

Optimal Filtering applied to 1998 Test Beam of Module 0

F. Camarena, J. Castelo, E. Fullana

*Instituto de Física Corpuscular
IFIC*

*Centro Mixto Univ. de Valencia - CSIC
Edificio Institutos de Investigación- Polígono la Coma S/N
VALENCIA, Spain*

ATL-TILECAL-2002-015
28/11/2002



Abstract

Optimal filtering is an algorithm that allows the reconstruction of energy and time for a photomultiplier multiple sampled signal, minimizing the noise coming from electronics and Minimum Bias events. This is anticipated to be the method used in ATLAS.

This note treats upon the application of the optimal filtering technic to real data from test beam and the comparison with the method used until now.

Introduction

The readout of the Tile Calorimeter uses a multiple sampling method to measure the signals coming from the PMTs [1]. The PMT signal is shaped and then, if the first level trigger accepts the signal, several samples on the pulse are digitized. After this, the digital information is pipelined in order to derandomize the events frequency. From these samples the amplitude and time information of the signal are reconstructed and passed on to further trigger levels and to the data banks. In order not to limit the bandwidth of the readout, i.e. the LVL1 trigger rate (100 KHz), the reconstruction of the samples to E & T is done in parallel using the last generation of Digital Signal Processors (DSPs), with up to 1200 MIPS and 8 ALUS [2]. Right after the digitization, the value of the single samples is only passed on for those events or for those channels where a reconstruction error or uncertainty is detected (possible if it is interesting) [3].

Without any noise, the height of the peak at the output of the shaper is proportional to the deposited energy. Multiple samplings allow to take into account the noises and recover some of the information lost when the pulse shape is distorted by noise. This is anticipated to be the method used in ATLAS.

In this note we will make a comparison between Optimal and Flat Filtering algorithms applied to test beam data. First, the data that we use for the study, 1998 test beam of Tile Cal experiment, is presented. In section 2, a brief description of the optimal filtering algorithm is exposed. Section 3 shows the results that the application of the algorithm to these real data produce. Finally, in section 4, the application of the technic to the ROD prototype present now in Valencia Laboratory and the future plans in this line of investigation are shown.

1 Test Beam considerations

In July 1998 test beam (figure 1), Module 0 was tested with 45 PMTs connected to 3-in-1 cards. 27 PMTs used 10 bit FERMI ADCs and 18 PMTs used 12 bit commercial ADCs (cells A7-A10, BC6-BC9 and D3) called DIRAC. In total there were 90 ADC channels. At the end of the test beam a new onboard 10 bit digitizers were installed inside the beam-right drawer.

Both, FERMI and DIRAC, systems had two independent gains per photomultiplier, low and high gain ($=64 \cdot \text{low gain}$). The FERMI(DIRAC) ADCs sample the signal every 25 ns. 40(15) samples per event per channel are recorded in the raw data. The first 10(5) samples contain no signal and they are used for the calculation of pedestal. The last 30(10) samples are considered as the signal window and they are used for the

evaluation of the signal [4]. The way to calculate the signal amplitude in FERMI(DIRAC) has been:

$$Q_{rec} = Filter(ch) * \sum_{n=1}^5 (Sample(ch)[N_{start} + n] - Ped(ch)) \quad (1)$$

The $Filter(ch)$ calibration coefficient is measured in the Charge Injection Calibration run (CIS) ($Filter(ch)=0.48$) [4]. This method is called *Flat Filtering*, and it has been the reconstruction method used in TileCal till the moment.

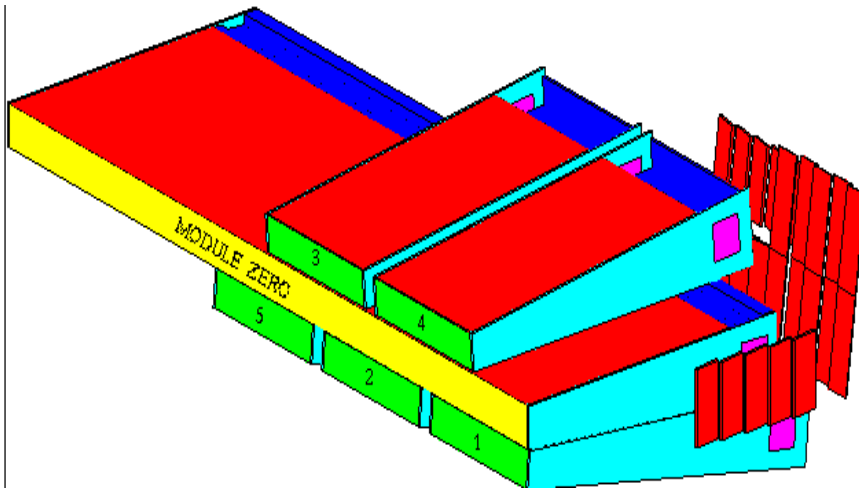


Figure 1: *Layout of module 0, old modules and scintillators in 1998 Test Beam.*

In this study, we have performed the optimal filtering offline reconstruction of a 300 GeV electron run knocking on the Module 0 with theta equal to -90° .

2 Optimal Filtering Algorithm

Optimal filtering is an algorithm that allows the reconstruction of energy and time for a photomultiplier multiple sampled signal, minimizing the noise contribution to these energy and time calculations. The advantages of OF versus FF method in sampling the peak of a shaped signal are: firstly, reduction of sensitivity to channel-to-channel variations in the pre-filter shaping parameters and secondly, the good performance over a wide range of operating conditions [5].

Mathematical expressions are:

$$u = \sum_i a_i S_i \quad i=1, \dots, 5$$

$$v = \sum_i b_i S_i \quad (2)$$

u will be the amplitude A of the signal, and v will evaluate to $A\tau$. S_i are the values of the samples containing the signal.

