FAMOS, a FAst MOnte-Carlo Simulation for CMS

F. Beaudette, CERN, Geneva, Switzerland

Abstract

An object-oriented FAst MOnte-Carlo Simulation (FAMOS) has recently been developed for the CMS (Compact Muon Solenoid) experiment to allow rapid analysis of all final states envisioned at the Large Hadron Collider while keeping a high degree of accuracy for the detector material description and the related particle interactions. After a brief description of the program architecture, the electromagnetic calorimeter fast simulation is emphasized. The material effects in the tracking device, the shower and detector simulation as well as the first results and optimizations are discussed.

DETAILED VS. FAST SIMULATION

In 2005, CMS will publish its Physics Technical Design Report (TDR). A large amount of simulated data will be required to carry out the analysis. The main tool up to now available to generate these events is the detailed simulation.

The CMS detailed simulation

The detailed simulation of CMS consists of three programs. The event generator simulates a proton-proton or heavy ions collision with a centre-of-mass energy of 14 TeV, then OSCAR (Object oriented Software for CMS Analysis and Reconstruction), computes the propagation and the interaction of the generated particles in the detector. The OSCAR program is based on Geant4 [1] and contains an accurate description of the detector geometry and of its materials. The digitization, i.e., the simulation of the read-out electronic response, is made by the reconstruction program ORCA (Object Oriented Reconstruction for CMS Analysis). In the process, pile-up events can be superimposed to the hard collision to simulate the multiple collisions occuring at each bunch crossing.

The typical timing of this "full" simulation is between four minutes for a $Z \rightarrow e^+e^-$ event and ten minutes for a $Z \rightarrow q\bar{q}$ event on a 1GHz computer.

With such a timing, it will be impossible to simulate all the samples needed for the physics TDR. As a result, a fast simulation able to generate quickly, large and reliable samples is needed.

FAMOS

Since November 2003, a particular effort has been set on the development of FAMOS, the CMS fast simulation. It has to be as accurate as possible with a timing below one second per event. One of the original feature of FAMOS with respect to other fast simulations (CMSJET,Atlfast [2]) is its full compatibility with the standard reconstruction software. Indeed, the simulated objects have the same format as in the reconstruction program. Thus, it is possible to run the standard reconstruction algorithms on the fast simulation output. As a result, the user analysis code can run on the fast simulation output with only some minor changes.

With FAMOS, the full chain is done at once, from the event generation to the reconstruction and analysis. It is currently interfaced only with Pythia [3], but a particle gun can also be used for detailed studies of single particles. It is also possible to use an external file produced by any generator as an input.

As far as the simulation is concerned, the organization of FAMOS reflects the different sub-detectors of CMS. The pixel and silicon detector simulation is obtained with a smearing of the generated particles according to a parameterization of the efficiency and resolution. The resulting tracks can be used as an input by the standard ORCA b-tagging algorithms to identify the heavy flavour decays.

Similarly, the muons are simulated according to a parameterization of the reconstruction efficiency with a smearing. Their isolation can be computed from the calorimeter and tracking information.

In the electromagnetic calorimeter (ECAL) and in the hadron calorimeter (HCAL), instead of directly simulating the reconstructed objects, as it is done in the tracker, the individual energy deposits (hits) are simulated. With these hits, the towers, the clusters, the jets and the missing transverse energy are reconstructed. A parameterization of the electromagnetic showers is used in the ECAL and in the preshower, whereas a smearing technique is, up to now, applied for the hadrons in the HCAL. The electron shower parametrization and ECAL simulation are briefly described in the following.

ECAL SIMULATION

For an accurate ECAL simulation, it is necessary to simulate the propagation of the particles in the tracking device as well as the material effects.

Material effect simulation

The material of the pixel and silicon trackers amount to about 1. radiation length at pseudorapidity $\eta = 1.2$. As a consequence, the particles traversing the tracker experience all sorts of interactions, e.g., Bremsstrahlung for electrons,

photon conversions, energy loss by ionisation and multiple scattering for all charged particles. These effects are all simulated in FAMOS, with visible effects on photon and electron clusters. To do so, each particle is propagated in the 4 T magnetic field in which the CMS tracking device is immersed. The intersections with the tracker layers are determined and the material effects are computed. It is assumed that the tracker is made of pure silicon with a simplified geometry to allow simple and fast formula to be used to compute all the material effects interactions. Thus, despite the high number of tracks per event, these calculations are very fast. On average, only 50 ms per event are required. As can been in Fig. 1, even without any tuning, the spectrum of the Bremsstrahlung photons emitted by 100 GeV electrons are in agreement with the detailed simulation.

