PROCEEDINGS OF SCIENCE

Panel Discussion on Future of Computing

Yasumichi Aoki,[a,*] **Marcello Dalmonte,**[b,*] **Zohreh Davoudi,**[c,d,*] **Steven Gottlieb,**[e,*] **Stefan Krieg**[f,g,*] **and Robert Mawhinney**[h,*]

[a] *RIKEN Center for Computational Science (R-CCS), 7-1-26, Minatojima Minamimachi, Kobe 650-0047, Japan*

[b] *The Abdus Salam International Center for Theoretical Physics (ICTP), Strada Costiera 11, Trieste I-34151, Italy*

[c] *Maryland Center for Fundamental Physics and Department of Physics, University of Maryland, College Park MD 20742, USA*

[d] *Joint Center for Quantum Information and Computer Science, University of Maryland, College Park MD 20742, USA*

[e] *Department of Physics, Indiana University, Bloomington, Indiana 47405, USA*

[f] *Jülich Supercomputing Center and Center for Advanced Simulation and Analytics (CASA), Forschungszentrum Jülich, 52428 Jülich, Germany*

[g] *Helmholtz-Institut für Strahlen- und Kernphysik, Rheinische Friedrich-Wilhelms-Universität Bonn, 53115 Bonn, Germany*

[h] *Department of Physics, Columbia University, 538 W 120th Street, New York, NY, 10027, USA*

E-mail: yasumichi.aoki@riken.jp, mdalmont@ictp.it, davoudi@umd.edu, sg@iu.edu, s.krieg@fz-juelich.de, rdm10@columbia.edu

We summarize the panel discussion on the future of computing held at LATTICE 2025 in Mumbai, India. The panel brought together experts in lattice gauge theory, high-performance computing, machine learning, and quantum computing to assess emerging opportunities and challenges for lattice field theory research. Topics included next-generation supercomputing architectures, GPU-accelerated systems, AI-driven hardware trends, quantum simulation of gauge theories, hybrid classical-quantum workflows, resource sharing, and workforce training. The panel emphasized that future progress in lattice gauge theory will depend on continued algorithmic innovation, close collaboration with hardware vendors and ML/AI and quantum-science communities, and flexible computing ecosystems combining CPUs, GPUs, and QPUs.

The 42nd International Symposium on Lattice Field Theory (LATTICE2025)
2-8 November 2025
Tata Institute of Fundamental Research, Mumbai, India

*Speaker

© Copyright owned by the author(s) under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0) All rights for text and data mining, AI training, and similar technologies for commercial purposes, are reserved. ISSN 1824-8039 . Published by SISSA Medialab.

PoS(LATTICE2025)025

1. Introduction

A fundamental part of progress in Lattice QCD is built upon access to and efficient utilization of ever-improving computational infrastructure. Faster and larger machines are regularly being brought on line, the performance of single GPUs continues to increase along with more bandwidth between GPUs, improved algorithms are a major research area and data storage allows efficient reuse of propagators and ensembles between projects and researchers. In addition, machine learning and quantum simulation are new methods being brought to bear on LQCD calculations: To help keep the community updated on this progress, there were four talks given in the Thursday morning plenary session on subjects in this area: 1) *Machine Learning for lattice gauge theories* by Urs Wenger; 2) *Machines and Exascale computing* by Stefan Krieg; 3) *Quantum Simulation of Gauge Theories for Particle and Nuclear Physics: Past, Today and Future* by Zohreh Davoudi and; 4) *Quantum simulation of lattice gauge theories* by Marcello Dalmonte. Writeups of their talks are part of these proceedings.

After these talks, a panel discussion on the Future of Computing was held with panelists representing classical computing and quantum computing. The panelists were:

- Classical Computing
  1. Yasumichi Aoki (RIKEN, Japan)
  2. Hemant Darbari (National Supercomputing Mission, India)
  3. Steven Gottlieb (Indiana University, USA)
  4. Stefan Krieg (Jülich Supercomputing Center, Germany)
  5. Jigar Halani (NVIDIA, India)
- Quantum Computing
  1. Marcello Dalmonte (ICTP, Italy)
  2. Zohreh Davoudi (University of Maryland, USA)
  3. R. Vijayaraghavan (experimentalist—quantum hardware, TIFR, India)

The panel was moderated by Robert Mawhinney of Columbia University, USA.

