FAST SIMULATION IN CMS/ATLAS

Andrea Perrotta (INFN Bologna)

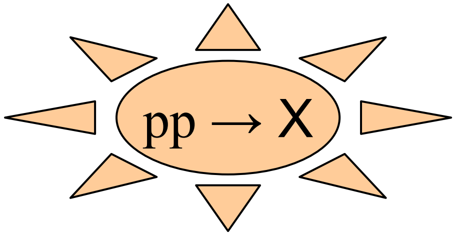
Secondo workshop INFN sui MC
Frascati, 22-24 Maggio 2006

Tutto quello che avreste sempre voluto sapere sulla simulazione veloce dei rivelatori a LHC ma non avete mai osato chiedere…

Outline

- What is a detector (full) simulation at LHC?
- What is a fast simulation at LHC?
- Fast simulation of the ATLAS detector
- Fast simulation of the CMS detector
- Comparisons
- Conclusions

What happens in the real life



Particles from the pp collision interact in the detector material

Sensitive areas of the detector produce electronic signals as result of the particle-material interaction

Fast hardware devices (Trigger) use (part of) those signals to decide whether to accept the event for the subsequent steps

DAQ system read out those electronic signals and write them into a support by using an appropriate format (Hits)

Hits are combined to produce intermediate analysis objects

Intermediate objects used to build higher level analysis objects

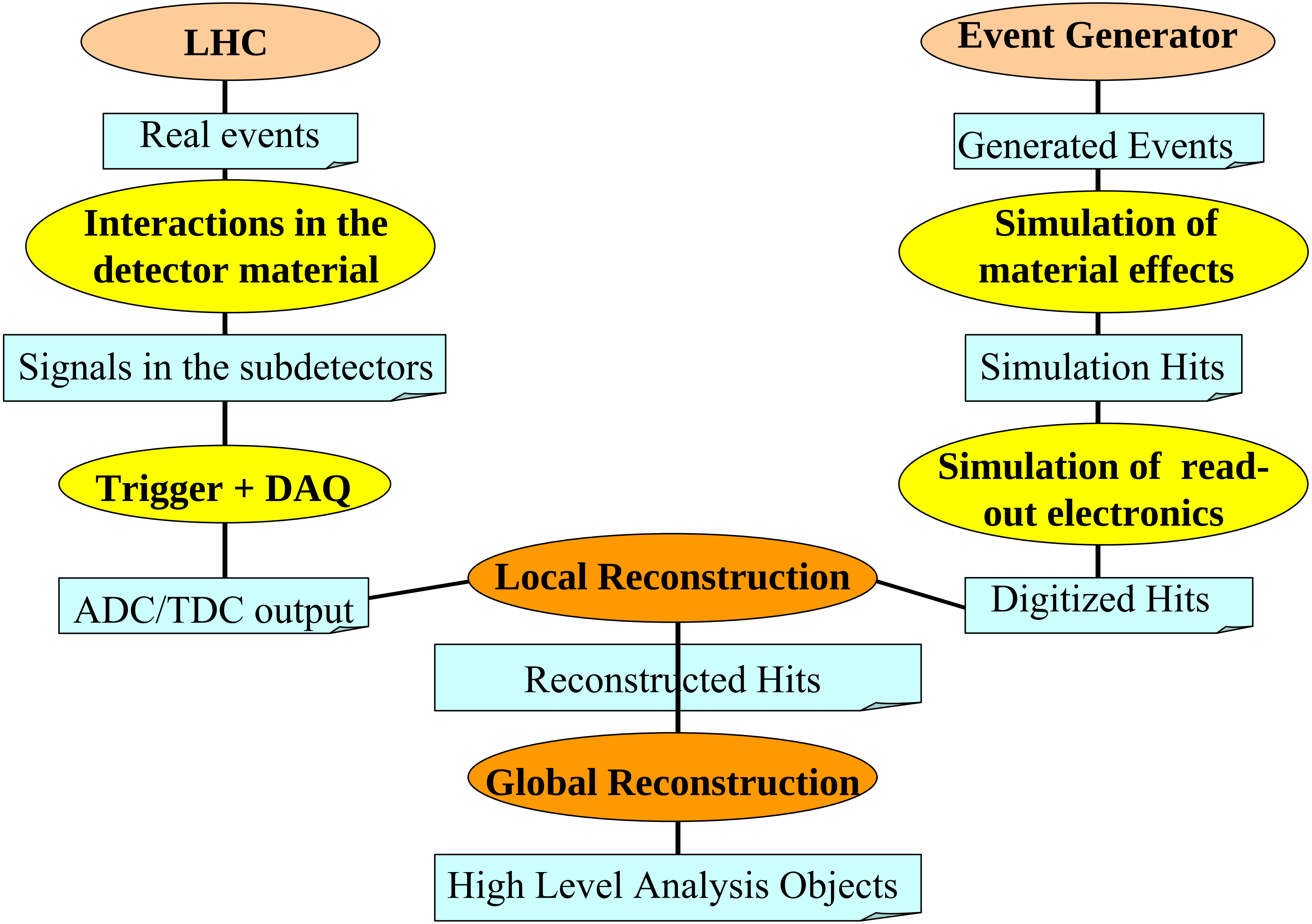
High level objects used for the final physics analysis

Detector (full) simulation

In the **detector simulation** one tries to simulate the result of all the physics processes and intermediate steps that lead from the four vectors (and vertices…) of the generated particles to the final analysis objects.

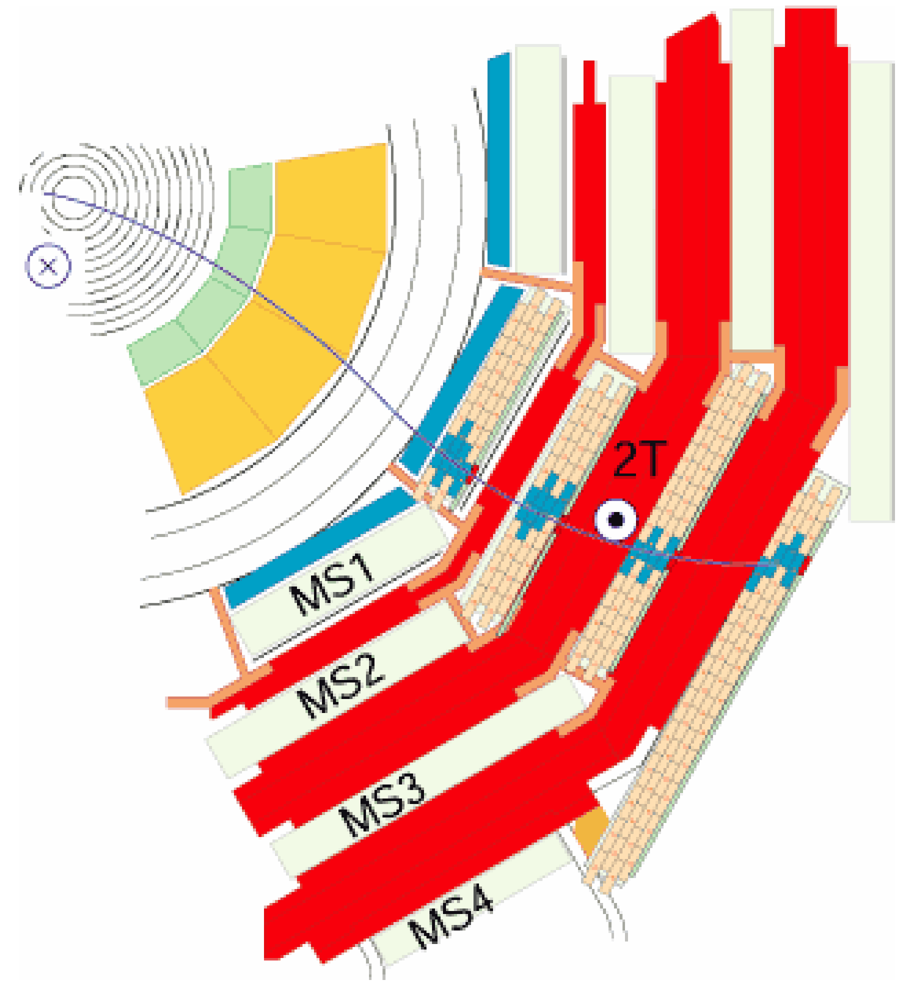
If we provide a simulation of the electronic signals as given by the true detector elements, the same **reconstruction** software as used in the real data can be applied to simulated data.

Effects as electronic noise in the detectors, event overlapping due to pile-up, dead-times, etc. can be well described with an accurate simulation to give realistic results for the reconstructed analysis objects.


LHC
Real events
Interactions in the detector material
Signals in the subdetectors
Trigger + DAQ
ADC/TDC output
Event Generator
Generated Events
Simulation of material effects
Simulation Hits
Simulation of read-out electronics
Digitized Hits
Local Reconstruction
Reconstructed Hits
Global Reconstruction
High Level Analysis Objects

SIMULATION OF THE DETECTOR

How does a detector simulation act on particles from the MC generator to produce the final analysis objects post reconstruction?

