Concepts and Plans towards fast large scale Monte Carlo production for the ATLAS Experiment

ACAT2013, Beijing

Elmar Ritsch (Univ. Innsbruck, CERN)
on behalf of the ATLAS collaboration

May 18, 2013

Overview

- **Current Status**

  from event generation to ROOT histograms

- **Fast Detector Simulation**

  fast simulation engines, Integrated Simulation Framework

- **Fast Digitization**

  from simulation to reconstruction input

- **Fast Reconstruction**

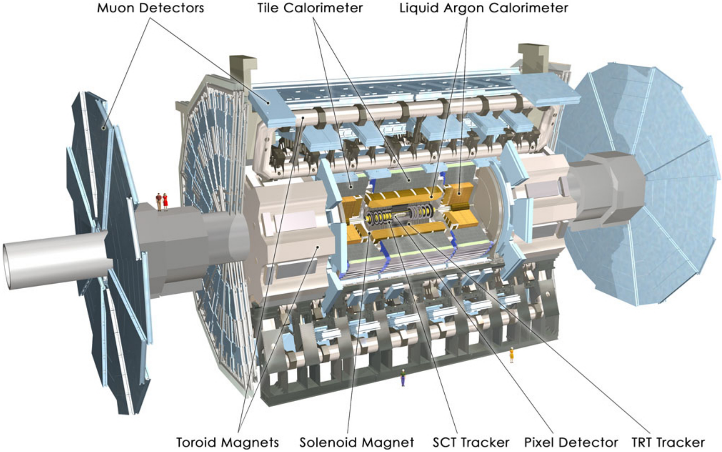
  truth based reconstruction

- **Fast Monte Carlo Production Chain**

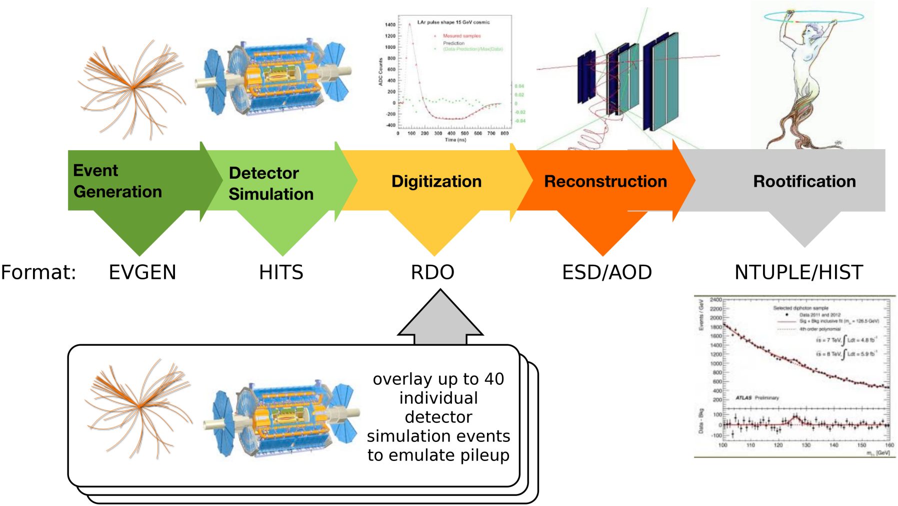
  combining the fast processing efforts

ATLAS Experiment at the LHC

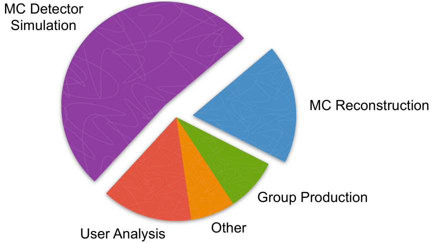
- **Inner Detector**: tracker with solenoid magnetic field
- **Calorimeter**: electromagnetic and hadronic
- **Muon Spectrometer**: muon tracker with toroid magnetic field



