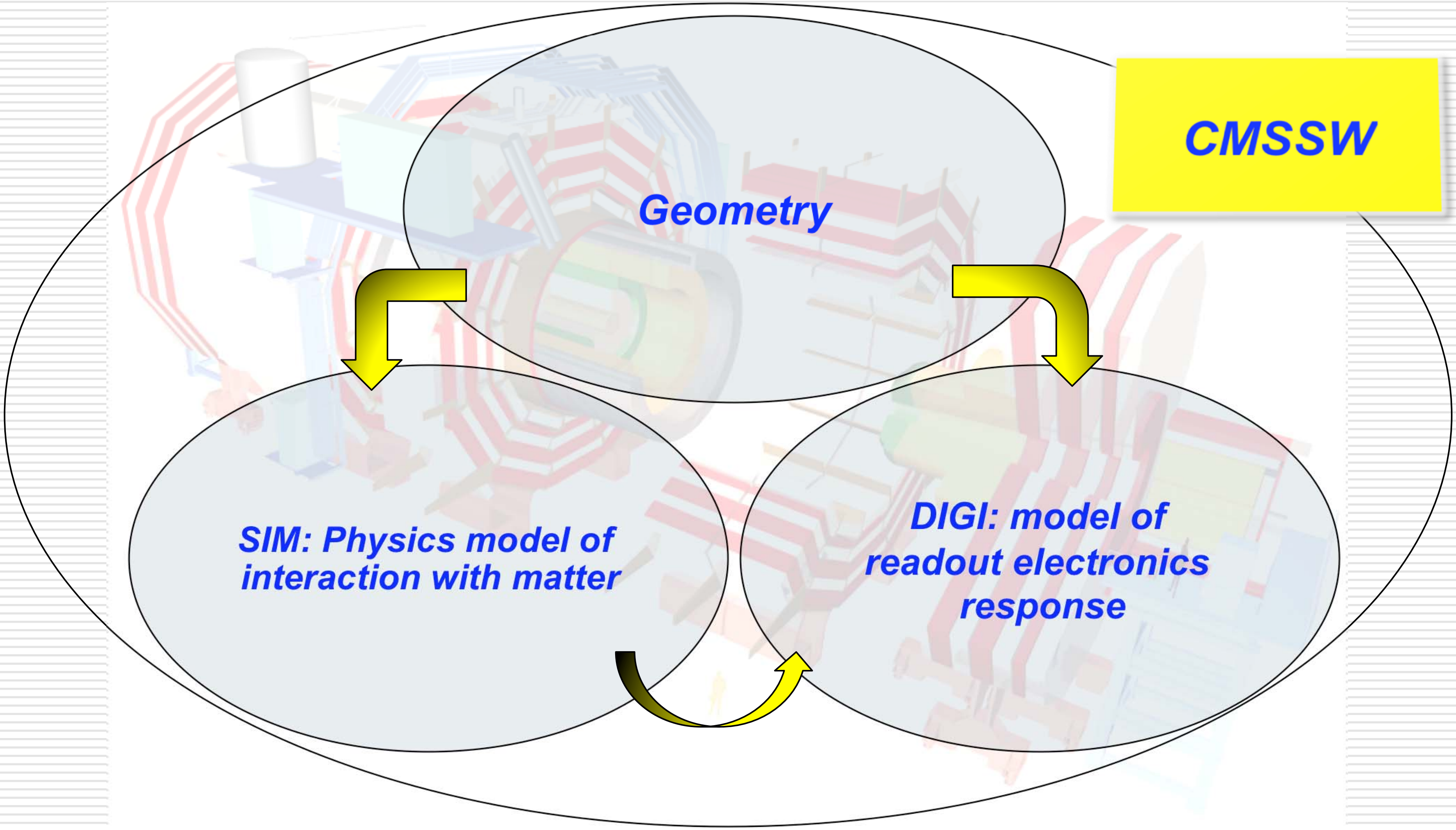
Tuning and optimization of the CMS simulation software

Fabio Cossutti - INFN Trieste
on behalf of the CMS Offline project

Computing in High Energy and Nuclear Physics 2009
Prague, 24 March 2009

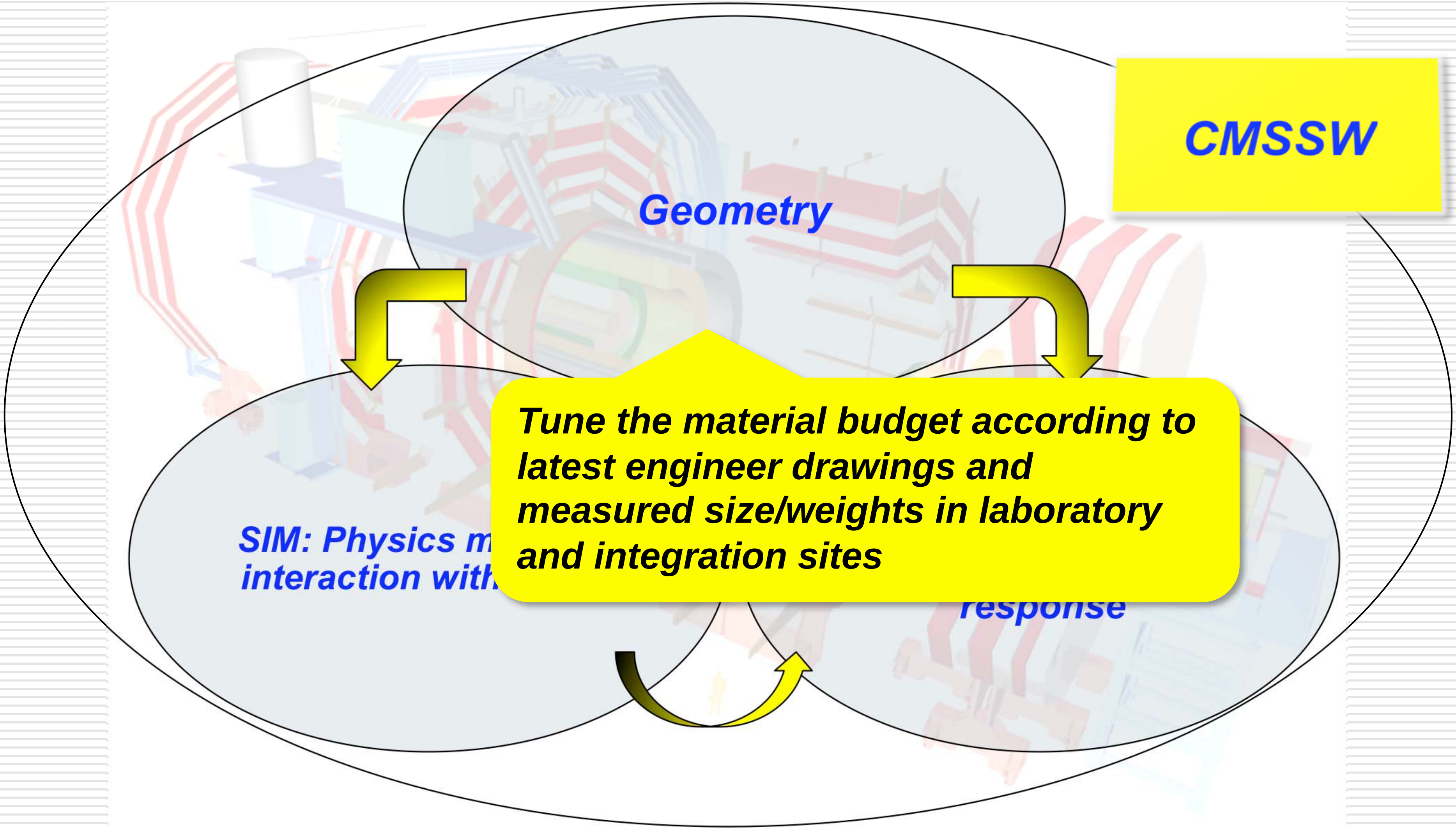
The CMS Simulation



Commissioning the simulation

- New target in recent times
  - from computing challenges and TDRs preparation towards comparison with real data and routine mass production dedicated to this task
- Validation and tuning on real data is a fundamental step
  - To achieve the best agreement and make the tool really usable and reliable for physics analysis
- Production mode implies computational optimization
  - To constantly fit the computing model constraint during code evolution
  - Physics tuning result often turns out to be expensive in terms of computing performances

