REVISTA Mundi
Engenharia, Tecnologia e Gestão
INSTITUTO FEDERAL Paraná

PERFORMANCE EVALUATION OF ENERGY RECONSTRUCTION METHODS IN HIGH ENERGY PHYSICS EXPERIMENTS

AVALIAÇÃO DE PERFORMANCE DE MÉTODOS DE RECONSTRUÇÃO DE ENERGIA EM EXPERIMENTOS DE FÍSICA DE ALTAS ENERGIAS

Juan Lieber Marin[1]

Eduardo Furtado de Simas Filho [2]

Bernardo Sotto-Maior Peralva [3]

Guilherme Inácio Gonçalves[4]

Luciano Manhães de Andrade Filho[5]

José Manoel de Seixas[6]

Abstract: The discovery of particles that shape our universe pushes the scientific community to increasingly build sophisticated equipment. Particle accelerators are one of these complex machines that put known particle beams on a collision course at speeds close to that of light. When collisions occur, subproducts are produced and measured by the calorimeter system, which entirely absorbs these subproducts. Typically, a high-energy calorimeter is highly segmented, comprising thousands of dedicated readout channels. The present work evaluates the performance of two energy reconstruction algorithms: the OF (Optimal Filter) and MAE (Multi-Amplitude Estimator), which was recently proposed to deal with the signal superposition (pile-up). In order to evaluate the energy estimation efficiency, artificial data were used, considering several pile-up levels. The statistics from the energy estimation is employed to compare the performance achieved by each method. A second analysis is made to quantify the MAE sensitivity to the pedestal parameter. The results show that the MAE method presents a better performance than the OF method and the usage of an uncalibrated pedestal value compromises the MAE performance.

***Keywords*:** High-energy physics. Parameter estimation. Optimal filtering. Signal pile-up.

[1]Doutorando em Engenharia Elétrica, PPGEE-UFBA, juan.lieber@ufba.br
[2]Doutor em Engenharia Elétrica, PPGEE-UFBA, eduardo.simas@ufba.br
[3]Doutor em Engenharia Elétrica, PMC-UERJ, bernardo@iprj.uerj.br
[4]Mestrando em Modelagem Computacional, PMC-UERJ, inacio.guilherme@gmail.com
[5]Doutor em Engenharia Elétrica, PGEE-UFJF, luciano.andrade@engenharia.ufjf.br
[6]Doutor em Engenharia Elétrica, PEE-COPPE-UFRJ, seixas@lps.ufrj.br

Resumo: A descoberta de partículas que compõem o nosso universo leva a comunidade científica a construir equipamentos cada vez mais sofisticados. Aceleradores de partículas são algumas dessas máquinas complexas que aceleram feixes de partículas com velocidade próxima à velocidade da luz. Quando as colisões ocorrem, os subprodutos são produzidos e medidos por um conjunto de detectores, com destaque para o sistema de calorimetria, que absorve completamente estes subprodutos (exceto muons e neutrinos). Tipicamente, um calorímetro de altas energias é altamente segmentado, contendo dezenas (ou centenas) de milhares de canais de leitura dedicados. O presente trabalho avalia a performance de dois algoritmos de reconstrução de energia: o OF (*Optimal Filter*) e o MAE (*Multi-Amplitude Estimator*), este último proposto recentemente para lidar com a sobreposição de sinais (empilhamento), que resulta da alta taxa de eventos. Para avaliar a eficiência da estimação de energia, dados sintéticos foram produzidos, considerando diversos níveis de empilhamento. As estatísticas da energia estimada são empregadas para comparar a performance alcançada por cada método. Uma segunda análise é feita para quantificar a sensibilidade do MAE com relação ao parâmetro de pedestal. Os resultados mostram que o método MAE apresenta uma melhor performance quando comparado ao método OF, mas o uso de um valor de pedestal descalibrado compromete a sua performance.

***Palavras-chave*:** Física de altas energias. Estimação paramétrica. Filtragem ótima. Empilhamento de sinais.

1 INTRODUCTION

Since the most remote times, the humanity searches for the universe origin and composition. After the electron discovery, a large number of subatomic particles have been identified and their properties were exhaustively explored (MARTIN, 2006). A number of modern experiments have been using high-energy particle colliders to understand the fundamental composition of the universe (GRIFFITHS, 2008). The Large Hadron Collider (EVANS; BRYANT, 2008) and the Tevatron (DROZHDIN; MOKHOV, 1998) are examples of these machines. To provide a precise particle identification, the experimental apparatus is composed by different subsystems. Among them, the calorimeter system plays an important role as it measures the energy of the incoming particles and participates intensively on particle identification tasks (WIGMANS, 2017).

