VETOR

Desconvolução de sinais para estimação de energia em sistemas de calorimetria operando a alta taxa de eventos☆

Signal deconvolution for energy estimation in calorimetry system operating at high event rate

Guilherme Barroso Morett[†], Bernardo Sotto-Maior Peralva, Gustavo Barbosa Libotte

Programa de Pós-Graduação em Modelagem Computacional, Universidade do Estado do Rio de Janeiro, Instituto Politécnico, Nova Friburgo - RJ - Brasil
[†]**Autor correspondente:** guilherme.morett@iprj.uerj.br

Resumo

Este artigo visa avaliar o desempenho de um método baseado na Desconvolução de Sinais para a estimação da energia em calorímetros de altas energias. O contexto do trabalho compreende a reconstrução dos sinais produzidos pelos canais de leitura de um sistema de calorimetria e advindos de colisões de partículas atômicas. O tratamento do conjunto de dados é realizado no formato contínuo ou sequenciado (*free-running*), conforme o funcionamento de experimentos modernos que operam em alta taxa de eventos e alta luminosidade. Esse cenário permite definir qual a quantidade de janelamento das amostras pode ser adotada como ideal, buscando aprimorar a eficiência da reconstrução do sinal adquirido pelos canais de leitura. Além disso, a técnica de validação cruzada *K-Fold* também é empregada para a análise estatística do erro de estimação. Dessa maneira, a comparação de eficiência do método proposto com a técnica linear atualmente utilizada em calorímetros modernos é apresentada considerando diferentes condições de operação. O estudo identificou os janelamentos ideais para os métodos comparados na estimação de amplitudes sob as condições analisadas, assim como revelou a independência do método de Desconvolução em relação ao empilhamento, tornando-o vantajoso em condições de alta luminosidade.

Palavras-chave

Reconstrução de sinal • Estimação de amplitude • Desconvolução de sinais • Alta luminosidade • Física experimental de altas energias

Abstract

This article aims to evaluate the performance of a method based on Signal Deconvolution for energy estimation in high-energy calorimeters. The study focuses on reconstructing signals produced by the readout channels of a calorimetry system resulting from atomic particle collisions. Data processing is performed in a continuous or sequential format (free-running), aligned with the operation of modern experiments characterized by high event rates and high luminosity. This context enables determining the optimal sample processing interval to enhance the efficiency of signal reconstruction by the readout channels. Additionally, the K-Fold cross-validation technique is employed for statistical error analysis. The study compares the efficiency of the proposed method with the linear technique currently used in modern calorimeters under different operating conditions. The findings identified the optimal intervals for the methods in amplitude estimation and highlighted the independence of the Deconvolution method from signal pile-up, making it advantageous in high-luminosity environments.

☆Este artigo é uma versão estendida do trabalho apresentado no XXVII ENMC Encontro Nacional de Modelagem Computacional e XV ECTM Encontro de Ciência e Tecnologia de Materiais, ocorridos em Ilhéus – BA, de 1 a 4 de outubro de 2024.

Keywords

Signal reconstruction • Amplitude estimation • Signal deconvolution • High Luminosity • Experimental high energy physics

1 Introdução

O campo de processamento de sinais teve seu desenvolvimento impulsionado no início do século XX. Atualmente, tal disciplina tem vasta aplicação, desde o setor de telecomunicações até diagnósticos médicos. Isso foi possível graças ao avanço computacional e tecnológico, o que permitiu trabalhar com uma elevada quantidade de dados gerados em um curto intervalo de tempo com resultados satisfatórios [1]. Um dos tópicos dessa área é a reconstrução de sinal, que visa estimar parâmetros de interesse a partir de dados de sinais corrompidos com ruídos. Pode-se dizer que o estado de arte é na física de altas energias, ambiente em que se deve ter cuidado para não perder informações provenientes de eventos raros.

Nos experimentos de física de altas energias, partículas atômicas, geralmente feixes de prótons ou íons pesados, são acelerados a uma velocidade muito próxima da luz e colididos em pontos específicos. Por ilustração, há os experimentos do Grande Colisor de Hádrons (*Large Hadron Collider - LHC*) [2] da Organização Europeia para a Investigação Nuclear (*l'Organisation Européenne pour la Recherche Nucléaire* - CERN). O conhecimento desenvolvido no CERN, o maior laboratório de física experimental do mundo, beneficia toda a sociedade [3].

Ao redor da região de choque, sofisticados sistemas de detecção, como os calorímetros, estão localizados. Esses dispositivos conseguem captar as energias em seus canais de leitura das partículas subatômicas que com eles interagem. Após o registro, o sinal inicialmente analógico é digitalizado para compor as amostras de pulsos de sinais. Tipicamente, calorímetros de altas energias possuem milhares de canais de leitura. As amostras temporais destes sinais são empregadas nos métodos lineares para a estimação dos parâmetros de interesse. Vale ressaltar que a amplitude do sinal de interesse é de suma importância, pois o seu valor é proporcional à energia depositada. Em seguida, o erro de estimação é calculado pela diferença entre a amplitude de referência e a estimada.

Segundo a nova atualização dos colisores [4], o cenário de operação é de alta luminosidade, em virtude da adoção de feixes de partículas cada vez mais densos. Por um lado, esse ambiente privilegia a investigação de fenômenos físicos que extrapolam aqueles já consolidados pela Física do Modelo Padrão, mas, por outro, intensifica a participação dos ruídos nas amostras, principalmente o de empilhamento de sinais (*pile-up effect*).

Os conceitos referentes ao sistema de calorimetria, assim como o conjunto de dados, técnicas e métodos adotados são explicados na seção de modelagem matemática. Já em resultados, há a exibição dos gráficos referentes à média e o desvio padrão obtidos pelo *K-Fold* para diferentes cenários de um canal de leitura. Logo é possível determinar qual janelamento pode ser definido como ótimo para a comparação entre os métodos pela avaliação do erro de estimação (*Mean Estimation Error* - EME) e desvio padrão médio para distintos casos de operação. Por fim, na parte de considerações finais, uma recapitulação do desempenho dos métodos é realizada, além de sugestões para melhorias e trabalhos futuros.

