

Estimação de Energia Assistida por Redes Neurais para o Calorímetro de Telhas do ATLAS no LHC

C. C. Dias, B. S. Peralva

Instituto Politécnico

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

dias.carlos@graduacao.uerj.br, bernardo.peralva@uerj.br

M. Hufnagel, A. M. Silva, L. M. A. Filho, A. S. Cerqueira

Programa de Pós-Graduação em Engenharia Elétrica

*Universidade Federal de Juiz de Fora
Juiz de Fora, MG, Brasil*

mateushufnagel@gmail.com, alessa.monay@engenharia.ufjf.br
luciano.andrade@ufjf.edu.br, augusto.santiago@ufjf.edu.br

J. M. de Seixas

*Laboratório de Processamento de Sinais (LPS)
Universidade Federal do Rio de Janeiro (COPPE/Poli)
Rio de Janeiro, RJ, Brasil
seixas@lps.ufrj.br*

Resumo—Em calorimetria de altas energias, a energia da partícula pode ser calculada através da estimação da amplitude do sinal produzido e digitalizado pelos canais de leitura. Entretanto, em condições de alta taxa de eventos e alta luminosidade, o sinal de resposta pode ser distorcido devido à presença de sinais proveniente de eventos adjacentes causando o efeito de empilhamento de sinais. Este efeito degrada a eficiência de métodos lineares tipicamente utilizados para estimar os parâmetros de interesse. Desta forma, este trabalho apresenta uma análise de desempenho de uma proposta de estimação da amplitude do sinal utilizando o método atualmente empregado no Calorímetro de Telhas (TileCal) do ATLAS no LHC, assistido por redes neurais artificiais. A rede neural tem a função de corrigir as não-linearidades introduzidas pelo empilhamento de sinais na estimativa linear. Os resultados, comparando diferentes métodos, mostram que a abordagem combinada aumenta consideravelmente a eficiência de estimação da energia em condições de empilhamento de sinais.

Palavras-chaves—Filtragem ótima, estimação de sinais, empilhamento de sinais, redes neurais, calorimetria de altas energias.

I. INTRODUÇÃO

Técnicas de inteligência computacional são vastamente empregadas no contexto de estimação de parâmetros em áreas como visão computacional, biometria, comunicações entre outras [1], [2]. Particularmente, devido à sua simplicidade, eficiência e resposta rápida, filtros lineares são tipicamente empregados para estimação de parâmetros de forma online [4]. Entretanto, em condições adversas de operação, em que componentes adicionais e não lineares estão presentes no ruído, o desempenho de métodos lineares é comprometido. Tipicamente, a solução deste problema envolve técnicas não-lineares para estimação de parâmetros [5], [6].

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

Em física experimental de altas energias, o uso de inteligência computacional está crescendo devido ao aumento da complexidade dos experimentos modernos e dos avanços nas áreas de processamento de sinais e hardware [7]. Dentre as ferramentas mais comuns utilizadas em física de altas energias, as Redes Neurais Artificiais (RNA) [8] podem ser encontradas em diversas aplicações tais como reconhecimento de padrões, reconstrução de traços, filtragem de dados e análises físicas [9], além de estimação de energia [10].

Neste contexto, o LHC (do inglês, *Large Hadron Collider*) [11] é o maior colisor de partículas e mais energético do mundo. No LHC, dois feixes de prótons são acelerados em sentidos opostos a um valor próximo da velocidade da luz e colidem em pontos específicos. Experimentos são construídos ao redor dos pontos de colisão e coletam a informação dos subprodutos provenientes de cada colisão. Dentre os experimentos do LHC, o ATLAS [12] é o maior e tem um extenso programa de física, sendo projetado para explorar todo o potencial de descoberta do LHC. Através das análises de dados do ATLAS, comprovações da existência do bóson de Higgs, partícula prevista pelo Modelo Padrão [13], foram observadas pelo experimento [14], [15].

A física de interesse é rara e necessita de uma quantidade imensa de dados para se poder inferir algum comportamento nos dados. Visando aumentar a estatística e, consequentemente, aumentar a probabilidade de se observar fenômenos desconhecidos, o programa do LHC prevê um aumento na luminosidade [16] dos feixes de prótons. O aumento da luminosidade significa que o feixe será mais denso e, assim sendo, mais interações próton-próton ocorrerão quando os feixes se cruzarem (colisão) [17]. Por outro lado, o sistema de calorimetria é fortemente afetado pelo efeito de empilhamento de sinais visto que a resposta produzida pelos canais de leitura dos calorímetros é mais lenta que a taxa de eventos do LHC. Consequentemente, o efeito do empilhamento de sinais poderá

ser observado numa mesma janela de leitura, dificultando a reconstrução do sinal de interesse associado a uma dada colisão.

