The TileCal Online Energy Estimation for the Next LHC Operation Period

B Sotto-Maior Peralva on behalf of the ATLAS Collaboration

Federal University of Juiz de Fora, Brazil

E-mail: `bernardo@cern.ch`

Abstract.
The ATLAS Tile Calorimeter (TileCal) is the detector used in the reconstruction of hadrons, jets and missing transverse energy from the proton-proton collisions at the Large Hadron Collider (LHC). It covers the central part of the ATLAS detector ($|\eta| < 1.6$). The energy deposited by the particles is read out by approximately 5,000 cells, with double readout channels. The signal provided by the readout electronics for each channel is digitized at 40 MHz and its amplitude is estimated by an optimal filtering algorithm, which expects a single signal with a well-defined shape. However, the LHC luminosity is expected to increase leading to pile-up that deforms the signal of interest. Due to limited resources, the current hardware setup, which is based on Digital Signal Processors (DSP), does not allow the implementation of sophisticated energy estimation methods that deal with the pile-up. Therefore, the technique to be employed for online energy estimation in TileCal for next LHC operation period must be based on fast filters such as the Optimal Filter (OF) and the Matched Filter (MF). Both the OF and MF methods envisage the use of the background second order statistics in its design, more precisely the covariance matrix. However, the identity matrix has been used to describe this quantity. Although this approximation can be valid for low luminosity LHC, it leads to biased estimators under pile-up conditions. Since most of the TileCal cell present low occupancy, the pile-up, which is often modeled by a non-Gaussian distribution, can be seen as outlier events. Consequently, the classical covariance matrix estimation does not describe correctly the second order statistics of the background for the majority of the events, as this approach is very sensitive to outliers. As a result, the OF (or MF) coefficients are miscalculated leading to a larger variance and biased energy estimator. This work evaluates the usage of a robust covariance estimator, namely the Minimum Covariance Determinant (MCD) algorithm, to be applied in the OF design. The goal of the MCD estimator is to find a number of observations whose classical covariance matrix has the lowest determinant. Hence, this procedure avoids taking into account low likelihood events to describe the background. It is worth mentioning that the background covariance matrix as well as the OF coefficients for each TileCal channel are computed offline and stored for both online and offline use. In order to evaluate the impact of the MCD estimator on the performance of the OF, simulated data sets were used. Different average numbers of interactions per bunch crossing and bunch spacings were tested. The results show that the estimation of the background covariance matrix through MCD improves significantly the final energy resolution with respect to the identity matrix which is currently used. Particularly, for high occupancy cells, the final energy resolution is improved by more than 20%. Moreover, the use of the classical covariance matrix degrades the energy resolution for the majority of TileCal cells.

1. Introduction

The ATLAS [1] Tile Calorimeter (TileCal) [2] is the detector used in the reconstruction of hadrons, jets and missing transverse energy from the proton-proton collisions at the Large Hadron Collider (LHC) [3]. It is composed by one central and two extended barrels covering the most central part of the ATLAS detector ($|\eta| < 1.6$). The energy deposited by the particles is read out by approximately 5,000 cells, with double readout channels. Figure 1 shows the TileCal readout granularity which consists of three radial layers (A,BC and D) with $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ (0.2×0.1 in outermost layer). There is also the E layer, which consists of four scintillator plates per module and are located in the extended barrels.

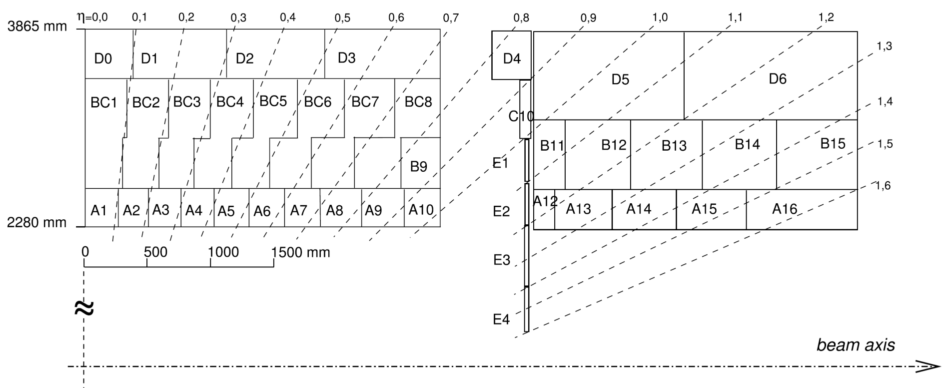

Figure 1. TileCal cell segmentation for half barrel and one extended partition.

