



ESTIMAÇÃO DE ENERGIA EM UM CALORÍMETRO FINAMENTE SEGMENTADO

Rodrigo Araujo Pereira

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

Orientadores: José Manoel de Seixas
Luciano Manhães de Andrade
Filho

Rio de Janeiro
Junho de 2014

ESTIMAÇÃO DE ENERGIA EM UM CALORÍMETRO FINAMENTE
SEGMENTADO

Rodrigo Araujo Pereira

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

Examinada por:

Prof. José Manoel de Seixas, D.Sc.

Prof. Luciano Manhães de Andrade Filho, D.Sc.

Prof. Amauri Oliveira, D.Sc.

Prof. Marcello Luiz Rodrigues de Campos, Ph.D.

Prof. Ronald Cintra Shellard, Ph.D.

RIO DE JANEIRO, RJ – BRASIL
JUNHO DE 2014

Pereira, Rodrigo Araujo

Estimação de Energia em um Calorímetro Finamente Segmentado/Rodrigo Araujo Pereira. – Rio de Janeiro: UFRJ/COPPE, 2014.

XVI, 116 p.: il.; 29,7cm.

Orientadores: José Manoel de Seixas

Luciano Manhães de Andrade Filho

Dissertação (mestrado) – UFRJ/COPPE/Programa de Engenharia Elétrica, 2014.

Referências Bibliográficas: p. 106 – 110.

1. Processamento de Sinais. 2. Física de Partículas.
3. Estimação de Parâmetro. I. Seixas, José Manoel de *et al.* II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Elétrica. III. Título.

À minha família.

Agradecimentos

Agradeço aos meus orientadores, Prof. Seixas e Prof. Luciano, pela orientação atenciosa, paciência e presente ao longo de todo o período de desenvolvimento deste trabalho. Também pelas valiosas discussões e pela confiança depositada em mim desde a graduação.

Agradeço também ao excepcional corpo docente e aos funcionários que formam a COPPE. A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela bolsa de estudos concedida.

Aos meus pais, José Luiz e Alzinete, por todo amor e carinho em todos os momentos de minha vida. E por terem me ensinado todo o valor da educação. Aos meus irmãos Renata e Clayton por todo apoio e boa vontade. Às minhas madrinhas Alzeny e Zilda por todo o carinho e a toda minha família.

A todos os amigos do Laboratório de Processamento de Sinais e da UFJF que me ajudaram durante o progresso deste trabalho, em especial, Bernardo, Luiz Eduardo, Luiz Évora e Natanael.

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

ESTIMAÇÃO DE ENERGIA EM UM CALORÍMETRO FINAMENTE SEGMENTADO

Rodrigo Araujo Pereira

Junho/2014

Orientadores: José Manoel de Seixas

Luciano Manhães de Andrade Filho

Programa: Engenharia Elétrica

Medir de forma precisa a energia de partículas altamente energéticas, tornou-se a principal maneira de se identificar os constituintes fundamentais que compõem a matéria. Em aceleradores de partículas como o LHC (*Large Hadron Collider*), detectores de partículas como o ATLAS (*A Toroidal LHC Apparatus*) estão localizados ao redor do ponto onde os feixes de partículas colidem e produzem uma série de partículas, algumas delas raras de se encontrar na natureza. No detector ATLAS, o responsável principal por detectar a energia das partículas que interagem de forma hadrônica, é o calorímetro hadrônico conhecido como TileCal. O calorímetro hadrônico do ATLAS (TileCal) é um calorímetro composto por mais de 10.000 canais de leitura que trabalham com uma taxa de eventos de 40 MHz. Durante o processo de leitura e digitalização da energia amostrada das partículas no calorímetro, ruídos indesejados podem ocorrer nestes canais de leitura. O ruído correlacionado entre canais, assim como canais com variância de ruído desequilibrados causam um impacto negativo na resolução de energia do TileCal. O presente trabalho propõe um método que tem por objetivo o uso de um estimador que minimiza a variância do ruído dos canais de uma célula para maximizar a relação sinal ruído da energia estimada pelas células de leitura do calorímetro, usando como vantagem o design de dupla leitura nas células do TileCal, minimizando assim o efeito do ruído na energia amostrada pelas células de leitura do calorímetro.

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

ENERGY ESTIMATION IN A FINELY SEGMENTED CALORIMETER

Rodrigo Araujo Pereira

June/2014

Advisors: José Manoel de Seixas

Luciano Manhães de Andrade Filho

Department: Electrical Engineering

Accurately measure the energy of highly energetic particles, has become the primary way to identify the principal constituents that make up matter. In particle accelerators like the LHC (Large Hadron Collider) particle detectors such as ATLAS (A Toroidal LHC Apparatus) are situated around the point where the beams collide and produce a wide amount of particles, some of them rare to find in the nature. In the ATLAS detector, the main responsible for detecting the energy of particles that interact in a hadronic form is the hadronic calorimeter called TileCal. The hadronic calorimeter of the ATLAS (TileCal) is a calorimeter composed of more than 10.000 readout channels working in a event rate of 40 MHz. Over the reading and digitalization period of the energy of the particles in the calorimeter, unwanted noise may corrupt these channels. The correlated noise between channels as well as channels with unbalanced noise variance cause a negative impact on the energy resolution of the TileCal. The dissertation presented proposes a method which aims to use an estimator that minimizes the variance of the noise of two channels from one cell to maximize the signal to noise ratio of the cells read the calorimeter, using the advantage of the design of double reading in the cells of TileCal, thus minimizing the effect of noise in the energy read by the calorimeter cells.

Sumário

Lista de Figuras	xii
Lista de Tabelas	xvi
1 Introdução	1
1.1 Tema	2
1.2 Objetivos	2
1.3 Justificativa	2
1.4 Organização da Dissertação	2
2 Física de Altas Energias	4
2.1 Física de Altas Energias	4
2.2 Aceleradores de Partículas	5
2.2.1 Aceleradores com Alvo Fixo e Colisores de Feixes	7
2.2.2 Aceleradores Lineares e Circulares	8
2.2.2.1 Linear	8
2.2.2.2 Circular	9
2.2.3 Luminosidade	10
2.3 O Modelo Padrão	12
2.3.1 Bósons	13
2.3.2 Férmions	13
2.3.3 Higgs	13
2.4 A Física além do Modelo padrão	14
3 Calorimetria	16
3.1 Calorimetria	16
3.2 Interação das Partículas com a Matéria	18
3.2.1 Excitação ou Cintilação	18
3.2.2 Ionização	19
3.2.3 Radiação Cherenkov	20
3.2.4 <i>Bremsstrahlung</i>	20
3.3 Calorímetros	21

3.4	Características e Propriedades dos Detectores	22
3.4.1	Eficiência do detector	22
3.4.2	Tempo de Resposta	22
3.4.3	Resolução Espacial	22
3.4.4	Resolução de Energia	23
3.4.5	Compensação	23
3.5	Calorímetros Eletromagnéticos e Hadrônicos	24
4	O Experimento LHC	28
4.1	CERN	28
4.2	LHC - <i>Large Hadron Collider</i>	29
4.2.1	Características Gerais do LHC	30
4.3	O Detector ATLAS	32
4.3.1	Sistema de Trigger e Aquisição de Dados	33
4.3.2	Coordenadas Geométricas do ATLAS	35
4.3.3	Sistema Magnético do ATLAS	36
4.3.4	Detectores de Traço - <i>Inner Detector</i>	37
4.3.5	O Detector PIXEL	38
4.3.6	O Detector SCT	38
4.3.7	O Detector TRT	38
4.3.8	O Espectrômetro de Múons	39
4.3.9	Calorímetro Eletromagnético	39
4.4	Calorímetro Hadrônico - TileCal	40
4.4.1	Segmentação do TileCal	42
4.4.2	Eletrônica de Leitura do TileCal	42
4.4.2.1	Fotomultiplicadores - PMT	43
4.4.2.2	Mixer	44
4.4.2.3	<i>High Voltage Dividers</i>	45
4.4.2.4	Placas 3-in-1	45
4.4.2.5	Placas Digitalizadoras	45
4.4.2.6	Placas de Interface	48
4.4.2.7	Eletrônica de Back-End	48
4.4.3	Fonte de Alimentação da Gaveta	49
4.4.4	Sistemas de Calibração do TileCal	51
5	Detecção e Estimação de Sinais	52
5.1	Histórico	52
5.2	Estimação de Parâmetros em Sinais	53
5.3	Desempenho de um estimador	55
5.3.1	Estimadores não tendenciosos	56

5.3.2	Critério da Mínima Variância	58
5.3.3	Precisão do Estimador	59
5.3.4	Limite Inferior de Cramer-Rao	60
5.4	Estimador Imparcial de Mínima Variância para o Modelo Linear . . .	63
5.4.1	Estimador de Mínima Variância para o Modelo Linear Corrompido por Ruído Colorido	64
5.5	Estimador de Máxima Verossimilhança	65
6	Redução de Ruído em Calorimetria	67
6.1	Ruído em Calorímetros	67
6.1.1	Ruído no TileCal	68
6.1.2	Ruído de Pile-up	70
6.2	Tratamento de Ruído em um Calorímetro de Argônio Líquido	71
6.3	Método do Chi-Quadrado para Descorrelação de Ruído em um Calorímetro Hadrônico	72
6.4	Correlação e Descorrelação de Ruído em Detectores Bolométricos . .	76
6.4.1	Correlação do Ruído	79
6.4.2	Descorrelação do Ruído	80
7	Método Proposto de Estimação de Energia para um Calorímetro Finamente Segmentado a Nível Celular	83
7.1	Estimação de Energia no Tilecal	83
7.1.1	Optimal Filter	84
7.2	Estimação de Energia nas Células do TileCal	85
8	Resultados	88
8.1	Resultados com Dados de Pedestal	88
8.2	Resultados com Dados de Colisão	94
8.2.1	Estimação de Energia em Células com Diferentes Algoritmos de Estimação de Amplitude dos Canais	95
9	Conclusões	104
	Referências Bibliográficas	106
A	Teoria do Limite Inferior de Cramer-Rao	111
A.1	Demonstração do Limite Inferior de Cramer-Rao para uma Função de um Parâmetro Escalar	111

B	Estimador de Mínima Variância para um Modelo Linear	114
B.1	Derivação do Estimador Imparcial de Mínima Variância para um Modelo Linear através do Teorema de Cramer-Rao	114
B.1.1	Extensão do Estimador Imparcial de Mínima Variância para um Modelo Linear Corrompido com Ruído Colorido	115

Lista de Figuras

2.1	Esquema Acelerador Linear.Fonte:[22]	9
2.2	Esquema do complexo de aceleradores do CERN.Fonte:[6]	10
2.3	Partículas do Modelo Padrão.Fonte:[24]	14
2.4	Estado atual e relações entre as Teorias da Física.Fonte:[25]	15
3.1	Calorímetro de Gelo de Laplace e Lavoisier.Fonte:[1]	17
3.2	Excitação atômica. Fonte:[2]	19
3.3	Ionização. Fonte:[3]	19
3.4	Radiação de Cherenkov e o cone formado no meio.Fonte:[4]	20
3.5	Bremsstrahlung. Fonte:[5]	21
3.6	Chuveiro EM.Fonte:[6]	25
3.7	Comparação do desenvolvimento da cascata EM e hadrônica. Fonte:[6]	26
3.8	Chuveiro hadrônico. Fonte:[7]	27
4.1	Complexo de Aceleradores do CERN. [8].	30
4.2	LHC e seus detectores [9].	32
4.3	ATLAS e seus principais detectores e dimensões.Fonte:[10].	33
4.4	Arquitetura do sistema de Trigger do ATLAS. Fonte:[10].	34
4.5	Sistema de coordenadas cartesianas do ATLAS.Fonte:[10].	36
4.6	Sistema de coordenadas cilíndricas do ATLAS.Fonte:[10].	37
4.7	<i>Barrel Toroid</i> do ATLAS.Fonte:[10].	37
4.8	Layout do <i>Inner Detector</i> do ATLAS.Fonte:[10].	39
4.9	Calorímetro Eletromagnético do ATLAS.Fonte:[10].	40
4.10	Calorímetro Hadrônico.Fonte:[10].	41
4.11	Disposição dos cintiladores e fibras óticas no TileCal.Fonte:[11].	42
4.12	Segmentação do TileCal.Fonte:[12].	43
4.13	Esquema de Leitura do TileCal.	44
4.14	Gaveta de um Módulo do TileCal.Fonte:[12]	45
4.15	Esquema de um bloco de PMT.Fonte:[12]	46
4.16	Fotomultiplicador. Fonte:[9].	46
4.17	.Resposta de uma PMT sem (esq.) e com (dir.) um mixer de luz. Fonte:[13].	47

4.18	<i>High Voltage Dividers</i> . Fonte:[13].	47
4.19	Diagrama da placa 3-in-1. Fonte:[14].	48
4.20	Diagrama da placa Digitalizadora. Fonte: [14].	49
4.21	Imagem de uma unidade de fonte de alimentação (LVPS) com a tampa removida. Fonte:[15].	50
4.22	Módulo de potência individual (versão 6.5.4). Fonte:[15].	50
5.1	Na esquerda a PDF de um estimador nao tendencioso (<i>unbiased</i>) e a direita um estimador tendencioso (<i>biased</i>). [16].	58
5.2	Em (a) temos um estimador MVU ($\hat{\theta}_3$), já em (b) não existe estimador MVU.	58
5.3	Em (a) temos um estimador MVU e ótimo, em (b) existe um estimador MVU, porém ele não é ótimo.	62
6.1	Na distribuição à esquerda, um fitting Gassiano para o ruído da célula. Na direita, o fitting considerando uma dupla-gaussiana.	69
6.2	Ruído correlacionado no módulo 48 do TileCal.	70
6.3	Relação entre o número de falhas nas LVPS e o aumento de luminosidade. Fonte:[15].	71
6.4	Grande câmara de teste com 200 chapas de aço com 1,5 mm de espessura, com lacunas de 2 mm. Fonte: [17].	72
6.5	Circuito de configuração detector-amplificador. Fonte: [17]	73
6.6	Correlação do ruído entre todos os canais do módulo LBA9 do TileCal antes da aplicação do método. Fonte: [18]	74
6.7	Correlação do ruído entre todos os canais do módulo LBA9 do TileCal após a aplicação do método. Fonte: [18]	75
6.8	Distribuição de ruído do canal 19 do módulo LBA23 do TileCal antes e depois da aplicação do método. Fonte: [18]	75
6.9	Distribuição de ruído do canal 47 do módulo LBA23 do TileCal antes e depois da aplicação do método. Fonte: [18]	76
6.10	Decaimento $0\nu\beta\beta$. Fonte: [19]	77
6.11	Torre do experimento CUORICINO dentro do sistema de resfriamento (esquerda). Uma Torre do experimento CUORE (centro). Uma matriz de torres experimento CUORE (direita). Fonte: [19] . . .	79
6.12	Espectro de potência do sinal e do ruído. Fonte: [19]	79
6.13	Correlação do ruído nos canais do experimento CUORICINO. Fonte: [19]	80

6.14	Espectro de potência original do ruído do canal 13 (linha sólida), o espectro de potência do canal 13 já descorrelacionado utilizando todos os demais canais (linha tracejada) e o espectro de potência do canal 13 descorrelacionado utilizando apenas 11 canais (linha vermelha). Fonte: [19]	82
6.15	Resolução do ruído dos canais do experimento CUORICINO. Fonte: [19]	82
7.1	Sinal característico do TileCal (em destaque, amplitudes calculadas no OF).	85
8.1	Corelação do ruído dos canais da célula da célula A9 do módulo LBA59.	89
8.2	Distribuição de ruído antes e depois da estimação de energia em célula da célula A9 do módulo LBA59.	90
8.3	Corelação do ruído dos canais da célula da célula A7 do módulo LBA23.	90
8.4	Distribuição de ruído antes e depois da estimação de energia em célula da célula A7 do módulo LBA23.	91
8.5	Corelação do ruído dos canais da célula da célula BC7 do módulo LBA51.	92
8.6	Distribuição de ruído antes e depois da estimação de energia em célula da célula BC7 do módulo LBA51.	93
8.7	RMS da distribuição do ruído das células do TileCal com e sem estimação em célula.	93
8.8	Distribuição dos pesos encontrados pelo estimador.	94
8.9	Distribuição da energia reconstruída das células do TileCal utilizando soma simples (SS).	95
8.10	Distribuição de energia na região de ruído das células do TileCal reconstruído pelos algoritmos OF2 e MF.	96
8.11	Comparação das distribuições de energia das células do TileCal utilizando diferentes algoritmos de reconstrução de amplitude.	97
8.12	Distribuição de energia da células do TileCal na região de ruído das células do TileCal reconstruído pelo método atual e utilizando estimação em célula.	98
8.13	Distribuição dos pesos calculados pelo algoritmo MVU para as células do TileCal.	98
8.14	Distribuição dos pesos calculados pelo algoritmo MLE para as células do TileCal.	99
8.15	Distribuição das variâncias dos estimadores e do limite inferior de Cramer-Rao.	100
8.16	Distribuição de energia (entre -10GeV e 5GeV) da células do TileCal reconstruído pelos métodos atuais (OF2 e MF) e utilizando estimação em célula (MVU e MLE).	101

8.17	Distribuição de energia (entre -10GeV e 65GeV) da células do TileCal reconstruído pelos métodos atuais (OF2 e MF) e utilizando estimação em célula (MVU e MLE).	102
8.18	Distribuição dos pesos calculados pelo algoritmo MVU para as células do TileCal.	102
8.19	Distribuição dos pesos calculados pelo algoritmo MLE para as células do TileCal.	103
8.20	Distribuição das variâncias dos estimadores e do limite inferior de Cramer-Rao.	103

Lista de Tabelas

2.1	Ordem de grandeza da energia mínima necessária e as ferramentas utilizadas para explorar distâncias Δx . Fonte:[6]	6
2.2	Características dos Aceleradores mais Recentes.	11
8.1	Pesos estimados para os canais da célula A9 do módulo LBA59.	89
8.2	Pesos estimados para os canais da célula A7 do módulo LBA23.	91
8.3	Pesos estimados para os canais da célula BC7 do módulo LBA51.	92
8.4	Informação dos <i>Runs</i> de colisão utilizados. Fonte: [20].	95

Capítulo 1

Introdução

O desafio de estar na fronteira de uma das mais intrigantes e complexas áreas do conhecimento humano, como a física de partículas, eleva a complexidade e o tamanho dos experimentos do *Large Hadron Collider* (LHC), localizado na Organização Européia para Pesquisa Nuclear (CERN), que se situa próxima à cidade de Genebra, na Suíça, e impele o desenvolvimento e a implementação de soluções robustas de engenharia. As condições extremas de operação dos equipamentos sujeitos à radiação e fortes campos magnéticos, juntamente com o elevado número de canais de leitura dos seus detectores e a dificuldade de manutenção periódica de alguns elementos, forçam o aperfeiçoamento dos componentes de hardware e das técnicas de software, visando garantir o bom funcionamento do experimento e a confiabilidade dos resultados esperados por ampla comunidade científica.

Em experimentos de física de altas energias, como os realizados no LHC, onde se se aceleram prótons, as partículas elementares que compõem a matéria são geradas como produto da colisão de dois feixes de partículas ou núcleos, acelerados através de fortes campos eletromagnéticos em direções opostas. Estas partículas, que são o objeto de estudo e análise, são criadas a partir da conversão da energia cinética - contida nos feixes acelerados - em massa. Uma importante maneira de as estudarmos é através da medição da energia que estas depositam em detectores posicionados ao redor do ponto de colisão. Assim, a energia é um elemento fundamental, tanto na criação, quanto na detecção das partículas nesses experimentos. A energia depositada por essas partículas nestes detectores é lida por centenas de milhares de canais e convertida em sinais digitais para serem armazenados e estudados posteriormente. Portanto, é primordial que existam ferramentas de processamento digital de sinais que contribuam para realizar uma aprimorada estimação, online e offline, dessa quantidade de energia.

1.1 Tema

À medida que a pesquisa em física de altas energias evoluiu, os experimentos necessários para seu avanço necessitaram de detectores cada vez maiores e mais segmentados, que utilizam um elevado número de canais de leitura com complexos sistemas eletrônicos expostos a extremas condições de operação. O uso de sinais digitais tornou-se fundamental para a identificação de partículas cada vez mais energéticas e raras, geradas nos experimentos. Desta maneira, os métodos de processamento de sinais digitais e estimação de parâmetros ganharam grande importância neste meio, garantindo a confiabilidade dos resultados encontrados.

1.2 Objetivos

Este trabalho tem por objetivo o desenvolvimento de uma abordagem que melhore a estimação de energia em um calorímetro utilizando conjuntos de canais que pertençam a um mesmo grupo de leitura, minimizando o ruído nesse conjunto e também identificando canais mortos no calorímetro.

1.3 Justificativa

Os diversos componentes que constituem o calorímetro em si e a cadeia de leitura dos sinais nele gerados sofrem constante degradação da radiação e são passíveis de falhas. Além disso, o acesso necessário para a manutenção destes equipamentos, durante o período de funcionamento é raro, chegando muitas vezes a apenas 1 vez por ano. Sendo assim, é importante que haja alternativas baseadas em software que minimizem o efeito das falhas e ruídos nos diversos canais de leitura do calorímetro. Melhorando assim a estimação de energia, que possui forte impacto na física e no *trigger* do detector.

1.4 Organização da Dissertação

No Capítulo 2 será apresentado um resumo da principal área de estudo deste trabalho, a física de partículas.

O Capítulo 3 apresenta uma descrição da calorimetria e das principais maneiras que podemos medir a energia das partículas. Também apresentamos os calorímetros e principais características dos detectores.

No Capítulo 4 será apresentado o CERN e os diversos detectores de partículas que o compõe. O calorímetro hadrônico do experimento ATLAS será descrito em detalhes por se tratar do ambiente onde este trabalho foi desenvolvido.

O Capítulo 5 apresenta um histórico do estudo de detecção e estimação de parâmetros em sinais. Uma ênfase maior é dada aos estimadores de parâmetros.

O Capítulo 6 apresenta a revisão bibliográfica feita no estudo de redução de ruído em calorímetros utilizando técnicas de processamento de sinais.

O Capítulo 7 apresenta o método proposto neste trabalho de estimação de energia utilizando os canais de uma célula em um calorímetro hadrônico finamente segmentado.

O Capítulo 8 apresenta os resultados obtidos com o método proposto neste trabalho.

O Capítulo 9 apresenta as conclusões e perspectivas futuras da estimação de energia em células.

O Apêndice A apresenta o desenvolvimento matemático do Teorema do Limite Inferior de Cramer-Rao.

O Apêndice B apresenta o desenvolvimento matemático do estimador imparcial de mínima variância.

Capítulo 2

Física de Altas Energias

Neste capítulo, será apresentada uma breve explicação e um histórico da física de altas energias, além de um resumo do modelo de referência, o modelo padrão.

2.1 Física de Altas Energias

A ideia de que toda a matéria do universo fosse composta por partículas mínimas e indivisíveis foi proposta pelos gregos a milhares de anos atrás. Eles deram a essas partículas o nome de átomos. Mais de dois mil anos após essa suposição, e embora a identidade dessas partículas fundamentais tenha sofrido numerosas revisões, a comunidade científica ainda acredita nela, e não mede esforços na tentativa de criar experimentos que comprovem a existência de novas partículas previstas por sofisticadas teorias físicas, que procuram descrever as forças e as partículas fundamentais da matéria.

No século XIX, os cientistas notaram que substâncias como o carbono e o oxigênio possuíam um limite mínimo para seu tamanho e atribuíram a eles o nome de átomos. O nome permaneceu, porém a ideia de que os átomos fossem indivisíveis já não existe mais. O trabalho conjunto de J. J. Thomson (descoberta do elétron), Ernest Rutherford (desenvolveu o conceito da estrutura do átomo), Niels Bohr e James Chadwick (provou a existência do nêutron). Contribuíram para o modelo do átomo como um núcleo composto por prótons e nêutrons envolvido por elétrons orbitantes, semelhante à estrutura do sistema solar. Durante algum tempo acreditou-se que os nêutrons, os prótons e os elétrons fossem as partículas fundamentais dos gregos, porém muitas questões ainda não haviam sido respondidas sobre a estrutura dos núcleos dos átomos, sua estabilidade, sua massa, entre outras. Os cientistas precisaram então desenvolver experimentos que os ajudassem a responder essas questões. Nesse contexto, surgiu a Física de Partículas, que estuda uma ampla variedade de fenômenos relacionados à estrutura da matéria e às interações de seus constituintes. As partículas e radiações são detectadas através de suas interações com a matéria.

O tipo de iteração característica dessas partículas com a matéria é a base para a construção dos detectores nesses experimentos [6]. No começo do século XX, os cientistas utilizavam pequenos detectores como, por exemplo, câmaras de nuvens e câmaras de bolhas para estudarem radiação e partículas que chegam à superfície da Terra, a maioria resultante das colisões dos raios cósmicos com as partículas de nossa atmosfera.

Devido à necessidade de se entender mais sobre a matéria, os cientistas precisaram reproduzir essas interações em ambientes controlados, surgindo então, os primeiros aceleradores de partículas. Os experimentos em física de partículas podem ser classificados nos que utilizam aceleradores e os que não utilizam aceleradores de partículas. Um exemplo de experimento que não utiliza aceleradores são os estudos de partículas resultantes da penetração de raios cósmicos em nossa atmosfera, que podem possuir centenas de vezes a energia de partículas estudadas em aceleradores porém são raros.

2.2 Aceleradores de Partículas

A necessidade do uso de aceleradores de partículas e o uso cada vez maior de energia para estudar a estrutura da matéria estão diretamente relacionados com o avanço da física no início do século XX, principalmente a mecânica quântica. A primeira razão para utilizarmos aceleradores de partículas está relacionado com a dualidade partícula-onda e ao princípio da incerteza [6].

Além das ondas, que algumas vezes possuem comportamento corpuscular, a mecânica quântica estendeu esse conceito da dualidade onda-partícula incluindo partículas que possuem um comportamento ondulatório. Com a descoberta de que a matéria possuía uma natureza dual, os cientistas foram forçados a fazer uma reavaliação da linguagem cinemática que usavam para descrever a posição e o momento linear de uma partícula. Isto porque a mecânica clássica Newtoniana, que descrevia uma partícula como um ponto no espaço usando um sistema de três coordenadas espaciais, já não poderia mais ser utilizada. Com a mecânica quântica muitos aspectos do comportamento de uma partícula só podem ser descritos em termos de probabilidades; dessa forma, uma partícula é modelada como uma entidade inerentemente espalhada que não pode ser descrita como um ponto com posição e velocidade determinadas.

Uma onda é caracterizada por seu comprimento de onda e uma partícula por sua energia e seu momento. Quanto maior a energia de uma partícula, menor o comprimento de onda associado a ela. Quantitativamente, a relação entre o comprimento de onda λ associado à partícula e o seu momento p é dado pela relação de de Broglie $\lambda = h/p = 2c\pi\hbar/pc$, onde h é a constante de Planck e $\hbar = h/2\pi$. Numericamente

temos:

$$\lambda(cm) = \frac{6.626 \times 10^{-27}(ergs)}{p} = \frac{1.24 \times 10^{-10}(MeVs)}{p(MeV/c)} \quad (2.1)$$

isto é, $\lambda(fm) = 1.24/p$ (GeV/c).

A partícula-onda pode ser vista como uma nuvem de probabilidade com dimensões comparáveis ao comprimento de onda. Nesta representação, a partícula não é um objeto com as dimensões da nuvem, na realidade, a partícula é menor e pode ser encontrada em algum ponto dentro dessa nuvem. O volume da nuvem pode ser reduzido pelo decréscimo do comprimento de onda, o que é equivalente ao aumento do momento da partícula e portanto sua energia. Sendo assim, ao aumentarmos a energia desse experimento podemos estudar objetos em dimensões cada vez menores. Em altas energias, o aspecto onda-partícula pode muitas vezes ser ignorado; a partícula pode ser considerada como uma pequena bola (ou um "spinning top", se seu *spin* for considerado). Neste caso, aceleradores são comparáveis a grandes microscópios.

O princípio da incerteza afirma que a posição x (com incerteza Δx) e momento p_x (com incerteza Δp_x) não pode simultaneamente ser conhecido com uma precisão melhor que $\Delta x \Delta p_x \approx \hbar/2$, onde Δx , Δp_x são o desvio padrão da posição e do momento, respectivamente uma relação para a energia pode ser obtida multiplicando c , $\Delta x \Delta E \approx \hbar c/2$, que, numericamente, é igual a:

$$\Delta E(MeV) \approx 1.973 \times 10^{-11}(MeVcm)/2\Delta x(cm). \quad (2.2)$$

Para explorar dimensões da ordem de Δx , uma energia de $E(MeV) \approx 2\Delta E(MeV) \approx \hbar c/\Delta x \approx 1.973 \times 10^{-11}(MeVcm)/\Delta x(cm)$ é necessária. A Tabela 2.1 lista algumas ordens de grandeza da magnitude mínima de energia necessária para explorar algumas dimensões Δx .

$\Delta x(cm)$	E	Ferramentas Utilizadas
10^{-5}	2eV	Microscópios
10^{-8}	2keV	Raios X
10^{-11}	2MeV	Raios γ
10^{-14}	2GeV	Aceleradores
10^{-16}	200GeV	Aceleradores
10^{-17}	2TeV	Aceleradores

Tabela 2.1: Ordem de grandeza da energia mínima necessária e as ferramentas utilizadas para explorar distâncias Δx . Fonte:[6]

A segunda razão para que aceleradores de partículas sejam amplamente utili-

zados em física experimental de partículas está ligada à capacidade de "criação de partículas". Esta característica está relacionada a outro importante avanço da física realizado por Einstein, que através de sua famosa equação $E = \gamma mc^2$ demonstrou que energia pode ser convertida em massa e massa em energia. Sendo assim, a energia total disponível em uma colisão de partículas em um acelerador pode se transformar em novas partículas, algumas tão instáveis que duram apenas um bilionésimo de bilionésimo de segundo antes de decaírem em partículas mais estáveis. Este processo de criação ocorre de acordo com certas regras, obedecendo precisas leis de conservação e justifica o uso de uma energia cada vez mais elevada no centro de massa de colisão dos aceleradores de partículas.

Em um acelerador de partículas, estuda-se a estrutura da matéria acelerando feixes de partículas como, por exemplo, elétrons, prótons e pósitrons a altas velocidades, utilizando poderosos campos eletromagnéticos. Quando o feixe de partículas atinge a energia desejada, ele é direcionado para colidir contra um alvo - o alvo pode ser do tipo fixo ou outro feixe de partículas acelerado em direção oposta - dependendo da proposta de estudo do acelerador. Os aceleradores também podem ser do tipo linear, por exemplo, o Stanford Linear Accelerator Center - SLAC [21], ou circular como, o Large Hadron Collider - LHC [8].

2.2.1 Aceleradores com Alvo Fixo e Colisores de Feixes

Em um acelerador que utiliza alvo fixo, um conjunto de partículas é enviado em direção a um alvo fixo para colidir com ele, o espalhamento das partículas resultantes da colisão fornece informações importantes sobre a estrutura do feixe e do alvo. Quando um feixe de partículas com massa m e energia E_{beam} colide com um alvo fixo consistindo de partículas com massa M , a energia disponível para a produção de novas partículas, nessa colisão, é igual a:

$$E_{cm} = \sqrt{m^2 + 2ME_{beam} + M^2} \quad (2.3)$$

para altas energias, $E_{beam} \gg m, M$, então E_{cm} torna-se $\sqrt{2ME_{beam}}$. Portanto, em aceleradores que utilizam alvo fixo, a maior parte da energia do feixe é usada para transferir momento para o alvo.

Em um acelerador de partículas que utiliza colisão de feixes, dois feixes contendo pacotes de partículas são acelerados e direcionados para colidirem com igual momento e sinais opostos, em pontos específicos. Esta geometria possui a vantagem de utilizar melhor a energia envolvida na colisão. Quando duas partículas idênticas, com energia E_{beam} cada, viajando em direções opostas, colidem entre si, a referência do centro de massa é a referência do laboratório. Neste caso, a energia disponível para a produção de novas partículas é igual ao total de energia envolvida na co-

lisão, ou seja, $2E_{beam}$. Notadamente, o método de colisão entre feixes de partículas tornou-se a geometria mais favorável para o estudo de novas partículas em energias elevadas. Nos anos 1960, experimentos pioneiros de colisão entre feixes colidiam elétrons e pósitrons. Entre esses experimentos destacam-se experimentos como o ADONE (Frascati, Itália), DORIS (DESY, Hamburg) e SPEAR (SLAC, Stanford).

2.2.2 Aceleradores Lineares e Circulares

Partículas carregadas eletricamente, usualmente elétrons e prótons, são aceleradas através de campos elétricos e magnéticos que podem ser constantes ou variáveis no tempo e espaço. Um acelerador é basicamente constituído de uma fonte de íons, um campo acelerador e um campo guia que força a partícula a se mover em uma órbita bem definida. O caminho por onde a partícula é acelerada se dá em uma região com alto vácuo a fim de reduzir as colisões das partículas com moléculas de gás residual e, portanto, reduzir a perda de energia ao longo do caminho. A fonte de íon consiste de uma pequena cavidade à baixa pressão, onde uma descarga elétrica ioniza continuamente gás hidrogênio então um campo elétrico é utilizado para carregar os elétrons e prótons através das próximas etapas do acelerador. Aceleradores podem ser classificados em lineares (conhecidos como LINAC) ou circulares. Em aceleradores lineares, as partículas são aceleradas em linha reta por campos elétricos. Em aceleradores circulares, um campo magnético é utilizado para forçar as partículas a circularem em uma órbita circular e a aceleração é realizada por um campo elétrico de rádio frequência, ou através do aumento no campo magnético.

2.2.2.1 Linear

Aceleradores lineares de prótons de baixa energia consistem de um conjunto de eletrodos cilíndricos conectados entre si e a uma tensão oscilante com frequência constante, como mostrado na Figura 2.1. Os prótons se movem da esquerda para a direita dentro de uma estrutura feita de sucessivos tubos lineares com campos de sinais opostos (o próton sempre vê uma tensão $V+$ no tubo a sua frente e uma tensão $V-$ no tubo por onde passou). Neste caso, os prótons podem atingir energias de até 20 MeV em aproximadamente 70 metros. Dentro dos tubos, os prótons estão em região livre de campo elétrico e, portanto, se movem com uma velocidade constante, já entre tubos vizinhos, as partículas são aceleradas por um campo elétrico. Se o comprimento do tubo for escolhido de forma que o tempo necessário para a partícula cruzá-lo seja igual a metade do período do oscilador, os prótons são acelerados uma segunda vez quando passarem para o próximo tubo, e assim sucessivamente eles serão acelerados. Os comprimentos dos tubos devem, portanto, aumentar de forma a levar em conta o aumento da velocidade do próton. No caso dos aceleradores

lineares de elétrons, os eletrodos são todos do mesmo comprimento, pois os elétrons, mesmo com baixa energia, viajam na velocidade da luz. A aceleração é realizada por ondas eletromagnéticas viajando ao longo do eixo do tubo com a fase da velocidade igual à velocidade do elétron. Somente elétrons em fase com o campo são acelerados. Aceleradores lineares possuem grande utilidade na medicina, por exemplo, na radioterapia que utiliza aceleradores lineares de feixes de elétrons e fótons de alta energia no tratamento de tumores.

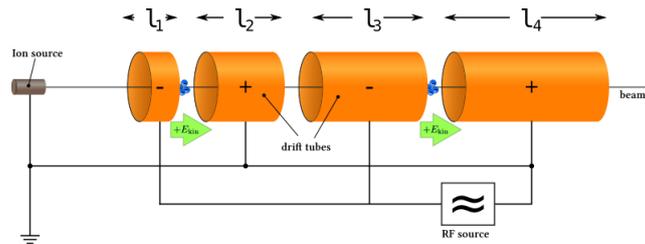


Figura 2.1: Esquema Acelerador Linear. Fonte: [22]

2.2.2.2 Circular

Aceleradores circulares do tipo *Synchrotrons* superaram outros aceleradores, por exemplo, *cyclotrons* e *synchrocyclotrons*. Nos aceleradores circulares do tipo *Synchrotrons*, o uso de grandes e caros campos magnéticos é desnecessário: as partículas são aceleradas através de campos elétricos em uma órbita de raio constante. O campo magnético necessário para defletir as partículas é produzido por um conjunto de magnetos localizados ao longo da circunferência do acelerador. Antes de serem injetadas no acelerador de raio maior, as partículas são aceleradas previamente por um acelerador eletrostático seguido por um acelerador linear. As partículas são aceleradas por campos elétricos de rádio frequência em cavidades ressonantes localizadas ao longo da circunferência do acelerador. Os *Synchrotrons* de elétrons possuem um RF constante, uma vez que elétrons se movem a velocidade constante e próxima a da luz no vácuo.

A Figura 2.2 mostra os diversos aceleradores utilizados no complexo de aceleradores no CERN para a colisão de elétrons (e^-) e pósitrons (e^+) no LEP e próton-próton (pp) no LHC. Cada acelerador do complexo eleva a energia das partículas por uma ordem de grandeza.

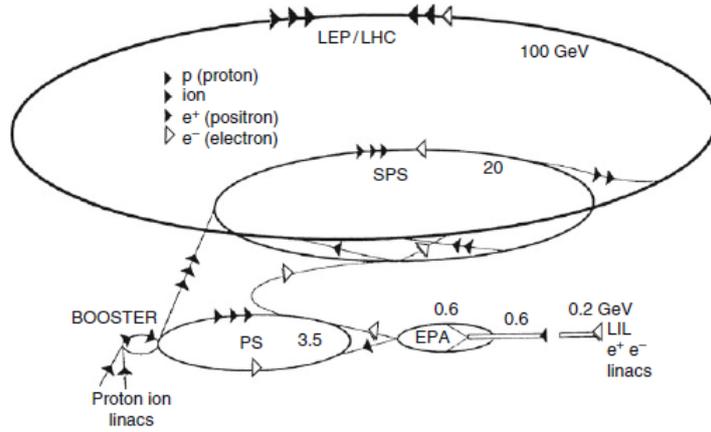


Figura 2.2: Esquema do complexo de aceleradores do CERN. Fonte: [6]

2.2.3 Luminosidade

Para obter altas energias no centro de massa, é essencial o uso de aceleradores circulares. Porém, colisores de feixes apresentam uma desvantagem em relação a aceleradores de alvo fixo, por produzirem um número menor de interações por unidade de tempo. A luminosidade (\mathcal{L}) de um colisor é definida como o número que, multiplicado pela seção total de um dado processo (σ), fornece o número total de colisões por unidade de tempo (N), isto é

$$N = \mathcal{L}\sigma. \quad (2.4)$$

A luminosidade pode ser expressa através de parâmetros do colisor. Seja R o raio do acelerador, n_p o número de partículas em cada pacote circulando em uma direção, N_p o número de pacotes, seja n_a o número de partículas em sentido oposto e N_a o número de pacotes de partículas em direção oposta, e r a média do raio transversal de cada pacote. Então a luminosidade pode ser descrita pela equação abaixo,

$$\mathcal{L} = \frac{fn_p n_a N_p N_a G}{4\pi r^2} \quad (2.5)$$

sendo f a frequência de circulação dos pacotes no acelerador, $f = 1/\tau$, onde τ é o período de revolução, $\tau = \frac{2\pi R}{c}$. O fator G leva em conta o comprimento finito de um pacote e usualmente é $G \approx 1$. Em 27 de janeiro de 1971, dois feixes de prótons colidiram pela primeira vez no primeiro colisor de prótons circular da história, o ISR no CERN. O ISR alcançou uma luminosidade de $2 \times 10^{31} \text{cm}^{-2} \text{s}^{-1}$, uma vez que a seção transversal próton-próton no ISR era de $5 \times 10^{-26} \text{cm}^2$, cerca de 10^6 interações

por segundo foram produzidas em cada região de interação.

