Development of Bayesian optimization tool and its applications in materials science

Ryo Tamura[a,b], Yuichi Motoyama[c], and Kazuyoshi Yoshimi[c]

[a]*International Center for Materials Nanoarchitectonics (WPI-MANA)*
National Institute for Materials Science
1-1 Namiki, Tsukuba, Ibaraki 305-0044, Japan
[b]*Graduate School of Frontier Sciences, The University of Tokyo*
5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8561, Japan
[c]*Institute for Solid State Physics, University of Tokyo, Chiba 277-8581, Japan*

1 Introduction

Recently, materials informatics has established an important position in materials science [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11]. As a feature of materials science, it is not always easy to obtain materials data whether by experiments or simulations. In contrast, when general machine learning and artificial intelligence techniques are applied to materials data, it is necessary to use a large amount of data to gather meaningful results. Therefore, to achieve success in materials informatics research, it is necessary to use some ingenuities in treating small-data problems. One strategy to overcome small-data problems is to start with a small amount of data and gradually increase the amount of data with the help of machine learning. This is called black-box optimization. One of the typical techniques is Bayesian optimization.

In this paper, we introduce a Bayesian optimization method that can effectively select the next candidate for an experiment or simulation with the help of machine learning. We also give examples of its application in materials informatics. In Bayesian optimization, Gaussian processes are used as a machine-learning surrogate model, which is constructed to predict the target material properties using already obtained materials data. Based on the predic-tions, the next candidate for the experiment or simulation is selected to improve the materials properties. By repeating these processes, *training* and *selection* using machine learning, materials with better properties can be found using the fewest number of experiments or simulations. Recently, many successful examples of Bayesian optimization have been reported in materials science [12, 13, 14, 15, 16].

2 Algorithm of Bayesian optimization

2.1 Problem setting

Let us explain an algorithm for Bayesian optimization. In materials research, the variables that describe materials are called descriptors, and include compositions, structures, processes, and simulation parameters. A descriptor vector with dimension d is denoted by $\mathbf{x} \in \mathrm{Re}^d$. Additionally, the materials property is expressed as y which is obtained through experiments or simulations when $\mathbf{x}$ is given. Evidently, the relationship between $\mathbf{x}$ and y cannot be explicitly expressed as a simple function in materials science. Thus, in materials informatics, we replace this relationship with a machine learning surrogate model. That is, the machine-learning surrogate model $f_{\mathrm{ML}}(\mathbf{x})$

is introduced, and y for any $\mathbf{x}$ is approximately obtained as follows:

$$y \simeq f_{\mathrm{ML}}(\mathbf{x}). \tag{1}$$

Using such a machine learning surrogate model, Bayesian optimization can find $\mathbf{x}$ with the better material properties y among the many candidates of $\mathbf{x}$. For example, a problem may be set up for Bayesian optimization as follows:

- There are N candidate materials expressed by $\{\mathbf{x}_i\}_{i=1,\ldots,N}$.

- Among $\{\mathbf{x}_i\}_{i=1,\ldots,N}$, we search for materials with a better material property value y. In addition, the number of experiments or simulations to obtain y is minimized as much as possible.

- Experiments or simulations for M candidate materials have been completed, and M material properties have been obtained. Currently, we have dataset $D = \{\mathbf{x}_k, y_k\}_{k=1,\ldots,M}$.

- We select the next candidate material expressed by $\mathbf{x}_{M+1}$ using $f_{\mathrm{ML}}(\mathbf{x})$. That is, the surrogate model is trained using the training data D, and the potential of a material is evaluated by using the trained model. To evaluate the potential, the acquisition function is introduced. Here, the Bayesian optimization uses a Gaussian process as a surrogate model.

- After evaluating the value of y_{M+1} on the selected material $\mathbf{x}_{M+1}$ by experiments or simulations, the $M+1$-th pair $(\mathbf{x}_{M+1}, y_{M+1})$ is added to D. Then, the next candidate material is selected by using the updated surrogate model.

