High-Dimensional Bayesian Optimization via Semi-Supervised Learning with Optimized Unlabeled Data Sampling

Yuxuan Yin [1] **Yu Wang** [1] **Peng Li** [1]

Abstract

We introduce a novel semi-supervised learning approach, named Teacher-Student Bayesian Optimization (TSBO), integrating the teacher-student paradigm into BO to minimize expensive labeled data queries for the first time. TSBO incorporates a teacher model, an unlabeled data sampler, and a student model. The student is trained on unlabeled data locations generated by the sampler, with pseudo labels predicted by the teacher. The interplay between these three components implements a unique *selective regularization* to the teacher in the form of student feedback. This scheme enables the teacher to predict high-quality pseudo labels, enhancing the generalization of the GP surrogate model in the search space. To fully exploit TSBO, we propose two optimized unlabeled data samplers to construct effective student feedback that well aligns with the objective of Bayesian optimization. Furthermore, we quantify and leverage the uncertainty of the teacher-student model for the provision of reliable feedback to the teacher in the presence of risky pseudo-label predictions. TSBO demonstrates significantly improved sample-efficiency in several global optimization tasks under tight labeled data budgets. The implementation is available at https://github.com/reminiscenty/TSBO-Official.

1. Introduction

Bayesian Optimization (BO) (Brochu et al., 2010) is widely adopted for black-box optimization, which is particularly useful when the objective function is expensive or impractical to evaluate directly. BO operates by constructing a surrogate model, e.g., a Gaussian Process (GP) (Seeger, 2004), of the objective function and then iteratively selecting the most promising locations for new labeled data query based on a criterion that balances exploration and exploitation (Kushner, 1964; Jones et al., 1998; Srinivas et al., 2010). Recent work has extended BO's applicability to high-dimensional tasks with various dimension reduction methods including linear embedding (Wang et al., 2016; Chen et al., 2020), nonlinear projections (Moriconi et al., 2020), and deep autoencoders (Kusner et al., 2017; Jin et al., 2018; Tripp et al., 2020; Chen et al., 2020; Grosnit et al., 2021; Maus et al., 2022; Chen et al., 2023b).

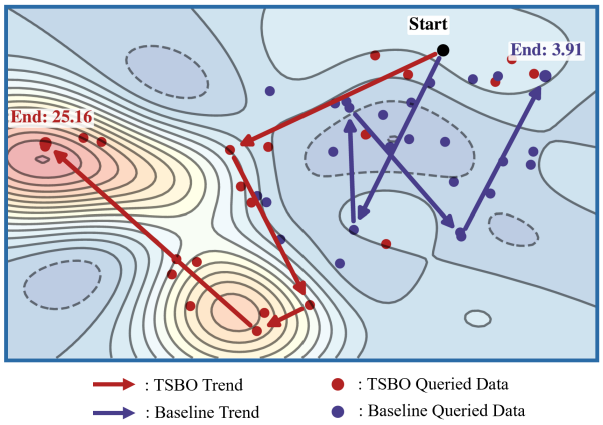

Figure 1: Visualization of queried data (dots) and trends (arrow sequences) on a high-dimensional molecule design task (Sterling & Irwin, 2015) to maximize the Penalized LogP score (Gómez-Bombarelli et al., 2018). Red and blue colors represent TSBO and a baseline (with vanilla BO), respectively. The evaluation budget is 450 in both approaches.

Despite these encouraging advances, labeled data acquisition remains inherently costly and presents a key bottleneck in BO across many application domains such as functional molecule design (Brown et al., 2019; Gao et al., 2022), structural optimization (Zoph et al., 2018; Lukasik et al., 2022), and failure analysis (Hu et al., 2018; Liang, 2019).

To address the challenge of data query efficiency in black-box optimization, we introduce a unified Semi-Supervised

[1]Department of Electrical and Computer Engineering, University of California, Santa Barbara, USA. Correspondence to: Peng Li <lip@ucsb.edu>.

Proceedings of the 41st International Conference on Machine Learning, Vienna, Austria. PMLR 235, 2024. Copyright 2024 by the author(s).

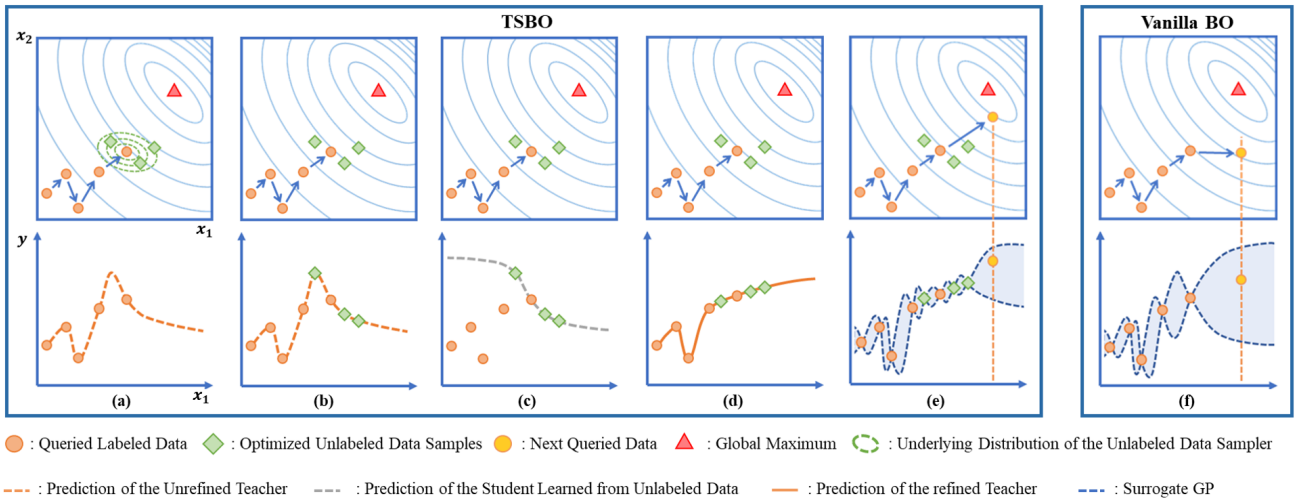

Figure 2: Illustrated example to demonstrate the interaction between the unlabeled data sampler, the teacher and the student employs selective regularization. (a): unlabeled data are sampled from regions with potentially high values. (b): the teacher predicts pseudo labels for unlabeled data. (c): the student learns from the unlabeled data and the predicted pseudo labels and is evaluated on labeled data as the feedback. (d): the teacher refines its prediction based on the feedback. (e): GP in TSBO fits on both the labeled data and unlabeled data with refined pseudo labels. (f): GP in vanilla BO fits only on the labeled data.

Learning (SSL) approach called Teacher-Student Bayesian Optimization (TSBO). TSBO is the first work integrating a teacher-student model into BO, and bridges the gap between SSL and the goals of BO by implementing a unique *selective regularization* mechanism, allowing the use of a subset of "potentially high-quality" unlabeled data from a vast sample space. This targeted utilization of cheap unlabeled data is optimized to serve the BO's optimization objective, steering it more effectively towards areas of high values with dramatically increased data query efficiency compared with baseline methods. An illustrated example on TSBO's sequential optimization process is visualized by UMAP (McInnes et al., 2018) and shown in Figure 1, which demonstrates the superior sample efficiency of TSBO.

At its core, TSBO incorporates a *teacher model*, a *student model*, and an *unlabeled data sampler*. It is the interplay of the three components that implement our targeted *selective regularization* to the teacher. Ultimately, the regularized teacher predicts high-quality pseudo labels [1], which supplement the queried labeled data to better train the standard GP surrogate model for more optimized new labeled data query during Bayesian optimization.

As illustrated in Figure 2, at each BO iteration, the unlabeled data sampler selects a set of *optimized* unlabeled data locations and passes them onto the teacher, which utilizes its current knowledge to prediction pseudo labels (Lee et al., 2013; Pham et al., 2021) for them. The student is then trained exclusively on the pseudo labels predicted by the teacher. Recognizing the fact the teacher's pseudo-label prediction can be misleading, the student is evaluated on the existing ground truth labeled data, and its performance is fed back to the teacher. Subsequently, the teacher is refined with the student feedback included as the selective regularization. Upon the completion of teacher-student interaction, the teacher refined with the regularization provides reliable pseudo-labels for training the BP surrogate model, bolstering its new data query capability even with limited labeled training data.

To fully exploit TSBO, it is essential to carefully design the key components involved in selective regularization. Instead of employing random sampling of unlabeled data, our unlabeled data sampler strategically places unlabeled data in regions of high-quality pseudo labels, and simultaneously encourages exploration towards the global optimum. Providing the student feedback to the teacher based on unlabeled data sampled under this strategy evaluates and refines the teacher in a way that well aligns with the overall objective of Bayesian optimization. We propose two optimized unlabeled data samplers: one based on the Extreme Value Theory (EVT) (Fisher & Tippett, 1928) and the other on a parameterized sampling distribution.

Furthermore, we boost the performance of TSBO by making our teacher-student model *uncertainty aware*. The teacher not only assigns pseudo labels to the selected unlabeled data but also quantifies their uncertainty. These uncertainties

[1] Although the term "label" often means the ground truth in classification problems, it is also widely used to represent observed values in BO tasks (Grosnit et al., 2021; Chen et al., 2020; Jean et al., 2018).

