [12, Cap. 34].

Há um aprofundamento a ser realizado quanto à Propriedade III. Dois casos peculiares de partículas interagindo com os calorímetros são os de neutrinos e múons. No caso de neutrinos, o desafio para sua detecção resulta da mínima seção de choque de interação com os núcleos do material por via da interação fraca, único meio de interação possível para essa partícula — para um neutrino de 10 GeV, a probabilidade de interação com um dos núcleos em 10 m de ferro sólido é $3,3 \times 10^{-10}$. Por isso, a estratégia para detecção de neutrinos é sempre indireta [39] e, em experimentos não dedicados a estudos específicos dessas partículas (iniciada historicamente pelo detector UA1 [53]), é feita por via de \cancel{E}_T . Nesse caso, assume-se que toda energia faltante é dada por neutrinos ao aplicar o conceito de conservação do momento transversal¹⁵ e, para isso, necessita da propriedade de hermeticidade do detector para as demais partículas, ou seja, ampla cobertura em η do detector, e conseqüentemente dos calorímetros¹⁶. Por outro lado, grande parte dos múons tipicamente produzidos nas colisões do LHC¹⁷ tem o processo de interação com os calorímetros predominado pelo processo de ionização. Um exemplo para a interação de múons com ferro, urânio e hidrogênio estão indicados na Figura 2.14. Como se observa, a E_c para múons é bastante superior a de elétrons e é melhor descrita pelo ponto em que a interação radioativa total ultrapassa a ionização, uma vez que a produção direta de pares de elétrons é predominante sobre o efeito de *bremsstrahlung*. Sem a geração de chuveiros de partículas, não ocorre grandes transferências de energia o que dificulta o emprego de calorimetria para detecção da partícula. Em termos numéricos, estimou-se durante estudo pré-Primeira Temporada de Coleta de Dados [56] que múons entre 10 e 300 GeV depositariam entre 2,3 a 2,9 GeV no *TileCal* do ATLAS (ver Subseção 4.2.3). Por isso, múons geralmente contam com seu próprio detector via espectrografia de traços, posicionado após os calorímetros.

¹⁵A conservação do momento longitudinal ao feixe não pode ser explorada devido à produção de partículas energéticas nessa direção [6].

¹⁶A detecção por calorimetria de neutrinos baseia-se na absorção do hádron gerado na interação do neutrino com o núcleo do material ativo. Dependendo da natureza do neutrino, o detector precisa ser sensível também a elétrons e múons. Ver experimentos [54, 55] para exemplos de calorimetria especializada para detecção de neutrinos por via de aceleradores (com feixes de neutrinos do SPS). [39]

¹⁷Vale ressaltar que múons mais energéticos têm grande importância para análise física, podendo advir de mecanismos em regiões de altas energias ainda não exploradas.

Capítulo 3

Levantamento Bibliográfico

Realiza-se, neste capítulo, um levantamento dos métodos aplicados na identificação de partículas em experimentos de física de altas energias através de aceleradores de partículas, focando na identificação de elétrons ou soluções envolvendo a aplicação de aprendizado de máquinas (Seção 3.1). Além disso, o principal desafio para a extensão do *NeuralRinger* para a atuação na análise em *offline* de elétrons do Experimento ATLAS é a fusão de informação de detectores com medidores de natureza distinta, de forma que as principais técnicas aplicadas no tema são levantadas na Seção 3.2.

3.1 Inteligência Computacional no Contexto de Física de Partículas de Altas Energias

A base das análises de reconstrução de eventos aplicaram, por gerações de experimentos, análises via cortes lineares. Entretanto, durante a preparação para a obtenção das colisões do maior acelerador predecessor ao LHC, realizou-se um grande esforço para a aplicação de aprendizado de máquinas na Física de Partículas [57], contando com grande participação do Fermilab na preparação de tutoriais e material introduzindo redes neurais em Física de Partículas [58]. Todavia, ao analisar a busca por elétrons em uma descoberta de grande importância na época [8, 9], não se encontra qualquer menção ao seu emprego.

Encontram-se aplicações contendo Aprendizado de Máquinas em Física de Altas Energias principalmente nos canais de busca física ou na discriminação de estados finais de jatos-b [10, 11, 59–101]. As principais técnicas aplicadas são *Boosted Decision Tree* (BDT), Rede Multicamadas de *Perceptrons* (MLP) e Rede Bayesiana. Um estudo [102] que se sobressai empregou redes neurais para identificação de fótons com base na informação de traço e depósito de energia.

Especificamente para identificação de elétrons, as primeiras propostas envol-

vendo a aplicação de redes neurais observadas envolveram as seguintes propostas: de aplicação na análise física e proposta para atuação na filtragem *online* de elétrons no $D\emptyset$ [103]; exercício considerando a separação elétron-fóton em dados do Experimento E734 [104] dedicado a medição de espalhamento elástico do neutrino do elétron. Os resultados mostram superioridade de eficiência de seleção com Rede Multicamadas de *Perceptrons* em relação às referências em todos os casos abordados. Esse estudo em elétrons desencadeou outro trabalho [105], que é a primeira referência a empregar informação anelada de calorimetria (empregada neste trabalho) que se tem conhecimento. Por consistência, as considerações sobre esse trabalho estão na Subseção 6.2.4.

O Experimento H1 [106] desenvolveu redes neurais [107–109] para a filtragem *online* em objetos físicos de interesse do experimento. A implementação foi realizada em *hardware* para a atuação no segundo nível de filtragem com $20\mu\text{s}$ de latência. O sistema operou em paralelo desde o início do experimento com outra decisão baseada em cortes lineares, complementando-a e possibilitando coletar estatística para processos de difícil análise, onde é citado o caso de processos de produção de mésons.

3.2 Fusão de Informação

O campo de Fusão de Informação é muito amplo e abarca uma grande variedade de problemas específicos que fogem do escopo deste trabalho. Ainda assim, para uma compreensão mais ampla da área, o levantamento bibliográfico estendeu-se para atender o surgimento histórico e as técnicas tradicionalmente empregadas da área (Subseção 3.2.1). Aborda-se uma aplicação promissora para esta área com diversos desafios em comum na Subseção 3.2.2. A revisão então se dirige à aplicação de técnicas de Inteligência Computacional neste campo (Subseção 3.2.3), segregando as mesmas conforme técnicas tradicionais (Tópico 3.2.3.1) e Aprendizado Profundo (Tópico 3.2.3.2).

3.2.1 Técnicas Tradicionais

Os termos *fusão de informação*¹ e *fusão de dados* são geralmente empregados para determinar o mesmo fim. A primeira definição (1985) para fusão de dados deu-se por [110, tradução própria] “um processo que lida com a associação, correlação, combinação de dados e informação de fontes únicas ou múltiplas para alcançar posição aperfeiçoada e identificar estimativas, e avaliação completa e oportuna de situações, e ameaças e sua significância. O processo é caracterizado pelo

¹O termo multisensor muitas vezes é encontrado nos títulos de trabalhos da área.

aperfeiçoamento contínuo das estimativas, avaliações e pela avaliação da necessidade de fontes adicionais, ou modificações do próprio processo, para obter resultados melhorados.”. Essa definição, proveniente de um léxico da marinha dos Estados Unidos, revela as origens (por volta de 1970 [111]) e correlação do campo com aplicação militar. De fato, o primeiro modelo e ainda empregado por alguns autores atualmente, *Joint Directors of Laboratories (JDL)*, leva o nome do grupo administrativo que existia para auxiliar na coordenação da pesquisa entre múltiplos laboratórios de defesa daquele país [112]. A sua especificidade dá-se pela construção de níveis hierárquicos que se comunicam de maneira específica através de funções claramente descritas, resultando em um método com uma estrutura complexa e ampla diversidade de conceitos, possivelmente seguindo a tendência de sua origem militar. Uma revisão do desenvolvimento do modelo JDL inicialmente proposto [113] encontra-se em [114], onde se observa o progresso com a introdução de níveis adicionais, novas categorias, extensões e variantes² focando em completar as lacunas de uma proposta focada na aplicação militar, de modo a atender outros casos. Essas modificações, por sua vez, discutem a questão da exploração e gestão da informação [115–117], considerando especialmente os aspectos da fusão de informação de alto nível [110, 118], ou seja, informação abstrata e simbólica como: ameaça, intenção, objetivo e interpretação de situação. As necessidades em torno de estender a aplicação militar para outros âmbitos exigiu, também, uma definição mais ampla, onde, geralmente, emprega-se “o processo de combinar dados ou informações para estimar ou prever estados de entidade” [119, tradução própria].

Atualmente, a abordagem ao problema de fusão da informação tende a concentrar-se na análise das técnicas dedicadas à solução dos problemas envolvidos, ao invés de tentar modelar o problema em grupos hierárquicos, como no caso do JDL [120]. O levantamento aqui realizado foca no contexto de identificação de eventos de interesse na área de Fusão de Informação, dando menor importância a outros problemas importantes nesta área, como rastreamento de objetos.

A abordagem clássica para os tópicos de Fusão de Informação dá-se por estratégia, probabilísticas com grande variedade [120–122]. A mais recente revisão da literatura [120] conclui que essas técnicas são mais simples e altamente aceitas no campo; porém, citam-se como desvantagens: dificuldade para obtenção das *prioris*;

²Observou-se um fato curioso enquanto lendo autores da época envolvidos na aplicação desta técnica. As referências notam que o desenvolvimento do campo iniciou-se na ponta dos sensores para progredir em direção à ponta da interação humana. Boa parte da discussão, durante a década de 2000, deu-se em torno da adição do usuário, em determinadas ocasiões, inclusive, para lidar com as limitações das técnicas de Aprendizado de Máquinas para tomada de decisão [110]. A referência menciona que o campo se iniciou na “ponta errada” em uma seção (com o título original de um artigo usado para a primeira edição do livro) nomeada de “Segredos Obscuros da Fusão de Dados” (tradução própria), o que evidencia o esforço por parte dos autores para lidar com a presença de usuários, geralmente como um supervisor do sistema. Não se pode negar que, com o surgimento da IoT, a ponta humana torna-se o centro da aplicação; porém, por motivo distinto.

as técnicas podem resultar em baixa acurácia; não há capacidade para lidar com incerteza; no caso de classificação, normalmente as classes são mutualmente exclusivas (não pode ocorrer simultaneamente observação de classes distintas); e saída sensível à presença de ruído. Uma solução mais eficiente à abordagem *bayesiana* dá-se pelo emprego do método de Dempster-Shaffer, baseado na Teoria de Evidência proposta pelo primeiro e formalizada matematicamente pelo segundo. Esse método resolve parte dos problemas citados ao considerar hipótese de solução mutualmente exclusiva, não necessitar a obtenção de *a priori* e obter uma representação explícita de incerteza [120, 121]. Aplicações dessa técnica para fusão de informação vêm de longa data [123].

Dentre as estratégias probabilísticas, destacou-se uma abordagem híbrida com a obtenção das *a priori* via redes neurais [124]. Neste caso, o estudo avaliou a identificação de três interlocutores com informação de áudio e vídeo e pronunciando o conjunto de letras $\{‘I’, ‘O’, ‘U’\}$, com 10 observações para o treinamento e 45 para teste. A rede neural é responsável pelo processamento da imagem, enquanto o áudio é processado por Modelo Oculto de *Markov* (HMM). O sistema de fusão é Bayesiano e tem como entrada o módulo de áudio, imagem e um terceiro item proveniente de um modelo VQ-HMM, representando a probabilidade do interlocutor ter pronunciado uma determinada letra. Quando os dados são avaliados considerando apenas as informações isoladas, ocorrem quatro erros no conjunto de teste tanto em áudio quanto em vídeo, a versão convencional com probabilidades estimadas obtém um (1) erro e nenhum erro ocorre para o método proposto.

3.2.2 Das Demandas da *Internet* das Coisas

Uma larga gama de aplicações podem ser atendidas empregando abordagens clássicas (incluindo modelos hierárquicos como JDL) mencionadas da subseção anterior, dentre elas [111]: vigilância, monitoração de tráfego (não se limitando ao terrestre), robótica, medicina, sentimento remoto/de imagem, transporte inteligente, biometria etc. Todavia, há grande movimentação no campo envolvendo aplicações destinadas a *Internet* das Coisas (IoT), sendo esperado um grande impacto dos avanços tecnológicos dos tempos atuais [120]. As técnicas desenvolvidas neste contexto estão mais alinhadas com o deste trabalho, uma vez que estima-se que a IoT seja uma das maiores produtoras de grande massa de dados no futuro próximo, esperando-se 50 bilhões de dispositivos conectados a ela ainda em 2020 [125]. Esses dispositivos, de diversos fornecedores, serão responsáveis pela produção dos dados de forma que o compartilhamento e colaboração de dados e outros recursos seria a chave para possibilitar ambientes sustentáveis e ubíquos como cidades e sociedades inteligentes. Assim, os dados obtidos pelos fornecedores são vistos como um serviço

que possibilita explorar por completo as benesses da IoT [126].

Para que a IoT seja possível, um dos componentes principais é a fusão da informação dos inúmeros sensores de distinta natureza, dando assim a importância do enfoque dado à aplicação nessa área. Por sua vez, esse é um desafio para a área que espera empregar técnicas de Inteligência Computacional, a fim de superar as demandas de um sistema confiável, com alta eficiência, provendo processamento *online* em um ambiente disponibilizando grande massa de dados para o ajuste das técnicas. Assim, fica evidente a correlação do contexto de aplicação de fusão de dados em IoT com o deste trabalho, especialmente ao considerar a analogia da informação obtida pelo Detector Interno (ID) com imagens, informação explorada na fusão de dados desde as suas primeiras aplicações.

Destacam-se outros desafios [110, 127] em IoT de interesse para o contexto deste trabalho:

- para disponibilização dos serviços de IoT, necessita-se que os sensores não tenham reposição de bateria durante seu tempo de vida. Com isso, populariza-se a demanda por sensores energeticamente eficientes. Dado que sensores de alta acurácia podem requisitar alta demanda de energia, a estratégia esperada é a aplicação de diversos sensores com menor precisão, evidenciando a importância da fusão de informação. Assim, possibilita-se realizar um paralelo entre os 100 milhões de canais do ID e 190 mil provenientes do Sistema de Calorimetria do ATLAS com o cenário esperado de IoT (ainda que os sensores do ATLAS sejam de alta precisão);
- a fusão de informação facilita lidar com grande massas de dados ao resumizá-los a partir de um grande número de sensores em informação acurada e precisa. De fato, esse é justamente o caso de aplicação deste trabalho, onde seria inviável (por limitação de armazenamento e, possivelmente, de processamento), lidar atualmente com toda a dimensão proveniente dos sensores crus do ATLAS. Os processamentos especialistas para geração das grandezas físicas (Seção 5.2) e o processamento anelado (Capítulo 6) procedem como o mencionado. Independente desses processamentos, essa estratégia está alinhada com a solução empregada neste trabalho com o uso de redes especialistas (Seção 7.3);
- os dados obtidos pelos sensores em IoT estão sujeitos a imprecisão repentina, podendo haver imprecisão ou incerteza relacionada ao valor obtido. Este é um caso conhecido em Redes de Sensores *Wireless* (WSN), de forma que essa imprecisão precisa ser levada em conta para o uso eficiente de algoritmos de fusão. No contexto deste trabalho, o mesmo efeito ocorre ao observar a não homogeneidade de resposta do detector em relação a posição de incidência da

partícula (seja pela quantidade de sensores disponíveis ou por sua qualidade), a alteração nos padrões conforme a energia e distorção de acordo com o empilhamento de sinais. Todavia, as variáveis relacionadas a essas distorções estão disponíveis para aplicação de soluções parametrizadas ou modelos específicos como a correção linear e *ensembles* baseados nessas variáveis, o que pode não ser o caso das aplicações de IoT;

- correlação dos dados e alinhamento, também comumente presente em WSN, que pode causar em uma sobre/subconfiança em um dos métodos usados para a fusão de dados. Nas aplicações específicas de IoT, o problema de alinhamento é também conhecido como um problema de registro e ocorre quando os dados dos sensores são transformados do plano local contendo cada um dos sensores para um plano comum anteriormente a fusão. Aqui, cabe ressaltar que tanto as medições de energia das células dos calorímetros quanto o *erro* sofrem correlação conforme a posição, resultando em um problema similar;
- parcela considerável das centenas ou milhares de sensores esperados em aplicações IoT podem gerar padrões redundantes³, que se empregados podem afetar a acurácia do método de fusão de dados. Assim, apenas os padrões relevantes devem ser selecionados para participar do processo de fusão. A análise de relevância dos padrões é igualmente importante no contexto deste trabalho, especialmente para a soma em anéis de energia;
- até o momento, não se tem conhecimento de uma abordagem ótima para o problema de fusão de informação, ainda que grandes avanços tenham sido obtidos com algoritmos com eficiência ótima.

3.2.3 Inteligência Computacional Aplicada a Fusão de Dados

Inteligência computacional aplicada a fusão de informação é um dos focos deste trabalho. Por isso, dividiu-se esta subseção em dois tópicos contendo técnicas tradicionais de Inteligência Computacional (Tópico 3.2.3.1) e seu recente subcampo (Tópico 3.2.3.2).

3.2.3.1 Técnicas Tradicionais

Técnicas já aplicadas na fusão de dados relacionadas a inteligência computacional envolvem [120] técnicas de Aprendizado de Máquinas como agrupamento [128], Rede

³Na referência, mencionam-se *padrões triviais* (*trivial features*). Preferiu-se traduzir conforme apresentado por adequar-se melhor ao contexto.

Neural Artificial (RNA) [128], Mapas Auto-Organizáveis (SOM) [128, 129], Máquina de Vetores de Suporte (SVM) [130]; Lógica *Fuzzy* [131]; e Estratégia Evolutiva [132].

No caso de SVM [130], uma de suas primeiras aplicações na área deu-se para a superação do problema do cálculo de *priori* em lógica Bayesiana [133] em aplicações distribuídas. Neste caso, as limitações de transmissão de dados impediam o envio completo da informação, de forma que se empregou o SVM para a compressão da informação em sua representação em *kernel*. Uma aplicação mais recente [134] envolvendo um sistema híbrido de SVM e Naïve Bayes para sensoriamento remoto envolvendo a fusão de Detecção de Luz e Alcance (LIDAR) e informação hiperespectral. LIDAR é uma técnica de sensoriamento remoto que determina a distância ao iluminar o alvo com um *laser* e examinar a luz refletida. Em cada conjunto, emprega-se um classificador SVM dedicado para então realizar a fusão de suas classificações via Naïve Bayes. A estratégia igualou ou superou os resultados dos classificadores isolados para todos os tipos de terreno e considerando dois critérios de medidas.

Em dados hiperespectrais, a aplicação de estratégias de *Random Florest* [135] tem sido mais eficiente [136] que SVM. Motivado por isso, no trabalho [136], empregou-se esta técnica para fusão de LIDAR e hiperespectral. Uma das vantagens do método é lidar com alta dimensionalidade e fornecer ordenamento da relevância dos padrões empregados. Observou-se que apenas metade dos 80 padrões auxiliavam na classificação, de forma que os mesmo foram removidos e possibilitaram a superação de resultados anteriores também aplicando a mesma técnica.

Avaliou-se a aplicação de k-Vizinhos Mais Próximos (KNN), Análise de Discriminante Linear (LDA) e Análise de Discriminante Quadrática (QDA) em [137] para fusão *online* de padrões para monitoração de tráfego. Os padrões envolviam medição de distância por *laser* e perfil de amplitude, sensor acústico, vídeo e sensoriamento de ambiente fornecendo temperatura, ruído acústico e incidência luminosa. As conclusões mostraram que as técnicas de LDA e QDA atendem a demanda de processamento *online*. Por outro lado, KNN não foi capaz de atender essas demandas por necessitar de grande quantidade de amostras para obter eficiência satisfatória, o que inviabilizaria a aplicação.

Empregam-se RNAs eficientemente para predição de longo e curto prazo da velocidade de vento [138–142]. A necessidade dessa aplicação auxilia no projeto ótimo de fazendas de vento, na gerência de energia e mercados de eletricidade reestruturados [138]. No caso de [138], realiza-se a fusão da saída de RNAs e abordagem especialista em sequências temporais de predição de diferentes intervalos (anual, mensal, diário, horário) a fim de obter previsão horária completa para o ano seguinte. A primeira rede (anual) operante emprega arquitetura Rede Não-Linear Autorregressiva com Entradas Exógenas (NARX) com duas camadas escondidas, usando como entrada períodos mensais de vento, e realiza a predição de 4 anos seguintes e com-

para o resultado com os 3 anos anteriores medidos por Erro Absoluto Médio (MAE), repetindo o procedimento de ajuste até que o valor seja inferior a um limiar. Duas MLP de duas camadas são empregadas para a predição mensal, cada uma se especializando para um semestre do ano, através de intervalos históricos mensais. Para predição diária, emprega-se abordagem especialista, enquanto a predição horária emprega outra MLP de duas camadas. A acurácia obtida é de 0.9 ms medidos por MAE. Em [139], empregou-se Redes Neurais Probabilísticas (PNNs) para classificar e abrigar dados crus de velocidade de vento nas turbinas de vento. Esses dados são empregados para a seleção de sinais de entrada para o modelo atuante final baseado em Rede Neural Recorrente de Valores Complexos (CRNN). O erro obtido avaliado por Erro Absoluto Percentual Médio (MAPE) é de 11,2 %.

Em predição de tempo de viagem, empregou-se MLP para predição de velocidade futura através de dados provenientes de Sensor Remoto de Tráfego por Micro-onda (RTMS) [143], enquanto que, para a fusão dos diversos sensores disponíveis, utilizaram-se duas estratégias especialistas. O Erro Quadrático Médio (MSE) obtido para ambos os métodos foi inferior a 10 %.

Múltiplas MLP em *hardware* foram empregadas por [144] para localização em ambiente interno via *bluetooth* empregando Indicador de Força de Sinal de Rádio (RSSI) e RF. Cada rede considerou a estimativa de posição para uma determinada rota motivado pela sensibilidade das medições de RSSI de acordo com a direção. O sistema inclui uma lógica heurística para navegação e obteve eficiência de 90 % de detecção da posição com 0,5 m de precisão.

3.2.3.2 Aprendizado Profundo

O ramo de Aprendizado Profundo (DL) [145] tem ganhado atenção significativa em diversas áreas, estando entre as maiores 10 tendências de 2016 pela Gartner [146]. Um dos fundadores da área, criador das Rede Neural Convolutiva (CNN) e atualmente diretor de Inteligência Artificial na empresa *Facebook* afirma que o DL encontrará sucesso em diversas áreas principalmente por dois fatores [147]: as técnicas são flexíveis e necessitam de pouca engenharia braçal e beneficiam-se intrinsecamente conforme a disponibilidade de recursos computacionais e dados. As aplicações com DL têm, aos poucos, encontrado o campo de fusão de dados.

Em [148], fusão de informação empregando duas RNAs de arquitetura distintas foram aplicadas para identificação de Objetos Independentes em Locomoção (Objeto Independente em Locomoção), com aplicação em automóveis autônomos. Uma rede MLP com uma camada escondida dedicou-se à construção de um mapa de verossimilhança para cada pixel da imagem a fim de identificar regiões candidatas a conter um IMO. Para realizar essa tarefa, empregam-se como padrões discriminante a disparidade estéreo (atual, anterior e do próximo *frame*), fluxo óptico e

coordenadas normalizadas do pixel. O reconhecimento dos veículos emprega uma rede convolutiva *LeNet* [149] com seis camadas. As saídas são seis mapas de verossimilhança contendo as classes esperadas, com as quais se constrói outro mapa responsável pela classe de ruído (sinalizando que não há um IMO). A fusão dos mapas provenientes das duas redes é realizada por técnica especialista. Por não haver alvos de referência para avaliar o modelo, os resultados são disponibilizados qualitativamente, onde o sistema demonstrou habilidade de encontrar, seguir e identificar IMO.

O problema de aprendizado multimodal foi abordado com DL em [150]. O objetivo é relacionar informações de fontes múltiplas que, no caso da aplicação, deram-se em obter as correlações de vídeo focando na movimentação de lábios e o áudio correspondente à fala. A tarefa é dividida em três etapas: aprendizado de padrões, treinamento supervisionado e teste. Três configurações de aprendizado foram consideradas: fusão multimodal, aprendizado de modalidade cruzada e aprendizado conjunto de representação. Em todos os casos, o aprendizado dos padrões emprega ambas modalidades, porém há distinção para a informação da informação disponível para o treinamento e teste: para este a informação de ambas modalidades está disponível em todas etapas; esse recebe uma modalidade no treino e outra no teste e; aquele emprega uma única modalidade para treino e teste. O objetivo do aprendizado cruzado de modalidade é verificar se os padrões conseguem capturar correlações entre as modalidades e, no caso do aprendizado conjunto, avaliar se a apresentação de uma representação conjunta consegue obter representações modais superiores a quando empregando a abordagem clássica.

O aprendizado da informação multimodal é realizado treinando duas *Restricted Boltzmann Machine* (RBM) [145] separadas para cada uma das modalidades de dados. Quando avaliando o treino de uma RBM multimodal, notou-se que poucos pesos internos tinham conexões entre ambas modalidades. Assim, a união é realizada através de *Deep Belief Networks* (DBN) a partir das camadas escondidas da RBM, conforme sugerido por [151], onde o treinamento é realizado de forma que o *autoencoder* seja capaz de reconstruir ambas modalidades. Quanto a parte do aprendizado, emprega a seguinte lógica para o caso de aprendizado de modalidade cruzada: dado que a otimização DBN não em seu objetivo a descoberta de correlações entre as modalidades, isso é realizado ao forçar no treinamento, inspirado por *denoising autoencoders*, empregar 1/3 do conjunto com valores zeros para as variáveis da primeira modalidade, outro 1/3 com zeros para as variáveis da segunda e o restante apresentando ambas modalidades. Já para os demais casos (treinamento único ou bimodal), o procedimento é trivial.

Para o treinamento de modalidade cruzada, os resultados para teste em vídeo mostram que o aprendizado bimodal, seja por treinamento único (64,4 %) ou cruzado

(59,2 %), superam os resultados do RBM (54,2 %) empregando vídeo e a referência (58,8 %) também foi superada. Em outro conjunto de dados o treinamento cruzado obteve resultados próximo ao DBN com treinamento único. A vantagem do treinamento cruzado é permitir o emprego do DBN em ambas modalidades, ainda que com um custo em performance quando comparando com o treinamento único. Cabe notar que os resultados em aprendizado único são interessantes para serem considerados caso se empregue DBN no *Ringer online*, ou seja, é possível que a aplicação de DBN empregando ambas modalidades de calorimetria e traço possibilite a obtenção de uma estratégia mais potente ainda que limitada à aplicação somente com calorimetria.

Outro resultado interessante dá-se pela concatenação dos padrões obtidos da RBM treinada somente em áudio para o modelo de operação nesta modalidade, atingindo uma performance superior a quaisquer outras configurações avaliadas quando adicionando ruído. O RBM apenas com áudio obtém a melhor eficiência (95,8 %) quando sem adição de ruído, seguido pela versão multimodal adicionado de RBM treinado unicamente por áudio (94,4 %).

No trabalho [152], avalia-se o emprego de CNN para fusão de dados multimodais combinando sinais contínuos e eventos discretos em base de dados contendo eventos em jogos e sinais fisiológicos. A aplicação direta dos eventos discretos na CNN piorou a eficiência do método em relação a modalidade contínua. Outra abordagem procurou adicionar uma CNN de única camada para discriminar os efeitos de eventos em sinais contínuos, porém essa abordagem também não obteve êxito. A solução encontrada deu-se pela aplicação de uma camada de votação que filtra a saída da CNN conectada ao sinal contínuo em função do acontecimento de eventos em uma determinada janela (no trabalho, 5, 10 e 15 s), resultando em um ganho de eficiência. Essa estratégia pode ainda obter melhor acurácia ao combiná-la com métodos tradicionais de votação por média.

O artigo [153] mostra que sensoriamento remoto pode ser empregado ao generalizar estruturas *encoder-decoder* do tipo *SegNet* [154] treinadas em conjuntos de imagens genéricos. Outras aplicações envolvendo DL incluem [120]: monitoração de falhas, reconhecimento em tempo real de dígitos, reconhecimento de atividade, fusão de dados para tráfego de rede e detecção de pedestres.

Todavia, DL contém seus problemas. O mais conhecido é o alto consumo de recursos de processamento e memória, havendo todo um novo mercado de *hardware* para atender suas necessidades de aplicação (placas de vídeo específicas para essa aplicação). O processo de *hardware* tendem a decrescer no futuro de forma que essa demanda pode ser desprezada. Quando considerando a aplicação em IoT, por sua vez, deve-se considerar que o processamento da fusão de dados pode ocorrer em dispositivos remotos. Além disso, DL é sujeito a presença de dados incoerentes, que

farão parte das aplicações de IoT. Assim, o desenvolvimento de DL para aplicações específicas de IoT ainda são uma questão em aberto.

Capítulo 4

O Experimento ATLAS

O Experimento ATLAS [31, 52, 155] (Figura 4.1) é o maior dos detectores que operam no LHC, medindo 45 metros de comprimento e 25 metros de altura e largura. Ele é um detector de propósito geral, registrando dados sobre os eventos de colisões de partículas que podem ser usados para estudos em diversas áreas da física. Cada um de seus subsistemas teve os componentes construídos por um grupo diferente pertencente às instituições colaboradoras. O Experimento ATLAS envolve cerca de 3.000 físicos de mais de 182 instituições e laboratórios de 38 países [28]. Esses números incluem a UFRJ, que participa através da COPPE/UFRJ, da Escola Politécnica e do Instituto de Física.

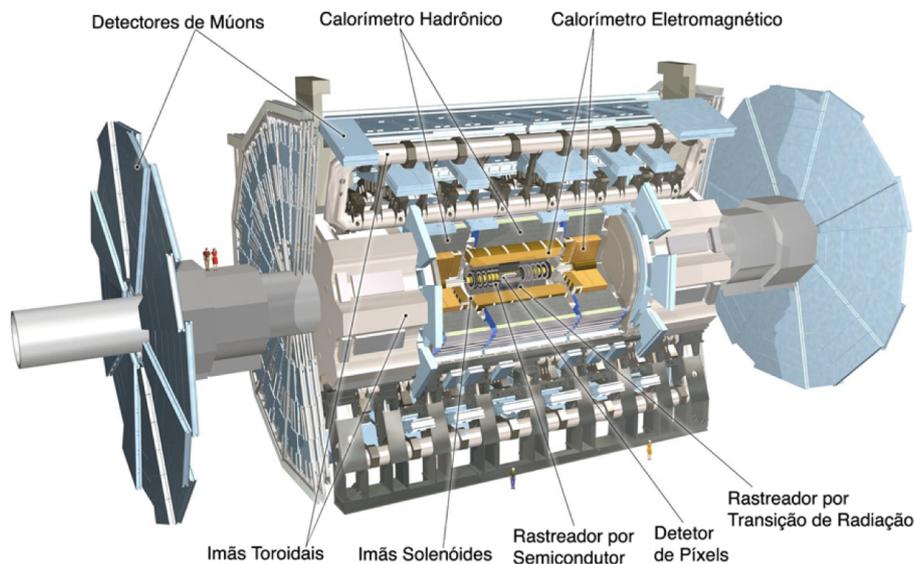


Figura 4.1: O Experimento ATLAS e seus principais sistemas de detectores. [156].

Constituindo o ATLAS estão três principais sistemas de detectores, ordenados respectivamente e radialmente a partir da origem no IP: ID¹, Sistema de Calorimetria e Espectrômetro de Múons. Resumidamente, o ID possibilita a identificação da tra-

¹O ID também é referido como Detector de Traços.

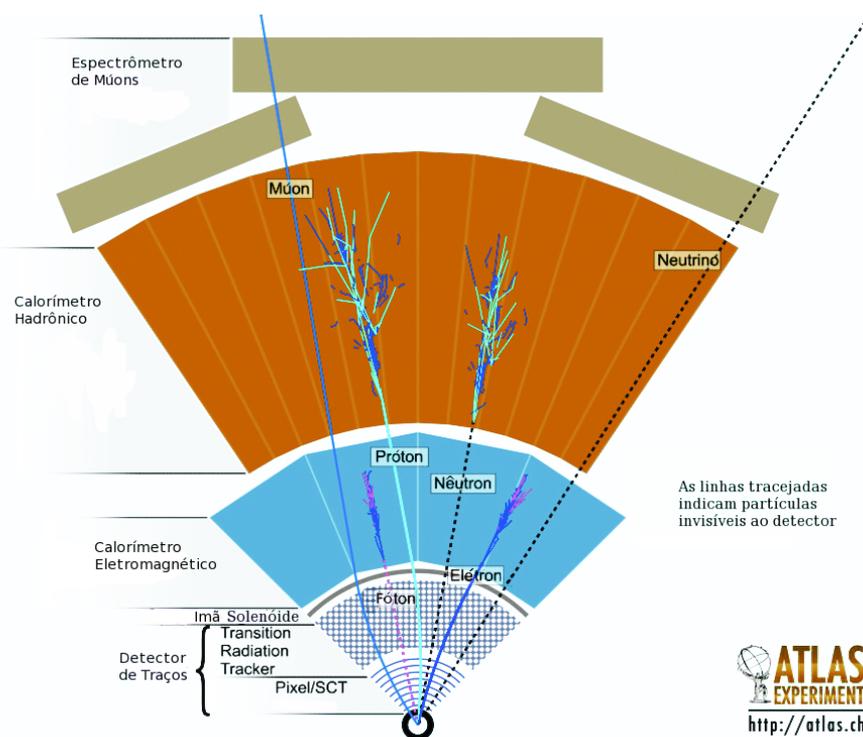


Figura 4.2: Um esboço contendo exemplos de interação ideal de partículas com os subsistemas do ATLAS. Apenas partículas carregadas eletricamente deixam traços no Detector Interno. Elétrons, pósitrons e fótons devem ser totalmente absorvidos pelo Calorímetro Eletromagnético (ECAL). Ao Calorímetro Hadrônico (HCAL), cabe a tarefa de absorver partículas com componentes hadrônicas, como nêutrons, prótons e outros mésons. Múons, mesmo contendo componentes eletromagnéticas, atravessam os calorímetros com facilidade, possuindo um subdetector dedicado — o Espectrômetro de Múons. Neutrinos não são detectados por nenhum dos subsistemas do ATLAS, e por esse motivo é necessário hermeticidade para encontrar falhas na conservação de momento indicando a presença de uma dessas partículas [157].

jetória e o momento de partículas carregadas, o Sistema de Calorimetria possibilita medições de energia através da absorção das partículas, enquanto o Espectrômetro de Múons é responsável exclusivamente pela detecção de múons obtendo informação de natureza idêntica àquela do ID, mas posterior ao Sistema de Calorimetria. Através das assinaturas feitas pelas partículas nesses subsistemas é possível realizar a reconstrução e identificação das mesmas. A Figura 4.2 contém um esboço de como as partículas interagem idealmente com esses subsistemas. A descrição aqui realizada do ATLAS enfoca nos sistemas fornecendo o núcleo de informação para reconstrução e, especificamente, a identificação de elétrons, constituídos do ID (Seção 4.1) e do Sistema de Calorimetria (Seção 4.2). Outros subsistemas são abordados menos detalhadamente na Seção 4.5.

No Capítulo 2, observou-se que o LHC opera em alta taxa de eventos a fim de permitir a observação dos processos raros dentre os objetivos dos experimentos, o que envolve a geração de uma grande massa de dados. Processam-se os resíduos das

colisões, constituídos em grande parte de física ordinária, em busca de eventos de interesse através do Sistema de Filtragem do ATLAS (Seção 4.4). A filtragem realizada por esse sistema possibilita obter taxa de eventos compatível com a capacidade de armazenamento e posterior processamento *offline* enquanto mantendo os eventos de interesse. Ainda assim, a informação resultante constitui-se de uma grande massa de dados, por exemplo, armazenaram-se 5,5 PB de informação crua dos sensores do ATLAS para o ano de 2017. O processamento *offline* dessa informação é realizado pelo Sistema de Reconstrução (*Offline*) de Física do ATLAS (Seção 4.3), que possui um sistema computacional próprio desenvolvido colaborativamente por milhares de especialistas durante os 25 anos do experimento e processado pela maior estrutura de computação em grade do mundo, a *LHC Computer Grid Project* (LCG). Em geral, a reconstrução da física empregada no Sistema de Filtragem do ATLAS é espelhada naquela disponível no Sistema de Reconstrução (*Offline*) de Física do ATLAS, porém com as degradações necessárias para que os métodos atendam as demandas de latência para operação *online*.

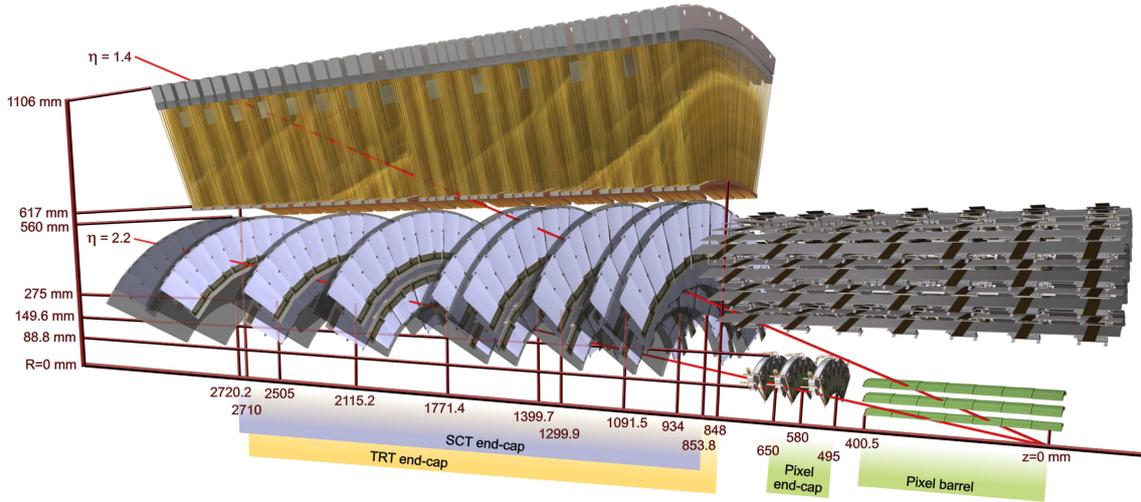
4.1 Detector de Traços

O ID [52, 158–161] (Figura 4.3) responsabiliza-se pela identificação das trajetórias (rastreamento) de partículas carregadas, medir o seu momento, o vértice primário (origem da colisão $p-p$), vértices secundários, parâmetro de impacto, e a distância entre o vértice primário e o ponto mais próximo do traço (d_0). Sua cobertura é simétrica em ϕ para $|\eta| < 2,5$ — região de precisão do detector. A limitação a essa faixa motiva-se na capacidade de resistência do ID à radiação, restringindo-se a região com incidência tolerável de radiação.

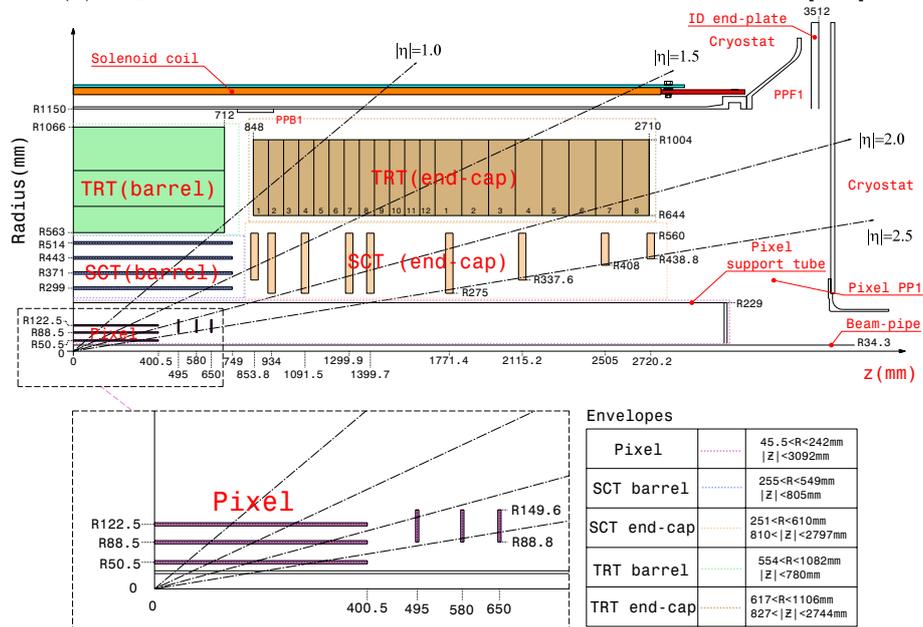
Realiza-se a medição do momento da deflexão das partículas pelo campo magnético (B) e sua resolução depende da resolução de posição (σ_p), da magnitude de B e do comprimento (l) do rastreador de acordo com

$$\frac{\delta p_T}{p_T} = \frac{p_T \sigma_p}{0,3Bl^2} [\text{GeV}, \text{T}, \text{m}]. \quad (4.1)$$

O campo magnético fornecido pelo Solenoide Central (CS) corresponde ao melhor compromisso possível (ver Subseção 4.1.4), restando dois parâmetros a serem explorados, o σ_p e l . A resolução de vértices é proporcional à capacidade de reconstrução da trajetória da partícula e, assim como o momento, é inversamente proporcional σ_p . Também é necessário uma baixa granularidade pois há uma grande densidade de traços gerados no LHC — supondo-se $\langle \mu \rangle = 19$ colisões médias ocorre a geração de ~ 10.000 traços no detector no intervalo de 100 ns [155]. Finalmente, deve evitar-se grande quantidades de material no ID para que seja possível realizar medições



(a) Seção de corte do Detector Interno e seus subsistemas [155].



(b) Esboço do Detector Interno e suas dimensões em função de η [155].

Figura 4.3: Esquema do Detector Interno, representado sem a *IBL*.

adiante dos calorímetros para minimizar a deposição de energia nos mesmos ou prejudicar o perfil dos chuveiros através da geração de chuveiros prematuros ao Sistema de Calorimetria do ATLAS.

O ID dispõe de dois medidores de alta precisão — Detector de Pixel e Detector de Rastreamento por Semicondutores (SCT) — que exploram a redução de σ_p , melhorando, assim, a resolução do detector. Por outro lado, o Detector de Rastreamento por Transição de Radiação (TRT) explora medição contínua do traço da partícula de forma a maximizar o l , resultando em melhor resolução de energia total. O TRT é especialmente importante para a identificação de elétrons por permitir discriminação elétron-píon. A resolução almejada durante a época de seu projeto para medição de momento é dada por

$$\sigma_{p_T}/p_T = 0,05 \% \times p_T \oplus 1 \%. \quad (4.2)$$

Atinge-se maior granularidade pelos pixels semicondutores, que estão instalados próximos ao vértice da colisão. Entretanto, o número de camadas de precisão são limitadas devido à grande quantidade de material introduzida — e assim evitar o processo de geração prematura de chuviscos, prejudicando o Sistema de Calorimetria — e ao seu alto custo. Dessa forma, o número de pontos obtidos com alta precisão são limitados. Tipicamente, uma partícula carregada ao transpassar seu material gera cerca de 41 pontos de posição na respectiva ordem de interação: quatro no Detector de Pixel (Subseção 4.1.1), oito no SCT (Subseção 4.1.2) e outros 30 no TRT (Subseção 4.1.3).

4.1.1 Detector de Pixel

O Detector de Pixel consiste de três camadas de detectores por sensores de pixel para fornecer cobertura da região $|\eta| < 2,5$. O tamanho de cada sensor é dado por $50 \times 400 \mu\text{m}$. Um total de 1744 sensores de pixel fornecem 92 milhões de canais de leitura dos quais 97,8 % [162] encontram-se operantes. Cada pixel fornece σ_p de $10 \mu\text{m}$ no plano $\eta \times \phi$ e $115 \mu\text{m}$ no eixo z . Sua três camadas encontram-se em uma distância radial média de 50,5, 88,5 e 122,5 mm. A carga coletada no sensor de pixel é obtida por medição da altura de pulso empregando a técnica Tempo-sobre-Limiar (ToT) [163].

Esse detector determina a resolução do ver Tabela 5.1 (d_0) e a habilidade do ID de encontrar partículas de curta vida como mésons-b e táons. Para a *Run 2*, adicionou-se a Camada-b Interna Removível (*IBL*) [164] em uma distância radial média de 33 mm, conferindo assim 4 pontos totais para o Detector de Pixels através de um ponto adicional mais próximo ao IP. Essa camada confere ao ATLAS uma melhor resolução para a obtenção de vértices e discriminação de mésons-b.

4.1.2 Detector de Rastreamento por Semicondutores (SCT)

Projetou-se o SCT para fornecer oito pontos de precisão por traço na região intermediária do ID, contribuindo tanto para a medição do momento, parâmetro de impacto e posição do vértice primário, quanto provendo bom reconhecimento de padrões pelo uso de alta granularidade. O seu barril possui quatro camadas duplas de microfibras de silicone. De um lado, as tiras são paralelas ao eixo z (direção do feixe) e um ângulo estéreo de 40 mrad para formar o outro lado. Cada componente mede $6,36 \times 6,40 \text{ cm}^2$ com 768 canais de leitura com $80 \mu\text{m}$ de afastamento. Conectam-se os leitores do SCT com *chips* de leitura binária, que não fornecem

informação sobre a carga coletada. A tampa possui construção similar, mas usa tiras cilíndricas, formando um conjunto alinhado radialmente. Ela possui cerca de 61 m^2 de detectores de silicone, com 6,3 M de canais (98,7 % disponíveis [162]), que permite distinguir dois traços separados por apenas $\sim 200\ \mu\text{m}$.

4.1.3 Detector de Rastreamento por Transição de Radiação (TRT)

Por outro lado, o TRT explora a medição contínua do traço com uma quantidade menor de material para a região $|\eta| < 2,0$. Uma partícula, ao percorrê-lo, fornece cerca de 30 pontos por traço, melhorando assim a resolução do momento ao explorar uma maior profundidade. Com o acréscimo desses pontos, o ID obtém um poder de curvatura (Bl^2) de $\sim 2\text{Tm}^2$. A coleta de sinal provém de tubos cilíndricos de deriva (do inglês, *cylindrical drift tubes*), que operam nas grandes taxas do LHC em virtude de seus pequenos diâmetros e do isolamento de seus sensores dentro de volumes individuais de gás. O barril contém cerca de 50 mil canais, divididos em dois no seu centro para reduzir a ocupação e sua leitura em cada ponta. A tampa contém 300 sensores radiais, com a leitura na ponta externa. As medições realizadas pelos canais são em formas de impulsos e dão uma resolução espacial de $170\ \mu\text{m}$ por canudo. Em 2017, 97,2 % dos canais de leitura encontravam-se operantes.

Esse detector possibilita identificação elétron-píon através da avaliação da energia depositada em seus tubos. Na *Run 1*, a composição da mistura de gás contida nos tubos era 70% Xe + 27% CO₂ + 3% O₂, o que permitiu ótima discriminação elétron-píon. Porém, com o vazamento do gás nos tubos e o alto preço do xenônio, um gás nobre raro na natureza, alterou-se para uma mistura com LAr, com menor potencial para discriminação. O ATLAS suplantou essa perda na resposta através da adição de uma nova grandeza discriminante (eProbabilityHT).

4.1.4 Solenoide Central (CS)

Um outro componente do ATLAS é o Sistema Supercondutor Magnético cujas dimensões totais são 26 m de comprimento e 20 m de diâmetro. Ele fornece, através do CS, um campo magnético central de 2 T com picos de 2,6 T no material supercondutor que possibilita a medição do momento de partículas carregadas pelo ID. Como o mesmo está localizado dentro do ECAL, necessita-se da minimização cuidadosa do volume de material para manter a performance desejada no calorímetro. Consequentemente, o ECAL e o CS dividem o mesmo recipiente de vácuo e, ainda, o CS é projetado da maneira mais fina possível sem que sejam sacrificadas a confiança e segurança operacional. Duas Tampas Toroidais (ECTs) estão inseridas nos finais do Barril do Toroide (BT) fazendo um envoltório externo aos calorímetros. Eles

estão inseridos de forma alinhada ao CS e fornecem o campo magnético necessário para o Espectrômetro de Múons como grandes toroides de núcleo a ar. Os picos de campo nos ECTs são de 3,9 T e de 4,1 T para o BT.

4.2 Sistema de Calorimetria

O Sistema de Calorimetria [52, 56] do ATLAS (Figura 4.4) consiste de calorímetros de amostragem com simetria e cobertura total em ângulo azimutal (ϕ). Todas as células dos calorímetros têm geometria aproximadamente retangular no plano $\eta \times \phi$. O ECAL do ATLAS cobre uma região $|\eta| < 3,2$, na qual o barril se estende na parcela central ($|\eta| < 1,475$) e a tampa cobre a região remanescente ($1,375 < |\eta| < 3,2$). Envolvendo o ECAL com mesma cobertura, encontra-se o HCAL. O barril do HCAL é composto pelo *TileCal*, contendo uma parte central ($|\eta| < 1,0$) e uma extensão ($0,8 < |\eta| < 1,7$). Realiza-se a cobertura final ($1,5 < |\eta| < 3,2$) da região de precisão por tampas (HEC) contendo LAr, devido à maior incidência radiação nessa região. Finalmente, emprega-se um calorímetro de menor precisão — Calorímetro Dianteiro (FCAL) — para cobrir a região mais próxima do tubo do feixe ($3,1 < |\eta| < 4,9$) e fornecer hermeticidade. A sua primeira camada atua como uma extensão ao ECAL, enquanto suas segunda e terceira camadas para a absorção de chuveiros HAD. As tecnologias empregadas em cada um dos calorímetros encontram-se nas subseções que se seguem.

A segmentação, granularidade e número de células de leitura (um total de $\sim 190k$ para os calorímetros, dos quais 170k provém do ECAL de precisão) em cada um dos calorímetros está disponível na Tabela 4.1. Em $|\eta| < 2,47$ aplica-se granularidade mais fina — compondo a região de precisão. A granularidade do HCAL é mais grosseira que a do ECAL devido à maior largura de chuveiros HAD. A extensão dos calorímetros de precisão até $|\eta| < 3,2$ permite o estudo de jatos para decaimentos do bóson W, enquanto no FCAL sua principal tarefa é fornecer hermeticidade para a obtenção de \cancel{E}_T , de modo que uma granularidade mais grosseira cumpre seu propósito. Ainda, observa-se um decaimento da granularidade conforme o acréscimo das camadas de segmentação longitudinais, o que ocorre devido a expansão da espessura lateral do chuveiro conforme a propagação do mesmo pelo calorímetro. Aplica-se, em geral, três amostragens longitudinais para obter, principalmente, melhor potencial de discriminação, mas também possibilitar a obtenção de melhor resolução em energia via calibração.

Tabela 4.1: Região de cobertura, granularidade, segmentação longitudinal e número de células de leitura dos calorímetros do ATLAS. Os valores das células de leitura em ϕ estão aproximados. O número total de células de leitura, incluindo ambas tampas, com (sem) os calorímetros pré-amostradores são 187.652 (178.308). [166]

Pré-Amostrador (PS)	Barril (<i>PreSamplerB</i>)	Tampa (<i>PreSamplerE</i>)
Cobertura	$ \eta < 1,52$	$1,5 < \eta < 1,8$
Segmentação Longitudinal	1 amostragem	1 amostragem
Granularidade ($\Delta\eta \times \Delta\phi$)	$0,025 \times 0,1$	$0,025 \times 0,1$
Células de Leitura	7.808	1.536 (total para ambos os lados)
Acordeão EM	Barril (EMB)	Tampa (EMEC)
	Cobertura Total	
	$ \eta < 1,475$	$1,375 < \eta < 3,2$
	Cobertura Parcial	
	Segmentação Longitudinal	
	3 amostragens	3 amostragens $1,5 < \eta < 2,5$
		2 amostragens $1,375 < \eta < 1,5$
		2 amostragens $2,5 < \eta < 3,2$
	Granularidade ($\Delta\eta \times \Delta\phi$)	
1 ^a amostra longitudinal	$0,003 \times 0,1$	$0,025 \times 0,1$ $1,375 < \eta < 1,5$
		$0,050 \times 0,1$ $1,375 < \eta < 1,425$
	$ \eta < 1,4$	$0,025 \times 0,1$ $1,425 < \eta < 1,5$
		$0,003 \times 0,1$ $1,5 < \eta < 1,8$
	$0,025 \times 0,025$	$0,004 \times 0,1$ $1,8 < \eta < 2,0$
		$0,006 \times 0,1$ $2,0 < \eta < 2,37$
$1,4 < \eta < 1,475$	$0,1 \times 0,1$ $2,37 < \eta < 3,2$	
	$0,025 \times 0,025$ $1,375 < \eta < 1,425$	
2 ^a amostra longitudinal	$ \eta < 1,4$	$0,025 \times 0,025$ $1,425 < \eta < 2,5$
		$0,050 \times 0,025$ $2,5 < \eta < 3,2$
	$1,4 < \eta < 1,475$	$0,1 \times 0,1$ $2,5 < \eta < 3,2$
3 ^a amostra longitudinal	$0,050 \times 0,025$	$0,050 \times 0,025$ $1,5 < \eta < 2,5$
N ^o de células de leitura	101.760	62.208 (total para ambos os lados)
Telhas Cintilantes HAD (<i>TileCal</i>)	Barril (<i>TileBar</i>)	Barril estendido (<i>TileExt</i>)
Cobertura Total	$ \eta < 1,0$	$0,8 < \eta < 1,7$
Segmentação Longitudinal	3 amostragens	3 amostragens
	Granularidade ($\Delta\eta \times \Delta\phi$)	
1 ^a e 2 ^a amostras longitudinais	$0,1 \times 0,1$	$0,1 \times 0,1$
3 ^a amostra longitudinal	$0,2 \times 0,1$	$0,2 \times 0,1$
N ^o de células de leitura	2.880	2.304 (total para ambos os lados)
Intermediário de Telhas (ITC)	Região do <i>Crack</i>	
Cobertura Total	$\sim 0,8 < \eta < 1,6$	
Segmentação Especial	Contém 6 células, dentre elas a barra cintiladora (Figura 4.8).	
Argônio Líquido HAD	Tampa (HEC)	
Cobertura Total	$1,5 < \eta < 3,2$	
Segmentação Longitudinal	4 amostragens	
	Granularidade ($\Delta\eta \times \Delta\phi$)	
	$0,1 \times 0,1$	$1,5 < \eta < 2,5$
		$2,5 < \eta < 3,2$
N ^o de células de leitura	3564 (total para ambos os lados)	
Calorímetro Dianteiro (FCAL)	Região Dianteira	
Cobertura Total	$3,1 < \eta < 4,9$	
Segmentação Longitudinal	3 amostragens	
Granularidade ($\Delta\eta \times \Delta\phi$)	$\sim 0,2 \times 0,2$	
N ^o de células de leitura	1762 (total para ambos os lados)	

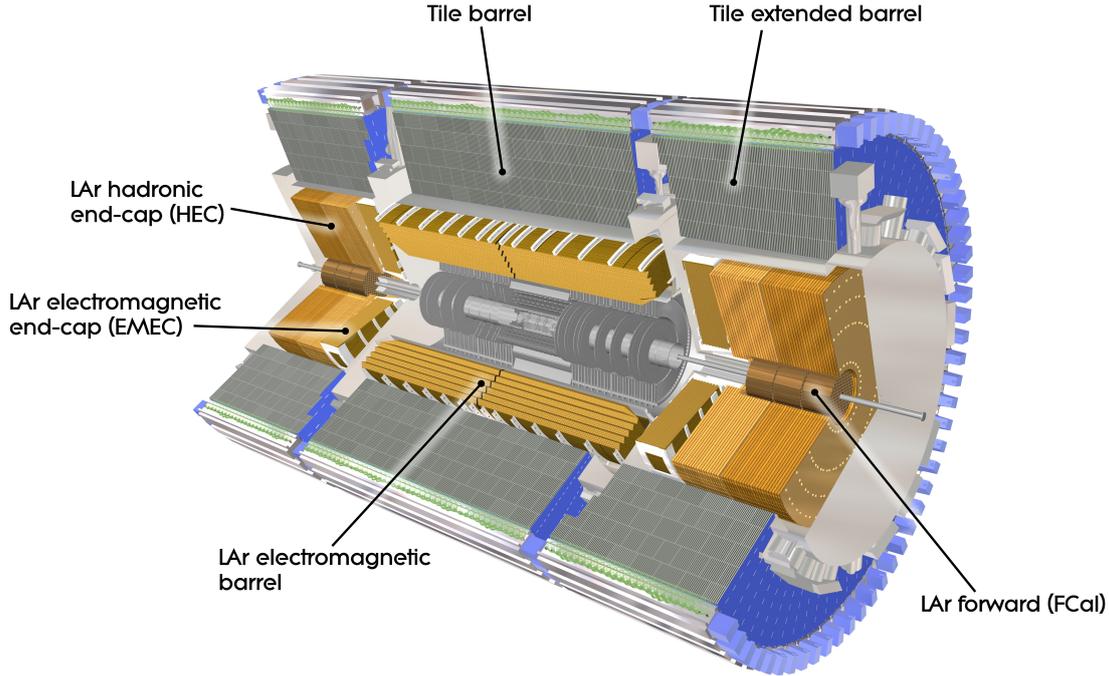


Figura 4.4: Os diversos calorímetros do ATLAS [165].

4.2.1 Calorímetro Eletromagnético (ECAL)

O ECAL de precisão, compostos por um barril (EMB) e tampa (EMEC), utilizam meio ativo de LAr com absorvedores compostos por eletrodos de cobre, enquanto o meio passivo é composto por placas de chumbo ($X_0 = 0,56$ cm e $\lambda_{int} = 17,59$ cm)². Projetou-se o ECAL para operar em uma larga faixa de dinâmica (50 MeV a 3 TeV), na qual a resolução de energia desses calorímetros atende

$$\frac{\delta E}{E} \approx \frac{10\%}{\sqrt{E}} \oplus 1\%. \quad (4.3)$$

A linearidade de resposta é superior a 0,5% na região de energia até 300 GeV, especialmente para obter ótima resolução ($\sim 1\%$) de massa de reconstrução dos decaimentos em elétrons e fótons do bóson de Higgs.

Por isso, escolheu-se LAr³ por comportamento linear e estabilidade de resposta temporal. Ainda, o LAr apresenta uma resistência a radiação intrínseca, necessário para esses calorímetros que estão na parte mais interna do detector. O sinal é gerado através da absorção das partículas geradas pelo chuveiro que causam a ionização do LAr. A ionização gera um par elétron-ion, e utiliza-se acoplamento capacitivo para direcionar as partículas carregadas para os eletrodos utilizando tensões na ordem de

²Retiraram-se os valores de comprimentos de [12].

³Para efeitos de comparação entre o meio ativo e passivo, os comprimentos do LAr são: $X_0 = 14,00$ cm e $\lambda_{int} = 85,77$ cm.

2 kV.

A estrutura dos calorímetros é no formato de acordeão (Figura 4.10a), que permite uma cobertura completa natural sem fissuras (*gaps*) em ϕ , assim como uma extração veloz dos sinais dos eletrodos na parte frontal e traseira. A estrutura em acordeão tem formato mais complexo na tampa devido à alteração das ondas do acordeão de acordo com a posição da célula. Esse fato, em conjunto com limitações construtivas, justificam a queda em granularidade em η na região na tampa (ver Tabela 4.1).

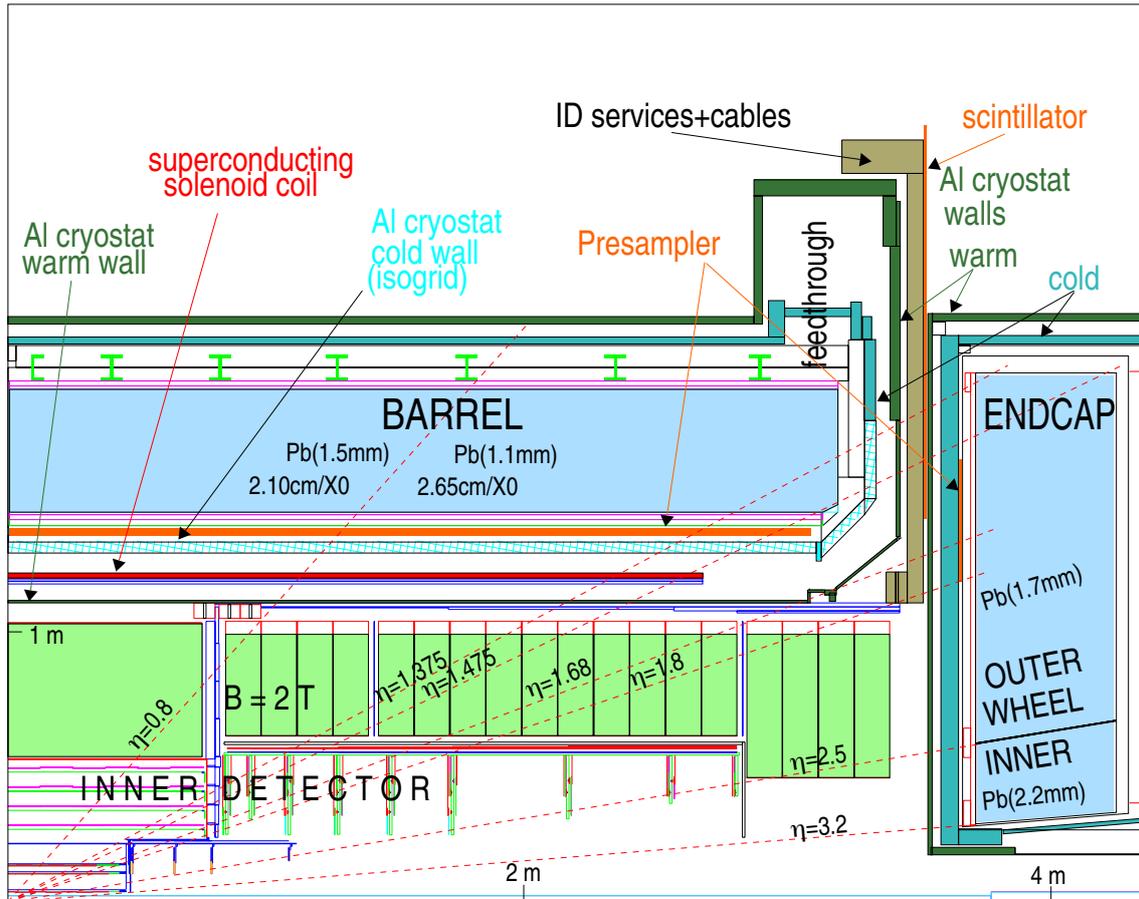


Figura 4.5: Seção de corte longitudinal do ECAL [56].

Um corte longitudinal do ECAL está disposto na Figura 4.5. Devido à complexidade da geometria do calorímetro, existem três regiões de fissuras (*crack*, em inglês), também simétricas em ϕ , nas quais a resposta do detector é degradada em comparação com o resto da cobertura. São elas:

- $|\eta| < 0,02$: entre a junção das duas metades do barril, existe uma pequena fissura totalizando 6 mm de LAr inativo;
- $1,37 < |\eta| < 1,54$: emprega-se a região de transição entre o barril e a tampa como trajeto para os cabos e serviços do ID. Essa região é a maior fissura do

detector, começando a degradar lentamente a resposta do detector em $\eta = 1,37$, e estende-se até cerca de $|\eta| < 1,68$. Todavia, por volta de $|\eta| < 1,54$ considera-se a resposta calorímetro do suficiente para estudo de física com precisão;

- $2,47 < |\eta| < 2,5$: na região de transição entre a tampa mais externa e interna do detector há uma pequena fissura de 3 mm que degrada a resposta do calorímetro. Há também uma degradação adicional causada por material morto (um anel de suporte mecânico) — material no qual a partícula interage porém que não contribui para geração de sinal — colocado na frente deste local.

A estrutura de acordeão tem como vantagem a flexibilidade de segmentação longitudinal e transversa, permitindo a implementação de camadas com diferentes granularidades. Ainda mais importante, essa estrutura permitiu ultrapassar uma limitação de tempo de resposta de calorímetros utilizando LAr e possibilitou seu emprego em ambientes de alta taxa de eventos como o LHC. Utilizam-se três camadas, que podem ser observadas na Figura 4.6, com as seguintes propriedades:

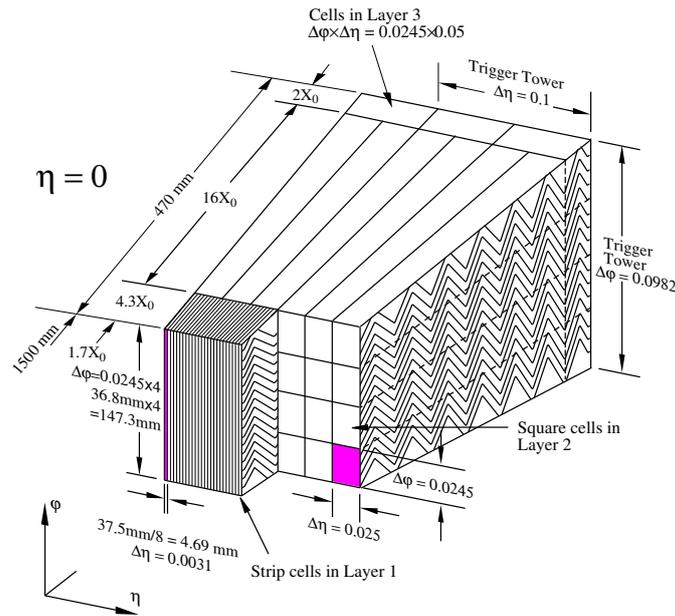


Figura 4.6: Esboço da granularidade das camadas do barril do ECAL em $|\eta| = 0$ e um esboço da geometria de acordeão [56].

- **primeira camada EM (EM1)**: composta por tiras (do inglês, *strips*) finas com grande granularidade em η , tendo como objetivo fazer uma boa leitura dessa posição. Isso é especialmente importante no caso de fótons, que não são medidos pelo detector de traços, e ao mesmo tempo para casar os traços das partículas com seus respectivos chuveiros. Conforme o η aumenta, há um decréscimo da granularidade devido ao fato das tiras não poderem ser feitas

com menos de 5 mm. A escolha de tiras mais grosseiras em ϕ quando comparadas com a segunda e terceiras camadas é consequência do campo magnético do CS espalhar em ϕ os chuveiros prematuramente iniciados anteriormente ao calorímetro. Além disso, a partícula π^0 é instável e decai próximo ao IP de acordo com $\pi^0 \rightarrow \gamma\gamma$ resultando em dois chuveiros EM com pequena separação devido ao pequeno ângulo formado entre os fótons no decaimento dessa partícula. Essa camada possui resolução suficiente para permitir tal separação;

- **segunda camada EM (EM2)**: responsável pela absorção da maior quantidade de energia. Ela é segmentada transversalmente em torres quadradas de $(\Delta\eta \times \Delta\phi) \approx (0,025 \times 0,025)$, que permitem um compromisso ótimo entre a contenção do perfil lateral do chuveiro com a contribuição do ruído por Empilhamento e eletrônico para a medição de energia;
- **terceira camada EM (EM3)**: tem a mesma granularidade que a central em ϕ , mas sua granularidade é duas vezes mais grosseira em η . A motivação da alteração de granularidade dá-se pela redução do número de canais. Essa camada possibilita a separação de chuveiros de altas energias, e contribui para a separação de γ /jato e elétron/jato, especialmente em menores energias ($E_T \lesssim 50$ GeV). No caso da tampa, para $|\eta| > 2,5$, emprega-se apenas duas camadas com granularidade mais grosseira uma vez que se está fora da região de precisão.

Para que o calorímetro possa realizar a medição com precisão da energia, deseja-se minimizar a quantidade de material anterior do mesmo. Além de deteriorar a resolução de energia, há impactos também na precisão da posição de impacto da partícula com o mesmo. Por mais que se tenha a quantidade de material colocado antes do calorímetro para minimizar esse efeito, ainda assim há a possibilidade de ocorrência de chuveiros prematuros. Consequentemente, adicionou-se um calorímetro especial (*pre-sampler*), contendo apenas material ativo, para estimar a perda de energia no material existente antes do ECAL, estando descrito na Subseção 4.2.2.

O ECAL contém profundidade superior (Figura 4.7) a $24 X_0$ ($26 X_0$) no barril (tampa). Tal comprimento permite absorção completa de energia de elétrons até 3 TeV.

4.2.2 Calorímetro Pré-Amostrador (PS)

O Calorímetro Pré-Amostrador (PS) não possui meio passivo, constituído apenas de uma fina camada de LAr na região de $|\eta| < 1,8$. Seu barril ($|\eta| < 1,52$) contém

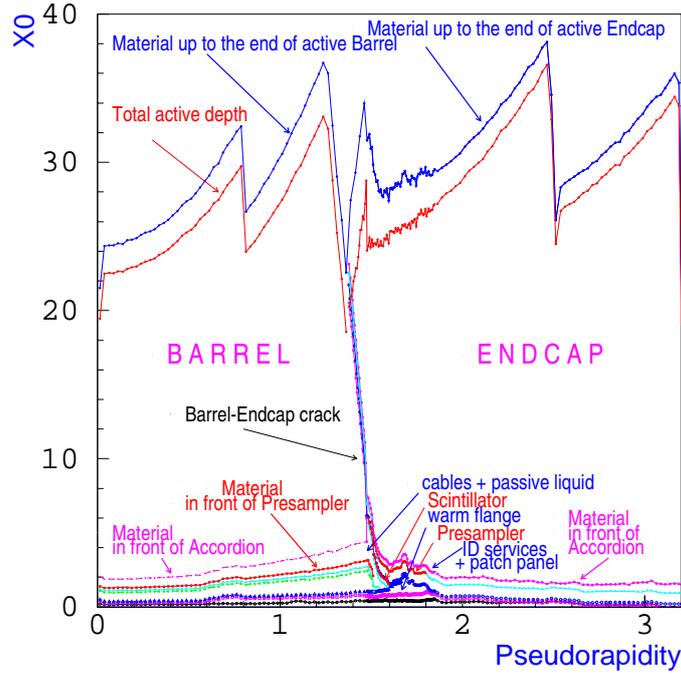


Figura 4.7: Simulação do comprimento total em termos de X_0 e em função de η transpassado por uma partícula até o final do ECAL. A quantidade de X_0 é discriminada em termos dos materiais disponíveis anteriores ao ECAL e, também, de suas amostragens longitudinais [56].

eletrodos perpendiculares ao feixe e um comprimento de 1,1 cm, enquanto na tampa ($1,5 < |\eta| < 1,8$) a configuração dos eletrodos é paralela e o comprimento de 0,5 cm. Sua função é absorver partículas de chuveiros formados antes do calorímetro do ATLAS pela interação das partículas com o material anterior ao calorímetro (observe os valores de X_0 desses materiais na Figura 4.7), como no CS, criostato e no ID. Com isso possibilita-se a realização de calibração da energia perdida pelas partículas nesse material. Além de $|\eta| > 1,8$, o PS não é mais necessário, dado que o material morto (material que não contribui para a detecção de energia da partícula) é reduzido e a energia total (E) depositada pelas partículas muito superior aquela depositada na região central para um mesmo E_T .

4.2.3 Calorímetro Hadrônico (HCAL)

O HCAL de precisão do ATLAS possui calorímetros com tecnologias distintas para a região central (Tópico 4.2.3.1 e Tópico 4.2.3.2) e tampa (Tópico 4.2.3.3).

4.2.3.1 Calorímetro Hadrônico de Telhas (*TileCal*)

No barril do HCAL, emprega-se como meio de amostragem telhas — originando a sua nomenclatura Calorímetro Hadrônico de Telhas (*TileCal*) — de cintiladores de plástico e aço como material passivo (o λ_{int} é aproximadamente igual ao do

chumbo, cerca de $16,8 \text{ cm}^2$). Ao invés do efeito capacitivo utilizado no LAr, seus cintiladores de plástico são excitados pelas partículas carregadas do chuveiro, de modo que são emitidos fótons capturados pelas fibras óticas posteriormente tendo seus sinais amplificados por Multiplicador de Fótons (PMT). O emprego da técnica de amostragem dá-se, pois, nesta região mais central do detector e após todo o LAr do ECAL, há menor incidência de radiação, o que possibilita utilizar um método com menor custo financeiro. Um esquema da disposição dos calorímetros do *TileCal* estão na Figura 4.8, enquanto um esquema da tecnologia empregada se encontra na Figura 4.10b.

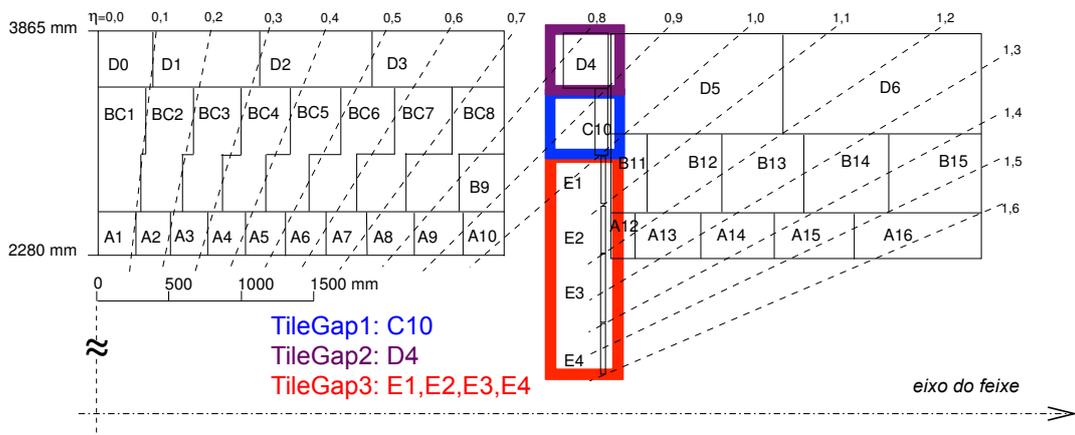


Figura 4.8: Corte esquemático das células do Barril (esquerda), Barril Estendido (direita) e Calorímetro Intermediário (células em destaque) do *TileCal*. As linhas horizontais indicam a segmentação longitudinal, já as linhas verticais definem fronteiras das células, formados por um grupamento ótico de fibras saindo dos cintiladores para as PMT. Linhas tracejadas representam um valor constante de η . As células da camada cintiladora estão destacadas em vermelho [167].

Compondo o *TileCal* estão o barril central, barril estendido e um calorímetro intermediário responsável por recuperar parte da informação perdida (Tópico 4.2.3.2). As coberturas de cada um dos calorímetros que o compõe e a disposição suas células podem ser observadas na Figura 4.5. O *TileCal* fornece três amostragens longitudinais para garantir melhor identificação das partículas e pela possibilidade de conseguir uma melhor resolução de energia através da calibração realizada pela ponderação do depósito em cada um das camadas. Projetou-se o calorímetro de forma a atender a resolução de energia do calorímetro necessária para possibilitar estudo de física contendo jatos duplos provenientes dos decaimentos do bóson W,

$$\frac{\delta E}{E} \approx \frac{50\%}{\sqrt{E}} \oplus 3\% \quad (4.4)$$

para $|\eta| < 3$. A linearidade de resposta do calorímetro mínima é de 2% cobrindo a escala até 4 GeV.

É necessário que o HCAL tenha no mínimo a espessura de $10 \lambda_{int}$ para a contensão completa dos chuveiros HAD, tanto para garantir resolução de energia, quanto reduzir o ruído causado no Espectrômetro de Múons. Os valores de λ_{int} para os diversos HCAL estão dispostos na Figura 4.9.

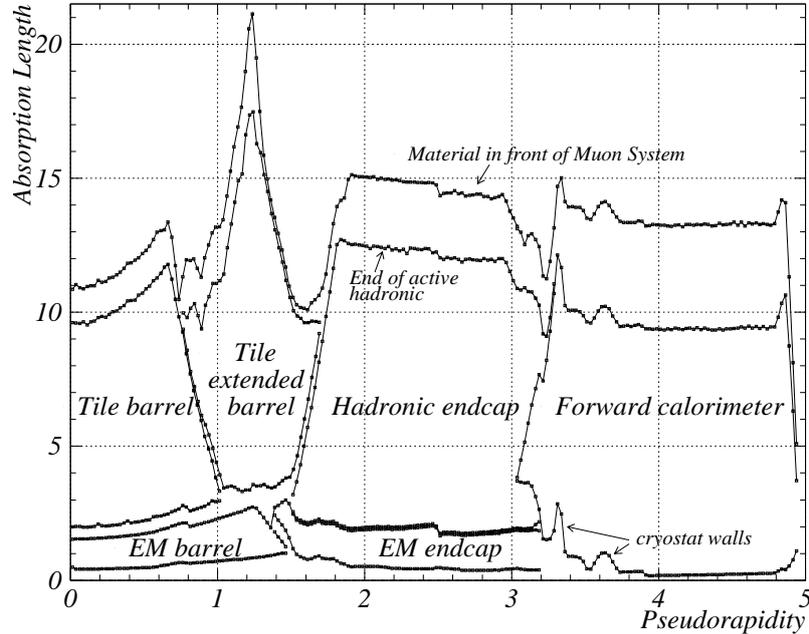


Figura 4.9: Simulação da profundidade em termos de λ_{int} e em função de η para os calorímetros compondo o Sistema de Calorimetria do ATLAS [56].

4.2.3.2 Calorímetro Intermediário de Telhas Cintilantes (ITC)

Na região de transição entre o barril e a tampa, para $|\eta| = 1,4$, a situação é particularmente crítica devido aos serviços e cabos para o ID, e uma camada cintiladora em $(1,0 < |\eta| < 1,6)$ é colocada entre os dois criostatos com o objetivo de recuperar parte da energia perdida e garantir a hermeticidade necessária para a obtenção de \cancel{E}_T . A barra cintiladora também permite melhorar a calibração de energia de jatos hadrônicos, elétrons e fótons. A barra cintiladora contém 4 células sem material passivo, nesse aspecto, operando de forma similar ao PS, ainda que as tecnologias empregadas sejam distintas. Há ainda duas células na região mais externa (ver Figura 4.8) responsáveis por garantir a quantidade mínima de λ_{int} do HCAL na região $0,8 < |\eta| < 1,0$. Assim, diferente das células compondo a barra cintiladora, essas células possuem material passivo que possibilita sua função. A barra cintiladora e essas duas células compõem o Calorímetro Intermediário de Telhas Cintilantes (ITC).

4.2.3.3 Tampas do Calorímetro Hadrônico (HEC)

Para a Tampa do Calorímetro Hadrônico (HEC) também se utiliza LAr como meio ativo devido à alta incidência de radiação, mas ao invés de chumbo, seu material passivo é o cobre ($X_0 = 1,43$ cm e $\lambda_{int} = 9,39$ cm²). Projetou-se sua estrutura como uma chapa plana demonstrada na Figura 4.10c. São utilizadas duas tampas para cada HEC contendo cada uma 32 módulos idênticos. Cada módulo consiste de 24 chapas de cobre para a primeira tampa, e 16 chapas para a segunda. Em ambos os casos as chapas de cobre estão separadas por uma fissura de 8,5 mm contendo LAr e três eletrodos. Essa estrutura foi escolhida principalmente por ter uma maior resistência a radiação e eficácia de custo, ainda fornecendo a cobertura espacial necessária.

