September 10, 2003

Distributions of Minimum Bias Current measurements in TileCal

X.Portell, I.Korolkov and M.Cavalli-Sforza

Abstract

The purpose of this note is to evaluate the spread of Minimum Bias (MB) currents in TileCal, when measured using the existing electronics. This information allows to estimate how many measurements in each cell are needed to obtain a statistical error of 1% of the mean, which is the goal of the MB monitoring system; in turn, this information will be used to design the MB data acquisition system.

A reliable estimate of the spread requires the use of signals from several billion bunch crossings, which is way in excess of the available sample of simulated MB collisions. Therefore the simulated MB signals in the TileCal readout cells are parametrized and the analytical forms are used to generate samples of MB signals as large as needed.

It is shown that at design luminosity even in the cells with the smallest currents a few tens of measurements should achieve the goal of a 1% error in monitoring the MB current.

ATL-TILECAL-2003-010
31 October 2003

1 Introduction: Monitoring Minumum Bias currents in TileCal

Inelastic proton-proton collisions at low momentum transfer, the Minimum Bias (MB) signal, will produce non-negligible backgrounds in TileCal, despite the shielding provided by Liquid Argon calorimeters. However, these processes can be used to continuosly monitor the response of all calorimeter cells during data-taking. Preliminary estimates of average MB energy depositions, energy spectra and photomultiplier (PMT) currents were given in the Tile calorimeter TDR [1] and led to the design of the TileCal calibration and monitoring electronics, which is described in the same TDR.

Let us recall the main features of MB signals given in the TileCal TDR. These results were obtained then (1995–1996) using pp collision simulation codes such as PYTHIA and full simulations of the ATLAS detector. Of course mean MB energy depositions per crossing are proportional to the LHC luminosity and are azimuthally symmetric (in the absence of polarization of the beams). In addition, the simulations show that they vary at most by a factor of 10 over the η range spanned by the Tile calorimeter, and that at the LHC design luminosity of $10^{34} cm^{-2} s^{-1}$

they amount to roughly 20 MeV per bunch crossing (BX) in the first depth sampling (A-cells) of the Barrel and Extended Barrel, dropping to a few tenths of an MeV in the last sampling (D-cells). At the nominal PMT gain and with photoelectron yields of $70/GeV/cell$ these energy deposition rates will give average PMT currents ranging from several hundred nA in the A-cells to only a few nA in some of the D-cells.

Minimum Bias current readout relies on the same electronics used to measure the radioactive source signals [2][3], but operates over a broader dynamic range in order to accomodate signals from all cells of the calorimeter. At design luminosity, the occupancies vary from about 1% to 40% and the MB signal energy spectrum has no sharp high-energy cutoff, because soft interactions smoothly merge into jet physics. Hence energy deposits may vary by orders of magnitude from bunch to bunch; however the only variable of interest for monitoring purposes is the time-average of the MB current, which is obtained by the MB monitoring electronics by "integrating" the current of each PMT over a time constant of up to 10 ms. More precisely, the MB signals generate a voltage across an operational amplifier equipped with a low-pass RC filter whose time constant can be programmed in the range between 0.28 and 10 ms.

The goal of the MB monitoring system is to detect in real time variations of more than 1% in the response of any channel of TileCal assuming stable luminosity conditions. The purpose of this study is to find how many MB current readouts are necessary to reach this level of repeatability considering fluctuations in the deposited energy. Hence, we will simulate MB energy deposits in the whole calorimeter and over very large numbers of bunch crossings, in order to obtain distributions of MB current measurements and evaluate their spread in all TileCal cells. These numbers must be known to specify the rate at which the whole calorimeter's MB signals (about 10000 PMTs) must be read out, to be able to see cell response anomalies in real time — which in our case we take to mean every few minutes.

Due to the (necessarily) very long time constants of the integrators, literally billions of MB interactions must be simulated in order to obtain distributions with reasonable statistics. Such numbers of MB events has not been generated to-date. In order to have enough statistics, we had to generate our own energy deposition spectra, cell by cell, from a parametrized distribution of the simulated MB energy depositions.

2 MB simulations and parametrization of energy spectra

We started from a data set of 5000 MB collisions, generated using PYTHIA5.7 with the Physics TDR layout. The full simulation code was GEANT3 with the GCALOR hadronic shower package. At the time when this study was begun, the much larger samples of MB events generated in 2002 in Mock Data Challenge 1 were not yet available. However even these much larger samples fall short of the needs of this simulation by orders of magnitude. Therefore the parametrization of energy spectra in order to obtain from them sufficient cell energy deposition statistics is still necessary.

For optimal use of the limited available MB event sample, cell energy deposition distributions were generated taking advantage of the ϕ symmetry of the process and of the symmetry around the $\eta = 0$ plane of the Tilecal cells. Specifically, energy depositions were folded into only one ϕ bin, and negative η signals were folded into the corresponding positive η plots. The layout of Barrel and Extended Barrel cells used in this MB data sample is shown in Fig. 1.

