arXiv:2603.00877v1 [cs.LG] 1 Mar 2026

主动流匹配

Yashvir S. Grewal [1 2] **Daniel M. Steinberg** [2] **Thang D. Bui** [1] **Cheng Soon Ong** [1 2] **Edwin V. Bonilla** [2]

摘要

离散扩散和流匹配模型通过并行迭代细化，捕捉高维目标景观中的复杂、非加性和非自回归结构。然而，它们隐式的生成特性使其无法直接与用于在线黑盒优化的严格变分框架相结合，例如变分搜索分布（VSD）和自适应采样条件化（CbAS）。我们提出主动流匹配（*AFM*），通过重新表述变分目标，使其能够作用于沿流的条件端点分布，从而在保留 VSD 和 CbAS 严谨性的同时，基于梯度引导流模型朝高适应度区域移动。我们利用自归一化重要性采样推导出前向和反向 Kullback-Leibler（KL）变体。在一系列在线蛋白质和小分子设计任务中，正向 KL 主动流匹配的表现始终与最先进基线方法相当，展现了其在实验预算紧张时有效平衡探索与利用的能力。

1. 引言

蛋白质设计等领域中高维目标景观所具有的复杂非加性交互作用（Starr & Thornton，2016年；Phillips，2008）限制了标准自回归（AR）模型的有效性。事实上，这些交互作用在蛋白质进化中被称为上位性，它们违背了 AR 模型所依据的序列因子分解；该模型在观察后续上下文之前，就会对每个词元作出确定。

另一种方法是离散流匹配（DFM），通过并行迭代细化生成序列（Austin 等人，2021；Gat 等人，2024）。在每一步中，词元更新都以完整的序列上下文为条件，使模型能够协调长程依赖关系，而序列的自回归因式分解可能无法捕捉这些关系。因此，DFM模型如今在具有挑战性的蛋白质和RNA 设计任务中，表现已达到或超过自回归基线方法（Alamdari 等人，2023年；Watson 等人，2023年；Ingraham 等人，2023年）。

然而，要将这些生成能力转化为实际发现，就必须满足有限的实验预算。设计空间不仅具有组合性（例如，$20^{20} \approx 10^{26}$ 对于含20个残基的肽），而且实验验证仍然成本高昂（生物物理学与 Core，2024；Core，2024）。为此，变分搜索分布（VSD，Steinberg 等人，2025年）和自适应采样条件化（CbAS，Brookes 等人，2019年）等主动生成框架会生成合理的序列并将概率质量集中于罕见的高适应度区域，同时保持多样性。这些方法在方法论上有充分依据，因为它们以近似概率推断为基础，但也有自身的局限：需要具备可处理的概率密度 $q_\phi(\mathbf{x})$，以便直接计算对数概率（CbAS）或估计得分函数梯度（VSD）。

因此，将 DFM模型的灵活性与 VSD 或 CbAS 等主动生成框架相结合，无疑是一个很有前景的研究方向。然而，最先进的 DFM 和离散扩散模型都是隐式生成器：它们无法给出离散序列上的归一化分布。因此，通常难以计算或求导 $\log q_\phi(\mathbf{x})$。更具体地说，对于 DFM，目前的公式无法提供简单的闭式质量函数（Lipman等人，2022；Gat 等，2024）；而对于离散扩散，精确计算 $\log q_\phi(\mathbf{x})$ 需要对数量呈指数级增长的腐蚀路径求和（Austin等人，2021）。因此，将需要 $\log q_\phi(\mathbf{x})$ 或其梯度$\nabla_\phi \log q_\phi(\mathbf{x})$ 的主动生成方法直接应用于这些生成模型颇具挑战。

主动流匹配（AFM）。我们通过重新构建变分目标，使其作用于流模型确实能够提供的条件端点分布，而非难解边缘分布，从而应对这一挑战。这使我们能够运用 VSD 和 CbAS 固有的、基于原则的探索-利用权衡，引导离散流模型朝高适应度区域发展。我们利用自归一化重要性采样推导了正向 KL 主动流匹配和反向 KL

[1]澳大利亚国立大学系统与社会学院计算机学院，澳大利亚 [2]澳大利亚联邦科学与工业研究组织 Data61。联系人：Yashvir Grewal<Yashvir.Grewal@anu.edu.au>。

预印本。*2026*年*3*月*3*日。

主动流匹配的两种变体，并引入一种混合提议分布，以平衡探索、利用和细化。在一系列在线蛋白质和小分子设计任务中，与现有基线方法相比，正向 KL 主动流匹配表现出竞争力；在许多情况下，即使实验预算有限，也能发现得分更高的设计方案。

2. 背景

2.1. 主动生成

考虑一个离散设计空间 $\mathcal{X}$（例如，$\mathcal{V}^L$，表示长度为 L、词汇表大小为 $\mathcal{V}$ 的序列），以及一个未知的适应度函数 $f: \mathcal{X} \to \mathbb{R}$。我们只能通过成本高昂的实验来评估 f，这些实验会返回含噪测量值：

$$y = f(\mathbf{x}) + \epsilon, \quad \epsilon \sim p(\epsilon), \quad \mathbb{E}[\epsilon] = 0.$$

我们的目标是发现超水平集$\mathcal{S}_\tau = \{\mathbf{x} \in \mathcal{X}: y \geq \tau\}$中的设计方案，其中阈值 τ是指定的。关键在于，$\mathcal{S}_\tau$ 在高维空间中呈指数级稀少，因此从任意固定先验 $p(\mathbf{x})$ 中进行随机采样都无济于事。Steinberg 等人（2025a）将其形式化为主动生成问题：从初始观测值 $\mathcal{D}_0 = \{(\mathbf{x}_n, y_n)\}_{n=1}^N$出发，在 R 轮中反复提出批次 $\mathcal{B}_r = \{\mathbf{x}_b\}_{b=1}^B$ 进行评估，并累积数据 $\mathcal{D}_r = \mathcal{D}_{r-1} \cup \{(\mathbf{x}_b, y_b)\}_{b \in \mathcal{B}_r}$。目标是学习一个生成模型 $q^\phi_{\mathcal{D}_r}(\mathbf{x}) \approx p(\mathbf{x} \mid y \geq \tau)$，使其将概率集中在 $\mathcal{S}_\tau$ 上，同时在批次评估中保持多样性。主动生成也用于黑盒优化，即 $\arg\max_{\mathbf{x}} f(\mathbf{x})$，方法是单调提高阈值，使得 $\tau_r > \tau_{r-1}$。

这带来了三个关键挑战：(i) 生成模型必须适应稀疏的序贯反馈，而不是固定数据集；(ii) 批次提议需要在探索与利用之间权衡，而监督学习中不存在这种权衡；(iii) $\mathcal{X}$ 的离散组合性质使直接基于梯度的搜索无法实现。VSD 和 CbAS 都采用密度估计方法来解决这一问题，其中 CbAS 使用前向 Kullback-Leibler（KL）散度，

$$\phi_r^* = \arg\min_\phi \mathrm{KL}\Big(p(\mathbf{x} \mid y \geq \tau) \| q^\phi_{\mathcal{D}_r}(\mathbf{x})\Big), \qquad (1)$$

而 VSD 使用反向 KL 散度，

$$\phi_r^* = \arg\min_\phi \mathrm{KL}\Big(q^\phi_{\mathcal{D}_r}(\mathbf{x}) \| p(\mathbf{x} \mid y \geq \tau)\Big). \qquad (2)$$

值得注意的是，CbAS 最初是为离线优化任务而设计的，但(Steinberg 等人，2025年等人已将其适配到在线场景中。

Steinberg 等人对 CbAS 和 VSD 的高效实现中，一个关键之处在于他们发现Steinberg 等人(2025年；b)条件集合 $\mathbb{I}\{y \geq \tau\}$ 可以通过二分类任务高效地近似，即通过类别概率估计器 $p(z = 1|x)$，其中 $z = 1$ 当 $y \geq \tau$ 时。

2.2. 离散流匹配

离散流匹配（Gat 等人，2024；Campbell 等人，2024）为离散序列的非自回归生成提供了一个有原则的框架。其关键见解是，尽管样本始终保持离散状态，概率质量却会沿着时间相关路径连续流动，从而能够通过对序列中所有位置进行并行更新来实现迭代细化。接下来，我们将简要概述离散流匹配；如需了解更多详情，请参阅Lipman 等人（2024）。

