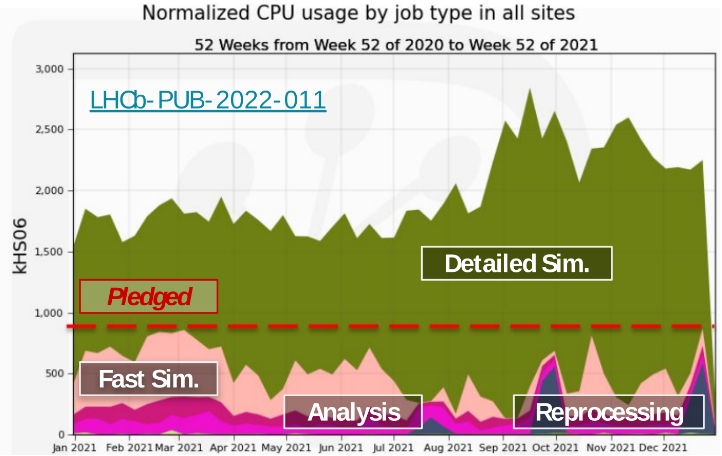
Fast simulation with generative models at the LHC

Liza Mijović, University of Edinburgh
on behalf of ALICE, ATLAS, CMS & LHCb

LHCP2024, Boston

Why fast simulation?

Because **we don't have computing resources** to produce sufficient samples for LHC analyses with detailed simulation.



Simulation with generative models

Generative models provide speed-up and high fidelity. Deployed in 2 ways:.

1) Fast simulation of detector components:

- **replace** detailed GEANT **simulation of bottle-neck detector**.
- For all experiments: first bottle-neck = **calorimeter.**

2) End-to-end fast simulation:

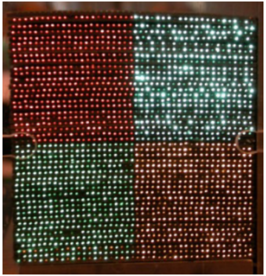
- **replace multiple simulation and reconstruction steps with fast approximations**, including generative models.
- Pro-s: cheaper (eg CPU, disk-space) con-s: lower fidelity & flexibility.

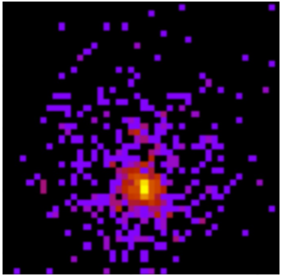
- **Finish off with cross-experiment software:**
  focus on those recently (co-)developed by the LHC experiments.

Fast Calorimeter simulation

Fast Simulation of ALICE ZDC

- Zero Degree Calorimeter (ZDC): measures energy of observers - particles that did not directly participate in collision.
- System of five sampling calorimeters placed at forward rapidity, both sides.
- Fast sim inputs: images with the dimensions of a ZDC fiber read-out (44x44).

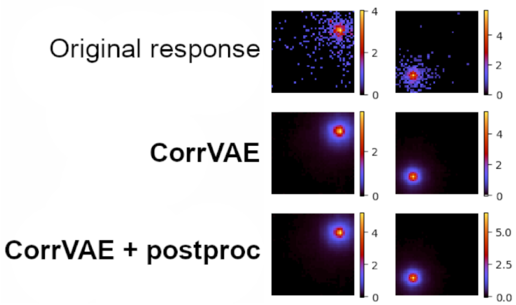




Source: arXiv:2306.13606. Thanks to Sandro Wenzel for inputs.

Fast Simulation of ALICE ZDC

- Objectives: speed-up, high fidelity of particular properties.
- Solution: based on CorrVAE - VAE with orthogonal latent spaces.