Usually, the calorimeter system comprises both electromagnetic and hadronic sections: for electrons and photons, the calorimeters have to provide the electromagnetic interaction information, and, for hadrons and jets, a longer and wider calorimeter is required (hadronic section) (FABJAN, 1994). The calorimeters usually comprises tens or hundred of thousands readout channel which provides good precision for particle position and incidence angle measurements. The signal generated by each readout channel depends of the calorimeter principle and a signal shaper is used to condition the signal, so that the measured energy can be retrieved simply from the signal amplitude. Therefore, an amplitude estimation algorithm is required for signal reconstruction.

In particle collider experiments, the data acquisition system clock is usually synchronized with the experiment collision period. For instance, the LHC (Large Hadron Collider) , which is operating now at CERN (European Center for Nuclear Research) and is the most powerful collider machine ever built, collides a bunch of protons each 25 nanoseconds. In such high event rate environment, although calorimeters are fast detectors, they would need a couple of bunch crossing periods to produce their final response. In case the particle collider luminosity is considerably high, the detector occupancy may make a signal from a given bunch crossing collision to suffer influence from signals coming from other bunch crossings, so that signal pileup may arise (POLUSHKIN, 2004). The presence of such out-of-time (OOT) signal may deteriorate the performance of the energy estimation task (PERALVA et al., 2017).

Energy reconstruction algorithms perform the energy estimation by taking the weighted sum of digitized signal samples. Also, the digitization process includes, a baseline (pedestal) in the electric pulse, which is calculated using a calibration procedure and its value is usually stored in a database system (WIGMANS, 2017).

One of the first methods proposed in the high energy physics experiment context is the Optimal Filter (OF) (CLELAND; STERN, 1994). This method aims to minimize the electronic noise contribution, which is usually Gaussian (KAY, 1993). However, when the pile-up effect increases, the OF method consider the pileup as an additive noise, although pileup noise is no longer Gaussian. Therefore, the pile-up effect makes the OF algorithm to be suboptimal in estimation performance.

An alternative method, called Multi-Amplitude Estimator (MAE), was proposed in (Peralva, 2013; FILHO et al., 2015) to handle the pile-up effect. In the MAE approach, the overlapped signals are no longer Incorporated in the signal model as an additive noise. Instead, a linear deconvolution transformation is applied to the digitized samples to recover the overlapped signal amplitudes and only the signal of interest (and its energy) is used for reconstruction from a particular collision. As a consequence, the noise contribution on the MAE signal model comes only from the electronic noise.

In this work, a performance comparison between OF and MAE method is presented. To this purpose, simulated data for several pile-up conditions were generated. It will be also evaluated the sensitivity of the MAE method to the pedestal. The text is organized as follows: Section 2 describes both OF and MAE methods. Section 3 presents details of the data used and the results of the performance analysis and the MAE pedestal sensitivity. Conclusions are presented in section 4.

2 ENERGY RECONSTRUCTION METHODS

Usually, the calorimeter systems operates with a fixed pulse shape. An example is the unipolar pulse, shown in Fig. 1.

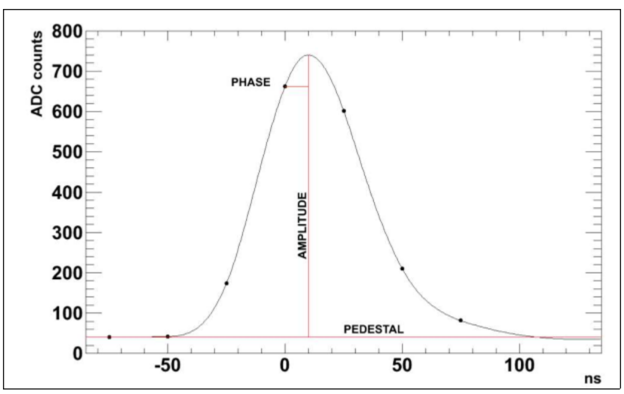

Figure 1: Typical unipolar pulse for energy reconstruction (extracted from (Peralva, 2013))

.

The energy diagram estimation process is shown in Fig. 2. A typical calorimeter data acquisition system operates sampling the received signal $\mathbf{y}(n)$, in order to estimate the signal amplitude A.