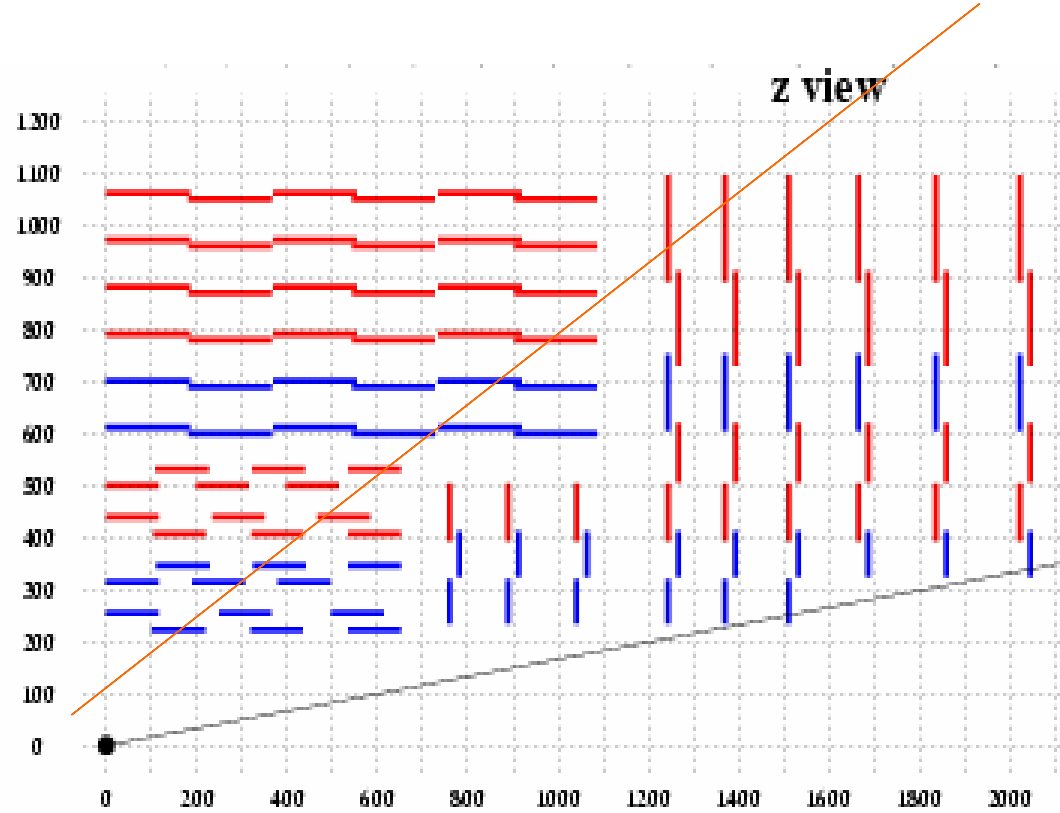


Example from an event with a muon in CMS:

- Several layers of different subdetectors crossed
- Passive material (cables, magnet, mechanical structure, …) crossed
- Non uniform magnetic field

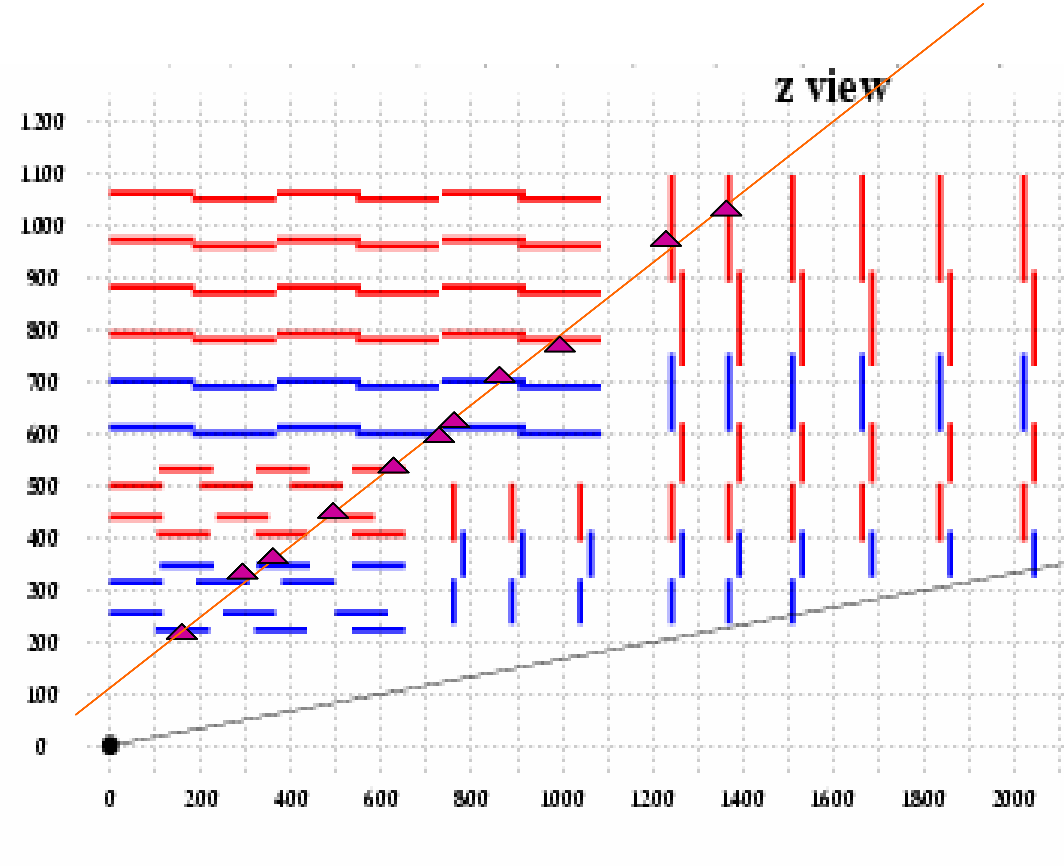
Simulation of the tracker system (schematic)

- A charged particle crosses the active layers (strips and pixels in CMS)



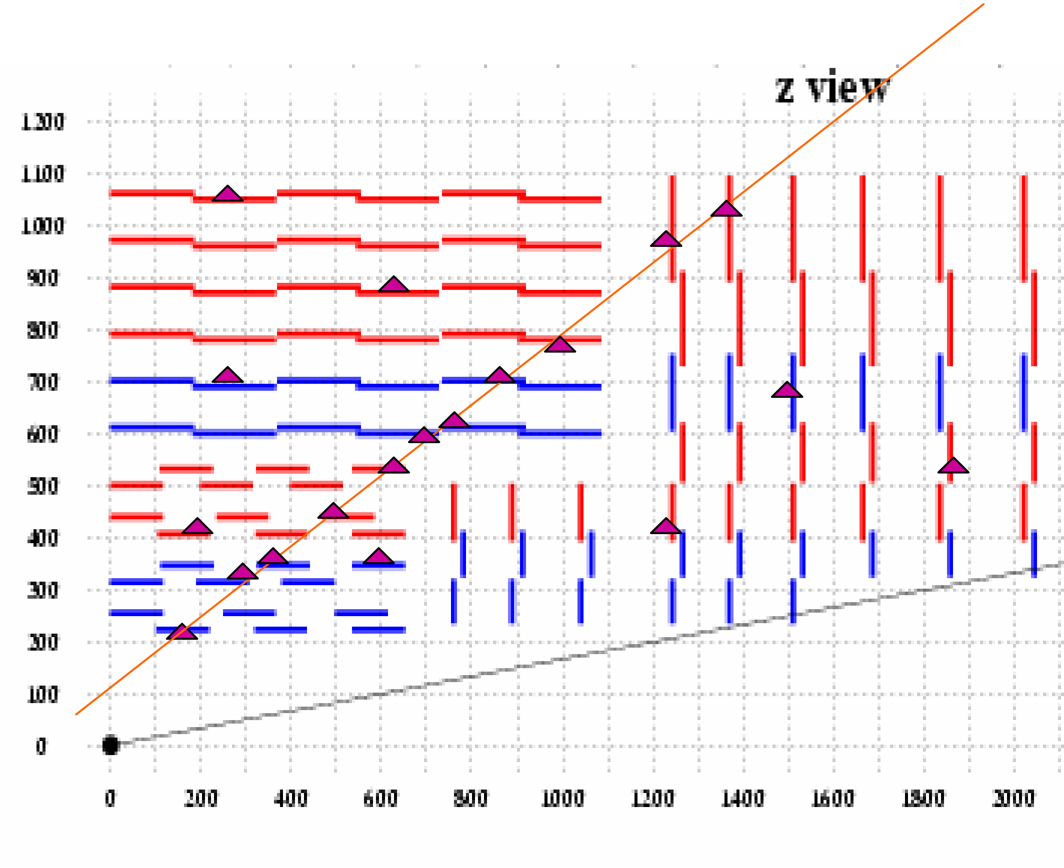
Simulation of the tracker system (schematic)

- A charged particle crosses the active layers (strips and pixels in CMS)
- (Within each layer) energy loss distributed along a path between entry and exit points
- Charges drift to the detector surface
- Gaussian noise is added