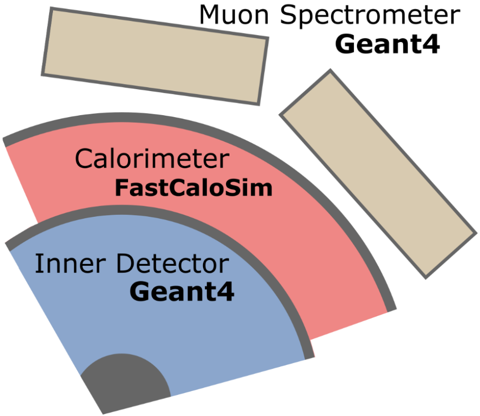
	MSE	Wasserstein
CorrVAE	**1.03**	**16.15**
CorrVAE + postproc	**1.18**	**3.83**
CVAE	1.02	6.35
CGAN	2.96	8.27
CGAN + reg + postproc	2.98	5.15

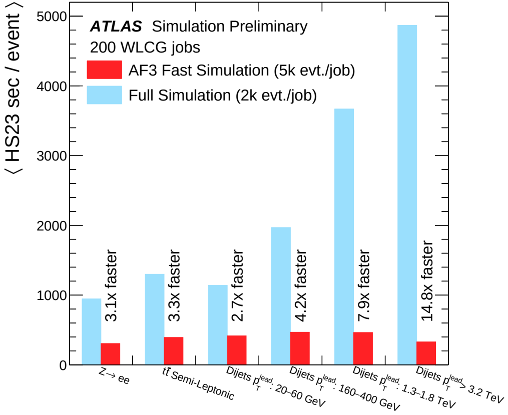
Prototype integrated in ALICE simulation work-flow, 100x speed-up of ZDC sim.

Source: arXiv:2405.14049, earlier work: arXiv:2306.13606, arXiv:2207.01561, arXiv:2006.06704. Thanks to Sandro Wenzel for inputs.

ATLAS Fast Calorimeter Sim.

- Problem: simulation dominates CPU resource usage, calorimeter: 80-90%
- ATLAS Fast Calorimeter Simulation: 3-15 times faster simulation.



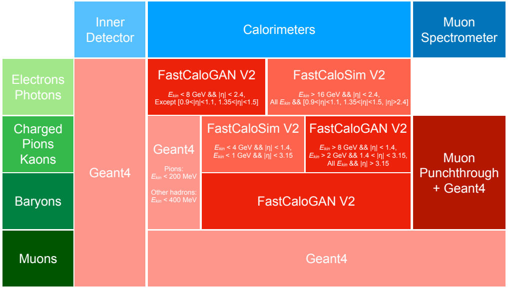


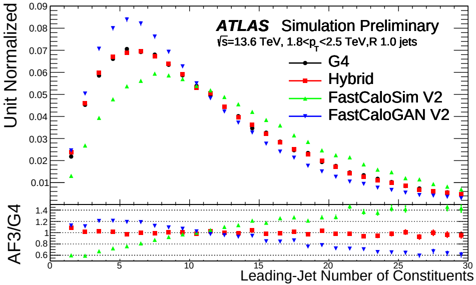
Sources: CERN-LHCC-2020-015, sketch: H. Ahmed's seminar, Figure: SIM-2023-005, ATLAS Simulation Public Results wiki.

ATLAS Calorimeter: Hybrid Sim.

Calorimeter fast simulation is a hybrid of two approaches:

- FastCaloSim: parametrised model.
- FastCaloGAN: trained to reproduce voxels and energies in layers as well as total energy in single step.

