3.6 Software Advancement Projects and Workshop Support

PHYSBO – optimization tools for PHYSics based on Bayesian Optimization–

Ryo TAMURA
International Center for Materials Nanoarchitectonics,
National Institute for Materials Science,
1-1 Namiki, Tsukuba, Ibaraki, 305-0044

Bayesian optimization (BO) is machine-learning based black-box optimization technique and has recently garnered significant attention in physics, chemistry, and materials science[1]. For example, for materials developments, the trial-and-error process to find better material is regarded as optimizing a black-box function where input is the composition, structure, and process and output will be materials property. In this algorithm, first, a Gaussian process regression that predicts the expected property and variance is constructed from the already observed input-output pairs. Next, through the trained Gaussian process, the probable input data that will yield the desired output value are also selected based on the acquisition function using expected property and variance. Then, the true output value for the selected candidate is obtained by experiments or simulations as black-box functions. BO repeats these processes to find better inputs. Although BO is powerful tool for black-box optimization, BO is generally computationally expensive in two parts: training Gaussian process regression and optimizing acquisition function.

COMBO (COMmon Bayesian Optimization) has been developed mainly for researchers in the materials science field[2]. In the Gaussian process, two hyperparameters, i.e., parameters whose values were given prior to learning, existed: the Gaussian kernel width and noise variance. Using the COMBO package, these hyperparameters were automatically determined by maximizing the Type-II likelihood. In addition, to avoid computationally expensive for training Gaussian process regression, COMBO achieves high scalability using Thompson sampling, random feature map, and one-rank Cholesky update,

To accelerate COMBO package further, PHYSBO (optimization tools for PHYSics based on Bayesian Optimization) package is developed as Python 3 code[3]. In PHYSBO, the massive parallelization using ISSP supercomputer can be performed for optimizing acquisition function, and then both computationally expensive parts in BO can be resolved. In addition, new function to perform multiobjective optimization is implemented.

In physics field, so far, BO has been applied to some problems such as autonomous X-ray scattering experiments [4], inverse scattering [5], crystal structure prediction [6], and effective model estimation[7]. Thus, PHYSBO package can accelerate these problems, and will solve more complex physical problems using supercomputer.

This package is developed with Tsuyoshi Ueno, Kei Terayama, Koji Tsuda, Yuichi Motoyama, Kazuyoshi Yoshimi, and Naoki Kawashima. We would like to thank the support from “Project for advancement of software usability in materials science” by The Institute for Solid State Physics, The University of Tokyo, for development of PHYSBO.

References

[1] K. Terayama, M. Sumita, R. Tamura, and K. Tsuda, Black-box optimization for automated discovery, Accounts of Chemical Research **54**, 1334 (2021).

[2] T. Ueno, T. D. Rhone, Z. Hou, T. Mizoguchi, and K. Tsuda, COMBO: An efficient Bayesian optimization library for materials science, Materials Discovery, **4**, 18 (2016).

[3] `https://www.pasums.issp.u-tokyo.ac.jp/physbo/`

[4] M. M. Noack, K. G. Yager, M. Fukuto, G. S. Doerk, R. Li, and J. A. Sethian, A kriging-based approach to autonomous experimentation with applications to x-ray scattering, Scientific Reports **9**, 1 (2019).

[5] R. Vargas-Hernández, Y. Guan, D. Zhang, and R. Krems, Bayesian optimization for the inverse scattering problem in quantum reaction dynamics, New Journal of Physics **21**, 022001 (2019).

[6] T. Yamashita, N. Sato, H. Kino, T. Miyake, K. Tsuda, and T. Oguchi, Crystal structure prediction accelerated by bayesian optimization, Physical Review Materials **2**, 013803 (2018).

[7] R. Tamura and K. Hukushima, Bayesian optimization for computationally extensive probability distributions, PLoS ONE **13**, e0193785 (2018).

Development of MateriApps Installer

Synge TODO[1,2], Yuichi MOTOYAMA[2], Kazuyoshi YOSHIMI[2], and Takeo KATO[2]
[1]*Department of Physics, University of Tokyo, Tokyo, 113-0033*
[2]*Institute for Solid State Physics, University of Tokyo, Kashiwa, 277-8581*

Nowadays, computer numerical computation is indispensable for theoretical researches in materials science. For the advancement of computational materials science, efficient algorithms for solving equations of materials science are essential. Many excellent applications based on state-of-the-art algorithms have been created so far. In 2013, we launched a portal site for materials science simulations, MateriApps [1], to disseminate information about the developed software to experimentalists and corporate researchers. We have been disseminating information about the application.

