Computing Challenges for Event Generators

J.T. Childers[1], M. Diefenthaler[2], S. Dytman[3], H. Gallagher[4], S. Gardiner[5],
S. Höche[5], P. Ilten[6], M. M. Kirsanov[7], F. Kling[8], N. Lurkin[9], S. Mrenna[5], and
S. Schuchmann[10]

[1]*Argonne National Laboratory, Lemont, IL, 60439*
[2]*Thomas Jefferson National Accelerator Facility, Newport News, VA 23606, USA*
[3]*Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, PA 15260, USA*
[4]*Department of Physics and Astronomy, Tufts University, Medford, MA 02155, USA*
[5]*Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA*
[6]*Department of Physics, University of Cincinnati, Cincinnati, OH 45221, USA*
[7]*Institute for Nuclear Research, 117312 Moscow, Russia*
[8]*SLAC National Accelerator Laboratory, Menlo Park, CA, 94025, USA*
[9]*Universite Catholique de Louvain, Louvain-la-Neuve, Belgium*
[10]*Johannes Gutenberg University, Mainz, Germany*

August 31, 2020

Event generators (EGs) are essential tools for interpreting data across all particle and nuclear physics frontiers. Based on current expectations and capabilities, these tools will soon become computational bottlenecks in physics analyses. The HEP Software Foundation (HSF) report [1] summarized the situation for ATLAS and CMS in the high luminosity LHC (HL-LHC) era. Here, we expand upon that report and summarize the computational concerns across the broader experimental community, including input from LHC (ATLAS [2], CMS [3], LHCb [4]), forward (FASER [5,6]), transverse (CODEX-b [7]), neutrino (DUNE [8], MicroBooNE [9], MINERvA [10], NOνA [11]), heavy ion (EIC [12]), and beam dump (NA62 [13] and NA64 [14]) experiments. This list is intentionally un-ordered, as all the following issues are priorities.

- **Improve the core EGs for accuracy and efficiency**: the emphasis of the HSF report is on perturbative matrix element calculations, as this is already becoming a computational bottleneck for both ATLAS and CMS. Significant effort has been invested into performing next-to-leading calculations with Herwig [15], MadGraph [16], and Sherpa [17]. However, the EGs will need either new or more efficient algorithms to handle rare processes such as forced fragmentation (LHCb, Belle II [18]), spin and color correlations in parton showers and hadronization (EIC), and matching/merging with higher order calculations (ATLAS, CMS). In some cases the algorithms for these improvements are already known, *e.g.* spin correlations in parton showers, but for others new algorithms must be developed. The inclusion of polarization and weighting schemes in hadronization is particularly challenging, while the complexity of merging algorithms needs to evaluated, also including the mixed effect of EW and QCD corrections.

- **Introduce new computing technologies:** The possibility of parallelization or the employment of specialized accelerators and GPUs for time-consuming portions of EGs (multi-parton interactions, hadronization, correlations) should be investigated. The time commitment for including these techniques is non-negligible, and so it is important that the entire particle physics community is surveyed for use-cases, prior to investing person power into particular techniques. Currently the LHC community utilizes standard CPU technologies in a *pleasingly parallel* paradigm. However, this environment is slowly shifting, *e.g.* the upgraded LHCb experiment will have an entirely GPU based trigger farm [19] that could be used for Monte Carlo generation when data taking is not underway.

- **Modernize and ensure sustainability of code**: for many communities, the core EG code is written in FORTRAN and is difficult to support, given the software expertise in the community. However, rewriting this code requires physics knowledge as well as both legacy and modern software expertise. As an example, the rewrite of PYTHIA [20] from FORTRAN to C++ needed three years of full time effort from a highly qualified expert [21]. Currently, the EIC utilizes a number of FORTRAN EGs developed for HERA but modern improvements from the LHC are not integrated. Similarly, the LHC experiments and the NA62 experiment utilize the FORTRAN based PHOTOS [22] for QED final state radiation, which is not usable in the multi-threaded environment needed for the HL-LHC era. Developing a consistent strategy across the community for ensuring code sustainability will provide both short and long-term benefits.

