Fast-Simulation of CMS: status and usage

Federica Primavera
University of Bologna & INFN (IT)
on behalf of the CMS collaboration

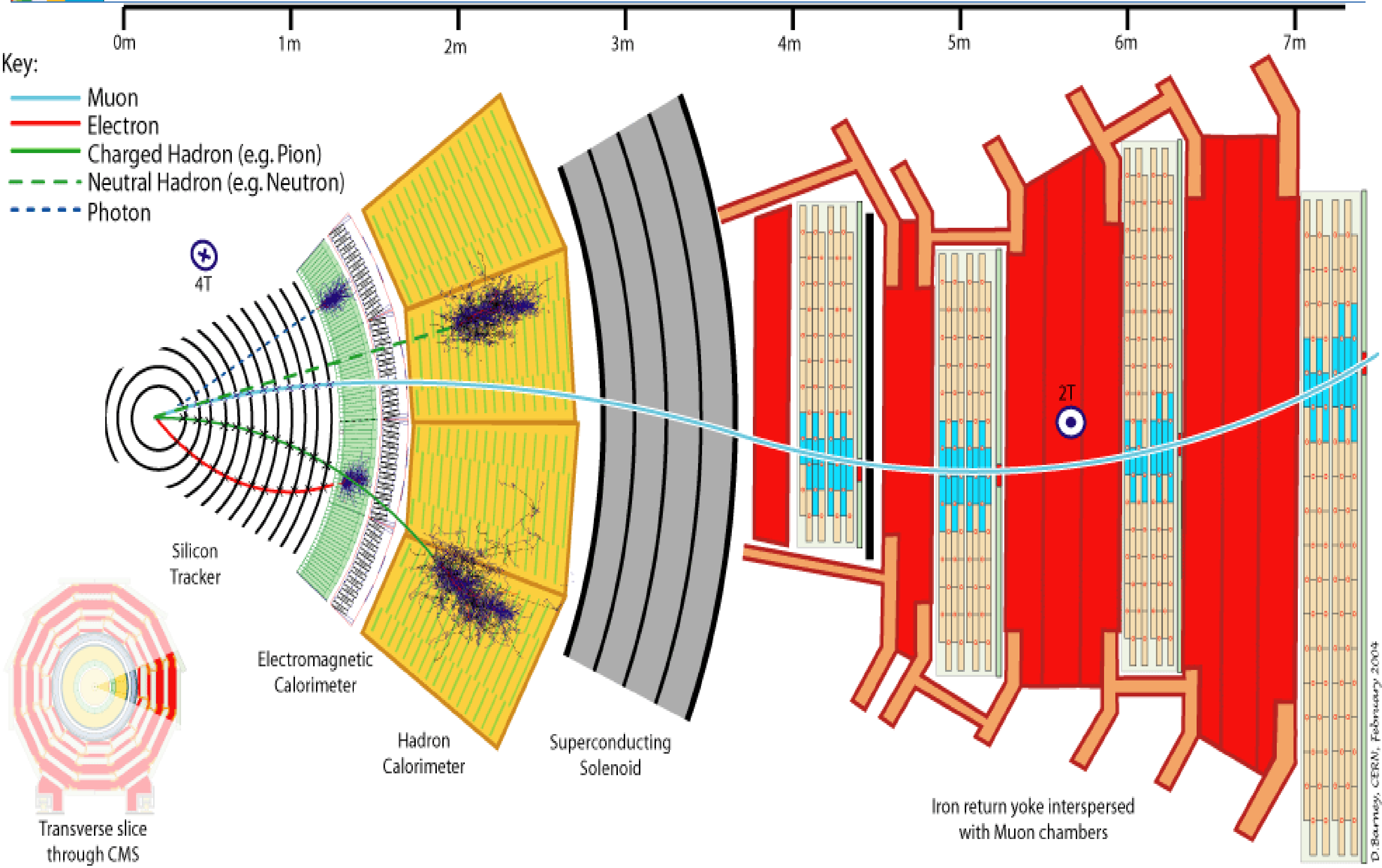
2nd Fast Monte Carlo Workshop in HEP

14-16 Jan 2014, DESY-Zeuthen

Outline

- Description of CMS
- What is the Fast Simulation:
  - detector response
  - digitization
  - reconstruction
  - plots of the performances
  - PileUp
- Usage and prospect
- Summary

CMS detector



Fast-Simulation in CMS

The Fast Simulation of CMS is an object-oriented subsystem of the general CMS C++ based software.

Goal:
event production rates are of the order of 100 times faster than the Geant4-based Full Simulation.

The Fast Simulation is alternative to the Geant4-based Full-Simulation

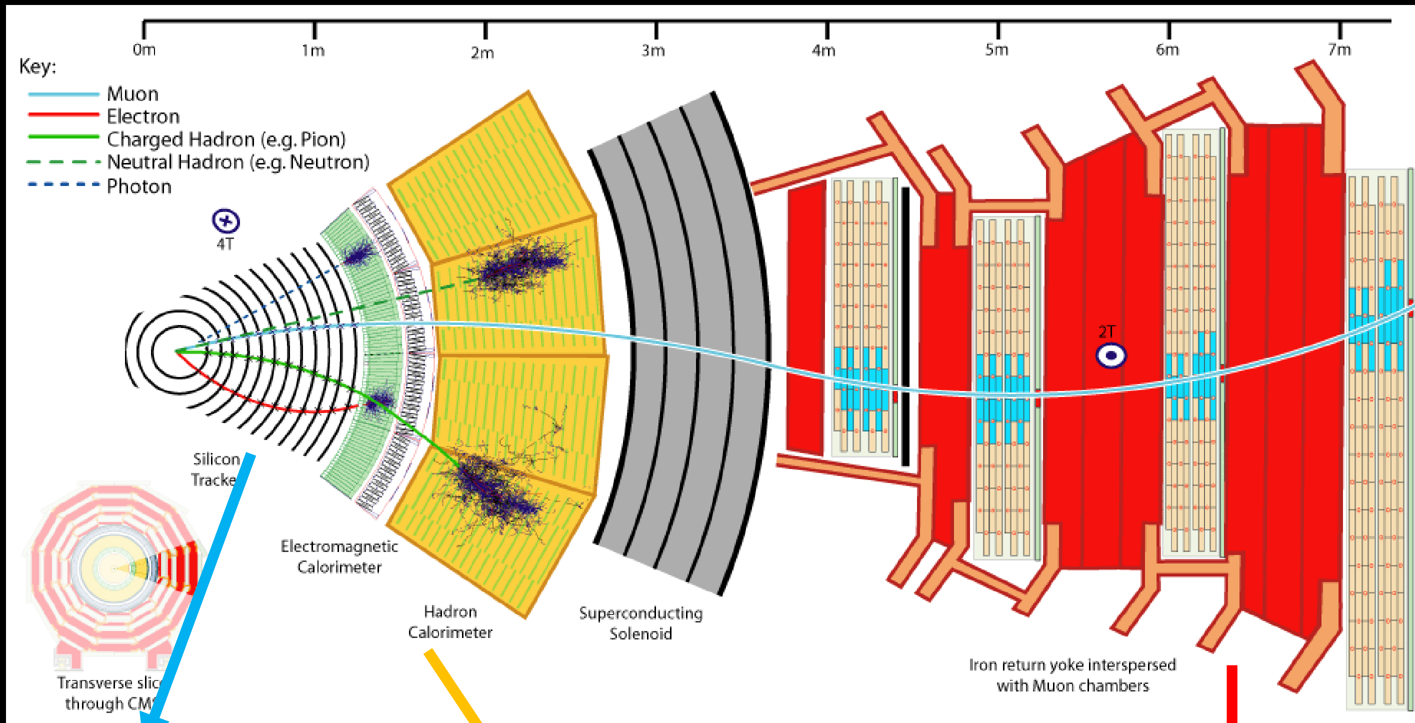
It is not an ab-initio simulation, but parametrises some material effects according to their known distributions.

The Full-Simulation is regularly used for validation and tuning

It is also different from PGS or Delphes, which are fully parametric simulations (it is not a smear gen-level to analysis-level).

Simulated interactions

All material effects are simulated when crossing a layer.



INNER TRACKER:
Electron bremsstrahlung
γ-conversion
Energy loss by ionization
Multiple scattering
Nuclear interactions.

CALORIMETERS:
electron, photon and
hadron showering.

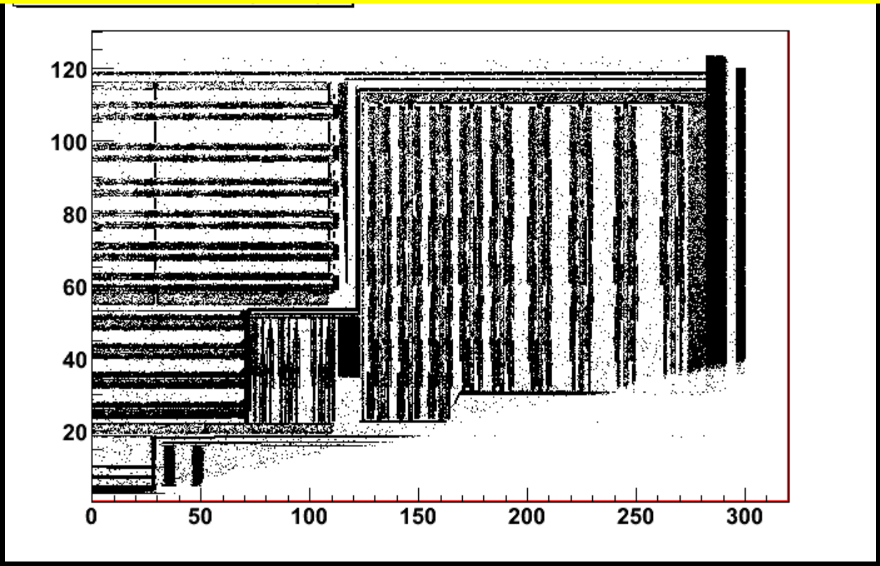
MUON CHAMBERS:
Energy loss by ionization
Multiple scattering

Detector simulation step: SimHits

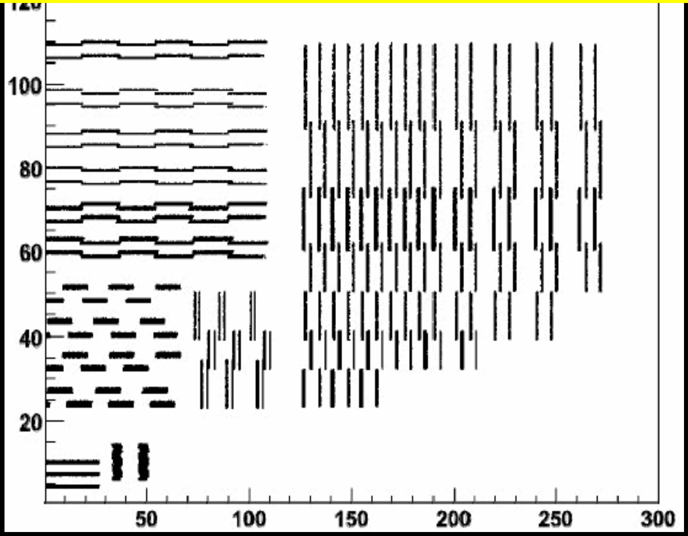
REALISTIC SIMULATION of low level object (hits and clusters) by:

- dedicated algorithms;
- simplified tracker geometry: - material is mapped onto layers;
  - exact description of sensitive elements by using reconstruction geometry
  - direct propagation between volume boundaries, but taking into account detailed magnetic field map;
  - dead modules are considered for the pixels (from conditions database).

CMS-tracker “tomography” with FullSim geometry



...and with reconstruction/FastSim geometry



Detector simulation step: SimHits

CALORIMETERS exploits a shower simulation à la GFLASH:

- **ECAL:** - treated as a homogeneous medium;
  - cracks, leakages, and magnetic field effects are taken as FullSim.

- **HCAL:** - hadronic response and resolution are tuned on the FullSim Single-Pion events;
  - extensively validation with test beam data from 2010;
  - new tuning in 2013 to account for the use of digitizers (at SimHits level instead of RecHits level to decouple from the electronic effects).

Muons are the only generated particles propagated in the **MUON CHAMBERS**:

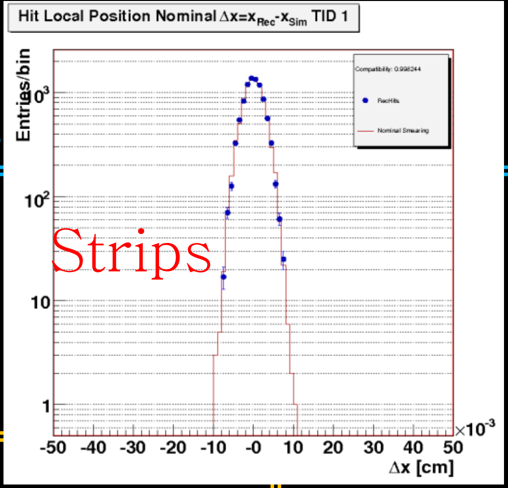
- multiple scattering and energy loss are taken into account ;
- muons from hadronic decays are propagated only if coming from decays inside the tracker volume. No late decays and no punch-through are propagated;
- calo-deposits are parametrized;
- no bremsstrahlung an δ-rays;
- standard geometry as FullSim is used.

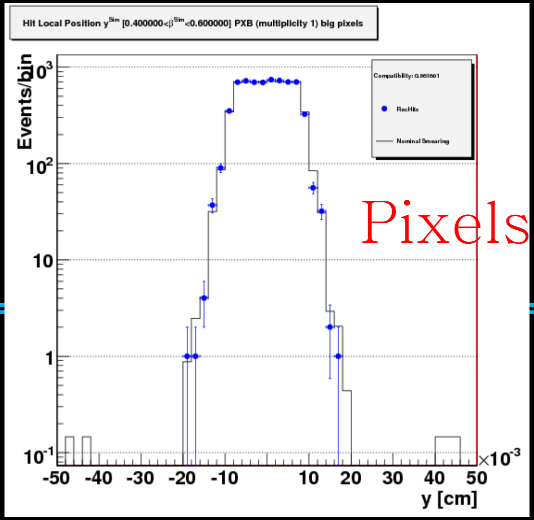
From SimHits to RecHits

TRACKER: SimHits are directly converted into RecHits after a position smearing.

