Snowmass 2021 Letter of Interest
Multi-scale QCD algorithms Beyond the Exascale

Richard C. Brower*, C.E. DeTar+, Kate Clark†, and Evan Weinberg†

*Department of Physics and Center for Computational Science, Boston University
†NVIDIA Corporation, Santa Clara, CA 95051, USA
+Department of Physics and Astronomy, Univ of Utah, Salt Lake City, UT 84112, USA

Introduction: [1]Through sustained advances in computer hardware and algorithms for nearly 50 years, now entering the Exascale Era, lattice QCD (LQCD) has become an indispensable high-precision tool in the arsenal of methods for exploring high-energy physics (HEP) beyond the Standard Model. With a nearly $O(1000)$-fold increase in HPC hardware each decade, matched roughly by advances in algorithms, finer and finer scales are revealed, increasing the advantages of multi-scale algorithms.

Indeed, the birth of QCD was intimately connected to multi-scale studies (aka `renormalization group`) introduced by Kenneth Wilson[1] in 1970, which bears a striking similarity to the coarse-grid projection of a multigrid solver. Applying scaling to non-Abelian gauge theories in 1973, Gross and Wilczek [2] suggested that QCD is a finite and exact quantum theory of the nuclear force. Next year in 1974, Kenneth Wilson[3] formulated the lattice-QCD computational paradigm to explain quark confinement with early numerical simulations[4] in 1979 demonstrating how computations could support this remarkable phenomena.

Multigrid Solver Example: How did Multigrid QCD algorithms get started? The classic multigrid problem of the free Laplacian is scale invariant (i.e. conformal) leading to a simple blocking scheme. With QCD's near scale invariance at short distances, early attempts in the 1990's applied these scaling ideas to a multigrid Dirac solver with modest success[5]. After a long period of frustration, a collaboration with applied mathematicians in 2010 produced a successful adaptive Galerkin multigrid for the Wilson discretization [6]. In retrospect, this adaptive blocking scheme may be viewed in the spirit of Wilson's original non-perturbative renormalized lattice QCD action. Only recently have multigrid algorithms been proposed for other two Staggered [7] and Domain Wall [8, 9] Dirac discretizations used most extensively in HEP lattice simulations. Surprisingly in spite of the equivalence of all 3 discretization in the continuum limit, each example required substantial changes to the Galerkin projectors.

The final step in proving the efficacy of each algorithm requires an optimized, fully parallel software implementation at scale, which in this era must be done in the context of heterogeneous

---

[1]**Topical Groups:**
(CompF2) Theoretical Calculations and Simulations and (TF05) Lattice Gauge Theory
Contact Information:
Richard Brower (Boston University) [email] brower@bu.edu

hardware. To accomplish this, each algorithm is built and tuned using the QUDA [10, 11], Chroma [12], and Grid [13] frameworks under active development in the Exascale project. This development pipeline from algorithmic design by physicists and mathematicians to testing at scale and, finally, to hardened software integrated into application codes takes, at best, several years. Multigrid is of course only one of several critical kernels in the full QCD code stack, as summarized in the report on `QCD in the transition to Exascale and beyond` [14].

After the initial discovery of a new algorithm, testing it at scale is often a source of delay. Asymptotically scaling theories are irrelevant unless the new algorithm can be proven to improve performance at scale on current hardware and even more so projected forward. A huge advantage at present to the Exascale team is the strong contributions of two theoretical physicists, Kate Clark and Evan Weinberg, former postdoctoral fellows at Boston University who are now full-time software engineers at NVIDIA with lead roles in developing the QUDA library for GPUs. The evolution of ideas and implementations thereof within the LQCD community demonstrates the huge advantage of a diverse team spanning physics, mathematics and engineering disciplines all of whom have sustained a collaboration over enough time to rapidly communicate across the disciplinary barriers.

Proposal: It is crucial to have a transition plan to support algorithmic research when the ECP support ends in 2023 and the first Exascale platforms are available. The experience of computing at the Exascale will stimulate the next round of multi-scale algorithm and software development. There are still more key kernels that are part of the QCD workflow which can benefit from multi-scale approaches, including the Hamiltonian integrator (aka Hybrid Monte Carlo, or HMC) used to generate lattice ensembles and the systematic generation of complex multi-quarks correlators. These challenges, among others, still need to be adressed. Outside the context of QCD, algorithms need to be redesigned for the study of strong dynamics beyond QCD near to the conformal limit[15], even eventually including curved manifolds in de-Sitter[16] and Anti-de-Sitter[17] space. Opportunities for multi-scale acceleration include:

1. Solving critical slowing down for Hybrid Monte Carlo (HMC) evolution [18, 19].
2. Eigenvalue deflation and data compression [20].
3. Renormalized actions, perhaps using Machine Learning[21].
4. Recursive methods for improved multi-quark operator bases.
5. Rapid equilibration to initiate the HMC Markov chain.

While each area may require separate methods in detail, the underlying mathematical structure and physical insight benefit from a broad and collective effort. To sustain this fundamental support for advances in lattice field theory, stable funding is required and the development of career paths for young recruits with new skills in the increasingly complex computational environment, whether it be machine learning or autotuning, tied to advanced computer technology. This all requires individuals who are able to participate in an interdisciplinary environment, quickly responding to the rapid evolution of modern software and applications codes on the highest-performing hardware platforms available.