Em geral, a maioria dos eventos produzidos nas interações não são de interesse, os chamados eventos de *minimum bias*. Por esta razão, atingir uma alta luminosidade é um dos principais objetivos nesse tipo de experimento, muitas vezes tão importante quanto atingir um elevado nível de energia. O LHC foi desenvolvido para atingir uma luminosidade recorde da ordem de 10^{32} a 10^{34} ($cm^{-2}s^{-1}$). Espera-se que através das melhorias técnicas que serão realizadas no LHC a luminosidade total de 10^{34} seja atingida a partir de 2015. A Tabela 2.2 mostra alguns dos principais aceleradores construídos nas últimas décadas e suas respectivas características como: energia no centro de massa, tipo de feixes e luminosidade.

Acelerador	Ano	Feixes	Energia(GeV)	$\mathcal{L}(cm^{-2}s^{-1})$
SPPS(CERN)	1981	$p\bar{p}$	630 – 900	6×10^{30}
Tevatron(FNAL)	1987	$p\bar{p}$	1800 – 2000	10^{31-32}
SLC(SLAC)	1989	e^+e^-	90	10^{31-32}
LEP(CERN)	1989	e^+e^-	90 – 200	10^{31-32}
HERA(DESY)	1992	ep	300	10^{31-32}
RHIC(BNL)	2000	pp, AA	200 – 500	10^{32}
LHC(CERN)	2010-2012	pp, pA, AA	$7K - 8K$	$\sim 0.76 \times 10^{34}$
SLHC(CERN)	2015	pp, pA, AA	$14K$	1×10^{34}

Tabela 2.2: Características dos Aceleradores mais Recentes.

Ao longo do século XX, o estudo das partículas resultantes dessas colisões permitiu aos cientistas a descoberta de diversas partículas elementares. Em 1968, experiências realizadas no *Stanford Linear Accelerator Center* para pesquisar a natureza microscópica da matéria [23] revelaram que os prótons e nêutrons não eram partículas indivisíveis, e sim formadas por três partículas ainda menores chamadas quarks, nome fictício retirado de uma passagem do livro *Finnegans Wake* de James Joyce, e dado pelo físico teórico Murray Gell-Mann que, anteriormente, já havia previsto teoricamente a existência desta partícula. Os estudos mostraram ainda que prótons e nêutrons são constituídos por dois tipos de quarks diferentes: *up* e *down*. O próton é formado por dois quarks *up* e um *down*, e o nêutron é formado por dois quarks *down* e um *up*. Tudo o que conhecemos no mundo e no universo parece ser formado pelas combinações de elétrons, quarks *up* e *down*. Até hoje, não há nenhuma indicação de que estas partículas sejam formadas por outras partículas ainda menores.

Os físicos continuam a provocar choques entre partículas usando novas tecnologias e cada vez mais energia, recriando condições nunca vistas antes. Dos resultados

dessas colisões, uma série de novas partículas fundamentais foi descoberta, fazendo parte de uma teoria conhecida como Modelo Padrão.

2.3 O Modelo Padrão

Um problema importante na física é entender os constituintes fundamentais da matéria e as interações fundamentais que ocorrem entre eles. Saber se estas interações são, na realidade, diferentes manifestações de uma única interação subjacente, que ainda tem de ser descoberta também é de interesse. Esta unificação das leis físicas, conhecida como Teoria de Tudo, é um dos grandes desafios da física - a questão da unificação foi abordada por Einstein no início do século passado embora ele não tenha alcançado uma conclusão.

O Modelo Padrão da Física de partículas, desenvolvido na década de 1970, é uma teoria quântica de campos que descreve os constituintes fundamentais da matéria e as interações entre eles. O modelo padrão separa as partículas em dois tipos fundamentais: os bósons e os férmions. São 12 constituintes fundamentais fermiônicos, sendo eles 6 léptons com *spin* 1/2 e energia do estado fundamental negativa, e 6 quarks com *spin* 1/2 e carga do estado fundamental positiva. Se incluirmos suas respectivas anti-partículas o número de férmions sobe para 24.

Quatro interações fundamentais foram identificados na escala de energia comum dos laboratórios até então, são elas: fraca, eletromagnética, forte e gravitacional. As interações fracas e eletromagnéticas foram unificados na teoria eletrofraca do Modelo Padrão - *Quantum ElectroDynamics* (QED) - e, juntamente com a teoria da Cromodinâmica Quântica - *Quantum ChromoDynamics* (QCD) - que descreve as interações fortes, formam o Modelo Padrão das interações eletrofraca e forte (*Standard Model* em inglês). Tanto a teoria QED quanto a QCD são teorias de calibre, isto é, elas modelam as forças entre férmions acoplando aos bósons responsáveis pelas forças. Cada Lagrangiana representante de um conjunto de bósons é invariante sob uma transformação de calibre (*gauge*), assim estes bósons mediadores são conhecidos como bósons de calibre. É comum descrever as transformações de *gauge* dos bósons de calibre utilizando um grupo unitário chamado grupo de calibre; o grupo de calibre da interação forte é o $SU(3)$, o da interação eletrofraca é o $SU(2) \times U(1)$. Sendo assim, muitas vezes o Modelo Padrão é representado utilizando-se esses grupos unitários como $SU(3) \times SU(2) \times U(1)$. As partículas que compõem o Modelo Padrão podem ser vistas na Figura 2.3.

2.3.1 Bósons

Os bósons possuem spin inteiro e são responsáveis por transmitir as interações na natureza, são eles: o fóton, mediador da interação eletromagnética; os bósons W^+ , W^- e Z que mediam a interação fraca¹; os oito tipos de glúons, responsáveis pela força forte²; e o bóson de Higgs, mais conhecido por ser o constituinte do campo de Higgs, responsável pela existência da massa inercial das demais partículas. Os Grávitons, bósons que acredita-se mediar a interação gravitacional, não são explicados no modelo padrão.

2.3.2 Férmions

Os férmions são partículas com spin semi-inteiro e basicamente correspondem às partículas que descrevem a matéria. Os férmions podem ser divididos em duas famílias: os quarks e os léptons, com três gerações de partículas. Como pode ser visto na Figura 2.3, existem seis quarks, que também são divididos em três gerações: *up/down*, *charm/strange* e *top/bottom*. Os léptons são: elétron, múon e tau, além de seus respectivos neutrinos. Toda a matéria hadrônica no universo é composta pela primeira geração de férmions, a segunda e terceira geração dos léptons são compostas por partículas raras e instáveis, que decaem rapidamente em partículas da primeira geração.

2.3.3 Higgs

Além dos constituintes fermiônicos e dos bósons que mediam as interações fundamentais, o Modelo Padrão requer a existência de um campo de Higgs escalar, necessário para explicar o mecanismo de quebra espontânea da simetria eletrofraca através do qual os férmions e os bósons intermediários W^+ , W^- e Z adquirem massa. Segundo a teoria, as partículas massivas adquirem sua massa através da interação delas no campo de Higgs, quanto maior o acoplamento da partícula com o campo de Higgs maior a massa gerada para a partícula.

A existência do campo de Higgs pôde ser comprovada através da procura de uma partícula correspondente associada a ele, o "bóson de Higgs". A detecção de bósons de Higgs provaria automaticamente que o campo de Higgs existe, contribuindo assim para a confirmação da Teoria Eletrofraca.

¹Força mais conhecida por ser responsável pela desintegração radioativa de alguns elementos, como o urânio.

²Força que mantém os quarks do tipo *up* e *down* unidos para formarem nêutrons e prótons.

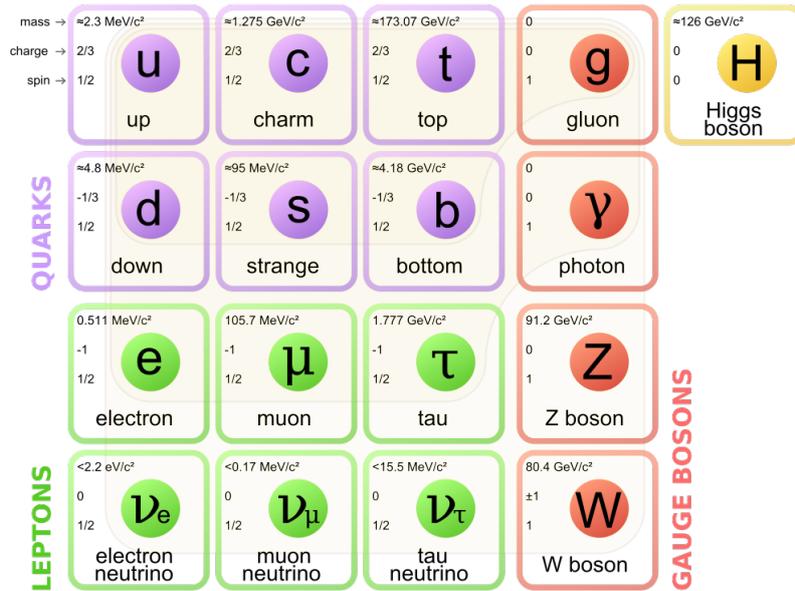


Figura 2.3: Partículas do Modelo Padrão. Fonte: [24]

2.4 A Física além do Modelo padrão

Apesar da Teoria do Modelo Padrão ser extraordinariamente bem sucedida, quando comparada com os resultados experimentais que temos hoje, existem indícios de que exista uma física além do Modelo Padrão. O modelo padrão possui algumas lacunas e não explica uma série de questões, como:

- Por que a existência de 3 famílias de quarks e léptons?
- Por que as cargas do elétron e do próton possuem o mesmo valor? E ainda, o porquê da carga elétrica ser quantizada.
- Por que a existência de 4 forças? Será que elas são manifestações de uma única força?
- Por que a assimetria da matéria e antimatéria?
- Qual a natureza da matéria escura e da energia escura?
- O Gráviton realmente existe?

A Figura 2.4 apresenta um organograma com teorias físicas já comprovadas (e seus respectivos proponentes) e também algumas teorias candidatas a explicar e unir as teorias já existentes. Todas essas questões ainda não respondidas tornam o estudo de física de partículas um campo ainda aberto e repleto de desafios interessantes.

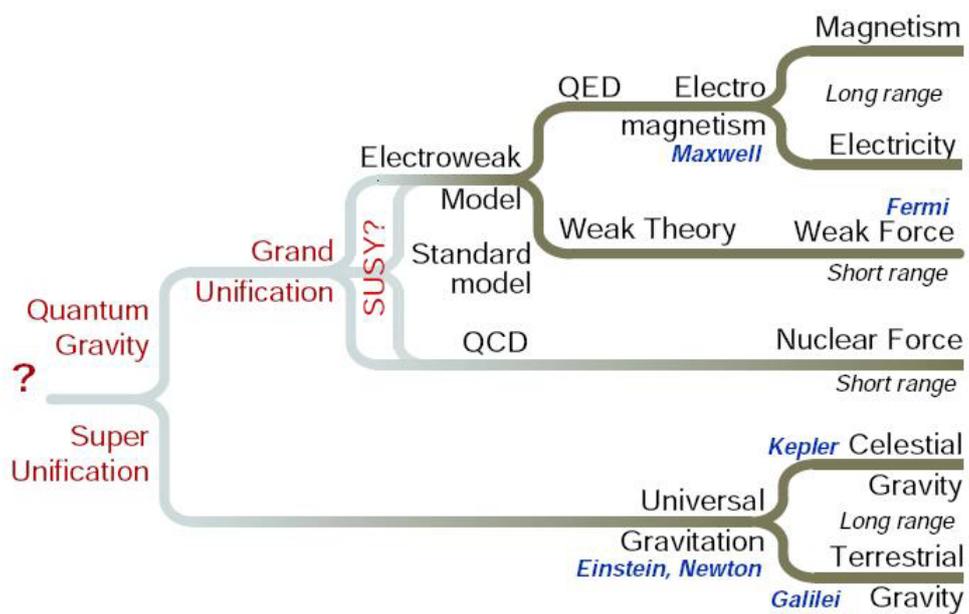


Figura 2.4: Estado atual e relações entre as Teorias da Física. Fonte: [25]

Capítulo 3

Calorimetria

Neste capítulo, será apresentada uma introdução ao estudo de calorimetria de altas energias, assim como as principais formas de interação entre as partículas elementares e a matéria. Também serão apresentadas algumas das principais características dos calorímetros.

3.1 Calorimetria

O termo calorimetria foi criado para se referir ao estudo da troca de energia entre corpos que se dava na forma de calor, e tem sua origem na termodinâmica. A essa transferência de energia térmica entre um corpo para o outro corpo dá-se o nome de calor. O estudo da troca de calor entre corpos teve seu início a partir do momento em que a ciência distinguiu de forma clara o conceito de calor e temperatura, antes disso existiam teorias que tentavam explicar a temperatura dos corpos, uma delas era a Teoria Calórica que acreditava que o calor era uma substância química que fluía através dos corpos. Essa distinção entre temperatura e calor ocorreu ainda no fim do século 18 com Joseph Black, cientista inglês conhecido por identificar o gás dióxido de carbono, e seus estudos pioneiros em calorimetria que introduziram o conceito de calor latente de vaporização e o conceito geral de capacidade calorífica (ou calor específico). Entre muitos estudos posteriores a respeito do calor, o mais conhecido e celebrado foi o trabalho conjunto de dois dos mais eminentes cientistas da época, Antoine Laurent Lavoisier e Pierre Simon Laplace, apresentado na *Académie Royale des Sciences* em Junho de 1783, onde eles apresentaram suas primeiras medições de calor feitas então no inverno anterior. Em 1784, eles publicaram a descrição do celebrado calorímetro de gelo desenvolvido por eles e de sua utilidade em uma variedade de estudos. O calorímetro desenvolvido por eles (ver Figura 3.1) é um receptáculo dividido em três seções, duas seções externas onde colocava-se gelo e uma seção interna onde se encontrava o objeto de estudo à alta temperatura, previamente aquecido. Quando o objeto com alta temperatura chegasse a uma tem-

peratura de equilíbrio com o restante do sistema, certa quantia de água formada pelo derretimento do gelo era drenada pela torneira que se encontrava abaixo do calorímetro, essa quantidade de água era pesada posteriormente e conhecendo-se o calor de fusão do gelo, o calor liberado por esse processo podia ser calculado pela massa de gelo fundido durante o processo.

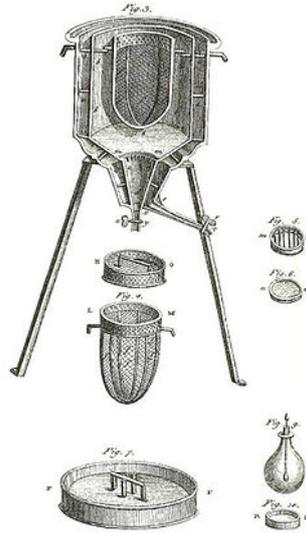


Figura 3.1: Calorímetro de Gelo de Laplace e Lavoisier.Fonte:[1]

Nos experimentos de física de altas energias, a energia de partículas e a medição de suas características são determinadas através da calorimetria. A absorção da energia da partícula é feita em um instrumento conhecido como calorímetro. Nos experimentos modernos, as partículas mais energéticas possuem quantidade de energia da ordem de TeV ($1TeV = 10^{12}eV$), por isso, o termo Altas Energias é apropriado para descrever esses experimentos. Dada a dificuldade em medir essa quantidade de energia diretamente, métodos sofisticados de detecção são necessários para ler e determinar as propriedades dessas partículas. Os experimentos modernos exigem que estes calorímetros sejam cada vez mais precisos e segmentados, possuindo uma grande quantidade de canais de leitura, os quais devem operar simultaneamente a uma alta taxa de eventos.

O estudo dos processos de interação entre a partícula incidente e a matéria é de fundamental importância, uma vez que as suas características estão na base dos métodos de detecção de partículas.

3.2 Interação das Partículas com a Matéria

Do ponto de vista epistemológico, observar ou detectar um objeto nada mais é do que detectar uma onda eletromagnética ou sonora refletida por esse objeto. Os objetos que observamos todos os dias são visíveis graças à luz refletida por eles. Sabemos que a luz nada mais é do que uma componente da radiação eletromagnética que pode ser detectada pelo olho humano. As dimensões das partículas físicas nesses experimentos, porém, são tais que as ondas eletromagnéticas não são perturbadas por elas, o comprimento de onda da luz visível estende-se do violeta ($400nm$) ao vermelho ($700nm$) e a dimensão de um átomo é de aproximadamente 0.1 nm . Portanto, percebemos que a única maneira de detectar tais partículas é através da radiação emitida quando de sua interação com a matéria. Uma partícula carregada com alta velocidade que se mova através de um meio perde energia quase que constantemente e é ligeiramente desviada de sua direção inicial. Estes efeitos são resultados de dois tipos de colisões que podem ocorrer entre a partícula e o meio:

- Colisão Inelástica, principal forma de perda de energia da partícula incidente, ocorre quando esta colide com os elétrons dos átomos do meio, particularmente com os de órbita mais externa, este tipo de colisão é responsável pela ionização e/ou excitação dos átomos do meio;
- Colisão elástica com o núcleo, menos frequente e que, na prática, não causa perda de energia mas sim o desvio da direção da partícula incidente;

Estas colisões ocorrem várias vezes ao longo do caminho percorrido pela partícula incidente no meio; a energia perdida em cada colisão representa uma pequena fração da energia cinética da partícula (algumas dezenas de eV), porém, devido ao elevado número de colisões que ocorrem em um meio denso, isto faz com que a energia total perdida pelas partículas no meio se torne mensurável.

Os principais mecanismos de interação das partículas com a matéria utilizados nos detectores são basicamente de dois tipos: cintilação (ou excitação) e a ionização. A radiação de Cherenkov também é utilizada em alguns tipos de detectores. Além disso, para partículas relativísticas mais leves, principalmente elétrons, o processo de perda de energia conhecido como *bremstrahlung* também pode ser considerado.

3.2.1 Excitação ou Cintilação

Na excitação, representada na Figura 3.2, parte da energia das partículas carregadas que atravessam a matéria do detector pode ser usada para ionizar o estado dos átomos ou moléculas do material do qual o detector é composto, ou levar estes átomos e moléculas a um estado de excitação. O efeito de cintilação está relacionado

a este último processo. O estado de excitação desses átomos e moléculas é instável, a excitação não consiste na remoção do elétron do átomo, apenas na sua elevação a um nível energético maior (estado excitado), saindo, portanto, do nível fundamental. Ao retornar ao nível fundamental, a energia perdida pelo elétron é emitida na forma de fótons que podem ser transformados em sinal elétrico. Quando a energia envolvida no processo é tal que os fótons emitidos estão na região do visível, este processo é chamado de fluorescência ou cintilação. Historicamente, a cintilação foi o primeiro processo utilizado para a geração de sinais em calorímetros.

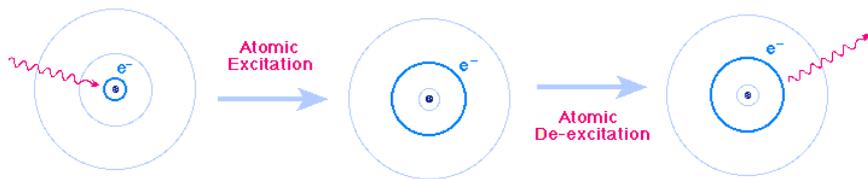


Figura 3.2: Excitação atômica. Fonte:[2]

3.2.2 Ionização

A ionização, vista na Figura 3.3, ocorre quando partículas carregadas que atravessam a matéria ionizam os átomos do qual essa matéria é feita. As radiações ionizantes são aquelas que possuem energia acima da energia de ligação dos elétrons do átomo com o núcleo. Dessa forma, a energia das radiações ionizantes é capaz de arrancar elétrons de seus orbitais, formando, assim, íons positivos. A coleção desses elétrons liberados pode ser usada como uma técnica de produção de sinal em uma grande variedade de detectores de partículas.

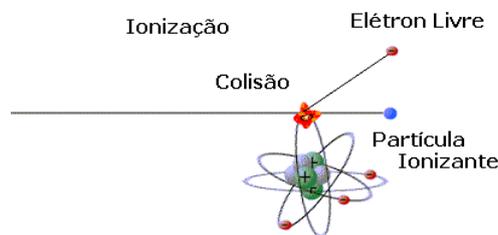


Figura 3.3: Ionização. Fonte:[3]

3.2.3 Radiação Cherenkov

A radiação de Cherenkov é uma radiação eletromagnética emitida quando uma partícula carregada atravessa um meio com velocidade maior do que a luz nesse determinado meio. As partículas carregadas polarizam as moléculas do meio que voltam para o seu estado fundamental rapidamente emitindo radiação no processo, o espectro visível da radiação exibe a característica de uma luz de tonalidade que depende do meio. Esta luz pode ser observada em abundância em ambientes altamente radioativos, como, por exemplo, em tanques de reatores nucleares. Outro local onde este fenômeno pode ser observado é nas profundezas do oceano, onde a luz azul observada é fruto da radiação causada por raios cósmicos penetrantes, uma ilustração deste fenômeno pode ser visto na Figura 3.4. A radiação de Cherenkov é utilizada em muitos experimentos, como, nos detectores localizados no observatório Pierre Auger na Argentina [26], que estuda raios cósmicos altamente energéticos. Outra aplicação interessante da radiação de Cherenkov é na detecção de neutrinos, como utilizado no detector de neutrinos *Super-Kamiokande* localizado no interior do monte *Ikenoyama* no Japão, onde 11 mil fotomultiplicadores são capazes de detectar e distinguir neutrinos do elétron e do múon [6]. Nesses experimentos a medição da velocidade da partícula pode ser obtida medindo-se o ângulo do cone de Cherenkov, da mesma forma como podemos medir a velocidade de uma embarcação através da esteira de onda formada por ele.

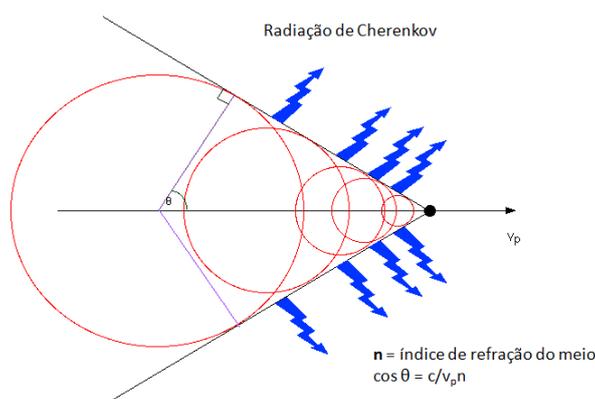


Figura 3.4: Radiação de Cherenkov e o cone formado no meio. Fonte:[4]

3.2.4 *Bremsstrahlung*

Este fenômeno, ilustrado na Figura 3.5, trata-se da emissão de um fóton a partir de um elétron defletido por um núcleo. É um processo que ocorre devido à interação

eletromagnética que produz uma grande perda de energia. O processo de *bremss-
trahlung* é a forma mais comum de perda de energia para partículas com velocidades
relativísticas.

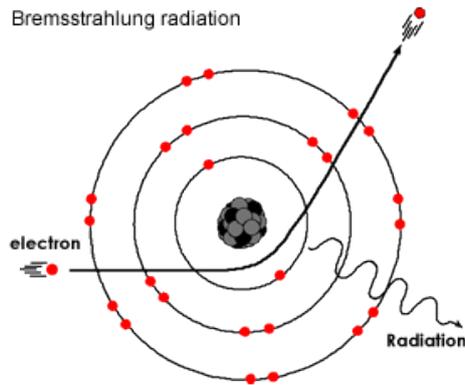


Figura 3.5: Bremsstrahlung. Fonte:[5]

3.3 Calorímetros

Calorímetros são usados, principalmente, para determinar a energia da partícula incidente, embora também se extraia a posição espacial e a direção de movimento, se o calorímetro for suficientemente bem segmentado. Em um calorímetro, a energia total da partícula é absorvida por um bloco de matéria e medida através de ionização ou excitação.

Em calorímetros eletromagnéticos a energia é transportada em uma cascata de elétrons, pósitrons e fótons. Por outro lado, em um calorímetro hadrônico, a energia se desenvolve através de uma cascata de píon-nucleon (prótons, nêutrons e também uma parte eletromagnética). Na cascata eletromagnética toda a energia de entrada é convertida em energia visível para o detector, que é em grande parte depositada por elétrons e pósitrons. Numa cascata hadrônica, apenas uma parte da energia que entra no calorímetro é vista como excitação ou ionização devido ao fato de a energia de ligação de um núcleo não ser insignificante no que diz respeito às energias das partículas e que algumas das partículas produzidas como múons, nêutrons e neutrinos escapam o volume fiducial de um calorímetro. Essa quantidade de energia invisível varia e é a causa de grandes flutuações nas medições da energia nos calorímetros Hadrônicos.

3.4 Características e Propriedades dos Detectores

Cada detector é sensível a um tipo de radiação e possui características específicas. Cada detector é projetado para atender os objetivos e requerimentos do experimento em questão. Não existe detector sensível a todo o tipo de radiação e energia. Os principais parâmetros que definem um calorímetro serão explicados a seguir.

3.4.1 Eficiência do detector

A eficiência do detector (ϵ) é a probabilidade com que o detector registra a radiação incidente nele, isto é, o quanto da radiação incidente será transformada em um pulso elétrico. Esta medida é dada pela razão entre o número de eventos gravados N_{rec} e o número de partículas N que atravessaram o detector, $\epsilon = N_{rec}/N$ ($0 \leq \epsilon \leq 1$). Seu valor é geralmente obtido através de métodos de simulação, como Monte Carlo, mas também pode ser medido usando-se um feixe de partículas conhecido.

3.4.2 Tempo de Resposta

Esta característica está conectada ao tempo intrínseco que o detector precisa para formar um sinal eletrônico, após a chegada da radiação, excluindo os atrasos incorporados por componentes eletrônicos e cabos. Para se obter uma boa resposta, o tempo de subida do pulso deve ser o mais breve possível. O tempo de resposta é usualmente uma distribuição gaussiana e a metade da largura à meia altura dessa distribuição pode ser considerada como a resolução do tempo. As resoluções de tempo estão abaixo de $1ns$ para cintiladores e detectores de Cherenkov; para alguns detectores a resposta ao tempo não é definida. A duração da formação do sinal é importante, pois durante esse tempo, um segundo evento não pode ser gravado. O *dead time* é o tempo entre a passagem de uma partícula e o momento em que o detector está pronto para gravar um novo evento - durante esse tempo o detector não está sensível a radiações. O comprimento do sinal, a eletrônica usada na leitura e o tempo de recuperação do detector influenciam no período de *dead time*.

3.4.3 Resolução Espacial

É uma figura de mérito importante que mede a precisão da reconstrução da posição da partícula carregada pelo calorímetro. Este parâmetro varia de cerca de $1\mu m$ para detectores de emulsão nuclear, $5 - 10\mu m$ para detectores de silício e até alguns centímetros para contadores de Cherenkov.

3.4.4 Resolução de Energia

Resolução de energia é considerada uma das principais características de performance de um calorímetro. A resolução de energia está relacionada com a capacidade do detector de distinguir duas energias próximas; este parâmetro determina a precisão que a energia desconhecida de uma partícula pode ser medida. Isto é determinado, experimentalmente, através da precisão com a qual a energia de partículas com energia conhecida são medidas no calorímetro. Nesses experimentos, a resolução de energia pode ser o principal fator que limita a precisão com que a massa da partícula pode ser determinada e o limite de separação entre partículas com massas semelhantes. Diversos fatores podem limitar a resolução de energia do calorímetro, dentre os quais o ruído eletrônico do sistema de leitura.

Uma parametrização geral da resolução de um calorímetro é dado pela Equação 3.1, onde o primeiro termo a é o termo estocástico, o termo b é o termo de ruído, devido ao ruído eletrônico e de *pile-up* e o último termo c é o termo constante devido a má inter-calibração entre as células e não-linearidades [27].

$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c. \quad (3.1)$$

3.4.5 Compensação

A resposta dos calorímetros para hádrons, em geral, é menor do que a resposta para elétrons com mesma energia. A razão de isto ocorrer se deve ao fato de que na absorção de chuveiros hadrônicos, uma fração significativa da energia transportada pelo chuveiro de partículas é invisível, ou seja, não contribui para o sinal de calorímetro. A principal fonte de energia invisível é a energia usada para liberar prótons e nêutrons dos núcleos. Contribuições adicionais menores vêm de neutrinos e múons que escapam ao detector. Algumas das consequências da não compensação dessa energia são: não-linearidade na resposta do calorímetro ao chuveiro hadrônico; degradação da resolução de energia; e também efeitos na forma do sinal do calorímetro hadrônico. O termo compensação tem sua origem nas primeiras tentativas de contornar os problemas mencionados. Ao aumentar seletivamente a energia depositada no calorímetro para chuveiros hadrônicos, pode-se compensar a energia invisível, e assim, reduzir ou, de preferência, eliminar as diferenças existentes entre os sinais de resposta do calorímetro para elétrons e hádrons de mesma energia. Existem três maneiras de realizar essa compensação:

- Reduzindo a resposta ao chuveiro eletromagnético(em): utilizando um material absorvedor com núcleo pesado, como o chumbo;
- Aumentar a resposta a chuveiros hadrônicos: utilizando um material absorve-

dor que sofra fissão nuclear induzido pela parte não-eletromagnética do chuveiro, fornecendo assim uma energia extra a parte não-em do chuveiro, como o urânio empobrecido U^{238} , como o calorímetro construído por Willis e Fabjan em 1977 [28];

- Compensação offline: utilizando pesos para compensar a energia invisível da parte hadrônica do chuveiro.

3.5 Calorímetros Eletromagnéticos e Hadrônicos

Com o aumento da energia envolvida nas colisões de partículas em grandes aceleradores, detectores maiores e mais complexos têm se tornado essenciais nesses experimentos, principalmente o papel dos calorímetros eletromagnéticos (ECAL) e calorímetros hadrônico (HCAL) tem se tornado mais crucial. A principal finalidade de um calorímetro eletromagnético é medir a energia de elétrons e fótons e identificá-los em meio a uma grande quantidade de hádrons. Além disso, enquanto a direção do elétron é dada pelos detectores de traços (*trackers*), a posição e direção dos fótons - pelo menos para aqueles sem uma identificação de conversão conhecida - são dadas pelo calorímetro. Elétrons e fótons são a chave para identificar processos físicos de interesse. Uma excelente resolução de energia e direção é obrigatória para a identificação de novos estados, como por exemplo, $Z^0 \rightarrow e^+e^-$ e $H \rightarrow \gamma\gamma$. Devemos ter em mente que calorímetros eletromagnéticos são essenciais na detecção de chuveiros hadrônico, sendo o primeiro calorímetro que as partículas atingem. Em um calorímetro eletromagnético, fótons e elétrons com alta energia produzem uma cascata eletromagnética através de sucessivos processos em que perdem energia e produzem mais pósitrons, elétrons e fótons (Figura 3.6). O efeito *bremstrahlung* é de longe o principal responsável pela perda de energia em elétrons altamente energéticos, efeito este que consiste na emissão de um fóton por um elétron quando defletido por um núcleo. Já o comportamento dos fótons com a matéria é muito diferente das demais partículas carregadas; os fótons não estão sujeitos às numerosas colisões inelásticas assim como as demais partículas; as principais interações dos fótons são através do efeito fotoelétrico, do efeito Compton e da produção de pares elétron-pósitron (e^-e^+).

- No Efeito fotoelétrico, um fóton é absorvido por um átomo de elétron com a consequente emissão de um elétron ($\gamma e^- \rightarrow e^-$).
- O efeito Compton basicamente consiste de uma colisão elástica entre um fóton e um elétron ($\gamma e^- \rightarrow \gamma e^-$).

- No processo de produção de pares elétron-positrón (e^-e^+), um fóton altamente energético se converte em um pósitron e um elétron ao interagir com um núcleo atômico ou elétron ($\gamma + Z \rightarrow Z + e^- + e^+$), onde Z é um núcleo atômico ou elétron. Fótons com energia da ordem de $5 - 10MeV$ criam pares e^-e^+ .

Chuveiros eletromagnéticos se diferenciam de outros chuueiros devido à grande multiplicação de partículas que ocorre em seu desenvolvimento (Figura 3.6); elétrons com energia da ordem de GeV podem irradiar milhares de fótons ao longo do detector. Os elétrons e pósitrons produzidos nesse processo podem perder energia através da irradiação de mais fótons, estes por sua vez podem produzir mais pares elétron-pósitron dando origem a um chuueiro de partículas que pode conter milhares de elétrons, pósitron e fótons, a Figura 3.6 ilustra o desenvolvimento de uma cascata eletromagnética. Devido a esse processo de formação do chuueiro de partículas, a medida que o chuueiro se desenvolve a energia do mesmo decai, até o ponto em que o chuueiro perca seu poder de penetração no calorímetro. Quanto maior a energia inicial do chuueiro de partícula, mais longa será a penetração do chuueiro de partículas.

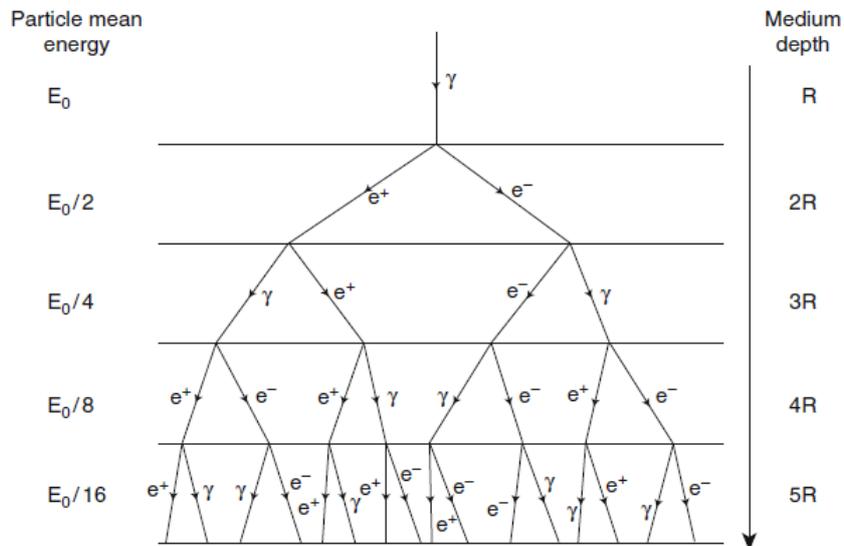


Figura 3.6: Chuueiro EM.Fonte:[6]

Os hádrons depositam energia na matéria através de uma série de colisões devido à interação forte e, portanto, através da ionização e excitação dos hádrons carregados (as primeiras interações fortes produzem muitos mésons π que sucessivamente interagem através da interação forte). A cascata hadrônica é mais larga e mais longa que a eletromagnética (ver Figura 3.7) e portanto um calorímetro hadrônico

possui dimensões maiores do que os eletromagnéticos. Contudo, existem grandes flutuações na cascata hadrônica devido a vários efeitos. O méson π^0 imediatamente decai em dois fótons 2γ , produzindo uma cascata eletromagnética onde a energia é depositada em um espaço estreito. Os mésons π^\pm decaem em um múon e no seu respectivo neutrino, o neutrino interage muito pouco com a matéria e passa através do calorímetro hadrônico sem interagir com ele (ver Figura 3.8). Estas características do chuveiro hadrônico causam uma perda de energia de cerca de 30%, além da perda de momento. Os múons, partículas parecidas com o elétron, porém 200 vezes mais pesadas, possuem um alto poder de penetração e passam tanto pelo calorímetro hadrônico quanto pelo eletromagnético produzindo um sinal mínimo de ionização, interagindo muito pouco com o meio. Muitos experimentos possuem instalados nas camadas mais externas do experimento detectores conhecidos como câmaras de múons, específicos para medir com mais precisão estas partículas.

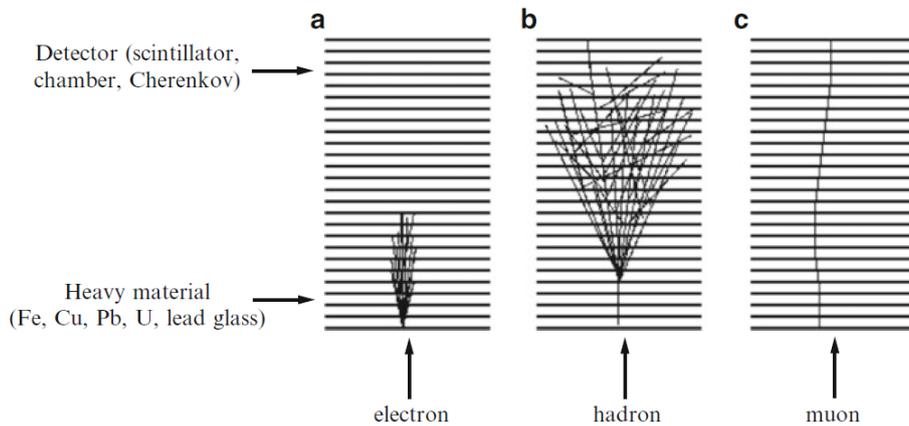


Figura 3.7: Comparação do desenvolvimento da cascata EM e hadrônica. Fonte:[6]

Os calorímetros são divididos em calorímetros homogêneos e de amostragem. Calorímetros homogêneos são constituídos por um material do tipo que combina propriedades de um absorvedor e um detector de partículas - isto significa que praticamente todo o volume do calorímetro é sensível a energia depositada - as funções de absorver as partículas e detectar os sinais produzidos neste processo são exercidas pelo mesmo material, que precisa ser de alta densidade para realizar a função de absorção de forma eficiente. Em calorímetros homogêneos, o bloco de absorção e o material ativo que detecta o chuveiro de partículas são idênticos, e portanto - com relação à resolução de energia - estes calorímetros fornecem a melhor performance.

