arXiv:2302.04040v2 [cs.LG] 2 Nov 2023

基于 GFlowNets 的样本高效多目标分子优化

Yiheng Zhu[1]*, **JialuWu**[2]*, **胡超文**[3], **严佳欢**[1],**Chang-YuHsieh**[2]†, **侯廷军**[2]†, **Jian Wu**[4,5,6]†[1]浙江大学计算机科学与技术学院[2]浙江大学药学院[3]浙江大学工程学院[4]浙江大学医学院附属第二医院[5]浙江大学公共卫生学院[6]浙江大学温州研究院 {zhuyiheng2020, jialuwu, ChaowenHu, jyansir, kimhsieh, tingjunhou, wujian2000}@zju.edu.cn

摘要

许多重要的科学问题都涉及设计具有理想性质的新分子，这可以表述为在离散化学空间中进行黑盒优化问题。在实际应用中，多个相互冲突的目标以及高成本评估（例如湿实验）使候选项的多样性至关重要。计算方法已取得初步成功，但在兼顾目标多样性和搜索空间多样性方面仍面临困难。为弥补这一不足，我们提出一种多目标贝叶斯优化（MOBO）算法，利用基于超网络的生成流网络（HN-GFN）作为采集函数优化器，旨在从近似帕累托前沿采样一批多样化的候选分子图。HN-GFN 使用单个偏好条件超网络，学习探索目标之间的各种权衡。我们进一步提出一种类事后离策略策略，在不同偏好之间共享高性能分子，以加快 HN-GFN 的学习。实验结果表明，HN-GFN 具备在不同偏好间泛化的充分能力。此外，在各种真实世界多目标贝叶斯优化场景中的实验表明，与现有方法相比，我们的框架在候选分子质量和样本效率方面总体表现更优。代码见 https://github.com/violet-sto/HN-GFN。

1 引言

设计具有所需性质的新型分子结构，也称为分子优化，是一项至关重要的任务，在从药物发现到材料设计等科学领域具有巨大的应用潜力。分子优化可自然地表述为针对离散化学空间的黑盒优化问题，而该空间的组合规模极其庞大 [52]。近年来，人们开始利用深度生成模型 [31]和组合优化算法 [62, 30], 等计算方法来促进优化。然而，大多数现有方法在现实场景中的适用性受到两项实际限制的制约：(i) 化学家通常希望同时优化多个目标性质 [60, 33]。例如，理想药物除了能有效抑制疾病相关靶点外，还应易于合成且无毒性。遗憾的是，由于各项目标往往相互冲突，在大多数情况下，并不存在单一的

*共同贡献。
†通讯作者。

第37届神经信息处理系统大会（NeurIPS 2023）。

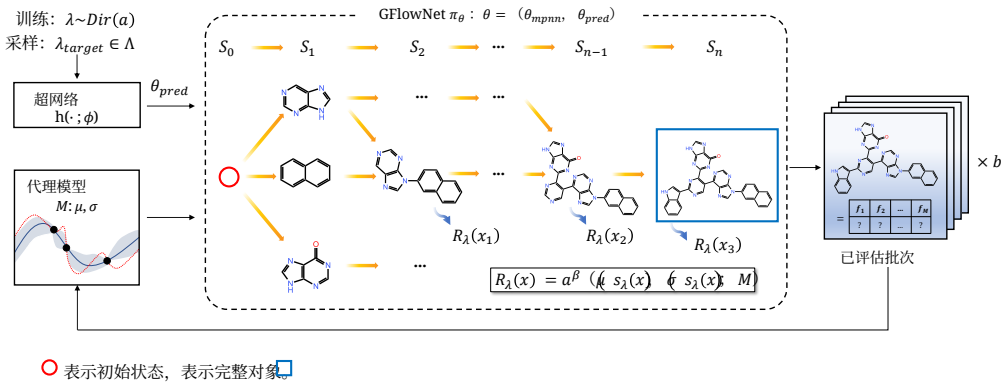


图1：使用代理模型 $\mathcal{M}$ 进行不确定性估计，并使用 HN-GFN 进行采集函数优化的分子优化多目标贝叶斯优化循环。在每一轮中，使用奖励函数 R_λ训练策略 π_θ，其中每次迭代均从 Dir(α) 中采样 λ。根据目标偏好向量 $\lambda_{\text{target}} \in \Lambda$，从近似帕累托前沿中采样一批新的候选项。

最优解，而是存在一组具有不同权衡的帕累托最优解 [18, 45]。(ii) 现实中的预言机（例如湿实验和高保真模拟）需要耗费大量成本来合成和评估分子 [22]。因此，预言机评估的次数通常极为有限。在此类场景中，候选项的多样性是一个关键考量因素。

贝叶斯优化（BO） [34, 54] 为全局优化昂贵的黑盒函数提供了一种高样本效率的框架。其核心思想是在观测数据集上构建一个易于评估的代理模型，以近似真实的目标函数（也称为预言机），并优化一个采集函数 （基于代理模型构建），从而获得效用较高且对下一轮评估具有信息价值的候选项。鉴于生化实验通常采用大批量、少轮次的设置 [3], ，因此优先采用批量贝叶斯优化，以缩短整个优化周期 [25]。多目标贝叶斯优化也受到了广泛关注，并通过有效优化可微采集函数 [12, 13]，在连续问题中取得了颇具前景的性能。然而，在离散问题中，多目标贝叶斯优化的应用相对较少，因为没有梯度可用于引导搜索空间中的搜索。

尽管可以采用现有的离散分子优化方法作为采集函数优化器来缓解这一问题，但它们很难同时考虑搜索空间和目标空间中的多样性：1）许多方法忽视了所提出候选项在搜索空间中的多样性 [30, 33]；2）许多多目标方法仅依赖预定义的标量化函数（例如均值），将多目标优化（MOO）问题转化为单目标问题 [61, 20]。由于代理模型无法准确复现预言机的全部行为，并且在优化之前（即使具备领域知识）也不清楚最优权衡，因此不仅希望提出能够为搜索空间带来额外信息的候选项，还希望探索更多潜在的、值得关注的权衡。为实现这一目标，我们探索如何扩展近期提出的生成流网络（GFlowNets） [6, 7], ——这是一类生成模型，旨在学习一种随机策略，以与奖励函数成比例的概率按序构建对象——从而促进多目标分子优化。与传统的组合优化方法相比，GFlowNets 擅长生成多样且高奖励的对象，这一点已在单目标问题中得到验证 [6, 28]。

在本研究中，我们提出了一种基于 GFlowNets 的多目标贝叶斯优化算法，用于实现样本高效的多目标分子优化。我们提出将基于超网络的生成流网络（HN-GFN）用作多目标贝叶斯优化中的采集函数优化器，从近似帕累托前沿中采样出多样化的候选批次。不同于以往工作中通常定义固定奖励函数的做法 [6], ，我们在奖励函数的分布上训练一个统一的 GFlowNet（由随机标量化 构成，并由偏好向量参数化），同时使用单个偏好条件超网络来控制策略。在采样候选项时，HN-GFN 通过改变输入偏好向量，灵活探索相互竞争的目标之间的各种权衡。受事后经验回放的启发 [2], ，我们进一步提出一种类事后离策略策略，在不同偏好之间共享高性能分子

并加快 HN-GFN 的学习。如我们报告的实验所示，我们首先通过实证验证 HN-GFN 能够在偏好向量之间实现泛化，随后将所提出的框架应用于现实场景，并证明其相较于现有方法的有效性和效率。我们的主要贡献总结如下：

• 我们提出一种基于 GFlowNet 的多目标贝叶斯优化算法，以促进现实场景中的分子优化。我们提出 HN-GFN，这是 GFlowNet 的一种条件变体，能够高效地从近似帕累托前沿中采样候选项。