Simulation tuning

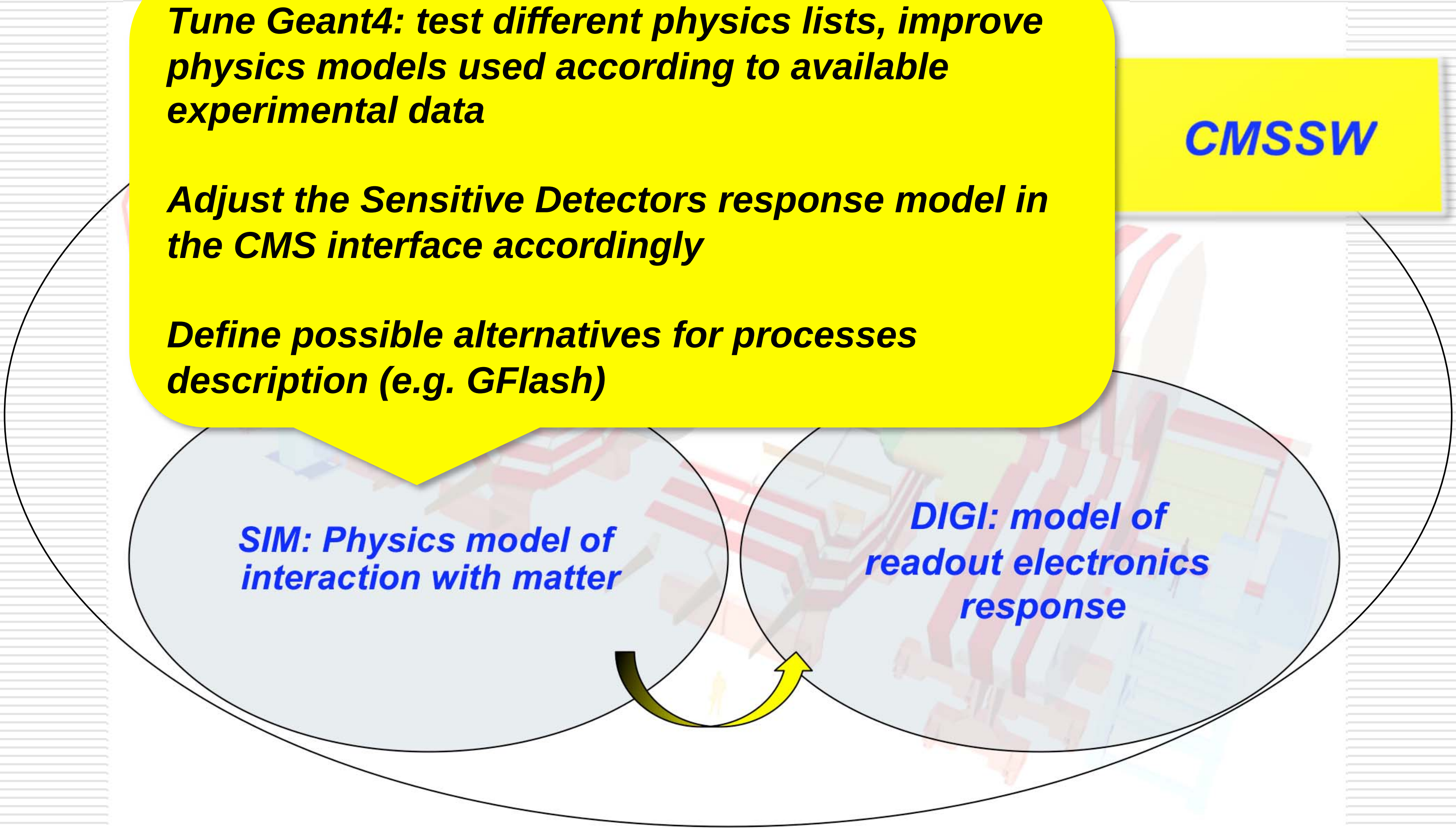


Simulation tuning

Tune Geant4: test different physics lists, improve physics models used according to available experimental data

Adjust the Sensitive Detectors response model in the CMS interface accordingly

Define possible alternatives for processes description (e.g. GFlash)

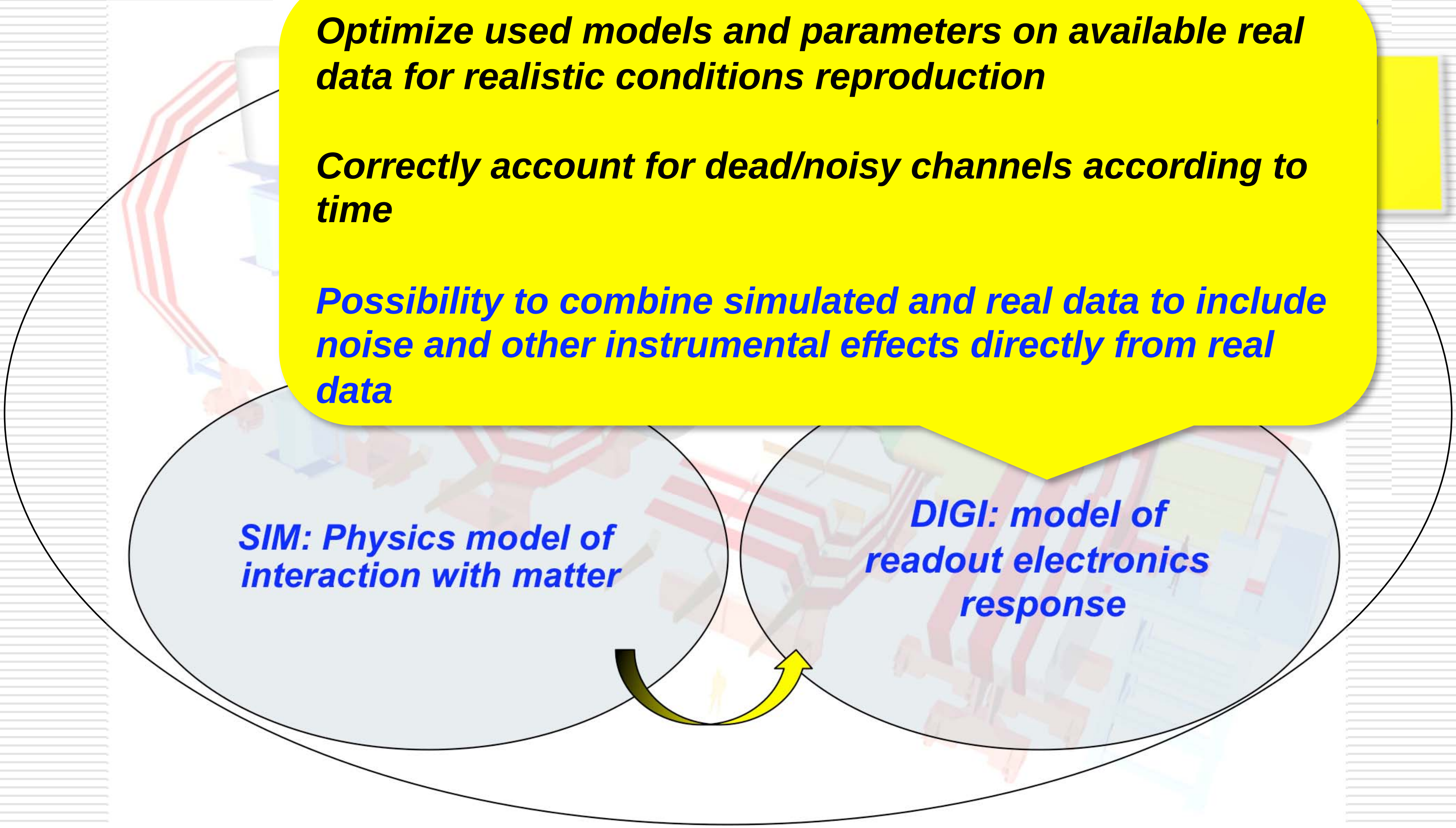


Simulation tuning

Optimize used models and parameters on available real data for realistic conditions reproduction

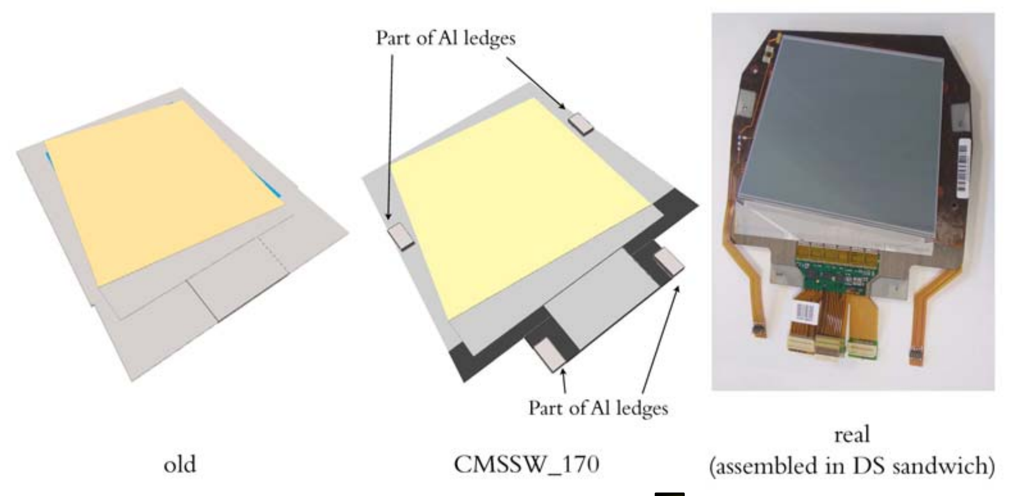
Correctly account for dead/noisy channels according to time

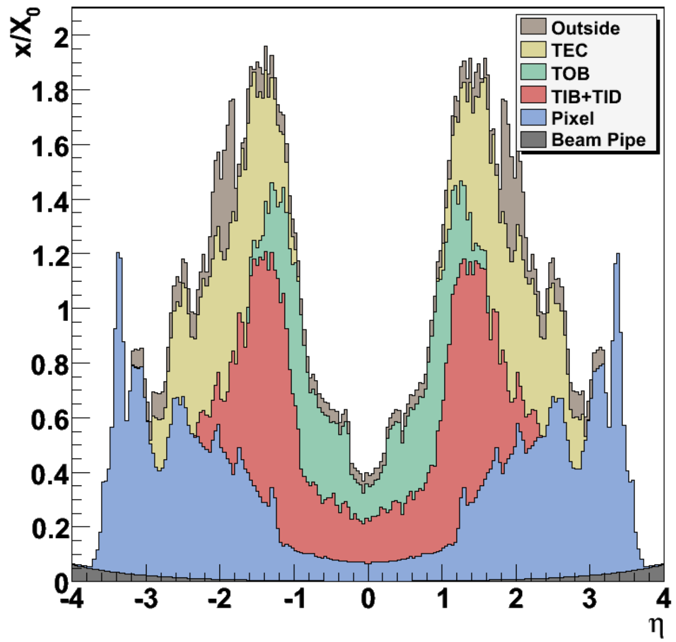
Possibility to combine simulated and real data to include noise and other instrumental effects directly from real data



Material budget: Tracker

- Long review of each component description:



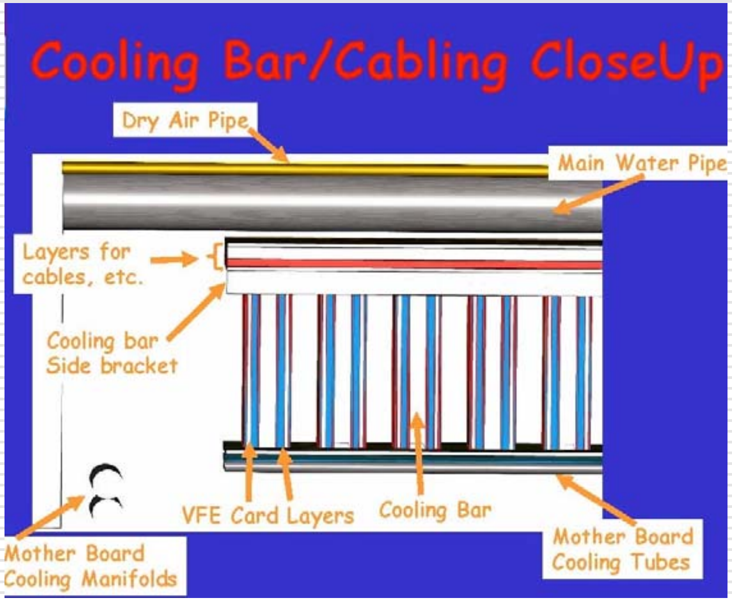


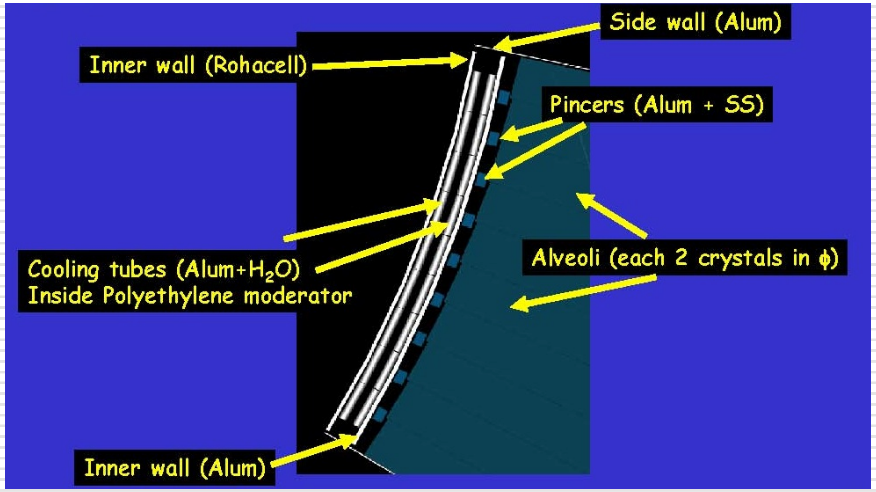
Tuning piece by piece

	Weight	Uncertainty	
TK in simulation	3963		
cooling fluid	-153	-53	77
pixel tube and beam pipe supports	45		
services from TST to PP1	670	-10	20
installation and packaging stuff	1648	-100	100
Total ***Preliminary***	6173	-163	197

Compare with measured weight at transport to P5: ~6350 Kg ➔ ~ 3% agreement

Material budget: ECAL

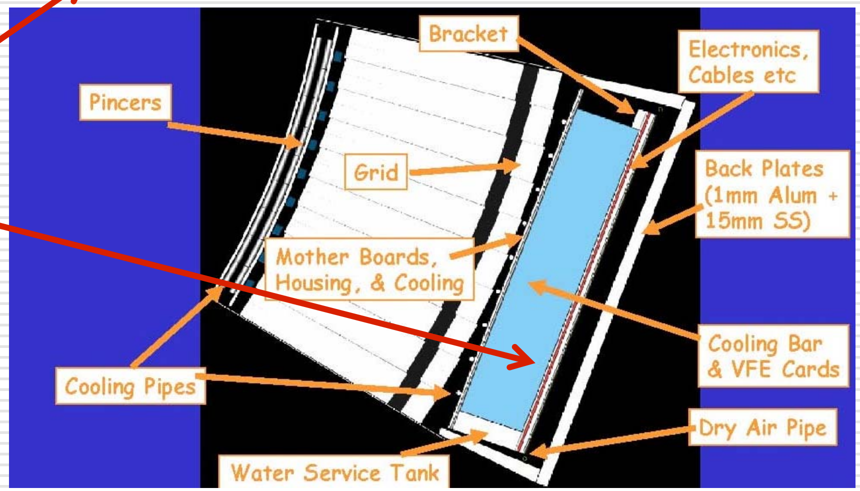




Crystals well known, tune passive material

In front crystals (for e.m. showers)

Behind crystals (for hadronic showers)

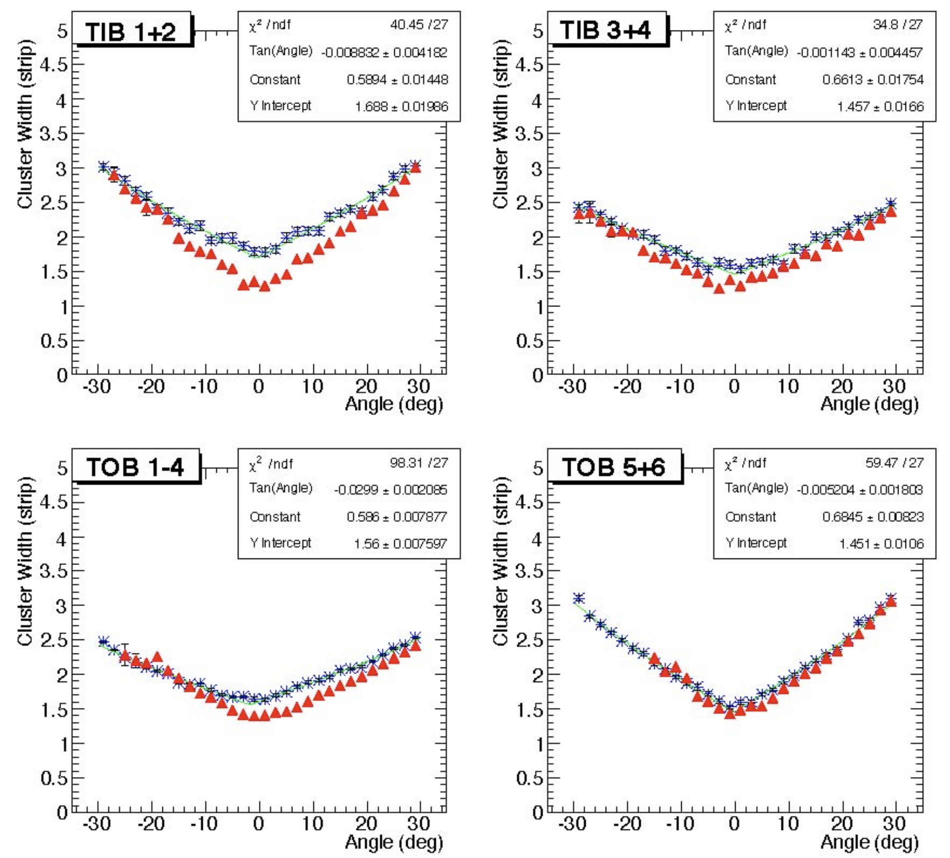


Passive material (cooling bars, electronic boards, cables) weights tuned against measurements in laboratory

Agreement within 3%

Tuning of electronics simulation

- Cosmic events useful to tune electronics response
  - Data-driven simulation of noise, sensitive detector characteristics,...
- Example: Strip Tracker

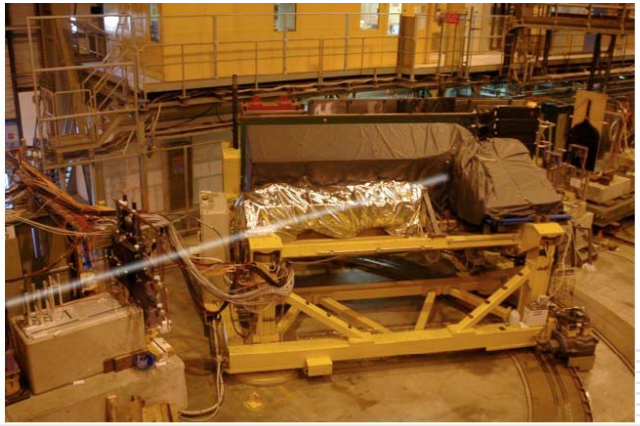


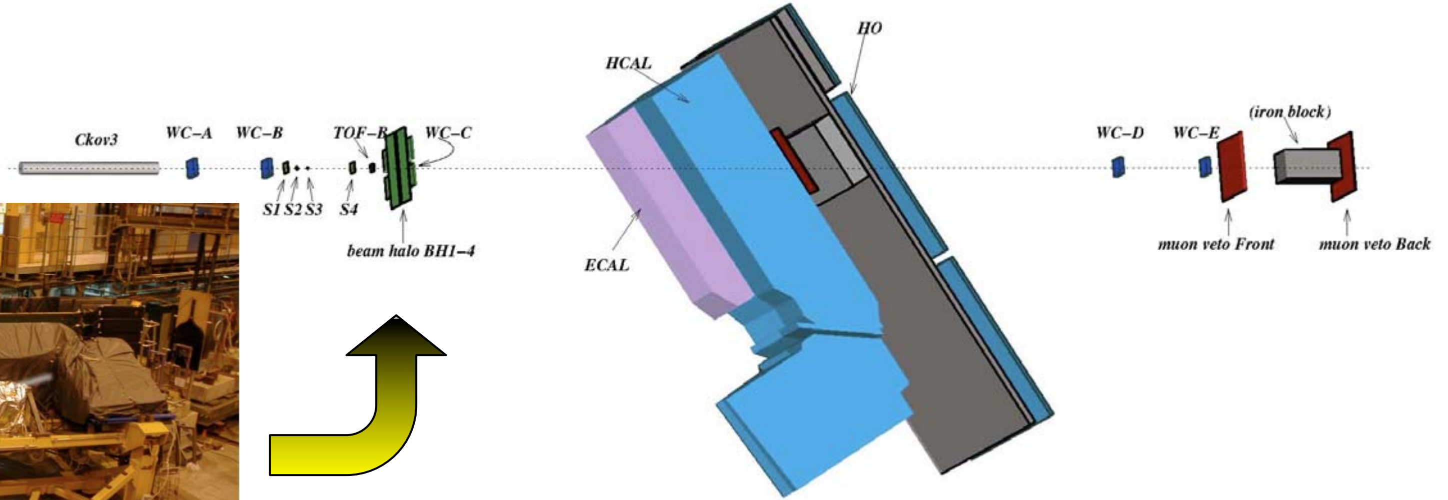
Adjust capacitive coupling between strips according to cosmic data taken at the integration facility

Effect on cluster width:
data (✖) vs tuned MC (▲)

Tuning of model of interaction with matter

- Mainly tuning of Geant4 itself
- Exploit available real data
  - E.g. test beams of calorimeters
  - Most results from 2006 ECAL (H4 beam line) and ECAL +HCAL (H2 beam line)





Dedicated simulation of the TB setup (beam line, trigger)

Problems in comparison with data

- At the beginning of 2008 a CMS Calorimetry Task Force was setup to work on tuning of simulation

