arXiv:2512.22597v1 [cs.LG] 27 Dec 2025

能量引导流匹配实现少步构象生成与基态识别

Guikun Xu[1], **Xiaohan Yi**[3], **Peilin Zhao**[1], ✉, **Yatao Bian**[2], ✉

[1]人工智能学院，上海交通大学，中国上海[2]计算机科学系，新加坡国立大学，新加坡[3]深圳国际研究生院，清华大学，中国深圳

{ : richxu945@sjtu.edu.cn

: peilinzhao@sjtu.edu.cn；ybian@nus.edu.sg

§ : https://github.com/Rich-XGK/EnFlow.git

摘要

利用基于物理的流程生成低能构象集合并识别基态构象，仍然需要大量计算。当前基于学习的方法往往存在范式割裂的问题：生成模型能够捕捉多样性，却缺乏可靠的能量校准；而确定性预测器以单一结构为目标，无法表征集合变异性。本文提出 **EnFlow**，这是一个统一框架，通过沿非高斯流匹配路径定义的能量引导采样方案，将流匹配（FM）与显式学习的能量模型相结合。在采样过程中融入能量梯度引导后，我们的方法能够将轨迹引向能量较低的区域，显著提升构象保真度，尤其是在少步情形下。学习到的能量函数还支持对生成的系综进行高效的基于能量的排序，从而准确识别基态。在 GEOM-QM9 和 GEOM-Drugs 上开展的大量实验表明，与最先进的方法相比，**EnFlow** 在仅需 1–2 个常微分方程步的情况下，同时提升了生成指标并降低了基态预测误差。

1 引言

给定一个指定原子和键类型的分子拓扑图，其原子的三维坐标定义了可能的分子构象集合。这些构象对于理解化学属性和生物功能至关重要 [1, 2]。识别能量有利的构象是计算化学和药物发现中的核心问题 [3]。传统方法，包括分子动力学模拟 [4, 5, 6] 和密度泛函理论优化 [7] ，计算量大且耗时，因此难以应用于大规模研究。在低能结构中，基态构象是势能面的全局最小值 [8] ，因而具有最高的热力学稳定性。它在决定结合亲和力、反应机理和光谱属性方面发挥着关键作用。因此，准确识别基态构象对于可靠预测和合理的分子设计至关重要。

近年来，人们持续致力于利用机器学习 [9, 10,11], ，尤其是深度生成建模 [12, 13, 14, 15, 16, 17, 18], ，以更高的计算效率 [14, 15, 16, 17, 18, 19]解决构象生成问题。尽管早期的基于学习的方法往往表现不佳 [20, 21, 22, 23, 24, 25],

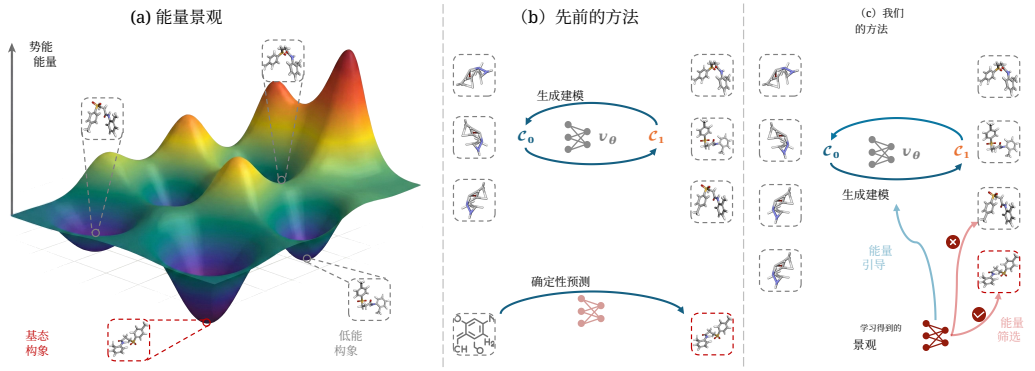


图 1：(a) 分子构象的势能景观：低能构象分布在局部极小值处，而基态构象对应全局最小值。(b) 现有生成方法能够采样出多样的低能构象，但难以可靠地识别基态；而确定性方法只预测一个接近基态的结构，无法捕捉集合变异性。(c) 我们的**EnFlow** 框架将这两种能力融为一体：通过利用一种学习得到的显式能量函数，同时引导生成过程，高效生成低能构象系综，并准确识别基态构象。

GeoDiff [26] 提出了针对分子结构的规范等变扩散 [14, 16, 27, 28] 公式。在此基础上，流匹配 [17]逐渐成为扩散模型 [16] 的一种替代方案，并展现出更高的采样效率和更快的生成速度 [17, 18]。后续研究进一步改进了流匹配流程，以在构象保真度与计算成本之间实现更有利的平衡 [29, 30]。

尽管在识别给定分子图的低能构象方面已取得显著进展，但要确定全局最小值，即基态构象，仍然颇具挑战。为应对这一挑战，近期研究采用确定性预测方法直接对基态进行建模，从而减少对随机采样及后续排序的依赖。在这些方法中， [31, 32, 33, 34, 35], 确定性神经网络，尤其是经过特定任务改造的图 Transformer [36] ，利用数据集进行训练，从中回归基态构象；这些数据集为每个分子提供一个基态标签。

如图 1(b) 所示，目前仍缺少一种实用方法，能够在单一统一框架内联合建模低能构象系综和特定基态。生成方法 [26, 29] 可以采样多样化的结构，但无法明确、校准地呈现能量景观，因此基态识别并不可靠。相比之下，预测方法 [32, 35], 会预测一个旨在近似基态的单一结构，因此无法捕捉集合变异性或量化不确定性。

为应对上述挑战，我们提出 **EnFlow**，如图 1(c) 所示。这是一个用于生成低能分子构象系综的统一框架，同时能够以经验证的快速学习方法认证基态构象。图 2 概述了 EnFlow：

(1) 提出了一种能量引导的流匹配框架，以便在非高斯流匹配路径上实现有原则的引导，为上述构象问题提供统一解决方案。不同于针对高斯路径流匹配和扩散模型开发的传统引导技术 [37, 38, 39], ，此类方法无法直接扩展到非高斯设置。在构象建模中，先验分布通常显著偏离高斯先验（例如谐波先验 [40]），这反映了对于捕捉分子结构至关重要的强领域特定结构归纳偏置，并因此诱导出与高斯路径机制存在根本差异的流动力学。(2) 构建了一种基于能量的引导机制，将能量信息直接纳入流动力学。因此，样本会被引导至低能量区域，从而显著提高样本保真度，尤其是在少步常微分方程采样下。(3) 该框架的核心组成部分是一个对构象能量景观进行建模的显式能量函数。该函数采用基于能量的

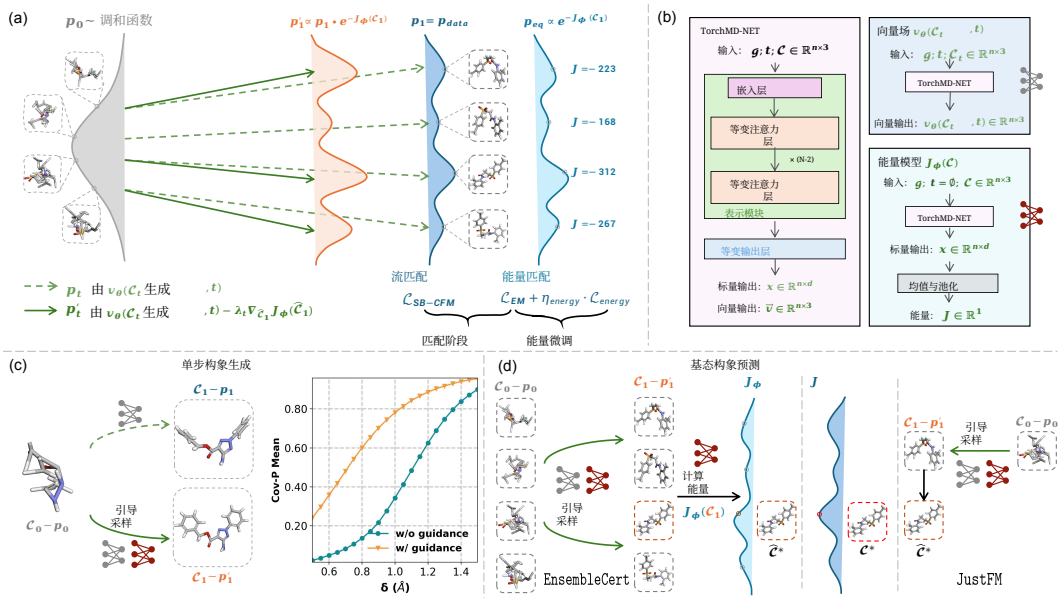


图2：**EnFlow** 示意图。(a)能量引导流匹配框架，其中通过能量匹配技术训练的能量基模型在流匹配过程中提供引导。(b) 向量场和能量模型的架构概览。为保证比较公平，我们采用了与 ET-Flow 相同的骨干架构 [29] (c) 能量引导采样实现改进的单步常微分方程采样示意图。(d) 使用经过充分训练的能量函数验证基态构象。

模型（能量基模型）[41], ，通过近期提出的能量匹配目标 [42]进行训练。除了捕捉全局能量结构外，能量基模型还经过进一步微调，以匹配分子级经验能量分布，从而解析同一分子不同构象之间细微的能量变化。(4) 除了改进系综生成，将强大的生成器与良好对齐的能量基模型相结合，还赋予 EnFlow 一个显式且经过校准的能量函数。这使得基态识别可以通过两种方式高效实现：按照预测能量对生成的构象进行排序，或通过所提出的能量引导采样过程直接进行一次性生成，为每个分子生成单个构象。在本工作中，所提出的框架采用 ET-Flow [29]实现，并以谐波先验 [40] 作为生成骨干，将谐波样本输运为物理上合理的分子构象。

GEOM-QM9 和 GEOM-Drugs 上的实证实验表明，在小规模和大规模设置下，EnFlow 在仅需 1–2 个常微分方程步的构象系综生成以及基态构象预测方面均达到最先进性能，显著优于现有方法。EnFlow 可作为量子力学的快速、可微替代模型，为具备生物学认知且热力学一致的分子设计铺平道路。

2 结果

尽管以往方法通常将构象系综生成或基态认证作为彼此独立的任务加以处理，但我们的方法在单一统一框架内将这两项任务结合起来。我们在 GEOM-QM9 和 GEOM-Drugs 数据集上报告了全面的实验结果 [43], 表明，所提出的 **EnFlow** 能够在两项任务上同时取得优异性能，从而凸显了所提出统一方法的有效性与实用性。如图 3 所示，在 GEOM-Drugs 上，与大多数现有方法相比，EnFlow 在仅使用相对较少神经网络参数的同时，首次在基态构象预测方面取得优异性能，并保持了均衡的系综生成能力。第 2.3 节给出了构象生成的详细结果，表明我们的方法在保持竞争力的同时，

在少步常微分方程采样下实现了显著改进。第 2.4 节表明，学习到的能量模型能够从生成的系综中有效识别最低能量构象，并将其作为预测的基态结构，从而带来显著的性能提升。第 2.5 节进一步提供了严谨分析，揭示学习到的能量函数能够主动引导构象趋向低能量区域，凸显了其在我们统一框架中的双重作用：实现引导生成，并作为用于基于能量的筛选的替代化学模型。更多实验与讨论见第 D.2 节。

2.1 符号与问题定义

主要符号： 三维分子形式化定义为 $\mathcal{M} := \{\mathcal{G}, \mathcal{C}\}$，其中 $\mathcal{G}$表示该分子的二维图表示，$\mathcal{C} \in \mathbb{R}^{n\times 3}$ 表示该分子在三维空间中的构象，具体包括每个原子的空间坐标。

问题定义： 分子构象生成任务旨在针对给定的分子图 $\mathcal{G}$，生成一组可能的低能稳定构象 $\mathcal{C}^1$、$\mathcal{C}^2$ 等。其目标是对条件分布 $p(\mathcal{C}|\mathcal{G})$进行建模。为简便起见，在以下各节中，我们使用记号 $p(\mathcal{C})$ 表示 $p(\mathcal{C}|\mathcal{G})$。分子基态构象预测任务旨在预测给定分子图 $\mathcal{G}$对应的最稳定构象 $\mathcal{C}^*$。

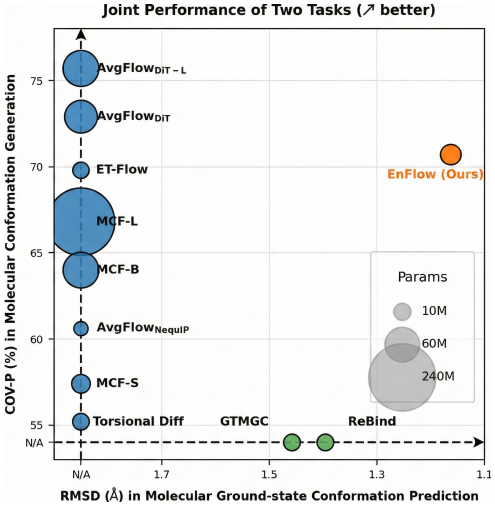


图 3：GEOM-Drugs 上分子构象生成与基态构象预测任务的联合表现。

2.2 数据集、指标和基线方法

数据集： 按照既有惯例，我们使用参考文献 [24, 44] 中定义的划分方案，在 **GEOM**数据集 [43] 上评估我们的方法。每个分子都关联有多个低能构象，每个构象均标注有玻尔兹曼能量及相应的玻尔兹曼权重。根据参考文献 [35], ，我们将每个分子中玻尔兹曼权重最高的构象定义为基态构象，并据此评估基态识别。GEOM-QM9 和 GEOM-Drugs 子集的测试集均包含 1,000 个分子。有关数据集的更多详情，请参见 A.2 节。

构象系综生成的指标： 此任务采用四种基于均方根偏差（RMSD）计算的成熟指标进行评估：分别针对召回率 和精确率 方案的覆盖率（COV）和平均最小均方根偏差（AMR），即 COV-R、AMR-R、COV-P 和 AMR-P。遵循以往研究 [24, 26, 45], ，设 R 为包含 K 个构象的参考集合，$\mathsf{R} = \{\mathcal{C}^1, \mathcal{C}^2, \ldots, \mathcal{C}^K\}$；G 表示由 $2K$ 个构象组成的生成集合，$\mathsf{G} = \{\hat{\mathcal{C}}^1, \hat{\mathcal{C}}^2, \ldots, \hat{\mathcal{C}}^{2K}\}$。召回率指标定义如下：

$$\mathrm{RMSD}(\mathcal{C}, \hat{\mathcal{C}}) = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \|\mathcal{C}_i - \hat{\mathcal{C}}_i\|^2}, \tag{1}$$

$$\mathrm{COV}(\mathsf{R}, \mathsf{G}) = \frac{|\{\mathcal{C}^k \in \mathsf{R} \mid \min_{\hat{\mathcal{C}}^j \in \mathsf{G}} \mathrm{RMSD}(\mathcal{C}^k, \hat{\mathcal{C}}^j) < \delta\}|}{|\mathsf{R}|}, \tag{2}$$

$$\mathrm{AMR}(\mathsf{R}, \mathsf{G}) = \frac{1}{|\mathsf{R}|}\sum_{\mathcal{C}^k \in \mathsf{R}} \min_{\hat{\mathcal{C}}^j \in \mathsf{G}} \mathrm{RMSD}(\mathcal{C}^k, \hat{\mathcal{C}}^j), \tag{3}$$

其中，δ 表示阈值，通常对于 GEOM-QM9 设为 0.5 埃，对于 GEOM-Drugs 设为 0.75 埃。覆盖率指标量化了参考构象在

R 中至少有一个生成构象位于阈值距离 δ 内的比例。AMR 指标计算 R 中每个参考构象与 G 中距离其最近的生成构象之间的平均最小均方根偏差。召回率指标评估它们的多样性，而精确率指标通过在上述公式中交换 R 和 G 的角色来获得，用于评估它们的质量。

基态构象预测的指标： 给定基态构象 $\mathcal{C}^*$ 及其对应的预测构象 $\hat{}\mathcal{C}^*$，根据已有研究 [32, 35], , 我们采用三个标准指标来评估预测性能： $\mathcal{C}$-RMSD、**D**-MAE 和 **D**-RMSE。首先，按照公式 1 计算预测构象与真实构象之间的 RMSD，得到 $\mathcal{C}$-RMSD。然后，令 $\mathbf{D}^{ij}$ 和 $\hat{}\mathbf{D}^{ij}$ 分别表示真实构象与预测构象的成对距离。这些距离之间的平均绝对误差（MAE）和均方根误差（RMSE）定义如下：

$$\mathbf{D}\text{-MAE}(\mathcal{C}^*, \hat{\mathcal{C}}^*) = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \left| \mathbf{D}_{ij}^* - \hat{\mathbf{D}}_{ij}^* \right|, \tag{4}$$

$$\mathbf{D}\text{-RMSE}(\mathcal{C}^*, \hat{\mathcal{C}}^*) = \sqrt{\frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\mathbf{D}_{ij}^* - \hat{\mathbf{D}}_{ij}^* \right)^2}. \tag{5}$$

基线方法： 对于分子构象生成，我们首先将方法的最佳配置与强基线方法进行比较，包括 GeoMol [24], GeoDiff [26],Torsional Diff [44], ET-Flow [45], MCF [46] 和 AvgFlow [30]。对于采用 1 步和 2 步 ODE 采样的生成，我们进一步与 MCF [46], ET-Flow [29] 和 AvgFlow [30]进行比较。对于基态构象预测，我们以 RDKit [47], GINE [48],GATv2 [49], GraphGPS [50], GTMGC [32], Torsional Diff [44] 和 ReBind [35], 作为基准，所有设置均遵循参考文献 Ref. [35]。为清晰起见，对于 ET-Flow [29] 类方法，ET-Flow 表示采用事后校正的模型，而 ET-Flow-$SO(3)$ 表示$SO(3)$-等变变体。遵循这一约定，我们将自己的模型称为**EnFlow** 和**EnFlow**-$SO(3)$。在 AvgFlow [30], 中，我们对几种 ODE 采样设置进行了基准测试，并采用 Reflow 和蒸馏 [18] 等技术来提升性能。因此，下标 {$_{\text{Reflow}}$ 和 {$_{\text{Distill}}$ 表示将这些技术应用于方法 {。

2.3 EnFlow 通过少量 ODE 步骤提升构象生成的保真度 s

针对低能构象集合生成任务，我们提出一种能量引导流匹配方案，将谐波先验 $p_0(\mathcal{C}_0)$ 输运至引导目标分布 $p_1'(\mathcal{C}) \propto p_1(\mathcal{C}_1)e^{-}J_\phi(c_1)$，而非原始数据分布 $p_1(\mathcal{C}_1)$。采用这一设计，从由 p_0 演化至 p_1' 的常微分方程轨迹视角来看，引导向量场 $v't(\mathcal{C}_t)$ 会受到抑制，不再朝高能区域移动。

