

Desempenho de Algoritmos de Estimação de Energia para o Calorímetro de Telhas do Experimento ATLAS [★]

Guilherme I. Gonçalves^{*} Bernardo S. Peralva^{*}
 Luciano M. A. Filho^{**} Augusto S. Cerqueira^{**}
 José M. de Seixas^{***}

^{*} Programa de Pós-graduação em Modelagem Computacional,
 Universidade do Estado do Rio de Janeiro, RJ (e-mail:
 inacio.guilherme@gmail.com, bernardo.peralva@uerj.br)

^{**} Programa de Pós-graduação em Engenharia Elétrica, Universidade
 Federal de Juiz de Fora, MG (e-mail: luciano.andrade@ufjf.br,
 augusto.santiago@ufjf.br)

^{***} Laboratório de Processamento de Sinais, Universidade Federal do
 Rio de Janeiro, RJ (e-mail: seixas@lps.ufrj.br)

Abstract: In signal estimation problems, the desired signal properties are calculated from models in order to study their characteristics. In calorimetry, for example, the particle energy is absorbed and sampled by the frontend electronics that provides a reference pulse whose amplitude is proportional to the deposited energy. However, when extremely rare physics events are envisaged in an experiment, signal pile-up may arise and in these conditions, the energy estimation performance may degrade. This work proposes a performance analysis of the methods currently available in the ATLAS Tile Calorimeter (TileCal), in the conditions of a state-of-art collider machine (LHC). Experimental data from LHC were used for energy estimation efficiency evaluation and the results show that such an efficiency varies according to the signal pile-up level observed in the readout cells.

Resumo: Em problemas de estimação de sinais, propriedades desejadas de um determinado sinal são calculadas a partir de modelos para que suas características possam ser estudadas. Em calorimetria, por exemplo, a energia da partícula é absorvida e amostrada pela eletrônica de leitura que fornece um pulso característico onde sua amplitude é proporcional à energia depositada. Entretanto, quando eventos extremamente raros são previstos, o empilhamento de sinais pode surgir e, em tais condições, o desempenho de estimação da energia pode ser degradado. Este trabalho propõe uma análise de eficiência dos métodos atualmente disponíveis no calorímetro de Telhas (TileCal) do ATLAS, nas condições encontradas em um colisor de partículas de última geração (LHC). Dados experimentais do LHC foram utilizados para avaliar a eficiência de estimação da energia e os resultados mostram que a eficiência varia de acordo com o nível de empilhamento de sinais observado nas células de leitura.

Keywords: Optimal filter, signal estimation, Wiener filtering, signal pile-up, high-energy calorimetry.

Palavras-chaves: Filtro ótimo, estimação de sinais, filtragem de Wiener, empilhamento de sinais, calorimetria de altas energias.

1. INTRODUÇÃO

Os grandes avanços nas áreas de instrumentação e experimentação trazem para ciência muitos benefícios, possibilitando a exploração de novos temas, trazendo novos desafios que instigam a curiosidade e motivam o progresso da pesquisa científica. Um dos campos de pesquisa que merece destaque é a área de estimação de sinais, onde são empregados métodos computacionais para inferir características

de modelos matemáticos, possibilitando reconstruir ou prever eventos e também proporcionar informações relevantes para tomada de decisões (Kay (1993); Van Trees (2013)).

Com uma ampla aplicabilidade, o problema de estimação de sinais em geral tem contribuído para diversas áreas como economia, biologia, química, telecomunicações, processamento de sinais, entre outros (Manoj Kumar Srivastava (2014)). Em física de altas energias, por exemplo, estimadores são projetados para a estimação de parâmetros como o momento e a energia de partículas geradas por colisões subatômicas (Wigmans (2017)).

^{*} O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

A instrumentação para experimentos em física de altas energias lida com enormes desafios tecnológicos, principalmente na área de processamento de sinais, haja vista que uma grande variedade de sensores operam em taxas cada vez maiores, resultando em uma quantidade cada vez maior de dados a serem processados em um curto intervalo de tempo. Neste contexto, experimentos modernos em física de altas energias se apoiam fortemente em seus sistemas de calorimetria, cuja principal função é absorver e amostrar a energia de partículas incidentes, sendo sua informação utilizada para a reconstrução de eventos e identificação de diferentes partículas (Wigmans (2017)).

Normalmente, um calorímetro é segmentado em inúmeras células de leitura (dezenas de milhares em calorímetros modernos), proporcionando resolução espacial ao detector. Os sinais produzidos por cada célula de leitura do calorímetro são utilizados para estimar a energia de partículas incidentes. Tipicamente, estes sinais são processados por um circuito eletrônico de conformação do pulso analógico, o que confere ao pulso uma forma fixa com uma amplitude proporcional à energia da partícula (Knoll (2010)). Desta forma, o problema de estimação de energia se dá pela determinação da amplitude do pulso produzido pelas células do calorímetro. A energia é a informação mais importante em sistemas de calorimetria de altas energias, sendo uma propriedade fundamental para a reconstrução de eventos e validação de modelos para fenômenos físicos.

Os métodos matemáticos para estimação de energia tipicamente empregados formulam o problema como sendo de estimação da amplitude de um pulso imerso em um ruído aditivo, onde a identificação dos canais com informação relevante é realizada através de um corte de energia na saída do estimador. Normalmente, na modelagem do problema adotam-se hipóteses que consideram uma forma fixa do pulso e o ruído aditivo como sendo Gaussiano (Fullana et al. (2006); Adzic et al. (2006)).

Entretanto, os experimentos que operam com uma alta taxa de eventos e buscam eventos raros operam em alta luminosidade (Herr and Muratori (2006)), onde o feixe de partículas é colimado de modo que um número maior de interações possa ocorrer a cada colisão, gerando mais sinais nos sistemas de leitura, como os calorímetros (Marshall (2014)). Conseqüentemente, o problema de estimação de energia apresenta uma complexidade maior devido ao efeito de empilhamento de sinais que pode ocorrer em condições de alta luminosidade, fenômeno em que dois ou mais pulsos ocorrem numa mesma janela de aquisição. Trabalhos recentes apresentam abordagens eficientes, baseadas na desconvolução dos sinais empilhados (Andrade et al. (2015); Barbosa et al. (2017)).