The importance of maintaining a skilled workforce and recruiting students via the University pipeline was considered by "ECP/Transition" subcommittee of the Advanced Scientific Computing Advisory Committee (ASCAC), chaired by Dr. Roscoe Giles of Boston University. That committee produced a draft report in Spring 2020 [22]. See Software LOI for more details.

References

[1] K. G. Wilson, "The Renormalization Group and Strong Interactions," Phys. Rev. D **3**, 1818 (1971). doi:10.1103/PhysRevD.3.1818

[2] D. J. Gross and F. Wilczek, "Ultraviolet Behavior of Nonabelian Gauge Theories," Phys. Rev. Lett. **30**, 1343 (1973). doi:10.1103/PhysRevLett.30.1343

[3] K. G. Wilson, "Confinement of Quarks," Phys. Rev. D **10**, 2445 (1974). doi:10.1103/PhysRevD.10.2445

[4] M. Creutz, L. Jacobs and C. Rebbi, "Monte Carlo Study of Abelian Lattice Gauge Theories," Phys. Rev. D **20**, 1915 (1979). doi:10.1103/PhysRevD.20.1915

[5] R. C. Brower, R. G. Edwards, C. Rebbi and E. Vicari, "Projective multigrid for Wilson fermions," Nucl. Phys. B **366**, 689 (1991). doi:10.1016/0550-3213(91)90035-V

[6] R. Babich, J. Brannick, R. C. Brower, M. A. Clark, T. A. Manteuffel, S. F. McCormick, J. C. Osborn and C. Rebbi, "Adaptive multigrid algorithm for the lattice Wilson-Dirac operator," Phys. Rev. Lett. **105**, 201602 (2010) doi:10.1103/PhysRevLett.105.201602 [arXiv:1005.3043 [hep-lat]].

[7] R. C. Brower, M. A. Clark, A. Strelchenko and E. Weinberg, "Multigrid algorithm for staggered lattice fermions," Phys. Rev. D **97**, no. 11, 114513 (2018) doi:10.1103/PhysRevD.97.114513 [arXiv:1801.07823 [hep-lat]].

[8] R. C. Brower, M. A. Clark, D. Howarth and E. S. Weinberg, "Multigrid for Chiral Lattice Fermions: Domain Wall," arXiv:2004.07732 [hep-lat].

[9] A. Yamaguchi and P. Boyle, "Hierarchically deflated conjugate residual," PoS LATTICE **2016**, 374 (2016) doi:10.22323/1.256.0374 [arXiv:1611.06944 [hep-lat]].

[10] R. Babich, R. Brower, M. Clark, S. Gottlieb, B. Joo and G. Shi, "Progress on the QUDA code suite," PoS LATTICE **2011**, 033 (2011). doi:10.22323/1.139.0033

[11] M. A. Clark, B. Joó, A. Strelchenko, M. Cheng, A. Gambhir and R. Brower, "Accelerating Lattice QCD Multigrid on GPUs Using Fine-Grained Parallelization," Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (SC '16), Article 68 (November, 2016) [arXiv:1612.07873 [hep-lat]].

[12] R. G. Edwards *et al.* [SciDAC and LHPC and UKQCD Collaborations], "The Chroma software system for lattice QCD," Nucl. Phys. Proc. Suppl. **140**, 832 (2005) doi:10.1016/j.nuclphysbps.2004.11.254 [hep-lat/0409003].

[13] D. Richtmann, P. A. Boyle and T. Wettig, PoS LATTICE **2018**, 032 (2019) doi:10.22323/1.334.0032 [arXiv:1904.08678 [hep-lat]].

[14] B. Joó *et al.* [USQCD Collaboration], "Status and Future Perspectives for Lattice Gauge Theory Calculations to the Exascale and Beyond," Eur. Phys. J. A **55**, no. 11, 199 (2019) doi:10.1140/epja/i2019-12919-7 [arXiv:1904.09725 [hep-lat]].

[15] T. Appelquist *et al.* [Lattice Strong Dynamics Collaboration], arXiv:2007.01810 [hep-ph].

[16] R. C. Brower, G. T. Fleming, A. D. Gasbarro, D. Howarth, T. G. Raben, C. I. Tan and E. S. Weinberg, arXiv:2006.15636 [hep-lat].

[17] R. C. Brower, C. V. Cogburn, A. L. Fitzpatrick, D. Howarth and C. I. Tan, arXiv:1912.07606 [hep-th].

[18] G. Cossu, P. Boyle, N. Christ, C. Jung, A. Jüttner and F. Sanfilippo, EPJ Web Conf. **175**, 02008 (2018) doi:10.1051/epjconf/201817502008 [arXiv:1710.07036 [hep-lat]].

[19] X. Y. Jin and J. C. Osborn, "Ensemble Quasi-Newton HMC," PoS LATTICE **2018**, 027 (2019) doi:10.22323/1.334.0027 [arXiv:1904.10039 [hep-lat]].

[20] M. A. Clark, C. Jung and C. Lehner, EPJ Web Conf. **175**, 14023 (2018) doi:10.1051/epjconf/201817514023 [arXiv:1710.06884 [hep-lat]].

[21] P. E. Shanahan, D. Trewartha and W. Detmold, "Machine learning action parameters in lattice quantum chromodynamics," Phys. Rev. D **97**, no. 9, 094506 (2018) doi:10.1103/PhysRevD.97.094506 [arXiv:1801.05784 [hep-lat]].

[22] ECP/Transition Subcommittee of the Advanced Scientific Computing Advisory Committee, "Draft ASCR/ECP Transition Report", (April, 2020).