• 我们提出一种类事后离策略策略，以加快 HN-GFN 的学习。

• 实验验证，与最先进的基线方法相比，我们基于 HN-GFN 的多目标贝叶斯优化算法能够更高效地找到高质量的帕累托前沿。

2 相关工作

分子优化。 分子优化已采用多种计算方法，主要可分为三类：1）**潜空间优化（LSO）**方法 [24, 31, 58, 66] 在变分自编码器（ VAE ）等生成模型学习得到的低维连续潜空间中进行优化 [36]。这些方法要求潜在表示具有判别性，但生成模型的训练与优化目标相互分离，这给优化带来了挑战 [58]。与在潜空间中搜索不同，组合优化方法在显式离散空间中直接搜索目标分子结构，并由 2）**进化算法** [30]和 3）**强化学习（RL）** [62] 引导搜索。然而，大多数以往方法仅关注单一性质，从药物相似性（QED） [8] 和合成可及性（SA） [19], 等非生物学性质，到衡量与蛋白质靶点结合能的生物学性质 [6]。近年来，多目标分子优化受到了广泛关注 [33, 61, 20]。例如，MARS [61] 采用马尔可夫链蒙特卡罗（MCMC）采样，寻找满足多种性质的新分子。然而，大多数方法都需要数量极其庞大的预言机调用，才能对分子进行即时评估 [33, 61]。相比之下，我们以样本高效的方式解决这一问题。

GFlowNet。 GFlowNets [6, 7] 旨在按奖励函数的比例采样复合对象，而非像强化学习（RL）中通常那样最大化奖励 [56]。由于目标相同，GFlowNets 与 MCMC方法相关；与此同时，它们通过训练生成模型，将采样（在不同模式间混合）的高昂成本摊销开来。GFlowNets 在多种应用中取得了显著进展，例如生物序列设计 [28], 、离散概率建模 [63], 以及贝叶斯结构学习 [15]。尽管条件GFlowNet的概念最初由 Bengio 等人讨论 [7], ，我们是首个针对多目标优化（MOO）研究并实例化这一概念的工作，同时 Jain 等人也开展了相关研究 [29]。与他们相比，我们精心设计了条件机制，并提出一种事后式离策略策略；在多目标优化领域（无论是强化学习还是 GFlowNet 文献中），这种策略都鲜有研究。附录 D 提供了详细比较。

离散空间中的贝叶斯优化。 近年来，贝叶斯优化在离散域中的应用日益广泛。与连续空间相比，在离散空间中构建代理模型和优化采集函数要困难得多。一种常见方法是借助生成模型将离散空间转换为连续空间 [24, 16, 44]。此外，也可以直接利用离散核定义高斯过程（GP） [46, 53] ，并通过进化算法求解采集函数优化问题 [35, 57]。

多目标贝叶斯优化。 贝叶斯优化已广泛应用于多目标优化问题，以高效优化多个相互竞争的昂贵黑盒函数。最受欢迎的方法主要基于超体积改进 [12], 随机标量化 [37, 49], 和熵搜索 [27, 5]。尽管已有若干方法考虑了并行评估 [11, 38] 和多样性 [38], ，但它们仅适用于连续域。

3 背景

3.1 问题表述

我们要解决的问题是在离散化学空间 $\mathcal{X}$中搜索能够最大化向量值目标函数 $f(x) = \big(f_1(x)\ f_2(x)\ldots, f_M(x)\big)$：$\mathcal{X} \to \mathbb{R}^M$ 的分子图 $x \in \mathcal{X}$，其中 f_m 是一个黑盒函数（也称为预言机），用于评估分子的某种性质。实际上，现实中的预言机评估成本极高，无论是通过高保真模拟还是湿实验进行评估都是如此。因此，我们提出尽可能减少预言机评估次数来进行优化，因为在这种情境下，样本效率至关重要。

由于不同目标之间可能相互矛盾，多目标优化问题通常不存在单一的最优解。因此，目标是恢复帕累托前沿——即一组帕累托最优解：在不使其他目标变差的情况下，无法进一步改善其中任何一个目标 [18, 45]。在最大化情形下，若解 $f(x)$ 满足帕累托支配另一个解 $f(x')$，当且仅当 $f_m(x) \geq f_m(x')\ \forall m = 1, \ldots\ldots, M$并且 $\exists m'$使得 $f_{m'}(x) > f_{m'}(x')$，则称解 $f(x) \succ f(x')$。若解 $f(x^*)$ 不被任何解帕累托支配，则称其为帕累托最优。帕累托前沿可表示为 $\mathcal{P}^* = \{f(x^*)\colon \{f(x)\colon f(x) \succ f(x^*)\} = \emptyset\}$。

有限近似帕累托前沿 $\mathcal{P}$ 的质量通常用超体积 [67] 来衡量——即由 $\mathcal{P}$ 支配且以给定参考点 $\boldsymbol{r} \in \mathbb{R}^M$为下界的空间的 M 维勒贝格测度 λ_M：$HV(\mathcal{P}, \boldsymbol{r}) = \lambda_M(\cup_{i=1}^{|\mathcal{P}|}[\boldsymbol{r}, y_i])$，其中 $[\boldsymbol{r}, y_i]$表示由 $\boldsymbol{r}$ 和 $y_i = f$界定的超矩形（x_i）。

3.2 批量贝叶斯优化

贝叶斯优化（BO） [54] 提供了一种基于模型的迭代框架，可高效地进行黑盒优化。给定观测数据集 $\mathcal{D}$，贝叶斯优化依靠贝叶斯代理模型$\mathcal{M}$ 来估计真实预言机评估的后验分布。借助代理模型，可构建采集函数 a：$\mathcal{X} \to \mathbb{R}$，为候选对象分配效用值，以决定接下来在预言机上评估哪些对象。与成本高昂的预言机相比，评估成本低廉的采集函数可以得到高效优化。我们考虑这样的情形：预言机的评估预算为 N 轮，每轮批次大小固定为 b。

具体而言，我们可以访问一个随机初始数据集 $\mathcal{D}_0 = \{(x_i^0, y_i^0)\}_{i=1}^n$，其中 $y_i^0 = f(x_i^0)$ 是真实的预言机评估。在每一轮 $i \in \{1, \ldots, N\}$ 中，通过最大化采集函数得到一个候选批次 $\mathcal{B}_i = \{x_j^i\}_{j=1}^b$，并在预言机 $y_j^i = f(x_j^i)$ 上并行评估。随后扩充观测数据集 $\mathcal{D}_{i-1}$，以供下一轮使用：$\mathcal{D}_i = \mathcal{D}_{i-1} \cup \{(x_j^i, y_j^i)\}_{j=1}^b$。

4 方法

本节介绍我们提出的基于超网络的生成流网络（HN-GFN）多目标贝叶斯优化算法，如图1所示。由于篇幅有限，详细算法见附录 A。我们的核心思路是将 GFlowNets 扩展为多目标贝叶斯优化的采集函数优化器，旨在从近似帕累托前沿中采样出具有多样性的候选批次。首先，我们介绍 GFlowNets 在分子设计中的应用，然后说明如何通过偏好条件超网络引导 GFlowNet，根据目标之间的各种权衡采样分子。接着，我们提出一种类事后离策略策略，以加快 HN-GFN 的学习。最后，我们介绍代理模型和采集函数。

4.1 预备知识

GFlowNets [6] 旨在学习一种随机策略 π，以顺序构造离散对象$x \in \mathcal{X}$，使其概率为 $\pi(x) \propto R(x)$，其中 $\mathcal{X}$是一个组合空间，且 R：$\mathcal{X} \to \mathbb{R}_{\geq 0}$ 是一个非负奖励函数。对象 $x \in \mathcal{X}$的生成过程可以表示为一系列离散的动作 $a \in \mathcal{A}$，这些动作逐步修改部分构造对象，该对象记作状态 $s \in \mathcal{S}$。设生成过程从一个特殊的初始状态 s_0 开始，并以表示对象已完成的特殊动作（$s = x \in \mathcal{X}$）结束，则对象x 的构造可以定义为一条完整轨迹 $\tau = (s_0 \to s_1 \to \cdots \to s_n = x)$。

按照基于片段的分子设计 [6, 61], , 分子图通过依次将一个片段连接到部分构造分子的一个原子上来生成，该片段从预定义的构件词汇表中选取。最大轨迹长度为 8，动作数量则根据状态在 100 到 2000 之间变化，因此 $|\mathcal{X}|$ 最多可达 10^{16}。多个动作序列可以导向同一状态，且不存在移除片段的动作，因此可能的动作序列空间可表示为一个有向无环图（DAG）$\mathcal{G} = (\mathcal{S}, \mathcal{E})$，其中 $\mathcal{E}$ 中的边是从一个状态到另一个状态的转移 $s \to s'$。为了学习上述期望的策略，Bengio 等人 [6] 提出将 DAG 结构视为一个流网络。

