Phase 1 综合实验报告

Trajectory-Free Loss 有效性验证

项目：从无标签数据学习交互粒子系统

Viska Wei

2026 年 1 月 28 日

实验结论

Gate-1 验证失败。弱形式 loss 存在**根本性的 identifiability 问题**，无法仅从分布演化学习势函数。

目录

1 研究背景与问题定义

1.1 研究动机

学习高维交互粒子系统的动力学是多个科学领域的基础任务。然而，在实际应用中常见的挑战是：数据仅包含在离散时间点采集的**无标签集合数据**（unlabeled ensemble data），缺少轨迹信息。这种数据缺失可能源于数据采集方法的限制或隐私保护需求。

1.2 数学模型

考虑由 N 个粒子组成的交互粒子系统（IPS），其动力学由以下随机微分方程（SDE）描述：

$$dX_t^i = -\frac{1}{N}\sum_{j\neq i}^{N} \nabla\Phi(X_t^i - X_t^j)dt + \nabla V(X_t^i)dt + \sigma dW_t^i, \quad i = 1, \ldots, N \tag{1}$$

其中：

- $X_t^i \in \mathbb{R}^d$ 是第 i 个粒子在时刻 t 的位置
- $\Phi : \mathbb{R}^d \to \mathbb{R}$ 是**交互势函数**（interaction potential）
- $V : \mathbb{R}^d \to \mathbb{R}$ 是**外势函数**（kinetic/external potential）
- W_t^i 是独立的标准布朗运动
- $\sigma > 0$ 是噪声强度

1.3 数据形式

观测数据由一系列时间快照的样本集合组成：

$$\mathcal{D} = \{X_{t_\ell}^{1:M}\}_{\ell=1}^{L}, \quad \text{其中} \quad X_t = (X_t^1, \ldots, X_t^N) \tag{2}$$

关键约束：数据缺少轨迹信息——位置 $X_{t_\ell}^{i,m}$ 和 $X_{t_{\ell+1}}^{i,m}$ 在数据中**不配对**，因为标签未知或它们可能来自不同的轨迹。

1.4 研究目标

目标：从无标签集合数据 $\mathcal{D}$ 中学习交互势函数 Φ 和外势函数 V。

挑战：传统方法（MLE、MSE）需要轨迹信息来计算 dX_t^i/dt，无法直接应用于无标签数据。

2 方法论：Trajectory-Free Loss

2.1 弱形式 PDE

论文提出使用经验分布的弱形式 PDE 来构建不需要轨迹的损失函数。设 $\mu_t^N(x) := \frac{1}{N}\sum_{i=1}^{N}\delta_{X_t^i - x}$ 为粒子在时刻 t 的经验分布，其演化满足：

$$\partial_t \mu_t^N = -\nabla \cdot \left[\mu_t^N \nabla(\Phi * \mu_t^N + V)\right] + \frac{\sigma^2}{2}\Delta\mu_t^N + \sigma m_{X_t} \tag{3}$$

2.2 原论文的损失函数

论文提出的 trajectory-free loss 形式为：

$$\mathcal{E}_{\mathcal{D}}(\Phi, V) = \frac{1}{M}\sum_{m,\ell} \mathcal{E}_{X^m_{t_\ell}, X^m_{t_{\ell+1}}}(\Phi, V) \tag{4}$$

其中每对时间快照的损失为：

$$\mathcal{E}_{X_{t_\ell}, X_{t_{\ell+1}}} = \underbrace{J_{\text{diss}}\Delta t}_{\text{耗散项}} + \underbrace{\sigma J_{\text{lap}}\Delta t}_{\text{扩散项}} - \underbrace{2\Delta E}_{\text{能量变化}} \tag{5}$$

各项的具体计算：

$$J_{\text{diss}} = \frac{1}{N}\sum_{i=1}^{N}\left|\nabla V(X_t^i) + \frac{1}{N}\sum_{j=1}^{N}\nabla\Phi(X_t^i - X_t^j)\right|^2 \tag{6}$$

$$J_{\text{lap}} = \frac{1}{N^2}\sum_{i,j=1}^{N}\left[\Delta V(X_t^i) + \Delta\Phi(X_t^i - X_t^j)\right] \tag{7}$$

$$E(t) = \frac{1}{N}\sum_{i=1}^{N} V(X_t^i) + \frac{1}{2N^2}\sum_{i,j}\Phi(X_t^i - X_t^j) \tag{8}$$

2.3 我们发现的公式错误

重要发现：Loss 公式系数错误

通过 Itô 引理重新推导，我们发现原论文公式 (5) 存在系数错误。

原公式：$\mathcal{L} = J_{\text{diss}} + \sigma \cdot J_{\text{lap}} - 2 \cdot dE$

正确公式：$R = J_{\text{diss}} - \frac{\sigma^2}{2} \cdot J_{\text{lap}} + dE = 0$

项	原公式	正确公式
Laplacian 系数	$+\sigma$	$-\sigma^2/2$
能量项系数	-2	$+1$

3 实验设计与实现

3.1 实验配置

3.2 神经网络实现

使用 PyTorch 实现势函数的神经网络表示：

- **V 网络**：$V_\theta : \mathbb{R}^d \to \mathbb{R}$，标准 MLP
- **Φ 网络**：$\Phi_\eta : \mathbb{R}^d \to \mathbb{R}$，强制对称性 $\Phi_\eta(x) = \frac{1}{2}[\tilde{\Phi}_\eta(x) + \tilde{\Phi}_\eta(-x)]$
- **导数计算**：使用自动微分（AD）计算 ∇V, $\nabla\Phi$, ΔV, $\Delta\Phi$

