jh170058-NAHI

Development of next-generation quantum material research platform

Takeo Hoshi（Tottori University）

The present project aims an international collaboration, mainly between Japan and US, in developing the next-generation quantum material research 'platform' for seamless research with scalable quantum material simulators and data scientific analyzers. In FY2017, our codes were optimized on Oakfores-PACS and shows a good strong scaling. Moreover, novel methodologies for large-scale linear-algebraic solvers were also developed. We started productive simulations, such as organic materials for flexible device. An international workshop was held at 4. Dec. 2017 on University of Tokyo. We found that we can share many issues between Japan and US, such as large-scale non-adiabatic quantum molecular dynamics simulation. Our progress will be reflected on several projects like Priority issue 7 of the post-K computer.

1. Basic Information

(1) Collaborating JHPCN Centers

University of Tokyo with Oakforest-PACS.

(2) Research Areas

- ☐ ●Very large-scale numerical computation
- ☐ Very large-scale data processing
- ☐ Very large capacity network technology
- ☐ Very large-scale information systems

(3) Roles of Project Members

Takeo Hoshi: (i) Combined research of simulation and data science, (ii) numerical methods

Aiichiro Nakano (dupty leader, U. Southern California): Discussion on the international collaboration between US and Japan

Yusaku Yamamoto (dupty leader, U. Electro-Communications) : numerical methods

Kengo Nakajima (U. Tokyo): Advices on HPC techniques on OFP

Kohei Shimamura (Kobe U.): Quantum material simulation

Shigenori Tanaka (Kobe U.): Quantum material simulation

Fuyuki Shimojo (Kumamoto U.): Quantum material simulation

Weichung Wang (National Taiwan U.): Numerical methods

Shuhei Kudo (U. Electro-Communications): Numerical methods

Wei Chien Liao (National Taiwan U): Numerical methods

Yuhsiang Tsai (National Taiwan U): Numerical methods

Yukiya Abe (Tottori U.): Quantum material simulations

Kentaro Oohira (Tottori U.): Combined research of simulation and data science

Tomoya Fukumoto (Tottori U.): Quantum material simulations

2. Purpose and Significance of the Research

The present project was organized for the international collaboration, mainly between Japan and US, for developing the next-generation quantum material research 'platform' capable of scientific and industrial

uses. The main collaborator in US is Nakano. Our final goal is to construct, on supercomputers, a comprehensive research 'platform' that contains several scalable quantum material simulators and data-scientific analyzers for academic and industrial researchers.

3. Significance as a JHPCN Joint Research Project

Since the above purpose requires the co-design research between application (simulation and data science), algorithm and architecture, a joint research between them is crucial.

4. Outline of the Research Achievements up to FY2016

The seeds of the present project are the novel scalable methods for large-scale quantum material simulations (electronic state calculations); (I) Hoshi and co-workers developed EigenKernel (https://github.com/eigenkernel/), a hybrid dense-matrix solver for generalized eigenvalue problem. EigenKernel is a general 'kernel' of the large-scale electronic state calculation. See Ref.(*1) at the end of this section. (II) Hoshi and co-workers developed also ELSES (= Extra-Large-Scale Electronic Structure calculation; http://www.elses.jp/), an order-N elelectronic state calculation code with ab-initio-based modeled (tight-binding) theory. See Ref.(*2) at the end of this section. ELSES was used also in industrial research by M. Ishida (Sumitomo Chemical Co. Ltd.) in the collaboration with Hoshi on the K computer (hp150066, hp160087, hp170083; Paper in preparation). (III) Shimojo, Nakano, Shimamura, and co-workers developed LDC-DFT(= Lean Divide-and-Conquer Density Functional Theory), an first-principle order-N electronic state calculation code, based on the domain-decomposition method. See Ref.(*3) at the end of this section.

(*1) H. Imachi and T. Hoshi, Hybrid numerical solvers for massively parallel eigenvalue computation and their benchmark with electronic structure calculations, J. Inf. Process. 24, pp. 164 -- 172 (2016).

(*2) T. Hoshi, H. Imachi, K. Kumahata, M. Terai, K. Miyamoto, K. Minami and F. Shoji, Extremely scalable algorithm for 10^8-atom quantum material simulation on the full system of the K computer, Proc. ScalA16 in SC16, pp.33-40 (2016).