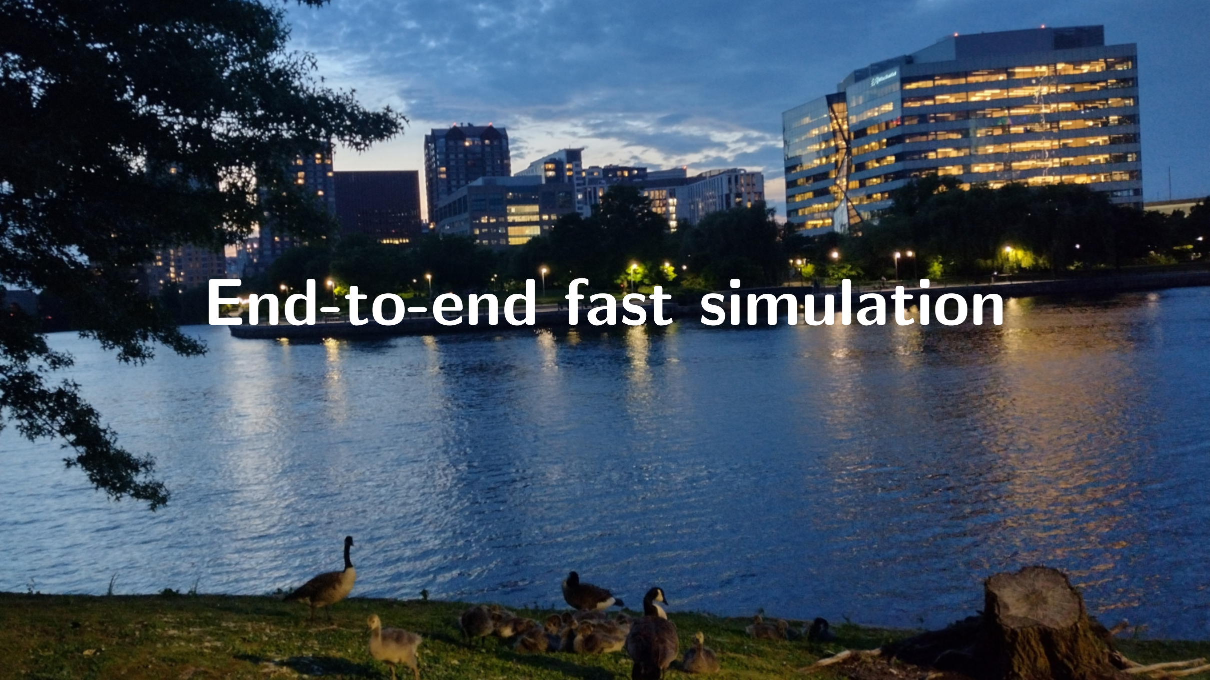




Deployed in ATLAS production work-flow, used for about 50% of Run2/3 events.

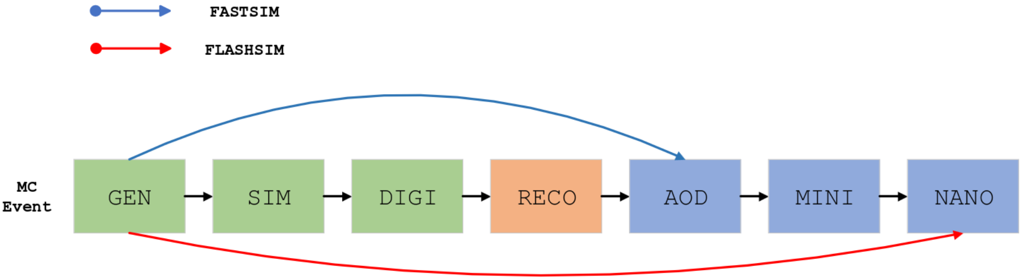
Sources: SIM-2023-04, SIM-2024-04. AtlFast3 paper: Comput Softw Big Sci 6, 7 (2022).

End-to-end fast simulation



CMS's FlashSim

- Event generator to nano-aod.
- Analysis-agnostic.

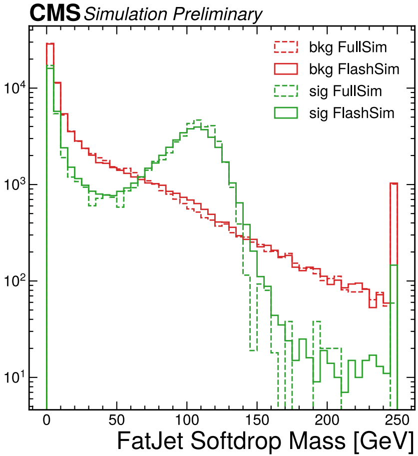


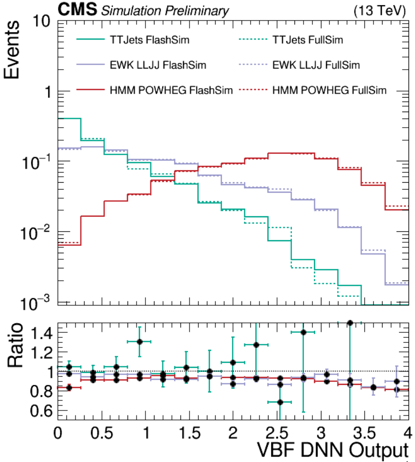
- Generative model: normalizing flow.
- One model per physics object (jet, e . . .), ran in a chain to capture correlations.