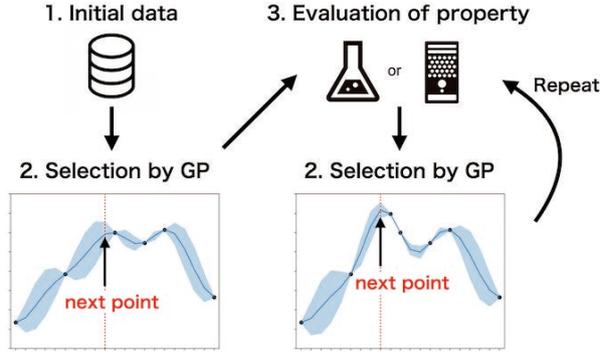The Bayesian optimization cycle is shown in Fig. 1.

Figure 1: Flow of Bayesian optimization. In step 2, the Gaussian process (GP) regression is trained by the already obtained data and the next point is selected based on the acquisition functions. In step 3, the materials property of the selected material is evaluated by experiments or simulations. By repeating of steps 2 and 3, we select better materials.

2.2 Gaussian process

In the Gaussian process which is used as the surrogate model in Bayesian optimization, when the training data D are given, the conditional probability distribution for y at $\mathbf{x}$ is given by

$$P(y|D) = N(\mu(\mathbf{x}), \sigma(\mathbf{x})). \tag{2}$$

Here, $\mu(\mathbf{x})$ is the mean of the predicted value by the Gaussian process regression when $\mathbf{x}$ is the input, and the variance $\sigma(\mathbf{x})$ is the uncertainty, when $D = \{\mathbf{x}_k, y_k\}_{k=1,\ldots,M}$. These are defined as follows:

$$\mu(\mathbf{x}) = \mathbf{k}^\top (K + \lambda I)^{-1} \mathbf{y}, \tag{3}$$

$$\sigma(\mathbf{x}) = k(\mathbf{x}, \mathbf{x}) + \lambda - \mathbf{k}^\top (K + \lambda I)^{-1} \mathbf{k}, \tag{4}$$

where λ is the hyperparameter, I is the identity matrix, and $\mathbf{y} = (y_1 \cdots y_M)^\top$. In addition,

$\mathbf{k}$ and K are defined as follows:

$$\mathbf{k} = \begin{pmatrix} k(\mathbf{x}_1, \mathbf{x}) & \cdots & k(\mathbf{x}_M, \mathbf{x}) \end{pmatrix}^\top, \quad (5)$$

$$K = \begin{pmatrix} k(\mathbf{x}_1, \mathbf{x}_1) & \cdots & k(\mathbf{x}_1, \mathbf{x}_M) \\ \vdots & \ddots & \vdots \\ k(\mathbf{x}_M, \mathbf{x}_1) & \cdots & k(\mathbf{x}_M, \mathbf{x}_M) \end{pmatrix}, \quad (6)$$

where $k(\mathbf{x}_i, \mathbf{x}_j)$ is a Gaussian kernel function,

$$k(\mathbf{x}_i, \mathbf{x}_j) = \exp\left[-\frac{1}{2\gamma^2} \|\mathbf{x}_i - \mathbf{x}_j\|^2 \right]. \quad (7)$$

Here, γ is a hyperparameter.

2.3 Acquisition functions

To evaluate the potential of the material, acquisition functions are introduced. There are many types of acquisition functions, such as the probability of improvement (PI) [17], the expected improvement (EI) [18], and the Thompson sampling [19]. Firstly, the PI represents the probability of exceeding the current maximum value of the already obtained y: $y_{\max} = \max_k y_k$. This score is defined as follows:

$$\begin{aligned} \mathrm{PI}(\mathbf{x}) &= \mathbb{P}\left[y(\mathbf{x}) > y_{\max} \right] \\ &= F\left(t(\mathbf{x}) \right), \end{aligned} \quad (8)$$

where $F(\cdot)$ is the cumulative distribution function of $\mathcal{N}(0,1)$ and $t(\mathbf{x}) = (\mu(\mathbf{x}) - y_{\max})/\sigma(\mathbf{x})$. Here, $\mu(\mathbf{x})$ and $\sigma(\mathbf{x})$ are obtained using the Gaussian process.