- Long list of discrepancies, when using Geant4.8.3p01, physics list QGSP_EMV:
    - Electromagnetic showers transverse profile reproduced at 0.5%, but fluctuating version after version
  - Most of the problems in the hadron interaction sector:
    - simulated showers are narrower and shorter
    - Mean energy from hadronic showers overestimated
    - Early interactions in ECAL of hadrons overestimated

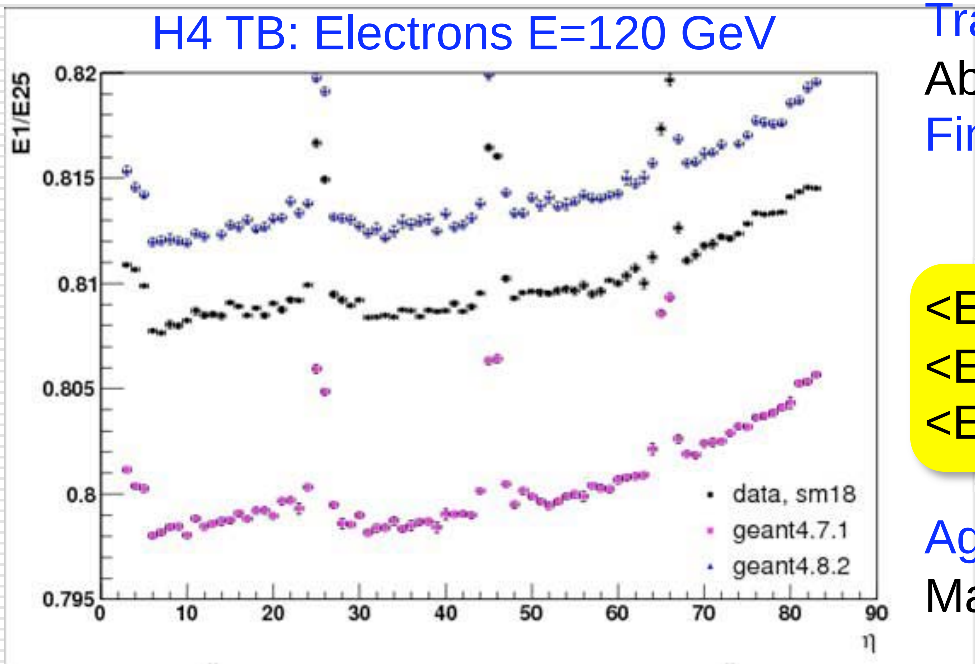
- Very close and productive collaboration with Geant4 development, driven also by CMS problems

Improving Geant4

- Move to G4.9.1p02 and then G4.9.2p01
- Adopt QGSP_BERT_EMV
  - Electromagnetic sector:
    - Improved multiple scattering, relativistic bremsstrahlung
    - Bremsstrahlung and pair production from hadrons
    - Better physics table interpolation
  - In CMS code: adopt Birk's law for scintillators saturation effects
  - Hadronic sector
    - Improved quasi-elastic scattering in QGS model at high energy
    - Improved quasi-elastic scattering in Bertini model at low energy
    - Bertini cascade improved in cross section
      - Many other improvements...

See also S. Banerjee's talk in this session

Electromagnetic showers



Transverse shape: η behavior ok

Absolute value fluctuating with versions

Final revision of multiple scattering in 9.2

separate tuning for particle species

$<E_1/E_{25}>_{DATA}$ = 0.8072±0.0005
$<E_1/E_{25}>_{9.1p01}$ = 0.8176±0.0003
$<E_1/E_{25}>_{9.2}$ = 0.8130±0.0003

Agreement at ~ 0.5% in G4.9.2

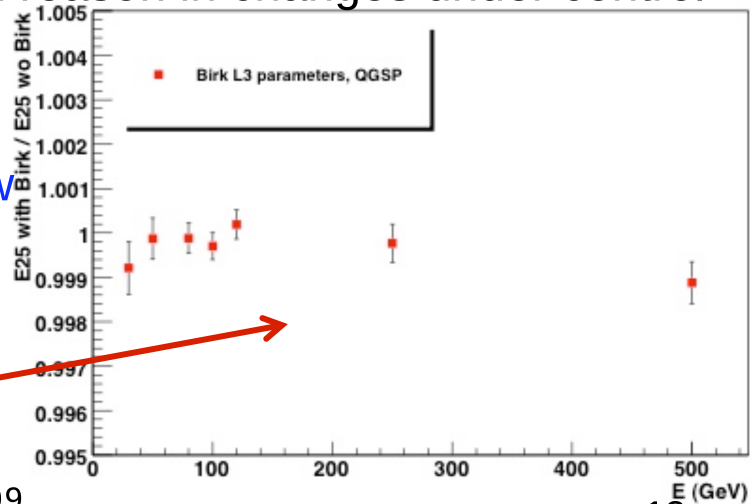
Main reason in changes under control

To improve description of hadronic shower in ECAL: need scintillation saturation, e.g. Birk's law

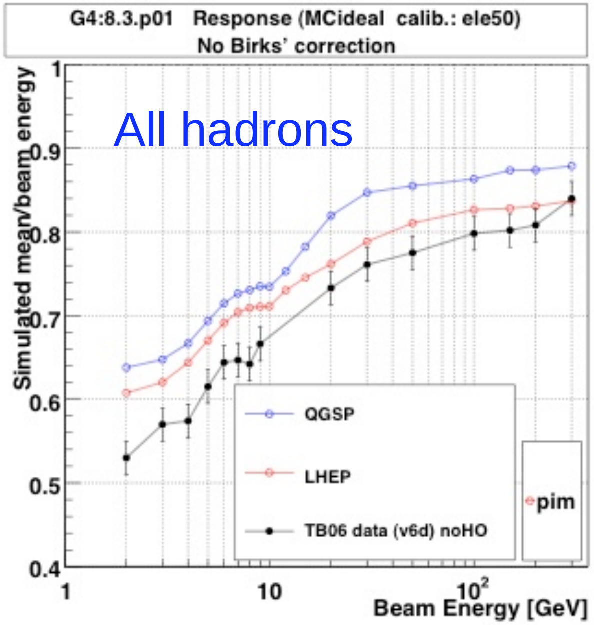
No measurement of parameters for $PbWO_4$

use BGO ones in L3 parameterization

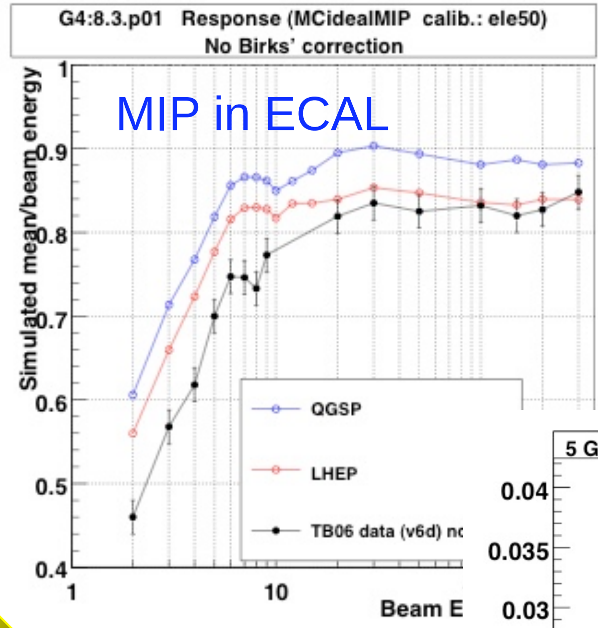
No effect on well understood e.m. showers



