

TÉCNICAS DE *DENOISING* PARA A ESTIMAÇÃO DE ENERGIA COM UM CALORÍMETRO FINAMENTE SEGMENTADO

RODRIGO A. PEREIRA*, BERNARDO S. PERALVA†, JOSÉ M. DE SEIXAS*

* *Universidade Federal do Rio de Janeiro*
Rio de Janeiro, Rio de Janeiro, Brasil

† *Universidade do Estado do Rio de Janeiro*
Nova Friburgo, Rio de Janeiro, Brasil

Emails: rodrigo@cern.ch, bernardo@iprj.uerj.br, seixas@lps.ufrj.br

Abstract— Particle accelerators such as the LHC (Large Hadron Collider) operate at high energy and luminosity allowing the study of rare particles in elementary dimensions. The LHC has experiments such as ATLAS (A Toroidal LHC Apparatus), which is located around the point where the beams of particles collide, producing a series of particles. This work proposes a minimum variance estimator to reduce the noise in the energy estimation of the more than 5,000 calorimeter cells that use a double readout approach and are responsible for detecting the energy of the hadronic particles interacting within ATLAS.

Keywords— Parameter Estimation, Calorimetry, Signal Processing.

Resumo— Aceleradores de partículas como o LHC (*Large Hadron Collider*) trabalham com altas energias e luminosidade permitindo o estudo de partículas raras e em dimensões ínfimas. O LHC possui experimentos como o ATLAS (*A Toroidal LHC Apparatus*) localizados ao redor do ponto onde os feixes de partículas colidem produzindo uma série de partículas. Este trabalho propõe um estimador de mínima variância para reduzir o ruído na estimação da energia de mais de 5.000 células do calorímetro responsável por detectar a energia das partículas que interagem de forma hadrônica no ATLAS, utilizando a abordagem de dupla leitura dessas células.

Palavras-chave— Estimação de Parâmetros, Calorimetria, Processamento de Sinais.

1 Introdução

Os avanços nas áreas de instrumentação eletrônica, processamento e armazenamento de dados impulsionam a grande difusão de sistemas eletrônicos assim como o desenvolvimento tecnológico. Existe, atualmente, uma tendência na utilização de um número crescente de sensores atuando com taxas de aquisição cada vez maiores. Consequentemente, observa-se um aumento da demanda para uma quantidade cada vez maior de dados a serem processados.

Os experimentos da física de altas energias envolvem justamente a aquisição e processamento de milhares de sinais a uma taxa de aquisição extremamente alta. Aceleradores de partículas atuam como grandes microscópios para o estudo de partículas subatômicas. Nestes experimentos, feixes de partículas são acelerados próximos à velocidade da luz colidindo em pontos específicos, em torno dos quais experimentos realizam as medições sobre os subprodutos resultantes das colisões. A alta energia e alta luminosidade (medida proporcional ao número de partículas e inversamente proporcional à seção transversal do feixe de partículas) envolvidas nesses experimentos permitem o estudo de partículas raras em dimensões da ordem de 10^{-20} m (Cottingham and Greenwood, 1998).

O LHC é o acelerador de partículas mais energético do mundo, e opera com sucesso no CERN desde 2010 colidindo prótons na taxa de 4×10^7 Hz e com energia variando entre 1,442 e $2,0828 \times 10^{-6}$

J. O ATLAS é o maior dos experimentos do LHC e possui um programa extenso de física, cobrindo todo o potencial de descoberta do LHC. Mais recentemente, comprovações da existência do bóson de Higgs (partícula prevista pelo Modelo Padrão (Cottingham and Greenwood, 1998)) foram observadas no ATLAS (ATLAS, 2012), que levou ao prêmio Nobel de física para seus proponentes.

No ATLAS, diversos sistemas são projetados para coletar o maior número de informações possíveis das colisões. Por exemplo, o ATLAS se apoia fortemente no seu sistema de calorimetria, que é responsável pela absorção e amostragem da energia das partículas incidentes. O ATLAS possui dois calorímetros, Eletromagnético e Hadrônico, que medem a energia de partículas que interagem de forma eletromagnética e hadrônica respectivamente. A informação da energia é extremamente importante para a reconstrução e identificação das partículas resultantes, bem como para o processo de filtragem online de eventos do ATLAS. Quando partículas são absorvidas pelo sistema de calorimetria, suas propriedades podem ser reconstruídas. A energia depositada em cada célula de leitura é medida individualmente através de sua eletrônica dedicada que fornece um sinal conformado e digitalizado. Do ponto de vista do sistema de aquisição de dados, cada célula corresponde a um canal de leitura individual.