Source: CMS-NOTE-2023-003, see also: EPJ Web of Conferences 295, 09020 (2024).

CMS's FlashSim: Performance

Early results: good object-level (left), and object correlations modelling (right: VBF H$\rightarrow \mu\mu$). Up to kHz event generation.

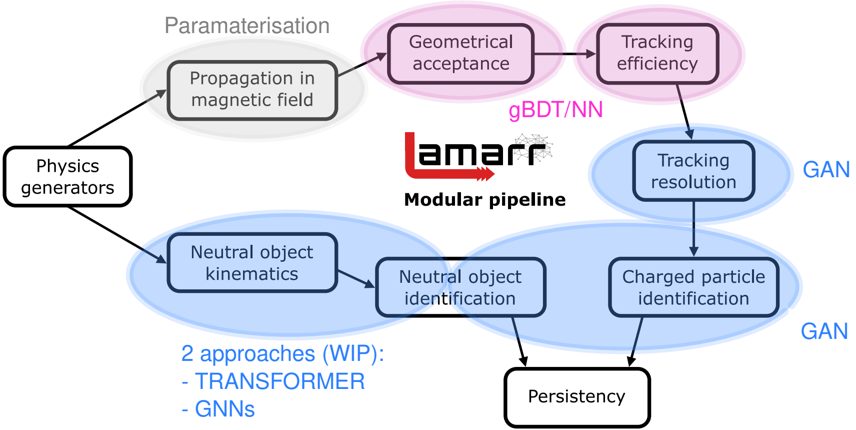




Source: CMS-NOTE-2023-003, see also: EPJ Web of Conferences 295, 09020 (2024).

LHCb's Lamarr

- Pipeline of modules parameterizing detector response & reconstruction.
- Several obtained by generative models.

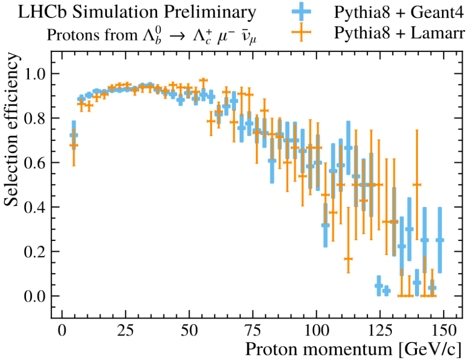


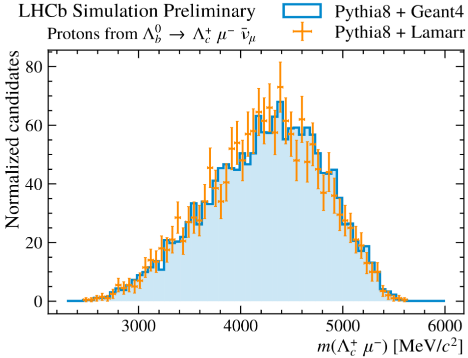
Source: arXiv:2309.13213, see also: arXiv:2303.11428.

LHCb's Lamarr : Performance

Validation:

- good results for PID (left) and analysis-level reconstructed quantities.
- Figures: $\Lambda_b^0 \to \Lambda_c^+ \mu^- X$, with $\Lambda_c^+ \to pK^-\pi^+$.





Achieves O(100) speed-up compared to GEANT-based chain.

Sources: LHCB-FIGURE-2022-014, PoS ICHEP2022 (2022) 233

Cross-experiment software

Case for cross-experiment software

- **Problem:** develop generative fast sim. of detector X of collider experiment Y.
- **Question:** is there cross-experiment software I can use? **Answer:** yes!

LHC fast simulations are solving **similar problems** with **similar techniques** - GANs, VAEs, NormalisingFlows. Compare to industry:

Diffusion (Images)

DALL·E 2

Imagen

stability.ai Stabe Diffusion

Autoregression (Text)

GPT-4

Bard

Meta

Stanford Alpaca

GANs, VAEs, NFs not favoured by industry.