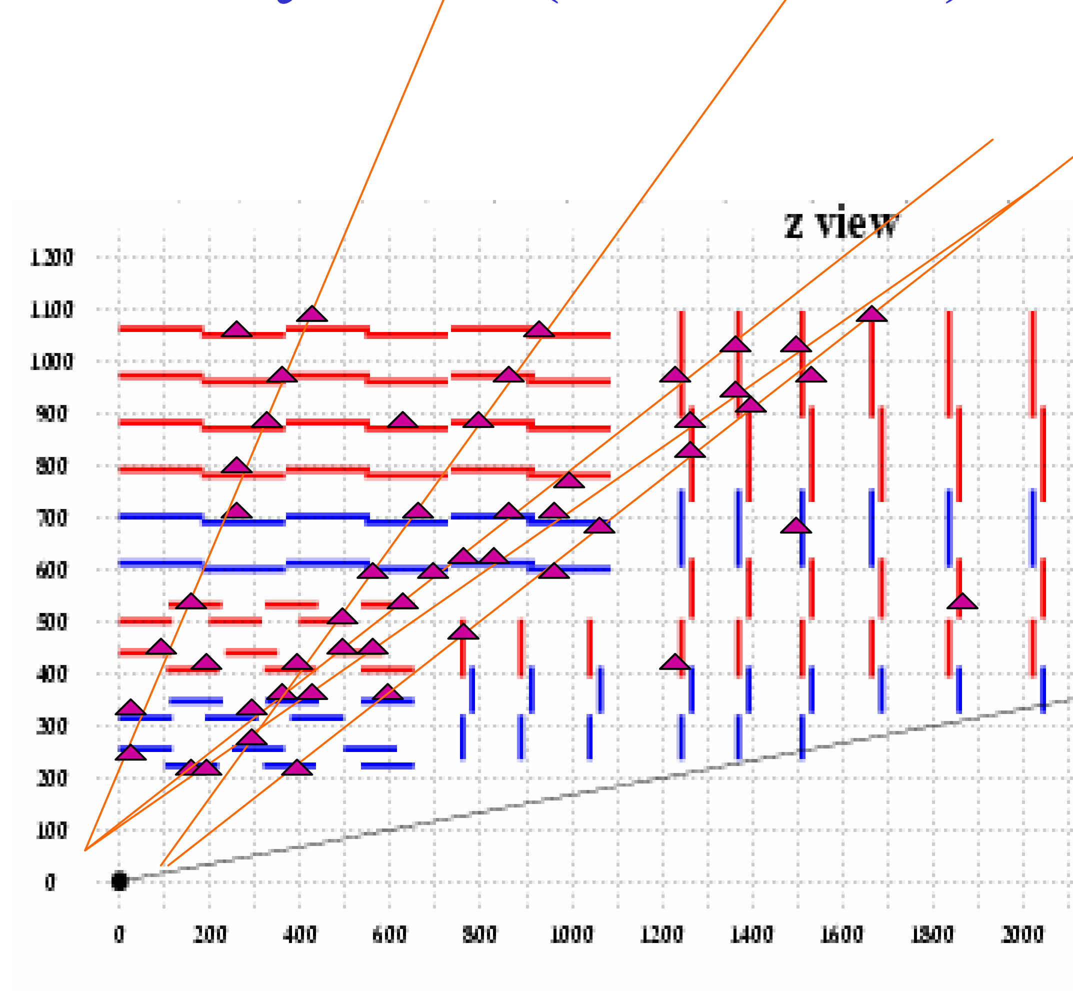
Simulation of the tracker system (schematic)

- A charged particle crosses the active layers (strips and pixels in CMS)
- (Within each layer) energy loss distributed along a path between entry and exit points
- Charges drift to the detector surface
- Gaussian noise is added
- Noise is also added to other channels



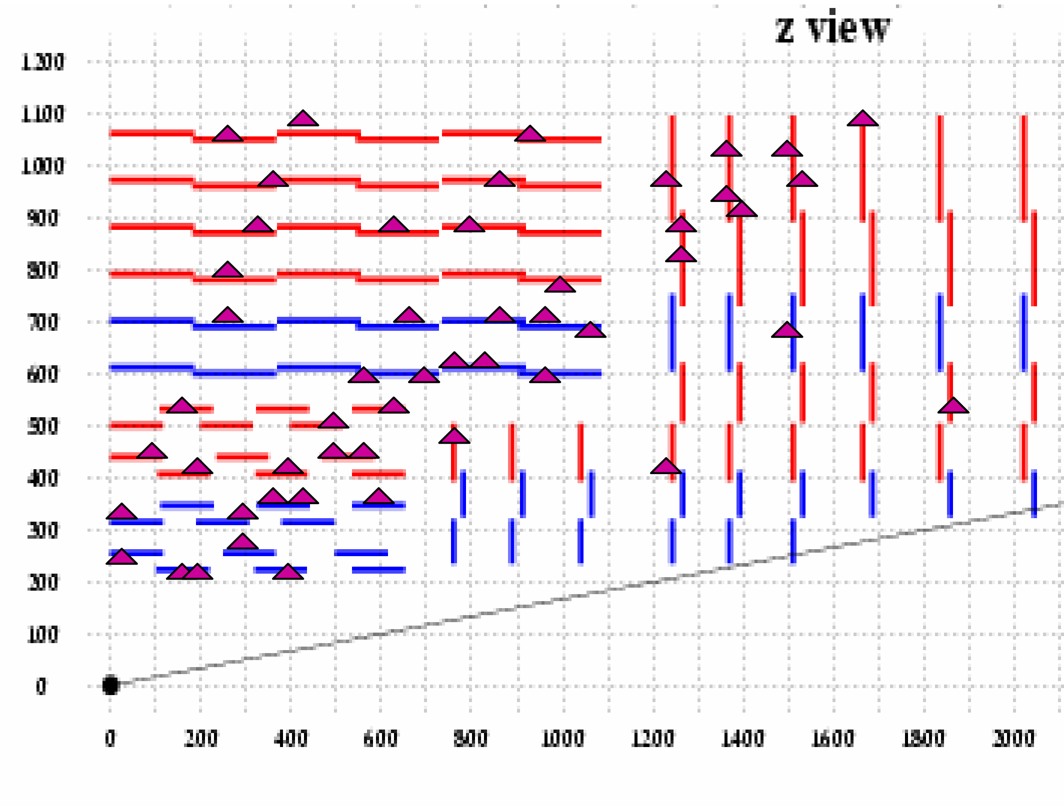
Simulation of the tracker system (schematic)

- A charged particle crosses the active layers (strips and pixels in CMS)
- (Within each layer) energy loss distributed along a path between entry and exit points
- Charges drift to the detector surface
- Gaussian noise is added
- Noise is also added to other channels
- Other particles from: same event, multiple interactions, pile-up
- Digitization



Reconstruction in the tracker system (schematic)

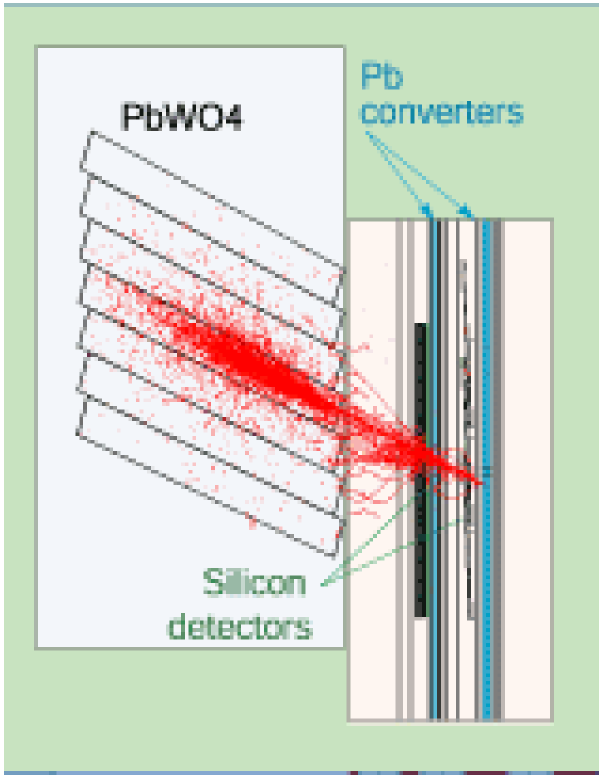
- Tracking algorithms with **pattern recognition** and **fit** of the tracks
- Take into account: curvature in the magnetic field, multiple scattering, material effects, etc.
- **Specialized approaches** for different use cases: low/high p_T tracks, searches for displaced vertices, etc.
- (As for the real data) exact 1-1 correspondence between generated charged particles and reconstructed tracks is lost: it can only be restored on probabilistic bases



Simulation of the calorimeters (schematic)

Sensitive volumes in **ECAL** and **HCAL**: the scintillating crystals and the silicon strips of the preshower detectors

Take into account:
• variation of light collection efficiency along the length of the crystal;
• modified crystal transparency with large integrated doses;
• noise;
• electronic thresholds;
• …



Calibration: test-beam and in situ

The whole charge collected in every single PM is read out together. In the **reconstruction**, exact 1-1 correspondence between generated particles and reconstructed showers is lost, and **cannot be restored**

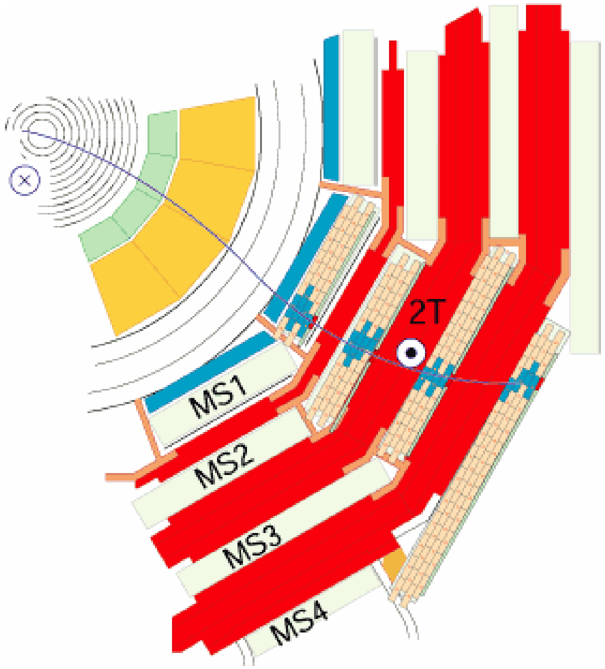
Simulation of the muon detectors (schematic)

Muon detectors are tracking devices

Muons produce ionization charge in the drift cells

Charges drift towards sense wires. Dependence on impact position, muon direction, residual magnetic field

Contributions from electronic noise, neutron background, muons from pileup event (in time or from a different BX)



Local reconstruction within a single superlayer

Global reconstruction, by correlating local tracks in the different substructures

Final muon reconstruction by matching with tracks in the inner tracking devices (plus calo signals, as mips).

