Detector Simulation in CMS

India
December, 2022

Sunanda Banerjee

Introduction (1)

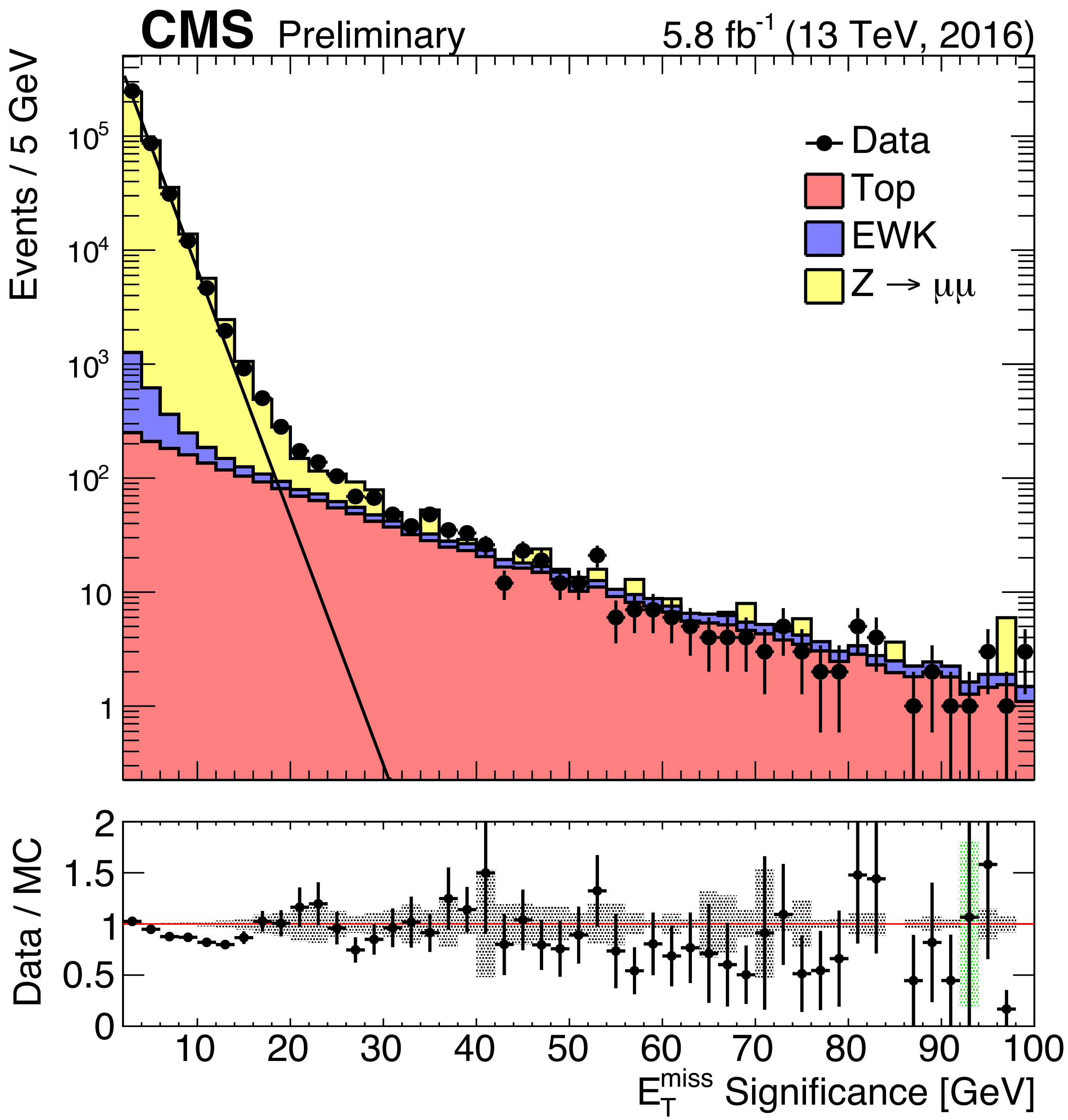
- Monte Carlo samples are used in CMS to
  - design the detector layout
  - develop reconstruction algorithms and trigger logic
  - generate large amounts of signal and background events for use in physics analyses
  - understand/demonstrate analysis procedures and methods based on data to derive calibrations, efficiencies, and resolutions for high-level physics objects
  - derive directly calibrations, efficiencies, and resolutions for high-level physics objects in cases where data are biased or not available

- A data-driven, realistic and accurate Monte Carlo is an essential tool for any high-energy physics experiment

- The simulation effort started in CMS using the toolkit Geant3 more than two decades ago, The current design of the simulation software has evolved through several generations. Two complementary approaches are available today:
  - start from the first principles (Full Simulation)
  - use a fast parametrisation (Fast Simulation)

Focus here on Full Simulation

Introduction (2)



- One needs to know the performance of all the detectors in order to understand missing transverse energy distribution
  - As can be seen from the plot, we understand this in terms of the production of electroweak processes and top production
    - This is a key achievement
- The reliability of MC prediction depends on
  - Goodness of Physics models of the event generator
  - Realistic description of the CMS detector and quality of models used in evaluating propagation of particles through the detector

Software Framework

- CMS started with a software framework written in Fortran77 and used the memory management package ZEBRA. It has now graduated to write most of its code base in C++.

- All application softwares of CMS are built on an Event Data Model (EDM). Simulation or Reconstruction softwares are no exception. They are built like any CMSSW application in the form of special shared libraries called plugins. In practice there is only one command one needs to know to run these applications:

  cmsRun <some-configuration-file>

  configuration-files are written in the python language

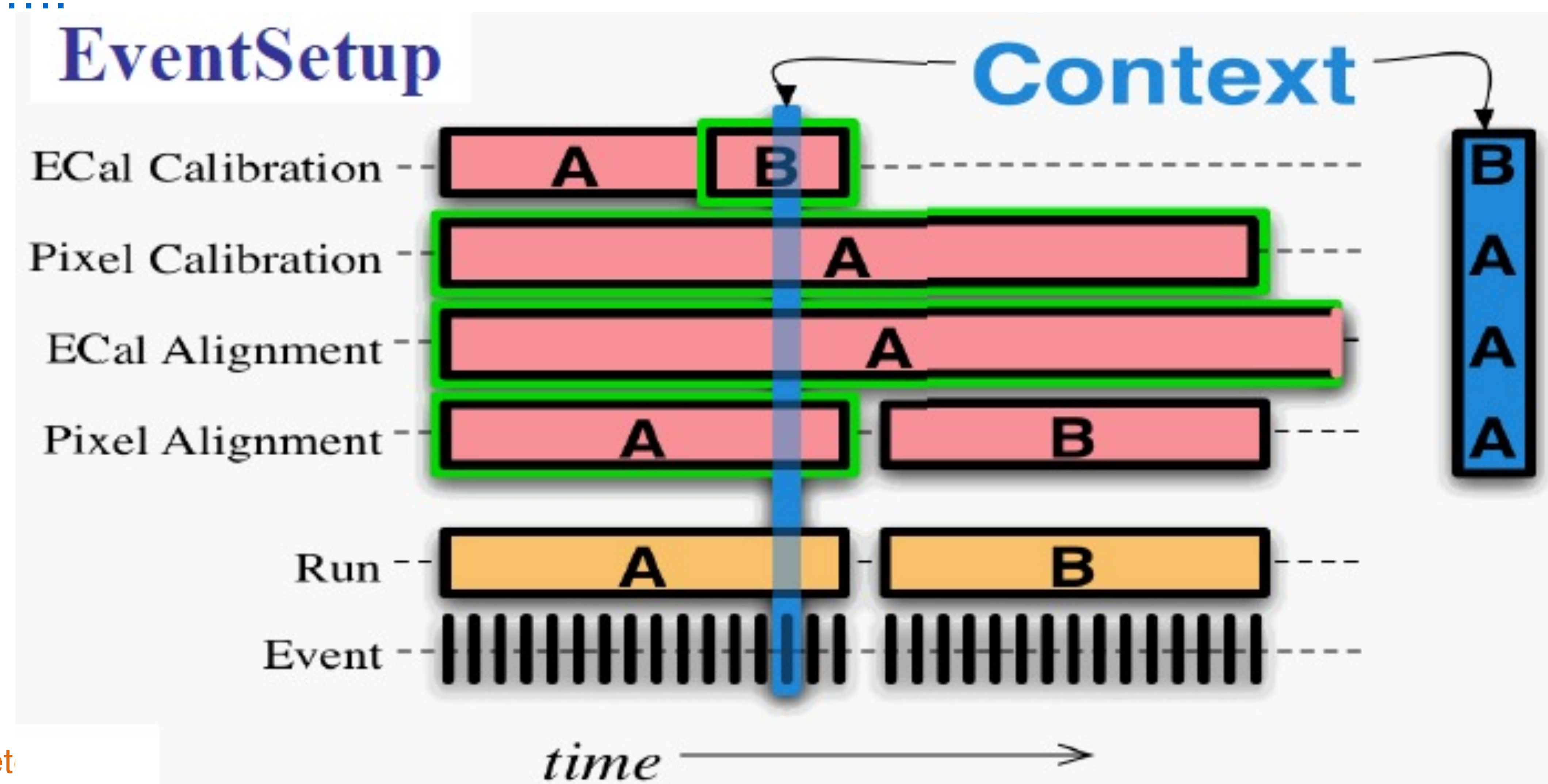
- There are two types of plugins:

  - Module Plugins — EDProducers, EDFilters, EDAnalyzers, EDLoopers, … These are the worker components of the framework.

  - Data Object Plugins — also known as “root dictionaries” because they can also be loaded directly into the “root” application. These are most of the products of the above work, and form the elements of the EDM.

Data Types

- There are two types of data:
  - The Event corresponds to all data belonging to a given bunch crossing in the Large Hadron Collider (LHC)
  - The EventSetup is the system which delivers all non-event data to the module plugins: detector geometry, magnetic field, calibration, alignment, ….



CMS Simulation Program

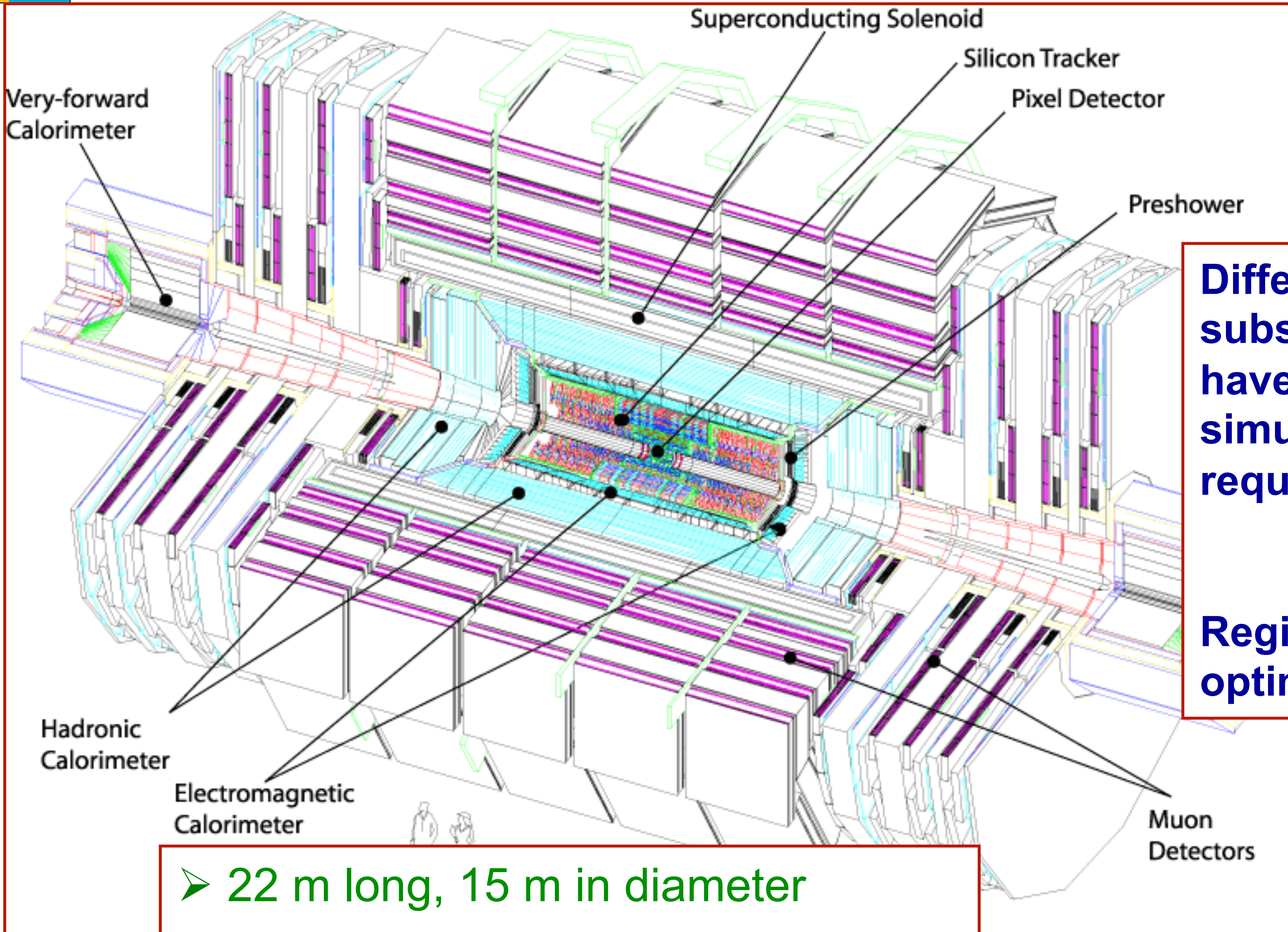
- Though in operation for a number of years, it's a live system – goals, requirements, tools evolve throughout the lifetime of the experiment

- CMS simulation program started with the Geant3 toolkit and is now based on the Geant4 toolkit. Geant4 provides the following functionalities:
  - physics processes: electro-magnetic and hadronic interactions.
  - tools for detector geometry and sensitive element response.
  - interfaces for tuning and monitoring particle tracking

- In addition CMS offline framework and Event Data Model:
  - manage application control at run time
  - rely on the concept of event processing module (EDProducer)
  - provide interface to common tools (generators, magnetic field, MC truth handling, infrastructure for hits, event mixing, digitization, …)
  - ensure provenance tracking and event immutability

- Emphasis given to simulation program:
  - agreement with the data through most reliable physics models
  - robustness, performant and adapted to modern technology
  - Extension to describe newer detectors (changes in current design and an eye for the future)

CMS Detector



Different subsystems have different simulation requirements

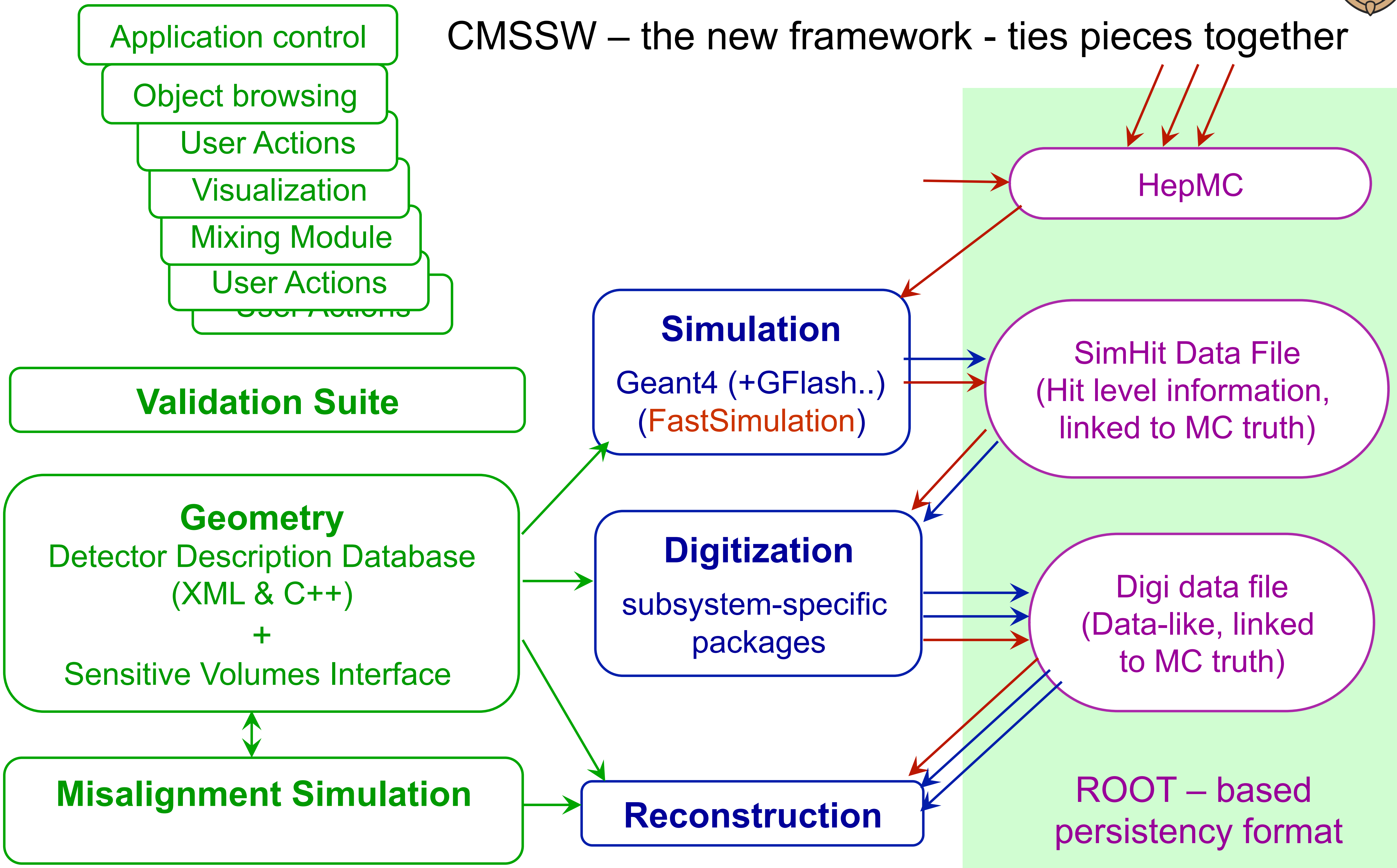
⇓

Region based optimization

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

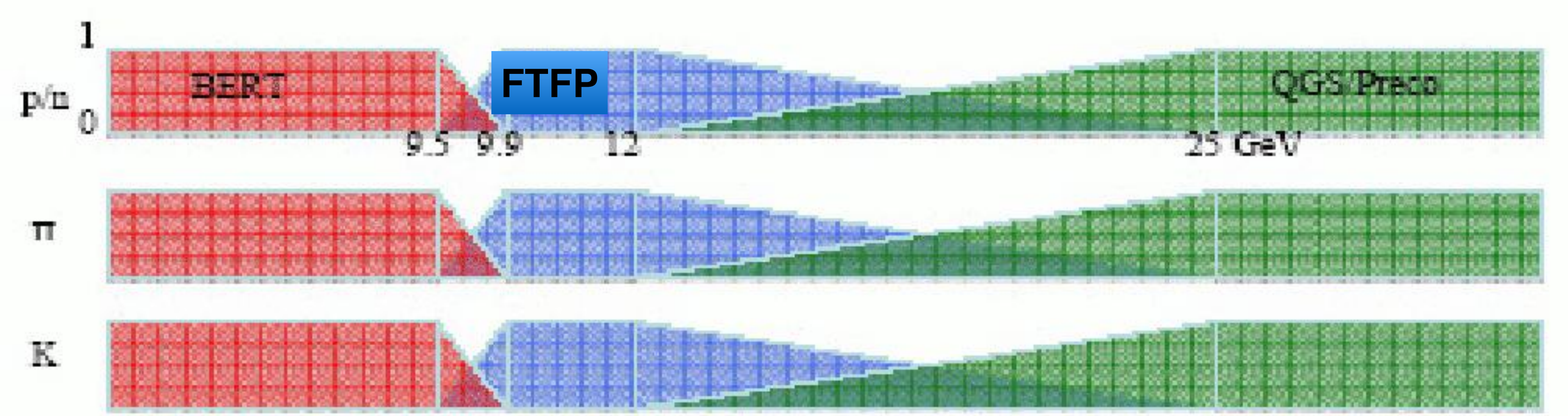
S. Banerjee

Simulation Software – CMS Solution



Interface with Geant4

- The detector geometry available through the EventSetup;is converted to Geant4 geometry
- Sensitive detectors get associated with geometrical volumes and defined through the XML configuration files
- Magnetic field is based on a dedicated geometry of magnetic volumes; and is provided by independent subsystem via EventSetup. Field selection, propagation tuning etc. are configurable at run time
- Variety of lists (QGSP_BERT, QGSP_FTFP_BERT, FTFP_BERT,…) for modelling physics processes exist and one is selected at run-time with appropriate production cuts and activation/tailoring of individual processes;
- Variety of Physics event generators (particle guns, Pythia, Herwig,…) provide generator information in HepMC format and are interfaced to the Event
- User actions allow access to Geant4 objects at any stage (run, event, track, step); used for tuning, diagnostics, custom bookkeeping
- Monte Carlo truth record with decay/interaction history of the generator's particles and selected tracks from Geant4 simulation



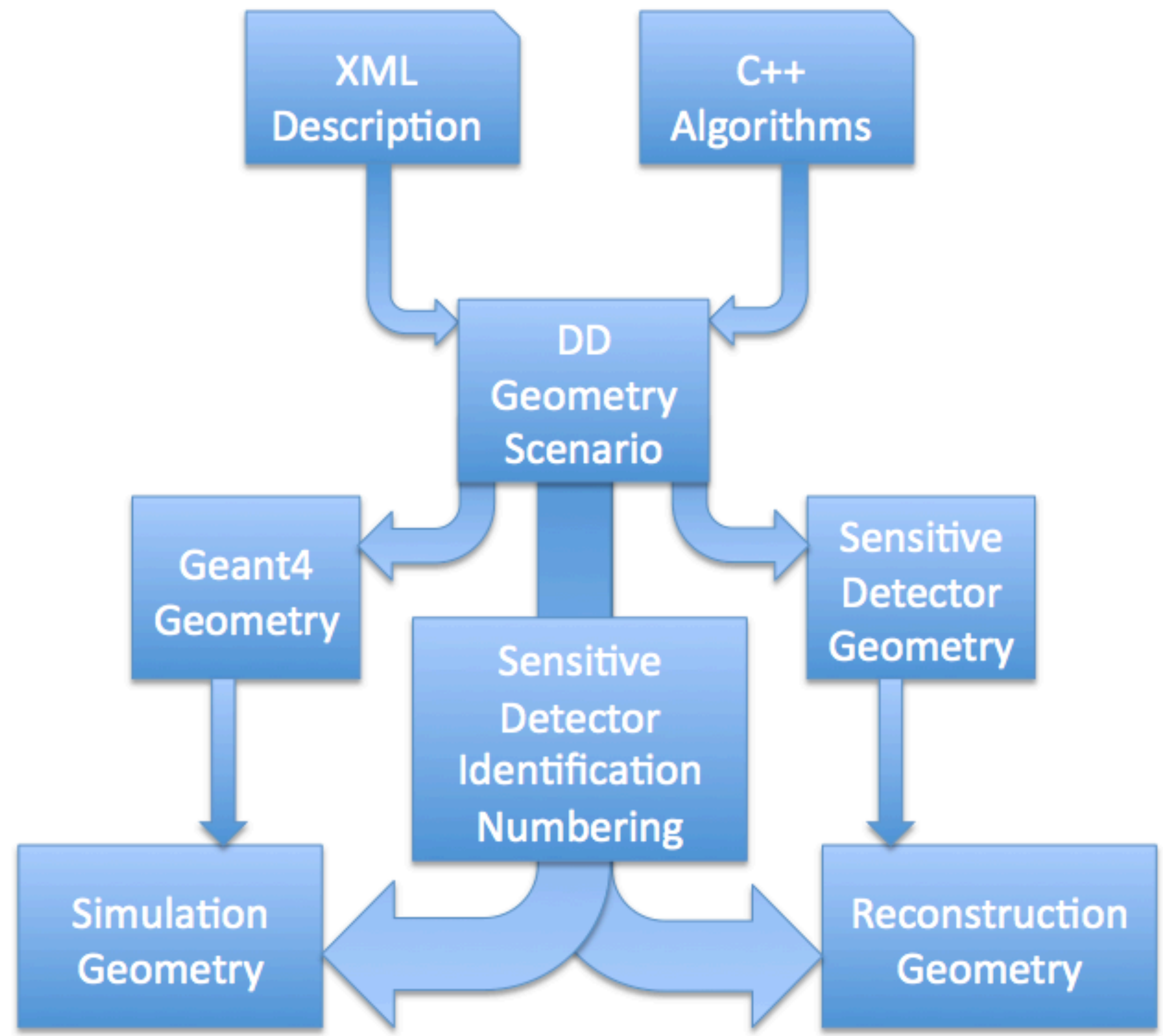
Physics Lists in CMS

- CMS used the physics list in the past for its Monte Carlo production
  - QGSP_FTFP_BERT_EML (with Geant4 versions 9.4.p02, 9.6.p02)
- CMS moved to multithreading mode from the beginning of Run2 (2015)
  - QGSP_FTFP_BERT_EML (with Geant4 version 10.0.p02)
- CMS moved to a new physics list for its 2017 MC production
  - FTFP_BERT_EMM (with Geant4 version 10.2.p02)

  and again moved to a new Geant4 version (10.4) for 2018 production, (10.7) for final Run2 analysis and (11.1) for Run3 applications
- FTFP_BERT is the recommended physics list from Geant4 collaboration (J.Allison *et al*. NIM A506, 2003, 250; NIM A835, 2016,186)
- The list QGSP_FTFP_BERT combines QGSP, FTFP, Bertini Cascade models for π/K/p/n with a fixed validity region:
  - Bertini Cascade valid at ≤ 8 GeV
  - FTFP valid between 6 and 25 GeV
  - QGSP valid at ≥ 12 GeV
- The list FTFP_BERT uses FTFP and Bertini Cascade models:
  - Bertini Cascade valid at ≤ 5 GeV
  - FTFP valid at ≥ 4 GeV
- EML, EMM specify the physics models for electromagnetic processes
  - EML utilizes a simplified multiple-scattering model for all detectors
  - EMM uses the detailed multiple scattering model for HCAL and HGCAL and the simplified one for other detectors (handling of multiple scattering is critical for sampling calorimeter)

