Portfolio-based Strategy for Bayesian Optimization for Autonomous Selection of Multiple Acquisition Functions and Hyperparameters

Runzhe Liang,a Haoyang Hu,a Zhihong Yuana,\*

aDepartment of Chemical Engineering, Tsinghua University, Beijing 100084, China

\*Corresponding author’s E-mail: zhihongyuan@mail.tsinghua.edu.cn

Abstract

Bayesian optimization has become an effective tool in chemical research in recent years. Acquisition functions act as a vital part of the Bayesian optimization framework to determine the most promising experimental point, however, it is difficult for most chemists to select appropriate acquisition functions for different chemical reactions. In this investigation, a novel portfolio-based strategy for Bayesian optimization (PCR-BO) is proposed, which consists of multiple acquisition functions and selects the most suitable one autonomously. Based on a multi-armed bandit strategy, the proposed method considers historical observations and recent progress together. Benchmark tests and case studies indicate that the proposed portfolio-based strategy can significantly outperform other portfolio strategies and approach the upper bound of its individual components, demonstrating its excellent adaptability and selectivity.

**Keywords**: Bayesian optimization, acquisition function, portfolio strategy, multi-armed bandit

* 1. Introduction

Experimental conditions play a critical role in the performance of chemical reactions. However, when faced with complex chemical systems or unclear reaction mechanisms, it can be a challenging task for chemists to select an appropriate combination of parameters. A traditional approach is to fit the response surface model by DoE. However, it strongly depends on the prior knowledge of chemists and the optimum cannot be guaranteed. Bayesian optimization, as a global search framework that can deal with expensive black-box functions (Shahriari et al., 2016), has attracted increasing attention in experimental condition screening. By incorporating and translating the uncertainty of regression models, Bayesian optimization combines existing observations with unknown information properly to drive towards the optimum step by step. Generally, for an unfamiliar reaction system, suitable combinations of experimental parameters (e.g., temperature, pressure, and concentration) can be found after several experiments by the guide of Bayesian optimization, leading to high yield, selectivity or reaction rate. At present, Bayesian optimization has proven its efficiency in several areas, such as organic synthesis, material design, photoreactions, and electrochemistry (Shields et al., 2021).

The key components of Bayesian optimization can be divided into two parts: building a probabilistic surrogate model and optimizing acquisition functions, wherein the Gaussian process (GP) is a common approach to build the surrogate model, and the acquisition function is utilized to determine the next experimental point by making a trade-off between expectations and variances. There exist multiple acquisition functions as reported in the literature, such as PI, EI, and UCB. Nevertheless, different acquisition functions may have distinctly different behavior when confronted with different chemical systems, and it is a non-trivial decision on which one should be selected. Furthermore, the suitable acquisition function may change as the optimization campaign progresses. Therefore, how to design and update the selection of acquisition functions is an essential issue to be addressed.

Inspired by the aforementioned problems, a novel portfolio-based strategy for Bayesian optimization, named PCR-BO (Portfolio-based strategy with Customized Reward function for Bayesian Optimization), is proposed in this work, which is developed based on a multi-armed bandit (MAB) strategy. Instead of using a single acquisition function, the portfolio-based strategy contains a set of acquisition functions and optional hyperparameters. At each iteration, the acquisition functions are selected randomly with different probabilities determined by cumulative rewards, and the appropriate one gradually emerges and iterates as the rewards accumulate. The following sections elaborate on the method of the proposed portfolio strategy and practical performance.

* 1. Method

The proposed strategy is inspired by GP-Hedge (Hoffman et al., 2011) and No-PASt-BO (Vasconcelos et al., 2019), and manages to tackle their undesirable properties while maintaining the simplicity of the algorithm. Compared to the previous investigations, two main contributions are included in the proposed method: 1) the variance of GP is considered when calculating the reward of each arm, and 2) the reward is set as the improvement upon the incumbent target. The overall framework is clarified as follows.

* + 1. Hierarchical hedging strategy

To realize the selection of different acquisition functions from the portfolio, a MAB strategy is employed in this investigation, in which each acquisition function is treated as a single “arm”. Given *K* available individual acquisition functions, *K* possible arms are formulated, with unknown and maybe dynamically variable reward probability distribution for each one. The task is to find the approach to select the arm combination that corresponds to the maximal cumulative rewards. Specifically, we are confronted with a trade-off between exploration and exploitation: Due to the unknown reward probability distribution, different arms need to be tested repeatedly to obtain the ground-truth distribution, which implies they should be chosen equally. However, to obtain the maximal cumulative rewards, existing information should be fully utilized, which means the best arm up to now should be chosen as much as possible.