GEOM-QM9 上的结果： **EnFlow** 在 GEOM-QM9 测试集上的构象生成 性能如表 1 所示。该表分别列出了公式 2 和公式 3 中定义的覆盖率和 AMR 指标，RMSD 阈值 δ 设为 0.5 埃。如表所示，EnFlow-$SO(3)$ 仅需*5* 个常微分方程采样步数 即可达到与先前基线方法相当的性能。采用 *50* 个常微分方程采样步数时，EnFlow-$SO(3)$ 在召回率指标上取得了与最佳 ET-Flow 相近的结果，同时在精确率指标上优于 ET-Flow；例如，其平均 COV-P 提高了 1.43%（与复现的 ET-Flow-$SO(3)^{\text{reproduced}}$相比则提高了 0.71%）。值得注意的是，仅采用*2* 个常微分方程采样步数时，EnFlow-$SO(3)$ 在召回率导向的 COV-R 和 AMR-R 指标上仍保持了相当的性能，同时在精确率导向的 COV-P 和 AMR-P 指标上取得了显著提升；例如，与 ET-Flow-$SO(3)$相比，平均 COV-P 提高了 2.10%，平均 AMR-P 降低了 0.037 埃（24.34%）。此外，在仅有 *1* 个常微分方程采样步数的极端情况下，EnFlow 显著优于所有现有方法，尤其是无引导 ET-Flow。与此前最先进水平的 AvgFlow$_{\text{NequIP-D}}$相比，在召回率指标上，尽管 EnFlow-$SO(3)$ 的 COV-R 仍有约 5% 的差距，

表 1：GEOM-QM9（$\delta = 0.5$埃）上的分子构象生成结果。

方法	ODE 步数	COV-R ↑		AMR-R ↓		COV-P ↑		AMR-P ↓	
		mean	中位数	mean	中位数	mean	中位数	mean	中位数
CGCF	1000	69.47	96.15	0.425	0.374	38.20	33.33	0.711	0.695
GeoDiff	1000	76.50	**100.00**	0.297	0.229	50.00	33.50	1.524	0.510
GeoMol	-	91.50	**100.00**	0.225	0.193	87.60	**100.00**	0.270	0.241
扭转扩散	20	92.80	**100.00**	0.178	0.147	92.70	**100.00**	0.221	0.195
MCF	1000	95.0	**100.00**	0.103	0.044	93.70	**100.00**	0.119	0.055
ET-Flow	50	**96.47**	**100.00**	**0.073**	0.047	94.05	**100.00**	0.098	0.039
ET-Flow-$SO(3)$	50	95.98	**100.00**	0.076	0.030	92.10	**100.00**	0.110	0.047
ET-Flow已复现	50	95.81	**100.00**	0.076	0.030	92.30	**100.00**	0.105	0.035
ET-Flow-$SO(3)$reproduced	50	95.69	**100.00**	0.079	0.028	94.77	**100.00**	0.088	0.033
AvgFlowNequIP	50	96.40	**100.00**	0.089	0.042	92.80	**100.00**	0.132	0.084
EnFlow	5	95.83	**100.00**	0.082	0.032	92.67	**100.00**	0.108	0.044
EnFlow	50	95.74	**100.00**	0.078	0.030	92.59	**100.00**	0.100	0.036
EnFlow-$SO(3)$	5	96.11	**100.00**	0.078	0.033	95.11	**100.00**	0.090	0.041
EnFlow-$SO(3)$	50	96.26	**100.00**	0.076	**0.028**	**95.48**	**100.00**	**0.083**	**0.034**
两步生成									
ET-Flow复现	2	96.27	**100.00**	0.119	0.071	90.41	**100.00**	0.182	0.140
ET-Flow-SO（3）reproduced	2	**96.70**	**100.00**	0.112	0.067	93.54	**100.00**	0.152	0.110
AvgFlowNequIP-R	2	95.90	**100.00**	0.151	0.104	87.70	**100.00**	0.236	0.207
EnFlow	2	94.81	**100.00**	0.112	**0.055**	92.58	**100.00**	0.138	0.075
EnFlow-$SO(3)$	2	95.49	**100.00**	**0.108**	0.058	**95.64**	**100.00**	**0.115**	**0.066**
***1*步生成**									
ET-Flow复现	1	50.47	50.00	0.483	0.505	36.38	25.00	0.550	0.580
ET-Flow-$SO(3)$reproduced	1	76.12	**100.00**	0.330	0.323	67.93	87.50	0.382	0.387
AvgFlowN equIP-D	1	**95.1**	**100.00**	0.220	0.195	84.80	**100.00**	0.304	0.283
EnFlow	1	89.43	**100.00**	0.215	0.178	88.09	**100.00**	0.256	0.227
EnFlow-$SO(3)$	1	90.45	**100.00**	**0.195**	**0.150**	**91.23**	**100.00**	**0.213**	**0.177**
EnFlow-$SO(3)_{\text{Reflow}}$	1	96.70	100.00	0.122	0.087	93.48	100.00	0.170	0.132

表注。大多数基线方法的结果均报告于相应文献中。方法 ET-Flowreproduced 和 ET-Flow-$SO(3)$reproduced 表示我们使用官方 ET-Flow [29] 实现复现得到的结果。关于复现问题的详细讨论见 D.1.3节。对于两步生成和多步生成，最佳结果以**粗体**标出，次佳结果以下划线标出。对于***1*步生成**，**EnFlow**-$SO(3)_{\text{Reflow}}$ 表示在 EnFlow 上采用可选 Reflow [18] 技术后获得的性能；在未使用 Reflow 的方法中，最佳和次佳结果分别以**粗体**和下划线标出。

但其平均 AMR-R 降低了 0.025 埃（11.36%），中位数 AMR-R 降低了 0.045 埃（23.08%）；在精确率指标上，EnFlow-$SO(3)$ 将平均 COV-P 提高了 6.43%，并将平均 AMR-P 降低了 0.091 埃（29.93%），中位数 AMR-P 降低了 0.070 埃（24.73%）。这些结果证明，我们提出的能量引导流匹配方案能够有效提高分子构象生成中少量常微分方程步数采样的保真度。此外，对采用 1 个常微分方程步数的 EnFlow-$SO(3)$ 应用 Reflow 技术 [18] 后，性能进一步提升，在所有指标上均取得新的最先进水平结果；例如，平均 COV-P 提高至 93.48%，平均 AMR-P 降至 0.170 埃。除了在固定 RMSD 阈值下评估这些指标外，考察覆盖率指标 COV-R 和 COV-P 随 RMSD 阈值变化的曲线，也能提供更全面的评估。如图8所示，我们针对在 1、2 和 5 个常微分方程采样步数下是否使用能量引导（即 EnFlow-$SO(3)$：**使用引导**）开展了消融研究，并与不使用能量引导的情况（即 ET-Flow-$SO(3)$：**不使用引导**）进行比较。结果一致表明，采用我们的能量引导后，在较小的 RMSD 阈值下，覆盖率指标显著提升，尤其是在步数极少的情况下（即 1 步和 2 步）。

GEOM-Drugs 上的结果： **EnFlow** 在 GEOM-Drugs 上的性能总结于表2中，其中所有指标均在 RMSD 阈值为 $\delta = 0.75$ 埃的条件下进行评估。该表报告了召回率导向和精确率导向评估下的覆盖率与 AMR 指标。如表所示，在进行*50* 个 *ODE* 采样步数时，相较于现有的扩散方法和基于流的方法，EnFlow 达到了具有竞争力的性能。具体而言，其均值 COV-R 达到 78.8%，均值 AMR-R 降至 0.475 埃，优于复现的 ET-Flow复现结果。在精确率导向指标方面，EnFlow 同样保持了较强的结果，在 COV-P 和 AMR-P 两项指标上均优于 ET-Flow复现结果，例如将均值 AMR-P 从 0.604 埃降至 0.590 埃。仅使用*5* 个 *ODE* 采样步数时，EnFlow 仍然取得了

表2：GEOM-Drugs 上的分子构象生成结果（$\delta = 0.75$埃）。

方法	ODE 步数	COV-R ↑		AMR-R ↓		COV-P ↑		AMR-P ↓	
		mean	中位数	mean	中位数	mean	中位数	mean	中位数
GeoDiff	1000	42.10	37.80	0.835	0.809	24.90	14.50	1.136	1.090
GeoMol	-	44.60	41.40	0.875	0.834	43.00	36.40	0.928	0.841
扭转扩散	20	72.70	80.00	0.582	0.565	55.20	56.90	0.778	0.729
MCF - S (13M)	1000	79.4	87.5	0.512	0.492	57.4	57.6	0.761	0.715
MCF - B (64M)	1000	84.0	91.5	0.427	0.402	64.0	66.2	0.667	0.605
MCF - L (242M)	1000	**84.7**	**92.2**	**0.390**	**0.247**	66.8	71.3	0.618	0.530
ET-Flow (8.3M)	50	79.53	84.57	0.452	0.419	74.38	81.04	0.541	0.470
ET-Flow复现 (8.3M)	50	79.54	85.00	0.470	0.444	69.79	75.53	0.604	0.538
ET-Flow - SS（8.3M）	50	79.62	84.63	0.439	0.406	75.19	81.66	0.517	**0.442**
ET-Flow - $SO(3)$ (9.1M)	50	78.18	83.33	0.480	0.459	67.27	71.15	0.637	0.567
AvgFlow$_{\text{NequIP}}$ （470 万）	102	76.8	83.6	0.523	0.511	60.6	63.5	0.706	0.670
AvgFlow$_{\text{DiT}}$ （52M）	100	82.0	86.7	0.428	0.401	72.9	78.4	0.566	0.506
AvgFlow$_{\text{DiT-L}}$ (64M)	100	82.0	87.3	0.409	0.381	**75.7**	**81.9**	**0.516**	0.456
EnFlow (16.6M)	5	77.2	82.3	0.499	0.479	70.0	76.5	0.607	0.541
EnFlow (16.6M)	50	78.8	84.6	0.475	0.455	70.7	76.9	0.590	0.521
两步生成									
MCF-B (64M)	2	46.7	42.4	0.790	0.791	21.5	13.2	1.155	0.715
MCF-L (242M)	2	54.2	54.4	0.752	0.746	25.7	18.8	1.119	1.115
ET-Flow (8.3M)	2	73.2	76.6	0.577	0.563	63.8	67.9	0.681	0.643
ET-Flow复现(8.3M)	2	72.3	76.9	0.592	0.583	58.3	60.4	0.733	0.699
AvgFlow$_{\text{NequIP-R重流}}$(4.7M)	2	64.2	67.7	0.663	0.661	43.1	38.9	0.871	0.853
AvgFlow$_{\text{DiT-Reflow}}$ (52M)	2	**75.7**	**81.8**	**0.545**	**0.533**	57.2	59.0	0.748	0.705
EnFlow (16.6M)	2	70.7	74.6	0.596	0.578	**69.1**	**75.7**	**0.623**	**0.575**
***1*步生成**									
MCF-B (64M)	1	22.1	6.9	0.962	0.967	7.6	1.5	1.535	1.541
MCF-L（242M）	1	27.2	13.6	0.932	0.928	8.9	2.9	1.511	1.514
ET-Flow（8.3M）	1	27.6	8.8	0.996	1.006	25.7	5.8	0.939	0.929
ET-Flow复现 （8.3M）	1	14.0	00.0	1.116	1.142	10.7	00.0	1.122	1.125
AvgFlow$_{\text{NequIP-（Reflow+蒸馏）}}$ （4.7M）	1	55.6	56.8	0.739	0.734	36.4	30.5	0.912	0.888
AvgFlow$_{\text{DiT-（Reflow+蒸馏）}}$ （52M）	1	**76.8**	**82.8**	**0.548**	**0.541**	**61.0**	**64.0**	**0.720**	**0.675**
EnFlow （16.6M）	1	53.1	50.00	0.802	0.773	54.3	56.3	0.773	0.743
EnFlow$_{\text{Reflow}}$ （16.6M）	1	74.3	80.0	0.566	0.548	60.6	63.5	0.719	0.666

表注。大多数基线方法的结果均报告于相应文献中。*1*步生成 和 两步生成 下 ET-Flow 的结果引自参考文献 [30]。方法 ET-Flow$_{\text{reproduced}}$ 表示我们使用官方 ET-Flow [29] 实现复现的结果。关于复现问题的详细讨论见 D.1.3节。对于 两步生成 和 多步生成，最佳结果以 **粗体**显示，次佳结果加下划线。对于 *1*步生成，EnFlow-$SO(3)_{\text{Reflow}}$ 表示在 EnFlow 上采用可选 Reflow [18] 技术后获得的性能；在未采用 Reflow 的方法中，最佳和次佳结果分别以**粗体** 和下划线标出。

具有良好的召回率–精确率权衡：其平均 COV-R 达到 77.2%，平均 AMR-R 为 0.499 埃；在精确率方面，平均 COV-P 达到 70.0%，平均 AMR-P 为 0.607 埃，显著优于 GeoDiff、GeoMol 和 Torsional Diffusion 等经典基线方法。

在少步场景下，EnFlow 展现出显著优势。在采用 *2* 个 *ODE* 采样步数时，EnFlow 的平均 COV-P 达到 69.1%，平均 AMR-P 为 0.623 埃，在精确率导向指标上优于所有竞争方法，包括 ET-Flow 和 AvgFlow变体，同时保持了较高的平均 COV-R（70.7%）。在 *1*步生成场景下，EnFlow 相较于先前方法取得了明显改进：与 ET-Flow复现结果相比，EnFlow 将平均 COV-P 从 14.0% 提升至 54.3%，并将平均 AMR-P 从 1.122 埃降至 0.773 埃。进一步将 Reflow 技术应用于我们的模型（**EnFlow**$_{\text{Reflow}}$）后，性能得到进一步提升，接近最强的 AvgFlow变体：**EnFlow**$_{\text{Reflow}}$ 的平均 COV-R 达到 74.3%，平均 AMR-R 为 0.566 埃，同时平均 COV-P 为 60.6%，平均 AMR-P 为 0.719 埃，其性能与 AvgFlow$_{\text{DiT-(Reflow+蒸馏)}}$ 相当，且无需额外的蒸馏阶段，并且参数量显著更少（16.6M 对 52M）。总体而言，在 GEOM-Drugs 上，EnFlow 在标准生成和少步生成设置下均展现出具有竞争力或更优的性能，尤其是在精确率导向指标和极低步数的 ODE 采样方面优势明显。如图9所示，我们在 1、2、5 和 50 个 ODE 采样步数下，对无引导基线（ET-Flow；**无引导**）和我们的能量引导模型（**EnFlow**；**有引导**）进行了消融研究。在所有设置下，能量引导均能持续提升 COV-R 和 COV-P，尤其是在较小的 RMSD 阈值下。该收益在极低步数场景中最为显著（即，

表 3：GEOM-Drugs 上的基态构象预测结果（埃）。

方法	mode	步数	D-平均绝对误差 ↓	D-均方根误差 ↓	$\mathcal{C}$-均方根偏差 ↓
RDKit-DG	-	1	1.181	2.132	2.097
RDKit-ETKDG	-	1	1.120	2.055	1.934
GINE	-	1	1.125	1.777	2.033
GATv2	-	1	1.042	1.662	1.901
GraphGPS（随机游走）	-	1	0.879	1.399	1.768
GraphGPS（标签传播）	-	1	0.815	1.300	1.698
GTMGC	-	1	0.823	1.319	1.458
扭转扩散	-	20	0.959	1.648	1.751
ReBind	-	1	0.776	<u>1.283</u>	1.396
ET-Flow	-	5	0.844±0.009	1.491±0.020	1.633±0.011
EnFlow	JustFM	5	0.793±0.011	1.412±0.024	1.550±0.029
EnFlow	EnsembleCert; $M = 5$	5	0.745	1.361	1.422
EnFlow	EnsembleCert; $M = 5$	50	0.714	1.296	1.338
EnFlow	EnsembleCert; $M = 50$	5	<u>0.703</u>	1.331	<u>1.312</u>
EnFlow	EnsembleCert; $M = 50$	50	**0.644**	**1.263**	**1.163**
最先进水平提升 ↑	-	-	17.01%	1.56%	16.69%

表注释。扭转扩散的结果取自参考文献 [35]。ET-Flow 和 EnFlow (JustFM) 的结果报告为 10 次独立运行的均值 ± 标准差，其中每次运行针对每个分子生成一个构象。值得注意的是，不包含基于能量的选择步骤的生成方法无法保证生成基态构象，这也是其与判别式预测方法之间存在性能差距的原因之一。尽管如此，我们仍将其相对于基态构象标签的单构象生成误差作为此处的代表性基线。对于 **EnFlow**，我们考虑了两种推理模式：

1.JustFM：使用我们的能量引导采样方案，为每个分子生成一个构象。2.EnsembleCert：为每个分子生成M 个构象，并选择能量最低的一个。当模式为 JustFM 时，引导强度设为 $\lambda_t = 0.5(1-t)^2$，以优先提高生成精确率。对于 EnsembleCert，我们使用较小的引导强度 $\lambda_t = 0.2(1-t)^2$，以提升生成构象集合的多样性，从而更有效地进行选择。

1 步和 2 步采样），其中引导模型相较于无引导基线提升了覆盖率曲线。这些结果进一步验证了我们的引导策略能够提升 GEOM-Drugs 中更复杂分子的少步构象生成效果。

2.4 EnFlow 实现稳健的基态构象选择

GEOM-Drugs 上的基态构象预测：表 3 报告了 GEOM-Drugs 上的基态构象预测结果。在现有方法中，ReBind 提供了最强的基线性能。我们首先在相同采样预算下比较单构象生成。在采样 5 步、每个分子生成一个构象的条件下，ET-Flow 的 **D**-MAE 为 $= 0.844 \pm 0.009$ 埃，**D**-RMSE 为 $= 1.491 \pm 0.020$ 埃，$\mathcal{C}$-RMSD 为 $= 1.633 \pm 0.011$ 埃；而 EnFlow (JustFM) 在所有三项指标上均有提升，分别达到 **D**-MAE $= 0.793 \pm 0.011$ 埃、**D**-RMSE $= 1.412 \pm 0.024$ 埃和 $\mathcal{C}$-RMSD $= 1.550 \pm 0.029$ 埃，凸显了即使不进行集成选择，能量引导采样也能发挥作用。所提出的 EnsembleCert 方案在集成规模较小、采样步数较少时进一步降低了误差，并且随着集成规模增大，性能持续提升。具体而言，在 $M = 50$ 且采样 50 步时，**EnFlow** 在三项指标上均取得了新的最先进性能（**D**-MAE $= 0.644$ 埃、**D**-RMSE $= 1.263$ 埃、 $\mathcal{C}$-RMSD $= 1.163$ 埃），相较此前最佳结果分别提升了 17.01%、1.56% 和 16.69%。图 10 进一步表明，增大集成规模 M 可持续降低**D**-MAE、**D**-RMSE 和 $\mathcal{C}$-RMSD；在 1–5 个常微分方程步骤下，将M 从 1 增至 20 时，提升尤为显著。值得注意的是，即使采样步数很少，只要集成规模足够大，性能也能接近采样步数多得多的情况，这表明学习到的能量函数$J_\Phi(\mathcal{C})$ 具有很强的判别能力，能够有效筛选并排序低能构象，从而实现 GEOM-Drugs 上的基态预测。