- **Standardize interfaces between EGs and other programs**: in many cases, specialized tools are required to calculate input for general EGs, and, in other cases, the output of EGs must be coordinated with other programs (for example, BSM interactions that can occur in detector simulations). There are already well defined standards within the LHC community for the former case (LHE [23] and SLHA [24]), but equivalent standards do not exist in the broader community. Establishing similar standards for EGs across the entire particle and nuclear physics community, *e.g.* lepton-nucleus collisions for neutrino physics, will help bring together currently disparate tools.

- **Advance tools for tuning**: while many aspects of EGs are based on first principles, the power of EGs lies in interfacing these calculations with phenomenological models that include tune-able parameters. Current tuning efforts have been concentrated on describing central collisions at the LHC. Even for this case, a full exploration of the many parameters and models has not been performed. Tuning studies are needed for forward particle production models, *e.g.* EPOS [25], PYTHIA, QGSJET [26], and SIBYLL [27,28]. A community-wide effort should be made to advance current tools (such as PROFESSOR [29] and machine learning based approaches [30,31]) and develop a framework to allow smaller collaborations to quickly and efficiently perform bespoke tunes. Such developments will enable physics analyses to:
  - quantify uncertainties in both the central region and forward direction simultaneously,
  - reduce these uncertainties with further tuning and improved models,
  - and determine if forward experiments such as FASER can help further reduce these uncertainties [32, 33].

  The impact of this campaign will be well outside LHC based physics, including critical support for modeling atmospheric neutrino production, *e.g.* IceCube [34]. In the context of neutrino-interaction generators such as GENIE [35], a standardized approach to systematic uncertainty parameterization will provide a consistent treatment of uncertainties between experiments.

- **Develop a strong community**: prior to the current focus on neutrino experiments, significant time for EG development was devoted to LHC applications. While this development continues in the HL-LHC era, it is important that LHC development also be coordinated with neutrino and nuclear EG physics. Many of the community tools developed for the LHC, *e.g.* HepMC [36], Rivet [37], and Professor, could also be used outside their LHC conception. Additionally, physics developments are applicable across the communities: the new parton Dire [38], driven by LHC physics requirements, can also be used in the context of nuclear physics. Similarly, development of generalized and transverse momentum parton distributions (GPDs and TMDs) can be used to help describe soft physics at the LHC. Another example is hadronic rescattering, which is relevant to beam dump experiments, neutrino experiments, and even LHC collisions. Indeed, the observation of the near-side ridge in both proton-proton and heavy ion collisions has already brought high energy and heavy ion models closer together, *i.e.* Angantyr [39].

- **Improve the culture**: software development, documentation, and maintenance of EGs is a critical aspect within particle and nuclear physics, but often is not recognized as such. While citation counts for EGs can be high, these metrics do not necessarily translate to career prospects for students and post-docs, or even funding opportunities for established researchers. Developing a particular aspect of an EG can require both physics and software expertise, along with significant person power. In many cases, this development does not translate to a significant number of publications or even community recognition. Indeed, working on EGs as a PhD student or post-doc can be detrimental to career development. Obtaining funding for EG projects is possible, but oftentimes falls between funding agencies, *e.g.* software for both high energy and nuclear physics, which complicates the situation. Furthermore, service work and maintenance, rather than just development of novel techniques, can be very difficult to support. While steps have been taken to remedy this situation, *e.g.* the Scientific Discovery Through Advanced Computing (SciDAC) program, a further paradigm shift is necessary, both with funding agencies and within academia.

Event generators play a key role in the particle and nuclear physics community, across a wide range of experiments from the LHC, to beam dump experiments, neutrino experiments, and the EIC. A cohesive approach will ensure the commonality between communities is utilized, and the development of novel models that unify different regimes of physics. More than ever, a strong effort with substantial support is required within the community to ensure that event generators are maintained and improved over the next decade.

References

[1] **HSF Physics Event Generator WG** Collaboration, S. Amoroso *et al.*, “Challenges in Monte Carlo event generator software for High-Luminosity LHC,” `arXiv:2004.13687 [hep-ph]`.

[2] **ATLAS** Collaboration, G. Aad *et al.*, “The ATLAS Experiment at the CERN Large Hadron Collider,” *JINST* **3** (2008) S08003.

[3] **CMS** Collaboration, S. Chatrchyan *et al.*, “The CMS Experiment at the CERN LHC,” *JINST* **3** (2008) S08004.