Alternativamente, uma abordagem baseada na filtragem de Wiener também tem sido testada para condições severas de empilhamento de sinais (Gonçalves et al. (2018); Oliveira Goncalves (2019)). Portanto, o objetivo deste trabalho é avaliar o desempenho dos métodos atualmente disponíveis para reconstrução da energia dos canais de leitura do Calorímetro de Telhas (TileCal, do inglês, *Tile Calorimeter*) (Collaboration (2010)) do ATLAS (do inglês, *A Toroidal LHC Apparatus*) (Collaboration (2008)), no contexto do maior acelerador de partículas do mundo, o LHC (Evans and Bryant (2008)).

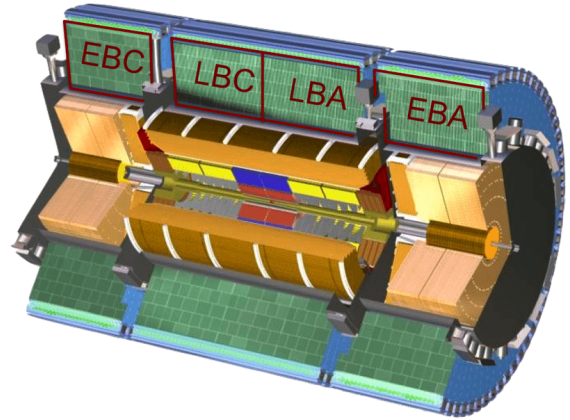


Fig. 1. Ilustração das partições do TileCal no ATLAS.

Na próxima seção o TileCal é brevemente apresentado. Na Seção 3 os métodos utilizados no TileCal para a estimação da energia são descritos. Os resultados descrevendo a análise de eficiência dos métodos utilizando dados reais adquiridos no ATLAS são apresentados na Seção 4. Por fim, na Seção 5 as conclusões deste trabalho são derivadas.

2. O TILECAL

Calorímetros de altas energias medem a energia que uma partícula perde quando esta interage com o detector. Geralmente, um calorímetro é projetado para detectar e absorver a maioria das partículas provenientes de uma colisão, forçando-as a depositar toda a sua energia no interior do detector (Wigmans (2017)).

O experimento ATLAS cobre um amplo espectro da física de interesse no LHC, e depende fortemente da calorimetria para o seu sistema de filtragem, o qual reduz o ruído de fundo que mascara os eventos de interesse, e para estimar a energia da partícula incidente. Através das informações adquiridas pelos calorímetros, é possível determinar propriedades importantes a respeito de uma partícula, o que impacta diretamente na identificação de diferentes partículas (Wigmans (2017)). O sistema de calorimetria é tipicamente seccionado de acordo com a interação: eletromagnética (elétrons e fótons) e hadrônica (prótons e nêutrons). No ATLAS, o TileCal é o principal calorímetro hadrônico do ATLAS, e fornece medidas precisas da energia de jatos, auxilia no cálculo do momento transversal faltante e assiste na detecção de múons. Objeto de estudo deste trabalho, o TileCal é composto por quatro partições (EBC, LBC, LBA e EBA) em que cada partição é dividida em 64 módulos com 32 a 48 canais de leitura por módulo, totalizando aproximadamente 10.000 canais de leitura (ver Fig. 1).

O princípio de funcionamento é baseado na técnica de amostragem que utiliza placas (ou telhas) de plástico cintilante como material ativo, intercalada com camadas de aço como material absorvente, conforme a ilustração de um módulo mostrada na Fig. 2. Ambos os lados de cada telha são conectados a um tipo específico de fibra ótica, chamada WLS (do inglês *Wavelength Shifting*), composta por um material foto-fluorescente que absorve um fóton de alta frequência e emite múltiplos fótons de baixa

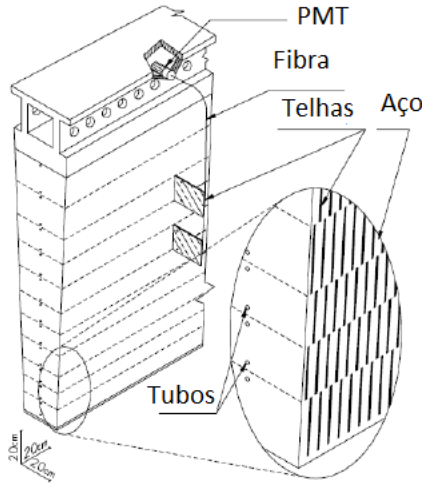


Fig. 2. Estrutura de um módulo do TileCal.

frequência. Múltiplas telhas são agrupadas em células e cada célula é conectada a duas fotomultiplicadoras, ou PMTs (do inglês *photo-multiplier tubes*).

Quando uma partícula atravessa o calorímetro, esta gera um chuva de partículas e perde energia ao atravessar o material passivo (aço como material absorvedor) e interage com o material ativo (telhas cintiladoras), emitindo luz. Essa luz é conduzida pelas fibras óticas e lidas por PMTs, que geram um pulso elétrico em resposta ao sinal luminoso. O pulso gerado pela PMT é condicionado por um circuito de conformação (*shaper*), o qual fornece um pulso com um formato conhecido e com uma amplitude proporcional à energia depositada (Anderson et al. (2005)). Este pulso analógico é digitalizado por um conversor Analógico-Digital (ADC) com frequência de amostragem de 40 MHz e uma janela com 7 amostras digitais (150 ns) cobre aproximadamente todo o pulso.

A energia depositada em cada célula do calorímetro pode ser calculada através da estimação correta da amplitude do pulso de resposta, sendo esta uma abordagem comumente observada nos calorímetros modernos. Fisicamente, apenas as células situadas na trajetória de uma partícula absorvem sua energia, portanto, devido à grande quantidade de canais, apenas as células com informação relevante são selecionadas e utilizadas na reconstrução de energia (Pastore (2016)).