Os requisitos de linearidade e de resolução de energia devem atender aqueles especificados para o *TileCal*. Ainda, um tempo de pico de sinal de ~ 40 ns deve ser atendido pelo LAr. O número de λ_{int} da HEC em função de η está na Figura 4.9.

Diferente do *TileCal*, a tampa do HEC realiza 4 amostragens longitudinais. Porém, normalmente as duas camadas centrais são agrupadas em uma única camada, de forma a manter a uniformidade de segmentação longitudinal do HCAL.

4.2.4 Calorímetro Dianteiro (FCAL)

Já o FCAL apresenta uma estrutura diferenciada para suportar o elevado índice de radiação próximo ao tubo do feixe. Em sua primeira camada, utiliza-se uma matriz de metal (cobre) absorvedor contendo buracos igualmente distribuídos. Coloca-se nesses buracos uma estrutura de hastes coaxiais e tubos, ambos novamente de cobre, separados por um pedaço preciso de fibra de plástico resistente a radiação. A matriz e o tubo compõem o material passivo, enquanto o espaço remanescente entre o tubo e a haste é preenchido por LAr, o meio de amostragem. O tubo está aterrado, enquanto as hastes estão em alta tensão, criando o efeito capacitivo. Essa estrutura fica mais facilmente compreendida na Figura 4.10d. O material das hastes e da matriz são substituídos de cobre para tungstênio ($X_0 = 0,350$ cm e $\lambda_{int} = 5,72$ cm²) com o objetivo de elevar a capacidade de absorção de partículas HAD nas segunda e terceira camadas. O acúmulo de íons de LAr nas fissuras limita a luminosidade máxima proporcionalmente a $1/g^2$, de forma que o comprimento das fissuras (g) devem ser os menores possíveis. As dimensões das mesmas são de 250-375 μ m para a seção EM e 500 μ m para a seção HAD.

A tarefa principal desse detector é a reconstrução de jatos e fornecer hermeticidade para a medição de \cancel{E}_T . Por isso, a resolução de energia para o HCAL na região dianteira pode ser menor, conforme

$$\frac{\delta E}{E} \approx \frac{100\%}{\sqrt{E}} \oplus 10\%; \quad (4.5)$$

para $3 < |\eta| < 5$.

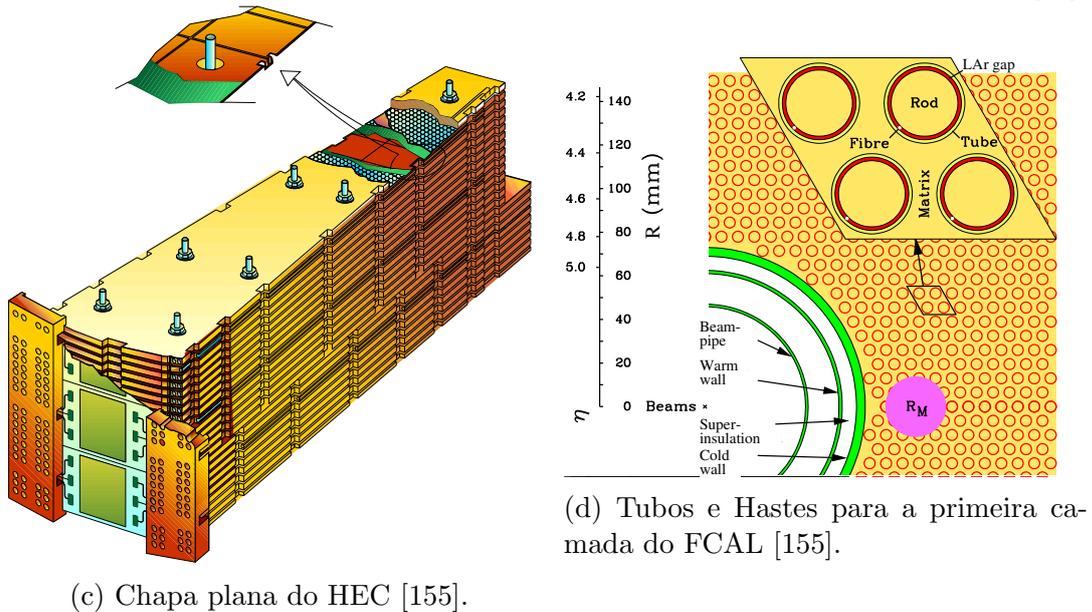
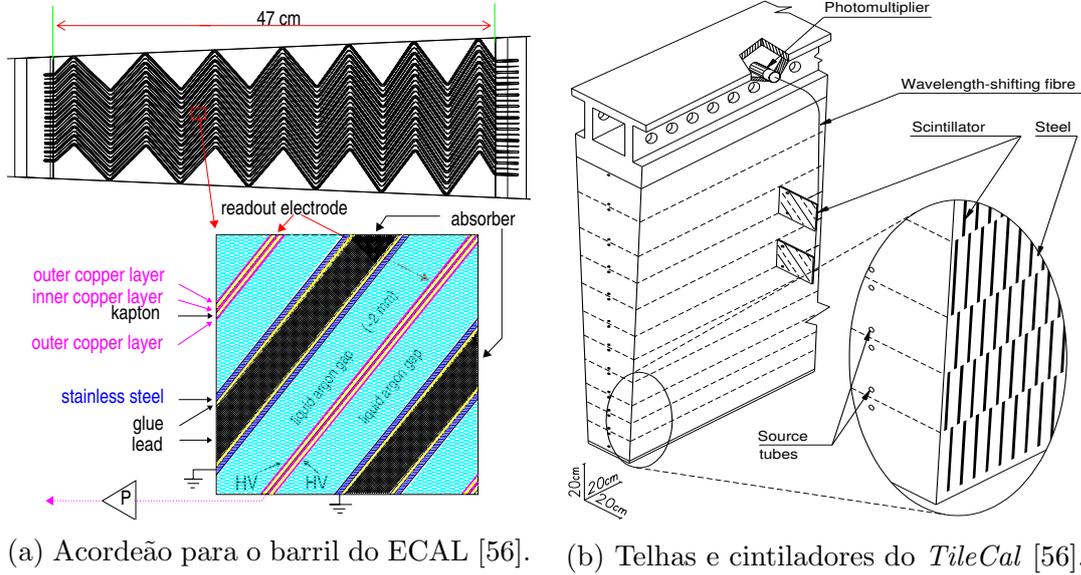


Figura 4.10: As diferentes estruturas dos subsistemas de calorimetria do ATLAS.

4.2.5 Sistema de Criogenia

Uma das desvantagens do emprego de calorímetros com LAr é dado pela necessidade de resfriamento. Para isso, o ATLAS dispõe de três criostatos individuais, um para o barril e dois para as tampas, que mantém o LAr em sua temperatura de operação (~ 80 K). A refrigeração ocorre através de nitrogênio líquido que circula em tubos no sistema de refrigeração em torno dos calorímetros. Há regulação

de pressão e fluxo para garantir apenas a potência necessária para manter o LAr nessa temperatura. Para minimizar o material, as paredes do criostato são feitas de alumínio. O criostato no barril envolve apenas o ECAL (o *TileCal* não utiliza LAr como meio passivo e por isso não necessita ser refrigerado), enquanto os criostatos na tampa envolvem o EMEC, HEC e o FCAL. A parede mais interna do criostato do barril utiliza uma estrutura *isogrid*, que tem $\sim 80\%$ de material a menos que uma placa sólida teria, mas rigidez suficiente para suportar a estrutura. Parte da interferência do material da estrutura de criogenia pode ser vista na Figura 4.5.

4.3 Sistema de Reconstrução (*Offline*) de Física do ATLAS

Os pequenos experimentos de Física de Partículas, como o caso de [168], não contavam com processamento computacional⁴. A necessidade de um Sistema de Reconstrução (*Offline*) de Física do ATLAS nos experimentos de física de partículas originou-se com o aumento da quantidade de dados a serem processados com o advento das Câmaras de Bolhas [169]. Assim, os primeiros *softwares* para a reconstrução *offline* da Física de Partículas originaram-se como uma solução para os desafios desses detectores, onde as primeiras versões de códigos eram responsáveis por fazer a reconstrução 3D das medidas digitalizadas das fotos (2D) retiradas no detector e então realizar um ajuste da cinemática do evento. Para mais detalhes da evolução dos Sistemas de Reconstrução (realizada focando experimentos de grande porte), dirigir-se a [1, 169]⁵.

Com o crescimento das análises, fez-se necessário o desenvolvimento de infra-estruturas dedicadas para atender as necessidades das análises físicas, o Sistema de Reconstrução (*Offline*) de Física do ATLAS. O mesmo realiza o escrutínio dos sinais do detector de forma a fornecer informação de alto nível para análise. O (re)processamento⁶ inicia-se a partir dos dados crus obtidos dos sensores do ATLAS selecionados pelo Sistema de Filtragem do ATLAS e informações sequencialmente de nível mais alto são geradas até que os objetos físicos básicos empregados nos canais

⁴Vale notar que, no ano do experimento (1957), os computadores eram extremamente caros e de difícil acesso [169]. Ainda, para dar uma noção do contexto, na data de publicação do trabalho (1957) ano, estava sendo lançada primeira versão comercial do Fortran.

⁵Ambas referências têm valor histórico, onde está disponibilizado a evolução e o legado em Fortran para a análise física em programas bastante utilizados na época, como o *PAW*, do mesmo autor que seu substituto *ROOT*, bastante utilizado atualmente para análise em Física de Altas Energias (HEP) [169].

⁶Realiza-se o processamento oficial quantas vezes forem necessárias em determinadas versões de *software* a fim de integrar melhorias no resultado a ser fornecido para análise. Normalmente, esse processo é chamado de reprocessamento, uma vez que os dados sempre são processados pela primeira vez em *offline* assim que liberado pelo grupo de plantonistas de qualidade de dados no P1.

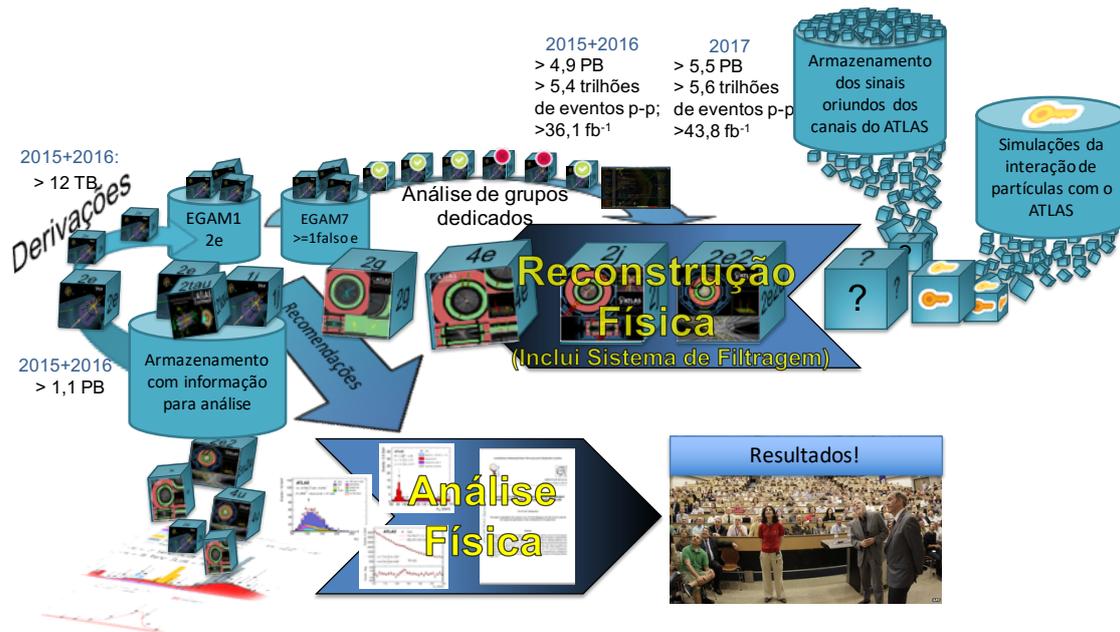


Figura 4.11: Esboço do fluxo da informação envolvendo a análise física contendo elétrons como parte dos canais avaliados no Experimento ATLAS.

de análise sejam gerados. Por exemplo, para o caso de elétrons (descrição completa na Seção 5.1), a informação tratada por algoritmos específicos de processamento do Sistema de Calorimetria e do ID são empregadas individualmente para a geração de informações de mais alto nível, como *clusters* e traços, respectivamente. O processo segue para fundir essa informação de mais alto nível em um candidato a elétron. Esses candidatos passam por uma série de processamentos, decorando-os com informações pertinentes, como tratamentos para melhoria de precisão (calibração de energia), geração de grandezas física e dos anéis em energia conforme implementado por este trabalho, decisão do classificadores e seus discriminantes etc.

No ATLAS, ambiente em *software* contendo o código de reconstrução está englobado na Infraestrutura *Athena*. Por sua vez, uma diversidade de outras funcionalidades também estão disponíveis nessa infraestrutura, englobando: o Sistema de Filtragem do ATLAS (Seção 4.4), integração de geradores de Monte Carlo (MC) para simulação de eventos de colisão, código de integração com o Geant para digitalização dos eventos simulados e, até mesmo, parte dos códigos de análise física. Os dados produzidos durante os processamento contém uma nomenclatura própria [170], onde o formato interessante para a análise se dá por AOD Multi-Ambiente (xAOD), um novo formato adicionado para a Segunda Temporada de Coleta de Dados que permite a aplicação da análise em ambiente do Infraestrutura *Athena* e no *Framework* de Análise ROOT (*ROOT*). As xAOD passam por derivações para geração de conjuntos de dados específicos e pré-selecionados para uma aplicação. No caso de elétrons,

derivações de interesse⁷ para avaliação de eficiência em elétrons:

- EGAM-1: dedicada a pré-seleção de eventos de cruzamento de pacotes contendo elétrons compatíveis com o decaimentos do bóson Z através do emprego da técnica Método *Tag and Probe* (T&P) (Seção 5.6). A pré-seleção compreende lógica inclusiva para os seguintes casos (todos exigindo que a massa invariante de um par, via soma dos quatro-momentos, seja superior a 50 GeV):
 - dois elétrons centrais ($|\eta| < 2,5$) de Carga Elétrica Oposta (OS), um deles aceito pelo qualidade *tight* das seleções *LLH* ou *CutBased* (Seção 5.3) e $p_T > 24,5$ GeV e outro aceito pelo qualidade *medium* por ao menos uma das seleções citadas e $p_T > 19,5$ GeV;
 - dois elétrons centrais OS qualidade *medium* pela seleção *LLH* ou *CutBased* com $E_T > 19,5$ GeV;
 - dois elétrons centrais, um qualidade *medium* *LLH* ou *CutBased* e outro $p_T > 6,5$ GeV;
 - um elétron qualidade *medium* *LLH* ou *CutBased* e um fóton com $E_T > 14,5$ GeV.
- EGAM-2: elétrons compatíveis com decaimentos de J/Ψ . Esses elétrons são empregados para o ajuste da seleção de elétrons em região de baixa cinemática ($E_T < 15$ GeV) e não são empregados atualmente para a versão *offline* de identificação de elétrons por limitações técnicas para sua produção (Tópico 7.3.4.1). O desenvolvimento do ajuste de modelos para esses dados está em andamento para versão *online*;
- EGAM-7: contém pré-seleção de eventos compatíveis com falsos elétrons. O critério aplicado é:
 - evento aceito por ao menos uma cadeia de suporte (Seção 4.4) destinada a coleta de elétrons e fótons e ocorre a reconstrução *offline* de ao menos um elétron central com $p_T > 4,5$.

A geração de derivações é extremamente importante para reduzir a banda e processamento de eventos desnecessários. Por exemplo, o reprocessamento mais recente de dados de 2016, em formato xAOD, contém 1,1 PB, enquanto a sua derivação EGAM-1 equivale somente à 12 TB desses dados. Apenas para comparação, a informação com os sinais oriundos dos canais do experimento ocupa 5,5 PB.

⁷Aqueles colaborando com o ATLAS e interessados em mais informação relevante às derivações de elétrons, dirigir-se a [171].

O processamento de toda essa informação requer um grande poder computacional, armazenamento e banda. Esses recursos são disponibilizados através de computação em nuvem pela a LCG [172], a maior grade computacional do mundo.

4.4 Sistema de Filtragem

O Sistema de Filtragem do ATLAS [173] realiza a seleção *online* de eventos para redução da taxa em nível viável de processamento e armazenamento *offline*, conforme esboçado na Figura 4.12.

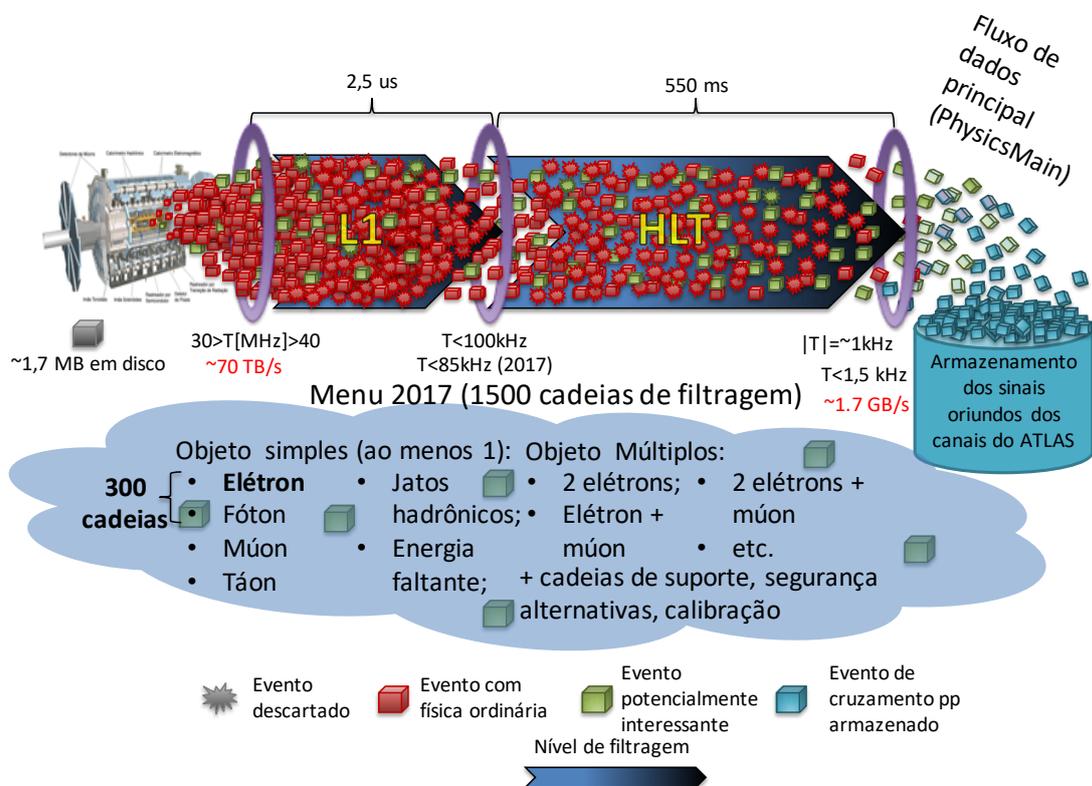


Figura 4.12: Esboço da atuação da filtragem *online* realizada no ATLAS. A nuvem ilustra a composição de objetos físicos de interesse, que coletivamente são denominados como o menu do SF_{on} . Eventos de colisão (não) gerando esses objetos são representados por cubos verdes (vermelho). Conforme o avanço das etapas da filtragem, o SF_{on} descarta sucessivamente maior quantidade de eventos ordinários e mantém grande parte dos eventos contendo objetos de interesse para o armazenamento e posterior processamento em *offline*.

A sua atuação ocorre como parte do Sistema de Qualidade e Filtragem de Dados (TDAQ) (Figura 4.13), consistindo de um nível baseado em *hardware* de primeiro nível (L1) que precede a atuação do Alto Nível de *Trigger* (HLT), implementado em *software*. Fazem parte do processo de tomada de decisão do L1 o Processador Central de Filtragem (CTP), que recebe informação de seleção do Primeiro Nível

de Filtragem para o Sistema de Calorimetria (*L1Calo*) e Primeiro Nível de Filtragem para o Espectrômetro de Múons (*L1Muon*), bem como de outros subsistemas como o Cintiladores de Filtragem para *Minbias* (MBTS), *LUCID* e Calorímetro Contador de Zero Grau (ZDC). No caso, o *L1Calo* e *L1Muon* são componentes de seleção especialista do L1 que realizam o processamento de informação unicamente do Sistema de Calorimetria e Espectrômetro de Múons, respectivamente. O CTP também se responsabiliza pela regularização da operação do L1 ao aplicar um limite para o tempo para a tomada de decisão, ignorar duas aceitações consecutivas do L1 (tempo morto simples, em 2015 de 100 ns) a fim de evitar sobreposição de janelas de leitura e restringem o número de aceitações em um número de cruzamentos de pacotes (tempo morto complexo) de forma a manter a operação dentro da capacidade dos *buffers* de *front-end*. Caso ocorrendo a aceitação por parte do L1, os eventos são armazenados no Sistemas de Leitura (ROS) e processados pelo HLT. O HLT obtém informação em Região de Interesse (RoI) do L1, que pode ser usada para reconstrução local pelos algoritmos de filtragem. Somente após a aceitação por parte do HLT que os eventos são transferidos para armazenamento.

Realizou-se uma série de atualizações durante a pausa de coleta de dados entre a Segunda Temporada de Coleta de Dados e Primeira Temporada de Coleta de Dados a fim de preparar o L1 para atuação em condições mais severas devido ao aumento da luminosidade. O *L1Calo* recebeu a implementação de novas correções de pedestal para aliviar o efeito de empilhamento, especialmente importante para melhor medição de \cancel{E}_T . No *L1Muon*, implementou-se uma lógica de coincidência⁸ entre *TileCal* e o Espectrômetro de Múons a fim de possibilitar redução da taxa de falso alarme. Melhorias de cobertura do *L1Muon* foram instaladas. Substituiu-se o CTP por uma versão que permite o dobro de limiares de decisão no L1 e a taxa de saída do L1 (HLT) para o fluxo de dados dedicado a física⁹ foi modificada do limite de 70 kHz (400 Hz) para 100 kHz (1 kHz, tolerando-se taxas de até 1,5 kHz em pico). Instalou-se uma nova lógica de L1 aplicando processamento topológico empregando de informação de calorimetria e do Espectrômetro de Múons (atuação a partir de setembro de 2016). Além disso, o HLT também foi afetado, onde se extinguiu a separação do mesmo em dois subníveis com fazenda própria dedicado para permitir maior flexibilidade durante operação. O tempo de latência para tomada de decisão foi alterado de 4 s [176] para 550 ms. Realizou-se, também, uma série de atualizações

⁸Contribuição sugerida e implementada pelo grupo brasileiro do ATLAS [174, 175]. Está programada sua atuação para 2018.

⁹O termo empregado no ATLAS é *PhysicsMain*. Há uma diversidade de outros fluxos de dados, em geral, empregados para a calibração, monitoração e verificação (*debug*) da operação *online*. Inclusive, o emprego do *PhysicsMain* foi parte das atualizações para a Segunda Temporada de Coleta de Dados, onde se separava os fluxos para cada assinatura de interesse do Sistema de Filtragem do ATLAS, o que gerava duplicação de eventos e requeria um aumento de 10 % no armazenamento e processamento durante a reconstrução.

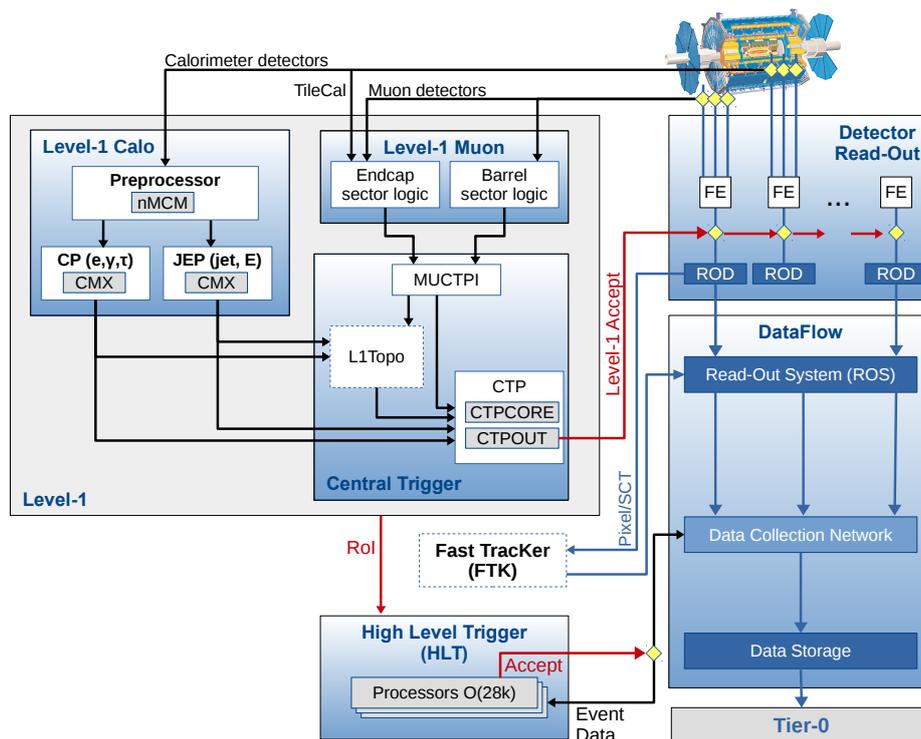


Figura 4.13: O TDAQ do ATLAS em atuação durante a Segunda Temporada de Coleta de Dados. Ênfase foi dada aos componentes relevantes para a filtragem. [173]

no código do HLT a fim de torná-los mais próximos da versão do *offline* e empregar novos sistemas disponíveis na Segunda Temporada de Coleta de Dados como a *IBL*. Algumas especificações importantes foram mantidas, como latência inferior a $2,5 \mu\text{s}$ para a atuação do L1.

A filtragem é organizada em cadeias de seleção dedicadas a coletar objetos físicos de interesse. As cadeias podem ser classificadas conforme:

- cadeias primária: empregada para análises físicas e tipicamente atuam sem a aplicação de um fator de pré-escala (evento enviado para o próximo nível somente após a n -ésima aceitação, onde n é o fator de pré-escala);
- cadeias de suporte: medição de eficiência para monitoração e atuando geralmente em pequenas taxas ($0,5 \text{ Hz}$ cada). Especificamente para o caso de elétrons, as cadeias de suporte incluem a atuação de seleções unicamente basea-

das em E_T ¹⁰ ou com a aplicação de seleção com critério minimamente restritivo (*veryloose*, Seção 5.3);

- cadeias alternativas: algoritmos experimentais, novos e alternativos de reconstrução comparados com as seleções de suporte e primárias, geralmente com grande sobreposição de decisão com as cadeias primárias;
- cadeias de segurança: usam seleções mais rígidas para substituição de cadeias primárias caso a taxa de operação esteja acima do esperado;
- cadeias de calibração: atuação geralmente com alta taxa porém armazenando informação relevante somente para calibração.

As cadeias primárias cobrem todas assinaturas relevantes para o programa de física do ATLAS, contendo dentre seus objetivos a medição precisa do MP (bósons Higgs, W e Z) e buscas por física além do MP (partículas pesadas, Teoria de Supersimetria (SUSY) e partículas exóticas). Os objetos físicos dentre os canais físicos explorados incluem elétrons, fótons, múons, táons, jatos-b e \cancel{E}_T , onde elétrons são de vital importância por participarem de canais de nova física, como o caso do $Z' \rightarrow ee$. As cadeias podem se dedicar exclusivamente a encontrar a presença da assinatura de um desses objetos ou buscar múltiplas assinaturas a fim de permitir uso eficiente da banda. Em 2017 [177], o *menu* (conjunto de cadeias) em operação conteve 1.500 cadeias das quais 350 destinavam-se a coleta de elétrons. Na Figura 4.14, sumariza-se as nuances operativas durante uma *run*.

Com o crescimento do empilhamento durante a Segunda Temporada de Coleta de Dados, outras atualizações foram realizadas durante 2016 [179] e 2017 [177], das quais aquelas afetando elétrons serão cobertas na Seção 5.7.

4.5 Outros Sistemas

Outros subsistemas especializados do ATLAS são:

- Espectrômetro de Múons: possibilita a identificação e medição com precisão de múons na região $|\eta| < 2,7$. É um espectrômetro magnético como o ID, contando com dois subsistemas de precisão e um dedicado para filtragem rápida;
- Serviços Eletrostáticos de Coleta de Feixes (BPTX): localizados a uma distância de $z = \pm 175$ m e responsável pela sincronização do Sistema de Filtragem do ATLAS e de monitoramento com os feixes;

¹⁰A nomenclatura empregada para a definição dessas cadeias é *etcut*.

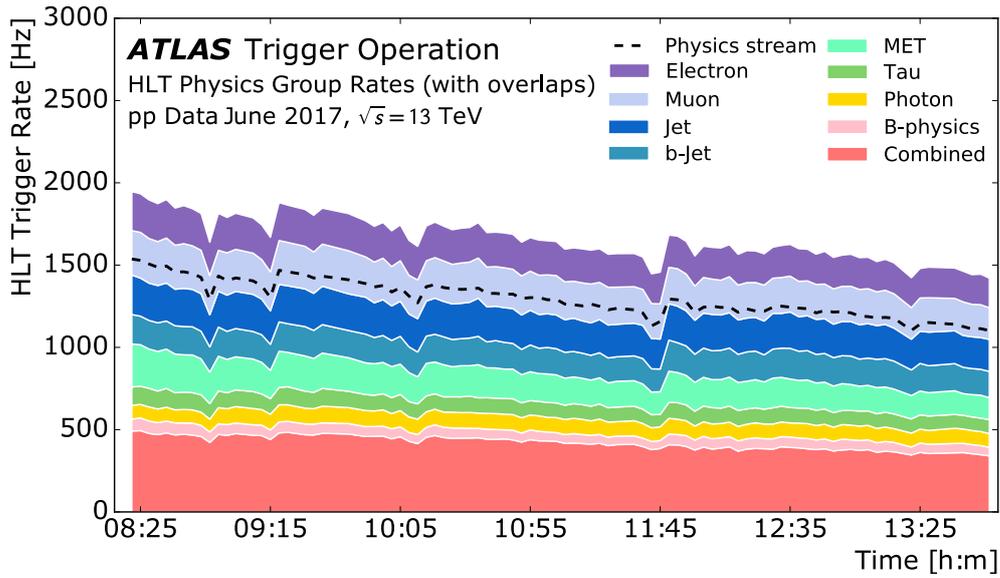


Figura 4.14: Taxas por grupos de eventos de interesse de filtragem no HLT em função do tempo em um período de coleta obtido em junho de 2017 e com pico de luminosidade de $1,53 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. Apresentam-se as taxas para os grupos individuais de filtragem específicos para a filtragem específica de objetos físicos. Cada grupo contém filtros únicos ou múltiplos para o mesmo objeto. Sobreposição são apenas levadas em conta para a taxa total de saída de física primária. A filtragem de física-B e estados leves são basicamente algoritmos de filtragem de múons; havendo um fluxo de dados dedicado aos mesmos e, portanto, a maioria dos eventos não são apresentados aqui. O grupo combinado apresenta cadeias para objetos diferentes, como combinações de elétrons, múons, táons, jatos e \cancel{E}_T . Padrões comuns para todas as taxas são o seu decaimento exponencial com a redução da luminosidade durante um preenchimento do LHC. As taxas aumentam periodicamente devido à mudança de fatores de pré-escala a fim de otimizar o uso de banda ou devido a reotimizações da luminosidade do LHC, enquanto declives são devidos a falhas no tempo de resposta, e picos causados por ruídos no detector. [178]

- Monitoramento das Condições dos Feixes (BCM): constituído de duas estações contendo sensores de diamante a uma distância de $z = \pm 1,84 \text{ m}$ e $\eta \approx 4,2 \text{ m}$;
- Contadores LUCID \tilde{C} erenkov cobrindo a região $5,4 < |\eta| < 5,9$, que juntos com os BCM são responsáveis pela medição precisa da luminosidade;
- ZDC: cobrindo uma região $|\eta| > 8,3$, responsável por medição de partículas neutras em 0° , tanto em colisões $p-p$ como Pb-Pb;
- Luminosidade Absoluta para o ATLAS (ALFA): cobrindo $6 < |\eta| < 13,5$, responsável por medir a seção de choque de colisões elásticas; e
- MBTS: que fornecem informação e dados ao Primeiro Nível de Filtragem, cobrindo a região de $2,1 < |\eta| < 3,8$.

Capítulo 5

Identificação de Elétrons no ATLAS

Os elétrons são objetos físicos presentes em diversos canais de física de interesse do ATLAS, dentre eles envolvendo a busca de processos físicos ainda mais raros que o bóson de Higgs, como o Z' . A reconstrução dos objetos físicos ocorre tanto em *offline* (Seção 5.1), para disponibilizar uma estratégia recomendada para as análises físicas, ou em cinco estágios sequenciais nas cadeias dedicadas à coleta de elétrons (Seção 5.7). As grandezas padrões empregadas pelo ATLAS para a discriminação de elétrons estão na Seção 5.2. As referências para identificação de elétrons estão na Seção 5.3. A essa decisão, pode-se aplicar complementarmente a exigência de isolamento dos elétrons (Seção 5.4), o que permite obter maior potencial de discriminação ao possibilitar a remoção de elétrons provenientes de outras origens distintas da física de interesse. Considerações são realizadas sobre as necessidades do experimento para a operação dos discriminadores em relação ao empilhamento (Seção 5.5). O principal método empregado para a avaliação de eficiência de elétrons está na Seção 5.6, que se caracteriza por uma pré-seleção *loose* de elétrons sem a aplicação de nenhum discriminador. As atualizações realizadas no SF_{on} durante a *Run 2* estão na Subseção 5.7.1, incluindo a necessidade para a mudança de paradigma nas cadeias de elétrons. A mudança de paradigma, para o emprego do *NeuralRinger*, está no cerne deste trabalho e por isso é realizada uma discussão em torno da questão do impacto envolvido nessa alteração (Subseção 5.7.2).

5.1 Reconstrução de Elétrons

A reconstrução de elétrons procede nos seguintes passos [180, 181]¹: reconstrução de *cluster* através de sementes; reconstrução de traços ajuste-fino de traço para candidatos a elétron; e reconstrução final dos candidatos.

O processo de construção de *cluster* é realizado para $|\eta| < 2,5$, onde é realizado a construção das torres, o Sistema de Calorimetria é dividido em uma grade de $0,025 \times 0,025$ (correspondente a granularidade da EM2), através da soma da energia das células das camadas longitudinais do ECAL contidas dentro da região. Em seguida, uma janela de 5×5 torres procura por uma região que ultrapasse um limiar para $E_T > 2,5$ GeV e que seja um máximo local. Quando encontrado, gera-se uma semente cuja posição é calculada através da ponderação de energia das células contidas em uma janela concêntrica a anterior porém menor (3×3 torres) para evitar o acúmulo de ruído, com as posições das células. Para remover duplicatas, ao se ocorrer múltiplas sementes dentro de uma região 2×2 torres, armazena-se apenas a semente mais energética. Os *clusters* são reconstruídos com seus tamanhos específicos dependente da partícula e sua incidência (3×7 para elétrons e fótons convertidos, 3×5 para fótons não-convertidos quando incidindo no barril e 5×5 na tampa). O processo inicia-se na EM2 através da construção do *cluster* a partir da posição da semente. Então o baricentro de energia é calculado e propagado para a EM1 e EM3, aonde são gerados os *cluster* para essas camadas. Para o PS emprega-se como centro o baricentro de energia da EM1.

Por outro lado, a reconstrução de traço ocorre em dois passos: reconhecimento de padrão e ajuste de traço. A reconstrução de elétrons possui um reconhecimento de padrão com algoritmo próprio para a perda de energia por *bremstrahlung* no material do ID, o que permite a recuperação de até 30 % da parcela de energia perdida. Essa algoritmo complementa a estratégia padrão de reconstrução de traços do ATLAS que assume como hipótese que a partícula atravessando o detector é um pión. O algoritmo inicia com sementes de traço (ao menos três pontos em diferentes camadas do SCT) de $p_T > 1$ GeV empregando a hipótese de píons para verificar a geração de traços com ao menos 7 pontos dentro de uma região de interesse contendo um *cluster* EM. Caso essa estratégia falhe, repete-se a reconstrução do traço com a hipótese de elétrons. Os traços reconstruídos de forma bem-sucedida por uma das estratégias anteriores são ajustados por um algoritmo [183] empregando a hipótese tomada para perda de energia. No caso da hipótese de pión seja tomada e a mesma falhe, reavalia-se o algoritmo com a hipótese de elétron. Essa estratégia possibilitou

¹Atualmente, há uma nova reconstrução com a construção de *clusters* dinâmicos [182] que melhoram a reconstrução de elétrons radioativos e fornecem resiliência ao efeito de empilhamento. Todavia, o emprego dos chamados “*superclusters*” não afetam a reconstrução da informação anelada de calorimetria.

integrar um algoritmo especializado na reconstrução de traços de elétrons em cima da base padrão de reconstrução de traços do ATLAS.

Em seguida, ocorre o casamento da informação através de critérios pouco exigentes (*loose*) através da extrapolação do traço até seu ponto de impacto com a EM2 e da posição do baricentro de energia calculado nessa camada. Para o casamento ocorrer, leva-se em consideração a curvatura do traço, o número de pontos de precisão no SCT. Os traços com número suficiente de pontos de precisão (4) são posteriormente reajustados através da aplicação de GSF otimizado, que considera perdas não-lineares devido ao efeito de *bremsstrahlung*.

Com isso, completa-se a reconstrução de candidatos a elétron. Caso tenham diversos traços casando com um *cluster*, a escolha do traço primário é feita por um algoritmo que computa a distância $\Delta_R = \sqrt{\Delta_\eta^2 + \Delta_\phi^2}$ utilizando diferentes hipóteses de momentos em conjunto com o número de pontos no Detector de Pixel e a exigência de ao menos um ponto no SCT.

5.2 As Grandezas Físicas para Discriminação de Elétrons

A partir dos candidatos a elétron obtidos, geram-se as grandezas físicas utilizadas para discriminação. O ATLAS emprega um conjunto de grandezas com alto teor discriminante na tarefa de identificação [180, 181] elétron-hádron, dentre elas constam sete grandezas relacionadas ao formato e largura do chuveiro de partículas, razões de energia provenientes do Sistema de Calorimetria. Outras seis grandezas são geradas a partir do conhecimento especialista da informação de traços e casamento (fusão da informação através de extrapolação do traço para o calorímetro para avaliação do Δ em posição) dessas duas informações. A Tabela 5.1 contém um sumário das grandezas empregadas pelo ATLAS. Um esboço da extração para candidatos a elétron com o detector encontra-se na Figura 5.1.

Dessas grandezas, adicionou-se a eProbabilityHT para a operação na *Run 2*. Isso ocorreu pois a mistura de gás empregada no TRT foi alterada para utilizar argônio ao invés de xenônio devido a vazamento do gás no tubo e por xenônio ter alto custo. A operação com argônio deteriora a capacidade de discriminação elétron-hádron do TRT e, por isso, empregou-se técnica de verossimilhança em cima das informações do TRT para recuperar essa perda.

Tabela 5.1: Definição das grandezas empregadas pelo ATLAS para discriminação de elétrons. Adaptado de [180]. Para mais detalhes, ver Seção 5.2.

Tipo	Descrição	Nome
Vazamento Hadrônico	Razão em E_T da primeira camada do calorímetro hadrônico para a E_T do <i>cluster</i> EM (empregado para a região $ \eta < 0,8$ ou $ \eta > 1,37$)	R_{had1}
	Razão em E_T no calorímetro hadrônico para a E_T a E_T do <i>cluster</i> EM (empregado para a região $0,8 < \eta < 1,37$)	R_{had}
Região Traseira do Calorímetro EM	Razão de energia da camada traseira para a energia total do calorímetro de acordeão EM. Essa variável é somente empregada abaixo de 100 GeV pois se sabe que ela é ineficiente em altas energias.	f_3
Camada Central do Calorímetro EM	Largura lateral do chuveiro, $\sqrt{(\sum E_i \eta_i^2)/(\sum E_i) - ((\sum E_i \eta_i)/(\sum E_i))^2}$, onde E_i é a energia e η_i a pseudorapidez da célula i e a soma é calculada dentro de uma janela com 3×5 células.	$w_{\eta 2}$
	Razão de energia em 3×3 células para a energia em 3×7 células centradas na posição do <i>cluster</i> do elétron	R_ϕ
	Razão de energia em 3×7 células para a energia em 7×7 células centradas na posição do <i>cluster</i> do elétron	R_η
Camada de Tiras do Calorímetro EM	Largura do chuveiro, $\sqrt{(\sum E_i (i - i_{max})^2)/(\sum E_i)}$, onde i percorre todas as tiras na janela $\Delta\eta \times \Delta\phi \approx 0,0625 \times 0,2$, correspondendo tipicamente a 20 tiras em η . i_{max} é o índice da tira com a maior deposição de energia.	w_{stot}
	Razão de energia da diferença entre os depósitos de maior e a de segunda maior energia no <i>cluster</i> para a soma dessas energias	E_{ratio}
	Razão da energia na camada de tiras para o total de energia no acordeão EM do calorímetro.	f_1
Condições do Traço	Número de pontos na camada mais interna de pixel; discriminante contra conversões de fótons.	n_{blayer}
	Número de pontos no detector de pixel.	n_{pixel}
	Número total de pontos nos detectores de pixel e SCT.	n_{Si}
	Parâmetro transverso do ponto de impacto em relação ao feixe de colisões	d_0
	Significância do parâmetro transverso do ponto de impacto definido como a razão de d_0 e sua incerteza	d_0/σ_{d_0}
Perda de momento do traço entre o ponto originário e final de medida divididos pelo momento original	$\Delta p/p$	
TRT	Probabilidade de verossimilhança baseada na radiação de transição no TRT	eProbabilityHT
	Transformação em eProbabilityHT: $TRT_PID = -\frac{\log((1.0/e^{ProbabilityHT})-1.0)}{15}$	TRT_PID
Casamento <i>cluster</i> -traço	$\Delta\eta$ entre a posição do <i>cluster</i> na camada de tiras e o traço extrapolado	$\Delta\eta_1$
	$\Delta\phi$ entre a posição do <i>cluster</i> na camada central e o traço extrapolado do ponto originário	$\Delta\phi_2$
	Definido como $\Delta\phi_2$, mas o momento do traço é reescalado para a energia do <i>cluster</i> antes de extrapolar o traço do ponto originário até a camada central do calorímetro	$\Delta\phi_{res}$
	Razão da energia do <i>cluster</i> para o momento do traço	E/p

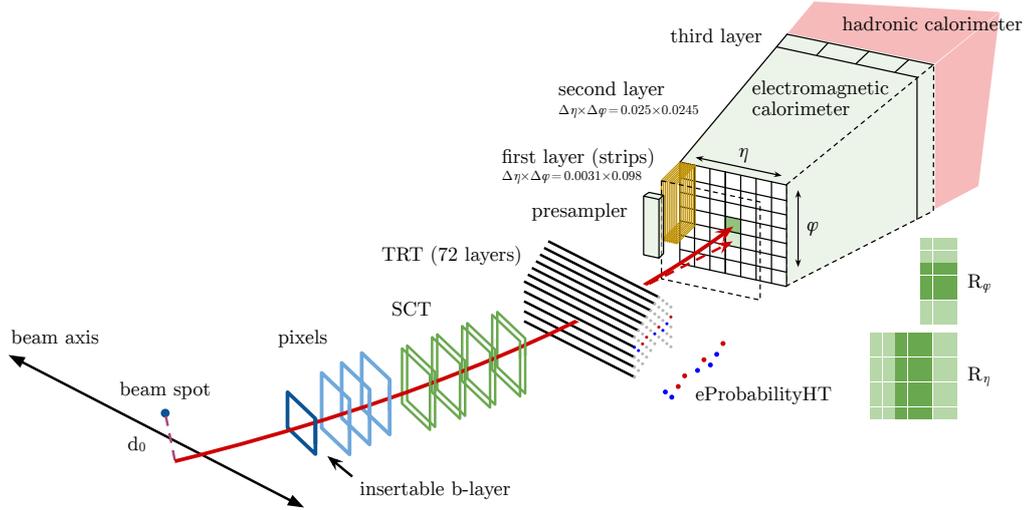


Figura 5.1: Esboço da região sensora do ATLAS originando as grandezas físicas com informação discriminante para a identificação de elétrons. [180]

Observou-se (Seção 2.4.2) que a resposta dos calorímetros depende da energia da partícula incidente. A própria natureza construtiva do calorímetro implica em alteração na resposta, considerando as diferentes granularidades em função da região em que a partícula incide e as flutuações na quantidade de material disponível para as amostragens longitudinais. Há também quantidade variável de material morto disponível anteriormente ao calorímetro e das regiões de *crack* (Figura 4.7 e 4.9). Essas variações construtivas ocorrem em função de η uma vez que o ATLAS optou por simetria em ϕ . Essa alteração tem influência no perfil das grandezas físicas geradas.

O impacto dessas variáveis no perfil da grandeza E_{ratio} está na Figura 5.2. Nessa grandeza, calcula-se a razão da diferença de energia das células de maior energia na EM1 pela sua soma. Espera-se que um verdadeiro elétron deposite a maior parte de sua energia em uma única célula da EM1 e pouca nas demais, uma vez que o chuveiro ainda não se desenvolveu, logo obtendo valores próximos a unidade. Já o ruído físico de elétrons, composto em sua maioria por jatos hadrônicos, uma vez iniciando o chuveiro nesta região, podem se expandir por uma região maior, tornando essa grandeza mais distante da unidade. O perfil do sinal é mais suscetível à alteração de acordo com a mudança dessas variáveis. Isso é esperado, uma vez que distorções irão afetar o perfil descrito pela área do calorímetro (de duas células) com maior impacto do que um perfil que naturalmente afeta uma área mais abrangente do calorímetro. Além disso, observa-se que elétrons possuem valores mais próximos a unidade na tampa, aonde a tira possui granularidade mais grossa e, logo, está sujeita a capturar maior energia nessa região. Isso, por sua vez, não garante maior poder de discriminação nessa região, necessariamente, já que os jatos hadrônicos

também tendem a obter valores igualmente mais próximos a unidade.

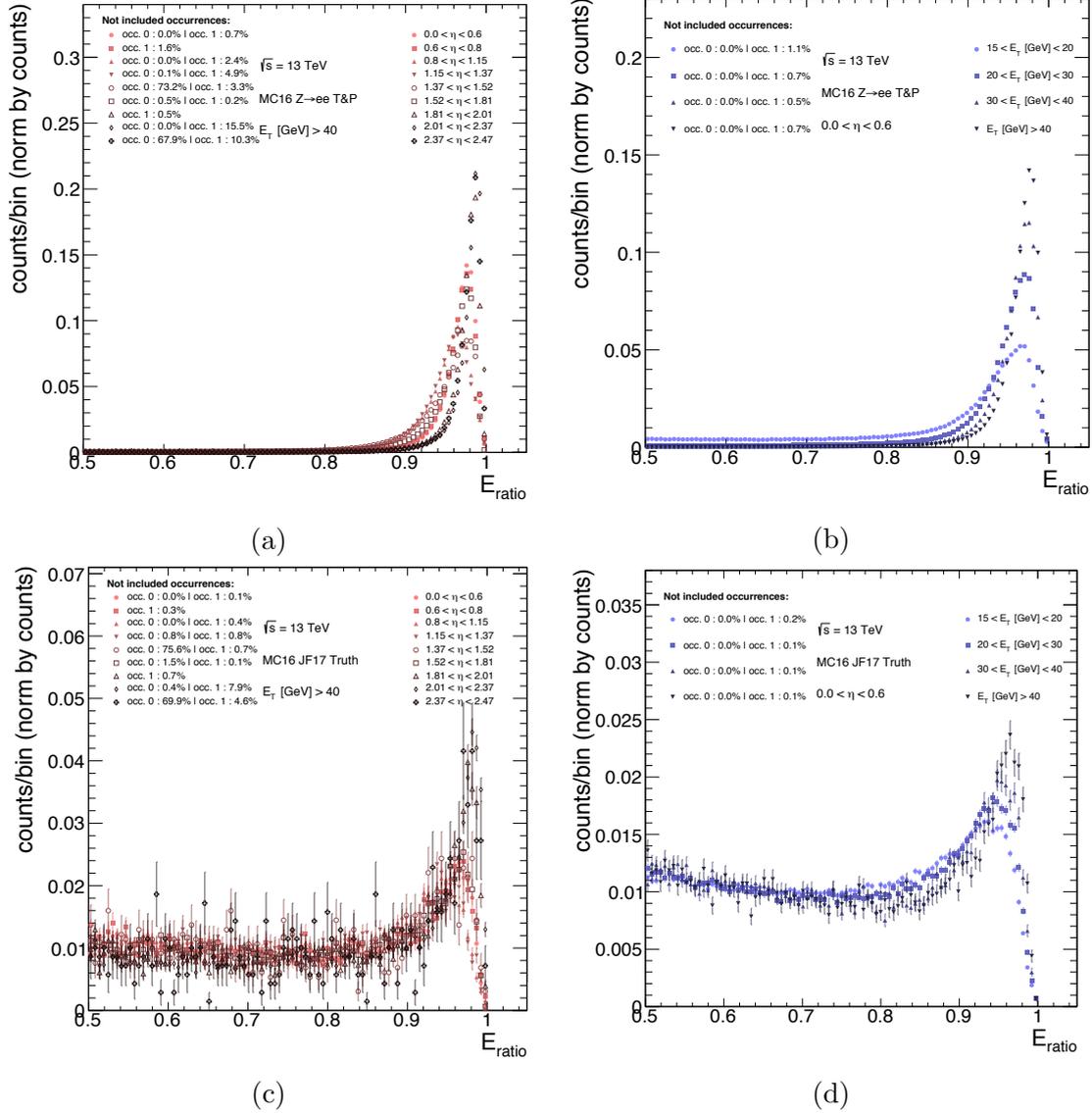


Figura 5.2: Perfil em dados de simulação com condições observadas durante colisões $p-p$ de 2016 para a grandeza E_{ratio} conforme reconstruída em *offline* para elétrons (a, b) para os de $Z \rightarrow ee$ (Seção 5.6) e ruído físico (c, d). Em (a, c), comparam-se os perfis para as regiões definidas em η para $p_T > 40$ GeV, já em (b, d) se comparam os perfis para regiões em E_T quando $0 < \eta < 0,6$. As observações estão normalizadas pelo total de entradas do histograma excluindo observações citadas no canto superior esquerdo das figuras. O E_{ratio} (Tabela 5.1) é calculado na EM1, assim essa grandeza possui limitações para sua reconstrução na região do *crack* e em $2,37 < |\eta| < 2,47$, uma vez que as tiras (células com fina granularidade em η) não estão disponíveis para essas regiões.

As grandezas de traço igualmente podem ser afetadas de acordo com essas variáveis. Um exemplo, para a grandeza TRT_PID está na Figura 5.3 para o perfil de elétrons. Observa-se que essa grandeza é menos discriminante para a região $0,6 < |\eta| < 0,8$,

onde os perfis de elétrons e ruído físico se encontram próximos de zero. Acredita-se que a região mais discriminante para essa grandeza dá-se por $1,52 < |\eta| < 1,81$ pela menor sobreposição das massas das distribuições observações. Por outro lado, a grandeza torna-se mais discriminante conforme o aumento do p_T , mostrando que a probabilidade de emissão de radiação aumenta conforme a energia da partícula.

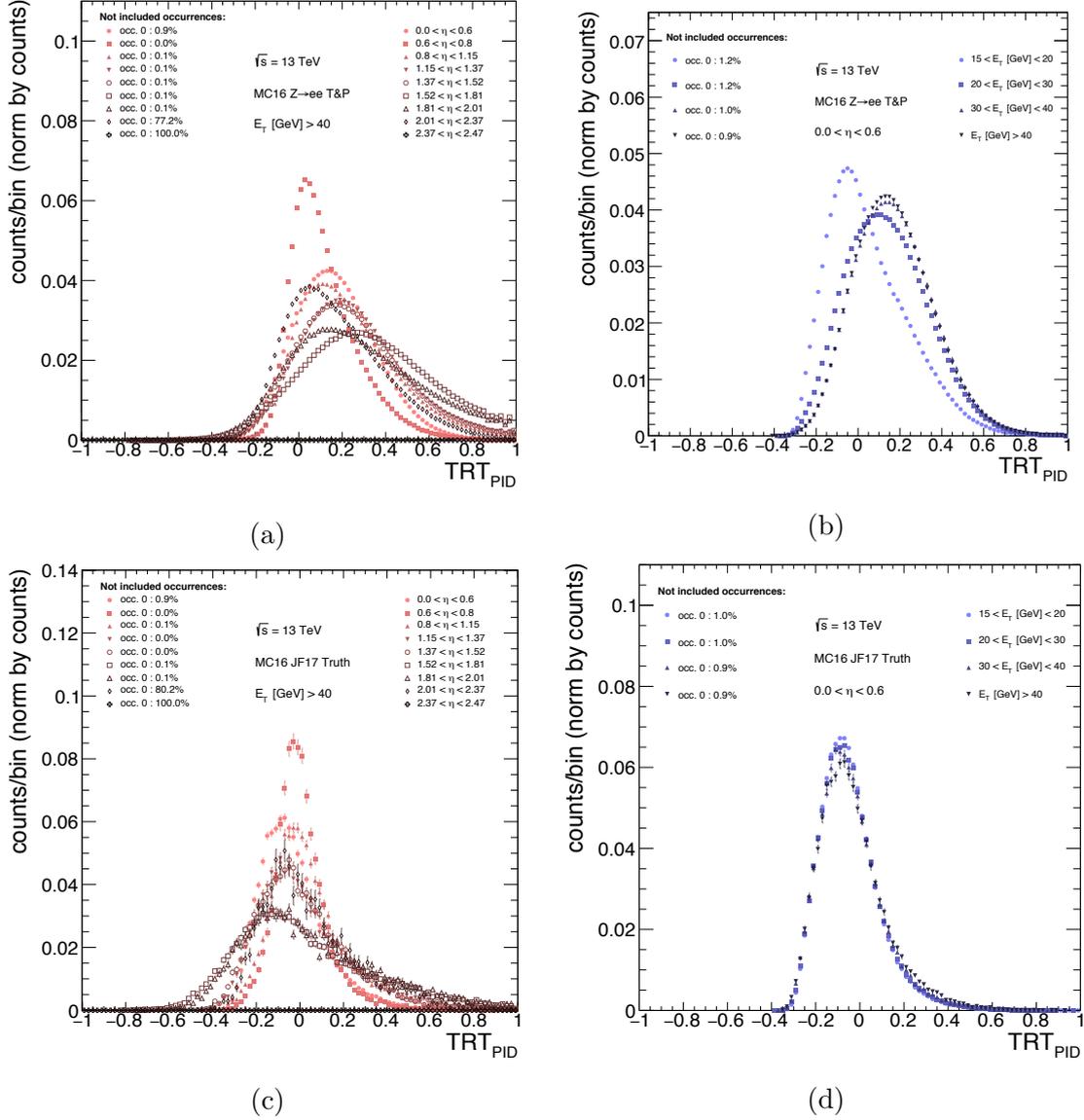


Figura 5.3: Perfil em dados de simulação com condições observadas durante colisões $p - p$ de 2016 para a grandeza TRT_{PID} conforme reconstruída em *offline* para elétrons (a, b) para os de $Z \rightarrow ee$ (Seção 5.6) e ruído físico (c, d). Em (a, c), comparam-se os perfis para as regiões definidas em η para $E_T > 40$ GeV, já em (b, d) se comparam os perfis para regiões em E_T quando $0 < \eta < 0,6$. As observações estão normalizadas pelo total de entradas do histograma excluindo observações citadas no canto superior esquerdo das figuras. O TRT_{PID} (Tabela 5.1) é obtido no TRT, assim essa grandeza perde seu poder discriminante para $|\eta| > 2,0$ (Subseção 4.1.3).

5.3 Estratégia para Discriminação e Avaliação de Performance

Mesmo após a aplicação do SF_{on} , há uma grande massa de dados contendo física ordinária que deve ser minerada a fim de prover informação útil para a análise. Na realidade, ainda que redundante, é importante frisar para aqueles não acostumados com o tema que a atuação do sistema desta maneira possibilita maior eficiência total, onde o SF_{on} atua literalmente como um filtro para atender as limitações técnicas de armazenar e processar por completo os dados. Assim, a busca por objetos de interesse ocorre no SR_{off} de forma a possibilitar a máxima eficiência do sistema composto pela decisão conjunta do SF_{on} e SR_{off} . Naquele, a decisão, uma vez realizada, é inexorável, enquanto o SR_{off} é flexível, onde cada análise avalia a estratégia que lhe for mais conveniente, e é possível o reprocessamento dos dados conforme a sua atualização.

Elétrons constituem-se de ótimos candidatos para a exploração de física no ATLAS. Os mesmos constituem uma assinatura mais limpa, com boa medição de suas propriedades, e participam de canais para a busca de nova física [184]. A esmagadora maioria do ruído desses elétrons são jatos hadrônicos LF, cujas assinaturas no ATLAS se diferenciam principalmente pela geração de chuveiros hadrônicos. Além disso, quando se refere à exploração de física com elétrons, geralmente se refere a elétrons isolados. Essa distinção se faz necessária devido às outras duas fontes de ruído físico, dadas por outros elétrons, não isolados, a partir de decaimentos de fótons convertidos ($\gamma \rightarrow e^-e^+$)², a segunda maior fonte de ruído físico, e jatos HF (como $b \rightarrow c^+e^- + \bar{\nu}_e$). Os fótons convertidos costumam apresentar maior d_0 (uma vez que o traço será distorcido pela trajetória de outra partícula), pior associação entre as informações de calorimetria e traço e pequenas distorções no desenvolvimento do chuveiro que podem ser empregadas para a distinção de elétrons (isolados). No caso de jatos HF, o elétron terá o desenvolvimento de componentes hadrônicas de chuveiro devido às demais partículas dessa natureza, mas também a presença de um pior d_0 .

Em determinadas análises, realiza-se uma comparação entre as seções de choque provenientes da predição em simulação com aqueles observados. Para que essa comparação seja possível, restrições são aplicadas na região geométrica onde o experimento é capaz de fornecer informação relevante (como a região de precisão $|\eta| < 2$, 47 para elétrons) e em regiões cinemáticas para limitar o impacto do ruído de fundo nas análises. Além disso, necessita-se levar em conta a influência dos detectores do experimento nas ineficiências de identificação. Em um dos tipos de medição (medições

²Outra assinatura que pode comportar-se de forma similar apresenta-se por decaimentos $\pi^0 \rightarrow e^+e^-\gamma$, conhecido por decaimento Dalitz

inclusivas fiduciais de seções de choque), isso pode ser realizado através de um fator de correção para colisão e simulação por regiões parciais do espaço das variáveis que influenciam a análise (espaço de fase). Esses fatores compreendendo as expectativas de distorções em eficiência em simulação e colisão são empregados para realizar o *unfolding* [185] da reconstrução nos dois casos e obter as distribuições fundamentais das grandezas em análise.

Assim, os métodos empregados nas análises físicas para identificação de objetos físicos costumam considerar medições em regiões do espaço de fase definido pelo plano $E_T \times \eta$. Uma vez que há alteração da resposta dos detectores e faça parte das análises o particionamento em regiões desse espaço de fase, é natural que os métodos de identificação considerem a aplicação de cortes de decisão e extração de parâmetros também em regiões.

Na Física de Altas Energias, é comum a criação de grandezas altamente discriminantes para a aplicação de cortes lineares (Seção 3). No ATLAS, essa estratégia (*CutBased*) foi a referência para as análises até 2012. Ainda assim, ela está disponível para o emprego nas análises para físicos que prefiram manter suas análises empregando essa estratégia por ser uma abordagem clássica ao problema. Essa estratégia emprega as grandezas descritas na Seção 5.2 com limiares definidos para regiões no espaço de fase com a grade definida na Tabela 5.2 e otimizadas em base de simulação³.

Porém, com o aumento do empilhamento e, portanto, demanda por abordagens mais eficientes, o ATLAS alterou sua referência para as análises envolvendo elétrons para uma estratégia multivariável a partir daquela data. Essa estratégia realiza a extração de estimativas das *pdf* marginais para essas grandezas físicas (\vec{x}) construídas usando KDE Gaussiano [186]. Os *kernels* resultantes (atualmente obtidos em dados de colisões de 2016), são aproximados por histogramas com fina granularidade (500 grupos) que são interpolados⁴ para dar os valores de probabilidade do candidato ser oriundo da classe de sinal (ruído) para a i -ésima grandeza ($P_{s(b),i}(x_i)$). O valor do discriminante é computado por

$$d_{\mathcal{L}} = \frac{\mathcal{L}_s}{\mathcal{L}_s + \mathcal{L}_b}, \quad (5.1)$$

onde

$$\mathcal{L}_{s(b)}(\vec{x}) = \prod_{i=1}^n P_{s(b),i}(x_i). \quad (5.2)$$

³O último ajuste para esse método foi realizado na base de dados *mc15c*, contendo um *plateau* de luminosidade até $\langle \mu \rangle = 60$.

⁴Essa estratégia foi adicionada relativamente recentemente no ATLAS e é referida pelo termo *smooth*. Anteriormente o valor armazenado no histograma era empregado para o cálculo da verossimilhança.

A decisão é avaliada no espaço de saída descrito por $d_{\mathcal{L}}$ através de um limiar linearmente corrigido para levar em conta o efeito de empilhamento (Seção 5.5). Adicionalmente, outros critérios com cortes rígidos podem ser empregados para melhorar a eficiência final. As regiões empregadas em η estão na Tabela 5.2. Realizou-se um esboço do comprimento do calorímetro em cada uma dessas regiões na Figura 5.4.

Tabela 5.2: Cobertura das regiões em η empregadas para as análises realizadas pela ATLAS. A descrição geral é apenas uma das propriedades da região, havendo outras propriedades que podem interferir na caracterização da partícula na região.

Nomenclatura	Cobertura	Descrição Geral
$\eta_{0,wh}$	$0 < \eta < 0,6$	Barril central
$\eta_{1,wh}$	$0,6 < \eta < 0,8$	TRT_PID com menor capacidade discriminante
$\eta_{2,wh}$	$0,8 < \eta < 1,15$	Barril estendido do <i>TileCal</i>
$\eta_{3,wh}$	$1,15 < \eta < 1,37$	Transição barril para região do <i>crack</i>
$\eta_{4,wh}$	$1,37 < \eta < 1,52$	Região do <i>crack</i>
$\eta_{5,wh}$	$1,52 < \eta < 1,81$	Recuperação após <i>crack</i> na tampa
$\eta_{6,wh}$	$1,81 < \eta < 2,01$	Granularidade da EM1 em η de 0,004
$\eta_{7,wh}$	$2,01 < \eta < 2,37$	Granularidade da EM1 em η de 0,006
$\eta_{8,wh}$	$2,37 < \eta < 2,47$	Alto η e ausência de tiras na EM1

Tabela 5.3: Fronteira das regiões em E_T empregadas pelo ATLAS para avaliação de eficiência, aplicação de limiares de decisão e ajuste dos modelos da *LLH*.

Fronteira das regiões em E_T [GeV]										
Avaliação de Performance ($Z \rightarrow ee$ T&P)										
15	20	25	30	35	40	45	50	60	80	150
Cortes de Discriminação										
15	20	25	30	35	40	45	50			
Ajuste de Modelo										
15	20	25	30	35	40					

5.4 Critério de Isolamento em Elétrons

Adicionalmente aos critérios de seleção de elétrons descritos na Seção 5.3, outro critério adicional considerando o conhecimento especialista pode ser aplicado a fim de obter melhor eficiência na seleção de elétrons de interesse. Isso se dá através de aplicação de cortes lineares em variáveis de isolamento, que provém da energia de partícula produzidas em torno do candidato a elétron e permitem a distinção de elétrons isolados (a partir de decaimentos de $W \rightarrow e\nu$ e $Z \rightarrow ee$) de outros, elétrons não-isolados, como aqueles originando de conversões de fótons produzidos em decaimentos hadrônicos, elétrons oriundos de decaimentos de sabores pesados

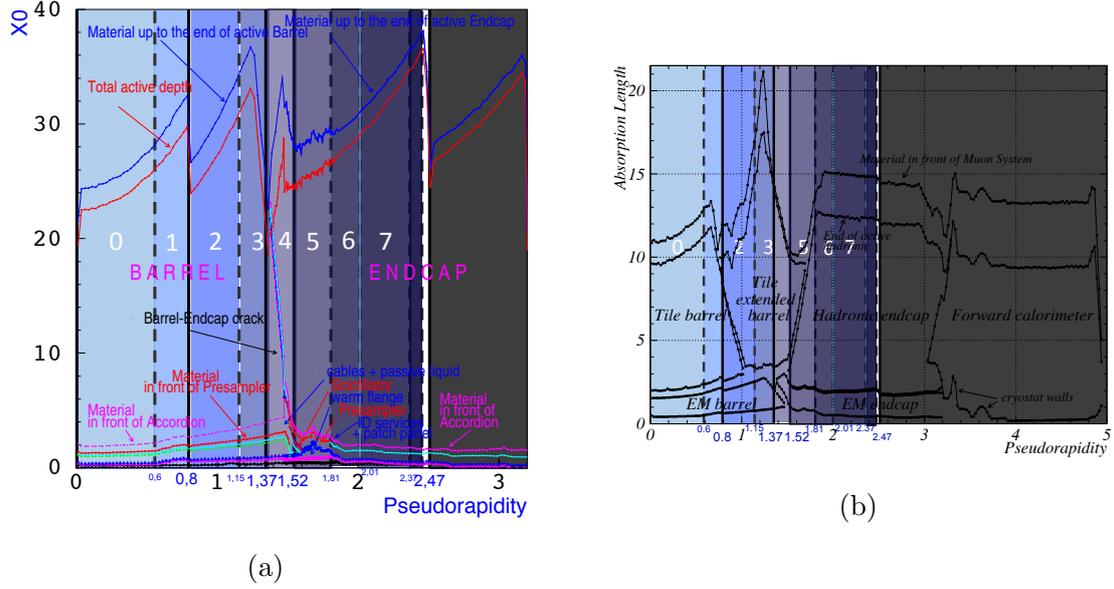


Figura 5.4: Esboço do comprimento simulado em unidades de X_0 do ECAL (a) e λ_{int} do Sistema de Calorimetria (b) contido nas regiões em η empregadas para medição de eficiência da tarefa de discriminação elétron-hádrão. Os índices dentro das faixas azuis indicam o número da região em η , respectivamente $\eta_{0,ilh}$, $\eta_{1,ilh}$, \dots , $\eta_{8,ilh}$. Omitiu-se a região $\eta_{8,ilh}$ em ambas figuras por questões visuais. Os limites em η no eixo x na cor azul descrevem as fronteiras das regiões. As linhas verticais contínuas indicam fronteiras comuns ou aproximadamente comuns à essa grade e aquela empregada para o *NeuralRinger online*. Já as linhas trajadas indicam fronteiras contidas somente nas análises do ATLAS.

(t etc.) e falso alarme de jatos LF [180]. Duas variáveis foram desenvolvidas pelo ATLAS para esse propósito:

- isolamento em *calorimetria* ($E_T^{\text{cone},\Delta R}$): definido pela soma em E_T dos *clusters* topológicos⁵ [166] dentro de um cone a ser definido por ΔR . A energia E_T contida em um *cluster* retangular de tamanho $\Delta\eta \times \Delta\phi = 0,125 \times 0,175$ em torno do baricentro de energia é subtraída. Uma correção dependente em E_T e η é aplicada para considerar o vazamento de energia. A contribuição de empilhamento e de eventos adjacentes é corrigida para cada evento usando a técnica em [187];
- isolamento variável em *traço* ($p_T^{\text{varcone},\Delta R}$): definido como a soma dos momentos de traços, satisfazendo os critérios de qualidade de traços, dentro de um cone de $\Delta R = \min(0,2,10[\text{GeV}/E_T])$ em torno do candidato a traço do elétron e originando do vértice primário, excluindo os traços associados ao

⁵ *Clusters* não retangulares, que crescem dinamicamente conforme a presença de células vizinhas considerando unidades de significância em relação ao ruído esperado nas células. Aplicam-se três limiares, um destinado à encontrar sementes para a geração de *clusters*, outro destinado ao crescimento do *cluster* e um terceiro limiar dedicado à extensão em que o algoritmo estende a busca para realizar a junção entre *clusters* vizinhos.

elétron (por exemplo, o traço principal do elétron e outros adicionais devido a fótons convertidos por *bremsstrahlung*).

5.5 Resiliência ao Efeito de Empilhamento

O congestionamento de sinal no detector afeta a reconstrução de elétrons, havendo maior probabilidade de sobreposição de sinais provenientes de partículas distintas em regiões próximas e, com isso, ocorrendo interferência durante o desenvolvimento dos sinais (empilhamento). Devido à alta taxa de eventos operada pelo LHC, o empilhamento de sinais não ocorre necessariamente de colisões $p - p$ oriundas do mesmo evento de cruzamento de pacotes do LHC já que o sinal (por mais que os detectores sejam otimizados para terem resposta rápida) desenvolve-se durante vários ciclos.

Há uma grande probabilidade do evento contaminando a janela de análise ser proveniente de jatos hadrônicos, uma vez que o LHC atua como uma fábrica de jatos hadrônicos (Figura 2.6). Como dito, o empilhamento afeta tanto elétrons quanto ruído físico, mas, do ponto de vista de discriminação, decorre uma descaracterização grave para elétrons tanto em calorimetria quanto no traço. No caso do calorímetro, uma contaminação da região em análise pela deposição de energia de outra partícula pode causar, dependendo de sua intensidade e posição de incidência, a descaracterização do processo de desenvolvimento de um chuveiro EM, que sofrerá com contribuições hadrônicas em regiões não centradas no baricentro de energia aumentando o perfil lateral. Longitudinalmente, ainda que essas contribuições sejam provenientes de partículas com, geralmente, baixo p_T , as mesmas podem se alastrar até o HCAL devido à sua natureza hadrônica. No traço, a presença de outra partícula próxima ao elétron pode torná-lo parecido com um jato, constituído de diversas partículas e normalmente contém pior precisão no traço reconstruído. Todavia, estudos anteriores mostram que o efeito de empilhamento afeta mais fortemente as grandezas de calorimetria R_{had} e R_η [181]. No caso de R_{had} isso se deve ao vazamento longitudinal de energia para o HCAL e em R_η pelo vazamento lateral [7].

Dado as considerações acima, há maior efeito nos padrões de elétrons, de forma que o efeito observado durante o desenvolvimento da LLH deu-se por uma queda (linear) de eficiência de detecção conforme o aumento da incidência de empilhamento. Cabe ressaltar que o empilhamento não afeta igualmente todos os eventos de cruzamento de pacotes. A partir de uma abordagem estatística, pode-se modelar uma maior probabilidade de ocorrência de empilhamento de acordo com estimativas de número médio de interações por cruzamento de pacote ($\langle \mu \rangle$) e número de vértices ($nvtx$)⁶. As análises durante o desenvolvimento mostraram que o $nvtx$ é um esti-

⁶Existem outros estimadores [35], como a ocupação do (calorímetro de) LAr e ocupação do

mador mais apropriado para caracterizar o efeito de empilhamento, havendo menor flutuação de eficiência em função do mesmo quando comparado com o emprego de $\langle\mu\rangle$. Todavia, O segundo não está disponível no SF_{on} por demandar alta latência para o seu calculo quando considerando as análises baseadas em ROI empregadas nas cadeias de elétrons (Seção 5.7). Assim, o $\langle\mu\rangle$, calculado *online* pelo *LUCID* (Seção 4.5), é empregado para a estimativa do empilhamento no SF_{on} .

Assim como as análises físicas consideram o efeito da alteração da resposta do ATLAS em função do espaço de fase, o efeito de empilhamento de sinais devem ser levado em conta nas análises. No caso da identificação de elétrons, a queda em eficiência apresentou-se linearmente até recentemente. A fim de lidar com isso, aplicou-se uma correção linear no limiar de decisão conforme o estimador de empilhamento. Até o final de 2017, essa correção era obtida através de uma reta para o balanceamento da eficiência de detecção calculada elétrons abaixo e acima de um valor do estimador. Mais recentemente, esse método foi atualizado para compreender um ajuste linear em função dos limiares responsáveis pela obtenção da eficiência de detecção desejada em função do estimador de empilhamento para uma grade espaçada de 1 unidade do estimador (Figura 7.4). O ajuste linear avalia [188]

$$\chi^2 = \frac{(y - f(x))^2}{e_y^2 + (0,5(e_{xl} + e_{xh})f'(x))^2}, \quad (5.3)$$

onde e_y é o erro inferior das ordenadas (superior) caso o $f(x)$ esteja abaixo (acima) de y e e_{xl} (e_{xh}) é o erro inferior (superior) nas abscissas. Essas correções causam um crescimento de probabilidade de falso alarme (P_F) em cortes mais exigentes (*medium* e *medium*), assim o processo passa por iterações manuais de relaxamento do coeficiente angular até que se atinja um equilíbrio entre a correção e um nível aceitável de crescimento de falso alarme.

5.6 Método *Tag and Probe*

O T&P [180]⁷ permite obter uma pré-seleção de amostras de interesse sem a aplicação de discriminadores a partir de conhecimento especialista. De forma genérica, consideram-se determinados processos físicos bem conhecidos que geram objetos de interesse, no caso elétrons, para aplicar uma série de critérios e, com isso, obter observações compatíveis com o processo físico de interesse. Isso possibilita lidar com a questão cega para análises em dados de colisão, ou seja, como determinar

TRT.

⁷Ao seguir as referências citando a técnica, encontra-se o trabalho [189] que se acredita ser a primeira referência empregando a técnica. Todavia, o trabalho não usa a terminologia aqui empregada, apenas se referindo ao método. Assim, preferiu-se manter como referência o trabalho mais recente de elétrons envolvendo a técnica para medição de eficiência de elétrons.

a eficiência ou realizar o ajuste de técnicas cujo objetivo é determinar a natureza das mesmas sendo que não se sabe a natureza das amostras. Ao se aplicar um outro discriminador para análise se torna tendencioso em relação a esse discriminador. Como as análises físicas realizam comparações entre simulações e dados de colisão, as medições de eficiência e ajustes de modelos em simulações empregam o mesmo procedimento para garantir uniformidade entre os casos.

No caso de elétrons, os decaimentos empregados pelo ATLAS são $Z \rightarrow ee$ (Figura 5.5), $Z \rightarrow ee\gamma$ e $J/\Psi \rightarrow ee$. A descrição limita-se ao método T&P para decaimentos de Z não radioativos ($Z \rightarrow ee$), a única empregada neste trabalho. Exige-se que o evento de cruzamento de pacotes contenha ao menos dois candidatos a elétrons reconstruídos pelo *offline*. Procura-se no evento por pares de elétrons que atendam o seguinte critério: um dos candidatos (*tag*) seja aceito como elétron ao empregar todo o conhecimento disponível da análise *offline*; enquanto o outro candidato (*probe*) precisa formar, em conjunto com o *tag*, as propriedades conhecidas da partícula. Este método fornece elétrons de maneira não-tendenciosa com grande ($> 99\%$) pureza para $E_T > 30$ GeV, porém está sujeito a contaminação das amostras.

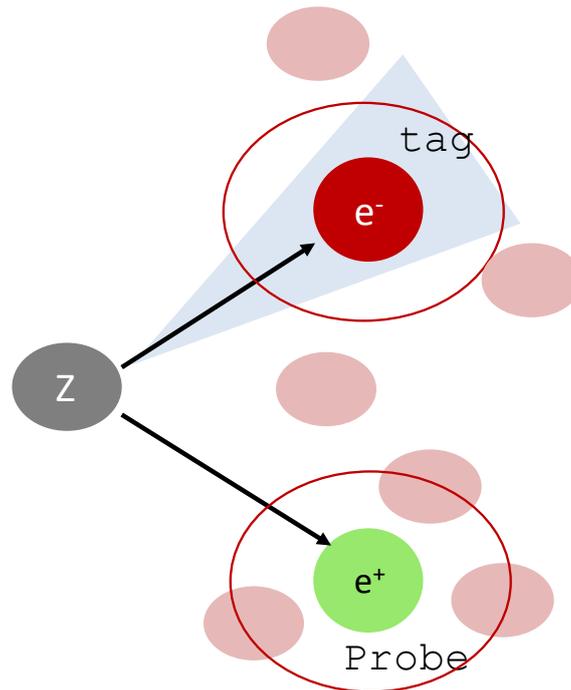


Figura 5.5: Esboço do T&P para o bóson Z decaindo em dois elétrons.

Neste trabalho, o método T&P implica, exceto quando especificado, na exigência dos seguintes critérios:

- exigência de ao menos dois candidatos a elétrons não oriundos de conversão de fótons;

- o evento é aceito pelas cadeias primárias dedicadas à coleta de eventos contendo ao menos um elétron de menor energia sem a aplicação de pré-escala. As mesmas correspondem aos seguintes critérios:
 - $E_T > 26$ GeV, aplicação de isolamento no HLT e critério critério de seleção de elétron *medium*;
 - $E_T > 60$ GeV e aplicando critério de seleção de elétron *medium*;
 - $E_T > 140$ GeV e aplicando critério de seleção de elétron *loose*;
 - $E_T > 300$ GeV, nenhum critério de seleção de elétrons.
- o *tag* não se encontra na região do *crack*;
- exige-se $E_T > 25$ GeV no *tag*;
- exige-se isolamento no *tag* em $p_T^{\text{varcone},\Delta R}$ com $\Delta R = 0,2$ e limiar de $0,1 \times p_T$ GeV (Seção 5.4);
- em dados de simulação, exige-se que o *tag* e o *probe* sejam elétrons de acordo com a verdade de MC;
- aplica-se isolamento do de jatos com $E_T > 20$ GeV e $\Delta R > 0,4$ (usando a seleção de jatos *offline*);
- o *tag* e o *probe* devem possuir OS;
- a massa obtida com o *tag* e o *probe* deve estar dentro da região $[80, 100]$ GeV.

5.7 Cadeias para Filtragem de Elétrons

Os objetos físicos de interesse do ATLAS são filtrados através de algoritmos de reconhecimentos de padrões. Os padrões, por sua vez, são obtidos através da reconstrução física que ocorre *online*. A fim de possibilitar atender a latência exigida pela alta taxa de eventos, a reconstrução é realizada em diversos estágios sequenciais possuindo maior profundidade de análise física e, igualmente, exigindo maior processamento. Os estágios de seleção podem fazer parte da estrutura base do TDAQ (Seção 4.4), ou por parte de uma especialização da cadeia para a coleta dos objetos físicos de interesse. No caso de elétrons, um esboço dos estágios aplicados na filtragem estão disponíveis na Figura 5.6.