Usualmente, a estimativa da energia é feita através da combinação linear das amostras digitais recebidas. Entretanto, a informação do empilhamento de sinais introduz não-linearidades no ruído, fazendo com que a eficiência destes métodos seja degradada. Neste sentido, trabalhos recentes [18] exploram o uso de redes neurais artificiais para assistir a estimativa linear. Neste caso, um perceptron multi-camadas é treinado para compensar o erro introduzido pela medida linear devido às componentes não-lineares provenientes do empilhamento de sinais. Desta forma, a RNA atua como um corretor não-linear, aplicando um ajuste fino na estimativa linear, mantendo-a disponível e preservada caso a correção linear não seja aplicada. Desta forma, este trabalho apresenta uma análise de eficiência do método linear atualmente em operação no Calorímetro de Telhas (TileCal) [19] do ATLAS no LHC quando o corretor não-linear, baseado em redes neurais artificiais, é utilizado para corrigir a saída linear. Diferentemente de trabalhos anteriores [18], este trabalho apresenta uma análise mais próxima das condições reais de aplicação, uma vez que utiliza dados reais adquiridos durante operação nominal do LHC. Além do método linear ser utilizado para comparação com o método combinado (OF+RNA), um terceiro método, baseado na filtragem de Wiener e introduzido recentemente no TileCal para canais com alto empilhamento de sinais [20], é também utilizado para efeitos comparativos. Vale ressaltar que o trabalho [20] apresenta uma comparação entre métodos lineares, e não discute o emprego de métodos não-lineares para a estimativa de energia no TileCal.

Na próxima seção, o TileCal é brevemente descrito. Na Seção III, os métodos lineares de estimativa de energia utilizados neste trabalho são apresentados, juntamente com a proposta do corretor não-linear baseado em redes neurais artificiais. Os resultados considerando dados reais adquiridos durante operação nominal do LHC no CERN são discutidos na Seção IV. Por fim, as conclusões são apresentadas na Seção V.

II. O TILECAL

Um sistema de calorimetria de altas energias tem o objetivo de medir a energia que uma partícula perde quando esta interage com os componentes do calorímetro. Tipicamente, o calorímetro é projetado para detectar a maioria das partículas provenientes de uma colisão, forçando-as a depositar toda a sua energia nos canais de leitura do calorímetro [21].

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 fenômenos [21]. Normalmente, o sistema de calorimetria é seccionado de acordo com a interação: eletromagnética (elétrons e fôtons) e hadrônica

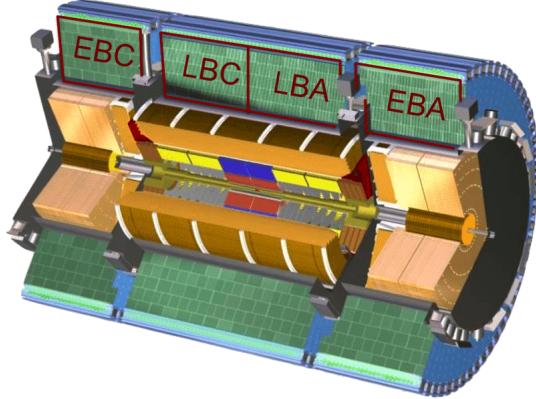


Figura 1: Ilustração das partições do TileCal no ATLAS.

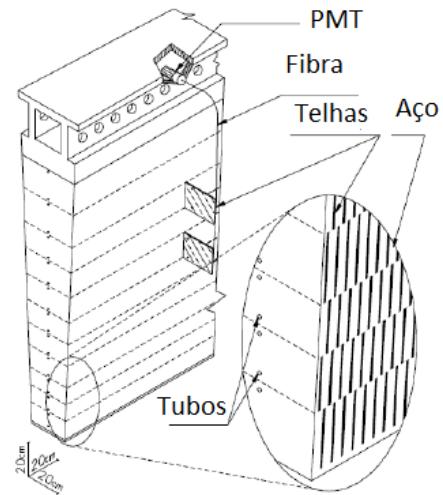


Figura 2: Estrutura de um módulo do TileCal.

(prótons e nêutrons). O Calorímetro de Telhas, ou TileCal, é o principal calorímetro hadrônico do ATLAS, e fornece medidas precisas da energia de jatos, auxilia no cálculo do momento transverso 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 do TileCal é baseado na técnica de amostragem que utiliza placas (ou telhas) de plástico cintilante como material ativo, intercaladas 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*), que transmite a luz para células fotomultiplicadoras, ou PMTs (do inglês, *Photo-Multiplier Tubes*).

A luz, lida pelas PMTs, gera um pulso elétrico em resposta

ao sinal luminoso. Tal pulso é condicionado por um circuito de conformação (chamado de *shaper*), produzindo um pulso de formato fixo e conhecido, o qual a amplitude é proporcional à energia depositada [22]. Desta forma, a energia pode ser recuperada através da estimativa da amplitude do sinal de resposta. O sinal analógico é digitalizado por um conversor Analógico-Digital (ADC) com frequência de amostragem de 40 MHz e uma janela de 150 ns (7 amostras digitais) é utilizada para representar todo o sinal de resposta. A partir das sete amostras temporais adquiridas, os métodos de estimativa operam calculando a amplitude do sinal digital recebido. Vale ressaltar que a correta estimativa da amplitude (energia) dos canais de leitura dos calorímetros impacta diretamente no sistema de filtragem e seleção de eventos, e consequentemente na identidade das partículas [23], [24].