The shape of the signal is known, so:

$$S_i = Ag(t_i - \tau) = Ag_i - A\tau g'_i + n_i \quad (3)$$

where n_i is the noise contribution and $g(t)$ is the shaper output waveform. From (2) and (3) equalities we obtain the next constraints:

$$\begin{aligned} \sum_i a_i g_i &= 1 & \sum_i a_i g'_i &= 0 \\ \sum_i b_i g_i &= 0 & \sum_i b_i g'_i &= -1 \end{aligned}$$

The variances of the parameters u and v are given by

$$Var(u) = \sum_{ij} a_i a_j \langle n_i n_j \rangle = \sum_{ij} a_i a_j R_{ij}$$

$$Var(v) = \sum_{ij} b_i b_j \langle n_i n_j \rangle = \sum_{ij} b_i b_j R_{ij}$$

where R_{ij} is the noise autocorrelation function evaluated at $(t_i - t_j)$.

We minimize the variances of u and v while satisfying the constraints using Lagrange multipliers. The function to minimize is:

$$I_u = \sum_{ij} R_{ij} a_i a_j - \lambda (\sum_i a_i g_i - 1) - \epsilon \sum_i a_i g'_i \quad (4)$$

The system to solve in order to obtain the a_i weights is:

$$2 \sum_j R_{jk} a_j + \lambda g_k + \epsilon g'_k = 0 \quad k = 1, \dots, n$$

$$\sum_j a_j g_j = 1 \tag{5}$$

$$\sum_j a_j g'_j = 0$$

Equation 5 is a normalization equation that fix the same units for the energy reconstructed than for the samples, it is ADC counts in our case.

Another Lagrange condition, equation 6, can be used in order to make the result insensitive to shifts in the baseline, which is easy to see by noting that $E = \sum_i a_i (S_i + B) = \sum_i a_i S_i$, where B is any constant. If a fluctuating baseline is observed, then it might be worthwhile to try this, but it will make the noise a little worse, as will any additional constraint.

$$\sum_j a_j = 0 \tag{6}$$

3 Optimal Filtering applied to Test Beam data.

It is necessary to know the waveform of the shaper evaluated in t_i , that is $g(t_i)$ and $g'(t_i)$, and also the noise autocorrelation function R_{ij} to solve the system of equations 5. This section describes the procedure to obtain this information as well as the problems we can find in the test beam context.

3.1 The Shaper Output Waveform

We have considered as first approximation the same shape form for all the channels (figure 2 left). Although in principle, it is clear that each channel could be characterized by a different shaper output waveform due to different electronic components and photomultipliers, this approximation is not unrealistic and surely in ATLAS, since in present experiments, e.g. NA48, channels with quite resemblance will be grouped with the same weights [6]. This assumption includes, also, the equivalence between the output waveforms when high or low gain is used. The effects of these considerations will be analyzed later.

The shaper output waveform function ($g(t)$), sampled each 2 ns, has been obtained from one channel using a digital oscilloscope. To perform the calculations in 5 as well as to obtain $g'(t)$ these numerical values (figure 2 left), and the lineal interpolation between them, have been used in order to perform the calculation in equations 5 as well as to obtain $g'(t)$. $g'(t)$ is shown in figure 2 right. The fluctuations observed are due to fluctuations in the baseline of the $g(t)$ shape. Such a high oscillations has not been observed in recent beam tests. It could be caused by the electronics of the particular channel analyzed with the scope. The units used to calculate the derivative of the waveform going into the calculation of the b_i are fixing the units of the time reconstruction, that are seconds in our case.

Figure 3 shows the coupling between data obtained from the scope and the corresponding to a test beam event.

Although the method of the digital oscilloscope is perfectly valid in order to obtain $g(t)$ in the context of the test beam, it will not be useful in ATLAS as there will be 10010 channels to characterize. The best alternative is to use the CIS system that allows the reconstruction of $g(t)$ function simultaneously for each channel through a fixed injected charge and incremental phase of 0.7 ns spanning 25 ns [7].

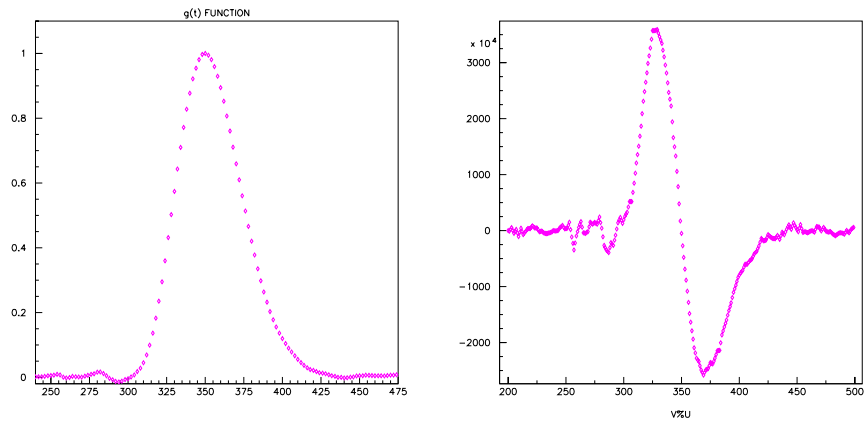


Figure 2: $g(t)$ normalized function and $g'(t)$.

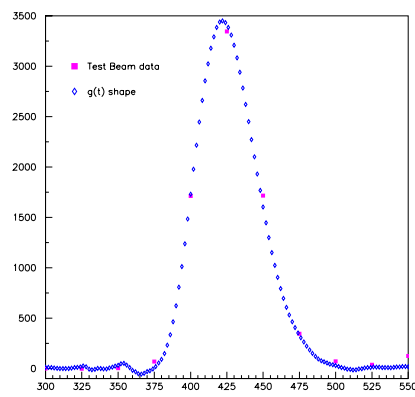


Figure 3: Adjustment of $g(t)$ with a test beam sampled event.

3.2 The Noise Autocorrelation Function

Since we are dealing with a system which is bandwidth limited by the shaper, data samples taken in a time which is short compared to the dwell time of the shaped signal are correlated. It is necessary, when considering the treatment of highly correlated data, to understand the autocorrelation function of the system, from which we can obtain the covariance matrix for the data samples [8]. The R_{ij} matrix elements can be understood in terms of the autocorrelation functions of the electronic noise, as pile-up is not present in the 1998 test beam. In ATLAS at low luminosity the electronic noise will dominate whereas at high luminosity the pile-up noise will do.