Entretanto, uma vez que a física de interesse é rara, uma grande quantidade de eventos é necessária para se obter estatística suficiente para inferir algum comportamento físico. Atualmente, o LHC está passando por uma grande atualização, denominada *Phase I Upgrade*, que deve aumentar a energia de colisão e o pico de luminosidade durante o próximo período de coleta de dados que ocorrerá entre os anos de 2021 e 2023 (*Run 3*) (Schmidt (2016)). A luminosidade é definida como um fator proporcional entre o número de eventos por segundo e a área do feixe de partículas, tendo unidade de $cm^2 s^{-1}$ (Herr and Muratori (2006)). O aumento da luminosidade significa que o diâmetro do feixe de prótons será maior, resultando em um feixe mais denso, o que aumenta o número de interações entre partículas por segundo dividido pela seção de choque do feixe. Com o aumento da luminosidade, mais interações

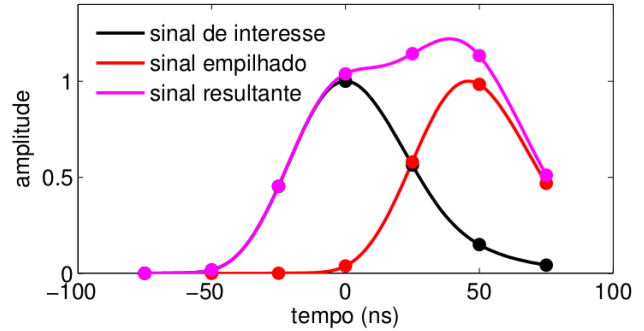


Fig. 3. Exemplo ilustrativo do efeito de empilhamento de sinais. Um sinal de interesse (preto) é adquirido e antes que a eletrônica produza o pulso, um outro sinal (vermelho) de uma colisão adjacente é observado, distorcendo o sinal resultante (magenta).

próton-próton ocorrerão quando os feixes de cruzam, ou seja, mais colisões serão geradas, produzindo mais dados e consequentemente aumentando a probabilidade de ocorrer eventos físicos de interesse. O número médio de interações por colisões $\langle \mu \rangle$ é utilizado para representar a quantidade de sinais que é gerado a cada colisão. Atualmente o LHC opera com $\langle \mu \rangle$ aproximadamente igual a 40, e este valor deve atingir 200 para o *Run 3*.

Os aumentos da taxa de eventos e luminosidade introduzem novos desafios para a eletrônica e para a área de processamento de sinais. Como consequência do aumento da luminosidade, a probabilidade da ocorrência de eventos em um mesmo canal de leitura do sistema de calorimetria do ATLAS também aumenta, acarretando o efeito de empilhamento de sinais, ilustrado pela Fig. 3. Este efeito ocorre quando a resposta do calorímetro é mais lenta que o intervalo entre as colisões (Andrade et al. (2015)), como é o caso do TileCal no ATLAS.

3. ESTIMAÇÃO DA ENERGIA NO TILECAL

O problema de reconstrução de energia em calorimetria é comumente abordado como um problema de estimação de parâmetros. Atualmente, o TileCal possui três métodos disponíveis para reconstrução da energia dos sinais provenientes das colisões: o método do filtro ótimo, o método COF, e o filtro de Wiener. Nesta seção, os métodos são descritos, destacando-se suas particularidades.

3.1 Filtro ótimo

O método do Filtro Ótimo (OF) busca minimizar a variância do ruído utilizando o conhecimento da forma do pulso (saída do circuito *shaper*) para estimar a amplitude do sinal, que é proporcional à energia. Como o ruído presente nos calorímetros do ATLAS é predominantemente eletrônico, este pode ser modelado por uma distribuição Gaussiana, o que faz este método ser amplamente utilizado neste experimento (Bertuccio et al. (1992); Cleland and Stern (1994); Delmastro (2003)).

A versão deste estimador implementada no TileCal é chamada de OF2 e está em operação desde 2014, sendo utilizado para para reconstrução *online* e *offline* de energia (Fullana et al. (2006)). Este método também é utilizado

em outros calorímetros, como no calorímetro de argônio líquido (LAr, do inglês *Liquid Argon Calorimeter*) (Cleland and Stern (1994)).

No método OF2 a amostra digital \mathbf{x} no instante k pode ser modelada por

$$x[k] = Ag[k - \tau] + n[k] + ped, \quad (1)$$

onde A é a amplitude do pulso, $g[k]$ o conjunto de valores do pulso de referência (*shaper*), $n[k]$ o ruído de fundo, τ o desvio de fase do sinal e ped o pedestal ou linha de base, um parâmetro constante adicionado ao sinal analógico antes de sua digitalização.

A amplitude do sinal de entrada é estimada através de um processo de baixo custo computacional, que utiliza uma operação de soma ponderada dada por

$$\hat{A}_{OF} = \sum_{k=0}^{N-1} x[k]w[k], \quad (2)$$

onde $w[k]$ são os coeficientes do filtro e $x[k]$ as amostras do sinal recebido.

Para calcular os coeficientes $w[k]$ do estimador OF2, o sinal digital recebido do TileCal pode ser aproximado por uma série de Taylor de primeira ordem, dada por

$$x[k] = Ag[k] - A\tau\dot{g}[k] + n[k] + ped, \quad (3)$$

onde $\dot{g}[k]$ representa a derivada do pulso de referência g , e $k = 0, 1, 2, \dots, N-1$.