设定。 考虑离散序列$\mathbf{x} \in \mathcal{V}^L$，其中每个 L位置取值于 $|\mathcal{V}|$ 词汇表中的一个值。生成式建模的任务是将来自简单源 分布 $p(\mathbf{x}_0)$ 的样本转换为来自数据分布 $q(\mathbf{x}_1)$ 的样本。源分布有两种自然选择：(i)均匀分布：对于每个位置，$p^{\text{unif}}(\mathbf{x}) = 1/|\mathcal{V}|$；(ii)掩码：$p^{\text{mask}}(\mathbf{x}) = \delta_M(\mathbf{x})$，将概率质量集中在特殊的掩码词元 M 上。对于单个词元位置，我们用 $\delta_y(x)$ 表示克罗内克δ函数，它表示一种点质量分布，将全部概率质量赋予状态y：若 $x = y$，则 $\delta_y(x) = 1$；否则为 0。对于序列，$\delta_{\mathbf{y}}(\mathbf{x})$ 当且仅当所有位置都匹配时才为 1：$\delta_{\mathbf{y}}(\mathbf{x}) = \prod_{i=1}^L \delta_{y^i}(x^i)$。

训练期间，我们会访问来自满足边缘约束 $\sum_{\mathbf{x}_1} \pi(\mathbf{x}_0, \mathbf{x}_1) = p(\mathbf{x}_0)$ 和 $\sum_{\mathbf{x}_0} \pi(\mathbf{x}_0, \mathbf{x}_1) = q(\mathbf{x}_1)$ 的耦合$\pi(\mathbf{x}_0, \mathbf{x}_1)$ 的成对样本 $(\mathbf{x}_0, \mathbf{x}_1)$。最简单的选择是独立耦合：$\pi(\mathbf{x}_0, \mathbf{x}_1) = p(\mathbf{x}_0)q(\mathbf{x}_1)$，不过最优传输耦合可能有助于改善训练（Pooladian 等，2023）。

概率路径。我们定义一个随时间变化的概率路径 $p_t(x)$，使其在 $t \in [0, 1]$ 上从源分布平滑插值到目标分布，并满足 $p_0 = p$ 和 $p_1 = q$。该路径通过条件概率路径$p_t(x \mid x_0, x_1)$构造，该路径在每个训练样本对之间进行插值。对于掩码源，一个自然的选择是凸插值：

$$p_t(x \mid x_0, x_1) = (1 - \kappa_t)\delta_{x_0}(x) + \kappa_t \delta_{x_1}(x) \qquad (3)$$

其中 $\kappa_t: [0, 1] \to [0, 1]$ 是一个单调递增的调度器，满足 $\kappa_0 = 0$和 $\kappa_1 = 1$。常见选择包括线性（$\kappa_t = t$）或二次（$\kappa_t = t^2$）调度。方程 3 描述了一个随机过程：在时间 t，我们以概率（$1-\kappa_t$）观测到源词元 x_0，以概率 κ_t 观测到目标词元 x_1。随着时间推移，目标逐渐显现。关键的是，时间 t的状态始终恰好是 $\{x_0, x_1\}$中的一个。插值发生在概率空间中，而非词元空间中。边缘路径可通过 $p_t(x) = \sum_{x_0,\ x_1} p_t(x \mid x_0,\ x_1)\pi(x_0,\ x_1)$获得。

连续时间马尔可夫链动力学。要从目标分布 q中采样，我们需要一种机制，使样本沿着概率路径从 $t=0$演化至$t=1$。在离散状态空间中，这种演化由连续时间马尔可夫链（Campbell 等人，2022）控制，其特征是一个随时间变化的速率矩阵。对于无穷小时间步长 dt，从状态 x_t 转移到状态 x' 的概率由速率 $R_t(x' \mid x_t)dt$给出。实际中，轨迹通过有限时间步长 h进行模拟：

$$p(x_{t+h}=x' \mid x_t)=\delta_{x_t}(x')+h\cdot u_t(x' \mid x_t)+o(h) \quad (4)$$

其中，$u_t(x' \mid x_t)$是概率速度，指向状态x'，满足 $\sum_{x'} u_t(x' \mid x_t)=0$且 $u_t(x' \mid x_t)\geq 0$，从而保证$x'\neq x_t$ 的概率转移有效。

从后验分布生成速度。Gat 等人的一个核心结果(2024) 表明，生成路径方程 3 的概率速度完全由后验分布 $p_{1|t}(x_1 \mid x_t)$决定；该分布用于估计给定当前状态 x_t时最终目标分布 x_1：

$$u_t(x' \mid x_t)=\frac{\dot{\kappa}_t}{1-\kappa_t}[p_{1|t}(x' \mid x_t)-\delta_{x_t}(x')] \quad (5)$$

其中，$\dot{\kappa}_t$ 是调度器的时间导数。项 $\delta_{x_t}(x')$确保各行之和为零，从而满足 Kolmogorov 前向方程。这种速度带来了优雅的解释：在每个时间步，概率质量都会流向可能成为目标的词元 x'（根据 $p_{1|t}$ 判断），并远离当前词元 x_t；其费率由信噪比 $\dot{\kappa}_t/(1-\kappa_t)$ 控制。在生成初期（ κ_t较小），演化缓慢且具有探索性；随着 $\kappa_t\to 1$，演化会加速，以锁定目标。

学习后验分布。 由于真实后验分布 $p_{1|t}(x_1 \mid x_t)$ 难以处理，我们训练一个参数化神经网络$q_\phi(x_1 \mid x_t,\ t)$来对其进行近似。标准训练目标函数通过最小化模型预测与真实目标分布 x_1之间的期望交叉熵来实现：

$$\mathcal{L}_{\mathrm{CE}}(\phi)=\mathbb{E}_{t,x_0,x_1,x_t}[-\log q_\phi(x_1 \mid x_t,t)]. \quad (6)$$

该目标函数对应一个去噪任务：(a) 采样一个时间 $t\sim\mathcal{U}[0,1]$ 和一对数据 $(x_0,\ x_1)\sim\pi$；(b) 从条件路径 x_t 中采样一个中间状态 $p_t(x \mid x_0,\ x_1)$，该路径见方程 3；(c) 更新 ϕ，以最大化给定带噪状态时干净目标分布 x_1 的似然 x_t。

采样算法。 在推理时，我们初始化 $x_0\sim p_0$，并利用欧拉离散化动力学迭代演化样本。对于时间步 $t=0,\ h,\ 2h,\ \ldots\ldots,\ Kh$，其中 $h=1/K$，状态 x_t 通过从分类转移核 $x_{t+h}\sim$Cat（$\delta_{x_t}(\cdot)+h\cdot u_t(\cdot \mid x_t)$）中采样而更新为 x_{t+h}。对于掩码源 $p_0(x)=\delta_M(x)$，该过程从全掩码序列开始。

随着 t 推进，速度场 u_t 将概率质量从掩码词元 M 移向由后验分布预测的特定词元，从而有效地对该序列进行“去掩码”。步数 K 在样本质量与计算成本之间进行权衡（Gat 等人，2024）。

扩展到序列。 对于长度为 L的序列，生成过程按位置分解，但通过速度相互耦合。具体而言，条件概率路径可以独立分解：

$$p_t(\mathbf{x} \mid \mathbf{x}_0,\mathbf{x}_1)=\prod_{i=1}^{L} p_t(x^i \mid x_0^i,x_1^i). \quad (7)$$

然而，学习到的速度会使这些位置相互耦合。近似后验分布 $q_\phi(x_1^i \mid \mathbf{x}_t,\ t)$ 会通过条件化于整个当前序列 $\mathbf{x}_t$，来预测位置 i 处的目标词元。这样，模型既能捕捉复杂的非局部依赖，又能保持 O 次并行生成步骤。与此相反，自回归模型需要 $O(L)$ 个顺序步骤。

隐式生成器问题。尽管离散流匹配擅长捕捉全局结构，但从根本上说，它是一个隐式生成器。模型通过优化交叉熵目标方程6 来学习条件后验分布 $q_\phi(\mathbf{x}_1 \mid \mathbf{x}_t,t)$，该后验分布定义了向量场。然而，这种形式无法给出边缘似然的闭式表达 $q_\phi(\mathbf{x})$。直接计算似然$\log q_\phi(\mathbf{x})$ 需要对从源到$\mathbf{x}$的所有可能随机轨迹进行边缘化：

$$q_\phi(\mathbf{x})=\sum_{\mathbf{x}_0} p(\mathbf{x}_0)\sum_{\text{paths }\gamma:\mathbf{x}_0\to\mathbf{x}} \mathbb{Q}_\phi[\gamma]. \quad (8)$$

对于离散序列，路径数量关于 L 和步数 K 呈组合爆炸式增长，因此这一求和无法计算。于是，梯度 $\nabla_\phi \log q_\phi(\mathbf{x})$ 也无法获得。

这种难以处理的问题与有原则的主动生成相冲突：VSD 需要 $\nabla_\phi \log q_\phi(\mathbf{x})$ 来估计证据下界（ELBO）梯度，而 CbAS 则需要涉及 $q_\phi(\mathbf{x})$ 的显式密度比。因此，标准主动生成框架在数学上与隐式流模型不兼容。接下来，我们通过重新表述这些目标函数，使其作用于条件端点分布来解决这一问题；流模型自然会提供这些分布。