2 Modelagem Matemática

2.1 Reconstrução de sinal em calorimetria de altas energias

O calorímetro é um sofisticado e sensível detector que registra em aproximadamente 10.000 canais de leitura, a energia de partículas subatômicas provenientes da colisão [5]. Geralmente, estes são divididos em dois tipos: eletromagnético e hadrônico [6]. A distinção entre eles se deve ao tipo de força de interação da partícula alvo. O que reflete na escolha do material ativo e passivo na sua composição. O material ativo desempenha a função de amostrar, enquanto que o passivo de absorver a energia e induzir interações secundárias entre as partículas.

O calorímetro eletromagnético é projetado para mensurar a energia de partículas que interagem pelo campo eletromagético, tais como fótons e elétrons. O material passivo adotado possui alta densidade como o aço ou tungstênio. O material ativo é composto por argônio líquido.

Por outro lado, o calorímetro hadrônico é arquitetado para operar com hádrons, partículas constituídas por quarks, como os prótons, nêutrons e píons. Analogamente ao eletromagético, um metal como o aço também é utilizado como material passivo. Entretanto, o material ativo é formado por telhas poliméricas cintilantes.

Depois de captado a energia, esta é convertida em amostras discretas de sinais que constituem o conjunto de dados de entrada para os métodos matemáticos com soluções numéricas. Os métodos são responsáveis pela reconstrução dos parâmetros do sinal.

2.2 Simulação do conjunto de dados

O conjunto de dados (*dataset*) foi simulado por meio da biblioteca *Python Pulse Generator* [7]. A taxa de colisão das partículas é próxima de 40 MHz, ou seja, a cada 25 ns ocorre um choque. Para cada instante de tempo de colisão, o simulador fornece o pulso de sinal em unidades Analog to Digital Converter (ADC Count, ou contagens de ADC), a amplitude de referência (em ADC Count) e a fase de referência (ns). Para a geração dos dados desse artigo, as seguintes configurações das distribuições dos parâmetros foram consideradas: exponencial de média 100 ADC Count para a amplitude; distribuição uniforme com valores inteiros no intervalo entre -5 e 5 para a fase; valor constante de 30 ADC Count para o pedestal e; nível de deformação de 0,01 ADC Count. No total, foram simulados 2.000.000 eventos. Vale observar que o valor do pedestal foi retirado dos pulsos de sinais gerados, por ser empregado somente na etapa de conversão do sinal analógico para digital.

O termo de ocupação se refere à probabilidade de um determinado canal de leitura receber informações dada uma certa colisão. A ocupação varia de 0 até 100 %, com incremento de 10 % nesse intervalo. Isso significa que, para ocupação 0, o canal não teve nenhum estímulo e a informação captada advém somente do ruído eletrônico. Por outro lado, para a ocupação 100 %, o canal foi constantemente bombardeado por deposição de energia e nessa ocupação o efeito de empilhamento tem maior participação.

A Tabela 1 representa a organização dos dados gerados pelo *Pulse Generator* dispostos em colunas.

Tabela 1: Exemplo de conjunto de dados simulados para a ocupação 100 %.

Tempo (ns)	Pulso de Sinal (ADC Count)	Amplitude ref. (ADC Count)	Fase ref. (ns)
0,0000	72,1957	14,1031	3,0000
25,0000	170,2104	76,0928	−3,0000
50,0000	216,7832	113,5893	2,0000
⋮	⋮	⋮	⋮
49999975,0000	187,1000	64,6319	1,0000

Fonte: O autor, 2024.

A curva de pulso de sinal de interesse gerado é similar a uma gaussina com uma duração de cerca de 150 ns [8]. O formato da curva é determinada pelos parâmetros da amplitude, fase e pedestal. A amplitude está associada ao pico, que por sua vez corresponde a amostra central, ou seja, aquela registrada no tempo zero. No entanto, em virtude da falta de sincronicidade, há uma defasagem em ns da amostra central em relação ao pico. A fase é justamente essa discrepância. Já o pedestal é a linha de base e não influencia no formato da curva, somente em seu deslocamento vertical.

Para a definição da curva de pulso de sinal são necessárias no mínimo sete amostras discretas de sinais, porque esse valor consegue abranger todo o tempo de extensão do pulso. Dessa análise advém o termo janelamento, devido ao fato que ele define a quantidade de elementos do espaço amostral empregados na estimação.

Os ruídos presentes nas amostras são constituídos por duas parcelas, uma referente à própria eletrônica do equipamento de medição e outra advinda do efeito de empilhamento de sinais (*pile-up effect*). O empilhamento pode ser do tipo *in-time pile-up*, quando ocorre concomitantemente com o pulso de interesse. Ou *out-of-time pile-up* que também ocorre na mesma janela temporal, só que com o pico deslocado no tempo. O efeito de empilhamento degrada o formato de um único pico da curva por haver a amostragem de colisões adjacentes em um mesmo canal de leitura, o que está representado em Fig. 1 para o tipo *out-of-time pile-up*, cujo pico está em + 50 ns.

2.3 Técnica de janelamento.

Os dados são gerados de maneira sequenciada (*free-running format*) a cada instante de colisão, conforme pode ser visto na Tabela 1. Sendo assim, previamente a aplicação do método é preciso organizar essas informações de acordo com o janelamento, que define a quantidade de amostras empregadas em cada processo de estimação. Nesse trabalho, foi adotado como 7 o valor inicial de janelamento e 19 o final. Por ilustração, para o janelamento sete, os sete primeiros pulsos de sinais definem o primeiro lote do espaço amostral, que está vinculado ao quarto elemento da amplitude de referência, alvo de estimação. Em seguida, o segundo lote é composto do segundo ao oitavo elemento e o quinto elemento passa a ser a amplitude de referência. O processo é repetido até atingir o pulso de sinal com índice igual à quantidade total de amostras menos o valor do janelamento desejado. A Fig. 2 denota essa estratégia para o janelamento 7. Vale ressaltar que quanto maior é o janelamento, maior é o custo computacional.

