arXiv:2305.02614v3 [cs.LG] 3 Feb 2024

基于半监督学习和优化的无标签数据采样的高维贝叶斯优化

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

摘要

我们提出了一种新颖的半监督学习方法，称为教师-学生贝叶斯优化（TSBO），首次将师生范式融入贝叶斯优化，以尽可能减少昂贵的有标签数据查询。TSBO由教师模型、无标签数据采样器和学生模型组成。学生模型利用采样器生成的无标签数据位置及教师预测的伪标签进行训练。这三个组成部分的相互作用通过学生反馈，对教师实施独特的选择性正则化。该方案使教师能够预测高质量的伪标签，从而提升高斯过程代理模型在搜索空间中的泛化能力。为充分发挥TSBO的作用，我们提出了两种优化的无标签数据采样器，以构建与贝叶斯优化目标高度一致的有效学生反馈。此外，我们量化并利用师生模型的不确定性，在伪标签预测存在风险时，为教师提供可靠反馈。在多个全局优化任务中，即使带标签数据预算有限，TSBO也显著提升了样本效率。

1. 引言

贝叶斯优化（BO）（Brochu 等人，2010年）被广泛应用于黑箱优化，尤其适用于目标函数评估成本高昂或难以直接进行的情况。贝叶斯优化通过构建目标函数的代理模型，例如高斯过程（GP）（Seeger，2004年），然后依据平衡探索与利用的准则，迭代选择最有希望的位置，以获取新的带标签数据（Kushner，1964年；Jones 等人，1998年；Srinivas 等人，2010年）。近期研究进一步拓展了贝叶斯优化在高维

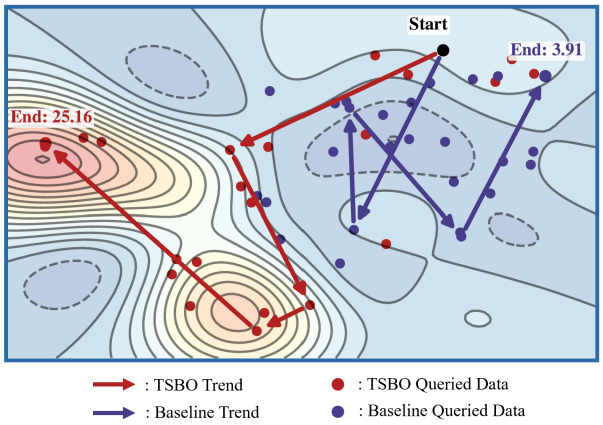


图 1：在高维分子设计任务中，为最大化惩罚对数P评分而查询的数据（圆点）及其趋势（箭头序列）的可视化（Sterling 与 Irwin，2015年）（G´omez-Bombarelli 等人，2018年）。红色和蓝色分别代表TSBO 和基线方法（采用原始贝叶斯优化）。两种方法的评估预算均为 450。

采用各种降维方法处理任务，包括线性嵌入（Wang 等人，2016；Chen 等人，2020）、非线性投影（Moriconi 等人，2020）和深度自动编码器（Kusner 等人，2017；Jin 等人，2018；Tripp 等人，2020；Chen 等人，2020；Grosnit 等人，2021；Maus 等人，2022；Chen 等人，2023b）。

尽管取得了这些令人鼓舞的进展，获取带标签数据的成本仍然很高，并且成为贝叶斯优化在许多应用领域中的关键瓶颈，例如功能性分子设计（Brown 等人，2019年；Gao 等人，2022年）、结构优化（Zoph 等人，2018年；Lukasik 等人，2022年）以及故障分析（Hu 等人，2018年；Liang，2019年）。

为应对黑盒优化中数据查询效率面临的挑战，我们提出了一种统一的半监督学习（SSL）方法，称为师生贝叶斯优化（TSBO）。TSBO 是首个将师生模型融入贝叶斯优化的工作，并通过实现一种独特的 选择性正则化 机制，弥合了半监督学习与贝叶斯优化目标之间的差距，使其能够利用海量

[1]美国加利福尼亚大学圣巴巴拉分校电子与计算机工程系。通信作者：Yuxuan Yin <y yin@ucsb.edu>。

初步工作。正在审稿。

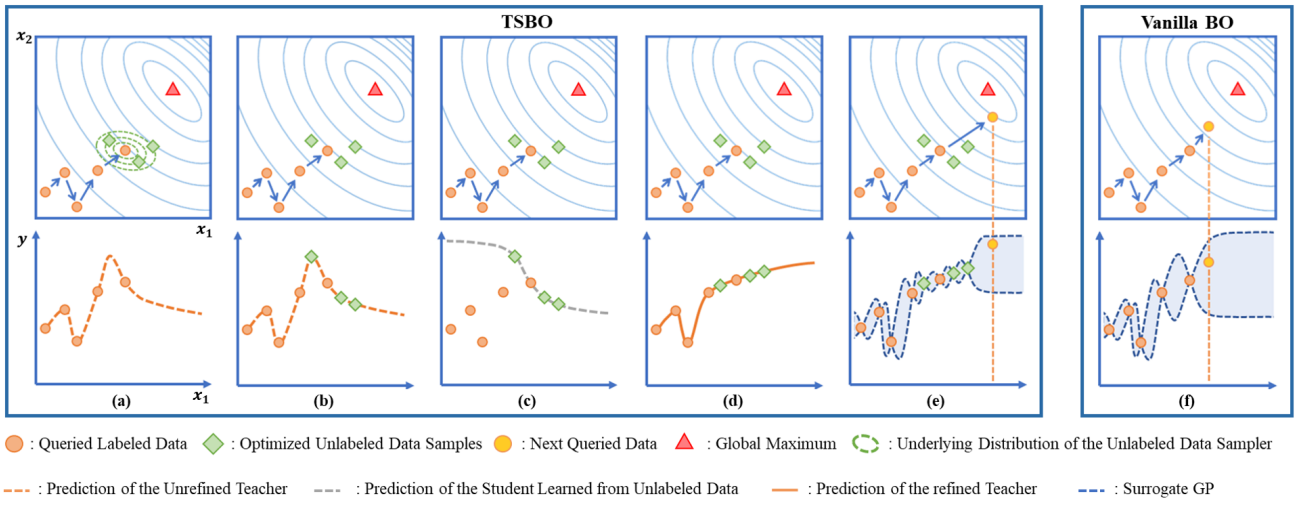


图 2：示意图展示了无标签数据采样器、教师模型和学生模型之间的交互，以及选择性正则化的应用。(a)：从可能具有较高取值的区域采样无标签数据。(b)：教师模型为无标签数据预测伪标签。(c)：学生模型从无标签数据及其预测伪标签中学习，并在带标签数据上进行评估，作为反馈。(d)：教师模型根据反馈改进其预测。(e)：TSBO 中的高斯过程同时使用带标签数据和带有改进后伪标签的无标签数据进行拟合。(f)：原始贝叶斯优化中的高斯过程仅使用带标签数据进行拟合。

样本空间中的部分无标签数据，这些数据成本低廉。针对性地利用这些数据，并使其服务于贝叶斯优化的优化目标，从而更有效地引导优化过程趋向高价值区域。与基线方法相比，该方法显著提高了数据查询效率。图 TSBO顺序优化过程的示意示例见图 1，展示了TSBO卓越的样本效率。

TSBO 的核心包含一个教师模型、一个学生模型和一个无标签数据采样器。这三个组件之间的相互作用，实现了我们针对教师模型的选择性正则化 。最终，经正则化的教师模型会预测高质量的伪标签[1]，这些伪标签与查询得到的带标签数据相结合，从而更好地训练标准高斯过程代理模型，以便在贝叶斯优化过程中更有效地查询新的带标签数据。

如图 2所示，在每次贝叶斯优化迭代中，无标签数据采样器会选取一组经过优化的无标签数据位置，并将其传递给教师模型。教师模型利用当前知识为这些位置预测伪标签（Lee 等人，2013年；Pham等人，2021年）。随后，学生模型仅使用教师预测的伪标签进行训练。考虑到教师模型预测的伪标签可能具有误导性，我们使用现有的真实值带标签数据评估学生模型，并将其表现反馈给教师模型。之后，教师模型会得到改进并将学生反馈作为选择性正则化纳入其中。师生交互完成后，经过正则化改进的教师模型能够提供可靠的伪标签，用于训练 BP代理模型，从而增强其查询新数据的能力，即使带标签训练数据有限，也能如此。

为充分发挥TSBO的作用，必须仔细设计选择性正则化所涉及的关键组成部分。我们没有对无标签数据进行随机采样，而是通过无标签数据采样器，有策略地将无标签数据放置在伪标签质量较高的区域，同时鼓励朝全局最优解进行探索。基于按此策略采样的无标签数据向教师模型提供学生反馈，可以评估并改进教师模型，使其与贝叶斯优化的整体目标高度一致。我们提出了两种优化的无标签数据采样器：一种基于极值理论（Fisher & Tippett，1928），另一种基于参数化采样分布。

此外，我们通过使师生模型具备不确定性感知能力来提升 TSBO 的性能。教师模型不仅为选定的无标签数据分配伪标签，还会量化其不确定性。在学生训练过程中会考虑这些不确定性，从而向教师模型提供更可靠的反馈，尤其是在存在不可靠伪标签预测风险的情况下。

