Fast Simulation with ATLAS

Digital Twins for Nuclear and Particle Physics
NP-Twins 2024

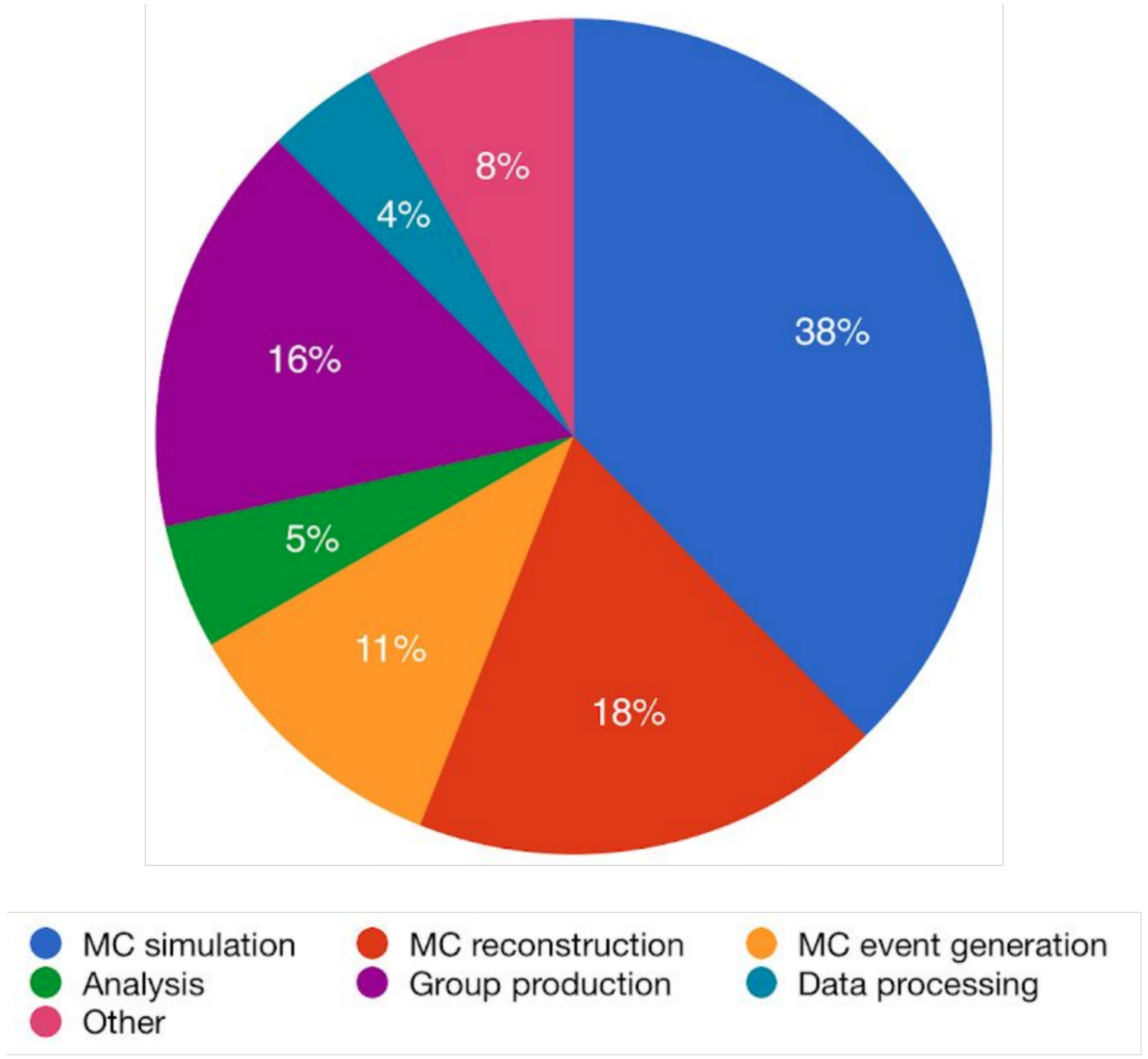
Federico A.G. Corchia
(University and INFN, Bologna, IT)
On Behalf of the ATLAS Collaboration

Genoa, 17 December 2024

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Introduction

Calorimeter Simulation Needs

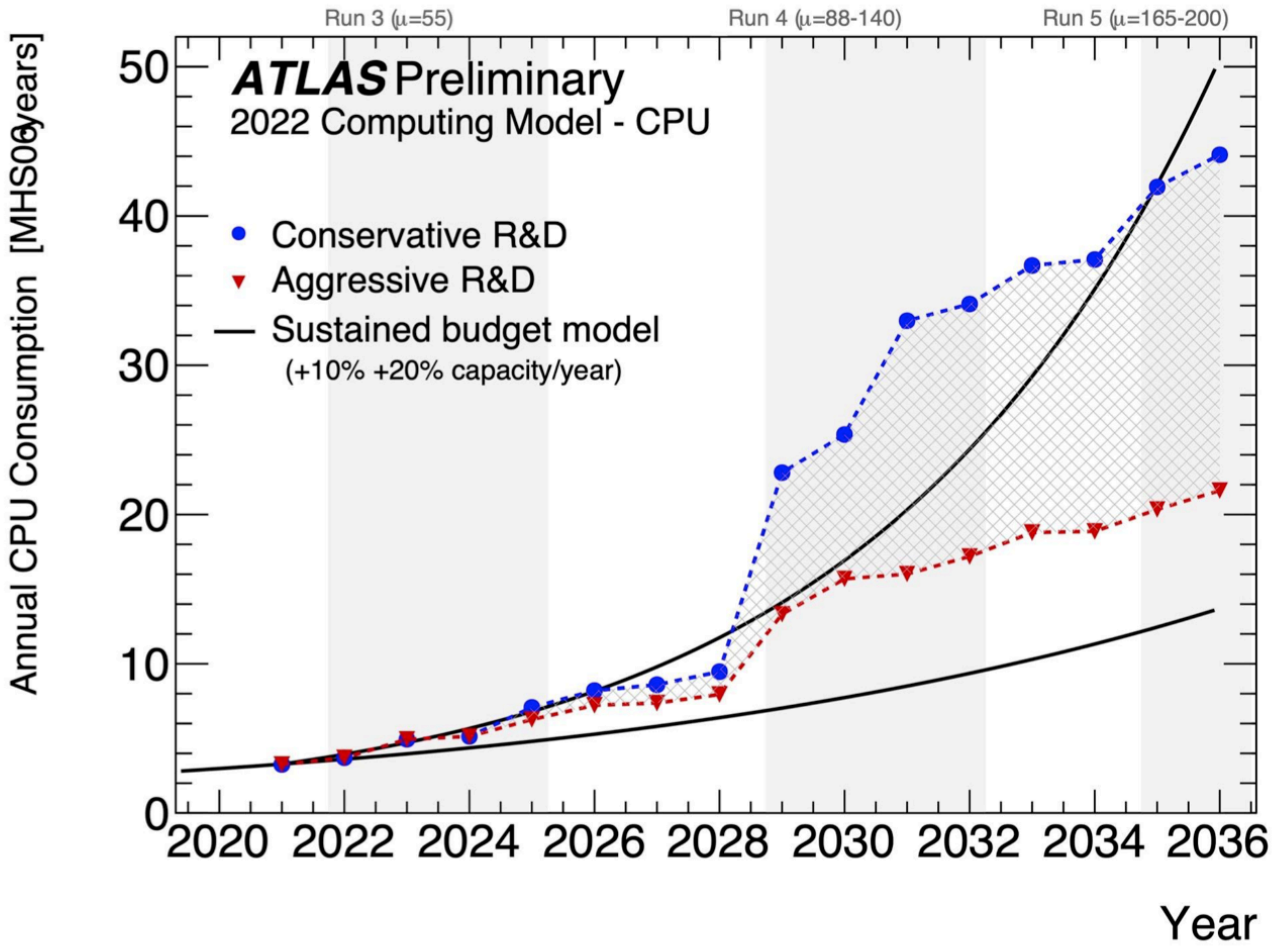


CPU-hours taken by ATLAS activities
The ATLAS Coll., ATLAS HL-LHC Computing Conceptual Design Report, CERN-LHCC-2020-015 (2020)

