The Fast Simulation of The CMS Experiment

Rahmat Rahmat
University of Mississippi, USA
E-mail: rahmat@fnal.gov

Rob Kroeger
University of Mississippi, USA
E-mail: kroeger@phy.olemiss.edu

Andrea Giammanco
Université Catholique de Louvain, Louvain-la-Neuve, Belgium
National Institute of Chemical Physics and Biophysics, Tallinn, Estonia
E-mail: andrea.giammanco@cern.ch

Abstract. The CMS collaboration has developed a fast Monte Carlo simulation of the CMS detector with event production rates \( \sim 100 \) times faster than the GEANT4-based simulation, with comparable accuracy for most of the physics analysis. We discuss basic technical principles of the CMS Fast Simulation and their implementation in the different components of the detector, and we show a few comparisons with the full GEANT4-based simulation and CMS data.

1. Introduction
A framework for fast simulation of particle interactions in the CMS detector FastSim has been developed, embedded in the overall simulation, reconstruction and analysis framework of CMS CMSSW[1]. The CMS fast simulation is intended to be used for most physics analysis, in particular those requiring the generation of many samples to span a wide phase-space region (e.g., SUSY searches) or analyses where background samples of manageable size resulting from large cross section processes can only be obtained by the selection of events based on final reconstructed objects.

The central feature of the CMS detector is a superconducting solenoid providing a field of 3.8 T. Located within the solenoid are the silicon pixel and strip tracker, the crystal electromagnetic calorimeter (ECAL) and the brass/scintillator hadron calorimeter (HCAL). Muons are measured in gas-ionization detectors embedded in the steel return yoke. In addition to the barrel and endcap detectors, a quartz-fiber Cherenkov detector extends the jet acceptance to \( |\eta| = 5 \), where the pseudorapidity \( \eta \) is defined as \( \eta = - \ln |\tan \frac{\theta}{2}| \), where \( \theta \) is the polar angle of the particle or jet trajectory with respect to the counterclockwise beam direction. Further forward, the CASTOR detector extends the calorimetric coverage in the pseudorapidity range \( 5.2 < |\eta| < 6.5 \). A much more detailed description of CMS can be found elsewhere [2].
The computer time needed to simulate and reconstruct an event in the fast simulation chain is about 2 orders of magnitude smaller than that needed in the full chain, based on GEANT4 [3] for the detector simulation (henceforth called full simulation), for a level of agreement aimed at the percent level or below.

As output, the fast simulation delivers a series of “high-level objects”, such as reconstructed hits for charged particles in the tracker and muon layers, energy deposits in calorimeter cells, which can then immediately be used as inputs of the same “higher-level algorithms” (track fitting, calorimeter clustering, $b$ tagging, electron identification, jet reconstruction and calibration, trigger algorithms, etc.) as in the full reconstruction and analysis chain.

In the following, after a brief summary of how the code deals with the simulation of particles passing through the main component of the CMS detector, a choice of plots and results will be listed to demonstrate the present status of the FastSim description when compared with the data taken in the first runs of LHC at 7 TeV.

2. The Physics Processes

The input of the fast simulation is a list of particles (originating from an event generator or a simple particle gun) characterized by their momentum and origin vertex, with mother and daughter relationships to follow the various decay chains in the event. Upon user request, each of the (quasi)-stable particles in this list is then propagated in the CMS magnetic field to the different layers of the various CMS subdetectors, which it may interact with. While propagating, these quasi-stable particles are also allowed to decay according to their known branching fractions and decay kinematics. The particles resulting from the interactions with the detector layers or from the decays in flight are added to the original list, and propagated/decayed in the same way. Events from pile-up interactions in the same bunch crossing as the original event are read from pre-generated files, added to the list according to a user-defined distribution, and follow the same treatment.

The interactions simulated at present in the fast simulation are

(i) electron Bremsstrahlung;
(ii) photon conversion;
(iii) charged particle energy loss by ionization;
(iv) charged particle multiple scattering;
(v) nuclear interactions;
(vi) electron, photon, and hadron showering.

The first five processes are simulated for particles traversing the thin layers of the tracker (Section 3.1), while the latter is parameterized in the electromagnetic (Section 3.2) and hadron (Section 3.3) calorimeters. Muons propagate through the tracker, the calorimeters and the muon chambers with multiple scattering and energy loss by ionization taken into account during the propagation (Section 3.4).

The probability of a nuclear interaction is calculated from the elastic and inelastic cross sections taken from experimental measurements [4]. When a nuclear interaction occurs, a simulated nuclear interaction event is randomly chosen from a library of 2.5 million events simulated with GEANT4, according to the hadron species and its energy, and a random rotation is performed around the hadron direction.

