arXiv:2601.02725v1 [math.PR] 6 Jan 2026

通过无 Hessian 高分辨率动力学对非对数凹密度进行采样

Xiaoyu Wang* Yingli Wang† Lingjiong Zhu‡2026年1月7日

摘要

我们研究通过无 Hessian 高分辨率（HFHR）动力学从定义在 $\mathbb{R}^d$ 上的目标分布 $\pi(q) \propto e^{-U(q)}$ 中进行采样的问题。其中，U 可以是非凸的。HFHR 动力学是一种二阶朗之万型过程，其唯一不变分布为 $e^{-U(q)-\frac{1}{2}|p|^2}$；当分辨率参数 $\alpha \to 0$ 时，该过程退化为动能朗之万动力学（KLD）。现有文献中关于 HFHR 动力学的理论仅限于强凸 U，尽管数值实验也表明其在非凸情形下颇具前景。我们重点研究 U 可以是非凸的情况下 HFHR 动力学的收敛性，以弥合理论与实践之间的差距。在对 U 作出标准的耗散性和光滑性假设后，我们采用反射/同步耦合方法。这使得在一种 Lyapunov 加权的 Wasserstein 距离下，只要 KLD 具有指数收缩性，对于所有充分小的 $\alpha > 0$，HFHR 半群也具有指数收缩性。我们进一步证明，在附加假设 ∇U 在无穷远处渐近线性增长的条件下，HFHR 动力学的收缩率严格优于 KLD，并且可以明确量化其增益。作为案例研究，我们验证了三个例子的假设及其带来的加速效果：多阱势、含 L^p 正则化项的贝叶斯线性回归，以及贝叶斯二分类。我们基于这些例子开展了数值实验，并额外使用经神经网络处理的真实数据，对贝叶斯逻辑回归进行实验。结果展示了基于 HFHR 动力学的算法的效率，并验证了其加速效果及相较于 KLD 的优越性能。

1 引言

我们考虑从目标分布中进行采样的问题

$$\pi(q) \propto e^{-U(q)}, \qquad q \in \mathbb{R}^d,$$

其中 $U: \mathbb{R}^d \to \mathbb{R}$ 是一个势函数。此类采样问题经常出现在贝叶斯统计、逆问题和现代机器学习中，例如高维模型的后验采样以及大规模优化的贝叶斯表述 [GCSR95、Stu10、ADFDJ03、TTV16、GGHZ21、GIWZ24、BGK+25]。

*金融科技重点领域，香港科技大学（广州），广州，广东，中华人民共和国；xiaoyuwang@hkust-gz.edu.cn

†通讯作者。上海财经大学数学学院，上海，中华人民共和国；naturesky1994@gmail.com

‡佛罗里达州立大学数学系，塔拉哈西，佛罗里达州，美利坚合众国；zhu@math.fsu.edu

一种经典方法基于过阻尼朗之万动力学 （OLD），

$$dq_t = -\nabla U(q_t)\,dt + \sqrt{2}\,dB_t, \tag{1.1}$$

其不变分布（在温和条件下）的密度为 $\pi(q) \propto e^{-U(q)}$；例如参见 [CHS87，HKS89]。实践中，可以通过欧拉–丸山格式模拟（1.1）

$$q_{k+1} = q_k - \eta\nabla U(q_k) + \sqrt{2\eta}\,\xi_{k+1}, \tag{1.2}$$

通常称为未校正朗之万算法（ULA） [Dal17, DM17, DM19], ，其中 ξ_k 是独立同分布（i.i.d.）的高斯随机向量 $\mathcal{N}(0, I_d)$。过去十年间，针对（1.2）在多种距离（全变差、Wasserstein、Kullback–Leibler、 χ^2、R´enyi）下，以及在随机梯度设置下，已经发展出精确的非渐近理论[Dal17, DM17, DM19, DK19, RRT17, BCM$^+$21, CMR$^+$21, ZADS23, CB18, EHZ22]。

为加速收敛，可以引入一个动量变量，并考虑动能朗之万动力学 （KLD，也称为欠阻尼或二阶朗之万动力学）[MSH02，Vil09，CCBJ18，CCA$^+$18，CLW21，CLW23，DRD20，GGZ20，MCC$^+$21，GGZ22]:

$$\begin{cases} dp_t = -\gamma p_t\,dt - \nabla U(q_t)\,dt + \sqrt{2\gamma}\,dB_t, \\ dq_t = p_t\,dt, \end{cases} \tag{1.3}$$

其中 $(B_t)_{t\geq 0}$ 是一个 d维布朗运动，而 $\gamma > 0$ 是摩擦参数。在温和的假设下，(1.3) 存在唯一的不变测度，其密度为 $\propto e^{-U(q)-\frac{1}{2}|p|^2}$，其$q$边缘分布与 π一致。如今我们已充分理解，无论在连续还是离散层面，动能朗之万动力学及其离散化算法都可以比过阻尼对应方法收敛得更快，并且对维度 d 和精度 ϵ [EGZ19，CLW23，CCBJ18，GGZ22]。

动能朗之万动力学与*Nesterov* 加速梯度（NAG）优化方法 [Nes83、Nes13、MCC$^+$21、GGZ22]密切相关。受 NAG 的高分辨率常微分方程（ODE）视角启发， [LZT22] 提出了无海森矩阵的高分辨率（HFHR）动力学，这是一种具有状态 $(q_t,\ p_t)$ 的 $\in \mathbb{R}^{2d}$维朗之万型动力学:

$$dq_t = (p_t - \alpha\nabla U(q_t))\,dt + \sqrt{2\alpha}\,dB_t^q, \tag{1.4}$$

$$dp_t = (-\gamma p_t - \nabla U(q_t))\,dt + \sqrt{2\gamma}\,dB_t^p, \tag{1.5}$$

其中， B^q、 B^p 是相互独立的 d维布朗运动，而 $\alpha > 0$ 是一个“分辨率”参数。形式上，当 $\alpha \to 0$ 该系统（1.4）–（1.5）退化为（1.3），而对于固定的 $\alpha > 0$ ，它保持密度为 $\propto e^{-U(q)-\frac{1}{2}|p|^2}$ [LZT22]的吉布斯测度。（1.4）–（1.5）中的漂移项仅依赖于 ∇U ，因此是“无海森矩阵”的；这与其他涉及 $\nabla^2 U$的 NAG 高分辨率 ODE 不同；例如参见 [SDJS22]。近期研究进一步利用了优化中 NAG 方法的相关性，设计出用于加速采样的梯度校正动力学，其中 HFHR 动力学是其特例 [ZOL25]。

数值实验 [LZT22] 表明，在非凸采样任务中，HFHR 动力学相较于动能朗之万动力学可以实现显著加速。然而，现有理论基本上仅适用于强凸（对数凹）势函数 [LZT22]。非凸情形要微妙得多：当 U 为非凸时，漂移项的雅可比矩阵存在扩张方向

，而朴素的Lyapunov估计可能无法在全局范围内控制动力学。与此同时，对于动能朗之万动力学（1.3），得益于 [Ebe16、EGZ19], 的工作，非凸景观中已有一种基于耦合的精细理论：他们构造了反射/同步耦合和加权 Wasserstein 距离，使得马尔可夫半群呈指数收缩。

这引出了以下问题：

(Q1) 对于非对数凹目标分布，能否证明 *HFHR* 动力学以指数速度收敛到平衡态，并且只需满足与动能朗之万动力学相同类型的 U 条件？

(Q2) *HFHR* 动力学是否确实能加速混合，也就是说，在某个合适的 *Wasserstein* 距离下，其收缩率是否严格优于动能朗之万动力学，至少在 $\alpha > 0$ 较小时如此？

本文旨在通过统一的耦合框架回答这两个问题。我们将 [Ebe16、EGZ19] 中的反射/同步耦合方法改用于 HFHR 动力学，并结合 [EGZ19], 所采用的 Lyapunov 加权距离，从而得到非渐近全局收缩性。分析精确揭示了式（1.4）–（1.5）中的额外无 Hessian 漂移如何影响 Lyapunov结构和 Wasserstein 收缩率。

我们的贡献概括如下。

(1) HFHR 动力学的 Lyapunov结构与全局收缩性。我们首先证明，在对可能非凸的 U 满足光滑性和耗散性假设（假设 2.1）的条件下，对于所有充分小的 $\alpha > 0$，动能朗之万李雅普诺夫函数 $\mathcal{V}_0$ 仍然是 HFHR 无穷小生成元 $\mathcal{L}_\alpha$ 的李雅普诺夫函数，因此已经可以推出 HFHR 动力学的非渐近指数收敛性，这与在 U 上满足同一组假设时的动能朗之万动力学相同；参见命题 2.2 和推论 3.9。更一般地，给定任意满足漂移条件（3.7）的李雅普诺夫函数 $\mathcal{V}$，我们将[Ebe16、EGZ19] 的反射/同步耦合推广到 HFHR 动力学，并构造一个 Lyapunov 加权半度量 $\rho_{\mathcal{V}}$，将相空间距离的凹函数与权重 $1 + \mathcal{V}(z) + \mathcal{V}(z')$相结合。我们证明，相关的加权 Wasserstein 距离 $\mathcal{W}_{\rho_{\mathcal{V}}}$ 在 HFHR半群下以显式收缩率 $c(\lambda) > 0$指数收缩；参见定理 3.8。

（2）精细李雅普诺夫函数与定量加速。我们为 HFHR 动力学构造了一种新颖的精细李雅普诺夫函数，其形式为 $\mathcal{V}_\alpha = \mathcal{V}_0 + \alpha\mathcal{M}$，其中 $\mathcal{V}_0$ 是动能朗之万李雅普诺夫函数，$\mathcal{M}$ 是无Hessian校正项。在附加假设下，即渐近地，∇U 在无穷远处呈线性增长（假设 4.1），我们证明 $\mathcal{V}_\alpha$ 可得到改进的漂移率 $\lambda_\alpha \geq \lambda + \Theta(\alpha)$（命题 4.7）；其中，$\lambda_\alpha$ 是参数为 α的 HFHR 动力学所对应的生成元/李雅普诺夫不等式中的漂移常数，而 λ表示基线（$\alpha = 0$）常数；这正是假设 2.1(iii)（耗散条件）中的 λ 。这将带来收缩率的严格提升。具体而言，记 c_0 和 c_α 分别为动能朗之万动力学和 HFHR 动力学的收缩率，我们证明（推论 4.13），对于所有充分小的 $\alpha > 0$，存在一个可显式计算的 $\kappa_{\text{global}} > 0$，使得

$$c_\alpha \geq c_0 + \kappa_{\text{global}}\,\alpha.$$

至关重要的是，我们证明了这种加速具有鲁棒性：无论收敛瓶颈由李雅普诺夫分支（从无穷远处的回归）还是度量分支（跨越能垒）决定，这一结论都成立。

）。这意味着，HFHR 动力学在加权 Wasserstein 距离 $\mathcal{W}_{\rho\nu_\alpha}$下的收缩率严格优于动能朗之万动力学 $\mathcal{W}_2$，因此在标准 2-Wasserstein 距离下也同样如此 $\mathcal{W}_2$ （推论4.14）。

(3) 案例研究。作为具体示例，我们研究三个势函数 U 通常为非凸的例子：多阱势（第 5.1 节）、带有 L^p 正则化项的贝叶斯线性回归（第 5.2 节）以及贝叶斯二分类（第 5.3 节）。对于所有这些例子，我们验证了假设 2.1 和 4.1 均成立。因此，第 3 节和第 4 节的所有理论结果均适用，这表明对于所有这些例子，HFHR 动力学的收缩率都严格优于动能朗之万动力学。

(4) 我们通过基于满足所有理论结果假设的这些例子开展数值实验，来说明我们的理论。具体而言，我们使用基于 HFHR 动力学和动能朗之万动力学离散化的算法，对多阱势（第6.1 节）、使用合成数据且带有 L^p 正则化项的贝叶斯线性回归（第 6.2 节），以及使用真实数据的贝叶斯二分类（第 6.3 节）开展实验。实验表明，与动能朗之万动力学相比，基于 HFHR 动力学的算法具有加速效果和更优的性能，从而验证了我们的理论发现。此外，我们还使用经神经网络处理的真实数据，对贝叶斯逻辑回归开展实验；这类数据可能不满足我们理论中的假设，但仍表现出优异的数值性能（第 6.4 节）。

我们强调，式 (2)-(3) 中为获得更高收缩率而采用的额外结构假设，仅适用于加速结果。对于较小的分辨率参数α，HFHR 动力学的基本指数收敛性已可在加权 Wasserstein 距离下得到，其所需的势函数 U 假设与动能朗之万情形相同。

2 预备知识

本节概述我们所研究的随机动力学的精确定义，介绍其无穷小生成元，并陈述推导所有结果所依据的势函数 U 的基本假设。全文中，我们在相空间 $\mathbb{R}^{2d}$ 内进行研究，其坐标为 $z = (q,\ p)$，其中 $q \in \mathbb{R}^d$ 表示位置，$p \in \mathbb{R}^d$ 表示动量。

2.1 HFHR 动力学与无穷小生成元

我们回顾式 (1.4)-(1.5，可知无海森矩阵高分辨率（HFHR）动力学由以下随机微分方程（SDE）定义：

$$\begin{aligned} dq_t &= (p_t - \alpha\nabla U(q_t))\ dt + \sqrt{2\alpha}\, dB_t^q, \\ dp_t &= (-\gamma p_t - \nabla U(q_t))\ dt + \sqrt{2\gamma}\, dB_t^p, \end{aligned} \tag{2.1}$$

其中 B^q 和 B^p 是 $\mathbb{R}^d$中的相互独立标准布朗运动，$\gamma > 0$是摩擦参数，而 $\alpha > 0$ 是分辨率参数。形式上，当 $\alpha \to 0$时，系统退化为动能朗之万动力学（1.3）；而对于固定的 $\alpha > 0$，系统保持密度为 $\propto e^{-U(q)-\frac{1}{2}|p|^2}$的吉布斯测度。（ $\mathcal{L}_\alpha$）的无穷小生成元作用于 C^2测试函数 φ： $\mathbb{R}^{2d} \to \mathbb{R}$其形式为

$$\mathcal{L}_\alpha\varphi(q,p) = (p - \alpha\nabla U(q)) \cdot \nabla_q\varphi + (-\gamma p - \nabla U(q)) \cdot \nabla_p\varphi + \alpha\Delta_q\varphi + \gamma\Delta_p\varphi. \tag{2.2}$$

对于微扰论证，将（2.2）中的漂移算子分解为以下形式会更方便:

$$\mathcal{A}_0 := p \cdot \nabla_q + (-\gamma p - \nabla U(q)) \cdot \nabla_p, \tag{2.3a}$$
$$\mathcal{A}' := -\nabla U(q) \cdot \nabla_q, \tag{2.3b}$$

其中 $\mathcal{A}_0$ 是动能朗之万漂移（$\alpha = 0$），而 $\mathcal{A}'$ 是 HFHR 动力学无穷小生成元中额外的无 Hessian 漂移项。采用（2.3a）-（2.3b）中的记号，我们可以将（2.2）重写为

$$\mathcal{L}_\alpha = \mathcal{A}_0 + \alpha\mathcal{A}' + \alpha\Delta_q + \gamma\Delta_p. \tag{2.4}$$

2.2 关于势函数的假设

现在陈述关于势函数 U 的主要假设。

假设 2.1。 存在常数 $L, A \in (0, \infty)$ 以及 $\lambda \in (0, 1/4]$，使得 U 满足:

(i) 下界与正则性: $U \in C^1(\mathbb{R}^d)$ 且 $U(q) \geq 0$ 对所有 $q \in \mathbb{R}^d$。

(ii) 利普希茨梯度: ∇U 是 L-利普希茨的:

$$|\nabla U(q) - \nabla U(q')| \leq L|q - q'|, \qquad q, q' \in \mathbb{R}^d. \tag{2.5}$$

(iii) 耗散性: U 满足漂移条件

$$\frac{1}{2} q \cdot \nabla U(q) \;\geq\; \lambda \left(U(q) + \frac{\gamma^2}{4}|q|^2 \right) - A, \qquad q \in \mathbb{R}^d. \tag{2.6}$$

假设 2.1 与 [EGZ19]中用于动能朗之万动力学的假设相同，尤其是，它已经蕴含了动能朗之万动力学的指数收敛。为方便起见，我们施加下界 $U \geq 0$，但也可以放宽为 U 有下界。条件 (2.5) 是 L的光滑性条件， U，这在朗之万文献中很常见[DK19, RRT17, EGZ19, DRD20, GGZ22, LZT22]。条件 (2.6) 是耗散条件，用于控制紧集之外 U 的增长；在势函数非凸时，朗之万文献中通常会假设这一条件及其变体 [RRT17, EGZ19, EGZ22]。

2.3 动能朗之万李雅普诺夫函数

本节回顾了[EGZ19]中为动能朗之万动力学引入的李雅普诺夫函数。定义 $\mathcal{V}_0$： $\mathbb{R}^{2d} \to \mathbb{R}$ 如下

$$\mathcal{V}_0(q, p) := U(q) + \frac{\gamma^2}{4} \left(|q + \gamma^{-1} p|^2 + |\gamma^{-1} p|^2 - \lambda |q|^2 \right), \tag{2.7}$$

其中 λ 如假设 2.1所述。令 $\mu_{\min}$ 和 $\mu_{\max}$ 分别表示该对称矩阵的最小和最大特征值

$$M := \frac{1}{4} \begin{pmatrix} \gamma^2(1 - \lambda) & \gamma \\ \gamma & 2 \end{pmatrix}, \tag{2.8}$$

使得

$$\mu_{\min} := \frac{1}{8}\left(\gamma^2(1-\lambda)+2-\sqrt{(\gamma^2(1-\lambda)-2)^2+4\gamma^2}\right),$$
$$\mu_{\max} := \frac{1}{8}\left(\gamma^2(1-\lambda)+2+\sqrt{(\gamma^2(1-\lambda)-2)^2+4\gamma^2}\right). \tag{2.9}$$

由于 $\lambda \le 1/4$, 我们有 $\det(M) = \frac{\gamma^2}{16}(1-2\lambda) > 0$，从而保证了 $\mu_{\min} > 0$。于是，对于所有 $(q,\ p) \in \mathbb{R}^{2d}$ ，

$$c_1\left(1+U(q)+|q|^2+|p|^2\right) \le 1+\mathcal{V}_0(q,p) \le c_2\left(1+U(q)+|q|^2+|p|^2\right), \tag{2.10}$$

均以显式常数成立

$$c_1 := \min(1,\mu_{\min}), \qquad c_2 := \max(1,\mu_{\max}). \tag{2.11}$$

此外，在假设2.1下， $\mathcal{V}_0$ 是动能朗之万无穷小生成元 $\mathcal{L}_0$的李雅普诺夫函数:

$$\mathcal{L}_0\mathcal{V}_0(q,p) \ \le\ \gamma\left(d+A-\lambda\,\mathcal{V}_0(q,p)\right), \tag{2.12}$$

其中， λ 和 A 是假设2.1（iii）中规定的常数；参见 [EGZ19，命题 2。4]特别地， $\mathcal{V}_0$ 已经足以推出动能朗之万动力学指数收敛至平衡态。

下文中，我们首先将在第3 节证明，对于 α 足够小时，未经改进的李雅普诺夫函数 $\mathcal{V}_0$ 仍然满足 HFHR 无穷小生成元 $\mathcal{L}_\alpha$ 的漂移条件，因而可推出 HFHR 动力学的指数收敛。在第4节中，我们随后构造一个改进的李雅普诺夫函数 $\mathcal{V}_\alpha = \mathcal{V}_0 + \alpha\mathcal{M}$ ，并在对 U的额外结构假设下，获得 HFHR 动力学的改进的漂移率和收缩率。

2.4 HFHR 动力学的基准李雅普诺夫漂移

现在我们记录一个简单的扰动结果，该结果表明，当 α 足够小时，动能朗之万李雅普诺夫函数 $\mathcal{V}_0$ 仍然满足关于 HFHR 无穷小生成元 $\mathcal{L}_\alpha$的李雅普诺夫漂移条件。

命题 2.2 （HFHR 动力学的基准李雅普诺夫漂移）。假设 *2.1* 成立，并令 $\mathcal{V}_0$ 按 （2.7）中的定义给出。则对于每个 $\alpha \ge 0$，*HFHR* 无穷小生成元 $\mathcal{L}_\alpha$（2.2）满足漂移不等式

$$\mathcal{L}_\alpha\mathcal{V}_0(q,p) \ \le\ \gamma\left(d+A_\alpha-\hat{\lambda}_\alpha\,\mathcal{V}_0(q,p)\right), \qquad (q,p)\in\mathbb{R}^{2d}, \tag{2.13}$$

其中

$$A_\alpha \ :=\ A+\frac{J_1}{\gamma}\,\alpha, \qquad \hat{\lambda}_\alpha \ :=\ \lambda-\frac{J_1}{\gamma}\,\alpha, \tag{2.14}$$

A 和 λ 是来自假设 *2.1* （*iii*）的常数，而常数 J_1 可以明确地选为

$$J_1 \ :=\ K_A+K_\Delta, \quad K_A \ :=\ \frac{1}{c_1}\left[\frac{\gamma^4}{4}(1-\lambda)^2+\frac{\gamma^2}{4}\right], \quad K_\Delta \ :=\ Ld+\frac{\gamma^2}{2}d(1-\lambda), \tag{2.15}$$

其中 $c_1 :=\min(1,\ \mu_{\min})$，其中 $\mu_{\min}$ 在 (2.9) 中给出了明确表达式。特别地，

$$\hat{\lambda}_\alpha \ =\ \lambda-\frac{\alpha}{\gamma}\left\{\frac{1}{c_1}\left[\frac{\gamma^4}{4}(1-\lambda)^2+\frac{\gamma^2}{4}\right]+Ld+\frac{\gamma^2}{2}d(1-\lambda)\right\}. \tag{2.16}$$

因此，如果我们选择

$$\alpha_0 := \frac{\gamma\lambda}{2}\left\{\frac{1}{c_1}\left[\frac{\gamma^4}{4}(1-\lambda)^2+\frac{\gamma^2}{4}\right]+Ld+\frac{\gamma^2}{2}d(1-\lambda)\right\}^{-1}, \tag{2.17}$$

ˆ于是 $\lambda_\alpha \geq \lambda/2 > 0$ 对所有 $\alpha \in [0, \alpha_0]$成立。

证明。 证明见附录 A.1。 □

3 全局收缩性：一般框架

本节将为 HFHR 动力学的几何遍历性建立一个一般框架。首先，我们定义反射–同步耦合及其对应的传输半度量。随后，我们证明一个“主定理”：若任意 李雅普诺夫函数满足漂移速率为 $\lambda > 0$的漂移条件，则动力学会以特定速率 $c(\lambda) > 0$ 收缩，且该速率可显式计算。最后，我们将这一框架应用于动能朗之万李雅普诺夫函数 $\mathcal{V}_0$ ，从而证明当分辨率参数 α 较小时，HFHR 动力学具有全局收缩性。

3.1 耦合构造

我们构造两个 HFHR 过程 $(z_t)_{t\geq 0} = (q_t,\ p_t)_{t\geq 0}$ 和 $(z'_t)_{t\geq 0} = (q'_t,\ p'_t)_{t\geq 0}$ 的耦合，它们由相同的参数 α, $\gamma > 0$驱动。令

$$\Delta q_t := q_t - q'_t, \qquad \Delta p_t := p_t - p'_t. \tag{3.1}$$

继 [Ebe16、EGZ19],的做法，我们定义有效速度差

$$\mathbf{R}_t := \Delta q_t + \gamma^{-1}\Delta p_t. \tag{3.2}$$

令 $e_t := \mathbf{R}_t/|\mathbf{R}_t|$ 若 $\mathbf{R}_t \neq 0$，否则固定任意一个单位向量。记 $\mathcal{P}_t := e_t e_t^\top$为到 e_t张成空间的正交投影。

耦合定义如下：两份过程均满足 HFHR 随机微分方程（2.1），并分别由满足以下条件的布朗运动（$B^q,\ B^p$）和（$B^{q'},\ B^{p'}$）驱动

$$dB_t^{q'} = dB_t^q, \qquad dB_t^{p'} = (I_d - 2\chi(t)\mathcal{P}_t)\, dB_t^p, \tag{3.3}$$

其中 $\chi(t) \in \{0, 1\}$ 是一个控制过程，在有效速度方向上的反射耦合（$\chi(t) = 1$）与同步耦合（$\chi(t) = 0$）之间进行插值。的精确选取$\chi(t)$取决于当前距离，将在命题3.7的证明中给出。

截断耦合族。 为便于后续使用，我们引入一个由截断参数 $\xi > 0$索引的近似粘滞耦合族。定义 $\chi_\xi(t)$: $= \mathbf{1}_{\{|\mathbf{R}_t|\geq\xi\}}$，并在（3.3）中使用 $\chi(t) = \chi_\xi(t$。极限粘滞耦合通过令 $\xi \downarrow 0$ 得到。

3.2 距离函数与可容许李雅普诺夫函数

接下来, 我们定义底层距离和 Lyapunov 加权半度量。令

$$L_{\mathrm{eff}}(\alpha) := (1+\alpha\gamma)L, \tag{3.4}$$

并固定一个松弛参数 $\eta_0 > 0$。定义度量权重

$$\theta := (1+\eta_0)\, L_{\mathrm{eff}}(\alpha)\, \gamma^{-2}. \tag{3.5}$$

那么，对于 $z=(q,p)$ 和 $z'=(q',p')$，我们定义

$$r(z,z') := \theta\,|q-q'| + \left|(q-q') + \gamma^{-1}(p-p')\right|, \tag{3.6}$$

其中，全文中 $|\cdot|$ 表示 $\mathbb{R}^d$ 上的欧几里得范数。

接下来，我们引入可容许李雅普诺夫函数类。

定义 3.1（可容许李雅普诺夫函数）。若函数 $\mathcal{V}\colon \mathbb{R}^{2d} \to [0,\infty)$ 相对于无穷小生成元 $\mathcal{L}_\alpha$ 是（λ, D）-可容许的，且其为 C^2（或 C^1 且导数局部利普希茨连续），并满足漂移不等式

$$\mathcal{L}_\alpha \mathcal{V}(q,p) \;\le\; \gamma\,(d + D - \lambda \mathcal{V}(q,p))\,, \qquad \text{for a.e. } (q,p)\in\mathbb{R}^{2d}, \tag{3.7}$$

对于某些常数 $\lambda > 0$ 和 $D \in \mathbb{R}$。

注 3.2。虽然生成元 $\mathcal{L}_\alpha$ 在（2.2）中定义，包含拉普拉斯算子 Δ_q，但并不要求 C^2 具有严格的正则性 U。在假设 2.1下， ∇U 是利普希茨连续的；根据拉德马赫定理， U 的二阶导数（因此 $\mathcal{V}_0$ 在（2.7）中的导数以及 $\mathcal{V}_\alpha$ 在（4.16）中的导数）几乎处处存在，且本质有界。因此，漂移不等式（3.7）应当理解为几乎处处意义下成立。

与 $\mathcal{V}$相关的梯度常数。对于形式为 $\mathcal{V} = \mathcal{V}_0 + \mathfrak{Q}$的李雅普诺夫函数，其中 $\mathfrak{Q}(z) := \frac{1}{2} z^\top \mathsf{A} z$，且 $\mathsf{A} := \begin{pmatrix} \mathsf{A}_{qq} & \mathsf{A}_{qp} \\ \mathsf{A}_{pq} & \mathsf{A}_{pp} \end{pmatrix}$，$\mathsf{A}_{pq}$，$\mathsf{A}_{qp} = \mathsf{A}_{pq}^\top$，定义

$$\bar{C}_{\mathcal{V}} := \max\left\{1, (2\theta)^{-1}\right\} + \frac{C_{\mathfrak{Q}}}{\gamma k_1}, \qquad C_{\mathfrak{Q}} := \|\mathsf{A}_{pp}\|_{\mathrm{op}} + \|\mathsf{A}_{pq}\|_{\mathrm{op}}, \tag{3.8}$$

其中 θ 是式（3.5）中的度量权重，而 k_1 来自引理3.3。

凹距离剖面。请注意，通过 $C_{\mathcal{V}}$ 该剖面 f_λ 取决于所选的李雅普诺夫函数 $\mathcal{V}$；我们在符号中省略这一依赖关系。固定参数 $\eta_0 > 0$、 $c > 0$和 $\varepsilon > 0$（稍后确定）。按照 [Ebe16、EGZ19], , 我们构造一个凹距离剖面 f_λ： $[0,\infty) \to [0,\infty)$，使其适配 (3.5) 中的度量权重 θ，具体如下。令 $R_1(\lambda) = R_1(\lambda;\ L_{\mathrm{eff}}(\alpha)) > 0$ 为将在定理 3.8 的证明中指定的截断半径。对于 $s \ge 0$，定义

$$\varphi_\lambda(s) := \exp\!\left(-\frac{1+\eta_0}{8}\, L_{\mathrm{eff}}(\alpha)\, s^2 - \frac{\gamma^2}{2}\,\varepsilon\,\bar{C}_{\mathcal{V}}\, s^2\right), \tag{3.9}$$

and

$$\Phi_\lambda(s) := \int_0^s \varphi_\lambda(x)\,dx. \tag{3.10}$$

接下来定义辅助修正因子

$$g_\lambda(r) := 1 - \frac{9}{4}\,c\,\gamma \int_0^r \Phi_\lambda(s)\,\left(\varphi_\lambda(s)\right)^{-1}\,ds, \tag{3.11}$$

最后设

$$f_\lambda(r) := \int_0^{r\wedge R_1} \varphi_\lambda(s)\,g_\lambda(s)\,ds, \qquad r \geq 0. \tag{3.12}$$

特别地， f_λ 在 $[0, R_1(\lambda)]$, 上单调递增且为凹函数，并在 $[R_1(\lambda)\ \infty)$上为常数。此外，对于 $r \in (0,$ $R_1(\lambda))$，几乎处处有 $f'_\lambda(r) = \varphi_\lambda(r)g_\lambda(r)$，且 $f''_\lambda(r) = \varphi'_\lambda(r)g_\lambda(r) + \varphi_\lambda(r)g'_\lambda(r)$。式（3.9）的选取旨在确保：在反射耦合起作用的短距离区域内，涉及 f''_λ的项会与区域估计中半度量漂移的“有害”漂移贡献相抵消；该贡献与 rf'^{-}_λ成正比（包括由$C_\mathcal{V}$控制的交叉变差项）。

李雅普诺夫加权半度量。给定一个$(\lambda,\ D)$-可容许李雅普诺夫函数 $\mathcal{V}$，我们将李雅普诺夫加权半度量定义为

$$\rho_\mathcal{V}(z, z') := f_\lambda\big(r(z, z')\big)\,\big(1 + \varepsilon\mathcal{V}(z) + \varepsilon\mathcal{V}(z')\big)\,, \qquad z, z' \in \mathbb{R}^{2d}, \tag{3.13}$$

其中 r 和 f_λ 定义于式（3.6）和式（3.12）中。概率测度 $\mu,\ \nu$ 在 $\mathbb{R}^{2d}$ 上的相应 Wasserstein 距离为

$$\mathcal{W}_{\rho_\mathcal{V}}(\mu, \nu) := \inf_{\Gamma\in\Pi(\mu,\nu)} \int_{\mathbb{R}^{2d}\times\mathbb{R}^{2d}} \rho_\mathcal{V}(z, z')\,\Gamma(dz, dz'),$$