表 1: 实验参数配置

参数	**值**	**说明**
维度 d	1	简化问题，验证方法可行性
粒子数 N	5–10	小规模系统
时间快照 L	10–20	离散时间点数量
样本数 M	30–100	每个时间点的独立样本数
噪声强度 σ	0.1	中等噪声水平
时间步长 dt	0.01	SDE 模拟步长
总时间 T	2.0	模拟总时长
外势 V	$V(x) = 0.5x^2$	简谐势（Harmonic）
交互势 Φ	$\Phi(r) = e^{-r^2/2}$	高斯排斥势
网络架构	MLP [64, 64]	两层全连接网络
激活函数	Tanh	平滑激活
优化器	Adam	学习率 0.001

3.3 评估指标

相对 L^2 误差：

$$\text{Error}(\Phi) = \sqrt{\frac{\sum_{l,m,i,j} |\Phi_{\hat{\eta}}(r_{lijm}) - \Phi_{\text{true}}(r_{lijm})|^2}{\sum_{l,m,i,j} |\Phi_{\text{true}}(r_{lijm})|^2}} \tag{9}$$

通过标准：$\text{Error}(V) < 10\%$ 且 $\text{Error}(\Phi) < 10\%$

4 实验结果

4.1 MVP-0.0: SDE 数据生成验证 PASS

目标：验证 Euler-Maruyama 模拟器的正确性。

方法：对 Ornstein-Uhlenbeck 过程（$V = 0.5x^2$, $\Phi = 0$），比较数值解与解析解。

表 2: SDE 验证结果

指标	**测量值**	**阈值**	**状态**
KL 散度	0.0005	< 0.01	PASS
方差相对误差	0.42%	$< 5\%$	PASS

结论：数据生成正确，可进行后续实验。

4.2 MVP-1.0: Joint V+Φ 学习 FAIL

目标：验证 trajectory-free loss 能否同时学习 V 和 Φ。

表 3: MVP-1.0 实验结果

配置	**V Error**	**Φ Error**	**Final Loss**	**Epochs**
hidden=[32,32], lr=0.01	162.31%	19.30%	~ 0	31
hidden=[64,64], lr=0.001	98.72%	94.10%	1.3×10^{-7}	32

关键发现

- Loss 快速收敛到 ~ 0（触发 early stopping）
- 但 V 和 Φ 形状完全错误
- V 学成了**凹函数**（真实是凸函数）
- **结论**：Loss→0 **不保证**正确解

4.3 MVP-1.0b: Φ-only 学习（已知 V）FAIL

目标：消除 V-Φ trade-off，固定 V 为真实值，仅训练 Φ。

表 4: MVP-1.0b 实验结果

配置	**Φ Error**	**Final Loss**	**Epochs**
N=5, L=10, M=30	78.13%	0.018	100

关键观察：

- 即使 V 已知，Φ 仍学不对
- Loss 没有收敛到 0（停在 0.018）
- 说明问题不仅是 V-Φ trade-off，loss 公式本身可能有问题

4.4 MVP-1.1: 添加 Identifiability 约束 FAIL

目标：通过添加约束解决 identifiability 问题。

约束设计：

1. $V(0) = 0$ —锚定 V 在原点
2. $\Phi(r_{\text{ref}}) = 0$ —Φ 在参考距离为 0
3. 梯度范数正则化

实现：

$$V(x) = V_{\text{raw}}(x) - V_{\text{raw}}(0) \tag{10}$$

$$\Phi(r) = \Phi_{\text{raw}}(r) - \Phi_{\text{raw}}(r_{\text{ref}}) \tag{11}$$

结论：简单约束无效。问题不在势函数的常数项，而在形状（梯度）。

表 5: MVP-1.1 实验结果

配置	V Error	Φ Error	Final Loss
N=10, L=20, M=100	110.28%	88.89%	~ 0

4.5 MVP-1.2: Loss 公式验证 PASS

目标：从理论推导检验 loss 公式的正确性。

方法：

1. 从 Itô 引理重新推导弱形式公式
2. 用真实势函数计算各项，验证残差

验证结果：残差随 $dt \to 0$ 线性收敛到 0。

表 6: 残差收敛验证

dt	Mean Residual
0.20	1.57×10^{-1}
0.10	7.65×10^{-2}
0.05	3.76×10^{-2}
0.02	1.49×10^{-2}