2.5 学习良好的能量景观引导构象趋向低能量 r 能量

我们的目标是学习形状良好的能量景观，以捕捉分子构象的能量变化，并从先验噪声（谐波先验样本）出发，沿近乎笔直的梯度方向趋向低能数据。为此，我们采用能量匹配

损失 [42]，用于正则化数据流形之外的区域，并使用能量回归损失来细化流形附近的能量景观。在图 11 中，我们在 GEOM-Drugs 上对 EnFlow 的三个变体进行了消融实验，使用 5 个 ODE 采样步数：(1) 无引导基线（ET-Flow；**无引导**）；(2) 仅使用 $\mathcal{L}_{\text{em}}$ 训练J_ϕ 的引导模型（**EnFlow**：**引导与** $\mathcal{L}_{\text{em}}$）；以及 (3) 同时使用两种损失训练的完全引导模型（**EnFlow**：**引导与**$\mathcal{L}_{\text{em}}$ **及** $\mathcal{L}_{\text{energy}}$）。我们为每个分子生成一个构象，并计算其与基态构象的误差，重复实验 10 次以获得统计数据。如主要的 $\mathcal{C}$-均方根偏差指标所示，完整模型（**引导与** $\mathcal{L}_{\text{em}}$ **及** $\mathcal{L}_{\text{energy}}$）的中位误差最低，且误差分布范围明显小于无引导基线和仅使用能量匹配的变体。**D**-MAE 和**D**-RMSE 也呈现类似趋势：完整模型使误差分布向较小值偏移。这些结果表明，组合训练目标促使J_ϕ 引导采样趋向低能构象，从而减小与基态结构之间的差异。此外，图 12 展示了随机选取的六个分子的学习到的能量景观，并通过对比集成采样期间是否使用引导（**使用引导**与**无引导**）消融了引导的影响。在图 12(b) 中，真实能量曲线与 J_ϕ 对真实构象的预测（均归一化至 $[0,1]$）表明，学习到的能量函数能够紧密追踪真实的能量变化。图 12(c) 和图 12(d) 进一步比较了使用 2 个和 5 个 ODE 采样步数生成的构象的预测能量分布：在两种情况下，引导都会使集成样本趋向较低能量，并缩小其分布范围；在 5 步设置下，这种集中趋势更加明显。这证实J_ϕ 能够有效引导采样趋向低能构象。

2.6 更多结果与分析

更多实验结果与分析见 D.2 节。具体而言，我们 (i) 在 D.2.1 节讨论引导调度 λ_t 的选择，(ii) 在 D.2.2 节论证进行能量匹配 训练的必要性，并且 (iii) 在 D.2.3 节展示，与其他向量场相比，EnFlow 中的能量引导向量场 在采样方面具有鲁棒性。

3 预备知识

3.1 条件流匹配

给定源分布$p_0 : \mathbb{R}^d \to \mathbb{R}_{\geq 0}$ 和目标分布$p_1 : \mathbb{R}^d \to \mathbb{R}_{\geq 0}$，概率流模型 [17, 51, 52] 引入一个随时间变化的向量场（VF）$v_t(x_t)$:$[0,1] \times \mathbb{R}^d \to \mathbb{R}^d$，它控制边缘概率路径$p_t(x_t)$：$[0,1] \times \mathbb{R}^d \to \mathbb{R}_{\geq 0}$的演化，从而促使$x_t$ 随着t 从 0 逐渐变为 1 而从p_0 转移至p_1。p_t 和v_t 都必须满足以下连续性方程：

$$\frac{\partial p_t(x_t)}{\partial t} + \nabla \cdot (p_t(x_t) v_t(x_t)) = 0. \tag{6}$$

流匹配（FM） [17, 51] 旨在学习一个由神经网络参数化的向量场$v_\theta(x_t,t)$，使其逼近真实向量场$v_t(x_t)$ ，方法是最小化以下流匹配损失：

$$\mathcal{L}_{\text{FM}} = \mathbb{E}_{t\sim\mathcal{U}(0,1),x_t\sim p_t}\left[\|v_\theta(x_t,t) - v_t(x_t)\|^2\right]. \tag{7}$$

然而，真实向量场$v_t(x_t)$ 在计算上难以处理。相比之下，定义一个条件向量场 $v_t(x_t|z)$在计算上是可行的，它会生成相应的条件概率路径 $p_t(x_t|z)$，其中z 充当条件变量，通常由一对源点和目标点表示，即x_0 和x_1 [52]。因此，条件流匹配损失 可表述为：

$$\mathcal{L}_{\text{CFM}} = \mathbb{E}_{t\sim\mathcal{U}(0,1),x_t\sim p_t(x_t|z),z\sim p(z)}\left[\|v_\theta(x_t,t) - v_t(x_t|z)\|^2\right]. \tag{8}$$

如参考文献 [17], 所示，两个损失的梯度等价，即 $\nabla_\theta\mathcal{L}_{\text{FM}} = \nabla_\theta\mathcal{L}_{\text{CFM}}$。因此，使用 $\mathcal{L}_{\text{CFM}}$训练后，学习到的向量场 $v_\theta(x_t,t)$ 可通过求解常微分方程 $dx_t = v_\theta(x_t,t)\cdot dt$ 来生成样本，其中 $x_0 \sim p_0(x_0)$。

3.2 能量引导流匹配

能量引导的一般流匹配：给定能量函数 J(x) : $\mathbb{R}^d \to \mathbb{R}$，能量引导的一般（任意源分布）流匹配旨在从引导分布 p′(x) ∝ p(x)e⁻J(x)中生成样本 x。这种方法等价于使用新的向量场 v′t 生成新的概率路径 p′t(x_t)，以此替代生成 p_t(x_t)的原始向量场 v_t 。

新的 VF v'_t 定义为：

$$v'_t(x_t) = v_t(x_t) + g_t(x_t). \tag{9}$$

定理 3.1（参考文献 [53]）。将引导 *VF* g_t（x_t）添加到原始 *VF* v_t（x_t）中，将形成 *VF* v'_t（x_t），只要 g_t（x_t）满足以下条件，即可生成 p'_t（x_t）= $\int p_t$（$x_t|z$）p'（z）dz：[1]

$$g_t(x_t) = \int \left(\mathcal{P} \frac{e^{-J(x_1)}}{Z_t(x_t)} - 1 \right) v_{t|z}(x_t|z) p(z|x_t) dz, \tag{10}$$

$$where\ Z_t(x_t) = \int \mathcal{P} e^{-J(x_1)} p(z|x_t) dz. \tag{11}$$

高斯路径流匹配（扩散）引导假设：独立耦合高斯流匹配假设源分布与目标分布相互独立，即 p(z) = p_0(x_0)p_1(x_1)，其中源分布为高斯分布，具体而言 p_0(x_0) = $\mathcal{N}$(x_0； μ, Σ)。文献中已充分确立，这种形式等价于采用不同噪声策略的扩散模型 [54, 55]。在此阶段，能量引导采取 g_t(x_t) ∝ $\nabla_{x_t} Z_t$(x_t) 的形式，这是扩散后验采样（DPS）框架 [56]中类似扩散模型引导技术的关键组成部分 [37, 56, 57] 。具体表示为 ∇_{x_t} log p′t(x_t) = ∇_{x_t} log p_t(x_t) − ∇_{x_t}J(x_t)，如参考文献 [39], 所述；或者等价地表示为 v″t (x_t) = (1 − w)v_t(x_t) + w[$a_t x_t$ + $b_t \nabla_{x_t}$J(x_t)]，其生成 p″(x) ∝ p(x)$^{1-w}$e⁻wJ(x)，如参考文献 [54]所述。此处，w 为系数，a_t、b_t 为时间调度。

（近似）非高斯路径流匹配的引导假设：为简化并使公式 10 更易处理，参考文献 [53] 提出通过泰勒展开对其进行近似，其依据是这样的直觉：$p(x_1|x_t)$ 的概率质量集中在其均值附近。首先，公式 11 中的归一化常数 $Z_t(x_t)$ 可近似为：

$$Z_t(x_t) = \int \mathcal{P} e^{-J(x_1)} p(z|x_t) dz \approx e^{-J(\hat{x}_1)}, \quad \hat{x}_1 = \mathbb{E}_{x_0, x_1 \sim p(z|x_t)}[x_1]. \tag{12}$$

接下来，引导向量场 $g_t(x_t)$ 可近似为（证明见参考文献 [53的 A.9 节）]):

$$g_t(x_t) \approx -\mathbb{E}_{z \sim p(z|x_t)} \left[(x_1 - \hat{x}_1) v_{1|t}(x_t|z) \right] \nabla_{\hat{x}_1} J(\hat{x}_1), \quad \hat{x}_1 = \mathbb{E}_{x_0, x_1 \sim p(z|x_t)}[x_1]. \tag{13}$$

对于非高斯路径流匹配（即 p_0 为非高斯分布），假设条件概率路径遵循广泛使用的仿射路径，即 $x_t = \alpha_t x_1 + \beta_t x_0 + \sigma_t \epsilon$，其中 $\epsilon \sim \mathcal{N}(0, I)$，且 σ_t、$\dot{\sigma}_t$ 均较小 [17, 52]。根据参考文献 [51], , $\hat{}x_1$ 可用向量场 v_t [53] 进行 x_1-参数化，表示为：

$$\hat{x}_1 \approx -\frac{\dot{\beta}_t}{\dot{\alpha}_t \beta_t - \dot{\beta}_t \alpha_t} x_t + \frac{\beta_t}{\dot{\alpha}_t \beta_t - \dot{\beta}_t \alpha_t} v_t(x_t). \tag{14}$$

通过将其与 $v^{1|t}(x_t|z) = \dot{\alpha}_t x_1 + \dot{\beta}_t x_0 + \dot{\sigma}_t \epsilon$ 相结合，近似引导 $g_t(x_t)$ 可以重新表述为 $p(x_1|x_t)$ 的协方差矩阵 $\Sigma_{1|t}$ 与梯度的乘积

[1]Z = ∫ e⁻J(x) p(x)dx 是归一化常数。 $\mathcal{P} = \pi'(x_0|x_1)\, \pi(x_0|x_1)$ 是反向耦合比，其中 $\pi'(x_0|x_1)$ 是新向量场的反向数据耦合，即在给定从目标分布中采样的 x_0 后，x_1 的分布。 $\mathcal{P} = 1$ 当 $\pi(x_0, x_1) = p(x_0)p(x_1)$ 时，即令 p(z) = p_0(x_0)p_1(x_1) 为独立耦合 [52, 53]。

$\nabla_{\hat{x}_1} J(\hat{x}_1)$（证明见参考文献 [53] 的附录 A.11）：

$$g_t(x_t) \approx -\frac{\dot{\alpha}_t \beta_t - \dot{\beta}_t \alpha_t}{\beta_t} \cdot \Sigma_{1|t} \cdot \nabla_{\hat{x}_1} J(\hat{x}_1). \tag{15}$$

如参考文献 [53] 所述， $\Sigma_{1|t}$ 可以近似为人工定义的调度方案。这种方法与传统扩散模型 [54, 55] 中用于非高斯路径流匹配的引导技术类似，得到以下表达式：

$$g_t(x_t) \approx -\lambda_t \cdot \nabla_{\hat{x}_1} J(\hat{x}_1). \tag{16}$$

因此，非高斯路径流匹配的能量引导向量场 $v'_t(x_t)$ 可近似为：

$$v'_t(x_t) \approx v_t(x_t) - \lambda_t \cdot \nabla_{\hat{x}_1} J(\hat{x}_1), \tag{17}$$

其中，建议 λ_t 随 $t \to 1$ 增大而衰减至零，$\hat{x}_1$ 由公式 **14** 给出。

4 方法

本节介绍所提出方法 **EnFlow** 的方法，该方法主要采用近期提出的能量匹配 技术 [42]，在流匹配训练过程中联合训练能量模型 $J_\phi(\mathcal{C})$ 。 [2] 在采样过程中，利用训练完成的能量模型$J_\phi(\mathcal{C})$ 引导反向生成过程。

4.1 用于构象生成的标准 SB-CFM（ET-Flow）

我们将 ET-Flow [29] 重新诠释为用于构象生成的无引导 FM 框架，采用标准的 薛定谔桥 条件流匹配 （SB-CFM）路径 [52] ，从非高斯先验，具体而言是 谐波先验 [40, 58]中生成构象。该方法能够以直接且高性能的方式生成构象。

谐波先验： 谐波先验 [40, 58]强制键连接的原子位置之间保持空间接近。它基于共价键构建，从而得到一个分子图 $\mathcal{G}$ 及其邻接矩阵 $A \in [0,1]^{n\times n}$，并据此推导出图拉普拉斯矩阵 $L = D - A$ ，其中 D 为度矩阵。谐波先验随后表示为 $p_0(\mathcal{C}_0) \propto \exp\left(-\frac{1}{2}\mathcal{C}_0{}^T L \mathcal{C}_0\right)$。

SB-CFM 路径： SB-CFM 路径由参考文献 [52], 引入，其形式化定义为$p(z) = \pi_{2\sigma^2}(\mathcal{C}_0, \mathcal{C}_1)$，其中 $\pi_{2\sigma^2}$ 表示熵正则化最优传输（OT）问题[3]，其代价为 $\|\mathcal{C}_0 - \mathcal{C}_1\|$ ，熵正则化参数为 σ^2 [59]。SB-CFM 路径的特征是尺度为 σ的布朗桥，用于连接构型 $\mathcal{C}_0$ 和$\mathcal{C}_1$。相应的条件流匹配 CFM 路径和生成向量场随后表示如下：

$$p_t(\mathcal{C}_t|\mathcal{C}_0, \mathcal{C}_1) = \mathcal{N}(\mathcal{C}_t; (1-t)\mathcal{C}_0 + t\mathcal{C}_1, t(1-t)\sigma^2 I), \tag{18}$$

$$v_t(\mathcal{C}_t|\mathcal{C}_0, \mathcal{C}_1) = \frac{1-2t}{2t(1-t)}\{\mathcal{C}_t - [(1-t)\mathcal{C}_0 + t\mathcal{C}_1]\} + (\mathcal{C}_1 - \mathcal{C}_0). \tag{19}$$

使用条件流匹配损失训练参数化向量场 $v_\theta(\mathcal{C}_t, t)$ 之后：

$$\mathcal{L}_{\text{SB-CFM}} = \mathbb{E}_{\mathcal{C}_t \sim p_t(\mathcal{C}_t|\mathcal{C}_0,\mathcal{C}_1), t\sim\mathcal{U}(0,1)}\left[\|v_\theta(\mathcal{C}_t, t) - v_t(\mathcal{C}_t|\mathcal{C}_0, \mathcal{C}_1)\|^2\right], \tag{20}$$

在初始条件为 $\mathcal{C}_0 \sim p_0(\mathcal{C}_0)$ 时，向量场 $v_\theta(\mathcal{C}_t, t)$ 可通过求解常微分方程 $d\mathcal{C}_t = v_\theta(\mathcal{C}_t, t) \cdot dt$ 来生成构象。

[2]在本节中，我们使用 $\mathcal{C} \in \mathbb{R}^{n\times 3}$ 表示数据空间，而非用于构象空间生成的 $x \in \mathbb{R}^d$ ，其中 ϕ 表示可训练参数。[3]在此背景下， $\mathcal{C}_0$ 和 $\mathcal{C}_1$ 并非独立样本对，而是根据最优传输代价 π [52] 采样得到。然而，参考文献 [53] 表明，对于相关的最优传输耦合，假设 $\mathcal{P} \approx 1$ 也是合理的（证明见参考文献 [53] 的 A.3 节）。

4.2 高效训练能量模型的能量匹配

基于能量的模型（EBM）： 基于能量的模型（EBM）是一类概率模型，它们通过为每个数据点关联一个能量值来定义数据的概率分布。能量函数 $J(\mathcal{C})$ 用于为概率较高的构型赋予较低的能量。这定义了一个玻尔兹曼分布 $p(\mathcal{C}) \propto e^{-J(\mathcal{C})}$。

能量匹配概述： 通过考察 Jordan–Kinderlehrer–Otto（JKO）方案的一阶最优性条件 [60, 61] [62], 参考文献 [42] 提出了一种新颖的框架，将能量模型的训练与最优传输（OT）框架相结合，并称之为能量匹配。其核心思想是学习一个与时间无关的标量势$J\phi(\mathcal{C})$ 它定义了一个玻尔兹曼分布 $p(\mathcal{C}) \propto e^{-J\phi(\mathcal{C})}$，该分布与数据分布$p_{data}(\mathcal{C})$相匹配。为了将带噪数据 $\mathcal{C}_0 \sim p_0(\mathcal{C}_0)$ 传输至数据分布 $p_{data}(\mathcal{C})$，参考文献 [42]提出了两个不同阶段：

•远离流形：在这一阶段， $\mathcal{C}_0$ 以类似最优传输的方式，以恒定速度确定性地向数据分布 $p_{data}(\mathcal{C})$ 传输。具体而言，$-\nabla_{\mathcal{C}_t} J\ \phi(\mathcal{C}'_t)\ \approx \mathcal{C}_{data}\ - \mathcal{C}_0$，其中采用插值 $\mathcal{C}'_t\ = (1-t)\mathcal{C}_0 + t\mathcal{C}_{data}$，且$\mathcal{C}_{data} \sim p_{data}(\mathcal{C})$。

• 接近流形： 在此阶段，样本向玻尔兹曼分布$p(\mathcal{C}) \propto e^{-J\phi(\mathcal{C})}$演化，从而紧密逼近数据分布。

4.3 使用 SB-CFM 联合训练 EBM

匹配阶段： 基于能量匹配 [42], 的成功，我们提出了一种联合训练框架，同时优化能量模型$J\phi(\mathcal{C})$ 和 SB-CFM 训练过程。在这一初始阶段，重点是 SB-CFM 和能量模型各自的匹配损失。在每个时间步长t，我们首先采样构象 $\mathcal{C}_0 \sim p_0(\mathcal{C}_0)$ 和 $\mathcal{C}_1 \sim p_1(\mathcal{C}_1)$，然后进行如下插值：

$$\mathcal{C}'_t = (1-t)\mathcal{C}_0 + t\mathcal{C}_1, \tag{21}$$

$$\mathcal{C}_t = \mathcal{C}'_t + \sigma\sqrt{t(1-t)} \cdot \epsilon, \quad \epsilon \sim \mathcal{N}(0, I). \tag{22}$$