O calorímetro hadrônico (TileCal) do ATLAS (ATLAS, 2008) é um calorímetro composto por mais de 5.000 células com dupla

leitura que operam numa taxa de eventos de 40 MHz (ATLAS, 2010). Esse modelo de leitura dupla é utilizado principalmente para permitir o cálculo exato da posição em que a partícula penetrou na célula e também para no caso de falha de um dos canais, podendo o canal restante substituir o canal defeituoso e evitar que a célula de leitura seja totalmente desativada (redundância). Adicionalmente, 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 (Wigmans, 2000); no detector CHORUS localizado no CERN (*Chorus Detector*, 2018); nas células de leitura do Telescópio Subaquático de Neutrinos situado no lago Baikal na Rússia (*Experiment Baikal*, 2018); assim como em células de leitura do detector SPS (*Super Proton Synchrotron*) no CERN (Wigmans, 2000).

Visando melhorar a estimativa da energia nas células de leituras do TileCal, o presente trabalho propõe um método que utiliza um estimador de mínima variância para reduzir o ruído na leitura de energia dos canais que formam cada célula de leitura no TileCal. O objetivo é maximizar a relação sinal ruído das células de leitura do calorímetro, visando aumentar a eficiência de estimação da energia. O método proposto consiste em utilizar a redundância de leitura existente nas células do TileCal para realizar a estimação de energia ao nível da célula do calorímetro, em contraposição ao método utilizado atualmente, que consiste simplesmente em somar com pesos iguais, os dois canais de leitura da célula. 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. Assim, pode-se 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 monitorização do TileCal.

Na próxima seção, o TileCal será apresentado assim como o método atualmente utilizado para estimação da energia de suas células. Na Seção 3 será descrito o método proposto que visa aumentar a relação sinal ruído entre os canais que compõem a mesma célula. O conjunto de dados utilizado para testar a eficiência e os resultados são apresentados na Seção 4. Finalmente, as conclusões e discussões são apresentadas na Seção 5.

2 O Calorímetro Hadrônico TileCal

O TileCal é o principal calorímetro hadrônico do ATLAS, e foi desenvolvido para ser um calorímetro de amostragem que utiliza aço como material absorvedor e telhas cintiladoras como meio ativo.

Ele é dividido em três partes cilíndricas: um barril central (dividido em duas partições LBA e LBC) de 5,64 metros de comprimento 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. No TileCal as telhas cintiladoras possuem diferentes tamanhos e são agrupadas compondo os canais de leitura. Tais telhas emitem uma cintilação proporcional a quantidade de energia da partícula que a atravessou. Cada partição do barril central e do barril estendido é dividida em 64 módulos onde cada módulo possui, respectivamente, em 23 e 16 células com dupla leitura, resultando em aproximadamente 10.000 canais (sinais).

A luz produzida nas telhas cintiladoras possui frequência na faixa do Ultra Violeta (UV) e sua intensidade é proporcional à energia depositada pela partícula. A cintilação produzida se propaga dentro da telha para suas bordas através de reflexões e esse feixe de luz é então transmitido através de fibras ópticas até um fotomultiplicador (*Photo Multiplier Tube - PMT*) que, então, transforma esse feixe de luz em um sinal elétrico (Anderson K., 1998). Na Figura 1 pode-se observar o processo de aquisição do sinal no TileCal. A partícula (*source path*) atravessa o calorímetro e telhas cintilantes amostram a energia na forma de luz, que é transmitida ao PMT. Na configuração de dupla leitura (*double readout*), cada telha de uma célula do TileCal possui em suas laterais fibras ópticas, que são conectadas a dois fotomultiplicadores para a coleta dos sinais de uma a mesma célula.

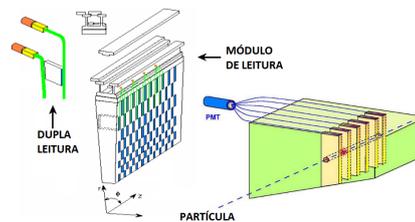


Figura 1: Princípio de aquisição e coleta do sinal no calorímetro de telhas do ATLAS (TileCal).

Cada módulo é segmentado radialmente (profundidade) em três camadas de células de leitura com fina granularidade. A segmentação de células do TileCal é representada na Figura 2 para um módulo do barril central e um módulo do barril estendido.