Hedge is a suitable algorithm to deal with this problem, which was also adopted by GP-Hedge and No-PASt-BO. Its main idea is to allocate different probabilities for each arm *i* for selection based on cumulative rewards. In the Bayesian optimization framework, each acquisition function *i* proposes an experimental point at each iteration *t*. The probability of each arm *i* can be determined by, where *η* is a hyperparameter and  is the cumulative rewards of the arm *i* up to the time *t*. In this way, the hedging algorithm enables possible exploration of every arm while focusing on the arms with high cumulative rewards.

* + 1. Reward customization

Based on the hedging strategy, how to measure and calculate the reward  of each step *t* is a crucial part of determining the probability of each arm. Previous investigations suggested that if the arm *i* proposed an experimental point , the expected value of the GP model at  could be a reasonable evaluation for the reward (i.e., ), which acts as a reasonable and natural approach.

However, it is flawed due to its simplistic formulation and inaccurate consideration. On the one hand, the reward function only focuses on the expectation, and the variance of the GP model is neglected. In Bayesian optimization settings, observations are usually rather limited, and exploration of the variances contributes to searching for possible global optimum. To take the variance into account, we propose to sample the reward function  from the posterior of GP, by which the expectation and variance can be handled together. The arm with the highest reward is selected at step *t*.

On the other hand, the formulation  is myopic and deviated from the ultimate target. Since we aim to reach the optimum within the limited number of iterations, it is desirable to make steady progress towards the optimum; that is, step-by-step improvements should be treasured during iterations. Therefore, in this investigation we propose to use the potential improvement , instead of pure expectation , as the reward function to make the selection concentrate on potential better value, where  represents the best observation up to the time *t* and  is the improvement function (i.e., ).

* + 1. Memory factor and reward normalization

Different from the vanilla MAB problem, one prominent feature of the portfolio strategy is that the optimal acquisition function may change and evolve as the optimization campaign progresses. Too distant reward evaluation hardly provides efficient information for the current decision. Therefore, a memory factor is introduced to reduce the influence of previous iterations and enables the optimizer to focus more on the recent reward. In detail, the cumulative reward can be calculated by the equation , where *m* is a memory hyperparameter and its range is [0, 1]. The recommended range for *m* in No-PASt-BO is [0.7, 1]; however, since the proposed portfolio strategy involves sampling from the posterior, more randomness may be introduced so that a smaller value of *m* is encouraged to recover from numerical fluctuation quickly. A recommended value for *m* in this investigation is 0.7.

The hedging framework also suffers from the scale of the reward function. Suppose the difference of  is huge for each arm *i*, the probability of one arm may occupy a dominant position over the others. On the contrary, if all remains unchanged except for scaling down  by a large number, the rewards may be close enough to make the portfolio degenerate into a completely random one. To avoid this undesirable behavior, the reward function is normalized as shown in Eq. (1):

|  |  |
| --- | --- |
|  | (1) |

In this way the reward  is constrained between -1 and 0. Then  is replaced by  to calculate the probability to realize the normalization of the reward function:

|  |  |
| --- | --- |
|  | (2) |

* + 1. The PCR-BO algorithm

The above is the basic idea of the PCR-BO algorithm. Its detailed realization is summarized in Algorithm 1.

|  |
| --- |
| **Algorithm 1** PCR-BO |
| **Input**: hyperparameter , |
| 1. Let |
| 1. **for**  **do** |
| 1. for each acquisition function  in the portfolio: |
|  |
| 1. Normalize rewards |
| 1. Nominate a point  from  with probability |
| 1. Evaluate objective function on point  and update the GP model |
| 1. Sample  from the GP posterior |
| 1. Update the rewards |
| 1. **end for** |

* 1. Case study

To test the performance of the proposed PCR-BO algorithm, a few case studies including several benchmark experiments and a simulated real-world application were carried out, and the performance of PCR-BO, GP-Hedge, No-PASt-BO, and the upper bound and average effect of individual acquisition functions was recorded for comprehensive comparison. The portfolio involved three common acquisition functions: PI, EI, and UCB, and various hyperparameters of these acquisition functions were included in the portfolio for optional selection:  and 0.001 for PI and EI,  and 3.0 for UCB. To ensure the reliability and stability of the test, 50 runs were executed for each method and case to obtain average results. For hyperparameters, *η* was set as 2.5 and *m* was set as 0.7 in the proposed method, while the hyperparameters in the other methods were consistent with the recommended values in their original references.