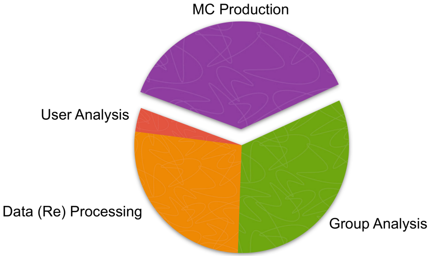
ATLAS Monte Carlo Production – Current Status



ATLAS grid CPU utilization:



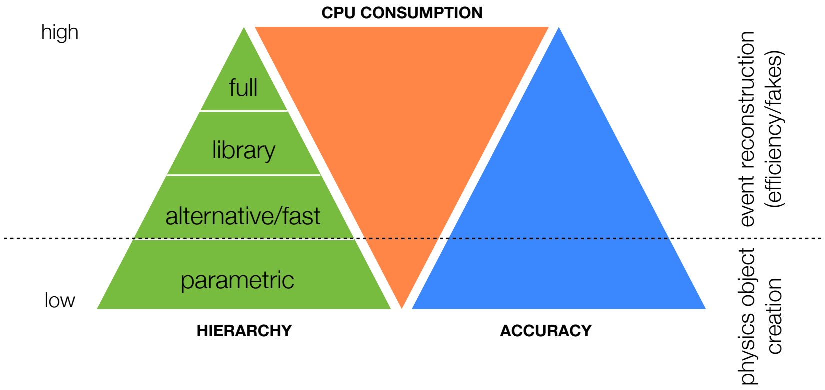
ATLAS grid disk utilization:



ATLAS Grid Usage 2012

- MC production dominates grid CPU usage
- MC production significantly contributes to grid disk usage
- Limits available MC sample sizes
- Some physics studies have significant limitations due to low MC statistics

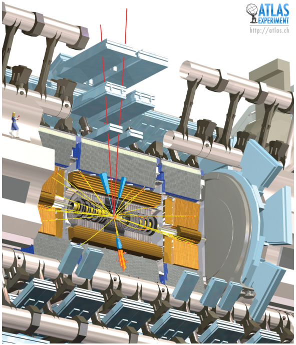
Simulation Hierarchy



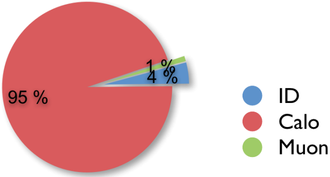
Simulation Hierarchy

- tradeoff between accuracy and simulation speed

Simulator: Geant4



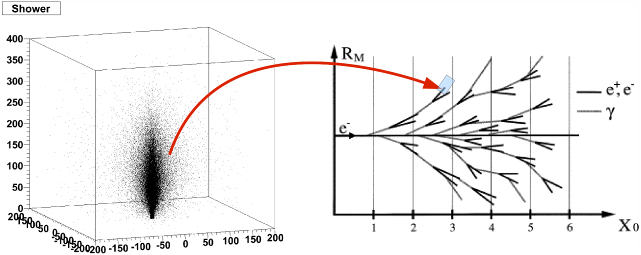
Geant4 simulation time per subdetector:



ATLAS Geant4

- Geant4 has a long history in ATLAS
- Close collaboration with Geant4 team
- Default full detector simulation in ATLAS
- Very successful, highly validated, very stable
- Very CPU intensive
  - Simulate many types of microscopic interactions for particles traversing material
  - $\sim$ 30M volumes
  - Most time spent in EM calorimeters

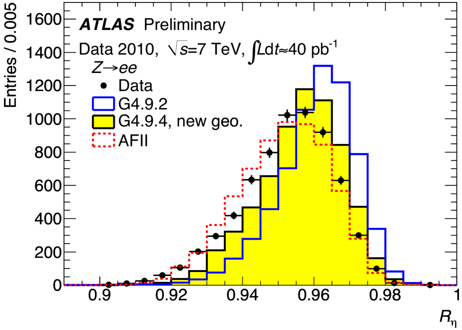
Simulator: Frozen Showers



ATLAS Frozen Showers

- Speed optimization for forward EM calorimeters
- Low energetic particles get replaced by pre-simulated EM showers
- Shower library generated with Geant4 simulation
- Used by default for *full simulation* MC campaigns

