arXiv:2601.10326v1 [math.ST] 15 Jan 2026

非线性偏微分方程模型中的梯度稳定性与相互作用粒子系统中的推断

Aur´elien Castre 和 Richard Nickl

剑桥大学

2026年1月16日

摘要

我们研究非线性偏微分方程的一般参数到解的映射 $\theta \mapsto \mathcal{G}(\theta)$，并介绍一种基于巴拿赫空间版本隐函数定理的方法，用于验证相关非线性逆问题的梯度稳定性条件 [35]，同时给出单射估计及相应的统计可识别性结果。我们通过两个例子说明这些方法：一个涉及非线性反应扩散系统，另一个涉及McKean–Vlasov相互作用粒子模型，二者均满足周期边界条件。我们应用所得结果，证明Langevin型算法可在多项式时间内收敛；该算法从相互作用粒子系统的离散聚合测量所得到的相互作用势后验测度中进行采样。

目录

1 引言 **2**

1.1 一般框架.2 1.2 主要结果.4 1.3 记号.6

2 计算非线性偏微分方程模型的梯度 7

2.1 反应扩散方程.7 2.2 McKean–Vlasov方程.9 2.2.1 映射 f 的性质.9 2.2.2 全局稳定性估计与伪线性化恒等式.11 2.2.3 线性化与高阶导数.14

3 基于朗之万型算法的多项式时间后验抽样 **16**

3.1 设定.17 3.1.1 逼近空间.17 3.1.2 似然函数的导数.18 3.2 似然函数的局部性质.19 3.2.1 局部平均曲率.19 3.2.2 局部正则性.20 3.3 主要结果.21 3.3.1 收缩速率.21 3.3.2 后验分布的对数凹近似.23 3.3.3 热启动ULA的多项式混合时间界.25

线性化麦基恩–弗拉索夫方程的偏微分方程估计 **26**

B 巴拿赫空间中的微积分 30B.1 其余证明.32

1 引言

当代计算数学中使用的许多算法都会用到定义在欧几里得空间 $\mathbb{R}^D$ 上的映射 $H:\ \mathbb{R}^D \to \mathbb{R}$ 的梯度 ∇H。这包括通过（随机）梯度下降法在其可能具有高维（$D \gg 1$）的定义域 H 上进行最小化，或从吉布斯型概率测度中生成随机样本。

$$\mathrm{d}\mu(\theta) \propto e^{-H(\theta)}\,\mathrm{d}\theta, \quad \theta \in \mathbb{R}^D,$$

通过朗之万型方法。在采样问题中，实际应用通常依赖基于梯度的马尔可夫链蒙特卡罗（MCMC）方法，例如梅特罗波利斯调整的朗之万算法（MALA） [6,40, 39]。这些方法已在广泛的应用中广受欢迎——不胜枚举，但这里提及非线性偏微分方程型逆问题和数据同化 [42, 13, 38, 23]，作为本文将研究的一个实例。近年来，人们利用 [4] 中发展的马尔可夫过程几何性质，对此类算法的理论收敛保证进行了广泛研究。值得注意的是，与凸优化 [8], 的情形类似，当 H 是凸的 [15, 19, 20, 12] 时，可以预期马尔可夫链在高维空间中快速混合并趋向其不变测度 μ。此外，已知凸性——或 μ 满足的相关泛函不等式，参见例如 [11] 及其中的参考文献——对于 MALA 等算法在高维空间中的多项式混合时间至关重要；而即使目标函数是单峰的 H，只要它不是全局凸的，运行时间也可能随相关参数呈指数级增长，参见 [5]。

在统计推断的背景下，$-H$ 通常是一个拟合泛函，例如（可能带惩罚项的）对数似然函数；而 μ 则是由 $\mathbb{R}^D$上的某个先验产生的后验测度，参见例如 [42, 31]。梯度 ∇H 由此决定费舍尔信息，进而决定真实值参数附近的局部平均曲率。在近期的研究 [35] 中，研究者采用了一种相关的梯度稳定性 概念，从而为朗之万型马尔可夫链蒙特卡洛算法建立了高概率多项式运行时间界，即使 H 不是凸的，但 μ 是后验测度时亦然。梯度稳定性已在多种涉及线性 偏微分方程的非线性情形中得到验证 [35, 31, 7, 2], ；但许多复杂模型涉及非线性偏微分方程，例如麦基恩–弗拉斯科方程或纳维–斯托克斯方程。对于这些情形，如何检验梯度稳定性，甚至如何推导梯度表达式（这对实际实现很重要），通常并不明确。

本文是对这一类问题的贡献：我们将介绍一种基本论证，其本质基于巴拿赫空间版本的隐函数定理，用于计算非线性偏微分方程的参数到解映射的梯度。我们将通过下文介绍的两个非线性偏微分方程原型示例，说明这种方法的工作原理，并利用所得结果表明，近期研究的相互作用粒子模型的贝叶斯推断算法在高维设置下可以通过热启动朗之万型方法实现计算。在此过程中，我们还提供了新的证明技术，用于推导非线性反问题中的全局单射性估计；特别地，这为建立 [33] 中麦肯–弗拉索夫系统的一些关键分析结果提供了一种替代方法。我们的思想还可以与用于建立反问题中正向映射的实解析性和全纯性的 [14] 中所发展的方法联系起来（例如，通过 [18] 中的定理 15.3）。

1.1 一般框架

我们将分析重点放在如下形式的能量 H 上

$$H(\theta) \coloneqq \frac{1}{2}\left(\sum_{i=1}^{N} |Y_i - \mathcal{G}(\theta)(X_i)|^2 + \theta^T \Sigma^{-1} \theta\right) \tag{1}$$

其中， $\Sigma = \Sigma(N, D)$ 是一个 $D \times D$ 对称半正定矩阵， $\mathcal{G}$ 是从 $\mathbb{R}^D$ 到空间 $L^2(\mathcal{X})$ 的某个映射；该空间由定义在欧几里得空间的有界子集 $\mathcal{X}$ 上的平方勒贝格可积函数构成，而 $(Y_i, X_i)_{i \le N}$ 是取值于 $\mathbb{R} \times \mathcal{X}$ 的随机变量。我们沿用反问题文献 [42]中的术语，将 $\mathcal{G}$ 称为“正向映射”。这类吉布斯测度在贝叶斯推断中自然地作为后验分布出现，其数据来自对 $\mathcal{G}(\theta)$ 的含噪点观测，先验则是协方差矩阵为 Σ 的中心高斯先验。更具体地说，若我们观测到数据 $Z_N = (Y_i, X_i)_{i \le N}$，且这些数据根据以下模型生成

$$Y_i = \mathcal{G}(\theta)(X_i) + \varepsilon_i, , \quad \varepsilon_i \overset{\text{iid}}{\sim} \mathcal{N}(0,1), \ X_i \overset{\text{iid}}{\sim} \mathrm{Unif}(\mathcal{X}),$$

and 令 Π 为先验分布，我们得到条件“后验”分布

$$\mathrm{d}\Pi(\theta | Z_N) \propto e^{\ell_N(\theta)} \, \mathrm{d}\Pi(\theta) = e^{-H(\theta)} \, \mathrm{d}\theta \tag{2}$$

其中 $\ell_N(\theta) = -\frac{1}{2} \sum_{i=1}^N |Y_i - \mathcal{G}(\theta)(X_i)|^2$ 是该模型的对数似然。我们用 P_θ^N 表示描述参数为 θ 时数据 Z_N 分布的乘积测度。关于这一标准框架的综述，可参见 [31]中的第 1 章。

上述“梯度稳定性”概念涉及对 $\mathcal{G}$ 在 θ 处的梯度 $\nabla \mathcal{G}_\theta$ 的识别，以及以下单射估计

$$\|\nabla \mathcal{G}_\theta(h)\|_{L^2(\mathcal{X})} \gtrsim \|h\|_{\mathbb{R}^D} \quad \forall h \in \mathbb{R}^D, \tag{3}$$

参见 [35, 第 3.1] 节或 [31, 第 3.2] 节。这一要求与底层回归模型在 θ 处逆费舍尔信息的存在性密切相关。对于涉及非线性参数识别问题的若干线性偏微分方程案例（例如薛定谔方程 [30, 35], 达西问题 [31, 2], 非阿贝尔 X-射线变换 [28, 29, 7], 扩散问题 [23]），以及反应扩散和纳维–斯托克斯方程等非线性偏微分方程的初值到解的映射（参见 [32, 25]），人们已计算出前向映射 $\mathcal{G}$ 的梯度。然而，这些证明无法直接推广到更困难的情形：例如，希望识别非线性偏微分方程的一般参数，如 McKean–Vlasov 方程 (6) 中的相互作用势 W。困难在于必须同时处理两个非线性：一个源自微分算子，另一个源自参数到解的映射。

为说明我们的方法，我们将在一般函数空间中进行分析，并将 θ 视为巴拿赫空间 Θ 中的一个元素。于是，通过适当的离散化过程（例如，使用小波基或傅里叶基／框架），可以将欧几里得空间 $\mathbb{R}^D$ 视为 Θ 的一个线性子空间。因此，给定两个巴拿赫空间 X 和 Y，我们将 $\mathcal{G}$ 视为从 Θ 到 X 的映射，并假设存在一个明确定义的二元函数 f，满足 $f\colon \Theta \times X \to Y$，从而可以通过以下方式隐式定义 $\mathcal{G}$

$$f(\theta, \mathcal{G}(\theta)) = 0, \quad \theta \in \Theta. \tag{4}$$

这类 f 可以适用于各种各样的问题——这里我们将通过两个非线性抛物型偏微分方程示例来说明这一框架。我们先从一个反应扩散模型开始，其中参数 θ 表示一个利普希茨反应项 $R\colon \mathbb{R} \to \mathbb{R}$，而非线性映射 $\mathcal{G}(\theta) =: u_R$ 定义为非线性抛物型偏微分方程的唯一解 $u_R = u$

$$\begin{cases} \partial_t u - \Delta u = R(u) & \text{on } (0,T) \times \mathbb{T}^d \\ u(0) = \phi & \text{on } \mathbb{T}^d \end{cases} \tag{5}$$

给定环面 $\mathbb{T}^d := \mathbb{R}^d / \mathbb{Z}^d$ 上的某个初始条件 ϕ，其中 Δ 为拉普拉斯算子。若令

$$f(R, u) = (\partial_t u - \Delta u - R(u), u(0) - \phi)$$

那么，对于适当选择的 Θ， X， Y，数对 (R, u_R) 恰好是 (4) 中 f 的零点。

第二个例子受到近期相互作用粒子模型统计推断方法研究热潮的启发，例如 [16, 3, 37, 33]。这里考虑一种称为 McKean–Vlasov方程的非线性偏微分方程，其中参数 θ 表示由 W 表示的“相互作用势”，而 $\mathcal{G}(\theta) =: \rho_W$ 定义为非线性、非局部福克-普朗克方程的唯一解 $\rho_W = \rho$ （其中 $\nabla\cdot$ 是散度算子）

$$\begin{cases} \partial_t \rho = \Delta\rho + \nabla \cdot (\rho \nabla W * \rho) & \text{on } [0,T] \times \mathbb{T}^d \\ \rho(0) = \phi & \text{on } \mathbb{T}^d \end{cases} \tag{6}$$

对于某个定义在 $\mathbb{T}^d$上的概率密度 ϕ。与上文类似，我们可以取

$$f(W, \rho) = (\partial_t \rho - \Delta\rho - \nabla \cdot (\rho \nabla W * \rho), \rho(0) - \phi)$$

在 (4) 中。我们在下文提出的技术远不止适用于这里讨论的两个例子，例如，还可用于纳维–斯托克斯方程 [34, 25]中出现的统计推断问题。

偏微分方程理论中常用的一种自然思路（例如，参见 [1]第 III 章）是对方程 (4) 关于 θ作形式求导。这得到一个恒等式

$$D\mathcal{G}(\theta) = -(D_2 f(\theta, \mathcal{G}(\theta)))^{-1} \circ D_1 f(\theta, \mathcal{G}(\theta)) \tag{7}$$

其中只涉及映射 f 的偏导数——这些偏导数通常可以轻松求得。若 f 连续可微，且 $D_2 f$ 是同胚，则可利用巴拿赫空间中的隐函数定理使上述公式严谨成立（例如，参见 [17, 定理 10.2.1]以及下文的定理 B.1）。由此，根据 (7) 以及 f的高阶可微性，可直接推出 $\mathcal{G}$ 的高阶可微性。下面我们将这一主要思路概括为如下断言。

命题。 设 f 满足(4) 。此外，假设对于某个 $k \geq 1$， f 在 ***Fr´echet*** 意义下是 C^k，并且对所有 $\theta \in \Theta$, $D_2 f(\theta, \mathcal{G}(\theta)) \colon X \to Y$都是同胚。最后，假设对任意 $\theta \in \Theta$，$u = \mathcal{G}(\theta)$ 都是 X 中方程 $f(\theta, u) = 0$的唯一解。那么， $\mathcal{G}$ 在 ***Fr´echet*** 意义下是 C^k，其导数由(7)给出。

由此可以从直觉上得出结论：验证梯度稳定性 (3) 将涉及对 (7) 中映射的单射估计，而这又归结为 f 的偏导数 $D_1 f$ 的单射性。对于适当定义域上的 $D_2 f$ 的单射性，已经是上述 $\mathcal{G}$ 导数公式中隐含的必要条件。我们还将在第2.2.2节说明，这些想法如何借助非线性反问题中常用的“伪线性化”恒等式，为非线性映射 $\mathcal{G}$ 提供单射性结果和稳定性估计 [28, 36]。

1.2 主要结果

对于反应扩散模型 (5)，我们将证明正向映射 $R \mapsto u_R$ 在 Fr´échet 意义下连续可微，并且其导数 Du_R 在点 R 处对应的线性映射可通过求解一个线性抛物型偏微分方程来计算。以下结果对此进行了总结。定理 2.1 给出了严格的表述及其证明。

断言 （反应-扩散正向映射的导数）。 若 $d \leq 3$ 且 $\phi \in H^1$，则对于任意 R、$H \in C_b^2$（$\mathbb{R}$）， u_R 关于 R、 $Du_R[H] := i_H = i$ 的 ***Fr´échet*** 导数是线性抛物型偏微分方程的唯一解

$$\begin{cases} \partial_t i - \Delta i - R'(u_R) i = H(u_R) & on\ (0,T) \times \mathbb{T}^d \\ i(0) = 0 & on\ \mathbb{T}^d. \end{cases}$$

现在我们讨论如何计算 McKean–Vlasov 模型（定义见式 (6)）中 $\mathcal{G}$ 的导数。我们的普适结果将给出函数空间中的 Fr´échet 导数（见定理 2.11），但对于使用基于梯度的算法从高维 Gibbs 测度（1）中进行采样的应用，我们先聚焦于这样的情形： W 被限制在由 Θ 构成的 D 维近似空间中。考虑 L^2 的三角函数基（见第 3.1.1 节），对于 $D \geq 2$ 为偶数的情形，将 $\mathbb{R}^D$ 中的每个向量与其对应的均值为零函数对应起来，该函数由其前 D 个傅里叶系数定义。通过这种方式，对于 $W \in \mathbb{R}^D$， ρ_W 是良定义的。对于任意固定的（t, x） $\in [0,T] \times \mathbb{T}^d$，通过 $\mathcal{G}^{t,\,x}(W) := \rho_W(t,\ x)W \in \mathbb{R}^D$ 定义映射 $\mathcal{G}^{t,\,x}$。然后， $\mathcal{G}^{t,\,x}$ 在点 $W \in \mathbb{R}^D$ 处的梯度可以通过求解一个线性抛物型偏微分方程来计算，如以下断言所述。完整细节见定理 2.11 和引理 3.2。

命题 （McKean–Vlasov 正向映射的梯度）。 假设 （6）中的概率密度 ϕ 满足：对某个 $\beta \geq 3+d$，有 $\phi \in H^\beta$。那么，对于任意 $H \in \mathbb{R}^D$ 以及（t、 x） $\in [0,T] \times \mathbb{T}^d$， $\mathcal{G}^{t,\,x}$ 关于 W 的梯度为 $H^T \nabla \mathcal{G}^{t,\,x}$（$W$） $= D\rho_W[H]$（t、 x），其中 $D\rho_W[H] = u_H = u$ 是下列线性抛物型偏微分方程的唯一解

$$\begin{cases} \partial_t u - \Delta u - \nabla \cdot (u \nabla W * \rho_W) - \nabla \cdot (\rho_W \nabla W * u) = \nabla \cdot (\rho_W \nabla H * \rho_W) & on\ [0,T] \times \mathbb{T}^d \\ u(0) = 0 & on\ \mathbb{T}^d. \end{cases}$$

现在我们转 向此前在（2）中定义的后验分布的随机样本生成问题 n

$$\mathrm{d}\Pi(W|Z_N) \propto e^{\ell_N(W)}\,\mathrm{d}\Pi(W)$$

，其中 $\mathcal{G}(\theta) = \rho_W$ 来自具有 $\mathcal{X} = [0,T] \times \mathbb{T}^d$ 的 McKean–Vlasov 模型。利用前述结果和 [35], 中的思想，我们可以验证梯度稳定性（3），并推导出 $-\ell_N$ 在“真实参数” W_0附近局部强凸，且该结论在离散化方案下平均成立。只要 ϕ 在（6）中的傅里叶模态衰减不太快，这一结论便成立；这也是相互作用粒子模型中的一个自然可辨识性假设，正如 [33]中所讨论的。更准确地说，若 $Z_N \sim P_{W_0}^N$，则平均海森矩阵的最小特征值有如下下界

$$\lambda_{\min}\left(\mathbb{E}_{W_0}[-\nabla^2 \ell(W)]\right) \geq c_D > 0$$

对势函数 W 在以 W_0 为中心的球内一致成立，该球的大小取决于维度 D，以及模型的可识别程度，而可识别程度由 ϕ 的傅里叶模态衰减情况决定。常数 c_D 的衰减速度至多为 D 的多项式阶。这正是定理 3.3 的内容。接着，沿用 [35, 7, 31], 的思路，我们在第 3.10 节引入一个全局对数凹的代理分布 Π_N，以 Wasserstein-2 距离近似 Π（$\cdot \mid Z_N$）。我们使用如下定义的非校正朗之万算法（ULA）从该代理分布中采样：

$$\vartheta_{k+1} = \vartheta_k + \gamma \nabla \log(\mathrm{d}\tilde{\Pi}_N(W)) + \sqrt{2\gamma}\xi_{k+1} \tag{8}$$

其中 ξ_k 独立同分布$\sim$ $\mathcal{N}(0, I_D)$，且 γ 是一个“步长”参数。每个梯度步都需要对非线性方程 (6) 进行（数值）求解，以及求解线性化抛物型偏微分方程 $D\rho_W[H]$；已有多种算法可用于求解（参考文献见 [33] ）。我们得到了链 (ϑ_k) 的 Wasserstein-2 混合时间上界；在“热启动”条件下，该上界随 N, D 呈多项式增长。完整细节见定理 3.11。

主张 (ULA 的多项式混合)。 假设式 (8) 中的链从 ϑ_0 出发，并且足够接近$W_0 \in H^\alpha$ ，其中某个 $\alpha > 0$ 足够大；同时， ϕ 在式 (6) 中的傅里叶模态不会衰减得过快。那么，在进一步的自然正则性假设下，对于任意期望的准确度 $\varepsilon > 0$，存在一个步长 $\gamma = \gamma(\varepsilon, D, N)$ 以及一个整数 k_{mix}，其关于 ε^{-1}、 D 和 N 的增长至多为多项式级别，使得对于所有 $k \geq k_{\mathrm{mix}}$，

$$\mathcal{W}_2^2(\mathcal{L}(\vartheta_k), \Pi(\cdot \mid Z_N)) \leq e^{-N^{d/(2(\alpha+1)+d)}} + \varepsilon$$

以 $P_{W_0}^N$-概率指数快速地趋于 *1*。

在上述断言中， $\mathcal{W}_2$ 是由下式定义的 Wasserstein-2 距离

$$\mathcal{W}_2^2(\mu,\nu) \coloneqq \inf_{\pi} \int_{\mathbb{R}^D\times\mathbb{R}^D} \|x-y\|_D^2 \,\mathrm{d}\pi(x,y) \tag{9}$$

其中 $\|\cdot\|_D$ is $\mathbb{R}^D$ 上的欧几里得范数，且下确界取遍 μ 与 ν的所有耦合 π 。

1.3 记号

集合 $\mathbb{N}$定义为 $\mathbb{N}\coloneqq\{1,2,\dots\}$，而 $\mathbb{Z}$表示全体整数。全文中，我们取d维平坦环面 $\mathbb{T}^d\coloneqq\mathbb{R}^d/\mathbb{Z}^d, d\in\mathbb{N}$作为空间定义域，并赋予其诱导的欧几里得度量。当存在常数 $c>0$使得 $A\le cB$时，我们记作 $A\lesssim B$；这里默认该常数 c仅依赖于固定参数。于是， $A\gtrsim B$ 表示 $B\lesssim A$ ，而$A\simeq B$ 表示 $A\lesssim B$ 和 $A\gtrsim B$同时成立。我们用 $A\lesssim_p B$ 表示隐含常数c 依赖于 p。我们遵循分析学中的标准约定：符号 C 表示一个正的常数，其值可能随行而变化。

