VOLUME 3 I FEBRUARY 2023

NOVEL MATERIALS DISCOVERY

NOMAD Center of Excellence

NEWSLETTER

EDITORIAL

2022 has been a busy year with remarkable progress in the NOMAD CoE developments. We have released libraries, organized tutorials and other events, and enjoyed fruitful interactions between the partners. We have also passed the midterm of the project, where Matthias Scheffler handed over the coordination to Claudia Draxl, now serving as the Deputy Coordinator.

In this release of our NOMAD CoE Newsletter, you will find a selection of topics: We report on the newest developments of ELPA and the first release of Green-X. You will find a summary on the status of LUMI, Europe's first pre-exascale machine, and a review of the NOMAD CoE Industry Workshop. Finally, we provide you additional information about our developments and our upcoming summer schools, workshops, and hackathons organized for the community.

We wish you a most joyful and successful 2023 and are looking forward to meeting you at one of our next events!

Claudia Draxl & Matthias Scheffler - on behalf of the entire NOMAD CoE team.

ELPA LIBRARY

Andreas Marek and Christian Carbogno

The solution of large eigenvalue problems plays a pivotal role for computational quantum mechanics and hence for the first-principles prediction of molecular and materials properties. For instance, density-functional theory (DFT) typically requires solving multiple symmetric / hermitian eigenvalue problems, which dominates large-scale simulations. Researchers and developers in the electronic-structure theory community are thus always in need of high-performance computing (HPC) eigenvalue solvers that are efficient, scalable, and as accurate as possible: The time-to-solution to directly solve an eigenvalue problem grows with N^3, in which N is the matrix dimension that effectively scales with the simulated system size. To enable calculations for real materials with thousands of atoms, excellent scalability to large node (and/or core) counts are thus required.

For more than a decade, the ELPA library[1] has been developed in a collaboration of mathematicians, computer scientists, HPC experts, and domain scientists in order to provide a fast and scalable eigenvalue-solver library for the electronic-structure community. For dense symmetric / hermitian eigenvalue problems, the ELPA library is nowadays the worldwide leading software package and the quasi standard for the solution of such problems in the overwhelming majority of electronic-structure-theory software packages.

One key point during the development of the ELPA library has always been to keep track of the advancements in HPC technology and to continuously port and optimize ELPA's eigenvalue solver routines for the next generation of HPC hardware in a timely fashion. This has allowed researchers to seamlessly exploit the ever-growing performance of newer HPC systems, be it due to increased machine sizes, due to the advent of new architectures, or due to improvements in the communication networks. With the first exascale HPC system installed at the Oak Ridge Leadership Computing Facility[2] in the US, the electronic-structure community

[1] See: https://elpa.mpcdf.mpg.de

[2] See: https://www.olcf.ornl.gov/frontier/

is eager to utilize this new generation of HPC installations. The huge increase in computational power allows first-principles research to address important, hitherto unsolved materials-science questions, for instance for sustainable energy production, energy storage, and energy saving. In all these cases, realistic and predictive calculations require an accurate thermodynamic description of the relevant materials, including defect formation, interface effects, and dynamical phase transitions. The system sizes necessary to accurately capture these effects involve several thousands of atoms and only become accessible on the latest generation of HPC exascale hardware.

To achieve these ambitious goals, updates and improvements of the ELPA library are necessary, since the new (pre)exascale HPC systems come with a disruptive change in HPC paradigms. Although GPUs have been around for many years to accelerate HPC calculations, the more traditional CPU-based HPC was still the most important aspect for reaching extraordinary computational efficiency. On the upcoming exascale systems, however, most of the computing power is delivered by GPUs. Accordingly, a massively parallel usage of thousands of GPUs is mandatory to efficiently harvest the potential computing power of the systems. This situation is even further complicated by the fact that there are three providers of HPC-GPUs (NVIDIA, AMD, and Intel), which all provide different hardware and different programming models for their accelerators. In order to provide an efficient, general library, developers thus have to port and optimize their software for all different vendors individually.