O pulso rápido gerado na saída do PMT é condicionado por um circuito de conformação (*shaper*) (Anderson K., 1998), o qual fornece um pulso com amplitude proporcional à energia depositada para cada canal de leitura. Os sinais analógicos são amostrados na taxa de 40 MHz e uma janela com 7 amostras (150 ns) é suficiente para cobrir todo o pulso. O pulso analógico e suas amostras,

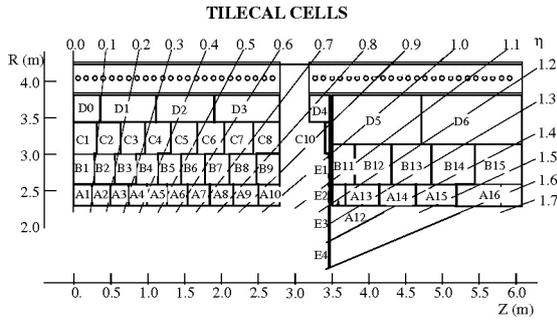


Figura 2: Vista esquemática da segmentação de células de um módulo do barril central ($\eta > 0$) e um módulo do barril estendido.

representadas por pontos, podem ser vistos na Figura 3, assim como os parâmetros pertinentes à reconstrução ao pulso (amplitude, fase e pedestal). Após a estimativa da amplitude (energia), os canais associados à mesma célula são somados compondo a informação final energia da célula do TileCal.

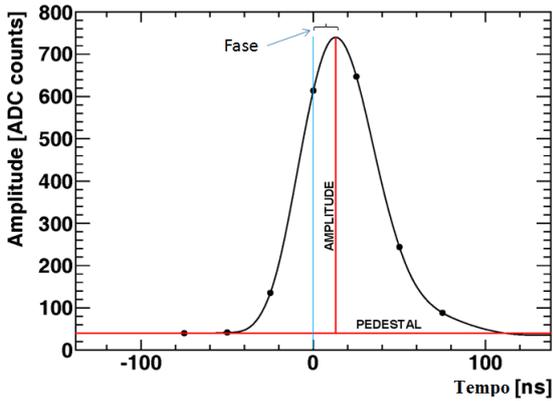


Figura 3: O sinal característico do TileCal e seus parâmetros. As sete amostras são representadas por pontos preenchidos no sinal.

2.1 Estimação de Energia

Todo sinal discretizado produzido pela eletrônica de leitura do TileCal mantém fixa a forma de pulso $h(t)$ (Figura 3), diferindo, essencialmente, em sua amplitude (A), o que facilita a estimação da energia, pois esta se baseia apenas na amplitude do pulso amostrado. 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*). Se o ruído eletrônico Gaussiano característico de cada canal ($w[n]$) também for incluído, pode-se descrever o sinal $s[n]$ de cada canal de leitura com a seguinte equação:

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

Para estimar as informações de interesse do

sinal amostrado (amplitude, pedestal e fase), existem algoritmos de estimação de amplitude que atuam de forma independente para o sinal de cada canal de leitura. Um filtro ótimo de estimação *Optimal Filtering* (OF2) (E. Fullana, 2006) realiza a estimação de energia com vistas a minimizar a variância da amplitude do sinal, realizando uma combinação linear das amostras digitais do sinal utilizando coeficientes calculados pelo método dos multiplicadores de Lagrange (G. Bertuccio and Sapietro, 1992). Esse estimador é atualmente utilizado para reconstruir a amplitude e a fase do sinal digitalizado de cada canal do TileCal.

Outro técnica de reconstrução de energia que apresenta bons resultados utiliza um Filtro Casado (MF, do inglês *Matched Filter*) (Trees, 1992) como método. Este algoritmo (Peralva, 2013), baseado no teste da razão de verossimilhança, pode ser adequado para todo o tipo de modelo de ruído de fundo.

O método aqui proposto tem por objetivo complementar a estimação da amplitude de energia realizada pelos algoritmos de reconstrução de energia disponíveis, atuando tanto com o OF2, quanto com o MF, que realizam a estimação de energia, individualmente em cada canal da célula através das amostras do sinal adquirido. Atualmente para se obter a energia da célula, as energias estimadas dos canais da célula são somadas utilizando pesos iguais (unitários):

$$E_{cell} = E_{pmt_1} + E_{pmt_2}. \quad (1)$$

Para realizar a estimação de energia a nível de célula, este trabalho propõe a utilização da técnica de *denoising* para minimizar o efeito do ruído produzido pelos canais de leitura que formam cada célula do calorímetro e otimizar a estimação de energia.