Source of Geometry

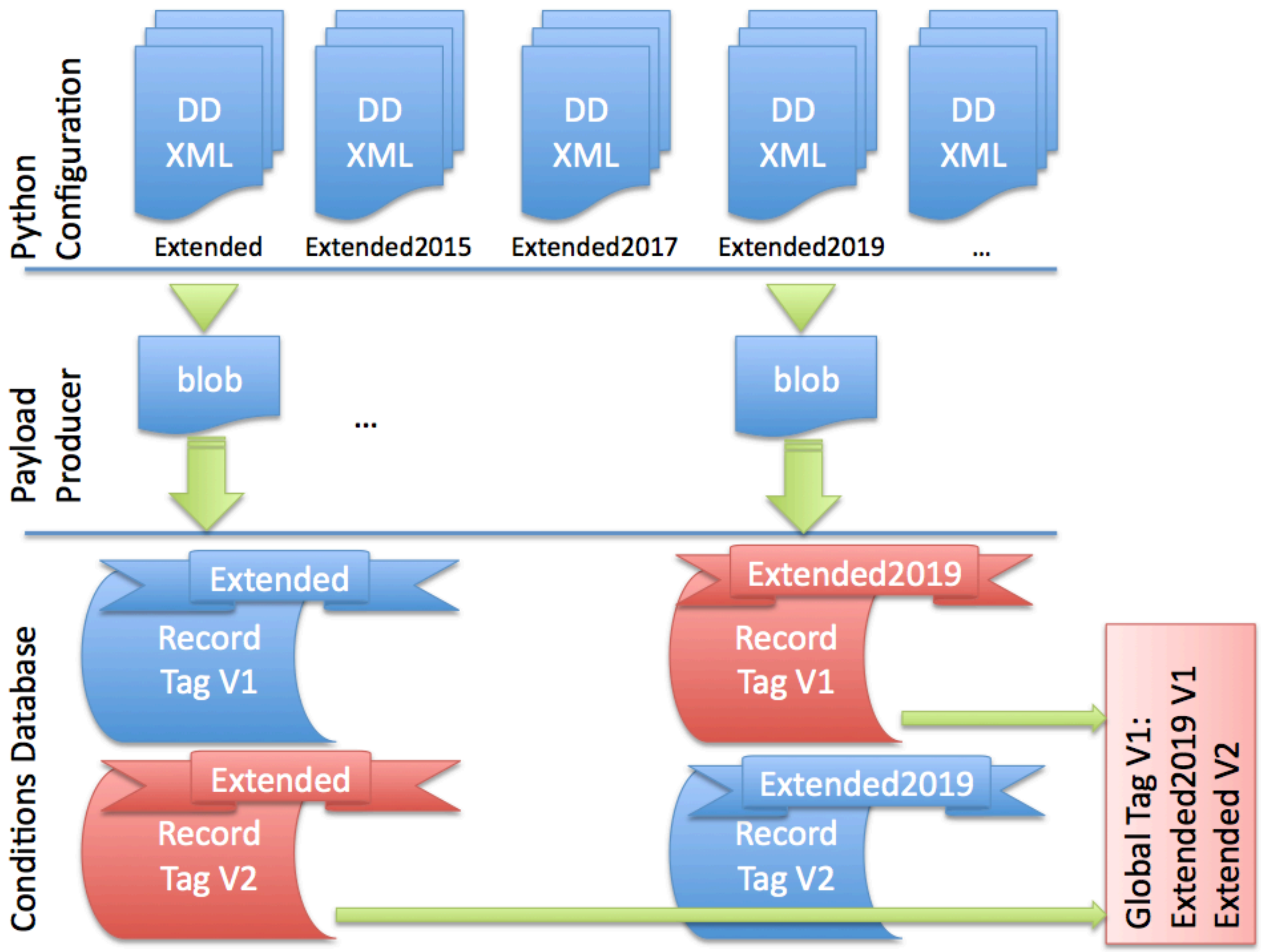
- CMS geometry has been defined through a Detector Description Language (DDL)
  - contains all static information: geometry, material, sub-detector specific constants
  - accessed by a single interface in all applications: simulation, reconstruction, analysis
  - realized through parameters defined using XML description and some specific C++ algorithms which utilizes some symmetry properties in hierarchical positioning
  - also used for CMS magnetic field geometry



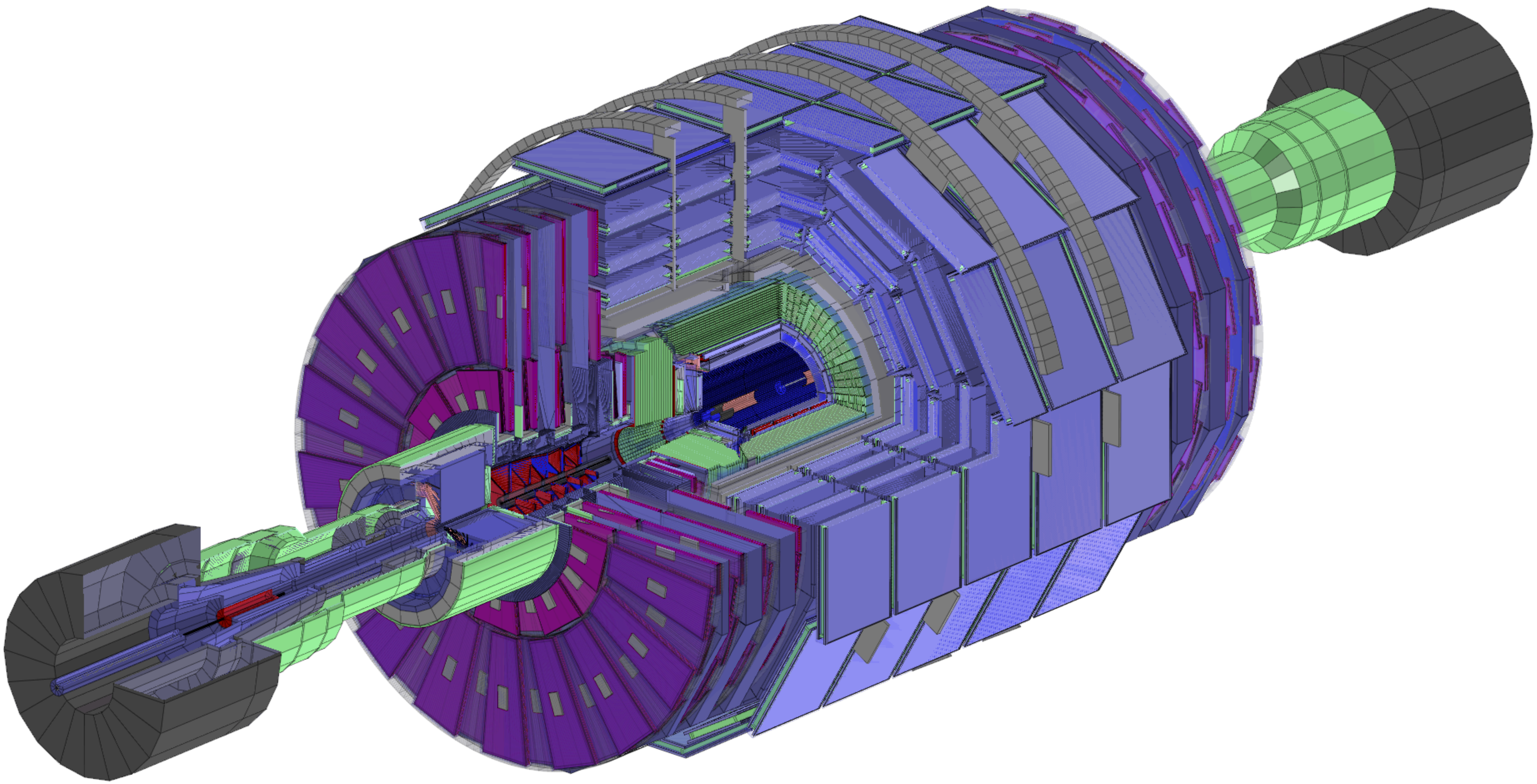
- CMS is in the process of replacing home-made DDL with CERN-IT supported software DD4Hep.

Views of Geometry

- A set of XML files which includes the parameters as well as reference of the C++ algorithms define a scenario
- Several such scenarios exist and each of them are put in the condition database as a blob
- Application programs can access either of the two sources
  - Production jobs always access the geometry files in the condition data base
  - XML files are used for the development of new geometries



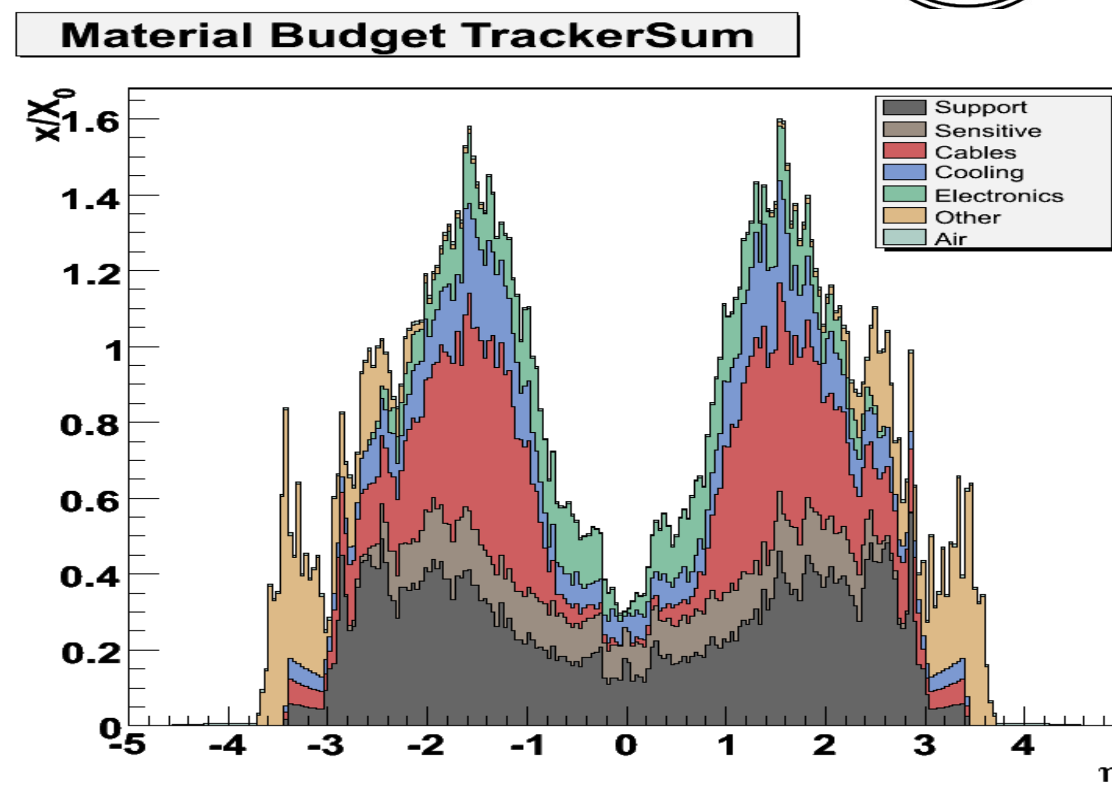
CMS Geometry

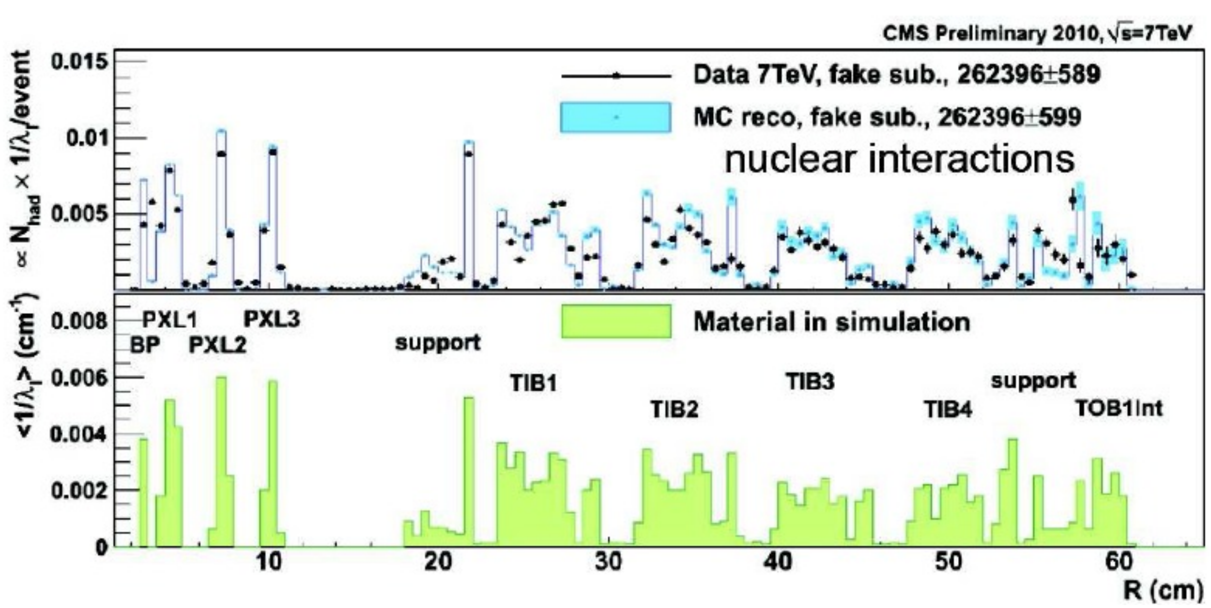


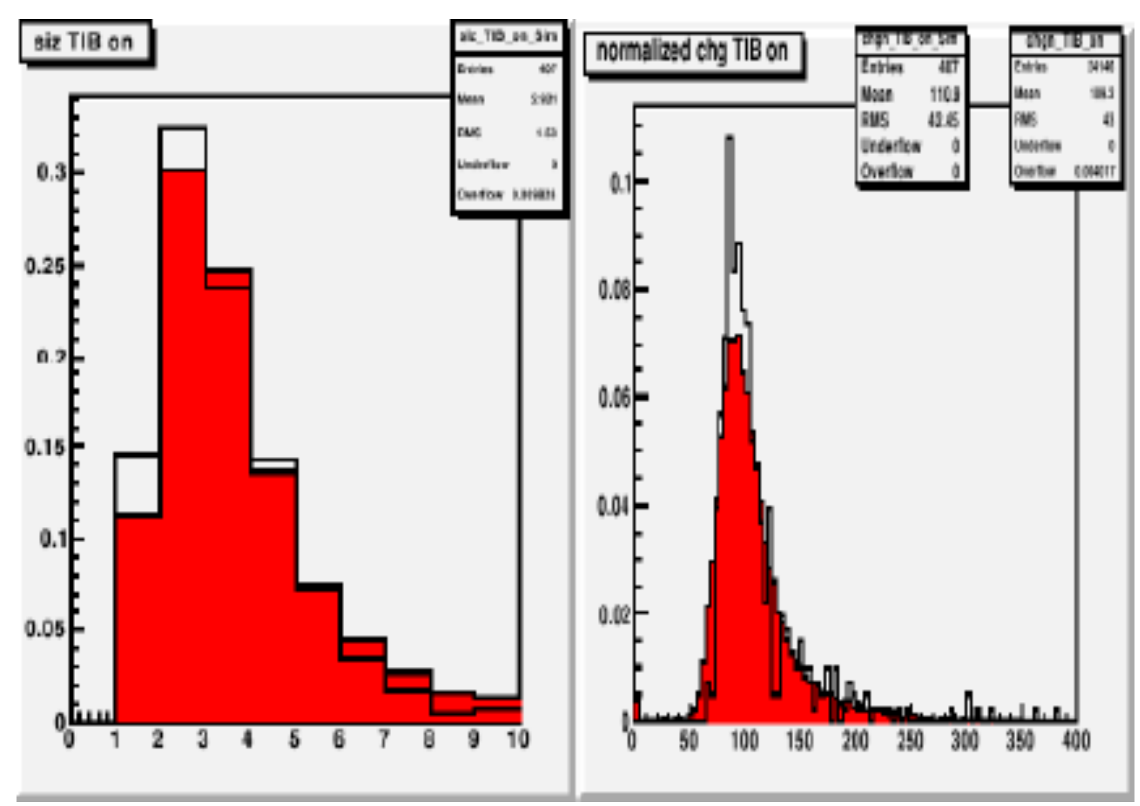
- Emphasis has always been to define sensitive part of the detector (and its proximities) in great detail
- For tracking detectors all heavy spots are defined accurately and the rest using average material definition

Tracker

- Demands a high degree of accuracy:
  - In the description of both active and passive components
    - review each component with full information from integration centres
    - verify by weighing all individual components and match them with the description in the geometry
  - Correct navigable Monte Carlo truth
  - Proper treatment of hard electron bremsstrahlung
- Final verifications:
  - Total weight of the tracker before insertion
  - Radiography using collision data (positions for γ-conversion and nuclear interaction)





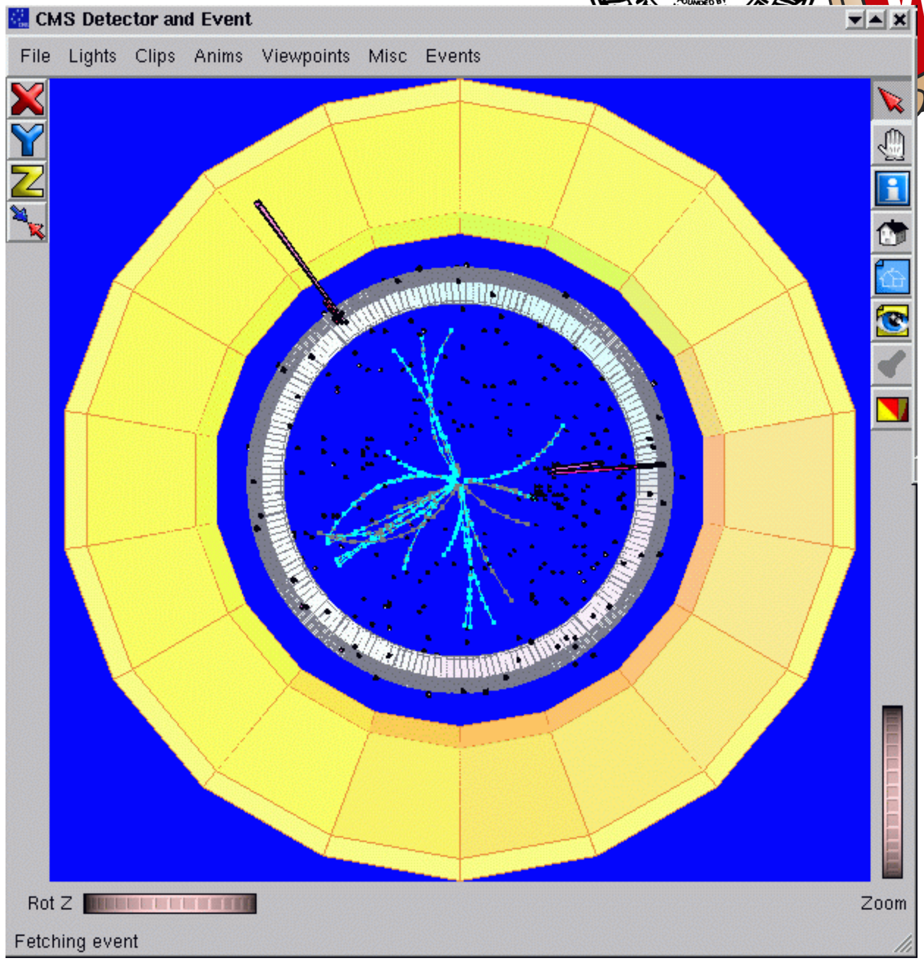


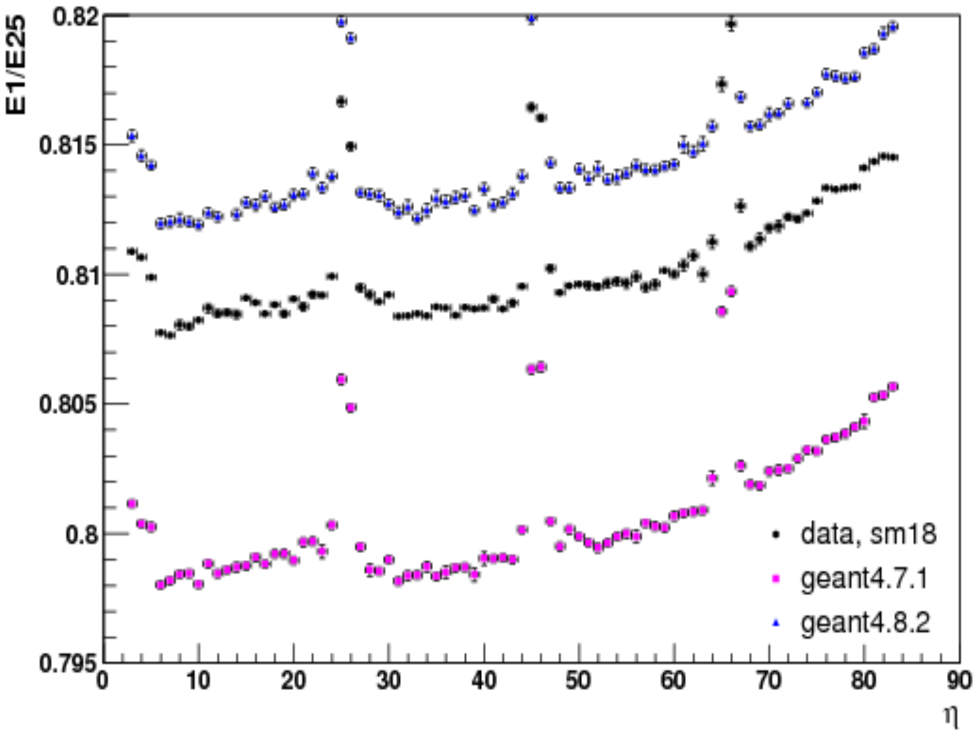
Electromagnetic Detector

- Accurate description of geometry and also material budget:
  - independent alignment of modules, super-crystals, wafers, ….
  - updated distribution of support, cooling and readout system

- Good/complete implementation of all physics processes to reproduce
  - transverse shower profile (containment, calibrations)
  - longitudinal shower profile (leakage, …)

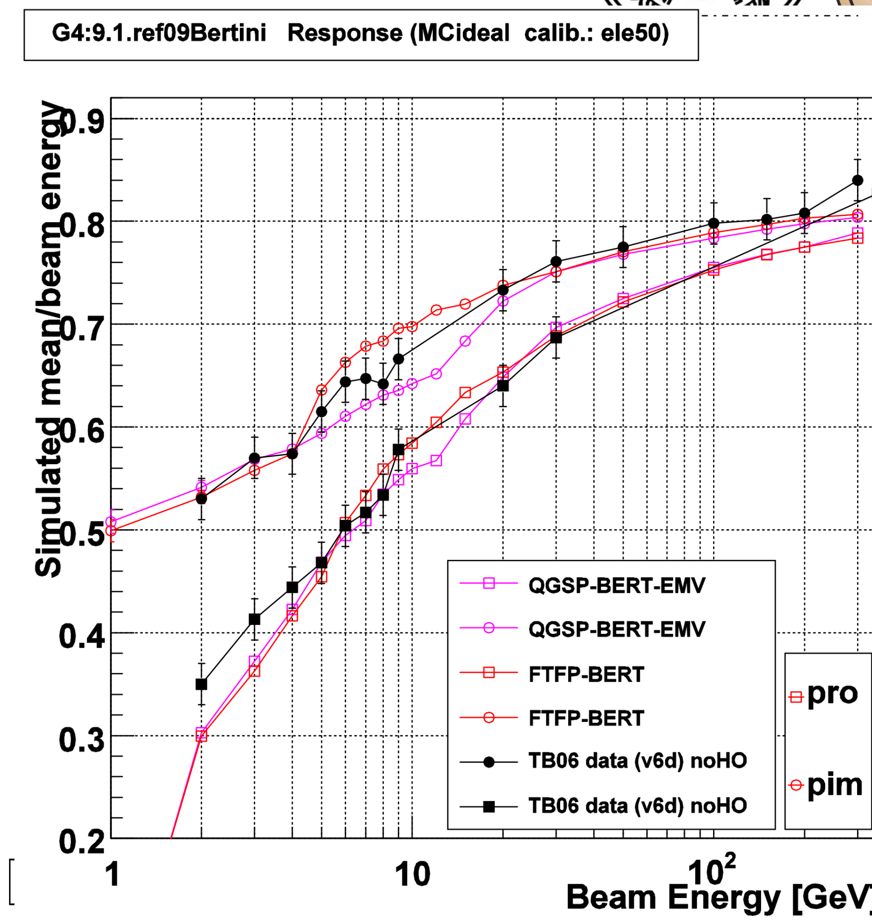
- Validated extensively with test beam (at the CERN H4 beam line) for energy measurement and transverse shower profiles.