- MC simulation of detector response is a **major computing challenge** at LHC experiments!
  - A **large part** of the resources taken by detector simulation is taken by **calorimeter simulation** (~80% of CPU consumption for $t\bar{t}$ processes)

Introduction

Calorimeter Simulation Needs in the Future



The ATLAS Coll., ATLAS Software and Computing HL-LHC Roadmap, CERN-LHCC-2022-005 (2022)

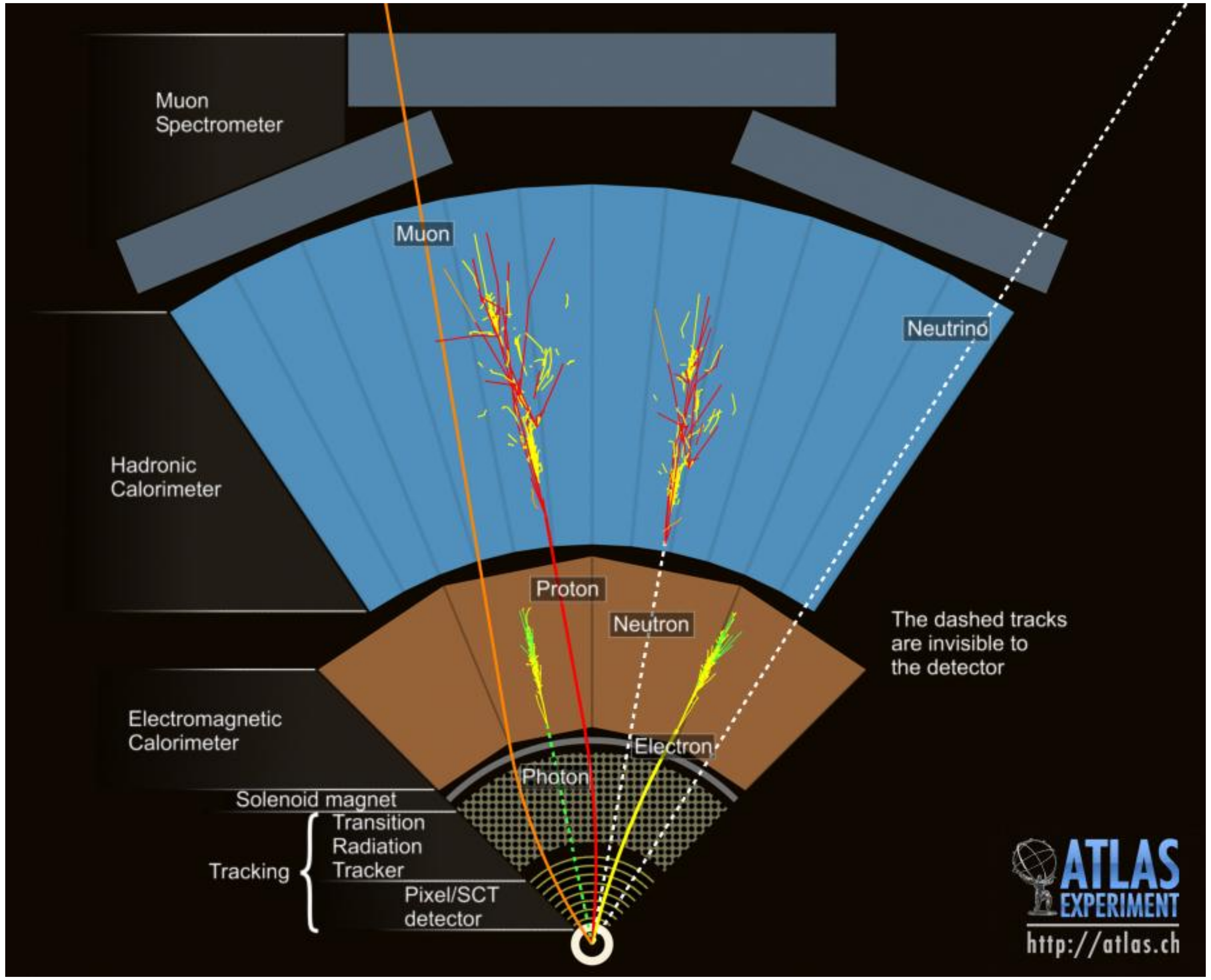
- CPU needs were expected to **increase** during Run 3 and in view of Run 4: significant R&D required to keep up!
- Solutions needed → **fast simulation**!

AtlFast3

Fast Simulation in ATLAS and How It Is Done

- **Fast simulation tools**: able to simulate calorimeter response faster than the ("traditional") full process simulation tool Geant4 but keeping high accuracy.

- **AtlFast3**: the fast simulation tool developed by ATLAS.
  - Introduced for Run 2 and further improved for Run 3. Already in production for Run 3;
  - It replaces the slow propagation and interactions of incident particles inside the calorimeter volume with the direct generation of energy deposits based on an underlying parametrisation;
  - Uses a simplified geometry of the calorimeter cells to simplify its complex and non-homogeneous structure.