对于 $1\le p\le+\infty$，令 $L^p(\mathbb{T}^d)$，或简写为 L^p ，表示通常的实值函数周期勒贝格空间。令 $\mathcal{D}(\mathbb{T}^d,\mathbb{C})$表示定义在 $\mathbb{R}^d$上的复值、 C^∞ 且 1-周期函数空间。再令$\mathcal{D}'(\mathbb{T}^d,\mathbb{C})$表示周期分布空间，即复向量空间$\mathcal{D}(\mathbb{T}^d,\mathbb{C})$的拓扑对偶。对于任意 $s\in\mathbb{R}$，我们定义 Sobolev空间 $H^s(\mathbb{T}^d)$，或简写为 H^s ，如下：

$$H^s=\left\{u\in\mathcal{D}'(\mathbb{T}^d,\mathbb{C}):\overline{\hat{u}_k}=\hat{u}_{-k}\text{ and }\sum_{k\in\mathbb{Z}^d}(1+|k|^2)^s\,|\hat{u}_k|^2<+\infty\right\}$$

其中，$\hat{u}_k\coloneqq\left\langle u,e^{2i\pi k\cdot x}\right\rangle$ 是分布 u 的 k^{th} 傅里叶系数，括号表示通常的对偶配对。赋予内积 $\langle u,v\rangle_{H^s}=\sum_{k\in\mathbb{Z}^d}(1+|k|^2)^s\hat{u}_k\hat{v}_k$, H^s后，它构成一个实的、可分的希尔伯特空间。我们进一步将齐次 Sobolev 空间 H^s 定义为 H^s 的闭子空间，其元素 $u\in H^s$ 满足$\hat{u}_0=0$。在 H^s上，我们有等价的内积 $\langle u,v\rangle_{\dot{H}^s}=\sum_{k\in\mathbb{Z}^d}|k|^{2s}\hat{u}_k\hat{v}_k$。更一般地，对于定义在 $\mathbb{T}^d$上的函数巴拿赫空间 B， $\dot{B}$表示由零均值函数构成的 B 的闭子空间。特别地，我们有 $\dot{L}^2=\dot{H}^0$，并令$\langle u,v\rangle\coloneqq\int u\bar{v}$，其中 $u,v\in L^2$。对于 $k\in\mathbb{N}$，空间 $W^{k,\infty}$是由 L^2 中具有直至 k 阶弱偏导数的函数构成的 Sobolev 空间。

对于 $T>0$，令 $\mathcal{X}=[0,T]\times\mathbb{T}^d$为配备均匀概率测度 $\lambda=\lambda_T$的时空柱 $\mathcal{X}$。我们用 $L^2(\mathcal{X},\lambda)$表示 $\mathcal{X}$上的平方可积函数空间。更一般地，令 B为巴拿赫空间，且 $1\le p\le\infty$。我们将 $L^p([0,T];B)$记作由强可测映射 $t\mapsto H(t)\in B$ 构成的 Bochner 空间，其中这些映射满足 $t\mapsto\|H(t)\|_B\in L^p([0,T])$。当 B是 $\mathbb{T}^d$ 上的空间函数空间时，我们通常将这样的 H 视为函数 $H:\mathcal{X}\to\mathbb{R}$ ，其满足 $H(t,\cdot)\in B$ ，对几乎处处的 t成立。对于整数 $m\ge0$，时间变量上的 Sobolev空间 $H^m([0,T];B)$ 定义为

$$H^m([0,T];B)=\left\{H\in L^2([0,T];B)\ :\ \partial_t^jH\in L^2([0,T];B)\text{ for }0\le j\le m\right\},$$

配备范数$\|H\|^2_{H^m([0,T];\,B)}=\sum_{j=0}^m\left\|\partial_t^jH\right\|^2_{L^2([0,T];\,B)}$。我们常用简写 H^m_TB表示 $H^m([0,T];B)$。对于任意两个巴拿赫空间 B_1， B_2，我们在交空间 $B=B_1\cap B_2$上赋予和范数 $\|\cdot\|_B=\|\cdot\|_{B_1}+\|\cdot\|_{B_2}$。回顾一下，只要 B_1和 B_2 连续嵌入到同一个 Hausdorff 拓扑向量空间（例如 $L^2([0,T],L^2)$）中，$(B,\|\cdot\|_B)$ 就构成一个巴拿赫空间。对于X 和 Y这两个赋范空间，当 X 连续嵌入于 Y时，我们记作 $X\hookrightarrow Y$。

我们还将反复用到抛物型偏微分方程理论中的以下观测：通过选取合适的光滑化序列，完全按照 [21, 第5章.9 节定理 4], 的证明，可以证明对于任意 $\gamma\ge0$ 和 $u\in L^2([0,T];H^{\gamma+1})\cap H^1([0,T];H^{\gamma-1})$，我们有 $u\in C([0,T];H^\gamma)$，并且

$$\sup_{0\le t\le T}\|u(t)\|_{H^\gamma}\lesssim_{\gamma,d,T}\|u\|_{L^2([0,T];H^{\gamma+1})}+\|u\|_{H^1([0,T];H^{\gamma-1})}. \tag{10}$$

2 非线性偏微分方程模型中的梯度计算

本节将说明，附录中回顾的非线性泛函分析和偏微分方程理论中的一般方法，如何应用于引言中介绍的两个抛物型方程 (5)、(6)。我们将展示如何推导 $\mathcal{G}$ 的表达式（如 (3) 中所示）并证明其梯度稳定性；此外，我们还将在第2.2.2节探讨一个相关问题：如何借助“伪线性化”思路证明 $\mathcal{G}$ 本身的全局稳定性估计。

2.1 反应扩散方程

令 $C_b^k(\mathbb{R})$ 为定义在 $\mathbb{R}$ 上的有界 k 次连续可微函数所构成的巴拿赫空间，这些函数的最高至 k 阶导数均有界，并配备范数 $\|f\|_{C^k} := \sup_{x\in\mathbb{R}} \sum_{i=0}^{k} |f^{(i)}(x)|$。固定 $\phi \in H^1$。对于 $R \in C_b^2(\mathbb{R})$，令 $u = u_R \in L^2([0,T];\, H^2) \cap H^1([0,T];\, L^2)$ 表示以下方程的唯一强解

$$\begin{cases} \partial_t u - \Delta u = R(u) & \text{in } (0,T)\times\mathbb{T}^d \\ u(0) = \phi & \text{in } \mathbb{T}^d \end{cases} \tag{11}$$

解 u_R 的存在性及其在 $C([0,T];\, L^2)$ 中的唯一性可由标准的不动点论证得出，参见例如 [21, 第536页及后续], ，其中该参考文献中的算子 A 将 u 指派给强 解（而非弱解），该解对应于初始条件为 ϕ 的线性问题 $\partial_t v - \Delta v = R(u)$；另见 [32, 定理 6]。细节从略。在应用中，u_R 描述某种化学物质（初始分布为ϕ）的演化，其中 $R(u)$ 表示该物质在时空各点的生成或消耗量。我们假设 $d \le 3$，以聚焦于主要思想。

定理 2.1 (反应扩散方程的线性化)。设 $d \le 3$。映射 $\mathcal{G}: C_b^2(\mathbb{R}) \to L^2([0,T];\, H^2) \cap H^1([0,T];\, L^2)$ 由 $\mathcal{G}(R) = u_R$ 定义，它是通过求解 (11) 得到的，在 *Fr´chet* 意义下是 C^1。此外，对于任意 $R \in C_b^2(\mathbb{R})$，其在 R处的 *Fr´chet* 导数，

$$D\mathcal{G}(R) : C_b^2(\mathbb{R}) \to L^2([0,T];H^2) \cap H^1([0,T];L^2)$$

由线性映射 $D\mathcal{G}(R)[H] =: i_H$ 给出，其中 $i = i_H \in L^2([0,T];\, H^2) \cap H^1([0,T];\, L^2)$ 是以下线性抛物型偏微分方程的唯一强解

$$\begin{cases} \partial_t i - \Delta i - R'(u_R) i = H(u_R) & in\ (0,T)\times\mathbb{T}^d \\ i(0) = 0 & in\ \mathbb{T}^d. \end{cases}$$

定理 *2.1* 的证明。 证明直接应用定理 B.1；我们先将记号调整为符合该定理的设定。令

$$E = C_b^2(\mathbb{R}), \quad F = L^2([0,T];H^2) \cap H^1([0,T];L^2), \quad G = L^2([0,T];L^2) \times H^1$$

并定义映射

$$\begin{aligned} f: \ E \times F \ &\longrightarrow \ G \\ (R,u) \ &\longmapsto (\partial_t u - \Delta u - R(u), u(0) - \phi)\,. \end{aligned} \tag{12}$$

函数 f 由 (10) 定义良好，且由于 $\phi \in H^1$ 和 $R \in C_b^2(\mathbb{R})$。为便于说明，我们仔细证明 f在 $E\times F$上是 C^1。考虑辅助映射 $g\colon F \to G$和 $h\colon E\times F \to L^2([0,T];\, L^2)$，定义如下

$$g(u) = (\partial_t u - \Delta u, u(0)), \quad h(R,u) = R(u).$$

那么，$f(R,\ u) = g(u) + (-h(R,\ u),\ 0) + (0,\ -\phi)$。映射 g显然是线性的，而且根据（10）可知它也是连续的。因此，由引理 B.2 可知，g 在 Fr´echet 意义下是光滑的。由于 $(0,\ -\phi)$ 只是一个常数，只需证明 h 是 C^1。对于任意 $R \in E$，注意到

$$\sup_{u\in F} \|h(R,u)\|_{L_T^2 L^2} \le \sqrt{T}\|R\|_{C^0}.$$

由于 h 关于 R 是线性的，我们得出 h 关于第一分量连续，且对 $u \in F$ 一致。由此我们也得出 h 关于第一分量是 C^∞，并且 $D_1h(R,\ u)[H] = H(u)$。接下来，对于 F 中任意的 $u \in F$ 和 $u_n \to u$，由微积分基本定理可得

$$\|R(u_n) - R(u)\|_{L^2_T L^2} \le \|R\|_{C^1}\|u_n - u\|_{L^2_T L^2} \to 0.$$

因此，h 在 $E \times F$上是联合连续的。现在我们计算 h 关于 u的导数。根据微积分基本定理连续应用两次，对于任意 $s \in F$，可得

$$\left|R(u+s) - R(u) - R'(u)s\right| \le \frac{1}{2}\|R\|_{C^2}|s|^2 \quad \text{a.e. in } [0,T]\times\mathbb{T}^d.$$

对时间和空间积分后，可得

$$\|R(u+s) - R(u) - R'(u)s\|_{L^2_T L^2} \le \frac{1}{2}\|R\|_{C^2}\|s^2\|_{L^2_T L^2} \le \frac{1}{2}\|R\|_{C^2}\|s\|_{L^\infty_T L^2}\|s\|_{L^2_T L^\infty}.$$

此外，由（10）可得 $\|s\|_{L^\infty_T L^2} \lesssim \|s\|_F$，由索伯列夫嵌入可得 $\|s\|_{L^2_T L^\infty} \lesssim \|s\|_F$（利用 $d \le 3$）。由于 $s \mapsto R'(u)s$ 是有界线性算子，我们 conclude that h 在第二个分量上可微，其导数为 $D_2h(R,\ u)[s] = R'(u)s$。为了证明 h 是 C^1，还需证明 D_2h在 $E \times F$上联合连续。设$(R,\ u) \in E \times F$，其中 $R_n \to R$属于 E，$u_n \to u$属于 F。那么，对于任意满足 $\|s\|_F \le 1$的$s \in F$，可得

$$\begin{aligned}\|R_n'(u_n)s - R'(u)s\|_{L^2_T L^2} &\le \|s\|_{L^\infty_T L^2}\|R_n'(u_n) - R'(u)\|_{L^2_T L^\infty}\\ &\lesssim \|s\|_F\left(\|R_n'(u_n) - R'(u_n)\|_{L^2_T L^\infty} + \|R'(u_n) - R'(u)\|_{L^2_T L^\infty}\right)\\ &\le \sqrt{T}\|R_n - R\|_E + \|R\|_{C^2}\|u_n - u\|_F \xrightarrow[n\to\infty]{} 0\end{aligned}$$

再次利用微积分基本定理，可得 R' 是利普希茨连续的；当 $d \le 3$ 时，利用索伯列夫嵌入$H^2 \hookrightarrow L^\infty$。由此证明了 h 是 C^1，进而也证明了 f 是 C^1。我们立即指出，根据多重线性映射的求导法则，对于任意 $R,\ H \in E$ 和 $u,\ s \in F$，

$$D_1 f(R,u)[H] = -(H(u),0), \quad D_2 f(R,u)[s] = (\partial_t s - \Delta s - R'(u)s, s(0)).$$

现在我们验证定理 B.1 的最后几个假设。对于任意 $R \in E$，$u_R \in F$（$R \in C_b^2$（$\mathbb{R}$）是全局利普希茨连续的），根据 f 的定义，$f(R,\ u_R) = 0$，并且 u_R 在 $C([0,T];\ L^2)$ 中是唯一的，因此根据（10），在 $F \hookrightarrow C([0,T];\ L^2)$ 中也是唯一的。最后，根据上方公式中 $D_2 f$ 的形式，$D_2 f(R,\ u_R)$ 定义了一个标准的线性抛物型初值问题，其理论已广为人知。特别地，根据线性抛物型方程的标准存在唯一性与正则性理论（例如参见 [21, 第 7] 章），$D_2 f(R,\ u_R)$：$F \to G$ 是一个双射，且其逆算子有界。根据定义，它是线性的，因此由开映射定理可知它是连续的。因此，我们可以应用定理 B.1，得出 $\mathcal{G}$ 在 Fr´echet 意义下是 C^1 的，并且其导数由定理陈述中的公式给出。 □

注 2.2 (梯度稳定性)。 正如引言中所述，梯度的一个关键性质对于模型的统计可识别性至关重要，这就是“梯度稳定性”(3)。下文将在 McKean–Vlasov 模型中进行完整分析。对于反应扩散模型，我们仅概述一些思路，留待未来工作中详细探究。定理 2.1 中的导数可写为

$$D\mathcal{G}(R)[H] = (D_2 f(R,u_R))^{-1}[H(u_R)].$$

由于 $D_2 f(R,u_R)$ 是一个线性同胚，因此利用插值估计，可以在适当的范数下对其逆算子给出下界。因此，单射性的主要挑战来自项 $H(u_R)$。为了证明当 $\|H(u_R)\|^2_{L^2_T L^2} = \int_0^T \int_{\mathbb{T}^d} |H(u_R(t,x))|^2 \,\mathrm{d}x\,\mathrm{d}t$ 为零时，H 必须消失，需要控制范围$\{u_R(t,x) : (t,x) \in \mathcal{X}\}$。例如，可以通过短时间论证实现这一点：由于 $t \mapsto u_R(t)$ 是连续的，在充分小的时间内，u_R 的范围仍然接近于初始条件 ϕ 的范围。通过规定 ϕ 的范围，随后或许便能在该集合上恢复 H。

2.2 McKean–Vlasov方程

在本节中，对于 $W \in \dot{W}^{2,\infty}$ 一个 均值为零的 相互作用势， ρ_W 表示由下式定义的 McKean–Vlasov 非线性非局部抛物型偏微分方程的唯一强解：

$$\begin{aligned} \partial_t \rho &= \Delta \rho + \nabla \cdot (\rho \nabla W * \rho) \text{ on } [0,T] \times \mathbb{T}^d \\ \rho(0) &= \phi, \end{aligned} \tag{13}$$

其中， ϕ 是定义在 $\mathbb{T}^d$ 上的某个概率密度，用于描述系统的初始状态。关于该方程如何作为扩散相互作用粒子系统的平均场极限导出，读者可参阅 [33] 的引言及其中的参考文献。为确保 ρ_W具有足够的正则性，我们将对 ϕ采用以下假设。

Assu假设 2.3。 初始条件 ϕ 属于 H^β ，其中 $\beta \geq 3$ +是某个偶整数 d.

在假设 2.3 下，对于 $W \in \dot{W}^2$， ∞， ρ_W 的存在性、唯一性和正则性均得到保证：特别地，方程（13）的唯一解 ρ_W属于 $L^2([0,T];\ H^{\beta+1}) \cap H^1([0,T];\ H^{\beta-1})$。关于这些结论的证明，请参阅[10], [9], [24, 定理 3.4], 和 [33, 定理 5] 。要求 β 为偶整数仅仅是为了方便，因为这样可以简化一些抛物型正则性的证明。

2.2.1 映射 f的性质

我们先固定一些记号。考虑空间

$$E := \dot{W}^{2,\infty},\ F = F_\beta := L^2([0,T]; H^{\beta+1}) \cap H^1([0,T]; H^{\beta-1}),\ G = G_\beta := L^2([0,T]; H^{\beta-1}) \times H^\beta. \tag{14}$$

其中我们回顾一下， F赋有和范数 $\|\cdot\|_{L^2_T H^{\beta+1}} + \|\cdot\|_{H^1_T H^{\beta-1}}$。 然后，定义 f： $E \times F \to G$及其分量 f_1， f_2，如下所示

$$f(W,\rho) := \left(\partial_t \rho - \Delta \rho - \nabla \cdot (\rho \nabla W * \rho),\, \rho(0) - \phi\right) =: (f_1(W,\rho), f_2(W,\rho)). \tag{15}$$

我们来说明 f 在其定义域上是良定义的。由式 (10) 可知，且由于 $\rho \in F$, ρ(0)是良定义的，并且属于 H^β。然后，令 $\mathcal{T}$为如下定义的三线性算子

$$\mathcal{T}(r,V,s) := \nabla \cdot (r \nabla V * s). \tag{16}$$

下面的引理证明，当 $V \in E$且 r， $s \in F$时， $\mathcal{T}$(r， V， s)属于 L^2([0,T]; $H^{\beta-1}$)。

引理 2.4。 对于 $V \in E$、 r、 $s \in F$ 以及整数 $\beta \geq 0$,

$$\|\mathcal{T}(r,V,s)\|_{L^2([0,T];H^{\beta-1})} \leq C \|r\|_F \|V\|_E \|s\|_F$$

其中 $C = C(\beta, d)$ > 0。

证明。 引理 A.1 方程 (56) 与 $k = \beta$ 和 $l = 0$ 给出

$$\|\mathcal{T}(r,V,s)\|_{L^2([0,T];H^{\beta-1})} \lesssim_{\beta,d} \sup_{0\leq t\leq T} \|r\|_{H^\beta} \|s\|_{L^2([0,T];H^\beta)} \|\nabla V\|_{L^\infty}.$$

根据 E和 F上的范数定义， $\|\nabla V\|_{L^\infty} \leq \|V\|_E$和$\|s\|_{L^2([0,T];\ H^\beta)} \leq \|s\|_F$。最后，利用 (10) 得到 sup $_{0\leq t\leq T} \|r\|_{H^\beta} \lesssim \|r\|_F$，从而完成证明。

□

根据上面的引理，我们推得 T(ρ， W， ρ) 属于 L^2([0,T]; $H^{\beta-1}$)，因此 f(W， ρ) 可良好定义为 G 的一个元素。现在我们转而证明， f 具有第1.1节框架中所要求的性质。

命题 2.5（f 的导数）。对于任意整数 $\beta \geq 0$，映射 $f\colon\ E \times F \to G$（在（15）中定义）在 *Fr´echet* 意义下是C^∞。对于任意 W、$H \in E$、ρ 和 $s \in F$，f 的一阶导数为：

$$\begin{aligned} D_1 f(W,\rho)[H] &= (-\nabla \cdot (\rho \nabla H * \rho), 0) \\ D_2 f(W,\rho)[s] &= (\partial_t s - \Delta s - \nabla \cdot (s \nabla W * \rho) - \nabla \cdot (\rho \nabla W * s), s(0)). \end{aligned}$$

f的二阶导数如下：对于任意的 W、H_1、$H_2 \in E$以及 ρ、s_1、$s_2 \in F$：

$$\begin{aligned} D^2_{1,1} f(W,\rho)[H_1,H_2] &= (0,0) \\ D^2_{1,2} f(W,\rho)[H_1,s_2] &= (-\nabla \cdot (s_2 \nabla H_1 * \rho) - \nabla \cdot (\rho \nabla H_1 * s_2), 0) \\ D^2_{2,2} f(W,\rho)[s_1,s_2] &= (-\nabla \cdot (s_1 \nabla W * s_2) - \nabla \cdot (s_2 \nabla W * s_1), 0), \end{aligned}$$

以及 $D^2_{2,\ 1} f = D^2_{1,\ 2} f$。

证明。见附录 B.1。 □

$\mathcal{G}$ 在 (7) 中的导数的一般形式，需要将适当空间之间的 $D_2 f$ 的逆算子的存在性作为一个关键要素。考察前一个命题可知，对于此处考虑的非线性抛物型方程，$(D_2 f)^{-1}$ 本质上是一个线性抛物型偏微分方程的解算子，因此可以应用标准技巧。值得注意的是，由于所得（此时为椭圆型）微分算子中存在零阶系数，相应稳态方程中适当逆算子的存在性并不那么显然。

命题 2.6。设 ϕ 对某个 $\beta \geq 3 + d$ 满足假设 *2.3*。对于任意 $W \in \dot{W}^{2,\ \infty}$，回顾一下，$\rho_W$表示 (13) 的唯一解。那么，映射

$$D_2 f(W,\rho_W) : L^2([0,T]; H^{\beta+1}) \cap H^1([0,T]; H^{\beta-1}) \to L^2([0,T]; H^{\beta-1}) \times H^\beta$$

由命题 *2.5* 可知，这是一个线性同胚。

命题 *2.6* 的证明。固定 $W \in E$，令 $B := D_2 f(W,\rho_W)$，并回忆一下，使用附录 A 中的记号，有

$$Bu = (\partial_t u - \Delta u - \nabla \cdot (u \nabla W * \rho_W) - \nabla \cdot (\rho_W \nabla W * u), u(0)) = ((\partial_t - \mathcal{L}_W) u, u(0)).$$