Exact 1-1 correspondence between muons from the generator and reconstructed muons is lost: it can only be restored on probabilistic bases (smaller combinatorial than in the inner tracker, however)

Simulation of the trigger

LHC collision rate: **$4 \cdot 10^7$ Hz**

Acceptable DAQ read-out and data storage rate: ***O*** **(100 Hz)**

Rate reduction performed in (at least) two consecutive steps:

- Dedicated boards for the on-line (hardware) first level trigger, **L1**
- (Almost) off-line analysis run on dedicated farms of pc's with fast reconstruction algorithms and decision functions for the High Level Trigger (**HLT**)

HLT decides which events store on a permanent support: others will get lost forever

From the point of view of the off-line analysis of simulated events, trigger decision (even HLT) cannot be evaluated on the same analysis objects used for the analysis: in order to obtain realistic performances, specialized **trigger objects** have to be considered instead

Summary of full simulation (I)

- High level of details and precision achievable with a “well done” full simulation

- Detector responses validated and tuned with:
  - Test beam data
  - In situ calibration data (cosmics, halo muons, ..)
  - Calibration data from LHC collisions ($Z \to \mu^+\mu^-, e^+e^-$; $\pi^0 \to \gamma\gamma$; …)

- Material budget estimated on data taken in the final configuration

Summary of full simulation (II)

The **drawbacks**:

- Long CPU time required for the whole chain. For a “typical” LHC high p_T pp collision in CMS processed with a 1 GHz Pentium III:
  - **Generation of MC events** → less than 100 ms / evt
  - **Simulation of the material effects** (GEANT4) → 100-200 s / evt
  - **Simulation of the read-out electronics** (digitization) → 1-10 s / evt, , depending on luminosity (pileup rate)
  - **Reconstruction of physics/analysis objects** → 10-100 s / evt

  Total **CPU time** before analysis of a collision can be started is 3-5 m

- Laborious to handle, even for “experts”:
  - Queues on the computer farms with different priorities inside the collaboration;
  - Version control, control on run-time parameters, …

FAST SIMULATION

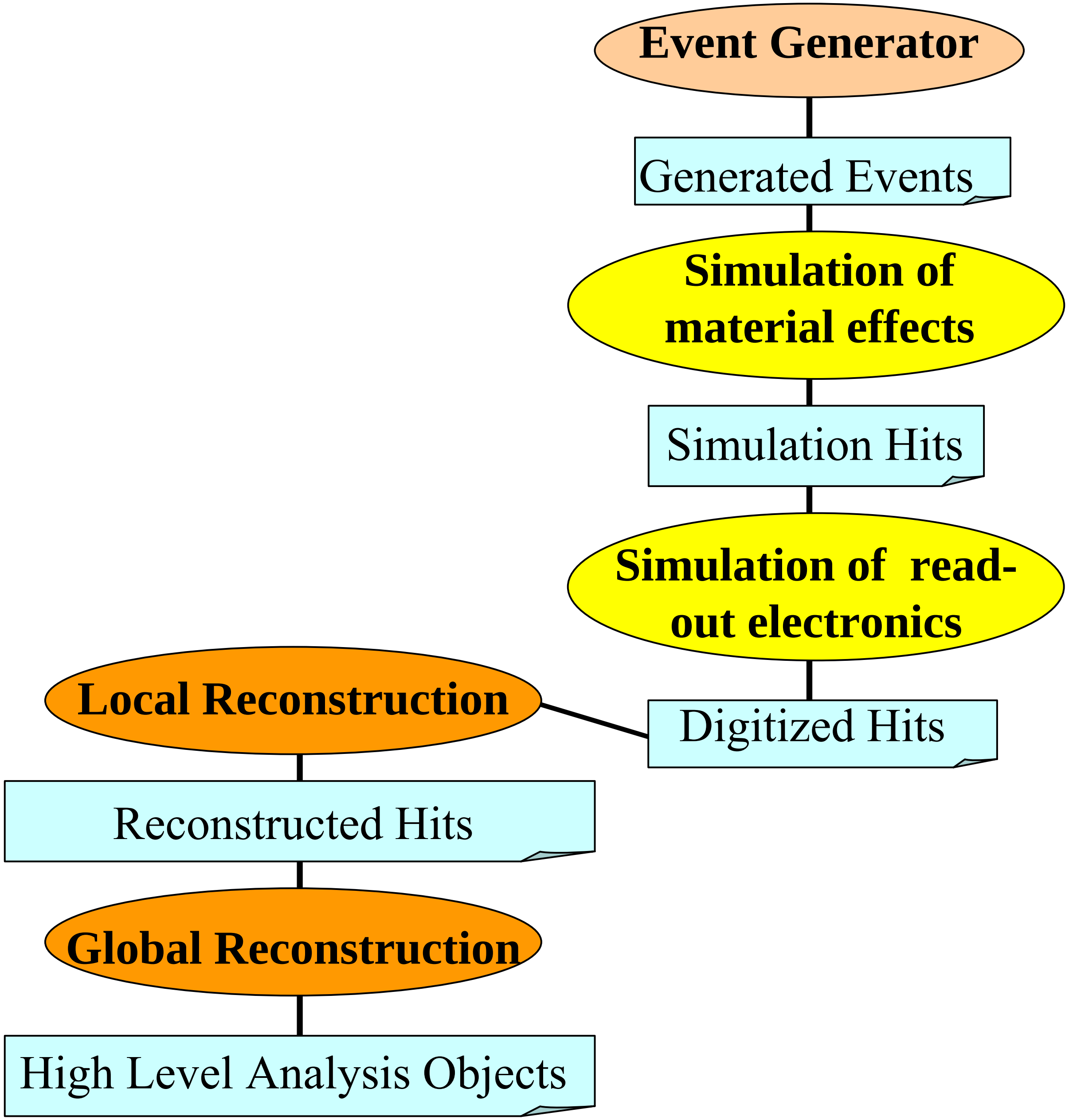
Fast simulation is <u>not</u> a replacement of the full detector simulation

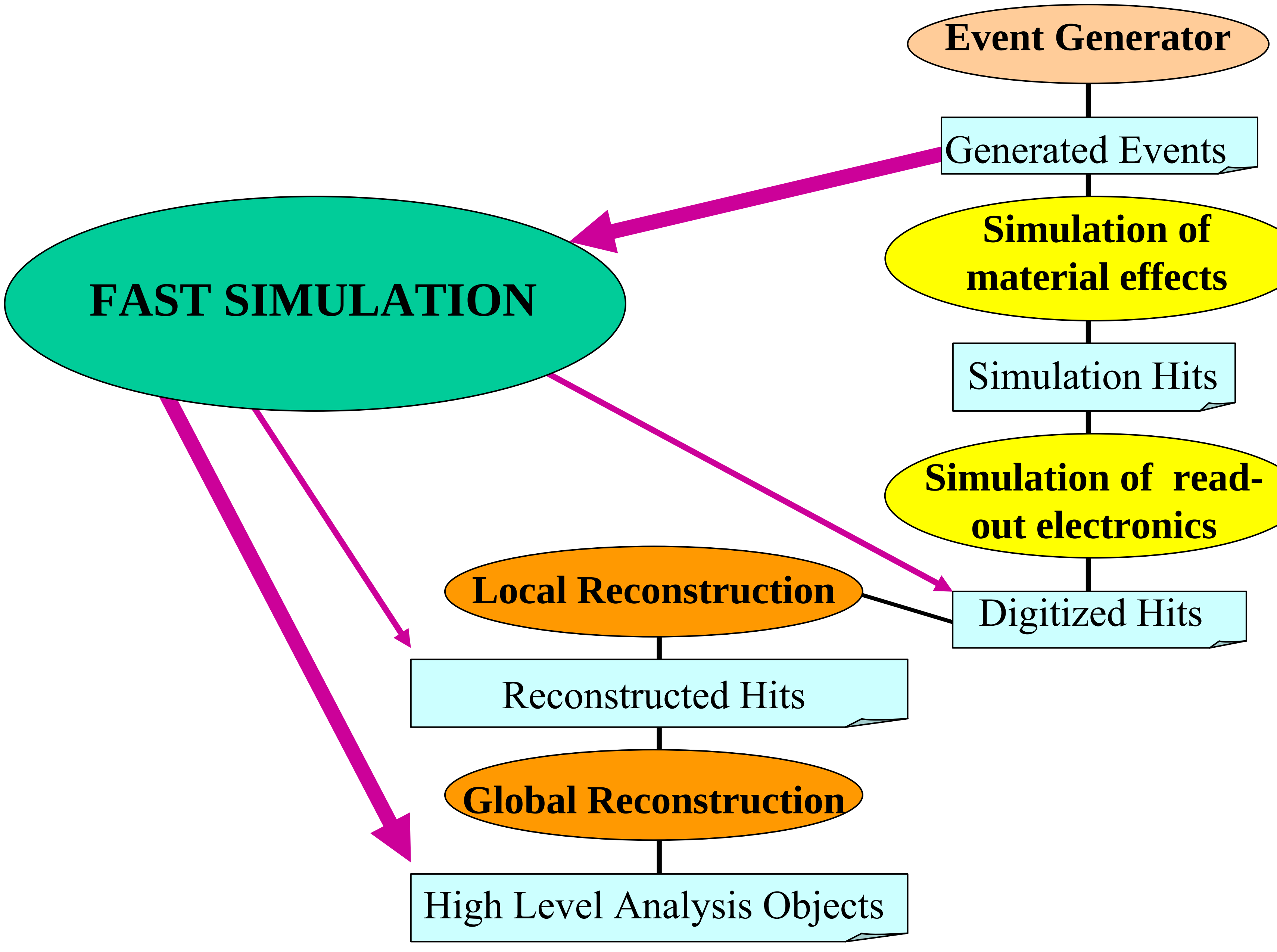
It is **tuned/validated with full simulation results** (while full simulation is tuned/validated with data, as test-beams or LHC collisions producing well known reference signals)