Pixels: resolution distributions extracted from FullSim (might be from data) as functions of cluster multiplicity and track incidence angle.

Strips: layer-dependent Gaussians.





CALORIMETERS:

Until CMSSW 5.3 (Run-I legacy release)

SimHits were converted into RecHits after a Gaussian energy smearing tuned on FullSim.

Since CMSSW 6.2 (Run-II release),

fully interfaced to FullSim digitizers.

MUON CHAMBERS: SimHits are fed into the digitizer modules (FullSim), and everything downstream is from standard reconstruction, including pattern recognition.

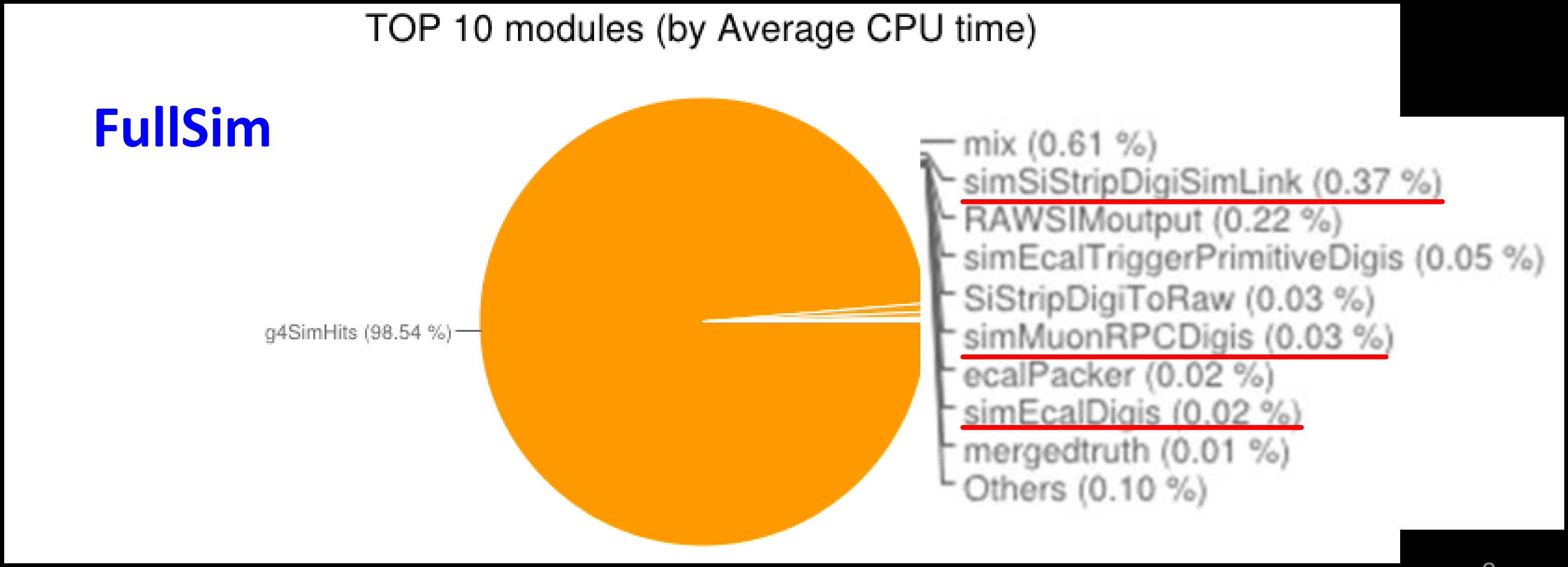
Digitizers

The digitizers are taken from the Full Simulation for the:

- muon system;

- electromagnetic calorimeter;
- hadronic calorimeter ;

- the tracker module is skipped for the moment, because is the most time consuming.



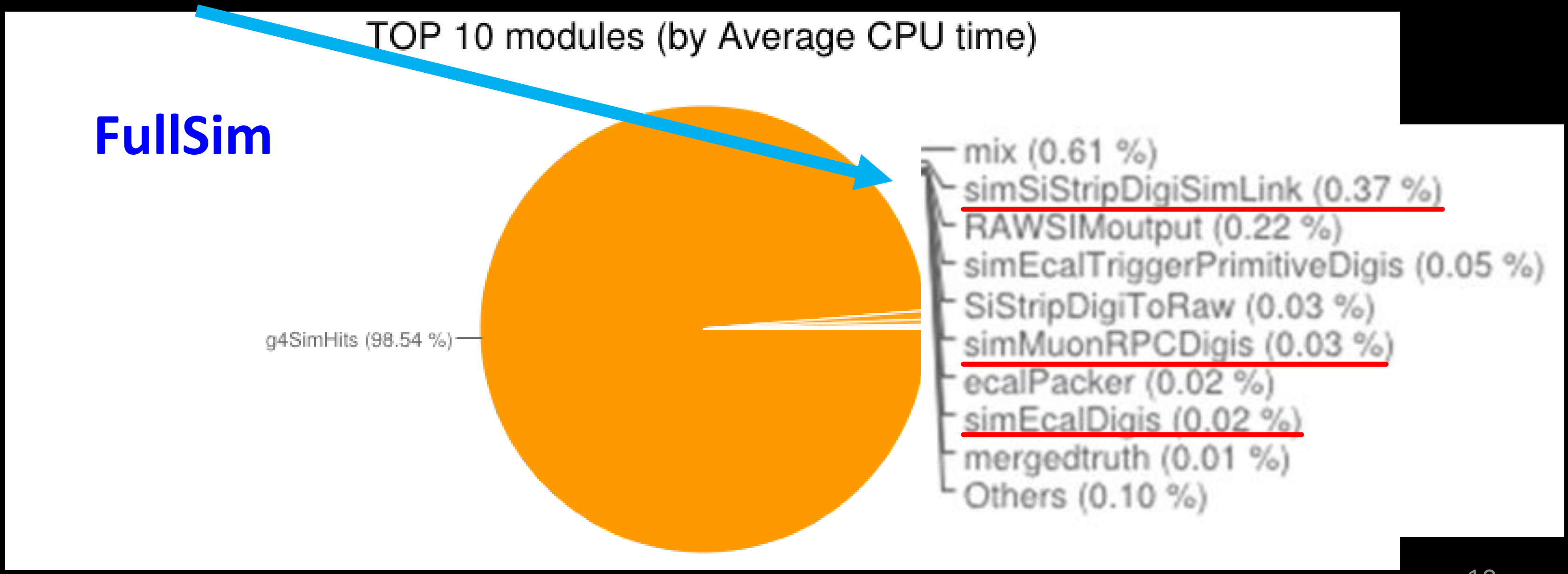
Digitizers

The digitizers are taken from the Full Simulation for the:

- muon system;

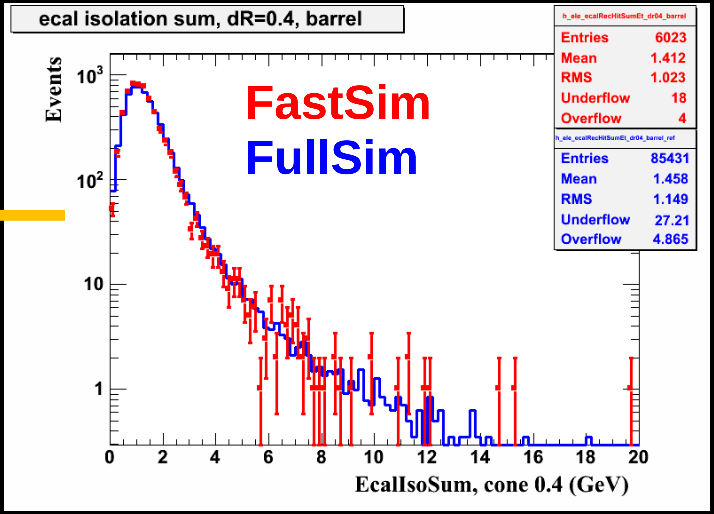
- electromagnetic calorimeter;
- hadronic calorimeter ;

- the tracker module is skipped for the moment, because is the most time consuming.

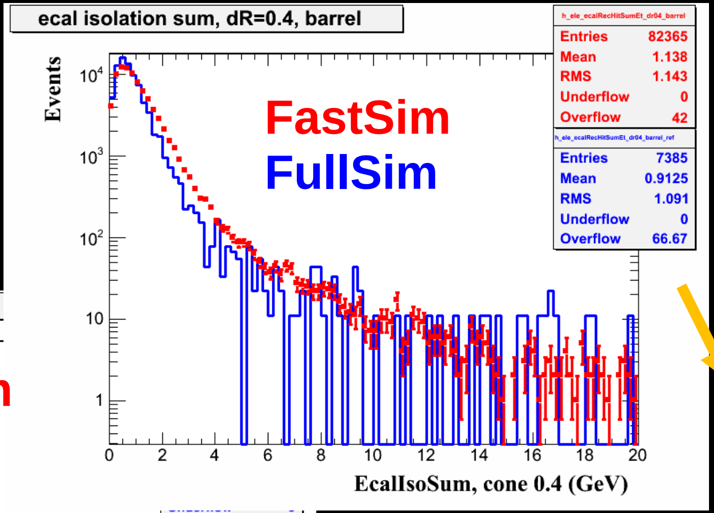


ECAL and HCAL digitizers

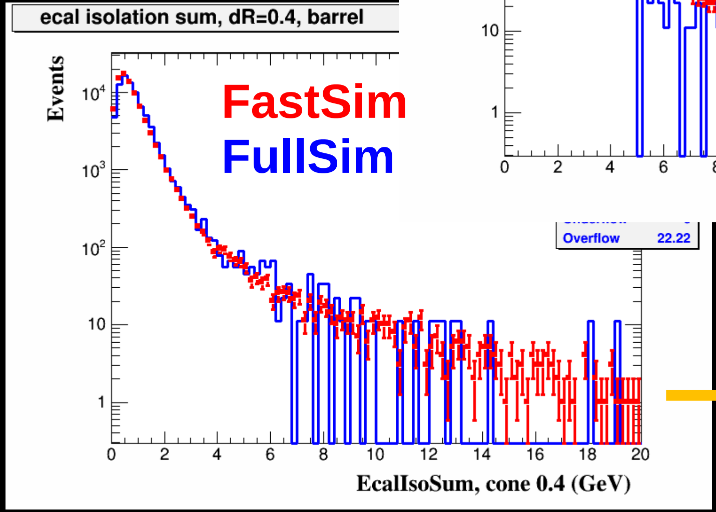
- Improvements in some physics observables.
- Less maintenance burden, as electronic effects (noise) don't need to be retuned to keep up with changes elsewhere.



CMSSW 4.4 (2011 release): Ecal noise applied in FastSim module tuned on FullSim; good agreement in noise-sensitive variables.



CMSSW 5.3 (2012 release): data and FullSim changes in selective redout and zero-suppression. FastSim specific noise model becomes inadequate.



CMSSW 6.2 (2013 release): after integration with FullSim digitizers, no need anymore for a FastSim-specific noise model.

Reco-objects

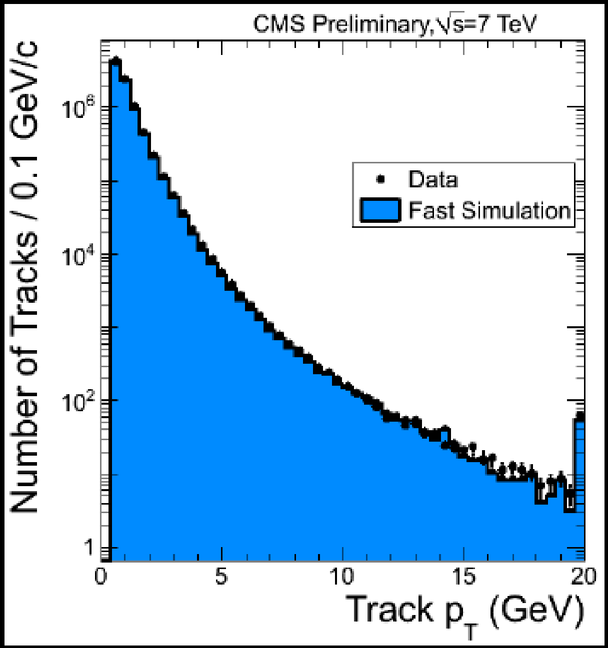
RECONSTRUCTION (more time consuming than simulation) exploits the same modules of FullSim:
- lepton reco;
- Particle Flow;
- jet finding;
- b-tagging;
- isolation;
- **exception is given for the tracking.**

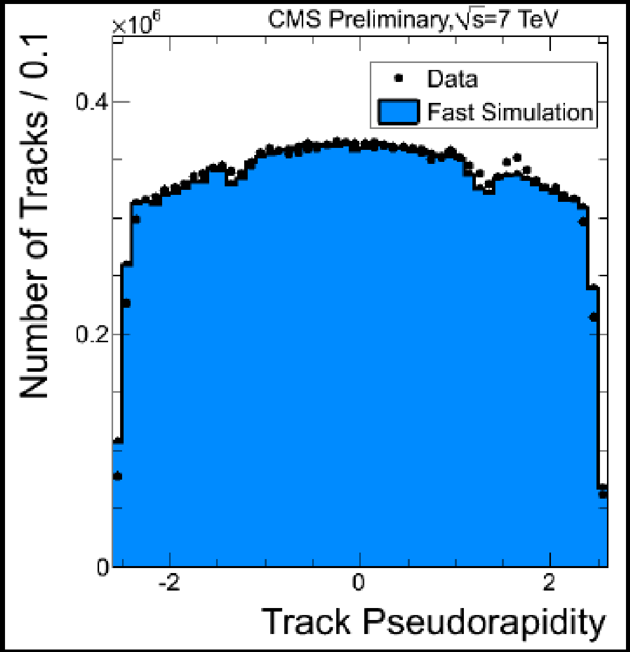
↓

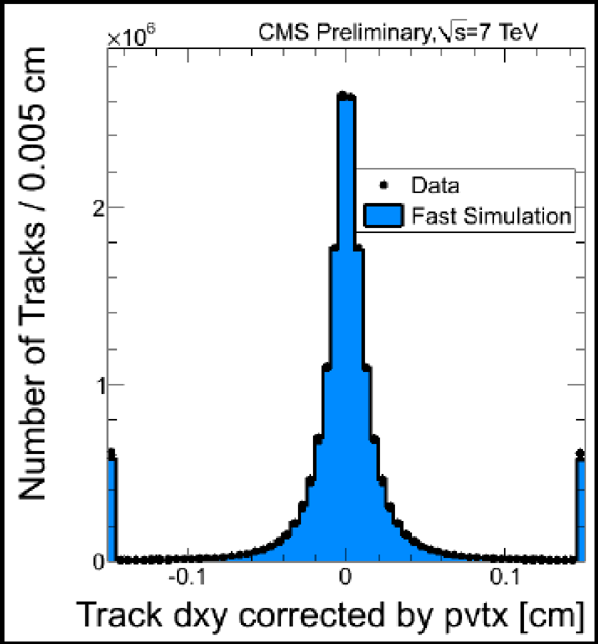
1. FAST TRACKING exploits the MC-truth:
   - it emulates the seeding efficiency using the hits from the charged particles of MC-truth;
   - the final track selection is done by the same modules as real tracking;
   - there are no fake tracks.