The ELPA library has been supporting massively parallel GPU-accelerated computations, albeit only for NVIDIA GPUs. In the last two years, we have started to port and to optimize the ELPA eigenvalue solvers also

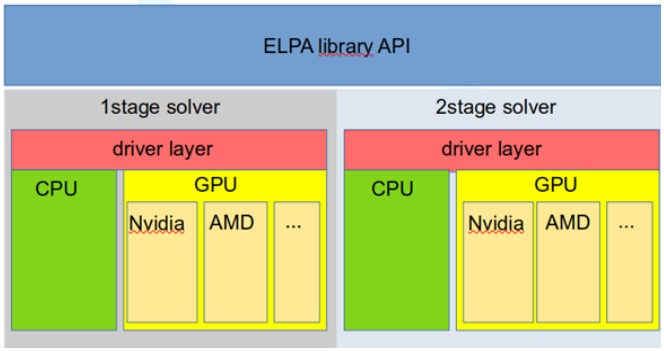

Figure 1: The new ELPA API layout allows one to use the same API calls for the one-stage and two-stage solvers. Within the solvers, a new "driver layer" was implemented which can switch between different CPU or GPU code paths. Thus, the GPU acceleration layer can now use different hardware (and programming models) for NVIDIA, AMD, and Intel GPUs. This driver layer is also easily extendable to support other hardware in the future.

for AMD and Intel GPUs by using the ROCm and Sycl programming models, respectively. This has been facilitated by implementing an abstraction level, the so-called "driver layer", which encapsulates vendor-specific aspects for the different GPU models (CUDA, ROCm, Sycl), and only exposes high-level routines to the users. In turn, this creates cleaner and better readable code that can be more easily maintained, updated, and used in production calculations. With the last release (2022.11.001), these new features are now officially accessible, so that all users can now profit from GPU acceleration for solving eigenvalue problems, regardless of which GPU architectures from which vendor they are actually using.

The described developments are the first key step to enabling routine ELPA usage on exascale systems. Already in 2021, the support of NVIDIA GPUs enabled large-scale production runs on 24,000 GPUs[3] on the Summit system.[4] This entailed the solution of a 1,769,472-dimensional matrix – to our knowledge the worldwide largest dense eigenvalue problem ever sought with a direct solver. Using the newest developments, the ELPA library is currently also being tested

[3] https://gitlab.fizyka.pw.edu.pl/wtools/wslda/-/wikis/Setting-up-diagonalization-engine#benchmarks-scalings

[4] https://www.olcf.ornl.gov/summit/

for AMD GPUs on the LUMI system at the CSCS in Finland.[5] So far, successful runs on up to 1,000 AMD Mi250x GPUs have been demonstrated, and it is planned to eventually exploit the new capabilities of the ELPA library to harvest the power of the complete GPU partition of LUMI. The experiences gained from these large-scale runs will be used to further optimize ELPA, so to provide even better performance to the community in the upcoming releases of the ELPA library. Eventually, this will allow researchers to unleash the computational power of the world's first exascale system *Frontier* for their DFT calculations.

GREEN-X, AN OPEN-SOURCE LIBRARY FOR TIME-FREQUENCY TRANSFORMS

Xavier Gonze, Maryam Azizi, and Claudia Draxl

Calculations beyond DFT (*GW* or RPA) have become part of the standard toolbox in computational condensed-matter physics, as the best approach for compute electronic properties, in particular band gaps. There are, however, severe bottlenecks which make such calculations cumbersome, especially for systems requiring large simulation cells. The main drawback is the increase of their computational cost with the fourth power of the number of atoms. To overcome this barrier, low-scaling algorithms - based on real-space sparseness and time-frequency transforms and exhibiting cubic scaling – have been proposed.[6,7]

So far, these have been implemented in a few codes but the routines are not freely available. Moreover, the accuracy of the provided time-frequency grids is insufficiently characterized. NOMAD CoE dedicates its Work Package 2 to this important topic, with the goal of providing a library (see Figure 2) of routines for operations on Green functions and associated quantities,

