Detector Simulation Overview

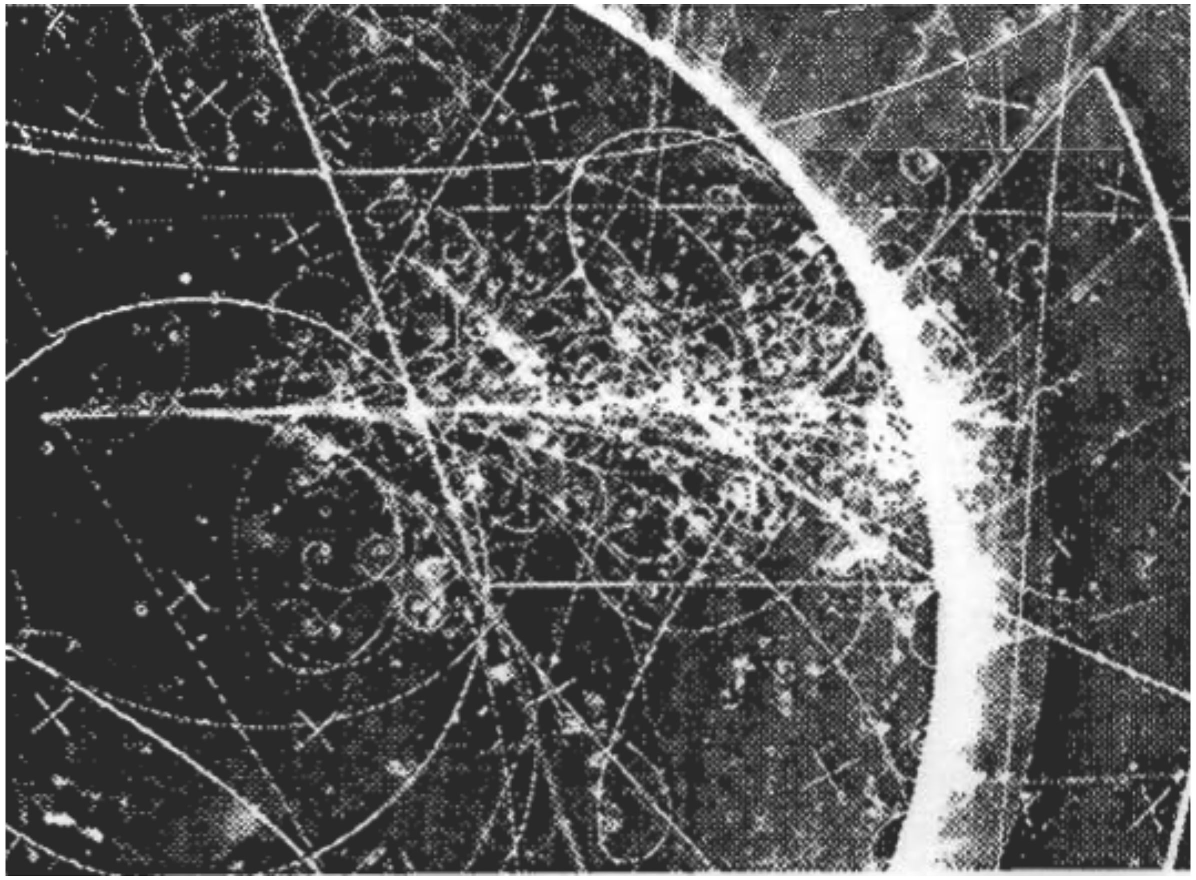
- Some History
- A Toolkit
- An Application
- Future

Bengaluru
November, 2022

Sunanda Banerjee

Introduction (I)

- High energy physics experiments during 60's and early 70's relied on visual detection techniques
  - Nuclear emulsions, spark chambers, bubble chambers, …



  - Natural intelligence was sufficient to identify particle trajectories
  - Relied on computers to take care of some numerical computation for the measurement of particle momenta, etc.
- This procedure was good enough for an understanding of not-so-rare processes which are produced abundantly in the experiments

Introduction (II)

- As knowledge and understanding in this field grew there were needs
  - to go at higher energy experiments through collisions of moving particles
    - proton-proton (ISR), electron-positron (ADONE), ..
  - to look for rarer and rarer processes
    - cross sections in mb→μb→nb→pb→fb

- Detection techniques changed with newer and newer technologies
  - multi-wire proportional chambers, drift chambers, Cerenkov detectors, crystal calorimeters, transition radiation detectors, ….
  - complex detector systems are built by combining detectors of different technologies

- To combat this, experiments are no longer done by a group consisting of human beings in single digits. It slowly grew to 2-digits, 3-digits, 4-digits
  - Even then natural intelligence was not sufficient to understand what the new detector systems provide
  - Model detectors in the computer and try to simulate what particles from an interaction will give signals in this detector system

Detector Simulation

Introduction (III)

- The first set of serious detector simulation code tried to simulate showers created by electrons, photons and then hadrons, nuclei → Transport code
  - EGS (Electron Gamma Shower) code was developed at SLAC during late 70's and early 80's
  - Neutron transport code existed even earlier (thanks to the building of nuclear reactors).
    - Extension of that to higher energy came through the HETC code (High Energy Transport Code) at Oakridge roughly at the same time
    - Also, a parametric approach, GHEISHA, was developed to satisfy experiments at PETRA

- Characteristics of these codes:
  - detector physics was well described through the current understanding of strong, electromagnetic and weak interactions
  - propagation of particles through various components of the detector was left to the ingenuity of the experimenters who were using these codes

- CERN was considering building the LEP at that time. Some of the CERN scientists thought of launching a software package which handles both detector physics and propagation of particles in the presence of an electromagnetic field

Geant3

Introduction (IV)

- During the '90s, experimental high energy physicists realised that they were using software tools which were outdated in the industry and future extensions of using that computing language (Fortran) would be difficult

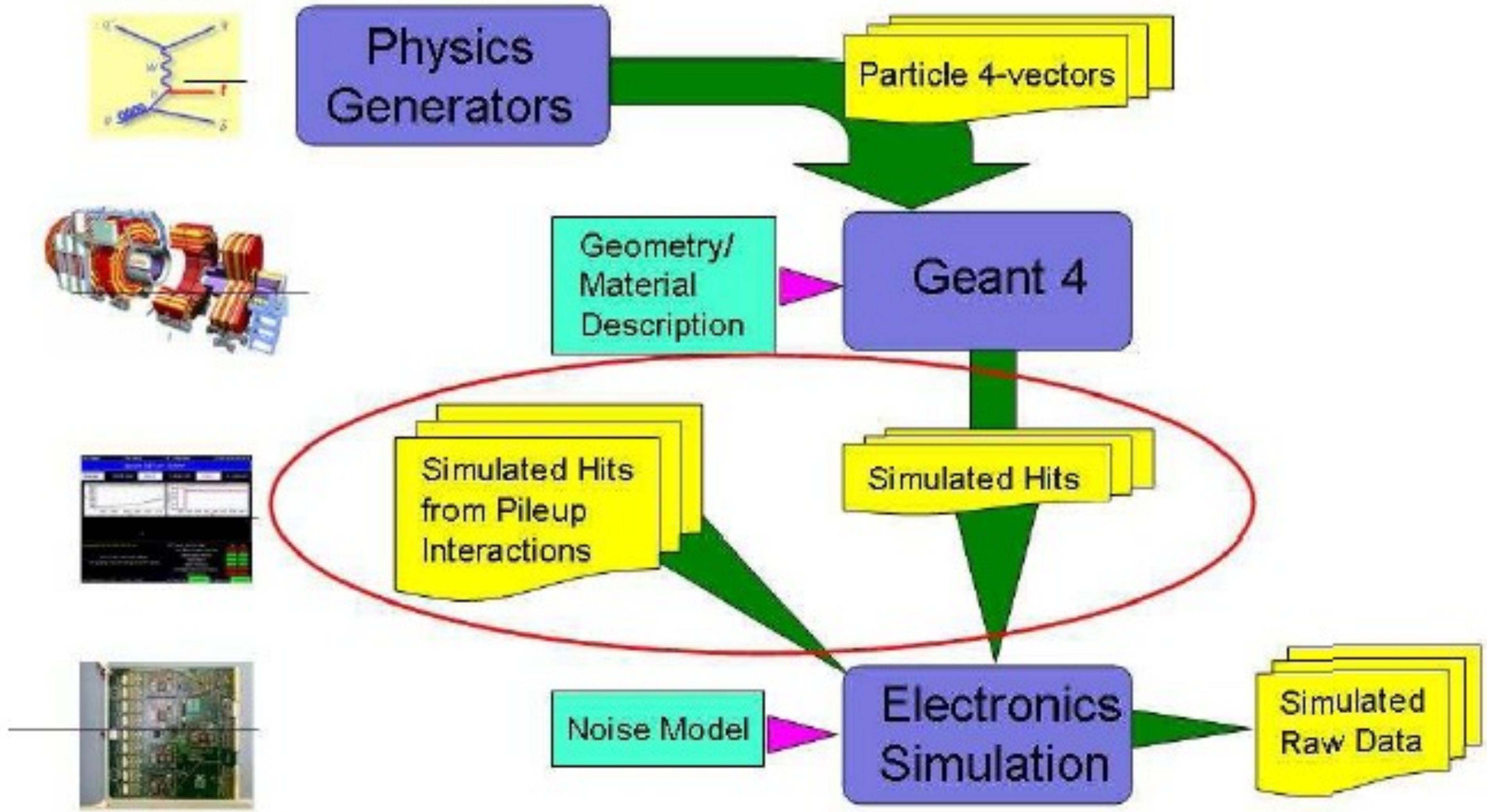
- Also, 4 major experiments were proposed to be built at the Large Hadron Collider (LHC) facility with a plan for using them for the next ~40-50 years
  - ATLAS, CMS, ALICE and LHC-b

- A new collaboration was formed jointly between CERN and KEK to design and implement simulation software in the newly adopted language (C++) and technology (Object Oriented programming)

geant4

- Three more software toolkits are also used in this field:
  - MARS: used for accelerator and beam-line designs
  - FLUKA: used for accelerators, beam-line designs as well as radiation tolerance studies of detectors
  - MCNP(X): one of the oldest codes and widely used in nuclear physics applications

Understanding Detector Performance



- Employ widely in physics analysis:
  - estimate various efficiencies, (ir)reducible backgrounds, determine systematic uncertainties,
- Also used to decide how the detector should be built

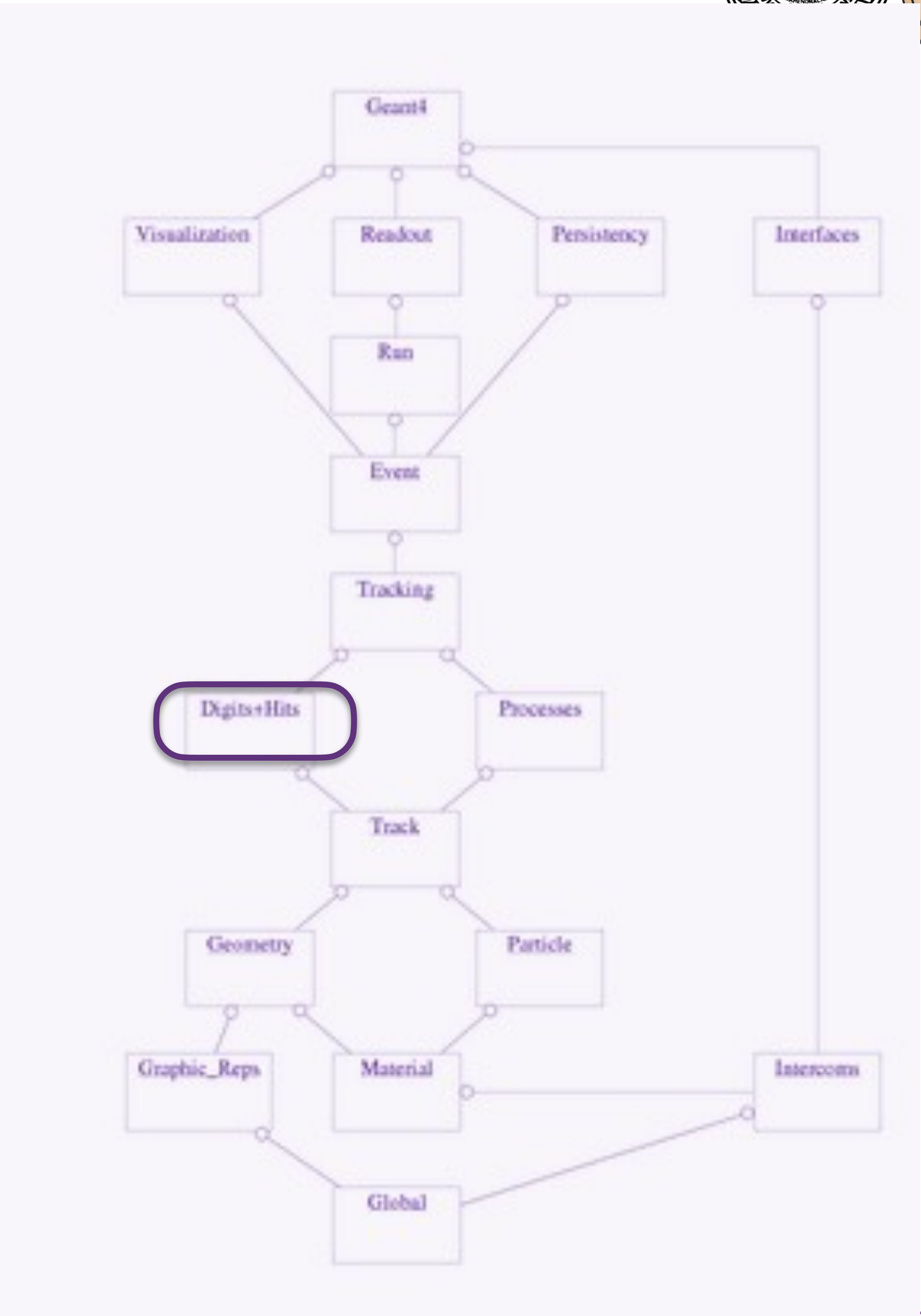
- The reliability of MC prediction depends on
  - the goodness of Physics models of the event generator
  - realistic description of the detector and quality of models used in evaluating propagation of particles through the detector

Geant4 (I)

- A mini-workshop was held in August 1993 with the participation of scientists from KEK and CERN to look into the possibility of making an Object-Oriented version of Geant
- A proposal, P58, was submitted to the CERN DRDC with the participation of 29 persons from 19 institutes in 9 countries
- The proposal was approved and the Geant4 collaboration was formed as a CERN project, RD44
- The project adopted the standard steps of an object-oriented program:
  - user requirement documents got benefits from the experience of Geant3 and other earlier simulation programs. The usage was extended to applications in space science and medical physics
  - major objects in the problem domain were identified using UML (Unified Modelling Language adopted from Booch methodology)
  - the central model was produced for each object leading to object diagrams and scenarios
  - closely coupled classes were clustered together to form class categories
  - the basic design of the Geant4 kernel was created by delegating detailed design works to each category group
  - the design work was completed by 1996 and the building of a prototype started

Geant4 (II)

- There is a total of 17 categories and the dependencies are one-sided
- Cyclic dependencies are avoided
- The category "Global" is at the bottom and "Geant4" at the top of the hierarchy
- There were intense discussions among group members in each of these categories to come out with a complete design



Geant4 (III)

- The alpha version of Geant4 was released in April 1997 which satisfy
  - comparable functionality to Geant3 for geometry, tracking, electromagnetic physics, high/medium energy hadronic physics
  - easily exchangeable physics models to satisfy the experimental condition
  - object persistency through ODBMS as advertised in a separate R&D project (RD45) of CERN

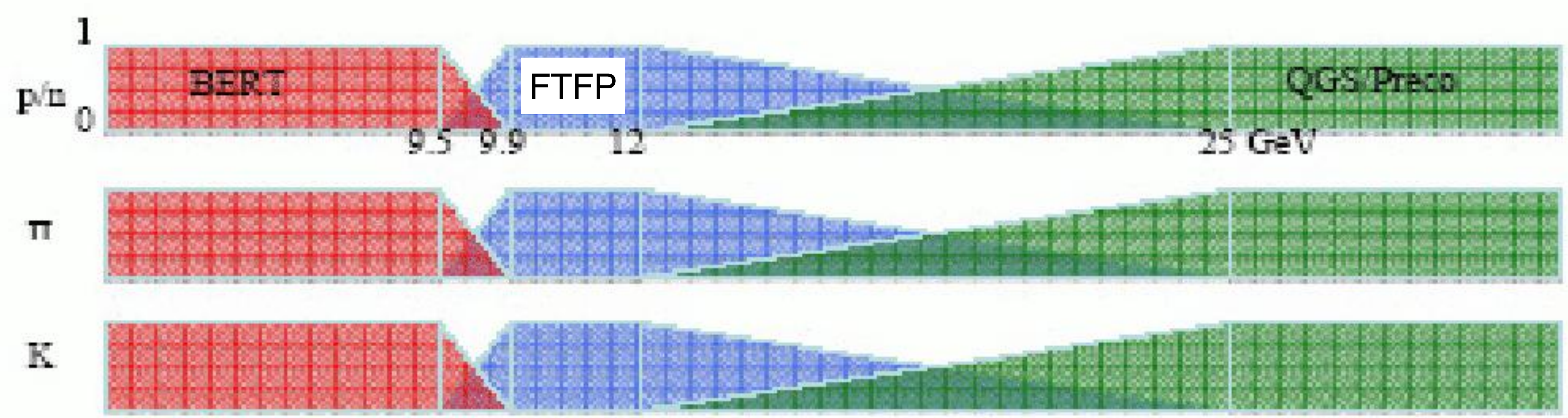
- The beta version of the code was released in July 1998 and major experiments were invited to test the code through an application software specific to that experiment. It consisted of
  - completed system with documentation, examples and tutorials
  - extended range of physics validation and in particular low energy neutron transfer for radiation studies
  - extension of geometry code to directly couple with CAD models