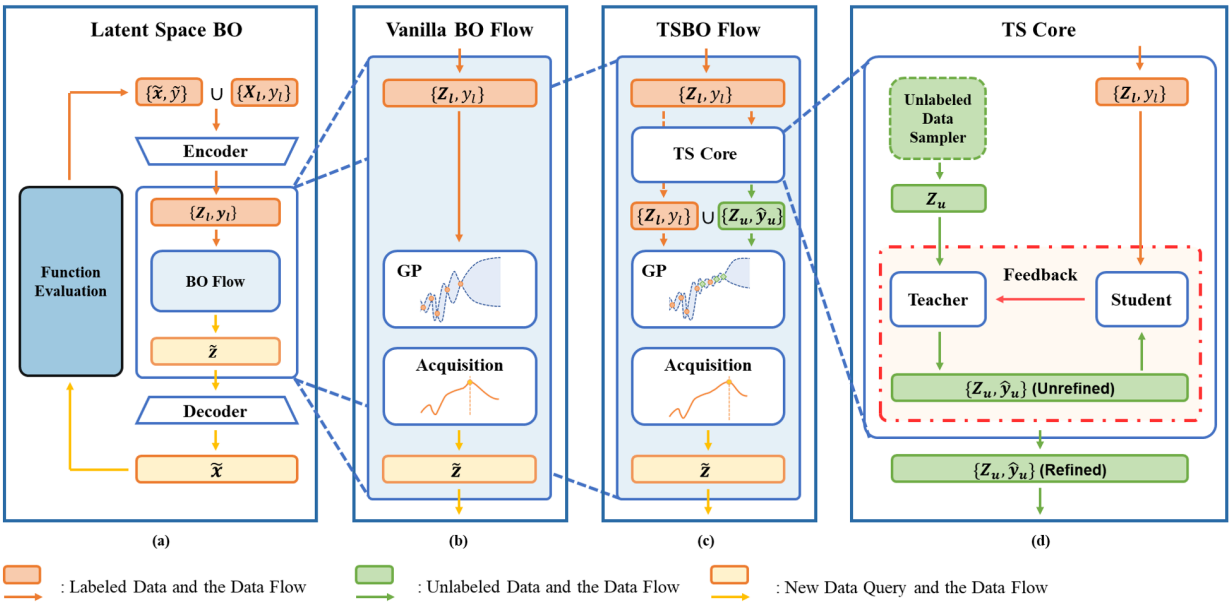

Figure 3: Overview of the TSBO framework. (a): the basic Latent Space BO architecture. (b): the vanilla BO flow utilizes only the encoded labeled data to train the GP model and query the next data. (c): TSBO flow incorporates a TS Core to provide additional high-quality unlabeled data to the GP model during each BO iteration.. (d): inside TS Core: the optimized unlabeled data sampler and the feedback from the student provides selective regularization to the teacher.

are taken into account during the student training process, thereby providing more reliable feedback to the teacher, especially in the presence of the risk associated with unreliable pseudo-label predictions.

We evaluate the proposed TSBO on various challenging high dimensional datasets and show superior data efficiency improvement. In a chemical design task (Sterling & Irwin, 2015) and an expression reconstruction task (Kusner et al., 2017), we achieve SOTA results compared to recent BO approaches.

2. Preliminaries

2.1. Bayesian Optimization (BO)

Bayesian Optimization (BO) (Brochu et al., 2010) is designed to identify the global maximum of an unknown, typically non-convex function $f : \mathcal{X} \to \mathbb{R}$, defined as:

$$\mathbf{x}^* = \underset{\mathbf{x}\in\mathcal{X}}{\operatorname{argmax}}\, f(\mathbf{x}) \tag{1}$$

where $\mathcal{X} \subseteq \mathbb{R}^D$ represents a D-dimensional input space.

BO approaches this challenge by iteratively selecting data points $\mathbf{x}$ for evaluation, building upon the outcomes of previously queried points. Within a single BO iteration, given a set of N evaluated examples $\{\mathbf{x}_i, \mathrm{y}_i\}_{i=1}^N = \{\mathbf{X}_l, \mathbf{y}_l\}$, where $\mathbf{X}_l$ is an $N \times D$ matrix of inputs and $\mathbf{y}_l$ is an $N \times 1$ vector of corresponding outputs, the next query point $\tilde{\mathbf{x}}$ is determined by:

$$\tilde{\mathbf{x}} = \underset{\mathbf{x}\in\mathcal{X}}{\operatorname{argmax}}\, \alpha\Big(Q\big(f|\mathbf{x}, \{\mathbf{X}_l, \mathbf{y}_l\}\big)\Big) \tag{2}$$

Here, $Q\big(f|\mathbf{x}, \mathbf{X}_l, \mathbf{y}_l\big)$ denotes the posterior distribution of f at $\mathbf{x}$ estimated by a probabilistic model Q, commonly termed as surrogate model(Frazier, 2018). The acquisition function α is designed to strike a balance between exploiting regions of known high values and exploring unknown spaces.

2.2. Latent Space BO

For high-dimensional challenges, direct application of Bayesian Optimization (BO) in the original space can lead to overfitting in the data query model Q, particularly under a constrained data query budget, a dilemma often referred to as the curse of dimensionality (Brochu et al., 2010). A viable remedy is latent space BO (Gómez-Bombarelli et al., 2018; Kusner et al., 2017), which conducts BO within a generative latent space $\mathcal{Z} \subseteq \mathbb{R}^d$ where $d \ll D$. This approach utilizes an encoder $\psi : \mathcal{X} \to \mathcal{Z}$ and a decoder $\varphi : \mathcal{Z} \to \mathcal{X}$ to navigate the query process. In latent space BO, the next query $\tilde{\mathbf{x}}, \tilde{\mathrm{y}}$, informed by the current labeled data $\mathbf{X}_l, \mathbf{y}_l$, involves:

- Fitting the data query model Q in $\mathcal{Z}$ using latent code pairs $\mathbf{Z}_l, \mathbf{y}_l$, where $\mathbf{Z}_l := \psi(\mathbf{X}_l)$;
- Identifying an optimal latent code $\tilde{\mathbf{z}}$ by max-

imizing the acquisition function α in $\mathcal{Z}$: $\tilde{\mathbf{z}} = \operatorname{argmax}_{\mathcal{Z}} \alpha\Big(Q\big(f|\mathbf{z}, \mathbf{Z}_l, \mathbf{y}_l\big)\Big)$;

- Decoding $\tilde{\mathbf{x}} = \varphi(\tilde{\mathbf{z}})$ and evaluating $\tilde{\mathrm{y}} = f(\tilde{\mathbf{x}})$.

The basic latent space BO framework is visualized in Figure 3 (a).

3. TSBO Problem Formulation

The TSBO framework builds upon the conventional latent space BO to enhance data efficiency in challenging high-dimensional tasks. While retaining the encoder-decoder structure and following the established data query procedure, TSBO additionally introduces a teacher model, a student model, and an unlabeled data sampler to employ selective regularization in the latent space between each data query. This key step enriches the surrogate model with a wealth of high-quality unlabeled data and their corresponding calibrated pseudo labels, thereby enhancing the model's predictive performance, as depicted in Figure 3 (c) and Figure 3 (d).

Let $T(\cdot; \boldsymbol{\theta_T})$ represent the teacher model, $S(\cdot; \boldsymbol{\theta_S})$ the student model, and $p_{\mathbf{z}_u}(\cdot; \boldsymbol{\theta_u})$ the distribution underlying the unlabeled data sampler. The set $\mathcal{D}_l := \{\mathbf{Z}_l, \mathbf{y}_l\}$ denotes the labeled data available at the current step. The interaction between these modules is formally described below and formulated as a bi-level optimization problem:

1. Unlabeled data $\mathbf{Z}_u$ is sampled from $p_{\mathbf{z}_u}(\cdot; \boldsymbol{\theta_u})$.
2. The teacher predicts pseudo labels on the unlabeled data samples: $\hat{\mathbf{y}}_u(\boldsymbol{\theta_T}) := \mathbb{E}(T(\mathbf{Z}_u; \boldsymbol{\theta_T}))$.
3. The student is trained with an unlabeled loss $\mathcal{L}_u$ defined on an unlabeled dataset $\mathcal{D}_u(\boldsymbol{\theta_T}) := \{\mathbf{Z}_u, \hat{\mathbf{y}}_u\}$:
$$\boldsymbol{\theta}^*_{\boldsymbol{S}}(\boldsymbol{\theta_T}) = \operatorname*{argmin}_{\boldsymbol{\theta_S}} \mathcal{L}_u\big(\mathcal{D}_u(\boldsymbol{\theta_T}); \boldsymbol{\theta_S}\big) \tag{3}$$
4. The trained student is evaluated on the labeled dataset $\mathcal{D}_l$ by a feedback loss $\mathcal{L}_f\big(\mathcal{D}_l; \boldsymbol{\theta}^*_{\boldsymbol{S}}(\boldsymbol{\theta_T})\big)$
5. The teacher is updated by optimizing the combination of a labeled loss $\mathcal{L}_l$ and the feedback loss $\mathcal{L}_f$:
$$\boldsymbol{\theta}^*_{\boldsymbol{T}} = \operatorname*{argmin}_{\boldsymbol{\theta_T}} \mathcal{L}_l(\mathcal{D}_l; \boldsymbol{\theta_T}) + \lambda \mathcal{L}_f\big(\mathcal{D}_l; \boldsymbol{\theta}^*_{\boldsymbol{S}}(\boldsymbol{\theta_T})\big) \tag{4}$$
where λ is a weighting parameter.

4. Detailed Design of TSBO Modules

4.1. Uncertainty-Aware Teacher-Student Model

The interplay between the teacher and student models is pivotal for refining the teacher's predictions and enhancing the accuracy of pseudo labels on the sampled unlabeled data. Crucially, we emphasize the importance of incorporating uncertainty awareness to safeguard the feedback mechanism's integrity. Inaccurate pseudo labels can disrupt the modeling process, leading to erroneous feedback. By quantifying prediction uncertainty, the student model can more prudently utilize pseudo labels, adjusting its reliance on them based on their associated uncertainty levels. As a result, we propose the uncertainty-aware teacher-student, a practical mechanism to utilize pseudo labels effectively while minimizing potential risks.

