Detector Simulation Challenges in Fermilab IF Experiments

V. D. Elvira[1], L. J. Fields[1], K. L. Genser[1], R. W. Hatcher[1], T. Junk[1], R. Kutschke[1], P. Lebrun[1], M. Mooney[2], B. Viren[2], …
[1]FNAL, [2]BNL

Introduction

The detector and beam simulation needs of Intensity Frontier (IF) experiments will grow substantially as several large liquid Argon detectors as well as the Muon g-2 [21] and Mu2e [22] experiments come online. These new experiments will continue the pursuit of precision measurements, measure remaining unknown parameters including neutrino CP violation, and open the possibility of observing new physics beyond the standard model. The demand for computer resources by IF experiments is expected to grow over the decades, and the experiments will need to adapt to new programing paradigms and evolving computing technology.

During the last few years, the Geant4 [23] Collaboration has dedicated significant efforts to improving the toolkit's computing performance in order to address HEP community needs. The introduction in Geant4 of event-level multithreading capabilities in 2013 brought significant memory savings for Geant4 applications. While per event time performance did not improve with multithreading, memory use improved significantly. Through code optimization and improvements to the Geant4 engine and physics algorithms, the time performance improvement in the 2010-2015 period has been of the order of 35%, for a simplified calorimeter simulation and a standalone application based on the CMS experiment geometry and magnetic field. Remarkably, the percentage time performance improvement during this period is in the double digits even as the physics models were improved considerably for accuracy, something that comes typically associated with a time performance penalty.

Most of the pressure on Geant4 to improve computing performance originates in the needs of the LHC experiments, which are spending more than 50% of the WLCG [1] grid resources on simulation. Through the life of the High Luminosity LHC (HL-LHC) run, they expect to collect at least 150 times more data than they have so far. Unfortunately, code optimization and transistor density growth are not enough to keep up with the increases in computing needs. Although transistor density growth is more or less keeping up with Moore's law, doubling every couple of years, clock speed has been flat since approximately 2003. Therefore, solutions must be found elsewhere, leveraging the core count growth in multicore machines, using new generation coprocessors, and re-engineering code along the lines of the new programming paradigm based on concurrency and parallel programming. These approaches are expected to yield significant gains in event throughput and should be pursued in parallel with an aggressive effort to improve physics models and understand the systematic uncertainties associated with the Geant4 predictions.

This document discusses some of the challenges related to detector simulations faced by experimenters at Fermilab working on Intensity Frontier experiments.

Challenges for the Intensity Frontier

Properly simulating a neutrino experiment like NOvA [2], MicroBooNE[3], MINERvA[4], or DUNE [5] involves a three-part software stack, the first and last of which are relevant to this paper; the second, event generators such as GENIE, is discussed elsewhere. The first element of the stack concerns the beamline and the neutrino flux prediction. Estimating the neutrino flux is a notorious problem because as weakly interacting particles, neutrinos offer no independent mechanisms for measuring the flux. A central problem is in the difficulty of accurately simulating the hadronic physics of meson production, especially in a thick target. Additionally, the simulated neutrino flux depends very strongly on minute details of the beamline, meaning that neutrino beam geometries require significantly more detail and accuracy than is generally required in detector simulations. Geant4 is the most commonly used software toolkit for this work, although FLUKA and MARS are also used. The final element of the stack is the detector simulation. Again there are numerous challenges in terms of efficiently simulating large numbers of events inside complex detector geometries and in handling subtle physics effects in the traversal of radiation across matter. Neutrino experiments in particular rely heavily on detector simulations to reconstruct neutrino energy, which requires accurate modeling of energy deposition by a variety of particles across a range of energies. Geant4 is the universal solution at this stage.

Muon experiments such as Muon g-2 and Mu2e also face large simulation challenges. Because they are searching for extremely rare effects, they must grapple with very low signal to background ratios and simulating low cross section background processes. Often the physics of these processes is not well understood, introducing large systematic uncertainties. Additionally, the size of the computational problem is a serious challenge, as large simulation runs are required to adequately sample all relevant areas of experimental phase space, even when techniques to minimize the required computations are used. Another aspect related to the background events is the need to simulate the effects of low energy neutrons, which requires large computational resources. For example, Mu2e used a total of ~60x10^6 CPU hours in their CD3 simulation campaign [6]. Geant4 is the primary simulation toolkit for all stages of these experiments.