随后，按如下方式计算相应的向量场：

$$S_t(\mathcal{C}'_t|\mathcal{C}_0, \mathcal{C}_1) = \mathcal{C}_1 - \mathcal{C}_0, \tag{23}$$

$$v_t(\mathcal{C}_t|\mathcal{C}_0, \mathcal{C}_1) = \frac{1-2t}{2t(1-t)}\{\mathcal{C}_t - \mathcal{C}'_t\} + S_t. \tag{24}$$

因此，SB-CFM 损失记为 $\mathcal{L}_{\text{SB}-\text{CFM}}$，其定义见公式 20，具体如下：

$$\mathcal{L}_{\text{SB}-\text{CFM}} = \mathbb{E}_{\mathcal{C}_t \sim p_t(\mathcal{C}_t|\mathcal{C}_0,\mathcal{C}_1), t\sim\mathcal{U}(0,1)} \left[\|v_\theta(\mathcal{C}_t, t) - v_t(\mathcal{C}_t|\mathcal{C}_0, \mathcal{C}_1)\|^2\right]. \tag{25}$$

然后按如下方式计算能量匹配损失：

$$\mathcal{L}_{\text{EM}} = \mathbb{E}_{\mathcal{C}_0\sim p_0(\mathcal{C}_0), \mathcal{C}_1\sim p_1(\mathcal{C}_1), t\sim\mathcal{U}(0,1)} \left[\| - \nabla_{\mathcal{C}_t} J_\phi(\mathcal{C}'_t) - S_t(\mathcal{C}'_t|\mathcal{C}_0, \mathcal{C}_1)\|^2\right]. \tag{26}$$

能量微调阶段： 在第二阶段，我们专注于对能量模型 $J\phi(\mathcal{C})$ 进行微调，以进一步增强其准确刻画分子构象能量景观的能力。值得注意的是，参考文献 [42]中提出的方法在能量模型 [63] 中采用传统的对比散度（CD）损失，以将能量模型隐式地向数据分布进行优化。然而，由于依赖基于 MCMC 的负采样，这一过程计算成本高昂且通常效率低下。在我们的场景中，数百万个分子对应着数百万个不同的能量景观，这导致使用 CD 损失进行训练时存在显著的不稳定性。得益于 GEOM 数据集 [43], 提供的高质量数据，每个分子构象都带有相应的玻尔兹曼能量标签，该标签通过高水平的 CREST+GFN2-xTB 计算并结合 DFT 微调获得（详见 A.2 节）。因此，我们提出直接

通过监督学习优化能量模型，使用预测能量值与真实能量标签之间的平均绝对误差（MAE）（按逐分子方式进行归一化）。该损失作为辅助微调目标，其权重因子 $\eta_{\text{能量}}$可控，同时在此阶段保留 $\mathcal{L}_{\text{EM}}$：

$$\mathcal{L}_{energy} = \mathbb{E}_{\mathcal{C}\sim p_{data}(\mathcal{C})}\left[\left|J_\phi(\mathcal{C}) - E_{true}(\mathcal{C})\right|\right], \tag{27}$$

$$\mathcal{L}_{fine\text{-}tune} = \mathcal{L}_{\text{EM}} + \eta_{energy}\cdot\mathcal{L}_{energy}. \tag{28}$$

该方法不仅提高了训练效率，还增强了能量模型捕捉分子构象能量景观的准确性，从而在实际应用中实现与真实能量分布更好的对齐。

模型架构：为确保公平性，我们采用与 ET-Flow [29], 中使用的相同模型架构，即 TorchMD-NET [64]，分别用于能量模型 $\mathrm{J}\phi$ 和向量场 v_θ。TorchMD-NET 为每个原子 i 生成标量特征 $\mathrm{x}_i \in \mathbb{R}^d$ 和向量特征 $\vec{\mathrm{v}}_i \in \mathbb{R}^3$，使得 $\mathrm{x}, \vec{\mathrm{v}} = \text{TorchMD-NET}(\mathcal{G}, \mathcal{C})$，其中 $\mathrm{x} \in \mathbb{R}^{n\times d}$ 且 $\vec{\mathrm{v}} \in \mathbb{R}^{n\times 3}$。对于能量模型，我们对 x 执行平均池化，并应用线性层以获得能量值，定义为 $\mathrm{J}\phi(\mathcal{C}) = \text{Mean-Pooing}(\mathrm{x}, \text{dim} = 0)\cdot \mathrm{W}$，其中 $\mathrm{W} \in \mathbb{R}^{d\times 1}$，且 $(\mathrm{x}, _) = \text{TorchMD-NET}\phi(\mathcal{G}, \mathcal{C})$。对于向量场，我们直接将向量特征 $\vec{\mathrm{v}}$ 用作输出，即 $\mathrm{v}_\theta = \vec{\mathrm{v}}$，其中 $(_, \vec{\mathrm{v}}) = \text{TorchMD-NET}_\theta(\mathcal{G}, \mathcal{C})$。详细信息请参阅参考文献 [29]的第 A.1 节和第 C 节。

4.4 用于构象生成的能量引导采样

此时，我们已经获得了训练良好的能量模型 $\mathrm{J}\phi(\mathcal{C})$ 和向量场 $\mathrm{v}_\theta(\mathcal{C}_t, \mathrm{t})$，后者生成概率路径 $\mathrm{p}_t(\mathcal{C}_t|\mathcal{C}_0, \mathcal{C}_1)$，并将非标准高斯先验 $\mathrm{p}_0(\mathcal{C}_0)$（即谐波先验）输运至真实的构象分布 $\mathrm{p}_1(\mathcal{C}_1)$。在采样过程中，能量模型可以使用式 9 中定义的引导向量场 $\mathrm{v}'\mathrm{t}(\mathcal{C}_t)$ 引导反向生成过程，将 $\mathcal{C}_0 \sim \mathrm{p}_0(\mathcal{C}_0)$ 输运至 $\mathcal{C}_1' \propto \mathrm{p}_1(\mathcal{C}_1)\mathrm{e}^{-}\mathrm{J}\phi(\mathcal{C}_1)$。

根据第 3.2 节和式 17，近似的引导向量场 $v'_t(\mathcal{C}_t)$ 可表示为：

$$v'_t(\mathcal{C}_t) \approx v_\theta(\mathcal{C}_t, t) - \lambda_t\cdot\nabla_{\hat{\mathcal{C}}_1} J_\phi(\hat{\mathcal{C}}_1), \quad \hat{\mathcal{C}}_1 \approx \mathcal{C}_t + (1-t)v_\theta(\mathcal{C}_t, t), \tag{29}$$

其中，λ_t 是一个衰减因子，当 $t \to 1$ 时趋近于零，而 $\hat{}\,\mathcal{C}_1$ 由式 14 中的重新参数化定义。采样过程可概括如下：

$$\mathcal{C}_{t+\Delta t} = \mathcal{C}_t + v'_t(\mathcal{C}_t)\cdot\Delta t, \tag{30}$$

其中 $\mathcal{C}_0 \sim p_0(\mathcal{C}_0)$。此过程持续进行，直至 $t = 1$，此时生成的构象 $\mathcal{C}_1'$ 预计将与从 $p_1(\mathcal{C}_1)$ 采样得到的真实构象非常接近，并且根据能量模型 $J\phi(\mathcal{C})$，其能量值更低，即 $\mathcal{C}_1' \propto p_1(\mathcal{C}_1)e^{-J\phi(\mathcal{C}_1)}$。整体采样过程总结于算法 2。此外，由于 λ_t 的调度被视为一个超参数，并会随着 $t \to 1$衰减至零，因此我们将关于 λ_t 调度的更详细讨论留至 D.1.4节。

为了实现更有效的单步采样，Reflow [18] 是一种后训练技术，可提升单步样本生成的质量。与 AvgFlow [30], 类似，我们也沿着 SB-CFM 路径应用 Reflow 过程，以进一步提升单步采样的性能。详细步骤见 D.1.1节。

4.5 基态构象认证

为提升所生成构象的实用价值，我们引入了基态构象认证步骤，旨在识别可能对应于全局最低能量态的构象。这项能力对于药物发现和材料设计等下游应用至关重要。我们考虑了两种互补的基态认证策略。(1) JustFM 模式：我们使用算法 2 中所述的能量引导采样过程，以单次生成的方式为每个分子直接生成一个构象，

其中，学习到的能量函数会明确引导采样轨迹朝低能量区域移动。(2) EnsembleCert 模式：另一种方式是，给定一个分子图，我们使用相同的能量引导采样过程生成一个构象系综，然后根据学习到的能量模型 $J_\phi(\mathcal{C})$对生成的构象进行排序，从而识别出基态候选。详细流程总结见算法 3。

5 局限性与结论

局限性。 尽管我们的方法在统一框架下的分子构象生成 和基态构象预测 方面均表现出色，但仍存在若干局限。首先，从常微分方程模拟的角度来看（步长 Δt），EnFlow 通过将谐波先验 样本引导至能量引导分布 $p_1(\mathcal{C}_1)e^{-J_\phi(\mathcal{C}_1)}$ 而非原始的$p_1(\mathcal{C}_1)$，实现了高质量的少步采样。然而，其推理时间显著长于无引导流匹配框架（例如 ETFlow [29]），这主要是因为计算能量模型梯度会带来额外开销。其次，尽管我们的方法借助学习到的能量景观显著降低了基态构象预测误差，但其推理延迟仍然较高，因为流匹配（或基于扩散）方法本质上需要迭代采样，而非仅进行一次前向计算。第三，使用非高斯先验（谐波先验）使得经典的高斯路径流匹配模型引导技术不再适用。因此，我们提出的能量引导机制依赖于参考文献 [53], 中近似的非高斯引导策略，其最优性和可控性仍有限。最后，尽管我们的框架首次以统一方式同时处理这两项任务，但在进一步提升特定任务的性能方面仍有很大空间。未来的工作将致力于克服这些局限，以增强所提方法的稳健性、效率和整体效果。

结论。 在这项工作中，我们对一种占主导地位的范式提出质疑：该范式将分子构象生成 与基态构象预测 视为彼此独立的任务。现有方法要么能够生成多样的构象，却无法评估其能量质量；要么以牺牲结构多样性为代价，预测单一的最低能量构象。为弥合这一差距，我们提出首个统一框架，可同时生成多样的构象并学习忠实的能量景观，从而可靠地区分并筛选基态构象。基于流匹配生成框架以及采用能量匹配 目标训练的基于能量的模型 ，我们提出一种适用于非高斯流匹配路径的能量引导流匹配方案，显著提升了从谐波先验 样本出发进行少步常微分方程采样的质量。在基态构象预测方面，学习到的能量景观可作为计算成本高昂的量子力学方法的快速、准确替代方案，从而能够高效认证最低能量构象。在多个基准测试上的大量实验表明，我们的方法在少步分子构象生成和基态构象预测两项任务中，均持续优于最先进方法。

参考文献

[1] Cristiano RW Guimarães、Alan M Mathiowetz、Marina Shalaeva、Gilles Goetz 和 Spiros Liras。利用三维性质刻画五规则性质空间之外的被动渗透。《化学信息与建模杂志》，52(4):882–890，2012。[2] Christof H Schwab。构象与三维药效团搜索。《今日药物发现：技术》，7(4):e245–e253，2010。[3] Paul CD Hawkins。构象生成：最新进展。《化学信息与建模杂志》，57(8):1747–1756，2017。[4] Andrew J Ballard、Stefano Martiniani、Jacob D Stevenson、Sandeep Somani 和 David J Wales。利用势能景观对自由能进行采样。《Wiley跨学科评论：计算分子科学》，5(3):273–289，2015。[5] Marco De Vivo、Matteo Masetti、Giovanni Bottegoni 和 Andrea Cavalli。分子动力学及相关方法在药物发现中的作用。《药物化学杂志》，59(9):4035–4061，2016。[6] Philipp Pracht、Fabian Bohle 和 Stefan Grimme。利用快速量子化学方法自动探索低能化学空间。《物理化学化学物理》，22(14):7169–7192，2020。[7] Robert G Parr、Shridhar R Gadre 和 Libero J Bartolotti。原子和分子的局域密度泛函理论。《美国国家科学院院刊》，76(6):2522–2526，1979。[8] P Muller 等。物理有机化学术语表。《纯粹与应用化学》，66(5):1077–1184，1994。[9] Zhi-Hua Zhou。机器学习。Springer Nature，2021。[10] Ethem Alpaydin。机器学习。MIT出版社，2021。[11] Christian Janiesch、Patrick Zschech 和 Kai Heinrich。机器学习与深度学习。《电子市场》，31(3):685–695，2021。[12] Taylan Cemgil、Sumedh Ghaisas、Krishnamurthy Dvijotham、Sven Gowal 和 Pushmeet Kohli。自编码变分自编码器。《神经信息处理系统进展》，33:15077–15087，2020。[13] Ian Goodfellow、Jean Pouget-Abadie、Mehdi Mirza、Bing Xu、David Warde-Farley、Sherjil Ozair、Aaron Courville 和 Yoshua Bengio。生成对抗网络。《ACM通讯》，63(11):139–144，2020。[14] Jascha Sohl-Dickstein、Eric Weiss、Niru Maheswaranathan 和 Surya Ganguli。利用非平衡热力学进行深度无监督学习。载于《国际机器学习会议》，第2256–2265页。PMLR，2015。[15] Yang Song 和 Stefano Ermon。通过估计数据分布的梯度进行生成建模。《神经信息处理系统进展》，32，2019。[16] Jonathan Ho、Ajay Jain 和 Pieter Abbeel。去噪扩散概率模型。《神经信息处理系统进展》，33:6840–6851，2020。[17] Yaron Lipman、Ricky TQ Chen、Heli Ben-Hamu、Maximilian Nickel 和 Matthew Le。用于生成建模的流匹配。载于《第十一届国际学习表征会议》，2023。[18] Xingchao Liu、Chengyue Gong 和 Qiang Liu。让流更直、更快：利用校正流学习生成和迁移数据。arXiv预印本 arXiv:2209.03003，2022。[19] Danilo Rezende 和 Shakir Mohamed。基于归一化流的变分推断。载于《国际机器学习会议》，第1530–1538页。PMLR，2015。[20] Gregor NC Simm 和 José Miguel Hernández-Lobato。用于分子距离几何的生成模型。arXiv预印本 arXiv:1909.11459，2019。[21] Minkai Xu、Shitong Luo、Yoshua Bengio、Jian Peng 和 Jian Tang。学习用于分子构象生成的神经生成动力学。arXiv预印本 arXiv:2102.10240，2021。

[22] Minkai Xu、Wujie Wang、Shitong Luo、Chence Shi、Yoshua Bengio、Rafael Gomez-Bombarelli 和 Jian Tang。通过双层规划实现分子构象生成的端到端框架。载于国际机器学习会议，第 11537–11547 页。PMLR，2021年。[23] Chence Shi、Shitong Luo、Minkai Xu 和 Jian Tang。学习用于分子构象生成的梯度场。载于国际机器学习会议，第 9558–9568 页。PMLR，2021年。[24] Octavian Ganea、Lagnajit Pattanaik、Connor Coley、Regina Barzilay、Klavs Jensen、William Green 和 Tommi Jaakkola。Geomol：用于三维构象体集合的扭转几何生成。神经信息处理系统进展，34：13757–13769，2021年。[25] Jiaqi Guan、Wesley Wei Qian、Wei-Ying Ma、Jianzhu Ma 和 Jian Peng。能量启发的分子构象优化。载于国际学习表征会议，2021年。[26] Minkai Xu、Lantao Yu、Yang Song、Chence Shi、Stefano Ermon 和 Jian Tang。GeoDiff：用于分子构象生成的几何扩散模型。*arXiv*预印本 *arXiv:2203 02923*，2022年。。[27] Nathaniel Thomas、Tess Smidt、Steven Kearnes、Lusann Yang、Li Li、Kai Kohlhoff 和 Patrick Riley。张量场网络：用于三维点云的旋转和平移等变神经网络。*arXiv*预印本 *arXiv:1802.08219*，2018。[28] Vıctor Garcia Satorras、Emiel Hoogeboom 和 Max Welling。E (n) 等变图神经网络。载于国际机器学习会议，第 9323–9332 页。PMLR，2021年。[29] Majdi Hassan、Nikhil Shenoy、Jungyoon Lee、Hannes Stärk、Stephan Thaler 和 Dominique Beaini。ET-Flow：用于分子构象生成的等变流匹配。神经信息处理系统进展，37：128798–128824，2024。[30] Zhonglin Cao、Mario Geiger、Allan Dos Santos Costa、Danny Reidenbach、Karsten Kreis、Tomas Geffner、Franco Pellegrini、Guoqing Zhou 和 Emine Kucukbenli。通过 SO (3*09785*，2025*2507*)-平均流匹配和 Reflow 高效生成分子构象。*arXiv*预印本 *arXiv:* 。。[31] Zhao Xu、Youzhi Luo、Xuan Zhang、Xinyi Xu、Yaochen Xie、Meng Liu、Kaleb Dickerson、Cheng Deng、Maho Nakata 和 Shuiwang Ji。Molecule3D：从分子图预测三维几何结构的基准。*arXiv*预印本 *arXiv:2110.01717*，2021年。[32] Guikun Xu、Yongquan Jiang、PengChuan Lei、Yan Yang 和 Jim Chen。GTMGC：使用图变换器预测分子的基态构象。载于第十二届国际学习表征会议，2023。[33] Shengjie Luo、Yixian Xu、Di He、Shuxin Zheng、Tie-Yan Liu 和 Liwei Wang。通过几何扩散桥连接几何状态。神经信息处理系统进展，37：109283–109322，2024。[34] Fanmeng Wang、Minjie Cheng 和 Hongteng Xu。WGFormer：由 Wasserstein 梯度流驱动、用于分子基态构象预测的 SE (3)-Transformer。载于第四十二届国际机器学习大会，2025。[35] Taewon Kim、Hyunjin Seo、Sungsoo Ahn 和 Eunho Yang。ReBind：通过基于力的图重连增强基态分子构象预测。载于第十三届国际学习表征会议，2025。[36] Chengxuan Ying、Tianle Cai、Shengjie Luo、Shuxin Zheng、Guolin Ke、Di He、Yanming Shen 和 Tie-Yan Liu。Transformer 在图表征方面的表现真的很差吗？神经信息处理系统进展，34：28877–28888，2021年。[37] Prafulla Dhariwal 和 Alexander Nichol。扩散模型在图像合成方面胜过生成对抗网络。神经信息处理系统进展，34：8780–8794，2021年。[38] Jonathan Ho 和 Tim Salimans。无分类器扩散引导。*arXiv*预印本 *arXiv:2207.12598*，2022年。