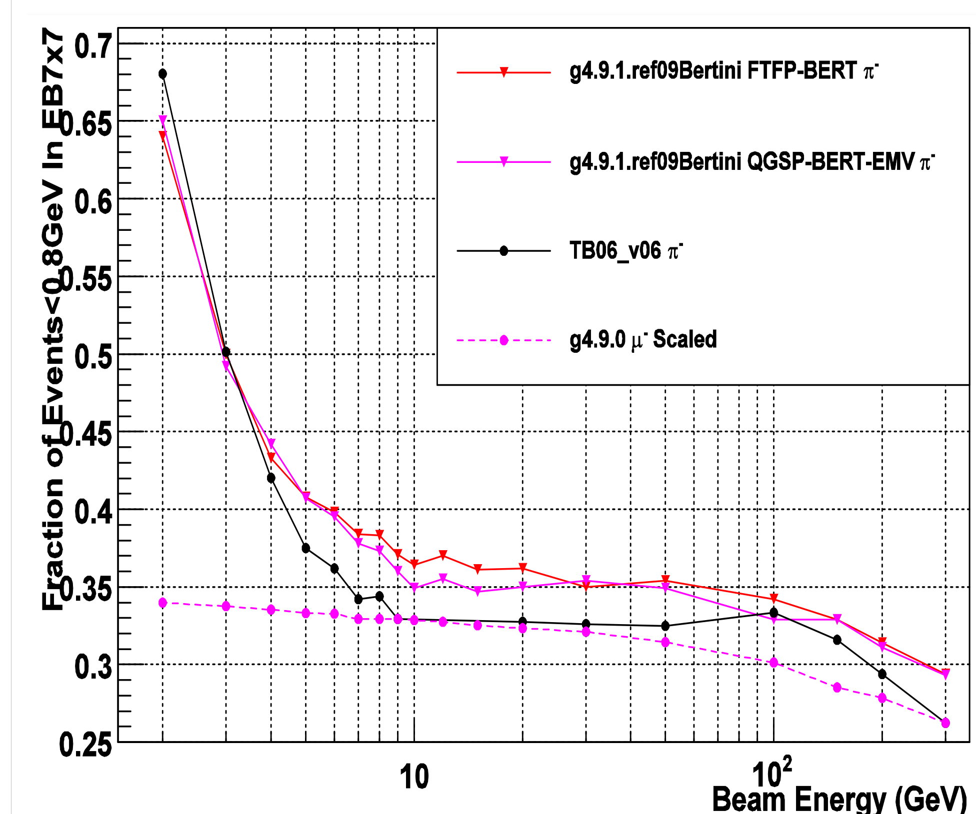




Hadron Calorimeter

- Reproduce accurately the measurements from test beam (at the CERN H2 beam line) done during 2002–2010 with different HCAL modules, preceded by real ECAL super module (super crystals) or their prototypes. Beams of π, e and μ were used over a large energy range.
- Studies on energy resolution, linearity in response, e/π ratio and shower profiles were instrumental in validating Geant4 hadronic physics models [parametric (LHEP), and microscopic (QGSP_FTFP_BERT, FTFP_BERT, …)]
- Faithful description of timings, noise, …
- Use of shower libraries, noise libraries, … wherever applicable



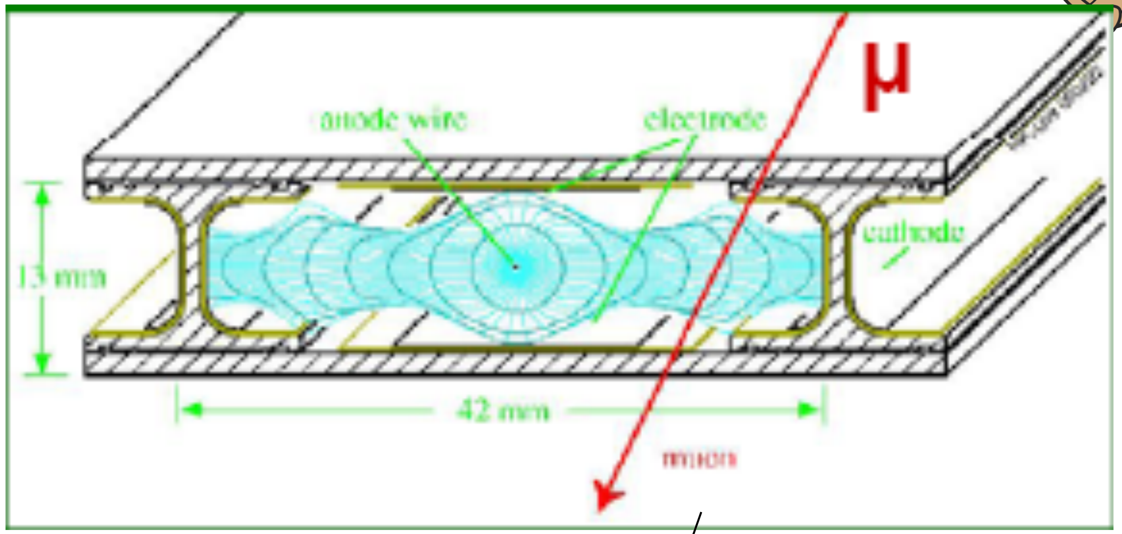


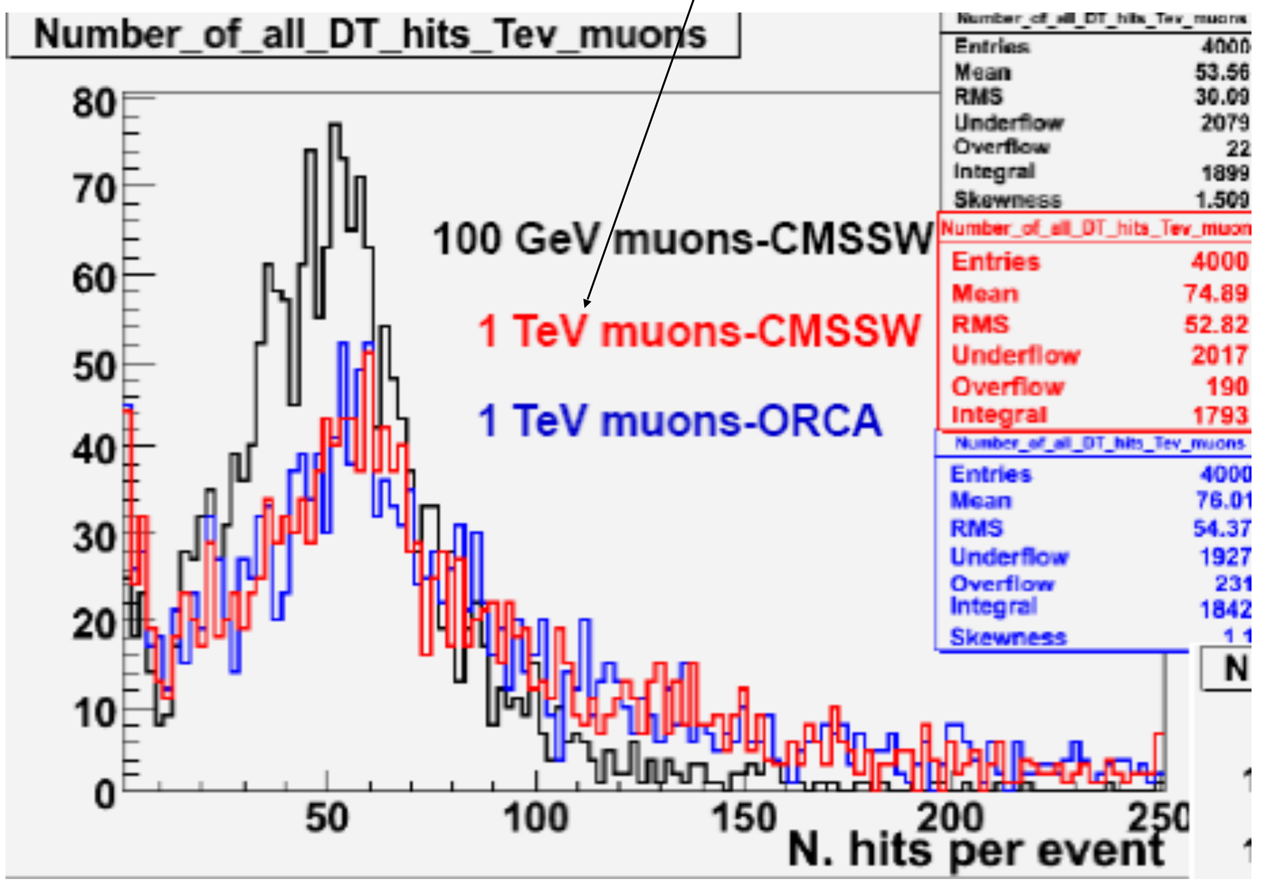
Muon System

- Geometry description of the barrel (drift tube chamber) and endcap (cathode-strip chamber) detectors were verified using the Cosmic data collected during MTCC, CRUZET, CRAFT, …

- Muon physics in Geant4 is extensively tested and validated in the energy range 10 GeV — 10 TeV
  - improved description of muon bremsstrahlung, μ-nuclear effects, …
  - better description of multiple scattering (in agreement with the data)

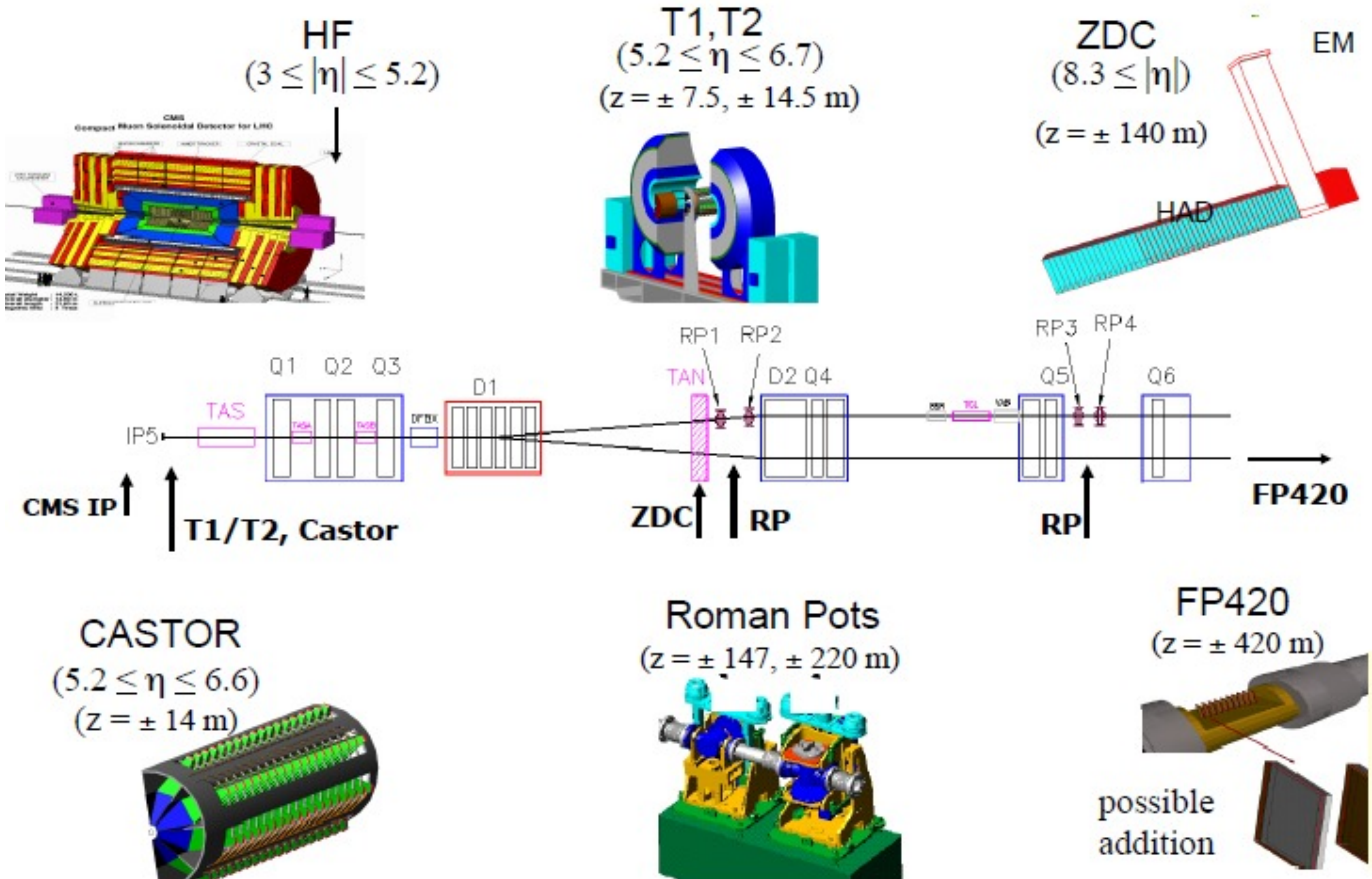
- Validate new descriptions with earlier simulation and with test data





Forward Detectors

- Beam pipe, shielding and forward detectors were described in detail with all technical knowledge available



Forward Detector Simulation

- Essential for diffractive and heavy ion physics

- Simulations of stand-alone systems have been compared with test beam studies regarding energy resolution, leakage, …

- Simulation of the central as well as very forward detectors (ZDC, Roman pots, FP420) is foreseen:
  - use a filter to separate particles from event generators to be processed through the central and very forward detectors
  - use a separate transport code Hector to transport particles within acceptance of the forward detectors close to the forward detectors
  - also obtain beam interactions from a library obtained using a separate simulation code with MARS
  - transport the particles in the central detector and also in the forward detector region using Geant4
  - combine all simulated hits to get information of a complete event

Run 3 Geometry

- Utilizes 3856 solids in 4173 Logical Volumes among 2.3 million touchables.
- Summary of solids used in touchables:

	Standard	Reflected
Box	1208k	434k
Tube	94k	1391
Trapezoid	240k	150k
Cone	1862	0
Polycone	426	32
Polyhedra	1449	0
Torus	128	32
UnionSolid	174k	0
SubtractionSolid	8289	468

- Overlap summary (for tolerance of 0.1 mm):

	Tracker	ECAL	PreSh.	HCAL	Muon	Infra.St	Total
DDD	0	0	4	0	0	0	4
DD4Hep	0	0	4	0	0	0	4

Task of the Digitizer

- Convert the energy deposit record in the sensitive detectors (tracking type and calorimetric type) to signals as expected from the real detectors
  - Two distinct types of actions required for the two detector types
  - Also hit information are different for tracking type and calorimetric type detectors
    - For calorimetric detectors, information exist for all individual readout unit in finite time slices (~1 ns). However the energy deposits are stored as cumulative total if the source is the same particle or one of its shower product.
    - For tracking detector, every crossing with the sensitive detector is stored as a hit and the sensitive detector does not have finer most cell identification (e.g. wafer rather than strip or pixel). Here store all information like entry, exit point, energy deposit, timing, ….
- Include the effect of pile-ups which could be in time (multiple interactions during the bunch crossing of interest) or out of time (due to interactions from previous or later bunch crossings which will interfere because of finite response time of the electronics)
  - These effects are common to all detector types and are handled using common interface

Digitization of Tracking Detectors

- The digitizer tries to take care of the following effects
  - Convert energy deposit to charge taking care of electrons produced, its fluctuation, drift in electric field (effect of Lorentz angle), …
  - Charge diffusion which may cause signal in multiple readout units
  - Smearing of the charge
  - Addition of noise
  - Take care of several effects
    - non-linearities and thresholds
    - miscalibration
    - noisy read out cells
    - inefficiencies and dead cells
    - saturation
    - aging and radiation damage
    - pulse shape
- Final output has digital information for every readout channel (with appropriate zero suppression algorithm)
  - a unique cell identifier
  - ADC and/or TDC information
- An external object to link the digit to simulation hit

Digitization of Calorimetric Detectors

- The digitizer tries to take care of the following effects (all calorimeters use either scintillation or Cernekov photons as primary signal)
  - Converts hit energy to photo electrons
  - Do photo statistics
  - Convert PE's to analog signal
  - Smear the charge
  - Add noise (due to photo-transducers)
  - Take care of non-linearity effects of the photo-transducers
  - Simulate the electronics taking care of
    - pulse shape
    - electronic noise
    - time slew effects
    - noisy or dead cells
    - radiation damage
- Again store digital information for every readout channel (with appropriate readout option)
  - Unique identifier of the readout channel
  - ADC (and TDC) information

Mixing Module

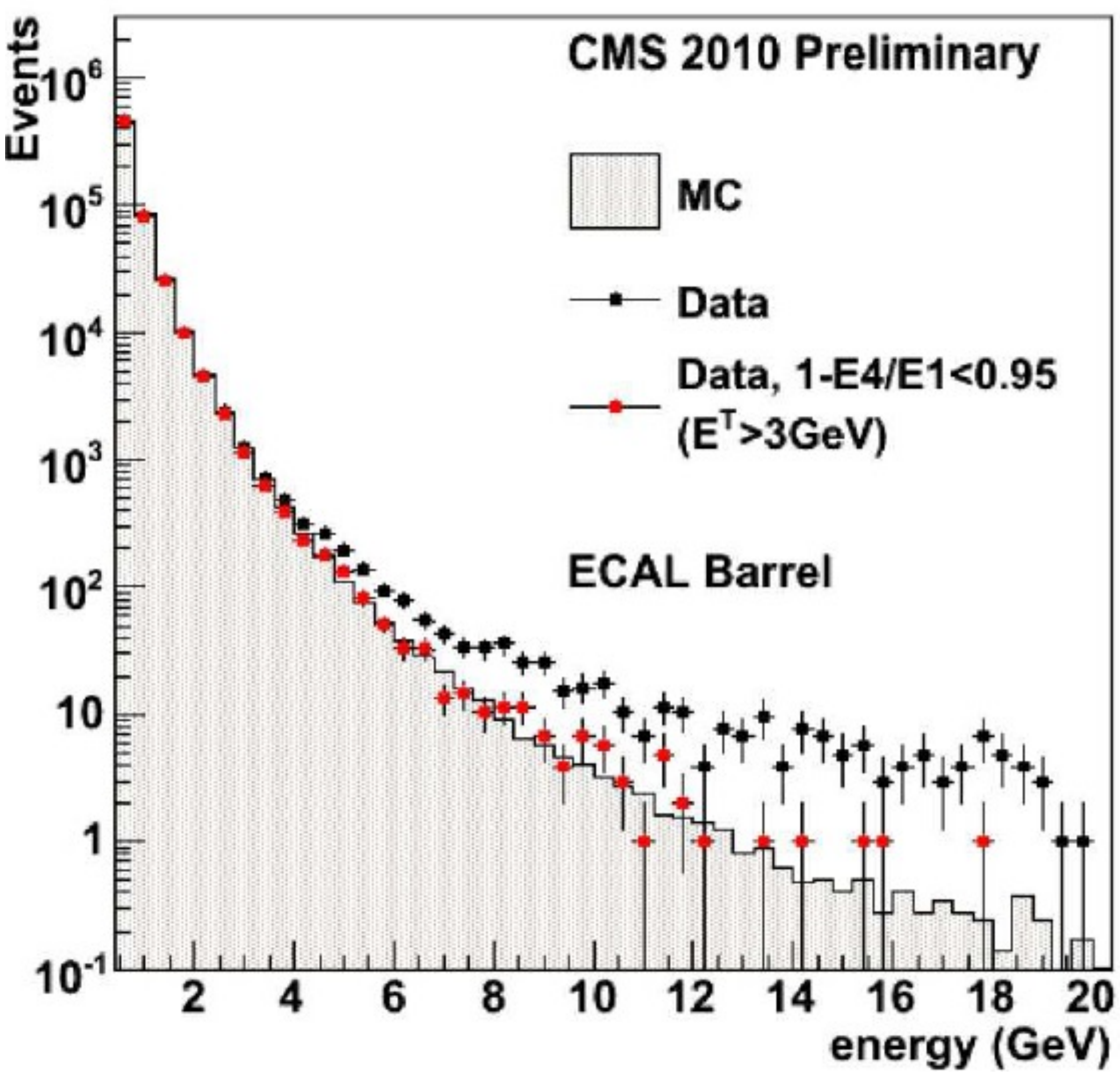
- The task of the mixing module is to add a given number of minimum bias events to the signal event to mimic in-time and out-of-time pile-up effects
- To get a coherent software scenario, the digitization process has to happen after the mixing is completed → mix the SimHit information
  - For high luminosity operation, an average number of in-time pile-ups is rather large and for many detectors (barrel muon in particular) one needs to consider a large number of bunch crossings
  - The mixing scenario was revised even during Run 1 operation to optimize the mixing stage
    - Each detector is called once for each event to be added to accumulate the hits
    - A final call is made when for the event the accumulation is completed and the digitization can be made
  - For high pileup runs a new approach is made using “premixed” events
    - A digitized sample for a certain running period is made with a given pileup scenario from a set of minimum bias events including in-time and out-of-time scenario
    - This “premixed sample” is used in the mixing module
    - The raw format of CMS hits is extended to ensure sufficient precision for making sums of small pulse heights in the Digi step