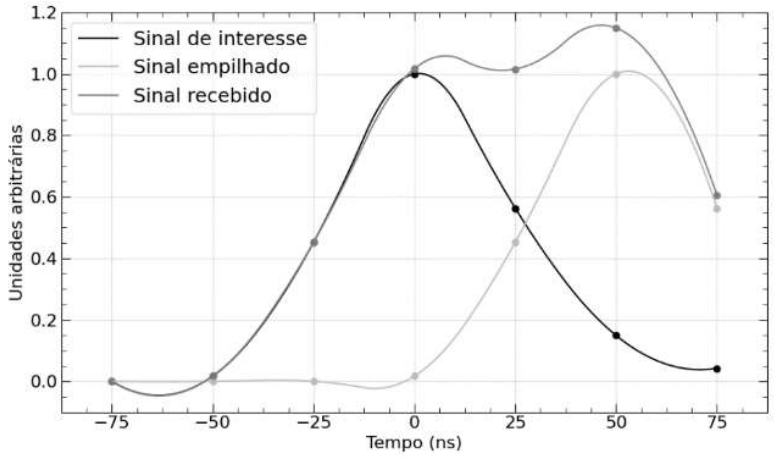


Figura 1: Efeito do empilhamento no formato da curva do pulso de sinal para sete amostras digitais.

Fonte: Rimes et al, 2024 [9].

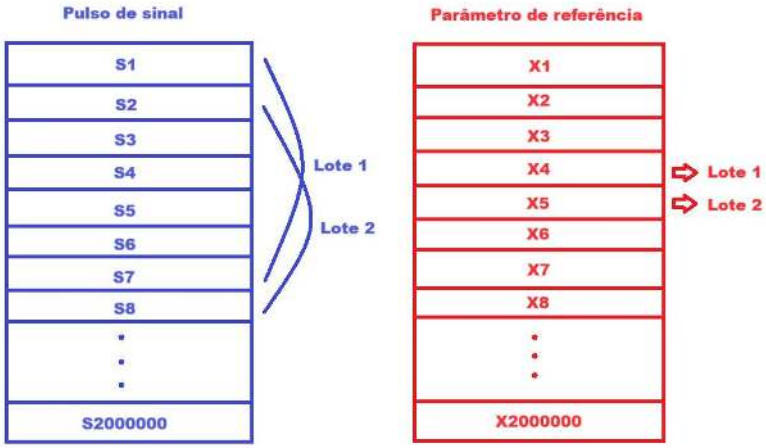


Figura 2: Ilustração da técnica de janelamento para 7 amostras.

Fonte: Os autores, 2024.

2.4 Técnica de validação cruzada *K-Fold*.

No campo do aprendizado de máquina, é comum dividir o conjunto em dados de treino e teste [10]. A Fig.3 ilustra o processo de validação cruzada *K-Fold* para apenas cinco blocos.

O grupo de treino é usado para o cálculo dos pesos. Para que, em seguida, os pesos sejam aplicados nos dados de teste para a estimação dos parâmetros. Além disso, para se ter uma melhor exploração estatística das informações contidas nos dados, estes são alocados em blocos indexados. De forma que cada bloco passa a ser um conjunto de dados. No final, após o processo iterativo dos blocos, tem-se um parecer estatístico da média e desvio padrão dos erros de estimação. Graficamente, os pontos e a barra de erro denotam a média e o desvio padrão dos resultados, respectivamente. Para esse trabalho, foi realizado o *K-Fold* com cem blocos (K = 100).

2.5 Método de Desconvolução de sinal

Ao considerar um sistema linear invariante no tempo discreto [11], a resposta do calorímetro $y[n]$ para a amostra n é interpretada como uma convolução entre a função pulso de referência $h[n]$ associada e a energia depositada no canal de leitura $a[n]$ somada com o ruído $w[n]$ [12]. Matematicamente tem-se a Eq. (1).

$$y[n] = \sum_{i=-\infty}^{+\infty} (h[i]\, a[n-i]) + w[n], \tag{1}$$

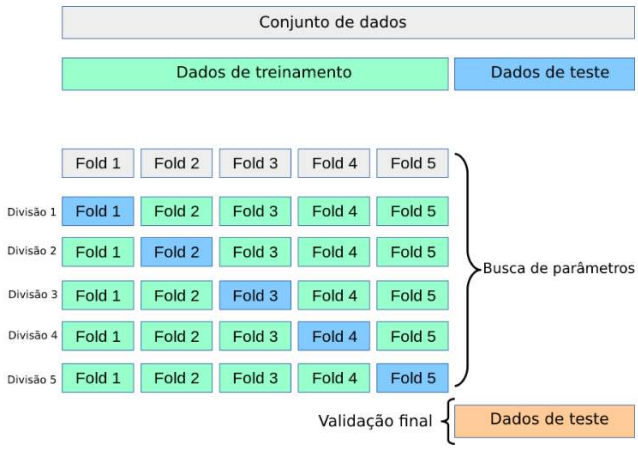


Figura 3: Técnica de validação cruzada *K-Fold* para cinco blocos.

Fonte: Scikit-learn, 2024 [10].

Essa expressão também pode ser vista como um sistema linear representado pela seguinte Eq. (2) matricial:

$$\mathbf{y}_{NX1} = \mathbf{H}_{NXP}\mathbf{a}_{PX1} + \mathbf{w}_{NX1}, \tag{2}$$

na qual $\mathbf{y}_{NX1}$ é o vetor das amostras de resposta do calorímetro, $\mathbf{H}_{NXP}$ é a matriz com os pulsos de referência deslocados, $\mathbf{a}_{PX1}$ é o vetor de interesse que contém os parâmetros e $\mathbf{w}_{NX1}$ é o ruído eletrônico combinado com o de empilhamento. Os índices das dimensões estão relacionados pela equação $P = D + N - 1$. O termo P se refere a quantidade deslocada do pulso de referência, D é o número de feixes envolvidos na colisão e N corresponde ao número de amostras adotado no processo. Note que $P > N$, o que torna o sistema a princípio indeterminado. Logo, o processo de desconvolução considera somente $P \leq N$ componentes do vetor $\mathbf{a}_{PX1}$.

A solução para o vetor da amplitude estimada $\hat{\mathbf{a}}$ é dada pela Eq. (3):

$$\hat{\mathbf{a}}_{\mathbf{P}} = \mathbf{G}_{\mathbf{P}}^{\mathbf{T}}\mathbf{y}, \tag{3}$$

portanto, o passo que falta é determinar a matriz $\mathbf{G}_{\mathbf{P}}$ e posteriormente sua transposta $\mathbf{G}_{\mathbf{P}}^{\mathbf{T}}$.