Three ways have been envisaged to evaluate the noise autocorrelation function and the relevance it has in the weight calculation:

1. Case A.- R_{ij} is computed considering our source of noise as thermal noise [5].
2. Case B.- R_{ij} is computed from test beam noise samples.
3. Case C.- R_{ij} is assimilated to *Dirac Delta*, meaning no correlation between samples.

Case A

R_{ij} can be evaluated from the next expression [5]:

$$R(t) = \sigma_T^2 \frac{\int h'(t+u)h'(u)du}{\int h'(u)h'(u)du} \quad (7)$$

where σ_T is the thermal noise, the variance of an individual sample S_i , and $h'(t) = \frac{dh(t)}{dt}$, where $h(t)$ (figure 4 left) is the δ -response of the shaper obtained from the digital scope.

In figure 4 right it can be seen the result of the thermal autocorrelation function versus time obtained from the upper expression.

The R_{ij} obtained is extracted from the values of figure 4 right, taking intervals of 25 ns.

1.	0.035	-0.441	-0.139	-0.008
0.035	1.	0.035	-0.441	-0.139
-0.441	0.035	1.	0.035	-0.441
-0.139	-0.441	0.035	1.	0.035
-0.008	-0.139	-0.441	0.035	1.

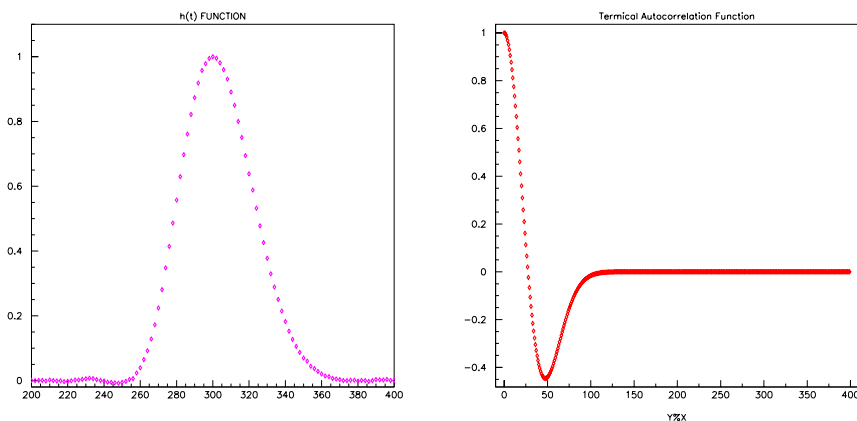


Figure 4: $h(t)$ and $R(t)$ functions

Case B

During a real experiment, the procedure to extract R_{ij} will be to evaluate the promedium $\langle n_i n_j \rangle$ directly from data without signal. In the case of our test beam, the promedium have been evaluated for each channel, 45 in total, and the averages are shown in the next matrix.

1.	-0.075	0.108	-0.101	0.092
-0.075	1.	-0.073	0.120	-0.087
0.108	-0.073	1.	-0.072	0.116
-0.101	0.120	-0.072	1.	-0.076
0.092	-0.087	0.116	-0.076	1.

The first thing we observe is that the values R_{11} to R_{15} do not correspond with the one we would expect for thermal noise. The reason could be a non-thermal source of noise. If we plot the covarianze of the 40 samples (figure 5) we see two peaks in the samples 15 and 30 approximately. This fact makes us think in the presence of an external clock as the source noise. As there are 25 ns between samples, the period of the clock would be $(30-15)*25$ ns that makes ~ 375 ns period of the clock. Of course with this hypothesis there would be peaks in the samples 0-15-30-45, probably there is one in first sample but it is hidden by the normalization factor and we can not check in the 45 because we just have 40 samples. We are suspicious that this noise is introduced by the Fermi adquisition system. This effect was already observed in the correlations of pedestals in the 1996 test beam of module 0, where also the Fermi system was used [9].

The source of noise exposed and other ones as contamination from real signal or skyshine could cause big errors in the calculation of R_{ij} with this procedure. As a consequence, in figure 6 is shown the big dispersion we obtain between the different channels. The effect of these anomalies will be analyzed in section 3.5.

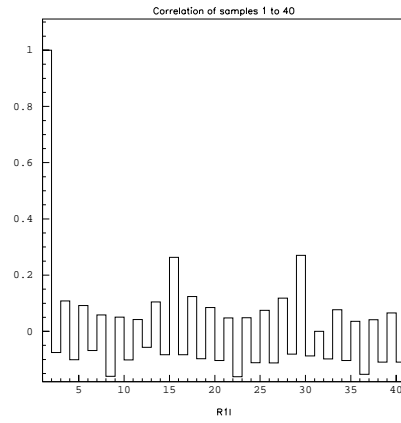


Figure 5: *Correlation between pedestal samples for one channel.*

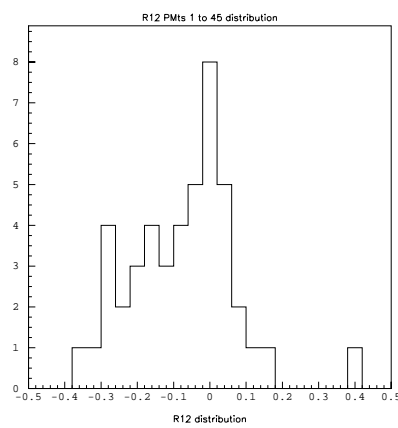


Figure 6: *R_{12} distribution for all the channels.*

Case C

In this section we assimilate R_{ij} to δ_{ij} . That supposes no correlation between samples, what is in fact no real in the experiment. Anyway, it is useful to prove that R_{ij} is not affecting drastically the energy and time reconstruction. This option is used in experiments like NA48.