A. Programa de atualização do ATLAS

O experimento ATLAS prevê um extenso programa de atualização, visando lidar com o aumento da luminosidade prevista pelo LHC nas próximas décadas. 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 $cm^2 s^{-1}$ [16]. 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 próton-próton ocorrerão quando os feixes se cruzam, 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. Até o momento, o LHC operou com $\langle \mu \rangle$ aproximadamente igual a 40, embora algumas tomadas de dados especiais tenham atingido valores próximos a $\langle \mu \rangle = 90$.

A próxima tomada de dados pelo LHC é prevista para acontecer em 2021 onde, em média, 80 interações proton-proton ocorrerão a cada colisão [25]. Nestas condições de operação muitos canais do TileCal já sofrem severamente com o efeito do empilhamento de sinais (ver Fig. 3), impondo enormes desafios aos sistemas de filtragem e aquisição de dados. Para as próximas fases de operação, já no contexto do HL-LHC (do inglês, *High-Luminosity LHC*) a luminosidade do LHC atingirá aproximadamente $\langle \mu \rangle = 200$ interações proton-proton por colisão, em média. Para lidar com esse cenário adverso de alta ocupação do detector, a eletrônica dos sistemas do ATLAS, incluindo o TileCal, será atualizada com dispositivos modernos [26], que suportam métodos avançados e sofisticados para o processamento digital de sinais.

III. ESTIMAÇÃO DA ENERGIA

O problema de reconstrução de energia em calorimetria é comumente abordado como um problema de estimativa de parâmetros. Atualmente, o TileCal utiliza o método do Filtro Ótimo (OF, do inglês, *Optimal Filter*) [27] para estimar a amplitude (energia) dos sinais provenientes das colisões. Além

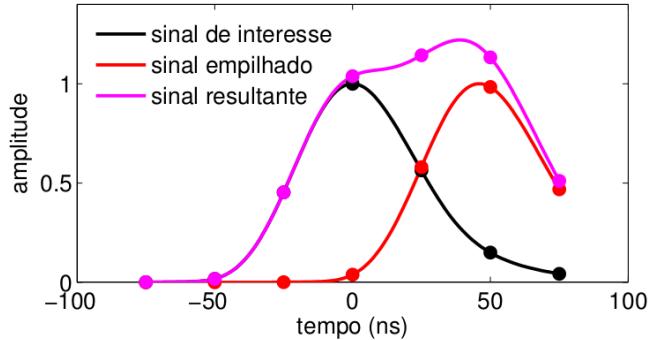


Figura 3: Ilustração do fenômeno de empilhamento de sinais no TileCal. Um sinal de interesse (preto) é adquirido e antes que a eletrônica produza o pulso, um outro sinal (vermelho) de uma colisão adjacente é observado, distorcendo o sinal resultante (magenta).

do OF, o TileCal possui um método alternativo para estimar a amplitude em condições severas de empilhamento de sinais baseado na filtragem de Wiener. Portanto, nesta seção, estes métodos são descritos, destacando-se as suas características principais.

A. Filtro ótimo

O método 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. Em condições de operação em baixa luminosidade, o ruído presente nos calorímetros do ATLAS é predominantemente eletrônico, o qual é tipicamente descrito por uma distribuição Gaussiana. Desta forma, este algoritmo é extensivamente utilizado em calorímetros modernos [28], [29].

No método OF, a amostra digital x no instante k pode ser modelada por

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

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

A amplitude do sinal de entrada é estimada através de uma 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 OF.

Para calcular os coeficientes $w[k]$ do estimador OF, o sinal digital recebido do TileCal é aproximado pela sua 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 normalizado g , e $k = 0, 1, 2, \dots, N - 1$.

Para garantir a imparcialidade do estimador, é imposto que o valor esperado de \hat{A}_{OF} seja igual a A . Logo, substituindo a (3) em (2) e considerando que a média do ruído eletrônico seja 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 imparcial, independente do pedestal e da fase, respectivamente, 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 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.

Os pesos \mathbf{w} são determinados minimizando a expressão da variância do estimador utilizando o método dos multiplicadores de Lagrange. Vale destacar que a matriz de covariância do ruído \mathbf{C} pode ser aproximada pela matriz identidade, caso o ruído seja descrito como sendo Gaussiano e descorrelacionado. Portanto, a solução deste sistema resulta no conjunto de pesos $w[k]$ do estimador OF que operam, atualmente, de forma *online* e *offline* no TileCal.

B. Corretor não-linear

O método OF opera de forma ótima em condições nas quais o ruído possa ser modelado por uma distribuição Gaussiana. Entretanto, em condições de empilhamento de sinais, tal aproximação não pode ser considerada, visto que o efeito de empilhamento de sinais introduz componentes não-Gaussianas no ruído, afetando o desempenho do método OF.