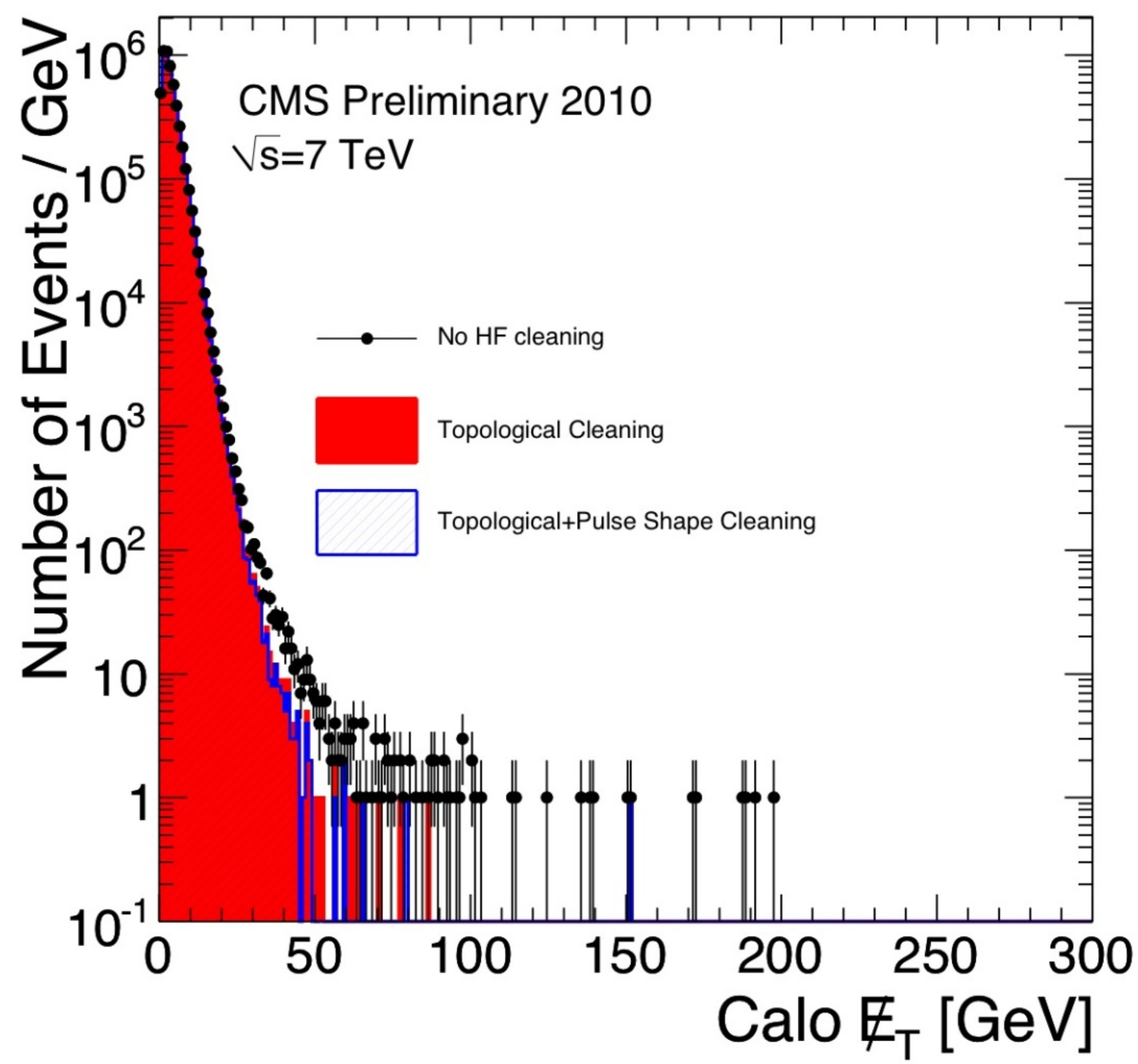
Data Mixer

- Instead of mixing a large number of minimum bias simulated events with a signal MC event, events from real collision is considered for the mixing process
- A random collision event represents the PU conditions of the machine at a given time. It includes properties of boys in-time and out-of-time PU interactions.
- CMS collects zero bias trigger events at the rate of 1 Hz and can be increased to 10 Hz if required
- To make appropriate usage of this one need to worry about
  - zero suppression effect in the detectors
  - mis-alignment in the detectors
    - simulation uses perfectly aligned detectors and effect of mis-alignment need to be taken care of by mapping each misaligned cell to a perfectly aligned cell
      - may lead to some edge effect in certain detectors
- The data mixer approach is tried for calorimetric objects
  - relative variation of this approach w.r.t. full simulation approach (without pre-mixed events) is significantly smaller (by a factor 5) than jet energy resolution

Use of Detector Simulation (I)

- 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

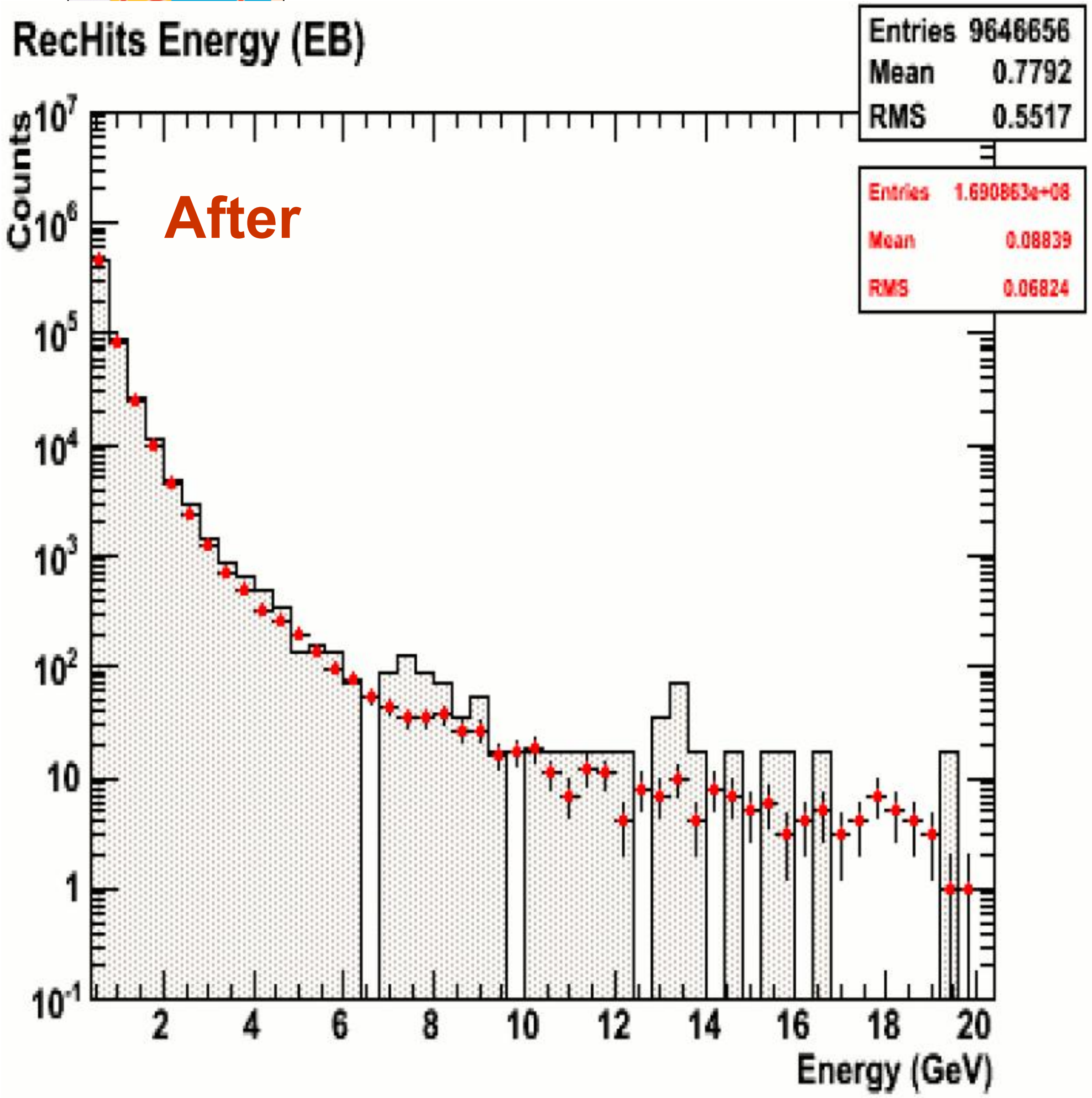


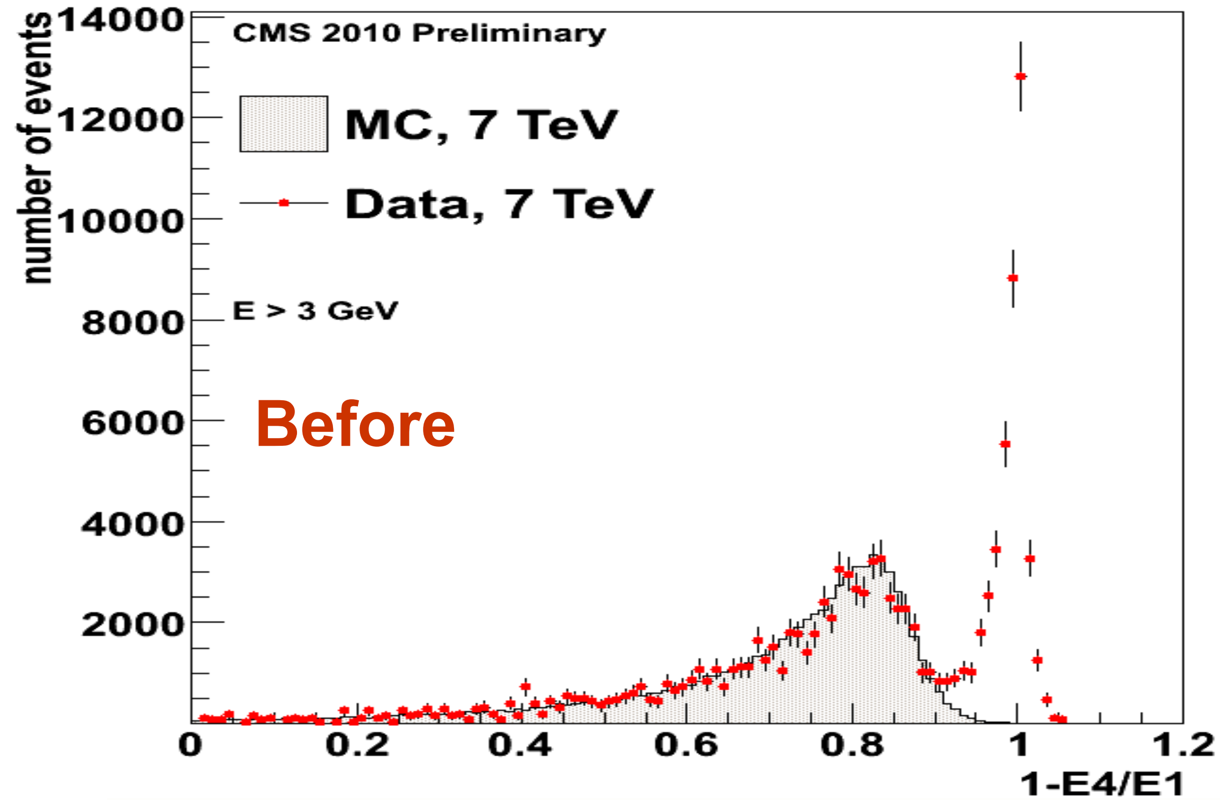


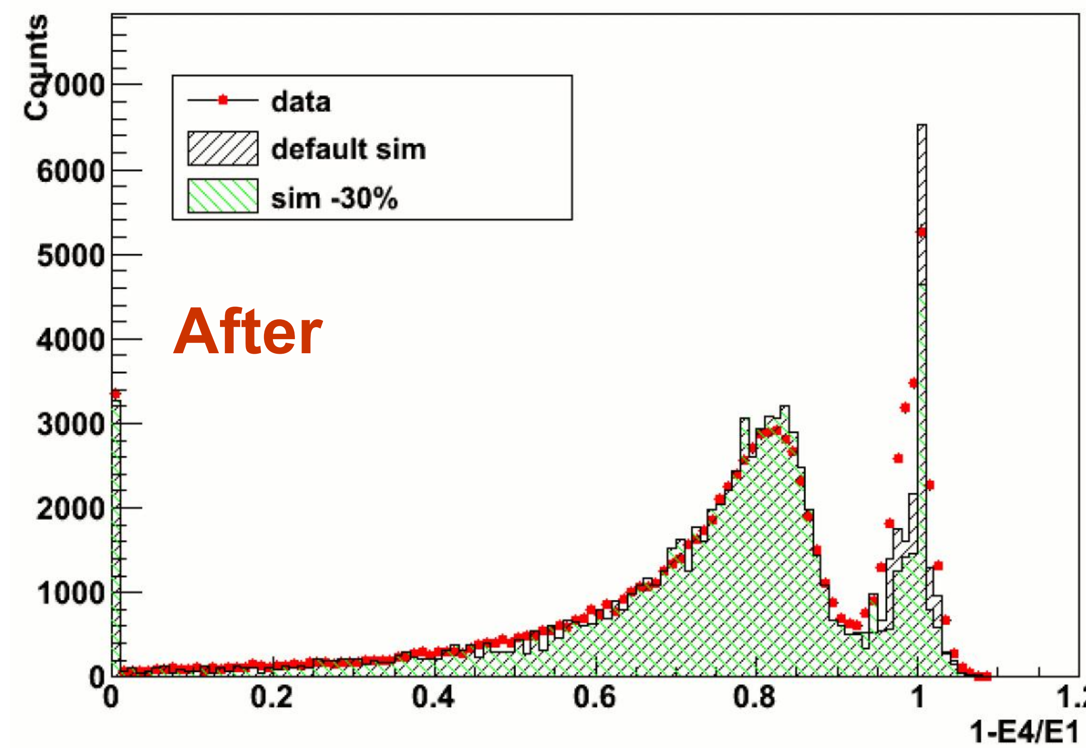
Use of Detector Simulation (II)

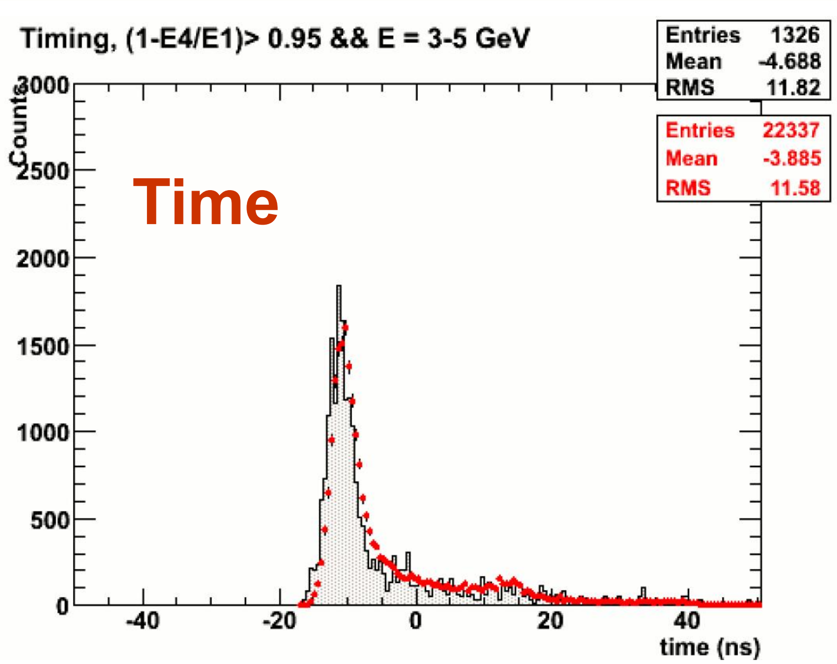
- Anomalously large signals were observed in the ECAL with the appearance of very large energy deposits in a single crystal
- These events are uniformly distributed in the barrel part where the readout utilizes APD. They are not seen in the endcap crystals which are readout using VPTs
- The rise time 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

Use of Detector Simulation (III)





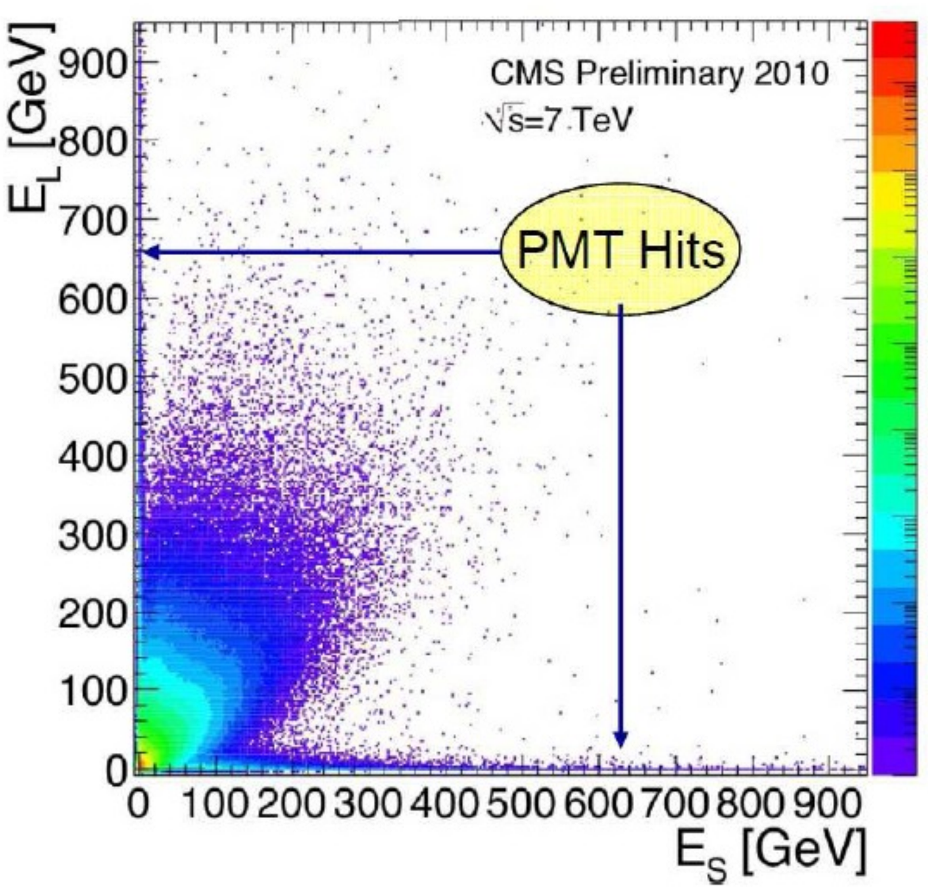


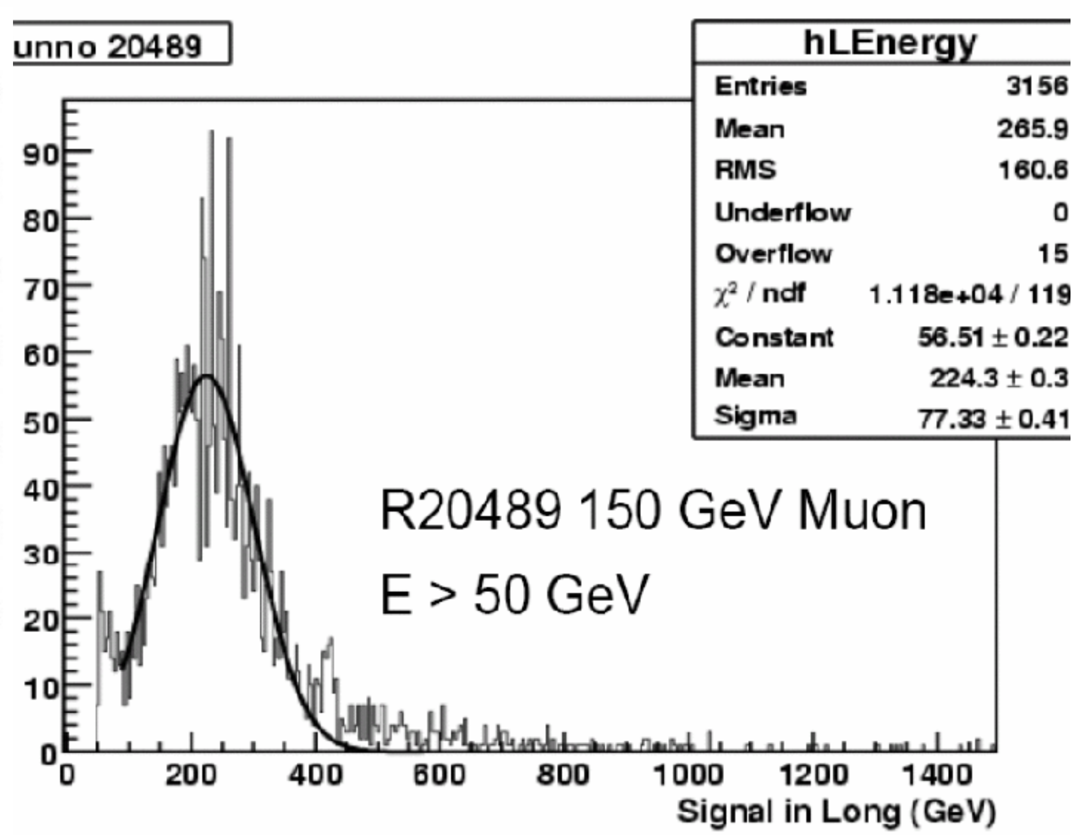


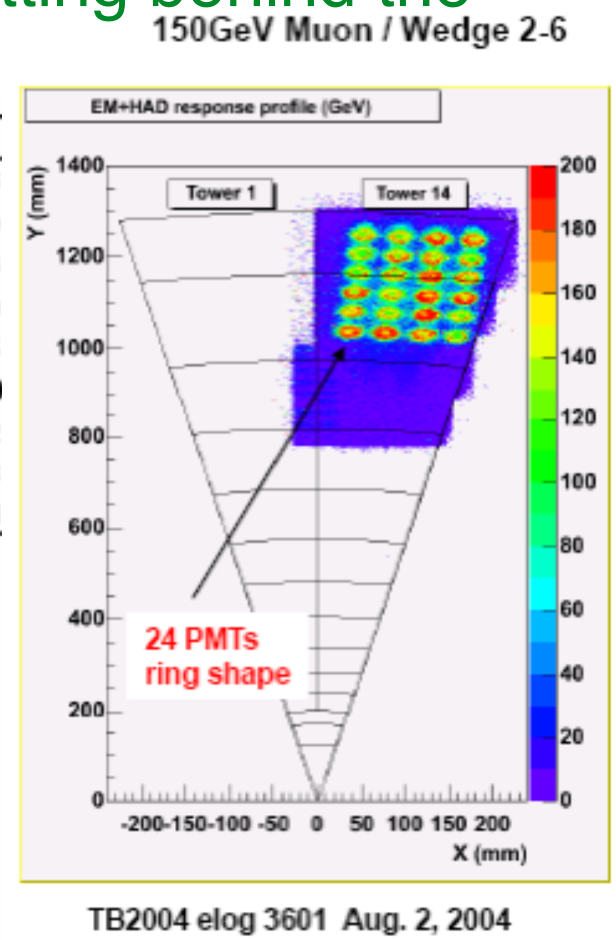
- Much improved understanding of the data

Use of Detector Simulation (IV)

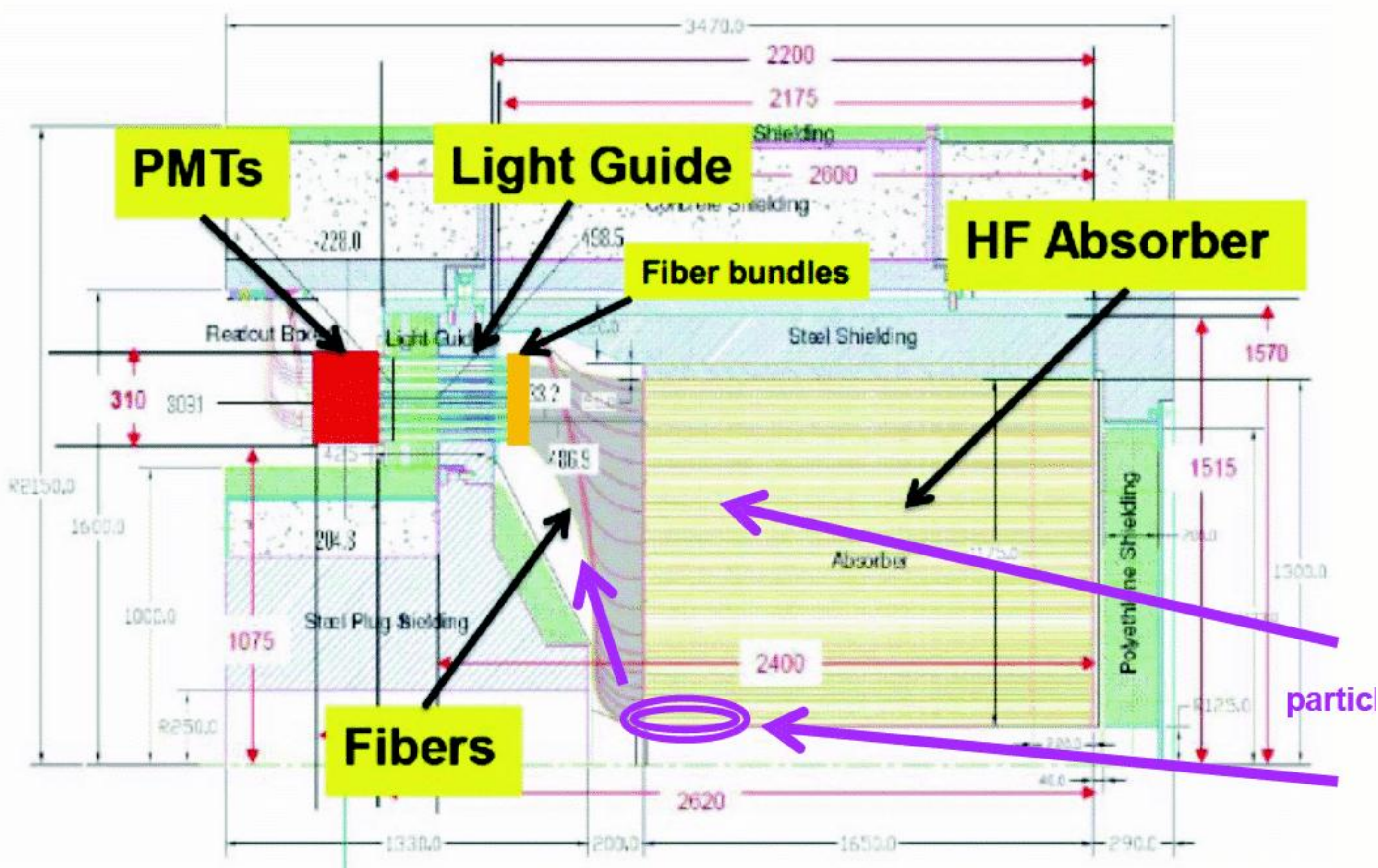
- Missing transverse energy is a key tool in the search for new particles in HEP. The long tails in the MET spectrum (which cannot be explained by simulation) were a worry.
- The events with large MET were having very high energy hits in the forward hadron calorimeter
- The large energy was seen in one 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