2. The option exists to run real tracking on the hits.

Tracking performances

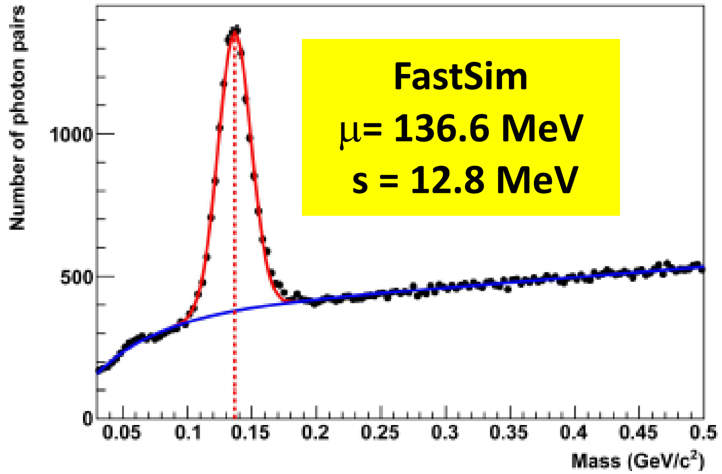




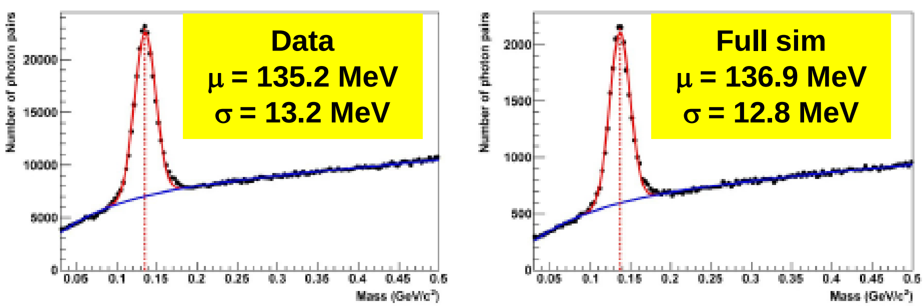


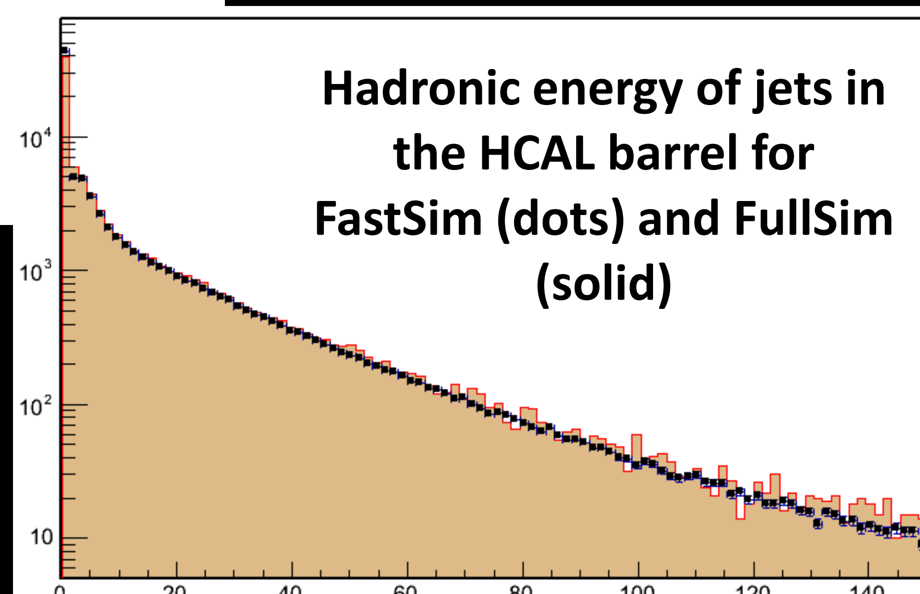
Excellent agreement with data after the basic quality cuts.

Calorimeters performances



Di-photon invariant mass in the ECAL barrel (photons reconstructed with the Particle Flow algorithm)





Hadronic energy of jets in the HCAL barrel for FastSim (dots) and FullSim (solid)

B-tagging performances

Same b-tagging modules are used as in data and in FullSim

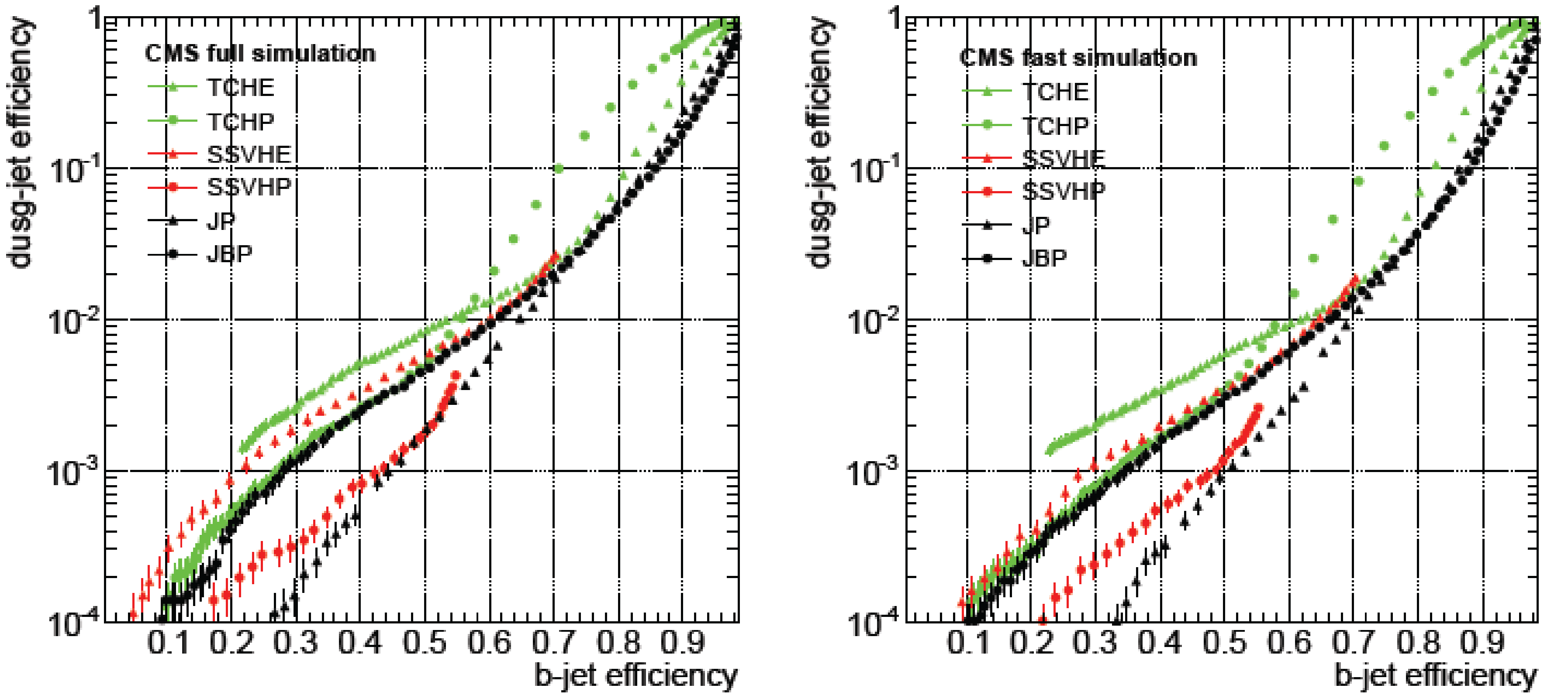


Figure 22: Light flavor mistag efficiency versus b-tagging efficiency in comparison for several b-tagging algorithms. On the left: full simulation, on the right: fast simulation.

B-tagging performances

Fake rate: generally lower | **Efficiency: generally higher**

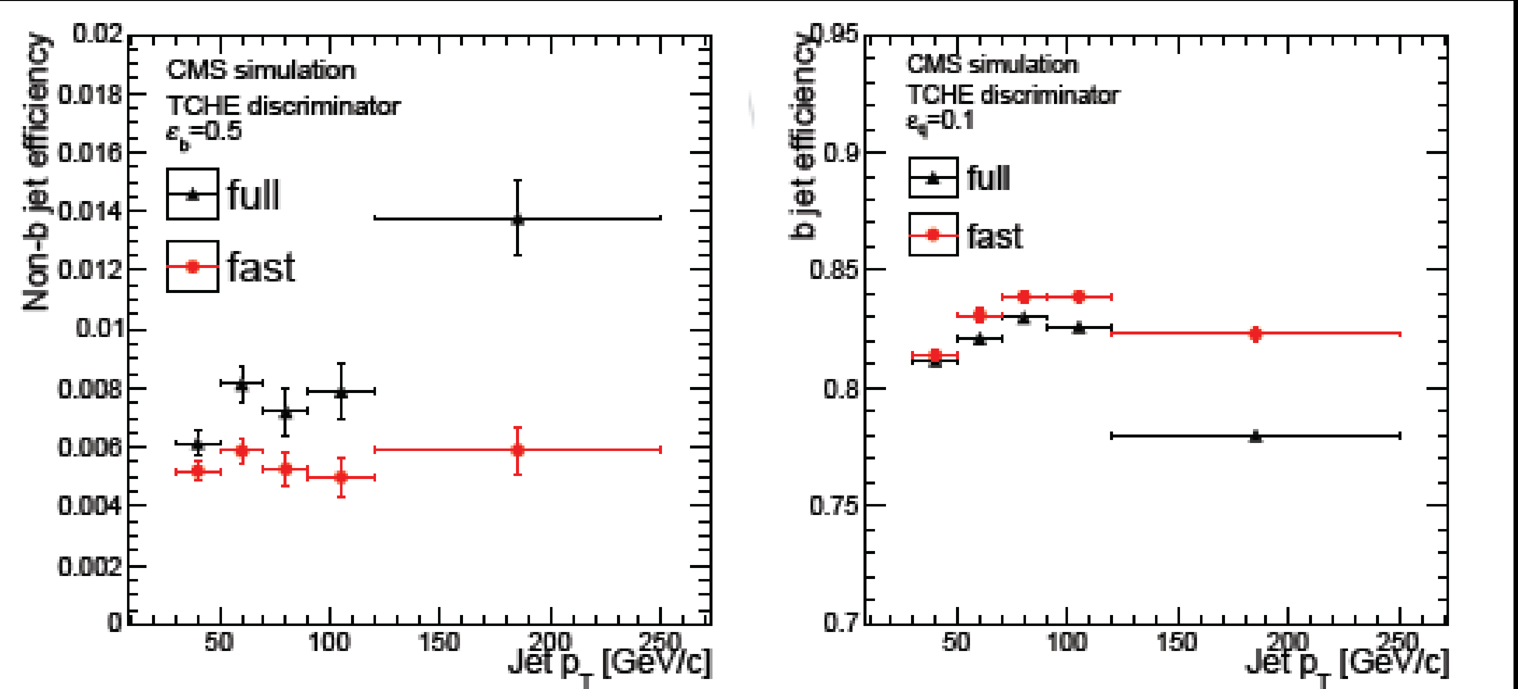
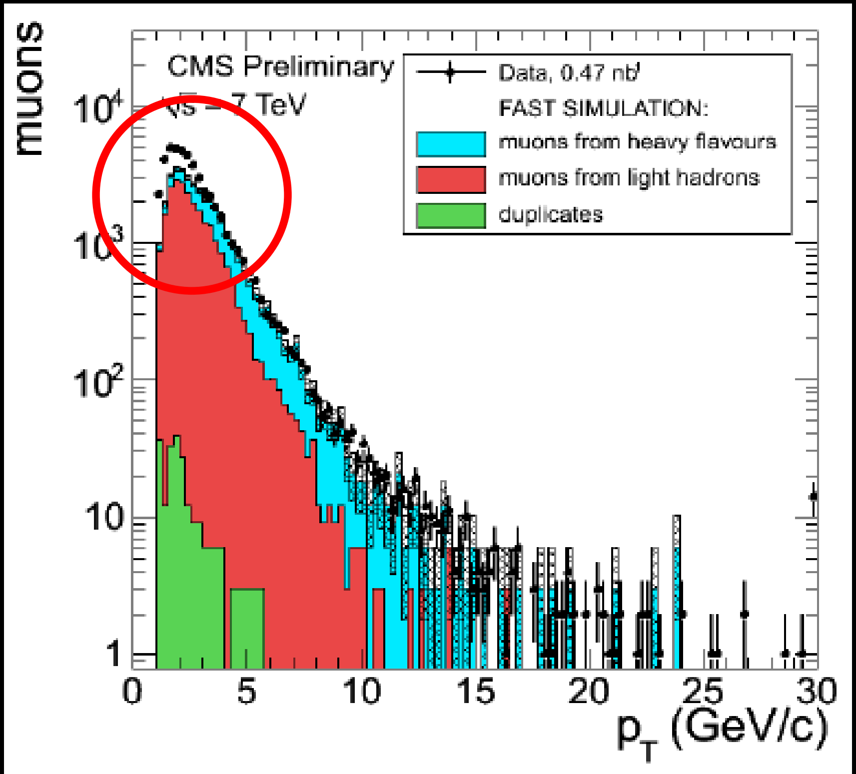
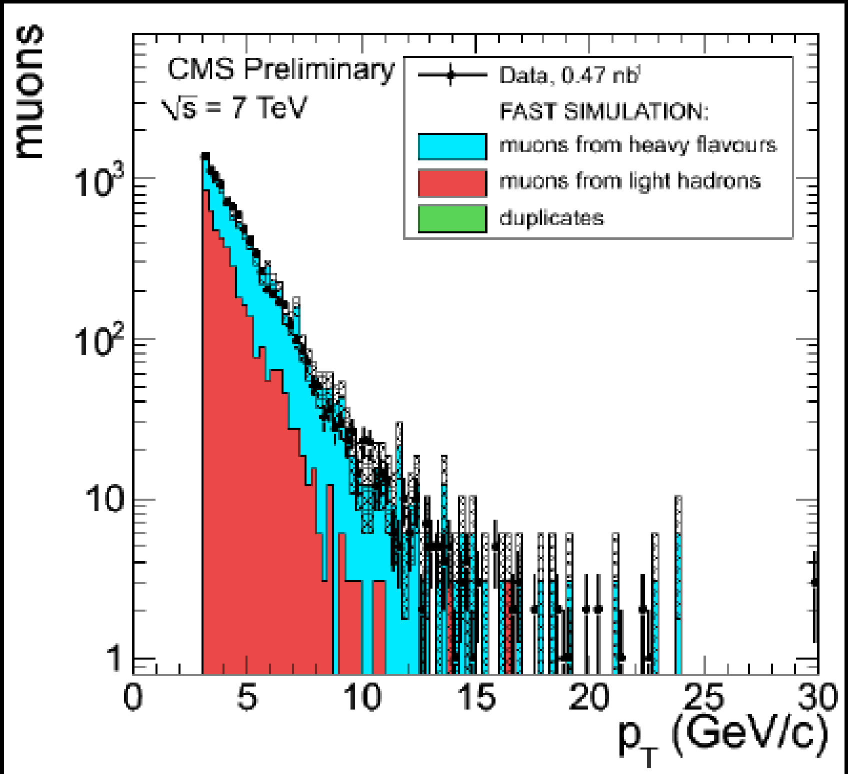