(*3) F. Shimojo, S. Hattori, R. K. Kalia, M. Kunaseth, W. Mou, A. Nakano, K. Nomura, S. Ohmura, P. Rajak, K. Shimamura, and P. Vashishta E A divide-conquer-recombine algorithmic paradigm for large spatiotemporal quantum molecular dynamics simulations, J. Chem. Phys. **140**, 18A529 (2014).

5. Details of FY2017 Research Achievements

5-1 Code optimization on OFP

Our codes were optimized on OFP. All the codes are written in Fortran90 with MPI/OpenMP parallelism. Since the algorithms are designed for massive parallelism, no drastic modification is needed. The benchmark test was carried out with upto 2048 nodes, a quarter of the whole system, since the maximum number of nodes in the present project is set to be 2048.

Figure 1 shows the strong scaling benchmark of optimized EigenKernel,

our generalized eigenvalue problem solver[23]. EigenKernel enables us several hybrid solver workflows that uses some of the three numerical libraries of ScaLAPACK, ELPA and EigenExa(*1). We solved a test problem with the matrix size of M=90000. As result, the two novel solvers of 'D' and 'F' that use ELPA or EigenExa routines (See (*1) for detail) show the strong scaling property with upto 2048 nodes, while the conventional (ScaLAPACK) solver shows a severe bottleneck in scalability.

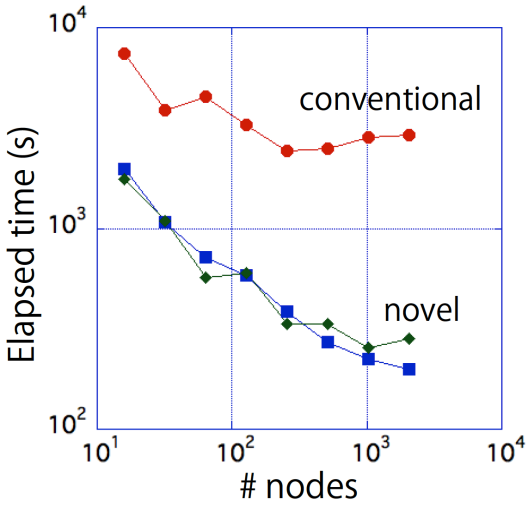

Fig. 1 Strong scaling benchmark on OFP of Eigenkernel, our generalized eigenvalue problem solver with the conventional (ScaLAPACK) solver (circle) and the two novel solvers of 'D'(square) and 'F'(diamond).

Figure 2 shows the strong scaling benchmark of ELSES with upto 2048 nodes [9]. The simulated system is a crystalline diamond sample with *N*=524,288 atoms in the simulation cell. The MPI communication time and barrier time are plotted, as well as the total time. The barrier time includes the time to wait for other processors. We found that the strong scaling property of the total time stems from the fact that the communication and barrier times consume only tiny fractions of the total time, as in the cases of the K computer (*1).

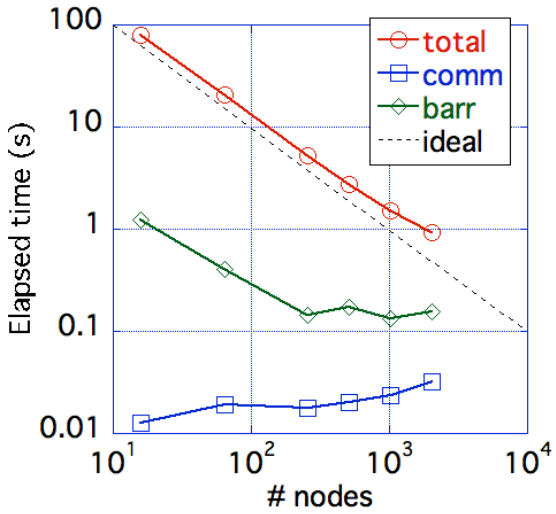

Fig. 2 Analysis of the elapsed time in the strong scaling benchmark of ELSES on OFP. The calculations were performed by the setting (B) in the job script file. The total time, the MPI communication time and the barrier time are plotted as 'total', 'comm' and 'barr', respectively. The ideal scaling line is drawn for eye guide.

Figure 3 shows the strong scaling benchmark of LDC-DFT with up to 512 nodes [24]. The total system contains liquid water consisting of 648 atoms in the cubic cell of side length 19.0 Å. The total system Ω was divided into 4×4×4 = 64 core domains $\Omega_{0\alpha}$.