The panel session had two parts: prepared statements or slides from the panelists and general questions from the moderator and audience to the panel. Because some of the panelists had just given presentations in this session, only some panelists had prepared materials. We turn first to these.

2. The FugakuNEXT Project: Yasumichi Aoki

Yasumichi Aoki of the RIKEN Center for Computational Science, Kobe outlined the development of FugakuNEXT, the next Japanese flagship supercomputer. It will be the successor to Japan's current flagship supercomputer, Fugaku.

The recent lineage of Japan's supercomputers is:

PoS(LATTICE2025)025

1. K computer (SPARC64 VIIIfx, 11 PFlops)

2. Fugaku (A64FX ARM, 540 PFlops)

3. FugakuNEXT (~ 2030)

The most significant architectural shift in FugakuNEXT is the introduction of NVIDIA GPUs alongside Fujitsu's ARM-based CPU "MONAKA-X", moving from a CPU-only to a heterogeneous CPU+GPU design. The development of FugakuNEXT began with a feasibility study in FY2022 and has proceeded through basic design and detailed design. Deployment phases are planned, targeting operational startup in 2030.

The system architecture of MONAKA-X is a many-core ARM ISA CPU using chiplets, with consideration for a Neural Processing Unit (NPU), based on ARM SME (low-precision matrix computing). A high-bandwidth, coherent CPU-GPU interconnect is planned, and NVIDIA's state-of-the-art GPU with advanced stacked memory and next-generation NVLink scale-up networking is under consideration. Final specifications will be determined in the detailed design phase.

The development philosophy for the project follows an "Application First" philosophy, leveraging existing Fugaku software assets while integrating AI-for-science capabilities. Lattice QCD (Bridge++ / DWF-HMC) is among the 14 Early Evaluation Applications, with key technical challenges including GPU porting and tuning, hierarchical memory/network utilization, high-speed low-precision arithmetic (e.g., Ozaki scheme for dense matrix multiplication), and AI integration.

3. Quantum Computing: Marcello Dalmonte and Zohreh Davoudi

Quantum computing and simulations are quickly developing paradigms that are particularly promising in the context of simulating "inherently quantum" computational problems—such as those relevant for chemistry, materials science, and nuclear and particle physics. For nuclear- and particle-physics applications, simulating lattice gauge theories is a frontier direction. While first lattice-gauge-theory studies were mostly focused on condensed-matter-inspired problems, there has been growing interest over the past decade in tackling gauge theories of relevance to the Standard Model with quantum computers. Progress has been steady: from the first key demonstration of gauge-theory dynamics in a trapped ion chain in 2016 [1], significant steps have been taken in the context of Abelian gauge theories: recent experiments have performed real-time simulations at couplings and volumes that begin to approach the classically challenging regimes. The reason is that classical methods either face severe sign problems or rapid entanglement growth, while quantum-computing methods evade these issues. Interestingly enough, the experimental progress has also stimulated new developments on classical algorithms for real-time dynamics, in particular in the context of tensor networks.

The present state of the art is that the community has evolved from building-block demonstrations to large-scale simulations on toy models. Most of the latter have been U(1) and Z_2 lattice gauge theories, coupled to either fermions or Higgs fields, although digital quantum simulations of non-Abelian gauge theories have been realized too. In order to move to richer scenarios, and in the long-term, to computational problems relevant to nuclear and particle physics, several challenges

PoS(LATTICE2025)025

have to be addressed. These range from developing novel strategies to ensure gauge-invariant dynamics in analog-simulation settings (as testified by the slow progress in the context of non-Abelian theories), to new Hamiltonian formulations and quantum algorithms that could be specifically tailored for present hardware capabilities.

Success will demand close collaboration between the lattice-gauge-theory and quantum-science communities, which is likely to yield unanticipated conceptual insights. The history of developments in lattice gauge theory using classical computers is a testimony to the importance of co-design processes in the quantum-computing era. In fact, since quantum computers enjoy several distinct architectures with unique physical properties, matching applications to hardware is most natural. Furthermore, while gauge theories present unique features associated with relativistic and continuum physics and with local-symmetry constraints, leveraging ideas and lessons learned in other quantum many-body simulations on the same hardware can often lead to faster progress.