我们在各种具有挑战性的高维数据集上评估了所提出的 TSBO，并展示了更高的数据效率

[1]虽然“标签”一词在分类问题中通常指真实值，但在贝叶斯优化任务中也广泛用于表示观测值（Grosnit 等人，2021年；Chen 等人，2020年；Jean等人，2018年）。

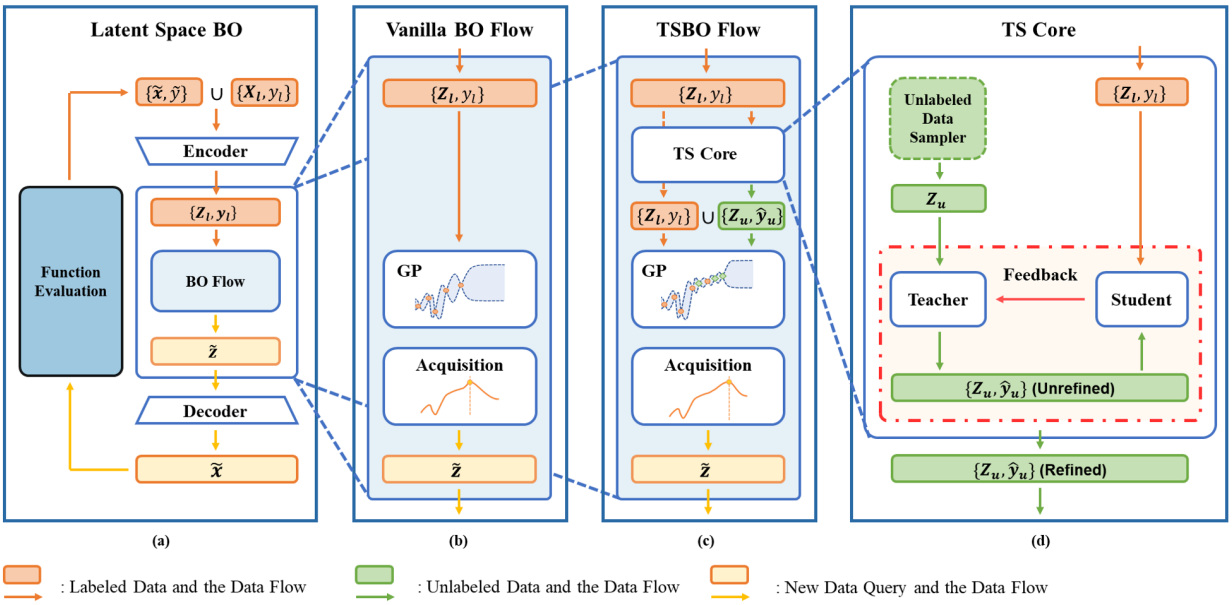


图 3：TSBO 框架概览。(a)：基本的潜空间贝叶斯优化架构。(b)：原始贝叶斯优化流程仅利用编码后的带标签数据训练高斯过程模型，并查询下一个数据点。(c)：TSBO 流程引入 TS 核心，在每次贝叶斯优化迭代期间向高斯过程模型提供额外的高质量无标签数据。(d)：TS 核心内部：优化后的无标签数据采样器以及来自学生模型的反馈，为教师模型提供选择性正则化。

提升。在化学设计任务（Sterling 与 Irwin，2015年）和表达式重构任务（Kusner 等人，2017年）中，与近期研究相比，我们分别仅用 3‰ 和 1% 的评估次数便取得了最先进的结果。

2. 预备知识

2.1. 贝叶斯优化（BO）

贝叶斯优化（BO）(Brochu 等人，2010年) 旨在确定一个未知的、通常为非凸函数的全局最大值 f： $\mathcal{X} \to \mathbb{R}$，定义如下：

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

其中， $\mathcal{X} \subseteq \mathbb{R}^D$ 表示一个 D维输入空间。

贝叶斯优化通过迭代选择数据点$\mathbf{x}$ 进行评估，并基于先前查询点的结果逐步构建。在单次贝叶斯优化迭代中，给定一组 N已评估样本 $\{\mathbf{x}_i,\ \mathbf{y}_i\}_{i=1}^N = \{\mathbf{X}_l,\ \mathbf{y}_l\}$，其中$\mathbf{X}_l$ 是一个输入矩阵，$\mathbf{y}_l$ 是对应输出的向量，则下一个查询点 $\tilde{}\mathbf{x}$ 由以下方式确定：

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

此处， $Q\big(\mathbf{x}|\mathbf{X}_l,\mathbf{y}_l\big)$ 表示由概率模型 Q 估计的$\mathbf{x}$的后验分布，通常称为代理模型(Frazier, 2018年)。采集函数 α旨在平衡对已知高值区域的利用与对未知空间的探索。

2.2. 潜空间贝叶斯优化

对于高维问题，在原始空间中直接应用贝叶斯优化（BO）可能导致数据查询模型过拟合 Q，尤其是在数据查询预算受限的情况下；这一困境通常被称为维数灾难（Brochu 等人，2010年）。一种可行的解决方案是潜空间贝叶斯优化（G´omez-Bombarelli 等人，2018年; Kusner 等人，2017年），它在生成式潜在空间 $\mathcal{Z} \subseteq \mathbb{R}^d$ 中 $d \ll D$进行贝叶斯优化。该方法利用编码器 ψ： $\mathcal{X} \to \mathcal{Z}$ 和解码器 φ： $\mathcal{Z} \to \mathcal{X}$来引导查询过程。在潜空间贝叶斯优化中，结合当前带标签数据$\mathbf{X}_l$，$\mathbf{y}_l$得到的下一个查询$\mathbf{x}$ $\tilde{}\mathbf{y}$包括：

- 在 $\mathcal{Z}$ 中使用潜代码对拟合数据查询模型 Q $\mathbf{Z}_l$,$\mathbf{y}_l$，其中 $\mathbf{Z}_l := \psi(\mathbf{X}_l)$;
- 通过最大化 α 中的采集函数来识别最优潜在代码 $\tilde{}\mathbf{z}$ $\mathcal{Z}$： $\tilde{}\mathbf{z} = \operatorname{argmax}_{\mathcal{Z}}\, \alpha\Big(Q\big(\mathbf{z}|\mathbf{Z}_l,\mathbf{y}_l\big)\Big)$;
- Decoding $\tilde{\mathbf{x}} = \varphi(\tilde{\mathbf{z}})$ and evaluating $\tilde{y} = f(\tilde{\mathbf{x}})$.

图 3 （a）展示了基本的潜空间贝叶斯优化框架。

3. TSBO 问题形式化

TSBO 框架在传统潜空间贝叶斯优化的基础上，提升了其在高维挑战性任务中的数据效率。该框架保留了编码器-解码器结构，并遵循既定的数据查询流程，同时TSBO 还引入了教师模型、学生模型和无标签数据采样器，以便在每次数据查询之间对潜在空间进行选择性正则化。这一关键步骤利用大量高质量无标签数据及其对应的校准伪标签丰富代理模型，从而提升模型的预测性能，如图 3 （c）和图 3（d）所示。

令 $T(\cdot;\boldsymbol{\theta_T})$ 表示教师模型， $S(\cdot;\boldsymbol{\theta_S})$ 表示学生模型，$p_{\mathbf{z}_u}(\cdot;\boldsymbol{\theta_u})$ 表示无标签数据采样器所依据的分布。集合 $\mathcal{D}_l := \{\mathbf{Z}_l, \mathbf{y}_l\}$ 表示当前步骤可用的带标签数据。以下将对这些模块之间的交互进行形式化描述，并将其表述为一个双层优化问题：

1. Unlabeled data $\mathbf{Z}_u$ is sampled from $p_{\mathbf{z}_u}(\cdot;\boldsymbol{\theta_u})$.

2. 教师模型在无标签数据样本上预测伪标签：$\hat{\mathbf{y}}_u(\boldsymbol{\theta_T}) := T(\mathbf{Z}_u;\boldsymbol{\theta_T})$。

3. 学生模型通过在无标签数据集 $\mathcal{D}_u(\boldsymbol{\theta_T})$ 上定义的无标签损失 $\mathcal{L}_u$ 进行训练：$= \{\mathbf{Z}_u, \hat{\mathbf{y}}_u\}$:

$$\boldsymbol{\theta}^*_{\boldsymbol{S}}(\boldsymbol{\theta_T}) = \underset{\boldsymbol{\theta_S}}{\operatorname{argmin}}\, \mathcal{L}_u\big(\mathcal{D}_u(\boldsymbol{\theta_T}); \boldsymbol{\theta_S}\big) \tag{3}$$

4. 通过反馈损失 $\mathcal{L}_f\big($对训练后的学生模型在有标签数据集$\mathcal{D}_l$ 上进行评估； $\boldsymbol{\theta}^*_{\boldsymbol{S}}(\boldsymbol{\theta_T})\big)$

5. 通过优化有标签损失 $\mathcal{L}_l$ 与反馈损失 $\mathcal{L}_f$的组合来更新教师模型：

$$\boldsymbol{\theta}^*_{\boldsymbol{T}} = \underset{\boldsymbol{\theta_T}}{\operatorname{argmin}}\, \mathcal{L}_l(\mathcal{D}_l;\boldsymbol{\theta_T}) + \lambda \mathcal{L}_f\big(\mathcal{D}_l; \boldsymbol{\theta}^*_{\boldsymbol{S}}(\boldsymbol{\theta_T})\big) \tag{4}$$