Next, the EI is the expected value of how much $y_{\max}$ updates when $\mathbf{x}$ is observed. This is defined as

$$\begin{aligned} \mathrm{EI}(\mathbf{x}) &= \mathbb{E}\left[\max(y(\mathbf{x}) - y_{\max}, 0) \right] \\ &= \sigma(\mathbf{x})\left[t(\mathbf{x}) F(t(\mathbf{x})) + f(t(\mathbf{x})) \right], \end{aligned} \quad (9)$$

where $f(\cdot)$ denotes the probability density of $\mathcal{N}(0,1)$.

Finally, we consider Thompson sampling, and the acquisition function using Thompson sampling is as follows:

$$\mathrm{TS}(\mathbf{x}) = \mathbf{w}^{*\top} \phi(\mathbf{x}). \quad (10)$$

Here, a coefficient vector $\mathbf{w}^*$ was sampled according to the posterior distribution in the Gaussian process. $\phi(\mathbf{x})$ is a random feature map and $\phi(\mathbf{x})^\top \phi(\mathbf{x})$ expresses the Gaussian kernel, approximately.

In Bayesian optimization, the material expressed by $\mathbf{x}$ with the largest value of the acquisition function is selected as the next candidate for the experiment or simulation. By using these acquisition functions, we can incorporate into the selection process not only the predicted value $\mu(\mathbf{x})$ but also information on the uncertainty of the prediction $\sigma(\mathbf{x})$. This makes the selection more effective than if the next candidate is randomly selected.

3 PHYSBO package

To realize fast and scalable Bayesian optimization, we developed PHYSBO (optimization tools for PHYSics based on Bayesian Optimization) [20, 21] which is a Python library developed under the support of the "Project for advancement of software usability in materials science" by the Institute for Solid State Physics, University of Tokyo. In PHYSBO, a random feature map, Cholesky decomposition, and Thompson sampling are used to accelerate the calculations required for Bayesian optimization. Note that PHYSBO is an updated version of COMBO [22] and includes some new functions.

The following are the features of PHYSBO.

- PHYSBO can be used to find better solutions for both single and multi-objective optimization problems.

- At each cycle in Bayesian optimization, a single proposal or multiple proposals can be selected.

- Parallel calculation can be performed. The results of parallelization performance are shown in Fig. 2.

- Detailed manual is provided.

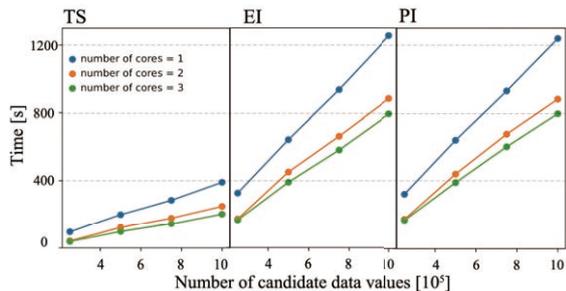PHYSBO is available from `https://github.com/issp-center-dev/PHYSBO`.

Figure 2: Selection time as a function of the number of candidate data values for different scores depending on the number of cores. As the number of cores increases, the selection times steadily decrease. Computing time is measured using CPU: 2.3 GHz Quad Core i7. Panels reproduced with permission from Ref. [20].

4 Examples in materials science

In this section, some results obtained by performing Bayesian optimization in materials science are introduced. In Secs. 4.1-4.3, single-objective problems are the focus for discussion. Sections 4.1 and 4.2 consider the cases where Bayesian optimization is combined with experiments, and in Sec. 4.3, the results where simulation data is used are introduced. In addition, the multi-objective optimization results are presented in Sec. 4.4.

4.1 Composition optimization for Li ion conductivity

Li-ion conductive oxides are attractive electrolytes for Li-ion secondary batteries owing to their stability and low toxicity. To increase the Li-ion conductivity of a material, doping of additives or mixing heterogeneous materials have been considered. However, it is difficult to determine the optimal chemical compositions us-ing experimental methods. This is because it is practically impossible to cover all compositions because of the large number of possibilities in the search space. Thus, to reduce the number of experiments, Bayesian optimization is a useful tool.