It usually emulates the combined result of detector simulation and reconstruction

It can be needed for:

- **Quick and approximate estimates of signal and background rates**
- **Scan of complex, multi-dimensional parameter spaces (e.g. SUSY)**
- **Fast development of analysis methods and algorithms**
- **Test of new MC or of new theoretical ideas in a realistic environment**
- **[Possible cross-check on some aspect of the full simulation]**
- **[Study of complex background processes with high cross-section at LHC]**

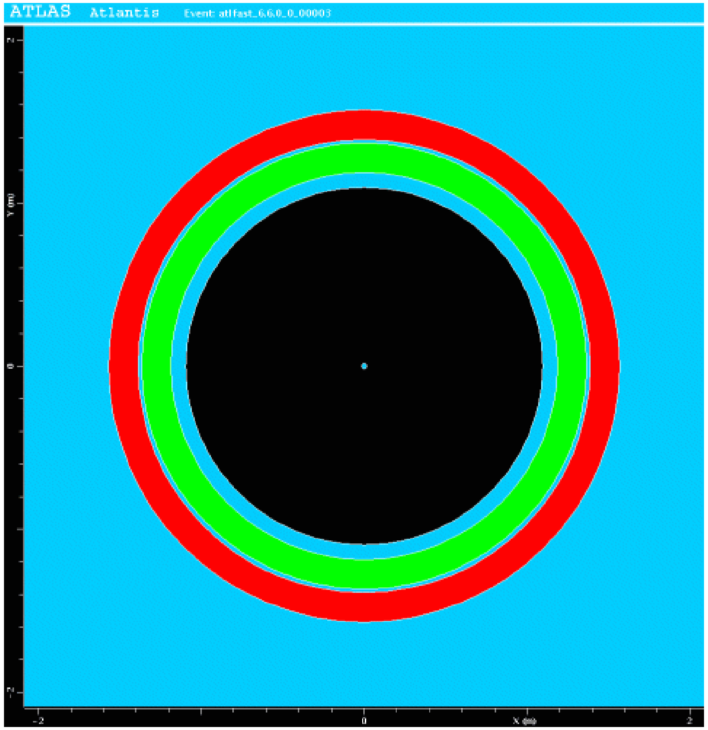

Event Generator
Generated Events
Simulation of material effects
Simulation Hits
Simulation of read-out electronics
Digitized Hits
Local Reconstruction
Reconstructed Hits
Global Reconstruction
High Level Analysis Objects


Event Generator
Generated Events
Simulation of material effects
Simulation Hits
Simulation of read-out electronics
Digitized Hits
FAST SIMULATION
Local Reconstruction
Reconstructed Hits
Global Reconstruction
High Level Analysis Objects

FAST SIMULATION IN ATLAS

ATLFAST is the package for fast simulation of ATLAS detector and physics analysis. It includes most crucial detector aspects:

- Jet reconstruction in the calorimeter, momentum/energy smearing for leptons and photons, magnetic fields effects and missing transverse energy.
- Provides, starting from the list of particles in the event, the list of reconstructed jets, isolated leptons, photons and muons and expected missing transverse energy.
- Provides also (optionally) the list of reconstructed charged tracks.

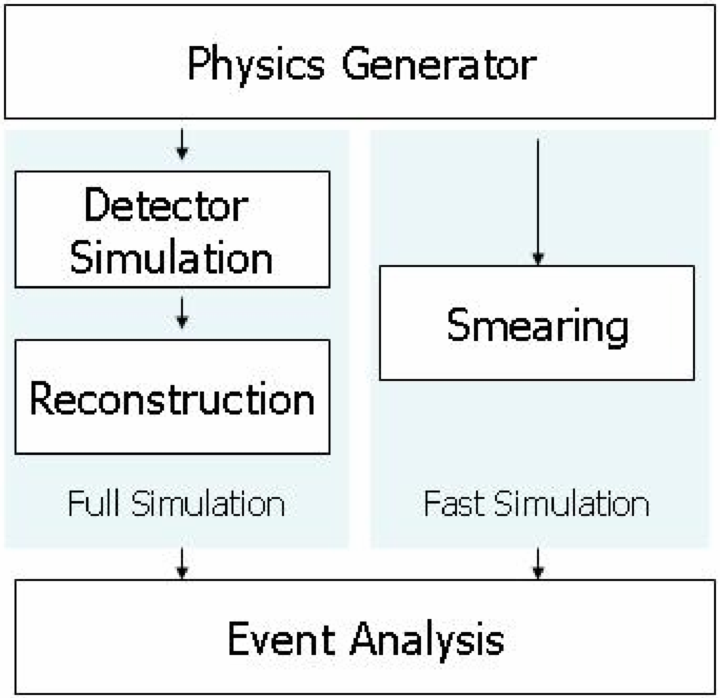


ATLFAST attempts to reproduce as well as possible the expected ATLAS detector mass resolution for important physics signals

No particle propagation, nor interaction with the detector material is simulated

Only the basic information on the detector geometry is used

Fast simulation is performed by **smearing the MC truth information directly with resolutions measured in full simulation studies**



Tracks simulation

Optionally, emulation of<track reconstruction is provided for charged particles inside the Inner Detector

It is obtained by **smearing three-momenta and impact parameters** as indicated in the full simulation studies

Different parameterizations of the smearing and reconstruction efficiency for muons, pions and electrons

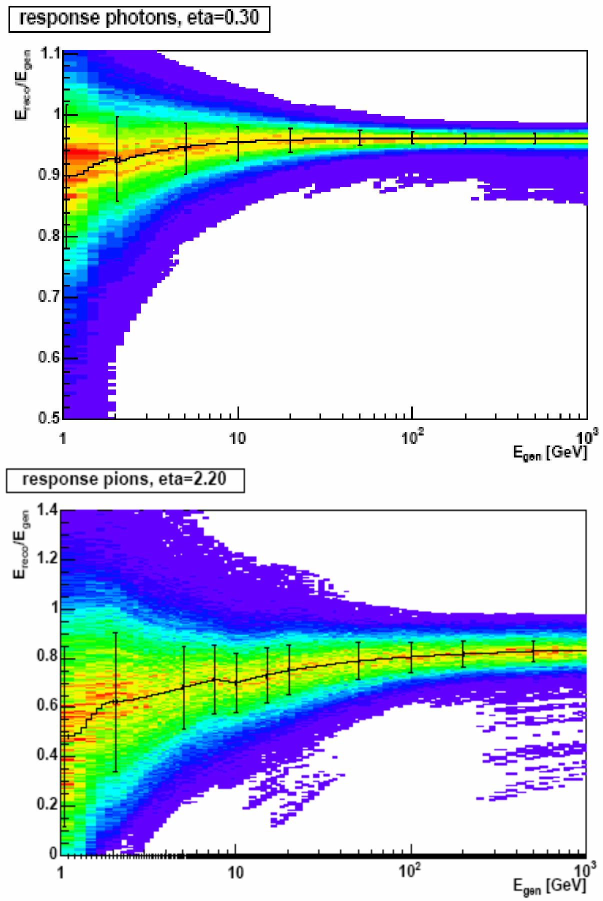
Calorimetric clusters (I)

Present implementation: all electron or photon energy deposited in one single ECAL cell; all hadron energy in one single HCAL cell.