马尔可夫流。Bengio 等人 [7] 首先在所有完整轨迹 $\mathcal{T}$ 的集合上定义轨迹流 $F: \mathcal{T} \to \mathbb{R}_{\geq 0}$，以衡量未归一化密度。随后可分别将边流 和状态流 定义为 $F(s \to s') = \sum_{s \to s' \in \tau} F(\tau)$ 以及 $F(s) = \sum_{s \in \tau} F(\tau)$。轨迹流 F 确定了一个概率测度 $P(\tau) = \frac{F(\tau)}{\sum_{\tau \in \mathcal{T}} F(\tau)}$。如果流 F 是马尔可夫的，则可将前向转移概率 P_F 计算为 $P_F(s'|s) = \frac{F(s \to s')}{F(s)}$。

流匹配。如果满足以下流一致性方程 $\forall s \in \mathcal{S}$，则流是一致的：

$$F(s) = \sum_{s' \in Pa_{\mathcal{G}}(s)} F(s' \to s) = R(s) + \sum_{s'' : s \in Pa_{\mathcal{G}}(s'')} F(s \to s'') \tag{1}$$

其中，$Pa_{\mathcal{G}}(s)$ 是 $\mathcal{G}$ 中 s 的一组父节点。正如 Bengio 等人 [6], 所证明的，如果对于非终止状态 s 满足流一致性方程 $R(s) = 0$，并且对于终止状态 x 满足 $F(x) = R(x) \geq 0$，则由前向转移概率 $\pi(s'|s) = P_F(s'|s)$ 定义的策略 π 会以概率 $\pi(x) \propto R(x)$ 采样对象 x。GFlowNets 提议使用具有足够容量的神经网络 $F_\theta(s, s')$ 来近似边流 $F(s \to s')$，从而在收敛时满足流一致性方程。为此，Bengio 等人 [6] 定义了一种类似时序差分的 [56] 学习目标，称为流匹配（FM）：

$$\mathcal{L}_\theta(s, R) = \left(\log \frac{\sum_{s' \in Pa_{\mathcal{G}}(s)} F_\theta(s', s)}{R(s) + \sum_{s'' : s \in Pa_{\mathcal{G}}(s'')} F_\theta(s, s'')} \right)^2 \tag{2}$$

可以使用任何具有全支持的探索策略 $\tilde{\pi}$ 来采样训练轨迹，并通过最小化 FM 目标函数 [6] 来获得一致的流 $F_\theta(s, s')$。因此，基于该近似流定义的策略 $\pi_\theta(s'|s) = P_{F_\theta}(s'|s) = \frac{F_\theta(s \to s')}{F_\theta(s)}$ 也可以按与奖励 $R(x)$ 成正比的概率 π_θ 采样对象 x。在实践中，训练轨迹从当前策略 P_{F_θ} 与允许动作 [6] 上的均匀分布的混合分布中采样。本研究采用 FM 目标函数，因为我们也考察了另一种轨迹平衡 [43]，但其在早期实验中的表现较差。请注意，未来的工作可以采用子轨迹平衡 [42] 等更先进的目标函数。

4.2 基于超网络的生成流网络

我们提出的 HN-GFN 旨在通过统一模型，从近似帕累托前沿中采样一批多样化的候选项。多目标优化的一种常见方法是将其分解为一组带有标量化函数的标量优化问题，并应用标准的单目标优化方法逐步逼近帕累托前沿 [37, 64]。本文考虑加权和 (WS)：$s_\lambda(x) = \sum_i \lambda_i f^i(x)$ 和切比雪夫 [45]：$s_\lambda(x) = \max_i \lambda_i f^i(x)$，其中 $\lambda = (\lambda_i, \cdots, \lambda_M)$ 是偏好向量，用于定义相互竞争的各项性质之间的权衡。

为支持并行评估，可以根据不同的标量化方法同时获得多个候选项。实际上，对于离散问题，这种方法很难随着目标数量的增加而高效扩展。以 GFlowNet 为例，我们需要针对每种奖励函数 $R_\lambda(x) = s_\lambda(x)$ 分别训练多个 GFlowNet，以覆盖目标空间：

$$\theta^*_\lambda = \arg\min_\theta \mathbb{E}_{s \in \mathcal{S}} \mathcal{L}_\theta(s, R_\lambda) \tag{3}$$

我们的核心动机是设计一个统一的 GFlowNet，使其能够根据不同的奖励函数（包括训练期间未见过的奖励函数）采样候选项。我们不再用一个

固定的偏好向量 λ 来定义奖励函数，而是提出在奖励函数分布 R_λ 上训练一个由偏好条件化的 GFlowNet，其中 λ 从单纯形 $S_M = \{\lambda \in \mathbb{R}^m | \sum_i \lambda_i = 1, \lambda_i \geq 0\}$ 中采样：

$$\theta^* = \arg\min_\theta \mathbb{E}_{\lambda \in S_M} \mathbb{E}_{s \in \mathcal{S}} \mathcal{L}_\theta(s, R_\lambda) \tag{4}$$

备注。 假设模型容量无限大，很容易证明，所提出的优化方案（公式 (4)）与原始方案（公式 (3)）同样强大，因为两个损失函数的解一致 [17]。然而，无限模型容量这一假设极为严格，几乎不可能成立，因此如何在实践中设计条件机制至关重要。

4.2.1 基于超网络的条件机制

我们提出通过超网络根据偏好向量对 GFlowNets 进行条件化 [26]。超网络是根据输入生成目标网络权重的网络。在标准 GFlowNet 中，流预测器 F_θ 由作用于分子片段图的消息传递神经网络（MPNN） [23] 参数化，并通过两个预测头，分别基于节点表示和图表示近似$F(s,\ s')$和 $F(s)$。这两个预测头由多层感知器（MLPs）参数化。为简洁起见，我们将其记作 $\theta = (\theta_{\text{mpnn}},\ \theta_{\text{pred}})$。

可以将 HN-GFN 的训练视为：训练一个智能体，使其在同一环境（状态空间$\mathcal{S}$ 和动作空间 $\mathcal{A}$）中执行多个对应于不同目标（奖励函数 R_λ）的策略。因此，我们提出仅使用超网络对预测头 θ_{pred}的权重进行条件化，同时共享 MPNN θ_{mpnn}的权重，从而获得泛化能力更强的状态表示。更确切地说，超网络 $h(\cdot;\phi)$以偏好向量 λ 作为输入，并输出预测头的权重 $\theta_{\text{pred}} = h(\lambda;\phi)$。超网络的参数 ϕ 像普通参数一样进行优化。遵循 Navon 等人 [47], 的方法，我们使用具有多个头的多层感知机对 h 进行参数化，每个头都为目标网络的不同层生成权重。

4.2.2 作为采集函数优化器

训练。 在每次迭代中，我们首先从狄利克雷分布 Dir(α)中采样一个偏好向量 λ。然后使用奖励函数 $R_\lambda(x) = a(\mu(s_\lambda(x))\ \sigma(s_\lambda(x)); \mathcal{M})$训练 HN-GFN，其中 μ 和 σ 分别是由 $\mathcal{M}$估计的后验均值和标准差。原则上，我们会在每轮之后重新训练 HN-GFN。我们还尝试了仅初始化超网络的参数，发现这种做法能带来相近的性能，且效率更高。

采样。 在每一轮 i中，我们使用训练好的 HN-GFN 采样一批具有多样性的 b 候选项。设 Λ^i 为 k 目标偏好向量的集合 $\lambda^i_{\text{target}}$。对于每个 $\lambda^i_{\text{target}} \in \Lambda^i$，我们采样 $\frac{b}{k}$ 个分子，并并行地在预言机上对其进行评估。我们只需通过从 Dir(α) 中采样来构造 Λ^i，但值得注意的是，这个先验分布也可以根据所关注的权衡自适应地定义。随着目标数量增加，我们会选择更大的 k，以覆盖目标空间。

4.3 类似事后经验的离策略策略

借助条件机制，HN-GFN 可以学习一系列策略以实现各种目标，也就是说，可以将针对特定偏好向量采样高奖励分子视为一个独立目标。正如 Jain 等人 [28], 通过实验证实的那样，由于 FM 目标函数是离策略 且离线的，只要全支持假设成立，我们就可以使用离线轨迹训练目标策略，以实现更好的探索。我们的关键洞见是，每个策略都可以从其他相似策略的宝贵经验（高奖励分子）中学习。

