

## **3.6 Software Advancement Projects and GPGPU Support**

# Development of numerical library $K\omega$ ver.1 and quantum lattice solver $\mathcal{H}\Phi$ ver.2

Takeo Hoshi<sup>1</sup>, Youhei Yamaji<sup>2</sup>, Mitsuaki Kawamura<sup>3</sup>, Kazuyoshi Yoshimi<sup>3</sup>, Takahiro Misawa<sup>3</sup>, Synge Todo<sup>4</sup>, Tomohiro Sogabe<sup>5</sup>, Naoki Kawashima<sup>3</sup>

<sup>1</sup> *Department of Applied Mathematics and Physics, Tottori University  
Koyama-Minami, Tottori 680-8550, Japan*

<sup>2</sup> *Department of Applied Physics, The University of Tokyo  
Bunkyo-ku, Tokyo, 113-8656, Japan*

<sup>3</sup> *The Institute for Solid State Physics, The University of Tokyo  
Kashiwa-shi, Chiba, 277-8581, Japan*

<sup>4</sup> *Department of Physics, The University of Tokyo  
Bunkyo-ku, Tokyo, 113-8656, Japan*

<sup>5</sup> *Department of Computational Science and Engineering, Nagoya University  
Chikusa-ku, Nagoya, 464-8603, Japan*

The two novel open-source softwares of (i) numerical library  $K\omega$  ver.1 [1] and (ii) quantum lattice solver  $\mathcal{H}\Phi$  ver.2 [2] were developed in Project for advancement of software usability in materials science [3] at the fiscal year of 2016. The project name is ‘shifted Krylov-subspace algorithm and novel solvers for computational condensed matter physics’. The project is an interdisciplinary one between computational material science and applied mathematics. The softwares are preinstalled on the supercomputer (Sekirei) at ISSP [3]. The two softwares are closely related, because  $K\omega$  is a set of numerical linear-algebraic routines and  $\mathcal{H}\Phi$  ver.2 supports the use of  $K\omega$  in the optical spectrum calculations.

$K\omega$  is a general numerical library for the Green’s function of

$$G^{ab}(\omega) = \langle a | \frac{1}{(E_0 + \omega + i\delta)I - H} | b \rangle, \quad (1)$$

where  $H$  is a large-scale complex-Hermitian or real-symmetric matrix and  $|a\rangle, |b\rangle$  are the input vectors. Traditionally, the problem was solved by the Lanczos-based algorithm. In the present solver, instead, the numerical solution

of Eq. (1) is obtained from the linear equation of

$$[(E_0 + \omega + i\delta)I - H] |x\rangle = |b\rangle \quad (2)$$

and the solution of  $|x\rangle$  is obtained by the novel iterative algorithm, called shifted Krylov-subspace algorithm [4, 5]. The algorithm was used for the excited spectrum of many-body states in a previous paper [5], which motivated us to the present project. The algorithm enables us to control the accuracy of the spectrum  $G^{ab}(\omega)$  at a specific frequently  $\omega$  [5], when one monitors the residual vector of Eq. (2). The method is general and was applied also to many other computational science fields, such as QCD[4], electronic structure calculations, transport calculation with non-equilibrium Green’s function theory.  $K\omega$  is a general numerical library and can be called, in principle, from any material simulator, as well as  $\mathcal{H}\Phi$  ver.2. Moreover, a mini-application is included in the package of  $K\omega$ , so that researchers can evaluate the numerical library before the use in their real researches.

$\mathcal{H}\Phi$  ver.2 is the latest version of  $\mathcal{H}\Phi$  [6]. The quantum lattice solver  $\mathcal{H}\Phi$  is a program

package based on exact diagonalization applicable to a broad range of quantum lattice models, including the Heisenberg model, the Kitaev model, the Hubbard model and the Kondo-lattice model. In  $\mathcal{H}\Phi$  ver.1, the Lanczos method for calculating the ground state and a few excited states, thermal pure quantum (TPQ) states [7] for finite-temperature calculations, and full diagonalization method for checking results of Lanczos and TPQ methods are implemented with an easy-to-use and flexible user interface.

The project in the 2016 fiscal year [3] has supported implementation of the Lanczos and shifted Krylov-subspace algorithm for calculating excitation spectra in the latest version  $\mathcal{H}\Phi$  ver.2. The  $\mathcal{H}\Phi$  ver.2 call subroutines for the shifted Krylov-subspace algorithm from the library  $K\omega$ .