One of the obstacles for users to start using published applications in materials science is installing software. MateriApps LIVE! [2] is an environment that allows users to quickly try out computational materials science applications on their laptops and other devices. MateriApps LIVE! is a Virtual Hard Disk Image (OVA) of VirtualBox that includes applications, OS (Debian GNU/Linux), editors, visualization tools, and other environments needed to get started with the tutorial. Using MateriApps LIVE!, it is possible to quickly set up a computing environment for participants in classes and software training sessions.

However, the environment provided by MateriApps LIVE! is not enough to proceed with full-scale simulations. Since MateriApps LIVE! runs as a virtual machine, its computational power is somewhat limited. To support users interested in larger-scale simulations, we had started the development of MateriApps Installer in 2013.

As FY2020 Project for Advancement of Software Usability in Materials Science (PASUMS), we have made several significant updates on MateriApps Installer: i) organized directory structure and scripts, ii) added full documentation and tutorials, iii) upgraded supported software, iv) supported new hardware [ISSP system B (ohtaka)], v) supported new compilers [GCC 10 and Intel oneAPI].

Version 1.0 of MateriApps Installer was released in March 2021, which includes install scripts for ALPS, ALPSCore, DSQSS, Quantum ESPRESSO, HΦ, Kω, LAMMPS, mVMC, OpenMX, RESPACK, and TeNeS. Also, it includes scripts for the following tools and libraries: Boost, CMake, Eigen3, FFTW, GCC, Git, GSL, HDF5, LAPACK, libffi, OpenBLAS, OpenMPI, OpenSSL, Python3, ScaLAPACK, Tcl/Tk, and zlib.

Using MateriApps Installer, the above material science applications have been preinstalled on the ISSP supercomputers (ohtaka and enaga). The source code of MateriApps Installer is freely available from GitHub [4]. MateriApps Installer is distributed under the GNU General Public License version 3 (GPLv3). However, the patch files for each software are distributed under the license of the software.

[1] `https://ma.issp.u-tokyo.ac.jp/`
[2] `https://cmsi.github.io/MateriAppsLive/`
[3] `https://www.pasums.issp.u-tokyo.ac.jp/mainstaller/`
[4] `https://github.com/wistaria/MateriAppsInstaller/`

Advancement of the experimental data analysis for 2D material structure

Takeo Hoshi

Department of Applied Mathematics and Physics, Tottori University,

4-101 Koyama-Minami, Tottori 680-8550, Japan.

The open source software 2DMAT, a data-analysis software of total-reflection high-energy positron diffraction (TRHEPD) experiment, was developed in Project for Advancement of Software Usability in materials science (PASUMS) at FY2020 [1]. TRHEPD is a novel experimental technique of structure determination of two-dimensional materials and is being conducted intensively at the Slow Positron Facility of High Energy Accelerator Research Organization (KEK) [2]. We developed a python-based data analysis software of TRHEPD at 2018 and 2019 [3]. The software uses an iterative optimization (Nelder-Mead) algorithm and the grid-base search algorithm. The software was used in several application studies [4, 5] in the collaboration with experimentalists, I. Mochizuki (KEK), A. Takayama (Waseda U) and their co-workers.

In the PASUMS project at FY2020, we reorganized the software and added the Bayesian optimization algorithm, realized by the PHYSBO library [6], and the replica-exchange Monte-Carlo algorithm. The software was called 2DMAT v.1 and was published at Feb. 2021 [1]. A hands-on seminar was held at 20. April 2021 and was filled to capacity with 30 participants (https://ccms.issp.u-tokyo.ac.jp/event/4570). Several application studies with 2DMAT are currently underway.

Our successor PASUMS project started in April 2021, so as to add several parallelized algorithms and support other experimental techniques of the structure determination of two-dimensional materials, such as surface X-ray diffraction experiment and low energy electron diffraction experiment.

References

[1] https://www.pasums.issp.utokyo.ac.jp/2DMAT

[2] (Review) Y. Fukaya, *et al*, J. Phys. D: Appl. Phys. 52, 013002 (2019).