* + 1. Benchmark test

Before large-scale application in practical scenarios, the statistical performance of the proposed method was determined via benchmark mathematical functions. Here four common benchmark functions were selected for testing, with different and multiple local optima: Branin, Hartman3, Ackley, and Hartman6, with 2, 3, 5, and 6 dimensions respectively. Figure 1 illustrates the comparison results of the logarithm of regret versus iteration for different methods, where regret denotes the difference between the actual optimum and the observed one.



Figure 1. Average results and 95% confidence intervals of PCR-BO, GP-Hedge, No-PASt-BO, and the upper bound and the average performance of individual acquisition functions in the portfolio for various benchmark functions. a) Branin, b) Hartman3, c) Ackley, d) Hartman6.

According to Figure 1, the proposed PCR-BO algorithm can outperform the other portfolio-based methods and approaches the best performance of individual acquisition functions in most cases. For the Ackley function, due to its significant multimodal characteristics (there exist thousands of local optima for the 5-dimensional Ackley function), it is difficult to determine the appropriate acquisition function, and wrong decisions may occur repeatedly during iteration. However, after wandering in the beginning, PCR-BO manages to find the suitable acquisition function to drive the regret decrease rapidly, especially after 60 iterations. In contrast, the regret of GP-Hedge and No-PASt-BO remains at high levels all along.

* + 1. Hydrogenation of p-chloronitrobenzene

It is an essential transformation from halogenated nitroaromatics to their corresponding aromatic anilines in the fine chemical industry. However, low selection and yield may occur due to incomplete conversion and dehalogenation when applied in the traditional batch mode. For better controllability and economic benefits, a simulated optimization campaign of the hydrogenation of *p*-chloronitrobenzene is conducted using continuous flow synthesis techniques, which provide uniform temperature distribution, intensified mass transfer, and easy implementation.

Previous investigations have determined the mechanism and kinetic parameters of the reaction (Duan et al., 2022). Figure 2a explains its competitive pathways, where apart from the target product, nitrobenzene and aniline as byproducts are generated via slow reaction steps. It implies that by precise control, it is feasible to enhance the yield of *p*-chloroaniline while keeping side reactions at low levels.



Figure 2. a) Reaction network for the hydrogenation of *p*-chloronitrobenzene. b) Average results and 95% confidence intervals for the hydrogenation of the *p*-chloronitrobenzene.

To maximize the production capability, the optimization objective is set to maximize the yield of *p*-chloroaniline per unit of time. Temperature *T*, pressure *P*, residence time *R,* and reagent concentration *C* are selected as decision variables with different value ranges: , , , . Furthermore, considering downstream separation costs, it is demanded that the yield of the target product must exceed 90 % after optimization, and moderate noise is added to the objective function to make it closer to reality.

Comparison results are shown in Figure 2b. Note that due to complex constraints and potential noise disturbance, the performance of different methods is close, making it hard to distinguish which acquisition function should be selected in portfolio-based strategies. However, PCR-BO is still the one that is the closest to the upper bound of individual acquisition functions, demonstrating its superior performance when faced with complex real-world applications.

* 1. Conclusions

A novel portfolio-based strategy for Bayesian optimization (PCR-BO), which assists chemists in selecting the appropriate acquisition function autonomously, is proposed in this investigation. Compared with existing methods, PCR-BO manages to tailor suitable reward functions and take into account the variances of GP models, forming a more complete hedging framework. Case studies demonstrate that PCR-BO can be close enough to the upper bound of its individual acquisition functions in most cases, and shows competitive performance when applied in practical applications.

References

B. Shahriari, K. Swersky, Z. Wang, R. P. Adams and N. de Freitas, 2016, Taking the Human Out of the Loop: A Review of Bayesian Optimization, Proc. IEEE, 104, 148-175.

B. J. Shields, J. Stevens, J. Li, M. Parasram, F. Damani, J. I. M. Alvarado, J. M. Janey, R. P. Adams and A. G. Doyle, 2021, Bayesian reaction optimization as a tool for chemical synthesis, Nature, 590, 89-96.

M. Hoffman, E. Brochu and N. De Freitas, 2011, Portfolio Allocation for Bayesian Optimization. In UAI, pp. 327-336.

T. D. P. Vasconcelos, D. A. de Souza, C. L. Mattos and J. P. Gomes, 2019, No-PASt-BO: Normalized portfolio allocation strategy for Bayesian optimization. In 2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI), pp. 561-568.

X. Duan, J. Yin, M. Huang, P. Wang and J. Zhang, 2022, Hydrogenation kinetics of halogenated nitroaromatics over Pt/C in a continuous Micro-packed bed reactor, Chem Eng Sci, 251.