受强化学习中的事后经验回放 [2] 启发，我们提出通过使用不同的偏好向量重新评估高性能分子，在各策略之间共享这些分子。由于可能的偏好向量有无穷多个，我们聚焦于目标偏好向量 Λ^i，它们基于采样候选项的目的而设定，并为每个 $\lambda^i_{\text{target}} \in \Lambda^i$ 构建一个回放缓冲区。在训练过程中采样若干轨迹后，我们将完整对象 x 及其对应的奖励 $R_{\lambda^i_{\text{target}}}(x)$存入回放缓冲区。

具体而言，在每次迭代中，我们首先从 Dir(α) 与定义在 Λ^i 上的均匀分布的混合分布中采样一个偏好向量：$(1-\gamma)$Dir(α) $+\ \gamma$均匀分布。若选择 Λ^i，我们就用对应回放缓冲区中奖励最高的分子离线轨迹构建训练批次的一半

。否则，我们改为加入来自当前观测数据集 $\mathcal{D}_i$ 的离线轨迹，以确保 HN-GFN 能够在观测到的帕累托集附近正确采样。我们的方法能够灵活权衡泛化与专门化。随着 γ 从 0 变化到 1，偏好向量的训练分布会从 Dir(α) 转变为 Λ^i。仅使用有限的目标偏好向量 Λ^i 训练 HN-GFN，可能导致泛化能力较差。在实践中，尽管我们只基于 Λ^i 采样候选项，但我们认为，保持泛化能力至关重要，这样才能利用训练好的 HN-GFN 自适应地探索各种偏好向量。详细算法见附录 A。

4.4 代理模型与采集函数

尽管高斯过程在连续空间中已得到广泛应用，但其计算效率会随着观测数量增加而显著降低，并且在离散空间中表现不佳 [57]。已有大量研究致力于高效训练非贝叶斯神经网络，以估计认知不确定性 [21, 39]。我们对广泛使用的方法进行了基准测试（见附录 C.3），并采用了一种灵活且有效的方法：证据深度学习 [1]。至于采集函数，我们使用置信上界 [55] 来纳入不确定性。具体而言，我们使用单个多任务消息传递神经网络对目标进行建模，并将采集函数应用于标量化结果。

5 实验

我们首先验证 HN-GFN 在单轮合成场景中对偏好向量具有足够的泛化能力。接下来，我们在多轮实际场景中评估基于 HN-GFN 的所提多目标贝叶斯优化算法的有效性；这类场景更贴近现实世界分子优化。此外，我们还开展了若干消融研究，以实证验证设计选择的合理性。

5.1 单轮合成场景

在这里，我们的目标是证明可以利用 HN-GFN 对具有偏好条件化属性分布的分子进行采样。HN-GFN 在 MOBO 之外作为独立的优化器使用，直接优化预言机评分的标量化结果。由于实际上无法根据需要多次调用预言机，我们将这一场景称为合成场景。为了更好地直观展示采样分子的属性分布随偏好权重变化的趋势，我们仅考虑两个目标：针对糖原合酶激酶3β（GNK3β）和 c-Jun氨基末端激酶3（JNK3） [40, 33]的抑制评分。

比较方法。 我们将 HN-GFN 与以下方法进行比较。偏好特定型 GFlowNets（**PS-GFN**）是一组原版无条件 GFlowNets，每个网络都针对特定的偏好向量分别进行训练。请注意，PS-GFN 被视为“黄金标准”，而非基线方法，因为它使用相同的偏好向量进行训练和评估。**Concat-GFN** 和 **FiLM-GFN**分别是基于拼接和 FiLM [51],的两种 GFlowNet 条件化变体。**MOEA/D** [64] 和 **NSGA-III** [14] 是两种同样融入偏好信息的多目标进化算法。我们在由 HierVAE [32], 学习得到的 32 维潜在空间上执行进化算法，其优化性能优于 JT-VAE [31]。

度量指标。 所有上述方法均在同一组由 5 个均匀间隔的偏好向量构成的集合上进行评估。对于每种基于 GFlowNet 的方法，我们针对每个偏好向量采样 1000 个分子作为解。我们使用以下度量指标比较上述方法：**超体积指标（HV）**衡量解的帕累托前沿所支配、且下界由偏好点（0, 0）限定的空间体积。**多样性（Div）**是 Morgan 指纹之间两两 Tanimoto 距离的平均值。**相关性（Cor）**是 GFlowNet 从外部测试集中采样到分子的概率与其各自奖励在对数域中的斯皮尔曼秩相关系数 [48]。更多详情请参见附录 B.1.2。简言之，HV 和 Div 衡量解的质量，而 Cor 衡量训练后的模型与给定偏好向量的契合程度。每项实验均使用 3 个随机种子重复进行。

实验结果。 如表 1 所示，HN-GFN 的表现优于基线方法，在所有度量指标上均取得了与 PS-GFN（黄金标准）相当的性能。与基于 GFlowNet 的

表 1：不同方法在合成场景中的评估。

方法	超体积（↑）	多样性（↑）	相关性（↑）
MOEA/D	0.182 ± 0.045	n/a	n/a
NSGA-III	0.364 ± 0.041	n/a	n/a
PS-GFN	0.545 ± 0.055	0.786 ± 0.013	0.653 ± 0.003
Concat-GFN	0.534 ± 0.069	0.786 ± 0.004	0.646 ± 0.008
FiLM-GFN	0.431 ± 0.045	0.795 ± 0.014	0.633 ± 0.009
HN-GFN	**0.550 ± 0.074**	**0.797 ± 0.015**	**0.666 ± 0.010**

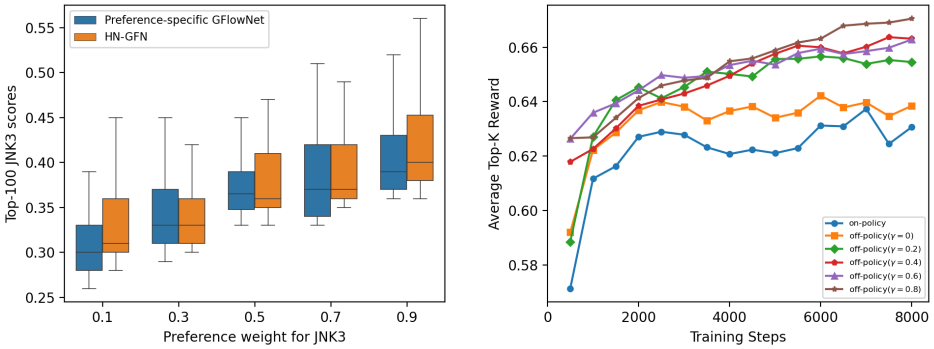


图2：**左**：JNK3 前100名得分的分布。**右**：HN-GFN 在优化 GSK3β 和 JNK3 的训练过程中，前20名奖励均值的变化。

方法相比，进化算法（MOEA/D 和 NSGA-III）未能找到高分分子，尤其是 MOEA/D。HN-GFN 的表现优于 Concat-GFN 和 FiLM-GFN，也是唯一能够达到 PS-GFN 性能的条件变体，这表明精心设计的基于超网络的条件机制具有优越性。在图2（左）中，我们展示了 HN-GFN 和 PS-GFN 在给定评估偏好向量集合的条件下采样得到的分子的经验性质（JNK3）分布。我们发现两种分布相似，且趋势一致：偏好权重越大，平均分数越高。相当的性能和一致的采样分布表明，HN-GFN 具备在偏好向量上进行泛化的充分能力。由于 PS-GFN 的运行时间和存储空间会随偏好向量数量线性增长，我们提出的统一 HN-GFN 为探索目标之间的各种权衡提供了一种高效得多的方法。