4.1.1. Uncertainty-Aware Teacher Model

Quantifying uncertainty accurately is inherently complex and computationally demanding. In TSBO, we adopt a heuristic approach inspired by (Nix & Weigend, 1994) using a Multilayer Perceptron (MLP) as the teacher model. This MLP is designed to output not only the predicted mean but also the diagonal covariance matrix for any given input $\mathbf{Z}$, thereby providing a probabilistic estimate of the output:

$$T(\mathbf{Z}; \boldsymbol{\theta_T}) := \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{\theta_T}}(\mathbf{Z}), \boldsymbol{\Sigma}_{\boldsymbol{\theta_T}}(\mathbf{Z})), \tag{5}$$
$$\text{where} \quad \boldsymbol{\Sigma}_{\boldsymbol{\theta_T}}(\mathbf{Z}) := \boldsymbol{\sigma}^{\mathbf{2}}_{\boldsymbol{\theta_T}}(\mathbf{Z}) \cdot \boldsymbol{I} \tag{6}$$

Training of the defined teacher model becomes fitting a parameterized Gaussian distribution by minimizing a negative log-likelihood (NLL) loss. Consequently, the labeled loss in Equation (4) is written as:

$$\mathcal{L}_l(\mathcal{D}_l; \boldsymbol{\theta_T}) := \mathrm{NLL}\big(T(\mathbf{Z}_l; \boldsymbol{\theta_T}), \mathbf{y}_l\big) \tag{7}$$

The pseudo label that incorporated with uncertainty from the teacher is generated as follows:

$$\begin{aligned} &\hat{\mathrm{y}}_u = \mathrm{y}_u + \epsilon_u(\mathbf{z}_u), \\ \text{where} \quad &\epsilon_u(\mathbf{z}_u) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\theta_T}}(\mathbf{z}_u)) \end{aligned} \tag{8}$$

4.1.2. Uncertainty-Aware Student Model

In our approach, the student model is implemented as a Gaussian Process (GP) as in Equation (9), which can efficiently incorporate the teacher's uncertainty predictions as prior knowledge. This configuration enables the student model not only to refine its predictions but also to provide precise feedback on the teacher model's mean predictions and associated uncertainties.

$$S(\cdot; \boldsymbol{\theta_S}) := \mathcal{GP}(\mathbf{0}, \kappa_0(\cdot, \cdot) + \sigma_0^2 \boldsymbol{I}) \tag{9}$$

where $\kappa_0(\cdot, \cdot))$ is a Radial Basis Function (RBF) kernel, σ_0^2 is an additive variance. The student model parameter $\boldsymbol{\theta_S}$ encloses the kernel parameters and the variance σ_0^2.

We propagate the teacher's uncertainty $\epsilon_u(\mathbf{z}_u)$ to the downstream training of the student GP model by forming a student's pseudo-label dependent prior: $\hat{\mathrm{y}}_u = \epsilon_u(\mathbf{z}_u) + \epsilon_{\kappa 0}$,

where $\epsilon_{\kappa 0} \sim \mathcal{N}(0, \kappa_0(\mathbf{z}_u, \mathbf{z}_u) + \sigma_0^2)$ is the student's vanilla prior variance in Equation (9). The optimization of student's variance $\epsilon_{\kappa 0}$ depends on the teacher's uncertainty estimation $\mathbf{\Sigma}_{\boldsymbol{\theta_T}}(\mathbf{z}_u)$.

Correspondingly, the student's covariance matrix $\mathbf{\Sigma}_u$ over the unlabeled dataset $\mathcal{D}_u(\boldsymbol{\theta_T})$ is computed as: $\mathbf{\Sigma}_{u_{ij}} = \mathbb{E}(\hat{\mathrm{y}}_{u_i} - \mathbb{E}\ \hat{\mathrm{y}}_{u_i})(\hat{\mathrm{y}}_{u_j} - \mathbb{E}\ \hat{\mathrm{y}}_{u_j}) = \kappa_0(\mathbf{z}_{u_i}, \mathbf{z}_{u_j}) + \delta_{ij}\sigma_0^2 + \delta_{ij}\mathbf{\Sigma}_{\boldsymbol{\theta_T}}(\mathbf{z}_{u_i})$, where δ represents the Kronecker delta. It is the sum of the uncertainty of the teacher and the student:

$$\mathbf{\Sigma}_u = \kappa_0(\mathbf{Z}_u, \mathbf{Z}_u) + \sigma_0^2 \boldsymbol{I} + \mathbf{\Sigma}_{\boldsymbol{\theta_T}}(\mathbf{Z}_u) \tag{10}$$

We optimize the uncertainty-aware student on the unlabeled dataset $\mathcal{D}_u(\boldsymbol{\theta_T})$ to minimize the NLL loss $\mathcal{L}_u$:

$$\mathcal{L}_u(\mathcal{D}_u(\boldsymbol{\theta_T}); \boldsymbol{\theta_S}) := \mathrm{NLL}\Big(\mathcal{N}\big(\mathbf{0}, \mathbf{\Sigma}_u\big), \hat{\mathbf{y}}_u\Big) \tag{11}$$

We denote the optimized student parameter as $\boldsymbol{\theta}_{\boldsymbol{S}}^*(\boldsymbol{\theta_T})$.

4.1.3. Derivation of the Feedback Loss

The student GP model which is optimized by minimizing the unlabeled data loss in Equation (11), is evaluated on the labeled dataset D_l. This evaluation yields a *posterior* prediction for the encoded inputs $\mathbf{Z}_l$. We focus on the posterior mean $\boldsymbol{\mu}_{\boldsymbol{\theta_S}}(\mathbf{Z}_l; \mathcal{D}_u(\boldsymbol{\theta_T}))$ to quantify the model's performance and define the feedback loss as the Mean Squared Error (MSE)[2] between this posterior mean and the true labels $\mathbf{y}_l$.

$$\mathcal{L}_f\big(\mathcal{D}_l; \boldsymbol{\theta}_{\boldsymbol{S}}^*(\boldsymbol{\theta_T})\big) := \mathrm{MSE}\Big(\boldsymbol{\mu}_{\boldsymbol{\theta}_{\boldsymbol{S}}^*(\boldsymbol{\theta_T})}(\mathbf{Z}_l; \mathcal{D}_u(\boldsymbol{\theta_T})), \mathbf{y}_l\Big) \tag{12}$$

The posterior mean has an explicit form which is derived as:

$$\boldsymbol{\mu}_{\boldsymbol{\theta}_{\boldsymbol{S}}^*(\boldsymbol{\theta_T})}(\mathbf{Z}_l; \mathcal{D}_u(\boldsymbol{\theta_T})) = \kappa_0(\mathbf{Z}_l, \mathbf{Z}_u)^T {\mathbf{\Sigma}_u}^{-1} \hat{\mathbf{y}}_u \tag{13}$$

Remark: The student GP model efficiently incorporates the teacher's uncertainty into its feedback, ensuring that its decisions are informed by this crucial context. Specifically, the model's posterior predictions show a diminished dependency on pseudo labels that are associated with high levels of uncertainty from the teacher: The teacher's predictive variance $\mathbf{\Sigma}_{\boldsymbol{\theta_T}}(\mathbf{z}_{u_i})$ for the i-th pseudo label is added to the i-th diagonal entry of the covariance matrix $\mathbf{\Sigma}_u$ in Equation (10). When this uncertainty is significantly greater than that of other pseudo labels, the corresponding diagonal element $\mathbf{\Sigma}_{u_{ii}}$ is much larger than the other diagonal elements. As a result, per Equation (13) the contributions of the i-th pseudo label in $\hat{\mathbf{y}}_u$ to the posterior mean predictions of the validation labels are considerably reduced.

[2]The feedback loss can be an MSE, a negative predictive marginal log-likelihood (Gneiting & Raftery, 2007) or a negative Mahalanobis distance (Bastos & O'hagan, 2009). For numerical stability, we choose the MSE in our work.

4.2. Optimized Unlabeled Data Sampling

Random sampling, often used in standard Semi-Supervised Learning (SSL) scenarios, falls short in Bayesian Optimization (BO) due to its non-optimized nature. The teacher model's pseudo labels for randomly chosen unlabeled data, particularly those far from the training dataset, are prone to low quality, characterized by small means or large variances. These pseudo labels can potentially mislead the student and steer it away from the global optimum. Consequently, an inaccurately informed student model will fail to provide the necessary feedback for the teacher model's refinement towards optimal solutions.

We argue that within the vast expanse of the sample space, only a specific subset of unlabeled data holds significant value and aligns with the BO objectives. Utilizing these unlabeled data should benefit the BO process most effectively. In light of this, the development of a targeted unlabeled data sampling strategy, one that efficiently identifies and selects these high-value samples, is crucial. To meet this objective, we propose two novel techniques designed to enhance the sampling distribution for unlabeled data, ensuring a more focused and effective optimization in BO settings.

Method 1: Extreme Value Theory (EVT) based Unlabeled Data Sampling The key idea is to place unlabeled data in regions of high-quality pseudo labels and at the same time encourage exploration towards the global optimum. To do so, we model the distribution of the part of the labeled data, which are extreme, i.e. with the best target values. EVT (Fisher & Tippett, 1928) states that if $\{\mathrm{y}_1, \cdots, \mathrm{y}_N\}$ are i.i.d. and as N approaches infinity, their maximum y^* follows a Generalized Extreme Value (GEV) distribution (Fisher & Tippett, 1928).

