Data-Analysis Software Framework 2DMAT and its Application to Experimental Measurements for Two-Dimensional Material Structures

An open-source data-analysis framework 2DMAT has been developed mainly for total-reflection high-energy positron diffraction (TRHEPD), a novel measurement technique for two-dimensional material structures. 2DMAT offers various analysis methods of optimization and Bayesian inference in parallel computations and realizes fast and reliable data analysis among personal computers and supercomputers. The data analysis was demonstrated for the structural analysis or the determination of the atomic positions of the Ge(001)-c4x2 surface. 2DMAT is general-purpose data analysis software and will be applied to other various advanced measurement experiments, such as reflection high-energy electron diffraction (RHEED), low-energy electron diffraction (LEED), low-energy positron diffraction (LEPD), surface X-ray diffraction (SXRD), and other problems.

We have developed an open-source framework 2DMAT [1] for the data analysis of measurement experiments, such as the structural analysis of two-dimensional (2D) MATerials [2]. 2D materials, ultra-thin films consisting of several atomic layers, are of great interest, for example, as a foundation of next-generation catalysts and devices since they express functions different from those of conventional (three-dimensional) materials. However, structural measurement techniques for 2D materials have not yet been established due to the difficulty of determining their atomic positions. The total-reflection high-energy positron diffraction (TRHEPD) [3] has rapidly progressed to overcome the above difficulty, at the Slow Positron Facility of the Institute Materials Structure Science, KEK. A conceptual diagram of the TRHEPD experiment is shown schematically in the left panel of Fig. 1. The experimental setup of TRHEPD is essentially the same as that for reflection high-energy electron diffraction (RHEED). TRHEPD is highly sensitive to the atomic positions at topmost and subsurface layers in sub-nanometer depth [4] since positron has positive charge and is repelled by the crystal potential energy of the bulk.

We demonstrated 2DMAT for analyzing TRHEPD experiments in five algorithms: Nelder-Mead optimization, grid search, Bayesian optimization, Bayesian inference by replica exchange Monte Carlo, and population annealing Monte Carlo methods. The program code, except Nelder-Mead optimization, is implemented in parallel computation and realizes fast and reliable data analysis among personal computers and supercomputers. 2DMAT requires sim-trhepd-rheed [5] for the analysis of the TRHEPD experiment. 2DMAT is a general-purpose data analysis framework, and the present version supports other advanced measurement experiments, such as RHEED, low-energy electron diffraction (LEED), low-energy positron diffraction (LEPD), and surface X-ray diffraction (SXRD), and other problems. The application studies with 2DMAT are found in the papers [1, 4, 6-10].

