An Object-Oriented Simulation Program for CMS

S. Abdoulline, *FNAL, Chicago, USA*, N. Amapane, *University of Torino, Italy*,
V. Andreev, *UCLA, USA*, P. Arce, *CIEMAT, Madrid, Spain*,
S. Banerjee, *Tata Institute, Bombay, India*, T. Boccali, *SNS, Pisa, Italy*,
M. Case, *UC Davis, USA*, A. De Roeck, *CERN, Geneva, Switzerland*,
S. Dutta, *University of Pisa, Italy*, G. Eulisse, *NorthEastern University, Boston, USA*,
D. Elvira, *FNAL, Chicago, USA*, A. Fanfani, *University of Bologna, Italy*,
F. Ferro, *INFN, Genova, Italy*, M. Liendl, *CERN, Geneva, Switzerland*,
S. Muzaffar, *NorthEastern University, Boston, USA*, A. Nikitenko, *Imperial College, London, UK*,
K. Lassila-Perini, *HIP, Helsinki, Finland*, I. Osborne, *NorthEastern University, Boston, USA*,
M. Stavrianakou, *FNAL, Chicago, USA*, T. Todorov, *IReS, Strasbourg, France*,
L. Tuura, *NorthEastern University, Boston, USA*,
H.P. Wellisch, *CERN, Geneva, Switzerland*, J. Weng, *Karlsruhe University, Germany*,
T. Wildish, *Princeton University, USA*, S. Wynhoff, *Princeton University, USA*,
M. Zanetti, *INFN, Italy*, A. Zhokin, *ITEP, Moscow, Russia*,
P. Zych, *Warsaw University, Poland*

Abstract

The CMS simulation package, OSCAR, is based on the Geant4 simulation toolkit and the CMS object-oriented framework for simulation and reconstruction. Geant4 provides a rich set of physics processes describing in detail electro-magnetic and hadronic interactions. It also provides the tools for the implementation of the full CMS detector geometry and the interfaces required for recovering information from the particle tracking in the detectors. This functionality is interfaced to the CMS framework, which, via its "action on demand" mechanisms, allows the user to selectively load desired modules and to configure and tune the final application. The complete CMS detector is rather complex with more than 1 million geometrical volumes. OSCAR has been validated by comparing its results with test beam data and with results from simulation with a Geant3-based program. It has been succesfully deployed in the 2004 data challenge for CMS, where more than 35 million events for various LHC physics channels were simulated and analysed.

INTRODUCTION

OSCAR [1] manages all CMS detectors, both central (Tracker, Calorimeters and Muons), Fig. 1 and forward (CASTOR calorimeter, Totem telescopes, and the Zero Degree Calorimeter, ZDC), as well as several test beam layouts and prototypes. It implements their sensitive detector behaviour, track selection mechanisms, hit collections and numbering schemes.

OSCAR is based on the Geant4 [2] toolkit and the CMS object-oriented framework COBRA [3], in particular the COBRA Mantis simulation specialization package [4].

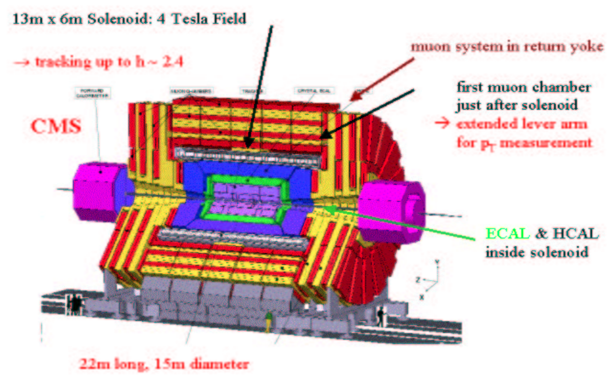

Figure 1: The CMS detector

The standard OSCAR application with the central systems in the CMS 4 Tesla magnetic field, and physics processes simulated with the Geant4 physics list, has been extensively validated in terms of physics and is regularly tested following the evolution of Geant4 and Mantis.

OVERVIEW, INTERFACES AND SERVICES

OSCAR application steering, persistency, histogramming, monitoring and other support services are handled by the COBRA/Mantis infrastructure.

Detector geometry construction is automated via Mantis interfaces that utilize the CMS Detector Description Database (DDD) [5] to convert the input XML files managed by the CMS Geometry project [5].

Generator input is handled entirely by COBRA/Mantis; the specific generator to be deployed and the event are specified at run-time with OSCAR datacards.

Magnetic field services are provided by COBRA/Mantis and the COBRA MagneticField package. The choice of field and the propagation parameters are specified at run-time with OSCAR datacards and DDD/XML files.