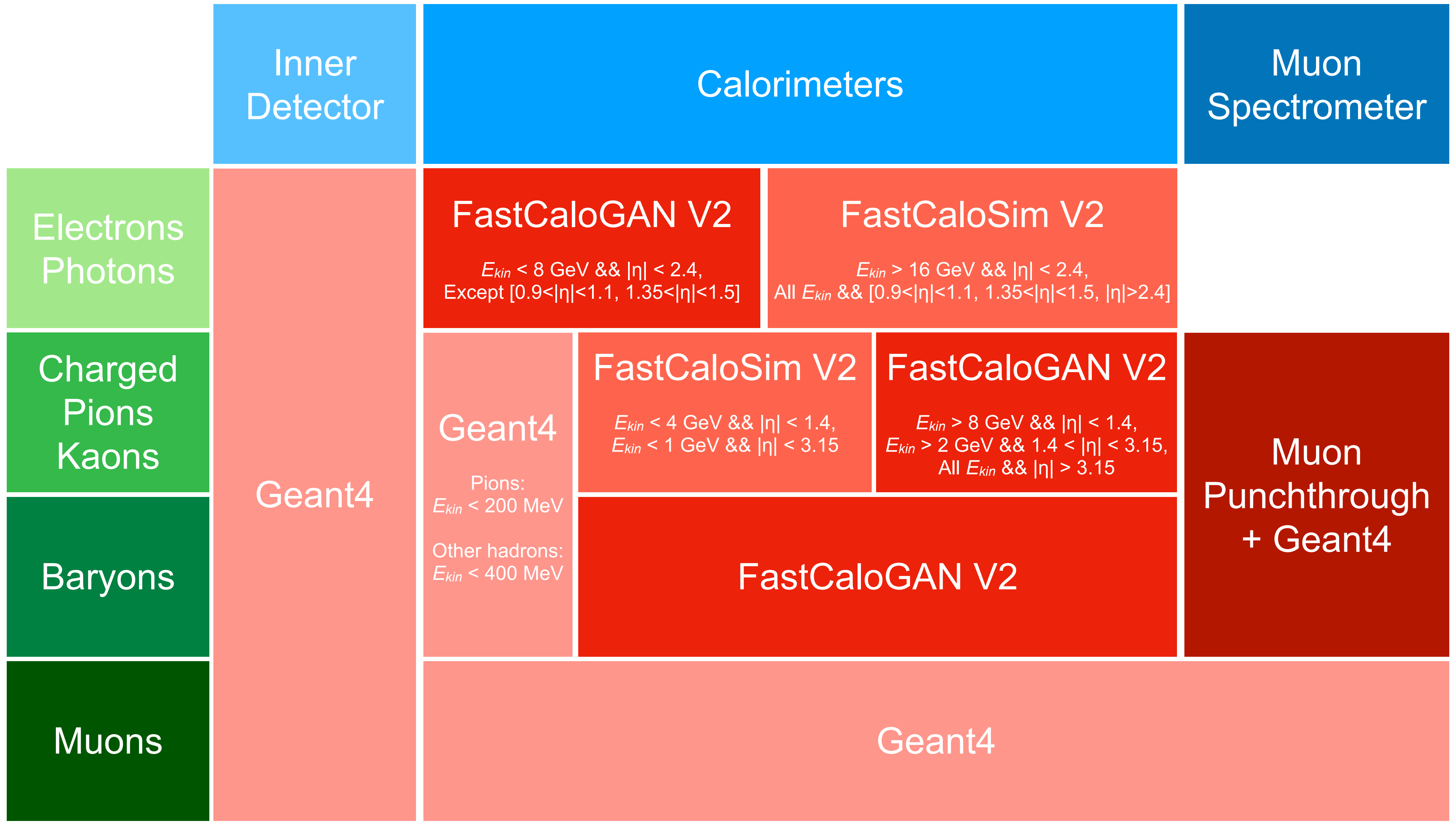


cds.cern.ch/record/1096081

AtlFast3

Subcomponents

- Employs two fast simulation approaches:
  - **FastCaloSim**: parametric approach for shower development;
  - **FastCaloGAN**: based on Generative Adversarial Networks (GANs).



ATLAS-SIM-2024-004

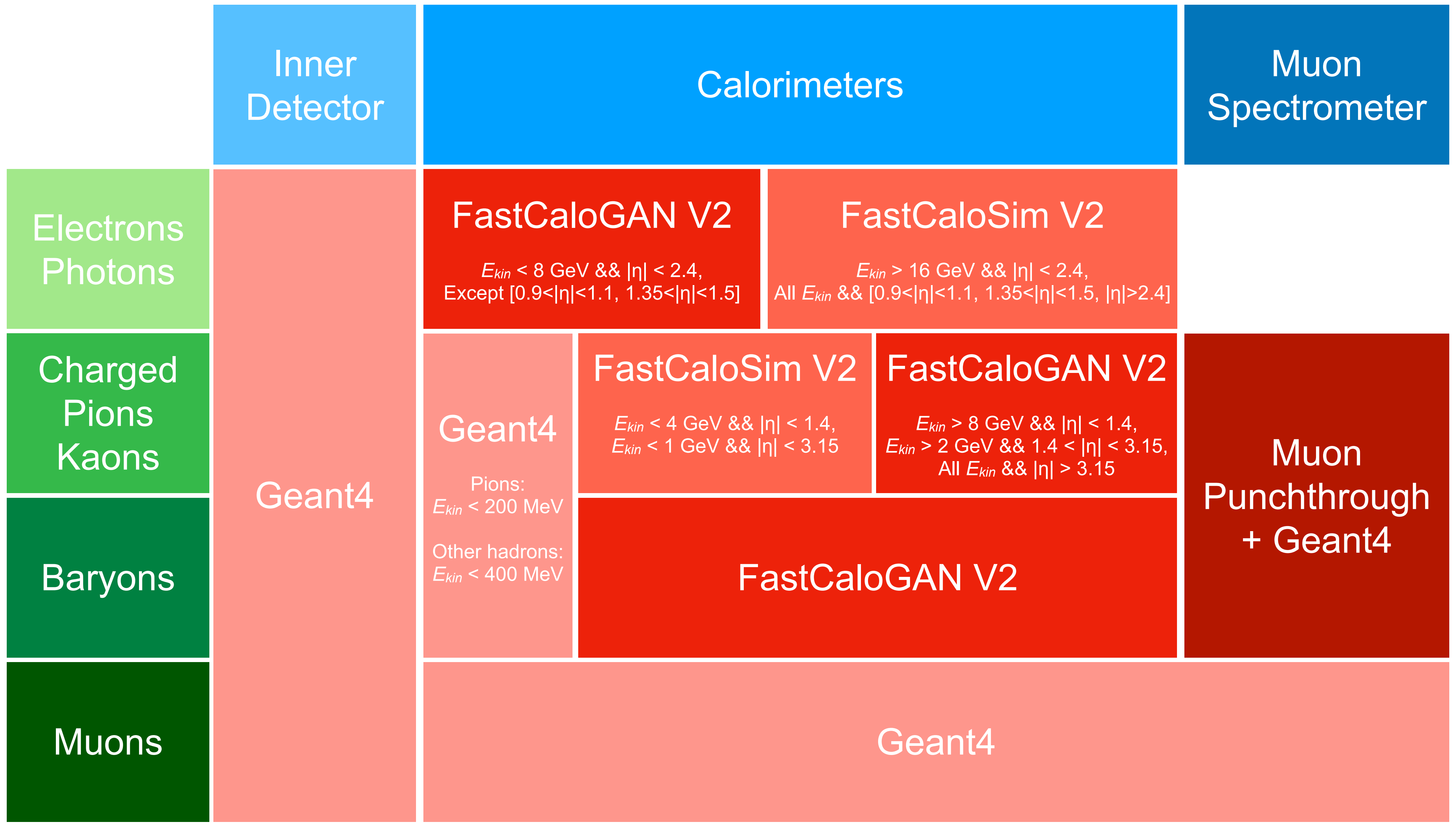
AtlFast3

Subcomponents (cont'd)

- AtlFast3 runs fast simulation either through the parametric approach (FastCaloSim) or through GANs (FastCaloGAN), depending on which one **returns the best simulation** (i.e. the most accurate with respect to Geant4) for the type of the particle initiating the shower and its energy.

Geant4 still used to simulate all particles in the inner detector, low energy hadrons in the calorimeters and muons

Muon Punchthrough (the spray of particles into the muon spectrometer resulting from late interacting high-energy hadrons) is modelled with a tool based on DNNs

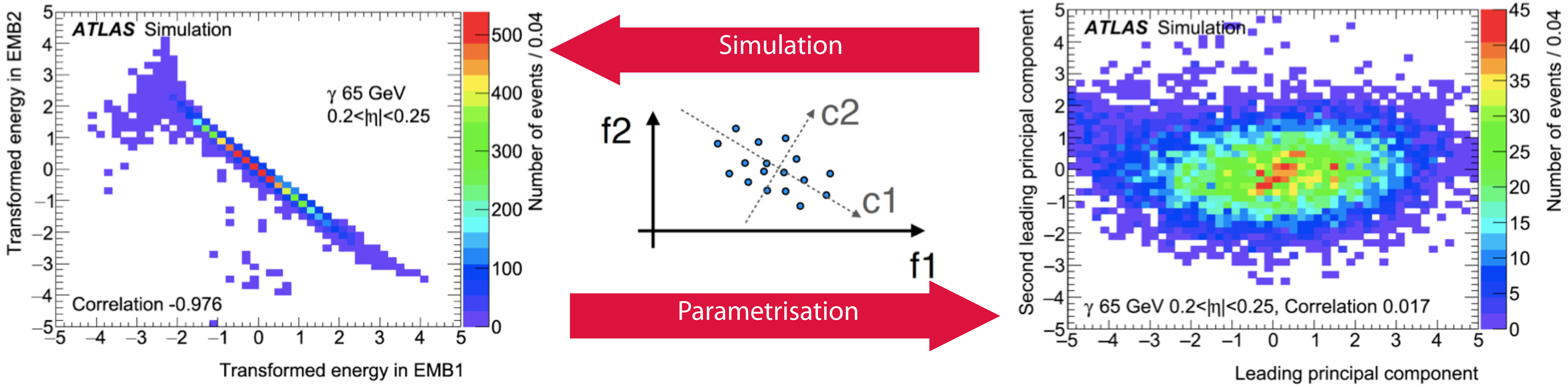


ATLAS-SIM-2024-004

FastCaloSim

- **Parametrised** modelling using Geant4 single photon, electron and charged pion samples;
- Parametrisation in 17 logarithmically spaced energy bins from 64 MeV to 4 TeV and 100 linearly spaced bins in |η| from 0 to 5;
- **Separate** parametrisation of longitudinal and lateral shower development;
- **Energy in layers decorrelated** through Principal Component Analysis (PCA);
- **Average lateral energy distribution parametrised** as 2D probability functions.

