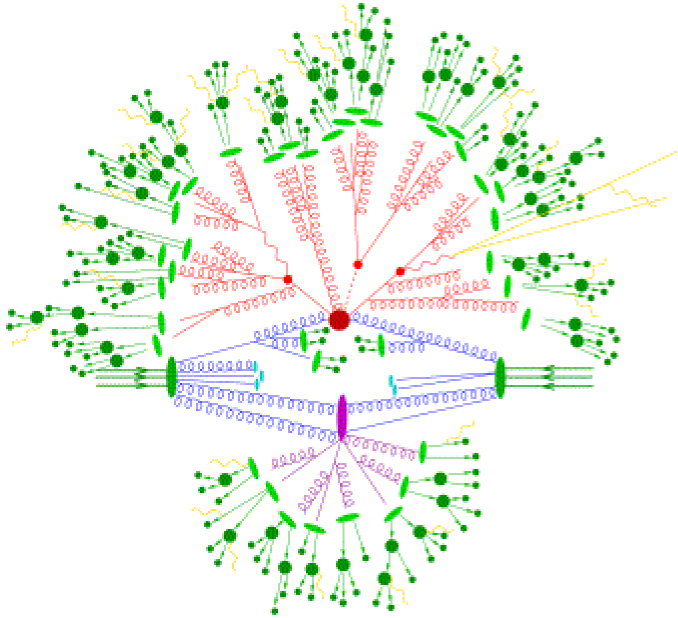
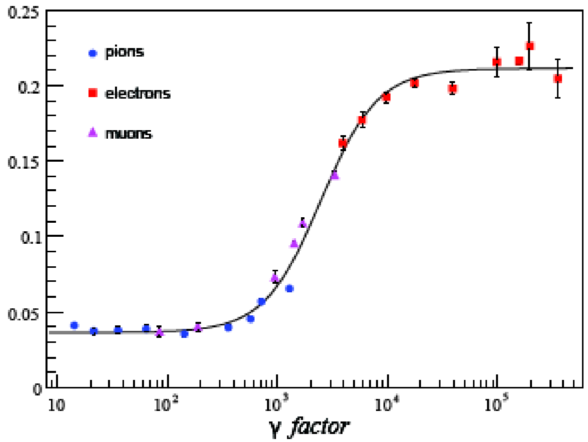
Fast Simulation

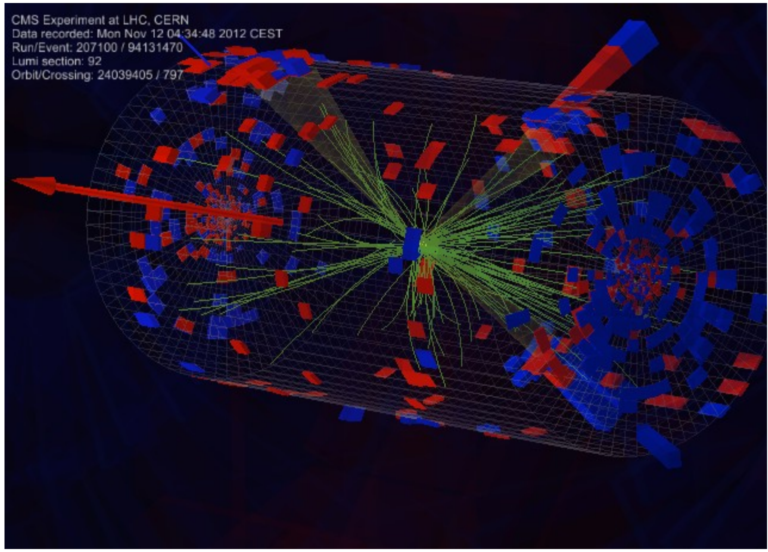
Thorsten Kuhl

- Introduction
- Concepts
- Summary

Simulation in high energy physics



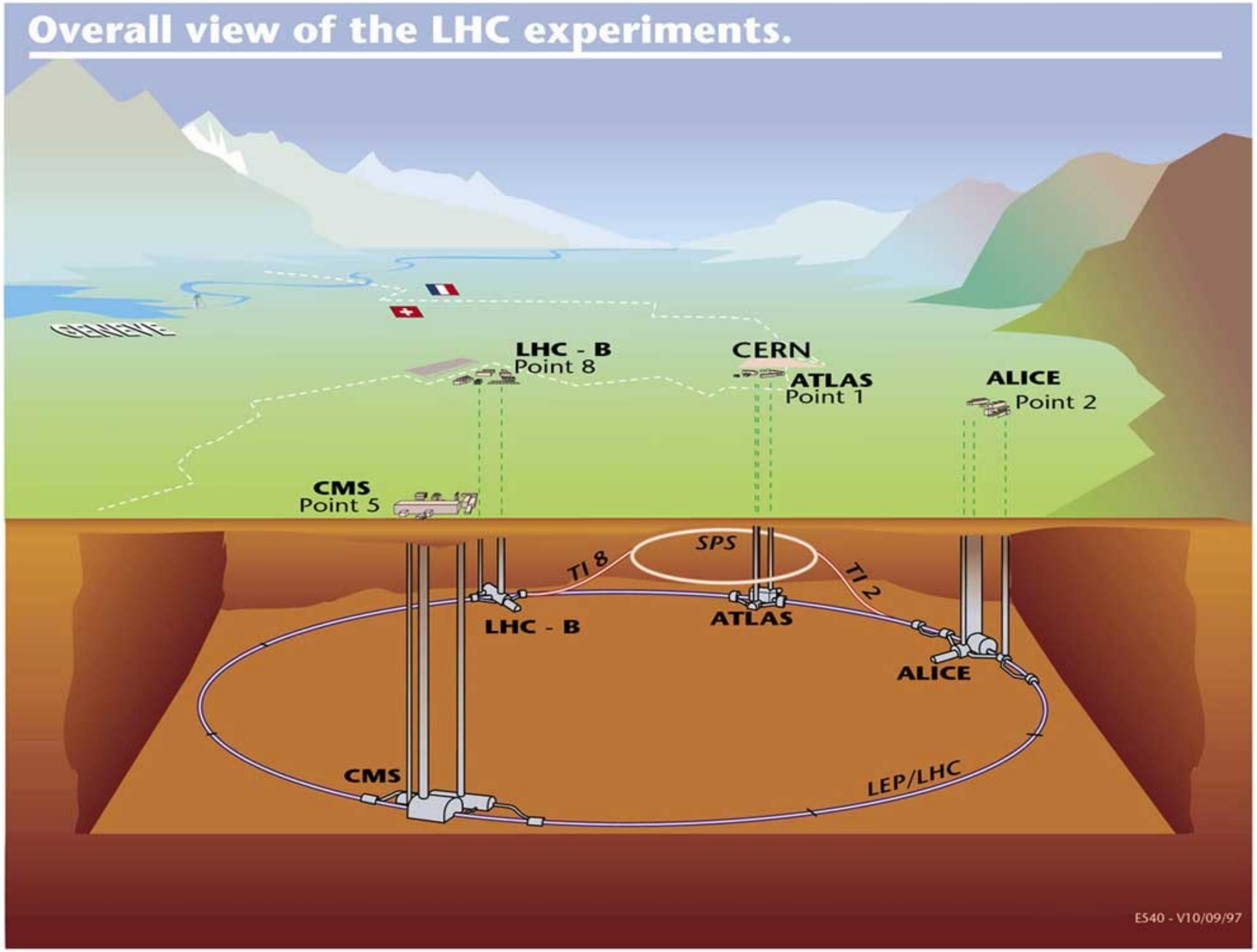




Generator + Detector response ~ recorded Data

- “Event Generation”: Simulation of the physics process
- “Event Simulation”: Response of the detector and its electronics to model the recorded data or prediction of future
- Workshop on modelling the detector response within time limitations:
  - As accurate as necessary
  - As fast as necessary