The new field design, for accessing a field map calculated with TOSCA, is based on a dedicated geometry of "magnetic volumes". Volume finding and interpolation within a volume are decoupled, facilitating optimization for simulation and reconstruction. The time spent in magnetic field query is 1.8-2 times less than with the FORTRAN/Geant3 implementation. Further improvement is expected by the mapping of Geant4 volumes to their corresponding magnetic ones and the deployment of configurable local field managers, yielding an overall performance improvement of the order of 5%, 10% being the time spent in propagation in the field with respect to the total event processing time.

The Mantis infrastructure for physics lists and production cuts is used to implement several kinds of physics lists with extensive run-time configurability: choice of physics subset (propagation only for debugging and material studies, standalone electromagnetic etc), choice of hadronic physics (LHEP, QGSP, QGSC and FTFP), activation of synchrotron radiation and gamma/e-nuclear process and customized electron bremsstrahlung treatment with new electron track creation for hard bremsstrahlung with configurable threshold, as well as the use of parameterized electromagnetic shower simulation with the G4Flash package.

Hit processing and collection utilize the COBRA/Mantis sensitive detector infrastructure and the COBRA Profound classes that are common for and sharable by all CMS applications. COBRA read-out factories handle the hit formatting as required for the subsequent digitization by the CMS ORCA [5] reconstruction program.

Several general purpose patterns and services such as the COBRA dispatcher-observer pattern and the DDD maps are used extensively to access, monitor and process various simulation observables as needed by the specific applications.

Visualization and event display are adderssed by OSCAR and Mantis extensions of the IGUANACMS framework [6], Fig. 2.

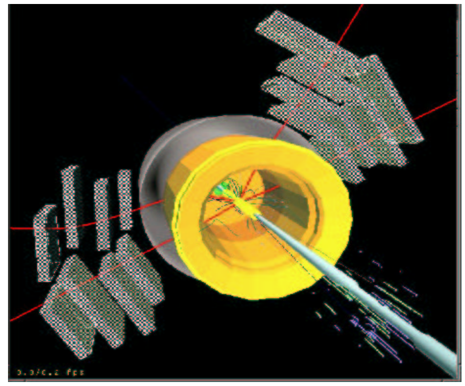

Figure 2: Higgs event displayed by IGUANACMS

OSCAR DETECTOR AND PHYSICS STUDIES

A detailed account of the CMS detector implementation in OSCAR and the very extensive physics validation studies performed and foreseen is outside the scope of this summary paper. It should be noted however that OSCAR has been validated and endorsed by the CMS tracker, electromagnetic (ECAL) and hadronic (HCAL) calorimeters, and the muon system, and is also adopted and used by the CMS forward detectors (the CASTOR calorimeter, the TOTEM telescopes, and the ZDC) which will play a key role in diffractive and heavy ion physics. A few notes and a tiny subset of the available results will have to suffice as token illustrations of the OSCAR physics scope and potential. Although not always explicitly mentioned, Geant4 simulation results are compared with test beam data, wherever available, and the corresponding CMSIM/Geant3 figures.

Tracker simulation has played a key role in the development and optimization of the simulation infrastructure and the validation process. The tracker material budget, which can only be correctly estimated with very detailed description of all active and passive detector components, directly affects the electromagnetic calorimeter physics performance and places stringent requirements on the accuracy of the detector description and geometry construction.

Correct, navigable Monte Carlo truth, for correct decay tree reconstruction, as well as the proper treatment of hard electron bremmsstrahlung are of vital importance in B-τ studies, in which the tracker plays a key role.

With the above requirements satisfied, tracker performance has been extensively validated in terms of tracking and hit distributions for single particles, minimum bias and physics events.

ECAL studies, comparing Geant4 simulation with CMSIM/Geant3 and test beam data have so far focused on energy, position resolution and shower studies, hadronic shower behaviour, preshower response, and level-1 electromagnetic trigger response.

They have allowed meaningful tuning of physics range cuts (bremsstrahlung recovery in just one inactive tracker component has been seen to affect energy resolution by as much as 10%) and contributed to the debugging and optimization of physics processes.

HCAL studies (energy resolution, shower profile, linearity and e/π ratio), especially in comparison with test beam data, stand-alone or in combined ECAL-HCAL runs, and in the context of the LCG simulation physics validation project, have been instrumental in hadronic physics validation, Fig. 3 and Fig. 4.

The 2002 HCAL test beam detector, with a total of 144 hadronic barrel and 16 outer hadronic channels in a combined setup with an electromagnetic calorimeter prototype, was exposed to beams of pions, electrons and muons over a large energy range. More than 100 M events were collected. The data were compared with Geant4 simulations using the hadronic physics parametric (LHEP) and microscopic (QGSP) models.