3. 主动流匹配

在确立离散流匹配模型能够提供条件端点分布、但缺乏可处理的边缘似然 $\log q_\phi(\mathbf{x})$之后，我们重新表述变分目标，使其沿流路径直接作用于这些条件分布。

3.1. 问题设定

我们考虑一个进行 R 轮次的主动生成场景。在第 r 轮，我们假设已收集到一个带标签数据集 $\mathcal{D}_r = \{(\mathbf{x}_n, y_n)\}_{n=1}^{N_r}$。我们的目标是学习一个参数为 ϕ 的主动流匹配模型，使其将概率质量集中在超水平集 $\mathcal{S}_\tau = \{\mathbf{x} \in \mathcal{X} : y \geq \tau\}$ 上，同时为下一批提案保持多样性 $\mathcal{B}_r$。该流由一个去噪神经网络参数化，该网络能够捕捉模型的条件端点分布 $q_t^\phi(\mathbf{x}_1 \mid \mathbf{x}_t)$；为简洁起见，我们省略了对 t 的条件设定。

此外，我们假设可以使用参数为 θ 的基础流。该模型采用与上述主动流相同的架构，并由参考分布 $q_t^\theta(\mathbf{x}_1 \mid \mathbf{x}_t)$ 参数化。它通常源自先前的优化状态，例如第 $(r-1)^{\text{th}}$ 轮得到的解，或当前主动流模型的滞后快照。该模型将用于反向 KL 目标和重要性采样的提议混合分布。最后，数据集 $\mathcal{D}_r$ 用于训练分类器，以估计适应度较高的概率 $p_{\mathcal{D}_r}(y \geq \tau \mid \mathbf{x})$，其中适应度阈值为预先设定的 τ；该概率用于对生成模型进行条件化。

关键见解。 标准变分方法（VSD、CbAS）会最小化涉及边缘分布 $q_\phi(\mathbf{x})$ 的散度。由于计算该边缘分布需要对所有离散路径进行难以处理的求和，我们将目标提升到条件流路径上。我们不再匹配 $t = 1$ 处的最终分布，而是匹配在流时间 $t \in [0, 1]$ 和中间状态 $\mathbf{x}_t$ 上取平均的条件分布。这种做法利用了流中易于计算的去噪器 $q_t^\phi(\mathbf{x}_1 \mid \mathbf{x}_t)$，同时在前向 KL 的情况下，仍能从所需的边缘分布中采样。接下来，我们将介绍所提出的方法及其实现细节。

3.2. 前向 KL 主动流匹配

遵循 CbAS（Brookes 等人，2019年），我们最小化真实条件端点分布与主动流匹配模型分布之间的前向 KL：

$$\begin{aligned}\mathcal{L}_{\text{fwd}}(\phi) &= \mathbb{E}_{t,\mathbf{x}_t \mid y \geq \tau}\left[\mathrm{KL}\Big(p_t(\mathbf{x}_1 \mid \mathbf{x}_t, y \geq \tau) \,\|\, q_t^\phi(\mathbf{x}_1 \mid \mathbf{x}_t)\Big)\right], \\ &= -\mathbb{E}_{t,\mathbf{x}_1 \mid y \geq \tau, \mathbf{x}_t \mid \mathbf{x}_1}\left[\log q_t^\phi(\mathbf{x}_1 \mid \mathbf{x}_t)\right] + C. \end{aligned} \tag{9}$$

这里 C 相对于 ϕ 为常数，且目标后验分布由基础流 θ 定义：

$$p_t(\mathbf{x}_1 \mid \mathbf{x}_t, y \geq \tau) \propto q_t^\theta(\mathbf{x}_1 \mid \mathbf{x}_t)\, p_{\mathcal{D}_r}(y \geq \tau \mid \mathbf{x}_1).$$

由于无法从 $p(\mathbf{x}_1 \mid y \geq \tau)$ 中采样，我们利用自归一化重要性采样（SNIS），并采用提议分布 $\mu(\mathbf{x}_1)$：

$$\mathcal{L}_{\text{fwd}}(\phi) \approx -\mathbb{E}_t\left[\frac{\sum_{k=1}^K w_k \log q_t^\phi(\mathbf{x}_{1,k} \mid \mathbf{x}_{t,k})}{\sum_{k=1}^K w_k}\right], \tag{10}$$

其中 $\{\mathbf{x}_{1,k}\}_{k=1}^K \sim \mu(\mathbf{x}_1)$，权重 $w_k = p_{\mathcal{D}_r}(y \geq \tau \mid \mathbf{x}_{1,k}) / \mu(\mathbf{x}_{1,k})$，$t \sim$ 均匀分布$[0, 1]$，以及 $\mathbf{x}_{t,k} \sim p_t(\mathbf{x}_t \mid \mathbf{x}_{1,k})$。我们接下来将证明，优化上述目标函数将产生来自所需边缘分布的样本。

定理 3.1（前向 KL AFM 的一致性）。设$p^*(\mathbf{x}) \propto p_{\text{prior}}(\mathbf{x})\, w(\mathbf{x})$ 为目标分布，其中$w(\mathbf{x}) = p(y \geq \tau \mid \mathbf{x})$ 且 p_{prior} 为均匀分布。在采用掩码源耦合与凸插值路径的标准 *DFM* 假设下（方程*3*），以及严格为正的调度器条件下，全局极小值 ϕ^* 所对应的正向 *KL* 主动流匹配目标函数（方程*10*）得到终端分布 $q_1^{\phi^*} = p^*$ 几乎处处成立。

证明概要。我们将重要性加权训练归结为 p^* 上的标准 DFM。关键恒等式为：

$$\mathbb{E}_{\mathbf{x} \sim p_{\text{prior}}}\left[w(\mathbf{x})\mathcal{L}_{\text{CE}}(\mathbf{x}; \phi)\right] = Z \cdot \mathbb{E}_{\mathbf{x} \sim p^*}\left[\mathcal{L}_{\text{CE}}(\mathbf{x}; \phi)\right] \tag{11}$$

其中 Z 为归一化，且 $\mathcal{L}_{\text{CE}}$ 见方程 6。由于 Z独立于 ϕ，在先验上最小化加权目标函数等价于在目标分布上最小化标准 DFM 损失。一致性由 Gat 等人（2024）：极小值点在 p^* 下恢复真实后验分布，并由此定义生成正确边缘路径的速度场。完整证明见附录 A.1。 □

注记 3.2。该结果表明，AFM 不需要新的流动力学理论。这些权重隐式地将训练数据从先验重新加权至 p^*。在实践中（算法1），我们使用来自混合提议分布 $\mu(\mathbf{x})$（方程 15）而非仅来自先验的样本来估计该目标函数。通过以重要性权重校正这些样本，我们得到目标分布对应目标函数的一致梯度估计，而无需从 p^* 本身获取样本。

3.3. 反向 KL 主动流匹配

遵循 VSD（Steinberg 等人，2025年），我们最小化：

$$\begin{aligned}\mathcal{L}_{\text{rev}}(\phi) &= \mathbb{E}_{t,\mathbf{x}_t \mid y \geq \tau}\left[\mathrm{KL}\Big(q_t^\phi(\mathbf{x}_1 \mid \mathbf{x}_t) \,\|\, p_t(\mathbf{x}_1 \mid \mathbf{x}_t, y \geq \tau)\Big)\right]. \\ &= \mathbb{E}_{t,\mathbf{x}_t \mid y \geq \tau, \mathbf{x}_1 \sim q_t^\phi(\cdot \mid \mathbf{x}_t)}\Big[g(\mathbf{x}_1, \mathbf{x}_t)\Big], \end{aligned} \tag{12}$$

其中 $g(\mathbf{x}_1, \mathbf{x}_t) = \log q_t^\phi(\mathbf{x}_1 \mid \mathbf{x}_t) - \log q_t^\theta(\mathbf{x}_1 \mid \mathbf{x}_t) - \log p_{\mathcal{D}_r}(y \geq \tau \mid \mathbf{x}_1)$。为了近似涉及 $\mathbf{x}_t \mid y \geq \tau$ 的期望，我们可以像上文一样使用 SNIS，得到：

$$\mathcal{L}_{\text{rev}}(\phi) \approx \mathbb{E}_t\left[\sum_{k=1}^K \widetilde{w}_k\, \mathbb{E}_{\mathbf{x}_1 \sim q_t^\phi(\cdot \mid \mathbf{x}_{t,k})}\big[g(\mathbf{x}_1, \mathbf{x}_{t,k})\big]\right], \tag{13}$$