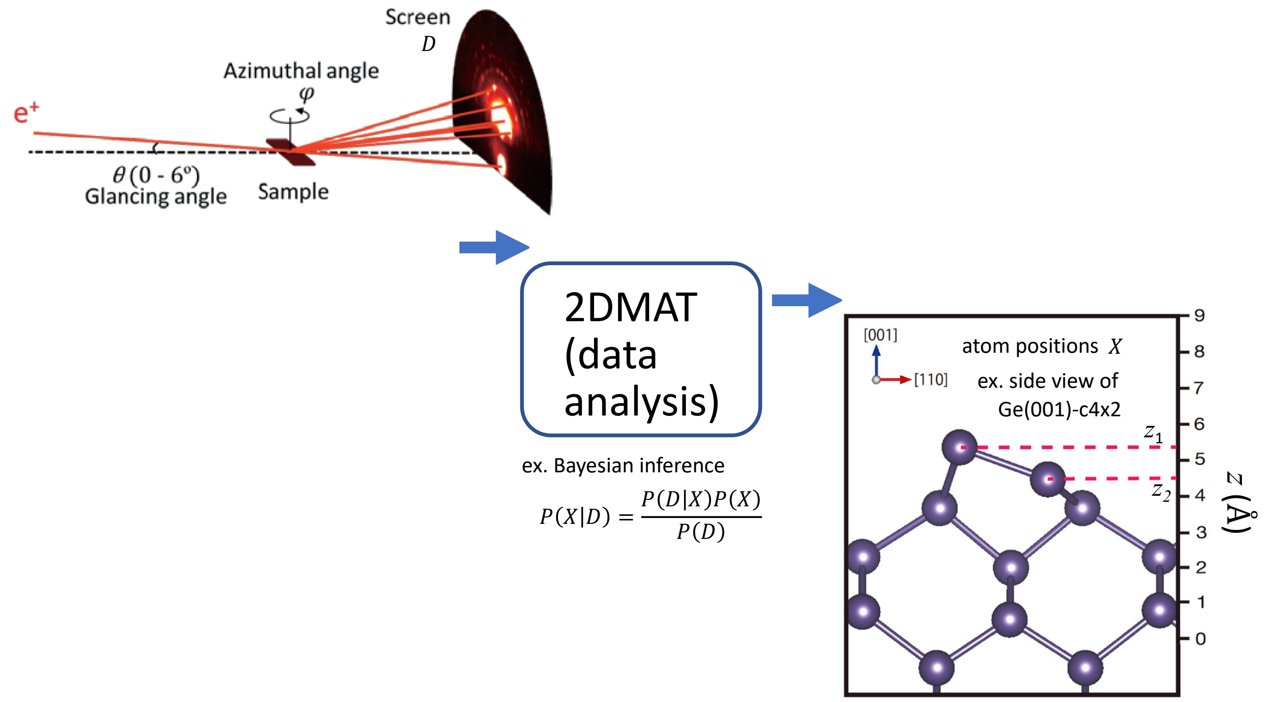

Figure 1: Workflow of the data analysis of the TRHEPD experiment by 2DMAT. (Left panel) Schematic figure of the TRHEPD experiment with diffraction data *D*. (Lower central panel) Bayesian inference with Bayes' theorem, a data analysis algorithm in 2DMAT. (Right panel) Example of 2D material: side view of the Ge(001)-c4x2 surface.