Para garantir a imparcialidade do estimador, ou seja, a ausência de tendências na estimação, é exigido que o valor esperado de \hat{A}_{OF} seja igual a A . Logo, substituindo a Equação (3) em (2) e considerando que a média do ruído é zero ($E\{n[k]\} = 0$), tem-se

$$E\{\hat{A}_{OF}\} = \sum_{k=0}^{N-1} (Aw[k]g[k] - A\tau w[k]\dot{g}[k] + w[k]ped), \quad (4)$$

e, portanto,

$$E\{\hat{A}_{OF}\} = A. \quad (5)$$

Para que o estimador seja independente do pedestal e da fase, as seguintes restrições são estabelecidas

$$\sum_{k=0}^{N-1} w[k]g[k] = 1, \quad (6)$$

$$\sum_{k=0}^{N-1} w[k]\dot{g}[k] = 0, \quad (7)$$

$$\sum_{k=0}^{N-1} w[k] = 0. \quad (8)$$

A primeira restrição garante a imparcialidade do estimador, enquanto a segunda e terceira restrições garantem, respectivamente, que o estimador seja imune à fase e à flutuações da linha de base, ou pedestal.

A variância do estimador é dada por

$$E\{(\hat{A}_{OF} - A)^2\} = \sum_{k=0}^{N-1} \sum_{j=0}^{N-1} w[k]w[j]C[k, j] = \mathbf{w}^T \mathbf{C} \mathbf{w}, \quad (9)$$

onde \mathbf{w} é o vetor de pesos do estimador e \mathbf{C} a matriz de covariância do ruído.

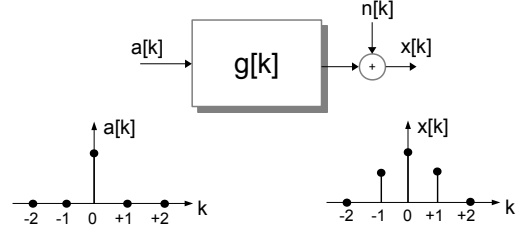


Fig. 4. Modelagem de um sistema linear para o problema de empilhamento.

Para determinar os pesos \mathbf{w} , é necessário minimizar a expressão da variância do estimador utilizando o método dos multiplicadores de Lagrange. Vale ressaltar que a matriz de covariância \mathbf{C} pode ser aproximada pela matriz identidade, caso o ruído seja modelado como sendo Gaussiano e decorrelacionado. A solução deste sistema resulta no conjunto de pesos $w[k]$ do estimador OF2 que operam, atualmente, de forma *online* e *offline* no TileCal.

3.2 O método COF

Um outro algoritmo também disponível para reconstrução *offline* no TileCal é o método COF (do inglês, *Constrained Optimal Filter*). Este estimador calcula uma transformação linear que recupera a amplitude dos sinais somados para uma determinada janela de leitura, para que o pulso central seja atribuído à colisão de interesse e possa ser dissociado e reconstruído (Andrade et al. (2015)).

Para tal, o método COF modela a deposição de energia em uma determinada célula do calorímetro como uma função delta de Kronecker (Oppenheim (1999)). Desta forma, quando aplicada a entrada de um sistema linear e invariante no tempo, a resposta deve ser o pulso de referência do TileCal, conforme mostrado na Fig. 4 (Andrade et al. (2015)).

Ou seja, considerando um conjunto $a[k]$ de deposições de energia, o sinal recebido pelo TileCal é dado por

$$x[k] = \sum_i (g[i]a[n-i]) + n[k], \quad (10)$$

em que $g[k]$ corresponde ao pulso de referência e $n[k]$ é o ruído. Desta forma, a estimação da energia depositada em uma determinada célula do calorímetro implica em desconvoluir a sequência $x[k]$ da resposta ao impulso $g[k]$. Aplicando um procedimento similar ao método OF, a amplitude pode ser dada por

$$\hat{\mathbf{a}}_j = \mathbf{U}_j^T \mathbf{x} \quad (11)$$

em que

$$\mathbf{U}_j = \mathbf{C}_j^{-1} \mathbf{G}_j (\mathbf{G}_j^T \mathbf{C}_j^{-1} \mathbf{G}_j)^{-1}. \quad (12)$$

A matriz \mathbf{G}_j é dada por

$$\mathbf{G}_j = \begin{pmatrix} g[3] & g[4] & g[5] & g[6] & 0 & 0 & 0 \\ g[2] & g[3] & g[4] & g[5] & g[6] & 0 & 0 \\ g[1] & g[2] & g[3] & g[4] & g[5] & g[6] & 0 \\ g[0] & g[1] & g[2] & g[3] & g[4] & g[5] & g[6] \\ 0 & g[0] & g[1] & g[2] & g[3] & g[4] & g[5] \\ 0 & 0 & g[0] & g[1] & g[2] & g[3] & g[4] \\ 0 & 0 & 0 & g[0] & g[1] & g[2] & g[3] \end{pmatrix}, \quad (13)$$

onde j é o número de colisões presentes na janela de leitura do calorímetro, e \mathbf{C}_j é a matriz de covariância. Quando

$j = N$, ou seja, quando o número de colisões é igual ao tamanho da janela de leitura, tem-se a melhor aproximação da desconvolução. Assim, $n[k]$ é composto pelo ruído eletrônico (considerado Gaussiano e descorrelacionado) e o estimador assumirá a forma

$$\hat{\mathbf{a}} = \mathbf{G}_j^{-1} \mathbf{x} \quad (14)$$

É importante ressaltar que é necessário subtrair o valor do pedestal do vetor \mathbf{x} antes de aplicar a Equação (14). O valor do pedestal pode ser obtido, por exemplo, através de um banco de dados que possua valores armazenados a partir de tomadas de dados de calibração do sistema de aquisição de dados. Também é válido notar que esta expressão não depende da matriz de covariância do ruído \mathbf{C} , sendo esta uma das vantagens do método COF com relação ao método OF.

Por fim, de posse do vetor de amplitudes estimadas $\hat{\mathbf{a}}$, o método COF aplica um corte linear para selecionar somente as amplitudes acima de um determinado patamar, definido no projeto do filtro. Esta etapa tem o objetivo de projetar novamente o COF através da Equação (12) evitando estimar sinais sem informação (ou seja, apenas ruído), e melhorando as estimações das amplitudes com informação relevante. Desta forma, o método COF se apresenta como uma melhor opção ao método OF, tanto na independência de informação sobre o empilhamento quanto menor erro de reconstrução de energia (Andrade et al. (2015)).