[39] Cheng Lu、Huayu Chen、Jianfei Chen、Hang Su、Chongxuan Li 和 Jun Zhu。用于离线强化学习中精确能量引导扩散采样的对比能量预测。载于国际机器学习会议，第 22825–22855 页。PMLR，2023。[40] Hannes Stärk、Bowen Jing、Regina Barzilay 和 Tommi Jaakkola。用于多配体对接和结合位点设计的谐波自条件流匹配。*arXiv*预印本 *arXiv:2310 05764*，2023。。[41] Yann LeCun、Sumit Chopra、Raia Hadsell、M Ranzato、Fujie Huang 等。基于能量的学习教程。结构化数据预测，1(0)，2006年。[42] Michal Balcerak、Tamaz Amiranashvili、Antonio Terpin、Suprosanna Shit、Lea Bogensperger、Se- bastian Kaltenbach、Petros Koumoutsakos 和 Bjoern Menze。能量匹配：统一流匹配与基于能量的生成建模。*arXiv*预印本 *arXiv:2504 10612*，2025。。[43] Simon Axelrod 和 Rafael Gomez-Bombarelli。GEOM：用于性质预测和分子生成的能量标注分子构象。*Scientific Data*，9(1)：185，2022年。[44] Bowen Jing、Gabriele Corso、Jeffrey Chang、Regina Barzilay 和 Tommi Jaakkola。用于分子构象生成的扭转扩散。神经信息处理系统进展，35：24240–24253，2022年。[45] Zhiguang Fan、Yuedong Yang、Mingyuan Xu 和 Hongming Chen。EC-conf：一种利用等变一致性进行分子构象生成的超快速扩散模型。*Journal of Cheminformatics*，16(1)：107，2024。[46] Yuyang Wang、Ahmed A Elhag、Navdeep Jaitly、Joshua M Susskind 和 Miguel Angel Bautista。吞下苦药：简化且可扩展的构象生成。*arXiv*预印本 *arXiv:2311.17932*，2023。[47] Greg Landrum 等。RDKit：用于化学信息学、计算化学和预测建模的软件套件。*Greg Landrum*，8(31.10)：5281，2013年。[48] Weihua Hu、Bowen Liu、Joseph Gomes、Marinka Zitnik、Percy Liang、Vijay Pande 和 Jure Leskovec。图神经网络预训练策略。*arXiv*预印本 *arXiv:1905 12265*，2019。。[49] Shaked Brody、Uri Alon 和 Eran Yahav。图注意力网络的注意力程度如何？*arXiv*预印本 *arXiv:2105.14491*，2021。[50] Ladislav Rampášek、Michael Galkin、Vijay Prakash Dwivedi、Anh Tuan Luu、Guy Wolf 和 Dominique Beaini。通用、强大且可扩展的图变换器设计方案。神经信息处理系统进展，35：14501–14515，2022年。[51] Yaron Lipman、Marton Havasi、Peter Holderrieth、Neta Shaul、Matt Le、Brian Karrer、Ricky TQ Chen、David Lopez-Paz、Heli Ben-Hamu 和 Itai Gat。流匹配指南与代码。*arXiv*预印本 *arXiv:2412.06264*，2024。[52] Alexander Tong、Kilian Fatras、Nikolay Malkin、Guillaume Huguet、Yanlei Zhang、Jarrid Rector- Brooks、Guy Wolf 和 Yoshua Bengio。利用小批量最优传输改进并泛化基于流的生成模型。*Transactions on Machine Learning Research*，第 1–34 页，2024年。[53] Ruiqi Feng、Chenglei Yu、Wenhao Deng、Peiyan Hu 和 Tailin Wu。论流匹配引导。*arXiv*预印本 *arXiv:2502.02150*，2025。[54] Qinqing Zheng、Matt Le、Neta Shaul、Yaron Lipman、Aditya Grover 和 Ricky TQ Chen。用于生成建模和决策的引导流。*arXiv*预印本 *arXiv:2311 13443*，2023。。[55] Nanye Ma、Mark Goldstein、Michael S Albergo、Nicholas M Boffi、Eric Vanden-Eijnden 和 Saining Xie。SiT：利用可扩展插值变换器探索基于流和扩散的生成模型。载于欧洲计算机视觉会议，第 23–40 页。Springer，2024。[56] Hyungjin Chung、Jeongsol Kim、Michael T Mccann、Marc L Klasky 和 Jong Chul Ye。面向一般含噪逆问题的扩散后验采样。*arXiv*预印本 *arXiv:2209.14687*，2022年。

[57] Jiaming Song、Qinsheng Zhang、Hongxu Yin、Morteza Mardani、Ming-Yu Liu、Jan Kautz、Yongxin Chen 和 Arash Vahdat。用于即插即用可控生成的损失引导扩散模型。载于国际机器学习会议，第 32483–32498 页。PMLR，2023年。[58] Bowen Jing、Ezra Erives、Peter Pao-Huang、Gabriele Corso、Bonnie Berger 和 Tommi Jaakkola。Eigenfold：基于扩散模型的生成式蛋白质结构预测。*arXiv* 预印本 *arXiv:2304.02198*，2023年。[59] Marco Cuturi。Sinkhorn 距离：以极速计算最优传输。神经信息处理系统进展，26，2013年。[60] Nicolas Lanzetti、Antonio Terpin 和 Florian Dörfler。Wasserstein 空间中的变分分析。*arXiv* 预印本 *arXiv:2406.10676*，2024年。[61] Nicolas Lanzetti、Saverio Bolognani 和 Florian Dörfler。Wasserstein 空间中优化的一阶条件。*SIAM* 数据科学数学杂志，7(1)：274–300，2025年。[62] Richard Jordan、David Kinderlehrer 和 Felix Otto。福克–普朗克方程的变分表述。*SIAM* 数学分析杂志，29(1)：1–17，1998年。[63] Geoffrey E Hinton。通过最小化对比散度训练专家乘积模型。神经计算，14(8)：1771–1800，2002年。[64] Philipp Thölke 和 Gianni De Fabritiis。TorchMD-NET：用于神经网络分子势的等变变换器。*arXiv* 预印本 *arXiv:2202.02541*，2022年。[65] Anthony K Rappé、Carla J Casewit、KS Colwell、William A Goddard III 和 W Mason Skiff。UFF：用于分子力学和分子动力学模拟的完整元素周期表力场。美国化学会志，114(25)：10024–10035，1992年。[66] Thomas A Halgren。默克分子力场。V. 使用实验数据、额外计算数据和经验规则扩展 MMFF94。计算化学杂志，17(5-6)：616–641，1996年。[67] Diederik P Kingma 和 Max Welling。自编码变分贝叶斯。*arXiv* 预印本 *arXiv:1312.6114*，2013年。[68] Leo Liberti、Carlile Lavor、Nelson Maculan 和 Antonio Mucherino。欧几里得距离几何及其应用。*SIAM* 综述，56(1)：3–69，2014年。[69] Yang Song、Jascha Sohl-Dickstein、Diederik P Kingma、Abhishek Kumar、Stefano Ermon 和 Ben Poole。通过随机微分方程进行基于分数的生成建模。*arXiv* 预印本 *arXiv:2011.13456*，2020年。[70] rdkit。RDKit：开源化学信息学工具，2016年。访问日期：2025-08-07。[71] Ashish Vaswani、Noam Shazeer、Niki Parmar、Jakob Uszkoreit、Llion Jones、Aidan N Gomez、Łukasz Kaiser 和 Illia Polosukhin。Attention Is All You Need。神经信息处理系统进展，30，2017年。[72] Yang Song、Prafulla Dhariwal、Mark Chen 和 Ilya Sutskever。一致性模型。载于国际机器学习会议，第 32211–32252 页。PMLR，2023年。[73] Guikun Xu、Yankai Yu、Yongquan Jiang、Yan Yang 和 Yatao Bian。CoFM：通过 SE(3) 不变潜在空间中的流匹配生成分子构象。载于*ICML 2025* 生成式人工智能与生物学（*GenBio*）研讨会，2025年。[74] Robin Rombach、Andreas Blattmann、Dominik Lorenz、Patrick Esser 和 Björn Ommer。利用潜在扩散模型进行高分辨率图像合成。载于*IEEE/CVF* 计算机视觉与模式识别会议论文集，第 10684–10695 页，2022年。[75] Hongkai Zheng、Wenda Chu、Austin Wang、Nikola Kovachki、Ricardo Baptista 和 Yisong Yue。集合卡尔曼扩散引导：一种用于逆问题的无导数方法。*arXiv* 预印本 *arXiv:2409.20175*，2024年。[76] Marcel Kollovieh、Marten Lienen、David Lüdke、Leo Schwinn 和 Stephan Günnemann。利用高斯过程先验进行流匹配，以实现概率时间序列预测。*arXiv* 预印本 *arXiv:2410.03024*，2024年。

[77] Shiyuan Zhang、Weitong Zhang 和 Quanquan Gu。用于离线强化学习的能量加权流匹配。*arXiv* 预印本 *arXiv:2503.04975*，2025年。[78] Philipp Pracht、Stefan Grimme、Christoph Bannwarth、Fabian Bohle、Sebastian Ehlert、Gereon Feldmann、Johannes Gorges、Marcel Müller、Tim Neudecker、Christoph Plett 等。CREST——用于探索低能分子化学空间的程序。《化学物理学杂志》，160(11)，2024年。[79] Kristof T Schütt、Huziel E Sauceda、P-J Kindermans、Alexandre Tkatchenko 和 K-R Müller。Schnet——面向分子和材料的深度学习架构。《化学物理学杂志》，148(24)，2018年。[80] Oliver T Unke 和 Markus Meuwly。PhysNet：用于预测能量、力、偶极矩和部分电荷的神经网络。《化学理论与计算杂志》，15(6)：3678–3693，2019年。[81] Jimmy Lei Ba、Jamie Ryan Kiros 和 Geoffrey E Hinton。层归一化。*arXiv* 预印本 *arXiv:1607.06450*，2016年。[82] Mostafa Dehghani、Josip Djolonga、Basil Mustafa、Piotr Padlewski、Jonathan Heek、Justin Gilmer、Andreas Peter Steiner、Mathilde Caron、Robert Geirhos、Ibrahim Alabdulmohsin 等。将视觉Transformer扩展至 220 亿参数。载于国际机器学习会议，第 7480–7512 页。PMLR，2023年。[83] Patrick Esser、Sumith Kulal、Andreas Blattmann、Rahim Entezari、Jonas Müller、Harry Saini、Yam Levi、Dominik Lorenz、Axel Sauer、Frederic Boesel 等。扩展校正流Transformer以实现高分辨率图像合成。载于第四十一届国际机器学习会议，2024年。[84] Max Welling 和 Yee W Teh。通过随机梯度朗之万动力学进行贝叶斯学习。载于第 *28* 届国际机器学习会议论文集（*ICML-11*），第 681–688 页，2011年。

附录

A 相关工作与数据集详情

A.1 相关工作

A.1.1 分子构象生成

分子构象生成 任务旨在根据分子的拓扑图预测其三维结构。针对传统方法的效率局限 [65, 66, 7]„ 研究者的关注显著转向了数据驱动方法，尤其是利用深度学习技术的方法。早期工作 [20, 21, 22, 23] 探索了变分自编码器 [67], 、归一化流 [19], 和 NCSN [15] 在生成原子距离矩阵方面的应用，随后通过距离几何方法将其转换为三维坐标 [68]。然而，这些方法的性能往往不尽如人意，生成的三维结构质量较低，不适用于实际应用。

尽管也有其他方法采用各种技术优化性能 [24, 25], ，近期进展促使更强大的扩散模型 [14, 16, 69]应运而生。GeoDiff [26] 率先应用 DDPM [16] 直接生成三维坐标。TorDiff [44] 首先尝试获取 RDKit 生成的 [70] 构象，并在其超环面表面上应用 DDPM。MCF [46] 通过扩展大型 Transformer [71] 而非同等规模的图神经网络 [27, 28, 64] 来进行模型选择，进一步提升了性能，但仍面临采样效率低的问题。

为提高采样效率，近期的基于常微分方程的方法（如一致性模型 [72] 和流匹配 [17, 18, 51, 52], [45, 29, 30, ）已被应用于生成快速且准确的构象。ET-Flow 通过薛定谔桥条件流匹配 [52], 将谐波先验输运至构象空间，实现了最先进的性能。AvgFlow [30] 则通过 Reflow 和蒸馏 [18]显著推动了 1(2) 步快速采样技术的发展。 73] [45] [40]

A.1.2 分子基态构象预测

受上述 A.1.1 节中的方法在不进行显式能量最小化的情况下生成构象这一观察的启发，近期研究逐渐转向分子基态构象的确定性预测。参考文献 [31] 中的工作首次引入了 Molecule3D 基准，其中每个分子都与一个高水平 DFT 优化的基态构象配对；然而，所提出的任务定义并不理想。GTMGC [32] 通过基于图 Transformer [36], 的更合适任务定义和强基线方法对该基准进行了补充，后续工作 [33, 34] 进一步提升了在 Molecule3D 上的性能，但其实验设置缺少与生成方法的比较，因此全面性受到限制。ReBind [35] 将实验扩展至更全面的 GEOM-Drugs [43]数据集：对于每个分子，选择其构象集合中玻尔兹曼权重最高的构象作为基态。随后，该方法重新评估了基线方法，并扩展了扭转扩散 [44] ，以实现更公平、更全面的比较。我们的方法与这些方法正交：我们专注于增强分子构象生成的生成建模，随后通过基于能量的选择确定基态。因此，遵循 ReBind [35], ，我们在 GEOM-Drugs 上评估基态构象预测，并报告 ReBind [35]中提供的基线结果。

A.1.3 扩散与流匹配中的引导

引导式生成技术，例如类别引导 [37, 38]和能量引导的 [39] 建模，在生成建模领域备受关注。这些方法旨在生成符合特定属性或类别的样本，而非仅仅从学习到的分布 [74, 57]中采样。这些技术已在标准扩散模型 [14, 15, 16, 69], 中得到充分发展，而这类模型近年来日益受到关注 [37, 38, 56, 74, 57, 75]。

鉴于流匹配 [17, 18, 51] 近来已成为扩散模型的一种更高效的替代方案，越来越多的研究开始探索这一设定下的引导机制 [54, 76, 77]。然而，现有研究大多考察的是采用高斯先验的流匹配，在概念上等同于扩散模型 [54], ，因此采用的引导策略也与扩散模型中的策略十分相似。相比之下，流匹配的一大独特优势在于，它能够将任意先验分布转换为目标分布，而不像扩散模型那样受限于高斯先验。这种灵活性表明，在非高斯先验下开发理论与实用引导方法，突破基于高斯的范式，仍有巨大的潜力有待发掘。在这一方向上，参考文献 [53] 的近期工作提出了一个理论框架，为此类探索奠定了基础。

A.2 数据集详情

在这项工作中，我们利用 **GEOM** [43] 数据集进行全面评估。下文将从四个方面概述其关键特征：分子来源、构象来源、构象能量标注，以及构象的玻尔兹曼权重。

分子来源： GEOM 汇集了来自 AICures 和多个 MoleculeNet 集合的类药化合物，以及 QM9 小分子数据集。使用 RDKit 生成规范 SMILES，将不同来源的属性合并到单个分子条目中；对簇 SMILES 进行脱盐处理，并在适当时进行质子化，以统一离子化状态。GEOM 还纳入了 MoleculeNet 中包含实验结合亲和力的 BACE 子集；由于规模原因，其余大多数生物物理数据集均未纳入。对于纳入的 MoleculeNet 数据集，真空构象体–旋转异构体系综的恢复率超过 98%；几乎所有 BACE 分子均生成了水溶剂系综。

构象来源： 对每个分子，先准备初始几何结构并进行优化，然后再进行 CREST 采样。CREST（构象体–旋转异构体系综采样工具）是一种自动化工作流，通过元动力学和半经验量子化学探索低能构象空间 [78]。类药分子先使用 RDKit 生成嵌入结构（多个候选构象），再经 MMFF 优化和筛选，随后进行 xTB 优化；能量最低的 xTB 构象用作 CREST 的初始构象。尽管 QM9 结构已通过 DFT 优化，但仍会使用 xTB 重新优化，以匹配 CREST 所采用的理论水平。随后，CREST 在 NVT恒温器下进行元动力学，并施加依赖历史的基于均方根偏差的偏置；沿轨迹发现的新结构会被加入参考构象。通过改变偏置参数进行多次元动力学运行，可进一步提高扭转空间的覆盖率，同时避免发生断键事件。

构象能量：大规模集合集的主要标签是使用 CREST 计算得到的 GFN2-xTB 能量。对于选定的子集（尤其是 BACE），还提供准确性更高的计算方案，包括 DFT 单点能，以及在某些情况下，包含振动贡献和溶剂化贡献的完整 CENSO 精修结果。

玻尔兹曼权重：令 $\lambda = 1/(k_{\mathrm{B}}T)$。GEOM 提供两种加权方案。在$CREST/xTB$ 方案中，每个构象i 根据其 GFN2-xTB 能量E_i，并结合明确的简并度因子d_i（用于计数化学等价的旋转异构体），被赋予一个近似权重，

$$w_i^{\mathrm{xTB}} = \frac{d_i\, e^{-\lambda E_i}}{\sum_j d_j\, e^{-\lambda E_j}},$$

该权重可作为自由能概率的近似值，但不包含平移、旋转和振动贡献。在$CENSO/DFT$ 方案中，根据各构象特有的自由能计算统计权重，

$$G_i = E_{\mathrm{gas}}^{(i)} + \Delta G_{\mathrm{solv}}^{(i)}(T) + G_{\mathrm{trv}}^{(i)}(T), \qquad w_i^{\mathrm{DFT}} = \frac{e^{-\lambda G_i}}{\sum_j e^{-\lambda G_j}},$$

且不显式使用简并度因子。为保证数值稳定性，计算前可将E_i 或G_i 替换为相对值（例如减去最小值）。

划分策略：遵循近期研究 [29, 30] 以及参考文献 [24,44], 中的划分方案，在两个标准 GEOM 子集上进行评估：**GEOM-Drugs** 和**GEOM- QM9**。GEOM-Drugs 包含约 30.4 万个类药分子（均值为 44 个原子），训练集/验证集/测试集分别包含 243,473/30,433/1,000 个分子。GEOM-QM9 包含约 12 万个小型有机分子（均值为 11 个原子），训练集/验证集/测试集分别包含 106,586/13,323/1,000 个分子。实际操作中，对于训练集和验证集，每个分子最多保留 30 个构象，即保留按玻尔兹曼权重排序最高的 30 个构象；无法通过 RDKit 处理的分子将被排除。