- This signalled the end of the R&D phase. The project RD44 was closed and the first production release of Geant4 was available in December 1998

Geant4 (IV)

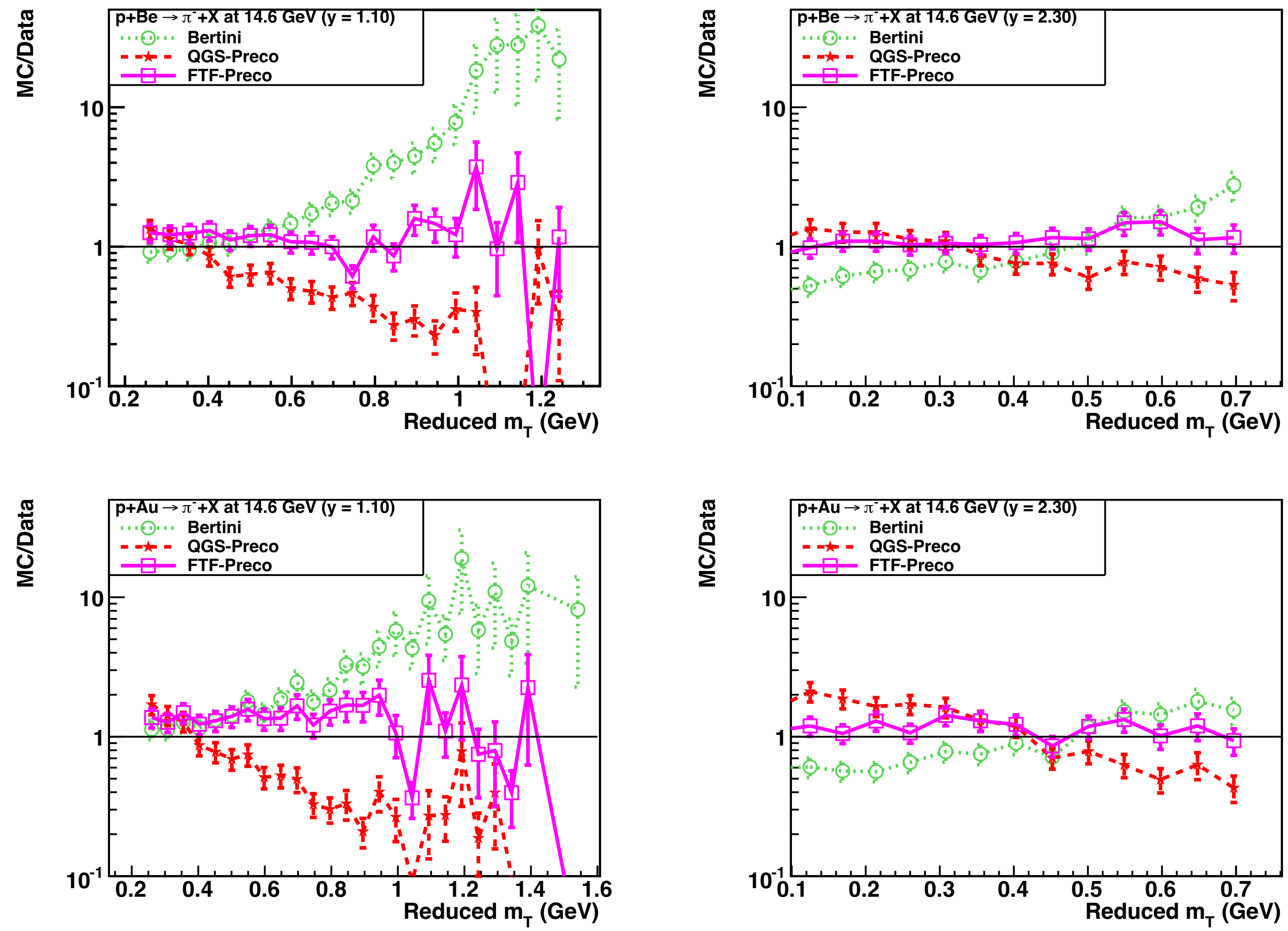
- Electromagnetic physics has been described in great detail and validated using predictions from the EGS4 shower code which is used as a standard in all simulation toolkits. At very low energies (below 10 keV) and more particularly for optical photons, the EM physics models need to be extended and introduced from the first principle.

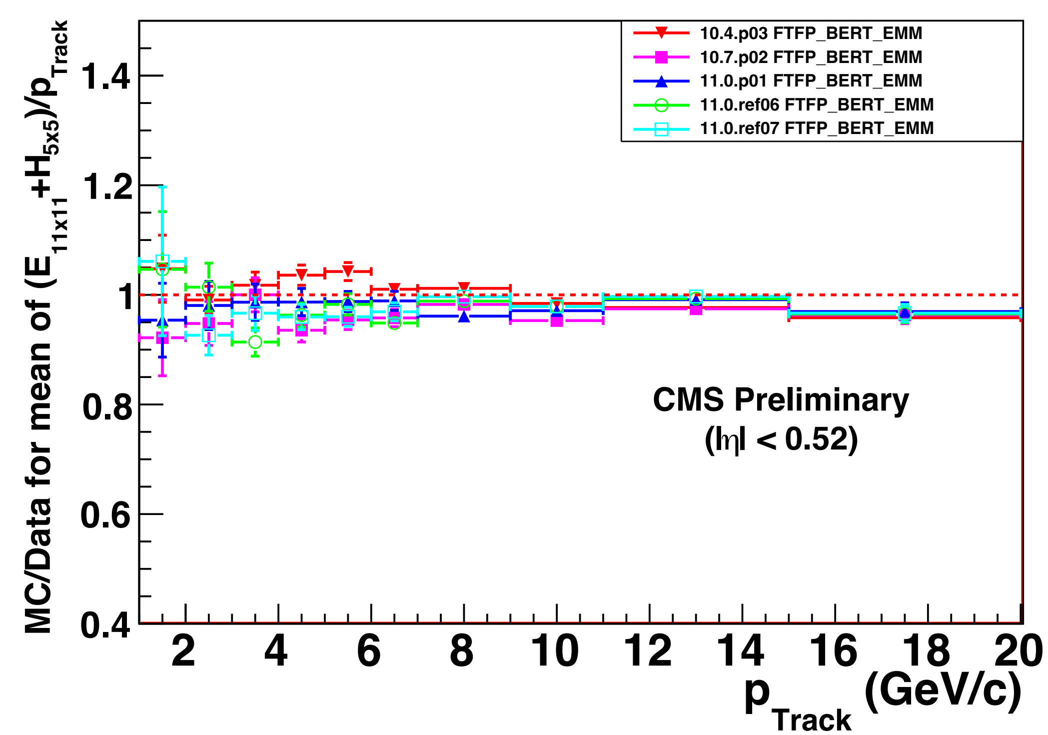
- The story is different for hadronic physics. No single model has been found which could explain hadronic physics from low energies (~ 10 keV) to the highest energy (~100 TeV) demanded by HEP experiments of today. The Geant4 team has taken the approach of using several models which provide predictions close to the data in different energy domains and combine them in one list, called "physics list" allowing a smooth transition scheme.



Geant4 (V)

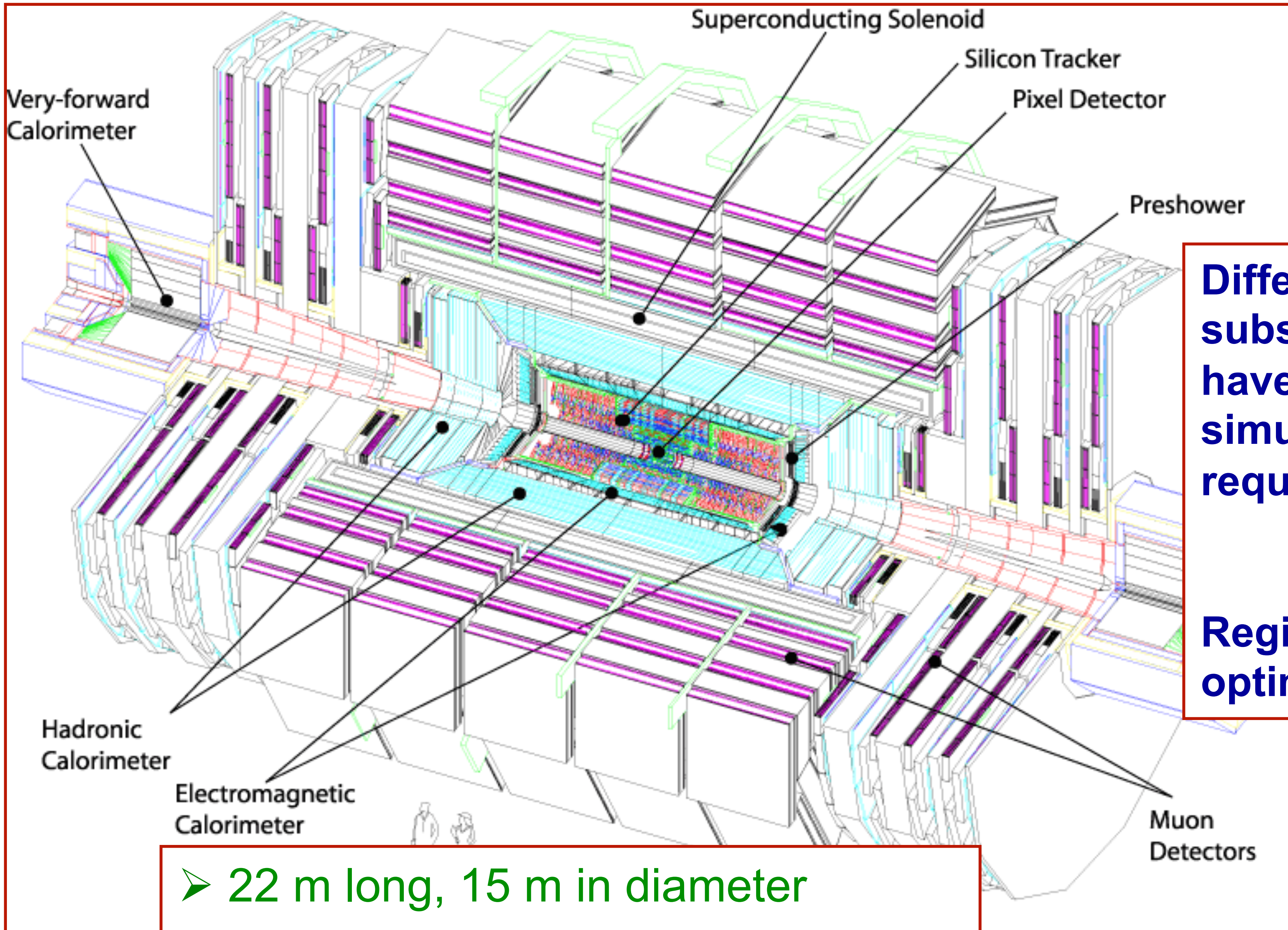
- It is extremely important to test the goodness of the model predictions by comparing them against data. Two types of comparisons are done:
  - against thin-target data: Several experiments have published data after making all detector-level corrections. Comparisons against those data do not require the predictions of the models to be dressed with the effect of other processes
  - against calorimeter data: These data are experiment specific and thus the experiments take the responsibility of carrying out the validation effort





Application to an Experiment

- All experiments at the LHC use detector simulation in detector design as well as in data analysis
- Concentrate on one experiment (CMS) here



Different subsystems have different simulation requirements

⇓

Region based optimization

- ➢ 22 m long, 15 m in diameter
- ➢ Over a million geometrical volumes
- ➢ Many complex shapes

CMS Detector Simulation (I)

- CMS formally started a collaboration in the early 1990s with the Evian meeting. The management at that time knew the importance of simulation and they started the effort immediately led by a Finnish group

- Soon other groups joined this activity and made the first version of a tool "CMSIM I" based on Geant3 (Fortran77 and ZEBRA-based tool developed for LEP). Some of the early collaborators came from ALEPH and they chose a memory management package BOSS

- When some US groups joined this effort, a fight began — about which memory management package to be used. At that time the software management delegated the decision to a task force and in a meeting at the University of California, Davis it was decided to use a single memory management package (ZEBRA) for the simulation of the CMS detector —> CMSIM

CMS Detector Simulation (II)

- CMSIM started as a Geant3-based software written in Fortran77. It relied on several detector experts: at least one from each major detector component: the Tracker, the electromagnetic calorimeter (ECAL), the hadron calorimeter (HCAL), the Muon detector system), and other components of the infrastructure (the beam pipe, the magnet, shieldings, forward detectors etc)

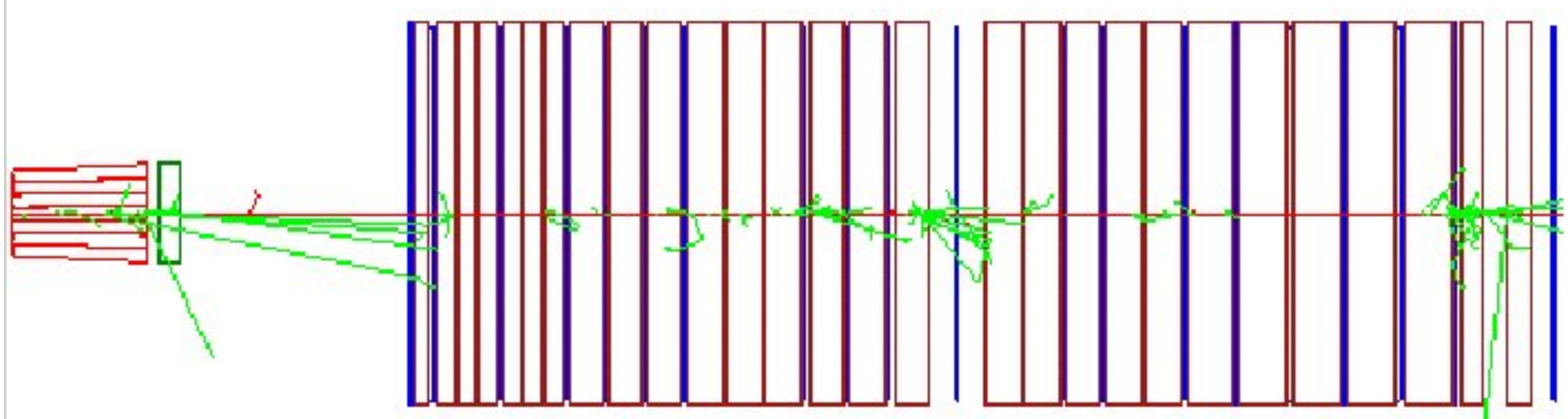
- CMSIM was a tool not just for detector simulation, but also for reconstruction. With the advent of C++, a part of the reconstruction code (for ECAL) was written in C++ and CMSIM became a hybrid system

- CMSIM was used in writing the Technical Proposal of CMS as well as in writing the Technical Design Reports of the major detector components: Magnet, Hadron Calorimeter, Muon Detectors, Electromagnetic Calorimeter, Tracker, Trigger Systems (L1 and HLT), Computing and Software.

- The geometry of all major detector components was described in detail in CMSIM and the future evolution of geometry description is derived from there.