5-2 Development of novel methods

Novel numerical methods were investigated so as to use in next-generation large-scale simulations; (i) linear-algebraic solver for intermediate eigenpairs in generalized

eigenvalue problems based on Sylvester's theorem of inertia [1], (ii) parallel block-Jacobi SVD solver [2], and (iii) efficient sapling method for numerical contour integral [3], (iv) an auto-optimization technique for contour-integral method for intermediate eigenpair calculation by Hoshi, Wang and coworkers (manuscript in preparation).

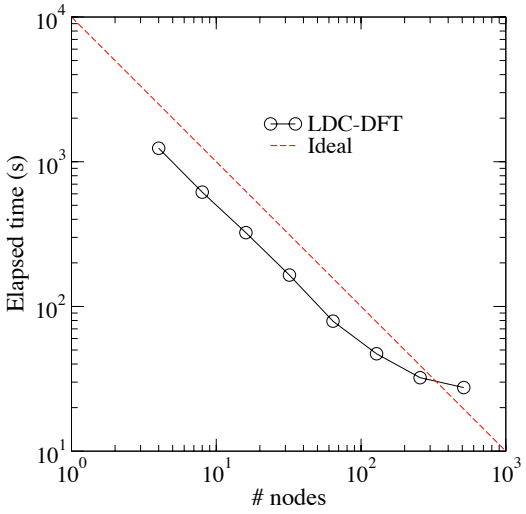

Fig. 3 Strong scaling benchmark of LDC-DFT on OFP. The ideal scaling line is drawn for eye guide.

5-3 Application I: Organic material for flexible device

Large-scale electronic state calculations were carried out by ELSES and EigenKernel for organic materials. Organic materials form the foundation of flexible device, one of next-generation Internet of Things (IoT) products. Since complicated disordered structure is crucial for device properties, such as device performance or device lifetime, large-scale simulations in 10-100 nm scales are required. Figure 4(a) shows such a flexible device (transistor) fabricated in Yamagata University (See (*4) at the end of this section). Figure 4(b) shows our result of semi-localized π electronic states in disordered pentacene thin-film, a proto-typical system of organic transistor. [7] The research is collaborated by H. Matsui (Yamagata University), an experimentalist. The semilocality of wavefunction is found in Fig.4(b).

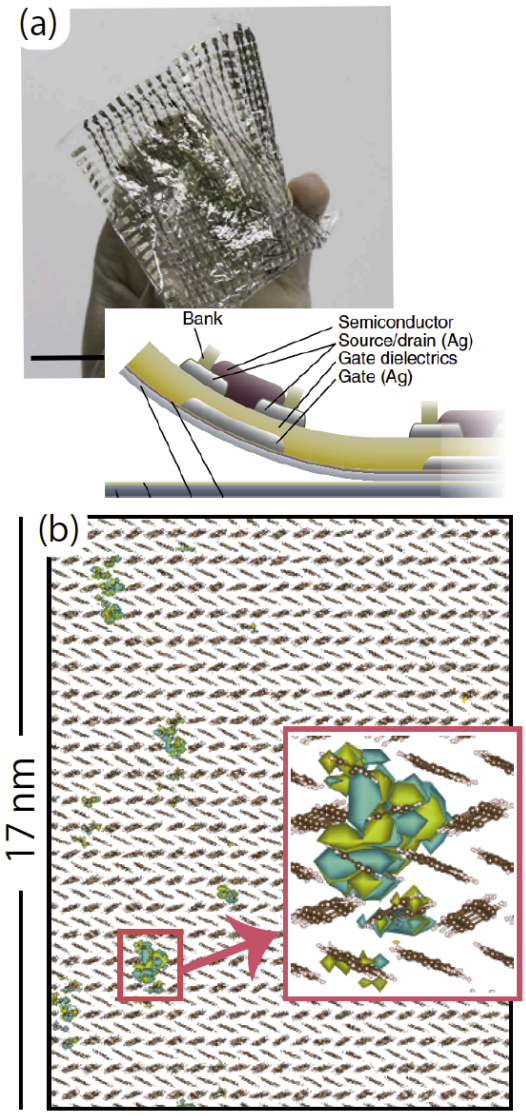

Fig. 4 (a) Example of organic device by Yamagata Univ. (*4) (b) Semi-localized π state of pentacene thin-film [x]

Here 'semi-locality' means that the wavefunction is extended among several tens of molecules but not in the whole system. Such semi-locality is crucial for the device performance and agree quantitatively with the experimental data of Electron Spin

Resonance (See (*5) at the end of this section).