3.3 Filtragem de Wiener

Uma terceira abordagem recentemente implementada e validada para estimação da energia nos canais de leitura do TileCal se baseia na filtragem de Wiener. Neste método, busca-se projetar um filtro linear digital $c[0], c[1], \dots, c[N-1]$ em que a saída $y[N]$ fornece uma estimativa da resposta desejada $d[n]$ (adquirida através de simulação), dado um sinal de entrada com N elementos $x[0], x[1], \dots, x[N-1]$. O projeto do filtro de Wiener busca minimizar o valor médio quadrático do erro de estimação, o que leva a um problema matematicamente mais tratável. Diferentemente do método OF, o filtro de Wiener considera as incertezas provenientes do sinal e ruído em seu processo de minimização. Em particular, este critério baseado no erro médio quadrático resulta em uma dependência de segunda ordem da função de custo sobre os coeficientes do filtro. Além disso, a função de custo tem um mínimo global distinto que define unicamente o projeto ótimo do filtro, no sentido estatístico (Haykin (2013)). Nesta abordagem, a saída do filtro é dada pelo somatório

$$y[n] = \sum_{k=0}^{N-1} c[k]x[n-k], \quad (15)$$

tal que o erro entre o valor desejado e o valor estimado $e[n] = d[n] - y[n]$ seja minimizado. Para otimizar o projeto do filtro, foi adotado o critério de se minimizar o erro médio quadrático. Para isso, define-se a seguinte função de custo

$$\mathbf{J} = E\{e[n]^2\}, \quad (16)$$

onde E representa o operador esperança.

Para que \mathbf{J} seja minimizado, é necessário calcular sua derivada em função dos coeficientes $c[k]$ e, para que a

função de custo \mathbf{J} atinja o seu mínimo, sua derivada deve ser igual a zero. Logo,

$$\sum_{i=0}^{N-1} c[i]E\{x[n-k]x[n-i]\} = E\{x[n-k]d[n]\}, \quad (17)$$

onde $k = 0, 1, \dots, N-1$. Da Equação (17), pode-se observar que:

- (1) O valor esperado $E\{x[n-k]x[n-i]\}$ é a função autocorrelação da entrada do filtro para o atraso $i-k$. Pode-se reescrever esta expressão da forma

$$R[i, k] = \frac{1}{N} \sum_{n=0}^{N-1} x[n-k]x[n-i]. \quad (18)$$

- (2) O valor esperado $E\{x[n-k]d[n]\}$ é a correlação cruzada entre a entrada do filtro e a saída desejada para o atraso $i-k$. Pode-se reescrever esta expressão da forma

$$p[k] = \frac{1}{N} \sum_{n=0}^{N-1} x[n-k]d[n]. \quad (19)$$

As Equações (18) e (19) são conhecidas como equações de Wiener-Hopf. Substituindo estas equações na Equação (17), obtém-se um sistema de equações lineares como condição necessária e suficiente para otimizar o filtro, como segue

$$\sum_{i=0}^{N-1} c[i]R[i, k] = p[k] \quad k = 0, 1, \dots, N-1. \quad (20)$$

Pode-se reescrever este sistema de equações na forma matricial

$$\mathbf{cR} = \mathbf{p}. \quad (21)$$

Finalmente, os pesos ótimos do filtro de Wiener podem ser expressos por

$$\mathbf{c} = \mathbf{R}^{-1}\mathbf{p}, \quad (22)$$

onde \mathbf{R} representa a matriz de autocorrelação das amostras temporais dos sinais de entrada (Equação (18)) e \mathbf{p} representa a matriz de correlação cruzada entre as amostras temporais dos sinais de entrada e os valores desejados para a saída do filtro (Equação (19)).

Vale ressaltar que o filtro de Wiener resulta no filtro ótimo no sentido da minimização do erro médio quadrático (dispersão do erro), levando-se em conta a estatística presente nos dados de entrada (sinal mais ruído). Entretanto, o filtro de Wiener considera que o ruído possui média zero, o que não corresponde ao caso do ruído de empilhamento de sinais unipolares. Para contornar o problema da média do ruído, este trabalho propõe a inclusão de um coeficiente adicional no processo de otimização do filtro de Wiener. Na presente proposta, um elemento adicional de valor constante igual a 1 (um) é adicionado a cada sinal de entrada como último elemento. Desta forma, o sinal de entrada possui $N+1$ elementos e o vetor de coeficientes também é acrescido de um elemento.

O objetivo da inclusão do elemento adicional é anular a componente independente no sinal no procedimento de otimização, absorvendo o valor médio do ruído a fim de compensar tal contribuição na medida da amplitude de uma dada janela de leitura. Como resultado, a estimação da amplitude \hat{A}_{FW} do filtro de Wiener proposto é dada pela soma dos produtos das amostras temporais do sinal

recebido e os N primeiros coeficientes de \mathbf{c} . Ao fim da operação, o último coeficiente $c[N]$ é somado ao resultado, compensando o valor médio do ruído conforme mostrado na Equação (23).

$$\hat{A}_{FW} = \left(\sum_{i=0}^{N-1} c[i]x[i] \right) + c[N] \quad (23)$$

Vale mencionar que a filtragem de Wiener tem natureza linear, baseado na operação de soma-produto similar aos métodos OF, atualmente empregados em calorimetria no ATLAS. Portanto, este método pode ser empregado para estimação da energia de forma *online* e *offline*.