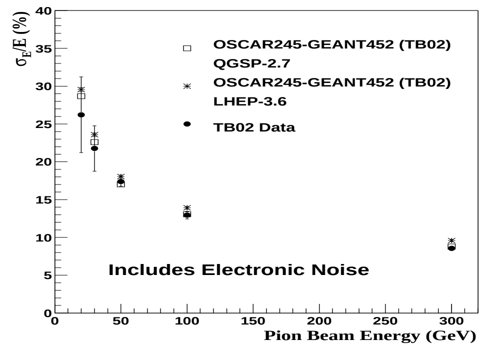

Figure 3: Energy resolution in the ECAL-HCAL test beam

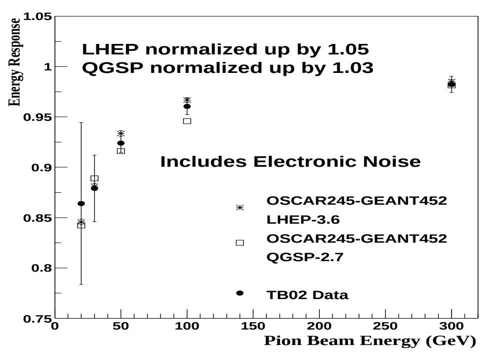

Figure 4: Energy response in the ECAL-HCAL test beam

The test beam comparisons take into account the effects of electronic noise and systematics. It is interesting to note that the electronic noise masks most of the non-Gaussian nature of the energy distributions for low energy pions. Its contribution to energy resolution measurements becomes increasingly relevant as the pion beam energy decreases, but it does not significantly affect linearity measurements. The size of systematic uncertainties associated with particle discoveries, mass, and cross section measurements is tightly associated with how accurately the simulations describe the actual performance of the detector in measuring physics objects.

The pion energy resolution and response linearity as a function of incident energy derived from the simulations are in good agreement with the data measurement within the large systematics uncertainties in the latter. Below 30 GeV, the uncertainties in the data are too large to provide information about deviations of the Monte Carlo model with respect to the data measurements.

Transverse and longitudinal shower profiles will be be compared with data from the 2004 test beam. These factors affect the calculations of calibration constants, total energy, missing transverse energy and hadronic shower punch-through. Pion showers predicted by Geant4 are narrower than those predicted by Geant3. Showers predicted by the QGSP physics list (version 2.7) are shorter than those predicted by the LHEP (version 3.6) list, with LHEP predictions being closer to those from Geant3/Gheisha. Comparisons with data will be critical for tuning and validating transverse shower profiles in Geant4.

The studies are continuing with particular emphasis on the simulation of shower development and the comparison of the Geant4 hadronic physics parametric (LHEP) or microscopic (QGSP) models.

Last but not least, the muon system detailed detector and physics validation in terms of tracking and hit distributions has been a sustained activity following the Geant4 tracking and muon physics evolution and driving the tuning of particle propagation in the magnetic field. The validation involves studies with single muon samples covering a large p_T spectrum, 2 TeV Drell-Yan pairs and physics events, most notably $H \to ZZ \to 4\ \mu$.

Forward detector simulation

As already mentioned, OSCAR is used for the CMS forward detectors CASTOR calorimeter, Totem telescopes and the ZDC.

The forward physics study program includes soft and hard diffraction studies, two-photon physics, cosmic ray studies and luminosity measurements. The combination of central and forward detectors allows measurements of hard diffraction processes to study the pomeron structure and the dynamics of diffraction.

The TOTEM system is designed to measure the total and elastic *pp* cross sections, and the diffraction dissociation. It consists of two telescopes, T1 and T2, to detect inelastic events and (to be added to the simulated setup) three Roman Pot stations, at 147, 180 and 215 m from the interaction point to measure protons scattered under very small angles.

The CASTOR calorimeter is designed to cover the same region as T2.

OSCAR simulations with minimum bias, diffractive and beam-gas events have been used for the TOTEM TDR [8] studies. CASTOR simulation with OSCAR is used for Cherenkov yield, energy resolution and shower profile studies and comparisons with the 2004 test beam data.

Parameterized electromagnetic shower simulation

The spatial energy distribution of electromagnetic showers is parameterized by three probability density functions [9]. This model has been used in the CMS G4Flash parameterized shower simulation for electrons and positrons.

Tuning is in progress; first comparisons with the full OSCAR physics simulation show good agreement (at the level of 1%) for the energy deposit, moderate (a few percent level) agreement for the longitudinal profile, getting worse at higher energies (above 200 GeV), and non-negligible differences (greater than 10%) in the radial profiles.

CPU comparisons between the full and parameterized simulation for a single electron indicate performance differences ranging from factor 2 to 3 at 1 to 10 GeV to factor 20 to 60 at 100 to 300 GeV.