Simulator: FastCaloSim

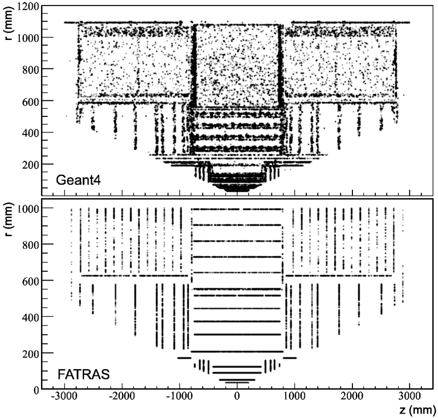


Energy Ratio R_η in a $\Delta\eta\Delta\phi = 3 \times 7$ cell cluster with respect to a 7×7 cell cluster size in the bulk EM calorimeter layer 2.

ATLAS FastCaloSim

- Parameterized calorimeter response
- Parameterization based on Geant4 full simulation
- Good description of EM processes
- Has been tuned to data for EM shower shapes

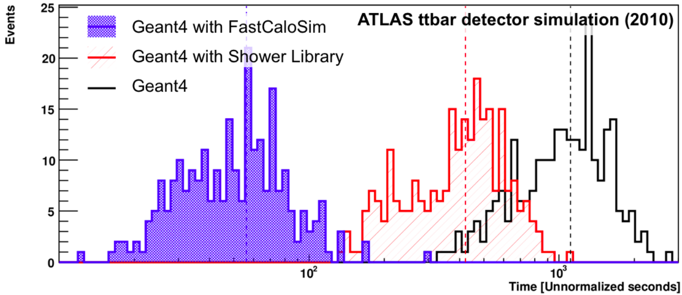
Simulator: Fatras



ATLAS Fatras

- Fast ATLAS tracker simulation
- For Inner Detector, Muon Spectrometer and muons in calorimeter
- Based on simplified detector geometry and physics interaction processes
- Dense material projected onto thin surfaces
- Ongoing work to access Geant4 hadronic interaction modules

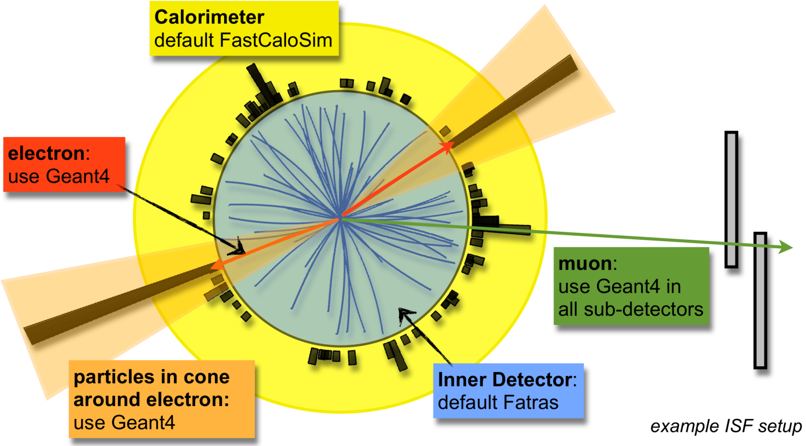
ATLAS Detector Simulation Time



ATLAS Simulation Time

- **ATLAS full detector simulation**: 2.7 B events simulated in 2012
  - Geant4 + FrozenShowers
  - Speedup ∼ 2 compared to Geant4 only
- **ATLAS fast detector simulation**: 1.8 B events simulated in 2012
  - Geant4 + FastCaloSim
  - Speedup ∼ 10 compared to Geant4+FrozenShowers