定理 A.2 与定理 A.3 结合可得，对于任意 $g \in L^2([0,T];\ H^{\beta-1})$，存在唯一的 $u \in L^2([0,T];\ H^{\beta+1}) \cap H^1([0,T];\ H^{\beta-1})$，使得 $Bu = (g, 0)$。为了构造完整的逆算子，令 $(g,h) \in L^2([0,T];\ H^{\beta-1}) \times H^\beta$，并回忆一下，根据迹定理，[26, 定理 I.3.2], 存在$v \in L^2([0,T];\ H^{\beta+1}) \cap H^1([0,T];\ H^{\beta-1})$，使得 $v(0) = h$（例如，v满足初始条件为 h的热方程）。选取这样的 v，并令 w_v 为以下方程的唯一解

$$Bw_v = (g - (\partial_t - \mathcal{L}_W) v, 0)$$

根据定理 A.2，该定义良好。接着，令 $u := v + w_v$，并注意到

$$Bu = Bv + Bw_v = (g,h).$$

注意，u 是唯一的，因为若 u_1 和 u_2 满足 $Bu_1 = Bu_2 = (g,\ h)$，那么 $B(u_1 - u_2) = (0,\ 0)$，这进而推出 $u_1 = u_2$，由定理 A.2 可知。由此，我们构造出了逆算子 $B^{-1}(g,\ h) := u$，现在来证明该逆算子有界。再次使用表示式 $u = v + w_v$，其中固定 v 为初始条件为 h 的热流，结合定理 A.3 中的正则性估计、热方程的标准正则性估计、引理 2.4 以及 [33, 定理 5]，可得

$$\begin{aligned} \left\| B^{-1}(g,h) \right\|_F = \|u\|_F &\leq \|v\|_F + \|w_v\|_F \\ &\lesssim \|h\|_{H^\beta} + \|g - (\partial_t - \mathcal{L}_W) v\|_{L^2([0,T];H^{\beta-1})} \\ &\lesssim \|h\|_{H^\beta} + \|g\|_{L^2([0,T];H^{\beta-1})} + \|v\|_F \lesssim \|(g,h)\|_G + \|h\|_{H^\beta} \leq C \|(g,h)\|_G \end{aligned}$$

其中 $C = C(d,\ T,\ \|W\|_E,\ \|\phi\|_{H^\beta})$。这证明了 B^{-1} 有界，因此根据开映射定理，命题得证。

□

2.2.2 全局稳定性估计与伪线性化恒等式

在转向线性化的计算之前，我们先简要考察如何利用方程 (4) 为正向映射 $\mathcal{G}\colon \Theta \to X$ 推导全局稳定性估计，其中 Θ, X 是巴拿赫空间——事实上，只要 Θ 是巴拿赫空间的开凸子集，以下论证便成立。按照第1.1节的记号，我们希望得到如下形式的不等式

$$\|\theta_2 - \theta_1\|_1 \lesssim \|\mathcal{G}(\theta_2) - \mathcal{G}(\theta_1)\|_2 \tag{17}$$

对于范数的适当选择 $\|\cdot\|_1$, $\|\cdot\|_2$。隐含的乘法常数要求对所有满足 $\|\theta_i\|_{\mathcal{R}} \le M$, $i \in \{1,2\}$ 的 θ_1, $\theta_2 \in \Theta$ 一致，其中 $\|\cdot\|_{\mathcal{R}}$ 是适当的正则化范数，例如 $\mathcal{R} = H^a$, $a \ge 0$。假设 f 在 $\Theta \times X$ 上是 C^1，那么巴拿赫空间中的微积分基本定理（例如 [27, 命题 A.2.3],）应用于 $u \mapsto f(\theta_1 + (\theta_2 - \theta_1)u,\ \mathcal{G}(\theta_1))$，并结合链式法则（例如 [17, 定理 8.2.1], ），可得

$$f(\theta_2, \mathcal{G}(\theta_1)) - f(\theta_1, \mathcal{G}(\theta_1)) = \int_0^1 D_1 f(\theta_u, \mathcal{G}(\theta_1))[\theta_2 - \theta_1]\,\mathrm{d}u$$

其中 $\theta_u = \theta_1 + (\theta_2 - \theta_1)u$ 位于连接 θ_1 与 θ_2 的线段上。类似地，我们得到

$$f(\theta_2, \mathcal{G}(\theta_2)) - f(\theta_2, \mathcal{G}(\theta_1)) = \int_0^1 D_2 f(\theta_2, \mathcal{G}_u)[\mathcal{G}(\theta_2) - \mathcal{G}(\theta_1)]\,\mathrm{d}u$$

其中 $\mathcal{G}_u = \mathcal{G}(\theta_1) + (\mathcal{G}(\theta_2) - \mathcal{G}(\theta_1))u$。将这两个恒等式相加，并利用对任意 $\theta \in \Theta$ （由 (4) 可知），$f(\mathcal{G}(\theta), \theta) = 0$，可得

$$\int_0^1 D_2 f(\theta_2, \mathcal{G}_u)[\mathcal{G}(\theta_2) - \mathcal{G}(\theta_1)]\,\mathrm{d}u = -\int_0^1 D_1 f(\theta_u, \mathcal{G}(\theta_1))[\theta_2 - \theta_1]\,\mathrm{d}u. \tag{18}$$

利用博赫纳积分与有界线性映射可交换这一性质（例如参见 [27, 命题 A.2.2]），我们可以将前一个恒等式改写为

$$\mathcal{A}(\mathcal{G}(\theta_2) - \mathcal{G}(\theta_1)) = \mathcal{B}(\theta_2 - \theta_1) \tag{19}$$

其中，我们将平均线性算子 $\mathcal{A}\colon X \to Y$ 和 $\mathcal{B}\colon \Theta \to Y$ 定义为

$$\mathcal{A} = \int_0^1 D_2 f(\theta_2, \mathcal{G}_u)\,\mathrm{d}u, \quad \mathcal{B} = -\int_0^1 D_1 f(\theta_u, \mathcal{G}(\theta_1))\,\mathrm{d}u.$$

在偏微分方程逆问题的背景下，这一关系意味着参数 $\theta_2 - \theta_1$ 的变化会通过线性偏微分方程 (19) 映射为观测 $\mathcal{G}(\theta_2) - \mathcal{G}(\theta_1)$ 的变化。沿用 [36, 第 13], 章的术语，可以将 (19) 称为伪线性化恒等式——关于其在 McKean–Vlasov方程中的应用，参见注记 2.10。特别地，(19) 表明，要获得稳定性估计 (17)，只需找到第三个范数 $\|\cdot\|_3$，使得 $\|\mathcal{A}u\|_3 \lesssim \|u\|_2$ 且 $\|\mathcal{B}v\|_3 \gtrsim \|v\|_1$。$\mathcal{A}$ 的上界通常很容易得到，因为 $D_2 f(\theta_2, \mathcal{G}_s)$ 本身已是从 X 到 Y 的有界算子。$\mathcal{B}$ 的下界则更为棘手，也是映射 $\theta \mapsto \mathcal{G}(\theta)$ 稳定性的关键。

注 2.7。原则上，只要 f 在路径的像上保持不变，就可以沿着从 $(\theta_1, \mathcal{G}(\theta_1))$ 到 $(\theta_2, \mathcal{G}(\theta_2))$ 的任意 C^1 路径应用微积分基本定理，从而得到类似于 (18) 的恒等式。请注意，对于参数 θ 在线性地出现在映射 f中的问题，上文所选的分段 C^1 路径尤为适用。事实上，此时 $D_1 f$ 在 θ 中为常数，因此 $\mathcal{B}$ 可简化为 $\mathcal{B} = -D_1 f(\mathcal{G}(\theta_1))$。不过，针对特定问题，其他路径选择可能更加合适。

作为这些一般思想的一个例证，我们给出了 [33, 稳定性估计2] 定理2] 的新证明，避免了其中使用的不动点迭代。为处理可辨识性问题，我们沿用 [33]的做法，并对初始条件 ϕ提出以下假设。

假设 2.8。

1. ϕ 在 $\mathbb{T}^d$ 上严格为正，且 下确界$_{x\in\mathbb{T}^d}\,\phi(x)\,\geq\phi_{\min}>0.$

2. 存在 $c_*>0$，使得对所有 $k\in\mathbb{Z}^d$ 以及某个 $\zeta>\beta+d/2$

$$|\widehat{\phi}_k|\geq c_*|k|^{-\zeta}. \tag{20}$$

满足假设 2.8 和 2.3 的密度，对于任意整数 $\beta>0$ 和 $\zeta>\beta+d/2$ 的取值确实存在：[33, Example 3] for $\zeta=2m$，$m\in\mathbb{N}$，对此进行了讨论；将 [33] 中的拉普拉斯型密度替换为适当贝塞尔势的周期化概率核（见 [41] 第 132 页），这一结论还可推广到任意 $\zeta>\beta+d/2$。对此假设的一种解释是，相互作用粒子系统的初始状态与均匀平衡态相差足够大（见 [9]），因而观测到的动力学具有信息量。若动力学从 $\phi=1$ 开始，则除第一项外，其所有傅里叶模态都消失，且式 (20) 不成立。然而，对于所有 W，这样的 ϕ 都是动力学 (13) 的稳态，这意味着对所有 t 和 W，$\rho_W(t)=1$，因此无法得到稳定性估计 —— 更多讨论见 [33]。

定理 2.9（稳定性估计）。设 ϕ 满足假设 *2.3*，其中某个参数为 $\beta\geq 3+d$；并满足假设 *2.8*，其中某个参数为 $\zeta>\beta+d/2$。对于任意 $W\in\dot{W}^{2,\infty}$，令 ρ_W 表示偏微分方程 (13) 在初始条件为 ϕ 时的唯一解。考虑由 (33) 中的次数至多为 K 的零均值三角多项式构成的空间 E_K。固定 $M>0$、$K\geq 1$，并设 $W_1\in\dot{W}^{2,\infty}$、$W_2\in E_K$ 满足 $\|W_1\|_{W^{2,\infty}}$、$\|W_2\|_{W^{2,\infty}}\leq M$。那么，存在一个常数 S，它依赖于 d、T、M、c_*、$\phi_{\min}$、$\|\phi\|_{H^d}$（但不依赖于 K），使得

$$\|W_2-W_1\|_{L^2}\leq\|W_1-W_{1,K}\|_{L^2}+SK^{3\zeta/2}\left(\|\rho_{W_2}-\rho_{W_1}\|_{L^2([0,T];L^2)}+\|\rho_{W_2}-\rho_{W_1}\|_{H^1([0,T];H^{-2})}\right)$$

其中 $W_{1,K}$ 是将 W_1 投影到空间 E_K 上的投影。

定理 *2.9* 的证明。定义两条 C^1 路径 $W(u)$ 和 $\rho(u)$，使得

$$W(u):=W_1+(W_2-W_1)\,u,\quad \rho(u)=\rho_{W_1}+(\rho_{W_2}-\rho_{W_1})\,u$$

对所有 $0\leq u\leq 1$ 均成立。由巴拿赫空间中的微积分基本定理（例如，参见 [27, 命题 A.2.3]，其中该参考文献中的巴拿赫空间 X 取为 $L^2([0,T];H^{\beta-1})\times H^\beta$），如同式 (18)，我们得到

$$\begin{aligned}\int_0^1 D_2f(W_2,\rho(u))[\rho_{W_2}-\rho_{W_1}]\,\mathrm{d}u&=-\int_0^1 D_1f(W(u),\rho_{W_1})[W_2-W_1]\,\mathrm{d}u\\&=(\nabla\cdot(\rho_{W_1}\nabla(W_2-W_1)*\rho_{W_1}),0)\end{aligned} \tag{21}$$

其中 D_if，$i=1,2$ 在命题 2.5 中给出。由此我们推得

$$\nabla\cdot(\rho_{W_1}\nabla(W_2-W_1)*\rho_{W_1})=A(\rho_{W_2}-\rho_{W_1})\text{ in }L^2([0,T];H^{\beta-1}) \tag{22}$$

其中 $A(\rho_{W_2}-\rho_{W_1}):=\int_0^1 D_2f_1(W_2,\rho(u))[\rho_{W_2}-\rho_{W_1}]\mathrm{d}u$。我们先对右侧求一个上界，再使用 [33] 中的反卷积论证证明左侧的一个下界。

定义中 A 的博赫纳积分在 $L^2([0,T];H^{\beta-1})$ 中收敛，而该空间连续嵌入到 $L^2([0,T];H^{-2})$，因此该积分也在 $L^2([0,T];H^{-2})$ 中收敛。因此，利用三角不等式、引理 A.1 的方程 (56)（取 $k=l=0$ 和 $\nu=1$）、索伯列夫嵌入以及 [33, 定理 5]，可得

$$\begin{aligned}\|A(\rho_{W_2}-\rho_{W_1})\|_{L^2([0,T];H^{-2})}&\leq\int_0^1\|D_2f_1(W_2,\rho(u))[\rho_{W_2}-\rho_{W_1}]\|_{L^2([0,T];H^{-2})}\,\mathrm{d}u\\&\lesssim_{d,T}\int_0^1\|\rho(u)\|_{L^\infty_TL^\infty}\,\mathrm{d}u\;\|\nabla W_2\|_{L^\infty}\|\rho_{W_2}-\rho_{W_1}\|_{L^2_TL^2\cap H^1_TH^{-2}}\\&\lesssim_{d,T,M}(\|\rho_{W_1}\|_{L^\infty_TH^d}+\|\rho_{W_2}\|_{L^\infty_TH^d})\|\rho_{W_2}-\rho_{W_1}\|_{L^2_TL^2\cap H^1_TH^{-2}}\\&\leq C\|\rho_{W_2}-\rho_{W_1}\|_{L^2_TL^2\cap H^1_TH^{-2}}\end{aligned}$$

对某个常数 $C = C(d,\ T,\ M,\ \|\phi\|_{H^d}) > 0$。这就证明了 (22) 右端所需的上界。

现在来证明式 (22) 左端的一个下界。为简化记号，以下记 $H := W_2 - W_1$，以及 $W := W_1$。因此，我们要寻找 $\|\nabla\cdot(\rho_W\nabla H * \rho_W)\|_{L^2_T H^{-2}}$ 的一个下界。由 [33, 定理 5], 可知，密度 ρ_W 一致地有正下界：$\inf_{x\in\mathbb{T}^d, t\in[0,T]} \rho_W(t,x) \geq \rho_{\min} > 0$ 因此，在固定 $t \in [0,T]$ 后，算子 $u \mapsto -\nabla\cdot(\rho_W(t)\nabla u)$ 一致椭圆。由标准椭圆理论（例如参见 [21, 第 6] 章），我们知道，对于任意 $\psi \in \dot{L}^2$，存在唯一的 $u \in \dot{H}^2$，使得

$$-\nabla\cdot(\rho_W(t)\nabla u) = \psi, \quad \|u\|_{H^2} \leq C_{\mathrm{reg}} \|\psi\|_{L^2} \tag{23}$$

其中 $C_{\mathrm{reg}} > 0$ 仅取决于 $d,\ \rho_{\min}$ 以及 $\sup_{0\leq t\leq T} \|\nabla\rho_W(t)\|_{L^\infty}$ 的一个上界（论证方式与 [31], 中命题 A.5.3 的证明相同，同时使用环面上的庞加莱不等式）。现在令 $\psi := \rho_W(t) * H \in \dot{L}^2$，并令 $u \in \dot{H}^2$ 为上述椭圆型偏微分方程所对应的解。利用分部积分和正则性估计（23），我们得到

$$\begin{aligned}
\|\nabla\cdot(\rho_W(t)\nabla\psi)\|_{H^{-2}} = \sup_{\|v\|_{H^2}\leq 1} \langle \nabla\cdot(\rho_W(t)\nabla\psi), v\rangle_{L^2} &\geq \left\langle \nabla\cdot(\rho_W(t)\nabla\psi), -\frac{u}{\|u\|_{H^2}} \right\rangle_{L^2} \\
&= \left\langle \psi, -\frac{\nabla\cdot(\rho_W(t)\nabla u)}{\|u\|_{H^2}} \right\rangle_{L^2} \\
&= \frac{\|\psi\|_{L^2}^2}{\|u\|_{H^2}} \geq \frac{1}{C_{\mathrm{reg}}} \|\psi\|_{L^2}.
\end{aligned}$$

利用 [33, 定理 5] 和索伯列夫嵌入，我们得到 C_{reg} 可以选取得仅取决于$d,\ T,\ M,\ \phi_{\min},\ \|\phi\|_{H^d}$。对 $[0,T]$ 上的估计进行积分可得

$$\|\nabla\cdot(\rho_W\nabla H * \rho_W)\|_{L^2_T H^{-2}} \geq C_{\mathrm{reg}}^{-1} \|H * \rho_W\|_{L^2_T L^2}. \tag{24}$$

现在，为方便读者起见，我们重复 [33, 定理 2] 中的反卷积论证。根据 [33, 定理 5], 我们知道 $t \mapsto \rho_W(t)$ 是从 $[0,T] \to L^1$ 到 $C = C(d,\ T,\ M,\ \|\phi\|_{H^1})$ 的利普希茨映射，其利普希茨常数为 $C = C(d,\ T,\ M,\ \|\phi\|_{H^1})$（因为 $\sup_t \|\rho_W'(t)\|_{L^2} \leq C$）。此外，$L^1$- 范数一致地控制傅里叶系数，因此对于所有 $0 \leq t \leq T$，我们得到

$$|\widehat{\rho}_W(t,k) - \widehat{\phi}_k| \leq \|\rho_W(t) - \rho_W(0)\|_{L^1} \leq Ct$$

因此，利用方程 (20)，我们推出

$$|\widehat{\rho}_W(t,k)| \geq c\,|k|^{-\zeta} - Ct.$$

因此，当 t 足够小时，我们得到 $\rho_W(t)$ 的傅里叶系数下界

$$\inf_{|k|\leq K} |\widehat{\rho}_W(t,k)| \geq \frac{c_*}{2} K^{-\zeta}, \quad \text{for any } t \leq \min\left(T, \frac{c_*}{2C} K^{-\zeta}\right) =: t_0.$$

这意味着对所有 $t \leq t_0$,

$$\|H * \rho_W(t)\|_{L^2}^2 \geq \sum_{|k|\leq K} |\widehat{h}_k|^2 |\widehat{\rho}_W(t,k)|^2 \geq \frac{c_*^2}{4} K^{-2\zeta} \|W_2 - W_{1,K}\|_{L^2}^2.$$

由于利普希茨常数 C 会随 T 增长，我们总可以将其取到足够大（且与 K 无关，因为 $K \geq 1$），使得 $t_0 = \frac{c_*}{2} K^{-\zeta}$。因此，

$$\|H * \rho_W\|_{L^2_T L^2}^2 \geq \int_0^{t_0} \|H * \rho_W(t)\|_{L^2}^2 \,\mathrm{d}t \geq t_0 \frac{c_*^2}{4} K^{-2\zeta} \|W_2 - W_{1,K}\|_{L^2}^2 \geq CK^{-3\zeta} \|W_2 - W_{1,K}\|_{L^2}^2$$

其中 $C = C(d, T, M, c_*, \phi_{\min}, \|\phi\|_{H^d}) > 0$。将这一估计代入（24）后，可得到（22）左侧所需的下界。将其与此前推导出的上界结合可得

$$\|W_2 - W_{1,K}\|_{L^2} \le SK^{3\zeta/2} \left(\|\rho_{W_2} - \rho_{W_1}\|_{L^2([0,T];L^2)} + \|\rho_{W_2} - \rho_{W_1}\|_{H^1([0,T];H^{-2})} \right).$$

注意到 $\|W_2 - W_1\|_{L^2} \le \|W_1 - W_{1,\ K}\|_{L^2} + \|W_2 - W_{1,\ K}\|_{L^2}$，证明即告完成。 □

注记 2.10。 [McKean–Vlasov方程的伪线性化恒等式。] 前面的证明（通过 (21)）也意味着，$v := \rho_{W_2} - \rho_{W_1}$ 是以下线性偏微分方程的解

$$\begin{cases} \partial_t v - \Delta v - \nabla \cdot (v \nabla W_2 * \overline{\rho}) - \nabla \cdot (\overline{\rho} \nabla W_2 * v) = \nabla \cdot (\rho_{W_1} \nabla (W_2 - W_1) * \rho_{W_1}) \\ v(0) = 0, \end{cases}$$

其中 $\bar{\rho} := (\rho_{W_1} + \rho_{W_2})/2$。利用附录 A 中的正则性理论，可以直接由此推导出正向估计，方法与后文的定理 2.12 类似。

2.2.3 线性化与高阶导数

利用命题 2.5 和 2.6 中确立的 f 的性质，我们现在可以应用定理 B.1，求得由 McKean–Vlasov方程 (13) 导出的正向映射 $W \mapsto \rho_W$ 的导数。

定理 2.11 （McKean–Vlasov方程的线性化）。 设 ϕ 满足某个$\beta \ge 3 + d$下的假设 *2.3*。对于任意 $W \in \dot{W}^{2,\ \infty}$，令 ρ_W表示偏微分方程 （13）在初始条件ϕ下的唯一解。那么，正向映射

$$\mathcal{G} : \dot{W}^{2,\infty} \to L^2([0,T]; H^{\beta+1}) \cap H^1([0,T]; H^{\beta-1})$$