3.3 The Optimal Weights

One of the problems we had during the analysis was that data from physics runs in beam tests are asynchronous. It means that the signals are sampled at start times generated through a 40 MHz clock, while the events arrive at random times. In the present system, the phase of the trigger is not synchronized with the ADC clock, so that the time jitter is uniformly distributed in a clock period and the samples will fall each time in different positions of the output waveform function [11]. It means that in system 5 we should evaluate $g(t)$ and $g'(t)$ in different positions in function of the phase, and consequently we would obtain different weights as phase function.

Our method takes into account this randomness by having 50 sets of 5 coefficients corresponding to 50 time bins of 2 ns each. However it remains a ± 1 ns uncertainty that we correct interpolating between the two closer values. In future applications at collider experiments, the clock will be phase locked with the beam crossing time so that the jitter will be much reduced and just a set of weights will be needed per channel.

Figure 7 shows the position of the samples in $g(t)$ for two different jitter events.

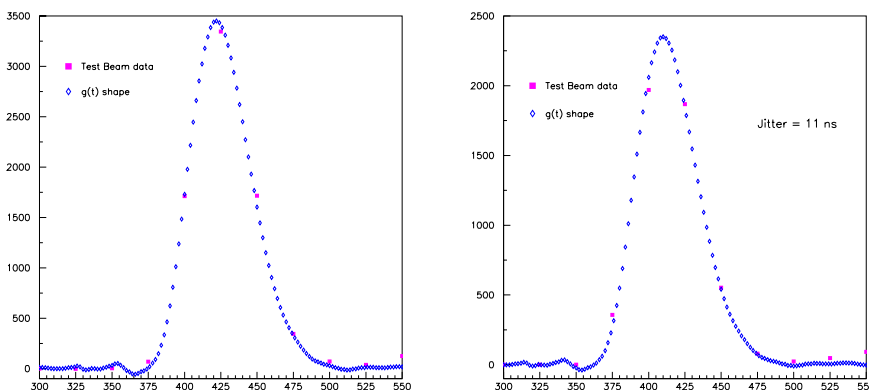


Figure 7: *Left: adjustment of $g(t)$ with a test beam sampled event with jitter=0. Right: effect of the jitter time*

Figures 8 to 10 show the look-up-table obtained for the optimal weights used to estimate the energy and $E\tau$ in the respective cases proposed. As expected, the third coefficient has the highest value when phase is close to zero, because in this case that the ratio signal/noise is the highest. The higher the luminosity, the higher is this third coefficient. As we would expect, when phase increases also coefficient 2 increases, because it is falling closer to the peak region, and analogously coefficient 3 decreases because it goes to the baseline region. Instead, the timing information does not lay in the third point (the peak) because at this point the shaping function is flat over many nanoseconds.

Next tables show numerical values of the weights when phase is equal to -3 ns, it is where the third sample falls closer to the peak.

	a_1	a_2	a_3	a_4	a_5
Case A	0.327	0.287	0.654	0.359	0.370
Case B	-0.028	0.243	0.688	0.389	0.063
Case C	-0.006	0.246	0.687	0.385	0.086

Table 1: a_i optimal weights associated to phase=-3 ns.

	b_1	b_2	b_3	b_4	b_5
Case A	1.63E-9	-2.16E-8	1.36E-9	1E-8	4.1E-9
Case B	-8.05E-10	-1.83E-8	-2.12E-9	1.41E-8	3.80E-9
Case C	-8.04E-10	-1.87E-8	-1.64E-9	1.33E-8	4.72E-9

Table 2: b_i optimal weights associated to phase=-3 ns.

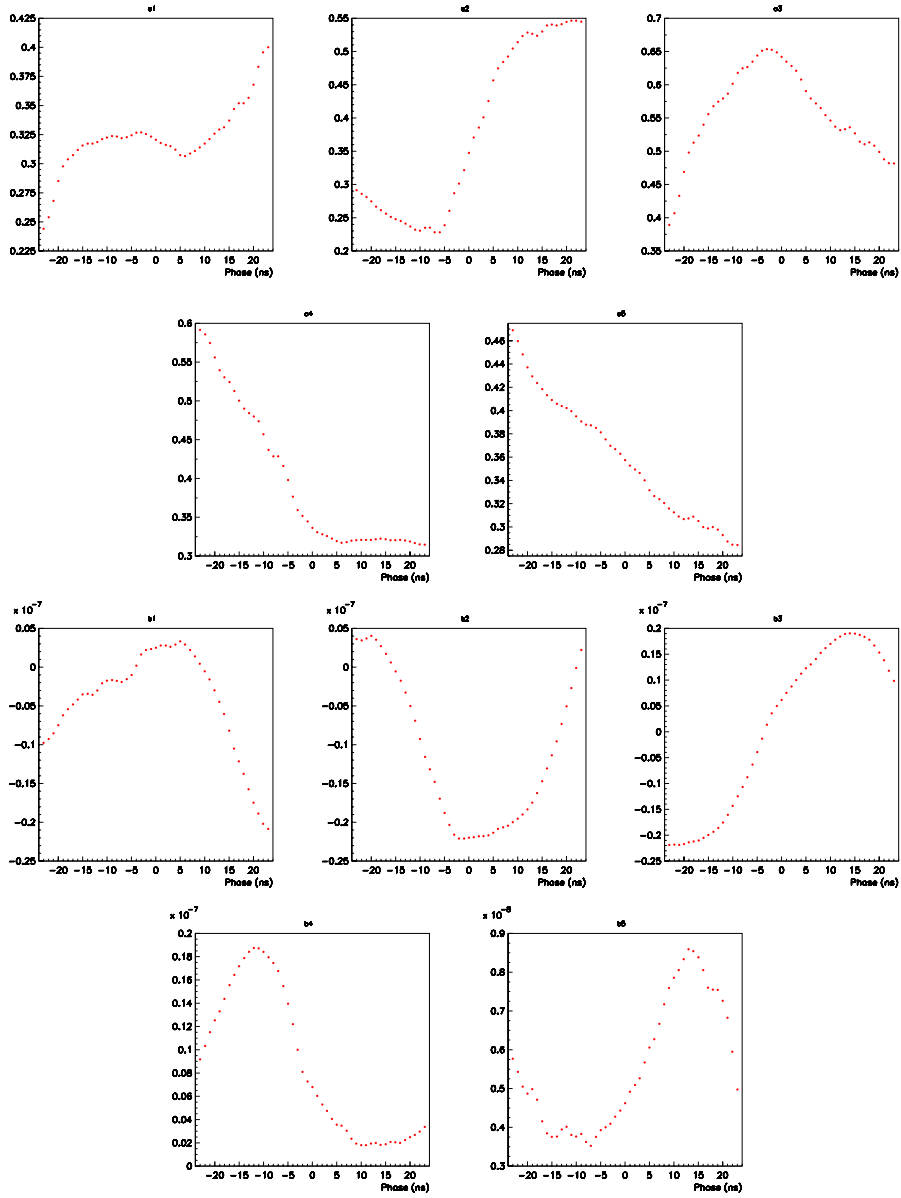


Figure 8: Values of the coefficients a_i and b_i as phase function. Case A.