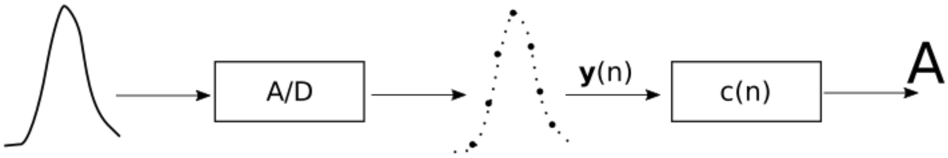

Figure 2: Block diagram of energy reconstruction process.

2.1 Optimal Filter

The Optimal Filter was proposed in (CLELAND; STERN, 1994) and it is based on a weighted sum of the incoming time samples. In order to compute the Optimal Filter coefficients, the OF method models the received signal as:

$$y(n) = Ah(n) + A\tau\dot{h}(n) + ped + w(n), \tag{1}$$

where A is the pulse amplitude, τ is the phase deviation, ped is the baseline added to the signal and $w(n)$ the electronic noise, usually modeled as Gaussian distribution. The signal $h(n)$ is the normalized reference pulse shape ($\dot{h}(n)$ is its time derivative). The amplitude A is obtained as follows:

$$\hat{A}_{OF} = \sum_{n=1}^{N} y(n)c(n), \tag{2}$$

where N is the number of digitized samples and $c(n)$ are the OF coefficients, which are calculated through an optimization process using the Lagrange Multipliers method, in order to minimize the noise contribution in the amplitude estimation procedure. For this purpose, the following constraints may be imposed:

$$\sum_{n=1}^{N} c(n)h(n) = 1, \tag{3}$$

$$\sum_{n=1}^{N} c(n)\dot{h}(n) = 0, \tag{4}$$

$$\sum_{n=1}^{N} c(n) = 0, \tag{5}$$

DOI: 10.21575/25254782rmetg2020vol5n21169

$$\sum_{n,j=1}^{N} c(n)C(n,j) - \lambda h(n) - \epsilon \dot{h}(k) - \kappa = 0. \tag{6}$$

The parameters λ, ϵ and κ are the Lagrange multipliers and $C(n, j)$ corresponds to the noise covariance matrix elements. Since the OF uses the noise covariance information to minimize the signal uncertainties, its precision is considerably degraded in signal pile-up conditions as signal pile-up is not properly modeled by Eq. 1. Fig. 3 illustrates the signal pile-up effect. The pulse in black corresponds to the signal of interest to be estimated, while the pulse in blue comes from a time-shifted collision that was acquired within a same readout window, distorting the output pulse (in red).

2.2 Multi-Amplitude Estimator

MAE computes a linear transformation that recovers the amplitude of the superimposed signals for a given readout window, so that the central pulse becomes assigned to the collision of interest and can be decoupled and reconstructed. In this context, the MAE method assumes that the energy deposition from each collision is an input of linear time-invariant system (OPPENHEIM; SCHAFER, 1989). The received signal results from the convolution of the received amplitude vector from the acquired time samples $a(n)$, the impulsive system response $h(n)$:

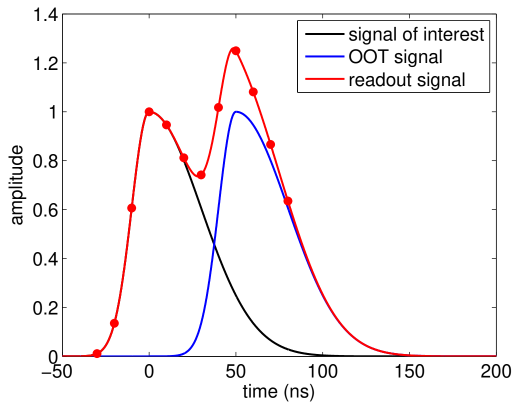

Figure 3: An illustration of the pile-up effect. The black pulse is the signal of interest and the blue one is the time-shifted signal. The resulting received pulse is in red (extracted from (PERALVA et al., 2017)).

$$y(n) = \sum_{i} (h(i)a(n-i)) + w(n), \tag{7}$$

where i is the convolutional auxiliary variable (for example, $0 \leq i \leq 6$) (OPPENHEIM; SCHAFER, 1989). Here, the noise $w(n)$ is Gaussian, since the pile-up contribution is already

