CSFI 2008

Calcolo Scientifico nella Fisica Italiana

Sessione di Fisica Nucleare e Subnucleare

Monte Carlo simulations in High Energy Experiments

G.Corti

Rimini 27-30 May

Senigallia 31 May

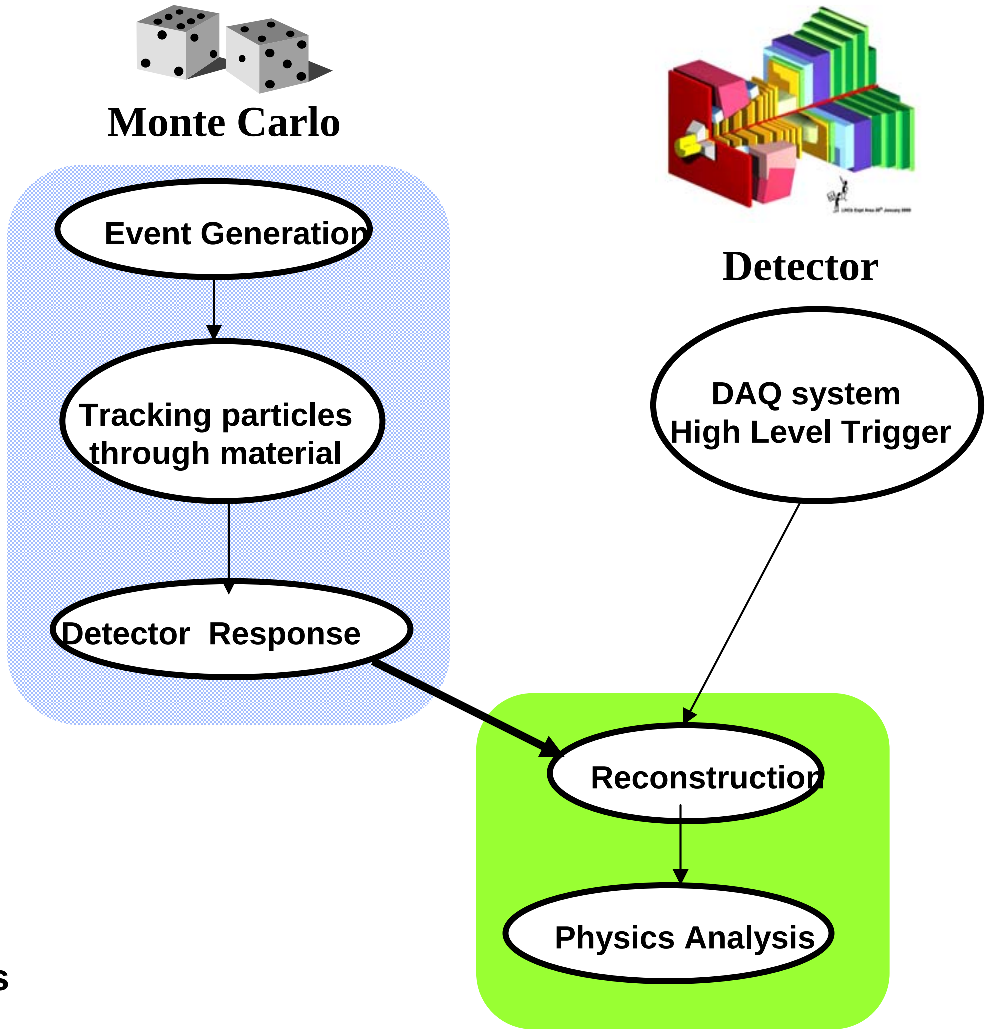
Outline

- **Introduction**
  - **Why and how Monte Carlo simulations are used in High Energy Experiments**
- **Walk through the various aspects covered by MC simulations in experiments**
  - **Generation**
  - **Tracking**
  - **Detector response**
- **How does it all fit together**
  - **LHCb example**
- **Statistics and CPU**

Introduction

- In High Energy Experiments when elementary particles collides in accelerators (for example)
  - unstable particles are created, these particles decay quickly.
  - it is necessary to reconstruct an “image” of the event through measurements by complex detectors comprising many subdetectors.

- The Monte Carlo simulation role is to mimic what happens in the spectrometers to understand experimental conditions and performance.
  - Monte Carlo data processed as real data in Reconstruction and Physics Analysis
    - BUT we know the “truth”
  - Comparing the simulation with what is measured in reality we can interpret the results



Why to use Monte Carlo simulations

- **Want to generate events in as much detail as Mother Nature**
  - **get average and fluctuations right**
  - **make random choices, ~ as in nature**
  - **an event with n particles involves O(10n) random choices**
    - **multiple variables: flavour, mass, momentum, spin, production vertex, lifetime,...**
  - **At LHC: ~ 100 charged and ~ 200 neutral (+ intermediate stages)**
  - **➔ several thousand choices**

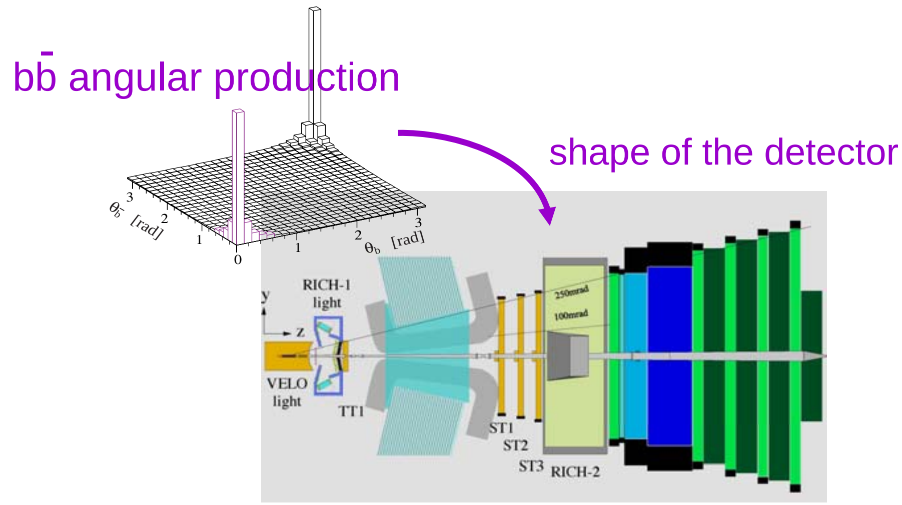
- **This applies also to the transport code through the spectrometer and the detectors response**
  - **want to "track" the particles in the geometrical setup and have them interact with the matter**
    - **energy loss, multiple scattering, magnetic field**
  - **want to simulate the detection processes and response of a given detector**
    - **ionization, scintillation, light**
  - **the interaction events are stochastic and so is the transport process**