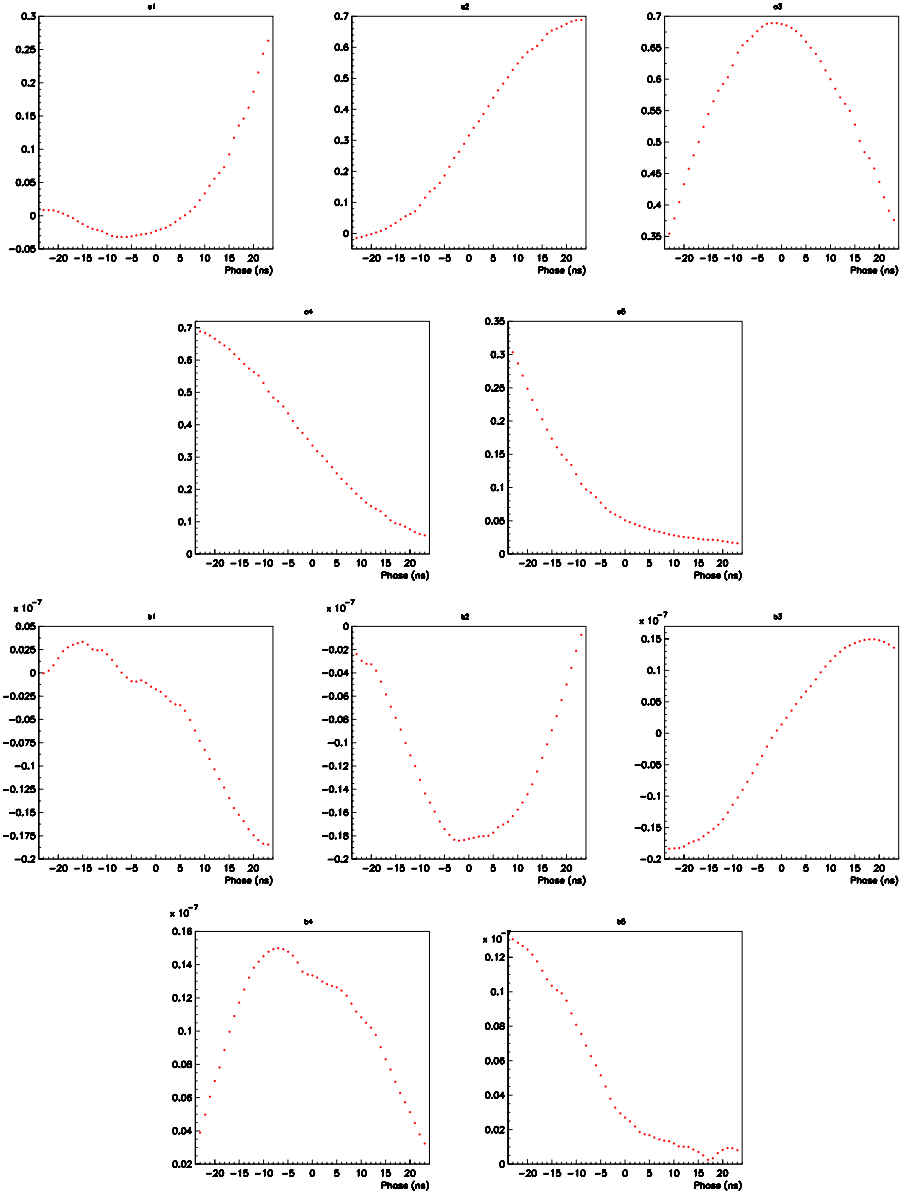


Figure 9: Values of the coefficients a_i and b_i as phase function. Case B.