其中 $\Pi(\mu,\ \nu)$ 是 μ 和 ν 的耦合集合。

3.3 半鞅分解与区域分析

我们研究由 HFHR 随机微分方程（2.1）和耦合噪声（3.3）定义的耦合 HFHR 过程。定义

$$Z_t := q_t - q'_t, \qquad W_t := p_t - p'_t, \qquad \mathbf{R}_t := Z_t + \gamma^{-1}W_t, \qquad r_t := r(z_t, z'_t). \tag{3.14}$$

给定一个（$\lambda,\ D$）-可容许李雅普诺夫函数 $\mathcal{V}$ 及其在式（3.9）–（3.12）中构造的相应剖面 f_λ，令

$$G_t := 1 + \varepsilon\mathcal{V}(z_t) + \varepsilon\mathcal{V}(z'_t), \qquad \rho_t := f_\lambda(r_t)\,G_t. \tag{3.15}$$

在进行 ρ_t 的漂移分析之前，我们先说明，底层距离 r 与相空间上的欧几里得度量等价。

引理 3.3（r 与欧几里得距离的等价性）。设 r 由（3.6）定义，即

$$r(z, z') = \theta|q - q'| + \left|(q - q') + \gamma^{-1}(p - p')\right|, \qquad z = (q, p),\ z' = (q', p') \in \mathbb{R}^{2d},$$

其中 $\theta = (1+\eta_0)\, L_{\mathrm{eff}}(\alpha)\, \gamma^{-2} > 0$。那么，对于所有 $z, z' \in \mathbb{R}^{2d}$,

$$k_1 \, |z - z'| \;\le\; r(z, z') \;\le\; k_2 \, |z - z'|, \tag{3.16}$$

其中常数 k_1、k_2 由下式明确给出

$$k_1 := \frac{\theta}{1 + \gamma(1+\theta)}, \qquad k_2 := \sqrt{(\theta+1)^2 + \gamma^{-2}}. \tag{3.17}$$

特别地， r 与 $\mathbb{R}^{2d}$上的欧几里得距离等价。

证明。 我们在附录 B.1中给出证明。 □

下一步是分析李雅普诺夫加权距离过程$\rho_t = f_\lambda(r_t)$的半鞅漂移。我们将反复使用这样一个事实：r 与欧几里得度量相当；此外，为了在距离漂移的$|Z_t|$-项中保持严格为正的耗散系数，我们还需要对 α 施加一个温和的小量条件。

关于 α 的小量条件。 固定一个参数 $\kappa_{\mathrm{adjust}} \in (0,\ 1)$。由于 HFHR 随机微分方程（$-\alpha\nabla U(q)$ 中的漂移项 $Z_t = q_t - q'_t$，差值 $-\alpha(\nabla U(q_t) - \nabla U(q'_t)$ 的动力学包含项 $\alpha L|Z_t|$。这会在 $d|Z_t|$的单侧估计中产生大小为 r_t的额外贡献。当我们将其转化为距离过程 $\eta_0/(1+\eta_0$的漂移界时，该项会降低动能朗之万情形中存在的基准耗散系数。因此，我们引入净 耗散参数

$$\delta_\alpha := \frac{\eta_0}{1+\eta_0} - \frac{\alpha L}{\gamma}. \tag{3.18}$$

条件 $\delta_\alpha > 0$意味着，额外的 HFHR 漂移不会压过基准收缩，因此 r_t 的漂移界在 $|Z_t|$ 中保留了严格耗散的线性项，且该项关于时间一致；这是建立区域收缩估计所必需的。在整个区域分析过程中，我们假设

$$\alpha \le (1 - \kappa_{\mathrm{adjust}}) \, \frac{\eta_0}{1+\eta_0} \, \frac{\gamma}{L}, \tag{3.19}$$

从而 $\delta_\alpha \ge \kappa_{\mathrm{adjust}} \frac{\eta_0}{1+\eta_0} > 0$。因此，在 r_t 的漂移界中，我们采用耗散系数 δ_α（或者在便于使用一致界时采用 $\kappa_{\mathrm{adjust}} \frac{\eta_0}{1+\eta_0}$）。

注 3.4。 在动能朗之万情形 $\alpha = 0$ 中，我们有 $\delta_\alpha = \eta_0/(1+\eta_0)$。因此，在界中可以取 $\kappa_{\mathrm{adjust}} \uparrow 1$（实际上也可以取 $\kappa_{\mathrm{adjust}} = 1$），从而在没有额外前置因子的情况下，恢复相应的动能朗之万收缩率。

接下来，对于任意固定的 $c \in \mathbb{R}$，我们推导 $e^{ct}\rho_t$ 的漂移分解。对于形式为 $\mathcal{V} = \mathcal{V}_0 + \mathfrak{Q}$（其中 $\mathfrak{Q}$ 为二次项）的李雅普诺夫函数 $\mathcal{V}$，我们确定关键的漂移系数，并将逐区域对其进行估计。

引理 3.5 （漂移分解）。回顾 δ_α 在 （3.18）中的定义，并假设

$$\delta_\alpha = \frac{\eta_0}{1+\eta_0} - \frac{\alpha L}{\gamma} > 0, \qquad i.e. \qquad \alpha < \frac{\eta_0}{1+\eta_0} \frac{\gamma}{L}.$$

固定 $\varepsilon > 0$ 和 $c \in \mathbb{R}$。回顾 (3.14) 和 (3.15) 可知

$$Z_t := q_t - q'_t, \quad W_t := p_t - p'_t, \quad \mathbf{R}_t := Z_t + \gamma^{-1} W_t, \quad r_t := \theta|Z_t| + |\mathbf{R}_t|, \tag{3.20}$$

$$G_t := 1 + \varepsilon \mathcal{V}(z_t) + \varepsilon \mathcal{V}(z'_t), \quad \rho_t := f_\lambda(r_t) G_t, \tag{3.21}$$

其中 $\theta := (1+\eta_0) L_{\text{eff}}(\alpha) \gamma^{-2}$ 对某个 $\eta_0 > 0$ 如 (3.5)所示。假设 $\mathcal{V}$ 是一个 (λ, D)-可容许李雅普诺夫函数，并且

$$\mathcal{V}(z) = \mathcal{V}_0(z) + \mathfrak{Q}(z), \tag{3.22}$$

其中 $\mathcal{V}_0$ 是动能李雅普诺夫函数 (2.7)，而 $\mathfrak{Q}$ 是定义在 $\mathbb{R}^{2d}$ 上的二次型（在 $z = (q,p)$中），即

$$\mathfrak{Q}(z) := \frac{1}{2} z^\top \mathsf{A} z, \qquad \mathsf{A} \in \mathbb{R}^{2d \times 2d} \text{ is symmetric.}$$

Then

$$e^{ct} \rho_t \le \rho_0 + \gamma \int_0^t e^{cs} K_s \, ds + M_t,$$

其中 M_t 是连续局部鞅，且 K_t 满足

$$\begin{aligned} K_t \le\ & 4\gamma^{-2} \left(\chi(t)\right)^2 f''_\lambda(r_t) G_t + \left(\theta|\mathbf{R}_t| - \delta_\alpha \, \theta |Z_t|\right) f'_{\lambda,-}(r_t) \, G_t \\ & + 4\varepsilon \bar{C}_{\mathcal{V}} \left(\chi(t)\right)^2 r_t f'_{\lambda,-}(r_t) + \gamma^{-1} \varepsilon f_\lambda(r_t) \left[\mathcal{L}_\alpha \mathcal{V}(z_t) + \mathcal{L}_\alpha \mathcal{V}(z'_t)\right] + \gamma^{-1} c f_\lambda(r_t) G_t, \end{aligned} \tag{3.23}$$

¯其中 $C_{\mathcal{V}}$ 定义见 (3.8)，且 k_1 是引理 *3.3*中的范数等价常数。

证明。 证明见附录 B.2。 □

注 3.6。 结构假设（3.22）是为分析 HFHR 动力学而设。在第 4 节中，我们将构造一个精细李雅普诺夫函数 $\mathcal{V}_\alpha = \mathcal{V}_0 + \alpha \mathcal{M}$，对应于选择 $\mathfrak{Q}(z) := \alpha \mathcal{M}(z)$。由于 $\mathcal{M}$ 将构造为二次多项式（见引理 4.3），其梯度是线性的，因此关于 $\nabla_p \mathfrak{Q}$的利普希茨条件在全局范围内自动成立。对于基线收敛结果（推论 3.9），我们只需取 $\mathfrak{Q} \equiv 0$ （从而 $C_{\mathfrak{Q}} = 0$）。

接下来，我们要估计状态空间不同区域中的 K_t 。我们区分小距离和中等距离，此时 f_λ 的凹性和反射耦合起主导作用；以及大距离，此时 $\mathcal{V}$ 的李雅普诺夫漂移项占主导。

命题 3.7 （区域压缩性）。假设假设 *2.1* 成立。固定 $\eta_0 > 0$ 和 $\kappa_{\text{adjust}} \in (0,1)$。假设小量条件 (3.19) 成立：

$$\alpha \le (1 - \kappa_{\text{adjust}}) \frac{\eta_0}{1+\eta_0} \frac{\gamma}{L}, \tag{3.24}$$

使得 $\delta_\alpha \ge \kappa_{\text{adjust}} \frac{\eta_0}{1+\eta_0} > 0$，其中 δ_α 定义见 (3.18)。设 $\mathcal{V}$ 是 (λ, D)-可容许李雅普诺夫函数，适用于 $\mathcal{L}_\alpha$，其含义见定义 *3.1*。进一步假设 $\mathcal{V}$ 是强制的：存在常数 $a_{\mathcal{V}} > 0$ 和 $b_{\mathcal{V}} \ge 0$ 使得

$$\mathcal{V}(z) \ \ge\ a_{\mathcal{V}} |z|^2 - b_{\mathcal{V}}, \qquad z \in \mathbb{R}^{2d}. \tag{3.25}$$

令 θ 如 *(3.5)* 所示，并令 $r(z,z')$ 如 *(3.6)* 所示。设 f_λ 由 *(3.9)–(3.12)* 定义，其中采用某个截断半径 $R_1(\lambda) = R_1(\lambda;\ L_{\text{eff}}(\alpha)) > 0$。

对于每个 $\xi > 0$，考虑通过选择 $\chi(t) = \chi_\xi(t) = \mathbf{1}_{\{|\mathbf{R}_t| \geq \xi\}}$ 在(3.3)中得到的截断耦合。令 $\left(z_t^\xi, z_t'^{,\xi}\right)$ 为所得的耦合过程，并定义

$$G_t^\xi := 1 + \varepsilon \mathcal{V}\left(z_t^\xi\right) + \varepsilon \mathcal{V}\left(z_t'^{,\xi}\right), \qquad \rho_t^\xi := f_\lambda\left(r\left(z_t^\xi, z_t'^{,\xi}\right)\right) G_t^\xi.$$

¯则存在常数 c_0、$\varepsilon_0 > 0$和 $C_{\mathrm{reg}} < \infty$（仅取决于 λ、D、η_0、γ、$L_{\mathrm{eff}}(\alpha)$、$C_{\mathcal{V}}$以及 f_λ的构造，但与 ξ无关），使得对于任意 $0 < c \leq c_0$ 和 $0 < \varepsilon \leq \varepsilon_0$，引理 K_t^ξ *3.5*（应用于 ρ_t^ξ）中的漂移系数满足

$$K_t^\xi \;\leq\; C_{\mathrm{reg}}\, \xi\, G_t^\xi, \qquad t \geq 0. \tag{3.26}$$

因此，对于每个 $t \geq 0$,

$$\mathbb{E}\left[e^{ct} \rho_t^\xi\right] \;\leq\; \mathbb{E}[\rho_0] + \gamma C_{\mathrm{reg}} \xi \int_0^t e^{cs} \mathbb{E}\left[G_s^\xi\right] ds, \tag{3.27}$$

从而 上极限$_{\xi \downarrow 0} \mathbb{E}\left[e^{ct} \rho_t^\xi\right] \leq \mathbb{E}[\rho_0]$。特别地，沿着 $\xi \downarrow 0$ 获得的任何极限（“黏滞”）耦合，在期望意义下都以 c 为收缩率。

证明。 证明见附录 B.3。 □

3.4 关于全局收缩的主定理

记 $(P_t^\alpha)_{t \geq 0}$ 为与 HFHR 动力学（2.1）相关的马尔可夫半群。我们有如下主定理，它表明 HFHR 动力学具有可显式计算的收缩率。

定理 3.8 （全局收缩主定理）。假设假设条件 *2.1* 成立。固定 $\kappa_{\mathrm{adjust}} \in (0,1)$ 并假设小量条件 (3.19) 成立，因此由 δ_α 在 (3.18) 中的定义可知，$\delta_\alpha \geq \kappa_{\mathrm{adjust}} \frac{\eta_0}{1+\eta_0}$。

令 $\mathcal{V}$ 为 *HFHR* 生成元 $\mathcal{L}_\alpha$ 的一个 (λ, D)-可容许李雅普诺夫函数，其含义见定义 *3.1*，其中 $\lambda \in (0, 1/4]$ 且 $d + D > 0$。此外，假设 $\mathcal{V}$ 具有强制性（含义见命题 *3.7*），即 (3.25)成立。

令 $\eta_0 > 0$ 为一个参数（将在下文的 (3.30)中明确选取），并回顾 θ 的定义，见 (3.5) 以及 $r(z, z')$ 的定义，见 (3.6)。令 $R_1(\lambda) > 0$ 满足

$$R_1^2(\lambda) \geq \frac{96(d + A)}{5\lambda(1 - 2\lambda)\gamma^2((1+\theta)^2 + \gamma^{-2})}. \tag{3.28}$$

令 f_λ 由 (3.9)–(3.12) 定义，并取 $R_1(\lambda)$，再令 $\rho_{\mathcal{V}}(z, z')$ 为 (3.13) 中定义的李雅普诺夫加权半度量:

$$\rho_{\mathcal{V}}(z, z') := f_\lambda(r(z, z'))\left(1 + \varepsilon \mathcal{V}(z) + \varepsilon \mathcal{V}(z')\right). \tag{3.29}$$

定义

$$\Lambda_0(\lambda) := \frac{L\, R_1^2(\lambda)}{8}, \qquad \eta_0 := \Lambda_0^{-1}(\lambda). \tag{3.30}$$

回顾一下 $L_{\mathrm{eff}}(\alpha) = (1+\alpha\gamma)L$，并令

$$\Lambda_\alpha(\lambda) := \frac{L_{\mathrm{eff}}(\alpha)\, R_1^2(\lambda)}{8}. \tag{3.31}$$

取

$$c(\lambda) := \frac{\gamma}{384} \min\left\{\widetilde{\Lambda}_{1,\alpha}(\lambda),\, \widetilde{\Lambda}_{2,\alpha}(\lambda),\, \widetilde{\Lambda}_{3,\alpha}(\lambda)\right\}, \tag{3.32}$$

其中

$$\begin{aligned}
\widetilde{\Lambda}_{1,\alpha}(\lambda) &:= \frac{\lambda L_{\mathrm{eff}}(\alpha)}{\gamma^2}, \quad \widetilde{\Lambda}_{2,\alpha}(\lambda) := \Lambda_\alpha^{1/2}(\lambda)\, e^{-\Lambda_\alpha(\lambda)}\, \frac{L_{\mathrm{eff}}(\alpha)}{\gamma^2}, \\
\widetilde{\Lambda}_{3,\alpha}(\lambda) &:= \kappa_{\mathrm{adjust}}\, \Lambda_\alpha^{1/2}(\lambda)\, e^{-\Lambda_\alpha(\lambda)}.
\end{aligned} \tag{3.33}$$

那么，*HFHR* 半群 P_t^α 在加权 *Wasserstein* 距离下呈指数收缩 $\mathcal{W}_{\rho_\mathcal{V}}$。具体而言，设 $c(\lambda)$ 由 (3.32) 定义。任取 $c \in (0, c(\lambda)]$ 并令

$$\varepsilon := \frac{4c}{\gamma(d+D)}. \tag{3.34}$$

Then

$$\mathcal{W}_{\rho_\mathcal{V}}(\mu P_t^\alpha, \nu P_t^\alpha) \le e^{-ct}\, \mathcal{W}_{\rho_\mathcal{V}}(\mu, \nu), \qquad t \ge 0, \tag{3.35}$$

对所有在 $\mathbb{R}^{2d}$ 上具有有限 $\mathcal{V}$ 阶矩的概率测度 μ, ν 均成立。特别地，取 $c = c(\lambda)$，即可得到一个显式的可容许收缩率。

证明。 证明见附录 B.4。 □

特别地，当 $\alpha = 0$ 时，我们有 $L_{\mathrm{eff}}(0) = L$，而收缩率 (3.32) 可化简为 [EGZ19 定理 2.3] 中得到的表达式。此外，(3.32) 中 $c(\lambda)$ 的显式形式已预示了~之后将被分别归入李雅普诺夫分支（记作 $\Lambda_{1,\alpha}$）和~ ~度量分支（记作 $\Lambda_{2,\alpha}$,$\Lambda_{3,\alpha}$）的两种机制。粗略而言，李雅普诺夫分支对应于李雅普诺夫漂移的贡献（通过 $\mathcal{V}$ 的 (λ, D)-可容许性），该贡献控制过程向高能量区域的偏离，并在过程充分远离时产生收缩；而度量分支对应于半度量 $\rho_\mathcal{V}$ 中体现的局部收缩机制（通过 f_λ 的凹性/平坦性设计，以及反射耦合与同步耦合），并在“邻近”区域占主导。因此，$c(\lambda)$ 可解释为平衡这两种效应后得到的有效全局收缩率：它是使李雅普诺夫漂移估计和局部度量估计同时成立的收缩率。

3.5 HFHR动力学的全局收敛

现在，我们将定理 3.8 应用于第 $\mathcal{V}_0$ 节中定义的动能朗之万李雅普诺夫函数 $\mathcal{V}_0$ defined in Section 2。根据命题 2.2，$\mathcal{V}_0$ 满足 HFHR 无穷小生成元所需的漂移条件 $\mathcal{L}_\alpha$，前提是 $\alpha \le \alpha_0$。因此，我们可以将一般主定理（定理 3.8）具体应用于基准李雅普诺夫函数，从而得到以下指数收敛结果。

推论 *3.9*（*HFHR* 动力学的全局收敛）。假设假设 *2.1* 成立。设 $\alpha \in [0, \alpha_0]$，其中 α_0 如命题 *2.2* 所述。取 $\kappa_{\text{adjust}} \in$ *(0, 1)*，并假设小量条件 *(3.19)* 对 $\eta_0 :=(\Lambda_0(\lambda_\alpha))^{-1}$ 成立，其中 $\Lambda_0(\cdot)$ 在 *(3.30)* 中定义。令 $\lambda_\alpha := \hat{\lambda}_\alpha > 0$，其中 $\hat{\lambda}_\alpha$ 在 *(2.16)* 中给出。令 $R_1(\lambda) > 0$ 满足 *(3.28)*，其中包含 $\lambda = \lambda_\alpha$ 和 $A = A_\alpha$；注意，$\eta_0 = (\Lambda_0(\lambda))^{-1}$。令 θ和 r 由 *(3.5)–(3.6)* 定义，令 f_λ 由 *(3.9)–(3.12)* 根据此 $R_1(\lambda)$ 定义，并选择任意满足定理 *3.8*（应用于 $\mathcal{V} = \mathcal{V}_0$ 和 $(\lambda, A) = (\lambda_\alpha, A_\alpha)$）条件的 $0 < c_\alpha \le c_*(\lambda_\alpha, A_\alpha)$ 和 $0 < \varepsilon_\alpha \le \varepsilon_*(\lambda_\alpha, A_\alpha)$。令 $\rho_{\mathcal{V}_0,\alpha}$ 表示相应的李雅普诺夫加权半度量 *(3.29)*。

那么，*HFHR* 动力学 (2.1) 存在唯一的不变概率测度 π_α，属于类$\{\mu : \int_{\mathbb{R}^{2d}} \mathcal{V}_0 \, d\mu < \infty\}$，并且对于所有概率测度 μ，ν满足 $\int_{\mathbb{R}^{2d}} \mathcal{V}_0 \, d\mu + \int_{\mathbb{R}^{2d}} \mathcal{V}_0 \, d\nu < \infty$，

$$\mathcal{W}_{\rho_{\mathcal{V}_0,\alpha}}(\mu P_t^\alpha, \nu P_t^\alpha) \le e^{-c_\alpha t} \mathcal{W}_{\rho_{\mathcal{V}_0,\alpha}}(\mu, \nu), \qquad t \ge 0.$$

特别地，

$$\mathcal{W}_{\rho_{\mathcal{V}_0,\alpha}}(\mu P_t^\alpha, \pi_\alpha) \le e^{-c_\alpha t} \mathcal{W}_{\rho_{\mathcal{V}_0,\alpha}}(\mu, \pi_\alpha), \qquad t \ge 0.$$

此外，c_α 的一个明确可行取值由 (3.32) 给出，见定理 *3.8*，其中 $(\lambda, A) = (\lambda_\alpha, A_\alpha)$。

证明。 证明见附录 B.5。 □

然而，我们注意到，由于采用了无 Hessian 漂移的微扰处理方法，(λ_α 2.16)中的漂移率可能低于基准速率 λ。因此，得到的收缩率 $c(\lambda_\alpha)$ 尚未体现出相较于动能朗之万动力学的加速效果。在第 4 节中，我们将用改进后的李雅普诺夫函数 $\mathcal{V}_\alpha = \mathcal{V}_0 + \alpha\mathcal{M}$ 替代 $\mathcal{V}_0$，并应用相同的抽象框架，以获得更高的收缩率。

3.6 从加权 Wasserstein 距离到 2-Wasserstein 距离

在本小节中，我们说明如何将加权 Wasserstein 距离$\mathcal{W}_{\rho_\mathcal{V}}$ 下的收缩转化为标准 Wasserstein 距离 $\mathcal{W}_2$ 下的收缩。回顾一下，主收缩定理（定理 3.8 ）是以加权代价 $\rho_\mathcal{V}$ 表述的，该代价由李雅普诺夫函数$\mathcal{V}$构造而成。在我们的应用中，将采用两种选择：（i）基准李雅普诺夫函数 $\mathcal{V} = \mathcal{V}_0$，对此我们可以给出显式的二次界；以及（ii）定理 3.8 中出现的一般化（λ, D）-李雅普诺夫函数$\mathcal{V}$ 。从加权 Wasserstein 距离 $\mathcal{W}_{\rho_\mathcal{V}}$ 过渡到标准 2-Wasserstein 距离 $\mathcal{W}_2$，只需满足 $\mathcal{V}$ 能够控制二阶矩（通过二次下界）；而（λ, D）-可容许性仅在后文用于为 $\mathbb{E}[\mathcal{V}(z_t)]$得到一致矩界。为使常数显式可见，我们首先记录规范选择 $\mathcal{V}_0$对应的二次界。首先，我们证明 $\mathcal{V}_0(q, p)$可以用具有显式系数的二次函数给出下界和上界。更准确地说，我们有以下引理。

引理 3.10。对于任意$(q, p) \in \mathbb{R}^{2d}$,

$$c_1' \left(1 + |q|^2 + |p|^2\right) \le 1 + \mathcal{V}_0(q, p) \le c_2' \left(1 + |q|^2 + |p|^2\right), \tag{3.36}$$

其中

$$c_1' := \min(1, \mu_{\min}), \qquad c_2' := \max(1, \mu_{\max}) + U(0) + \frac{L}{2} + \frac{1}{2}|\nabla U(0)|, \tag{3.37}$$

其中 $\mu_{\min}$、$\mu_{\max}$ 是 M （定义见 （2.8）） 的特征值，显式公式见 （2.9）。

证明。证明见附录B.6。 □

特别地，（3.36）意味着 $\mathcal{V}_0$ 控制了 $\mathbb{R}^{2d}$上的欧几里得二阶矩，这正是比较加权 Wasserstein 距离 $\mathcal{W}_{\rho_\mathcal{V}}$ 与标准 2-Wasserstein 距离 $\mathcal{W}_2$ 所需的条件（下面我们将其应用于 $\mathcal{V}=\mathcal{V}_0$，更一般地，也可应用于任何 $\mathcal{V}$满足二次下界的情形）。此外，在 r （3.6）中定义的底层相空间距离等价于 $\mathbb{R}^{2d}$上的欧几里得距离：

$$k_1\,|(q,p)-(q',p')| \;\le\; r\big((q,p),(q',p')\big) \;\le\; k_2\,|(q,p)-(q',p')|, \qquad (q,p),(q',p')\in\mathbb{R}^{2d}, \tag{3.38}$$

其中 k_1， $k_2>0$ 是仅依赖于 θ 和 γ 的显式常数（见（3.17））。不等式（3.38）中的第一个不等式直接由定义（3.6）得出，而不等式（3.38）中的第二个不等式则源于以下事实： r 等价于欧几里得距离；见引理3.3。

下一个引理量化了 $\rho_\mathcal{V}$ 对二次传输代价的控制作用。

引理 3.11。假设假设条件*2.1* 成立。固定 $\eta_0>0$，并令 f_λ 由（3.9）–（3.12）定义，截断半径为 R_1（λ）>0。此外，假设 [1]

$$g_*(\lambda):=\inf_{0\le s\le R_1(\lambda)} g_\lambda(s) \;>\; 0. \tag{3.39}$$

令 $\mathcal{V}\colon \mathbb{R}^{2d}\to[1,\infty)$ *satisfy the quadratic lower bound:*

$$|z|^2\le C_V\,(1+\mathcal{V}(z))\,, \qquad z\in\mathbb{R}^{2d}, \tag{3.40}$$

对于某个常数 $C_V\in(0,\infty)$。对于 $\varepsilon\in$（$0,1$], 令 $\rho_\mathcal{V}$（z， z'）为定义于（3.13）中的 *Lyapunov* 加权半度量：

$$\rho_\mathcal{V}(z,z'):=f_\lambda(r(z,z'))\,\big(1+\varepsilon\mathcal{V}(z)+\varepsilon\mathcal{V}(z')\big)\,.$$

那么存在 $C_\rho<\infty$，使得对于所有满足 $\int_{\mathbb{R}^{2d}}\mathcal{V}\,d\mu+\int_{\mathbb{R}^{2d}}\mathcal{V}\,d\nu<\infty$、定义在 $\mathbb{R}^{2d}$ 上的概率测度 μ、 ν，

$$\mathcal{W}_2^2(\mu,\nu)\le C_\rho\,\mathcal{W}_{\rho_\mathcal{V}}(\mu,\nu).$$

更明确地，可以取

$$C_\rho:=\frac{1}{\varepsilon}\max\left\{\frac{k_1^{-2}R_1(\lambda)}{g_*\,c_r},\;\frac{4C_V}{c_0}\right\}, \tag{3.41}$$

其中

$$c_r:=\inf_{0\le s\le R_1(\lambda)}\varphi_\lambda(s), \qquad c_0:=f_\lambda(R_1(\lambda))=\int_0^{R_1(\lambda)}\varphi_\lambda(s)g_\lambda(s)\,ds,$$

且 $k_1=\frac{\theta}{1+\gamma-(1+\theta)}$ 来自引理 *3.3*。特别地，当 $\mathcal{V}=\mathcal{V}_0$ 时，引理 *3.10* 表明，取 $C_V=1/c_1'$。(3.40)

证明。 证明见附录 B.7。 □

结合引理3.11与收缩（3.35），我们得到以下推论，它将 $\mathcal{W}_{\rho_\mathcal{V}}$ 中的指数收缩提升为标准 2-Wasserstein 距离下的估计$\mathcal{W}_2$。

[1]例如，(3.32) 中的显式参数选择即可确保这一正性条件；参见定理 3.8 的证明中区域 II 的验证，其中我们证明 $g_*\ge 1/2$。

推论 3.12 （ $\mathcal{W}_2$中的指数收缩）。 在定理 *3.8*的假设下，设$c > 0$ 和$\varepsilon = \gamma \frac{4c}{d+D}$（）如（3.34）所示，并设 $\rho_{\mathcal{V}}$ 由（~~3.29~~）定义。设 $C_\rho < \infty$ 为引理 *3.11* 中的常数（用此 ε计算）。那么，对于所有满足 μ、ν 且$\int_{\mathbb{R}^{2d}} \mathcal{V}\, d\mu + \int_{\mathbb{R}^{2d}} \mathcal{V}\, d\nu < \infty$ 的概率测度，以及所有 $t \geq 0$,

$$\mathcal{W}_2(\mu P_t^\alpha, \nu P_t^\alpha) \;\leq\; C_\rho^{1/2}\, e^{-ct/2}\, (\mathcal{W}_{\rho_{\mathcal{V}}}(\mu, \nu))^{1/2}.$$

特别地，若 π_α 是满足 $\int_{\mathbb{R}^{2d}} \mathcal{V}\, d\pi_\alpha < \infty$ 的不变概率测度，则取 $\nu = \pi_\alpha$ 可得 μP_t^α 在 $\mathcal{W}_2$ 中指数收敛至 π_α 。

证明。 证明见附录 B.8。 □

结合推论 3.12 与基准李雅普诺夫函数 $\mathcal{V}_0$ 以及推论 3.9给出的不变测度的存在性与唯一性，我们得到以下基准 $\mathcal{W}_2$ 收敛估计。

推论 *3.13*（$\mathcal{W}_2$ 中的基准指数收敛）。在推论 *3.9* 的假设下，应用定理 *3.8*，其中 $\mathcal{V} = \mathcal{V}_0$ 和 *(*λ*,* D*)* $=$ *(*$\hat{}\lambda_\alpha$*,* A_α*)*。设 $c_\alpha > 0$ 和$\varepsilon_\alpha = \gamma \frac{4c_\alpha}{(d+A_\alpha)}$ 为所得的 $\hat{}$ 参数（ c_α 的一个可行显式选择由 *(3.32)* 给出，其中 $\lambda = \lambda_\alpha$ 和 $D = A_\alpha$），并设 $\rho_{\mathcal{V}_0,\alpha}$ 为相应的加权半度量。设 $C_{\rho,\alpha}$ 为引理 *3.11* 中与 $\mathcal{V}_0$ 对应、并使用 $\varepsilon = \varepsilon_\alpha$ 计算得到的常数。则对所有 $t \geq 0$ 以及任意满足 $\int_{\mathbb{R}^{2d}} \mathcal{V}_0\, d\mu < \infty$ 的概率测度 μ,

$$\mathcal{W}_2(\mu P_t^\alpha, \pi_\alpha) \;\leq\; C_{\rho,\alpha}^{1/2}\, e^{-\frac{1}{2}c_\alpha t} \left(\mathcal{W}_{\rho_{\mathcal{V}_{0,\alpha}}}(\mu, \pi_\alpha)\right)^{1/2}.$$

证明。 证明见附录 B.9。 □

4 加速分析

推论 3.9 表明，当采用基准李雅普诺夫函数 $\alpha \in [0, \alpha_0]$ 和 $\alpha \leq (1 - \kappa_{\text{adjust}}) \frac{\eta_0}{1+\eta_0} \frac{\gamma}{L}$时，HFHR 动力学以指数速度收敛。然而，相应的李雅普诺夫漂移率 $\lambda_\alpha = \lambda - \mathcal{O}(\alpha)$（命题 2.2）略小于 λ，因此定理 3.8 中可选取的收缩参数未必优于未扰动情形 $\alpha = 0$。本节将表明，在 U上增加结构条件后，可以构造一个改进的李雅普诺夫函数 $\mathcal{V}_\alpha$ ，其漂移率随 α呈一阶增长，并重新应用第 3 节的抽象收缩框架，得到加速收敛界。