Para tal feito, a restrição de que a esperança do vetor da amplitude estimada deve ser igual ao vetor da amplitude de referência é imposta através da Eq. (4),

$$E\{\hat{\mathbf{a}}_{\mathbf{P}}\} = \mathbf{G}_{\mathbf{P}}^{\mathbf{T}}E\{\mathbf{y}\} = \mathbf{a}_{\mathbf{P}}, \tag{4}$$

Sabe-se que o vetor de ruídos $\mathbf{w}_{\mathbf{P}}$ apresenta média nula, o que torna a expressão para a esperança do vetor de pulsos de sinais $\mathbf{y}$ simplesmente o produto entre a matriz dos pulsos de referência $\mathbf{H}_{\mathbf{P}}$ pelo vetor da amplitude real $\mathbf{a}_{\mathbf{P}}$, o que é denotado pela Eq. (5),

$$E\{\mathbf{y}\} = E\{\mathbf{H}_{\mathbf{P}}\mathbf{a}_{\mathbf{P}} + \mathbf{w}_{\mathbf{P}}\} = \mathbf{H}_{\mathbf{P}}\mathbf{a}_{\mathbf{P}}. \tag{5}$$

Ao substituir o resultado da Eq. (5) na Eq. (4) encontra-se uma relação entre as matrizes $\mathbf{G}_{\mathbf{p}}$ e $\mathbf{H}_{\mathbf{P}}$ por meio da identidade $\mathbf{I}_{\mathbf{P}}$ mostrada pela Eq. (6):

$$\mathbf{G}_{\mathbf{P}}^{\mathbf{T}}\mathbf{H}_{\mathbf{P}} = \mathbf{I}_{\mathbf{P}}. \tag{6}$$

Em outras palavras, a Eq. (6) estabelece um princípio de ortogonalidade entre o produto escalar das linhas da matriz $\mathbf{G}_{\mathbf{P}}$ e as colunas da matriz $\mathbf{H}_{\mathbf{P}}$, o que pode ser representado pela função delta de Kronecker δ_{ij}, segundo a Eq. (7):

$$\mathbf{g}_{\mathbf{i}}^{\mathbf{T}} \cdot \mathbf{h}_{\mathbf{j}} = \delta_{ij}. \tag{7}$$

Essa ortogonalidade implica na independência desse método ao efeito de empilhamento de sinais. Dessa maneira para o caso $N = P$, o vetor da amplitude estimada $\hat{\mathbf{a}}_{\mathbf{P}}$ é calculado pelo produto da matriz de desconvolução $\mathbf{H}_{\mathbf{P}}^{-\mathbf{1}}$ pelo vetor de pulsos de sinais $\mathbf{y}$ por meio da Eq. (8):

$$\hat{\mathbf{a}}_{\mathbf{P}} = \mathbf{H}_{\mathbf{P}}^{-\mathbf{1}}\mathbf{y}. \tag{8}$$

A matriz $\mathbf{H}_{\mathbf{P}}$ é composta pelos valores da função pulso de referência de acordo com o janelamento adotado que determina o número de amostras N empregado para cada processo de estimação. Por exemplo, ao adotar o janelamento 7, o mais simples de todos, o formato da matriz é apresentado abaixo [13].

$$\mathbf{H_P} = \begin{bmatrix} g_4 & g_5 & g_6 & g_7 & 0 & 0 & 0 \\ g_3 & g_4 & g_5 & g_6 & g_7 & 0 & 0 \\ g_2 & g_3 & g_4 & g_5 & g_6 & g_7 & 0 \\ g_1 & g_2 & g_3 & g_4 & g_5 & g_6 & g_7 \\ 0 & g_1 & g_2 & g_3 & g_4 & g_5 & g_6 \\ 0 & 0 & g_1 & g_2 & g_3 & g_4 & g_5 \\ 0 & 0 & 0 & g_1 & g_2 & g_3 & g_4 \end{bmatrix}. \tag{9}$$

Na álgebra linear, essa estrutura em que os valores das diagonais são constantes é conhecida como Toeplitz [14]. Os valores da função pulso de referência $g(t)$ para o janelamento 7 estão compreendidos no intervalo de [-75,0;75,0] ns, respectivamente: $g_1 = 0,00002304, g_2 = 0,01722640, g_3 = 0,45244500, g_4 = 1,00000000, g_5 = 0,56330700, g_6 = 0,14933500, g_7 = 0,04235980$. Ou seja, a matriz $\mathbf{H_P}$ é computada de maneira offline.

Após obtido o vetor de amplitudes estimadas, este passa por uma etapa condicional, na qual são eliminadas as linhas da matriz $\mathbf{H_P}$ associadas as amplitudes menores que 4,5 ADC Count (três vezes o valor do desvio padrão do ruído eletrônico igual a 1,5 ADC Count). O processo de estimação é efetuado novamente, desta vez calculando a pseudo-inversa da matriz $\mathbf{H_P}$.

Por fim, a amplitude estimada a_{est} é considerada como o elemento central do vetor da amplitude estimada $\hat{\mathbf{a}}_\mathbf{P}$, que sempre será mantido. Sendo assim, o erro de estimação e_{est} é obtido pela diferença entre a amplitude estimada menos a de referência a_{ref} dada pela Eq. (10):

$$e_{est} = a_{est} - a_{ref}. \tag{10}$$

2.6 Método do Melhor Estimador Linear Não Enviesado - BLUE 2

A segunda versão do BLUE busca estimar somente a amplitude central [12]. A Desconvolução de Sinal se aplica a esse método quando $P = 1$. Nesse caso simplificado, o vetor de pesos $\mathbf{G_1}$ é calculado pela Eq. (11):

$$\mathbf{G_1} = \frac{\mathbf{C^{-1}h}}{\mathbf{h^T C^{-1} h}}, \tag{11}$$

o vetor $\mathbf{h}$ é formado pelos valores do pulso de referência não deslocados, segundo o janelamento desejado.

Logo, o valor da amplitude central estimada é encontrado pelo produto interno expresso pela Eq. (12):

$$a_{est} = \mathbf{G_1^T y}. \tag{12}$$

Novamente a Eq. (10) é usada para a obtenção do erro de estimação da amplitude.