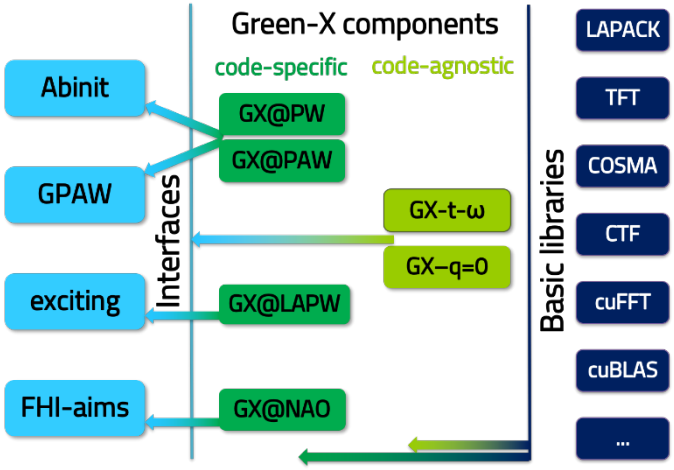

Figure 2: Structure of the library Green-X with its code-family specific parts (dark green) and code-agnostic parts (light green). The NOMAD demonstrator codes on the left stand for representatives of the considered code families.

called Green-X. NOMAD researchers, together with international collaborators, have recently released the time-frequency component of Green-X. It is easily linkable to any DFT code (all-electron or planewave methods) and freely available at NOMAD's GitHub.[8]

The time-frequency Fourier transforms deal with functions that exhibit long tails as well as very localized features in imaginary time and frequency, as shown in Figure 3. As such, their treatment using usual Fast Fourier transforms requires a huge number of sampling points. Instead, a nonuniform Fourier transform approach has

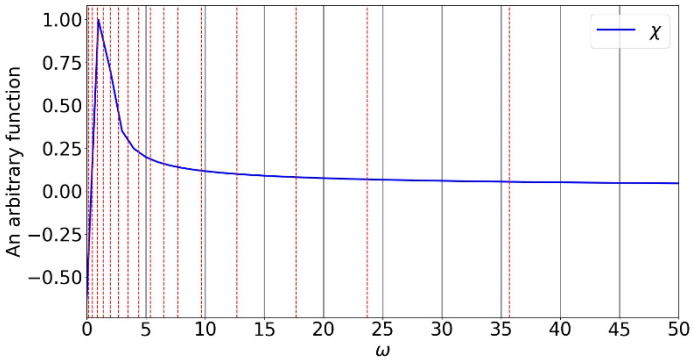

Figure 3: A typical function with a long tail as well as localized features to be Fourier transformed. Black vertical lines indicate usual uniform sampling, red dotted lines nonuniform sampling.

[5] https://www.lumi-supercomputer.eu/

[6] H.N. Rojas *et al.*, Phys. Rev. Lett. **74**, 1827 (1995).

[7] M. Kaltak, J. Klimeš, and G. Kresse, J. Chem. Theory Comput. 10, 2498 (2014); Phys. Rev. B **90**, 054115 (2014).

[8] https://github.com/nomad-coe/greenX

been developed in recent years,[7,9] cutting down the computational requirements by more than one order of magnitude, both in CPU time and memory.

Our library can produce a series of time-frequency grids for different numbers of points and in a broad energy range, which have already been successfully used in a prototype cubic-scaling *GW* implementation[10].

Green-X has a layered design with both; specific modules for different code families but also several functionalities common to all of them. In addition to the time-frequency component, Green-X will soon also include space transformations, sparse or full basis-set dependent transformations, and solutions of the Poisson equation.