B 所提出的算法

我们为本方法中开发的主要算法提供伪代码。整体训练过程见算法 1。用于生成构象的能量引导采样过程详见算法 2。最后，基态构象认证过程见算法 3。

算法 1 基于 SB-CFM 的能量模型联合训练

1: 输入：初始 $p_0(\mathcal{C}_0)$、$p_1(\mathcal{C}_1)$、$J\phi(\mathcal{C})$（能量模型）、$v_\theta(\mathcal{C}_t, t)$（速度场）、 σ （噪声尺度）以及学习率 η_{CFM}、 η_{EM} 和 η_{CD}。2: 输出：训练后的参数 ϕ 和 θ。3: 匹配阶段：4: 当尚未收敛时，循环执行 5: 采样 $\mathcal{C}_0 \sim p_0(\mathcal{C}_0)$、 $\mathcal{C}_1 \sim p_1(\mathcal{C}_1)$、$t \sim \mathcal{U}(0,1)$ 和噪声 $\epsilon \sim \mathcal{N}(0, I)$。6: 计算 $\mathcal{C}'_t = (1-t)\mathcal{C}_0 + t\mathcal{C}_1$; $S_t(\mathcal{C}'_t|\mathcal{C}_0,\mathcal{C}_1) = \mathcal{C}_1 - \mathcal{C}_0$。7: 计算 $\mathcal{C}_t = \mathcal{C}'_t + \sigma\sqrt{t(1-t)}\cdot\epsilon$; $v_t(\mathcal{C}_t|\mathcal{C}_0,\mathcal{C}_1) = \frac{1-2t}{2t(1-t)}\{\mathcal{C}_t - \mathcal{C}'_t\} + S_t$。8: 计算 $\mathcal{L}_{SB\text{-}CFM} = \|v_\theta(\mathcal{C}_t, t) - v_t(\mathcal{C}_t|\mathcal{C}_0,\mathcal{C}_1)\|^2$; $\mathcal{L}_{EM} = \| - \nabla_{\mathcal{C}_t} J\phi(\mathcal{C}'_t) - S_t(\mathcal{C}'_t|\mathcal{C}_0,\mathcal{C}_1)\|^2$。9: 更新 $\theta \leftarrow \theta - \eta_{CFM}\nabla_\theta \mathcal{L}_{SB\text{-}CFM}$; $\phi \leftarrow \phi - \eta_{EM}\nabla_\phi \mathcal{L}_{EM}$。10: 循环结束 11: 能量微调阶段（冻结 θ）：12: 当尚未收敛时，循环执行 13: 采样 $\mathcal{C}_0 \sim p_0(\mathcal{C}_0)$、 $\mathcal{C}_1 \sim p_1(\mathcal{C}_1)$ 和 $t \sim \mathcal{U}(0,1)$。14: 计算 $\mathcal{C}'_t = (1-t)\mathcal{C}_0 + t\mathcal{C}_1$; $S_t(\mathcal{C}'_t|\mathcal{C}_0,\mathcal{C}_1) = \mathcal{C}_1 - \mathcal{C}_0$。15: 计算 $\mathcal{L}_{EM} = \| - \nabla_{\mathcal{C}_t} J\phi(\mathcal{C}'_t) - S_t(\mathcal{C}'_t|\mathcal{C}_0,\mathcal{C}_1)\|^2$。16: 计算 $\mathcal{L}_{energy} = \mathbb{E}_{\mathcal{C}\sim p_{data}(\mathcal{C})}\left[|J\phi(\mathcal{C}) - E_{true}(\mathcal{C})|\right]$。17: 更新 $\phi \leftarrow \phi - (\eta_{EM}\nabla_\phi \mathcal{L}_{EM} + \eta_{energy}\nabla_\phi \mathcal{L}_{energy})$。18: 循环结束 19: 返回：训练后的能量模型参数 ϕ 和速度场参数 θ。

算法 2 用于构象生成的能量引导采样

1：输入：初始先验 $p_0(\mathcal{C}_0)$、训练好的能量模型 $J\phi(\mathcal{C})$、训练好的速度场 $v_\theta(\mathcal{C}_t, t)$、衰减因子 λ_t、采样步数 N，以及时间步长 $\Delta t = 1/N$。2：输出：生成的构象 $\mathcal{C}'_1$。3：设置 $t = 0$；初始化 $\mathcal{C}_t = \mathcal{C}_0$。4：对于 $i = 0$ 到 $N - 1$，执行 5：设置 $t = i \cdot \Delta t$。6：从 $\mathcal{C}_0 \sim p_0(\mathcal{C}_0)$ 中采样。7：计算 $\hat{\mathcal{C}}_1 = \mathcal{C}_t + (1-t)v_\theta(\mathcal{C}_t, t)$。8：计算引导向量场：$v'(\mathcal{C}_t) = v_\theta(\mathcal{C}_t, t) - \lambda_t \cdot \nabla \hat{\mathcal{C}}_1 J\phi(\hat{\mathcal{C}}_1)$。9：更新构象：$\mathcal{C}_{t+\Delta t} = \mathcal{C}_t + v'(\mathcal{C}_t)\cdot\Delta t$。10：结束循环。11：返回：生成的构象 $\mathcal{C}'_1 = \mathcal{C}_t$。

C 模型架构

本节介绍本研究采用的可学习神经网络架构。为确保与 ET-Flow 框架 [45], 保持一致，我们采用 TorchMD-NET [64]的架构设计。需要强调的是，本节绝大多数描述和技术细节均改编自 ET-Flow 资料 [45], ，并针对我们的设置作了少量修改。我们衷心感谢 ET-Flow [45] 的作者所作出的开源贡献。

C.1 架构

图 4(a) 中改进后的 TorchMD-NET 架构由两个主要部分组成：表示层和输出层。表示层采用了 TorchMD-NET [64] 等变变换器架构中的嵌入层和等变注意力更新层的改进版本。输出层采用了文献 [79] 中的门控等变块。由于 ET-Flow [45] 使用的网络比 TorchMD-NET [64] 论文提出的网络更大，因此对其作了一些修改以稳定训练。此外，由于输入结构是先验采样结构与实际构象之间的插值，因此必须确保当插值中有两个原子彼此非常接近时，网络仍具有数值稳定性。

嵌入层： 嵌入层将每个原子的物理和化学性质映射到学习表示空间中，同时捕获局部原子特征和几何邻域信息。对于一个含有N 个原子的分子中的第i个原子，我们通过以下方式计算不变嵌入x_i

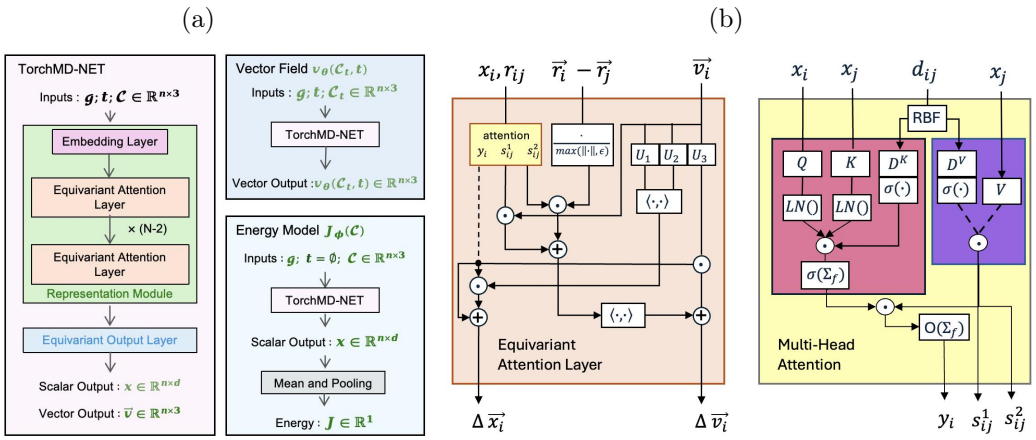


图 4：本文中的模型架构。(a) 遵循 ET-Flow [45], , 主要的可学习神经网络是 TorchMD-NET [64], , 经过修改后将时间作为额外输入特征，如左图所示。右图展示了 TorchMD-NET 如何同时用作向量场模型和能量模型。(b) TorchMD-NET 中等变注意力层和多头注意力 组件的细节，图示改编自 ET-Flow [45] 资源。

具体过程如下：

$$z_i = \text{embed}^{\text{int}}(z_i) \tag{31}$$

$$h_i = \text{MLP}(z_i) \tag{32}$$

其中，z_i 是原子序数，h_i 表示原子属性。多层感知机将原子属性投影为维度为d_h的特征向量。

接下来，我们计算一个能够捕捉局部原子环境的邻域嵌入n_i：

$$n_i = \sum_{j=1}^{N} \text{embed}^{\text{nbh}}(z_j) \cdot g(d_{ij}, l_{ij}). \tag{33}$$

这里，$\text{embed}^{\text{nbh}}(z^j)$ 为邻近原子序数提供独立的嵌入，d^{ij} 是原子i 和j之间的距离，l^{ij} 对边特征进行编码（来自基于半径的图或分子键）。交互函数$g(d^{ij},l^{ij})$ 结合距离和边信息：

$$g(d_{ij}, l_{ij}) = W^F \left[\phi(d_{ij}) e_1^{\text{RBF}}(d_{ij}), \ldots, \phi(d_{ij}) e_K^{\text{RBF}}(d_{ij}), l_{ij}\right] \tag{34}$$

算法 3 基态构象认证

1：输入：分子图 $\mathcal{G}$、训练好的能量模型 $J\phi(\mathcal{C})$、采样器（算法 2）、集成规模 M、模式 $\in$ { JustFM、EnsembleCert}。2：输出：经认证的基态构象 $\hat{\mathcal{C}}^*$。3：如果模式 = JustFM，则 4：运行一次算法 2，使用 $\mathcal{G}$ 得到 $\hat{\mathcal{C}}^*$。5：否则，如果模式 = EnsembleCert，则 6：初始化 $\mathcal{S} = \emptyset$。7：对于 m $= 1$ 到 M，执行 8：运行算法 2，使用 $\mathcal{G}$ 得到 $\mathcal{C}^{(m)}$。9：$\mathcal{S} \leftarrow \mathcal{S} \cup \{\mathcal{C}^{(m)}\}$。10：结束循环。11：$m^* \leftarrow \arg\min^{m\in\{1,\ldots,M\}} J\phi(\mathcal{C}^{(m)})$；12：$\hat{\mathcal{C}}^* \leftarrow \mathcal{C}^{(m^*)}$。13：结束判断。14：返回：预测的基态构象 $\hat{\mathcal{C}}^*$。

Name	描述	范围
手性	手性标签	{未指定，四面体顺时针与逆时针，其他}
度	成键邻居数	$\{x: 0 \leq x \leq 10, x \in \mathbb{Z}\}$
电荷	原子形式电荷	$\{x: -5 \leq x \leq 5, x \in \mathbb{Z}\}$
num_H	氢原子总数	$\{x: 0 \leq x \leq 8, x \in \mathbb{Z}\}$
自由基电子数_自由基_电子	自由基电子数	$\{x: 0 \leq x \leq 4, x \in \mathbb{Z}\}$
杂化	杂化类型	{sp, sp^2, sp^3,sp^3d, sp^3d^2, 其他}
芳香性	是否位于芳香环上	{True, False}
位于_环中	是否在环中	{True, False}

表 4：包含的原子特征。

其中 e_k^{RBF}是遵循 [80], 的K 个指数径向基函数，而 $\phi(d^{ij})$ 是一个平滑截断函数：

$$\phi(d_{ij}) = \begin{cases} \frac{1}{2}\left(\cos(\frac{\pi d_{ij}}{d_{\text{cutoff}}} + 1)\right), & \text{if } d_{ij} \leq d_{\text{cutoff}} \\ 0, & \text{otherwise} \end{cases} \tag{35}$$

最后，我们通过线性投影将所有特征合并到原子的嵌入中：

$$x_i = W^C \left[\text{embed}^{\text{int}}(z_i), h_i, t, n_i\right], \tag{36}$$

其中 t 表示时间步， $[\cdot,\cdot]$ 表示拼接。所得嵌入 $x_i \in \mathbb{R}^d$ 作为网络后续层的输入。

注意力机制： 多头点积注意力操作使用原子特征 x_i、原子属性 h_i、时间步 t 和原子间距离 d^{ij} 来计算注意力权重。输入的原子级特征 x_i 与原子属性 h_i 和时间步 t 通过多层感知机进行混合，然后使用层归一化 [81]进行归一化。为了计算注意力矩阵，将原子间距离 d^{ij} 投影为二维滤波器 D^K 和 D^V ：

$$\begin{aligned} D^K &= \sigma\left(W^{D^K} e^{RBF}(d_{ij}) + b^{D^K}\right) \\ D^V &= \sigma\left(W^{D^V} e^{RBF}(d_{ij}) + b^{D^V}\right) \end{aligned} \tag{37}$$

随后对原子级特征进行线性投影，并执行层归一化操作，以得到查询Q 向量和键K 向量。值向量V 仅通过对原子级特征进行线性投影得到。对 Q、K 向量应用层归一化（也称为 QK归一化）已证明能够在将网络扩展到大量参数时，稳定注意力矩阵 [82, 83] 中的未归一化值。随后，将Q 向量和K 向量与距离滤波器D^K结合，在特征维度上执行点积运算：

$$Q = \text{LayerNorm}(W^Q x_i), \quad K = \text{LayerNorm}(W^K x_i), \quad V = W^V x_i \tag{38}$$

$$\text{dot}(Q, K, D^K) = \sum_k^F Q_k \cdot K_k \cdot D_k^K. \tag{39}$$

将上述点积运算得到的矩阵经过非线性函数处理，并使用余弦截断函数 $\phi(d^{ij})$ （类似于嵌入层）进行加权，即可得到注意力矩阵；该函数可确保仅当两个原子之间的距离小于指定截断距离时，注意力权重才非零：

$$A = \text{SiLU}(\text{dot}(Q, K, D^K)) \cdot \phi(d_{ij}). \tag{40}$$

使用值向量V 和距离滤波器D_V，沿特征维度拆分，得到 3 个大小相同的滤波器，

$$s_{ij}^1, s_{ij}^2, s_{ij}^3 = \text{split}(V_j \cdot D_{ij}^V). \tag{41}$$

随后通过线性投影将注意力矩阵与向量s_{ij}^3 结合，得出原子级特征$y_i = W^O\left(\sum_j^N A_{ij} \cdot s_{ij}^3\right)$。注意力操作的输出包括$y_i$ （原子级特征）和两个标量滤波器s_{ij}^1 与s_{ij}^2 （边级特征）。

更新层：更新层计算注意力块中原子之间的相互作用，并利用输出更新标量特征x_i 和向量特征 $\vec{v}_i$。首先，将注意力机制的标量特征输出y_i 拆分为三个特征（q_i^1、q_i^2、q_i^3），其中q_i^1 和q_i^2 用于更新标量特征，如下所示，

$$\Delta x_i = q_i^1 + q_i^2 \cdot \langle U_1 \vec{v}_i \cdot U_2 \vec{v}_i \rangle, \tag{42}$$

其中，$\langle U_1 \vec{v}_i \cdot U_2 \vec{v}_i \rangle$ 表示向量特征 $\vec{v}_i$经矩阵 U_1,U_2线性投影后的内积。

边向量更新包含两个部分。首先，我们计算一个向量 $\vec{w}_i$，对于每个原子，该向量由所有邻居的向量特征加权和以及边向量的截断范数组成：

$$\vec{w}_i = \sum_j^N s_{ij}^1 \cdot \vec{v}_j + s_{ij}^2 \cdot \frac{\vec{r}_i - \vec{r}_j}{\max(\|\vec{r}_i - \vec{r}_j\|, \epsilon)}, \tag{43}$$

$$\Delta \vec{v}_i = \vec{w}_i + q_i^3 \cdot U_3 \vec{v}_i \tag{44}$$

其中，U_1 和 U_3 是针对向量特征 $\vec{v}_i$的特征维度的投影矩阵。在这一层中，我们将范数的最小值截断为 $\epsilon = 0.01$，以防止先验中两个原子的位置采样得过于接近时出现数值过大的情况。

$SO(3)$ **更新层：**我们还通过在式 43 中添加一个额外的叉积项，设计了一种 $SO(3)$ 等变架构，如下所示，

$$\vec{w}_i = \sum_j^N s_{ij}^1 \cdot \vec{v}_j + s_{ij}^2 \cdot \frac{\vec{r}_i - \vec{r}_j}{\max(\|\vec{r}_i - \vec{r}_j\|, \epsilon)} + s_{ij}^4 \cdot \left(\vec{v}_j \times \frac{\vec{r}_i - \vec{r}_j}{\max(\|\vec{r}_i - \vec{r}_j\|, \epsilon} \right), \tag{45}$$

其中，s_{ij}^4 是通过修改注意力层中的拆分操作（式 41）得到的，其中值向量 V 和距离滤波器 D_V 被投影到 4 个大小相等的滤波器中，而非 3 个。

输出层： 输出层由 [79] 中的门控等变块组成。给定原子标量特征 x_i 和向量特征 $\vec{v}_i$，每个块中的更新定义如下：

$$x_{i,\text{updated}}, \vec{w}_i = \text{split}(\text{MLP}([x_i, U_1 \vec{v}_i])) \tag{46}$$

$$\vec{v}_{i,\text{updated}} = (U_2 \vec{v}_i) \cdot \vec{w}_i \tag{47}$$

Here，U_1 和U_2 是沿特征维度作用的线性投影矩阵。我们的修改 n
is t是在多层感知机中使用层归一化，以提高训练稳定性。

C.2 输入特征化

遵循 ET-Flow [45], , 使用 RDKit [47] 描述符计算原子（节点）特征，详见表 4。为构建边特征和边索引，我们结合使用全局的基于半径的边与局部分子图边，这与参考文献 [44] 中的方法类似。

D 更多细节与实验

D.1 实现细节

D.1 Reflow 技术

继 AvgFlow [30], 采用 Reflow 技术 [18]，通过单个常微分方程步骤提升构象采样质量之后，我们也在实验中实现了 Reflow 流程。具体而言，完成主要训练阶段后，我们首先从谐波先验分布中，为训练集和验证集中的每个分子随机采样固定数量的先验样本 $\mathcal{C}_0'$。随后，我们采用固定数量的常微分方程步骤生成相应的构象样本 $\mathcal{C}_1'$。最后，使用耦合对 $(\mathcal{C}_0', \mathcal{C}_1')$，通过 Reflow 损失对训练好的模型进行进一步微调：

$$\mathcal{C}_t' = (1-t)\mathcal{C}_0' + t\mathcal{C}_1' + \sigma\sqrt{t(1-t)} \cdot \epsilon, \quad \epsilon \sim \mathcal{N}(0, I). \tag{48}$$

$$v_t(\mathcal{C}_t' | \mathcal{C}_0', \mathcal{C}_1') = \frac{1-2t}{2t(1-t)} (\sigma\sqrt{t(1-t)} \cdot \epsilon) + (\mathcal{C}_1' - \mathcal{C}_0'). \tag{49}$$

$$\mathcal{L}_{Reflow} = \mathbb{E}_{\mathcal{C}_t' \sim p_t(\mathcal{C}_t' | \mathcal{C}_0', \mathcal{C}_1'), t \sim \mathcal{U}(0,1)} \left[\|v_\theta(\mathcal{C}_t', t) - v_t(\mathcal{C}_t' | \mathcal{C}_0', \mathcal{C}_1')\|^2 \right]. \tag{50}$$

	1 个常微分方程步骤	2 个常微分方程步骤	5 个常微分方程步骤	50 个常微分方程步骤
GEOM-Drugs	$0.5(1-t)^2$	$0.3(1-t)^2$	$0.2(1-t)^2$	$0.1(1-t)^2$
GEOM-QM9	$0.4(1-t)^2$	$0.3(1-t)^2$	$0.2(1-t)^2$	$0.1(1-t)^2$