[4] **LHCb** Collaboration, J. Alves, A.Augusto *et al.*, "The LHCb Detector at the LHC," *JINST* **3** (2008) S08005.

[5] **FASER** Collaboration, A. Ariga *et al.*, "Technical Proposal for FASER: ForwArd Search ExpeRiment at the LHC," `arXiv:1812.09139 [physics.ins-det]`.

[6] **FASER** Collaboration, H. Abreu *et al.*, "Detecting and Studying High-Energy Collider Neutrinos with FASER at the LHC," *Eur. Phys. J. C* **80** no. 1, (2020) 61, `arXiv:1908.02310 [hep-ex]`.

[7] G. Aielli *et al.*, "Expression of Interest for the CODEX-b Detector," `arXiv:1911.00481 [hep-ex]`.

[8] **DUNE** Collaboration, R. Acciarri *et al.*, "Long-Baseline Neutrino Facility (LBNF) and Deep Underground Neutrino Experiment (DUNE): Conceptual Design Report, Volume 1: The LBNF and DUNE Projects," `arXiv:1601.05471 [physics.ins-det]`.

[9] **MicroBooNE** Collaboration, R. Acciarri *et al.*, "Design and Construction of the MicroBooNE Detector," *JINST* **12** no. 02, (2017) P02017, `arXiv:1612.05824 [physics.ins-det]`.

[10] **MINERvA** Collaboration, L. Aliaga *et al.*, "Design, Calibration, and Performance of the MINERvA Detector," *Nucl. Instrum. Meth. A* **743** (2014) 130–159, `arXiv:1305.5199 [physics.ins-det]`.

[11] **NOvA** Collaboration, D. Ayres *et al.*, "NOvA: Proposal to Build a 30 Kiloton Off-Axis Detector to Study $\nu_\mu \to \nu_e$ Oscillations in the NuMI Beamline," `arXiv:hep-ex/0503053`.

[12] A. Accardi *et al.*, "Electron Ion Collider: The Next QCD Frontier: Understanding the glue that binds us all," *Eur. Phys. J. A* **52** no. 9, (2016) 268, `arXiv:1212.1701 [nucl-ex]`.

[13] **NA62** Collaboration, E. Cortina Gil *et al.*, "The Beam and detector of the NA62 experiment at CERN," *JINST* **12** no. 05, (2017) P05025, `arXiv:1703.08501 [physics.ins-det]`.

[14] S. Andreas *et al.*, "Proposal for an Experiment to Search for Light Dark Matter at the SPS," `arXiv:1312.3309 [hep-ex]`.

[15] J. Bellm *et al.*, "Herwig 7.0/Herwig++ 3.0 release note," *Eur. Phys. J. C* **76** no. 4, (2016) 196, `arXiv:1512.01178 [hep-ph]`.

[16] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations," *JHEP* **07** (2014) 079, `arXiv:1405.0301 [hep-ph]`.

[17] **Sherpa** Collaboration, E. Bothmann *et al.*, "Event Generation with Sherpa 2.2" *SciPost Phys.* **7** no. 3, (2019) 034, `arXiv:1905.09127 [hep-ph]`.

[18] **Belle-II** Collaboration, W. Altmannshofer *et al.*, "The Belle II Physics Book," *PTEP* **2019** no. 12, (2019) 123C01, `arXiv:1808.10567 [hep-ex]`. [Erratum: PTEP 2020, 029201 (2020)].

[19] R. Aaij *et al.*, "Allen: A high level trigger on GPUs for LHCb," *Comput. Softw. Big Sci.* **4** no. 1, (2020) 7, `arXiv:1912.09161 [physics.ins-det]`.

[20] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, "An introduction to PYTHIA 8.2" *Comput. Phys. Commun.* **191** (2015) 159–177, `arXiv:1410.3012 [hep-ph]`.

[21] T. Sjöstrand, "The PYTHIA Event Generator: Past, Present and Future," *Comput. Phys. Commun.* **246** (2020) 106910, `arXiv:1907.09874 [hep-ph]`.

[22] E. Barberio and Z. Was, "PHOTOS: A Universal Monte Carlo for QED radiative corrections. Version 2.0" *Comput. Phys. Commun.* **79** (1994) 291–308.