Hadronic showers: energy response

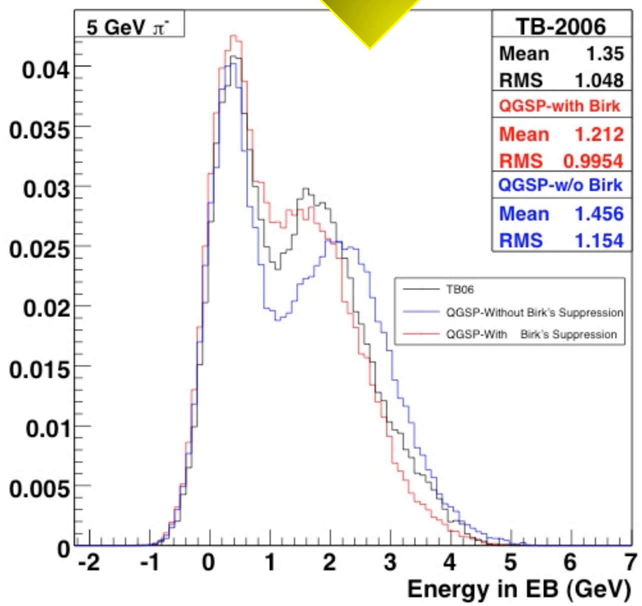


All hadrons



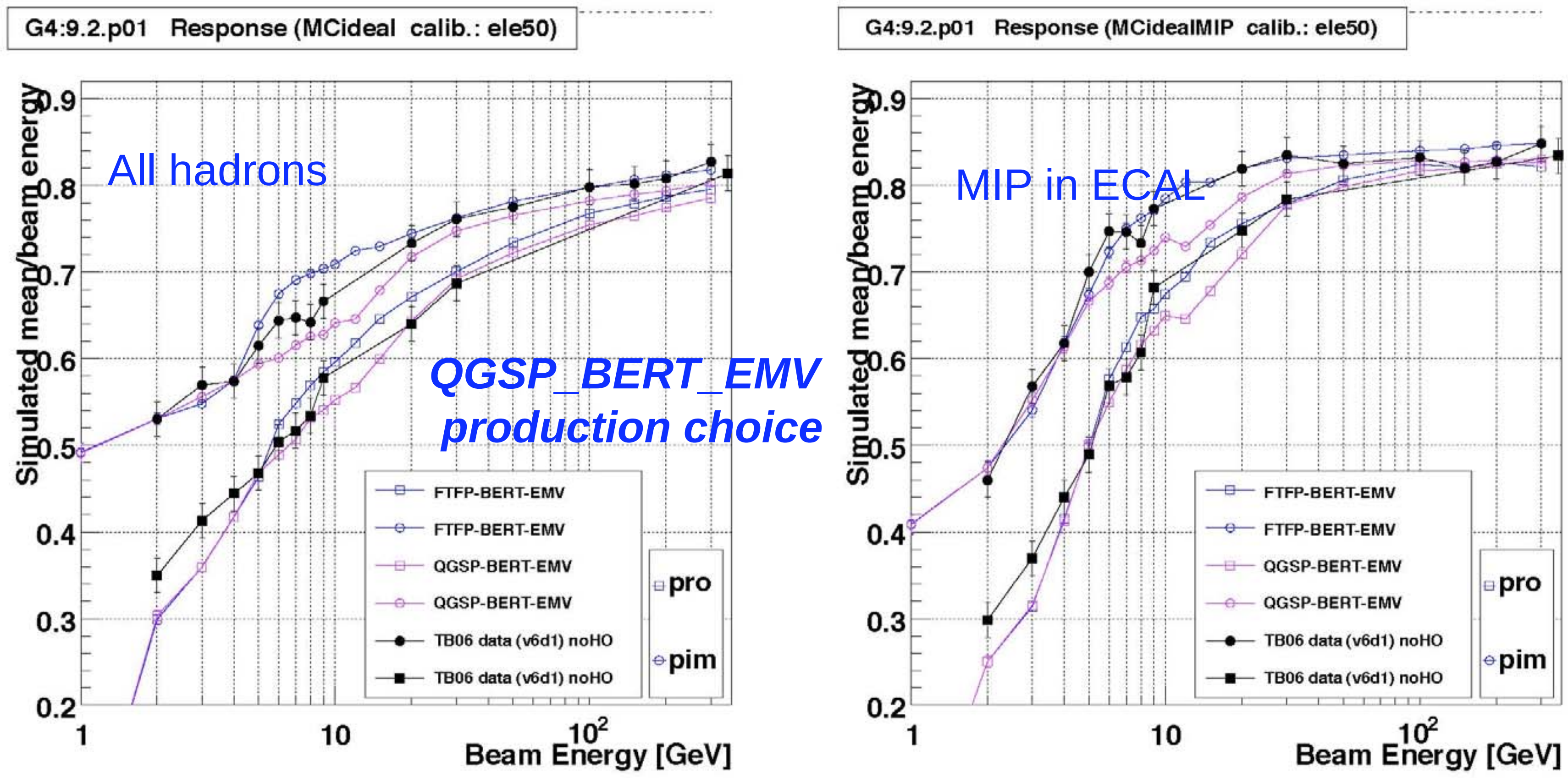
MIP in ECAL

Saturation effect in scintillator can partly explain behavior at low energy

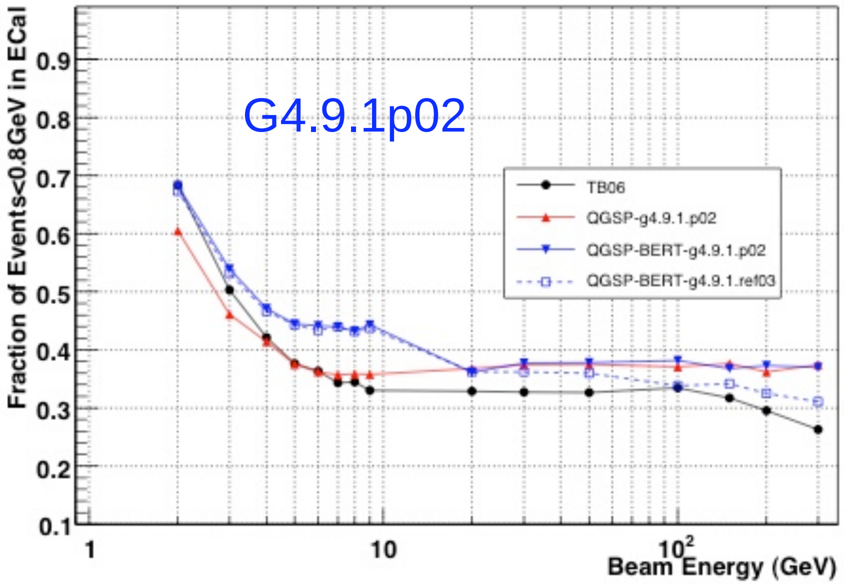


In G4.8.3p01: no used physics list (LHEP, QGSP) could reproduce data

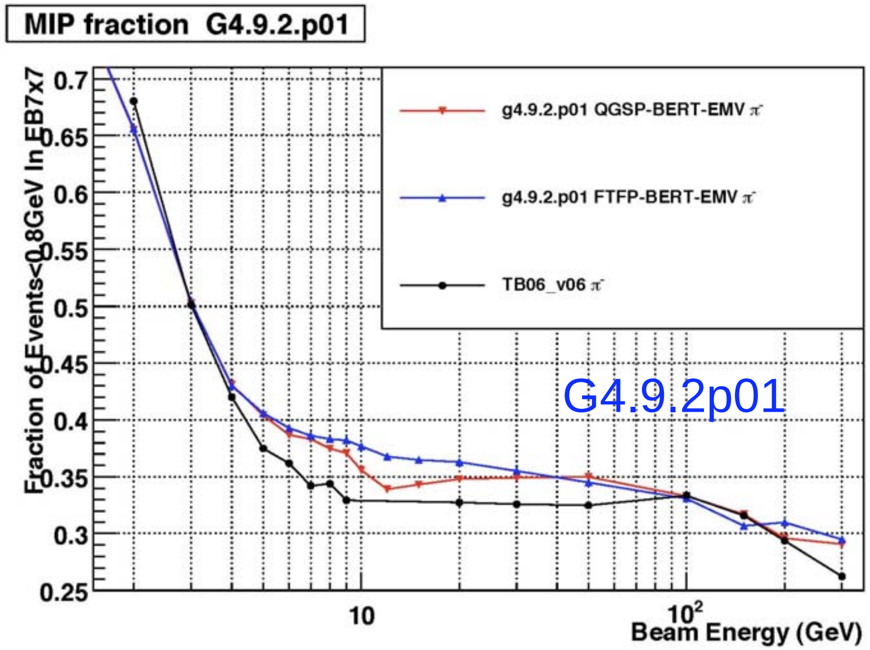
Hadronic showers: energy response



Hadronic showers: MIP in ECAL

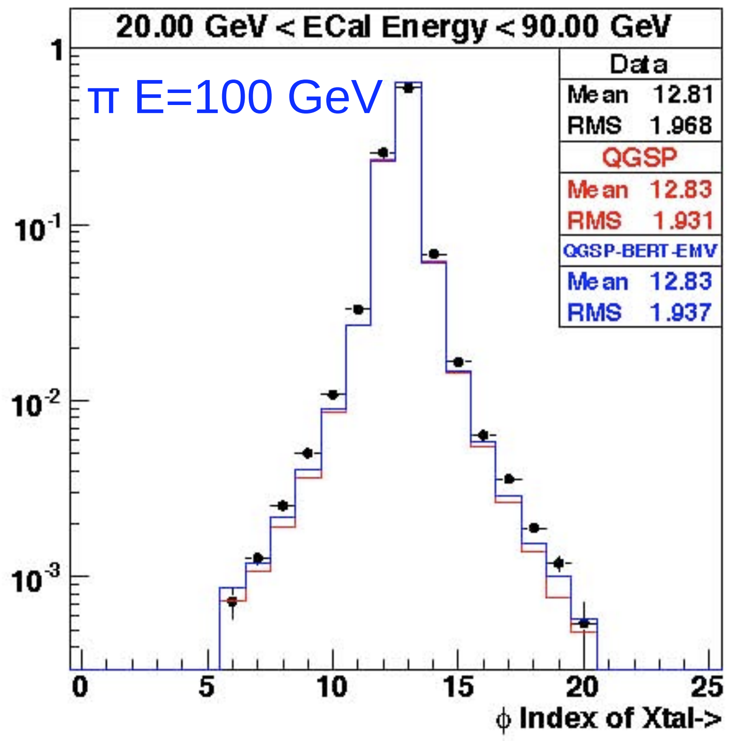


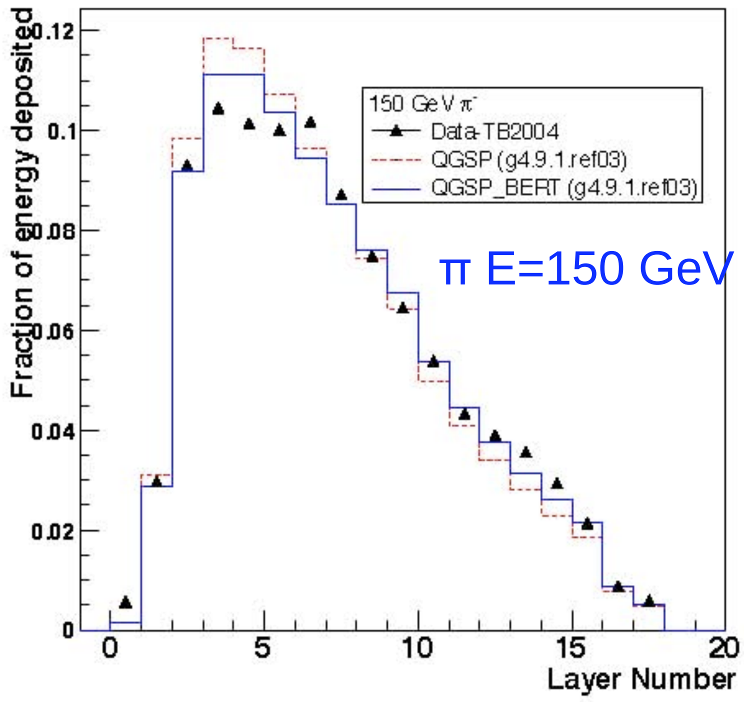
Discrepancy at high energy:
Pion bremsstrahlung