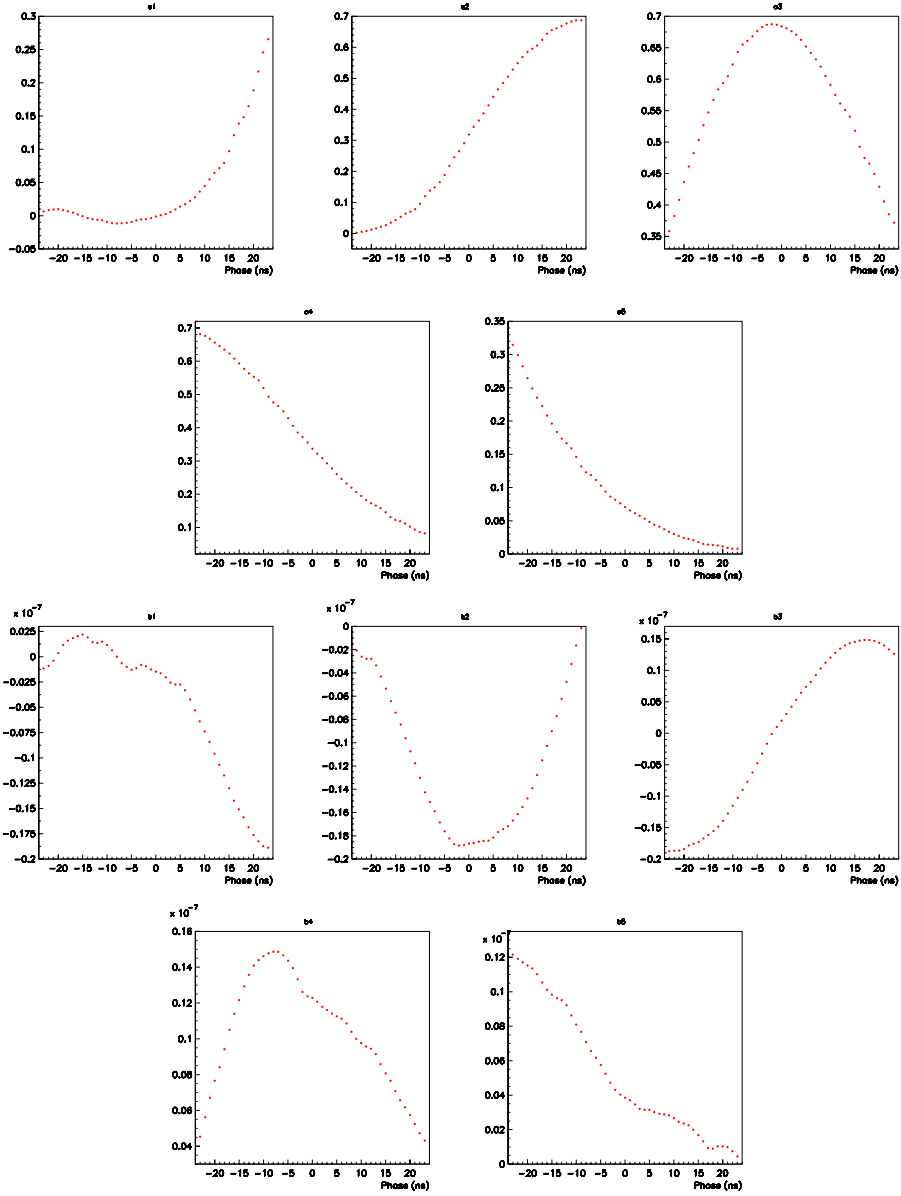


Figure 10: Values of the coefficients a_i and b_i as phase function. Case C.

3.4 Energy and Noise Reconstruction

In this section the results for the energy reconstruction are presented in comparison with the reconstruction obtained with Flat Filtering method. As it can be observed in figures 11 and 12 the three possible R_{ij} studied provide good energy reconstruction, pointing out that, at first order, its effect is not drastic, as can be the ignorance of the output waveform. Results are presented in GeV using the absolute electromagnetic scale (0.87 pC/GeV) [10].

Study of electrons in test beam is important to find absolute energy scale, the pC/GeV conversion factor. Resolution of electrons is interesting in order to compare with GEANT3 and GEANT4 simulations.

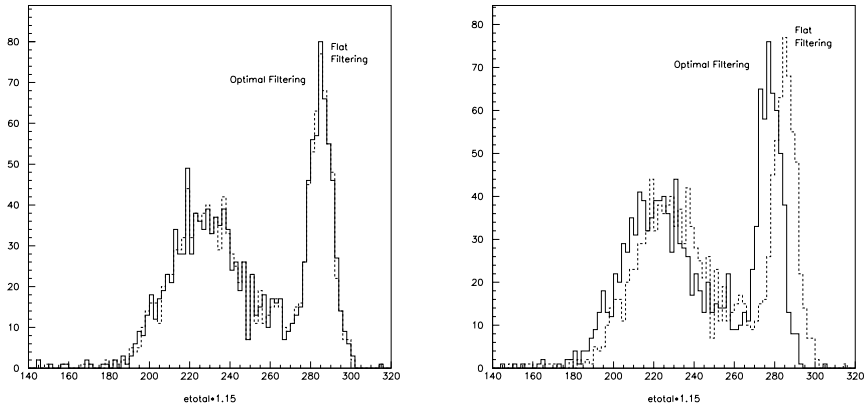


Figure 11: *Energy reconstruction (in GeV) obtained for cases A and B.*

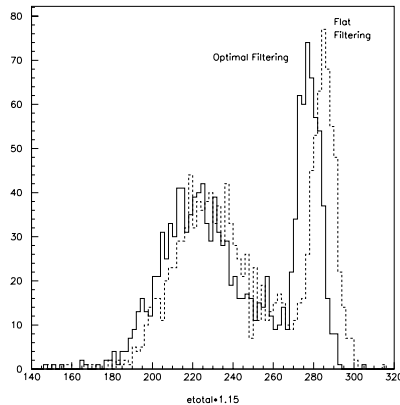


Figure 12: *Energy reconstruction (in GeV) obtained for case C.*

In table 3 are shown the resolution obtained with the different methods for 300 GeV electrons. Resolution remains constant and it is coherent

with results obtained in the last beam tests [12], where the values obtained could be parametrized between:

$$\frac{\sigma}{E} = \frac{39\%}{\sqrt{E}} + 0.6\%$$

and

$$\frac{\sigma}{E} = \frac{30\%}{\sqrt{E}} + 0.2\%$$

in function of the pseudorapidity.

Method	μ (GeV)	σ (GeV)	$\frac{\sigma}{E}$ (%)
Flat Filtering	285.1	5.88	2.1
Case A	285.0	5.73	2.0
Case B	277.6	5.58	2.0
Case C	277.6	5.53	2.0

Table 3: *Resolution obtained with the four methods.*

Any of the three possibilities applied in the study improve the RMS noise respect to the value obtained with *Flat Filtering*, which is clearly expected as it is one of the requirements we impose to the equations to calculate the weights. The minimization of electronic RMS noise is very important in order to establish a cut for the energy deposited in a cell. The thinner the gaussian noise, the easier distinguish between noise or energy from a low pt jet deposited in a cell.

