The new ATLAS Fast Calorimeter Simulation

ICHEP 2016

Flavia de Almeida Dias

on behalf of the ATLAS Collaboration

05 - August - 2016
• Main subdetectors
  ➡ **Inner Detector** => Silicon and transition radiation technologies, in solenoidal magnetic field
  ➡ **Calorimeters** => Liquid Argon (LAr) Electromagnetic (EM) calorimeter (in central and forward regions) and hadronic calorimeter (tile in central and LAr in forward region)
  ➡ **Muon system** => spectrometer in toroidal magnetic field
Simulation is responsible for a very substantial fraction of ATLAS computing workload

- In full simulation with Geant4, the calorimeter accounts ~95% of the simulation time
- The inner detector follows with ~4%, while the muon systems take ~1% of the total simulation time
Simulation Hierarchy Pyramid

- Simulate interactions of particles with sensitive and non-sensitive detector material
- Produce sensitive detector hits with position and deposited energy information => input to detector electronics modeling (digitisation)
- More accurate simulation means slower simulation

➡ tradeoff between accuracy and speed

high

full

library

alternative/fast

parametric

HIERARCHY

low

ACCURACY

event reconstruction (efficiency/fakes)

physics object creation
Simulation Software in ATLAS

- Development of alternative software to the full Geant4 simulation in order to cope with very big datasets.

- Geant4 / Fluka, Flugg / Geant3
- Frozen Showers
- AF2 (Atlfast2) / AF2F (Atlfast2F)
- Atlfast

used in analysis

used extensively for the TDR (late 1990’s)
ISF - Integrated Simulation Framework

• Combines different simulation approaches in ATLAS into one framework
  ➡ Output format is always the same independent of simulation chosen
  ➡ Configuration is done at one central place and standardized
  ➡ Fast and full simulation setup can be mixed and used alongside

• Compatible with multithreading and multiprocessing
AF2 - Fast Calorimeter Simulation

- Replacement of calorimeter simulation with **FastCaloSim**
  - Parametrisation of calorimeter response based on Geant4
  - Tuned to data
  - Requires dedicated calibration
- Relative CPU speed-up:
  - ~ 20 with the AF2 setup
  - ~ 500 comparing only the calorimeter
- Used in production for physics groups in ATLAS
Current FastCaloSim

- Energy and shape parametrisation for single particles ($\gamma$, e, $\pi^{\pm}$), in E-$\eta$ grid
- Deposit of the particle energy in each calorimeter layer in one step, following correct:
  - total energy distribution
  - longitudinal shower depth
  - energy fractions per calorimeter layer
  - correlation of energy fractions per layer
  - overall lateral shower shape description

- Limitations
  - less fluctuations, leading to poor shower substructure description
  - simplified geometry
New FastCaloSim Parametrisation

- Re-simulate single particles ($\gamma$, e, $\pi^\pm$) on a fine energy and $\eta$ grid (approx 15x100 points)
  - Using full Geant4 simulation, current ATLAS geometry
  - Save detailed spatial information ($x,y,z,t,E$) of the developed shower

- Re-derive the energy parametrisation (longitudinal) and shower shape (lateral)
  - Reduce the amount of information to a compact form
  - Use multivariate analysis (TMVA) regression to approximate histograms
  - Add lateral shower fluctuations

- Assignment of hits to cells overcoming simplified geometry drawback
Energy Parametrisation

- **Aim**: Fast and accurate parametrisation, and use as little memory space as possible
- **1st Principal Component Analysis (PCA)**: Separate the events in 10 bins in the leading components

**2nd PCA:**

- **G4 Inputs**: Energy fractions $f$
  - Total energy $\rightarrow N$ inputs
- **Regression Training**
  - TMVA method MLP
  - Store weights in root files
- **To be stored**:
  - 10 PCA Matrices
  - Mean and RMS of PCA output data
  - MLP weights
Energy Parametrisation (II)

- From random numbers and stored weights and PCA matrix
Lateral Shower Shape Parametrisation

- Get the **average shower shape** per layer and per PCA bin

Binning defined iteratively in \((\alpha, dr)\) using mm units to match the calorimeter quantities

Number of hits per bin for the neural network (NN) fit - kept as constant as possible for the NN fit stability
Lateral Shower Shape Parametrisation (II)

- The binned energy fraction distribution is fit with a Neural Network (NN).
- Store TMVA weights instead of histograms.
Lateral Shape Fluctuations

- Introduction of random fluctuations in the cell energy deposit
  - Produce more clusters in the shower
  - Better description of the cluster level variables w.r.t. Geant4 simulation
New FastCaloSim Prototype

- Liquid argon calorimeter: accordion shape
  - Not described by simplified geometry
  - Assign hits to cells taking the wiggles into account
- Putting all ingredients together into a prototype

Diagram:
- Simplified Geometry Hit Assignment
- Fixed Hit Assignment

Flowchart:
- Input ISF Particle
  - Rand #
  - Decision: E param
  - E param inputs
  - Decision: Shape param
  - Shape param inputs
  - Assign to cells
  - Output HitCollection
  - Hits
Summary and Outlook

- Fast simulation essential for ATLAS physics programme
  - About 5-7 billion events are simulated in ATLAS per year, out of which 50% are produced using fast simulation

- ATLAS is developing and maintaining both full and fast simulation software
  - Decide, case-by-case, what is the desired tradeoff between accuracy and speed

- Fast Calorimeter Simulation is crucial for fast simulation, and a new parametrisation is ongoing
  - Update to benefit from knowledge acquired during Run 1, and to address known shortcomings of the current software
  - Use of neural network techniques to improve the parametrisation and to optimise information storage
  - Prototype to undergo first tests and tuning in the upcoming fall/winter
Backup Slides
New FastCaloSim - Scheme - Parametrisation

ESD, reco cells → G4 hits → position G4 hits

ESD reader

Muon punch through

Matching + adjusting reco cell ↔ G4 hits ↔ G4 position hits

Energy param (reco cells)

Bins leading variable

Correlated E param in bins

Average shape (G4 position hits)

TMVA regression

Average shape function

Cell fluctuations

Cluster fluctuations

Automatic tune
New FastCaloSim - Scheme - Simulation

- Find $E, \eta$ parametrisation
- $N$ random numbers
- Find $E_i$ for each layer $i$
- Energy simulation
- Leading var bin
- Shape for each layer $i$
- Find shape $E, \eta, ID$, leading var
- Random cell deposit following average shape
- Define random subclusters