定义为 $\mathcal{G}（W）= \rho_W$，在 ***Fr´echet*** 意义下是无穷次可微的。

对于任意 W、$H \in \dot{W}^{2,\ \infty}$，其一阶 ***Fr´echet*** 导数 $D\rho_W[H]$ 是满足线性抛物型偏微分方程的唯一 $v \in L^2（[0,T]；\ H^{\beta+1}）\cap H^1（[0,T]；\ H^{\beta-1}）$

$$\begin{aligned} (\partial_t - \mathcal{L}_W)\, v &= \nabla \cdot (\rho_W \nabla H * \rho_W) \\ v(0) &= 0 \end{aligned} \tag{25}$$

其中 $\mathcal{L}_W = \Delta[\cdot] + \nabla \cdot （[\cdot]\nabla W * \rho_W）+ \nabla \cdot （\rho_W \nabla W * [\cdot]）$。

对于任意 $W, H_1, H_2 \in \dot{W}^{2,\ \infty}$，其二阶 ***Fr´échet*** 导数 $D^2\rho_W[H_1, H_2]$ 是满足线性抛物型偏微分方程的唯一解$v \in L^2（[0,T]；\ H^{\beta+1}）\cap H^1（[0,T]；\ H^{\beta-1}）$

$$\begin{aligned} (\partial_t - \mathcal{L}_W)\, v = {} & \nabla \cdot (D\rho_W[H_2] \nabla H_1 * \rho_W) + \nabla \cdot (\rho_W \nabla H_1 * D\rho_W[H_2]) \\ & + \nabla \cdot (D\rho_W[H_1] \nabla H_2 * \rho_W) + \nabla \cdot (\rho_W \nabla H_2 * D\rho_W[H_1]) \\ & + \nabla \cdot (D\rho_W[H_1] \nabla W * D\rho_W[H_2]) + \nabla \cdot (D\rho_W[H_2] \nabla W * D\rho_W[H_1]) \\ v(0) = {} & 0. \end{aligned} \tag{26}$$

定理 *2.11* 的证明。我们将定理 B.1 应用于 $E = W^{2,\ \infty}_{0}$, $F = L^2([0,T]; H^{\beta+1}) \cap H^1([0,T]; H^{\beta-1})$,$G = L^2([0,T]; H^{\beta-1}) \times H^{\beta}$, $U = E \times F$, $u = \mathcal{G}$, $V = E$。命题 2.5 和 2.6 确保满足关于 f 的条件。根据 [10], [24, 定理 3.4], ，我们知道对于每个 W，解 ρ_W 在F中是唯一的。因此，定理 B.1 的所有条件均已满足。

将命题 2.5 中 f 的导数表达式代入方程 (64)，即可得到 $\mathcal{G}$ 的一阶和二阶导数所满足的偏微分方程。

□

下面的正向 Lipschitz 估计是我们所用技巧的直接推论。它为 [33, 引理 3. (ii)] 提供了一种新的证明，并得到了稍好一些的界（右侧使用 $H^{-(\beta+1)}$-范数，而非 $H^{-\beta}$）。

定理 2.12 (正向估计)。 设 ϕ 满足假设 *2.3*，其中某个 $\beta \geq 3+d$。对于任意 $W \in \dot{W}^{2,\infty}$，令 ρ_W 表示偏微分方程 (13) 在初始条件 ϕ下的唯一解。设 $M > 0$ 和 $W_1, W_2 \in \dot{W}^{2,\infty}$满足 $\|W_1\|_{W^{2,\infty}} + \|W_1\|_{W^{2,\infty}} \leq M$。那么，存在一个常数 $L > 0$，其取决于d、T、M、$\|\phi\|_{H^{d+1}}$，使得

$$\|\rho_{W_2} - \rho_{W_1}\|_{L^2([0,T];L^2)} \leq L \|W_2 - W_1\|_{H^{-(\beta+1)}}.$$

定理 *2.12* 的证明。根据巴拿赫空间中的微积分基本定理（例如 [27, 命题 A.2.3]）和定理 2.11,

$$\rho_{W_2} - \rho_{W_1} = \int_0^1 D\rho_{W_1+(W_2-W_1)u}[W_2 - W_1]\,\mathrm{d}u$$

积分在 $L^2([0,T];\, H^{\beta+1}) \cap H^1([0,T];\, H^{\beta-1})$ 中收敛，因此在 $L^2([0,T];\, L^2)$ 中也收敛。取范数并应用三角不等式可得

$$\|\rho_{W_2} - \rho_{W_1}\|_{L^2([0,T];L^2)} \leq \int_0^1 \|D\rho_{W_u}[H]\|_{L^2([0,T];L^2)}\,\mathrm{d}u \tag{27}$$

其中 $W_u := W_1 + (W_2 - W_1)u$ 且 $H := W_2 - W_1$。利用 $D\rho_W[H]$ 满足方程 (25)，再结合定理 A.4 中在 $s = 2$ 下的正则性估计，可得对于足够大的某个 $L > 0$，

$$\begin{aligned}
\|D\rho_{W_u}[H]\|_{L^2([0,T];L^2)} &\lesssim \|\nabla \cdot (\rho_{W_u} \nabla H * \rho_{W_u})\|_{L^2([0,T];H^{-2})} \\
&\lesssim \sup_{0\leq t\leq T} \|\rho_{W_u}\|_{C^1} \|\rho_{W_u}\|_{L^2_T H^{\beta+1}} \|H\|_{H^{-(\beta+1)}} \\
&\leq L \|H\|_{H^{-(\beta+1)}}
\end{aligned}$$

其中，第二个不等式用到了引理 A.1 的方程 (58) 及 $\nu = \beta + 1$，最后一个不等式用到了索伯列夫嵌入和 [33, 定理 5]。常数 L 可取得足够大，使其仅依赖于 d, T, M, $\|\phi\|_{H^d}$（而不依赖于 s），因为 $\|W_u\|_{W^{2,\infty}} \leq 2M$。因此，将此估计代回 (27) 即完成证明。 □

$W \mapsto \rho_W$ 的高阶导数也可以在适当的范数下得到界，如下文引理所示。虽然我们只会在 $k = 1,2,3$ 的情形下使用这一结果，但这里仍给出一般性陈述。

引理 2.13. 设 ϕ 满足某个 $\beta \geq 3 + d$下的假设 *2.3*。对于任意 $W \in \dot{W}^{2,\infty}$，令 ρ_W表示偏微分方程 (13) 在初始条件 ϕ下的唯一解。设 $k \geq 1$ 为自然数，$M > 0$、$\frac{d}{2} < \gamma \leq \beta$和 $W \in \dot{W}^{2,\infty}$满足$\|W\|_{\dot{W}^{2,\infty}} \leq M$。则存在 $C = C(k, d, \gamma, T, M, \|\phi\|_{H^\gamma}) > 0$，使得对于所有$h_1$、.....、$h_k \in E$,

$$\left\| D^k \rho_W[h_1, \ldots, h_k] \right\|_{L^2([0,T];H^{\gamma+1})\cap H^1([0,T],H^{\gamma-1})} \leq C \|h_1\|_{L^2} \cdots \|h_k\|_{L^2}.$$

证明。 见附录 B.1。 □

对于合适的 $a_k > 0$，上述引理 2.13 右侧的界可以放宽为适用于较弱的 H^{-a_k}范数，但后文不会用到这一点。我们以验证梯度稳定性条件 (3) 结束本节。

定理 2.14 （梯度稳定性）。 设初始条件 ϕ 满足假设 *2.3* 和 *2.8*，其中参数为 $\beta \geq 3 + d$和 $\zeta > \beta + d/2$。设 $W \in \dot{W}^{2,\infty}$满足$\|W\|_{W^{2,\infty}} < M$，并将偏微分方程 (13) 在初始条件 ϕ下的唯一解记为 ρ_W。考虑由 (33) 中次数至多为 $K \geq 1$ 的零均值三角多项式构成的空间 E_K。那么，存在常数 $C = C(d, T, M, \beta, c_*, \phi_{\min}, \|\phi\|_{H^\beta}) > 0$（与 K 无关），使得对任意 $H \in E_K$,

$$\|D\rho_W[H]\|_{L^2([0,T];L^2)} \geq C K^{-3\zeta} \|H\|_{L^2}.$$

定理 *2.14* 的证明。 根据定理 2.11， $D\rho_W[H]$ 满足线性偏微分方程 (25)，因此

$$(\partial_t - \mathcal{L}_W)\, D\rho_W[H] = \nabla \cdot (\rho_W \nabla H * \rho_W) \text{ in } L^2([0,T]; H^{\beta-1}) \tag{28}$$

其中我们回顾一下， $\mathcal{L}_W = \Delta[\cdot] + \nabla \cdot ([\cdot]\nabla W * \rho_W) + \nabla \cdot (\rho_W \nabla W * [\cdot])$。正如我们在定理 2.9 的证明中所示（见（24）及其后的计算），右端的强迫项有如下下界：

$$\|\nabla \cdot (\rho_W \nabla H * \rho_W)\|^2_{L^2_T H^{-2}} \gtrsim \|H * \rho_W\|^2_{L^2_T L^2} \geq C K^{-3\zeta} \|H\|^2_{L^2}, \quad H \in \dot{E}_K$$

当 $C = C(d,\ T,\ M,\ c_*,\ \phi_{\min},\ \|\phi\|_{H^d}) > 0$。因此，只需证明对所有 $H \in E_K$

$$\|(\partial_t - \mathcal{L}_W)\, D\rho_W[H]\|^2_{L^2_T H^{-2}} \lesssim \|D\rho_W[H]\|_{L^2_T L^2} \|H\|_{L^2}. \tag{29}$$

利用 $\mathcal{L}_W$ 的定义、 $k = l = 0$ 和 $\nu = 1$ 时的引理 A.1 方程（56）、索伯列夫嵌入以及 [33, 定理 5], ，最后再利用插值（例如 [26, 第 4 章，命题 2.1]），我们得到，对于任意充分光滑的 u,

$$\begin{aligned}
\|(\partial_t - \mathcal{L}_W)u\|^2_{L^2_T H^{-2}} &\lesssim \|\partial_t u\|^2_{L^2_T H^{-2}} + \|\Delta u\|^2_{L^2_T H^{-2}} + \|\nabla \cdot (u \nabla W * \rho_W)\|^2_{L^2_T H^{-2}} + \|\nabla \cdot (\rho_W \nabla W * u)\|^2_{L^2_T H^{-2}} \\
&\lesssim \|u\|^2_{H^1_T H^{-2}} + \|u\|^2_{L^2_T L^2} (1 + \|\nabla W\|^2_{L^\infty} \|\rho_W\|^2_{L^\infty_T L^\infty}) \\
&\lesssim \|u\|^2_{H^1_T H^{-2}} + \|u\|^2_{L^2_T L^2} \\
&\leq C \|u\|_{L^2_T L^2} (\|u\|_{H^2_T H^{-2}} + \|u\|_{L^2_T L^2})
\end{aligned} \tag{30}$$

其中 $C = C(d,\ T,\ M,\ \|\phi\|_{H^d})$。将上述结果应用于 $u = D\rho_W[H]$，我们得出，证明（29）归结为证明

$$A + B := \|D\rho_W[H]\|_{H^2_T H^{-2}} + \|D\rho_W[H]\|_{L^2_T L^2} \lesssim \|H\|_{L^2}.$$

所需的关于 B 的界可直接由引理 2.13 应用于 $k = 1$ 和 $\gamma = d$ 得出。接下来只需对 A 给出界。为此，再次回顾 $D\rho_W[H]$ 满足偏微分方程 (28)，因此定理 A.5 的正则性估计、引理 A.1（取 $k = 1$ 和 $\nu = 0$）、索伯列夫嵌入以及 [33, 定理 5] 可得

$$\begin{aligned}
A = \|D\rho_W[H]\|_{H^2_T H^{-2}} &\lesssim \|\nabla \cdot (\rho_W \nabla H * \rho_W)\|_{H^1_T L^2} \\
&\lesssim \|\nabla \cdot (\rho_W \nabla H * \rho_W)\|_{L^2_T L^2} + \left\| \frac{\mathrm{d}}{\mathrm{d}t} \nabla \cdot (\rho_W \nabla H * \rho_W) \right\|_{L^2_T L^2} \\
&\lesssim_{d,T} \|H\|_{L^2} \|\rho_W\|_{H^1_T C^1} \|\rho_W\|_{L^\infty_T H^2} \leq C \|H\|_{L^2}
\end{aligned}$$

其中 $C = C(d,\ T,\ M,\ \beta,\ \|\phi\|_{H^\beta})$。由此完成了 (29) 的证明，进而完成了该命题的证明 . □

我们没有尝试优化前述结果中的梯度稳定性指数（因为下文不会用到它）。必要时，按照 [21, 第 7.1.3 节定理 6], 中的方法使用更高阶的时间正则性估计，可以对 $1/2 < \nu < 1$ 得到稍好的 $\|D\rho_W[H]\|_{L^2_T L^2} \geq C K^{-3\zeta/(2\nu)} \|H\|_{L^2}$。

3 多项式 时间后验抽样的 Langevin型算法 s

相互作用粒子系统的推断算法近来受到广泛研究，参见例如[16, 3, 37, 33] 及其中的参考文献。 [33] 提出的方法是，利用贝叶斯非线性反问题中的技术，从McKean–Vlasov方程 (6) 的解的含噪测量中推断相互作用势W ； 该方法将这些技术应用于 (1) 中由.....所导出的正向映射 $\theta \mapsto \mathcal{G}(\theta)$ [42, 28, 31]

$W \mapsto \rho_W$（因此此处为 $\theta = W$）。[33] 证明，对于 W，由高斯过程先验得到的后验均值向量在适当假设下可为 W 提供统计一致估计量。实践中，这类估计量通过马尔可夫链蒙特卡洛方法计算。本小节旨在运用第 2.2 节中的技术，证明引言中提到的未校正离散朗之万算法可用于近似此类后验均值，并且无论要求何种准确度，所需迭代次数都随相关参数呈多项式增长。由于非线性映射 $\theta \mapsto \mathcal{G}(\theta)$对应的后验吉布斯测度 (2) 不具备全局对数凹性，因此这种多项式运行时间算法的存在并非显而易见——关于在缺乏针对 H的强假设时，马尔可夫链蒙特卡洛算法的尺度行为的负面结果，参见 [5] 。

3.1 设定

我们首先确定本节中会用到的常数。回顾一下，$d \geq 1$ 表示环面 $\mathbb{T}^d$ 的维数，相互作用势 W 定义在该环面上。参数 β 表示初始条件 ϕ 的索伯列夫正则性，如假设 2.3 中所述；而 ζ 刻画了关于 ϕ的傅里叶模态下界 (20)。

我们再引入两个实值常数：α表示生成回归数据的真实值 W_0 的索伯列夫正则性（见下文 (46)）；w是一个缩放指数，用于确定 (39) 中定义的球 $\mathcal{B}_{\mathbf{r}}$ 的半径$\mathbf{r} \simeq D^{-w}$。可以将 $\mathcal{B}_{\mathbf{r}}$ 视为真实值 W_0 周围的局部区域，在该区域内，后验分布 (2) 将以较高的频率学派概率被证明是对数凹的。假设这些参数满足以下条件（不难看出，该条件组非空）：

$$\begin{cases} \beta \geq 4 + d \quad \text{and is an even integer}, \\ \alpha > 12\beta + 6d - 1, \\ \beta + \frac{d}{2} < \zeta < \frac{\alpha+1}{12}, \\ \frac{6\zeta}{d} < w < \frac{1}{d} \min\left(\frac{(\alpha+1)(\beta-2)}{\beta} - \frac{3\zeta}{2}, \alpha + 1 - 6\zeta \right). \end{cases} \tag{31}$$

回顾一下，对于任意这样的 β 和 ζ，都可以构造一个 ϕ，使其满足假设 2.3 和 2.8，具体讨论见假设 2.8 之后。

3.1.1 逼近空间

我们首先通过张量化一维实傅里叶基，构造 $L^2(\mathbb{T}^d)$ 的标准正交基。具体而言，对于 $k = (k_1, \ldots, k_d) \in \mathbb{Z}^d$, $x = (x_1, \ldots, x_d) \in \mathbb{T}^d$ 和 $m \in \mathbb{Z}$, $y \in \mathbb{T}$，定义

$$\tau_k(x) \coloneqq \prod_{j=1}^{d} T_{k_j}(x_j), \qquad T_m(y) \coloneqq \begin{cases} \sqrt{2}\cos(2\pi m y) & \text{for } m > 0, \\ 1 & \text{for } m = 0, \\ \sqrt{2}\sin(2\pi m y) & \text{for } m < 0. \end{cases} \tag{32}$$

对于 $K \geq 1$，我们令 E_K 表示次数至多为 K 的零均值三角多项式空间，并将其用作 W 的离散化空间。也就是说，

$$\dot{E}_K \coloneqq \operatorname{span}_{\mathbb{R}} \left\{ \tau_k : k \in \mathbb{Z}^d, k \neq 0, |k| \leq K \right\} = \mathfrak{R} \left(\operatorname{span}_{\mathbb{C}} \left\{ e_k : k \in \mathbb{Z}^d, k \neq 0, |k| \leq K \right\} \right) \tag{33}$$

其中 $e_k : x \mapsto e^{2i\pi k \cdot x}$，$\mathfrak{R}$ 表示实部。第二个等式很容易利用三角函数的标准积化和差恒等式证明。离散化空间$\dot{E}_K$ 的维数记为 D。空间 $\dot{E}_K$ 和 $\mathbb{R}^D$ 同构，因此下文中我们将默认把它们视为同一空间。

请注意，采用这一基时，截断发生的层级与主要参考文献 [33], 中所用的基 $(e_k)_k$ 的截断层级相同，也就是说，对于任意（实值） $f \in L^2$ 以及任意截断层级 K，

$$\sum_{|k|\le K} \langle f, e_k\rangle_{L^2}\, e_k = \sum_{|k|\le K} \langle f, \tau_k\rangle_{L^2}\, \tau_k.$$

这意味着，当 $f \in \dot{E}_K$ 时，其 Sobolev范数可由其 L^2-范数乘以一个依赖于 K 的惩罚项来控制；这一事实我们将在下文中反复使用。更准确地说，对于任意 $s \ge 0$ 和 $f \in E_K$，有：

$$\|f\|_{H^s}^2 \lesssim \sum_{|k|\le K} |k|^{2s} |\langle f, e_k\rangle|^2 \le K^{2s}\|f\|_{L^2}^2. \tag{34}$$

注释3.1。回忆一下，$D = \dim(\dot{E}_K) = \#\{k \in \mathbb{Z}^d : |k| \le K\} - 1$。通过用边长为 1 的立方体覆盖 $\mathbb{R}^d$ 中的单位球，我们可知，当 $K \ge \sqrt{d}/2$ 时，

$$\omega_d(K - \sqrt{d}/2)^d \le D(K) + 1 \le \omega_d(K + \sqrt{d}/2)^d$$

其中 ω_d 是单位球的体积。由此可立即得出，当 $K \to \infty$ 时，$D(K) \sim \omega_d K^d$，并且

$$D(K) \simeq_d K^d \text{ for any } K \ge 1. \tag{35}$$

我们将始终利用这一等价关系，在以 D 表述的第 5] 章条件与以 K^d 表述的条件之间进行转换。

3.1.2 似然函数的导数

设 ϕ 满足某个 $\beta \ge 3 + d$的假设 2.3。对于任意 $W \in \dot{W}^{2,\infty}$，令 ρ_W 表示满足初始条件 ϕ的偏微分方程 (13) 的唯一解。定义每个单个观测的对数似然 $(Y, t, X) \in \mathbb{R} \times [0,T] \times \mathbb{T}^d$ 如下

$$\ell(W) := -\frac{1}{2}|Y - \rho_W(t,X)|^2, \quad W \in \dot{E}_K, \tag{36}$$

其梯度为

$$\nabla\ell(W) := \lim_{h\to 0}\left(\frac{\ell(W + h\tau_k) - \ell(W)}{h}\right)_{\substack{|k|\le K\\ k\neq 0}} =: \left(\frac{\partial \ell}{\partial W_k}(W)\right)_{\substack{|k|\le K\\ k\neq 0}}$$

我们将其视为 $\mathbb{R}^D$中的一个向量。海森矩阵 $\nabla^2\ell$ 类似地定义为二阶偏导数组成的 $D \times D$ 矩阵。这些导数的定义适用于从 E_K 到 $\mathbb{R}$的任意函数。

根据 [33, 定理 5], 定理 2.11 和 Sobolev嵌入，解 ρ_W 及其更高阶的 Fr´echet 导数 $D^k\rho_W[H_1, \dots, H_k]$, $k = 1$，2，均有代表元在 (t, x) 上联合连续，因此点值是良定义的。因此，为了使 $\nabla\ell$ 和 $\nabla^2\ell$ 存在，只需映射

$$\mathcal{G}^{t,x} : \dot{E}_K \to \mathbb{R}, \quad W \mapsto \rho_W(t,x) \tag{37}$$

具有梯度和海森矩阵，这正是下一个引理要证明的内容。

引理 3.2. 设 ϕ 满足假设 *2.3*，其中参数为 $\beta \ge 3 + d$。对于任意 $W \in \dot{W}^{2,\infty}$，令 ρ_W 表示偏微分方程 (13) 在初始条件 ϕ 下的唯一解，并令 $D^k\rho_W$ 表示根据定理 *2.11* 得到的 $W \mapsto \rho_W$关于 k 阶 *Fr´echet* 导数。对于任意 $(t, x) \in [0,T] \times \mathbb{T}^d$以及 $W \in \dot{W}^{2,\infty}$，式 *(37)* 中 ^{x} 关于 $\mathcal{G}^t$ 的梯度 $\nabla\mathcal{G}^{t,x}(W)$和海森矩阵 $\nabla^2\mathcal{G}^{t,x}(W)$ 均存在，且可表示为

$$h_1^T\nabla\mathcal{G}^{t,x}(W) = D\rho_W[h_1](t,x), \quad h_1^T\nabla^2\mathcal{G}^{t,x}(W)h_2 = D^2\rho_W[h_1,h_2](t,x) \tag{38}$$