The ATLAS Coll.,
AtlFast3: The Next Generation of Fast Simulation in ATLAS, Comput Softw Big Sci 6, 7 (2022)



FastCaloGAN

Basics on GANs

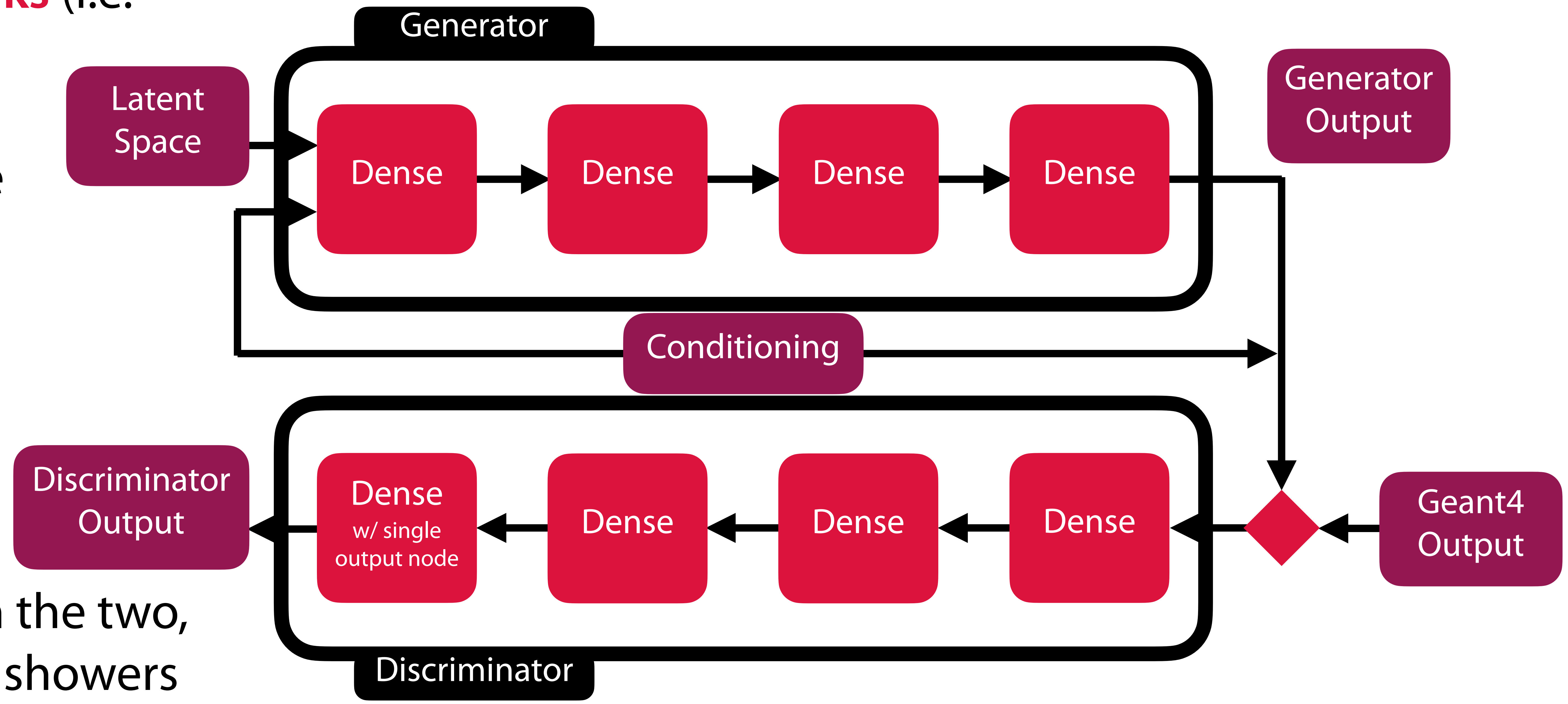
- Based on **Wasserstein GANs** with gradient penalty
  I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin and A. Courville, "Improved Training of Wasserstein GANs", arXiv:1704.00028 [cs.LG] (2017)

- **Simultaneous training of two neural networks** (i.e. the typical GAN structure):

  - **Generator**: aims at generating sample the most similar to Geant4 datasets;

  - **Discriminator**: aims at distinguishing Geant4 samples from the ones produced by the generator.

- When it has reached an **equilibrium** between the two, FastCaloGAN is ready to simulate calorimeter showers and it does so much faster than Geant4, keeping good accuracy.



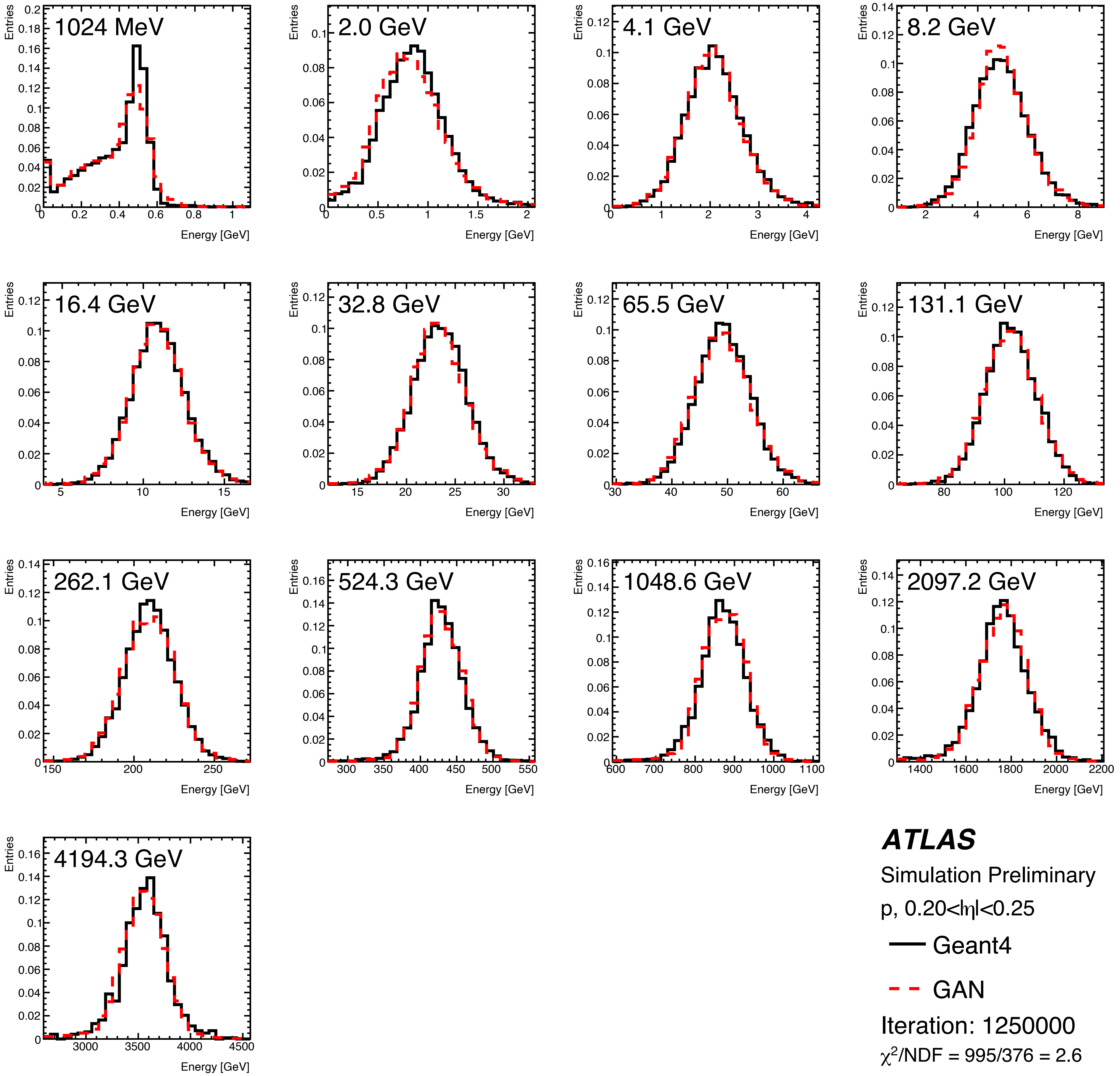
FastCaloGAN

Further Details

- Trained for e^-, γ, p, $\pi^\pm$ as shower-initiating particles;
- Trained on each of the 100 bins in $|\eta|$ and conditioned on truth momentum;
- Calorimeter hits are grouped into voxels (= 3D bins; their granularity is optimised and finer than the one of the calorimeter cells, which improves modelling). FastCaloGAN is trained to reproduce voxels and energies in layers as well as the total energy in a single step;
- Architecture of the network and hyperparameters have been optimised.