算法 1 r 轮主动流匹配（AFM）

输入： 数据 $\mathcal{D}_r$，阈值 τ，提议分布 μ，批量大小 K
输入： 主动流 q_ϕ，分类器 $p_{\mathcal{D}}(y \geq \tau \mid \mathbf{x})$，基础流 p_θ，轮次 r
输入： 目标函数 $\in$ {前向 KL 散度，反向 KL，对称 KL 散度}
每个梯度步执行
  采样 $\{\mathbf{x}_{1,\ k}\}_{k=1}^{K} \sim \mu(\mathbf{x}_1)$
  计算 $w_k \equiv p_{\mathcal{D}}(y \geq \tau \mid \mathbf{x}_{1,\ k})/\mu(\mathbf{x}_{1,\ k})$
  归一化 $\widetilde{w}_k = w_k / \sum_j w_j$
  采样 $t \sim$ 均匀分布$[0,1]$
  采样$\mathbf{x}_{t,\ k} \sim p_t(\cdot \mid \mathbf{x}_{1,\ k})$对所有 k (见方程 3)
  若前向 KL 散度 **则**
    $\mathcal{L} \leftarrow$方程10
  else
    对所有 k，采样 $\mathbf{x}'_{1,k} \sim q_\phi(\cdot \mid \mathbf{x}_{t,k}, t)$
    $\mathcal{L} \leftarrow$ 方程 13 **若** REV-KL **否则**方程14
  结束条件判断
  更新 $\phi \leftarrow \phi - \eta \nabla_\phi \mathcal{L}$
结束 for
输出： 更新后的流 q_ϕ

其中 $\{\mathbf{x}_{1,k}\}_{k=1}^{K} \sim \mu(\mathbf{x}_1)$、$w_k = p_{\mathcal{D}_r}(y \geq \tau \mid \mathbf{x}_{1,k})/\mu(\mathbf{x}_{1,k})$，以及 $\widetilde{w}_k = w_k / \sum_j w_j$。遗憾的是，与正向 KL 主动流匹配不同，我们尚未能证明反向 KL 主动流匹配能够对目标边缘分布给出一致估计。

3.4. 对称 KL 主动流匹配

我们通过最小化对称 KL 散度来结合这两个目标函数：

$$\mathcal{L}_{\text{sym}}(\phi) = \mathcal{L}_{\text{fwd}}(\phi) + \mathcal{L}_{\text{rev}}(\phi), \tag{14}$$

其中 $\mathcal{L}_{\text{fwd}}$ 与 $\mathcal{L}_{\text{rev}}$ 分别定义于式 10 和 13 。对称 KL 从正向项继承模式覆盖行为，并从反向项继承寻模行为。所提方法的实现总结于算法1。

3.5. 提议分布设计

正向和反向 KL AFM 的一个关键组成部分，是为重要性采样选择提议分布 $\mu(\mathbf{x}_1)$。一个好的提议分布必须在三个相互竞争的目标之间取得平衡：(i) 广泛覆盖序列空间，以免错过有前景的区域；(ii) 集中于高适应度的设计方案，以降低估计量方差；(iii) 计算上易于处理。我们引入了一种三分量混合，以兼顾每个目标。

混合定义。 在主动生成轮次 r，我们将提议分布定义为：

$$\mu(\mathbf{x}_1) = \alpha_0\, p_0(\mathbf{x}_1) + \alpha_{\text{flow}}\, q_1^\theta(\mathbf{x}_1) + \alpha_{\text{rbuff}} \sum_{j=1}^{J} \pi_j\, \delta_{x^{(j)}}(\mathbf{x}_1), \tag{15}$$

其中，α_0、α_{flow}和 α_{rbuff} 是满足约束 $\alpha_0 + \alpha_{\text{flow}} + \alpha_{\text{rbuff}} = 1$的非负混合系数；$p_0$ 是序列上的均匀先验；q_1^θ 是基础流 θ（来自上一轮次）的边缘终点分布；$\{\mathbf{x}^{(j)}\}_{j=1}^{J}$ 是一个包含高适应度分数序列及其归一化权重的回放缓冲区：

$$\pi_j = \frac{\exp(\gamma\, y^{(j)})}{\sum_{j'=1}^{J} \exp(\gamma\, y^{(j')})}, \tag{16}$$

其中，$y^{(j)}$ 表示观测到的适应度分数，$\gamma > 0$ 是控制对顶尖序列集中程度的温度参数。

先验分量：权重简化。当提议分布为先验时，$\mu(\mathbf{x}_1) = p_0(\mathbf{x}_1)$，重要性权重会大幅简化。未归一化后验为：

$$\tilde{p}(\mathbf{x}_1 \mid y \geq \tau) = p(y \geq \tau \mid \mathbf{x}_1) \cdot p_0(\mathbf{x}_1), \tag{17}$$

其中，$p(y \geq \tau \mid \mathbf{x}_1)$是分类器似然，$p_0(\mathbf{x}_1)$是先验。对于自归一化重要性采样，我们需要未归一化权重：

$$w(x_1) = \frac{\tilde{p}(\mathbf{x}_1 \mid y \geq \tau)}{\mu(\mathbf{x}_1)} = p(y \geq \tau \mid \mathbf{x}_1). \tag{18}$$

也就是说，先验会完全抵消于重要性权重比中，无论 $p_0(\mathbf{x}_1)$ 是否为均匀分布，重要性权重都会简化为分类器输出。这无需计算先验密度 $p_0(\mathbf{x}_1)$，与（Brookes & Listgarten，2018 ）中的策略类似。

流分量：蒙特卡洛积分。从上一次更新得到的流中采样，会使提议分布集中在模型已经识别为有希望的区域附近，从而实现局部细化。然而，与先验和经验回放缓冲区不同，流的边缘密度 $q_1^\theta(\mathbf{x}_1)$无法写成闭式。因此，我们通过蒙特卡洛积分对其进行估计。

回想一下，基础流 θ 通过学习到的条件分布 $p_\theta(\mathbf{x}_1 \mid \mathbf{x}_0, t = 1)$定义了一个从源 p_0 到数据的随机过程。端点处的边缘密度为：

$$q_1^\theta(\mathbf{x}_1) = \int p_0(\mathbf{x}_0)\, p_\theta(\mathbf{x}_1 \mid \mathbf{x}_0, t = 1)\, d\mathbf{x}_0, \tag{19}$$

其中，θ 表示上一次更新得到的流参数。由于离散状态空间使该积分

难以计算，我们使用 N 蒙特卡罗样本$\{\mathbf{x}_0^{(n)}\}_{n=1}^N \sim p_0$ 进行近似：

$$q_1^\theta(\mathbf{x}_1) \approx \frac{1}{n}\sum_{n=1}^{N} p_\theta(\mathbf{x}_1 \mid \mathbf{x}_0^{(n)}, t=1). \tag{20}$$

这个蒙特卡罗估计量会引入方差，但在实践中，我们发现这种近似足够准确，能够有效地进行训练。

回放缓冲区组件：根据经验观测得到的均一权重。回放缓冲区 $\{\mathbf{x}_1^{(j)}\}_{j=1}^J$ 存储了先前轮次中的序列，这些序列是我们直接观测到的 $f(\mathbf{x}_1^{(j)}) > \tau$，因此确认它们满足约束。当使用提议分布 $\mu(\mathbf{x}_1) = \sum_{j=1}^J \pi_j \delta_{\mathbf{x}_1^{(j)}}(\mathbf{x}_1)$ 完全从回放缓冲区中采样一个批次时，未归一化的重要性权重为：

$$w(\mathbf{x}_1^{(j)}) = \frac{p(y \geq \tau \mid \mathbf{x}_1^{(j)}) \cdot p_0(\mathbf{x}_1^{(j)})}{\pi_j} = \frac{p_0(\mathbf{x}_1^{(j)})}{\pi_j}. \tag{21}$$

如果缓冲区中所有序列的先验概率均匀，或者我们设置 $\pi_j \propto p_0(\mathbf{x}_1^{(j)})$，那么所有权重都近似相等。经过 SNIS 归一化后，得到 $\widetilde{w}_k = 1/K$，从而实现重加权带来的零方差。在实践中，我们采用适应度加权采样 $\pi_j \propto \exp(\gamma y^{(j)})$，以集中采样得分最高的已观测序列，同时保持低方差的梯度估计。

实际实现。虽然公式 15 定义了我们的提议密度，但对每个样本评估完整混合分布 $\mu(\mathbf{x}_1)$都需要计算所有分量密度，代价很高。相反，在每次迭代中，我们选择单个分量 $c \sim$ 分类分布$(\alpha_0, \alpha_{\text{flow}}, \alpha_{\text{rbuff}})$，并从该分量中抽取整个批次。随后仅相对于所选分量密度计算重要性权重：先验为 $w(\mathbf{x}_1) = p_{\mathcal{D}_r}(y \geq \tau \mid \mathbf{x}_1)$，流为$w(\mathbf{x}_1) = p_{\mathcal{D}_r}(y \geq \tau \mid \mathbf{x}_1)/q_1^\theta(\mathbf{x}_1)$，回放为$w(\mathbf{x}_1) = 1$。这种按分量的方案在显著降低计算开销的同时保持一致性。