Although the computing needs of the intensity frontier are currently small compared to the LHC experiments, they are expected to grow substantially in the next decade, and it is imperative that the intensity frontier experiments make use of modern computing architectures and algorithms. This transition will be complicated by the fact that the intensity frontier is fractured into many small experiments, some of which have as few as 50 collaborators, limiting the manpower available for updating detector simulations to use new tools. Efforts such as LArSoft[7] have recognized this challenge and aim to pool resources of many experiments for the benefit of all.

In addition to substantial increases in data size, the intensity frontier faces a need for substantial improvements in the precision of detector and neutrino beam simulations. All experiments rely heavily on Geant4 physics models that were tuned for the needs of the LHC experiments. Geant4-based simulations of neutrino fluxes are accurate at the level of ~30%. Currently, each neutrino experiment must conduct a time-consuming program of data taking and/or incorporation of external data constraints, which can improve the flux accuracy to ~10%. Similar procedures are required to improve and quantify uncertainties on detector simulations. A program of thin and thick target measurements is planned to further improve neutrino flux and detector simulations. When available, this data must be used to tune Geant4. It would be good to have a common effort devoted to this, possibly within Geant4, for the use of all experiments, rather than the experiment-by-experiment way data is currently incorporated into simulations.

Another important aspect of neutrino detector simulations, particularly in liquid Argon, is simulating the creation and propagation of electrons and ions, which is often done outside Geant4. In the case of surface-level detectors, this simulation must include space charge effects, which may impact not only particle propagation in a given event, but can have a cumulative effect across events, modifying the electric field seen by traversing particles depending on the event sequence and the beam intensity. Subsequent signal creation in the sense wires and the digitalization of the analog wire and photodetector signals, although often experiment specific, pose another challenge as it may be computationally intensive, particularly for large detectors like DUNE with many wires and sensors that must record signals continuously. Non-uniformity of the liquid Argon purity (and thus electron lifetime) throughout the TPC may need to be modeled in simulation if observed in data, and this could lead to similar technical and computational difficulties as with the space charge effect.

Of particular need for LArTPC detectors are accurate electrostatic field calculation and drift simulation. Although the calculations do not need to be done often, and do not require as much CPU as the simulations in which they are used, fields calculated at mm precision are needed to produce the average detector response functions which are required to be deconvolved from the digitized waveforms in order to gain a proper measure of the distribution of drifting electrons. Today, performing these calculations using a 2D approximation to the geometry are fairly routine but work is needed to understand if full 3D field calculations will be needed to for valid results. Such calculations are very difficult to do with both fine precision and over large volumes using conventional Finite Element Methods. Boundary Element Method provides promise for better scaling but still requires substantial RAM and CPU (See e.g [11] and [12]). Whether 2D or 3D field responses are used, the drift simulation has to be carefully written and optimized as the required fine-grain and nested convolutions can easily exhaust the RAM and CPU available on typical compute hosts.

How to Meet the Challenges?

Neutrino-Community Efforts

Efforts to address common needs of the liquid Argon experiments including the post Geant4 phase of simulation are undertaken in projects like LArSoft[7] and Qscan[8]. (See e.g. [5] sections 8.3 and 8.4 and references therein.) Detectors with large number of wires and photo-sensors may naturally lend themselves to parallelization techniques, provided that the simulation frameworks are able to handle such cases in an efficient way. Additionally, the Wire Cell Toolkit [9] provides simulation, noise filtering, signal processing and reconstruction algorithms for LArTPC detectors. It includes experimental support for parallel processing following the data flow processing paradigm. More work and input from parallel processing experts is needed to bring this support to maturity.