The signal provided by the readout electronics [4] for each channel is digitized at 40 MHz and its amplitude is estimated by an optimal filtering algorithm, which expects a single signal with a well-defined shape. With the increase of the luminosity during Run2, the pile-up signal deforms the signal of interest. For illustration purposes, Figure 2 shows the readout window with the expected signal pulse shape (black line) and an out-of-time signal (red line). The resultant signal (magenta line) corresponds to the signal that is acquired by the front-end electronics.

Due to limited resources, the current Digital Signal Processors (DSP) based hardware setup does not allow the implementation of very sophisticated energy estimation methods, therefore, the online energy estimation must be based on fast filters such as the Optimal Filter (OF) [5]. The OF method uses the background second order statistics in its design, more precisely the covariance matrix. The identity matrix was used during Run1 to describe this quantity, although this approximation (that is valid for low luminosity LHC) can lead to larger variance and biased estimators under pile-up conditions, as it will be shown in Section 3.2.

This work evaluates the use of the background covariance matrix in the design of the OF weights to improve energy estimation performance under pile-up conditions. Next section describes the TileCal energy estimation algorithm and how we plan to improve it for LHC Run2. Section 3 shows the results using a simulated data set, where different versions of the OF method are evaluated. Finally, the conclusions are presented in Section 4.

2. TileCal Energy Estimation

The OF method was implemented to operate both online and offline in TileCal during Run1. The algorithm is based on a weighted sum of the received Analog to Digital Converter (ADC)

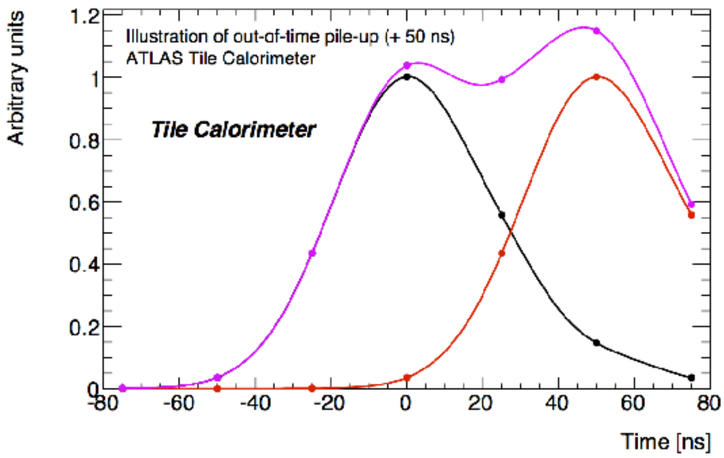

Figure 2. Illustration of the pile-up effect in TileCal.

samples aiming at minimizing the variance of the signal amplitude (see Equation1).

$$\hat{A} = \sum_{i=1}^{N} s_i w_i \tag{1}$$

The vectors **s** and **w** correspond to the received ADC samples and the OF weights, respectively. The parameter N is the number of samples available.

The variance of the amplitude parameter to be minmized is given by:

$$var(\hat{A}) = \mathbf{w'Cw} \tag{2}$$

where the matrix **C** corresponds to the background covariance matrix.

The current implementation of the OF in TileCal is called OF2, and it performs the optimization procedure subjected to the following constraints:

$$\sum_{i=1}^{N} g_i w_i = 1 \tag{3}$$

$$\sum_{i=1}^{N} g_i' w_i = 0 \tag{4}$$

$$\sum_{i=1}^{N} w_i = 0 \tag{5}$$

where the vectors **g** and **g**' are the TileCal reference pulse shape and its derivative, respectively.

The first constraint (Equation 3) regards the energy scale factor, while the additional second and third constraints (Equations 4 and 5) are added to make the estimation procedure immune against phase and baseline fluctuations, respectively.