FastCaloGAN

Further Details (cont'd)



Fast (GAN) vs full (Geant4, i.e. "traditional") simulation of energy for protons.
Remarkable agreement!

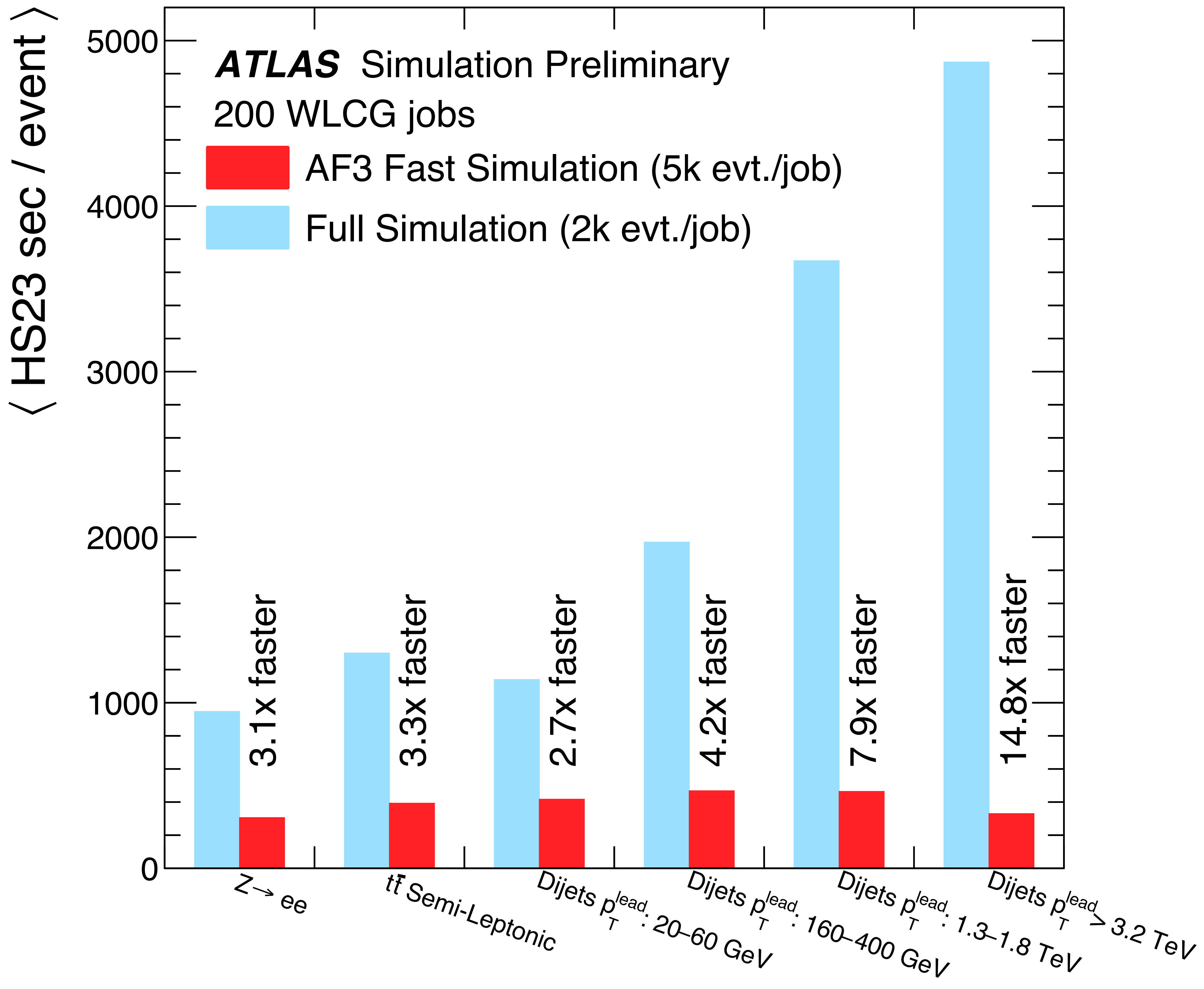
ATLAS-SIM-2023-004

Performance

Speedup

- AtlFast3 is **3** (for Z → ee events) **to 15** (for high-p_T dijet events) **times faster** than the Geant4 simulation of the ATLAS Run 3 detector for 2023 data taking (= mc23c production campaign)!

Simulation time in AtlFast3 dominated by full simulation of the inner detector



ATLAS-SIM-2023-005

Performance

Physics

- Very accurate modelling of the leading cluster energy with FastCaloGAN, as well as of the number of constituents for the jet and substructure variables in dijet events with the hybrid approach (FastCaloSim+FastCaloGAN combined);

- For most observables used in Physics analyses, AtlFast3 and Geant4 **agree within a few %**;

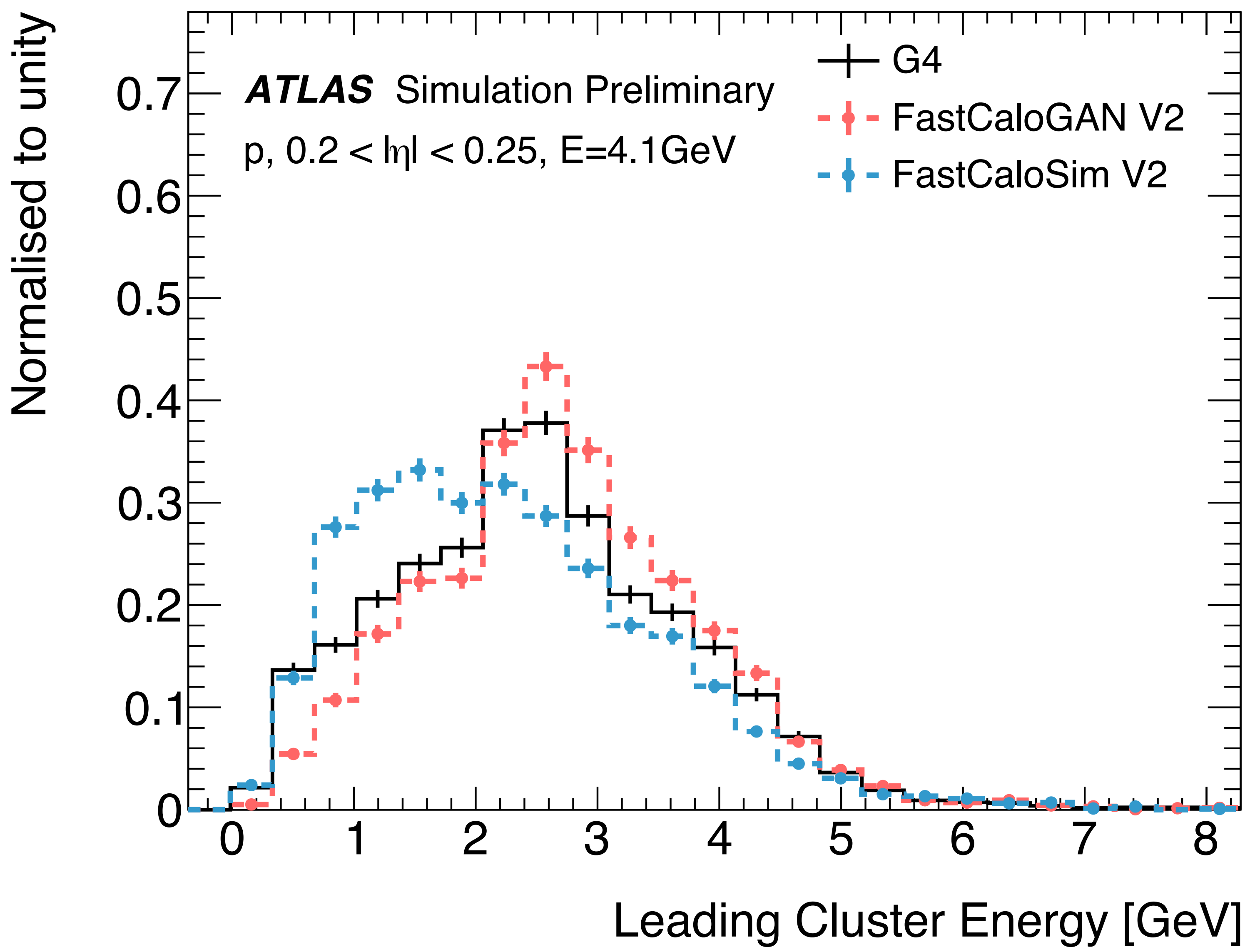
- AtlFast3 can be used **for almost every analysis** (not only signal but also background). Very high precision measurements might not be able to use AtlFast3, but for Run 4 **>90-95%** of analyses will be required to use fast simulation, as there will not be the CPU capacity to allow full simulation.