DOI: 10.21575/25254782rmetg2020vol5n21169

included in the signal model, not contributing to the noise term. Therefore, to obtain an estimation for the deposited energy, it is necessary to apply a deconvolution process to estimate the $a(n)$ coefficients. The method equations are:

$$\hat{\mathbf{a}}_p = \mathbf{G}_p^T \mathbf{y}, \tag{8}$$

where:

$$\mathbf{G}_p = \mathbf{C}_p^{-1} \mathbf{H}_p (\mathbf{H}_p^T \mathbf{C}_p^{-1} \mathbf{H}_p)^{-1}, \tag{9}$$

and

$$\mathbf{H}_p^T = \begin{bmatrix} h(3) & h(4) & h(5) & h(6) & 0 & 0 & 0 \\ h(2) & h(3) & h(4) & h(5) & h(6) & 0 & 0 \\ h(1) & h(2) & h(3) & h(4) & h(5) & h(6) & 0 \\ h(0) & h(1) & h(2) & h(3) & h(4) & h(5) & h(6) \\ 0 & h(0) & h(1) & h(2) & h(3) & h(4) & h(5) \\ 0 & 0 & h(0) & h(1) & h(2) & h(3) & h(4) \\ 0 & 0 & 0 & h(0) & h(1) & h(2) & h(3) \end{bmatrix}, \tag{10}$$

where p ($0 \leq p \leq N$) is the number of collisions within a readout window and $\mathbf{C}$ is the noise correlation matrix. The algorithm performs in two steps: firstly, considering $p = N$, the method achieves the best approximation for the deconvolution process, estimating N signals within the readout window. It is worth mentioning that if the $\mathbf{C}$ matrix can be neglected (white noise, $\mathbf{C} \approx \mathbf{I}$), the energy estimation process for N signals becomes:

$$\hat{\mathbf{a}} = \mathbf{H}_p^{-1} \mathbf{y}. \tag{11}$$

Once the N amplitudes are estimated, the MAE method compares all the $a(n)$ values with a predefined threshold (usually associated to the electronic noise variance) and the amplitudes above the threshold are selected. Thus, $\mathbf{H}_p^T$ is computed again to improve the signal deconvolution process, calculating only the components that probably have information of interest and not only noise contribution. Before computing the (11), the method subtract the pedestal value, which is usually retrieving such value from a database

3 RESULTS

The performance comparison analysis uses simulated data, using the Calo Pulse Kit package, available in (GONCALVES, 2019). In this dataset, several pileup levels are simulated by setting the probability of an incident particle to hit a given cell, which defines the cell occupancy. The greater this probability is more severe is the pileup effect and, consequently, the distortion in the received signal. The occupancy range is 10% to 90%. For a fixed occupancy, there are 30 subsets, each one containing 142857 pulses, which have each one seven samples. The deposited energy in each pulse follows an exponential distribution with rate parameter equal to 480 MeV. The simulation considers a single readout channel, which generates unipolar pulses corrupted by a white Gaussian noise with $\sigma = 12$MeV. The pedestal value added to all pulses is 20 ADC counts.

3.1 Performance Comparison

In order to provide a numerical comparison, the percentage deviation of the RMS estimation error will be employed, considering both methods:

$$\sigma_z = 100 \times \frac{\sigma_{OF} - \sigma_{MAE}}{\sigma_{OF}} \tag{12}$$

where σ_{OF} the RMS estimation error associated to the OF performance and σ_{MAE} the RMS estimation error associated to the MAE performance. When $\sigma_z > 0$ the MAE method exhibits better performance with respect to the OF algorithm and $\sigma_z < 0$, otherwise. The Fig. 4 shows the σ_z evaluation for all the occupancy values. Each boxplot was generated using the 30 subsets, each one with 142857 pulses.

In Fig. 4, it is possible to see that the MAE method produces a smaller RMS value than the OF algorithm, due to the positive values of the σ_z. It can be noticed that the σ_z achieves the maximum value around to 30% occupancy value. As the occupancy keeps increasing, smaller is the difference between both methods in term s of estimation performance. For occupancy values between 20% and 30%, the MAE achieves a performance approximately 20% times better than the OF method. This is an expected result, since the MAE incorporates the pile-up effect in its signal model.

3.2 Pedestal Sensitivity

The Fig. 5 shows the σ_z measurement as a function of the pedestal. Here, the occupancy value was kept fixed fixed in 30%. To explore the MAE sensitivity to the pedestal value, a deviation range from -15 to 20 around the reference value was applied. For each pedestal value, a single dataset with occupancy equals to 30% was used, considering all the 142857 pulses.