Discrepancy below 10 GeV:
Quasi-elastic scattering in Bertini cascade

Hadronic shower shapes

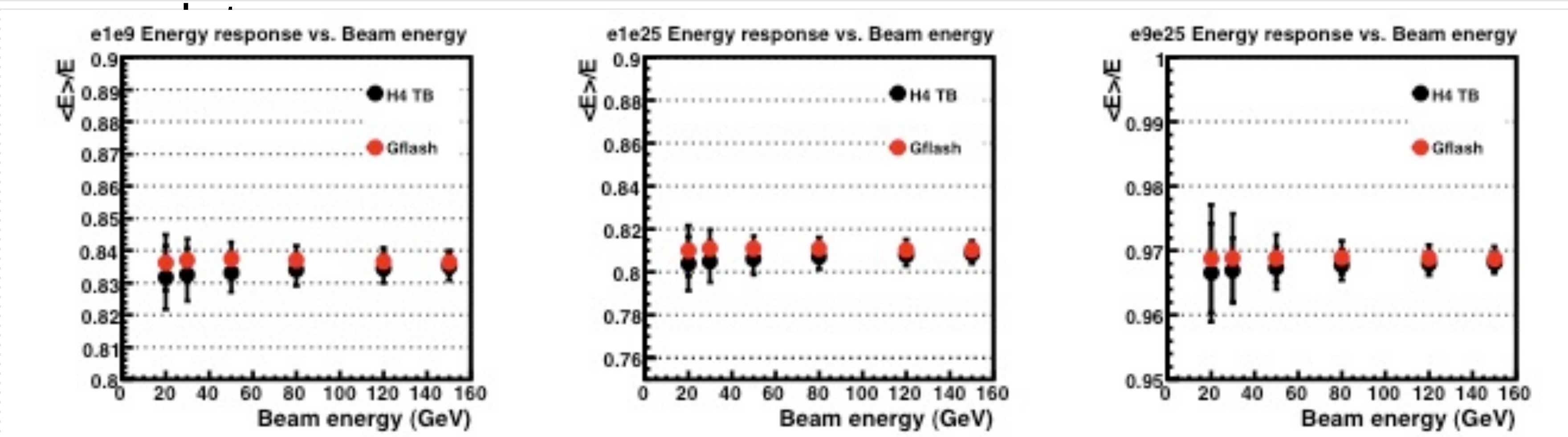




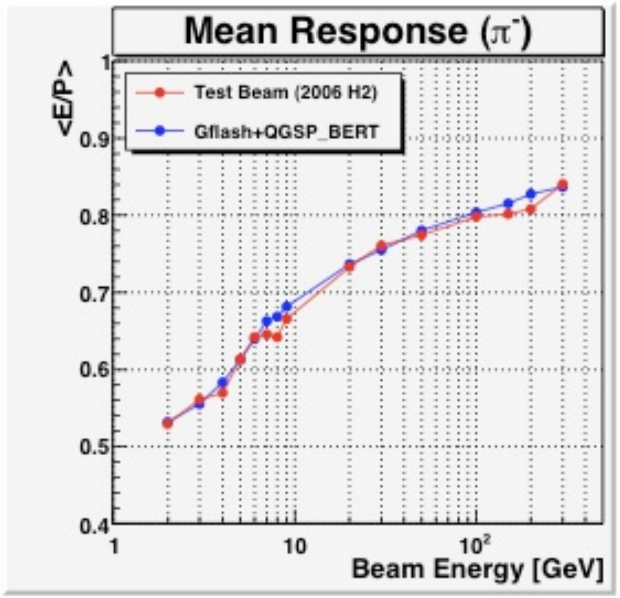
The adoption of the Bertini cascade definitely improve the agreement of the shapes with data

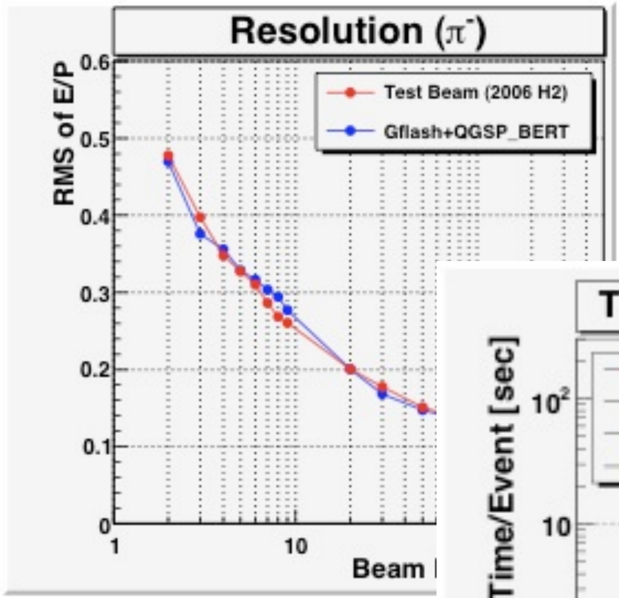
GFlash

- Add the possibility to replace the G4 calorimetric showers with a fast and easily tunable parameterization
  - Fully interfaced with the standard Geant4 simulation
  - Adapt the H1 original approach to CMS geometry, both for the homogeneous and the sampling part
  - Both electromagnetic and hadronic showers, tuned on TB

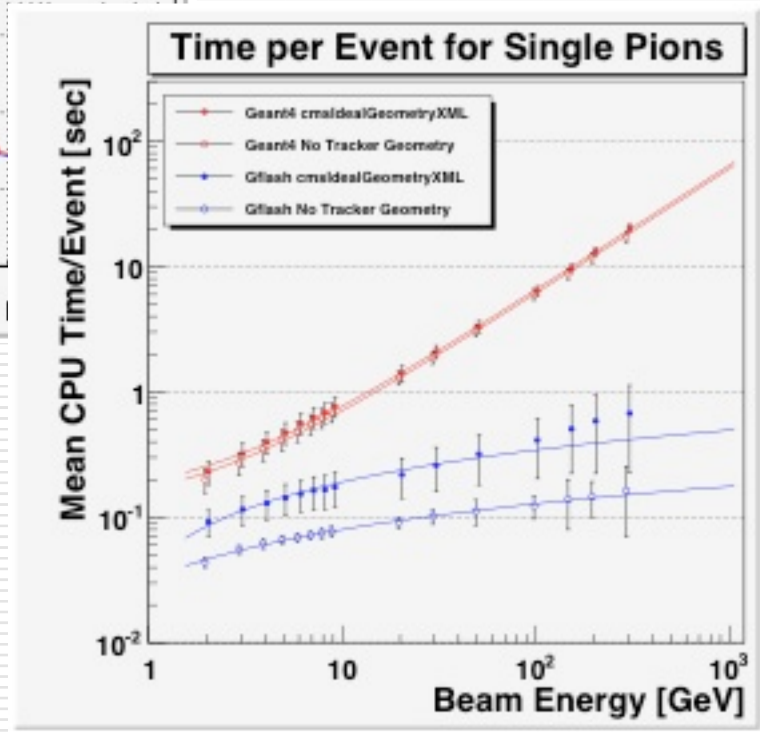


GFlash





Main reason to study this approach:
easily tunable on data



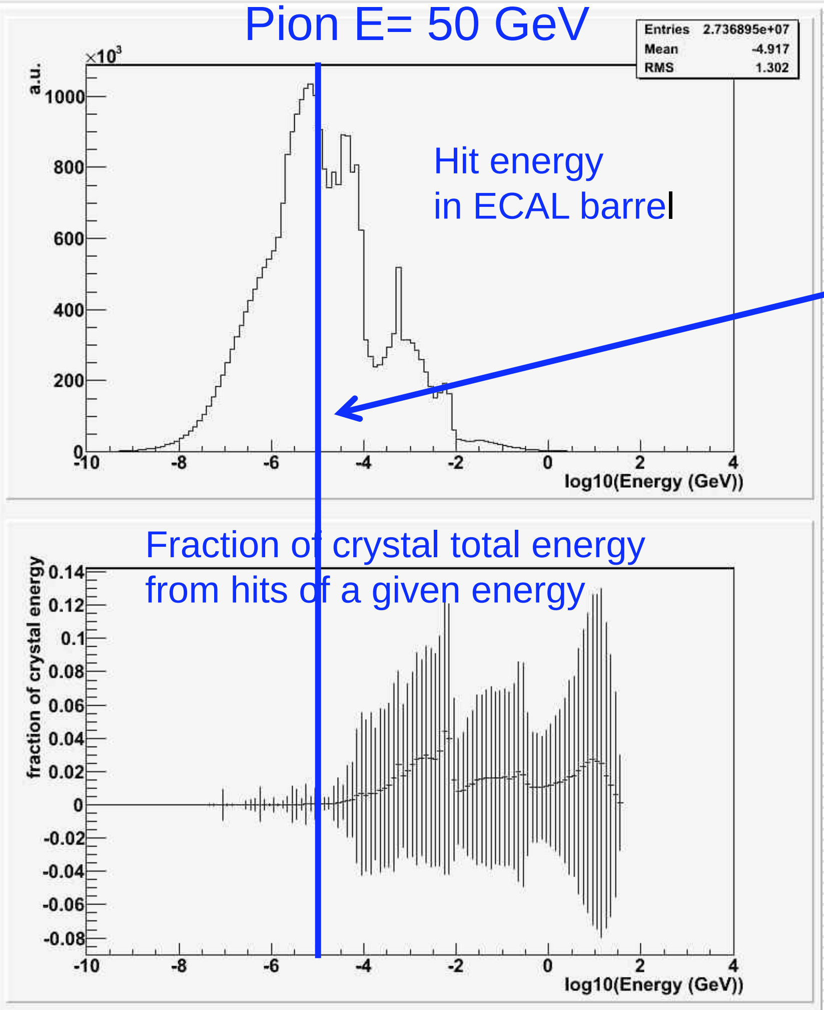
Potential high gain in simulation speed
In complex events: it depends on how the energy cut to trigger the shower is set

Performance issues

- Extensive studies of the code performances with many different tools (see P. Elmer's talk)
  - CPU, memory footprint, dynamic allocations, output size...
  - Constant monitoring during development
  - Keep the memory footprint in the 1 GB constraint from Computing TDR
    - Important for high multiplicity events, pileup, heavy ions collisions
  - Additional problems from physics tuning:

CPU (relative)	QGSP_EMV	QGSP_BERT_EMV	QGSP_BERT
Minimum Bias	1.00	1.41	1.58
TTbar	1.00	1.46	1.69
Event size (relative)			
Minimum Bias	1.00	1.52	1.52
TTbar	1.00	1.77	1.72