4. 相关工作

ProGen2（Nijkamp 等人，2023年）等自回归模型和 Tranception（Notin 等人，2022）在适应度预测方面表现出色，但其从左到右的因子分解限制了上位性耦合和灵活的条件化。离散扩散通过并行去噪来解决这一问题：D3PMs（Austin 等人，2021）将扩散推广到有限状态空间；MDLM（Sahoo 等人，2024）将吸收态扩散与掩码语言建模联系起来；EvoDiff（Alamdari 等人，2023年）则实现了无需结构信息、以基序为条件的蛋白质生成。离散流匹配（Gat 等人，2024；Campbell 等人，2024）通过条件端点分布$q_\phi(\mathbf{x}_1 \mid \mathbf{x}_t, t)$学习离散空间上的连续概率流。关键在于，所有并行生成模型都是隐式的：精确的边缘概率评估需要对指数级数量的路径求和（Austin 等人，2021），这促使我们将其重新表述为可处理的条件分布。

主动生成方法在测量数据上训练代理模型 $f_\psi(\mathbf{x})$，并搜索高分设计（Yang 等人，2019年）。两种变分框架直接以 $p(\mathbf{x} \mid y \geq \tau)$为目标：CbAS（Brookes 等，2019）通过重要性加权再训练最小化正向 KL；VSD（Steinberg 等人，2025年）使用得分函数梯度最小化反向 KL。两者都需要易于处理的 $q_\phi(\mathbf{x})$——用于密度比（CbAS）或得分梯度（VSD）——而隐式流模型并不具备这一点。我们通过在 $q_t^\phi(\mathbf{x}_1 \mid \mathbf{x}_t)$上重新表述目标函数来扩展这些原则。LaMBO-2（Gruver 等人，2023年）则在主动生成设定中，使用期望改进（EI）或期望超体积改进（EHVI）采集函数来引导扩散生成骨干网络。

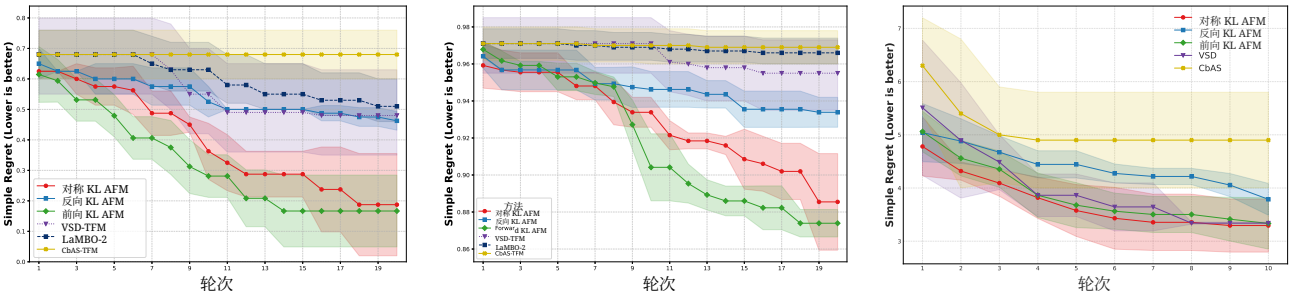
分类器（Dhariwal &Nichol，2021）和无分类器（Ho& Salimans，2022）引导通过得分扰动来引导连续扩散。离散扩展需要松弛：Gumbel-Softmax（Jang 等，2017）支持可微采样，但会产生有偏梯度；直通估计器（Bengio 等，2013）让梯度通过离散操作传播；基于连续时间马尔可夫链的引导（Nisonoff 等，2025）利用单词元转移。AFM 在根本上有所不同：我们将适应度纳入条件终点上的训练目标函数，从而无需松弛或修改采样过程即可获得精确的离散样本。

5. 实验

我们在五个蛋白质设计任务和一个小分子任务上评估 AFM，涵盖不同的序列长度、适应度景观和预言机特征。我们的实验旨在回答两个问题：（i）在无法获得易于处理的边缘似然时，能否引导隐式生成模型趋向高适应度区域？（ii）AFM 的前向 KL、反向 KL 和对称 KL 变体在性能上有何差异？

5.1. 任务与指标

Ehrlich 合成景观（Stanton等人，2024）。Ehrlich 函数是在离散序列上程序化生成的闭式测试函数，用于模拟生物物理序列优化问题的关键几何性质，包括崎岖性、约束和上位性。每个实例通过汇总基序类约束的满足程度，为有效序列定义得分 $f(\mathbf{x}) \in [0,1]$；无效序列则得到 $f(\mathbf{x}) = -1$。每个实例都存在一个



*(a)*Ehrlich景观， $L = 32$。 *(b)*Ehrlich景观， $L = 64$。 *(c)*AAV景观（$L = 28$）。

图 *1.* 预算约束 优化在 Ehrlich 合成景观和 AAV 衣壳设计任务上的表现。我们报告了简单的 regr et

已知全局最优$\mathbf{x}^\star$，且 $f(\mathbf{x}^\star) = 1.0$。我们以固定长度 $L \in \{32, 64\}$ 实例化 Ehrlich 函数。我们在包含 5,000 条无标签有效序列的池上进行预训练，以学习序列空间的可行域。我们执行 $R = 20$ 轮优化，批量大小为 $B = 128$。这些任务使我们能够在已知全局最优的受控环境中比较不同方法。

腺相关病毒 (AAV) (Bryant等人，2021)。我们还按照 Stein- berg等人（2025年）的黑盒优化设定，评估 AAV 衣壳设计任务的性能。目标是最大化一个 28 个氨基酸的序列（20^{28} 搜索空间）的适应度，该序列对应于参与病毒包装的 AAV衣壳蛋白区域。我们使用 Kirjner等人 （2024）提供的卷积神经网络（CNN）预言机作为真实适应度函数。已知的全局最大值 $y^* \approx 19.54$ 使得能够测量遗憾值。我们进行优化，持续 $R = 10$ 轮次，批量大小为 $B = 128$。

基于结构的蛋白质设计。 我们考虑一个具有两个目标函数的基于结构的蛋白质优化任务：(i) 提高热力学稳定性（$-\Delta G$）以及 (ii) 增加溶剂可及表面积（SASA）。黑盒预言机为 FoldX 分子模拟软件（Schymkowitz等人，2005年），由 POLI 库（Gonz´alez-Duque等人，2024）封装。我们使用 mRouge 红色荧光蛋白（$L = 228$）作为基础蛋白质。我们进行优化，持续 $R = 10$ 轮次，批量大小为 $B = 8$，并且每轮最多允许三处突变。

分子对接（F2/凝血酶）。我们使用 Dockstring 基准测试（Garc´ıa-Orteg´on 等人，2022），评估以凝血酶（F2）为目标的分子对接，分子以 SELFIES 表示（Krenn 等人，2020）。我们优化 $f(\mathbf{x}) = s_{\text{dock}}(\mathbf{x}) - 7.5(1 - \text{定量类药性估计}(\mathbf{x}))$，将对接分数与类药性惩罚相结合（Bickerton等人，2012）。从 ∼260k 个具有预先计算分数的分子开始，每种方法每轮提出 $B = 32$ 个分子，共进行 $R = 100$ 轮（3,200 次预言机调用）。我们将其与 VSD 进行比较；我们未能在此任务上稳定 CbAS。

评估指标。所有结果均为五次独立运行的平均值（Dockstring 为三次）。对于已知全局最优值的任务（Ehrlich、AAV），我们报告简单后悔值：$\text{reg}_r = f(\mathbf{x}^\star) - \max_{s \le r} \max_{\mathbf{x} \in \mathcal{B}_s} f(\mathbf{x})$（越低越好）。对于最优值未知的任务（FoldX、Dockstring），我们报告找到的最佳分数： $I_r = \max_{s \le r} \max_{x \in \mathcal{B}_s} f(\mathbf{x})$（越高越好）。