OSCAR PERFORMANCE

In the current version OSCAR performance compares to that of the Geant3-based CMSIM simulation program. In terms of CPU, OSCAR event processing is within 50% of that for CMSIM. This is partly accounted for by the more conservative (lower) production cuts, most notably in the tracker system, which will be optimized for performance at a later stage, partly by the current Geant4 implementations of electromagnetic physics and propagation in field, which are in the process of being profiled and optimized, already with significant improvements with respect to earlier versions, and last but not least by CMS specific implementations both in the Mantis simulation framework and OSCAR itself.

Improvements in the implementation of calorimeter hit processing and track selection have resulted in dramatic performance improvements for heavy ion simulation.

The generator-to-Geant4 interface and the Monte-Carlo truth processing in Mantis, may also be further optimizable, both in terms of logic and technical implementation.

OSCAR performance in terms of memory, for standard *pp*-physics events at least, is already quite satisfactory. The typical footprint of 110 MB (resident set) per *pp*-physics event compares well to the 100 MB footprint of CMSIM/Geant3.

This is not yet the case for heavy ion events, where CPU performance and memory optimization can be coupled and further implementation improvements are being evaluated, aiming at bringing the current range of 0.6 to 2 GB for full event simulation to below 0.5 GB, the CMSIM/Geant3 footprint for segmented-event processing.

Event reproducibility, which lends credence to simulation, is also indispensable for debugging of very rare crashes and understanding the effect of subtle changes in physics and tracking. It has been achieved after in-depth memory debugging for both OSCAR/Mantis and Geant4.

Physics initialization on a P4 2.8 GHz takes 20-30 seconds with cross-section retrieval from file, 100-120 seconds without, for a setup with 317 materials and a full production physics list. In terms of memory, the cross-section tables require about 55 MB. These figures represent factor 2 improvement in Geant4 from release 5.2.p02 to 6.2.p01.

In terms of robustness, experience has been very positive, both for OSCAR and Geant4. A recent OSCAR public release with Geant4 6.2.p01 has successfully passed a stress-test with 800 K single particle and 300 K full QCD events. A small rate of 1 /10000 crashes in *pp* events was observed in the DC04 production with earlier versions of OSCAR and Geant4 (5.2.p02).

CPU and memory performance, robustness and reproducibility tests, as well as regression tests using the OVAL tool suite, are integrated in the OSCAR release procedure. The former are done for minimum bias and $H \rightarrow ZZ \rightarrow$ 2 e 2 μ events, include testing of the storage and retrieval of physics tables, and OSCAR operation in production mode. Major new features are incorporated in the test suite as required.

Several customized detector applications for the various calorimeter test beams and the forward systems are provided for further testing by the relevant detector and physics groups.

SUMMARY AND CONCLUSIONS

In CMS, OSCAR, the OO simulation program based on the Geant4 toolkit, has successfully replaced its FORTRAN/Geant3 predecessor. It has been validated and adopted by all CMS detector and physics groups. It has proven robust and performant, easily extensible and configurable.

OSCAR 2.4.5 based on Geant4 5.2.p02 has been in production since November 2003; for over 10 months at the time of writing. It has been the longest-used version of any software package in production and accounts for 35 M of 85 M physics events.

As of September 2004, CMS has entered sustained-mode production for 10 M physics events/month pushed through the full chain (simulation, digitization, reconstruction and DST production). OSCAR is expected to meet this challenge and provide the data for the CMS physics TDR.

REFERENCES

[1] OSCAR, Object oriented Simulation for CMS Analysis and Reconstruction, http://cmsdoc.cern.ch/oscar/

[2] S. Agostinelli et al., "Geant4: a simulation toolkit", NIM A 506 (2003), 250-303
Geant4, http://wwwinfo.cern.ch/asd/geant4/geant4.html

[3] V. Innocente, et al., "CMS Software Architecture: Software framework, services and persistency in high level trigger, reconstruction and analysis", Computer Physics Communications 140 (2001) 31-44

[4] M. Stavrianakou et al.,"Mantis: the Geant4-based simulation specialization of the CMS COBRA framework", CHEP'04, Interlaken, Switzerland, September 2004

[5] CMS Object-Oriented projects, http://cmsdoc.cern.ch/cmsoo/cmsoo.html

[6] V. Innocente, G. Eulisse, S. Muzaffar, I. Osborne, L. Taylor, L.A. Tuura, "Composite Framework for CMS Applications", CHEP'04, Interlaken, Switzerland, September 2004

[7] Validation of Geant4 Physics Using the CMS HCAL Test Beam 2002, D. Elvira, CMS Note 2005/xxx *(to be published)*

[8] TOTEM TDR, CERN-LHCC-2004-002

[9] hep-ex/0001020, G. Grindhammer and S. Peters