A informação relevante para a identificação e, logo, filtragem de elétrons provém do Sistema de Calorimetria (Seção 4.2) e do ID (Seção 4.1). O processamento deste exige maior esforço computacional que aquele devido a dois fatores: há a disponibilidade de ordens superiores de canais a serem acessados no ID, comparando

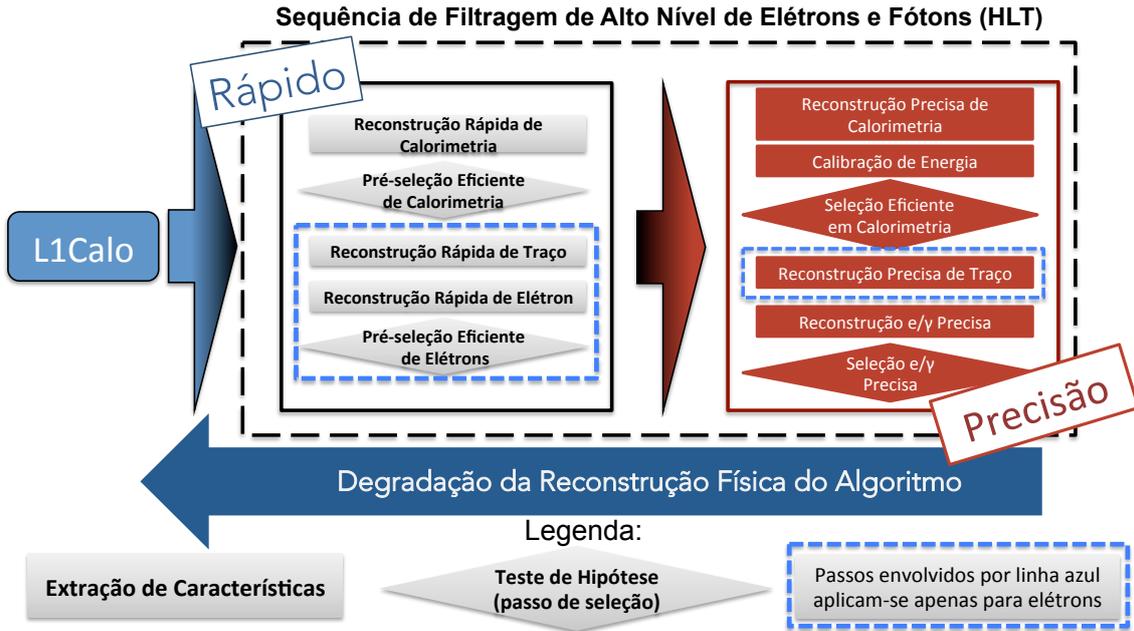


Figura 5.7: A sequência de reconstrução e seleção empregada em cadeias de seleção de elétrons e fótons durante a Segunda Temporada de Coleta de Dados. As cadeias de fótons não executam as etapas relacionadas a reconstrução e seleção com informação de traço (envolvidas por linha tracejada em azul). Um total de 4 (5) decisões são aplicadas para fótons (elétrons), cada nível reconstruindo em maior grau a informação a fim de reduzir a taxa disponível a ser processada no próximo estágio. Reconstrução e seleção em baixa latência ($2,5 \mu\text{s}$) é realizada pelo *L1Calo* no Sistema de Calorimetria. Posteriormente, o HLT opera em uma etapa rápida e eficiente para então atuar a etapa precisa com a aplicação de algoritmos totalmente baseados na estratégia de seleção *offline*. Em cada etapa, aplica-se para as cadeias de elétrons um estágio de pré-seleção avaliando somente informação de calorimetria a fim de remover eventos adicionais antes do processamento da informação do ID. As etapas de seleção são mencionadas através das seguintes nomenclaturas, ordenados conforme a aplicação na cadeia: FastCalo, FastElectron, HLTCalo e HLT.

algoritmos específicos, como o exemplo da avaliação por fótons convertidos (fótons que interagem com o material do ID e se convertem em elétrons).

O processamento no *L1Calo* (Figura 5.8) ocorre via algoritmo de janela deslizante para a separação de RoIs no Sistema de Calorimetria. Para tal fim, a informação empregada pelo *L1Calo* faz mão de Torre de Filtragem (TT) que se expandem lateralmente em uma região de $0,1 \times 0,1$ no plano $\eta \times \phi$ e por todo o calorímetro longitudinalmente. O algoritmo do *L1Calo* busca pelo máximo local de energia com a deposição de energia calculada na região central EM calculada em 2×2 TT no ECAL. De posse das RoI, o *L1Calo* aplica um corte em E_T que pode ser variável conforme η ⁸. Outros dois critérios podem ser aplicados para cadeias de elétrons

⁸As seguintes letras são adicionadas para indicar as possíveis variações de seleção do *L1Calo*: variação do limiar aplicado em E_T conforme η (V); aplicação de limiar em isolamento eletromagnético

destinadas a coleta :

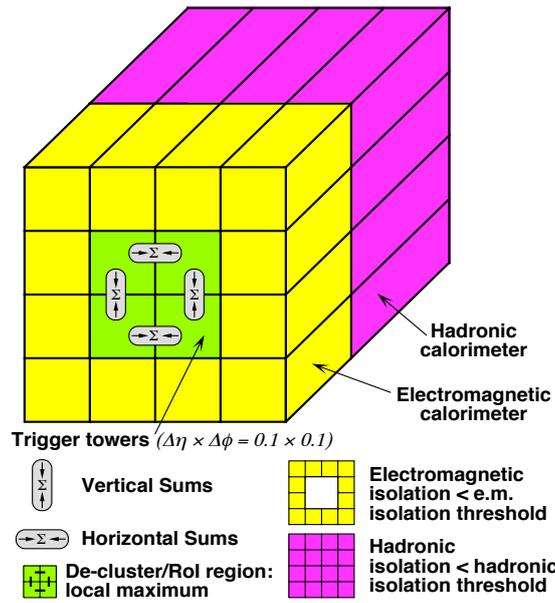


Figura 5.8: A lógica da janela deslizante aplicada no L1 para seleção de elétrons e fótons. Na figura estão assinaladas a região central (verde), a região de isolamento EM (amarelo) e a região central e de isolamento HAD (roxo). O valor de energia das TT na região central é somada para fornecer a energia local [190].

- isolamento eletromagnético, a energia depositada atrás da região central EM, computada em 2×2 TT no HCAL; e
- isolamento hadrônico, similar ao isolamento EM, mas para as torres equivalentes no HCAL.

RoI aceitas pelo *L1Calo* são examinadas posteriormente no HLT por cadeias específicas. As lógicas aplicadas no HLT dependem do menu empregado, podendo haver uma coletânea de cadeias alimentadas por uma mesma RoI aceita pelo *L1Calo*. Uma seleção em calorimetria compõe o primeiro estágio da etapa rápida do HLT. Aplica-se um corte em energia reconstruída em uma calibração⁹ menos precisa exigindo um limiar de 3 GeV abaixo daquele empregado na etapa precisa, a fim de evitar a perda prematura de elétrons devido à menor precisão. Essa energia é computada em *clusters* retangulares de células da EM2 contendo 3×7 (dimensões no plano $\eta \times \phi$) no barril ($\eta < 1,4$) e 5×5 na região remanescente. A seleção inicialmente proposta para a Segunda Temporada de Coleta de Dados incluía a aplicação de cortes lineares em três grandezas físicas de calorimetria altamente discriminantes (R_{had} , R_η e E_{ratio} , definições e considerações na Seção 5.2). A seguir, aplica-se a

(I); aplicação de limiar em isolamento hadrônico (H).

⁹Inclusive, o código de calibração foi implementado em 2010 pelo autor deste trabalho como parte de suas primeiras tarefas no ATLAS.

reconstrução e extrapolação simples de traços até a EM2 exigindo que ao menos um traço possua $\Delta\eta < 0,2$ da posição de incidência de sua extrapolação em relação à medição obtida no *cluster* e $p_T > 1$ GeV.

O estágio preciso aplica reconstrução baseada em algoritmos de reconstrução *offline* (Seção 5.1 e 5.3). Assim, o estágio limitado à calorimetria realiza calibração multivariável como avaliada pelo *offline* [191]. O limiar aplicado pela cadeia nesse estágio é 1 GeV inferior daquele que se deseja coletar elétrons para estudo *offline*. Durante a operação em 2015 e 2016, aplicou-se a estratégia *LLH* para as cadeias primárias de elétrons em cima das grandezas de calorimetria. Isso se deu para aumentar a harmonia entre a seleção *online* e a estratégia *offline* enquanto possibilitando melhor alocação da banda ao possuir menor taxa de falso alarme para a mesma eficiência de sinal. Extrapolação precisa de traços ocorre em seguida, exigindo $\Delta\eta < 0,05$ e $\Delta\phi < 0,05$. A decisão final emprega a estratégia *LLH* com as mesmas grandezas físicas que o *offline*, exceto a d_0 que foi removida por questões de robustez sem perda de eficiência.

As figuras 5.9a e 5.9b evidenciam o efeito na taxa das cadeias em função da luminosidade para os estágios de seleção do *L1Calo* e HLT, respectivamente. A utilização do *LLH* forneceu parte da redução da taxa necessária, porém outras atualizações foram sendo adicionadas a fim de aliviar a sempre crescente luminosidade de operação do LHC. Abordam-se as atualizações realizadas durante o desenvolvimento na Subseção 5.7.1. Uma dessas atualizações envolve diretamente o desenvolvimento deste trabalho, havendo uma discussão sobre as questões na Subseção 5.7.2.

5.7.1 Atualizações durante a *Run 2*

Após as implementações das melhorias para a operação na Segunda Temporada de Coleta de Dados, a cadeia para seleção de elétrons manteve-se operante sem modificações nos algoritmos em atuação durante 2015 e 2016 (Figura 5.10). Todavia, estudos ocorridos durante 2016 e no início de 2017 mostraram que o consumo necessário para processamento nas condições de empilhamento esperados para 2017 ($\langle\mu\rangle = 60$ colisões) não era sustentável e impediriam a operação do ATLAS em uma luminosidade de $2,0 \times 10^{34} \text{cm}^{-2}\text{s}^{-1}$. As condições tornaram-se ainda mais severas devido ao fato da reposição de parte da fazenda de processamento do HLT não ter sido realizada, uma vez que isso exigiria uma elevação de consumo de energia de 20 %. Isso levou ao escrutínio do consumo de processamento do menu e dos algoritmos individuais compondo-o para garantir a operação do Sistema de Filtragem do ATLAS em ordem de luminosidade sem precedentes.

Ao considerar ganhos de processamento na fazenda, os maiores impactos provêm de melhorias em estágios iniciais de seleção das cadeias que podem ser realizadas em

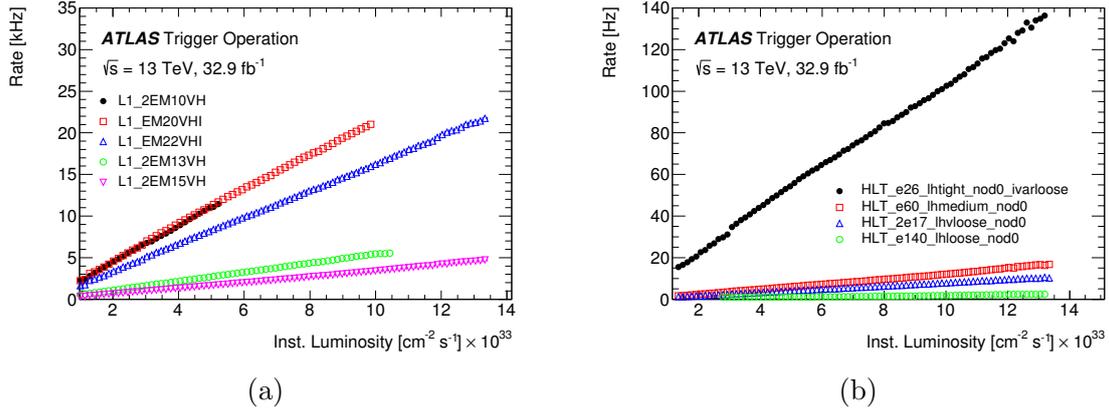


Figura 5.9: Taxa de saída do *L1Calo* (a) e HLT (b) para cadeias de seleção de elétrons em função da luminosidade instantânea não calibrada computada *online* durante o período completo de coleta de colisões $p - p$ com centro de massa de 13 TeV em 2016. As taxas mostradas em ambos os casos consideram cadeias sem aplicação de pré-escala. Todas as cadeias mostram dependência linear em função da luminosidade. Em (a), as cadeias cujas taxas estão representadas em pontos vermelhos e azul dedicam-se a coleta de eventos com ao menos um elétron, enquanto as demais cadeias exigem a presença de ao menos dois elétrons. Em (b), todas as cadeias aplicam o critério de isolamento calculado como a soma do p_T de todos os traços dentro de um cone variável em torno do elétron, excluindo o seu traço principal, divididos pelo E_T do *cluster*. $\sum p_T/E_T < 0,1$ também é aplicado para a cadeia de menor E_T (pontos pretos). A cadeia representada por pontos em azul dedica-se a coleta de eventos contendo ao menos dois elétrons, enquanto as demais dedicam-se a coleta de eventos com um único elétron. [192]

dois aspectos: tornar os algoritmos mais rápidos e/ou mais eficientes. O objetivo era obter eficiência suficiente para evitar a elevação do limiar em energia, o que seria prejudicial para a física.

No *L1Calo*, adicionou-se um novo critério de operação, mais seletivo que possibilitou a redução da taxa em 10,8 % da cadeia com uma perda de 1 % na eficiência para o principal item do *L1Calo* empregado no menu. No *FastCalo*, considerou-se alterar a estratégia *CutBased* nas três grandezas físicas (Seção 5.7) para o emprego de uma abordagem aplicando Aprendizado de Máquina (*ensemble* de MLP) alimentada por informação de calorimetria formatada em anéis concêntricos, o Algoritmo *NeuralRinger* desenvolvido como parte desse trabalho (Seção 7.1). O emprego do Algoritmo *NeuralRinger* no estágio *online* de pré-seleção rápida baseada em informação de calorimetria (*FastCalo*) possibilitava obter redução equivalente ou ligeiramente superior à seleção empregando *LLH* em grandezas de calorimetria no *HLTCalo* enquanto atuando com eficiência equivalente de seleção. Porém, por atuar posteriormente ao *FastElectron*, a atuação do *HLTCalo* implicava em uma redução de apenas 9 ms/evento (valores centrais) e implicava em uma perda de

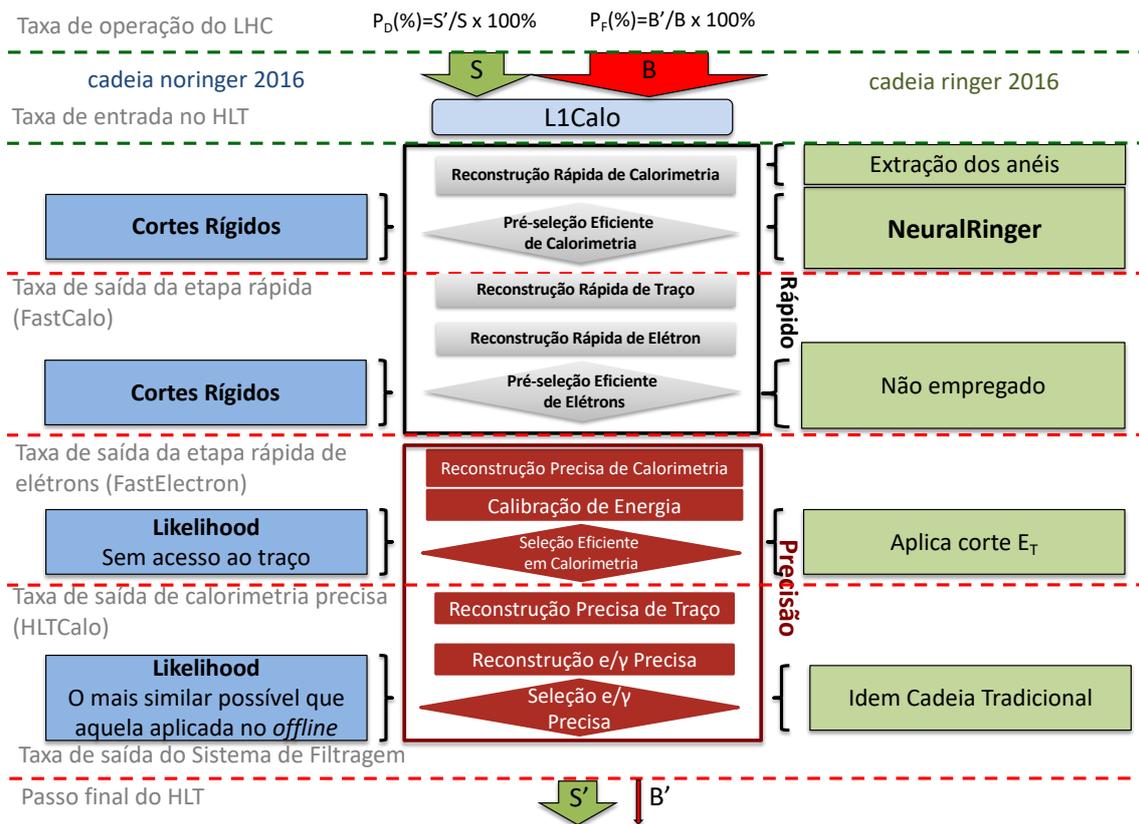


Figura 5.10: Estágios de seleção avaliados pela cadeia original (*noringer*, azul) considerando o menu de 2016. A cadeia proposta (*ringer*, verde) encontrava-se em desenvolvimento neste estágio.

4 % de eficiência. Essa redução pífia de processamento se dava pela execução após o FastElectron, um algoritmo extremamente exigente (em 2016, 55 ms/evento). O emprego do *NeuralRinger* possibilitava uma redução de 40 ms/evento exigindo apenas 0,15 ms/chamada, em comparação com os 1,7 ms/chamada do HLTCalo. Considerando que era necessário reduzir o processamento, não optar pelo *NeuralRinger* exigiria o emprego de cortes mais rígidos no *CutBased* atuando no FastCalo e implicaria na perda de eficiência das cadeias. Assim, o *NeuralRinger* apresentava-se como uma solução que possibilitava a manutenção de eficiência e atender as demandas de processamento.

Dado o grande impacto em processamento e a necessidade do HLT, colocou-se a cadeia com o Algoritmo *NeuralRinger* para comissionamento e posterior atuação como parte das cadeias primárias para filtragem de elétrons. O FastElectron foi otimizado para empregar os cortes na ordem mais discriminante e o arranjo do menu foi alterado de forma a possibilitar que as cadeias compartilhassem dos mesmos objetos de traço mesmo quando estágios anteriores da cadeia empregassem lógicas distintas, resultando em grande ganho de latência (consumo de um fator de 4× menor de tempo). Além disso, ganhos de redução na taxa também foram obtidos ao

empregar o ajuste da LLH em dados de colisão. A lógica de seleção durante 2017 para a nova cadeia (*ringer*) e a duplicata mantendo o método de *CutBased* estão na Figura 5.11.

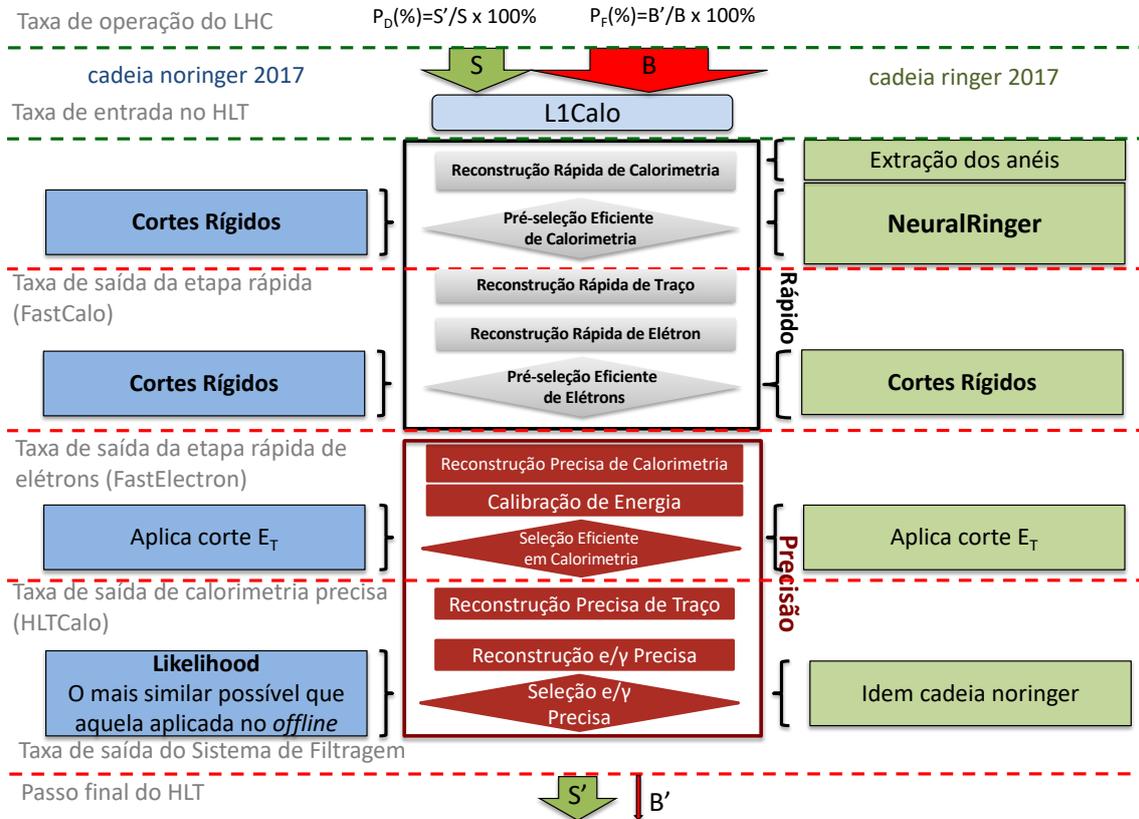


Figura 5.11: Estágios de seleção avaliados pela cadeia atuante em 2017 (*ringer*, verde) e sua versão duplicada com o método *CutBased* atuando no FastCalo (*noringer*, azul).

Durante 2017, identificou-se um problema no setor 16L2 do LHC. Um esquema alternativo (8b4e) foi aplicado para contorná-lo, o que reduziu o número de pacotes preenchidos no feixe para 70 % de sua capacidade nominal. A fim de compensar a perda em luminosidade devido à menor taxa de cruzamento de pacotes, aumentou-se a concentração dos feixes (emitância) e adicionou-se mais prótons nos pacotes (intensidade do feixe), o que possibilitou atingir luminosidade recorde de $\mathcal{L} = 2,08 \times 10^{34} \text{cm}^{-2}\text{s}^{-1}$. A decisão deu-se pelo nivelamento da luminosidade no início da *run* para o valor de $1,56 \times 10^{-34} \text{cm}^{-2}\text{s}^{-1}$, implicando em $\langle \mu \rangle = 60$ colisões, até que a luminosidade naturalmente atingisse valores menores ao limiar de nivelamento. A técnica de nivelamento consiste-se em reduzir a luminosidade no IP do experimento para que os seus picos estejam dentro da capacidade de atuação do experimento, permitindo, assim que o experimento obtenha maior luminosidade integrada (estatística) final durante as [193].

5.7.2 Discussão

A atuação *online* de um método não empregado para a análise *offline* gera receio quanto à desarmonia do método com as análises. A referência e última decisão nas análises físicas é dada pelo *offline*, porém a eficiência do sistema como um todo tem contribuições dos algoritmos atuando no SF_{on} (Figura 5.12). Esses algoritmos introduzem efeitos sistemáticos que são de difícil análise [194], uma vez que grande parte dos eventos rejeitados pelos algoritmos não estão disponíveis para análises (salvo aqueles provenientes de cadeias de suporte, que são executados com fatores de pré-escala e, logo, fornecem estatística reduzida). Os métodos *CutBased* e *LLH* são estratégias conhecidas e empregadas na Física de Altas Energias em longa data [39]. Esse não é o caso de técnicas de Aprendizado de Máquinas em assinaturas menos complexas¹⁰, como no caso de elétrons (Seção 3.1). Nesses casos, postergou-se ao máximo o emprego de técnicas que não sejam empregadas nas análises em *offline*, a fim de evitar a análise de impactos nas análises físicas, que exigiriam grande esforço. Por sua vez, manter o paradigma anterior atuante nas cadeias de elétrons implicava em impactos negativos na análise física, uma vez que as limitações em processamento implicavam na utilização de um menu coletando objetos físicos de maior energia ou a utilização de fatores de pré-escala e, portanto, deteriorando a capacidade do ATLAS nas análises empregando tais objetos.

O fato de não se ter o desenvolvimento e compreensão dos comportamentos da técnica nas análises físicas é parte das contribuições deste trabalho¹¹. Com a entrada da operação *online*, essa demanda se tornou ainda mais importante. Isso deveria levar em conta as necessidades desse ambiente de análise com grande exigência quanto

¹⁰Quando em comparação com demais assinaturas constituídas de subobjetos múltiplos ou provenientes de detectores com menor precisão, como o caso de jatos hadrônicos, táons etc.

¹¹Na realidade, essas excursões se iniciaram em 2010 na primeira estadia deste autor no CERN e foi fortemente baseada em um esboço inicial de implementação realizado por Danilo Enoque Ferreira de Lima. Em 2011, as primeiras análises foram reveladas no ATLAS, onde se observou, em dados de simulação, que a técnica tinha potencial quando a comparação com o algoritmo *CutBased* da época [195]. Todavia, essas análises iniciais esbarravam na limitação de conhecimento especialista. Em 2012, o trabalho foi descontinuado durante o mestrado deste autor que abordou outro tema e retomado em 2014, com o início deste trabalho de doutoramento. As limitações de conhecimento especialista foram atacadas a fim de possibilitar sua viabilidade em um ambiente complexo e altamente exigente quanto aos métodos e impactos envolvidos nas análises, devido à precisão necessária para identificar e compreender os mecanismos físicos que regem o nosso Universo. O conhecimento obtido durante esse processo se mostrou extremamente valioso para as nas análises durante o desenvolvimento da versão *online*, que demandou muito esforço em conjunto com João Victor da Fonseca Pinto e que impactou parcialmente no avanço de sua contraparte no *offline*, uma vez que o primeiro se dava como uma demanda prioritária e imediata dado os pontos trazidos nesta subseção. Igualmente, o desenvolvimento da técnica *offline*, baseado plenamente em Inteligência Computacional, sem agregar o devido contexto físico, é irrelevante para esse ambiente, sendo necessário integrar as duas pontas de forma a obter uma solução prática e, portanto, exigindo muito esforço multidisciplinar. Dentre o esforço para as diversas frentes de análise, inclui-se todo o ferramental desenvolvido para atender as demandas computacionais de uma aplicação envolvendo grande massa de dados em um ambiente altamente dinâmico como o ATLAS.

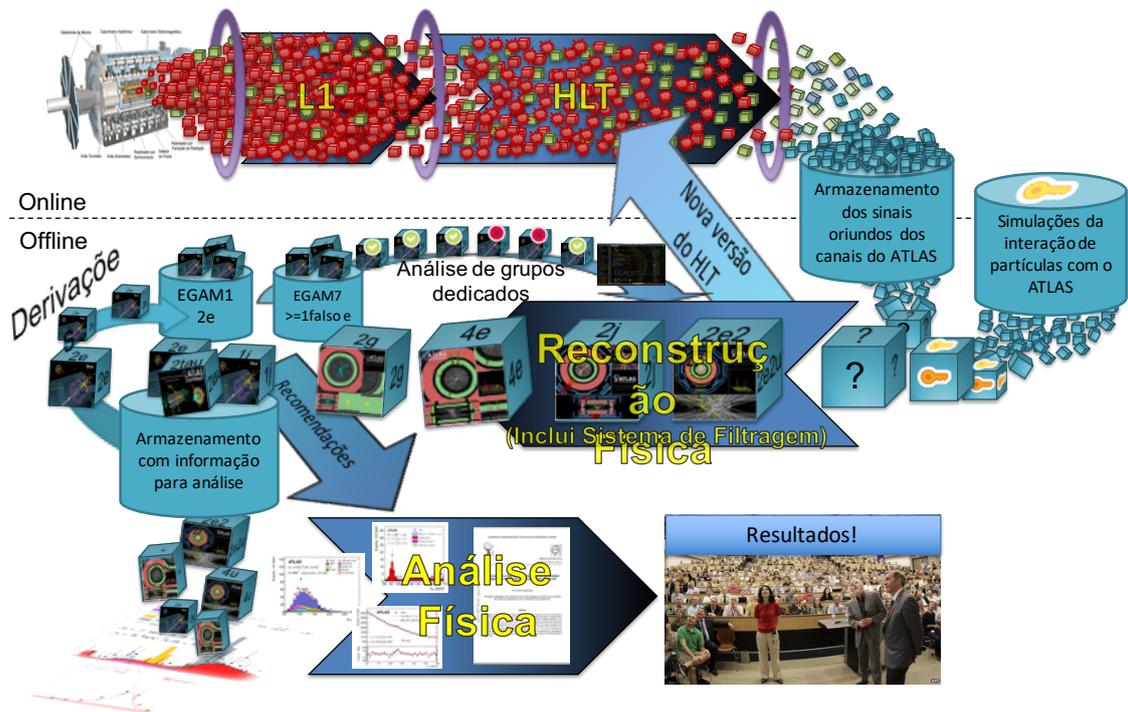


Figura 5.12: Esboço da dinâmica dos sistemas atuando desde a tomada de dados até os resultados das análises finais do Experimento ATLAS.

aos métodos empregados e a sua precisão. Além disso, a versão *offline* precisa estender a técnica para o seu pleno potencial empregando toda a informação disponível, de modo a alcançar eficiência ótima. Isso envolve um cenário complexo com diversos problemas de interesse do campo da Inteligência Computacional, como: fusão de informação; raros eventos de interesse; grande massa de dados; alta dimensionalidade e alteração dos padrões conforme variáveis descrevendo as condições de observação (posição, energia e empilhamento). O método para a extensão da técnica para o Sistema de Reconstrução (*Offline*) de Física do ATLAS está descrita na Seção 7.3 e os resultados na Seção 8.4.

A fim de dar luz à questão do impacto que o *NeuralRinger* introduziria no sistema, realizou-se um estudo inicial em uma *run* de 2016, onde não se observou qualquer indício de alteração no perfil da distribuição das grandezas físicas quando computadas no *offline* por parte da introdução da técnica na estrutura da cadeia. Contudo, havia uma série de limitações nesse estudo, citando estas como as principais: a estatística disponível para análise era limitada a uma única *run* e, a mesma, continha dados já filtrados pela cadeia anterior, ou seja, aplicando *CutBased* no FastCalo. Isso implicava na introdução de tendência no estudo, que reduzia à observação de amostras adicionais coletadas pela nova cadeia em decisões mutualmente exclusivas com o paradigma anterior à estatística menosprezável¹² obtida em cadeias de

¹²Método similar àquele descrito na Seção 7.2 foi empregado, ou seja, exigindo a presença de

suporte. Para prover precisão suficiente na análise, pediu-se ao ATLAS para manter operante uma única duplicata da cadeia primária dedicada à coleta de, ao menos, um elétron com a menor energia possível para o critério *medium* usando o paradigma anterior. A motivação para essa escolha e o método empregado para análise estão descritos em detalhes na Seção 7.2, já os respectivos resultados na Seção 8.3.

dois elétrons pelo método T&P, um deles passando o critério *medium* e outro o critério *loose*. As cadeias de suporte garantem a presença de um único candidato a elétron minimamente seletivo com pré-escala; logo, resultando em poucas observações que, eventualmente, acabaram por ser descartadas pelas cadeias primárias.

Capítulo 6

Soma em Anéis para a Descrição de Chuveiros de Partículas

Outra forma de representar e capturar a informação lateral e longitudinal do desenvolvimento do chuveiro dá-se através da descrição por via da soma das células do Sistema de Calorimetria por anéis concêntricos. Isso fornece compactação da informação para lidar com a alta dimensionalidade, enquanto mantém o núcleo da informação discriminante geralmente explorada na Física de Altas Energias. O algoritmo atualmente empregado e a ideia envolvida estão descritos na Seção 6.1. Na Seção 6.2, apresenta-se um histórico do emprego dessa informação na área, onde é descrito o desenvolvimento do método Algoritmo *NeuralRinger*, proposto para a filtragem *online* de elétrons no Experimento ATLAS.

6.1 Informação Anelada de Calorimetria

Durante o levantamento da interação de partículas com o calorímetro (Seção 2.4.2), mostrou-se que duas informações do chuveiro de partículas são comumente utilizadas para a seleção de elétrons: as larguras lateral e longitudinal do chuveiro. O ATLAS também explora essa informação (Subseção 5.2).

Como se viu na Subseção 2.4.2, boa parte (90 a 99 %) da energia de chuveiros EM fica contida dentro de um cilindro (de raio 1 a $3,5 R_M$). Além disso, também se observou que esses chuveiros se alastram por uma região lateralmente mais ampla do calorímetro conforme o seu desenvolvimento longitudinal. Assim, uma descrição razoável do chuveiro dá-se através de um formato cônico. Uma alternativa às grandezas físicas tradicionalmente empregadas utiliza esse conceito para obter sua base representativa, formando de anéis concêntricos de deposição de energia [196]. O conceito é somar a energia das células dentro da região delimitada pelo anel para compor a informação anelada. Ao realizar isso para uma janela do calorímetro

contendo a energia do chuveiro, detém-se uma representação que mantém a largura lateral do chuveiro. Adicionalmente, se tal processo for repetido para todas as amostragens longitudinais obtidas dentro da janela, isso permite caracterizar a largura longitudinal, bem como as nuances da evolução lateral do chuveiro. Assim, obtém-se uma base representativa que mantém as características físicas do processo de interação do chuveiro, mas com menor dimensão que a base constituída pelas células de leitura [196].

O algoritmo de anelamento (*Ringer*), conforme empregado atualmente¹, opera desta forma (acompanhar descrição observando Figura 6.1):

- dada uma posição central $c_{a,l}$ na l -ésima camada, emprega-se seu ponto no espaço $\eta \times \phi$ descrito por $(\eta_{a,l}; \phi_{a,l})$;
- obter-se todas as células dentro da região Θ_{Ringer} no plano $\eta \times \phi$. Atualmente, emprega-se a região $(\Delta_{\eta,Ringer}; \Delta_{\phi,Ringer}) = (0, 2; 0, 2)$ centrada em $c_{a,l}$;
- o anel $R_{n,l}$ é o subconjunto de Θ_{Ringer} formado por todas as células $c_{n,l}$. Calcula-se n por

$$n = \left\lceil \max \left(\frac{|\eta_{i,l} - \eta_{a,l}|}{h_{\eta,l}}, \frac{|\phi_{i,l} - \phi_{a,l}|}{h_{\phi,l}} \right) \right\rceil \quad \forall c_{i,l} \in \Theta_{RoI,l}, \quad (6.1)$$

onde (análogo para ϕ quando adequado): $\eta_{i,l}$ é o centro da célula $c_{i,l}$ em η ; $h_{\eta,l}$ é o passo adotado em η para a l -ésima camada, por sua granularidade; $l \in \{\text{PS, EM1, EM2, EM3, HAD1, HAD2, HAD3}\}$; e $n \in \{0, \dots, (N_l - 1)\}$, caso contrário se descarta a célula. Ressalta-se que as regiões descritas são retangulares, assim como são as células no plano $\eta \times \phi$;

- a grandeza $(r_{n,l})$ descrita pelo anel $R_{n,l}$ representa (aproximadamente, conforme discussão na Seção 9.6) a energia do chuveiro depositada na região delimitada pelo anel, dada por

$$r_{n,l} = \sum_{i=0}^{\forall c_{i,l} \in R_{n,l}} E_{i,l} \quad (6.2)$$

onde $E_{i,l}$ é a energia da célula $c_{i,l}$. Caso um anel seja dado por um conjunto vazio (ou seja, nenhum centro de célula dentro de sua região), assume-se valor nulo para sua respectiva grandeza.

¹A partir de ANJOS [197], implementado no Infraestrutura *Athena* por TORRES [198] e sofrendo atualizações de Danilo Enoque, Dhiana Deva e, mais recentemente, João Victor da Fonseca Pinto [199] para a versão *online*. O primeiro esboço da versão *offline* foi feito por Danilo Enoque para a EDM da Primeira Temporada de Coleta de Dados. Posteriormente, implementado pelo autor deste trabalho. A atualização para a Segunda Temporada de Coleta de Dados está descrita na Subseção 7.3.4.

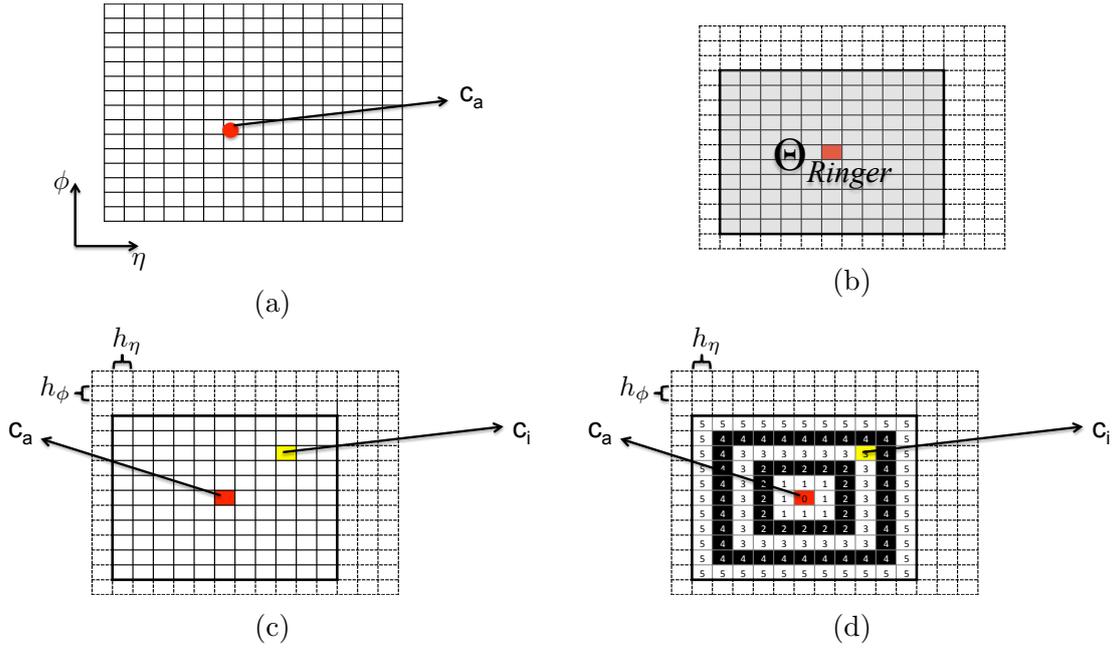


Figura 6.1: Esquema do funcionamento (não otimizado) do algoritmo *Ringer* para extração da base de informação anelada de calorimetria. A grade representa os limites de células para uma dada camada do calorímetro, da qual o algoritmo emprega o centro c_a (em vermelho) e está computando o índice n para o qual a célula c_i (amarelo) pertence. O grupo de células para o qual o algoritmo realiza os cálculos estão dentro de Θ_{Ringer} , que é a janela de busca. Indicam-se os valores comumente empregados para h_η e h_ϕ , ajustados para os valores típicos de granularidade de uma dada camada. O resultado final dos índices n do subconjunto R_n para o qual as células c_i pertencem estão em (d).

As células dos calorímetros compoem a l -ésima camada virtual descrita pelo algoritmo anelador, os passos $h_{\eta,l}$ e $h_{\phi,l}$, seu número de anéis (N_l), a divisão em seção para essas camadas e a notação empregada pelo trabalho para as grandezas representativas da base anelada estão disponíveis na Tabela 6.1. Na tabela, a nomenclatura ou acrônimos usados para determinar as amostragens de células que constituem uma camada no algoritmo anelador foram determinados pelo ATLAS e encontram-se disponíveis na Tabela 4.1. Os índices adicionais aqui presentes indicam a contagem da amostra, começando pela amostragem mais próxima ao feixe. Na convenção do ATLAS empregada no *Athena*, inicia-se a contagem dos calorímetros hadrônicos em 0. Por outro lado, para os calorímetros eletromagnéticos, a primeira amostra é indicada pelo índice 1. Além disso, o ITC é uma extensão do *TileCal* na região do *crack* que contém uma segmentação especial disponível na Figura 4.8. Para evitar essa não uniformidade, a convenção empregada no caso do algoritmo anelador inicia a contagem em 1 tanto as camadas eletromagnéticas, quanto as hadrônicas. Para o pré-amostrador, omitir-se-á o índice da representação dessa camada no algoritmo. O mesmo é considerado como pertencente a seção EM para efeitos do algoritmo

anelador. Por fim, para referir-se a um anel da camada empregar-se-á a notação “XYZZ”, onde “X” é o índice no anel na camada, “YY” indica a seção e “Z” a amostragem longitudinal nessa seção (exemplo: 1EM2 simboliza o primeiro anel da segunda camada eletromagnética).

Tabela 6.1: Descrição das amostragens longitudinais dos calorímetros compondo as camadas e seções no algoritmo de anelamento. Veja texto para detalhes.

Seção	Camada	Barril	ITC	Tampa	$h_{\eta,l}$	$h_{\phi,l}$	N_l
EM	PS	<i>PreSamplerB</i>		<i>PreSamplerE</i>	0,025	0,1	8
	EM1	EMB1		EMEC1	0,003	0,1	64
	EM2	EMB2		EMEC2	0,025	0,025	8
	EM3	EMB3		EMEC3	0,050	0,025	8
HAD	HAD1	<i>TileBar0</i> <i>TileExt0</i>	<i>TileGap3</i>	HEC0	0,1	0,1	4
	HAD2	<i>TileBar1</i> <i>TileExt1</i>	<i>TileGap1</i>	HEC1 HEC2	0,1	0,1	4
	HAD3	<i>TileBar2</i> <i>TileExt2</i>	<i>TileGap2</i>	HEC3	0,2	0,1	4

6.2 Histórico

Apresenta-se um histórico da aplicação de informação anelada de calorimetria no ATLAS, com a proposta inicial (Subseção 6.2.1) e contribuições posteriores (Subseção 6.2.2). Após o levantamento do histórico, realiza-se uma discussão das informações na Subseção 5.7.

Esses trabalhos se estendem por um longo período de tempo, onde a primeira publicação ocorreu em 1996 e a última contribuição aqui abordada como histórica, encerrou-se no final da Primeira Temporada de Coleta de Dados. Em todos os trabalhos, abordou-se a identificação para atuação no primeiro estágio de seleção do HLT, onde se emprega apenas a informação de calorimetria, correspondentes ao atual FastCalo. Na época (atualização descrita na Seção 4.4), o termo equivalente para esse estágio era *T2Calo*, uma vez que naquele período, o TDAQ do ATLAS considerava a execução do HLT em dois níveis sequenciais (na ordem de execução) que contavam com *software* e processamento dedicados: o Segundo Nível de Filtragem (L2) e o Filtro de Eventos (EF).

6.2.1 Proposta inicial

A proposta inicial que originou a aplicação do *Ringer* para o ATLAS [196]² surge de autores que fizeram parte do projeto do calorímetro SPACAL [201–204]. Esse calorímetro foi projetado durante a virada da década de 1990 para aplicação em ambientes de alta taxa de colisões, como o caso do *Superconducting Super Collider* (SSC)³ e do LHC. O calorímetro continha 155 células hexagonais (devido às condições experimentais) distribuídas lateralmente, de forma a gerar sete anéis concêntricos (novamente hexagonais). Devido a uma peculiaridade do arranjo experimental para a coleta dos sinais analógicos, estudos realizados pelo experimento (primeira publicação em 1989) contavam somente os sinais da célula central e das células dos dois anéis subsequentes estavam disponíveis individualmente. Para os demais cinco anéis, unicamente a informação da energia depositada nas células constituindo o anel encontrava-se acessível, de forma que as análises foram adaptadas para serem avaliadas nessas condições.

Outro conceito importante desenvolvido no SPACAL também foi empregado para a proposta de filtragem de elétrons. Os estudos realizados nesse calorímetro [202, 203] para avaliar a discriminação elétron–píon empregaram o valor R_p com valores unitários de α e β em dados de feixes experimentais e aplicando um corte linear para a discriminação. A abordagem proposta atingia os requisitos do sistema de filtragem de experimentos no LHC. Com esse método, o calorímetro atingido por elétrons e píons de 80 GeV obteve fatores de rejeição de píons (taxa de falso alarme) de $1,5 \times 10^{-4}$ para uma taxa alvo de identificação de elétrons fixada em 99 %. O valor R_p é dado conforme

$$R_p = \frac{\sum_{i=0}^{\forall(c_i) \in \Theta_{RoI}} E_i^\alpha r_i^\beta}{\sum_{i=0}^{\forall(c_i) \in \Theta_{RoI}} E_i^\alpha}, \quad (6.3)$$

onde c_i indica uma i -ésima célula qualquer, Θ_{RoI} constitui o conjunto de células pertencentes a RoI, E_i é a energia da i -ésima célula e r_i é a distância Euclidiana do centro da i -ésima célula ao baricentro de energia do chuveiro. Esse mapeamento representa uma boa medida do valor efetivo da largura lateral do chuveiro [196, 203, 204].

Todavia, a proposta para atuação em etapa equivalente ao atual FastCalo de um dos experimentos do LHC [196], com acesso apenas a informação de calorimetria, considerou diversas possibilidades para a utilização dessa informação para alimentar

²O mesmo trabalho cita [200] (1993) que também trata desse assunto, porém, a escolha baseou-se no maior alcance e extensão de [196].

³Descontinuou-se a construção do SSC posteriormente, no final de 1993.

uma MLP com $\tanh(\cdot)$ como função de ativação, um único neurônio na camada de saída e implementada através da versão disponível na época do pacote de redes neurais *JETNET* [57].

O espaço de entrada do estudo consistiu-se de 11×11 células retangulares $(0,02 \times \pi/300)$ uniformemente distribuídas no plano $\eta \times \phi$, tanto para uma seção eletromagnética quanto hadrônica — seções que foram simuladas estando perfeitamente alinhadas espacialmente na coordenada longitudinal. A quantidade de células empregadas foi determinada em [200], onde se notou que a maioria das demais células externas a essa região, cobrindo aproximadamente $0,4 \times 0,2$, continham valores nulos e não contribuíam para a discriminação, o que possibilitou a redução do espaço de entrada para essa subregião do espaço de disponível da RoI, que na época foi simulada contendo 22×22 células. Não obstante, o estudo operou com a soma das células na seção eletromagnética e hadrônica por ausência de simulações com acesso a essa separação. Assim, esse estudo não explorou a informação longitudinal e concentrou-se apenas na informação discriminante lateral dos chuveiros.

Estudou-se a aplicação de filtragem de elétrons, na qual se propagou, nesse calorímetro, eventos simulados dessas partículas e seu respectivo ruído de física esperado (~ 1600 elétrons e ~ 1000 jatos hadrônicos) em dois cenários: sem efeito de empilhamento e com a adição de uma média de 20 eventos *minimum bias*. Essas observações constituem partículas aprovadas pela respectiva versão do algoritmo de filtragem do L1 do ATLAS da época do estudo.

Assim, adotou-se MLP [128] para discriminação com $\tanh(\cdot)$ como função de ativação nas camadas escondida e de saída, essas contendo 10 e um único neurônio, respectivamente. Treinaram-se redes utilizando validação cruzada dividindo o espaço em 10 subconjuntos com parcela iguais para os dados de cada classe, sorteando 10 casos com 5 subconjuntos para ajuste (conjunto de treino) das redes neurais e os remanescentes para avaliação de eficiência (conjunto de teste). Os resultados publicados no trabalho constituem-se do valor médio das eficiências nos 10 sorteios. Avaliaram-se diversas transformações para o espaço de características:

- **Análise matricial de energia celular:** as células contidas no espaço de entrada são normalizadas pela norma-1 (energia total na região) e então propagadas para uma MLP com 10 neurônios na camada escondida;
- **Informação Anelada de Calorimetria:** aproveitou-se a ideia de anéis de calorimetria empregada em [204] como um procedimento para redução do espaço de entrada. Definiu-se uma normalização sequencial com o objetivo de aumentar a informação dos anéis mais externos, justamente os responsáveis pela descrição da largura lateral do chuveiro. Para isso, cada anel recebe sequencialmente valores menores de normalização, conforme: $E_{norm,1} = E_{norm1} - E_1$ e

$E_{norm,i} = E_{norm,i-1} - E_i$; onde E_{norm1} é o fator de normalização por norma-1 e E_i é a energia do anel. A normalização sequencial, que emprega conhecimento especialista, mostrou-se mais eficiente quando comparada à norma-1 e, por isso, empregou-a em todas análises envolvendo informação anelada. Estudou-se as seguintes topologias dos anéis: *a*) retangular, que se adequa à formulação realizada em (6.1); e *b*) circular, através do ponderamento das energias das células para reproduzir anéis com topografia aproximadamente circulares. Além disso, avaliou-se a topologia retangular com a adição de duas variáveis do primeiro nível de filtragem que são empregadas para o cálculo do isolamento EM (mais detalhes na Seção 5.7);

- **Valor R_p como generalização para normalização por norma-1:** observa-se que o valor R_p é uma generalização para a norma-1. Assim, para a análise matricial de energia celular, avaliou-se uma grade com 400 pontos na região $0.1 \leq \alpha \leq 2$ e $0.5 \leq \beta \leq 4$ com o objetivo de minimizar $\sum_i^{P_{D,i} \in \{90,90.5,\dots,99\}\%} P_{F,i}$, onde $P_{D,i}$ e $P_{F,i}$ são, respectivamente, a probabilidade de detecção e falso alarme para um mesmo limiar de decisão.
- **Mapeamento R_p com centroide na célula mais energética:** transforma-se o valor R_p em um mapeamento de $\mathbb{R}^2 \rightarrow \mathbb{R}$ conforme,

$$R_{p,i} = \frac{E_i^\alpha r_i^\beta}{\sum_{i=0}^{\forall(c_i) \in \Theta_{RoI}} E_i^\alpha}, \quad (6.4)$$

- **Mapeamento R_p com centroide no baricentro de energia:** redefina-se o cálculo de $r_{(i,bari)} = \sqrt{(x_i - x_{bari})^2 + (y_i - y_{bari})^2}$, onde:

$$(x, y)_{bari} = \frac{\sum_{i=1}^{\forall(c_i) \in \Theta_{RoI}} (x, y)_i E_i}{\sum_{i=1}^{\forall(c_i) \in \Theta_{RoI}} E_i}; \quad (6.5)$$

- **Anéis R_p :** realiza-se uma aproximação do mapeamento R_p ao aplicar o mapeamento no espaço de fase descrito pela energia dos anéis e suas respectivas distâncias em relação ao centro. Também se considerou a aplicação do mapeamento descrito por (6.4) antes da transformação em anéis, porém os resultados se mantiveram (apesar da maior necessidade de processamento por essa solução);
- **MLP com Treinamento Especialista para mapeamento R_p (Rede-Rp):** neste caso, a rede neural é adaptada no intuito de ser capaz de otimi-

zar, durante o treinamento, os valores de α e β aplicando retropropagação do erro até uma camada inicial cujos neurônios são meramente responsáveis pela aplicação do mapeamento R_p , assim tendo como entrada os valores de energia e distância euclidiana de cada uma das células do espaço de entrada. No caso aplicado, manteve-se $\beta = 2$ fixo, estando disponível a formulação matemática para o ajuste de α da Rede-Rp no apêndice de [196]. Chama-se atenção do leitor ao fato dessa configuração permitir o ajuste de α individualmente para cada célula.

Obtiveram-se as melhores eficiências de P_F sucedendo esta ordem decrescente (com suas respectivas topologias neurais e demais parâmetros, se aplicável), quando operando em $P_D = 97\%$:

1. mapeamento R_p com centroide no baricentro de energia: $121 - 10 - 1^4$, $\alpha = 0,6$ e $\beta = 1,3$;
2. Rede-Rp: $121 - 10 - 1$, $\langle \alpha \rangle_{(R_0; \dots; R_5)} = (0, 1; 0, 4; 0, 5; 0, 8; 0, 9; 1, 0)$, onde $\langle \alpha \rangle_{R_i}$ representa o valor central de α para o i -ésimo anel;
3. anéis em combinação com as variáveis para cálculo de isolamento EM: $8 - 4 - 1$, em conjunto com avaliação matricial de energia celular: $121 - 10 - 1^4$; e
4. anéis circulares: $6 - 3 - 1$.

Todos⁵ os resultados superaram o algoritmo de referência, que aplicou cortes lineares duas grandezas: valor R_p com $\alpha = 1$ e $\beta = 2$ e a fração de energia depositada nas 5 células centrais compondo uma cruz divididas pela energia total composta em uma região circular com raio de 6 células [205, 206].

Ademais, o estudo também ressalta a convergência da Rede-Rp para valores médios de α maiores para as células englobadas nos anéis mais externos (ver resultados da Rede-Rp no item 2), o que, como notado pelo estudo, está de acordo — uma vez que α controla (exponencialmente) o crescimento do fator de normalização — com os resultados obtidos com a normalização sequencial aplicada nos anéis.

6.2.2 Contribuições Posteriores

A seguir, encontram-se os trabalhos de pesquisa concluídos com contribuições para a identificação de elétrons com informação de calorimetria e aprendizado de máquinas no ATLAS.

⁴Não se encontrou diretamente no texto a explicitação desse valor. Porém, comenta-se no apêndice que a Rede-Rp, que tem como espaço de entrada 121 dimensões, empregou 10 neurônios na camada escondida. Por isso, estipula-se que o valor de 10 neurônios tenha sido aplicado em todos os casos com 121 entradas.

⁵Exceto anéis circulares quando operando em $P_D = 96\%$.

6.2.2.1 ANJOS [197]

Nesse trabalho, finalizado em 2006, o ATLAS encontrava-se em período preparatório para o início da coletada de dados — programada, naquele estágio, para 2007 [207] —; assim os esforços se encontravam focalizados para o desenvolvimento do TDAQ. Como parte dessa diligência, decorreu a produção do Primeiro Desafio de Dados do ATLAS [207]. Essa foi a primeira grande produção de dados simulados do ATLAS, que ocorreu através da jovem LCG⁶ durante o período 2002–2003. O processamento relacionado à produção de dados, tanto simulados quanto de colisão (se aplicável), referidos nos trabalhos subsequentes (incluindo o presente trabalho), deram-se através dessa grade computacional.

Os dados empregados continham 100 mil⁷ amostras de decaimentos $Z \rightarrow ee$ e 10 milhões de amostras de jatos duplos⁷. Empregou como base de dados amostras de $Z \rightarrow ee$ e jatos duplos com energia do processo rígido⁸ da colisão superior a 25 GeV. Aplicou-se nos dados a filtragem do *L1Calo*, exigindo *clusters* eletromagnéticos de, ao menos, 10 GeV na região de precisão do detector⁹ ($|\eta| < 2.5$), isolamento eletromagnético de 4 GeV e isolamento hadrônico de 2 GeV. Após a aplicação das exigências, foram obtidos 22.000 elétrons e 7.000 jatos. Metade dos dados compuseram o conjunto de treino e o restante para avaliação de eficiência de classificação.

Destaca-se que, diferente de [196], a simulação já continha a geometria do calorímetro que seria instalada no ATLAS, o que possibilitou avaliar melhor a evolução do chuveiro ao acessar a informação longitudinal através do seccionamento disponibilizado pelo Sistema de Calorimetria, bem como uma melhor disposição da informação lateral, devido a granularidade mais fina do calorímetro do ATLAS, quando comparado ao hipotético do estudo inicial.

O trabalho realizou uma otimização por busca exaustiva por via de uma grade de ($\sim 770k$) cortes lineares sequenciais em 4 variáveis R_η , E_{ratio} , $E_{T,EM}$ (*cluster*) e $E_{T,HAD1}$ que foi utilizada como referência. A performance para o SP máximo (SP_{max}) é dado por ($P_{D,spmax} = 91,62\%$; $P_{F,spmax} = 10,45$).

Com as variáveis normalizadas por esferização¹⁰ (empregada em todas análises nessa pesquisa), realizou-se o treinamento através do método *Resilient Back-propagation* (RPROP) para ajuste (batelada de 22.000 observações por época) de uma rede neural com arquitetura MLP 4 – 4 – 1 com função de ativação $\tanh(\cdot)$. A

⁶Para o primeiro relatório técnico, realizado durante a época de [197], ver [208].

⁷Encontrou-se o tamanho dos conjuntos de dados sem pré-filtragem pelo L1 em [209, 210].

⁸Ver Tópico 2.3.2.2.

⁹Como todos os demais estudos. Até o momento, o projeto não se estende além da região de precisão do ATLAS.

¹⁰Retira-se a média e divide-se pelo desvio padrão. Técnica também conhecida por *z-score*.

performance obtida é ($P_{D,sp_{max}} = 92,38\%$; $P_{F,sp_{max}} = 9,05\%$).

Aperfeiçoou-se essa configuração através da adição de mais variáveis: a) 4 amostragens longitudinais no ECAL (dentre elas a E_{EM1} definida na Seção 5.2); b) 4 amostragens longitudinais no HCAL; c) $w_{\eta 2}$; d) vazamento lateral em η na EM2 (R_{η}). A performance para operação em SP_{max} foi ($P_{D,sp_{max}} = 94,79\%$; $P_{F,sp_{max}} = 7,93\%$) para a arquitetura 14 – 10 – 1, que indica a configuração desse espaço de entrada como uma melhor base para classificação. Quando utilizando a base obtida após a aplicação de Análise de Componentes Principais (PCA) [211], essa performance se traduziu em ($P_{D,sp_{max}} = 94,75\%$; $P_{F,sp_{max}} = 7,59\%$), ligeiramente superior a base original, com arquitetura 14 – 12 – 1.

Aplicou-se análise de relevância [212] para as 14 variáveis, sendo avaliada com base na alteração de performance para MSE e índice soma-produto (SP) (Seção 7.4). Reportam-se, aqui, somente os resultados para SP , uma vez que essa medida é melhor para problemas de classificação. As cinco variáveis mais discriminantes foram estas, em ordem decrescente: R_{η}^{11} , E_{ratio} , E_{EM1} , $w_{\eta 2}$ e $E_{HAD,1}$. Aplicou-se a técnica de poda de rede neural, onde se obteve as eficiências ($P_{D,sp_{max}} = 93,94\%$; $P_{F,sp_{max}} = 9,09\%$) e arquitetura 5 – 5 – 1.

O estudo aplicou também análise de compactação na base de PCA, comparando a mesma em relação a poda por análise de relevância. Quando reduzindo o espaço de características para 11 dimensões, não houve diferença de performance entre a escolha por base na retenção de energia por PCA ou relevância, porém quando reduzindo para 7 ou 5 dimensões, a escolha por relevância foi crucial para uma deterioração diminuta da performance.

Então, o trabalho voltou-se para a implementação do algoritmo de anelamento, que foi realizada atendendo às necessidades de implementação do segundo nível de filtragem. Essa implementação guiou todas as demais versões do algoritmo de anelamento, inclusive aquela que atualmente é utilizada para o Sistema de Reconstrução (*Offline*) de Física do ATLAS, descrita na Seção 6.1. Aplicam-se as ressalvas mencionadas quanto à precisão do algoritmo de gerar anéis discutidas na Capítulo 9. Todavia, cabe a diferença para a determinação do centroide para a computação dos anéis explicada na Subseção 7.3.4. A normalização sequencial desenvolvida em [196] (ver p. 114) foi empregada nos anéis, que alimentaram uma rede neural. A melhor topologia para a rede neural foi 100 – 12 – 1 com a performance de ($P_{D,sp_{max}} = 96,21\%$; $P_{F,sp_{max}} = 3,44\%$).

Em seguida, aplicou-se análise de relevância por MSE e SP . Dessa análise de

¹¹A variável definida pelo autor como $R_{7 \times 7}^{3 \times 7}$ teve relevância praticamente nula, o que indica uma possível repetição dessa variável no espaço de entrada tendo em vista o conhecido alto teor discriminante da mesma.

relevância, mantiveram-se somente os anéis com relevância positiva ($62 - 12 - 1$), o que resultou em uma eficiência de ($P_{D,spmax} = 96,89\%$; $P_{F,spmax} = 2,93\%$). Algumas outras configurações foram avaliadas ao utilizar somente parte dos primeiros anéis mais relevantes e aglutinação das camadas hadrônicas. A que teve melhor eficiência ($P_{D,spmax} = 97,59\%$; $P_{F,spmax} = 3,22\%$) utilizou apenas os 6 anéis mais relevantes da EM1, 4 anéis da EM2 e somou os anéis longitudinais das camadas HAD de forma a obter uma única representação longitudinal em anéis HAD, totalizando 14 anéis.

6.2.2.2 TORRES [198]

O trabalho de pesquisa seguinte finalizou-se em 2010. Os anéis também foram gerados com o mesmo algoritmo descrito na Seção 6.1. Durante essa época, o LHC já se encontrava no período inicial da *Run 1*, de forma que o trabalho implementou a versão inicial de operação do algoritmo anelador no *Athena*.

A base de dados continha elétrons igualmente distribuídos entre $7 < E_T[\text{GeV}] < 80$ e jatos hadrônicos com ao menos 17 GeV, ambos com $\mathcal{L} = 2 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$. Aplicou-se a filtragem do *L1Calo* para obter dois conjuntos de dados:

- (a) *e10* (470k elétrons e 315k jatos): filtraram-se os eventos aplicando um corte em $E_T > 10 \text{ GeV}$ pelo *L1Calo*;
- (b) *e15i* (140k elétrons e 13k jatos): idem *e10*, porém com os seguintes cortes: $E_T > 15 \text{ GeV}$; isolamento eletromagnético em 3 GeV; isolamento hadrônico em 1 GeV; e limite para vazamento hadrônico de 2 GeV.

O trabalho contém uma caracterização dos anéis no conjunto *e10*, onde se percebeu um pequeno pico na deposição média de energia na HAD3. Isso ocorreu pela contribuição da barra cintiladora que, pelo aspecto construtivo dos anéis, contribui para aquela camada. Realizou-se a correlação entre os anéis, onde percebeu-se correlação entre anéis na EM1 afastados por $k \times h_{\eta,l}$ (Seção 6.1). O autor identificou a causa como decorrência da diferença entre a granularidade empregada pelo algoritmo e aquela disponível no calorímetro. Uma análise das consequências e proposta de soluções está disponível no Seção 9.6 e 9.7.

Avaliou-se a normalização a ser empregada. A normalização sequencial apresentada pelo trabalho possui dois termos adicionais: E_{stop} , parâmetro delimitando o valor mínimo (empregou-se $E_{stop} = 100 \text{ MeV}$) para a normalização dos anéis; e E_{thres} , no qual, caso a norma-1 seja inferior, evita-se a aplicação ($E_{thres} = 0,01 \text{ MeV}$) da normalização. Dentre todas normalizações avaliadas, obtiveram-se os melhores valores de SP_{max} conforme: a esferização, seguida pela norma-1, norma-2 e sequencial. O trabalho manteve a normalização por norma-1 para as demais análises, por ser de mais simples avaliação e estar próxima em termos de SP_{max} da esferização.

Realizou-se uma discussão quanto à normalização sequencial, onde se esperava que o emprego de conhecimento especialista possibilitasse melhor eficiência final de classificação da mesma em relação à norma-1 devido à adição de informação especialista — a normalização sequencial acentua a informação discriminante do perfil lateral do chuveiro —. Acreditou-se em uma necessidade de melhor avaliação do ajuste dos parâmetros E_{stop} e E_{thres} .

Uma análise do efeito da inclusão das células do ITC utilizando todos os dados mostrou uma leve variação ao empregar, ou não, essas células na análise. Ao empregar o cintilador, houve um ganho marginal em termos de SP_{max} .

Esse estudo introduziu a avaliação de parada da rede neural utilizando o critério SP que mostrou ser um melhor critério para discriminação. O método de treinamento também foi o RPROP. Esse critério é utilizado pelo trabalho atual (Seção 7.3). Outros parâmetros para treinamento foram indicados por esse trabalho: 6.000 eventos de batelada; 10.000 épocas máximas; e sem parada prematura de treinamento por especialização. O trabalho empregou validação cruzada com 12 subconjuntos, dos quais 6 compuseram o treino e 6 teste/validação (ou seja, o mesmo conjunto em ambos os casos).

Duas abordagens nos dados *e15i* realizaram-se para a obtenção das bases que compuseram o espaço de características via pré-processamentos estatísticos PCA [211], Componentes Principais de Discriminação (PCD) [213] e Análise de Componentes Independentes (ICA) [214]:

- segmentada: aplica-se o pré-processamento para cada uma das camadas do algoritmo de anelamento;
- não-segmentada: o pré-processamento é avaliado em cima de toda a base formada pelos anéis.

Em PCA, a melhor eficiência ($P_{D,sp_{max}} = 96,75 \pm 0,10$; $P_{F,sp_{max}} = 6,81 \pm 0,35$) deu-se por uma abordagem segmentada retendo 99 % da energia. A análise segmentada por PCD mostrou que as camadas com maior informação relevante para discriminação são, em ordem decrescente, a EM1 e EM2; onde se observou que isso também é refletido nas variáveis do $T2Calo$ dessas camadas. Isso ocorre igualmente para as camadas hadrônicas, que tem o menor potencial discriminante em ambos os casos. A operação da rede neural treinada com as projeções por PCD não apresentou grandes vantagens quando comparada com a base representada pelos anéis — estando dentro da margem de incerteza.

A análise de ICA empregou o algoritmo JADE [214]. Uma série de compactações por PCD e PCA foram avaliadas antes de obter a projeção na representação independente. A melhor performance foi a retenção de 99 % de energia por PCA através de

abordagem não-segmentada para então projetar na base independente. Essa abordagem (36 – 10 – 1) obteve os melhores resultados para o ponto de operação em SP_{\max} ($P_{D,sp_{\max}} = 96,86 \pm 0,52$; $P_{F,sp_{\max}} = 6,50 \pm 0,39$). Esse também foi o caso quando operando em $P_D = 97,34\%$, cujo P_F foi de $7,13 \pm 0,31\%$. A versão segmentada dessa abordagem encontrou-se dentro da incerteza e bastante próxima da performance média não-segmentada.

O estudo também realizou análise por relevância nos anéis empregando como critério a performance em SP em todas as bases, incluindo a base de anéis puros. Treinaram-se redes em cima das bases podadas, cuja melhor performance ($P_{D,sp_{\max}} = 96,81 \pm 0,25$; $P_{F,sp_{\max}} = 7,03 \pm 0,26$) também foi a aplicação de ICA na base de PCA não-segmentada contendo 99 % de energia (28 – 10 – 1).

A última abordagem empregou classificação linear (Fisher [215]), cuja performance mais relevante foi ($P_{D,sp_{\max}} = 96,49 \pm 0,23$; $P_{F,sp_{\max}} = 7,54 \pm 0,21$) aplicando ICA na base composta por PCD não segmentada (8 – 1).

Considerou-se o impacto computacional, onde as abordagens segmentadas, apesar de não atingirem a performance ótima, operaram próximos às suas contrapartes não-segmentada e possuem menor complexidade computacional quando o número de multiplicação e acumulação é considerado.

Finalmente, validou-se o algoritmo em raios cósmicos, onde apenas aqueles com perfil muito próximos de elétrons foram aprovados. O algoritmo anelador mostrou capacidade superior ao *T2Calo* para rejeição de raios cósmicos.

6.2.2.3 SIMAS FILHO [216]

Cerca de um semestre depois, no final de 2010, finalizou-se outra contribuição que se dedicou ao estudo de pré-processamento não-linear. As projeções em bases ocorreram em cima dos anéis normalizados por norma-1 para os dados simulados e sequencial para os dados de colisão — acrescenta-se aqui a informação de que essa normalização sequencial empregava E_{stop} e E_{thres} , valores idênticos aqueles mencionados por [198]. A motivação para a diferença deu-se pelo fato da versão do algoritmo em operação no Sistema de Filtragem do ATLAS empregar essa normalização na época (herança dos estudos de [197]), que armazenavam a base normalizada e impossibilitava a análise em outras normalizações¹². Aplicaram-se três variantes de NLICA:

- *Post Nonlinear model* (PNL) [217]: onde se assume que as contribuições não-lineares ocorrem na base das misturas ao invés das fontes;

¹²Isso foi posteriormente corrigido, em março de 2011.

- SOM [218]: extração de SOM [129]: mapa com duas dimensões para obter uma base composta pelas coordenadas do neurônio vencedor que constitui informação independente. Três algoritmos de ajustes foram avaliados: LVQ [219], não-supervisionado e supervisionado (adicionando a informação da classe durante o ajuste, e removendo essas dimensões para operação);
- ICA Local: avaliou-se também uma abordagem local para a ICA linear, empregando a sequência $R \rightarrow PCD \rightarrow ICA \rightarrow MLP$ em grupos obtidos via $k - means$ [215].

Inclui-se a esses pré-processamentos uma análise por ICA linear com a sequência descrita para ICA Local. Essas informações alimentaram uma MLP com duas camadas.

Ao mesmo tempo, propôs-se uma abordagem utilizando *classificadores especialistas* para cada camada longitudinal obtida pelo Algoritmo *NeuralRinger*. Essa abordagem contrasta com aquela utilizada por [198], no qual se aplicou a segmentação longitudinal somente para a obtenção da base compondo o espaço de características via *pré-processamento* — em [216], aplica-se tanto o classificador quanto o pré-processamento em bases especialistas. Para a obtenção do espaço de saída final, o trabalho concatenou as saídas de cada um dos classificadores segmentados e os alimentou para uma rede MLP com duas camadas. A essas avaliações de discriminadores, acrescentam-se cortes lineares por discriminante de Fisher.

Realizou-se uma análise de relevância por camada, substituindo toda uma dada camada na base dos anéis pela média para todos os pré-processamentos avaliados. Em todos os casos de pré-processamentos, observou-se maior relevância para a EM1, em especial para o algoritmo ICA Local. Ao retirar determinadas camadas da operação do *NeuralRinger* sem adição de pré-processamento estatístico, observou-se que a rede operava *melhor* sem a HAD2 e HAD3.

Para o conjunto *e10*, a melhor performance ($P_{D,sp_{max}} = 99,60 \pm 0,43 \%$, $P_{F,sp_{max}} = 1,50 \pm 0,08 \%$) deu-se para o ICA Local, usando um *cluster* para cada classe. Quando em operação para $P_D = 92,5 \%$, a melhor abordagem deu-se por uma rede obtendo as saídas dos classificadores segmentados por camada, com $P_F = 0,61 \pm 0,07 \%$. O mesmo se repetiu para o conjunto *e15i*, porém com menores performances devido à eliminação dos candidatos mais simples uma vez que o corte do *L1Calo* estava sendo aplicado.

Uma análise em quadrantes realizada em função das variáveis η e E_T entre o Algoritmo *NeuralRinger* e o *T2Calo* mostrou que a melhor performance daquele se deveu, principalmente, por recuperação de elétrons de baixa energia e na região do *crack*.

Ao aplicar os algoritmos ajustados em dados de simulação em uma base composta por dados de colisão, obteve-se performance baixas para todas as configurações, que também ocorreu para o Segundo Nível de Filtragem empregando informação do Sistema de Calorimetria (*T2Calo*) (esse fato será discutido no Subseção 6.2.3). Reajustou-se os discriminadores em dados de colisão utilizando o *offline* como referência — empregando *medium* para compor a base de dados de elétrons e não-*loose* para hádrons —, onde a melhor performance ($P_{D,spmax} = 87,61 \pm 0,30$ %; $P_{F,spmax} = 14,22 \pm 0,17$ %) se deu pela ICA Local. A rede ajustada em cima das saídas das redes segmentadas em dados de colisão mostrou boa performance ($P_{D,spmax} = 89,44 \pm 0,18$; $P_{F,spmax} = 15,80 \pm 0,10$), e superior a do algoritmo *NeuralRinger*. Um fato importante nessa análise mostra que ao remover-se as camadas EM3, HAD2 e HAD3 (com um decréscimo mínimo, inclui-se também o PS), obtém-se performance praticamente igual a da rede operando com todas as camadas.

6.2.2.4 XAVIER [175]

Aproximadamente no mesmo período em que [175] finalizou seu trabalho, o LHC entrega suas últimas colisões $p - p$ da *Run 1*. Assim, nesse trabalho já havia sido coletada estatística suficiente para uma análise de eficiência do *NeuralRinger* em dados de colisão, onde empregou quatro ocorridas durante outubro de 2011. Todavia, essa estatística estava polarizada pela decisão do *T2Calo* — apenas eventos aceitos pelo algoritmo do ATLAS estavam armazenadas com anéis, de forma que o trabalho somente pode ajustar discriminadores em estatística aceita pelo algoritmo de seleção de elétrons¹³. A pior dificuldade para esse problema se dá pela estatística obtida para o ruído, onde essa polarização significa que apenas eventos previamente aceitos como sinal pelo SF_{on} contém anéis. Ainda assim, como as condições de operação na época não eram tão exigentes quanto as atuais, isso permitia o SF_{on} operar razoavelmente *loose* e, em decorrência, haver alguma estatística disponível de hádrons para compor a base de ruído.

Nesse trabalho, empregou-se pela primeira vez a a técnica de T&P (Seção 5.6) para avaliação de performance do algoritmo *NeuralRinger*. Todavia, o trabalho não aplicou subtração de ruído no método T&P, ao invés disso, aplicou-se o critério *medium* do algoritmo de seleção de elétrons do *offline* nos *probes*. Manteve-se, também, a avaliação realizada pelos trabalhos [198, 216] empregando diretamente como referência as decisões do algoritmo *offline*. Além disso, utilizou-se também uma base de dados simulada que continha eventos $p - p$ a 14 TeV, como os demais estudos anteriores. O trabalho ressalta que essa faixa de energia difere daquela de

¹³Ainda restam alguns detalhes técnicos sobre quais cadeias contendo anéis foram executadas em cada uma das *runs*, o que se adicionará na versão final deste trabalho.

operação do LHC na *Run 1* (7 e 8 TeV). A base de dados de simulação continha luminosidade de $10^{-34}\text{cm}^{-2}\text{s}^{-1}$. Aplicou-se a otimização do *NeuralRinger* para duas bases de dados com as filtragens do *L1Calo* para as cadeias *e10* e *e22* da época. Em cada uma dessas bases de dados, aplicou-se uma seleção de eventos baseada no *offline* e outra na *truth* — casamento da informação do candidato de elétron com a informação do gerador de eventos simulados. Os trabalhos anteriores apenas consideraram a informação da *truth* em suas análises de simulação.

Os resultados obtidos para dados de simulação foram, em valores de eficiência de T&P para a referência *offline*, de ($P_{D,spmax} = 96,86\%$; $P_{F,spmax} = 0,25\%$) para a base *e10* e ($P_D = 97,11\%$; $P_F = 0,66\%$) para base *e22*. A performance do *T2Calo* foi respectivamente ($P_D = 95,27\%$; $P_F = 6,22\%$) e ($P_D = 95,21\%$; $P_F = 3,25\%$). Já em dados de colisão, com um algoritmo treinado em T&P, as eficiências foram ($P_{D,spmax} = 97,18 \pm 0,11\%$; $P_{F,spmax} = 1,89 \pm 0,56\%$) para o discriminador ajustado com os *probes* e ($P_{D,spmax} = 98,30 \pm 0,11\%$; $P_{F,spmax} = 5,05 \pm 0,84\%$). Essas performances foram novamente superiores aquelas do *T2Calo*, respectivamente ($P_{D,spmax} = 96,35 \pm 0,11\%$; $P_{F,spmax} = 6,66 \pm 0,56\%$) e ($P_{D,spmax} = 96,59 \pm 0,12\%$; $P_{F,spmax} = 6,27 \pm 0,84\%$).

6.2.3 Discussão

Uma primeira consideração importante antes da discussão deste levantamento bibliográfico concerne o progresso gradual dos simuladores de Monte Carlo, o que impede a comparação de performance entre os diversos resultados apresentados em diferentes bases de dados. Esse fator é ainda mais impactante após o início da Primeira Temporada de Coleta de Dados, onde o ATLAS notou, com as primeiras colisões em 7 TeV, que as distribuições com simulações das grandezas padrões continham desvios, especialmente em R_η e $w_{\eta 2}$ (ou seja, informação lateral do chuveiro), em relação às suas representações obtidas em dados de colisão [220]¹⁴. As simulações tiveram correções após a coleta desses primeiros dados. Ainda assim, existem problemas na modelagem por simulação [180, 181], o que mostra a importância da utilização de dados de colisão para o ajuste das técnicas empregadas, estratégia atualmente empregada pelo ATLAS para o Sistema de Reconstrução (*Offline*) de Física do ATLAS. Em 2017, o SF_{on} também utilizará versões ajustadas em dados de colisão.

Durante o trabalho de [196], observaram-se, em todos os casos para a otimização de α e β , valores ótimos diferentes para os cenários com ou sem a presença do efeito

¹⁴Iniciou-se, neste ano, uma força tarefa [221] que se dedica a melhorar a descrição de simulação para o comprimento lateral do chuveiro. Em especial, há uma limitação na representação da correlação cruzada entre células do calorímetro. Esse efeito afeta os anéis e deve ser mantido em mente quando considera-se a aplicação de classificadores ajustados por simulação.

de empilhamento. Por exemplo, na MLP alimentada por mapeamento- R_p , obtém-se $(\alpha; \beta) = (0, 9; 1, 6)$ quando sem empilhamento e $(\alpha; \beta) = (0, 8; 0, 7)$ no outro caso. Isso revela que uma operação por mapeamento/valor R_p pode obter melhor performance ao empregar valores adaptativos de α e β de acordo com o estimador de empilhamento utilizado.

Os anéis mais relevantes (Tabela 6.2) entre os diversos trabalhos mostram uma determinada discrepância na ordem em que aparecem. No entanto, há informações convergentes: a EM1 é a camada mais relevante, em especial seus primeiros dois a três anéis; outra camada importante é a HAD1, destacando-se o seu segundo anel; as camadas HAD2, HAD3 e EM3 são menos relevantes. O fato da rede ter pouca dependência na EM3 é um fato bastante benéfico para a operação do *NeuralRinger* e deve ser explorado, possivelmente removendo essa camada (ou somando ela com a EM2). A motivação para isso dá-se pela dificuldade enfrentada pelo ATLAS com a grandeza f_3 [180, 181], que tem importante relevância para discriminação de elétrons; porém, em altas energias ($E_T > 80$ GeV), gera *redução* na performance de classificação por haver deposição de energia de elétrons nessa região. Removeu-se essa variável para operação em altas energias, uma vez que há pouca estatística nessa região para ajuste das técnicas. Nesses casos, a operação ocorre por discriminadores ajustados pelos elétrons concentrados na região disponível de mais alta de energia, cujo pico ocorre em torno de 50 GeV (exemplo na Figura 7.8a). Já que o *NeuralRinger* não se baseia nessa camada em baixas energias, isso permite acreditar que a retirada dessa informação seja possível sem grandes impactos em elétrons menos energéticos, caso essa camada cause perda de eficiência em altas energias. Todavia, uma análise nesse sentido faz-se necessária. O PS, ainda que a função principal seja recuperar a energia depositada no ID, aparece eventualmente como um dos mais relevantes. Uma explicação possível pode se dar com a semelhança do mesmo com calorímetros *pre-shower* [41], possibilitando adicionar a informação discriminante na maior probabilidade de chuviros EM ocorrerem prematuramente ao calorímetro.