LUMI NEARING PRODUCTION

Yussi Heikonen

LUMI will be first of the three European High-Performance Computing Joint Undertaking (EuroHPC JU) funded pre-exascale supercomputers in full production. The system is located at CSC's energy efficient datacentre in Kajaani, Finland, where the waste heat is reused by local households. LUMI is jointly owned by EuroHPC JU and a CSC led consortium of ten countries (Finland, Belgium, Czech Republic, Denmark, Estonia, Iceland, Norway, Poland, Sweden, and Switzerland). For more information, please see https://lumi-supercomputer.eu/. LUMI is an HPE Cray EX supercomputer with 2,560 GPU nodes with one AMD Trento CPU and four AMD MI250X GPUs per node. The CPU partition has 1,536 nodes with two AMD Milan processors per node. The committed Linpack performance of LUMI is 375 Pflops. For more details, please see https://docs.lumi-supercomputer.eu/hardware/.While raw computing power is always a high priority, versatility and the ability to handle a wide variety of workflows were key targets of LUMI's design brief. Accordingly, to complement the main computing engine, the GPU partition (LUMI-G), there is a substantial CPU partition (LUMI-C), a container cloud partition (LUMI-K) and an interactive analysis/post-processing partition (LUMI-D). The computing power is supported by a multi-tier storage system with flash (LUMI-F), fast parallel disk (LUMI-P) and object storage (LUMI-O) components. A high-speed interconnect handles the communication and data transfers within the system. As a result, LUMI is a powerful research instrument not only for traditional HPC simulations but for AI and data-analysis applications as well.

LUMI is being delivered in phases. The first phase consisting of the CPU partition and most of the storage functionality has been available since January 2022. The GPU partition is currently in an extended beta-testing phase until mid-late January, and its acceptance tests are being finalized. In November 2022, LUMI reached position number 3 on the Top500 and number 7 on the Green500 lists. The general availability for the full system is expected in early 2023, once the acceptance tests have been completed.

The LUMI computing and storage resources are divided within the consortium and JU proportionally to the respective investments. Hence, the share of the JU is 50%. Users of the consortium countries can apply resources through their national calls and the JU organizes call of their own. There are various access modes from regular access targeted for typical project use to development and benchmarking access for developing and testing applications. For details, see https://lumi-supercomputer.eu/get-started/. LUMI's computing power comes mainly from the GPU partition, whose

[9] J. Wilhelm, P. Seewald, and D. Golze, J. Chem, Theory Comput. **17**, 1662 (2021).

[10] M. Azizi, J. Wilhelm, D. Golze, M. Giantomassi, A. Buccheri, C. Draxl, A. Gulans, and X. Gonze, to be published.

