Searching for suitable hyperparameters in Bayesian optimization for material synthesis: three-dimensional case

Tokyo Tech. [1], JST-PRESTO. [2], AIST[3]

◦(D) Han Xu[1], Takefumi Kimura[1], Ryo Nakayama[1], Ryota Shimizu[1,2], Yasunobu Ando[3], Nobuaki Yasuo[1], Masakazu Sekijima[1], Taro Hitosugi[1]

E-mail: xu.h.ai@m.titech.ac.jp

[Introduction]: Conventional materials exploration requires trial-and-error in multi-dimensional searching space. However, in general, it is difficult for humans to optimize multi-dimensional synthesis conditions. Bayesian optimization has aroused the attention of material researchers as an efficient method to explore novel functional materials. In order to find the optimal material synthesis conditions with the minimum number of experiments, we have already investigated the optimum hyperparameters of the kernel function and the acquisition function in Bayesian optimization using 1D model functions.[1] However, the optimum hyperparameters in the 3D case are still not investigated. In this study, we determined the appropriate hyperparameters that cost the least number of experiments for optimization in 3D synthesis conditions through simulation, and compared the effect of the process window on searching speed.

[Method]: To mimic the actual material synthesis, we assumed thin film deposition using sputtering with 3D synthesis parameters; temperature from 200 °C to 1200 °C, oxygen partial pressure (P_{O2}) from 1×10^{-4} Pa to 11×10^{-4} Pa and power from 100 W to 200 W. The searching space comprises 51 grids in each dimension (Total grids: 51^3). 3D model functions are generated by adding four Gaussian functions with peak height 1.2, 0.7, 0.6, and 0.3, which correspond to one global peak, two local peaks and one background peak respectively (Fig. 1). Here, we defined the standard deviation of global maximum peaks (σ_{gauss}) as process window (P_w) of the model function. In this definition, the synthesis conditions within P_w provide the physical property value beyond ca. 90% of maximum value. We generated three kinds of model functions with σ_{gauss} = 3, 5, 7 grid in each dimension, which correspond to P_w set = [60 °C, 0.6×10^{-4} Pa, 6 W], [100 °C, 1×10^{-4} Pa, 10 W], and [140 °C, 1.4×10^{-4} Pa, 14 W] respectively. For each P_w set, 100 model functions with different peak positions are generated. The number of trials required to find the synthesis conditions within P_w in 90 out of 100 model functions ($N_{90\%}$) is used as an index to compare the optimization speed. The simulation is implemented in Python by using GPy package. We investigated the effects of hyperparameters in acquisition function (ξ) and radial basis function kernel (variance and lengthscale) on $N_{90\%}$.

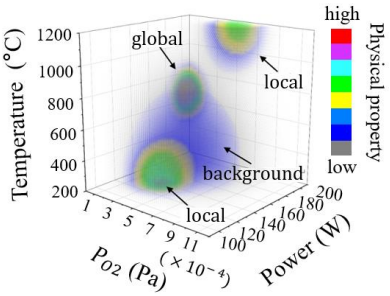

Fig. 1. 3D model function for the simulation of material synthesis

[**Results**]: Fig. 2 shows the histogram when P_w set = [100 °C, 1×10^{-4} Pa, 10 W], variance = 0.1, lengthscale = 1 grid, and acquisition function is expected improvement (EI) with hyperparameter ξ = 0.01. These parameters realize the lowest $N_{90\%}$ of 73 for P_w set = [100 °C, 1×10^{-4} Pa, 10 W]. This combination of hyperparameters is also the optimum at the other two different process windows P_w set = [60 °C, 0.6×10^{-4} Pa, 6 W] and P_w = [140 °C, 1.4×10^{-4} Pa, 14 W]

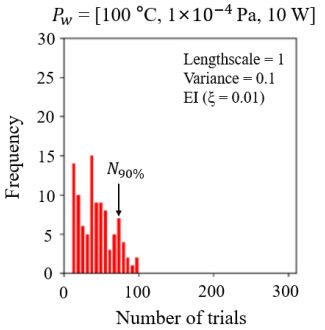

Fig. 2. Histogram for number of synthesis required for optimization

1. Nakayama, Hitosugi *et al.*, the 68th JSAP spring meeting, 10a-Z32-10 (2021)