4.1 结构条件与精细化李雅普诺夫函数

我们引入额外的结构假设，以实现李雅普诺夫漂移的一阶改进。

假设 4.1 （渐近线性梯度）。 存在常数 $C_{\text{linear}} > 0$、对称正定矩阵 $Q_\infty \in \mathbb{R}^{d\times d}$以及非增函数 ϱ：$[0, \infty)$ $\to [0, \infty)$ ，满足 ϱ（r） $\to 0$ 当 $r \to \infty$ 时，因此

$$|\nabla U(q) - Q_\infty q| \leq \varrho(|q|)\, |q|, \qquad |q| \geq C_{\text{linear}}. \tag{4.1}$$

假设 4.1 意味着 ∇U 在一致的相对意义下渐近线性：比值$|\nabla U(q) - Q_\infty q|/|q|$ 在 $|q| \to \infty$时趋于零。沿射线对（4.1）积分可得二次尾部行为 $U(q) = \frac{1}{2} q^\top Q_\infty q + o(|q|^2)$，当 $|q| \to \infty$时成立。该假设涵盖的一类典型情形是 $U(q) = \frac{1}{2} q^\top Q_\infty q + W(q)$，其中 $\nabla W(q) = o(|q|)$，当 $|q| \to \infty$时成立。我们将在第 5 节中验证，包含多阱势（第 5.1 节）、带有 L^p正则化项的贝叶斯线性回归（第 5.2 节）以及贝叶斯二分类（第 5.3 节）在内的示例均满足假设 4.1。

为说明改进后的李雅普诺夫函数构造，我们首先记录额外 HFHR 漂移项 $\mathcal{A}'$ 在（2.3b）作用于基准李雅普诺夫函数 $\mathcal{V}_0$时的精确贡献。

引理 4.2（相互作用漂移的精确分解）。令 $\mathcal{V}_0$ 和 $\mathcal{A}'$ 分别由 （2.7） 和（2.3b） 定义。那么，对于任意势函数 $U \in C^1(\mathbb{R}^d)$，

$$\mathcal{A}'\mathcal{V}_0(q,p) = -|\nabla U(q)|^2 - \frac{\gamma^2}{2}(1-\lambda)\,\nabla U(q)\cdot q - \frac{\gamma}{2}\,\nabla U(q)\cdot p. \tag{4.2}$$

证明。 证明见附录 C.1。 □

直观而言，在假设 4.1下，相互作用漂移 $\mathcal{A}'\mathcal{V}_0$ 在空间尾部的表现类似于 $\mathcal{V}_0$ 的负倍数，这启发我们引入一个校正子 $\mathcal{M}$，以实现一致的漂移增益。这也启发了下一个引理中引入的结构条件和常数 c_{imp}。

引理 4.3（一阶改进）。假设 U 满足假设 *2.1* 和假设 *4.1*，并令 $\mathcal{V}_0$ 为（2.7）中定义的动能朗之万李雅普诺夫函数。那么存在一个函数 $\mathcal{M}$: $\mathbb{R}^{2d} \to \mathbb{R}$，满足以下性质：

(i) 增长性与正则性。 函数 $\mathcal{M}$ 为 C^2，且至多具有二次增长：

$$|\mathcal{M}(q,p)| \;\le\; C_{\mathcal{M}}\left(1 + |q|^2 + |p|^2\right), \qquad (q,p) \in \mathbb{R}^{2d}, \tag{4.3}$$

其中

$$C_{\mathcal{M}} := \frac{\|\mathsf{K}\|_{\mathrm{op}}}{2} < \infty, \tag{4.4}$$

其中 K 见 (4.15)，且其一阶导数至多具有线性增长：

$$|\nabla_q \mathcal{M}(q,p)| + |\nabla_p \mathcal{M}(q,p)| \;\le\; C_{\mathcal{M}}\,(1 + |q| + |p|), \qquad (q,p) \in \mathbb{R}^{2d}. \tag{4.5}$$

此外，

$$|\Delta_q \mathcal{M}(q,p)| + |\Delta_p \mathcal{M}(q,p)| \;\le\; C_\Delta\,(1 + \mathcal{V}_0(q,p)), \qquad (q,p) \in \mathbb{R}^{2d}, \tag{4.6}$$

其中

$$C_\Delta := 2d\,\|\mathsf{K}\|_{\mathrm{op}} < \infty. \tag{4.7}$$

(ii) 一阶改进。 存在显式常数 $c_{\mathrm{imp}} > 0$ 和 $C_{\mathrm{imp}} \ge 0$，使得

$$\mathcal{A}_0\mathcal{M}(q,p) + \mathcal{A}'\mathcal{V}_0(q,p) \;\le\; C_{\mathrm{imp}} - \underline{c}_{\mathrm{imp}}\,\mathcal{V}_0(q,p), \qquad (q,p) \in \mathbb{R}^{2d}, \tag{4.8}$$

其中， $\mathcal{A}_0$ 和 $\mathcal{A}'$ 定义见(2.3a)–(2.3b)。此外，常数c_{imp} 可取为

$$\underline{c}_{\mathrm{imp}} := \frac{3}{8}\cdot\frac{a_{\min} + 1 - \sqrt{(a_{\min}-1)^2 + \gamma^2}}{a_{\max} + 1 + \sqrt{(a_{\max}-1)^2 + \gamma^2} \;+\; 8\delta_U(R_0)}, \tag{4.9}$$

and

$$C_{\mathrm{imp}} := \sup_{p,q\in\mathbb{R}^d:|q|\le R_0} \left\{\mathcal{A}_0\mathcal{M}(q,p) + \mathcal{A}'\mathcal{V}_0(q,p) + \underline{c}_{\mathrm{imp}}\mathcal{V}_0(q,p)\right\} < \infty, \tag{4.10}$$

其中

$$a_{\min} := \lambda_{\min}(Q_\infty) + \frac{\gamma^2}{2}(1-\lambda), \qquad a_{\max} := \lambda_{\max}(Q_\infty) + \frac{\gamma^2}{2}(1-\lambda), \tag{4.11}$$

且 $\delta_U(R)$ 定义为

$$\delta_U(R) := \sup_{|q|\ge R} \frac{\left|U(q) - \frac{1}{2}\langle Q_\infty q, q\rangle\right|}{1+|q|^2}, \qquad R \ge 1.$$

这一选择足够，因为在 $\{|q| \ge R_0\}$ 上，我们有一致负漂移 $\mathcal{A}_0\mathcal{M} + \mathcal{A}'\mathcal{V}_0 \le -c_{\mathrm{imp}}\mathcal{V}_0$。最后，截断半径 R_0 明确定义如下。令 Q_∞ 和 $\varrho(\cdot)$ 如假设*4.1*所述。定义尾部模量

$$\rho_\nabla(R) := \sup_{|q|\ge R} \frac{|\nabla U(q) - Q_\infty q|}{|q|}, \qquad R \ge 1. \tag{4.12}$$

注意，根据假设 *4.1*，$\rho_\nabla(R) \le \varrho(R)$ 对所有 $R \ge C_{\mathrm{linear}}$均成立。于是令

$$\rho_\star := \frac{-A + \sqrt{A^2 + \frac{5}{4}\underline{a}}}{2}, \quad A := 2\left(\|\mathsf{K}_{pq}\|_{\mathrm{op}} + \|\mathsf{K}_{pp}\|_{\mathrm{op}}\right) + 4\lambda_{\max}(Q_\infty) + \gamma^2|1-\lambda| + \gamma, \tag{4.13}$$

and

$$R_0 := \inf\{R \ge \max\{1, C_{\mathrm{linear}}\} : \rho_\nabla(R) \le \rho_\star\}, \tag{4.14}$$

其中 $\underline{a}$ 由证明中的 (C.20) 给出。

特别地，$\mathcal{M}$ 可以选为关于$(q,p) \in \mathbb{R}^{2d}$ 的二次多项式，使得 $\mathcal{M}(z) = \frac{1}{2}z^\top \mathsf{K} z$，其中

$$\mathsf{K} = \begin{pmatrix} \mathsf{K}_{qq} & \mathsf{K}_{qp} \\ \mathsf{K}_{pq} & \mathsf{K}_{pp} \end{pmatrix}, \tag{4.15}$$

是对称的，且K 是下式的解

$$B^\top \mathsf{K} + \mathsf{K} B = C_{B_1}, \qquad B_1(z) = \frac{1}{2} z^\top C_{B_1} z,$$

with

$$B_1(q,p) := -Q(q,p) - \left(\frac{1}{2}|p|^2 + \frac{\gamma}{2}\langle q,p\rangle + \frac{1}{2}\left\langle \left(Q_\infty + \frac{\gamma^2}{2}(1-\lambda)I_d\right) q, q\right\rangle\right).$$

等价地，$C_{B_1} = \nabla^2 B_1$ 是与二次型 B 相关联的显式对称矩阵 1.

证明。证明见附录C.2。 □

备注 4.4。引理 4.3 仅用于实现李雅普诺夫漂移率的一阶改进（并通过主收缩定理，改进 Wasserstein 收缩率），适用于 HFHR 动力学。仅凭假设2.1并使用未经校正的李雅普诺夫函数 $\mathcal{V}_0$，即可得到基本的几何遍历性和收缩性。在第 5 节中，我们验证假设 4.1，并通过若干代表性示例说明引理 $\mathcal{M}$ 4.3 中二次校正项的构造，包括多阱势（第 5.1 节）、带有 L^p 正则化项的贝叶斯线性回归（第 5.2 节）以及贝叶斯二分类（第 5.3 节）。

在引理 4.3 的假设下，定义精细李雅普诺夫函数

$$\mathcal{V}_\alpha(q,p) := \mathcal{V}_0(q,p) + \alpha\mathcal{M}(q,p). \tag{4.16}$$

现在我们证明， $\mathcal{V}_\alpha$ 满足 HFHR 生成元 $\mathcal{L}_\alpha$ 的一个严格改进的一阶速率李雅普诺夫漂移条件，其中 α_*。回顾式（2.4），记

$$\mathcal{L}_\alpha = \mathcal{L}_0 + \alpha\mathcal{A}' + \alpha\Delta_q, \tag{4.17}$$

其中， $\mathcal{L}_0 = \mathcal{A}_0 + \gamma\Delta_p$ 是动能朗之万生成元，且 $\mathcal{A}_0$、 $\mathcal{A}'$ 定义见式（2.3a）–（2.3b）。

在分析 $\mathcal{V}_\alpha$ 的漂移项之前，我们先验证扰动项 $\alpha\mathcal{M}$ 不会改变李雅普诺夫函数的整体增长：当 α 足够小时， $\mathcal{V}_\alpha$ 与 $\mathcal{V}_0$ 仍然等价，且其等价常数可以显式给出。

引理 4.5 （ $\mathcal{V}_\alpha$ 与 $\mathcal{V}_0$的等价性）。假设假设条件*2.1* 和引理*4.3*成立。设 $C_\mathcal{M}$ 如式（4.4）所示，并定义

$$\alpha_* := \frac{c_1}{2C_\mathcal{M}}. \tag{4.18}$$

那么，对所有 $\alpha \in [0,\alpha_*]$ 以及所有 $(q,\ p)\ \in \mathbb{R}^{2d}$,

$$\frac{1}{2}\left(1+\mathcal{V}_0(q,p)\right) \ \le\ 1+\mathcal{V}_\alpha(q,p)\ \le\ \frac{3}{2}\left(1+\mathcal{V}_0(q,p)\right). \tag{4.19}$$

证明。证明见附录C.3。 □

在建立了增长界之后，我们现在转向对漂移项的分析。下面的引理给出了生成元作用 $\mathcal{L}_\alpha\mathcal{V}_\alpha$ 按 α的幂展开的形式，使我们能够分离出导致加速的一阶贡献。

引理 4.6 （ $\mathcal{V}_\alpha$的漂移展开）。 在假设 *2.1* 和引理 *4.3*的条件下，设 $\mathcal{V}_\alpha = \mathcal{V}_0 + \alpha\mathcal{M}$，其中 $\mathcal{M}$ 是引理 *4.3*中构造的二次多项式。设 $c_{\text{imp}} > 0$ 和 $C_{\text{imp}} \ge 0$ 是式 (4.8)–(4.9)中的显式常数。_设 c_1 是式 (2.10)中的常数。定义

$$K_\Delta := dL + \frac{\gamma^2}{2}d|1-\lambda|, \qquad \textit{so that}\quad |\Delta_q\mathcal{V}_0(q,p)| \le K_\Delta \quad \textit{for a.e. } (q,p)\in\mathbb{R}^{2d}.$$

此外，由于 $\mathcal{M}(z) = \frac{1}{2}z^\top \mathsf{K} z$，其中 K 如式 (4.15)所示，

$$\Delta_p\mathcal{M}(q,p) = \operatorname{tr}(\mathsf{K}_{pp}) \quad \textit{and} \quad \Delta_q\mathcal{M}(q,p) = \operatorname{tr}(\mathsf{K}_{qq}) \qquad \textit{for all } (q,p)\in\mathbb{R}^{2d}.$$

那么对于所有 $\alpha \in [0,1]$ 以及所有 $(q,p)\ \in \mathbb{R}^{2d}$,

$$\mathcal{L}_\alpha\mathcal{V}_\alpha(q,p) \le \gamma(d+A) - \lambda\mathcal{V}_0(q,p) + \alpha\left(C_1 - \underline{c}_{\text{imp}}\mathcal{V}_0(q,p)\right) + C_2\alpha^2(1+\mathcal{V}_0(q,p)), \tag{4.20}$$

其中我们可以选择

$$C_1 := C_{\text{imp}} + \gamma\operatorname{tr}(\mathsf{K}_{pp}) + K_\Delta, \qquad C_2 := |\operatorname{tr}(\mathsf{K}_{qq})| + 3\,C_\mathcal{M}\frac{(L+|\nabla U(0)|)}{c_1}, \tag{4.21}$$

其中 $C_\mathcal{M}$ 是式 (4.4)中的常数。

证明。 证明见附录 C.4。 □

基于这一展开以及在引理 4.5中确立的 $\mathcal{V}_\alpha$ 与 $\mathcal{V}_0$ 的等价性，我们现在可以陈述本节的主要结果之一：改进后的李雅普诺夫函数 $\mathcal{V}_\alpha$ 在较小的情况下可带来严格更优的漂移率 $\alpha_\circ$

命题 4.7（HFHR 动力学的增强漂移率）。 假设条件 *2.1* 和引理 *4.3*成立，并令 $\mathcal{V}_\alpha = \mathcal{V}_0 + \alpha\mathcal{M}$ 如 (4.16)中所定义。令 $\alpha_0 > 0$ 为 (2.17)中的显式常数（来自命题 *2.2*）。令 $C_\mathcal{M}$ 为 (4.3)中的常数，且 c_1 为 (2.10)中的常数。定义

$$\widetilde{C}_\mathcal{M} := \frac{C_\mathcal{M}}{c_1}, \qquad \textit{so that} \qquad |\mathcal{M}| \le \widetilde{C}_\mathcal{M}\,(1 + \mathcal{V}_0). \tag{4.22}$$

令 C_1、C_2 为引理 *4.6*中的显式常数。定义

$$\delta := \underline{c}_{\mathrm{imp}} - \lambda\,\widetilde{C}_\mathcal{M}, \tag{4.23}$$

$$C_\lambda := C_2 + \widetilde{C}_\mathcal{M}\,\underline{c}_{\mathrm{imp}}, \tag{4.24}$$

其中 λ 为假设条件 *2.1**(iii)* 中的耗散常数。定义有效漂移常数 A'_α（注意：A'_α 对应于定义 A中的“*3.1*”）为

$$A'_\alpha := A + \frac{\alpha}{\gamma}\left[C_1 + \alpha C_2 + \widetilde{C}_\mathcal{M}(\lambda + \underline{c}_{\mathrm{imp}})\right]. \tag{4.25}$$

那么存在一个显式的 $\alpha_1 \in (0, \alpha_0]$，使得对于所有 $\alpha \in (0, \alpha_1]$，函数 $\mathcal{V}_\alpha$ 对于 $\mathcal{L}_\alpha$ 是$(\lambda_\alpha, A'_\alpha)$-可容许的，其漂移率为

$$\lambda_\alpha \;\ge\; \lambda + \delta\,\alpha - C_\lambda\,\alpha^2. \tag{4.26}$$

事实上，α_1 可以显式地选为

$$\alpha_1 := \min\{\alpha_0,\ 1,\ \alpha_*,\ \alpha_{\mathrm{pos}}\}, \tag{4.27}$$

其中 α_* 由 (4.18) 给出，且 $\alpha_{\mathrm{pos}} > 0$ 是满足以下条件的任意常数：$\lambda + \delta\alpha - C_\lambda\alpha^2 \ge \lambda/2$ 对所有 $\alpha \in (0, \alpha_{\mathrm{pos}}]$均成立，例如，可以取：

$$\alpha_{\mathrm{pos}} := \begin{cases} \min\left\{1,\ \frac{\delta + \sqrt{\delta^2 + 2C_\lambda\lambda}}{2C_\lambda}\right\}, & C_\lambda > 0, \\ \min\left\{1,\ \frac{\lambda}{2\max\{1, |\delta|\}}\right\}, & C_\lambda = 0. \end{cases} \tag{4.28}$$

证明。 证明见附录 C.5。 □

4.2 收缩率的加速

现在，我们将命题 4.7 与主定理 3.8 结合起来，得到 HFHR 动力学的加速收缩率。首先，我们回顾一下定理 3.8给出的收缩率

$$c(\lambda) = \frac{\gamma}{384}\min\left\{\widetilde{\Lambda}_{1,\alpha}(\lambda),\ \widetilde{\Lambda}_{2,\alpha}(\lambda),\ \widetilde{\Lambda}_{3,\alpha}(\lambda)\right\}.$$

~ ~ ~第一项 $\Lambda_{1,\ \alpha}$ 在（3.33）中对应李雅普诺夫分支，而 $\Lambda_{2,\ \alpha}$ 和 $\Lambda_{3,\ \alpha}$ 在（3.33）中则是度量分支。令 c_0 ：$= c$（λ）表示动能朗之万动力学的收缩率

在 $\alpha = 0$处，其中 λ 定义见（2.12）。根据 [EGZ19，式（2.18）]，，我们假设在 $\alpha = 0$ 处李雅普诺夫分支处于激活状态，即

$$c_0 = \frac{\gamma}{384}\,\widetilde{\Lambda}_{1,0}(\lambda), \qquad \widetilde{\Lambda}_{1,0}(\lambda) \le \widetilde{\Lambda}_{2,0}(\lambda), \quad \widetilde{\Lambda}_{1,0}(\lambda) \le \widetilde{\Lambda}_{3,0}(\lambda).$$

令 $c_\alpha := c(\lambda_\alpha)$ 表示采用改进的李雅普诺夫函数 $\mathcal{V}_\alpha$时 HFHR 动力学的收缩率，其中 λ_α 定义见（4.26）。

由于全局收缩率 $c(\lambda)$ 是李雅普诺夫分支 $\widetilde{\Lambda}_{1,\alpha}$ 和 $\widetilde{\Lambda}_{2,\alpha}$、$\widetilde{\Lambda}_{3,\alpha}$ 度量分支中的最小值，因此我们分别分析改进漂移 λ_α 对这些分支的影响。特别是，李雅普诺夫分支会随 λ_α 直接改善。对于度量分支，只有满足额外的定量条件时才会有所改善；我们将在下文验证，当 α 足够小时，这些条件成立。

4.2.1 度量分支上的加速

现在我们研究李雅普诺夫速率提升至 λ_α 时度量分支的表现。回顾（3.4），对于 HFHR 动力学，有效利普希茨常数为 $L_{\text{eff}}(\alpha) = (1+\alpha\gamma)L$。为匹配 [EGZ19]，中的尺度关系，我们将几何常数固定为未受扰动情形 $\alpha = 0$。特别地，由于 $L_{\text{eff}}(0) = L$，我们设 $\theta_0 := L\gamma^{-2}$ 并定义

$$J_2 := \frac{12}{5}\left(1 + 2\theta_0 + 2\theta_0^2\right)\frac{d + A}{\gamma^2(1 - 2\lambda)}. \tag{4.29}$$

我们重点关注度量参数的主导项。下述定理表明，若$\delta > \gamma\lambda$，则度量分支上的收缩率会严格提高。

定理 4.8 **(度量分支加速)。** 假设命题 *4.7* 的条件成立。设λ_α 为命题 *4.7*给出的漂移率，即对所有 $\alpha \in (0,\ \alpha_1]$,

$$\lambda_\alpha \ \ge\ \underline{\lambda}_\alpha := \lambda + \delta\,\alpha - C_\lambda\,\alpha^2,$$

其中 δ 和 C_λ 是式 (4.23)–(4.24)中的显式常数。此外，假设

$$D := \delta - \gamma\lambda > 0,$$

其中 λ是假设 *2.1**(iii)* 中的耗散常数。定义适用于 $\alpha \ge 0$ 的（与维数无关的）度量参数函数如下：

$$\Lambda_\alpha(\lambda) := J_2\,\frac{(1+\alpha\gamma)L}{\lambda}. \tag{4.30}$$

令 $\Lambda_0 := \Lambda_0(\lambda) = J_2 L/\lambda$ 并假设 $\Lambda_0 > \frac{1}{2}$。定义

$$h(\Lambda) := \sqrt{\Lambda}\,e^{-\Lambda}, \tag{4.31}$$

$$M_h := \sup_{\Lambda\in[\Lambda_0/2,\Lambda_0]}\left|\frac{\Lambda^2 - \Lambda - 1/4}{\Lambda^{3/2}}e^{-\Lambda}\right|. \tag{4.32}$$

那么存在一个显式常数

$$\alpha_{\text{metric,acc}} := \min\left\{\alpha_1,\ 1,\ \frac{D}{4C_\lambda},\ \sqrt{\frac{\lambda}{2C_\lambda}},\ \frac{4\lambda}{D},\ \frac{8\lambda^2\sqrt{\Lambda_0}e^{-\Lambda_0}\left(1-\frac{1}{2\Lambda_0}\right)}{J_2 L\, D\, M_h}\right\}, \tag{4.33}$$

约定当 $C_\lambda = 0$ 时，省略涉及 C_λ 的项，因此对所有 $\alpha \in (0, \alpha_{\text{metric,acc}}]$，以下结论成立：

(i) 度量参数递减。

$$\Lambda_\alpha(\lambda_\alpha) \le \Lambda_0 - c_\Lambda\,\alpha, \qquad c_\Lambda := \frac{1}{8}\,J_2 L\,\frac{D}{\lambda^2} > 0.$$

(ii) 度量分支递增。 存在显式常数 $c_2, c_3 > 0$，使得

$$\widetilde{\Lambda}_{2,\alpha}(\lambda_\alpha) \ge \widetilde{\Lambda}_{2,0}(\lambda)\,(1+c_2\alpha), \qquad \widetilde{\Lambda}_{3,\alpha}(\lambda_\alpha) \ge \widetilde{\Lambda}_{3,0}(\lambda)\,(1+c_3\alpha),$$

其中

$$\widetilde{\Lambda}_{2,\alpha}(\lambda) := h(\Lambda_\alpha(\lambda))\,\frac{L_{\text{eff}}(\alpha)}{\gamma^2}, \qquad \widetilde{\Lambda}_{3,\alpha}(\lambda) := \kappa_{\text{adjust}}\,h(\Lambda_\alpha(\lambda)),$$

并且 $L_{\text{eff}}(\alpha) = (1+\alpha\gamma)L$。更准确地说，可以取任意满足以下条件的 $c_3 < c_3^*$

$$c_3^* := \frac{1}{2}\left(1-\frac{1}{2\Lambda_0}\right)c_\Lambda,$$

然后令 $c_2 := \gamma + c_3$。

证明。 证明见附录 C.6。 □

注 4.9. 定理 4.8 中出现的抽象条件 $\delta > \gamma\lambda$ 纯粹是定量性的：它在考虑有效利普希茨常数 $L_{\text{eff}}(\alpha) = (1+\gamma\alpha)L$ 随度量参数 $\Lambda_\alpha(\lambda) = J_2 L_{\text{eff}}(\alpha)/\lambda$ 而增加的 $\mathcal{O}(\alpha)$ 后，将一阶漂移改进的强度与基准李雅普诺夫速率进行比较。我们在第 5 节中针对多阱势函数（在适当选择 γ 和 λ 的情况下，见第 5.1 节）、带有L^p 正则化项的贝叶斯线性回归（第 5.2 节）以及贝叶斯二分类（第 5.3 节），明确验证了这一条件。这表明，对于这些例子，HFHR 动力学不仅改进了李雅普诺夫分支，也改进了支配跨越能垒的度量分支。

4.2.2 李雅普诺夫分支上的加速

为证明加速，我们首先验证，当 α 较小时，收缩率 c_α 仍由~李雅普诺夫分支决定，即在 $c(\lambda_\alpha)$ 的定义中，最小值仍在 Λ_1 处取得，$\alpha(\lambda_\alpha)$。在这种情况下， λ_α 的增益直接转化为收敛速度的提升。事实上，~根据（3.32）中的定义以及Λ_1 在α （3.33）中的表达式，我们有

$$c_\alpha = \frac{\gamma}{384}\,\widetilde{\Lambda}_{1,\alpha}(\lambda_\alpha) = \frac{\gamma}{384}\,\frac{\lambda_\alpha L_{\text{eff}}(\alpha)}{\gamma^2}.$$

为证明李雅普诺夫分支上的加速，我们首先需要确保收敛瓶颈仍由无穷远处的漂移项决定，而不会转而由度量耦合机制决定。以下引理保证了当扰动参数足够小时，这种稳定性得以保持。

引理 4.10 (活跃分支的连续性)。 设 $\alpha_1 > 0$ 如命题 4.7所述。假设在 $\alpha = 0$ 李雅普诺夫分支严格占优:

$$\widetilde{\Lambda}_{1,0}(\lambda) < \widetilde{\Lambda}_{2,0}(\lambda), \qquad \widetilde{\Lambda}_{1,0}(\lambda) < \widetilde{\Lambda}_{3,0}(\lambda).$$

其中 λ 是 $\alpha = 0$ 处的基准漂移率。进一步假设 $\alpha \mapsto \Lambda_{i,\alpha}(\lambda_\alpha)$ 在$[0, \alpha_1]$ 上对 $i = 1, 2,3$连续。定义

$$\Delta(\alpha) := \min\left\{\widetilde{\Lambda}_{2,\alpha}(\lambda_\alpha), \widetilde{\Lambda}_{3,\alpha}(\lambda_\alpha)\right\} - \widetilde{\Lambda}_{1,\alpha}(\lambda_\alpha).$$

并令

$$\alpha_{\text{branch}} := \sup\{\alpha \in (0, \alpha_1] : \ \Delta(\beta) > 0 \ \textit{for all} \ \beta \in [0, \alpha]\}. \tag{4.34}$$

那么 $\alpha_{\text{branch}} > 0$ ，且对所有 $\alpha \in (0, \alpha_{\text{branch}}]$ ，李雅普诺夫分支始终处于激活状态，即

$$c_\alpha = c(\lambda_\alpha) = \frac{\gamma}{384}\,\widetilde{\Lambda}_{1,\alpha}(\lambda_\alpha).$$

证明。 证明见附录 C.7。 □

注 4.11。 李雅普诺夫分支 $\Lambda_{1,\alpha}$ 在 $\alpha = 0$ 处处于激活状态这一假设，一般而言并非自动成立，应视为动力学的一项结构条件。它描述的是这样一种情形：全局收敛速率确实由无穷远处的漂移（由 Lyapunov 参数 λ刻画）所控制，而局部收缩机制（由 $\Lambda_{2,\alpha}$,$\Lambda_{3,\alpha}$刻画）相对较快。特别是，对于具有显著能垒的强亚稳态目标（例如经典双阱势）， [EGZ19]中的显式界表明，收缩率往往由与跨越能垒相关的度量分支主导，而非李雅普诺夫分支。我们的引理表明，只要 李雅普诺夫分支在 $\alpha = 0$处严格处于激活状态，它就会在所有充分小的 $\alpha > 0$下保持激活。

既然活跃分支的连续性已得到确立，我们现在就可以将改进的李雅普诺夫漂移项 λ_α 直接转化为收敛速率的定量加速。

定理 4.12 （Lyapunov 分支加速）。 设 $c_0 = c(\lambda)$ 为定理 $\alpha = 0$ *3.8* 给出的在 处的动能朗之万动力学收缩率。进一步假设引理 *4.10* 的假设成立（特别是，李雅普诺夫分支在 严格 活跃），并令 $\alpha = 0$，且$\alpha_{\text{branch}} \in (0, \alpha_1]$ 为 (4.34)中定义的阈值。

令 $c_\alpha := c(\lambda_\alpha)$ 表示通过对 应用定理 $\mathcal{V} = \mathcal{V}_\alpha$*3.8* 得到的 *HFHR* 动力学收缩率。则存在

$$\alpha_{\text{branch,acc}} := \min\left\{\alpha_{\text{branch}}, 1, \frac{\kappa}{C'}\right\} \in (0, \alpha_{\text{branch}}],$$

其中

$$C' = \frac{L}{384\gamma}(1+\gamma)\left(C_2 + \widetilde{C}_{\mathcal{M}}\,\underline{c}_{\text{imp}}\right), \tag{4.35}$$

其中 C_2 在 (4.21) 中定义， $C_{\mathcal{M}}$ 在 (4.22) 中定义，而 $\underline{c}_{\text{imp}}$ 在 (4.9) 中定义，使得对所有 $\alpha \in (0, \alpha_{\text{branch, acc}}]$,

$$c_\alpha \ \geq \ c_0 + \kappa\,\alpha, \tag{4.36}$$

其中

$$\kappa := \frac{L(\delta + \gamma\lambda)}{768\,\gamma}, \tag{4.37}$$

以及 $\delta > 0$ 以及 $C_\lambda \geq 0$ 是命题中的常数 4.7。

证明。 我们在附录 C.8 中给出证明。 □

4.2.3 全局加速

基于度量分支（定理 4.8）和李雅普诺夫分支（定理 4.12）的加速，我们得出以下全局加速结果。

推论 4.13 (HFHR 动力学的全局加速)。假设命题 *4.7* 的条件成立，并且

$$\delta > \gamma\lambda \qquad and \qquad \Lambda_0 := \Lambda(\lambda) = \frac{J_2 L}{\lambda} > \frac{1}{2},$$

其中 λ 是假设*2.1(iii)* 中的耗散常数，因此定理 *4.8* 适用。令 α_1如命题 *4.7*中所述，$\alpha_{\text{branch, acc}} \in (0,\ \alpha_1]$ 为定理 *4.12*中的阈值，并令 $\alpha_{\text{metric, acc}} \in (0,\ \alpha_1]$ 为定理 *4.8*中的阈值。定义

$$\alpha_{\text{global}} := \min\{\alpha_{\text{branch,acc}}, \alpha_{\text{metric,acc}}\}.$$