Performance optimization



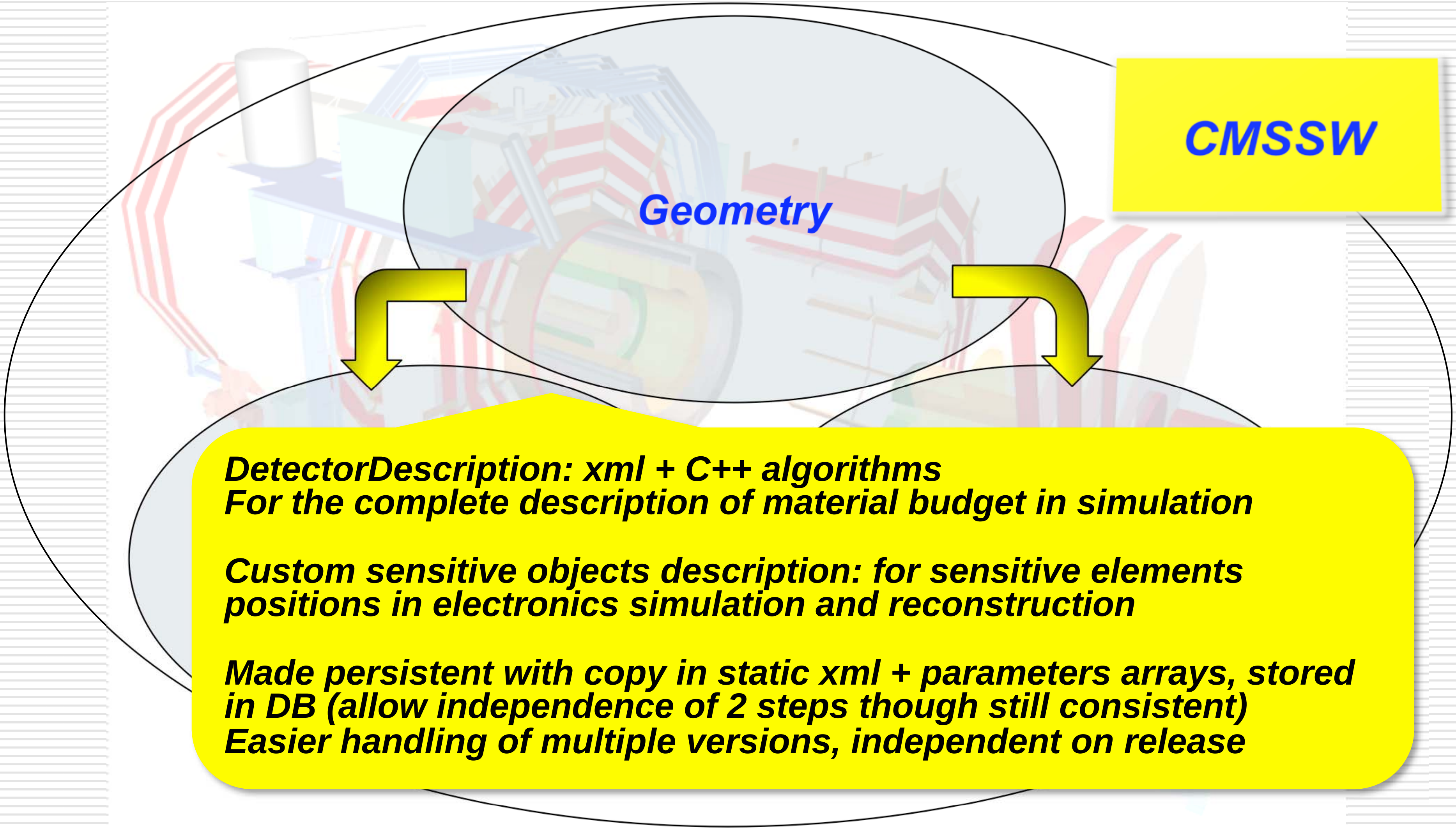
- Main problem in Bertini: many very low energy hits
  - Cut those with no impact on observable energy
  - Reduce by 20-30% overall simulation output size
- Move the track and calorimeter hits management on a primary-by-primary basis
  - Reuse of memory released at every new track/hit
  - Gain in memory footprint: O(> 50 MB) on TTbar run

Conclusions

- An extensive program of tuning of the full simulation has been carried on by CMS
  - Basis for next round of tuning on real collision data as soon as available
  - Very fruitful collaboration with Geant4 developers
- Attention to balance physics improvements and computing performances

BACKUP SLIDES

The CMS Simulation



DetectorDescription: xml + C++ algorithms
For the complete description of material budget in simulation

Custom sensitive objects description: for sensitive elements positions in electronics simulation and reconstruction

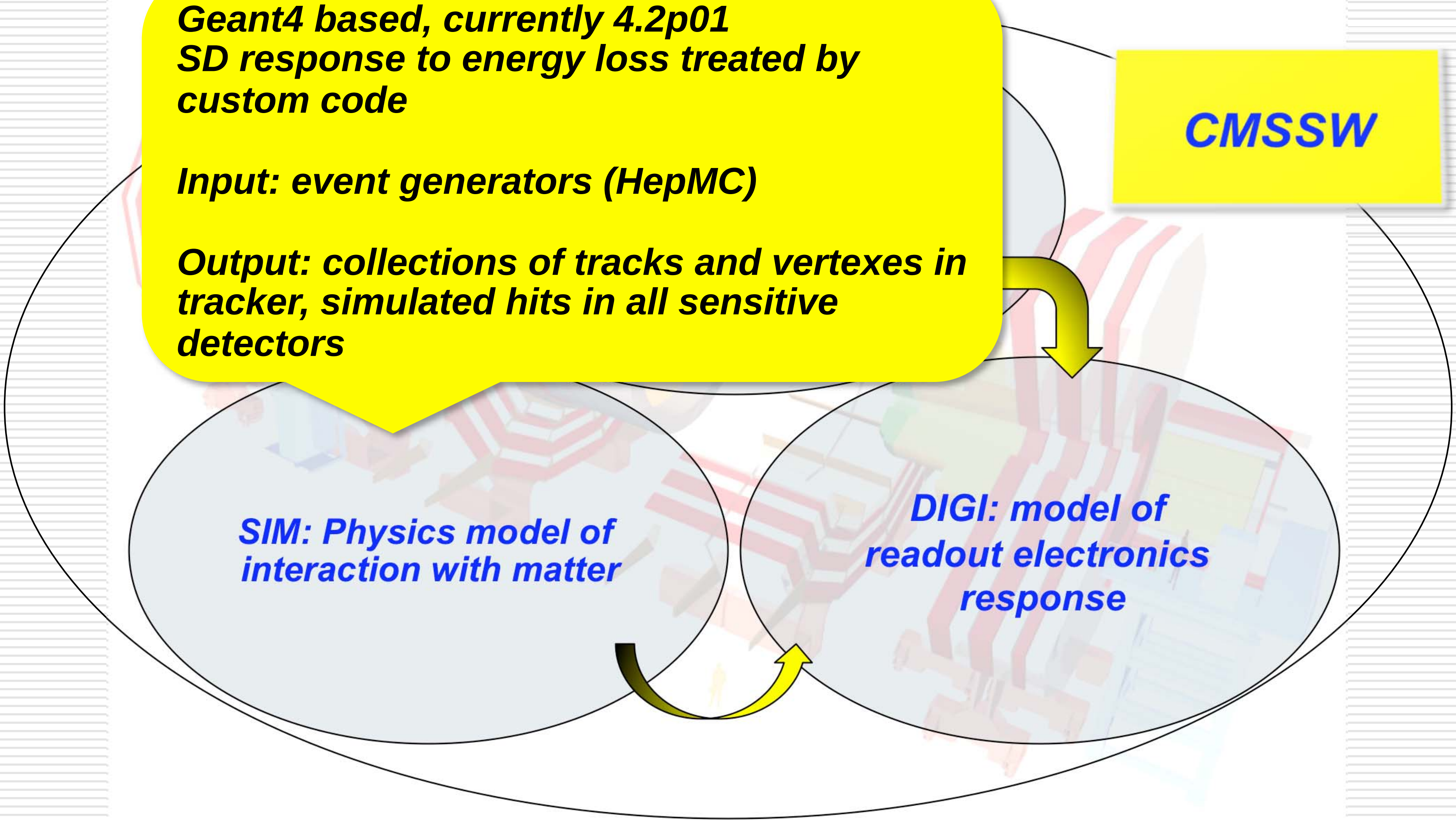
Made persistent with copy in static xml + parameters arrays, stored in DB (allow independence of 2 steps though still consistent)
Easier handling of multiple versions, independent on release

The CMS Simulation

Geant4 based, currently 4.2p01
SD response to energy loss treated by custom code

Input: event generators (HepMC)

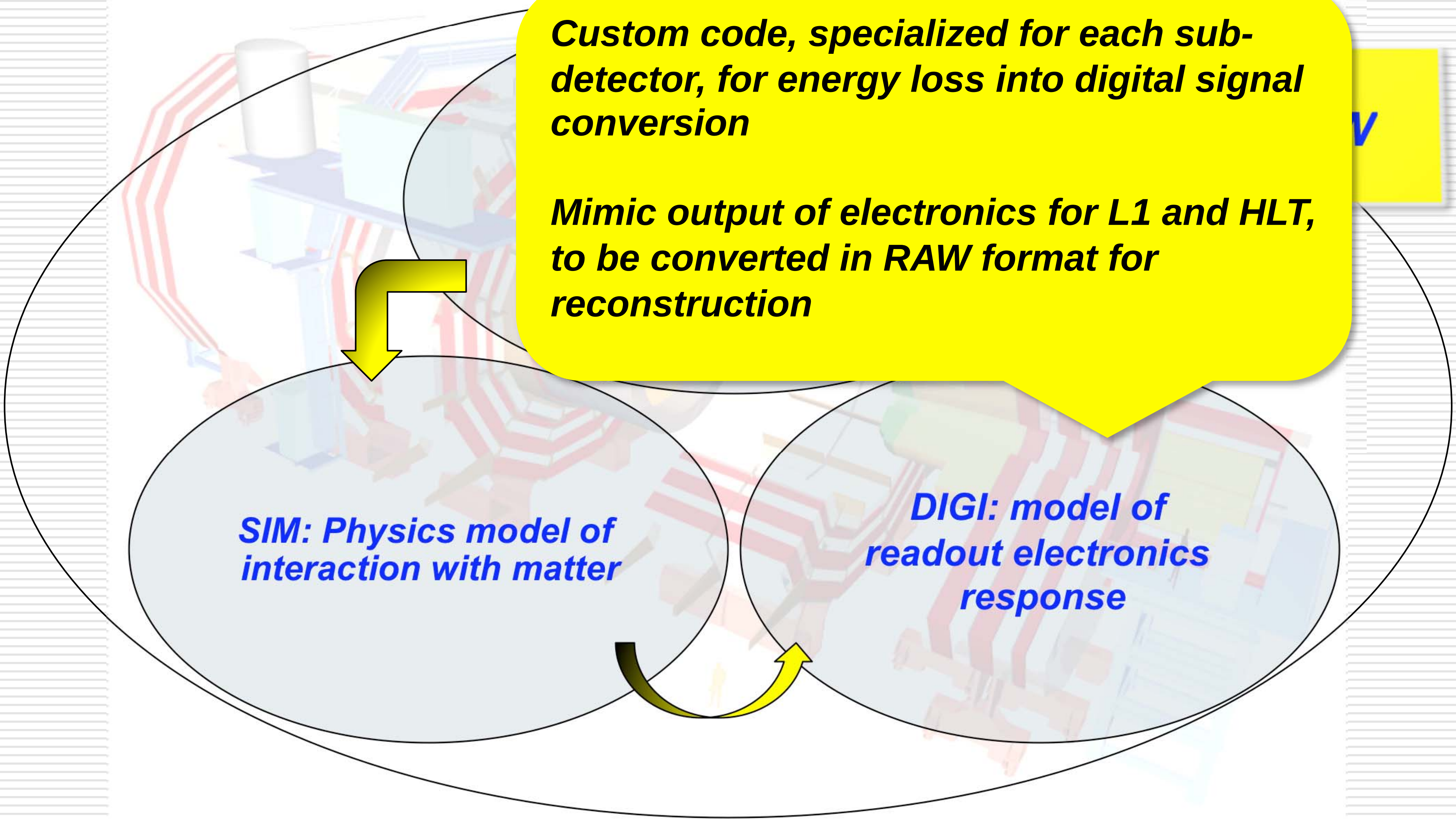
Output: collections of tracks and vertexes in tracker, simulated hits in all sensitive detectors