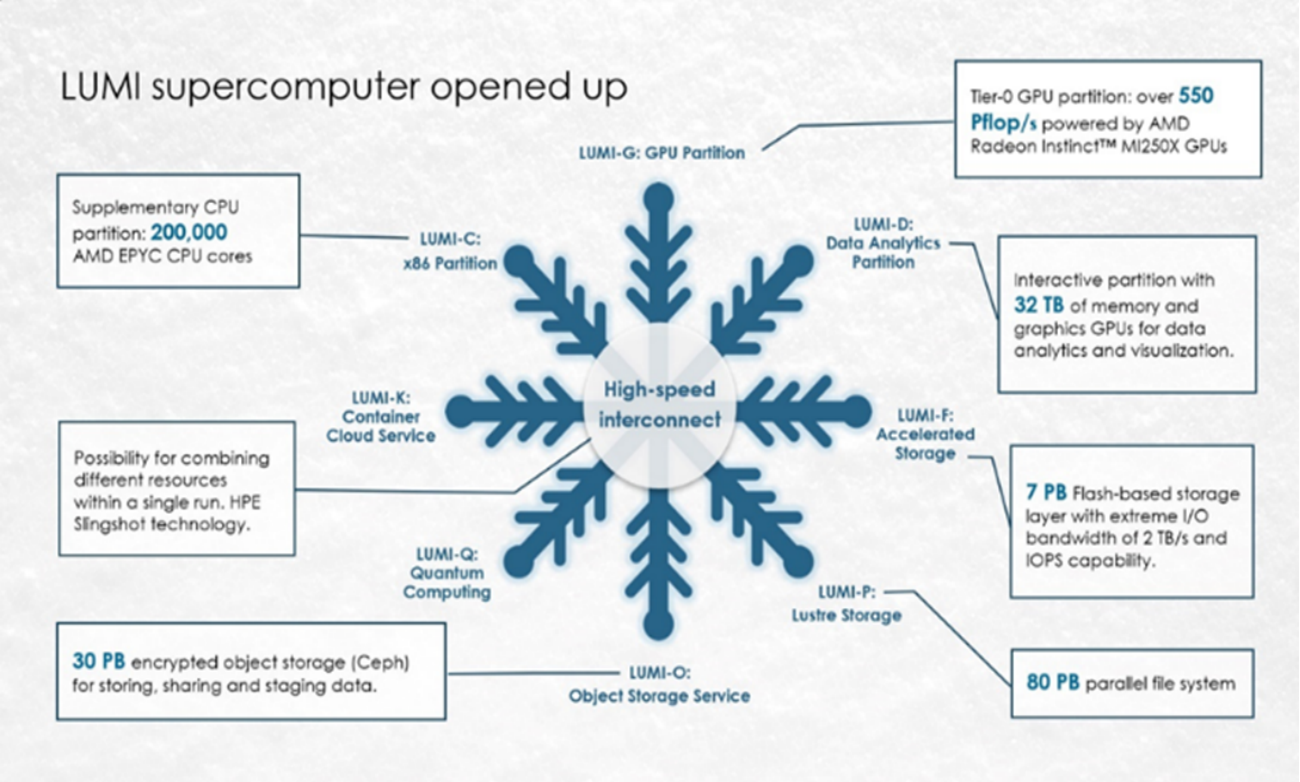

Figure 4: About LUMI - Europe's first pre-exascale machine at CSC in Finland.

AMD hardware may require some porting effort from the application developer. HIP is AMD's native GPU programing language which is almost one-to-one with CUDA. Hence, C++ based CUDA codes can typically be easily converted to HIP (hipified) using a source-to-source translator script. Regarding directive-based GPU programming approaches, OpenMP offloading is supported for C++ and Fortran, while OpenACC is only supported for Fortran. Porting Fortran-based CUDA codes is also possible, but requires more work.

LUMI user support is provided through a distributed team that can be contacted through https://lumi-supercomputer.eu/user-support/need-help/. LUMI documentation is available at https://docs.lumi-supercomputer.eu/.

NOMAD INDUSTRY WORKSHOP

Gian-Marco Rignanese

The 2022 NOMAD Industry Workshop took place as a closed meeting in Berlin, involving several representatives of our Industry Advisory Board comprised in NOMAD CoE's *second-shell* of industrial partners, together with NOMAD's Principal Investigators (PIs).

Previous similar meetings had, in fact, proven to be very valuable and constructive in explaining NOMAD's progress towards exascale computing and in handling extreme-scale data, and in learning about current and future needs of different industrial domains.

The main event was organized by Gian-Marco Rignanese (Université Catholique de Louvain, UCL) and Matthias Scheffler (Fritz-Haber Institute Berlin, FHI). Gabriele Mogni and Sebastian Kokott provided technical support.

The workshop took place on June 24-25, 2022 at the IRIS building of the Humboldt-Universität zu Berlin (HUB). The scientific part comprised 11 talks, of which 7 were given by NOMAD PIs, and 4 by representatives of different NOMAD industrial partners. In addition,

there were several discussion sessions, highlighting that our industry partners very much value the training that NOMAD provides to young researchers who can then join their company with skills in simulation and machine-learning. They also very much appreciate that NOMAD makes a vast amount of data available which they can use for machine-learning and/or as a benchmark for their own calculations. Also, the work on the efficient scaling of simulation and machine-learning codes (including beyond DFT) and the development of automatic workflows for high-throughput calculations is regarded as very beneficial for them. Moreover, it was suggested to carry out a use case in collaboration between the NOMAD CoE and industry, exploiting the NOMAD tools. This would indeed be very convincing for other industrial sectors. All this is very much in line with what NOMAD CoE is currently doing. To stay aligned with the needs of industry, another Industry Workshop will be organized at the end of 2023 or the beginning of 2024.