The GEV distribution (Fisher & Tippett, 1928) is defined as

$$p_{\mathrm{y}^*}(\mathrm{y}^*) = \mathbb{I}_{\{\xi \neq 0\}} \left(1 + \xi \bar{\mathrm{y}}\right)^{-\frac{1}{\xi}} e^{-(1+\xi\bar{\mathrm{y}})^{-\frac{1}{\xi}}} + \mathbb{I}_{\{\xi = 0\}} e^{-\bar{\mathrm{y}}} e^{-e^{-\bar{\mathrm{y}}}}, \tag{14}$$

where $\bar{\mathrm{y}} := (\mathrm{y}^* - a)/b$ defined by 3 learnable parameters of the GEV distribution: a location coefficient $a \in \mathbb{R}$, a scale value $b > 0$, and a distribution shape parameter $\xi \in \mathbb{R}$.

Following (Gomes & Guillou, 2015), we fit a GEV distribution p_{y^*} with parameters estimated by maximizing the likelihood of the joint distribution of several extreme labels, which are top K target value points in our design.

This GEV distribution captures the distribution of the best-observed target values as seen from the current evaluated data. As such, generating unlabeled data whose predicted labels follow the GEV distribution allows us to start out from the region of the existing extreme labeled data while exploring points with potentially even greater target values due to the random nature of the sampling process. Once the GEV distribution p_{y^*} is fitted, we adopt a specific MCMC

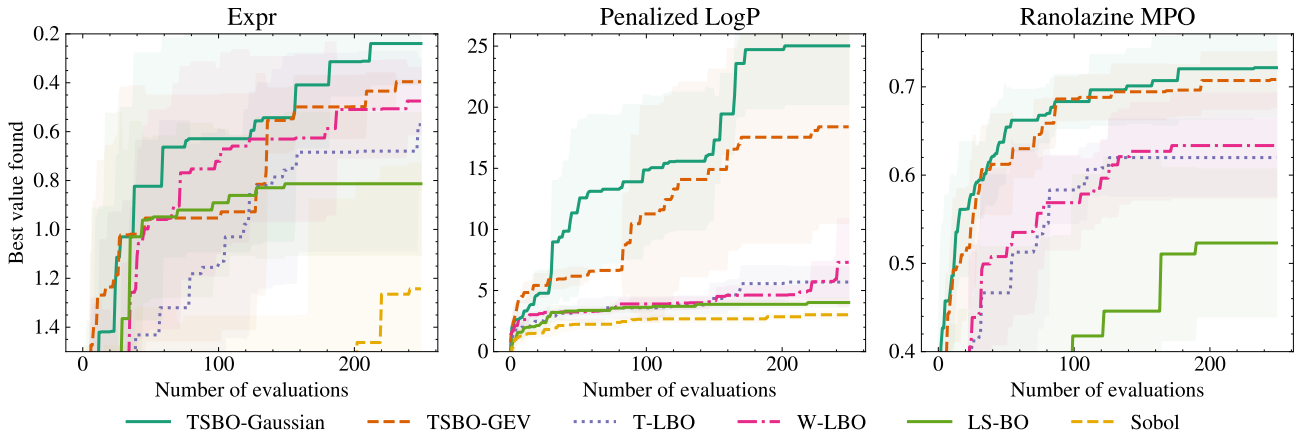

Figure 4: Comparison between the mean performance and standard deviations between 4 LSO baselines and TSBO.

method tailored for GEV (Hu et al., 2019) to draw samples from it.

Method 2: Unlabeled Data Sampling Distribution Learned from Student's Feedback While the proposed GEV distribution approach offers a theoretically sound method for generating unlabeled data, its practical effectiveness is constrained by the computationally intensive nature of the MCMC sampling technique.

To circumvent the computational burden associated with MCMC, we endeavor to identify an alternative approach for sampling unlabeled data, denoted as $\mathbf{z}_u$, from a distribution $p_{\mathbf{z}_u}(\cdot;\boldsymbol{\theta_u})$ parameterized $\boldsymbol{\theta_u}$.

In designing and optimizing the parameterized distribution, we adhere to two key principles: First, the unlabeled data should be sampled from potential high value regions, which is likely to be around the current best value point. Second, we need to exclude low-quality unlabeled data samples that could mislead the teacher-student feedback loop, as indicated by a corresponding large feedback loss.

In our design, we choose the Gaussian distribution as the unlabeled data distribution $p_{\mathbf{z}_u}(\cdot;\boldsymbol{\theta_u}) := \mathcal{N}(\boldsymbol{\mu_u}, \boldsymbol{\sigma}_{\boldsymbol{u}}^2\boldsymbol{I})$, whose mean is initialized by the current best target value point, and initial variance is set to 0.1. This allows $p_{\mathbf{z}_u}(\cdot;\boldsymbol{\theta_u})$ to explore around potential high-value regions.

After initialization, $p_{\mathbf{z}_u}(\cdot;\boldsymbol{\theta_u})$ is optimized to minimize the feedback loss to avoid the generation of low-quality unlabeled data:

$$\boldsymbol{\theta}_{\boldsymbol{u}}^* = \underset{\boldsymbol{\theta_u}}{\operatorname{argmin}}\, \mathbb{E}_{\mathbf{Z}_u\sim p_{\mathbf{z}_u}} \mathcal{L}_f\big(\mathcal{D}_l; \boldsymbol{\theta}_{\boldsymbol{S}}^*(\boldsymbol{\theta_T}), \mathbf{Z}_u\big). \tag{15}$$

We adopt the reparametrization trick (Kingma & Welling, 2013) for optimization. By introducing a random vector $\mathbf{r} \sim p_{\mathbf{r}} := \mathcal{N}(\mathbf{0}, I)$ and a mapping function $g(\mathbf{r};\boldsymbol{\theta_u}) = \boldsymbol{\mu_u} + \mathbf{r}\otimes\boldsymbol{\sigma_u}$, we can sample unlabeled data $\mathbf{z}_u := g(\mathbf{r};\boldsymbol{\theta_u})$. With this, the gradient for updating $\boldsymbol{\theta_u}$ can be written as follows:

$$\begin{aligned}&\nabla_{\boldsymbol{\theta_u}} \mathbb{E}_{\mathbf{Z}_u\sim p_{\mathbf{z}_u}} \mathcal{L}_f\big(\mathcal{D}_l; \boldsymbol{\theta}_{\boldsymbol{S}}^*(\boldsymbol{\theta_T}), \mathbf{Z}_u\big)\\ &= \nabla_{\boldsymbol{\theta_u}} \mathbb{E}_{\mathbf{R}\sim p_{\mathbf{r}}} \mathcal{L}_f\big(\mathcal{D}_l; \boldsymbol{\theta}_{\boldsymbol{S}}^*(\boldsymbol{\theta_T}), g(\mathbf{R};\boldsymbol{\theta_u})\big)\end{aligned} \tag{16}$$

where $\mathbf{R} \in \mathbb{R}^{M\times d}$ is a batch of M samples $\{\mathbf{r}_i\}_{i=1}^M$. We incorporate the update of $\boldsymbol{\theta_u}$ to the alternating one-step scheme for $\boldsymbol{\theta_S}$ and $\boldsymbol{\theta_T}$, as detailed in Appendix A.

5. Experiments

We employ multiple challenging blackbox optimization datasets to demonstrate TSBO's superior sample efficiency compared to recent baselines. Our results highlight that the proposed selective regularization forms the foundation of TSBO's enhanced performance. Through detailed ablation studies, we quantify the distinct contributions of each component within TSBO, providing insights into how these components collectively influence its overall performance.

5.1. Experimental Settings

We conduct experiments on three challenging high-dimensional global optimization tasks based on two datasets. The first dataset comprises 40,000 single-variable arithmetic expressions, and is employed for an **arithmetic expression reconstruction** task (Kusner et al., 2017). The second ZINC250K dataset (Sterling & Irwin, 2015), consisting of 250,000 molecules, is used for two **chemical design** tasks with two objective molecule profiles: the penalized water-octanol partition coefficient (Penalized LogP) (Gómez-Bombarelli et al., 2018) and the Ranolazine Multi-Property Objective (Ranolazine MPO) (Brown et al., 2019). Detailed details of these tasks can be found in Appendix C.1.

Table 1: Mean and standard deviation of the best value found after 250 data queries.