2.1. Pile-up events

With an average of $\approx 2$ pp interactions per bunch crossing during the 2010 run, $\approx 10$ in 2011, and possibly a factor 4-5 more in 2012, event pile-up is a crucial ingredient of any LHC simulation. Only in-time pile-up is simulated in FastSim, i.e., additional collisions in the same
bunch crossing as the primary (hard) event. Out-of-time pile-up (i.e., signals coming from bunch crossings before and after the event of interest) is not simulated at the time of writing, but its inclusion is in progress and the deployment of this development is planned for the analysis of 2012 data.

Pile-up is simulated by overlaying minimum bias pp collisions, randomly chosen from a pregenerated sample of one million events. In order to avoid correlations between event subsets overlapped to the same pile-up event sequence, a pile-up sample is never reused in the same order. We are working to get the pile-up distribution treatment fully consistent with the full simulation one.

Two options are available for the distribution of the number of pile-up interactions:

- Poisson distribution (default);
- arbitrary distribution of the number of additional interaction events defined by a user.

In the first case the user must only provide the desired average number of pile-up interactions, while in the second case a normalized vector is used as probability distribution. In both cases the number of actual interactions is stored per each event, allowing to reweight a posteriori to match the pile-up profile observed in data.

3. Detector Simulation

3.1. Tracker Response

A simplified version of the tracker geometry is used, made of over 30 thin nested cylinders representing the sensitive layers of, from inside to outside, the pixel detector (3 barrel layers and 2 endcap disks), the 4 tracker inner barrel layers, the 3 tracker inner disks, the 6 tracker outer barrel layers and the 9 tracker endcap disks, interleaved with non-instrumented cylinders with dead material (cables, support, etc.) [2]. The material, assumed to be pure silicon, is also assumed to be uniformly distributed over each cylinder barrel (respectively, endcap). A “radiography” in the transverse plane of this simplified geometry is shown in Figure 1a, obtained from the vertices of converted photons in a large number of simulated events. For comparison, the higher level of details present in the full simulation is visible in Figure 1b.

![Fast Tracker radiography](images/fast_tracker_radiography.png) ![Full Tracker radiography](images/full_tracker_radiography.png)

**Figure 1.** A radiography of a quarter of the simulated tracker geometry in the (a) fast and (b) full simulation.

The thickness of each layer was tuned in the fast simulation to reproduce the number of Bremsstrahlung photons, in full simulation, above a certain energy threshold radiated by energetic electrons traversing this layer. The same level of agreement between the fast and full simulations is obtained for each single layer. After tuning, the total number of radiation lengths traversed in the tracker reaches $1.42X_0$ at a pseudorapidity $\eta = 1.65$, in agreement with
the full geometry. This agreement demonstrates in addition that the Bremsstrahlung model implemented in FastSim reproduces that of the full GEANT4 [3] simulation.

The complete magnetic field map is used for the track propagation between two surfaces. While being propagated in the magnetic field through the tracker layers, charged particles experience multiple scattering and energy loss by ionization. The intersections between the modified trajectories and each tracker layer define the position of “simulated hits”. Each simulated hit is turned with a certain efficiency to a “reconstructed hit”, the position of which is obtained from a Gaussian smearing of the simulated hit position. In the silicon strip tracker, the Gaussian resolution in each of the two directions (longitudinal and transverse to the beam direction), obtained from a fit of the residuals with respect to the reconstructed charged particle tracks in the full simulation, is treated as a constant for each layer. In the pixel detector, the Gaussian resolution in each of the two directions is parameterized according to the pixel cluster size (itself generated according to its fully simulated $\eta$-dependent distribution) and on the incident angle of the particles with respect to the layer.

3.2. Calorimeter Response to Electrons and Photons

The showers of electrons and photons which impinge on the ECAL are simulated using the Grindhammer parameterization [6][5] [7], as if the ECAL were an homogeneous medium. This approximation is realistic because the CMS calorimeter is made of contiguous crystals. In this parameterization, an electron shower consists of several thousands energy spots, longitudinally distributed according to a $\Gamma$ function, the parameters of which fluctuate from one shower to the other. The deposited energy is integrated over $2X_0$-thick longitudinal slices, including uncertainties due to the limited photo-statistics and the longitudinal non-uniformity in the crystals.