Integrated Simulation Framework (ISF)



example ISF setup

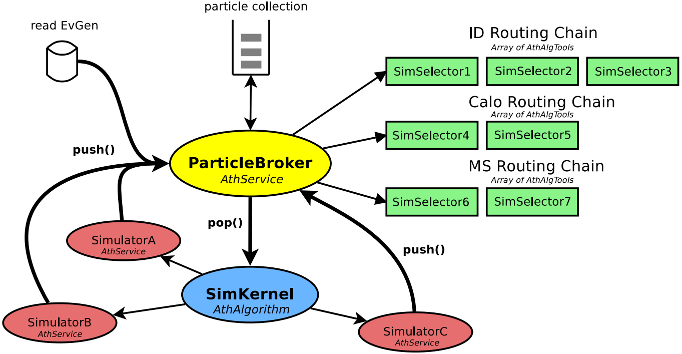
ISF Vision

- Combine different simulators in one framework
- Flexible rules for particle→simulator assignments

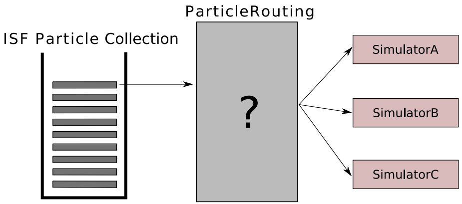
ISF Core Design

Main Components

- **SimKernel**: responsible for sending particles to simulators
  - Algorithm implementing the main particle loop
- **ParticleBroker**: stores particles and determines which simulator should be used for each particle
  - Uses *RoutingChain* to determine appropriate simulator
  - Separate RoutingChains for each sub-detector



Particle Routing



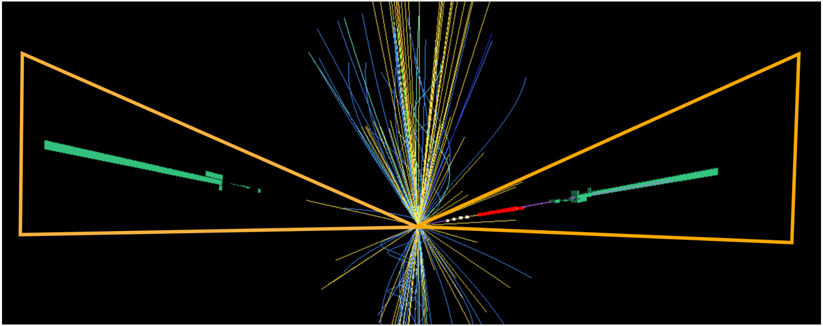
ISF main purpose is to route particles through sub-detectors and different simulators

- Book-keeping necessary

Basic Router Requirements

- Static routing rules: e.g. using kinematic parameters or particle type
- Dynamic routing rules: consider other particles in the event
- Simple to configure
- Intuitive, no deep technical knowledge of the ISF should be needed to understand the routing behaviour

ISF Usecase: H$\rightarrow\gamma\gamma$ fast simulation



- Very high statistics needed by this physics group for background shape studies
- ISF setups based on partial event simulation
  - Simulate only generator photons and particles in cone around them
- various ISF setups currently studied:
  - Geant4 inside cones
  - Geant4 inside cones and FastCaloSim for calorimeter
  - Fatras for whole Inner Detector and FastCaloSim for calorimeter
  - Fatras only inside cones and FastCaloSim for calorimeter

ISF Simulation Time Measurements

ISF Simulation Setup	Speedup	Accuracy
Full Geant4	**1**	best possible
Geant4 with FastCaloSim	**~25**	approximated calorimeter
Fatras with FastCaloSim	**~750**	all subdetectors approximated
Fatras with FastCaloSim only simulating particels inside cones around photons	**~3000**	all subdetectors approximated + partial event simulated