Figure 24: Comparison of the b-tagging performance between full and fast detector simulation for the track counting high efficiency algorithm. Left: mistag rate versus jet p_T at fixed b-tag efficiency of 50%. Right: b-tag efficiency versus jet p_T at fixed light flavor mistag rate of 10%.

The lower fake rate is attributed to the absence of fake tracks, while the absence of cluster merging (important source of track inefficiency in dense high-momentum jets) gives an higher efficiency.

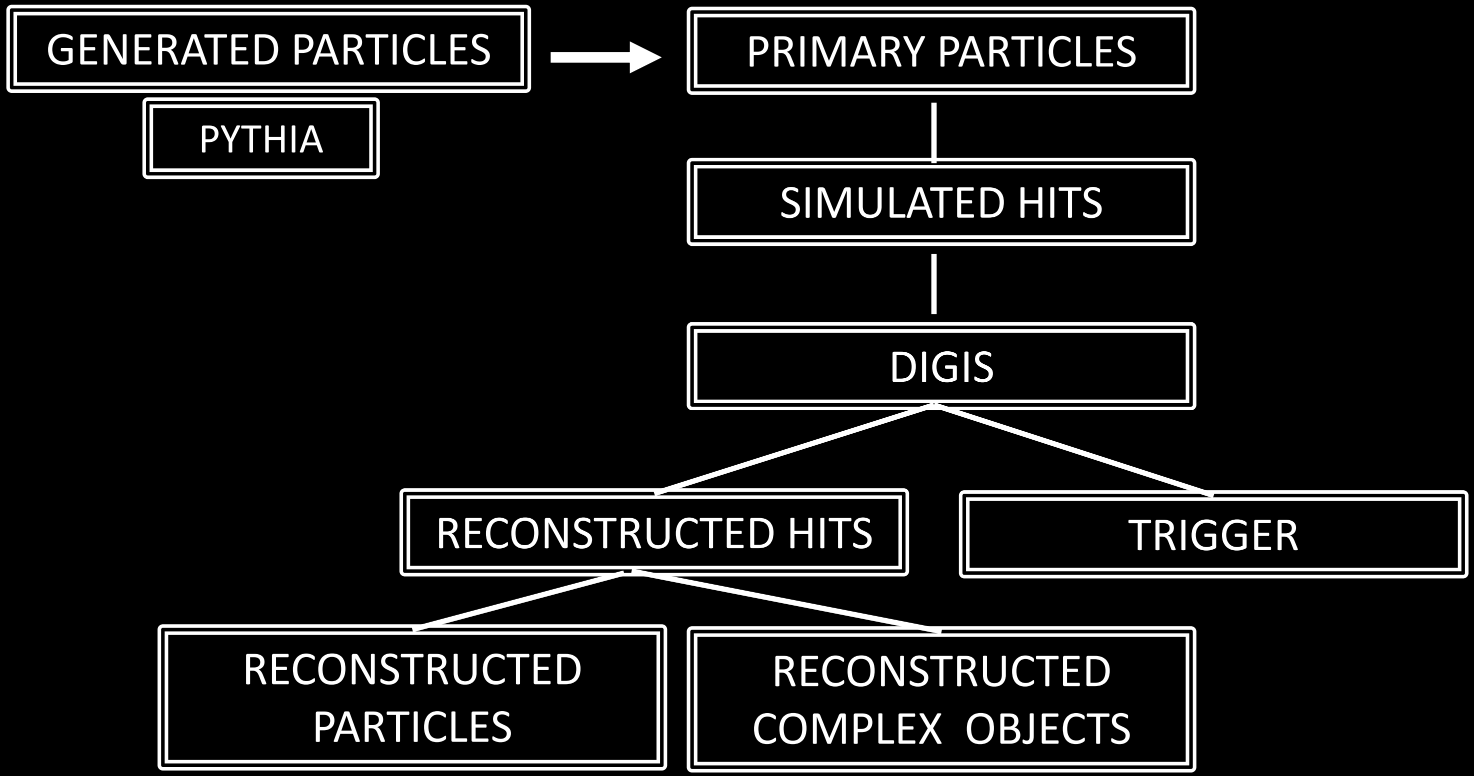
Muon performances





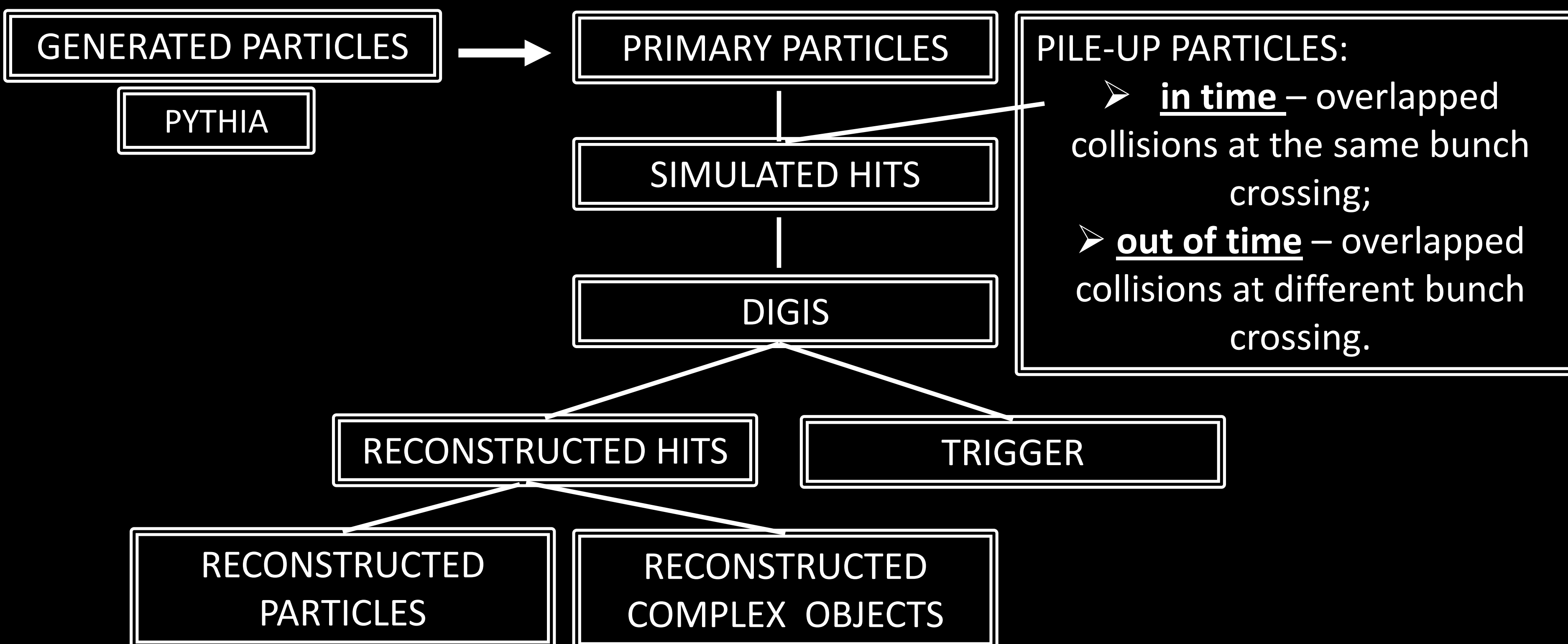
Low momentum discrepancy is because in Fast Sim there are no muons from hadronic punch-through. Excellent agreement after the tight-muon selection is applied, as commonly done in the high-pt analyses.

Fast Simulation workflow



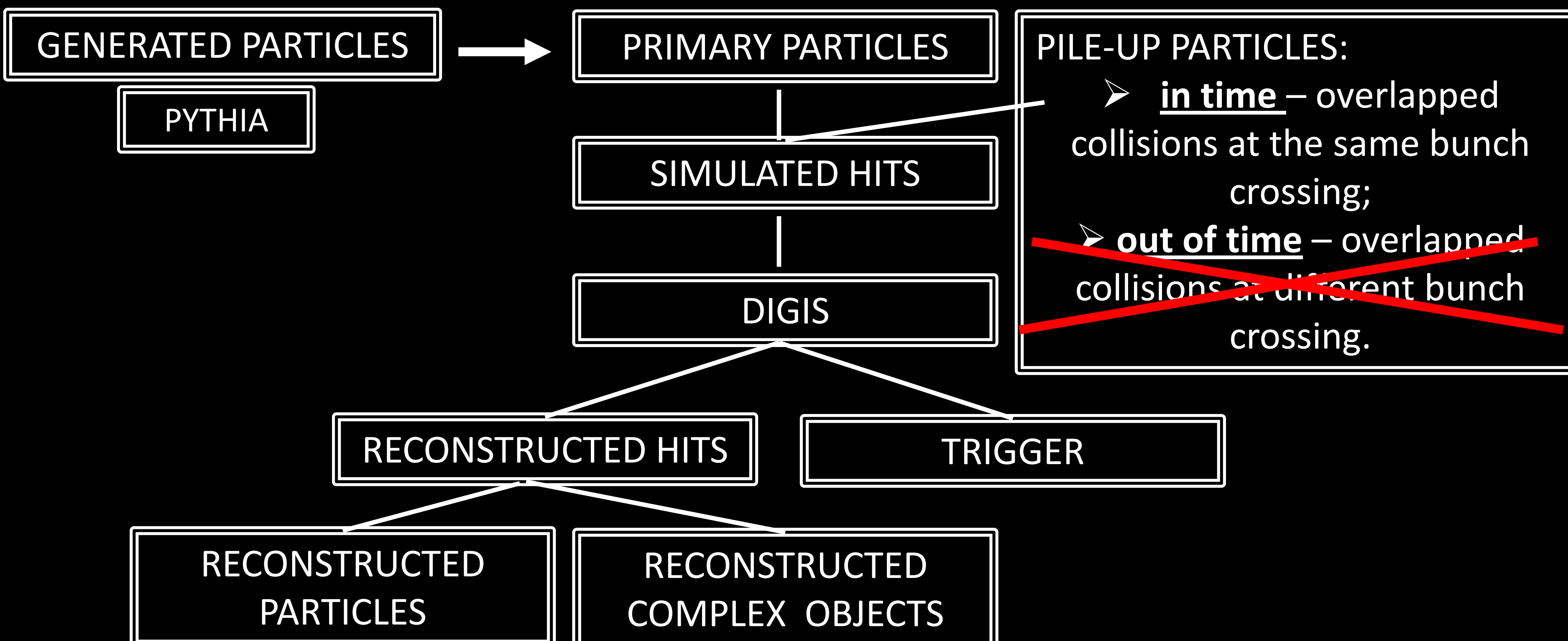
Fast Simulation workflow + PU

Default PU mixing in FastSim is performed at gen-level mixing:
It works very well at low PU but has cpu-time and memory issues with large PU and it is unable to simulate OOT.



Fast Simulation workflow + PU

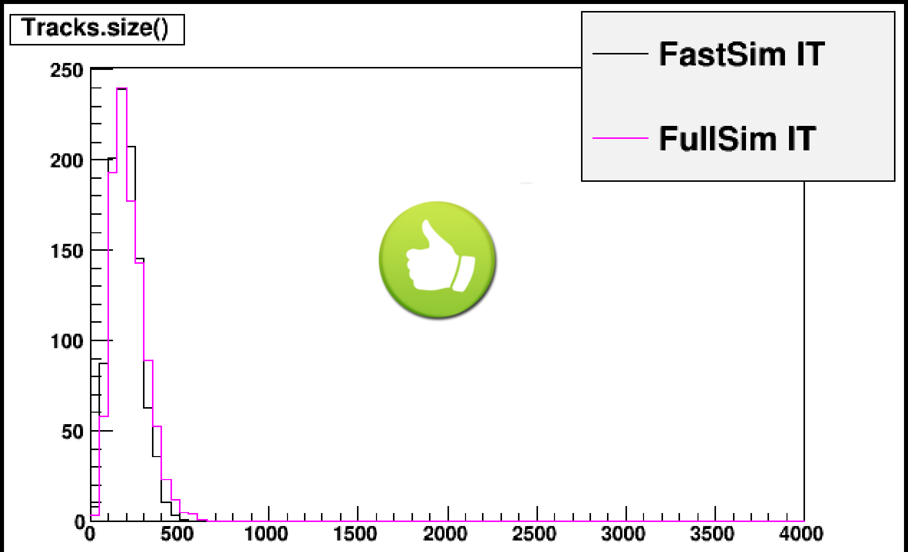
Default PU mixing in FastSim is performed at gen-level mixing:
It works very well at low PU but has cpu-time and memory issues with large PU and it is unable to simulate OOT.