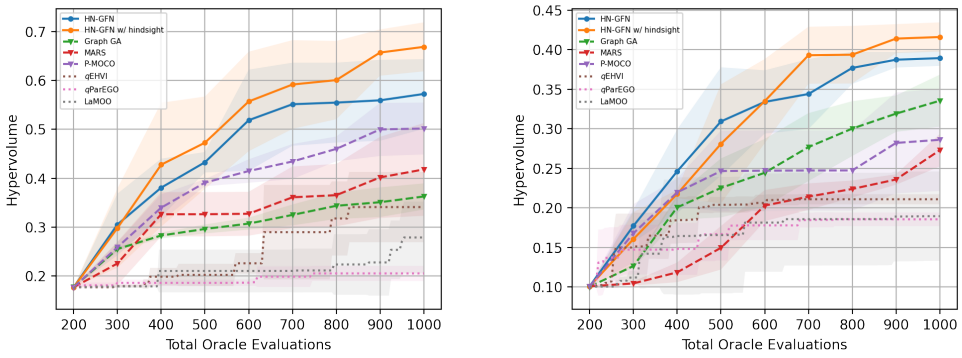
5.2 多目标贝叶斯优化

接下来，我们将在实际场景中评估 HN-GFN 作为多目标贝叶斯优化中采集函数优化器的有效性。我们考虑以下不同规模的目标组合：

- GNK3β+JNK3：联合抑制阿尔茨海默病相关靶点 GNK3β 和 JNK3。
- GNK3β+JNK3+定量估计药物相似性+合成可及性：联合抑制 GNK3β 和 JNK3，同时具有类药性且易于合成。

我们对 SA 评分（初始范围为 10 到 1）进行重新缩放，使上述所有性质的取值范围均为 [0,1]，且数值越高越好。对于这两种组合，我们考虑从 $|\mathcal{D}_0| = 200$ 个随机分子开始，并以批量大小 $b = 100$ 向预言机进一步查询 $N = 8$ 轮。每项实验均使用 3 个随机种子重复进行。

基线方法。 我们将 HN-GFN 与三种具有代表性的潜空间优化方法（q**ParEGO** [37]，q**期望超体积改进** [12]，以及 **LaMOO** [65]）进行比较，同时还与多种先进的组合优化



(a) GSK3β +JNK3

(b) GSK3β + JNK3 + 定量估计药物相似性 + 合成可及性

图 3：多目标贝叶斯优化循环中的优化性能（超体积）。

表 2：多目标贝叶斯优化场景中不同方法的多样性。

	多样性（↑）	
	GSK3β + JNK3	GSK3β +JNK3 + 定量估计药物相似性 +合成可及性
图遗传算法	0.347 ± 0.059	0.562 ± 0.031
MARS	0.653 ± 0.072	**0.754 ± 0.027**
P-MOCO	0.646 ± 0.008	0.350 ± 0.130
HN-GFN	**0.810 ± 0.003**	0.744 ± 0.008
带有事后信息的 HN-GFN	0.793 ± 0.007	0.738 ± 0.009

方法进行比较：**图遗传算法** [30] 是一种遗传算法，**MARS** [61]是一种马尔可夫链蒙特卡洛采样方法，而**P-MOCO** [41] 是一种多目标强化学习方法。更多详情请参见附录 B.2。

实验结果。图 3 展示了 HN-GFN 的性能（超体积）显著优于基线方法，尤其是在采用所提出的类事后离策略策略进行训练时。在两种目标组合下，HN-GFN 均大幅超越了表现最佳的基线方法 P-MOCO 和图遗传算法（分别为 0.67 对 0.50，以及 0.42 对 0.34）。此外，HN-GFN 的样本效率更高，仅用一半数量的预言机评估就达到了基线方法的性能。与潜空间优化方法相比，所有组合优化方法都带来了更大的性能提升，这表明直接在离散空间中进行优化颇具前景。表 2 报告了多样性（在 100 个候选项的候选批次中计算，并对各轮结果取平均值）。为公平比较，我们省略了潜空间优化方法，因为受内存限制，它们仅支持批量大小为 5 的 160 轮实验。与有时会提出相当相似候选项的图遗传算法和 P-MOCO 相比，HN-GFN 优异的优化性能可归因于其能够采样出多样化的候选批次。另一个有趣的观察是，在更具挑战性的设定（GNK3β+JNK3+QED+SA）中，MARS 通过马尔可夫链蒙特卡洛采样生成了多样化的候选项，却未能找到高质量的帕累托前沿，这表明 HN-GFN 比 MARS 更善于找到高奖励模式。计算成本详见附录 B.4。

5.3 消融实验

类事后离策略策略的效果。在 MOBO 的第一轮中，对于每个 $\lambda_{\text{target}} \in \Lambda$，我们每隔 500 个训练步数采样 100 个分子，并计算 Λ 上的 Top-20 奖励均值。图2（右）显示，类事后离策略策略显著提高了平均奖励，表明在各个策略之间共享高性能分子能够有效加快 HN-GFN 的训练。另一方面，进一步增加 γ 仅带来轻微提升。因此，为了在泛化与专门化之间取得理想的权衡，我们选择 $\gamma = 0.2$ 。

Dir() 的效果。这里我们考虑更具挑战性的 GNK3β+JNK3+QED+SA 组合，其中各项性质的优化难度差异很大。表3显示，偏向较难性质的分布能够带来更好的优化性能。在早期实验中，我们发现，如果 Λ^i 的分布固定，HN-GFN 对 α的变化相当稳健。

标量化函数的影响 s_λ。 在表 3 中，我们观察到 WS 得到的帕累托前沿优于切比雪夫。尽管遇到非凸帕累托前沿时，简单的 WS 通常并不适用 [10], ，但我们发现，使用 HN-GFN 进行优化时，WS 效果良好；HN-GFN 能够采样多样化的高奖励候选项，并可能到达帕累托前沿的非凸部分。此外，我们推测，切比雪夫所产生的非光滑奖励景观更难优化。

表 3：对 α 和 GNK3β+JNK3+定量估计药物相似性+合成可及性所做的消融研究。

	α			标量化函数	
	(1,1,1,1)	(3,3,1,1)	(3,4,2,1)	WS	切比雪夫
HV	0.312 ± 0.039	0.385 ± 0.018	**0.416 ± 0.023**	**0.416 ± 0.023**	0.304 ± 0.075
Div	**0.815 ± 0.015**	0.758 ± 0.018	0.738 ± 0.009	**0.738 ± 0.009**	0.732 ± 0.014

6 结论

我们提出了一种用于样本高效多目标分子优化的多目标贝叶斯优化算法。该算法利用基于超网络的生成流网络（HN-GFN），从近似帕累托前沿中采样一个具有多样性的候选批次。此外，我们还提出了一种类事后离策略策略，以提升优化性能。我们的算法在合成场景和实际场景中均优于现有方法。**未来工作**包括将该算法扩展到其他离散优化问题，例如生物序列设计和神经架构搜索。**所提出的**HN-GFN 的一个局限是，其计算成本高于其他免训练优化方法。然而，与真实应用中评估候选项所需的成本相比，模型训练所产生的成本微不足道。我们认为，候选项的质量更高远比降低成本重要。

致谢与经费披露

本研究部分得到中国国家重点研发计划（项目编号：2018AAA0102102）、国家自然科学基金委员会（项目编号：62176231、62106218、82202984、92259202、62132017 和 U22B2034）以及浙江省重点研发计划（项目编号：2023C03053）的资助。侯廷军的研究部分得到国家自然科学基金委员会（项目编号：22220102001）的资助。

参考文献

[1] Alexander Amini、Wilko Schwarting、Ava Soleimany 和 Daniela Rus。深度证据回归。神经信息处理系统进展，33:14927–14937，2020年。[2] Marcin Andrychowicz、Filip Wolski、Alex Ray、Jonas Schneider、Rachel Fong、Peter Welinder、Bob McGrew、Josh Tobin、OpenAI 的 Pieter Abbeel 和 Wojciech Zaremba。事后经验回放。神经信息处理系统进展，30，2017年。[3] Christof Angermueller、David Dohan、David Belanger、Ramya Deshpande、Kevin Murphy 和 Lucy Colwell。用于生物序列设计的基于模型的强化学习。载于国际学习表征会议，2020年。URL https://openreview. net/forum?id=HklxbgBKvr。[4] Maximilian Balandat、Brian Karrer、Daniel Jiang、Samuel Daulton、Ben Letham、Andrew G Wil- son 和 Eytan Bakshy。Botorch：用于高效蒙特卡洛贝叶斯优化的框架。神经信息处理系统进展，33:21524–21538，2020年。