于是，对所有 $\alpha \in (0, \alpha_{\text{global}}]$，*HFHR* 动力学的收缩率严格优于动能朗之万动力学，即

$$c_\alpha \ \geq \ c_0 + \kappa_{\text{global}}\,\alpha,$$

其中

$$\kappa_{\text{global}} := \min\{\kappa,\ c_0 c_2,\ c_0 c_3\} \ > \ 0.$$

这里 $c_0 = c(\lambda)$ 是 $\alpha = 0$处的收缩率，κ 是定理 *4.12*中的 *Lyapunov* 分支加速常数，而 $c_2,\ c_3 > 0$ 是定理 *4.8*中的度量分支改进常数。

证明。 证明见附录 C.9。 □

推论 4.13 表明，对于所有足够小的 $\alpha > 0$，HFHR 动力学在加权 Wasserstein 距离 $\mathcal{W}_{\rho_{\mathcal{V}_\alpha}}$下的收缩率严格优于动能朗之万动力学。最后，我们证明，这种加速并非加权 Wasserstein 距离造成的假象，而是直接体现为标准 2-Wasserstein 距离 $\mathcal{W}_2$下的加速。

推论 4.14 (2-Wasserstein 距离下的加速)。 假设推论 *4.13*中的设定成立，并进一步将定理 *3.8* 中的漂移参数选为

$$\lambda_\alpha := \underline{\lambda}_\alpha = \lambda + \delta\alpha - C_\lambda \alpha^2 \qquad (\alpha > 0),$$

其中 λ 是假设 *2.1(iii)* 中的耗散常数，δ、C_λ 来自命题 *4.7*。令 c_0和 c_α 分别为推论 $\mathcal{W}_{\rho_{\mathcal{V}_\alpha}}$ *4.13*给出的收缩率，并设

$$\alpha_{\text{W2}} := \min\{\alpha_{\text{global}},\ \alpha_{\text{pos}}\},$$

其中 α_{global} 是推论 *4.13* 中的阈值，α_{pos} 是 (4.28) 中的显式常数（因此$\lambda_\alpha \geq \lambda/2$ 对所有 $\alpha \in (0,\ \alpha_{\text{pos}}]$）。

定义区间

$$I_\lambda := [\lambda_-, \lambda_+] \quad with \quad \lambda_- := \frac{\lambda}{2}, \qquad \lambda_+ := \lambda + \delta\,\alpha_{\text{W2}}.$$

对于每个 $\lambda \in I_\lambda$，令 $R_1(\lambda)$ 为满足 *(3.28)* 的任意可容许截断半径（其中 $\eta_0(\lambda)$、$= (\Lambda(\lambda))^{-1}$ 和 f_λ 取定理 *3.8* 中对应的取值），并令 φ_λ、Φ_λ、g_λ、f_λ 按照 *(3.9)– (3.12)* 定义，其中使用此 $R_1(\lambda)$。定义显式极值常数

$$R_1(\lambda)^+ := \sup_{\lambda\in I_\lambda} R_1(\lambda), \qquad c_r^- := \inf_{\lambda\in I_\lambda} \inf_{0\le s\le R_1(\lambda)} \varphi_\lambda(s), \qquad g_*^- := \inf_{\lambda\in I_\lambda} \inf_{0\le s\le R_1(\lambda)} g_\lambda(s),$$

and

$$c_0^- := \inf_{\lambda\in I_\lambda} f_\lambda(R_1(\lambda)) = \inf_{\lambda\in I_\lambda} \int_0^{R_1(\lambda)} \varphi_\lambda(s) g_\lambda(s)\, ds.$$

假设 $g_*^- > 0$ （例如，通过(3.32)中的显式参数选择即可确保这一点；参见定理(3.39) 证明中对*3.8*的验证）。

此外，根据引理*4.5* 以及 $\mathcal{V}_0$的二次增长，固定显式常数$C_V^{\mathrm{unif}} > 0$ 和$C \ge 0$，使得

$$|z|^2 \le C_V^{\mathrm{unif}} (1 + \mathcal{V}_\alpha(z)) + C, \qquad z \in \mathbb{R}^{2d},\ \alpha \in (0, \alpha_{\mathrm{W2}}].$$

令 $k_1(\lambda)$ 为引理*3.3*中与 $\theta(\lambda)$ 在(3.5)中的选择相对应的范数等价常数（其中 $\eta_0(\lambda)$），并定义

$$k_1^- := \inf_{\lambda\in I_\lambda} k_1(\lambda) > 0.$$

最后，令$\varepsilon_\alpha = \gamma \frac{4c_\alpha}{(d+A'_\alpha)}$ 为定理*3.8* 中的参数（应用时取 $\mathcal{V} = \mathcal{V}_\alpha$ 以及$(\lambda, D) = (\lambda_\alpha, A'_\alpha)$），并定义显式下界

$$\varepsilon^- := \inf_{\alpha\in(0,\alpha_{\mathrm{W2}}]} \varepsilon_\alpha > 0.$$

（例如，可以取$\varepsilon^- = \gamma \frac{4c_0}{(d+A^+)}$，~~其中~~ $A^+ :=$ 上确界$_{\alpha\in(0,\alpha_{\mathrm{W2}}]} A'_\alpha < \infty$，使用 $c_\alpha \ge c_0$ 推论*4.13*中的结果。）令

$$C_\rho^{\mathrm{unif}} := \frac{1}{\varepsilon^-} \max\left\{ \frac{(k_1^-)^{-2} R_1(\lambda)^+}{g_*^- c_r^-}, \frac{4C_V^{\mathrm{unif}}}{c_0^-} \right\}. \tag{4.38}$$

那么对于所有 $\alpha\in(0, \alpha_{\mathrm{W2}}]$、所有 $t \ge 0$，以及所有概率测度 μ，ν具有有限 $\mathcal{V}_\alpha$ 阶矩，

$$\mathcal{W}_2(\mu P_t^\alpha, \nu P_t^\alpha) \le \left(C_\rho^{\mathrm{unif}}\right)^{1/2} e^{-c_\alpha^{(2)} t} \left(\mathcal{W}_{\rho\mathcal{V}_\alpha}(\mu,\nu)\right)^{1/2}, \tag{4.39}$$

其中 $c_\alpha^{(2)} := \frac{1}{2} c_\alpha$。此外，在 $\mathcal{W}_2$ 中可实现加速，并具有显式增益：

$$c_\alpha^{(2)} \ge c_0^{(2)} + \kappa^{(2)}\, \alpha, \qquad \alpha \in (0, \alpha_{\mathrm{W2}}],$$

其中 $c_0^{(2)} := \frac{1}{2} c_0$，以及 $\kappa^{(2)} := \frac{1}{2}\kappa_{\mathrm{global}}$。

证明。 证明见附录 C.10。 □

5 案例研究

本节通过三个具体的非凸示例来说明我们的普遍结果：多阱势（第 5.1 节）、带有 L^p 正则化项的贝叶斯线性回归（第 5.2 节），以及贝叶斯二分类（第 5.3 节）。我们验证这些示例均满足假设 2.1，回顾动能朗之万动力学的基准收缩估计，然后为这些具体示例构造显式的二次校正项 $\mathcal{M}$，并由此得到一个改进常数，该常数可能大于抽象理论给出的通用下界。最后，我们证明，对于所有足够小的 $\alpha > 0$，HFHR 动力学所达到的收缩率严格优于动能朗之万动力学。

5.1 多阱势

本节研究一个高维非凸势函数的例子。具体而言，我们考虑一个 d维的多阱势函数 $U: \mathbb{R}^d \to \mathbb{R}$ 它由若干相互独立的一维双阱势函数求和构造而成 [EGZ19，例 1。1]。设 $z = (q, p) \in \mathbb{R}^{2d}$。定义

$$U(q) := \sum_{i=1}^{d} v(q_i), \tag{5.1}$$

其中 $v: \mathbb{R} \to \mathbb{R}$ 是按分量定义的势函数，形式如下：

$$v(s) := \begin{cases} \frac{1}{2}(|s| - 1)^2, & |s| \geq \frac{1}{2}, \\ \frac{1}{4} - \frac{1}{2}s^2, & |s| \leq \frac{1}{2}. \end{cases} \tag{5.2}$$

势函数 U具有 2^d位于 $(\pm 1, \ldots, \pm 1)$ 的局部极小值，并且是高维多模态分布采样的经典基准。下面验证 U 满足我们理论所要求的所有结构假设，即假设 2.1 和4.1。

命题 5.1 (假设验证)。 固定 $\gamma > 0$。势函数 U 定义于 (5.1)，满足假设*2.1* 和假设 *4.1*。具体而言：

(a) 正则性与利普希茨梯度： U 是 $C^1(\mathbb{R}^d)$，且 ∇U 是 L-利普希茨的，其中 $L = 1$。

(b) 耗散性：耗散条件(2.6) 对任意 $\lambda \in \left(0, \frac{1}{4+\gamma^2}\right]$均成立，常数$A$ 随 d线性变化。更准确地说，可以取

$$A := d\, A_1(\gamma), \qquad A_1(\gamma) := \frac{\gamma^4 + 6\gamma^2 + 16}{4(\gamma^4 + 10\gamma^2 + 24)}.$$

特别地，数值 $\lambda = 1/(4 + \gamma^2)$ 只是一个方便的上界；取更小的λ 仍然有效，且 A保持不变。

(c) 渐近线性漂移：对于任意 $|q| \geq \sqrt{d}$,

$$|\nabla U(q) - q| \leq \varrho(|q|)\, |q|, \qquad \varrho(r) := \sqrt{d}/r,$$

因此，对于每个固定的 d， $\varrho(r) \to 0$ 如 $r \to \infty$。

证明。 证明见附录 D.1。 □

由于我们已在命题 5.1 中验证，势函数 U 满足假设2.1 和4.1，因此第3 节和第 4 节的理论结果均适用。不过，由于多阱势 U （5.1）具有非常特殊的结构，我们可以利用多阱势 U （5.1）的特殊尾部结构，构造显式的二次校正项 $\mathcal{M}$ ，从而获得更精确的加速结果。虽然引理 4.3 已保证在抽象的渐近线性漂移条件下校正子的存在，但当前这个可分离的多阱模型允许我们进行更精确的构造：通过匹配 U 在无穷远处的二次部分，我们选择 $\mathcal{M}$，使主导相互作用漂移 $\mathcal{A}'\mathcal{V}_0$ 被抵消（余项一致有界）。由此得到：（i）一个闭式改进常数 c_{imp} ；（ii）关于辅助常数（例如 $C_{\mathcal{M}}$ 和二阶误差系数）的显式维度控制界，而一般性的存在性论证无法给出这些结果。

下面构造的显式校正子与一般理论一致：引理 4.3通过无穷远处二次型的李雅普诺夫方程刻画 $\mathcal{M}$，而我们对$\mathcal{M}$ 的选择，正是针对当前各向同性／可分离情形特化的一个此类二次解。具体而言，它可视为引理 4.3中可容许校正项类的一个具体代表，其选择旨在使处理更简便，并使常数完全显式。

命题 *5.2*（显式一阶改进）。固定 $\gamma > 0$ 和维度 $d \in \mathbb{N}$。令 $\lambda \in (0, 1/4]$ 为假设 *2.1* 中的参数。考虑 *(5.1)–(5.2)* 中的多阱势函数 $U(q) = \sum_{i=1}^{d} v(q_i)$，其中 $L = 1$ 且 $\nabla U(0) = 0$。定义二次校正项 $\mathcal{M} : \mathbb{R}^{2d} \to \mathbb{R}$ 如下

$$\mathcal{M}(q,p) := \frac{2+\gamma^2}{4\gamma}|q|^2 + \frac{1}{2\gamma}|p|^2. \tag{5.3}$$

Let

$$B := 1 + \frac{\gamma^2}{2}(1-\lambda), \qquad c_{\mathrm{imp}} := \frac{2\sqrt{B}}{2\sqrt{B}+\gamma} \in (0,1).$$

则有如下结论。

(i) 一阶改进不等式。存在一个常数 $C^{(d)}{}_{\mathrm{imp}} < \infty$ （其随 d 至多线性增长），使得对所有$(q,p) \in \mathbb{R}^{2d}$,

$$\mathcal{A}_0\mathcal{M}(q,p) + \mathcal{A}'\mathcal{V}_0(q,p) \le -c_{\mathrm{imp}}\,\mathcal{V}_0(q,p) + C^{(d)}_{\mathrm{imp}}. \tag{5.4}$$

特别地，对于一个显式常数 $C^{\mathrm{MW}}_{\mathrm{imp}}(\gamma,\lambda)$，可以取 $C^{(d)}{}_{\mathrm{imp}} = C^{\mathrm{MW}}_{\mathrm{imp}}(\gamma,\lambda)\, d$。

(ii) 二次下界与一致增长常数 $\widetilde{C}^{\mathrm{MW}}_{\mathcal{M}}$。定义

$$c_1^{\mathrm{MW}} := \frac{1}{8}\left(\gamma^2(1-\lambda) + 2 - \sqrt{(\gamma^2(1-\lambda)-2)^2 + 4\gamma^2}\right) > 0, \tag{5.5}$$

使得

$$\mathcal{V}_0(q,p) \ge c_1^{\mathrm{MW}}\,(|q|^2 + |p|^2), \qquad (q,p) \in \mathbb{R}^{2d}. \tag{5.6}$$

此外，由

$$\widetilde{C}^{\mathrm{MW}}_{\mathcal{M}} := \frac{2+\gamma^2}{4\gamma\, c_1^{\mathrm{MW}}}, \tag{5.7}$$

我们有逐点界

$$|\mathcal{M}(q,p)| \le \widetilde{C}_{\mathcal{M}}^{\mathrm{MW}}\,\mathcal{V}_0(q,p) \le \widetilde{C}_{\mathcal{M}}^{\mathrm{MW}}\,(1+\mathcal{V}_0(q,p)). \tag{5.8}$$

(iii) 二阶余项与漂移率展开。设

$$\mathrm{Err}^{(d)}(q,p) := |\mathcal{A}'\mathcal{M}(q,p)| + |\Delta_q\mathcal{M}(q,p)|. \tag{5.9}$$

那么存在 e 显式常数 $C_2^{\mathrm{MW}} \ge 0$ （与维数无关）和 $C^{(d)}{}_2{}^{\mathrm{MW}} = \mathcal{O}(d)$ 使得 t

$$\mathrm{Err}^{(d)}(q,p) \le C_2^{\mathrm{MW}}\,\mathcal{V}_0(q,p) + C_2^{(d),\mathrm{MW}}, \qquad (q,p)\in\mathbb{R}^{2d}. \tag{5.10}$$

特别地，可以取

$$C_2^{\mathrm{MW}} := 2\widetilde{C}_{\mathcal{M}}^{\mathrm{MW}}, \qquad C_2^{(d),\mathrm{MW}} := \frac{2+\gamma^2}{2\gamma}\,d.$$

最后，定义

$$\delta_{\mathrm{MW}} := c_{\mathrm{imp}} - \lambda\,\widetilde{C}_{\mathcal{M}}^{\mathrm{MW}}, \qquad C_{\lambda,\mathrm{MW}} := C_2^{\mathrm{MW}} + \widetilde{C}_{\mathcal{M}}^{\mathrm{MW}}\,c_{\mathrm{imp}}. \tag{5.11}$$

那么命题 4.7 适用（若基准常数需要，取 λ充分小），且改进的漂移率满足

$$\lambda_\alpha \;\ge\; \lambda + \delta_{\mathrm{MW}}\cdot\alpha - C_{\lambda,\mathrm{MW}}\cdot\alpha^2. \tag{5.12}$$

特别地，当 $\delta_{\mathrm{MW}} > 0$ 时，对所有充分小的 $\alpha > 0$， $\lambda_\alpha > \lambda$ 。

证明。 证明见附录 D.2。 □

要将推论 4.13 的定量加速结果应用于多阱模型，需要满足一维条件 $\delta_{\mathrm{MW}} > \gamma\lambda$。下一个引理表明，该条件并不严格：对于任意固定的 $\gamma > 0$，都可以通过将耗散参数 λ选得足够小来满足该条件。

引理 *5.3*（定量条件 $\delta_{\mathrm{MW}} > \gamma\lambda$ 的可行性）。固定 $\gamma > 0$，并考虑一维双阱势函数 $v : \mathbb{R}\to\mathbb{R}$，其定义见 *(5.2)*。那么，对于任意 $\lambda \in (0, 1/4]$，假设 *2.1(iii)*（耗散性）均成立（加性常数依赖于 λ）。此外，令 δ_{MW} 如 *(5.11)* 中所定义，则存在 $\lambda_\star(\gamma) \in (0, 1/4]$，使得对于每个 $\lambda \in (0, \lambda_\star(\gamma)]$，我们都有

$$\delta_{\mathrm{MW}} > \gamma\lambda.$$

证明。 证明见附录 D.3。 □

根据命题 $\mathcal{M}$中二次校正项的显式构造5.2，我们给出关于 d 维多阱势函数的主要加速结果。在继续之前，令 $\rho_{\alpha,\ 1}$为推论 4.13中使用的一维代价，并定义 $\mathbb{R}^{2d}$ 上的张量化代价如下：

$$\rho_{\alpha,d}(z,z') := \sum_{i=1}^{d}\rho_{\alpha,1}(z_i,z_i'), \qquad z_i = (q_i,p_i)\in\mathbb{R}^2.$$

令 $\mathcal{W}_{\rho_{\alpha,d}}$ 为由 $\rho_{\alpha,d}$诱导的 Wasserstein 距离。那么，我们有如下结果。

定理 *5.4*（多阱势的 *HFHR* 加速）。考虑 *(5.1)* 中的可分势$U(q) = \sum_{i=1}^{d} v(q_i)$，并令 $P^{\alpha,(d)}_t$ 为相应 d 维 *HFHR* 动力学的半群。固定 $\gamma > 0$，并选择 $\lambda \in (0, \lambda_\star(\gamma)]$，如引理 *5.3* 所示，使得 $\delta_{\mathrm{MW}} > \gamma\lambda$。此外，假设多阱模型满足推论 *4.13* 中其余的一维定量条件（特别是 $\Lambda_0 > 1/2$）。那么，存在显式常数 $\alpha_{\mathrm{MW}} > 0$和 $\kappa_{\mathrm{MW}} > 0$（与 d 无关），其仅取决于一维双井模型（以及 γ），使得对于任意 $d \geq 1$ 和任意 $\alpha \in (0, \alpha_{\mathrm{MW}}]$,

$$\mathcal{W}_{\rho_{\alpha,d}}\left(\mu P_t^{\alpha,(d)}, \nu P_t^{\alpha,(d)}\right) \leq e^{-(c_0+\kappa_{\mathrm{MW}}\alpha)t}\, \mathcal{W}_{\rho_{\alpha,d}}(\mu,\nu), \qquad t \geq 0,$$

对于 $\mathbb{R}^{2d}$ 上的所有概率测度 μ、ν，其中 c_0 表示在 $\alpha = 0$ 处、与代价函数 $\rho_{0,\,1}$相关联的一维动能朗之万收缩率。此外，可以显式地选取

$$\alpha_{\mathrm{MW}} := \min\left\{\alpha^{(1)}_{\mathrm{branch,acc}},\ \alpha^{(1)}_{\mathrm{metric,acc}}\right\},$$

其中 $\alpha^{(1)}$ 分支，加速度 和 $\alpha^{(1)}$ 度量，加速度 分别是定理 *4.12* 和定理 *4.8* 中给出的显式阈值，在一维模型下取值（使用 $L = 1$、$\delta = \delta_{\mathrm{MW}}$ 和 $C_\lambda = C_{\lambda,\ \mathrm{MW}}$）。类似地，显式增益为

$$\kappa_{\mathrm{MW}} := \kappa^{(1)}_{\mathrm{global}} = \min\left\{\kappa^{(1)},\ c_0\, c_2^{(1)},\ c_0\, c_3^{(1)}\right\},$$

其中 $\kappa^{(1)}$ 是定理 *4.12* 在维度 1下的 *Lyapunov* 分支增益，且 $c_2^{(1)}$、$c_3^{(1)}$ 是定理 *4.8* 在维度 1下的度量分支改进常数。

特别是，使用命题 *5.2*中的显式校正子，可以取

$$\kappa^{(1)} = \frac{L(\delta_{\mathrm{MW}} + \gamma\lambda)}{768\,\gamma},$$

（多阱模型中取 $L = 1$），其中

$$\delta_{\mathrm{MW}} := c_{\mathrm{imp}} - \lambda\, \widetilde{C}^{\mathrm{MW}}_{\mathcal{M}}, \qquad C_{\lambda,\mathrm{MW}} := C_2^{\mathrm{MW}} + \widetilde{C}^{\mathrm{MW}}_{\mathcal{M}}\, c_{\mathrm{imp}}.$$

证明。 证明见附录 D.4。 □

5.2 贝叶斯线性回归

本节研究贝叶斯线性回归问题的示例。给定输入数据$X \in \mathbb{R}^{n\times d}$和输出数据 $y \in \mathbb{R}^n$，我们考虑以下目标函数 $U:\ \mathbb{R}^d \to \mathbb{R}$其中正则化函数为 $g:\ \mathbb{R}^d \to \mathbb{R}$ 在贝叶斯线性回归任务 [Hof09]:

$$U(q) = \frac{1}{2\sigma^2}|y - Xq|^2 + g(q), \tag{5.13}$$

使得 $\nabla U(q) = -\frac{X^\top(y-Xq)}{\sigma^2} + \nabla g(q)$，其中参数 $\sigma > 0$。特别地，我们考虑L^p 正则化。使用平滑的 L^p 正则化，可以从贝叶斯视角将其解释为对回归系数施加先验。当 $p < 2$时，此类先验在高斯分布与拉普拉斯分布之间插值；参见 [PS10、GZL+13]，以了解全面综述。

此外， [PC08、PSW14] 中的贝叶斯套索是对应于 $p \to 1$ 的一个特例。我们将正则化函数 g 取为如下的 L^p 函数：

$$g(q) := \iota \sum_{i=1}^{d} (q_i^2 + \varepsilon^2)^{p/2}, \qquad 1 < p < 2, \tag{5.14}$$

其中， $\iota > 0$ 是正则化参数， $\varepsilon^2 > 0$ 是自调节参数。由于 $1 < p < 2$，正则化项 $g(q)$ 以及势函数 $U(q)$ 通常是非凸的。我们作出如下假设。

假设 5.5。 假设对于某个 $m > 0$， $X^\top X \succ m I_d$ 。

请注意，假设 5.5较为温和，并且在文献中经常被采用；例如，参见 [MBM18] 中的假设 9。接下来，我们将证明，在假设 5.5 下，带有 L^p 正则化项（5.14）的贝叶斯线性回归问题（5.13）满足我们的理论所要求的假设 2.1 和 4.1。

命题 *5.6*（带平滑 L^p 正则化项的贝叶斯线性回归满足既定假设）。固定 $1 < p < 2$、 $\sigma > 0$、 $\iota > 0$ 和 $\varepsilon > 0$。令 $X \in \mathbb{R}^{n \times d}$、 $y \in \mathbb{R}^n$ ，并按照（*5.13*）定义 U，其中带有 L^p 正则化项 $g(q)$，如（*5.14*）所示。假设假设 *5.5* 成立。记$M := \|X^\top X\|_{\mathrm{op}}$。则：

(a) $U \in C^\infty(\mathbb{R}^d)$ 以及 $U \geq 0$。此外， ∇U 是 L-利普希茨的，且满足 $L := \frac{M}{\sigma^2} + \iota\, p\, \varepsilon^{p-2}$。

(b) U 满足假设 *2.1-(iii)* 意义下的耗散性。特别地，

$$\langle \nabla U(q), q \rangle \geq \frac{m}{2\sigma^2} |q|^2 - \frac{|X^\top y|^2}{2m\sigma^2} \quad \textit{for all } q \in \mathbb{R}^d.$$

(c) U 满足假设 *4.1* ，其中 $Q_\infty := \frac{1}{\sigma^2} X^\top X$，且函数

$$\varrho(r) := \frac{c_0^{\mathrm{LR}}}{r} + c_1^{\mathrm{LR}}\, r^{p-2}, \qquad r \geq C_{\mathrm{linear}},$$

其中例如可取 $C_{\mathrm{linear}} := 1$ 以及

$$c_0^{\mathrm{LR}} := \frac{|X^\top y|}{\sigma^2} + \iota p \sqrt{d}\, \varepsilon^{p-1}, \qquad c_1^{\mathrm{LR}} := \iota p\, d^{\frac{2-p}{2}}. \tag{5.15}$$

证明。 证明见附录 D.5。 □

与可分离多阱模型（其中 $\nabla U(q) - Q_\infty q$ 一致有界）不同，平滑后的 L^p 正则化项会产生一个次线性但无界的余项 $|\nabla U(q) - Q_\infty q| = \mathcal{O}(|q|^{p-1})$，当 $|q| \to \infty$时。因此，结合李雅普诺夫方程校正项的一阶改进引理4.3，并利用针对（ $\rho_\nabla(\cdot)$ 和 $\delta_U(\cdot)$ 的显式界（专用于（5.13）–（5.14））， 即可得到显式的加速常数。

命题 5.7 （引理 4.3在贝叶斯线性回归中的显式常数）。 假设命题 *5.6* 中的设定，并令 λ 为假设 *2.1*中的耗散参数。令 $Q_\infty := \frac{1}{\sigma^2} X^\top X$， $b := \frac{1}{\sigma^2} X^\top y$，以及 $M := \|X^\top X\|_{\mathrm{op}}$。则有以下结论。

(*i*) Q_∞ 的谱界。

$$\lambda_{\min}(Q_\infty) = \frac{1}{\sigma^2}\lambda_{\min}\left(X^\top X\right) \geq \frac{m}{\sigma^2}, \qquad \lambda_{\max}(Q_\infty) = \frac{1}{\sigma^2}\lambda_{\max}\left(X^\top X\right) = \frac{M}{\sigma^2}.$$

(*ii*) 显式尾模量 ρ_∇ 和 δ_U。取 c_0^{LR}， c_1^{LR}如（5.15）所示，对于任意 $R \geq 1$,

$$\rho_\nabla(R) := \sup_{|q|\geq R} \frac{|\nabla U(q) - Q_\infty q|}{|q|} \leq \frac{c_0^{\mathrm{LR}}}{R} + c_1^{\mathrm{LR}}\, R^{p-2}, \tag{5.16}$$

且对于任意 $R \geq 1$,

$$\delta_U(R) := \sup_{|q|\geq R} \frac{\left|U(q) - \frac{1}{2}\langle Q_\infty q, q\rangle\right|}{1+|q|^2} \leq \frac{|b|}{R} + \iota\, d^{1-\frac{p}{2}}\, R^{p-2} + \frac{\iota\, d\, \varepsilon^p + \frac{1}{2\sigma^2}|y|^2}{R^2}. \tag{5.17}$$

(*iii*) 显式可容许截断半径 R_0。令K为引理*4.3* 中的李雅普诺夫方程矩阵（定义如下），并令 $\rho_\star$ 如（4.13）所示。由于 c_0^{LR}， $c_1^{\mathrm{LR}} \geq 0$ 且 $p-2<0$，式（5.16）的右端关于 R递减。因此，取

$$R_0 := \max\left\{1,\ C_{\mathrm{linear}},\ \frac{c_0^{\mathrm{LR}}}{\rho_\star},\ \left(\frac{c_1^{\mathrm{LR}}}{\rho_\star}\right)^{\frac{1}{2-p}}\right\} \tag{5.18}$$

可确保 ρ_∇（R_0）$\leq \rho_\star$，从而（4.14）成立。

(*iv*) 二次　校正子，采用与一般理论相同的李雅普诺夫方程。L　*et*

$$B := \begin{pmatrix} 0 & I_d \\ -Q_\infty & -\gamma I_d \end{pmatrix}$$

为无穷远处的线性化动理朗之万漂移矩阵，并令C_{B_1} 为引理 $\nabla^2 B_1$ 中的显式对称矩阵*4.3*。那么校正子可取为

$$\mathcal{M}(z) = \frac{1}{2} z^\top \mathsf{K} z, \qquad z = (q,p) \in \mathbb{R}^{2d},$$

其中K是方程的（唯一）对称解

$$B^\top \mathsf{K} + \mathsf{K} B = C_{B_1},$$

等价地，可由积分表示给出

$$\mathsf{K} = \int_0^\infty e^{tB^\top}\, C_{B_1}\, e^{tB}\, dt. \tag{5.19}$$

(*v*) 一阶改进常数 c_{imp} 以及 C_{imp} 的显式上界。设

$$a_{\min} := \lambda_{\min}(Q_\infty) + \frac{\gamma^2}{2}(1-\lambda), \qquad a_{\max} := \lambda_{\max}(Q_\infty) + \frac{\gamma^2}{2}(1-\lambda),$$

因此由 *(i)* 可得，

$$a_{\min} \geq \frac{m}{\sigma^2} + \frac{\gamma^2}{2}(1-\lambda), \qquad a_{\max} \leq \frac{M}{\sigma^2} + \frac{\gamma^2}{2}(1-\lambda).$$

因此，由引理 *4.3* 可得改进不等式 *(4.8)*，其中 c_{imp} 按照 *(4.9)* 选取；$\delta_U(R_0)$ 可由 *(5.17)* 显式界定（其中 $R = R_0$ 取自 *(5.18)*）。特别地，可得到一个完全显式的正下界 $c_{\mathrm{imp}} > 0$，其取决于 $(m, M, \sigma, \lambda, p, \varepsilon, |y|, |X^\top y|, \gamma)$。此外，引理 *4.3* 中相应的常数 C_{imp} 有限。

(vi) 漂移率展开的一个方便的显式余项界。将 K 写成分块形式 $\mathsf{K} = \begin{pmatrix} \mathsf{K}_{qq} & \mathsf{K}_{qp} \\ \mathsf{K}_{pq} & \mathsf{K}_{pp} \end{pmatrix}$ KK，并令

$$k_q := \|\mathsf{K}_{qq}\|_{\mathrm{op}} + \|\mathsf{K}_{qp}\|_{\mathrm{op}}, \qquad b_0 := |\nabla U(0)| = |b|.$$

利用 $|\nabla U(q)| \leq L|q| + b_0$（其中 L 取自命题 5.6*(a)*），可得对所有 $(q,p) \in \mathbb{R}^{2d}$，

$$|\mathcal{A}'\mathcal{M}(q,p)| \leq k_q \left(\frac{3L+1}{2}|q|^2 + \frac{L+1}{2}|p|^2 + b_0^2 \right), \qquad |\Delta_q \mathcal{M}(q,p)| = \mathrm{tr}(\mathsf{K}_{qq}) \leq d\,\|\mathsf{K}_{qq}\|_{\mathrm{op}}.$$

令 $c_1 > 0$ 为满足 $\mathcal{V}_0(q,p) \geq c_1(|q|^2 + |p|^2)$的（二次）下界常数；必要时，可通过加上加性常数来平移 U。那么，“误差项” $\mathrm{Err}^{(d)}(q,p) := |\mathcal{A}'\mathcal{M}(q,p)| + |\Delta_q \mathcal{M}(q,p)|$ 满足

$$\mathrm{Err}^{(d)}(q,p) \leq C_2^{\mathrm{LR}}\ \mathcal{V}_0(q,p) + C_2^{(d),\mathrm{LR}}, \tag{5.20}$$

采用明确的选择

$$C_2^{\mathrm{LR}} := \frac{k_q}{c_1} \max\left\{ \frac{3L+1}{2}, \frac{L+1}{2} \right\}, \qquad C_2^{(d),\mathrm{LR}} := k_q b_0^2 + d\,\|\mathsf{K}_{qq}\|_{\mathrm{op}}.$$