The weights $\boldsymbol{w}$ can be found by solving the following matrix system using Lagrange multipliers:

$$\begin{pmatrix} C_{1,1} & \dots & C_{1,7} & -g_1 & -g_2' & -1 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \\ C_{7,1} & \dots & C_{7,7} & -g_7 & -g_7' & -1 \\ g_1 & \dots & g_7 & 0 & 0 & 0 \\ g_1' & \dots & g_7' & 0 & 0 & 0 \\ 1 & \dots & 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_7 \\ \lambda \\ \xi \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \tag{6}$$

where λ, ξ, v are the Lagrange multipliers.

Since the electronic noise can be modeled as an uncorrelated Gaussian process, the covariance matrix **C** was replaced by an identity matrix during LHC Run1 data taking.

If the third constraint (Equation 5) is removed from the optimization procedure, we call this method OF1. The difference is that OF2 computes the baseline value in an event-by-event basis, while OF1 relies on the stability of the baseline, and it subtracts a fixed value from each incoming ADC sample (see Equation 7).

$$\hat{A} = \sum_{i=1}^{N} (s_i - ped) w_i \tag{7}$$

The constant value *ped* is computed through a special run and stored in a data base.

It is worth mentioning that OF1 achieves similar performance as the Gaussian Matched Filter (MF) based approach, which was recently proposed for TileCal energy estimation [6]. The MF method is developed from the likelihood ratio test and, consequently, can be designed for any background model. However, for non-Gaussian background, its design leads to a nonlinear digital circuit, which is difficult to implement.

2.1. Online Energy Estimation for Run2

As the pile-up introduces correlation between the samples, the use of the background covariance matrix is expected to improve the performance of the OF method under pile-up conditions. Moreover, in TileCal, the occupancy of most of the readout cells are low, and the pile-up signal is considered as outlier for the majority of the cells. Additionally, since the covariance matrix is very sensitive to outliers, alternative ways of computing this quantity must be considered.

The classical covariance matrix estimation [7] takes into account the whole dataset regarless the presence of outliers, leading to the miscalculation of the OF weights. As a result, the OF weights can be miscalculated leading to a larger variance and biased energy estimator.

Alternatively, the Minimum Covariance Determinant Estimator (MCDE) [8] provides a more careful way of computing the covariance matrix. The MCDE algorithm takes randomly a subset of the noise events and computes its classical covariance matrix and its determinant. The algorithm repeats this procedure several times and it selects the subset that resulted in the lowest determinant. This subset contains the events that has the lowest covariance between the samples (ADC digits) and therefore it consists of the most probable events, disgarding most of the outliers (high energy pile-up in this case).

3. Results

This section presents the data set used and the performance evaluation is carried out in terms of the estimation error for both OF1 and OF2 methods. The OF1 and OF2 weights were designed using the identity matrix as well as the covariance matrix, computed using the MCDE algorithm.

3.1. Data set

The data set consists of 100,000 events for TileCal E3 ($\eta = 1.3$) and E4 ($\eta = 1.5$) cells, which are the highest occupancy cells in TileCal. The events are full ATLAS Monte Carlo simulation of Minimum Bias (MB) events with 25 ns of bunch spacing and an averaged number of $p-p$ interactions per bunch crossing ($<\mu>$) of 40 [9]. Only MB pile-up signals (both in-time and out-of-time) and electronic noise are present in the events.

3.2. Performance evaluation

The designed OF1 and OF2 were applied to the data set described previously. Since the events considered above comprise only background, the energy estimation is expected to be centered at zero, with the smallest dispersion as possible. These distributions can be seen as the error introduced by each method to the final energy estimate, under this condition of pile-up. Figure 3 shows the energy estimation for the E3 cell.

