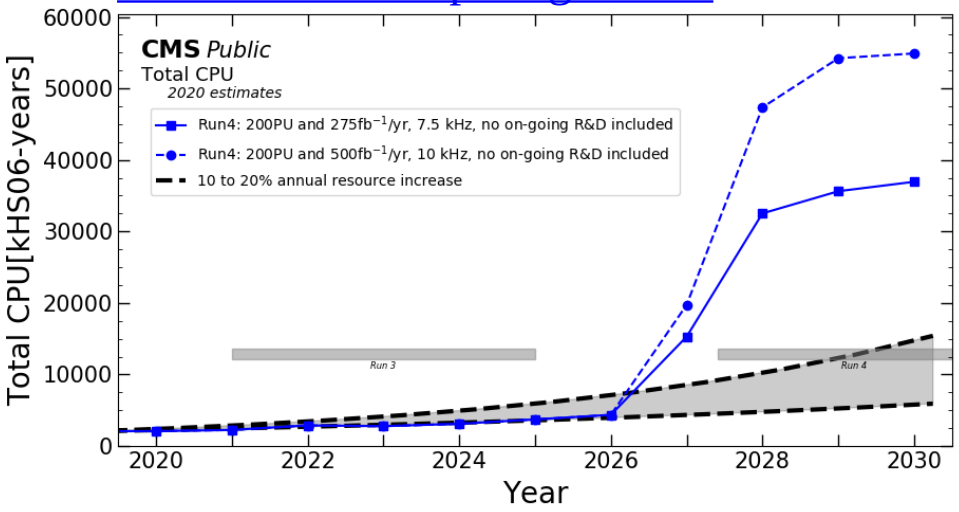
Machine Learning for Detector Simulations

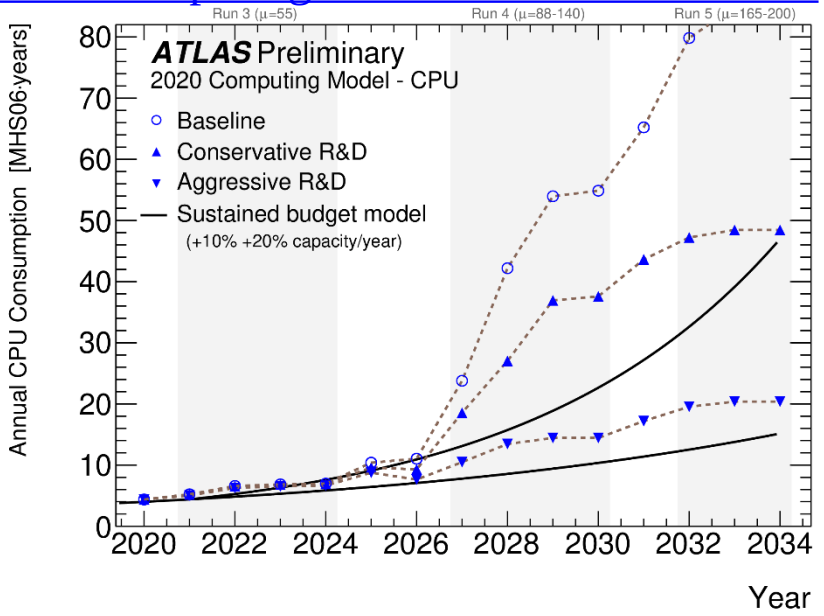
Kevin Pedro (FNAL)
[CMS ML4Sim Convener]
October 7, 2020

Motivation

CMSOfflineComputingResults



AtlasComputingandSoftwarePublicResults



- Beginning of Run 2: full detector simulation (Geant4) took ~40% (plurality) of grid CPU resources for CMS and ATLAS [arXiv:1803.04165]
  - Detector upgrades for HL-LHC: increased complexity [arXiv:2004.02327]
- Reconstruction CPU usage scales superlinearly with pileup

➢ Simulation needs to deliver more events w/ more complexity while using smaller fraction of CPU