To investigate the efficiency of Bayesian optimization, we attempted to optimize the composition of the ternary oxide solid electrolyte Li_3PO_4-Li_3BO_3-Li_2SO_4 [23, 24]. The mixed material was stable under atmospheric conditions and could be combined by sintering with both positive and negative electrode materials. As the initial traial, we prepared 15 samples with a composition ratio interval of 25%. Using these 15 data points, a ternary component heat map of Li-ion conductivity was drawn by the Gaussian process regression, which is shown in Fig. 3 (a). We can see that in samples 7 and 8, the Li-ion conductivity increased. Bayesian optimization was performed using this as the initial state, and 10 samples were further synthesized. Using all data, the ternary component heat map is shown in Fig. 3 (b). We obtained an optimum polycrystalline material with a composition of 25:14:61 (mol %) for Li_3PO_4-Li_3BO_3-Li_2SO_4, which is three times the Li-ion conductivity (4.9×10^{-4} S/cm at 300 °C) compared with the highest case in binary compounds.

4.2 Process optimization for gas atomization

The use of metal 3D printing is rapidly advancing in aerospace engines, and low-cost production and a ready supply of superalloy powders have become essential in the industry. In particular, for high-pressure turbine disks, which are important components of engines, it is necessary to produce high-quality superalloy powders with a high sphericity, uniform structure, high yield, and low cost. Most of these powders are produced by gas atomization, and the desired powder is produced using the industrial equipment used in the manufacturing process.
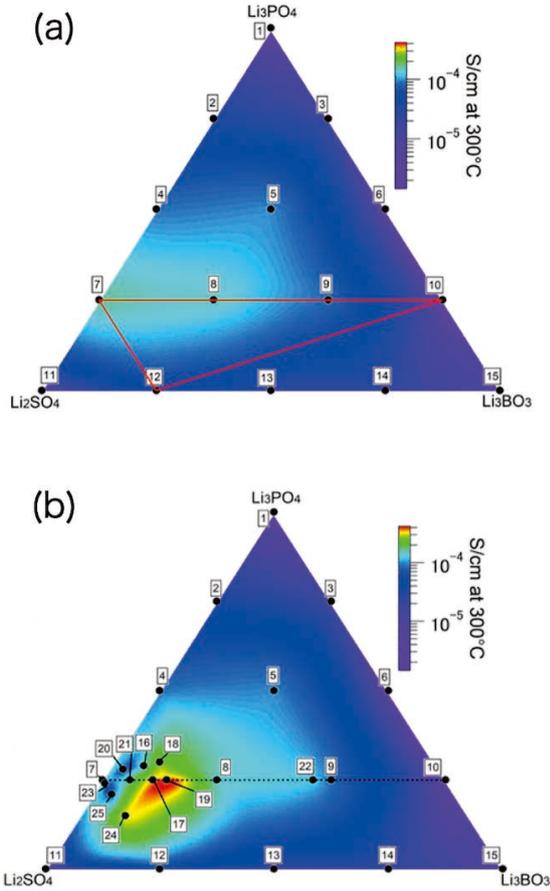
Figure 3: Heat maps of the Li-ion conductivity interpolated by the Gaussian process for (a) initial 15 data and (b) 25 data after Bayesian optimization. Panels reproduced with permission from Ref. [23].

To do this it is necessary to optimize multiple process conditions such as the melting temperature of the metal and gas pressure. Even with the expertise of specialists, this optimization task requires enormous cost, time, and human resources.

In this study, Bayesian optimization was used to optimize the gas atomization process without the use of expert knowledge or historical data (Fig. 4 (a)) [25]. A Ni-Co based superalloy for high-pressure turbine disks with excellent heat resistance was chosen for investigation. The target powder specifications are a size of 53 μm or less, which is suitable for turbine disk fabrication. The process of maximizing the yield of these powers was explored. Melting temperature of the metal and gas pressure were chosen from the various gas atomization processes as a focus for the optimization.