最后，定义（显式）增长常数

$$\widetilde{C}_{\mathcal{M}}^{\mathrm{LR}} := \frac{\|\mathsf{K}\|_{\mathrm{op}}}{2c_1}$$

（参见 （4.22））， 并令

$$\delta_{\mathrm{LR}} := \underline{c}_{\mathrm{imp}} - \lambda\,\widetilde{C}_{\mathcal{M}}^{\mathrm{LR}}, \qquad C_{\lambda,\mathrm{LR}} := C_2^{\mathrm{LR}} + \widetilde{C}_{\mathcal{M}}^{\mathrm{LR}}\,\underline{c}_{\mathrm{imp}}. \tag{5.21}$$

那么 *Pro* 位置 4.7 成立（当 α 充分小时），并给出漂移率展开

$$\lambda_\alpha \ \geq\ \lambda + \delta_{\mathrm{LR}} \cdot \alpha - C_{\lambda,\mathrm{LR}} \cdot \alpha^2.$$

特别地，对于所有充分小的 $\alpha > 0$，只要 $\delta_{\mathrm{LR}} > 0$，就有 $\lambda_\alpha > \lambda$。

证明。 证明见附录 D.6。 □

最后，我们确认，通过将耗散参数 λ 选得足够小，可以满足加速所需的定量条件 $\delta_{\mathrm{LR}} > \gamma\lambda$。

引理 5.8 （$\delta_{\mathrm{LR}} > \gamma\lambda$ 的可行性，具有显式的 $\lambda_\star(\gamma)$）。 固定 $\gamma > 0$，并考虑假设 (5.13)–(5.14) 下的贝叶斯线性回归势函数 5.5。令

$$Q_\infty := \frac{1}{\sigma^2} X^\top X, \quad m_\infty := \lambda_{\min}(Q_\infty) = \frac{\lambda_{\min}(X^\top X)}{\sigma^2} \geq \frac{m}{\sigma^2}, \quad M_\infty := \lambda_{\max}(Q_\infty) = \frac{\lambda_{\max}(X^\top X)}{\sigma^2}.$$

Set

$$\bar{\lambda} := \min\left\{\frac{1}{4}, \frac{m}{2\sigma^2}\right\}, \qquad R := \max\{1, C_{\mathrm{linear}}\}.$$

定义显式尾界

$$\delta_U^+ := \frac{|X^\top y|}{\sigma^2 R} + \iota\, d^{1-\frac{p}{2}} R^{p-2} + \frac{\iota\, d\, \varepsilon^p + \frac{1}{2\sigma^2}|y|^2}{R^2}, \tag{5.22}$$

以及谱代理量

$$a_{\min}^- := m_\infty + \frac{\gamma^2}{2}(1-\bar{\lambda}), \qquad a_{\max}^+ := M_\infty + \frac{\gamma^2}{2}. \tag{5.23}$$

令 $c_1 = c_1(\gamma, \lambda) > 0$ 为基准李雅普诺夫函数的显式二次下界常数（相差一个加性常数），即

$$c_1 := \frac{1}{8}\left(\gamma^2(1-\bar{\lambda}) + 2 - \sqrt{(\gamma^2(1-\bar{\lambda}) - 2)^2 + 4\gamma^2}\right). \tag{5.24}$$

令 $B := \left(\begin{smallmatrix} 0 & I_d \\ -Q_\infty & -\gamma I_d \end{smallmatrix}\right)$ 为无穷远处的线性漂移矩阵。定义

$$\eta := \frac{\gamma - \sqrt{(\gamma^2 - 4m_\infty)_+}}{2} > 0, \qquad C_B := 1 + \frac{\gamma}{2\sqrt{m_\infty}} + \sqrt{M_\infty} + \frac{1}{\sqrt{m_\infty}}. \tag{5.25}$$

最后，定义

$$C_{B_1}^+ := 2\left(1 + \gamma + M_\infty + \frac{\gamma^2}{2}\right), \qquad \widetilde{C}_{\mathcal{M}}^+ := \frac{1}{2c_1} \cdot \frac{C_B^2}{2\eta}\, C_{B_1}^+, \tag{5.26}$$

以及显式下界

$$\underline{c}_{\mathrm{imp}}^- := \frac{3}{8} \cdot \frac{a_{\min}^- + 1 - \sqrt{(a_{\min}^- - 1)^2 + \gamma^2}}{a_{\max}^+ + 1 + \sqrt{(a_{\max}^+ - 1)^2 + \gamma^2} \;+\; 8\,\delta_U^+}. \tag{5.27}$$

那么，显式选择

$$\lambda_\star(\gamma) := \min\left\{\bar{\lambda}, \frac{\underline{c}_{\mathrm{imp}}^-}{\gamma + \widetilde{C}_{\mathcal{M}}^+}\right\} \tag{5.28}$$

将得到以下性质。

（*i*）对于任意 $\lambda \in (0, \lambda]$，命题 5.6（*b*）中的耗散不等式蕴含了假设 2.1（*iii*），其中 λ（加性常数取决于 λ）。

(*ii*) 对于每个 $\lambda \in (0, \lambda_\star(\gamma)]$, , 定量条件

$$\delta_{\mathrm{LR}} > \gamma\lambda$$

成立，其中

$$\delta_{\mathrm{LR}} := \underline{c}_{\mathrm{imp}}(\lambda) - \lambda\,\widetilde{C}^{\mathrm{LR}}_{\mathcal{M}}(\lambda),$$

以及 $\underline{c}_{\mathrm{imp}}(\lambda)$、$\widetilde{C}^{\mathrm{LR}}_{\mathcal{M}}(\lambda)$ 是引理 *4.3* 应用于当前模型时出现的常数。

证明。 证明见附录 D.7。 □

将命题 5.7 与推论 4.13 结合（在维度为 1 时，必要时再进行张量化），可得到显式常数 $\alpha_{\mathrm{LR}} > 0$ 和 $\kappa_{\mathrm{LR}} > 0$，使得 HFHR 收缩率满足 $c_\alpha \geq c_0^{\mathrm{LR}} + \kappa_{\mathrm{LR}}\alpha$ 对所有 $\alpha \in (0, \alpha_{\mathrm{LR}}]$，只要 $\lambda \in (0, \lambda_\star(\gamma)]$。

定理 5.9 （贝叶斯线性回归的 HFHR 加速）。 考虑由 (5.13)–(5.14) 定义的贝叶斯线性回归问题，并假设 *5.5*成立。令 P_t^α 为相应 *HFHR* 动力学的半群。令 $\rho_{\mathcal{V}_\alpha}$ 为推论 *4.13* 中使用的李雅普诺夫加权半度量（利用全局 *Lipschitz* 常数 L和李雅普诺夫函数 $\mathcal{V}_\alpha$构造），并令 $\mathcal{W}_{\rho_{\mathcal{V}_\alpha}}$ 为相应的 *Wasserstein* 距离。

固定 $\gamma > 0$，并按照引理 *5.8* 选择耗散参数 $\lambda \in (0, \lambda_\star(\gamma)]$，使得$\delta_{\mathrm{LR}} > \gamma\lambda$。此外，假设推论 *4.13* 的定量条件成立（特别是，$\Lambda_0 > 1/2$）。那么存在显式常数 $\alpha_{\mathrm{LR}} > 0$ 和 $\kappa_{\mathrm{LR}} > 0$，它们取决于模型参数（$X, y, \sigma, \iota, p, \varepsilon$）以及 γ，使得对于任意 $\alpha \in (0, \alpha_{\mathrm{LR}}]$,

$$\mathcal{W}_{\rho_{\mathcal{V}_\alpha}}(\mu P_t^\alpha, \nu P_t^\alpha) \;\leq\; e^{-(c_0^{\mathrm{LR}} + \kappa_{\mathrm{LR}}\alpha)t}\,\mathcal{W}_{\rho_{\mathcal{V}_\alpha}}(\mu, \nu), \qquad t \geq 0,$$

对于所有具有有限 *Lyapunov* 矩的概率测度 μ, ν，其中 c_0^{LR}表示参数为 $\alpha = 0$时动能朗之万动力学的收缩率。此外，可以显式地选取

$$\alpha_{\mathrm{LR}} := \min\left\{\alpha^{\mathrm{LR}}_{\mathrm{branch,acc}},\ \alpha^{\mathrm{LR}}_{\mathrm{metric,acc}}\right\},$$

其中 $\alpha^{\mathrm{LR}}_{\mathrm{branch,\ acc}}$ 以及 $\alpha^{\mathrm{LR}}_{\mathrm{metric,\ acc}}$ 分别是定理 *4.12* 和定理 *4.8*中的显式阈值，使用命题 *5.6* 和命题 *5.7*中的全局常数计算得到。类似地，显式增益为

$$\kappa^{\mathrm{LR}}_{\mathrm{global}} := \min\left\{\kappa_{\mathrm{LR}},\ c_0^{\mathrm{LR}}\,c_2^{\mathrm{LR}},\ c_0^{\mathrm{LR}}\,c_3^{\mathrm{LR}}\right\},$$

其中 κ_{LR} 是定理 *4.12*中的 *Lyapunov* 分支增益，c_2^{LR}，c_3^{LR}是定理 *4.8*中的度量分支改进常数。特别是，利用命题 *5.7*中的显式常数，

$$\kappa_{\mathrm{LR}} = \frac{L(\delta_{\mathrm{LR}} + \gamma\lambda)}{768\,\gamma},$$

其中

$$\delta_{\mathrm{LR}} := \underline{c}_{\mathrm{imp}} - \lambda\,\widetilde{C}^{\mathrm{LR}}_{\mathcal{M}}, \qquad C_{\lambda,\mathrm{LR}} := C_2^{\mathrm{LR}} + \widetilde{C}^{\mathrm{LR}}_{\mathcal{M}}\,\underline{c}_{\mathrm{imp}},$$

并且 $L = \frac{\|X^\top X\|_{\mathrm{op}}}{\sigma^2} + \iota p \varepsilon^{p-2}$ 是命题 *5.6*中的全局 *Lipschitz* 常数。

证明。 我们在附录 D.8中给出证明。 □

5.3 贝叶斯二分类

本节考虑一个带有数据$\{(x_i,\ y_i)\}_{i=1}^n$的二分类任务的贝叶斯表述，其中 $x_i \in \mathbb{R}^d$是特征向量，$|x_i| < \infty$，而 $y_i \in \{0,1\}$是标签。在分类任务中，我们的目标是学习形如 $\mathbb{P}(y_i = 1 \mid x_i,\ q) = h(\langle q,\ x_i\rangle)$的预测模型。设 $h:\ \mathbb{R} \to \mathbb{R}$ 为预测函数，$\varphi:\ \mathbb{R} \to \mathbb{R}_+$为损失函数。在此设定下，我们可以将势函数 $U:\ \mathbb{R}^d \to \mathbb{R}$ 写为

$$U(q) = \frac{1}{n}\sum_{i=1}^{n} \varphi(y_i - h(\langle q, x_i\rangle)) + \frac{\iota}{2}|q|^2, \tag{5.29}$$

相应的采样目标是吉布斯后验 $\pi(q) \propto \exp(-U(q))$。对样本迭代 K 步后，分类器可表示为 $\hat{}y = \mathbf{1}_{\{h(\langle \bar{q}_K, \hat{}x\rangle) \geq 1/2\}} \in \{0,1\}$，其中 $\bar{q}_K$ 是 M 条链在第 K 个迭代点处的平均值，$\hat{}x \in \mathbb{R}^d$ 是用于预测 $\hat{}y \in \{0,1\}$ 的给定新特征。我们作出以下假设，另见 [GGZ22]，中的假设 12、[FSS18] 中的假设 2，以及 [MBM18] 中的假设9。

假设 5.10。假设以下条件成立。

- $B_x := \text{最大值}_{1\leq i\leq n} |x_i| < \infty$。
- $h \in C^2$ 使得 $H_1:\ = \text{上确界}_x |h'(x)| < \infty$ 且 $H_2:\ = \text{上确界}_x |h''(x)| < \infty$。
- $\varphi \geq 0$ 和 $\varphi \in C^2$ 使得 $\Phi_1:\ = \sup_x |\varphi'(x)| < \infty$ 且 $\Phi_2:\ = \sup_x |\varphi''(x)| < \infty$。

假设 5.10 较为宽松，许多 φ，h的选择都能满足该假设。例如，遵循[MBM18、FSS18、GGZ22]，，我们考虑 $h(z):\ = z$ 采用 Tukey 双平方损失：

$$\varphi_{\text{Tukey}}(t) := \begin{cases} 1 - \left(1 - (t/t_0)^2\right)^3 & \text{for } |t| \leq t_0, \\ 1 & \text{for } |t| > t_0. \end{cases} \tag{5.30}$$

于是，假设 5.10 得以满足，且势函数 $U(q$通常是非凸的。关于 φ 在无穷远处有界或缓慢增长的非凸例子，也已在 [FSS18、MBM18]中有所研究。接下来，我们证明在假设5.10下，贝叶斯二分类问题（5.29）满足我们的理论所要求的假设 2.1 和4.1 。

命题 5.11 （贝叶斯二分类势函数满足既定假设）。考虑 U 中的势函数 (5.29)。假设假设 *5.10*成立。则：

(a) $U \in C^2(\mathbb{R}^d)$ *and* $U \geq 0$. *Moreover,* ∇U *is* L*-Lipschitz with* $L := \iota + \left(\Phi_2 H_1^2 + \Phi_1 H_2\right) B_x^2$.

(b) U 满足假设 *2.1-(iii)* 意义下的耗散性。特别地，

$$\langle \nabla U(q), q\rangle \geq \frac{\iota}{2}|q|^2 - \frac{C_0^2}{2\iota} \quad \textit{for all } q \in \mathbb{R}^d,$$

其中可取 $C_0:\ = \Phi_1 H_1 B_x$。

(c) U 满足假设*4.1*，其中

$$Q_\infty := \iota I_d, \qquad \varrho(r) := \frac{C_0}{r}, \quad r \geq 1.$$

证明。 证明见附录 D.9。 □

与采用平滑 L^p 正则化的贝叶斯线性回归（其中 $|\nabla U(q) - Q_\infty q| = \mathcal{O}(|q|^{p-1})$ 无界）相比，本文的分类势函数满足余项一致有界：

$$\nabla U(q) = \iota q + r(q), \qquad |r(q)| \le C_0 := \Phi_1 H_1 B_x.$$

引理 $\rho_\nabla(\cdot)$ 和 $\delta_U(\cdot)$ 中的尾模量在引理 4.3 中具有尤其简单的显式界，因此可以得到一个显式的 λ 取值范围，以确保满足定量度量分支条件$\delta_{\mathrm{BC}} > \gamma\lambda$。

命题 5.12 (贝叶斯二分类中引理4.3 的显式常数)。 假设满足命题 5.11中的设定，并令 $\lambda \in (0, 1/4]$ 表示假设 2.1 中出现的耗散参数（而非式 ι 中的岭系数 (5.29)）。令 $Q_\infty := \iota I_d$ 且 $C_0 := \Phi_1 H_1 B_x$。则有以下结论。

(i) ***Spectral bounds for*** Q_∞***:*** $\lambda_{\min}(Q_\infty) = \lambda_{\max}(Q_\infty) = \iota$.

(ii) 显式尾部模量 ρ_∇。对于任意 $R \ge 1$,

$$\rho_\nabla(R) := \sup_{|q| \ge R} \frac{|\nabla U(q) - Q_\infty q|}{|q|} \le \frac{C_0}{R}. \tag{5.31}$$

(iii) 显式尾部模量 δ_U。对于任意 $R \ge 1$,

$$\delta_U(R) := \sup_{|q| \ge R} \frac{\left|U(q) - \frac{1}{2}\langle Q_\infty q, q\rangle\right|}{1 + |q|^2} \le \frac{C_0}{R} + \frac{A_\varphi}{R^2}, \tag{5.32}$$

其中 $A_\varphi := |\varphi(0)| + \Phi_1(1 + |h(0)|)$。

(iv) 显式可容许截断半径 R_0。令 $\rho_\star$ 按 (4.13) 定义。由于 $R \mapsto C_0/R$单调递减，取值

$$R_0 := \max\left\{1,\ C_{\mathrm{linear}},\ \frac{C_0}{\rho_\star}\right\} \tag{5.33}$$

可确保 $\rho_\nabla(R_0) \le \rho_\star$，因此 (4.14) 成立。

(v) 通过与一般理论相同的李雅普诺夫方程构造二次校正项。令

$$B := \begin{pmatrix} 0 & I_d \\ -Q_\infty & -\gamma I_d \end{pmatrix} = \begin{pmatrix} 0 & I_d \\ -\iota I_d & -\gamma I_d \end{pmatrix}$$

并令 C_{B_1} 为显式对称矩阵 $\nabla^2 B_1$，来自引理4.3。则可以取$\mathcal{M}(z) = \frac{1}{2} z^\top \mathsf{K} z$，对任意 $z = (q, p) \in \mathbb{R}^{2d}$，其中 K 是以下方程的（唯一）对称解 $B^\top \mathsf{K} + \mathsf{K} B = C_{B_1}$，等价地可表示为

$$\mathsf{K} = \int_0^\infty e^{tB^\top} C_{B_1} e^{tB}\, dt. \tag{5.34}$$

(vi) 一阶改进常数 c_{imp} （显式下界）。对于_

$$a_{\min} = a_{\max} =: a(\lambda) = \iota + \frac{\gamma^2}{2}(1-\lambda),$$

引理 *4.3*推出 (4.8) 其中

$$\mathfrak{c}_{\mathrm{imp}} := \frac{3}{8} \cdot \frac{a(\lambda) + 1 - \sqrt{(a(\lambda)-1)^2 + \gamma^2}}{a(\lambda) + 1 + \sqrt{(a(\lambda)-1)^2 + \gamma^2} \; + \; 8\,\delta_U(R_0)} \; > \; 0, \tag{5.35}$$

其中 $\delta_U(R_0)$ 可利用 (5.32) 和 R_0 由 (5.33)显式界定。引理C_{imp} ***4.3*** 中的 是有限的。

(vii)命题 *4.7*的漂移率展开界。 将K 写成分块形式$\mathsf{K} = \begin{pmatrix} \mathsf{K}_{qq} & \mathsf{K}_{qp} \\ \mathsf{K}_{pq} & \mathsf{K}_{pp} \end{pmatrix}$并令

$$k_q := \|\mathsf{K}_{qq}\|_{\mathrm{op}} + \|\mathsf{K}_{qp}\|_{\mathrm{op}}, \qquad b_0 := |\nabla U(0)|.$$

根据命题 *5.11*，有 $\nabla U(q) = \iota q + r(q)$ 其中 $|r(q)| \le C_0$ 对所有 q均成立，因此 $b_0 \le C_0$并且

$$|\nabla U(q)| \le \iota|q| + C_0 \le L|q| + b_0,$$

其中 $L = \iota + (\Phi_2 H_1^2 + \Phi_1 H_2) B_x^2$ 是命题 *5.11****(a)*** 中的全局 ***Lipschitz*** 常数。由于 $\mathcal{M}(z) = \frac{1}{2} z^\top \mathsf{K} z$，有 $\nabla_q \mathcal{M}(q,\ p) = \mathsf{K}_{qq} q + \mathsf{K}_{qp} p$，因此 $|\nabla_q \mathcal{M}(q,\ p)| \le k_q(|q| + |p|)$。因此

$$|\mathcal{A}'\mathcal{M}(q,p)| = |\langle \nabla U(q), \nabla_q \mathcal{M}(q,p)\rangle| \le k_q(L|q| + b_0)(|q| + |p|) \le k_q\left(\frac{3L+1}{2}|q|^2 + \frac{L+1}{2}|p|^2 + b_0^2\right),$$

而且

$$|\Delta_q \mathcal{M}(q,p)| = \mathrm{tr}(\mathsf{K}_{qq}) \le d\,\|\mathsf{K}_{qq}\|_{\mathrm{op}}.$$

设 $c_1 > 0$ 为一个二次下界常数，使得对U进行加性常数平移后，

$$\mathcal{V}_0(q,p) \ge c_1\left(|q|^2 + |p|^2\right), \qquad (q,p) \in \mathbb{R}^{2d}. \tag{5.36}$$

那么“误差项”

$$\mathrm{Err}^{(d)}(q,p) := |\mathcal{A}'\mathcal{M}(q,p)| + |\Delta_q \mathcal{M}(q,p)|$$

对于所有$(q,p) \in \mathbb{R}^{2d}$，均满足，

$$\mathrm{Err}^{(d)}(q,p) \le C_2^{\mathrm{BC}}\ \mathcal{V}_0(q,p) + C_2^{(d),\mathrm{BC}}, \tag{5.37}$$

取如下明确的选择

$$C_2^{\mathrm{BC}} := \frac{k_q}{c_1} \max\left\{\frac{3L+1}{2}, \frac{L+1}{2}\right\}, \qquad C_2^{(d),\mathrm{BC}} := k_q b_0^2 + d\,\|\mathsf{K}_{qq}\|_{\mathrm{op}} \ \le \ k_q C_0^2 + d\,\|\mathsf{K}_{qq}\|_{\mathrm{op}}.$$

最后定义（显式的）增长常数

$$\widetilde{C}_{\mathcal{M}}^{\mathrm{BC}} := \frac{\|\mathsf{K}\|_{\mathrm{op}}}{2c_1}$$

（参见 （4.22））， 并令

$$\delta_{\mathrm{BC}} := \underline{c}_{\mathrm{imp}} - \lambda\,\widetilde{C}^{\mathrm{BC}}_{\mathcal{M}}, \qquad C_{\lambda,\mathrm{BC}} := C^{\mathrm{BC}}_{2} + \widetilde{C}^{\mathrm{BC}}_{\mathcal{M}}\,\underline{c}_{\mathrm{imp}}. \tag{5.38}$$

然后证明 位置 *4.7* 适用 （对于 α 足够小的情况）， 并给出漂移率展开

$$\lambda_\alpha \;\geq\; \lambda + \delta_{\mathrm{BC}}\cdot\alpha - C_{\lambda,\mathrm{BC}}\cdot\alpha^2.$$

特别地，对于所有足够小的 $\alpha > 0$， 只要 $\delta_{\mathrm{BC}} > 0$，就有 $\lambda_\alpha > \lambda$ 。

证明。 我们在附录 D.10 中给出证明。 □

最后，我们确认，通过使耗散参数 λ 相对于岭系数 ι 足够小，可以满足加速所需的定量条件 $\delta_{\mathrm{BC}} > \gamma\lambda$ 。

引理 5.13 （可行性 $\delta_{\mathrm{BC}} > \gamma\lambda$ ， 具有显式 $\lambda_\star(\gamma)$ ）。 假设命题 *5.11* 的设定，并固定 $\gamma > 0$。令

$$\bar{\lambda} := \min\left\{\frac{1}{4},\ \frac{\iota}{2}\right\}, \qquad R := \max\{1, C_{\mathrm{linear}}\}.$$

定义

$$\delta_U^+ := \frac{C_0}{R} + \frac{A_\varphi}{R^2}, \qquad C_0 := \Phi_1 H_1 B_x, \qquad A_\varphi := |\varphi(0)| + \Phi_1(1 + |h(0)|). \tag{5.39}$$

Let

$$a^- := \iota + \frac{\gamma^2}{2}(1 - \bar{\lambda}).$$

令 $c_1 = c_1(\gamma, \lambda) > 0$ 为 $\mathcal{V}_0$ 的基准二次下界常数（差一个加性常数），即

$$c_1 := \frac{1}{8}\left(\gamma^2(1-\bar{\lambda}) + 2 - \sqrt{(\gamma^2(1-\bar{\lambda}) - 2)^2 + 4\gamma^2}\right). \tag{5.40}$$

令 $B = \begin{pmatrix} 0 & I_d \\ -\iota I_d & -\gamma I_d \end{pmatrix}$ 并定义显式衰减代理量

$$\eta := \frac{\gamma - \sqrt{(\gamma^2 - 4\iota)_+}}{2} > 0, \qquad C_B := 1 + \frac{\gamma}{2\sqrt{\iota}} + \sqrt{\iota} + \frac{1}{\sqrt{\iota}}.$$

另定义

$$C^+_{B_1} := 2\left(1 + \gamma + \iota + \frac{\gamma^2}{2}\right), \qquad \widetilde{C}^+_{\mathcal{M}} := \frac{1}{2c_1}\cdot\frac{C_B^2}{2\eta}\,C^+_{B_1}.$$

最后定义显式下界

$$\underline{c}^-_{\mathrm{imp}} := \frac{3}{8}\cdot\frac{a^- + 1 - \sqrt{(a^- - 1)^2 + \gamma^2}}{a^- + 1 + \sqrt{(a^- - 1)^2 + \gamma^2} \;+\; 8\,\delta_U^+}. \tag{5.41}$$

那么， 显式的选择

$$\lambda_\star(\gamma) := \min\left\{\bar{\lambda},\ \frac{\underline{c}^-_{\mathrm{imp}}}{\gamma + \widetilde{C}^+_{\mathcal{M}}}\right\} \tag{5.42}$$

足以保证对于每个 $\lambda \in (0,\ \lambda_\star(\gamma)]$ ， 我们都有

$$\delta_{\mathrm{BC}} > \gamma\lambda, \qquad \delta_{\mathrm{BC}} := \underline{c}_{\mathrm{imp}}(\lambda) - \lambda\,\widetilde{C}^{\mathrm{BC}}_{\mathcal{M}}(\lambda),$$

其中， $\underline{c}_{\mathrm{imp}}(\lambda)$ 和 $\widetilde{C}^{\mathrm{BC}}_{\mathcal{M}}(\lambda)$ 是将引理 *4.3* 应用于 (5.29)时得到的常数。

证明。 证明见附录 D.11。 □

结合推论 5.12 中得到的显式一阶改进与引理 5.13中的参数选择策略，我们得到贝叶斯二分类背景下 HFHR 动力学的以下定量加速结果。

定理 5.14 （贝叶斯二分类的 HFHR 加速）。 考虑势函数 (5.29) 所定义的贝叶斯二分类问题，并假设 *5.10*成立。令 P_t^α 为相应 *HFHR* 动力学的半群。令 $\rho_{\mathcal{V}_\alpha}$ 为推论 *4.13* 中使用的李雅普诺夫加权半度量（通过全局 *Lipschitz* 常数 L和李雅普诺夫函数$\mathcal{V}_\alpha$构造），并令 $\mathcal{W}_{\rho_{\mathcal{V}_\alpha}}$ 为相应的 *Wasserstein* 距离。

固定 $\gamma > 0$，并按照引理 *5.13* 选择耗散参数 $\lambda \in (0, \lambda_\star(\gamma)]$，使定量条件 $\delta_{\mathrm{BC}} > \gamma\lambda$ 成立。进一步假设推论 *4.13* 的定量条件成立（特别是 $\Lambda_0 > 1/2$）。那么存在显式常数 $\alpha_{\mathrm{BC}} > 0$ 和 $\kappa_{\mathrm{BC}} > 0$，它们取决于模型参数（$B_x, H_1, H_2, \Phi_1, \Phi_2, \iota$）以及 γ，使得对每个$\alpha \in (0, \alpha_{\mathrm{BC}}]$,

$$\mathcal{W}_{\rho_{\mathcal{V}_\alpha}}(\mu P_t^\alpha, \nu P_t^\alpha) \;\le\; e^{-(c_0^{\mathrm{BC}} + \kappa_{\mathrm{BC}}\alpha)t}\, \mathcal{W}_{\rho_{\mathcal{V}_\alpha}}(\mu, \nu), \qquad t \ge 0,$$

对于所有具有有限 *Lyapunov* 矩的概率测度 μ、ν，其中 c_0^{BC} 表示在 $\alpha = 0$ 处动能朗之万动力学的收缩率。此外，可以显式地选取

$$\alpha_{\mathrm{BC}} := \min\left\{\alpha^{\mathrm{BC}}_{\mathrm{branch,acc}},\ \alpha^{\mathrm{BC}}_{\mathrm{metric,acc}}\right\},$$

其中 $\alpha^{\mathrm{BC}}_{\mathrm{branch、acc}}$ 和 $\alpha^{\mathrm{BC}}_{\mathrm{metric、acc}}$ 分别是根据命题 *5.11* 和命题 *5.12*中的全局常数计算得出的定理 *4.12* 和定理 *4.8*中的显式阈值。类似地，显式增益为

$$\kappa^{\mathrm{BC}}_{\mathrm{global}} := \min\left\{\kappa_{\mathrm{BC}},\ c_0^{\mathrm{BC}}\, c_2^{\mathrm{BC}},\ c_0^{\mathrm{BC}}\, c_3^{\mathrm{BC}}\right\},$$

其中 κ_{BC} 是定理 *4.12*中的 *Lyapunov* 分支增益，而 c_2^{BC}、c_3^{BC}是定理 *4.8*中的度量分支改进常数。特别是，利用命题 *5.12*中的显式常数，可以取

$$\kappa_{\mathrm{BC}} = \frac{L(\delta_{\mathrm{BC}} + \gamma\lambda)}{768\,\gamma},$$

其中 $L = \iota + (\Phi_2 H_1^2 + \Phi_1 H_2) B_x^2$ 是命题 *5.11* 中的全局 *Lipschitz* 常数，且

$$\delta_{\mathrm{BC}} := \underline{c}_{\mathrm{imp}} - \lambda\, \widetilde{C}^{\mathrm{BC}}_{\mathcal{M}}, \qquad C_{\lambda,\mathrm{BC}} := C_2^{\mathrm{BC}} + \widetilde{C}^{\mathrm{BC}}_{\mathcal{M}}\, \underline{c}_{\mathrm{imp}}.$$

证明。 证明见附录 D.12。 □

6 数值实验

在本节中，我们对无 *Hessian* 高分辨率蒙特卡洛（HFHRMC）进行数值实验。该方法基于式（2.1）中 HFHR 动力学的 Euler–Maruyama 离散化。HFHRMC 的迭代点定义如下:

$$\begin{aligned} q_{k+1} &= q_k + (p_k - \alpha\nabla U(q_k))\,\eta + \sqrt{2\alpha\eta}\,\xi^q_{k+1}, \\ p_{k+1} &= p_k + (-\gamma p_k - \nabla U(q_k))\,\eta + \sqrt{2\gamma\eta}\,\xi^p_{k+1}, \end{aligned} \tag{6.1}$$

其中 ξ_k^q， ξ_k^p 是独立同分布的高斯随机向量 $\mathcal{N}$（0， I_d），且 ξ_k^q， ξ_k^p 彼此独立。我们还使用动力学朗之万蒙特卡罗（KLMC）进行实验，该方法基于式（1.3）中动能朗之万动力学的 Euler–Maruyama 离散化，其迭代点定义如下：

$$x_{k+1} = x_k + v_k\eta, \tag{6.2}$$

$$v_{k+1} = v_k + (-\gamma v_k - \nabla U(x_k))\,\eta + \sqrt{2\gamma\eta}\xi_{k+1}, \tag{6.3}$$