Calorímetros de amostragem são feitos de componentes ativos, como por exemplo cintiladores, intercalados com camadas absorvedoras (chumbo em calorímetros eletromagnéticos e ferro em calorímetros hadrônicos). Calorímetros homogêneos

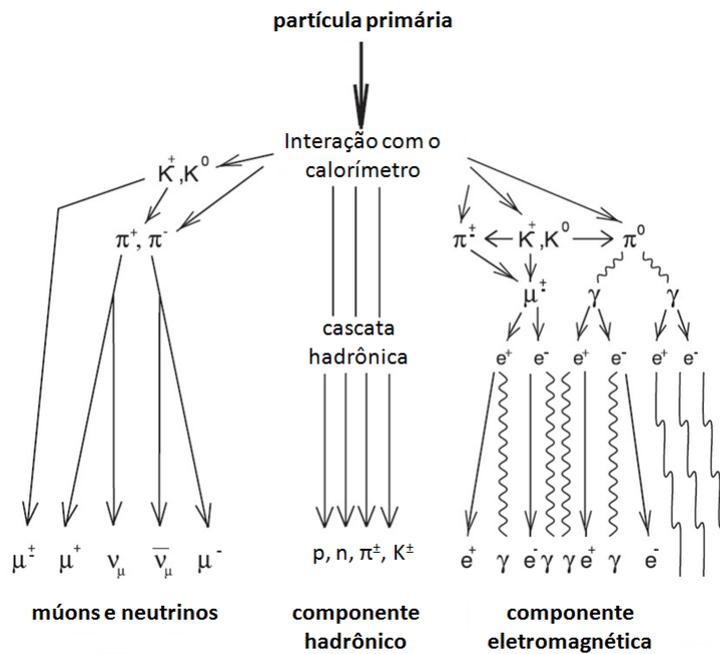


Figura 3.8: Chuveiro hadrônico. Fonte: [7]

possuem geralmente melhor performance do que os de amostragem, porém são mais caros de se produzir. Praticamente todos os calorímetros hadrônicos são calorímetros de amostragem, devido ao tamanho que deve possuir para medir a cascata hadrônica.

Capítulo 4

O Experimento LHC

A seguir, será apresentado o CERN e seu experimento mais importante na atualidade, o acelerador de partículas LHC. O maior detector do LHC, o ATLAS e os seus subdetectores serão explicados com mais detalhes.

4.1 CERN

A Organização Européia para Pesquisa Nuclear (antigo Conselho Europeu para Pesquisa Nuclear, em francês, *Conseil Européen pour la Recherche Nucléaire* - CERN) foi fundada em 1954, localizado na fronteira Franco-Suíça próximo da cidade de Genebra, na Suíça [29]. Foi idealizado por um grupo de cientistas pioneiros europeus que, após a Segunda Guerra Mundial, idealizaram a criação de um laboratório Europeu de física de partículas que pudesse impulsionar a ciência na Europa. Os cientistas franceses Lew Kowarski, Raoul Dautry e Pierre Auger, Edoardo Amaldi da Itália e Niels Bohr da Dinamarca estavam entre esses pioneiros. O físico francês Louis de Broglie apresentou a primeira proposta oficial para a criação de um laboratório europeu de física na abertura da Conferência Européia Cultural em Laussane em 9 de Dezembro de 1949. Tal laboratório tinha não somente a intenção de reunir cientistas europeus, mas também permitir que os países compartilhassem os custos elevados das instalações nucleares necessárias para o avanço da física. O CERN é o maior centro de pesquisa de física de partículas do mundo. Ficou conhecido mundialmente por suas inúmeras experiências com detectores de partículas, a descoberta dos bósons W e Z e os diversos prêmios Nobel ganhos por seus pesquisadores. Em 1994, o CERN Council aprovou a construção do *Large Hadron Collider* (LHC) [29], um novo acelerador de partículas a ser construído no complexo de aceleradores localizados no CERN, com propriedades radicalmente melhoradas em relação aos aceleradores já construídos em todo o mundo. Entre essas propriedades estão a alta taxa de eventos, para aumentar a probabilidade de se estudar eventos escassos e raros. As características do LHC foram projetadas para que seus experimentos possam

estudar algumas previsões ainda não verificadas experimentalmente; o acelerador necessitou de um enorme esforço e avanço tecnológico para os seus componentes, resultando em diversos avanços em áreas correlatas à física. Um dos resultados mais aguardados pela comunidade científica com o estudo das colisões de partículas no LHC é a comprovação ou não da existência do bóson de Higgs. Em julho de 2012, pesquisadores do CERN anunciaram ao mundo a descoberta de uma partícula com a mesma massa prevista teoricamente para o Bóson de Higgs, a probabilidade desta partícula não ser o bóson de Higgs é de 1 em 300 milhões. Em outubro de 2013, o Prêmio Nobel de Física foi concedido ao belga François Englert e ao britânico Peter Higgs por seus trabalhos teóricos sobre como as partículas adquirem massa, propostos separadamente em 1964. Entre 1996 e 1998, quatro experimentos - ALICE [30], ATLAS [10], CMS [31] e LHCb [32] - receberam permissão oficial para serem construídos em quatro pontos ao longo da circunferência do LHC. Anos depois dois experimentos menores também foram incorporados ao complexo do LHC, o experimento TOTEM [33] próximo ao CMS e o LHCf [34] próximo ao ATLAS. O complexo do CERN (Figura 4.1) é constituído por uma série de aceleradores menores responsáveis por elevar a energia do feixe de partículas. Cada acelerador injeta o feixe de partículas no próximo acelerador e este, por sua vez, fica responsável por elevar ainda mais a energia do feixe ao máximo possível e a assim sucessivamente. O LHC é o último estágio dessa cadeia de aceleradores e é o responsável final por elevar a energia do feixe de partículas a uma energia recorde de 7 TeV.

4.2 LHC - *Large Hadron Collider*

O LHC é, hoje, o maior acelerador de partículas do mundo. Ele está localizado a cerca de 100 metros de profundidade em um túnel com 26,7 km de circunferência. O túnel que abriga o LHC - a maior obra de engenharia civil realizada na Europa antes do túnel do Canal da Mancha - foi escavado na década de 1980 para abrigar o maior acelerador de elétron-pósitron já construído até então, chamado *Large Electron-Positron* (LEP) [35]. Os responsáveis por detectar as partículas no LEP foram os detectores ALEPHI, DELPHI, L3 e OPAL. Durante 11 anos, o LEP forneceu dados para diversos experimentos, principalmente o estudo da interação eletrofraca; medidas realizadas com o LEP também provaram que existem apenas três gerações de partículas da matéria. Em 2 de novembro de 2000, o LEP foi oficialmente aposentado para dar lugar ao LHC.

Em Setembro de 2008 o LHC iniciou a operação de circulação de feixes de próton, porém, 9 dias depois, devido ao aumento da energia dos feixes, um problema entre uma das conexões elétricas de super-condutores dos dipolos magnéticos do LHC provocou uma explosão do gás Hélio do sistema de resfriamento dos dipolos, o que

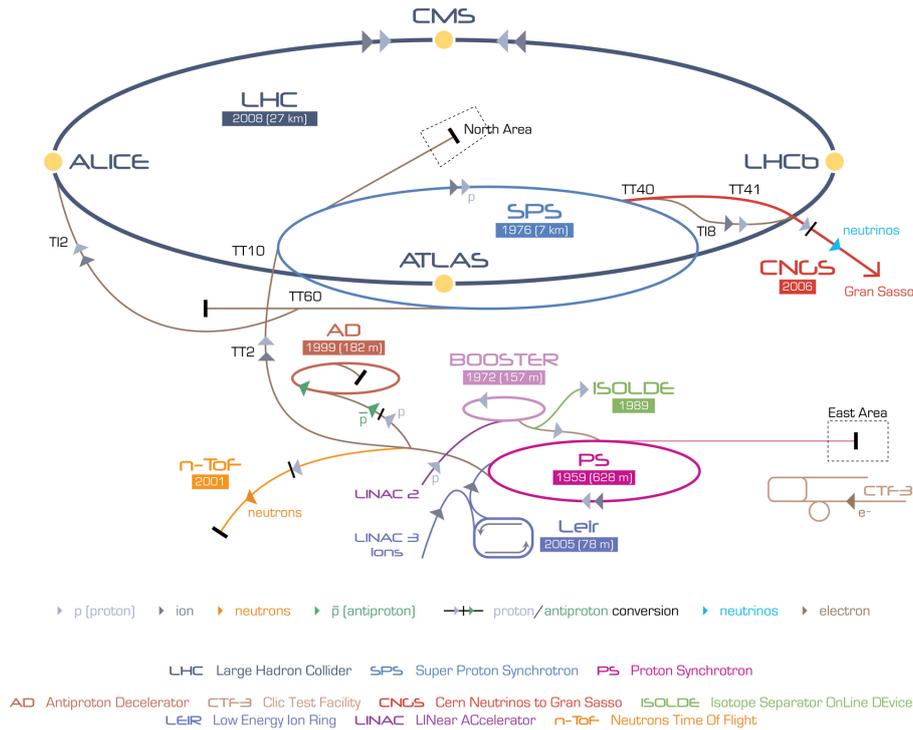


Figura 4.1: Complexo de Aceleradores do CERN. Fonte: [8]

causou uma interrupção de 14 meses nas operações no LHC. Em Novembro de 2009, após um ano de reparos, o LHC voltou a circular feixes de partículas, e já com algumas semanas de funcionamento se tornou o maior acelerador de partículas do mundo realizando colisões entre prótons com energia total de 2,2 TeV no centro de massa. No dia 30 de Março de 2010 realizou uma colisão próton-próton com energia de 7 TeV no centro de massa.

4.2.1 Características Gerais do LHC

Cada feixe de partículas em intensidade completa consiste de 2808 pacotes (*bunches*) se movendo no túnel circular. No caso de um feixe de prótons, cada pacote contém cerca de $1,15 \times 10^{11}$ dessas partículas, sendo que cada feixe de prótons possui poucos centímetros de comprimento com uma seção transversa de dimensões da ordem de 1 milímetro, mas, no ponto de colisão, passa a ter 16 microns. As partículas no LHC se movem com velocidade ultra-relativística alcançando 0,999999991 da velocidade da luz (7 TeV). O total de energia nominal de um feixe circulando no LHC chega a 362MJ ($2808 \text{ pacotes} \times 1,15 \times 10^{11} \text{ prótons} @ 7\text{TeV} = 2808 \times 1,15 \times 10^{11} \times 1,602 \times 10^{-19} \text{ Joules} = 362\text{MJ}$ por feixe), equivalente a 77,4 Kg de TNT. A distância entre um feixe e outro é de 7,5 metros. Uma vez que a luz leva cerca de 25 ns para percorrer essa distância e os prótons estão pratica-

mente se movendo na velocidade da luz, o valor nominal para que cada feixe passe pelos pontos de colisões é de cerca de 25 ns ou 40 milhões de vezes por segundo, este valor nominal é conhecido como *bunch spacing* do LHC. A performance do LHC é dada por dois parâmetros principais: sua luminosidade e a energia no centro de massa do feixe de colisão. O LHC foi construído para ser capaz de alcançar uma luminosidade nominal de $10^{34} \text{cm}^{-2} \text{s}^{-1}$ e uma energia no centro de massa de 14 TeV (7 TeV em cada feixe).

Porém, devido às limitações de interconexões entre os magnetos, a energia foi limitada a 3,5 TeV em 2010/2011 e a 4 TeV em 2012. O pico máximo de luminosidade obtida no ano de 2012, pelo LHC foi de $0,75 \times 10^{34} \text{cm}^{-2} \text{s}^{-1}$ e essa luminosidade foi alcançada operando com um *bunch* de 50 ns ao invés do valor nominal de 25 ns. Esse valor só foi possível com emitâncias¹ muito menores do que a nominal e *bunch currents* de mais de 30% superior ao nominal [36]. O número máximo de prótons por pacotes utilizado, nos anos de 2011 e 2012, foi de 1.380 prótons/pacote. A luminosidade assim como a estrutura do feixe (25ns ou 50ns *bunch spacing*) e o número de *bunches*, são determinados, em grande medida, pelo conjunto de aceleradores injetores do LHC (LINAC, Booster, PS e SPS). Um grande esforço está sendo realizado para entender e melhorar os parâmetros do feixe no complexo de injetores do LHC, com um impacto direto sobre o desempenho futuro do LHC, para que o acelerador alcance os parâmetros nominais para o qual foi projetado. Um programa de melhoria ambicioso está a caminho durante a próxima década para melhorar ainda mais os parâmetros do feixe que podem ser entregues ao LHC, bem como a outros experimentos de física no CERN.

Os quatro detectores mais importantes do LHC (ATLAS, CMS, ALICE e LHCb) implementam sistemas síncronos de aquisição para operar com eventos na taxa de 40 MHz:

- O detector de partículas ATLAS (*A Toroidal LHC Apparatus*), o maior dos quatro detectores principais do LHC, foi construído numa colaboração entre 150 institutos de pesquisa de 37 países dentre os quais a Universidade Federal do Rio de Janeiro participa através do Laboratório de Processamento de Sinais e do Instituto de Física. Seus propósitos e constituição serão explicados em detalhe na seção 4.3.
- O detector CMS (*Compact Muon Solenoid*) é um experimento de propósito geral para estudos do bóson de Higgs, de partículas supersimétricas e de física de íons pesados.

¹Grandeza relacionada com a dispersão óptica das partículas que formam o feixe. Chama-se de emitância do feixe a região que as partículas ocupam no espaço de fase.

- O ALICE (*A Large Ion Collider Experiment*) é o único experimento do LHC inteiramente dedicado à física de colisões nucleares. Seu objetivo é estabelecer e estudar a formação do plasma de quarks e glúons.
- O LHCb (*Large Hadron Collider beauty experiment for precision measurements of CP-violation and rare decays*) é um experimento cuja principal finalidade é o estudo da violação de CP (Carga-Paridade) no setor do quark b, o detector também estuda decaimentos raros dos mésons B.

O presente trabalho está relacionado com o experimento ATLAS, que será descrito com mais detalhes a seguir.

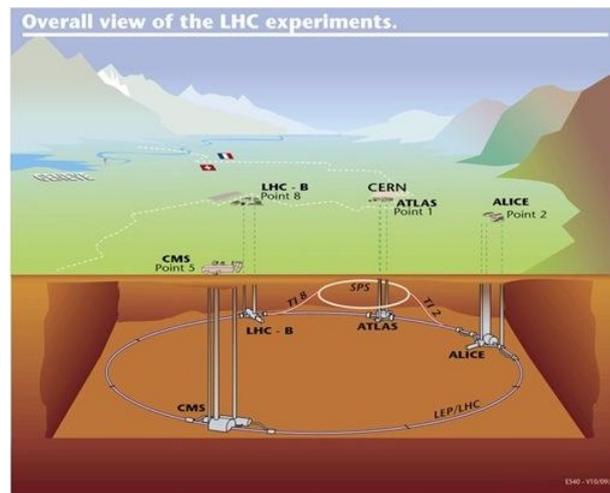


Figura 4.2: LHC e seus detectores. Fonte:[8]

4.3 O Detector ATLAS

O detector de partículas ATLAS (*A Toroidal LHC Apparatus*) é um experimento de propósito geral para medir e detectar o resultado das colisões entre feixes de partículas no LHC. Está localizado em um dos pontos de colisão do LHC, possui 44m de comprimento, 25m de altura e pesa aproximadamente 7000 toneladas, peso equivalente ao da Torre Eiffel, sendo assim a maior máquina já construída pelo Homem (Figura 4.3). O ATLAS possui, como objetivo, estudar o maior número de fenômenos físicos passíveis de serem gerados nas colisões no LHC. Seus objetivos principais são aprender sobre as forças básicas que moldaram o nosso Universo desde o início do tempo, e que irão determinar seu destino. Os principais estudos são sobre a origem da massa (bóson de Higgs), dimensões extras do Universo, buracos negros microscópicos e as provas para os candidatos a matéria escura no Universo.

O detector é composto por diversos sub-sistemas, cada um com características e funcionalidades diferentes, esses são:

- *Solenoid e Toroid Magnets*: sistema magnético desenvolvido para contribuir na medição do momento das partículas [37].
- *Inner Detector*: sub-detector composto por outros três sub-detectores (*Pixel Detector*, *Semiconductor Tracker* (SCT) e *Transition Radiation Tracker* (TRT)) responsáveis por medir o momento das partículas carregadas [37].
- *Muon Spectrometer*: sub-detector desenvolvido para identificar, filtrar e medir o momento dos múons [37].
- *Calorímetro Eletromagnético e Hadrônico*: o Calorímetro Eletromagnético e o Calorímetro Hadrônico (TileCal) do ATLAS são subdetectores desenvolvidos para medir a energia das partículas que reagem de forma eletromagnética e hadrônica respectivamente [37].

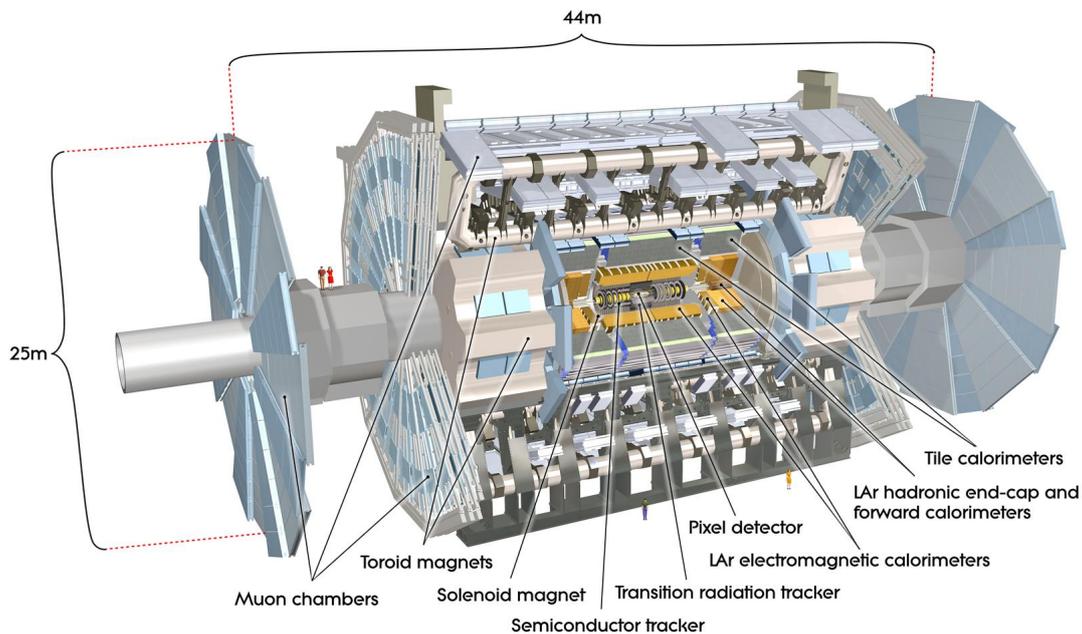


Figura 4.3: ATLAS e seus principais detectores e dimensões. Fonte:[10].

4.3.1 Sistema de Trigger e Aquisição de Dados

Devido a grande quantidade de colisões produzidas no LHC, foi necessário construir um sofisticado sistema de filtragem para esses eventos com a intenção

de selecionar somente os eventos de interesse. Três níveis de filtragem online (*trigger*) permitem uma operação coerente entre diferentes sub detectores, formando o sistema o de *Trigger*, que opera em conjunto com o sistema de *Data Acquisition* (DAQ) e o *Detector Control System* (DCS). O Sistema de *trigger* tem a responsabilidade de escolher quais eventos físicos serão considerados de interesse para o programa de pesquisa do ATLAS. Todos os eventos não selecionados são descartados para sempre, permitindo assim reduzir a taxa inicial de aquisição de cerca de 1GHz para 100Hz. Esta redução da taxa de eventos permite que os dados físicos possam ser armazenados e analisados *offline*. Para um evento ser considerado de interesse, é preciso ser definido pelo programa de física do ATLAS e identificado pelo seu sistema de *trigger*. A assinatura de interesse dos eventos, no sistema de *trigger*, é alcançada usando o estado final de vários objetos de física: léptons carregados com baixa ou alta p_T , jatos com alto p_T (originados por quarks e glúons), bósons eletrofracos de Gauge (W, Z, γ) e por último a energia transversa faltante. A informação necessária para construir esses objetos de interesse dos detectores do ATLAS é construída em três níveis de seleção de eventos online : nível1(LVL1), nível2(LVL2)

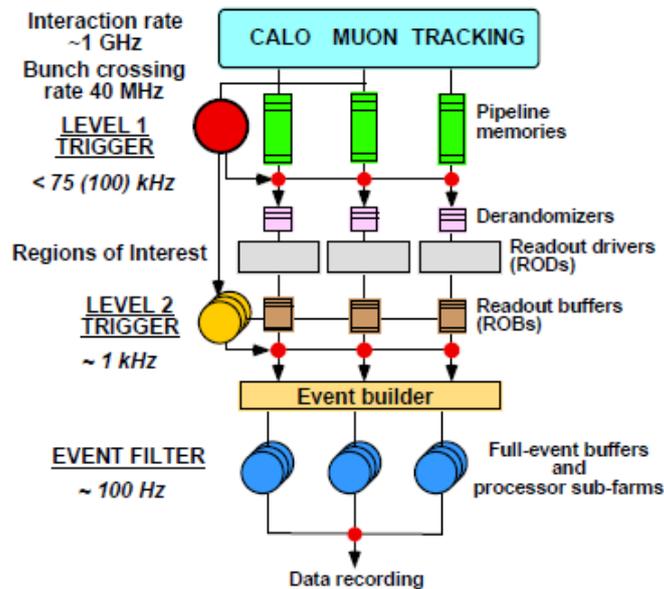


Figura 4.4: Arquitetura do sistema de Trigger do ATLAS. Fonte:[10].

O sistema de seleção do LVL1 é baseado em *hardware* e faz a decisão inicial baseando-se na informação de granularidade reduzida fornecida por um sub-conjunto de detectores (detectores de múons e calorímetros). O LVL1 aceita dados destes detectores na taxa completa de luminosidade. A latência, tempo para formar e distribuir a decisão do nível 1, é de $2\mu s$ e a máxima saída é limitada a 75 KHz

(podendo chegar a 100 KHz), devido à capacidade do sistema de leitura dos sub-detecores e do nível 2 do *trigger*. Um requisito essencial no LVL1 é identificar unicamente o *bunch-crossing* que tiver um intervalo de 25ns. No caso das câmaras de múons, o tamanho físico do espectrômetro de múons implica em um tempo de vôo comparável ao período de *bunch-crossing*. Para o calorímetro, o tempo de formação do pulso do sinal ultrapassa o tempo de *bunch-crossing*. Durante o processamento no nível 1 do *trigger*, todos os dados dos canais do detector são conservados em memórias *pipeline*; todos os dados selecionados pelo nível 1 são lidos no sistema de eletrônica de *front-end* dos detectores em *Read Out Drivers* (ROD) e, então, encaminhados para os *Read Out Buffers* (ROB) onde eles são armazenados até o nível 2 tomar sua decisão. Para minimizar a latência, somente dados das regiões de interesse, definidas pelo nível 1, são transferidas para os processadores do *trigger*. O nível 2 do *trigger* reduz a taxa de cerca de 100 kHz (saída do LVL1) para 1 kHz, com um latência variando de 1 a 10 ms, dependendo do evento. Após um evento ser selecionado, os dados são enviados para o *Event Filter* (EF) através do *event builder*. O EF emprega algoritmos e métodos *offline* adaptados para trabalharem *online*, e usa a mais recente informação e mapas de campo magnético. O EF faz a seleção final de eventos físicos que serão escritos e armazenados permanentemente no banco de dados do ATLAS, para subsequente análise completa com algoritmos diversos. A taxa de saída de dados do nível 2 deve ser reduzida para aproximadamente 100 Hz que corresponde a um armazenamento de 10-100 MB/s se todo o dado do evento for gravado. O sistema DAQ controla a distribuição de dados da ROD para o banco de armazenamento, também controla a aquisição de dados e monitora todo experimento.

4.3.2 Coordenadas Geométricas do ATLAS

O sistema de coordenadas do ATLAS utiliza o sistema de coordenadas da mão-direita: o eixo x aponta para o centro da circunferência do LHC (ver Figura 4.2), e o eixo z acompanha sua circunferência e o eixo y é ligeiramente inclinado em relação a vertical. O sistema de coordenadas cartesianas do ATLAS é mostrado na Figura 4.5. A maioria das análises físicas e reconstruções de trajetórias feita no ATLAS utiliza o sistema de coordenadas cilíndricas, onde o eixo das abscissas representa a coordenada z , o eixo da ordenada representa os raios (ρ) dos cilindros e o ângulo azimutal ϕ completam o sistema de três coordenadas desse sistema, como mostrado na Figura 4.6, onde também podemos observar o ângulo polar θ (Ângulo entre o momento da partícula e o eixo do feixe), que fornece através da equação 4.1, a pseudo-rapidez² (η) da partícula, e o momento da partícula p_t perpendicular ao eixo

²Coordenada espacial comumente usada em experimentos de física de partículas para descrever o ângulo de uma partícula em relação ao eixo do feixe.

z.

$$\eta(\theta) = -\log\left(\tan\frac{\theta}{2}\right) \quad (4.1)$$

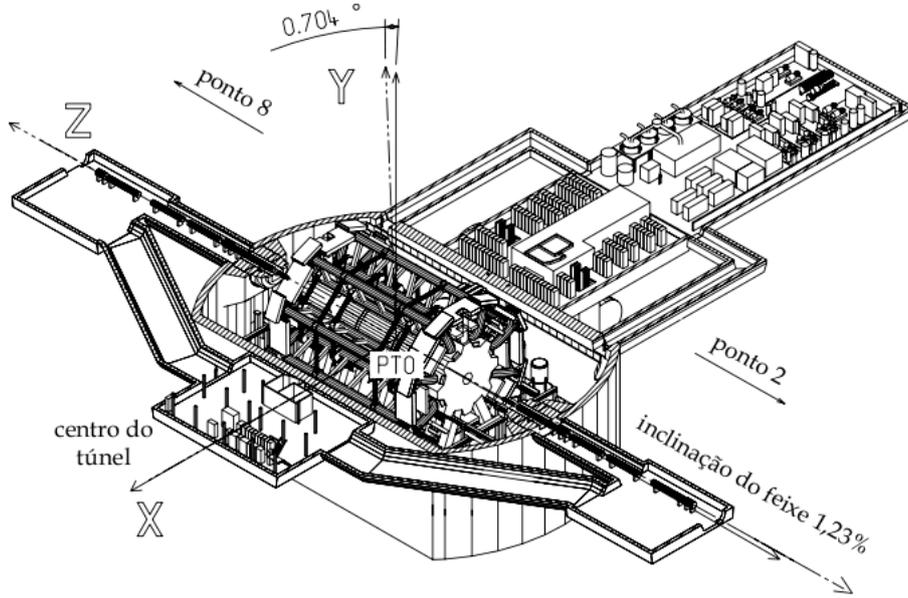


Figura 4.5: Sistema de coordenadas cartesianas do ATLAS.Fonte:[10].

4.3.3 Sistema Magnético do ATLAS

O sistema magnético do ATLAS é composto por quatro subsistemas: *central solenoid*, *barrel toroid* e dois *end-cap toroids* [37]. O Solenóide Central (CS) do ATLAS tem comprimento de 5,3 m e um diâmetro de 2,4m, pesa cerca de 5,7 toneladas e opera com uma corrente de 7.600 A. Ele foi projetado para fornecer um campo magnético de 2 Tesla com picos de 2,6 Tesla. O *Barrel Toroid* (BT) é composto por oito bobinas supercondutoras, cada uma com 25 metros de comprimento e 5 metros de largura, agrupada em forma de um toróide, que são sustentada por 16 anéis de suporte (ver Figura 4.7). Cada bobina pesa cerca de 40 toneladas. O conjunto total do *Barrel Toroid* tem um diâmetro exterior de 20 metros e pesa aproximadamente 830 toneladas. O toróide opera com uma corrente de 20.500A e o campo magnético produzido pode chegar a 4 Tesla de pico. Os dois *End-Cap Toroids* (ECT) estão inseridos no *Barrel Toroid* e encontram-se nas extremidades do Solenóide Central. Cada ECT possui um comprimento de 5m, um diâmetro de 10.7m e pesa cerca de 240 toneladas. Os *End-Cap Toroids* estão conectados em série com o *Barrel Toroid* e operam igualmente com uma corrente de 20.500 A, fornecendo um campo magnético

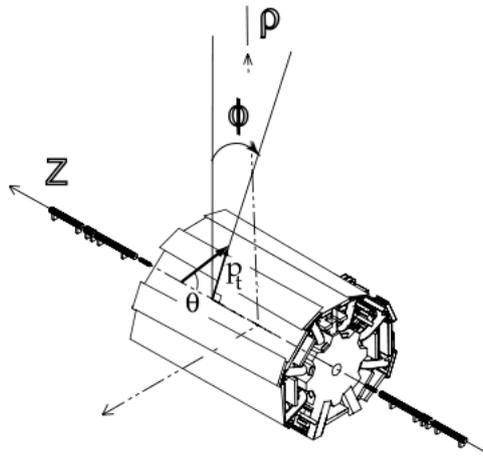


Figura 4.6: Sistema de coordenadas cilíndricas do ATLAS.Fonte:[10].

de 4 Teslas.

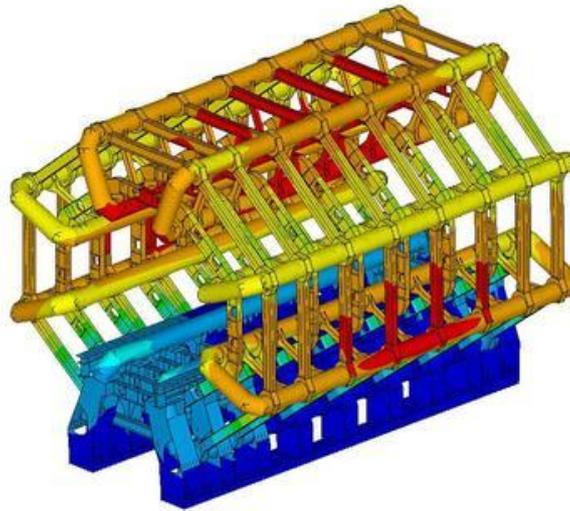


Figura 4.7: *Barrel Toroid* do ATLAS.Fonte:[10].

4.3.4 Detectores de Traço - *Inner Detector*

O *Inner Detector* é formado por um conjunto de três detectores: o *Pixel Detector*, o *Semiconductor Tracker* (SCT) e *Transition Radiation Tracker* (TRT) [37]. Por ser o detector mais próximo ao vértice de colisão das partículas, é de sua responsabilidade medir as direções, o momento e as cargas das partículas eletricamente

carregadas. O momento e o vértice de colisão requerem uma alta precisão e as medidas devem ser feitas por detectores com fina granularidade, devido à grande quantidade de partículas liberadas em cada colisão. Uma breve descrição dos subdetectores que formam o *Inner Detector* será dada a seguir. Na Figura 4.8, podemos observar cada um dos subdetectores descritos.

4.3.5 O Detector PIXEL

O detector de Pixel foi desenvolvido para fornecer uma alta precisão e granularidade o mais próximo possível do ponto de colisão, o sistema ao todo possui cerca de 140 milhões de elementos de detecção. A segmentação bidimensional de seus sensores fornece pontos no espaço sem nenhuma das ambiguidades associadas com geometrias projetivas, porém, requer o uso de avançadas técnicas de eletrônica e conexões para a leitura.

4.3.6 O Detector SCT

O SCT é um detector de tiras de silício cujo barril utiliza oito camadas de detectores de microtubos de silício para fornecer pontos precisos nas coordenadas $R\phi$ e z utilizando um pequeno ângulo estéreo para fornecer a medida em z . O detector contém $61m^2$ de detectores de silício, com 6,2 milhões de canais de leitura. No detector SCT, trajetórias de partículas podem ser distinguidas se separadas por uma distância maior que aproximadamente $200\mu m$.

4.3.7 O Detector TRT

O TRT é baseado no uso de detectores de tubo que podem operar em altas taxas de eventos, como ocorre no LHC, em virtude do seu pequeno diâmetro e ao isolamento do sensor elétrico dentro de cada volume individual de gás. O detector possui cerca de 500.000 tubos preenchidos com uma mistura de gases, principalmente Xenônio (Xe) que ao ser atravessado por uma partícula eletricamente carregada geram, por ionização, uma corrente elétrica que é coletada através de um fio no interior de cada tubo. Essa leitura permite a detecção da partícula que ionizou o gás no interior do tubo. Cada tubo do detector possui um diâmetro de 4mm e em seu interior encontra-se um fio banhado a ouro de $30\mu m$ de diâmetro responsável pela leitura; cada tubo possui, no máximo, 144cm de comprimento. O detector possui um total de 420.00 canais de leitura que adicionam ao sistema de reconhecimento de padrões cerca de 36 pontos por trajetórias, permitindo que métodos simples de reconstrução de trajetórias sejam implementados.

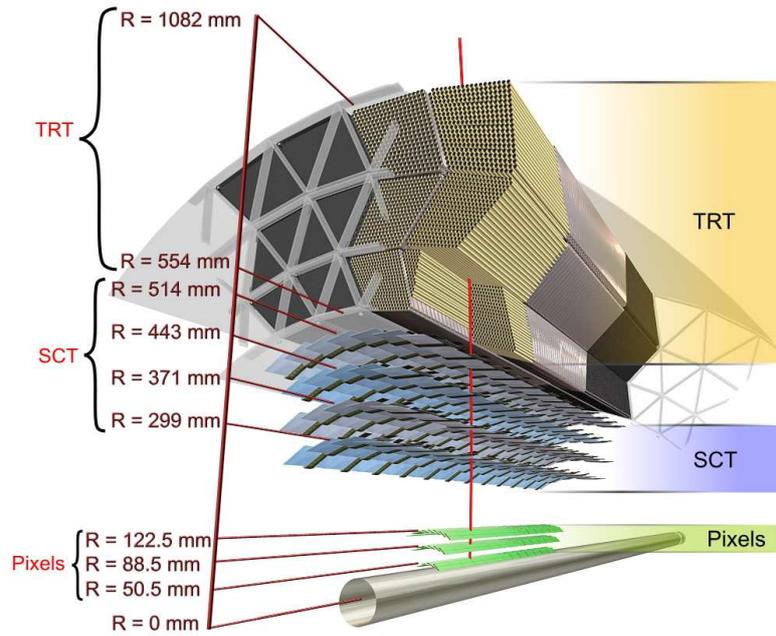


Figura 4.8: Layout do *Inner Detector* do ATLAS.Fonte:[10].

4.3.8 O Espectrômetro de Múons

O espectrômetro de múons localiza-se ao redor dos calorímetros do ATLAS e é responsável por medir as trajetórias dos múons para determinar sua direção, sua carga e seu momento. As câmaras de múons, também conhecida por MDT (*Monitored Drift Tubes*), são os elementos de detecção do espectrômetro e são formadas por milhares de tubos de metal preenchidos com gás e com um fio em seu centro. Este detector possui o mesmo princípio de detecção do detector TRT.

4.3.9 Calorímetro Eletromagnético

O calorímetro eletromagnético possui a função de observar e amostrar os componentes eletromagnéticos do decaimento das partículas. Isto é crucial para a separação das partículas que produzem chuveiros eletromagnéticos daquelas que produzem jatos de partículas hadrônicas. O calorímetro eletromagnético do ATLAS é um calorímetro de amostragem que utiliza Argônio líquido (LAR) como material sensível e chumbo como material absorvedor. Sua estrutura interna possui a forma de um acordeão. Essa geometria fornece ao calorímetro uma completa simetria em Φ , eliminando assim zonas mortas. O calorímetro EM é dividido em um barril ($|\eta| < 1.475$) e dois *end-caps* ($1.375 < |\eta| < 3.2$), como pode ser visto na Figura 4.9.

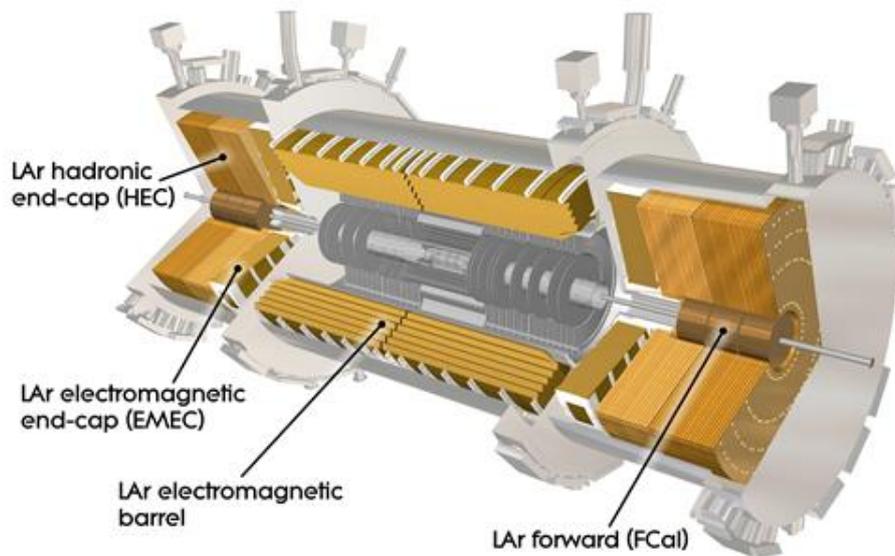


Figura 4.9: Calorímetro Eletromagnético do ATLAS.Fonte:[10].

4.4 Calorímetro Hadrônico - TileCal

O TileCal, principal calorímetro hadrônico do ATLAS, foi desenvolvido para ser um calorímetro de amostragem que utiliza aço como material absorvedor e telhas cintiladoras, que são lidas por fibras óticas conhecidas como *wavelength shifting fibers* (WLS) como meio ativo. Ele é dividido em três partes ao longo do eixo z: um barril central (dividido em duas partições LBA e LBC) de 5,64 metros de comprimento que cobre a região de $|\eta| < 1.0$, e dois barris estendido (EBA e EBC) de 2,91 metros de comprimento, com raio interno de 2,28 metros e raio externo de 4,23 metros, que cobre a região de $0.8 < |\eta| < 1.7$. A existência desse vão entre o barril central e os barris externos se faz necessário para que cabos e equipamentos dos detectores internos do ATLAS possam ser montados. Os barris do TileCal são mostrados na Figura 4.10.