Method	RMS (MeV)
Flat Filtering	152
Case A	126
Case B	106
Case C	106

Table 4: *Electronic noise obtained with the four methods.*

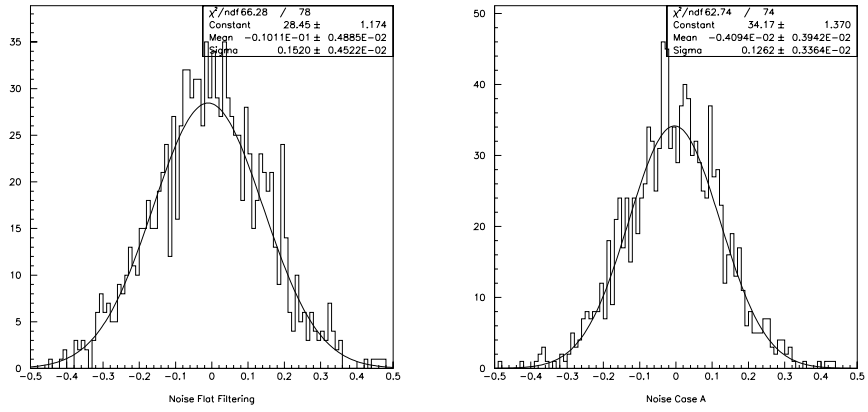


Figure 13: *Noise reconstruction with Flat Filtering and Optimal Filtering in case A.*

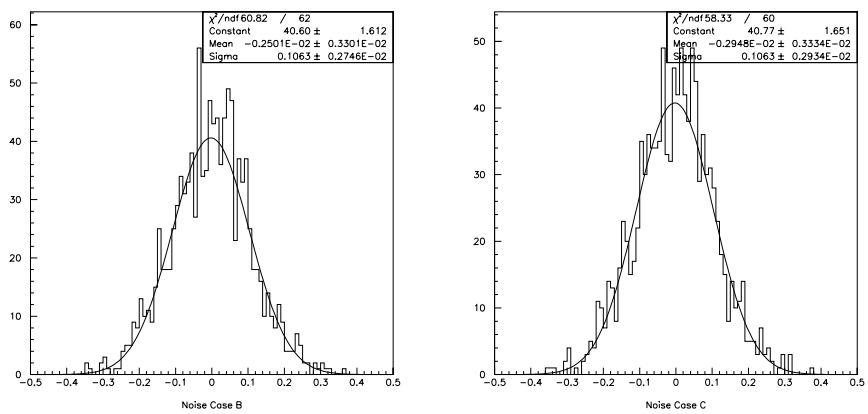


Figure 14: *Noise reconstruction with Optimal Filtering in cases B and C.*

3.5 Time Reconstruction

Optimal Filtering allows the time reconstruction of the event. Figure 15 left shows this time reconstruction applied to the testbeam. The meaning of the time reconstruction is different in the context of the test beam than in the real experiment. In a real experiment the trigger of the digitizers will be synchronized with the bunch crossing one. Therefore the time variable, reconstructed for each channel, means the possible delays introduced by a few factors as: development of the shower in the calorimeter, length of the fibers of each channel or delays coming from the electronic components. In our actual prototype in the test beam this synchronization was not possible, this is the reason of the look-up-tables used for the energy reconstruction. A consequence of using the look-up-table is that the meaning of the time variable change. Now we are using b_i weights to recover this time, but these weights have been evaluated in a phase time that incorporates all the delay causes: unsynchronization, shower development, electronics, fibers etc... so that the result we would expect for τ is a distribution centered in zero. Obviously this is not the case observed in figure 15 left, which shape makes us suspect about some systematics.

Figure 15 right shows a clear dependence of the time reconstruction with the phase, where we would expect a narrow distribution centered at zero. That could indicate a change in the quality of our fit as a function of samples phase, as is proved by figure 16 left, where χ^2 represents the difference between the assumed and measured waveform (the residuals, $\sum_1^5 \frac{(Sample_i - g(t_i))^2}{5}$) as a phase function. We can find that when the samples fall on the regions where the residuals are largest, we get the biggest spread in the values of τ , indicating that certain portions of the waveform are not well described by the model. This systematic is consequence of two factors: regarding the same waveform for all channels and the fluctuations in the baseline of our $g(t)$ estimation. Whether it helps to introduce a waveform for each channel can only be answered experimentally.

In order to have an estimation of the time resolution that we can get in the test beam context using optimal filtering and the method proposed to correct the effect of the unsynchronization, in figure 16 right is presented the distribution of τ obtained straightening the plot shown in figure 15 right. The non-gaussian tails that we can see in the time distribution could arise mostly from imperfections in the correction function used to straighten out the curve, but there could be many other causes at the level of a few nanoseconds. It would be required a counter in the beam with good timing properties to give an independent measure of the event timing in order to do a good study of this problem.