其中， ξ_k 是独立同分布的高斯随机向量 $\mathcal{N}$(0， I_d)。

在以下各节中，我们将使用 HFHRMC 和 KLMC 进行数值实验。首先，我们将针对一个玩具示例，即式（5.1）中的多阱势情形（第6.1节），进行数值实验。接下来，我们将研究带有 L^p 正则化项的贝叶斯线性回归（第6.2节）。我们还将把这些算法应用于贝叶斯二分类（第6.3节）。在所有这些示例中，势函数 U 均为非凸，并满足假设2.1和假设4.1。最后，我们将研究另一个数值示例：带岭正则化项的贝叶斯逻辑回归，其中势函数 U 是非凸的，且可能不满足假设2.1 和4.1（第6.4节）。

6.1 多阱势

本节基于一个玩具示例开展数值实验，即第 5.1 节中讨论的多阱势，该势函数同时满足假设 2.1 和假设 4.1。我们考虑维度为 $d = 8$ 的多阱势，并为 HFHRMC (6.1) 选取 α 的不同取值：0.01、0.05、0.1、0.2、0.5、0.8、1.0；对于 (6.1) 中的 HFHRMC 和 (6.2) 中的 KLMC，均选取 $\gamma = 2.0$ 。我们用步长 $\eta = 10^{-3}$ 让这两种算法各迭代 10000 步，并在 $M = 2000$ 条链上进行计算。所得结果如图1所示，其中 x 轴表示迭代点 k， y 轴表示算法生成的经验分布与吉布斯分布之间的 Wasserstein 距离的对数。

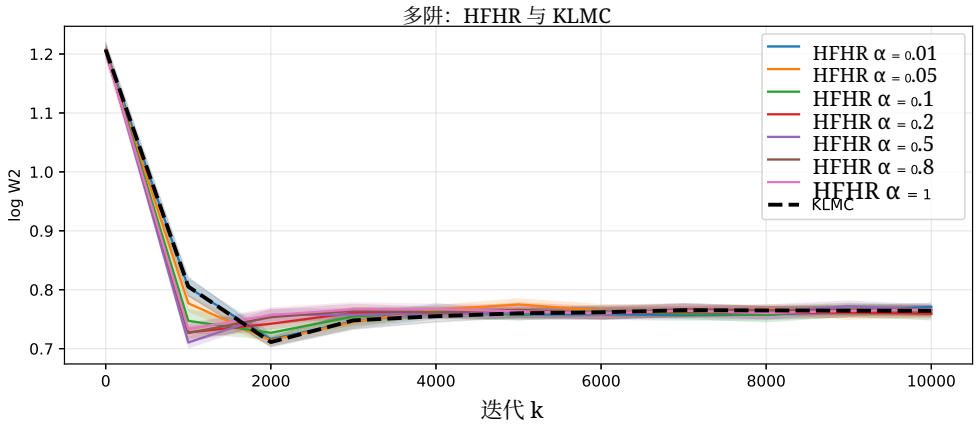

图1： $d = 8$维度下的多阱势。

从图1 可以看出，在这个多阱势示例中，HFHRMC 的表现优于 KLMC。我们发现，对于 $\alpha = 0.01$，HFHRMC 和 KLMC 的收敛性能相当。然而，当

α时。具体而言，我们观察到，增大 α 会加快 HFHRMC 的收敛，这与我们在第 4 节中的理论一致。当 α 趋近于 1 时，HFHRMC 的收敛速度会变慢，这对应于较小的收缩率，如推论3.13所示。

6.2 使用合成数据进行贝叶斯线性回归

本节中，我们考虑如下贝叶斯线性回归模型：

$$y_j = x_j^\top \beta_* + \delta_j, \quad \delta_j \sim \mathcal{N}(0, \sigma^2), \quad x_j \sim \mathcal{N}\left(0, 0.5^2 I_d\right), \quad j = 1, \ldots, n, \tag{6.4}$$

其中，$\beta_* = [1.0, -0.5, 0.7, 1.2, -3.0, 5.4]^\top$ 是一个固定的真实系数向量。我们的目标是从由 $\pi(q) \propto \exp\{-U(q)\}$给出的后验分布中采样，其中 $U(q)$是负对数后验，即带有我们将要选定的正则化项的平方损失。为了展示算法的收敛性能，我们根据以下公式计算第 k个迭代点处的均方误差：均方误差$_k := \frac{1}{n}\sum_{j=1}^{n}\left(y_j - (x_j)^\top q_k\right)^2$；经过 M 条链的 K次迭代后，参数的均值为 $\bar{q}_K = \frac{1}{M}\sum_{m=1}^{M} q_k^{(m)}$。

我们采用第 5.2 节中介绍的带有 L^p 正则化项的贝叶斯线性回归，并考虑式 (5.13) 中的目标函数。如第 5.2 节所述，相应的目标函数 U 满足假设 2.1 和假设 4.1。将 HFHRMC (6.1) 和 KLMC (6.2) 中的参数设为 $\alpha = 0.1$、$\gamma_{\text{HFHRMC}} = 1.0$、$\gamma_{\text{KLMC}} = 10.0$，并将线性回归 (5.13) 中的参数设为 $\sigma = 0.4$、$\lambda = 0.1$、$\varepsilon = 0.001$、$p = 1.2$ 后，我们取$n = 1000$ 个样本、运行 $M = 10$ 条链，选择 $\eta = 10^{-4}$ 并运行 10000 步，得到图2所示的图。如图所示，与使用 $\gamma_{\text{KLMC}} = 10.0$的 KLMC 相比，使用 $\gamma_{\text{HFHRMC}} = 1.0$ 的 HFHRMC 收敛速度显著更快，且均方误差更低。后者表现出振荡行为，并且直到大约 6000 步后才开始收敛。还值得注意的是，当 γ_{KLMC} 设为 1.0 时，KLMC 无法收敛。

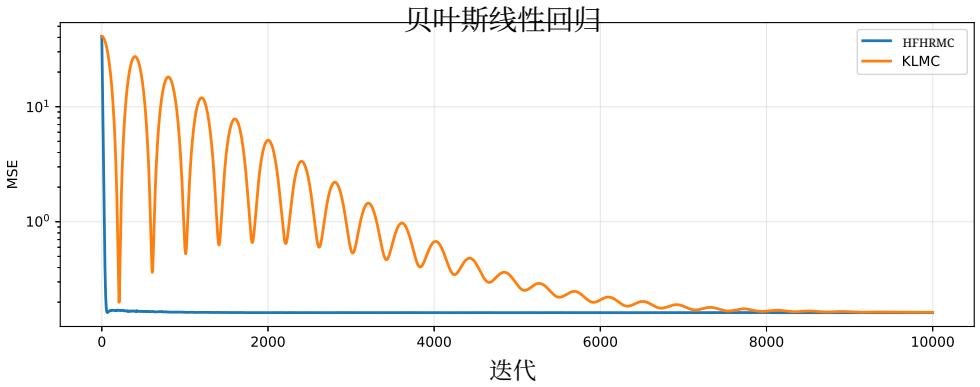

图2：贝叶斯线性回归。

6.3 使用真实数据进行贝叶斯二分类

在本节中，我们测试算法在贝叶斯二分类问题中的表现，问题定义见（5.29），并采用（5.30）中的 Tukey 双平方损失。这些问题在第 5.3 节中介绍。如

第5.3节所述，相应的目标函数 U 满足假设2.1 和4.1。我们将 HFHRMC 和 KLMC 算法应用于这项基于真实数据（乳腺癌[2]）的贝叶斯二分类任务。威斯康星乳腺癌（诊断）数据集包含 $n = 569$ 个样本和$d = 30$ 个实值特征。二元响应表示肿瘤是恶性（标记为 1）还是良性（标记为 0）。我们将数据集划分为训练集和测试集（70/30）。目标是实现二分类：给定 $x \in \mathbb{R}^{30}$，我们能够预测 $y \in \{0, 1\}$。

为了展示算法的收敛性能，我们首先计算参数的均值，计算时使用 K 条链上的迭代点 M。然后，对于一个测试特征 $\hat{x}$，我们计算预测标签 $\hat{y} = \mathbf{1}_{\{h(\langle \bar{q}_K, \hat{x} \rangle) \geq 1/2\}} \in \{0, 1\}$，其中 h 是第5.3节中定义的预测函数。分类性能通过以下形式的测试准确率进行评估：

$$\mathrm{Acc} := \frac{1}{n_{\mathrm{test}}} \sum_{i=1}^{n_{\mathrm{test}}} \mathbf{1}_{\hat{y}_i = y_i}. \tag{6.5}$$

为进行实验，我们使用 U 式（5.29）中贝叶斯二分类问题的目标函数，并采用式（5.30）中的 Tukey 双平方损失；同时选择 HFHRMC（6.1）和 KLMC（6.2）中的参数，使得 $\alpha = 0.05$, $\gamma = 1.0$，并选择采用 Tukey 双平方损失的二分类问题中的参数，使得 $\iota = 0.05$ 且 $t_0 = 2.0$。此外，我们采用 $M = 50$ 条链，并选择 $\eta = 10^{-4}$，每条链运行 20000 步。因此，我们得到图3。

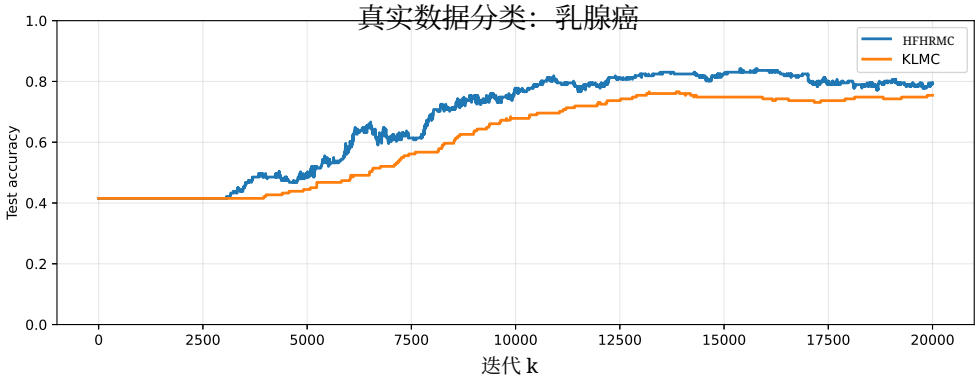

图 3：贝叶斯二分类。

我们可以从图 3 中观察到，在贝叶斯二分类任务中，HFHRMC 的测试准确率约为 80%，高于 KLMC；此外，我们还观察到，HFHRMC 的收敛速度略快于 KLMC。

6.4 使用真实数据的贝叶斯逻辑回归

在本节中，我们考虑使用神经网络处理的真实数据（鸢尾花[3]）进行贝叶斯逻辑回归。鸢尾花数据集包含 $n = 150$ 个样本，每个样本具有 $d = 4$ 个实值特征。为适配贝叶斯二元逻辑回归框架，我们从数据集中选取变色鸢尾和维吉尼亚鸢尾这两个类别，并将观测值重新标记为 $y_i \in \{0, 1\}$ 和 $x_i \in \mathbb{R}^4$。我们将数据集划分为

[2]乳腺癌 - UCI 机器学习库，https://archive.ics.uci.edu/dataset/14/breast+癌症[3]鸢尾花 - UCI 机器学习库，https://archive.ics.uci.edu/dataset/53/iris

训练集和测试集（80/20），我们的目标是对给定特征和参数向量 $q \in \mathbb{R}^d$ 时标签的条件分布建模为 $\mathbb{P}(y_i = 1|x_i, q) = \sigma\left(q^\top x_i\right)$，其中 Sigmoid 函数为$\sigma(z) = \frac{1}{1+e^{-z}}$。我们对回归系数施加高斯先验 $q \sim \mathcal{N}(0, \iota^{-1} I_d)$，使其对应的吉布斯势为 $\pi(q) \propto e^{-U(q)}$，其中 $U(q) = \frac{1}{n}\sum_{i=1}^n \Big(\log\left(1 + e^{q^\top x_i}\right) - y_i\, q^\top x_i\Big) + \frac{\iota}{2}|q|^2$，第一项为负对数熵损失，第二项为岭正则项。

对于新的特征向量 $\hat{x}$，后验预测概率为 $\bar{p}(\hat{x}) = \mathbb{E}_{q\sim\pi}\left[\sigma\left(q^\top \hat{x}\right)\right]$，可通过 M 条链以 $\frac{1}{M}\sum_{m=1}^M \sigma\left((q^{(m)})^\top \hat{x}\right)$ 的形式近似，预测标签为 $\hat{y} = \mathbf{1}_{\{\bar{p}(\hat{x}) \geq 1/2\}}$。我们处理一个前馈神经网络，并使用来自吉布斯后验的 HFHRMC 和 KLMC 样本来计算预测量。

利用神经网络处理真实数据的贝叶斯逻辑回归研究也见于 [BCKW15，OSK+19，GNZZ25]。尽管在存在神经网络的情况下，似乎难以验证假设 2.1 和 4.1，但我们仍将展示所提算法的效率。具体而言，我们考虑一个具有$L = 3$ 个隐藏层的全连接前馈神经网络，每个隐藏层的神经元数量均为 $N_{\text{neurons}} = 32$，该神经网络由 $\theta = (W_1, \ldots\ldots, W_L) \in \mathbb{R}^D$ 参数化，其中 $W_\ell \in \mathbb{R}^{m_{\ell-1}\times m_\ell}$，且 $m_0 = d$和 $m_L = 1$。为保证势函数的光滑性，我们采用高斯平滑 ReLU 激活函数，使得 $\phi_\nu(z) = \mathbb{E}_{\xi\sim\mathcal{N}(0,\,1)}[(z + \nu\xi)^+]$。可以验证，其导数有界且利普希茨连续。网络前向映射递归定义如下

$$h_0(x) = x, \quad h_\ell(x;\theta) = \phi_\nu\left(h_{\ell-1}(x;\theta)W_\ell\right), \quad \ell = 1, \ldots, L-1,$$

现在，预测概率由参数为 θ 的 Sigmoid 函数给出，使得$p_\theta(x) = \sigma(z_\theta(x)) = \frac{1}{1+e^{-z_\theta(x)}}$. 因此，我们将势函数定义为

$$U(\theta) = \frac{1}{n}\sum_{i=1}^n \mathcal{L}\left(y_i, p_\theta(x_i)\right) + \frac{\iota}{2}|\theta|^2,$$

其中，损失函数为，损失函数为 $\mathcal{L}(y, p_\theta(x)) = -y \log p_\theta(x) - (1-y)\log(1-p_\theta(x))$，而二次项 $\frac{\iota}{2}|\theta|^2$ 对应于高斯先验 $\theta \sim \mathcal{N}(0, \iota^{-1} I_d)$。

为进行二元逻辑回归任务，我们选择高斯平滑因子 $\nu = 32$、岭强度 $\iota = 10^{-3}$，以及 HFHRMC 和 KLMC 的参数，其中 $\alpha = 1.0$ 和 $\gamma = 2.0$，随后运行算法 $M = 50$ 条链，迭代 5000 次，步长为 $\eta = 0.5 \times 10^{-4}$，得到图 4。

图中的左侧曲线4 是通过式（6.5）计算得到的测试准确率，图中的右侧曲线4是预测后验的对数损失。从图中可以看出，HFHRMC 和 KLMC 都取得了较高的预测准确率；此外，HFHRMC 实现了加速，表现更优，其对数损失下降得更快。

7 结论

本文对无海森矩阵高分辨率（HFHR）动力学进行了理论分析，用于从具有非凸势函数的目标分布 $\pi(q) \propto e^{-U(q)}$ 中采样。尽管 HFHR 动力学在各种场景中都展现了良好的实证效果，但现有理论大多局限于强凸情形。我们的工作通过在非凸情形下建立收敛保证，弥合了理论与实践之间的差距。通过采用反射/同步

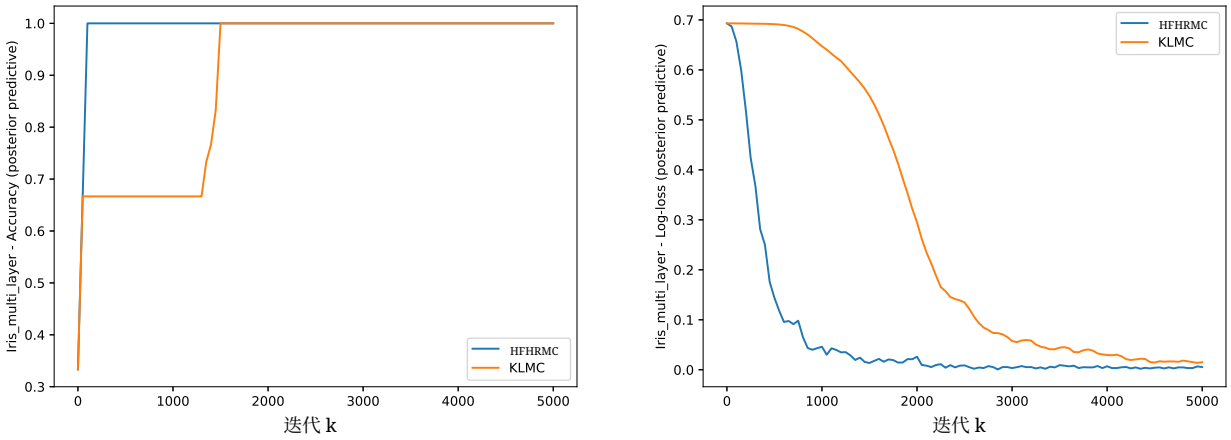

图4：经前馈神经网络处理的贝叶斯逻辑回归，包含 $L = 3$ 层。

耦合框架并构造适当的李雅普诺夫函数，在光滑性和耗散性假设下，我们证明对于所有充分小的分辨率参数 $\alpha > 0$，HFHR 半群在 Lyapunov 加权 Wasserstein 距离下呈指数收缩。尤为重要的是，我们不止证明了基本的收敛性，还展示了定量加速。在额外假设渐近地 ∇U 在无穷远处呈线性增长的条件下，我们证明 HFHR 动力学的收缩率严格优于动能朗之万动力学。我们建立了显式的、关于α 呈线性的加速优势；当收敛受李雅普诺夫漂移（从无穷远处的回归）限制时如此，当收敛主要受度量耦合（跨越能垒）支配时亦如此。我们通过三个具体例子展示了这些理论结果：多阱势、带有 L^p 正则化项的贝叶斯线性回归，以及贝叶斯二元分类。我们基于这些例子进行了数值实验，并额外考察了一个使用神经网络处理真实数据的贝叶斯逻辑回归例子。数值实验验证了理论结果，并展示了基于 HFHR 动力学的算法的效率。数值结果表明，与动能朗之万动力学相比，HFHR 动力学具有加速效果和更优的性能。

致谢

Xiaoyu Wang 获得广州-香港科技大学（广州）联合资助计划（编号：2024A03J0630）和广州市金融科技前沿研究重点实验室的资助。Lingjiong Zhu 部分获得美国国家科学基金会 DMS-2053454 和 DMS-2208303 项目资助。

参考文献

[ADFDJ03] Christophe Andrieu、Nando De Freitas、Arnaud Doucet 和 Michael I Jordan。机器学习中的马尔可夫链蒙特卡罗导论。机器学习，50(1):5–43，2003。

[BCKW15] Charles Blundell、Julien Cornebise、Koray Kavukcuoglu 和 Daan Wierstra。神经网络中的权重不确定性。载于国际机器学习会议，第 37 卷，第 1613–1622 页。PMLR，2015。[BCM+21] Mathias Barkhagen、Ngoc Huy Chau、 ́Eric Moulines、Mikl ́os R ́asonyi、Sotirios Sabanis 和 Ying Zhang。关于对数凹情形下依赖数据流的随机梯度朗之万动力学。Bernoulli，27(1):1–33，2021。[BGK+25] Waheed U. Bajwa、Mert G¨urb¨uzbalaban、Mustafa Ali Kutbay、Lingjiong Zhu 和 Muhammad Zulqarnain。DIGing-SGLD：时变网络上的去中心化、可扩展朗之万采样。arXiv:2511.12836，2025。[CB18] Xiang Cheng 和 Peter L. Bartlett。朗之万马尔可夫链蒙特卡罗在 KL 散度下的收敛。载于第 29 届算法学习理论国际会议论文集（ALT），第 83 卷，第 186–211 页。PMLR，2018。[CCA+18] Xiang Cheng、Niladri S. Chatterji、Yasin Abbasi-Yadkori、Peter L. Bartlett 和 Michael I. Jordan。非凸情形下朗之万动力学的精确收敛速率。arXiv:1805.01648，2018。[CCBJ18] Xiang Cheng、Niladri S Chatterji、Peter L Bartlett 和 Michael I Jordan。欠阻尼朗之万马尔可夫链蒙特卡罗：非渐近分析。载于学习理论会议，第 75 卷，第 300–323 页。PMLR，2018。[CHS87] Tzuu-Shuh Chiang、Chii-Ruey Hwang 和 Shuenn Jyi Sheu。用于全局优化的扩散， $\mathbb{R}^n$。SIAM 控制与优化期刊，25(3):737–753，1987。[CLW21] Yu Cao、Jianfeng Lu 和 Lihan Wang。欠阻尼朗之万动力学随机算法的复杂度。数学科学通讯，19(7):1827–1853，2021。[CLW23] Yu Cao、Jianfeng Lu 和 Lihan Wang。欠阻尼朗之万动力学的显式 L^2-收敛速率估计。理性力学与分析档案，247(90):1–34，2023。[CMR+21] Ngoc Huy Chau、 ́Eric Moulines、Miklos R ́asonyi、Sotirios Sabanis 和 Ying Zhang。关于依赖数据流下的随机梯度朗之万动力学：完全非凸情形。数据科学数学 SIAM 期刊，3(3):959–986，2021。[Dal17] 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（统计方法论），79(3):651–676，2017。[DK19] Arnak S. Dalalyan 和 Avetik G. Karagulyan。梯度不准确时朗之万蒙特卡洛的简明保证。随机过程及其应用，129(12):5278–5311，2019。[DM17] Alain Durmus 和 Eric Moulines。未经调整的朗之万算法的非渐近收敛分析。应用概率年刊，27(3):1551–1587，2017。

[DM19] Alain Durmus 和 Eric Moulines。通过未校正朗之万算法进行高维贝叶斯推断。《Bernoulli》，25(4A)：2854–2882，2019。[DRD20] Arnak S Dalalyan 和 Lionel Riou-Durand。使用动理学朗之万扩散从对数凹密度中采样。《Bernoulli》，26(3)：1956–1988，2020。[Ebe16] Andreas Eberle。扩散过程的反射耦合与收缩率。《概率论及相关领域》，166(3)：851–886，2016。[EGZ19] Andreas Eberle、Arnaud Guillin 和 Raphael Zimmer。朗之万动力学的耦合与定量收缩率。《概率年刊》，47(4)：1982–2010，2019。[EHZ22] Murat A Erdogdu、Rasa Hosseinzadeh 和 Matthew S. Zhang。朗之万蒙特卡洛在卡方散度和 Rényi 散度下的收敛性。载于：第25届国际人工智能与统计会议论文集，第151卷，第8151–8175页。PMLR，2022。[FSS18] Dylan J Foster、Ayush Sekhari 和 Karthik Sridharan。非凸学习与优化中梯度的一致收敛。载于：神经信息处理系统进展，第31卷。Curran Associates, Inc.，2018。[GCSR95] Andrew Gelman、John B Carlin、Hal S Stern 和 Donald B Rubin。《贝叶斯数据分析》。Chapman & Hall/CRC Press，1995。[GGHZ21] Mert G¨urb¨uzbalaban、Xuefeng Gao、Yunhan Hu 和 Lingjiong Zhu。去中心化随机梯度朗之万动力学与哈密顿蒙特卡洛。Journal of Machine Learning Research，22(239)：1–69，2021。[GGZ20] Xuefeng Gao、Mert G¨urb¨uzbalaban 和 Lingjiong Zhu。打破可逆性可加速朗之万动力学，实现全局非凸优化。载于：神经信息处理系统进展（NeurIPS），第33卷，第17850–17862页。Curran Associates, Inc.，2020。[GGZ22] Xuefeng Gao、Mert G¨urb¨uzbalaban 和 Lingjiong Zhu。随机梯度哈密顿蒙特卡洛在非凸随机优化中的全局收敛：非渐近性能界与基于动量的加速。《运筹学》，70(5)：2931–2947，2022。[GIWZ24] Mert G¨urb¨uzbalaban、Mohammad Rafiqul Islam、Xiaoyu Wang 和 Lingjiong Zhu。广义 EXTRA 随机梯度朗之万动力学。arXiv 预印本 arXiv:2412.01993，2024。[GNZZ25] Mert Gurbuzbalaban、Hoang M Nguyen、Xicheng Zhang 和 Lingjiong Zhu。锚定朗之万算法。arXiv 预印本 arXiv:2509.19455，2025。[GZL+13] Pinghua Gong、Changshui Zhang、Zhaosong Lu、Jianhua Huang 和 Jieping Ye。用于非凸正则化优化问题的一般迭代收缩阈值算法。载于：国际机器学习会议论文集，第37–45页。PMLR，2013。

[HJ12] Roger A Horn 和 Charles R Johnson。《矩阵分析》。剑桥大学出版社，2012。[HKS89] Richard A Holley、Shigeo Kusuoka 和 Daniel W Stroock。谱隙的渐近性质及其在模拟退火理论中的应用。《泛函分析杂志》，83(2)：333–347，1989。[Hof09] Peter D Hoff。《贝叶斯统计方法入门》，第580卷。施普林格，2009。[LZT22] Ruilin Li、Hongyuan Zha 和 Molei Tao。用于采样的无海森矩阵高分辨率内斯捷罗夫加速。载于：国际机器学习会议论文集，第162卷，第13125–13162页。PMLR，2022。[MBM18] Song Mei、Yu Bai 和 Andrea Montanari。非凸损失的经验风险景观。《统计年刊》，46(6A)：2747–2774，2018。[MCC+21] Yi-An Ma、Niladri S. Chatterji、Xiang Cheng、Nicolas Flammarion、Peter L. Bartlett 和 Michael I. Jordan。基于梯度的马尔可夫链蒙特卡罗是否存在类似内斯捷罗夫加速的方法？《Bernoulli》，27(3)：1942–1992，2021。[MSH02] Jonathan C Mattingly、Andrew M Stuart 和 Desmond J Higham。随机微分方程及其近似的遍历性：局部利普希茨向量场与退化噪声。《随机过程及其应用》，101(2)：185–232，2002。[Nes83] Yurii Nesterov。求解凸规划问题的一种方法及其收敛率 $o(1/k^2)$。苏联科学院院刊，269：543–547，1983。[Nes13] Yurii Nesterov。《凸优化导论：基础课程》，第87卷。施普林格科学与商业媒体，2013。[OSK+19] Kazuki Osawa、Siddharth Swaroop、Mohammad Emtiyaz Khan、Anirudh Jain、Runa Eschenhagen、Richard E Turner 和 Rio Yokota。运用贝叶斯原理的实用深度学习。载于：神经信息处理系统进展，第32卷。Curran Associates, Inc.，2019。[PC08] Trevor Park 和 George Casella。《贝叶斯套索》。《美国统计学会杂志》，103(482)：681–686，2008。[PS10] Nicholas G Polson 和 James G Scott。全局收缩，局部作用：稀疏贝叶斯正则化与预测。《贝叶斯统计》，9(501-538)：105，2010。[PSW14] Nicholas G Polson、James G Scott 和 Jesse Windle。《贝叶斯桥》。《英国皇家统计学会杂志B辑：统计方法学》，76(4)：713–733，2014。[RRT17] Maxim Raginsky、Alexander Rakhlin 和 Matus Telgarsky。通过随机梯度朗之万动力学进行非凸学习：非渐近分析。载于：2017年学习理论会议论文集，第65卷，第1674–1703页。PMLR，2017。

[SDJS22] Bin Shi、Simon S. Du、Michael I. Jordan 和 Weijie J. Su。通过高分辨率微分方程理解加速现象。*Mathematical Program-ming*，195：79–148，2022。[Stu10] Andrew M Stuart。逆问题：贝叶斯视角。*Acta Numerica*，19：451–559，2010。[TTV16] Yee Whye Teh、Alexandre H Thiery 和 Sebastian J Vollmer。随机梯度朗之万动力学的一致性与涨落。*Journal of Machine LearningResearch*，17(1)：193–225，2016。[Vil09] C´edric Villani。次强制性。*Memoirs of the American Mathematical Society*，202(950)：iv+141，2009。.[ZADS23] Ying Zhang、Omer Deniz Akyildiz、Theodoros Damoulas 和 Sotirios Sabanis。非凸优化中局部条件下随机梯度朗之万动力学的非渐近估计。*Applied Mathematics & Optimization*，87：25，2023。[ZOL25] Xinzhe Zuo、Stanley Osher 和 Wuchen Li。用于采样的梯度调整欠阻尼朗之万动力学。*SIAM/ASA Journal on Uncertainty Quantification*，13(4)：1735–1765，2025。

第 A 节 结果的证明（第 2 节）

A.1 命题 2.2 的证明

证明。 利用分解（2.4），我们有

$$\mathcal{L}_\alpha \mathcal{V}_0 = \mathcal{L}_0 \mathcal{V}_0 + \alpha\, \mathcal{A}'\mathcal{V}_0 + \alpha\, \Delta_q \mathcal{V}_0. \tag{A.1}$$

由（2.12），我们有

$$\mathcal{L}_0 \mathcal{V}_0(q,p) \;\le\; \gamma\,(d + A - \lambda\, \mathcal{V}_0(q,p))\,. \tag{A.2}$$

接下来需要控制（ $\mathcal{A}'\mathcal{V}_0$ 和 $\Delta_q \mathcal{V}_0$ 中的扰动项（A.1）。

首先，我们希望获得 $\mathcal{A}'\mathcal{V}_0$的显式上界。根据定义（2.7），我们得到

$$\nabla_q \mathcal{V}_0(q,p) = \nabla U(q) + \frac{\gamma^2}{2}\left(q + \gamma^{-1}p\right) - \frac{\gamma^2\lambda}{2}\, q = \nabla U(q) + \frac{\gamma^2}{2}\left((1-\lambda)q + \gamma^{-1}p\right).$$

因此

$$\mathcal{A}'\mathcal{V}_0(q,p) = -\nabla U(q)\cdot \nabla_q \mathcal{V}_0(q,p) = -|\nabla U(q)|^2 - \frac{\gamma^2}{2}\,\nabla U(q)\cdot\left((1-\lambda)q + \gamma^{-1}p\right).$$

利用柯西–施瓦茨不等式和杨氏不等式，结合

$$a := |\nabla U(q)|, \qquad b := \frac{\gamma^2}{2}\left|(1-\lambda)q + \gamma^{-1}p\right|,$$

可得

$$\frac{\gamma^2}{2}\left|\nabla U(q)\right|\left|(1-\lambda)q+\gamma^{-1}p\right| \;\le\; \frac{1}{2}a^2+\frac{1}{2}b^2=\frac{1}{2}|\nabla U(q)|^2+\frac{\gamma^4}{8}\left|(1-\lambda)q+\gamma^{-1}p\right|^2.$$

因此

$$\mathcal{A}'\mathcal{V}_0(q,p) \;\le\; -|\nabla U(q)|^2+\frac{1}{2}|\nabla U(q)|^2+\frac{\gamma^4}{8}\left|(1-\lambda)q+\gamma^{-1}p\right|^2 \;\le\; \frac{\gamma^4}{8}\left|(1-\lambda)q+\gamma^{-1}p\right|^2.$$

接下来，利用

$$\left|(1-\lambda)q+\gamma^{-1}p\right|^2\le 2(1-\lambda)^2|q|^2+2\gamma^{-2}|p|^2,$$