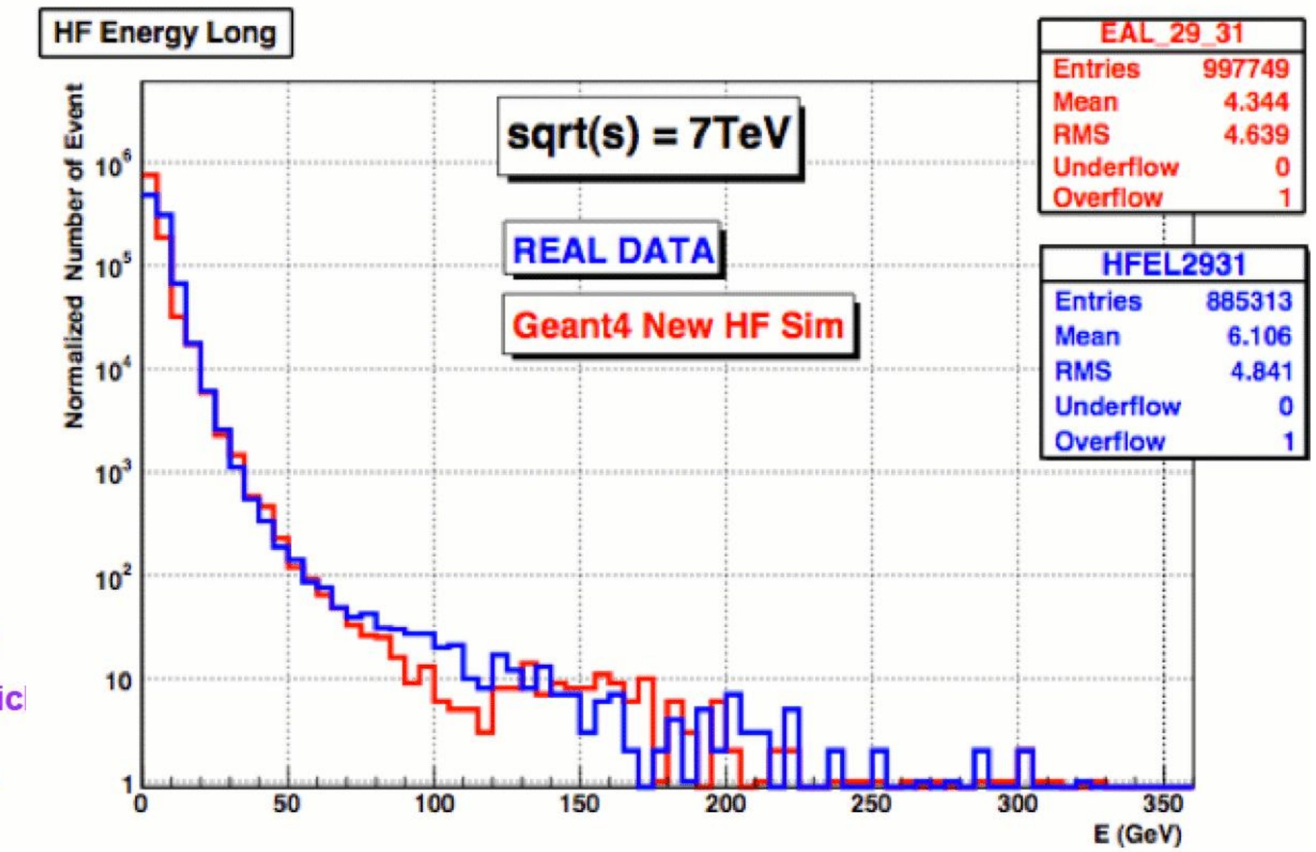


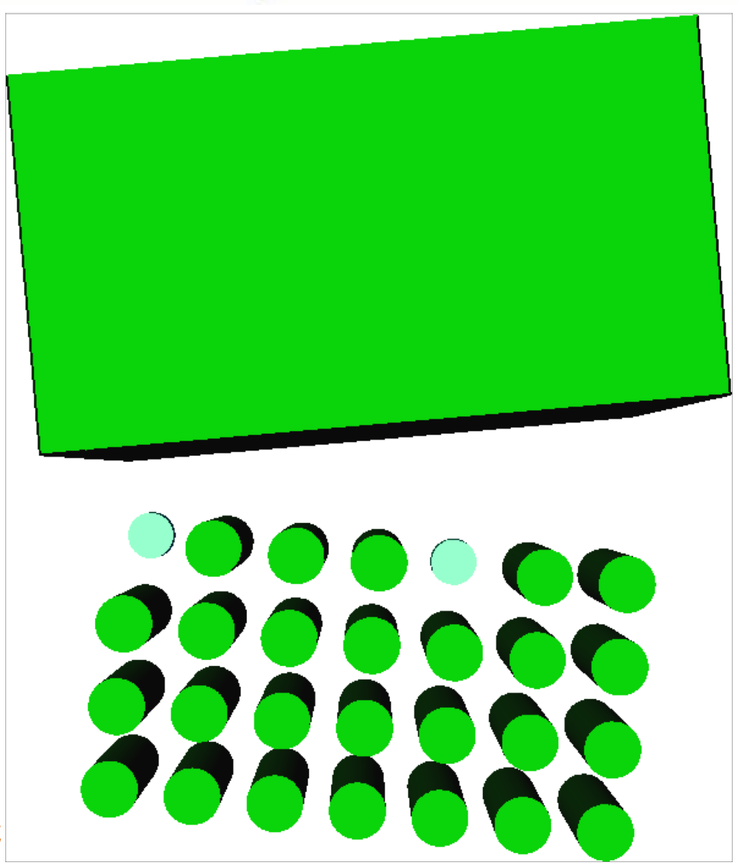




Use of Detector Simulation (V)



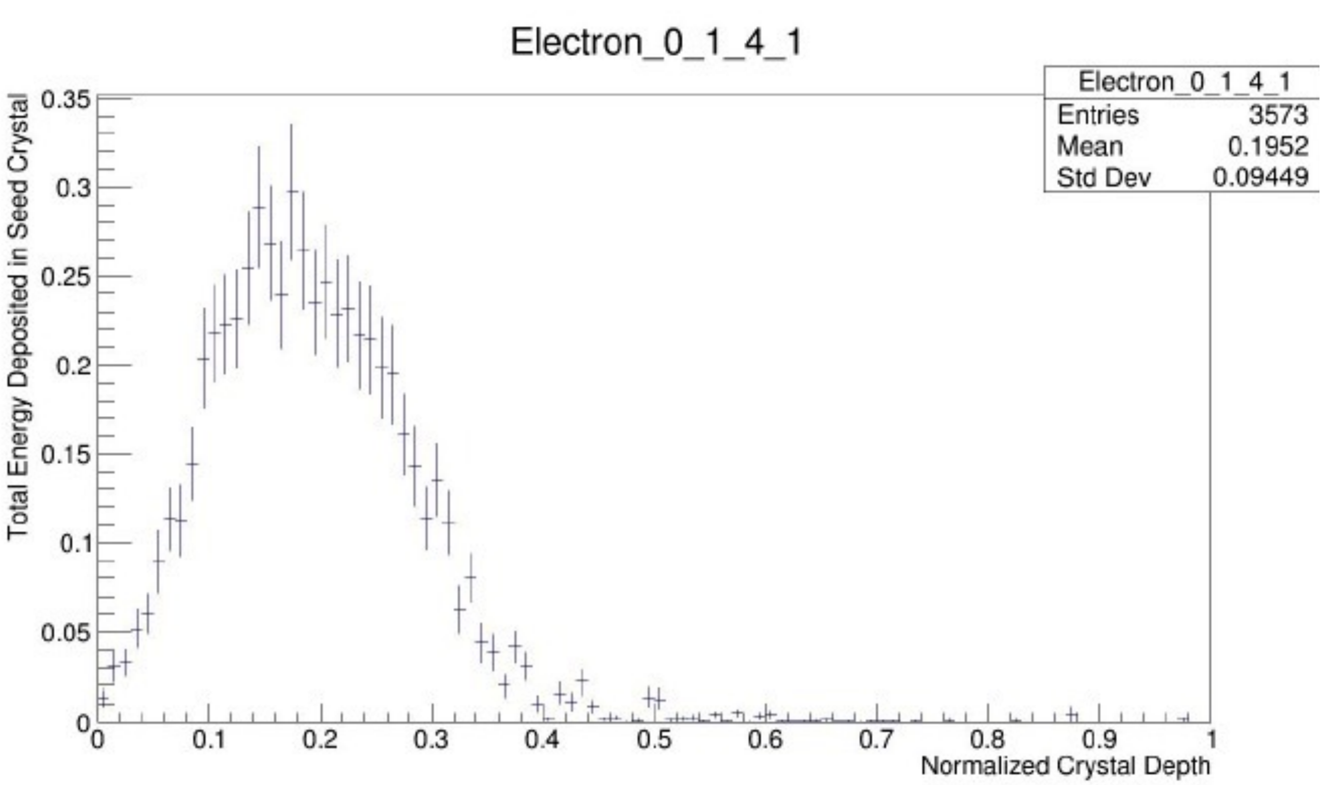


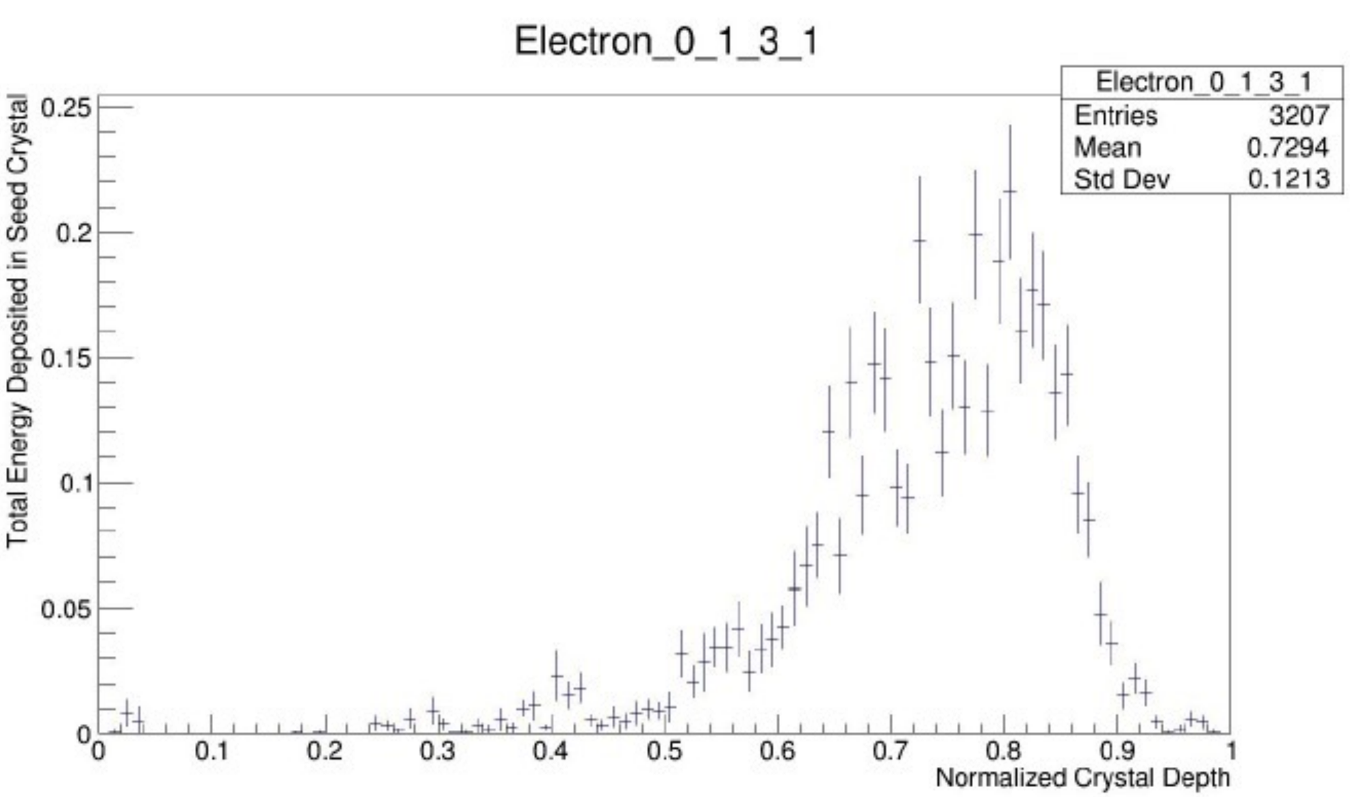


- Describe the PMTs behind HF and declare the photocathodes as Sensitive Detector
- Also, fibre bundles are described in the geometry and hits in the fibre bundle are associated with a given readout channel
- Energy spectrum, as well as anomalous hits, are well reproduced in the simulation
- The dominant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

Some Bugs Live Long

- Study of radiation damage in ECAL required to look into the longitudinal shower development in the crystals
- There is a facility to provide the shower depth in units of radiation length as the depth index of PCaloHit
- The current implementation of shower depth calculation showed peculiar profiles for 2 different sets of crystals for electrons and photons





How Depth is Computed

- The crystals are defined as trapezoid with the crystal axis along z-axis
- Computation of depth involves
  - Transform the hit point to the local frame of reference
  - Add the crystal length to the z-coordinate

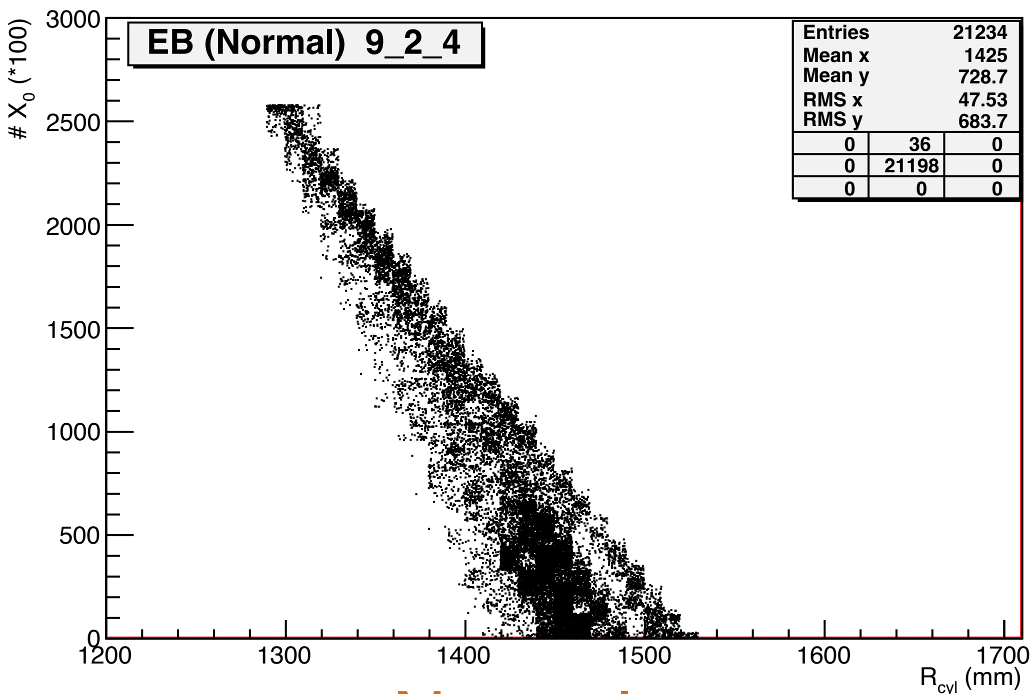
```
G4LogicalVolume* lv = stepPoint->GetTouchable()->GetVolume(0)->GetLogicalVolume();
G4ThreeVector localPoint = setToLocal(stepPoint->GetPosition(),stepPoint->GetTouchable());
double crlength = crystalLength(lv);
double dapd = 0.5 * crlength - localPoint.z();
if (dapd <= 100.) weight = 1.0 + slopeLY - dapd * 0.01 * slopeLY;
else              weight = 1.0;

double depth = 0.5 * crlength + localPoint.z();
```

- The same calculation is used to compute distance of the hit point from the optical transducer. This helps in estimating non-uniformity of light collection efficiency
- In the simulation of fast timing option for PhaseII studies, the method getLayerIDForTimeSim(…) is invoked - but there the calculation is done by testing the occurrence of “refl” in the name of the volume

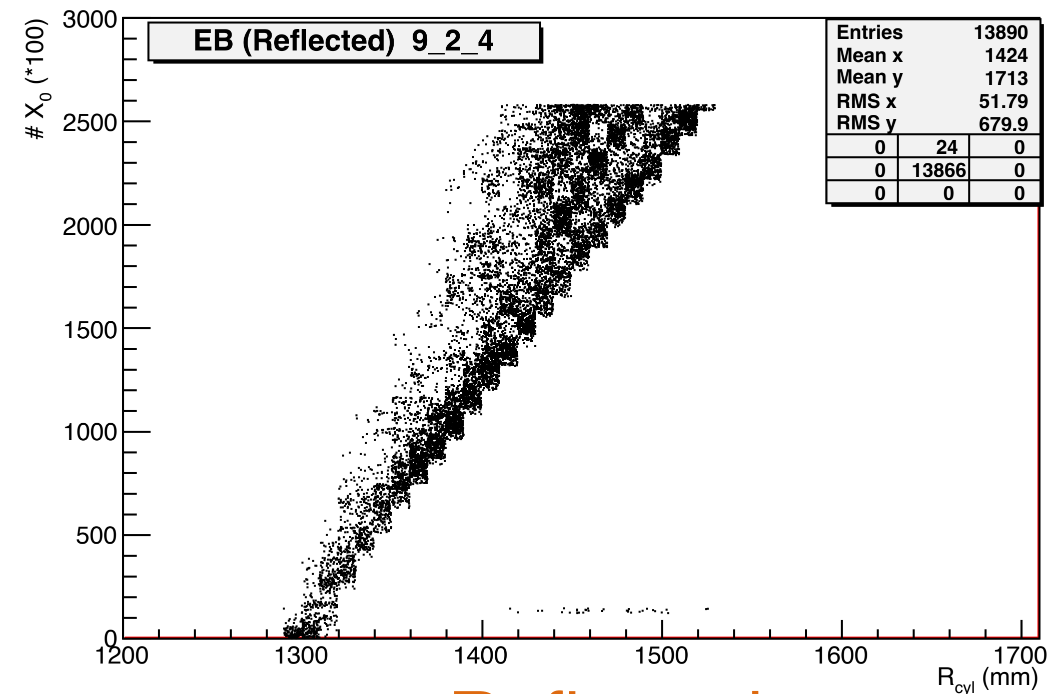
Depth Computation in the Buggy Version

- Generate di-muon events of fixed p_T in the region $|\eta| < 3.0$
- Compute depth in units of X_0 and plot it as a function of
  - R_{cyl} for crystals in the barrel region
  - $|z|$ for crystals in the endcap region