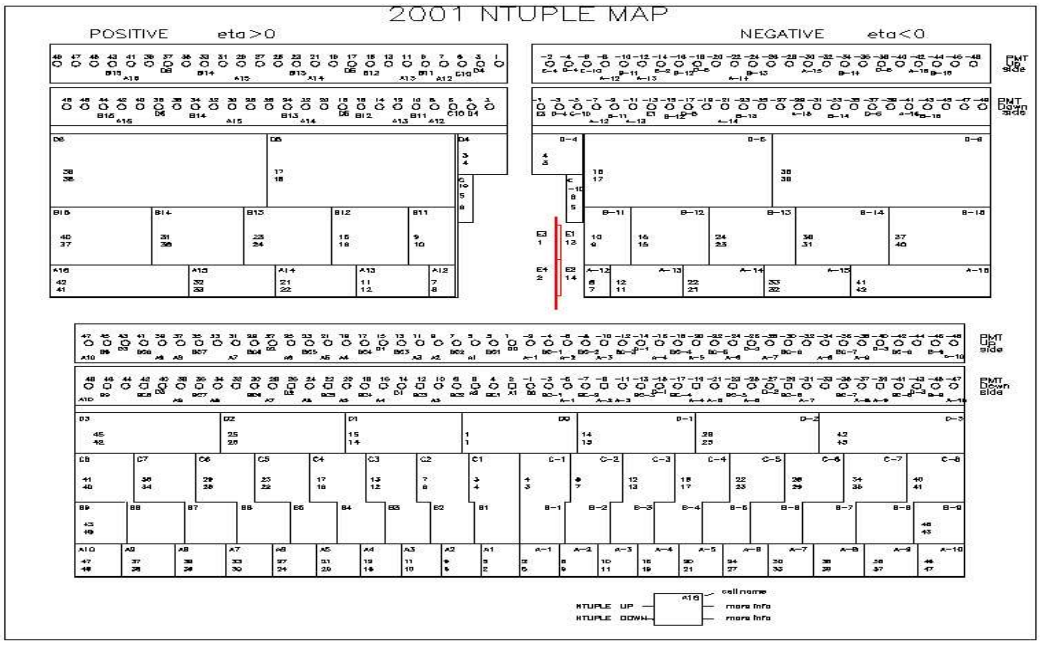

Figure 1: Cell locations and labels for the TileCal Barrel and Extended Barrels

The parametrization of the spectra required considerable care. First, for the results to be useful in predicting the spread of the MB measurements, the parametrizations must be reasonably accurate. We required mean values of the parametrizations to agree with the original distributions within 20%, and generally obtained better than 20% agreement. Second, to make the simulation task more manageable it is very desirable to to parametrize the energy deposit distribution in all cells with the same analytical form, while keeping the specified accuracy. The following parametrization meets both goals:

$$P_1 \cdot e^{-P_2 \cdot E} + P_3 \cdot e^{-P_4 \cdot E} + P_5 \cdot (E + 1\mathrm{GeV})^{-6} \tag{1}$$

This function was fit to the symmetrized MB energy distributions of each cell over the interval from 1 MeV to 4 GeV. The contribution to MB currents of energy deposits of less than 1 MeV is negligible; above 4 GeV, there is almost no statistics to guide the fit. The first term dominates in the 0.001–0.2 GeV region (see Fig. 2), the second term from 0.2 to 0.9 GeV, and the third above

0.9 GeV. The inverse-power-law form of the third term is motivated by the behavior of QCD-like distributions; the 1 GeV shift is to keep it from diverging at zero energy.

As shown in Fig. 2, above 1 GeV events are very few and most histogram bins have no contents. Therefore a binned maximum likelihood fit was made to the cell energy distributions, using bins of 5 MeV. We checked that the fitted distributions do not overrepresent any region of the energy spectrum by comparing the mean energy values of the simulated and fit MB energy deposition spectra in each of the three energy regions mentioned above (0.001–0.2 GeV, 0.2–0.9 GeV, and above) and found good agreement. The fit results are summarized in Fig. 3, where each point represents the total energy of the MB simulated events or of the fitted MB energy distribution for the sum of the all 64x2 cells at a given $|\eta|$ for each of the three depth samplings of TileCal. In most cases, the means of the MB simulated energy depositions agree with the parametrized energy depositions to better than 10%; the few cells in which a larger discrepancy is seen are those that detect very small energies because of their location in the calorimeter. Also note that the parametrized mean energy depositions tend to be a few percent higher than in the MC events, a slight overestimate that we consider acceptable.

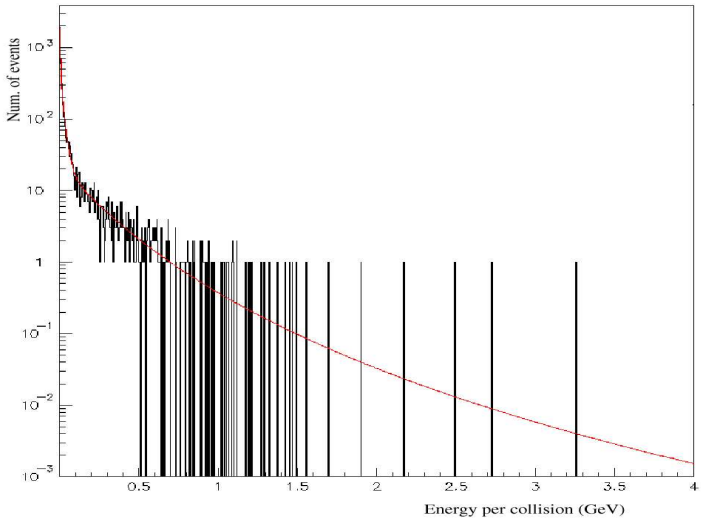

Figure 2: MB energy distribution for cells A1 and A-1 — spanning $|\eta| \leq 0.1$, shown in logarithmic scale in the region 1MeV – 4GeV

In the Table 1 we present the range of the parameters that appear in equation (1).