3 Técnicas de Denoising

Embora seja difícil projetar um estimador não-tendencioso de mínima variância MVU (em inglês, *Minimum Variance Unbiased*), muitos problemas de estimação em processamento de sinais podem ser representados através de um modelo linear em que o estimador MVU possa ser obtido (Kay, 1993). Tal modelo pode ser definido da seguinte forma:

$$x[n] = A + B\theta + w[n] \quad n = 0, 1, 2, \dots, N-1 \quad (2)$$

sendo $w[n]$ um ruído branco Gaussiano, B a inclinação do modelo linear e A o *offset*. Pode-se reescrever o modelo linear na forma matricial, onde θ é o parâmetro a ser estimado, da seguinte forma:

$$x = H\theta + w \quad (3)$$

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

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

Sendo a matriz H conhecida como matriz de observação e o vetor de ruído caracterizado como $w \sim N(0, \sigma^2 I)$. O estimador MVU é determinado através da restrição de igualdade do limite inferior de Cramer-Rao CRLB (em inglês, Cramer-Rao Lower Bound) (Kay, 1993). Pelo Teorema $\hat{\theta} = g(x)$ será o estimador MVU de θ se,

$$\frac{\partial \ln p(x; \theta)}{\partial \theta} = I(\theta)(g(x) - \theta). \quad (4)$$

Derivando o logaritmo neperiano da função de verossimilhança do modelo linear e assumindo que $H^T H$ é inversível, chegamos a seguinte equação:

$$\frac{\partial \ln p(x; \theta)}{\partial \theta} = \frac{H^T H}{\sigma^2} [(H^T H)^{-1} H^T x - \theta] \quad (5)$$

que possui exatamente a forma da Equação 4, de onde obtém-se o estimador MVU do modelo linear e sua variância.

$$\hat{\theta} = (H^T H)^{-1} H^T x \quad (6)$$

$$C_{\hat{\theta}} = \sigma^2 (H^T H)^{-1}. \quad (7)$$

3.1 Minimização de Ruído por Mínima Variância para um Modelo Linear Corrompido por Ruído Colorido

Pode-se generalizar o resultado anterior e encontrar um método para a minimização do ruído baseado no estimador MVU para um conjunto de dados corrompido por um ruído colorido. Assim, o modelo geral assume que o ruído tem a seguinte característica

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

onde C não é uma matriz identidade. O estimador é determinado através do uso de uma abordagem onde utiliza-se um branqueamento do conjunto de dados. Sendo a matriz de covariância C positiva definida, sua matriz inversa é positiva definida e pode ser fatorada utilizando a decomposição de Cholesky (Kay, 1993) da seguinte forma:

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

onde D é uma matriz inversível $N \times N$. A matriz D age como uma matriz de branqueamento quando aplicada ao ruído. Aplicando a matriz D

ao modelo linear do conjunto de dados, obtém-se a generalização do estimador MVU para ruído colorido:

$$\hat{\theta} = (H^T C^{-1} H)^{-1} H^T C^{-1} x. \quad (8)$$

3.2 Minimização de Ruído por Máxima Verossimilhança

Este método alternativo de *denoising* para o TleCal se baseia no princípio da máxima verossimilhança (MLE em inglês), representando uma das abordagens mais populares para se obter estimadores práticos e desejável em diversas situações onde o estimador MVU não pode ser encontrado ou não existe.

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

$$\mathcal{L}(\theta) = \prod_{i=1}^N f(x_i, \theta) \quad (9)$$

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

$$x[n] = \theta + w[n]$$

onde θ é o parâmetro a ser estimado e $w[n]$ é um ruído branco gaussiano com variância σ_i , sua p.d.f (função densidade de probabilidade) será:

$$f(x_i; \theta) = \frac{1}{(2\pi\sigma_i^2)} \exp\left[-\frac{(x_i - \theta)^2}{2\sigma_i^2}\right]. \quad (10)$$

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

$$\mathcal{L}(\theta) = \prod_{i=1}^N \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[-\frac{(x_i - \theta)^2}{2\sigma_i^2}\right]. \quad (11)$$

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

$$\ln p(x; \theta) = \sum_{i=1}^N \ln \frac{1}{\sqrt{2\pi}\sigma_i} - \frac{1}{2} \sum_{i=1}^N \frac{(x_i - \theta)^2}{\sigma_i^2}. \quad (12)$$