CMS Detector Simulation (III)

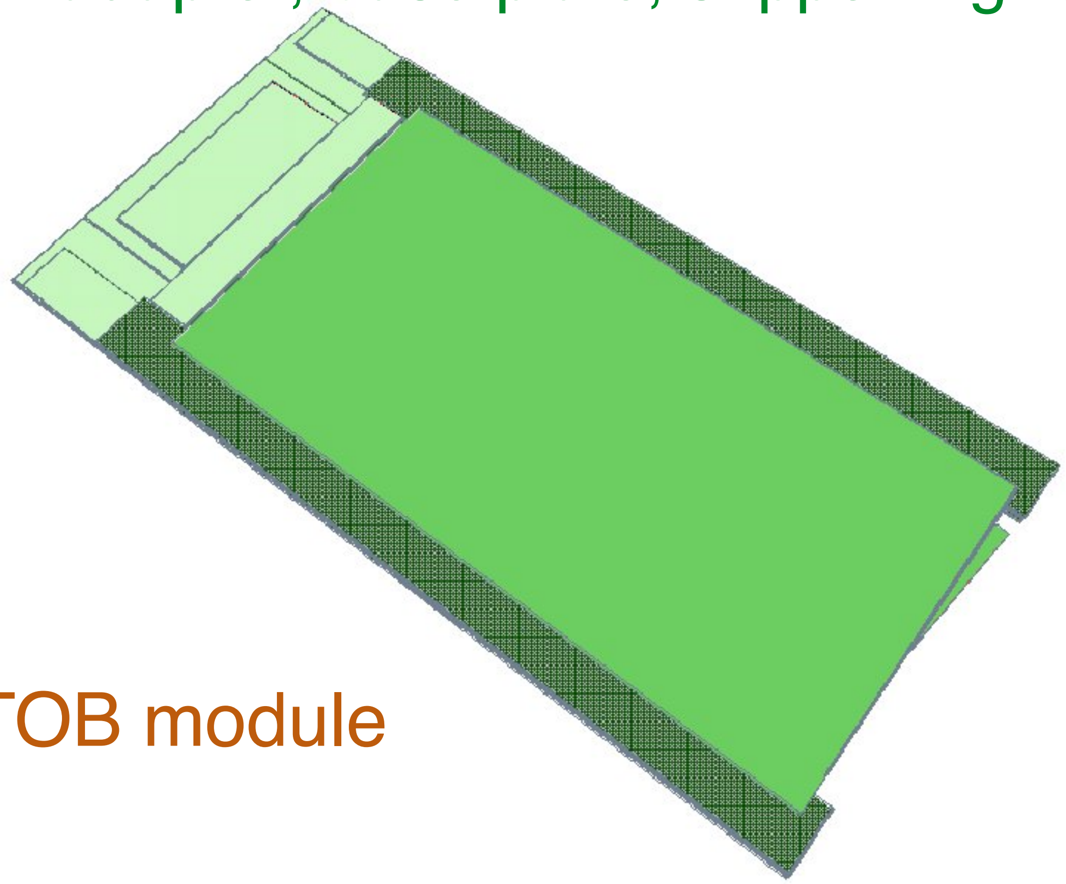
- The design of a complex experiment relies on understanding individual detector components often with test beams
- One such activity happened during 1995-1996 with the hanging file structure of alternate layers of scintillator and absorber in presence of magnetic field (along and perpendicular to beam direction)



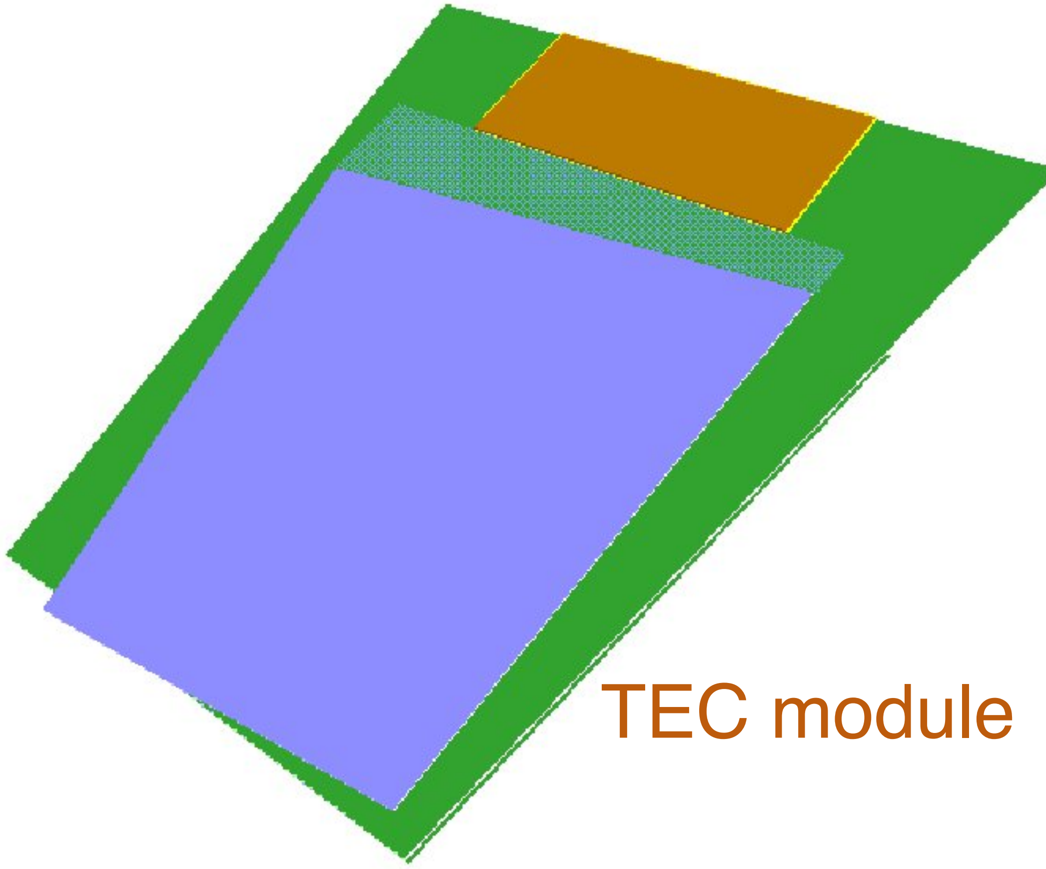
- CMS learnt from this test beam studies: the effect of B-field in the potential endcap calorimeter (only scintillation brightness) and the barrel component (combination of scintillation brightness and increased path length)
- This standalone simulation code was first written using Geant3. It was then written in C++ and became the first contribution of CMS to the Geant4 effort in the form of an advanced example (Composite Calorimeter)

CMS Detector Simulation (IV)

- As knowledge of the future tracker grew, a concrete proposal of the tracker geometry description was required. This was done in a combined effort of three individuals from Europe and Asia
- All components were described in detail: silicon wafer, readout chip, pitch adaptor, base plate, supporting rods, cables, etc.



TOB module

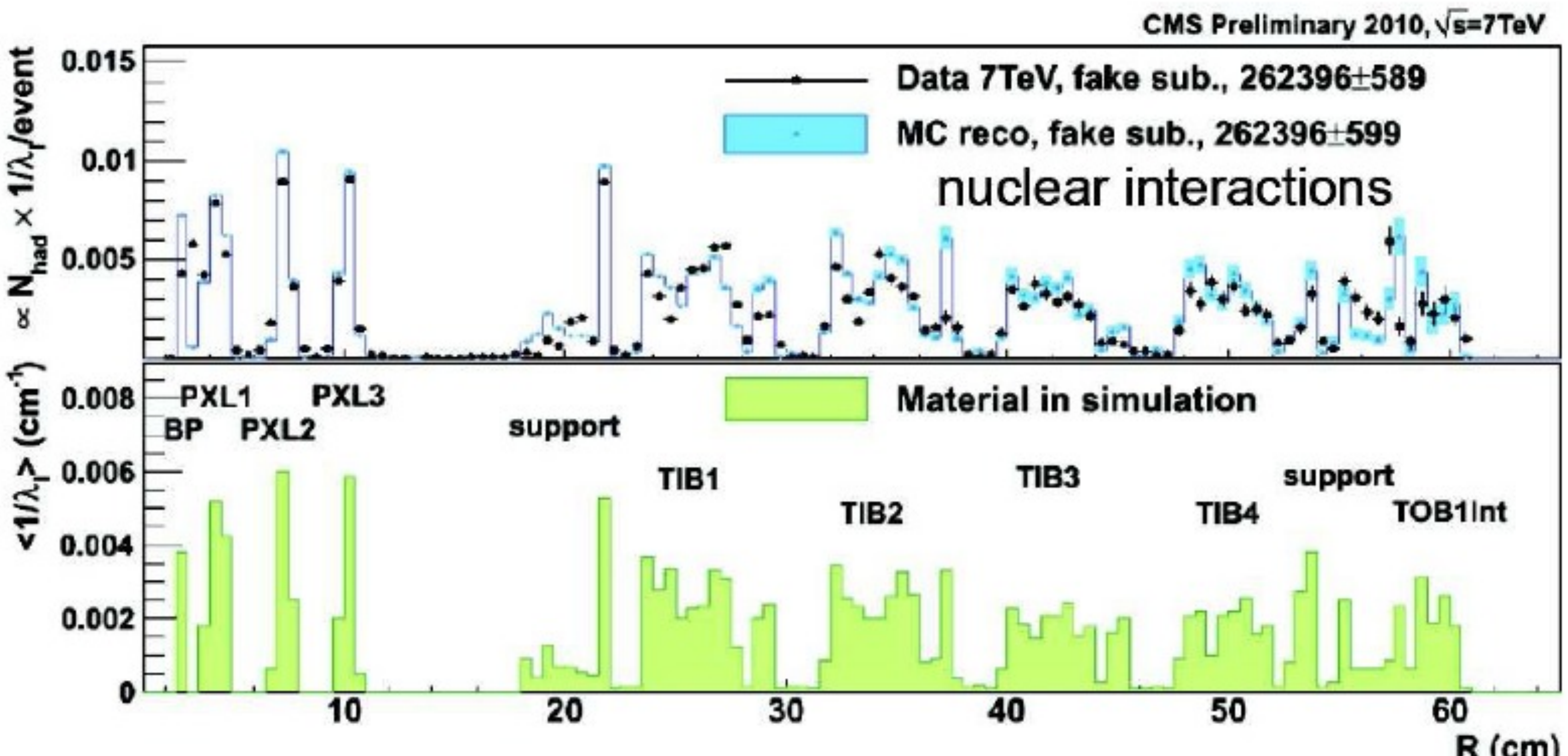


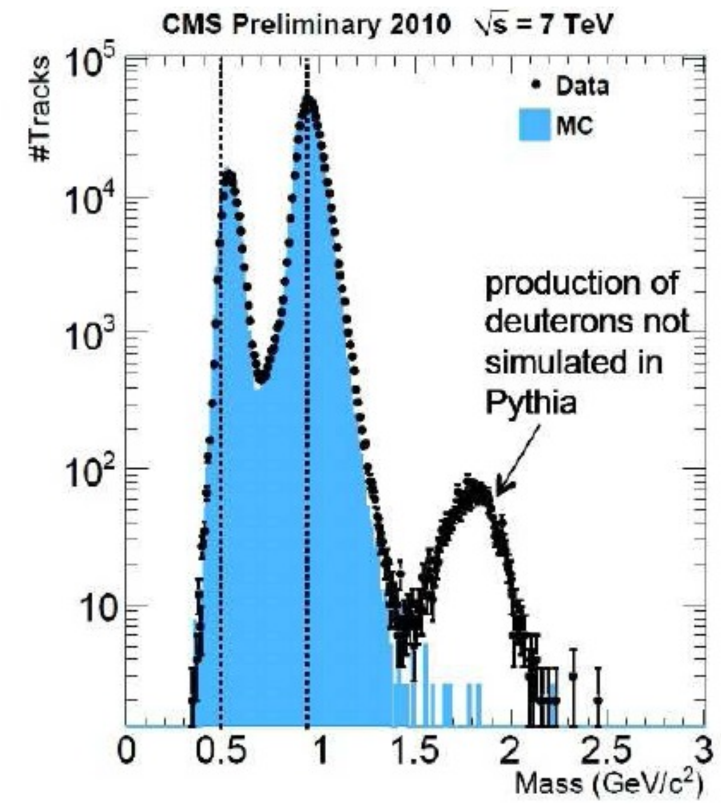
TEC module

- The components are then combined to make rods, petals, complete sub-detectors, and finally the Tracker

CMS Detector Simulation (V)

- The goodness of this geometric description was validated using the first collision data.
  - The first test was to check the weights of the components and the combined system — comparison with the measurements yields a near-perfect agreement
  - The second comparison was done from a study of conversion vertices (photons and nuclear interaction)

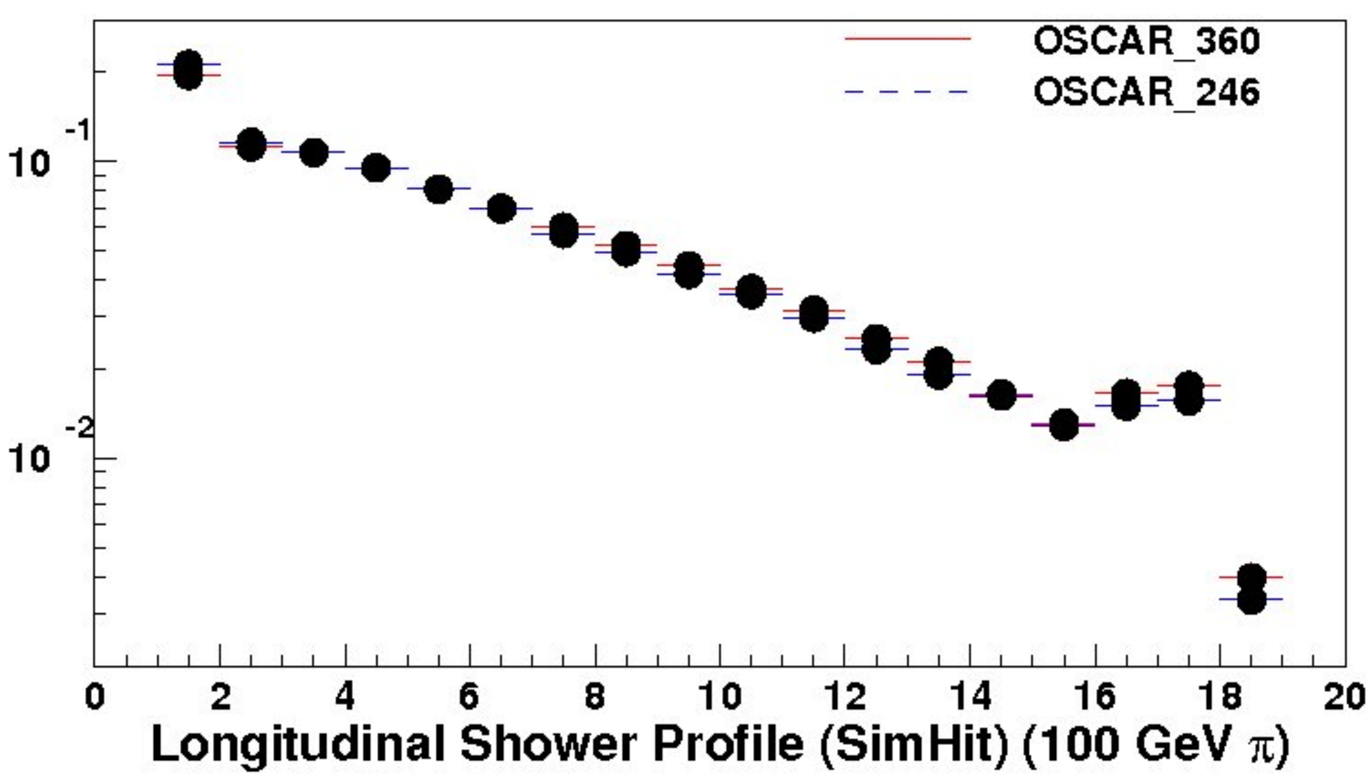


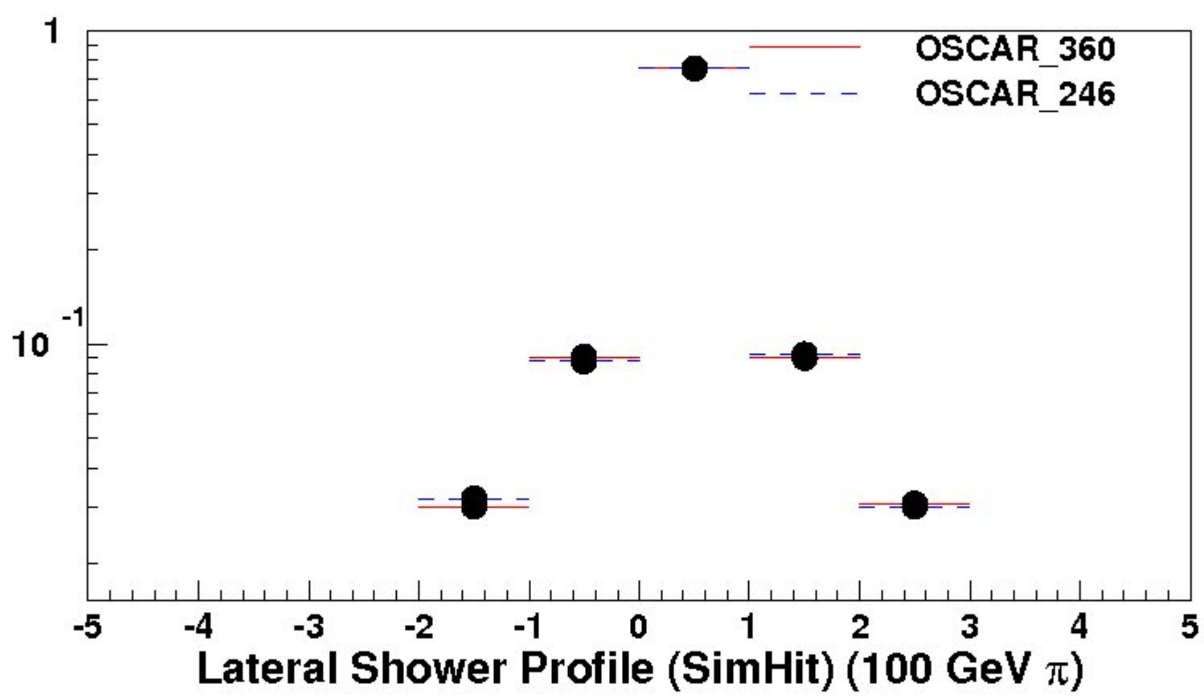


- An interesting comparison shows the inadequacy of event generators

CMS Detector Simulation (VI)