LHC



Proton proton Collider

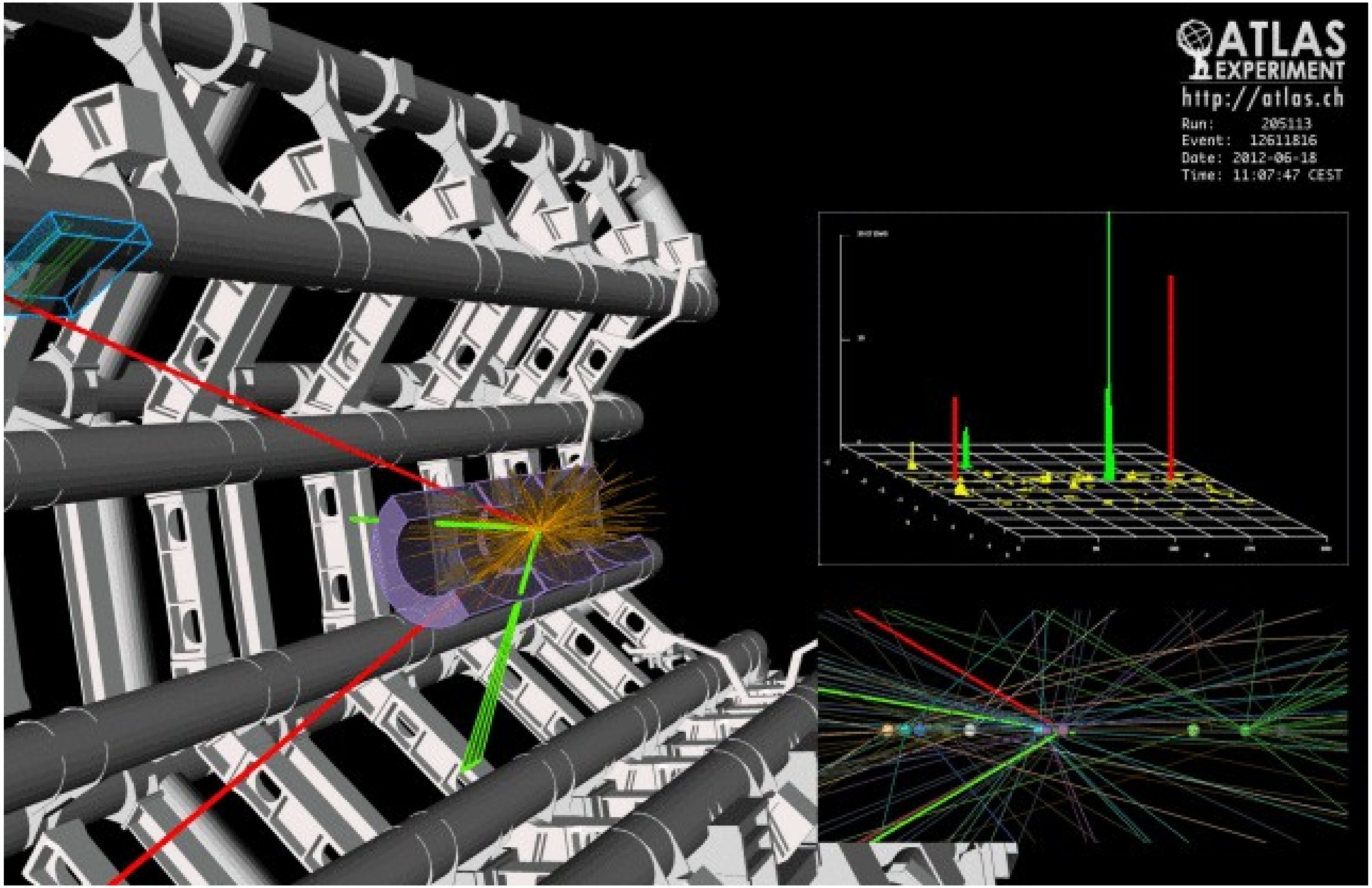
ECM: 14 GeV

2011: 6 TeV ($\sim$5 fb^{-1})
2012: 7 TeV $\sim$15 fb^{-1}

Discovery of the Higgs Boson and new physics beyond the SM

Simulation for experimental data unfolding and comparison with theory
Upgrade simulation for future improvements of detector and machine