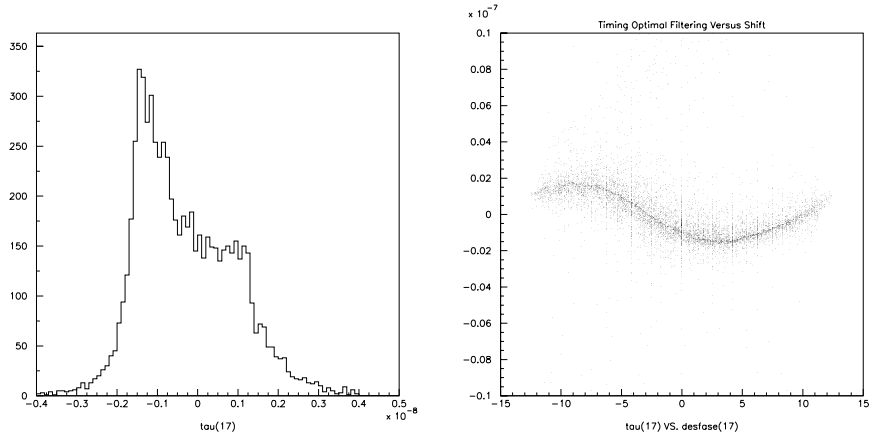


Figure 15: *Time reconstruction and dependency of it in function of the phase*

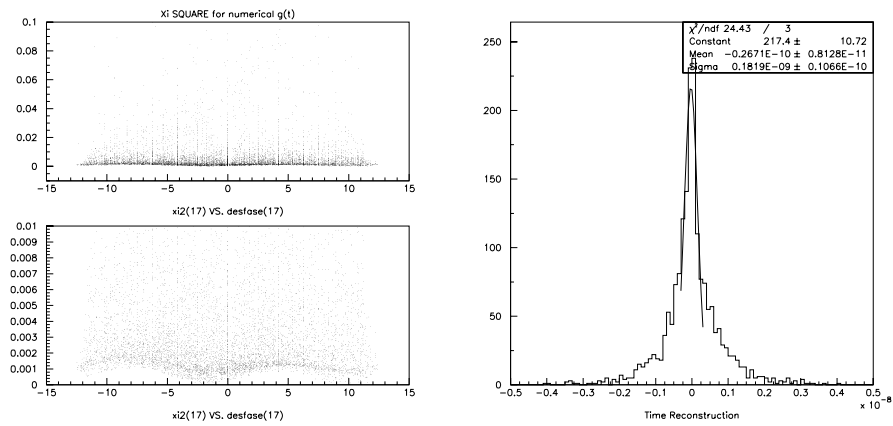


Figure 16: χ^2 and time resolution once subtracted the systematic behaviour

4. Hardware implementation of optimal filtering algorithms

At Valencia laboratory there are several facilities for implementing the optimal filtering algorithm over the ROD prototype hardware. The basic set-up we have is based in VME Crates with two ROD prototypes modules and single board computers running data acquisition software for acquire the data processed in each module. The ROD modules are motherboard 9U VME format with several programmable devices (FPGAs) for data management, VME communication and trigger information reception. These motherboards can hold up to four processing units which are responsible for raw data processing coming from the front end boards electronics of the detector and output the processed data (with optimal filtering algorithm) to the next level trigger stage of the read out system of the global ATLAS detector.

Each processing unit is based in Texas Instruments DSP TMS320C6202 running at 250MHz with an instruction cycle of 4ns and 8 general purpose and specific arithmetic and logic units running 32 bits instructions in parallel with the aid of a previous pipeline. Around this core there are two FPGAs (for input and communications management), one dual port memory for buffering the input events, and one FIFO memory for store the output processed and formatted events.

Actually, several studies have been done for processing with this processing unit (in conjunction with the other hardware) up to 45 channels (one barrel/drawer of the calorimeter, 256 in total for complete readout) in less than 10 μ s (final ATLAS level 1 trigger rate 100kHz). The processing is done based on the simulation studies described in this note, but over the hardware, testing it at full speed. There is software developed for the data acquisition system running in linux, and the DSP with commercial compilers. The optimal filtering implementation in the DSP has been based in two input programming languages: C and assembler. Of course a low level programming language (the assembler) reports more speed than the high level (C), but paying the price of a difficult input and maintainability. In principle the assembler based compilation promise a optimal filtering processing over 45 channels in less than the 10 μ s limit.

The aim of the present study is the evaluation of the time consumption by this hardware setup in order to obtain the magnitudes associated with optimal filtering technic (E, t, and χ^2). The algorithm runs with arranged event files taken from testbeam data. The data acquisition software takes this file and stores it in the DSP memory. The DSP runs the optimal filtering algorithm calculating the energy, time, and χ^2 for 45 channels using the weights calculated with method described in this note. Finally several histograms are filled for each channel and the processed data is acquired by the data acquisition software and stored for a final offline analysis of processed data.

Conclusions

A method to correct data asynchronization in test beam has been proposed in order to be able to do the energy and time reconstruction with the optimal filtering technic. Results of energy, noise, and time has been provided for the first time using real data from a TileCal test beam, showing three possible alternatives to evaluate the optimal weights in function of the selected autocorrelation matrix. The existence of some extra correlation in noise events that affects the result of R_{ij} has also been detected. The importance of a good knowledge of the shaper waveform has been figured out in the existence of systematic errors in the time reconstruction. Optimal filtering algorithm improves the energy reconstruction respect to the flat filtering method even with the worse-case hypothesis of one shape form for all the channels. Moreover with OF it is also possible to make time reconstruction if set up conditions are suitable. All these advantages just for a negligible increase of calculation time allow us to foresee a good performance in the future experiment.

Acknowledgements

We would like to express our gratitude to W.E. Cleland for sharing with us his wide knowledge and experience; with kindness and friendship.

Also we would like to thank G. Unal for his good suggestions in base of his experience in NA48 experiment and, A. Solodkov and B. Stanek for their help in questions related with the test beam, and R. Teuscher for the information about the pulse shapes.

Bibliography

- [1] The Tilecal Collaboration, ATLAS Tile Calorimeter Technical Design Report, CERN/LHCC96-42(1996).
- [2] J. Castelo, ROD Processing Unit Performance. Tilecal General Meeting Presentation. CERN, Feb. 2002.
- [3] L. Hervás, A New Reconstruction Algorithm for Calorimeter Signals. ZEUS-Note 96-003. February, 1996.
- [4] A. Solodkov, I. Efthymiopoulos, The TILECAL Program for Test Beam Data Analysis. User Manual (1998).
- [5] W.E. Cleland and E.G. Stern, Signal processing considerations for liquid ionization calorimeters in a high rate environment. Nucl. Instr. and Meth. in Phys. Res. A 338 (1994) 467-497.
- [6] Internal communication with G. Unal.
- [7] A. Rios, Estudio del Sistema de Adquisición de Datos del Experimento ATLAS/LHC del CERN. Aplicación al detector Tilecal. Tesis Doctoral (1999).
- [8] E. Gatti et al., Nucl. Instr. and Meth. A 274 (1989) 469.
- [9] J. Sjolín, Electronics Evaluation, Jet Reconstruction and a Study of GMSB in ATLAS. ATL-TILECAL-2000-015. June, 2000.
- [10] Taskforce discussion in TileCal Week, June 2001.
- [11] S. Agnvall et al., Evaluation of FERMI Read-out of the ATLAS Tilecal Prototype. CERN-PPE/97-144 (1997).
- [12] Y. Kulchitsky, Electron Energy Linearity and Resolution for TileCal Extended Barrel Modules: eta-scan and tile-scan. Talk in the TileCal week (February, 2002)