O TileCal é dividido em seções transversais, sendo que cada seção é dividida em três camadas de células (A, BC e D) de diferentes tamanhos, como será explicado, com maior detalhes, na subseção 4.4.1. Cada célula representa um conjunto de cintiladores agrupados de forma a medir, em conjunto, a energia de uma partícula que passe através da célula. Uma novidade em calorimetria, que é proposta pelo TileCal, é a disposição dos cintiladores perpendicularmente ao feixe de colisão (ao

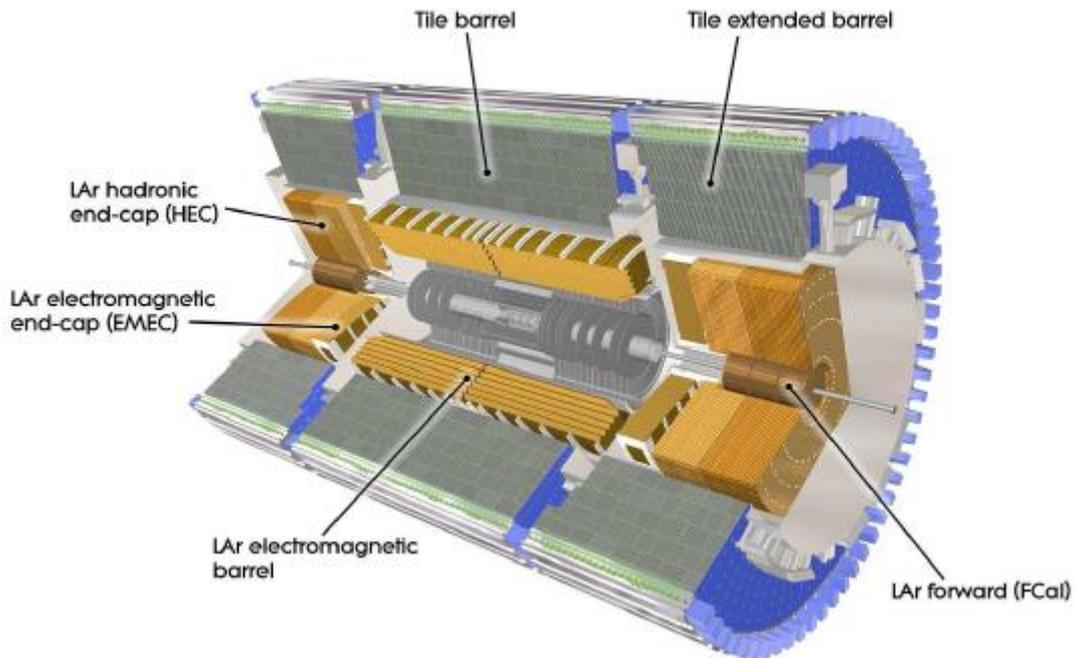


Figura 4.10: Calorímetro Hadrônico.Fonte:[10].

eixo z), como pode ser visto na Figura 4.11, que permitiu uma excelente resolução em ϕ .

O responsável por amostrar a energia das partículas hadrônicas que atravessam o TileCal é o cintilador. O cintilador é um material de poliestireno em formato de telha (*Tile*) dopado com material cintilante que, ao interagir com uma partícula, transforma a energia depositada pela partícula em luz, através da cintilação. A luz produzida nas telhas cintiladoras está na faixa do Ultra Violeta (UV) e sua intensidade é proporcional à energia depositada pela partícula. A luz se propaga dentro da telha para suas bordas através de reflexões. Esse feixe de luz é então transmitido através de uma fibra óptica até um fotomultiplicador (*Photo Multiplier Tube* - PMT) que, então, transforma esse feixe de luz em um sinal elétrico. Esse sinal elétrico é enviado então para o sistema de *trigger* onde sinais de interesse são separados dos sinais ordinários. As células do TileCal (telhas cintiladoras agrupadas em um conjunto) estão acopladas em fibras ópticas em cada uma de suas duas laterais, e estas são conectadas a dois fotomultiplicadores que leem a mesma célula. Esta configuração foi realizada para obter redundância no sinal coletado. A configuração e disposição de um módulo do TileCal podem ser vistas na Figura 4.11.

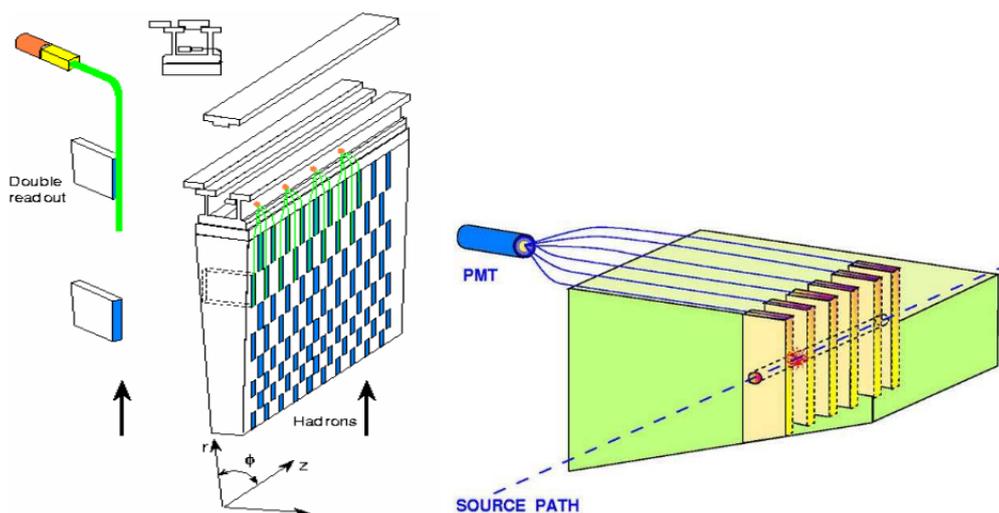


Figura 4.11: Disposição dos cintiladores e fibras óticas no TileCal. Fonte:[11].

4.4.1 Segmentação do TileCal

O TileCal é composto por três cilindros: um barril central (LB - *Long Barrel*) e dois barris estendidos (EB - *Extended Barrel*). Cada metade do LB e cada um dos EB forma uma partição independente de leitura. O LB é dividido em duas partes (LBA e LBC) com 22 células de leitura em cada, e cada EB possui 14 células de leitura. Cada partição é dividida azimutalmente (em ϕ) em 64 módulos, com uma segmentação de 0,1 radianos. Cada módulo do TileCal possui 48 posições para PMT, porém somente 45 PMTs são utilizadas no barril central e 32 PMTs nos barris estendidos. Com exceção das células D0, E1, E2, E3 e E4, que só possuem uma PMT para leitura, as demais células são lidas por duas PMT. O mapa completo de um dos módulos do TileCal pode ser visto na Figura 4.12.

4.4.2 Eletrônica de Leitura do TileCal

O sistema eletrônico de leitura formata, amplifica e digitaliza os sinais provenientes do sistema óptico de leitura (telhas cintiladoras e fotomultiplicadores). O sinal resultante é amostrado por um ADC de 10 bits na frequência de *bunch crossing* do LHC, de 25 ns. A Figura 4.13 mostra a cadeia simplificada do sistema de leitura do calorímetro TileCal, desde a aquisição do sinal de luz produzido pelas partículas que atravessam o detector nos elementos óticos, até a digitalização dos dados para seu armazenamento.

O sistema eletrônico de *front-end* (FE) e os fotomultiplicadores ficam localizados

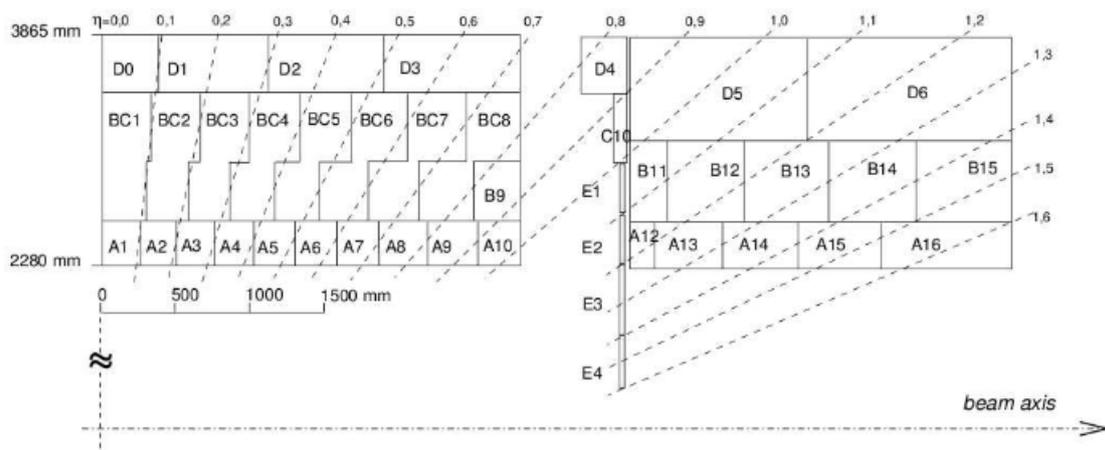


Figura 4.12: Segmentação do TileCal.Fonte:[12].

em um compartimento removível, localizado no raio externo de cada módulo de cada barril do detector. Esta estrutura sólida feita de aço é conhecida como gaveta (*Drawer*) e permite o acesso fácil para a reparo e substituição (veja a Figura 4.14). Um sistema contendo duas gavetas é conhecido como super-gaveta e possui cerca de 3 metros de comprimento. Uma única super-gaveta é capaz de ler um módulo completo de um barril estendido; para os módulos de barril central, são necessárias duas super-gavetas. Cada super-gaveta pode suportar até 48 blocos de PMT além de abrigar placas para providenciar a leitura, amplificação e digitalização do sinal físico.

Cada bloco de PMT no TileCal, como mostrado na Figura 4.15, representa um dos mais de 10.000 canais de leitura do detector. Os principais elementos desse bloco são: um fotomultiplicador, um *Mixer*, um *HV Divider* e uma placa 3-in-1. O restante da eletrônica que não se encontra nas super-gavetas, fica localizado na sala de eletrônica do ATLAS, onde se encontram: a fonte de alimentação de Alta Voltagem, o nível 1 de *trigger* (LVL1), os *Read-Out Drivers* (ROD) e os elementos de controle para os sistemas de calibração.

4.4.2.1 Fotomultiplicadores - PMT

O PMT utiliza-se do efeito fotoelétrico e da emissão secundária para amplificar um feixe de luz incidente e de pouca energia. O efeito fotoelétrico, cuja explicação satisfatória foi dada por Albert Einstein em 1905, consiste na emissão de elétrons por um material quando excitado por uma radiação eletromagnética como a luz. O efeito da emissão secundária de elétrons é um fenômeno onde elétrons adicionais, chamados de elétrons secundários, são emitidos da superfície de um material quando

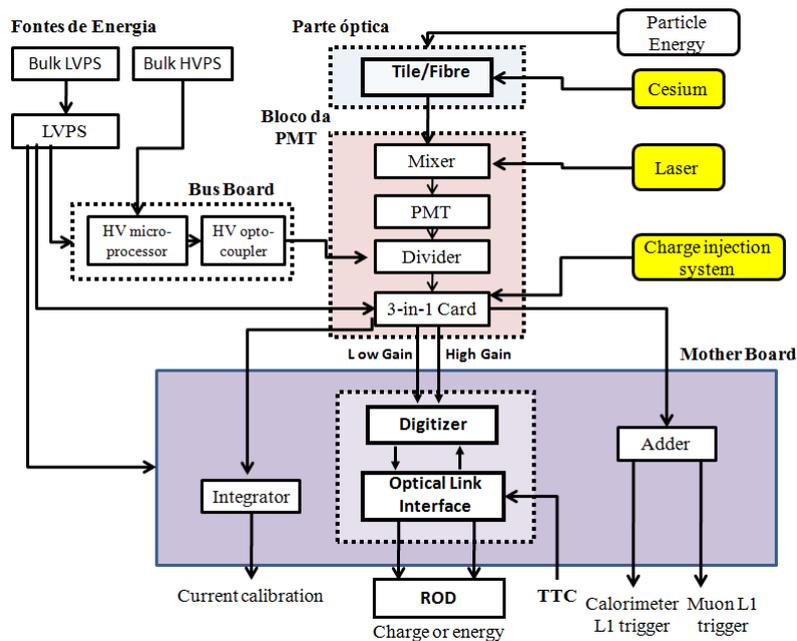


Figura 4.13: Esquema de Leitura do TileCal.

uma partícula carregada, como íons ou elétrons, incide em sua superfície com energia suficiente. O fotomultiplicador consiste de uma série de eletrodos, que são chamados dinodos, e um anodo. Um feixe de luz quando incide em um fotocátodo depositado na superfície de entrada do fotomultiplicador, este quando excitado pelo feixe libera elétrons pelo efeito fotoelétrico. Esses elétrons são direcionados para incidir em eletrodos onde são multiplicados pelo processo de emissão secundária. Esses elétrons em um número maior são acelerados até o segundo dinodo e, assim, sucessivamente multiplicando o número de elétrons produzidos a cada estágio. Finalmente, este acúmulo de carga chega ao anodo produzindo um pulso de corrente que pode assim ser medido pela eletrônica do experimento. Esse processo pode ser visto na Figura 4.16.

Os fotomultiplicadores são capazes de multiplicar o sinal produzido pela luz incidente em milhões de vezes, permitindo que pequenos fluxos de fótons sejam detectados. No TileCal, o fotomultiplicador utilizado é o Hamamatsu R7877 [13].

4.4.2.2 Mixer

A maior parte da resposta do fotomultiplicador depende da superfície do fotocátodo. Um misturador de luz é responsável por misturar a luz vinda de todas as fibras em um feixe, de modo que não haja nenhuma correlação entre a posição da fibra e a área do fotocátodo que recebe a luz, otimizando assim a uniformidade da

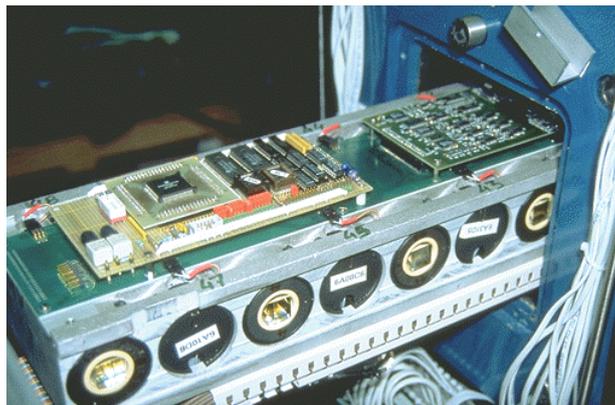


Figura 4.14: Gaveta de um Módulo do TileCal. Fonte:[12]

detecção da luz. A Figura 4.17 mostra a diferença entre a detecção de luz por uma PMT ao utilizar um Mixer.

4.4.2.3 *High Voltage Dividers*

O objetivo principal deste elemento (ver Figura 4.18) é partilhar a alta tensão entre os dinodos do PMT. A placa também serve como uma tomada para permitir a conexão do PMT com a eletrônica de *front-end* sem qualquer interligação de fios. Este esquema minimiza a capacitância entre o PMT e a eletrônica e é importante para reduzir o ruído e as conexões instáveis.

4.4.2.4 **Placas 3-in-1**

As principais funções analógicas da cadeia de leitura do TileCal são efetuadas por esta placa [13]. É nesta placa que um circuito de *shaper* LC passivo de 7 pólos condiciona os pulsos elétricos vindos do fotomultiplicador, para que possam ter uma forma padronizada (pulso padrão do TileCal) tal que a energia total da partícula que é medida no PMT seja proporcional a amplitude do pulso. Essas placas também fornecem ganhos Alto (*High Gain*) e Baixo (*Low Gain*) ao pulso antes deles irem para as placas digitalizadoras; além disso, é nas placas 3-in-1 que se encontra o sistema de calibração, através de injeção de cargas (TileCIS) e a integração lenta dos sinais do PMT para monitoramento e calibração, com fonte de Césio (Cs^{137}).

4.4.2.5 **Placas Digitalizadoras**

Sinais de pulso rápido das placas 3-in-1 são digitalizados nas placas digitalizadoras e enviados para um pipeline digital. Após a recepção do sinal de validação (L1A) do nível 1 do *trigger* LVL1, as placas digitalizadoras capturam uma estrutura de

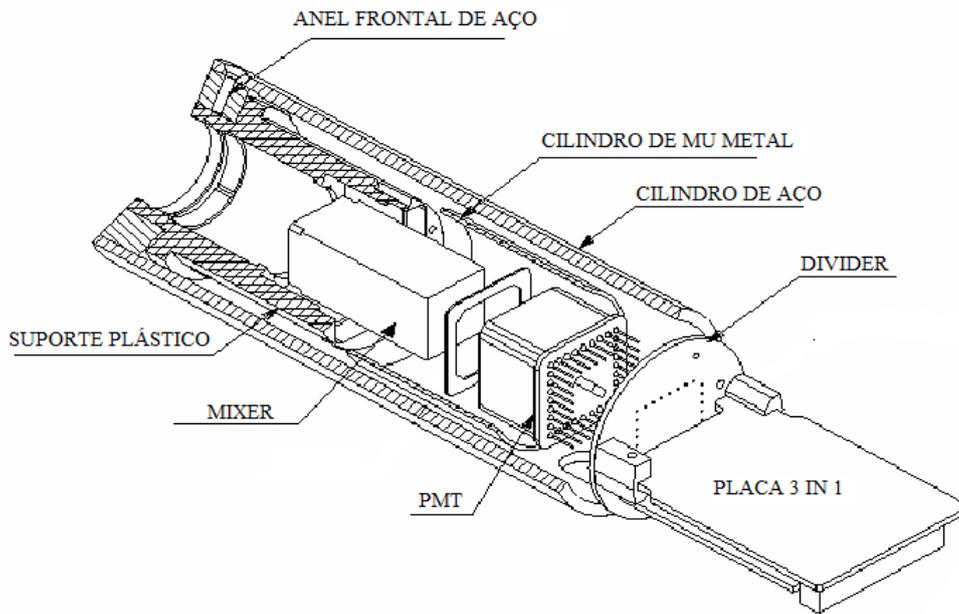


Figura 4.15: Esquema de um bloco de PMT. Fonte:[12]

eventos consistindo de uma sequência de digitalizações. Os eventos (*data frames*) são armazenados localmente e em fila para transmissão para o link de interface. Os dados são apresentados para as placas digitalizadoras pelo sistema 3-in-1, que fornece duas versões de cada sinal, um com ganho alto e um com ganho baixo (sendo 64 a razão entre esses ganhos). Estes dados analógicos são digitalizados em intervalos de 25 ns e armazenados temporariamente. Os digitalizadores são formados, prin-

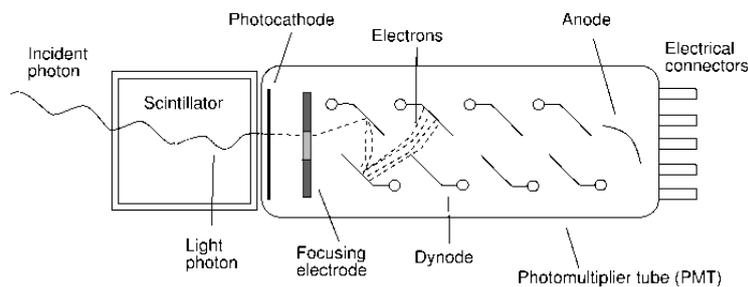


Figura 4.16: Fotomultiplicador. Fonte:[9].

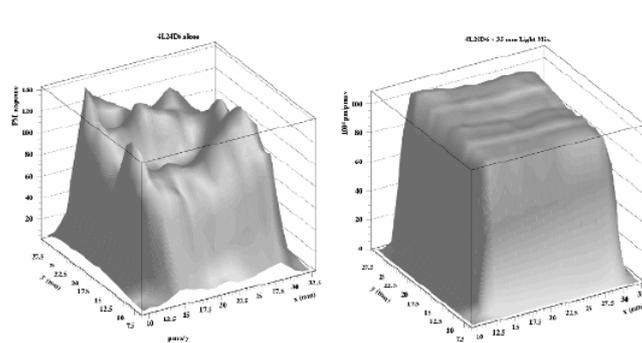


Figura 4.17: Resposta de uma PMT sem (esq.) e com (dir.) um mixer de luz. Fonte:[13].



Figura 4.18: *High Voltage Dividers*. Fonte:[13].

principalmente por ADCs comerciais, TTCrx³ para o sinal de TTC⁴ (Timing, Trigger and Control) e chips ASIC TileDMU (Unidades de Gestão de Dados) customizados para as necessidades do TileCal. Após receber o sinal de L1A, a sequência de amostras digitais é enviada pela placa digitalizadora para a transmissão para a interface link. O TileDMU é responsável pela reformatação e reordenamento dos dados digitalizados e para serem enviados para as ligações de interface. Cada TileDMU gere 3 ADCs, e cada placa digitalizadora possui 2 TileDMUs. Portanto, existem oito placas de digitalizadoras para cada super-gaveta do barril central (capaz de gerir até 48 canais, porém apenas 45 são realmente utilizadas), e 6 para as super-gavetas do barril estendido (espaço para até 36 canais, porém apenas 32 são usadas). Um

³O TTCrx é um circuito integrado de aplicação específica (ASIC em inglês) projetado pelo CERN EP Microelectronics group.

⁴O sistema de TTC fornece: os sinais de Clock e controle para a sincronização dos elementos do detector com o *clock* da máquina; o sinal de aceitação do Trigger; a identificação da informação de eventos; e o controle de dados.

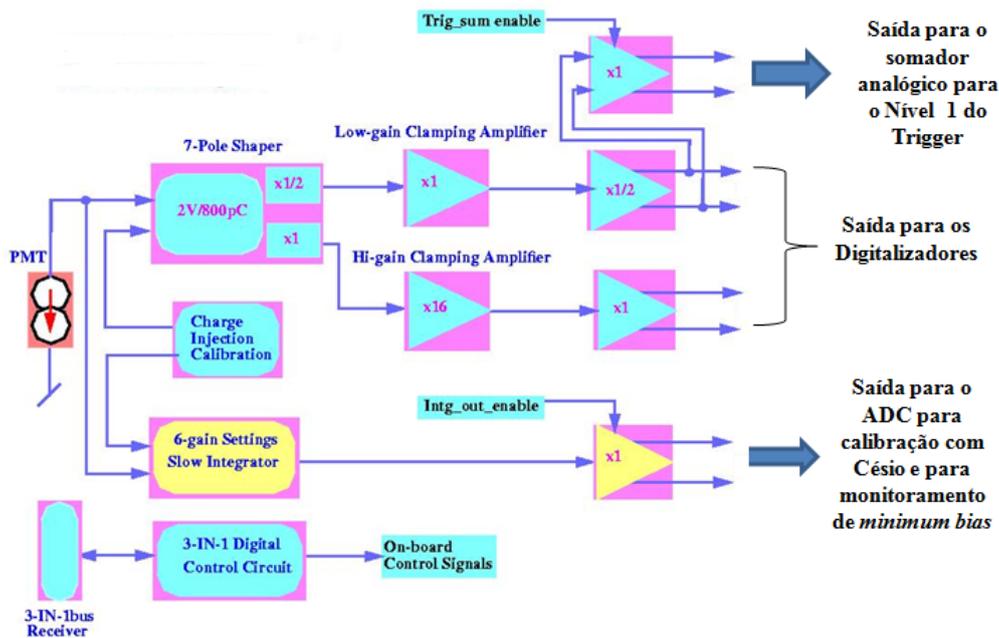


Figura 4.19: Diagrama da placa 3-in-1. Fonte: [14].

diagrama mostrando os componentes e o caminho do sinal nos digitalizadores pode ser visualizado na Figura 4.20.

4.4.2.6 Placas de Interface

As placas de interface tem duas funcionalidades principais, a primeira é receber as informações TTC (realizado através de duas fibras ópticas) e distribuí-lo para os digitalizadores equipados com o *chip* TTCrx. A segunda tarefa é a de receber os dados serializados das 8 placas digitalizadoras de uma gaveta, desrandomizá-los e enviá-los, através de uma ligação óptica, para a ROD. Assim como nas células do TileCal, a dupla leitura é utilizada também na interface link com o objetivo de prover redundância na comunicação, duas fibras ópticas transmitindo o mesmo sinal TTC e as informações de leitura.

4.4.2.7 Eletrônica de Back-End

A placa de *Read-Out Driver*, conhecida como ROD, é o elemento central da eletrônica de *back-end* do TileCal [13]. A ROD é um módulo VME 9U, que recebe como entrada as fibras provenientes das placas de interface link, contendo as amostras digitalizadas do pulso eletrônico a uma taxa de 100 kHz. Uma ROD pode trabalhar com até 8 fibras em sua entrada, provenientes de diferentes super-gavetas. Sendo assim, 32 ROD são necessárias para ler todo o TileCal. A ROD processa os dados

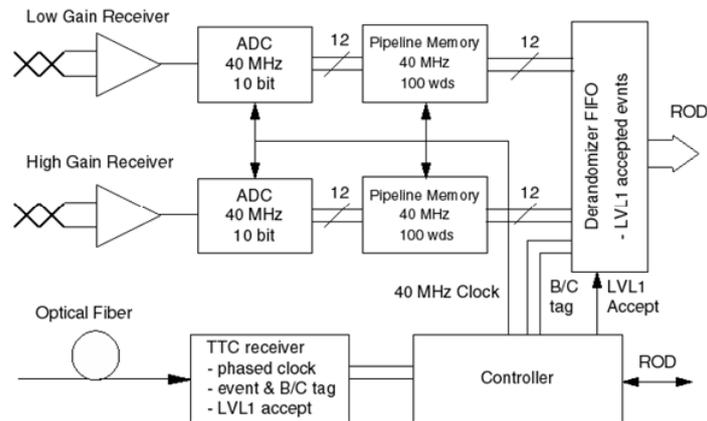


Figura 4.20: Diagrama da placa Digitalizadora. Fonte: [14].

online e é responsável pela detecção de erros antes de transmitir os dados para o nível 2 do *trigger* (LVL2).

4.4.3 Fonte de Alimentação da Gaveta

Toda a eletrônica das gavetas é alimentada por fontes chaveadas desenvolvidas especificamente para este fim, chamadas de *Finger Low Voltage Power Supply* (fLVPS ou LVPS) [38]. Um exemplo é visto na Figura 4.21. Cada fonte de alimentação fornece energia para um único módulo do detector. A fonte de alimentação consiste de um envoltório contendo oito módulos de potência ou "bricks" (Figura 4.22), fornecendo uma gama de correntes e tensões diferente para cada componente de uma gaveta. Os diferentes módulos de potência utilizam a mesma topologia e projeto básico, incluindo o mesmo projeto da placa de circuito impresso. Cada módulo é configurado especialmente para a carga especial que atende.

Existe uma "placa fria", que corre lateralmente através do meio da caixa da fonte de alimentação, arrefecida por água, que fornece o arrefecimento térmico para a unidade de energia. Os componentes de alta potência do módulo de potência são conectados termicamente a esta placa fria. A LVPS também possui uma placa de interface chamada ELMB-MB [39]. Esta placa contém uma placa ELMB que possui conversores analógico-digital (ADC), conversores digital-analógico (DAC), uma interface de I/O digital, e uma interface para o sistema de controle CANBUS do calorímetro. Existem vários aspectos no projeto dessas fontes que são relevantes. Uma vez que as fontes estão localizadas no próprio detector, elas devem ser capazes de funcionar sob um forte campo magnético, por esta razão não é possível a utilização de fontes lineares para esta aplicação uma vez que os núcleos dos trans-

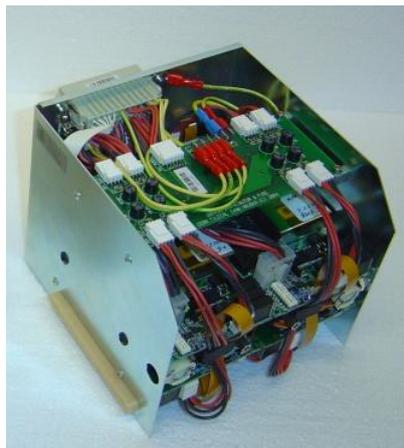


Figura 4.21: Imagem de uma unidade de fonte de alimentação (LVPS) com a tampa removida. Fonte: [15].



Figura 4.22: Módulo de potência individual (versão 6.5.4). Fonte:[15].

formadores iriam saturar [38]. Outro desafio no projeto da LVPS é manter baixo o ruído, conduzido e irradiado pelas fontes chaveadas, para a eletrônica de *front-end* analógica que processa sinais de baixa amplitude. Os módulos de potência chaveiam a 300 KHz, e produzem harmônicos através de toda a faixa espectral de interesse na eletrônica de *front-end*. Como a fonte esta localizada no próprio detector, ela estará sujeita a radiação, incluindo radiação gama, prótons e nêutrons. O tamanho compacto da caixa e o fato dela ser fechada representa um desafio em matéria de arrefecimento. A utilização da placa fria, juntamente com os contatos térmicos prudentemente colocados entre os módulos, visa resolver este problema. Por último, o acesso para o detector é muito difícil, podendo haver apenas uma oportunidade por ano para realizar operações de manutenção e reparos para o detector. Isto significa que o projeto dessas fontes precisa ser de alta confiabilidade.

4.4.4 Sistemas de Calibração do TileCal

Diversos sistemas são utilizados para calibrar e monitorar a performance do calorímetro hadrônico do ATLAS. Os sistemas a serem testados podem ser divididos em três partes: parte óptica (telhas cintiladoras e fibras óticas); os fotomultiplicadores (PMTs); e toda a eletrônica de leitura que molda, amplifica e digitaliza o sinal vindo das PMTs. Os principais sistemas de calibração e monitoramento são:

- O sistema de Laser: o sistema de Laser foi desenvolvido para calibrar e monitorar a resposta dos fotomultiplicadores, em geral, a estabilidade do seu ganho e sua linearidade [40].
- Sistema de fonte radioativa de Césio (Cs^{137}): foi desenvolvido para verificar a qualidade e uniformidade da resposta óptica de cada célula do calorímetro, e também para equalizar a resposta ao sinal de todas as células [40].
- Sistema de Injeção de Cargas (CIS): o sistema de injeção de cargas é usado para calibrar a resposta da parte eletrônica de leitura dos fotomultiplicadores.

No esquema apresentado na Figura 4.13, os boxes amarelos representam as diferentes calibrações existentes na cadeia de leitura eletrônica do TileCal.

Capítulo 5

Detecção e Estimação de Sinais

A seguir, será apresentada uma breve introdução da história da teoria de detecção e estimação de sinais. Será enfatizada, neste capítulo, a teoria de estimação de parâmetros em sinais utilizando estimadores de mínima variância, assim como os critérios utilizados para se medir a performance e precisão de um estimador.

5.1 Histórico

O período de tempo de grande desenvolvimento na história da teoria de detecção e estimação de sinais nos remete aos anos da Segunda Guerra Mundial. As primeiras observações em detecção e estimação de sinais estão relacionadas à história e desenvolvimento do Radar, isto porque foi onde os problemas de detectar sinais corrompido por ruído e estimar seus parâmetros começaram a aparecer. Nesse contexto um conjunto de temas que se sobrepõem foram sendo desenvolvidos até chegarmos às técnicas que conhecemos hoje em dia: a teoria de detecção de sinais, teoria da estimação de parâmetros dos sinais e a teoria do ruído. O estudo da teoria de ruído criou as fundações para o trabalho estatístico e, em muitos trabalhos iniciais sobre ruído, a teoria de detecção e estimação estiveram estreitamente entrelaçadas.

A teoria de ruído, inicialmente incluía duas principais áreas distintas: o ruído físico, tendo a ver com as origens físicas do ruído elétrico e caracterização de suas propriedades; e parte da teoria matemática de processos estocásticos, que é útil na descrição do comportamento e características do ruído elétrico em circuitos e dispositivos.

A teoria física do ruído foi iniciada em definitivo em meados de 1920 com o trabalho de W. Schottky "*Small-Shot Effect and Flicker Effect*" [41] e por H. Nyquist com o trabalho "*Thermal Agitation of Electricity in Conductors*" [42]. O tema se tornou uma teoria mais madura por volta dos anos 1940 com uma grande quantidade de confirmações experimentais; livros como "*Threshold Signals*" de J. L. Lawson e G. E. Uhlenbeck [43], "*Statistical Communication Theory*" de D. Middleton [44] são

exemplos do amadurecimento da teoria. Considerado como um marco divisório no estudo de ruído, os artigos de S.O. Rice que foram reunidos "*Mathematical Analysis of Random Noise*" [45] publicados em 1944 e 1945, juntaram grande parte de anos de resultados de trabalhos notáveis em teoria de ruído e contribuíram com muitas ideias e resultados novos.

O desenvolvimento do Radar durante a Segunda Guerra Mundial forçou um rápido desenvolvimento da teoria de detecção. A maioria dos trabalhos publicados nesta área, a partir de meados dos anos 1940 para cerca de 1950, surgiu a partir de estudos de radar em tempo de guerra. O campo de detecção de sinal e teoria de estimativa de parâmetros começou a mudar de rumo nos anos 1953-1954. Em primeiro lugar, novas áreas de aplicação destes métodos se multiplicaram e cresceram rapidamente como, por exemplo, investigação sísmica, sonares, tomografia, física de partículas entre outras. A teoria de detecção começou a deixar de ser estritamente uma teoria de radares. Em segundo lugar, houve uma infusão de algumas técnicas matemáticas bastante sofisticadas, diferente do que tinha sido usado anteriormente. Novos estudos matemáticos na área surgiram e relevantes livros apareceram investigando a nova e promissora área de estudo. O primeiro destes foi o livro de H. Cramer, "*Mathematical Methods of Statistics*" publicado em 1946 [46]. Outros livros mais específicos também surgiram como: "*Sequential Analysis Statistical*" (1947) [47] e "*Decision Functions*" (1950) de A. Wald [48]; "*Stochastic Processes*" (1953) de J.Doob [49]; e "*Probability Theory*" (1955) de M. Loève [50]. Além destes livros, apareceu em 1950 a tese de U. Grenander "*Stochastic Processes and Statistical Inference*" [51]. Este foi um trabalho marcante, pois trazia, de forma sistemática, compreensiva e de forma bastante rigorosa a teoria de processos estocásticos e a inferência estatística - conforme indicado no título. Com este vasto material de apoio na área de teoria da medida e análise funcional, o estudo da Teoria da Detecção se tornou muito mais disponível e, ao mesmo tempo, ficou realmente viável para um engenheiro, físico, ou matemático aplicado trabalhar em problemas em teoria de detecção ou teoria de ruído com rigor matemático e sofisticação.

5.2 Estimação de Parâmetros em Sinais

A Teoria moderna de estimação pode ser encontrada no coração de muitos sistemas eletrônicos de processamento de sinais desenvolvidos para extrair informações. Estes sistemas incluem, Radar, Sonar, Voz, Imagem, Biomedicina, Comunicações, Controle e Sismologia. Todos compartilham de um mesmo contratempo em comum, a necessidade de estimar o valor de um grupo de parâmetros geralmente corrompidos por ruído, tipicamente aditivo. Em alguns desses exemplos, um ser humano pode interpretar e discernir o sinal desejado em meio ao ruído. O real problema, então, é

desenvolver um mecanismo automático capaz de realizar a mesma tarefa.

O sinal medido como resultado desses sistemas pode ser expresso como $x(t)$, que é uma combinação do sinal desejado $s(t; \theta)$ e o ruído aditivo $w(t)$. Assim, $x(t)$ pode ser expresso como:

$$x(t) = s(t; \theta) + w(t) \quad (5.1)$$

Devido ao uso de computadores para amostrar e armazenar o conjunto de dados podemos expressar o conjunto de dados no domínio do tempo discreto, através de amostras do sinal observado. Uma combinação do sinal real e o ruído é expressa como:

$$x(n) = s(n; \theta) + w(n) \quad (5.2)$$

Matematicamente, nós temos um conjunto de dados com N-pontos $x[0], \dots, x[N - 1]$ que depende de um parâmetro desconhecido θ . O objetivo então é determinar o parâmetro θ baseado no conjunto de dados definindo um estimador através de alguma função $g(x)$, tal que:

$$\hat{\theta} = g(x[0], x[1], \dots, x[N - 1]). \quad (5.3)$$

Como os dados são inerentemente aleatórios, para se determinar bons estimadores, o primeiro passo é ter um bom modelamento matemático dos dados; a melhor maneira de se fazer isso é descrevendo o conjunto de dados por sua função de densidade de probabilidade (PDF) ou $p(x[0], x[1], \dots, x[n]; \theta)$. A função PDF é parametrizada pelo parâmetro desconhecido θ , ou seja, temos uma classe de pdfs em que cada uma é diferente, devido a um valor diferente de Teta.

Por exemplo, dado $N = 1$ e o parâmetro θ sendo a média, então uma PDF dos dados pode ser dada por:

$$p(x[0]; \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{1}{2\sigma^2} (x[0] - \theta)^2 \right] \quad (5.4)$$

O desempenho de qualquer estimador obtido será criticamente dependente dos pressupostos escolhidos para a PDF. Podemos apenas esperar que o estimador obtido seja robusto, e que pequenas mudanças na PDF não afetem severamente o desempenho do estimador. Estimação baseada em PDF é denominada como estimação clássica, na qual os parâmetros de interesse assumem-se ser determinísticos mas desconhecidos [52]. Em um problema real onde não é dada uma PDF, devemos escolher uma PDF que seja matematicamente tratável e consistente com as restrições do problema e com o conhecimento prévio que temos a respeito do sistema em questão.

Quando conhecemos *a priori* alguma característica a respeito do parâmetro a ser estimado, como o alcance deste parâmetro, podemos incorporar esse conhecimento prévio na solução do problema. Para incorporar esta informação, podemos assumir que o parâmetro (θ) não é mais um parâmetro determinístico e sim uma variável aleatória e atribuir-lhe uma PDF uniforme dentro do intervalo de alcance conhecido, e todo estimador produzirá valores dentro deste intervalo. O parâmetro que estamos tentando estimar será então visto como uma realização da variável θ . Tal abordagem é conhecida como estimação Baesiana e é descrita pela PDF conjunta:

$$p(x, \theta) = p(x|\theta)p(\theta) \quad (5.5)$$

onde $p(\theta)$ é a PDF *a priori*, representando nosso conhecimento acerca de θ antes que qualquer dado tenha sido observado, e $p(x|\theta)$ é a PDF condicional, resumando nosso conhecimento fornecido pelo conjunto de dados \mathbf{x} condicionado ao conhecimento de θ [52].