3 Resultados e discussão

Os códigos em Python elaborados na IDE Visual Studio Code, assim como os demais resultados, podem ser acessados no repositório do GitHub [15].

A Fig. 4 exibe o comportamento da média do erro de estimação da amplitude pelo K-Fold para diversos janelamentos ao longo das ocupações para a Desconvolução de Sinais.

O método de Desconvolução apresenta o menor valor de erro para o janelamento 17. Aliás, neste tamanho de janela os erros para todas as ocupações estão mais concentrados. É bastante notório o efeito do janelamento no valor do erro, até porque o janelamento inicial 7 é o que apresenta a maior média. E a medida que vai aumentando a quantidade de janelamento, o erro decai e se mantém mais estável. A exceção a regra é a ocupação 0 que é praticamente imune a variação do janelamento.

Por sua vez, no método BLUE 2, a média do erro de estimação é menor e mais uniforme para todas as ocupações no janelamento 11, conforme pode ser visto na Fig. 5. O pior caso agora não corresponde ao janelamento 7, mas ao 9, que além de apresentar valores maiores de média possuiu maior diferença dos resultados entre as ocupações, salvo novamente a ocupação 0.

Com respeito a dispersão em torno da média, a Fig. 6 demostra o procedimento do desvio padrão pelo K-Fold ao variar a quantidade do janelamento ao longo das ocupações para o método da Desconvolução de Sinais.

O método de Desconvolução de Sinais possui um intervalo de desvio padrão em torno de 1 a 45 ADC Count. Nota-se que o pior caso é para o janelamento 7. Inicialmente, o decaimento do valor da dispersão é bem acentuado entre os janelamentos 7 e 9 e depois há uma certa oscilação até que haja estabilização a partir do janelamento 17. Dessa forma, o janelamento 17 é o ótimo, tanto pela análise da média como pelo desvio padrão do erro de estimação.

Já pelo método BLUE 2, representado pela Fig. 7, a queda do desvio padrão com o aumento do janelamento é bem mais suave e os valores começam a se estabilizar a partir do 15. Embora, pela análise do erro, o janelamento

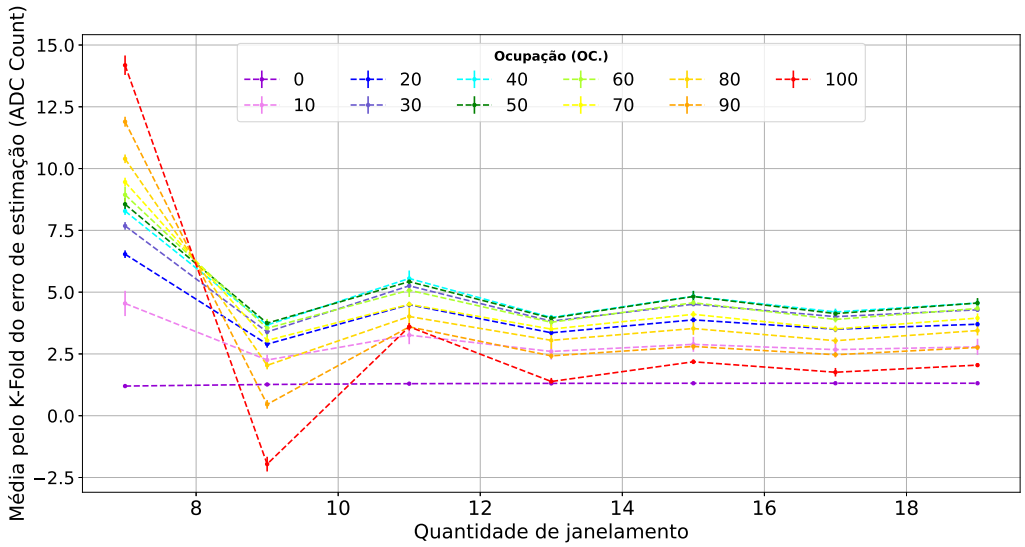


Figura 4: Média do erro de estimação da amplitude pela Desconvolução de Sinais.

Fonte: Os autores, 2024.

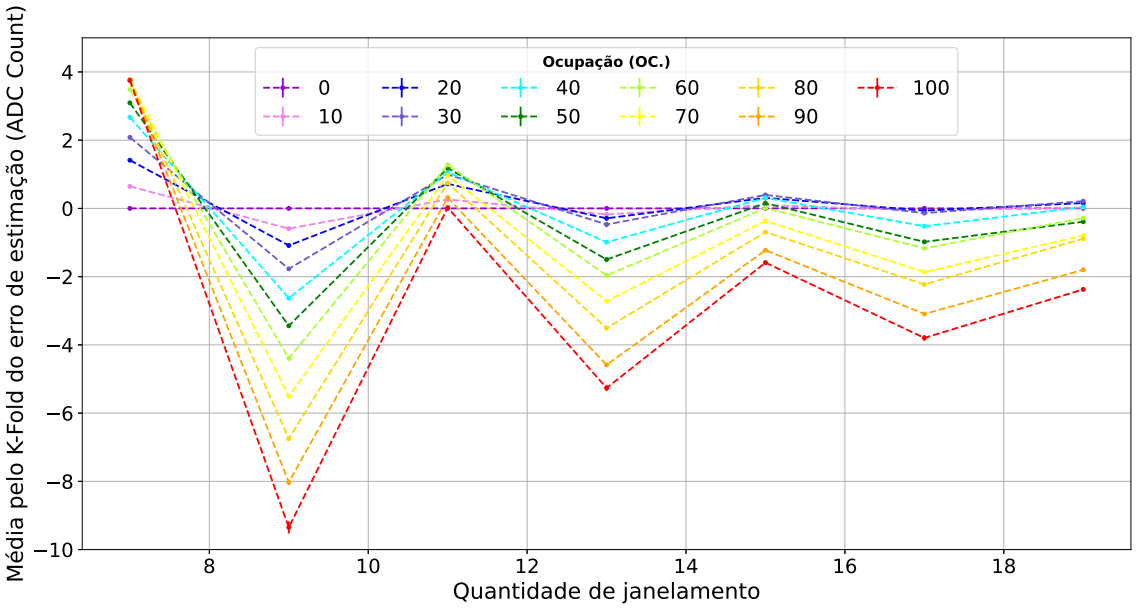


Figura 5: Média do erro de estimação da amplitude pelo BLUE 2.