Derivando parcialmente essa função em relação ao parâmetro desconhecido, obtém-se

$$\frac{\partial \ln p(x; \theta)}{\partial \theta} = \sum_{i=1}^N \frac{x_i - \theta}{\sigma_i^2}.$$

Igualando a zero a derivada, encontra-se o valor ótimo do parâmetro conforme

$$\hat{\theta} = \frac{\sum_{i=1}^N \frac{1}{\sigma_i^2} x_i}{\sum_{i=1}^N \frac{1}{\sigma_i^2}}. \quad (13)$$

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

3.3 Estimação de Energia nas Células do TileCal

Uma vez que a estimação de energia do sinal de cada canal de leitura do TileCal é realizada de forma independente, existe uma abordagem ótima para maximizar a relação sinal-ruído da célula. A Equação 14 indica que os pesos a_1 e a_2 , utilizados na soma de energia de uma célula, podem ser diferentes, visando minimizar a contribuição de ruído. Para manter o patamar de energia da célula, a soma dos pesos deve ser igual a dois.

$$E_{cell_{opt}} = a_1 \times E_{pmt_1} + a_2 \times E_{pmt_2}. \quad (14)$$

A informação do ruído pode ser usada para encontrar o peso ótimo para cada canal. Para tal, o ruído é adquirido e as técnicas de *denoising* são aplicadas para determinar os pesos a_1 e a_2 . Vale ressaltar que o cálculo destes pesos é realizado de maneira *offline*, utilizando tomadas de dados específicas para calibração (sem colisões). Uma vez calculados os pesos ótimos para cada canal, a estimação em célula pode ser realizada *offline* ou *online*, contanto que haja disponibilidade de recursos de *hardware* na eletrônica de leitura do TileCal numa estimação *online*.

A técnica baseada em MVU leva em consideração uma possível correlação entre o ruído dos canais e pode ser generalizada para o caso onde o ruído aditivo é colorido. Com o MVU pode-se encontrar a energia ótima da célula (\widehat{E}_{cell}) utilizando a Equação 8 para o caso de $N = 2$ da seguinte maneira, onde 1 se refere uma matriz de referência [11]:

$$\widehat{E}_{cell} = 2 \frac{1^T C^{-1}}{1^T C^{-1} 1} x, \quad (15)$$

em que a constante 2 da equação possui a função de garantir o nível de energia na célula. O parâmetro 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 de cada canal podem ser obtidos da seguinte forma:

$$a_{1,2} = \frac{\sigma_{w_{2,1}}^2 - C_{w_1 w_2}}{\sigma_{w_1}^2 + \sigma_{w_2}^2 - 2C_{w_1 w_2}}. \quad (16)$$

Já a técnica baseada em MLE assume que os ruídos dos canais de leitura da célula são decorrelacionados, aditivos, brancos e Gaussianos. O MLE atribui um peso maior ao canal que possuir a menor variância de ruído (Knoll, 1979). Os pesos para cada canal serão obtidos da seguinte maneira:

$$a_j = \frac{1}{\sigma_{w_j}^2} \left(\sum_{i=1}^2 \frac{1}{\sigma_{w_i}^2} \right)^{-1} \quad j = 1, 2. \quad (17)$$

4 Resultados

Nos resultados a seguir, a estimação de sinal realizada utilizando as técnicas de *denoising* propostas estarão identificadas pelas siglas **MLE** e **MVU**, já o método de referência, com a simples soma dos canais de uma célula (pesos unitários), será identificado com a sigla **SS-OF2** (quando o método de estimação de energia for o OF2) e **SS-MF** (quando utiliza-se o método MF).

4.1 Conjuntos de Dados Utilizados

Dois conjuntos diferentes de dados foram utilizados. Inicialmente, para testar a capacidade de redução de ruído e a viabilidade de implementação dos métodos propostos, a estimativa de *denoising* é aplicada em banco de dados contendo apenas o ruído eletrônico dos canais das células. Periodicamente, esses dados são coletados com a intenção de caracterizar o ruído eletrônico dos canais do TileCal. Essa tomada de dados ocorre com o detector em pleno funcionamento, porém, na ausência de colisões de partículas, de modo a garantir que os dados sejam puramente de ruído. Neste trabalho, um conjunto de dados com aproximadamente 20.000 eventos foi utilizado para cada canal do TileCal. Foi realizada uma separação dos dados em dois conjuntos: um conjunto de teste (para analisar a generalidade da estimação de *denoising*) correspondendo a 20% dos eventos e um conjunto de treino correspondendo a 80% dos eventos, que foi utilizado para calcular os pesos através dos estimadores propostos.