3 Distributions of MB energy per bunch crossing

The next task is to generate distributions of MB energy deposits *per bunch crossing* (BX) for each cell type. We assume the design luminosity $\mathcal{L} = 10^{34} cm^2 s^{-1}$, the inelastic cross-section $\sigma_{pp}^{inel} = 70mb$, and a bunch crossing rate $f = 31.6$ MHz, which takes into account the fraction of empty bunches. Then the mean number of collisions per bunch crossing is $M = 23$.

Parameter	Range of values
P1	250–16000 GeV^{-1}
P2	250–1300 GeV^{-1}
P3	6–1350 GeV^{-1}
P4	20–80 GeV^{-1}
P5	0.3–64 GeV^{5}

Table 1: Ranges of the parameters in equation (1) for all TileCal cells

The simulated MB events allow to estimate the probability *per collision* that each cell be hit. This is

$$P_{col} = \frac{\text{entries}}{5000 * 64 * 2}$$

where entries is the number of energy deposits greater than 1 MeV in a cell; the denominator accounts for the 5000 MB collisions and the adding together of the cells with phi and eta symmetry (the cell D0 has been treated differently). Then an estimate of the mean occupancy of each cell type (the probability that it be hit per bunch crossing) is

$$P_{BX} = 1 - (1 - P_{col})^{23}.$$

To generate MB energy spectra per bunch crossing, for each cell type, energies were randomly generated from the parametrizations of Eq. 1. Each energy point is the sum of m terms, with m randomly extracted from a Poisson distribution of mean M=23, and each term added or not according to the hit probability P_{col}. At the same time, the occupancy per individual bunch crossing is calculated.

The obtained energy distributions in four characteristic cells, that span the eta range and belong to all three samplings, are shown in Fig. 4. The plots in Fig. 5 are crosschecks on the procedure just described. The occupancies estimated with the formula given above for P_{BX} are compared to the simulation average values of the occupancies per individual bunch crossing obtained from the parametrized distributions. In the right-hand part of the Fig. 5, the estimates of the average energies per BX from 5000 MB events are compared to the results from the parametrizations of the energy spectra. In the two sets of three plots the cells are grouped according to the samplings A, BC and D (including ITC cells in the latter). In almost all cases the agreement between estimates based on the MB event sample and the results of the higher-statistics event generation using the parametrizations is excellent.

It should be pointed out that the mean energies/BX obtained here are roughly 30% lower than those estimated for the TileCalorimeter TDR (see also [4]). The differences may be associated to differences in the generation codes (PYTHIA) and interaction codes (GEANT3) used to produce the two MB event samples. Also note that in the TileCal TDR only 17.5 collisions/BX were assumed whereas the simulation given here assume 23 collisions/BX.

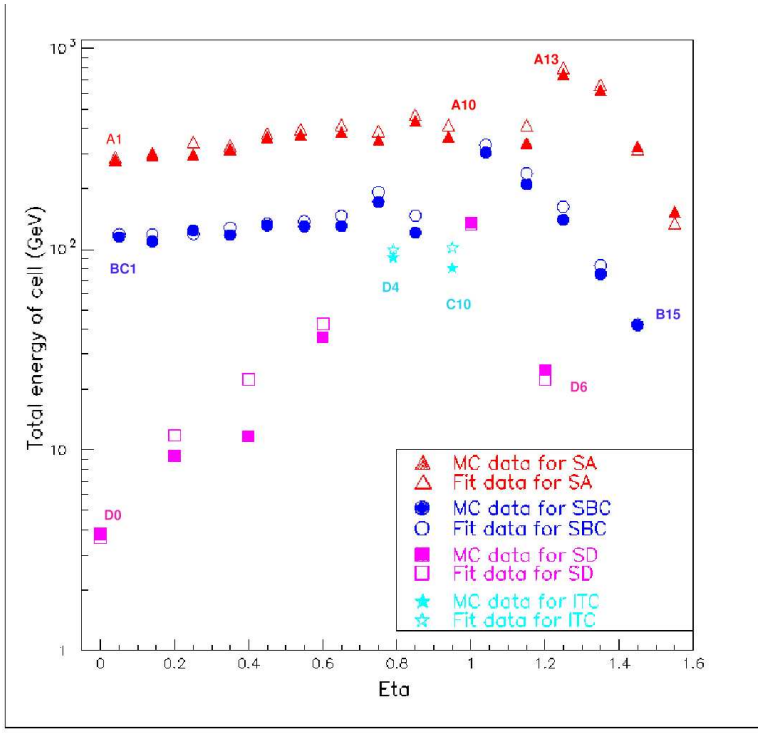

Figure 3: Total energy in 64×2 cells of same $|\eta|$ over 5000 MB collisions, compared to the parametrisation of equation (1). Cells of different depth samplings (Samplings A, BC and D and ITC cells) are shown separately (energy deposits of less than 1 MeV were excluded from the sums and further analysis)

4 Simulating MB currents

Next, we are going to simulate the formation of the MB signal in the TileCal "integrators" mentioned in the Introduction, by generating as many MB energy depositions as necessary to reach a stable signal on the integrator output. Once a stable signal is reached, distributions of such signals will be generated in order to evaluate their spread and from that estimate how many measurements are needed to obtain a statistical error of 1% of the mean.