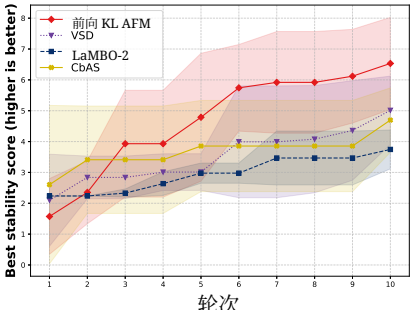
5.2. 基线与实现

我们将三种 AFM 变体（前向 KL、反向 KL、对称 KL）与三种基线方法进行比较：VSD（Steinberg 等人，2025年）、CbAS（Brookes 等人，2019年）和 LaMBO- 2（Gruver 等人，2023年）。VSD 最小化基于反向 KL 散度的目标函数，而 CbAS 最小化前向 KL；两者都使用具有可处理似然的自回归模型。LaMBO-2 使用离散扩散，并在去噪网络的潜在空间中通过基于梯度的引导，将生成引向高适应度区域。对于 VSD 和 CbAS，我们采用Steinberg 等人(2025年)[1] 公开实现中的 Transformer 主干网络；他们证明，这种网络优于 LSTM 对应模型。对于 LaMBO-2，我们使用POLI库提供的官方实现。[2] 在突变约束设置下，LaMBO-2 提供了内置功能，可控制生成过程中的编辑操作数量。对于 VS D 和 CbAS，我们遵循（Steinberg 等人，2025年）使用掩码 Transformer（mTFM）主干网络，该网络天然支持编辑操作受限的生成。AFM 的超参数选择见附录A.2。

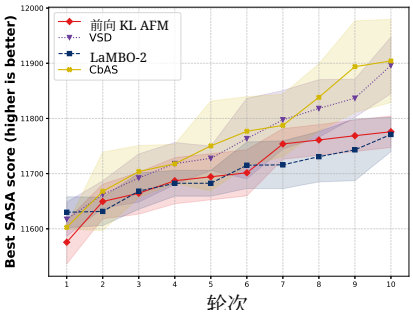
5.3. 结果

在 Ehrlich-32 和 Ehrlich-64 上，正向 KL 主动流匹配收敛到接近最优解的速度最快，对称 KL

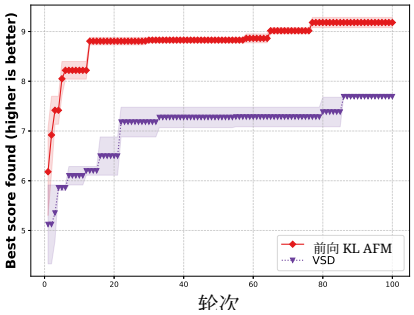
[1]https://github.com/csiro-funml/ 变分搜索[2] https://github.com/prescient-design/ 皮层



(a) FoldX 稳定性。



(b) FoldX 溶剂可及表面积。



(c) F2 分子对接。

图 *2.* 使用 FoldX 预言机进行基于结构的蛋白质设计。我们绘制了找到的最佳得分与预言机调用次数之间的关系；

主动流匹配紧随其后，而反向 KL 主动流匹配则落后于这两种方法。VSD 的收敛速度较慢，且在较长序列（Ehrlich-64）上差距进一步扩大，这表明其难以捕捉长程上位性相互作用。LaMBO-2 在两种序列长度下的收敛速度也都较慢，表现不佳。CbAS 表现出很强的利用倾向，未能探索最初发现的序列以外的区域，并过早收敛到次优区域。在 AAV 衣壳设计任务中，正向 KL 和对称 KL 主动流匹配均表现出色，取得了最低的遗憾值，VSD 紧随其后。反向 KL 主动流匹配在这项任务中表现不佳，收敛效果相对较差。鉴于反向 KL 和对称 KL 主动流匹配在 Ehrlich 和 AAV 任务中的表现相对较弱，我们在其余实验中聚焦于正向 KL 主动流匹配。在稳定性优化任务中，正向 KL 主动流匹配表现最佳，比其他方法更快地发现高稳定性蛋白质变体。然而，在溶剂可及表面积优化任务中，正向 KL 主动流匹配表现不佳，逊于 VSD 和 CbAS。在 F2 分子对接任务中，正向 KL 主动流匹配显著优于 VSD，在整个优化过程中取得了明显更高的对接分数。

6. 局限性

主动流匹配依赖于对高适应度概率 $p(y \geq \tau \mid \mathbf{x})$ 的准确估计，因此其性能对分类器质量十分敏感，尤其是在标注数据稀少的早期轮次。尽管自归一化重要性采样渐近地能得到一致估计量，但当提议分布与目标分布的重叠度较低时，有限样本方差可能很高，因此需要谨慎设计提议分布并确保良好的混合。此外，我们的反向 KL 散度表述缺乏与前向 KL 变体同等的理论一致性保证；实证结果也显示出寻模行为，可能导致过早收敛。

7. 结论

我们提出了主动流匹配，这是一种原则性框架，用于将隐式离散流模型与变分主动生成目标函数相结合。通过重新表述基于 KL 散度的优化准则，使其作用于流中易于处理的条件端点分布，AFM 无需访问难以处理的边缘似然，即可微调离散流，使其趋向高适应度区域。我们的前向 KL 散度表述具有一致性保证，并且在实证中展现出良好的探索—利用权衡；在评估预算有限的情况下，它在多种蛋白质设计基准测试中均优于最先进基线方法。

更广泛地说，AFM 弥合了现代非自回归生成模型与以概率论为基础的黑箱优化框架（如 VSD 和 CbAS）之间的差距。这为解决（Steinberg 等人，2025年）所阐述的主动生成问题提供了另一种方法。这一联系开辟了若干值得探索的方向，包括更有效的方差缩减策略、向多目标优化和约束优化的扩展，以及在生物序列设计之外的应用。我们认为，主动流匹配是朝着将表达能力强的隐式生成模型与严格的决策理论目标相统一迈出的一步，可用于离散高维空间中的黑箱优化。

参考文献

Alamdari, S., Thakur, N., van den Berg, R., Lu, A. X., Fusi, N., Louber, A. P. 和 Madani, A. 基于进化扩散的蛋白质生成：序列就是你所需要的一切。*bioRxiv*，第 2023.09.11.556673 页，2023年。

Austin, J., Johnson, D. D., Ho, J., Tarlow, D. 和 van den Berg, R. 离散状态空间中的结构化去噪扩散模型。*arXiv*，2021 年。

Bengio, Y.、L´eonard, N. 与 Courville, A. 估计

或通过随机神经元传播梯度以实现条件计算，2013。

Bickerton, G. R., Paolini, G. V., Besnard, J., Muresan, S. 和 Hopkins, A. L. 定量研究药物的化学美感。《自然·化学》，4(2):90–98，2012年。

Biophysics, S. 和 Core, B. 表面等离子体共振（SPR）费用，2024。

Brookes, D. H. 和 Listgarten, J. 通过自适应采样进行设计。*arXiv* 预印本 *arXiv:1810.03714*，2018年。

Brookes, D. H., Park, H. 和 Listgarten, J. 通过自适应采样条件化实现稳健设计。载于国际机器学习会议，第773–782页，2019年。

Bryant, D. H., Bashir, A., Sinai, S., Jain, N. K., Ogden, P. J., Riley, P. F., Church, G. M., Colwell, L. J. 和 Kelsic, E. D. 利用机器学习对AAV衣壳蛋白进行深度多样化。《自然·生物技术》，39:691–696，2021年。doi: 10.1038/s41587-020-00793-4。

Campbell, A., Benton, J., De Bortoli, V., Rainforth, T., Deligiannidis, G., and Doucet, A. 离散去噪模型的连续时间框架。神经信息处理系统进展，35:28266–28279，2022。

Campbell, A., Yim, J., Barzilay, R., Rainforth, T., and Jaakkola, T. 离散状态空间上的生成流：实现多模态流及其在蛋白质协同设计中的应用。*arXiv* 预印本 *arXiv:2402.04997*，2024。

Core, D. B. I. A. 预订、政策和费率，2024。表面等离子体共振/生物层干涉每小时费率；样品制备/数据分析费用。Dhariwal, P. and Nichol, A. 扩散模型在图像合成方面胜过生成对抗网络。载于神经信息处理系统进展，第34卷，第8780–8794页，2021。Garc´ıa-Orteg´on, M., Simm, G. N., Tripp, A. J., Hern´andez- Lobato, J. M., Bender, A., and Bacallado, S. DOCK- STRING：简便的分子对接可为配体设计提供更优的基准。*J. Chem. Inf. Model.*，62(15): 3486–3502，2022。

Gat, I.、Remez, T.、Shaul, N.、Kreuk, F.、Chen, R. T. Q.、Synnaeve, G.、Adi, Y. 与 Lipman, Y.。离散流匹配。载于神经信息处理系统进展，2024。

Gonz´alez-Duque, M.、Bartels, S. 与 Michael, R.。Poli：离散序列目标库。载于神经信息处理系统进展，第37卷，2024。

Gruver, N., Stanton, S., Frey, N. C., Rudner, T. G. J., Hotzel, I., Lafrance-Vanasse, J., Rajpal, A., Cho, K. 和 Wil- son, A. G. 使用引导式离散扩散进行蛋白质设计。收录于 神经信息处理系统进展，2023年。