Visando tratar as não-linearidades presentes no ruído devido ao empilhamento de sinais, uma estratégia combinada, que aplica um ajuste na medida linear produzida pelo método OF (ou outro método linear) foi testada no TileCal [18] alcançando resultado promissores. Nesta proposta, um corretor não-linear, baseado em redes neurais, é treinado com a função de corrigir as incertezas no método linear que são provenientes do empilhamento de sinais. Deste modo, tal sistema combinado não visa estimar a energia a partir das amostras digitais, mas sim fornecer um ajuste fino ao método linear já estabelecido. A

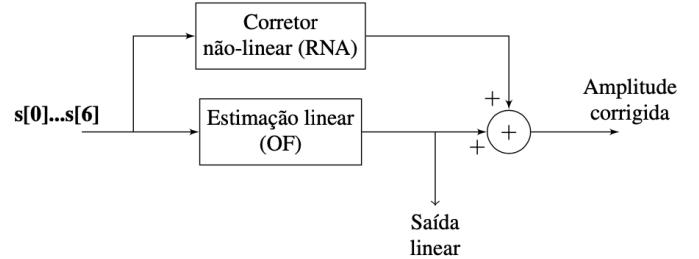


Figura 4: Sistema baseado no corretor não-linear utilizado para assistir a medida linear.

Fig. 4 descreve a proposta híbrida para estimação da amplitude em condições severas de empilhamento de sinais, em que as amostras digitais recebidas, representadas pelo vetor s , correspondem à entrada da rede neural e do método linear. Neste trabalho, esta abordagem é chamada de OF+RNA e é testada utilizando dados reais adquiridos durante operação nominal do LHC em 2018.

Vale ressaltar que para condições de baixa ocupação do canal de leitura (pouco ou nenhum empilhamento de sinais), a contribuição do corretor não-linear deve ser pequena, ou nula, e a estimativa final da amplitude deve ser igual ou próxima à estimativa linear.

C. Filtragem de Wiener

Uma terceira abordagem recentemente implementada e validada para estimação da energia nos canais de leitura do TileCal com alta ocupação (muito empilhamento de sinais) se baseia na filtragem de Wiener [20]. 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 estimativa, 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. No filtro de Wiener, a saída do filtro é dada pelo somatório

$$y[n] = \sum_{k=0}^{N-1} c[k]x[n-k], \quad (10)$$

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

$$\mathbf{J} = \mathbb{E}\{e[n]^2\}, \quad (11)$$

onde \mathbb{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,

$$\sum_{i=0}^{N-1} c[i] \mathbb{E}\{x[n-k]x[n-i]\} = \mathbb{E}\{x[n-k]d[n]\}, \quad (12)$$

onde $k = 0, 1, \dots, N - 1$. Da (12), pode-se observar que:

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

$$R[i, k] = \frac{1}{N} \sum_{n=0}^{N-1} x[n-k]x[n-i]. \quad (13)$$

- 2) O valor esperado $\mathbb{E}\{x[n-k]d[n]\}$ 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

$$p[k] = \frac{1}{N} \sum_{n=0}^{N-1} x[n-k]d[n]. \quad (14)$$

As equações definidas em (13) e (14) são conhecidas como equações de Wiener-Hopf. Substituindo estas equações em (12), obtém-se um sistema de equações lineares como condição necessária e suficiente para otimizar o filtro, como segue

$$\sum_{i=0}^{N-1} c[i]R[i, k] = p[k] \quad k = 0, 1, \dots, N - 1. \quad (15)$$

Pode-se reescrever este sistema de equações na forma matricial

$$\mathbf{c}\mathbf{R} = \mathbf{p}. \quad (16)$$

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

$$\mathbf{c} = \mathbf{R}^{-1}\mathbf{p}, \quad (17)$$

onde \mathbf{R} representa a matriz de autocorrelação das amostras temporais dos sinais de entrada (expressão mostrada em (13)) 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 (ver equação (14)).

Vale ressaltar que o filtro de Wiener considera que o ruído possui média zero, o que não corresponde ao caso do ruído de empilhamento de sinais unipolares, como no TileCal. Para contornar o problema da média do ruído, um coeficiente adicional, de valor constante igual a 1 (um), é adicionado ao processo de otimização do filtro de Wiener. 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 suprimir 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. Sendo assim, a estimativa da amplitude \hat{A}_{FW} do filtro de Wiener do TileCal 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 em (18).

$$\hat{A}_{FW} = \left(\sum_{i=0}^{N-1} c[i]x[i] \right) + c[N] \quad (18)$$

IV. RESULTADOS

Os resultados são apresentados na unidade de contagem de ADC (do inglês, *analog to Digital Converter*). Para o TileCal, uma contagem de ADC corresponde a aproximadamente 12 MeV.

A. Banco de dados

Dados reais de colisão próton-próton adquiridos em 2018 pelo LHC (último período de tomada de dados) foram utilizados para análise de eficiência de estimativa da amplitude (energia) dos métodos descritos. Visto que este trabalho visa avaliar o desempenho de métodos operando em condições adversas de empilhamento de sinais, foi utilizado um conjunto de dados com número médio de interações próton-próton $\langle \mu \rangle$ igual a 90. Estes dados foram produzidos de forma excepcional no LHC, uma vez que, tipicamente, as tomadas de dados do último ano de operação atingiram valores de $\langle \mu \rangle$ próximos a 40.