Method	Expression (↓)	Penalized LogP (↑)	Ranolazine MPO (↑)
Sobol (Owen, 2003)	1.261±0.689	3.019±0.296	0.260±0.046
LS-BO (Gómez-Bombarelli et al., 2018)	0.579±0.356	4.019±0.366	0.523±0.084
W-LBO (Tripp et al., 2020)	0.475±0.137	7.306±3.551	0.633±0.059
T-LBO (Grosnit et al., 2021)	0.572±0.268	5.695±1.254	0.620±0.043
TSBO-GEV	0.396±0.070	18.40±7.890	0.708±0.032
TSBO-Gaussian	**0.240±0.168**	**25.02±4.794**	**0.744±0.030**

Table 2: A broader comparison on the Chemical Design Task to maximize the Penalized LogP

Method	n_{Init}	n_{Query}	Penalized LogP (↑)	Top 1 Penalized LogP (↑)
MolDQN (Zhou et al., 2019)	250,000	$\geq 5,000$	N/A	11.84
W-LBO (Tripp et al., 2020)	200	500	12.09±7.576	21.74
	250,000	500	N/A	27.84
T-LBO (Grosnit et al., 2021)	200	500	10.82±4.688	16.45
	250,000	500	26.11	29.06
LOL-BO (Maus et al., 2022)	250,000	500	27.53±2.393	N/A
PG-LBO (Chen et al., 2023b)	1000	500	23.23±2.913	N/A
TSBO	200	500	**28.04±3.731**	**32.92**

Table 3: A broader comparison on the Expression Task

Method	n_{Init}	n_{Query}	Expression (↓)
W-LBO	100	500	0.386±0.016
	40000	500	0.314±0.1436
T-LBO	100	500	0.475±0.172
PG-LBO	100	500	0.358±0.195
TSBO	100	250	**0.240±0.168**

Baseline Methods TSBO is benchmarked against 5 VAE-based latent space optimization baselines: LS-BO (Gómez-Bombarelli et al., 2018), W-LBO (Tripp et al., 2020), T-LBO (Grosnit et al., 2021), LOL-BO (Maus et al., 2022), and PG-LBO (Chen et al., 2023b). LS-BO performs BO in the latent space with a fixed pre-trained Variational Autoencoder (VAE) (Kingma & Welling, 2013); W-LBO periodically fine-tunes the VAE with current labeled data; T-LBO introduces deep metric learning to W-LBO by additionally minimizing the triplet loss of the labeled data; LOL-BO optimizes the GP surrogate and VAE simultaneously via maximizing the joint likelihood; PG-LBO calibrates the VAE to improve the GP surrogate's accuracy on labeled data and to minimize the reconstruction loss on synthetic unlabeled data. Additionally, we include a random search algorithm Sobol (Owen, 2003) and a reinforcement learning approach MOLDQN (Zhou et al., 2019) for reference.

TSBO's Configurations Two variants of TSBO are built on top of the baseline T-LBO (Grosnit et al., 2021). We denote TSBO with the optimized Gaussian distribution based unlabeled data sampling by **TSBO-Gaussian**, and that with the GEV distribution for sampling unlabeled data by **TSBO-GEV**. The full model setup can be found in Appendix C.

5.2. Sample Efficiency of TSBO

We set the initial amount of labeled data for starting off each BO run to 100 for the arithmetic expression reconstruction task and 200 for the two chemical design tasks, respectively. Figure 4 and Table 1 compare TSBO with four BO baselines and Sobol. Both TSBO-GEV and TSBO-Gaussian consistently outperform T-LBO and other baselines across all tasks within 250 additional function evaluations (new data queries). Notably, TSBO-Gaussian drastically surpasses all baselines within the first 50 new queries, and finds further improved target values subsequently. These results underscore the superior sample efficiency of TSBO and the effectiveness of the proposed selective regularization.

We present a broader comparison on two optimization tasks with different data query budget settings. Table 2 and Table 3 respectively show the results of the expression and chemical design task, where n_{Init} is the number of initial

labeled data used, and n_{Query} the number of new queries. TSBO attains the SOTA performance, surpassing all baseline models and achieving a remarkable reduction in the utilization of labeled data by up to 364.2x.

5.3. Generalization Improvement of Data Query GP Model of TSBO

The generalization of the GP data query model is key to the overall BO performance. Here, we demonstrate improved generalization resulting from incorporating the pseudo labels predicted by the TSBO teacher as additional GP training data. We assess the GP's accuracy across the whole search space (global) and in the high target value region (local) with a total of 100 test points. For the global assessment, we sample test data in the latent space from the VAE prior $\mathcal{N}(\mathbf{0}, \boldsymbol{I})$. The local assessment samples test data from a Gaussian distribution centered at the point corresponding to the best target value found with a small standard deviation of 0.01. The assessments are conducted after TSBO completes the final (250th) data query on the two Chemical Design tasks with 200 initial labeled molecules.

Table 4 reports the negative log-likelihood (NLL) loss of the posterior prediction of a vanilla GP fitted exclusively on labeled data vs. that of the TSBO GP fitted on both the labeled data and unlabeled data with pseudo labels predicted by the teacher, both evaluated on the test data. We use the abbreviations PL for pseudo-label, P-LogP for Penalized LogP, and R-MPO for Ranolazine MPO, respectively in the table. The TSBO GP shows superior global and local generalization, both of which may boost TSBO's sample efficiency. Local GP generalization is more critical for Bayesian optimization as it helps query data with a higher target value. The improvement in local GP generalization brought by TSBO is more pronounced, specifically on the Penalized LogP task.

5.4. Ablation Study

We conduct an ablation study to assess the efficacy of various ingredients of TSBO, namely: 1) selective regularization to the teacher, 2) uncertainty awareness of the teacher-student model, 3) optimized unlabeled data sampling. The experiments are conducted on the Chemical Design task, following the settings at the beginning of Section 5.2.

The results of the ablation study are presented in Table 5, where we denote uncertainty awareness by "UA". The effectiveness of each of the three techniques listed above is manifested by a drastic performance drop resulting from its removal from TSBO.

In addition, Appendix D demonstrates the robustness of our approach to the value of the feedback weight λ, and Appendix E shows the impact of the validation data selection.

Table 4: The NLL loss of data query model of TSBO on the Chemical Design Task

Data Query Model	Test Region	P-LogP	R-MPO
GP w/o PL	Global	0.881	-1.504
GP with PL	Global	**0.863**	**-2.019**
GP w/o PL	Local	4.228	-2.086
GP with PL	Local	**1.388**	**-2.391**

Table 5: TSBO's ablation on the Chemical Design Task

TSBO Variant	Penalized LogP (↑)
TSBO w Random Sampler	4.881±1.416
TSBO w/o UA	12.568±7.965
TSBO w/o Student's Feedback	17.557±6.998
TSBO	25.020±4.794

6. Related Work

Semi-Supervised Learning (SSL) Beyond the scope of Bayesian optimization, many SSL techniques have been developed to serve the general goal of reducing expensive labeled data use (Zhang et al., 2021; Wang et al., 2022). SSL often involves consistency regularization (Rasmus et al., 2015), maintaining the consistency of the model's predictions on unlabeled data under perturbations of either data (Miyato et al., 2019; Xie et al., 2020a; Berthelot et al., 2019; 2020; Sohn et al., 2020; Chen et al., 2023a) or the model (Laine & Aila, 2017; Tarvainen & Valpola, 2017; Xie et al., 2020b; Pham et al., 2021). The family of model perturbation approaches makes use of two separate models with one acting as the teacher and the other a student, where the student learns from pseudo labels (Lee et al., 2013) predicted by the teacher (Pham et al., 2021). The proposed TSBO is the first work integrating one of the above SSL approaches, i.e., the teacher-student paradigm into BO while introducing the *selective regularization* that is optimized for black-box global optimization.

High-dimensional Bayesian Optimization in a Latent Space Operating in a reduced dimensional latent space (Deshwal & Doppa, 2021; Notin et al., 2021) is essential for making BO applicable to high-dimensional optimization problems (Eriksson et al., 2019; Nayebi et al., 2019; Letham et al., 2020; Eriksson & Jankowiak, 2021; Papenmeier et al., 2022). For this, it is a common practice to employ a dimension reduction model in the form of linear or nonlinear projections or VAE.

Early latent space BO methods use a fixed dimension reduction model during the adaptive data query process. These models are either randomly initialized (Wang et al., 2016),

or pre-trained on unlabeled data (Kusner et al., 2017; Jin et al., 2018; Alperstein et al., 2019). Recent advances adapt the latent space, for example, by retraining the dimension reduction model with an additional loss defined on available labeled data alone (Eissman et al., 2018; Tripp et al., 2020; Grosnit et al., 2021; Maus et al., 2022), or in conjunction with sampled synthetic data (Chen et al., 2020; 2023b). W-LBO (Tripp et al., 2020) retrains the VAE-based dimension reduction with a weighted evidence lower bound (ELBO) on labeled data whose weight is proportional to the label value. T-LBO (Grosnit et al., 2021) introduces a triplet loss (Xing et al., 2002) to pull labeled data with similar target values together in the latent space. LOL-BO (Maus et al., 2022) simultaneously optimizes the GP surrogate and VAE, and adopts the local approach TurBO (Eriksson et al., 2019) in the latent space. PG-LBO (Chen et al., 2023b) updates the VAE using an MSE loss of the GP surrogate on labeled data cast in the latent space and weighted retraining (Tripp et al., 2020) on heuristically sampled unlabeled data with pseudo labels (Lee et al., 2013).

While the aforementioned approaches focus on optimizing dimension reduction, TSBO takes on an orthogonal approach to boost the generalization of the GP surrogate by utilizing high-quality "pseudo" training data predicted by the well-regularized teacher model. The introduced selective regularization of the teacher model is instrumented by the uncertainty-aware teacher-student interaction and a systematically optimized unlabeled data sampler, all tailored for Bayesian optimization. It is worth noting that the dimension reduction techniques employed in the surveyed BO methods may be integrated into our TSBO framework.

7. Discussion

The proposed (TSBO) approach presents the first work integrating teacher-student based semi-supervised learning to enable sample-efficient Bayesian optimization. TSBO incorporates uncertainty-aware teacher-student interactions and optimized unlabeled data sampling, which collectively implement the selective regularization to the teacher. This makes it possible to enhance the generalization of the GP data query model by leveraging high-quality pseudo labels predicted by the teacher. TSBO achieves superior performance in comparison with other competitive latent space BO algorithms under tight labeled data budgets.

Opportunities exist for further improvement of TSBO in the future. For example, a more rigorous treatment of uncertainty quantification of the teacher-student model using techniques such as deep ensembles (Lakshminarayanan et al., 2017) and Bayesian neural networks (Hernandez-Lobato & Adams, 2015) may be explored to better mitigate the risk of predicted pseudo labels.