New parameterization studied and ready to be implemented:

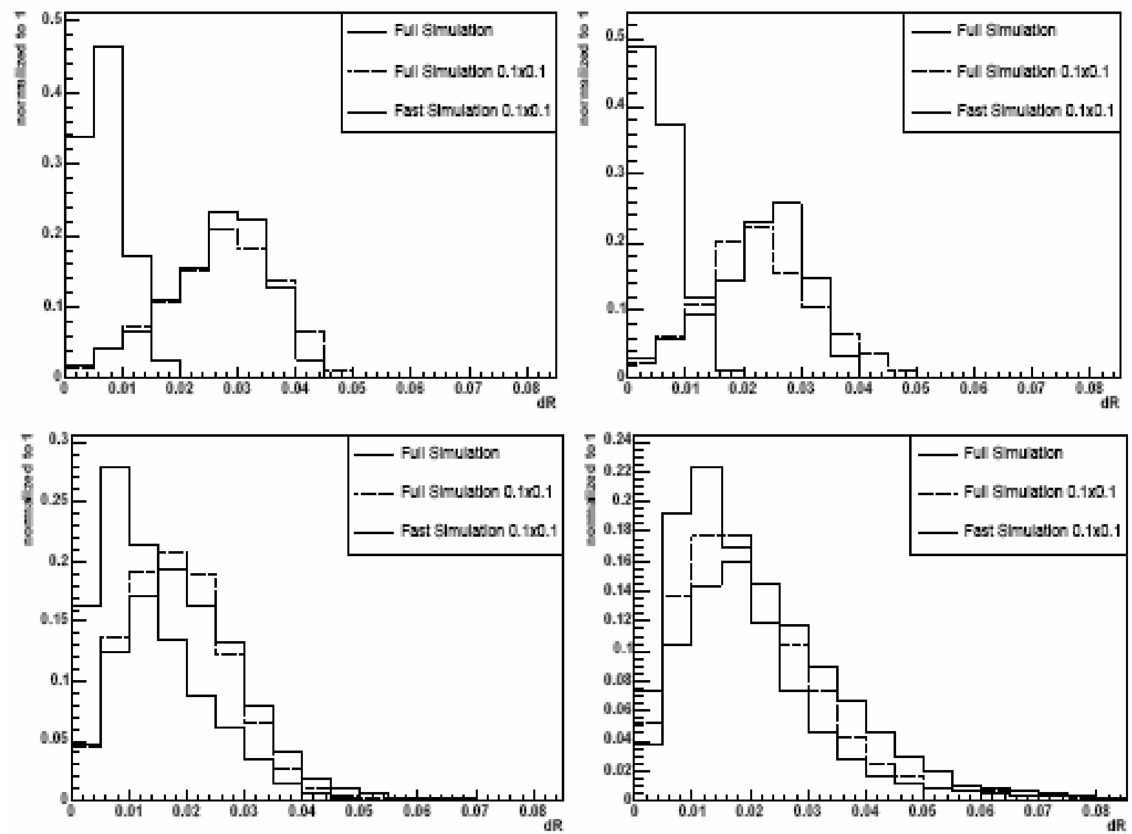
- E_T of all undecayed particles summed up in **cells** of given $(\Delta\eta x\Delta\varphi)=(0.1x0.1)$ granularity (Level 1 trigger tower granularity, coarser than full simulation); **longitudinal segmentation** limited to the separation between ECAL and HCAL
- Effect of 2T **B-field** on the position of charged particles taken into account
- Calorimetric **clusters** searched for starting from those cells
- Appropriate **energy smearing** and **reconstruction efficiency** applied **after cluster identification (from MC truth)** as electron, photon or hadron (pion)

Calorimetric clusters (II)



Response to photons and pions in ATLFAST (color) and full GEANT simulation (mean values and widths)

Energy flow around photon (top) and pion (bottom) directions, for 50 GeV particles at η=0.2 (left) and η=2.0 (right)



Jets

- Calo-clusters not associated with isolated *e* or γ smeared with a parameterized gaussian resolution
- b, c and jets are tagged (according to MC truth) and are smeared with **different parameterizations**
- Effect of **B-field** and pileup (at **high luminosity**) included in the parameterization
- Reconstruction and tagging efficiencies not in ATLFAST: they can be applied at a later stage:

WH production

$H \rightarrow bb$

$M_H = 100\ GeV/c^2$

WH , $H \rightarrow b\bar{b}$	Without B-field Without smearing	With B-field Without smearing	With B-field With smearing
$\langle p_T^{b\text{-}jet} / p_T^{b\text{-}quark} \rangle$	0.83	0.80	0.80
$\varepsilon_{b\text{-}jet}^{recon}$	83%	80%	81%
$\langle m_{b\bar{b}} \rangle$	87.5 GeV	82.6 GeV	82.0 GeV
$\varepsilon_{bin} = 82 \pm 20$ GeV	86.5%	84%	83%

Muons

Three options for the parameterization of the momentum resolution, depending on which detectors are used for muon reconstruction:

- muon system stand-alone
- inner detector stand-alone
- combined

Muons can be flagged as isolated or non-isolated

Muon-tagging efficiency not in ATLFAST (can be applied at a later stage)

Trigger

Primitive trigger routine

Not meant to cover all ATLAS triggers and levels

Aim at eliminating events which have essentially no chance of passing ATLAS level-1 and level-2 trigger

Three classes of trigger particles, whose efficiencies are parameterized in the low- and high- luminosity case:

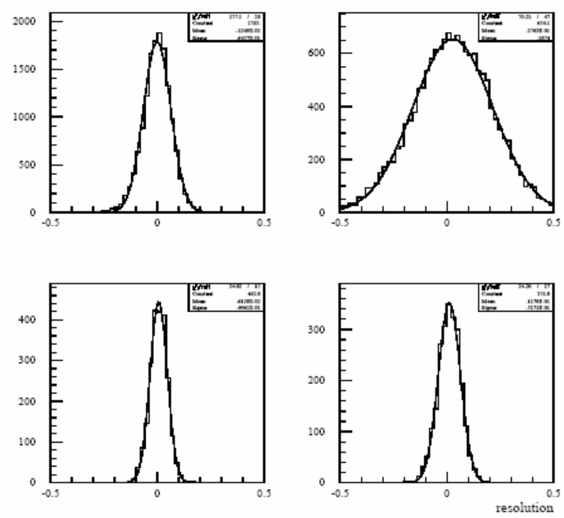
- isolated electrons and photons
- muons
- jets

Pileup

Pileup events not simulated in **ATLFAST**

Smearing of jets because of pileup parameterized as function of the luminosity

Low lumi High lumi



$40 < p_T(jet) < 50$ GeV/c

$200 < p_T(jet) < 250$ GeV/c

Parameterization of trigger selection allows for two low- and high- luminosity options

Timing

The approach chosen in **ATLFAS**T of relying on parameterizations of the properties of the final reconstructed analysis objects, without simulated interactions nor reconstruction, allows very fast processing

Four to five orders of magnitude gain with respect to the full simulation and reconstruction

FAST SIMULATION IN CMS

FAMOS (**FA**st **MO**nte-Carlo **S**imulation) is the package for fast simulation of particle interactions in the CMS detector.

Based on the present CMS framework (CARF): work is ongoing to migrate to the new framework, CMSSW. Basic features will remain unchanged, however.

The output of FAMOS is supposed to be as close as possible to the output of the full simulation (OSCAR) and full reconstruction (ORCA) of CMS. It **delivers the same physics objects** (calorimeter hits and clusters, tracker hits and reconstructed tracks, etc...), with **identical interface**. They can then be **used as inputs of the same higher-level analysis algorithms** (b-tagging algorithms, electron candidates, jet clustering, lepton isolation, etc...), or as starting point to new algorithm development.

Interactions simulated in FAMOS

Particles in FAMOS are propagated in the magnetic field through the inner trackers and until the entrance in the calorimeters

The following interaction are simulated in the tracker material:

- **Electron bremsstrahlung**
- **Photon conversion**
- **Charged particles energy loss by ionization**
- **Charged particles multiple scattering**

Only in the ECAL and HCAL:

- **Electron, photon and hadron showering**

No nuclear interactions simulated so far. It implies:

→ hadronic showers never initiated before the calorimeters

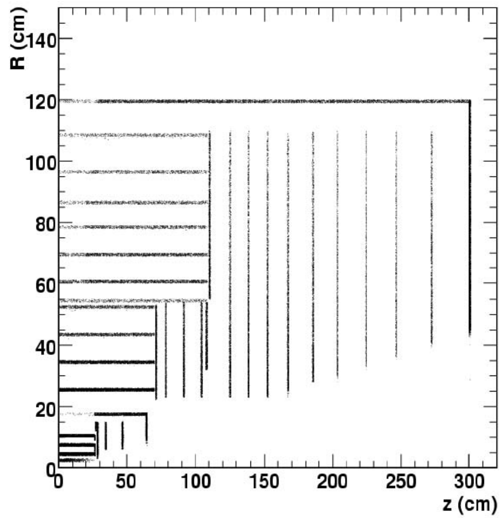
→ lower number of secondary vertices

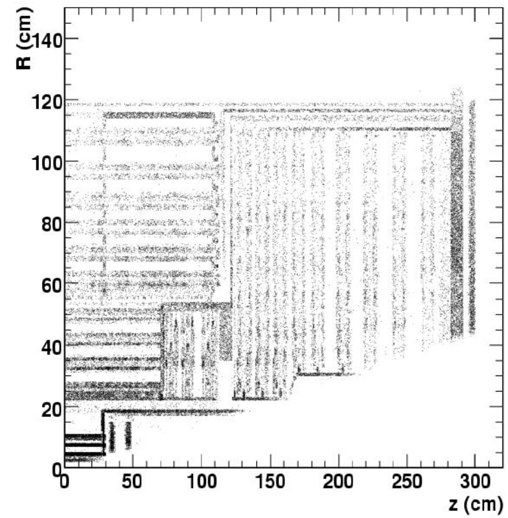
→ different b-tagging significance wrt full simulation: correspondence with full simulation has to be restored by retuning it “at hand”

Tracker (I)

Charged particles in FAMOS traced through a **simplified detector geometry**:

FAMOS

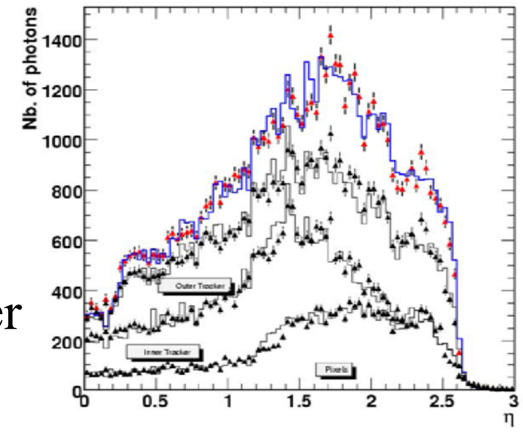




OSCAR

Pure silicon assumed as the sole tracker material

Thickness of all layers **tuned on the number of Bremsstrahlung photons** with E_γ>500 MeV radiated by energetic electrons traversing such layer



Tracker (II)

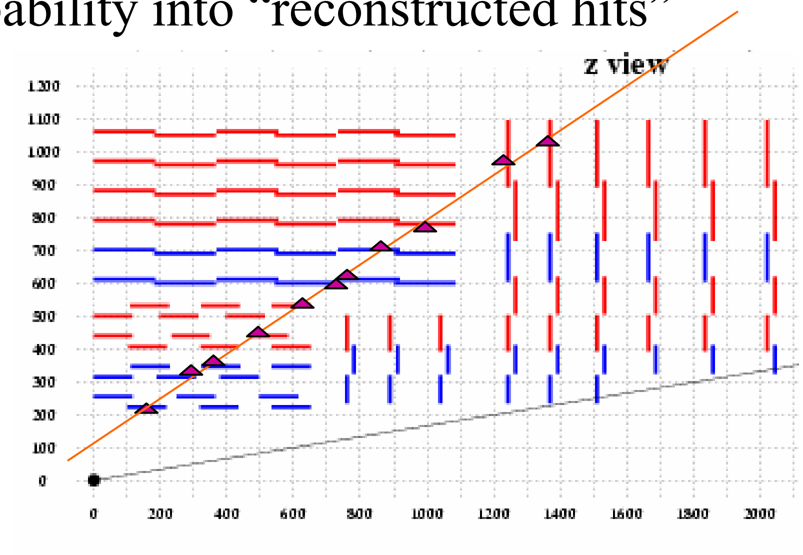
Charged particles in FAMOS are propagated in the magnetic field through the tracker layers

Multiple scattering and energy loss by ionization taken into account

Intersection between simulated trajectories and tracker layers give the “simulated hits”

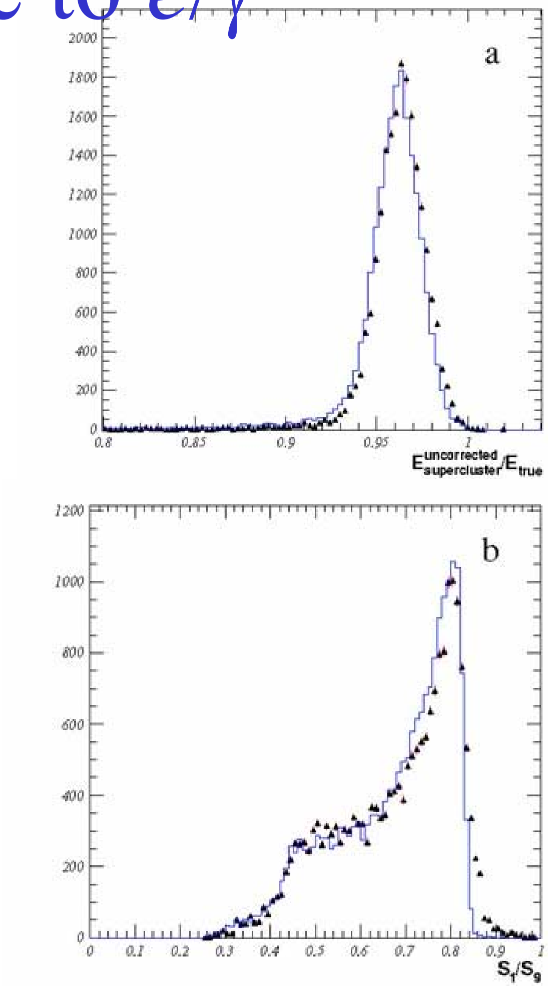
“Simulated hits” are turned with a given probability into “reconstructed hits”

No pattern recognition is performed, just a fit on the “reconstructed hits” belonging to the track, with the same fitting algorithm as the complete reconstruction



Calorimeter response to *e/γ*

- Shower developed (*following the Grindhammer parameterization*) as if the ECAL were an **homogeneous medium**.
- Energy deposits **sliced longitudinally**
- In each slice, **energy spots distributed in space according to the radial profile and placed in the actual crystal geometry**
- Simulate: **leakage**, **gaps between ECAL modules**, shower enlargement due to **B-field**
- **Electronic noise** and **zero suppression**
- Preshower considered in the endcaps
- **Leakage** propagated into HCAL
- **Clustering as in the complete reconstruction**

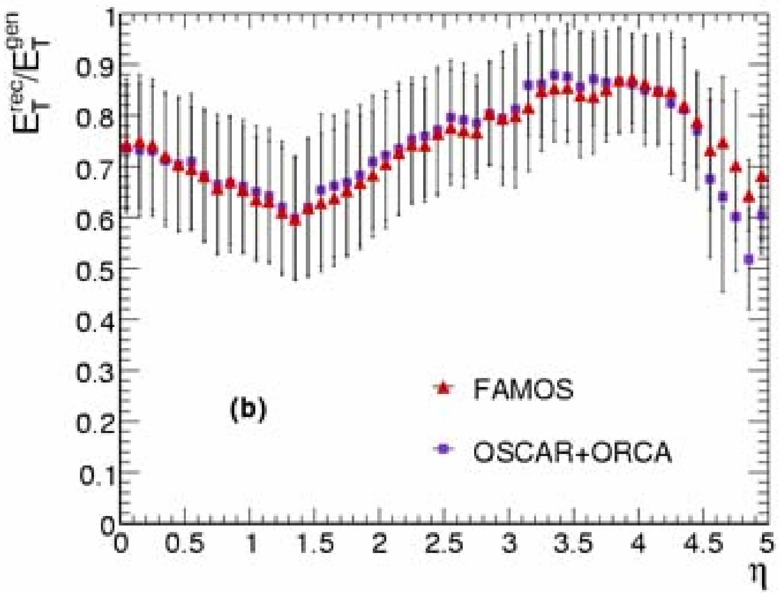


Electron reconstructed total energy and transverse energy profile in FAMOS (dots) and OSCAR+ORCA (histogram)

Calorimeter response to hadrons

- Charged and neutral hadrons **propagated to the ECAL and HCAL/HF** entrance
- **Energy response derived from full simulation of single pions** at fixed p_T values between 2 and 300 GeV/c (Gaussian mean value and width)
- Smeared energy distributed in calorimeters using **parameterized longitudinal and lateral shower profiles**, using an approach similar to GFLASH
- Other hadrons treated as **pions** with the same p_T

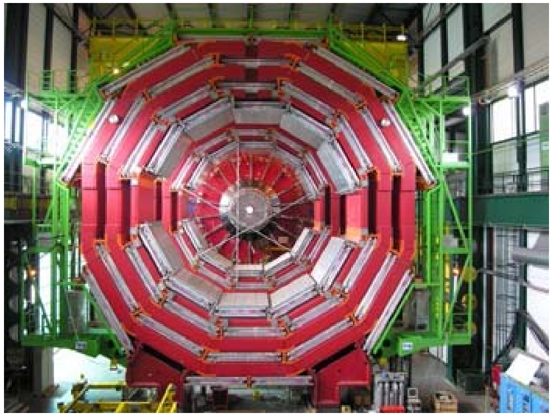
$80 < E_T < 120$ GeV

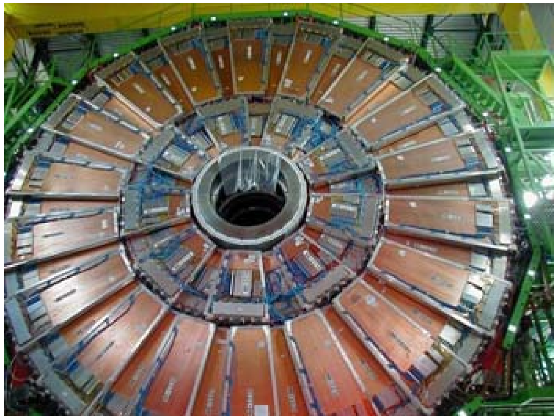


Ratio of reconstructed to generated jet E_T as function of pseudorapidity

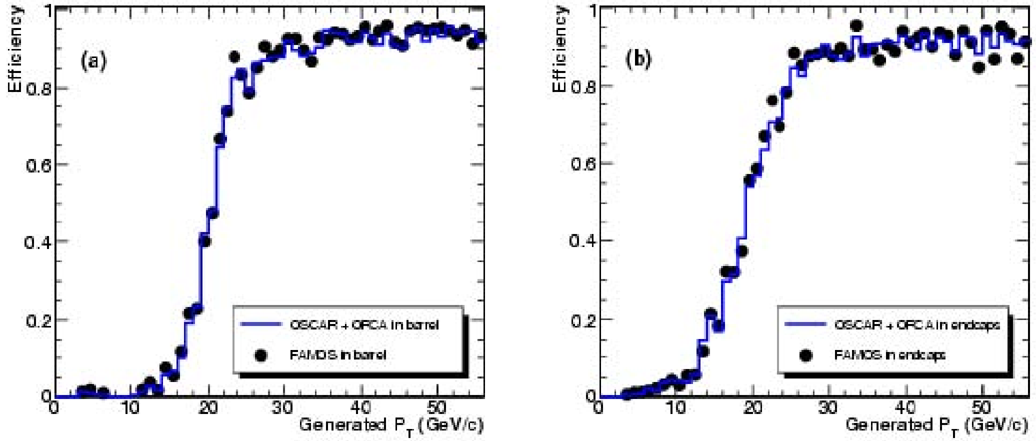
Muons (I)