- Once the beta version of Geant4 was available, CMS used this and made the first prototype software "OSCAR" to simulate the CMS detector. It came in two versions:
  - A standalone version which only tested simulation and had no connection with the reconstruction code ORCA
  - A combined version using common tools from the reconstruction toolkit (COBRA) and thus could be tested with the reconstruction





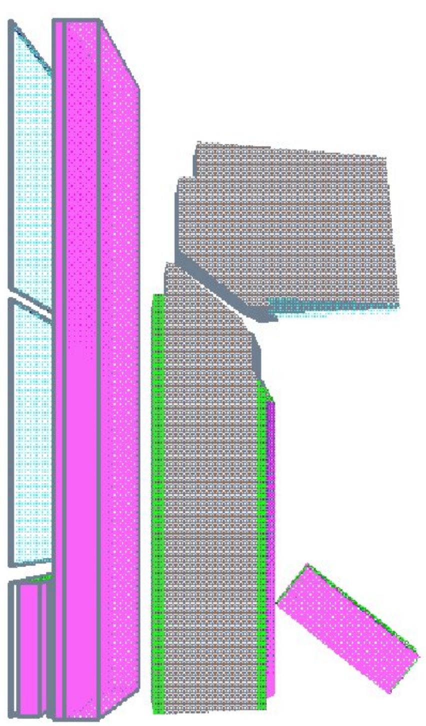
CMS Detector Simulation (VII)

- The late arrival of the reconstruction code for the Physics TDR prompted an in-depth review of the CMS software
- The new event data model, proposed by the Fermilab team, was adopted and the first version of CMSSW was released
- The simulation code was taken care of by an international team from Greece, India, Italy and Spain. A special tool was built using XML to facilitate the geometry description, and this approach was later taken up by the Geant4 team.
- For a brief period of time, two Geant4 experts were part of the simulation effort of CMS and CMS benefited a lot due to their insights.
- The code is optimized to consume less CPU time and memory (~8 times faster than the ATLAS full simulation code)

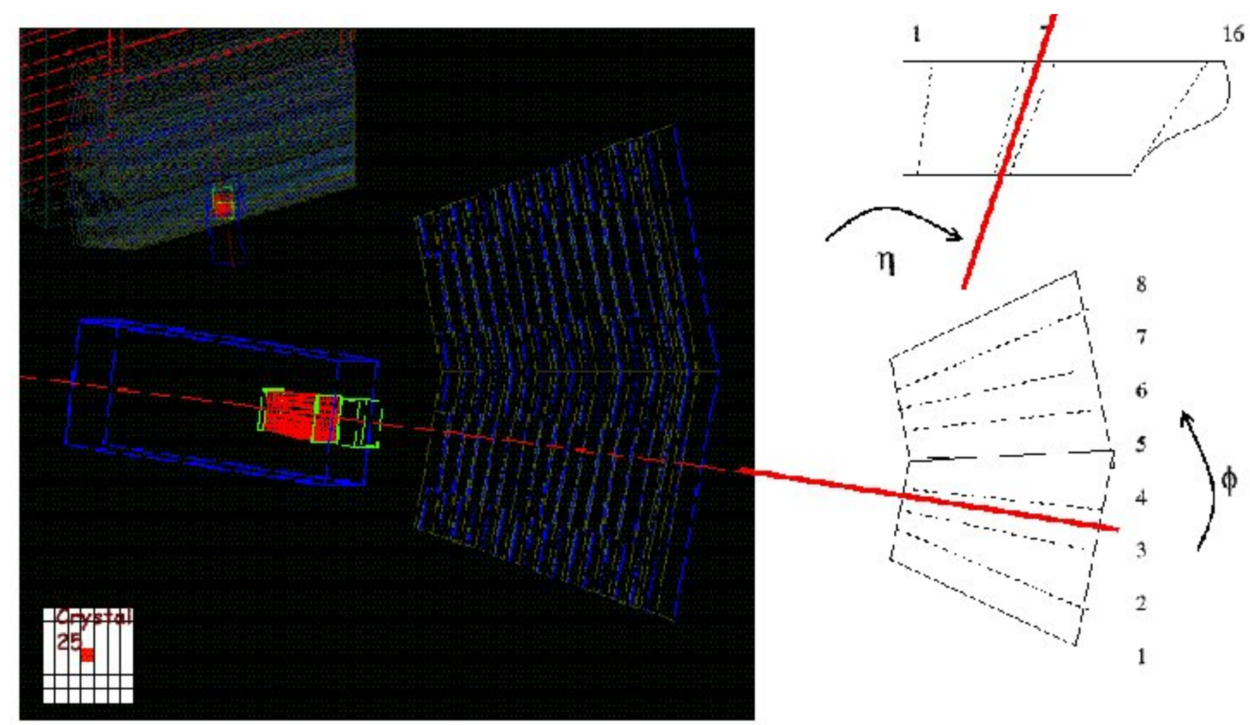
CMS Detector Simulation (VIII)

- The calorimeter teams made a continuous effort in learning about EM and hadronic showers. The major simulation effort went to describe the test beam setups of 2002, 2004, 2006 and 2007 in the H2 beam line of SPS

- CMS built a table to accommodate one endcap and two barrel modules of HCAL with the corresponding structure of HO and a mockup structure for the magnet coil. This could be preceded by a module mimicking ECAL or a real EB supermodule

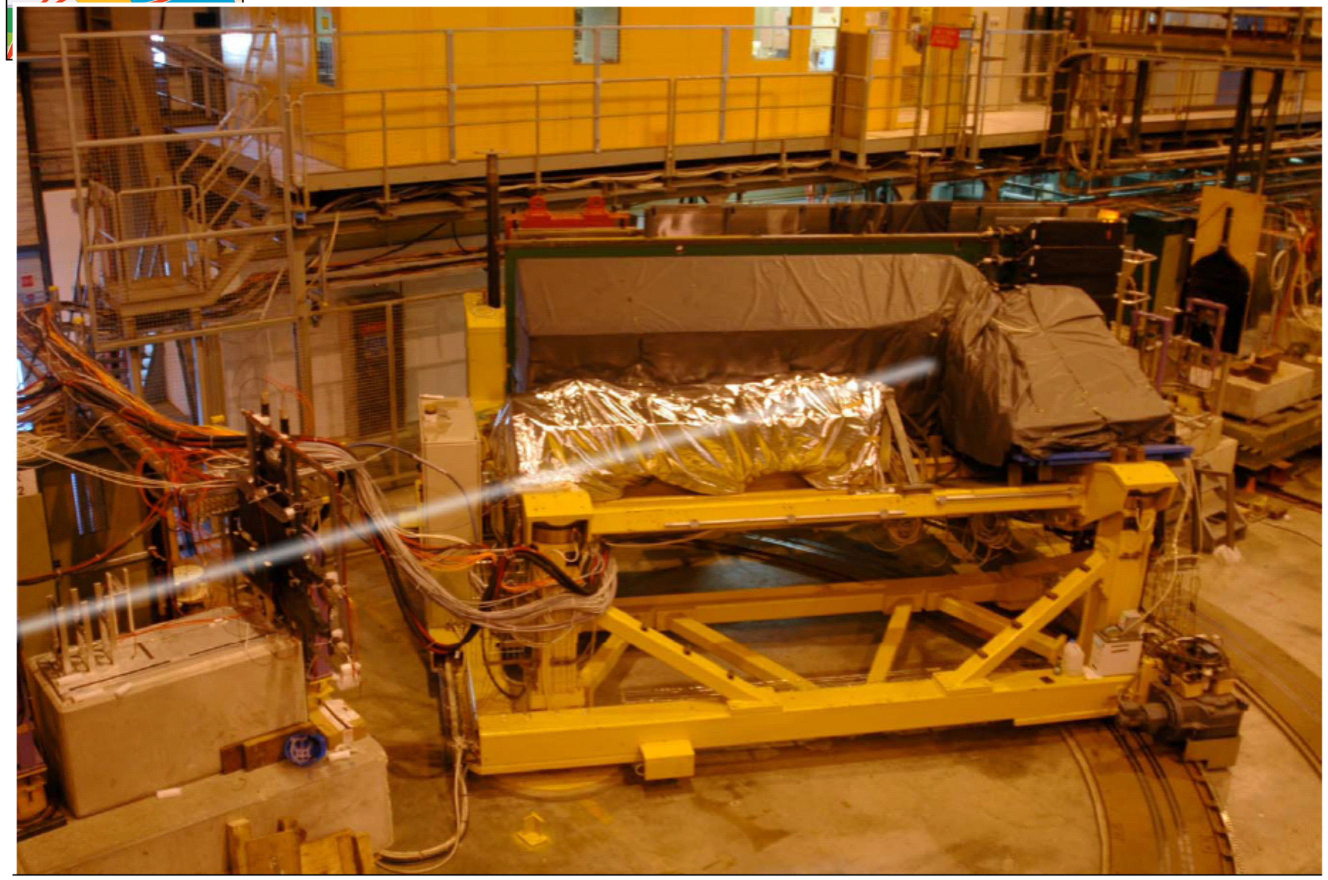


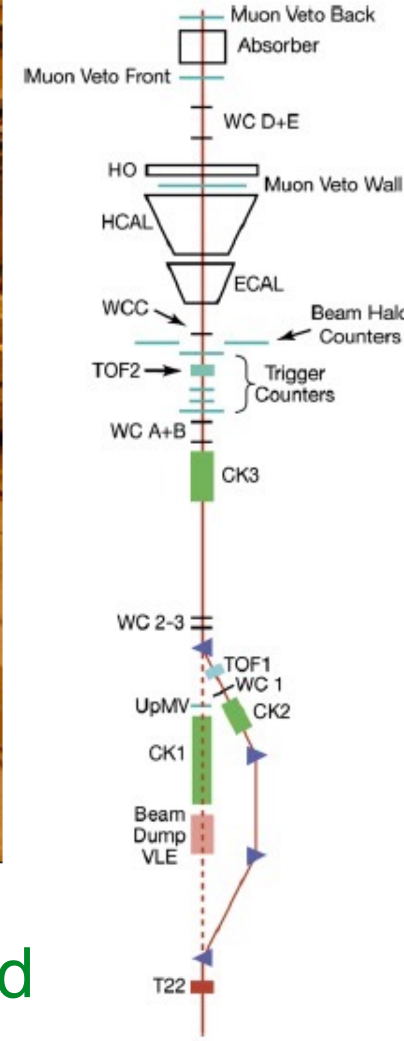
2004 Setup

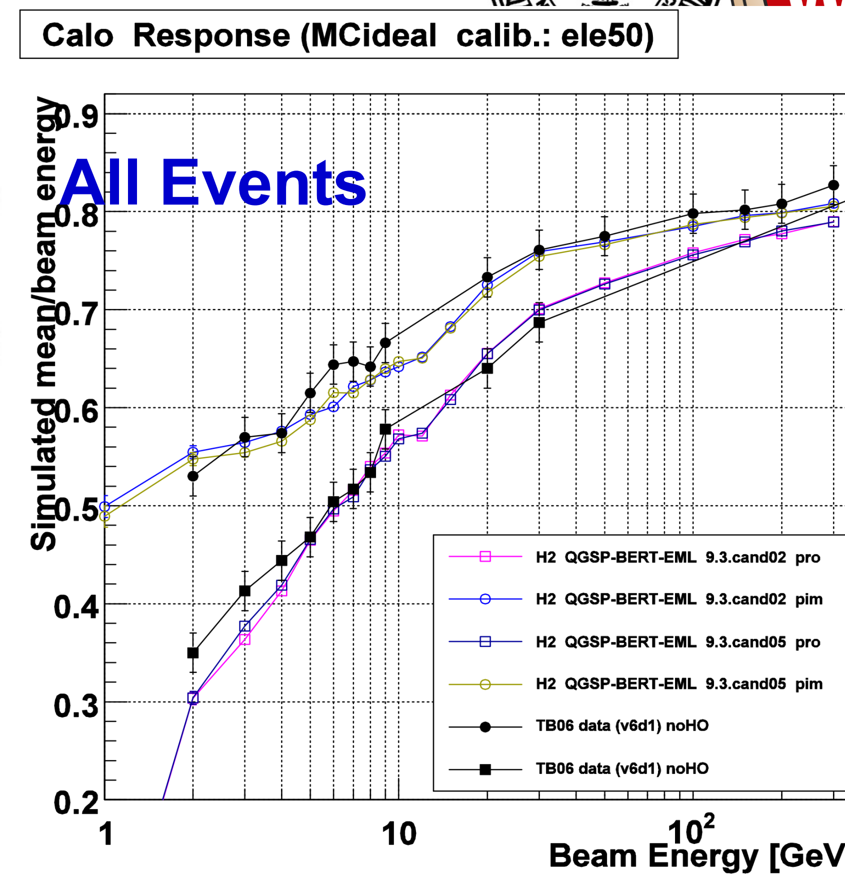


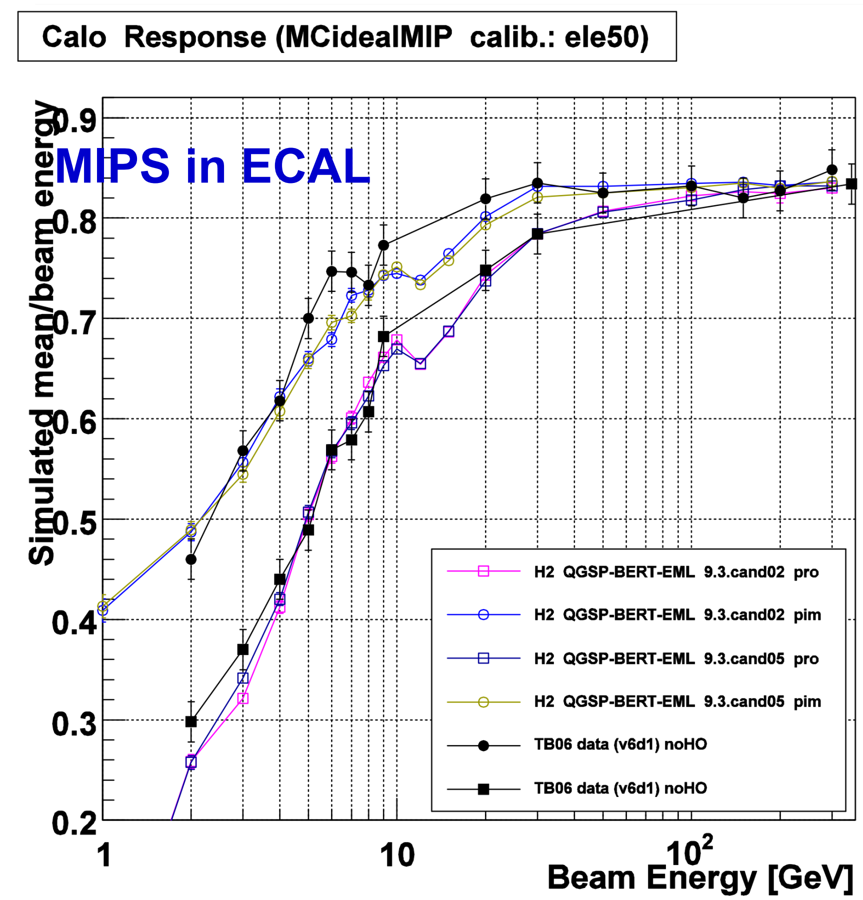
2002 Setup

CMS Detector Simulation (IX)