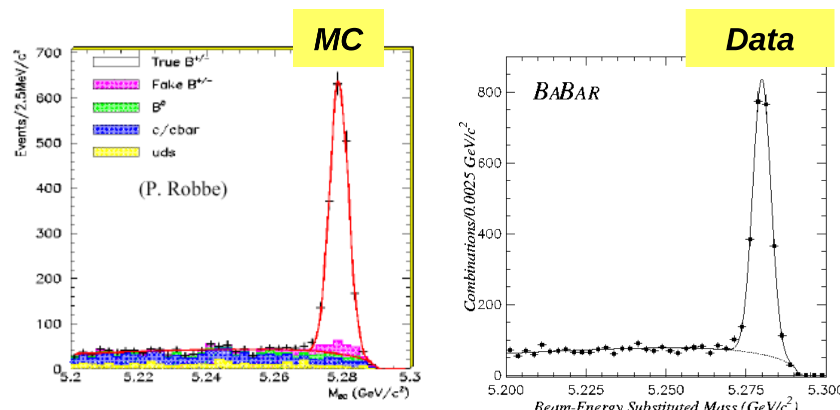
- **A problem well suited for Monte Carlo method simulations**
  - **computational algorithms relying on repeated random sampling to compute their results**

- **In fact a Monte Carlo simulation in a High Energy Experiment is a collection of different Monte Carlo, each specialized in a given domain, working together to provide the whole picture**

How are MC simulations used ?

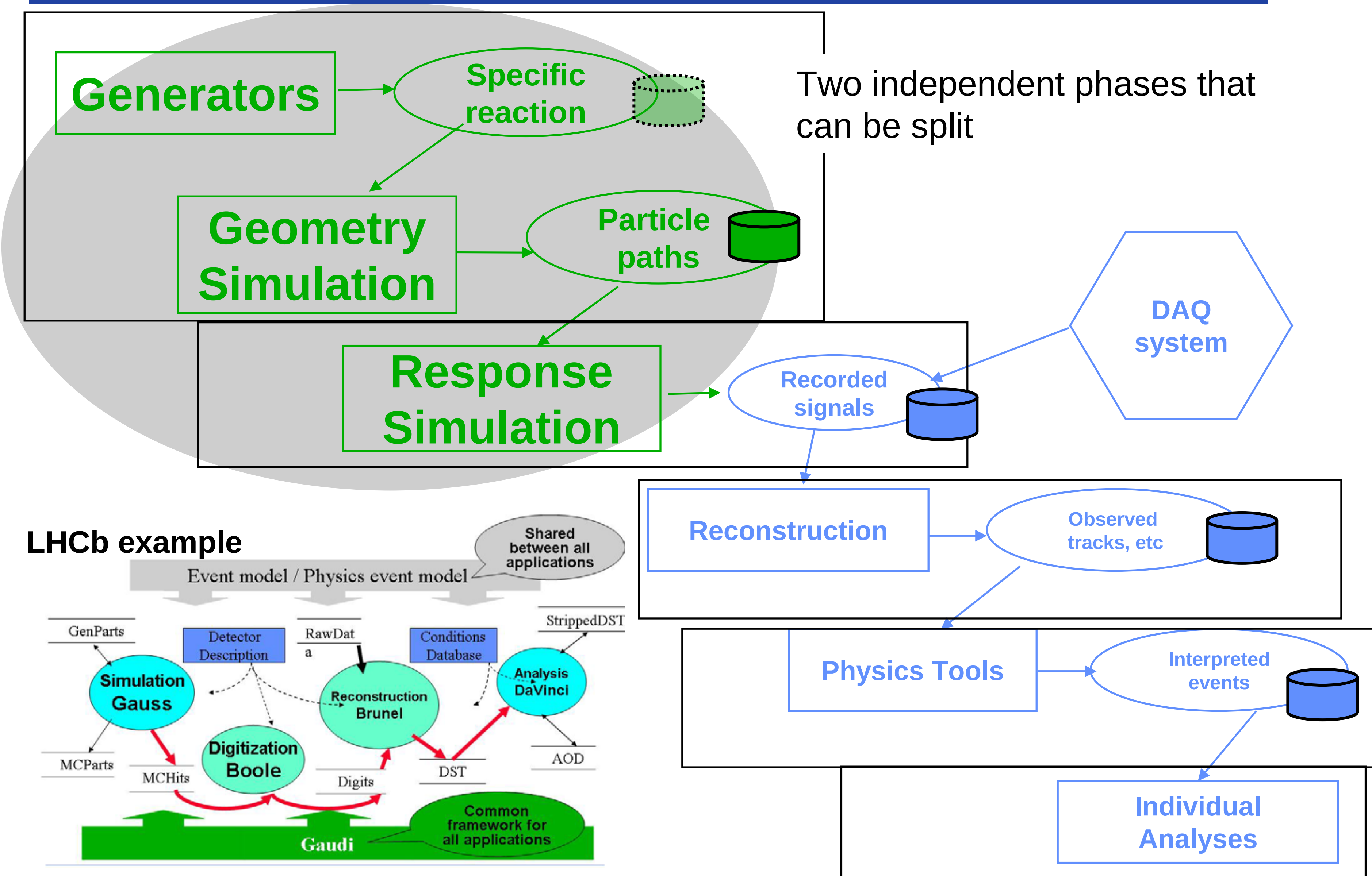
- **Detailed simulations are part of HEP physics**
  - **Simulations are present from the beginning of an experiment**
    - **Simple estimates needed for making detector design choices**
    - **Develop reconstruction and analysis programs**
    - **Evaluate physics reach**
  - **We build them up over time**
    - **Adding/removing details as we go along**
  - **We use them in many different ways**
    - **Detector performance studies**
    - **Providing efficiency, purity values for analysis**
    - **Looking for unexpected effects, backgrounds**
    - **When theory is non well known compare to various models and accounts for different detector "acceptance"**