- FAMOS muons are **not propagated** to the muon chambers (DT, CSC, RPC)
- **Calorimeters response** to muons tabulated in similar way as for hadrons
- Response of the muon chambers **parameterized on a sample of single muons** (with $2 < p_T < 1000$ GeV/c, uniform on the η and φ acceptance of the muon chambers) to reproduce efficiencies and resolutions (Gaussian distribution assumed) of the full simulation
- Different parameterizations for:
  - **L1 trigger**
  - **HLT muons**
  - **Global** (i.e. finally offline reconstructed) **muons**
- HLT and global muons look for a correlation with the reconstructed tracks

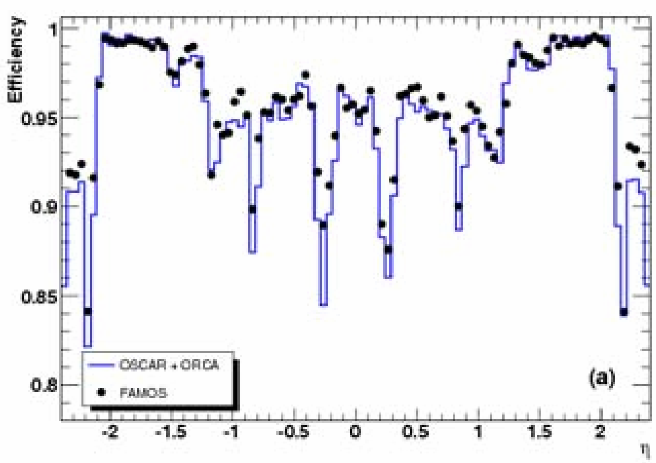




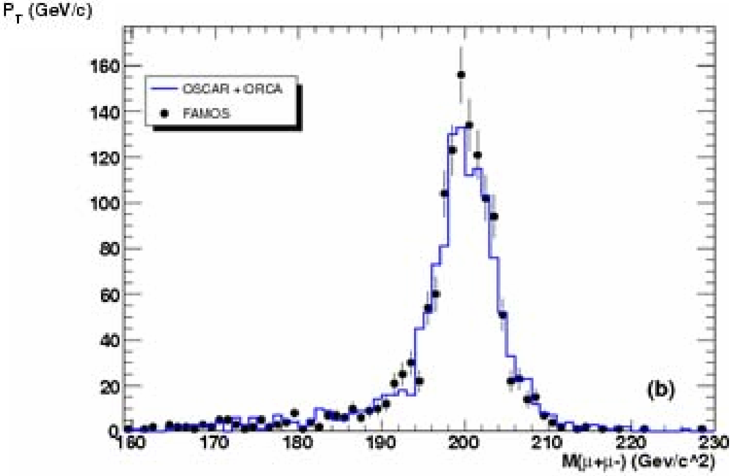
Muons (II)



"Turn on" curve for muons when the ***L1 single muon*** *p_T threshold is at 20 GeV*



Algorithmic efficiency of HLT muons as function of η



Di-muon mass in $h/A \to \mu\mu$
Selection efficiencies in fast and full simulations agree within ±1.5 %

Trigger

- L1 and HLT trigger signals and primitives produced "as a byproduct" of the fast simulation of the corresponding subdetectors
- Decision functions reconstructed with the very same logic as in the real data.

Pileup

- **In-time** pileup MB events **superimposed** to the signal event, their particles treated as all other particles in the event

- No out-of-time pileup considered

Timing

A complete, high p_T event takes a couple of seconds to be simulated and reconstructed with **FAMOS** (about 1s in FAMOS itself, the rest in the analysis and framework overhead) \, slightly more with pileup

More than **two orders of magnitude** gain with respect to the full simulation and reconstruction

COMPARISON OF THE FAST SIMULATIONS IN ATLAS AND IN CMS

Fast simulation in ATLAS focuses on:

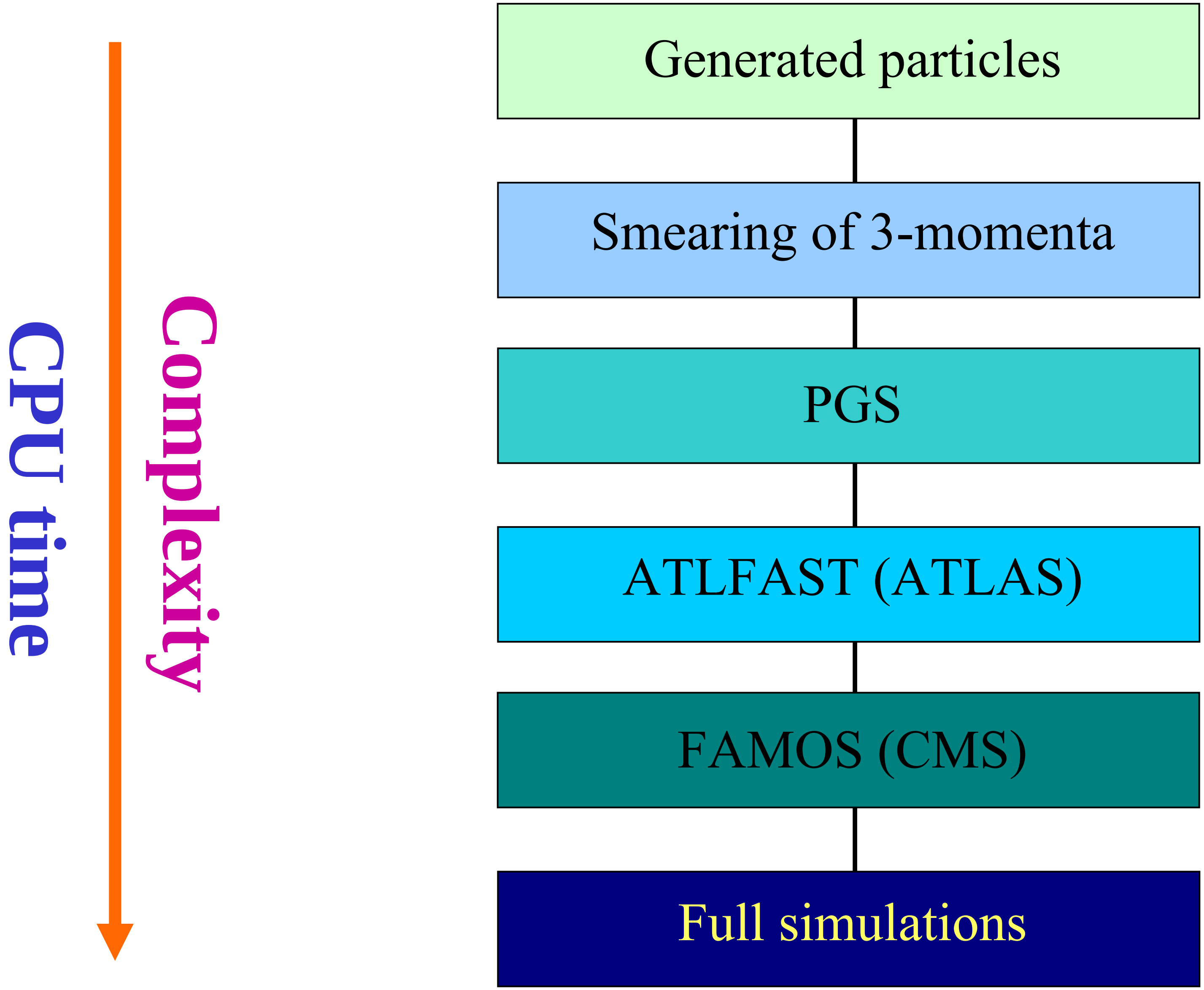
- **simplicity**
- **velocity**

without sacrifying too much the agreement with the results of the full simulation

Fast simulation in CMS focuses on:

- **intermodularity with full simulation and reconstruction**
- **best possible reproduction of full simulation results**

without sacrifying too much the velocity (CPU time)


CPU time
Complexity
Generated particles
Smearing of 3-momenta
PGS
ATLFAST (ATLAS)
FAMOS (CMS)
Full simulations

Conclusions

- Both ATLAS and CMS have developed programs for the Fast Simulation, to help analysis of LHC collision events
- Different approaches in the two cases:
  - **FAMOS** in CMS points to have results as close as possible to full simulation ones
  - **ATLFAST** in ATLAS points to have results as fast as possible
- Extensively used already for PTDR studies
- Validated, maintained and kept up to date
- Fast simulations of LHC detectors can become the entry point for phenomenologists wanting to test their ideas and MC in a realistic LHC environment. Not meant for public use, however: interaction with experimental collaborations is mandatory!

Thanks to Patrick Janot and Giacomo Polesello!