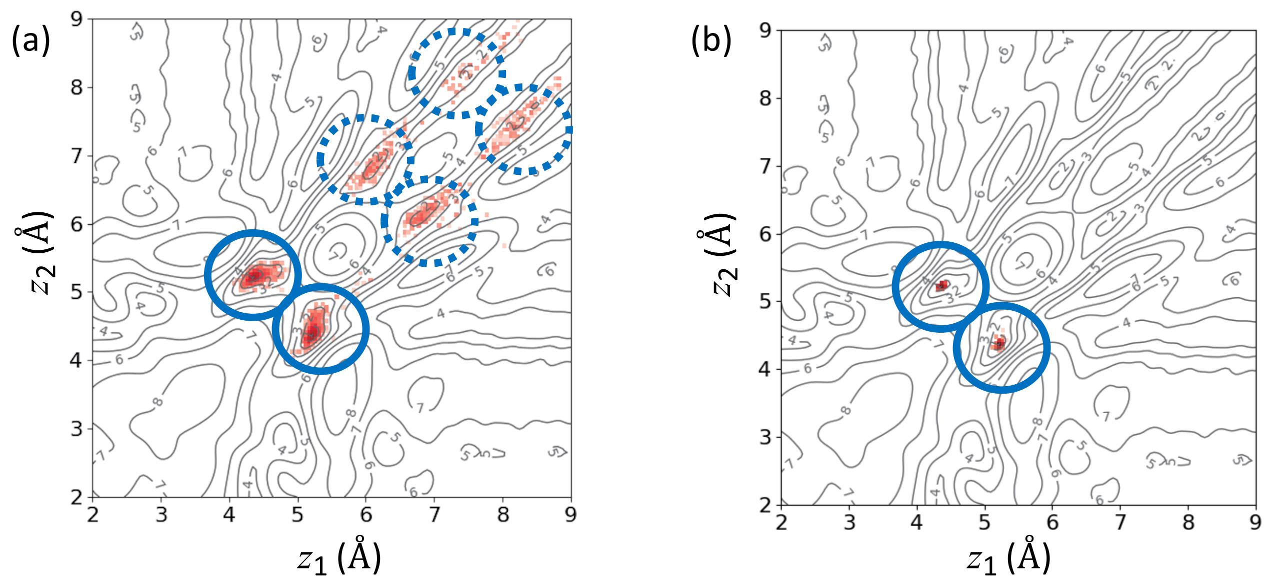

Figure 2: The histograms of the posterior probability density $P(X|D)$ for the atomic positions $X \equiv (z_1, z_2)$ of the Ge(001)-c4x2 surface [1]. The true solutions are indicated by the solid line circles, while the local solutions are indicated by the dashed line circles. The uncertainty parameter τ is set to be high and low values in (a) and (b), respectively [1]. The contour is drawn for the objective function $f(X)$ that represents the residual difference between the experimental and calculated diffraction data. The true solution(s) is(are) defined as $X^* \equiv \mathrm{argmin}_x f(X)$.

Hereinafter, the global analysis with the Monte Carlo methods [1] is focused on. A global analysis finds the solution without any guesses of researchers. Figure 1 shows the workflow of the data analysis. The data analysis is based on Bayes' theorem for calculating the posterior probability density, where the atomic positions *X* and the observed diffraction data *D* are vector quantities. The theory contains the uncertainty parameter τ that represents the measure of the uncertainty (or noise) in the observed diffraction data [1]. Figure 2 demonstrates the analysis result of the TRHEPD experiment for the Ge(001)-c4x2 surface shown in the right panel of Fig. 1. The *z* coordinates of the first and second surface atoms, z_1 and z_2, are chosen as the variable set $X \equiv (z_1, z_2)$. The calculated posterior probability density $P(X|D) = P(z_1,z_2|D)$ is drawn as a histogram in Fig. 2, where the uncertainty parameter τ is set to be high and low values in (a) and (b), respectively [1]. It is confirmed that the calculated posterior probability density $P(z_1,z_2|D)$ can detect solutions, and distinguish between the local and true solutions by tuning the value of the parameter τ, which demonstrates the reliability of the data analysis by 2DMAT.

REFERENCES

[1] https://www.pasums.issp.u-tokyo.ac.jp/2DMAT/
[2] Y. Motoyama, K. Yoshimi, I. Mochizuki, H. Iwamoto, H. Ichinose and T. Hoshi, *Comp. Phys. Commun.* **280**, 108465 (2022).
[3] Y. Fukaya, A. Kawasuso, A. Ichimiya and T. Hyodo, *J. Phys. D* **52**, 013002 (2019).
[4] T. Hoshi, D. Sakata, S. Oie, I. Mochizuki, S. Tanaka, T. Hyodo and K. Hukushima, *Comp. Phys. Commun.* **271**, 108186 (2022).
[5] https://github.com/sim-trhepd-rheed/sim-trhepd-rheed/; T. Hanada, Y. Motoyama, K. Yoshimi and T. Hoshi, *Comp. Phys. Commun.* **277**, 108371 (2022).
[6] K. Tanaka, I. Mochizuki, T. Hanada, A. Ichimiya, T. Hyodo and T. Hoshi, *JJAP Conf. Series* **9**, 011301 (2023).
[7] Y. Tsujikawa, M. Horio, X. Zhang, T. Senoo, T. Nakashima, Y. Ando, T. Ozaki, I. Mochizuki, K. Wada, T. Hyodo, T. Iimori, F. Komori, T. Kondo and I. Matsuda, *Phys. Rev. B* **106**, 205406 (2022).
[8] Y. Tsujikawa, M. Shoji, M. Hamada, T. Takeda, I. Mochizuki, T. Hyodo, I. Matsuda and A. Takayama, *Molecules* **27**, 4219 (2022).
[9] H. Kohashi, H. Iwamoto, T. Fukaya, Y. Yamamoto and T. Hoshi, *JSIAM Letters* **14**, 13 (2022).
[10] K. Tanaka, T. Hoshi, I. Mochizuki, T. Hanada, A. Ichimiya and T. Hyodo, *Acta. Phys. Pol. A* **137,** 188 (2020).

BEAMLINE

SPF-A3

T. Hoshi[1, 2], Y. Motoyama[3], K. Yoshimi[3] and I. Mochizuki[2]
([1]Tottori Univ., [2]KEK-IMSS-SPF, [3]The Univ. of Tokyo)