其中 λ 是一个加权参数。

4. TSBO 模块的详细设计

4.1. 不确定性感知教师-学生模型

教师模型与学生模型之间的相互作用，对于改进教师模型的预测以及提高采样无标签数据上伪标签的准确性至关重要。尤为重要的是，我们强调纳入不确定性感知，以维护反馈机制的完整性。不准确的伪标签可能会扰乱建模过程，导致错误的反馈。通过量化预测的不确定性，学生模型能够更审慎地使用伪标签，并根据其相关的不确定性水平调整对伪标签的依赖程度。因此，我们提出不确定性感知教师-学生机制，这是一种实用机制，能够有效利用伪标签，同时最大限度地降低潜在风险。

4.1.1. 不确定性感知教师模型

准确量化不确定性本质上十分复杂且计算开销巨大。在TSBO中，我们采用一种受（Nix & Weigend，1994）启发的启发式方法，以多层感知机（MLP）作为教师模型。该 MLP 旨在针对任意给定输入$\mathbf{Z}$，不仅输出预测均值，还输出对角协方差矩阵，从而提供输出的概率估计：

$$T(\mathbf{Z};\boldsymbol{\theta_T}) := \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{\theta_T}}(\mathbf{Z}), \boldsymbol{\Sigma}_{\boldsymbol{\theta_T}}(\mathbf{Z})), \tag{5}$$

$$\text{where} \quad \boldsymbol{\Sigma}_{\boldsymbol{\theta_T}}(\mathbf{Z}) := \boldsymbol{\sigma}^{\mathbf{2}}_{\boldsymbol{\theta_T}}(\mathbf{Z}) \cdot \boldsymbol{I} \tag{6}$$

训练所定义的教师模型，就是通过最小化负对数似然（NLL）损失来拟合参数化的高斯分布。因此，公式（4）中的有标签损失写为：

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

结合教师模型不确定性的伪标签生成如下：

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

4.1.2. 不确定性感知学生模型

在我们的方法中，学生模型采用高斯过程（GP）实现，如公式（9）所示，它能够高效地将教师模型的不确定性预测作为先验知识纳入其中。这种配置使学生模型不仅能够改进自身预测，还能针对教师模型的均值预测及相关不确定性提供精确反馈。

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

其中， $\kappa_0(\cdot,\ \cdot))$ 是径向基函数（RBF）核函数， σ_0^2是加性方差。学生模型参数 $\boldsymbol{\theta_S}$包含核参数和方差 σ_0^2。

我们通过构造依赖于学生模型伪标签的先验，将教师模型的不确定性 ϵ_u（$\mathbf{z}_u$）传播到学生高斯过程模型的下游训练中：$\hat{\mathrm{y}}_u = \epsilon_u$（$\mathbf{z}_u$）$+ \epsilon_{\kappa 0}$，其中 $\epsilon_{\kappa 0} \sim \mathcal{N}$（$\mathbf{0}$，$\kappa_0$（$\mathbf{z}_u$，$\mathbf{z}_u$）$+ \sigma_0^2$）是方程（9）中的学生模型原始先验方差。学生模型方差的优化 $\epsilon_{\kappa 0}$ 取决于教师模型的不确定性估计$\boldsymbol{\Sigma}_{\boldsymbol{\theta_T}}$（$\mathbf{z}_u$）。

相应地，学生模型在无标签数据集 $\mathcal{D}_u$（$\boldsymbol{\theta_T}$）上的协方差矩阵 $\boldsymbol{\Sigma}_u$ 计算如下： $\boldsymbol{\Sigma}_{u_{ij}} =$

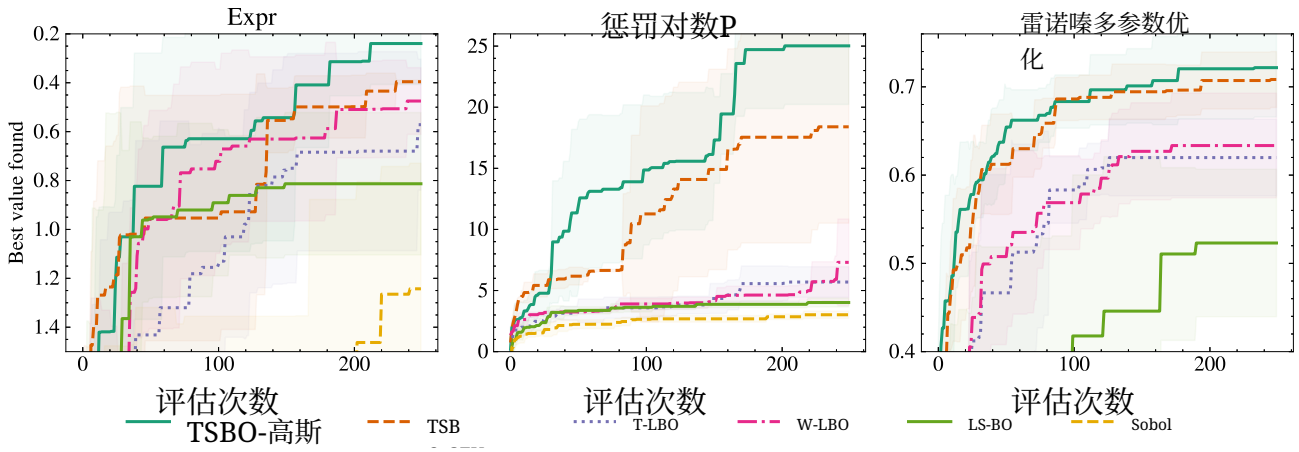


图 4：比较 4 个 LSO 基线方法与 TSBO 的平均性能和标准差 .

$\mathbb{E}(\hat{\mathrm{y}}_{u_i} - \mathbb{E}\,\hat{\mathrm{y}}^{u_i})(\hat{\mathrm{y}}_{u_j} - \mathbb{E}\,\hat{\mathrm{y}}^{u_j}) = \kappa_0(\mathbf{z}^{u_i}, \mathbf{z}^{u_j}) + \delta_{ij}\sigma_0^2 + \delta_{ij}\boldsymbol{\Sigma}_{\boldsymbol{\theta_T}}(\mathbf{z}^{u_i})$，其中 δ 表示克罗内克δ。它是教师模型和学生模型的不确定性之和：

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

我们在无标签数据集 $\mathcal{D}_u(\boldsymbol{\theta_T})$ 上优化不确定性感知学生的参数 $\boldsymbol{\theta_S}$，以最小化负对数似然损失$\mathcal{L}_u$：

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

4.1.3. 反馈损失的推导

通过最小化方程（11中的无标签数据损失进行优化的学生高斯过程模型，在有标签数据集 D_l 上进行评估。该评估得出针对编码输入后验预测$\mathbf{Z}_l$。我们关注后验均值 $\boldsymbol{\mu}_{\boldsymbol{\theta_S}}(\mathbf{Z}_l; \mathcal{D}_u(\boldsymbol{\theta_T}))$，以量化模型性能，并将反馈损失定义为该后验均值与真实标签[2] 之间的均方误差（MSE）$\mathbf{y}_l$。

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

后验均值具有如下显式形式：

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

备注： 学生GP模型能够高效地将教师模型的不确定性纳入反馈，确保其决策充分考虑这一关键背景。具体而言，模型的后验预测对与教师模型高度不确定性相关的伪标签的依赖程度有所降低：教师模型针对第 $\boldsymbol{\Sigma}_{\boldsymbol{\theta_T}}(\mathbf{z}^{u_i})$ 个 i伪标签的预测方差会被加到i协方差矩阵 $\boldsymbol{\Sigma}_u$ 中第（10）式的第 $\boldsymbol{\Sigma}_{u_{ii}}$ 个对角元素上。当该不确定性显著高于其他伪标签的不确定性时，相应的对角元素会远大于其他对角元素。因此，根据式（13），第 i个伪标签在 $\hat{}\mathbf{y}_u$ 中对验证标签后验均值预测的贡献会显著降低。

4.2. 优化的无标签数据采样

随机采样常用于标准的半监督学习（SSL）场景，但由于其缺乏优化，在贝叶斯优化（BO）中效果不佳。教师模型为随机选取的无标签数据生成的伪标签，尤其是那些远离训练数据集的数据，质量往往较低，表现为均值较小或方差较大。这些伪标签可能会误导学生模型，使其偏离全局最优解。因此，信息不准确的学生模型无法向教师模型提供必要的反馈，帮助其改进并找到最优解。

我们认为，在广阔的样本空间中，只有特定子集的无标签数据具有重要价值，并且与贝叶斯优化目标相一致。利用这些无标签数据能够最有效地促进贝叶斯优化过程。因此，开发一种有针对性的无标签数据采样策略，以高效识别并筛选出这些高价值样本，至关重要。为实现这一目标，我们提出了两种新颖的技术，旨在优化无标签数据的采样分布，从而确保在贝叶斯优化场景中实现更加聚焦且高效的优化。

[2]反馈损失可以是均方误差、负的预测边际对数似然（Gneiting 和 Raftery，2007年）或负的马氏距离（Bastos 和 O'hagan，2009年）。为保证数值稳定性，我们在本文中选择均方误差。