As an example tractable by  $\mathcal{H}\Phi$  ver.2, we show excitation spectra of an *ab initio* spin hamiltonian of an iridium oxides  $\text{Na}_2\text{IrO}_3$  [8], which is a so-called Kitaev material and a typical example of frustrated magnets due to magnetic anisotropy. In Fig.1, dynamical spin structure factors calculated for a 24-site cluster of the *ab initio* spin hamiltonian of an iridium oxides  $\text{Na}_2\text{IrO}_3$  are shown. The continuum in the spectra is the hallmark of the proximity to the Kitaev's quantum spin liquid. The controlled accuracy of the shifted Krylov-subspace algorithm safely resolves the detailed continuum spread over the wide range of frequency, where the typical exchange energy scale of  $\text{Na}_2\text{IrO}_3$  is 30 meV.

For a future outlook, the successor project in the fiscal year of 2017 [3] will support implementation of real-time dynamics, finite-temperature excitation spectra, and interfaces for data science approaches.

## References

[1] <https://github.com/issp-center-dev/Komega>

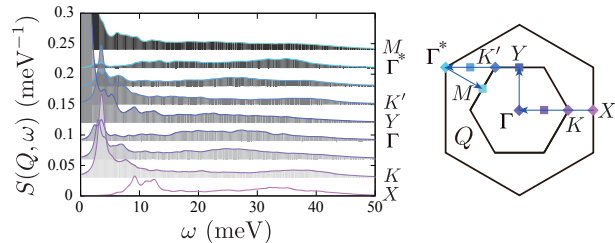


Figure 1: Calculated dynamical structure factors  $S(Q, \omega)$  of  $\text{Na}_2\text{IrO}_3$  [8]. The spectra are vertically shifted depending on  $Q$ . The Brillouin zone of  $\text{Na}_2\text{IrO}_3$  is shown in the right panel.

[2] <https://github.com/QLMS/HPHi>

[3] <http://www.issp.u-tokyo.ac.jp/supercom/softwaredev>

[4] A. Frommer, BiCGStab( $l$ ) for families of shifted linear systems, *Computing* 70, 87 (2003).

[5] S. Yamamoto, T. Sogabe, T. Hoshi, S.-L. Zhang and T. Fujiwara, ‘Shifted conjugate orthogonal conjugate-gradient method and its application to double orbital extended Hubbard model’, *J. Phys. Soc. Jpn.* 77,114713 (2008).

[6] M. Kawamura, K. Yoshimi, T. Misawa, Y. Yamaji, S. Todo, N. Kawashima, ‘Quantum Lattice Model Solver  $\mathcal{H}\Phi$ ’, arXiv:1703.03637, *Computer Physics Communications*, in press.

[7] S. Sugiura and A. Shimizu, ‘Thermal Pure Quantum States at Finite Temperature,’ *Phys. Rev. Lett.* 108, 240401 (2012).

[8] Y. Yamaji, Y. Nomura, M. Kurita, R. Arita, and M. Imada, ‘First-principles study of the honeycomb-lattice iridates  $\text{Na}_2\text{IrO}_3$  in the presence of strong spin-orbit interaction and electron correlations,’ *Phys. Rev. Lett.* 113, 107201 (2014).

# Development of open-source software for Many-variable variational Monte Carlo (mVMC) method

Takahiro Misawa

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

The variational Monte Carlo (VMC) method is a commonly used method for treating the low-energy effective models of quantum many-body systems such as the Hubbard model and the Heisenberg model [1]. In the VMC method, the many-body correlation factors such as the Gutzwiller and the Jastrow factors are introduced to the one-body wave functions. By optimizing the variational parameters in the many-body wave functions to lower the energy, we can obtain the wave functions for the ground states of the quantum many-body systems.

Because the VMC does not suffer from the negative sign problems, the VMC method is widely used for analyzing the low-energy effective models for quantum many-body systems including the strongly correlated electrons systems. However, it has been pointed out that accuracy of the conventional VMC method is not high due to the strong limitation of the implemented wave functions in the conventional VMC method. To overcome the limitation of the conventional VMC, we have introduced many variational parameters in the one-body part of the wave functions (typically more than ten thousand variational parameters) and have succeeded in relaxing the limitations of the wave functions. We also employ the quantum number projection to recover the symmetries inherent to the systems [2]. By combining the many-variable variational wave functions and the quantum number projections,

we have shown that the highly accurate calculations is possible for the wide range of the quantum many-body systems [2, 3]. We call this method *many-variable variational Monte Carlo* (mVMC) method to distinguish the conventional VMC method, whose number of variational parameters is limited to more or less one hundred. Although the optimization of the many variational parameters is difficult in general, it is shown that the optimization of more than ten thousand variational parameters is possible by using the stochastic reconfiguration (SR) method proposed by S. Sorella [4], which is based on the imaginary-time evolution of the wave functions.