ggF Higgs $\rightarrow \gamma\,\gamma$ sample, no pileup

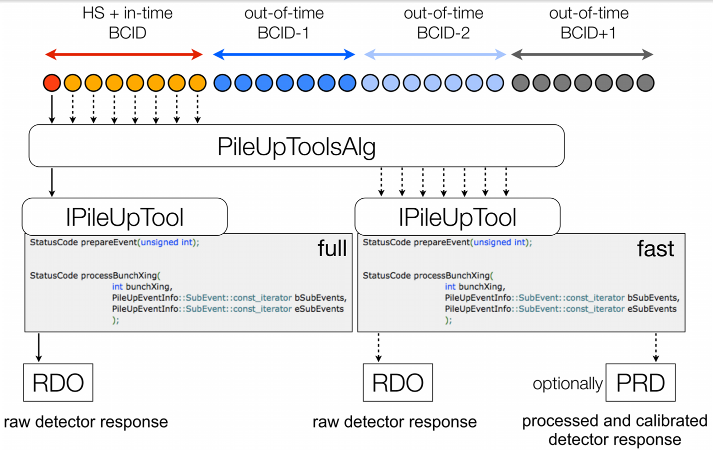
ISF Simulation Time

- significant speedup possible with **fast simulation engines**
- **partial event simulation** gives additional significant speedup
  - also reduces output filesize

Fast Digitization

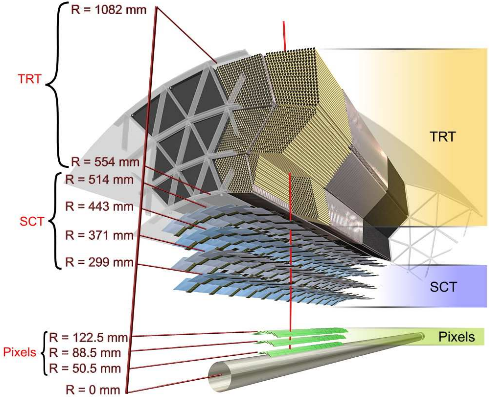
Motivation

- detector simulation orders of magnitude faster (up to factor 3000)
- digitization and reconstruction becoming limiting factors in MC production chain



HS: Hard Scatter (Event), BCID: Bunch Crossing Identifier, RDO: Raw Data Object, PRD: Prepared Raw Data

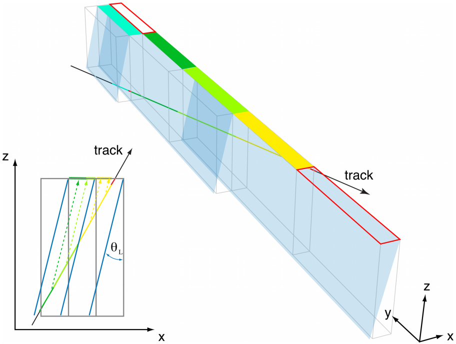
Fast Digitization in the ATLAS Inner Detector

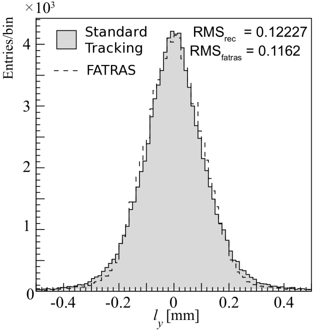


ATLAS Inner Detector Digitization

- Two different technologies:
- Silicon tracker → Pixel and silicon microstrip SCT Detector
- Transition radiation tracker with drift tubes → TRT Detector

Fast Digitization in Silicon Detectors



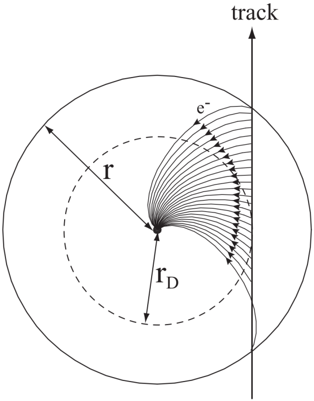


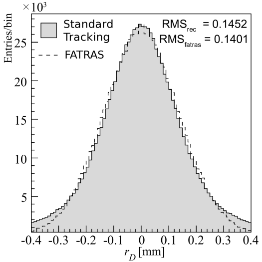
Comparison of Pixel cluster position residuals l_y between historic FATRAS fast digitization module and standard reconstruction.