Algumas diferenças podem ser citadas como, por exemplo, no caso de [197], no qual as redes mostram uma concordância entre os anéis imediatamente posteriores aos anéis centrais, havendo relevância praticamente nula para o anel central. Já em [198] e [175], os anéis centrais aparecem como informação mais relevante. Isso pode ser uma consequência da base de representação, onde o primeiro empregou a normalização sequencial e os demais a norma-1. Em alguns casos, a EM1 aparece como mais relevante, em outros a HAD1. A HAD1 é a mais relevante quando o filtro do *L1Calo* não inclui isolamento hadrônico¹⁵, assim há grande quantidade de hádrons com vazamento para o HCAL, geralmente de baixa energia e, portanto, contidos na primeira camada. Em [197], percebeu-se que uma melhor representação ocorria ao

¹⁵A cadeia *e22*, empregada em [175] aplicava isolamento HAD no *L1Calo*.

somar as três informações longitudinais. Ao realizar essa abordagem, incrementa-se as contribuições hadrônicas, o que motiva uma avaliação dessa abordagem para a implementação do SF_{on} , onde grande parte dos dados também contém contribuições HAD. Além disso, os resultados de [175] para diferentes cortes no $L1Calo$ mostram que as informações relevantes alteram-se conforme a energia da partícula. Há motivos para acreditar-se que o mesmo deve ocorrer em η , devido à diferença de granularidade em cada uma das regiões e a ocorrência de regiões anômalas (Seção 9.6). Assim, há motivação para realizar a análise de relevância em múltiplos subespaços de fase ($E_T \times \eta$) e compreender a informação relevante em cada um desses espaços.

Tabela 6.2: Anéis ou camadas mais relevantes obtidas para as contribuições realizadas para o *NeuralRinger*. A informação mais relevante está em ordem decrescente. Para notação dos anéis e camadas, ver Tabela 6.1. A abreviação seg. é empregada para representar uma abordagem segmentada.

Autor	Base de dados	Critério	Abordagem	Informação Mais Relevante
Dados de Simulação				
ANJOS [197]	Zee/JF17 ATLAS DC1 (2002)	MSE	MLP	2EM1, 2EM2, 3HAD1, 3EM3, 2EM3, 3EM2 e 2PS
ANJOS [197]	Zee/JF17 ATLAS DC1 (2002)	<i>SP</i>	MLP	3HAD1, 2HAD1, 2EM1, 3HAD1, 3EM1, 4EM1 e 2EM2
TORRES [198]	Single.e/JF17 e10 (2010)	<i>SP</i>	MLP	1PS, 2PS, 2EM1, 1EM1, 3EM1, 4EM1 e 6EM1 (+ 5 anéis mais externos da EM2 e os dois primeiros anéis da HAD1 e o último da EM3, com menor relevância)
SIMAS FILHO [216]	Single.e/JF17 e10 (2010)	<i>SP</i>	MLP seg.	EM1, EM2, HAD1
SIMAS FILHO [216]	Single.e/JF17 e15i (2010)	<i>SP</i>	MLP seg.	EM1, PS, EM2
XAVIER [175]	Zee/JF17 - e10 (2011)	<i>SP</i>	MLP	2HAD1, 1HAD1, 1EM2, 1EM1, 2EM1, 1PS, 2PS
XAVIER [175]	Zee/JF17 - e22 (2011)	<i>SP</i>	MLP	1EM2, 2HAD1, 2EM2, 5EM2, 1PS, 1EM1, 2EM1
Dados de Colisão				
SIMAS FILHO [216]	155697, 158582 - L1EM2 (2010)	<i>SP</i>	MLP seg.	EM1, EM2, PS, HAD1
XAVIER [175]	191715, 191190 - T&P (2011)	<i>SP</i>	MLP	4EM2, 2HAD1, 1EM1, 2EM1, 1HAD1, 3EM2
XAVIER [175]	191715, 191190 - T&P (2011)	<i>SP</i>	MLP seg.	EM1, EM2, HAD1
XAVIER [175]	190236 - 191920 - <i>offline</i> (2011)	<i>SP</i>	MLP	2HAD1, 4EM2, 1EM1, 2EM1, 1PS, 2PS, 2HAD2
XAVIER [175]	190236 - 191920 - <i>offline</i> (2011)	<i>SP</i>	MLP seg.	EM1, EM2, HAD1, PS

Quanto a análise de normalização realizada em [198], percebe-se, ao analisar a performance obtida em função de E_T da partícula (ver Figura 6.2), que a esferização, apesar de obter a melhor performance, cria uma polarização na performance conforme a concentração dos dados. Os hádrons do conjuntos de dados aplicados se concentravam em baixa energia, onde ocorre uma brusca queda em P_D para essa normalização. Já as demais normalizações baseadas em energia (norma-1, norma-2) — e independentes de estatística — não tem esse efeito. Isso deve ser tomado em consideração caso adotar essa normalização para todo o espaço de fase. Esse efeito

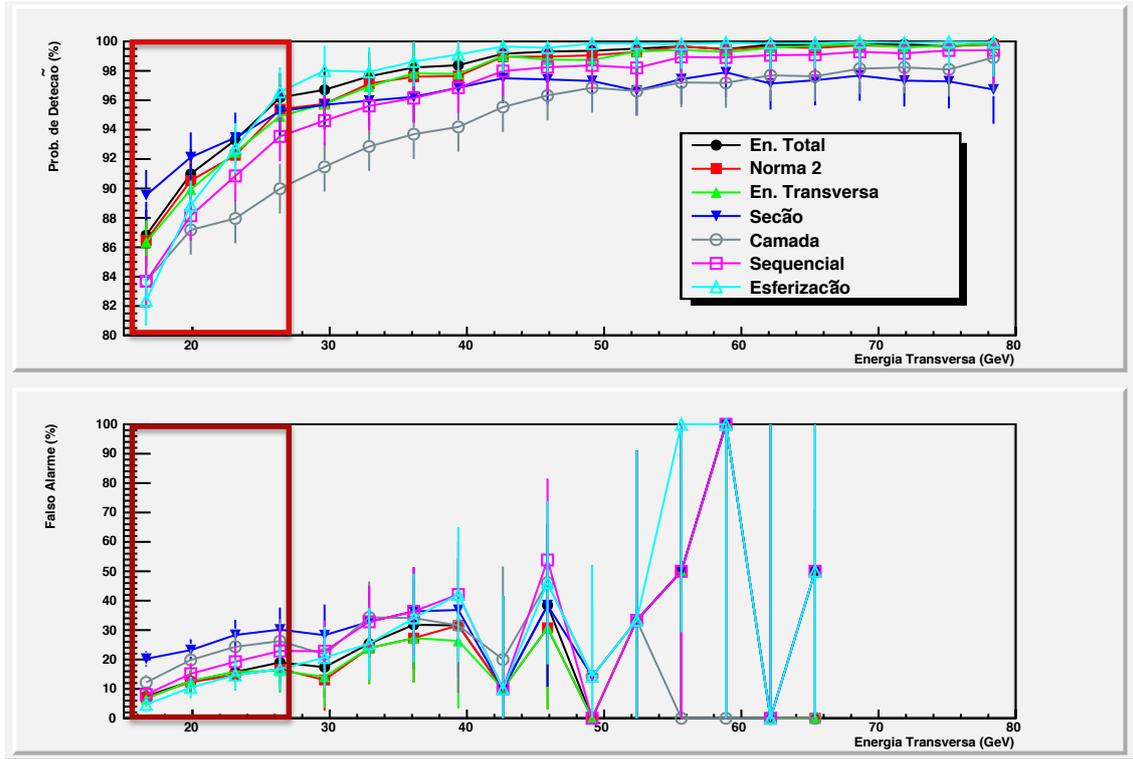


Figura 6.2: Variação da probabilidade de detecção (em cima) e do falso alarme (embaixo) de diferentes normalizações em função da energia transversa do evento. Aplicaram-se as normalizações na base de informação anelada de calorimetria para formar o espaço de características que alimentou uma MLP não-segmentada recebendo todo o espaço de fase obtido pela região de precisão do ATLAS para um conjunto de dados simulados empregados por [198]. Destaca-se a região na qual ocorre uma queda brusca de eficiência para a detecção da normalização por esferização enquanto as normalizações norma-1 (En. Total), norma-2 não possuem essa queda e obtêm valores similares de P_D [198].

é especialmente prejudicial para o SF_{on} , onde se aplica o algoritmo *NeuralRinger* como estágio de pré-seleção e se deve evitar ao máximo perdas de P_D para todo o espaço de fase. Esse comportamento igualmente motiva a análise por subespaços de fase, para obter performance ótima em cada um dos mesmos e evitar polarizações como essas em regiões com maior disponibilidade de estatística.

A análise de correlação não-linear via informação mútua feita por [198] mostrou forte dependência lateral entre os anéis iniciais até ligeiramente centrais de elétrons. Nesses anéis, há dependência longitudinal também. Na EM2, há forte correlação para todos anéis exceto o último. Por outro lado, para hádrons a dependência lateral é menos pronunciada, bem como o caso da longitudinal. Essa menor correlação não-linear é curiosa e uma melhor investigação permitiria melhorar a compreensão do desenvolvimento do chuveiro. Acredita-se que isso seja reflexo da adição de flutuações independentes, como a f_{inv} e o desenvolvimento de chuveiros EM.

Quanto a análise em quadrantes realizada por [216], cabem algumas consi-

derações. A análise considerou os pontos que mais se ressaltavam nas diferenças de performance, porém, o foco para análise deve ser na importância das regiões para o estudo da física. No caso, os elétrons de baixa energia recuperados estavam abaixo da energia da cadeia de operação (10/15 GeV) ou na região do *crack*. No primeiro caso, eles não compunham informação destinada para captura pela cadeia (e seriam aniquilados no estágio seguinte de operação da cadeia), enquanto no segundo aquela região é de importância secundária (e ignorada por análises físicas de precisão, como no caso da reconstrução de H por elétrons como estados finais [10, 184]) devido à baixa resolução de energia. Assim, as análises futuras devem focar na região de precisão do ATLAS com a descrição das nuances operativas e comparação da resposta do algoritmo para dados de simulação e de colisão é prioritária para o êxito do algoritmo — ainda que garantindo performance razoável na região do *crack*, especialmente no SF_{on} . Mesmo desconsiderando esses fatores, o Algoritmo *NeuralRinger* possibilitava um acréscimo de $P_{D,NeuralRinger} \gtrsim P_{D,T2Calo} + 3\%$ dos elétrons e queda de $P_{F,NeuralRinger} \lesssim P_{F,T2Calo} - 10\%$ na taxa de falso alarme para toda faixa dinâmica em E_T , o que demonstra o potencial que o algoritmo possuía em dados simulados.

Por outro lado, como dito, sabe-se que os dados simulados não representavam bem os dados de colisão, o que é uma das causas para a baixa performance de ambos algoritmos em [216]. Mas além disso, as duas *runs* empregadas para estudo pelo trabalho foram de período muito inicial, com baixíssima luminosidade e cortes do Sistema de Filtragem do ATLAS muito baixos (L1_EM2). Isso significa que esses eventos estavam justamente na pior faixa (2–15 GeV) de operação do calorímetro¹⁶, onde ocorre grande depreciação na performance mesmo após todo o progresso realizado pelo ATLAS nesses intervalos. A estratégia inicialmente adotada pelo ATLAS durante a coleta de dados de 2010 foi remediativa, uma vez que esses desvios ainda não eram completamente compreendidos. A solução deu-se pela adaptação dos limites de decisão dos algoritmos em operação no Sistema de Filtragem do ATLAS colocando os cortes nas caudas das distribuições obtidas em dados de colisão (via decaimentos dos bósons Z e W).

6.2.4 Emprego de Anéis Além do ATLAS

O estudo [105] propôs o emprego de descrição de calorimetria em anéis para aplicação na filtragem de elétrons no Experimento $D\theta$ enquanto esse experimento ainda estava em sua fase de projeto. A publicação desse trabalho (de 1989) antecede a proposta inicial para o ATLAS (Subseção 6.2.1). O mesmo adveio como uma continuação de um estudo envolvendo a viabilidade de aplicação de MLP para a identificação de elétrons [103]. A nomenclatura empregada pelos autores para a

¹⁶Inclusive, as análises do Sistema de Reconstrução (*Offline*) de Física do ATLAS só consideram elétrons acima de 4 GeV atualmente. [180, 181]

descrição em anéis de energia deu-se por *bins* radiais, possivelmente fazendo uma analogia às células de leitura do calorímetro com um histograma de duas dimensões. Posteriormente, o trabalho [222] (1990), constituindo de um tutorial para aplicação de redes neurais para físicos, fez referência direta ao trabalho descrevendo essa descrição usando o termo *anéis concêntricos*. O trabalho no qual a proposta inicial do ATLAS (Subseção 6.2.1) se baseou [204] não faz qualquer menção a esses trabalhos. Igualmente, não se há conhecimento por parte dos autores da proposta inicial para o ATLAS, que participaram do desenvolvimento do calorímetro em [204], de que se tenha ocorrido influência desses trabalhos para o emprego de informação anelada. Observa-se que, no caso de [204], a motivação deu-se por razão distinta: redução do número de canais avaliados em um protótipo de calorímetro.

Em [105], a arquitetura MLP empregando informação anelada utilizou função de ativação logística¹⁷. O conjunto de dados de sinal foi composto pela simulação de decaimentos $Z \rightarrow ee$ interagindo, por via do pacote de simulação Geant, com o calorímetro de urânio-LAr do $D\emptyset$, esse contendo células retangulares ocupando uma região de 0.1×0.1 no plano $\eta \times \phi$. Utilizaram-se as primeiras cinco camadas longitudinais do calorímetro e quatro anéis compondo um espaço de característica com dimensão 4×5 . Para o conjunto de ruído de fundo, simularam-se dois jatos hadrônicos de alto momento ($80 < p_T < 120$ GeV), o que completou a base para treinamento¹⁸. A arquitetura da rede foi $20 - 8 - 2$ (20 entradas) com alvos $(1, 0)$ para elétrons e $(0, 1)$ no outro caso. Avaliou-se a rede em um conjunto independente de teste, obtendo a performance de $(P_D = 90,2\%; P_F = 0,9\%)$, valor superior a de um algoritmo baseado em cortes que operou em $(P_D = 75\%; P_F = 5\%)$. A melhor performance também se confirmou para jatos menos energéticos, onde o estudo avaliou casos até 40 GeV (com a rede treinada para jatos de $80 < p_T < 120$ GeV).

¹⁷O texto refere apenas a uma função sigmoïdal, porém cita que as saídas são limitadas em $(0, 1)$, o que leva a acreditar que a função de ativação logística tenha sido empregada.

¹⁸Não foi possível localizar a amostragem para treinamento.

Capítulo 7

Algoritmo *NeuralRinger*

Neste capítulo, encontram-se a descrição do método envolvendo a aplicação e análise do *NeuralRinger* para a seleção *offline* de elétrons do Experimento ATLAS (Seção 7.3), conforme proposto e desenvolvido por este trabalho. Nesse caso, ocorre a necessidade de fusão da informação proveniente do ID, onde a estratégia abordada foi o ajuste por redes neurais especialistas (Subseção 7.3.1). Para que a técnica possa ser empregada na análise física, realizou-se o desenvolvimento de uma infraestrutura completa nesse ambiente para a obtenção da informação anelada e a aplicação dos classificadores (Subseção 7.3.4). O desenvolvimento do *NeuralRinger* no *offline* deve ser uma generalização de sua contraparte atuante na filtragem *online* e, portanto, absorver atualizações realizadas nesse ambiente. Por isso, descreve-se anteriormente o método para atuação *online* (Seção 7.1). O desenvolvimento do mesmo não pode negligenciar as necessidades finais do experimento (Subseção 5.7.2). Os métodos estatísticos desenvolvidos por este trabalho a fim de elucidar os impactos causados na referência empregada para a seleção de elétrons nas análises físicas, em atuação *offline*, estão descritos na Seção 7.2.

7.1 *NeuralRinger Online*

Os trabalhos anteriores (Seção 6.2) mostraram a boa capacidade do método *NeuralRinger* para atuação na discriminação de elétrons aplicando apenas com informação de calorimetria no estágio equivalente ao atual FastCalo. Na época, a atuação de um único modelo não-linear para tomada de decisão de posse da informação normalização por norma-1 dos anéis era suficiente para disponibilizar uma técnica superior à referência da época, baseada em cortes lineares nas grandezas físicas (*CutBased*).

Com a demanda por estratégias mais eficientes devido à sempre crescente luminosidade do LHC, o ATLAS alterou sua estratégia para o uso de estratégia multivariável (Seção 5.3) para as análises *offline*, ainda no final da *Run 1*. Durante a

Primeira Parada Técnica (TS1), essa estratégia, baseada em verossimilhança (*LLH*), foi estendida para atuação no estágio preciso do HLT, onde se costuma empregar técnicas com a mesma abordagem que o *offline*. Todavia, essas alterações não atingiram o FastCalo, de forma que, apesar da redução da taxa de saída do HLT, vital para manter o HLT atendendo às demandas da física e dos recursos disponíveis, essas alterações tiveram pouco impacto na demanda por processamento do HLT. O *CutBased* continuou atuando no estágio rápido com um corte pouco restritivo a fim de evitar a perda prematura de elétrons devido às limitações de uma estratégia linear.

Esse cenário se mostrou extremamente favorável para a alternativa que o *NeuralRinger* representava, havendo uma demanda que já se desenhava em 2014 e, com as previsões de limitação de processamento em 2016, se tornaram ainda mais urgentes.

Existe forte motivação por parte da física para a extração de modelos por região específica das variáveis impactando na resposta das análises (Seção 5.3). As análises físicas envolvendo a comparação da seção de choque do esperado teórico, obtido através de simulação, com o observado experimentalmente geralmente consideram a divisão do espaço definido do plano (espaço de fase) composto pela cinemática e por um limite geométrico em regiões. A definição das regiões em η guiaram-se por fronteiras geométricas resultando em alguma alteração na instrumentação do experimento, seja pela quantidade, qualidade ou total ausência. Já as fronteiras em função de cinemática da partícula são definidas em função da influência dessa variável no desenvolvimento do chuveiro.

Outro fator que influencia nessa definição é a necessidade de estatística necessária para o objetivo final com o qual se emprega a informação. Para a análise e definição de limiares de decisão, regiões mais finas são definidas quando em comparação com aquelas definidas para o ajuste de modelos. Atualmente, somente os limites das regiões de E_T sofrem alteração. No caso da *LLH*, as regiões para cada caso estão disponíveis na Tabela 5.2 e 5.3.

Havia, assim, um forte indício de que o emprego de um *ensemble* de classificadores para tomada de decisão em informação anelada seria positivo para a estratégia *NeuralRinger*, o que motivou o seu desenvolvimento. O emprego de um *ensemble* foi benéfico para possibilitar uma resposta mais suave do classificador em função das regiões do espaço de fase.

A disponibilidade de uma grande massa de dados para o ajuste desses modelos exigiu o desenvolvimento de uma infraestrutura computacional integrada às fontes de processamento, capaz de atender a demanda envolvida. Essa infraestrutura foi desenvolvida em grupo com a atuação e supervisão do autor deste trabalho (Seção 7.4).

Além disso, também se fez necessário o desenvolvimento de um método de clas-

sificação resiliente ao empilhamento. A estratégia baseou-se no conhecimento especialista obtido para a referência *LLH* (Seção 5.5), onde a queda de eficiência era corrigida através de um limiar adaptativo em função de um estimador de empilhamento. O emprego de um classificador resiliente ao empilhamento fornece uma vantagem do *NeuralRinger* em relação ao *CutBased* disponível, uma vez que este não possui correção.

Adiante se encontram a descrição do método empregado para a tomada de decisão realizado nessa infraestrutura (Subseção 7.1.1) e aquele empregado para a análise (Subseção 7.1.2).

7.1.1 Método para Tomada de Decisão

Os parâmetros envolvidos no desenvolvimento do modelo para a tomada de decisão (comparar a descrição desta subseção com a Tabela 7.1) partiram do conhecimento agregado pelo histórico de aplicação do *NeuralRinger* (Seção 6.2). Porém, houve uma série de atualizações para o desenvolvimento dos modelos finais envolvidos.

O método inicia-se através da escolha da seleção de amostras para o treinamento. A fim de possibilitar uma maior redução de processamento no SF_{on} , escolheu-se realizar o ajuste dos modelos em decaimentos de $Z \rightarrow ee$, que fornece estatística para o ajuste e treinamento de elétrons para todas as cadeias primárias atuando no ATLAS sem a aplicação de fator de pré-escala. Os conjuntos deram-se por simulações (*mc15c*) desenvolvidas para o ajuste de modelos do SF_{on} para 2017 contendo decaimentos $Z \rightarrow ee$ para a coleta de elétrons e de jatos hadrônicos com um filtro exigindo que os mesmos contenham no mínimo 17 GeV. Aplicou-se o método T&P e a verdade de MC para selecionar os elétrons. A seleção de eventos para o ruído físico está sujeita a menor contaminação e, por isso, aplicou-se somente a remoção de eventuais decaimentos contendo elétrons isolados nos canais $Z \rightarrow ee$ e $W \rightarrow e\nu$. O número de amostras em cada região do espaço de fase está disponível na Figura 7.1.

Em seguida, a definição das regiões do espaço de fase para a tomada de decisão é realizada em dois níveis: ajuste do modelo, e ajuste do patamar de decisão. Empregou-se uma grade retangular em ambos os casos. As regiões definidas basearam-se na geometria do ATLAS, conforme avaliada pela estratégia de *CutBased* atuando no FastCalo tanto em E_T quanto em η . O *CutBased* emprega cortes definidos em função de regiões idênticas àquelas para os limiares de decisão empregados na *LLH* em *offline*² disponível na Tabela 5.3 e 5.2. Todavia, como foi realizado pela

¹Devido a observações de RoI obtidas pelo *L1Calo* na borda da região do detector gerando eventuais candidatos com $|\eta| > 2,5$ quando computado pelo FastCalo ou por algum erro na configuração do menu.

²Na realidade, essa grade provém do método *CutBased offline*, que governou o desenvolvimento

Tabela 7.1: Parâmetros empregados para ajuste, escolha de topologia do modelo e escolha de modelo para operação para as redes compondo o *ensemble* para atuação no Sistema de Filtragem do ATLAS (SF_{on}). Exceto quanto especificado, os parâmetros empregados foram os valores padrões das técnicas.

Critério	Valor
Composição do <i>ensemble</i>	
Modelo	MLP totalmente conectada, 1 camada escondida ($\tanh(\cdot)$) e 1 neurônio na camada escondida ($\tanh(\cdot)$)
Regiões definindo a Estatística de Ajuste	$\eta : 0 \rightarrow 0,8 \rightarrow 1,37 \rightarrow 1,54 \rightarrow 2,5$
Extrapolação de Atuação	E_T [GeV] : $15 \rightarrow 20 \rightarrow 30 \rightarrow 40 \rightarrow 50 \rightarrow \infty$
Eventuais observações ¹ fora da região empregam a rede com a região de fase vizinha (menor distância euclidiana no plano $E_T \times \eta$)	
Treinamento das Redes	
Núcleo para o Ajuste	<i>FastNet</i>
Dados Empregados para Ajuste	Simulação contendo colisões $p - p$ (25 ns) à 13 TeV no centro de massa em um <i>plateau</i> de $\langle \mu \rangle$ entre 0 e 60 colisões médias (<i>mc15c</i>)
Normalização	Divisão pelo absoluto da soma em energia dos anéis
Função de Custo para Ajuste	MSE
Método de Retropropagação	RPROP (com parâmetro $\eta^+ = 1,1$)
Alvo de Elétron/Ruído Físico	+1/-1
Tamanho da Batelada	Amostragem da menor classe
Limite de Épocas de Treinamento	∞
Critério de Parada	Multiparada (apenas considerando a rede obtida com P_D próximo ao da referência)
Referência de P_D e P_F	Eficiências no estágio HLTCalo
Neurônios na Camada Escondida	5-20
Inicializações	100
Método de Validação Cruzada	<i>jack-knife</i>
Número de Subconjuntos	10 (9 compondo a base de treino, 1 para validação/teste)
Método de Divisão em Subconjuntos	Divisão em regiões uniformes sem a aplicação de permutação aleatória
Extração da ROC [223]	1.000 pontos uniformemente espaçados entre os alvos do modelo
Avaliação do Treino e Escolha de Modelo para Operação	
Escolha Ponto de Operação	Ponto da ROC mais próximo a referência (P_D)
Escolha da Melhor Inicialização	Menor P_F (conjunto de teste) dentro de uma flutuação de até $\epsilon = 0,2\%$ em relação à referência (P_D)
Escolha da Topologia do Modelo	Análise Manual (<i>box plot</i> e gráfico da média e RMS), definido cinco neurônios na camada escondida para todos modelos compondo <i>ensemble</i> (ver texto)
Escolha da Modelo de Operação	Menor P_F (toda estatística) dentro de uma flutuação de até $0,2\%$ em relação à referência (P_D)
Tomada de Decisão	
Modelo para o Ajuste do Limiar	Modelo de Operação
Dados Empregados para Ajuste	Colisões $p - p$ de 2016 à 13 TeV no centro de massa (GRL: v88), exceto <i>run</i> de referência (ver texto)
Método de Validação Cruzada	Nenhum
Regiões definindo a Estatística de Ajuste	$\eta : 0 \rightarrow 0,8 \rightarrow 1,37 \rightarrow 1,54 \rightarrow 2,37 \rightarrow 2,5$
Extrapolação de Atuação	E_T [GeV] : $15 \rightarrow 20 \rightarrow 30 \rightarrow 40 \rightarrow 50 \rightarrow \infty$
Estratégia para Tomada de Decisão	Conforme definido para os modelos neurais
Método de Ajuste	Ajuste linear em função de $\langle \mu \rangle$ até 40 colisões médias
Alvo do Ajuste	Reta Equilibrando a Eficiência com parâmetros [0, 20, 40]
	Manutenção do P_D na saída do HLT em relação às taxas obtidas pela cadeia de referência

LLH, necessitou-se do agrupamento dessas regiões em grupos maiores. O objetivo é possibilitar maior estatística no ajuste da maior quantidade de parâmetros envolvida técnica *LLH*.

Signal/Background available statistics

η					
3 : 1.54->2.50	s: 60725 b: 822873	s: 326499 b: 482897	s: 493423 b: 117075	s: 452193 b: 39815	s: 226161 b: 34941
2 : 1.37->1.54	s: 8970 b: 130035	s: 54147 b: 73299	s: 90713 b: 19024	s: 64745 b: 6563	s: 36901 b: 6236
1 : 0.80->1.37	s: 36366 b: 529623	s: 231170 b: 316623	s: 349752 b: 79293	s: 325101 b: 27275	s: 179527 b: 25302
0 : 0.00->0.80	s: 71169 b: 775904	s: 327843 b: 469003	s: 502084 b: 118154	s: 475286 b: 42051	s: 259527 b: 37883
	0 : 15->20 [GeV]	1 : 20->30 [GeV]	2 : 30->40 [GeV]	3 : 40->50 [GeV]	4 : 50->500000 [GeV]
					E_T

Figura 7.1: Número de amostras disponível por região do espaço de fase empregado para o ajuste dos modelos neurais a partir de dados de simulação disponibilizados para o desenvolvimento das estratégias de seleção do SF_{on} para 2017. Ver texto para os critérios de seleção.

vidos nos modelos neurais, quando em comparação com um único limiar de corte empregado na estratégia *CutBased*. Assim, optou-se por dividir o espaço em quatro regiões em η (Figura 7.2 e Tabela 7.1): barril central, barril estendido do *TileCal*, região do *crack* e tampa; e cinco regiões cinemáticas ao agrupar as regiões de forma para conter o alcance em intervalos de uma (1) dezena de GeV para as regiões acima de 20 GeV. Os valores para definir a região em que a partícula pertence provêm da reconstrução do FastCalo e, portanto, estão sujeitos a flutuações substancialmente maiores que aqueles do estágio preciso e do *offline*. Essas regiões são empregadas durante todo o processo de treinamento e escolha dos modelos compondo o *ensemble* de operação, e totalizam 20 modelos neurais. Os modelos foram designados para operarem exatamente na região em que os mesmos foram ajustados, exceto em caso de partículas que eventualmente estejam fora do espaço de fase, onde se emprega a rede ajustada no espaço de fase mais próximo (distância euclidiana) ao ponto de interação da partícula.

O ajuste das redes ocorre com o algoritmo RPROP [224] (Apêndice C) iniciadas aleatoriamente conforme [225]. O critério de parada empregado é o multiparada (Seção 7.4), onde se obtém para o treinamento três redes para pontos de parada avaliando a especialização no conjunto de validação por critérios diferentes durante o treinamento. Os critérios são o SP_{max} , máximo em P_D para uma referência de P_F e mínimo em P_F para uma referência de P_D . O objetivo é obter melhorias para os modelos quando em operação com probabilidade de detecção de sinal (P_D) ou P_F próximos ao do HLTCalo (referência). Inicialmente, a referência empregada foi a eficiência obtida pelo algoritmo *CutBased* no FastCalo; porém, percebeu-se

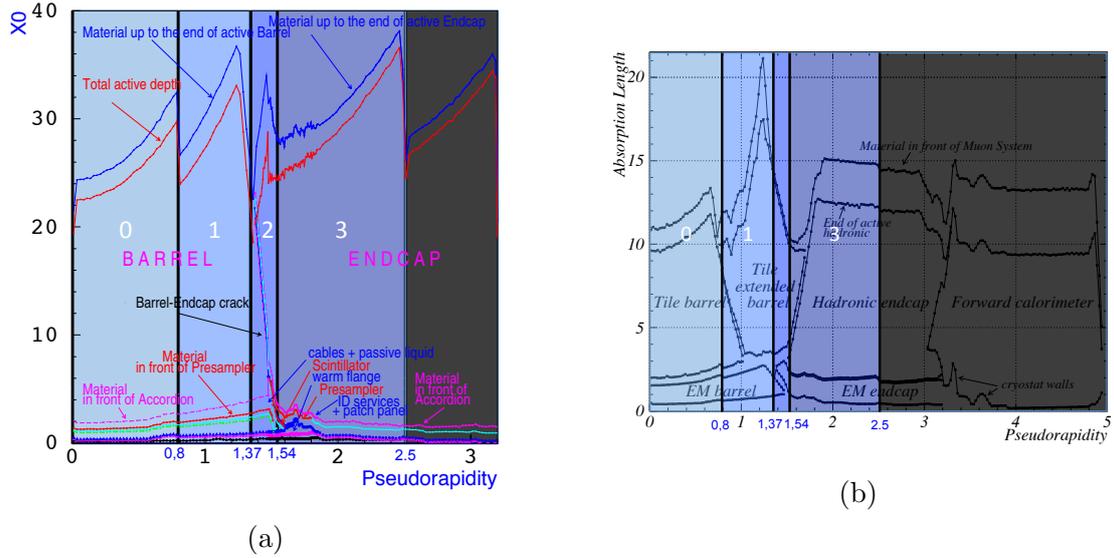


Figura 7.2: Esboço do comprimento simulado em unidades de X_0 do ECAL (a) e λ_{int} do Sistema de Calorimetria (b) contido nas regiões em η empregadas para análise da informação anelada de calorimetria. Os índices dentro das faixas azuis indicam o índice da região em η empregada para o ajuste de redes neurais, respectivamente barril central, barril estendido do *TileCal*, região do *crack* e tampa. Omitiu-se por questões visuais o índice da região do *crack* em (b). Os limites em η no eixo x na cor azul descrevem as fronteiras das regiões.

que o *NeuralRinger* era capaz de substituir toda a sequência de seleção composta por $\text{FastCalo}(\text{CutBased}) \rightarrow \text{FastElectron} \rightarrow \text{HLTCalo}$ (majoritariamente baseada apenas em calorimetria). Assim, a cadeia *ringer* deu-se conforme a sequência de cortes disponível na Figura 5.10. A referência durante os ajustes se davam pela eficiência obtida nas cadeias após o estágio HLTCalo, com o objetivo de obter a melhor redução de processamento na fazenda do HLT. Como se desejava que o *NeuralRinger* não impactasse na eficiência do SF_{on} , apenas a rede obtida para P_D similar ao da referência foi avaliada.

Para a validação cruzada [128], existem dois métodos implementados, ambos estratificados, na infraestrutura de ajuste (Seção 7.4): *jack-knife* e *bootstrap*. Optou-se pelo primeiro devido à sua menor demanda computacional dado a grande massa de dados. Além disso, o treinamento foi realizado em dados de simulação para posteriormente terem seus resultados avaliados em dados de colisão. Por esse motivo, empregou-se como conjunto de validação (parada para evitar especialização durante treinamento) o conjunto de teste de eficiência da rede. A estratégia para evitar performance subótima devido à convergência de primeira ordem do algoritmo RPROP, deu-se por força bruta, repetindo o processo de treinamento 100 vezes com novas inicializações para cada uma das configurações. Todas as redes ajustadas e respectivos históricos são salvos para possibilitar a avaliação do treinamento posteriormente. Obtêm-se os valores de P_D e P_F através de uma *Receiver Operation Curve* (ROC)

empregando 1.000 *bins* linearmente distribuídos dentre o suporte do espaço de saída. Das 100 inicializações, apenas aquela com o menor P_F no conjunto de validação é empregada para o cálculo da eficiência de validação cruzada, obtida em termos de P_D , P_F e

$$SP = \sqrt{\sqrt{P_D(1 - P_F)} \frac{P_D + (1 - P_F)}{2}}. \quad (7.1)$$

A escolha da topologia é realizada subjetivamente através da comparação das eficiências de validação cruzada em gráfico por *boxplot*. Avaliou-se a escolha da topologia, porém se observou pouca alteração em eficiência conforme a escolha dos neurônios para todas as regiões do espaço de fase. Dado a operação *online* e a melhor capacidade de generalização de redes com menor neurônios, manteve-se a configuração com o menor número de neurônios para compor o *ensemble*. Esse resultado foi integrado ao método posteriormente para reduzir a demanda computacional e possibilitar um menor ciclo de treinamento. Um pequeno ciclo de treinamento é extremamente importante para viabilizar a sintonia da técnica com as necessidades do experimento, em especial ao considerar que os ajustes finos ocorrem próximos aos prazos finais, uma vez que somente nessa etapa grande parte da estrutura que atuará na tomada de dados está disponível para ser avaliada.

Por sua vez, a escolha da rede de operação é realizada avaliando em todas as redes treinadas com a topologia escolhida, inclusive todas inicializações, aquela que obtém menor P_F quando operando no P_D da referência, ambos calculados de posse de toda a estatística a fim de evitar o emprego de um subconjunto com estatística polarizada na escolha. Assim, encerra-se o processo de ajuste dos modelos.

Cabe a realização da otimização dos limiares de operação. Dado as limitações da simulação em modelar a interação das partículas com a matéria do detector, o ajuste é calculado em dados de colisão. O impacto na saída das redes neurais (Figura 7.3) compondo o *ensemble* estão de acordo com o conhecimento da física (Seção 5.5). No caso, empregou-se os dados de colisão obtidos durante 2016 que continham as condições esperadas para boa parte da tomada de dados durante 2017, removendo-se a *run* (de referência) empregada para o comissionamento da versão em *software* durante a coleta de dados.

Para conferir resiliência ao efeito de empilhamento, considerou-se a estratégia empregada no *offline* e estágio preciso do HLT (Seção 5.5), referencias para o Fast-Calo e que empregavam estratégia multivariada. Os mesmos realizam um ajuste linear em função de um estimador de empilhamento, no SF_{on} dado por $\langle \mu \rangle$. No caso do *NeuralRinger*, removeu-se a função de transferência do neurônio de saída a fim de evitar a propagação de não-linearidades adicionais para o ajuste linear. Essa estratégia se mostrou fundamental para possibilitar um comportamento linear na

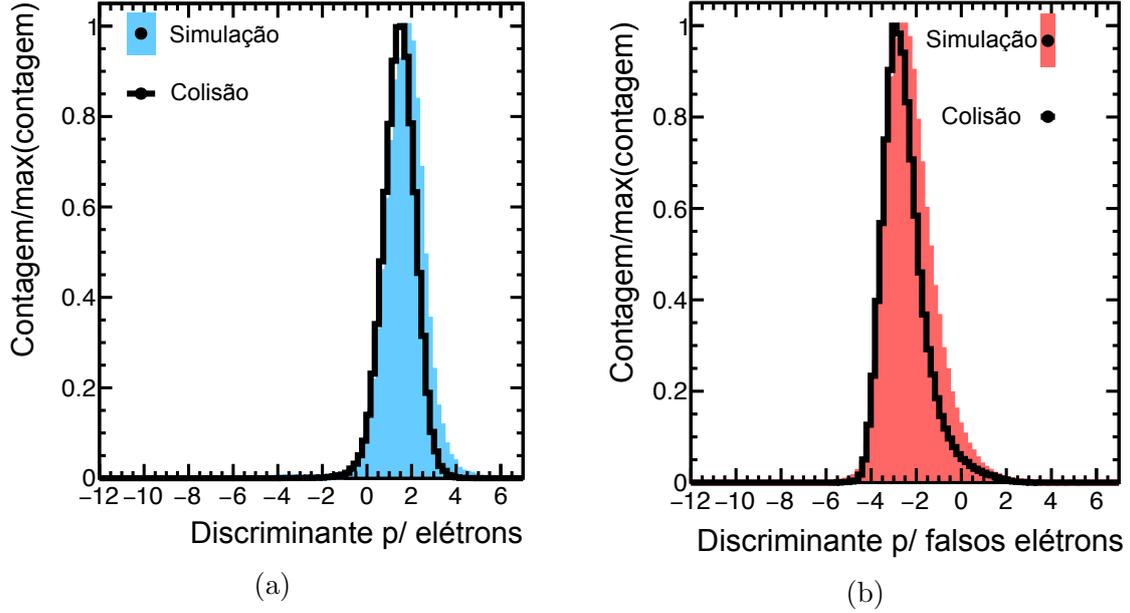


Figura 7.3: Saída da rede neural compondo o *ensemble* neural que entrou em operação em 2017 com a função de ativação não-linear removida após o estágio de treinamento para elétrons (a) e ruído físico (b) para a região $0 < |\eta| < 0,8$ e $15 < E_T < 20$. As envoltórias em preto são obtidas das saídas da rede calculadas em dados de colisão de 2016, exceto a *run* de referência, aplicando o método T&P e exigindo o critério *vloose* computado pelo *offline* nos para remoção de impurezas. As áreas preenchidas em azul (elétrons) e vermelho (ruído físico) foram calculadas nos dados de simulação empregados para o ajuste dos modelos sem a separação em subconjuntos.

saída da rede neural para os critérios mais rígidos: *medium* e *medium*.

Esse ajuste é realizado para todas as regiões do espaço de fase em cima da rede escolhida para operação. A Figura 7.4 apresenta um exemplo de correção, onde se empregou o método de balanceamento de eficiência (Seção 5.5) excluindo regiões com pouca estatística. Um limite superior para o ajuste em $\langle \mu \rangle = 40$ foi colocado a fim de evitar que a correção gerasse demasiado aumento no P_F ³. Com isso, encerra-se o método de desenvolvimento da tomada de decisão.

7.1.1.1 Discussão

Cabe uma consideração sobre o conhecimento especialista obtido sobre a importância do emprego de T&P em simulação e dados de colisão durante o desenvolvimento do ajuste para a aplicação no ano seguinte.

O desenvolvimento de um ajuste para seleção no SF_{on} é um processo complexo. O leitor atento percebeu que a estratégia natural para o desenvolvimento em dados de MC em simulação seria o emprego direto da verdade de simulação para a captura

³Posteriormente esse limite se mostrou ortodoxo, podendo ter sido estendido até $\langle \mu \rangle = 60$.

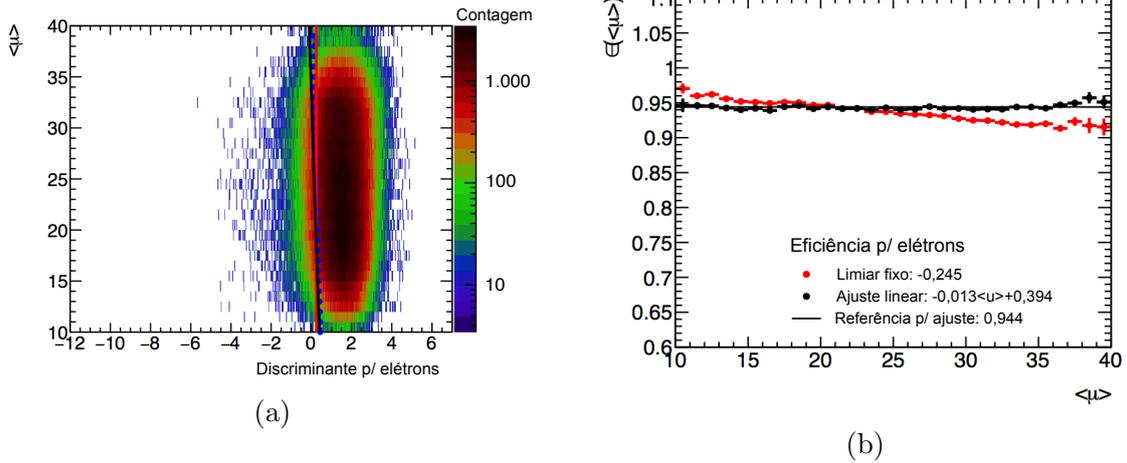


Figura 7.4: Saída (a) e P_D (b) da rede neural compondo o *ensemble* neural que entrou em operação em 2017 com a função de ativação não-linear removida após o estágio de treinamento para elétrons em função de $\langle \mu \rangle$ $0,8 < |\eta| < 1,37$ e $20 < E_T < 30$. Os elétrons são obtidos pelo método T&P em dados de colisão de 2016, exceto *run* de referência, e exigindo o critério *vloose* computado pelo *offline* nos para remoção de impurezas.

dos decaimentos de Z, sendo redundante o emprego de T&P quando avaliando dados de simulação já que a verdade de MC coletaria todos esses eventos. O T&P ainda resultaria em perda de estatística, logo, não haveria qualquer motivação para o seu emprego dado que se sabe a chave de MC. Todavia, a chave de MC não é perfeita devido à complexidade do processo da simulação. Durante o processo, há uma enorme quantidade de partículas geradas devido às interações com o material do ATLAS. O processamento completo dessa sequência seria um processo muito custoso e complexo para precisar quais foram exatamente as regiões que se originaram devido a um objeto físico gerado durante as colisões. O que se realiza, na realidade, é um casamento em posição entre o traço do candidato a elétron com a informação de MC e, portanto, sujeita a erros. Ao aplicar o T&P em conjunto com a verdade de MC, reduz-se essas ocorrências.

No caso de colisão, a motivação é clara para o emprego de T&P uma vez que não se sabe a origem da partícula e esse método permite obter uma pré-seleção de elétrons. Porém, para o ajuste de novos modelos para operação no SF_{on} a partir de dados coletados com os métodos de seleção anteriores, esse método possui um aspecto muito importante de ser observado. Primeiro, boa parte da taxa coletada de elétrons decorre de cadeias exigindo a existência de ao menos um elétron. Segundo, as cadeias de objetos múltiplos, como aquelas exigindo ao menos dois elétrons, só fazem sentido quando dedicadas à coleta de elétrons de menor energia que aquelas dedicadas à coleta de ao menos um elétron. Logo, quando exige-se no T&P (Seção 5.6) que o evento tenha sido coletado a partir das cadeias principais para a

coleta, elimina-se a estatística proveniente dessas cadeias, restando somente objetos provenientes das cadeias primárias dedicadas à coleta de ao menos um elétron. Eventualmente, os dois elétrons do T&P foram aceitos quando analisados pela cadeia primária, porém isso não gera qualquer tendência para a estatística de treino, uma vez que a decisão da cadeia sobre o segundo elétron é irrelevante para o método que exige somente a presença de ao menos um elétron. O aspecto a ser observado é que, caso a massa de dados selecionada para o treino não fosse proveniente do método T&P, haveria uma quantidade maior de dados com os padrões reconhecidos pela cadeia anterior. Considere-se o caso, por exemplo, que os dados fossem selecionados aplicando somente a decisão do *offline*. Nesse caso, os padrões selecionados para o ajuste dos modelos estão sujeitos a uma dependência direta da decisão formada pelo sistema *offline* e acrescentado da decisão do SF_{on} que operou durante a coleta de dados. Assim, isso poderia criar uma tendência no treinamento, uma vez que o modelo observaria com maior frequência padrões aceitos pelo antigo paradigma do SF_{on} que esteja em sincronia com a atual versão do *offline* (que espera-se que seja grande, já que o *offline* analisa o evento em maior profundidade). Com o emprego do T&P, essa tendência nas amostras é removida.

Outro ponto interessante a ser observado é o caso da coleta de dados de colisão ter sofrido alterações durante o período, como foi o caso da intervenção realizada ao alterar a cadeia para o emprego do *ringer* em 2017. A exigência do evento ser aceito por uma cadeia é avaliado pela reconstrução do SF_{on} em *offline* com a sua (mais recente) versão disponível durante o reprocessamento. Isso significa que, caso o SF_{on} tenha sofrido alterações durante o processo, somente as amostras sendo aceitas pela última versão da cadeia são mantidas para a análise *offline*. Assim, decorre a importância dos métodos utilizados terem alta sintonia de decisão para garantir menor perda de amostras obtidas com versões anteriores do SF_{on}.

7.1.2 Método de Análise

As análises do SF_{on}, exceto quando avaliado pelos plantonistas durante a coleta de dados, ocorrem em *offline* após o reprocessamento dos dados. Durante o reprocessamento, automaticamente são realizadas análises para a monitoração de eficiências do SF_{on} através das principais estratégias definidas para a sua medição. Os mesmos são dados pela aplicação de T&P para decaimentos $Z \rightarrow ee$ ($J/\Psi \rightarrow ee$) para a região $E_T > 15$ GeV ($E_T < 15$ GeV). Essas análises são tendenciosas em relação ao atual *offline*, onde se emprega o critério equivalente do *offline* para o qual a cadeia está dedicada para coleta (*offline medium* no *probe* para cadeias *medium*). Com isso, obtém-se a medição de eficiência em relação à referência atual das análises física. Como essas análises são para monitoração, apenas uma parcela pequena do menu

de cadeias são avaliados a fim de reduzir o impacto de processamento e tamanho dos dados.

O método para a análise de eficiência de novos métodos para atuação nas cadeias de elétrons no HLT estão sujeitos aos mesmos passos, de forma que as versões desenvolvidas para o SF_{on} sempre estão avaliadas em relação à referência da análise física. Decorre, portanto, que se uma nova técnica tiver maior discordância com o *offline*, a mesma terá menor eficiência de detecção de elétrons durante o processo de avaliação, garantindo que esse comportamento seja levado em consideração.

Outra análise interessante, aqui referida por análise de quadrante, considera a ocorrência dos casos mutualmente exclusivos referidos na Seção 7.2. Naquela análise, o impacto é avaliado ao modificar a sistemática para seleção do evento contendo o(s) par(es) T&P. Porém, neste caso, os casos mutualmente exclusivos são considerados na análise T&P aplicando as seleções da cadeias no *probe* com o método descrito acima.

7.2 Análise de Impacto na Coleta de Elétrons

A discussão da necessidade desta análise está disponível na Subseção 5.7.2, aqui se limitando a considerar os pontos envolvidos para o desenvolvimento de um método capaz de avaliar o impacto da alteração da estrutura da cadeia no perfil das amostras avaliadas pelo SR_{off} .

Assim como a discussão envolvendo a questão da introdução de tendências na massa de dados empregada para o ajuste de modelos no Tópico 7.1.1.1, há uma série de nuances envolvidas para possibilitar esta análise. Primeiro, como ponto trivial, é necessário que hajam as duas versões da cadeia atuando para a coleta de dados, caso contrário não se haverá acesso aos padrões rejeitados por uma dessas cadeias. Cabe então decidir qual a estrutura do menu a ser empregado durante a coleta de dados de forma que seja possível realizar esse estudo. Para isso, pondera-se:

1. A cadeia primária sem a aplicação de fator de pré-escala de menor energia (26 GeV) no ATLAS para coleta de elétrons emprega o critério *medium*, de forma que a escolha de outro critério de seleção implicaria na aplicação de fatores de pré-escala para coletar elétrons com mesma energia;
2. Da mesma forma, o critério *medium* opera com menor eficiência e, portanto, está sujeito a maior margem de discordância entre as cadeias de natureza distintas. Logo, espera-se que os impactos apareçam primeiro para esse critério, precisando de menor estatística para que possam ser observados;
3. Uma vez que o critério *medium* está sujeito à maior margem de discordância, não há motivos para acreditar que um impacto pior ocorreria para um critério

menos exigente, especialmente quando considerando que os diversos critérios para a cadeia de uma dada natureza foram ajustados com a mesma abordagem (mesmos dados, seleção, algoritmos, parâmetros, modelos etc., exceto pelo alvo definido para a tomada de decisão);

4. A avaliação em cadeias de maior energia causaria maior alteração da sistemática no T&P empregada nas análises *offline* (cadeia primária de menor energia de 26 GeV);
5. Ao considerar os itens anteriores, não há motivos para crer que o emprego de mais duplicações de cadeias agregariam valor ao estudo, porém reduziriam os ganhos de processamento trazidos pela atuação da nova técnica;
6. A quantidade de estatística disponível para o estudo é vital para possibilitar a observação de pequenas nuances e distorções no perfil dos elétrons coletados, sendo prioritário manter a atuação da duplicada sem pré-escala ou com a menor pré-escala possível.

Assim, optou-se por priorizar o emprego de uma única duplicata de cadeia *medium* com a menor energia possível sem a aplicação de pré-escala e para a coleta de eventos contendo ao menos um elétron. A fim de possibilitar melhores ganhos, a cadeia escolhida para ser duplicada considerou a coleta de ao menos um elétron com $E_T > 28$ GeV.

Cabe definir a estratégia para a análise. A identificação de elétrons no SR_{off} é governada pela LLH , cujos parâmetros são obtidos através de estimativas das *pdf* marginais para as 13 grandezas padrões (Seção 5.2) empregadas para o cálculo do discriminante. Assim, uma alteração nos parâmetros dessas marginais em relação àquele obtido pela cadeia *noringer* (Figura 5.11) implica em uma alteração do comportamento da LLH nas análises *offline*, que pode ser interessante ou não para a física. Desta forma, a análise pode ser realizada ao buscar por impactos nas marginais das grandezas físicas, não havendo necessidade de se preocupar com uma análise multivariável em cima dessas grandezas devido à LLH assumir independência estatística.

A extração do KDE das grandezas ocorre através de histogramas obtidos pelo método T&P (Seção 5.6) em dados de colisão, sujeito às considerações realizadas no Tópico 7.1.1.1. Para retirada das contaminações, uma abordagem ligeiramente tendenciosa é aplicada ao exigir o critério *vloose* para a limpeza dos . Replicou-se todo o método para a construção dos histogramas empregado na obtenção do KDE, porém alterando a abordagem do T&P para, ao invés de exigir a aceitação do evento com as cadeias primárias, conforme na descrição do T&P, considerar apenas os casos das cadeias duplicadas operando no menu de 2017.

Uma estratégia possível seria abordar o problema como um teste de ajuste, uma abordagem de grande popularidade e que consideram que o emprego de estatísticas bastante conhecidas, como χ^2 [226], Kolmogorov-Smirnov (KS) [227], Anderson-Darling (AD) [228]. No caso do χ^2 , essa estratégia pode ser estendida para computar os resíduos de cada observação em relação ao modelo. Todavia, assume-se que o modelo não possui incerteza estatística e ignora-se a dependência estatística entre as cadeias. A abordagem via os testes estatísticos podem ser estendida para avaliar se observações divididas em dois grupos foram retirados da mesma população [226, 229, 230]. Por exemplo, para a estatística χ^2 , há estratégias consolidadas⁴ para aplicação em diversas áreas, como ciências sociais [233], medicina [234] e, inclusive, Física de Altas Energias [235]. Essas estratégias se baseiam na análise de testes de contingência, sendo também referidas por teste de homogeneidade ou teste de consistência. Em boa parte dos casos, o que se deseja avaliar é se duas amostragens obtidas independentemente provêm de uma mesma população.

Esse não é o caso do problema em questão. As amostras aqui vêm de um *único* grupo de observação que está sujeito a dois processos de danificação [236] de amostras distintos representado pelas decisões das cadeias. Os casos abordados [226, 233] para essa solução consideram, geralmente, se os processos de danificação são independentes estatisticamente, mas não se a distribuição das amostras sobreviventes em dois grupos de observações permanecem inalteradas. Espera-se que as cadeias sejam altamente dependentes, dado que elas foram designadas para realizar a mesma tarefa: identificar elétrons.

As observações podem ser de quatro tipos mutualmente exclusivos, caso conhecido por ser modelado por uma binomial bivariável do tipo 1 [227] (no limite de grandes amostragens, uma Gaussiana bivariável) ou através de uma distribuição multinomial, nos grupos descritos na Tabela 7.2.

Tabela 7.2: Casos eventualmente exclusivos para a atuação das cadeias *ringer* e *no-ringer* em paralelo durante a tomada de dados.

Caso I	ambas cadeias aceitam o elétron
Caso II	apenas a cadeia <i>ringer</i> aceita o elétron
Caso III	apenas a cadeia <i>no-ringer</i> aceita o elétron (ou apenas <i>ringer</i> rejeita)
Caso IV	nenhuma cadeia aceita o elétron

Assim, para possibilitar a análise completa, construiu-se os histogramas obtidos com o T&P conforme empregado para a construção dos KDE, porém dividindo a abordagem do T&P para considerar os quatro casos mutualmente exclusivos. Com

⁴A ideia desse teste veio 11 anos [231] após a criação do teste para avaliação de ajuste por χ^2 , pelo próprio PEARSON e, da mesma forma, posteriormente aprimorada para lidar com a questão do número de graus de liberdade [232].

isso, obtêm-se quatro histogramas com as observações equivalentes para cada um dos casos alterando a lógica exigida pelas cadeias no método T&P conforme esses casos.

As observações do caso I são o ideal, uma vez que esse tipo de observação não pode gerar distorção no perfil das grandezas do SR_{off} . As distorções podem ocorrer a partir do caso II e III, enquanto o caso IV só pode ser observado devido às cadeias de suporte ou a partir de elétrons encontrados pelo *offline* que não foram reconstruídos pelo *trigger*. Essas observações não são necessárias para responder a questão de qual o impacto na distribuição original, cabendo avaliar somente as relações entre os casos I, II e III.

A estratégia por avaliação de resíduos considera a comparação dos casos compostos pelas marginais da binomial bivariável ao avaliar o caso I+II (elétron aceito pelo *ringer* independente da decisão do *noringer*) em relação ao I+III (elétron aceito pela cadeia *noringer* independente da decisão do *ringer*). Essa análise foi realizada em termos de resíduos em função da estatística χ alterada para manter o direção da distorção, conforme

$$\chi_{i,j}^s = \frac{(r_{i,j} - n_{i,j})}{\sqrt{n_{i,j}}}, \quad (7.2)$$

onde r_i (n_i) representam o número de observações do histograma obtido pela cadeia *ringer* (*noringer*) para a i -ésima região no plano $E_T \times \eta$ e $r_{i,j}$ ($n_{i,j}$) é o número de observações no j -ésimo grupo do histograma r_i (n_i) correspondente.

No teste de ajuste de modelos, não se considera a possibilidade de haver uma quantidade diferente entre o número de observações e aquele esperado no modelo. Isso ocorre devido a pequenas alterações na eficiência das cadeias já que há diferença entre as grades do espaço de fase e flutuações estatísticas no processo de ajuste tentando alcançar a manutenção do P_D . Assim, essa abordagem gerará resíduos não nulos quando computados por (7.2) que não implicam em uma alteração no perfil e podem ser modelados por

$$\frac{\mu_{i,j}}{\sigma_{ref}} = \frac{(r_i - n_i) \times \frac{n_{i,j}}{n_i}}{\sigma_{n_{i,j}} \approx \sqrt{n_{i,j}}}, \quad (7.3)$$

onde $\sigma_{n_{i,j}}$ pode ser considerada como a unidade para medição dos resíduos e será simplesmente referida por σ quando se referindo aos resultados por essa abordagem. A fim de beneficiar-se de aproximação de erro Gaussiano (empregando as recomendações conservadoras de [40]) os grupos do histograma são unidos para obter um mínimo de 30 contagens. A natureza descontínua da contagem de eventos limita a precisão dos erros esperados nas caudas das distribuições.

Essa abordagem permite uma avaliação qualitativa dos resíduos a fim de se

encontrar um padrão de distorção que fuja daquele esperado pelo modelo em (7.3).

Por outro lado, uma análise sem integrar I para a estatística permite avaliar diretamente uma comparação entre as amostras que possam causar distorção. Uma estratégia pela aplicação de estatística como o χ^2 entre os dois casos ignoraria a dependência gerada entre as contagens devido a natureza mutuamente exclusiva dos eventos. Considerou-se abordar esse problema empregando o valor de $-np_{(II)}p_{(III)}$ [227] para a variância e avaliar a estatística χ^2 não-central com correlação [237], porém isso exigiria avaliar até que ponto todas as hipóteses estatísticas se mantinham.

Uma estratégia mais simples dada a demanda por uma resposta da análise para avaliar o impacto durante a coleta de dados deu-se por simulação ao se basear em exemplos na Física de Altas Energias [238]. Todavia, o número de histogramas foi reduzido de 10^7 para 10^5 uma vez que não se fazia necessário a precisão de 5σ nas medições, acreditando ser suficiente medições com 2σ de precisão para avaliar a presença de distorção. Apesar da distribuição da cadeia *noringer* ser a referência, a mesma está sujeita à flutuação estatística. Assim, ao considerar que ambas observações vêm da mesma população, uma estratégia melhor se dá por empregar a estimativa de máxima verossimilhança dada por [239]

$$\hat{p}_{i,j} = \frac{r_{i,j} + o_{i,j}}{r_i + o_i}. \quad (7.4)$$

Assim, geraram-se dois histogramas mutuamente exclusivos 100.000 vezes para representar a distribuição esperada de em caso de similaridade devido a erros multinomiais. A estatística empregada foi a divergência Kullback-Leibler (D_{KL}) obtida por [240]

$$D_{KL}(p_i||q_i) = \sum_j (p_{i,j} - q_{i,j}) \ln \left(\frac{p_{i,k}}{q_{i,j}} \right). \quad (7.5)$$

A fim de possibilitar o acesso das contribuições individuais trazendo maiores distorções, os j -ésimos componentes calculados para a D_{KL} no histograma são separados para avaliação.

7.3 *NeuralRinger Offline*

O *NeuralRinger* foi proposto inicialmente para a operação no SF_{on}, porém, o desenvolvimento de novos métodos na área, especialmente para os níveis mais altos de seleção, ocorre no caminho inverso, ou seja, o desenvolvimento é realizado para a reconstrução *offline* e, posteriormente, o método é adaptado para atuar na filtragem *online*, sofrendo as degradações necessárias para atender as demandas de latência.

Do ponto de vista da física, esse caminho é a abordagem lógica, já que o SR_{off} é a referência responsável pela tomada de decisão final para o reconhecimento de padrões dos objetos físicos que serão empregados nas análises físicas. Além disso, o fato de não haver uma versão atuante do método no SR_{off} gerou a necessidade da análise de impacto, conforme discutido na Subseção 5.7.2 e que consiste de uma análise meticulosa com diversos pontos a serem considerados (Seção 7.2). Ainda, o ambiente *offline* consiste de um local propício para a avaliação da técnica sobre outra ótica, sem a avaliação tendenciosa em função de outros métodos, o que permite uma compreensão melhor de suas nuances operativas. Isso permite compreender as capacidades e efeitos da técnica para a atuação nas análises físicas. Para que isso seja possível, há necessidade de que o desenvolvimento ocorra na infraestrutura comum (Seção 4.3) de código (para reconstrução, simulação, análise etc.) do ATLAS, o *Athena*. A descrição da infraestrutura implementada encontra-se na Subseção 7.3.4.

A atuação do método no ambiente de análise *offline* exige que o mesmo disponha uma solução completa, diferente de sua atuação no Sistema de Filtragem do ATLAS, onde o mesmo atua em um ambiente híbrido e conta com a decisão posterior da *LLH* no estágio final do HLT. A fim de explorar todo o potencial discriminante da instrumentação do ATLAS, exige-se a fusão da informação proveniente do Sistema de Calorimetria com aquela do ID. Decidiu-se explorar o ajuste por redes especialistas para cada representação de informação, conforme descrito na Subseção 7.3.1. O método de ajuste para a tomada de decisão é descrita na Subseção 7.3.2 e o método de análise encontra-se na Subseção 7.3.3.

7.3.1 Ajuste de Redes Especialistas para Fusão de Informação

A alteração mais clara e urgente para a atuação do método no SR_{off} é a necessidade de avaliar uma estratégia para a fusão de informação. A estratégia que se decidiu desenvolver se baseou em [241], onde se explorou o ajuste de MLP especialistas a segmentação longitudinal por seções (ECAL, HCAL) do Sistema de Calorimetria do ATLAS. Ainda que essa estratégia seja interessante, o foco deste trabalho será no emprego do ajuste por redes especialistas para a fusão das diferentes representações de informação. Existem três representações possíveis para representações dos padrões discriminantes para o ajuste de redes especialistas:

- soma em anel concêntrico de energia (anel), representação parte do *Neural-Ringer* em atuação no SF_{on} e, portanto, deve ser parte de sua versão para a proposta no SR_{off} ;
- grandezas físicas de traço e de fusão especialista calorimetria-traço ($track^+$),

acrescenta informação discriminante proveniente do ID não empregada no *NeuralRinger online*;

- grandezas físicas de calorimetria (grandezas *shower*), fornece uma representação complementar de calorimetria aos anéis, contendo uma grandeza não-linear construída a partir das células de leitura ($w_{\eta 2}$), três grandezas que não podem ser construídas a partir de combinações lineares dos anéis (R_η , R_ϕ , E_{ratio}) e outras duas grandezas que podem ser obtidas através de combinação linear dos anéis (f_1 e f_3).

As etapas do ajuste de redes especialistas para fusão estão na Figura 7.5. O objetivo é realizar a fusão de informação no espaço latente especialista obtido na camada escondida das MLP após o treinamento de cada representação especialista da informação. Assim, após o ajuste dos parâmetros das redes neurais especialistas, removem-se seus neurônios de saída para conectar suas camadas escondidas a uma rede de fusão, composta por uma camada escondida e um único neurônio de saída próprios. Acredita-se que a fusão no espaço latente é benéfico para o processo devido à alta dependência dos anéis, o que dificultaria o processo de fusão ao haver um maior número de padrões para o ajuste dos discriminadores.

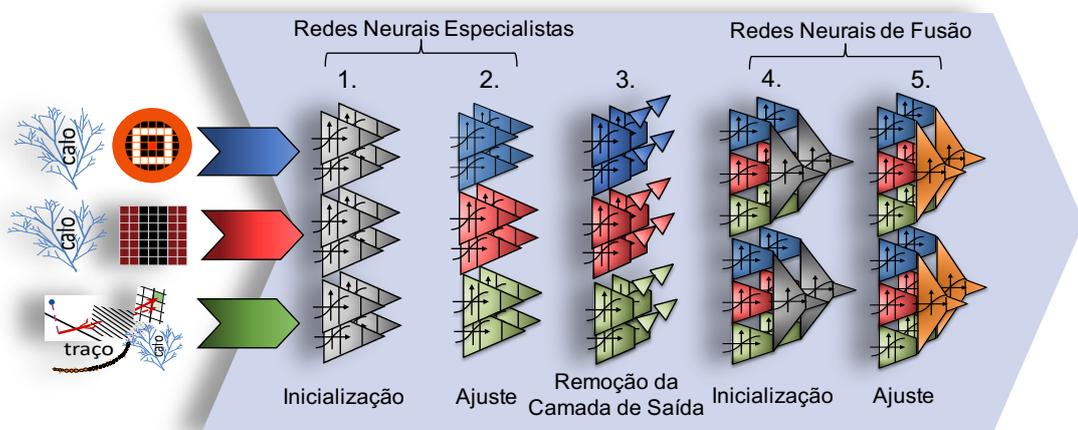


Figura 7.5: Esboço de um dos ajustes de redes neurais especialistas realizado neste trabalho para comporem o *ensemble* de redes neurais do *NeuralRinger offline* e suas respectivas etapas. O processo começa a inicialização e ajuste das redes neurais especialistas em somente uma das representações da informação. Em seguida, remove-se a camada de saída de cada uma das redes especialistas para então iniciar e ajustar a rede de fusão, totalmente conectada às sinapses das camadas escondidas das redes especialistas.

Além do emprego de outras representações de informação, há outra diferenciação em relação à [241], que realiza o ajuste da rede de fusão com as sinapses das redes especialistas congeladas. Dado que as representações de informações disponíveis

podem ter forte dependência entre alguns padrões, liberaram-se as sinapses das redes especialistas para sofrerem um ajuste fino durante o treinamento das redes de fusão. Esse processo se mostrou benéfico para possibilitar a captura de espaços latentes especialistas mais discriminante para a fusão da informação e está de acordo com a importância que se vem dando ao processo de ajuste fino em DL. Em [151, 242], avaliou-se que há ganhos consideráveis ao aplicar o ajuste fino supervisionado dos padrões obtidos não supervisionadamente através de *Stacked Autoencoders* e DBN. Ainda que a aplicação deste trabalho seja diferente, já que as sinapses da rede especialistas foram obtidos supervisionadamente, há indícios em [150] que o acesso à informação conjunta permite a obtenção de padrões mais poderosos (não supervisionadamente) e supervisionadamente (Tópico 3.2.3.2).

7.3.2 Método de Ajuste para a Tomada de Decisão

A seleção de dados para o ajuste deu-se na base de dados de simulação dedicada a representar as condições observadas durante as colisões $p - p$ à 13 TeV no centro de massa de 2016 (*mc16a*). Quando se refere em simulações para um dado perfil de colisões na Física de Altas Energias, faz-se menção a simulações que contêm o perfil de luminosidade instantânea gerado de forma que o mesmo se equivalha ao das colisões. Isso tem consequência direta no perfil de empilhamento simulado, como no caso do *nvtx* empregado no SR_{off} (Figura 7.6). De posse das observações realizadas no Tópico 7.1.1.1, a seleção de eventos aplica o método T&P para a coleta de elétrons além da verdade de simulação. Para a seleção de ruído físico, somente a verdade de simulação. A informação anelada é reconstruída pela infraestrutura desenvolvida para atuação nas análises *offline* (Subseção 7.3.4). O número de amostras em cada região do espaço de fase está disponível na Figura 7.7. O perfil para os elétrons selecionados para E_T e η estão na Figura 7.8.

O desenvolvimento do ajuste das redes ocorre em estágios separados, sendo possível aplicar parâmetros específicos para cada um dos ajustes. Há, todavia, dependência durante o desenvolvimento das etapas de fusão que necessitam do ajuste finalizado das etapas especialistas anteriores. As tabelas 7.3–7.6 contêm a descrição dos valores empregados em cada um dos estágios do desenvolvimento até a obtenção do modelo mais complexo responsável pela fusão de todas as informações.

Boa parte dos parâmetros empregados são idênticos àqueles citados na Subseção 7.1.1, dedicando-se aqui a descrever somente o que é distinto no método. A primeira questão para diferenciação (além do fato dos mesmos serem provenientes de uma reconstrução mais precisa realizada em *offline*) dá-se pela grade para o ajuste e atuação dos modelos no *ensemble*, onde se empregou a mesma da referência (*LLH*). O *ensemble* obtido atua com 36 modelos neurais oriundos de 4 regiões em E_T e 9

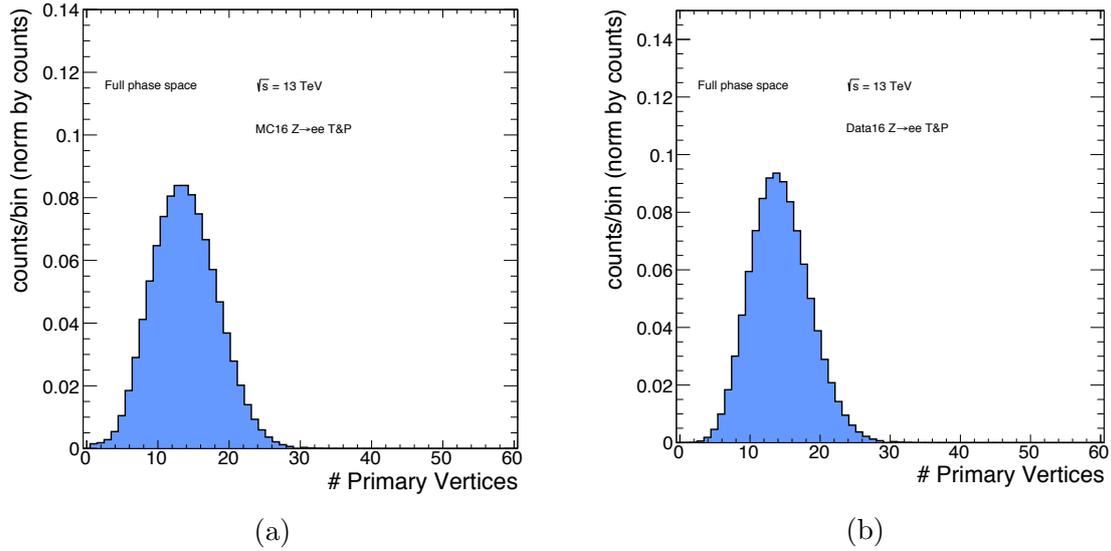


Figura 7.6: Perfil de $nvtx$ para as simulações de MC (a) empregadas no estudo para o *NeuralRinger offline* (mc16a) e dos dados de colisão (b) coletados pelo ATLAS durante o ano de 2016 aplicando o método T&P para decaimentos $Z \rightarrow ee$ e exigindo, no caso de colisão, o critério *vloose* para remoção de impurezas. Em ambos os casos, os elétrons são coletados na sua respectiva região de precisão ($|\eta| < 2,47$) e para $E_T > 15$ GeV.

Signal/Background available statistics

η	0: 15->20 [GeV]	1: 20->30 [GeV]	2: 30->40 [GeV]	3: 40->50000 [GeV]
8: 2.37->2.47	s: 28011 b: 190689	s: 127422 b: 108744	s: 245241 b: 25074	s: 321795 b: 16236
7: 2.01->2.37	s: 112770 b: 753840	s: 536946 b: 429288	s: 1103685 b: 97104	s: 1562916 b: 62166
6: 1.81->2.01	s: 64092 b: 415152	s: 319005 b: 235359	s: 694842 b: 54213	s: 1062351 b: 34851
5: 1.52->1.81	s: 93663 b: 759504	s: 486054 b: 419529	s: 1129038 b: 92529	s: 1831269 b: 57987
4: 1.37->1.52	s: 43962 b: 482994	s: 249855 b: 265143	s: 628254 b: 61113	s: 1055067 b: 37140
3: 1.15->1.37	s: 69927 b: 619308	s: 409842 b: 314826	s: 1064988 b: 69873	s: 1729662 b: 45978
2: 0.80->1.15	s: 112686 b: 871245	s: 689847 b: 487542	s: 1978338 b: 113178	s: 3018663 b: 73500
1: 0.60->0.80	s: 70551 b: 491217	s: 405981 b: 279036	s: 1243887 b: 65640	s: 1854609 b: 43359
0: 0.00->0.60	s: 231150 b: 1422279	s: 1494546 b: 805794	s: 3735015 b: 191943	s: 5849436 b: 126174

Figura 7.7: Número de amostras disponível por região do espaço de fase empregado para o ajuste dos modelos neurais a partir de dados de simulação disponibilizados para o desenvolvimento de *ensemble* neurais para a atuação no Sistema de Reconstrução (*Offline*) de Física do ATLAS (SR_{off}).

regiões em η . No *offline*, normalmente se emprega o p_T para a definição da grade, porém a influência do desenvolvimento a partir do *NeuralRinger online* fez com que se utilizasse o E_T ao invés do p_T , já que o método é governado pela descrição em anéis de energia. Apenas extrapolação para a região $2,47 < |\eta| < 2,5$ é realizada, já que essa região não é empregada no treinamento a fim de evitar contaminação por

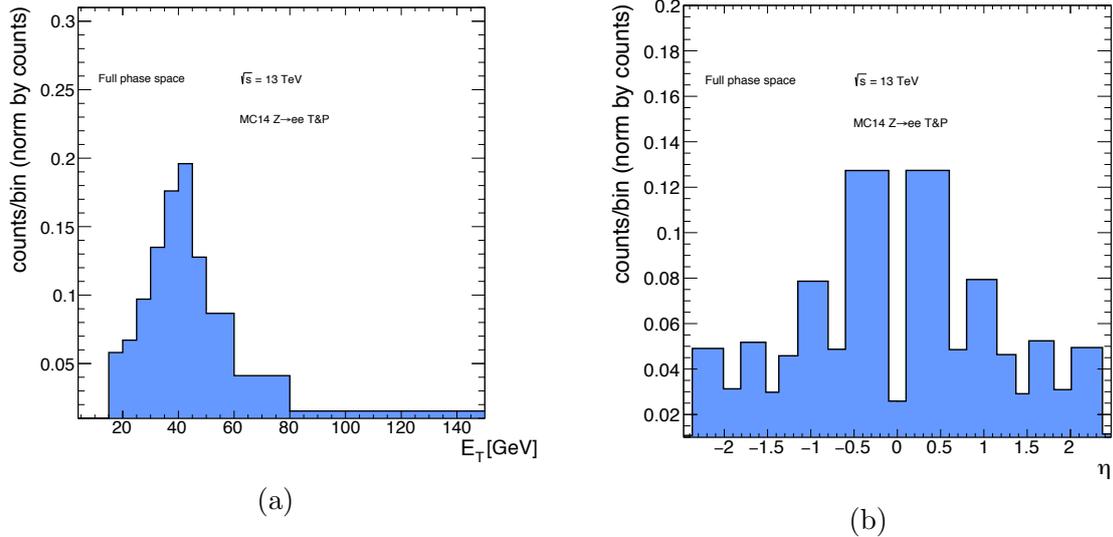


Figura 7.8: Perfil de E_T (a) e η para as simulações de MC empregadas neste trabalho (mc16a) aplicando o método T&P para decaimentos $Z \rightarrow ee$.

amostras obtidas na região de um pequeno *crack* do ECAL (Subseção 4.2.1).

O treinamento empregou como critério de parada a avaliação do SP no conjunto de validação desenvolvido em [198]. Isso se deve pela menor demanda do mesmo de memória em relação ao critério multiparada, um aspecto importante quando considerando que a massa de dados para o treinamento é cerca de 20 vezes maior que aquela empregada no SF_{on} . O consumo de memória limita o acesso aos nós menos sofisticados da LCG, reduzindo a capacidade de processamento disponível a uma fração dos sítios da nuvem e ao Lobo Carneiro. A fim de reduzir a necessidade de processamento, a abordagem inicial englobou apenas topologia mais simples, abordando somente 5–10 neurônios na camada escondida.

Outra vantagem para o emprego da função não-linear ($\tanh(\cdot)$) após o treinamento para uma função linear mostrou-se benéfica para possibilitar uma melhor alocação dos 1.000 pontos empregados para a construção da ROC, já que as distribuições no espaço de saída com a função linear de ativação possuem acessar melhor a dispersão dos dados. Com isso, aloca-se melhor os recursos da grade de busca, a fim de obter melhor precisão para o ajuste de operação. Considerou-se o ajuste para operação similar às referências *LLH* e *CutBased* em seus valores de P_D e P_F em toda a estatística disponível para o ajuste.

A escolha de topologia inicialmente era realizada através da avaliação gráfica. Todavia, com o grande número de topologias a serem avaliadas (36) para cada modalidade de informação, necessitou-se automatizar esse processo. A estratégia mais simples se deu pela escolha da topologia com melhor eficiência quando avaliados em toda estatística disponível. Acredita-se que essa escolha não polariza a validação

cruzada nos modelos de fusão criados a partir desses modelos, uma vez que esses modelos não tiveram acessos a estatística além daquela do conjunto de treino e validação, somente a escolha de sua topologia que poderia igualmente ser determinada visualmente a partir de qualquer critério subjetivo.

Diferente do ajuste de limiares realizado no SF_{on} que ocorreu posteriormente à escolha da topologia, a determinação dos parâmetros envolvidos no ajuste do limiar de decisão participa do processo de validação cruzada para avaliar a generalização e impacto desta técnica na eficiência final do classificador. E com isso define-se o treinamento para as MLP especialistas para descrição de energia em anéis de calorimetria.

O mesmo foi observado quando realizando o treinamento das redes nas grandezas de calorimetria (Tabela 7.5). A estratégia para a normalização foi a mesma que aquela empregada para o traço a fim de manter sua representação idêntica àquela empregada pela *LLH*. Por se constituírem essencialmente de razões em energia, boa parte das mesmas não precisaram da aplicação de um fator para normalização. O objetivo principal deste treinamento é verificar se a adição dessa informação aumenta a concordância das decisões do *ensemble* com a *LLH*.

A estratégia para a extração de informação em traço é similar (Tabela 7.4), somente sendo alterada a normalização para lidar com a descrição por grandezas de traços e fusão especialista calorimetria-traço via casamento de posição. A normalização deu-se pela aplicação de simples fatores de escala para o ajuste de algumas grandezas para o alcance dinâmico da função de ativação. Observou-se que o emprego de um número superior de neurônios na camada escondida para as grandezas de traços era benéfico para a rede. A hipótese que se trabalha para explicar esse fenômeno se dá pela natureza altamente compactada das grandezas físicas, de forma que ao inflar o espaço, seja possível produzir outros padrões envolvidos na informação do traço que sejam a partir dessa informação compactada por conhecimento possível. Assim, manteve-se as topologias avaliadas de 5 – 10 neurônios na camada escondida.

De posse da topologia e redes especialistas em cada sorteio, realiza-se o agrupamento das mesmas com a melhor inicialização para compor as sinapses de entrada do modelo a ser adicionado à parte dedicada a fusão da informação (Tabela 7.6). Realizam-se 100 inicializações da rede de fusão a partir da estrutura das redes especialistas e o procedimento se repete para a determinação da rede de operação, limiar de operação e obtenção das eficiências em de validação cruzada. Essa estratégia é aplicada para o treinamento dos *ensembles* neurais para fusão de informação de calorimetria e grandezas traço, bem como a fusão de toda as representações de informação disponíveis.

A eficiência desse ajuste será mantida para avaliação nas análises. A proposta

para atuação no SR_{off} é ajustada em base da fusão dessas duas redes e comparada com as referências LLH . Como se deseja reduzir ao máximo a dissonância dos métodos, considerou-se, também, a fusão da informação das grandezas físicas de calorimetria a fim de verificar se essa abordagem aumenta a harmônia das decisões. Além disso, as grandezas físicas de calorimetria possuem informações complementares aos anéis, como o caso de E_{ratio} , R_η e R_ϕ , que não podem ser construídas diretamente a partir dos anéis, $w_{\eta 2}$, uma grandeza com manipulação não-linear a partir da energia das células.

7.3.3 Método de Análise

Uma das vantagens para as análises do método em *offline* dá-se pela a avaliação do método com uma lógica de seleção dedicada, em contraste com as cadeias de filtragem onde o *NeuralRinger* atua como um método híbrido. Além disso, não se realiza a análise do método atuando em conjunto com a referência *offline*, como no caso das análises das cadeias de filtragem (Subseção 7.1.2).