Higgs-Decay



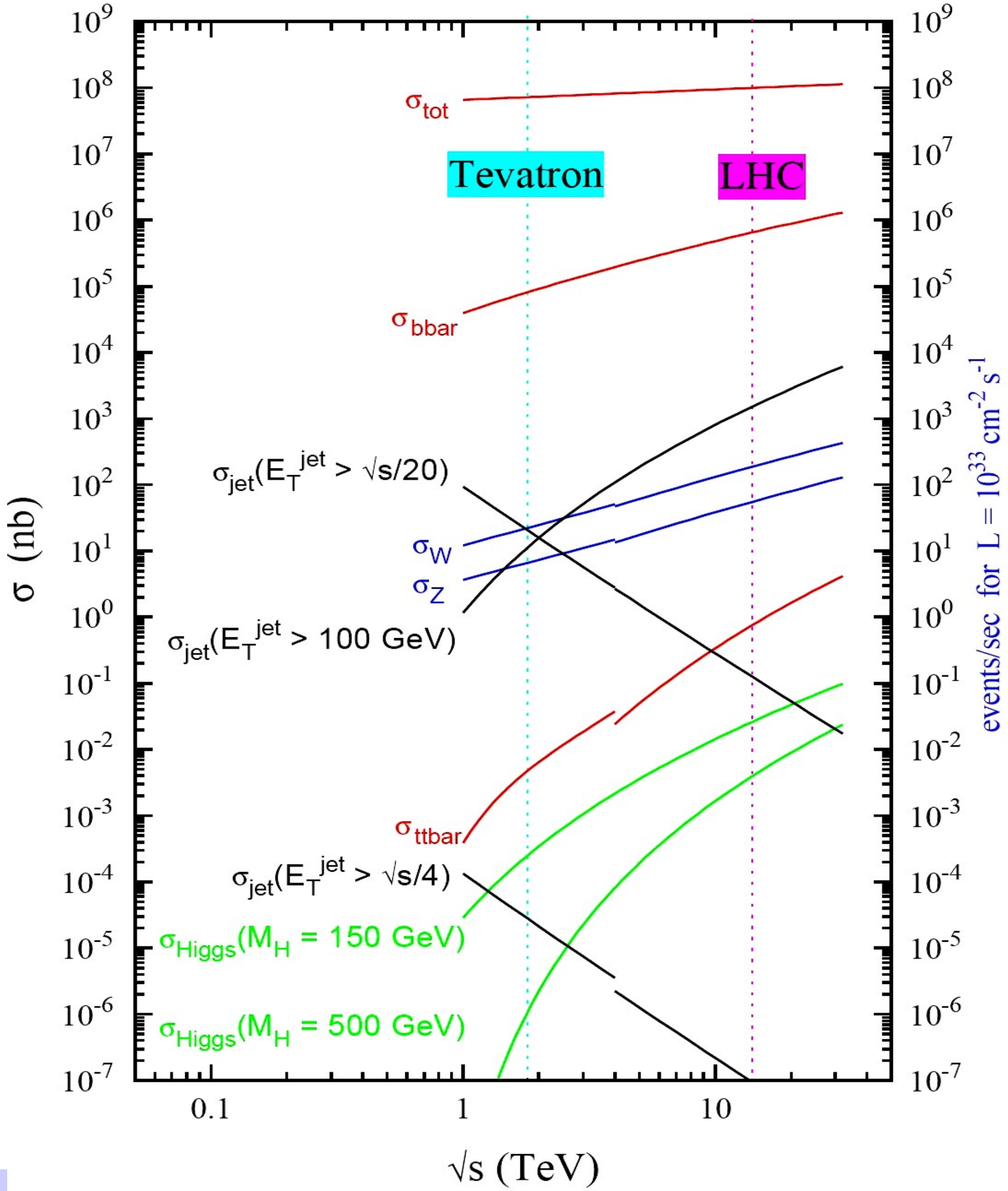
$H \rightarrow Z^* + Z$

$Z \rightarrow e^- + e^+$

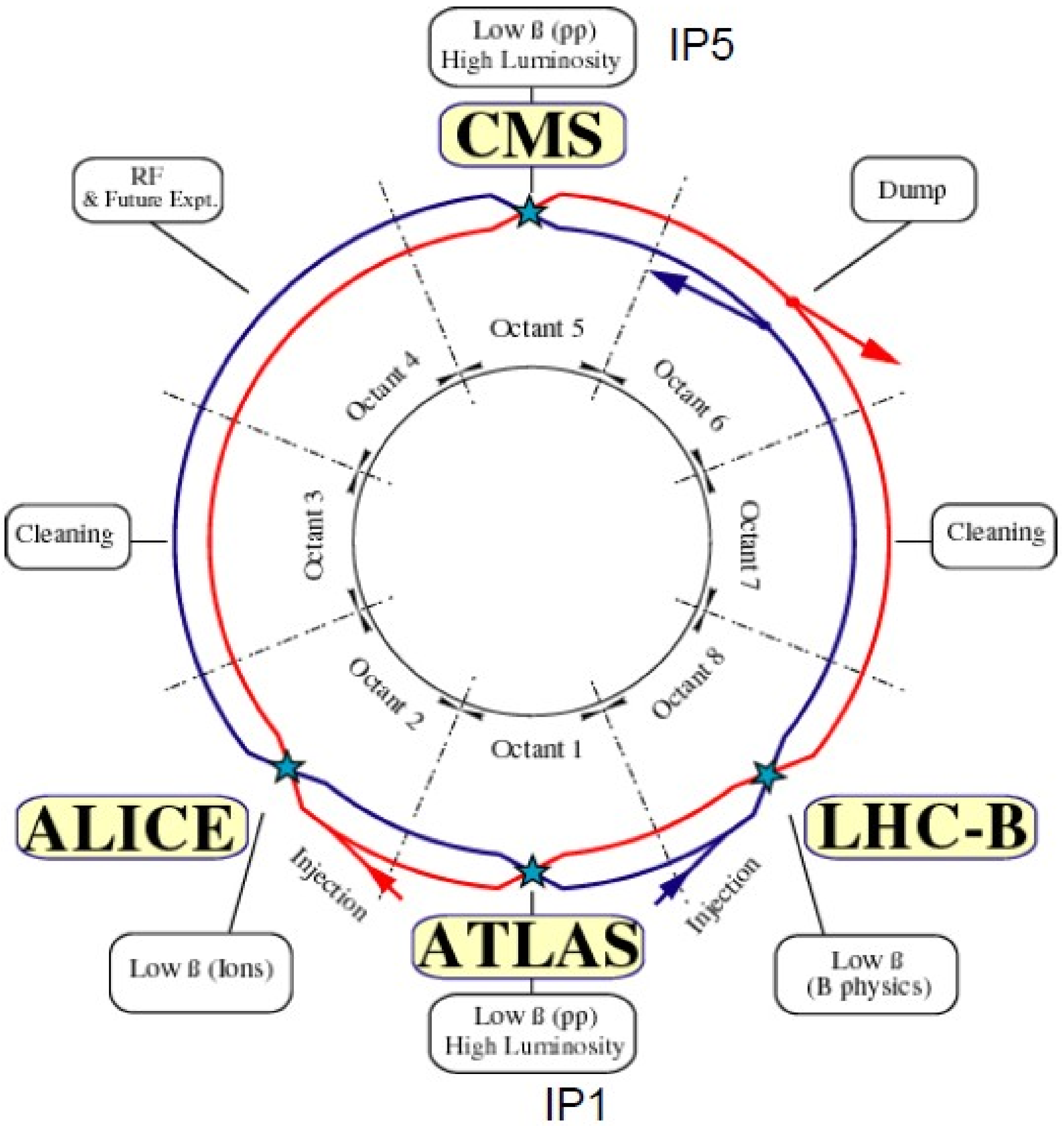
$Z^* \rightarrow \mu^- + \mu^+$

LHC physics processes

- Higgs is a rare process at the LHC:
  - Every 10 seconds a event, but accessible channels are suppressed by branching ratio
  - W/Z/top suppressed by factor of 100000 and more
- Similar situation in other searches
- Need high background suppression factors to enhance a sub-sample with these events → tails modelling in the background samples are important, need huge statistics
- Not that easy to access channels (H → WW,bb, often uses multivariate method to extract signal)
- SUSY scans



Machine parameters



LHC design parameters

c.m. energy = 14 TeV
luminosity $=10^{34}$ $cm^{-2}s^{-1}$

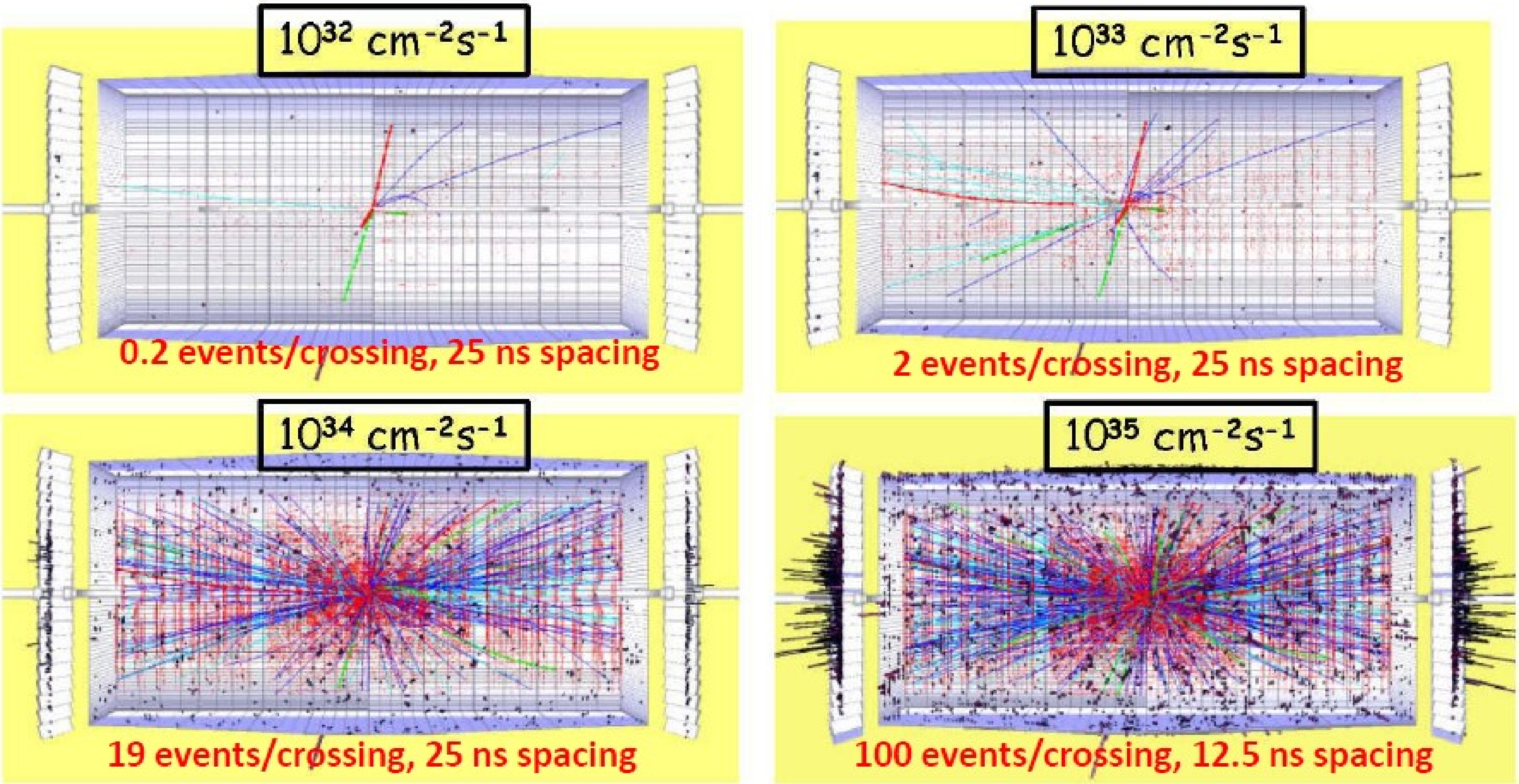
→ ~20 events/bx

1.15×10^{11} p/bunch
2808 bunches/beam

$\gamma\varepsilon = 3.75$ μm
$\beta^* = 0.55$ m
$\theta_c = 285$ μrad
$\sigma_z = 7.55$ cm
$\sigma^* = 16.6$μm (IP1 & 5)