O segundo conjunto de dados utilizado é composto por dados reais produzidos pelas colisões de partículas *pp* (próton-próton) no experimento ATLAS. Esses dados contêm, além do ruído do canal, o sinal de interesse, que é a energia depositada em cada célula TileCal. Foi utilizada uma tomada de dados de colisão *pp* com 290.634 eventos com energia máxima de 3,9 GeV no centro da colisão. A estimação ao nível de cada célula ocorre utilizando a amplitude do sinal estimada por um dos algoritmos de estimação existentes no TileCal

(OF2 ou MF). Foi aplicado os pesos para cada canal, estimados utilizando somente a informação do ruído.

4.2 Resultados com Dados de Ruído

A célula A9 do módulo LBA59 do TileCal foi escolhida como exemplo, pois, nesta tomada de dados o ruído de seus canais apresentaram correlação e um dos canais apresentou alta variância de ruído. Tal observação pode ser constatada na Figura 4 onde a correlação entre o ruído do canal 1 (eixo das abscissas) e o ruído do canal 2 (eixo das coordenadas). Os pesos de *denoising* de cada canal da célula são apresentados na Tabela 1.

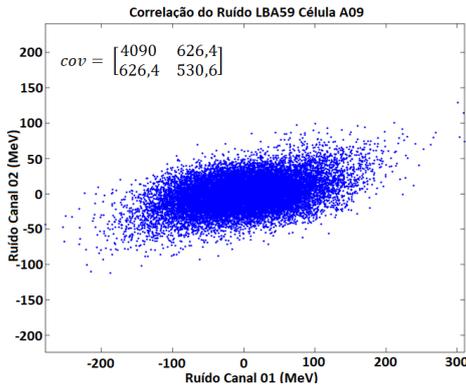


Figura 4: Correlação do ruído dos canais da célula A9 do TileCal.

Tabela 1: Pesos Ótimos estimados para os canais da célula A9.

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

Na Figura 5 pode-se observar a distribuição de energia do ruído da célula considerando a abordagem atual (soma simples não otimizada, em azul) e os métodos propostos em que os pesos associados a energia de cada canal é otimizado utilizando MVU (em vermelho) e MLE (em magenta).

Na Figura 5 apresenta-se o resultado da estimação realizada em um conjunto de dados de uma célula contendo apenas a componente de ruído eletrônico do canal. No eixo das abscissas exibe-se o valor de energia do ruído medido em MeV ($1 \text{ eV} = 1,602 \times 10^{-19} \text{ joules}$), e, no eixo das coordenadas, exibe-se a quantidade de medições (eventos) realizadas.

Como pode ser observado, devida a alta variância do ruído do canal 1 da célula, a estimação de *denoising* para este canal por ambos os métodos MVU e MLE (Tabela 1) diminuiu significativamente a contribuição deste canal na estimação

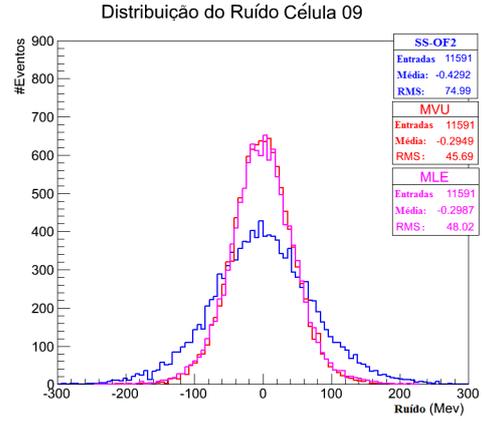


Figura 5: Distribuição de ruído antes e depois da aplicação dos métodos de *denoising* propostos na célula A9, comparando com modelo OF2 com pesos unitários.

da energia da célula, haja vista seu elevado ruído. Por outro lado, o canal 2 da mesma célula teve sua contribuição praticamente dobrada. Nesta condição, os métodos MVU e MLE mostraram-se mais eficiente em mascarar o canal ruidoso, diminuindo assim a variância do ruído total da célula, como mostrado na Figura 5.