Ainda assim, no campo da Física de Partículas, não é incomum a avaliação do comportamento dos métodos na própria massa de dados em que se empregou para o ajuste. Acredita-se que o efeito de generalização não seja levado em conta pois o ajuste dos métodos ocorrem em grande massas de dados, aonde cuidado é tomado para que haja estatística suficiente para a caracterização dos processos a partir de conhecimento especialista.

As análises avaliando as eficiências em função das variáveis compondo o espaço de fase, do estimador de empilhamento e o quadrante (Subseção 7.1.2) foram realizadas considerando esse contexto na Física de Partículas. As comparações de eficiência, todavia, são realizadas em validação cruzada para cada espaço de fase e em uma versão resumida agrupando todas as regiões em um valor global.

Além disso, existem quatro critérios de operação disponíveis para a atuação no *offline*. A análise de todos eles tornaria este trabalho demasiadamente extenso e foge do seu escopo trazer todos os pontos de operação do algoritmo. Como a análise física em busca de física nova envolve critérios mais relaxados devido à sua natureza ainda mais rara que o Higgs (os critérios atuais foram otimizados considerando explorar esse processo), as análises envolverão apenas o critério *vloose*.

7.3.4 Infraestrutura para Atuação no ATLAS

Como as análises ocorrem colaborativamente e em grandes massas de dados, havia a necessidade do desenvolvimento de uma infraestrutura em *software* que possibilitasse a sua avaliação dentro do ambiente comum de análise. Essa infraestrutura também é essencial para possibilitar a adição da informação anelada nos reproces-

samento dos dados de colisão e produções de simulação (Subseção 7.3.4.1) e, com isso, tornar viável o desenvolvimento do *NeuralRinger* para o SR_{off} .

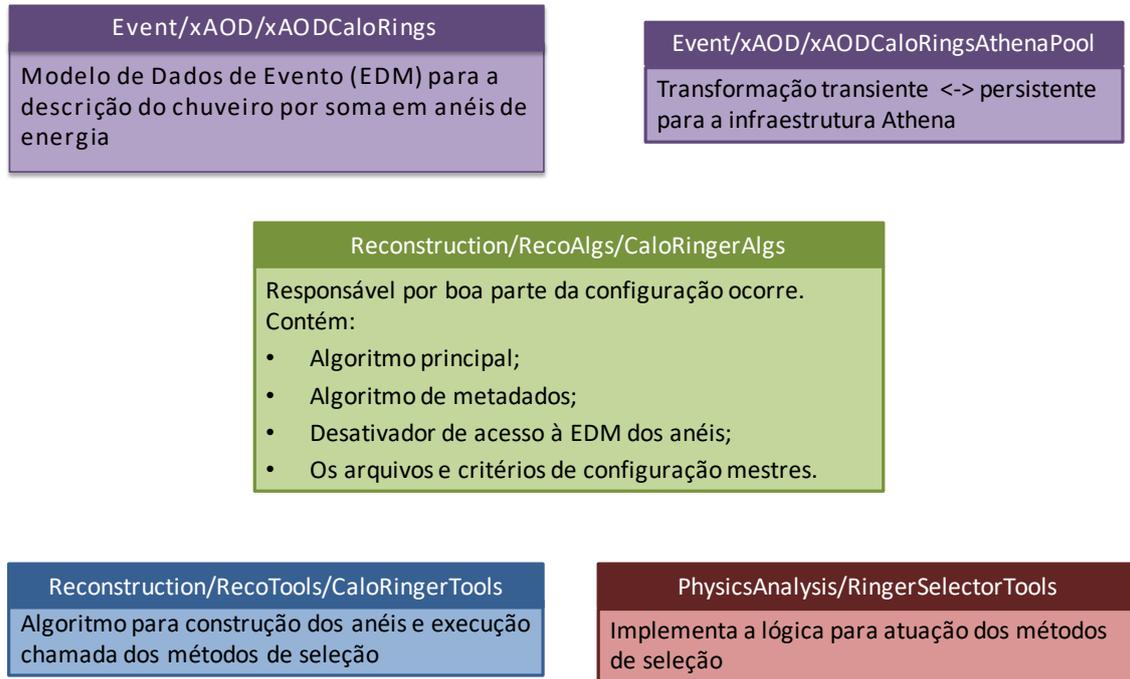


Figura 7.9: Módulos implementados para a geração da informação anelada e aplicação de seleção a partir de técnicas de Inteligência Computacional dentro da infraestrutura comum (*Athena*) do ATLAS.

Implementaram-se cinco módulos (Figura 7.9) baseados no código desenvolvido para a reconstrução e identificação de elétrons. Uma descrição mais aprofundada da infraestrutura pode ser encontrada no Apêndice B. Basicamente, a implementação da estrutura de representação dos dados encontra-se no módulo *xAODCaloRings*. O módulo *xAODCaloRingsAthenaPool* é responsável pela sua leitura/escrita em mídia persistente. A configuração e execução do algoritmo principal é realizada no módulo *CaloRingerAlgs* que chama a execução de ferramentas, disponíveis no módulo *CaloRingerTools*, dedicadas à implementação da lógica para cada objeto físico a ser analisado. Atualmente somente elétrons contêm lógica atuando durante a reconstrução, onde se decora a informação anelada de calorimetria previamente reconstruídos e aplica-se as lógicas de seleção para os diversos critérios a serem aplicados nas análises implementados no módulo *RingerSelectorTools*.

Apesar do esforço exigido⁵, isso possibilitou grande flexibilidade para o emprego da infraestrutura. A mesma encontra-se disponível completamente aparte do código de elétrons e, com isso, pode ser empregada para a identificação com o emprego de técnicas de Inteligência Computacional de outros objetos físicos com ou sem in-

⁵Para comparação, a versão em cima da estrutura do código disponível exigiu 2 semanas para implementação, enquanto a construção da infraestrutura completa precisou de ao menos 3 meses para obter a primeira versão, não contando com o esforço para manutenção e atualização.

formação anelada de calorimetria. Exemplos de objetos que poderiam ser explorados são táons e fótons.

No que tange à reconstrução da informação anelada de calorimetria, atualmente, existem somente essas duas diferenças entre o algoritmo disponível implementado e aquele disponível no SF_{on} :

1. O centro para construção dos anéis no SR_{off} (*offline*) é o baricentro de energia do *cluster* do elétron EM em todas as camadas, enquanto o algoritmo no SF_{on} utiliza como centro a célula mais energética para as camadas EM e o centro da EM2 para as camadas HAD;
2. A energia das células são calibradas pelo SR_{off} , enquanto a energia crua é empregada no SF_{on} .

Todavia, o ambiente de reconstrução *offline* exige uma revisão do método para atender o novo cenário de aplicação, onde há maior capacidade de processamento e outras representações de informação disponíveis. Considerações sobre os aprimoramentos e os desdobramentos envolvidos estão no Capítulo 9.

Por isso, durante o desenvolvimento levou-se em conta a possibilidade de explorar outras técnicas como o desenvolvimento com DL e aplicação de pré-processamentos estatísticos.

Finalmente, esse ambiente também foi interligado, como parte deste desenvolvimento, com a infraestrutura de análise de elétrons o que possibilita a obtenção das eficiências do método para comparação das nuances da proposta com as referências disponíveis.

7.3.4.1 Otimização para Produção Oficial de Dados

Restrições técnicas precisaram ser superadas para a utilização da infraestrutura no processamento de dados do ATLAS. Os dados no formato de análise, resultado final dos reprocessamentos, exigem uma quantidade razoável de recursos para armazenamento, ocupando ordens na mesma escala dos sinais crus provenientes do ATLAS (em 2017: 5,5 PB para os dados crus e 1,1 PB para a informação de análise). Por isso, o ATLAS tem realizado um grande esforço no sentido de reduzir o tamanho dos arquivos, especialmente aqueles dedicados a análise (xAOD). Igualmente, o processamento dessa grande massa de dados exige uma grande quantidade de recursos, onde o processamento conjunto de elétrons e fótons é o segundo item de maior relevância neste aspecto, sendo alvo de otimizações para redução de processamento. Assim, a adição da infraestrutura para o processamento da informação anelada para elétrons apontava na direção contrária dos esforços do ATLAS, onde os seguintes

obstáculos deviam ser considerados para que a proposta fosse algo palpável para o emprego na produção oficial:

- reduzir o impacto em tamanho de disco no formato de análise (xAOD);
- tornar o impacto em exigência computacional do algoritmo praticamente nulo, seja em termos de processamento quanto em consumo de memória transiente.

A otimização foi feita com foco no conhecimento especialista obtido. Os estudos e ajuste de modelos no contexto ocorrem através da técnica T&P (Seção 5.6), onde elétrons compatíveis com o bóson Z cobrem boa parte do alcance cinemático envolvido nas análises ($E_T > 15$ GeV). Esse alcance cinemático inclui todas as cadeias primárias de elétrons sem a aplicação de fatores de pré-escala do ATLAS e, assim, também foi considerado para o ajuste das primeiras versões do *online* do Algoritmo *NeuralRinger* ajustadas em dados de colisão durante a mesma época, assim já havia conhecimento adquirido por parte do grupo no emprego desta técnica. Além disso, boa parte dos falsos candidatos à elétrons encontram-se em energias inferiores aos estudos nessa região cinemática e são a maior parte da informação gerada para análise quando desconsiderando este aspecto. De base dessas considerações, decidiu-se realizar o processamento da informação anelada de calorimetria apenas para candidatos a elétrons com $E_T > 14$ GeV, o que garante a energia mínima para o método T&P baseado na massa do bóson Z de 15 GeV. A ideia de empregar a margem de 1 GeV deu-se para garantir que alterações na calibração de energia não impactassem na faixa contendo anéis. Com isso, há uma grande redução no impacto de disco do algoritmo e de processamento. Para maximizar ainda mais esse impacto, retirou-se o processamento do classificador, mantendo apenas a geração da informação anelada.

Assim, para a execução em 1.000 eventos em dados de colisão da *run* 311481⁶ obteve-se um aumento médio mínimo no tamanho da xAOD (0,07 %) e o impacto de processamento para a etapa mais pesada computacionalmente da reconstrução (*RAWtoESD*) um acréscimo médio de 0,46 % de processamento. O algoritmo exigiu um acréscimo de consumo de memória em pico de 5 MB e em média de 6 MB. Uma vez que o impacto é mínimo, o ATLAS adicionou o algoritmo de informação anelada de calorimetria para execução e estará inclusa nos próximos reprocessamentos de dados realizados. Isso abrange dados de colisão e de simulação, de forma que isso possibilitará a reconstrução da massa do bóson Z , com a extração dos fatores de escala (do inglês *scale factors*) entre a performance em simulação/colisão, com um mapeamento da incerteza sistemática/estatística da estratégia de identificação de elétrons com o mesmo ferramental do ATLAS, o que validará os resultados.

⁶Empregando a *AMI-tag* f641, adicionada do algoritmo *Ringer*, no lb0140 dessa *run*.

7.4 Infraestrutura para Ajuste de Modelos

Devido à grande massa de dados, mesmo treinamentos de MLP com uma única camada escondida e com algoritmo de convergência de primeira ordem (Apêndice C) decorre na necessidade de forte poder computacional. Para isso, a infraestrutura de ajuste de modelos [243] permite o processamento distribuído em duas grandes fontes de poder computacional⁷: a LCG [172] e o supercomputador Lobo Carneiro [244]. Flexibilidade e velocidade costumam ser aspectos ortogonais em computação, sendo impossível obter o caso ótimo em ambas. Por isso, a infraestrutura é integrada com dois *backends*:

- *FastNet* [245]: infraestrutura implementada em C++ para treinamento de rede neural desenvolvido por [198] e que permite, assim, alta velocidade apenas dependendo na presença de um compilador. Possibilita atualmente o treinamento via RPROP [224] e gradiente descendente [215]. Alterou-se o código de forma adaptá-lo para operação na LCG, realizando sua configuração em *python*. Essa interligação entre *python* e C++ foi otimizada para minimizar o consumo de memória, item importante para possibilitar o uso de todo o potencial de processamento da LCG. Adicionou-se a capacidade de congelamento de sinapses para possibilitar o ajuste de redes de fusão. Outro critério adicionado foi a parada múltipla (critério multiparada) por P_D , P_F e SP_{\max} , onde só se encerra o treinamento quando os três critérios atingiram um número máximo de falhas no conjunto de validação, retornando em um único treinamento três redes neurais, conforme:
 - P_D : época da rede com o menor P_F quando operando com o mesmo P_D de uma dada referência;
 - P_F : idem rede P_D , porém com a lógica invertida;
 - SP_{\max} : época em que o conjunto de teste obteve o ponto de SP_{\max} . O índice é calculado, quando maximizado, fornece um equilíbrio entre as taxas de detecção das duas classes dado sua componente de média geométrica enquanto maximiza as mesmas por via de sua componente de média aritmética.
- *keras* [246]: atualmente empregada para o aprendizado de máquinas do ATLAS, essa infraestrutura permite a aplicação de uma larga gama de funções de treinamento, incluindo aprendizado profundo (*deep learning*), e critérios para avaliação de performance durante o treinamento. No momento, esse *backend* está disponível para utilização somente no supercomputador Lobo

⁷Dado que o usuário tenha acesso as mesmas.

Carneiro. Sua performance de processamento é bastante inferior a do *Fast-Net* (ao menos 20 vezes mais exigente computacionalmente para algoritmos de convergência de primeira ordem), porém, graças ao Lobo Carneiro, há disponibilidade de poder computacional para usufruir de sua flexibilidade.

Tabela 7.3: Parâmetros empregados para ajuste, escolha de topologia do modelo e escolha de modelo para operação no *ensemble* do *NeuralRinger* ou para a extração das redes especialistas em informação anelada para atuação Sistema de Reconstrução (*Offline*) de Física do ATLAS (SR_{off}). Exceto quanto especificado, os parâmetros empregados foram os padrões empregados nas técnicas empregadas.

Critério	Valor
Composição das <i>ensemble</i>	
Modelo	MLP totalmente conectada, 1 camada escondida ($\tanh(\cdot)$) e 1 neurônio na camada escondida ($\tanh(\cdot)$)
Regiões definindo a Estatística de Ajuste	$ \eta : 0 \rightarrow 0,6 \rightarrow 0,8 \rightarrow 1,15 \rightarrow 1,37 \rightarrow 1,52 \rightarrow 1,81 \rightarrow 2,01 \rightarrow 2,37 \rightarrow 2,47$ (idem grade <i>LLH</i> , Tabela 5.2) E_T [GeV] : $15 \rightarrow 20 \rightarrow 30 \rightarrow 40 \rightarrow \infty$ (idem grade <i>LLH</i> , Tabela 5.3)
Extrapolação de Atuação	Observações em $2,47 < \eta < 2,5$ empregam rede treinada para $2,37 < \eta < 2,47$ com a sua região cinemática equivalente. Observações fora da região cinemática são rejeitadas.
Treinamento das Redes	
Núcleo para o Ajuste	<i>FastNet</i>
Dados Empregados para Ajuste	Simulação contendo colisões $p - p$ (25 ns) à 13 TeV no centro de massa com as condições observadas em 2016 (<i>mc16a</i>)
Representação da Informação	Descrição em anéis de energia
Normalização	Divisão pelo absoluto da soma em energia dos anéis
Função de Custo para Ajuste	MSE
Método de Retropropagação	RPROP (com parâmetro $\eta^+ = 1,1$)
Alvo de Elétron/Ruído Físico	+1/-1
Tamanho da Batelada	Amostragem da menor classe
Limite de Épocas de Treinamento	∞
Critério de Parada	SP_{max}
Referência de P_D e P_F	<i>LLH</i> e <i>CutBased</i>
Neurônios na Camada Escondida	5–10
Inicializações	100
Método de Validação Cruzada	<i>jack-knife</i>
Número de Subconjuntos	10 (9 compondo a base de treino, 1 para validação/teste)
Método de Divisão em Subconjuntos	Divisão em regiões uniformes sem a aplicação de permutação aleatória
Extração da ROC	1.000 pontos uniformemente espaçados entre $[-12, 12]$ no espaço sem aplicação da função de ativação
Avaliação do Treino e Escolha de Modelo para Operação	
Escolha Ponto de Operação	Ponto da ROC mais próximo a referência (P_D ou P_F)
Escolha da Melhor Inicialização	Menor P_F ou maior P_D (conjunto de teste) dentro de uma flutuação de até $\epsilon = 0,2$ % em relação ao ponto de atuação da referência (P_D ou P_F)
Escolha da Topologia do Modelo	Escolha da topologia com a melhor eficiência de operação (ver texto)
Escolha da Modelo de Operação	Menor P_F ou maior P_D (toda estatística) dentro de uma flutuação de até $0,2$ % em relação à referência (P_D ou P_F)
Tomada de Decisão	
Modelo para o Ajuste do Limiar	Ajuste realizado em cada sorteio no conjunto de treino e avaliado no teste
Dados Empregados para Ajuste	Idem ajuste do modelo
Método de Validação Cruzada	Idem ajuste do modelo
Escolha do Limiar para Operação	Calculado novamente em toda massa de dados
Regiões definindo a Estatística de Ajuste	$ \eta : 0 \rightarrow 0,6 \rightarrow 0,8 \rightarrow 1,15 \rightarrow 1,37 \rightarrow 1,52 \rightarrow 1,81 \rightarrow 2,01 \rightarrow 2,37 \rightarrow 2,47$ (idem grade <i>LLH</i> , Tabela 5.2) E_T [GeV] : $15 \rightarrow 20 \rightarrow 25 \rightarrow 30 \rightarrow 35 \rightarrow 40 \rightarrow 45 \rightarrow 50 \rightarrow \infty$ (idem grade, Tabela 5.3)
Extrapolação de Atuação	Conforme definido para os modelos neurais
Estratégia para Tomada de Decisão	Ajuste linear em função de <i>nvtx</i> até 35 vértices
Método de Ajuste	Ajuste linear empregando observações na região $nvtx \in [0, 30]$
Alvo do Ajuste	Manutenção do P_D (ou do ponto de P_D para o P_F ajustado) da operação da referência

Tabela 7.4: Parâmetros empregados para ajuste, escolha de topologia do modelo e escolha de modelo para operação no *ensemble* neural ou para a extração das redes especialistas em grandezas de traço e fusão especialista calorímetro-traço para atuação Sistema de Reconstrução (*Offline*) de Física do ATLAS (SR_{off}). Os demais critérios e valores foram os mesmos que o ajuste para a informação anelada.

Critério	Valor
Treinamento das Redes	
Representação da Informação	Grandezas físicas oriundas do ID e do casamento especialista de posição com a extrapolação do traço para o calorímetro e a posição do baricentro de energia
Normalização	Escala das grandezas por fatores
$\Delta\eta_1$	0,05
$\Delta\phi_2$	0,05
$\Delta p/p$	1
d_0/σ_{d_0}	6
d_0	0,2
TRT_PID	1

Tabela 7.5: Parâmetros empregados para ajuste, escolha de topologia do modelo e escolha de modelo para operação em um *ensemble* neural ou para a extração das redes especialistas em grandezas de calorimetria para atuação Sistema de Reconstrução (*Offline*) de Física do ATLAS (SR_{off}). Os demais critérios e valores foram os mesmos que o ajuste para a informação anelada.

Critério	Valor
Treinamento das Redes	
Representação da Informação	Grandezas físicas oriundas do Sistema de Calorimetria
Normalização	Escala das grandezas por fatores
R_{had}	0,1
R_η	1
E_{ratio}	1
R_ϕ	1
$w_{\eta 2}$	0,02
f_1	0,6
f_3	0,04

Tabela 7.6: Parâmetros ou critérios envolvidos especificamente para o treinamento de modelos neurais para fusão de informação para comporem o *ensemble*.

Critério	Valor
Treinamento das Redes	
Representação da Informação	Fusão de representação em anéis de energia com grandezas físicas oriundas do ID e casamento especialista de posição com a extrapolação do traço para o calorímetro e a posição do baricentro de energia
Normalização	Mantém normalização para cada modalidade
Estratégia de Inicialização	Cada sorteio recebe rede a melhor inicialização da rede especialista no sorteio equivalente para a topologia escolhida
Ajuste fino das sinapses especialistas	Sim

Capítulo 8

Resultados

Apresentam-se os resultados da proposta deste trabalho, a avaliação do *NeuralRinger* para atuação na seleção *offline* de elétrons (Seção 8.4). A avaliação dos impactos da atuação do *NeuralRinger* na filtragem *online* de elétrons nas amostras coletadas e possíveis desdobramentos na atuação da referência para as análises físicas, realizadas em *offline*, encontram-se na Seção 8.3. A análise do impacto foi um ponto essencial para possibilitar a atuação do *NeuralRinger* na filtragem *online*, todavia aqui se reportam somente os resultados finais, com toda a estatística disponível para colisões $p - p$ à 13 TeV ocorridas em 2013. Como parte dos requisitos de ambas análises, cabe compreender a atuação da cadeia contendo o *NeuralRinger* durante o ano de 2017 (Seção 8.2). A nomenclatura empregada na discussão dos resultados está descrita na Seção 8.1.

8.1 Nomenclatura Empregada

Na parte referente à avaliação dos resultados no SF_{on} , referir-se-á à cadeia empregando o *ensemble* neural, a partir de informação anelada de calorimetria para tomada de decisão no FastCalo, como cadeia *ringer* e sua versão empregando cortes lineares em três grandezas físicas (R_{had} , R_{η} e E_{ratio}) como cadeia *noringer*. Mais detalhes referentes ao menu de 2017 na Figura 5.11.

As análises realizadas para o *offline* empregaram as representações descritas na Seção 7.3.1. Essa informação é utilizada por classificadores, conforme os seguintes grupos:

- *NN*: *ensemble* de 36 redes neurais composto por unidade única MLP com uma camada escondida;
- *ExNN*: *ensemble* de 36 redes neurais composto por unidades MLP especialistas com uma única camada escondida, alimentando unidades de rede de fusão;

- *LLH*: método de verossimilhança de referência.

A configuração de atuação do classificador é determinada segundo a nomenclatura *classificador(infos)*, onde *infos* são as representações da informação apresentadas por somas das informações alimentadas para um determinado tipo de classificador. Por exemplo: *ExNN(rings + track⁺)* refere-se às redes especialistas tomando decisão em representação com os anéis de energia e as grandezas físicas extraídas do ID e do casamento especialista em posição.

8.2 Identificação *Online* de Elétrons

Os resultados para a atuação do *NeuralRinger online* nas cadeias destinadas à filtragem de elétrons com $E_T > 15$ GeV durante o ano de 2017 dividem-se em dois grupos, ambos avaliados a partir de reprocessamentos das decisões do SF_{on} ocorridas em *offline*. Em um primeiro estágio, o método encontrava-se em desenvolvimento e os resultados foram fornecidos por reprocessamento. Os reprocessamentos decorrem em base de uma *run* de referência para avaliações oficiais de eficiência e de comissionamento de novas versões de código para atuação no HLT durante a coleta (*online*). Essa *run* foi removida do ajuste realizado no limiar para evitar a influência dos dados empregados no ajuste na análise de eficiência. Os dados de simulação empregados nas análises são os mesmos que aqueles utilizados para o ajuste dos modelos. Conforme mencionado na Subseção 7.3.3, acredita-se que eventuais influências devido à especialização sejam mínimas, uma vez que há cuidado da área de aplicação em obter uma estatística suficiente para caracterizar o processo físico com precisão.

A Tabela 8.1 mostra a superioridade da cadeia *ringer* em relação à versão que atuou em operação durante 2016. A cadeia proposta possibilita uma redução de $\sim 3\times$ (valores centrais) em relação à cadeia original em dados de colisão (*enhanced bias*) e $\sim 2\times$ em simulação no que tange a probabilidade de falso alarme, mantendo a taxa de detecção desejadas na saída do HLT. Em ambos casos, a eficiência no nível HLT para MC permanece praticamente inalterada, enquanto em colisão há uma queda de 0,4 (valor central) em eficiência. Posteriormente, observou-se que essa queda se dava devido à alteração da saída neural para a região $2,37 < |\eta| < 2,47$, que se deve à ausência das células com alta granularidade (tiras) na EM1. Assim, esse efeito foi posteriormente recuperado ao introduzir um limiar dedicado à essa região.

Esses resultados são ainda mais relevantes quando atentando para a necessidade de processamento na fazenda em cada nível. Após o cenário de atualizações na cadeia de elétrons descrito na Subseção 5.7.1, a diferença de latência dos resultados para a cadeia primária de elétrons de menor energia (26 GeV) implicava uma redução pela

Tabela 8.1: Evolução da eficiência para os estágios de seleção do HLT para as cadeias *ringer* e *noringer* em previsões *offline* para o menu de 2016 em dados de colisão e simulação. O denominador é comum para todos os estágios de seleção, composto pelo número de RoI com casamento a um elétron reconstruído pelo *offline*. Em ambos os casos, as cadeias dedicam-se a coleta de elétrons com $E_T > 28$ GeV e aplicam o critério *medium*. Para as avaliações de eficiência de detecção, empregou-se a seleção de eventos pelo método T&P. Especificamente para as amostras provenientes para a coleta de ruído em dados de colisão, as mesmas são provenientes do fluxo de dados *enhanced bias*, contendo cadeias de suporte dedicadas à coleta de ruído físico. A fim de eliminar possível contaminação de sinal nesses dados, aplicou-se a seleção de falsos candidatos ao inverter a lógica T&P (evento não participa como *probe* caso forme um par) ao alterar o critério aplicado no *tag* para *medium* e a lógica inversa de seleção para o critério *vloose* nos . A seleção de eventos de ruído em simulação aplicou a remoção de decaimentos em elétrons isolados.

Cadeia	FastCalo [%]	FastElectron [%]	HLTCalo [%]	HLT [%]
<i>Z</i> → <i>ee</i> T&P <i>run 311244</i> (detecção)				
cadeia <i>noringer</i>	96,77 ± 0,03	96,70 ± 0,03	95,84 ± 0,05	89,74 ± 0,07
cadeia <i>ringer</i>	96,23 ± 0,03	96,16 ± 0,03	95,31 ± 0,04	89,32 ± 0,08
<i>enhanced bias run 311244</i> (falso alarme)				
cadeia <i>noringer</i>	16,03 ± 0,02	15,70 ± 0,02	12,90 ± 0,02	0,57 ± 0,00
cadeia <i>ringer</i>	5,41 ± 0,02	5,12 ± 0,02	3,40 ± 0,01	0,54 ± 0,00
<i>Z</i> → <i>ee</i> T&P <i>MC15</i> (detecção)				
cadeia <i>noringer</i>	96,10 ± 0,01	96,09 ± 0,01	95,37 ± 0,01	89,06 ± 0,01
cadeia <i>ringer</i>	96,11 ± 0,01	96,09 ± 0,01	95,38 ± 0,01	89,12 ± 0,01
<i>JF17 MC15</i> (falso alarme)				
cadeia <i>noringer</i>	8,04 ± 0,01	7,93 ± 0,01	6,54 ± 0,00	0,20 ± 0,00
cadeia <i>ringer</i>	3,69 ± 0,01	3,59 ± 0,01	2,55 ± 0,00	0,19 ± 0,00

metade (200→100 ms/evento) quando empregando o *NeuralRinger*. Para a cadeia primária de menor energia para a coleta de elétrons duplos, a queda na taxa foi de 9700 Hz para 4900 Hz.

Esses resultados motivaram o comissionamento da cadeia no menu de elétrons a fim de substituir a cadeia *noringer* pela versão empregado o *ensemble* neural. A operação durante o comissionamento mostrou-se conforme as previsões nos reprocessamentos de forma que o algoritmo entrou em operação após a TS1¹ que se encerrou na primeira semana de julho de 2017. Os resultados mencionados adiante referem-se à atuação do algoritmo conforme avaliado pela monitoração ocorria após a coleta dos dados (Subseção 7.1.2).

A mudança após a TS1 para o emprego das cadeias do *Ringer* ocorreu de forma que não houve impacto significativo na eficiência (Figura 8.1). As diferenças de

¹Parada técnica 1, em inglês.

eficiência na região $2,37 < |\eta| < 2,47$ ocorreram devido à mudança de eficiência do *L1Calo*. A redução na taxa de falso alarme (Figura 8.2), por sua vez, deu-se conforme o esperado.

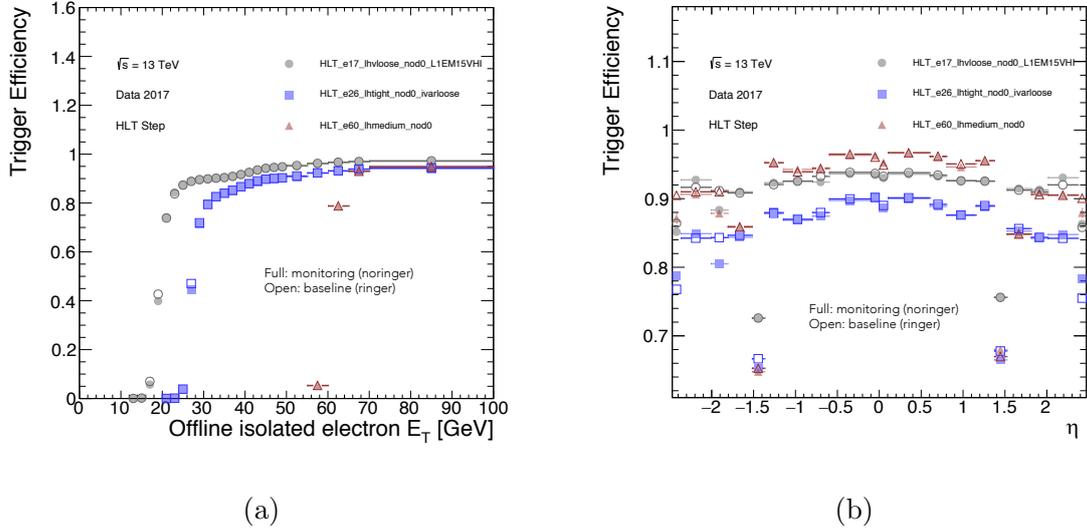


Figura 8.1: Comparação da taxa de detecção na saída do HLT para cadeias primárias (cinza: 17 GeV, azul: 26 GeV, vermelho: 60 GeV) empregando o método *CutBased* (marcadores fechados) antes da TS1 com o método *NeuralRinger* (marcadores abertos) após a TS1 para dados coletados até outubro de 2017.

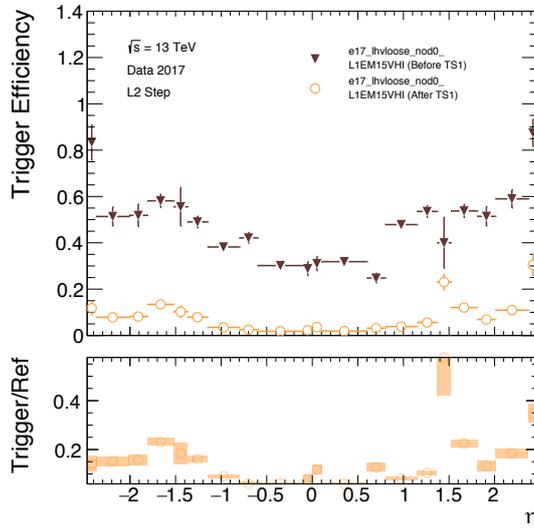


Figura 8.2: Comparação da taxa de falso alarme no estágio FastCalo para uma cadeia primária com emprego de pré-escala, utilizando o método *CutBased* (marcadores fechados), antes da TS1 com o método *NeuralRinger* (marcadores abertos) após a TS1 para dados coletados até outubro de 2017.

Para a análise de quadrante considerando a grandeza R_η (Figura 8.3a), observa-se

que há grande concordância entre as duas cadeias. Eventuais discordâncias ocorrem dentro da região em que há grande concordância entre os algoritmos, com uma ligeira tendência à coleta mais próxima a unidade pela cadeia *ringer*. Esse valor é o ideal para essa grandeza, região representada por pouco vazamento lateral de energia na EM2. Comportamento similar ocorre para a grandeza R_{had} (Figura 8.3b) que representa o vazamento hadrônico (vazamento de energia para o HCAL), mostrando que a cadeia *ringer*, apesar de não obter essas grandezas apresentadas durante o ajuste dos modelos neurais, conseguiu capturar a estrutura do desenvolvimento do chuveiro a partir dos anéis e mostra uma ligeira tendência a coletar elétrons de melhor qualidade, conforme a descrição nessas duas grandezas físicas, quando avaliado usando como referência o *offline*.

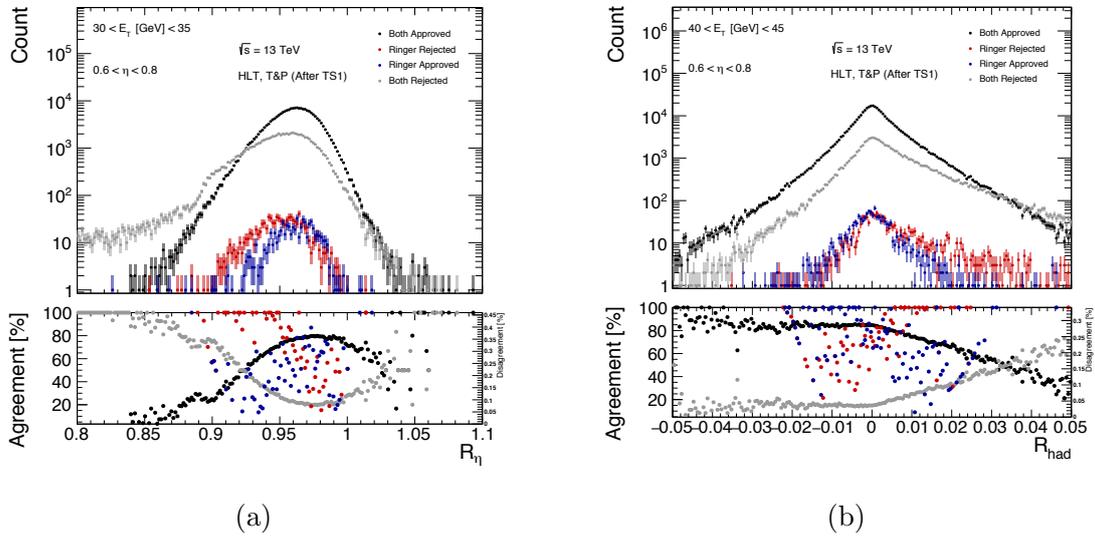


Figura 8.3: Análise de quadrante em para as grandezas físicas R_η (a) e R_{had} (b) nas as cadeias *ringer* e *noringer* atuando no critério *medium* e avaliada em dados coletados até outubro de 2017. Na figura inferior, está a razão entre os casos mutuamente exclusivos pelo total de observações. Há duas escalas, a da esquerda refere-se à concordância entre as cadeias (caso I e IV) e a da direita refere-se à discordância (caso II e III). A descrição dos casos encontra-se na Tabela 7.2.

8.3 Análise de Impacto na Coleta de Elétrons

Inicia-se esta seção considerando a eficiência (Figura 8.4) das cadeias dedicadas ao estudo de impacto (Subseção 7.2). Observa-se que apesar das eficiências serem bastante próximas, há pequena diferença das mesmas que está sujeita à área de região do espaço de fase. Com isso, as cadeias possuem diferenças na quantidade de amostras coletadas para cada região do espaço de fase.

Tendo isso em mente, aborda-se a avaliação do impacto considerando o caso

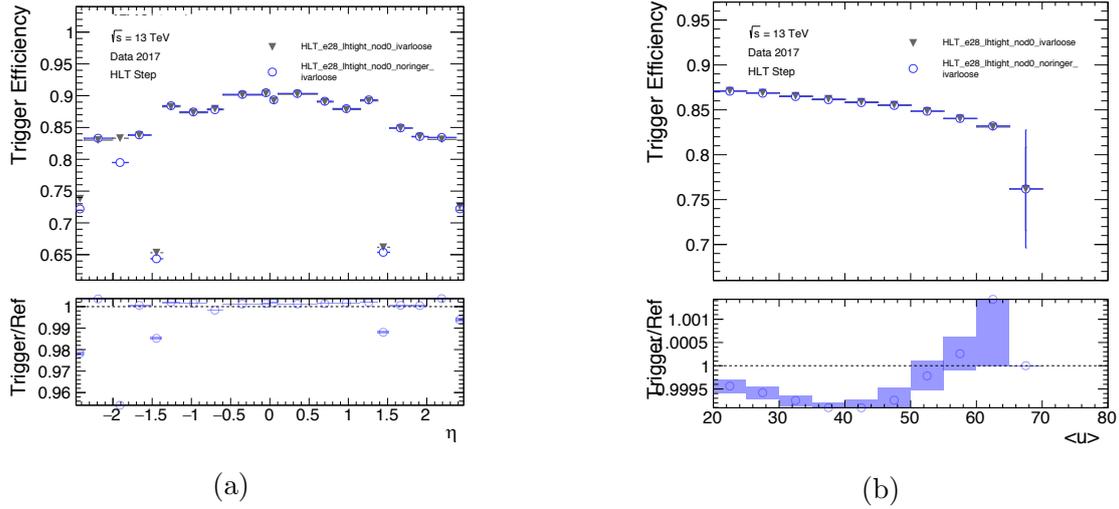


Figura 8.4: Eficiências (parte superior) e sua razão (parte inferior) na saída do HLT em função de η (a) e $\langle \mu \rangle$ (b) para cadeias de seleção *ringer* e *noringer* duplicadas para a análise de impacto durante toda a tomada de dados para colisões $p - p$ à 13 TeV.

I+II (*ringer*) em relação à referência, caso I+II (*noringer*) conforme especificados na Tabela 7.2. Na Figura 8.5, observa-se que R_η contém sistematicamente mais amostras em relação ao modelo no pico da distribuição, comportamento que se mantém no decaimento da distribuição à direita do pico. Por sua vez o oposto é observado para o decaimento à esquerda do pico. Cabe notar que esse desvio é de ordem muito inferior ao resíduo devido às diferenças de eficiência, que, por sua vez, já é pequeno ($< 1\sigma$) em ~ 400.000 amostras coletadas no pico. Essa análise qualitativa sugere haver uma distorção, ainda que mínima, no mesmo sentido que aquele observado na análise de quadrante. Quando observando o R_{had} , há uma massa de resíduos no pico com maior contagem para a cadeia *ringer* porém o mesmo não é tão claro quanto para R_η . Porém, observa-se uma ligeira tendência à coletar menos amostras na região em que o valor descrito para R_{had} caracteriza padrões mais frequentes de ruído físico.

A análise de impacto mostra que apesar de alterar a estratégia, um resultado similar é observado àquele na análise de quadrante. Cabe ressaltar aqui a diferença: na análise de impacto, o par coletado para compor uma observação de T&P precisa ser proveniente de um evento de cruzamento de pacotes $p - p$ que tenha sido aceito pela cadeia de seu respectivo histograma, enquanto a análise de quadrante aplica a seleção no *probe* que foi selecionado pelo mesmo critério da cadeia em análise. No caso do impacto, o *probe* está sujeito à seleção pelo critério *vloose*. Assim, a diferença entre as duas abordagens quando considerando os critérios aqui estão resumidos na Tabela 8.2. Percebe-se que os casos são diferentes e levam a resultados diferentes.

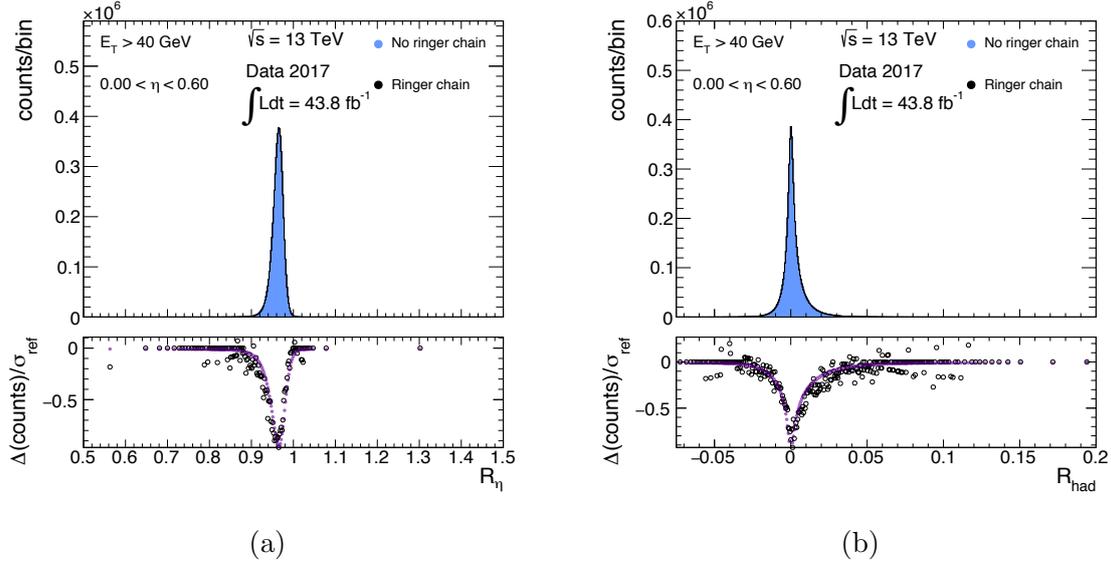


Figura 8.5: Análise de impacto no *offline* ao alterar a sistemática do método T&P para as grandezas físicas R_η (a) e R_{had} (b) para as cadeias *ringer* e *noringer* em dados coletados para o ano de 2017. Na parte superior, o histograma conforme reconstruído pelo *offline* para a geração do KDE. Na parte inferior, a análise de resíduo empregando como estatística χ^s (preto) e o modelo esperado em caso de não distorção em relação à referência (*noringer*) conforme computado por (7.3).

Todavia, como as distorções são muito pequenas, esse efeito não fica evidente.

Tabela 8.2: Diferenças de critérios exigidos no T&P entre as análises de impacto e de quadrante avaliados neste trabalho.

	Impacto	Quadrante
evento $p - p$	caso mutuamente exclusivo ou cadeia em análise	cadeias primárias (<i>ringer</i>) sem a aplicação de pré-escala
<i>tag</i>	caso mutuamente exclusivo ou cadeia em análise	cadeias primárias (<i>ringer</i>) sem a aplicação de pré-escala
<i>probe</i>	<i>LLH vloose</i>	caso mutuamente exclusivo no critério <i>medium</i> e <i>LLH medium</i>

Uma maneira melhor de acessar a informação que pode causar distorção foi discutida na definição do método, através da análise somente dos casos II (*ringer* aceita e *noringer* rejeita) e III (lógica inversa ao caso II). A análise de impacto através de divergência KL, conforme disponível na Figura 8.6, permite observar com maior facilidade o impacto. Observa-se, o fenômeno de distorção com muito maior facilidade ao comparar os dois histogramas. Uma análise proporcional à frequência de observação pode ser realizada ao comparar os marcadores circulares com os triângulos na mesma região. Observa-se com clareza que a cadeia *ringer* coleta sistematicamente mais amostras quando R_η torna-se superior a um valor

determinado valor ($R_\eta = 0,957$).

Para a grandeza R_{had} (Figura 8.7), há um maior acúmulo de amostras em torno do valor ideal para essa variável (0) por parte da cadeia *ringer*. Por sua vez, há oscilação estatística em torno do ponto em que ocorre mudança para o acúmulo, não sendo possível indicar com precisão o *bin* onde ocorre a mudança para que a cadeia *ringer* passe a coletar menos amostras. Todavia, é possível notar que há uma menor coleta por parte da cadeia *ringer* na cauda à direita da na figura. O teste em divergência considera que há distorção no perfil dos dois histogramas observados. Grande parte das outras grandezas apresentam o mesmo comportamento, porém, em alguns casos de regiões com menor estatística (como as regiões de menor cinemática, $15 < E_T \text{ GeV} < 20$), não há capacidade de acordo com o teste em divergência-KL para afirmar sobre a existência de distorção na distribuição.

Assim, pode-se concluir da análise de impacto que as distorções são de ordem bastante inferior aos resíduos oriundos de pequenas diferenças em eficiência ($< 1\%$) para a região do espaço de fase avaliada. As distorções nas grandezas mostram que a cadeia *ringer* possibilitou a coleta de amostras com melhor qualidade de acordo com a descrição das grandezas físicas, apesar do método não ter sido ajustado em posse dessas grandezas. Isso mostra o potencial da abordagem empregando a informação anelada e *ensemble* neural em capturar a essência do desenvolvimento do chuveiro.

8.4 Identificação *Offline* de Elétrons

Os resultados de eficiência para o critério *vloose* por espaço de fase mostram (Tabela 8.3) pontos interessantes a serem assinalados. A eficiência para a composição $NN(\text{rings})$ mostra melhoria na eficiência conforme se aumenta a energia. Isso se deve ao melhor desenvolvimento do chuveiro para essas regiões, onde o elétron desenvolve um chuveiro mais fino em relação ao daqueles desenvolvidos por ruído de fundo. Há, também, menor flutuação estatística (Seção 4.2) por parte do Sistema de Calorimetria para essa região. O comportamento oposto é observado para a composição $NN(\text{track}^+)$, onde há uma piora substancial na eficiência conforme o aumento em energia do objeto. O ID realiza a construção de boa parte de suas grandezas (exceto TRT_PID) em base de pontos no espaço que perdem resolução (Seção 4.1) conforme o aumento de energia. Como o traço do elétron está isolado (na ausência de empilhamento), a reconstrução do traço é mais precisa que o ruído físico, oriundo de mais de uma partícula. Com a perda de precisão, perde-se a capacidade de capturar esse padrão. A comparação dos resultados na tabela do modelo $NN(\text{rings})$ com o $NN(\text{track}^+)$ permite observar que a representação em anéis é mais discriminante para boa parte do espaço de fase, exceto pela região do *crack* ($1,37 < |\eta| < 1,52$) do Sistema de Calorimetria na região de menor energia.

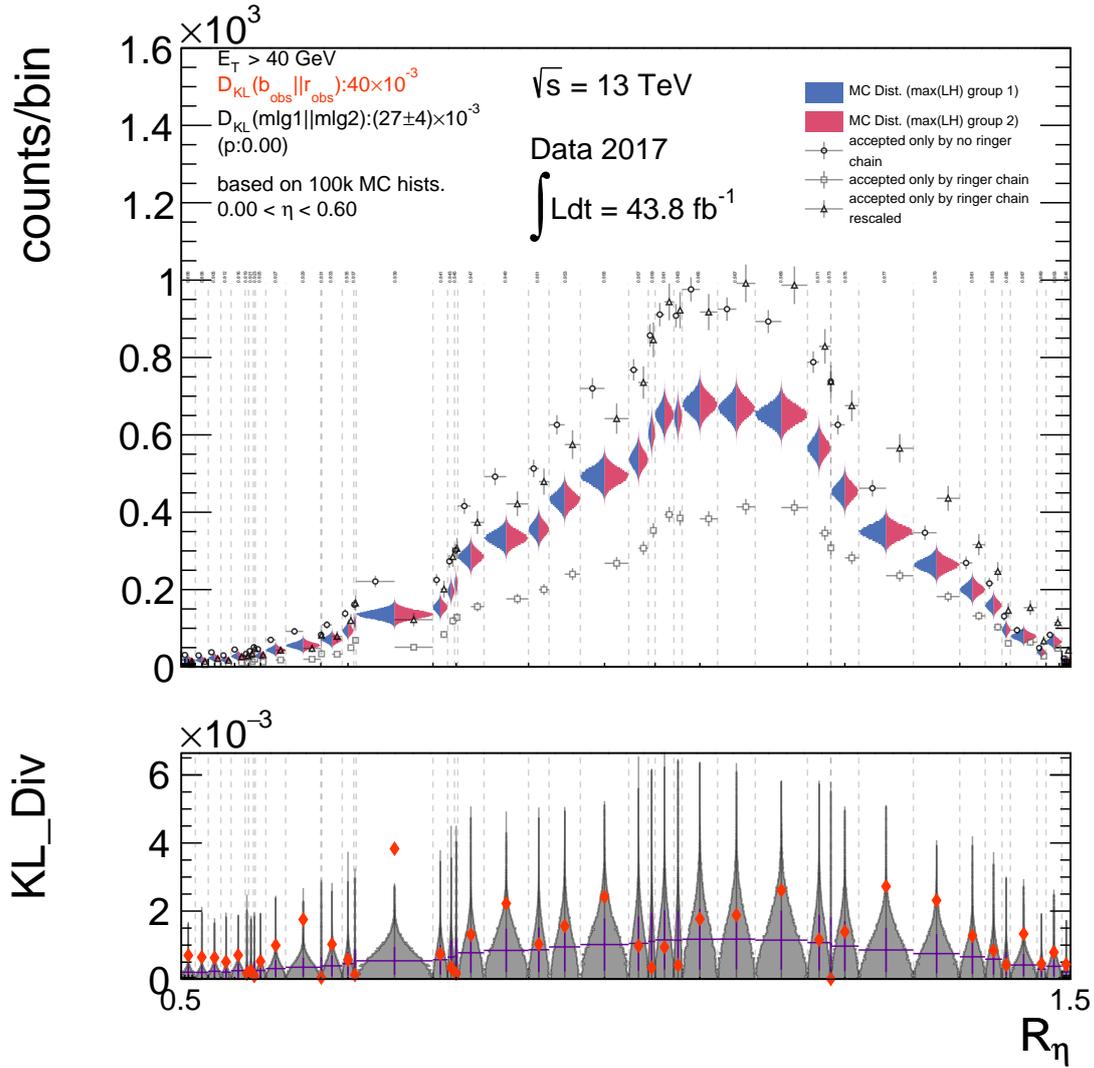


Figura 8.6: Análise de impacto para a grandeza R_η reconstruída pelo *offline* considerando as amostras coletadas durante 2017, exclusivamente pela cadeia *ringer* ou *noringer*. O eixo das abscissas sofre uma transformada discreta ponderando o tamanho do *bin* do histograma pela divergência observada. No canto superior esquerdo estão a divergência observada ($D_{KL}(b_{obs}||r_{obs})$) e a divergência média e seu desvio padrão para os dois grupos de histogramas gerados via simulação mutuamente exclusivos a partir da estimativa de máxima verossimilhança das observações. A distribuição unidimensional das simulações estão em azul e vermelho para cada *bin* do histograma. A partir dessas distribuições, computa-se a contribuição da divergência-KL na parte inferior cuja distribuição obtida está na parte inferior da figura em cinza. A divergência entre a observação de entradas exclusivas para um *bin* no histograma proveniente da cadeia *ringer* (quadrado na figura superior) e *noringer* (círculo na figura superior) geram a contribuição individual da divergência-KL, em laranja na figura inferior. Como o número de observações nos dois histogramas é diferente, na figura superior está disponível um reescalonamento do histograma *ringer* de forma que a área dos dois histogramas seja a mesma. As linhas pontilhadas indicam os limites do *bin* e os números acima do máximo do histograma o centro dos *bins*.

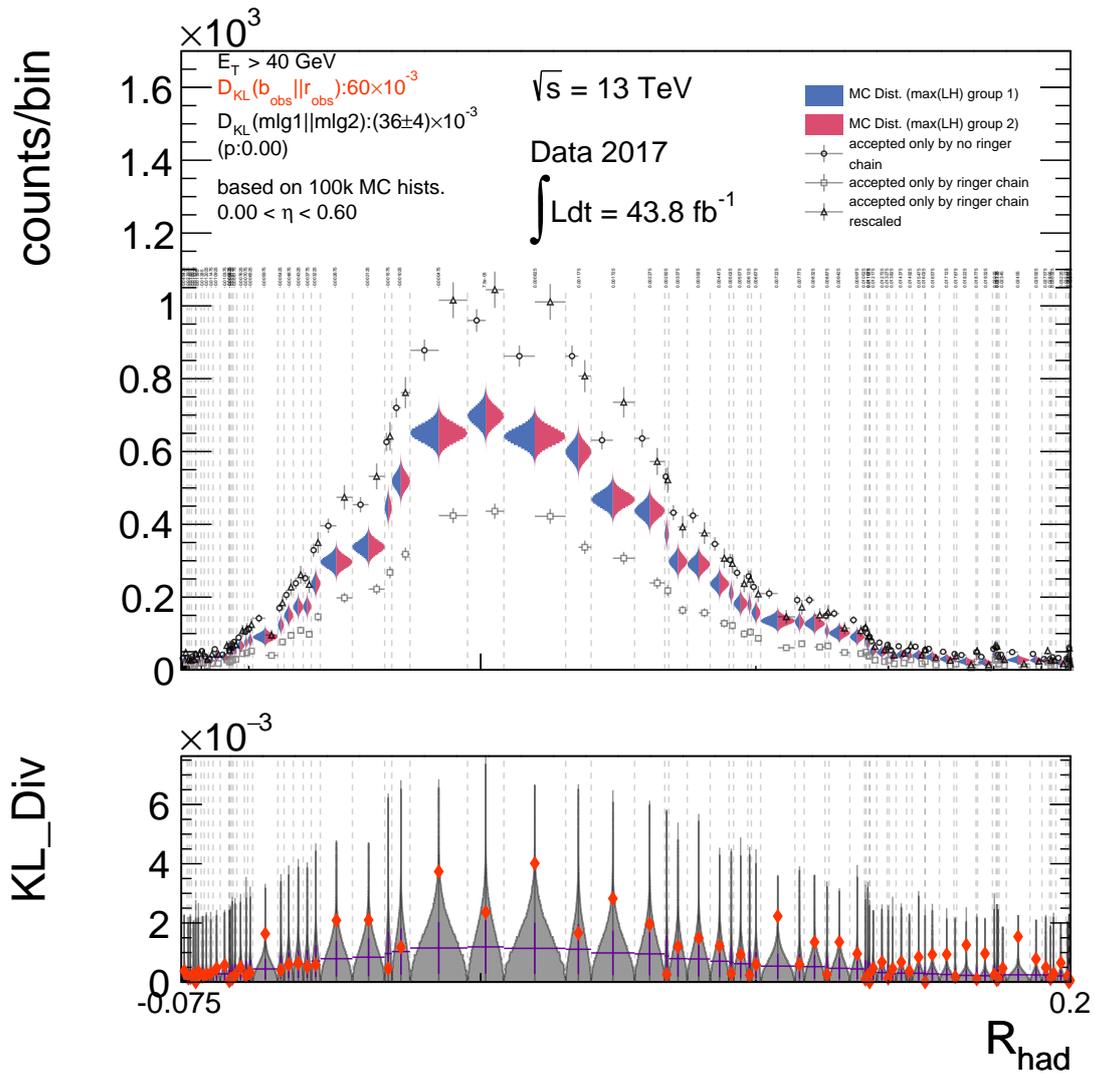


Figura 8.7: Análise de impacto para a grandeza R_{had} reconstruída pelo *offline* considerando as amostras coletadas durante 2017, exclusivamente pela cadeia *ringer* ou *noringer*.

Vale ressaltar que o *crack* não atinge o ID. Justamente pelo fato da informação de calorimetria se tornar mais relevante conforme o aumento de energia, a disparidade de eficiência do $NN(rings)$ em relação aos demais classificadores empregando fusão de informação do Sistema de Calorimetria e do ID reduz.

Tabela 8.3: Valores de eficiência (P_D , SP e P_F) para a referência (LLH) medida em toda estatística disponível e do ensemble neural operando no ponto mais próximo ao P_D (coluna ressaltada em verde) da referência e apresentado como o valor central e 1 RMS das eficiências de validação cruzada. Nomenclatura empregada disponível na Seção 8.1.

kinematic region			$15 < E_T[\text{GeV}] < 20$			$20 < E_T[\text{GeV}] < 30$			$30 < E_T[\text{GeV}] < 40$			$E_T[\text{GeV}] > 40$		
Det. Region	Method	Type	P_D [%]	SP [%]	P_F [%]	P_D [%]	SP [%]	P_F [%]	P_D [%]	SP [%]	P_F [%]	P_D [%]	SP [%]	P_F [%]
$0.00 < \eta < 0.60$	$LLH(\text{shower} + \text{track}^+)$	Reference	95.41	96.81	1.78	97.20	97.77	1.66	98.41	98.25	1.91	98.74	98.51	1.72
	$NN(\text{rings})$	Cross Validation	95.31±0.04	95.04±0.02	5.24±0.04	97.14±0.01	96.12±0.02	4.88±0.05	98.37±0.01	97.09±0.02	4.18±0.03	98.71±0.01	97.78±0.01	3.16±0.02
	$NN(\text{track}^+)$	Cross Validation	95.40±0.03	94.38±0.02	6.64±0.05	97.18±0.01	94.96±0.03	7.23±0.06	98.40±0.00	94.55±0.07	9.21±0.14	98.73±0.01	93.81±0.08	10.98±0.16
	$ExNN(\text{rings} + \text{track}^+)$	Cross Validation	95.35±0.02	97.05±0.01	1.23±0.03	97.14±0.01	98.04±0.00	1.06±0.01	98.38±0.00	98.73±0.01	0.92±0.03	98.72±0.00	98.95±0.01	0.81±0.02
	$ExNN(\text{rings} + \text{shower} + \text{track}^+)$	Cross Validation	95.34±0.03	97.09±0.02	1.14±0.02	97.15±0.01	98.08±0.01	0.98±0.02	98.38±0.01	98.76±0.01	0.87±0.03	98.72±0.00	98.99±0.01	0.74±0.03
$0.60 < \eta < 0.80$	$LLH(\text{shower} + \text{track}^+)$	Reference	95.26	96.43	2.38	96.99	97.55	1.88	98.28	98.13	2.01	98.58	98.41	1.76
	$NN(\text{rings})$	Cross Validation	95.15±0.04	95.09±0.02	4.97±0.04	96.91±0.02	96.01±0.01	4.90±0.02	98.24±0.01	97.26±0.02	3.70±0.04	98.55±0.01	97.97±0.02	2.61±0.04
	$NN(\text{track}^+)$	Cross Validation	95.21±0.06	93.55±0.04	8.09±0.11	96.98±0.01	94.41±0.04	8.14±0.09	98.27±0.01	93.80±0.09	10.57±0.18	98.58±0.01	93.08±0.18	12.25±0.36
	$ExNN(\text{rings} + \text{track}^+)$	Cross Validation	95.14±0.04	96.92±0.02	1.28±0.03	96.95±0.02	97.91±0.01	1.12±0.03	98.24±0.00	98.61±0.02	1.02±0.04	98.55±0.00	98.91±0.02	0.73±0.04
	$ExNN(\text{rings} + \text{shower} + \text{track}^+)$	Cross Validation	95.16±0.04	96.97±0.03	1.21±0.03	96.96±0.02	97.96±0.01	1.05±0.03	98.24±0.01	98.64±0.02	0.97±0.05	98.56±0.01	98.93±0.02	0.69±0.05
$0.80 < \eta < 1.15$	$LLH(\text{shower} + \text{track}^+)$	Reference	94.92	96.34	2.24	96.50	97.25	2.01	97.85	97.98	1.89	98.34	98.36	1.62
	$NN(\text{rings})$	Cross Validation	94.77±0.03	94.13±0.03	6.51±0.07	96.42±0.02	95.48±0.04	5.45±0.07	97.78±0.01	96.64±0.01	4.50±0.03	98.28±0.01	97.51±0.01	3.26±0.03
	$NN(\text{track}^+)$	Cross Validation	94.91±0.04	93.69±0.04	7.51±0.09	96.47±0.02	94.65±0.03	7.16±0.06	97.84±0.01	94.36±0.06	9.05±0.12	98.33±0.01	93.78±0.08	10.67±0.16
	$ExNN(\text{rings} + \text{track}^+)$	Cross Validation	94.83±0.02	96.69±0.02	1.43±0.04	96.44±0.01	97.67±0.01	1.09±0.01	97.80±0.01	98.35±0.01	1.10±0.03	98.30±0.01	98.71±0.02	0.88±0.04
	$ExNN(\text{rings} + \text{shower} + \text{track}^+)$	Cross Validation	94.84±0.02	96.75±0.01	1.31±0.03	96.44±0.02	97.70±0.01	1.02±0.02	97.80±0.01	98.38±0.01	1.04±0.03	98.30±0.00	98.73±0.02	0.84±0.03
$1.15 < \eta < 1.37$	$LLH(\text{shower} + \text{track}^+)$	Reference	93.41	95.65	2.09	95.72	96.88	1.96	97.41	97.66	2.10	97.98	97.95	2.07
	$NN(\text{rings})$	Cross Validation	93.24±0.04	93.97±0.04	5.30±0.05	95.59±0.02	95.26±0.02	5.08±0.05	97.31±0.01	96.47±0.03	4.37±0.07	97.89±0.01	97.37±0.03	3.14±0.06
	$NN(\text{track}^+)$	Cross Validation	93.44±0.04	93.71±0.03	6.02±0.06	95.69±0.01	94.72±0.02	6.25±0.05	97.39±0.01	94.32±0.11	8.69±0.21	97.96±0.01	93.70±0.17	10.46±0.33
	$ExNN(\text{rings} + \text{track}^+)$	Cross Validation	93.34±0.03	96.08±0.03	1.13±0.04	95.64±0.02	97.31±0.01	1.00±0.02	97.34±0.01	98.10±0.02	1.13±0.04	97.92±0.01	98.49±0.02	0.93±0.04
	$ExNN(\text{rings} + \text{shower} + \text{track}^+)$	Cross Validation	93.31±0.06	96.10±0.03	1.07±0.04	95.63±0.02	97.34±0.01	0.93±0.02	97.34±0.02	98.12±0.02	1.11±0.05	97.92±0.00	98.52±0.01	0.88±0.03
$1.37 < \eta < 1.52$	$LLH(\text{shower} + \text{track}^+)$	Reference	84.50	90.57	3.15	88.71	92.44	3.75	93.93	93.73	6.46	95.35	93.87	7.60
	$NN(\text{rings})$	Cross Validation	84.17±0.11	89.03±0.06	5.97±0.05	88.39±0.07	91.35±0.07	5.64±0.13	93.72±0.03	93.02±0.16	7.68±0.30	95.15±0.02	94.05±0.10	7.05±0.21
	$NN(\text{track}^+)$	Cross Validation	84.37±0.09	89.85±0.05	4.49±0.05	88.57±0.05	91.49±0.02	5.55±0.04	93.83±0.03	90.93±0.12	11.93±0.25	95.26±0.02	87.60±0.18	19.74±0.36
	$ExNN(\text{rings} + \text{track}^+)$	Cross Validation	84.25±0.08	91.47±0.03	1.00±0.04	88.47±0.05	93.69±0.02	0.95±0.02	93.77±0.02	95.97±0.03	1.80±0.07	95.20±0.03	96.59±0.06	2.00±0.12
	$ExNN(\text{rings} + \text{shower} + \text{track}^+)$	Cross Validation	84.20±0.12	91.49±0.07	0.91±0.03	88.48±0.05	93.73±0.03	0.87±0.02	93.76±0.03	96.12±0.03	1.50±0.05	95.19±0.02	96.80±0.04	1.57±0.08
$1.52 < \eta < 1.81$	$LLH(\text{shower} + \text{track}^+)$	Reference	92.49	95.02	2.41	94.78	96.03	2.71	95.85	96.69	2.47	96.76	97.12	2.51
	$NN(\text{rings})$	Cross Validation	92.25±0.04	92.95±0.04	6.35±0.07	94.62±0.03	94.42±0.04	5.79±0.09	95.71±0.02	95.40±0.02	4.91±0.06	96.65±0.02	96.49±0.05	3.68±0.09
	$NN(\text{track}^+)$	Cross Validation	92.33±0.03	93.33±0.03	5.67±0.05	94.73±0.02	94.27±0.05	6.18±0.10	95.81±0.01	94.23±0.07	7.33±0.15	96.73±0.01	93.18±0.08	10.30±0.17
	$ExNN(\text{rings} + \text{track}^+)$	Cross Validation	92.32±0.05	95.45±0.02	1.37±0.02	94.67±0.02	96.69±0.01	1.26±0.02	95.75±0.02	97.29±0.02	1.16±0.04	96.68±0.01	97.73±0.03	1.21±0.05
	$ExNN(\text{rings} + \text{shower} + \text{track}^+)$	Cross Validation	92.31±0.05	95.50±0.03	1.27±0.03	94.67±0.02	96.74±0.02	1.16±0.03	95.76±0.02	97.33±0.01	1.08±0.03	96.67±0.01	97.79±0.02	1.10±0.04
$1.81 < \eta < 2.01$	$LLH(\text{shower} + \text{track}^+)$	Reference	91.73	94.87	1.93	92.80	95.46	1.85	94.34	96.32	1.68	95.54	96.98	1.58
	$NN(\text{rings})$	Cross Validation	91.44±0.05	93.38±0.02	4.66±0.02	92.57±0.06	94.20±0.05	4.16±0.06	94.12±0.03	95.29±0.03	3.53±0.05	95.38±0.01	96.50±0.02	2.37±0.04
	$NN(\text{track}^+)$	Cross Validation	91.65±0.05	92.72±0.03	6.20±0.06	92.74±0.02	93.36±0.04	6.02±0.08	94.31±0.02	93.52±0.03	7.27±0.07	95.51±0.01	93.68±0.05	8.13±0.11
	$ExNN(\text{rings} + \text{track}^+)$	Cross Validation	91.54±0.07	95.04±0.04	1.40±0.02	92.66±0.03	95.76±0.02	1.09±0.03	94.22±0.03	96.61±0.02	0.97±0.04	95.42±0.02	97.30±0.01	0.80±0.03
	$ExNN(\text{rings} + \text{shower} + \text{track}^+)$	Cross Validation	91.52±0.06	95.09±0.03	1.27±0.04	92.67±0.03	95.81±0.02	0.99±0.03	94.22±0.03	96.62±0.02	0.95±0.04	95.42±0.02	97.31±0.02	0.78±0.04
$2.01 < \eta < 2.37$	$LLH(\text{shower} + \text{track}^+)$	Reference	90.78	93.97	2.78	90.85	94.28	2.22	92.21	95.18	1.80	93.95	95.95	2.02
	$NN(\text{rings})$	Cross Validation	90.50±0.06	93.09±0.02	4.29±0.03	90.57±0.05	93.43±0.03	3.67±0.08	91.93±0.05	94.48±0.03	2.93±0.02	93.67±0.04	95.66±0.02	2.33±0.01
	$NN(\text{track}^+)$	Cross Validation	90.79±0.06	90.96±0.03	8.86±0.07	90.79±0.05	92.02±0.03	6.74±0.08	92.18±0.03	92.67±0.03	6.83±0.06	93.93±0.02	93.05±0.03	7.81±0.07
	$ExNN(\text{rings} + \text{track}^+)$	Cross Validation	90.59±0.03	94.50±0.02	1.51±0.02	90.68±0.04	94.79±0.02	1.01±0.02	92.04±0.05	95.54±0.02	0.90±0.03	93.80±0.03	96.48±0.02	0.80±0.03
	$ExNN(\text{rings} + \text{shower} + \text{track}^+)$	Cross Validation	90.60±0.05	94.55±0.02	1.42±0.03	90.65±0.04	94.82±0.02	0.91±0.01	92.04±0.04	95.55±0.02	0.87±0.04	93.80±0.01	96.48±0.01	0.81±0.03
$2.37 < \eta < 2.47$	$LLH(\text{shower} + \text{track}^+)$	Reference	91.59	86.53	18.39	89.73	93.27	3.12	91.21	93.95	3.28	92.75	94.95	2.83
	$NN(\text{rings})$	Cross Validation	91.19±0.13	92.51±0.05	6.16±0.04	89.34±0.09	92.21±0.04	4.87±0.06	90.83±0.08	93.55±0.05	3.69±0.06	92.44±0.04	94.81±0.03	2.78±0.04
	$NN(\text{track}^+)$	Cross Validation	91.56±0.08	89.46±0.06	12.62±0.15	89.63±0.05	90.67±0.04	8.28±0.09	91.14±0.06	91.08±0.07	8.98±0.14	92.67±0.04	91.66±0.09	9.34±0.19
	$ExNN(\text{rings} + \text{track}^+)$	Cross Validation	91.32±0.09	94.45±0.05	2.37±0.06	89.46±0.06	94.05±0.04	1.23±0.02	90.97±0.05	94.82±0.03	1.26±0.05	92.53±0.03	95.79±0.04	0.89±0.05
	$ExNN(\text{rings} + \text{shower} + \text{track}^+)$	Cross Validation	91.38±0.07	94.56±0.05	2.20±0.06	89.48±0.07	94.12±0.04	1.13±0.03	90.98±0.04	94.85±0.03	1.20±0.07	92.57±0.03	95.81±0.03	0.90±0.06

Tabela 8.4: Valores de eficiência (P_D , SP e P_F) para a referência (LLH) medida em toda estatística disponível e do *ensemble* neural operando no ponto mais próximo ao P_D (coluna ressaltada em verde) da referência e apresentado como o valor central e 1 RMS das eficiências de validação cruzadas integrando cada região do espaço de fase. A incerteza de validação cruzada é desprezível. Nomenclatura empregada disponível na Seção 8.1.

Method	Type	P_D [%]	SP [%]	F_a [%]
$LLH(shower + track^+)$	Reference	96.82	97.14	2.54
$NN(rings)$	CrossValidation	96.72±0.00	95.80±0.00	5.12±0.00
$NN(track^+)$	CrossValidation	96.80±0.00	94.77±0.00	7.24±0.00
$ExNN(rings + track^+)$	CrossValidation	96.75±0.00	97.76±0.00	1.22±0.00
$ExNN(rings + shower + track^+)$	CrossValidation	96.75±0.00	97.81±0.00	1.13±0.00

Observa-se uma redução no falso alarme, ao realizar-se a fusão da informação no *ensemble*, via redes especialistas, partindo de 5,12 % (7,24 %) em P_F , quando somente os anéis (grandezas de traços) são empregados para todo o espaço de fase (Tabela 8.3), para atingir 1,22%. Essa redução de contaminação de ruído para as análises físicas se distribui em todas as regiões do espaço de fase, geralmente resultando em um fator de 50 % (valor central) na contaminação em relação à referência. Essa redução se distribui com certa uniformidade para todas as regiões, exceto para a região do *crack* e a região mais externa do detector ($2,37 < |\eta| < 2,47$). Ainda que a simulação seja bastante fiel para boa parte das regiões do espaço de fase, há especialmente uma limitação de representação para essas regiões. A LLH , por sua vez, foi ajustada em dados de colisão, e portanto não obtém o ganho em eficiência devido à essa limitação de representar toda a complexidade do desenvolvimento do chuveiro nessas regiões. Assim, é possível que esse ganho não ocorra conforme obtido em simulações quando avaliando dados de colisão. Além disso, em geral, os elétrons no *crack* não participam das análises físicas devido à sua pior reconstrução. A incerteza estatística é pequena o suficiente para dar margem aos resultados, porém, como dito, a LLH foi ajustada para operação em dados de colisão e cabendo uma comparação em dados de colisão e de posse de toda a incerteza envolvida para a sistemática do processo de medição de eficiência por T&P. A estratégia de adição das grandezas físicas na fusão da informação fornece uma melhoria considerável de eficiência (1,22 \rightarrow 1,13 % em P_F para todo o espaço de fase) que, igualmente, distribui-se por com certa uniformidade por todas as regiões. A hipótese para justificar esse ganho dá-se da seguinte forma: apesar dos anéis se apresentarem como uma generalização² de boa parte das grandezas físicas de calorimetria, a grandeza E_{ratio} não é bem aproximada pelos anéis uma vez que ela considera a soma em diversas células ao entorno do 1EM1. Com isso, acrescenta-se informação discriminante, além do fato dessas representações conhecidas por sua capacidade discriminante serem apresentar diretamente para a rede sem

²Ou seja, é possível obter variáveis muito próximas às grandezas a partir dos anéis a partir da soma e razão de seus valores.

a necessidade da mesma compor padrões parecidos via minimização do MSE.

Ao avaliar esses resultados em função das variáveis compondo o espaço de fase (Figura 8.8), percebe-se que o perfil de eficiência contém oscilações inferiores à 1 % em função de E_T para todos os casos, porém para a região central de operação essa oscilação é inferior à 0,5 % em sua maioria. Observa-se que os *ensembles* mantêm a tendência de aumento em eficiência em maiores energia sem aumento de falso alarme, o que pode ser um indício dos mesmos terem capturado melhor a essência do chuvaire nessa região. Isso se confirmando, seria bastante interessante para a análise física, uma vez que a física rara deve decair em elétrons de alta energia. Em η , observa-se uma ligeira tendência dos *ensembles* a operarem com um P_D ligeiramente ($\approx 0,5\%$) mais baixo em algumas regiões na tampa ($|\eta| > 1,52$).

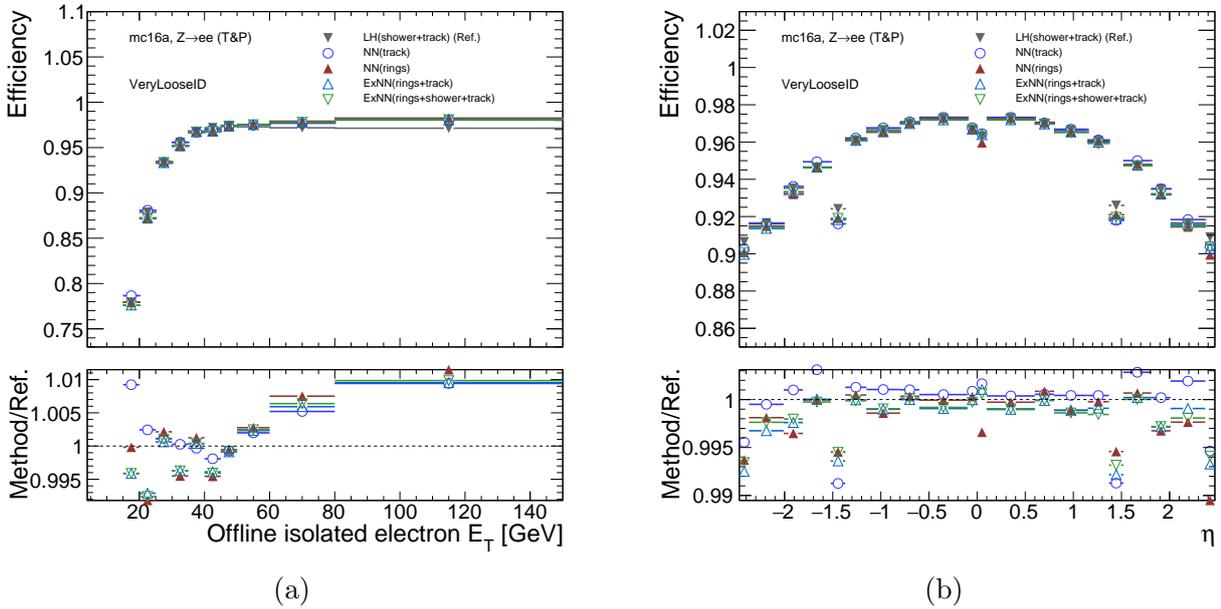


Figura 8.8: Eficiência em T&P (figura superior) e sua razão (inferior) em função de E_T (a) e η para as estratégias de seleção de elétrons desenvolvidas para atuação no SR_{off} e a atual referência atuando no critério *loose*. Nomenclatura empregada disponível na Seção 8.1.

O falso alarme (Figura 8.9), por sua vez, confirma o efeito observado para os *ensembles* explorando apenas uma única representação de informação. O *ensemble* em anéis mostra uma tendência de se aproximar ao P_F da referência conforme o aumento do E_T , obtendo eficiência próxima a da referência em alta energia (> 80 GeV), enquanto aquele alimentado pelas grandezas de traço perde sua capacidade discriminante. As versões com fusão de informação, além de apresentarem a redução de falso alarme já mencionada, mostram um perfil mais uniforme que a referência para ambas variáveis.

Quanto à resiliência para o efeito de empilhamento (Figura 8.10), a referência possui um efeito mais pronunciado de queda em função do empilhamento em comparação com os *ensembles*, apesar da correção. É possível que isso tenha ocorrido pelo fato dos parâmetros

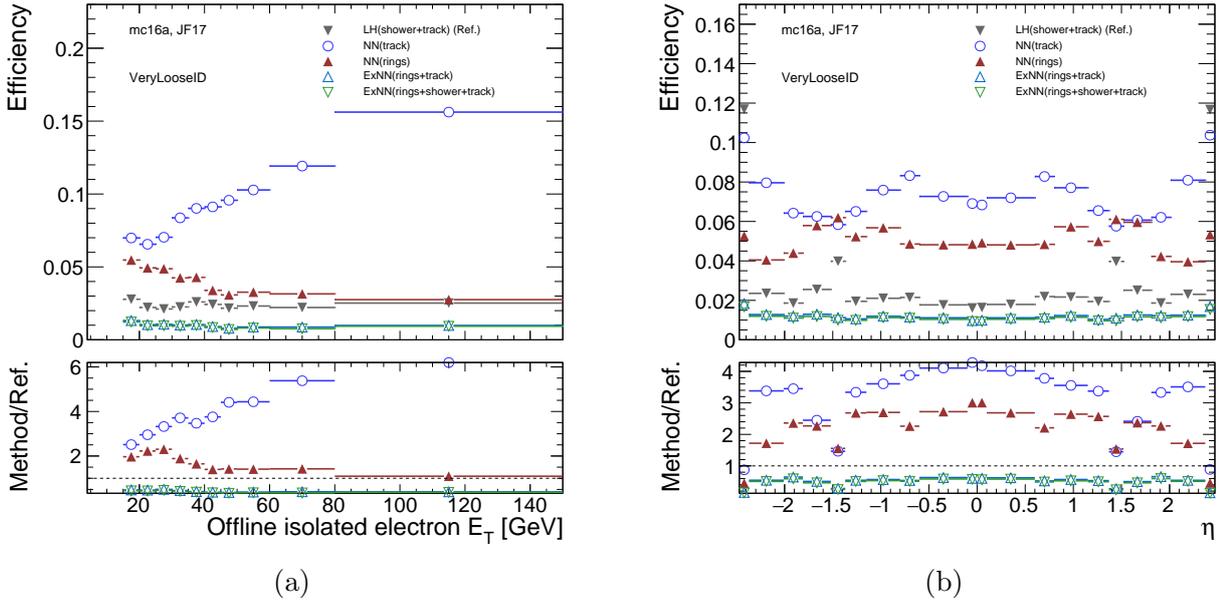


Figura 8.9: Falso alarme (figura superior) e sua razão (inferior) em função de E_T (a) e η para as estratégias de seleção de elétrons desenvolvidas para atuação no SR_{off} e a atual referência atuando no critério *vloose*. Nomenclatura empregada disponível na Seção 8.1.

do ajuste da *LLH* terem sido calculados em dados de colisão. Observou-se, durante o desenvolvimento do *NeuralRinger* para atuação no SF_{on} que a simulação apresenta menor influência na eficiência em função do empilhamento, porém as medições no SF_{on} são realizadas em função da própria *LLH*, então não é claro até que ponto a natureza operativa da *LLH* interferiu nesse aspecto. Somente após a medição da eficiência em colisão será possível tirar conclusões sobre esse aspecto. O ajuste no *ensemble* acessando somente informação de calorimetria sofre com um aumento considerável de P_F em função de $nvtx$, enquanto o *ensemble* em traço, apesar de possuir pior operação, é mais estável. Há uma tendência de aproximação de P_F da *LLH* e dos *ensembles* com fusão para maior empilhamento, sendo possível ser do mesmo motivo do cálculo do parâmetro de correção em colisão.