LHC physics processes

event pile up in detector



$p_t > 1$ GeV/c cut, i.e. all soft tracks removed

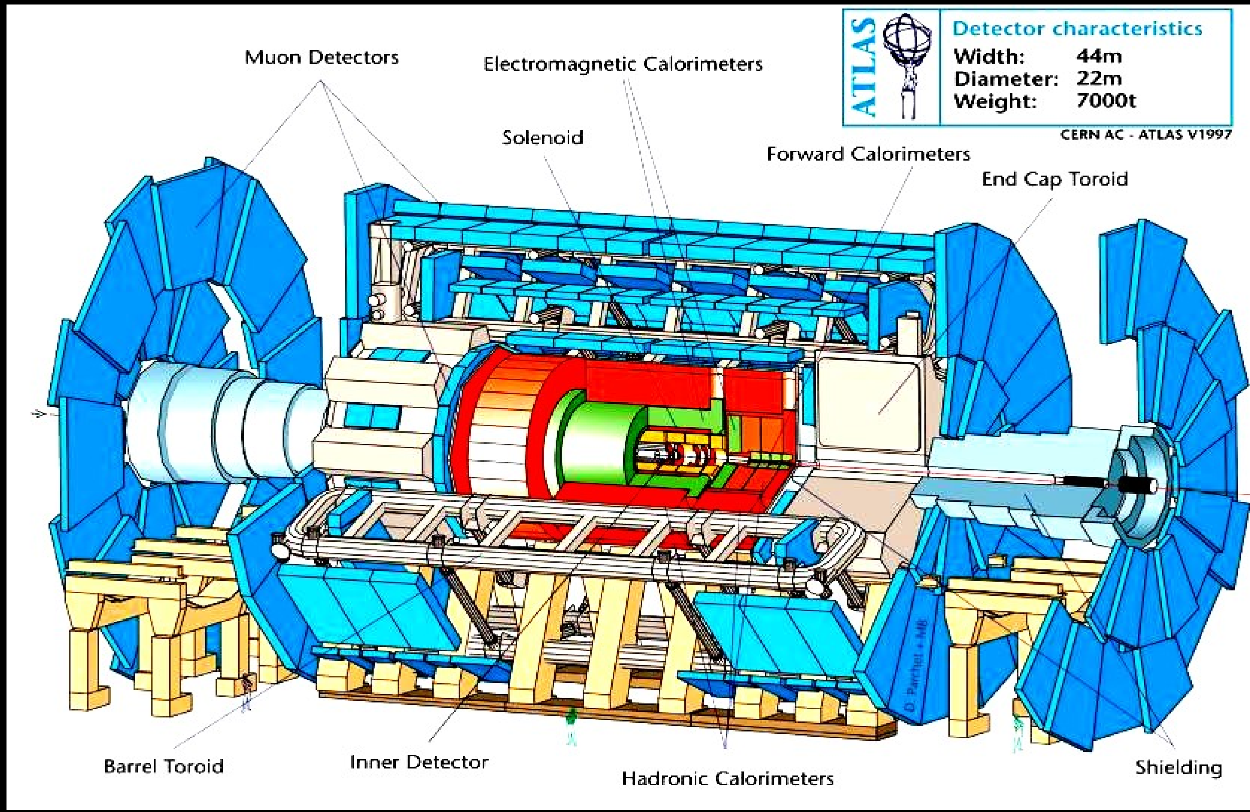
historical simulation I. Osborne

This cut removes already 90% of pile up

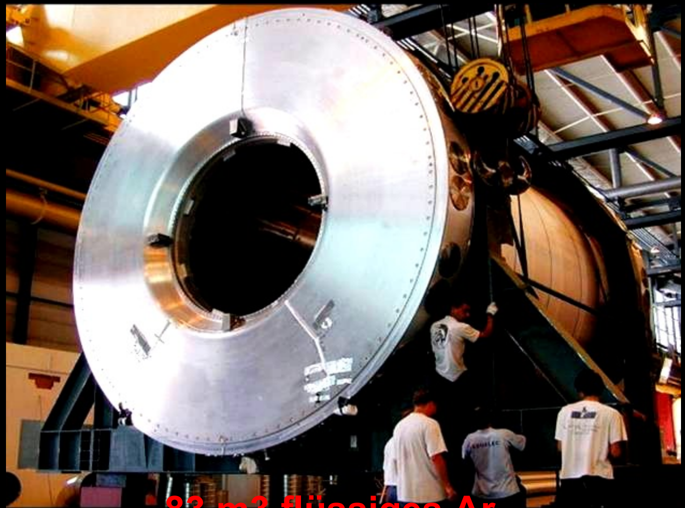
Cathedrals of Physics

ATLAS

Forward-Muon-Chambers

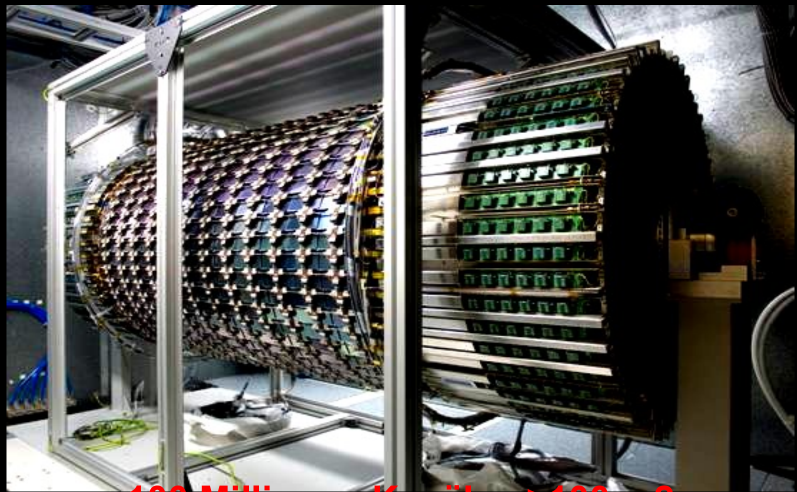


Liquid-Argon-Calorimeter



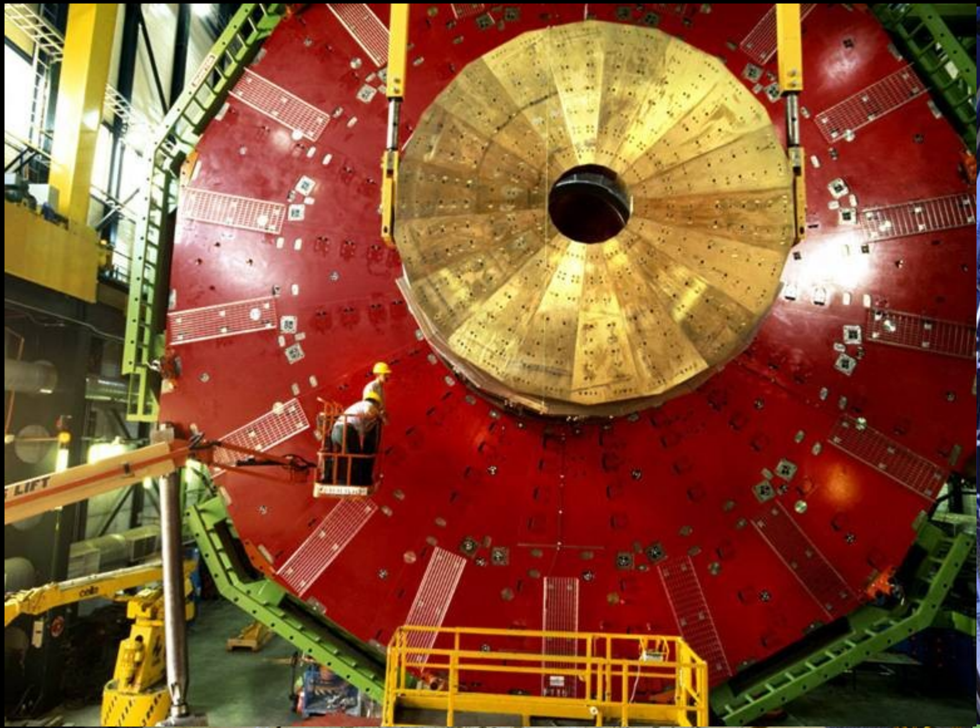
83 m3 flüssiges Ar

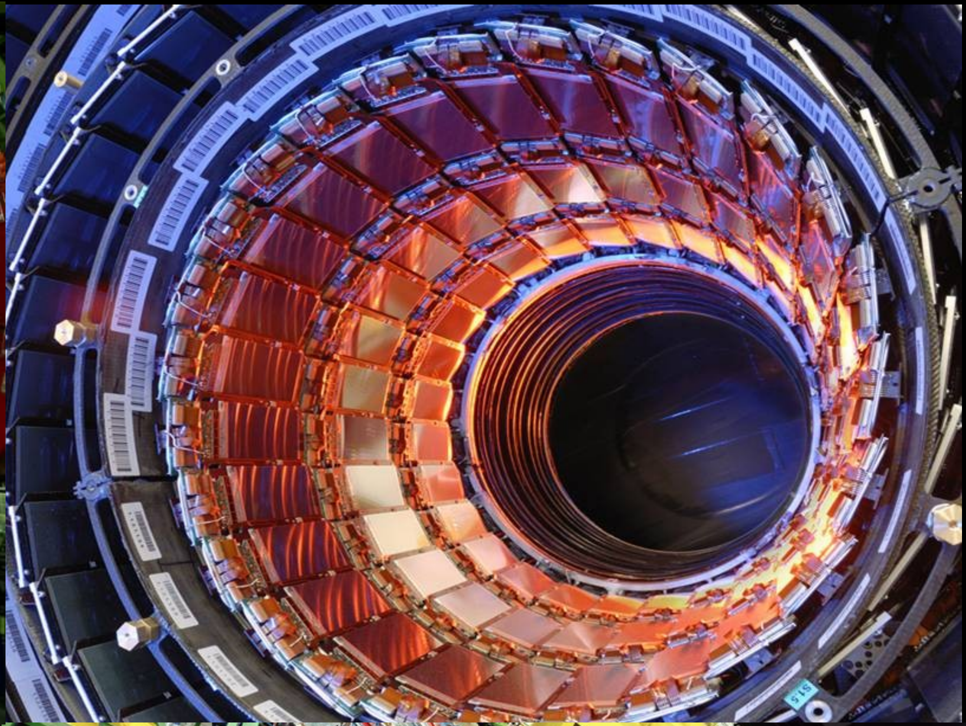
Silicon-Tracker

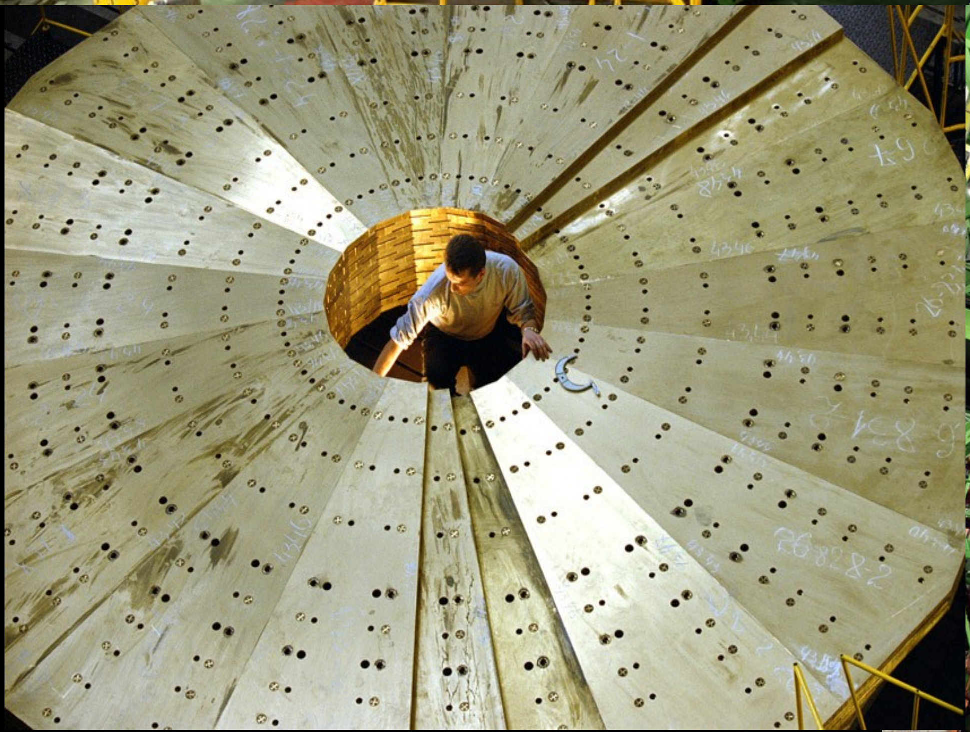


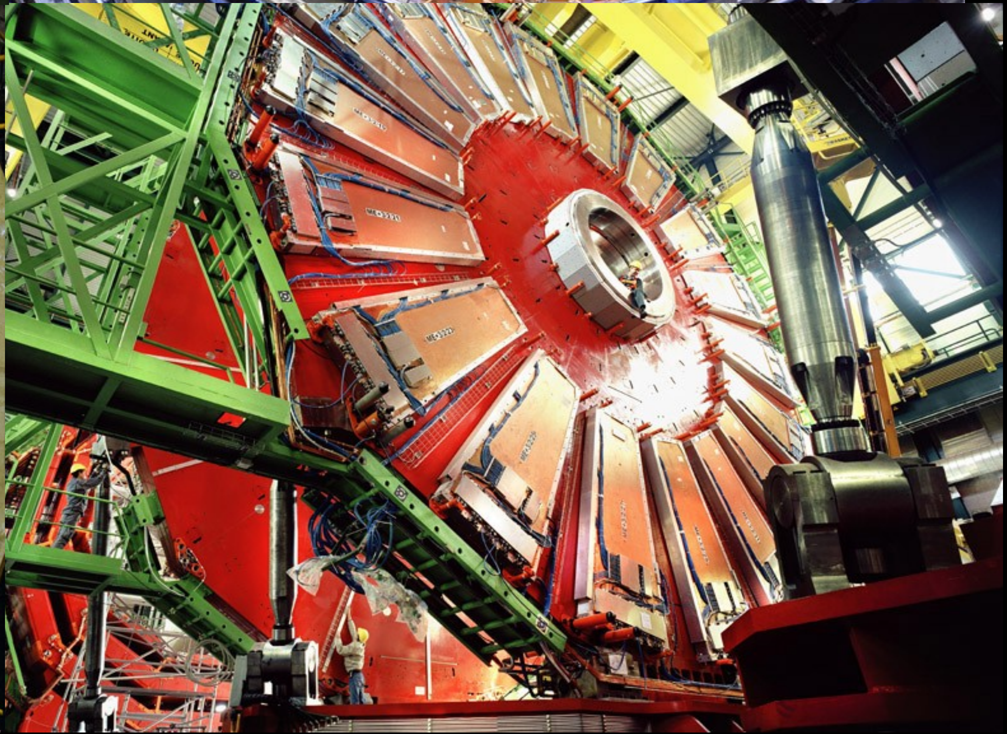
~100 Millionen Kanäle >100 m2


CMS
ME&RE+1/3/02
ME&RE+1/3/01
ME&RE+1/3/36









Detector Parameters

	ATLAS	CMS
Diameter (m) x length (m)	22 x 46	15 x 20
Magnetic field for tracking	2	4
Weight (tons)	7000	12500
σ/p_T tracker	$5 \cdot 10^{-4}\, p_T \oplus 0.001$	$1.5 \cdot 10^{-4}\, p_T \oplus 0.005$
σ/E EM cal	$10\%/\sqrt{E(GeV)} \oplus 0.7\%$	$3\%/\sqrt{E(GeV)} \oplus 0.5\%$
σ/E HAD cal	$50\%/\sqrt{E(GeV)} \oplus 3\%$	$100\%/\sqrt{E(GeV)} \oplus 5\%$
σ/p_T Muon tracker (@ 1 TeV/c)	10%	5% (with tracker)
Lepton Energy scale	0.02%	0.05%
Jet Energy Scale	1%	3%