In each slice, the energy spots are distributed in space according to the radial profile of the Grindhammer parameterization and placed into the actual crystal geometry, under the realistic assumption that no energy is lost in the small inter-crystal gaps. The time used in this step is kept to a reasonable value by limiting the two-dimensional spot-crystal assignment to a small $7\times7$ crystal grid (and even smaller for low energy electrons) in a plane perpendicular to the shower longitudinal development. The energy collection simulation is then refined by simulating a number of effects such as rear and front leakage, energy losses in the gaps between ECAL modules, and shower enlargement due to the magnetic field.

In front of the ECAL endcaps, electrons may first cross the preshower. In this case, the corresponding showers are developed from the preshower entrance, and the energy deposit in each layer is converted into a number of MIPs (with related statistical uncertainties), assigned to the relevant strips according to the shower radial profile. Very energetic electrons (above several hundred GeV) can extend their shower substantially beyond the ECAL. A $2X_0$-thick gap between the rear side of the crystals and the entrance of the HCAL is assumed, in which all the energy integrated from the $\Gamma$-distribution tail is lost. The rest is assigned to the HCAL towers according to the shower radial profile at this depth, and the energy of each spot is longitudinally distributed based on Gamma function.

The Grindhammer parameterization only applies to electrons. Photons are first converted in the ECAL (or preshower) material at a varying depth (according to the number of radiation lengths traversed). Each of the resulting $e^+e^-$ pairs gives rise to 2 separate showers generated as explained above along the same longitudinal direction and, therefore, with the same transverse crystal grids.

Finally, at rapidity values not covered by the electromagnetic calorimeter ($|\eta| > 3$), electrons and photons are propagated directly to the forward hadron calorimeter. Here, the detector response is evaluated from the full simulation of electrons with energies of 30, 100, 300, 1000 and 3000 GeV as a function of pseudorapidity, in a way similar to that explained in Section 3.3.
When all electrons and photons are processed, the electronic noise is simulated and the zero suppression applied. At last (when the hadron showers are simulated as well, as is explained in Section 3.3), a list of reconstructed hits is built and stored in a format readable by the clustering algorithms, electron reconstruction, etc. Altogether, the various cluster reconstructed energies in the fast and the full simulations agree at the level of the permil in the ECAL barrel, and at the percent level in the endcaps, for energies ranging from 1 GeV to 1 TeV.

3.3. Calorimeter Response to Hadrons

Charged and neutral hadrons are also propagated to the ECAL, HCAL and HF entrance after their interactions with the tracker layers. Their energy response is derived from the full simulation in the following way. Single charged pions are fully simulated for $P_T$ values of 2, 5, 10, 30, 50, 100 and 300 GeV/c, uniformly distributed in pseudorapidity between $-5.0$ and $+5.0$. The reconstructed energy is collected in $5 \times 5$ HCAL tower matrices and in the corresponding $25 \times 25$ ECAL crystals. The energy distributions are then sliced into $\eta$ bins of 0.1. These distributions are fitted to a Gaussian, the mean value and the sigma of which are tabulated as a function of the energy and pseudorapidity, used in turn to smear the hadron energy response in the fast simulation. Linear interpolation is used for $P_T$ in the range from 2 to 300 GeV/c, and extrapolation for particles with transverse momentum outside this range.

This smeared energy is then distributed in the calorimeters using parameterized longitudinal and shower profiles, with shower-to-shower variations, following an approach similar to that of GFLASH [7] [8].

Calorimetric jets are reconstructed from the energy deposits (reconstructed hits) in each HCAL tower and the corresponding $5 \times 5$ ECAL crystal window [9], with the Iterative Cone algorithm [10] and a cone size $R = 0.5$ in the $(\eta, \phi)$ plane. No jet energy corrections are applied at this level. Figure 2 shows the ratio (mean value and Gaussian sigma) of the reconstructed jet transverse energy to that of the generated jet as a function of jet pseudorapidity and transverse momentum. Despite the disagreement for isolated charged particles, the agreement between the fast and the full simulations is satisfactory. More details about the FastSim performance in jet reconstruction, including more complex algorithms, are given in Section 4.4.

![Figure 2](image-url)  
**Figure 2.** Ratio of the reconstructed to the generated jet $E_T$ as a function of jet pseudorapidity (a) and transverse momentum (b), in the fast (triangles) and the full (squares) simulations.
3.4. Detector Response to Muons

A muon, either coming from the main interaction vertex or produced afterwards in the decay of a jet particle, is propagated in the CMS magnetic field through the calorimeters, the solenoid and the muon chambers.

At the time of writing, the implementation of the calorimeter response to muons is done independently from the propagation of the muon track, and in a way very much similar to that for pions (Section 3.3), with the only difference that the response (mean value and Gaussian sigma) of fully simulated muons is tabulated for transverse momenta of 10, 30, 100, 300 GeV/c, and in $\eta$ regions corresponding to barrel, endcap and forward hadron calorimeters.