B →J/Ψ K*

Traditional flow of simulated data and applications



Experiments simulation software

- **HEP experiments have their own software frameworks and use external packages developed in the physics community for Generators and for Transport in the detectors**
  - **Athena (ATLAS), CMSSW (CMS), Gauss (LHCb), VMC (ALICE, CBM@GSI, Minos), bbsim (BaBar), etc.**

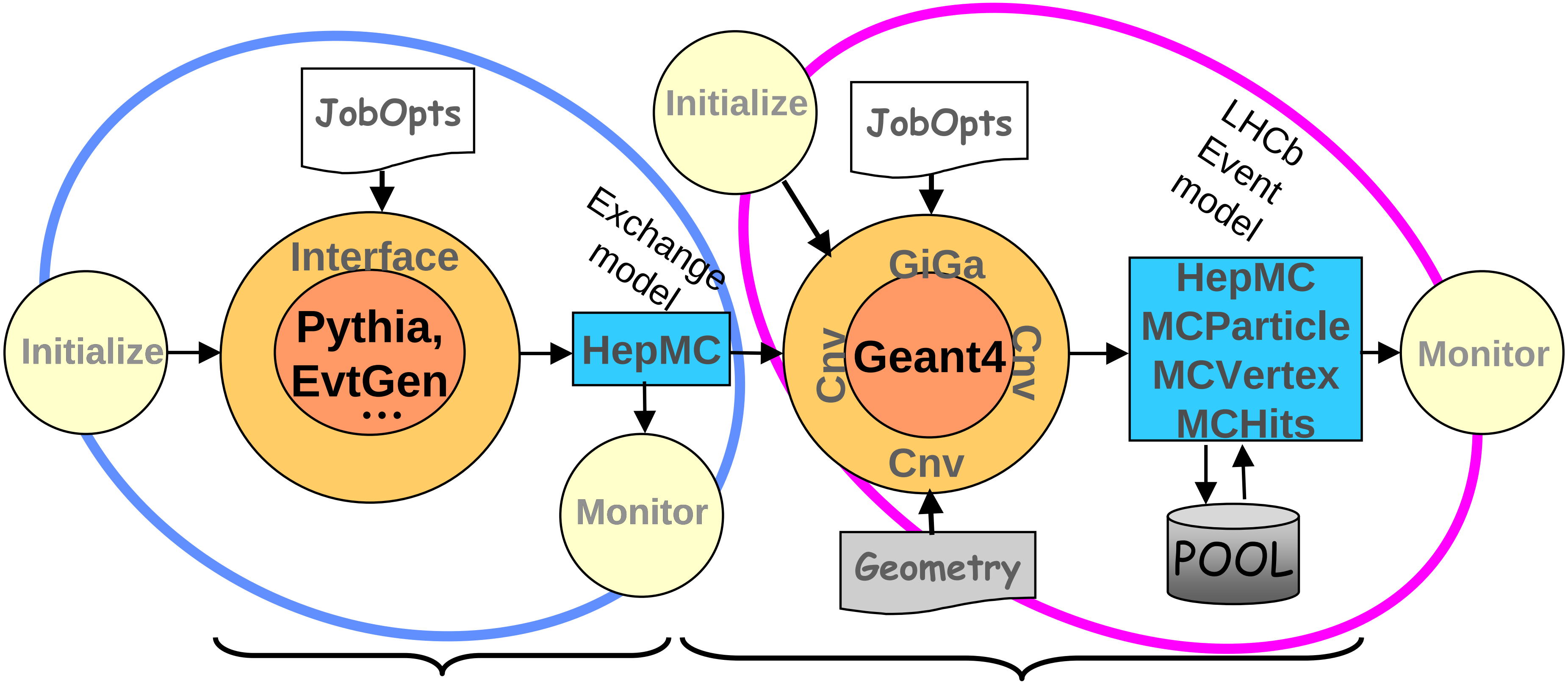
- **Response of the detectors is often in-”house” and requires detectors experts**
  - **tuned first with test beam data, then with measurements in the experiment**

- **Experiment applications provide**
  - **Interfaces to event generators**
  - **Interface to transport code**
  - **Event model for MC truth and Persistency**
    - **Access to snap-shots of process to understand what happened**
  - **Histograms, messaging**
  - **Physicists in the experiments are shielded from Generators and transport code (eg. Geant4) to different degrees**
    - **different in the various experiments**
    - **different for different roles**

Structure of a simulation application

Gauss the LHCb simulation application as example

Two INDEPENDENT phases normally run in sequence in a single job



Event Generation

primary event generator
specialized decay package
pile-up generation

Detector Simulation

geometry of the detector (LHCb → Geant4)
tracking through materials (Geant4)
hit creation and MC truth information (Geant4 → LHCb)