4. RESULTADOS

Dados reais de colisão próton-próton adquiridos em 2018 pelo LHC (último período de tomada de dados) foram utilizados para análise de eficiência dos métodos descritos (Martínez (2016)). Para observar o comportamento do filtro para diferentes intensidades de empilhamento de sinais, foram utilizados conjuntos de dados de baixa luminosidade, com número médio de interações por colisão de $\langle \mu \rangle \approx 30$, $\langle \mu \rangle \approx 40$ e $\langle \mu \rangle \approx 50$, e um conjunto com $\langle \mu \rangle \approx 90$, o qual representa faixas de alta luminosidade previstas para os próximos anos de operação do LHC. Cada conjunto contém cerca de 1 milhão de eventos. O conjunto $\langle \mu \rangle \approx 40$ foi utilizado para projetar os estimadores e aplicá-los nos conjuntos de $\langle \mu \rangle \approx 30$ e $\langle \mu \rangle \approx 50$ visto que as condições de ruído são próximas, e visando aproveitar o maior número de eventos. Já para a condição de $\langle \mu \rangle \approx 90$, o conjunto foi dividido igualmente em dois subconjuntos, de desenvolvimento (utilizado para projetar os estimadores), e de teste (utilizado para avaliar a eficiência dos mesmos).

Vale ressaltar que os dados utilizados foram obtidos de tomadas de dados chamadas *ZeroBias stream*, onde somente *triggers* aleatórios são utilizados. Portanto, apenas informação de ruído eletrônico e de empilhamento é adquirida. Em outras palavras, nestes eventos não é esperado observar nenhum sinal de interesse na janela de aquisição, constituindo apenas dados de ruído. Portanto, o valor da média e RMS da estimação da energia destes eventos representam, respectivamente, a tendência, ou viés, de estimação e variância associado a cada algoritmo.

O filtro de Wiener é projetado a partir de um conjunto de dados composto por sinais de interesse imersos em ruído, bem como os valores conhecidos de amplitude dos respectivos sinais. Portanto, um simulador foi desenvolvido e validado para produzir um conjunto de dados de sinais de interesse com o valor das amplitudes conhecidas. Para simular deformações associadas à tolerância dos componentes eletrônicos de condicionamento do pulso, cada amostra do sinal é somada a um valor aleatório modelado por uma distribuição Gaussiana de média zero e desvio padrão de 1% sobre o valor da amostra. Um desvio de fase modelado por uma distribuição uniforme entre $[-1, +1]$ ns também foi introduzido no sinal de interesse, simulando incertezas da fase na aquisição (Fullana et al. (2006)). A amplitude segue uma distribuição uniforme no intervalo de $[0; 1023]$ contagens de ADC, visto que o conversor analógico-digital

dos canais de leitura do TileCal possui 10 bits (Anderson et al. (2005)). Desta forma, cada valor de amplitude tem a mesma probabilidade de ocorrência. Por fim, os sinais gerados foram somados aos eventos do conjunto utilizados para projetar os coeficientes do filtro de Wiener.

4.1 Análise de eficiência

Para analisar a eficiência dos filtros para condições severas de empilhamento de sinais, as células mais afetadas pelo efeito de empilhamento no TileCal foram utilizadas. A Fig. 5 mostra as distribuições de energia reconstruídas pelo filtro de Wiener, COF e OF2, considerando os diferentes valores de $\langle \mu \rangle$. Pode-se observar que o histograma para o filtro de Wiener apresenta menor dispersão seguido pelo método COF. Este resultado mostra que para canais com alto empilhamento o filtro de Wiener apresenta desempenhos promissores.

A Fig. 6 mostra a evolução da média e do desvio padrão das distribuições mostradas na Fig. 5. Novamente, o filtro de Wiener apresenta menor média e RMS comparado aos filtros COF e OF2 nas células de leitura de maior ocupação. Por exemplo, para a condição de $\langle \mu \rangle = 90$, o filtro de Wiener introduz uma melhora de aproximadamente 35% e 20% em relação aos filtros COF e OF2, respectivamente.

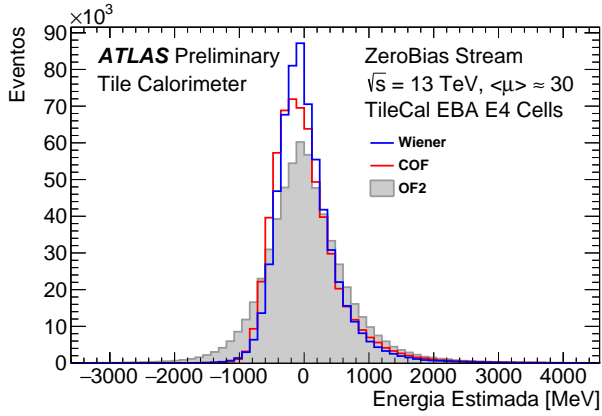
A fim de verificar a eficiência de estimação em outros canais do TileCal, um módulo completo foi utilizado (Módulo 01 da partição EBA). Nesta análise, a eficiência dos três métodos são comparadas. A Fig. 7a mostra diferença da média da distribuição de energia estimada entre o filtro de Wiener e COF, adotando o método OF2 como referência. Valores positivos representam valores de média superiores ao OF2 e valores negativos representam valores menores. Já a Fig. 7b mostra a diferença relativa percentual do desvio padrão da distribuição de energia dos métodos, adotando o OF2 como referência e considerando um módulo apenas.

Mais uma vez, nota-se que as melhorias mais significativas obtidas pelo uso do filtro de Wiener em respeito ao COF e OF2, são visíveis para os canais 0 e 1 (células de maior ocupação no TileCal), células severamente afetadas pelo efeito de empilhamento de sinais para o nível de luminosidade $\langle \mu \rangle \approx 90$. Para o restante dos canais, o filtro COF se mostrou mais eficiente, superando o filtro de Wiener e OF2, melhorando o RMS em aproximadamente 25% com respeito ao método OF2.