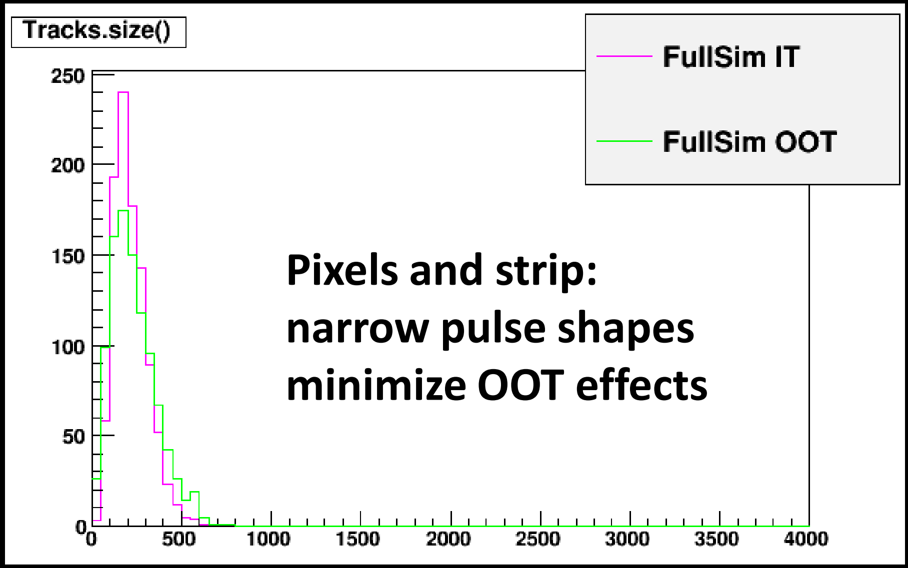


Pile up mixing

There is a work ongoin towards a **new mixing where:**

- **Calo & Muons** SimHits are mixed after the detector simulation, as in FullSim:
  - this profits from the integration with the digitizers
  - it allows for the same treatment of OOT PU as in FullSim
- **Tracker** tracks are mixed after reconstruction step, and only in-time (OOT would cost too much in cpu time.

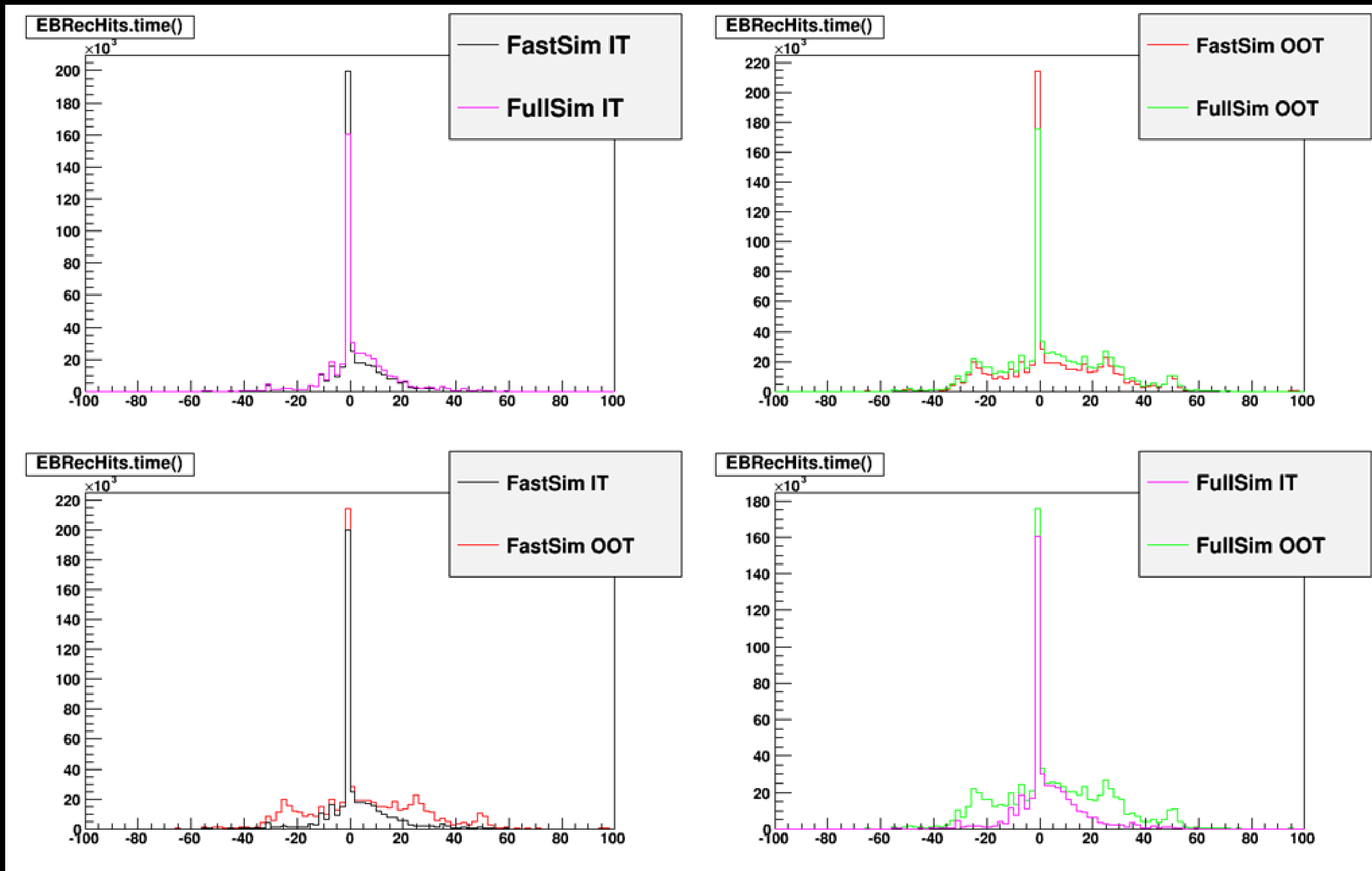




- Appealing possibility: mix FastSim signal with FullSim minimum bias events.

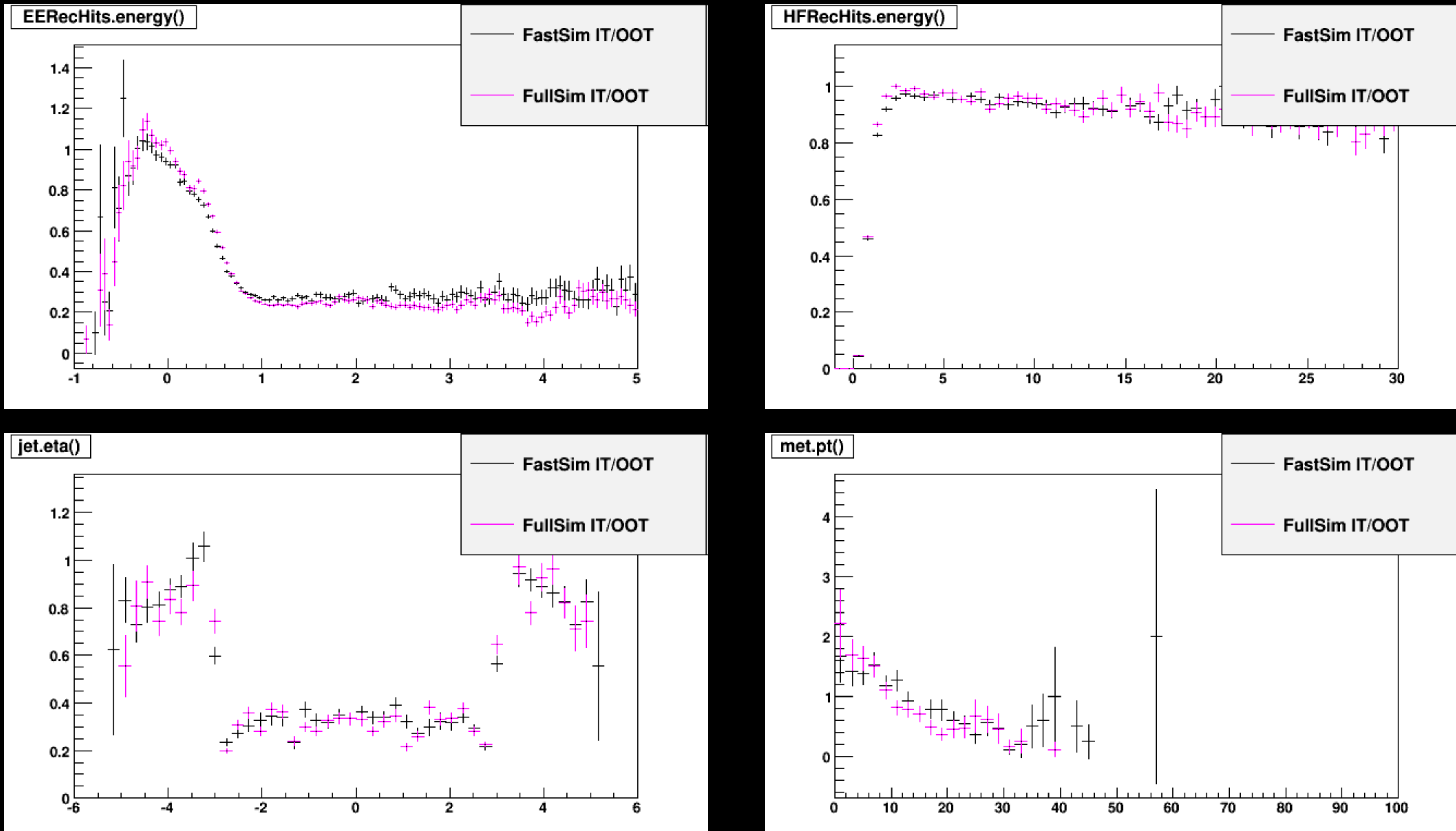
New mixing status: calorimeters

Plots in this and next slides are for single mu + <PU>~10, OOT 25 ns



Habemus cumulum intempestivum! (There is out-of-time pile-up)

PU performances: ratios IT/OOT



The agreement is very good, but the new mixing option is not available by default because the correct propagation of the mixed tracks to the rest of reconstruction is still work in progress.

Example of usage at CMS

- Centralized high-statistics productions:

– SUSY “simplified model scans”, used as signal samples for all publications of the CMS SUSY group since 2011 (FullSim used to validate a few points in the scan);

– Top (pair and single), W+jets, Z+jets samples with non-default parameters (masses, Pythia tune, QCD scales), for the evaluation of systematic errors.

- Private productions

– Signals of interest of one/few groups (e.g., black holes);

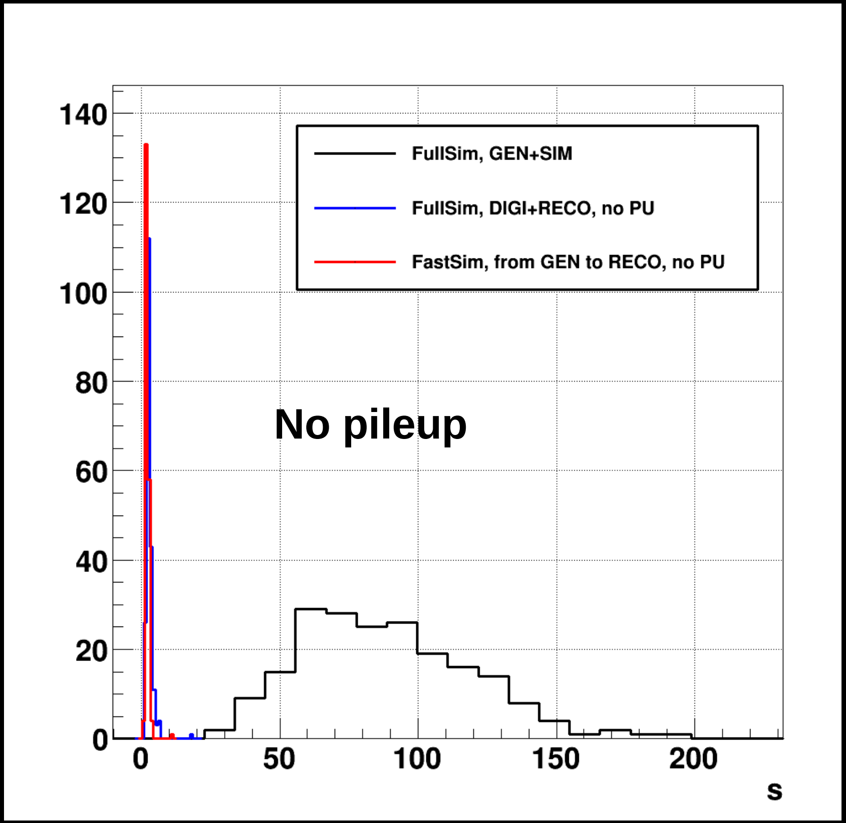
– Template extraction for scans of parameter of interest;

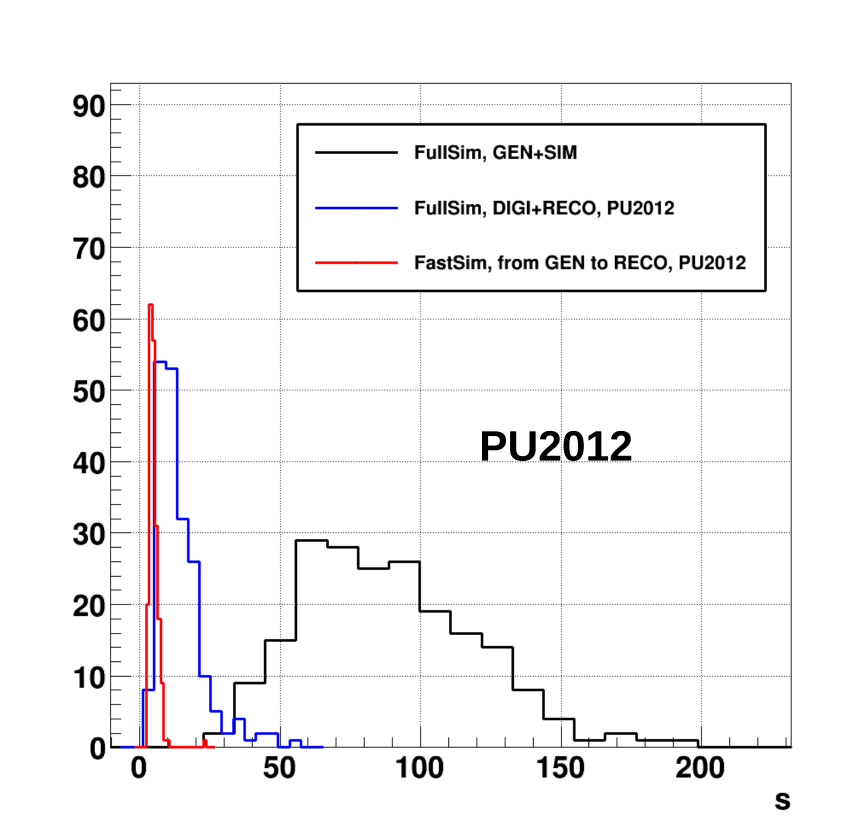
– MVA training;

– Enhance statistics of interesting events inside large backgrounds (filtering at RECO level is possible).