At the moment, only the multiple scattering and the energy loss by ionization are taken into account as physics processes for the muon in the fast simulation: therefore, no bremsstrahlung or delta-ray production are simulated. The actual geometry of the CMS muon chambers (DT, CSC and RPC) is taken from a database, and simulation hits are positioned in the detector whenever the track trajectory crosses an active layer of those chambers. Then, these simHits are digitized in the same way as in the full simulation chain, and the resulting digis (raw data equivalent) are fed to the normal local and global muon reconstruction packages, to end up with the final muon objects to be used in the physics analyses.

4. Event Reconstruction and FastSim Performance

4.1. Track Reconstruction and Performance

To save execution time, a “fast tracking” is performed in FastSim, that differs from the standard tracking [11]:

- the seeding efficiency is emulated: it is checked that at least one combination of hits fulfills the seed selection criteria;
- the hits belonging to a given track are fit as in the standard tracking;
- the pattern recognition is also emulated (outlier rejection): hits with large contributions to the $\chi^2$ are removed, and the fit is redone.

As a result of the fact that no real pattern recognition is performed, there are no fake tracks (the fraction of fakes among the tracks with at least 8 hits in the standard tracking is well below 1%). This procedure, together with the fact that two hits simulated at the same position are never merged into a single reconstructed hit, may have to be revised in a very high luminosity scenario. The good agreement with the current data is shown in Fig. 3 for a few representative variables.

4.2. Electron and Photon Reconstruction and Performance

Electrons are reconstructed by matching an ECAL supercluster with a track reconstructed by a Gaussian Sum Filter [12]. Photon reconstruction instead vetoes the presence of charged particles matched to the supercluster.

Figure 4 and performance plots from Ref. [13] demonstrate the impressive quality of the fast simulation of the response to low-energy photons by means of the Particle Flow algorithm. The $\pi^0$ resonance is reconstructed in the mass distribution of pairs of photons in $|\eta| < 1$, of energy above 0.4 GeV, whose summed energy is larger than 1.5 GeV. Peak position and resolution in FastSim are consistent with both data and full simulation [14].

1 Only within the Tracker volume: decays outside the Tracker outer radius are not considered, and therefore no punch-through can be simulated so far in the Fast Simulation
Figure 3. Comparison between data and fast simulation at 7 TeV for the transverse momentum (a), the pseudorapidity (b), the azimuthal angle (c), the transverse (d) and longitudinal (e) impact parameter with respect to the leading primary vertex, for tracks passing the quality criteria of Ref. [11].

4.3. Muon Reconstruction and Performance

Several categories of muon candidates are reconstructed in CMS [15], among them:

- **Tracker Muons**: tracks reconstructed in the inner tracker, matched to at least one segment in the muon chambers.
- **Global Muons**: combined fit of hits in the inner tracker and in the muon chambers.

Typical high $P_T$ analyses, for example in Higgs or SUSY searches or in the studies of the properties of W, Z and top quark, require a tighter selection. Hereafter we define a prototypical **Tight Muon** selection as follows:

- the candidate belongs to both the Tracker Muon and Global Muon categories;
- $P_T > 3$ GeV/c;
- matched to at least two segments in the muon chambers;
- normalized $\chi^2$ of the Global Muon less than 10;
- at least one muon chamber hit included in the global fit;
- at least ten tracker hits included in the global fit, out of which at least one in the pixel system;
- transverse impact parameter < 2 mm with respect to the leading primary vertex.

Transverse momentum, pseudorapidity, transverse and longitudinal impact parameters are shown in Figure 5 for Tight Muons with 7 TeV data compared to fast simulation.
4.4. Jet Reconstruction and Performance

Three main jet type reconstructions are performed in CMS [9], according to the kind of input objects to be clusterized: calorimetric jets (“CaloJets”) are built from the clustering of calorimetric cells; track-corrected-jets (or jets-plus-tracks, “JPTJets”) complement the calorimetric information with a-posteriori corrections based on the associated tracks; and particle-flow jets (“PFJets”) are built from particles identified by the Particle Flow algorithm. Several clustering algorithms and resolution parameters are available for all these jet types.

The electromagnetic and hadronic components of CaloJets are also well simulated. The PFJets composition is also well reproduced for what concerns the charged and the neutral electromagnetic components.