Let us first review the formulas that tie the MB deposition rate to the integrator output voltage and the PMT currents. At every bunch crossing, the MB energy deposit in a cell will increase the output voltage of the integrator, which will then exponentially decrease with a time constant $\tau = R_{eff}C$, where R_{eff} and C are the effective resistance and the capacitance in the integrator feedback loop. Using energy units (independent of conversion factors) for the MB integrator output, we write

$$E(t_N) = \sum_{n=0}^{N} E_n e^{\frac{-(t_n - t_0)}{\tau}}$$

where E_n is the MB energy deposit at bunch crossing n and $t_n - t_0$ is the time since t_0, when

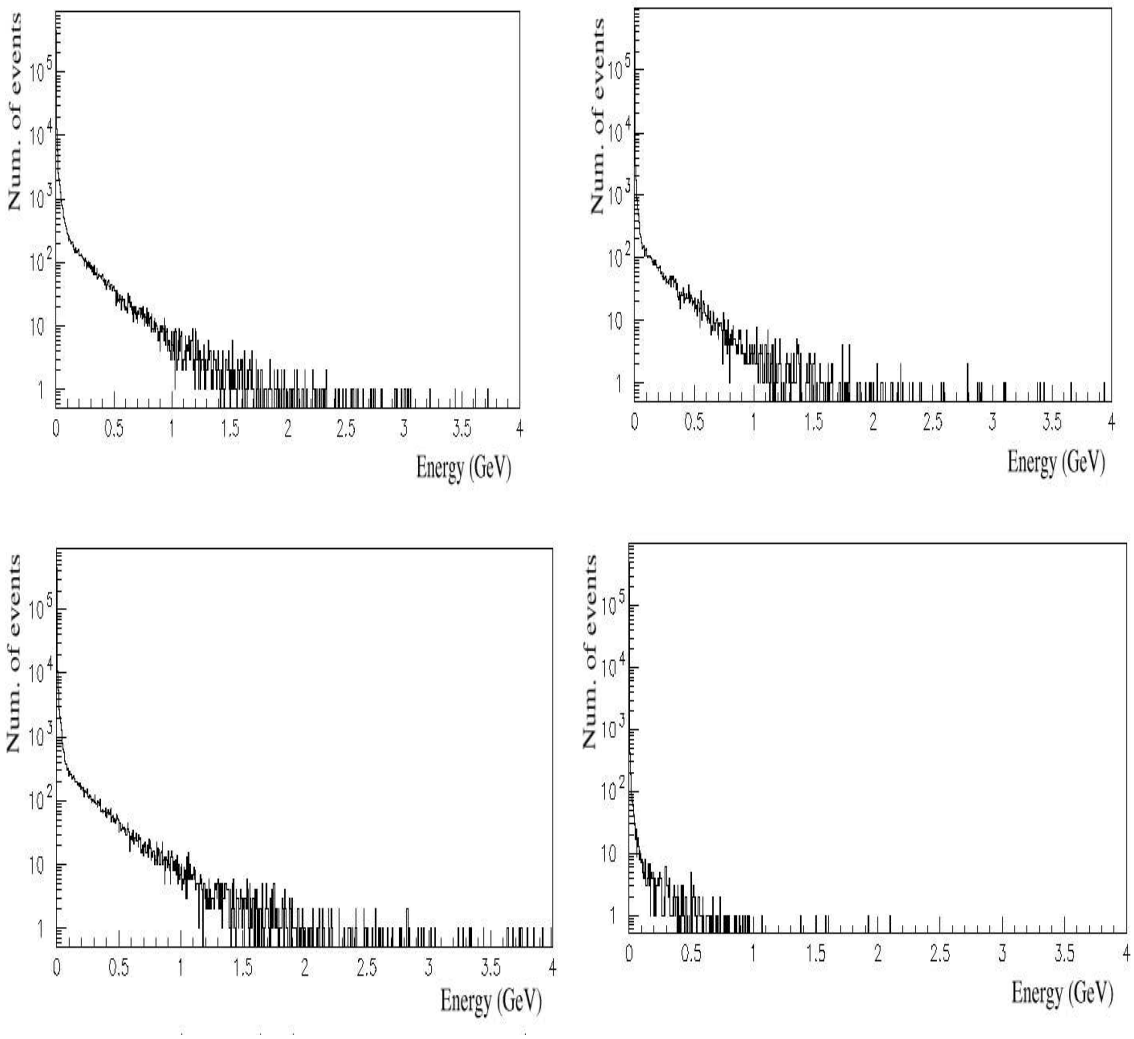

Figure 4: Energy spectra (in 5 MeV bins) generated according to the parametrization of Eq.1 (top-left: cells A1; top-right: cells A16; bottom-left: cells B11; bottom-right: cells D0)

$E_0 = 0$. As $n = (t_n - t_0) \cdot f$, we can rewrite the last expression as:

$$E_N \equiv E(t_N) = \sum_{n=0}^{N} E_n e^{-\frac{n}{\tau f}}. \tag{2}$$

Introducing the average energy deposit per BX (the mean value $< E >$ of the MB energy spectrum of the cell) we can write for the output signal when $N \to \infty$:

$$E_\infty =< lim_{N\to\infty} E(t_N) >=< E > \cdot \frac{1}{1 - e^{-1/\tau f}} \approx < E > \tau f$$

where $\tau f >> 1$ for all values of τ, as will be shown below. Note that we take $f = 31.6$ MHz to account for the fact that only 2808 out of 3564 bunches will be "filled" (the effect of the empty bunches on the signal asymptote stability is negligible, because the proton revolution time of about $90\mu s << \tau$).