(*4) K. Fukuda, et al., Fully-printed high-performance organic thin-film transistors and circuitry on one-micron-thick polymer films, Nature Comm. 5, 4147 (2014).
(*5) H. Matsui, et al., Distribution of Localized States from Fine Analysis of Electron Spin Resonance Spectra in Organic Transistors, Phys. Rev. Lett. 104, 056602 (2010).

5-4 Application II: Combined study between HPC (large-scale simulation) and data science

A combined study between HPC (large-scale simulation) and data science was carried out for disordered organic polymer, Poly(para-phenylene ethynylene) (PPE). [25, 26] We simulated 40,000 polymer samples with 1200 atoms. The spatial extent (participation ratio) of electronic wavefunctions were used as descriptors and the principal component analysis was carried out, as shown in Fig. 5. We found that the spatial extention of π wavefuntion, thus the device performance, is determined by the two structural factors of the polyer type (linear or zigzag) and the disorder of dihedral angles between benzene rings. The observation is consistent to the experimental result (*6).

It is noted that the principal component analysis was carried out by our Python code. Since Oakforest-PACS supports Python with various modules, the data scientific analysis can be carried out easily.

(*6) J. Terao, et al., Design principle for increasing charge mobility of π-conjugated polymers using regularly localized molecular orbitals, Nature Comm. 4, 1691 (2013).

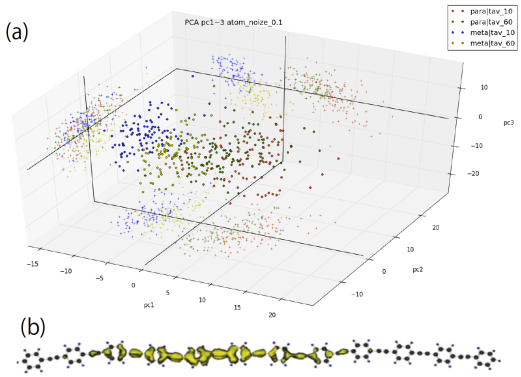

Fig. 5 (a) Principal component analysis of organic polymer with the descriptor of spatial extent of wavefuntions. The 40,000 polymer samples are plotted on the three dimensional data space by the first, second and third principal components. (b) A close-up of a calculated polymer. The drawn region is about 12 % of the whole polymer.

5-5 Application III: Research for origin of life

In this origin of life research field, quantum simulation reproducing alkaline hydrothermal vent environments is strongly required because the chemical reactions occurred there are believed to have achieved the emergence of primitive metabolism and cell replication on the early ocean. However, the functions of transition metal catalysts such as iron sulfides and the spatial gradients of temperature and pH must essentially be reproduced for the simulation.

Here, we tried to model the alkaline hydrothermal vent environment as shown in Fig. 6 and perform *ab initio* molecular dynamics (MD) simulation using LDC-DFT code [4, 12, 24]. The total system was

divided into 2×2×18 = 72 core domains. With the number of division we set, it took more than 30 seconds per 1 MD step, so long time MD calculation is not realistic. We thus must further proceed with optimization for such as the number of divisions. On the other hand, we obtained information on electronic states (such as potential energies and atomic forces) for a lot of local minimum atomic configurations. By using these information for parameter fitting of Density functional tight-binding method and Artificial neural network atomic potential, we are also preparing such methods with lower calculation cost as well as LDC-DFT [28].

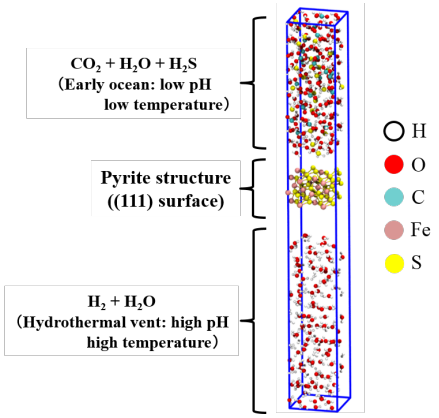

Fig. 6 Initial atomic configuration of the system consisting of 996 atoms in the rectangular cell of (L_x, L_y, L_z) = (10.828, 10.828, 100) Å ^{3}. This model consists of upper "early ocean" domain ($24CO_2$ + $24H_2S$ + 114 H_2O), pyrite slab domain (consisting of 72 atoms), and bottom "hydrothermal vent" domain ($48H_2$ + 114 H_2O).