Fonte: Os autores, 2024.

ótimo seja o 11, pelo estudo da dispersão o ideal seria o 15, visto que também apresenta valores médios aceitáveis e menor desvio padrão.

Observe que os melhores resultados foram para a ocupação 0, no entanto este é um caso a parte, no qual o canal de leitura não recebeu nenhum estímulo proveniente da colisão. Logo, todo o processo de estimação considerou apenas o ruído da eletrônica do equipamento que possui uma distribuição gaussiana bem controlada. Em contraste com o ruído de empilhamento, que pode ser descrito como um somatório de exponenciais.

Em relação aos histogramas do erro de estimação da amplitude, a Desconvolução apresenta o formato esperado do histograma com a média em torno de zero e uma dispersão tolerável em torno desta para ocupação 50 e janelamento ideal 17, por exemplo para a Fig. 8.

Similarmente, o mesmo acontece para o BLUE 2, só que com valores menores tanto para a média como para o desvio padrão, o que pode ser notado na Fig. 9,.

A explicação desse resultado é de que o método BLUE 2 em sua composição considera os dados de ruídos ao

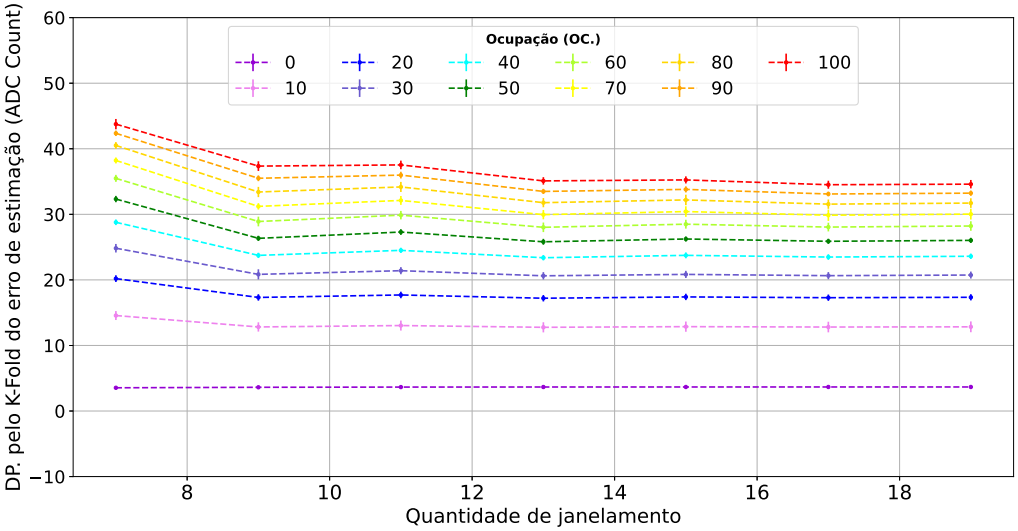


Figura 6: Desvio padrão do erro de estimação da amplitude pelo Desconvolução de Sinais.

Fonte: Os autores, 2024.

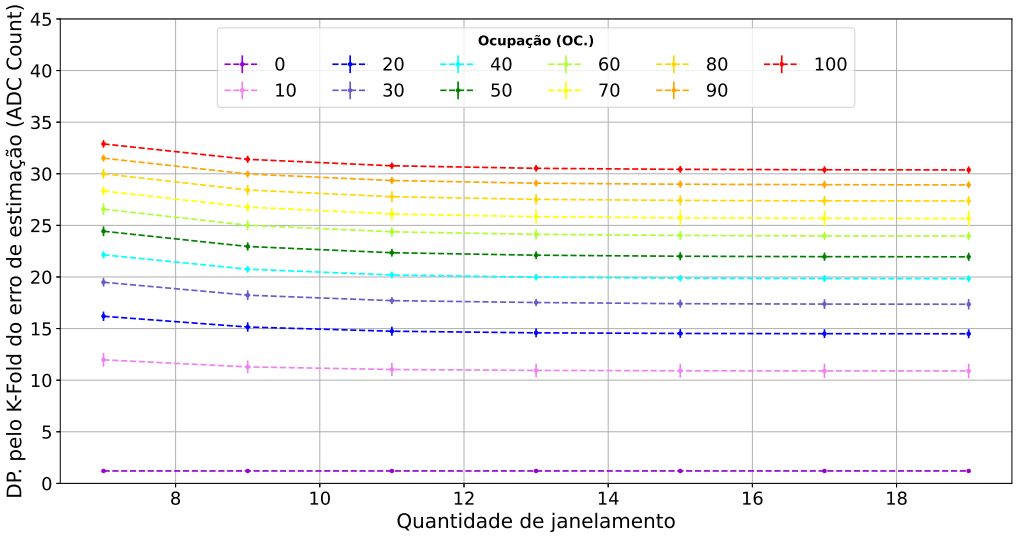


Figura 7: Desvio padrão do erro de estimação da amplitude pelo BLUE2.

Fonte: Os autores, 2024.

calcular a matriz de covariância. Em contrapartida, o método de Desconvolução é imune ao empilhamento para o caso em que $N = P$.

Isso é mais evidente nas Figs. 10 e 11 pois, embora os valores do erro de estimação médio $\left(\bar{E}_{est} = \frac{1}{N}\sum_{i=1}^{N}[e_{est}]\right)$ sejam menores no BLUE 2 para todas as ocupações, não há uma diferença significativa para o de Desconvolução que possui a grande vantagem de não necessitar dos dados de ruídos.