The project for advancement of software usability in materials science in the 2016 fiscal year has supported the release of the mVMC as the open-source software [5]. The project has especially supported the improvement of the user interfaces of mVMC and extensions of the wave functions. The latest version of the mVMC is preinstalled on the system B (sekirei) of the supercomputers at ISSP.

By using the software, users can easily start the mVMC calculation by preparing one input file, whose length is typically less than ten lines, for the standard models in the strongly correlated electron systems. As an example, we show the input file for the  $4 \times 4$  Hubbard model as below:

```
L           = 4
```

```

W      = 4
Lsub   = 2
Wsub   = 2
model  = "FermionHubbard"
lattice = "tetragonal"
U      = 4.0
t      = 1.0
nelec  = 16

```

Here  $L$  and  $W$  are a linear extents of square lattice, in  $x$  and  $y$  directions, respectively. The nearest neighbor transfer (on-site Coulomb repulsion) is represented by  $t$  ( $U$ ). Details of the other keywords in the input files can be found in the manuals of mVMC [5]. In Fig. 1, we show the optimization process of the mVMC. As shown in this figure, by preparing the above simple input file, users can easily obtain the wave functions for the ground state whose accuracy is  $E_{\text{mVMC}} - E_{\text{exact}} \sim 10^{-3} \times t$ . By changing the keywords in the input file slightly, users can easily perform the mVMC calculations for the other standard models such as the Heisenberg model and the Kondo-lattice model. It is also easy to treat more general models in the quantum many-body systems by properly preparing the input files. We hope that the mVMC becomes useful software and will be widely used by the broad spectrum of the researchers in the condensed matter physics as well as the other fields such as the quantum chemistry.

## References

- [1] For a review, see C. Gros, Ann. Phys. 189, 53 (1989).
- [2] D. Tahara and M. Imada, J. Phys. Soc. Jpn. 77, 114701 (2008).
- [3] T. Misawa and M. Imada, Phys. Rev. B 90, 115137 (2014).
- [4] S. Sorella, Phys. Rev. B 64, 024512 (2001).

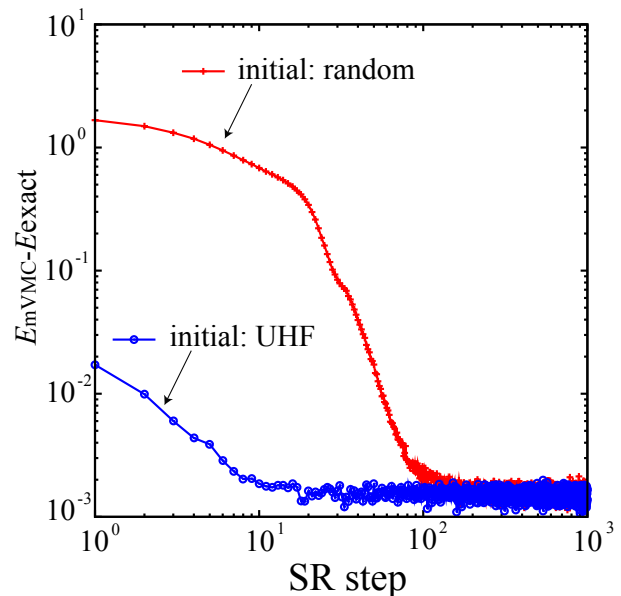


Figure 1: Optimization process of mVMC for  $4 \times 4$  Hubbard model. The internal energy per site calculated by mVMC (exact diagonalization) is represented by  $E_{\text{mVMC}}$  ( $E_{\text{exact}}$ ). Exact diagonalization is done by using  $\mathcal{H}\Phi$  [6]. We take two different initial states: One is the random initial states (red crosses) whose initial variational parameters are generated by the random numbers and the variational parameters of the another one is generated by results of the unrestricted Hartree-Fock (UHF) calculations (blue circles). Although the two different initial states give nearly the same converged results, the convergence is faster by taking UHF initial states.

- [5] <http://ma.cms-initiative.jp/en/application-list/mvmc/mvmc>
- [6] <http://qlms.github.io/HPhi/index.html>

# ISSP support service of software optimization for the GPGPU architectures: A case study for molecular dynamics simulation software MODYLAS

Yoshimichi ANDOH

*Center of Computational Science, Graduate School of Engineering, Nagoya University  
Furo-cho, Chikusa-ku, Nagoya 464-8603*

With the GPGPU support service by the ISSP in the 2015 and 2016 fiscal years, our originally developed general-purpose molecular dynamics (MD) simulation software MODYLAS[1] has been optimized for the system B in the ISSP supercomputer center by using the OpenACC directive-based extension for the Fortran language.