Event generation

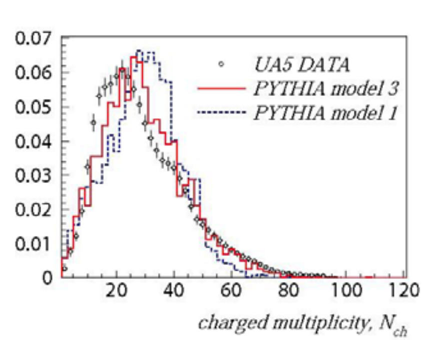
- **Many programs available in the physics community to generate primary collisions**
  - **specialized for hard processes, resonance, decays, parton showers, hadronization, etc.**
    - **often best at given task, but not always directly usable by experiments**
  - **general-purpose**
    - **PYTHIA (T.Sijostrand et al.)**
    - **HERWIG (G.Marchesini et al.)**
    - **ISAJET (H.Baer et al.)**
    - **SHERPA (F.Kraus et al.)**
  - **LCG Generators Service at CERN provides libraries for many of them**
    - **http://lcgapp.cern.ch/project/simu/generator/**
- **Experiments generally use one generator for massive production and make smaller data sets with others**
    - **Pythia/Herwig have different hadronisation mechanism (clusters as opposed to strings) for example**

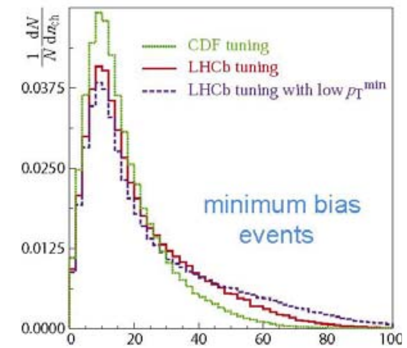
  and specialized codes when necessary
    - **BaBar and LHCb use EvtGen (D.Lange and A.Ryd) for b-hadrons decays**
    - **ATLAS and CMS use MCatNLO (S.Frixione, B.Webber) , AlpGen (M.Mangano et al) for matrix elements to feed to general purpose**
    - **ALICE uses Hiijing (M.Gyulassy and X.-N. Wang) for Pb-Pb interaction**

Generator phase in practice

- **In LHCb as example in Gauss proton-proton interactions at E_{cm}= 14 TeV are generated using Pythia 6**
  - **First simulate the hard processes in colliding particles, then the subsequent hadronisation**
  - **LHAPDF for Parton Density Functions**

  particles are then decayed by EvtGen (+PHOTOS and Pythia)
- **Various type of events generated**
  - **Minimum bias**
    - **includes hard QCD processes, single and double diffractive events**
  - **Signal b and bb events**
    - **obtained from minimum bias events with b or b-hadron**
  - **Cosmics and "particle guns (ie. a given particle with given kinematic)"**

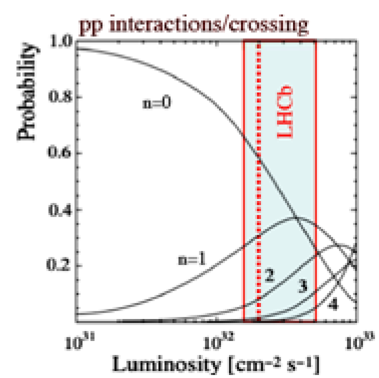




Choose Pythia setting that would agree with experiments measured particle multiplicities