First, as initial data, superalloy powders were produced by three different processes, and the yield under 53 μm was evaluated by sieve classification. Using these data as the training data, a Gaussian process was trained to predict the yields of the candidate processes and estimate the errors. Here, in the candidates processes, the melting temperature is $T = 50 \times i + 1400$ [°C] $(i = 1, ..., 5)$ and the gas pressure is $P = i + 4$ [MPa] $(i = 1, ..., 5)$. Based on the results of the prediction, we selected the process with the highest acquisition function and conducted powder production using the selected process. By repeating this optimization loop thrice, we succeeded in determining the process conditions to obtain the target powder finer than 53 μm in approximately 78% yield, when it is typically a 10-30% yield. The yields obtained from these six trials and unit costs of the raw materials are shown in Fig. 4 (b). By performing Bayesian optimization, the yield was improved, and the unit cost of the raw material was cheaper. The superalloy powder produced by the process found by Bayesian optimization succeeded in reducing the cost by approximately 72% compared with commercially available powder (values estimated from the unit cost of raw materials).

4.3 Model parameter estimation

In condensed matter physics, knowing the effective model of a material is one method used to understand the target material [27, 28]. Bayesian statistics are useful for solving inverse problems [29], in which an effective model is derived that explains the given simulation results or experimental results. Based on Bayesian statistics, plausible model parameters in the effective model that explain the given physical quantity are determined by
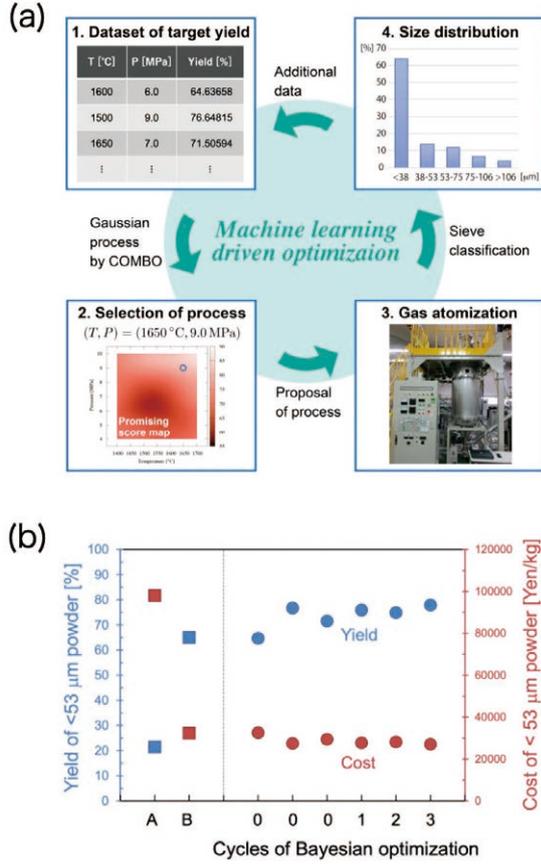
Figure 4: (a) Cycle of Bayesian optimization in process optimization of gas atomization. (b) Yields and the unit cost depending on the cycles. Values obtained for commercial gas atomized powders prepared by Company A and Company B [26] are also plotted. Panels reproduced with permission from Ref. [25].

maximizing the posterior distribution, which is defined as

$$P(\mathbf{x}|y^{\text{target}}) \propto \exp[-E(\mathbf{x})], \qquad (11)$$

where the energy function $E(\mathbf{x})$ as a function of model parameters $\mathbf{x}$ is given by

$$E(\mathbf{x}) = 1/(2\sigma^2)(y^{\text{target}} - y^{\text{cal}}(\mathbf{x}))^2 - \log[P(\mathbf{x})], \quad (12)$$

where y^{target} and $y^{\text{cal}}(\mathbf{x})$ are the sets of physical quantities given as targets and those calculated from the effective model characterized by $\mathbf{x}$, respectively. In addition, $P(\mathbf{x})$ is the prior distribution that expresses prior knowledge of the model parameters $\mathbf{x}$. To estimate an effective model, the maximizer of the posterior distribution should be determined. This problem can be solved using Bayesian optimization, where the value of y is generated from the posterior distribution.