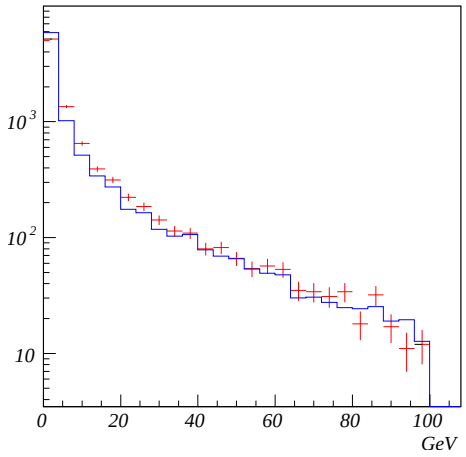

Figure 1: Photon energy spectrum emitted by 100 GeV electrons. The full curve is the fast simulation and the crosses represent the detailed simulation. The vertical scale is arbitrary.

The history of the material effects is recorded in a container which is used afterward by the calorimeter simulation. It will also soon be used by the tracker simulation. Indeed, the momentum and its resolution for electrons are correlated with the Bremsstrahlung photon spectrum.

Once all particles are extrapolated to the ECAL entrance, it is possible to simulate the ECAL response.

Electron shower simulation

The strategy for the electromagnetic particles simulation in the ECAL proceeds in two steps. First, the shower is simulated in a homogeneous medium, then the generated shower is placed in the detector geometry.

Because the ECAL is made of crystals, generating the showers in a homogeneous medium is relevant. For this purpose, the latest Grindhammer [4] electron shower parametrization has been used. This simulation has been developed in the nineties for the H1 experiment where it has been intensively tested on real and simulated data as part of the GFlash package [5]. With this procedure, the detailed simulation is not the starting point of the fast simulation development, but is only needed for the tuning.

In this electron simulation algorithm, each shower consists of hundreds or thousands of energy spots. Since, their number depends almost linearly on the energy, the low energy photon showers generation is not too time consuming. The longitudinal profile of each shower is represented by a gamma distribution. It is split into several longitudinal slices, and the total amount of energy in each slice is computed. The shower-to-shower fluctuations are included at this level, as well as photo-statistics and longitudinal non-uniformity fluctuations. In each slice, the spots are distributed along the radial profile uniformly in φ, taking into account the correlation between the longitudinal and radial fluctuations. The radial profile is modeled by the sum of too similar functions $f(r, R)$ one representing the core of the shower and the other the tail,

$$f(r, R) = \frac{2rR}{(r^2 + R^2)^2} \quad ,$$

where r is the distance to the centre of the shower and R is the mean radius of the core or the tail of the shower. The relative weights of the core and the tail terms as well as the associated R value highly depend on the longitudinal depth.

Even if many spots are generated, this algorithm is fast. A 40 GeV electron shower is simulated in 12 ms. The generated shower must now be transported into the calorimeter geometry.

Detector simulation

The ECAL is made of $\mathrm{PbWO_4}$ crystals. Each crystal has a truncated pyramid shape, is 25 $\mathrm{X_0}$ long, 1 Moliere radius wide. The crystals are arranged with a pseudo-projective geometry both in η and in φ. The barrel part of the detector covers the $|\eta| < 1.48$ region where the two end caps cover the $1.48 < |\eta| < 3.0$ regions. Several essential effects must be taken into account in the detector simulation.

- There are 0.3 mm gaps between the crystals and 7 mm cracks between the modules and the super-modules. The energy deposits in the cracks may be lost whereas there is no evidence of a similar effect in the gaps.

- The electromagnetic showers are not fully contained in the 25 $\mathrm{X_0}$ thick crystals, which leads to a rear leakage. A so-called front leakage also exists at large $|\eta|$ in the barrel, due to the specificity of the geometry.

- The CMS calorimeters are inside the magnetic field. The magnetic field is responsible for an enlargement of the electromagnetic showers, especially in the central region.

- The electronic noise as well as the zero-suppression must be simulated.

Each energy spot of the generated shower is affected to a crystal in the detector. Since the real geometry has to be used, a 3D treatment can be very time consuming. With the standard tools, the treatment of the 10000 spots of a 40 GeV shower exceeds 700 ms, whereas the total simulation has to stay below 1 s / event.