Ho, J. 与 Salimans, T.。无分类器扩散引导。*arXiv*预印本 *arXiv:2207.12598*，2022。

Ingraham, J. B.、Baranov, M.、Costello, Z.、Barber, K. W.、Wang, W.、Ismail, A.、Frappier, V.、Lord, D. M.、Ng-Thow-Hing, C.、Vlack, E. R. V.、Tie, S.、Xue, V.、Cowles, S. C.、Leung, A.、Rodrigues, J. V.、Morales-Perez, C. L.、Ayoub, A. M.、Green, R.、Puentes, K.、Oplinger, F.、Panwar, N. V.、Obermeyer, F.、Root, A. R.、Beam, A. L.、Poelwijk, F. J. 与 Grigoryan, G.。用可编程生成模型照亮蛋白质空间。《自然》，623(7988):1070–1078，2023年。

Jang, E., Gu, S., and Poole, B. 使用 Gumbel-softmax 进行分类分布重参数化。载于国际学习表征会议，2017。

Kirjner, A., Yim, J., Samusevich, R., Bracha, S., Jaakkola, T. S., Barzilay, R., and Fiete, I. R. 通过平滑适应度景观改进蛋白质优化。载于国际学习表征会议，2024。

Krenn, M., H¨ase, F., Nigam, A., Friederich, P., and Aspuru-Guzik, A. 自引用嵌入字符串（SELFIES）：一种 100% 稳健的分子字符串表示。《机器学习：科学与技术》，1(4):045024，2020。Lipman, Y., Chen, R. T. Q., Ben-Hamu, H., Nickel, M., and Le, M. 用于生成建模的流匹配。*arXiv*，2022。Lipman, Y., Havasi, M., Holderrieth, P., Shaul, N., Le, M., Karrer, B., Chen, R. T. Q., Lopez-Paz, D., Ben-Hamu, H., and Gat, I. 流匹配指南与代码，2024。

Nijkamp, E., Ruffolo, J. A., Weinstein, E.N., Naber, C., and Madani, A. Progen2：探索蛋白质语言模型的边界。*Cell Systems*，14(11):968–978，2023年。

Nisonoff, H., Xiong, J., Alber, S., and Listgarten, J. 解锁离散状态空间扩散与流模型的引导方法。载于国际学习表征会议，2025。

Notin, P., Dias, M., Frazer, J., Marchena-Hurtado, J., Gomez, A., Marks, D., and Gal, Y. Tranception：结合自回归变换器与推理时检索的蛋白质适应度预测。载于国际机器学习会议，第16990–17017页。PMLR，2022。

Phillips, P. C. 上位性——基因相互作用在遗传系统的结构与演化中发挥的关键作用。《自然综述：遗传学》，9(11):855–867，2008。

Pooladian, A.-A., Ben-Hamu, H., Domingo-Enrich, C., Amos, B., Lipman, Y., 和 Chen, R. T. 多样本流匹配：利用小批量耦合使流变直。*arXiv* 预印本 *arXiv:2304.14772*，2023年。

Sahoo, S. S., Arriola, M., Schiff, Y., Gokaslan, A., Marro- quin, E., Chiu, J. T., Rush, A., and Kuleshov, V. 简单有效的掩码扩散语言模型。载于神经信息处理系统进展，第37卷，2024。Schymkowitz, J., Borg, J., Stricher, F., Nys, R., Rousseau, F., and Serrano, L. FoldX网络服务器：一个在线力场。《核酸研究》，33:W382–W388，2005。Stanton, S., Alberstein, R., Frey, N., Watkins, A., and Cho, K. 生物物理序列优化算法的闭式测试函数，2024。Starr, T. N. and Thornton, J. W. 蛋白质演化中的上位性。《蛋白质科学》，25(7):1204–1218，2016。Steinberg, D. M., Oliveira, R., Ong, C. S., and Bonilla, E. V. 变分搜索分布。载于国际学习表征会议，2025年。Steinberg, D. M., Wijesinghe, A., Oliveira, R., Koniusz, P., Ong, C. S., and Bonilla, E. V. 帕累托集的摊销主动生成。arXiv:2510.21052，2025年。Watson, J. L., J uergens, D., Bennett, N. R., Trippe, B. L., Yim, J., Eisenach, H. E., Ahern, W., Borst, A. J., Ragotte, R. J., Milles, L. F., Wicky, B. I. M., Hanikel, N., Pellock, S. J., Courbet, A., Sheffler, W., Wang, J., Venkatesh, P., Sappington, I., Torres, S. V., Lauko, A., Bortoli, V. D., Mathieu, E., Ovchinnikov, S., Barzilay, R., Jaakkola, T. S., DiMaio, F., Baek, M., and Baker, D. 利用 RFdiffusion 从头设计蛋白质结构与功能。《自然》，620(7976):1089–1100，2023。Yang, K. K., Wu, Z., and Arnold, F. H. 机器学习引导的蛋白质工程定向进化。《自然·方法》，16(8):687–694，2019。

A. 附录。

A.1. 定理的证明3.1

定理 3.1。 设 $p^*(\mathbf{x}) \propto p_{\text{prior}}(\mathbf{x})w(\mathbf{x})$ 为目标分布，其中 $w(\mathbf{x}) = p(y \geq \tau \mid \mathbf{x})$ 且 p_{prior} 为均匀分布。在标准 *DFM* 假设（掩码源耦合、凸插值路径和严格为正的调度器）下，前向 *KL* 主动流匹配目标函数的全局极小值点 ϕ^* 所得到的终端分布 $q_1^{\phi^*} = p^*$ 几乎处处成立。

证明。 证明分为两个阶段：首先，我们证明前向 KL 主动流匹配目标函数等价于在目标分布 p^* 上训练的标准离散流匹配（DFM）目标函数。其次，我们借助已有的 DFM 一致性结果，证明这将导向正确的终端分布。

1. 目标函数等价性。

前向 KL 主动流匹配目标函数（正文中的方程 10）用于估计先验分布上的以下期望损失：

$$\mathcal{L}_{\text{AFM}}(\phi) = \mathbb{E}_{t\sim\mathcal{U}[0,1]}\mathbb{E}_{\mathbf{x}_1\sim p_{\text{prior}}}\left[w(\mathbf{x}_1)\cdot \ell_\phi(\mathbf{x}_1, t)\right], \tag{22}$$

其中，在$\mathbf{x}_0 = m$ 确定的掩码源耦合下，

$$\ell_\phi(\mathbf{x}_1, t) = \mathbb{E}_{\mathbf{x}_t\sim p_t(\cdot|m,\mathbf{x}_1)}\left[-\log q_\phi(\mathbf{x}_1 \mid \mathbf{x}_t, t)\right] \tag{23}$$

是单个数据点的标准流 匹配损失，针对时刻 t 的单个数据点$\mathbf{x}_1$ 。我们可以将对$\mathbf{x}_1$ 的外层期望重写为：

$$J(\phi) = \sum_{\mathbf{x}_1\in\mathcal{X}} p_{\text{prior}}(\mathbf{x}_1)w(\mathbf{x}_1)\ell_\phi(\mathbf{x}_1), \tag{24}$$

其中，我们已将时间期望并入 ℓ_ϕ。

根据定义，目标分布为 $p^*(\mathbf{x}_1) = \frac{1}{Z}p_{\text{prior}}(\mathbf{x}_1)w(\mathbf{x}_1)$，这意味着恒等式 $p_{\text{prior}}(\mathbf{x}_1)w(\mathbf{x}_1) = Zp^*(\mathbf{x}_1)$ 成立。将其代入目标函数：

$$J(\phi) = \sum_{\mathbf{x}_1\in\mathcal{X}} Zp^*(\mathbf{x}_1)\ell_\phi(\mathbf{x}_1) = Z\cdot\mathbb{E}_{\mathbf{x}_1\sim p^*}\left[\ell_\phi(\mathbf{x}_1)\right]. \tag{25}$$

由于 $Z = \mathbb{E}_{p_{\text{prior}}}[w(\mathbf{x})]$ 是一个与 ϕ 无关的常数，因此最小化 $\mathcal{L}_{\text{AFM}}(\phi)$ 等价于最小化标准 DFM目标 $\mathcal{L}_{\text{DFM}}(\phi; p^*) = \mathbb{E}_{\mathbf{x}_1\sim p^*}[\ell_\phi(\mathbf{x}_1)]$。

实现说明：在算法1中，我们使用混合提议分布 $\mu(\mathbf{x}_1)$ （方程15）来近似该目标函数。通过使用来自混合分布的重要性权重，经验估计量是 $J(\phi)$的一致估计量，从而保留了此处推导出的理论最小值。