[23] J. Alwall *et al.*, "A Standard format for Les Houches event files," *Comput. Phys. Commun.* **176** (2007) 300–304, `arXiv:hep-ph/0609017`.

[24] B. Allanach *et al.*, "SUSY Les Houches Accord 2," *Comput. Phys. Commun.* **180** (2009) 8–25, `arXiv:0801.0045 [hep-ph]`.

[25] T. Pierog, I. Karpenko, J. M. Katzy, E. Yatsenko, and K. Werner, "EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider," *Phys. Rev.* **C92** (2015) 034906, `arXiv:1306.0121 [hep-ph]`.

[26] S. Ostapchenko, "Monte Carlo treatment of hadronic interactions in enhanced Pomeron scheme: I. QGSJET-II model," *Phys. Rev. D* **83** (2011) 014018, `arXiv:1010.1869 [hep-ph]`.

[27] A. Fedynitch, F. Riehn, R. Engel, T. K. Gaisser, and T. Stanev, "Hadronic interaction model sibyll 2.3c and inclusive lepton fluxes," *Phys. Rev.* **D100** no. 10, (2019) 103018, `arXiv:1806.04140 [hep-ph]`.

[28] F. Riehn, R. Engel, A. Fedynitch, T. K. Gaisser, and T. Stanev, "The hadronic interaction model Sibyll 2.3c and extensive air showers," `arXiv:1912.03300 [hep-ph]`.

[29] A. Buckley, H. Hoeth, H. Lacker, H. Schulz, and J. E. von Seggern, "Systematic event generator tuning for the LHC," *Eur. Phys. J. C* **65** (2010) 331–357, `arXiv:0907.2973 [hep-ph]`.

[30] A. Andreassen and B. Nachman, "Neural Networks for Full Phase-space Reweighting and Parameter Tuning," *Phys. Rev. D* **101** no. 9, (2020) 091901, `arXiv:1907.08209 [hep-ph]`.

[31] P. Ilten, M. Williams, and Y. Yang, "Event generator tuning using Bayesian optimization," *JINST* **12** no. 04, (2017) P04028, `arXiv:1610.08328 [physics.data-an]`.

[32] **FASER** Collaboration, "FASER 2: Forward Search Experiment at the HL LHC." `https://www.snowmass21.org/docs/files/summaries/EF/SNOWMASS21-EF9_EF6-NF3_NF6-RF6_RF0-CF7_CF0-AF5_AF0_FASER2-038.pdf`.

[33] **FASER** Collaboration, "FASERν 2: A Forward Neutrino Experiment at the HL LHC." `https://www.snowmass21.org/docs/files/summaries/NF/SNOWMASS21-NF10_NF6-EF6_EF9-IF0_FASERnu2-006.pdf`.

[34] **IceCube** Collaboration, M. Aartsen *et al.*, "IceCube-Gen2: A Vision for the Future of Neutrino Astronomy in Antarctica," `arXiv:1412.5106 [astro-ph.HE]`.

[35] C. Andreopoulos, C. Barry, S. Dytman, H. Gallagher, T. Golan, R. Hatcher, G. Perdue, and J. Yarba, "The GENIE Neutrino Monte Carlo Generator: Physics and User Manual," `arXiv:1510.05494 [hep-ph]`.

[36] A. Buckley, P. Ilten, D. Konstantinov, L. Lönnblad, J. Monk, W. Porkorski, T. Przedzinski, and A. Verbytskyi, "The HepMC3 Event Record Library for Monte Carlo Event Generators," `arXiv:1912.08005 [hep-ph]`.

[37] C. Bierlich *et al.*, "Robust Independent Validation of Experiment and Theory: Rivet version 3," *SciPost Phys.* **8** (2020) 026, `arXiv:1912.05451 [hep-ph]`.

[38] S. Höche and S. Prestel, "The midpoint between dipole and parton showers," *Eur. Phys. J. C* **75** no. 9, (2015) 461, `arXiv:1506.05057 [hep-ph]`.

[39] C. Bierlich, G. Gustafson, L. Lönnblad, and H. Shah, "The Angantyr model for Heavy-Ion Collisions in PYTHIA8," *JHEP* **10** (2018) 134, `arXiv:1806.10820 [hep-ph]`.