- **Simulate running conditions**

  Luminous region with smearing of primary vertex according to bunch sizes

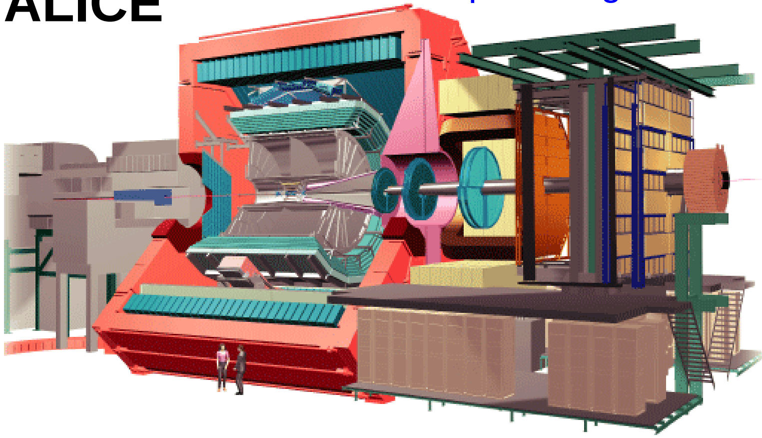


Pileup (number of pp collisions per bunch) at running luminosity

Detectors simulation, eg. the LHC experiments

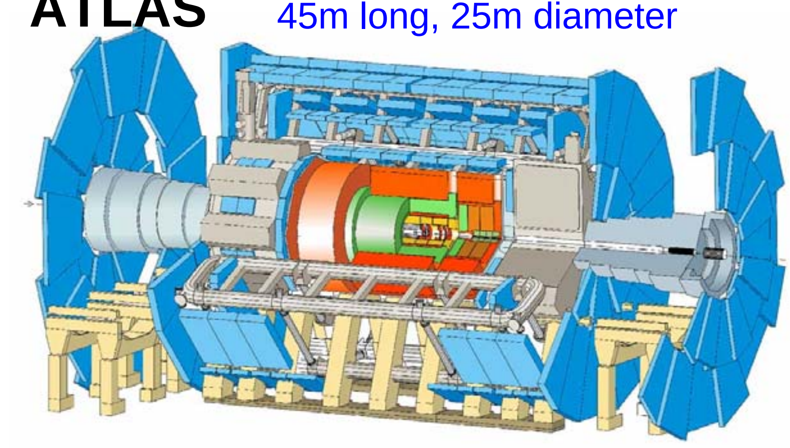
ALICE

Solenoid and dipole magnets

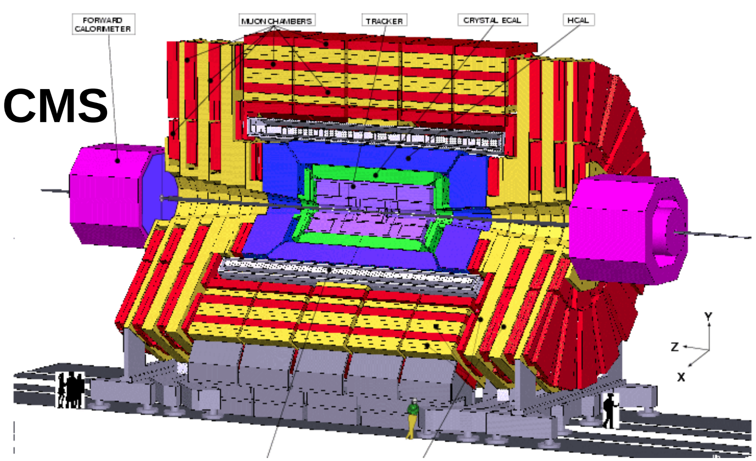


ATLAS

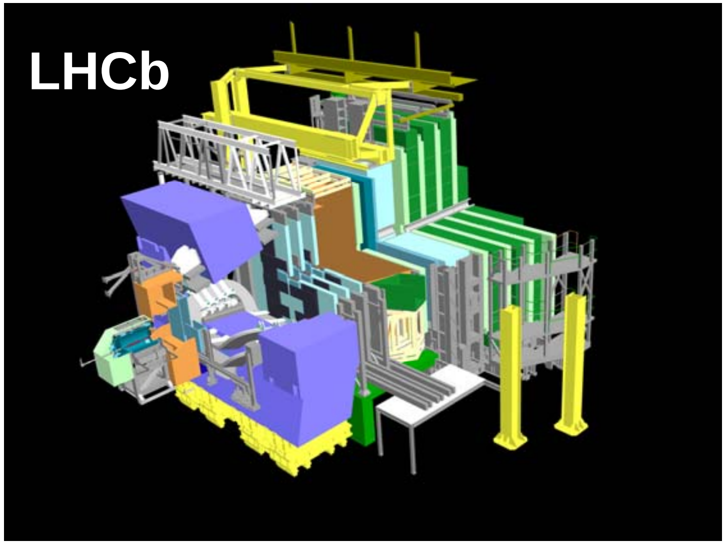
Toroid and Solenoid magnets
45m long, 25m diameter



CMS



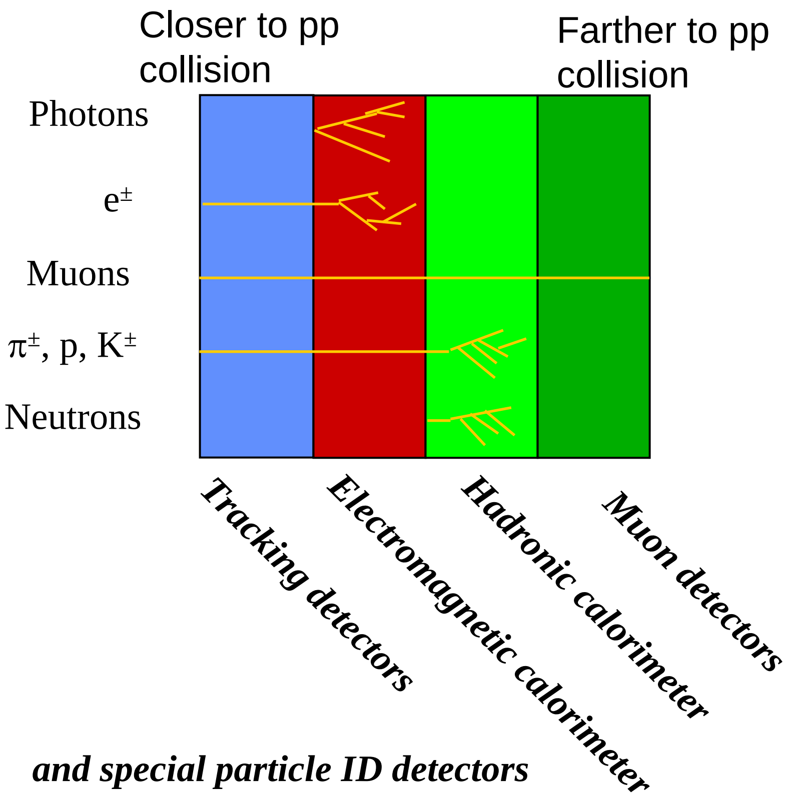
Solenoid magnet
22m long, 15m diameter



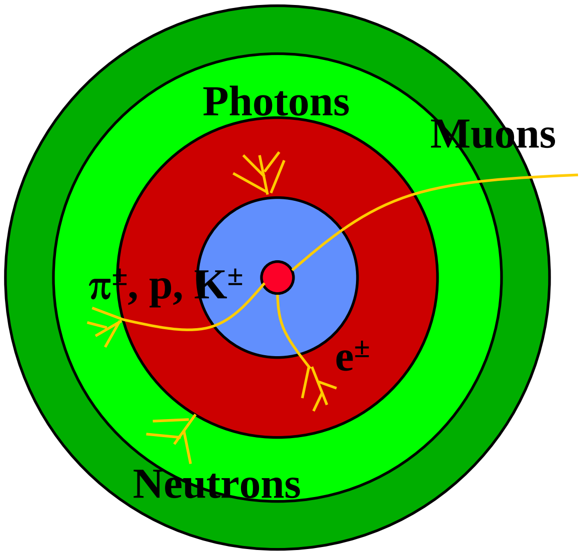
Dipole magnet
22m long, up to 10 m height