The whole event finished with a concert by the Berlin Philharmonic Orchestra at the Waldbühne, one of the most impressive open-air concert arenas in Europe.

Workshop Program

Day 1: Talks by NOMAD CoE PIs

14:00	Welcome address and overview talk	Matthias Scheffler (FHI)
14:30	OPTIMADE: A Common REST API for Materials Databases Interoperability	Gian-Marco Rignanese (UCL)
15:00	Reducing the carbon footprint by AI-guided workflows; the NOMAD AI Toolkit for academia and industry	Luca Ghiringhelli (HUB)
15:30	Complex and universal workflows for materials design	Geoffroy Hautier (UCL and Dartmouth College)
16:00	Coffee break	
16:30	Exascale computing in Europe and its impact on industry	Erwin Laure (MPCDF)
17:00	Green-X library	Claudia Draxl (HUB)
17:30	Coupled cluster theory for solids	Andreas Grüneis (TU Wien)
18:00	General discussion	

Day 2: Talks by NOMAD CoE industry partners

09:00	Accuracy and extrapolation of 0K DFT calculations at GTT-Technologies	Moritz to Baben (Managing Director, GTT-Technologies)
09:30	The industrial value of materials modeling: data, simulations, and understanding	Erich Wimmer (Chief Scientific Officer, Materials Design, Inc.)
10:00	Discussion	
10:30	Coffee break	
11:00	Insights into battery materials from atomistic modeling	Ansgar Schäfer (BASF)
11:30	A data management platform to enable material informatics	Davide di Stefano (Senior Project Manager, Ansys)
12:00	Final Discussion	
20:15	Berlin Philharmonics concert at the Waldbühne Berlin	

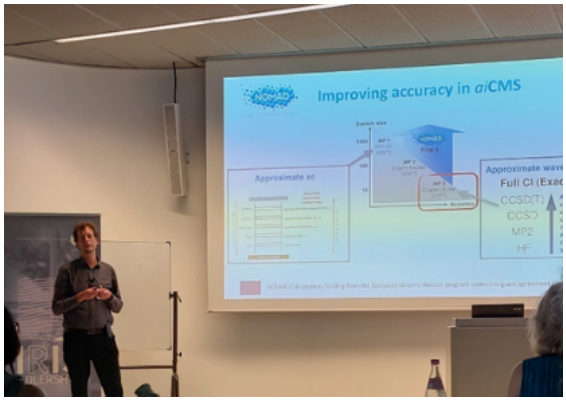

Figure 5: Andreas Grueneis presenting his talk on coupled-cluster theory.

Figure 6: Dinner at the end of Day 1.

Figure 7: Coffee break during Day 2.

OUR LIBRAIRES

Information about our publicly released libraries can be found on our webpage: https://www.nomad-coe.eu/nomad-coe/outreach-nomad-coe/libraries-nomad-coe

OUR EVENTS

Keep track on our upcoming events, like the NOMAD CoE hackathons on the coupled-cluster methods in Vienna (Feburary 2023) and on workflows in Lungby (March 2023), the NOMAD CoE summer schools on "Artificial Intelligence for Materials Science in the Exascale Era" in Cap Roig, Spain (May 2023) and on "Exascale challenges related to Green-function based methods and advanced xc functionals" in Paphos, Cyprus (October 2023). For more detailed information and other events, see https://www.nomad-coe.eu/nomad-coe/events-nomad-coe.

INTERNSHIPS FOR FEMALE RESEARCHERS

You are a female junior scientist at master's or PhD level or early postdoc? Join us for an internship to explore NOMAD CoE's exascale efforts! Have a look at https://www.nomad-coe.eu/nomad-coe/outreach-nomad-coe/opportunities-women-nomad-coe

GET IN TOUCH WITH US

https://www.nomad-coe.eu/

@NoMaDCoE

The NOMAD Laboratory

The NOMAD LinkedIn page

Facebook page

contact@nomad-coe.eu

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 951786.