表 1：250 次数据查询后找到的最佳值的均值和标准差。

方法	表达式 (↓)	惩罚对数P（↑）	雷诺嗪多参数优化 (↑)
Sobol（Owen，2003）	1.261±0.689	3.019±0.296	0.260±0.046
LS-BO(G´omez-Bombarelli 等人, 2018年)	0.579±0.356	4.019±0.366	0.523±0.084
W-LBO (Tripp 等人，2020年)	0.475±0.137	7.306±3.551	0.633±0.059
T-LBO（Grosnit 等人，2021年）	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-高斯	**0.240±0.168**	**25.02±4.794**	**0.744±0.030**

方法 1：基于极值理论（EVT）的无标签数据采样 核心思想是将无标签数据放置在高质量伪标签所在的区域，同时鼓励向全局最优解进行探索。为此，我们对带标签数据中极端部分的分布进行建模，即具有最佳目标值的数据。极值理论（Fisher & Tippett，1928年）指出，如果 $\{y_1, \cdots, y_N\}$是独立同分布的，且当 N 趋于无穷大时，其最大值y^*服从广义极值分布（GEV）（Fisher& Tippett，1928年）。我们在附录中讨论了广义极值分布（GEV）的优化方法和马尔可夫链蒙特卡罗（MCMC）采样方法（Andrieu 等人，2003年）（Hu 等人，2019年）A.1。

方法 2：通过学生反馈学习无标签数据采样分布 尽管所提出的广义极值分布方法在理论上为生成无标签数据提供了可靠的方法，但其实际效果受限于马尔可夫链蒙特卡罗采样技术的高计算开销。

为避免马尔可夫链蒙特卡罗带来的计算负担，我们尝试寻找一种替代方法来采样无标签数据，记作 $\mathbf{z}_u$，使其来自由 $p_{\mathbf{z}_u}(\cdot;\boldsymbol{\theta_u})$ 参数化的分布 $\boldsymbol{\theta_u}$。我们提出通过最小化反馈损失 $\mathcal{L}_f$来优化 $p_{\mathbf{z}_u}$。直观而言，较大的 $\mathcal{L}_f$ 意味着使用的无标签数据伪标签质量较差，这可能会误导师生模型。实践中，我们采用重参数化技巧（Kingma 与 Welling，2013年）来优化 $\boldsymbol{\theta_u}$，并将其整合到师生模型的训练循环中。详情见附录 A.2。

5. 实验

我们采用多个具有挑战性的黑盒优化数据集，展示 TSBO 相较于近期基线方法更高的样本效率。我们的结果表明，所提出的选择性正则化构成了 TSBO 性能提升的基础。通过详细的消融研究，我们量化了 TSBO 中各组件的独特贡献，并深入分析了这些组件如何共同影响其整体性能。

5.1. 实验设置

我们基于两个数据集，在三个具有挑战性的高维全局优化任务上开展实验。第一个数据集包含 40,000 个单变量算术表达式，用于**算术表达式重构**任务（Kusner 等人，2017年）。第二个 ZINC250K 数据集（Sterling 与 Irwin，2015年）包含 250,000 个分子，用于两个**化学设计**任务，具有两种目标分子特征：惩罚水-辛醇分配系数（惩罚对数P）（G´omez-Bombarelli 等人，2018年）和雷诺嗪多参数优化（雷诺嗪 MPO）（Brown 等人，2019年）。这些任务的详细信息见附录D.1。

基线方法TSBO 与 5 种基于 VAE 的潜在空间优化基线方法进行比较：LS-BO（G´omez- Bombarelli 等人，2018年）、W-LBO（Tripp 等人，2020年）、T-LBO（Grosnit 等人，2021年）、LOL-BO（Maus 等人，2022年）以及 PG-LBO（Chen 等人，2023b）。LS-BO 使用固定的预训练变分自编码器（VAE）在潜在空间中执行贝叶斯优化（Kingma 与 Welling，2013年）；W-LBO 定期使用当前带标签数据对 VAE 进行微调；T-LBO 在 W-LBO 的基础上引入深度度量学习，并额外最小化带标签数据的三元组损失；LOL-BO 通过最大化联合似然，同时优化高斯过程代理模型和 VAE；PG-LBO 对 VAE 进行校准，以提高高斯过程代理模型在带标签数据上的准确性，并最小化合成未标记数据上的重构损失。此外，我们还将随机搜索算法 Sobol（Owen，2003）和强化学习方法 MOLDQN（Zhou 等人，2019）作为参照。

TSBO的配置 TSBO 的两个变体 均基于基线方法 T-LBO（Grosnit 等人，2021年）。我们将基于优化后的高斯分布进行无标签数据采样的 TSBO **称为 TSBO-高斯**，将采用广义极值分布对无标签数据进行采样的变体称为 **TSBO-GEV**。完整的模型设置见附录D。

5.2. TSBO 的样本效率

对于算术表达式重构任务，我们将每次贝叶斯优化（BO）运行启动时的初始带标签数据量设定为 100。

表 2：在化学设计任务上最大化惩罚对数P的更广泛比较

方法	n_{Init}	n_{Query}	惩罚对数P（↑）	排名第 1 的惩罚对数P（↑）
MolDQN（Zhou 等人，2019年）	250,000	≥ 5,000	N/A	11.84
W-LBO（Tripp 等人，2020年）	200	500	12.09±7.576	21.74
	250,000	500	N/A	27.84
T-LBO（Grosnit 等人，2021年）	200	500	10.82±4.688	16.45
	250,000	500	26.11	29.06
LOL-BO（Maus 等人，2022年）	250,000	500	27.53±2.393	N/A
PG-LBO（Chen 等人，2023b）	1000	500	23.23±2.913	N/A
TSBO	200	500	**28.04±3.731**	**32.92**

表 3：表达式任务的更广泛比较

方法	n_{Init}	n_{Query}	表达式（↓）
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**

任务以及两个化学设计任务分别为 200。图 4 和表 1 对比了 TSBO 与四个 BO 基线方法及 Sobol。TSBO-GEV 和 TSBO- 高斯在所有任务中，均在额外进行 250 次函数评估（新数据查询）的范围内，持续优于 T-LBO 和其他基线方法。值得注意的是，TSBO-高斯在最初的 50 次新数据查询中便大幅超越所有基线方法，并在随后找到了进一步优化的目标值。这些结果凸显了 TSBO 卓越的样本效率，以及所提出的选择性正则化方法的有效性。

我们进一步比较了两项在不同数据查询预算设置下的优化任务。表2和表3 分别展示了表达式任务和化学设计任务的结果，其中 n_{Init} 表示使用的初始带标签数据数量，n_{Query} 表示新增查询数量。TSBO 取得了最先进性能，超越所有基线模型，并显著减少了带标签数据的使用量，最高可达 364.2 倍。

5.3. TSBO 的数据查询高斯过程模型的泛化能力提升

GP 数据查询模型的泛化能力是整体贝叶斯优化性能的关键。在此，我们展示了将 TSBO 教师模型预测的伪标签纳入 GP 训练数据后，模型的泛化能力得到了提升。我们在整个搜索空间（全局）和目标值较高的区域（局部）中，使用总计 100 个测试点评估 GP 的准确性。全局评估时，我们从变分自编码器的先验$\mathcal{N}(\mathbf{0}, \boldsymbol{I})$中，在潜在空间内抽取测试数据。局部评估时，我们从高斯分布中抽取测试数据，该分布以当前找到的目标值最佳点所对应的位置为中心，标准差较小，为 0.01。在两项化学设计任务中，TSBO 完成最后一次（第 250 次）数据查询后进行评估；每项任务均使用 200 个初始带标签分子。

表4 报告了仅使用带标签数据拟合的原始高斯过程与使用带标签数据和教师模型预测的伪标签无标签数据共同拟合的TSBO 高斯过程的后验预测负对数似然（NLL）损失；两者均在测试数据上评估。表中分别使用 PL 表示伪标签、P-LogP 表示惩罚对数P、R-MPO 表示雷诺嗪多参数优化。TSBO 高斯过程的全局和局部泛化能力均更强，这可能提升TSBO的样本效率。对于贝叶斯优化而言，局部高斯过程泛化能力更为关键，因为它有助于查询目标值更高的数据。TSBO 带来的局部高斯过程泛化能力提升更为显著，尤其是在惩罚对数P任务中。

5.4. 消融研究

我们进行了消融研究，以评估TSBO各组成部分的有效性，具体包括：1）对教师模型进行选择性正则化；2）教师模型与学生模型对不确定性的感知；3）优化的无标签数据采样。实验在化学设计任务上进行，遵循第5.2节开头所述的设置。

消融研究的结果列于表5中，其中我们用“UA”表示不确定性感知。移除上述三种技术中的任意一种，都会导致TSBO的性能大幅下降，这体现了每种技术的有效性。

表 4：TSBO 在化学设计任务中数据查询模型的负对数似然损失

数据查询模型	测试区域	P-LogP	R-MPO
不带伪标签的高斯过程	全局	0.881	-1.504
带PL的高斯过程	全局	**0.863**	**-2.019**
不带PL的高斯过程	局部	4.228	-2.086
带PL的高斯过程	局部	**1.388**	**-2.391**