EB

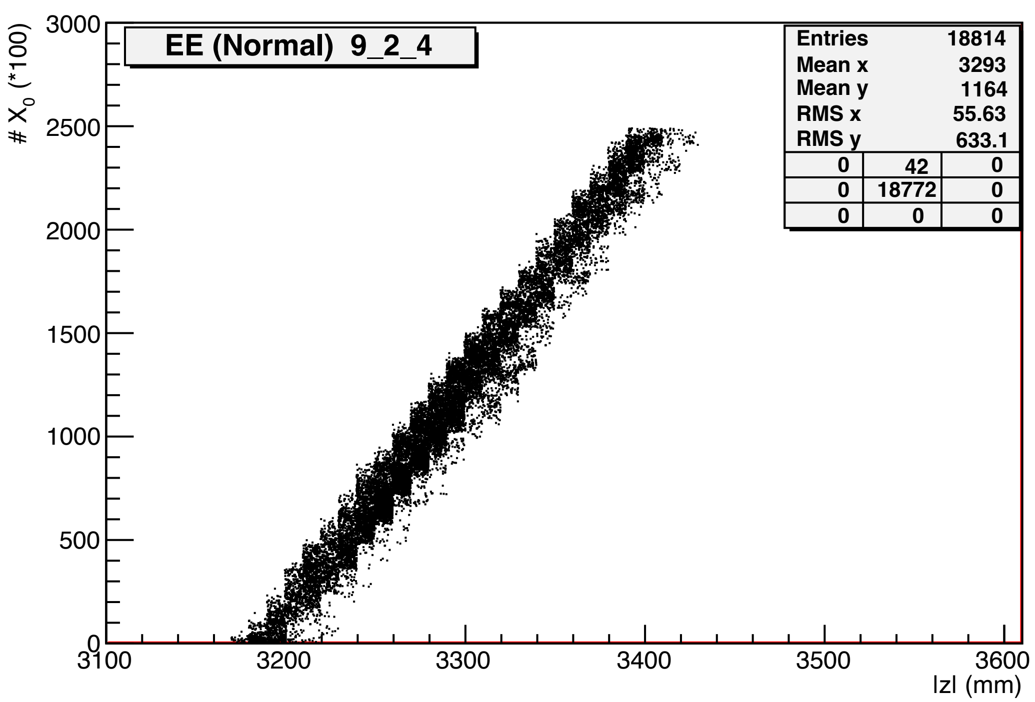


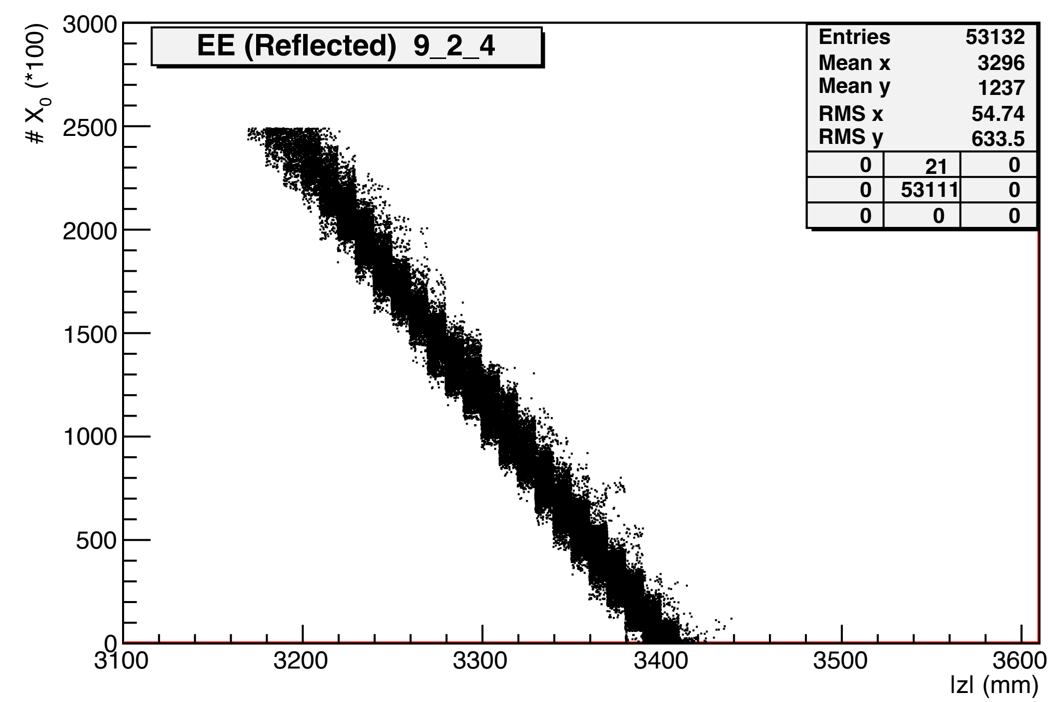
Normal



Reflected

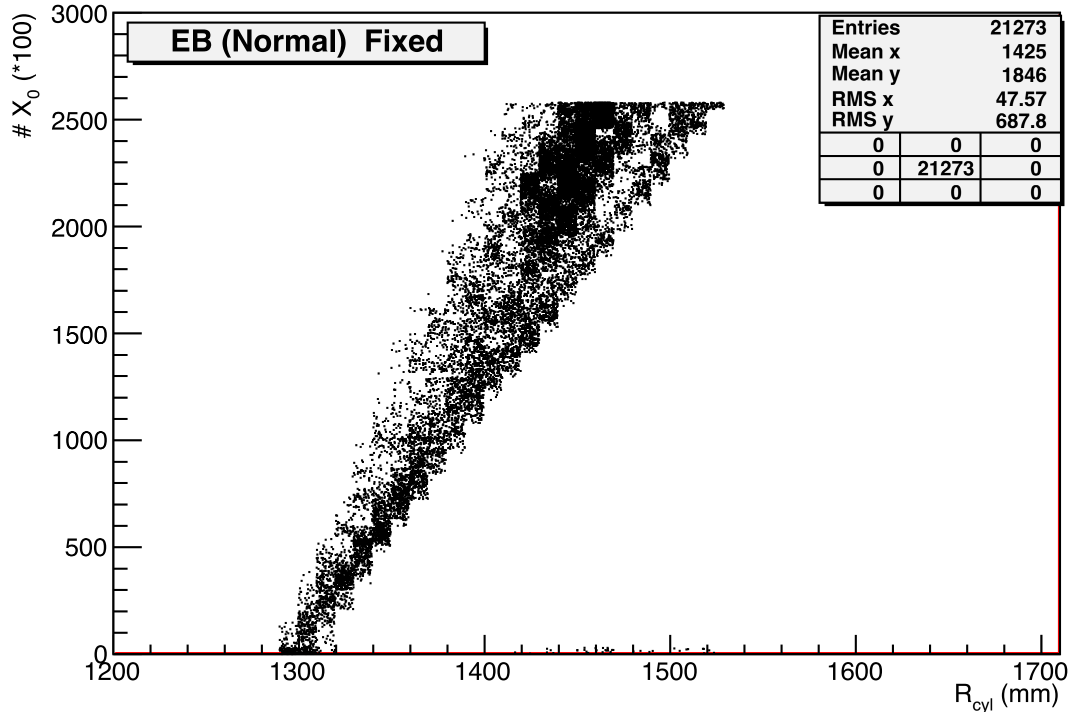
EE



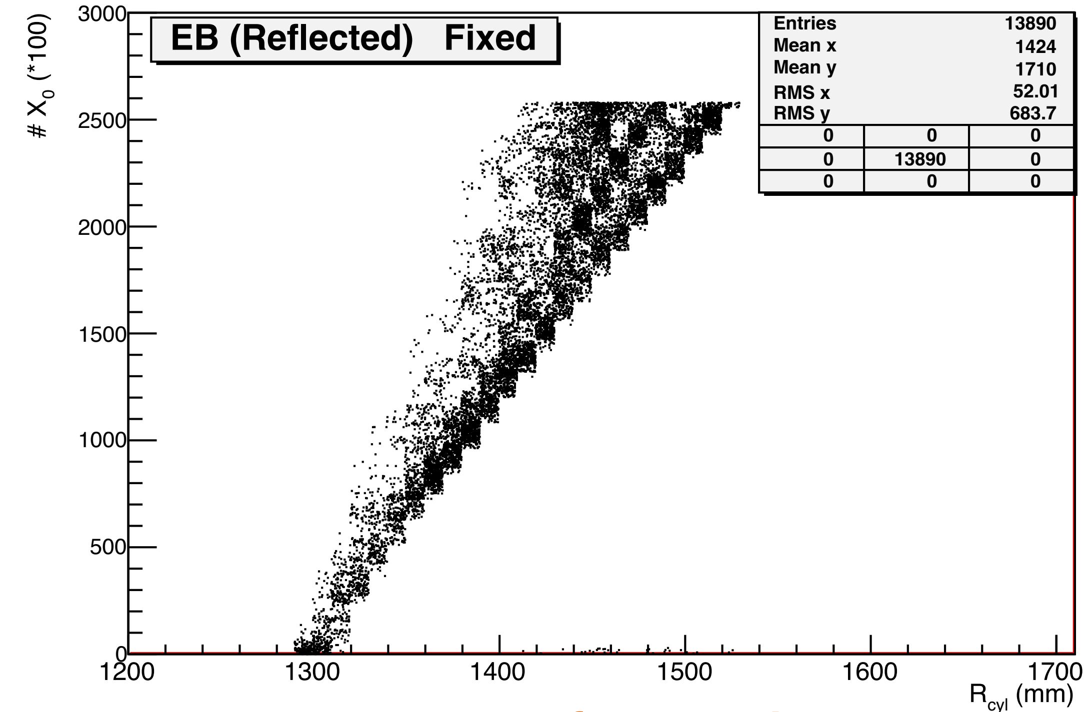


Fix Cures the Issue

- Treat the normal and “refl” volumes differently both for EB and EE

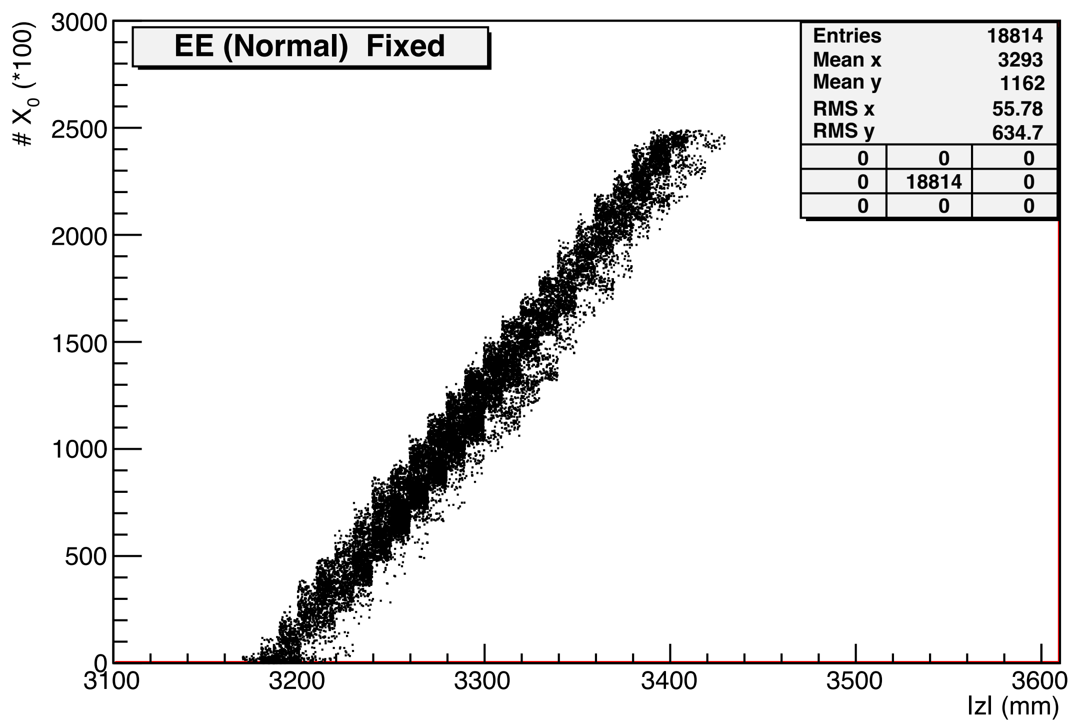


EB

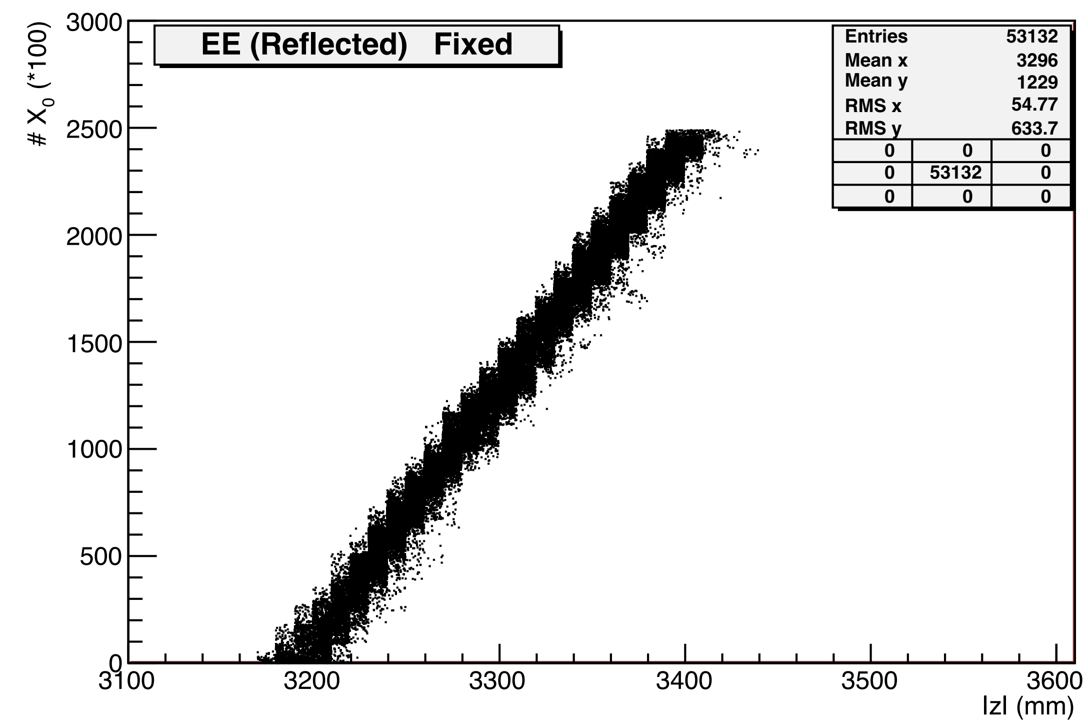


Normal

Reflected



EE



Future of Detector Simulation (I)

- 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

Future of Detector Simulation (II)

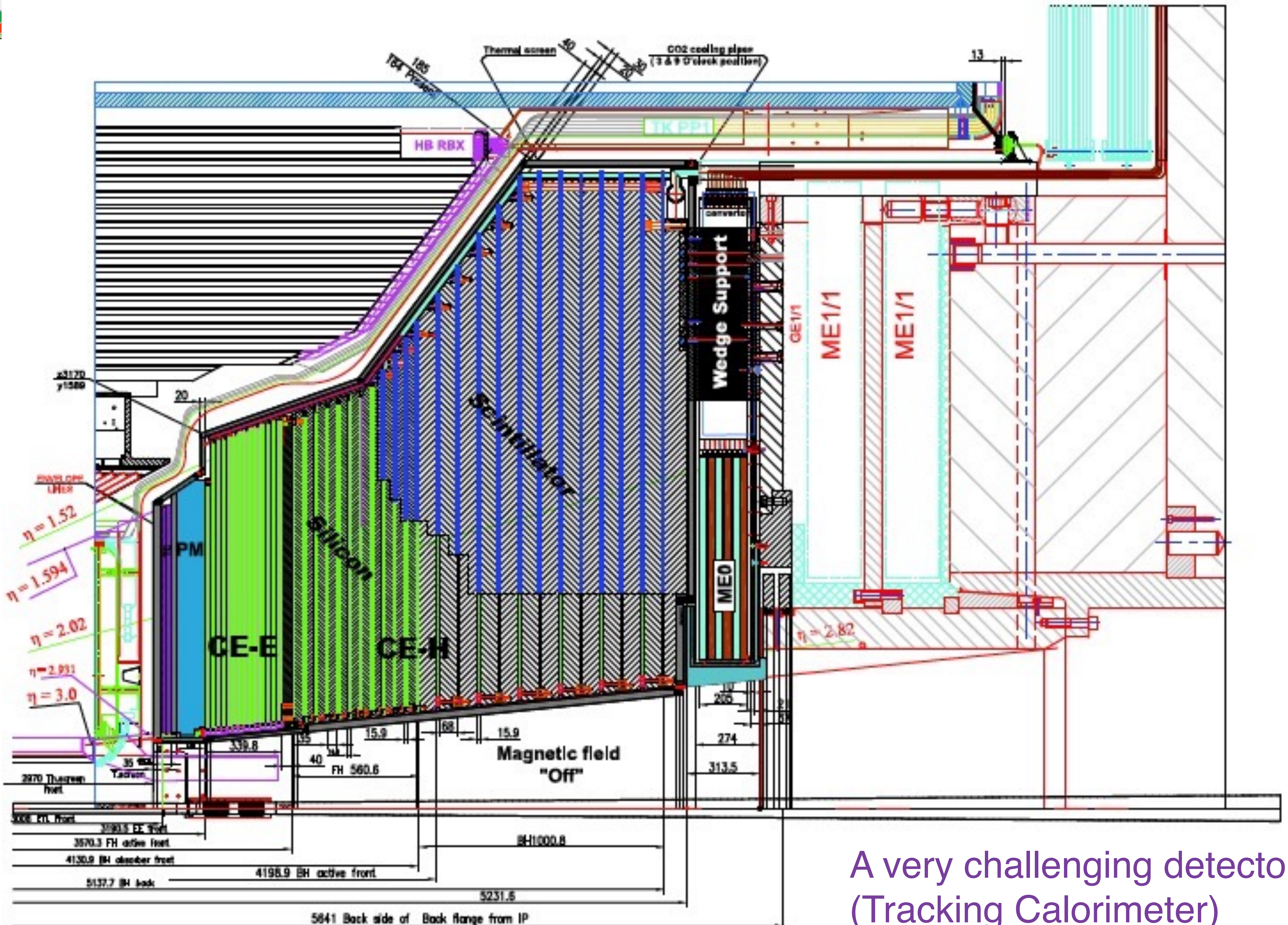
- Many detector elements will be unusable during the high luminosity runs of the LHC — some detectors are damaged due to radiation, and some detectors will suffer due to higher occupancy

- 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 for charged particles 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

The New Layout of Endcap

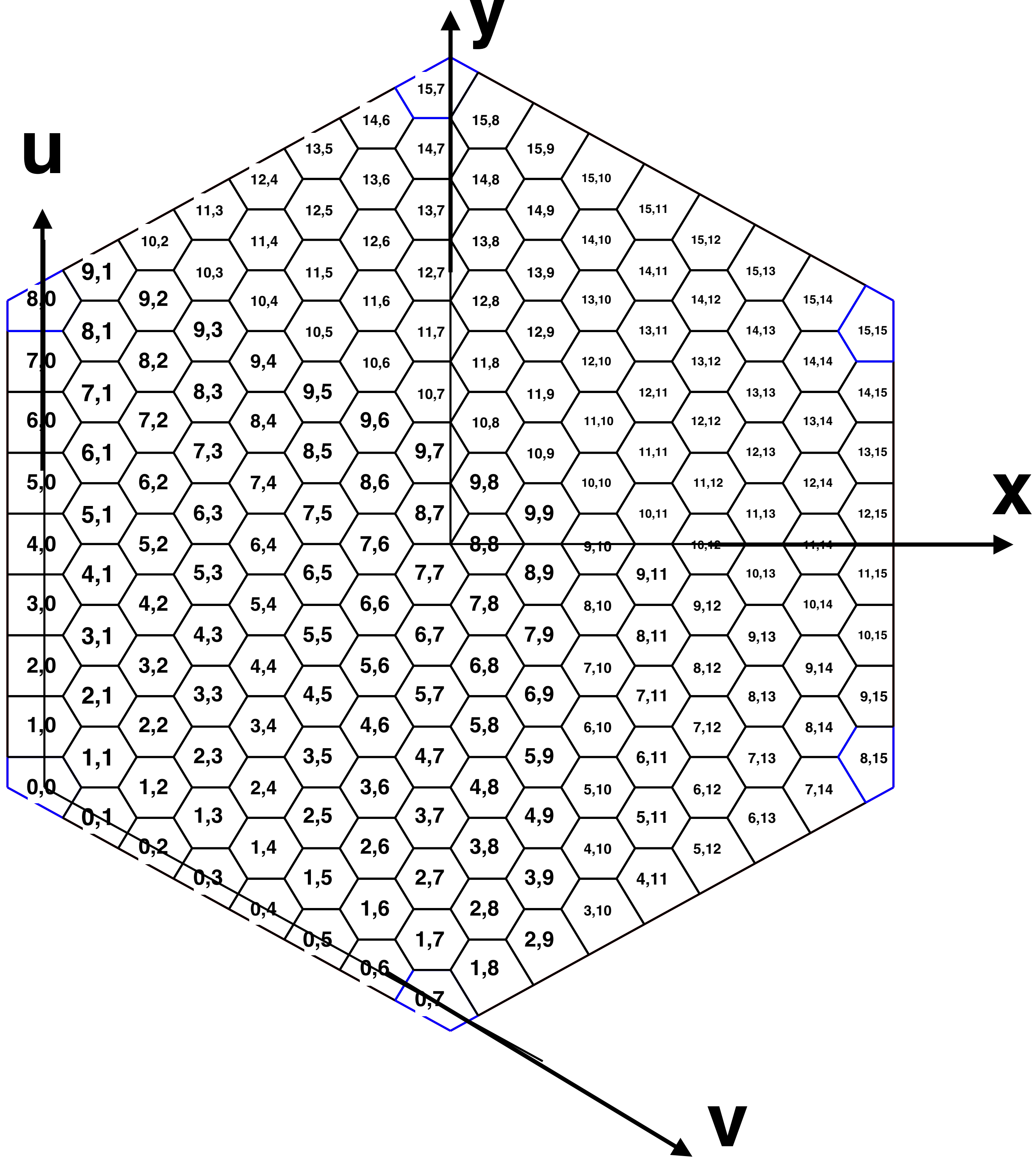


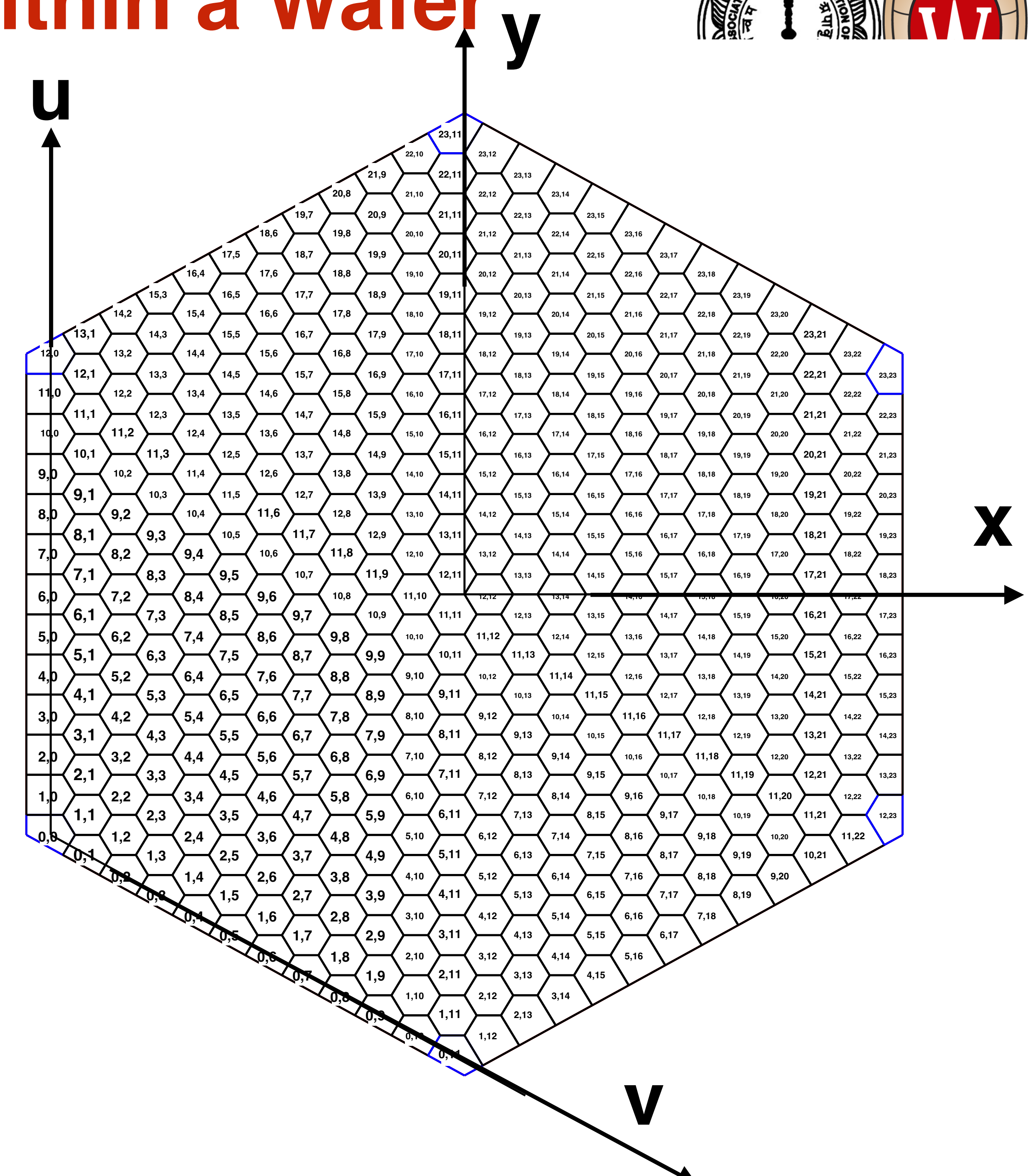
A very challenging detector
(Tracking Calorimeter)

Tasks to be Undertaken for HGCAL

- Update longitudinal structure to match the final engineering description
- Describe correctly the “mixed scintillator and silicon” layers allowing missing tiles
- Update silicon wafer and associated layer (base plate, read-out board, …) sizes, introduce inter-wafer gaps, and mouse bites
- Describe the final cell layout; inclusion of “calibration cells”
- Update the final “incomplete wafers” geometry at the boundaries
- Provide more realistic active area coverage at the inner/outer boundary and allow flexibility to adjust the coverage
- Introduce inter-cassette gaps
- Describe mechanical structures on the inner and outer boundaries (support cone, thermal shield etc.)

Cell Definition within a Wafer



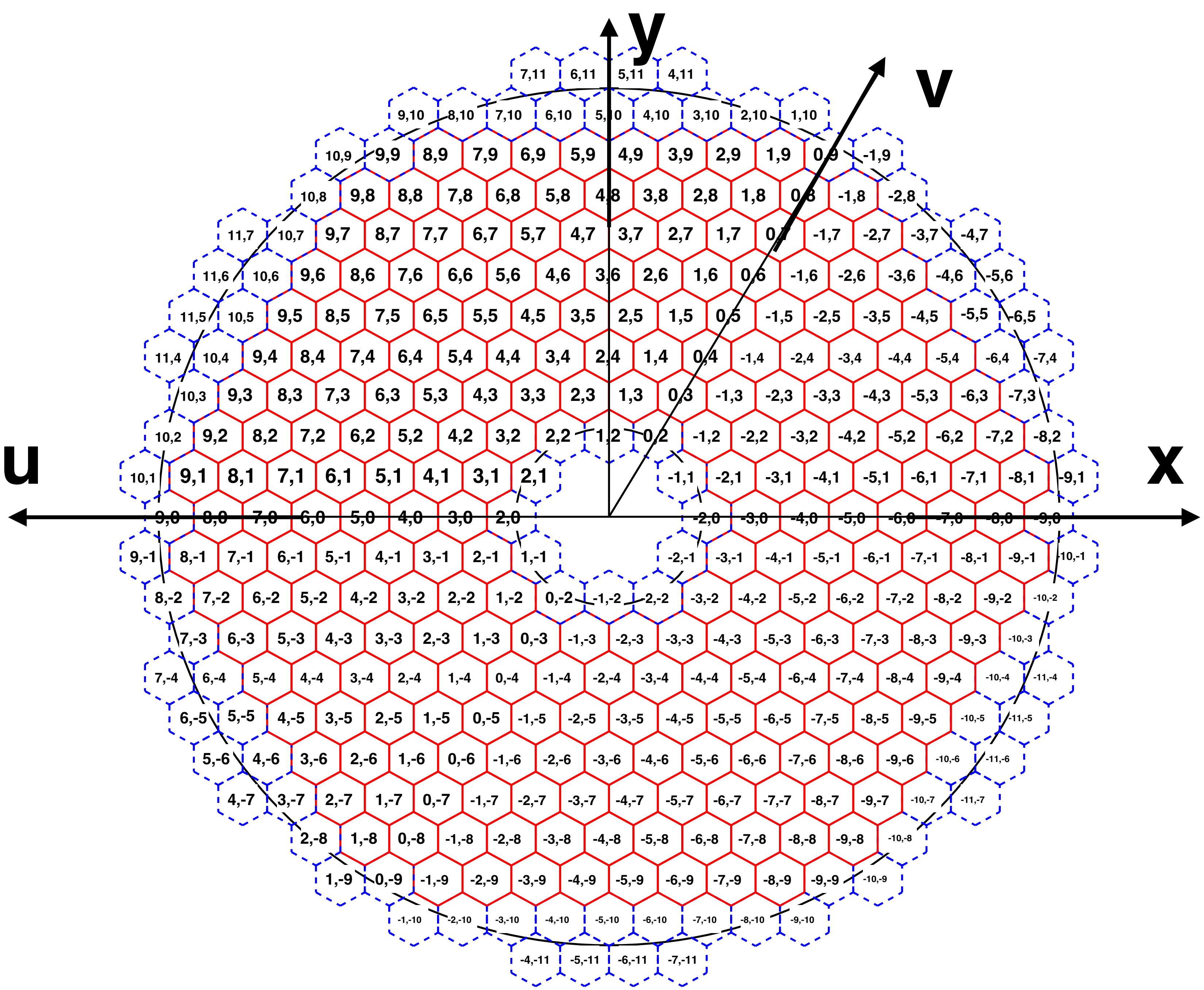


- Cells are numbered by u,v coordinates and can be transformed to row,column indices by trivial linear expressions.