Transport through detectors

Complex detectors and a large variety of physics processes that need to be simulated



and special particle ID detectors (eg. Ring Imaging Cerenkov)



At LHC particle energies range from few MeV to TeV

What is signal for one detector can be background for another

HEP Experiments use in their framework packages developed in the physics community for transport of particles: GEANT4 and/or FLUKA

Detector simulation: transport codes

- **GEANT4 is a C++ toolkit to track particles through the detector developed in the Physics community**
  - **GEANT4 International Collaboration ~ 10 year old**
  - **used in HEP, nuclear physics, heavy ion physics, cosmic ray physics, astrophysics, space science and medical applications**
  - **GEANT4 is the successor of GEANT3, the world-standard toolkit for HEP detector simulation**
  - **http://geant4.web.cern.ch/geant4/**

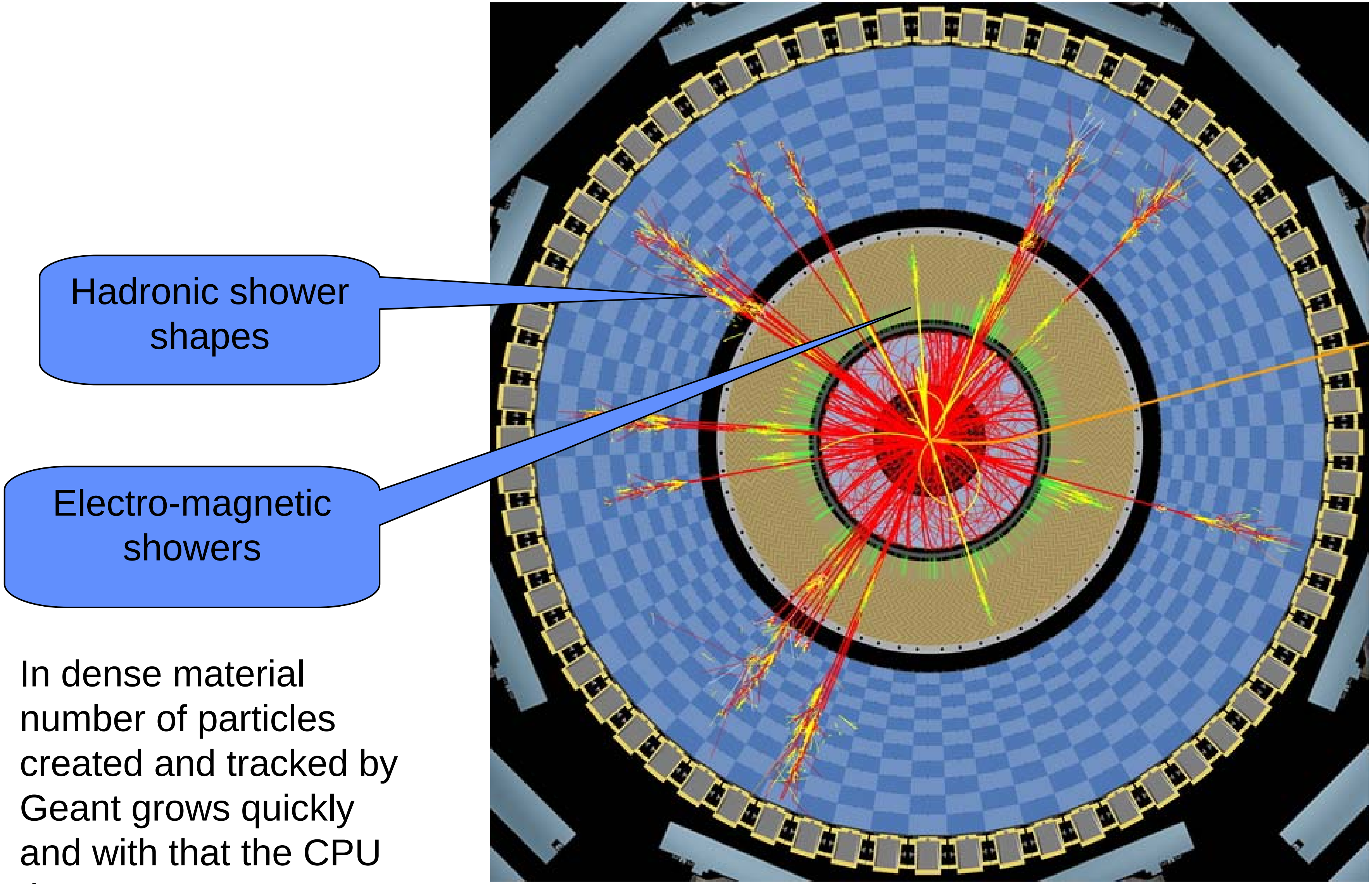
- **All LHC experiments use GEANT4 for transporting particles in the experimental setup and simulating the physics processes that can occur**
  - **Navigation in EM fields**
  - **Physics processes for a variety of particles at different energies**

- **ALICE also uses FLUKA (A.Ferrari et al)**
  - **FORTRAN-based**
  - **Couples low energy neutron transport and particle transport**
  - **Other LHC experiments use it radiation levels**

Detector simulation: Geant 4

- **GEANT4 coverage of physics comes from mixture of theory-driven, parameterized, and empirical formulae. Both cross-sections and models (final state generation) can be combined in arbitrary manners as necessary.**
  - **Standard and Low energy EM processes, Hadronic processes, Optical photon processes, Decay processes, etc.**
- **Each particle is moved in steps (few microns to cm)**
  - **For each step:**
    - **The material the particle is in is found**
    - **The energy loss and directional change due to multiple scattering is applied**
    - **The effects of electric and magnetic fields on charged particles are calculated**
    - **The particle might generate photons by bremstrahlung or Cherenkov radiation**
    - **The particle might decay in flight**
    - **The particle might collide with a nucleus and generate more particles**
    - **If the particle enters or leaves a detector it is recorded**