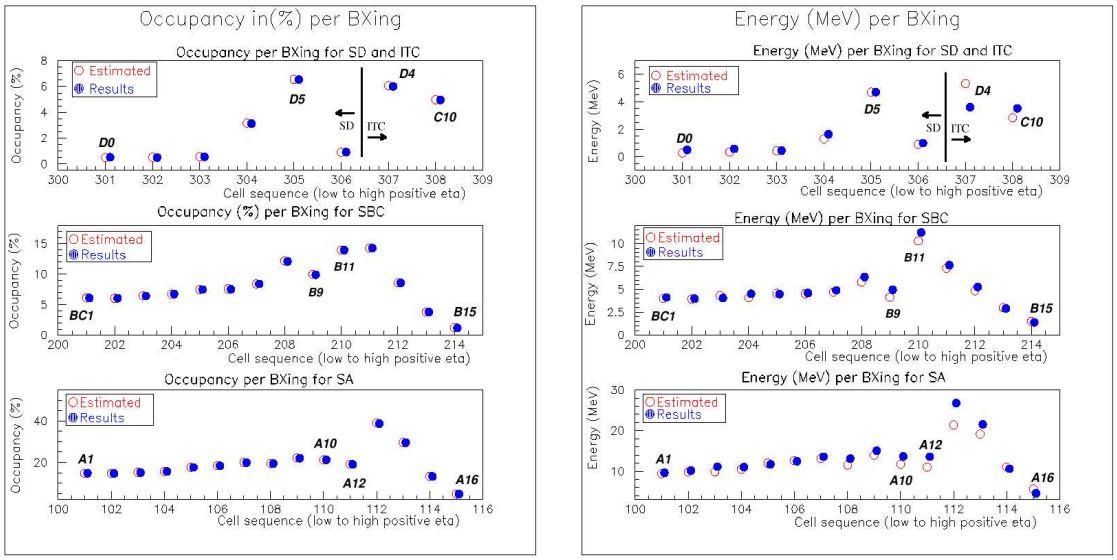

Figure 5: Occupancies and energies per bunch crossing, for all cells, using the MB collisions and using the parametrized energy spectra.

CELL	$\tau = RC$ (ms)
A1	$\tau_2 = 2.5$
A12	$\tau_1 = 0.3$
A16	$\tau_1 = 0.3$
BC1	$\tau_4 = 5.3$
B11	$\tau_1 = 0.3$
B15	$\tau_6 = 9.9$
D0	$\tau_6 = 9.9$
D2	$\tau_6 = 9.9$
D5	$\tau_4 = 5.3$

Table 2: Measured τ values for different cells

The asymptotic value of the charge on the integrator capacitor is given by

$$Q_\infty = k \cdot E_\infty.$$

The conversion factor k is given by $k = \frac{neG}{2}$, where n is the number of photoelectrons per GeV deposited in a TileCal *cell*, e is the electron charge, G is the photomultiplier gain, and the factor of two in the denominator is there because each cell is read out by two photomultipliers (k is set to 1.2 pC/GeV per channel at the test beam).

Finally, the time-averaged current input from each PMTs is:

$$I = \frac{Q_\infty}{C} \cdot \frac{1}{R_{eff}} = \frac{E_\infty \cdot k}{\tau} \tag{3}$$

and the corresponding average output voltage is $V = IR_{eff}$. For the cells exposed to the larger MB energies, the value of R_{eff} is chosen to comply with $V < 5V$, the input range of the ADC; for the cells with smaller average signals, R_{eff} is chosen to be large enough to have a signal well above noise. The time constants τ used for a few characteristic cells are given in Table 2. The values were measured using integrator calibration signals, as reported in the Appendix.

Using the procedure described in the previous sections to generate very large numbers of MB signals/BX in the Tilecal cells, and the expression (2), we obtain asymptotic signal values as shown in Fig. 6, in which the signal buildup is calculated over five time constants. The fluctuations in the MB current when plateau is reached are due to the fluctuations in the primary energy deposits. We generate 1000 signals for each TileCal cell, always starting from zero energy and accumulating signals over five time constants. The signal distributions for a few characteristic cells are given in Fig. 7. The distributions are reasonably symmetric, however moderate tails are visible on the high side. In the gaussian approximation, we can estimate how many such measurements are needed for their mean to approximate the true mean with 1% statistical error. Using

$$\frac{\sigma}{<\mu>} = \frac{RMS}{<\mu> \cdot \sqrt{N}} \leq 1\%$$

where $<\mu>$ stands for the mean value and N is the number of measurements one obtains