Gaussino core simulation software

- Gaussino: simulation framework with only experiment-independent components.
- Deployed by LHCb as Gauss-on-Gaussino; Gauss: LHCb-specific simulation sw components, based on Gaussino core functionality.

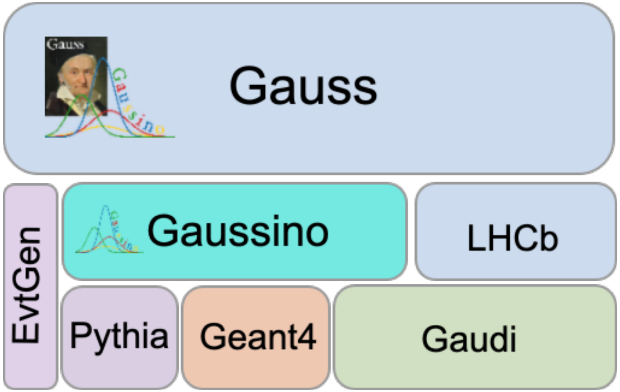
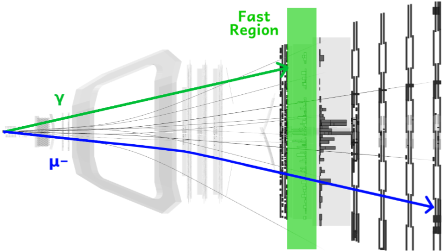
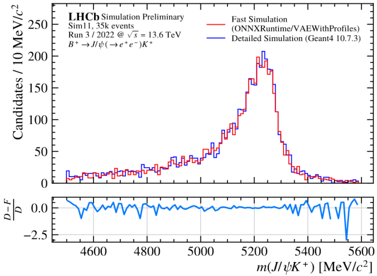


Image source: https://doi.org/10.22323/1.414.0225

Gaussino for generative simulation

- *Problem: develop generative fast sim. of detector X of collider experiment Y.*
- X = EM calorimeter, Y = LHCb.
- Gaussino enables use of CaloChallenge as starting point.
- Experiment-agnostic: train & choose model on CaloChallenge data, ML model services, interface fast simulation with GEANT, distributed computing ...
- LHCb-specific: re-train model (VAE), plug in at right position.



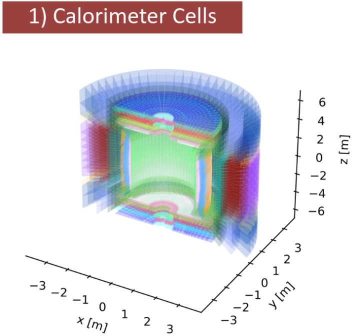


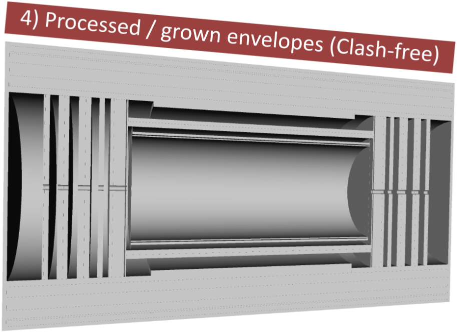
Sources: CHEP2023, ACAT2024. See also: arXiv:2207.06329, arXiv:1812.01319, EPJ Web of Conf. 214, 02040 (2019).

Simplified detector geometry

- *Problem: develop generative fast sim. of detector X of collider experiment Y.*
- X = inner tracker, Y = ATLAS.
- Want: configurable simplified geometry which is clash-free.

Available in pygeosimplify - initially developed for ATLAS fast calo sim.

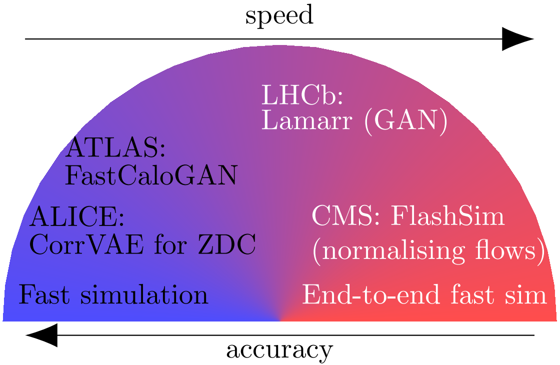




Source: J. Beirer's at ACAT2024, pygeosimplify.