One key aspect that is particularly interesting in the mid-term is how quantum error correction will interface with lattice field theories. This question has begun to be explored; the resource theory for quantum error correction—non-stabilizerness, also known as magic—has only recently received attention in the quantum many-body and lattice gauge theory contexts. In fact, we do not presently know to what extent interesting problems in lattice gauge theory will be computationally challenging for quantum computers. Addressing such quantum computational complexity questions is of paramount importance, as it parallels the intense efforts in understanding classical computational complexity, e.g., the sign problem and its relation to physical phenomena.

Ultimately, even when quantum-hardware systems reach the reliability and scalability needed to solve meaningful scientific problems, they may not fully replace classical-hardware systems. The more likely scenario will be a computing ecosystem that incorporates CPUs, GPUs, and QPUs (quantum processing units). The lattice-gauge-theory workflow can greatly benefit from these hybrid systems: Classical Monte Carlo computations may inform state preparation strategies on quantum hardware, for which no proven quantum advantage is expected, but real-time dynamical tasks can be offloaded to QPUs. Classical resources can further assist in storing and analyzing quantum measurements at scale.

Last but not least, as the many scientific applications of quantum computing become evident and more communities begin to realize the potential of quantum computing in boosting their efforts, the demand for quantum hardware increases. In fact, this situation is already affecting accessibility and affordability of quantum computing across all fields. The lattice-field-theory community has a long tradition in implementing best practices in acquiring, sharing, and preserving computational resources. Such practices will likely need to be updated and expanded in a quantum-computing ecosystem, particularly in the near-term era where fair and efficient distribution of existing resources at national and international levels has not been fully realized. Workforce education and training remain particularly important goals; standard lattice-gauge-theory training will need to be augmented with theoretical, algorithmic, and practical modules in quantum computing to prepare early-career researchers to leverage this emerging technology.

Many recent reviews provide a comprehensive view of the developments and offer the relevant references for the results and statements mentioned here [2–11]. Furthermore, two plenary presentations at the conference, along with several parallel talks, outline the current state of the field of quantum simulation and computing for lattice field theory. We refer the reader to the proceedings.

PoS(LATTICE2025)025

4. Planned US Computers: Steven Gottlieb

As Yogi Berra is reputed to have said, "It is tough to make predictions, especially about the future." Nevertheless, I will try. I will first summarize what I know about the upcoming Department of Energy and National Science Foundation funded supercomputers. I would then like to shift to longer term issues.

The Department of Energy has two *leadership* computer centers. The Argonne Leadership Computer Center (ALCF) currently runs Aurora and new computers called Equinox and Solstice have been announced. The Oak Ridge Leadership Computer Center (OLCF) runs Frontier. The new computers there are called Lux and Discovery. For NSF, the Texas Advanced Computing Center (TACC) runs Frontera and the new computer is called Horizon. The National Energy Research Scientific Computing center (NERSC) is not a leadership center, but its computer Perlmutter is an important resource in the US for lattice QCD calculations. The next computer there is called Doudna in honor of the Nobel Laureate who invented CRISPR.

The plans for Horizon (TACC) include 2,000 GPU nodes with 4,000 Blackwell GPUs with an 800 Gb/s Infiniband fat-tree network. There will also be 4,750 CPU nodes with two 88 core Vera devices. The network will be based on 400 Gb/s Infiniband. In addition there will be a 400 PB 100% solid state VAST file system. At the time of the conference, an NVL72 Blackwell queue for development, was expected "very soon," and early access to the GPU part of Horizon was expected in March. With the high demand for Blackwell hardware, that has not happened, but TACC does have at least one NVL72 cabinet. As of mid-May, I am still waiting to try it. More details can be found in an August, 2025 **talk** by Dan Stanzione, Director of TACC.

The Doudna computer at NERSC will be a Vera-Rubin system with NVIDIA Quantum-X800 Infiniband 800 Gb network. The vendor is Dell, and many details have not been disclosed. The machine was slated for delivery in late 2026, expected to be run for early science in 2027, and allocated under the ERCAP program in 2028. This **article** by Timothy Morgan has more information.