The OF method is not affected by a pedestal value changing, since the Eq. 2 doesn't subtract the pedestal value from received signal. Therefore, in this pedestal sensitivity analysis, the evaluation of the σ_z represents only a changing in the σ_{MAE}. In the Fig. 5, the maximum

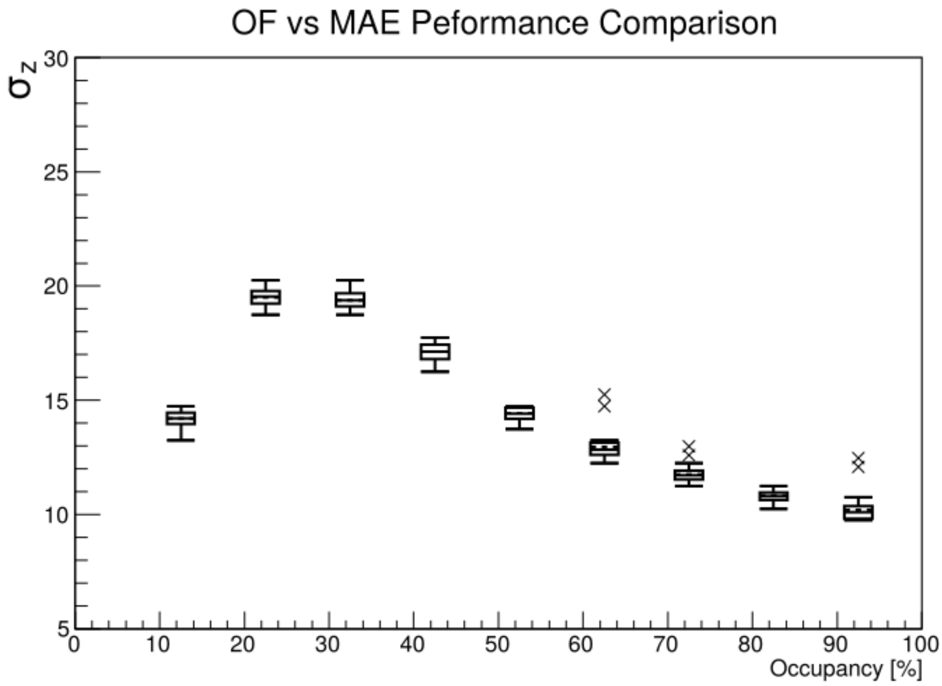

Figure 4: Comparison between the OF and MAE methods. A positive value of σ_z indicates that the MAE presented a better performance than the OF.

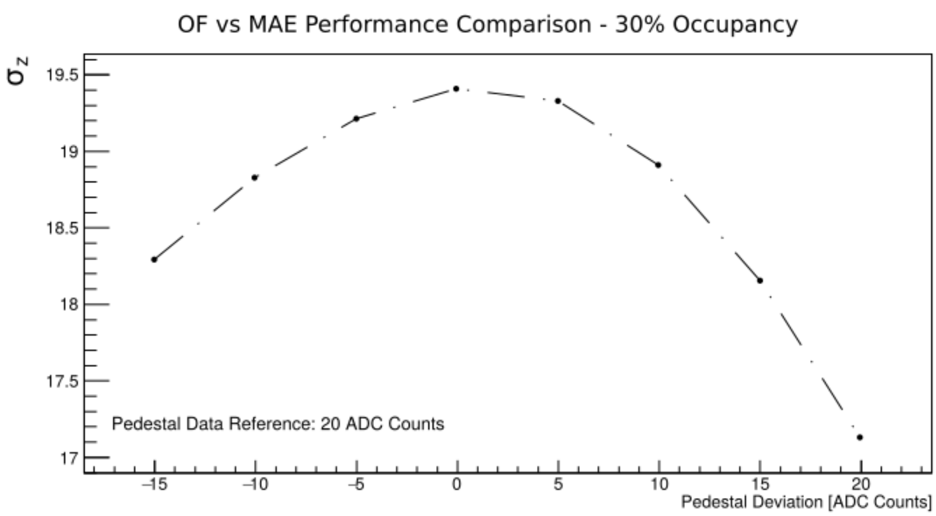

Figure 5: Sensitivity of the MAE method to an uncalibrated pedestal value. For reference, the data were generated using a pedestal equals to 20 ADC Counts.