我们得到

$$\mathcal{A}'\mathcal{V}_0(q,p) \;\le\; \frac{\gamma^4}{4}(1-\lambda)^2|q|^2+\frac{\gamma^2}{4}|p|^2.$$

由（2.10）和 $U\ge 0$，对于某个 $c_1>0$，

$$c_1\left(1+|q|^2+|p|^2\right) \;\le\; 1+\mathcal{V}_0(q,p),$$

这意味着

$$|q|^2+|p|^2 \;\le\; \frac{1}{c_1}\left(1+\mathcal{V}_0(q,p)\right).$$

因此

$$\mathcal{A}'\mathcal{V}_0(q,p)\le\left[\frac{\gamma^4}{4}(1-\lambda)^2+\frac{\gamma^2}{4}\right]\left(|q|^2+|p|^2\right)\le\frac{1}{c_1}\left[\frac{\gamma^4}{4}(1-\lambda)^2+\frac{\gamma^2}{4}\right]\left(1+\mathcal{V}_0(q,p)\right). \tag{A.3}$$

通过引入

$$K_A:=\frac{1}{c_1}\left[\frac{\gamma^4}{4}(1-\lambda)^2+\frac{\gamma^2}{4}\right], \tag{A.4}$$

我们得出

$$\mathcal{A}'\mathcal{V}_0(q,p) \;\le\; K_A\left(1+\mathcal{V}_0(q,p)\right). \tag{A.5}$$

接下来，我们推导出 $\Delta_q\mathcal{V}_0$的一个显式界。再次由（2.7）可得，

$$\Delta_q\mathcal{V}_0(q,p)=\Delta U(q)+\frac{\gamma^2}{2}d(1-\lambda).$$

由于 ∇U 是常数为 L的全局 Lipschitz 函数， U 的海森矩阵的算子范数几乎处处有界于 L，因此 $|\Delta U(q)|\le Ld$。所以

$$|\Delta_q\mathcal{V}_0(q,p)| \;\le\; Ld+\frac{\gamma^2}{2}d(1-\lambda)=:K_\Delta. \tag{A.6}$$

利用 $1+\mathcal{V}_0\ge 1$，由（A.6）可得

$$|\Delta_q\mathcal{V}_0(q,p)| \;\le\; K_\Delta\left(1+\mathcal{V}_0(q,p)\right), \qquad K_\Delta:=Ld+\frac{\gamma^2}{2}d(1-\lambda). \tag{A.7}$$

结合（A.5）和（A.7），我们得到

$$\mathcal{A}'\mathcal{V}_0(q,p) + \Delta_q \mathcal{V}_0(q,p) \;\le\; J_1\left(1 + \mathcal{V}_0(q,p)\right), \tag{A.8}$$

其中 J_1 定义见（2.15）。

因此，结合（A.2）中 $\mathcal{L}_0$ 的界和（A.8）中的扰动项：

$$\begin{aligned}\mathcal{L}_\alpha \mathcal{V}_0(q,p) &\le \gamma\left(d + A - \lambda \mathcal{V}_0(q,p)\right) + \alpha J_1\left(1 + \mathcal{V}_0(q,p)\right) \\ &= \gamma(d+A) - \gamma\lambda\mathcal{V}_0(q,p) + \alpha J_1 + \alpha J_1 \mathcal{V}_0(q,p).\end{aligned} \tag{A.9}$$

为了将其写成标准漂移形式 $\gamma(d + A_\alpha - \lambda_\alpha \mathcal{V}_0)$，我们将常数项与$\mathcal{V}_0$ 项分组。将整个表达式提取出因子 γ，得到：

$$\begin{aligned}\mathcal{L}_\alpha \mathcal{V}_0(q,p) &\le \left[\gamma(d+A) + \alpha J_1\right] - \left[\gamma\lambda - \alpha J_1\right]\mathcal{V}_0(q,p) \\ &= \gamma\left(d + A + \frac{J_1}{\gamma}\alpha\right) - \gamma\left(\lambda - \frac{J_1}{\gamma}\alpha\right)\mathcal{V}_0(q,p).\end{aligned}$$

这与所需的不等式（2.13）一致，其中 $A_\alpha := A + \frac{J_1}{\gamma}\alpha$ 且 $\lambda_\alpha := \lambda - \frac{J_1}{\gamma}\alpha$。通过代入 λ_α 的表达式，即可直接得到其显式的 展开式以及 α_0 的取值。具体而言，为确保 $\lambda_\alpha \ge \lambda/2$，我们要求 $\frac{J_1}{\gamma}\alpha \le \frac{\lambda}{2}$，这等价于 $\alpha \le \frac{\gamma\lambda}{2J_1}$，对应于（2.17）中 α_0 的定义。证明完毕。 □

B 第3节结果的证明

B.1 引理3.3的证明

证明。令 $\Delta z := z - z' = (\Delta q, \Delta p)$。我们用 $|\cdot|$ 表示 $\mathbb{R}^{2d}$ 上的标准欧几里得范数。注意，$|\Delta z| \le |\Delta q| + |\Delta p|$ 和 $|\Delta z|^2 = |\Delta q|^2 + |\Delta p|^2$。

上界（k_2）：根据三角不等式以及 $r(z,z')$ 的定义，我们有

$$r(z,z') \le \theta|\Delta q| + |\Delta q| + \gamma^{-1}|\Delta p| = (\theta+1)|\Delta q| + \gamma^{-1}|\Delta p|.$$

应用柯西 –施瓦茨不等式于向量 $((\theta+1)\ \gamma^{-1})$ 和 $(|\Delta q|, |\Delta p|)$，得到

$$r(z,z') \le \sqrt{(\theta+1)^2 + \gamma^{-2}}\sqrt{|\Delta q|^2 + |\Delta p|^2} = k_2\,|\Delta z|.$$

下界（k_1 ）：根据 $r(z,z')$ 的定义，我们可以立即对 Δq 进行显式控制：

$$|\Delta q| \le \frac{1}{\theta} r(z,z'). \tag{B.1}$$

为控制 Δp，我们将其改写为 $\Delta p = \gamma\left((\Delta q + \gamma^{-1}\Delta p) - \Delta q\right)$。利用三角不等式：

$$|\Delta p| \le \gamma\left(\left|\Delta q + \gamma^{-1}\Delta p\right| + |\Delta q|\right).$$

由于 $\left|\Delta q + \gamma^{-1}\Delta p\right| \le r(z,z')$（在 r 的定义中去掉第一个非负项，并使用（附录 B.1），我们得到

$$|\Delta p| \le \gamma\left(r(z,z') + \frac{1}{\theta}r(z,z')\right) = \frac{\gamma(1+\theta)}{\theta} r(z,z'). \tag{B.2}$$

最后，利用基本不等式 $|\Delta z| \le |\Delta q| + |\Delta p|$，将（附录 B.1）和（附录 B.2）相加：

$$|\Delta z| \le \left(\frac{1}{\theta} + \frac{\gamma(1+\theta)}{\theta}\right) r(z,z') = \frac{1+\gamma(1+\theta)}{\theta} r(z,z').$$

重新整理可得

$$r(z,z') \ge \frac{\theta}{1+\gamma(1+\theta)} |\Delta z| = k_1 \, |\Delta z|.$$

至此，证明完毕。 □

附录 B.2 证明引理3.5

证明。固定 $\varepsilon > 0$ 和 $c \in \mathbb{R}$。回顾适用于 $z_t = (q_t, p_t)$ 和$z'_t = (q'_t, p'_t)$ 的耦合 HFHR 动力学 (2.1)，并回顾 (3.20) 中的 $Z_t := q_t - q'_t$, $W_t := p_t - p'_t$ 和 $\mathrm{R}_t := Z_t + \gamma^{-1} W_t$。令$e_t :=\mathrm{R}_t/|\mathrm{R}_t|$ （若 $\mathrm{R}_t \neq 0$），否则令其为任意单位向量，并令 $\mathcal{P}_t := e_t e_t^\top$。回顾 (3.3)，耦合定义为

$$dB_t^{q'} = dB_t^q, \qquad dB_t^{p'} = (I_d - 2\chi(t)\mathcal{P}_t)\, dB_t^p,$$

其中控制过程为 $\chi(t) \in \{0,1\}$。

步骤 *1*：差分动力学与 $\mathbf{R}_t$ 的噪声。 由（2.1），

$$dZ_t = \big(W_t - \alpha(\nabla U(q_t) - \nabla U(q'_t))\big)\, dt,$$

and

$$dW_t = \big(-\gamma W_t - (\nabla U(q_t) - \nabla U(q'_t))\big)\, dt + \sqrt{2\gamma}\, d\widetilde{B}_t, \qquad d\widetilde{B}_t := dB_t^p - dB_t^{p'} = 2\chi(t)\mathcal{P}_t\, dB_t^p.$$

因此

$$d\mathbf{R}_t = -\frac{1+\alpha\gamma}{\gamma} \left(\nabla U(q_t) - \nabla U(q'_t)\right) dt + 2\sqrt{2}\, \gamma^{-1/2} \chi(t) \mathcal{P}_t\, dB_t^p.$$

由于 $\mathcal{P}_t$ 投影到 $\mathrm{span}\{e_t\}$ 上，噪声沿着方向 e_t作用，因此 $d|\mathbf{R}_t|$中的 Itô 修正项为零。特别地，

$$d|\mathbf{R}_t||_{\text{noise}} = \langle e_t,\, d\mathbf{R}_t\rangle_{\text{noise}} = 2\sqrt{2}\, \gamma^{-1/2} \chi(t)\, \langle e_t, dB_t^p\rangle, \qquad d\langle |\mathbf{R}|\rangle_t = 8\, \gamma^{-1}\, (\chi(t))^2\, dt.$$

步骤 *2*：的漂移界 r_t。 回顾式（3.20），可得

$$r_t := \theta |Z_t| + |\mathbf{R}_t|, \qquad \theta := (1+\eta_0) L_{\text{eff}}(\alpha) \gamma^{-2}.$$

回顾式（ δ_α 3.18）中 $\delta_\alpha := \frac{\eta_0}{1+\eta_0} - \frac{\alpha L}{\gamma}$ 的定义。在整个证明过程中，我们假设$\delta_\alpha > 0$ （等价地，$\alpha < \frac{\eta_0}{1+\eta_0} \frac{\gamma}{L}$）。

利用 $dZ_t = (W_t - \alpha(\nabla U(q_t) - \nabla U(q'_t)))\, dt$ 以及单边利普希茨界

$$\left\langle \frac{Z_t}{|Z_t|},\, \nabla U(q_t) - \nabla U(q'_t) \right\rangle \le L|Z_t|,$$

结合针对 W_t–贡献的标准动能估计（取 $\theta = (1+\eta_0)L_{\text{eff}}(\alpha)\gamma^{-2}$），我们得到有限变差不等式

$$dr_t \;\le\; \gamma\,(\theta|\mathbf{R}_t| - \delta_\alpha\,\theta|Z_t|)\;dt + dM_t^{(r)}, \tag{B.3}$$

其中，连续局部鞅 $M^{(r)}$ 为

$$M_t^{(r)} := 2\sqrt{2}\,\gamma^{-1/2}\int_0^t \chi(s)\,\langle e_s, dB_s^p\rangle.$$

此外，

$$d\langle r\rangle_t = d\langle|\mathbf{R}|\rangle_t = 8\,\gamma^{-1}\,(\chi(t))^2\;dt. \tag{B.4}$$

（恒等式（B.4）由以下事实可得： Z_t 在我们的耦合下没有噪声。）

步骤 *3*：对 $f_\lambda(r_t)$ 应用 *Meyer–Itˆo* 公式。令 f_λ 为（3.9）–（3.12）中的凹轮廓函数。利用 Meyer–Itˆo 公式，我们得到

$$df_\lambda(r_t) = f'_{\lambda,-}(r_t)\,dr_t + \frac{1}{2}f''_\lambda(r_t)\,d\langle r\rangle_t + dM_t^{(f)}, \tag{B.5}$$

其中有一个连续局部鞅

$$M_t^{(f)} := 2\sqrt{2}\,\gamma^{-1/2}\int_0^t f'_{\lambda,-}(r_s)\chi(s)\,\langle e_s, dB_s^p\rangle.$$

将（B.3）–（B.4）代入（B.5）得

$$df_\lambda(r_t) \;\le\; \gamma\left[4\gamma^{-2}\,(\chi(t))^2\,f''_\lambda(r_t) + (\theta|\mathbf{R}_t| - \delta_\alpha\,\theta|Z_t|)\,f'_{\lambda,-}(r_t)\right]\,dt + dM_t^{(f)}. \tag{B.6}$$

步骤 *4*： G_t 的动力学和乘积法则。设 $\mathcal{V}$ 是（λ，D）-可容许的，并由（3.21）回忆可知， $G_t := 1 + \varepsilon\mathcal{V}(z_t) + \varepsilon\mathcal{V}(z'_t)$ 且 $\rho_t := f_\lambda(r_t)\,G_t$。由伊藤公式，

$$dG_t = \varepsilon\left(\mathcal{L}_\alpha\mathcal{V}(z_t) + \mathcal{L}_\alpha\mathcal{V}(z'_t)\right)\,dt + dM_t^{(G)}, \tag{B.7}$$

其中， $M^{(G)}$ 是如下连续局部鞅：

$$\begin{aligned} M_t^{(G)} := {}& \varepsilon\sqrt{2\alpha}\int_0^t \left\langle \nabla_q\mathcal{V}(z_s) + \nabla_q\mathcal{V}(z'_s),\, dB_s^q\right\rangle \\ & + \varepsilon\sqrt{2\gamma}\int_0^t \langle \nabla_p\mathcal{V}(z_s),\, dB_s^p\rangle + \varepsilon\sqrt{2\gamma}\int_0^t \left\langle \nabla_p\mathcal{V}(z'_s),\, dB_s^{p'}\right\rangle. \end{aligned} \tag{B.8}$$

利用耦合关系 $dB_t^{p'} = (I_d - 2\chi(t)\mathcal{P}_t)\,dB_t^p$， p-噪声部分等价地可写为

$$\varepsilon\sqrt{2\gamma}\int_0^t \left\langle \nabla_p\mathcal{V}(z_s) + (I_d - 2\chi(s)\mathcal{P}_s)^\top\nabla_p\mathcal{V}(z'_s),\, dB_s^p\right\rangle. \tag{B.9}$$

将伊藤乘积法则应用于 $e^{ct}\rho_t = e^{ct}f_\lambda(r_t)G_t$，得到

$$d(e^{ct}\rho_t) = ce^{ct}\rho_t\,dt + e^{ct}G_t\,df_\lambda(r_t) + e^{ct}f_\lambda(r_t)\,dG_t + e^{ct}\,d\langle f_\lambda(r), G\rangle_t. \tag{B.10}$$

将（B.6）和 dG_t （B.7）中的表达式代入（B.10），得到

$$d(e^{ct}\rho_t) \le e^{ct}\gamma\Big[4\gamma^{-2}\left(\chi(t)\right)^2 f_\lambda''(r_t)G_t + \left(\theta|\mathbf{R}_t| - \delta_\alpha\,\theta|Z_t|\right) f_{\lambda,-}'(r_t)G_t \\ + \gamma^{-1}\varepsilon f_\lambda(r_t)\left(\mathcal{L}_\alpha\mathcal{V}(z_t) + \mathcal{L}_\alpha\mathcal{V}(z_t')\right) + \gamma^{-1}c f_\lambda(r_t)G_t\Big]dt \\ + e^{ct}\,d\langle f_\lambda(r), G\rangle_t + dM_t, \tag{B.11}$$

其中 M_t 是连续局部鞅

$$M_t := \int_0^t e^{cs}G_s\,dM_s^{(f)} + \int_0^t e^{cs}f_\lambda(r_s)\,dM_s^{(G)}. \tag{B.12}$$

步骤 5：界定交叉变差项。只有 p 分量中的噪声会对交叉变差 $d\langle f_\lambda(r)\ G\rangle_t$ 产生贡献。利用耦合关系 $dB_t^{p'} = (I_d - 2\chi(t)\mathcal{P}_t)\ dB_t^p$ 以及鞅部分的显式表达式，直接计算可得：

$$d\langle f_\lambda(r), G\rangle_t = 4\varepsilon(\chi(t))^2 f_{\lambda,-}'(r_t)\langle e_t, \nabla_p\mathcal{V}(z_t) - \nabla_p\mathcal{V}(z_t')\rangle\,dt. \tag{B.13}$$

我们利用以下结构来估计梯度差 $\mathcal{V} = \mathcal{V}_0 + \mathfrak{Q}$。首先，考虑基准函数 $\mathcal{V}_0$。由（2.7）可知， $\mathcal{V}_0$ 关于 p 的梯度是线性的：

$$\nabla_p\mathcal{V}_0(q,p) = p + \frac{\gamma}{2}q.$$

因此，差为

$$\nabla_p\mathcal{V}_0(z_t) - \nabla_p\mathcal{V}_0(z_t') = \Delta p_t + \frac{\gamma}{2}\Delta q_t. \tag{B.14}$$

回顾 $\mathbf{R}_t = \Delta q_t + \gamma^{-1}\Delta p_t$，这意味着 $\Delta p_t = \gamma(\mathbf{R}_t - \Delta q_t)$。代回此式 ：

$$\Delta p_t + \frac{\gamma}{2}\Delta q_t = \gamma(\mathbf{R}_t - \Delta q_t) + \frac{\gamma}{2}\Delta q_t = \gamma\mathbf{R}_t - \frac{\gamma}{2}\Delta q_t. \tag{B.15}$$

取 t 对（B.14）-（B.15）中的范数进行处理，并与距离 $r_t = \theta|\Delta q_t| + |\mathbf{R}_t|$ ：

$$|\nabla_p\mathcal{V}_0(z_t) - \nabla_p\mathcal{V}_0(z_t')| \le \gamma|\mathbf{R}_t| + \frac{\gamma}{2}|\Delta q_t| \le \gamma\max\left\{1, \frac{1}{2\theta}\right\}(|\mathbf{R}_t| + \theta|\Delta q_t|) = \gamma\max\left\{1, (2\theta)^{-1}\right\}r_t. \tag{B.16}$$

接下来，对于扰动项 $\mathfrak{Q}(z) = \frac{1}{2}z^\top\mathsf{A}z$，我们显式计算 p-梯度。相对于 $z = (q,p)$，将 A 写成分块形式，

$$\mathsf{A} = \begin{pmatrix}\mathsf{A}_{qq} & \mathsf{A}_{qp}\\ \mathsf{A}_{pq} & \mathsf{A}_{pp}\end{pmatrix}, \qquad \mathsf{A}_{qp} = \mathsf{A}_{pq}^\top,$$

我们有

$$\nabla_p\mathfrak{Q}(q,p) = \mathsf{A}_{pq}\,q + \mathsf{A}_{pp}\,p.$$

因此

$$\left|\nabla_p\mathfrak{Q}(z_t) - \nabla_p\mathfrak{Q}(z_t')\right| \le \|\mathsf{A}_{pq}\|_{\mathrm{op}}\,|\Delta q_t| + \|\mathsf{A}_{pp}\|_{\mathrm{op}}\,|\Delta p_t| \le \left(\|\mathsf{A}_{pq}\|_{\mathrm{op}} + \|\mathsf{A}_{pp}\|_{\mathrm{op}}\right)|z_t - z_t'|.$$

利用范数等价 $r_t \geq k_1|z_t - z_t'|$ （引理 3.3），我们得到

$$|\nabla_p \mathfrak{Q}(z_t) - \nabla_p \mathfrak{Q}(z_t')| \leq \frac{C_{\mathfrak{Q}}}{k_1}\, r_t, \qquad C_{\mathfrak{Q}} := \|\mathsf{A}_{pp}\|_{\mathrm{op}} + \|\mathsf{A}_{pq}\|_{\mathrm{op}}. \tag{B.17}$$

合并估计式（B.16）和（B.17）可得

$$|\nabla_p \mathcal{V}(z_t) - \nabla_p \mathcal{V}(z_t')| \leq \left(\gamma \max\left\{1, (2\theta)^{-1}\right\} + \frac{C_{\mathfrak{Q}}}{k_1}\right) r_t = \gamma \bar{C}_{\mathcal{V}}\, r_t.$$

将此界代入交叉变差项（B.13）可得

$$|d\langle f_\lambda(r), G\rangle_t| \leq 4\gamma\, \varepsilon \bar{C}_{\mathcal{V}}\, (\chi(t))^2\; r_t\, f'_{\lambda,-}(r_t)\, dt.$$

这与引理 3.5 中所述的形式一致。

将这一项并入（B.11）中的漂移项后，我们得到

$$e^{ct}\rho_t \leq \rho_0 + \gamma \int_0^t e^{cs} K_s\, ds + M_t,$$

其中 M_t 是（B.12）中定义的连续局部鞅，而 K_t 满足（3.23）。 □

B.3 命题 3.7 的证明

证明。固定 $\xi > 0$ 并简记 $(z_t, z_t') = \left(z_t^{\xi}, z_t'^{,\xi}\right)$, $r_t = r(z_t, z_t')$, $G_t = 1 + \varepsilon\mathcal{V}(z_t) + \varepsilon\mathcal{V}(z_t')$,$\rho_t = f_\lambda(r_t)G_t$, and $\chi(t) = \chi_\xi(t)$. 始终假设 (3.24) 成立，因此$\delta_\alpha \geq \kappa_{\text{adjust}} \frac{\eta_0}{1+\eta_0} > 0$。由引理 3.5，对任意 $c \in \mathbb{R}$,

$$e^{ct}\rho_t \leq \rho_0 + \gamma \int_0^t e^{cs} K_s\, ds + M_t, \tag{B.18}$$

其中 M_t 是一个连续局部鞅，而 K_t 由引理 3.5中的右端给出上界。我们在两个区域 K_t 和 $r_t \leq R_1(\lambda)$ 以及 $r_t > R_1(\lambda)$ 上界定

1）区域 $r_t \leq R_1(\lambda)$。进一步划分为事件 $\{|\mathbf{R}_t| \geq \xi\}$（反射激活）和 $\{|\mathbf{R}_t| < \xi\}$（反射未激活）。

(i) 若 $r_t \leq R_1(\lambda)$ 且 $|\mathbf{R}_t| \geq \xi$，则 $\chi(t) = 1$。在 $(0, R_1(\lambda))$ 上，f_λ 是 C^2。此外，f_λ的构造（在 φ_λ中取 3.9）确保了 f''_λ-项、“坏”的线性漂移项以及交叉变差贡献之和严格为负。更准确地说，存在$C_{\text{conc}} > 0$（仅取决于轮廓构造），使得对于该事件上几乎处处的 t，

$$4\gamma^{-2} f''_\lambda(r_t)\, G_t + (\theta|\mathbf{R}_t| - \delta_\alpha\, \theta|Z_t|)\, f'_{\lambda,-}(r_t)\, G_t + 4\varepsilon\bar{C}_{\mathcal{V}}\, r_t f'_{\lambda,-}(r_t) \leq -C_{\text{conc}}\, f_\lambda(r_t)\, G_t.$$

利用 (λ, D)-可容许性，并将 $0 < c \leq c_0$ 和 $0 < \varepsilon \leq \varepsilon_0$ 取得足够小（使得剩余的 ε-项和 c-项受到控制），即可得到此处的 $K_t \leq 0$。特别地，$K_t \leq C_{\text{reg}}\xi G_t$ 成立。

(ii) 若 $r_t \leq R_1(\lambda)$ 且 $|\mathbf{R}_t| < \xi$，则 $\chi(t) = 0$。在这种情况下，f''_λ 项和 $(\chi(t))^2 r_t f'_{\lambda,-}(r_t)$ 项均消失。此外，由于 $|\mathbf{R}_t| < \xi$ 且 $r_t \leq R_1(\lambda)$，我们得到粗略上界

$$(\theta|\mathbf{R}_t| - \delta_\alpha\, \theta|Z_t|)\, f'_{\lambda,-}(r_t)\, G_t \leq \theta\, |\mathbf{R}_t| \sup_{[0,R_1(\lambda)]} f'_{\lambda,-}\, G_t \leq C\, \xi\, G_t.$$

其中 C 与 ξ 无关。其余的 Lyapunov 项和 c 项可由 G_t 的常数倍界定（因为 f_λ 在 $[0, R_1(\lambda)]$ 上有界），因此增大 C_{reg} 后，可将它们吸收到 $C_{\mathrm{reg}}\xi G_t$ 中。因此，在此事件上，$K_t \le C_{\mathrm{reg}}\xi G_t$ 也成立。

2) 区域 $r_t > R_1(\lambda)$。根据构造，f_λ 在 $[R_1(\lambda)\ \infty)$ 上为常数，因此 $f'_{\lambda,\ -}(r_t) = 0$ 以及 $f''_\lambda(r_t) = 0$ 在 $\{r_t > R_1(\lambda)\}$ 上几乎处处成立。因此，引理 3.5 可归结为

$$K_t \le \gamma^{-1}\varepsilon f_\lambda(r_t)\left[\mathcal{L}_\alpha \mathcal{V}(z_t) + \mathcal{L}_\alpha \mathcal{V}(z'_t)\right] + \gamma^{-1} c f_\lambda(r_t) G_t.$$

利用 $(\lambda,\ D)$-可容许性，

$$\mathcal{L}_\alpha \mathcal{V}(z_t) + \mathcal{L}_\alpha \mathcal{V}(z'_t) \le 2\gamma(d+D) - \gamma\lambda\left(\mathcal{V}(z_t) + \mathcal{V}(z'_t)\right).$$

Thus

$$K_t \le f_\lambda(r_t)\left[2\varepsilon(d+D) + \gamma^{-1}c + \varepsilon(\gamma^{-1}c - \lambda)\left(\mathcal{V}(z_t) + \mathcal{V}(z'_t)\right)\right].$$

选择 $c_0 < \gamma\lambda$ 使得 $\gamma^{-1}c - \lambda < 0$。由于根据引理 3.3，$r_t > R_1(\lambda)$ 蕴含 $|z_t - z'_t| \gtrsim r_t$，因此 $|z_t|, |z'_t|$ 中至少有一个是 $\gtrsim r_t$，而强制性 (3.25) 给出 $\mathcal{V}(z_t) + \mathcal{V}(z'_t) \gtrsim r_t^2$。将 $R_1(\lambda)$（在构造中已是一个自由截断参数）取得足够大，负项便占主导，于是我们得到 $K_t \le 0$ 在 $\{r_t > R_1(\lambda)\}$ 上成立，从而再次得到 $K_t \le C_{\mathrm{reg}}\xi G_t$。

合并两个区域得到（3.26）。对（B.18）取期望（通过局部化消去 M_t），得到（3.27）。最后，对于每个固定的 t，$\sup_{s\le t}\mathbb{E}[G_s^\xi] < \infty$ 且不会随 $\xi \downarrow 0$ 增大而发散（边缘分布对应相同的 HFHR 动力学）。因此令 $\xi \downarrow 0$，可得 $\limsup_{\xi\downarrow 0}\mathbb{E}[e^{ct}\rho_t^\xi] \le \mathbb{E}[\rho_0]$。证明完毕。

□

B.4 定理 3.8 的证明

证明。令 $(Z_t,\ Z'_t)$ 为引理 3.5 （反射／同步切换）中使用的耦合，并设 $\rho_t := \rho_{\mathcal{V}}(Z_t,\ Z'_t)$，$r_t := r(Z_t,\ Z'_t)$。根据引理 3.5，对于任意 $c > 0$，只要 K_t 中的漂移项 $d(e^{ct}\rho_t) = e^{ct}K_t\,dt + dM_t$ 满足 $K_t \le 0$ a.s.，过程 $e^{ct}\rho_t$ 就是一个超鞅。我们取

$$\varepsilon := \frac{4c}{\gamma(d+D)}. \tag{B.19}$$

我们在三个区域中验证 $K_t \le 0$。

区域 I： $r_t \ge R_1(\lambda)$ **（大距离）**。由于 f_λ 在 $[R_1(\lambda)\ \infty)$ 上为常数，我们有 $f'_\lambda = f''_\lambda = 0$。利用 $(\lambda,\ D)$-可容许性，

$$\mathcal{L}_\alpha \mathcal{V} \le \gamma(d+D) - \gamma\lambda\mathcal{V}.$$

选择 $R_1(\lambda)$ 于（3.28）意味着，只要 $r_t \ge R_1(\lambda)$，

$$\mathcal{V}(Z_t) + \mathcal{V}(Z'_t) \ \ge\ \frac{12}{5}\frac{d+D}{\lambda}, \tag{B.20}$$

因此

$$\mathcal{L}_\alpha \mathcal{V}(Z_t) + \mathcal{L}_\alpha \mathcal{V}(Z'_t) \ \le\ -\frac{1}{6}\gamma\lambda\left(\mathcal{V}(Z_t) + \mathcal{V}(Z'_t)\right). \tag{B.21}$$

利用（附录 B.21）以及 $\varepsilon = 4c/(\gamma(d+D))$ 如（附录 B.19）中所述，我们在该区域内得到 $K_t \le 0$，条件是

$$c \le \frac{\gamma}{16}\lambda.$$

区域 II： $r_t < R_1(\lambda)$ **且反射处于激活状态。**在此事件下，$\chi(t) = 1$。在引理 3.5中，该项

$$-\delta_\alpha\,\theta|Z_t|\,f'_{\lambda,-}(r_t)\,G_t$$

是非正的，因此可以省略。φ_λ 的构造确保对所有 $r \in (0, R_1(\lambda))$，以下抵消条件成立：

$$4\gamma^{-2}\varphi'_\lambda(r) + \left(\theta + 4\varepsilon\bar{C}_{\mathcal{V}}\right) r\,\varphi_\lambda(r) \le 0. \tag{B.22}$$

注意，该条件使用距离 r 定义，以涵盖最坏情况漂移，因为 $|\mathbf{R}_t| \le r_t$。回忆引理3.5中方程（3.23）给出的 K_t 界。由于 $G_t \ge 1$ 且 $f'_{\lambda,-}(r_t) = \varphi_\lambda(r_t)g_\lambda(r_t) \ge 0$，我们可以将交叉变差项乘以 G_t，从而得到其上界：

$$4\varepsilon\bar{C}_{\mathcal{V}} r_t f'_{\lambda,-}(r_t) \le 4\varepsilon\bar{C}_{\mathcal{V}} r_t f'_{\lambda,-}(r_t) G_t.$$

由于 $r_t < R_1(\lambda)$ $r \in (0, R_1(\lambda))$ $f'_{\lambda,-}(r) = \varphi_\lambda(r)\,g_\lambda(r)$ $f''_\lambda(r) = \varphi'_\lambda(r)\,g_\lambda(r) + (\,)$，对几乎处处的 $(\,,\,(\,))$，有 $(\,)(\,)(\,)$ 且 $(\,)(\,)(\,)(\,)(\,)$。将这些恒等式代入（。），得到 3.23

$$\begin{aligned} K_t \le{}& \left[4\gamma^{-2}\varphi'_\lambda(r_t) + \theta|\mathbf{R}_t|\,\varphi_\lambda(r_t) + 4\varepsilon\bar{C}_{\mathcal{V}}\,r_t\,\varphi_\lambda(r_t)\right] g_\lambda(r_t)\,G_t \\ &+ 4\gamma^{-2}\varphi_\lambda(r_t)g'_\lambda(r_t)G_t + \gamma^{-1}\varepsilon f_\lambda(r_t)\left[\mathcal{L}_\alpha\mathcal{V}(z_t) + \mathcal{L}_\alpha\mathcal{V}(z'_t)\right] + \gamma^{-1}c f_\lambda(r_t)G_t. \end{aligned}$$