Our MD software assumes that calculation system is in the condensed phase, and contains the long-range Coulombic interaction under three dimensional periodic boundary condition. The Coulombic interactions are strictly and efficiently calculated by the fast multipole method[2]. Most time consuming parts in such MD calculations (hot spots) are pairwise additive calculations of the interaction with nearby located atoms (p2p), and the translation from multipole to local expansions (M2L). With preceding optimization with the OpenACC, these hot spots had been already accelerated at the kickoff of the support service. So, the optimizations in this service were mainly applied to the secondary hot spots.

In the 2015 fiscal year, the conversion from point charges to multipoles (p2M) and evaluation of the electric field on particle by local expansions (L2p) were successfully accelerated by nearly 25 times, inserting the OpenACC directives into these secondary hot spots. Further, transfer style of the data of coordinates and force field parameters from the main memory to the GPGPU's on-board memory, which

is needed in the first hot spot, was improved. With these tunings, total calculation time for one MD cycle was considerably reduced, if compared with the code before the tunings.

In the 2016 fiscal year, we tried to further accelerate the p2p code by utilizing the Newton's third law, and to introduce task parallelization between CPU and GPGPU. In the former, the atomic operations at the summation of the calculated atom pair forces drastically reduced calculation efficiency. In the latter, grammatical restriction in the OpenACC prevented us from implementing the task parallelization straightforwardly. Consequently, despite our efforts, further acceleration of the code on the GPGPU was not achieved.

In this support program, code structure was changed only locally under a constraint of work time. For further speed-up of the code, it is essential to propose new arithmetic algorithms suited for the parallelization on GPGPU architectures, and if necessary, to implement them with the CUDA platform.

## References

- [1] Y. Andoh, *et al.* : J. Chem. Theory Comput. **9** (2013) 3201.
- [2] L. F. Greengard : *The Rapid Evaluation of Potential Fields in Particle Systems* (MIT Press, 1988).

# Implementation of GPGPU codes in Real-space Car-Parrinello Molecular Dynamics (RS-CPMD)

Yasuteru SHIGETA and Jun-ichi IWATA

*Center for Computational Sciences,*

*University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8577, Japan.*

*Graduate School of Engineering, The University of Tokyo, Hongo 7-3-1, Tokyo, Japan.*

The real space density functional theory (RSDFT) developed by one of us has enabled static calculations of several tens of thousands of silicon atoms on K computer [1, 2]. However, in order to describe physical properties of a given system, entropic contribution becomes important as the number of particles increases. Molecular dynamics (MD) simulations are usually used to take it into account. However, it is difficult to handle chemical reactions with classical MD. Nevertheless, the first-principles MD (FPMD) also has a drawback in required computational costs. In order to avoid them, Car and Parrinello have proposed an efficient method to avoid solving the self-consistent field calculations in FPMD [3].

We have implemented CPMD modules in RSDFT (RS-CPMD), which enables us to simulate thousand atoms FPMD on K computer. In this report, performances of a GPU version of RSCPMD under the support of ISSP (SGI Jpn and Ark Information Systems) is presented. The most severe bottleneck in RS-CPMD is evaluation of inner products between wave functions and its velocities ( $O(N_{\text{grid}}^3)$ ).

Therefore, in the implementation of the last year, we made GPU modules for this inner product calculation. Since BLAS3 routines can be used for this operation, three or four-fold speed-up can be achieved by adopting GPU. In this year, GPU versions of several routines in a force evaluation were implemented. By changing the numbers of MPIs and threads, we have tested our programs on ISSP system B 64 nodes (512 cores and 128 GPUs) for an organic solvent (1664 atoms) as a target system. For the fastest case it took 6.28 sec for one CPMD step, which is comparable with 5.54 sec with 4096 cores in K computer. For Ewald and local force calculations, drastic speed-up was achieved. While the performance of the GPU version for non-local forces is not high due to a difficulty in load balance and multiple internode communications.

## References

- [1] Y. Hasegawa *et al.*, *J. High Perform. Comp. Appl.* **28**, 335 (2014).
- [2] <https://github.com/j-iwata/RSDFT>.
- [3] R. Car, M. Parrinello, *Phys. Rev. Lett.* **55**, 2471 (1985)