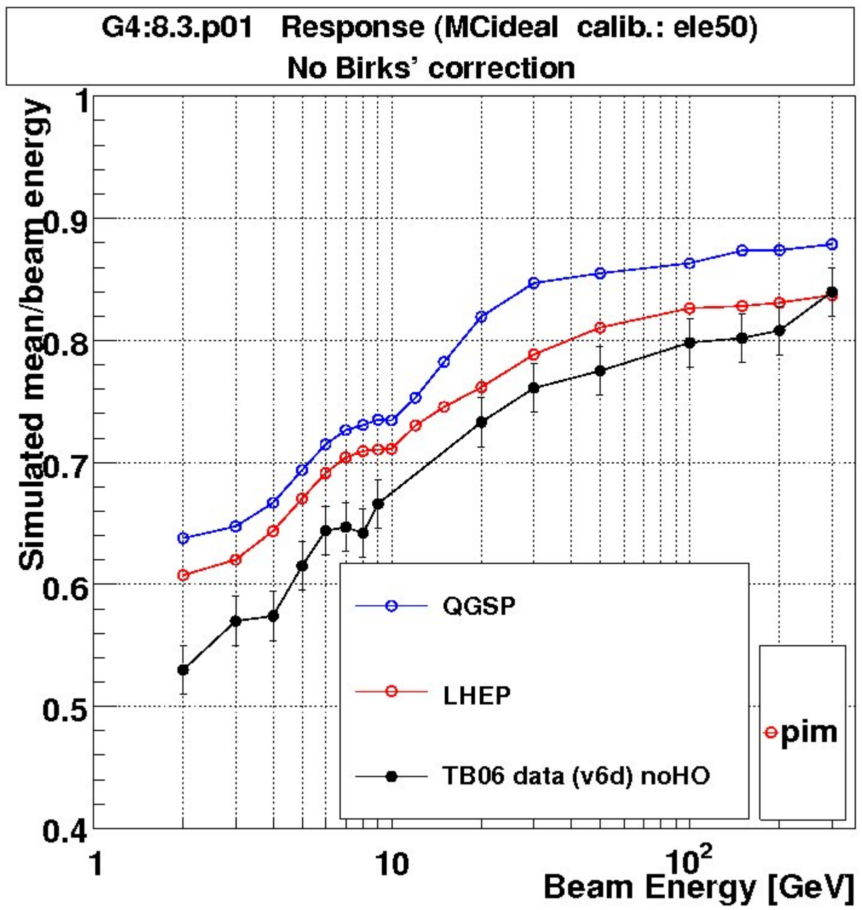


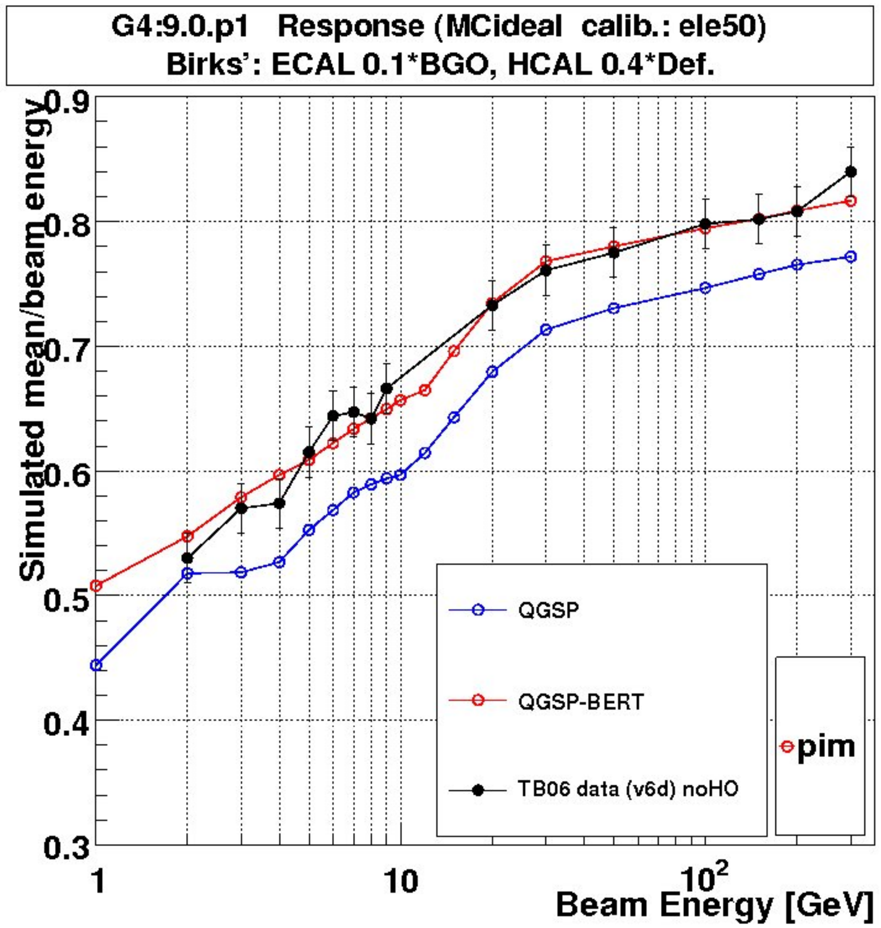


- In 2006, the test beam setup was equipped with a sophisticated particle identification system to distinguish among pions, kaons, (anti)protons of the incident beam. Also, the beam energy could be brought down to 1 GeV so that the shower energy spectrum can be measured for identified hadrons over a large energy region

CMS Detector Simulation (X)

- Geant4 came up with 2 major physics lists to describe hadronic showers
  - LHEP: A parametrized model based on GHEISHA code due to Ham Fesefeldt
  - QGSP: Starts with a Quark_GLuon string model with a better theoretical insight

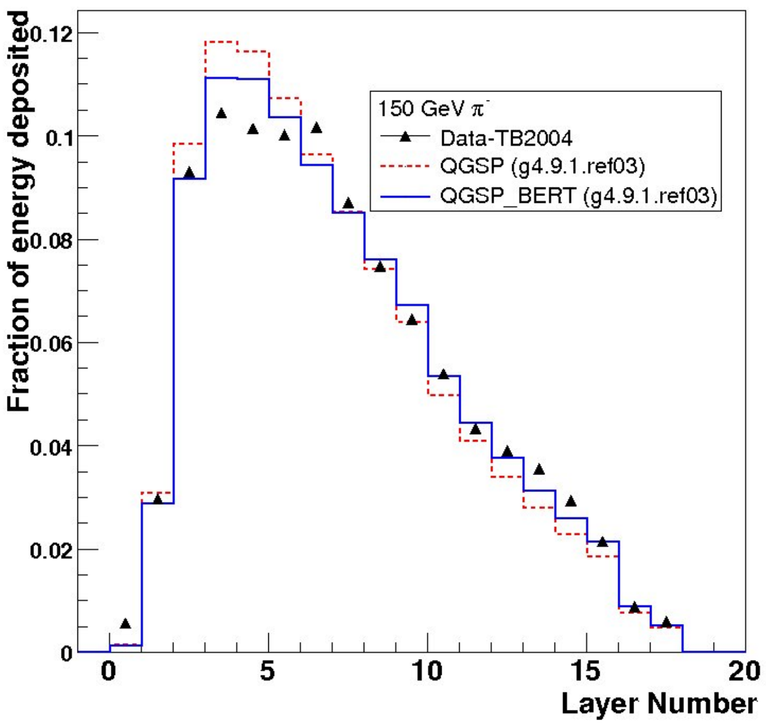


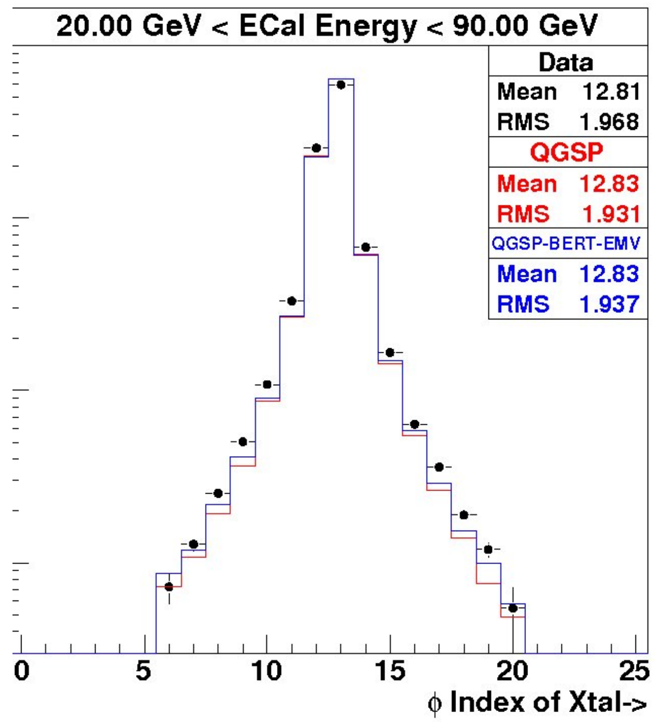


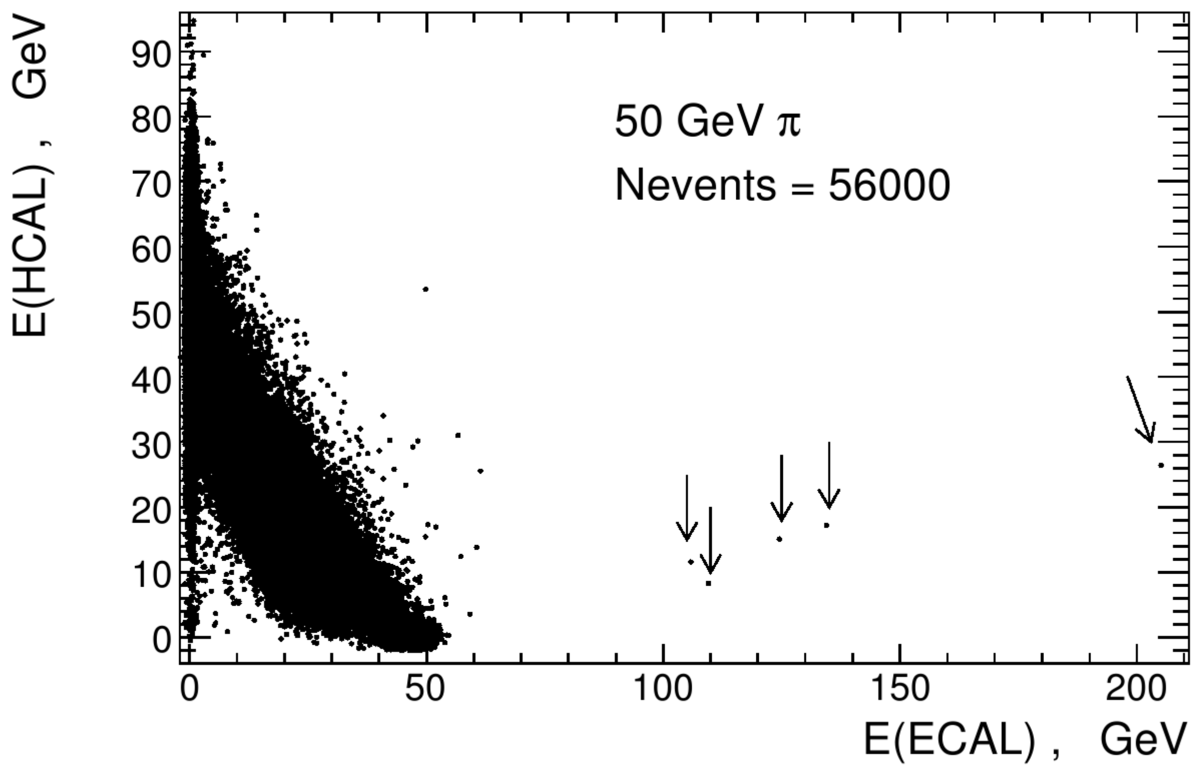
- Confronted with data, the QGSP model needed 2 serious extensions:
  - At low energies, QGSP has to rely on a cascade model due to Bertini
  - It is required to introduce the effect of saturation in producing scintillation light, often called Birk's correction, for both crystals and plastic components

CMS Detector Simulation (XI)

- There was a task force to investigate the goodness of CMS simulation. It examined lateral, longitudinal shower profiles for EM as well as hadronic showers and leant a bit about large energy deposits in the ECAL