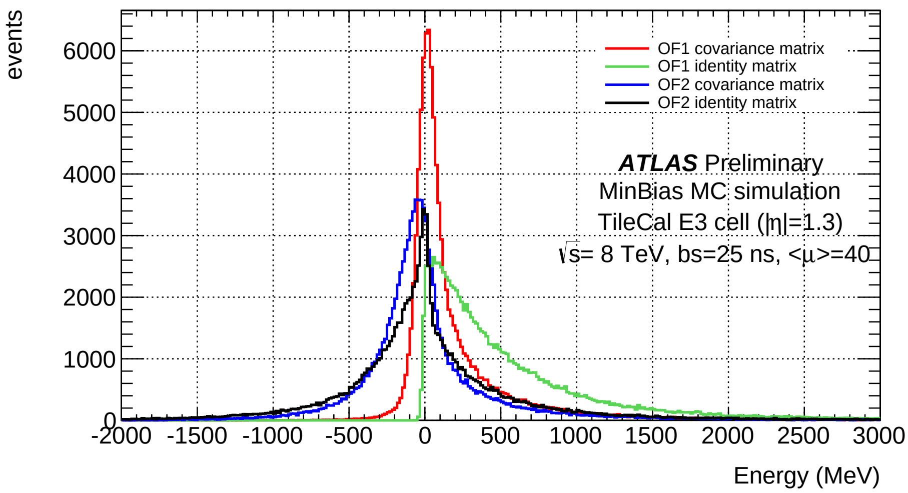

Figure 3. Energy distribution for cell E3 using both OF1 and OF2.

It can be noticed that the use of the covariance matrix in the computation of the OF1 weights shows the best performance in terms of estimation error (smaller dispersion). Additionally, the positive tail in the distribution is due to the high order statistics from the in-time pile-up signals that can not be accessed through the covariance matrix alone.

Figure 4 shows the energy distribution for the E4 cell which is the highest occupancy channel in TileCal. In order to summarize the parameters of each distribution, Table 1 shows the mean and RMS values for both E3 and E4 cells.

Table 1. Summary of the mean and RMS values of the energy distributions from E3 and E4 cells (in MeV).

	OF1 covariance		**OF1 identity**		**OF2 covariance**		**OF2 identity**	
	mean	RMS	mean	RMS	mean	RMS	mean	RMS
E3	18.28	113.84	74.04	282.69	-55.23	201.06	-43.97	299.55
E4	145.79	311.26	971.25	687.82	-163.13	424.67	-98.75	674.40

As expected, the OF method becomes biased and it also increases its variance under pile-up, and the best performance is again achieved by OF1 using the covariance matrix (smaller

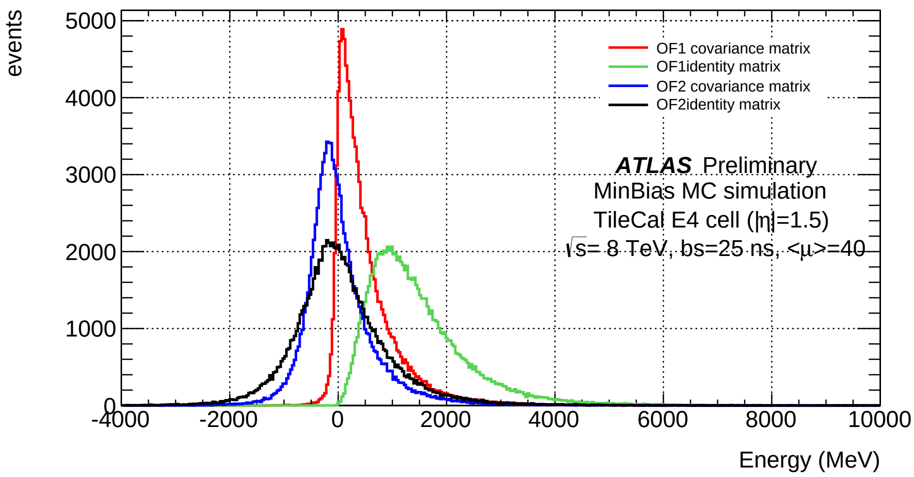

Figure 4. Energy distribution for cell E4 using both OF1 and OF2.

dispersion). In the case of these two highest occupancy cells, the bias (mean value) could be stored in a data base and used to be subtracted from each estimate for correction. This bias is not present for the other TileCal cells. It is worth mentioning that the mean and RMS were computed similarly as the covariance matrix, using a robust approach [8].

Figure 5 summarizes the improvement in the energy resolution (RMS of the energy distribution) using the covariance matrix with respect to the identity matrix for all cells in the central and extended barrels. The improvement is computed according to the following equation:

$$RI(\%) = 100 - \frac{RMS_{covariance}}{RMS_{identity}} \times 100 \tag{8}$$

For cells located in the layers BC and D, which are not highly affected by the pile-up, the improvement is minimal for TileCal central barrel ($|\eta| < 1$). In those cells, the identity approximation remains a reasonable description of the background for the majority of the events considered in this pile-up condition. While for the cells located in layers A and E in the extended barrel regions, the improvement becomes substantial, especially for the E cells which are closer to the beam and more exposed to the pile-up.