CPU time performance

- Latest CMSSW pre-release; numbers are in **seconds/event**
- "PU2012" is the PU profile of 2012 data: only in-time PU is officially used for FastSim, also OOT for FullSim.
- Machine: 64 bits, Scientific Linux 5.9, Intel(R) Xeon(R) CPU L5640 @ 2.27GHz





Summary

- The event production rates are about 100 times faster than the Geant4-based Full Simulation one with nonetheless comparable accuracy for most of the physics objects typically considered in the analyses.

- The use of Fast Simulation is important in particular for planning the forthcoming analyses at higher Pile-Up and luminosity of LHC;

- The current development is aimed to obtain a tighter integration with FullSim (common use of the digitizers, similar treatment of pileup) and more flexibility.

- The new mixing option will be available as soon as the technical issues will be solved.

Summary

- The event production rates are about 100 times faster than the Geant4-based Full Simulation one with nonetheless comparable accuracy for most of the physics objects typically considered in the analyses.

- The use of Fast Simulation is important in particular for planning the forthcoming analyses at higher Pile-Up and luminosity of LHC;

- The current development is aimed to obtain a tighter integration with FullSim (common use of the digitizers, similar treatment of pileup) and more flexibility.

- The new mixing option will be available as soon as the technical issues will be solved.

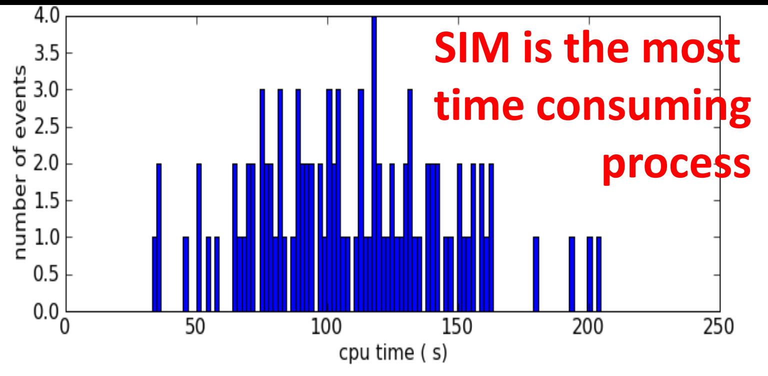
Dank! ☺

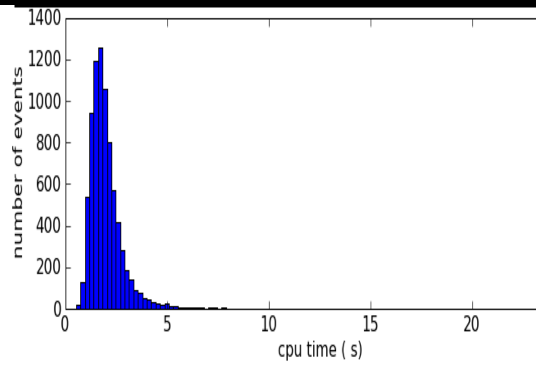
Back up

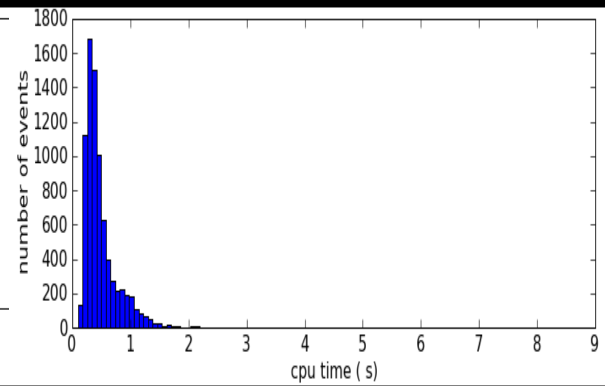
CPU time: FullSim vs FastSim

TTbar events **without Pile-Up** processed by **– slc5_amd64_gcc472 – 8 cores**

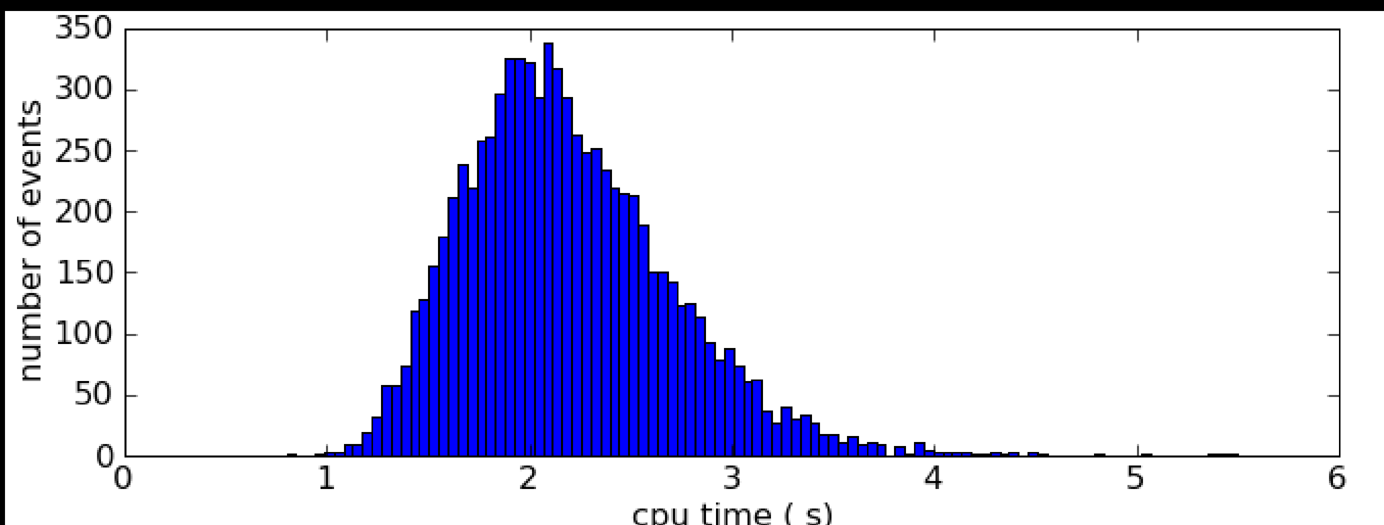
GEN, SIM,DIGI,L1,DIGI2RAW | RAW2DIGI,RECO | HLT







GEN, FASTSIM, HLT



Digitization

The digitization corresponds to the process of simulating the electronics signal read out by the detector. The software is subdivided into 3 domains:
SimTracker, **SimCalorimetry** and **SimMuon.**

For each one is simulated:

the **Analog to Digital Converter (ADC)** module

Subdetectors response proportional to the particle energy

Quantized signal

Threshold

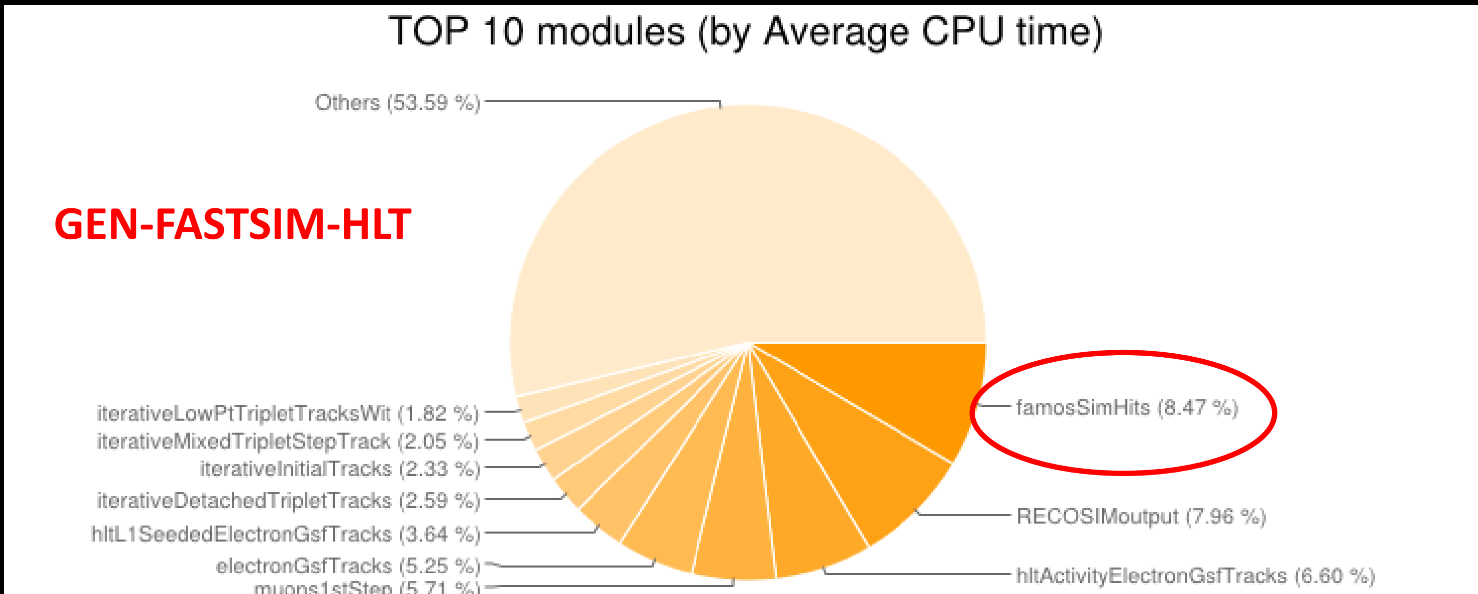
the **noise** from the electronic devices, other effects as **saturation** and **cross-talk** and the **Zero Suppression.**

Channels giving signal comparable with noise are ignored

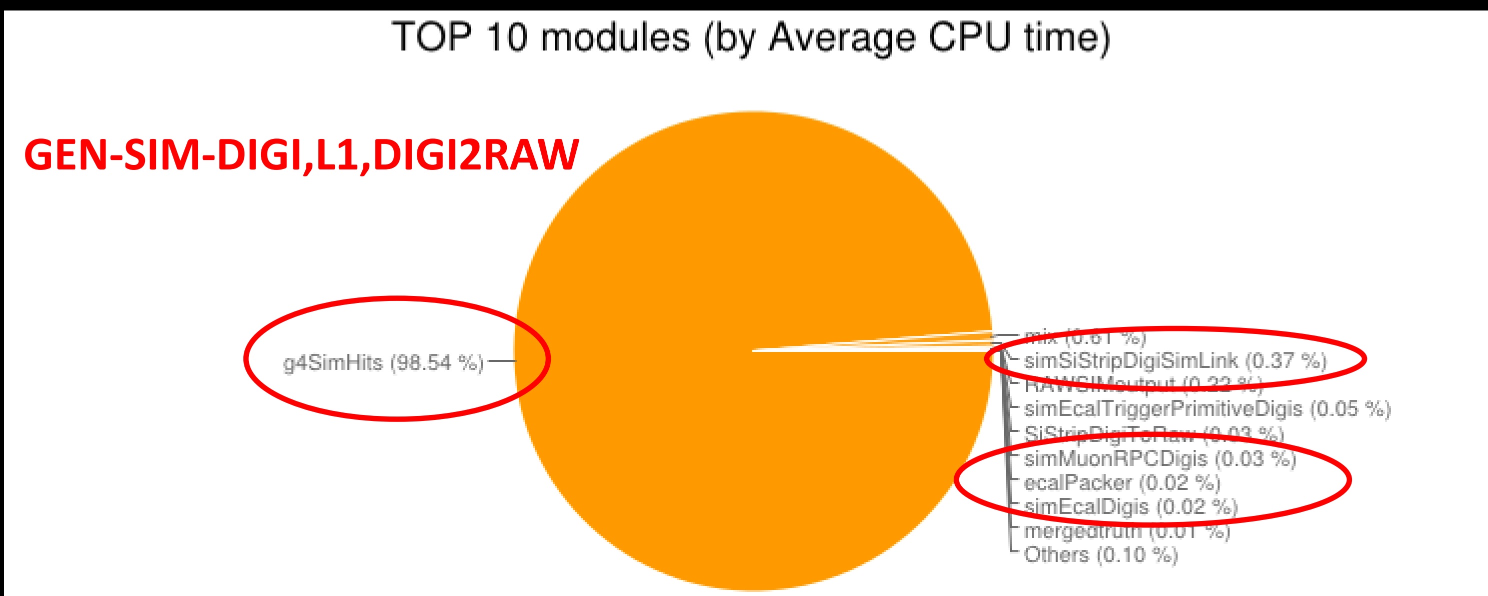
Why do we use the digitizers?

To obtain the same data format to give as input to the local reconstruction modules, and to simulate the electronic effects (especially noise) in the same realistic way as FullSim.
[see the EM isolation in the next plots]

Timing



In FullSim the timing for digis is dominated by the tracker.