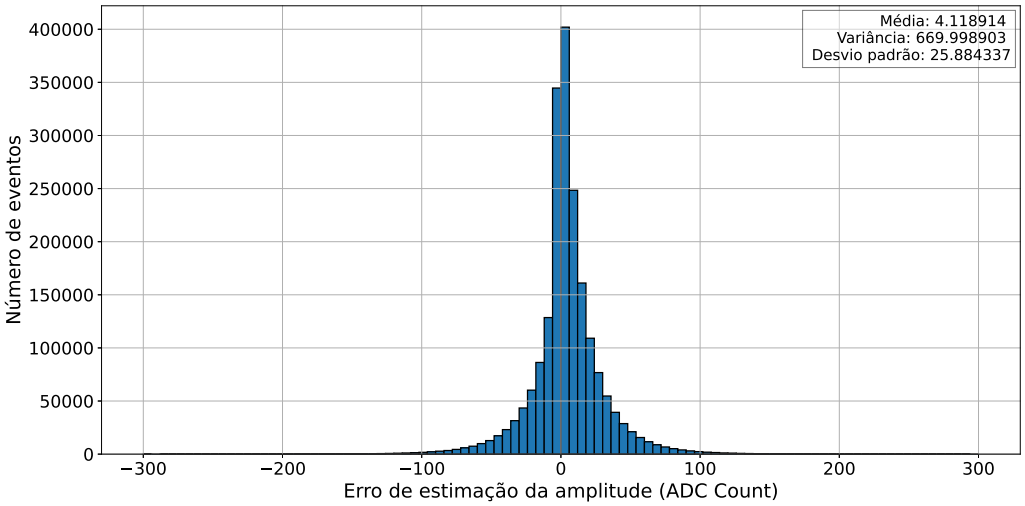


Figura 8: Histograma do erro de estimação da amplitude pelo método de Desconvolução de Sinais para a ocupação 50 % e janelamento ideal 17.

Fonte: Os autores, 2024.

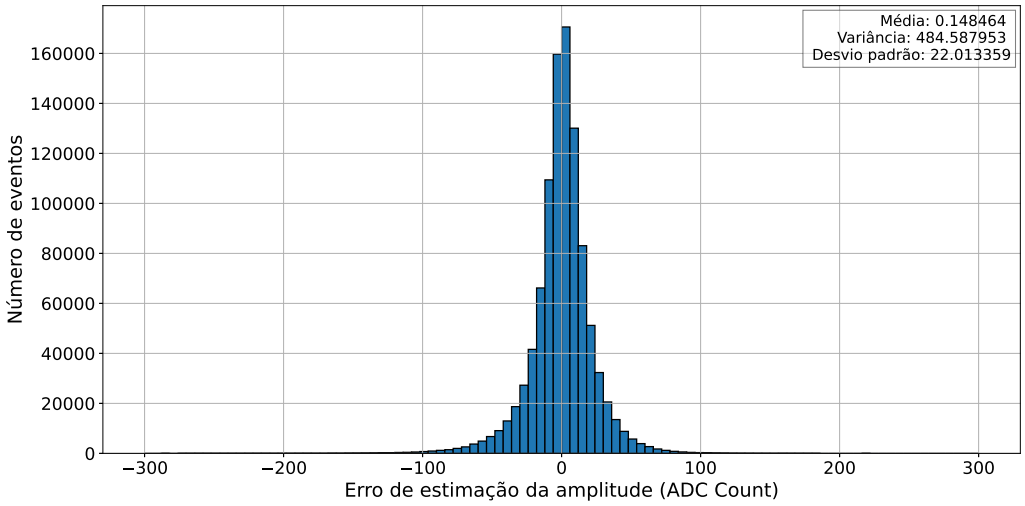


Figura 9: Histograma do erro de estimação da amplitude pelo método BLUE2 para a ocupação 50 % e janelamento ideal 15.

Fonte: Os autores, 2024.

4 Conclusão

Os métodos lineares de Desconvolução de Sinal e a segunda versão do Melhor Estimador Linear Não Enviesado (BLUE 2) apresentaram resultados satisfatórios para a estimação da amplitude para somente um canal de leitura. Os resultados para o janelamento ideal para os respectivos métodos foram 17 e 15. Uma das grandes vantagens do Desconvolução é sua independência ao efeito de empilhamento que é mais agravante com o o ambiente de alta luminosidade presentes nos experimentos de física de altas energias. O estudo deste artigo foi direcionado apenas para um canal de leitura. No entanto, como sugestão para uma abordagem mais próxima da realidade pode ser a implementação desses métodos por meio do "simulador *Lorenzetti*" que é capaz de emular os efeitos físicos das colisões das partículas, como o decaimento. Além de fornecer uma parecer mais geral do detector ao analisar os demais canais de leitura.

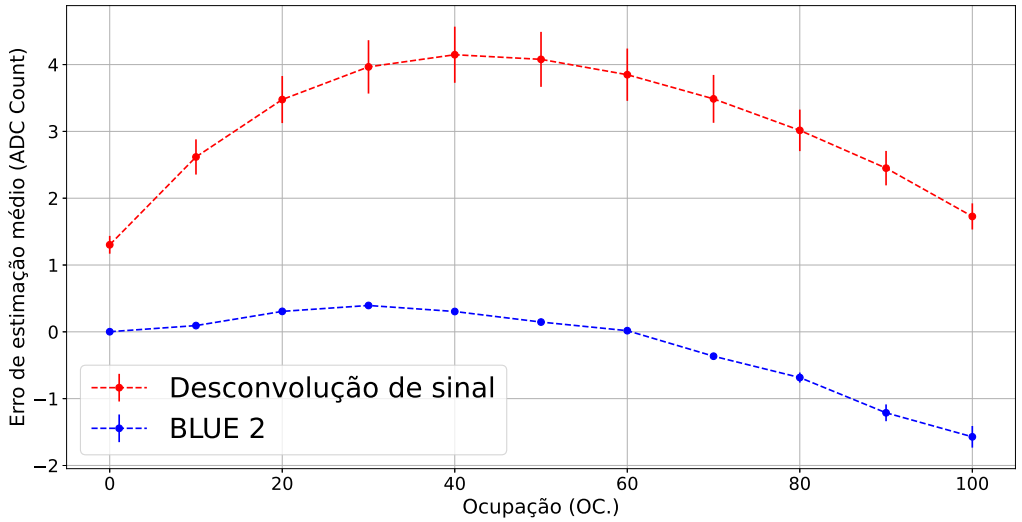


Figura 10: Comparação entre os métodos dos valores do erro de estimação médio da amplitude para todas as ocupações.

Fonte: Os autores, 2024.