5.3 Desempenho de um estimador

Uma vez obtido um estimador, várias questões vêm a mente: quão perto este estimador está do valor real de \mathbf{A} ? Será este o melhor estimador do parâmetro? Considerando, como exemplo, um conjunto de medidas $x[n]$ em um experimento de medida de um nível um nível DC (\mathbf{A}) imerso em um ruído $w[n]$ (com média zero), podemos modelar o conjunto de dados corrompido com ruído aditivo como:

$$x[n] = A + w[n]. \quad (5.6)$$

Baseado no conjunto de dados $x[n]$, o objetivo é estimar \mathbf{A} da melhor maneira possível. Intuitivamente, uma vez que \mathbf{A} é o nível médio de $x[n]$, parece ser razoável estimar \mathbf{A} através da média de $x[n]$:

$$\hat{A} = \frac{1}{N} \sum_{n=0}^{N-1} x[n] \quad (5.7)$$

Um outro estimador que poderíamos considerar para este exemplo seria pegar a primeira amostra do conjunto de dados $\tilde{A} = x[0]$.

Para avaliar a performance de cada estimador devemos fazer isto estatisticamente. Uma das possibilidades seria a de repetir a experiência que gerou os dados, e aplicar cada estimador para cada conjunto de dados gerados, e então, apreciar qual estimador produziu uma estimativa melhor na maioria dos casos. Esta solução, porém, é suscetível a questionamentos e ao ceticismo quanto a quantidade de vezes que devemos repetir o processo a fim de obter uma conclusão quanto a escolha do

melhor estimador. Para provar que um dos estimadores é melhor do que o outro podemos demonstrar que a variância de um dos estimadores é menor.

As premissas de modelagem que empregaremos neste exemplo são de que o ruído $w[n]$, além de ter média zero, é decorrelacionado com variância igual a σ^2 . Então, primeiro demonstramos que a média de cada estimador é o valor verdadeiro \mathbf{A} e que na média, os estimadores encontrarão o valor verdadeiro do parâmetro.

$$E(\hat{A}) = E\left(\frac{1}{N} \sum_{n=0}^{N-1} x[n]\right) = \frac{1}{N} \sum_{n=0}^{N-1} E(x[n]) = A, \quad (5.8)$$

$$E(\tilde{A}) = E(x[0]) = A, \quad (5.9)$$

onde o operador $E[X]$ representa a média de uma variável aleatória X , ou seja, $\bar{X} = E[X]$. Então os estimadores na média alcançam o valor verdadeiro de \mathbf{A} . Agora calculamos a variância dos estimadores:

$$var(\hat{A}) = var\left(\frac{1}{N} \sum_{n=0}^{N-1} x[n]\right) = \frac{1}{N^2} \sum_{n=0}^{N-1} var(x[n]) = \frac{1}{N^2} N \sigma^2 = \frac{\sigma^2}{N},$$

e,

$$var(\tilde{A}) = var(x[0]) = \sigma^2.$$

Comparando as variâncias dos dois estimadores propostos temos então que:

$$var(\tilde{A}) > var(\hat{A}).$$

Sendo assim, como a variância do estimador \hat{A} é menor, podemos concluir que ele é mais preciso em medir o valor real do parâmetro do experimento em questão, sendo o melhor estimador dentre os propostos. Como o estimador é uma variável aleatória, sua performance só pode ser descrita completamente de forma estatística ou por sua PDF.

5.3.1 Estimadores não tendenciosos

Além de possuir baixa variância, outra característica importante que procuramos quando buscamos um estimador está relacionada a sua tendenciosidade, ou seja, é importante que o estimador seja não tendencioso (*unbiased*). Para um estimador ser não tendencioso, significa que, em média, o estimador produzirá o verdadeiro valor do parâmetro desconhecido.

Uma vez que o valor do parâmetro pode, em geral, estar em qualquer lugar dentro de um intervalo $a < \theta < b$, sua não tendenciosidade afirma que não importa

qual seja o valor verdadeiro de θ , a nossa estimativa irá produzir este valor na média. Matematicamente, um estimador não é tendencioso se:

$$E(\hat{\theta}) = \theta, \quad a < \theta < b \quad (5.10)$$

No exemplo dado da Seção 5.3, de um nível DC em meio a ruído Gaussiano aditivo, os dois estimadores apresentados não eram tendenciosos, como demonstrado pelas equações 5.8 e 5.9.

O fato de um estimador não ser tendencioso não significa necessariamente que ele é um bom estimador, porém, nós garantimos que na média ele irá atingir o valor verdadeiro. Estimadores tendenciosos são caracterizados por um erro sistemático, e um "bias" estocástico sistemático no estimador sempre resultará em uma péssima estimação do parâmetro.

Uma estimação não tendenciosa tem uma propriedade interessante que pode ser explorada quando possuímos mais de um estimador. Quando ocorrem múltiplas estimações do mesmo parâmetro, $(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_n)$, uma maneira razoável de combinar estas estimações, a fim de encontrar uma estimação melhor, é realizando a média destas estimações:

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n \hat{\theta}_i \quad (5.11)$$

Assumindo que os estimadores não são tendenciosos, possuem a mesma variância e são descorrelacionados, o valor esperado da combinação de diferentes estimações e sua variância serão dadas por:

$$E(\hat{\theta}) = \theta,$$

e

$$\begin{aligned} \text{var}(\hat{\theta}) &= \frac{1}{n^2} \sum_{i=1}^n \text{var}(\hat{\theta}_i) \\ &= \frac{\text{var}(\hat{\theta}_i)}{n}, \end{aligned}$$

assim sendo, quanto maior o número de estimativas combinadas, mais a variância irá diminuir. Contudo se os estimadores forem tendenciosos, isto é:

$$E(\hat{\theta}_i) = \theta + b_i(\theta),$$

então o valor esperado para estes estimadores será:

$$E(\hat{\theta}) = \frac{1}{n} \sum_{i=1}^n E(\hat{\theta}_i) = \theta + b(\theta).$$

Portanto, a não ser no caso raro onde o *bias* da combinação dos estimadores se cancelem, em geral, os estimadores tendenciosos combinados nunca alcançarão o valor real do parâmetro a ser estimado.

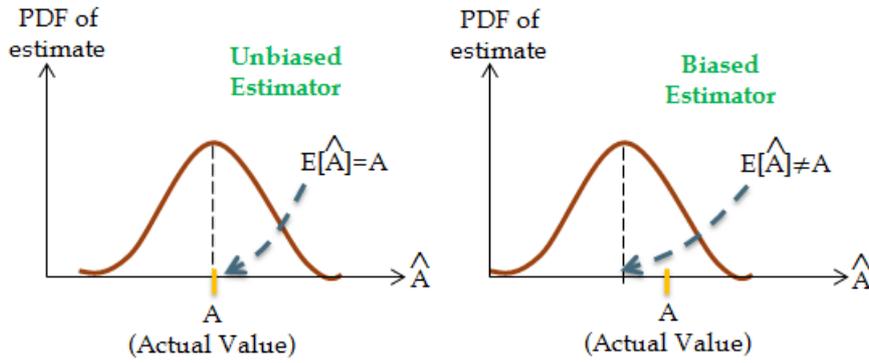


Figura 5.1: Na esquerda a PDF de um estimador não tendencioso (*unbiased*) e a direita um estimador tendencioso (*biased*). Fonte:[16]

5.3.2 Critério da Mínima Variância

Na busca por um estimador que seja ótimo, precisamos adotar algum critério de avaliação. Uma abordagem muito utilizada é a de restringir o *bias* a zero e encontrar o estimador que minimize a variância da estimação. Tal estimador é denominado como estimador imparcial de variância mínima (MVU em inglês).

A questão que surge então é saber se existe um estimador MVU, ou seja, um estimador imparcial com variância mínima para o intervalo de θ procurado.

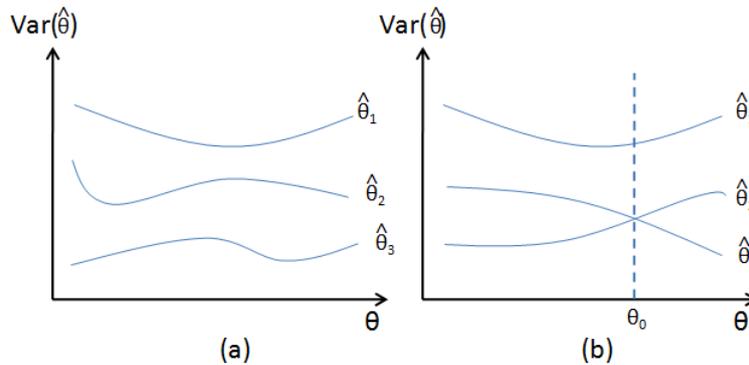


Figura 5.2: Em (a) temos um estimador MVU ($\hat{\theta}_3$), já em (b) não existe estimador MVU.

Duas situações possíveis são ilustradas na Figura 5.2. Se existem três estimadores cujos desvios são mostrados na Figura 5.2 (a), então claramente $\hat{\theta}_3$ é o estimador MVU. No entanto, se a situação na Figura 5.2 (b) existe, então não há estimador MVU já que para $\theta_2 < \theta_0$, $\hat{\theta}_2$ é melhor, enquanto para $\hat{\theta} > \theta_0$, $\hat{\theta}_3$ é melhor. No primeiro caso o estimador é por vezes referido como um estimador uniforme de variância mínima para enfatizar que a variância é menor para todo Teta. Em geral,

o estimador MVU não existe sempre para toda a faixa dinâmica.

5.3.3 Precisão do Estimador

Uma vez que toda a nossa informação está incorporada nos dados observados e que a PDF está relacionada a esses dados, a precisão da estimativa depende diretamente da PDF. Não espera-se ser possível estimar um parâmetro com qualquer grau de precisão se a PDF depender fracamente deste parâmetro ou, em caso extremo, se a PDF não depender dele. Portanto, quanto mais a PDF sofre influência deste parâmetro, melhor deve ser nossa capacidade de estimação.

Utilizando como exemplo uma única observação $x[0]$ de um experimento onde desejamos estimar a variável DC (A) corrompida com ruído $w[0]$ (com média zero e variância σ^2) :

$$x[0] = A + w[0]$$

podemos notar que a variância da amostra é igual a σ^2 ; portanto, quanto menor a variância, melhor será a estimativa do parâmetro. Um estimador não tendencioso conhecido para este exemplo é dado por $\hat{A} = x[0]$ (ver Seção 5.3).

Quando uma PDF é vista como uma função do parâmetro desconhecido então ela é denominada como função de verossimilhança (*likelihood function*). A agudez da função de verossimilhança determina o quão preciso podemos estimar o parâmetro desconhecido. Para quantificar esta noção observamos que a agudez é efetivamente medida pela função negativa da segunda derivada do logaritmo da função *likelihood* no seu pico. Na geometria, a agudez de uma curva é medida através de sua curvatura; como na geometria diferencial a curvatura está relacionada com a segunda derivada de uma função, efetivamente podemos medir a agudez da função de verossimilhança através da negativa de sua segunda derivada.

Revisitando o exemplo anterior, sendo a PDF do conjunto de observação dada por:

$$p_i(x[0]; A) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp \left[-\frac{1}{2\sigma_i^2} (x[0] - A)^2 \right] \quad (5.12)$$

seu logaritmo natural será:

$$\ln p(x[0]; A) = -\ln\sqrt{2\pi\sigma^2} - \frac{1}{2\sigma^2} (x[0] - A)^2 \quad (5.13)$$

então seu gradiente, também conhecido como *score*, dado pela sua primeira derivada é:

$$\frac{\partial \ln p(x[0]; A)}{\partial A} = \frac{1}{\sigma^2} (x[0] - A) \quad (5.14)$$

e calculando o negativo da segunda derivada (a curvatura da PDF) teremos:

$$-\frac{\partial^2 \ln p(x[0]; A)}{\partial A^2} = \frac{1}{\sigma^2} \quad (5.15)$$

Portanto, podemos observar que a curvatura da PDF aumenta conforme sua variância diminui. Um estimador para este exemplo é dado por $\hat{A} = x[0]$ e sua variância será igual a σ^2 , então para este exemplo temos que:

$$\text{var}(\hat{A}) = \frac{1}{-\frac{\partial^2 \ln p(x[0]; A)}{\partial A^2}}, \quad (5.16)$$

assim sendo, quanto maior a curvatura menor a variância do estimador.

5.3.4 Limite Inferior de Cramer-Rao

Em um cenário prático, onde buscamos um estimador ótimo para encontrar o parâmetro no conjunto de dados em nosso problema, podemos nos deparar com a questão de saber qual dentre os estimadores encontrados é de fato o melhor estimador imparcial e com variância mínima; para isto podemos comparar a variância destes estimadores entre si e encontrar o estimador MVU dentre os estimadores que possuímos. Porém, como garantir que o estimador imparcial de mínima variância encontrado é de fato o melhor estimador possível para resolver nosso problema? Poderá existir outro estimador MVU melhor do que o encontrado que nos forneça uma estimacão melhor? Dado estes questionamentos, encontrar um limite inferior para a variância de qualquer estimador imparcial revela-se extremamente útil na prática. E para isto o Limite Inferior de Cramer-Rao (CRLB em inglês) vem em nossa assistência.

Na melhor das hipóteses, calcular um limite inferior na variância nos permite saber que um estimador é de fato o melhor estimador MVU que existe para dado problema, este será o caso em que o estimador atinge o limite mínimo de Cramer-Rao para todos os valores do parâmetro desconhecido. Na pior das hipóteses, ele fornece uma referência contra a qual podemos comparar o desempenho de qualquer estimador. Além disso, possuir um limite inferior pode nos precaver da impossibilidade de encontrarmos algum estimador com variância menor do que a do limite de Cramer-Rao. Apesar de existir outros limites para a variância, por exemplo, McAulay e Hofstetter (1971); Kendall e Stuart (1979); Sidman (1970); Ziv e Zakai (1969); o limite de Cramer-Rao é de longe o mais utilizado em estimacão de sinais devido a sua praticidade de ser determinado.

Assumindo que a PDF do conjunto de dados $p(x; \theta)$ satisfaça a condição de regularidade, onde o valor esperado da primeira derivada do logaritmo da função de verossimilhança é zero:

$$E \left[\frac{\partial \ln p(x; \theta)}{\partial \theta} \right] = 0$$

então a variância de um estimador imparcial ($\hat{\theta}$) deve satisfazer a seguinte condição:

$$\text{var}(\hat{\theta}) \geq \frac{1}{-E \left[\frac{\partial^2 \ln p(x; \theta)}{\partial \theta^2} \right]} \quad (5.17)$$

onde a derivada é calculada em função de θ e o valor esperado tomado em relação a $p(x; \theta)$. Além disso, podemos encontrar um estimador imparcial que alcance o limite de CR para todo θ se e somente se

$$\frac{\partial \ln p(x; \theta)}{\partial \theta} = I(\theta)(g(x) - \theta), \quad (5.18)$$

para alguma função $g(\cdot)$ e $I(\cdot)$. Este estimador, será o estimador MVU se $\hat{\theta} = g(x)$, e a mínima variância for igual a $1/I(\theta)$. As demonstrações dessas condições do Teorema de Cramer-Rao encontram-se no Apêndice A.

O limite de Cramer-Rao pode ser escrito de uma forma diferente da Equação 5.17. Esta forma difere sutilmente e pode ser útil em muitas situações e é obtida a partir da identidade a seguir (ver Apêndice A):

$$E \left[\left(\frac{\partial \ln p(x; \theta)}{\partial \theta} \right)^2 \right] = -E \left[\frac{\partial^2 \ln p(x; \theta)}{\partial \theta^2} \right], \quad (5.19)$$

sendo assim

$$\text{var}(\hat{\theta}) \geq \frac{1}{E \left[\left(\frac{\partial \ln p(x; \theta)}{\partial \theta} \right)^2 \right]}. \quad (5.20)$$

O denominador da Equação 5.17 também é conhecido como Informação de Fisher $I(\theta)$ para o conjunto de dados x :

$$I(\theta) = -E \left[\frac{\partial^2 \ln p(x; \theta)}{\partial \theta^2} \right]$$

assim, quando o limite de CR é alcançado a variância do estimador será proporcional ao inverso da Informação de Fisher, e quanto mais informação menor será o limite. Assim sendo, o limite de CR possui as propriedades essenciais de medida de uma informação:

- Não negativa, devido a Equação 5.19;
- Aditiva para observações independentes;

Esta última propriedade tem como consequência o fato de que o limite de Cramer-Rao, para N observações independentes e igualmente distribuídas (IID), ser $1/N$ vezes menor do que para apenas uma observação.

Um estimador não tendencioso é tido como um estimador MVU ótimo para o conjunto de dados, se e somente se, sua variância alcançar o limite inferior de Cramer-Rao para todo θ . Na Figura 5.3 (a) temos três estimadores, porém apenas um deles é um estimador MVU de fato e como sua variância é igual ao limite de CR o estimador $\hat{\theta}_1$ é por definição o estimador ótimo para este parâmetro; já na Figura 5.3 (b) observamos novamente três estimadores e somente um estimador MVU, porém, como a variância deste estimador não coincide com o limite inferior de CR, ele não é o estimador ótimo, mas ainda assim dentre todos os demais estimadores ele é o que melhor estima o parâmetro do conjunto de dados.

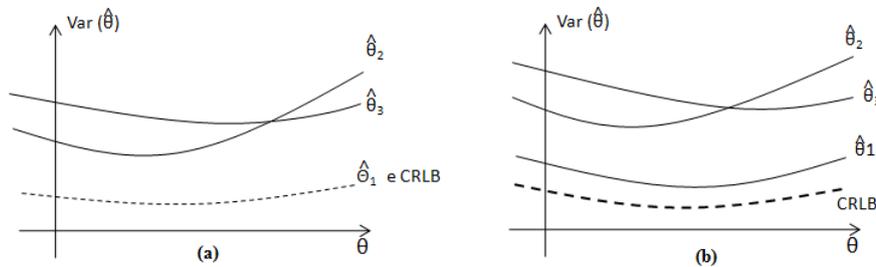


Figura 5.3: Em (a) temos um estimador MVU e ótimo, em (b) existe um estimador MVU, porém ele não é ótimo.

Dessa forma, o limite inferior de Cramer-Rao para o exemplo de uma amostra de um nível DC em meio a ruído ($x[0] = A + w[0]$), conhecendo-se as Equações 5.15 e 5.17, será dado por:

$$\text{var}(\hat{A}) \geq \sigma_w^2.$$

Generalizando o exemplo anterior para o caso de múltiplas observações ($x[n] = A + w[n]$) e utilizando a propriedade da aditividade do limite de CR, obtemos o CRLB:

$$\text{var}(\hat{A}) \geq \frac{\sigma_w^2}{N}. \quad (5.21)$$

Assim, a variância de qualquer estimador de nível DC, dado um conjunto de observações, será sempre maior ou igual ao limite inferior de Cramer-Rao.

5.4 Estimador Imparcial de Mínima Variância para o Modelo Linear

Determinar um estimador não tendencioso de mínima variância é em geral uma tarefa difícil. Contudo, como muitos problemas de estimação em processamento de sinais podem ser representados através de um modelo linear, o estimador MVU pode ser obtido. O modelo linear para os dados é definido da seguinte forma:

$$x[n] = A + Bn + w[n]$$

sendo $w[n]$ um ruído branco Gaussiano, B a inclinação e A o ponto de interseção. Podemos reescrever o modelo linear na forma matricial da seguinte forma:

$$x = H\theta + w \quad (5.22)$$

onde

$$\begin{aligned} x &= [x[0]x[1]\dots x[N-1]]^T \\ w &= [w[0]w[1]\dots w[N-1]]^T \\ \theta &= [AB]^T \end{aligned}$$

e

$$H = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ 1 & N-1 \end{bmatrix}$$

A matriz H é conhecida como matriz de observação. O vetor de ruído é caracterizado como $w \sim N(0, \sigma^2 I)$. Uma das maneiras mais utilizadas de se encontrar um estimador MVU é através do Teorema de CRLB (uma outra maneira é utilizando o Teorema de Rao-Black-Lehmann-Scheffe). Através da condição de igualdade do teorema de CRLB, o estimador $\hat{\theta} = g(x)$ será um estimador MVU se satisfizer a condição:

$$\frac{\partial \ln p(x; \theta)}{\partial \theta} = I(\theta)(g(x) - \theta). \quad (5.23)$$

Derivando a função de *likelihood* do modelo linear e assumindo que $H^T H$ é inversível, chegamos a seguinte equação (para a demonstração ver Apêndice A):

$$\frac{\partial \ln p(x; \theta)}{\partial \theta} = \frac{H^T H}{\sigma^2} [(H^T H)^{-1} H^T x - \theta] \quad (5.24)$$

que possui exatamente a forma da Equação 5.23 com

$$\begin{aligned}\hat{\theta} &= (H^T H)^{-1} H^T x, \\ I(\theta) &= \frac{H^T H}{\sigma^2}.\end{aligned}$$

A única restrição que se deve ter atenção é a condição de $H^T H$ ser inversível.

Sendo assim, dado um conjunto de dados que possa ser modelado através da Equação 5.22, onde x é um vetor de observações $N \times 1$, H uma matriz conhecida de observações $N \times p$ (com $N > p$) e posto p , θ um vetor $p \times 1$ de parâmetros a serem estimados, e w um vetor de ruído $N \times 1$ com PDF $N(0, \sigma^2 I)$, então o estimador de mínima variância é dado por:

$$\hat{\theta} = (H^T H)^{-1} H^T x, \quad (5.25)$$

e sua matriz de covariância será

$$C_{\hat{\theta}} = I^{-1}(\theta) = \sigma^2 (H^T H)^{-1}. \quad (5.26)$$

5.4.1 Estimador de Mínima Variância para o Modelo Linear Corrompido por Ruído Colorido

Podemos generalizar o resultado anterior e encontrar um estimador MVU para um conjunto de dados corrompido por um ruído colorido¹. Assim, o modelo geral assume que o ruído tem a seguinte característica

$$w \sim n(0, C),$$

onde C não é necessariamente uma matriz identidade. Para determinarmos o estimador podemos utilizar uma abordagem onde realizemos um branqueamento do conjunto de dados. Uma vez que a matriz de covariância C seja definida positiva, sua matriz inversa C^{-1} é positiva definida e pode ser fatorada utilizando a decomposição de Cholesky da seguinte forma:

$$C^{-1} = D^T D,$$

onde D é uma matriz inversível $N \times N$. A matriz \mathbf{D} age como uma matriz de branqueamento quando aplicada ao ruído (ver Apêndice B). Aplicando a matriz D ao conjunto de dados modelado segundo a Equação 5.22, obtemos a generalização do estimador MVU para ruído colorido (ver Apêndice B):

$$\hat{\theta} = (H^T C^{-1} H)^{-1} H^T C^{-1} x. \quad (5.27)$$

¹Ruído que apresenta componentes frequenciais em uma ou algumas frequências.

5.5 Estimador de Máxima Verossimilhança

Um estimador alternativo também foi considerado. Esse estimador baseado no princípio da máxima verossimilhança (MLE em inglês) é uma das abordagens mais populares para se obter estimadores práticos e é desejável em diversas situações onde o estimador MVU não pode ser encontrado ou não existe.

Para se determinar um parâmetro desconhecido α , o princípio básico desse estimador é encontrar o ponto de máximo da função de verossimilhança ($\mathcal{L}(\alpha)$) de um dado experimento que possui uma função densidade probabilidade $f(x; \alpha)$. O valor mais provável de α é chamado de solução de máxima verossimilhança α^* .

$$\mathcal{L}(\alpha) = \prod_{i=1}^N f(x_i, \alpha) \quad (5.28)$$

Quando temos uma aplicação em que cada medida feita x_i apresenta uma distribuição Gaussiana com um erro de medição diferente σ_i , ou seja:

$$x[n] = A + w[n]$$

onde A é o parâmetro a ser estimado e $w[n]$ é um ruído branco gaussiano com variância σ_i , sua PDF será:

$$f(x_i; A) = \frac{1}{(2\pi\sigma_i^2)} \exp \left[-\frac{(x_i - A)^2}{2\sigma_i^2} \right]. \quad (5.29)$$

Para N medidas x_i , a função de verossimilhança será dada por:

$$\mathcal{L}(A) = \prod_{i=1}^N \frac{1}{\sqrt{2\pi}\sigma_i} \exp \left[-\frac{(x_i - A)^2}{2\sigma_i^2} \right]. \quad (5.30)$$

Com isso, o logaritmo natural da função de verossimilhança é expressa por

$$w = \sum_{i=1}^N \ln \frac{1}{\sqrt{2\pi}\sigma_i} - \frac{1}{2} \sum_{i=1}^N \frac{(x_i - A)^2}{\sigma_i^2}, \quad (5.31)$$

derivando parcialmente essa função em relação ao parâmetro desconhecido obtemos

$$\frac{\partial w}{\partial A} = \sum_{i=1}^N \frac{x_i - A}{\sigma_i^2}.$$

Igualando a zero a derivada, obtemos o valor ótimo do parâmetro

$$\hat{A} = \frac{\sum_{i=1}^N \frac{1}{\sigma_i^2} x_i}{\sum_{i=1}^N \frac{1}{\sigma_i^2}}. \quad (5.32)$$

Portanto, a medida ótima é obtida ponderando-se as medidas individuais de acordo com o inverso do quadrado dos seus erros. Se todas as medidas possuírem o mesmo erro associado, então a medida ótima será a média aritmética das medidas individuais.

Capítulo 6

Redução de Ruído em Calorimetria

Neste capítulo, será apresentada a revisão da literatura dos métodos de redução de ruído adotado em experimentos de física de altas energias envolvendo calorimetria. Para tanto, foi realizado um estudo das principais técnicas utilizadas para mitigar o efeito dessas flutuações nesses tipos de experimentos.

6.1 Ruído em Calorímetros

Com o passar dos anos, a necessidade de investigarmos, com cada vez mais profundidade, a natureza da matéria fez com que a faixa de energia e a complexidade dos calorímetros envolvidos nesses experimentos se tornassem cada vez maiores, o que aumenta a exigência em desempenho para os sinais medidos. Em calorímetros de amostragem com meio ativo, que utilizam coleta direta da carga produzida, como na câmara de ionização (*Liquid Argon*) ou modo proporcional (câmaras de fios), os sinais, normalmente, acumulam alguns picocoulombs de carga por GeV de energia do chuveiro de partículas, havendo inevitavelmente uma contribuição de ruído eletrônico nos sinais. Isto significa que a energia coletada, mesmo na ausência de um chuveiro de partículas, sofre flutuações de evento para evento. O desvio padrão destas flutuações (σ_{noise}) é medido em unidades de carga.

Uma vez que o calorímetro mede a energia do chuveiro de partículas na mesma unidade, o termo de ruído é equivalente a uma certa quantidade de energia, geralmente da ordem de MeV. O valor preciso da contribuição do ruído depende de uma série de fatores, como o tempo de coleta de energia por evento, a capacitância do detector e propriedades da amplificação eletrônica da carga coletada. Uma vez que o desvio padrão do ruído eletrônico corresponde a uma certa quantidade de energia, a contribuição do ruído na resolução de energia do calorímetro é proporcional a E^{-1} ;

essa dependência com a energia mostra que este termo de ruído domina a resolução de energia em baixas energias, prejudicando a qualidade da medição de energia do calorímetro nessa região [27].

Em calorímetros de cintilação, onde os sinais dos fotomultiplicadores são digitalizados por meio de conversores analógico-digitais (*ADC*), na ausência de sinal na PMT, o ADC pode acumular uma certa quantidade de carga, resultando em um pedestal no sinal. Para encontrar os sinais resultantes da energia das partículas, o pedestal deve ser removido do sinal bruto. Contudo, o pedestal pode exibir flutuações devido a uma variedade de fontes de ruído e problemas eletrônicos, por exemplo, o mal casamento de impedância, mal aterramento das fontes de energia, entre outros [27]. Esta flutuação do pedestal afeta a resolução de energia dos calorímetros baseados em cintilação da mesma maneira que em calorímetros baseados em ionização. Parte dessas flutuações pode ser minimizada projetando-se, de maneira correta, os componentes eletrônicos envolvidos na leitura dos sinais; porém, dada a grande quantidade de canais de leitura, a extrema condição de operação dos componentes sujeitos a alto campos magnéticos e à radiação e a impossibilidade de se realizar uma manutenção periódica, a presença de ruído nas saídas dos canais torna-se um fato comum.

6.1.1 Ruído no TileCal

Antes do início das colisões no LHC, no fim de 2009, a descrição do ruído do Calorímetro Hadrônico do ATLAS sofreu uma substancial mudança de perfil, apresentando componentes não-Gaussianas. A distribuição do ruído, que esperava-se ser Gaussiana, mostrou uma característica de dupla-Gaussiana; a descrição do ruído do TileCal teve que ser alterada em diversos pacotes de análises e ferramentas de calibração. A Figura 6.1 mostra a distribuição de ruído em uma célula de um módulo do TileCal, onde a característica de dupla-Gaussianidade é observada: o *fitting* de dupla-gaussiana foi o mais adequado para descrever a distribuição de ruído. O principal motivo dessa mudança de perfil no ruído foi uma indesejável correlação do ruído (ver Figura 6.2) entre os canais de leitura em módulos do TileCal [18]. O ruído correlacionado no TileCal se manifestava em regiões de 6 ou 12 canais, que é a extensão respectiva dos digitalizadores e das placas-mãe de cada gaveta. Existem áreas na eletrônica de front-end (FE) da gaveta no TileCal que são mais sensíveis à interferência eletromagnética das *Low Voltage Power Supply* (LVPS), como os canais nas bordas entre duas gavetas [18].

Os módulos de potência das LVPS que alimentam as gavetas, versão 6.5.4 (ver Figura 4.22), foram produzidos e instalados entre 2006-2007. Durante o comissionamento do detector, as LVPS apresentaram alguns problemas de desempenho e

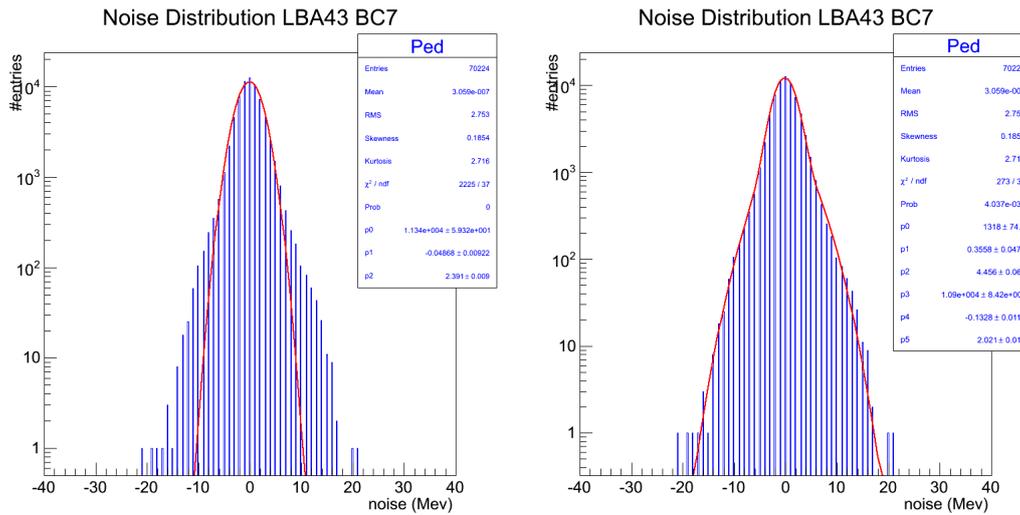


Figura 6.1: Na imagem a esquerda um fitting Gassiano na distribuição de ruído da célula. Na imagem a direita o fitting de dupla-gaussiana.

confiabilidade que preocuparam a comunidade do TileCal; porém, não comprometiam os requisitos mínimos exigidos. Destacamos como alguns dos problemas: o *design* das placas dos módulos da LVPS; solda e montagem; estabilidade; falhas espontâneas [15]. No entanto, uma vez que as colisões iniciaram com crescente luminosidade, o sistema demonstrou sensibilidade para falhas (*trips*), a uma taxa altamente correlacionada com o aumento da luminosidade das colisões (ver Figura 6.3).

Preocupada com a confiabilidade a longo prazo e com o aumento da luminosidade nas colisões futuras no LHC, a colaboração decidiu que a melhor maneira de tentar resolver esses problemas seria o desenvolvimento de novas versões dos módulos da LVPS; assim sendo, desde Junho de 2009 (antes mesmo do início do período de circulação de feixes no LHC), a colaboração TileCal iniciou um programa para o desenvolvimento de novas versões da LVPS. Os problemas críticos que precisaram ser resolvidos foram: redução do ruído; travamento dos opto-isoladores; melhor gerenciamento térmico dos módulos; uma melhor proteção contra descargas eletrostáticas (ESD) dos circuitos integrados (ICs) dos módulos de potência da LVPS [15]. Diversas versões aprimoradas da LVPS foram desenvolvidas e testadas a partir de dezembro de 2009 (versão 7.1.0 com 10 módulos produzidos), e durante a parada técnica de 2013, milhares de módulos de potência versão 7.5.1 foram instalados nos módulos do detector.

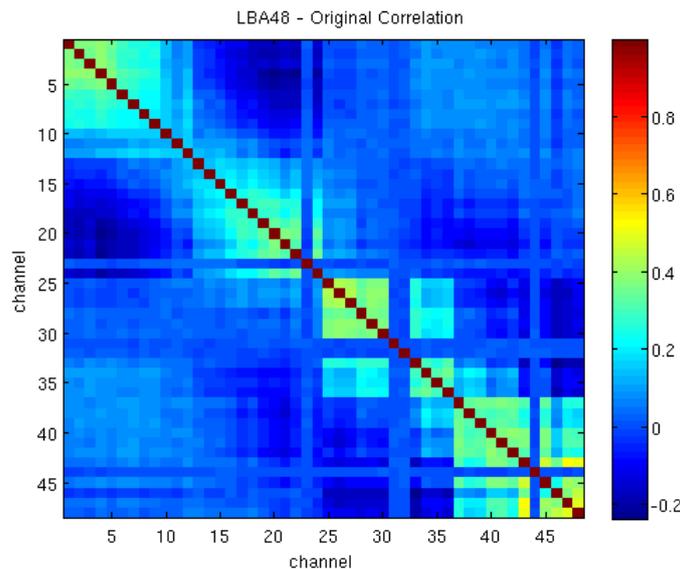


Figura 6.2: Ruído correlacionado no módulo 48 do TileCal.

6.1.2 Ruído de Pile-up

Além das fontes de ruído decorrentes da eletrônica de leitura e da fonte de energia da gaveta, o sinal gerado por um evento de colisão de um pacote de partículas no TileCal pode sofrer a interferência de flutuações proveniente do empilhamento de sinais (*Pile-up*). Estas flutuações podem ser de dois tipos: *In Time Pile-up* e *Out of Time Pile-up*. Na primeira, o evento formador do sinal de interesse sofre a interferência de sinais produzidos por partículas que emergem de vértices secundários dentro do mesmo *bunch spacing*. Na segunda, ocorre uma sobreposição de sinais oriundos de diferentes colisões no detector. Isto ocorre quando a janela de tempo entre as colisões é menor do que o tempo que a eletrônica de leitura do detector necessita para gerar o sinal correspondente de cada colisão. Esta sobreposição de impulsos deslocados no tempo resulta em mais uma fonte de ruído que compromete o sinal do detector. O *pile-up* também aumenta conforme aumenta a luminosidade dos feixes, pois aumenta o número de eventos por colisão e, assim, a ocupação do detector. O ruído de *pile-up* causa um impacto não-gaussiano na distribuição de energia.

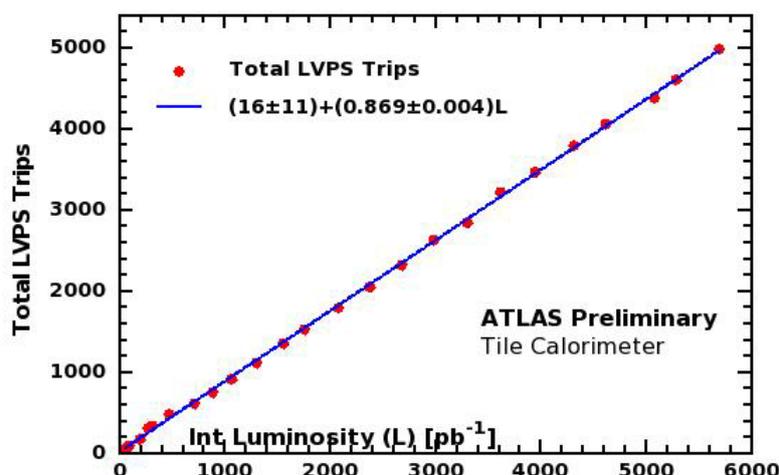


Figura 6.3: Relação entre o número de falhas nas LVPS e o aumento de luminosidade. Fonte:[15].

6.2 Tratamento de Ruído em um Calorímetro de Argônio Líquido

Um dos trabalhos mais antigos e importantes, onde o problema do ruído em calorimetria aparece de maneira direta como um empecilho ao desenvolvimento do campo de estudo, foi realizado por W. J. Willis e V. Radeka, intitulado "Liquid-Argon Ionization Chambers as Total-Absorption Detectors" [17]; o trabalho de 1974 apresenta as principais características do então novo tipo de detector para a medida de energia através de absorção total, baseado em múltiplas placas de câmaras de íons, chamado câmara de Argônio líquido. O ruído nesse tipo de experimento pode ter várias origens; por exemplo, ruído do amplificador, *pile-up* (devido às partículas de fundo ou em casos de detectores de múltiplas partículas, o vazamento de energia de uma partícula de um dado evento, vazando o espaço de domínio de outra partícula). Já em 1974, os autores do trabalho afirmam que, para medir a carga de grandes câmaras de ionização com variações insignificantes de ruído eletrônico, um sistema otimizado de processamento de sinal seria essencial e discutem os elementos básicos para atingirem uma medição ótima da carga no calorímetro apresentado.