结论：正确公式为 $R = J_{\text{diss}} - \frac{\sigma^2}{2} J_{\text{lap}} + dE = 0$。

4.6 MVP-1.2b: 修正公式训练 FAIL

目标：用修正后的公式重新训练。

表 7: MVP-1.2b 实验结果

配置	V Error	Φ Error	Final Loss
修正公式	97.62%	128.35%	~ 0

结论：修正公式后仍然失败。问题是根本性的 identifiability，不是公式 bug。

5 结果汇总

6 根因分析

6.1 为什么弱形式 loss 无法学习势函数？

6.1.1 数学层面

弱形式公式 $J_{\text{diss}} - \frac{\sigma^2}{2} J_{\text{lap}} + dE = 0$ 是一个**能量平衡方程**。

表 8: Phase 1 所有实验结果汇总

MVP	名称	V Error	Φ Error	Loss	状态
0.0	SDE 数据生成	-	-	-	PASS
1.0	Joint V+Φ	162%	94%	~ 0	FAIL
1.0b	Φ-only (原公式)	-	78%	0.018	FAIL
1.1	约束版	110%	89%	~ 0	FAIL
1.2	Loss 公式验证	-	-	-	PASS
1.2b	修正公式	98%	128%	~ 0	FAIL

对于**任何**满足这个平衡的 (V, Φ) 对都成立，不仅仅是真实的势函数。

数学表述:

- 设 $(V_{\text{true}}, \Phi_{\text{true}})$ 是真实势函数
- 设 $F = \nabla V_{\text{true}} + \nabla \Phi_{\text{true}} * \mu$ 是真实的力场
- 存在无穷多个 (V', Φ') 使得 $\nabla V' + \nabla \Phi' * \mu = F$
- 这些 (V', Φ') 给出相同的分布演化，弱形式 loss 无法区分

6.1.2 物理直觉

粒子只" 感受" 到总力 $F = -\nabla V - \nabla \Phi * \mu$。

仅从粒子运动无法确定:

- 多少力来自外势 V
- 多少力来自交互势 Φ

类比: 测量物体加速度只能得到合力，无法确定各分力。

6.1.3 为什么约束无效？

表 9: 约束效果分析

约束	作用	为什么无效
$V(0) = 0$	固定 V 的常数项	不影响梯度 ∇V
$\Phi(r_{\text{ref}}) = 0$	固定 Φ 的常数项	不影响梯度 $\nabla \Phi$
梯度正则化	限制梯度大小	不能保证正确方向

需要的约束: 直接约束势函数的**形状**，而非常数。

表 10: Gate-1 评估结果

标准	要求	实际	状态
V 相对误差	$< 10\%$	98–162%	FAIL
Φ 相对误差	$< 10\%$	19–128%	FAIL
Loss 收敛	是	是	PASS (但不保证正确)

7 结论与建议

7.1 Gate-1 评估

Gate-1 最终结论

FAIL —弱形式 loss 存在根本性 identifiability 问题，当前形式下无法工作。

7.2 对 Route A 的影响

当前结论：Route A（NN + Trajectory-free loss）在纯弱形式下**不可行**。

7.3 可能的修复方向

1. **RKHS 正则化**：论文提到的”automatic reproducing kernel” 可能提供 identifiability
2. **多系统联合学习**：不同 V 的系统共享 Φ，增加约束
3. **部分已知信息**：假设 V 已知，仅学习 Φ
4. **强约束**：势函数参数化（如已知函数形式，仅学参数）

7.4 建议的下一步

表 11: 下一步任务优先级

优先级	任务	说明	预期效果
P0	MVP-1.3: Φ-only 简化	假设 V 完全已知	消除一半 trade-off
P1	RKHS 正则化	实现 automatic kernel	可能提供 identifiability
P1	多系统联合	共享 Φ 学习	增加约束
P2	Route B: Kernel	替代方案	理论保障更好

8 附录

8.1 关键公式推导

正确的弱形式 loss（从 Itô 引理推导）：

$$R = \underbrace{\langle |\nabla V + \nabla\Phi * \mu|^2, \mu\rangle \Delta t}_{J_{\text{diss}}} - \underbrace{\frac{\sigma^2}{2}\langle \Delta V + \Delta\Phi * \mu, \mu\rangle \Delta t}_{J_{\text{lap}}} + \underbrace{[E(t+\Delta t) - E(t)]}_{dE} = 0 \tag{12}$$

能量定义:

$$E(t) = \langle V, \mu_t\rangle + \langle \Phi * \mu_t, \mu_t\rangle = \frac{1}{N}\sum_i V(X_t^i) + \frac{1}{N^2}\sum_{i,j}\Phi(X_t^i - X_t^j) \tag{13}$$

8.2 代码文件索引

文件	说明
`core/trajectory_free_loss.py`	Loss 函数（已修正）
`core/nn_models.py`	神经网络模型
`core/sde_simulator.py`	SDE 模拟器
`scripts/train_nn.py`	训练脚本
`scripts/train_nn_constrained.py`	约束版训练脚本
`scripts/verify_loss_formula.py`	Loss 验证脚本

报告生成时间：*2026-01-28*

总实验数：*6*

总耗时：约 *4* 小时