EM and hadronic showers in material



Atlas event from their publicity page

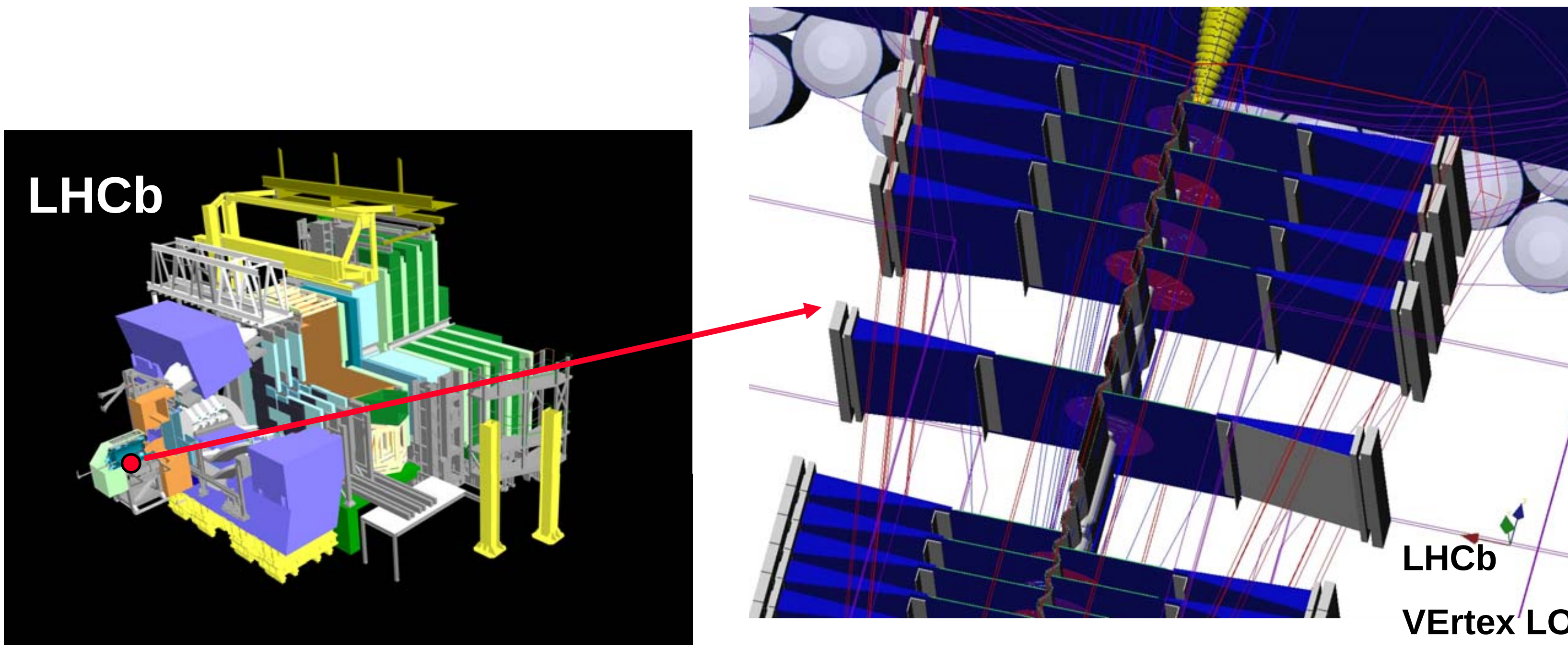
In dense material number of particles created and tracked by Geant grows quickly and with that the CPU time!

Geometry modeling

Need accurate modeling to have the most accurate results but more volumes, more memory, more CPU time

A wide variety of dimensions in an HEP experiment

Choose depending on relevance to physics study



LHCb VErtex LOcator

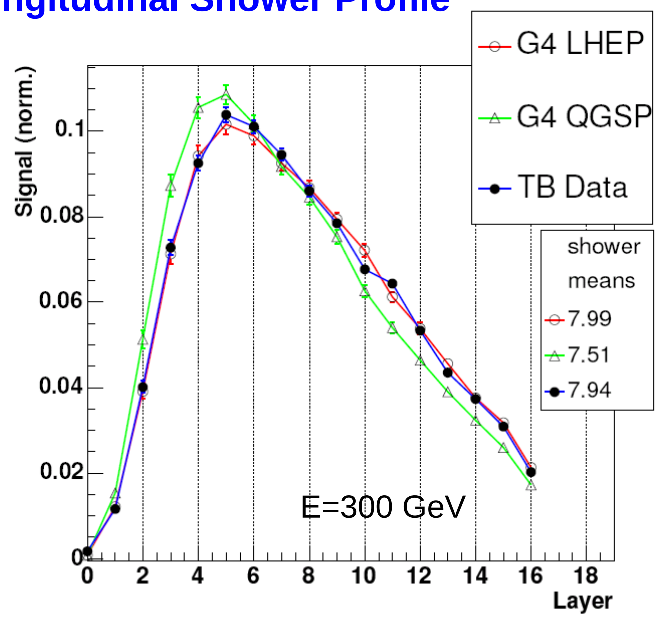
For trackers detailed description of all active and passive components; material budget

Physics processes and cuts

Need to describe physics processes as accurate as needed by resolution of detectors