DOI: 10.21575/25254782rmetg2020vol5n21169

value of σ_z is associated to the real pedestal value added to each pulse (20 ADC Counts). As the pedestal value used in the MAE equations changes, the performance of the MAE method gets worse, showing the sensitivity of the pedestal value used to compute Eq. 11. A wrong pedestal value represents a problem in the value stored in the database system. Usually calibration procedures are responsible to keep the database well updated to avoid a bad performance of the energy reconstruction methods. Therefore, the good performance of the MAE method relies on an accurate pedestal estimation from specifics runs.

4 CONCLUSIONS

Searching for new physics, modern high event rate experiments pushing the calorimeters design to the limit in order to deal with the immense amount of data that is produced, moreover when out of time pileup effects arise in signal reconstruction. The standard OF method is not capable of handling the high pileup scenario, while the recently proposed MAE method was actually designed to deal with such conditions, incorporating the signal pileup in its signal description model. This work presented a performance comparison between both the OF and MAE methods, using a high pile-up simulation dataset and considering a single calorimeter readout channel. It was shown that the MAE algorithm improves considerably the energy estimation efficiency with respect to the current OF approach. It was also observed that poor pedestal value estimation degrades the MAE performance, so that the MAE method relies on pedestal runs for producing good pedestal values for the thousands of calorimeter channels usually employed in modern experiments.

ACKNOWLEDGMENTS

The authors thanks to CNPq, RENAFAE (MCTIC), FAPERJ, FAPEMIG and FAPESB for the financial support. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

REFERENCES

CLELAND, W.; STERN, E. 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*, v. 338, n. 2, p. 467 – 497, 1994. ISSN 0168-9002. Disponível em: <http://www.sciencedirect.com/science/article/pii/0168900294913323>.

DROZHDIN, A. I.; MOKHOV, N. V. Tevatron Run-II Beam Collimation System. 1998.

EVANS, L.; BRYANT, P. LHC machine. *Journal of Instrumentation*, IOP Publishing, v. 3, n. 08, p. S08001–S08001, aug 2008.

FABJAN, C. W. *Detectors for elementary particle physics*. [S.l.: s.n.], 1994. 77-116 p.

FILHO, L. de A. et al. Calorimeter response deconvolution for energy estimation in high-luminosity conditions. *IEEE Transactions on Nuclear Science*, v. 62, p. 1–9, 11 2015.

GONCALVES, G. I. *Calo Pulse Kit*. 2019. Disponível em: <https://gitlab.cern.ch/ginaciog/calopulsekit>.

GRIFFITHS, D. J. *Introduction to elementary particles; 2nd rev. version*. New York, NY: Wiley, 2008. (Physics textbook). Disponível em: <https://cds.cern.ch/record/111880>.

KAY, S. *Fundamentals of Statistical Signal Processing: Detection theory*. PTR Prentice-Hall, 1993. (Fundamentals of Statistical Signal Processing). ISBN 9780133457117. Disponível em: <https://books.google.com.br/books?id=aFwESQAACAAJ>.

MARTIN, B. R. *Nuclear and Particle Physics*. [S.l.]: John Wiley & Sons, Ltd, 2006.

OPPENHEIM, A. V.; SCHAFER, R. W. *Discrete-time signal processing*. [S.l.]: Prentice Hall, 1989.

PERALVA, B. et al. Nonlinear correction for an energy estimator operating at severe pile-up conditions. In: ____. [S.l.: s.n.], 2017. v. 664, p. 87–107.

Peralva, B. S. The tilecal energy reconstruction for collision data using the matched filter. In: *2013 IEEE Nuclear Science Symposium and Medical Imaging Conference (2013 NSS/MIC)*. [S.l.: s.n.], 2013. p. 1–6. ISSN 1082-3654. Available at <https://cds.cern.ch/record/1629575>.

POLUSHKIN, V. *Nuclear Electronics: Superconducting Detectors and Processing Techniques*. Wiley, 2004. ISBN 9780470857687. Disponível em: <https://books.google.hn/books?id=88LWf2pPjbYC>.

WIGMANS, R. *Calorimetry: Energy Measurement in Particle Physics*. [S.l.]: Clarendon Press, 2017. (International series of monographs on physics). ISBN 9780198502968.

Edição especial - XXII ENMC (Encontro Nacional de Modelagem Computacional) e X ECTM (Encontro de Ciência e Tecnologia dos Materiais)

Enviado em: 31 mar. 2020

Aceito em: 13 mai. 2020

Editor responsável: Rafael Alves Bonfim de Queiroz