对于任意 h_1、$h_2 \in \mathbb{R}^D$。

引理 *3.2* 的证明。由 $D\rho_W$ 和 $D^2\rho_W$ 的线性性，只需证明当 $h \to 0$ 时，$\left(\mathcal{G}^{t,x}(W+h\tau_k) - \mathcal{G}^{t,x}(W)\right)/h$ 的极限存在，且等于 $D\rho_W[\tau_k](t,x)$。回顾一下，由 (10) 和索伯列夫嵌入可知，$L^2_T H^{\beta+1} \cap H^1_T H^{\beta-1} \hookrightarrow L^\infty_T L^\infty$。因此，

$$\begin{aligned}\left|\mathcal{G}^{t,x}(W+h\tau_k) - \mathcal{G}^{t,x}(W) - hD\rho_W[\tau_k](t,x)\right| &\le \|\rho_{W+h\tau_k} - \rho_W - D\rho_W[h\tau_k]\|_{L^\infty([0,T],L^\infty)} \\ &\lesssim \|\rho_{W+h\tau_k} - \rho_W - D\rho_W[h\tau_k]\|_{L^2_T H^{\beta+1}\cap H^1_T H^{\beta-1}} \\ &= o(\|h\tau_k\|_{W^{2,\infty}}) = o(h)\end{aligned}$$

其中最后一行用到了如下事实：根据定理 2.11，$D\rho_W$ 是 $\mathcal{G}: W^{2,\infty} \to L^2([0,T],H^{\beta+1}) \cap H^1([0,T], H^{\beta-1})$的 Fr´echet 导数。除以 $|h|$ 即可完成证明。 □

3.2 似然函数的局部性质

3.2.1 局部平均曲率

给定真实参数 $W_0 \in \dot{H}^{\alpha+1}$，定义$W_{0,K}$ 为其在 (33) 中向 $\dot{E}_K$ 的投影，并取$\mathbf{r} > 0$，使得球

$$\mathcal{B}_{\mathbf{r}} = \{W \in \dot{E}_K : \|W - W_{0,K}\|_{L^2} \le \mathbf{r}\}. \tag{39}$$

若 $\|W_0\|_{H^{\alpha+1}} \le M$，则 $\mathcal{B}_{\mathbf{r}}$中$\mathbf{r} = rD^{-w}$包含于一个半径为 $\bar{M}(d,r,M)$的 $W^{2,\infty}$-球内，该结论对K 和 W_0 一致成立：事实上，若 $W \in \mathcal{B}_{\mathbf{r}}$，则利用索伯列夫嵌入、(34) 和 (35)，

$$\|W\|_{W^{2,\infty}} \lesssim_d \|W_{0,K}\|_{H^{\alpha+1}} + \|W - W_{0,K}\|_{H^{2+d}} \lesssim_d M + rK^{2+d-wd} \le C_d(M+r) =: \bar{M} \tag{40}$$

其中我们利用了根据（31）中常数的选取，$wd > 6\zeta > \beta > d+2$。下面的定理表明，在局部高维区域 $\mathcal{B}_{\mathbf{r}}$ 内，负对数似然函数平均而言是凸的——特别是，它验证了 [35, 方程（43）]给出的关键曲率条件。请注意，稳定性指数自然取决于底层反卷积问题的“病态性”程度，该程度由假设 2.8 中的常数 ζ 决定。

定理 *3.3*（局部平均曲率）。按照（*31*）选取 α、β、ζ 和 w。设初始条件 ϕ 在此处对 β 和 ζ 的选取满足假设 *2.3* 和 *2.8*。对于任意 $W \in \dot{W}^{2,\infty}$，令 ρ_W 表示偏微分方程（*13*）在初始条件 ϕ 下的唯一解。令 $M > 0$ 和 $W_0 \in \dot{H}^{\alpha+1}$ 满足$\|W_0\|_{W^{2,\infty}} \le M$。那么，存在依赖于 d、T、M、β、c_*、$\phi_{\min}$、$\|\phi\|_{H^\beta}$ 的 $0 < r_{\max} \le 1$ 和 $c_0 > 0$，使得对于任意$0 < r \le r_{\max}$，$\mathbf{r} = rD^{-w}$、$\|\rho_{W_0} - \rho_{W_{0,K}}\|_{L^2([0,T];\,L^2)} \le \mathbf{r}$，我们有

$$\inf_{W \in \mathcal{B}_{\mathbf{r}}} \lambda_{\min}\left(\mathbb{E}_{W_0}[-\nabla^2 \ell(W)]\right) \ge c_0 D^{-6\zeta/d},$$

其中，$\mathcal{B}_{\mathbf{r}}$ 如（39）中所示；对于 $K \ge 1$，$\lambda_{\min}$ 表示括号内矩阵的最小特征值，期望是在（46）中的 $P^1_{W_0}$ 所对应的概率律下取得的。

定理 *3.3* 的证明。令 $W \in \mathcal{B}_{\mathbf{r}}$，并回忆（40）可知$\|W\|_{W^{2,\infty}} \le M$其中某个 $M = M(M,d) > 0$，因为 $r \le r_{\max} \le 1$。根据（36）中 ℓ 的定义，对于任意 $W \in E_K$，我们有

$$\begin{aligned}\nabla \ell(W) &= (Y - \rho_W(t,X))\nabla \mathcal{G}^{t,X}(W), \\ \nabla^2 \ell(W) &= -\nabla \mathcal{G}^{t,X}(W)(\nabla \mathcal{G}^{t,X}(W))^T + (Y - \rho_W(t,X))\nabla^2 \mathcal{G}^{t,X}(W).\end{aligned}$$

因此，利用 (38)，并对 $P^1_{W_0}$ from (46) 取期望，

$$h^T \mathbb{E}_{W_0}[-\nabla^2 \ell(W)]h = \frac{1}{T}\|D\rho_W[h]\|^2_{L^2_T L^2} - \frac{1}{T}\left\langle \rho_W - \rho_{W_0}, D^2\rho_W[h,h]\right\rangle_{L^2_T L^2}$$

对于任意 $h \in \mathbb{R}^d$。现在，通过对第一项应用定理 2.14，并对第二项应用柯西-施瓦茨不等式，我们得到

$$h^T \mathbb{E}_{W_0}[-\nabla^2 \ell(W)]h \gtrsim_T CK^{-6\zeta}\|h\|_{L^2}^2 - \|\rho_W - \rho_{W_0}\|_{L_T^2 L^2} \left\|D^2 \rho_W[h,h]\right\|_{L_T^2 L^2}.$$

但是根据定理 2.12，关于 $\|\rho_{W_0} - \rho_{W_{0,K}}\|_{L^2([0,T];L^2)}$ 的假设、(35) 式以及 $W \in \mathcal{B}_{\mathbf{r}}$ 这一事实，我们得到

$$\|\rho_W - \rho_{W_0}\|_{L_T^2 L^2} \le \|\rho_W - \rho_{W_{0,K}}\|_{L_T^2 L^2} + \|\rho_{W_{0,K}} - \rho_{W_0}\|_{L_T^2 L^2} \lesssim_d L\|W - W_{0,K}\|_{L^2} + rK^{-dw} \lesssim_d r(1+L)K^{-dw}.$$

此外，根据关于 $k=2$ 和 $\gamma = d$ 的引理 2.13，我们看到

$$\left\|D^2 \rho_W[h,h]\right\|_{L_T^2 L^2} \le C'\|h\|_{L^2}^2.$$

将这些估计结合起来，我们得到

$$h^T \mathbb{E}_{W_0}[-\nabla^2 \ell(W)]h \gtrsim_{d,T} (CK^{-6\zeta} - C'r(1+L)K^{-wd})\|h\|_{L^2}^2.$$

进一步利用 $wd > 6\zeta$ 可知，存在足够小的 $0 < r_{\max} \le 1$ 以及依赖于 $d, T, M, \beta, \|\phi\|_{H^\beta}$ 的 $c_0 > 0$，使得对于任意 $0 < r \le r_{\max}$，

$$h^T \mathbb{E}_{W_0}[-\nabla^2 \ell(W)]h \ge c_0 K^{-6\zeta}\|h\|_{L^2}^2.$$

最后，对所有满足 $\|h\|_{L^2} = 1$ 和 $W \in \mathcal{B}_{\mathbf{r}}$ 的 $h \in \mathbb{R}^d$ 取下确界，并应用 (35) 式，即可完成证明。

□

3.2.2 局部正则性

对于 $G\colon \dot{E}_K \to \mathbb{R}$，定义 $C^{2,1}(\mathcal{B})$-范数为

$$\|G\|_{C^{2,1}(\mathcal{B})} = \sup_{\substack{w,w' \in \mathcal{B} \\ w \ne w'}} \left\{ |G(w)| + \|G(w)\|_D + \left\|\nabla^2 G(w)\right\|_D + \frac{\left\|\nabla^2 G(w) - \nabla^2 G(w')\right\|_D}{\|w - w'\|_D} \right\}$$

其中，$\|\cdot\|_D$ 在作用于向量时表示 $\mathbb{R}^D$ 上的欧几里得范数，在作用于矩阵时表示由欧几里得范数诱导的算子范数。

命题 3.4 (局部正则性)。选取 α、β、ζ和 w，如(31)所示。令 ϕ 在此处选择的 β下满足假设 *2.3*。令 $M > 0$、$K \ge 1$和 $W_0 \in \dot{H}^{\alpha+1}$满足$\|W_0\|_{W^{2,\infty}} \le M$。那么，存在$c_1(d、T、M、\|\phi\|_{H^d}) > 0$，使得在为(39) 中的球 $\mathcal{B}_{\mathbf{r}}$选取半径 $\mathbf{r} = rD^{-w}$，且$0 < r \le 1$任意时，

$$\sup_{(t,x)\in[0,T]\times\mathbb{T}^d} \left\|\mathcal{G}^{t,x}\right\|_{C^{2,1}(\mathcal{B}_{\mathbf{r}})} \le c_1.$$

命题 *3.4* 的证明。我们首先用 $W \mapsto \rho_W$ 的 Fr´echet 导数表示 $\sup_{(t,x)\in[0,T]\times\mathbb{T}^d} \left\|\mathcal{G}^{t,x}\right\|_{C^{2,1}(\mathcal{B}_{\mathbf{r}})}$。利用 (38)，可得以下恒等式：

$$\sup_{t,x} \left|\mathcal{G}^{t,x}(W)\right| = \|\rho_W\|_{L^\infty([0,T];L^\infty)}, \tag{41}$$

$$\sup_{t,x} \left\|\nabla \mathcal{G}^{t,x}(W)\right\|_D = \sup_{h\in\mathbb{R}^d:\|h\|_D=1} \|D\rho_W[h]\|_{L^\infty([0,T];L^\infty)}, \tag{42}$$

$$\sup_{t,x} \left\|\nabla^2 \mathcal{G}^{t,x}(W)\right\|_D = \sup_{h_1,h_2\in\mathbb{R}^d:\|h_1\|_D=\|h_2\|_D=1} \left\|D^2\rho_W[h_1,h_2]\right\|_{L^\infty([0,T];L^\infty)}. \tag{43}$$

回顾一下，根据索伯列夫嵌入，对于任意$\gamma > \frac{d}{2}$，有 $L^2([0,T];\, H^{\gamma+1}) \cap H^1([0,T];\, H^{\gamma-1}) \hookrightarrow L^\infty([0,T];\, L^\infty)$。为简便起见，我们取 $\gamma = d$。因此，除去一个仅依赖于 d和 T的常数，上述显示中的范数可以替换为 $L^2([0,T];\, H^{d+1}) \cap H^1([0,T];\, H^{d-1})$。然后，由于[33,定理 5],，(41) 有界；由引理 2.13，(42) 和 (43) 有界。回顾一下，根据 (40)，$\|W\|_{W^{2,\,\infty}} \le M(d,\ M)$，因此最终的常数依赖于 $d,\ T,\ M,\ \|\phi\|_{H^d}$。

对于最后一个 Lipschitz项，令 $X = L^2([0,T];\, H^{d+1}) \cap H^1([0,T];\, H^{d-1})$，并像之前一样记为，

$$\sup_{t,x} \left\|\nabla^2 \mathcal{G}^{t,x}(W_2) - \nabla^2 \mathcal{G}^{t,x}(W_1)\right\|_D \lesssim_{d,T} \sup_{h_1,h_2 \in \mathbb{R}^d : \|h_1\|_D = \|h_2\|_D = 1} \left\|(D^2\rho_{W_2} - D^2\rho_{W_1})[h_1,h_2]\right\|_X .$$

现在，我们使用中值定理。根据巴拿赫空间中的微积分基本定理（例如，[27,命题 A.2.3]），

$$D^2\rho_{W_2} - D^2\rho_{W_1} = \int_0^1 D^3 \rho_{W_1 + (W_2 - W_1)u}[W_2 - W_1]\,\mathrm{d}u$$

根据定理 2.11，在 X 中收敛。但由于 $W_1, W_2 \in \mathcal{B}_{\mathbf{r}}$，方程 (40) 表明

$$\|W_1 + (W_2 - W_1)u\|_{W^{2,\infty}} \lesssim \bar{M}.$$

在 X 中取范数，并利用三角不等式和引理 2.13（其中 $k = 3$ 和 $\gamma = d$），我们得到

$$\left\|(D^2\rho_{W_2} - D^2\rho_{W_1})[h_1,h_2]\right\|_X \le C(d,T,M,\|\phi\|_{H^d}) \|h_1\|_D \|h_2\|_D \|W_2 - W_1\|_D$$

对 h_1, h_2, W_1, W_2 取上确界后，即得证明。

□

3.3 主要结果

3.3.1 收缩速率

我们首先引入一个速率 δ_N，以及与之对应的指数 $\eta < 1$:

$$\delta_N := N^{-\frac{\alpha+1}{2(\alpha+1)+d}}, \quad \eta := \frac{\beta-2}{\beta} - \frac{3\zeta}{2(\alpha+1)}. \tag{44}$$

注意，由于 (31) 中常数的选取，η 为（严格）正值。还要注意，该速率满足 $N\delta_N^2$ 趋于 $+\infty$。

我们考虑一种截断高斯先验，类似于 [33, 例 2]，其支撑集为 E_K:

$$W(x) = \frac{1}{\sqrt{N}\delta_N} \sum_{\substack{|k| \le K \\ k \ne 0}} \frac{1}{(1+|k|^2)^{(\alpha+1)/2}} g_k \tau_k(x), \qquad g_k \overset{\text{iid}}{\sim} \mathcal{N}(0,1), \tag{45}$$

其中 τ_k 来自 (32)，并令 $\Pi = \Pi_N := \text{Law}(W)$。我们从一个随机设计回归模型中观测数据

$$Y_i = \rho_W(t_i, X_i) + \varepsilon_i, \quad \varepsilon_i \overset{\text{iid}}{\sim} \mathcal{N}(0,1), \quad i = 1, \dots, N \tag{46}$$

其中，ρ_W 表示满足初始条件 ϕ 的偏微分方程 (13) 的唯一解，其中 $X_i \overset{\text{独立同分布}}{\sim}$ 服从 Unif($\mathbb{T}^d$)，且 $t_i \overset{\text{独立同分布}}{\sim}$ 服从 Unif($[0,T]$)。参数 W 下数据向量 $Z_N := \{(Y_i,\ t_i,\ X_i)\}_{i=1}^N$ 的分布记为 P_W^N。由这些数据 Z_N 得到的后验分布如 [31], 第 1.2.3 节所述。

$$\mathrm{d}\Pi(W \mid Z_N) = \frac{e^{\ell_N(W)}\,\mathrm{d}\Pi(W)}{\int_{\dot{E}_K} e^{\ell_N(w)}\,\mathrm{d}\Pi(w)}, \quad \ell_N(W) := -\frac{1}{2}\sum_{i=1}^N |Y_i - \rho_W(t_i, X_i)|^2, \quad W \in \dot{E}_K. \tag{47}$$

现在我们引入两个假设，应将其视为关于二元组 $(K,\ N)$ 以及真实值 W_0 的条件。在高维情形下（K 因而 $D(K)$ 很大），自然会预期 W_0 可由其在 E_K 上的投影 $W_{0,\ K}$ 充分逼近。例如，若选择截断值 K 随 N 足够快地增长，或 W_0 本身带限，就可以验证这一点，见注 3.8。由于也可以通过其他方式验证这一点，我们选择将偏差的控制与其余论证分开，并将其作为一项假设提出；但请参见注 3.8。

假设 3.5（控制偏差）。给定某个 $c_{\mathrm{err}} > 0$，假设

$$\left\|\rho_{W_0} - \rho_{W_{0,K}}\right\|_{L^2(\mathcal{X},\lambda)} \leq \frac{\delta_N}{2}, \tag{48}$$

以及

$$\|W_0 - W_{0,K}\|_{L^2} \leq c_{\mathrm{err}}\,(\delta_N)^{\eta}\,. \tag{49}$$

假设 3.6（先验光滑性）。给定某个 $c_{\mathrm{pr}} > 0$，假设

$$D \leq c_{\mathrm{pr}} N \delta_N^2.$$

首先，我们证明一个适用于当前情形的 [33, 定理 3] 版本。具体而言，我们证明，即使将后验分布限制在 W 上，只要其索伯列夫范数关于 K，N一致有界，后验分布仍会收缩至真实值附近；这类集合称为“正则化集合”。

定理 *3.7*（后验收缩）。按照（*31*）选择 α、β 和 ζ。设初始条件 ϕ 满足针对所选 β 的假设 *2.3* 和 *2.8*，并满足 ζ。考虑由（*33*）定义的空间 E_K，其元素为次数至多为 K 的零均值三角多项式；并考虑（*44*）中的速率 (δ_N, η)。设后验分布$\Pi(\ \cdot \mid Z_N)$如（*47*）所示，由（*45*）中的先验 Π_N和（*46*）中的数据 $Z_N = \{(Y_i, t_i, X_i)\}_{i=1}^N$得到。令 a、$M > 0$、b 大于某个普适常数，并令 $W_0 \in \dot{H}^{\alpha+1}$、$\|W_0\|_{H^{\alpha+1}} \leq M$、$K \geq 1$、$N \in \mathbb{N}$满足假设 *3.5* 和假设 *3.6*。

因此，存在一个常数 $m > 0$，其取值取决于 a、b、c_{err}、c_{pr}、α、β、d、c_*、T、M、$\phi_{\min}$、$\|\phi\|_{H^\beta}$，使得

$$P_{W_0}^N \left(\Pi(W \in \dot{E}_K : \|W - W_0\|_D \leq m(\delta_N)^{\eta}, \|W\|_{H^{\alpha+1}} \leq m | Z_N) \geq 1 - e^{-aN\delta_N^2}\right) \geq 1 - e^{-bN\delta_N^2}.$$

注 3.8（截断序列 $K = K_N$ 的选择）。令 $M > 0$ 和 $W_0 \in \dot{H}^{\alpha+1}$ 满足 $\|W_0\|_{H^{\alpha+1}} \leq M$，并回顾 L（d，T，M，$\|\phi\|_{H^d}$）为定理 2.12 中的利普希茨常数。定义

$$c_{\min} \coloneqq (2ML)^{\frac{1}{\alpha+1+\beta+1}}, \quad K_N = c(N\delta_N^2)^{1/d} \text{ for any } c \geq c_{\min}$$

然后，利用 (31) 中对常数的选择以及定理 2.12，可以验证，对于任意 $N \in \mathbb{N}$，二元组（N，K_N）满足假设 3.5 和 3.6，其中 c_{pr} 和 c_{err} 的取值仅取决于 α、d 和 c。因此，定理 3.7 的结论对 $K = K_N$、$D = D_N$ 成立，并且存在一个常数 $m > 0$，其取值仅取决于 a，b，α，β，d，c_*，T，M，$\phi_{\min}$，$\|\phi\|_{H^\beta}$。

定理 *3.7* 的证明。令 a, b, M 如定理所述。我们先像[33, 定理 1]中那样证明一个正向收缩率。也就是说，存在足够大的 $\tilde{m}(a, b, c_{\mathrm{pr}}, \alpha, \beta, d, T, M, \|\phi\|_{H^\beta}) > 0$

$$P_{W_0}^N \left(\Pi(\mathcal{E}_N | Z_N) \geq 1 - e^{-aN\delta_N^2}\right) \geq 1 - e^{-bN\delta_N^2}$$

其中

$$\mathcal{E}_N \coloneqq \left\{W \in \dot{E}_K : \|\rho_W - \rho_{W_0}\|_{L^2([0,T];L^2)} \leq \tilde{m}\delta_N, \|W\|_{H^{\alpha+1}} \leq \tilde{m}\right\}.$$

该界由 [31, 定理 2.2.2] 推出，更具体地说，是由 [31, 练习 2.4.3]推出的。ρ_W 的有界性证明与 [33, 定理 1]，中的证明相同，而 $W \mapsto \rho_W$ 的利普希茨连续性为：$\dot{L}^2 \to$

$L^2(\mathcal{X})$ 由定理 2.12 推出。此外，正则化集合可以如上选取，因为如果$\|W_1\|_{L^2} \le M\delta_N$，那么当 $\tilde{m}$ 足够大时，

$$\|W_1\|_{H^{\alpha+1}} \le K^{\alpha+1}\|W_1\|_{L^2} \lesssim_d c_{\mathrm{pr}}(N\delta_N^2)^{\frac{\alpha+1}{d}} M\delta_N = c_{\mathrm{pr}}M \le \tilde{m}$$