[3] K. Tanaka, *et al*, Acta. Phys. Pol. A 137, 188 (2020); JJAP Conf. Series, in press; https://arxiv.org/abs/2002.12165

[4] T. Hoshi, *et al*, Submitted; Preprint: https://arxiv.org/abs/2103.04875

[5] M. Hamada , *et al*, Poster presentation, JPS meeting, online, Mar. 2021.

[6] T. Ueno, *et al*, Materials Discovery 4, 18 (2016); https://www.pasums.issp.u-tokyo.ac.jp/physbo

Report of CCMS hands-on sessions in the 2020 fiscal year

Kota IDO
Institute for Solid State Physics, University of Tokyo
Kashiwa-no-ha, Kashiwa, Chiba 277-8581

In the 2020 fiscal year, Center for Computational Materials Science (CCMS)[1] in the Institute for Solid State Physics (ISSP) held one hands-on session using the ISSP supercomputer, which is shown in Table 1. In this report, we briefly summarize the hands-on session in Table 1.

TeNeS is open-source software based on the tensor network method[3, 4]. Users can solve two-dimensional quantum lattice models using the corner transfer matrix renormalization group method for a projected entangled pair state. Since TeNeS supports MPI/OpenMP hybrid parallelization, this software efficiently works on massively parallel machines such as enaga. Development of TeNeS was supported by "Project for advancement of software usability in materials science" (PASUMS) [5] in the 2019 fiscal year. In the hands-on session, Tsuyoshi Okubo and co-developers explained the basics of the tensor network method and gave a tutorial of TeNeS. Materials for this hands-on session are available on the official site (in Japanese) [6].

Date	Software	Main lecturer	Website
Nov. 10	TeNeS	T. Okubo	[2]

Table 1: List of CCMS hands-on sessions using ISSP supercomputer in the 2020 fiscal year.

References

[1] https://ccms.issp.u-tokyo.ac.jp

[2] https://ccms.issp.u-tokyo.ac.jp/event/3844

[3] https://www.pasums.issp.u-tokyo.ac.jp/tenes/en

[4] https://ma.issp.u-tokyo.ac.jp/en/app/2291

[5] https://www.pasums.issp.u-tokyo.ac.jp/

[6] https://www.pasums.issp.u-tokyo.ac.jp/tenes/doc/paper

Supercomputer course of Computational Materials Design (CMD®) workshop

Masaaki GESHI[1], Yoshitada MORIKAWA[2], Tomoya ONO[3]

[1]*Institute for NanoScience Design, Osaka University,*
1-2 Machikaneyama, Toyonaka, Osaka 560-0043, Japan

[2]*Department of Precision Engineering, Graduate School of Engineering,Osaka University,*
2-1 Yamada-oka, Suita, Osaka 565-0871, Japan

[3]*Department of Electrical and Electronic Engineering,*
Graduate School of Engineering, Kobe University,1-1 Rokkodaicho, Nada, Kobe 657-8501, Japan

The 37th Computational Materials Design (CMD®) workshop (CMD37) has been held from August 31st to September 4th and the 38th CMD® workshop (CMD38) has been done from February 22nd to February 26th. Both were held online. In this workshop we have the supercomputer course to train up human resources to advance research by using system B supercomputer of ISSP, the University of Tokyo.

In CMD37 eight participants took the supercomputer course and got a tutorial on STATE-Senri developed by Y. Morikawa. After giving the introductory lecture of large-scale computing by M. Geshi and explaining how to use the supercomputer of ISSP and also how to use STATE-Senri, calculation models on each research subject of the participants were built and their calculations were carried out. Concrete subjects were molecular adsorption on solid surfaces, chemical reactions at electrode interfaces, adsorption and diffusion of atoms on graphene, electronic structures of oxides for ion batteries and so on. The participants performed the calculations and examined the results.

In CMD38 four participants took the supercomputer course and used the supercomputer of ISSP. They got a tutorial on RSPACE developed by T. Ono and M. Uemoto. After giving the introductory lecture of large-scale computing by M. Geshi and describing the calculation method of electronic states and electron conduction property using RSPACE, exercises published in the manual were carried out. Then, electronic state calculations were carried out on a plurality of molecular systems, and the electronic density distribution was visualized. Finally, the atomic structure optimization of graphene blisters, SiC/SiO_2 interfaces, fullerenes, and ZnO was carried out.

We would like to thank the organizers of the MateriApps workshop at the Institute for Solid State Physics, the University of Tokyo, for providing us with useful information for holding the workshop online.