Fast Digitization in Silicon

- Estimate charge deposition per read-out channel
- Project simulated track length in silicon onto read-out surface
- Lorentz angle drift correction ($\vec{E} \times \vec{B}$)
- Multiple scattering of drifting charge carriers

Fast Digitization in the TRT Detector





Comparison of TRT drift radius residuals r_D between historic FATRAS fast digitization module and standard reconstruction.

Fast Digitization in TRT

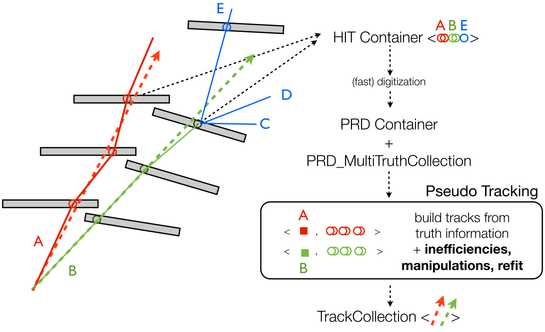
- Create *prepared raw data* (PRD) objects from simulation hits

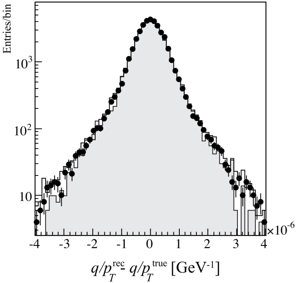
$\rightarrow$ Compute **closest approach radius**

$\rightarrow$ Determine **measurement uncertainty**

- Parameterized response of transition radiation (used for particle identification)

Fast Reconstruction



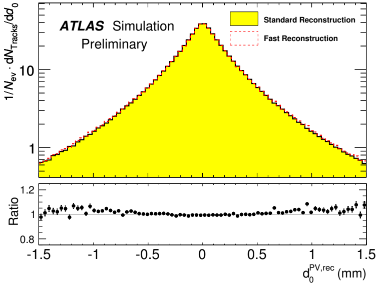


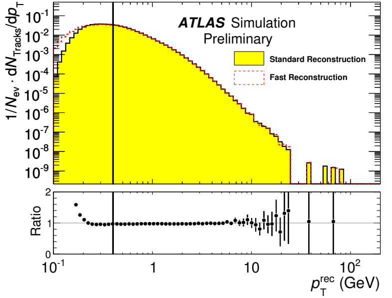
Comparison of q/p_T resolution between historic FATRAS fast refitting mode (dots), standard reconstruction (line) and a parameter smearing approach (grey area), for single muons with $p_T = 100$ GeV.

Fast Tracking

- Track seeding using Monte Carlo truth information
    - Skipping pattern recognition, track seeding and ambiguity treatment
    - → Most time consuming part in standard reconstruction
- Correct for inefficiencies (especially in low momentum regime)

Fast Reconstruction – Performance

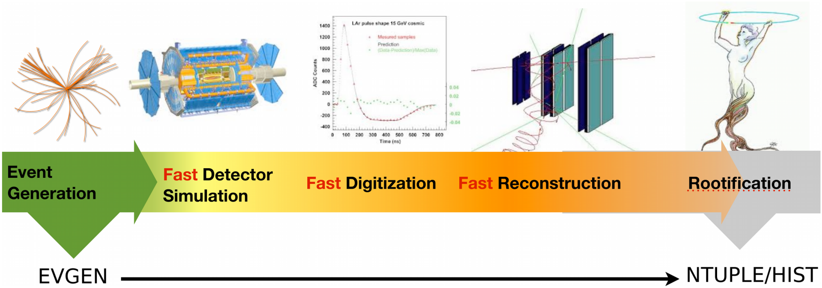




Fast Reconstruction

- Very good agreement with standard reconstruction (8 TeV minimum bias simulation shown)
- Significant speedup
- Standard data and MC processing requires $p_T > 400$ MeV
  $\rightarrow$ Difference in very low momentum regime insignificant