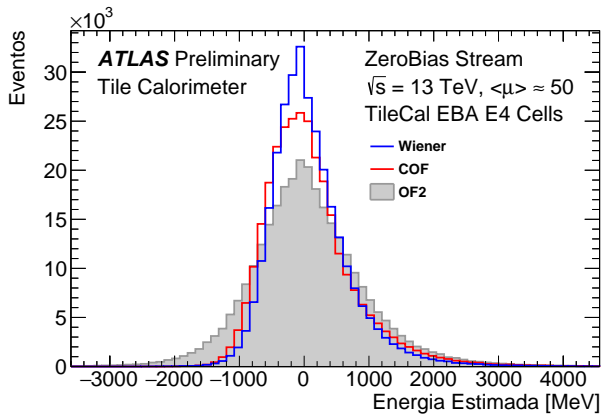
5. CONCLUSÃO

Este trabalho abordou o problema de estimação de parâmetros aplicado a reconstrução de energia no calorímetro de Telhas (TileCal) do ATLAS, no LHC, operando em condições severas de empilhamento de sinais. Calorímetros modernos operando em alta taxa de eventos são afetados pelo efeito de empilhamento de sinais. Este fenômeno introduz novos desafios para a estimação de energia, uma vez que métodos determinísticos que consideram uma forma fixa do pulso proveniente do calorímetro imerso em ruído Gaussiano têm suas eficiências degradadas.

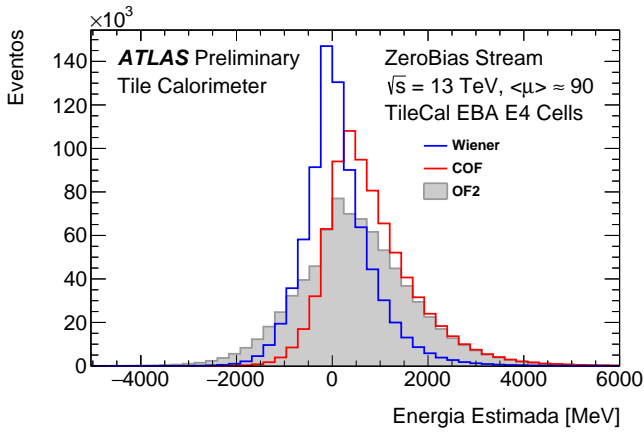
Utilizando dados reais de colisão de células severamente afetadas pelo efeito de empilhamento de sinais, observou-se um ganho em torno de 30% em resolução de estimação



(a)



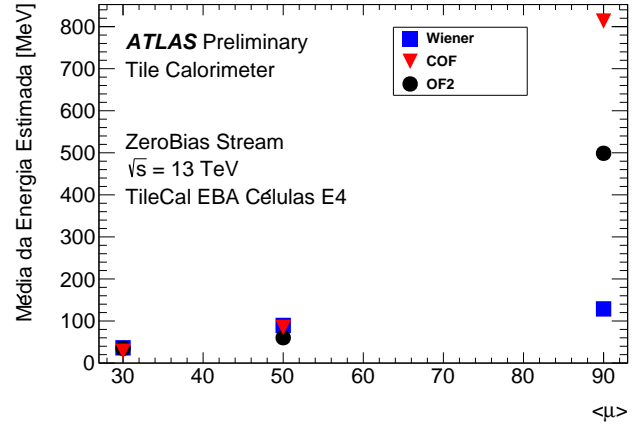
(b)



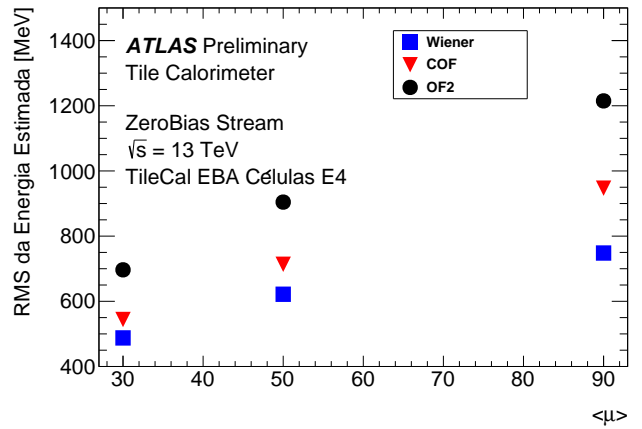
(c)

Fig. 5. Distribuição de energia reconstruída para (a) $\langle \mu \rangle \approx 30$, (b) $\langle \mu \rangle \approx 50$ e (c) $\langle \mu \rangle \approx 90$.

comparado ao método atualmente utilizado (método OF2) no TileCal. Para alta luminosidade, cenário previsto para os próximos anos de operação do LHC, o ganho alcançou cerca de 40%. Também foi observado que para faixas de baixa e média ocupação, o método COF apresenta um desempenho superior para estimação de energia, melhorando a resolução de estimação em aproximadamente 25%.



(a)



(b)

Fig. 6. Evolução da (a) média e do (b) desvio padrão das distribuições de energia reconstruídas com dados reais.

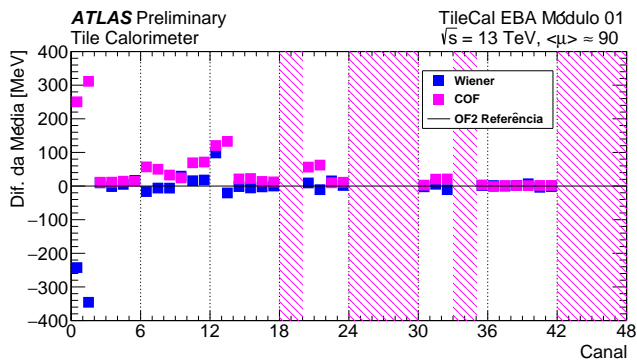
Visto que o foco deste trabalho consistiu em analisar a parcialidade e variância dos estimadores, os esforços atuais estão sendo empregados para se obter conjuntos de dados com eventos selecionados a fim de se analisar o impacto da estimação de energia nos objetos da Física, como a reconstrução de partículas.