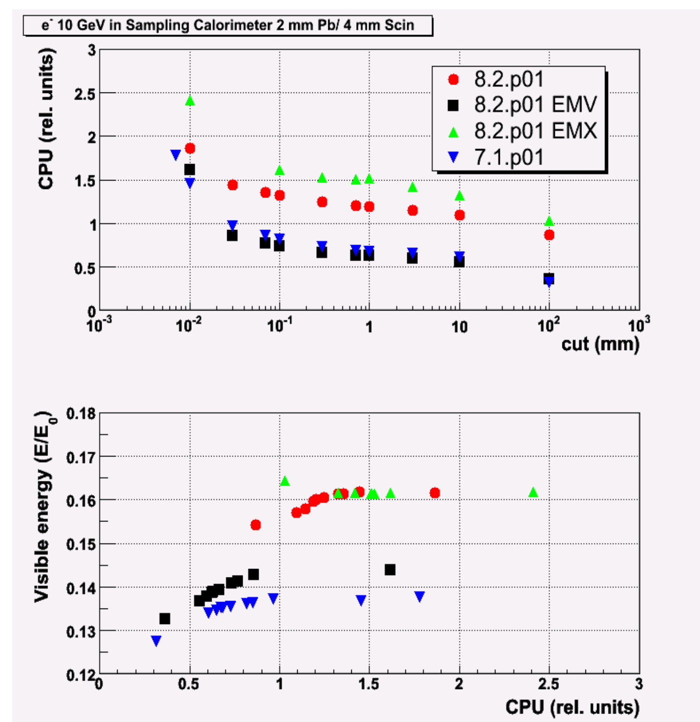
Choose appropriate physics models and set cuts: more accuracy → more CPU

CMS: Pion HCAL Longitudinal Shower Profile



In LHCb full simulation QGSP takes ~ 10% longer that LHEP

Timing Performance vs Range Cut



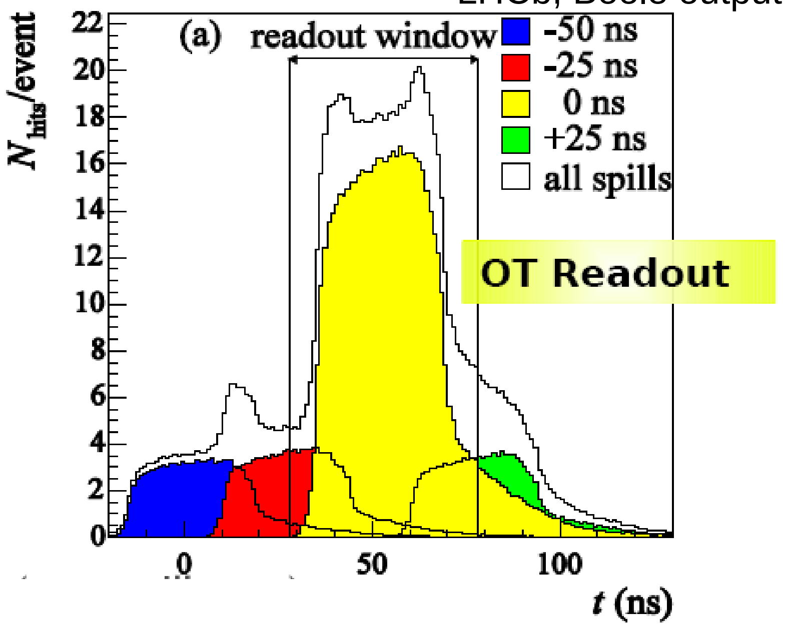
V. Ivanchenko, Geant4

Simulation of Detector response (aka digitization)

- Simulation of detector response transforming hits in sensitive detectors to produce digitized data and provide them in DAQ-like format is provided by separate application
  - Convenience
  - Flexibility
- Each detector has its own detailed detector response simulation and imperfections
  - Detections efficiencies and resolutions are adjusted according to test beam data
  - Electronic noise and cross-talk for each specific detector are added
  - Time information is correctly taken into account by all detectors

- Handling of spill-over effects (data from adjacent beam crossings)
  - A detector is sensitive to previous or subsequent bunch crossings depending on its electronics

LHCb, Boole output



In LHCb last proceessing of Boole is the L0 harwdare trigger emulation

Statistics and CPU time

- **Statistics required of the same order of real events**
  - **for LHC feasible for signal (O(10)-O(10^6)/year), even many times more.**
  - **while impossible for all pp collisions (O(10^7)-O(10^8)/sec) → months of running on the GRID to generate few seconds generic b events of LHC(b) data taking**
- **Simulations are the most CPU time consuming applications of HEP experiments**

  in LHCb the simulation takes 50 + 1 KSI2k sec/event (Gauss + Boole)
  while the reconstruction takes 2.5 KSI2k sec/event

- **Transport through the detectors is the longest**

  in LHCb: generator phase 1/100th of simulation phase

- **The CPU time depends on the complexity of the primary events**

  LHCb: minimum bias 20 KSI2k sec/event
  generic b events 50 KSI2k sec/event

  ALICE: pp collisions 39 KSI2k sec/event
  Pb-Pb collisions 7000 KSI2k sec/event with tricks to save time

- **Require a production system and GRID resources to simulate events**
  - **in LHC experiment computing model Tier2 are used to produce them and Tier1 to store the output data**
  - **ensure non-overlap of samples with control of random number generators seeds**

Simplify simulation by crossing levels

Generators → **Specific reaction** → **parameterized simulation** → **Interpreted events** → **Individual Analyses**

eg. Atlfast (ATLAS)

- **Advantages:**
  - **Fast and flexible for "what if" analysis studies**
  - **Retains flexibility to choose generators**
- **Disadvantages**
  - **Often not sufficiently realistic**
  - **Only certain information, tools available**
- **Can use similar techniques at other levels**

Conclusions

- **Events studied in High Energy Physics experiments at more and more complex. With LHC very high energy will be reached.**

- **Simulations of such events are an integral part of an experiment in all of its phases from design to physics analysis**

- **Simulations applications are a complex system of dedicated software developed by the physics community and by the experiments**

- **Considerable computing resources are (and will be) necessary: the GRID is already extensively used for this**