Acknowledgement

This material is based upon work supported by the National Science Foundation under Grant No. 1956313, and by the U.S. Department of Energy, through the Office of Advanced Scientific Computing Research's "Data-Driven Decision Control for Complex Systems (DnC2S)" project under award number DE-SC0021321.

Impact Statement

The proposed TSBO has the potential for significant positive impacts in various domains. By effectively finding the optimum compared with baselines on multiple datasets, TSBO offers a promising solution to enhance the efficiency and efficacy of optimization processes given limited labeled data and evaluation budgets. For instance, in engineering and manufacturing, TSBO can facilitate outlier detection, failure analysis, and the design of more efficient processes, leading to increased productivity and reduced resource consumption. By enabling faster and more accurate optimization, TSBO can ultimately benefit society as a whole.

Even though TSBO holds great promise, it is important to acknowledge and mitigate potential negative impacts. One concern is the overreliance on automated optimization algorithms, which could lead to a decreased emphasis on human intuition and creativity. TSBO should be used as a supportive tool that enhances human decision-making rather than replacing it entirely. Additionally, there is a risk of bias in the optimization process if the training data used for the teacher model contain inherent biases. Careful attention must be given to the training data to ensure fair and unbiased optimization results.

In conclusion, TSBO offers significant potential for broad impact in optimization tasks. By improving the efficiency and efficacy of optimization processes, TSBO can accelerate the discovery of optimal solutions, benefiting various industries and ultimately improving the well-being of individuals and society at large. However, it is important to consider and mitigate potential negative impacts, such as overreliance on automation and the risk of bias, to ensure that TSBO is used responsibly and ethically. With proper safeguards and considerations, TSBO can be a valuable tool that enhances human expertise and drives advancements in optimization across diverse domains.

References

Alperstein, Z., Cherkasov, A., and Rolfe, J. T. All smiles variational autoencoder. *arXiv preprint arXiv:1905.13343*, 2019.

Bastos, L. S. and O'hagan, A. Diagnostics for gaussian process emulators. *Technometrics*, 51(4):425–438, 2009.

Berthelot, D., Carlini, N., Goodfellow, I., Papernot, N., Oliver, A., and Raffel, C. A. Mixmatch: A holistic approach to semi-supervised learning. In Wallach, H., Larochelle, H., Beygelzimer, A., d'Alché-Buc, F., Fox, E., and Garnett, R. (eds.), *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.

Berthelot, D., Carlini, N., Cubuk, E. D., Kurakin, A., Sohn, K., Zhang, H., and Raffel, C. Remixmatch: Semi-supervised learning with distribution alignment and augmentation anchoring, 2020.

Brochu, E., Cora, V. M., and De Freitas, N. A tutorial on bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. *arXiv preprint arXiv:1012.2599*, 2010.

Brown, N., Fiscato, M., Segler, M. H., and Vaucher, A. C. Guacamol: benchmarking models for de novo molecular design. *Journal of chemical information and modeling*, 59(3):1096–1108, 2019.

Chen, H., Tao, R., Fan, Y., Wang, Y., Wang, J., Schiele, B., Xie, X., Raj, B., and Savvides, M. Softmatch: Addressing the quantity-quality trade-off in semi-supervised learning, 2023a.

Chen, J., Zhu, G., Yuan, C., and Huang, Y. Semi-supervised embedding learning for high-dimensional bayesian optimization. *arXiv preprint arXiv:2005.14601*, 2020.

Chen, T., Duan, Y., Li, D., Qi, L., Shi, Y., and Gao, Y. Pg-lbo: Enhancing high-dimensional bayesian optimization with pseudo-label and gaussian process guidance. *arXiv preprint arXiv:2312.16983*, 2023b.

Deshwal, A. and Doppa, J. Combining latent space and structured kernels for bayesian optimization over combinatorial spaces. In Ranzato, M., Beygelzimer, A., Dauphin, Y., Liang, P., and Vaughan, J. W. (eds.), *Advances in Neural Information Processing Systems*, volume 34, pp. 8185–8200. Curran Associates, Inc., 2021.

Eissman, S., Levy, D., Shu, R., Bartzsch, S., and Ermon, S. Bayesian optimization and attribute adjustment. In *Proc. 34th Conference on Uncertainty in Artificial Intelligence*, 2018.

Eriksson, D. and Jankowiak, M. High-dimensional Bayesian optimization with sparse axis-aligned subspaces. In de Campos, C. and Maathuis, M. H. (eds.), *Proceedings of the Thirty-Seventh Conference on Uncertainty in Artificial Intelligence*, volume 161 of *Proceedings of Machine Learning Research*, pp. 493–503. PMLR, 27–30 Jul 2021.

Eriksson, D., Pearce, M., Gardner, J., Turner, R. D., and Poloczek, M. Scalable Global Optimization via Local Bayesian Optimization. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.

Fisher, R. A. and Tippett, L. H. C. Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Mathematical Proceedings of the Cambridge Philosophical Society*, 24(2):180–190, 1928. doi: 10.1017/S0305004100015681.

Frazier, P. I. A tutorial on bayesian optimization. *arXiv preprint arXiv:1807.02811*, 2018.

Gao, W., Fu, T., Sun, J., and Coley, C. Sample efficiency matters: a benchmark for practical molecular optimization. *Advances in Neural Information Processing Systems*, 35:21342–21357, 2022.

Gneiting, T. and Raftery, A. E. Strictly proper scoring rules, prediction, and estimation. *Journal of the American statistical Association*, 102(477):359–378, 2007.

Gomes, M. I. and Guillou, A. Extreme value theory and statistics of univariate extremes: a review. *International statistical review*, 83(2):263–292, 2015.

Gómez-Bombarelli, R., Wei, J. N., Duvenaud, D., Hernández-Lobato, J. M., Sánchez-Lengeling, B., Sheberla, D., Aguilera-Iparraguirre, J., Hirzel, T. D., Adams, R. P., and Aspuru-Guzik, A. Automatic chemical design using a data-driven continuous representation of molecules. *ACS central science*, 4(2):268–276, 2018.

Grosnit, A., Tutunov, R., Maraval, A. M., Griffiths, R.-R., Cowen-Rivers, A. I., Yang, L., Zhu, L., Lyu, W., Chen, Z., Wang, J., et al. High-dimensional bayesian optimisation with variational autoencoders and deep metric learning. *arXiv preprint arXiv:2106.03609*, 2021.

Hernandez-Lobato, J. M. and Adams, R. Probabilistic back-propagation for scalable learning of bayesian neural networks. In Bach, F. and Blei, D. (eds.), *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pp. 1861–1869, Lille, France, 07–09 Jul 2015. PMLR.

Hu, H., Li, P., and Huang, J. Z. Parallelizable bayesian optimization for analog and mixed-signal rare failure detection with high coverage. In *2018 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, pp. 1–8. IEEE Press, 2018. doi: 10.1145/3240765.3240835.

Hu, H., Shah, M., Huang, J. Z., and Li, P. Global adversarial attacks for assessing deep learning robustness. *arXiv preprint arXiv:1906.07920*, 2019.

Jean, N., Xie, S. M., and Ermon, S. Semi-supervised deep kernel learning: Regression with unlabeled data by minimizing predictive variance. *Advances in Neural Information Processing Systems*, 31, 2018.

Jin, W., Barzilay, R., and Jaakkola, T. Junction tree variational autoencoder for molecular graph generation. In *International conference on machine learning*, pp. 2323–2332. PMLR, 2018.

Jones, D. R., Schonlau, M., and Welch, W. J. Efficient global optimization of expensive black-box functions. *Journal of Global optimization*, 13(4):455–492, 1998.

Kingma, D. P. and Welling, M. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.

Kushner, H. J. A New Method of Locating the Maximum Point of an Arbitrary Multipeak Curve in the Presence of Noise. *Journal of Basic Engineering*, 86(1):97–106, 03 1964. ISSN 0021-9223. doi: 10.1115/1.3653121.

Kusner, M. J., Paige, B., and Hernández-Lobato, J. M. Grammar variational autoencoder. In *International conference on machine learning*, pp. 1945–1954. PMLR, 2017.

Laine, S. and Aila, T. Temporal ensembling for semi-supervised learning, 2017.

Lakshminarayanan, B., Pritzel, A., and Blundell, C. Simple and scalable predictive uncertainty estimation using deep ensembles. In *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.

Lee, D.-H. et al. Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks. In *Workshop on challenges in representation learning, ICML*, volume 3, pp. 896. Atlanta, 2013.

Letham, B., Calandra, R., Rai, A., and Bakshy, E. Re-examining linear embeddings for high-dimensional bayesian optimization. In Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M., and Lin, H. (eds.), *Advances in Neural Information Processing Systems*, volume 33, pp. 1546–1558. Curran Associates, Inc., 2020.

Liang, X. Image-based post-disaster inspection of reinforced concrete bridge systems using deep learning with bayesian optimization. *Computer-Aided Civil and Infrastructure Engineering*, 34(5):415–430, 2019.

Lukasik, J., Jung, S., and Keuper, M. Learning where to look–generative nas is surprisingly efficient. *arXiv preprint arXiv:2203.08734*, 2022.

Maus, N., Jones, H., Moore, J., Kusner, M. J., Bradshaw, J., and Gardner, J. Local latent space bayesian optimization over structured inputs. *Advances in Neural Information Processing Systems*, 35:34505–34518, 2022.

McInnes, L., Healy, J., and Melville, J. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018.

Miyato, T., Maeda, S.-I., Koyama, M., and Ishii, S. Virtual adversarial training: A regularization method for supervised and semi-supervised learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(8): 1979–1993, 2019. doi: 10.1109/TPAMI.2018.2858821.

Moriconi, R., Deisenroth, M. P., and Sesh Kumar, K. High-dimensional bayesian optimization using low-dimensional feature spaces. *Machine Learning*, 109(9): 1925–1943, 2020.