4.3 Dados de Colisões

Na Figura 6 apresenta-se a distribuição de energia das células do TileCal em uma ampla faixa de energia utilizando diferentes abordagens. Neste resultado os métodos de *denoising* foi combinado com a estimação de amplitude por MF. Uma cauda negativa nas distribuições de energia pode ser observada. Tal cauda deve-se ao ruído eletrônico dos canais, assim como pela presença de sinais pertencentes a colisões adjacentes (empilhamento de sinais) como resultado do aumento crescente da luminosidade do LHC.

Pode-se observar que a cauda positiva estimada com *denoising* não sofre nenhuma alteração significativa se comparada a cauda do método de estimação simples SS-MF. Entretanto, há uma diminuição significativa da cauda negativa nas distribuições de energia, em relação ao método SS-OF2. Isso demonstra que houve uma diminuição do ruído na estimação de energia das células. Essa diminuição deve-se a mitigação de canais com alta variância de ruído em determinadas células, através da ponderação realizada com os métodos MVU ou MLE.

Nas Figuras 7 e 8, a correlação entre a energia das células do TileCal obtida pelo método de soma simples (eixo das coordenadas), e no eixo das abscissas a energia da célula é mostrada após o uso dos métodos MLE e MVU respectivamente, apresenta-se também o resultado da regressão linear (parâmetros p_0 e p_1) realizada no conjunto

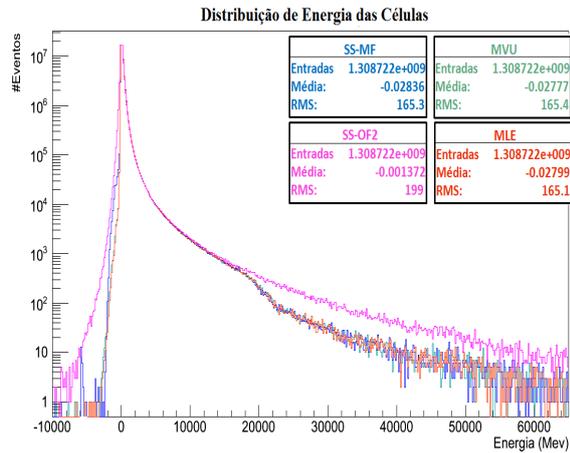


Figura 6: Distribuições de energia das células do TileCal resultado de colisões entre prótons.

de dados. Pode-se observar novamente em ambos os casos a redução da quantidade de células com energia negativa, e comparando as correlações e os parâmetros da regressão linear, observa-se uma dispersão menor ao utilizar o método MLE.

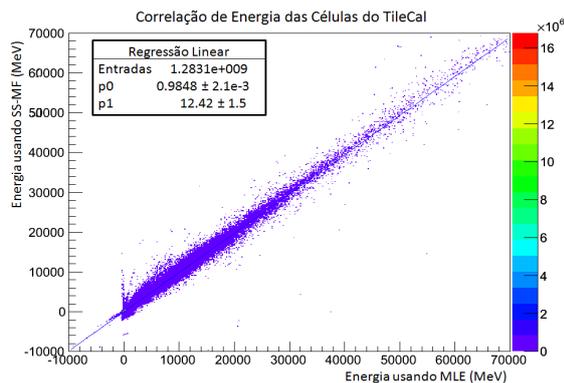


Figura 7: Correlação entre a energia das células usando o método de soma simples e usando método de *denoising* MLE.

Nas figuras 9 e 10 observa-se as distribuições dos pesos para os canais das células utilizados na estimação de *denoising* pelos estimadores MVU e MLE. Vale ressaltar que o valor dos pesos para SS-OF2 e SS-MF é igual a um. Pode-se observar que para ambos os enfoques, alguns canais receberam pesos muito próximos a 2 enquanto outros canais receberam pesos praticamente iguais a zero, mostrando assim que alguns canais foram, praticamente, excluídos na estimação da energia na célula. Adicionalmente, nota-se que, em média, os enfoques MVU e MLE para o *denoising* de canais produzem pesos similares para os dois canais de leitura, 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 di-

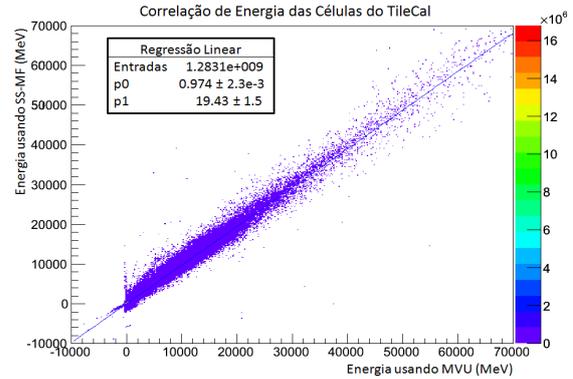


Figura 8: Correlação entre a energia das células usando o método de soma simples e usando método de *denoising* MVU.

ferença indica que algumas células apresentaram uma correlação de ruído relevante entre o ruído dos canais.