$$N \geq \Big(\frac{RMS}{<\mu>}\Big)^2 \cdot 10^4. \tag{4}$$

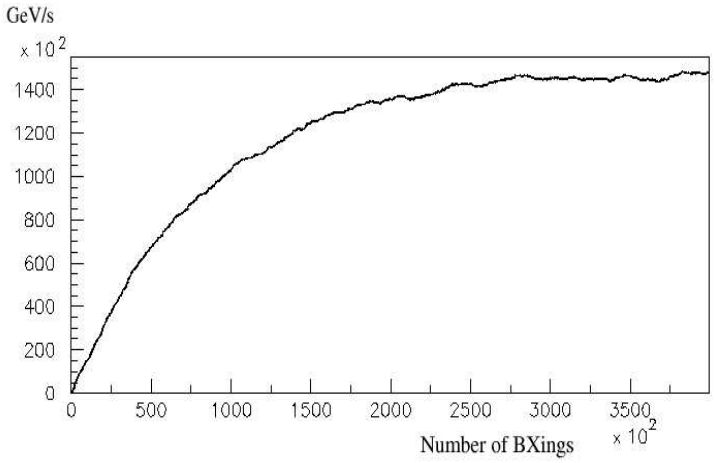

Figure 6: Simulation of signal buildup over five time constants for one PMT in cell A1.

The number of measurements needed to reach 1% precision are given at the end of the next section.

5 An aside: Can we avoid summing billions of MB energies?

The procedure of adding up roughly 10^9 numbers - each of them randomly generated from the parametrized energy distributions described in Section 2 - for every different TileCal cell is quite CPU- intensive. This begs the question of whether this large numerical calculation can be entirely avoided, by applying the Central Limit Theorem (CLT) to the distributions in expression (2). The question is reinforced by the fact that the distributions of simulated measurements in Fig.7 appear almost Gaussian. Every entry in these plots is the sum of a large number of random variables, described by the parametrizations of (1), which do have finite mean and RMS values as required in order to apply the CLT.

The Central Limit Theorem states that the sum of n random variables, each with mean μ_j and variances σ_j^2, not necessarily equal, approaches a gaussian distribution with mean $\sum_j \mu_j$ and variance $\sum_j \sigma_j^2$. In our case, the n variables are the terms in the sums of equation (2). The mean and variances of our signal distributions should be given by

$$\mu = \sum_{n=0}^{N} \mu_n e^{-n\lambda} \qquad \sigma^2 = \sum_{n=0}^{N} \sigma_n^2 e^{-2n\lambda}$$

where $\lambda = 1/\tau f$ is the inverse number of bunch crossings per integrator time constant and n runs over the number of BX considered. For large N, we can write:

$$\lim_{N\to\infty} \mu = \mu_0 \frac{1}{1-e^{-\lambda}} \approx \frac{\mu_0}{\lambda}$$