As a sample problem, the estimation of magnetic interactions of the spin-1/2 Heisenberg chain model at 12 sites from a magnetization curve was set as a challenge. Here, three types of magnetic interactions that need to be estimated are considered: nearest neighbor J_1, next-nearest neighbor J_2, and 3rd-nearest neighbor J_3. Thus, the Hamiltonian can be written as

$$\mathcal{H} = \sum_{i=1}^{12} J_1 \mathbf{S}_i \cdot \mathbf{S}_{i+1} + J_2 \mathbf{S}_i \cdot \mathbf{S}_{i+2} + J_3 \mathbf{S}_i \cdot \mathbf{S}_{i+3}, (13)$$

where $\mathbf{S}_i$ denotes the vector of the spin operator at i-th site. In addition, a periodic boundary condition is imposed as follows: $\mathbf{S}_i = \mathbf{S}_{i+12}$.

In the model parameter estimation, firstly, the target magnetization curve $\{m^{\text{target}}(h_j)\}_{j=1,\dots,N_h}$ is given. Here, h_j is the magnetic field and N_h is the number of magnetizations with different h_j. Instead of the value of posterior distribution itself, we used the difference between the target and calculated magnetization curves as y in Bayesian optimization, which is given as

$$y = -\sum_{j=1}^{N_h}[m^{\text{target}}(h_j) - m(h_j; J_1, J_2, J_3)]^2, (14)$$

where $\{m(h_j; J_1, J_2, J_3)\}_{j=1,\dots,N_h}$ is the calculated magnetization curve when the interaction values (J_1, J_2, J_3) are given. By searching for (J_1, J_2, J_3) such that y is maximized, the difference between the target and calculated magnetization curves is minimized, and we can estimate the model parameters that describe the target magnetization curve.

We prepared the target magnetization curve using the Hamiltonian with $(J_1, J_2, J_3) = (1.0, 0.5, 0.3)$; that is, these values are the solution to the model estimation problem. We used $\mathcal{H}\Phi$ package [30] to calculate the magnetization curve, which is a quantum lattice solver that uses the exact diagonalization method. In Bayesian optimization, the search space is set as a grid in which each magnetic interaction is discretized between 0 and 2 in increments of 0.1 for J_1, J_2, and J_3. Thus, the total number of input candidates is $21^3 = 9261$. The best difference as a function of the number of cycles in Bayesian optimization is shown in Fig. 5. Using Bayesian optimization, we demonstrated that better parameters for an effective model can be found.
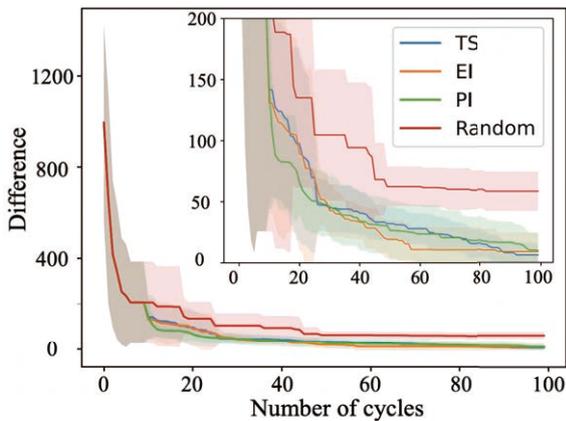

Figure 5: Best difference $(= -y)$ during the optimization as a function of the number of cycles. The results are compared with the results of random sampling. The inset is an enlarged view. Panels reproduced with permission from Ref. [20].

4.4 Materials screening for multi-objective optimization

Finally, the results of multi-objective optimization problems using Bayesian optimization are introduced. The purpose of multi-objective optimization is to find as many Pareto solutions as possible. Here, the Pareto solution is defined as the optimal solution obtained by varying the balance of the objective functions. A metric for determining whether better Pareto solutions are found is the dominated region, which is the area dominated by Pareto solutions in the objective function space. The target problem is materials screening, where we search for materials with a large bandgap and a large electronic dielectric constant. In general, these properties exhibit a trade-off tendency, in other words, as the bandgap increases, the electronic dielectric constant decreases. In order to address this problem, we used the semiconductor database from Ref. [31], and found the number of materials is 1277. First principles simulation results are presented in this database. We use the compositional descriptors generated by Magpie [32] as $\mathbf{x}$.