The longitudinal segmentation of the shower generation, however, makes a 2D treatment natural. Indeed, at a given depth, the algorithm gives the position and the energy of the spots in the plane orthogonal to the particle direction. It is thus sufficient to determine once for all, the intersections of the crystals with this plane. The pads obtained are approximative squares. They are simple enough to allow a quick determination of the spot-to-crystal associations. Moreover, instead of the full calorimeter, a limited area can be used to reduce the number of calculations. By default, a 7x7 crystal window is used (Fig. 2).

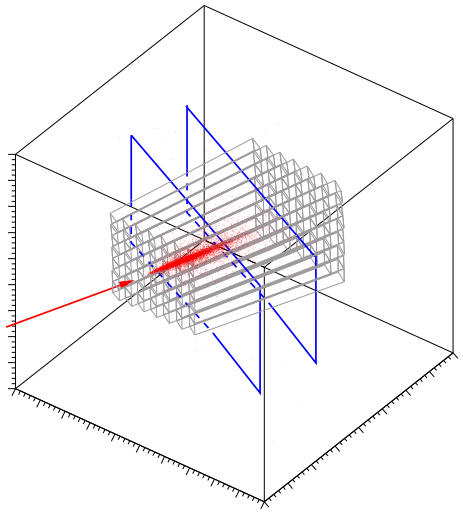

Figure 2: A 40 GeV electron shower in a 7x7 crystal window grid. Two planes orthogonal to the particle direction are represented. Their intersections with the crystals are approximate chessboards.

The 2D treatment also allows the detector effects to be easily included. The rear leakage is automatically taken into account and the gaps/cracks difference can be implemented as follows. Because no energy is lost in the gaps, they are simply filled by moving the boarder of one of the neighbouring crystals. In contrast, new pads, attached to those of the neighbouring crystals, are created for the cracks, with a certain energy loss. The shower enlargement due to the magnetic field is taken into account by shrinking the grid. The front leakage can be also included at this level.

At each depth of the longitudinal segmentation, the grid is determined and reorganized as explained above. Then, each spot is quickly attributed to a pad. Since each pad corresponds to a crystal, the total amount of energy in each crystal is readily computed. The electronic noise is added, the zero-suppression is applied, and the result is turned into standard hits that can be used to reconstruct electromagnetic clusters.

The Grindhammer parameterization applies only to electron showers. The photons are first converted into two electrons at a variable depth and the previous treatment is applied to the two electrons. The same grid is used for both of them to reduce the number of calculations.

First results

Photons (unconverted after the tracker material) with a transverse energy of 35 GeV in the central calorimeter have been used to compare the reconstructed clusters in the detailed and fast simulations.

The comparison of the overall energy in the cluster is shown in Fig. 3. Even with essentially no tuning, the agreement between both simulations in terms of absolute energy and resolutions is quite satisfactory.

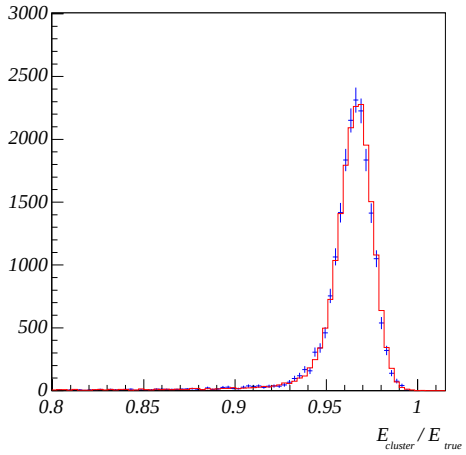

Figure 3: Reconstructed cluster energy over true energy. The full curve is the detailed simulation and the crosses represent the fast simulation. The vertical scale is arbitrary.

The number of crystals in a cluster (Fig. 4) is sensitive to the transverse shape of the shower. The ratio of the energy in the most energetic crystal over the nine most energetic crystal, S_1/S_9 is an other measurement of the transverse shape of the shower (Fig. 5).

As it can be seen on Figs. 4 and 5, both variables compare well in the full and the fast simulations. The showers are, however, slightly too large in FAMOS. Some tuning is consequently necessary. The behaviour of the shower close to the cracks will also need to be adjusted to the detailed simulation, or better, to the test-beam data.