Vale ressaltar que os dados reais utilizados foram obtidos de tomadas de dados chamadas de *Zero Bias stream*, onde somente *triggers* aleatórios são utilizados. Portanto, é esperado que apenas informação de ruído eletrônico e de empilhamento seja adquirida. Em outras palavras, nestes eventos não é esperado observar nenhum sinal de interesse na janela de aquisição, constituindo apenas dados de ruído. Adicionalmente, para efeito de comparação, sinais de dois canais de leitura do TileCal foram utilizados, um com muito empilhamento de sinais (canal E4) e outro com pouco empilhamento de sinais (canal D4). Ao todo, para cada canal de leitura, aproximadamente 800.000 eventos foram utilizados. A Fig. 5 mostra as distribuições das amostras temporais dos sinais dos canais E4 e D4, onde pode ser observado que o efeito de empilhamento de sinais produz um ruído com uma maior dispersão e maior assimetria.

O conjunto foi dividido em dois subconjuntos onde 60% foi atribuído ao conjunto de desenvolvimento, utilizado para projetar os métodos. O restante, 40% dos eventos, foram destinados ao conjunto de teste, utilizado para avaliar a eficiência dos métodos.

O método OF necessita do conhecimento do pulso de referência, e da matriz de covariância do ruído eletrônico, estimado através do conjunto de desenvolvimento. Por outro lado, tanto para o projeto do filtro de Wiener e do método linear assistido por redes neurais, um sinal de interesse conhecido (alvo) deve ser adicionado a cada evento do conjunto de desenvolvimento. 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

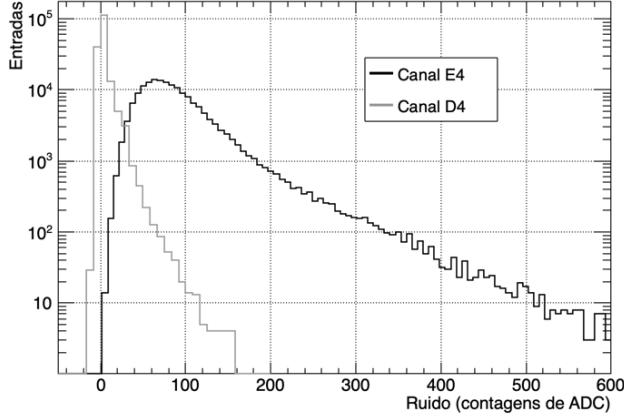


Figura 5: Ruído eletrônico e de empilhamento presente em dois canais com diferentes ocupações no TileCal.

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 [27]. 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 [22]. Desta forma, cada valor de amplitude tem a mesma probabilidade de ocorrência. Por fim, os sinais gerados foram somados aos eventos do conjunto de desenvolvimento utilizados para projetar os coeficientes do filtro de Wiener e da rede neural do método OF+RNA.

Para o conjunto de teste, também foi adicionado um sinal de interesse a cada evento, com as mesmas características utilizadas para o conjunto de treinamento. Entretanto, a fim de caracterizar o comportamento mais apropriado para a deposição de energia, foi utilizado um modelo exponencial de média 90 para a amplitude do sinal de interesse, ao invés de uniforme.

B. Projeto da RNA

De forma similar à apresentada em [18], o projeto do corretor não-linear foi baseado em uma rede neural artificial MLP (do inglês, *Multi-Layer Perceptron*), com duas camadas escondidas com oito neurônios cada, e uma única camada de saída com um neurônio. Todas as camadas permitem a inclusão de um neurônio *bias*, conforme mostrado na Fig. 6. A função de ativação utilizada para os neurônios da camada escondida foi a ReLU (do inglês, *Rectified Linear Unit*). Já para o neurônio da camada de saída, foi utilizada a função linear. A função de perda escolhida foi o MSE (do inglês, *Mean Squared Error*) enquanto o algoritmo otimizador aplicado foi o Adam [30]. A escolha desta topologia foi baseada em análises de complexidade e eficiência de estimativa da amplitude, ou seja, a eficiência (dispersão do erro de estimativa) se mantém

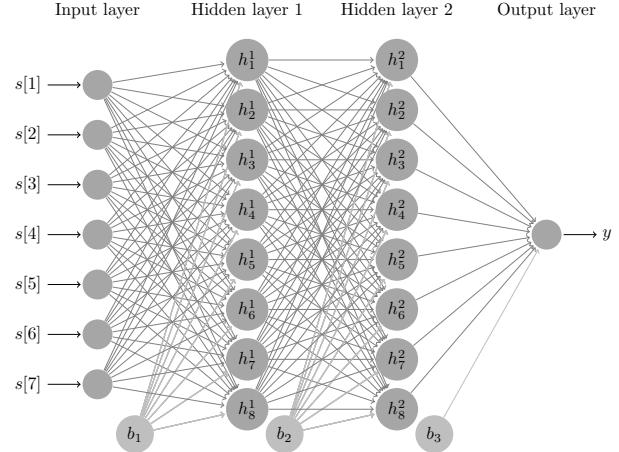


Figura 6: Topologia da rede neural artificial utilizada como corretor não-linear.

constante para um número maior de neurônios na camada intermediária.