Fast ATLAS MC Chain



Fast MC Chain

- From event generator output to ROOT ntuples/histograms in one go
- Combining **Fast Simulation**, **Fast Digitization** and **Fast Reconstruction**
- No intermediate output files (minimizes file I/O overhead and disk space)

- Estimated event time: **few seconds per event !**
  - → Allows for private MC production in larger scales

Summary and Outlook

Summary

- Developing and maintaining **full and fast detector simulation**
  - Geant4, FrozenShowers, FastCaloSim, Fatras
- **ISF** – Integrated Simulation Framework
  - Mixing of different simulation engines, partial event simulation, region of interest simulation
- **Fast digitization** approaches
  - From simulated hits to calibrated tracking information (eg. drift circles, silicon hit clusters)
- **Fast reconstruction** approaches
  - Track seeding using MC truth

Outlook

- **Fast Monte Carlo chain**
  - Combining fast simulation/digitization/reconstruction efforts
  - From four vectors to ROOT histograms in one go
  - Estimated processing time: **few seconds per event**

Backup

Static-only Router Examples

Two Examples

- Particle Type Selector: send all muons to *SimulatorA*
- Kinematic Particle Selector: send all high η particles to *SimulatorB*

Pros

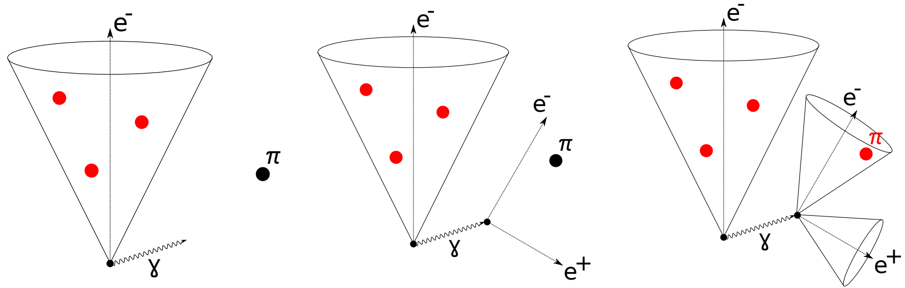
- **order independent**
- fully **consistent**: with the knowledge of particles after event simulation, the exact same decisions would have been made
- **intuitive** for the user
- **single pass** (each particle only simulated once)

Keep in Mind

- selector decisions may contradict each other

$\rightarrow$ selectors need to be defined in a priority list

Dynamic Router Example: Cone Selector



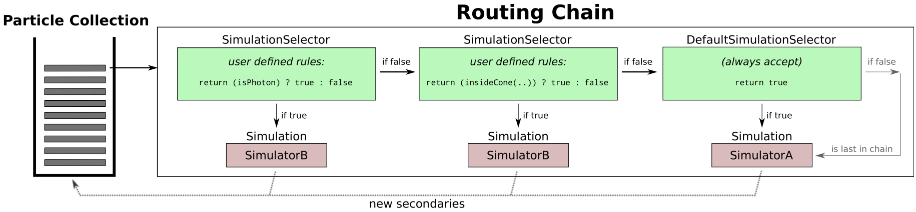
Dynamic Cone Selector

- the dynamic selector registers a cone for each new electron in the event
- all particles inside a cone are to be simulated in a certain simulation

Attention!