Nair, V. and Hinton, G. E. Rectified linear units improve restricted boltzmann machines. In *Icml*, 2010.

Nayebi, A., Munteanu, A., and Poloczek, M. A framework for Bayesian optimization in embedded subspaces. In Chaudhuri, K. and Salakhutdinov, R. (eds.), *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 4752–4761. PMLR, 09–15 Jun 2019.

Nix, D. and Weigend, A. Estimating the mean and variance of the target probability distribution. In *Proceedings of 1994 IEEE International Conference on Neural Networks (ICNN'94)*, volume 1, pp. 55–60 vol.1, 1994. doi: 10.1109/ICNN.1994.374138.

Notin, P., Hernández-Lobato, J. M., and Gal, Y. Improving black-box optimization in vae latent space using decoder uncertainty. In Ranzato, M., Beygelzimer, A., Dauphin, Y., Liang, P., and Vaughan, J. W. (eds.), *Advances in Neural Information Processing Systems*, volume 34, pp. 802–814. Curran Associates, Inc., 2021.

Owen, A. B. Quasi-monte carlo sampling. *Monte Carlo Ray Tracing: Siggraph*, 1:69–88, 2003.

Papenmeier, L., Nardi, L., and Poloczek, M. Increasing the scope as you learn: Adaptive bayesian optimization in nested subspaces. In Koyejo, S., Mohamed, S., Agarwal,

A., Belgrave, D., Cho, K., and Oh, A. (eds.), *Advances in Neural Information Processing Systems*, volume 35, pp. 11586–11601. Curran Associates, Inc., 2022.

Pham, H., Dai, Z., Xie, Q., and Le, Q. V. Meta pseudo labels. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 11557–11568, 2021.

Rasmus, A., Berglund, M., Honkala, M., Valpola, H., and Raiko, T. Semi-supervised learning with ladder networks. In Cortes, C., Lawrence, N., Lee, D., Sugiyama, M., and Garnett, R. (eds.), *Advances in Neural Information Processing Systems*, volume 28. Curran Associates, Inc., 2015.

Seeger, M. Gaussian processes for machine learning. *International journal of neural systems*, 14(02):69–106, 2004.

Sohn, K., Berthelot, D., Carlini, N., Zhang, Z., Zhang, H., Raffel, C. A., Cubuk, E. D., Kurakin, A., and Li, C.-L. Fixmatch: Simplifying semi-supervised learning with consistency and confidence. In Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M., and Lin, H. (eds.), *Advances in Neural Information Processing Systems*, volume 33, pp. 596–608. Curran Associates, Inc., 2020.

Srinivas, N., Krause, A., Kakade, S., and Seeger, M. Gaussian process optimization in the bandit setting: No regret and experimental design. In *Proceedings of the 27th International Conference on International Conference on Machine Learning*, ICML'10, pp. 1015–1022, Madison, WI, USA, 2010. Omnipress. ISBN 9781605589077.

Sterling, T. and Irwin, J. J. Zinc 15–ligand discovery for everyone. *Journal of chemical information and modeling*, 55(11):2324–2337, 2015.

Tarvainen, A. and Valpola, H. Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results. In Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., and Garnett, R. (eds.), *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.

Tripp, A., Daxberger, E., and Hernández-Lobato, J. M. Sample-efficient optimization in the latent space of deep generative models via weighted retraining. *Advances in Neural Information Processing Systems*, 33:11259–11272, 2020.

Wang, Y., Chen, H., Fan, Y., SUN, W., Tao, R., Hou, W., Wang, R., Yang, L., Zhou, Z., Guo, L.-Z., Qi, H., Wu, Z., Li, Y.-F., Nakamura, S., Ye, W., Savvides, M., Raj, B., Shinozaki, T., Schiele, B., Wang, J., Xie, X., and Zhang, Y. Usb: A unified semi-supervised learning benchmark for classification. In Koyejo, S., Mohamed, S., Agarwal, A., Belgrave, D., Cho, K., and Oh, A. (eds.), *Advances in Neural Information Processing Systems*, volume 35, pp. 3938–3961. Curran Associates, Inc., 2022.

Wang, Z., Hutter, F., Zoghi, M., Matheson, D., and De Feitas, N. Bayesian optimization in a billion dimensions via random embeddings. *Journal of Artificial Intelligence Research*, 55:361–387, 2016.

Xie, Q., Dai, Z., Hovy, E., Luong, T., and Le, Q. Unsupervised data augmentation for consistency training. In Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M., and Lin, H. (eds.), *Advances in Neural Information Processing Systems*, volume 33, pp. 6256–6268. Curran Associates, Inc., 2020a.

Xie, Q., Luong, M.-T., Hovy, E., and Le, Q. V. Self-training with noisy student improves imagenet classification. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020b.

Xing, E., Jordan, M., Russell, S. J., and Ng, A. Distance metric learning with application to clustering with side-information. *Advances in neural information processing systems*, 15, 2002.

Zhang, B., Wang, Y., Hou, W., WU, H., Wang, J., Okumura, M., and Shinozaki, T. Flexmatch: Boosting semi-supervised learning with curriculum pseudo labeling. In Ranzato, M., Beygelzimer, A., Dauphin, Y., Liang, P., and Vaughan, J. W. (eds.), *Advances in Neural Information Processing Systems*, volume 34, pp. 18408–18419. Curran Associates, Inc., 2021.

Zhou, Z., Kearnes, S., Li, L., Zare, R. N., and Riley, P. Optimization of molecules via deep reinforcement learning. *Scientific reports*, 9(1):1–10, 2019.

Zoph, B., Vasudevan, V., Shlens, J., and Le, Q. V. Learning transferable architectures for scalable image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 8697–8710, 2018.

Algorithm 1 Bi-Level Optimization of the Teacher-Student Model

Input: Epochs L, feedback weight λ, teacher $T(\cdot;\boldsymbol{\theta}_{\boldsymbol{T}}^0)$, student $S(\cdot;\boldsymbol{\theta}_{\boldsymbol{S}}^0)$, labeled data $\mathcal{D}_l$, unlabeled data $\mathbf{Z}_u$
Output: Pseudo labels $\hat{\mathbf{y}}_u$
for $i = 1$ to L **do**
  Generate pseudo labels: $\hat{\mathbf{y}}_u \leftarrow \boldsymbol{\mu}_{\boldsymbol{\theta_T}}(\mathbf{Z}_u;\boldsymbol{\theta}_{\boldsymbol{T}}^{i-1})$
  Update the student model: $\boldsymbol{\theta}_{\boldsymbol{S}}^i \leftarrow \boldsymbol{\theta}_{\boldsymbol{S}}^{i-1} - \eta_S \cdot \nabla_{\boldsymbol{\theta}_{\boldsymbol{S}}^{i-1}} \mathcal{L}_u(\mathcal{D}_u(\boldsymbol{\theta}_{\boldsymbol{T}}^{i-1});\boldsymbol{\theta}_{\boldsymbol{S}}^{i-1})$
  Fix $\boldsymbol{\theta}_{\boldsymbol{S}}^i$, and update the teacher model: $\boldsymbol{\theta}_{\boldsymbol{T}}^i \leftarrow \boldsymbol{\theta}_{\boldsymbol{T}}^{i-1} - \eta_T \cdot \nabla_{\boldsymbol{\theta}_{\boldsymbol{T}}^{i-1}} \{\lambda \mathcal{L}_f(\mathcal{D}_l;\boldsymbol{\theta}_{\boldsymbol{S}}^i) + \mathcal{L}_l(\mathcal{D}_l;\boldsymbol{\theta}_{\boldsymbol{T}}^{i-1})\}$
end for
Predict pseudo labels: $\hat{\mathbf{y}}_u \leftarrow \boldsymbol{\mu}_{\boldsymbol{\theta_T}}(\mathbf{Z}_u;\boldsymbol{\theta}_{\boldsymbol{T}}^L)$

A. Alternating One-step Update Rule

To solve it in a computationally efficient way, we adopt an alternating one-step gradient-based approximation method from Pham et al. (2021). When unlabeled data $\mathbf{Z}_u$ are sampled from the distribution $p_{\mathbf{z}_u}(\cdot|\boldsymbol{\theta}_{\boldsymbol{u}})$, we adopt the reparameterization trick to optimize $\boldsymbol{\theta}_{\boldsymbol{u}}$. In the ith training iteration of the teacher-student, we update $\boldsymbol{\theta}_{\boldsymbol{u}}^{i-1}$ with a learning rate η_u:

- Sample $\mathbf{Z}_u$ with reparameterization trick: $\mathbf{Z}_u = g(\mathbf{R}, \boldsymbol{\theta}_{\boldsymbol{u}}^{i-1})$ where $\mathbf{R} \sim p_{\mathbf{r}}$;
- Update the student model: $\boldsymbol{\theta}_{\boldsymbol{S}}^i \leftarrow \boldsymbol{\theta}_{\boldsymbol{S}}^{i-1} - \eta_S \cdot \nabla_{\boldsymbol{\theta}_{\boldsymbol{S}}^{i-1}} \mathcal{L}_u(\mathcal{D}_u(\boldsymbol{\theta}_{\boldsymbol{T}});\boldsymbol{\theta}_{\boldsymbol{S}}^{i-1})$;
- Fix $\boldsymbol{\theta}_{\boldsymbol{S}}^i$, and update the teacher model: $\boldsymbol{\theta}_{\boldsymbol{T}}^i \leftarrow \boldsymbol{\theta}_{\boldsymbol{T}}^{i-1} - \eta_T \cdot \nabla_{\boldsymbol{\theta}_{\boldsymbol{T}}^{i-1}} \{\lambda \mathcal{L}_f(\mathcal{D}_l;\boldsymbol{\theta}_{\boldsymbol{S}}^i) + \mathcal{L}_l(\mathcal{D}_l;\boldsymbol{\theta}_{\boldsymbol{T}}^{i-1})\}$.
- Fix $\boldsymbol{\theta}_{\boldsymbol{S}}^i$, and update $\boldsymbol{\theta}_{\boldsymbol{u}}^i$: $\boldsymbol{\theta}_{\boldsymbol{u}}^i \leftarrow \boldsymbol{\theta}_{\boldsymbol{u}}^{i-1} - \eta_u \nabla_{\boldsymbol{\theta}_{\boldsymbol{u}}^{i-1}} \mathcal{L}_f(\mathcal{D}_l;\boldsymbol{\theta}_{\boldsymbol{S}}^i)$.