Summary

Snapshot of fast simulation with generative models we've discussed:



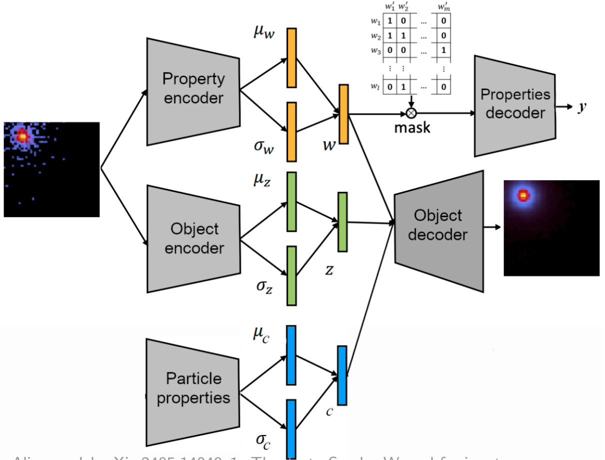
- More need for speed and accuracy $\Rightarrow$ more development needed.
- Developments will benefit from common software, including experiment-agnostic simulation software by LHC experiments.

See also: https://g4fastsim.web.cern.ch/


Extra

ALICE ZDC CorrVAE Model

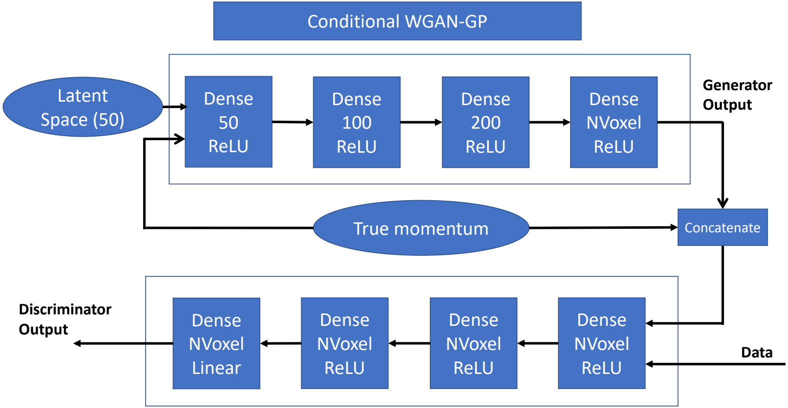
- CorrVAE model - VAE with orthogonal latent spaces (w,z).
- Additional latent space (c) conditions the response acc. to particle data.



Sources: CorrVAE: arXiv:2210.01796, Alice model arXiv:2405.14049v1. Thanks to Sandro Wenzel for inputs.

ATLAS FastCaloGAN Model

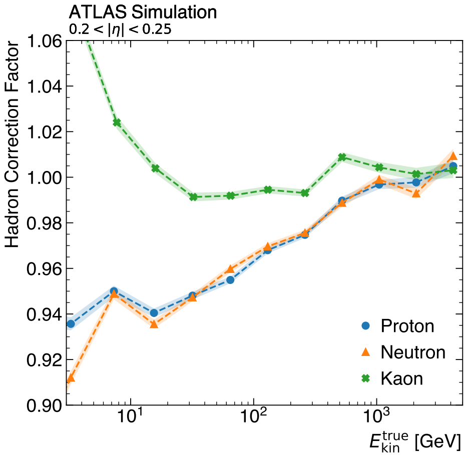
- 300 GANs: one per particle type (3) and per $|\eta|$ slice (100).
- Pre-trained for fixed energy (E=32 GeV), subsequently trained for other E.
- Training time per GAN: 8h on NVIDIA V100, for 1M epochs.