- Linear equations relate (u,v) and (x,y)
- Maximum number of cells along u/v:
  - 16 for coarse cells
  - 24 for fine cells

Wafers in a Layer

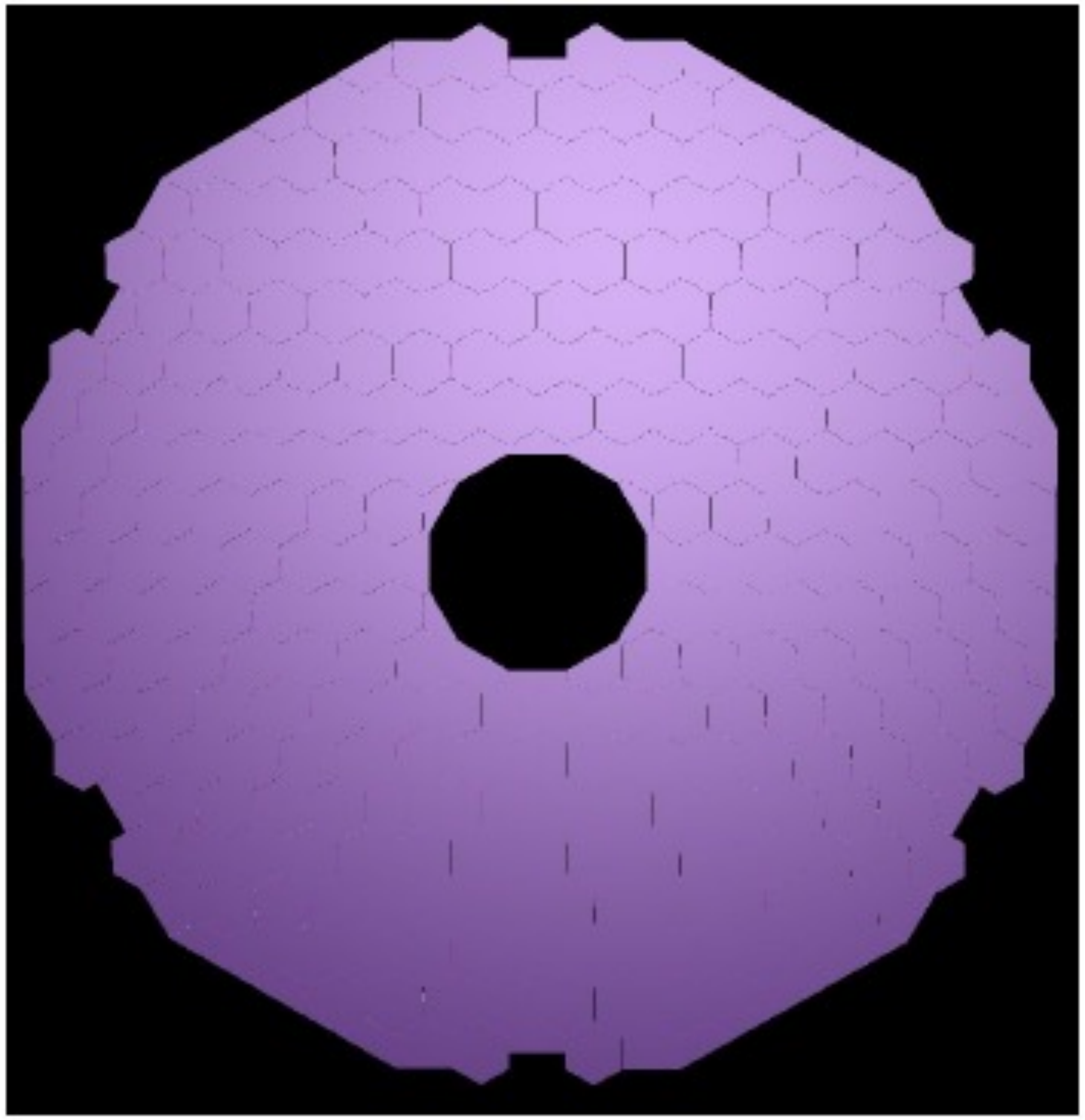
- Positioning wafers in a plane is done using full and partial wafers
  - Real wafers where all 6 corners are contained within the boundary
  - Inner cell structure is not defined in partial wafer geometry
    - assigned at run time during simulation

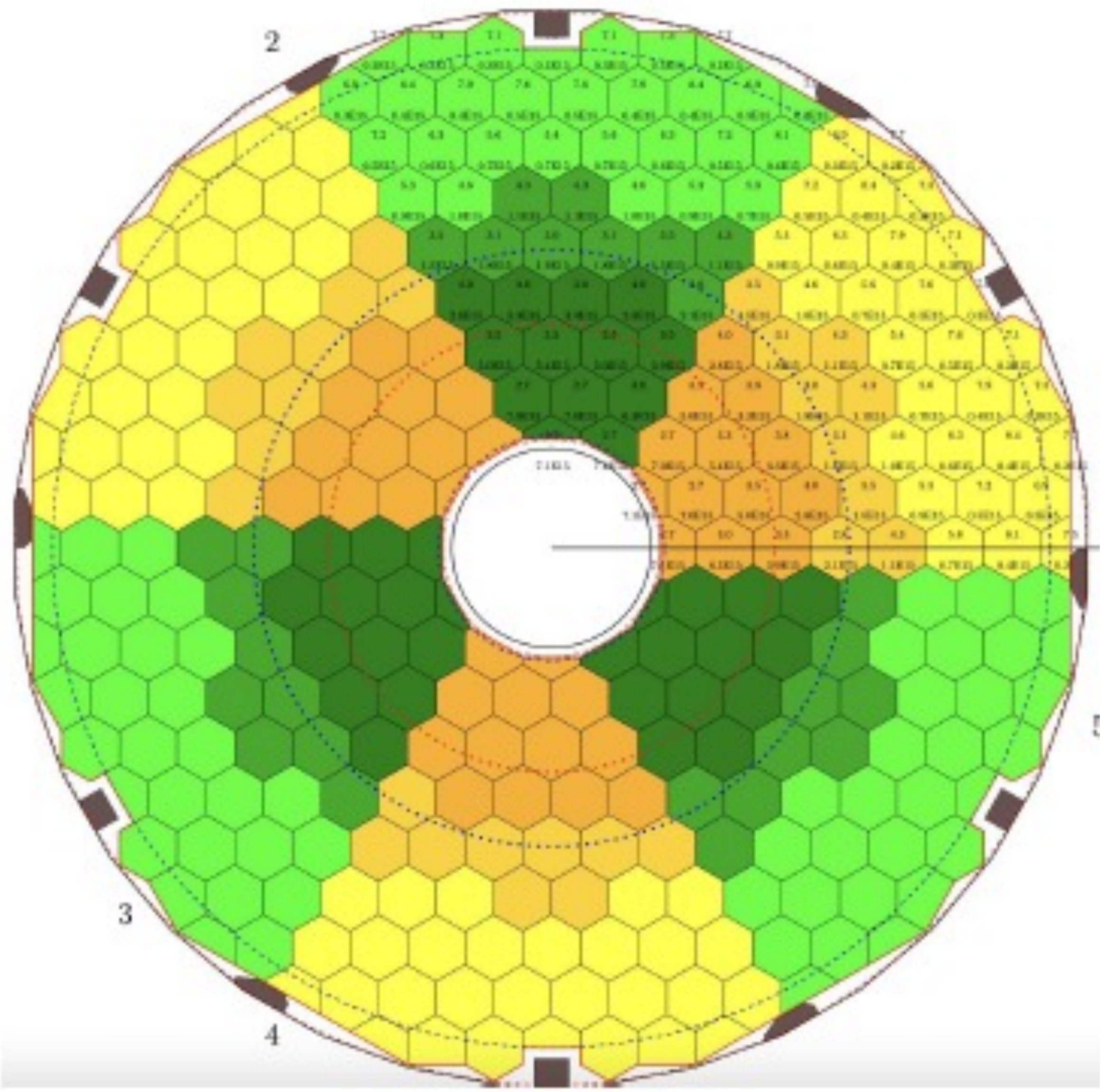


- The center is defined on the beam axis for wafer-centering layers (all EE and some HE layers)
- Some HE layers enjoy corner centering (2-types with the centre shifted up and down along y-axis)
- There is a simple linear relation to go from (u,v) to (x,y) scheme using

Wafer Type

- Three different wafer types are there to handle different radiation environment:
- The boundaries for the 3 types depend on |z| position of the layer and the boundary is determined from an independent study using FLUKA

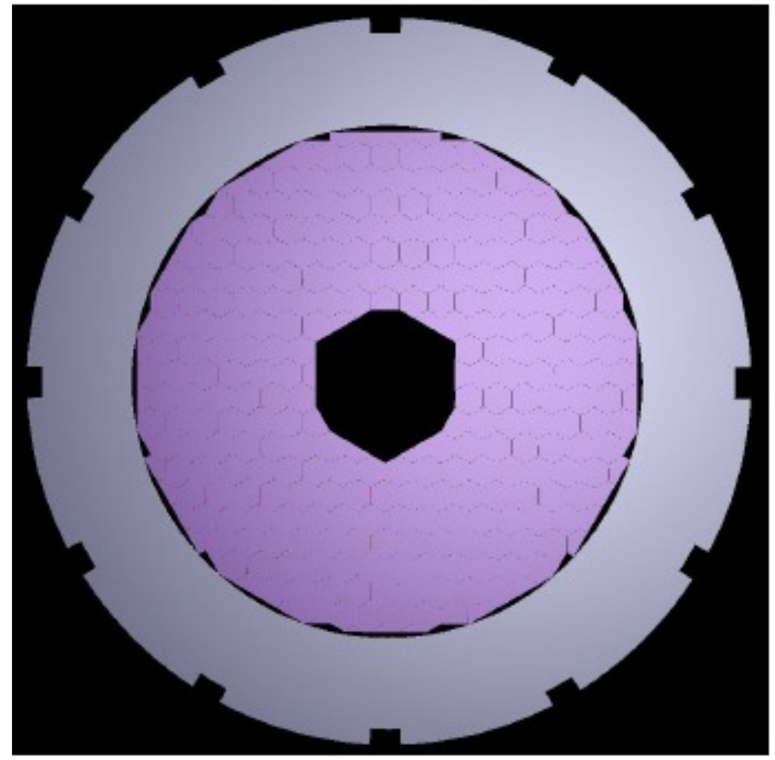


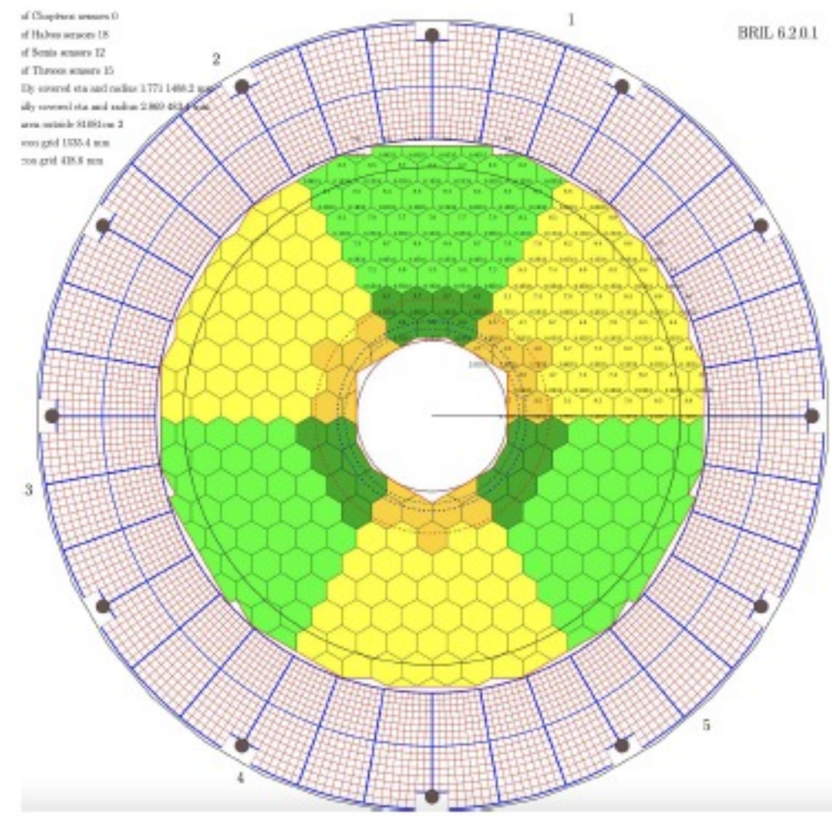


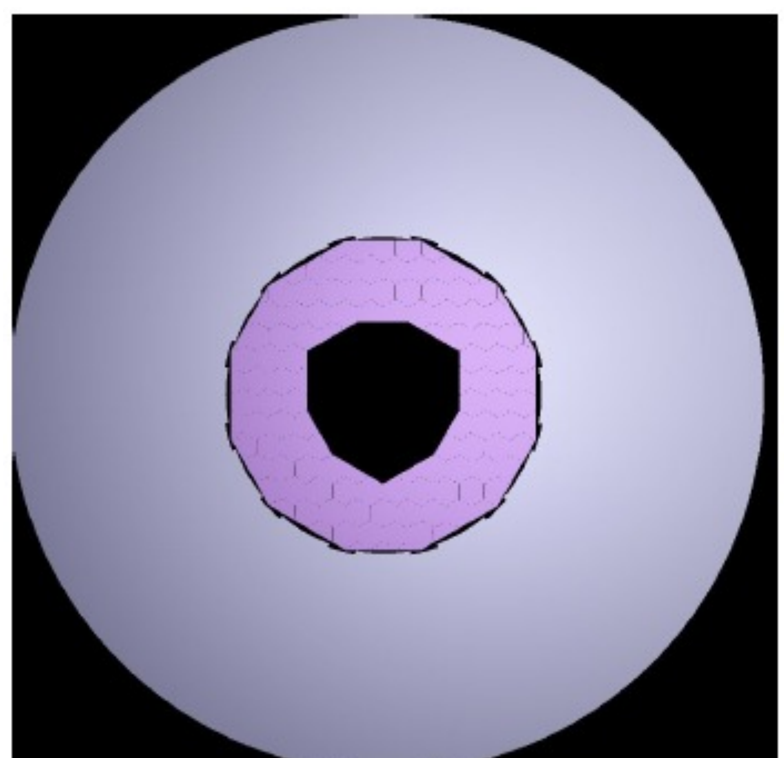
- Geometry description (left) reproduces the engineering condition (right)

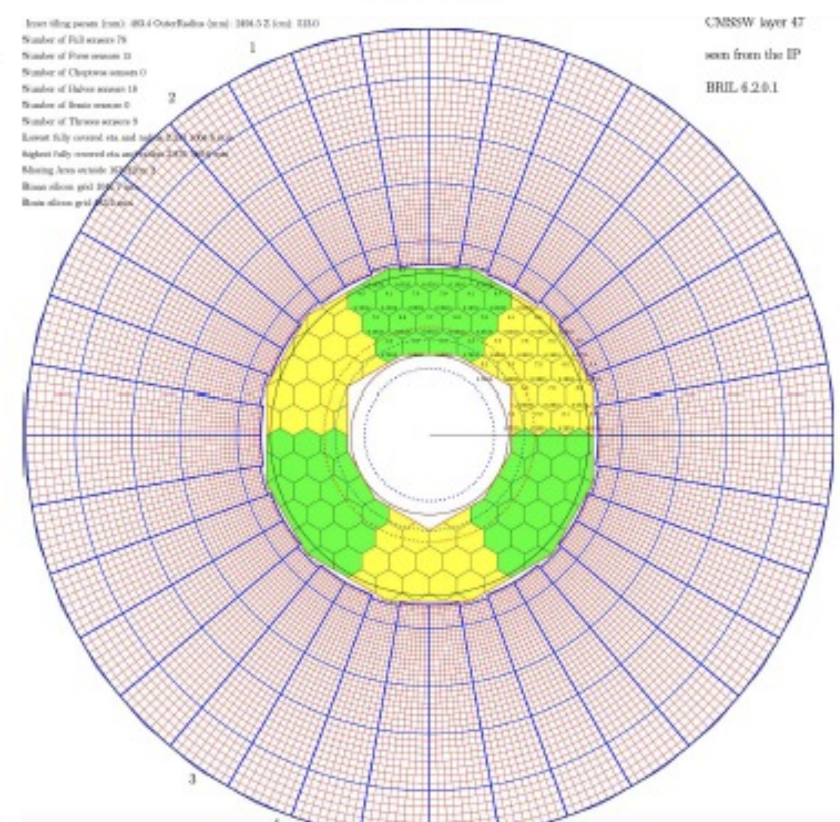
The Mixed Section

- The mixed section is divided into two virtual parts one containing wafers and the other scintillators
  - The bottom part contains wafers with base plates, PCB's, .. and wafers (full and partial)
  - The top part contains scintillator tiles, cables, connectors, PCB, …
    - The scintillator layer is defined as partial tubes according to the presence or absence of tiles









Phase 2 Geometry (one proposed scenario)

- Utilizes 12415 solids in 12613 Logical Volumes among 13.0 million touchables
- Summary of solids used in touchables:

	Standard	Reflected
Box	1236k	429k
Tube	57.9k	755
Trapezoid	158k	141k
Cone	1862	0
Polycone	206	0
Polyhedra	1572	0
ExtrudedPolygon	10845k	0
TruncatedTube	92	0
UnionSolid	614	0
SubtractionSolid	173k	594
IntersectionSolid	72	0

- Overlap summary (tolerance of 0.1 mm):

	Tracker	ECAL	PreSh.	HCAL	Muon	Infra.St	Total
DDD	0	0	0	0	0	0	0
DD4Hep	0	0	0	0	0	0	0

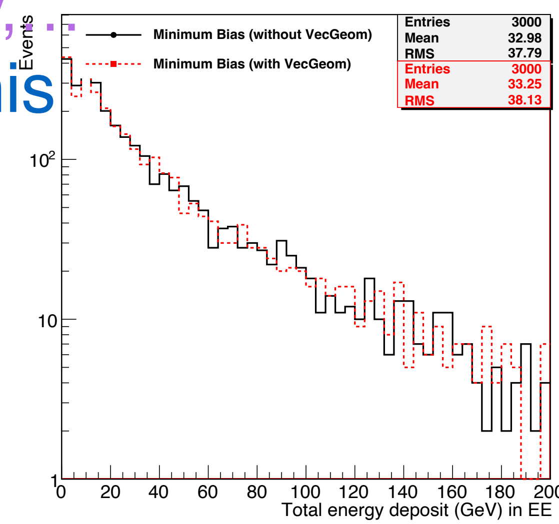
Progress of Basic Toolkit

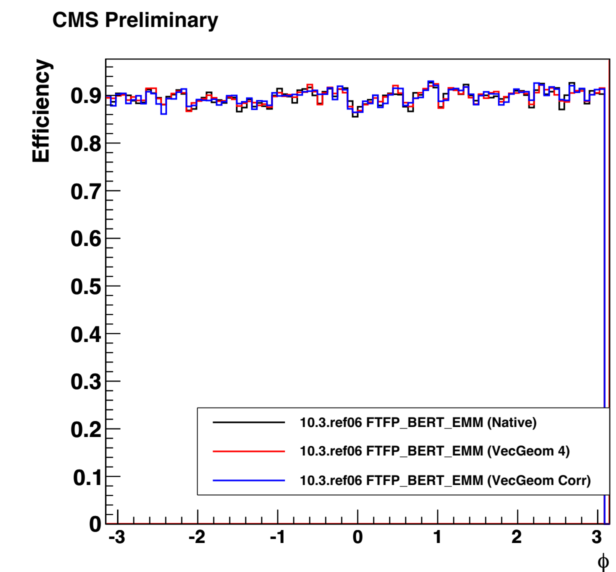
- Computing moves from single-threading to multi-threading and this advantage is already utilized by Geant4 and all application codes (ATLAS, CMS, LHC-b)
- There was an R&D effort to make use of vectorisation
  - The geometry and tracking code in EM field was rewritten to enable effective vectorisation
  - A new approach is 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 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; yet to provide a stable and well-tested toolkit

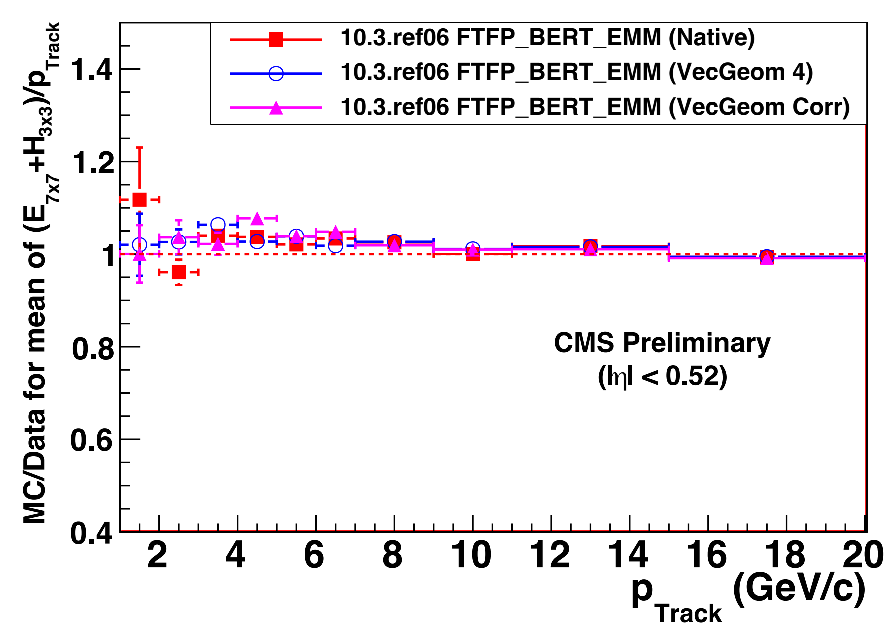
Use of these Improvements

- 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







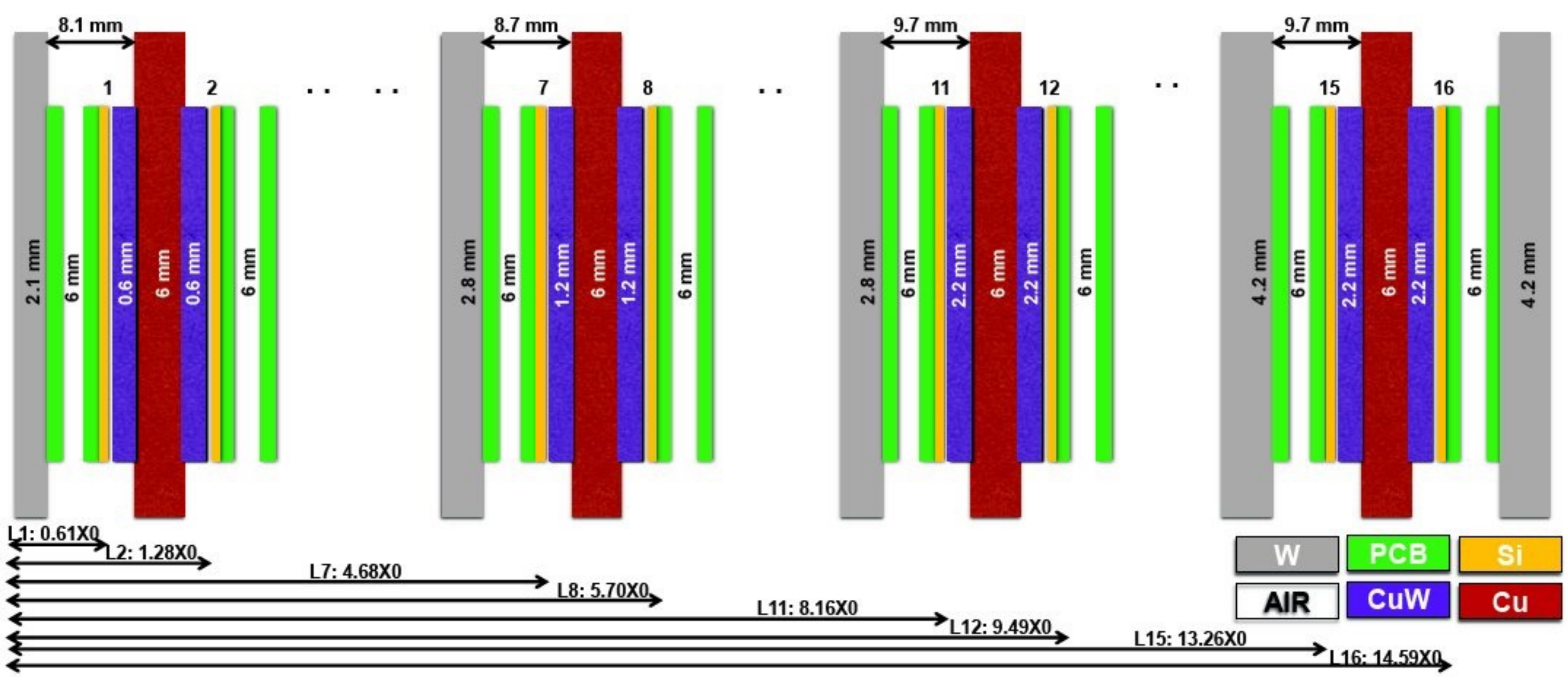
Summary

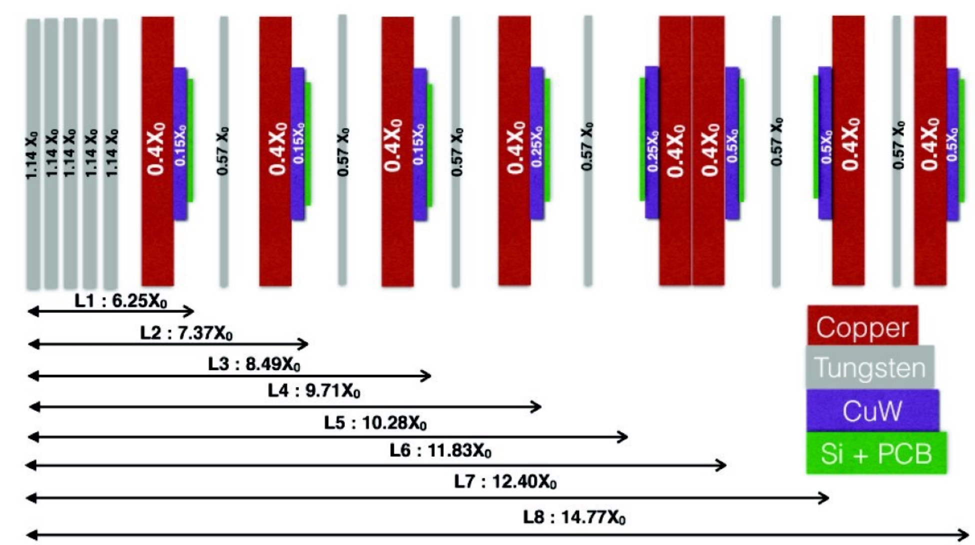
- Detector simulation is an essential tool for modern-day nuclear and high-energy physics experiments
- CMS has gone through a series of developments to meet the challenges:
  - Accuracy in the predictions
  - Performance in speed and in memory
  - Robust against unusual circumstance
  - Extension for the changing detector
- The ultimate test of detector simulation is how it performs against the data
  - Test beam data with identified particle types are used as one source of validation while isolated charged particles from collision data are used as a second source
  - There is a good agreement between data and Monte Carlo for the new version of the physics list (FTFP_BERT_EMM) to be used by CMS for its future event production using Geant4 version 11.1.p0X
  - Validation of physics within Geant4 is continued using CMS data from new test beam data
- CMS tries to achieve a well-tested Monte Carlo program to get better results from the CMS experiment