Um grande calorímetro, como o mostrado no artigo, envolve um grande número de placas (Figura 6.4) e, por conseguinte, uma grande capacitância. O detector de teste descrito neste artigo tem uma capacitância de cerca de $10^5 pF$. A medição de carga é realizada observando-se uma fração da carga total produzida pela ionização, que é desviada para o amplificador de carga de medição. O melhor dispositivo

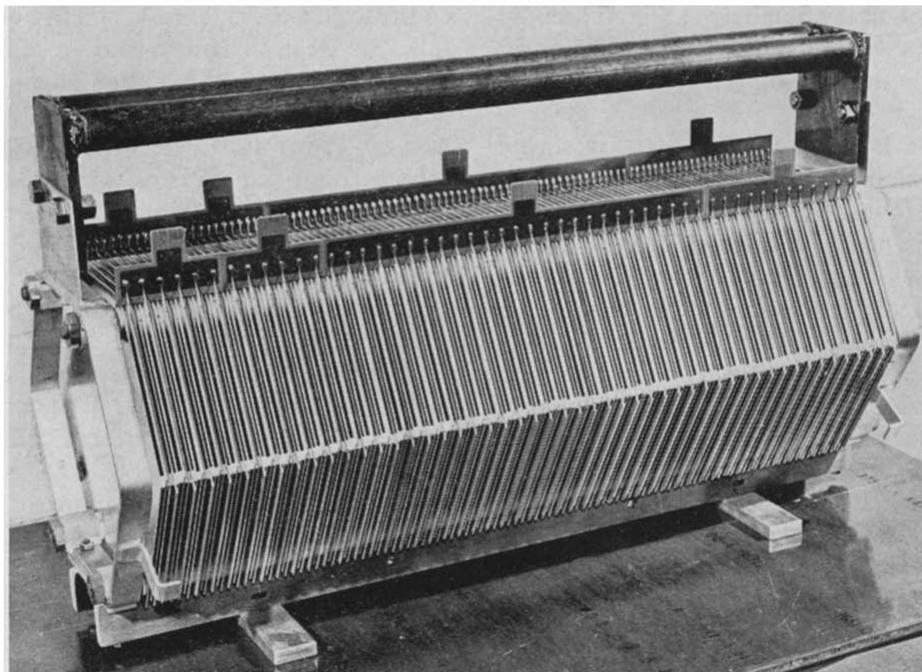


Figura 6.4: Grande câmara de teste com 200 chapas de aço com 1,5 mm de espessura, com lacunas de 2 mm. Fonte: [17].

amplificador de baixo ruído disponível na época tratava-se do transistor de efeito de campo (FET), que possuía uma capacitância de entrada da ordem de 10 pF. Conectado diretamente à câmara de ionização, tal amplificador "veria" somente uma parte em 10^4 da carga total, o que não é o bastante, comparado ao ruído do amplificador. Portanto, esta desigualdade é um caso óbvio de descasamento de impedância entre o detector e o amplificador. Um melhor compartilhamento de carga (e uma maior razão de sinal-ruído) pode ser alcançada se o detector e amplificador estiverem casados através do uso de um transformador, como pode ser visto na configuração mostrada na Figura 6.5.

Portanto, o trabalho neste artigo se concentra na realização do projeto de um transformador que faça o casamento da capacitância entre o detector e o circuito amplificador, para melhorar a relação entre sinal e ruído do detector.

6.3 Método do Chi-Quadrado para Descorrelação de Ruído em um Calorímetro Hadrônico

O trabalho "*Correlated noise unfolding on a Hadronic Calorimeter*" realizado por M. C. N. Fiolhais [18], propôs um método de descorrelação do ruído dos canais do

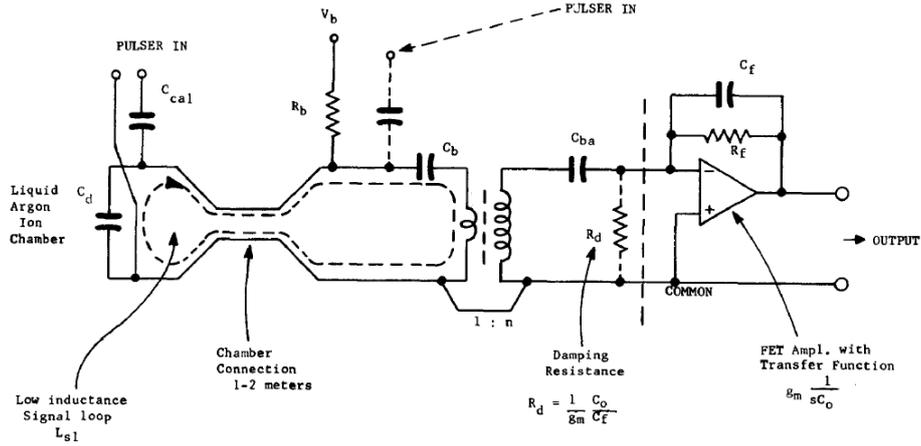


Figura 6.5: Circuito de configuração detector-amplificador. Fonte: [17]

TileCal.

Como visto na sub-Seção 6.1.1, o problema do ruído correlacionado no TileCal desempenhou um papel importante na degradação nas flutuações do pedestal do TileCal. Para solucionar este problema, o trabalho considerou uma abordagem geral com base em princípios fundamentais que não dependem da fonte específica do ruído. A abordagem apresentada assumiu que a medição de ruído observado (x_i), em um canal i de um dado módulo do TileCal, é uma combinação de componentes de ruído (β_i), somado a uma contribuição que depende da resposta de todos os demais canais de leitura do módulo como um todo e provavelmente dominada pelos canais vizinhos mais próximos. A abordagem mais simples para descrever a medição do ruído no canal i foi considerar x_i como sendo uma combinação linear entre a componente de ruído intrínseco (β_i) e uma soma ponderada dos sinais de todos os outros canais de leitura (N) no módulo, ou seja:

$$x_i = \beta_i + \sum_{j \neq i}^N \alpha_{i,j} x_j, \quad (6.1)$$

onde x_i e x_j são sinais de ruído dos canais i e j respectivamente. Estes sinais referem-se à amplitude do sinal já reconstruídos pelo método de *Optimal Filter* não iterativo [53]. Os parâmetros $\alpha_{i,j}$ garantem que o ruído medido de outros canais seja levado em conta com pesos diferentes. Dado o fato de que os sinais de pedestal subtraídos ao redor de zero são usados, os valores β_i deverão ser zero. O ruído medido de cada canal pode ser comparado com o modelo apresentado usando o método de chi-quadrado:

$$\chi^2 = \sum_{Events} \frac{\left[x_i - \left(\beta_i + \sum_{k \neq i}^N \alpha_{i,k} x_k \right) \right]^2}{\sigma_i^2}, \quad (6.2)$$

que pode ser minimizado (individualmente para cada canal) com respeito a $\alpha_{i,j}$ e β_i do modelo,

$$\frac{\partial \chi_i^2}{\partial \alpha_{i,1}} = \frac{\partial \chi_i^2}{\partial \alpha_{i,2}} = \dots = \frac{\partial \chi_i^2}{\partial \alpha_{i,N}} = \frac{\partial \chi_i^2}{\partial \beta_i} = 0 \quad (6.3)$$

Esta técnica assume que as correlações de ruído são as mesmas, independentemente da amplitude do sinal, as informações da correlação do ruído (incorporado na matriz α) pode ser extraído a partir da primeira das sete amostras digitalizadas e da própria matriz α :

$$\begin{pmatrix} 0 & \alpha_{1,2} & \dots & \alpha_{1,N} \\ \alpha_{2,1} & 0 & \dots & \alpha_{2,N} \\ \dots & \dots & \dots & \dots \\ \alpha_{N,1} & \alpha_{N,2} & \dots & 0 \end{pmatrix}$$

Os valores de $\alpha_{i,j}$ e β_i são obtidos através do processo de minimização. A reconstrução do sinal no canal i (s_i^{rec}) é realizado através da remoção do offset estimado durante o processo de minimização β_i e aplicando a matriz α para os valores medidos de todos os outros canais do módulo de acordo com,

$$s_i^{rec} = s_i - (\alpha_{i,1}x_1 + \alpha_{i,2}x_2 + \dots + \beta_i + \dots + \alpha_{i,N}x_N) \quad (6.4)$$

Se o método proporciona uma boa descrição das correlações no padrão de ruído, pode-se esperar que a distribuição de ruído se torne mais estreita depois de corrigir quaisquer efeitos indesejáveis, ou seja, mais próxima da distribuição do ruído intrínseco. A correlação do nível de ruído foi estudada utilizando 10.000 eventos a partir de um evento de pedestal de 2009.

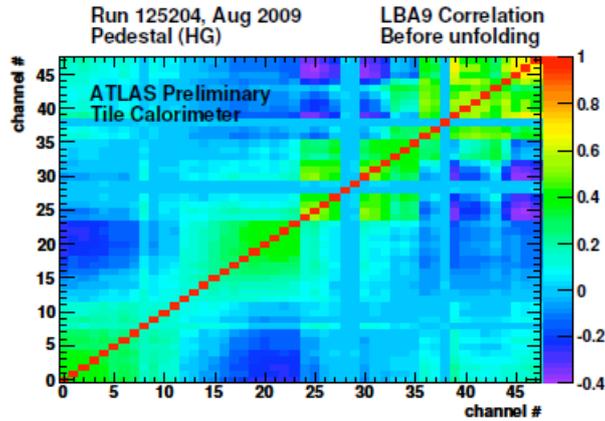


Figura 6.6: Correlação do ruído entre todos os canais do módulo LBA9 do TileCal antes da aplicação do método. Fonte: [18]

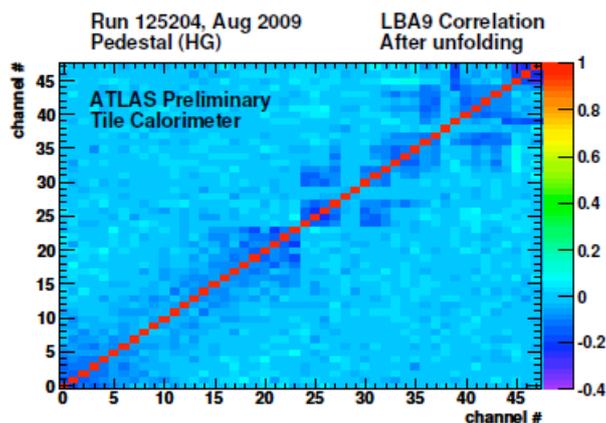


Figura 6.7: Correlação do ruído entre todos os canais do módulo LBA9 do TileCal após a aplicação do método. Fonte: [18]

Na Figura 6.6, as correlações entre os canais são claramente visíveis em sinais de *high-gain* para o módulo LBA 9 do TileCal. As regiões com altos e baixos valores de correlação estão refletindo a configuração de hardware do TileCal. A Figura 6.7 mostra que as correlações foram significativamente reduzidas após a aplicação do método de decorrelação de chi-quadrado.

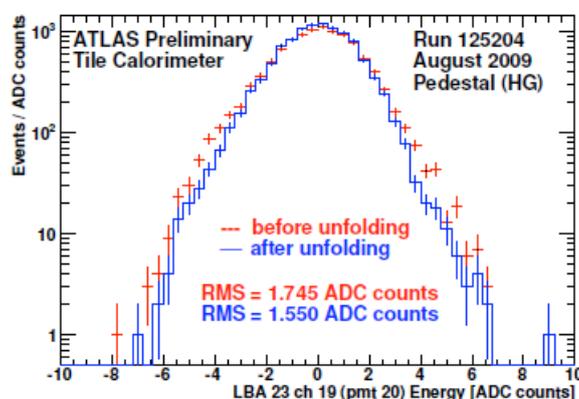


Figura 6.8: Distribuição de ruído do canal 19 do módulo LBA23 do TileCal antes e depois da aplicação do método. Fonte: [18]

Na Figura 6.8, o pedestal reconstituído do canal 19 do módulo LBA23 em ADC *counts* antes (pontos vermelhos) e depois (linha azul) de se aplicar a decorrelação de chi-quadrado. O canal 19 é um exemplo de um canal com pouco ruído correlacionado e pode-se observar que a distribuição permanece decorrelacionada após a aplicação do método. Já o canal 47, visto na Figura 6.9, é um exemplo de um canal com alta correlação de ruído com os demais canais. Podemos observar que a cauda é significativamente reduzida, após a aplicação do método.

O método proposto neste trabalho para remover o componente de ruído corre-

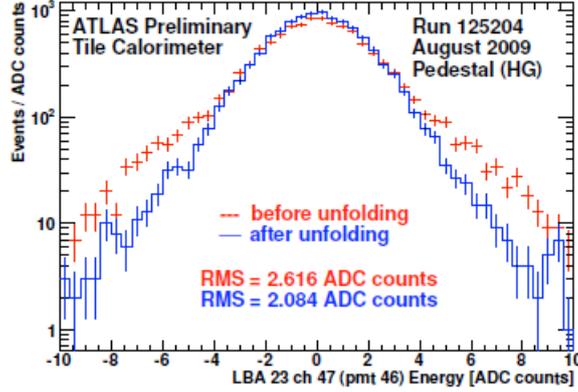


Figura 6.9: Distribuição de ruído do canal 47 do módulo LBA23 do TileCal antes e depois da aplicação do método. Fonte: [18]

lacionado do TileCal foi testado e apresentou bons resultados na descorrelação dos canais em módulos do TileCal em eventos de pedestal.

6.4 Correlação e Descorrelação de Ruído em Detectores Bolômétricos

Neste trabalho, realizado por C. Mancini-Terracciano e M. Vignatia [19], foi utilizado uma técnica de *denoising* em frequência para descorrelacionar o ruído em um detector de bolômetros utilizado para detectar quantidades muito pequenas de energia. Bolômetros são detectores mediados de *phonon* (quantum de energia vibracional) nos quais a energia das interações de partículas é convertida em energia térmica e medida através do aumento da temperatura. São usados em experimentos de física de partículas para procurar processos raros, como decaimentos beta duplo sem neutrinos ($0\nu\text{DBD}$) e interações da matéria escura. Eles apresentam uma excelente resolução de energia, alguns keV, ao longo de um intervalo de energia que se estende a partir de alguns keV até vários MeV. No entanto, a resolução pode ser limitada pelo ruído induzido pelas vibrações do aparelho mecânico. Nas matrizes de bolômetros parte deste ruído é correlacionado entre bolômetros diferentes e pode ser removida utilizando um algoritmo de descorrelação de multicanal.

O experimento CUORE (*Cryogenic Underground Observatory for Rare Events*) vai procurar o decaimento beta duplo sem neutrinos de Te^{130} (um isótopo relativamente abundante do elemento telúrio), utilizando uma matriz de 988 bolômetros de 750g cada [19]. Operando a uma temperatura de cerca de 10 mK, estes detectores apresentam uma ótima resolução de energia. O decaimento beta duplo sem neutrino ($0\nu\beta\beta$) é um processo teórico que pode, se comprovado, ocorrer muito raramente em

alguns elementos onde o neutrino (ν) emitido pelo decaimento beta de um nêutron ($n \rightarrow p + e^- + \nu$) é absorvido no decaimento beta simultâneo de outro nêutron ($\nu + n \rightarrow p + e^-$) (ver Figura 6.10). Isto pode ocorrer apenas se o neutrino for a sua própria anti-partícula e, portanto, uma partícula denominada "Majorana", ao contrário de todos os outros constituintes conhecidos da matéria (por exemplo, os elétrons e quarks), que são conhecidos como partículas de Dirac. O decaimento $0\nu\beta\beta$ é o único caminho viável para descobrir se o neutrino é uma partícula de Majorana. Se observado, esta descoberta terá enormes consequências para a física nuclear, para a física de partículas, astrofísica e para a cosmologia. Uma observação dos decaimentos $0\nu\beta\beta$ também tem o potencial para determinar a massa absoluta de neutrinos, que até então é desconhecida. O decaimento $0\nu\beta\beta$, portanto, oferece uma oportunidade única para sondar a natureza fundamental dos neutrinos.

Um detector utilizando 62 bolômetros, conhecido como CUORICINO, foi testado no *Laboratori Nazionali del Gran Sasso* (LNGS) na Itália entre 2003 e 2008, e comprovou a viabilidade técnica bolométrica do cristal de TeO_2 [19]. O experimento CUORE está em construção.

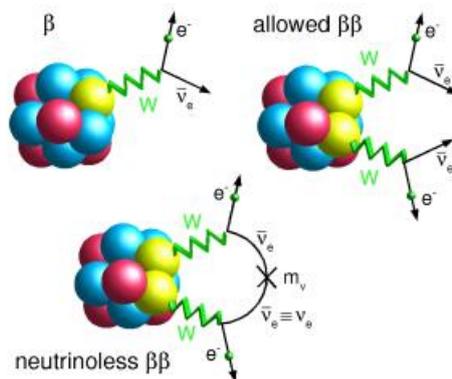


Figura 6.10: Decaimento $0\nu\beta\beta$. Fonte: [19]

As vibrações do aparelho criogênico (que mantém o sistema a baixas temperaturas) introduzem um ruído que, limita a resolução de energia a baixas energias e o limiar de energia. Uma vez que todos bolômetros estão contidos na mesma estrutura, parte do ruído de vibração deverá ser correlacionado. Foi apresentado então um método para estimar o ruído correlacionado entre os diferentes bolômetros, e um método para removê-lo. A aplicação aos dados do experimento CUORICINO mostraram que o ruído correlacionado é visível e que pode ser eficientemente removido. A estrutura do detector CUORE será diferente da estrutura do experimento CUORICINO, e, portanto, um ruído de vibração diferente é esperado. Os bolômetros CUORICINO e CUORE são compostos de duas partes principais, um cristal de TeO_2 e um termistor transmutado de nêutrons dopado com germânio (Ge-NTD).

O cristal possui a forma de cubo ($5 \times 5 \times 5 \text{ cm}^3$) e o termistor está acoplado ao cristal, atuando como um termômetro. Quando a energia da partícula é liberada no cristal, sua temperatura aumenta e altera a resistência do termistor. Para ler o sinal, o termistor é polarizado com uma corrente constante, a qual é fornecida por um gerador de tensão e uma resistência de carga em série com o termistor. A resistência do termistor varia no tempo, com a temperatura, e com a tensão entre o sinal do bolômetro. O valor da resistência de carga é escolhida para ser muito maior do que a do termistor, de maneira que a tensão através do termistor é proporcional à sua resistência. Os cabos que ligam o termistor para a eletrônica introduz uma capacitância parasita não negligenciável.

A resposta típica dos bolômetros para partículas que incidem sobre o cristal é da ordem de 100 mV/MeV . A largura de banda de frequência de sinal varia de $0\text{-}10 \text{ Hz}$, enquanto que os componentes de ruído estendem-se a frequências mais elevadas.

O sinal é amplificado, filtrado com um filtro ativo de Bessel com 6 pólos com uma frequência de corte de 12 Hz e, em seguida, adquirido com um ADC de 18 bits com amostragem frequência de 125 Hz . O ganho dos amplificadores varia entre $500\text{-}10.000 \text{ V/V}$ e é ajustado para cada bolômetro para regular os sinais ao alcance do ADC, que é de $\pm 10.5 \text{ V}$ [19].

Os 62 bolômetros do experimento CUORICINO foram dispostos em uma torre de 13 andares: 11 desses andares foram compostos por quatro bolômetros de $5 \times 5 \times 5 \text{ cm}^3$ cada, e dois andares foram compostos por nove pequenos bolômetros de $3 \times 3 \times 6 \text{ cm}^3$ cada. O experimento CUORE usará apenas cristais de $5 \times 5 \times 5 \text{ cm}^3$ ordenados em 19 torres de 13 andares cada (ver Figura 6.11). A eletrônica de *front-end*, que fornece a alimentação, as resistências de carga e o amplificador, é colocada fora do criostato, em temperatura ambiente.

Vibrações da estrutura do detector geram dois tipos de ruído: térmico e microfônico. O ruído térmico é devido às vibrações dos cristais que induzem flutuações de temperatura dos próprios cristais. O ruído no microfone é devido às vibrações dos fios que conectam o termistor à tomada do criostato. O espectro de potência do ruído típico é mostrado na Figura 6.12, os picos na figura são atribuídos a microfonia, enquanto a parte contínua é atribuída às vibrações do cristal.

O sistema aquisição de dados ligados às placas ADC implementa dois tipos de *trigger* de software. O primeiro *trigger* é acionado quando um sinal é detectado em um bolômetro, o segundo dispara aleatoriamente para adquirir ondas sonoras. Para estimar a correlação de ruído, a aquisição foi programada para adquirir simultaneamente dados de todos os bolômetros do experimento quando o *trigger* de sinal ou o ruído é disparado em um bolômetro.

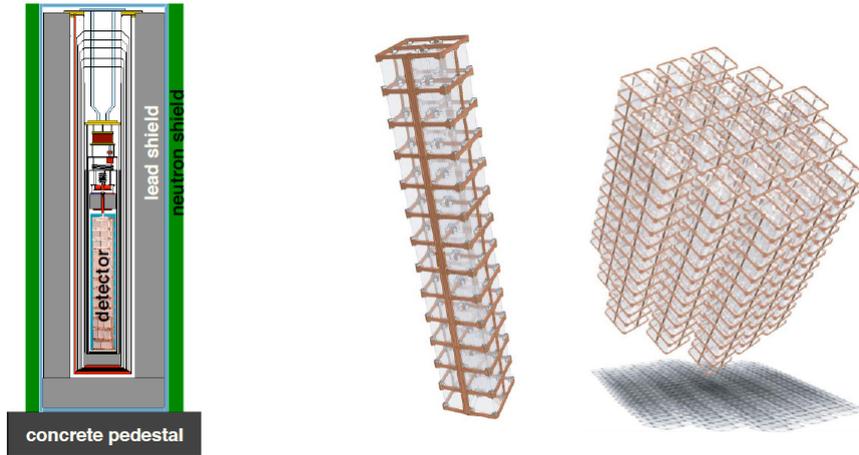


Figura 6.11: Torre do experimento CUORICINO dentro do sistema de resfriamento (esquerda). Uma Torre do experimento CUORE (centro). Uma matriz de torres experimento CUORE (direita). Fonte: [19]

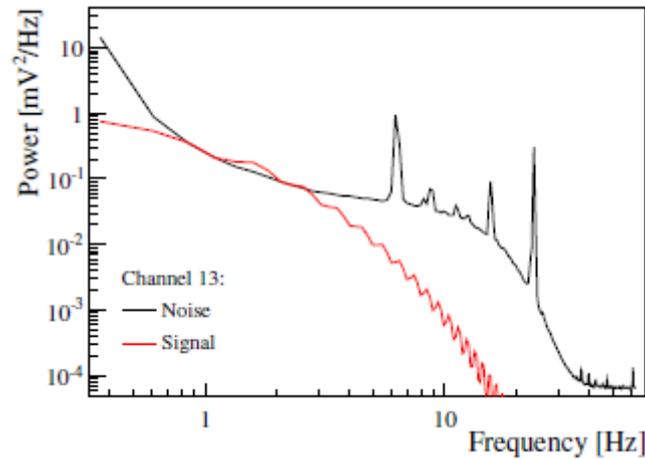


Figura 6.12: Espectro de potência do sinal e do ruído. Fonte: [19]

6.4.1 Correlação do Ruído

O espectro de potência de ruído é estimado a partir de um grande número de formas de ondas adquiridas com o *trigger* aleatório, removendo aquelas que, por acaso, contêm sinais. O espectro de potência $N_i(\omega)$ de cada canal do bolômetro i é calculado como:

$$N_i(\omega) = E [n_i(\omega) \times n_i^*(\omega)], \quad (6.5)$$

onde $n_i(\omega)$ é a transformada discreta de Fourier (DFT) da forma de onda do ruído. A covariância entre o ruído dos canais dos bolômetros i e j é estimada da seguinte forma:

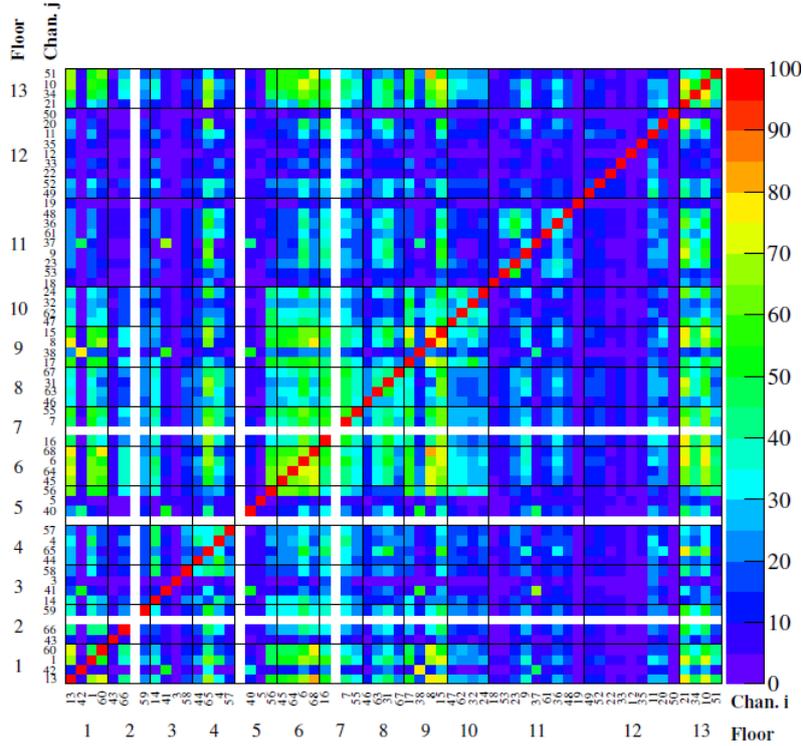


Figura 6.13: Correlação do ruído nos canais do experimento CUORICINO. Fonte: [19]

$$C_{i,j}(\omega) = E [n_i(\omega) \times n_j^*(\omega)] . \quad (6.6)$$

6.4.2 Descorrelação do Ruído

Nos bolômetros de TeO_2 o ruído é puramente aditivo ao sinal, significando que a forma de onda $f(t)$ observada na presença do sinal de amplitude A pode ser expressa como:

$$f(t) = A \times s(t) + n(t), \quad (6.7)$$

onde $s(t)$ é a forma de onda do sinal. Uma vez que o ruído em diferentes bolômetros é parcialmente correlacionado, é possível remover parte de $n(t)$ de $f(t)$ usando a forma de onda de outros bolômetros, nos quais somente ruído está presente. Assumindo que cada componente de frequência de $n_i(\omega)$ é normalmente distribuído, a distribuição de probabilidade multidimensional dos ruídos em todos os bolômetros, ou seja, $n^{\rightarrow} = \{n_1(\omega), n_2(\omega), \dots, \}$ pode ser escrita como:

$$P[\vec{n}(\omega)] \propto \exp\left(-\frac{1}{2} \vec{n}^\dagger(\omega) C^{-1}(\omega) \vec{n}(\omega)\right), \quad (6.8)$$

onde C é a matriz de covariância. A distribuição de probabilidade do ruído no bolômetro i , $n_i(\omega)$, pode ser obtida integrando o ruído de outros bolômetros:

$$P[n_i(\omega)] \propto \exp \left[-\frac{C_{ii}^{-1}(\omega)}{2} \left(n_i(\omega) + \sum_{j \neq i} \frac{C_{ij}^{-1}(\omega)}{C_{ii}^{-1}(\omega)} n_j(\omega) \right)^2 \right]. \quad (6.9)$$

A equação acima representa uma Gaussiana com média igual a $\left[-\sum_{j \neq i} \frac{C_{ij}^{-1}(\omega)}{C_{ii}^{-1}(\omega)} n_j(\omega) \right]$ e variância $1/C_{ii}^{-1}$. O valor descorrelacionado de $n_i(\omega)$ pode ser obtido para cada forma de onda como:

$$n_i^d(\omega) = n_i(\omega) + \sum_{j \neq i} \frac{C_{ij}^{-1}(\omega)}{C_{ii}^{-1}(\omega)} n_j(\omega). \quad (6.10)$$

E seu espectro de potência é esperado ser de:

$$N_i^d(\omega) = \frac{1}{C_{ii}^{-1}(\omega)}. \quad (6.11)$$

Uma forma de onda genérica $f_i(t)$ no bolômetro i , contendo somente ruído ou ruído mais sinal, pode ser descorrelacionada usando o ruído dos outros bolômetros como:

$$f_i^d(\omega) = f_i(\omega) + \sum_{j \neq i} \frac{C_{ij}^{-1}(\omega)}{C_{ii}^{-1}(\omega)} n_j(\omega) = As_i(\omega) + n_i^d(\omega). \quad (6.12)$$

Uma vez que a matriz de covariância é estimada a partir dos dados, as formas de onda de cada bolômetro são descorrelacionadas utilizando a Equação 6.12 e o efeito no espectro de potência pode ser previsto pela Equação 6.11.

Os dados coletados pelo experimento CUORICINO, foram divididos em dois grupos. O primeiro grupo de dados foi utilizado para se estimar a matriz de correlação dos canais do detector e o espectro de potência esperado do ruído, o segundo grupo de dados foi utilizado para verificar se os resultados da aplicação do algoritmo de descorrelação da forma de onda foi considerado consistente com as expectativas. Na Figura 6.14 pode-se observar que o espectro de potência do ruído diminuiu em baixas frequências e os picos microfônicos de ruído foram eliminados. Porém, ao utilizar todos os demais canais para descorrelacionar cada canal tornou o custo computacional do método elevado uma vez que deve-se calcular a DFT de cada canal utilizado. Como o experimento CUORE tem 988 bolômetros ao invés de 62 do experimento CUORICINO, os autores testaram a descorelação de um canal utilizando um número reduzido de bolômetros. Para isso eles selecionaram 11 canais mais correlacionados ao canal 13 para realizar a descorrelação e calcularam o espectro de potência esperado do ruído do canal 13 descorrelacionado (linha vermelha). O resultado obtido foi equivalente aqueles obtidos utilizando todos os demais canais na descorrelação.

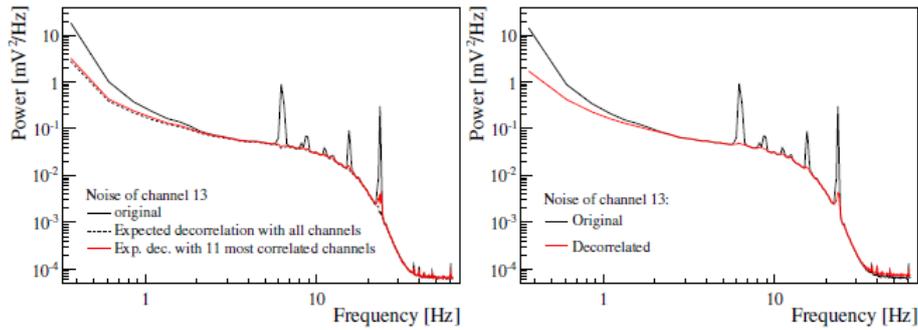


Figura 6.14: Espectro de potência original do ruído do canal 13 (linha sólida), o espectro de potência do canal 13 já decorrelacionado utilizando todos os demais canais (linha tracejada) e o espectro de potência do canal 13 decorrelacionado utilizando apenas 11 canais (linha vermelha). Fonte: [19]

O espectro de ruído obtido na forma de onda decorrelacionada com a Equação 6.12 é mostrado na Figura 6.14 (à direita), a variância do ruído foi reduzida de 2.0 mV para 1.2 mV (equivalente a 3.6 keV e 2.2 keV). A Figura 6.15 mostra a resolução de energia de todos os canais do detector antes e depois da decorrelação, em alguns canais a redução superou 50%.

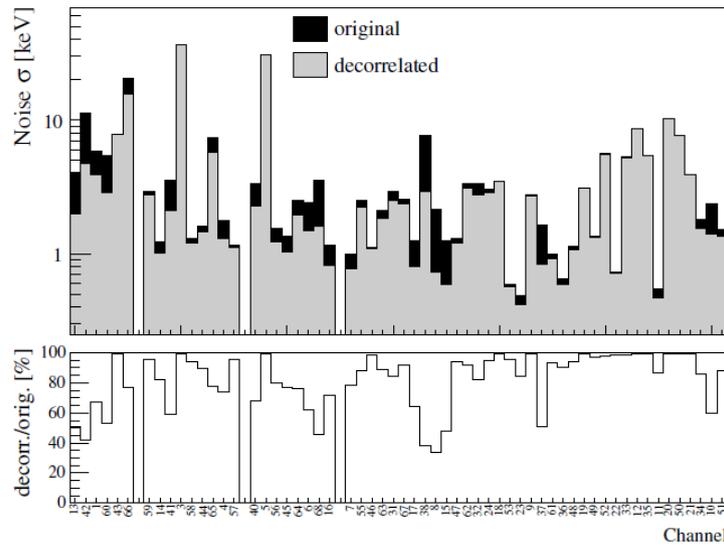


Figura 6.15: Resolução do ruído dos canais do experimento CUORICINO. Fonte: [19]

Portanto, o artigo demonstrou que o ruído correlacionado entre os 62 bolômetros do experimento CUORICINO, causado por vibrações da estrutura do detector, pode ser removido de forma eficiente através de uma técnica de processamento de sinal digital usada para a decorrelação de ruído.

Capítulo 7

Método Proposto de Estimação de Energia para um Calorímetro Finamente Segmentado a Nível Celular

Neste capítulo será apresentado uma breve descrição da maneira como é feita a estimação de amplitude de energia nos canais do TileCal. Apresentamos também o método desenvolvido neste trabalho para melhorar a estimação de energia utilizando as células do TileCal.

7.1 Estimação de Energia no Tilecal

Todo sinal produzido pela placa 3-in-1 do TileCal possui a mesma forma de pulso $h(t)$ (Figura 7.1), diferindo apenas em sua amplitude (A), além disso, para garantir que a conversão analógico-digital do sinal seja sempre positiva, o sinal adquirido recebe a adição de uma linha de base conhecida como pedestal (ped). Os sinais são digitalizados em 7 amostras em intervalos de tempo regulares (25ns no TileCal). Se incluirmos também o ruído eletrônico Gaussiano característico de cada canal ($w[n]$), podemos escrever o sinal de cada canal de leitura da seguinte forma:

$$S[n] = Ah[n] + w[n] + ped.$$

Sendo assim, o objetivo dos algoritmos de reconstrução de energia presentes no TileCal é estimar o valor de A dado $S[n]$. O algoritmo denominado *Optimal Filtering* (OF) é o algoritmo padrão utilizado para reconstruir a amplitude e a fase do sinal digitalizado de cada canal do TileCal. Outro algoritmo de reconstrução de energia que apresentou bons resultados utiliza o Filtro Casado (algoritmo identificado pela sigla MF) como método, mais detalhes deste algoritmo pode ser encontrado em [54].

7.1.1 Optimal Filter

Este algoritmo utiliza uma combinação linear das amostras digitais do sinal usando pesos calculados pelo método dos multiplicadores de Lagrange para reconstruir a amplitude e a fase. Os detalhes do método podem ser encontrados em [53]. O processo utilizado para calcular a fase e a amplitude com o algoritmo é dado pelas seguintes equações :

$$A = \sum_{i=1}^7 a_i S_i, \quad (7.1)$$

$$\tau = \frac{1}{A} \sum_{i=1}^7 b_i S_i, \quad (7.2)$$

onde S_i representa a i -ésima amostra digital do sinal. O pedestal, pode ser estimado como a média entre a primeira e a última amostra do sinal, ou calculado também através de uma combinação linear das amostras através da equação:

$$p = \sum_{i=1}^7 c_i S_i. \quad (7.3)$$

A amplitude A , é a distância entre o pedestal e o pico máximo do sinal reconstruído. A energia é proporcional a A , e uma constante de calibração medida para cada canal deve ser aplicada para se obter a energia.

A fase τ , é definida como o tempo entre o pico do sinal (Figura 7.1) e o tempo esperado do pulso. Este tempo de referência é calculado para cada canal em sistemas de calibração, levando em consideração o tempo de voo das partículas do ponto de interação e o comprimento de deslocamento do comprimento de onda de fibras ópticas. Os pesos, a_i e b_i são obtidos a partir da forma do pulso do sinal dos fotomultiplicadores e da correlação de ruído entre as amostras digitais. A maneira de calcular os pesos visa minimizar o efeito do ruído na reconstrução da amplitude e do tempo em cada canal. No algoritmo denominado OF2 a combinação linear das constantes são calculadas de tal maneira que a subtração do pedestal não é necessária. Já na versão do algoritmo conhecida como OF1, o pedestal é calculado pela média da primeira e última amostra do sinal, e então subtraído das amostras. A energia reconstruída usada no *Trigger* e nos softwares *offline* do ATLAS é encontrada através da seguinte equação:

$$E_{canal} = A \times C_{ADC \rightarrow pC} \times C_{pC \rightarrow GeV} \times C_{C_s} \times C_{Laser}.$$

Onde A é a amplitude representando a energia medida dada em *ADC counts*, $C_{ADC \rightarrow pC}$ é uma constante de conversão de unidade (energia para carga), definida através da calibração de CIS (*Charge Injection System*). A constante $C_{pC \rightarrow GeV}$ é

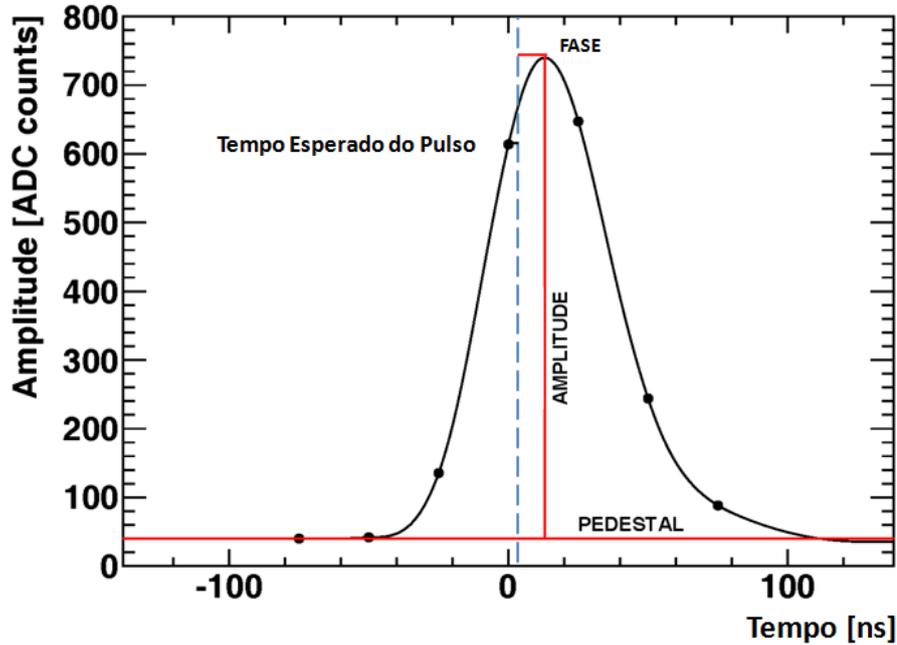


Figura 7.1: Sinal característico do TileCal (em destaque, amplitudes calculadas no OF).

outro fator de conversão (carga para GeV) definida através de calibrações de *testbeam* utilizando um feixe de elétrons com momento conhecido. Já a constante C_{C_s} é utilizado para corrigir não-uniformidades residuais após a equalização de ganhos em todos os canais, e é realizado pelo sistema de calibração de Césio; o termo C_{Laser} , não implementado atualmente, permitem corrigir não-linearidades futuras na resposta do PMT medida pelo sistema de calibração a laser.

