Geant4 for Future Lepton Colliders

Alberto Ribon
CERN EP-SFT

ECFA Workshop on Simulation, 1-2 February 2022

Geant4 in a Nutshell

- Monte Carlo particle transportation toolkit
  - ~28 years of development, ~130 contributors (~30 FTE, ~40% high-energy), ~1.7 M lines of code, ~20’000+ paper citations
- Used for designing experiments and for data analysis
- Multi-disciplinary applications
  - HEP, space science, medical physics, engineering, nuclear physics, *etc.*
- Used in production by LHC experiments
  - ...and in several other mission-critical projects, in various domains
- Innovative and flexible from the outset
  - From OO C++ to multi-threading and now to tasking mechanism
- Large, growing testing and physics validation suite
  - Major effort, essential for the quality and success of Geant4

Geant4 Road Map

- Provide **better simulations**, in all aspects
  - More precise; faster; smaller (in memory); more robust; easier to use; capable of more things; covering more use-cases
  - Exploit advances in both computing technologies and physics modeling
- Priorities driven by **user needs**
  - For the EP-SFT Geant4 team, the top priority is the support of CERN projects – in particular the LHC experiments – for simulation
    - Recent examples: extension to charm and bottom hadron nuclear physics (G4 10.7); on-going work to transport light hyper-nuclei and anti-hyper-nuclei (ALICE request)
- Method: calibrate on **thin-**target data, validate on **thick-**target
  - Simulation of electromagnetic and hadronic showers in calorimeters is the most important and challenging aspect (for both physics and computing) for HEP applications

Validation of Simulation

- **Thin-target** data
  - Tuning of models' parameters is done only with these data (*e.g.* NA61/SHINE, NA49, HARP, *etc.*)
- **Thick-target** data (*i.e.* EM & HAD showers in calorimeters)
  - **Calorimeter test-beam data** : cleanest, single-particle
    - No MC Event Generator and simple geometry
    - So far mostly from ATLAS, CMS and CALICE test-beams
  - **Collider data** : complex, multi-particle environment
    - Convolution of generation, pile-up overlay, G4 physics, geometry, digitization

Note: Typical iterative pattern : observe discrepancies in showers; look at relevant thin-target data; try to improve the simulation; validate with test-beam data

Precision, Speed, Systematic Errors

- Precision (accuracy) and speed (event throughput) of simulation contribute to the systematic error of data analyses
  - For a given integrated luminosity and computing budget, the optimal simulation is the fastest one with a precision good enough to not dominate the overall systematic error
- As the experimental data sample (luminosity) increases, both more precise simulations and larger MC samples are needed
  - Sometimes, the speed of simulation is considered uniquely, or as the top priority, forgetting or reducing the role of the other side of the coin
- There is a continuum of possibilities between “full” and “fast” sim
  - It is not binary: many types of alternative models, simplifications, tricks both at the level of Geant4 toolkit and of application

Parallelism Opportunities in Detector Simulations

- Run: a large number of similar and independent events
  - This is the “embarrassingly parallel” part of Monte Carlo simulation
  - **Event-level parallelism** : already solved by multi-threading
- Event: a large number of similar but not fully independent tracks
  - They are independent only once they are created, but they are created randomly during the transport of previously generated particles
  - **Track-level parallelism** : difficult, on-going effort in the last ~8 years
- Challenges of track-level parallelism for generic HEP simulations
  - Stochastic nature of Monte Carlo methods; very complex geometries; a hundred different particle types; energies from TeV down to at least keV; several dozens of different physics models, with non-trivial algorithms; a number of physics tables to access randomly at run-time; event reproducibility; *etc*...

Simulation Challenges for High Precision Experiments

- HL-LHC and future Lepton Colliders will make possible **high-precision measurements**, which require **large samples of very accurate simulations**
  - Otherwise systematic errors would dominate over statistical errors
- A big challenge for both MC Event Generators and Geant4
  - From the physics side: for HEP experiments, the physics of generators is the interesting part, while the physics of Geant4 has to be good enough to guarantee the needed precision of simulated detector response
  - From the computational side: for a fully simulated event, Geant4 dominates the CPU time; but both more complex computations and more generated events are needed to ensure higher order precision...
- For Geant4, we are working in 3 directions: **improvements & optimizations**, **fast simulation**, **R&D**

1. Improvements and Optimizations

- Keep improving the precision of physics models
  - The most precise physics configuration as a “reference”
    - Regardless of its speed; a kind of “calibration line” for simulation!
- Keep improving the computing performance of simulation
  - Everywhere: geometry, transportation, stepping algorithm, physics models, cross sections, user-actions, scoring, *etc.*
  - By: code revision; better algorithms; memory layout; *etc.*
  - Experiment-specific optimizations (thresholds, cuts, options, parameters)
- Keep enlarging (and using!) the testing and validation suite
  - Larger code-coverage; more use-cases; higher statistics
  - Migrate test-beam calorimeter set-ups in our validation repository
    - ***geant-val***.*cern.ch*

2. Fast Simulation

- Improve and extend traditional HEP fast-sim techniques
  - Such as shower parameterizations and shower libraries
    - Mostly electromagnetic showers; to explore also for hadronic showers
  - Make available in Geant4 the best algorithms developed by the experiments, and develop general tools for tuning of parameters
- Investigate novel HEP fast-sim techniques
  - Machine Learning generation of showers
    - Use full and detailed simulation for training
  - Biasing (*a.k.a.* variance reduction) techniques
    - *E.g.* Russian roulette for low-energy neutrons and gammas
    - ...