Para cada canal (E4 e D4), a mesma arquitetura da rede neural foi treinada separadamente utilizando o respectivo conjunto de desenvolvimento em que o alvo corresponde à diferença entre o valor verdadeiro conhecido da amplitude do sinal, e a estimativa linear produzida pelo método OF. Vale ressaltar que 20% dos dados do conjunto de desenvolvimento foi utilizado para a validação da rede neural treinada.

C. Análise de eficiência

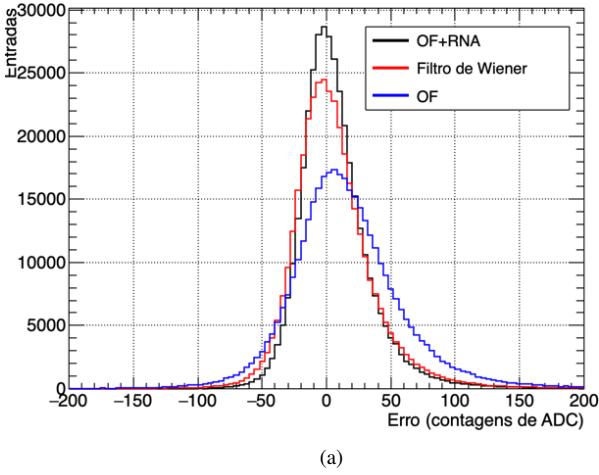
Para avaliar a eficiência dos métodos, O erro de estimação (E_r) foi calculado subtraindo-se o valor de amplitude estimado $\hat{A}[i]$ pelo valor de referência $d[i]$ para cada sinal i do conjunto de teste:

$$E_r[i] = \hat{A}[i] - d[i] \quad (19)$$

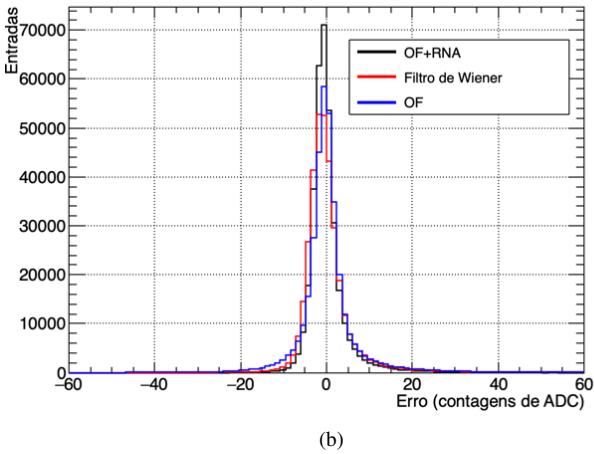
Os valores de média e RMS (do inglês, *Root Mean Square*) da distribuição do erro de estimação do conjunto de teste foram utilizados. Estes valores representam, respectivamente, a parcialidade (ou tendência) de estimativa e variância associado a cada algoritmo. A Fig. 7 mostra as distribuições do erro de estimação dos conjuntos de teste para os canais E4 e D4 quando o método OF+RNA, o método OF e o Filtro de Wiener são aplicados. A Tab. I apresenta os valores de média e RMS (do inglês, *Root Mean Square*) de cada uma das distribuições de erro dos canais E4 e D4.

Para o canal E4, pode-se observar que o histograma para o método OF+RNA apresenta a menor dispersão (menor erro) e, portanto, a melhor eficiência. Comparando com o método linear OF, a abordagem combinada OF+RNA apresenta uma melhora 35% no RMS do erro. Já para o canal D4, o método OF+RNA não apresenta um desempenho muito superior aos outros métodos considerando os valores absolutos. Isto pode ser justificado pelo fato de o canal D4 não sofrer severamente do efeito de empilhamento de sinais, e os métodos lineares

conseguem um desempenho próximo ao ótimo. Entretanto, em valores relativos, a melhora introduzida pela RNA continua sendo de aproximadamente 35% em relação ao método OF puro (sem a RNA corretora). Adicionalmente, conforme esperado, o Filtro de Wiener apresenta um desempenho superior ao método OF, mostrando que seu procedimento de otimização é mais amplo que o utilizado pelo método OF.



(a)



(b)

Figura 7: Erro de estimação considerando (a) o canal de maior ocupação no TileCal (canal E4) e (b) um canal de baixa ocupação (canal D4).

Tabela I: Média e RMS dos histogramas de erro de estimação (em contagens de ADC).