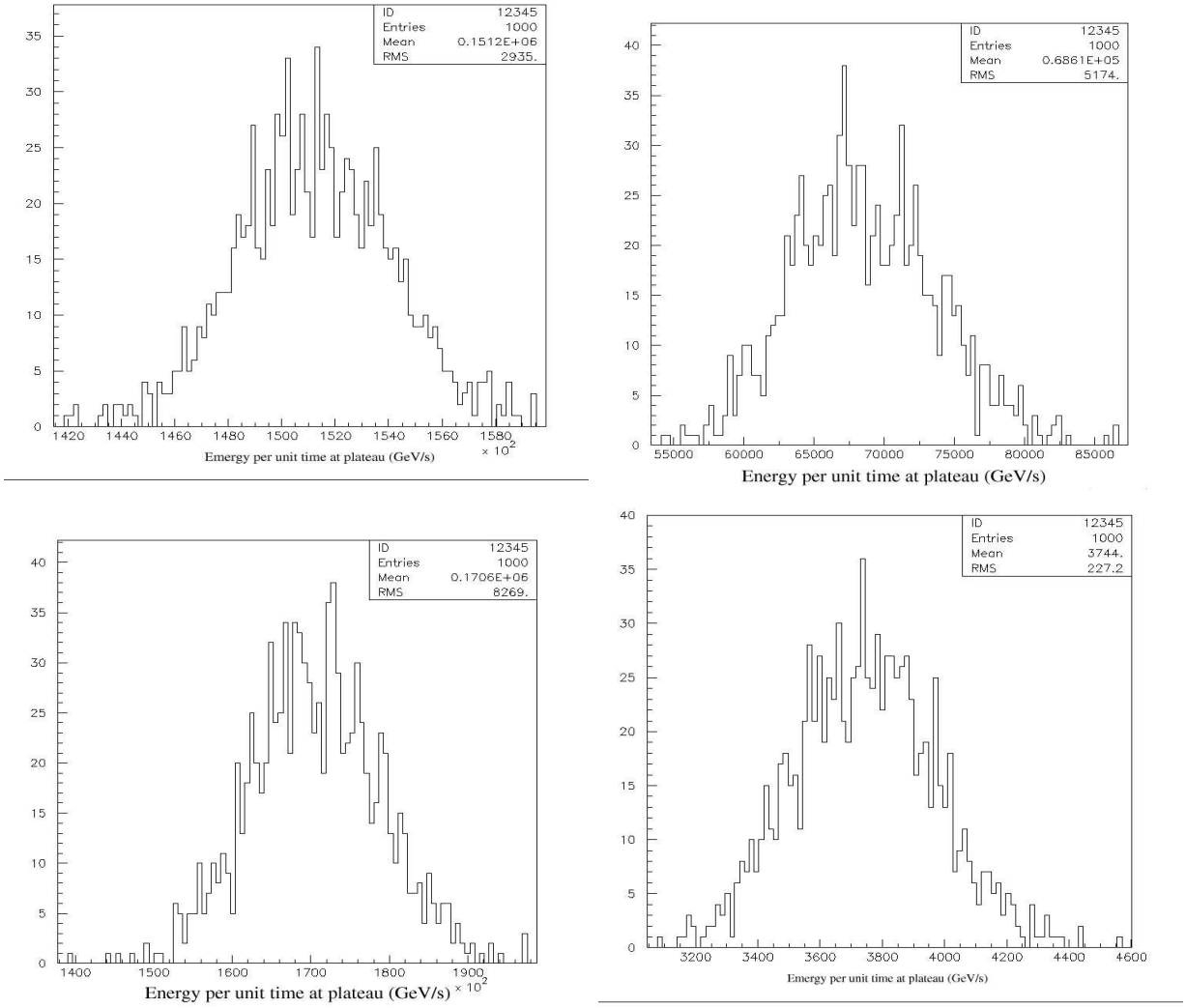

Figure 7: Distributions of signals simulated over five time constants for a few representative cells (top-left: A1; top-right: A16; bottom-left: B11; bottom-right: D0)

CELL	μ (MeV)	$\left(\frac{\sigma}{\mu}\right)_{numeric}$	$\left(\frac{\sigma}{\mu}\right)_{CLT}/\left(\frac{\sigma}{\mu}\right)_{num}$	$\frac{N_{num}}{\text{PMT}}$
A1	9.64	0.0194	0.95	**4**
A12	13.55	0.0516	0.96	**27**
A16	4.56	0.0936	0.91	**88**
BC1	4.10	0.0218	0.89	**5**
B11	11.21	0.0577	0.95	**33**
B15	1.38	0.0284	0.90	**7**
D0	0.50	0.0607	0.45	**8**
D2	0.45	0.0694	0.47	**11**
D5	4.70	0.0196	0.92	**4**

Table 3: Number of measurements per PMT needed to achieve 1% accuracy level and comparison between numerical method and prediction of CLT

$$\lim_{N\to\infty}\sigma^2 = \sigma_0^2\frac{1}{1-e^{-2\lambda}} \approx \frac{\sigma_0^2}{2\lambda}$$

where $\mu_0 = \mu_n$ and $\sigma_0 = \sigma_n$ for all n. Using this result we proceed to estimate the means and RMS (labeled here as σ) of the distributions in MB measurements from their parametrized expressions, using for each cell the number of BX per time constant N_B. Then from the CLT we are led to expect that

$$\frac{\sigma}{\mu} = \sqrt{\lambda}\frac{\sigma_0}{\mu_0}\frac{1}{\sqrt{2}} = \frac{\sigma_0}{\mu_0}\frac{1}{\sqrt{N_B}}\frac{1}{\sqrt{2}}.$$

In the table (3) we give the numerical results for σ/μ calculated from the CLT and numerically (from the distributions exemplified in Fig. 7), for a few characteristic and a few extreme cells.

Several observations can be made:
(1) regarding the numerical simulations, it can be seen that even in the worst cases - which correspond to cells that have low occupancy and low signal level, and are therefore exposed to larger fluctuations - a few tens of MB measurements should allow to reach the desired statistical error on the mean of 1%. See Fig.5.
(2) the σ/μ values predicted with the CLT are always smaller than obtained by the numerical simulation, and the discrepancy is entirely correlated with μ, being larger when μ is smaller. This is not altogether a surprise, however its causes are not clear, particularly considering the very large number of terms in the MB energy measurements performed by the integrators.

6 Conclusions

We have developed a method to simulate Minimum Bias energy measurements with the existing TileCal electronics. A realistic numerical simulation of signal formation requires billions of bunch crossings; therefore we have parametrized the simulated energy deposition spectra and generated large random samples of MB signals extracted from these spectra.
The generated distributions of MB signal measurements are fairly well-behaved, and indicate that

with a few tens of measurements one may obtain 1% statistical precision in measuring the MB response of all cells of TileCal, which we take to be adequate for monitoring the response of the calorimeter.
This conclusion is subject to the uncertaintes typical of the MB models (and less critically, to the accuracy of the particle interaction codes). However, the results are encouraging and are being used by the authors of this Note in specifying the MB data acquisition system.

References

[1] ATLAS Collaboration, ATLAS Tile Calorimeter Technical Design Report, CERN/LHCC/96-42, 1996

[2] G.Blanchot *et al.*, Cell Intercalibration and Response Uniformity Studies Using a Movable Cs137 Source in the TILECAL 1994 Prototype, N-44, 20-12-1994

[3] E.Starchenko *et al.*, Cesium Monitoring System for Atlas Tile Hadron Calorimeter, NIM *A* **494** 381-384, 2002

[4] G.Blanchot *et al.*, Long Term Monitoring of Tilecal Response in ATLAS: Design Considerations, N-45, 20-12-1994

Appendix: *In situ* measurements of the 3in1 integration time constants

Using a JFET input operation amplifier with a low-pass RC filter, the broad spectrum of pileup signals in the Tile Calorimeter is converted to a quasi-DC voltage. The output voltage scale is set by remotely selected resistor network which range from $R_{eff} = 2.8$ to $R_{eff} = 102$ MΩ. The integration time constants scale proportionally to the selected gains. The output voltages are digitized by dedicated ADCs with an adjustable sampling rate. Hereafter we present the first direct measurements of the integration time constants of the RC filter circuits, based on a significant sample of 3in1 cards.