The number of Pareto solutions obtained by performing Bayesian optimization with multiple objectives as a function of the number of cycles is plotted in Fig. 6 (a). For Bayesian optimization, three types of scores were used for multi-objective optimization, called TS, EHVI, and HVPI. The results show that Bayesian optimization was able to find many more Pareto solutions than random sampling. The area of the dominated region was calculated for the results after 200 cycles. A large value of this area indicates that better Pareto solutions are found. Figure 6 (b) shows a violin plot for this area using 10 independent runs. Bayesian optimization produced a significantly larger dominated area than random sampling.

5 Summary

We introduced a Bayesian optimization technique and demonstrated its application in materials science. Bayesian optimization is a machine learning method that can effectively select the next experiments or simulations. The efficiency of Bayesian optimization is recognized in materials science, and thus, many
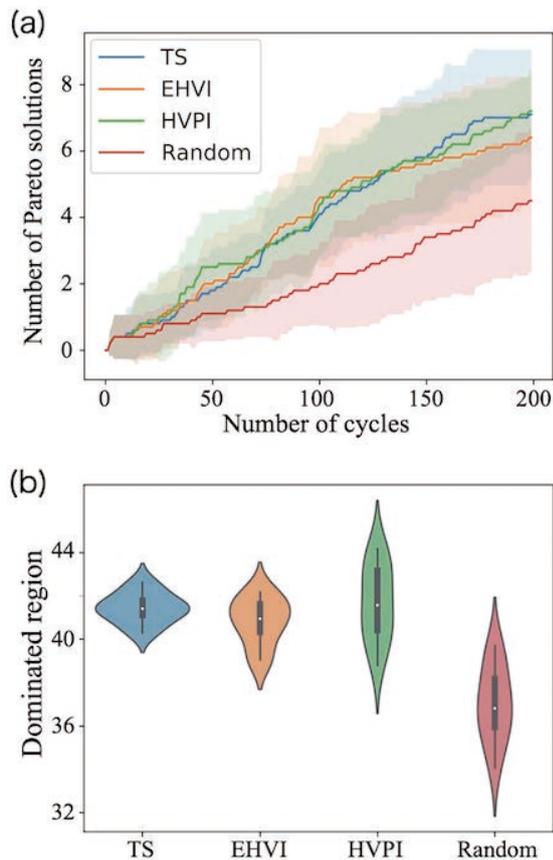
Figure 6: (a) Number of Pareto solutions found as a function of the number of cycles by PHYSBO for a multi-objective optimization problem. (b) Dominated region depending on the method when the number of cycles is 200. Panels reproduced with permission from Ref. [20].

problems can be solved using this technique in the future. In particular, using the PHYSBO package, massive parallel calculations can be performed using supercomputers, and we expect that more complicated problems can be solved.

Acknowledgements

References

[1] A. Seko, A. Togo, H. Hayashi, K. Tsuda, L. Chaput, and I. Tanaka, Phys. Rev. Lett. **115** (2015) 205901.

[2] S. Ju, T. Shiga, L. Feng, Z. Hou, K. Tsuda, and J. Shiomi, Phys. Rev. X **7** (2017) 021024.

[3] V. Stanev, C. Oses, A.G. Kusne, E. Rodriguez, J. Paglione, S. Curtarolo, and I. Takeuchi, npj Comput. Mater. **4** (2018) 1.

[4] K. Terayama, R. Tamura, Y. Nose, H. Hiramatsu, H. Hosono, Y. Okuno, and K. Tsuda, Phys. Rev. Mater. **3** (2019) 033802.

[5] T. Fukazawa, Y. Harashima, Z. Hou, and T. Miyake, Phys. Rev. Mater. **3** (2019) 053807.