It has been checked that the position resolutions are very well reproduced in the fast simulation. Altogether, the fast simulation, with the original electron shower parametrization and a $2\,X_0$ longitudinal segmentation gives very satisfying results with a timing of 44 ms per photon.

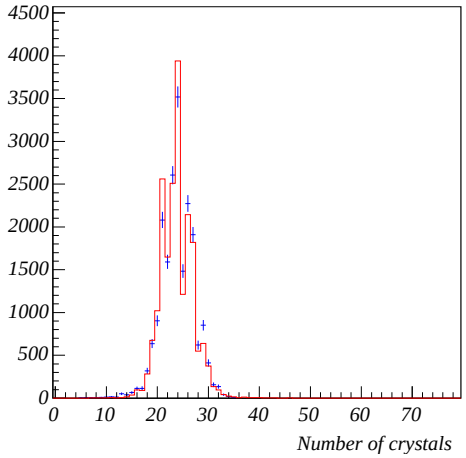

Figure 4: Number of crystals in a cluster. The full curve is the detailed simulation and the crosses represent the fast simulation. The vertical scale is arbitrary.

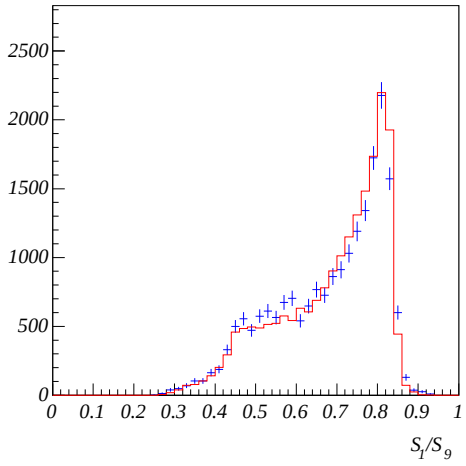

Figure 5: S_1/S_9. The full curve is the detailed simulation and the crosses represent the fast simulation. The vertical scale is arbitrary.

Timing vs. tuning

The timing can be improved as follows. The size of the grid could be reduced, but it has been checked that a 7x7 crystal grid is the optimal one. It is however possible to save time in the shower generation process. First, the size of the longitudinal steps can be increased to $5\,X_0$ instead of $2\,X_0$. This reduces the number of grids which have to be calculated. Then, the number of spots in the centre of each shower can be divided by a factor of 10, without any impact on the previous distributions, since the energy of each spot is multiplied accordingly. A 12 ms per photon timing is achieved, which corresponds to 600 ms per $Z \rightarrow e^+e^-$ event, e.g., 400 times faster than the detailed simulation.

The significant difference in the timing between the single photons and the electrons from Z is mainly due to the large number of radiated photons emitted by each electron. The whole sophisticated treatment described previously is applied to each of them. A simplified treatment can be applied to these photons, the size of the grid can, for instance, be reduced. A factor larger than 1000 in speed with respect to the full simulation can certainly be achieved.

CONCLUSION

The fast simulation of CMS has been presented. The ECAL simulation is currently the most achieved part. It is four hundred times faster than the detailed simulation but there is plenty of room for improvement. A factor in excess of 1000 is within reach. This simulation is accurate even with a very preliminary tuning. A shower parameterization is being developed for the HCAL.

The FAMOS program is also an user-friendly tool. Since the full chain from the event generation to the analysis is done at once, and with the help of the large number of examples included in the package, it is possible to get results within a few hours. As a consequence, FAMOS will certainly soon be a good way for the new user to gets started with the CMS software.

The first release aimed at physics is scheduled for December 2004.

ACKNOWLEDGMENTS

I would to thank P. Bloch, P. Janot and A. de Roeck who gave me the opportunity to participate in this conference.

REFERENCES

[1] `http://wwwasd.web.cern.ch/wwwasd/geant4/geant4.html`

[2] S. Abdullin *et al.*, CMSJET, CMS TN/ 94-180, `http://www.hep.ucl.ac.uk/atlas/atlfast/`

[3] `http://www.thep.lu.se/~torbjorn/Pythia.html`

[4] G. Grindhammer, S. Peters, The Parameterized Simulation Of Electromagnetic Showers In Homogoneous And Sampling Calorimeters. hep-ex/0001020 (1993)

[5] G. Grindhammer et al., The fast simulation of electromagnetic and hadronic showers, Nucl. Instrum. Methods Phys. Res., Sect. A 290, 469 (1990)