这里用到了 $W_1 \in E_K$ 、假设 3.6、方程 (34)、(35)，以及 (44) 中 δ_N 的定义。通过检查定理 1.3.2 和 2.2 的证明，以及练习 5.4 的证明，可知这些事件发生的概率至少为 $1 - e^{-bN\delta_N^2}$ 。2]：在该练习中，只要 $K = K(b)$ 足够大，就可以通过在伯恩斯坦不等式中令 $x = bN\sigma^2 + \log(2)$，使 b 任意大。证明定理 1.3.2] 时，利用定理 7.1 控制第一类检验错误，使其达到所需的指数速率。对于足够大的 m ，[22, 成立。

然后，我们应用定理 2.9 中的稳定性估计，将速率传递到 $\|W - W_0\|_{L^2}$。我们按照定理 3] 的证明进行时间插值，从而得到：在上述事件 $\mathcal{E}_N$ 上，

$$\|W - W_0\|_{L^2} \lesssim_{d,\beta,T} \|W_0 - W_{0,K}\|_{L^2} + \tilde{m}SK^{3\zeta/2}\delta_N^{\frac{(\beta-2)}{\beta}}.$$

现在使用假设 3.5、方程 (49) 和假设 3.6，得出当 m足够大时 ，

$$\|W - W_0\|_{L^2} \le m(\delta_N)^{\eta}$$

因为根据 (44) 中 η 的定义，$(N\delta_N^2)^{\frac{3\zeta}{2d}}\delta_N^{\frac{\beta-2}{\beta}} = \delta_N^{\eta}$。 □

注 3.9 （更快的收缩率）。请注意，[33, 定理 3] 得到的收缩率比前一个证明中的更快。虽然下文不会利用这一点，但这里要指出，我们的方法也能得到这一收缩率，事实上，通过定理 2.12 中的利普希茨估计，还能得到略快一些的收缩率。具体来说，若定义一个新的速率和一个新的指数

$$\tilde{\delta}_N := N^{-\frac{\alpha+1+\beta+1}{2(\alpha+1)+2(\beta+1)+d}}, \quad \tilde{\eta} := \frac{\beta-2}{\beta} - \frac{3\zeta}{2(\alpha+1)+2(\beta+1)}$$

那么根据定理 2.12， $W \mapsto \rho_W$ 仍然是从 $\dot{H}^{-(\beta+1)}$ 到 $L^2(\mathcal{X})$ 的利普希茨连续函数；并且如果 $\|W_1\|_{H^{-(\beta+1)}} \le M\delta_N$，那么当 $\tilde{m}$ 足够大时，仍然成立：

$$\|W_1\|_{H^{\alpha+1}} \le K^{\alpha+1+\beta+1}\|W_1\|_{H^{-(\beta+1)}} \lesssim_d c_{\mathrm{pr}}(N\tilde{\delta}_N^2)^{\frac{\alpha+1+\beta+1}{d}} M\tilde{\delta}_N = c_{\mathrm{pr}}M \le \tilde{m}$$

这里我们用到了 $W_1 \in E_K$，以及假设 3.6（用 δ_N 替代 δ_N）和方程 (34)、(35)。因此，只要假设 3.5 和 3.6 对 $(\delta_N,\ \tilde{\eta})$（而非 $(\delta_N,\ \eta)$）成立，且先验 (45) 按 δ_N而非 δ_N进行缩放，定理 3.7 的结论便以更快的速率 $(\delta_N)^{\eta}$成立。事实上，这一论证还意味着 $\|\rho_W - \rho_{W_0}\|_{L_T^2 L^2}$ 的后验收缩率为 δ_N，我们认为这是极小极大最优的。

3.3.2 后验分布的对数凹近似

这里我们回顾替代后验分布 $\Pi(\cdot \mid Z_N)$ 的定义，该定义与 [31, 第 5.1 节1]完全相同。令

$$\mathbf{r} := rD^{-w} \text{ for some } 0 < r \le r_{\max} \tag{50}$$

其中 $r_{\max}$ 如定理 3.3 所示， w 如式 (31) 所示。然后，令 $W_{\mathrm{init}} \in E_K$ 为一个“热启动”初始值，使得

$$\|W_{\mathrm{init}} - W_{0,K}\|_D \le \frac{\mathbf{r}}{8}. \tag{51}$$

我们需要两个辅助函数： $g_{\mathbf{r}}$ （具有强凸尾部）和 $\alpha_{\mathbf{r}}$ （截断函数），定义如下。对于某个光滑且关于 0 对称的 φ： $\mathbb{R} \to [0,\infty)$ ，其支集包含于 $[-1,1]$，且积分为 $\int_{\mathbb{R}} \varphi(x)\,\mathrm{d}x = 1$，我们首先定义光滑化核 $\varphi_h(x)$ ：$= h^{-1}\varphi(x/h)$ $h > 0$。然后定义 $\tilde{\gamma}_{\mathbf{r}}$， $\gamma_{\mathbf{r}}$： $\mathbb{R} \to \mathbb{R}$，如下：

$$\tilde{\gamma}_{\mathbf{r}}(t) := \begin{cases} 0 & \text{if } t < 5\mathbf{r}/8 \\ (t - 5\mathbf{r}/8)^2 & \text{if } t \geq 5\mathbf{r}/8 \end{cases}, \quad \gamma_{\mathbf{r}}(t) := \left[\varphi_{\mathbf{r}/8} * \tilde{\gamma}_{\mathbf{r}}\right](t).$$

此外，设 $\alpha\colon\ [0,\infty) \to [0,1]$ 为光滑函数，并满足：当 $t \in [0,3/4]$ 时，$\alpha(t) = 1$；当 $t \in [7/8,\infty$ 时，$\alpha(t) = 0$）。最后，我们定义 $g_{\mathbf{r}}\colon\ E_K \to [0,\infty)$ 和 $\alpha_{\mathbf{r}}\colon\ E_K \to [0,1]$，如下：

$$g_{\mathbf{r}}(W) := \gamma_{\mathbf{r}}\left(\|W - W_{\text{init}}\|_D\right), \quad \alpha_{\mathbf{r}}(W) = \alpha\left(\frac{\|W - W_{\text{init}}\|_D}{\mathbf{r}}\right).$$

令 $\lambda_{\min} := \max\left(N\log(N)/\mathbf{r}^2, CN(c_1+1)(1+\mathbf{r}^{-2})\right)$ 其中 C 按命题 5.1.2] 中的方式选取，而 c_1 在命题 3.4 中引入。现在，对于 (47) 中的 ℓ_N 以及某个 $\lambda \geq \lambda_{\min}$，定义替代似然函数

$$\tilde{\ell}_N(W) = \alpha_{\mathbf{r}}(W)\ell_N(W) - \lambda g_{\mathbf{r}}(W), \quad W \in \dot{E}_K,$$

它在 $-\ell_N$ “尾部”进行了“凸化”。于是，替代后验分布 $\Pi(\cdot \mid Z_N)$ 通过其密度定义

$$\mathrm{d}\tilde{\Pi}(W \mid Z_N) = \frac{e^{\tilde{\ell}_N(W)}\,\mathrm{d}\Pi(W)}{\int_{\dot{E}_K} e^{\tilde{\ell}_N(w)}\,\mathrm{d}\Pi(w)}, \tag{52}$$

该函数在全局上强对数凹。请注意，该替代函数取决于初始值 W_{init} 的选择；在下文定理中，可取为 $W_{\text{init}} = W_{0,K}$。对于下文关于马尔可夫链蒙特卡罗运行时间的应用，需要找到可行的初始值。这个问题本身也值得研究，将在其他工作中探讨。

定理 *3.10*（对数凹的 *Wasserstein* 近似）。选取 α、β、ζ 和 w，如 *(31)* 所示。令初始条件 ϕ 对此处选取的 β 和 ζ 满足假设 *2.3* 和 *2.8*。设后验分布 $\Pi(\cdot \mid Z_N)$ 由 *(47)* 中的先验 Π_N（见 *(45)*）和数据 $Z_N = \{(Y_i, t_i, X_i)\}_{i=1}^N$（见 *(46)*）给出。回顾 *(44)* 中的速率 $\tilde{\ }(\delta_N, \eta)$，并令替代后验分布 $\Pi(\cdot \mid Z_N)$ 如 *(52)* 所示，其中某个 W_{init} 满足 *(51)*。令 $M > 0$、$\|W_0\|_{H^{\alpha+1}} \leq M$、$K \geq 1$、$N \in \mathbb{N}$、$W_0 \in \dot{H}^{\alpha+1}$ 使得假设 *3.5* 和 *3.6* 成立，并且 $\|\rho_{W_0} - \rho_{W_0,\kappa}\|_{L^2([0,T];L^2)} \leq r$ 如 *(50)* 中所定义。

那么，存在 $C_{\text{pr}}(d, w, \alpha, \beta, \zeta, r)$，以及关于 W_0, K, N 一致的某些 $b, c > 0$ 和一个通用常数 $C > 0$，使得如果 c_{pr} 在假设 (3.6) 中满足 $c_{\text{pr}} \leq C_{\text{pr}}$，

$$P_{W_0}^N\left(\mathcal{W}_2^2\left(\Pi(\cdot \mid Z_N), \tilde{\Pi}(\cdot \mid Z_N)\right) \leq e^{-N\delta_N^2}\right) \geq 1 - C\left(e^{-bN\delta_N^2} + e^{-cND^{-12\zeta/d}}\right) \xrightarrow[N\to\infty]{} 1$$

其中 $\mathcal{W}_2$ 是在 (9) 中定义的 *Wasserstein-2* 距离。

定理 *3.10* 的证明。我们将验证 [31, 定理 5.1.3] 的假设对于 $\mathbf{r} = rD^{-w}$ 成立，其中 $0 < r \leq r_{\max}$ 如定理 3.3 中所述。我们首先验证条件 5.1.1 的子条件，其中 [31] 中的 $\Theta = \mathscr{R} = H^{\alpha+1}$ 与 $\alpha + 1 > 2 + d/2 > 1 + d/2$。

条件 2.1.1 对于 $\kappa = \beta + 1$ 由 [33, 定理 5] （结合通常的索伯列夫嵌入）和定理 2.12 得到满足，因此对于 $\kappa = 0$ 也成立。

由于问题具有反卷积性质，条件 2.1.4 严格来说并不成立。不过，可以用定理 3.7 替代，该定理给出的收缩率为 $(\delta_N)^\eta$；此外，由于 (31) 中常数的选取，还得到 $0 < \eta \leq 1$。在 [31, 定理 5.1.3], 的证明中，只需将项 III 的界中的第一个公式替换为定理 3.7 的结果，即可得到相同的结论。

如命题 3.4 所示，条件 3.2.1 对 $\kappa_2 = 0$ 成立。

根据定理 3.3 以及 [31, 条件 3.2.2], 下方的注释，条件 3.2.2 对 $\kappa_0 = 6\zeta/d$ 和 $\kappa_1 = 0$ 成立；此外，根据假设，$\|\rho_{W_0} - \rho_{W_{0,K}}\|_{L^2([0,T];L^2)} \leq \mathbf{r}$。

现在我们来验证 [31, 条件 5.1.1]中的方程 (5.5)。根据 (31) 中常数的选取以及假设 3.6,

$$\log(N)^2(\delta_N)^\eta D^w \xrightarrow[N\to+\infty]{} 0, \quad \log(N)^2\delta_N D^{w+6\zeta/d} \xrightarrow[N\to+\infty]{} 0$$

˜特别地，存在 $N_{\min}(d,w,\alpha,\beta,\zeta,r)$，使得对所有 $N \geq N_{\min}$，$\mathbf{r} \geq \delta_N \log N$。于是，可以将 $C_{\mathrm{pr}}(d,w,\alpha,\beta,\zeta,r)$ 选得足够小，以控制项 $N \leq N_{\min}$，从而使得只要 $c_{\mathrm{pr}} \leq C_{\mathrm{pr}}$，˜
$\mathbf{r} \geq \delta_N \log N$，对所有 $N \in \mathbb{N}$ 均成立。

最后，$\lambda \geq N\log(N)D^{\kappa_1}/\mathbf{r}^2$ 中的条件 [31, 定理 5.1.3]（在该参考文献中，λ 记作 K）根据式 (52) 中替代函数的构造而得到满足。另请注意，根据假设 3.6 以及式 (31) 中常数的选取，$ND^{-12\zeta/d} \xrightarrow[N\to\infty]{} +\infty$。 □

3.3.3 热启动 ULA 的多项式混合时间界

我们先简要介绍非校正朗之万算法（ULA），定理 3.11 将对其进行研究。给定一个具有正密度 p 且定义在 $\mathbb{R}^D$上的目标测度，该算法源自对朗之万扩散进行欧拉–丸山离散化

$$\mathrm{d}L_t = \nabla \log p(L_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t$$

该扩散是一个连续时间马尔可夫过程，其不变测度的密度为 p。在我们的情形中，感兴趣的˜密度为 $p = \mathrm{d}\Pi_N$，即式 (52) 中替代后验分布的密度。由于该密度是对数凹的，我们可以像 [20] 中那样，利用现有的非校正朗之万算法理论得到混合时间界。接着，我们可以利用替代后验分布在瓦瑟斯坦距离意义下近似真实后验这一性质（定理 3.10），进而得到真实后验的混合时间界。这正是下文定理 3.11 的研究对象。

在将 $\dot{E}_K$ 与 $\mathbb{R}^D$通常地对应起来后，我们可以将式 (45) 中的先验 W 视为一个协方差矩阵为

$$\Sigma := (N\delta_N^2)^{-1}\Sigma', \quad W \sim \mathcal{N}(0,\Sigma) \tag{53}$$

维高斯向量，其中 Σ' 是一个对角矩阵，其对角元素为 $(1+|k|^2)^{-(\alpha+1)}$，其中 $k \in \mathbb{Z}^d \setminus \{0\}$, $|k| \leq K$按重数计。于是，替代后验概率密度为

$$\mathrm{d}\tilde{\Pi}_N(W) = \frac{1}{Z} e^{\tilde{\ell}_N(W) - \frac{1}{2}W^T\Sigma^{-1}W}, \quad W \in \mathbb{R}^D$$

归一化常数为 Z。对于满足式 (51) 的某个初始值 W_{init} 以及步长 $\gamma > 0$，我们定义 ULA 马尔可夫链 $(\vartheta_k)_{k\geq 0} = (\vartheta_k(W_{\mathrm{init}},\ \gamma))_{k\geq 0}$ 如下：

$$\begin{aligned} \vartheta_0 &= W_{\mathrm{init}}, \\ \vartheta_{k+1} &= \vartheta_k + \gamma\left[\nabla\tilde{\ell}_N(\vartheta_k) - \Sigma^{-1}\vartheta_k\right] + \sqrt{2\gamma}\xi_{k+1} \quad \text{for } k = 0,1,\ldots \end{aligned} \tag{54}$$

其中 $\xi_k \overset{\text{独立同分布}}{\sim} N(0,\ I_D)$。注意，方括号中的漂移项恰好是 $\nabla\log(\mathrm{d}\,\tilde{\Pi}_N(W))$.

定理 3.11 (瓦瑟斯坦混合时间)。 假设定理 *3.10* 的假设成立，并且存在一个满足式 (51) 的可行初始值 W_{init}，用于式 (54) 中的马尔可夫链 $(\vartheta_k)_{k\geq 0}$。此外，假设对于定理 *3.10* 中的某个 $0 < c_{\mathrm{pr}} \leq C_{\mathrm{pr}}$，假设 *3.6* 成立。

那么，对于任意给定的准确度 $\varepsilon > 0$，存在一个步长 $\gamma = \gamma(\varepsilon, D, N)$ 和一个整数$k_{\mathrm{mix}} = O(\varepsilon^{-b_1}N^{b_2}D^{b_3})$，其随 ε^{-1}、D 和 N 呈多项式增长，使得对于所有 $k \geq k_{\mathrm{mix}}$，

$$P_{W_0}^N\left(\mathcal{W}_2^2(\mathcal{L}(\vartheta_k),\Pi(\cdot \mid Z_N)) \leq e^{-N^{d/(2(\alpha+1)+d)}} + \varepsilon\right) \geq 1 - C\left(e^{-bN\delta_N^2} + e^{-cND^{-12\zeta/d}}\right) \xrightarrow[N\to\infty]{} 1$$

对于某些 b_1、b_2、b_3、b、c、$C > 0$。

定理 *3.11* 的证明。 注意，由于假设 3.6 结合（31）中的常数选择意味着$D \lesssim ND^{-12\zeta/d}$， [31, 定理 5.21] 的结论成立（此外，这是因为我们在替代函数的构造中对 λ 的选择足够大）。我们选取事件 $\mathscr{E} = \mathscr{E}_{\mathrm{conv}} \cap \mathscr{E}_{\mathrm{wass}}(\rho)$，其中 $\rho = e^{-N\delta_N^2}$，并应用该定理。通过先将 γ 取得足够小，再将 k 取得足够大，我们可以使 [31, 定理 5.21] 中方程（5.21）右侧小于 $\rho + \varepsilon$。下界 k_{mix} 随着N 和 D 呈多项式增长，这是因为（53）中定义的 Σ 的特征值关于 N 和 D呈多项式缩放；另见（35）。

因此，如果我们能够证明 $\mathscr{E}$ 以所需的概率成立，证明便告完成。根据同样的观察 $D \lesssim ND^{-12\zeta/d}$， [31, 定理 3.2.3] 的结论也同样成立。因此，结合定理 3.10，

$$P_{W_0}^N(\mathscr{E}_{\mathrm{conv}} \cap \mathscr{E}_{\mathrm{wass}}(e^{-N\delta_N^2})) \geq 1 - C\left(e^{-bN\delta_N^2} + e^{-cND^{-12\zeta/d}} + e^{-N/8}\right).$$

最后一项 可以忽略，并可并入常数 C中。这就证明了该定理。 □

注释 3.12。 定理 3.11 还揭示了计算后验均值向量$\Pi(W \mid Z_N)$的复杂度。事实上，只要存在可行的初始值，任何后验分布的利普希茨泛函都可以通过沿着形式为 ϑ_k 的链计算遍历平均，在多项式时间内近似得到

$$\frac{1}{J}\sum_{k=J_{\mathrm{in}}+1}^{J_{\mathrm{in}}+J} H(\vartheta_k)$$

其中 J_{in} 是预热期， H 是我们关注的利普希茨泛函。详情见 [31,定理 5.2.2]。

线性化 McKean–Vlasov 方程的偏微分方程估计

在本节中，对于某个固定的 $T > 0$，我们研究以下关于$\mathcal{X} = [0,T] \times \mathbb{T}^d$的柯西问题的存在性和正则性：

$$(\partial_t - \mathcal{L}_W)\, u = f, \quad u(0) = 0 \tag{55}$$

其线性微分算子为

$$\mathcal{L}_W = \Delta[\cdot] + \nabla \cdot ([\cdot]\nabla W * \rho_W) + \nabla \cdot (\rho_W \nabla W * [\cdot]).$$

这里我们假设$\|W\|_{W^{2,\infty}} \leq M$其中 $M > 0$ 为某个固定量，且 ρ_W 是非线性偏微分方程 **(13)** 的给定解——事实上， ρ_W 可以替换为任何索伯列夫范数具有适当界的函数。源函数 f 依赖于时间和空间，其条件将在后文说明。我们先给出一个辅助引理。

引理 A.1. 回顾(16)，三线性算子 $\mathcal{T}$ 定义为 $\mathcal{T}(f,V,g) := \nabla \cdot (f\nabla V * g)$。对于任意使下列表达式有意义的 f、V、 g ， 我们有以下估计：

1. 对任意整数 $k \geq 0$、 $0 \leq l \leq k$ 以及任意 $\nu \in \mathbb{R}$，

$$\|\mathcal{T}(f,V,g)\|_{H^{k-1}} \lesssim_{k,d} \min\left\{\|\nabla V\|_{W^{l,\infty}} \|f\|_{H^k} \|g\|_{W^{k-l,1}}, \|V\|_{H^\nu} \|f\|_{W^{k,\infty}} \|g\|_{H^{k+1-\nu}}\right\}. \tag{56}$$

特别是，由于 $\mathbb{T}^d$ 有界，

$$\|\mathcal{T}(f,V,g)\|_{H^{k-1}} \lesssim_{k,d} \|\nabla V\|_{W^{l,\infty}} \|f\|_{H^k} \|g\|_{H^{k-l}}.$$

2. 对任意实数 $\gamma > \frac{d}{2}$ 和 ν，

$$\|\mathcal{T}(f,V,g)\|_{H^{\gamma-1}} \lesssim_{\gamma,d} \|f\|_{H^\gamma} \|g\|_{H^{\gamma+1-\nu}} \|V\|_{H^\nu}. \tag{57}$$

3. 对任意实数 ν,

$$\|\mathcal{T}(f,V,g)\|_{H^{-2}} \le \|f\|_{W^{1,\infty}} \|V\|_{H^{-\nu}} \|g\|_{H^{\nu}}. \tag{58}$$

证明。 该证明是标准的，包含对环面上索伯列夫函数卷积所使用的傅里叶分析不等式的应用，具体如下

$$\|u*v\|_{H^{a+b}} \le \|u\|_{H^a}\|v\|_{H^b},\ a,b\in\mathbb{R}, \tag{59}$$

我们仅概述主要思路。首先注意到，对于任意 $\mu\in\mathbb{R}$,

$$\|\mathcal{T}(f,V,g)\|_{H^{\mu-1}} = \|\nabla\cdot(f\nabla V * g)\|_{H^{\mu-1}} \le \|f\nabla V * g\|_{H^{\mu}}$$