A análise de quadrante realizada para os *ensembles* de fusão (comparar Figura 8.11a e 8.11b) não mostram diferenças nítidas no perfil de coleta de elétrons em função de R_η quando comparando a versão adicionada das grandezas de calorimetria em relação a versão fundido apenas os anéis ao traço. Isso pode ser decorrência do critério em análise, onde espera-se que as diferenças na saída das redes do *ensemble* ocorram em regiões menos bem representadas. Ou seja, o deslocamento de uma saída ruim no *ensemble* para a versão acessando apenas anéis ao traço seja deslocado para um região melhor no *ensemble* fundido toda a informação, mas não que esse deslocamento ocorra em regiões em que a saída neural está próxima do alvo, uma vez que se espera que ambas tenham conseguido capturar os padrões das informações discriminantes lhes apresentadas que se sobrepõe para a versão obtendo informação extra

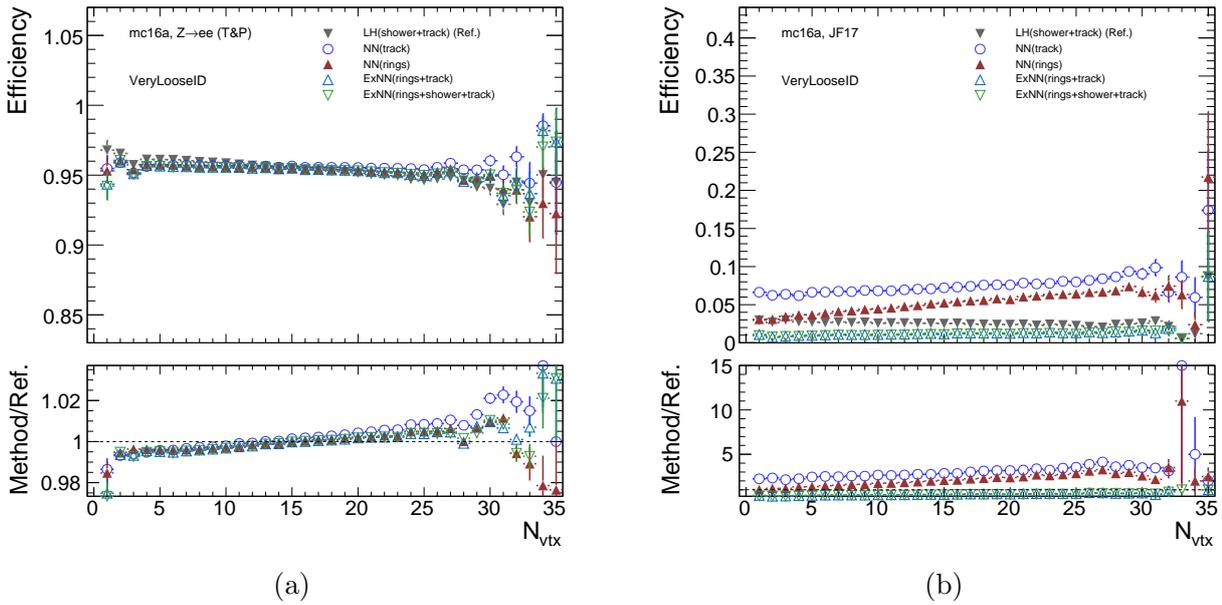


Figura 8.10: Figura superior: Eficiência em T&P (a) e falso alarme (b) em função de n_{vtx} para as estratégias de seleção de elétrons desenvolvidas para atuação no SR_{off} e a atual referência atuando no critério *vloose*. Na figura inferior estão as razões das eficiências em relação à referência. Nomenclatura empregada disponível na Seção 8.1.

de grandeza de calorimetria. Quando comparando o perfil das ocorrências de discordância entre os *ensembles* e a referência, há uma ligeira tendência de coletar amostras com melhor qualidade pela descrição da grandeza física (valores mais próximos a 1 são melhores elétrons por conterem menor vazamento de energia lateral).

Já a análise de quadrante em ruído físico (Figura 8.12)³ possui muito poucas amostras de falso alarme para inferir qualitativamente uma diferença de comportamento. A única informação clara já foi inferida quando analisando o falso alarme em termos de eficiência: o *ensemble* tende a aceitar menos falsos elétrons.

³Como a informação dos dois *ensembles* são muito próximas, para evitar redundância manteve-se somente o caso com menor P_F .

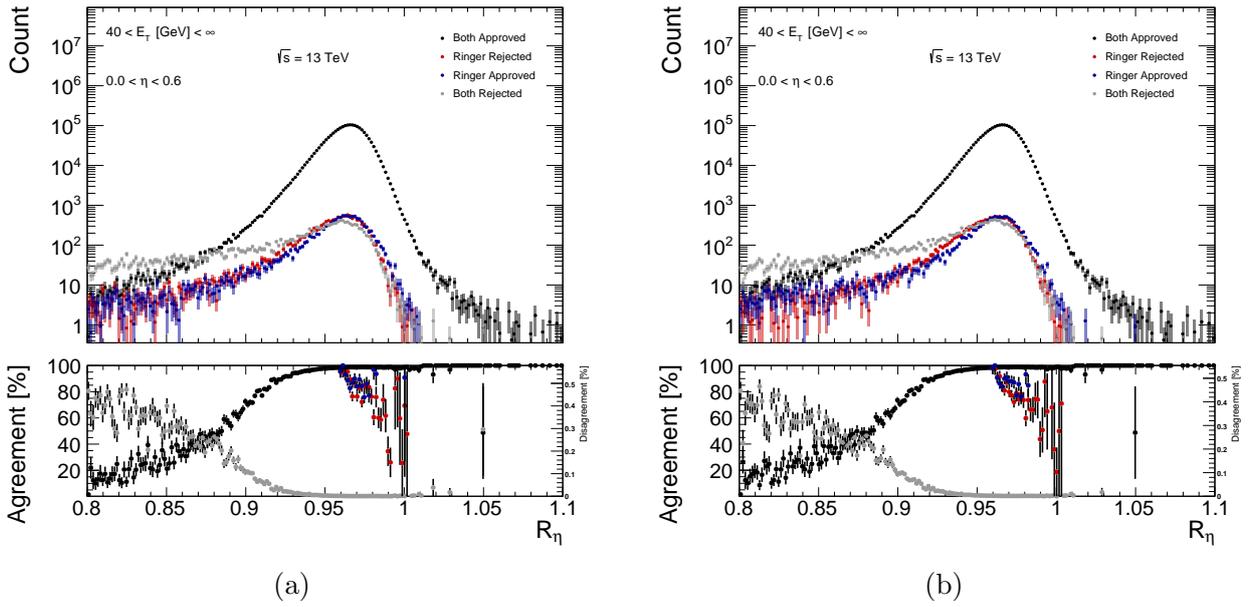


Figura 8.11: Análise de quadrante em para a grandeza física R_η de $ExNN(rings + track^+)$ (a) e $ExNN(rings + shower + track^+)$ em relação à referência (LLH). Na figura inferior, está a razão entre os casos mutuamente exclusivos pelo total de observações. Há duas escalas, a da esquerda refere-se à concordância entre as cadeias (caso I e IV) e a da direita refere-se à discordância (caso II e III).

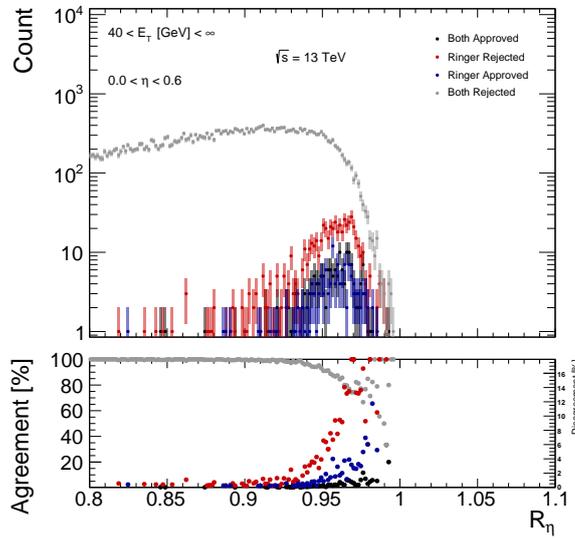


Figura 8.12: Análise de quadrante em ruído físico para a grandeza física R_η de $ExNN(rings + shower + track^+)$ em relação à referência (LLH). Na figura inferior, está a razão entre os casos mutuamente exclusivos pelo total de observações. Há duas escalas, a da esquerda refere-se à concordância entre as cadeias (caso I e IV) e a da direita refere-se à discordância (caso II e III).

Capítulo 9

Limitações e Aperfeiçoamentos

Deu-se prioridade ao desenvolvimento das técnicas principais deste trabalho, envolvendo grande necessidade de processamento. Isso exigiu forte demanda computacional, dividida com o desenvolvimento do método para atuação no *trigger*, de forma que o levantamento de limitações e aperfeiçoamentos visando à melhoria de eficiência para a aplicação do método em *offline* não puderam ser completamente avaliadas. Em certos casos, a avaliação dessas técnicas envolve a melhoria dos padrões dos anéis ou a consideração de outras representações a serem empregadas ou adicionadas, de forma que seria necessário sua inclusão em reprocessamentos oficiais. Para que houvesse justificativa para a exploração desses novos padrões, primeiro seria necessário o desenvolvimento completo da análise em *offline* e, dado a demanda por parte da atuação do método do Sistema de Filtragem do ATLAS, isso não pode ser realizado.

Assim, realiza-se uma consideração quanto à abordagem para simplificar e reduzir a necessidade de processamento no desenvolvimento dos métodos (Seção 9.2) via menor emprego de informação a ser processada durante os ajustes.

Outra parte que deve ser considerada se dá pela questão da informação disponível para o processo de treinamento (Seção 9.1), com maior impacto quando realizando o ajuste em dados de colisão. A técnica proposta pode auxiliar na abordagem para lidar com o processamento de grande massas de dados e, por isso, esse tema é descrito anteriormente.

Igualmente, o pré-processamento atualmente empregado se limita à aplicação unicamente de conhecimento especialista: anelamento e normalização por energia. Melhorias podem ser realizadas no método através da obtenção de um melhor espaço de características. Avaliações iniciais envolvem pré-processamento estatístico (no caso explorou-se ICA, baseando-se nos resultados anteriores de [198], Seção 9.4) e a compressão do espaço representado na base da informação anelada através da aplicação de Aprendizado Profundo por via da técnica *Stacked Autoencoder* (Seção 9.3).

Os trabalhos anteriores de pesquisa (ver Subseção 6.2) consideraram melhorias na representação da base anelada de calorimetria através de processamento estatístico e conhecimento

especialista, porém melhorias na própria representação da informação desse detector ainda não foi profundamente analisada. Tal fato é uma decorrência da demanda bastante restritiva em termos de esforço computacional do SF_{on} , em especial em seu estágio rápido, que colocou essa análise em segundo plano. Como se verá a seguir, a representação adotada, apesar de uma boa abordagem inicial para o SF_{on} , possui limitações que podem ser aprimoradas para atuação no SR_{off} e, inclusive, estendidas posteriormente para atuação no próprio SF_{on} . Isso inclui a consideração do tamanho da janela (Seção 9.5), o tratamento de regiões anômalas, onde o algoritmo descreve anéis em regiões do calorímetro nas quais células possuem granularidade distinta àquela empregada no algoritmo (Seção 9.6), e outras considerações sobre a região do *crack*. Além disso, este trabalho não está limitado à descrição por informação anelada de calorimetria, sendo proposto outra abordagem à questão das regiões anômalas através da descrição por densidade de energia (Seção 9.7).

9.1 Estratégia para Homogeneização dos Subconjuntos de Validação Cruzada

A informação proveniente do ATLAS para identificação de elétrons está sujeita a uma série de influências além da natureza do objeto físico desenvolvendo os sinais em seus diversos sensores. A física de interesse é rara e, ainda que as observações empregadas para o ajuste dos modelos não o sejam, a perda de eventuais elétrons por não capturar pequenas nuances durante o processo de ajuste tem impacto na física do experimento. Ou seja, alta eficiência não basta caso os elétrons de física de interesse sejam observados através de pequenas nuances que não estavam presentes durante o ajuste e avaliação de generalização das técnicas. Por isso, acredita-se que a homogeneização dos dados nos subconjuntos possa auxiliar no processo de captura ao garantir que amostras constituindo um determinado grupo estejam distribuídas de maneira equilibrada nos subconjuntos.

Isso é ainda mais importante para o ajuste em dados de colisão, dado a estrutura temporal empregada para a coleta. Esses ajustes permitem uma maior eficiência ao não estarem sujeitos às limitações de simulações por MC de representarem completamente o desenvolvimento do chuveiro e as respostas dos detectores para o desenvolvimento desse sinal.

O método empregado para a obtenção de grupos representando os dados baseou-se na extração de SOMs, conforme:

- Normalização dos dados por norma-1, aplica-se PCA (com a prévia remoção da média) para obter compressão da informação em 95 % dos dados a fim de agilizar o processo de treinamento (mínimo de cinco componentes). Os valores obtidos de componentes para cada região do espaço de fase estão disponíveis na Tabela 9.1 para elétrons e na Tabela 9.2 para falsos elétrons;

- A quantidade de unidades ($\lceil 5 \cdot \sqrt{N} \rceil$), dimensão no mapa (distribuição do número de unidades nas dimensões proporcional à relação da variância dos dois componentes principais), inicialização (distribuição linear na região definida entre o máximo e o mínimo da distribuição das observações na projeção das componentes principais) é realizado com os valores padrões do programa disponibilizado por [247];
- Aplica-se o treinamento por bateladas do SOM com treliça hexagonal com função de vizinhança Gaussiana com o mesmo programa [247]. Empregam-se 1.000 épocas de treinamento com ajuste grosso ($\sigma = \lceil \max(dim)/2 \rceil$, onde dim é a dimensão do mapa) e outras 500 épocas para ajuste fino ($\sigma = 2$). Esses valores foram alterados pois a heurística comumente empregada não se adequa quando $N \rightarrow \infty$, caso de grande massa de dados.

Como parâmetros de avaliação da capacidade do SOM de representar o espaço original, empregou-se as medidas erro de quantização (Q_e) e erro topográfico (T_e) [128]. No caso,

$$Q_e = \frac{1}{N} \sum_i^N \|\vec{x}_i - \vec{w}_{BMU}\| \quad (9.1)$$

onde \vec{x}_i são as coordenadas no espaço de entrada para a i -ésima observação e \vec{w}_{BMU} é o centroide do neurônio do SOM mais competitivo (no caso, com menor distância euclidiana), referido como Melhor Unidade Máquina (BMU). Por outro lado,

$$T_e = \frac{1}{N} \sum_i^N u(\vec{x}_i) \quad (9.2)$$

onde $u(\cdot)$ é uma função mapeando neurônios se o primeiro e o segundo BMU obtidos para uma observação são vizinhos no mapa (vizinhança aqui determina que os mesmos estão conectados pela treliça definida pelo mapa) em 0 para indicar a não ocorrência de erro e 1 no caso oposto. Logo, o T_e dedica-se a representar a continuidade dos neurônios no espaço de maior dimensão e Q_e a dispersão das observações em relação aos seus neurônios.

- Realizam-se 100 ajustes por k -means de $2\text{-min}(100, \sqrt{N})$ centróides na base composta pelos centroides do mapa e mantém-se somente aquele com menor MSE para cada número de grupos empregados no método k -means. Avaliam-se os índices

$$DB = \frac{1}{N_C} \sum_{i=1}^{N_C} \max_{i \neq l} \left(\frac{s(i) + s(l)}{d_{ce}(i, l)} \right), \text{ onde} \quad (9.3a)$$

$$d_{inter}(c) = \frac{1}{N_c} \sum_{k=1}^{N_c} \|\vec{x}_k - \vec{x}_c\|. \quad \text{E} \quad (9.3b)$$

$$d_{intra}(c_1, c_2) = \|\vec{x}_{c_1} - \vec{x}_{c_2}\|. \quad \text{E} \quad (9.3c)$$

$$Sil_o = \frac{b(o) - a(o)}{\max(a(o), b(o))} \quad (9.4)$$

onde N_C é o número de *clusters*, N_c é o número de observações no c -ésimo *cluster*, $d_{inter}(\cdot)$ é a dispersão das observações de um *cluster*, $d_{intra}(\cdot, \cdot)$ é uma medida de dissimilaridade entre os centroides de dois *clusters*, $a(o)$ é a média de uma medida de dissimilaridade da observação o para todos objetos do c -ésimo *cluster* a que ela pertence e $b(o)$ é distância mínima da média dessa medida do objeto i para os demais *clusters* diferentes de c . Em todos casos, as medidas foram avaliadas em base da métrica euclidiana.

Percebe-se que o índice de Davies-Bouldin (DB) [248] fornece uma medida de dissimilaridade dos *clusters* enquanto o índice de silhueta (Sil) [249], ao contrário, confere uma medida de similaridade. Como Sil fornece uma medida para cada ponto, normalmente se emprega a média de todas silhuetas para sintetizar sua distribuição em um único valor.

Caso não se encontre um mínimo índice DB (DB_{min}) ou silhueta média máxima ($\langle Sil \rangle_{max}$) inferior a cinco unidades do valor máximo de centroides avaliados pelo algoritmo k -means, o processo continua computando o algoritmo k -means com mais centroides, em bateladas de 5 em 5, até que o processo atenda essa condição.

Os valores obtidos durante a aplicação do método para o agrupamento dos dados estão na Tabela 9.1 (9.2) para elétrons (ruído físico). Observa-se que os T_e estão altos para todos os mapas ($T_e > 0,5$ para a maioria dos mapas, tanto em sinal e ruído). Como a porcentagem de BMU vazios está baixa, acredita-se que ao aumentar as dimensões do mapa se obtenha menores valores de T_e . Todavia, o processo de ajuste dos mapas demandou grande custo computacional¹ com as dimensões aplicadas no método, de forma que isso exigiria ainda mais recursos. Para os mapas obtidos, observou-se que, em geral, os índices encontram baixo número de *clusters* com BMU no mapa.

Para elétrons, os mapas apresentam padrões interessantes. Em geral, aqueles extraídos para a região da tampa mostram um claro perfil de alteração das componentes principais em relação à posição de incidência da partícula (exemplo na Figura 9.1). Esse comportamento está de acordo com aquele observado nas grandezas físicas (Seção 5.2), onde as mesmas

¹Três semanas empregando 32 núcleos físicos com configurações atuais.

Tabela 9.1: Resultados para o ajuste de SOM na informação anelada de calorimetria processada por PCA e de k -means para o agrupamento dos neurônios do mapa em dados de simulação compatíveis com as condições de colisão de 2016 para amostras de elétrons pré-selecionadas pelo método T&P. Uma série de propriedades e medidas envolvidas estão disponíveis, dentre eles a dimensão do mapa (Dim.), o número de componentes da compressão por PCA em 95 % de energia (n_{comp}), erro de quantização (Q_e), erro topográfico (T_e), porcentagem de BMU do mapa vazias (\varnothing_{bmu} [%]), índice de Davis-Bouldin mínimo (DB_{min}) com o respectivo número de $clusters$ empregados no k -means e de $clusters$ com BMU atingidas (k_{bmu}), o mesmo para a máxima silhueta média ($\langle Si \rangle_{max}$), e o número de neurônios escolhidos por análise gráfica de DB e Sil com os respectivos valores obtidos para os índices. Em alguns casos, manteve-se a escolha através de $\langle Si \rangle_{max}$, indicado por esse símbolo no local referente à escolha.

	$0.0 < \eta < 0.8$	$0.8 < \eta < 1.37$	$1.37 < \eta < 1.54$	$1.54 < \eta < 2.37$	$2.37 < \eta < 2.47$	
$15.0 < E_T[\text{GeV}] < 20.0$	Dim.	89×87	90×72	62×54	86×84	48×46
	n_{comp}	16	16	14	16	13
	Q_e	0.06	0.06	0.07	0.06	0.05
	T_e	0.8	0.8	0.7	0.7	0.6
	\varnothing_{bmu} [%]	0	0	0.030	0	0
	$DB_{min}(k, k_{bmu})$	1.0 (33, 7)	1.0 (23, 7)	1.0 (2, 2)	1.2 (7, 7)	1.2 (37, 37)
	$\langle Si \rangle_{max}(k, k_{bmu})$	1.0 (19, 2)	1.0 (2, 2)	1.0 (2, 2)	0.35 (4, 4)	0.4 (7, 7)
	$DB \langle Si \rangle(k_c, k_{c,bmu})$	2.2 0.21 (52, 21)	2.1 0.29 (18, 5)	$\langle Si \rangle_{max}$	1.2 0.33 (7, 7)	1.4 0.4 (7, 7)
	$20.0 < E_T[\text{GeV}] < 30.0$	Dim.	121×109	117×94	113×52	116×109
n_{comp}		14	10	2 (5)	13	10
Q_e		0.04	0.04	0.021	0.04	0.029
T_e		0.8	0.8	0.8	0.7	0.6
\varnothing_{bmu} [%]		0	0	0.09	0	0
$DB_{min}(k, k_{bmu})$		0.7 (2, 2)	1.0 (82, 3)	0.7 (2, 2)	1.1 (8, 5)	1.2 (8, 8)
$\langle Si \rangle_{max}(k, k_{bmu})$		1.0 (2, 2)	1.0 (31, 2)	1.0 (2, 2)	0.4 (9, 6)	0.4 (6, 6)
$DB \langle Si \rangle(k_c, k_{c,bmu})$		1.8 0.25 (16, 16)	1.0 1.0 (57, 2)	$\langle Si \rangle_{max}$	2.5 0.34 (19, 7)	1.2 0.4 (6, 6)
$30.0 < E_T[\text{GeV}] < 40.0$		Dim.	136×117	121×109	543×13	128×119
	n_{comp}	10	10	1 (5)	8	7
	Q_e	0.026	0.030	0.018	0.025	0.015
	T_e	0.8	0.8	0.8	0.10	0.5
	\varnothing_{bmu} [%]	0	0	0.10	0.013	0
	$DB_{min}(k, k_{bmu})$	1.2 (4, 4)	1.0 (25, 8)	0.4 (2, 2)	1.1 (9, 9)	1.1 (24, 24)
	$\langle Si \rangle_{max}(k, k_{bmu})$	0.4 (2, 2)	0.19 (94, 26)	1.0 (2, 2)	0.5 (3, 3)	0.5 (6, 6)
	$DB \langle Si \rangle(k_c, k_{c,bmu})$	1.2 0.4 (4, 4)	$\langle Si \rangle_{max}$	$\langle Si \rangle_{max}$	1.1 0.4 (9, 9)	1.1 0.5 (6, 6)
	$40.0 < E_T[\text{GeV}] < 50.0$	Dim.	137×113	118×108	517×13	141×104
n_{comp}		7	8	1 (5)	6	6
Q_e		0.017	0.023	0.021	0.011	0.012
T_e		0.7	0.8	0.8	0.6	0.5
\varnothing_{bmu} [%]		0	0	0.030	0.007	0
$DB_{min}(k, k_{bmu})$		1.5 (5, 5)	1.3 (8, 8)	0.5 (2, 2)	1.0 (40, 6)	1.0 (26, 26)
$\langle Si \rangle_{max}(k, k_{bmu})$		0.4 (2, 2)	0.34 (2, 2)	1.0 (2, 2)	0.33 (74, 35)	0.5 (6, 6)
$DB \langle Si \rangle(k_c, k_{c,bmu})$		2.8 0.30 (12, 12)	1.7 0.22 (14, 14)	$\langle Si \rangle_{max}$	1.0 0.25 (40, 6)	1.1 0.5 (6, 6)
$E_T[\text{GeV}] > 50.0$		Dim.	118×94	148×61	90×55	101×99
	n_{comp}	4 (5)	2 (5)	3 (5)	6	6
	Q_e	0.08	0.08	0.030	0.014	0.012
	T_e	0.32	0.4	0.7	0.6	0.5
	\varnothing_{bmu} [%]	0	0	0.06	0	0
	$DB_{min}(k, k_{bmu})$	0.7 (12, 12)	0.5 (3, 3)	0.9 (79, 3)	1.1 (13, 13)	1.0 (34, 34)
	$\langle Si \rangle_{max}(k, k_{bmu})$	0.7 (2, 2)	0.8 (2, 2)	1.0 (96, 3)	0.4 (4, 4)	0.5 (6, 6)
	$DB \langle Si \rangle(k_c, k_{c,bmu})$	0.7 0.6 (12, 12)	0.6 0.8 (2, 2)	$\langle Si \rangle_{max}$	1.2 0.4 (4, 4)	1.1 0.5 (6, 6)

Tabela 9.2: Resultados para o ajuste de SOM na informação anelada de calorimetria processada por PCA e de k -means para o agrupamento dos neurônios do mapa em dados de simulação compatíveis com as condições de colisão de 2016 para amostras de ruído físico. Uma série de propriedades e medidas envolvidas estão disponíveis, dentre eles a dimensão do mapa (Dim.), o número de componentes da compressão por PCA em 95 % de energia (n_{comp}), erro de quantização (Q_e), erro topográfico (T_e), porcentagem de BMU do mapa vazias ($\emptyset_{bmu}[\%]$), índice de Davis-Bouldin mínimo (DB_{min}) com o respectivo número de $clusters$ empregados no k -means e de $clusters$ com BMU atingidas (k_{bmu}), o mesmo para a máxima silhueta média ($\langle Sil \rangle_{max}$), e o número de neurônios escolhidos por análise gráfica de DB e Sil com os respectivos valores obtidos para os índices. Em alguns casos, manteve-se a escolha através de $\langle Sil \rangle_{max}$, indicado por esse símbolo no local referente à escolha.

		$0.0 < \eta < 0.8$	$0.8 < \eta < 1.37$	$1.37 < \eta < 1.54$	$1.54 < \eta < 2.37$	$2.37 < \eta < 2.47$
$15.0 < E_T[\text{GeV}] < 20.0$	Dim.	71×57	66×54	49×44	75×53	41×31
	n_{comp}	22	4 (5)	14	19	20
	Q_e	0.07	0.018	0.11	0.07	0.08
	T_e	0.7	0.7	0.7	0.7	0.6
	$\emptyset_{bmu}[\%]$	0	0.06	0	0	0
	$DB_{min}(k, k_{bmu})$	1.4 (46, 46)	1.0 (56, 5)	1.2 (9, 5)	0.8 (7, 2)	1.4 (73, 73)
	$\langle Sil \rangle_{max}(k, k_{bmu})$	0.4 (2, 2)	1.0 (2, 2)	1.0 (2, 2)	1.0 (49, 3)	0.35 (2, 2)
	$DB \langle Sil \rangle (k_c, k_{c,bmu})$	1.6 0.25 (8, 8)	$\langle Sil \rangle_{max}$	$\langle Sil \rangle_{max}$	$\langle Sil \rangle_{max}$	1.5 0.24 (15, 15)
$20.0 < E_T[\text{GeV}] < 30.0$	Dim.	59×51	96×27	57×28	58×51	34×28
	n_{comp}	20	2 (5)	8	20	19
	Q_e	0.07	0.019	0.06	0.07	0.07
	T_e	0.7	0.7	0.7	0.7	0.6
	$\emptyset_{bmu}[\%]$	0	0.08	0.13	0	0
	$DB_{min}(k, k_{bmu})$	1.5 (15, 15)	0.5 (2, 2)	1.0 (59, 4)	1.0 (2, 2)	1.3 (65, 65)
	$\langle Sil \rangle_{max}(k, k_{bmu})$	0.4 (2, 2)	1.0 (2, 2)	1.0 (74, 4)	1.0 (2, 2)	0.4 (2, 2)
	$DB \langle Sil \rangle (k_c, k_{c,bmu})$	1.6 0.24 (10, 10)	$\langle Sil \rangle_{max}$	$\langle Sil \rangle_{max}$	$\langle Sil \rangle_{max}$	$\langle Sil \rangle_{max}$
$30.0 < E_T[\text{GeV}] < 40.0$	Dim.	40×37	37×33	95×8	41×34	23×20
	n_{comp}	17	16	1 (5)	18	18
	Q_e	0.06	0.06	0.027	0.06	0.07
	T_e	0.6	0.6	0.6	0.6	0.5
	$\emptyset_{bmu}[\%]$	0	0	1.4	0	0
	$DB_{min}(k, k_{bmu})$	1.4 (15, 15)	1.0 (52, 9)	0.33 (2, 2)	1.4 (19, 19)	1.3 (67, 67)
	$\langle Sil \rangle_{max}(k, k_{bmu})$	0.4 (3, 3)	1.0 (2, 2)	1.0 (2, 2)	0.32 (2, 2)	0.4 (2, 2)
	$DB \langle Sil \rangle (k_c, k_{c,bmu})$	1.4 0.27 (13, 13)	$\langle Sil \rangle_{max}$	$\langle Sil \rangle_{max}$	1.4 0.25 (19, 19)	1.4 0.25 (12, 12)
$40.0 < E_T[\text{GeV}] < 50.0$	Dim.	30×29	30×24	55×8	31×26	18×15
	n_{comp}	16	15	1 (5)	16	17
	Q_e	0.06	0.05	0.023	0.06	0.07
	T_e	0.6	0.5	0.5	0.5	0.4
	$\emptyset_{bmu}[\%]$	0	0.14	1.6	0	0
	$DB_{min}(k, k_{bmu})$	1.3 (13, 13)	1.0 (54, 8)	0.5 (2, 2)	1.3 (97, 97)	1.2 (41, 41)
	$\langle Sil \rangle_{max}(k, k_{bmu})$	0.4 (2, 2)	0.9 (2, 2)	1.0 (71, 3)	0.4 (2, 2)	0.4 (2, 2)
	$DB \langle Sil \rangle (k_c, k_{c,bmu})$	1.3 0.27 (13, 13)	1.0 -0.05 (54, 8)	$\langle Sil \rangle_{max}$	1.5 0.27 (8, 8)	$\langle Sil \rangle_{max}$
$E_T[\text{GeV}] > 50.0$	Dim.	29×28	28×25	33×12	30×26	17×15
	n_{comp}	13	17	7	14	15
	Q_e	0.05	0.05	0.032	0.05	0.06
	T_e	0.5	0.5	0.4	0.5	0.4
	$\emptyset_{bmu}[\%]$	0	0	1.3	0	0
	$DB_{min}(k, k_{bmu})$	1.3 (13, 13)	1.0 (57, 4)	0.5 (2, 2)	1.3 (96, 96)	1.2 (41, 41)
	$\langle Sil \rangle_{max}(k, k_{bmu})$	0.4 (2, 2)	1.0 (2, 2)	1.0 (2, 2)	0.4 (2, 2)	0.4 (2, 2)
	$DB \langle Sil \rangle (k_c, k_{c,bmu})$	1.4 0.31 (9, 9)	1.0 0.20 (57, 4)	$\langle Sil \rangle_{max}$	1.4 0.27 (10, 10)	1.4 0.28 (12, 12)

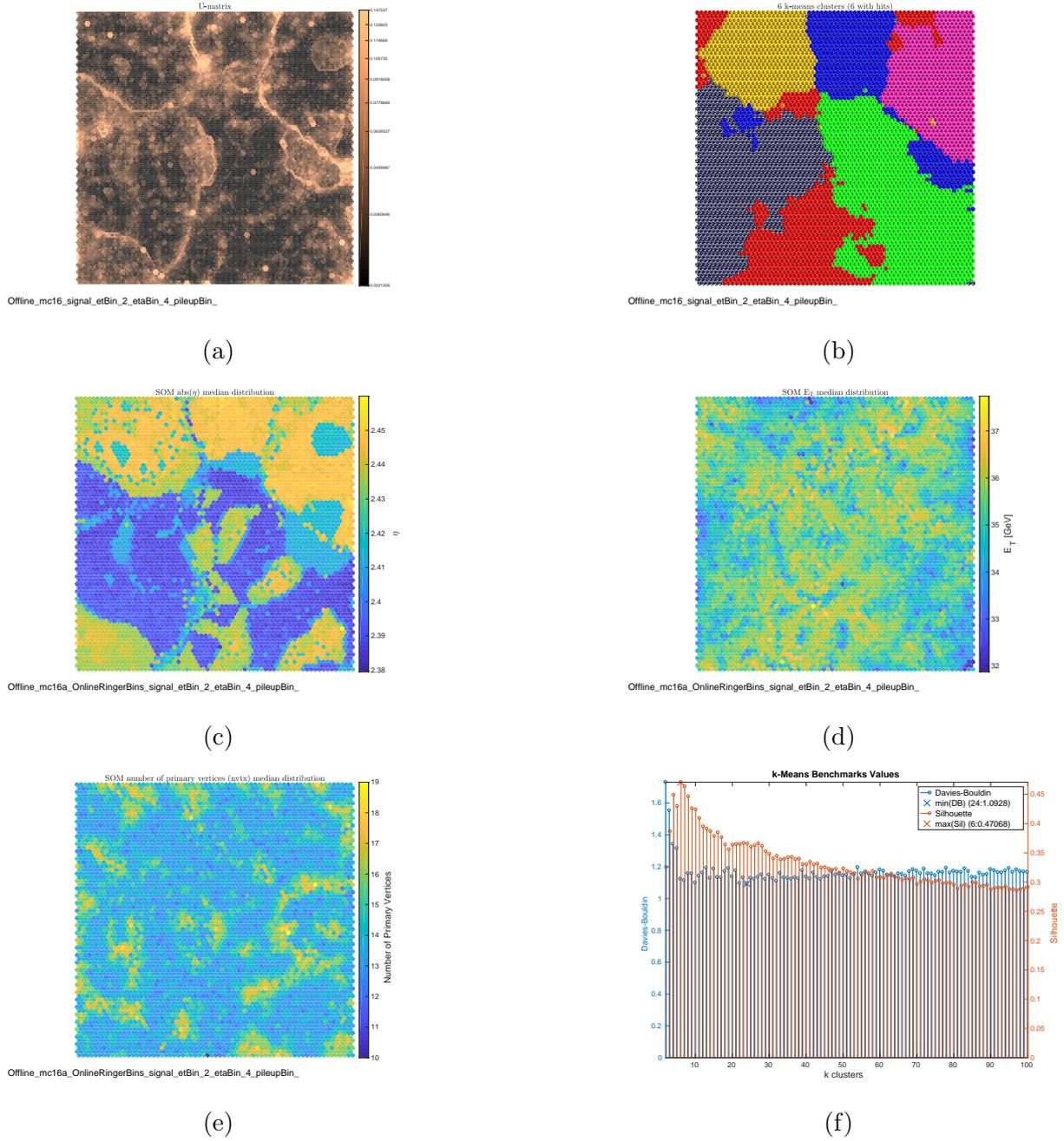


Figura 9.1: Mapa obtido para simulações em condições compatíveis com colisões de 2016 contendo decaimentos $Z \rightarrow ee$ aplicando o método T&P na região $2,37 < \eta < 2,47$ e $30 < E_T < 40$ GeV. As figuras mostram a respectiva matriz-u (a), *clusters* para o método de *k-means* para os neurônios do mapa (b), mediana em η (c), E_T (d) e n_{vtx} (e) para as distribuições das observações contidas em cada neurônio do mapa e o índice de *Sil* média e *DB* para cada configuração avaliada do método *k-means* (f). A fim de melhorar visualmente a percepção das bordas na matriz-U, aplica-se uma transformação nos limites do eixo z conforme $z_{min} = \max(\min(\frac{z_{max}}{1,5\sqrt{10}}, 1, 5z_{min}), z_{min})$ e $z_{max} = \min(10z_{min}, z_{max})$.

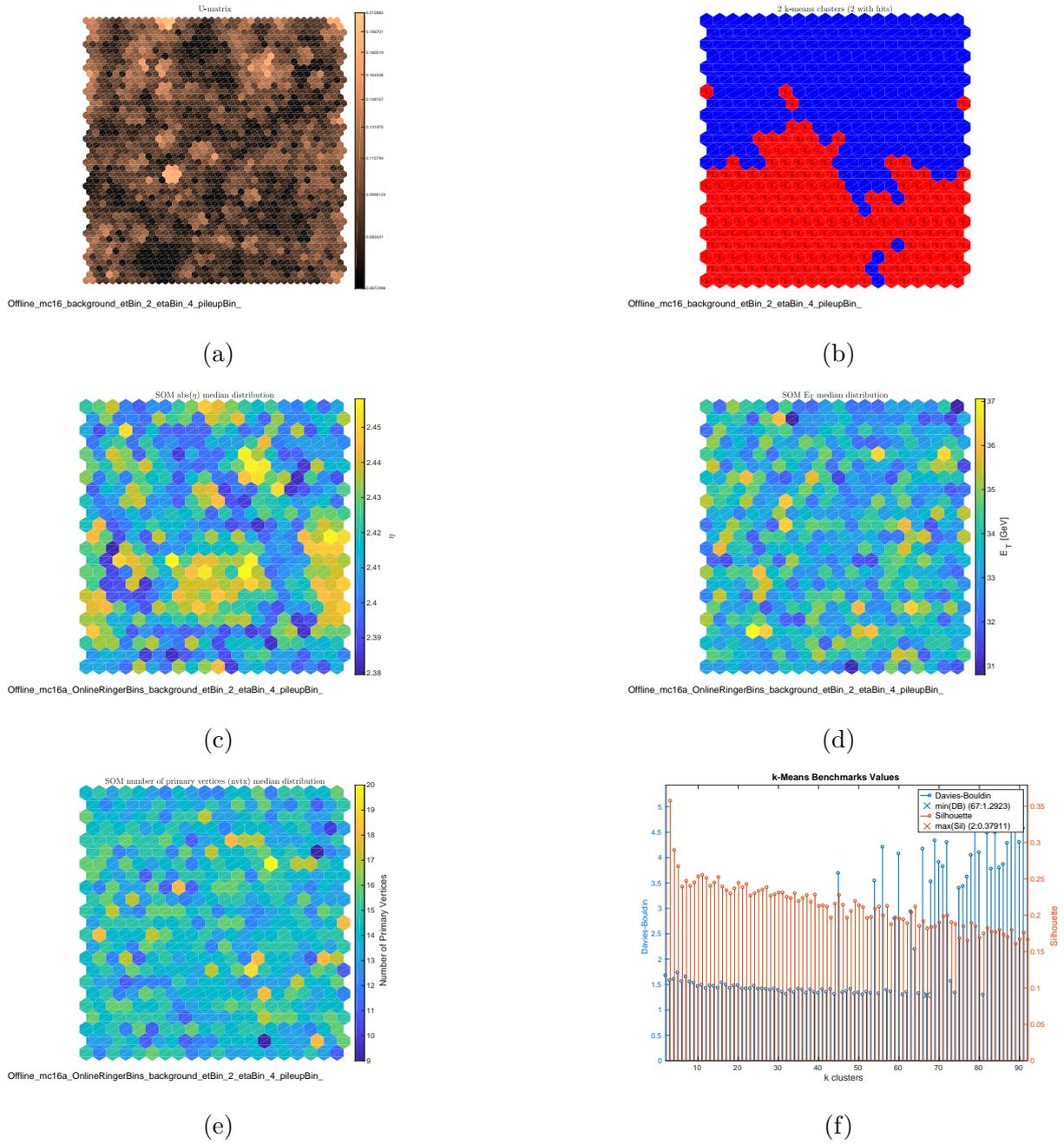


Figura 9.2: Mapa obtido para simulações em condições compatíveis com colisões de 2016 contendo decaimentos $Z \rightarrow ee$ aplicando o método T&P na região $2,37 < \eta < 2,47$ e $30 < E_T < 40$ GeV. As figuras mostram a respectiva matriz-u (a), *clusters* para o método de *k-means* para os neurônios do mapa (b), mediana em η (c), E_T (d) e n_{vtx} (e) para as distribuições das observações contidas em cada neurônio do mapa e o índice de *Sil* média e *DB* para cada configuração avaliada do método *k-means* (f). A fim de melhorar visualmente a percepção das bordas na matriz-U, aplica-se uma transformação nos limites do eixo z conforme $z_{min} = \max(\min(\frac{z_{max}}{1,5\sqrt{10}}, 1, 5z_{min}), z_{min})$ e $z_{max} = \min(10z_{min}, z_{max})$.

sofrem alteração de perfil conforme as regiões de sua incidência no detector². A variável com perfil mais nítido para o agrupamento nos mapas dá-se por $|\eta|$. O impacto por $nvtx$, quando avaliado por mediana, não apresenta neurônios além do alcance de $10 \leq nvtx \leq 19$, apesar de haver um quantidade razoável de estatística abaixo de 10 (Figura 7.6a). Isso pode ser um indício de que esses eventos não apresentam grande influência nos anéis abaixo do limiar inferior. Indo além, a variação do perfil devido à influência cinemática tende ligeiramente a formar mais neurônios com mediana superior a 35 GeV, isso se deve ao fato da maior concentração das observações nessa região do espaço de fase (Figura 7.8a).

Igualmente, os mapas mostram que o ruído físico (Figura 9.2) formam grupos menos dependentes conforme as condições das observações em função dessas variáveis. Há concordância, portanto, com o treinamento não-supervisionado obtido nos mapas e o conhecimento especialista agregado nas grandezas físicas (Seção 5.2). O motivo dá-se pelo fato do Chuveiro de jatos alastrar-se por uma região maior, logo, há menor influência no perfil dos anéis já que boa parte deles geralmente são afetados. Outro aspecto que fica evidente nas figuras se dá pela formação de regiões com dispersões não contínua no mapa, justificado pelo alto T_e obtido na convergência do treinamento.

Apesar dos mapas apresentarem alto T_e , os mesmos foram capazes de capturar o comportamento esperado por parte do conhecimento especialista. Acredita-se, ainda havendo possivelmente melhorias a ser obtidas no treinamento, que os mesmos podem auxiliar a distribuir a informação nos subconjuntos a serem aplicados pelo método de validação cruzada.

9.2 Estratégia para Redução da Demanda Computacional para o Ajuste de Modelos em Grande Massa de Dados

Apesar de todo poder computacional disponível para o ajuste dos modelos (LCG, Lobo Carneiro e *cluster* local do LPS), uma das grandes limitações para o desenvolvimento e avaliação de novas abordagens se dá pelo grande esforço de processamento necessário nessa etapa. Ainda que a LCG tenha realizado movimentações para o emprego de *hardware* com maior capacidade de processamento para o ajuste das técnicas de Aprendizado de Máquinas, não há garantias quanto ao impacto disso na capacidade de processamento já que é incerto a parcela desses recursos que serão disponibilizadas para o uso dos membros. Além disso, há motivo para crer que a tendência seja de que a estatística disponível cresça de forma mais rápida que os recursos destinados ao processamento, dado que os modelos se tornam mais complexos ao mesmo tempo que mais observações estão disponíveis. Seja por questões

²Ressaltando que os mapas não tiveram acesso a essas variáveis durante o treinamento, apenas aos anéis pré-processados por PCA conforme ressaltado durante a descrição do método.

imediatas ou previsão futura, é necessário uma estratégia para empregar os recursos de forma inteligente.

A abordagem que se deseja explorar é através da captura da informação relevante a ser apresentada para o ajuste dos modelos. Isso pode se dar por dois caminhos, primeiro reduzir aquela que é supérflua para obter menor demanda computacional, ou explorar a ideia de *bagging* [250] para determinar grupos que irão compor o subconjunto para participar do processo de treinamento. Em ambos os casos pretende-se realizar apenas o treinamento para o subconjunto determinado nos dados, enquanto o restante é mantido para o teste do modelo.

Os mapas obtidos para a Seção 9.1 podem ser empregados para determinar um valor mínimo de Sil para o qual as observações não serão removidas, já nos outro caso, subamostra-se o espaço mantendo as amostras retiradas para o teste.

O método desenvolvido para a análise do impacto (Seção 7.2) pode ser empregado ao considerar seu potencial de avaliação de distorção nos perfis das grandezas com alto poder de síntese da interação das partículas com o ATLAS. Nesse caso, poderia empregar-se uma abordagem baseada em *bagging*, ao subdividir o espaço em grupos. Esses grupos podem ser retirados da amostragem e verificar se os mesmos causam impacto no perfil original. Se for o caso, adiciona-se esse subgrupo para o treinamento, caso oposto o mesmo fica reservado para subamostragem. Com isso, espera-se manter a estatística disponível ao acessar o impacto causado pela remoção das amostras no perfil da física.

9.3 Ajuste Envolvendo Aprendizado Profundo por *Stacked Autoencoder*

Para possibilitar a compressão não-linear da informação anelada de calorimetria e, com isso, obter uma descrição mais discriminante do desenvolvimento do chuveiro, pode-se aplicar aprendizado profundo por *Stacked Auto-encoder* (SAE). A estratégia permite [151, 251, 252] obter uma estrutura discriminante em menor dimensão através de uma abordagem não-supervisionada. Ainda que existam críticas quanto a essa abordagem em comparação com o aprendizado profundo supervisionado via regularização por *dropout* ou por normalização de batelada [145], preferiu-se essa abordagem em relação ao DL por treinamento supervisionado com regularização pois deseja-se explorar as representações encontradas dos anéis pelo *Stacked Autoencoder* e verificar a informação capturada do desenvolvimento do chuveiro.

O desenvolvimento é realizado para a região $0 < |\eta| < 0,6$ e $30 < E_T[\text{GeV}] < 40$ pela a mesma conter boa quantidade de estatística para sinal e ruído físico e se consistir de uma região com boa resposta do ATLAS. O treinamento do SAE ocorre pelo método *greedy-layer wise* com batelada determinada pelo número de observações da classe de menor estatística,

critério de parada de 30 épocas e limita-se o número de épocas em 2.000. O otimizador avalia a função objetivo dada pela minimização do MSE usando o algoritmo [253] a fim de melhorar a convergência. A mudança do algoritmo de gradiente descendente para esse otimizador deu-se necessário pela alta incidência de convergência para mínimos locais do primeiro, como mostrado na Figura 9.3. Mantém-se a normalização por norma-1 para não alterar a estrutura do desenvolvimento lateral e longitudinal do chuveiro.

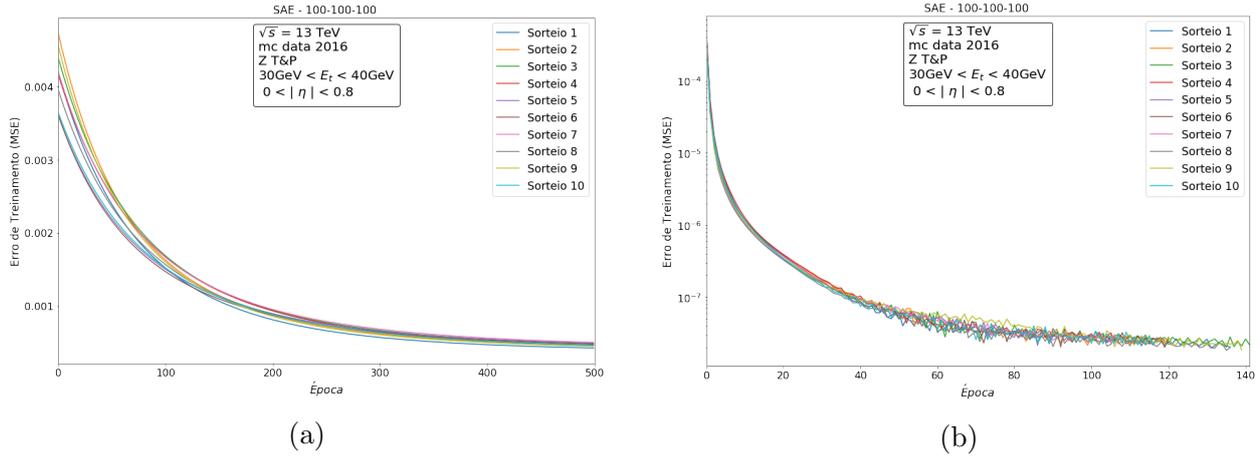


Figura 9.3: Evolução do MSE em função das épocas de treinamento da primeira camada do SAE, com 100 neurônios na camada escondida, quando empregando a otimização por gradiente descendente (a) e ADAM (b).

Como a representação da informação de cada dimensão é dada por energia, é possível realizar a soma das mesmas para verificar como as segmentações longitudinais do Sistema de Calorimetria estão sendo reconstruídas pelo SAE. Para verificar a reconstrução, empregou-se informação mútua normalizada para a região $[0, 1]$ [254]. Observa-se (Figura 9.4) que a HAD3 é a primeira a perder capacidade de reconstrução conforme a redução do número de neurônios no *encoder*, o que se mostra um bom comportamento já que essa camada captura apenas ruído na dada energia ao se posicionar no final do calorímetro. Com uma configuração empregando menos de 40 neurônios na camada escondida, o SAE começa a deteriorar a informação da EM3, que fornece informação similar a HAD1 para essa energia. Posteriormente, em torno de 25 neurônios, há SAE deterioração da informação da seção HAD. Ao minimizar o MSE, essa informação constitui de uma menor perda na função objetivo que a informação na seção EM, dado há maior concentração de captura da energia nessa seção. Comportamento similar ocorre para o PS, que captura apenas eventuais desenvolvimentos de chuveiro prematuros ao calorímetro. Por conter maior parcela da energia dos eventos, a EM2 é a última camada a sofrer deterioração.

Quando observando a curva de SP em função dos neurônios empregados no SAE, observa-se que a queda em eficiência se intensifica a partir de 25 neurônios, onde há maior perda da informação hadrônica. Na região do gráfico superior a 25 neurônios, percebe-se que

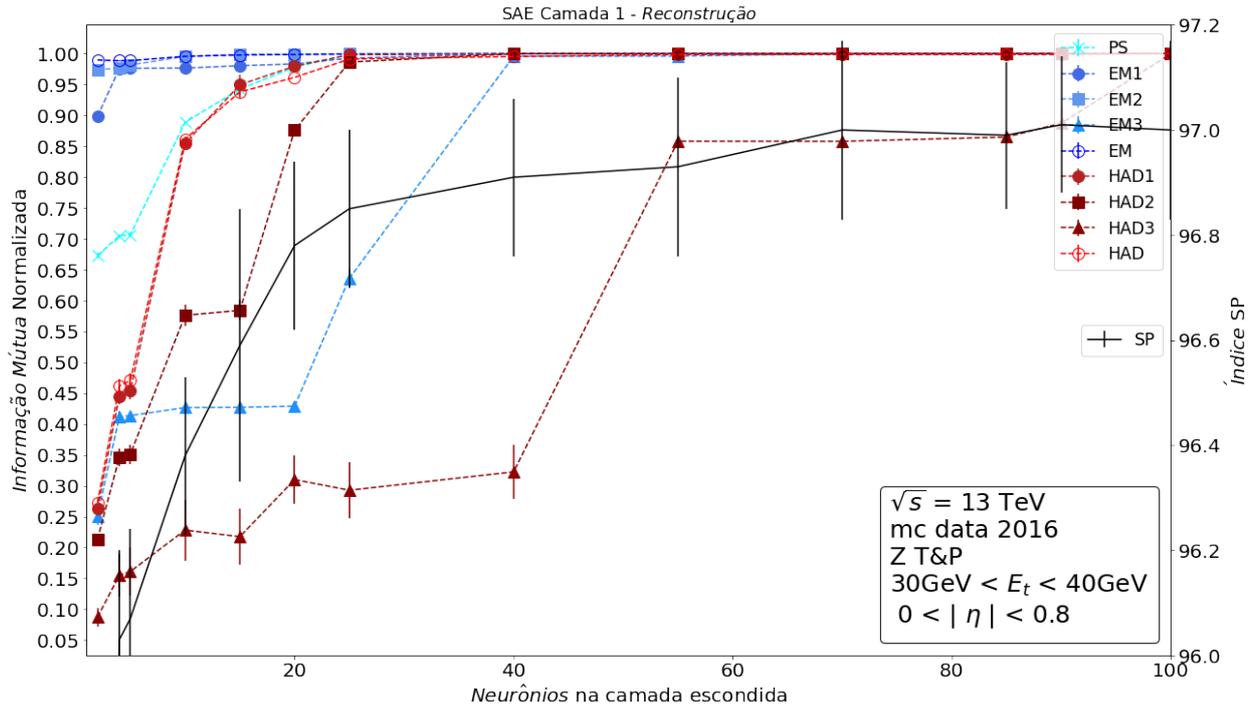


Figura 9.4: Informação mútua para as segmentações longitudinais em camadas ou seções do Sistema de Calorimetria para a aplicação da primeira camada do SAE em função do número de neurônios empregados. No eixo da direita, o valor de SP_{\max} medidos com validação cruzada por k -fold para um classificador MLP de uma camada escondida alimentado pelo espaço na camada interna do SAE.

uma perda de 20 % em informação mútua para a HAD3 não representa em perda de poder discriminante. Todavia, em 55 neurônios, pode haver um indício de que perdas em torno de uma unidade de ponto percentual em outras camadas podem ser prejudiciais para o SAE quando empregando apenas uma única camada. Espera-se, com a adição de mais camadas, melhorar a capacidade de reconstrução do SAE com um menor número de neurônios internos, o que pode possibilitar um espaço mais discriminante.

Por sua vez, a aplicação do SAE não precisa ser direcionada totalmente para a aplicação em classificação. A grande massa de dados disponível no ATLAS é um fator que dificulta para explorar outras representações da informação. O SAE pode ser uma solução de compressão para possibilitar o acesso à informações como as células de leitura ou até mesmo os pulsos dos seus canais de leitura no nível de reconstrução de elétrons.

9.4 Pré-processamento por ICA

Uma maneira de obter um espaço latente possivelmente mais discriminante dá-se pelo emprego de ICA. Esse processamento baseia-se em encontrar uma nova base de representação buscando obter independência estatística entre as populações na nova representação. No caso

da informação de calorimetria, estaria buscando encontrar as fontes originais pela descrição do desenvolvimento do chuveiro ou, ao menos, àsquelas mantidas após o pré-processamento por anelamento da informação: o desenvolvimento longitudinal e lateral do chuveiro. Espera-se que ao capturar essas fontes, o espaço contenha a essência discriminante por operar com as diferenças intrínsecas da geração dos chuveiros EM e HAD.

O trabalho anterior [198] (Seção 6.2) mostrou que essa técnica consiste de uma ferramenta poderosa a ser aliada ao pré-processamento especialista. As técnicas de ICA geralmente empregadas exigem que o número de fontes independentes a ser determinado seja igual às dimensões do espaço de entrada. Todavia, não se acredita que esse seja o caso da informação anelada. Como salientado por [198], apesar de não se ter conhecimento preciso do número de fontes envolvidas no processo de desenvolvimento de chuveiro, acredita-se que o número de componentes independentes sejam menores que a dimensão de anéis. Percebeu-se, nesse trabalho, que os anéis são uma base altamente correlacionada, inclusive não-linearmente, tanto lateral quanto longitudinalmente. A abordagem segmentada por camadas força com que o processamento ocorra acessando somente a informação lateral do desenvolvimento, possibilitando essa síntese em um estágio dedicado. Uma abordagem interessante, ainda não explorada pelos estudos anteriores (Seção 6.2), daria-se por aplicar o processamento novamente para acessar a informação longitudinal em uma base já sintetizada lateralmente.

O ferramental envolvido para a avaliação dessa análise já se encontra desenvolvido, porém pelas limitações de processamento para o desenvolvimento dos ajustes e análises dos classificadores para atuação no *offline*, ainda não estão disponíveis resultados com o rigor necessário para a divulgação científica.

9.5 Considerações sobre a Janela

O tamanho da janela do algoritmo anelador é atualmente invariante no espaço de fase $E_T \times \eta$ por herança dos primeiros estudos realizados sobre o tema. Os mesmos basearam-se na literatura disponibilizada pelo ATLAS em simulações pré-*Run 1*, que estimavam ser suficiente uma janela cobrindo $0,4 \times 0,4$ no plano $\eta \times \phi$ para conter a maior parte da deposição de energia do chuveiro. Ainda assim, nenhum estudo foi realizado para verificar se essa janela é, de fato, a melhor janela do ponto de vista de classificação. Um exemplo disso pode ser encontrado na estratégia anterior do ATLAS para criação de *clusters* EM com tamanho variável, 3×7 (em unidades de 0.025×0.025 no plano $\eta \times \phi$, ver Seção 5.1) para a região central e 5×5 na tampa. Essa variação em ϕ é decorrência da deflexão de elétrons no forte campo magnético criado pelo CS.

Além disso, o efeito de empilhamento pode alterar o tamanho ótimo para a janela devido à menor relação sinal-ruído nas regiões mais afastadas do centro de deposição de energia. Todavia, a informação lateral provém justamente dessa região de forma que eventos com

menor incidência de empilhamento poderiam explorar essa região. Assim, o tamanho da janela pode considerar regiões em função do empilhamento, o que exigiria a expansão do espaço de fase nessa dimensão. Por sua vez, isso teria implicações nas análises físicas, ainda tendo que ser avaliado a sua viabilidade.

Outra abordagem especialista ao tema seria o emprego dos *superclusters* [182] para obter anéis com energia corrigida para o empilhamento.

9.6 Tratamento de Regiões Anômalas

Uma avaliação aprofundada de (6.1) evidencia uma limitação do algoritmo atual de representar com precisão o processo de interação da partícula ao longo do Sistema de Calorimetria. A formulação do algoritmo decorre em diferença na área dos calorímetros compondo os anéis caso a granularidade não se adeque *perfeitamente* àquela empregada nos parâmetros do algoritmo em toda a RoI, ocorrências referidas aqui como regiões anômalas. Isso é uma consequência da descrição discretizada no plano $\eta \times \phi$, obtendo todo o valor de deposição de energia sumarizado em alguma posição (no algoritmo atual, centro) da célula. Essa limitação foi observada durante sua proposição, em [197], que empregou essa abordagem por uma questão de desempenho e simplicidade para aplicação no Sistema de Filtragem do ATLAS. Todavia, quando se trata do ambiente do Sistema de Reconstrução (*Offline*) de Física do ATLAS, essa preocupação se extingui. Indo além, há motivos para se acreditar que soluções desenvolvidas a seguir podem ser aplicadas sem comprometer o desempenho do algoritmo.

Ao empregar valores nulos na ausência de células, não há qualquer diferenciação para a informação descrita nessa base entre o caso em que um dado anel obteve baixo valor de energia e daquele no qual nenhuma célula se compreendeu na região por ele delimitada. Devido à natureza de dispersão do chuveiro, espera-se que haja deposição contínua no calorímetro, gradativamente menor conforme a distância ao baricentro de energia diminui, o que não é representado ao inserir valores nulos nesses anéis. Com isso, perde-se a coerência na descrição do chuveiro, causando que cada dimensão do espaço de entrada (energia dos anéis) não descreva a mesma informação do calorímetro dependendo da posição em $|\eta|$ que a partícula incide.

Evidentemente, na região anômala, a área coberta no calorímetro por cada anel varia conforme a área das células a eles pertencentes. No caso da anomalia se dar em regiões com granularidade maior do que aquela empregada pelo algoritmo (o que ocorre atualmente), haverão anéis cobrindo área nula e outros cobrindo área maior que a eles destinada. Isso ocorre para cada uma das ocasiões de troca de granularidade descrita na Tabela 4.1, onde os anéis representam informações diferentes para cada uma dessas regiões. Um exemplo de como isso afeta os anéis estão disponíveis para a Figura 9.5 para a região em $2,37 < \eta < 2,47$,

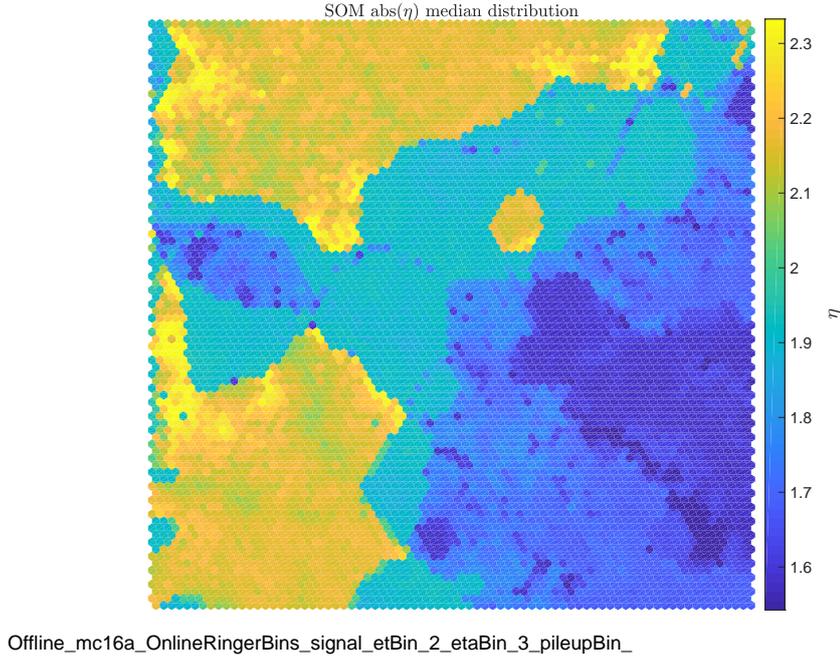


Figura 9.5: Mapa extraído conforme indicado na Figura 9.1c, porém alterando a região em η para $1,54 < |\eta| < 2,37$.

contendo anomalias na parte inferior aonde estão disponíveis as tiras da EM1 e na parte superior, aonde se deixa a região de precisão do Sistema de Calorimetria. Uma região de cobertura mais ampla em η está disponível na Figura 9.5, onde observa-se a presença de *plateaus* em três níveis dessa variável, totalmente governados em grupos no qual a borda da janela do algoritmo, de 0,2 em η atinge a região anômala. Assim, essas diferenças descritivas aumentam a complexidade do espaço de entrada aplicado para o discriminador e, mesmo que haja estatística suficiente para representar bem essas anomalias, bem como que o discriminador seja capaz de se adequar a elas, faz-se necessário corrigir tais falhas na representação do chuveiro.

A seguir, descreve-se cada um dos casos de regiões anômalas:

Caso 1 - Incidência da partícula em região com granularidade distinta: nesse caso, o baricentro de energia da partícula incide, por exemplo, na região da tampa, enquanto a granularidade assumida para o algoritmo é dado pelo valor da granularidade do barril que é diferente para a l -ésima camada em questão. A solução para esse caso é trivial, bastando ajustar a geração dos anéis para obter valores de $h_{\eta,l}$ e $h_{\phi,l}$ conforme as granularidades conhecidas e disponíveis na Tabela 4.1. Desta forma, o algoritmo computaria $h(\eta_{a,l})_{\eta,l}$ e $h(\eta_{a,l})_{\phi,l}$.

Caso 2 - Uma determinada camada simplesmente não existe na posição incidência

da partícula: como no caso anterior, porém uma camada não tem células disponível. A solução mais simples é a remoção dos anéis pertencentes a essa camada naquela subregião, porém, perde-se a informação caso a janela seja grande o suficiente a alcançar células de outras regiões. Na região de precisão estudada por este trabalho, isso só ocorre para a região do *crack*, de forma que não há prioridade em avaliar se há ganhos ao recuperar a informação dessas células. Assim, a solução a ser considerada é a remoção da camada;

Caso 3 - Incidência em uma região de transição de granularidade: nesse caso, células com distintos tamanhos ($h_{\eta,l,1}$ e $h_{\eta,l,2}$), ou até mesmo ausentes, compõem a janela para a extração dos anéis numa dita região anômala ($\Theta_{l,anom}$) na l -ésima camada. Prevêem-se duas soluções:

Solução 3.1 - *Ponderação da energia anômala na base dos anéis:* A solução apresentada restringe-se apenas para as ocorrências em que o eixo η^3 sofre alteração de granularidade em $\Theta_{l,anom}$. Para isso, define-se

$$h_{\eta,l,min} = \min(h_{\eta,l,1}, h_{\eta,l,2}), \quad (9.5a)$$

$$h_{\eta,l,max} = \max(h_{\eta,l,1}, h_{\eta,l,2}). \quad (9.5b)$$

Percebem-se dois subcasos para a granularidade $h_{\eta,l}$ empregada pelo algoritmo de anelamento em (6.1):

a) $h_{\eta,l} = h_{\eta,l,min}$;

b) $h_{\eta,l} = h_{\eta,l,max}$.

Em ambas possibilidades, obtêm-se $N_{l,\eta}$ correções cíclicas, cuja base em η é $ciclo_{(l,\eta)}$, a serem aplicadas para um anel $R_{l,n}$ que está dentro da s -ésima subregião, conforme

$$ciclo_{(l,\eta)} = \frac{mmc_{\eta}}{h_{\eta,l}}, \quad (9.6a)$$

$$N_{l,\eta} = \frac{\eta_{(\Theta_{RoI,l}^*)}}{ciclo_{(l,\eta)}}, \quad (9.6b)$$

$$s = \left\lfloor \frac{\eta_{R_{l,n}} - \eta_{(base,\Theta_{l,anom})}}{mmc_{\eta}} \right\rfloor \quad \forall R_{l,n} \in \Theta_{l,anom}, \quad (9.6c)$$

onde mmc_{η} é o menor múltiplo comum entre $h_{\eta,l,1}$ e $h_{\eta,l,2}$; $\eta_{(\Theta_{RoI,l}^*)}$ é o limite total coberto para a janela de anelamento na direção de $\Theta_{RoI,l}$ partindo do centro (no caso atual, $\eta_{\Theta_{RoI,l}^*} = 0, 2$); $\eta_{R_{l,n}}$ é o valor em η da fronteira externa do n -ésimo anel;

³A formulação aqui se limita ao caso de η pois engloba praticamente todas regiões com anomalias na granularidade (ver Tabela 4.1). Apenas as transições em $|\eta| = 2, 5|_-$ para $2, 5|_+$ (o que se inicia em $|\eta| = 2, 3$, para a janela atualmente empregada) apresentam anomalia na granularidade em ϕ .

$\eta_{(base, \Theta_{anom})}$ é o valor de η para a região anômala que pode assumir um dos valores de transição de granularidade descritos na Tabela 4.1.

Nos casos ocorrendo a), a solução pode ser obtida através de uma analogia. Considere-se que um y -ésimo grupo de células em Θ_{anom} , com o mesmo η , represente um baú com $b = k \times h_{\eta,l,max}$ bolas brancas — representando uma unidade quântica energética — , onde k é obtido por

$$k = \frac{\min(\{h_{(\eta,l,min)} \times k \in \mathbb{N}^+\})}{h_{\eta,l,min}}. \quad (9.7)$$

O anel pode ser apresentado por uma caixa na qual se deseja colocar $c = k \times h_{\eta,l,min}$ bolas, ou seja, distribuir igualmente a fração de energia para os anéis englobados no y -ésimo grupo de células. Nota-se que a nomenclatura baú é empregada para o repositório com maior capacidade de armazenamento, enquanto a caixa fica com menor capacidade.

Assim, cada bola branca representa uma operação de soma de energia a um dado anel, equivalente numericamente à fração ($f = \frac{1}{b}$) de energia ($E_{anom,y}$) do y -ésimo grupo de células no qual há maior interseção com o anel. Portanto, o problema análogo consiste-se em preencher todas as b caixas pertencentes a um ciclo com as bolas disponíveis nos c baús. Frisa-se que nesse problema o reservatório contendo as bolas é dado pelos baús.

Além disso, observa-se que, ao empregar o algoritmo descrito por (6.1), os anéis contendo o centro do y -ésimo grupo de células estarão preenchidos por $E_{anom,y}$, de forma que é necessário retirar a fração energética contida nesse anel que será repartida nos demais anéis. No problema análogo, isso significa que essas caixas estariam previamente preenchidas com b bolas brancas por um agente externo, que representa o algoritmo de anelamento. Sabe-se que esse agente preencheu as caixas que receberão a $v = \lceil b/2 \rceil$ -ésima bola contida no baú com b bolas, de forma que se é possível pintá-las de vermelho para sinalizar uma operação especial. A operação simbolizada é a retirada de b bolas brancas da caixa, e a adição de uma bola branca equivalente aquela que agora está pintada de vermelho que não entra na caixa.

A separação do problema em operações de retirada e adição facilita o preenchimento das matrizes correspondendo as correções. As correções equivalentes ao repartimento das frações de energias do grupo de células anômalos a serem repartidas entre os anéis ($\lambda_{y,+}$), ou, no problema análogo, a quantidade de bolas brancas a serem adicionadas

nas caixas para o y -ésimo baú dá-se por

$$\lambda_{y,+} = \left\{ \mathbf{0}^{1 \times \left\lceil \frac{by}{c} \right\rceil}; c - r_y; \underbrace{c; \dots; c}_{\left\lfloor m = \frac{b-c-r_y}{c} \right\rfloor \times}; \overbrace{b - ((m+1) \times c - r_y)}^{\lambda_{(y;m_0+m+1)}}; \mathbf{0}^{1 \times b - (m+m_0+1)} \right\}^T, \quad (9.8a)$$

$$r_{y+1} = c - \lambda_{(y;m_0+m+1)} = (m+2) \times c - (r_y + b), \quad (9.8b)$$

onde r_{y+1} é o número de bolas que precisam ser retirados do próximo baú para completar a última caixa preenchida para o y -ésimo baú e, analogamente, r_y as bolas retiradas do y -ésimo baú enquanto preenchendo as caixas com o baú anterior. Também, pela natureza do problema, $r_0 = 0$.

Ainda, as correções equivalentes à retirada de bolas para o y -ésimo baú são dadas por $\lambda_{(y,-)} = \mathbf{0}^{b \times 1}$, exceto por $\lambda_{y,x,-}$, caixa na posição $x_{y,-}$ que recebe a bola vermelha, conforme

$$x_y = \frac{v - r_y}{c}, \quad (9.9a)$$

$$\lambda_{y,x,-} = b. \quad (9.9b)$$

O total de bolas brancas a serem movidas em cada caixa é dado por (9.10), que possui valores negativos equivalente a uma operação de retorno de uma bola ao baú para que essa seja colocada em outra caixa.

$$\lambda_y = \lambda_{y,+} - \lambda_{y,-}. \quad (9.10)$$

Assim, escreve-se a matriz final de preenchimento das caixas por

$$\mathbf{\Lambda} = \left\{ \lambda_0; \lambda_1; \dots \lambda_b \right\}, \quad (9.11)$$

que representa a fração de energia a ser corrigida em cada anel pertencente a $\Theta_{l,anom}$. A propriedade $|\lambda_y| = 0$, consequência da conservação de energia, pode ser empregada para verificação da solução.

Finalmente, os valores de energia corrigidos nos anéis contidos no s -ésimo ciclo desse subcaso dá-se por

$$\mathbf{R}'_{l,s} = \mathbf{R}_{l,s} + \mathbf{\Lambda} \times \mathbf{E}_{anom,s}, \quad (9.12)$$

onde $\mathbf{E}_{anom,s}$ é um vetor de dimensão $b \times 1$, que contém a energia de todos os grupos de células, ordenados de forma crescente, no s -ésimo ciclo da região anômala, começando pelo grupo mais próximo ao centro. A mesma ordem é empregada para os anéis $\mathbf{R}_{l,s}$,

com dimensão $c \times 1$.

Por outro lado, os casos b) são inseridos como consequência da solução apresentada para o Caso 1, não estando presentes atualmente no algoritmo pois $h_{l,\eta}$ e $h_{l,\phi}$ atualmente adotados correspondem a menor granularidade disponível no calorímetro.

Mantendo a analogia proposta para o problema anterior, agora se necessita preencher os baús com as caixas. Dessa forma, as caixas contém c bolas e nos baús deseja-se colocar b bolas, estando disponíveis b caixas e um total de c baús, c e b sendo obtidos pela multiplicação da constante k em (9.7) como no subcaso **a**).

A solução para a quantidade de bolas a serem colocadas no y -ésimo baú dá-se por $\lambda'_{y,+} = \lambda_{y,+}^T$, onde $\lambda_{y,+}$ é dado por (9.8).

Contudo, o algoritmo já havia preenchido os anéis com as energias dos grupamentos de células, de forma que novamente é necessário realizar a retirada de energia desses anéis para equilibrar a energia repartida. Os casos em que as baús encontram-se preenchidos ocorrem para a $v = \lceil \frac{c}{2} \rceil$ -ésima bola sendo retirada da caixa, porém contendo c bolas. Uma forma simples para representar as operações de retirada $\lambda_{y,-}^T$ é obtida através de uma representação dependente em $\lambda_{y,+}^T$, conforme

$$\lambda_{(x,y,-)} = \begin{cases} 0 & \text{se } \lambda_{(x,y,+)} < \lceil \frac{c}{2} \rceil \\ c & \text{o.c.} \end{cases} \quad (9.13)$$

Basta, então, aplicar (9.10) para obter λ_y . A matriz Λ possui agora c vetores concatenados conforme descrito em (9.11), exceto atentando para esse detalhe. Finalmente, a correção aplicada dá-se por (9.12), onde os vetores $\mathbf{R}_{l,s}$ e $\mathbf{E}_{anom,s}$ tem as dimensões $b \times 1$ e $c \times 1$, respectivamente. Exemplos de matrizes de correção a serem aplicadas encontram-se na Tabela 9.3.

Solução 3.2 - Ponderação da energia anômala na base das células: Outra solução, que permite a construção de anéis com outras geometrias — como realizado para formação de anéis circulares em [196] (ver 6.2.1) — e também robusta a alterações de granularidade na região aplicando o algoritmo, consiste em realizar a ponderação da área da célula na região do anel. Nota-se que essa solução requer o mapeamento detalhado da área de cada uma das células compondo o calorímetro para o cálculo da ponderação célula a célula dentro da janela.

As correções propostas para o Caso 1 e Caso 2 são importantes especialmente para a redução do armazenamento em disco necessário para as análises de performance e ajuste da estratégia de seleção por anelamento⁴. Ao aplicar essas propostas, decorre naturalmente

⁴Existe um grande esforço por parte no ATLAS para a redução do tamanho em disco dos arquivos, contando com uma força tarefa exclusiva. Qualquer redução do número de grandezas empregas é de grande interesse para evitar que o algoritmo sofra com impedimentos técnicos envolvendo a realização de análises e

a necessidade de aplicação de classificadores especialistas para cada uma das subregiões com granularidade específica descritas na Tabela 4.1. Não se espera obter alteração na performance ao se adotar essas soluções nos espaços de entrada previamente alimentando classificadores especialistas, visto que a única alteração na representação do espaço de entrada se dá pela remoção das dimensões de anéis redundantes (constantes e nulos).

Por outro lado, a correção para o Caso 3 tem impacto direto na representatividade da base composta pelo perfil energético dos anéis. Espera-se, assim, que haja impacto na performance em decorrência de uma descrição mais fidedigna dessa base em relação ao desenvolvimento do chuveiro. É importante realçar que a análise dessa solução deve ser realizada somente para as partículas incidindo na região com transição, uma vez que apenas⁵ essa descrição sofrerá alteração. Ainda, a solução Solução 3.1 permite a uniformização da deposição de energia nos anéis com baixo custo computacional, já que permite o cálculo prévio da matriz Λ descrita em (9.11). Dois exemplos de Λ calculados estão disponíveis na Tabela 9.3.

Tabela 9.3: Exemplos de matrizes de correção de região anômala para a EM1 na alteração de granularidade ocorrendo em $|\eta| = 1,8$. Essas matrizes devem ser aplicadas para os anéis na região anômala $\Theta_{anom} = \{|\eta| > 1,8\}$ para o caso $h_\eta = 0,003$ e $\Theta_{anom} = \{|\eta| < 1,8\}$ quando $h_\eta = 0,004$, assumindo-se que a solução para o Caso 1 esteja sendo adotada. As matrizes são aplicadas a cada $\Delta\eta = ciclo_\eta = 0,012$, composta por 4 anéis para o primeiro caso e 3 anéis no segundo, até que todos anéis em Θ_{anom} tenham sido corrigidos.

	$h_\eta = 0,003$	$h_\eta = 0,004$
$\Lambda_{EM1, \eta =1,5,+}$	$\begin{pmatrix} 3 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix}$	$\begin{pmatrix} 3 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 1 & 3 \end{pmatrix}$
—	—	—
$\Lambda_{EM1, \eta =1,5,-}$	$\begin{pmatrix} 4 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix}$	$\begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$
=	=	=
$\Lambda_{EM1, \eta =1,5}$	$\begin{pmatrix} -1 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & -1 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & -1 & -1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$

ajuste dos classificadores. Isso se aplica para ambos ambientes Sistema de Filtragem do ATLAS e Sistema de Reconstrução (*Offline*) de Física do ATLAS.

⁵Obviamente, no intuito de validar o algoritmo, é necessário verificar se a descrição de outras regiões sofre alteração.

9.7 Descrição por Densidade de Energia

Em alternativa, acredita-se que uma melhor abordagem ao problema, para a aplicação no Sistema de Reconstrução (*Offline*) de Física do ATLAS, seja o emprego de uma outra representação distinta daquela descrita em (6.1). A ideia consiste em realizar o ajuste do perfil de energia com base nas amostragens discretas (empregando, por exemplo, KDE bidimensional) para cada partícula incidindo no calorímetro, obtendo, assim, a densidade de energia do chuveiro para cada uma das camadas do Sistema de Calorimetria. O espaço discreto a ser ajustado pela técnica pode empregar os valores calibrados de posição e energia da célula $(\eta; \phi; E)$, de forma a melhor representar a distribuição de deposição energética no calorímetro. Ao chegar no detalhamento que a densidade energética do chuveiro proporcionada, cabe estudar se as nuances obtidas pelo ajuste do chuveiro poderiam auxiliar na classificação.

Essa hipótese é feita com base no levantamento da literatura de calorimetria, sumarizada no Tópico 2.4.2.1. Observou-se que o desenvolvimento de chuveiros é dado por um processo estocástico, tanto EM quanto HAD. Notou-se que há maior flutuação para os chuveiros HAD, que se desenvolvem com maior irregularidade no seu formato e multiplicidade de partículas. Aumentando a complexidade do processo, os chuveiros HAD costumam conter componentes EM que também são descritas por processos estocásticos, com maior dependência nas primeiras interações ocorridas durante a formação do chuveiro. Não obstante, a resposta dos calorímetros do ATLAS não são iguais para chuveiros EM e HAD, aumentando as flutuações para medições de partículas HAD. A descrição por densidade de energia do chuveiro — em conjunto com a qualidade do ajuste (*gof*) ou outra descrição de desvio, para sumarizar flutuações não captadas no ajuste — possibilita acesso a essas flutuações⁶. Caso haja estatística suficiente para a caracterização delas, tal informação possibilitaria a discriminação entre um elétron e um hádron com mesma largura lateral e longitudinal, como nos casos da existência de maior assimetria na deposição de energia ou maior flutuação da energia a nível de célula. Isso adicionaria informação independente à largura e profundidade atualmente disponível na informação anelada em (6.1), além de não apresentar regiões anômalas. Obviamente, isso está limitado a capacidade dos calorímetros de representarem, com sua granularidade e amostragem longitudinal, tal irregularidade. Todavia, mesmo que eventualmente os casos de irregularidade não sejam capturados, ou que a irregularidade não seja uma particularidade dominante para hádrons nos eventos em que a informação lateral e longitudinal sejam insuficientes para discriminação, essa informação permite um acesso maior ao perfil do chuveiro. No trabalho [196], observaram-se resultados ligeiramente superiores para o emprego da informação no nível das células, o que leva a acreditar que essa base também seja mais discriminante que os anéis (Subseção 6.2.1). Por fim, ainda que não se deseje abordar o tema

⁶Outro método que também possibilitaria o acesso a flutuações seria aplicar o MSE em células acima de um limiar compondo os anéis.

no desenvolvimento deste trabalho, percebe-se que tal descrição possa ser extremamente útil para calibração da energia do *cluster* por permitir acessar assimetrias do perfil de deposição de energia, de interesse para essa correção.