	Canal E4		Canal D4	
	média	RMS	média	RMS
OF+RNA	6,45	27,95	0,19	5,10
Filtro de Wiener	5,33	32,00	0,31	6,65
OF	15,19	42,80	0,43	8,00

A Fig. 8 mostra as contribuições produzidas pela rede neural do método OF+RNA. Conforme esperado, pode-se notar que para um canal com muito empilhamento (canal E4) a rede

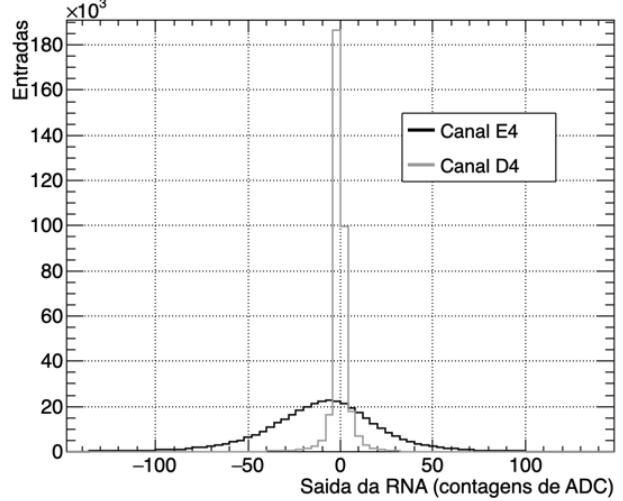


Figura 8: Saídas do corretor não-linear baseado em redes neurais artificiais considerando os canais E4 (alta ocupação) e D4 (baixa ocupação).

neural fornece valores com uma dispersão muito maior que para um canal com pouco empilhamento de sinais (canal D4). O RMS do histograma da saída da rede neural para o canal D4 foi 6,91 contagens de ADC, enquanto para o canal E4, a distribuição da saída da rede neural apresentou um valor de RMS igual a 33,61 contagens de ADC. Ou seja, nestas condições de operação, o RMS da contribuição da RNA para um canal de alta ocupação é aproximadamente 5 vezes maior quando comparado à um canal de baixa ocupação, como o canal D4.

A fim de verificar a linearidade e a parcialidade do método OF+RNA, a Fig. 9 mostra o erro relativo, em percentual, em função do valor de referência da amplitude. Pode-se notar que o erro relativo se dispersa em torno do zero, indicando imparcialidade, e tende ao valor zero quando o valor de amplitude de referência aumenta, mostrando a resposta linear do método.

V. CONCLUSÃO

Este trabalho abordou o problema de estimativa de parâmetros aplicado a reconstrução de energia no calorímetro de Telhas (TileCal) do ATLAS, no LHC, operando em condições severas de empilhamento de sinais. Métodos lineares operam de forma subótima visto que o efeito inevitável do empilhamento de sinais introduz componentes não-lineares no ruído. Considerando o programa de atualização do ATLAS, onde novos dispositivos de processamento digital de sinais serão empregados, o uso de técnicas avançadas de processamento de dados poderá ser empregado.

Utilizando dados reais de colisão de células severamente afetadas pelo efeito de empilhamento de sinais, observou-se um ganho em torno de 35% em resolução de estimativa (RMS do erro de estimativa) utilizando o método assistido por redes neurais (OF+RNA), quando comparado ao método atualmente

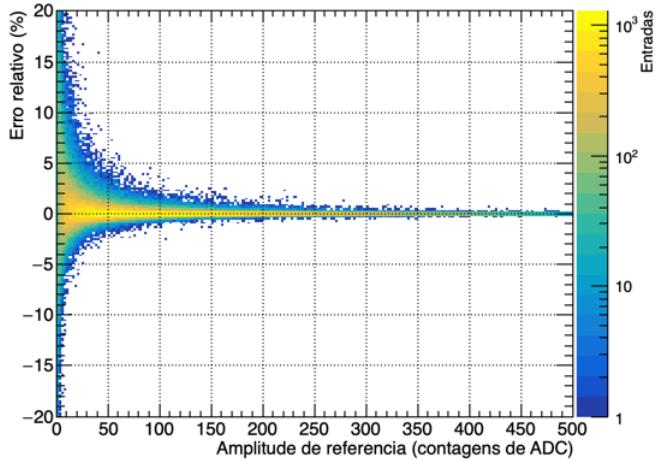


Figura 9: Erro relativo considerando o método combinado (OF+RNA), mostrando o seu comportamento linear em função da amplitude do sinal recebido.

utilizado (método OF) no TileCal. Já para o método linear baseado na filtragem de Wiener, alternativo para o método OF, o método OF+RNA apresentou um desempenho próximo de 10% superior em termos do RMS do erro de estimação.

Os esforços atuais estão sendo empregados para testar métodos baseados em aprendizado profundo e também de se analisar o impacto da estimativa de energia dos métodos nos objetos da Física, como a reconstrução de partículas.