Performance

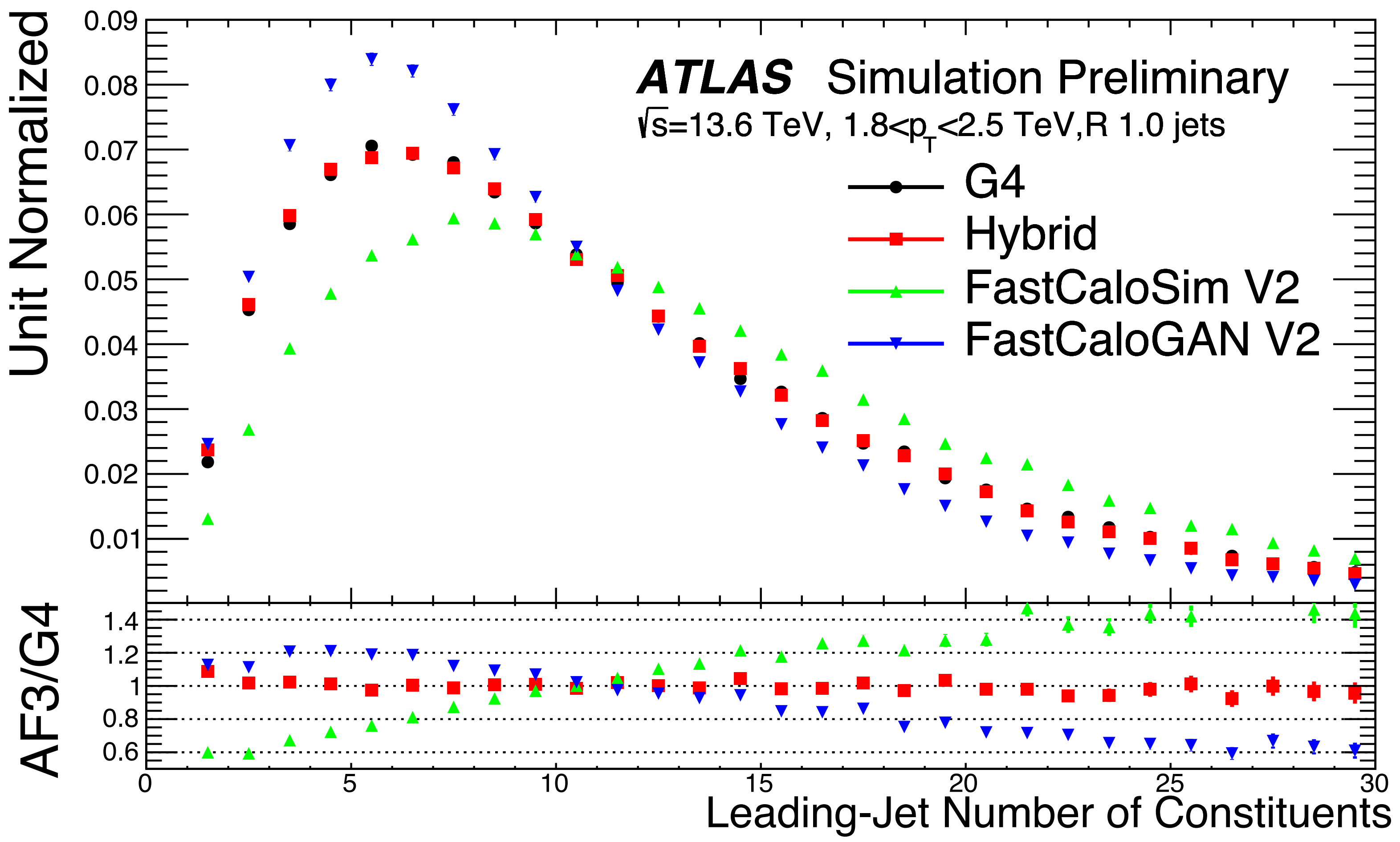
Physics (cont'd)