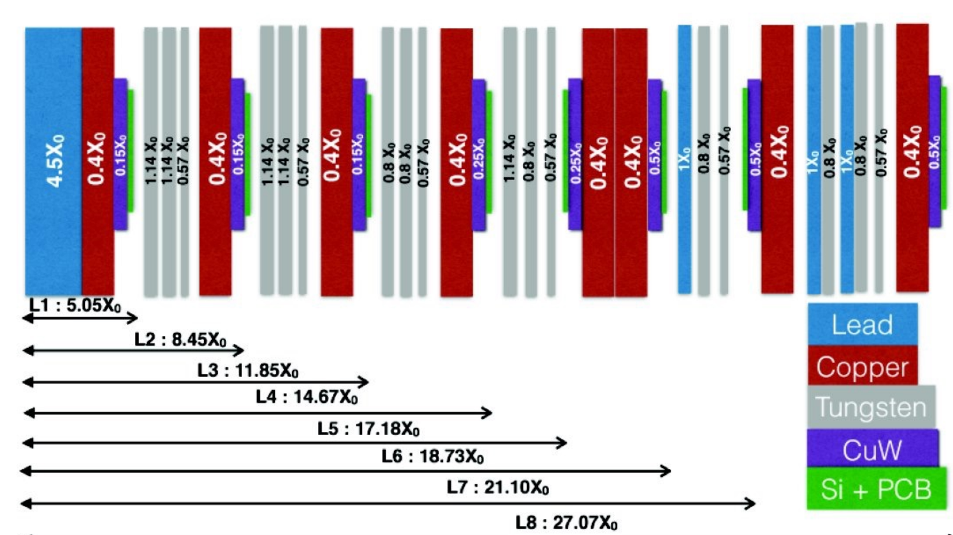
Additional Slides

Test Beam with HGCal Prototype

- Prototype of the electromagnetic component of high granularity calorimeter is exposed to electron beams at Fermilab (up to 32 Gev) and at CERN (up to 250 GeV)
  - 16 modules at Fermilab in a setup of ~14.6 X_0
  - 8 modules at CERN in two setups of 14.8 X_0 and 27.1 X_0
- All three setups are simulated using Geant4 version 10.2.p02 and physics list FTFP_BERT_EMM

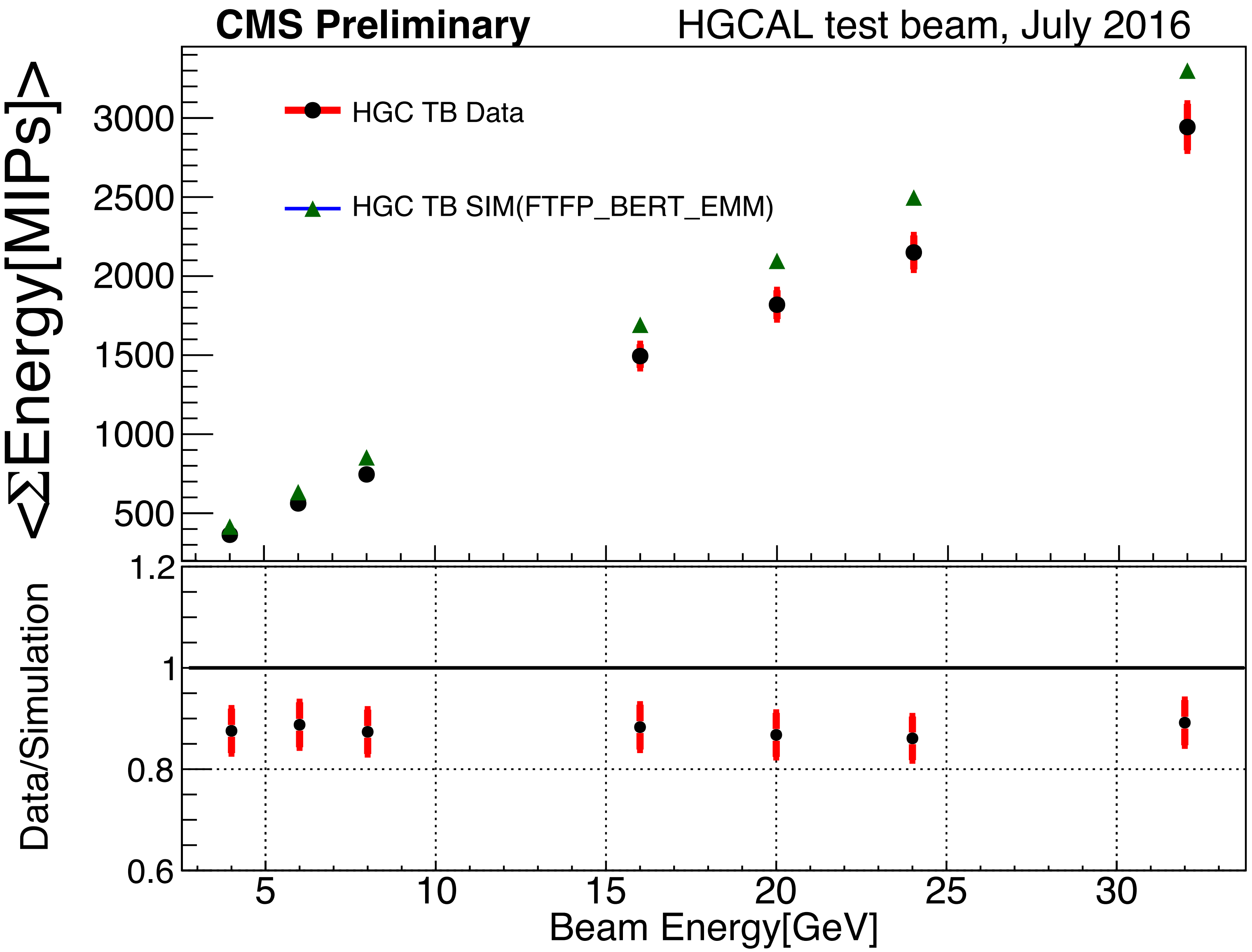


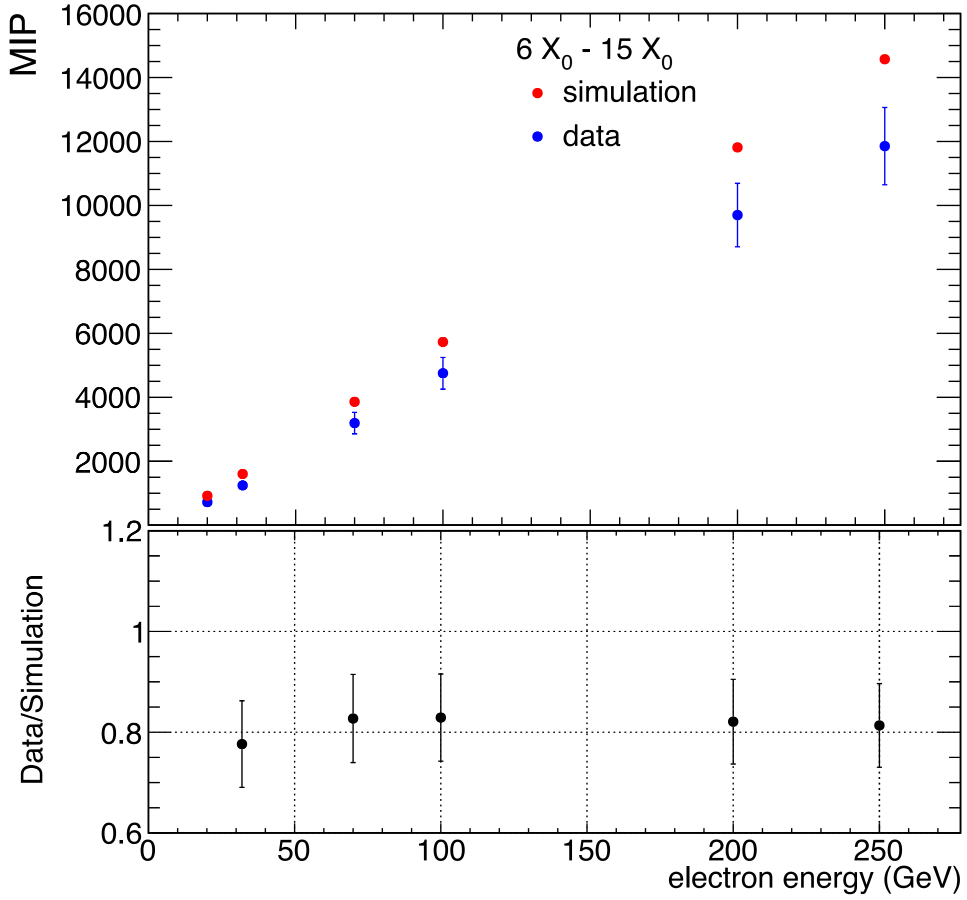


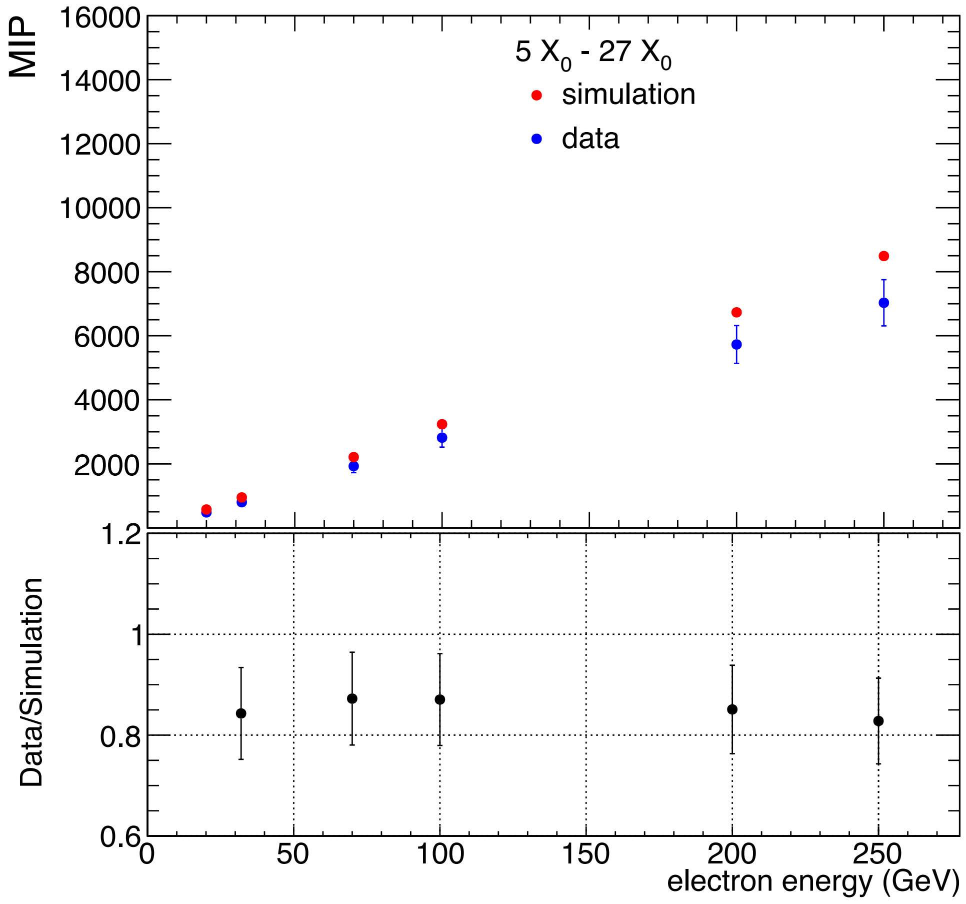


Response to Electrons

- Both data and MC samples are calibrated with MIP
- Energy deposit in each layer is weighted by a factor depending on material in front
- Linearity is observed in data and MC
- Energy scales are ~10% different in the two cases

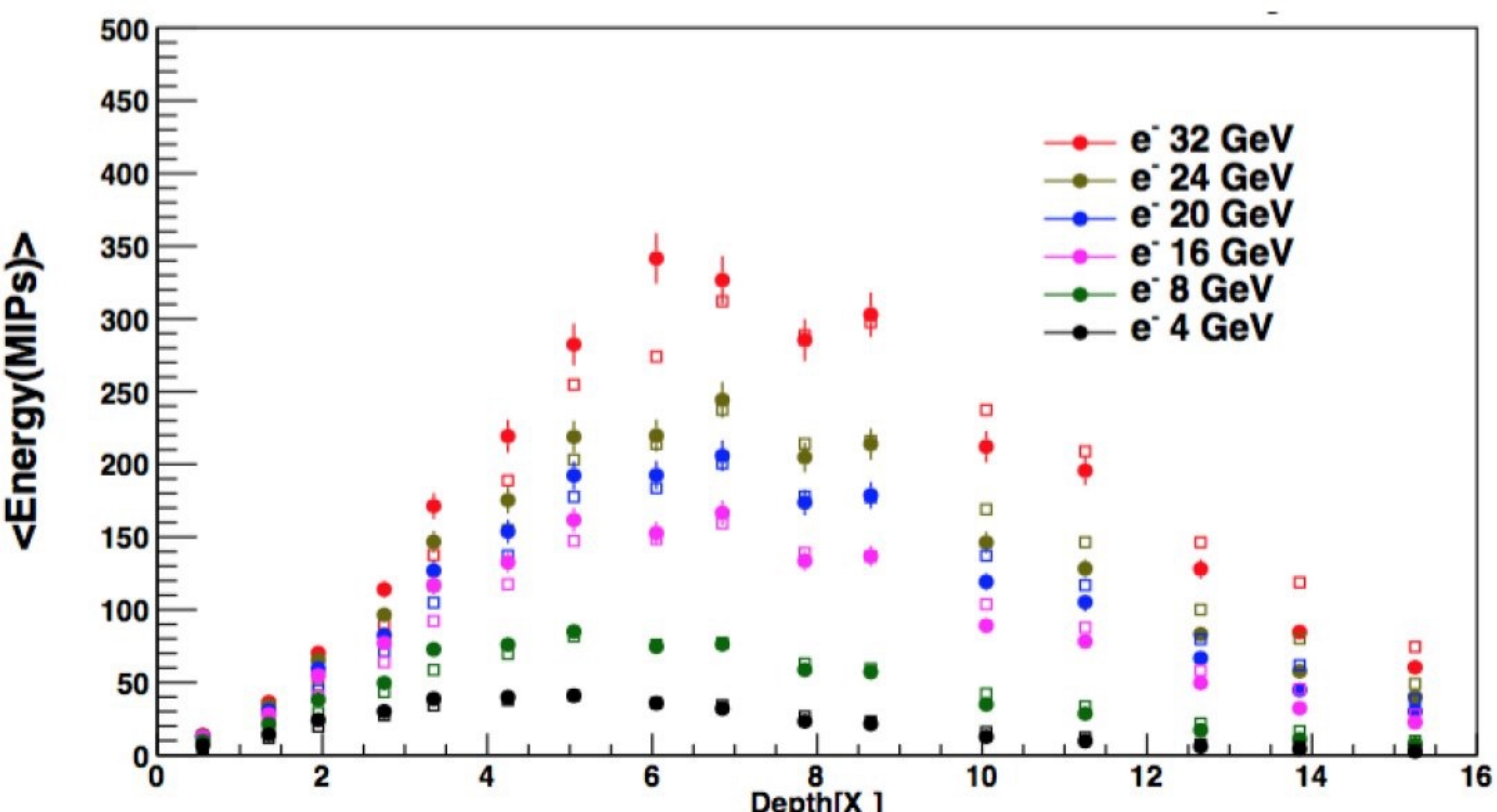


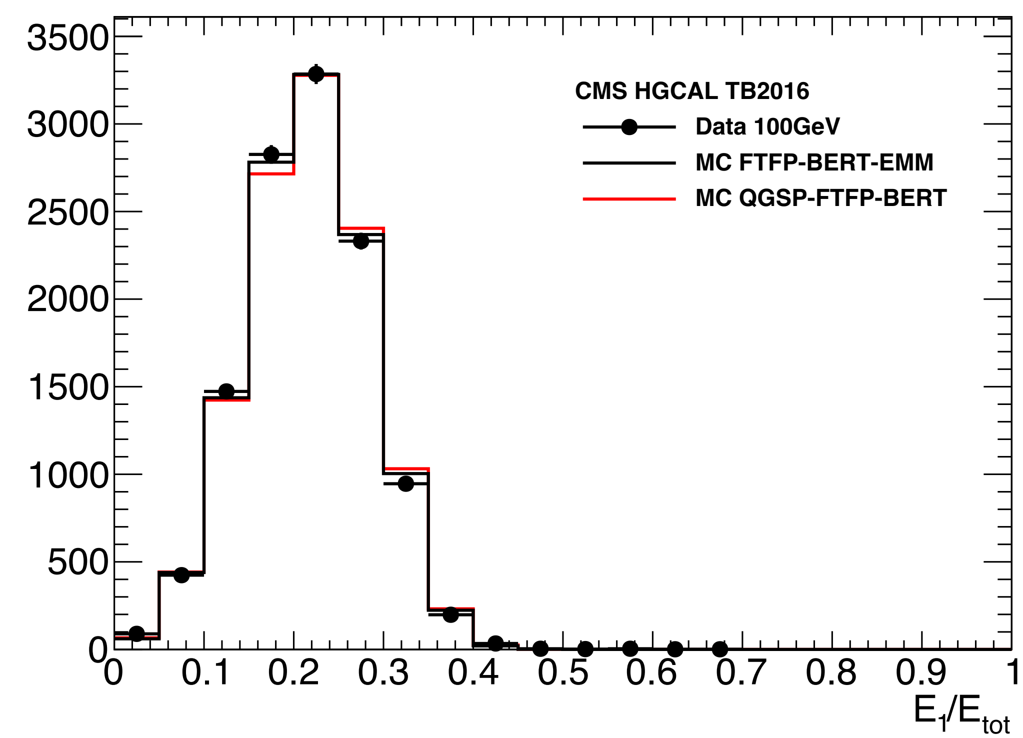




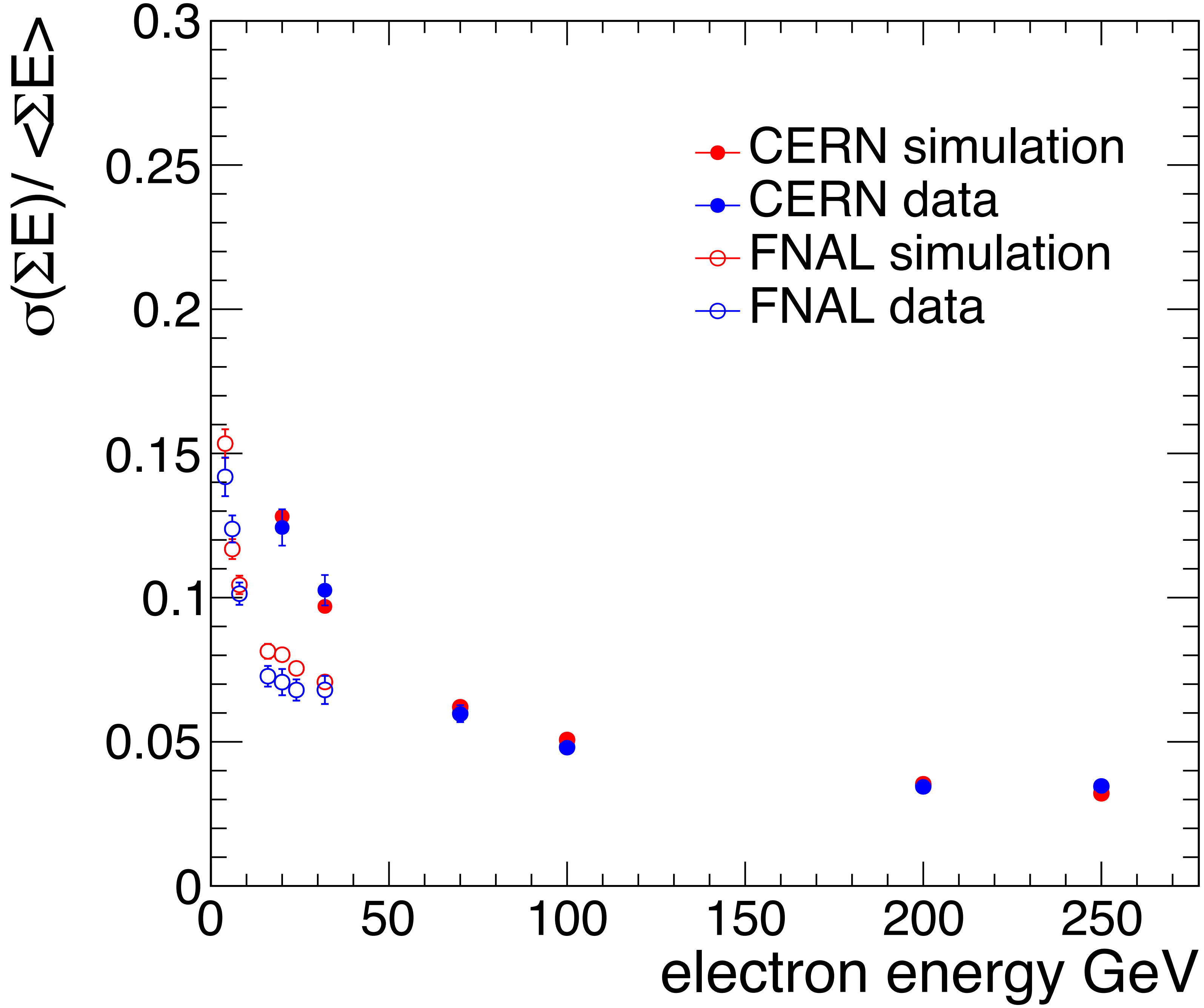
Shower Shapes

- Shower shapes are compared
  - Longitudinal shower profile studied from mean energy as a function of depth measured in units of X_0
  - Lateral profile measured from energy ratio of the central crystal to the total
- Fairly good agreement observed between data and MC





Energy Resolution



- Energy resolution compares well between data and MC