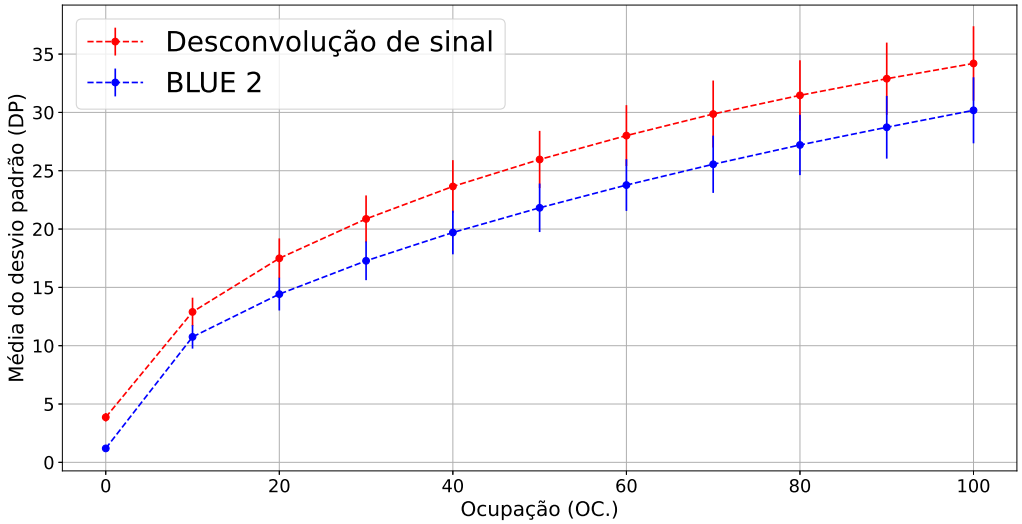


Figura 11: Comparação entre os métodos dos valores do desvio padrão do erro de estimação da amplitude para todas as ocupações.

Fonte: Os autores, 2024.

Agradecimentos

Os autores agradecem a FAPERJ, CNPq, CAPES, RENAFAE, e FAPEMIG pelo apoio. GBL é apoiado pela Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ), projeto nº E-26/210.430/2024. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Referências

[1] N. Nebeker, “Fifty years of signal processing: The IEEE signal processing society and its technologies 1948-1998,” *The IEEE Signal Processing Society*, vol. 20, 1998. Disponível em: https://signalprocessingsociety.org/uploads/history/history.pdf

[2] A. Lopes e M. L. Perrey, "Faq-lhc the guide," CERN, Genève , Switzerland, Relatório técnico 1, 2022. Disponível em: https://cds.cern.ch/record/2255762/files/CERN-Brochure-2017-002-Eng.pdf

[3] A. Rassat e M. Cirilli. (2017) From the web to a start-up near you. Página online CERN Courier. Disponível em: https://cerncourier.com/a/from-the-web-to-a-start-up-near-you/

[4] P. L. Rocca e F. Riggi, "The upgrade programme of the major experiments at the large hadron collider," *Journal of Physics: Conference Series*, vol. 515, no. 1, p. 012012, 2014. Disponível em: https://dx.doi.org/10.1088/1742-6596/515/1/012012

[5] The ATLAS Collaboration, "The atlas experiment at the cern large hadron collider," *Journal of Instrumentation*, vol. 3, no. 08, p. S08003, 2008. Disponível em: https://dx.doi.org/10.1088/1748-0221/3/08/S08003

[6] S. de Miranda Rimes, B. S.-M. Peralva, L. M. de Andrade Filho, e J. M. de Seixas, "Qualidade de dados baseada em probabilidades para reconstrução da energia em calorimetria de altas energias," *The Journal of Engineering and Exact Sciences*, vol. 9, no. 10, pp. 17 808–17 808, 2023. Disponível em: https://doi.org/10.18540/jcecvl9iss10pp17808

[7] G. I. Gonçalves, B. S.-M. Peralva, J. M. de Seixas, L. M. de Andrade Filho, e A. S. Cerqueira, "Performance of optimal linear filtering methods for signal estimation in high-energy calorimetry," *Journal of Control, Automation and Electrical Systems*, vol. 33, no. 5, pp. 1601–1611, 2022. Disponível em: https://doi.org/10.1007/s40313-022-00907-0

[8] P. Klimek e O. behalf of the ATLAS Tile Calorimeter group, "Signal reconstruction performance with the atlas hadronic tile calorimeter," *Journal of Physics: Conference Series*, vol. 404, no. 1, p. 012046, 2012. Disponível em: https://dx.doi.org/10.1088/1742-6596/404/1/012046

[9] S. de Miranda Rimes, B. S.-M. Peralva, L. M. de Andrade Filho, e J. M. de Seixas, "Qualidade de dados baseada em probabilidades para reconstrução da energia em calorimetria de altas energias," *The Journal of Engineering and Exact Sciences*, vol. 9, no. 10, pp. 17 808–17 808, 2023. Disponível em: https://doi.org/10.18540/jcecvl9iss10pp17808

[10] ScikitLearn. (2024) Cross-validation: evaluating estimator performance. Explicação técnica da validação cruzada K-Fold por meio da biblioteca sicikit-learn. Disponível em: https://scikit-learn.org/stable/modules/cross_validation.html

[11] A. V. Oppenheim, *Discrete-time signal processing*. Pearson Education India, 1999, ISBN: 0131988425.

[12] L. M. de A. Filho, B. S. Peralva, J. M. de Seixas, e A. S. Cerqueira, "Calorimeter response deconvolution for energy estimation in high-luminosity conditions," *IEEE Transactions on Nuclear Science*, vol. 62, no. 6, pp. 3265–3273, 2015. Disponível em: https://doi.org/10.1109/TNS.2015.2481714

[13] S. d. M. Rimes, "Filtragem inversa não-linear para estimação de sinais em calorímetros operando a alta taxa de eventos," Dissertação de mestrado, Universidade do Estado do Rio de Janeiro (UERJ), 2021. Disponível em: https://www.bdtd.uerj.br:8443/handle/1/16415

[14] F. Szabo, *The linear algebra survival guide: illustrated with Mathematica*. Academic Press, 2015.

[15] G. Morett. (2024) Repositório do método de desconvolução de sinais. GitHub. Disponível em: https://github.com/MorettGuilherme?tab=repositories