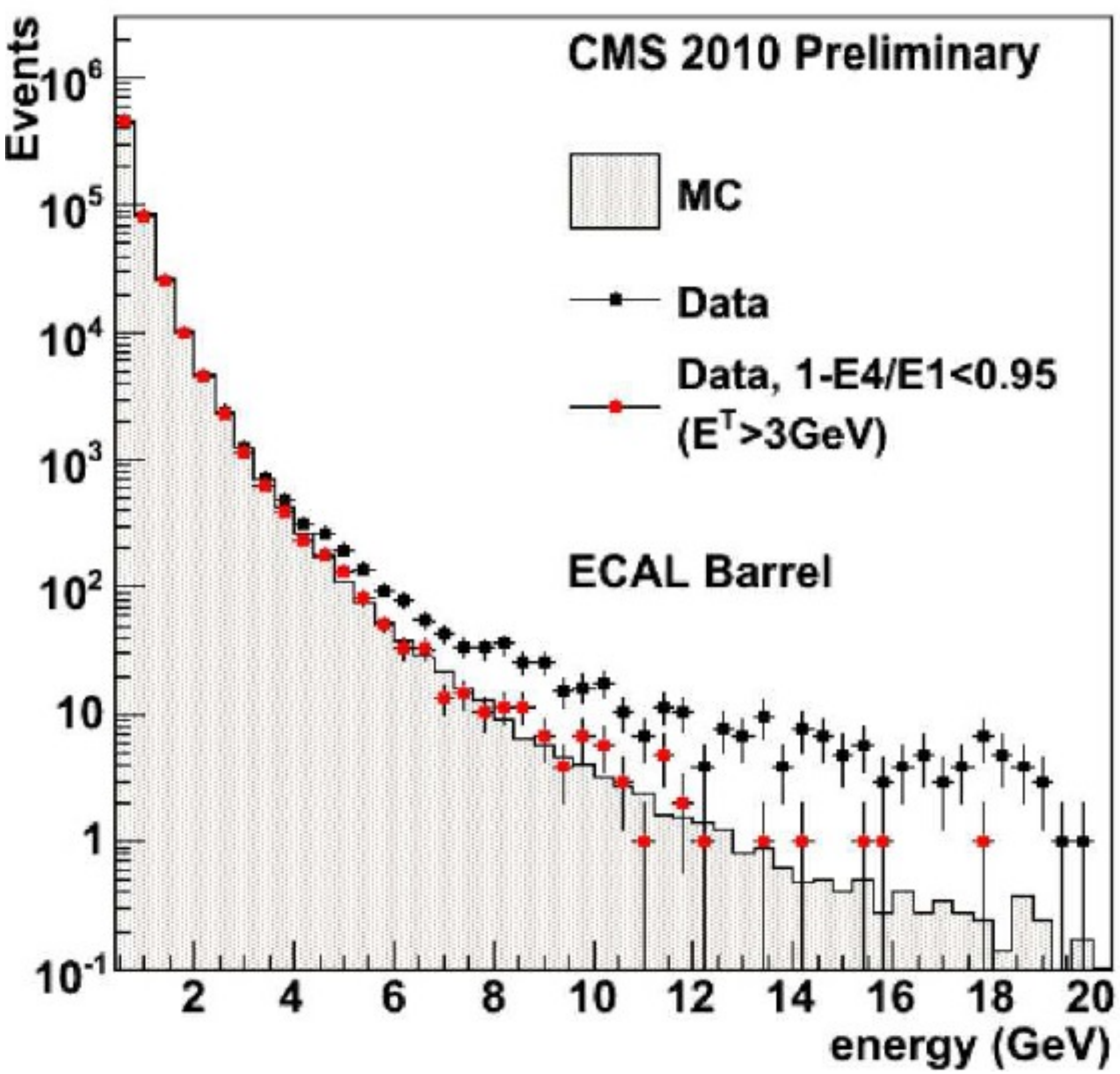


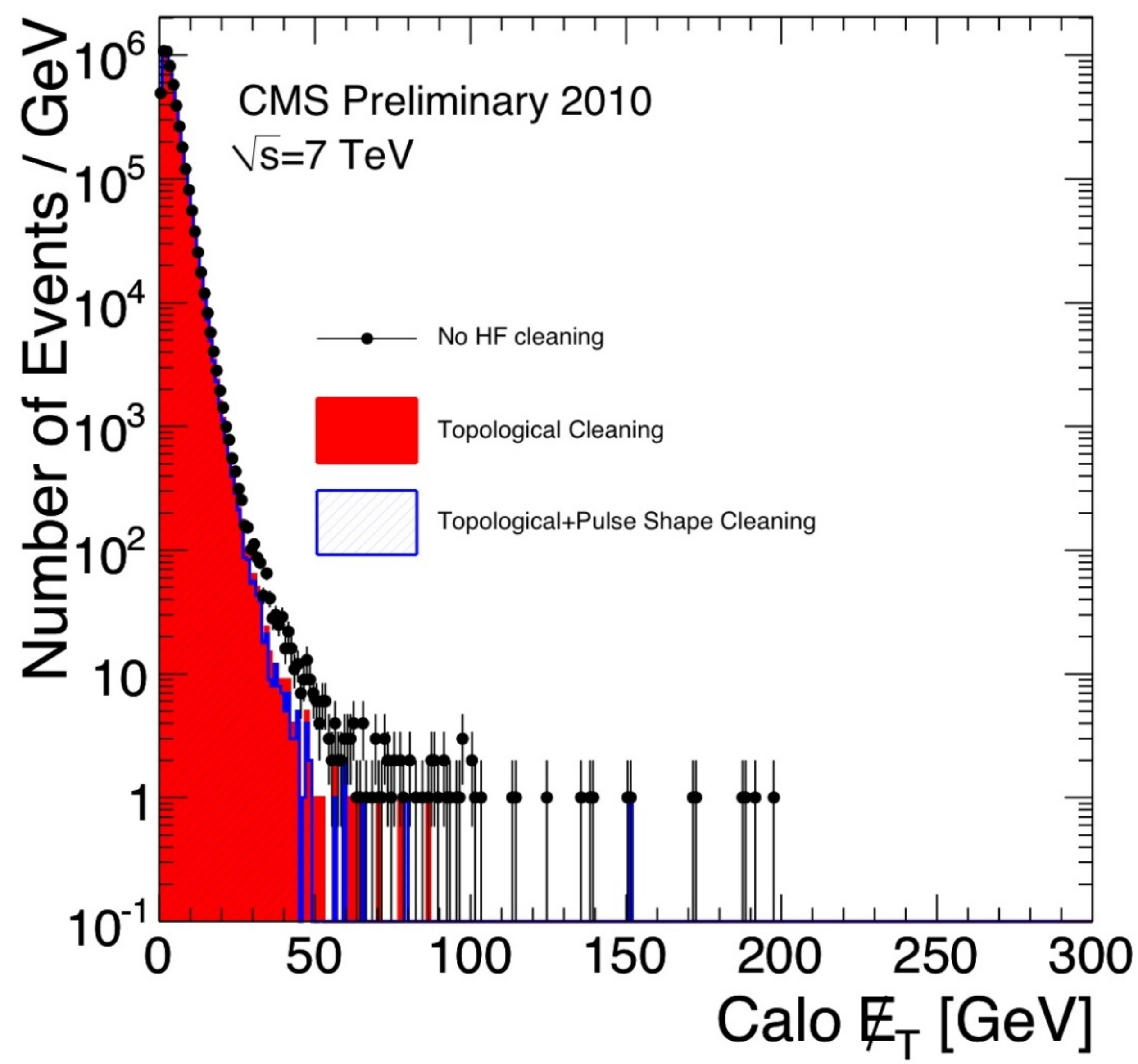
- The team recommended the use of the physics list QGSP_FTFP_BERT and also the use of Birk's law

- There was a concern about the utilization of the combined calorimeter to hadrons which was initially ignored before CMS was shocked by the first collision data at high energy (7 TeV)

CMS Detector Simulation (XII)

- CMS had serious issues in measuring calorimetric energies during the first phase of high energy (7 TeV) data taking:
  - long-tail at high energy in the energy measurement of ECAL
  - long-tail in the missing transverse energy in the data which could not be explained by MC

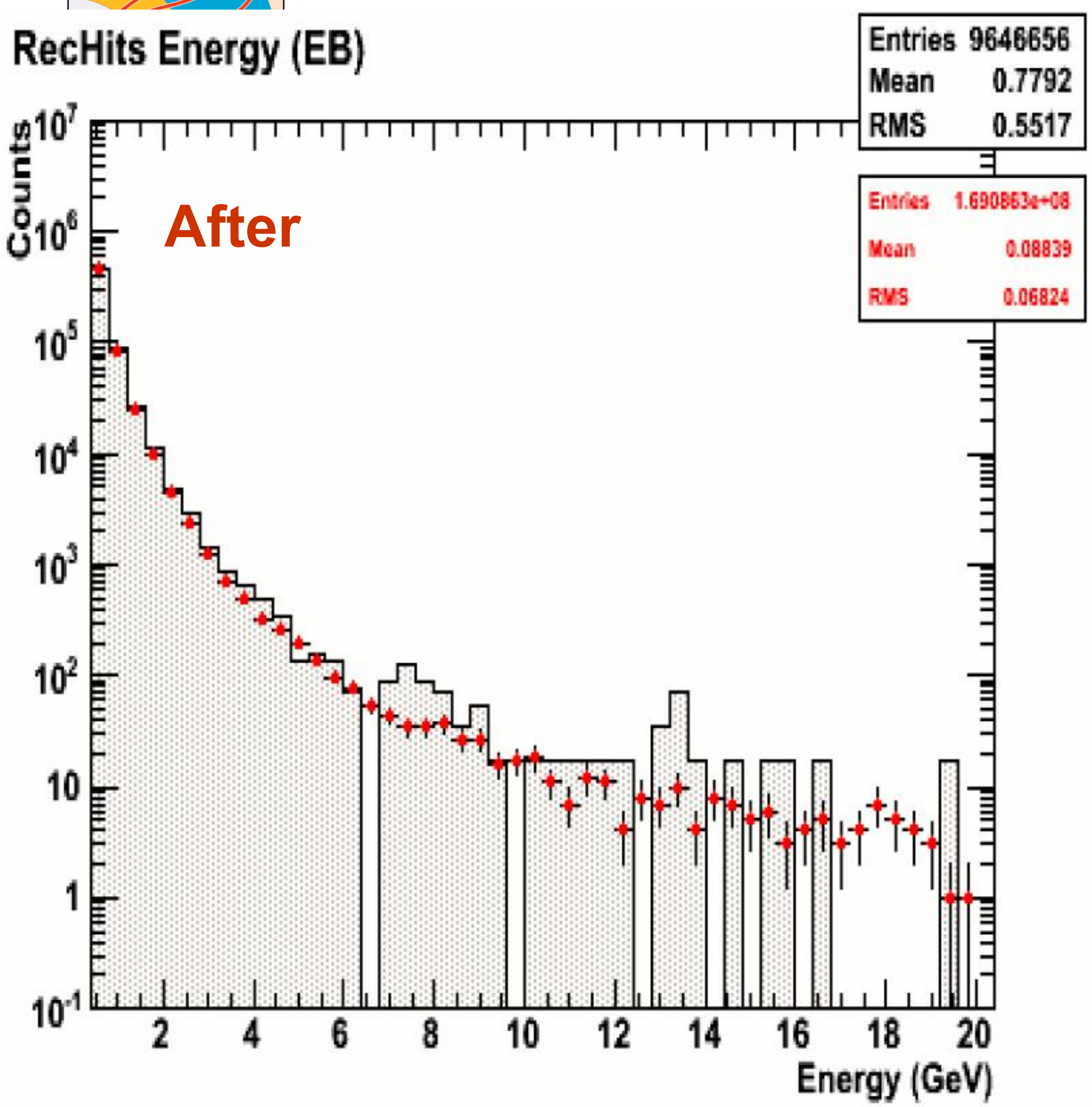


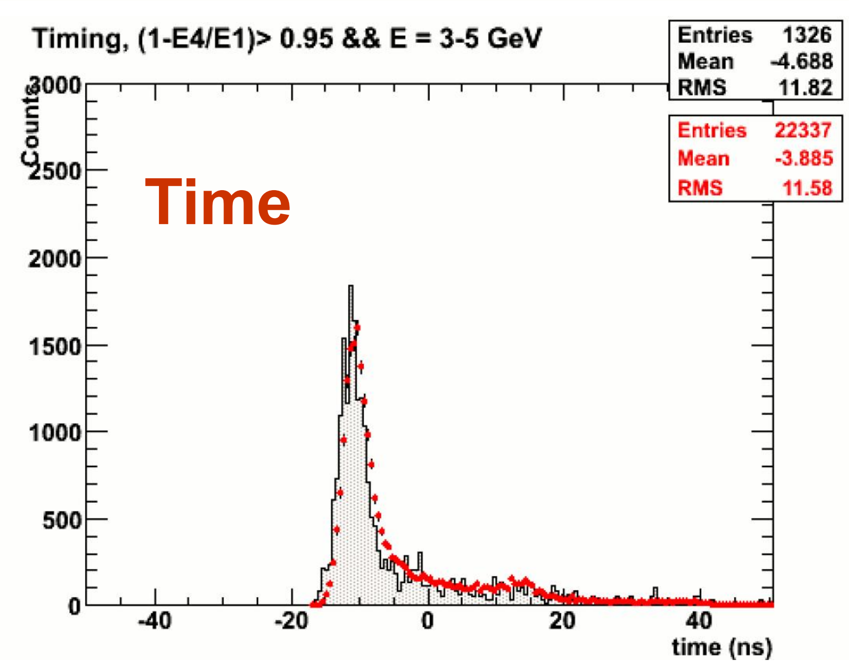


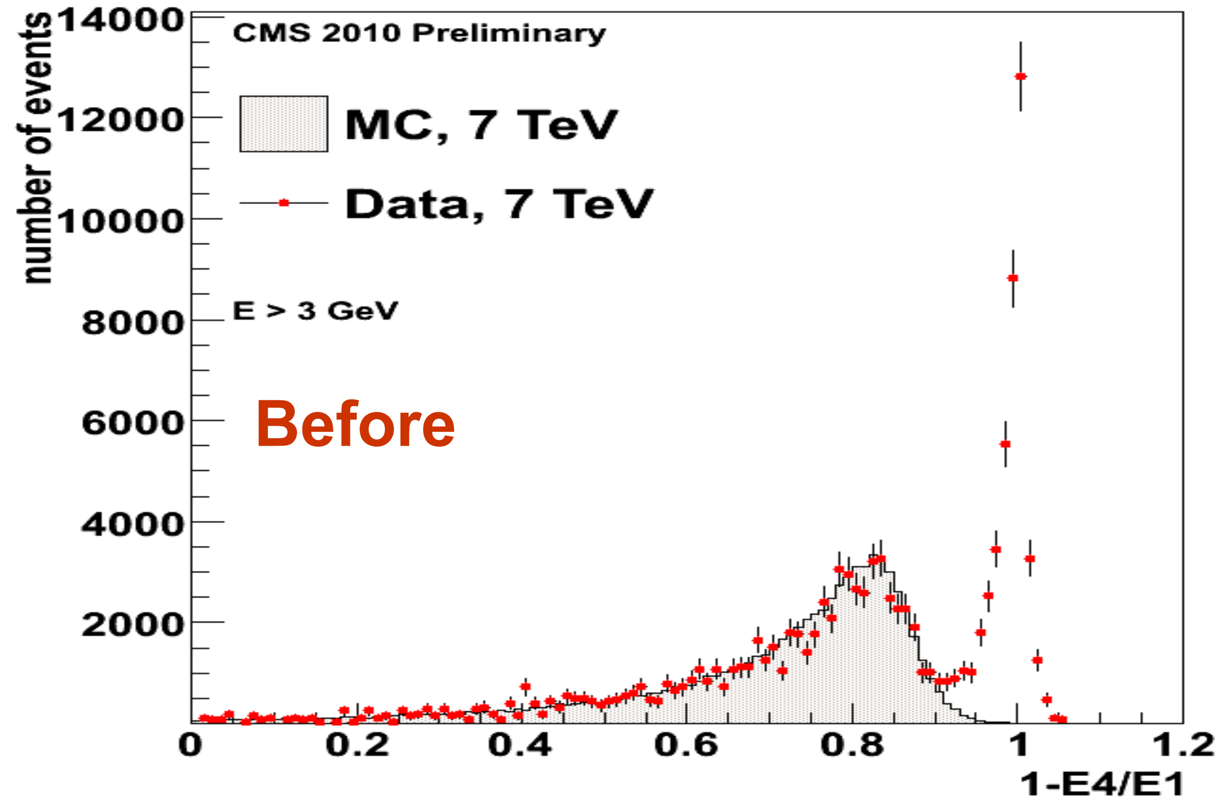
CMS Detector Simulation (XIII)

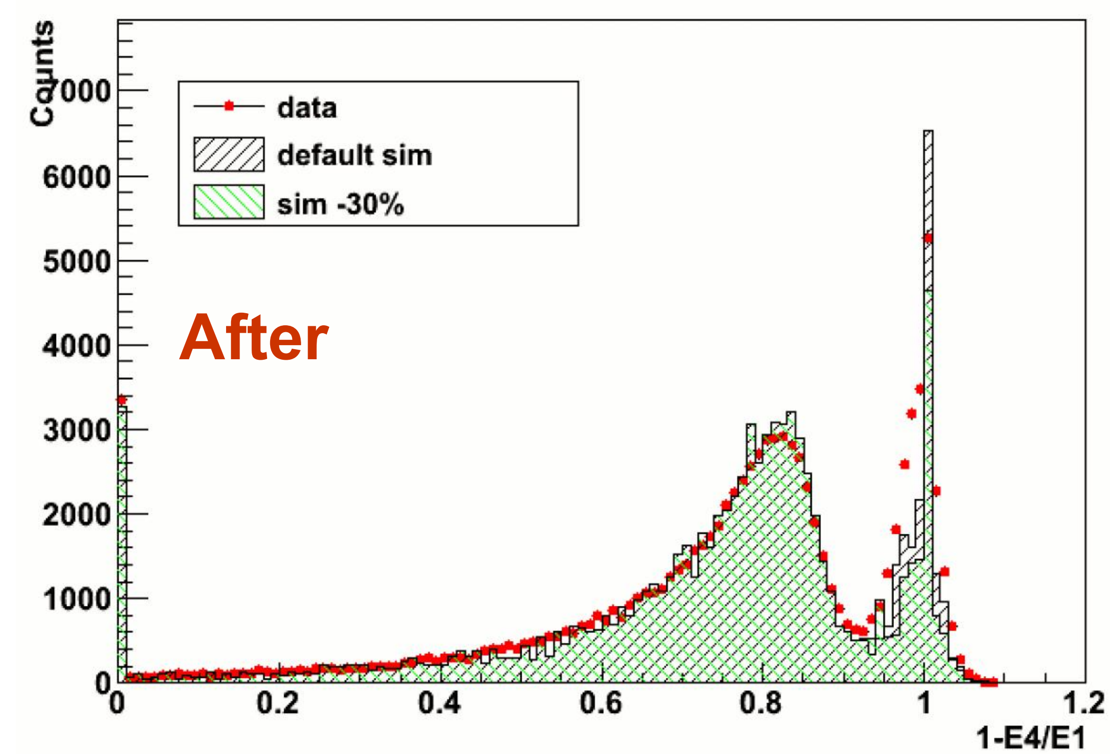
- Anomalously large signals were observed in the ECAL with the appearance of very large energy deposits in a single crystal
- These events are uniformly distributed in the barrel part where the readout utilizes APD. They are not seen in the endcap crystals which are readout using VPTs
- The rise time of the electronic pulse is consistent with an instantaneous signal from the APD, not the typical decay spectrum of the crystal
- The rate is roughly proportional to the minimum bias rate
- They are not observed during the Cosmic Ray runs, only during the collision and in the test-beam runs with incident hadron beams
- The simulation code was changed to treat the crystals and APD volumes as independent detectors. Energy gain in each gets different gain factors
- The simulated rate for energy deposits in a single APD volume above a threshold matched the rate in the data
- The simulation could also match the energy spectrum for the passage of single muons in the detector
- Time distribution also matches between data and simulation
- It was concluded that the anomalous hits are due to the energy loss of heavily ionizing particles (protons or ions) in the APD

CMS Detector Simulation (XIV)