- decision on pion depends on the **simulation order**
- if π simulated before conversion: **inconsistent selector decision**

ISF Routing Chain: Functionality



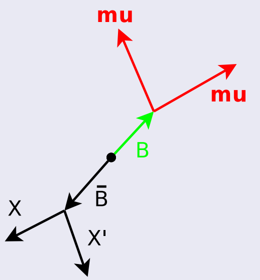
How the Routing Chain works

- dynamic `SimulationSelector` rules **can depend on initial EvGen particles**

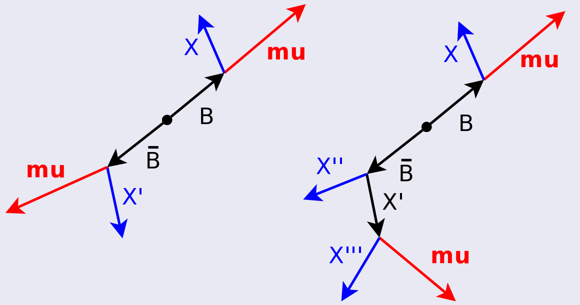
1. a particle is taken from the particle collection
2. the `SimulationSelectors` are asked in a specific order whether they would select the particle
3. in case a `SimulationSelector` does not take the particle, it will be handed over to the next in the chain
4. the first `SimulationSelector` which returns `true` decides that the particle will be sent to the simulation attached to this `SimulationSelector`

ISF Usecase: B-Physics fast simulation

Signal (green B)



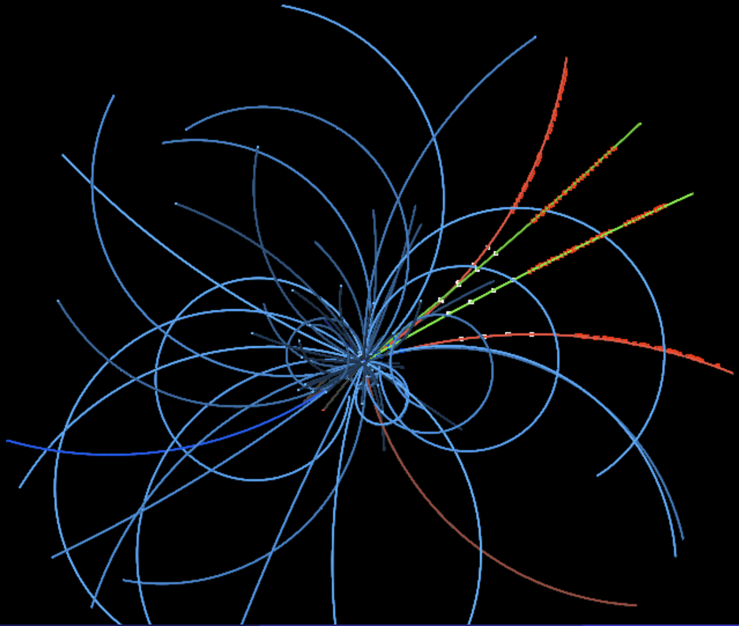
Background (two muons)



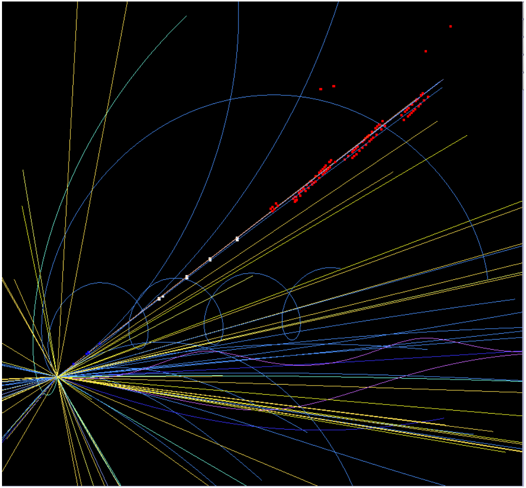
- very high stats needed for background shape studies
- **ISF simulation setups:**
  - only simulate particles that **have a B ancestor** (red and blue particles)

1. all particles→Geant4
2. muons→Geant4 , other particles→Fatras
3. all particles→Fatras
4. muons→Fatras (drop other particles)

ISF Usecase: B-Physics fast simulation



ISF Usecase: τ Tracking Performance



- study tracking performance of hadronic τ decays
- main motivation: clean simulation output, ie sensitive detector hits only from relevant decay products