On October 28, 2025, there was an announcement of the new computers at ALCF and OLCF. Equinox was expected to be available in early 2026 at ALCF with 10,000 NVIDIA Blackwell GPUs. Solstice, with 100,000 Blackwell GPUs, I later learned cannot be installed without a new building. I do not have a time line for its operation. At OLCF, the Lux AI cluster was announced to start operations in early 2026. It uses an AMD EPYC CPU, AMD Instinct MI355X GPUs, and an AMD Pensando network. Discovery is scheduled to be delivered in 2028. It will contain AMD EPYC Venus CPUs and AMD Instinct MI430X GPUs (4 per node), but the number of nodes was not announced. There will be an HPE Slingshot network. The **Argonne** and **ORNL** announcements have more details.

To consider the longer term prospects, I reached out to John Shalf, Director of Computer Science at Lawrence Berkeley National Laboratory. John was heavily involved with the Exascale Computing Project. He sent me links to both an interesting talk and an article from 2019. The talk is called **Computing Beyond the End of Moore's Law**. It contains many ideas including co-design efforts, chiplets, and domain specific architectures. John also discusses the difference between how hyperscalars and leadership centers acquire new hardware, mixed (low) precision arithmetic, and new materials for chips, a development that may be two decades away. You will find additional

PoS(LATTICE2025)025

information in the article. *The Future of Computing Beyond Moore's Law*, J. Shalf, Phil. Trans. R. Soc. A 378: 20190061, **available here**.

After the lattice conference, I met John in person at Supercomputing '25 in St. Louis. John was awarded the Seymour Cray Award. As John reminisced about his career, he said, "We used to call the people who worked at Cray Crayons. What do we call the people who work at NVIDIA?" I shouted from the audience, "Rich." I later emailed John to tell him I was the one who shouted and he replied, "And thanks for the shout out. I couldn't have had a better call-out if I'd planted it myself." Also at Supercomputing, my fellow panelist Stefan Krieg and I were treated to dinner at an expensive steakhouse by NVIDIA's Kate Clark. We did our best to reduce NVIDIA's quarterly profit, but barely made a dent.

5. Configuring a Supercomputing Center: Stefan Krieg

I see a lot of options going forward. Selecting the right architecture for a new system has not become simpler in the recent years. From different accelerators to the amount and type of CPUs in a system, network and memory bandwidths —- hitting the balance right is hard. And outside pressures concerning the HPL number to be delivered are not simplifying matters very much. A pragmatic approach for a computer center is to look at one's application portfolio and estimate how this can realistically be evolved to run well on a new system. But one also needs to consider system operation, middleware, and, crucially, the user facing software infrastructure of the selected architecture. A reasonable measure for the potential of the application portfolio and user facing software stack is to set up a benchmark suite. From the perspective of our field, it must include LQCD codes, as was the case for the JUPITER procurement. The acceptance tests ensure that, at least for the benchmarks, the system and its stack work correctly and are performant.

A new option entering the game is Quantum Computing (QC). Many QC algorithms are, by nature, hybrid classical-quantum methods. It is, therefore, natural to consider a QC device as an add-on to a classical supercomputer. The combined system could then be able to balance the needs of any particular application between the QC and classical part. In Jülich, we refer to such "compute elements" as modules, of which there can be a wide variety, from an AI accelerator module to a neuromorphic or photonic one. If the connectivity between these modules is sufficiently performant, such that one could consider the whole thing to be one system, we refer to it as "Modular Supercomputing Architecture". Assembling these different architectures could enable a higher level of optimization for some workloads, in particular when energy efficiency is concerned. If, however, industry keeps building "GigaWatt-Factories", the impact may be somewhat limited. . .

6. Summary of Open Panel Discussion

During the open discussion, a number of questions were considered by the panel and some panelists elaborated on topics discussed in the previous parts of this report.