[5] Syrine Belakaria、Aryan Deshwal 和 Janardhan Rao Doppa。用于多目标贝叶斯优化的最大值熵搜索。神经信息处理系统进展，32，2019年。[6] Emmanuel Bengio、Moksh Jain、Maksym Korablyov、Doina Precup 和 Yoshua Bengio。基于流网络的生成模型，用于非迭代多样化候选生成。神经信息处理系统进展，34:27381–27394，2021年。[7] Yoshua Bengio、Salem Lahlou、Tristan Deleu、Edward J Hu、Mo Tiwari 和 Emmanuel Bengio。Gflownet基础。机器学习研究期刊，24(210):1–55，2023年。[8] G Richard Bickerton、Gaia V Paolini、Jérémy Besnard、Sorel Muresan 和 Andrew L Hopkins。药物化学美感的量化。《自然·化学》，4(2):90–98，2012年。[9] Julian Blank 和 Kalyanmoy Deb。PyMOO：Python中的多目标优化。《*IEEE Access*》，8:89497–89509，2020年。[10] Stephen Boyd、Stephen P Boyd 和 Lieven Vandenberghe。凸优化。剑桥大学出版社，2004年。[11] Eric Bradford、Artur M Schweidtmann 和 Alexei Lapkin。采用高斯过程、谱采样和遗传算法的高效多目标优化。《全球优化期刊》，71(2):407–438，2018年。[12] Samuel Daulton、Maximilian Balandat 和 Eytan Bakshy。用于并行多目标贝叶斯优化的可微期望超体积改进。神经信息处理系统进展，33:9851–9864，2020年。[13] Samuel Daulton、Maximilian Balandat 和 Eytan Bakshy。使用期望超体积改进对多个含噪目标进行并行贝叶斯优化。神经信息处理系统进展，34:2187–2200，2021年。[14] Kalyanmoy Deb 和 Himanshu Jain。使用基于参考点的非支配排序方法的进化多目标优化算法，第一部分：求解具有盒约束的问题。*IEEE Transactions on Evolutionary Computation*，18(4):577–601，2013年。[15] Tristan Deleu、António Góis、Chris Emezue、Mansi Rankawat、Simon Lacoste-Julien、Stefan Bauer 和 Yoshua Bengio。使用生成流网络进行贝叶斯结构学习。*arXiv*预印本，2022年。[16] Aryan Deshwal 和 Jana Doppa。结合潜在空间与结构化核，对组合空间进行贝叶斯优化。神经信息处理系统进展，34:8185–8200，2021年。[17] Alexey Dosovitskiy 和 Josip Djolonga。只需训练一次：深度网络的损失条件训练。载于国际学习表征会议，2019年。[18] Matthias Ehrgott。多准则优化，第491卷。施普林格科学与商业媒体，2005年。[19] Peter Ertl 和 Ansgar Schuffenhauer。基于分子复杂度和片段贡献估算类药分子的合成可及性评分。*Journal of cheminfor-matics*，1(1):1–11，2009年。[20] Tianfan Fu、Wenhao Gao、Cao Xiao、Jacob Yasonik、Connor W Coley 和 Jimeng Sun。用于分子优化的可微支架树。载于国际学习表征会议，2022年。[21] Yarin Gal 和 Zoubin Ghahramani。随机失活作为贝叶斯近似：在深度学习中表征模型不确定性。载于国际机器学习会议，第1050–1059页。PMLR，2016年。[22] Wenhao Gao、Tianfan Fu、Jimeng Sun 和 Connor Coley。样本效率至关重要：实用分子优化的基准测试。神经信息处理系统进展，35:21342–21357，2022年。

[23] Justin Gilmer、Samuel S Schoenholz、Patrick F Riley、Oriol Vinyals 和 George E Dahl。用于量子化学的神经消息传递。载于国际机器学习会议，第1263–1272页。PMLR，2017年。[24] Rafael Gómez-Bombarelli、Jennifer N Wei、David Duvenaud、José Miguel Hernández-Lobato、Benjamín Sánchez-Lengeling、Dennis Sheberla、Jorge Aguilera-Iparraguirre、Timothy D Hirzel、Ryan P Adams 和 Alán Aspuru-Guzik。利用分子数据驱动的连续表示实现自动化学设计。*ACS Central Science*，4(2):268–276，2018年。[25] Javier González、Zhenwen Dai、Philipp Hennig 和 Neil Lawrence。通过局部惩罚实现批量贝叶斯优化。载于人工智能与统计学，第648–657页。PMLR，2016年。[26] David Ha、Andrew Dai 和 Quoc V Le。超网络。*arXiv*预印本 *arXiv:1609.09106*，2016年。[27] Daniel Hernández-Lobato、Jose Hernandez-Lobato、Amar Shah 和 Ryan Adams。用于多目标贝叶斯优化的预测熵搜索。载于国际机器学习会议，第1492–1501页。PMLR，2016年。[28] Moksh Jain、Emmanuel Bengio、Alex Hernandez-Garcia、Jarrid Rector-Brooks、Bonaventure FP Dossou、Chanakya Ajit Ekbote、Jie Fu、Tianyu Zhang、Michael Kilgour、Dinghuai Zhang 等。利用 GFlowNets 进行生物序列设计。载于国际机器学习会议，第9786–9801页。PMLR，2022年。[29] Moksh Jain、Sharath Chandra Raparthy、Alex Hernandez-Garcia、Jarrid Rector-Brooks、Yoshua Bengio、Santiago Miret 和 Emmanuel Bengio。多目标 GFlowNets。*arXiv*预印本 *arXiv:2210.12765*，2022年。[30] Jan H Jensen。用于探索化学空间的基于图的遗传算法、生成模型与蒙特卡洛树搜索。《化学科学》，10(12):3567–3572，2019年。[31] Wengong Jin、Regina Barzilay 和 Tommi Jaakkola。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。[32] Wengong Jin、Regina Barzilay 和 Tommi Jaakkola。利用结构基元层次化生成分子图。载于国际机器学习会议，第4839– 4848页。PMLR，2020年。[33] Wengong Jin、Regina Barzilay 和 Tommi Jaakkola。利用可解释子结构进行多目标分子生成。载于国际机器学习会议，第4849–4859页。PMLR，2020年。[34] Donald R Jones、Matthias Schonlau 和 William J Welch。昂贵黑盒函数的高效全局优化。《全球优化期刊》，13(4):455–492，1998年。[35] Kirthevasan Kandasamy、Willie Neiswanger、Jeff Schneider、Barnabas Poczos 和 Eric P Xing。结合贝叶斯优化与最优传输的神经架构搜索。神经信息处理系统进展，31，2018年。[36] Diederik P Kingma 和 Max Welling。自编码变分贝叶斯。*arXiv*预印本 *arXiv:1312.6114*，2013年。[37] Joshua Knowles。ParEGO：一种结合在线景观近似的混合算法，用于昂贵的多目标优化问题。*IEEETransactionsonEvolutionary Computation*，10(1):50–66，2006年。[38] Mina Konakovic Lukovic、Yunsheng Tian 和 Wojciech Matusik。基于多样性引导并采用批量评估的多目标贝叶斯优化。神经信息处理系统进展，33:17708–17720，2020年。[39] Balaji Lakshminarayanan、Alexander Pritzel 和 Charles Blundell。利用深度集成进行简单且可扩展的预测不确定性估计。神经信息处理系统进展，30，2017年。