Hybrid approach = combining FastCaloSim+FastCaloGAN

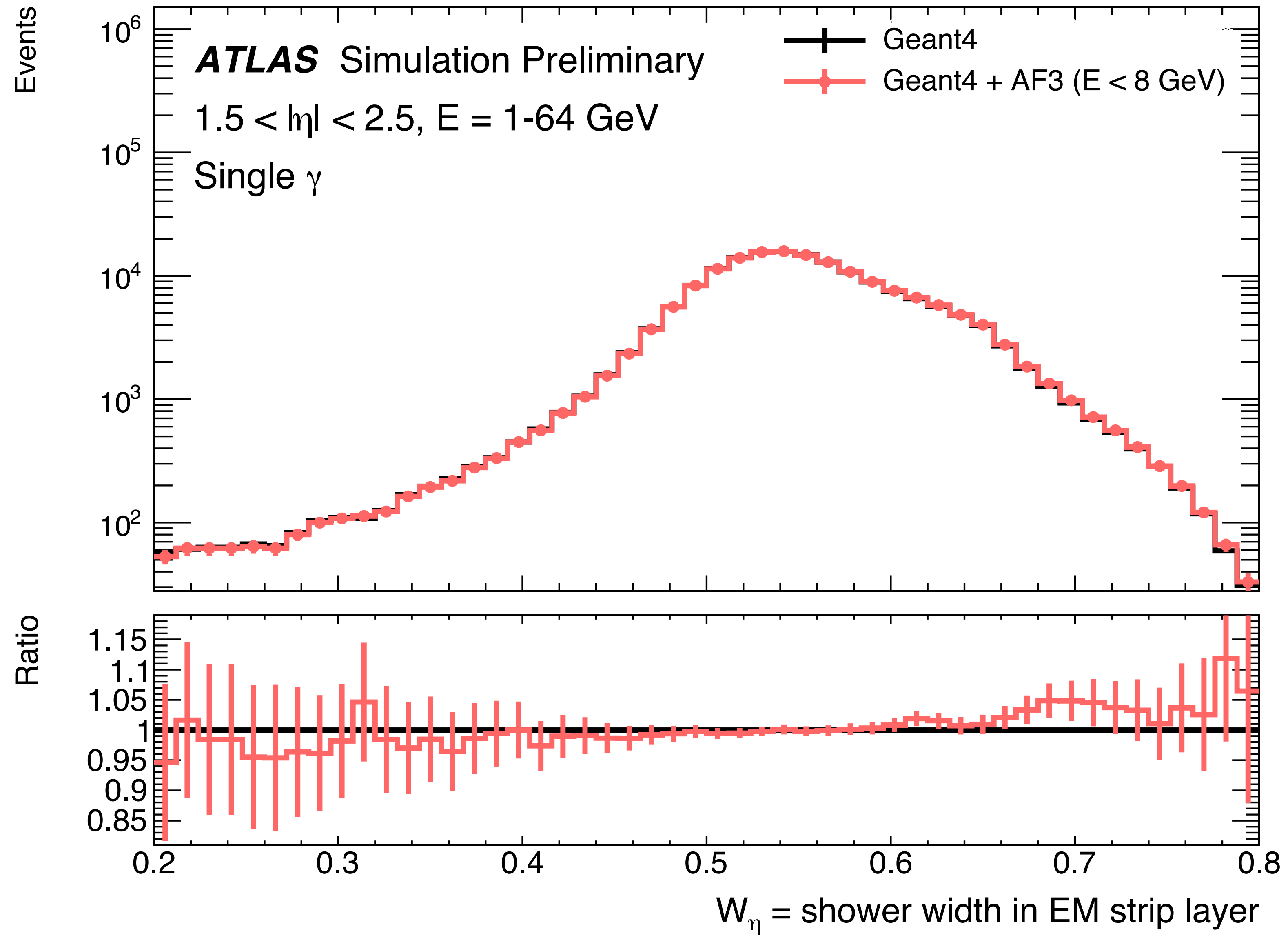
ATLAS-SIM-2023-004

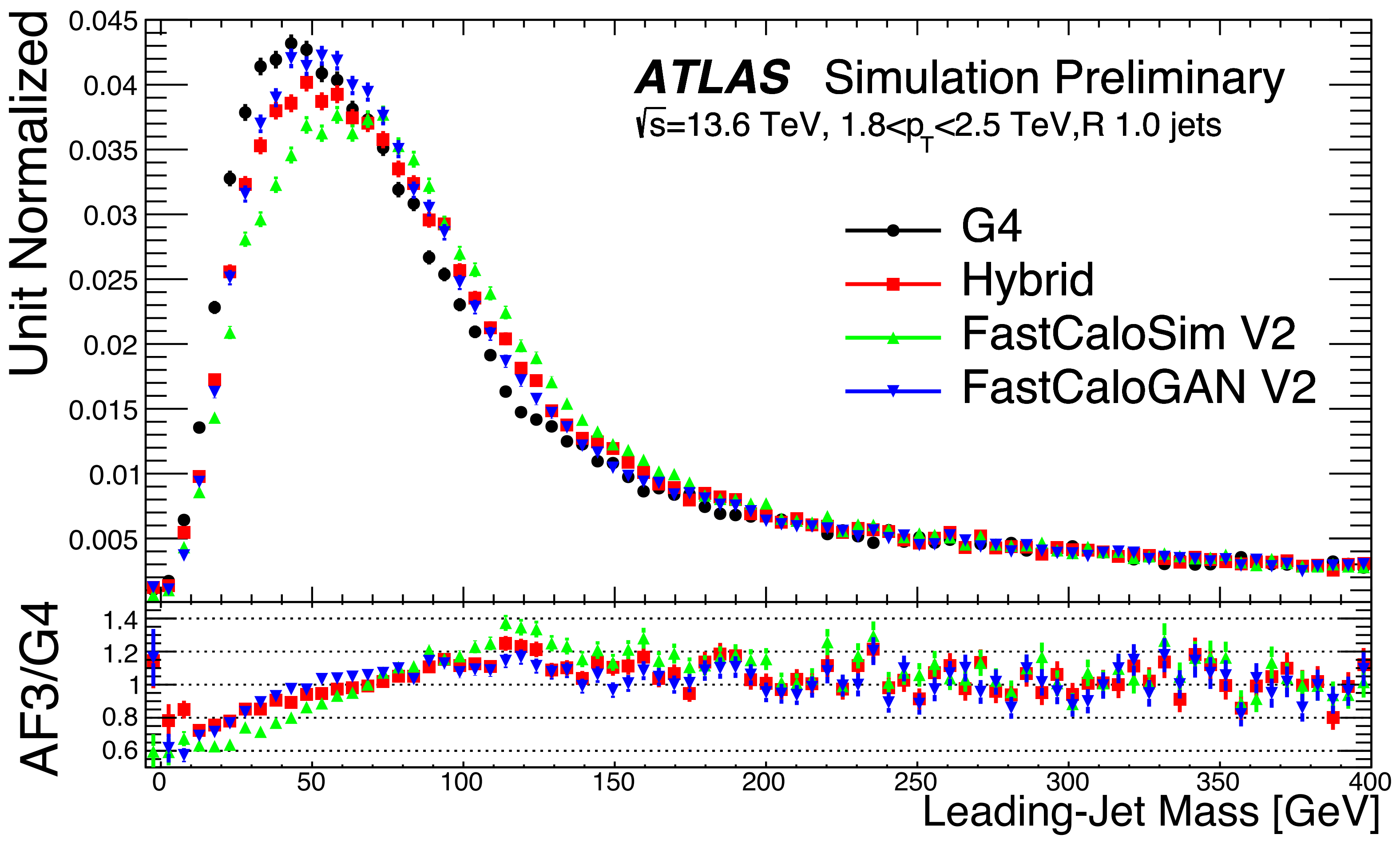


ATLAS-SIM-2023-004

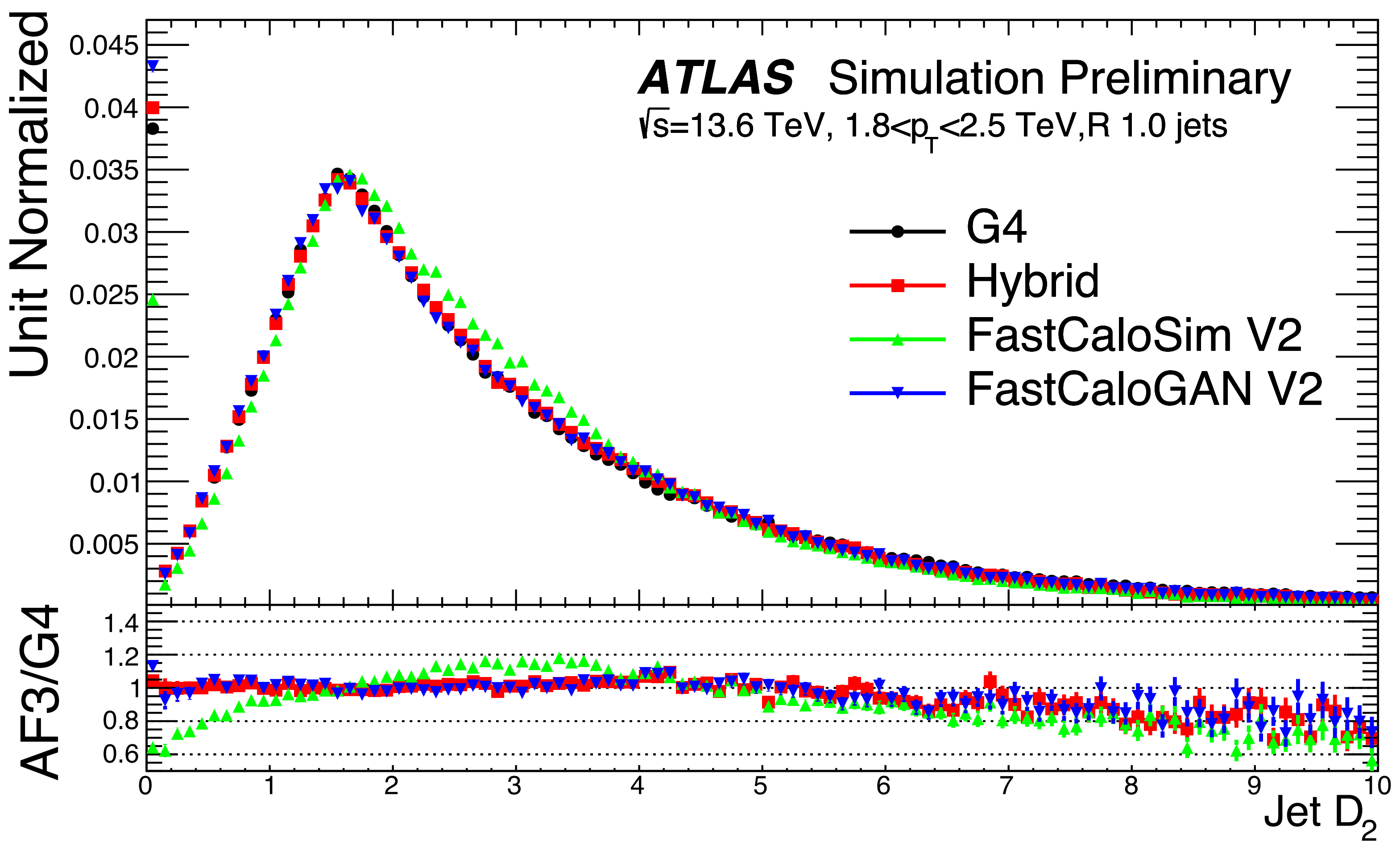


ATLAS-SIMU-2024-08





ATLAS-SIM-2023-004



ATLAS-SIM-2023-004

Looking Ahead

Towards Run 4

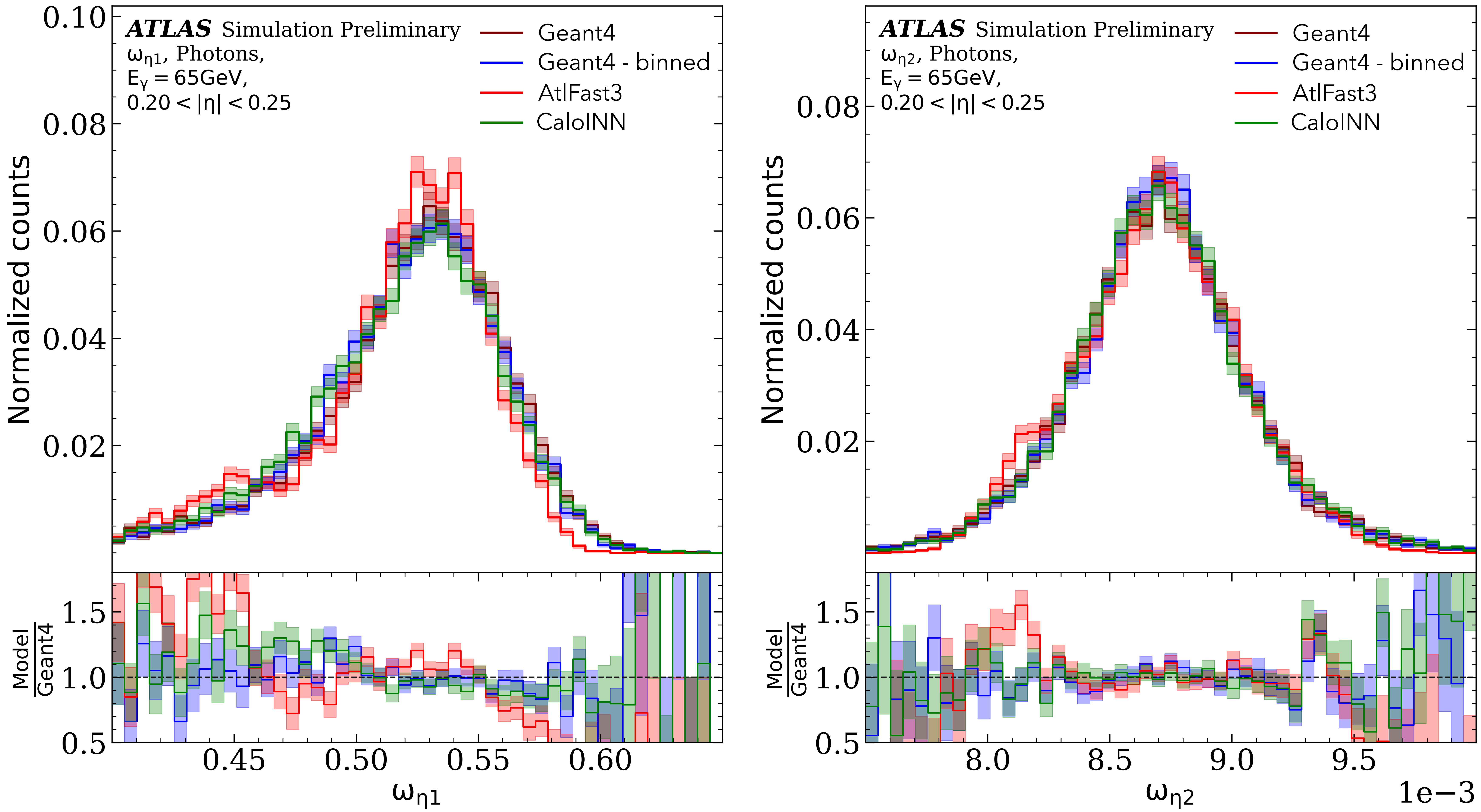
- After successful insertion into Run 3 production, further development is underway in view of the final part of Run 3 and Run 4:
  - tests on **improved** and finer voxelisation that reduces bias due to calorimeter geometry (fineness actually not that easy to handle!);
  - research **diffusion** models and **Invertible Neural Networks** (INNs). If we find them to be appropriate, we can add other subcomponents to AtlFast3/4 (or even replace GANs)!
- Current results show that **GANs still have potential** and compete well with Invertible Neural Networks, **which also yield very good results**.

Looking Ahead

Towards Run 4 (cont'd)

ATLAS-SIMU-2024-10

Voxelisation has been fully re-optimised and it was shown that the old voxelisation introduces significant Physics artifacts. INNs have been shown to accurately reproduce the Geant4 distributions with the optimised voxelisation.



Looking Ahead

An Extra Help: FastCaloGANtainer

More here:
F.A.G. Corchia, L. Rinaldi, M. Franchini,
Recent Advances in the GAN-based Fast Calorimeter Simulation of the ATLAS Experiment