7.2 Estimação de Energia nas Células do TileCal

Como visto no Capítulo 4, as bordas das telhas cintiladoras são "lidas" por conjuntos de fibras ópticas e por dois PMT diferentes, ou seja, sempre que uma partícula atravessa uma célula do TileCal e deposita parte de sua energia nesta célula nós temos duas observações do mesmo evento realizados por PMT independentes. O método proposto neste trabalho consiste em utilizar a redundância de leitura existente nas células do TileCal para realizar a estimação de energia ao nível da célula em contraposição ao método utilizado atualmente que consiste em apenas somar com pesos iguais os dois canais de leitura da célula. O design de dupla leitura em células de um calorímetro já foi utilizado em outros experimentos de física de altas energias, como nas células de leitura do detector de partículas ZEUS localizado no acelerador

de partículas DESY na Alemanha; esse modelo de leitura dupla é utilizado principalmente para permitir o cálculo exato da posição em que a partícula penetrou a célula e também para no caso de falha de um dos canais o canal restante substituir o canal defeituoso, evitando assim que a célula de leitura seja totalmente desativada. Para realizar a estimação de energia a nível de célula, propomos a utilização de um estimador para minimizarmos o ruído dos canais de leitura e otimizarmos a energia estimada da célula.

Na prática o procedimento proposto busca encontrar pesos diferentes para cada um dos dois canais da célula de leitura levando-se em conta a característica do ruído e a possível correlação existente em cada par de canais de uma célula; dessa maneira podemos minimizar a contribuição de um canal ruidoso em uma célula e até mesmo eliminar canais defeituosos que por ventura não tenham sido identificados e suprimidos pelo sistema calibração e monitoramento do TileCal. O procedimento proposto tem por objetivo complementar a estimação da amplitude de energia realizada pelos algoritmos de reconstrução de energia disponíveis, *Optimal Filter* e *Matched Filter*, que realizam a estimação de energia individualmente em cada canal da célula através das 7 amostras do sinal. Atualmente para se obter a energia da célula, a energia estimada dos dois canais da célula são somados:

$$E_{cell} = E_{pmt_1} + E_{pmt_2}.$$

Porém, se o ruído dos canais forem gaussianos e possuírem variâncias diferentes sendo ou não correlacionados, existe então uma abordagem ótima para maximizar a relação sinal-ruído da célula melhorando assim sua resolução de energia:

$$E_{cell_{opt}} = a_1 \times E_{pmt_1} + a_2 \times E_{pmt_2},$$

sendo,

$$a_1 + a_2 = 2, \tag{7.4}$$

para manter o patamar de energia da célula.

Podemos encontrar o peso para cada canal utilizando a informação do pedestal estimado por um dos algoritmos de estimação de amplitude de cada canal ou através de algum *Run* de pedestal realizado periodicamente no TileCal. O cálculo para se encontrar os pesos tem de ser feito de maneira *offline* uma vez que é preciso de estatística para realizar o cálculo da matriz de covariância (no caso do estimador MVU) e da variância (no caso do estimador MLE), mas a ponderação dos canais pode ser realizada de maneira *online* contanto que haja disponibilidade nas placas de *Read-Out Drivers* (ROD) da eletrônica de *back-end* do TileCal.

O estimador MVU leva em consideração uma possível correlação entre o ruído dos canais, e como visto no Capítulo 5, pode ser generalizado para abranger o caso

onde o ruído aditivo é colorido. Com o estimador MVU podemos encontrar a energia ótima da célula utilizando a Equação 5.27 para o caso de $N = 2$ da seguinte maneira:

$$\widehat{E}_{cell} = 2 \frac{1^T C^{-1}}{1^T C^{-1} 1} x, \quad (7.5)$$

onde a constante dois multiplicando a equação possui a função de garantir o nível de energia na célula total na célula, x é o vetor com a energia estimada dos dois canais e C é a matriz de covariância do ruído dos canais. Sendo $\sigma_{w_j}^2$ a variância do ruído do canal j e $C_{w_1 w_2}$ a correlação entre o ruído dos canais, os pesos da soma ponderada são obtidos da seguinte forma:

$$a_{1,2} = \frac{\sigma_{w_{2,1}}^2 - C_{w_1 w_2}}{\sigma_{w_1}^2 + \sigma_{w_2}^2 - 2C_{w_1 w_2}}. \quad (7.6)$$

A variância do estimador MVU será dada por:

$$\sigma_{MVU}^2 = \frac{1}{1^T C^{-1} 1} = \frac{\sigma_{w_1}^2 \sigma_{w_2}^2 - C_{w_1 w_2}}{\sigma_{w_1}^2 + \sigma_{w_2}^2 - 2C_{w_1 w_2}} \quad (7.7)$$

O estimador MLE assume que o ruído dos canais de leitura da célula são descorrelacionados, aditivo, branco e gaussiano. O estimador MLE tem por objetivo encontrar e atribuir um peso maior a medida que possuir a menor variância de ruído e um peso menor a medida que possuir uma variância maior [55]. Como visto na Seção 5.5, no estimador MLE os pesos para cada canal serão obtidos da seguinte maneira:

$$a_j = \frac{1}{\sigma_{w_j}^2} \left(\sum_{i=1}^2 \frac{1}{\sigma_{w_i}^2} \right)^{-1} \quad j = 1, 2. \quad (7.8)$$

$$a_{1,2} = \frac{\sigma_{w_{2,1}}^2}{\sigma_{w_1}^2 + \sigma_{w_2}^2}$$

A variância do estimador MLE será dada por:

$$\sigma_{MLE}^2 = a_1^2 \times \sigma_{w_1}^2 + a_2^2 \times \sigma_{w_2}^2 \quad (7.9)$$

$$\sigma_{MLE}^2 = \left(\frac{\sigma_{w_2}^2}{\sigma_{w_1}^2 + \sigma_{w_2}^2} \right)^2 \times \sigma_{w_1}^2 + \left(\frac{\sigma_{w_1}^2}{\sigma_{w_1}^2 + \sigma_{w_2}^2} \right)^2 \times \sigma_{w_2}^2 = \frac{\sigma_{w_2}^4 \sigma_{w_1}^2 + \sigma_{w_1}^4 \sigma_{w_2}^2}{(\sigma_{w_1}^2 + \sigma_{w_2}^2)^2}$$

$$\sigma_{MLE}^2 = \frac{\sigma_{w_1}^2 \sigma_{w_2}^2 (\sigma_{w_1}^2 + \sigma_{w_2}^2)}{(\sigma_{w_1}^2 + \sigma_{w_2}^2)^2} = \frac{\sigma_{w_1}^2 \sigma_{w_2}^2}{(\sigma_{w_1}^2 + \sigma_{w_2}^2)} \quad (7.10)$$

Capítulo 8

Resultados

Neste capítulo, serão apresentados os resultados obtidos com a implementação da estimação de energia em nível das células no TileCal.

8.1 Resultados com Dados de Pedestal

Para testar a capacidade de redução de ruído e a viabilidade de implementação do método proposto, inicialmente aplicamos a estimação em banco de dados de *runs* de pedestal. Periodicamente são realizados tomadas de dados com a intenção de caracterizar o ruído eletrônico do TileCal. Essa tomada de dados ocorre com o detector em pleno funcionamento, porém, na ausência de colisões para garantir que os dados sejam puramente ruído, esse conjunto de evento é conhecido como *run* de pedestal. Nesta etapa do trabalho utilizamos um *Run* de pedestal obtido em Outubro de 2011 com aproximadamente 20.000 eventos, sendo estes eventos o número de observações realizadas para cada canal do TileCal durante a duração do *Run*.

Ao utilizar *runs* de pedestal onde temos apenas a componente do ruído do calorímetro, realizamos uma separação dos dados em dois conjuntos, um de teste correspondendo a 60% dos eventos e um de treino correspondendo a 40% dos eventos. O conjunto de treino foi utilizado para calcularmos a matriz de covariância do ruído dos canais das células e posteriormente calcularmos os pesos adequados para cada canal utilizando os estimadores MVU e MLE. Uma vez obtido os pesos dos canais, aplicamos a soma ponderada aos canais da célula e medimos o RMS da distribuição do ruído para compararmos com o RMS do ruído quando utilizamos apenas a soma simples (com pesos iguais a um).

A seguir, apresentamos exemplos de células com diferentes tipos de relações entre o ruído de seus canais, e o resultado obtido quando aplicamos a soma ponderada utilizando os estimadores comparado ao uso da soma simples entre os canais da célula. A Figura 8.1 mostra um plot com a correlação entre o ruído dos canais da

célula A9 do módulo LBA59 do TileCal. Neste caso o ruído dos canais da célula estão correlacionados, e principalmente um dos canais apresentou alta variância de ruído como podemos observar na Figura 8.1 através do valor de sua variância na matriz de covariância.

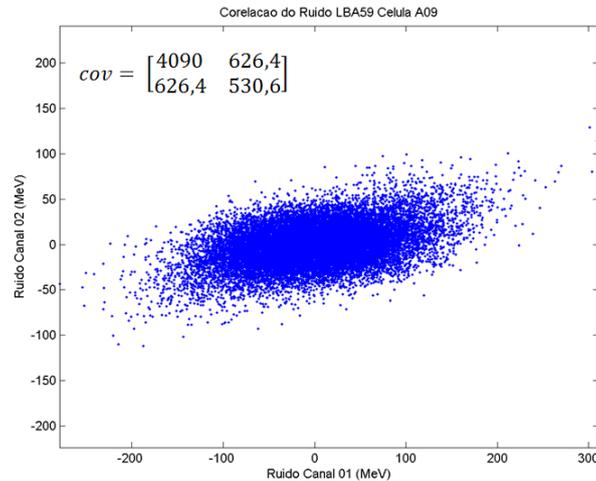


Figura 8.1: Correlação do ruído dos canais da célula da célula A9 do módulo LBA59.

A Figura 8.2 mostra três distribuições de ruído da mesma célula A9 do módulo 59 da partição LBA do TileCal. A distribuição em azul mostra o ruído da célula onde os canais foram somados com peso igual a um. As distribuições em vermelho e magenta apresentam o ruído da célula após a aplicação da soma ponderada pelos pesos calculados pelos estimadores MVU e MLE respectivamente.

Neste caso, devido a alta variância do ruído do canal 1 da célula, o peso calculado para este canal por ambos os estimadores (Tabela 8.1) diminuiu muito a contribuição deste canal na célula, enquanto isso o canal 2 da célula teve sua contribuição praticamente dobrada. Ou seja, o estimador realizou o mascaramento do canal defeituoso que apresentou alta variância de ruído e com isso diminuiu a variância do ruído na célula como visto na Figura 8.2.

Estimador	Peso Canal 1	Peso Canal 2
MVU	0,0568	1,9432
MLE	0,2296	1,7704

Tabela 8.1: Pesos estimados para os canais da célula A9 do módulo LBA59.

No caso da Figura 8.3 apresentamos um exemplo onde o ruído dos canais da célula são pouco correlacionados, mas ainda assim a variância do ruído de um dos

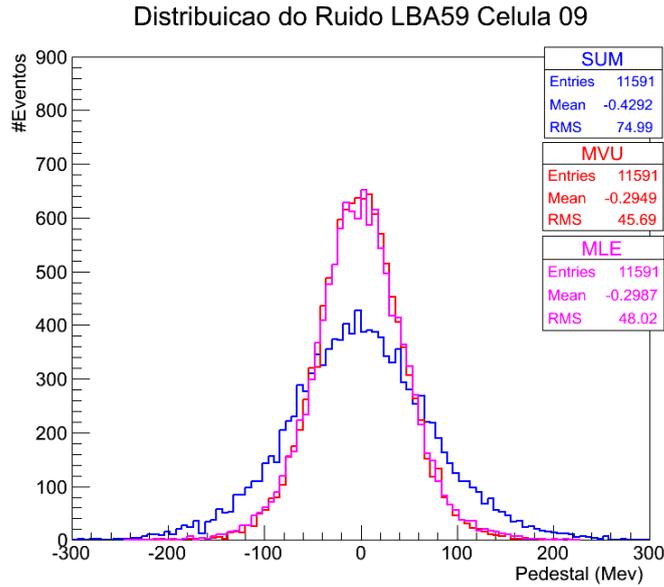


Figura 8.2: Distribuição de ruído antes e depois da estimação de energia em célula da célula A9 do módulo LBA59.

canais difere muito de seu par na célula. Assim, nesse caso também houve uma significativa redução do rms da distribuição do ruído da célula ao utilizarmos a estimação de energia proposta utilizando os pesos calculados pelos estimadores MVU e MLE, como podemos ver na Figura8.4.

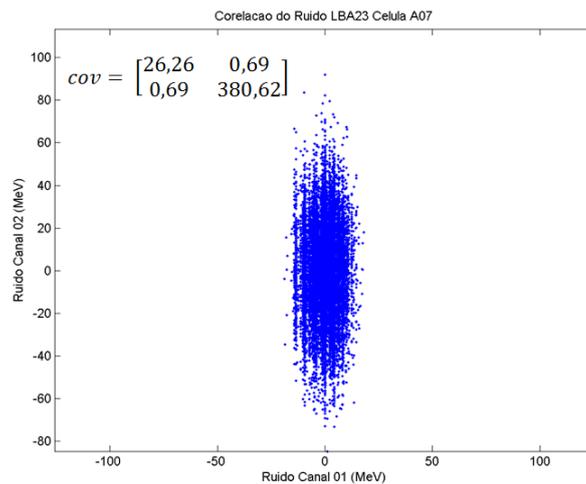


Figura 8.3: Corelação do ruído dos canais da célula da célula A7 do módulo LBA23.

Já na Figura 8.5, apresentamos um caso onde o ruído dos canais da célula apre-

Estimador	Peso Canal 1	Peso Canal 2
MVU	1,868	0,132
MLE	1,87	0,129

Tabela 8.2: Pesos estimados para os canais da célula A7 do módulo LBA23.

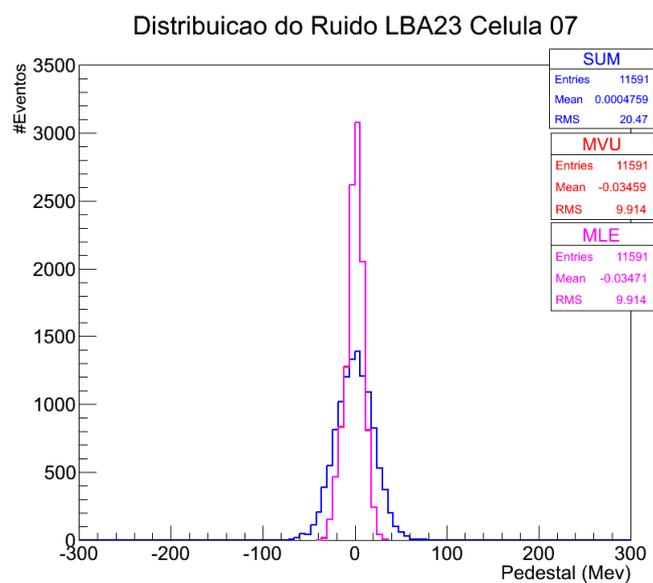


Figura 8.4: Distribuição de ruído antes e depois da estimação de energia em célula da célula A7 do módulo LBA23.

sentam uma correlação forte, porém, apresentam ruído com variâncias semelhantes. Portanto, os pesos calculados para os canais foram próximos a um (peso padrão), ainda assim observamos que o RMS da distribuição de ruído da célula apresentou uma melhora (Figura 8.6).

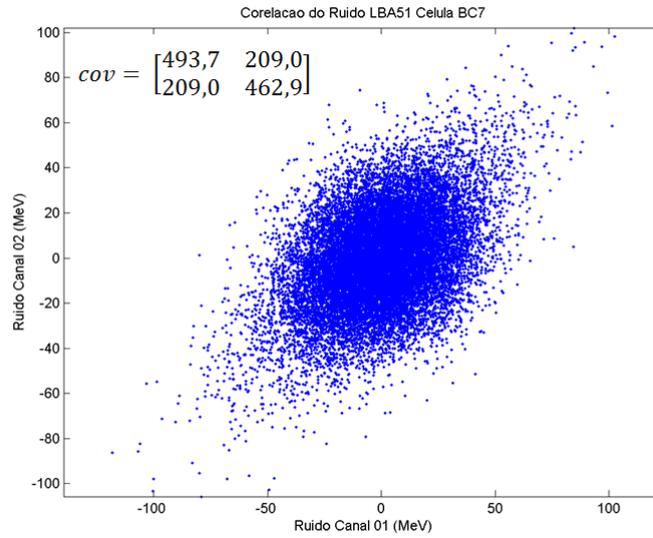


Figura 8.5: Corelação do ruído dos canais da célula da célula BC7 do módulo LBA51.

Estimador	Peso Canal 1	Peso Canal 2
MVU	0,9427	1,0573
MLE	0,9678	1,0322

Tabela 8.3: Pesos estimados para os canais da célula BC7 do módulo LBA51.

Na Figura 8.7 podemos observar uma comparação entre o rms de todas as células de todos os módulos do TileCal, antes (eixo vertical) e depois (eixo horizontal) de utilizarmos a estimação baseada em célula, uma reta vermelha com coeficiente angular de 45 graus foi utilizada para melhor comparação. Podemos concluir pela Figura 8.7 que o método proposto, em geral, sempre melhora ou iguala o RMS do ruído nas células se comparado ao método atual; podemos observar também que em alguns casos a diminuição do ruído foi significativa, estes casos em geral correspondem a células onde um dos canais apresentavam alta variância de ruído se comparado a seu par na célula, esses casos representam canais defeituosos que não foram identificados por nenhum sistema antes da reconstrução dos dados.

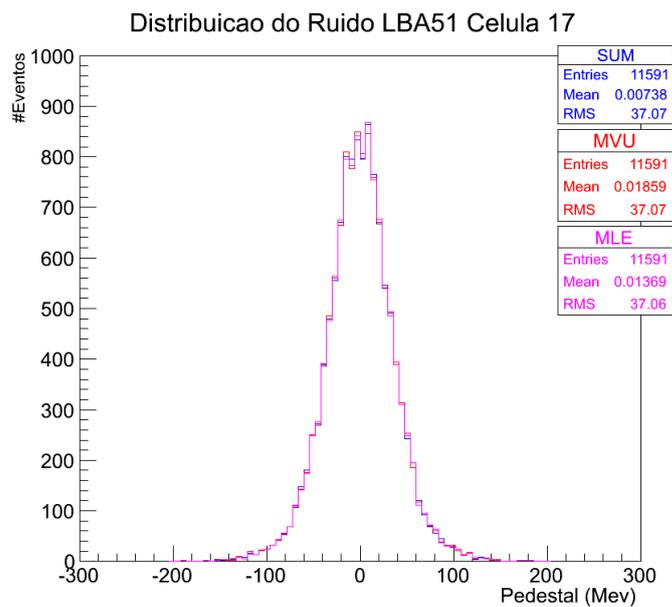


Figura 8.6: Distribuição de ruído antes e depois da estimação de energia em célula da célula BC7 do módulo LBA51.

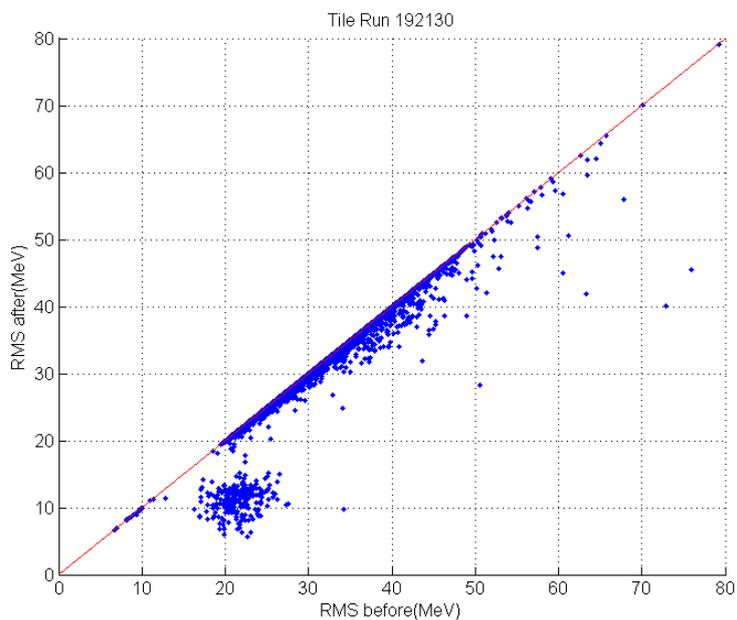


Figura 8.7: RMS da distribuição do ruído das células do TileCal com e sem estimação em célula.

A Figura 8.8 apresenta dois histogramas com a distribuição dos pesos encon-

trados para canal das célula de cada módulo do TileCal. Podemos observar que os histogramas são complementares uma vez que a soma dos pesos deve atender a relação definida na Equação 7.4. Em destaque mostramos que em certos casos o valor do peso do canal ficou próximo a zero. Ao realizarmos essa abordagem utilizando dados puramente de ruído, constatamos que a utilização do estimador de mínima variância para calcular pesos variáveis para os canais apresentou potencial na redução da variância do ruído nas células. Após essa etapa, passamos a realizar a estimação de energia utilizando dados reais de colisões realizadas no CERN com sinais representando a energia de partículas incidentes no TileCal.

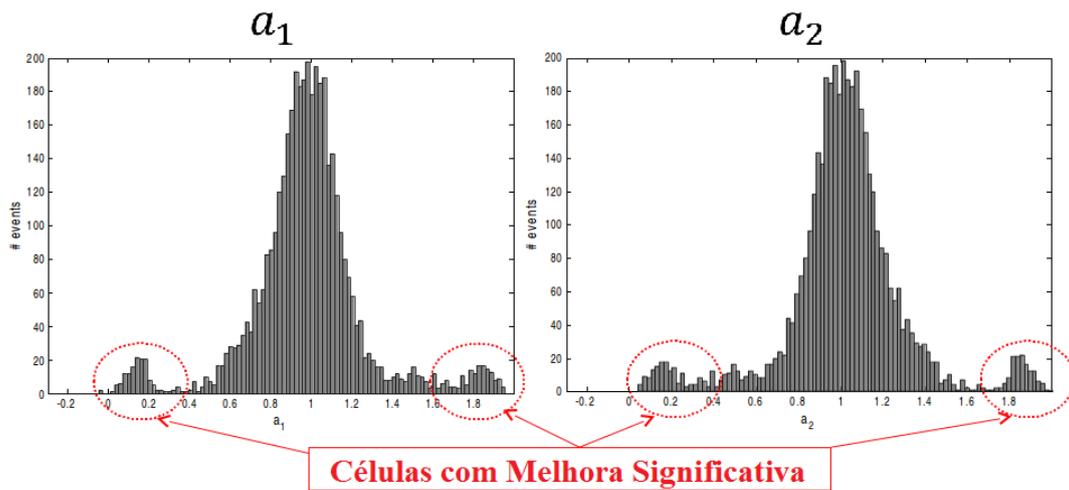


Figura 8.8: Distribuição dos pesos encontrados pelo estimador.

8.2 Resultados com Dados de Colisão

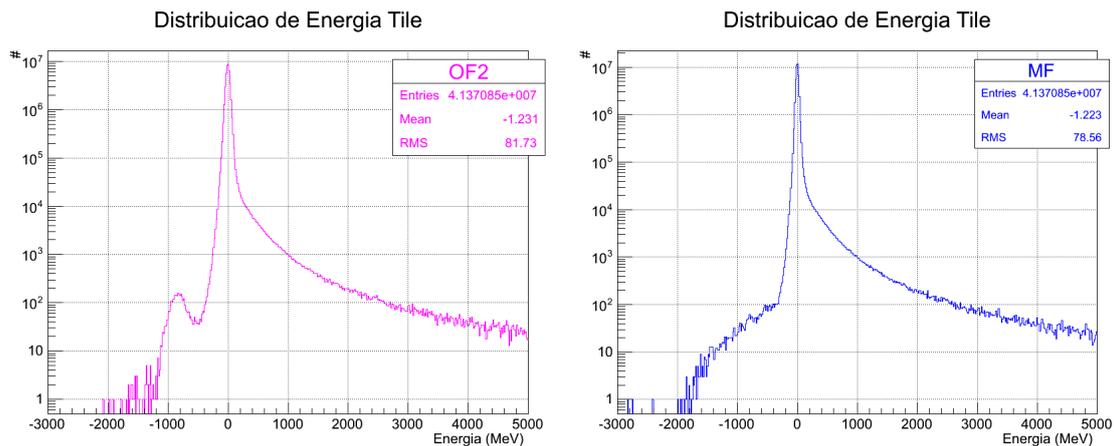
Nesta etapa da pesquisa, passamos a aplicar a estimação efetivamente em dados produzidos em colisões de partículas contendo energia de partículas depositada nas células do calorímetro hadrônico do ATLAS. Como visto no capítulo 7, a estimação em célula ocorre utilizando a amplitude do sinal estimada por um dos algoritmos de estimação de amplitude existentes no TileCal. Para calcular os pesos de cada canal foi utilizado o valor de pedestal estimado para cada canal em cada evento. Para esta etapa do trabalho foram utilizado dois *Runs* de colisão *pp* (próton-próton) listados na Tabela 8.4 juntamente com algumas de suas características.

Run	Data	Eventos Reconstruídos	Energia Máxima (GeV)	\mathcal{L} ($10^{32} cm^{-2} s^{-1}$)
167776	Out/2010	9.302	3500	1,8
216399	Dez/2012	290.634	3999	1,7

Tabela 8.4: Informação dos *Runs* de colisão utilizados. Fonte:[20].

8.2.1 Estimação de Energia em Células com Diferentes Algoritmos de Estimação de Amplitude dos Canais

A estimação em célula proposta neste trabalho deve ser utilizada em conjunto com um dos algoritmos de estimação de amplitude do pulso do TileCal. Uma vez que a estimação em célula utiliza o valor estimado da amplitude do pulso canal para realizar a estimação, foi importante determinar qual seria o melhor algoritmo de reconstrução de amplitude para utilizar juntamente com o método proposto.



(a) Energia reconstruída com o algoritmo *Optimal Filter 2*. (b) Energia reconstruída com o algoritmo *Matched Filter*

Figura 8.9: Distribuição da energia reconstruída das células do TileCal utilizando soma simples (SS).

A Figura 8.9 mostra a distribuição de energia das células do TileCal do *Run* 167776, utilizando como estimador de amplitude o algoritmo OF2 (Figura 8.9a) e utilizando o algoritmo de Filtro Casado-MF (Figura 8.9b). Podemos observar uma cauda negativa em ambas as distribuições de energia, evidentemente que essa cauda não representa uma deposição negativa de energia na célula do calorímetro. A cauda negativa observada em ambas as distribuições deve-se em parte ao ruído eletrônico dos canais assim como a presença de sinais fora de fase. Outra razão para o aparecimento dessa cauda negativa esta associado ao algoritmo OF2 de estimação de amplitude do canal, limitações no cálculo dos pesos das amostras realizados pelo

algoritmo permite que alguns dos pesos sejam negativos. Uma vez que haja sinais fora de fase onde o pico se encontra em uma destas amostras com peso negativo, provoca o aparecimento de uma estimaco de amplitude negativa. Na estimaco de amplitude realizada pelo algoritmo MF tambm pode surgir uma cauda negativa, como podemos ver na Figura 8.9b. Porm, nesse caso a principal razo sera a presena de canais com alto nvel de rudo, isto porque neste algoritmo todos os pesos calculados so positivos, portanto mesmo na presena de sinais fora de fase a estimaco da amplitude no sera negativa.

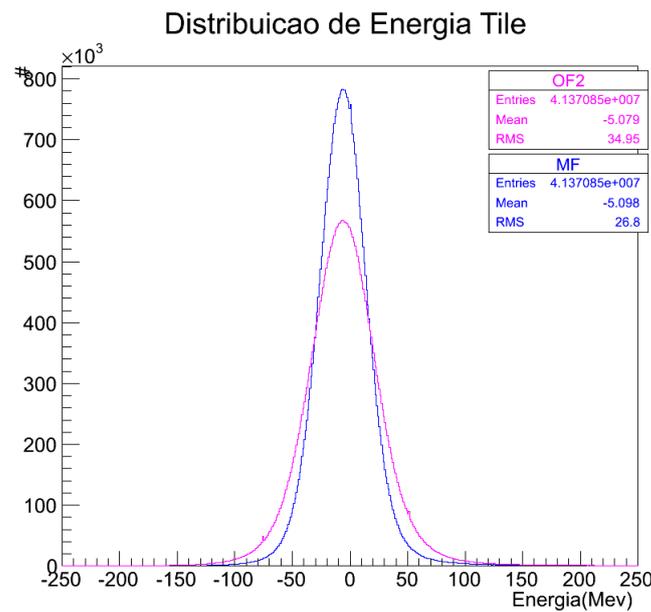
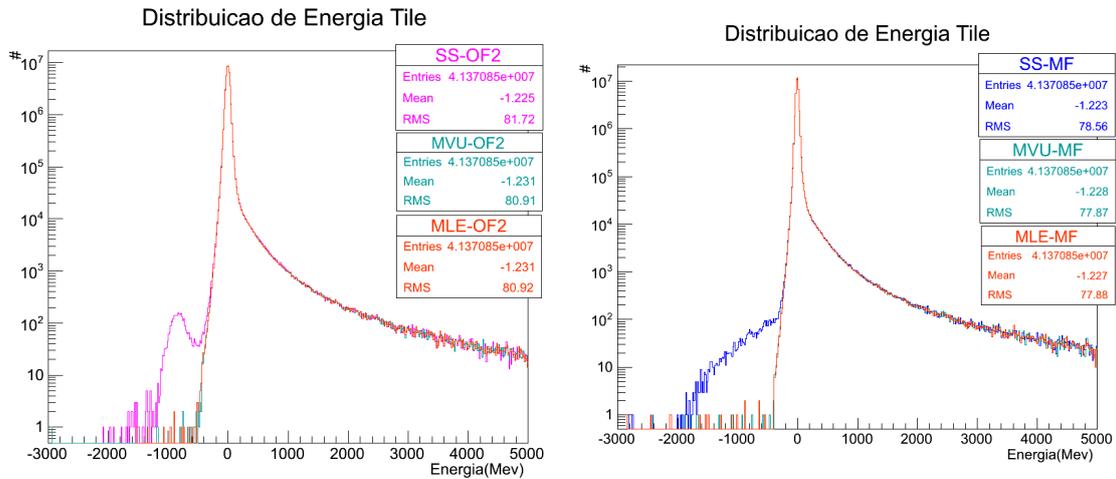


Figura 8.10: Distribuico de energia na regio de rudo das clulas do TileCal reconstrudo pelos algoritmos OF2 e MF.

A Figura 8.10 mostra a regio com maior presena de rudo (entre -100 e 100 MeV), podemos observar que a estimaco utilizando o algoritmo MF apresenta uma varincia menor em relao a estimaco com o algoritmo OF2. Portanto, o algoritmo MF tende a realizar uma estimaco melhor da energia do canal em relao ao algoritmo OF2. Sabendo dessa superioridade do algoritmo MF, esperava-se que a utilizao do mtodo proposto de estimaco em clula apresentasse um resultado melhor utilizando a estimaco de amplitude realizada por este algoritmo. Para certificar esta hiptese realizamos a estimaco em clula utilizando os estimadores MVU e MLE tanto na energia estimada pelo algoritmo OF2 quanto pelo algoritmo MF. A Figura 8.11 apresenta as distribuo de energia das clulas, a comparao  feita utilizando o mtodo atual de soma simples (SS) e utilizando soma ponderada com pesos estimados pelo estimador MVU ou MLE, na Figura 8.11a foi utilizado a am-

plitude estimada pelo algoritmo OF2 e na Figura 8.11b a amplitude estimada pelo algoritmo MF.



(a) Estimação em célula em conjunto com o algoritmo *Optimal Filter 2*. (b) Estimação em célula em conjunto com o algoritmo *Matched Filter*

Figura 8.11: Comparação das distribuições de energia das células do TileCal utilizando diferentes algoritmos de reconstrução de amplitude.

Como previsto, os resultados da estimação em célula utilizando o algoritmo de estimação de amplitude MF foi superior. A partir desses resultados, e sabendo-se que o algoritmo MF será amplamente utilizado pela colaboração do TileCal como algoritmo de reconstrução de energia nos canais, passamos a utilizar os estimadores propostos neste trabalho para a estimação em célula em conjunto com este método de reconstrução.

A Figura 8.12 apresenta o resultado da estimação em célula utilizando os estimadores MVU e MLE, em comparação com o método utilizado atualmente onde os canais da célula são somados. Constatamos que o RMS da distribuição de energia diminui na região onde encontrarmos maior concentração de ruído, a diminuição da cauda negativa da distribuição confirma que houve diminuição do ruído, assim como observado nos testes realizados utilizando somente dados de pedestal visto na Seção 8.1. A diminuição da cauda negativa deve-se a mitigação de canais com alta variância de ruído em determinadas células através da ponderação realizada com os estimadores MVU ou MLE.

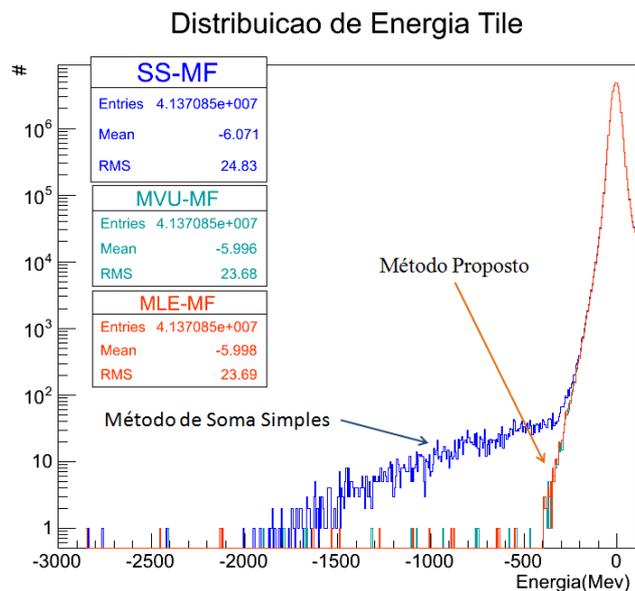


Figura 8.12: Distribuição de energia da células do TileCal na região de ruído das células do TileCal reconstruído pelo método atual e utilizando estimação em célula.

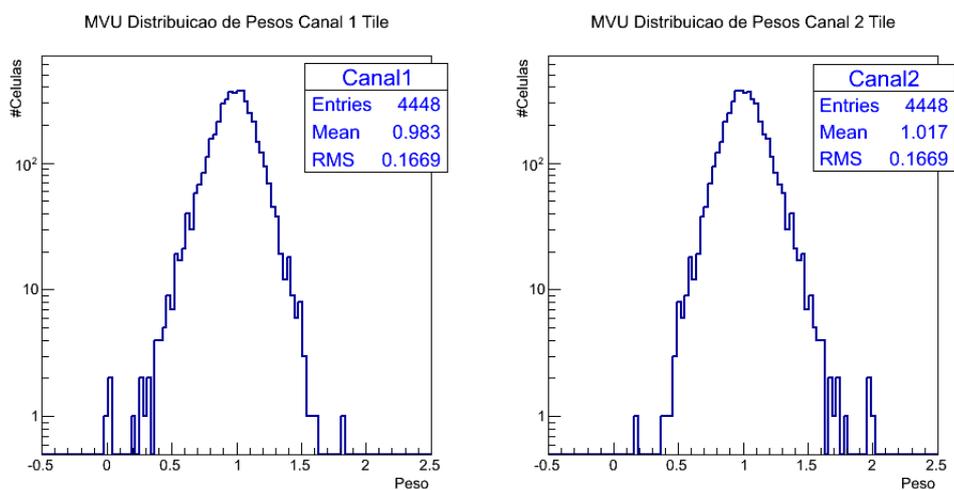


Figura 8.13: Distribuição dos pesos calculados pelo algoritmo MVU para as células do TileCal.

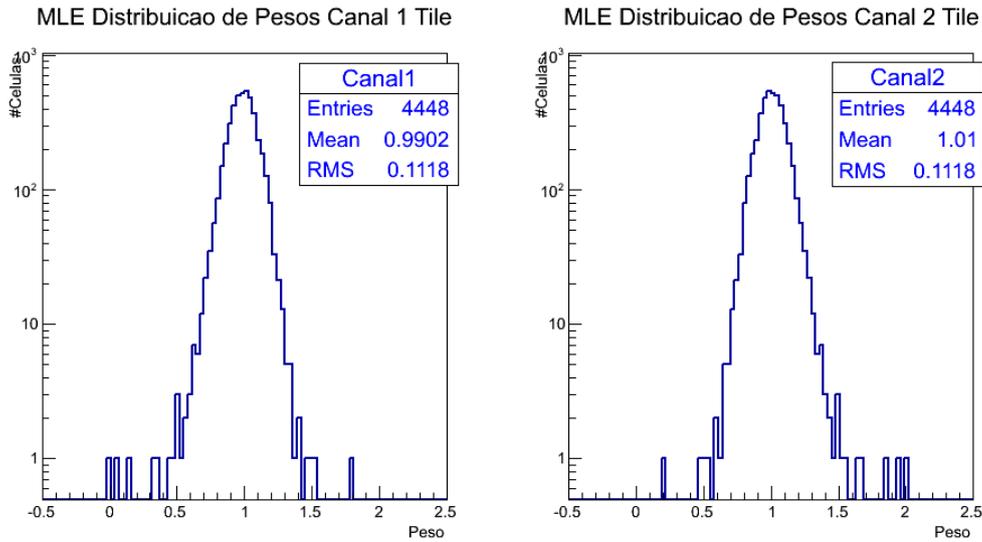


Figura 8.14: Distribuição dos pesos calculados pelo algoritmo MLE para as células do TileCal.