[40] Yibo Li、Liangren Zhang 和 Zhenming Liu。利用条件图生成模型进行多目标从头药物设计。《化学信息学期刊》，10(1):1–24，2018年。[41] Xi Lin、Zhiyuan Yang 和 Qingfu Zhang。面向神经多目标组合优化的帕累托集学习。载于国际学习表征会议，2021年。[42] Kanika Madan、Jarrid Rector-Brooks、Maksym Korablyov、Emmanuel Bengio、Moksh Jain、Andrei Cristian Nica、Tom Bosc、Yoshua Bengio 和 Nikolay Malkin。通过部分轨迹学习 GFlowNets，以提升收敛性和稳定性。载于国际机器学习会议，第23467–23483页。PMLR，2023年。[43] Nikolay Malkin、Moksh Jain、Emmanuel Bengio、Chen Sun 和 Yoshua Bengio。轨迹平衡：改进 GFlowNets 中的信用分配。*arXiv*预印本 *arXiv:2201.13259*，2022年。[44] Natalie Maus、Haydn T Jones、Juston S Moore、Matt J Kusner、John Bradshaw 和 Jacob R Gardner。面向结构化输入的局部潜在空间贝叶斯优化。*arXiv*预印本 *arXiv:2201.11872*，2022年。[45] Kaisa Miettinen。非线性多目标优化，第12卷。施普林格科学与商业媒体，2012年。[46] Henry Moss、David Leslie、Daniel Beck、Javier Gonzalez 和 Paul Rayson。BOSS：字符串空间上的贝叶斯优化。神经信息处理系统进展，33:15476–15486，2020年。[47] Aviv Navon、Aviv Shamsian、Ethan Fetaya 和 Gal Chechik。利用超网络学习帕累托前沿。载于国际学习表征会议，2021年。[48] Andrei Cristian Nica、Moksh Jain、Emmanuel Bengio、Cheng-Hao Liu、Maksym Korablyov、Michael M Bronstein 和 Yoshua Bengio。评估 GFlowNets 在分子设计中的泛化能力。载于*ICLR2022*药物发现机器学习，2022年。[49] Biswajit Paria、Kirthevasan Kandasamy 和 Barnabás Póczos。使用随机标量化进行多目标贝叶斯优化的灵活框架。载于人工智能中的不确定性，第766–776页。PMLR，2020年。[50] Adam Paszke、Sam Gross、Francisco Massa、Adam Lerer、James Bradbury、Gregory Chanan、Trevor Killeen、Zeming Lin、Natalia Gimelshein、Luca Antiga 等。PyTorch：一种命令式风格、高性能的深度学习库。神经信息处理系统进展，32，2019年。[51] Ethan Perez、Florian Strub、Harm de Vries、Vincent Dumoulin 和 Aaron Courville。FiLM：利用通用条件层进行视觉推理。载于*AAAI*人工智能会议，2018年。[52] Pavel G Polishchuk、Timur I Madzhidov 和 Alexandre Varnek。基于 GDB-17 数据估算类药化学空间的规模。《计算机辅助分子设计杂志》，27(8):675–679，2013年。[53] Binxin Ru、Xingchen Wan、Xiaowen Dong 和 Michael Osborne。通过结合魏斯费勒-莱曼核的贝叶斯优化实现可解释的神经架构搜索。*arXiv*预印本 *arXiv:2006.07556*，2020年。[54] Bobak Shahriari、Kevin Swersky、Ziyu Wang、Ryan P Adams 和 Nando De Freitas。让人类退出循环：贝叶斯优化综述。《*IEEE*会刊》，104 (1):148–175，2015年。[55] Niranjan Srinivas、Andreas Krause、Sham Kakade 和 Matthias Seeger。多臂老虎机设定下的高斯过程优化：无遗憾性与实验设计。载于第*27*届国际机器学习会议论文集，第1015–1022页，2010年。[56] Richard S Sutton 和 Andrew G Barto。强化学习：导论。麻省理工学院出版社，2018年。

[57] Kevin Swersky、Yulia Rubanova、David Dohan 和 Kevin Murphy。在离散空间上的摊销贝叶斯优化。载于人工智能不确定性会议，第769–778页。PMLR，2020年。[58] Austin Tripp、Erik Daxberger 和 José Miguel Hernández-Lobato。通过加权再训练，在深度生成模型的潜在空间中实现样本高效的优化。神经信息处理系统进展，33：11259–11272，2020年。[59] Austin Tripp、Gregor NC Simm 和 José Miguel Hernández-Lobato。重新审视从头分子设计基准。载于*NeurIPS 2021* 科学人工智能研讨会，2021年。[60] Robin Winter、Floriane Montanari、Andreas Steffen、Hans Briem、Frank Noé 和 Djork-Arné Clevert。在连续潜空间中高效进行多目标分子优化。《style id='11'>化学科学，10(34)：8016–8024，2019年。[61] Yutong Xie、Chence Shi、Hao Zhou、Yuwei Yang、Weinan Zhang、Yong Yu 和 Lei Li。MARS：用于多目标药物发现的马尔可夫分子采样。载于国际学习表征会议，2021年。[62] Jiaxuan You、Bowen Liu、Zhitao Ying、Vijay Pande 和 Jure Leskovec。用于目标导向分子图生成的图卷积策略网络。神经信息处理系统进展，31，2018年。[63] Dinghuai Zhang、Nikolay Malkin、Zhen Liu、Alexandra Volokhova、Aaron Courville 和 Yoshua Bengio。用于离散概率建模的生成流网络。*arXiv* 预印本 *arXiv:2202.01361*，2022年。[64] Qingfu Zhang 和 Hui Li。MOEA/D：一种基于分解的多目标进化算法。*IEEE Transactions on Evolutionary Computation*，11(6)：712–731，2007年。[65] Yiyang Zhao、Linnan Wang、Kevin Yang、Tianjun Zhang、Tian Guo 和 Yuandong Tian。通过学习空间划分实现多目标优化。载于国际学习表征会议，2022年。URL https://openreview.net/forum?id=FlwzVjfMryn。[66] Yiheng Zhu、Zhenqiu Ouyang、Ben Liao、Jialu Wu、Yixuan Wu、Chang-Yu Hsieh、侯廷军和Jian Wu。MolHF：一种用于分子图生成的分层归一化流。载于第三十二届人工智能国际联合会议论文集，第5002–5010页，2023年8月。[67] Eckart Zitzler 和 Lothar Thiele。多目标进化算法：一项比较案例研究及强度帕累托方法。*IEEE Transactions on Evolutionary Computation*，3(4)：257–271，1999年。

A 算法

算法1描述了所提出的多目标贝叶斯优化算法的整体框架，其中采用HN-GFN作为采集函数优化器。算法2描述了多目标贝叶斯优化中HN-GFN的训练过程。

算法1基于 HN-GFN 的多目标贝叶斯优化

输入：预言机 $f = (f_1, \ldots, f_M)$、初始数据集 $\mathcal{D}_0 = \{(x_i^0, f(x_i^0))\}_{i=1}^n$、采集函数 a、狄利克雷分布的参数 α、轮数 N、批次大小 b初始化：代理模型 $\mathcal{M}$、HN-GFN 的参数 π_θ for $i = 1$to N do 在数据集 $\mathcal{D}_{i-1}$ 上拟合代理模型 $\mathcal{M}$采样目标偏好权重集合 $\Lambda \sim \text{Dir}(\alpha)$ 使用奖励函数 $R_\lambda(x)$ 训练 $\pi_\theta = a(\mu(s_\lambda(x))\ \sigma(s_\lambda(x)); \mathcal{M}) \triangleright$ 算法2 采样查询批次 $\mathcal{B}_i = \{x_j^i\}_{j=1}^b$ 基于 $\lambda_{target} \in \Lambda$使用 f评估批次 $\mathcal{B}_i$，并扩充数据集 $\mathcal{D}_{i+1} = \mathcal{D}_i \cup \{(x_j^i, f(x_j^i))\}_{j=1}^b$ end for

---

算法2 使用类事后离策略策略的 HN-GFN 训练过程 输入：可用数据集 $\mathcal{D}_i$、奖励函数 R、小批量大小 m、目标偏好向量集合Λ、类事后策略的比例 γ、重放缓冲区 $\{\mathcal{R}_\lambda\}_{\lambda\in\Lambda}$while 未收敛do 标记 $\sim$ 伯努利分布(γ) if 标记 $= 1$then$\lambda \sim \Lambda$从重放缓冲区 $\mathcal{R}_\lambda$中采样 $\frac{m}{2}$ 条轨迹 $\lambda \sim$elseαDir(α) 从可用数据集 $\mathcal{D}_i$中采样 $\frac{m}{2}$ 条轨迹$\theta =$ end if$\theta_{\text{mpnn}}(h,\lambda(\phi; \frac{m}{2}))$ 从策略 $\tilde{\pi}$ 中采样 x 条轨迹，并将终止状态存储在 $\mathcal{R}_\lambda$ 中 $\lambda \in \Lambda$ 对所有 R_λ，计算小批量中每条轨迹的终止状态 x 上的奖励 x使用关于公式（4）的随机梯度下降步骤更新参数 θ_{mpnn} 和 ϕ end while

—

—

B 实现细节

B.1 实验设置

B.1.1 分子领域

按照 Bengio 等人 [6], 的方法，分子通过逐步连接片段来生成。片段从预定义的构建单元词汇表中选取，并连接到部分构建的分子中的一个原子上。最大轨迹长度为 8，动作数量随状态变化，在 100 到 2000 之间，因此最多可达 $|\mathcal{X}|\ 10^{16}$。我们采用 Xie 等人发布的性质预测模型 [61] ，评估生成分子对 GSK3β 和 JNK3 的抑制能力。