3. R&D on Accelerators and new Architectures

- Vectorization
  - GeantV prototype: vector transport, to use CPU vector registers more efficiently; concluded with no gains from vectorization observed (*arXiv:2005.00949)*
- Prototype simulation on GPU
  - On-going work (AdePT, Celeritas) to demonstrate :
    - Offload of EM shower simulation in calorimeters to GPUs
      - Favorable environment: many tracks, with few materials and only 3 particle types (e-, e+, γ)
    - Needed step to see if we can achieve any significant speed-up in the simulation of such showers, before considering to port the whole Geant4 on CUDA (huge effort with a priori unclear benefits)
- Looking for benchmarking new promising possibilities
  - *E.g.* Fujitsu A64FX
  - ...

Geant4 Simulations for Future Lepton Colliders (1/2)

- We expect a large convergence of detector simulation needs between HL-LHC and future Lepton Collider experiments
  - For both physics and computing aspects
- Still, it is likely that the following physics aspects will play a bigger role for future Lepton Collider experiments
  - Gamma/lepton-nuclear interactions
    - Large number of gammas, electrons and positrons from beam-beam dynamics
  - Decays of secondary particles produced in the detector
    - Limited number of channels and simple decay treatment in Geant4
  - Rare & higher-order physics processes happening in the detector
    - Detector background that is either not included in Geant4 or not properly modeled

Geant4 Simulations for Future Lepton Colliders (2/2)

- Gamma/lepton-nuclear interactions
  - Recent model improvements (*e.g.* Giant Dipole resonance region)
  - On-going work to provide an alternative approach in Geant4
  - Strong interest from the EIC community, help may come from them
- Decays of secondaries
  - Use external MC generators to handle also the decay of secondaries
    - New example in G4 11.0: *examples / extended / eventgenerator / pythia / py8decayer/* showing how to use Pythia8 for decaying tau, charm and bottom hadrons in Geant4
- **Rare & higher-order physics processes**
  - Some progress over the years… priorities and help are welcome !
    - *E.g.* μ+ μ- pair-production by γ conversion available in G4 10.1 (2014)
      e+ e- pair-production by e- and e+ radiation available in G4 10.3 (2016)
      τ+ τ- pair-production by e+ annihilation available in G4 11.0 (2021)
    - Coming next: positron annihilation into 3 gammas ; bremsstrahlung on atomic electrons ; radiative correction to ionization

A few Suggestions

- Use the latest Geant4 version
  - Benefit from bug fixes, improved physics models, better computing performance
    - See back-up slides for the Geant4 release policy, our quality assurance practice, and how new features are introduced
- Use Geant4 as natively as possible
  - Use reference physics lists – we recommend FTFP_BERT
  - Avoid extra “mini-frameworks” on top of Geant4 when not necessary
    - *E.g.* you can have sophisticated mixing of full and fast simulation, even in the same event, with existing Geant4 interfaces, no need of extra “mini-frameworks” for fast-sim !
- Contribute to the validation of Geant4 with test-beams
  - Not only through the validation results, but by sharing with us the test-beam simulation set-up and experimental data
    - On-going activity (Lorenzo Pezzotti) to integrate in ***geant-val****.cern.ch* useful calorimeter test-beams, old and new, *e.g.* : ATLAS HEC, ATLAS TileCal, CALICE, Dual Readout, *etc.*

Summary & Outlook

- As of today, Geant4 can be used, with reasonable confidence, for designing experiments at future Lepton Colliders
  - To the best of our knowledge, no obvious missing main / leading components or physics processes
- Substantial progress on both physics accuracy and computing performance is needed already for HL-LHC
  - Future Lepton Colliders will largely benefit from these
- Full-sim as much as possible, fast-sim as much as needed
  - Fast simulation will likely play an increasing role, but full simulation will remain essential for some analyses and also for tuning / traning fast-sim
- Let’s start working together for the best detector simulations!
  - In particular on test-beams and rare & higher-order physics processes

Back-up

Geant4 Release Policy

- One public release per year – in December
  - Major releases (*e.g.* 11.0) : user interfaces can be changed
    - User-application code might need to be modified
  - Minor releases (*e.g.* 10.7) : user interfaces are preserved
- A beta "pre-view" release in June
  - To get early feedback from users before the public release of December
- Monthly development snapshots
  - Available on CVMFS, to get early feedback from users
- Patches of public releases
  - Typically a few per year, according to needs/requests; mostly bug-fixes
    - New features are never introduced, but non-trivial corrections can be included

Quality Assurance

- Large, growing testing suite
  - Unit tests
  - Integration tests
  - Reproducibility tests
  - Regression tests
  - Physics validation bechmarks
  - Computing – CPU and memory – benchmarks
  - Sub-set of these tests are run “continuously” for each git merge request; a wider set of tests are run every night (before merging on the master); an even wider set of tests are run for monthly development versions; full set of tests, with higher statistics, are run for public releases & patches

==> *Major investment to ensure production quality of Geant4*

New Features in Geant4

- To preserve its “production” quality, new features in Geant4 are introduced gradually and conservatively
  - Typically, a new feature is first added but not used in physics lists (used only by developers in special tests)
  - Then, a new feature can be switched on in physics lists, but by default it is switched off
  - Sometimes, a new feature is switched on by default only in so-called “experimental” physics lists (*e.g.* FTFP_BERT_TRV) – these are meant for testing, not for large productions
  - Finally, when a reasonable confidence is reached on a new feature, this is switched on by default in the main reference physics lists (*e.g.* FTFP_BERT), but with the possibility to switch it off
  - For a new physics model, older, competitive models are kept for a while, and eventually removed only later on – in a major release – if not useful (even for evaluation of systematic uncertainties)