Absolute luminosity	< 5%	< 5%
Total cost (MSfr)	550	550

- Using different and complementary technologies, the two large LHC experiments arrive at similar overall performances.
- The higher magnetic field in CMS has advantages (better p_T resolution) and disadvantages (lower tracking efficiency).
- The CMS crystal calorimeter has a superior energy resolution but no longitudinal sampling.
- The MUON acceptance is larger in ATLAS, but in CMS it has a simpler geometry and a uniform magnetic field.

Summary: LHC physics requirements

- Interesting SM processes are highly suppresses vs QCD:
  - W/Z: ~10^5
  - Top: 10^8
- New physics is rare:
  - SUSY and (SM-like) Higgs production even more suppressed

    → high background suppression necessary, simulation of tails

→ Machine with high luminosity → a lot of events, pile up

  - 20 events/bx every 25ns → thousands of tracks/clusters in the detector

→ High granular detector:

  - a lot of channels
  - high time resolution (only 25ns between bunch crossings)
  - High energy/momentum resolution
  - Physics fakes
- Highest precision but fast?

Simulation schema

Generator event: Hard interaction, Fragmentation and UE,
Short lifetime decays

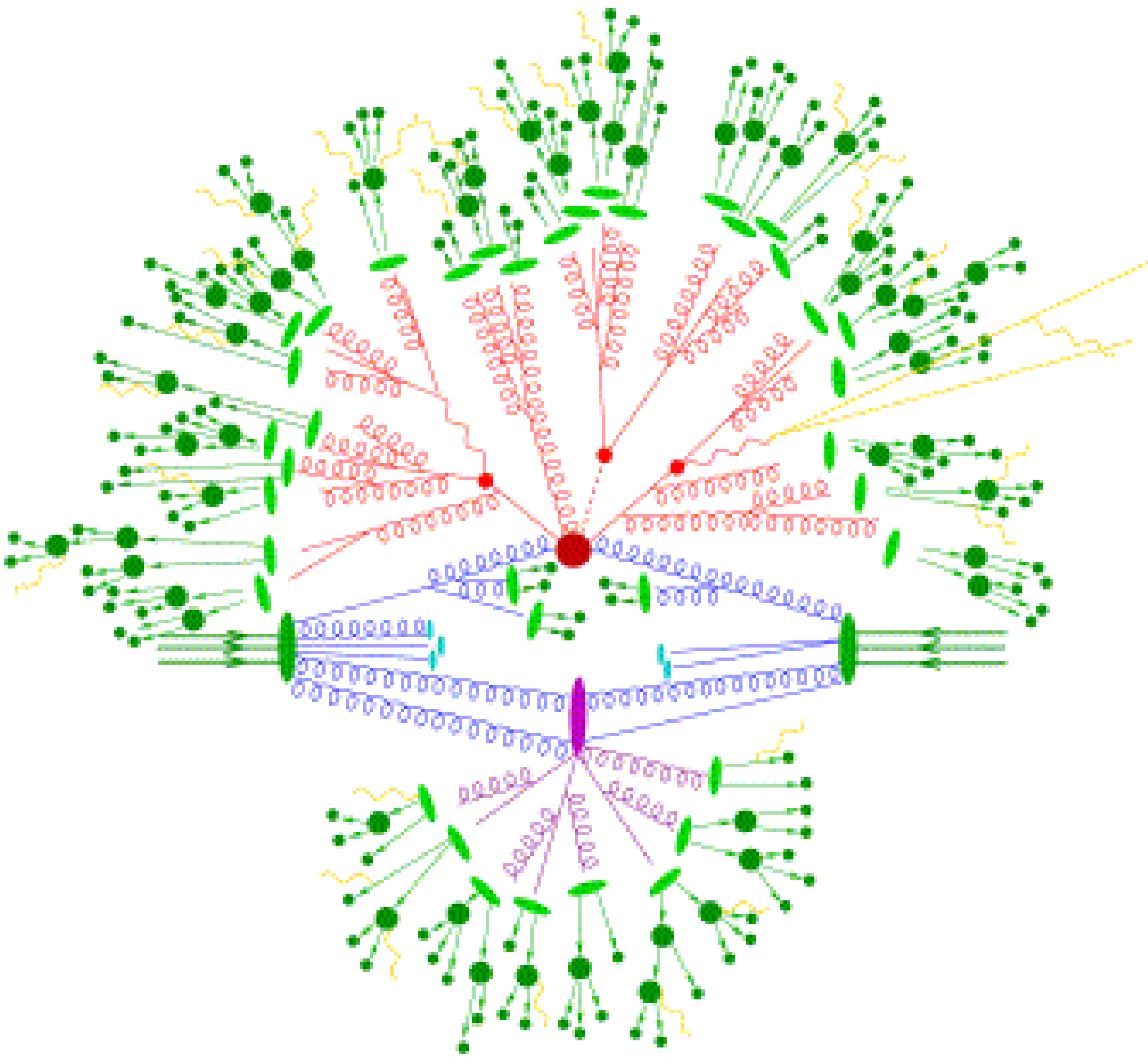
?

High level analysis based on objects (4-vectors)

First step of simulation: Generator

Not part of the workshop but for completeness

- Partons from Protons: PDF and underlying event
- Hard interaction/matrix element
- Parton shower (radiation of gluons/quarks), matching to ME
- Fragmentation and decays
- Done in a complete framework (Sherpa, Vanilla Pythia) or in sequence of packages with defined interface: MC@NLO/Powheg/Alpgen/Mathgraph+Pythia6/8 /Herwig (++) etc.
- Some detector simulations have Generator included, other use standard inputs (more flexible if you use already ~a dozen different generators for ttbar only including new physics in ttbar)



Truth Level analyses

Generator event: Hard interaction, Fragmentation and UE, Short lifetime decays

- 4-vector smearing (a lot of generators using matching, no good description of hadronic environment (not that good)

- Particle truth analyis: Applying jet finding after fragmentation/hadronisation after defining leptons → “rivet analysis”; Very usable for theorist to compare generator with unfolded data