Naturalmente, essa última abordagem é mais densa computacionalmente, bem como exige maior armazenamento em disco quando comparada com a atual, sendo razoável avaliá-la inicialmente para o Sistema de Reconstrução (*Offline*) de Física do ATLAS. Já as correções para o algoritmo atual usando a Solução 3.2 para o Caso 3, mesmo que provavelmente menos custosa que o ajuste do perfil do chuveiro, ainda exige a computação da ponderação para cada célula, provavelmente sendo mais indicada para uma futura implementação no estágio de precisão do HLT. A solução por interpolação, caso recupere a classificação de partículas nessa região, é plausível de ser aplicada no FastCalo com acesso apenas a informação de calorimetria, ao notar que as matrizes de correção para cada caso podem ser facilmente pré-computadas para cada um dos casos, exigindo somente a obtenção de E_{anom} para a aplicação das matrizes e obtenção dos valores corrigidos. Assim, acredita-se que seja possível a implementação dessa solução sem comprometer a aplicabilidade do algoritmo na etapa rápida do Sistema de Filtragem do ATLAS.

9.8 Adição de Informações de Outras Naturezas a partir do Sistema de Calorimetria

A informação temporal do pulso do sinal coletado nas células do calorímetro não é, atualmente, explorada. Essa informação, com a devida taxa de amostragem, é discriminativa e independente do formato do chuveiro [202]. Todavia, o ECAL do ATLAS possui taxa de amostragem bastante inferior à de [202], o que provavelmente impossibilita acessar esse nível de informação. Além disso, o acesso à informação do pulso amostrado para as células não é de fácil acesso na estrutura do Sistema de Reconstrução (*Offline*) de Física do ATLAS atual — ainda que seja possível obtê-la. Outra informação temporal que é de mais fácil acesso é o tempo de voo da partícula (ver Propriedade VI na p. 36), porém não se acredita que essa informação possa trazer ganhos uma vez que ambas partículas na escala de energia têm velocidade relativística elevada. Assim, com base na literatura, não se espera obter grandes ganhos na capacidade de discriminação ao explorar essa informação. Por outro lado, não há impedimento de uma rápida verificação dessa informação dependendo das demais prioridades.

9.9 Região do *Crack*

Atualmente, o algoritmo implementado no SR_{off} emprega as células provenientes da barra cintiladora ($0 < |\eta| < 1,6$) para a construção dos anéis, que é uma herança de sua primeira implementação por [197]⁷. Um estudo realizado por [198] mostrou uma leve melhora⁸ quando adicionada essa informação.

Durante a implementação do algoritmo para o SF_{on} , notou-se que a ordem de adição das células compondo o ITC estavam empregadas como se as mesmas representassem sucessivas amostragens longitudinais, ou seja: *TileGap1* era computado na camada HAD1; *TileGap2* em HAD2 e *TileGap3* em HAD3. Todavia, ao observar a Figura 4.8, percebe-se que essas células não representam informação longitudinal, em especial as células da camada cintiladora. Assim, fica evidente o fato das análises de [216] e [198] para a deposição de energia em função de η indicarem grande vazamento para a seção hadrônica na região do *crack* que se concentra na camada HAD3 — a mesma provém da camada cintiladora. Como não se encontrou qualquer menção da motivação para essa configuração, desde a primeira versão implementada (ver Subseção 7.3.4) para o SR_{off} empregou-se a configuração descrita na Tabela 6.1, onde as células da camada cintiladora pertencem à camada HAD1, a célula C10 à HAD2 e célula D3 à HAD3.

Uma avaliação detalhada para a região do *crack* do calorímetro, quando se tratando de identificação de elétrons, é de importância secundária, em decorrência dos estudos de física geralmente não empregarem essa informação na análise. Os elétrons nessa região, mesmo que corretamente identificados, estão em região com resposta bastante deteriorada que prejudicam a resolução dos resultados e por isso costumam serem descartados. Todavia, a barra cintiladora expande-se além da região do *crack* e cabe analisar se a adição das células E1 e E2 para a camada HAD1 não dificulta a formação de padrões. Para evitar a adição de informação não alterada na análise, o que obumbraria modificações na região afetada, o estudo deve restringir-se à região do espaço de fase afetada.

⁷Essa informação não consta explicita nesse trabalho, porém pode ser encontrada em [198, p. 228].

⁸Dentro da barra de erro, porém todos os dados foram empregados, gerando sobreposição estatística dos resultados. Uma análise empregando somente o espaço de fase sofrendo impacto pode fornecer uma melhor sensibilidade para as alterações de eficiência como será detalhado adiante no texto.

Capítulo 10

Conclusões

O campo de Física de Altas Energias fornece um ambiente propício para desenvolvimento e avaliação de técnicas de Inteligência Computacional, por envolver desafios como grande massas de dados, fusão de informação, ambiente de alta taxa de eventos, empilhamento de sinais, alta dimensionalidade e raros eventos de interesse.

O atual ciclo de coleta de dados no LHC é um exemplo desse cenário, onde a sempre crescente taxa de eventos fez com que a taxa de empilhamento de sinais alcançasse um nível em que técnicas de referência, empregadas por mais de 50 anos na área, atingissem limitações para atender às demandas de eficiência. Há, assim, uma necessidade por métodos mais eficientes. Esse foi o caso do Experimento ATLAS, que realizou uma série de atualizações em seu *trigger* para garantir o bom funcionamento durante a coleta de dados. Uma das propostas (*NeuralRinger*) para a redução de processamento na cadeia de elétrons, disponibilizada no ATLAS pela COPPE/UFRJ a partir de 1991, enfrentou resistência para seu emprego pela mesma não ser desenvolvida a partir da referência *offline* empregada nas análises realizadas como parte dos objetivos do experimento. Há receio, em casos como esse, em relação ao comportamento da técnica, uma vez que não há avaliação de seus efeitos e capacidades nas análises físicas. Nesse caso, é possível que o método possua algum efeito sistemático desconhecido, implicando, no pior cenário, em uma distorção na medição de propriedades ou incapacidade de observar determinados processos físicos. Outra questão que surge, no mesmo sentido, corresponde à harmonia dos dois sistemas após a intervenção no *trigger*, de forma que a mudança para o emprego de uma nova filtragem que deteriore o perfil dos elétrons selecionados pode implicar na perda de estatística total disponível para as análises.

A fim de lidar com a primeira limitação, este trabalho propôs o *NeuralRinger* para atuação no ambiente de análise *offline*. O desenvolvimento conforme realizado inverte a lógica da área, que parte das avaliações da capacidade da técnica para a análise física e então deteriora a mesma para atender as demandas do *trigger*. Como vantagem, o desenvolvimento a partir do *trigger* está em sintonia com a maturação das técnicas, que vão se tornando mais complexas conforme se adquire conhecimento no problema. O ambiente do *trigger* pode

conter nichos para a aplicação de novas ideias que agreguem para o experimento mesmo que as técnicas a partir dessas ideias ainda não tenham atingido o seu ápice para a aplicação em *offline*. Em decorrência disso, o maturamento da técnica permite que o *offline* usufrua das atualizações obtidas por parte das análises para a atuação no *trigger*, enquanto eventuais desenvolvimentos por parte das análises de sua atuação em *offline* nem sempre podem ser estendidos para atuação no *trigger*. Ainda assim, o aprofundamento do método para sua aplicação em *offline* é importante para garantir seu maior alinhamento com as necessidades finais do experimento. Desta forma, encontrar nichos no *trigger* pode ser um modo para tornar viável a maturação de técnicas nesse ambiente, todavia as necessidades do experimento não podem ser negligenciadas.

É natural, assim, que o desenvolvimento do Algoritmo *NeuralRinger* para atuação em *offline* absorva todas as atualizações realizadas para o *trigger*. Os importantes desenvolvimentos realizados para o *trigger* adicionados na sua atuação *offline* foram:

- emprego de um *ensemble* de redes neurais específicos por regiões no plano $E_T \times \eta$ a fim de aliviar a distorção nos padrões causados por essas variáveis e facilitar o processo de aprendizado;
- alteração da função de ativação na camada de saída para a operação da rede neural após o seu treinamento para o emprego de função linear, com o objetivo de evitar propagar não-linearidade adicional para a técnica de correção linear da eficiência em função de um estimador do empilhamento.

Cabe ressaltar que cada ambiente contém suas necessidades, o que resultou no emprego de um total de modelos diferentes para cada caso. Enquanto o *ensemble* para atuação no *trigger* foi definido para empregar 20 modelos, a versão *offline* desenvolvida contém as regiões (36) empregadas pela técnica de referência nesse ambiente a fim de facilitar a comparação entre os métodos.

Como parte dos desenvolvimentos específicos necessários para que a técnica seja viável no *offline*, está a adição da informação discriminante proveniente do ID. Diferente do *trigger*, onde *NeuralRinger* compõe uma solução híbrida com a complementação da decisão em estágio posterior com a aplicação de uma técnica de verossimilhança acessando grandezas físicas de ambos Sistemas de Calorimetria e do ID, o *offline* requer que a técnica atue como uma solução integral. O uso do ajuste de MLP especialistas para as representações das informações e posterior fusão é uma solução que permite expandir a solução disponível para a atuação no *offline* enquanto mantendo a técnica atuante no *trigger* como um caso específico daquele em análise no *offline*. A importância disso decorre por estes motivos:

- se o método empregado no *offline* não contiver o *trigger* como caso específico, a técnica proposta para o *offline* é um novo método e retornam-se às questões iniciais;

- uma das motivações para o uso de *ensembles* com métodos distintos dá-se pela complementação das lacunas das técnicas. Como as técnicas atuam em um sistema em série, o uso de métodos similares em ambos sistemas tendem a reduzir a exigência de lacunas distintas nos mesmos e, com isso, aumentando a eficiência total do sistema.

Consideraram-se diversos cenários onde o *ensemble* é alimentado por diferentes combinações das representações da informação disponível: soma em anéis concêntricos de energia, grandezas *shower* e *track*⁺. A fusão das *track*⁺ mostrou-se extremamente benéfica para o *NeuralRinger*, que obteve eficiência superior à atual referência, baseada em verossimilhança, em simulações, contendo condições de empilhamento compatíveis com colisões de 2016, ao reduzir a taxa de falso alarme de 2,54 % para 1,22 % (incerteza estatística desprezível) quando operando com taxa de detecção de elétrons similar para o critério *vloose*. A adição das grandezas *shower* possibilitou um acréscimo na performance ao reduzir a taxa de falso alarme para 1,13 %, onde se acrescenta informação discriminante não acessível quando empregando os anéis, como o caso da variável não linear ($w_{\eta 2}$), ou de razões de energia que não podem ser obtidas por manipulação dos anéis (E_{ratio} , R_{η} , R_{ϕ}). Os ganhos em eficiência distribuem-se com certa uniformidade por todo o plano $E_T \times \eta$, um comportamento interessante para evitar introduzir alterações nas análises físicas atuais, se não pela redução de impurezas nas mesmas. Não se observou grande alterações no quadrante ao adicionar as grandezas *shower*, onde há, em ambos casos, uma pequena tendência de coletar melhores amostras conforme as grandezas físicas. Esses resultados em simulação mostram um ótimo potencial para o método por não haver grandes mudanças no perfil dos elétrons selecionados, contudo com o benefício de maior pureza para a análise.

A obtenção desses resultados só foram possíveis devido ao desenvolvimento de uma infraestrutura completa para a atuação do *NeuralRinger* na reconstrução *offline*. Atualmente, estão disponíveis 10,5 PB de dados, correspondentes ao período de coleta entre 2015–2017, contendo a informação de calorimetria em anéis para candidatos a elétrons acima de 14 GeV. Os resultados mencionados são motivantes, porém ainda não são suficientes para a aplicação da técnica por não avaliar os efeitos e capacidades da técnica nas análises *offline*. A disponibilidade desses dados possibilitam a realização desses estudos, cuja especificação está na Seção 10.1. Isso mostra a importância do desenvolvimento para a atuação em *offline*, uma vez que permite um diálogo próximo com as necessidades das análises do experimento.

Outro aspecto a ser tratado é a avaliação dos impactos da intervenção no *trigger* nas análises físicas do experimento. Essa avaliação é um dos pontos centrais, onde se desenvolveram métodos estatísticos para avaliar alterações no perfil das grandezas físicas empregadas pela referência utilizada nas análises físicas. Os métodos permitiram elucidar que as distorções são pequenas ($< 1\sigma$) e ocorrem majoritariamente devido a pequenas diferenças na eficiência de detecção de elétrons entre a cadeia *ringer* e a sua duplicata aplicando o método substituído (baseado em cortes rígidos). Além disso, a avaliação direta dos casos que causam

distorção permitiu observar que, apesar de mínima, as distorções causam a coleta de amostras com melhor perfil conforme descrito pelas grandezas físicas empregadas para a tomada de decisão. Isso implica que essas amostras tendem, de acordo com as marginais, a serem aceitas pelo método sem a necessidade de um novo ajuste. Ainda assim, o resultado mais relevante ocorre justamente pelas distorções serem mínimas, o que torna simples a adição das estatísticas antes e após a intervenção durante a *Run 2*.

Esses resultados foram essenciais para possibilitar a adoção no *trigger* do ATLAS, em julho de 2017, do *NeuralRinger* para a cadeia de elétrons destinado à coleta de elétrons com $E_T > 15$ GeV, decorrendo no primeiro modelo neural atuando em filtragem *online* como método definitivo para a coleta de dados em grandes experimentos de Física de Altas Energias. Em valores estimados, a intervenção causou a redução de $\sim 1/4$ do processamento proveniente de cadeias de elétrons e fótons. Esse impacto foi ainda superior quando considerando que esse resultado contém incluso o processamento de cadeias de fótons e de elétrons com $E_T < 15$ GeV, que não sofreram alteração em sua estrutura. Um exemplo é a cadeia principal de elétrons de menor energia que sofreu uma redução na latência de 200 ms para 100 ms (valores centrais). A contribuição para a redução no processamento da fazenda foi um ponto importante dado que o mesmo tornou-se um gargalo para a tomada de dados durante a coleta de dados, sendo assim fundamental, em conjunto com demais aprimoramentos, para a boa campanha durante 2017 mesmo com o problema ocorrido em um setor no LHC que tornou o empilhamento ainda mais crítico. Assim, este trabalho fornece ferramental para o desenvolvimento e implementação bem-sucedida de métodos a partir do *trigger*.

10.1 Perspectivas e Trabalhos Futuros

A partir dos resultados obtidos neste trabalho, há grande motivação para continuar o desenvolvimento do método para aplicação nas análises físicas de elétrons e, possivelmente, outras partículas.

No caso de elétrons, ainda se faz necessário que o atual modelo de *ensembles* com fusão de informação de calorimetria e traços via redes especialistas, desenvolvido para a análise *offline*, precisa ter sua eficiência avaliada de forma a disponibilizar todos os parâmetros necessários para a sua aplicação nas análises físicas. Isso significa avaliar sua eficiência em colisão, a relação de eficiência do método em simulação e colisão, e a respectiva incerteza sistemática desse processo. Para que o mesmo possa ser aplicado de forma eficiente na busca por nova física, faz-se necessário cobrir toda a região cinemática. Especial atenção deve ser dada para possibilitar a aplicação na escala TeV, onde o desenvolvimento do chuveiro sofre alteração e, por isso, necessita-se especializar a técnica para essa região que conta com baixa estatística (todavia, aonde parte da nova física pode ser observada). Com a obtenção dos parâmetros na região atual de cinemática, pode ser possível adicionar a extração dos anéis

para a região de baixa cinemática (< 15 GeV).

Encontra-se em desenvolvimento, em outra tese de doutoramento do grupo brasileiro, a avaliação da tomada de decisão por verossimilhança que pode responder questões interessantes sobre a informação anelada.

Como parte do pós-doutoramento do autor deste trabalho em uma colaboração da COPPE/UFRJ com a Universidade *Pierre-et-Marie-Curie*, na França, pretende-se explorar anéis axiais (descrição dos anéis limitada pelos eixos no plano $\eta \times \phi$) para a calibração em energia de elétrons a fim de obter, com essa descrição, acesso à assimetria na deposição de energia. Outro tópico que se explorará está envolvida diretamente na análise física, onde se realizará a busca por Partículas Compatíveis com Axions (ALP) [255] através de decaimentos em dois fótons. Em um dos cenários esperados, o decaimento a ser estudado pode ocorrer somente quanto a mesma já se encontra transpassando o Sistema de Calorimetria, podendo ocorrer tanto nas camadas longitudinais mais internas ou externas. Assim, o desenvolvimento do chuveiro de partículas ocorrerá com uma estrutura distinta de fótons oriundos de decaimentos ocorrendo aproximadamente no ponto de colisão, havendo a necessidade de preparar o modelo para esse cenário. Os anéis fornecem uma estrutura ideal para essa análise, uma vez que os mesmos são construídos sequencialmente em cada camada longitudinal, sendo possível a extração de modelos dedicados para cada um dos possíveis cenários de decaimentos. Com isso, pretende-se contribuir para o desenvolvimento a análise física de um potencial candidato a descrição de matéria escura.

Um trabalho de simples implementação e que pode possibilitar melhor eficiência se dá por alterar a estrutura do *ensemble* para aplicar redes especialistas por camadas ou seções do calorímetro. Também no âmbito de aprimorar o *ensemble*, pode-se realizar a otimização do tamanho da janela do algoritmo anelador para as regiões do espaço de fase. Na mesma linha, atualmente a grade para atuação do *ensemble* foi definida conforme as fronteiras geométricas do ATLAS e por limites empíricos de faixa de energia. Os SOM ajustados neste trabalho mostraram capacidade de capturar, de forma não-supervisionada, a influência dessas variáveis no mapa, o que pode indicar uma maneira de definir a estrutura do *ensemble* através da minimização das mesmas no SOM.

Atualmente, encontra-se em desenvolvimento um novo ajuste do *NeuralRinger* atuando no SF_{on} para atuação em 2018 a partir de dados de colisão, onde se espera obter melhores eficiências. Em paralelo, desenvolve-se o ajuste de modelos para elétrons de baixa energia (via T&P para J/Ψ) para possibilitar a integração de cadeias *ringer* de baixa energia ao menu.

A atuação do método no ambiente *offline* exige reavaliar as limitações e possíveis aprimoramentos no método, que fizeram parte das contribuições deste trabalho. A aplicação desses aprimoramentos causaria maiores divergências entre os métodos nos dois ambientes, o que faria com que o desenvolvimento *offline* perdesse seu foco de preencher as lacunas do desen-

volvimento a partir do *trigger*. As análises estão próximas de atingir seu êxito, contendo todo o ferramental e dados necessários, como mencionado no início dessa seção. Uma vez atingido esse patamar, há espaço para a aplicação dos aprimoramentos e superação das limitações desenvolvidos neste trabalho. Para o caso do tratamento de regiões anômalas, necessita-se implementar as correções propostas e realizar o estudo do impacto dessas correções no *NeuralRinger*. As mesmas podem ser especialmente importantes no caso da aplicação da descrição por anéis axiais para a calibração de energia. O desenvolvimento do *NeuralRinger* ocorre em grande massas de dados, de forma que a estrutura desenvolvida para avaliar o impacto pode ser empregada para obter amostras com informação distinta.

Trabalhos futuros podem empregar uma abordagem por agrupamento, como SOM para realizar a redução da estatística. No caso do SOM, cabe ajustar o tamanho do mapa e o processo de treinamento a fim de buscar melhores ajustes dos mapas nos dados. Outra maneira de realizar a compressão dos dados a ser explorada é via PCD, dado que se deseja trabalhar com a informação discriminante nos dados e que mostrou, em trabalhos anteriores [198], ser benéfico para o caso do Algoritmo *NeuralRinger*. No caso da informação reduzida, o balanceamento via homogeneização dos subconjuntos de validação cruzada deverá ser considerado para evitar polarização do discriminador. Outro ponto a continuar o desenvolvimento é a busca por um espaço latente mais discriminante via SAE e ICA. A normalização especialista para obtenção de um espaço de características mais discriminante e com ótimos resultados, conforme mencionado no levantamento de trabalhos anteriores (Seção 6.2), também pode ser explorada. Esse é o caso do mapeamento-Rp, anéis-Rp e normalização sequencial. No caso da otimização parâmetros α e β para os dois primeiros casos, a otimização considerando os mesmos como uma função dos estimadores do empilhamento pode possibilitar melhor eficiência. Já no caso da normalização sequencial, é necessário o ajuste de seus parâmetros que poderiam contar com Técnicas de Enxames ou Computação Evolutiva.

Outras representações de informação podem ser exploradas para identificação de elétrons. Há motivação por parte de trabalhos anteriores para explorar a informação de calorimetria diretamente no nível de célula. Inclusive, o ATLAS está motivado à empregar essa descrição diretamente para alimentar técnicas de Aprendizado de Máquinas. Outra descrição interessante a ser explorada, idealizada neste trabalho, é a descrição por densidade de energia, que permitiria lidar com as regiões anômalas do Sistema de Calorimetria do ATLAS. Além disso, a informação do desenvolvimento temporal do pulso do sinal nos canais do calorímetro podem adicionar informação discriminante que pode ser explorada. Uma estratégia para possibilitar o acesso a essas informações para o ajuste de classificadores pode se dar através do emprego de SAE para a compactação dessa informação em disco.

Avaliações de aprimoramentos, novas técnicas e novas representações de informação, que se mostrem eficientes e com comportamento mais adequado para atender as demandas das análises do experimento, podem ser adaptadas para a sua aplicação no *trigger*. Neste caso,

o ciclo de desenvolvimento seria um processo de duas vias, iniciando-se em um nicho de aplicação no *trigger*, desenvolvendo-se para a atuação no *offline*, enquanto absorvendo os aprimoramentos do *trigger*. Como os ambientes de desenvolvimento encontram-se integrados, os desenvolvimentos do *offline* podem ser avaliados, e eventualmente degradados, para a atuação no *trigger*. Acredita-se que o ciclo de desenvolvimento tenderá, naturalmente, após a maturação do método para atuação em *offline*, para o padrão da área. Um exemplo seria o desenvolvimento do HLT empregando o ajuste de redes especialistas caso o *NeuralRinger* se adeque para atuar como referência nas análises físicas.

Referências Bibliográficas

- [1] MOUNT, R. P. “Off-line computing for experimental high-energy physics”, *Reports on Progress in Physics*, v. 55, n. 9, pp. 1385, 1992. Disponível em: <<http://stacks.iop.org/0034-4885/55/i=9/a=001>>. Acesso em: 21 de abril de 2018. 8, 76
- [2] ENCICLOPÉDIA BRITANNICA. “Biografia de Demócritos de Abdera”. Disponível em: <<https://www.britannica.com/biography/Democritus>>. Acesso em: 21 de abril de 2018. 9
- [3] NOBEL MEDIA AB. “Biografia de Thompson - Prêmio Nobel”. Disponível em: <http://nobelprize.org/nobel_prizes/physics/laureates/1906/thomson-bio.html>. Acesso em: 03 de março de 2017. 9
- [4] MARTIN, B. R. *Nuclear and particle physics: an introduction*. 2 ed. Hoboken, NJ, Wiley, 2009. 9, 12
- [5] COTTINGHAM, W. N., GREENWOOD, D. A. *An Introduction to the Standard Model of Particle Physics*. 2 ed. Cambridge, Cambridge Univ. Press, 2007. 9, 11, 12, 13
- [6] TULLY, C. C. *Elementary particle physics in a nutshell*. Princeton, NJ, Princeton Univ. Press, 2011. Disponível em: <<https://cds.cern.ch/record/1417476>>. Acesso em: 21 de abril de 2018. 20, 22, 24, 33, 34, 47
- [7] WHALEN, K. *From discovery to precision measurements: electron identification, electron-like backgrounds, and measurement of the differential fiducial cross-sections of the Higgs boson in the four-lepton decay channel with the ATLAS detector*. Tese de Doutorado, Carleton U., Jan 2015. Disponível em: <<https://cds.cern.ch/record/1995030>>. Acesso em: 21 de abril de 2018. Apresentado em: 19 de Dezembro de 2014. 9, 95
- [8] D0 COLLABORATION. “Observation of the top quark”, *Phys. Rev. Lett.*, v. 74, pp. 2632–2637, 1995. doi: 10.1103/PhysRevLett.74.2632. 9, 48

- [9] CDF COLLABORATION. “Observation of top quark production in $\bar{p}p$ collisions”, *Phys. Rev. Lett.*, v. 74, pp. 2626–2631, 1995. doi: 10.1103/PhysRevLett.74.2626. 9, 48
- [10] ATLAS COLLABORATION. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”, *Phys. Lett. B*, v. 716, pp. 1, 2012. doi: 10.1016/j.physletb.2012.08.020. 9, 12, 13, 18, 19, 48, 128
- [11] CMS COLLABORATION. “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”, *Physics Letters B*, v. 716, n. 1, pp. 30 – 61, 2012. ISSN: 0370-2693. doi: <http://dx.doi.org/10.1016/j.physletb.2012.08.021>. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0370269312008581>>. Acesso em: 21 de abril de 2018. 9, 12, 13, 18, 48
- [12] PATRIGNANI, C., et al. “Review of Particle Physics”, *Chin. Phys.*, v. C40, n. 10, pp. 100001, 2017. doi: 10.1088/1674-1137/40/10/100001. 10, 11, 25, 26, 27, 28, 31, 34, 38, 39, 40, 41, 43, 44, 46, 67
- [13] BEISER, A. *Concepts of Modern Physics*. 6 ed. New York, NY, McGraw-Hill, 2003. 11, 13
- [14] HIGGS, P. “Broken symmetries, massless particles and gauge fields”, *Physics Letters*, v. 12, n. 2, pp. 132 – 133, 1964. ISSN: 0031-9163. doi: [http://dx.doi.org/10.1016/0031-9163\(64\)91136-9](http://dx.doi.org/10.1016/0031-9163(64)91136-9). Disponível em: <<http://www.sciencedirect.com/science/article/pii/0031916364911369>>. Acesso em: 21 de abril de 2018. 12
- [15] HIGGS, P. W. “Broken Symmetries and the Masses of Gauge Bosons”, *Phys. Rev. Lett.*, v. 13, pp. 508–509, Oct 1964. doi: 10.1103/PhysRevLett.13.508. Disponível em: <<http://link.aps.org/doi/10.1103/PhysRevLett.13.508>>. Acesso em: 21 de abril de 2018. 12
- [16] ENGLERT, F., BROUT, R. “Broken Symmetry and the Mass of Gauge Vector Mesons”, *Phys. Rev. Lett.*, v. 13, pp. 321–323, Aug 1964. doi: 10.1103/PhysRevLett.13.321. Disponível em: <<http://link.aps.org/doi/10.1103/PhysRevLett.13.321>>. Acesso em: 21 de abril de 2018. 12
- [17] GURALNIK, G. S., HAGEN, C. R., KIBBLE, T. W. B. “Global Conservation Laws and Massless Particles”, *Phys. Rev. Lett.*, v. 13, pp. 585–587, Nov 1964. doi: 10.1103/PhysRevLett.13.585. Disponível em: <<http://link.aps.org/doi/10.1103/PhysRevLett.13.585>>. Acesso em: 21 de abril de 2018. 12
- [18] CERN. “Portal oficial do CERN”. Disponível em: <<http://cern.ch/>>. Acesso em: 05 de março de 2017. 12, 13

- [19] EVANS, L., BRYANT, P. “LHC Machine”, *JINST*, v. 3, pp. S08001, 2008. doi: 10.1088/1748-0221/3/08/S08001. 13, 14, 19, 20
- [20] APOLLINARI, G., BÉJAR ALONSO, I., BRÜNING, O., et al. *High-Luminosity Large Hadron Collider (HL-LHC): Preliminary Design Report*. CERN Yellow Reports: Monographs. Geneva, CERN, 2015. doi: 10.5170/CERN-2015-005. Disponível em: <<https://cds.cern.ch/record/2116337>>. Acesso em: 21 de abril de 2018. 13
- [21] LHC. “Portal oficial de comissionamento do LHC”. . Disponível em: <<https://lhc-commissioning.web.cern.ch/lhc-commissioning/schedule/LHC-long-term.htm>>. Acesso em: 05 de março de 2017. 13
- [22] NOBEL MEDIA AB. “Prêmio Nobel de Física em 2013”. 2014. Disponível em: <http://www.nobelprize.org/nobel_prizes/physics/laureates/2013/>. Acesso em: 21 de abril de 2018. 13
- [23] LHC. “Portal oficial de divulgação do LHC”. . Disponível em: <<https://lhc-machine-outreach.web.cern.ch/>>. Acesso em: 06 de março de 2017. 14, 16
- [24] MARCASTEL, F. “CERN’s Accelerator Complex. La chaîne des accélérateurs du CERN”, Oct 2013. Disponível em: <<https://cds.cern.ch/record/1621583>>. Acesso em: 21 de abril de 2018. 15
- [25] CMS COLLABORATION. “The CMS experiment at the CERN LHC”, *Journal of Instrumentation*, v. 3, n. 08, pp. S08004, 2008. Disponível em: <<http://stacks.iop.org/1748-0221/3/i=08/a=S08004>>. Acesso em: 21 de abril de 2018. 15
- [26] LHCb COLLABORATION. “The LHCb Detector at the LHC”, *Journal of Instrumentation*, v. 3, n. 08, pp. S08005, 2008. Disponível em: <<http://stacks.iop.org/1748-0221/3/i=08/a=S08005>>. Acesso em: 21 de abril de 2018. 15
- [27] ALICE COLLABORATION. “The ALICE experiment at the CERN LHC”, *JINST*, v. 3, pp. S08002, 2008. doi: 10.1088/1748-0221/3/08/S08002. 15
- [28] ATLAS COLLABORATION. “Portal oficial do ATLAS”. . Disponível em: <<http://atlas.cern>>. Acesso em: 06 de março de 2017. 18, 59
- [29] ACOSTA, D., AMBROGLINI, F., BARTALINI, P., et al. *The Underlying Event at the LHC*. Relatório Técnico CMS-NOTE-2006-067. CERN-CMS-NOTE-2006-067, CERN, Geneva, June 2006. 17

- [30] FIELD, R. “Min-Bias and the Underlying Event at the LHC”. In: *Physics in Collisions*, Vancouver, British Columbia, August 2011. Disponível em: <http://www.phys.ufl.edu/~rfield/cdf/PIC2011_RickField_8-29-11.ppt>. Acesso em: 21 de abril de 2018. 19
- [31] ATLAS COLLABORATION. *ATLAS detector and physics performance: Technical Design Report, 2*. Relatório técnico, Geneva, 1999. Disponível em: <<https://cds.cern.ch/record/391177>>. Acesso em: 21 de abril de 2018. 18, 59
- [32] ATLAS COLLABORATION. “Resultados públicos do Experimento ATLAS para o Modelo Padrão”. . Disponível em: <<https://twiki.cern.ch/twiki/bin/view/AtlasPublic/StandardModelPublicResults>>. Acesso em: 06 de março de 2017. 20, 21
- [33] ATLAS COLLABORATION. “Measurements of $W\gamma$ and $Z\gamma$ production in pp collisions at $\sqrt{s}=7$ TeV with the ATLAS detector at the LHC”, *Phys. Rev. D*, v. 87, pp. 112003, Jun 2013. doi: 10.1103/PhysRevD.87.112003. Disponível em: <<http://link.aps.org/doi/10.1103/PhysRevD.87.112003>>. Acesso em: 21 de abril de 2018. 21
- [34] BELOV, P., BRITZGER, D., CAMARDA, S., et al. “Parton distribution functions at LO, NLO and NNLO with correlated uncertainties between orders”, *The European Physical Journal C*, v. 74, n. 10, pp. 3039, Sep 2014. ISSN: 1434-6052. doi: 10.1140/epjc/s10052-014-3039-4. Disponível em: <<https://doi.org/10.1140/epjc/s10052-014-3039-4>>. Acesso em: 03 de março de 2018. 20
- [35] “Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC”, *Eur. Phys. J. C*, v. 76, n. arXiv:1608.03953. CERN-EP-2016-117. 12, pp. 653. 71 p, Aug 2016. Disponível em: <<https://cds.cern.ch/record/2208146>>. Acesso em: 21 de abril de 2018. Todas figuras incluindo as auxiliares estão disponíveis em <http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/DAPR-2013-01>. 20, 22, 95
- [36] ATLAS COLLABORATION. “Resultados públicos de luminosidade do Experimento ATLAS para a *Run 2*”. . Disponível em: <<https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2>>. Acesso em: 06 de março de 2017. 22
- [37] ATLAS COLLABORATION. “Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV Using the ATLAS Detector at the LHC”, *Eur. Phys. J.*, v. C71, pp. 1630, 2011. doi: 10.1140/epjc/s10052-011-1630-5. 23

- [38] ATLAS COLLABORATION. “Resultados públicos de luminosidade do Experimento ATLAS para a *Run 1*” . . Disponível em: <<https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults>>. Acesso em: 06 de março de 2017. 23
- [39] GRUPEN, C., SHWARTZ, B. *Particle Detectors*. Cambridge monographs on particle physics, nuclear physics, and cosmology. 2 ed. Cambridge, Cambridge Univ. Press, 2008. Disponível em: <<https://cds.cern.ch/record/1110991>>. Acesso em: 21 de abril de 2018. 25, 26, 27, 28, 29, 32, 33, 34, 36, 37, 38, 39, 40, 43, 44, 45, 47, 106
- [40] KNOLL, G. F. *Radiation Detection and Measurement*. 3 ed. New York, NY, Wiley, 2000. Disponível em: <<https://cds.cern.ch/record/441925>>. Acesso em: 21 de abril de 2018. 25, 29, 143
- [41] WIGMANS, R. *Calorimetry: energy measurement in particle physics*. 2 ed. Oxford, Oxford University Press, 2017. 25, 34, 125
- [42] AGOSTINELLI, S., ALLISON, J., AMAKO, K., et al. “Geant4—a simulation toolkit”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, v. 506, n. 3, pp. 250 – 303, 2003. ISSN: 0168-9002. doi: [https://doi.org/10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8). Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168900203013688>>. Acesso em: 21 de abril de 2018. 25
- [43] PARTICLE PHYSICS GROUP. “Permissão para uso das figuras do *Particle Physics Group*” . . Disponível em: <<http://pdg.lbl.gov/2017/figures/figures.html>>. Acesso em: 03 de março de 2018. 26
- [44] GRUPEN, C. “Electromagnetic Interactions of High Energy Cosmic Ray Muons”, *Fortschritte der Physik*, v. 24, n. 3, pp. 127–209, 1976. ISSN: 1521-3979. doi: 10.1002/prop.19760240302. Disponível em: <<http://dx.doi.org/10.1002/prop.19760240302>>. Acesso em: 21 de abril de 2018. 27
- [45] GERHARDT, L., KLEIN, S. R. “Electron and photon interactions in the regime of strong Landau-Pomeranchuk-Migdal suppression”, *Phys. Rev. D*, v. 82, pp. 074017, Oct 2010. doi: 10.1103/PhysRevD.82.074017. Disponível em: <<http://link.aps.org/doi/10.1103/PhysRevD.82.074017>>. Acesso em: 21 de abril de 2018. 28
- [46] WIGMANS, R. “Advances in hadron calorimetry”, *Annu. Rev. Nucl. Part. Sci.*, v. 41, n. CERN-PPE-91-39, pp. 133–185. 69 p, Feb 1991. Disponível em: <<http://cds.cern.ch/record/1110991>>. Acesso em: 21 de abril de 2018. 29

cern.ch/record/217893>. Acesso em: 21 de abril de 2018. 32, 33, 34, 35, 39, 40, 42, 44, 45, 46

- [47] FABJAN, C. W., GIANOTTI, F. “Calorimetry for Particle Physics”, *Rev. Mod. Phys.*, v. 75, n. CERN-EP-2003-075, pp. 1243–1286. 96 p, Oct 2003. Disponível em: <<https://cds.cern.ch/record/692252>>. Acesso em: 21 de abril de 2018. 33, 34, 35, 38, 39, 40, 41, 42, 46
- [48] LONGO, E., SESTILI, I. “Monte Carlo calculation of photon-initiated electromagnetic showers in lead glass”, *Nuclear Instruments and Methods*, v. 128, n. 2, pp. 283 – 307, 1975. ISSN: 0029-554X. doi: [http://dx.doi.org/10.1016/0029-554X\(75\)90679-5](http://dx.doi.org/10.1016/0029-554X(75)90679-5). Disponível em: <<http://www.sciencedirect.com/science/article/pii/0029554X75906795>>. Acesso em: 21 de abril de 2018. 37
- [49] NELSON, W. R., HIRAYAMA, H., ROGERS, D. W. O. *The EGS4 code system*. Stanford, CA, SLAC, 1985. Disponível em: <<http://cds.cern.ch/record/107578>>. Acesso em: 21 de abril de 2018. 37, 38
- [50] PUZO, P. “ATLAS calorimetry”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, v. 494, n. 1–3, pp. 340 – 345, 2002. ISSN: 0168-9002. doi: [http://dx.doi.org/10.1016/S0168-9002\(02\)01490-0](http://dx.doi.org/10.1016/S0168-9002(02)01490-0). Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168900202014900>>. Acesso em: 21 de abril de 2018. Anais do *8th International Conference on Instrumentation for Colliding Beam Physics*. 43
- [51] PARTICLE PHYSICS GROUP. “Propriedades atômico-nucleares de materiais para mais de 300 materiais”. . Disponível em: <<http://pdg.lbl.gov/2017/AtomicNuclearProperties/>>. Acesso em: 03 de março de 2017. 44
- [52] ATLAS COLLABORATION. *ATLAS detector and physics performance: Technical Design Report, 1*. Technical Design Report ATLAS. Geneva, CERN, 1999. Disponível em: <<https://cds.cern.ch/record/391176>>. Acesso em: 21 de abril de 2018. 46, 59, 61, 65
- [53] TIMMER, J. “The UA1 detector”, 1983. Disponível em: <<https://cds.cern.ch/record/149406>>. Acesso em: 21 de abril de 2018. 47
- [54] CDHS COLLABORATION. “Página oficial do experimento de neutrinos CDHS”. Disponível em: <<http://knobloch.home.cern.ch/knobloch/cdhs/cdhs.html>>. Acesso em: 23 de fevereiro de 2017. Também conhecido por WA1. 47

- [55] DIDDENS, A., JONKER, M., PANMAN, J., et al. “A detector for neutral-current interactions of high-energy neutrinos”, *Nuclear Instruments and Methods*, v. 178, n. 1, pp. 27 – 48, 1980. ISSN: 0029-554X. doi: [http://dx.doi.org/10.1016/0029-554X\(80\)90854-X](http://dx.doi.org/10.1016/0029-554X(80)90854-X). Disponível em: <<http://www.sciencedirect.com/science/article/pii/0029554X8090854X>>. Acesso em: 21 de abril de 2018. 47
- [56] *ATLAS Calorimeter Performance: Technical Design Report*. Technical Design Report ATLAS. Geneva, CERN, 1996. 47, 65, 68, 69, 71, 73, 75
- [57] PETERSON, C., ROGNVALDSSON, T., LONNBLAD, L. “JETNET 3.0: A Versatile artificial neural network package”, *Comput. Phys. Commun.*, v. 81, pp. 185–220, 1994. doi: 10.1016/0010-4655(94)90120-1. 48, 114
- [58] DENBY, B. H. “Tutorial on neural network applications in high-energy physics: A 1992 perspective”. In: *2nd International Workshop on Software Engineering, Artificial Intelligence and Expert Systems for High-energy and Nuclear Physics (AIHENP 92) La Londe Les Maures, France, January 13-18, 1992*, pp. 287–325, 1992. Disponível em: <http://lss.fnal.gov/cgi-bin/find_paper.pl?conf-92-121>. Acesso em: 21 de abril de 2018. 48
- [59] CDF COLLABORATION. “Measurement of the top quark mass with dilepton events selected using neuroevolution at CDF”, *Phys. Rev. Lett.*, v. 102, pp. 152001, 2009. doi: 10.1103/PhysRevLett.102.152001. 48
- [60] CDF COLLABORATION. “Top Quark Mass Measurement in the Lepton plus Jets Channel Using a Modified Matrix Element Method”, *Phys. Rev.*, v. D79, pp. 072001, 2009. doi: 10.1103/PhysRevD.79.072001.
- [61] CDF COLLABORATION. “Measurement of the single top quark production cross section and $|V_{tb}|$ in 1.96 TeV $p\bar{p}$ collisions with missing transverse energy and jets and final CDF combination”, *Phys. Rev. D*, v. 93, pp. 032011, Feb 2016. doi: 10.1103/PhysRevD.93.032011. Disponível em: <<http://link.aps.org/doi/10.1103/PhysRevD.93.032011>>. Acesso em: 21 de abril de 2018.
- [62] AALTONEN, T. A., et al. “Search for the production of ZW and ZZ boson pairs decaying into charged leptons and jets in $p\bar{p}$ collisions at $\sqrt{s}=1.96$ TeV”, *Phys. Rev.*, v. D88, n. 9, pp. 092002, 2013. doi: 10.1103/PhysRevD.88.092002.
- [63] CDF COLLABORATION. “Measurement of the top quark mass in the all-hadronic mode at CDF”, *Physics Letters B*, v. 714, n. 1, pp. 24 – 31, 2012. ISSN: 0370-2693. doi: <http://dx.doi.org/10.1016/j.physletb.2012.06.007>. Disponível em: <<http://>>

www.sciencedirect.com/science/article/pii/S0370269312006338>. Acesso em: 21 de abril de 2018.

- [64] CDF COLLABORATION. “Search for the standard model Higgs boson produced in association with a W^\pm boson with 7.5 fb^{-1} integrated luminosity at CDF”, *Phys. Rev.*, v. D86, pp. 032011, 2012. doi: 10.1103/PhysRevD.86.032011.
- [65] FREEMAN, J., JUNK, T., KIRBY, M., et al. “Introduction to HOBIT, a b-Jet Identification Tagger at the CDF Experiment Optimized for Light Higgs Boson Searches”, *Nucl. Instrum. Meth.*, v. A697, pp. 64–76, 2013. doi: 10.1016/j.nima.2012.09.021.
- [66] CDF COLLABORATION. “Measurement of the $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96 \text{ TeV}$ using events with large Missing E_T and jets”, *Phys. Rev.*, v. D84, pp. 032003, 2011. doi: 10.1103/PhysRevD.84.032003.
- [67] CDF COLLABORATION. “Search for Standard Model Higgs Boson Production in Association with a W Boson at CDF”, *Phys. Rev.*, v. D85, pp. 052002, 2012. doi: 10.1103/PhysRevD.85.052002.
- [68] FREEMAN, J., LEWIS, J. D., KETCHUM, W., et al. “An Artificial neural network based b jet identification algorithm at the CDF Experiment”, *Nucl. Instrum. Meth.*, v. A663, pp. 37–47, 2012. doi: 10.1016/j.nima.2011.10.024.
- [69] CDF COLLABORATION. “Search for single top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96 \text{ TeV}$ in the missing transverse energy plus jets topology”, *Phys. Rev.*, v. D81, pp. 072003, 2010. doi: 10.1103/PhysRevD.81.072003.
- [70] CDF COLLABORATION. “Observation of Single Top Quark Production and Measurement of $|\text{V}_{tb}|$ with CDF”, *Phys. Rev.*, v. D82, pp. 112005, 2010. doi: 10.1103/PhysRevD.82.112005.
- [71] CDF COLLABORATION. “A Search for the Higgs Boson Using Neural Networks in Events with Missing Energy and b-quark Jets in p anti-p Collisions at $\sqrt{s} = 1.96\text{-TeV}$ ”, *Phys. Rev. Lett.*, v. 104, pp. 141801, 2010. doi: 10.1103/PhysRevLett.104.141801.
- [72] CDF COLLABORATION. “New Suppressed Decays of B_s^0 Mesons at CDF”, *PoS*, v. ICHEP2010, pp. 211, 2010.
- [73] CDF COLLABORATION. “Search for WW and WZ Production in Lepton Plus Jets Final State at CDF”, *Phys. Rev.*, v. D79, pp. 112011, 2009. doi: 10.1103/PhysRevD.79.112011.

- [74] CDF COLLABORATION. “Measurement of the fraction of $t\bar{t}$ production via gluon-gluon fusion in $p\bar{p}$ collisions at $\sqrt{s} = 1.96\text{-TeV}$ ”, *Phys. Rev.*, v. D79, pp. 031101, 2009. doi: 10.1103/PhysRevD.79.031101.
- [75] CDF COLLABORATION. “Search for top-quark production via flavor-changing neutral currents in W+1 jet events at CDF”, *Phys. Rev. Lett.*, v. 102, pp. 151801, 2009. doi: 10.1103/PhysRevLett.102.151801.
- [76] CDF COLLABORATION. “Search for Standard Model Higgs Boson Production in Association with a W Boson using a Neural Network Discriminant at CDF”, *Phys. Rev.*, v. D80, pp. 012002, 2009. doi: 10.1103/PhysRevD.80.012002.
- [77] CDF COLLABORATION. “Search for the Higgs boson produced with $Z \rightarrow \ell^+\ell^-$ in $p\bar{p}$ collisions at $\sqrt{s} = 1.96\text{ TeV}$ ”, *Phys. Rev. Lett.*, v. 101, pp. 251803, 2008. doi: 10.1103/PhysRevLett.101.251803.
- [78] CDF COLLABORATION. “Search for SM Higgs in the $WH \rightarrow \ell\nu b\bar{b}$ Channel using $2fb^{-1}$ ”. In: *Proceedings, 43rd Rencontres de Moriond on Electroweak Interactions and Unified Theories: La Thuile, Italy, March 1-8, 2008*, 2008. Disponível em: <<https://inspirehep.net/record/786158/files/arXiv:0805.2803.pdf>>. Acesso em: 21 de abril de 2018.
- [79] CDF COLLABORATION. “Search for Hadronic Decays of W and Z Bosons in Photon Events in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96\text{-GeV}$ ”, *Phys. Rev.*, v. D80, pp. 052011, 2009. doi: 10.1103/PhysRevD.80.052011.
- [80] CDF COLLABORATION. “Top Quark Mass Measurement Using a Matrix Element Method with Quasi-Monte Carlo Integration”. In: *Proceedings, 34th International Conference on High Energy Physics (ICHEP 2008): Philadelphia, Pennsylvania, July 30-August 5, 2008*, 2008. Disponível em: <http://lss.fnal.gov/cgi-bin/find_paper.pl?conf-08-468>. Acesso em: 21 de abril de 2018.
- [81] CDF COLLABORATION. “Measurement of the Single Top Quark Production Cross Section at CDF”, *Phys. Rev. Lett.*, v. 101, pp. 252001, 2008. doi: 10.1103/PhysRevLett.101.252001.
- [82] CDF COLLABORATION. “Search for Technicolor Particles Produced in Association with W Boson at CDF”. In: *Hadron collider physics. Proceedings, 19th Symposium, HCP2008, Galena, USA, May 27-31, 2008*, 2008. Disponível em: <http://lss.fnal.gov/cgi-bin/find_paper.pl?conf-08-278>. Acesso em: 21 de abril de 2018.

- [83] CDF COLLABORATION. “Search for Standard Model Higgs Boson Production in Association with a W Boson at CDF”, *Phys. Rev.*, v. D78, pp. 032008, 2008. doi: 10.1103/PhysRevD.78.032008.
- [84] CDF COLLABORATION. “Measurement of the $p\bar{p} \rightarrow t\bar{t}$ production cross-section and the top quark mass at $\sqrt{s} = 1.96$ -TeV in the all-hadronic decay mode”, *Phys. Rev.*, v. D76, pp. 072009, 2007. doi: 10.1103/PhysRevD.76.072009.
- [85] CDF COLLABORATION. “Single Top Results from CDF”. In: *Electroweak Interactions and Unifield Theories: Proceedings, 42nd Rencontres de Moriond, La Thuile, Italy, March 10-17, 2007*, pp. 117–122, 2007. Disponível em: <<https://inspirehep.net/record/752175/files/arXiv:0706.0282.pdf>>. Acesso em: 21 de abril de 2018.
- [86] CDF COLLABORATION. “Search for single top production using multivariate analyses at CDF”. In: *SUSY 2007 Proceedings, 15th International Conference on Supersymmetry and Unification of Fundamental Interactions, July 26 - August 1, 2007, Karlsruhe, Germany*, pp. 426–429, 2007. doi: 10.2172/920439. Disponível em: <http://lss.fnal.gov/cgi-bin/find_paper.pl?conf-07-557>. Acesso em: 21 de abril de 2018.
- [87] CDF COLLABORATION. “Search for single-top production at CDF”. In: *Proceedings, 42nd Rencontres de Moriond on QCD and High Energy Hadronic Interactions: La Thuile, Italy, March 17-24, 2007*, 2007. Disponível em: <http://lss.fnal.gov/cgi-bin/find_paper.pl?conf-07-152>. Acesso em: 21 de abril de 2018.
- [88] CDF COLLABORATION. “Search for Standard Model Higgs Bosons Produced in Association with W Bosons”, *Phys. Rev. Lett.*, v. 100, pp. 041801, 2008. doi: 10.1103/PhysRevLett.100.041801.
- [89] CDF COLLABORATION. “Measurement of the cross section for $t\bar{t}$ production in $p\bar{p}$ collisions using the kinematics of lepton + jets events”, *Phys. Rev.*, v. D72, pp. 052003, 2005. doi: 10.1103/PhysRevD.72.052003.
- [90] ODORICO, R. “A comparative study of the D0 neural-network analysis of the top quark non-leptonic decay channel”, *EPL (Europhysics Letters)*, v. 55, n. 3, pp. 324, 2001. Disponível em: <<http://stacks.iop.org/0295-5075/55/i=3/a=324>>. Acesso em: 21 de abril de 2018.
- [91] ABAZOV, V. M., et al. “Measurement of the inclusive $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV and determination of the top quark pole mass”, *Phys. Rev.*, v. D94, pp. 092004, 2016. doi: 10.1103/PhysRevD.94.092004.

- [92] FRANC, J., BOUŘ, P., ŠTĚPÁNEK, M., et al. “New Statistical Techniques in the Measurement of the inclusive Top Pair Production Cross Section”. In: *Proceedings, 7th International Workshop on Top Quark Physics (TOP2014): Cannes, France, September 28-October 3, 2014*, 2014. Disponível em: <<https://inspirehep.net/record/1332508/files/arXiv:1412.1076.pdf>>. Acesso em: 21 de abril de 2018.
- [93] D0 COLLABORATION. “Search for the rare decay $B_s \rightarrow \mu\mu$ ”, *Phys. Rev.*, v. D87, n. 7, pp. 072006, 2013. doi: 10.1103/PhysRevD.87.072006.
- [94] D0 COLLABORATION. “Search for single top quarks in the tau+jets channel using 4.8 fb⁻¹ of p p-bar collision data”, *Phys. Lett.*, v. B690, pp. 5–14, 2010. doi: 10.1016/j.physletb.2010.05.003.
- [95] D0 COLLABORATION. “Top Production Cross Sections at D0”. In: *Proceedings, 17th International Workshop on Deep-Inelastic Scattering and Related Subjects (DIS 2009): Madrid, Spain, April 26-30, 2009*, 2009. Disponível em: <http://lss.fnal.gov/cgi-bin/find_paper.pl?conf-09-365>. Acesso em: 21 de abril de 2018.
- [96] D0 COLLABORATION. “Search for flavor changing neutral currents via quark-gluon couplings in single top quark production using 2.3 fb⁻¹ of $p\bar{p}$ collisions”, *Phys. Lett.*, v. B693, pp. 81–87, 2010. doi: 10.1016/j.physletb.2010.08.011.
- [97] D0 COLLABORATION. “Measurement of the t-channel single top quark production cross section”, *Phys. Lett.*, v. B682, pp. 363–369, 2010. doi: 10.1016/j.physletb.2009.11.038.
- [98] D0 COLLABORATION. “Search for Single Top Quark Production at the D0 Experiment using Bayesian Neural Networks”. In: *24th Lake Louise Winter Institute: Fundamental Interactions (LLWI 2009) Lake Louise, Alberta, Canada, February 16-21, 2009*, 2009. Disponível em: <<https://inspirehep.net/record/821793/files/arXiv:0906.0119.pdf>>. Acesso em: 21 de abril de 2018.
- [99] D0 COLLABORATION. “Observation of Single Top Quark Production with the D0 Detector”, *PoS*, v. EPS-HEP2009, pp. 356, 2009.
- [100] D0 COLLABORATION. “Evidence for production of single top quarks”, *Phys. Rev.*, v. D78, pp. 012005, 2008. doi: 10.1103/PhysRevD.78.012005.
- [101] D0 COLLABORATION. “A search for the standard model Higgs boson in the missing energy and acoplanar b-jet topology at $\sqrt{s}=1.96$ ”, *Phys. Rev. Lett.*, v. 101, pp. 251802, 2008. doi: 10.1103/PhysRevLett.101.251802. 48

- [102] CDF COLLABORATION AND D0 COLLABORATION. “Single Top Quark Measurements at the Tevatron”. In: *Proceedings, 7th International Workshop on Top Quark Physics (TOP2014): Cannes, France, September 28-October 3, 2014*, 2014. Disponível em: <<https://inspirehep.net/record/1331789/files/arXiv:1412.0500.pdf>>. Acesso em: 21 de abril de 2018. 48
- [103] CUTTS, D., HOFTUN, J. S., SORNBORGER, A., et al. “Neural networks for event filtering at D0”, *Computer Physics Communications*, v. 57, n. 1-3, pp. 478–482, 1989. 49, 128
- [104] ABE, K., AHRENS, L. A., AMAKO, K., et al. “Determination of $\sin^2\theta_W$ from measurements of differential cross sections for muon-neutrino and -antineutrino scattering by electrons”, *Phys. Rev. Lett.*, v. 62, pp. 1709–1712, Apr 1989. doi: 10.1103/PhysRevLett.62.1709. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevLett.62.1709>>. Acesso em: 21 de abril de 2018. 49
- [105] CUTTS, D., HOFTUN, J. S., SORNBORGER, A., et al. “The use of neural networks in the D0 data acquisition system”, *IEEE Transactions on Nuclear Science*, v. 36, n. 5, pp. 1490–1493, Oct 1989. ISSN: 0018-9499. doi: 10.1109/23.41089. 49, 128, 129
- [106] H1 COLLABORATION. “The H1 detector at HERA”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, v. 386, n. 2, pp. 310 – 347, 1997. ISSN: 0168-9002. doi: [https://doi.org/10.1016/S0168-9002\(96\)00893-5](https://doi.org/10.1016/S0168-9002(96)00893-5). Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168900296008935>>. Acesso em: 21 de abril de 2018. 49
- [107] KÖHNE, J., FENT, J., FRÖCHTENICHT, W., et al. “Realization of a second level neural network trigger for the H1 experiment at HERA”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, v. 389, n. 1, pp. 128 – 133, 1997. ISSN: 0168-9002. doi: [https://doi.org/10.1016/S0168-9002\(97\)00062-4](https://doi.org/10.1016/S0168-9002(97)00062-4). Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168900297000624>>. Acesso em: 21 de abril de 2018. New Computing Techniques in Physics Research V. 49
- [108] ZIMMERMANN, J., KIESLING, C. “Neural networks for the H1 experiment”. In: *IEEE Symposium Conference Record Nuclear Science 2004.*, v. 3, pp. 1869–1872 Vol. 3, Oct 2004. doi: 10.1109/NSSMIC.2004.1462609.

- [109] KIESLING, C., JANAUSCHEK, L., PLACAKYTE, R., et al. “The H1 neural network trigger”. In: *Intelligent Sensors, Sensor Networks and Information Processing Conference, 2004. Proceedings of the 2004*, pp. 319–324. IEEE, 2004. 49
- [110] LIGGINS, M. E., HALL, D. L., LLINAS, J. *Handbook of multisensor data fusion: theory and practice*. Electrical engineering and applied signal processing series. 2 ed. Hoboken, NJ, Taylor & Francis Ltd, 2008. 49, 50, 52
- [111] MOLINA, J. M. “Computational Intelligence and Data Fusion”, Apresentado em: *XIII Brazilian Congress on Computational Intelligence*, 2017. Disponível em: <<http://cbic2017.org/public/pdf/cbic-2017-keynote-molina.pdf>>. Acesso em: 02 de março de 2018. 50, 51
- [112] HALL, D., JORDAN, J. *Human-Centered Information Fusion*. Artech House Electronic Warfare Library. Artech House, 2010. ISBN: 9781596934351. 50
- [113] KESSLER, O., ASKIN, K., BECK, N., et al. *Functional Description of the Data Fusion Process*. Relatório técnico, Office of Naval Technology Data Fusion Development Strategy, 1991. 50
- [114] BLASCH, E., STEINBERG, A., DAS, S., et al. “Revisiting the JDL model for information Exploitation”. In: *Information Fusion (FUSION), 2013 16th International Conference on*, pp. 129–136. IEEE, 2013. 50
- [115] BLASCH, E., PLANO, S. “DFIG level 5 (user refinement) issues supporting situational assessment reasoning”. In: *Information Fusion, 2005 8th International Conference on*, v. 1, pp. xxxv–xliii. IEEE, 2005. 50
- [116] BLASCH, E., KADAR, I., SALERNO, J., et al. “Issues and Challenges in Situation Assessment (Level 2) fusion”. In: *Signal Processing, Sensor Fusion, and Target Recognition XV*, v. 6235, p. 623510. International Society for Optics and Photonics, 2006.
- [117] BLASCH, E., SALERNO, J., KADAR, I., et al. “Resource management coordination with level 2/3 fusion issues and challenges [Panel Report]”, *IEEE Aerospace and Electronic Systems Magazine*, v. 23, n. 3, pp. 32–46, 2008. 50
- [118] BLASCH, E., BOSSÉ, É., LAMBERT, D. A. *High-level information fusion management and systems design*. Artech House, 2012. 50
- [119] WALTZ, E. L. *Information warfare principles and operations*. Artech House, Inc., 1998. 50

- [120] ALAM, F., MEHMOOD, R., KATIB, I., et al. “Data fusion and IoT for smart ubiquitous environments: A survey”, *IEEE Access*, v. 5, pp. 9533–9554, 2017. 50, 51, 53, 57
- [121] CASTANEDO, F. “A review of data fusion techniques”, *The Scientific World Journal*, v. 2013, 2013. 51
- [122] HALL, D. L., LLINAS, J. “An introduction to multisensor data fusion”, *Proceedings of the IEEE*, v. 85, n. 1, pp. 6–23, 1997. 50
- [123] BUEDE, D. M. “Shafer-Dempster and Bayesian reasoning: a response to ‘Shafer-Dempster reasoning with applications to multisensor target identification systems’”, *IEEE Transactions on Systems, Man, and Cybernetics*, v. 18, n. 6, pp. 1009–1011, 1988. 51
- [124] PAN, H., LIANG, Z.-P., ANASTASIO, T. J., et al. “A hybrid NN-Bayesian architecture for information fusion”. In: *Image Processing, 1998. ICIP 98. Proceedings. 1998 International Conference on*, v. 1, pp. 368–371. IEEE, 1998. 51
- [125] EVANS, D. “The Internet of Things: How the Next Evolution of the Internet is Changing Everything”, v. 1, pp. 1–11, 01 2011. 51
- [126] MEHMOOD, R., ALAM, F., ALBOGAMI, N. N., et al. “UTiLearn: A personalised ubiquitous teaching and learning system for smart societies”, *IEEE Access*, v. 5, pp. 2615–2635, 2017. 52
- [127] MITCHELL, H. B. *Multi-sensor data fusion: an introduction*. Springer Science & Business Media, 2007. 52
- [128] HAYKIN, S. *Neural Networks and Learning Machines*. 3 ed. Upper Saddle River, NJ, USA, Prentice Hall, 2008. 53, 54, 114, 135, 177
- [129] KOHONEN, T., OJA, E., SIMULA, O., et al. “Engineering applications of the self-organizing map”, *Proceedings of the IEEE*, v. 84, n. 10, pp. 1358–1384, 1996. 54, 122
- [130] SCHOLKOPF, B., SMOLA, A. J. *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT press, 2001. 54
- [131] ROSS, T. J. *Fuzzy logic with engineering applications*. John Wiley & Sons, 2009. 54
- [132] EIBEN, A., SMITH, J. *Introduction to Evolutionary Computing*. Natural Computing Series. Springer, 2003. ISBN: 9783540401841. 54

- [133] CHALLA, S., PALANISWAMI, M., SHILTON, A. “Distributed data fusion using support vector machines”. In: *Information Fusion, 2002. Proceedings of the Fifth International Conference on*, v. 2, pp. 881–885. IEEE, 2002. 54
- [134] BIGDELI, B., SAMADZADEGAN, F., REINARTZ, P. “A decision fusion method based on multiple support vector machine system for fusion of hyperspectral and LIDAR data”, *International Journal of Image and Data Fusion*, v. 5, n. 3, pp. 196–209, 2014. doi: 10.1080/19479832.2014.919964. Disponível em: <<https://doi.org/10.1080/19479832.2014.919964>>. Acesso em: 03 de março de 2018. 54
- [135] BREIMAN, L. “Random forests”, *Machine learning*, v. 45, n. 1, pp. 5–32, 2001. 54
- [136] MERENTITIS, A., DEBES, C., HEREMANS, R., et al. “Automatic fusion and classification of hyperspectral and LiDAR data using random forests”. In: *Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International*, pp. 1245–1248. IEEE, 2014. 54
- [137] STARZACHER, A., RINNER, B. “Evaluating KNN, LDA and QDA Classification for embedded online Feature Fusion”. In: *Intelligent Sensors, Sensor Networks and Information Processing, 2008. ISSNIP 2008. International Conference on*, pp. 85–90. IEEE, 2008. 54
- [138] AZAD, H. B., MEKHILEF, S., GANAPATHY, V. G. “Long-term wind speed forecasting and general pattern recognition using neural networks”, *IEEE Transactions on Sustainable Energy*, v. 5, n. 2, pp. 546–553, 2014. 54
- [139] LIU, Z., GAO, W., WAN, Y.-H., et al. “Wind power plant prediction by using neural networks”. In: *Energy Conversion Congress and Exposition (ECCE), 2012 IEEE*, pp. 3154–3160. IEEE, 2012. 55
- [140] WU, L., PARK, J., CHOI, J., et al. “A study on wind speed prediction using artificial neural network at Jeju Island in Korea”. In: *Transmission & Distribution Conference & Exposition: Asia and Pacific, 2009*, pp. 1–4. IEEE, 2009.
- [141] ZHANG, F., CHO, K., CHOI, J., et al. “A study on wind speed prediction using artificial neural network at Jeju Island in Korea II”. In: *Circuits and Systems (MWS-CAS), 2011 IEEE 54th International Midwest Symposium on*, pp. 1–4. IEEE, 2011.
- [142] FONTE, P. M., SILVA, G. X., QUADRADO, J. “Wind speed prediction using artificial neural networks.” *WSEAS Transactions on Systems*, v. 4, n. 4, pp. 379–384, 2005. 54

- [143] CHEN, D., ZHANG, K., LIAO, T. “Practical travel time prediction algorithms based on neural network and data fusion for urban expressway”. In: *Natural Computation (ICNC), 2010 Sixth International Conference on*, v. 4, pp. 1754–1758. IEEE, 2010. 55
- [144] ALTINI, M., BRUNELLI, D., FARELLA, E., et al. “Bluetooth indoor localization with multiple neural networks”. In: *Wireless Pervasive Computing (ISWPC), 2010 5th IEEE International Symposium on*, pp. 295–300. IEEE, 2010. 55
- [145] GOODFELLOW, I., BENGIO, Y., COURVILLE, A. *Deep learning*. 2016. Disponível em: <<http://www.deeplearningbook.org>>. Acesso em: 21 de abril de 2018. 55, 56, 184
- [146] GARTNER, INC. “Top 10 Strategic Technology Trends - 2016”. Disponível em: <<http://www.gartner.com/smarterwithgartner/top-ten-technology-trends-signal-the-digital-mesh/>>. Acesso em: 03 de março de 2018. 55
- [147] LECUN, Y., BENGIO, Y., HINTON, G. “Deep learning”, *nature*, v. 521, n. 7553, pp. 436, 2015. 55
- [148] CHUMERIN, N., VAN HULLE, M. M. “Cue and sensor fusion for independent moving objects detection and description in driving scenes”. In: *Signal Processing Techniques for Knowledge Extraction and Information Fusion*, Springer, pp. 161–180, 2008. 55
- [149] LECUN, Y., BOTTOU, L., BENGIO, Y., et al. “Gradient-based learning applied to document recognition”, *Proceedings of the IEEE*, v. 86, n. 11, pp. 2278–2324, 1998. 56
- [150] NGIAM, J., KHOSLA, A., KIM, M., et al. “Multimodal Deep Learning”. In: *Proceedings of the 28th International Conference on International Conference on Machine Learning, ICML’11*, pp. 689–696, USA, 2011. Omnipress. ISBN: 978-1-4503-0619-5. Disponível em: <<http://dl.acm.org/citation.cfm?id=3104482.3104569>>. Acesso em: 21 de abril de 2018. 56, 147
- [151] HINTON, G. E., SALAKHUTDINOV, R. R. “Reducing the dimensionality of data with neural networks”, *science*, v. 313, n. 5786, pp. 504–507, 2006. 56, 147, 184
- [152] MARTÍNEZ, H. P., YANNAKAKIS, G. N. “Deep Multimodal Fusion: Combining Discrete Events and Continuous Signals”. In: *Proceedings of the 16th International Conference on Multimodal Interaction, ICMI ’14*, pp. 34–41, New York, NY,

- USA, 2014. ACM. ISBN: 978-1-4503-2885-2. doi: 10.1145/2663204.2663236. Disponível em: <<http://doi.acm.org/10.1145/2663204.2663236>>. Acesso em: 21 de abril de 2018. 57
- [153] AUDEBERT, N., SAUX, B. L., LEFÈVRE, S. “Semantic Segmentation of Earth Observation Data Using Multimodal and Multi-scale Deep Networks”, *CoRR*, v. abs/1609.06846, 2016. Disponível em: <<http://arxiv.org/abs/1609.06846>>. Acesso em: 21 de abril de 2018. 57
- [154] BADRINARAYANAN, V., KENDALL, A., CIPOLLA, R. “SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation”, *CoRR*, v. abs/1511.00561, 2015. Disponível em: <<http://arxiv.org/abs/1511.00561>>. Acesso em: 21 de abril de 2018. 57
- [155] ATLAS COLLABORATION. “The ATLAS Experiment at the CERN Large Hadron Collider”, *Journal of Instrumentation*, v. 3, n. 08, pp. S08003, 2008. Disponível em: <<http://stacks.iop.org/1748-0221/3/i=08/a=S08003>>. Acesso em: 21 de abril de 2018. 59, 61, 62, 75
- [156] ATLAS COLLABORATION. “Computer generated image of the whole ATLAS detector”. Mar 2008. Disponível em: <<https://cds.cern.ch/record/1095924>>. Acesso em: 21 de abril de 2018. 59
- [157] ATLAS COLLABORATION. “Event Cross Section in a computer generated image of the ATLAS detector.” Mar 2008. Disponível em: <<http://cds.cern.ch/record/1096081>>. Acesso em: 21 de abril de 2018. 60
- [158] *Studies of the ATLAS Inner Detector material using $\sqrt{s} = 13$ TeV pp collision data*. Relatório Técnico ATL-PHYS-PUB-2015-050, CERN, Geneva, Nov 2015. Disponível em: <<http://cds.cern.ch/record/2109010>>. Acesso em: 21 de abril de 2018. 61
- [159] “A study of the material in the ATLAS inner detector using secondary hadronic interactions. A measurement of the material in the ATLAS inner detector using secondary hadronic interactions”, *JINST*, v. 7, n. arXiv:1110.6191. CERN-PH-EP-2011-147, pp. P01013. 22 p, Oct 2011. Disponível em: <<https://cds.cern.ch/record/1394292>>. Acesso em: 21 de abril de 2018.
- [160] ATLAS COLLABORATION. *ATLAS Inner Detector: Technical Design Report, 1*. Technical Design Report ATLAS. Geneva, CERN, 1997.
- [161] ATLAS COLLABORATION. *ATLAS Inner Detector: Technical Design Report, 2*. Technical Design Report ATLAS. Geneva, CERN, 1997. 61

- [162] ATLAS COLLABORATION. “ATLAS Run-2 Detector Status”. 2018. Disponível em: <<https://twiki.cern.ch/twiki/bin/view/AtlasPublic/ApprovedPlotsATLASDetector>>. Acesso em: 4 de março de 2018. 63, 64
- [163] ATLAS COLLABORATION. “ATLAS pixel detector electronics and sensors”, *Journal of Instrumentation*, v. 3, n. 07, pp. P07007, 2008. Disponível em: <<http://stacks.iop.org/1748-0221/3/i=07/a=P07007>>. Acesso em: 21 de abril de 2018. 63
- [164] ATLAS COLLABORATION. “ATLAS Insertable B-Layer Technical Design Report”. ATLAS-TDR-19, 2010. Disponível em: <<http://cds.cern.ch/record/1291633>>. Acesso em: 21 de abril de 2018. 63
- [165] ATLAS COLLABORATION. “Computer Generated image of the ATLAS calorimeter”, Mar 2008. Disponível em: <<http://cds.cern.ch/record/1095927>>. Acesso em: 21 de abril de 2018. 67
- [166] ATLAS COLLABORATION. “Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1”, *Eur. Phys. J.*, v. C77, pp. 490, 2017. doi: 10.1140/epjc/s10052-017-5004-5. 66, 94
- [167] VAN WOERDEN, M. C., WILKENS, H. G., SANTONI, C. *Performance study of the ATLAS Tile Calorimeter using single muon events from $W \rightarrow \mu\nu$ in 2010-2012 collision data*. Relatório Técnico ATL-COM-TILECAL-2015-008, CERN, Geneva, Feb 2015. Disponível em: <<https://cds.cern.ch/record/1988094>>. Acesso em: 21 de abril de 2018. 72
- [168] GARWIN, R. L., LEDERMAN, L. M., WEINRICH, M. “Observations of the Failure of Conservation of Parity and Charge Conjugation in Meson Decays: the Magnetic Moment of the Free Muon”, *Phys. Rev.*, v. 105, pp. 1415–1417, Feb 1957. doi: 10.1103/PhysRev.105.1415. Disponível em: <<http://link.aps.org/doi/10.1103/PhysRev.105.1415>>. Acesso em: 21 de abril de 2018. 76
- [169] MCCUBBIN, N. A. “Off-line Software in HEP: Experience and Trends”, *15th CERN School of Computing*, 1993. Disponível em: <<https://cds.cern.ch/record/400708>>. Acesso em: 21 de abril de 2018. 76
- [170] ALBRAND, S., CHAPMAN, J., COTE, D., et al. *ATLAS Dataset Nomenclature*. Relatório Técnico ATL-GEN-INT-2007-001. ATL-COM-GEN-2007-003, CERN, Geneva, Nov 2007. Disponível em: <<https://cds.cern.ch/record/1070318>>. Acesso em: 21 de abril de 2018. 77

- [171] ATLAS COLLABORATION. “egamma xAOD derivations”. . Disponível em: <<https://twiki.cern.ch/twiki/bin/view/AtlasProtected/EGammaxAODDerivations>>. Acesso em: 05 de março de 2018. 78
- [172] BIRD, I., BUNCIC, P., CARMINATI, F., et al. *Update of the Computing Models of the WLCG and the LHC Experiments*. Relatório Técnico CERN-LHCC-2014-014. LCG-TDR-002, Apr 2014. Disponível em: <<https://cds.cern.ch/record/1695401>>. Acesso em: 21 de abril de 2018. 79, 155
- [173] ATLAS COLLABORATION. *Performance of the ATLAS Trigger System in 2015*. Relatório Técnico arXiv:1611.09661. CERN-EP-2016-241, CERN, Geneva, Nov 2016. Disponível em: <<https://cds.cern.ch/record/2235584>>. Acesso em: 21 de abril de 2018. Todas figuras incluindo auxiliares estão disponíveis em: <http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/TRIG-2016-01/>. 79, 81
- [174] GAMA, R. G. *Filtragem Online Baseada na Fusão de Informação de Detectores Finalemente Segmentados*. Tese de doutorado em engenharia elétrica, COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2017. 80
- [175] XAVIER, T. C. *Identificação Online de Sinais Baseada em Calorimetria de Altas Energias e com Fina Segmentação*. Tese de doutorado em engenharia elétrica, COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2012. 80, 123, 125, 126
- [176] “Performance of the ATLAS Trigger System in 2010”, *The European Physical Journal C*, v. 72, n. 1, pp. 1849, Jan 2012. ISSN: 1434-6052. doi: 10.1140/epjc/s10052-011-1849-1. Disponível em: <<https://doi.org/10.1140/epjc/s10052-011-1849-1>>. Acesso em: 21 de abril de 2018. 80
- [177] COLLABORATION, A. *Trigger Menu in 2017*. Relatório Técnico ATL-COM-DAQ-2018-011, CERN, Geneva, Feb 2018. Disponível em: <<https://cds.cern.ch/record/2304969>>. Acesso em: 21 de abril de 2018. 82
- [178] ATLAS COLLABORATION. “Trigger Operation Public Results”. 2018. Disponível em: <<https://twiki.cern.ch/twiki/bin/view/AtlasPublic/TriggerOperationPublicResults>>. Acesso em: 4 de março de 2018. 83
- [179] *Trigger Menu in 2016*. Relatório Técnico ATL-DAQ-PUB-2017-001, CERN, Geneva, Jan 2017. Disponível em: <<https://cds.cern.ch/record/2242069>>. Acesso em: 21 de abril de 2018. 82
- [180] ATLAS COLLABORATION. *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*. Relatório Técnico

ATLAS-CONF-2016-024, CERN, Geneva, Jun 2016. Disponível em: <<https://cds.cern.ch/record/2157687>>. Acesso em: 21 de abril de 2018. 85, 86, 87, 88, 94, 96, 124, 125, 128

- [181] ANASTOPOULOS, C., UNAL, G., BENHAR NOCCIOLI, E., et al. *Supporting document on electron identification and efficiency measurements using the 2015 LHC proton-proton collision data*. Relatório Técnico ATL-COM-PHYS-2016-041, CERN, Geneva, Jan 2016. Disponível em: <<https://cds.cern.ch/record/2125283>>. Acesso em: 21 de abril de 2018. Nota interna. 85, 86, 95, 124, 125, 128
- [182] *Electron and photon reconstruction and performance in ATLAS using a dynamical, topological cell clustering-based approach*. Relatório Técnico ATL-PHYS-PUB-2017-022, CERN, Geneva, Dec 2017. Disponível em: <<http://cds.cern.ch/record/2298955>>. Acesso em: 21 de abril de 2018. 85, 188
- [183] CORNELISSEN, T. G., ELSING, M., GAVRILENKO, I., et al. “The global χ^2 track fitter in ATLAS”, *Journal of Physics: Conference Series*, v. 119, n. 3, pp. 032013, 2008. Disponível em: <<http://stacks.iop.org/1742-6596/119/i=3/a=032013>>. Acesso em: 21 de abril de 2018. 85
- [184] BRENDLINGER, K. *Physics with Electrons in the ATLAS Detector*. Tese de Doutorado, Pennsylvania U., Aug 2016. Disponível em: <<https://cds.cern.ch/record/2228644>>. Acesso em: 21 de abril de 2018. Apresentado em: 21 de junho de 2016. 91, 128
- [185] COWAN, G. D. *Statistical data analysis*. Oxford, Oxford Univ. Press, 1998. Disponível em: <<https://cds.cern.ch/record/358560>>. Acesso em: 21 de abril de 2018. 92
- [186] HOECKER, A., SPECKMAYER, P., STELZER, J., et al. “TMVA - Toolkit for Multivariate Data Analysis”, *PoS*, v. ACAT, pp. 040, 2007. 92
- [187] CACCIARI, M., SALAM, G. P. “Pileup subtraction using jet areas”, *Phys. Lett.*, v. B659, pp. 119–126, 2008. doi: 10.1016/j.physletb.2007.09.077. 94
- [188] BRUN, R., RADEMAKERS, F. “ROOT — An object oriented data analysis framework”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, v. 389, n. 1, pp. 81 – 86, 1997. ISSN: 0168-9002. doi: [https://doi.org/10.1016/S0168-9002\(97\)00048-X](https://doi.org/10.1016/S0168-9002(97)00048-X). Disponível em: <<http://www.sciencedirect.com/science/article/>

pii/S016890029700048X>. Acesso em: 21 de abril de 2018. *New Computing Techniques in Physics Research* V. 96

- [189] CDF COLLABORATION. “First measurements of inclusive W and Z cross sections from Run II of the Tevatron collider”, *Phys. Rev. Lett.*, v. 94, pp. 091803, 2005. doi: 10.1103/PhysRevLett.94.091803. 96
- [190] ATLAS COLLABORATION. *ATLAS Level-1 Calorimeter Trigger Algorithms*. Relatório Técnico ATL-DAQ-2004-011. CERN-ATL-DAQ-2004-011, CERN, Geneva, Sep 2004. Disponível em: <<https://cds.cern.ch/record/792528>>. Acesso em: 21 de abril de 2018. 101
- [191] *Electron and photon energy calibration with the ATLAS detector using data collected in 2015 at $\sqrt{s} = 13$ TeV*. Relatório Técnico ATL-PHYS-PUB-2016-015, CERN, Geneva, Aug 2016. Disponível em: <<https://cds.cern.ch/record/2203514>>. Acesso em: 21 de abril de 2018. 102
- [192] ATLAS COLLABORATION. “Public Egamma Trigger Plots for Collision Data”. 2018. Disponível em: <<https://twiki.cern.ch/twiki/bin/view/AtlasPublic/EgammaTriggerPublicResults>>. Acesso em: 4 de março de 2018. 103
- [193] MURATORI, B., PIELONI, T. “Luminosity levelling techniques for the LHC”. In: *Proceedings, ICFA Mini-Workshop on Beam-Beam Effects in Hadron Colliders (BB2013): CERN, Geneva, Switzerland, March 18-22 2013*, pp. 177–181, 2014. doi: 10.5170/CERN-2014-004.177. Disponível em: <<https://inspirehep.net/record/1323174/files/arXiv:1410.5646.pdf>>. Acesso em: 21 de abril de 2018. 105
- [194] DENBY, B. H. “A bat’s ear view of neural nets in physics”, *Nucl. Instrum. Meth.*, v. A389, pp. 8–11, 1997. doi: 10.1016/S0168-9002(97)00030-2. 106
- [195] FREUND, W. S. *Algoritmo neural para detecção de partículas em um calorímetro de altas energias*. Trabalho de conclusão de curso de graduação em engenharia elétrica, Poli/UFRJ, Rio de Janeiro, RJ, Brasil, 2011. 106
- [196] SEIXAS, J., CALOBA, L., SOUZA, M., et al. “Neural second-level trigger system based on calorimetry”, *Computer Physics Communications*, v. 95, n. 2–3, pp. 143 – 157, 1996. ISSN: 0010-4655. doi: [http://dx.doi.org/10.1016/0010-4655\(96\)00012-4](http://dx.doi.org/10.1016/0010-4655(96)00012-4). Disponível em: <<http://www.sciencedirect.com/science/article/pii/0010465596000124>>. Acesso em: 21 de abril de 2018. 109, 110, 112, 113, 116, 117, 118, 124, 193, 195