表 5：不同 ODE 采样步数下的引导调度 λ_t 。

值得注意的是，我们仅对向量场模型 v_θ（即无引导路径）应用 Reflow 技术。完成重流微调后，我们仍像之前一样，使用 Alg. 2 中描述的相同引导采样算法进行推理，但采用经过 Reflow 的新向量场模型。

D.1.2 训练细节

对于 GEOM-Drugs，在第一个匹配阶段，EnFlow 使用 6 张 NVIDIA A100 GPU 训练 500 个 epoch，批次大小为 24，每个 epoch 最多 5,000 个批次。在第二个能量微调阶段，EnFlow 再次训练 500 个 epoch，批次大小为 48，每个 epoch 最多 5,000 个批次，使用 4 张 NVIDIA A100 GPU。两个阶段均使用 AdamW 优化器和余弦退火学习率调度，学习率范围为 1×10^{-8} 至 5×10^{-4}。对于 GEOM-QM9，在第一个匹配阶段，EnFlow 使用 8 张 NVIDIA A100 GPU 训练 1,000 个 epoch，批次大小为 48，每个 epoch 最多 1,000 个批次。在第二个能量微调阶段，EnFlow 训练 1,000 个 epoch，批次大小为 64，每个 epoch 同样最多 1,000 个批次，使用 4 张 NVIDIA A100 GPU。对于 GEOM-QM9，我们同样使用 AdamW 优化器和余弦退火学习率调度，学习率范围为 1×10^{-8} 至 7×10^{-4}。对于两个数据集，均根据最低验证损失选择最佳检查点。

D.1.3 ET-Flow 的复现问题

由于 EnFlow 基于 ET-Flow [45], , 我们已尽一切努力复现 ET-Flow 在 GEOM-Drugs 和 GEOM-QM9 上报告的结果。然而，如正文中的表1和表2所示，我们复现的结果与原报告结果之间存在明显的性能差距，尤其是在 GEOM-Drugs 上。ET-Flow 的官方 GitHub 仓库（ https://github.com/shenoynikhil/ETFlow.git）中提出了若干相关问题，作者也已承认这些问题。我们感谢作者在复现过程中给予的积极沟通与支持，并将对这一差距进行更深入的调查作为未来工作的开放问题。

D.1.4 引导调度的选择

根据公式29的表述，我们使用引导向量场$v'_t(\mathcal{C}_t)$进行能量引导采样。引导强度 λ_t 被设为一个时间相关衰减因子，并在$t \to 1$时趋于零。在实验中，我们采用了一个简单的二次调度，如表5所示。关于这些选择的更多讨论见第 D.2.1 节。

D.2 更多实验结果

D.2.1 关于引导调度的观察

在 D.1.4节中，我们概述了实验中采用的引导调度。一个自然的问题是，为何在 5 步和 50 步常微分方程采样器中选择相对较小的引导强度。为论证这一设计选择，我们在两种设置下，对一系列引导强度 λ_t 进行了消融研究。定量结果和指标曲线汇总于图 5。

消融结果表明，引导强度的大小决定了明显的权衡。随着引导强度增加，召回率指标（COV-R 和 AMR-R）持续下降，而精确率导向指标（COV-P 和 AMR-P）则大幅提升，且在较小的 RMSD 阈值 δ 下提升最为显著。同时，平均预测能量 $J\phi$ 显著降低，表明更强的引导会使采样过程偏向构象景观中的低能量区域。因此，高能量构象所占比例会减少，而生成结构则与低能量参考构象更加一致。

这一现象表明，增大引导强度会使采样器更加保守：它会优先生成能量上更有利的结构，从而提高精确率，但代价是结构多样性降低，导致在不同 RMSD 阈值下的召回率下降。由于

我们的首要目标是仅用极少的 ODE 步数获得高质量构象，因此在 5 步和 50 步设置下，我们采用中等引导强度。这样可以在召回率和精确率之间取得更均衡的权衡，从而能够与现有基线方法进行公平比较，同时仍能发挥能量引导采样的优势。

D.2.2 为什么能量模型需要进行能量匹配训练

由于我们的主要目标是利用能量引导向量场 $v'_t(\mathcal{C}_t, t)$ 生成引导概率路径 $p'_t(\mathcal{C}_t)$，而不是依赖生成无引导路径 $p_t(\mathcal{C}_t)$ 的原始无引导向量场 $v_\theta(\mathcal{C}_t, t)$，因此，学习到的能量模型必须提供准确且校准良好的梯度。两个向量场都将同一个源分布（谐波先验 $p_0(\mathcal{C}_0)$）输运至不同的目标：引导情形下的 $p_1(\mathcal{C}_1)\, e^{-J_\phi(\mathcal{C}_1)}$，以及无引导情形下的 $p_1(\mathcal{C}_1)$。虽然能量回归损失 $\mathcal{L}_{\text{energy}}$ 促使模型 $J_\phi(\mathcal{C})$ 逼近绝对能量，却无法保证能量差得到保留，而能量差对于产生正确的梯度方向至关重要。相比之下，能量匹配损失 $\mathcal{L}_{\text{em}}$ 会显式对齐预测的能量差与真实能量差，从而为调和先验样本提供忠实的梯度，引导其趋向低能构象。

图 5 的实证结果证实了能量匹配 [42] 的必要性。对于 2 步和 5步常微分方程采样器，仅使用无引导向量场$v(\mathcal{C}_t,t)$ 会带来较高的召回率（COV-R 和 AMR-R），但精确率（COV-P 和 AMR-P）明显较低，表明能量一致性较差。通过$v'_t(\mathcal{C}_t,t)$ 添加引导且仅使用 $\mathcal{L}_{\text{energy}}$，只能带来有限的提升，反映出梯度方向不可靠。相比之下，使用 $\mathcal{L}_{\text{energy}}$ 和 $\mathcal{L}_{\text{em}}$训练能量模型，则能持续提升所有指标：面向召回率和精确率的分数同步提高，且不同 RMSD 阈值下的曲线均呈现明显而稳定的提升。这些结果表明，能量匹配 对于获得可靠的能量梯度以及实现有效的能量引导 ODE 采样至关重要。

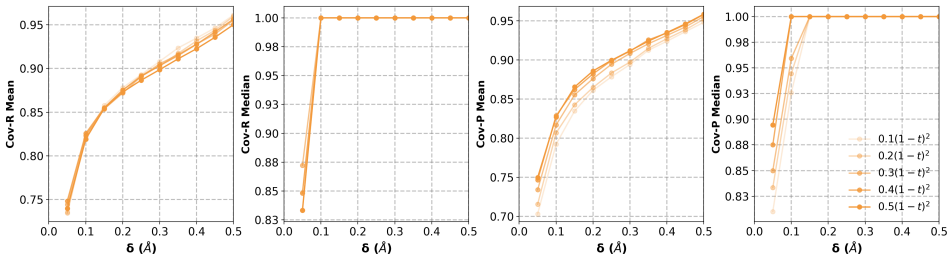
D.2.3 采样向量场选择的消融实验

为进一步验证我们提出的引导向量场v'的稳健性，我们通过消融研究比较了用于 ODE 采样的三种向量场：（1）无引导向量场$v_\theta(\mathcal{C}_t,t)$，对应于 ET-Flow [29]中的纯流匹配框架；（2）纯能量梯度场 $-\nabla_{\mathcal{C}_t} J_\phi(\mathcal{C}_t)$，表示参考文献 [18] 中的直线路径流匹配形式，其向量场与时间无关，或者等价地说，是无噪声的确定性朗之万动力学 [84]（即梯度下降）；以及（3）我们提出的能量引导向量场$v_\theta(\mathcal{C}_t,t) - \lambda_t \nabla_{\mathcal{C}_1} J_\phi(\hat{\mathcal{C}}_1)$。定量结果和指标轨迹见图 7。

图 7 中的消融结果一致表明，在不同数据集和 ODE采样预算下，我们提出的能量引导向量场均具有优越性。与无引导的流匹配向量场 $v_\theta(\mathcal{C}_t,t)$相比，我们的方法取得了显著更低的 AMR-R 和 AMR-P 值，表明生成的构象与参考结构更加接近。同时，覆盖率指标（COV-R 和 COV-P）保持相当水平或略有提升，说明引入能量引导并未损害多样性。相比之下，纯能量梯度场 $-\nabla_{\mathcal{C}_t} J_\phi(\mathcal{C}_t)$ 在覆盖率和准确性方面均出现明显退化，证实仅依赖能量基模型会导致过度平滑和较差的生成行为。值得注意的是，我们的方法始终产生低得多的平均预测能量 J_ϕ，体现了其将样本引导至具有物理意义的低能量区域的能力。在 GEOM-QM9 上进行两步和五步采样，以及在 GEOM-Drugs 上进行五步采样时，这些趋势依然保持不变，表明所提出的引导向量场能够提升基于ODE的采样的稳定性和保真度，尤其是在少步数场景下。

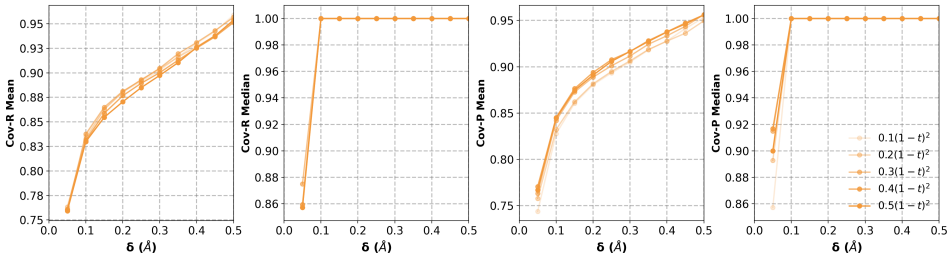
(a) GEOM-QM9 上的 5步常微分方程采样

λ_t	COV-R ↑		AMR-R ↓		COV-P ↑		AMR-P ↓		J_ϕ ↓
	mean	中位数	mean	中位数	mean	中位数	mean	中位数	mean
$0.1(1-t)^2$	96.11	100.00	0.084	0.036	94.82	100.00	0.098	0.044	0.880
$0.2(1-t)^2$	95.97	100.00	0.083	0.037	95.22	100.00	0.093	0.042	0.611
$0.3(1-t)^2$	95.53	100.00	0.084	0.032	95.56	100.00	0.085	0.038	0.406
$0.4(1-t)^2$	95.64	100.00	0.083	0.033	95.83	100.00	0.081	0.033	0.243
$0.5(1-t)^2$	95.01	100.00	0.086	0.035	95.86	100.00	0.081	0.032	0.148



(b) 在 GEOM-QM9 上进行 50 个常微分方程步骤采样

λ_t	COV-R ↑		AMR-R ↓		COV-P ↑		AMR-P ↓		J_ϕ ↓
	mean	中位数	mean	中位数	mean	中位数	mean	中位数	mean
$0.1(1-t)^2$	95.89	100.00	0.080	0.032	95.33	100.00	0.085	0.034	0.166
$0.2(1-t)^2$	95.23	100.00	0.080	0.031	94.96	100.00	0.084	0.029	0.097
$0.3(1-t)^2$	95.67	100.00	0.077	0.030	95.64	100.00	0.078	0.028	0.051
$0.4(1-t)^2$	95.13	100.00	0.081	0.030	95.60	100.00	0.076	0.026	0.009
$0.5(1-t)^2$	95.39	100.00	0.083	0.031	95.61	100.00	0.078	0.031	-0.023



(c) 在 GEOM-Drugs 上进行 5步常微分方程采样

λ_t	COV-R ↑		AMR-R ↓		COV-P ↑		AMR-P ↓		J_ϕ ↓
	mean	中位数	mean	中位数	mean	中位数	mean	中位数	mean
$0.1(1-t)^2$	78.62	84.21	0.492	0.474	68.49	73.82	0.620	0.565	9.197
$0.2(1-t)^2$	77.24	82.31	0.499	0.479	69.97	76.47	0.607	0.541	7.254
$0.3(1-t)^2$	76.36	80.86	0.508	0.485	71.15	77.57	0.598	0.530	5.753
$0.4(1-t)^2$	75.74	80.00	0.516	0.496	72.23	78.92	0.916	0.515	4.631
$0.5(1-t)^2$	74.47	78.15	0.530	0.502	72.51	79.57	0.605	0.510	3.811

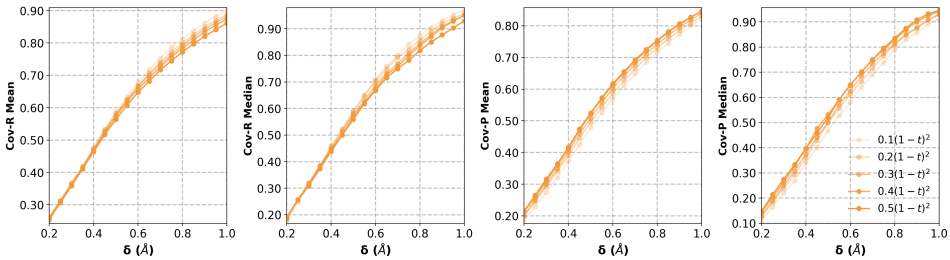
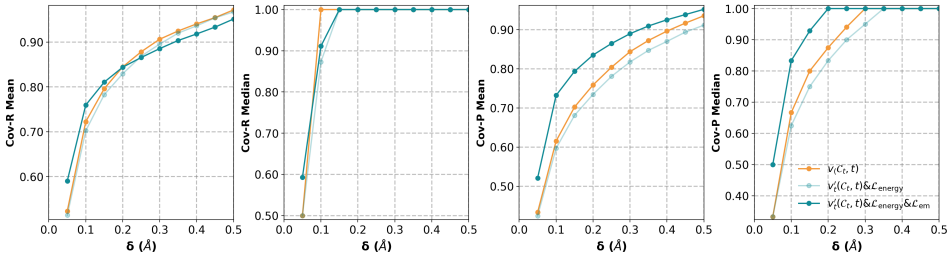


图5：在 GEOM-QM9 上进行 5 步 (a) 和 50 步 (b) ODE采样，以及在 GEOM-Drugs 上进行 5 步 (c) ODE采样时，不同引导强度 λ_t 的消融研究。表格报告了 GEOM-QM9 在固定均方根偏差阈值 $\delta = 0.5$ 埃、GEOM-Drugs 在固定均方根偏差阈值 $\delta = 0.75$ 埃时的召回率和精确率指标，以及平均预测能量 $J\phi$。图中展示了这些指标随均方根偏差阈值 δ的变化情况。

(a) 在 GEOM-QM9 上进行 2 个常微分方程步骤的采样

方法	COV-R ↑		AMR-R ↓		COV-P ↑		AMR-P ↓	
	mean	中位数	mean	中位数	mean	中位数	mean	中位数
$v_(\mathcal{C}_t, t)$	97.18	100.00	0.107	0.062	93.54	100.00	0.151	0.107
$v't(\mathcal{C}_t,t)$ & $\mathcal{L}_{\text{energy}}$	96.73	100.00	0.112	0.068	91.14	100.00	0.174	0.122
$v't(\mathcal{C}_t,t)$ & $\mathcal{L}_{\text{energy}}$ & $\mathcal{L}_{\text{em}}$	95.14	100.00	0.109	0.060	95.18	100.00	0.118	0.067



（b）GEOM-QM9 上的 5步常微分方程采样

方法	COV-R ↑		AMR-R ↓		COV-P ↑		AMR-P ↓	
	mean	中位数	mean	中位数	mean	中位数	mean	中位数
$v_(\mathcal{C}_t, t)$	96.02	100.00	0.084	0.038	94.48	100.00	0.103	0.053
$v't(\mathcal{C}_t,t)$ & $\mathcal{L}_{\text{energy}}$	95.82	100.00	0.084	0.037	94.59	100.00	0.100	0.048
$v't(\mathcal{C}_t,t)$ & $\mathcal{L}_{\text{energy}}$ & $\mathcal{L}_{\text{em}}$	96.28	100.00	0.081	0.037	95.39	100.00	0.088	0.038

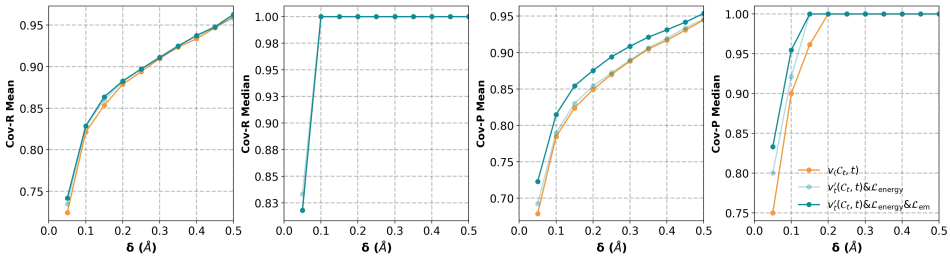
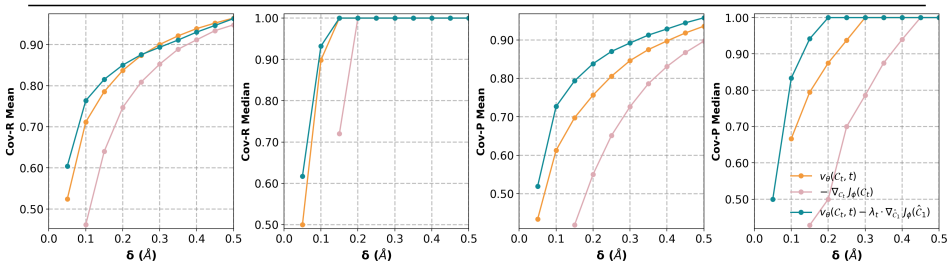


图6：关于能量匹配 训练必要性的消融研究。对于 GEOM-QM9 上的 2 步（a）和 5 步（b）常微分方程采样，表格报告了在固定均方根偏差阈值 $\delta = 0.5$ 埃下的召回率和精确率指标，图中则展示了这些指标随均方根偏差阈值 δ 的变化情况。

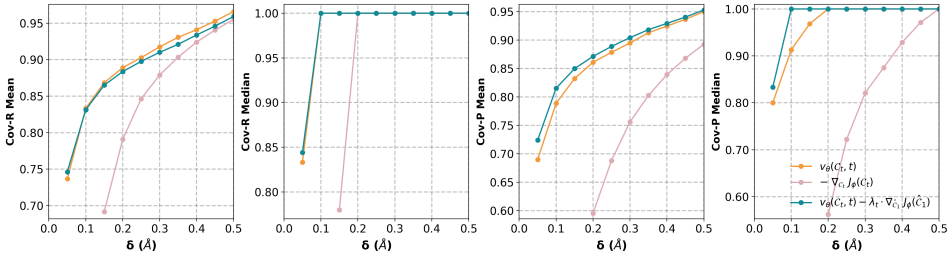
（a）在 GEOM-QM9 上进行 2 个常微分方程步数采样

向量场	COV-R ↑		AMR-R ↓		COV-P ↑		AMR-P ↓		J_ϕ ↓
	mean	中位数	mean	中位数	mean	中位数	mean	中位数	mean
$v_\theta(\mathcal{C}_t, t)$	96.46	100.00	0.110	0.066	93.56	100.00	0.150	0.110	10.182
$-\nabla_{\mathcal{C}_t} J\phi(\mathcal{C}_t)$	94.82	100.00	0.167	0.130	89.67	100.00	0.246	0.208	36.150
$v_\theta(\mathcal{C}_t, t) - \lambda_t \cdot \nabla_{\mathcal{C}_t} J\phi(\hat{\mathcal{C}}_1)$	96.33	100.00	0.103	0.055	95.74	100.00	0.116	0.066	5.579