AGRADECIMENTOS

Os autores agradecem a FAPERJ, FAPEMIG, CAPES, CNPq e RENAFAE 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

- [1] Labati R., Genovese A., Piuri V. and Scotti F., *Low-cost volume estimation by two-view acquisitions: A computational intelligence approach*, International Joint Conference on Neural Networks (IJCNN), pp. 18 (2012).
- [2] Kay, S. M., *Fundamentals of Statistical Signal Processing, Estimation Theory*, Prentice Hall, (1993).
- [3] Vizireanu, D.N. and Halunga, S.V., *Simple, fast and accurate eight point samplitude estimation method of sinusoidal signals for DSP based instrumentation*, Journal of Instrumentation, vol.7 (2012).
- [4] Pincibono B. and Duvault P., *Optimal Linear-Quadratic Systems for Detection and Estimation*, IEEE Transactions on Information Theory, vol.34, pp.304-311 (1988).
- [5] Pincibono B. and Chavalier P., *Widely linear estimation with complex data*, IEEE Transactions on Signal Processing, vol.43, pp.2030-2033 (1995).
- [6] Bos A., *Parameter Estimation for Scientists and Engineers*, Wiley-Interscience, 1st edition (2007).
- [7] Wang D. et al., *Readout electronics upgrade on ALICE/PHOS detector for Run 2 of LHC*, Journal of Instrumentation, vol.10, C02025 (2015).
- [8] Haykin, S., *Neural Networks: A Comprehensive Foundation*, Prentice Hall, 1998.
- [9] Denby B., *Neural networks in high energy physics: A ten year perspective*, Computer Physics Communications, vol. 119, no. 23, pp.219-231 (1999).
- [10] Gleyzer S.V. and Prosper H., *An artificial neural network based algorithm for calorimetric energy measurements in CMS*, Proceedings of XII Advanced Computing and Analysis Techniques in Physics Research, p.91 (2008).
- [11] Evans, L. and Bryant, P., (editors), *LHC Machine*, J. Instrum., vol. 3, p. S08001, 2008.
- [12] The ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, J. Instrum., vol. 3, p. S08003, 2008.
- [13] Cottingham W. N. and Greenwood, D. A., *An Introduction to the Standard Model of Particle Physics*, Cambridge University Press, 1998.
- [14] The ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Physics Letters B, v. 716, pp. 1-29, 2012.
- [15] The ATLAS and CMS Collaborations, *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV*, Journal of High Energy Physics volume 2016, Article number: 45 (2016).
- [16] Herr, W. and Muratori, B., *Concept of luminosity*. doi:10.5170/CERN-2006-002.361. URL <https://cds.cern.ch/record/941318> (2006).
- [17] Ruggiero, F., *LHC Accelerator R&D and Upgrade Scenarios*, The European Physical Journal C - Particles and Fields, v. 34, pp. 433-442 (2004).
- [18] Peralva, B. et al., *Correção não-linear na estimativa de energia para um calorímetro de altas energias*, Congresso Brasileiro de Inteligência Computacional, Curitiba, <https://doi.org/10.21528/CBIC2015-161> (2015).
- [19] The ATLAS Collaboration, *Readiness of the ATLAS Tile Calorimeter for LHC collisions*, EPJC 70, pp.1193-1236 (2010).
- [20] Gonçalves, G.I., et al., *Desempenho de Algoritmos de Estimação de Energia para o Calorímetro de Telhas do Experimento ATLAS*, Congresso Brasileiro de Automática, <https://doi.org/10.48011/asba.v2i1.1125> (2020).
- [21] R. Wigmans, *Calorimetry: Energy Measurement in Particle Physics*, 2nd Edition, Oxford University Press (2017).
- [22] Anderson, K., et al., *Front-end Electronics for the ATLAS Tile Calorimeter*, Proceedings of Fourth Workshop on Electronics for LHC Experiments, Rome (1998).
- [23] Kordas, K., *The ATLAS data acquisition and Trigger: Concept, design and status*, Nuclear Physics B - Proc. Suppl., vol. 172, pp.178-182, 2007.
- [24] The ATLAS collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. C 77 (2017).
- [25] Schmidt, B., *The High-Luminosity upgrade of the LHC: Physics and Technology Challenges for the Accelerator and the Experiments*, J. Phys.: Conf. Ser. 706 022002 (2016).
- [26] Scuri, F., *Upgrade of the ATLAS Tile Calorimeter for the High Luminosity LHC*, J. Phys.: Conf. Ser. **1162** 012017 (2019).
- [27] Fullana, E., et al., *Digital Signal Reconstruction in the ATLAS Hadronic Tile Calorimeter*, IEEE Transaction On Nuclear Science, v. 53, number 4, 2006.
- [28] Bertuccio, G., Gatti, E., Sapietro, M., *Sampling and optimum data processing of detector signals*, Nuclear Instruments and Methods in Physics Research A, v. 322, pp.271-279, 1992.
- [29] Cleland, W. and Stern, E., *Signal processing considerations for liquid ionization calorimeters in a high rate environment*, Nuclear Instruments and Methods in Physics Research A 338 467-497 (1994). [https://doi.org/10.1016/0168-9002\(94\)91332-3](https://doi.org/10.1016/0168-9002(94)91332-3).
- [30] Kingma, D. and Ba, J., *Adam: A method for stochastic optimization*, International Conference on Learning Representations, 12 (2014).