5-6 International workshop

An half-day workshop 'Development of next-generation quantum material research platform' (Next QUMAT2017) was organized for collaboration mainly between Japan and USA and held on 4. Dec. 2017 at University of Tokyo [21]. The workshop contains one plenary talk by A. Nakano [11] and many other talks in the present members. [12-14] We shared many issues for simulation methods and target materials between Japan and US on next-generation supercomputers. We found that ELSES and LDC-DFT are complementary; ELSES can handle larger systems but is limited in its accuracy because of the use of modeled (tight-binding) theory, while LDC-DFT is based on first principle. Therefore, the combined studies between the two method will be fruitful. In particular, we found that an important topic is large-scale non-adiabatic quantum molecular-dynamics simulation, in which electronic states are excited. We started to collaborate in the topic and will discuss further. [6]

6. Progress of FY2017 and Future Prospects

In FY2017, we optimized our simulation codes on OFP for good strong scaling. After that we started a productive simulation on OFP. In addition, we started a collaboration between Japan and US for large-scale simulation. Our progress will be reflected on several projects, such as Priority issue 7 of the post-K computer.

7. List of Publications and Presentations

(1) Journal Papers

[1] D. Lee, T. Hoshi, T. Sogabe, Y. Miyatake, S.-L. Zhang, 'Solution of the k-th eigenvalue problem in large-scale electronic structure calculations', submitted; Preprint: http://arxiv.org/abs/1710.05134

[2] S. Kudo, Y. Yamamoto, M. Bečka and M. Vajteršic, 'Performance analysis and optimization of the parallel one-sided block Jacobi SVD algorithm with dynamic ordering and variable blocking', Concurrency and Computation: Practice and Experience, Vol. 29, Issue 9, 24pp (2017).

[3] H. Kohashi, K. Sugita, M. Sugihara and T. Hoshi, 'Efficient methods for computing integrals in electronic structure calculations', JSIAM Letters 9, pp.81-84 (2017).

[4] K. Shimamura, F. Shimojo, A. Nakano, and S. Tanaka, 'Meteorite impacts on ancient oceans opened up multiple NH_3 production pathways', Phys. Chem. Chem. Phys. 19, pp. 11655-11667 (2017).

[5] (Japanese journal) 島村孝平，下條冬樹，中野愛一郎，最近の研究から「非平衡不規則系における化学反応の理解に向けた第一原理分子動力学手法の開発」分子シミュレーション研究会会誌「アンサンブル」19, pp. 123-129 (2017).

(2) Conference Papers

(None)

(3) Oral Presentations

[6] (symposium talk, planned) T. Hoshi, 'Organic device material research by the combination of large-scale massively-parallel electronic state calculation and data-driven science', in Materials Genome Towards Exascale, Spetses island, Greece, organized by R. Kalia, A. Nakano, P. Vashishta (USC) and E. Kaxiras (Harvard), Greece, 10-15, June, 2018.

[7] (Japanese) 星 健夫，安部 友樹也，桑田亨成，角田皓亮，藤田貴敏，松井弘之，乱れたペンタセン薄膜系の大規模電子状態計算，応用物理学会春季学術講演会，大阪，2018 年 3 月 17−20 日.

[8] (symposium talk) D. Lee, T. Hoshi, T. Sogabe, Y. Miyatake, and S.-L. Zhang, 'Solving the k-th eigenvalue problem in large-scale electronic structure Calculations', Symposium 'Innovative methods for high performance iterative solvers' in 18th SIAM Conference on Parallel Processing for Scientific Computing (SIAM-PP18), Tokyo, 7-10, Mar. 2018.

[9] (symposium talk) T. Hoshi, '100-nano-meter-scale electronic structure calculation for organic device materials', Symposium 'Large-scale Electronic Structure Calculations', in 18th SIAM Conference on Parallel Processing for Scientific Computing (SIAM-PP18), Tokyo, 7-10, Mar. 2018.

[10] T. Hoshi, 'Eigenvalue computation and related numerical methods for large-scale quantum material simulations ', International Workshop on Eigenvalue Problems: Algorithms; Software and Applications, in Petascale Computing (EPASA2018), Tsukuba, 5-6, Mar. 2018.

[11] A. Nakano, 'Quantum and reactive molecular dynamics simulations on the next-generation US supercomputers', Next QUMAT2017, Tokyo, 4, Dec. 2017 [21].

[12] K. Shimamura, 'Application of ab initio molecular dynamics simulation to investigate the origin of life', Next QUMAT2017, Tokyo, 4, Dec. 2017 [21].