One important topic that received considerable attention centered on how the push to support AI and its need for lower precision hardware for training would impact vendors' support for high performance in the single and double precision formats needed for traditional high-performance computing (HPC). Dr. Darbari stressed India's broad initiative, spanning chip development to the

PoS(LATTICE2025)025

selection of the system site for large supercomputing centers. In this complete pipeline, they need to accommodate AI machines going to small numbers of bits in their floating point representations and are striving to find the right balance between AI and more traditional HPC needs. Jigar Halani of NVIDIA said that NVIDIA was well aware of the continued need of some applications for high performance in single and double precision formats and that their products would have offerings where the customer could balance the performance between traditional precisions and reduced precisions, depending on their need. Current LQCD algorithms, particularly Dirac equation solvers, have parts that can be done in reduced precision. Reduced precision calculations also impact the bandwidth requirements on networks connecting computational elements.

The discussion of precision and the role of traditional HPC in influencing vendors' offerings led to similar comments in a larger context, *i.e.*, HPC and traditional scientific algorithms may have less influence on the hardware development, network options or overall machine architecture for computational platforms being developed today. Since the time of the conference, the U.S. Department of Energy announced the Genesis Mission, where it has identified an initial 26 national science and technology challenges to accelerate U.S. innovation and leadership. Steven Gottlieb's section details the new hardware being brought online in the U.S. to support AI and traditional HPC and the Genesis Mission is directly funding scientists to include AI in their computational efforts.

With the rapid evolution of the capabilities of foundation models across the spectrum of human activities, and the attendant large investment in data centers by major corporations, it is difficult to foresee all the impacts on LQCD calculations. Coupled with the continued development of different physical systems to support quantum computing, the future for LQCD calculations is full of promise and uncertainty. Continuing to adapt our algorithms and methods, along with advancing our theoretical understanding of lattice quantum field theory, promises to be an exciting intellectual activity with real impact on understanding the experimental results from nuclear, particle and astroparticle physics experiments.

References

[1] E. A. Martinez et al. Real-time dynamics of lattice gauge theories with a few-qubit quantum computer. *Nature*, 534:516–519, 2016. doi: 10.1038/nature18318.

[2] Monika Aidelsburger et al. Cold atoms meet lattice gauge theory. *Phil. Trans. Roy. Soc. Lond. A*, 380:20210064, 2021. doi: 10.1098/rsta.2021.0064.

[3] M. C. Bañuls et al. Simulating Lattice Gauge Theories within Quantum Technologies. *Eur. Phys. J. D*, 74(8):165, 2020. doi: 10.1140/epjd/e2020-100571-8.

[4] Erez Zohar. Quantum simulation of lattice gauge theories in more than one space dimension—requirements, challenges and methods. *Phil. Trans. A. Math. Phys. Eng. Sci.*, 380 (2216):20210069, 2021. doi: 10.1098/rsta.2021.0069.

[5] Natalie Klco, Alessandro Roggero, and Martin J. Savage. Standard model physics and the digital quantum revolution: thoughts about the interface. *Rept. Prog. Phys.*, 85(6):064301, 2022. doi: 10.1088/1361-6633/ac58a4.

PoS(LATTICE2025)025

[6] Christian W. Bauer et al. Quantum Simulation for High-Energy Physics. *PRX Quantum*, 4(2): 027001, 2023. doi: 10.1103/PRXQuantum.4.027001.

[7] Alberto Di Meglio et al. Quantum Computing for High-Energy Physics: State of the Art and Challenges. *PRX Quantum*, 5(3):037001, 2024. doi: 10.1103/PRXQuantum.5.037001.

[8] Christian W. Bauer, Zohreh Davoudi, Natalie Klco, and Martin J. Savage. Quantum simulation of fundamental particles and forces. *Nature Rev. Phys.*, 5(7):420–432, 2023. doi: 10.1038/s42254-023-00599-8.

[9] Lena Funcke, Tobias Hartung, Karl Jansen, and Stefan Kühn. Review on Quantum Computing for Lattice Field Theory. *PoS*, LATTICE2022:228, 2023. doi: 10.22323/1.430.0228.

[10] Zohreh Davoudi. TASI/CERN/KITP Lecture Notes on "Toward Quantum Computing Gauge Theories of Nature". 7 2025. URL https://arxiv.org/pdf/2507.15840.

[11] Jad C. Halimeh, Niklas Mueller, Johannes Knolle, Zlatko Papić, and Zohreh Davoudi. Quantum simulation of out-of-equilibrium dynamics in gauge theories. 9 2025. URL https://arxiv.org/abs/2509.03586.

PoS(LATTICE2025)025