The data were taken in November 2002 on six production Super-Drawers, altogether equipped with 244 3in1 cards. The data taking included feeding a voltage step into a selected 3in1 channel and sampling the exponentially rising voltage by the ADC. To have more samples of the initial signal rise, the data were taken at three ADC sampling rates: 110, 100, and 90 Hz. A typical integrator response to the injected step function is shown in Fig. 8.

After correcting for pedestals the data were fitted to the function:

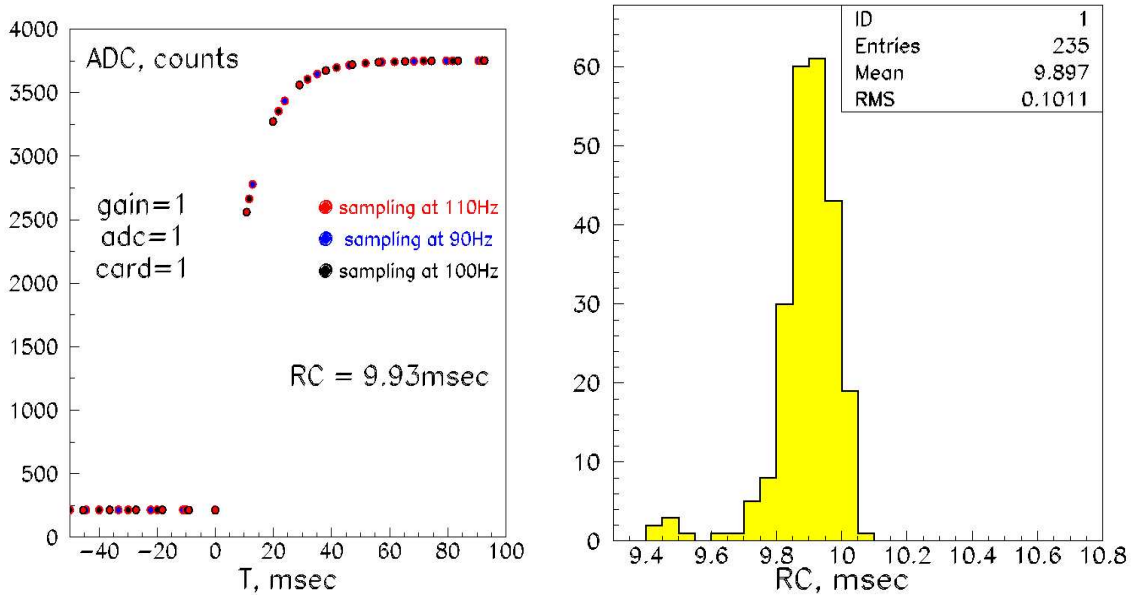

Figure 8: Left: Response to injected step voltage; Right: Distribution of the fitted integration time constants

$$Amp \cdot \left(1 - exp^{-\frac{Del+t}{RC}}\right)$$

where t is time in *msec* since the first sampling. The parameters of the fit are: *Amp* - amplitude of the integrator response (ADC counts), *Del* - delay between charge injection and the first sampling (*msec*), and *RC* - integration time constant (*msec*). The data errors to evaluate χ^2 were estimated from the pedestal fluctuations. The χ^2 was minimized using the MINUIT package.

Based on the quality of the fit several channels were excluded from further analysis. Detailed studies of these cases traced the data quality problem in these channels down to mis-functioning of the CANbus communication link. Thus the integration time constants for these cards are not expected to be any different from the ones of the main sample. The distribution of the integration time constants for 235 cards and the highest gain (and RC time constant) is shown on Figure 8. The systematic error on the measurement was estimated from several data runs taken in different conditions, such as with the photomultiplier HV set ON. Convoluted with systematic error of the fit it resulted in 3% overall systematic error of the measurement.

The results for the integration time constants and the respective gains, measured independently, are summarized in the following table, where the time constants corresponding to the lowest resistor values (Gain 5 and 6) could not be extracted :

Gain 1 :	$(9.90 \pm 0.10 \pm 0.30)$ *msec*	(102.1 ± 1.0) $M\Omega$
Gain 2 :	$(7.40 \pm 0.08 \pm 0.22)$ *msec*	(76.1 ± 0.8) $M\Omega$
Gain 3 :	$(5.31 \pm 0.05 \pm 0.16)$ *msec*	(54.9 ± 0.8) $M\Omega$
Gain 4 :	$(2.70 \pm 0.03 \pm 0.08)$ *msec*	(29.0 ± 0.5) $M\Omega$

Gain 5 :	$(26.1 \pm 0.4)\ M\Omega$
Gain 6 :	$(2.82 \pm 0.03)\ M\Omega$

Here the first error represents the channel to channel variation in the measured values and the second one, when given, stands for the systematic uncertainty of the method.