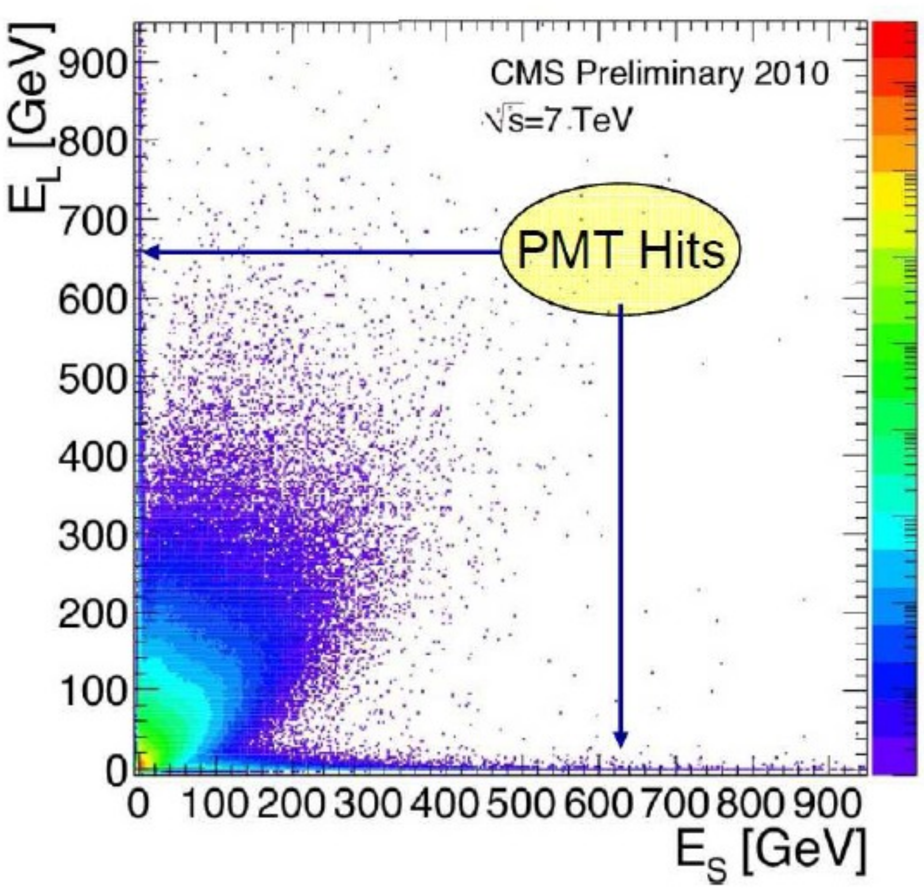


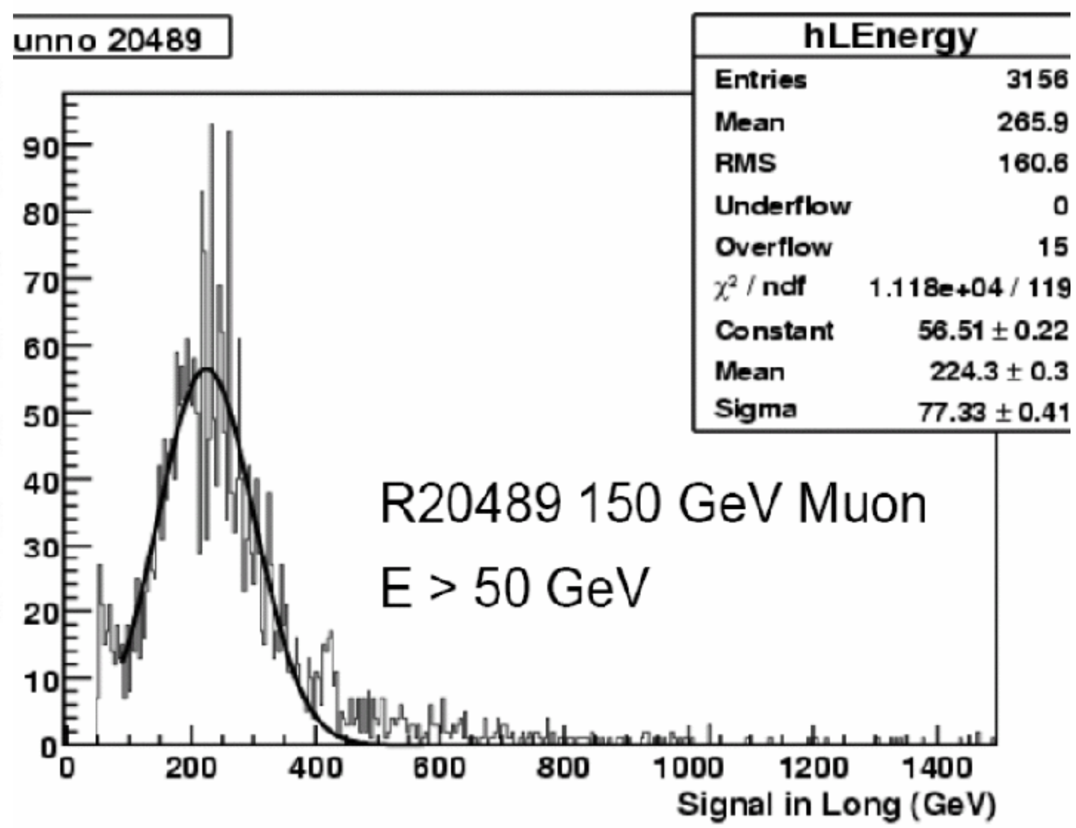
- Much improved understanding of the data

S. Banerjee

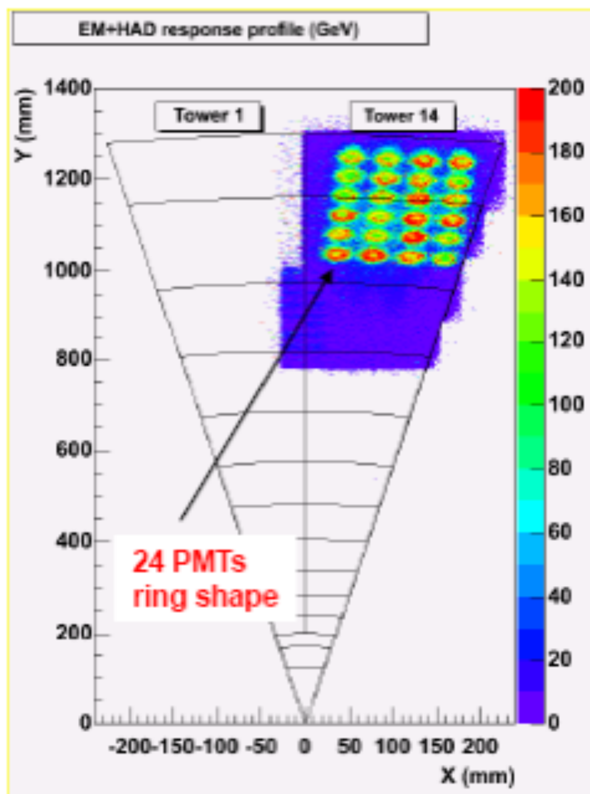
CMS Detector Simulation (XV)

- Missing transverse energy is a key tool in the search for new particles in HEP. The long tails in the MET spectrum (which cannot be explained by simulation) were a worry.
- The events with large MET were having very high energy hits in the forward hadron calorimeter
- The large energy was seen in one type of fibre (either long or short) covering the same phase space (in η and φ)
- Even muons in the test beam runs gave rise to large pulses. These large energy deposits were identified with direct hits to the PMT sitting behind the absorber



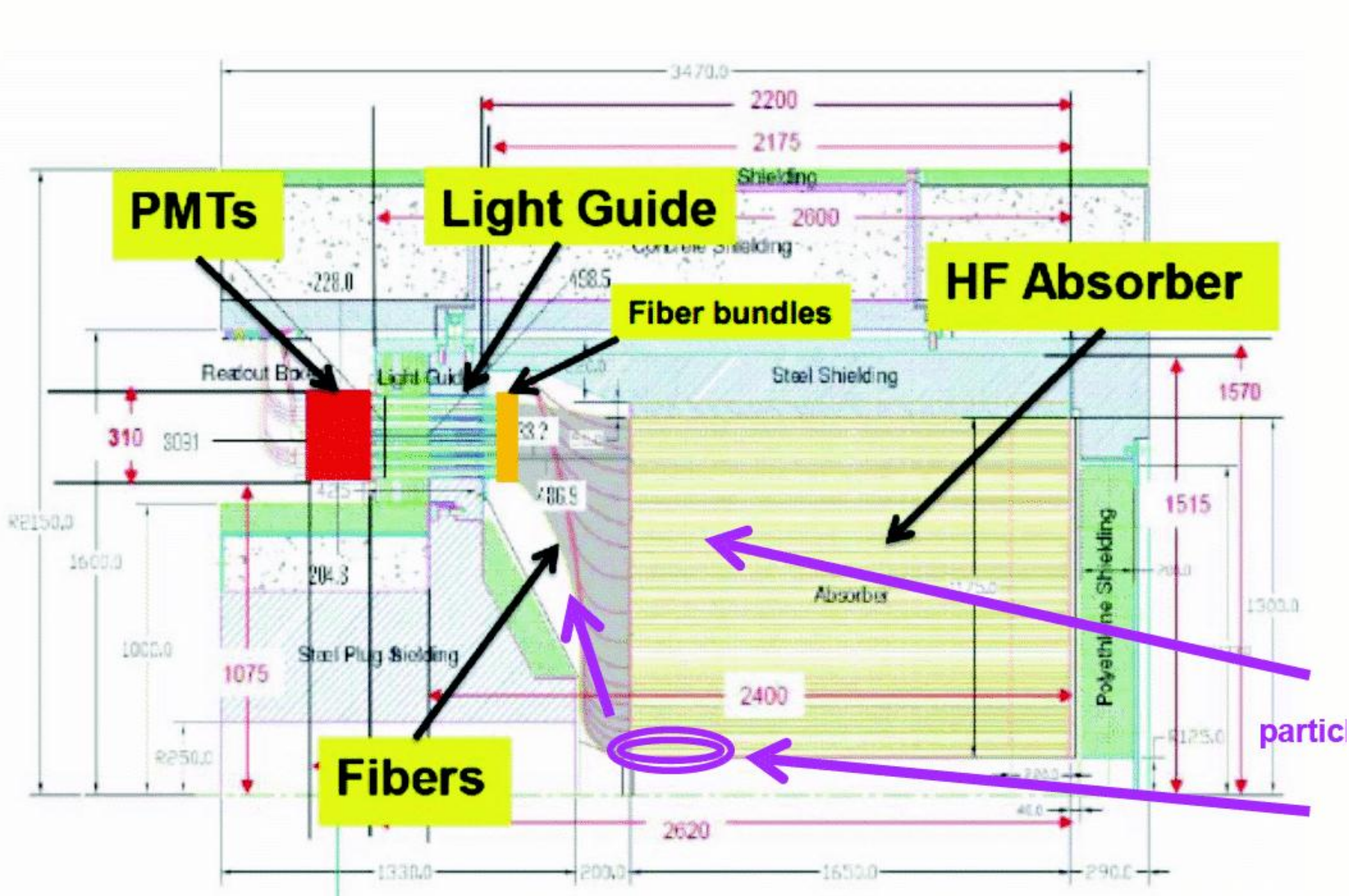


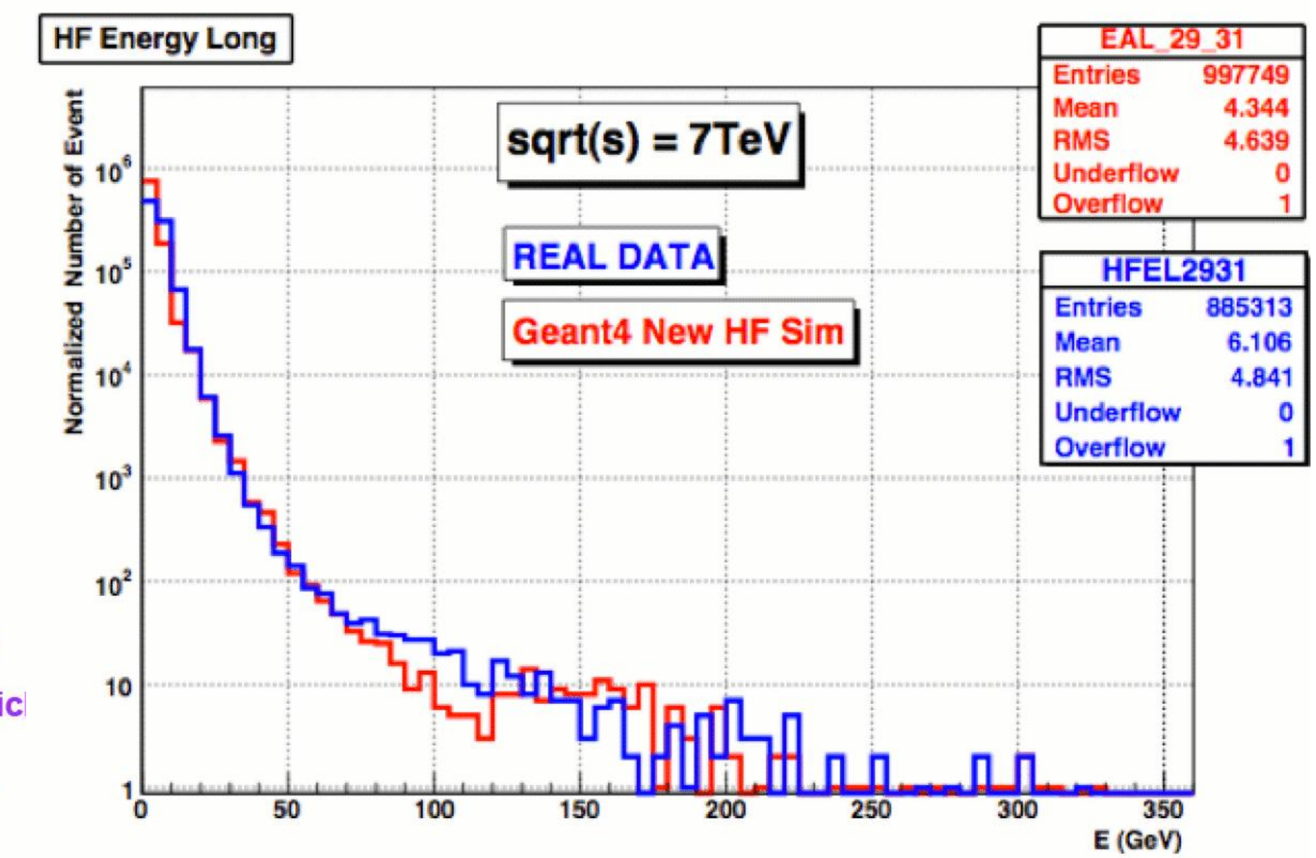
150GeV Muon / Wedge 2-6

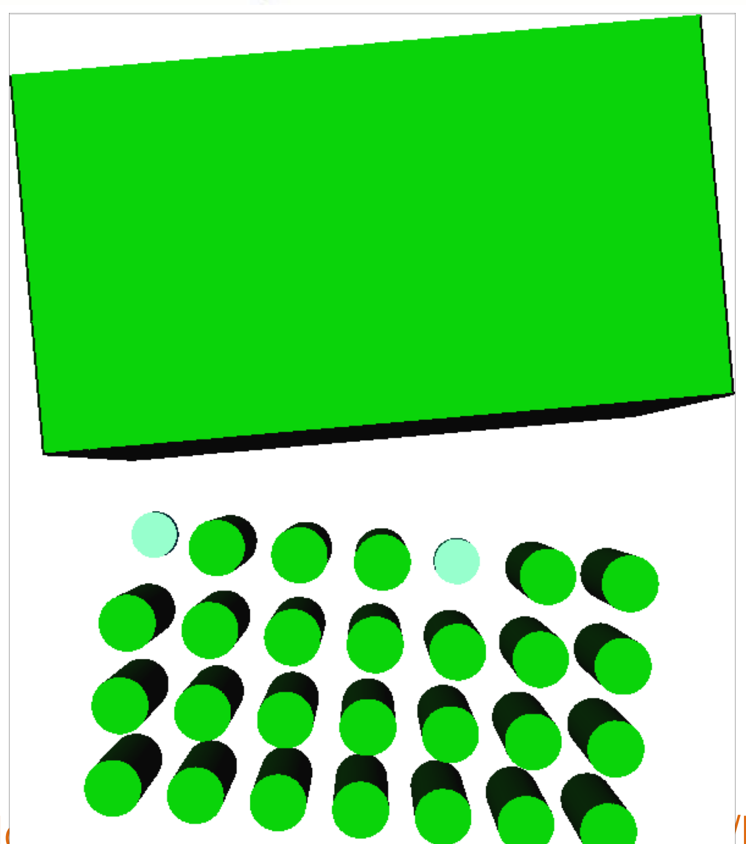


TB2004 elog 3601 Aug. 2, 2004

CMS Detector Simulation (XVI)







- Describe the PMTs behind HF and declare the photocathodes as Sensitive Detector
- Also, fibre bundles are described in the geometry and hits in the fibre bundle are associated with a given readout channel
- Energy spectrum, as well as anomalous hits, are well reproduced in the simulation
- The dominant source of these hits is muons from decays in flight and hadron shower punch through

CMS Detector Simulation (XVII)

- CMS detector is not a static object. It has evolved over the years
  - During LS1, the beam pipe was changed in view of a modified pixel detector
    - Some of the forward detectors were modified
      - Beam scintillator (BSC was removed), Beam Halo Monitor (BHM) and Pixel luminosity monitor (PLT) prototype were introduced
    - The PMTs for HF were replaced with a new set (the single anode is changed to a set of four)
    - HO readout system was changed from the use of HPD to SiPM
    - Totem and CASTOR detectors were decommissioned (partly here and partly during run2)
  - HF readout was modified to have a single cell being readout twice
  - The pixel detectors (both barrel and endcap) were modified
  - Readout box for HE started using SiPM and the number of depth segments was significantly modified
  - Some demonstration chambers for the first layer of the GEM detector were introduced
  - During LS2, some more changes are foreseen
    - Readout boxes for HB also use SiPM and have more depth segments
    - The first station of the GEM detector is now complete
    - Demonstration chambers for the second station of GEM detectors and the lowest rings of forward-backwards RPC detectors (stations 3 and 4) are inserted
    - A new beam pipe is put in and the shielding structure is modified
    - A new detector in the position of Totem T2