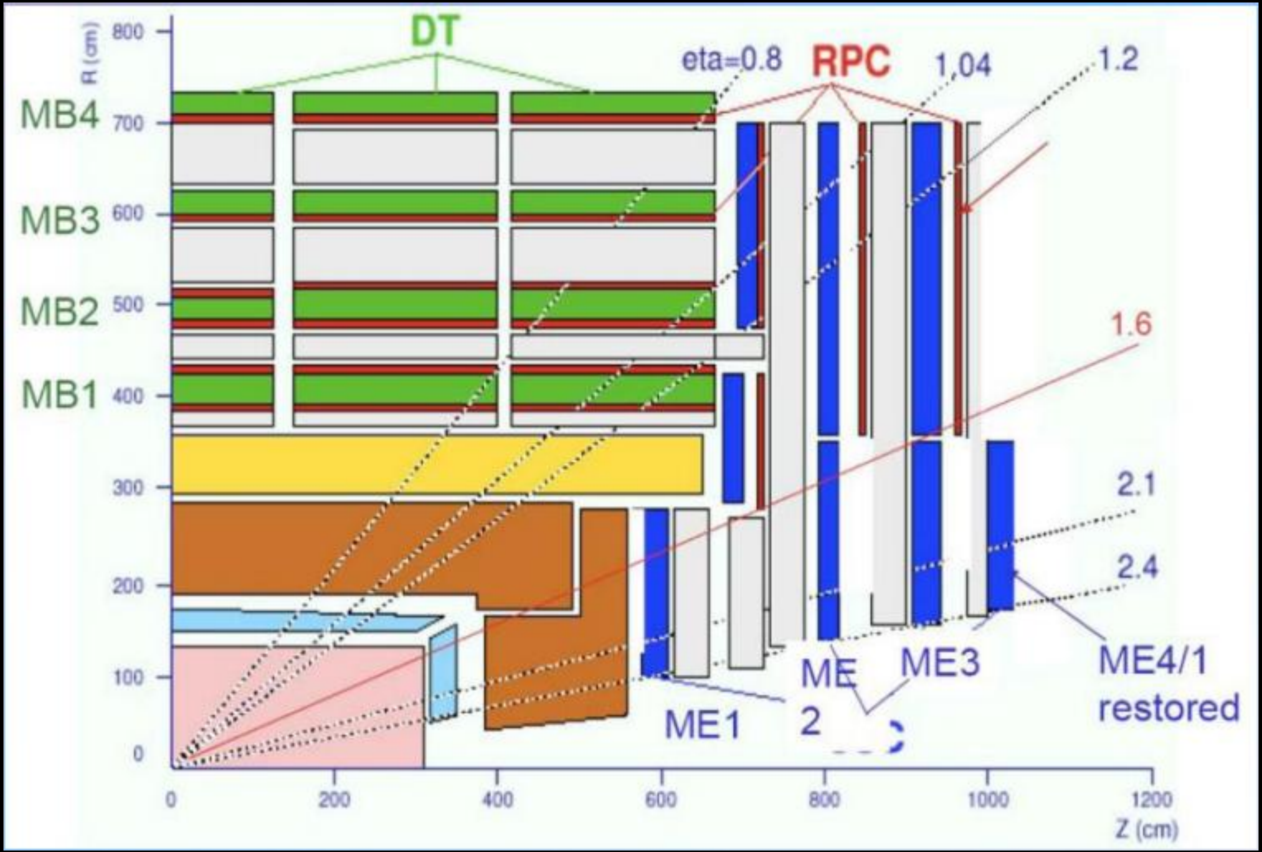


The digitization for calorimeters and muon chambers is not so time consuming.

Muon digitizer

The digitizers for the muon chambers were implemented since the beginning in Fast Simulation because: - is not so time consuming;
- was convenient to have a uniform code and tuning with FullSim;
- in particular was convenient for the trigger emulator.

the good agreement with FullSim and data is shown in A.Giammanco's talk



BARREL:
Drift-Tubes (DT),
Resistive-Plate-Chambers (RPC);

ENDCAP:
CSC and RPC.

DT, CSC, and RPC are composed by several chambers in turn containing strips or wires.

Calorimeter digitizers: Ecal

The digitizers for the calorimeters have just been implemented in FastSim (for the moment it can switch on separately Ecal and Hcal).

The process is similar in both calorimeters so an unified framework is used. The digis sequence of the whole Calorimeter system is divided in **ecalDigiSequence** and **hcalDigiSequence**;

Ecal digitizer:

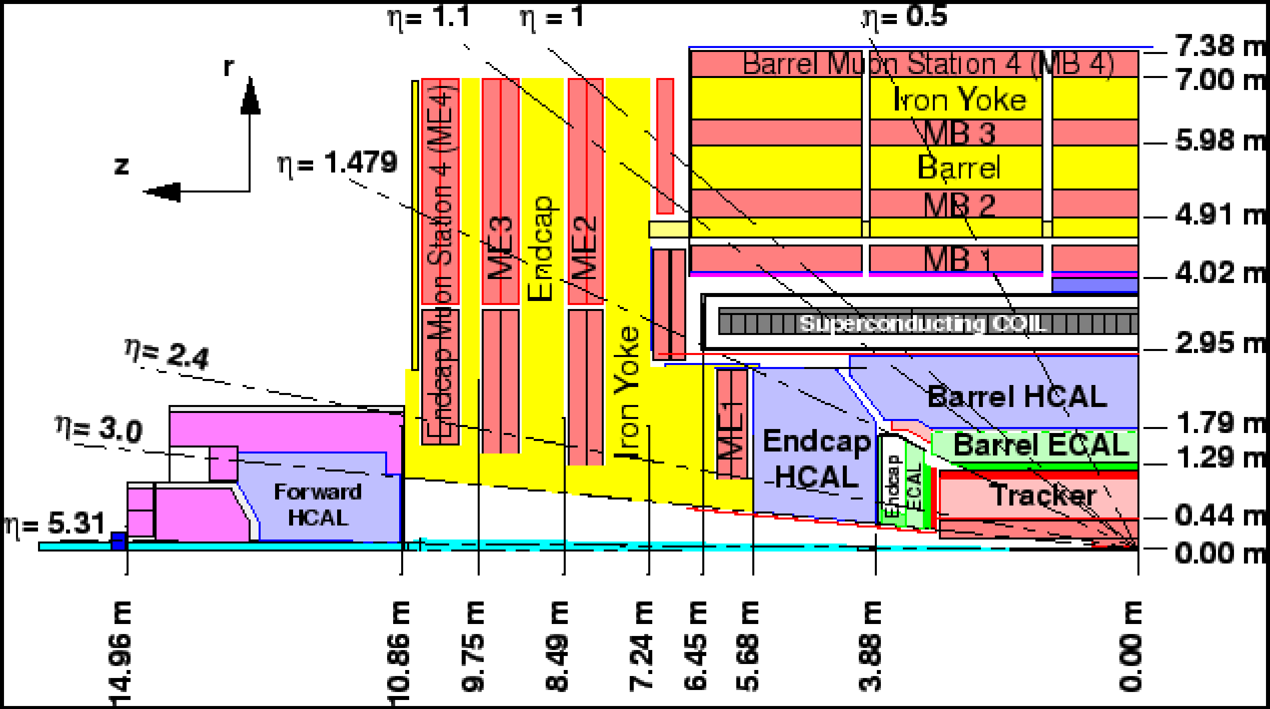
1. **conversion of the SimHit energy deposition in the channels into the electronics signal;**
2. **simulation of the noise, in EcalBarrel, EcalEndcap, and ES (Preshower);**
3. **modeling of the trigger primitives;**
4. **application of the zero suppression;**

Each module comes with its set of configurable parameters.

Calorimeter digitizers: Hcal

… for Hcal is a bit more complicated:

An HCAL Digi represents the signal in one readout channel, and consists of ten coded integers (time slice), each representing charge deposited in a 25 ns time bin.
The SimHits are the total signal in a given readout channel in a nanosecond.



Different electronic read out gives different units for these SimHits:
GeV for HB, HE, and HO, and in units of photoelectrons for HF.

Calorimeter digitizers: Hcal

Hcal digitizer:

1. conversion of the SimHits amplitudes into photoelectrons:
   - taking into account the gains;
   - modelling the channel response as well as a ring-dependent sampling factor.

2. The photoelectrons are then subject to Poisson statistics, and have their timing adjusted to correct for time of flight.

3. Conversion of each group of photoelectrons into an electronics pulse (units of fc);

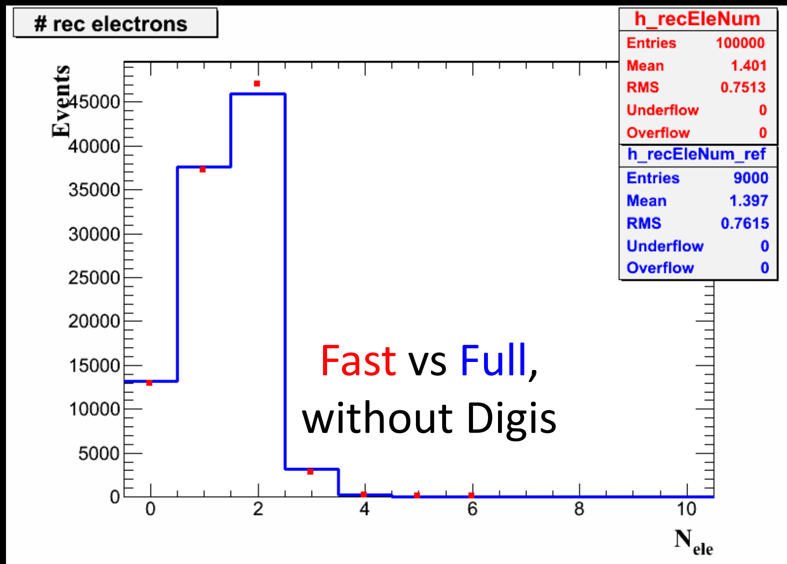
4. Simulation of the noise.

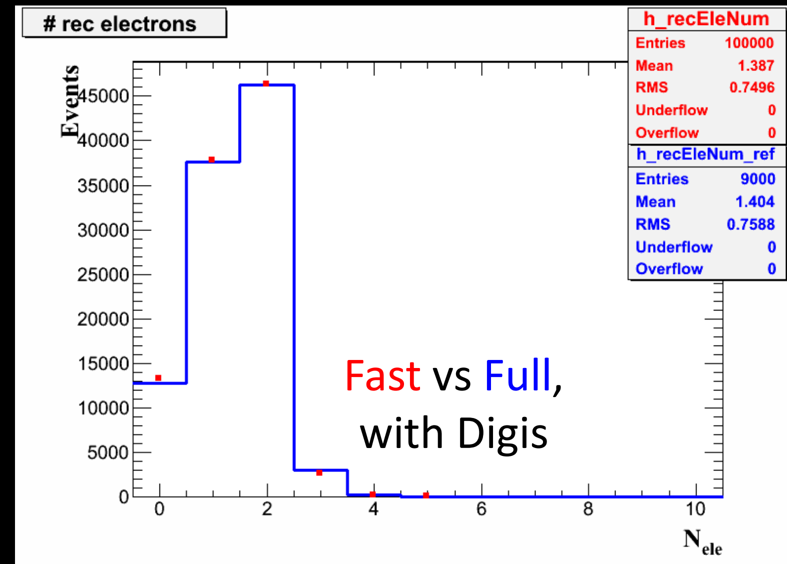
5. Trigger primitives.

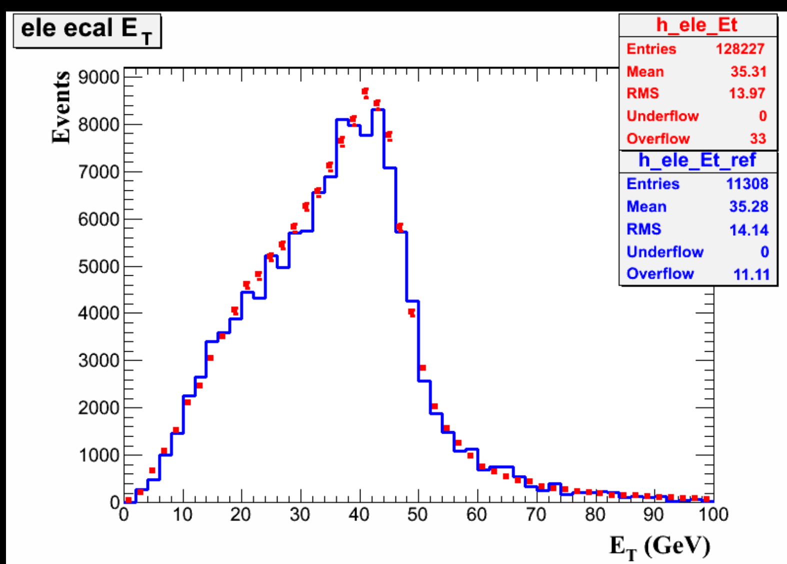
6. Zero suppression.