[13] *K. Nakajima*, 'Application development framework for manycore

architectures in Post-K/Post-Moore era', Next QUMAT2017, Tokyo, 4, Dec. 2017 [21].

[14] T. Hoshi, 'Development of next-generation quantum material research platform', Next QUMAT2017, Tokyo, 4, Dec. 2017 [21].

[15] (invited talk, Japanese) 星健夫, 「100 ナノメートルスケール電子状態計算とデータ科学による有機材料研究」, 山形大学有機エレクトロニクス研究センター講演会,山形, 2017 年 10 月 18 日.

[16] (symposium talk, Japanese) 星健夫, 「大規模電子状態計算と機械学習の融合による有機材料研究」, 日本物理学会シンポジウム 「理論による表面・界面・ナノ構造の微視的構造と物性の予測：現状と展望」, 岩手大学, 2017 年 9 月 21-24 日.

[17] (invited talk, Japanese) 星健夫, 「大規模量子 MD とデータ科学による有機デバイス材料研究」, 薬物 - 標的親和性計算の新潮流 ～古典 MD から量子 MD へ～, 情報計算化学生物学会,大阪, 2017 年 9 月 1 日.

[18] (Japanese) 星健夫, 安部友樹也, 大平健太郎, 福島孝治, 藤田貴敏, 望月祐志,「大規模電子状態計算とデータ科学による有機デバイス材料研究」, 日本応用物理学会学術講演会, 福岡, 2017 年 9 月 5-8 日.

[19] MD ALAM, T. Fujita, T. Hoshi, 'Excited-state calculation for large size C60/Pentacene interface based on the fragment molecular orbital method', JSAP Meeting, Fukuoka, 5-8., Sep. 2017.

[20] (symposium talk) T. Hoshi, ' Extremely large quantum material simulations with novel linear algebraic algorithms and massively parallel supercomputers ', Mini symposium: Computing bulks of inner eigenpairs of large sparse matrices: from application and algorithms to performance and software, The Platform for Advanced Scientific Computing Conference (PASC17), Switzerland, 26-28., Jun. 2017.

(4) Others

[21] (organized workshop) 'Development of next-generation quantum material research platform' (Next QUMAT2017) organized by T. Hoshi, A. Nakano, J. Iwata, *K. Nakajima*, University of Tokyo, 4. Dec. 2017, sponsored by the present project and Priority issue 7 of the Post-K computer. URL: https://sites.google.com/site/nextqumat2017/

[22] (poster) T. Hoshi, H. Imachi, Y. Abe, K. Oohira, K. Hukusima, 'Principal component analysis with large-scale electronic state calculation for ultra-flexible device materials', 18th SIAM Conference on Parallel Processing for Scientific Computing (SIAM-PP18), Tokyo, 7-10, Mar. 2018.

[23] (poster) T. Hoshi, T. Fukumoto, T. Fukaya, Y. Yamamoto, 'Analysis and prediction of the performance in generalized eigenvalue solvers on Oakforest-PACS ', International Workshop on Eigenvalue Problems: Algorithms; Software and Applications, in Petascale Computing (EPASA2018), Tsukuba, 5-6, Mar. 2018.

[24] (poster) K. Shimamura, 'Large scale ab initio calculation using LDC-DFT algorithm on many-core processor architectures', HPC Asia 2018, Tokyo, 28-31, Jan. 2018.

[25] (poster, Japanese)(学生優秀発表賞) 大平健太郎, 星健夫, 福島孝治, ‘データ科学と大規模電子状態計算による有機高分子探索’, 第 31 回分子シミュレーション討論会, 金沢, 2017 年 11 月 29 日-12 月 1 日.

[26] (poster) T. Hoshi, H. Imachi, K. Oohira, Y. Abe, and K. Hukushima, 'Principal component analysis with electronic wavefunctions for exploration of organic polymer device materials', International Meeting on High-Dimensional Data-Driven Science, Kyoto, 10-13, Sep. (2017).

[27] (poster) 島村孝平, 下條冬樹, 中野愛一郎, ‘非平衡不規則系で起こる化学反応の理解を目的とした第一原理分子動力学法の開発とその応用’, 第 2 回 CDMSI(ポスト「京」重点課題（7）)研究会, 東京, 2017 年 07 月 11 日-12 日.

[28] (poster) 島村孝平, ‘アルカリ性深海熱水孔環境を通して考える生命の起源’, 量子生命科学研究会第２回学術集会, 東京, 2018 年 05 月 10 日.