- [197] ANJOS, A. R. D. *Sistema Online de Filtragem em um Ambiente com Alta Taxa de Eventos*. Tese de doutorado em engenharia elétrica, COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2006. 110, 117, 121, 125, 126, 188, 197
- [198] TORRES, R. C. *Sistema Online de Filtragem em um Ambiente com Alta Taxa de Eventos e Fina Granularidade*. Tese de doutorado em engenharia elétrica, COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2010. 110, 119, 121, 122, 123, 125, 126, 127, 149, 155, 175, 187, 197, 203
- [199] PINTO, J. V. F. *Sistema neural para filtragem online em um detector finamente segmentado*. Trabalho de conclusão de curso de graduação em engenharia elétrica, Poli/UFRJ, Rio de Janeiro, RJ, Brasil, 2016. 110
- [200] SEIXAS, J. M., CALOBA, L. P., SOUZA, M. N., et al. “A second-level trigger system based on calorimeters and using neural networks for feature extraction and electron/jets discrimination”. In: *4th International Conference on Calorimetry in High-energy Physics La Biodola, Italy, September 19-25, 1993*, pp. 254–258, 1993. 113, 114
- [201] DESALVO, R., HARTJES, F., HENRIQUES, A., et al. “A novel way of electron identification in calorimeters”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, v. 279, n. 3, pp. 467 – 472, 1989. ISSN: 0168-9002. doi: [https://doi.org/10.1016/0168-9002\(89\)91293-X](https://doi.org/10.1016/0168-9002(89)91293-X). Disponível em: <<http://www.sciencedirect.com/science/article/pii/016890028991293X>>. 113
- [202] ACOSTA, D., BUONTEMPO, S., CALÔBA, L., et al. “Electron-pion discrimination with a scintillating fiber calorimeter”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, v. 302, n. 1, pp. 36 – 46, 1991. ISSN: 0168-9002. doi: [http://dx.doi.org/10.1016/0168-9002\(91\)90489-D](http://dx.doi.org/10.1016/0168-9002(91)90489-D). Disponível em: <<http://www.sciencedirect.com/science/article/pii/016890029190489D>>. Acesso em: 21 de abril de 2018. 113, 196
- [203] ACOSTA, D., BUONTEMPO, S., CALÔBA, L., et al. “Localizing particles showering in a Spaghetti Calorimeter”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, v. 305, n. 1, pp. 55 – 70, 1991. ISSN: 0168-9002. doi: [http://dx.doi.org/10.1016/0168-9002\(91\)90519-V](http://dx.doi.org/10.1016/0168-9002(91)90519-V). Disponível em: <<http://www.sciencedirect.com/science/article/pii/016890029190519V>>. Acesso em: 21 de abril de 2018. 113

- [204] ACOSTA, D., AVELLINO, P., BUONTEMPO, S., et al. “The performance of a lead/scintillating-fiber calorimeter at LHC/SSC compatible gate widths”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, v. 314, n. 3, pp. 431 – 449, 1992. ISSN: 0168-9002. doi: [http://dx.doi.org/10.1016/0168-9002\(92\)90235-V](http://dx.doi.org/10.1016/0168-9002(92)90235-V). Disponível em: <http://www.sciencedirect.com/science/article/pii/016890029290235V>>. Acesso em: 21 de abril de 2018. 113, 114, 129
- [205] BOCK, R., GHEORGHE, A., KLYUCHNIKOV, G., et al. *A second-level trigger, based on calorimeter only*. Relatório Técnico ATL-DAQ-92-007. ATL-D-PN-7, CERN, Geneva, Oct 1992. Disponível em: <https://cds.cern.ch/record/685721>>. Acesso em: 21 de abril de 2018. 116
- [206] HAUSER, R., LEGRAND, I. *EAST note 94-37: Algorithms in second-level triggers for ATLAS and benchmark results*. Relatório Técnico ATL-DAQ-94-027. ATL-D-PN-27, CERN, Geneva, Dec 1994. Disponível em: <https://cds.cern.ch/record/685743>>. Acesso em: 21 de abril de 2018. 116
- [207] ATLAS DATA CHALLENGE TEAM. “ATLAS Data Challenge 1”, *CoRR*, v. cs.DC/0306052, 2003. Disponível em: <http://arxiv.org/abs/cs.DC/0306052>>. Acesso em: 21 de abril de 2018. 117
- [208] ECK, C., KNOBLOCH, J., ROBERTSON, L., et al. *LHC computing Grid: Technical Design Report. Version 1.06 (20 Jun 2005)*. Technical Design Report LCG. Geneva, CERN, 2005. Disponível em: <https://cds.cern.ch/record/840543>>. Acesso em: 21 de abril de 2018. 117
- [209] ATLAS COLLABORATION. “Conjuntos de Dados de Alta Amostragem para o Desafio de Dados do ATLAS”. . Disponível em: http://atlas.web.cern.ch/Atlas/GROUPS/SOFTWARE/DC/DC1/DC1_1/highStat/high_statistics_samples.html>. Acesso em: 15 de fevereiro de 2017. 117
- [210] ATLAS COLLABORATION. “Conjuntos de Dados de Média Amostragem para o Desafio de Dados do ATLAS”. . Disponível em: http://atlas.web.cern.ch/Atlas/GROUPS/SOFTWARE/DC/DC1/DC1_1/mediumStat/medium_statistics_samples.html>. Acesso em: 15 de fevereiro de 2017. 117
- [211] JOLLIFFE, I. T. *Principal component analysis; 2nd ed.* Springer Series in Statistics. New York, Springer, 2002. 118, 120
- [212] GRUBER, A., MOCK, J., FENT, J., et al. “Performance of backpropagation networks in the second-level trigger of the H1-experiment”. In: *3rd International Workshop*

on Software Engineering, Artificial Intelligence and Expert systems for High-energy and Nuclear Physics (AIHENP 93) Oberammergau, Germany, October 4-8, 1993, pp. 429–434, 1993. 118

- [213] SOARES-FILHO, W., SEIXAS, J., CALOBA, L. P. “Enlarging neural class detection capacity in passive sonar systems”. In: *Circuits and Systems, 2002. ISCAS 2002. IEEE International Symposium on*, v. 3, pp. III–III. IEEE, 2002. 120
- [214] HYVÄRINEN, A., KARHUNEN, J., OJA, E. *Independent component analysis*, v. 46. New Jersey, NJ, USA, John Wiley & Sons, 2004. 120
- [215] DUDA, R. O., HART, P. E., STORK, D. G. *Pattern classification*. 2 ed. New York, NY, USA, Wiley, 2001. Disponível em: <<https://cds.cern.ch/record/683166>>. Acesso em: 21 de abril de 2018. 121, 122, 155
- [216] SIMAS FILHO, E. F. D. *Análise Não-Linear de Componentes Independentes para uma Filtragem Online Baseada em Calorimetria de Alta Energia e com Fina Segmentação*. Tese de doutorado em engenharia elétrica, COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2010. 121, 122, 123, 126, 127, 128, 197
- [217] TALEB, A., JUTTEN, C. “Source separation in post-nonlinear mixtures”, *IEEE Transactions on Signal Processing*, v. 47, n. 10, pp. 2807–2820, Oct 1999. ISSN: 1053-587X. doi: 10.1109/78.790661. 121
- [218] PAJUNEN, P., HYVÄRINEN, A., KARHUNEN, J. “Nonlinear blind source separation by self-organizing maps”. In: *In Proc. Int. Conf. on Neural Information Processing*. Citeseer, 1996. 122
- [219] KOHONEN, T. “Improved versions of learning vector quantization”. In: *1990 IJCNN International Joint Conference on Neural Networks*, pp. 545–550 vol.1, June 1990. doi: 10.1109/IJCNN.1990.137622. 122
- [220] ATLAS COLLABORATION. *Expected electron performance in the ATLAS experiment*. Relatório Técnico ATL-PHYS-INT-2010-126, CERN, Geneva, Nov 2010. Disponível em: <<https://cds.cern.ch/record/1309907>>. Acesso em: 21 de abril de 2018. 124
- [221] ATLAS COLLABORATION. “Egamma Collision and MC Data Agreement Task Force 2017”. . Disponível em: <<https://twiki.cern.ch/twiki/bin/view/AtlasProtected/EGammaDataMCAGreementTaskForce2017>>. Acesso em: 2 de março de 2017. 124

- [222] DENBY, B. *Neural network tutorial for high energy physicists*. Relatório técnico, Fermi National Accelerator Lab., Batavia, IL (USA), 1990. 129
- [223] VAN TREES, H. L., BELL, K. L. *Detection, estimation, and filtering theory*. Detection, estimation, and modulation theory. 2 ed. Hoboken, NJ, Wiley, 2013. 133
- [224] RIEDMILLER, M., BRAUN, H. “A Direct Adaptive Method for Faster Backpropagation Learning: The RPROP Algorithm”. In: *Proceedings of the IEEE International Conference on Neural Networks*, pp. 586–591, San Francisco, USA, 1993. 134, 155, 256
- [225] NGUYEN, D., WIDROW, B. “Improving the Learning Speed of 2-Layer Neural Networks by Choosing Initial Values of the Adaptive Weights”. In: *Proceeding of the International Joint Conference on Neural Networks*, v. 3, pp. 21–26, San Diego, USA, 1990. 134
- [226] GREENWOOD, P. E., NIKULIN, M. S. *A guide to chi-squared testing*. Wiley Series in Probability and Statistics. New York, NY, Wiley, 1996. 142
- [227] STUART, A., ORD, J. K. *Distribution Theory*, v. 1, *Kendall’s Advanced Theory of Statistics*. 6 ed. 605 Third Avenue, New York, NY, 10158, John Wiley & Sons, 1994. 142, 144
- [228] ANDERSON, T. W., DARLING, D. A. “Asymptotic theory of certain”goodness of fit”criteria based on stochastic processes”, *The annals of mathematical statistics*, pp. 193–212, 1952. 142
- [229] LOPES, R. “A two-dimensional Kolmogorov-Smirnov test. The two-dimensional Kolmogorov-Smirnov test”, *PoS*, v. ACAT, pp. 045, 2007. Disponível em: <<http://cds.cern.ch/record/1116570>>. Acesso em: 21 de abril de 2018. 142
- [230] PORTER, F. C. “Testing Consistency of Two Histograms”, 2008. 142
- [231] PEARSON, K. “On the probability that two independent distributions of frequency are really samples from the same population”, *Biometrika*, v. 8, n. 1/2, pp. 250–254, 1911. 142
- [232] CRAMÉR, H. *Mathematical methods of statistics (PMS-9)*, v. 9. Princeton university press, 2016. 142
- [233] WICKENS, T. D. *Multiway contingency tables analysis for the social sciences*. Psychology Press, 2014. 142

- [234] MA, C., SHAN, G., LIU, S. “Homogeneity test for correlated binary data”, *PloS one*, v. 10, n. 4, pp. e0124337, 2015. 142
- [235] GAGUNASHVILI, N. D. “Comparison of weighted and unweighted histograms”, *ArXiv Physics e-prints*, maio 2006. arXiv:physics/0605123. 142
- [236] RAO, C. R. “On discrete distributions arising out of methods of ascertainment”, *Sankhyā: The Indian Journal of Statistics, Series A*, pp. 311–324, 1965. 142
- [237] STUART, A., ORD, J. K., ARNOLD, S. *Classical Inference & the Linear Model*, v. 2, *Kendall’s Advanced Theory of Statistics*. 6 ed. 198 Madison Avenue, New York, NY, 10016, Oxford University Press, 1999. 144
- [238] GROSS, E., VITELLS, O. “Trial factors for the look elsewhere effect in high energy physics”, *Eur. Phys. J.*, v. C70, pp. 525–530, 2010. doi: 10.1140/epjc/s10052-010-1470-8. 144
- [239] GAGUNASHVILI, N. D. “Chi-square tests for comparing weighted histograms”, *Nuclear Instruments and Methods in Physics Research A*, v. 614, pp. 287–296, mar. 2010. doi: 10.1016/j.nima.2009.12.037. Acesso em: 21 de abril de 2018. 144
- [240] CICHOCKI, A., ZDUNEK, R., PHAN, A. H., et al. *Nonnegative matrix and tensor factorizations: applications to exploratory multi-way data analysis and blind source separation*. John Wiley & Sons, 2009. 144
- [241] SOARES, F., DAMAZIO, D., SEIXAS, J. “Particle discrimination using matched filters and expert neural networks”. In: *Electronics, Circuits and Systems, 1999. Proceedings of ICECS’99. The 6th IEEE International Conference on*, v. 1, pp. 595–598. IEEE, 1999. 145, 146
- [242] LAMBLIN, P., BENGIO, Y. “Important gains from supervised fine-tuning of deep architectures on large labeled sets”. In: *NIPS* 2010 Deep Learning and Unsupervised Feature Learning Workshop*, 2010. 147
- [243] FREUND, W. S., PINTO, J. V. D. F. “Ringer Project”. 2017. Disponível em: <<https://github.com/joaoVictorPinto/RingerProject>>. Acesso em: 21 de abril de 2018. 155, 255
- [244] NACAD. “Recurso computacional do NACAD - COPPE/UFRJ: Lobo Carneiro”. Disponível em: <<http://www.nacad.ufrj.br/pt/recursos/sgiiicex>>. Acesso em: 22 de fevereiro de 2017. 155
- [245] TORRES, R. C. “Fastnet”. Disponível em: <<https://github.com/rctorres/fastnet>>. Acesso em: 21 de abril de 2018. 155, 255

- [246] CHOLLET, F. “Keras”. 2015. Disponível em: <<https://github.com/fchollet/keras>>. Acesso em: 21 de abril de 2018. 155
- [247] KOHONEN, T. *MATLAB implementations and applications of the self-organizing map*. Unigrafia Oy, Helsinki, Finland, 2014. 177
- [248] DAVIES, D. L., BOULDIN, D. W. “A cluster separation measure”, *IEEE transactions on pattern analysis and machine intelligence*, , n. 2, pp. 224–227, 1979. 178
- [249] ROUSSEEUW, P. J. “Silhouettes: A graphical aid to the interpretation and validation of cluster analysis”, *Journal of Computational and Applied Mathematics*, v. 20, pp. 53 – 65, 1987. ISSN: 0377-0427. doi: [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7). Disponível em: <<http://www.sciencedirect.com/science/article/pii/0377042787901257>>. Acesso em: 21 de abril de 2018. 178
- [250] BREIMAN, L. “Bagging Predictors”, *Machine Learning*, v. 24, n. 2, pp. 123–140, Aug 1996. ISSN: 1573-0565. doi: 10.1023/A:1018054314350. Disponível em: <<https://doi.org/10.1023/A:1018054314350>>. Acesso em: 21 de abril de 2018. 184
- [251] BALDI, P. “Autoencoders, unsupervised learning, and deep architectures”. In: *Proceedings of ICML Workshop on Unsupervised and Transfer Learning*, pp. 37–49, 2012. 184
- [252] ERHAN, D., BENGIO, Y., COURVILLE, A., et al. “Why does unsupervised pre-training help deep learning?” *Journal of Machine Learning Research*, v. 11, n. Feb, pp. 625–660, 2010. 184
- [253] KINGMA, D. P., BA, J. “Adam: A method for stochastic optimization”, *arXiv preprint arXiv:1412.6980*, 2014. 185
- [254] JOE, H. “Relative entropy measures of multivariate dependence”, *Journal of the American Statistical Association*, v. 84, n. 405, pp. 157–164, 1989. 185
- [255] BAUER, M., NEUBERT, M., THAMM, A. “LHC as an Axion Factory: Probing an Axion Explanation for $(g-2)_\mu$ with Exotic Higgs Decays”, *Phys. Rev. Lett.*, v. 119, n. 3, pp. 031802, 2017. doi: 10.1103/PhysRevLett.119.031802. 202

Apêndice A

Produção Acadêmica

A.1 Artigos Publicados em Revista

1. ATLAS COLLABORATION. “Search for $W' \rightarrow tb$ decays in the hadronic final state using pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”, *Phys. Lett.*, B781, pp. 327–348, 2018. doi: 10.1016/j.physletb.2018.03.036.

A search for W' -boson production in the $W' \rightarrow t\bar{b} \rightarrow q\bar{q}'b\bar{b}$ decay channel is presented using 36.1 fb^{-1} of 13 TeV proton-proton collision data collected by the ATLAS detector at the Large Hadron Collider in 2015 and 2016. The search is interpreted in terms of both a left-handed and a right-handed chiral W' boson within the mass range 1-5 TeV. Identification of the hadronically decaying top quark is performed using jet substructure tagging techniques based on a shower deconstruction algorithm. No significant deviation from the Standard Model prediction is observed and the results are expressed as upper limits on the $W' \rightarrow t\bar{b}$ production cross-section times branching ratio as a function of the W' -boson mass. These limits exclude W' bosons with right-handed couplings with masses below 3.0 TeV and W' bosons with left-handed couplings with masses below 2.9 TeV, at the 95

2. ATLAS COLLABORATION. “Search for High-Mass Resonances Decaying to $\tau\nu$ in pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector”, *Phys. Rev. Lett.*, v. 120, n. 16, pp. 161802, 2018. doi: 10.1103/PhysRevLett.120.161802.

A search for high-mass resonances decaying to $\tau\nu$ using proton-proton collisions at $\sqrt{s} = 13$ TeV produced by the Large Hadron Collider is presented. Only τ -lepton decays with hadrons in the final state are considered. The data were recorded with the ATLAS detector and correspond to an integrated luminosity of 36.1 fb^{-1} . No statistically significant excess above the Standard Model expectation is observed; model-independent upper limits are set on the visible $\tau\nu$ production cross section. Heavy W' bosons with masses less than 3.7 TeV in the Sequential Standard Model and masses less than 2.2-

3.8 TeV depending on the coupling in the non-universal G(221) model are excluded at the 95

3. ATLAS COLLABORATION. “Search for the Standard Model Higgs boson produced in association with top quarks and decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”, *Phys. Rev. D*, 2017. doi: 10.1103/PhysRevD.97.072016.

A search for the Standard Model Higgs boson produced in association with a top-quark pair, $t\bar{t}H$, is presented. The analysis uses 36.1 fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV, collected with the ATLAS detector at the Large Hadron Collider in 2015 and 2016. The search targets the $H \rightarrow b\bar{b}$ decay mode. The selected events contain either one or two electrons or muons from the top-quark decays, and are then categorized according to the number of jets and how likely these are to contain b -hadrons. Multivariate techniques are used to discriminate between signal and background events, the latter being dominated by $t\bar{t} + \text{jets}$ production. For a Higgs boson mass of 125 GeV, the ratio of the measured $t\bar{t}H$ signal cross-section to the Standard Model expectation is found to be $\mu = 0.84_{-0.61}^{+0.64}$. A value of μ greater than 2.0 is excluded at 95% confidence level while the expected upper limit is $\mu < 1.2$ in the absence of a $t\bar{t}H$ signal.

4. ATLAS COLLABORATION. “Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector”, *Phys. Rev., D97*, n. 7, pp. 072003, 2018. doi: 10.1103/PhysRevD.97.072003

A search for the associated production of the Higgs boson with a top quark pair ($t\bar{t}H$) is reported. The search is performed in multilepton final states using a dataset corresponding to an integrated luminosity of 36.1 fb^{-1} of proton–proton collision data recorded by the ATLAS experiment at a center-of-mass energy $\sqrt{s} = 13$ TeV at the Large Hadron Collider. Higgs boson decays to WW^* , $\tau\tau$, and ZZ^* are targeted. Seven final states, categorized by the number and flavor of charged-lepton candidates, are examined for the presence of the Standard Model Higgs boson with a mass of 125 GeV and a pair of top quarks. An excess of events over the expected background from Standard Model processes is found with an observed significance of 4.1 standard deviations, compared to an expectation of 2.8 standard deviations. The best fit for the $t\bar{t}H$ production cross section is $\sigma(t\bar{t}H) = 790_{-210}^{+230} \text{ fb}$, in agreement with the Standard Model prediction of $507_{-50}^{+35} \text{ fb}$. The combination of this result with other $t\bar{t}H$ searches from the ATLAS experiment using the Higgs boson decay modes to $b\bar{b}$, $\gamma\gamma$ and $ZZ^* \rightarrow 4\ell$, has an observed significance of 4.2 standard deviations, compared to an expectation of 3.8 standard deviations. This provides evidence for the $t\bar{t}H$ production mode.

5. ATLAS COLLABORATION. “Search for electroweak production of supersymmetric states in scenarios with compressed mass spectra at $\sqrt{s} = 13$ TeV with the ATLAS de-

tector”, *Phys. Rev.*, D97, n. 5, pp. 052010, 2018. doi: 10.1103/PhysRevD.97.052010.

A search for electroweak production of supersymmetric particles in scenarios with compressed mass spectra in final states with two low-momentum leptons and missing transverse momentum is presented. This search uses proton-proton collision data recorded by the ATLAS detector at the Large Hadron Collider in 2015–2016, corresponding to 36.1 fb^{-1} of integrated luminosity at $\sqrt{s} = 13 \text{ TeV}$. Events with same-flavor pairs of electrons or muons with opposite electric charge are selected. The data are found to be consistent with the Standard Model prediction. Results are interpreted using simplified models of R-parity-conserving supersymmetry in which there is a small mass difference between the masses of the produced supersymmetric particles and the lightest neutralino. Exclusion limits at 95% confidence level are set on next-to-lightest neutralino masses of up to 130 GeV for Higgsino production and 170 GeV for wino production, and sleptons masses of up to 180 GeV for pair production of sleptons. In the compressed mass regime, the exclusion limits extend down to mass splittings of 3 GeV for Higgsino production, 2.5 GeV for wino production, and 1 GeV for slepton production. The results are also interpreted in the context of a radiatively-driven natural supersymmetry model with non-universal Higgs boson masses.

6. ATLAS COLLABORATION. “Measurement of the production cross section of three isolated photons in pp collisions at $\sqrt{s} = 8 \text{ TeV}$ using the ATLAS detector”, *Phys. Lett.*, v. B781, pp.55–76, 24p, 2018. doi: 10.1016/j.physletb.2018.03.057.

A measurement of the production of three isolated photons in proton–proton collisions at a centre-of-mass energy $\sqrt{s} = 8 \text{ TeV}$ is reported. The results are based on an integrated luminosity of 20.2 fb^{-1} collected with the ATLAS detector at the LHC. The differential cross sections are measured as functions of the transverse energy of each photon, the difference in azimuthal angle and in pseudorapidity between pairs of photons, the invariant mass of pairs of photons, and the invariant mass of the triphoton system. A measurement of the inclusive fiducial cross section is also reported. Next-to-leading-order perturbative QCD predictions are compared to the cross-section measurements. The predictions underestimate the measurement of the inclusive fiducial cross section and the differential measurements at low photon transverse energies and invariant masses. They provide adequate descriptions of the measurements at high values of the photon transverse energies, invariant mass of pairs of photons, and invariant mass of the triphoton system.

7. ATLAS COLLABORATION. “Search for heavy resonances decaying into WW in the $e\nu\mu\nu$ final state in pp collisions at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS detector. Search for heavy resonances decaying into WW in the $e\nu\mu\nu$ final state in pp collisions at $\sqrt{s} = 13$

TeV with the ATLAS detector”, *Eur. Phys. J.*, v. C78, n. 1, pp.24., 2018. doi: 10.1140/epjc/s10052-017-5491-4.

A search for neutral heavy resonances is performed in the $WW \rightarrow e\nu\mu\nu$ decay channel using pp collision data corresponding to an integrated luminosity of 36.1 fb^{-1} , collected at a centre-of-mass energy of 13 TeV by the ATLAS detector at the Large Hadron Collider. No evidence of such heavy resonances is found. In the search for production via the quark–antiquark annihilation or gluon–gluon fusion process, upper limits on $\sigma_X \times B(X \rightarrow WW)$ as a function of the resonance mass are obtained in the mass range between 200 GeV and up to 5 TeV for various benchmark models: a Higgs-like scalar in different width scenarios, a two-Higgs-doublet model, a heavy vector triplet model, and a warped extra dimensions model. In the vector-boson fusion process, constraints are also obtained on these resonances, as well as on a Higgs boson in the Georgi–Machacek model and a heavy tensor particle coupling only to gauge bosons.

A.2 Artigos Submetidos em Revista

1. ATLAS COLLABORATION. “Search for heavy particles decaying into top-quark pairs using lepton-plus-jets events in proton–proton collisions at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS detector”, *Submetido para: Eur. Phys. J.*, 2018.

A search for new heavy particles that decay into top-quark pairs is performed using data collected from proton–proton collisions at a centre-of-mass energy of 13 TeV by the ATLAS detector at the Large Hadron Collider. The integrated luminosity of the data sample is 36.1 fb^{-1} . Events consistent with top-quark pair production are selected by requiring a single isolated charged lepton, missing transverse momentum and jet activity compatible with a hadronic top-quark decay. Jets identified as likely to contain b -hadrons are required to reduce the background from other Standard Model processes. The invariant mass spectrum of the candidate top-quark pairs is examined for excesses above the background expectation. No significant deviations from the Standard Model predictions are found. Exclusion limits are set on the production cross-section times branching ratio for hypothetical Z' bosons, Kaluza–Kein gluons and Kaluza–Klein gravitons that decay into top-quark pairs.

2. ATLAS COLLABORATION. “A search for lepton-flavor-violating decays of the Z boson into a τ -lepton and a light lepton with the ATLAS detector”, *Submetido para: PRD*, 2018.

Direct searches for lepton flavor violation in decays of the Z boson with the ATLAS detector at the LHC are presented. Decays of the Z boson into an electron or muon and a hadronically decaying τ -lepton are considered. The searches are based on a data

sample of proton–proton collisions collected by the ATLAS detector in 2015 and 2016, corresponding to an integrated luminosity of 36.1 fb^{-1} at a center-of-mass energy of $\sqrt{s} = 13 \text{ TeV}$. No significant excess of events above the expected background is observed, and upper limits on the branching ratios of lepton-flavor-violating decays are set at the 95% confidence level: $\mathcal{B}(Z \rightarrow e\tau) < 5.8 \times 10^{-5}$ and $\mathcal{B}(Z \rightarrow \mu\tau) < 2.4 \times 10^{-5}$. This is the first limit on $\mathcal{B}(Z \rightarrow e\tau)$ with ATLAS data. The upper limit on $\mathcal{B}(Z \rightarrow \mu\tau)$ is combined with a previous ATLAS result based on 20.3 fb^{-1} of proton–proton collision data at a center-of-mass energy of $\sqrt{s} = 8 \text{ TeV}$ and the combined upper limit at 95% level is $\mathcal{B}(Z \rightarrow \mu\tau) < 1.3 \times 10^{-5}$.

3. ATLAS COLLABORATION. “Search for pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS detector”, *Submitted para: JHEP*, 2018.

A search for Higgs boson pair production in the $b\bar{b}b\bar{b}$ final state is carried out with up to 36.1 fb^{-1} of LHC proton–proton collision data collected at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS detector in 2015 and 2016. Three benchmark signals are studied: a spin-2 graviton decaying into a Higgs boson pair, a scalar resonance decaying into a Higgs boson pair, and Standard Model non-resonant Higgs boson pair production. Two analyses are carried out, each implementing a particular technique for the event reconstruction that targets Higgs bosons reconstructed as pairs of jets or single boosted jets. The resonance mass range covered is 260–3000 GeV. The analyses are statistically combined and upper limits on the production cross section of Higgs boson pairs times branching ratio to $b\bar{b}b\bar{b}$ are set in each model. No significant excess is observed; the largest deviation of data over prediction is found at a mass of 280 GeV, corresponding to 2.3 standard deviations globally. The observed 95% confidence level upper limit on the non-resonant production is 13 times the Standard Model prediction.

4. ATLAS COLLABORATION. “Search for R-parity-violating supersymmetric particles in multi-jet final states produced in p - p collisions at $\sqrt{s} = 13 \text{ TeV}$ using the ATLAS detector at the LHC”, *Submitted para: Phys. Lett.*, 2018.

Results of a search for gluino pair production with subsequent R-parity-violating decays to quarks are presented. This search uses 36.1 fb^{-1} of data collected by the ATLAS detector in proton-proton collisions with a center-of-mass energy of $\sqrt{s} = 13 \text{ TeV}$ at the LHC. The analysis is performed using requirements on the number of jets and the number of jets tagged as containing a b -hadron as well as a topological observable formed by the scalar sum of masses of large-radius jets in the event. No significant excess above the expected Standard Model background is observed. Limits are set on the production of gluinos in models with the R-parity-violating decays of either the gluino itself (direct decay) or the neutralino produced in the R-parity-conserving

gluino decay (cascade decay). In the gluino cascade decay model, gluinos with masses between 1000 GeV and 1875 GeV are excluded at 95% confidence level, depending on the neutralino mass. For the gluino direct decay model, the 95% confidence level upper limit on the cross section times branching ratio varies between 0.80 fb at $m_{\tilde{g}} = 900$ GeV and 0.011 fb at $m_{\tilde{g}} = 1800$ GeV.

5. ATLAS COLLABORATION. “Search for low-mass dijet resonances using trigger-level jets with the ATLAS detector in pp collisions at $\sqrt{s}=13$ TeV”, *Submitted para: Phys. Rev. Lett.*, 2018.

Searches for dijet resonances with sub-TeV masses using the ATLAS detector at the Large Hadron Collider can be statistically limited by the bandwidth available to inclusive single-jet triggers, whose data-collection rates at low transverse momentum are much lower than the rate from Standard Model multijet production. This Letter describes a new search for dijet resonances where this limitation is overcome by recording only the event information calculated by the jet trigger algorithms, thereby allowing much higher event rates with reduced storage needs. The search targets low-mass dijet resonances in the range 450-1800 GeV. The analyzed dataset has an integrated luminosity of up to 29.3 fb^{-1} and was recorded at a center-of-mass energy of 13 TeV. No excesses are found; limits are set on Gaussian-shaped contributions to the dijet mass distribution from new particles and on a model of dark-matter particles with axial-vector couplings to quarks.

6. ATLAS COLLABORATION. “Search for a heavy Higgs boson decaying into a Z boson and another heavy Higgs boson in the $\ell\ell b\bar{b}$ final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”, *Submitted para: Phys. Lett.*, 2018.

A search for a heavy neutral Higgs boson, A , decaying into a Z boson and another heavy Higgs boson, H , is performed using a data sample corresponding to an integrated luminosity of 36.1 fb^{-1} from proton-proton collisions at $\sqrt{s} = 13$ TeV recorded in 2015 and 2016 by the ATLAS detector at the Large Hadron Collider. The search considers the Z boson decaying to electrons or muons and the H boson into a pair of b -quarks. No evidence for the production of an A boson is found. Considering each production process separately, the 95% confidence-level upper limits on the $pp \rightarrow A \rightarrow ZH$ production cross-section times the branching ratio $H \rightarrow b\bar{b}$ are in the range of 14-830 fb for the gluon-gluon fusion process and 26-570 fb for the b -associated process for the mass ranges 130-700 GeV of the H boson and process for the mass ranges 130-700 GeV of the H boson and 230-800 GeV of the A boson. The results are interpreted in the context of the two-Higgs-doublet model.

7. ATLAS COLLABORATION. “Search for Higgs boson decays into pairs of light

(pseudo)scalar particles in the $\gamma\gamma jj$ final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”, *Submetido para: Phys. Lett.*, 2018.

This Letter presents a search for exotic decays of the Higgs boson to a pair of new (pseudo)scalar particles, $H \rightarrow aa$, with a mass in the range 20-60 GeV, and where one of the a bosons decays into a pair of photons and the other to a pair of gluons. The search is performed in event samples enhanced in vector-boson fusion Higgs boson production by requiring two jets with large invariant mass in addition to the Higgs boson candidate decay products. The analysis is based on the full dataset of pp collisions at $\sqrt{s} = 13$ TeV recorded in 2015 and 2016 with the ATLAS detector at the CERN Large Hadron Collider, corresponding to an integrated luminosity of 36.7 fb^{-1} . The data are in agreement with the Standard Model predictions and an upper limit at the 95% confidence level is placed on the production cross section times the branching ratio for the decay $H \rightarrow aa \rightarrow \gamma\gamma gg$. This limit ranges from 3.1 pb to 9.0 pb depending on the mass of the a boson.

8. ATLAS COLLABORATION. “Search for top squarks decaying to tau sleptons in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”, *Submetido para: PRD*, 2018.

A search for direct pair production of top squarks in final states with two tau leptons, b -jets, and missing transverse momentum is presented. The analysis is based on proton-proton collision data at $\sqrt{s} = 13$ TeV corresponding to an integrated luminosity of 36.1 fb^{-1} recorded with the ATLAS detector at the Large Hadron Collider in 2015 and 2016. Two exclusive channels with either two hadronically decaying tau leptons or one hadronically and one leptonically decaying tau lepton are considered. No significant deviation from the Standard Model predictions is observed in the data. The analysis results are interpreted in terms of model-independent limits and used to derive exclusion limits on the masses of the top squark \tilde{t}_1 and the tau slepton $\tilde{\tau}_1$ in a simplified model of supersymmetry with a nearly massless gravitino. In this model, masses up to $m(\tilde{t}_1) = 1.16$ TeV and $m(\tilde{\tau}_1) = 1.00$ TeV are excluded at 95% confidence level.

9. ATLAS COLLABORATION. “Search for flavour-changing neutral current top-quark decays $t \rightarrow qZ$ in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”, *Submetido para: JHEP*, 2018.

A search for flavour-changing neutral-current processes in top-quark decays is presented. Data collected with the ATLAS detector from proton-proton collisions at the Large Hadron Collider at a centre-of-mass energy of $\sqrt{s} = 13$ TeV, corresponding to an integrated luminosity of 36.1 fb^{-1} , are analysed. The search is performed using top-quark pair events, with one top quark decaying through the $t \rightarrow qZ$ ($q = u, c$)

flavour-changing neutral-current channel, and the other through the dominant Standard Model mode $t \rightarrow bW$. Only Z boson decays into charged leptons and leptonic W boson decays are considered as signal. Consequently, the final-state topology is characterized by the presence of three isolated charged leptons (electrons or muons) and at least two jets, one of the jets originating from a b -quark. The data are consistent with Standard Model background contributions, and at 95% confidence level the search sets observed (expected) upper limits of 1.7×10^{-4} (2.4×10^{-4}) on the $t \rightarrow uZ$ branching ratio and 2.4×10^{-4} (3.2×10^{-4}) on the $t \rightarrow cZ$ branching ratio, constituting the most stringent limits to date.

10. ATLAS COLLABORATION. “Search for pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”, *Submetido para: JHEP*, 2018.

A search for Higgs boson pair production in the $b\bar{b}b\bar{b}$ final state is carried out with up to 36.1 fb^{-1} of LHC proton-proton collision data collected at $\sqrt{s} = 13$ TeV with the ATLAS detector in 2015 and 2016. Three benchmark signals are studied: a spin-2 graviton decaying into a Higgs boson pair, a scalar resonance decaying into a Higgs boson pair, and Standard Model non-resonant Higgs boson pair production. Two analyses are carried out, each implementing a particular technique for the event reconstruction that targets Higgs bosons reconstructed as pairs of jets or single boosted jets. The resonance mass range covered is 260–3000 GeV. The analyses are statistically combined and upper limits on the production cross section of Higgs boson pairs times branching ratio to $b\bar{b}b\bar{b}$ are set in each model. No significant excess is observed; the largest deviation of data over prediction is found at a mass of 280 GeV, corresponding to 2.3 standard deviations globally. The observed 95% confidence level upper limit on the non-resonant production is 13 times the Standard Model prediction.

11. ATLAS COLLABORATION. “Search for the Decay of the Higgs Boson to Charm Quarks with the ATLAS Experiment”, *Submetido para: Phys. Rev. Lett.*, 2018.

A direct search for the Standard Model Higgs boson decaying to a pair of charm quarks is presented. Associated production of the Higgs and Z bosons, in the decay mode $ZH \rightarrow \ell^+ \ell^- c\bar{c}$ is studied. A dataset with an integrated luminosity of 36.1 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV recorded by the ATLAS experiment at the LHC is used. The $H \rightarrow c\bar{c}$ signature is identified using charm-tagging algorithms. The observed (expected) upper limit on $\sigma(pp \rightarrow ZH) \times \mathcal{B}(H \rightarrow c\bar{c})$ is 2.7 ($3.9_{-1.1}^{+2.1}$) pb at the 95% confidence level for a Higgs boson mass of 125 GeV, while the Standard Model value is 26 fb.

12. ATLAS COLLABORATION. “Measurements of Higgs boson properties in the dipho-

ton decay channel with 36 fb^{-1} of pp collision data at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS detector”, *Submetido para: Phys. Rev. Lett.*, 2018.

Properties of the Higgs boson are measured in the two-photon final state using 36 fb^{-1} of proton–proton collision data recorded at $\sqrt{s} = 13 \text{ TeV}$ by the ATLAS experiment at the Large Hadron Collider. Cross-section measurements for the production of a Higgs boson through gluon–gluon fusion, vector-boson fusion, and in association with a vector boson or a top-quark pair are reported. The signal strength, defined as the ratio of the observed to the expected signal yield, is measured for each of these production processes as well as inclusively. The global signal strength measurement of 0.99 ± 0.14 improves on the precision of the ATLAS measurement at $\sqrt{s} = 7$ and 8 TeV by a factor of two. Measurements of gluon–gluon fusion and vector-boson fusion productions yield signal strengths compatible with the Standard Model prediction. Measurements of simplified template cross sections, designed to quantify the different Higgs boson production processes in specific regions of phase space, are reported. The cross section for the production of the Higgs boson decaying to two isolated photons in a fiducial region closely matching the experimental selection of the photons is measured to be $55 \pm 10 \text{ fb}$, which is in good agreement with the Standard Model prediction of $64 \pm 2 \text{ fb}$. Furthermore, cross sections in fiducial regions enriched in Higgs boson production in vector-boson fusion or in association with large missing transverse momentum, leptons or top-quark pairs are reported. Differential and double-differential measurements are performed for several variables related to the diphoton kinematics as well as the kinematics and multiplicity of the jets produced in association with a Higgs boson. No significant deviations from a wide array of Standard Model predictions are observed.

13. ATLAS COLLABORATION. “Search for photonic signatures of gauge-mediated supersymmetry in 13 TeV pp collisions with the ATLAS detector.”, *Submetido para: Phys. Rev. Lett.*, 2018.

A search is presented for photonic signatures, motivated by generalized models of gauge-mediated supersymmetry breaking. This search makes use of proton-proton collision data at $\sqrt{s} = 13 \text{ TeV}$ corresponding to an integrated luminosity of 36.1 fb^{-1} recorded by the ATLAS detector at the LHC, and explores models dominated by both strong and electroweak production of supersymmetric partner states. Experimental signatures incorporating an isolated photon and significant missing transverse momentum are explored. These signatures include events with an additional photon or additional jet activity not associated with any specific underlying quark flavor. No significant excess of events is observed above the Standard Model prediction, and 95% confidence-level upper limits of between 0.083 fb and 0.32 fb are set on the visible cross section of contributions from physics beyond the Standard Model. These results are interpreted

in terms of lower limits on the masses of gluinos, squarks, and gauginos in the context of generalized models of gauge-mediated supersymmetry, which reach as high as 2.3 TeV for strongly produced and 1.3 TeV for weakly produced supersymmetric partner pairs.

14. ATLAS COLLABORATION. “Search for a Structure in the $B_s^0\pi^\pm$ Invariant Mass Spectrum with the ATLAS Experiment”, *Submetido para: Phys. Rev. Lett.*, 2018.

A search for the narrow structure, $X(5568)$, reported by the D0 COLLABORATION in the decay sequence $X \rightarrow B_s^0\pi^\pm$, $B_s^0 \rightarrow J/\psi\phi$, is presented. The analysis is based on a data sample recorded with the ATLAS detector at the LHC corresponding to 4.9 fb^{-1} of pp collisions at 7 TeV and 19.5 fb^{-1} at 8 TeV. No significant signal was found. Upper limits on the number of signal events, with properties corresponding to those reported by D0, and on the X production rate relative to B_s^0 mesons, ρ_X , were determined at 95% confidence level. The results are $N(X) < 382$ and $\rho_X < 0.016$ for B_s^0 mesons with transverse momenta above 10 GeV, and $N(X) < 356$ and $\rho_X < 0.017$ for transverse momenta above 15 GeV. Limits are also set for potential $B_s^0\pi^\pm$ resonances in the mass range 5550 MeV to 5700 MeV.

15. ATLAS COLLABORATION. “Search for light resonances decaying to boosted quark pairs and produced in association with a photon or a jet in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”, *Submetido para: Phys. Lett.*, 2018.

This Letter presents a search for new light resonances decaying to pairs of quarks and produced in association with a high- p_T photon or jet. The dataset consists of proton-proton collisions with an integrated luminosity of 36.1 fb^{-1} at a centre-of-mass energy of $\sqrt{s} = 13$ TeV recorded by the ATLAS detector at the Large Hadron Collider. Resonance candidates are identified as massive large-radius jets with substructure consistent with a particle decaying into a quark pair. The mass spectrum of the candidates is examined for local excesses above background. No evidence of a new resonance is observed in the data, which are used to exclude the production of a lepto-phobic axial-vector Z' boson. These results improve upon the limits on light dijet resonances obtained at lower centre-of-mass energies.

16. ATLAS COLLABORATION. “Measurements of $t\bar{t}$ differential cross-sections of highly boosted top quarks decaying to all-hadronic final states in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector”, *Submetido para: PRD*, 2018.

Measurements are made of differential cross-sections of highly boosted pair-produced top quarks as a function of top-quark and $t\bar{t}$ system kinematic observables using proton-proton collisions at a center-of-mass energy of $\sqrt{s} = 13$ TeV. The data set corresponds to an integrated luminosity of 36.1 fb^{-1} , recorded in 2015 and 2016 with

the ATLAS detector at the CERN Large Hadron Collider. Events with two large-radius jets in the final state, one with transverse momentum $p_T > 500$ GeV and a second with $p_T > 350$ GeV, are used for the measurement. The top-quark candidates are separated from the multijet background using jet substructure information and association with a b -tagged jet. The measured spectra are corrected for detector effects to a particle-level fiducial phase space and a parton-level limited phase space, and are compared to several Monte Carlo simulations by means of calculated χ^2 values. The cross-section for $t\bar{t}$ production in the fiducial phase-space region is $292 \pm 7(\text{stat}) \pm 76(\text{syst})$ fb, to be compared to the theoretical prediction of 384 ± 36 fb.

17. ATLAS COLLABORATION and CMS COLLABORATION. “Combination of inclusive and differential $t\bar{t}$ charge asymmetry measurements using ATLAS and CMS data at $\sqrt{s} = 7$ and 8 TeV”, *Submetido para: Journal of High Energy Physics*, 2017.

This paper presents combinations of inclusive and differential measurements of the charge asymmetry (A_C) in top quark pair ($t\bar{t}$) events with a lepton+jets signature by the ATLAS and CMS Collaborations, using data from LHC proton-proton collisions at centre-of-mass energies of 7 and 8 TeV corresponding to integrated luminosities of about 5 and 20 fb^{-1} for each experiment, respectively. The resulting combined LHC measurements of the inclusive charge asymmetry are $A_C^{\text{LHC7}} = 0.005 \pm 0.007$ (stat) \pm 0.006 (syst) at 7 TeV and $A_C^{\text{LHC8}} = 0.0055 \pm 0.0023$ (stat) \pm 0.0025 (syst) at 8 TeV. These values, as well as the combination of A_C measurements as a function of the invariant mass of the $t\bar{t}$ system at 8 TeV, are consistent with the respective standard model predictions.

A.3 Artigos Publicados em Anais de Congresso

1. SANTOS, M., SOUZA, E., FARIAS, P., SIMAS, E., FREUND, W. S., SEIXAS, J. M. “Máquinas de Aprendizado Extremo para Classificação Online de Eventos no detector ATLAS”. XXXV Simpósio Brasileiro de Telecomunicações e Processamento de Sinais, 2017, São Pedro. In: *Anais do XXXV Simpósio Brasileiro de Telecomunicações e Processamento de Sinais*, 2017.

O ATLAS é um dos detectores do LHC (Large Hadron Collider), e está localizado no CERN (Organização Europeia para a pesquisa Nuclear). Para adequada caracterização das partículas é preciso realizar uma precisa medição do perfil de deposição de energia à medida que ocorrem interações com o detector. No ATLAS os calorímetros são responsáveis por realizar a estimativa da energia das partículas e, neste sentido, utilizam mais de 100.000 sensores. Um dos discriminadores para a detecção online de elétrons utilizados no ATLAS é o Neural Ringer, no qual o perfil de deposição de energia

é utilizado como entrada para um classificador neural tipo perceptron de múltiplas camadas. Este trabalho propõe o uso de Máquinas de Aprendizado Extremo (ELM) em substituição às redes do tipo *multilayer perceptron* no Neural Ringer. Os resultados obtidos de uma base de dados simulados apontam para uma significativa redução do tempo de treinamento, com desempenho de classificação semelhante.

2. FREUND, W. S., SOBRINHO, G. S., SEIXAS, J. M. “Ensemble Neural para Identificação de Partículas baseada na Informação de um Calorímetro Finamente Segmentado.” In: *Annals of XIII Brazilian Congress on Computational Intelligence*, 2017.

A Ciência de Dados tem natureza multidisciplinar e possibilita estender soluções de engenharia para outras áreas. Uma aplicação em Física de Partículas reflete desafios comuns de detecção de eventos de interesse, como empilhamento de sinais, alta taxa, eventos raros e alteração dos padrões dos eventos conforme posição de incidência no detector e intensidade do sinal. No caso, um método de *ensemble* de redes neurais tornou-se, em 2017, parte da seleção *online* de eventos de interesse de um dos maiores experimentos da atualidade (ATLAS), onde o *ensemble* atua como um classificador de partículas. Neste trabalho, avalia-se a extensão da técnica para atuação *offline* através da comparação com a técnica vigente, deteriorada para ambas acessarem o mesmo nível de informação, em dados simulados. A proposta obtém redução de falso alarme a, no mínimo, 70 % (valor central) da referência para toda a faixa de energia para detecção similar.

3. PINTO, J. V. F., FREUND, W. S., SEIXAS, J. M. “Sistema de filtragem Online Utilizando um Ensemble de Redes Neurais e Informação de Calorimetria para Operar em Altas Taxas de Eventos.” In: *XIII Brazilian Congress on Computational Intelligence*, 2017.

Diversas aplicações na engenharia estão relacionadas em ambientes cujo a alta taxa de eventos e a rara ocorrência são uma dificuldade a ser superada. Este artigo tem como proposta apresentar a nova estratégia de seleção de elétrons no sistema de filtragem Online do detector ATLAS, no CERN, utilizando técnicas multivariadas, como Redes Neurais, para selecionar eficientemente e reduzir o custo de processamento na fazenda de computadores durante a filtragem. Dentro desse contexto, serão apresentadas as estratégias de treinamento das redes e correção das eficiências mediante ao crescimento do empilhamento de eventos ocasionado pelo aumento da luminosidade das colisões para o cenário de 2017.

4. OLIVEIRA, A. C. N., FREUND, W. S., SEIXAS, J. M. “Redes Neurais Especialistas para a Identificação de Partículas em um Calorímetro Finamente Segmentado.” In: *XIII Brazilian Congress on Computational Intelligence*, 2017.

Apresenta-se uma proposta de aplicação de Ciência de Dados no contexto de Física de Partículas, aplicando um ensemble de redes neurais para efetuar a discriminação de partículas a partir da fusão de dados de diferentes conjuntos de sensores finamente segmentados. A técnica procura identificar raros eventos de interesse durante fase de operação online, dentre um grande conjunto total de eventos. Ao comparar a proposta estendendo a versão em vigor, observa-se um ganho de 1% em detecção e redução de 1,5% (valores centrais) em falso alarme para atuação na filtragem online de elétrons.

5. FREUND W. S., PINTO J. V. F., SEIXAS M. J. “Redes neurais aplicadas na identificação de partículas em um detector finamente segmentado”, XXI Congresso Brasileiro de Automática.

A física de altas energias proporciona um ambiente contendo diversos desafios comuns a sistemas inteligentes, com requisitos extremos na eficiência de identificação de raros eventos de interesses. No caso do experimento ATLAS, no CERN, os eventos são gerados em alta taxa (40 MHz), e aplica-se uma filtragem online para reduzi-la a um nível viável (1 kHz), havendo a manutenção dos eventos de interesse, que serão analisados em offline. Uma proposta alternativa para a seleção de eventos contendo elétrons, baseada em redes neurais, é descrita e analisada neste trabalho, utilizando dados de simulação, obtendo-se uma redução de 50% para o ruído de fundo, enquanto mantém-se a taxa de detecção dos eventos de interesse praticamente inalterada no final da filtragem.

6. PINTO, J. V. F., FREUND, W. S., DAMAZIO, D., SEIXAS, J. M., ATLAS COLLABORATION. “Ring-shaped Calorimetry Information for a Neural Egamma Identification with ATLAS Detector”. In: *17th International Workshop on Advanced Computing and Analysis Techniques in Physics Research*, 2016, Val Paraiso. Journal of Physics: Conference Series, 2016. v. 762.

This article presents the identification process of electrons based only on calorimeter information. It is based on the usage of ring-shaped description for a region of interest of the calorimeter which explores the shower shape propagation throughout the ATLAS calorimeters. This information is fed into a multivariate discriminator, currently an artificial neural network, responsible for hypothesis testing. The concept is evaluated for online selection (trigger), used for reducing storage rate into viable levels while preserving collision events containing desired signals. Preliminary results from Monte Carlo simulation data indicate that the background rejection can be reduced by as much as 50 % over the current method used in the High-Level Trigger, allowing for high-latency reconstruction algorithms such as tracking to run at a later stage of the trigger.

A.4 Resumos Publicados em Anais de Congresso

1. OLIVEIRA, A. C. N., FREUND, W. S., SEIXAS, J. M., BEGALLI, M. “Identificação Neural de elétrons no Experimento ATLAS a partir da Fusão de Informações de Traços e Calorimetria”. In: XXXVIII Encontro Nacional de Física de Partículas e Campos, 2017, Passa Quatro. Anais do XXXVIII Encontro Nacional de Física de Partículas e Campos, 2017.

O Experimento ATLAS está em operação durante seu segundo período de coleta de dados (Run 2), onde o acelerador LHC tem incrementado sucessivamente sua luminosidade de operação. Em 2016, atingiram-se valores de luminosidade em pico de $1,4 \times 10^{34} \text{cm}^{-2} \text{s}^{-1}$, o equivalente a 44 interações inelásticas médias por evento de cruzamento pacotes próton-próton. Com o aumento da luminosidade, decorre um aumento de empilhamento de sinais nos detectores e, conseqüentemente, acarreta em maior exigência das técnicas aplicadas para identificação de partículas, como na tarefa de discriminação elétron-hádron.

Os elétrons interagem com dois detectores do ATLAS: Detector Interno (ID), e o Sistema de Calorimetria. O ID permite observar a trajetória de partículas carregadas, enquanto o Sistema de Calorimetria mede a energia da partícula original através de um processo destrutivo de multiplicação de partículas chamado chuveiro.

Este trabalho desenvolve uma estratégia para discriminação elétron-hádron offline no ATLAS baseada na representação de informação anelada de calorimetria, o algoritmo Ringer, em conjunto com 6 grandezas físicas atualmente empregadas pelo ATLAS para descrição das medições do ID, como propriedades do traço, combinação da informação de traço-calorimetria e medições de perdas por bremsstrahlung. A perspectiva de análise da informação anelada calorimetria possibilita a redução da dimensionalidade do problema ao explorar a geometria aproximadamente cônica do desenvolvimento do chuveiro de partículas para discriminar sinal (elétrons) de ruído (hádrons) através da informação lateral e longitudinal retida na soma de energia em anéis concêntricos ao baricentro de energia da propagação do chuveiro.

Os anéis e as grandezas físicas são alimentados a conjuntos de redes neurais especialistas e ajustados separadamente para lidar com a natureza fundamentalmente distinta dos sinais de traço e de calorimetria. Essas redes são treinadas por regiões do espaço de fase $E_T \times \eta$ e então combinadas através da conexão de suas camadas escondidas em uma terceira rede neural, a qual, após o ajuste de seus parâmetros, realiza a decisão final.

Avaliou-se a eficiência da proposta em simulações contendo decaimentos $Z \rightarrow ee$ e seu respectivo ruído de fundo, segundo as condições de operações da Run 2 durante 2016,

comparando o resultado com a técnica vigente.

2. SILVEIRA, R. C., FREUND, W. S., SEIXAS, J. M., BEGALLI, M. “Identificação de Elétrons no ATLAS Baseada em Pré-Processamento Segmentado da Informação de Calorimetria”. In: XXXVIII Encontro Nacional de Física de Partículas e Campos, 2017, Passa Quatro.

O LHC acelera partículas e colide feixes de prótons a uma taxa de 40 MHz, com energia de 13 TeV no centro de massa. Estas colisões são observadas em alta luminosidade ($1,4 \times 10^{34} \text{cm}^{-2}\text{s}^{-1}$ em pico), gerando um ambiente propício para o estudo de eventos raros. Nesse contexto, têm-se elétrons como parte dos portadores de física de interesse, permitindo aprofundar o conhecimento da estrutura da matéria.

Dentre os experimentos do LHC, destaca-se neste trabalho o ATLAS, que se apoia fortemente na informação proveniente do seu sistema de calorimetria para reconstrução de elétrons contendo 4 camadas eletromagnéticas e 3 hadrônicas. Durante o processo de interação das partículas com o material do calorímetro, ocorre a produção de chuveiros de partículas secundárias que geram sinal proporcional a energia da partícula incidente. Outras partículas, como hádrons, possuem perfil de deposição de energia similar ao de elétrons, sendo necessária a tarefa de discriminação elétron-hádron para o sucesso do experimento.

Para realizar a identificação de elétrons, utiliza-se o algoritmo Neural Ringer (NR), que se encontra em operação desde este ano como etapa rápida de pré-seleção do Alto Nível de Filtragem (online) do ATLAS. O NR formata a informação de calorimetria em 100 anéis concêntricos de deposição de energia, mantendo a informação discriminante do chuveiro, e utiliza um conjunto de redes neurais artificiais para a tomada de decisão operando conforme regiões do espaço de fase.

Neste trabalho, propõe-se aprimorar a estratégia empregada no NR através do emprego de pré-processamento por ICA contemplando a segmentação longitudinal em camadas. O algoritmo de ICA empregado (JADE) tem como propriedade a utilização de número de fontes originais idêntico às misturas. Como supõe-se que o chuveiro de partículas tenha um número menor de fontes do que o observado, emprega-se anteriormente Análise de Componentes Principais (PCA) com a finalidade de compactação de dados, preservando a maior parte de energia do sinal em menos componentes e possibilitando a redução do número de componentes extraídas no final do processo.

Para a extração não-segmentada, aplicando um corte em 99% de energia em dados gerados por simulação de Monte Carlo, identificaram-se 22 componentes principais, enquanto que a abordagem segmentada decorreu em 48 componentes, indicando a maior necessidade de informação longitudinal para descrever a deposição de energia

das camadas. Este resultado preliminar indica que ao segmentar em camadas, preserva-se mais informação discriminante para descrever os anéis do calorímetro em relação a abordagem não-segmentada.

3. FREUND, W. S.; SOBRINHO, G. S., SEIXAS, J. M., BEGALLI, M. “Identificação Neural de elétrons com informação anelada de calorimetria”. In: XXXVIII Encontro Nacional de Física de Partículas e Campos, 2017, Passa Quatro.

O LHC opera com alta taxa de colisões (40 MHz em pico), o que exigiria um fluxo de dados de 70 TB/s de armazenamento e torna inviável o armazenamento de todos os eventos. Por outro lado, grande parte das colisões não possui informação física relevante, assim Sistema de Filtragem é encarregado pela redução da taxa de armazenamento enquanto mantendo apenas eventos de colisões promissoras para posterior análise off-line. A filtragem é feita em cadeias sequenciais, crescente em latência e profundidade de reconstrução física. O primeiro nível é implementado em hardware, responsável pela redução da taxa para 100 kHz com uma latência menor que $2,5 \mu\text{s}$. Em seguida, o High-Level Trigger (HLT), implementado em software, limita a taxa de saída para 1 kHz, com uma latência alvo de 550 ms

Elétrons são encontrados como decaimentos finais em canais de física de interesse, tornando assim a tarefa de discriminação essencial. Tal tarefa se resume em discerní-los de hádrons que interagem de forma similar com o detector. Essas partículas interagem com o Detector de Traços (ID), responsável pela reconstrução da trajetória da partícula, e o Sistema de Calorimetria que, através de um processo contínuo de multiplicação de partículas (chuveiro), gera sinal proporcional a energia da partícula original. Esse processo tem em seu formato informação discriminante para a tarefa de identificação. A informação do ID demanda maior poder computacional para processamento, assim a operação no Sistema de Filtragem conta com uma etapa de pré-seleção somente com a informação de calorimetria. Neste ano, o ATLAS adotou uma estratégia para discriminação elétron-hádron (Ringer) baseada em informação de calorimetria e um conjunto de redes neurais MLP padrões por região do espaço de fase $E_T \times \eta$ para pré-seleção de elétrons no HLT. .

Este trabalho estende a proposta para a operação no ambiente off-line. No algoritmo, emprega-se a representação da informação discriminante por informação anelada de calorimetria, na qual se explora a geometria cônica dos chuveiros ao formar anéis com as células do calorímetro ao entorno do seu centro de deposição de energia. A grandeza representada é a soma de energia das células que formam os anéis.

Apresenta-se uma comparação da proposta com o método vigente (LH) em operação off-line empregado no experimento para dados simulados com as condições das colisões obtidas durante 2016. Como parte do desenvolvimento do algoritmo proposto,

degradou-se o LH para acessar somente a informação de calorimetria, possibilitando uma comparação dos métodos, no critério medium, acessando a mesma parcela de informação. Constatou-se que, para a mesma taxa de detecção de elétrons, o Ringer apresenta taxa de falso alarme entre 40-80% daquelas obtidas pela técnica LH com apenas informação de calorimetria.

4. PINTO, J. V. F., FREUND, W. S., SEIXAS, J. M., DAMAZIO, D., BEGALLI, M. “Informação anelada de calorimetria para o Trigger do ATLAS”. In: XXXVIII Encontro Nacional de Física de Partículas e Campos, 2017, Passa Quatro.

O LHC continua a progressão da luminosidade dos eventos de colisões próton-próton, o que dificulta a reconstrução correta de eventos em decorrência do empilhamento de sinais nos detectores dos experimentos. Esse é o caso da tarefa de discriminação de elétron-jato no ATLAS, de vital importância para o programa de física do experimento, em especial durante seu processamento *trigger*, que é responsável pela redução da taxa de armazenamento de eventos para nível viável de armazenamento e processamento *off-line*.

A filtragem ocorre em cadeias sequenciais, crescente em latência e profundidade de reconstrução da física. O primeiro nível, implementado em *hardware*, seleciona eventos promissores em uma taxa até 100 kHz com uma latência inferior a 2,5 μ s. Posteriormente, opera via *software* o *High-Level Trigger* com uma latência média de 550 ms e taxa de saída para 1,5 kHz. O HLT opera em duas etapas, aplicando uma etapa rápida, necessária para a redução da latência total das cadeias, e uma etapa precisa para a decisão final. No caso de elétrons, a informação discriminante é oriunda do Sistema de Calorimetria e do Detector Interno (ID), porém ambas etapas possuem uma pré-seleção empregando apenas a informação de calorimetria (*FastCalo*, no caso da etapa rápida) devido à alta demanda de processamento da informação do ID.

Neste trabalho, continuou-se o desenvolvimento do algoritmo *Neural Ringer* para o *FastCalo*. Esse algoritmo emprega soma de energia em anéis concêntricos à posição de incidência da partícula. Os anéis são propagados para um conjunto de redes neurais especialistas por regiões do espaço de fase $\eta \times E_T$. Para dar resiliência ao empilhamento de sinais, os limiares de decisão aplicados são ajustados conforme o número de colisões inelásticas médias. Avaliou-se a performance do algoritmo em *runs* de referência de 2016, onde se observou que as cadeias empregando a proposta causam reduções entre 2,8 a 6,4 da taxa de ruído após o *FastCalo*, dependendo do critério aplicado, sem a alteração na taxa de detecção de elétrons no final da cadeia em relação a estratégia originalmente empregada. Em termos de latência, há uma redução de 159 ms para 121 ms. Esses resultados, em conjunto com crescimento de processamento causado pelo aumento previsto da luminosidade, motivaram o ATLAS a empregar o *Neural*

Ringer em todas suas cadeias dedicadas a filtragem de elétrons. O algoritmo encontra-se comissionado, período encerrado no final de junho de 2017.

5. PINTO, J. V. F.; FREUND, W. S.; SEIXAS, J. M.; DAMAZIO, D. “Data Collision Analysis to an Alternative High Level Trigger in ATLAS Detector”. In: Encontro de Física de 2016, Natal.

O método proposto explora uma alternativa para identificação de elétrons e fótons na pré-seleção realizada no HLT utilizando descrição anelada das células do calorímetro e alimentando com essa informação discriminadores multivariáveis responsáveis pela seleção, atualmente redes neurais. Resultados preliminares em dados de colisão mostram que a cadeia de filtragem exigindo elétrons com no mínimo 24 GeV obtém uma redução relativa de ruído de fundo de 60 %, com respeito àquela empregada pela estratégia convencional, quando medindo a eficiência da cadeia nesse estágio de pré-seleção, o que impacta na redução de latência da cadeia.

6. PINTO, J. V. F., FREUND, W. S., SEIXAS, J. M., DAMAZIO, D. O. Informação anelada de calorimetria para identificação neural de elétrons e fótons no detector ATLAS. In: XXXVI Encontro Nacional de Física de Partículas e Campos, 2015, Caxambu (MG).

Neste trabalho, examina-se a proposta de extração anelada de calorimetria para a identificação de partículas no Canal e/γ em condições de alta luminosidade. Essa informação, que explora a geometria do formato do chuveiro, é propagada para um discriminador multivariável, atualmente redes neurais, responsável pela tomada de decisão. A proposta é realizada tanto para o ambiente de reconstrução offline, após o armazenamento dos dados, quanto no âmbito do sistema de filtragem online, necessário para reduzir a taxa de armazenamento a níveis viáveis, para preservar apenas física de interesse. A filtragem é realizada em três níveis sequenciais, tratamento que permite descartar falsos candidatos de simples análise nos níveis inferiores, possibilitando uma análise mais extensa em candidatos complexos. O Neural Ringer opera no segundo nível de filtragem. Quando contrapondo resultados com a abordagem de referência utilizada pelo ATLAS, no segundo nível de filtragem, baseada em cortes, identificou-se a capacidade em reduzir em até 70 % a taxa de falsos candidatos aceitos pelo Sistema de Filtragem enquanto se mantenha a mesma eficiência de identificação de elétrons obtida pela abordagem de referência.

A.5 Qualificação para Autoria no Experimento ATLAS

Para obter o direito de ser autor de publicações no ATLAS, os membros do ATLAS precisam realizar uma tarefa de necessidade do experimento. Esse processo é coordenado por dois supervisores, um local, José Manoel de Seixas (COPPE/UFRJ), e outro indicado pelo ATLAS, Aranzazu Ruiz Martinez (*Carleton University*, Ottawa, Canada). O autor deste trabalho encerrou o seu processo de qualificação na data 28/07/2017, com o seguinte comentário pelo revisor técnico indicado pelo ATLAS:

The L2 ringer neural network developed by the Rio group was implemented as part of the electron trigger sequence to be used by default in 2017 data taking. During these months, the ringer algorithm was tuned to reproduce the same efficiency as obtained with the L2 calorimeter cut-based selection used up to now, while achieving a higher early rejection at L2 and therefore reducing the CPU cost. Werner in collaboration with the Rio group provided new ringer tunes in a timely manner and had an outstanding contribution to the electron trigger commissioning with the first 2017 data. Duplicated ringer chains have been extensively validated in the last months with successful results. The electron triggers with $E_T > 15$ GeV switched to ringer in rel 21.1.8 deployed online at P1 after TS1. All the performance studies have been presented and discussed in several egamma trigger meetings. I would like to thank Werner for his huge effort during this period and I highly recommend him for qualification.

A.6 Educação e Popularização de Ciência e Tecnologia

- Divulgação do trabalho no jornal impresso e digital do Estado de São Paulo. “UFRJ cria filtro de dados para superacelerador de partículas”, 2018. Disponível em: <http://ciencia.estadao.com.br/noticias/geral,ufrj-cria-filtro-de-dados-parasuperacelerador-de-particulas>. Acesso em: 21 de abril de 2018.
 - Época Negócios. “Grupo da UFRJ cria filtro de dados para acelerador de partículas”.
 - Jornal da Ciência. “Universidade cria filtro de dados para superacelerador de partículas”.
 - Jornal do Brasil Online. “Grupo da UFRJ cria filtro de dados para acelerador de partículas”.

- UOL Notícias. “Grupo da UFRJ cria filtro de dados para acelerador de partículas”.
- Folha Vitória. “Grupo da UFRJ cria filtro de dados para acelerador de partículas”.
- Divulgação do trabalho com entrevista do professor e orientador do autor do trabalho, José Manoel de Seixas, no jornal digital do Sputnik. “Pesquisadores brasileiros colaboram para a compreensão científica da ‘Partícula de Deus’”. Disponível em: https://br.sputniknews.com/ciencia_tecnologia/2018032310810328-ciencia-pesquisa-origem-universo-universidades-fisica/. Acesso em: 21 de abril de 2018.
- Divulgação do trabalho no jornal digital da EBC. Maior laboratório de partículas do mundo usa sistema atualizado da UFRJ. Disponível em: <http://agenciabrasil.ebc.com.br/pesquisa-e-inovacao/noticia/2018-03/maior-laboratorio-de-particulas-do-mundo-usa-sistema-atualizado>. Acesso em: 21 de abril de 2018. Reproduções por outros agentes de mídia:
 - Isto É. “Maior laboratório de partículas do mundo usa sistema atualizado da UFRJ”.
 - Jornal do Brasil Online. “Maior laboratório de partículas do mundo usa sistema atualizado da UFRJ”.
 - O Dia Online. “Ciência da UFRJ é destaque no mundo”.
 - Cenário MT. “Maior laboratório de partículas do mundo usa sistema atualizado da UFRJ”.
 - Inovação Tecnológica. “LHC vai adotar software modernizado por pesquisadores brasileiros”.
- Divulgação do trabalho no jornal digital da COPPE/UFRJ. “Atlas adota sistema desenvolvido pela COPPE/UFRJ”, 2018. Disponível em: <http://www.coppe.ufrj.br/pt-br/planeta-coppe-noticias/noticias/atlas-adota-sistema-desenvolvido-pela-coppe>. Acesso em: 21 de abril de 2018. Reproduções por outros agentes de mídia:
 - FAPERJ. “Atlas adota sistema desenvolvido pela COPPE/UFRJ”.

Apêndice B

Detalhes da Infraestrutura em Atuação na Reconstrução de Física em *Offline*

O esquema da implementação completa encontra-se na Figura B.1. Apesar da complexidade, basicamente o fluxo ocorre da seguinte forma. Após a reconstrução de estados finais contendo elétrons e fótons, o algoritmo *ElectronReaderTool* procura em memória todos os candidatos a elétrons reconstruídos que executa uma sequência de ferramentas. O mesmo executa o algoritmo *ElectronRingerBuilderTool* e realiza a decoração de elétrons com essa informação. Posteriormente, a lógica segue com a execução do algoritmo de seleção de elétrons que se encontra na ferramenta *AsgElectronRingerSelectorTool*. A mesma pode empregar outras modalidades de informação além da descrição em anéis. As configurações dos parâmetros para a tomada de decisão encontram-se em arquivos *ROOT*, havendo um arquivo dedicado para os parâmetros do classificador e um outro para a descrição da região do espaço de saída onde as observações nela contida constituem a uma dada classe. Isso permite, que um mesmo classificador aplique diversos critérios de seleção. Além disso, as ferramentas de leitura de objetos físicos são manipuladas pelo algoritmo *CaloRingerAlgorithm*. A fim de realizar a análise de outros objetos físicos, basta a adição da ferramenta implementada para um objeto físico para execução nesse algoritmo. Como fótons também desenvolvem chuveiros EM, atualmente também se encontra disponível essa ferramenta para esses estados finais.

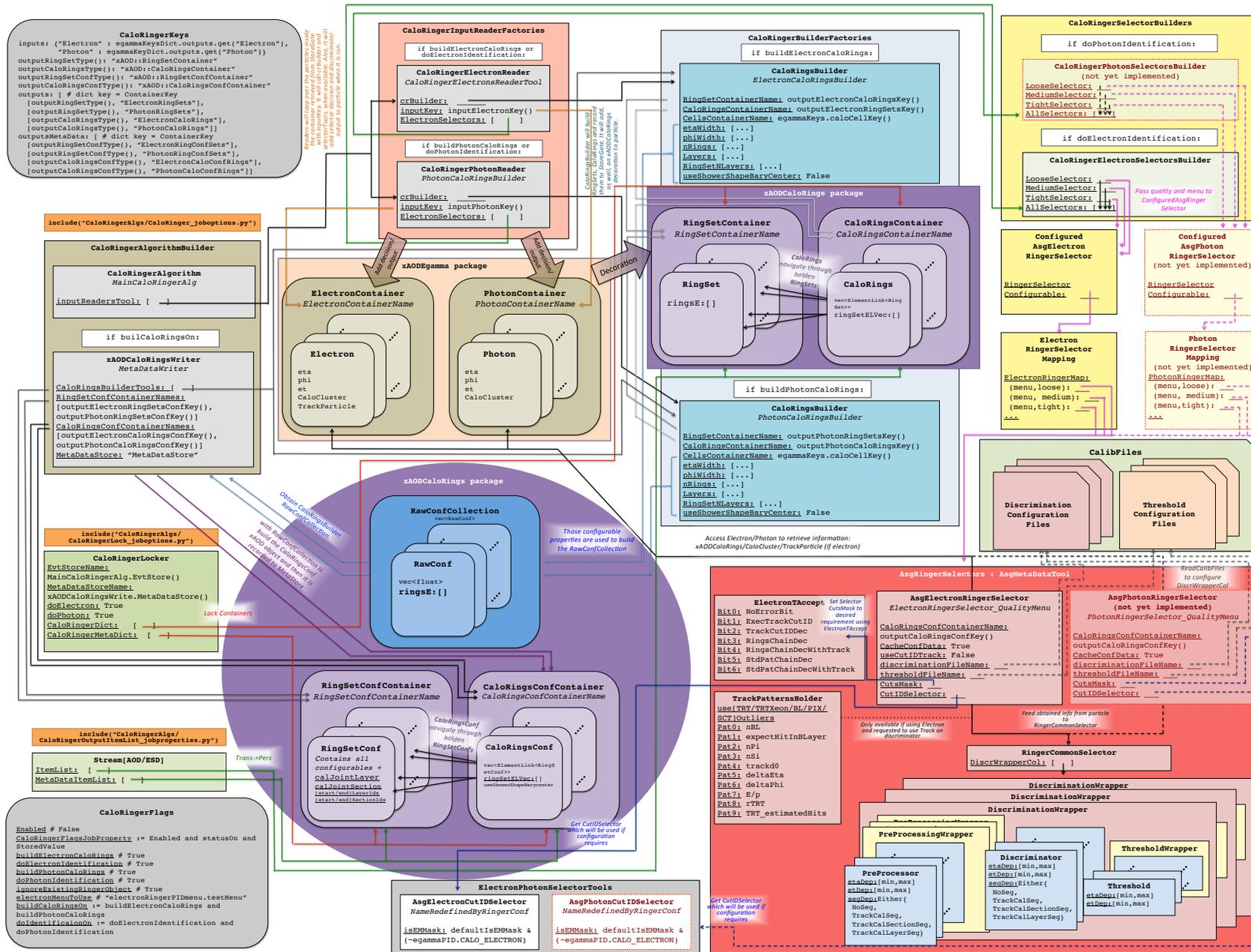
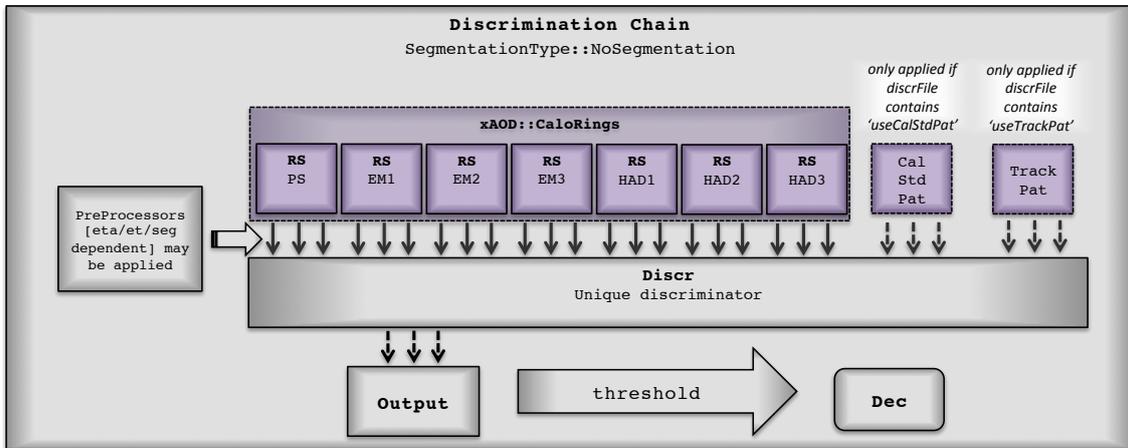
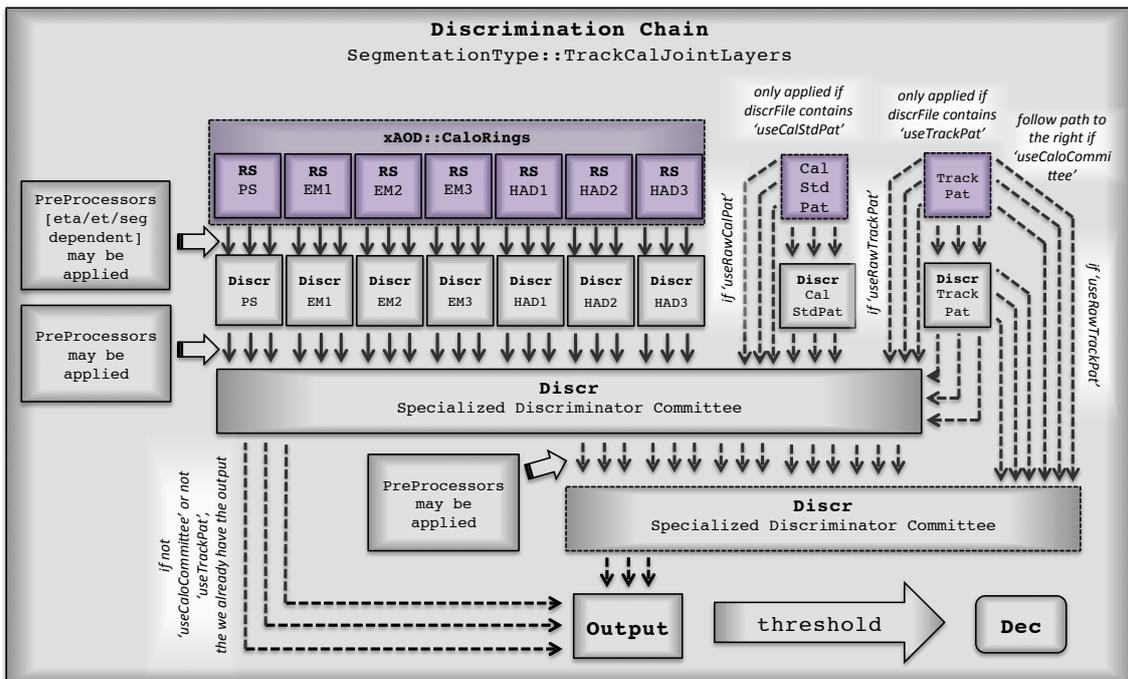


Figura B.1: Esquema da infraestrutura implementada para o algoritmo *NeuralRinger* no Sistema de Reconstrução (*Offline*) de Física do ATLAS.

Quanto à lógica de seleção, duas possibilidades para os casos extremos de fluxo de informação para comporem os espaços de entrada dos *ensemble* neurais encontram-se nas Figuras B.2a e B.2b.



(a)



(b)

Figura B.2: Fluxo de informação de classificação para o caso de aplicação menos segmentado (a) e mais segmentado (b). Os discriminadores e pré-processadores aplicados podem ser especializados em uma região do espaço de fase. A informação anelada está no objeto *xAOD::CaloRings*, que possibilita o acesso a cada uma das camadas (RS). A informação das grandezas de física calorimetria estão disponíveis no objeto *CalorimeterStandardPatterns* (CalStdPat), já as grandezas de traço e de fusão de especialista estão representadas no objeto *TrackPatterns* (TrackPat). A adição das grandezas padrões são controladas via configuração.

Apêndice C

Pseudo-código do Algoritmo RPROP

Empregou-se o RPROP (Algoritmo 1) no ajuste de todos os modelos MLP realizados nesta tese. Isso envolveu os modelos compondo os *ensembles* com MLP de uma camada escondida alimentados pela informação de calorimetria formatada em anéis atuando na filtragem *online* de elétrons (Subseção 7.1.1) ou nas redes de fusão para atuação na identificação *offline* (Subseção 7.3.2). A implementação foi realizada por [245] e a versão utilizada encontra-se em [243].

Algoritmo 1 O RPROP com os parâmetros padrões conforme estabelecido por [224]

```

1: procedure RPROP( $w, \Delta_0 = 0, 1, \eta^+ = 1, 2, \eta^- = 0, 5, \Delta_{\max} = 50, 0, \Delta_{\min} =$ 
    $1e - 6$ )
2:   for  $t \leftarrow 0$  do
3:     for all  $w$  do                                      $\triangleright w$ : matriz com as sinapses
4:                                              $\triangleright w_{ij}$ : sinapse na  $i$ -ésima linha e  $j$ -ésima coluna
5:       if  $t == 0$  then
6:          $\Delta_{ij}(t) = \Delta_0$ 
7:       end if
8:       if  $\left(\frac{\delta E}{\delta w_{ij}}(t-1) \times \frac{\delta E}{\delta w_{ij}} > 0\right)$  then            $\triangleright E$ : erro da saída
9:          $\Delta_{ij}(t) \leftarrow \min(\Delta_{ij}(t-1) \times \eta^+, \Delta_{\max})$ 
10:         $\Delta w_{ij}(t) \leftarrow -\text{sign}\left(\frac{\delta E}{\delta w_{ij}}(t)\right)\Delta_{ij}(t)$ 
11:         $w_{ij}(t+1) \leftarrow w_{ij}(t) + \Delta w_{ij}(t)$ 
12:       else if  $\left(\frac{\delta E}{\delta w_{ij}}(t-1) \times \frac{\delta E}{\delta w_{ij}} < 0\right)$  then
13:          $\Delta_{ij}(t) \leftarrow \max(\Delta_{ij}(t-1) \times \eta^-, \Delta_{\min})$ 
14:          $\Delta w_{ij}(t) \leftarrow -\text{sign}\left(\frac{\delta E}{\delta w_{ij}}(t)\right)\Delta_{ij}(t)$ 
15:          $w_{ij}(t+1) \leftarrow w_{ij}(t) - \Delta w_{ij}(t)$ 
16:       else
17:          $\Delta w_{ij}(t) \leftarrow -\text{sign}\left(\frac{\delta E}{\delta w_{ij}}(t)\right)\Delta_{ij}(t)$ 
18:          $w_{ij}(t+1) \leftarrow w_{ij}(t) + \Delta w_{ij}(t)$ 
19:       end if
20:     end for
21:     if condições de convergência atendidas then
22:       break
23:     end if
24:   end for
25: end procedure

```