B. Dynamic Selection of Validation Data

It is attempting to use the full set of available labeled data $\mathcal{D}_l$ as the validation set $\mathcal{D}_v$ to assess the student, as proposed in (Pham et al., 2021) for image classification proposes. However, it is not always optimal under the setting of BO, whose objective is to find the global optimum using an overall small amount of labeled data. Hence, the assessment of the student, which provides feedback to the teacher, shall be performed in a way to improve the accuracy of the teacher-student model in predicting the global optimum. Since the majority of labeled data $\mathcal{D}_l$ are used in training the teacher, the quality of pseudo labels around $\mathcal{D}_l$ is high. Thus, validating the student using $\mathcal{D}_l$ may lead to a low averaged loss, however, which is not necessarily indicative of the model's capability in predicting the global optimum. Our empirical study shows that the performance of `TSBO` improves as the validation data are chosen to be the ones with higher target values. This is meaningful in the sense that assessing the student in regions with target values closer to the global optimum provides the best feedback to the teacher for improving its accuracy at places where it is most needed. We adopt a practical way to dynamically choose $\mathcal{D}_v$ at each BO iteration: the subset of $\mathcal{D}_l$ with the K highest label values. For this, we apply a fast sorting algorithm to rank all labeled data. Appendix E demonstrates the effectiveness of the proposed selection approach for $\mathcal{D}_v$.

C. Experimental Details

C.1. High-dimensional Optimization Tasks

Arithmetic Expression Reconstruction Task: The objective is to discover a single-variable arithmetic expression $\mathbf{x}^*$ = `1 / 3 + v + sin( v * v )`. For an input expression $\mathbf{x}$, the objective function is a distance metric $f(\mathbf{x}) = \max\{-7, -\log(1 + \text{MSE}(\mathbf{x}(\mathbf{v}) - \mathbf{x}^*(\mathbf{v}))\}$, where $\mathbf{v}$ are 1,000 evenly spaced numbers in $[-10, 10]$. A grammar VAE (Kusner et al., 2017) with a latent space of dimension 25 is adopted. It is pre-trained on a dataset of 40,000 expressions (Kusner et al., 2017).

Chemical Design Task: The purpose of this task is to design a molecule with a required molecular property profile. The objective profiles considered are 1) the penalized water-octanol partition coefficient (Penalized LogP) (Gómez-Bombarelli et al., 2018), and 2) the Ranolazine Multiproperty Objective (Ranolazine MPO) (Brown et al., 2019). A Junction-Tree VAE (Jin et al., 2018) with a latent space of dimension 56 and pre-trained on the ZINC250k dataset (Sterling & Irwin, 2015).

For each task, prior to optimization, a VAE is pre-trained using unlabeled data through the maximization of the ELBO

(Kingma & Welling, 2013), and all methods employ this pre-trained VAE at the outset of optimization.

C.2. TSBO's Model Architecture and Hyperparameters

In TSBO, the teacher model is a multilayer perceptron (MLP) with 5 hidden layers and ReLU activation (Nair & Hinton, 2010). The output dimension is 2. The student model is a standard GP with an RBF kernel.

For the purpose of reproducibility, we provide a comprehensive account of the hyperparameters employed in all our experiments using TSBO. Our approach is based on T-LBO, and thus we adopt the default hyperparameters of VAE as suggested by (Grosnit et al., 2021). The remaining hyperparameters, specific to TSBO, are presented in Table 6.

Table 6: Hyperparameters of TSBO

Group	Hyperparameter Name	Expression	Chemical Design
Common	Number of training steps in each BO iteration	20	20
	Number of warm-up steps	2,000	2,000
	Feedback weight	10^{-1}	10^{-1}
	Number of validation data	10	30
	Number of sampled unlabeled data	100	300
	Acquisition Function	EI	EI
	Acquisition optimizer	LBFGS	LBFGS
Teacher	Learning rate	10^{-3}	10^{-4}
	Batch size of labeled data	256	32
	Optimizer	Adam	Adam
Student	Kernel	RBF	RBF
	Prior mean	Constant	Constant
	Learning rate	10^{-2}	10^{-2}
	Optimizer	Adam	Adam
Data Query GP	Kernel	RBF	RBF
	Prior mean	Constant	Constant

C.3. Training of VAE in TSBO

Although TSBO stands as a general BO framework, it has been seamlessly integrated into T-LBO (Grosnit et al., 2021), a state-of-the-art VAE-based BO method, to facilitate a fair comparison. The training approach for the VAE remains unaltered, aligning with T-LBO's methodology:

- Pretrain: Before the first BO iteration, the VAE is trained on the dataset in an unsupervised way to maximize the ELBO (Kingma & Welling, 2013);
- Fine-tune: After each 50 BO iteration, the VAE is trained on all labeled data for 1 epoch to both maximize the ELBO and minimize the triplet loss which penalizes data having similar labels located far away. The weight of triplet loss is set to 10 in the Expression task and 1 in the Chemical Design task.

The training schemes of all models proposed in TSBO and the VAE are decoupled, rendering T-LBO an apt baseline for validating TSBO's sample efficacy.

D. The Influence of the Feedback Weight in TSBO

We analyze the influence of the selection of the feedback weight λ. Our experiments demonstrate that in a large range of λ, TSBO consistently outperforms the baseline T-LBO, underscoring the robustness of our approach to this hyper-parameter.

As shown in Table 7, while the selection of λ in $\{0.001.0.01, 0.1, 1\}$ has an impact on TSBO's performance, for each considered λ, TSBO-Gaussian consistently outperforms T-LBO, indicating that our success is not contributed to a deliberate λ selection.

Table 7: The ablation test of the weight of the feedback loss

Method	λ	Expression (↓)	Penalized LogP (↑)	Ranolazine MPO (↑)
T-LBO	-	0.572±0.268	5.695±1.254	0.620±0.043
TSBO-Gaussian	0.001	0.240±0.168	21.106±8.960	0.713±0.021
	0.01	0.433±0.260	21.384±1.533	0.720±0.040
	0.1	0.450±0.130	25.021±4.794	0.744±0.030
	1	0.474±0.113	21.122±7.494	0.712±0.023

E. Additional Ablation Test of Validation Data Selection

In order to verify the effectiveness of the proposed dynamic selection of validation data $\mathcal{D}_v$ in `TSBO`, where $\mathcal{D}_v$ is the subset of $\mathcal{D}_l$ with the K highest label values, we conduct an ablation study to compare it (TSBO-Gaussian) with two non-optimized $\mathcal{D}_v$ selection strategies: random K examples uniformly sampled from $\mathcal{D}_l$ (TSBO-Gaussian-ValRand), and the current labeled dataset $\mathcal{D}_l$ (TSBO-Gaussian-ValAll). These three variants of `TSBO` are measured on the Chemical Design task, and we report their average best values of 5 runs starting with 200 initial labeled data and a new data query budget of 250. K is set to 30.

As shown in Table 8, despite each variant of `TSBO` outperforms the baseline T-LBO, both TSBO-Gaussian-ValRand and TSBO-Gaussian-ValAll are less competitive than TSBO-Gaussian in finding the global maximum. This result meets our expectations: the student's feedback on those examples with the K highest label values is more beneficial for the training of the teacher-student model, and eventually better facilitating the search for the maximum.

Table 8: Comparison of validation data selection of `TSBO` on the Chemical Design task

Method	Validation Data Selection	Penalized LogP (↑)
T-LBO (Grosnit et al., 2021)	-	5.70
TSBO-Gaussian-ValRand	Random 30	21.60
TSBO-Gaussian-ValAll	All labeled data	22.65
TSBO-Gaussian	Top 30	**25.02**

F. Robustness of `TSBO` to Noisy Labels

In this section, we demonstrate the robustness of `TSBO` in noisy environments. We compare `TSBO` with the other baselines on the Chemical Design task, where all labels are subject to additional i.i.d. zero-mean Gaussian noises, with a standard deviation (std) of 0.1. All methods start with 200 initial labeled data and query 250 new examples. We report the mean and the standard deviation of the best values found by each method among 5 runs in Table 9.

Table 9: Labels with white Gaussian noises on the Chemical Design Task

method	Best Penalized LogP (↑)	
	Noise variance=0	Noise variance=0.1
Sobol	3.019±0.296	3.06±0.38
LS-BO	4.019±0.366	4.41±0.68
W-LBO	7.306±3.551	4.44±0.27
T-LBO	5.695±1.254	4.22±0.68
TSBO-GEV	18.40±7.89	20.73±6.24
TSBO-Gaussian	**25.02±4.794**	**25.97±4.82**

While nearly all of the baselines, especially T-LBO, exhibit decreases in the noisy scenario compared to the results obtained

without observation noise, TSBO shows no performance deterioration. This phenomenon demonstrates the noise-resistant capability of the proposed uncertainty-aware teacher-student model.