*SNIS*一致性。 在实践中，我们使用自归一化重要性采样，这会引入阶为$O(1/K)$的有限样本偏差。然而，SNIS 是一致的：当 $K\to\infty$时，估计量收敛到真实期望，并且经验目标的极小值点收敛到 ϕ^*。

2. 极小值点的一致性。

既然已经确定我们实际上是在最小化标准离散流匹配目标函数，其所对应的目标分布为 p^*，该结果便由Gat 等人（2024）中建立的一致性保证得出：

- 最优后验恢复（命题5）： 交叉熵损失 $\mathcal{L}_{\text{DFM}}(\phi; p^*)$的全局极小值 ϕ^* 恢复了目标分布的真实词元级条件后验：$p_{\phi^*}(x_1^i \mid \mathbf{x}_t, t) = p^*(x_1^i \mid \mathbf{x}_t, t)$，对每个词元 $i \in [N]$成立。
- 条件速度（定理 3）： 对于凸插值路径族，由该后验分布通过方程 5 构造的概率速度场生成条件概率路径 $p_t(\cdot \mid m, \mathbf{x}_1)$。
- 边缘速度（定理 2）： 边缘速度场通过在后验分布 $p_t(\mathbf{x}_0, \mathbf{x}_1 \mid \mathbf{x}_t)$ 下对条件速度取平均而得到，并生成边缘概率路径 p_t^*。

- 终端分布：迭代应用采样规则 $x^i_{t+h} \sim \delta_{\mathbf{x}^i_t}(\cdot) + h\, u^i_t(\cdot, \mathbf{x}_t)$，即可从 p^*_t 中生成样本。根据插值路径的边界条件，$p^*_1 = p^*$，因此终端分布满足 $q_1^{\phi^*} = p^*$。

因此，生成分布几乎处处与目标分布 p^* 匹配。 □

A.2. 实现细节

A.2.1. EHRLICH 任务超参数

我们采用Stanton等人（2024）提供的实现。任务定义的序列长度为 $L = 32$ （或 $L = 64$），字母表大小为 $|\mathcal{V}| = 20$。

表 *1.* Ehrlich实验的超参数和动态调度。

参数	值／调度
生成流与分类器架构	
架构	双向长短期记忆网络
隐藏维度	256
嵌入维度	128
层	3
丢弃	0.1
时间调度器	二次（t^2）
主动生成循环（第 r 轮次）	
每轮梯度步数	2，000
优化器	AdamW
学习率	$1 \times 10^{-4} \times 0.95^r$
阈值 τ	90^{th} 百分位数（动态）
混合系数（第 r 轮）	
回放（α_{rb}）	$\min(0.4, 0.05 \times r)$
流（α_{flow}）	$(1 - \alpha_{\text{rb}}) \times 0.1$（早期） $(1 - \alpha_{\text{rb}}) \times 0.95$（晚期）
先验（α_0）	$1 - \alpha_{\text{rb}} - \alpha_{\text{flow}}$
缓冲区温度（γ）	0.3

提议调度。 重要性采样提议分布 $\mu(\mathbf{x}_1)$ 会在两种机制下动态调整。在早期轮次，当回放缓冲区较小时，我们主要从先验中采样，以促进广泛探索。一旦回放缓冲区超过批量大小，流分量便占据主导地位（非回放概率的 $\alpha_{\text{flow}} \approx 95\%$），从而能够在有前景的区域进行局部细化。在整个训练过程中，回放缓冲区的占比（α_{rb}）从 0% 线性增加至 40%，逐步将重点转向利用已确认的高适应度序列。

A.3. FoldX 任务超参数

FoldX 任务采用与 Ehrlich 相同的生成流和分类器架构（表 1），序列长度为 $L = 228$，字母表大小为 $|\mathcal{V}| = 21$（20 种氨基酸加一个填充标记）。

源分布。 与 Ehrlich 任务不同，在 FoldX 任务中，我们限制流可以对源序列进行的突变数量。流的源 x_0 从回放缓冲区采样，而不是从均匀先验采样。这使得训练聚焦于已知高适应度区域周围的局部细化。

约束采样。 我们通过修改欧拉采样器来跟踪相对于参考序列 x_{ref} 的编辑操作，从而实施突变预算 $K \leq 3$ 。每一步都按预测速率的比例对位置和目标标记进行采样，

表 2. FoldX 实验的超参数。未列出的参数与表 1相同。

参数	值／调度
混合系数（第 r 轮）	
回放（α_{rb}）	0.3（固定）
流（α_{flow}）	$(1-\alpha_{\text{rb}}) \times 0.15$ （$r < 5$）
	$(1-\alpha_{\text{rb}}) \times 0.35$ （$r \geq 5$）
先验（α_0）	$1 - \alpha_{\text{rb}} - \alpha_{\text{flow}}$
缓冲区温度（γ）	0.2
突变预算 K	3

同时屏蔽自转移和填充词元。突变达到 K 次后，或不再存在有利转移时，流程即告终止。

A.4. AAV 衣壳设计超参数

对于 AAV 任务（$L = 28, |\mathcal{V}| = 20$），我们在生成流和分类器中都采用基于 Transformer 的架构。真实适应度使用预训练的卷积神经网络（CNN）预言机进行评估。

预言机详情。适应度景观由一个 CNN 预言机定义，该预言机使用来自Bryant 等人（2021）的数据集进行训练。为便于评估，分数将反归一化至原始尺度。

提议策略。与 Ehrlich 任务不同，AAV 实验在整个优化过程中使用固定的重放比率（$\alpha_{\text{rb}} = 0.3$）。与 Ehrlich 任务一样，一旦回放缓冲区累积了足够的样本（非回放概率为 $\alpha_{\text{flow}} \approx 66\%$），流分量便会占据主导地位。

A.5. Dockstring 分子对接超参数

对于 Dockstring 任务，分子以 SELFIES 字符串表示，词汇表大小为 $|\mathcal{V}| \approx 100$（因数据集而异）。生成流采用基于 LSTM 的架构，CPE 则采用 LSTM 分类器。

目标函数。遵循 Garc´ıa-Orteg´on 等人（2022），我们通过加入定量类药性估计惩罚项来优化对接分数：$f(\mathbf{x}) = s_{\text{dock}}(\mathbf{x}) - \lambda(1 - \text{定量类药性估计}(\mathbf{x}))$ 其中 $\lambda = 7.5$。

提议分布调度。与蛋白质任务不同，分子对接任务在整个优化过程中采用以回放为主的策略。回放缓冲区以初始池中排名前 2,500 的分子作为种子。回放的贡献比例在 100 轮次内从 95% 降至 70%，而流的贡献比例则相应地从 10% 升至 40%。这一保守的调度方案反映了分子优化更大的搜索空间和更稀疏的奖励景观。

表3. AAV衣壳优化的超参数。

参数	值／调度
生成式流模型	
架构	Transformer
层	5
注意力头	8
隐藏维度	256
丢弃	0.1
分类器模型	
架构	Transformer 编码器
层	3
注意力头	8
嵌入维度	128
前馈维度	512
丢弃	0.25
主动生成循环（轮次 r）	
梯度步数	1,000
优化器	AdamW
学习率	$1 \times 10^{-4} \times 0.95^{r}$
阈值 τ	13.5（初始），14.0（$r \geq 2$） 15.0（$r \geq 5$），15.5（$r \geq 6$）
混合系数（轮次 r）	
回放（α_{rb}）	0.3（固定）
流（α_{flow}）	$(1 - \alpha_{\text{rb}}) \times 0.35$（早期） $(1 - \alpha_{\text{rb}}) \times 0.95$（晚期）
先验（α_0）	$1 - \alpha_{\text{rb}} - \alpha_{\text{flow}}$
缓冲区温度（γ）	0.4

表 4. Dockstring（凝血因子II／凝血酶）实验的超参数 .

参数	值／调度
生成式流模型	
架构	双向长短期记忆网络
隐藏维度	512
嵌入维度	256
层	4
丢弃	0.1
时间调度器	二次（t^2）
分类器（*CPE*）模型	
架构	双向长短期记忆网络
隐藏维度	256
嵌入维度	128
层	3
丢弃	0.25
主动生成循环（第 r 轮）	
预训练步数	5，000
每轮梯度步数	1，000
优化器	AdamW
学习率（预训练）	1×10^{-3}
学习率（重要性加权）	1×10^{-4}
阈值调度	
初始 τ	80^{th} 百分位数
最终 τ	95^{th} 百分位数
混合系数（轮次 r, $R = 100$）	
回放（α_{rb}）	$0.95 \to 0.70$（线性）
流（α_{flow}）	$0.10 \to 0.40$（线性）
先验（α_0）	0。0（固定）
缓冲区温度（γ）	0.5
缓冲区大小	1,500