As Figuras 8.13 e 8.14 apresentam as distribuições dos pesos para os canais das células do TileCal utilizados na estimação de energia em célula pelos estimadores MVU e MLE respectivamente. Notamos em ambos os resultados dos estimadores que alguns canais de algumas células receberam pesos muito próximos a 2 enquanto outros canais receberam pesos praticamente igual a zero, mostrando assim que alguns canais tiveram sua participação reduzida na estimação da energia na célula. Comparando os resultados obtidos pelos estimadores MVU e MLE, notamos que em média os estimadores calcularam pesos muito próximos para os dois canais, mas o RMS da distribuição dos pesos do estimador MVU é maior do que do estimador MLE. Como o estimador MVU leva em consideração a correlação entre o ruído dos canais de uma célula ao estimar o peso de cada canal, essa diferença indica que algumas células apresentaram uma correlação entre o ruído dos canais.

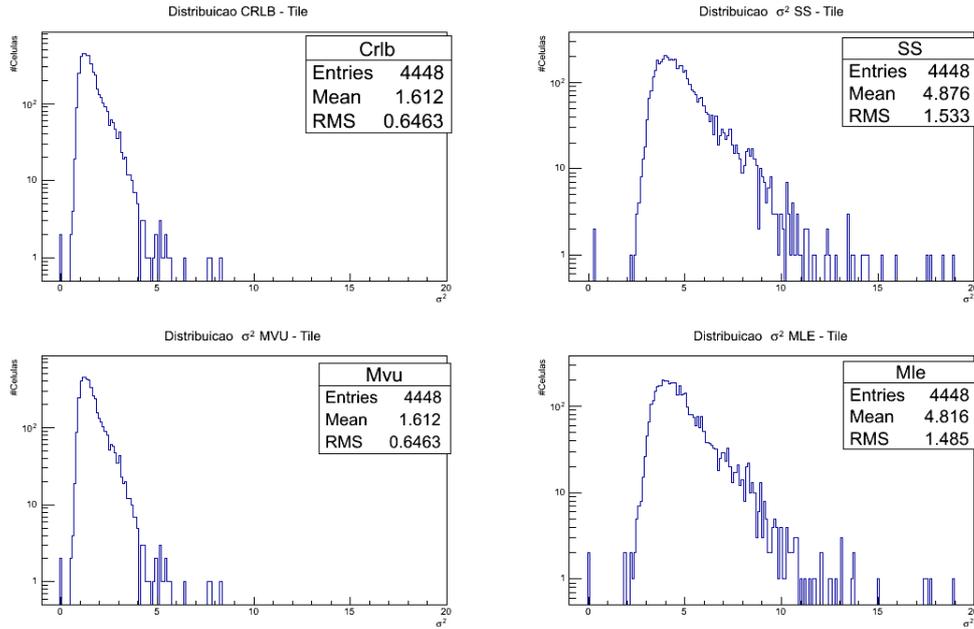


Figura 8.15: Distribuição das variâncias dos estimadores e do limite inferior de Cramer-Rao.

A Figura 8.15 mostra a distribuição da variância dos estimadores e a distribuição do limite inferior de Cramer-Rao. O estimador MVU proposto é o estimador de mínima variância para o caso de um modelo linear com ruído colorido e portanto sua variância se iguala ao limite inferior de Cramer-Rao. Isto se deve ao fato de que o estimador MVU é obtido através da condição de igualdade do teorema de CRLB (Equação 5.23) como demonstrado no Apêndice B. Podemos observar que o estimador MLE também obtém em média uma variância menor do que a método atual (SS).

Para certificarmos os resultados obtidos com o *Run* de 2010, realizamos a estimação em células utilizando os estimadores MVU e MLE também em um *Run* de 2012, onde temos um número ainda maior de eventos reconstruídos.

A Figura 8.16 mostra a distribuição de energia das células do TileCal no *Run* de 2012. Podemos observar novamente a diminuição da cauda negativa da distribuição de energia das células do TileCal, confirmando assim a eficácia da realização da estimação de energia em canais de uma mesma célula.

A Figura 8.17 apresenta a distribuição de energia das células do TileCal onde também apresentamos toda a cauda positiva de energia onde temos predominância de sinal, podemos observar que a cauda positiva não sofre nenhuma alteração significativa com o uso dos estimadores propostos neste trabalho.

As Figuras 8.18 e 8.19 apresentam novamente as distribuições dos pesos para

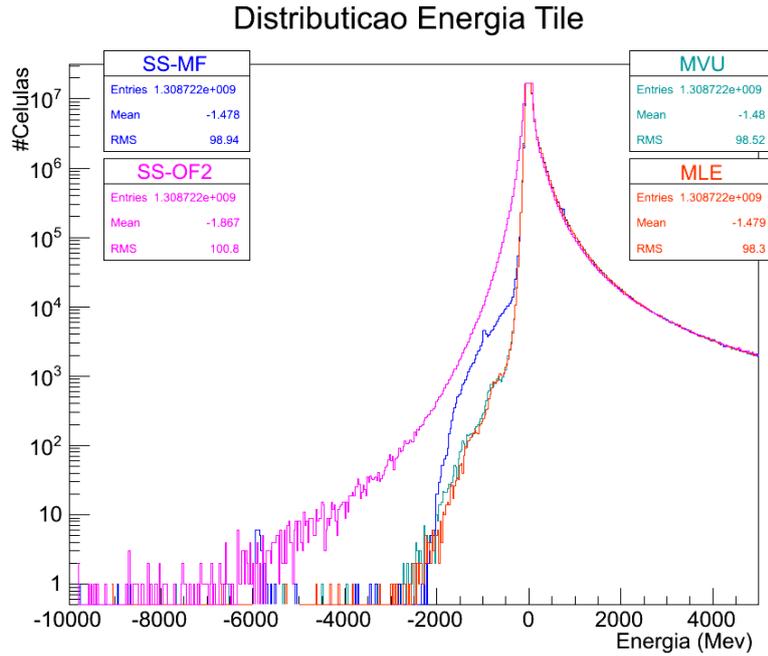


Figura 8.16: Distribuição de energia (entre -10GeV e 5GeV) da células do TileCal reconstruído pelos métodos atuais (OF2 e MF) e utilizando estimação em célula (MVU e MLE).

os canais das células utilizados na estimação de energia pelos estimadores MVU e MLE. Assim como visto no *Run* de 2010 a média dos pesos fica em torno de 1 como esperado, e comparando os pesos calculados pelos estimadores podemos confirmar através da variância da distribuição dos pesos a tendência do estimador MVU em atribuir pesos mais incisivos para cada canal.

A Figura 8.20 mostra a distribuição da variância dos estimadores para o *Run* de 2012 e a distribuição do limite inferior de Cramer-Rao. Podemos observar que o estimador MLE obtém uma média menor do que a método atual (SS).

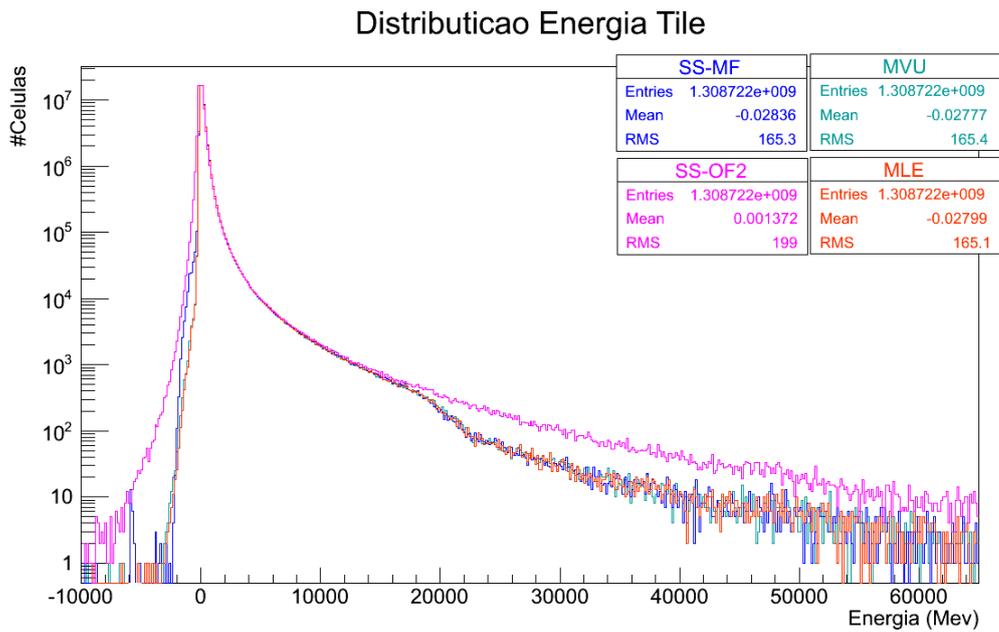


Figura 8.17: Distribuição de energia (entre -10GeV e 65GeV) da células do TileCal reconstruído pelos métodos atuais (OF2 e MF) e utilizando estimação em célula (MVU e MLE).

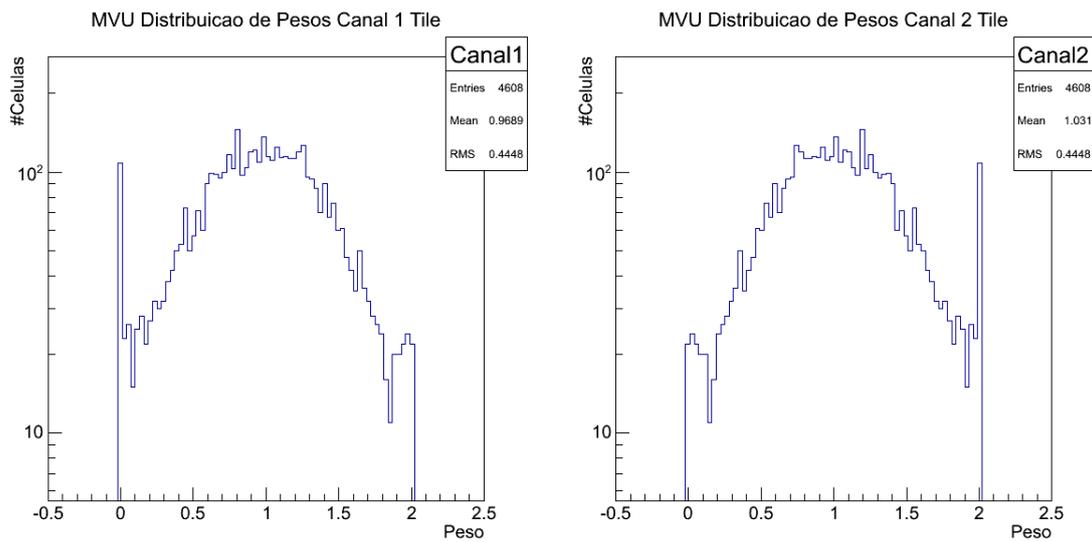


Figura 8.18: Distribuição dos pesos calculados pelo algoritmo MVU para as células do TileCal.

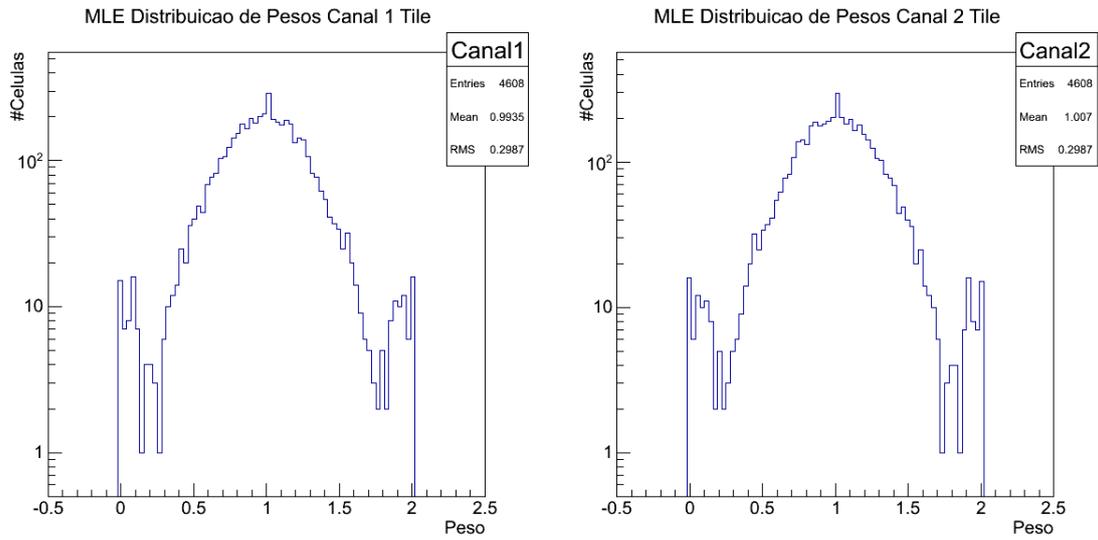


Figura 8.19: Distribuição dos pesos calculados pelo algoritmo MLE para as células do TileCal.

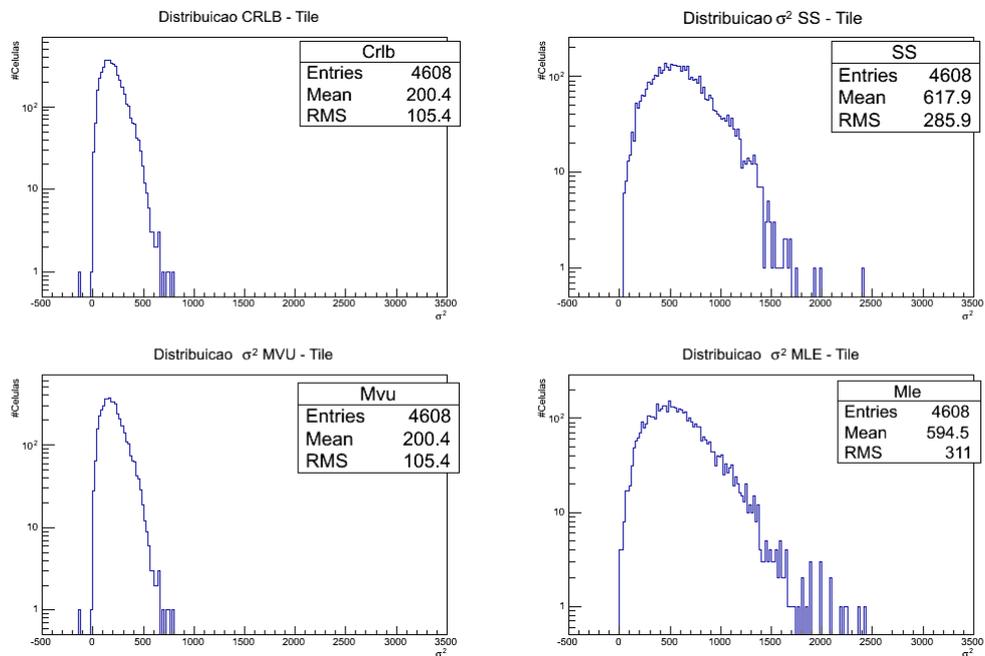


Figura 8.20: Distribuição das variâncias dos estimadores e do limite inferior de Cramer-Rao.

Capítulo 9

Conclusões

Foi apresentado nesta dissertação, um método de estimação de energia em um calorímetro finamente segmentado que utiliza os canais de uma mesma célula de leitura. O método desenvolvido utiliza um estimador de mínima variância, com o objetivo de otimizar a estimação de energia nas células do TileCal. O trabalho apresentou duas propostas diferentes de estimadores como ferramenta para realização da estimação em célula, o estimador *Minimum Variance Unbiased* (MVU) e o estimador *Maximum Likelihood* (MLE). Os dois estimadores são estimadores de mínima variância e diferem apenas na abordagem que ambos dão ao problema. O estimador MLE considera que o ruído dos canais de leitura são descorrelacionados, aditivo, branco e gaussiano e o estimador MVU leva em consideração uma possível correlação entre o ruído dos canais. Por esta razão, o estimador MLE possui uma implementação mais simples do ponto de vista computacional. Os estimadores obtiveram uma boa eficiência na minimização da distribuição de ruído da célula utilizando um conjunto de dados contendo pedestal. Além deste resultado, foi possível constatar, através da análise metódica dos pesos atribuídos pelos estimadores aos canais das células, que os estimadores conseguem mascarar canais que apresentavam uma alta variância de ruído. Estes canais, em geral, são canais defeituosos que eventualmente não entraram na lista de canais defeituosos antes da reconstrução de dados. Portanto, o emprego dos estimadores também pode contribuir para a qualidade de dados na identificação de canais defeituosos.

Aplicamos o método em dois conjuntos de dados reais de colisão de partículas, um do ano de 2010, onde a maioria das fontes de alimentação dos módulos eram versões antigas e conseqüentemente a correlação de ruído entre os canais era alta, e outro do ano de 2012 onde a maioria dos módulos já possuíam novas versões da fonte de alimentação. Observamos que a cauda negativa de ruído nas distribuições de energia diminuiu de forma significativa em comparação com a distribuição de energia usando o método atual de soma dos canais da célula sem ponderação. A distribuição dos

pesos encontrados pelos estimadores apontou que em geral a média dos pesos estão próximas a um, também foi possível perceber que os pesos calculados pelo estimador MVU possui uma variância maior, indicando que existe uma correlação do ruído entre canais de uma mesma célula. A comparação das variâncias dos estimadores MVU, MLE e do método atual, mostra que o estimador MVU atingi o limite inferior de Cramer-Rao, e que a variância do estimador MLE esta mais próxima do limite do que a variância da estimacão de energia utilizada atualmente. A distribuicão de energia na região onde o sinal predomina (acima de 500 MeV) coincide com a distribuicão atual. O método de estimacão apresentado neste trabalho pode ser empregado na estimacão de energia do TileCal em um futuro upgrade das *Read-Out Drivers* (ROD). Além disso, o método proposto de soma ponderada de canais para diminuicão de ruído pode ser empregado em outras áreas do TileCal onde haja a soma de canais de leitura, como por exemplo, as torres de *trigger* do detector, onde células de um mesmo módulo de leitura que estão na mesma região em η são somadas.

Referências Bibliográficas

- [1] “CalorimetroGelo”. Disponível em: <<http://en.wikipedia.org/wiki/Thermochemistry>>. (Acessado em Novembro de 2013).
- [2] “The Bohr Model”. Disponível em: <<http://csep10.phys.utk.edu/astr162/lect/light/bohr.html>>. (Acessado em Março de 2014).
- [3] “Radiação”. Disponível em: <<http://www.fisica.net/aplicada/biofisica/radiacao.php>>. (Acessado em Março de 2014).
- [4] “Conceptual Explanation of the Cherenkov Effect”. Disponível em: <<http://www.lepp.cornell.edu>>. (Acessado em Março de 2014).
- [5] “The Bremsstrahlung Effect”. Disponível em: <<http://www.hardhack.org.au/book/export/html/76>>. (Acessado em Março de 2014).
- [6] SYLVIE BRAIBANT, GIORGIO GIACOMELLI, M. S. *Particles and fundamental interactions: an introduction to particle physics*. Springer, 2012.
- [7] “Chuveiro Hadrônico”. Disponível em: <<http://www.lip.pt/~jespada/>>. (Acessado em Março de 2014).
- [8] “CERN”. <http://www.cern.ch/>, . (Acessado em Agosto de 2013).
- [9] NETTLE, P. “Raios Cósmicos”. <http://www.lip.pt/>. (Acessado em 28 Agosto de 2013).
- [10] “Atlas”. <http://atlas.web.cern.ch/Atlas>. (Acessado em Julho de 2013).
- [11] DI GIROLAMO, B. “An overview of the ATLAS TILECAL hadronic calorimeter”, *ATLAS Internal Note*, September 1996.
- [12] “The ATLAS Collaboration, The ATLAS Experiment at the CERN Large Hadron Collider”, *Journal of Instrumentation*, v. JINST, n. 3 S08003, 2008.
- [13] BIOT, J. A. V. *TileCal Read-Out Drivers Production and Firmware Developments*. Relatório técnico, IFIC - Universitat de Valencia.

- [14] “The University of Chicago”. <http://hep.uchicago.edu/atlas/tilecal/>. Disponível em: <<http://hep.uchicago.edu/atlas/tilecal/>>. (Acessado em Fevereiro de 2014).
- [15] *Design of a New Switching Power Supply for the ATLAS TileCal Front-End Electronics*, Setembro de 2012. Argonne National Laboratory, the Argonne HEP Electronics Group.
- [16] “GaussianWaves - Signal Processing Simplified”. Disponível em: <<http://www.gaussianwaves.com/>>. (Acessado em Março de 2014).
- [17] WILLIS W.J., R. V. “Liquid-Argon Ionization Chambers as Total-Absorption Detectors”, *Nuclear Instruments and Methods*, , n. 120, 1974.
- [18] CALHEIROS, F., HRUSKA, I. *Correlated noise unfolding on a Hadronic Calorimeter, ATLAS-TILECAL-PROC-2011-003*. Relatório técnico, CERN, Geneva Switzerland, .
- [19] MANCINI-TERRACCIANO, C. “Noise correlation and decorrelation in arrays of bolometric detectors”, *JINST*, 2012.
- [20] “ATLAS Run Query”. Disponível em: <<http://atlas-runquery.cern.ch/>>. (Acessado em Março de 2013).
- [21] “SLAC”. Disponível em: <<http://www.slac.stanford.edu/>>. (Acessado em Junho de 2014).
- [22] “Linear accelerator”. Disponível em: <<http://en.wikipedia.org/wiki/Particleaccelerator>>. (Acessado em Novembro de 2013).
- [23] GREENE, B. *O Universo Elegante*. Cia da Letras, 1999.
- [24] “AstroPT”. <http://astropt.org/blog/>. Disponível em: <<http://astropt.org/blog/>>. (Acessado em Fevereiro de 2014).
- [25] *A FÍSICA DO LHC: Modelo Padrão e Além e Principais Resultados Experimentais (Carla Gobel)*, Fevereiro de 2013. PUC-Rio.
- [26] “Pierre Auger”. <http://www.auger.org/>. (Acessado em Março de 2014).
- [27] WIGMANS, R. *Calorimetry: Energy Measurements in Particle Physics*. Clarendon Press, 2000.
- [28] FABJAN, C., STRUCZINSKI, W., WILLIS, W., et al. “Iron Liquid - Argon and Uranium Liquid - Argon Calorimeters for Hadron Energy Measurement”, *Nucl.Instrum.Meth.*, v. 141, pp. 61. doi: 10.1016/0029-554X(77)90747-9.

- [29] “The history of CERN”. <http://timeline.web.cern.ch/timelines/The-history-of-CERN>, . (Acessado em Abril de 2014).
- [30] “ALICE”. <http://aliweb.cern.ch/>. (Acessado em Junho de 2014).
- [31] “CMS”. <http://cms.web.cern.ch/>. (Acessado em Junho de 2014).
- [32] “LHCb”. <https://lhcb.web.cern.ch/lhcb/>, . (Acessado em Junho de 2014).
- [33] “TOTEM”. <http://totem.web.cern.ch/Totem/>. (Acessado em Abril de 2014).
- [34] “LHCf”. <http://home.web.cern.ch/about/experiments/lhcf>, . (Acessado em Janeiro de 2014).
- [35] “LEP”. <http://home.web.cern.ch/about/accelerators/large-electron-positron-collider>. (Acessado em Janeiro de 2014).
- [36] SCHMIDT, R. “LHC - CHALLENGES IN HANDLING BEAMS EXCEEDING 100 MJ”, *Proceedings of HB2012, Beijing-China, 2012*.
- [37] COLLABORATION, A. *ATLAS Detector and Physics Performance - Technical Design Report Vol.I*. Relatório técnico, CERN.
- [38] DRAKE, G. *Design of a new switching power supply for the ATLAS TileCAL front-end electronics*. Relatório técnico, Argonne National Laboratory-ANL.
- [39] CALHEIROS, F., HRUSKA, I. *Application of ELMB for steering LV power supply for TILECAL, ATLAS internal note, ATL-COM-2003 001*. Relatório técnico, CERN, Geneva Switzerland (2003), .
- [40] GREEN, D. *At the leading edge the ATLAS and CMS LHC experiments*. World Scientific, 2010.
- [41] SCHOTTKY, W. “Small-shot effect and flicker effect”, *Phys.Rev.*, vol. 28,p. 74, 1926.
- [42] H, N. “Thermal agitation of electricity in conductors”, *Phys.Rev.*, vol. 32,p. 110, 1927.
- [43] E G.E. UHLENBECK, J. L. *Threshold Signals(MIT Rad.Lab.)*, v. 24. New York, NY:McGraw-Hill, 1956.
- [44] MIDDLETON, D. *Statistical Communication Theory*. New York, NY:McGraw-Hill, 1960.

- [45] RICE, S. “Mathematical analysis of random noise”, *Bell Syst. Tech. J.*, v. 23 e 24, pp. 282 e 46, 1944 e 1945.
- [46] CRAMER, H. *Mathematical Methods of Statistics*. Princeton University Press, New Jersey, 1946.
- [47] A.WALD. *Sequential Analysis*. New York, NY: Wiley, 1947.
- [48] A.WALD. *Statistical Decision Functions*. New York, NY: Wiley, 1950.
- [49] J.DOOB. *Stochastic Processes*. New York, NY: Wiley, 1953.
- [50] M.LOÉVE. *Probability Theory*. New York, NY D. Von Nostrand, 1955.
- [51] GRENDER, U. *Stochastic processes and statistical inference*, v. 1. Arkiv fur Matematik, 1950.
- [52] KAY, S. M. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, 1993.
- [53] P. ADRAGNA, E. A. “Testbeam studies of production modules of the ATLAS Tile Calorimeter”, *nucl. Inst. Meth*, v. A606, 2009.
- [54] PERALVA, B. S.-M. *Detecção de Sinais e Estimacão de Energia para Calorimetria de Altas Energias*. Tese de Mestrado, UFJF, Juiz de Fora, MG, Brasil, 2012.
- [55] KNOLL, G. F. *Radiation Detection and Measurement*. New York: Wiley, 1979.
- [56] MANCINI-TERRACCIANO, C. “INSPIRE High Energy Physics information system”. <http://inspirehep.net>, . (Acessado em Julho de 2013).
- [57] MANCINI-TERRACCIANO, C. “SciVerse”. <http://www.sciencedirect.com>, . (Acessado em Julho de 2013).
- [58] MANCINI-TERRACCIANO, C. “CUORE-Cryogenic Underground Observatory for Rare Events”. <http://crio.mib.infn.it/wigmi/pages/cuore.php>, . (Acessado em Julho de 2013).
- [59] “A Aventura das Partículas”. Disponível em: <http://www.aventuradasparticulas.ift.unesp.br>. (Acessado em Novembro de 2013).
- [60] V., R. “LOW-NOISE TECHNIQUES IN DETECTORS”, *Brookhaven National Laboratory*, 1973.

- [61] J.C. STREET, E. S. “New Evidence for the Existence of a Particle of Mass Intermediate Between the Proton and Electron”, *Physical Review*, v. 52, pp. 1003–1004, 1937.
- [62] GRUPEN C., S. B. *Particle Detectors*. Cambridge University Press.
- [63] COTTINGHAM, W., GREENWOOD, D. *An Introduction to the Model of Particle Physics*. Cambridge University Press, 2001.

Apêndice A

Teoria do Limite Inferior de Cramer-Rao

Neste apêndice encontra-se a demonstração do Teorema do Limite inferior de Cramer-Rao para uma função de um parâmetro escalar, assim como a demonstração para que a condição de regularidade seja verdadeira.

A.1 Demonstração do Limite Inferior de Cramer-Rao para uma Função de um Parâmetro Escalar

Considerando uma função de um parâmetro escalar com $g(\theta) = \alpha$ (sendo α um escalar qualquer) e que os estimadores do parâmetro ($\hat{\alpha}$) sejam imparciais, ou seja:

$$E[\hat{\alpha}] = \alpha = g(\theta) \implies \int \hat{\alpha} p(x; \theta) dx = g(\theta) \quad (\text{A.1})$$

Visto que, o valor esperado de uma função de x , $g(x)$, dado que x tenha p.d.f $f(x)$, é dado pelo produto interno:

$$E[g(x)] = \int_{-\infty}^{\infty} g(x) f(x) dx$$

Examinando a condição de regularidade do teorema:

$$E \left[\frac{\partial \ln p(x; \theta)}{\partial \theta} \right] = 0, \quad (\text{A.2})$$

Notamos que

$$\begin{aligned} \int \frac{\partial \ln p(x; \theta)}{\partial \theta} p(x; \theta) dx &= \int \frac{\partial p(x; \theta)}{\partial \theta} dx \\ &= \frac{\partial}{\partial \theta} \int p(x; \theta) dx \end{aligned}$$

$$\begin{aligned}
&= \frac{\partial 1}{\theta} \\
&= 0
\end{aligned}$$

Assim, a condição de regularidade será satisfeita se a ordem da integral e derivada for intercambiável, ou seja, se a regra de Leibniz for realizável.

Derivando ambos os lados da Equação A.1 em relação a θ e trocando a ordem da integração com a derivada temos:

$$\int \hat{\alpha} \frac{\partial p(x; \theta)}{\partial \theta} dx = \int \hat{\alpha} \frac{\partial \ln p(x; \theta)}{\partial \theta} p(x; \theta) dx = \frac{\partial g(\theta)}{\partial \theta} \quad (\text{A.3})$$

Uma vez que pela condição de regularidade (equação A.2) podemos ter:

$$\alpha E \left[\frac{\partial \ln p(x; \theta)}{\partial \theta} \right] = \int \alpha \frac{\partial \ln p(x; \theta)}{\partial \theta} p(x; \theta) dx = 0 \quad (\text{A.4})$$

podemos usar essa condição de regularidade, para reescrever a equação A.3 da seguinte forma:

$$\int (\hat{\alpha} - \alpha) \frac{\partial \ln p(x; \theta)}{\partial \theta} p(x; \theta) dx = \frac{\partial g(\theta)}{\partial \theta}. \quad (\text{A.5})$$

Agora aplicando a desigualdade de Cauchy-Schwarz:

$$\left[\int w(x)g(x)h(x)dx \right]^2 \leq \int w(x)g^2(x)dx \int w(x)h^2 dx$$

considerando $w(x) = p(x; \theta)$, $g(x) = \hat{\alpha} - \alpha$, e $h(x) = \frac{\partial \ln p(x; \theta)}{\partial \theta}$, teremos:

$$\left(\frac{\partial g(\theta)}{\partial \theta} \right)^2 \leq \int (\hat{\alpha} - \alpha)^2 p(x; \theta) dx \int \left(\frac{\partial \ln p(x; \theta)}{\partial \theta} \right)^2 p(x; \theta) dx \quad (\text{A.6})$$

podemos observar que a equação pode ser reescrita de forma simplificada dado que:

$$\begin{aligned}
\int (\hat{\alpha} - \alpha)^2 p(x; \theta) dx &= \text{var}(\hat{\alpha}) \\
\int \left(\frac{\partial \ln p(x; \theta)}{\partial \theta} \right)^2 p(x; \theta) dx &= E \left[\left(\frac{\partial \ln p(x; \theta)}{\partial \theta} \right)^2 \right]
\end{aligned}$$

então a equação A.6 será igual a:

$$\text{var}(\hat{\alpha}) \geq \frac{\left(\frac{\partial g(\theta)}{\partial \theta} \right)^2}{E \left[\left(\frac{\partial \ln p(x; \theta)}{\partial \theta} \right)^2 \right]} \quad (\text{A.7})$$

Podemos demonstrar que:

$$E \left[\left(\frac{\partial \ln p(x; \theta)}{\partial \theta} \right)^2 \right] = -E \left[\frac{\partial^2 \ln p(x; \theta)}{\partial \theta^2} \right]$$

utilizando novamente a condição de regularidade (Equação A.2), temos:

$$\begin{aligned} -E \left[\frac{\partial^2 \ln p(x; \theta)}{\partial \theta^2} \right] &= \int \frac{\partial \ln p(x; \theta)}{\partial \theta} \frac{\partial \ln p(x; \theta)}{\partial \theta} p(x; \theta) dx \\ &= E \left[\left(\frac{\partial \ln p(x; \theta)}{\partial \theta} \right)^2 \right] \end{aligned}$$

Sendo assim,

$$\text{var}(\hat{\alpha}) \geq \frac{\left(\frac{\partial g(\theta)}{\partial \theta} \right)^2}{-E \left[\left(\frac{\partial^2 \ln p(x; \theta)}{\partial \theta^2} \right) \right]} \quad (\text{A.8})$$

Sendo assim, a Equação A.8 é a inequação do teorema de Cramer-Rao para o caso geral de uma função de um parâmetro escalar $g(\theta)$ como queríamos demonstrar. No caso de $g(\theta) = \theta$ então a equação se torna:

$$\text{var}(\hat{\theta}) \geq \frac{1}{-E \left[\left(\frac{\partial^2 \ln p(x; \theta)}{\partial \theta^2} \right) \right]} \quad (\text{A.9})$$

assim como visto na Equação 5.17.

A condição de igualdade para que um estimador imparcial possa se encontrado e que alcance o limite de CR para todo θ é dada por:

$$\frac{\partial \ln p(x; \theta)}{\partial \theta} = \frac{1}{c} (\hat{\alpha} - \alpha),$$

onde c pode depender de θ , mas não pode depender de x . Se $\alpha = g(\theta) = \theta$, teremos:

$$\frac{\partial \ln p(x; \theta)}{\partial \theta} = \frac{1}{c(\theta)} (\hat{\theta} - \theta),$$

para determinarmos $c(\theta)$, derivando ambos os lados da equação obtemos:

$$\frac{\partial^2 \ln p(x; \theta)}{\partial \theta^2} = -\frac{1}{c(\theta)} + \frac{\partial \left(\frac{1}{c(\theta)} \right)}{\partial \theta} (\hat{\theta} - \theta),$$

estimando o valor esperado da equação e sabendo que $E[\hat{\theta}] = \theta$, logo

$$-E \left[\frac{\partial^2 \ln p(x; \theta)}{\partial \theta^2} \right] = \frac{1}{c(\theta)},$$

finalmente conseguimos encontrar a relação de $c(\theta)$:

$$\begin{aligned} c(\theta) &= \frac{1}{-E \left[\frac{\partial^2 \ln p(x; \theta)}{\partial \theta^2} \right]} \\ &= \frac{1}{I(\theta)}. \end{aligned}$$

Apêndice B

Estimador de Mínima Variância para um Modelo Linear

Neste apêndice encontra-se a demonstração do Estimador de Mínima Variância para o modelo linear, e a extensão desse estimador para o caso de ruído colorido.

B.1 Derivação do Estimador Imparcial de Mínima Variância para um Modelo Linear através do Teorema de Cramer-Rao

Dado o modelo Linear da forma

$$x[n] = H\theta + w[n] \quad (\text{B.1})$$

Determinamos o estimador MVU através da restrição de igualdade do CRLB. Pelo Teorema proposto na Seção A.1, $\hat{\theta} = g(x)$ será o estimador MVU se

$$\frac{\partial \ln p(x; \theta)}{\partial \theta} = I(\theta)(g(x) - \theta),$$

para alguma função g . Dado a seguinte função de verossimilhança e seu logaritmo natural do modelo Linear

$$p(x; H\theta) = \prod_{n=0}^{N-1} \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[\frac{1}{2\sigma^2} (x[n] - H\theta)^2 \right]$$
$$\ln p(x; H\theta) = -\ln(2\pi\sigma^2)^{N/2} - \frac{1}{2\sigma^2} (x - H\theta)^T (x - H\theta).$$

Derivando o logaritmo da função de verossimilhança obtemos

$$\frac{\partial \ln p(x; \theta)}{\partial \theta} = \frac{\partial}{\partial \theta} \left[-\ln(2\pi\sigma^2)^{N/2} - \frac{1}{2\sigma^2} (x - H\theta)^T (x - H\theta) \right]$$

$$\begin{aligned}
&= -\frac{1}{2\sigma^2} \frac{\partial}{\partial \theta} [x^T(x - H\theta) - (H\theta)^T(x - H\theta)] \\
&= -\frac{1}{2\sigma^2} \frac{\partial}{\partial \theta} [x^T x - 2x^T H\theta + \theta^T H^T H\theta]
\end{aligned}$$

conhecendo as seguintes propriedades de derivação de matrizes

$$\begin{aligned}
\frac{\partial X^T A}{\partial X} &= \frac{\partial A^T X}{\partial X} = A, \text{ onde } \mathbf{A} \text{ não depende de } \mathbf{X}, \\
\frac{\partial X^T B X}{\partial X} &= (B + B^T)X, \text{ onde } \mathbf{B} \text{ não depende de } \mathbf{X}.
\end{aligned}$$

Uma vez que

$$\begin{aligned}
\frac{\partial x^T H\theta}{\partial \theta} &= H^T \theta, \\
\frac{\partial \theta^T H^T H\theta}{\partial \theta} &= 2H^T H\theta,
\end{aligned}$$

nós teremos

$$\frac{\partial \ln p(x; \theta)}{\partial \theta} = \frac{1}{\sigma^2} [H^T x - H^T H\theta].$$

Assumindo que $H^T H$ é inversível obtemos

$$\frac{\partial \ln p(x; \theta)}{\partial \theta} = \frac{H^T H}{\sigma^2} [(H^T H)^{-1} H^T x - \theta],$$

que possui a mesma forma da restrição de igualdade do Teorema de CRLB, de onde podemos inferir o estimador MVU do modelo linear, a informação de Fisher e consequentemente sua variância.

$$\hat{\theta} = (H^T H)^{-1} H^T x \tag{B.2}$$

B.1.1 Extensão do Estimador Imparcial de Mínima Variância para um Modelo Linear Corrompido com Ruído Colorido

Dado o ruído da forma

$$w \sim \mathcal{N}(0, C)$$

Uma vez que a matriz de covariância \mathbf{C} seja positiva definida, podemos decompor C^{-1} da seguinte forma:

$$C^{-1} = D^T D$$

onde \mathbf{D} é uma matriz inversível $N \times N$ que atua como matriz de branqueamento quando aplicada a \mathbf{w} uma vez que:

$$E[(Dw)(Dw)^T] = DCD^T = D^{-1}D^{T^{-1}}D^T = I.$$

Podemos branquear nosso modelo linear de dados

$$x = H\theta + w$$

para

$$\begin{aligned} x' &= Dx \\ &= DH\theta + Dw \\ &= H'\theta + w' \end{aligned}$$

o ruído será branqueado uma vez que $w' = N(0, I)$. Substituindo o modelo linear obtido na equação do estimador MVU para o modelo linear teremos

$$\begin{aligned} \hat{\theta} &= (H^T H)^{-1} H^T x \\ &= (H^T D^T D H)^{-1} H^T D^T D x \end{aligned}$$

sabendo-se que $C^{-1} = D^T D$, obtemos o estimador

$$\hat{\theta} = (H^T C^{-1} H)^{-1} H^T C^{-1} x. \quad (\text{B.3})$$