这将始终是下面各个证明的第一步。

1. 对于最小值的第一部分，我们可以估计

$$\|f\nabla V * g\|_{H^k} \lesssim_{k,d} \|f\|_{H^k}\|\nabla V * g\|_{W^{k,\infty}} \lesssim_{k,d} \|\nabla V\|_{W^{l,\infty}}\|f\|_{H^k}\|g\|_{W^{k-l,1}}$$

如所要求的。对于最小值的第二部分，我们注意到

$$\|f\nabla V * g\|_{H^k} \lesssim_{k,d} \|f\|_{W^{k,\infty}}\|\nabla V * g\|_{H^k} \le \|f\|_{W^{k,\infty}}\|V * g\|_{H^{k+1}}$$

然后结合 $a=\nu$ 和 $b=k+1-\nu$ 使用（59）即可得出结论。

2. 由于 $\gamma > d/2$，逐点乘子在 H^γ 上是利普希茨的，因此

$$\|f\nabla V * g\|_{H^\gamma} \lesssim_{\gamma,d} \|f\|_{H^\gamma}\|V * g\|_{H^{\gamma+1}}.$$

结合 (59) 和 $a=\gamma+1-\nu$、 $b=\nu$ 可得结论。

3. 令 $\psi := V * g$，并回顾

$$\|f\nabla\psi\|_{H^{-1}} = \sup_{\substack{\varphi\in H^1\\ \|\varphi\|_{H^1}\le 1}} \langle f\nabla\psi,\varphi\rangle_{L^2}.$$

积分　　分部积分、乘积法则和柯西-施瓦茨不等式可得

$$|\langle f\nabla\psi,\varphi\rangle_{L^2}| \le |\langle\psi\nabla f,\varphi\rangle| + |\langle f\psi,\nabla\varphi\rangle| \le \|f\|_{W^{1,\infty}}\|\psi\|_{L^2}\|\phi\|_{H^1}.$$

因此，对所有 $\varphi\in H^1$ 取上确界后，我们得到

$$\|f\nabla V * g\|_{H^{-1}} \le \|f\|_{W^{1,\infty}}\|V * g\|_{L^2}.$$

利用 $a=\nu$ 和 $b=-\nu$中的 (59)，即可得到结论。

□

定理 A.2。 设 $M>0$、$W\in\dot{W}^{2,\infty}$满足$\|W\|_{W^{2,\infty}}\le M$，并令出现在 $\mathcal{L}_W$中的 ρ_W，即(55) 中的唯一解，记为 (13)。令 ϕ（ ρ_W的初始条件）对$\beta\ge 3+d$满足假设 *2.3*。则对于任意 $f\in L^2([0,T];\ L^2)$，存在 u 的唯一强解 u，满足(55)，且

$$u\in L^2([0,T];H^2)\cap H^1([0,T];L^2).$$

此外，存在一个常数 $C=C(d,\ T,\ M,\ \|\phi\|_{H^1})$ 使得对于 $u'=\mathrm{d}u\,\mathrm{d}t$　　—

$$\|u\|_{L^2([0,T];H^2)} + \|u\|_{L^\infty([0,T];H^1)} + \|u'\|_{L^2([0,T];L^2)} \le C\|f\|_{L^2([0,T];L^2)}.$$

证明。 我们将通过推导完整证明所需的先验估计，说明如何处理非局部项 $\nabla\cdot(\rho_W\nabla W*[\cdot])$，主要是将 [21, 第 7.1.3 节定理 5] 改编到我们的情形。随后，标准的 Galerkin 逼近论证可得出所需函数空间中的存在性和正则性，详见 [21] 。下文将 (55) 的一个光滑解记为 u。

将 (55) 乘以 u 并分部积分，得到

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u\|_{L^2}^2+\|\nabla u\|_{L^2}^2=-\langle u\nabla W*\rho_W,\nabla u\rangle_{L^2}-\langle\rho_W\nabla W*u,\nabla u\rangle_{L^2}+\langle f,u\rangle_{L^2}.$$

利用柯西-施瓦茨不等式以及引理 A.1 的证明中方程 (56)（取 $k=0$ 和 $l=0$），可得

$$-\langle u\nabla W*\rho_W,\nabla u\rangle_{L^2}-\langle\rho_W\nabla W*u,\nabla u\rangle_{L^2}\lesssim_d\|\nabla W\|_{L^\infty}\|\rho_W\|_{L^2}\|u\|_{L^2}\|\nabla u\|_{L^2}.$$

采用与 [33, 定理 5], 相同的论证可得 $\sup_{0\le t\le T}\|\rho_W(t)\|_{L^2}\le C(T,\ M,\ \|\phi\|_{L^2})$。结合这两个界，并利用杨氏乘积不等式吸收一阶项，得到

$$\frac{\mathrm{d}}{\mathrm{d}t}\|u\|_{L^2}^2+\|\nabla u\|_{L^2}^2\le C\left(\|u\|_{L^2}^2+\|f\|_{L^2}^2\right)$$

其中 $C=C(T,\ M,\ \|\phi\|_{L^2})<\infty$。由此，利用格朗沃尔引理并从 0 积分至 T，即可直接得出

$$\sup_{0\le t\le T}\|u(t)\|_{L^2}^2+\|u\|_{L_T^2H^1}^2\le C\|f\|_{L_T^2L^2}^2 \tag{60}$$

其中 $C=C(T,\ M,\ \|\phi\|_{L^2})<\infty$。接下来，将 (55) 乘以 $\partial_t u=:u'$ 并再次分部积分:

$$\|u'\|_{L^2}^2+\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\nabla u\|_{L^2}^2=\left\langle\nabla\cdot(\rho_W\nabla W*u),u'\right\rangle_{L^2}+\left\langle\nabla\cdot(u\nabla W*\rho_W),u'\right\rangle_{L^2}+\left\langle f,u'\right\rangle_{L^2}. \tag{61}$$

同样，柯西- 柯西-施瓦茨不等式和引理 A.1 的方程 (56)，结合 $k=1$ 和 $l=0$，得到 u s

$$\left\langle\nabla\cdot(\rho_W\nabla W*u),u'\right\rangle_{L^2}+\left\langle\nabla\cdot(u\nabla W*\rho_W),u'\right\rangle_{L^2}\lesssim_d\|\nabla W\|_{L^\infty}\|\rho_W\|_{H^1}\|u\|_{H^1}\left\|u'\right\|_{L^2}.$$

根据 [33, 定理 5],，我们有 $\sup_{0\le t\le T}\|\rho_W(t)\|_{H^1}\le C(d,\ T,\ M,\ \|\phi\|_{H^1})$；将 u' 项移至左边并应用杨氏乘积不等式整理后，得到

$$\|u'\|_{L^2}^2+\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\nabla u\|_{L^2}^2\lesssim\|u\|_{H^1}^2+\|f\|_{L^2}^2.$$

对时间积分，并使用第一个估计 (60)、 $u(0)=0$ 以及 Galerkin 逼近的（时间和空间）正则性，我们得到第二个估计

$$\|u'\|_{L_T^2L^2}^2+\sup_{0\le t\le T}\|u(t)\|_{H^1}^2\le C\|f\|_{L_T^2L^2}^2 \tag{62}$$

其中 $C=C(d,\ T,\ M,\ \|\phi\|_{H^1})<\infty$。

$L_T^2H^2$ 中的正则性现在可由前面的估计直接推出。将 (55) 乘以 Δu 并积分

$$\|\Delta u\|_{L^2}^2=\left\langle u',\Delta u\right\rangle_{L^2}-\langle\nabla\cdot(u\nabla W*\rho_W),\Delta u\rangle_{L^2}-\left\langle\nabla\cdot(\rho_W\nabla W*u),\Delta u\right\rangle_{L^2}-\langle f,\Delta u\rangle_{L^2}.$$

如前所述，利用引理A.1 的方程 (56)，以及柯西-施瓦茨不等式和杨氏乘积不等式，我们可得

$$\|\Delta u\|_{L^2}^2\lesssim\|u'\|_{L^2}^2+\|u\|_{H^1}^2+\|f\|_{L^2}^2.$$

现在回顾一下 (60) 和 (62) 中关于 u和 u' 的界，并直接对时间积分，得到

$$\|\Delta u\|_{L_T^2L^2}^2\le C\|f\|_{L_T^2L^2}^2$$

其中 $C=C(d,\ T,\ M,\ \|\phi\|_{H^{d+1}})$。由于我们知道，对于足够光滑的 u， $\|\Delta u\|_{L^2}=\left\|D^2u\right\|_{L^2}$（例如利用普朗歇雷尔恒等式），因此可以将上述估计与 (60) 结合，得到 $L_T^2H^2$ 中的正则性，从而完成证明。

□

定理 A.3 (更高阶空间正则性)。 假设定理 *A.2* 的条件成立，并考虑该定理给出的 (55) 的强解 u。那么，若进一步有 $f \in L^2([0,T];\ H^k)$，其中某个$0 \le k \le \beta - 1$,

$$u \in L^2([0,T]; H^{k+2}) \cap H^1([0,T]; H^k).$$

此外，存在一个常数 $C = C(d, k, T, M, \|\phi\|_{H^{k+1}})$ 使得

$$\|u\|_{L^2([0,T];H^{k+2})} + \|u\|_{L^\infty([0,T];H^{k+1})} + \left\|u'\right\|_{L^2([0,T];H^k)} \le C \|f\|_{L^2([0,T];H^k)}$$

其中 $u' = \mathrm{d}u/\mathrm{d}t$。

证明。 只需将定理 A.2 的证明中的 L^2 内积替换为 H^k 内积，即可重复该证明。对于 $k \le \beta - 1$, [33, , 由于定理 5] 保证

$$\sup_{0\le t\le T} \|\rho_W\|_{H^{k+1}} \le C(d, k, T, M, \|\phi\|_{H^{k+1}}).$$

□

定理 A.4 (较弱范数下的估计)。 假设定理 *A.2* 的条件成立，并考虑该定理给出的 (55) 的强解 u。那么，对于任意 $1 \le s < \beta + 1 - \frac{d}{2}$,

$$\|u\|_{L^2([0,T];H^{-s+2})} + \|u\|_{L^\infty([0,T];H^{-s+1})} + \left\|u'\right\|_{L^2([0,T];H^{-s})} \le C \|f\|_{L^2([0,T];H^{-s})}.$$

其中 $C = C(s, d, T, M, \|\phi\|_{H^k}),\quad k = \min(\beta, \lfloor s - 1 + d/2 \rfloor + 1)$。

证明。 重复定理 A.2 证明中的估计，这次在 H^{-s}中取内积。下面概述主要步骤。由柯西-施瓦茨不等式,

$$-\langle \rho_W \nabla W * u, \nabla u\rangle_{H^{-s}} - \langle u \nabla W * \rho_W, \nabla u\rangle_{H^{-s}} \le \|\nabla u\|_{H^{-s}} \left(\|\rho_W \nabla W * u\|_{H^{-s}} + \|u \nabla W * \rho_W\|_{H^{-s}}\right).$$

为了界定括号内的各项，我们使用对偶性 $\|u\|_{H^{-s}} = \sup_{\|v\|_{H^s}\le 1} \langle u, v\rangle_{L^{2\circ}}$ 任取$v \in H^s$，并注意到，利用引理 A.1 方程 (56) 的证明，其中 $k = s$ 和 $l = 1$,

$$\langle u \nabla W * \rho_W, v\rangle_{L^2} = \langle u, v \nabla W * \rho_W\rangle_{L^2} \le \|u\|_{H^{-s}} \|v \nabla W * \rho_W\|_{H^s} \le \|u\|_{H^{-s}} \|W\|_{W^{2,\infty}} \|\rho_W\|_{H^{s-1}} \|v\|_{H^s}.$$

类似地，令 $\check{W}(x) := W(-x)$，并利用富比尼定理,

$$\langle \rho_W \nabla W * u, v\rangle_{L^2} = \left\langle u, \nabla \check{W} * (\rho_W v)\right\rangle_{L^2} \le \|u\|_{H^{-s}} \left\|\nabla \check{W} * (\rho_W v)\right\|_{H^s} \le \|u\|_{H^{-s}} \|W\|_{W^{2,\infty}} \|\rho_W\|_{C^{s-1}} \|v\|_{H^s}.$$

由于根据假设， $s - 1 + d/2 < \beta$，索伯列夫嵌入结合 [33, 定理 5] 可推出

$$\sup_{0\le t\le T} \|\rho_W\|_{C^{s-1}} \le C(s, d, T, M, \|\phi\|_{H^k}) < \infty,$$

其中 $k = \min(\beta,\ \lfloor s - 1 + d/2 \rfloor + 1)$。因此，对所有满足 $\|v\|_{H^s} \le 1$的 $v \in H^s$ 取上确界，可得

$$\|\rho_W \nabla W * u\|_{H^{-s}} + \|u \nabla W * \rho_W\|_{H^{-s}} \le C \|u\|_{H^{-s}},$$

其中 $C = C(d,\ s,\ T,\ M,\ \|\phi\|_{H^\beta})$。在定理 A.2 的证明中对其余步骤重复这些估计，即可得到所需结果。

□

定理 A.5 (改进的时间正则性)。 假设定理 *A.2* 的假设成立，并考虑该定理给出的方程 (55) 的强解 u。若进一步有 $f' \in L^2([0,T];\ L^2)$，则,

$$u' \in L^2([0,T]; H^1) \cap H^1([0,T]; H^{-1}).$$

此外，存在一个常数 $C = C(d, T, M, \|\phi\|_{H^1})$ 使得

$$\left\|\frac{\mathrm{d}u}{\mathrm{d}t}\right\|_{L^2([0,T];H^1)} + \left\|\frac{\mathrm{d}u}{\mathrm{d}t}\right\|_{L^\infty([0,T];L^2)} + \left\|\frac{\mathrm{d}^2u}{\mathrm{d}t^2}\right\|_{L^2([0,T];H^{-1})} \le C \|f\|_{H^1([0,T];L^2)}$$

证明。我们基本上重复 [21, 第 7.1.3 节定理 5], 的证明中的步骤 4 至 7，使用引理 A.1 方程 (56) 对非局部项进行界定，正如在定理 A.2 中所做的那样。

在整个证明过程中，我们用 f' 表示 f 的时间导数。我们给出先验估计，并参见 [21, Sec. 7.1.3 Theorem 5] 的证明，了解完整的伽辽金论证。首先，对方程关于时间求导，并与某个 u' 取 L^2 内积，随后分部积分，得到

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left\|u'\right\|_{L^2}+\left\|\nabla u'\right\|_{L^2}^2=&\left\langle f',u'\right\rangle\\&-\left\langle\rho_W'\nabla W*u,\nabla u'\right\rangle-\left\langle\rho_W\nabla W*u',\nabla u'\right\rangle\\&-\left\langle u'\nabla W*\rho_W,\nabla u'\right\rangle-\left\langle u\nabla W*\rho_W',\nabla u'\right\rangle.\end{aligned}$$

由 [33, Theorem 5], 可知，$\sup_{0\le t\le T}\|\rho_W'(t)\|_{L^2}\le C(d,\ T,\ M,\ \|\phi\|_{H^1})$，且 $\sup_{0\le t\le T}\|\rho_W(t)\|_{L^2}\le C(M,\ \|\phi\|_{L^2},\ T)$。结合这些估计与引理 A.1 的方程 (56)，取 $k=0$ 和 $l=0$，可得

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left\|u'\right\|_{L^2}+\left\|\nabla u'\right\|_{L^2}^2\lesssim\left\|f'\right\|_{L^2}\left\|u'\right\|_{L^2}+\|u\|_{L^2}\left\|\nabla u'\right\|_{L^2}+\left\|u'\right\|_{L^2}\left\|\nabla u'\right\|_{L^2}.$$

回顾定理 A.2 可知，$\sup_{0\le t\le T}\|u\|_{L^2}\le C(M,\ \|\phi\|_{H^1},\ d,\ T)\ \|f\|_{L^2_TL^2}$，并利用杨氏乘积不等式，关于时间积分后得到

$$\sup_{0\le t\le T}\left\|u'\right\|_{L^2}^2+\int_0^T\left\|u'(t)\right\|_{H^1}^2\mathrm{d}t\lesssim\|f\|_{H^1_TL^2}^2+\left\|u'(0)\right\|_{L^2}^2.$$

现在，利用偏微分方程 (55)，我们得到 $u'(0)=f(0)$，后者由索伯列夫嵌入保证良定义（例如参见 [21, Sec. 5.9.2. Theorem 2]）。这意味着 f 是关于时间连续的、取值于 L^2 的函数，且 $\sup_{0\le t\le T}\|f(t)\|_{L^2}\lesssim_T\|f\|_{H^1_TL^2}$。最终得到

$$\sup_{0\le t\le T}\left\|u'(t)\right\|_{L^2}^2+\int_0^T\left\|u'(t)\right\|_{H^1}^2\mathrm{d}t\le C\|f\|_{H^1_TL^2}^2$$

其中 $C=C(d,\ T,\ M,\ \|\phi\|_{H^1})$。

现在，我们回到偏微分方程，对时间求导，并与某个 $v\in H^1$ 取 L^2 内积:

$$\begin{aligned}\left\langle u'',v\right\rangle=&\left\langle f',v\right\rangle-\left\langle\nabla u',\nabla v\right\rangle\\&-\left\langle\rho_W'\nabla W*u,\nabla v\right\rangle-\left\langle\rho_W\nabla W*u',\nabla v\right\rangle-\left\langle u'\nabla W*\rho_W,\nabla v\right\rangle-\left\langle u\nabla W*\rho_W',\nabla v\right\rangle.\end{aligned}$$

再次利用 [33, Theorem 5] 以及引理A.1、方程 (56) 中的估计和柯西-施瓦茨不等式，我们得到

$$\left\langle u'',\frac{v}{\|v\|_{H^1}}\right\rangle\lesssim\left\|u'\right\|_{H^1}+\|u\|_{L^2}+\left\|f'\right\|_{L^2}.$$

因此，根据 H^{-1} 作为对偶空间的定义，

$$\left\|u''\right\|_{H^{-1}}\lesssim\left\|u'\right\|_{H^1}+\|u\|_{L^2}+\left\|f'\right\|_{L^2}.$$

对时间进行积分并使用之前的估计，即可得到所需结果。 □

B 巴拿赫空间中的微积分

本节汇总一些巴拿赫空间中的微积分结果。

定理 B.1. 设 E、F 和 G 为巴拿赫空间，$U \subset E \times F$ 为 $E \times F$ 的一个开子集。设 $f\colon U \subset E \times F \to G$ 在 $k \geq 1$ 上按 *Fr´echet* 意义为 C^k。设 $V \subset E$ 为 E 的一个开子集，并假设存在一个函数 $u\colon V \to F$，使得对所有 $x \in V$，

1. $(x, u(x)) \in U$, *2.* $f(x, u(x)) = 0$, *3.* $D_2 f(x, u(x))\colon F \to G$ 是一个线性同胚。

此外，假设 u 是连续的，或者对于所有 $x \in V$，方程 $f(x, v) = 0$ 的唯一解是 $(x, v) \in U\,(x, u(x))$。那么，u 在 *Fr´echet* 意义下是 C^k 的，其导数 $Du(x) \in \mathcal{L}(E, F)$ 由下式给出

$$Du(x)[h] = -(D_2 f(x, u(x)))^{-1} D_1 f(x, u(x))[h] \tag{63}$$

对于任意 $x \in V$ 和 $h \in E$。

对于 $k \geq 2$，二阶导数 $D^2 u(x) \in \mathcal{L}(E \times E, F)$ 由下式给出

$$\begin{aligned} D^2 u(x)[h_1, h_2] = -(D_2 f(x, u(x)))^{-1} \Big(& D^2_{1,1} f(x, u(x))[h_1, h_2] \\ & + D^2_{1,2} f(x, u(x))[h_1, Du(x)[h_2]] \\ & + D^2_{2,1} f(x, u(x)) \cdot (Du(x)[h_1], h_2) \\ & + D^2_{2,2} f(x, u(x)) \cdot (Du(x)[h_1], Du(x)[h_2]) \Big) \end{aligned} \tag{64}$$

对于任意 $x \in V$ 和 $h_1, h_2 \in E$。

对于更大的 $k \geq 2$，一般公式为

$$D^k u(x)[h_1, \ldots, h_k] = -(D_2 f(x, u(x)))^{-1} \sum_{\substack{\pi \in \Pi_k \setminus \{\pi_0\} \\ \pi =: \{P_1, \ldots, P_m\}}} D^{|\pi|} f(x, u(x))[H_{P_1}, \ldots, H_{P_m}] \tag{65}$$

其中 Π_k 是 $\{1, \ldots, k\}$ 的所有划分的集合，$\pi_0 := \{\{1, \ldots, k\}\}$ 并且

$$H_{P_j} := \begin{cases} (h_i, Du(x)[h_i]) & \text{if } |P_j| = 1 \text{ and } P_j = \{i\}, \\ \left(0, D^{|P_j|} u(x)[(h_i)_{i \in P_j}]\right) & \text{if } |P_j| \geq 2. \end{cases}$$

定理 *B.1* 的证明。首先，我们证明 u 是 C^1。为此，任取一个 $x_0 \in V$。我们先应用巴拿赫空间隐函数定理，参见例如 [17, 定理（10.2.1）]。由此可知，存在一个 x_0 的连通开邻域 U_{x_0}，且 U_{x_0} 包含于 V 中；还存在唯一的连续函数 $\tilde{u}\colon U_{x_0} \to F$，使得对所有 $x \in U_{x_0}$，$(x, \tilde{u}(x)) \in U$ 且 $f(x, \tilde{u}(x)) = 0$。此外，$\tilde{u}$ 在 U_{x_0} 中是 C^1 的。