(b) 在 GEOM-QM9 上进行 5步常微分方程采样

向量场	COV-R ↑		AMR-R ↓		COV-P ↑		AMR-P ↓		J_ϕ ↓
	mean	中位数	mean	中位数	mean	中位数	mean	中位数	mean
$v_\theta(\mathcal{C}_t, t)$	96.58	100.00	0.079	0.037	95.00	100.00	0.097	0.048	1.218
$-\nabla_{\mathcal{C}_t} J\phi(\mathcal{C}_t)$	95.50	100.00	0.148	0.116	89.30	100.00	0.246	0.204	16.338
$v_\theta(\mathcal{C}_t, t) - \lambda_t \cdot \nabla_{\mathcal{C}_t} J\phi(\hat{\mathcal{C}}_1)$	95.92	100.00	0.082	0.034	95.32	100.00	0.089	0.038	0.608



（c）在 GEOM-Drugs 上进行 5步常微分方程采样

向量场	COV-R ↑		AMR-R ↓		COV-P ↑		AMR-P ↓		J_ϕ ↓
	mean	中位数	mean	中位数	mean	中位数	mean	中位数	mean
$v_\theta(\mathcal{C}_t, t)$	78.88	85.71	0.488	0.463	66.73	71.78	0.643	0.580	12.174
$-\nabla_{\mathcal{C}_t} J\phi(\mathcal{C}_t)$	73.62	80.22	0.574	0.554	53.68	51.68	3.451	0.760	81.589
$v_\theta(\mathcal{C}_t, t) - \lambda_t \cdot \nabla_{\mathcal{C}_t} J\phi(\hat{\mathcal{C}}_1)$	77.69	83.33	0.497	0.469	70.41	76.01	0.599	0.531	7.244

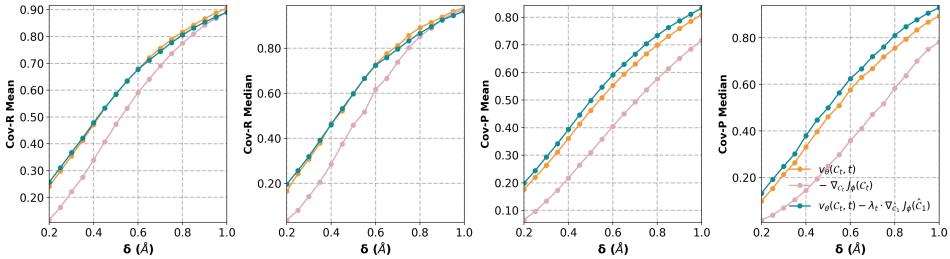


图 7：针对 GEOM-QM9 上 2 步（a）和 5 步（b）常微分方程采样，以及 GEOM-Drugs 上 5 步（c）常微分方程采样，对不同类型向量场进行消融研究。表格报告了 GEOM-QM9 在固定均方根偏差阈值 $\delta = 0.5$ 埃、GEOM-Drugs 在固定均方根偏差阈值 $\delta = 0.75$ 埃时的召回率和精确率指标，以及平均预测能量 $J\phi$。图中展示了这些指标随均方根偏差阈值 δ 的变化情况。

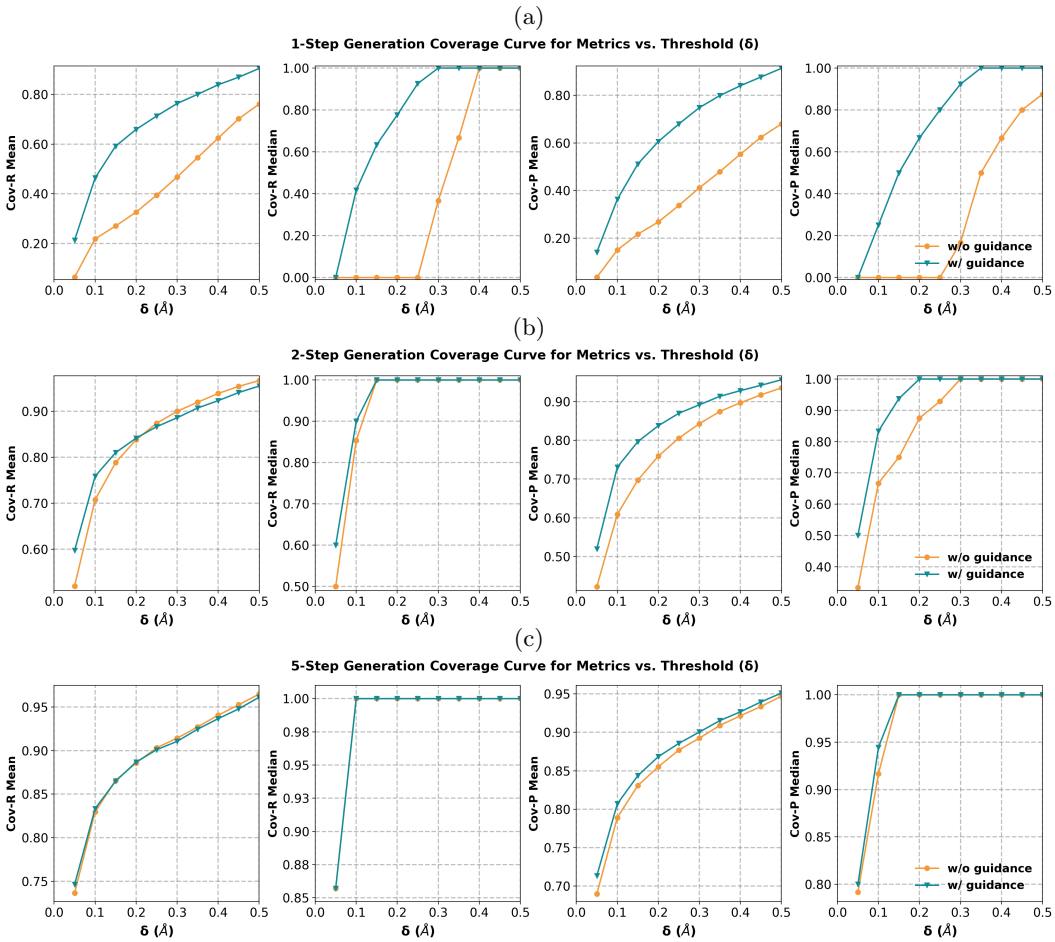


图 8：在 GEOM-QM9 数据集上进行常微微分方程采样时，覆盖率（%）与阈值 δ 的消融研究，分别采用 1 (a)、2 (b) 和 5 (c) 步数，并比较无引导基线（ET-Flow；**无引导**）与引导模型（**EnFlow**；**有引导**）。

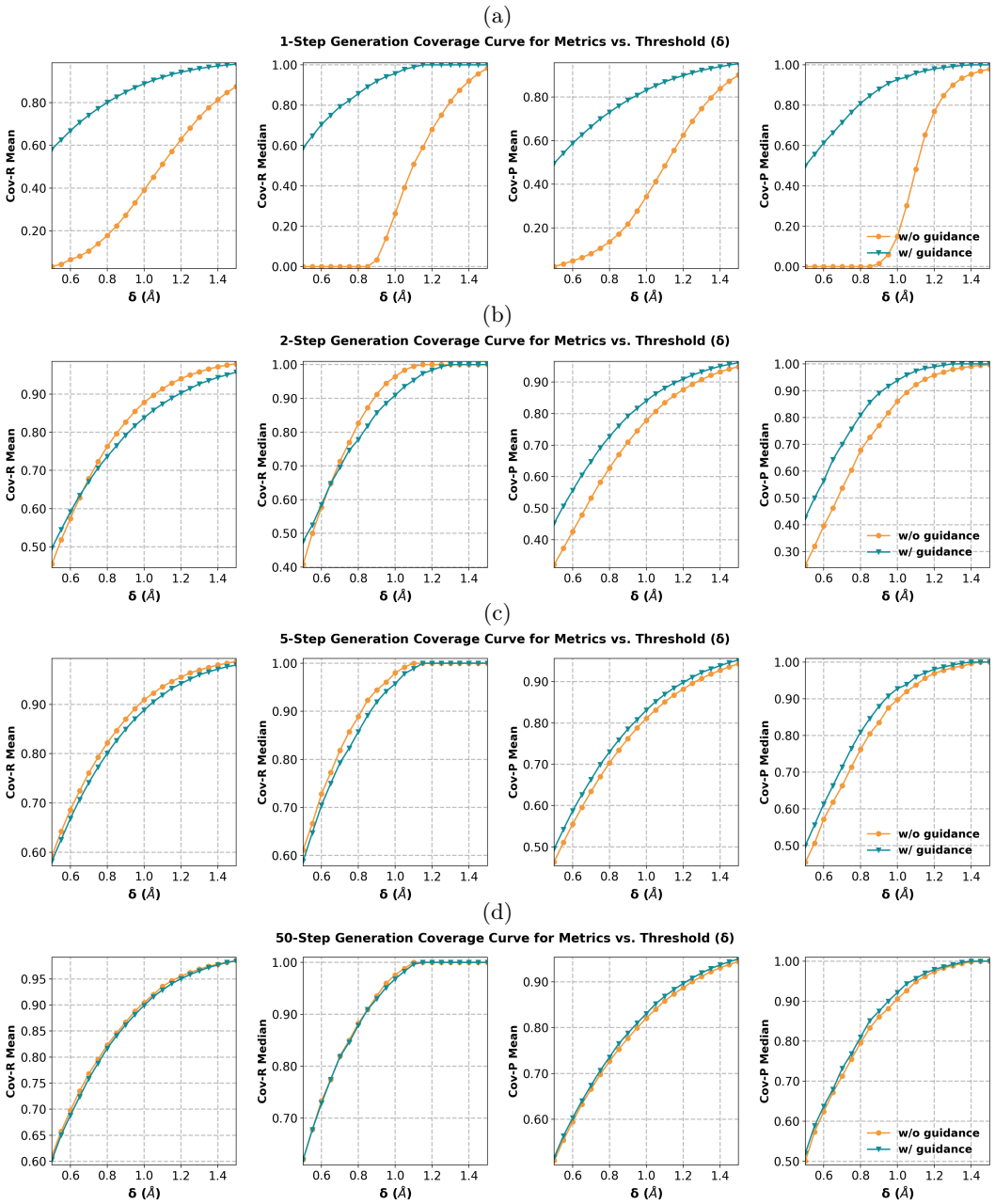


图9：在 GEOM-Drugs 数据集上，对采用 1（a）、2（b）、5（c）和 50（d）步的 ODE采样进行覆盖率（%）与阈值 δ 的消融研究，比较无引导基线（ET-Flow；**无引导**）与引导模型（**EnFlow**；**有引导**）。

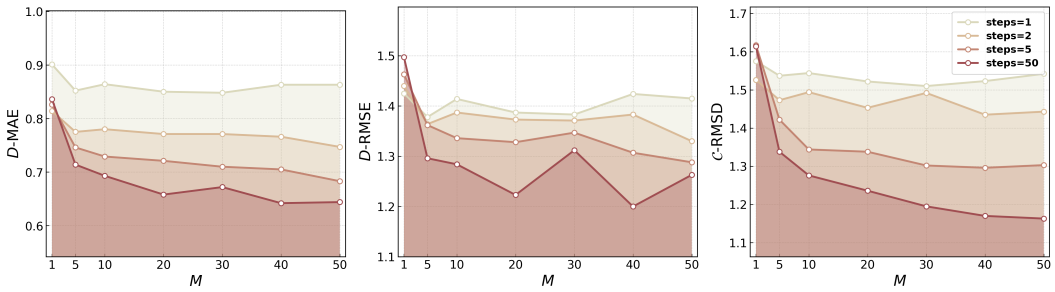


图 10：在 GEOM-Drugs 数据集上，针对基态构象预测的 EnsembleCert 模式消融实验。在 1、2、5 和 50 个 ODE 采样步数下，集成规模 $M = 1$、5、10、20、50 的影响。从左到右：**D**-MAE（埃）、**D**-RMSE（埃）、 $\mathcal{C}$-RMSD（埃）。

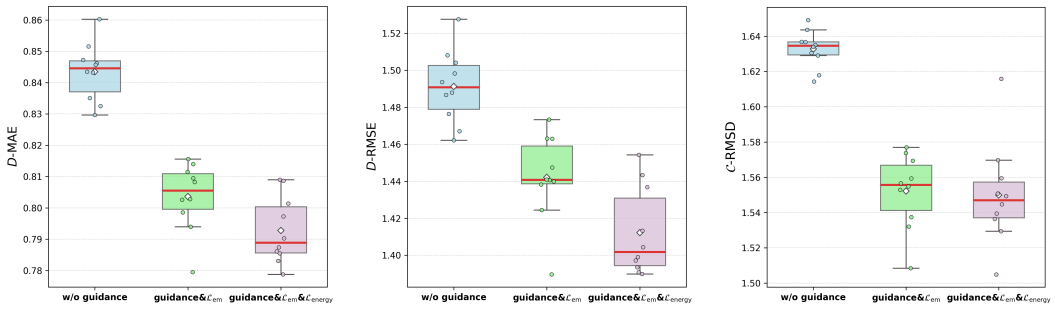


图 11：采用 JustFM 模式和 5 步 ODE 采样时，三种设置下基态构象预测性能的箱线图：**(1)** 无引导基线（ET-Flow；**无引导**）；**(2)** 仅进行能量匹配的引导模型（**引导**& $\mathcal{L}_{\text{em}}$）；**(3)** 同时进行能量匹配和能量微调的完全引导模型（**EnFlow**；**引导**& $\mathcal{L}_{\text{em}}$ &$\mathcal{L}_{\text{energy}}$）。

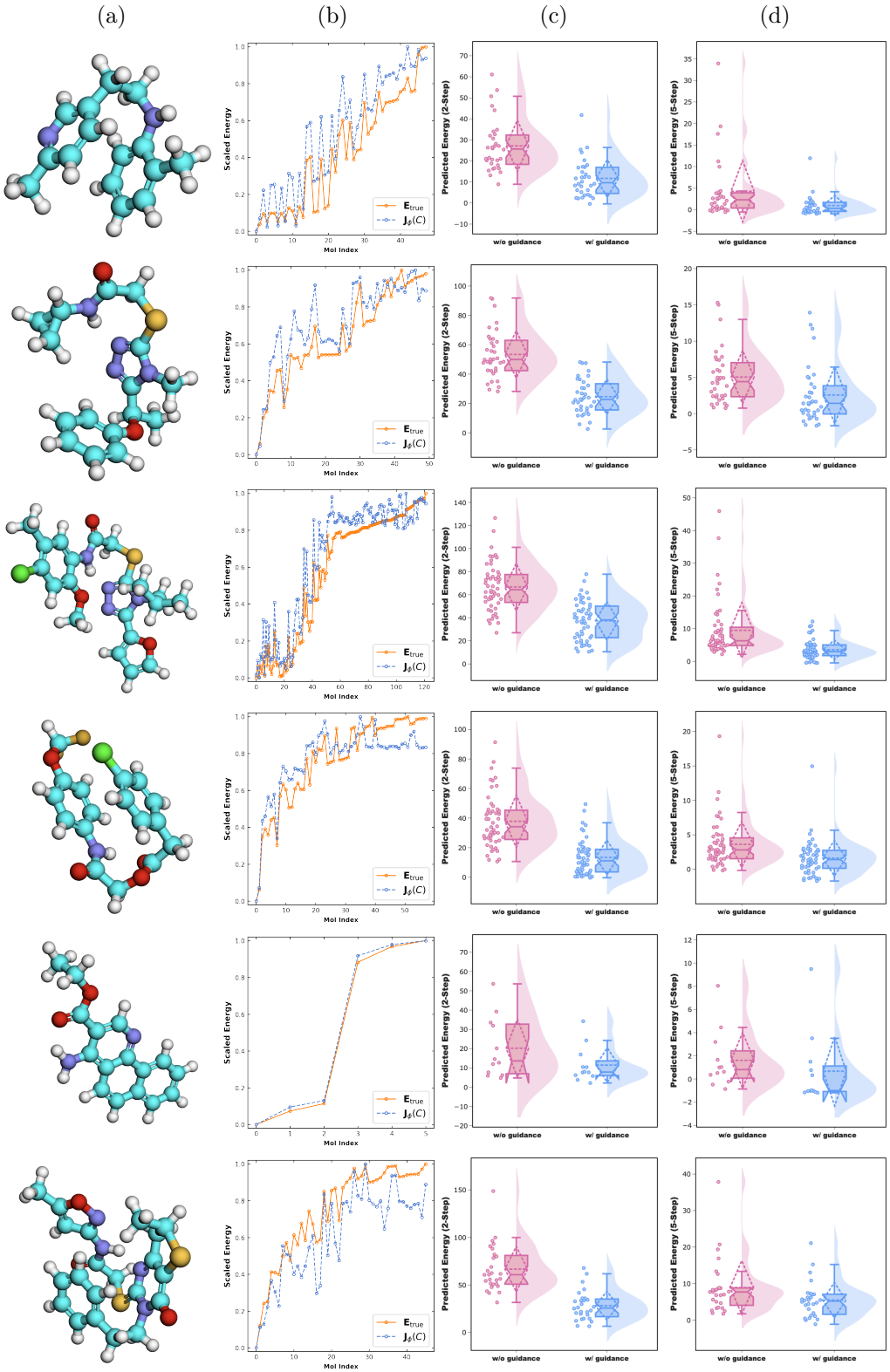


图12：（a）GEOM-Drugs 数据集中的六个代表性分子。（b）它们的能量景观（默认按数据集中的真实玻尔兹曼权重对分子排序）：$\mathbf{E}_{\text{true}}$ 表示通过高水平量子化学方法计算得到的数据集真实能量值，$\mathbf{J}_\phi(C)$ 表示我们的能量基模型 $\mathbf{J}_\phi$所学习到的能量景观。二者都归一化到范围 $[0,\,1]$，以便比较其变化。（c）–（d）对于每个分子，沿两条采样轨迹生成的构象的预测能量，采样步数分别为 2 步（c）和 5 步（d）。无引导基线（ET-Flow；**无引导**）以红色显示，引导模型（**EnFlow**；**有引导**）以蓝色显示。34