[6] K. Kitai, J. Guo, S. Ju, S. Tanaka, K. Tsuda, J. Shiomi, and R. Tamura, Phys. Rev. Res. **2** (2020) 013319.

[7] R. Katsube, K. Terayama, R. Tamura, and Y. Nose, ACS Mater. Lett. **2** (2020) 571.

[8] R. Tamura, M. Watanabe, H. Mamiya, K. Washio, M. Yano, K. Danno, A. Kato, and T. Shoji, Sci. Technol. Adv. Mater. **21** (2020) 540.

[9] K. Terayama, M. Sumita, R. Tamura, and K. Tsuda, Acc. Chem. Res. **54** (2021) 1334.

[10] I. Ohkubo, Z. Hou, J. N. Lee, T. Aizawa, M. Lippmaa, T. Chikyow, K. Tsuda, and T. Mori, Mater. Today Phys. **16** (2021) 100296.

[11] R. Tamura, Y. Takei, S. Imai, M. Naka-hara, S. Shibata, T. Nakanishi, and M. Demura, Sci. Technol. Adv. Mater. Methods **1** (2021) 152.

[12] M. M. Noack, K. G. Yager, M. Fukuto, and G. S. Doerk, R. Li, and J. A. Sethian, Sci. Rep. **9** (2019) 1.

[13] R. Vargas-Hernández, Y. Guan, D. Zhang, and R. Krems, New J. Phys. **21** (2019) 022001.

[14] T. Yamashita, N. Sato, H. Kino, T. Miyake, K. Tsuda, and T. Oguchi, Phys. Rev. Mater. **2** (2018) 013803.

[15] F. Häse, L. M. Roch, C. Kreisbeck, and A. Aspuru-Guzik, ACS Cent. Sci. **4** (2018) 1134.

[16] K. Terayama, K. Tsuda, and R. Tamura, Jpn. J. Appl. Phys. **58** (2019) 098001.

[17] H. J. Kushner, J. Basic Eng. **86** (1964) 97.

[18] J. Mockus, V. Tiesis, and A. Zilinskas, Towards Global Optimization **2** (1978) 2.

[19] S. Q. Yahyaa and B. Manderick, The 23rd European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (2015).

[20] Y. Motoyama, R. Tamura, K. Yoshimi, K. Terayama, T. Ueno, and K. Tsuda, arXiv: 2110.07900 (2021).

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

[22] T. Ueno, T. D. Rhone, Z. Hou, T. Mizoguchi, and K. Tsuda, Mater. Discov. **4** (2016) 18.

[23] K. Homma, Y. Liu, M. Sumita, R. Tamura, N. Fushimi, J. Iwata, K. Tsuda, and C. Kaneta, J. Phys. Chem. C **124** (2020) 12865.

[24] M. Sumita, R. Tamura, K. Homma, C. Kaneta, and K. Tsuda, Bull. Chem. Soc. Jpn. **92** (2019) 1100.

[25] R. Tamura, T. Osada, K. Minagawa, T. Kohata, M. Hirosawa, K. Tsuda, and K. Kawagishi, Mater. Des. **198** (2021) 109290.

[26] C. Lehnert, B. Sitzmann, F. Pfahls, H. Franz, and M. Hohmann, The Special Steel **65** (2016) 10.

[27] R. Tamura and K. Hukushima, Phys. Rev. B **95** (2017) 064407.

[28] R. Tamura, K. Hukushima, A. Matsuo, K. Kindo, and M. Hase, Phys. Rev. B **101** (2020) 224435.

[29] R. Tamura and K. Hukushima, PLOS ONE **13** (2018) e0193785.

[30] M. Kawamura, K. Yoshimi, T. Misawa, Y. Yamaji, S. Todo, and N. Kawashima, Comput. Phys. Commun. **217** (2017) 180.

[31] A. Takahashi, Y. Kumagai, J. Miyamoto, Y. Mochizuki, and F. Oba, Phys. Rev. Materials **4** (2020) 103801.

[32] L. Ward, A. Agrawal, A. Choudhary, and C. Wolverton, npj Comput. Mater. **2** (2016) 16028.