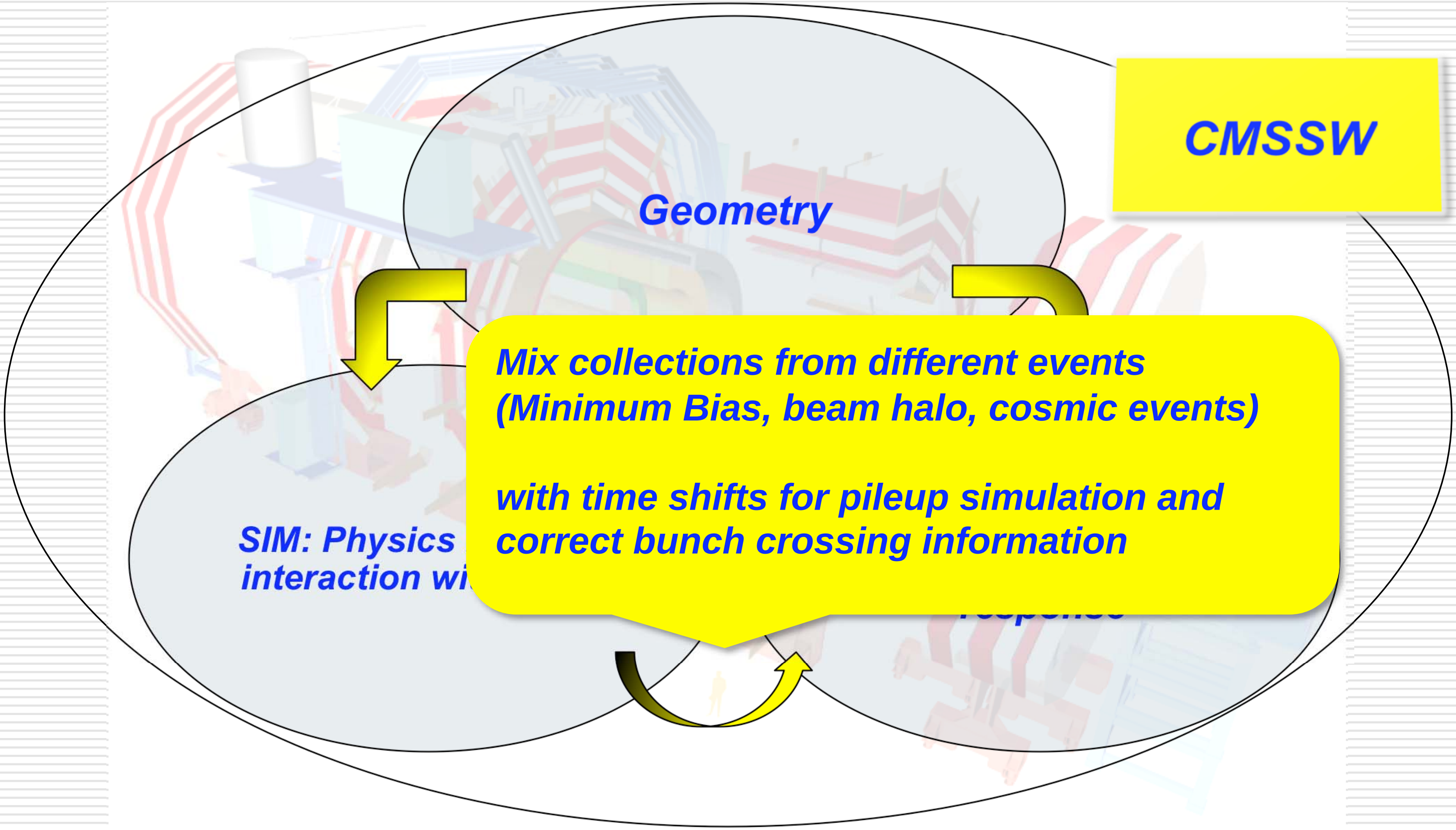
The CMS Simulation

Custom code, specialized for each sub-detector, for energy loss into digital signal conversion

Mimic output of electronics for L1 and HLT, to be converted in RAW format for reconstruction



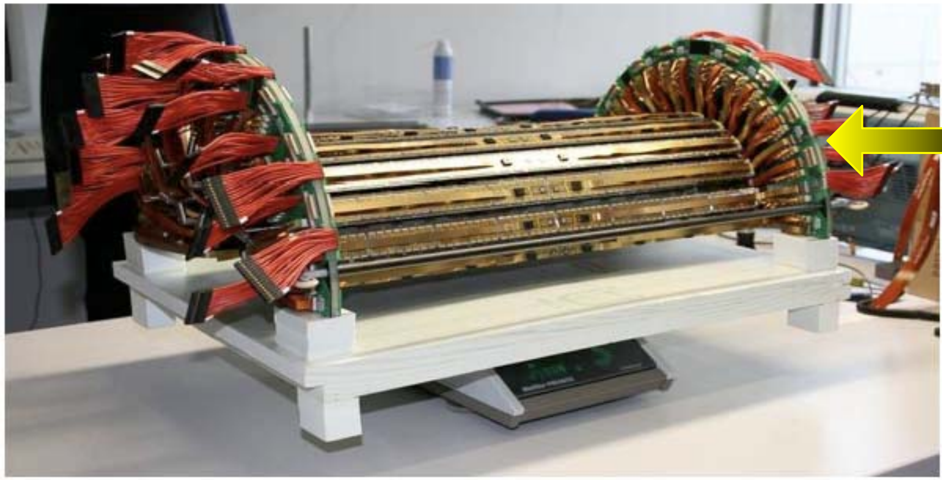
The CMS Simulation



Mix collections from different events
(Minimum Bias, beam halo, cosmic events)

with time shifts for pileup simulation and correct bunch crossing information

Material Budget: Pixel

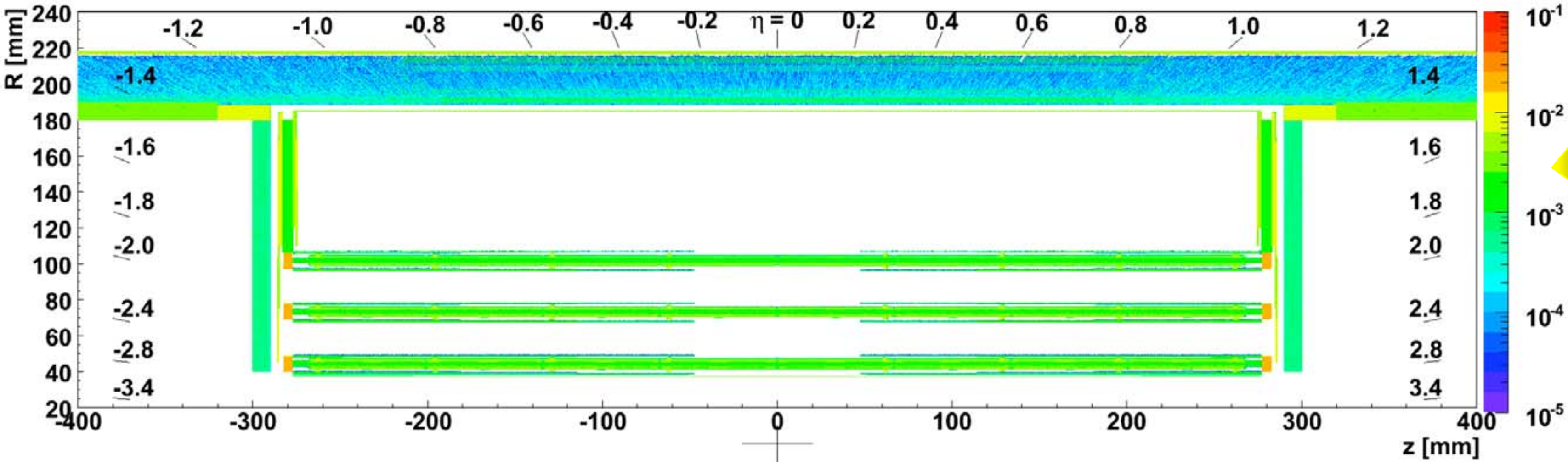


Preliminary

Measured: 2598 g
Simulation: 2455 g

(no cables, coolant)

Radiation length distribution seen by a particle



Simulation computing performances

Measurements performed on CERN build machine
Intel Xeon 5160 @ 3 GHz (dual processor dual core)
RAM 8 GB
1 core used, the other kept busy with scimark benchmarking code

Using Geant4 9.2

	SIM (CPU s/ev)	**DIGI (CPU s/ev)**
Single muon p_t=10 GeV	0.53	0.95
Single electron E=1 TeV	115.01	0.92
Single pion E=1 TeV	70.80	1.03
Minimum bias	12.00	1.02
TTbar	104.09	1.95