- Usage of FastCaloGAN requires its GANs to be trained → large resources needed! → **we should also use other resources** than the ones commonly used at CERN (the CERN batch system LXBATCH, the Worldwide LHC Computing Grid)!
- FastCaloGANtainer **containerises FastCaloGAN training** for distribution on other resources. Based on the OS and software of LXBATCH, it is **independent** from the system where it is installed.
- Tested on Leonardo (the 9th most powerful cluster in the world, at CINECA, Bologna, IT) and on the HPC cluster at INFN-CNAF (Bologna, IT, close to WLCG INFN-T1). x3-4 speedup on Leonardo and x2-3 on CNAF-HPC for training with respect to LXBATCH (where training takes ~12 h for π and 30-31 h for γ). **Usage of supercomputers brings great advantage!**
- To do: distribution on other resources (also cloud), architectures (ARM) and for more particle types, code optimisation (e.g. to take even more advantage of multi-CPU/GPU nodes).

Conclusion

- AtlFast3 is the state-of-the-art fast simulation in ATLAS, able to simulate a broad range of Physics processes with high precision.
- AtlFast3 for Run 3 reaches high precision for many objects (improvement in Physics performance compared to predecessor) and 3-15 times CPU speedup.
- Essential to meet the computational requirements of the future runs of the LHC, as well as Physics modelling accuracy needs.
- Is being used in many Physics analyses.

Thank you!

Backup

Definitions

- $\omega_{\eta 1}$ = reconstructed lateral shower width in the strip layer of the ATLAS calorimeter;
- $\omega_{\eta 2}$ = reconstructed lateral shower width in the EMB2 layer of the ATLAS calorimeter.