The simulation program is also not a static object. It supports multiple scenarios

CMS Detector Simulation (XVIII)

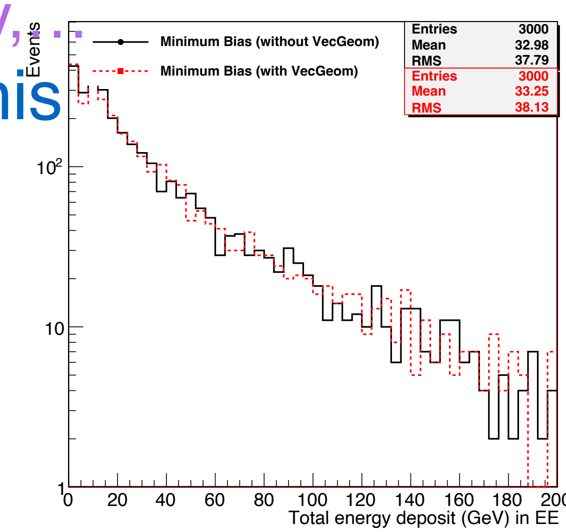
- Many detector elements will be unusable during the high luminosity runs of the LHC — some detectors are damaged due to radiation, and some detectors will suffer due to higher occupancy
- A major change is foreseen in the CMS detector
  - A new tracker will replace the present pixel and strip detectors
  - The barrel calorimeters (both ECAL and HCAL) will have new electronics to extract timing information with much better resolution
  - Layers of detectors will be introduced in the barrel and endcap to provide timing information with high precision
  - The muon detectors will improve solid angle coverage by completing the second and the zeroth GEM station, and the detectors in the lowest rings of the third and the fourth RPC stations
  - The endcap calorimeter (both ECAL and HCAL) system will be replaced by high granularity calorimeter utilizing silicon and scintillator detectors
- Many of these changes require verification by exposing prototype detectors in the test beam facilities
- The simulation program not only takes care of the modified CMS detector, also the individual test beam scenarios

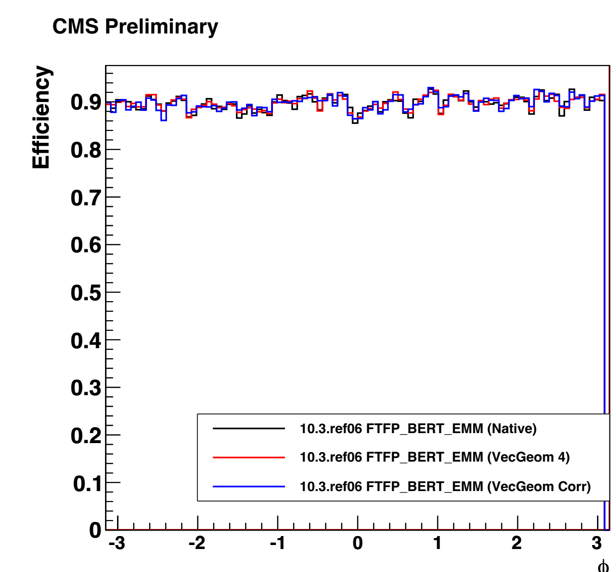
Progress of Basic Toolkit

- Computing moves from single-threading to multi-threading and this advantage is already utilized by Geant4 and all application codes (ATLAS, CMS, LHC-b)
- There was an R&D effort to make use of vectorisation
  - The geometry and tracking code in the EM field was rewritten to enable effective vectorisation
  - A new approach was made in tracking by basketizing particles to be tracked in the same volume
  - EM physics code was rewritten to match the physics performance of the scalar version
  - Adopted by CMS and verified the performance (physics + computing)
  - Observed a speed up by a factor of 2 — this is identified to be due to better algorithm and proper packaging of the code
  - The experience was transmitted to Geant4 (use some of the new codes and packaging)
- A new effort has now started to make the simulation code run on heterogeneous architectures
  - Utilise the benefits of CPUs (efficient in branch prediction and instruction prefetching) as well as those from GPUs (hundreds and thousands of simple cores and efficient in single instruction multiple data handling)
  - Also, improve the physics predictions to move to higher energies (100-1000 TeV)
  - Good progress is observed; not yet ready to provide a stable and well-tested toolkit

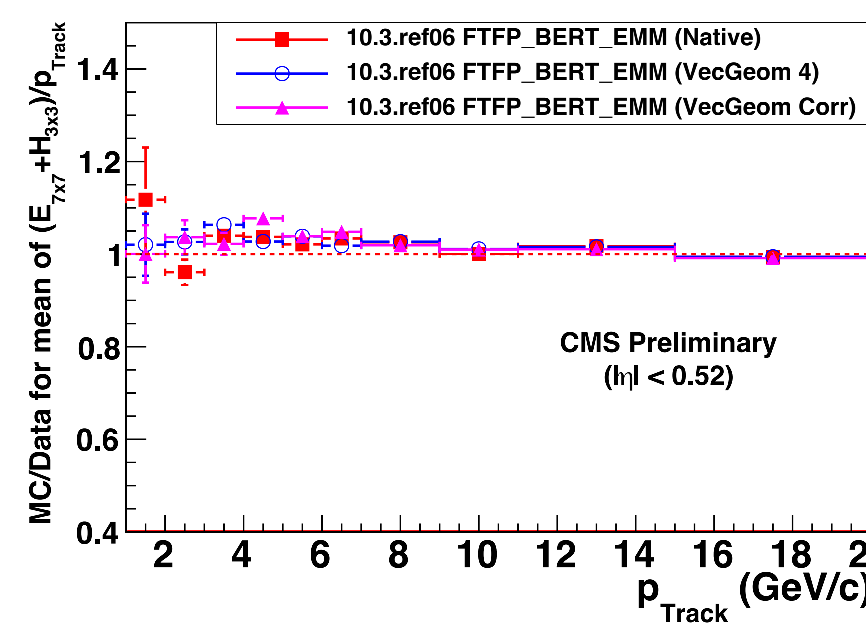
Use of these Improvements

- CMS closely monitors the developments within Geant4 and utilizes some of the improvements on a regular basis
- For example, the alternate geometry code (VecGeom) developed initially for a vectorized version is well integrated with Geant4
  - It showed significant improvement in performance for CMS simulation
  - Physics predictions were also examined:
    - Calorimetric measurement, track efficiency, ...
  - CMS was the first experiment to adopt this





	RSS (Native) (GB)	CPU (wrt Native)	RSS (GB)
Muon	0.49	0.984	0.49
Muon	0.49	0.945	0.52
Pion (Barrel)	0.60	0.959	0.55
Pion	0.55	0.953	0.60
Elec (Barrel)	0.51	0.998	0.55
Elec	0.51	0.983	0.50
Minimum	0.59	0.900	0.58
t-tbar	0.64	0.918	0.62



Last Word

- Detector simulation has been living with the HEP community for the past several decades

- The general toolkits, originally designed for High Energy and Nuclear Physics, find their utilities in space science, medical physics as well

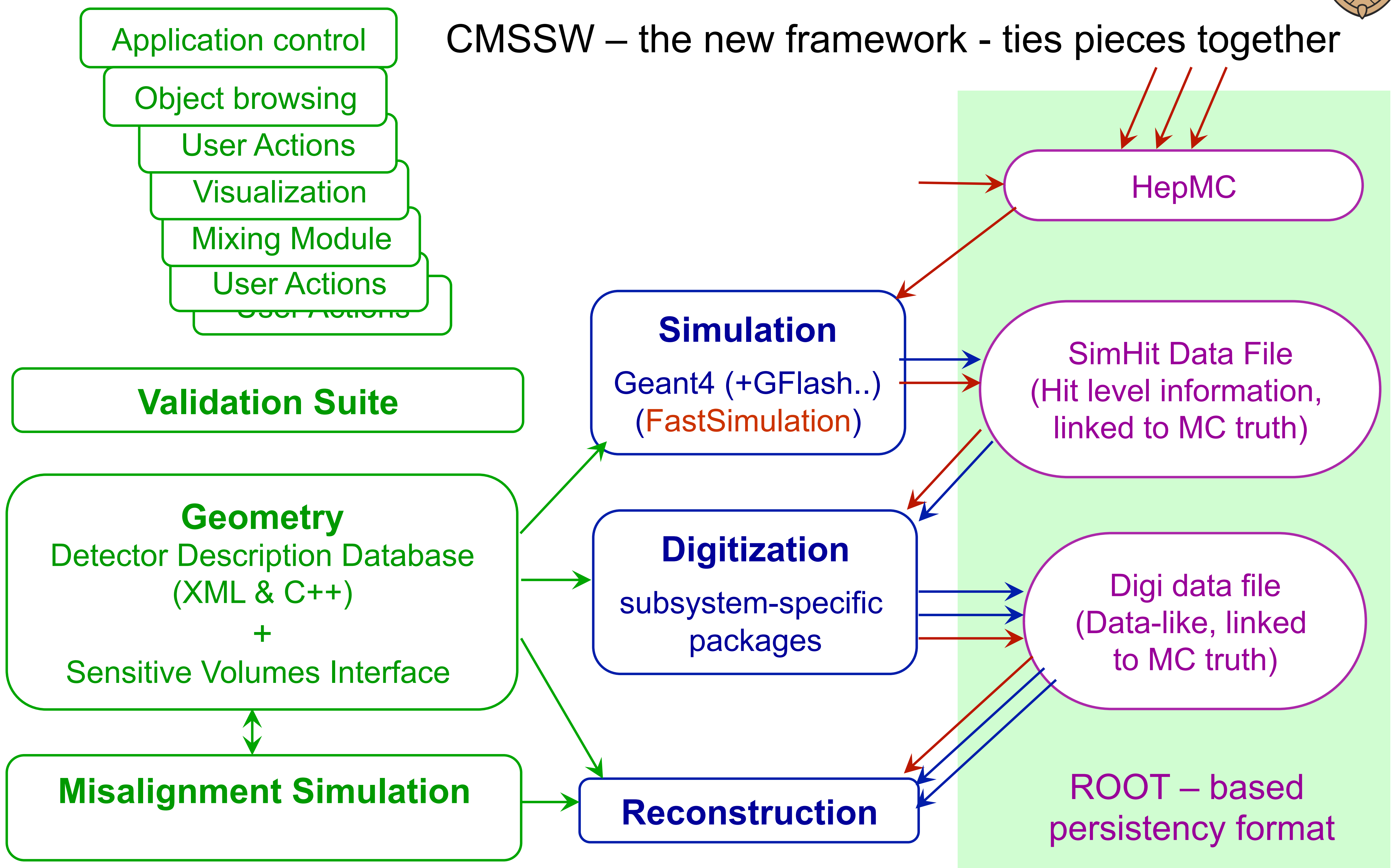
- The HEP experiments use these tools
  - to design an experiment
  - to understand the anomalies observed in the experiment
  - to interpret the data
    - testing the physics models
    - finding evidence of new physics
    - discovering particles yet unseen
    - completing all measurements by helping estimation of the uncertainties

- The field of computation is changing with time. Need new ideas in hardware & software:
  - use of GPU’s, FPGA’s, …
  - use of artificial intelligence

- The tools, as well as the application software, have to keep up with technological advances

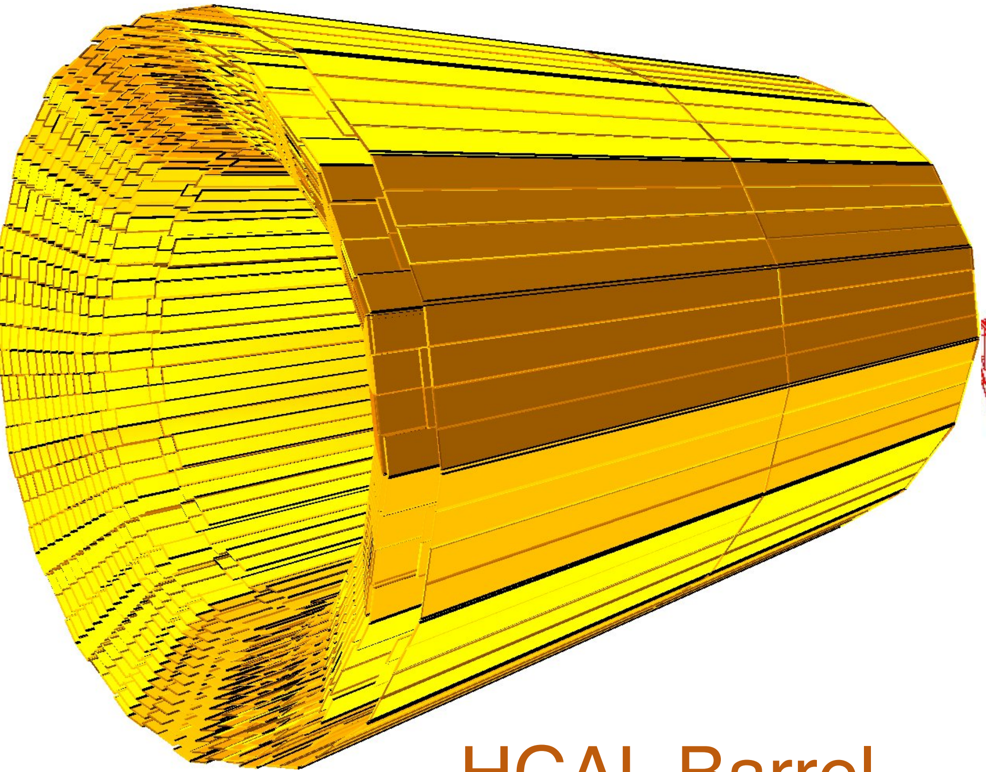
Additional Slides

Simulation Software – CMS Solution

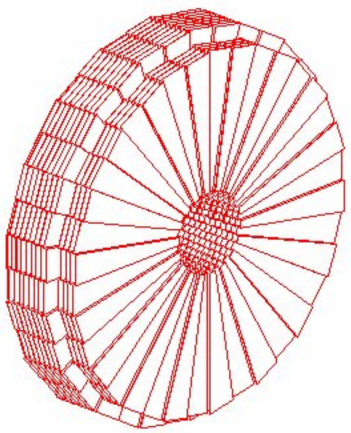


CMS Detector Simulation (V')

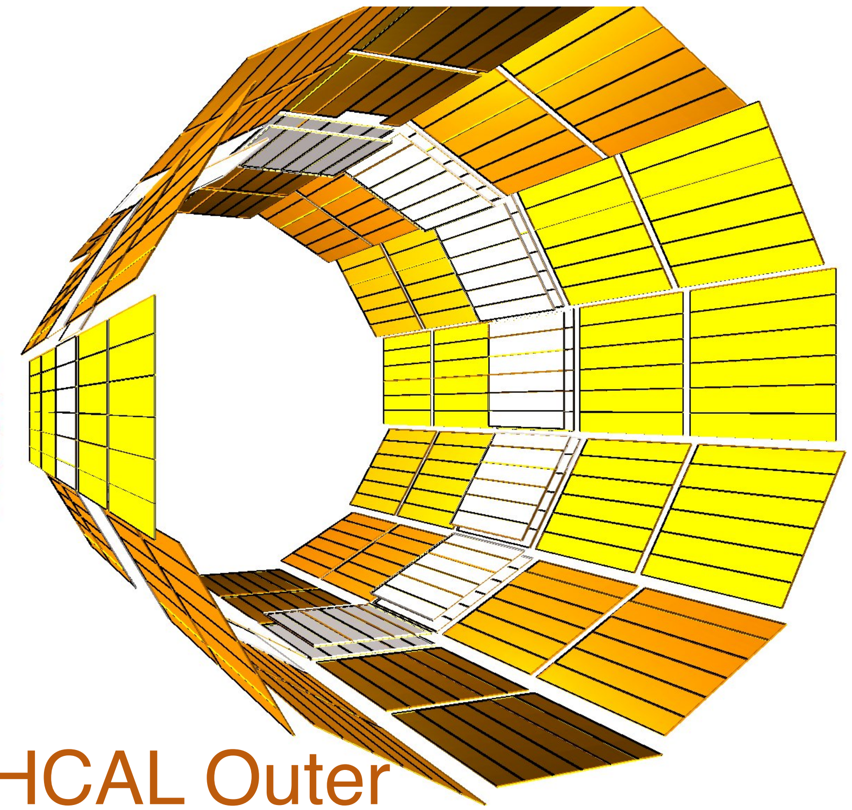
- All other detector components were also described in great detail



HCAL Barrel



HCAL Endcap



HCAL Outer

CASTOR
(QUARTZ)