表5：TSBO 在化学设计任务上的消融实验

TSBO 变体	惩罚对数P（↑）
采用随机采样器的 TSBO	4.881±1.416
不含不确定性感知的 TSBO	12.568±7.965
不含学生反馈的 TSBO	17.557±6.998
TSBO	25.020±4.794

此外，附录E 展示了我们的方法对反馈权重取值的稳健性 λ，附录F 则展示了验证数据选择的影响。

6. 相关工作

半监督学习（SSL）除贝叶斯优化之外，研究者还开发了许多半监督学习技术，以实现减少对昂贵带标签数据使用量这一普遍目标（Zhang 等人，2021年；Wang 等人，2022年）。半监督学习通常涉及一致性正则化（Rasmus 等人，2015年），即在对数据（Miyato 等人，2019年；Xie 等人，2020a；Berthelot 等人，2019年；2020年；Sohn 等人，2020年；Chen 等人，2023a）或模型（Laine 和 Aila，2017年；Tarvainen 和 Valpola，2017年；Xie 等人，2020b；Pham等人，2021年）进行扰动时，保持模型对无标签数据预测的一致性。模型扰动方法这一系列技术采用两个独立模型，其中一个充当教师模型，另一个充当学生模型；学生模型从教师模型预测的伪标签中学习（Lee 等人，2013年；Pham等人，2021年）。本文提出的 TSBO 首次将上述半监督学习方法之一——师生范式——引入贝叶斯优化，并提出针对黑盒全局优化而优化的选择性正则化。

潜在空间中的高维贝叶斯优化 在降维后的潜在空间中运行（Deshwal 和 Doppa，2021年；Notin 等人，2021年）对于使贝叶斯优化适用于高维优化问题至关重要（Eriksson 等人，2019年；Nayebi 等人，2019年；Letham 等人，2020年；Eriksson 和 Jankowiak，2021年； Papenmeier 等人，2022年）。为此，通常采用线性或非线性投影，或变分自编码器形式的降维模型。

早期的潜空间贝叶斯优化方法在自适应数据查询过程中使用固定的降维模型。这些模型要么经过随机初始化（Wang 等人，2016），要么在无标签数据上预训练（Kusner 等人，2017年；Jin 等人，2018年；Alperstein 等人，2019年）。近期的进展会对潜在空间进行调整，例如，使用仅基于可用带标签数据定义的额外损失函数重新训练降维模型仅使用带标签数据（Eissman 等人，2018年；Tripp 等人，2020年；Grosnit 等人，2021年；Maus 等人，2022年），或将其与采样合成数据结合使用（Chen 等人，2020年；2023b）。W-LBO（Tripp 等人，2020年）通过在带标签数据上使用加权证据下界（ELBO）来重新训练基于 VAE 的降维模型，其中权重与标签值成正比。T-LBO（Grosnit 等人，2021年）引入三元组损失（Xing 等人，2002年），将目标值相似的带标签数据拉近至潜在空间中的相邻位置。LOL-BO（Maus 等人，2022年）同时优化高斯过程代理模型和 VAE，并在潜在空间中采用局部方法 TurBO（Eriksson 等人，2019年）。PG-LBO（Chen 等人，2023b）使用高斯过程代理模型在映射至潜在空间的带标签数据上的均方误差损失更新 VAE，并通过加权再训练（Tripp 等人，2020年）处理基于启发式方法采样且带有伪标签的无标签数据（Lee 等人，2013年）。

前述方法侧重于优化降维，而 TSBO 则采用了一种正交的方法，通过利用经过良好正则化的教师模型预测出的高质量“伪”训练数据，提升高斯过程代理模型的泛化能力。教师模型引入的选择性正则化通过不确定性感知教师-学生机制以及经过系统优化的无标签数据采样器来实现，这些设计均针对贝叶斯优化量身打造。值得注意的是，所调查的贝叶斯优化方法中采用的降维技术也可以集成到我们的 TSBO 框架中。

7. 讨论

所提出的 TSBO 方法首次将基于教师-学生的半监督学习融入贝叶斯优化，从而实现高样本效率的贝叶斯优化。TSBO 引入了不确定性感知教师-学生交互和优化的无标签数据采样，二者共同对教师实施选择性正则化。借助教师预测的高质量伪标签，该方法能够增强 GP 数据查询模型的泛化能力。在带标签数据预算紧张的情况下，与其他具有竞争力的潜在空间贝叶斯优化算法相比，TSBO 取得了更优的性能。

未来仍有机会进一步改进 TSBO。例如，可以对教师-学生模型的不确定性量化进行更严格的处理，

采用深度集成（Lakshminarayanan 等人，2017年）和贝叶斯神经网络（Hernandez-Lobato 和 Adams，2015年）等技术，以更好地降低预测伪标签所带来的风险。

参考文献

Alperstein, Z.、Cherkasov, A. 和 Rolfe, J.T. 全笑脸变分自编码器。*arXiv* 预印本 *arXiv:1905.13343*，2019年。

Andrieu, C.、deFreitas, N.、Doucet, A. 和 Jordan, M. I. 面向机器学习的马尔可夫链蒙特卡罗简介。*MachineLearning*，50(1–2)：5–43，2003年。

Bastos, L. S. 和 O'hagan, A. 高斯过程模拟器的诊断。*Technometrics*，51(4)：425–438，2009年。

Berthelot, D., Carlini, N., Goodfellow, I., Papernot, N., Oliver, A. 和 Raffel, C. A. Mixmatch：一种半监督学习的整体方法。载于 Wallach, H., Larochelle, H., Beygelzimer, A., d'Alch´e-Buc, F., Fox, E. 和 Garnett, R.（编），神经信息处理系统进展，第 32 卷。Curran Associates, Inc.，2019年。

Berthelot, D., Carlini, N., Cubuk, E. D., Kurakin, A., Sohn, K., Zhang, H. 和 Raffel, C. Remixmatch：结合分布对齐与增强锚定的半监督学习，2020年。

Brochu, E., Cora, V. M. 和 De Freitas, N. 关于昂贵代价函数的贝叶斯优化教程，并将其应用于主动用户建模和分层强化学习。*arXiv* 预印本 *arXiv:1012.2599*，2010年。

Brown, N., Fiscato, M., Segler, M. H. 和 Vaucher, A. C. Guacamol：用于从头分子设计的模型基准测试。《化学信息与建模杂志》，59(3)：1096–1108，2019年。

Chen, H., Tao, R., Fan, Y., Wang, Y., Wang, J., Schiele, B., Xie, X., Raj, B., andSavvides, M. Softmatch：解决半监督学习中的数量与质量权衡，2023a。

Chen, J.、Zhu, G.、Yuan, C. 和 Huang, Y. 面向高维贝叶斯优化的半监督嵌入学习。*arXiv* 预印本 *arXiv:2005.14601*，2020年。

Chen, T.、Duan, Y.、Li, D.、Qi, L.、Shi, Y. 和 Gao, Y. Pg-lbo：利用伪标签和高斯过程引导增强高维贝叶斯优化。*arXiv* 预印本 *arXiv:2312.16983*，2023b。

Deshwal, A. 和 Doppa, J. 结合潜在空间与结构化核函数，实现组合空间上的贝叶斯优化。载于 Ranzato, M.、Beygelzimer, A.、Dauphin, Y.、Liang, P. 和 Vaughan, J. W.（编），神经信息处理系统进展，第34卷，第8185–8200页。Curran Associates, Inc.，2021年。Eissman, S.、Levy, D.、Shu, R.、Bartzsch, S. 和 Ermon, S. 贝叶斯优化与属性调整。载于第*34*届人工智能不确定性会议论文集，2018年。

Eriksson, D. 和 Jankowiak, M. 基于稀疏轴对齐子空间的高维贝叶斯优化。载于 de Campos, C. 和 Maathuis, M. H.（编），第三十七届人工智能不确定性会议论文集，机器学习研究论文集第 161 卷，第 493–503 页。PMLR，2021年7月27–30日。

Eriksson, D., Pearce, M., Gardner, J., Turner, R. D. 和 Poloczek, M. 通过局部贝叶斯优化实现可扩展的全局优化。载于神经信息处理系统进展，第 32 卷。Curran Associates, Inc.，2019年。

Fisher, R. A. 和 Tippett, L. H. C. 样本中最大或最小成员频率分布的极限形式。剑桥哲学学会数学学报，24(2):180–190，1928年。doi: 10.1017/S0305004100015681。

Frazier, P. I. 贝叶斯优化教程。*arXiv*预印本 *arXiv:1807.02811*，2018年。

Gao, W., Fu, T., Sun, J. 和 Coley, C. 样本效率至关重要：实用分子优化基准。神经信息处理系统进展，35:21342–21357，2022年。

Gneiting, T. 和 Raftery, A. E. 严格适当的评分规则、预测与估计。《美国统计学会杂志》，102(477):359–378，2007年。

G´omez-Bombarelli, R., Wei, J. N., Duvenaud, D., Hern´andez-Lobato, J. M., S´anchez-Lengeling, B., She- berla, D., Aguilera-Iparraguirre, J., Hirzel, T. D., Adams, R. P. 和 Aspuru-Guzik, A. 利用分子的数据驱动连续表征进行自动化学设计。《*ACS Central Science*》，4(2):268–276，2018年。