4.5. Missing Energy Reconstruction and Performance

Missing Energy Reconstruction can be performed quite well by Fast Simulation as shown in Figure 6 and Figure 7. More plots are available in [16].
4.6. L1 and HLT Integration

The CMS trigger system is organized in two steps [17]. The first step (Level-1 Trigger, or “L1”) is designed to reduce the rate of events accepted for further processing to less than 100 kHz. The second step (High-Level Trigger, or “HLT”) is designed to reduce this rate to approximately 100 Hz. The HLT is internally organized into L2 and L3.

In the fast simulation, L1 trigger primitives are built in the ECAL, HCAL and Muon systems starting from the detector hits, and are used to generate the L1 decision functions as for full simulation and real data. Those L1 primitives serve then as seeds for the subsequent L2/L3 objects, which build up the HLT decision functions. Although there is some customized, FastSim-specific reconstruction of a few L3 objects, the HLT uses the same menus and algorithms as in the real data chain.

5. Conclusions

The distributions obtained with the Fast Simulation of minimum bias events have been compared with the real collision data collected by CMS. The CMS Fast Simulation has proven to be a very useful tool for analysis and algorithm design in CMS. Due to its efficient processing and ease of use, it will also be an invaluable tool for LHC physics analysis, either for estimating the contributions of large cross-section backgrounds or in the generation of elaborate sets of physics signals.

Acknowledgements

Andrea Giammanco wishes to thank the Estonian Academy of Science for supporting this work with the Mobilitas Top Researcher Grant MTT59.
Figure 6. First row: MET in the calorimetric (a), track-corrected (b), and particle flow (c) definitions, in data, fast and full simulation. Second row: MET component along the $x$ or $y$ axes. Third row: $\sum E_T$.

[1] CMS Collaboration, “CMS technical design report, volume I: Detector performance and software”, Technical Report CERN-LHCC-2006-001, CERN, (2006).
[2] CMS Collaboration, “The CMS experiment at the CERN”, JINST 03 (2008) S08004.
[3] GEANT4 Collaboration, “GEANT4: A Simulation”, Nucl.Instrum.Meth. A506 263 (2003)
[4] Particle Data Group, “Review of Particle Physics”, J. Phys. G37 (2010) 075021.
[5] G. Grindhammer, M. Rudowicz, and S. Peters, “THE FAST SIMULATION OF ELECTROMAGNETIC AND HADRONIC”, Nucl. Instrum. Meth. A29 (1990) 469.
[6] G. Grindhammer and S. Peters, “The parameterized simulation of electromagnetic showers in homogeneous and sampling”, (1993). arXiv:hep-ex/0001020.
[7] R. Rahmat, R. Kroeger, “HF GFlash”, Physics Procedia, Volume 37, pp. 340-346 (2012)
[8] J. Allison, K. Amako, J. Apostolakis et.al “Geant4 developments and applications”,IEEE Trans.Nucl.Sci. 53 (2006) 270.
[9] CMS Collaboration, “CMS Jet Performance in pp Collisions at $\sqrt{s} = 7$ TeV”, CMS Physics Analysis Summary CMS-PAS-JME-10-003 (2010).
[10] UA1 Collaboration, “Observation of Jets in High Transverse Energy Events at the CERN Proton - antiproton”, Phys.Lett. B123 (1983) 115.
[11] CMS Collaboration, “CMS Tracking Performance Results from early LHC”, Eur. Phys. J. C70 (2010)
[12] CMS Collaboration, “Electron reconstruction and identification at $\sqrt{s} = 7$ TeV”, CMS Physics Analysis Summary CMS-PAS-EGM-10-004(2010)
[13] CMS Collaboration, “Photon reconstruction and identification at $\sqrt{s} = 7$ TeV”, CMS Physics Analysis Summary CMS-PAS-EGM-10-005 (2010)
[14] CMS Collaboration, “Commissioning of the Particle-Flow Reconstruction in Minimum-Bias and Jet Events from pp Collision at $\sqrt{s} = 7$ TeV”, CMS Physics Analysis Summary CMS-PAS-PFT-10-002(2010)
[15] CMS Collaboration, “Performance of muon identification in pp Collision at $\sqrt{s} = 7$ TeV”, CMS Physics Analysis Summary CMS-PAS-MUO-10-002(2010)
[16] CMS Collaboration, “Missing Transverse Energy Performance in Minimum-Bias and Jet Events from Proton-Proton Collisions at $\sqrt{s} = 7$ TeV”, CMS Physics Analysis Summary CMS-PAS-JME-10-004 (2010).
Figure 7. MET resolution versus $\sum E_T$ in data, fast and full simulation.

[17] CMS Collaboration, “CMS. The TriDAS project. Technical design report, vol. 1: The trigger systems”, Technical Report CERN-LHCC-2000-038, CERN, (2000).