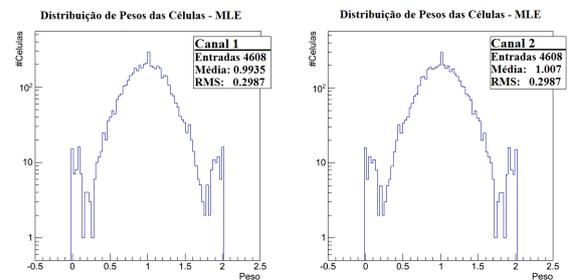


Figura 9: Distribuição dos pesos calculados pelo algoritmo MLE para as células do TileCal.

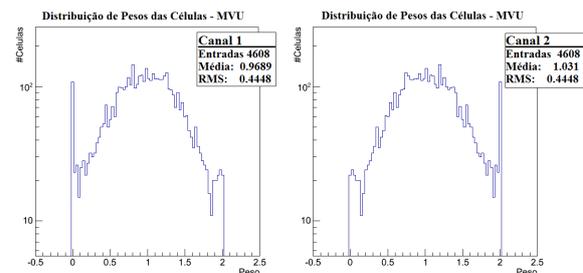


Figura 10: Distribuição dos pesos calculados pelo algoritmo MVU para as células do TileCal.

5 Conclusões

Este artigo apresentou um método de estimação de energia que utiliza uma estimação de *denoising* para canais que compõem cada célula de leitura de um calorímetro finamente segmentado, utilizando estimadores de mínima variância para otimizar a estimação de energia nas células. O método foi testado em dados de colisão e os resultados mostram que o uso de técnicas para *denoising* é capaz

de mitigar o ruído de canal, ao atribuir pesos ótimos para cada canal de leitura. Aplicações em qualidade de dados, para eliminar canais ruidosos da reconstrução dos eventos de colisão, também são possíveis com as técnicas de *denoising* aqui propostas.

Agradecimentos

Os autores gostariam de agradecer ao apoio dado pelas seguintes instituições: CNPq, CAPES, FAPERJ e RENAFEA. Os autores também agradecem ao grupo de pesquisadores do ATLAS e TileCal pelos dados e discussões frutíferas.

Referências

- Anderson K., e. (1998). Front-end electronics for the atlas tile calorimeter, *Technical report*, Proceedings of Fourth Workshop on Electronics for LHC Experiments.
- ATLAS, C. (2008). *The ATLAS Experiment at the CERN Large Hadron Collider*, Technical Report JINST 3 S08003.
- ATLAS, C. (2010). *Readiness of the ATLAS Tile Calorimeter for LHC collisions*, Technical Report EPJC 70.
- ATLAS, C. (2012). Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc, *Physics Letters B* **716**: 1–29.
- Chorus Detector* (2018). <http://www.nu.to.infn.it/exp/all/chorus/>. Acessado em Março de 2018.
- Cottingham, W. N. and Greenwood, D. A. (1998). An Introduction to the Standard Model of Particle Physics, Cambridge University Press.
- E. Fullana, e. (2006). Digital signal reconstruction in the atlas hadronic tile calorimeter, *IEEE Transaction On Nuclear Science* **53**(4).
- Experiment Baikal* (2018). <http://www.inr.ru/>. Acessado em Março de 2018.
- G. Bertuccio, E. G. and Sapietro, M. (1992). Sampling and optimum data processing of detector signals, *Nuclear Instruments and Methods in Physics Research A* **322**: 271–279.
- Kay, S. M. (1993). *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice Hall.
- Knoll, G. F. (1979). *Radiation Detection and Measurement*, New York: Wiley.
- Peralva, B. S. (2013). *The TileCal energy reconstruction for collision data using the matched filter*, Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC) .
- Trees, H. L. V. (1992). *Detection, Estimation, and Modulation Theory: Radar-Sonar Signal Processing and Gaussian Signals in Noise*, Krieger Publishing Co., Inc., Melbourne, FL, USA.
- Wigmans, R. (2000). *Calorimetry: Energy Measurements in Particle Physics*, Clarendon Press.