B.1.2 度量指标

多样性。 多样性（Div）是 Morgan 指纹之间两两 Tanimoto 距离的平均值。在合成场景中，对于每个偏好向量，我们采样 1000 个分子，计算排名前100的分子中的 Div，并报告各偏好下的平均值。在多目标贝叶斯优化（MOBO）中，Div 是按每轮 100 个候选项的批次计算的，因为图遗传算法和 MARS 不受偏好向量控制。我们认为，这一指标可能更符合这些方法在生物学或化学实验中的实际用法。

相关性。 相关性（Cor）是指 GFlowNet 从外部测试集采样分子的概率与其相应奖励在对数域中的斯皮尔曼秩相关系数：Cor $=$ Spearman's $\rho_{\log(\pi(x)),\log(R(x))}$。外部测试集通过两个步骤获得：首先，根据构件数量均匀生成一个包含 300K 个分子的随机数据集；接着，分别从这 300K 个分子中采样属性分布均匀、对应于 GSK3β 和 JNK3 的测试集。

B.2 基线方法

所有基线方法均基于公开发布的源代码实现，并针对我们的多目标贝叶斯优化场景进行了调整。进化算法（MOEA/D 和 NSGA-III）使用 PyMOO [9]实现，潜空间优化方法（qParEGO、 q期望超体积改进和 LaMOO）使用 BoTorch [4]实现。在多目标贝叶斯优化场景中，LaMOO 和图遗传算法使用期望超体积改进作为采集函数。对于所有基于高斯过程的方法，每个目标均由一个独立的高斯过程建模。

B.3 HN-GFN

我们使用 PyTorch [50]实现所提出的 HN-GFN。关键超参数的取值见表 4。

代理模型：在实验中，我们采用 12 层 MPNN 作为代理模型的基础架构。在多目标贝叶斯优化中，我们使用批量大小为 64 的数据训练单个多任务消息传递神经网络，并采用 Adam优化器、0.1 的丢弃率和 1e-6 的权重衰减率。我们采用早停来提升泛化能力。

HN-GFN：HN-GFN由一个标准GFlowNet和一个偏好条件超网络组成。GFlowNet的架构是一个10层MPNN，超网络则是一个具有多个头的3层多层感知机，每个头都为目标网络的不同层生成权重。HN-GFN使用Adam优化器进行训练，以优化流匹配目标。

表 4：现实世界 MOBO 实验中使用的超参数。

超参数	GSK3β +JNK3	GSK3β + JNK3 +定量估计药物相似性 + 合成可及性
代理模型		
隐藏层大小	64	64
学习率	2.5e-4	1e-3
λ 用于证据回归	0.1	0.1
迭代次数	10000	10000
早停耐心值	500	500
随机失活	0.1	0.1
权重衰减	1e-6	1e-6
采集函数（*UCB*）		
β	0.1	0.1
HN-GFN		
学习率	5e-4	5e-4
奖励指数	8	8
奖励范数	1.0	1.0
轨迹小批量大小	8	8
离线小批量大小	8	8
事后回顾 γ	0.2	0.2
均匀策略系数	0.05	0.05
GFlowNet 隐藏层大小	256	256
超网络隐藏层大小	100	100
训练步数	5000	5000
α	(1,1)	(3,4,2,1)

B.4 实测运行时间

在相同的计算设施上使用1块Tesla V100 GPU进行效率比较。在多目标贝叶斯优化中，三种潜空间优化方法（即 qParEGO、 q期望超体积改进和LaMOO）的运行时间约为3小时，而图遗传算法的优化速度快得多，仅需13分钟。相比之下，基于深度学习的离散优化方法的时间复杂度要高得多。MARS需要32小时，而我们提出的HN-GFN需要10小时。采用类似事后回顾的训练策略后，HN-GFN的运行时间将增加约33%。
然而，从更宏观的角度来看，与在实际应用中评估候选分子所需的时间相比，模型训练耗时可能微不足道。因此，我们认为，候选项的高质量（即多目标贝叶斯优化算法的性能）比降低训练成本更为重要。

C 其他结果

C.1 多目标贝叶斯优化实验中的采样分子

我们给出了在 GSK3β +JNK3 + 定量估计药物相似性 + 合成可及性优化设置下，由 HN-GFN 从帕累托前沿采样得到的一些分子示例（图 4）。每个分子下方的数字分别表示 GSK3β、JNK3、定量估计药物相似性和合成可及性评分。

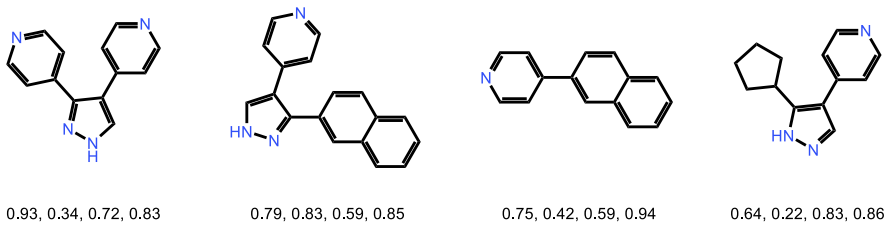


图 4：HN-GFN 从近似帕累托前沿采样得到的分子。

C.2 合成场景

如图 5 所示，Top-100 GSK3β 分数的分布在偏好特定的生成流网络和我们提出的 HN-GFN 中呈现出一致的趋势，尽管这一趋势不如 JNK3 性质中的趋势显著。

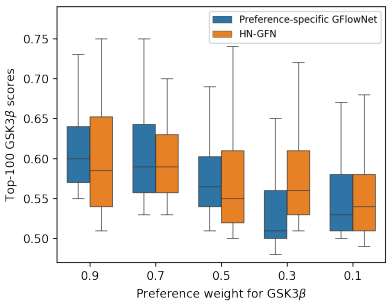


图 5：比较使用特定于偏好向量的 GFlowNets 和 HN-GFN 采样得到的 Top-100 GSK3β 分数分布。

C.3 代理模型的影响

我们进行了消融实验，以研究不同代理模型的有效性。我们考虑以下三种代理模型：证据回归 [1], 深度集成 [39], 以及基于 Tanimoto 核的高斯过程 [59]。如表 5 所示，证据回归的优化性能优于深度集成。虽然证据回归和高斯过程的超体积相当，但证据回归能够提出更多样化的候选项。此外，我们认为，相较于证据回归和深度集成，高斯过程在离散空间中的灵活性较低，因为需要根据数据结构设计不同的核。

表 5：多目标贝叶斯优化场景中不同代理模型的评估

	GSK3β +JNK3		GSK3β + JNK3 + 定量估计药物相似性 +合成可及性	
	HV	Div	HV	Div
HN-GFN（证据型）	**0.669 ± 0.061**	0.793 ± 0.007	0.416 ± 0.023	0.738 ± 0.009
HN-GFN（集成）	0.583 ± 0.103	**0.797 ± 0.004**	0.355 ± 0.048	**0.761 ± 0.012**
HN-GFN（高斯过程）	0.662 ± 0.054	0.739 ± 0.008	**0.421 ± 0.037**	0.683 ± 0.018

D 与并行工作的比较

尽管 Bengio 等人 [7], 从理论上讨论了帕累托生成流网络的概念，但我们是首批针对多目标优化研究并实例化这一概念的团队之一，并解决了原始理论阐述中未充分讨论的实际挑战，以期让高样本效率的分子优化成为现实。我们深入研究了条件机制（第 5.1 节）和代理模型（附录 C.3）的影响。此外，我们还精心提出了一种类事后离策略策略（第 4.3 节），而该策略在多目标优化领域（无论是强化学习还是 GFlowNet 文献中）鲜有研究。

我们注意到，Jain 等人 [29], 提出了一项并行研究，同样采用 GFlowNets 解决多目标优化问题。在 Jain 等人 [29], 的研究中，他们针对单轮场景和贝叶斯优化（BO）场景，分别提出了两种变体：MOGFN-PC 和 MOGFN-AL。事实上，当 HN-GFN 在多目标贝叶斯优化之外作为独立优化器使用时（第 5.1 节），除采用不同的条件机制外，它与 MOGFN-PC 类似。至于多目标贝叶斯优化，MOGFN-AL 是一种标准GFlowNet，其奖励函数定义为多目标采集函数（NEHVI [13]）。除了扩展 GFlowNet 外，我们还精心提出了一种类事后离策略策略（第 4.3 节），而该策略在多目标优化领域（无论是强化学习还是 GFlowNet 文献中）鲜有研究。