Data-taking at current experiments, such as MicroBooNE, will greatly inform the simulation needs of the IF community moving forward. This might be most relevant for LArTPC technology as it is the newest and least familiar. For example, if significant non-uniformity of purity in data is observed in MicroBooNE, simulation of non-uniform purity will be a high priority simulation need of future LArTPC experiments.

Geant4 Improvements

Efforts are also underway by the Physics and Detector Simulation (PDS) group at Fermilab to improve the accuracy of Geant4. While much of this effort was motivated by requests of from Intensity Frontier experiments, especially the variation of the model parameters, it will be of use to the entire HEP community. This work includes:

- Validation of models: Continuing the Geant4 validation efforts, Fermilab (with CERN and SLAC) is leading the development of a validation database known as DoSSiER [10] that contains data from experiments measuring particle cross sections (e.g. NA61). It was started as a Geant4 project, but is now also used by GeantV and should be suitable for other MC toolkits such as e.g. GENIE.
- Development of new physics lists: Geant4 physics lists utilized by IF experiments have been largely designed for the needs of LHC experiments. This is problematic for a number of areas of Intensity Frontier Simulation. Geant4 simulations of neutrino beams disagree with existing hadron-production data by up to 40% in some areas of phase space. In light of this, a new physics list ("NuBeam") has been created in Geant4, aimed at meeting the specific needs of neutrino beam simulations. A ShieldingM physics list (a variant of Shielding) was created to address the needs of Mu2e.
- Variation of model parameters: Equally as critical as having an accurate beam or detector simulation is having an ability to quantify that accuracy. Neutrino experiments go to great effort to estimate and propagate uncertainties on Geant4 models to physics measurements. This is currently done experiment-by-experiment, but an initiative within Geant4 was recently begun to provide methods of accessing and changing model parameters. Additionally, the PDS group at Fermilab (with Geant4 collaboration) is developing procedures for tuning these model parameters to data, with the aim of producing uncertainties and covariance matrices on these parameters that can then be propagated to physics measurements [24].

Conclusion

Detector simulations are a critical part of all modern HEP experiments, as essential to our measurements as the detectors themselves. In the near term, the intensity and energy frontier experiments face somewhat different needs -- the LHC experiments must cope with increases in already large data volumes and CPU and memory requirements. The intensity frontier experiments would benefit from the improvements in the speed and efficient memory utilization of the calculations, but also need substantial improvements in simulation of the physics of detectors and neutrino beams. Efforts are already underway within the Geant4, GeantV, LArSoft and other collaborations to address these needs, and it is extremely important to both frontiers that these efforts be continued and expanded. While specific efforts are often done in service of one part of the HEP community, they will benefit the entire community, and should be executed with that fact in mind.

References

[21] http://muon-g-2.fnal.gov
[22] http://mu2e.fnal.gov
[23] Recent developments in Geant4, J. Allison et al., Nucl. Instr. and Meth. A835, (2016),186, http://inspirehep.net/record/1488031
[1] WLCG http://wlcg.web.cern.ch
[2] https://www-nova.fnal.gov
[3] http://www-microboone.fnal.gov
[4] Nucl. Instr. and Meth. A743 (2014) 130, https://arxiv.org/abs/1305.5199, http://minerva.fnal.gov
[5] Long-Baseline Neutrino Facility (LBNF) and Deep Underground Neutrino Experiment (DUNE) Conceptual Design Report, Volume 4 The DUNE Detectors at LBNF. https://arxiv.org/abs/1601.02984
[6] OSG All Hands 2016, R. Culbertson MU2E use of OSG
[7] LArSoft: The Liquid Argon Software Collaboration. http://larsoft.org.
[8] D. Lussi, Study of the response of the novel LAr LEM-TPC detector exposed to cosmic rays and a charged particle beam. PhD thesis, ETH Zürich, 2013.
[9] http://www.phy.bnl.gov/wire-cell/
[10] DoSSiER: Database of Scientific Simulation and Experimental Results, http://inspirehep.net/record/1500313
[11] http://garfield.web.cern.ch/garfield
[12] https://github.com/brettviren/larf
[24] http://inspirehep.net/record/1501398