Grosnit, A., Tutunov, R., Maraval, A. M., Griffiths, R.-R., Cowen-Rivers, A. I., Yang, L., Zhu, L., Lyu, W., Chen, Z., Wang, J., 等. 使用变分自编码器和深度度量学习进行高维贝叶斯优化。*arXiv* 预印本 *arXiv:2106.03609*，2021年。

Hernandez-Lobato, J. M. 和 Adams, R. 用于可扩展贝叶斯神经网络学习的概率反向传播。载于 Bach, F. 和 Blei, D.（编），第*32*届国际机器学习会议论文集，机器学习研究论文集第37卷，第1861–1869页，法国里尔，2015年7月7–9日。PMLR。

Hu, H.、Li, P. 和 Huang, J. Z. 通过高覆盖率的可并行贝叶斯优化进行模拟与混合信号罕见故障检测。载于 *2018 IEEE/ACM* 国际计算机辅助设计会议（*IC-CAD*），第 1–8 页。IEEE出版社，2018年。doi: 10.1145/3240765。3240835。

Hu, H.、Shah, M.、Huang, J. Z. 和 Li, P. 用于评估深度学习鲁棒性的全局对抗攻击。*arXiv* 预印本 *arXiv:1906.07920*，2019年。Jean, N.、Xie, S. M. 和 Ermon, S. 半监督深度核学习：通过最小化预测方差，利用无标签数据进行回归。神经信息处理系统进展，第 31 卷，2018年。Jin, W.、Barzilay, R. 和 Jaakkola, T. 用于分子图生成的连接树变分自编码器。载于国际机器学习会议，第 2323–2332 页。PMLR，2018年。Jones, D. R.、Schonlau, M. 和 Welch, W. J. 昂贵黑箱函数的高效全局优化。《全球优化杂志》，13(4):455–492，1998年。Kingma, D. P. 和 Welling, M. 自编码变分贝叶斯。*arXiv* 预印本 *arXiv:1312.6114*，2013年。Kushner, H. J. 噪声环境下定位任意多峰曲线最大值点的新方法。《基础工程杂志》，86(1):97–106，1964年 03 月。ISSN 0021-9223。doi：10.1115/1.3653121。Kusner, M. J.、Paige, B. 和 Hern´ández-Lobato, J. M. 语法变分自编码器。载于 国际机器学习会议，第 1945–1954 页。PMLR，2017年。Laine, S. 和 Aila, T. 面向半监督学习的时间集成，2017年。Lakshminarayanan, B.、Pritzel, A. 和 Blundell, C. 使用深度集成进行简单且可扩展的预测不确定性估计。载于神经信息处理系统进展，第 30 卷。Curran Associates, Inc.，2017年。Lee, D.-H. 等。伪标签：一种简单高效的深度神经网络半监督学习方法。载于表征学习挑战研讨会，*ICML*，第 3 卷，第 896 页。亚特兰大，2013年。

Letham, B., Calandra, R., Rai, A. 和 Bakshy, E. 重新审视用于高维贝叶斯优化的线性嵌入。载于 Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M. 和 Lin, H.（编），神经信息处理系统进展，第 33 卷，第 1546–1558 页。Curran Associates, Inc.，2020年。Liang, X. 使用深度学习和贝叶斯优化对钢筋混凝土桥梁系统进行基于图像的灾后检查。计算机辅助土木与基础设施工程，34(5):415–430，2019年。Lukasik, J., Jung, S. 和 Keuper, M. 学会观察何处——生成式 NAS 的效率出奇地高。*arXiv* 预印本 *arXiv:2203.08734*，2022年。Maus, N., Jones, H., Moore, J., Kusner, M. J., Bradshaw, J. 和 Gardner, J. 面向结构化输入的局部潜在空间贝叶斯优化。神经信息处理系统进展，35:34505–34518，2022年。Miyato, T., Maeda, S.-I., Koyama, M. 和 Ishii, S. 虚拟对抗训练：一种用于监督学习和半监督学习的正则化方法。*IEEE* 模式分析与机器智能汇刊，41(8): 1979–1993，2019年。doi: 10.1109/TPAMI.2018.2858821。Moriconi, R., Deisenroth, M. P. 和 Sesh Kumar, K. 利用低维特征空间进行高维贝叶斯优化。机器学习，109(9): 1925–1943，2020年。Nair, V. 和 Hinton, G. E. 修正线性单元提升受限玻尔兹曼机的性能。载于*Icml*，2010年。Nayebi, A., Munteanu, A. 和 Poloczek, M. 嵌入子空间中贝叶斯优化的框架。载于 Chaudhuri, K. 和 Salakhutdinov, R.（编），第 *36* 届国际机器学习会议论文集，机器学习研究论文集第 97 卷，第 4752–4761 页。PMLR，2019年6月9–15日。Nix, D. 和 Weigend, A. 估计目标概率分布的均值和方差。载于 *1994* 年 *IEEE* 神经网络国际会议（*ICNN'94*）论文集，第 1 卷，第 55–60 卷，第 1 册，1994年。doi: 10.1109/ICNN.1994.374138。Notin, P., Hern´ández-Lobato, J. M. 和 Gal, Y. 利用解码器不确定性改进 VAE 潜在空间中的黑箱优化。载于 Ranzato, M., Beygelzimer, A., Dauphin, Y., Liang, P. 和 Vaughan, J. W.（编），神经信息处理系统进展，第 34 卷，第 802–814 页。Curran Associates, Inc.，2021年。Owen, A. B. 准蒙特卡洛采样。蒙特卡洛光线追踪：*Siggraph*，1:69–88，2003年。

Papenmeier, L., Nardi, L. 和 Poloczek, M. 随着学习不断拓展范围：嵌套子空间中的自适应贝叶斯优化。载于 Koyejo, S., Mohamed, S., Agarwal, A., Belgrave, D., Cho, K. 和 Oh, A.（编），神经信息处理系统进展，第 35 卷，第 11586–11601 页。Curran Associates, Inc.，2022年。

Pham, H., Dai, Z., Xie, Q., 和 Le, Q. V. 元伪标签。载于*IEEE/CVF*计算机视觉与模式识别会议论文集，第 11557–11568 页，2021年。

Rasmus, A., Berglund, M., Honkala, M., Valpola, H., and Raiko, T. 基于梯形网络的半监督学习。载于 Cortes, C., Lawrence, N., Lee, D., Sugiyama, M., and Garnett, R.（编），神经信息处理系统进展，第28卷。Curran Associates, Inc., 2015年。Seeger, M. 用于机器学习的高斯过程。国际神经系统杂志，14(02):69–106，2004年。Sohn, K., Berthelot, D., Carlini, N., Zhang, Z., Zhang, H., Raffel, C. A., Cubuk, E. D., Kurakin, A., and Li, C.-L. Fixmatch：通过一致性和置信度简化半监督学习。载于 Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M., and Lin, H.（编），神经信息处理系统进展，第33卷，第596–608页。Curran Associates, Inc., 2020年。Srinivas, N., Krause, A., Kakade, S., and Seeger, M. 赌博机设定下的高斯过程优化：无遗憾与实验设计。载于第*27*届国际机器学习会议论文集，ICML'10，第1015–1022页，美国威斯康星州麦迪逊，2010年。Omnipress。ISBN 9781605589077。Sterling, T. and Irwin, J. J. Zinc 15——人人皆可进行配体发现。《化学信息与建模杂志》，55(11):2324–2337，2015年。Tarvainen, A. and Valpola, H. 平均教师是更好的榜样：权重平均一致性目标提升半监督深度学习效果。载于 Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., and Garnett, R.（编），神经信息处理系统进展，第30卷。Curran Associates, Inc., 2017年。Tripp, A., Daxberger, E., and Hern´andez-Lobato, J. M. 通过加权再训练，在深度生成模型的潜在空间中进行样本高效优化。神经信息处理系统进展，33:11259–11272，2020年。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：用于分类的统一半监督学习基准。载于 Koyejo, S., Mohamed, S., Agarwal, A., Belgrave, D., Cho, K., and Oh, A.（编），神经信息处理系统进展，第35卷，第3938–3961页。Curran Associates, Inc., 2022年。

Wang, Z., Hutter, F., Zoghi, M., Matheson, D., 和 De Feitas, N. 通过随机嵌入在十亿维空间中进行贝叶斯优化。人工智能研究期刊，55:361–387，2016年。

Xie, Q., Dai, Z., Hovy, E., Luong, T., 和 Le, Q. 用于一致性训练的无监督数据增强。载于 Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M., 和 Lin, H.（编），神经信息处理系统进展，第33卷，第6256–6268页。Curran Associates, Inc., 2020年a。Xie, Q., Luong, M.-T., Hovy, E., 和 Le, Q. V. 通过带噪学生模型的自训练提升ImageNet分类性能。载于*IEEE/CVF*计算机视觉与模式识别会议（*CVPR*）论文集，2020年6月b。Xing, E., Jordan, M., Russell, S. J., 和 Ng, A. 将距离度量学习应用于结合辅助信息的聚类。神经信息处理系统进展，15，2002年。Zhang, B., Wang, Y., Hou, W., WU, H., Wang, J., Okumura, M., 和 Shinozaki, T. FlexMatch：通过课程式伪标签增强半监督学习。载于 Ranzato, M., Beygelzimer, A., Dauphin, Y., Liang, P., 和 Vaughan, J. W.（编），神经信息处理系统进展，第34卷，第18408–18419页。Curran Associates, Inc., 2021年。Zhou, Z., Kearnes, S., Li, L., Zare, R. N., 和 Riley, P. 通过深度强化学习优化分子。*Scientific reports*，9(1):1–10，2019年。Zoph, B., Vasudevan, V., Shlens, J., 和 Le, Q. V. 学习可迁移架构以实现可扩展的图像识别。载于*IEEE*计算机视觉与模式识别会议论文集，第8697–8710页，2018年。