4. Conclusions

We present a study on the use of the background covariance matrix in the algorithm that will be used for TileCal online energy estimation during the next LHC run. The background covariance matrix can be used to reduce the uncertanties and bias due to the pile-up under high luminosity conditions. The results show that the OF using a fixed baseline value and the correct background covariance matrix presented the best performance in terms of estimation error. Most of the cells located in the central barrel did not show improvements due to their relatively low occupancy in the considered scenario. However, in some E cells in the extended barrel, the improvement is large as 60% with respect to the identity matrix.

Both the baseline value and the background covariance matrix are computed offline and stored in database. The OF weights are also computed offline and loaded in the DSPs for the online energy estimation.

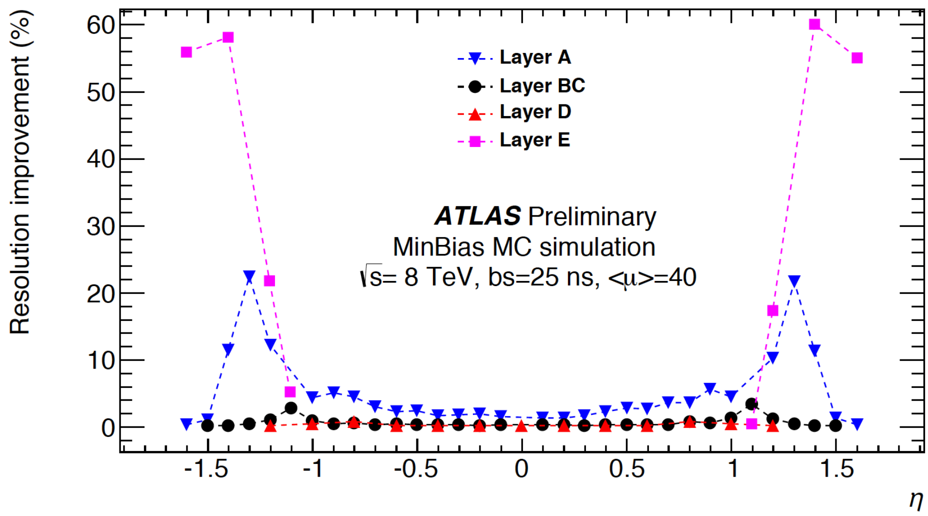

Figure 5. Improvement in function of η by using OF1 designed with the covariance matrix with respect to OF1 designed with an identity matrix.

Acknowledgments

We would like to thank the brazilian agencies CAPES, CNPq, FAPEMIG, FAPERJ and RENAVAE for the financial support. Also, we are very thankful to our colleagues from the ATLAS collaboration for the fruitful discussions concerning this work.

References

[1] ATLAS Collaboration 2008 *The ATLAS Experiment at the CERN Large Hadron Collider* JINST **3** S08003
[2] ATLAS Collaboration 2010 *Readiness of the ATLAS Tile Calorimeter for LHC collisions* EPJC **70** pp 1193-1236
[3] Evans L and Bryant P (editors) 2008 *LHC Machine* JINST **3** S08001
[4] Anderson K *et al* 1998 *Front-end Electronics for the ATLAS Tile Calorimeter* Proceedings of Fourth Workshop on Electronics for LHC Experiments
[5] Fullana E *et al* 2006 *Digital Signal Reconstruction in the ATLAS Hadronic Tile Calorimeter* IEEE Transaction On Nuclear Science **53** 4
[6] Peralva B *et al* 2013 *The TileCal Energy Estimation for Collision Data Using the Matched Filter* IEEE Nuclear Science Symposium and Medical Imaging Conference pp 1-6
[7] Peebles P 2000 *Probability, Random Variables, and Random Signal Principles* McGraw-Hill
[8] Rousseeuw P and Dreissen K 1999 *A Fast Algorithm for the Minimum Covariance Determinant Estimator* Technometrics **41** 3
[9] Chapman J 2011 *ATLAS Simulation Computing Performance and Pile-Up Simulation in ATLAS* LPCC Detector Simulation Workshop