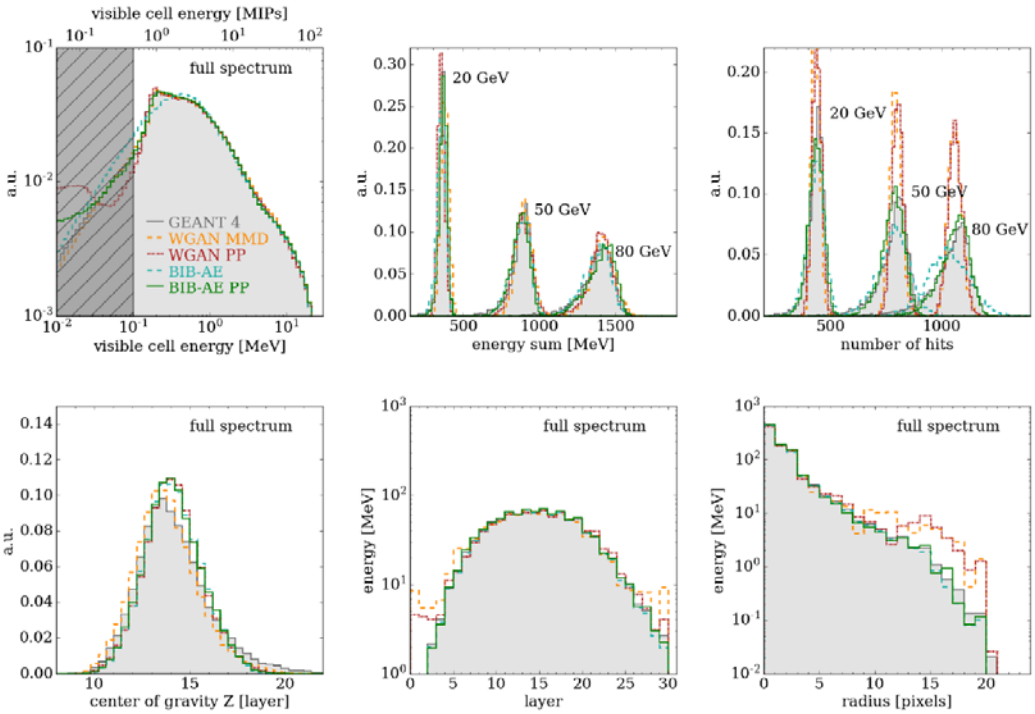
  - Similar challenges in other frontiers (neutrino, cosmic, …)

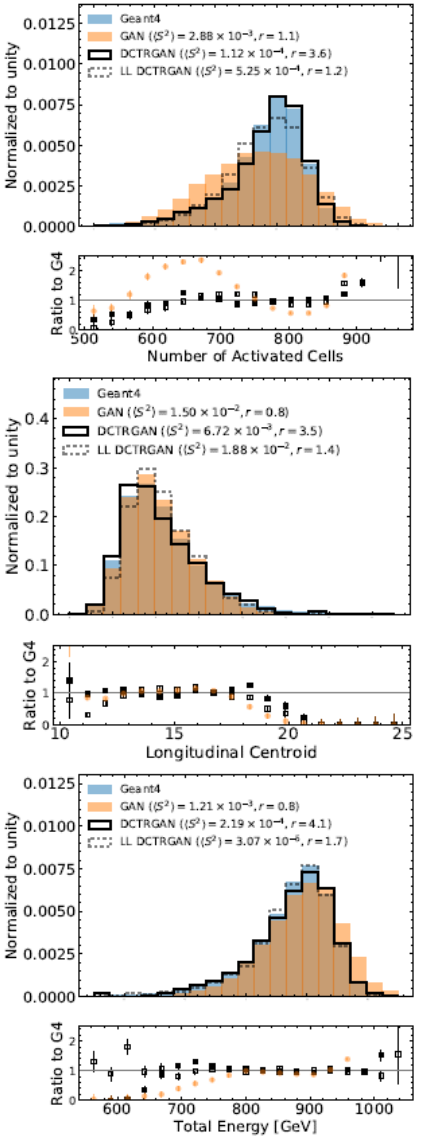
- ML may be able to surpass limitations of technical improvements to existing simulation software/algorithms/etc.

Perspective

- Physics fidelity is *crucial*
  - Without this, quoted speedups are *meaningless*
- Basic GANs pose convergence and reliability concerns
  - If massive training datasets are needed, net CPU benefits may be reduced
  - Learning more about how to improve GAN architectures and training
- Other approaches are worth considering:
  - (V)AEs, CNNs & GNNs, FCNs (regression), new architectures
- ML in detector simulation provides natural avenue to utilize *heterogeneous computing resources* (including HPCs)
  - e.g. inference as a service with GPUs, FPGAs, etc.
- Need to balance *tradeoff*: exploration/novelty vs. production-readiness
  - Limited-author papers are nice
  - Collaborations need solutions implemented and tested before Run 4 starts
- End-to-end ML (generation → reconstruction): interesting alternative, but even more concerns about reliability, accuracy, resources for training…

Examples





Optimized autoencoder architecture with post-processing network better than (some) GANs [arXiv:2005.05334]

But some new techniques can improve GAN results significantly [arXiv:2009.03796]

Many more projects on arXiv or in development!

Backup

Glossary

- GAN: Generative Adversarial Network
- VAE: Variational AutoEncoder
- CNN: Convolutional Neural Network
- GNN: Graph Neural Network
- FCN: Fully Connected Network

- CPU: Central Processing Unit
- GPU: Graphics Processing Unit
- FPGA: Field-Programmable Gate Array
- HPC: High Performance Computing