AGRADECIMENTOS

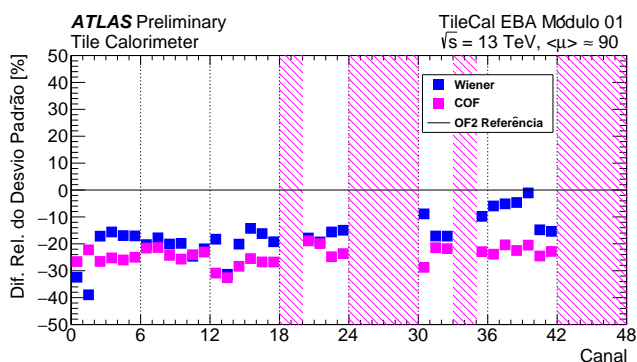
Os autores agradecem a FAPERJ, FAPEMIG, CNPq e RENAFAP pelo suporte para a realização deste trabalho. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

REFERÊNCIAS

- Adzic, P. et al. (2006). Reconstruction of the signal amplitude of the CMS electromagnetic calorimeter. *Eur. Phys. J. C*, 46S1, 23-35. doi:10.1140/epjcd/s2006-02-002-x.
- Anderson, K., Gupta, A., Merritt, F., Oreglia, M., Pilcher, J., Sanders, H., Shochet, M., Tang, F., Teuscher, R., Wu, H., et al. (2005). Design of the front-end analog



(a)



(b)

Fig. 7. Variação da (a) diferença da média e do (b) desvio relativo do desvio padrão da distribuição de energia para o filtro de Wiener e COF, adotando o método OF2 como referência. Apenas o módulo 1 da partição EBA do TileCal é considerado. Áreas hachuradas correspondem aos canais não instrumentados.

electronics for the ATLAS tile calorimeter. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 551(2-3), 469–476.

Andrade, Filho, L.M., Peralva, B.S., de Seixas, J.M., and Cerqueira, A.S. (2015). Calorimeter response deconvolution for energy estimation in high-luminosity conditions. *IEEE Transactions on Nuclear Science*, 62(6), 3265–3273.

Barbosa, D.P., d. A. Filho, L.M., Peralva, B.S., Cerqueira, A.S., and de Seixas, J.M. (2017). Sparse representation for signal reconstruction in calorimeters operating in high luminosity. *IEEE Transactions on Nuclear Science*, 64(7), 1942–1949.

Bertuccio, G., Gatti, E., Sampietro, M., Rehak, P., and Rescia, S. (1992). Sampling and optimum data processing of detector signals. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 322(2), 271–279.

Cleland, W. and Stern, E. (1994). Signal processing considerations for liquid ionization calorimeters in a high rate environment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 338(2-3), 467–497.

Collaboration, A. (2008). The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3, S08003. 437 p. doi:10.1088/1748-0221/3/08/S08003.

Collaboration, A. (2010). Readiness of the ATLAS Tile Calorimeter for LHC collisions. *Eur. Phys. J. C*, 70, 1193–1236.

Delmastro, M. (2003). A stand-alone signal reconstruction and calibration algorithm for the ATLAS electromagnetic calorimeter. In *2003 IEEE Nuclear Science Symposium. Conference Record (IEEE Cat. No. 03CH37515)*, volume 2, 1110–1114. IEEE.

Evans, L.R. and Bryant, P. (2008). LHC Machine. *JINST*, 3, S08001. 164 p. doi:10.1088/1748-0221/3/08/S08001.

Fullana, E., Castelo, J., Castillo, V., Cuenca, C., Ferrer, A., Higon, E., Iglesias, C., Munar, A., Poveda, J., Ruiz-Martinez, A., et al. (2006). Digital signal reconstruction in the ATLAS hadronic tile calorimeter. *IEEE transactions on nuclear science*, 53(4), 2139–2143.

Gonçalves, G.I., Peralva, B.S., d. A. Filho, L.M., Cerqueira, A.S., and de Seixas, J.M. (2018). Estimação de energia para o calorímetro de telhas do ATLAS baseada na filtragem de Wiener-Hopf. In *Anais do XXI Encontro Nacional de Modelagem Computacional*, 1–5.

Haykin, S.O. (2013). *Adaptive filter theory*. Pearson.

Herr, W. and Muratori, B. (2006). Concept of luminosity. doi:10.5170/CERN-2006-002.361. URL <https://cds.cern.ch/record/941318>.

Kay, S.M. (1993). *Fundamentals of statistical signal processing*. Prentice Hall signal processing series. Prentice Hall PTR, Upper Saddle River, NJ.

Knoll, G.F. (2010). *Radiation Detection and Measurement*. John Wiley & Sons.

Manoj Kumar Srivastava, Abdul Hamid Khan, N.S. (2014). *Statistical inference*. PHI Learning.

Marshall, Z. (2014). Simulation of pile-up in the ATLAS experiment. *Journal of Physics: Conference Series*, 513(2), 022024. doi:10.1088/1742-6596/513/2/022024.

Martínez, A.R. (2016). The run-2 ATLAS trigger system. *Journal of Physics: Conference Series*, 762, 012003. doi:10.1088/1742-6596/762/1/012003.

Oliveira Gonçalves, D. (2019). Energy reconstruction of the ATLAS Tile Calorimeter under high pile-up conditions using the Wiener Filter. Technical Report ATL-TILECAL-PROC-2019-002, CERN, Geneva. URL <https://cds.cern.ch/record/2674807/>.

Oppenheim, A.V. (1999). *Discrete-time signal processing*. Pearson Education India.

Pastore, F. (2016). The atlas trigger system: Past, present and future. *Nuclear and Particle Physics Proceedings*, 273-275, 1065 – 1071. doi:https://doi.org/10.1016/j.nuclphysbps.2015.09.167. 37th International Conference on High Energy Physics (ICHEP).

Schmidt, B. (2016). The high-luminosity upgrade of the LHC: Physics and technology challenges for the accelerator and the experiments. *Journal of Physics: Conference Series*, 706, 022002. doi:10.1088/1742-6596/706/2/022002.

Van Trees, Harry L; Bell, K.L. (2013). *Detection, estimation, and modulation theory, part I: detection, estimation, and linear modulation theory*. John Wiley & Sons.

Wigmans, R. (2017). *Calorimetry: Energy measurement in particle physics*. Oxford University Press.