算法 1师生模型的双层优化

训练轮次 L，反馈权重 λ，教师模型 $T(\cdot;\ \boldsymbol{\theta}_{\boldsymbol{T}}^{0})$，学生模型 $S(\cdot;\ \boldsymbol{\theta}_{\boldsymbol{S}}^{0})$，带标签数据 $\mathcal{D}_l$，无标签数据 $\mathbf{Z}_u$输出：伪标签 $\hat{\mathbf{y}}_u$对于 $i=1$到 L 执行生成伪标签：$\hat{\mathbf{y}}_u \leftarrow \boldsymbol{\mu}_{\boldsymbol{\theta_T}}(\mathbf{Z}_u;\boldsymbol{\theta}_{\boldsymbol{T}}^{i-1})$ 更新学生模型：$\boldsymbol{\theta}_{\boldsymbol{S}}^{i} \leftarrow \boldsymbol{\theta}_{\boldsymbol{S}}^{i-1} - \eta_S \cdot \nabla_{\boldsymbol{\theta}_{\boldsymbol{S}}^{i-1}} \mathcal{L}_u(\mathcal{D}_u(\boldsymbol{\theta}_{\boldsymbol{T}}^{i-1});\ \boldsymbol{\theta}_{\boldsymbol{S}}^{i-1})$ 固定 $\boldsymbol{\theta}_{\boldsymbol{S}}^{i}$，并更新教师模型：$\boldsymbol{\theta}_{\boldsymbol{T}}^{i} \leftarrow \boldsymbol{\theta}_{\boldsymbol{T}}^{i-1} - \eta_T \cdot \nabla_{\boldsymbol{\theta}_{\boldsymbol{T}}^{i-1}} \{\lambda \mathcal{L}_f(\mathcal{D}_l;\ \boldsymbol{\theta}_{\boldsymbol{S}}^{i}) + \mathcal{L}_l(\mathcal{D}_l;\ \boldsymbol{\theta}_{\boldsymbol{T}}^{i-1})\}$循环结束预测伪标签：$\hat{\mathbf{y}}_u \leftarrow \boldsymbol{\mu}_{\boldsymbol{\theta_T}}(\mathbf{Z}_u;\boldsymbol{\theta}_{\boldsymbol{T}}^{L})$

A. 无标签数据采样的技术细节

A.1. 基于极值理论的无标签数据采样

广义极值分布（Fisher &Tippett，1928）定义为

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

其中 $\bar{\mathrm{y}} := (\mathrm{y}^* - a)/b$ 由广义极值分布的 3 个可学习参数定义：位置系数 $a \in \mathbb{R}$、尺度值 $b > 0$ 和分布形状参数 $\xi \in \mathbb{R}$。

我们拟合一个广义极值分布 p_{y^*}，其参数通过最小化若干极值标签的负对数似然损失进行估计。该广义极值分布捕捉了从当前已评估数据中观察到的最佳目标值的分布。因此，生成其预测标签遵循广义极值分布的无标签数据，使我们能够从现有极值带标签数据所在的区域出发，同时由于采样过程的随机性，探索目标值可能更高的点。一旦广义极值分布 p_{y^*} 拟合完成，我们采用一种针对广义极值分布定制的特定马尔可夫链蒙特卡罗方法（Hu 等人，2019年）从中进行采样。

A.2. 从学生反馈中学习的无标签数据采样分布

我们采用重参数化技巧（Kingma 与 Welling，2013年）作为首选采样策略。通过引入一个随机向量 $\mathbf{r} \in \mathcal{R} \subseteq \mathbb{R}^d$和一个映射函数 $g(\cdot;\boldsymbol{\theta_u}): \mathcal{R} \to \mathcal{Z}$，其中 $g(\mathbf{r};\ \boldsymbol{\theta_u}) \sim p_{\mathbf{z}_u}$当 $\mathbf{r} \sim p_{\mathbf{r}}$时，我们可以高效地采样无标签数据 $\mathbf{z}^u := g(\mathbf{r};\boldsymbol{\theta_u})$，使用 $p_{\mathbf{r}}$，一种已知且可便捷采样的分布，例如高斯分布。

通过最小化反馈损失来学习参数化采样分布，是一种合理的选择。较大的反馈损失表明所使用的无标签数据伪标签质量较差，这可能会误导师生模型。我们优化 $\boldsymbol{\theta_u}$ 以最小化反馈损失 $\mathcal{L}_f$:

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

更新 $\boldsymbol{\theta_u}$ 的梯度可使用重参数化技巧表示如下:

$$\nabla_{\boldsymbol{\theta_u}} \mathbb{E}_{\mathbf{Z}_u \sim p_{\mathbf{z}_u}} \mathcal{L}_f\left(\mathcal{D}_l; \boldsymbol{\theta_S}, \mathbf{Z}_u\right) = \nabla_{\boldsymbol{\theta_u}} \mathbb{E}_{\mathbf{R} \sim p_{\mathbf{r}}} \mathcal{L}_f\left(\mathcal{D}_l; \boldsymbol{\theta_S}, g(\mathbf{R};\boldsymbol{\theta_u})\right) \tag{16}$$

其中 $\mathbf{R} \in \mathbb{R}^{M\times d}$ 是一个由以下内容组成的批次： M 样本 $\{\mathbf{r}_i\}_{i=1}^{M}$。我们将 $\boldsymbol{\theta_u}$ 的更新纳入针对$\boldsymbol{\theta_S}$ 和 $\boldsymbol{\theta_T}$的交替单步方案中，详见附录 B。

B. 交替单步更新规则

为以计算高效的方式求解该问题，我们采用了Pham等人 (2021年提出的交替单步基于梯度的近似方法。当无标签数据 $\mathbf{Z}_u$ 从分布 $p_{\mathbf{z}_u}(\cdot|\boldsymbol{\theta_u})$中采样时，我们采用重参数化技巧来优化 $\boldsymbol{\theta_u}$。在师生模型的第 i 次训练迭代中，我们以学习率$\boldsymbol{\theta}_{\boldsymbol{u}}^{i-1}$更新 η_u:

• 使用重参数化技巧对样本进行采样：$\mathbf{Z}_u$ ， $\mathbf{Z}_u = g(\mathbf{R}\ \boldsymbol{\theta}_{\boldsymbol{u}}^{i-1})$，其中$\mathbf{R} \sim p$r</style};

- 更新学生模型： $\boldsymbol{\theta}_{\boldsymbol{S}}^{i} \leftarrow \boldsymbol{\theta}_{\boldsymbol{S}}^{i-1} - \eta_S \cdot \nabla_{\boldsymbol{\theta}_{\boldsymbol{S}}^{i-1}} \mathcal{L}_u(\mathcal{D}_u(\boldsymbol{\theta}_{\boldsymbol{T}}); \boldsymbol{\theta}_{\boldsymbol{S}}^{i-1})$;
- 修复 $\boldsymbol{\theta}_{\boldsymbol{S}}^{i}$，并更新教师模型： $\boldsymbol{\theta}_{\boldsymbol{T}}^{i} \leftarrow \boldsymbol{\theta}_{\boldsymbol{T}}^{i-1} - \eta_T \cdot \nabla_{\boldsymbol{\theta}_{\boldsymbol{T}}^{i-1}} \{\lambda \mathcal{L}_f(\mathcal{D}_l;\ \boldsymbol{\theta}_{\boldsymbol{S}}^{i}) + \mathcal{L}_l(\mathcal{D}_l;\ \boldsymbol{\theta}_{\boldsymbol{T}}^{i-1})\}$。
- 修复 $\boldsymbol{\theta}_{\boldsymbol{S}}^{i}$，并更新 $\boldsymbol{\theta}_{\boldsymbol{u}}^{i}$： $\boldsymbol{\theta}_{\boldsymbol{u}}^{i} \leftarrow \boldsymbol{\theta}_{\boldsymbol{u}}^{i-1} - \eta_u \nabla_{\boldsymbol{\theta}_{\boldsymbol{u}}^{i-1}} \mathcal{L}_f(\mathcal{D}_l;\ \boldsymbol{\theta}_{\boldsymbol{S}}^{i})$。