现在，若 u 连续，则根据隐函数定理的唯一性部分，$u = \tilde{u}$ 属于 U_{x_0}。这意味着 u 在 U_{x_0} 中是 C^1 的，特别是 u 在 x_0 处可微。由于 x_0 是任意的，我们得到 u 在 V 中可微。接着，由于 u 在 U_{x_0} 中是 C^1 的，特别是其导数在 x_0 处连续。同样，x_0 是任意的，因此 Du 在 V 中连续。反之，若我们假设对每个 $x \in V$，$u(x)$ 是 $f(x, u(x)) = 0$ 的唯一解，那么 $u = \tilde{u}$ 也必定属于 U_{x_0}，并且根据前述推理，u 再次是 C^1 的。

现在我们知道 u 是 C^1，第一个公式（63）可直接通过对方程 $f(x, u(x)) = 0$ 关于 x 求导并应用链式法则得到。然后，我们可以利用（63）通过简单的自举论证得出 u 是 C^k。事实上，假设对某个 $l < k$，u 是 C^l。由于 f 是 C^k，链式法则给出 $(D_2 f(x, u(x)))^{-1} D_1 f(x, u(x))$ 是 C^l。但这意味着 Du 本身是 C^l，即 u 是 C^{l+1}。由归纳法即可证明这一结论。一旦我们知道 u 是 C^k，对 $D^k(f \circ g)(x) = 0$ 应用 Fa`a di Bruno 多元公式并将涉及分拆 π_0 的项单独提取出来，即可得到公式（65）（因此也得到公式（64）；这里令 $g(x) = (x, u(x))$。 □

引理 B.2。 设 $f\colon\ E_1 \times \cdots \times E_m \to G$ 是一个连续多重线性映射，即 f 是 m-线性且连续的。那么， f 在 *Fr´echet* 意义下是 C^∞ 。

*Pro*引理 *B.2* 的证明。 参见 [17, 第 VIII 章], 中的 (8.1.4)，并通过归纳法进行。 □

B.1 其余证明

命题 *2.5* 的证明。 我们首先证明 f 是 C^∞。连续多重线性映射在 Fr´échet 意义下是 C^∞ 的，参见引理 B.2。考虑辅助映射 $g\colon\ F \times E \times F \to G$ 以及 $h\colon\ F \to G$ ，定义为

$$h(\rho) \coloneqq (\partial_t \rho - \Delta\rho, \rho(0)), \quad g(\rho, W, u) \coloneqq (-\nabla \cdot (\rho \nabla W * u), 0) = (-\mathcal{T}(\rho, W, u), 0).$$

根据定义， h 是线性的。我们证明它是连续的，即有界:

$$\|h(\rho)\|_G \le \|(\partial_t - \Delta)\rho\|_{L^2_T H^{\beta-1}} + \|\rho(0)\|_{H^\beta} \le \|\rho\|_{H^1_T H^{\beta-1}} + \|\rho\|_{L^2_T H^{\beta+1}} + C\big(\|\rho\|_{H^1_T H^{\beta-1}} + \|\rho\|_{L^2_T H^{\beta+1}}\big)$$

其中使用了 (10)。这证明了 $\|h(\rho)\|_G \lesssim \|\rho\|_F$ ，因此 h 是 C^∞。然后，根据定义， g 是三重线性的，并由引理 2.4 知其连续，因此也是 C^∞。最后， f 分解为

$$f(W, \rho) = h(\rho) + g(\rho, W, \rho) + (0, -\phi) \tag{66}$$

这表明 f 作为光滑映射之和是光滑的。

f 的导数表达式也可由分解式（66）得出。根据链式法则，一阶导数为

$$Df(W, \rho)[H, s] = Dh(\rho)[s] + Dg(\rho, W, \rho)[s, H, s] + 0$$

因为映射（W， ρ） $\mapsto \rho$和（W， ρ） $\mapsto$ （ρ， W， ρ）是连续线性映射，因此它们的导数就是其自身。现在，连续多重线性映射的求导法则给出:

$$Df(W, \rho)[H, s] = h(s) + g(s, W, \rho) + g(\rho, H, \rho) + g(\rho, W, s)$$

因此

$$D_1 f(W, \rho)[H] = g(\rho, H, \rho), \quad D_2 f(W, \rho)[s] = h(s) + g(s, W, \rho) + g(\rho, W, s) \tag{67}$$

这恰好就是命题中给出的表达式。

对于二阶导数，我们分别对关于 W 和 ρ的表达式（67）求导。由于 $D_1 f$不依赖于 W ，我们得到$D^2_{1,1} f = 0$，并且与上面类似

$$D^2_{2,1} f(W, \rho)[s, H] = g(\rho, H, s) + g(s, H, \rho) = D^2_{1,2} f(W, \rho)[H, s]$$

此外，（67）中的第二个方程关于 ρ已经是线性的，因此我们得到

$$D^2_{2,2} f(W, \rho)[s_1, s_2] = g(s_1, W, s_2) + g(s_2, W, s_1),$$

如所需。 □

引理 *2.13* 的证明。 回忆定义 $\mathcal{T}(r,\ V,\ s) \coloneqq \nabla \cdot (r \nabla V * s)$，见式（16）。在整个证明中，我们将使用记号 $\|\cdot\|_\gamma \coloneqq \|\cdot\|_{L^2([0,T];\ H^{\gamma+1}) \cap H^1([0,T],\ H^{\gamma-1})}$ 我们采用归纳法。情形$k = 1$ 可由定理 A.3 中的正则性估计结合引理 A.1 的方程（57）（其中 $\nu = 0$ ）以及式（10）直接得到:

$$\|D\rho_W[h_1]\|_\gamma \lesssim \|\mathcal{T}(\rho_W, h_1, \rho_W)\|_{L^2([0,T], H^{\gamma-1})} \lesssim \|h_1\|_{L^2} \|\rho_W\|^2_\gamma .$$

由于 $\|\rho_W\|_\gamma \le C(d,\ \gamma,\ T,\ M,\ \|\phi\|_{H^\gamma})$，这就证明了基例。

现在，令 $k \geq 2$ 并假设引理对所有 $l < k$都成立。公式（65）允许我们用低阶项表示 $D^k \rho_W$。具体而言， $D^k \rho_W[H_1, \ldots, H_k]$ 满足方程（55），其右端强迫项为F，初始条件 ψ 为

$$(F, \psi) \coloneqq - \sum_{\substack{\pi \in \Pi_k \setminus \{\pi_0\} \\ \pi =: \{P_1, \ldots, P_m\}}} D^{|\pi|} f(W, \rho_W)[H_{P_1}, \ldots, H_{P_m}]$$

其中我们回顾一下， Π_k是 $\{1, \ldots, k\}$的所有划分的集合， $\pi_0 \coloneqq \{\{1, \ldots, k\}\}$ 以及

$$H_{P_j} \coloneqq \begin{cases} (h_i, D\rho_W[h_i]) & \text{if } |P_j| = 1 \text{ and } P_j = \{i\}, \\ \left(0, D^{|P_j|}\rho_W \cdot [h_i]_{i \in P_j}\right) & \text{if } |P_j| \geq 2. \end{cases}$$

特别地，我们注意到，由于 $\pi \neq \pi_0$，求和中的各项包含 f 的二阶或更高阶导数。命题 2.5 随即表明，$D^2_{1,\ 2} f$ 关于 W 是常数，关于 ρ是线性的；而$D^2_{2,\ 2} f$ 关于 ρ 是常数，关于 W是线性的。因此，我们得到 $D^4 f = 0$, $D^3_{1,\ 1,\ 1} f = D^3_{1,\ 1,\ 2} f = D^3_{2,\ 2,\ 2} f = 0$，以及

$$D^3_{1,2,2} f(W, \rho)[h, s_1, s_2] = -\left(\mathcal{T}(s_1, h, s_2) + \mathcal{T}(s_2, h, s_1), 0\right).$$

换言之， $\psi = 0$ 和 F 是以下所有项的和

$$\mathcal{T}(D^{|P_2|}\rho_W[h_i]_{i \in P_2}, h_j, D^{|P_3|}\rho_W[h_i]_{i \in P_3}) + \mathcal{T}(D^{|P_3|}\rho_W \cdot [h_i]_{i \in P_3}, h_j, D^{|P_2|}\rho_W \cdot [h_i]_{i \in P_2})$$

对所有大小为 3 的划分以及所有满足 $P_1 = \{j\}$ 的 $j = 1$、.....、 k（来自$D^3_{1,\ 2,\ 2} f$的项），

$$\mathcal{T}\left(D^{k-1}\rho_W[h_i]_{i \neq j}, h_j, \rho_W\right) + \mathcal{T}\left(\rho_W, h_j, D^{k-1}\rho_W[h_i]_{i \neq j}\right)$$

对所有 $j = 1$、.....、 k（来自$D^2_{1,\ 2} f$的项），以及，

$$\mathcal{T}(D^{|P_1|}\rho_W[h_i]_{i \in P_1}, W, D^{|P_2|}\rho_W[h_i]_{i \in P_2}) + \mathcal{T}(D^{|P_2|}\rho_W[h_i]_{i \in P_2}, W, D^{|P_1|}\rho_W[h_i]_{i \in P_1})$$

对于所有大小为 2 的划分（来自 $D^2_{2,\ 2} f$ 的项）。

现在我们可以使用定理 A.3 得到

$$\left\| D^k \rho_W[h_1, \ldots, h_k] \right\|_{L^2([0,T]; H^{\gamma+1}) \cap H^1([0,T], H^{\gamma-1})} \leq C \left\| F \right\|_{L^2([0,T], H^{\gamma-1})}$$

其中 $C = C(d,\ \gamma,\ T,\ M,\ \|\phi\|_{H^\gamma})$。为完成归纳步骤并得出证明结论，我们使用三角不等式将 F 拆分为它的所有项。结合 (10)，我们使用引理 A.1 的方程（57）和 $\nu = 0$ 来控制每一项。

来自 $D^3_{1,\ 2,\ 2} f$ 的项：

$$\begin{aligned} &\left\| \mathcal{T}\left(D^{|P_2|}\rho_W[h_i]_{i \in P_2},\, h_j,\, D^{|P_3|}\rho_W[h_i]_{i \in P_3}\right) \right\|_{L^2([0,T], H^{\gamma-1})} \\ &\qquad \lesssim \|h_j\|_{L^2} \sup_{0 \leq t \leq T} \left\| D^{|P_2|}\rho_W[h_i]_{i \in P_2} \right\|_{H^\gamma} \left\| D^{|P_3|}\rho_W[h_i]_{i \in P_3} \right\|_{L^2([0,T]; H^{\gamma+1})} \\ &\qquad \lesssim \|h_j\|_{L^2} \left\| D^{|P_2|}\rho_W[h_i]_{i \in P_2} \right\|_{\gamma} \left\| D^{|P_3|}\rho_W[h_i]_{i \in P_3} \right\|_{\gamma} \\ &\qquad \leq C \|h_1\|_{L^2} \ldots \|h_k\|_{L^2} \end{aligned}$$

其中最后一个不等式用到了归纳假设。其余各项可用类似方法界定。

来自 $D^2_{1,\ 2} f$ 的项：

$$\begin{aligned} \left\| \mathcal{T}\left(D^{k-1}\rho_W[h_i]_{i \neq j}, h_j, \rho_W\right) \right\|_{L^2([0,T], H^{\gamma-1})} &\lesssim \|h_j\|_{L^2} \sup_{0 \leq t \leq T} \left\| D^{k-1}\rho_W[h_i]_{i \neq j} \right\|_{H^\gamma} \|\rho_W\|_{L^2([0,T]; H^{\gamma+1})} \\ &\lesssim \|h_j\|_{L^2} \left\| D^{k-1}\rho_W[h_i]_{i \neq j} \right\|_{\gamma} \leq C \|h_1\|_{L^2} \ldots \|h_k\|_{L^2} \end{aligned}$$

其中，第二个不等式用到了 [33, 定理 5]，最后一个不等式用到了归纳假设。其他项也可以用类似的方法进行界定。

来自$D^2_{2、2}f$的项：

$$\begin{aligned}
&\left\|\mathcal{T}(D^{|P_1|}\rho_W[h_i]_{i\in P_1}, W, D^{|P_2|}\rho_W[h_i]_{i\in P_2})\right\|_{L^2([0,T],H^{\gamma-1})} \\
&\qquad \lesssim \|W\|_{L^2} \sup_{0\le t\le T} \left\|D^{|P_1|}\rho_W[h_i]_{i\in P_1}\right\|_{H^\gamma} \left\|D^{|P_2|}\rho_W[h_i]_{i\in P_2}\right\|_{L^2([0,T];H^{\gamma+1})} \\
&\qquad \lesssim \|W\|_{W^{2,\infty}} \left\|D^{|P_1|}\rho_W[h_i]_{i\in P_1}\right\|_{\gamma} \left\|D^{|P_2|}\rho_W[h_i]_{i\in P_2}\right\|_{\gamma} \\
&\qquad \le C\,\|h_1\|_{L^2}\dots\|h_k\|_{L^2}
\end{aligned}$$

其中，我们利用定义域的有界性写出$\|W\|_{L^2} \le \|W\|_{W^{2、\infty}}$，并在最后一个不等式中使用了归纳假设。其他各项可类似地得到有界性。

□

致谢。作者感谢 ERC 高级研究员资助（UKRI G116786）以及英国工程与自然科学研究委员会项目资助 EP/V026259 的资助。RN 感谢 Jan Bohr 指出了参考文献 [1],，两位作者也感谢 Greg Pavliotis 的多次讨论。

参考文献

[1] Serge Alinhac 和 Patrick G´erard。《伪微分算子与纳什-莫泽定理》，数学研究生教材第 82 卷。美国数学会，普罗维登斯，罗得岛州，2007 年。由 Stephen S. Wilson 从 1991 年法文原版翻译。[2] Randolf Altmeyer。高维广义线性模型中基于抽样的后验推断的多项式时间保证。《即将发表于 *Annales de l'Institut Henri Poincar´e，arXiv* 预印本 *arXiv:2208.13296*》，2022 年。[3] Chiara Amorino、Denis Belomestny、Vytaut¨e Pilipauskait¨e、Mark Podolskij 和 Shi-Yuan Zhou。麦基恩-弗拉索夫随机微分方程的非参数估计的反卷积多项式速率。《概率论与相关领域》，2024 年。[4] Dominique Bakry、Ivan Gentil 和 Michel Ledoux。《马尔可夫扩散算子的分析学与几何》，第 348 卷。施普林格科学与商业媒体，2013 年。[5] Afonso S. Bandeira、Antoine Maillard、Richard Nickl 和 Sven Wang。高斯先验中的自由能垒以及高维单峰分布冷启动马尔可夫链蒙特卡罗失效。《哲学学会学报：皇家学会 *A* 辑》，381(2247)：论文编号 20220150，29 页，2023 年。[6] Julian Besag。关于 U. Grenander 和 M. I. Miller 的"复杂系统中的知识表征"的评论。《皇家统计学会杂志 *B* 辑》，56：591–592，1994 年。[7] Jan Bohr 和 Richard Nickl。高维后验测度的对数凹近似以及非线性反问题中的稳定性。《*Annales de l'Institut Henri Poincar´e，Probabilit´es et Statistiques*》，60(4)：2619–2667，2024 年。[8] S. Boyd 和 L. Vandenberge。《凸优化》。剑桥大学出版社，2004 年。

[9] J. A. Carrillo, R. S. Gvalani, G. A. Pavliotis 和 A. Schlichting. 环面上麦基恩–弗拉索夫方程的长期行为与相变。《数学分析与力学文献》，2020年。

[10] Bernard Chazelle、Quansen Jiu、Qianxiao Li 和 Chu Wang. 意见动力学中含噪共识模型极限方程的适定性。微分方程杂志，269(11):9942– 9977，2020年。[11] Sinho Chewi、Murat A. Erdogdu、Mufan Li、Ruoqi Shen 和 Matthew S. Zhang. 从庞加莱不等式到对数索博列夫不等式的朗之万蒙特卡罗分析。Found. Comput. Math.，25(4):1345–1395，2025年。[12] Sinho Chewi、Chen Lu、Kwangjun Ahn、Xiang Cheng、Thibaut Le Gouic 和 Philippe Rigollet. 梅特罗波利斯调整的朗之万算法的最优维数依赖性。载于学习理论会议。PMLR，2021年。[13] S. L. Cotter、G. O. Roberts、A. M. Stuart 和 D. White. 函数的马尔可夫链蒙特卡洛方法：改进旧算法以提升速度。统计科学，28(3):424–446，2013年。[14] Dinh D˜ung、Van Kien Nguyen、Christoph Schwab 和 Jakob Zech. 高斯随机场输入偏微分方程不确定性量化中的解析性与稀疏性，数学讲义第2334卷。Springer, Cham， [2023] ©2023年。[15] Arnak S Dalalyan. 从光滑对数凹密度进行近似采样的理论保证。英国皇家统计学会杂志B辑：统计方法学，79(3):651–676，2017年。[16] Laetitia Della Maestra 和 Marc Hoffmann. 相互作用粒子系统的非参数估计：麦基恩–弗拉索夫模型。概率论及相关领域，182(1):551–613，2022年。[17] J. Dieudonn´e. 现代分析基础。纯粹与应用数学。学术出版社，1969年。[18] Simon Duane、Anthony D. Kennedy、Brian J. Pendleton 和 Duncan Roweth. 混合蒙特卡罗。物理快报B辑，195(2):216–222，1987年。[19] A. Durmus 和 ´E. Moulines. 未调整朗之万算法的非渐近收敛分析。应用概率年刊，27(3):1551–1587，2017年。[20] A. Durmus 和 ´E. Moulines. 通过未调整朗之万算法进行高维贝叶斯推断。《伯努利》，25(4A):2854–2882，2019年。[21] L. C. Evans. 偏微分方程。数学研究生教材。美国数学学会，2010年。[22] E. Gin´e 和 R. Nickl. 无限维统计模型的数学基础。剑桥大学出版社，2016年。[23] Matteo Giordano 和 Sven Wang. 低频扩散数据的统计算法：偏微分方程方法。《统计年刊》，53(3):1150–1175，2025年。[24] R. S. Gvalani. 多粒子系统：相变、特征极限与梯度流。博士论文，伦敦帝国理工学院，2020年。[25] D. Konen 和 R. Nickl. 基于二维纳维–斯托克斯方程的数据同化：后验测度的最优高斯渐近性。arXiv 预印本，arXiv:2507.18279，2025年。[26] J.-L. Lions 和 E. Magenes. 非齐次边值问题及其应用，第2卷，数学科学基础丛书第182卷。施普林格出版社，柏林、海德堡，1972年。

[27] W. Liu 和 M. R¨ockner。《随机偏微分方程：导论》。Universitext。Springer，Cham，2015年。[28] Francois Monard、Richard Nickl 和 Gabriel P. Paternain。噪声非阿贝尔 X 射线变换的一致反演。《纯粹与应用数学通讯》，74(5):1045–1099，2021年。[29] Fran¸cois Monard、Richard Nickl 和 Gabriel P. Paternain。高斯过程先验下非线性反问题贝叶斯不确定性量化的统计保证。《统计年鉴》，49(6):3255–3284，2021年。[30] R. Nickl。统计反问题的伯恩斯坦–冯·米泽斯定理 I：薛定谔方程。《欧洲数学会杂志》，22(8):2697–2750，2020年。[31] R. Nickl。《贝叶斯非线性统计反问题》。欧洲数学会出版社，2023年。[32]R. Nickl。时间演化方程的伯恩斯坦–冯·米泽斯定理。《arXiv预印本》，arXiv:2407.14781，2024年。[33] R. Nickl、G. A. Pavliotis 和 K. Ray。麦基恩–弗拉索夫模型中的贝叶斯非参数推断。《统计年刊》，2025年。[34] R. Nickl 和 E. S. Titi。关于高斯过程先验下数据同化的后验一致性：二维纳维–斯托克斯方程。《统计年刊》，52(4):1825–1844，2024年。[35] R. Nickl 和 S. Wang。通过朗之万型算法对高维后验测度进行多项式时间计算。《欧洲数学会杂志》，26(3):1031–1112，2024年。[36] Gabriel P. Paternain、Mikko Salo 和 Gunther Uhlmann。《几何反问题——侧重二维》，剑桥高等数学研究丛书第204卷。剑桥大学出版社，剑桥，2023年。附有 Andr´as Vasy 撰写的序言。[37] Grigorios A. Pavliotis 和 Andrea Zanoni。遍历麦基恩随机微分方程的线性化及其应用。《非线性》，38(8):085008，2025年。[38] Sebastian Reich 和 Colin Cotter。《概率预测与贝叶斯数据同化》。剑桥大学出版社，纽约，2015年。[39] G. O. Roberts 和 J. S. Rosenthal。朗之万扩散离散近似的最优尺度调整。《皇家统计学会学报：B辑（统计方法论）》，60(1):255–268，1998年。[40] Gareth O. Roberts 和 Richard L. Tweedie。朗之万分布及其离散近似的指数收敛。《伯努利》，2(4):341–363，1996年。[41] Elias M. Stein。《奇异积分与函数的可微性》，普林斯顿数学丛书第30卷。普林斯顿大学出版社，新泽西州普林斯顿，1970年。[42] A. M. Stuart。反问题：贝叶斯视角。《数值数学学报》，19:451–559，2010年。

Department of Pure Mathematics & Mathematical StatisticsUniversity of Cambridge, 英国剑桥

电子邮箱：nickl@maths.cam.ac.uk, afc51@cam.ac.uk