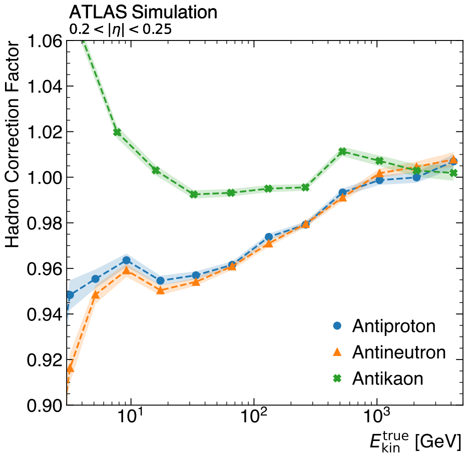


Source: AtlFast3 paper: Comput Softw Big Sci 6, 7 (2022). Nb: Run3 FastCaloSim uses an improved model.

ATLAS FastCaloSim: Corrections

- Hadrons not used in GAN training use correction factors (figures).
- Correction factors also deployed to improve energy resolution.





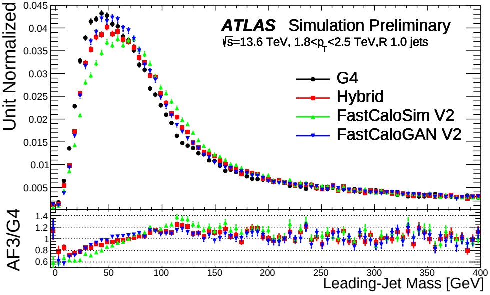
Source: AtlFast3 paper: Comput Softw Big Sci 6, 7 (2022). Nb: Run3 FastCaloSim uses dedicated GANs for protons.

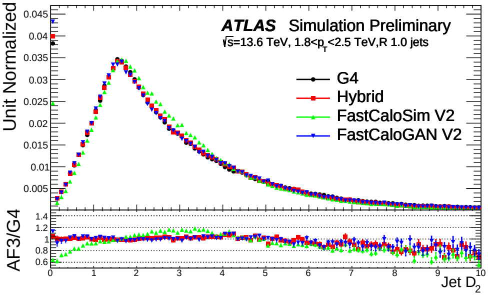
ATLAS FastCaloSim: Performance

Calorimeter fast simulation is a hybrid of two approaches:

- FastCaloSim: parametrised model.
- FastCaloGAN: trained to reproduce voxels and energies in layers as well as total energy in single step.

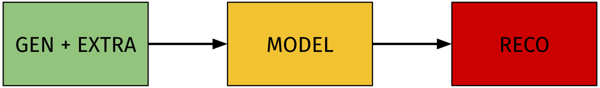
Details of high-pt large-R jet sub-structure remain challenging.





Source: SIM-2023-004, further plots in AtlFast3 paper: Comput Softw Big Sci 6, 7 (2022).

CMS's FlashSim: Inputs & Target



Inputs (Gen level):

Gaussian Noise +

- kinematics
- other relevant information
  - flavor of jet
  - gen status flag: isTauDecayProduct, isPrompt, etc...
- Pile-up vars: Pileup_nTrueInt
- physics-informed information:
  - closest gen muon
  - closest gen jet

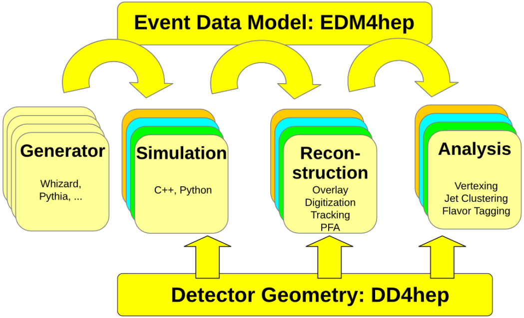
Targets (Reco level):

- 4-vector
- other relevant information
  - tagging of Jets
  - PF Isolation
  - DISCRETE status flag: isGlobal, isPFCand, etc...
  - DISCRETE distributions: nConstituents, nCharged...

Source: F. Vaselli's talk at HSF Simulation meeting

Gaussino in Key4Hep

- *Problem: develop generative fast sim. of detector X of collider experiment Y.*
- X = EM calorimeter, Y = future collider experiment, eg CEPC
- Gaussino integrated as simulation model in Key4Hep software stack.
- Key4Hep: reduce overhead by sharing common sw components, preserving and integrating functionality of iLCSoft, FCCSW, CEPCSW.



Source: CHEP2023.