C. 验证数据的动态选择

它试图将全部可用的带标签数据 $\mathcal{D}_l$ 用作验证集 $\mathcal{D}_v$ 来评估学生模型，正如（Pham等人，2021年）针对图像分类所提出的方法。然而，在贝叶斯优化的设定下，这并不总是最优的，因为贝叶斯优化的目标是在带标签数据总量较少的情况下找到全局最优解。因此，对学生模型的评估应以能够提高师生模型预测全局最优解准确性的方式进行，因为该评估会向教师模型提供反馈。由于大多数带标签数据 $\mathcal{D}_l$ 用于训练教师模型，因此 $\mathcal{D}_l$ 附近的伪标签质量较高。因此，使用 $\mathcal{D}_l$ 来验证学生模型可能会得到较低的平均损失，但这并不一定能反映模型预测全局最优解的能力。我们的实证研究表明，当验证数据选择目标值较高的数据时，TSBO 的性能会得到提升。从这个意义上说，在目标值更接近全局最优解的区域评估学生模型，能够为教师模型提供最佳反馈，从而提高其在最需要的地方的准确性。我们采用一种实用方法，在每次贝叶斯优化迭代中动态选择$\mathcal{D}_v$：从 $\mathcal{D}_l$ 中选取具有 K 最高标签值的子集。为此，我们应用快速排序算法对所有带标签数据进行排序。附录 F 展示了所提出选择方法对于 $\mathcal{D}_v$的有效性。

D. 实验细节

D.1. 高维优化任务

算术表达式重构任务：目标是发现一个单变量算术表达式 x* = 1 / 3 + v + sin(v * v)。对于输入表达式 x，目标函数是一个距离度量f(x) $=\max\{-7, -\log(1+$均方误差(x(v) $-$ x*(v))$\}$，其中 v 是 $[-10, 10]$ 中均匀分布的 1,000 个数值。采用潜在空间维度为 25 的语法变分自编码器（Kusner 等人，2017）。该模型在包含 40,000 个表达式的数据集上进行预训练（Kusner 等人，2017）。

化学设计任务：该任务旨在设计一种具有所需分子性质特征的分子。所考虑的目标特征包括：1）惩罚性水辛醇分配系数（惩罚对数P）（G´omez-Bombarelli 等人，2018年），以及 2）雷诺嗪多属性目标（雷诺嗪多参数优化）（Brown 等人，2019年）。采用一种连接树变分自编码器（Jin 等人，2018年），其潜在空间维度为 56，并在 ZINC250k 数据集（Sterling 与 Irwin，2015年）上进行了预训练。

对于每个任务，在开始优化之前，先通过最大化证据下界对变分自编码器进行预训练，使用无标签数据（Kingma 与 Welling，2013年），并且所有方法都在优化开始时使用这个预训练的变分自编码器。

D.2. TSBO 的模型架构和超参数

在TSBO中，教师模型是一个具有 5 个隐藏层并采用 ReLU 激活函数的多层感知机（MLP）（Nair 与 Hinton，2010年）。输出维度为 2。学生模型是采用径向基函数核的标准高斯过程。

为确保可复现性，我们全面介绍了所有使用TSBO进行的实验所采用的超参数。我们的方法基于 T-LBO，因此采用了（Grosnit 等人，2021年）建议的变分自编码器默认超参数。专用于TSBO的其余超参数列于表6。

D.3. TSBO 中变分自编码器的训练

尽管TSBO 作为一种通用的贝叶斯优化框架，已被无缝集成到 T-LBO（Grosnit 等人，2021年）这一最先进的基于变分自编码器的贝叶斯优化方法中，以便进行公平比较。变分自编码器的训练方法保持不变，与 T-LBO 的方法保持一致：

- 预训练：在首次贝叶斯优化迭代之前，以无监督方式在数据集上训练变分自编码器，以最大化证据下界（Kingma 与 Welling，2013年）；
- 微调：每进行 50 次贝叶斯优化迭代后，都会使用所有带标签数据训练变分自编码器 1 个周期，以最大化证据下界并

组	超参数	表达式	化学设计
通用	每次贝叶斯优化迭代的训练步数	20	20
	预热步数	2,000	2,000
	反馈权重	10^{-1}	10^{-1}
	验证数据数量	10	30
	采样的未标注数据数量	100	300
	采集函数	EI	EI
	采集优化器	L-BFGS	L-BFGS
教师模型	学习率	10^{-3}	10^{-4}
	标注数据批量大小	256	32
	优化器	Adam	Adam
学生模型	核函数	RBF	RBF
	先验均值	常数	常数
	学习率	10^{-2}	10^{-2}
	优化器	Adam	Adam
数据查询高斯过程	核函数	RBF	RBF
	先验均值	常数	常数

表6：超参数

表7：反馈损失权重的消融测试

方法	λ	表达式 (↓)	惩罚对数P (↑)	雷诺嗪多参数优化（↑）
T-LBO	-	0.572±0.268	5.695±1.254	0.620±0.043
TSBO-高斯	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

最小化三元组损失，该损失会惩罚相似标签的数据彼此距离较远的情况。表达式任务中的三元组损失权重设为 10，化学设计任务中的权重设为 1。

TSBO 中提出的所有模型与变分自编码器的训练方案彼此解耦，因此 T-LBO 是验证 TSBO 样本效能的合适基线。

E. TSBO 中反馈权重的影响

我们分析了反馈权重 λ 的选择所产生的影响。实验表明，在 λ 的较大范围内，TSBO 的表现始终优于基线 T-LBO，这凸显了我们的方法对该超参数的鲁棒性。

如表 7所示，尽管 λ 在 $\{0.001.0.01, 0.1, 1\}$ 中的选择会影响 TSBO的性能，但对于每个考虑的 λ，TSBO-Gaussian 的表现始终优于 T-LBO，这表明我们的成功并非源于刻意进行λ 选择。

F. 验证数据选择的额外消融测试

为验证 TSBO 中所提出的动态验证数据选择 $\mathcal{D}_v$ 的有效性（其中 $\mathcal{D}_v$ 是 $\mathcal{D}_l$ 中标签值最高的 K 个样本的子集），我们开展了一项消融研究，将其（TSBO-Gaussian）与两种未经优化的$\mathcal{D}_v$ 选择策略进行比较：从 $\mathcal{D}_l$ 中均匀随机抽取 K 个样本（TSBO-Gaussian-ValRand），以及使用当前带标签数据集 $\mathcal{D}_l$（TSBO-Gaussian-ValAll）。我们在化学设计任务上评估了这三种 TSBO 变体，并报告了 5 次运行的最佳值均值；每次运行均以 200 个初始带标签数据和 250 的新增数据查询预算开始。 K 设为 30。

如表8所示，尽管TSBO 的每个变体都优于基线 T-LBO，但 TSBO-Gaussian-ValRand 和 TSBO-Gaussian-ValAll 在寻找全局最大值方面都不如 TSBO-Gaussian 有竞争力。这一结果符合我们的

预期：学生模型对具有 K 最高标签值的那些样本的反馈，更有利于教师-学生模型的训练，最终也更有助于搜索最大值。

表8：化学设计任务中 TSBO 验证数据选择的比较

方法	验证数据选择	惩罚对数P（↑）
T-LBO(Grosnit 等人，2021年)	-	5.70
TSBO-Gaussian-ValRand	随机30	21.60
TSBO-Gaussian-ValAll	所有带标签数据	22.65
TSBO-高斯	前 30 名	**25.02**

G. TSBO 对噪声标签的鲁棒性

本节展示了TSBO在噪声环境中的鲁棒性。我们在化学设计任务上将TSBO 与其他基线方法进行比较，其中所有标签都受到额外的独立同分布零均值高斯噪声影响，标准差（std）为 0.1。所有方法均以 200 个初始带标签数据开始，并查询 250 个新样本。我们在表 9中报告了各方法在 5 次运行中找到的最佳值的均值和标准差。

表 9：化学设计任务中带有白高斯噪声的标签

方法	最佳惩罚对数P（↑）	
	噪声方差=0	噪声方差=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-高斯	**25.02±4.794**	**25.97±4.82**

与未加入观测噪声时的结果相比，几乎所有基线方法，尤其是 T-LBO，在噪声场景下的表现都有所下降，而 TSBO 的性能却没有退化。这一现象证明了所提出的不确定性感知教师-学生模型具有抗噪能力。

H. 更广泛影响

所提出的 TSBO 有望在各个领域产生显著的积极影响。TSBO 在多个数据集上相较于基线方法能够有效地找到最优解，为如何在带标签数据和评估预算有限的情况下提高优化过程的效率和效能提供了有前景的解决方案。例如，在工程和制造业中，TSBO 可助力异常值检测、故障分析以及更高效流程的设计，从而提高生产力并减少资源消耗。通过实现更快速、更准确的优化，TSBO 最终能够造福整个社会。

尽管 TSBO 前景广阔，但仍需承认并减轻其可能带来的负面影响。一个值得关注的问题是过度依赖自动化优化算法，这可能会导致人们对人类直觉和创造力的重视程度下降。TSBO 应作为辅助工具，用于增强而非完全取代人类决策。此外，如果教师模型所用的训练数据存在固有偏差，优化过程就可能产生偏见。必须谨慎处理训练数据，以确保优化结果公平且无偏。

总之，TSBO 在优化任务中具有广泛影响的巨大潜力。通过提高优化过程的效率和效能，TSBO 可以加速发现最优解，使各行各业受益，并最终改善个人和整个社会的福祉。然而，必须考虑并减轻潜在的负面影响，例如过度依赖自动化和偏见风险，以确保负责任且合乎伦理地使用 TSBO。通过采取适当的保障措施并进行充分考量，TSBO 可以成为增强人类专业能力、推动不同领域优化发展的宝贵工具。