High level analysis

Simulation

Generator event: Hard interaction, Fragmentation and UE, Short lifetime decays

Full simulation: detailed complete detector simulation for all generator objects

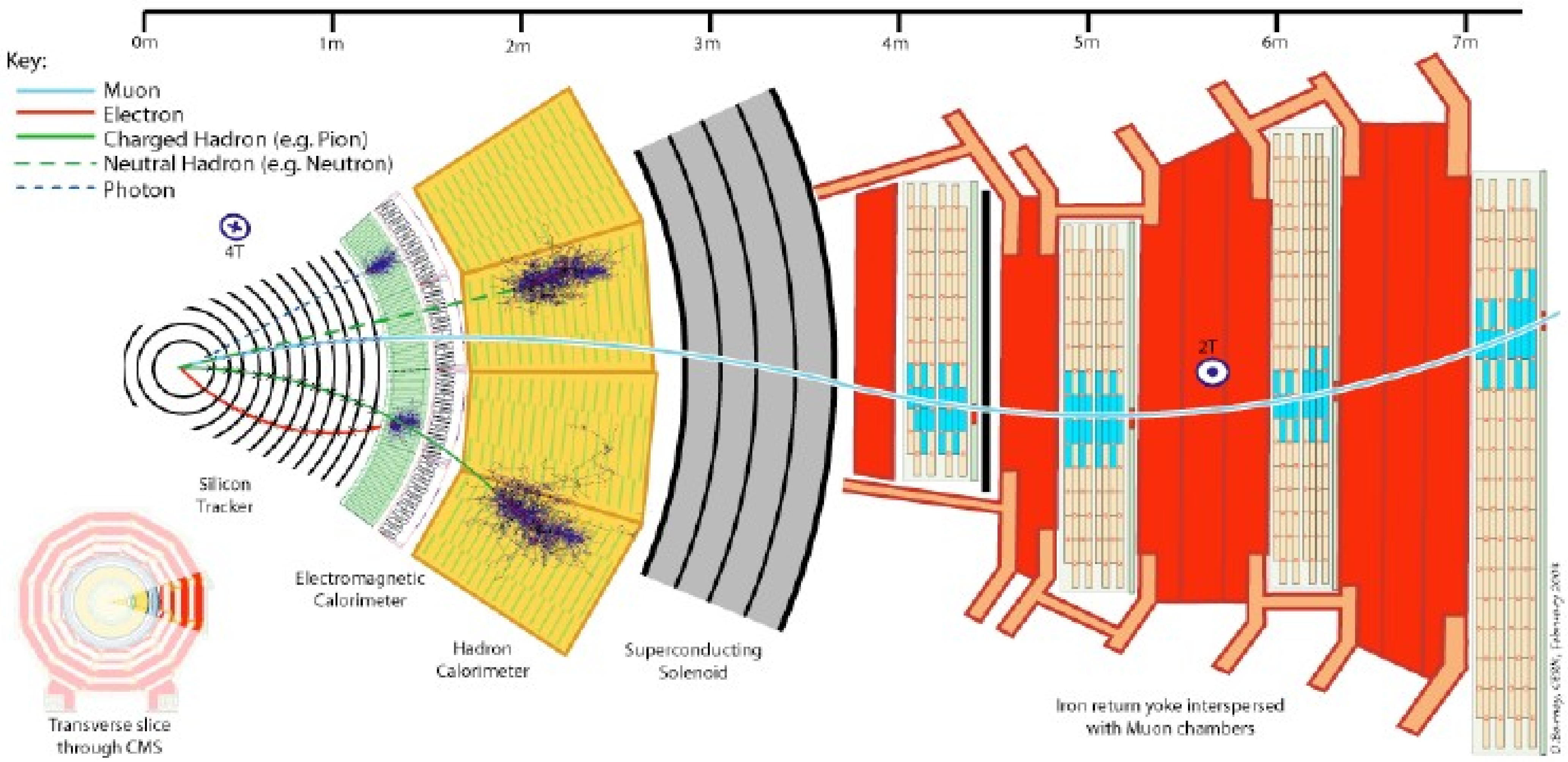
Smearing of particle truth jets or jet constituents and leptons

High level analysis

Full detector simulation

- Detailed detector description:
  - Data base including every object bigger then O(mm)
  - Sensitive and dead material
  - Possibility to enable misalignment
- Detailed simulation of interaction of particles with detector material (Geant 4) by stepping trough
  - Long physics lists including most processes, Energy loss in material, Multiple scattering, bremsstrahlung, Radiation, Hadronic interactions
  - Calorimeter (shower) simulation: tracking every particle in a shower over a low threshold
- Result: energy deposit in sensitive volumes

Detector Parameters



Tracking Calorimeters Coil Muon Chambers

Detector electronics response/noise/pile up

- Translation of the G4 energy deposit into electronic signals from the detectors and overlay of pile up noise etc
- We cannot see trajectory of every single particle in a shower...
  - Dead material
  - Granularity
  - Noise/sensitive threshold etc.
- Overlay of hard interaction with other primary interactions
  - Pile up due to multiple interactions in detector volume
  - Pile up from limited time resolution of the electronics
- Digitalisation:
  - Response of electronics for the energy deposit in sensitive volume with more/less fancy models
  - Electronic noise/inefficiency/timewalk
  - Model for time resolution/pile up in detector

Object reconstruction

- Run reconstruction the same algorithms like for real data:
  - Clustering of raw data to hits and clusters
  - Track/e-m. Cluster reconstruction
- Advantage of Full simulation:
  - Based only on knowledge of fundamental processes
  - Reconstruction can be done in the same way like for data
  - Maybe best you can do to predict future experiments or learn about the physics of your detector
- Disadvantage:
  - Slow (Atlas: O(5-10) minutes for big events(ttbar)
  - Not perfect: Still often need data driven adjustments

Full simulation

- Advantage of Full simulation:
  - Based only on knowledge of fundamental processes
  - Reconstruction will be done in the same way like for data
  - Maybe best you can learn about the physics of your detector
- Disadvantage:
  - Slow: (Atlas: O(5-10) minutes for big events(ttbar)
    - In a lot of physics case you need huge number of simulated events to populate physics tails/fake distributions
    - Already very difficult to match the data statistics at the LHC
  - Not perfect: Still often need data driven adjustments

Simulation

Generator event: Hard interaction, Fragmentation and UE, Short lifetime decays

Geant 4 using detailed detector geometry

Digitalisation: Electronic response of detector channels

Reconstruction of basic objects (tracks, leptons, jets)

Smearing of truths jets and leptons or all truth particles to simulate detector response

High level analysis

Very fast simulation

- Smearing of objects (electrons, muons, jets or jet constituents = tracks/cluster) using simple parametrisation of the detector response
- Advantage:
  - Very fast
- Disadvantage:
  - Simple gauss does not provide tails … need better parametrisation for rare processes
  - Inefficiencies (can be parametrized, too)
  - Overlapping objects, fakes, hit sharing, overlapping events?
- Better:
  - Parametrisation based on Fullsim/testbeam data, allow more tails in function
  - Calculation of full covariance matrix using simply detector geometry model

Very fast simulation

- Why does a smearing work at all → simple simulation for multiple scattering (on Andys Mac)
- (works perfect if nature would have only one physics process for tracks)

Simulation

Generator event: Hard interaction, Fragmentation and UE,
Short lifetime decays

Geant 4 using detailed detector geometry

Digitalisation:
Electronic response of detector channels

Reconstruction of basic objects (tracks, leptons, jets)

Very Fastsim:

Parametric smearing of basic generator objects

High level analysis

Simulation

Generator event: Hard interaction, Fragmentation and UE, Short lifetime decays

Geant 4 using detailed detector geometry

Digitalisation: Electronic response of detector channels

Reconstruction of basic objects (tracks, leptons, jets)

Mixed simulation: Geant4, hit/shower based fastsim, data or data driven (tau embedding (Atlas), Real data Pile up/noise (D0)

Very Fastsim:

Parametric smearing of basic generator objects

High level analysis

Workshop is about different mixture/compromises/solutions

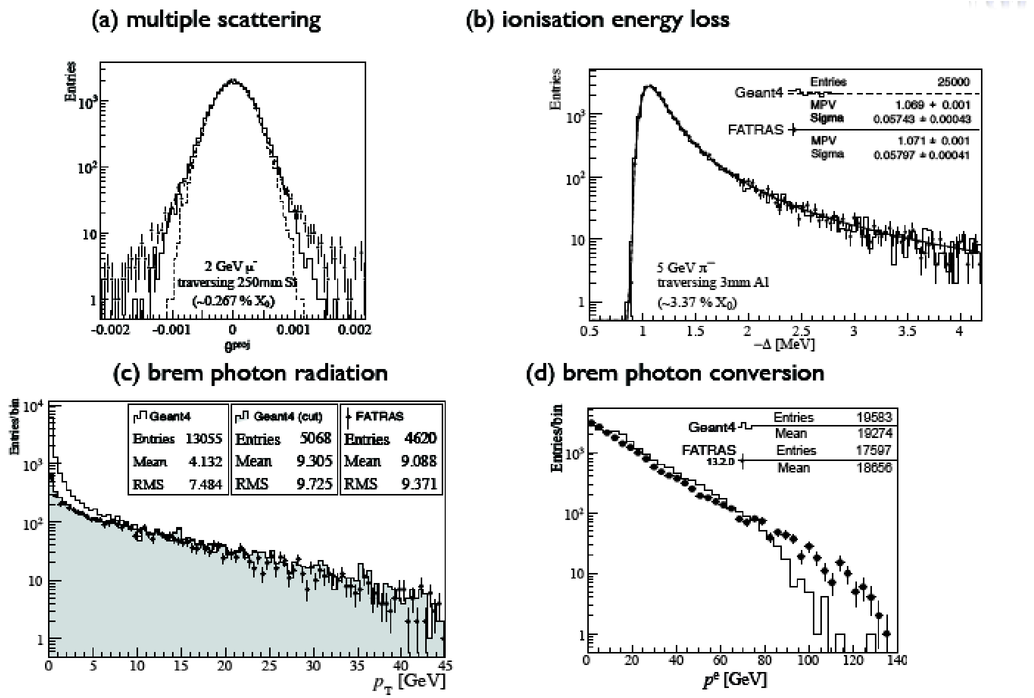
Different Flavors of Atlas Simulation

	Atlfast I	Atlfast II		Full
		Atlfast IIF	Atlfast II	
ID	parameterised track perigee's	FatrasID	full simulation digitisation reconstruction	full simulation digitisation reconstruction
Calo	parameterised clusters	FastCaloSim	FastCaloSim muons: full	full/frozen G4 digitisation reconstruction
MS	parameterised track perigee's	Atlfast I FatrasMS (exp.)	full simulation digitisation reconstruction	full simulation digitisation reconstruction
rel. gain factor timing	~ 1000	~ 100	~ 10	~ 1

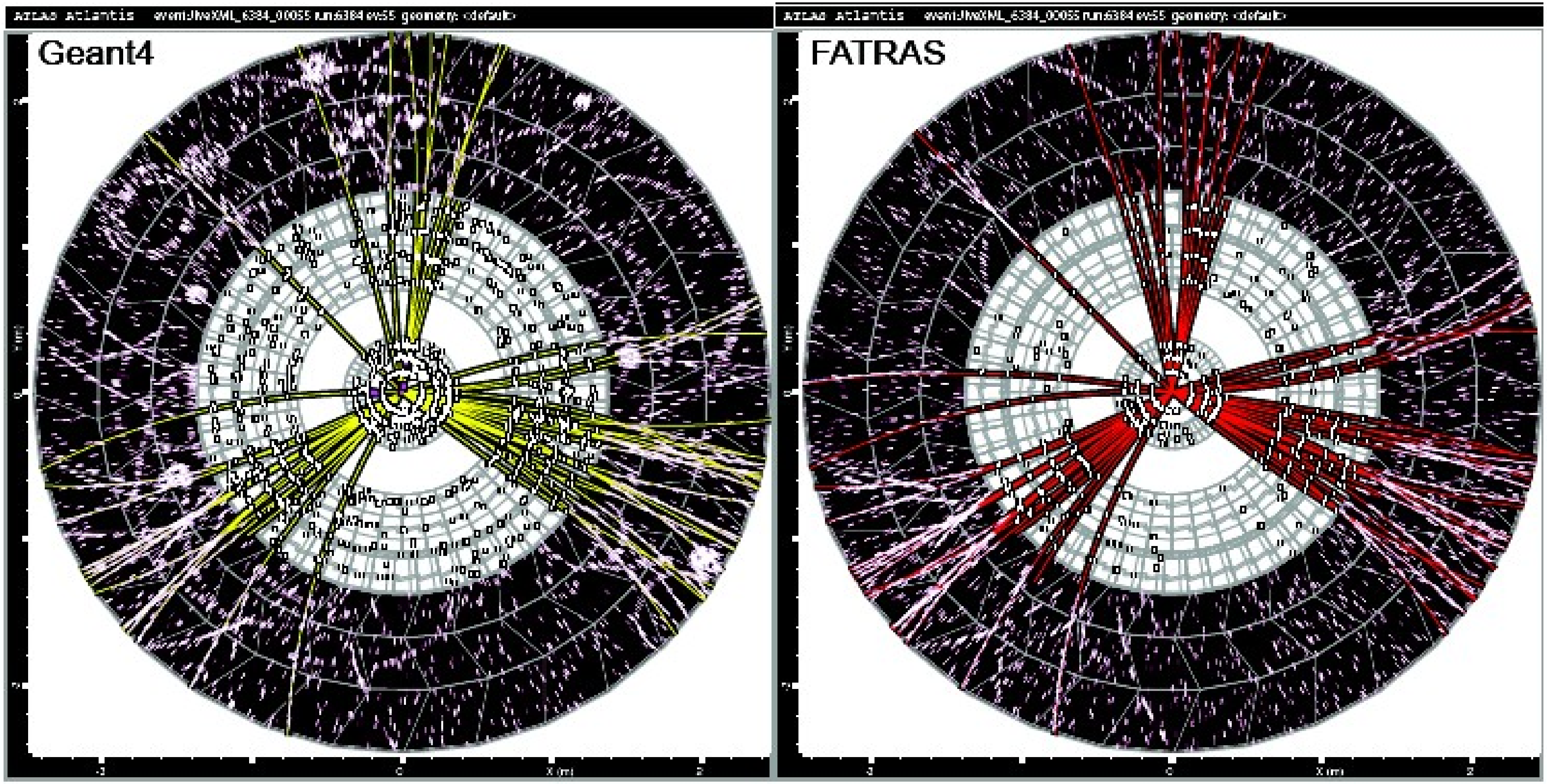
- What we call full simulation use frozen shower for forward calorimeter
- Our “Atlfast II” simulation has full simulated tracks (its close to CMS full)

Tracking concepts

- Tracking concepts:
  - simple track parametrisation (based on any input, knowledge)
  - Inverse Kalman filter: covariance matrix smearing based on geometry, X0 distribution
  - Very fast simulation of all physics processes using simplified G4 similar models for material interactions, digitalisation and clustering



And it looks not that different



~ 100 sec

~ 1 sec

Concepts (II)

- Calorimeter concepts:
  - parametric smearing, clustering
  - frozen shower
- Frameworks:
  - Switching between the different approaches, optimize for physics case
  - Interfaces to the outer world
  - Flexible detector geometry: allow to test your preferred detector concept
- Upgrade: how to simulate the detector/physics of the future
- Event overlay: pile up/noise from simulation or real data, tau embedding
  - Generator still do not simulate pile up perfect, new models from astroparticle physics seem to be better, but not that perfect like data events
- Free programs of the market (Delphes, SGV) vs embedded Atlas/CMS fast simulation
  - Fast estimate for what we get in future
  - Chance for theory peoples to get a reasonable estimate for predictions

Summary

- LHC: very rare signal processes
- Huge number of background processes (W/Z+jets ~270M events last year):
    - Not easy to provide statistics with Full Geant4 simulation
    - Generator cuts are not always good/possible
- LHC experiments use already a mixture of fast and full simulation:
    - The simulation has to speed up with data
    - Fast simulation gets more and more important
    - Need concepts to switch simulation depending on physics signal
        - Specially needed for simulation of fakes
- Upgrade/Future experiments:
    - Possibility to switch geometries for upgrade simulation
    - Fast simulation a good tool to estimate a real physics performance on benchmark processes for different scenarios
- Free fast simulations: possibility for theory peoples to get a good guess