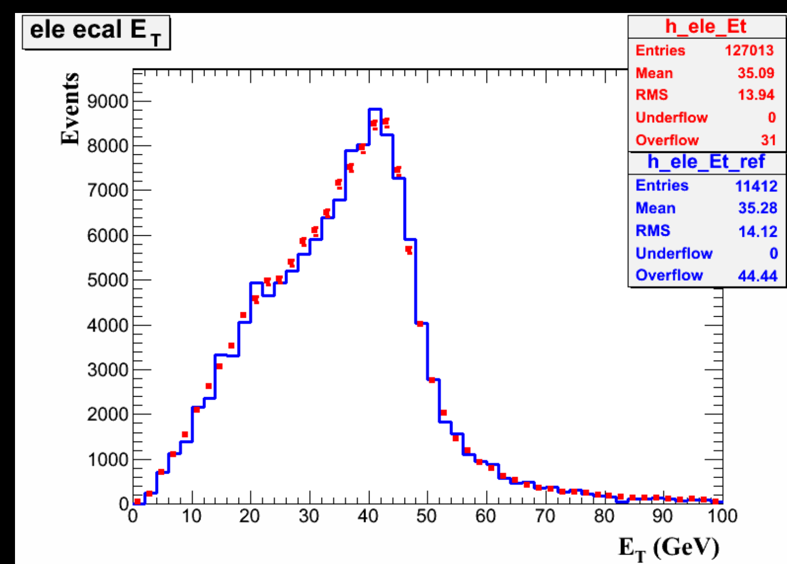
Ecal digitizers

Electrons in Z→ee events



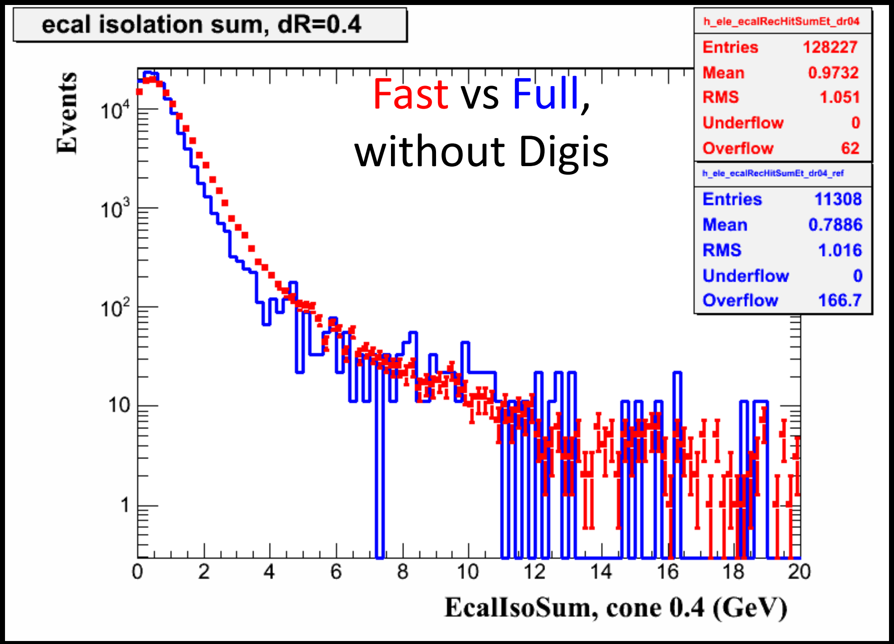


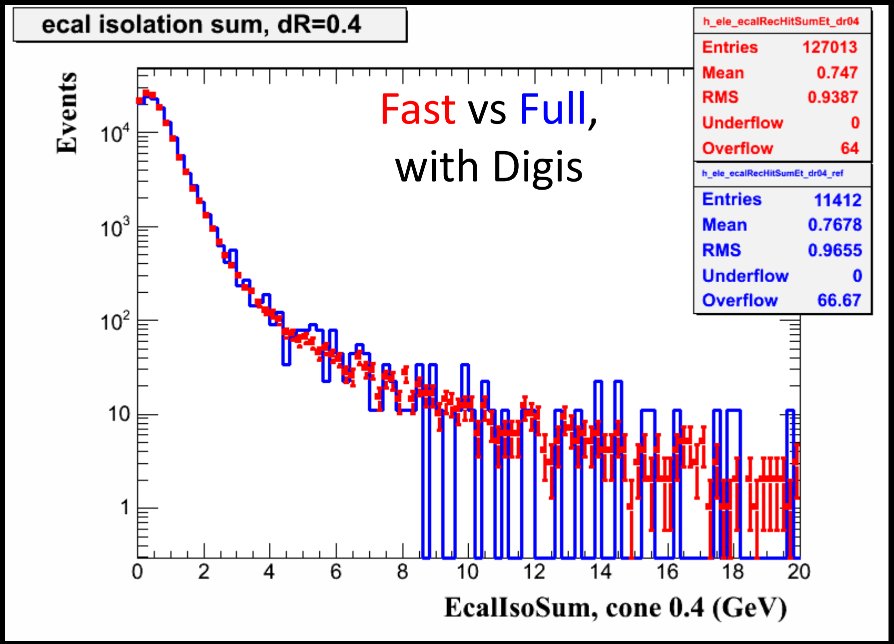




Ecal digitizer

Electrons in Z→ee events

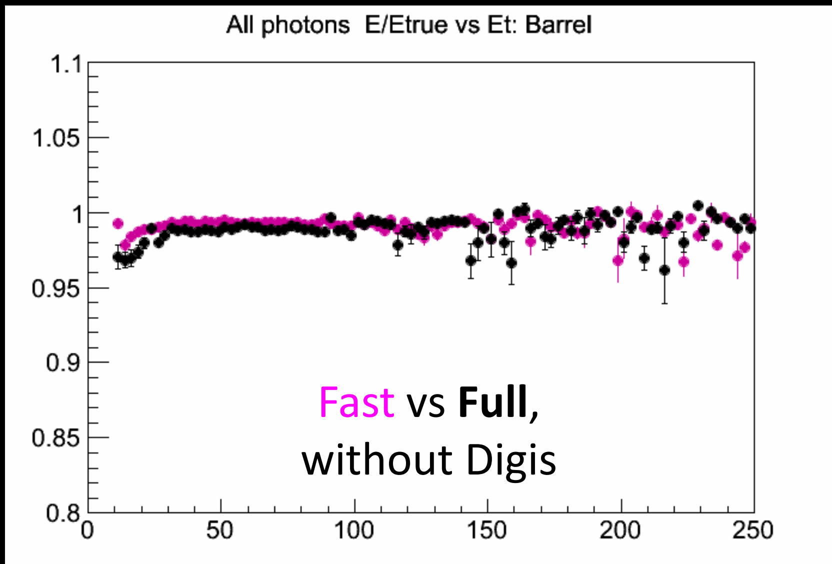


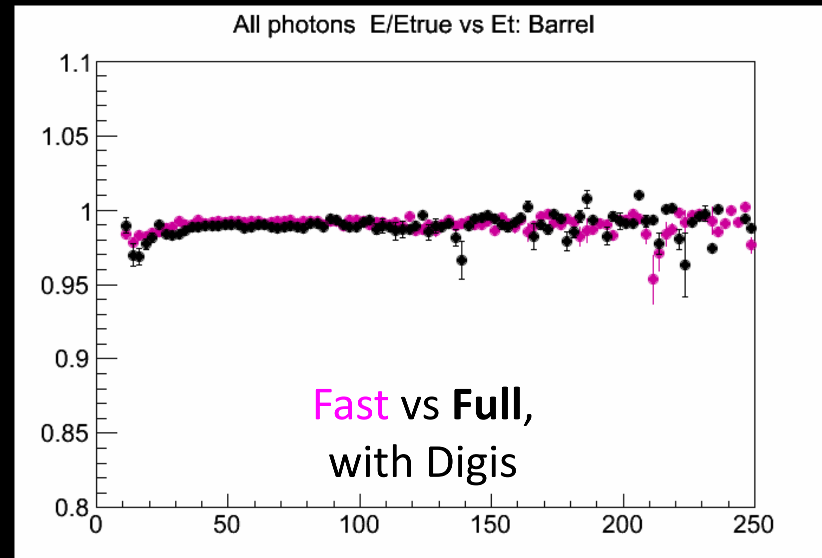


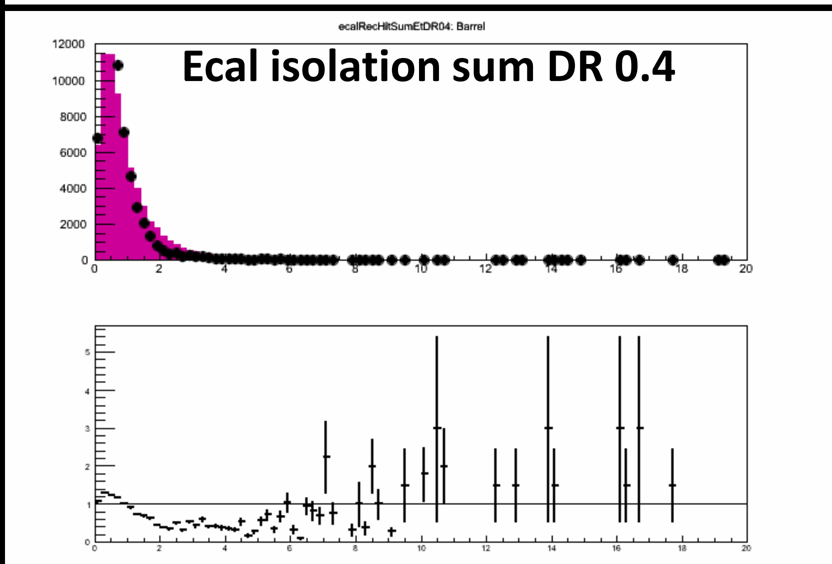
All Fast-Full comparisons stayed the same or improved.
Noise-related ones improved a lot

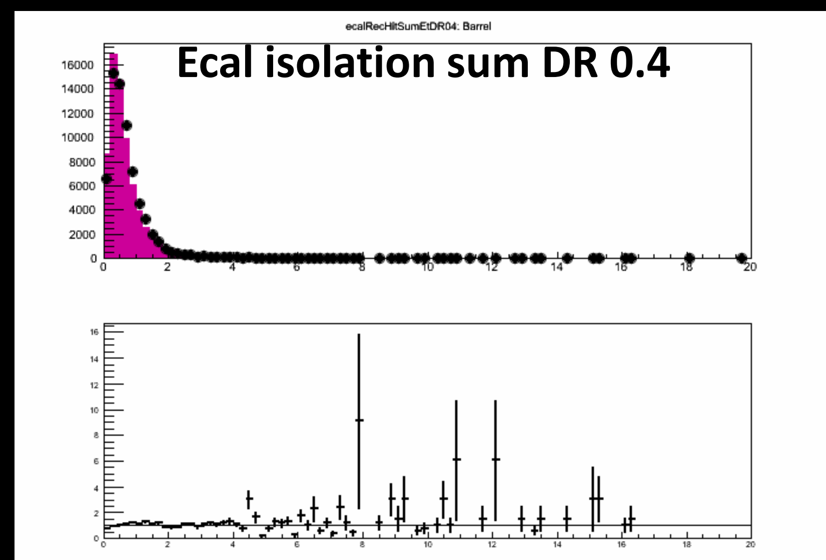
Ecal digitizers

Photons in H→γγ events



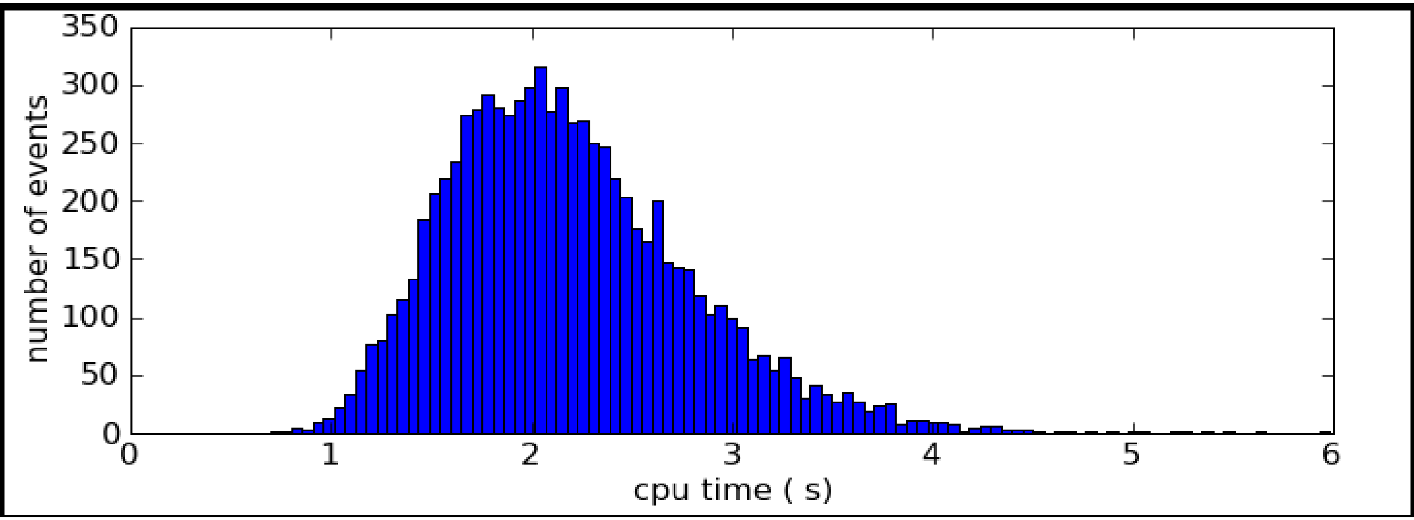




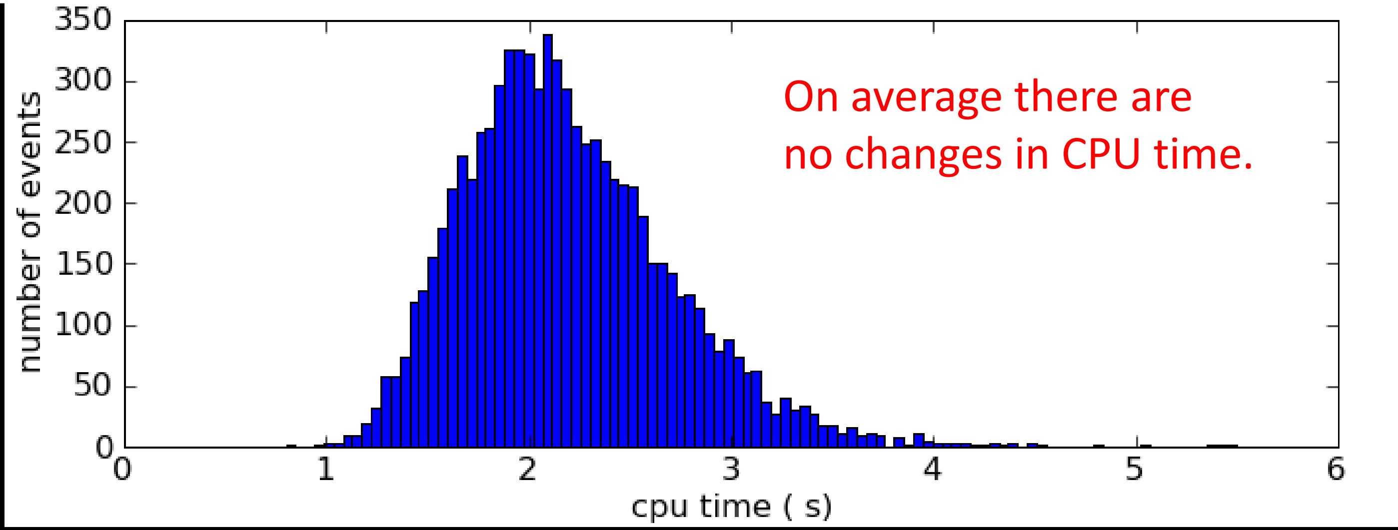


CPU time: FastNoDigis vs FastDigis

GEN-FASTSIM-HLT : without Ecal Digis



with Ecal Digis



Muon digitizer: DT

The DT digitizer simulates the signal collection in a single chamber:

- gas ionization;
- electrons propagation in the electric field geometry;
- drift velocity of electrons (400 ns);
- dead time;
- propagation of the signal along the anodic wire to the front end modules;