利用 $|\mathbf{R}_t| \le r_t$，括号项的上界为

$$\left[4\gamma^{-2}\varphi'_\lambda(r_t) + (\theta + 4\varepsilon\bar{C}_{\mathcal{V}})\,r_t\,\varphi_\lambda(r_t)\right] g_\lambda(r_t),$$

根据（B.22），该项非正，因此可以舍去。对于剩余项，根据（λ，D）-可容许性（3.7），

$$\mathcal{L}_\alpha\mathcal{V}(z) \le \gamma(d + D - \lambda\mathcal{V}(z)) \le \gamma(d + D),$$

因此 $\mathcal{L}_\alpha\mathcal{V}(z_t) + \mathcal{L}_\alpha\mathcal{V}(z'_t) \le 2\gamma(d + D)$。结合 $\varepsilon = \frac{4c}{\gamma(d+D)}$，可得

$$\gamma^{-1}\varepsilon f_\lambda(r_t)\cdot 2\gamma(d + D) = 8\gamma^{-1}c\,f_\lambda(r_t) \le 8\gamma^{-1}c\,f_\lambda(r_t)\,G_t,$$

因此

$$K_t \le 4\gamma^{-2}\varphi_\lambda(r_t)\,g'_\lambda(r_t)\,G_t + 9\gamma^{-1}c\,f_\lambda(r_t)\,G_t.$$

根据 g_λ 在（3.11）中的定义，

$$4\gamma^{-2}\varphi_\lambda(r)\,g'_\lambda(r) = -9\gamma^{-1}c\,\Phi_\lambda(r),$$

并且由于 $f_\lambda(r) \le \Phi_\lambda(r)$ 对于 $r \in [0, R_1(\lambda)]$，我们得出 $K_t \le 0$，只要 $g_\lambda(r)$ 在 $[0, R_1(\lambda)$ 上 $\ge 1/2]$,，也就是说

$$\frac{9}{4}\,c\,\gamma\int_0^{R_1(\lambda)}\frac{\Phi_\lambda(s)}{\varphi_\lambda(s)}\,ds \;\le\; \frac{1}{2}.$$

与 [EGZ19中的定理 2 一样，3], 上述结论在以下条件下成立

$$c \;\le\; \frac{\gamma}{384}\min\left\{\sqrt{\Lambda_\alpha(\lambda)}\,e^{-\Lambda_\alpha(\lambda)}\frac{L_{\text{eff}}(\alpha)}{\gamma^2},\;\sqrt{\Lambda_\alpha(\lambda)}\,e^{-\Lambda_\alpha(\lambda)}\right\}.$$

区域 III： $r_t < R_1(\lambda)$ **且同步耦合处于激活状态。** 在此情形下，$\chi(t) = 0$ 且反射噪声项消失。引理 3.5 中的漂移界包含耗散部分

$$-\delta_\alpha\,\theta|Z_t|\,f'_\lambda(r_t)\,G_t,$$

从而得到约束条件

$$c \;\le\; \frac{\gamma}{18}\,\delta_\alpha \inf_{s\in(0,R_1(\lambda)]} \frac{s\,\varphi_\lambda(s)}{\Phi_\lambda(s)}.$$

利用 $\delta_\alpha \ge \kappa_{\text{adjust}}\,\frac{\eta_0}{1+\eta_0}$ 并选择 $\eta_0 = (\Lambda_0(\lambda))^{-1}$，我们对高斯比值估计如下。由于 $s \mapsto s\varphi_\lambda(s)/\Phi_\lambda(s)$ 在 $(0,\ R_1(\lambda)]$ 上递减

$$\inf_{s\in(0,R_1(\lambda)]} \frac{s\,\varphi_\lambda(s)}{\Phi_\lambda(s)} = \frac{R_1(\lambda)\,\varphi_\lambda(R_1(\lambda))}{\Phi_\lambda(R_1(\lambda))}.$$

此外，$\Phi_\lambda(R_1(\lambda)) \le \int_0^\infty \varphi_\lambda(s)\,ds = \frac{\sqrt{\pi}}{2}\Big(\mathcal{L}_{\text{eff}}\,\alpha^{8}()\,R_1^2(\lambda)\Big)^{1/2}$，因此

$$\inf_{s\in(0,R_1(\lambda)]} \frac{s\,\varphi_\lambda(s)}{\Phi_\lambda(s)} \;\ge\; \frac{2}{\sqrt{\pi}}\,\sqrt{\Lambda_\alpha(\lambda)}\,e^{-\Lambda_\alpha(\lambda)}.$$

因此，只需施加

$$c \;\le\; \frac{\gamma}{18}\,\delta_\alpha\,\frac{2}{\sqrt{\pi}}\,\sqrt{\Lambda_\alpha(\lambda)}\,e^{-\Lambda_\alpha(\lambda)}.$$

取三个区域中可接受界限的最小值，得到（3.32），从而得到收缩估计。

□

B.5 推论 3.9的证明

证明。 根据命题 2.2，对于 $\alpha \in [0,\alpha_0]$，我们有

$$\mathcal{L}_\alpha \mathcal{V}_0 \le \gamma\,(d + A_\alpha - \lambda_\alpha \mathcal{V}_0)\,.$$

因此，$\mathcal{V}_0$ 按定义 3.1 的意义是 $(\lambda_\alpha, A_\alpha)$-可容许的。固定 $\kappa_{\text{adjust}} \in (0,1)$，并假设 (3.19) 成立（其中 $\eta_0 = (\Lambda_0(\lambda))^{-1}$ 取为定理 3.8 中的选择）。因此，定理 3.8 适用于 $\mathcal{V} = \mathcal{V}_0$，并对相应的半度量 $\rho_{\mathcal{V}_0,\alpha}$ 给出收缩估计

$$\mathcal{W}_{\rho_{\mathcal{V}_0,\alpha}}(\mu P_t^\alpha, \nu P_t^\alpha) \le e^{-c_\alpha t}\,\mathcal{W}_{\rho_{\mathcal{V}_0,\alpha}}(\mu,\nu), \qquad t \ge 0.$$

接下来，我们推导不变测度的存在性与唯一性，以及向其指数收敛。令

$$\mathcal{P}_{\mathcal{V}_0}(\mathbb{R}^{2d}) := \left\{\mu \text{ probability measure on } \mathbb{R}^{2d} : \int_{\mathbb{R}^{2d}} \mathcal{V}_0\,d\mu < \infty\right\},$$

配备 $\mathcal{W}_{\rho_{\mathcal{V}_0,\ \alpha}}$ 与 [EGZ19 中的推论 2.6]，$\left(\mathcal{P}_{\mathcal{V}_0}(\mathbb{R}^{2d})\mathcal{W}_{\rho_{\mathcal{V}_0,\ \alpha}}\right)$ 是完备的，且李雅普诺夫漂移蕴含沿半群的矩控制：对于 $\mu \in \mathcal{P}_{\mathcal{V}_0}$，

$$\sup_{t\ge 0}\int_{\mathbb{R}^{2d}} \mathcal{V}_0\,d(\mu P_t^\alpha) \le \max\left\{\int_{\mathbb{R}^{2d}} \mathcal{V}_0\,d\mu,\ \frac{d+A_\alpha}{\lambda_\alpha}\right\} < \infty.$$

固定 $\mu_0 \in \mathcal{P}_{\mathcal{V}_0}(\mathbb{R}^{2d})$ 并设 $\mu_t := \mu_0 P_t^\alpha$。对于 $s > t$，根据半群性质和收缩，

$$\mathcal{W}_{\rho_{\mathcal{V}_0,\alpha}}(\mu_s, \mu_t) = \mathcal{W}_{\rho_{\mathcal{V}_0,\alpha}}\left((\mu_0 P_{s-t}^\alpha)P_t^\alpha, \mu_0 P_t^\alpha\right) \le e^{-c_\alpha t}\,\mathcal{W}_{\rho_{\mathcal{V}_0,\alpha}}(\mu_0 P_{s-t}^\alpha, \mu_0).$$

上述一致矩界以及$\rho_{\mathcal{V}_0,\ \alpha}$蕴含 $\sup_{u\ge 0}\mathcal{W}_{\rho_{\mathcal{V}_0,\ \alpha}}(\mu_0 P_u^\alpha,\ \mu_0) < \infty$(见 [EGZ19，推论 2.6])。因此，$(\mu_t)_{t\ge 0}$ 关于度量 $\mathcal{W}_{\rho_{\mathcal{V}_0,\ \alpha}}$构成一个柯西族；由空间 $\left(\mathcal{P}_{\mathcal{V}_0}(\mathbb{R}^{2d})\mathcal{W}_{\rho_{\mathcal{V}_0,\ \alpha}}\right)$的完备性可知，它收敛到某个 $\pi_\alpha \in \mathcal{P}_{\mathcal{V}_0}($ $\mathbb{R}^{2d})$。

极限 π_α 是不变的：对于任意 $t \ge 0$,

$$\pi_\alpha P_t^\alpha = \lim_{s\to\infty} \mu_s P_t^\alpha = \lim_{s\to\infty} \mu_{s+t} = \pi_\alpha.$$

由收缩可得唯一性：若 π'_α 是 $\mathcal{P}_{\mathcal{V}_0}$中的另一个不变测度，则

$$\mathcal{W}_{\rho_{\mathcal{V}_0,\alpha}}(\pi_\alpha, \pi'_\alpha) = \mathcal{W}_{\rho_{\mathcal{V}_0,\alpha}}(\pi_\alpha P_t^\alpha, \pi'_\alpha P_t^\alpha) \le e^{-c_\alpha t}\mathcal{W}_{\rho_{\mathcal{V}_0,\alpha}}(\pi_\alpha, \pi'_\alpha),$$

令 $t \to \infty$ 并得到 $\pi_\alpha = \pi'_\alpha$。取 $\nu = \pi_\alpha$ 即可得到所述的收敛到平衡态。 □

B.6 引理3.10 的证明

证明。在假设 A 假设 2.1(ii)， ∇U 是 L-利普希茨的，这意味着满足二次增长界 d

$$U(q) \le U(0) + |\nabla U(0)||q| + \frac{L}{2}|q|^2. \tag{B.23}$$

结合（B.23）与 $\mathcal{V}_0$的显式二次型，我们得到

$$\begin{aligned} c_1'\left(1+|q|^2+|p|^2\right) &\le 1 + \mathcal{V}_0(q,p) \\ &\le \left(\max(1,\mu_{\max}) + \sup_{q\in\mathbb{R}^d}\frac{U(q)}{1+|q|^2}\right)\left(1+|q|^2+|p|^2\right) \\ &\le \left(\max(1,\mu_{\max}) + \sup_{q\in\mathbb{R}^d}\frac{U(0)+|\nabla U(0)||q|+\frac{L}{2}|q|^2}{1+|q|^2}\right)\left(1+|q|^2+|p|^2\right) \\ &\le c_2'\left(1+|q|^2+|p|^2\right), \end{aligned} \tag{B.24}$$

其中 $c_1' = \min(1,\ \mu_{\min})$ 且 $c_2' = \max(1,\ \mu_{\max}) + U(0) + \frac{L}{2} + \frac{1}{2}|\nabla U(0)|$，其中 $\mu_{\min}$和 $\mu_{\max}$分别是对称矩阵 M 的最小和最大特征值，该矩阵在（2.8）中定义，即对应于（q，p）中二次型的矩阵，见（2.7）；显式公式见（2.9）。证明完毕。

□

附录 B.7 引理 3.11 的证明

证明。令 Γ 为（μ，ν）的任意耦合。由引理 3.3,

$$|z - z'|^2 \le k_1^{-2}\left(r(z,z')\right)^2.$$

令 $r := r(z,\ z')$。由于 φ_λ 是正的且非递增的，并且 $g_\lambda(s)$ 在 $[0, R_1(\lambda)$ 上的 $\ge g_*\,]$,，对于 $0 \le r \le R_1(\lambda)$，我们有

$$f_\lambda(r) = \int_0^r \varphi_\lambda(s) g_\lambda(s)\,ds \ge g_* \int_0^r \varphi_\lambda(s)\,ds \ge g_*\, c_r\, r.$$

此外，根据定义，对于所有 $r \geq R_1(\lambda)$，都有 $f_\lambda(r) = f_\lambda(R_1(\lambda)) = c_0$。

情况 *1*： $r \leq R_1(\lambda)$。于是 $r^2 \leq R_1(\lambda) r$ 因此

$$|z - z'|^2 \leq k_1^{-2} r^2 \leq k_1^{-2} R_1(\lambda) r \leq \frac{k_1^{-2} R_1(\lambda)}{g_* c_r} f_\lambda(r).$$

由于 $\mathcal{V} \geq 1$，我们有 $1 \leq 1 + \mathcal{V}(z) + \mathcal{V}(z')$，因此

$$|z - z'|^2 \leq \frac{k_1^{-2} R_1(\lambda)}{g_* c_r} f_\lambda(r) \left(1 + \mathcal{V}(z) + \mathcal{V}(z')\right). \tag{B.25}$$

情况 *2*： $r > R_1$（λ）。利用（3.40），

$$|z - z'|^2 \leq 2|z|^2 + 2|z'|^2 \leq 4C_V \left(1 + \mathcal{V}(z) + \mathcal{V}(z')\right).$$

由于在 $\{r > R_1(\lambda)\}$ 上，$f_\lambda(r) \geq c_0$，我们得到

$$|z - z'|^2 \leq \frac{4C_V}{c_0} f_\lambda(r) \left(1 + \mathcal{V}(z) + \mathcal{V}(z')\right). \tag{B.26}$$

结合两种情况（B.25）-（B.26）可得

$$|z - z'|^2 \leq C\, f_\lambda(r) \left(1 + \mathcal{V}(z) + \mathcal{V}(z')\right), \quad C := \max\left\{ \frac{k_1^{-2} R_1(\lambda)}{g_* c_r}, \frac{4C_V}{c_0} \right\}.$$

最后，由于 $\varepsilon \leq 1$ 和 $\mathcal{V} \geq 0$，

$$1 + \mathcal{V}(z) + \mathcal{V}(z') \leq \frac{1}{\varepsilon} \left(1 + \varepsilon \mathcal{V}(z) + \varepsilon \mathcal{V}(z')\right),$$

结合 $\rho_{\mathcal{V}}(z, z')$ 在（3.13）中的定义，可得

$$|z - z'|^2 \leq \frac{C}{\varepsilon} \rho_{\mathcal{V}}(z, z').$$

对 Γ 积分，并对所有耦合取下确界，得到 $\mathcal{W}_2^2(\mu, \nu) \leq C_\rho \mathcal{W}_{\rho_{\mathcal{V}}}(\mu, \nu)$，其中$C_\rho := C/\varepsilon$。证明完毕。 □

B.8 推论 3.12 的证明

证明。设 $c > 0$ 和 $\varepsilon = \gamma \frac{4c}{d + D()}$ 如（3.34）所示，并令 C_ρ 为引理 3.11 中基于此 ε计算得到的常数。将引理 3.11 应用于这对 $(\mu P_t^\alpha, \nu P_t^\alpha)$，得到

$$\mathcal{W}_2^2(\mu P_t^\alpha, \nu P_t^\alpha) \leq C_\rho \mathcal{W}_{\rho_{\mathcal{V}}}(\mu P_t^\alpha, \nu P_t^\alpha).$$

利用压缩性质（3.35），

$$\mathcal{W}_{\rho_{\mathcal{V}}}(\mu P_t^\alpha, \nu P_t^\alpha) \leq e^{-ct} \mathcal{W}_{\rho_{\mathcal{V}}}(\mu, \nu),$$

我们得到

$$\mathcal{W}_2^2(\mu P_t^\alpha, \nu P_t^\alpha) \leq C_\rho\, e^{-ct} \mathcal{W}_{\rho_{\mathcal{V}}}(\mu, \nu).$$

两边开平方即可得到所述界。最后一个结论可通过在 π_α 存在且满足 $\int_{\mathbb{R}^{2d}} \mathcal{V}\, d\pi_\alpha < \infty$时选取 $\nu = \pi_\alpha$ 得出。由此证明完毕。 □

B.9 推论的证明 3.13

ˆ证明。 根据命题 2.2， $\mathcal{V}_0$ 是（λ_α， A_α）-可容许的。应用定理 3.8 ，取 $\mathcal{V} = \mathcal{V}_0$ 和 ˆ(λ， D) = (λ_α， A_α)，可得收缩率 $c_\alpha > 0$ 以及相应的选取

$$\varepsilon_\alpha = \frac{4c_\alpha}{\gamma(d + A_\alpha)}.$$

令 $\rho_{\mathcal{V}_0,\ \alpha}$ 为相应的加权半度量，并令 $C_{\rho,\ \alpha}$ 表示引理3.11中与 $\mathcal{V}_0$ 相关且用 $\varepsilon = \varepsilon_\alpha$计算得到的常数。于是，对任意具有有限 ν -矩的$\mathcal{V}_0$，推论 3.12 给出：

$$\mathcal{W}_2(\mu P_t^\alpha, \nu P_t^\alpha) \le C_{\rho,\alpha}^{1/2} e^{-\frac{1}{2}c_\alpha t} \left(\mathcal{W}_{\rho_{\mathcal{V}_0,\alpha}}(\mu,\nu)\right)^{1/2}.$$

由推论 3.9可知，不变测度 π_α 存在且唯一，并满足 $\int_{\mathbb{R}^{2d}} \mathcal{V}_0\, d\pi_\alpha < \infty$。取$\nu = \pi_\alpha$ 并利用 $\pi_\alpha P_t^\alpha = \pi_\alpha$ 即可完成证明。 □

C 第4节结果的证明

C.1 引理4.2的证明

证明。 回顾式（2.3b，相互作用算子定义为 $\mathcal{A}' = -\nabla U$（q） $\cdot \nabla_q$，因此对任意光滑测试函数 f，

$$\mathcal{A}' f(q,p) = -\nabla U(q) \cdot \nabla_q f(q,p).$$

将其应用于李雅普诺夫函数 $\mathcal{V}_0$ （2.7）中定义的函数，可得：

$$\mathcal{V}_0(q,p) = U(q) + \frac{\gamma^2}{4}\left(|q + \gamma^{-1}p|^2 + |\gamma^{-1}p|^2 - \lambda|q|^2\right).$$

*Ste*步骤 *1*：计算 q–梯度： $\mathcal{V}_0$。首先，对 $\mathcal{V}_0$ 关于 q 求导：

$$\nabla_q \mathcal{V}_0(q,p) = \nabla U(q) + \frac{\gamma^2}{4}\left(2(q + \gamma^{-1}p) - 2\lambda q\right),$$

因为 $|\gamma^{-1}p|^2$ 不依赖于 q。因此，

$$\nabla_q \mathcal{V}_0(q,p) = \nabla U(q) + \frac{\gamma^2}{2}\left(q + \gamma^{-1}p - \lambda q\right) = \nabla U(q) + \frac{\gamma^2}{2}\left((1-\lambda)q + \gamma^{-1}p\right).$$

步骤 *2*：将 $\mathcal{A}'$ 应用于 $\mathcal{V}_0$。根据 $\mathcal{A}'$的定义，我们得到

$$\begin{aligned}
\mathcal{A}'\mathcal{V}_0(q,p) &= -\nabla U(q) \cdot \nabla_q \mathcal{V}_0(q,p) \\
&= -\nabla U(q) \cdot \left[\nabla U(q) + \frac{\gamma^2}{2}\left((1-\lambda)q + \gamma^{-1}p\right)\right] \\
&= -|\nabla U(q)|^2 - \frac{\gamma^2}{2}(1-\lambda)\,\nabla U(q)\cdot q - \frac{\gamma}{2}\nabla U(q)\cdot p.
\end{aligned}$$

这正是所声称的恒等式（4.2）。 □

C.2 引理的证明4.3

证明。 在整个证明过程中，我们记 $z = (q, p) \in \mathbb{R}^{2d}$，并使用记号 $\langle x, y\rangle = x^\top y$ 表示欧几里得内积。

步骤 *1*：极限 *Ornstein–Uhlenbeck* 算子，以及在 U 无穷远处的二次控制。 根据假设 4.1，存在一个对称正定矩阵 $Q_\infty \in \mathbb{R}^{d\times d}$ 和一个非增函数 $\varrho\colon [0,\infty) \to [0,\infty)$，使得 $\varrho(r) \to 0$ 当 $r \to \infty$ 时成立

$$|\nabla U(q) - Q_\infty q| \le \varrho(|q|)\,|q|, \qquad |q| \ge C_{\text{linear}}. \tag{C.1}$$

定义 $r(q) := \nabla U(q) - Q_\infty q$ 并且，对于 $R \ge 1$，回顾式（4.12）中尾部模量的定义：

$$\rho_\nabla(R) := \sup_{|q|\ge R} \frac{|r(q)|}{|q|}.$$

那么，当 $R \ge$ 取最大值时，$\rho_\nabla(R) < \infty$，并且由于对于 $R \ge C_{\text{linear}}$，$\rho_\nabla(R) \le \varrho(R)$，我们有$\rho_\nabla(R) \to 0$ 当 $R \to \infty$ 时。此外，$\rho_\nabla(\cdot)$ 是非递增的。

现在我们推导 $U(q) - \frac{1}{2}\langle Q_\infty q, q\rangle$ 在无穷远处的二次控制。由于 U满足假设2.1, ∇U 是连续的。因此

$$B_{\text{lin}} := \sup_{|x|\le C_{\text{linear}}} |r(x)| < \infty.$$

任取满足 $q \ne 0$ 的 $q \in \mathbb{R}^d$，并记 $\theta := q/|q| \in \mathbb{S}^{d-1}$。定义

$$g_\theta(s) := U(s\theta) - \frac{1}{2}\langle Q_\infty(s\theta), s\theta\rangle, \qquad s \ge 0.$$

根据微积分基本定理,

$$g_\theta(|q|) - g_\theta(0) = \int_0^{|q|} \langle r(s\theta), \theta\rangle\, ds.$$

在 C_{linear} 处分段，我们得到：

$$|g_\theta(|q|) - g_\theta(0)| \le \int_0^{C_{\text{linear}}} |r(s\theta)|\, ds + \int_{C_{\text{linear}}}^{|q|} |r(s\theta)|\, ds \le C_{\text{linear}} B_{\text{lin}} + \int_{C_{\text{linear}}}^{|q|} \rho_\nabla(s)\, s\, ds,$$

其中我们对 $s \ge C_{\text{linear}}$ 使用了 $|r(s\theta)| \le \rho_\nabla(s)\,|s\theta| = \rho_\nabla(s)\,s$。

因此，对所有 $|q| \ge C_{\text{linear}}$,

$$\left|U(q) - U(0) - \frac{1}{2}\langle Q_\infty q, q\rangle\right| \le C_{\text{linear}} B_{\text{lin}} + \int_{C_{\text{linear}}}^{|q|} \rho_\nabla(s)\, s\, ds. \tag{C.2}$$

因此，对于每个 $R \ge \max\{1, C_{\text{linear}}\}$ 以及每个 $|q| \ge R$,

$$\begin{aligned}\int_{C_{\text{linear}}}^{|q|} \rho_\nabla(s) s\, ds &= \int_{C_{\text{linear}}}^{R} \rho_\nabla(s) s\, ds + \int_R^{|q|} \rho_\nabla(s) s\, ds \\ &\le \int_{C_{\text{linear}}}^{R} \rho_\nabla(s) s\, ds + \rho_\nabla(R) \int_R^{|q|} s\, ds \le \int_{C_{\text{linear}}}^{R} \rho_\nabla(s) s\, ds + \frac{1}{2}\rho_\nabla(R)|q|^2,\end{aligned}$$

利用 ρ_∇ 是非递增的这一点。除以 $1+|q|^2$ 并对 $|q|\ge R$ 取上确界，得到

$$\delta_U(R) := \sup_{|q|\ge R} \frac{\left|U(q) - \frac{1}{2}\langle Q_\infty q, q\rangle\right|}{1+|q|^2} \le \frac{|U(0)| + C_{\text{linear}} B_{\text{lin}} + \int_{C_{\text{linear}}}^R \rho_\nabla(s) s\, ds}{1+R^2} + \frac{1}{2}\rho_\nabla(R). \tag{C.3}$$

最后，还需证明

$$\frac{1}{R^2}\int_{C_{\text{linear}}}^R \rho_\nabla(s)\, s\, ds \xrightarrow[R\to\infty]{} 0. \tag{C.4}$$

固定任意 $\varepsilon>0$，并选择 $S\ge C_{\text{linear}}$ 使得 $\rho_\nabla(S)\le\varepsilon$（这是可能的，因为 $\rho_\nabla(R)\to 0$）。于是对所有 $R\ge S$,

$$\begin{aligned}\frac{1}{R^2}\int_{C_{\text{linear}}}^R \rho_\nabla(s)\, s\, ds \le& \frac{1}{R^2}\int_{C_{\text{linear}}}^S \rho_\nabla(s)\, s\, ds + \frac{1}{R^2}\int_S^R \rho_\nabla(s)\, s\, ds \\ \le& \frac{1}{R^2}\int_{C_{\text{linear}}}^S \rho_\nabla(s)\, s\, ds + \varepsilon\cdot\frac{R^2-S^2}{2R^2} \le \frac{1}{R^2}\int_{C_{\text{linear}}}^S \rho_\nabla(s)\, s\, ds + \frac{\varepsilon}{2}.\end{aligned}$$

令 $R\to\infty$ 得到上极限 $\limsup_{R\to\infty}\frac{1}{R^2}\int_{C_{\text{linear}}}^R \rho_\nabla(s)\, s\, ds \le \varepsilon/2$，由于 $\varepsilon>0$ 是任意的，极限为 0，命题（C.4）得证。

接下来，引入“极限”动力学 Ornstein–Uhlenbeck 漂移算子

$$\mathcal{A}_\infty f(q,p) := \langle p, \nabla_q f(q,p)\rangle - \langle \gamma p + Q_\infty q, \nabla_p f(q,p)\rangle, \qquad (q,p)\in\mathbb{R}^{2d},$$

并写成

$$\mathcal{A}_0 = \mathcal{A}_\infty + \mathcal{A}_{\text{pert}}, \qquad \mathcal{A}_{\text{pert}} f(q,p) := -\langle r(q), \nabla_p f(q,p)\rangle.$$

步骤 2： $\mathcal{A}_\infty$ 在二次多项式上可逆。令 Q_2 表示 $\mathbb{R}^{2d}$上的二次多项式向量空间。对于 $\mathcal{M}(z) = \frac{1}{2}z^\top \mathsf{K} z$，其中 $\mathsf{K}=\mathsf{K}^\top$，有

$$(\mathcal{A}_\infty \mathcal{M})(z) = \frac{1}{2} z^\top \left(B^\top \mathsf{K} + \mathsf{K} B\right) z, \qquad B := \begin{pmatrix} 0 & I_d \\ -Q_\infty & -\gamma I_d \end{pmatrix}.$$

由于 Q_∞ 是正定的，且 $\gamma>0$， B 是 Hurwitz 矩阵。因此，李雅普诺夫方程 $B^\top \mathsf{K} + \mathsf{K} B = C$对于任意对称的 C（例如，参见 [HJ12]）都有唯一的对称解。因此，线性映射$\mathsf{Q}_2 \ni \mathcal{M} \mapsto \mathcal{A}_\infty \mathcal{M} \in \mathsf{Q}_2$ 是一个同构。

步骤 3：一个 对 $\mathcal{A}'\mathcal{V}_0$ 的显式展开及一个显式上界。回顾（2.7）可知

$$\mathcal{V}_0(q,p) = U(q) + \frac{\gamma^2}{4}\left(|q+\gamma^{-1}p|^2 + |\gamma^{-1}p|^2 - \lambda|q|^2\right).$$

直接计算可得显式的 q–梯度

$$\nabla_q \mathcal{V}_0(q,p) = \nabla U(q) + \frac{\gamma^2}{2}(1-\lambda)\, q + \frac{\gamma}{2}\, p. \tag{C.5}$$

根据引理 4.2，将 $\nabla U(q) = Q_\infty q + r(q)$ 代入（4.2）。定义二次型

$$Q(q,p) := -|Q_\infty q|^2 - \frac{\gamma^2}{2}(1-\lambda)\,\langle Q_\infty q, q\rangle - \frac{\gamma}{2}\,\langle Q_\infty q, p\rangle, \tag{C.6}$$

以及余项:

$$\mathcal{R}(q,p) := -2\langle Q_\infty q, r(q)\rangle - |r(q)|^2 - \frac{\gamma^2}{2}(1-\lambda)\,\langle r(q), q\rangle - \frac{\gamma}{2}\,\langle r(q), p\rangle. \tag{C.7}$$

Then

$$\mathcal{A}'\mathcal{V}_0(q,p) = Q(q,p) + \mathcal{R}(q,p). \tag{C.8}$$

现在我们给出尾部中 $|\mathcal{R}|$ 的上界。利用 $|Q_\infty q| \le \lambda_{\max}(Q_\infty)|q|$、初等界 $|q| \le |z|$, $|q||p| \le \frac{1}{2}(|q|^2+|p|^2) \le |z|^2$，以及尾部估计 $|r(q)| \le \rho_\nabla(|q|)\,|q|$（其对$|q| \ge C_{\text{linear}}$有效），对于所有满足 $|q| \ge \max\{1, C_{\text{linear}}\}$ 的情形以及所有 $p \in \mathbb{R}^d$,

$$|\mathcal{R}(q,p)| \le \rho_1(|q|)\,(1+|z|^2), \tag{C.9}$$

其中 ρ_1 被显式定义为

$$\rho_1(r) := \left(4\lambda_{\max}(Q_\infty) + \gamma^2|1-\lambda| + \gamma\right)\rho_\nabla(r) + (\rho_\nabla(r))^2, \qquad r \ge 0. \tag{C.10}$$

特别地，当 $r \to \infty$ 时， $\rho_1(r) \to 0$ 。

步骤 4：构造 $\mathcal{M}$ 并得到显式下界 c_{imp}。定义一个二次型

$$B_1(q,p) := -Q(q,p) - \left(\frac{1}{2}|p|^2 + \frac{\gamma}{2}\langle q,p\rangle + \frac{1}{2}\left\langle \left(Q_\infty + \frac{\gamma^2}{2}(1-\lambda)I_d\right)q, q\right\rangle\right). \tag{C.11}$$

令 $C_{B_1} := \nabla^2 B_1$ ，使得 $B_1(z) = \frac{1}{2}z^\top C_{B_1} z$。根据步骤 2，存在唯一的二次多项式$\mathcal{M} \in \mathbb{Q}_2$ ，使得

$$\mathcal{A}_\infty \mathcal{M}(z) = B_1(z), \qquad z \in \mathbb{R}^{2d}. \tag{C.12}$$

等价地，记 $\mathcal{M}(z) = \frac{1}{2}z^\top \mathsf{K} z$ ，其中 $\mathsf{K} = \mathsf{K}^\top$，则矩阵 K 是李雅普诺夫方程的唯一对称解

$$B^\top \mathsf{K} + \mathsf{K} B = C_{B_1}, \qquad B := \begin{pmatrix} 0 & I_d \\ -Q_\infty & -\gamma I_d \end{pmatrix}. \tag{C.13}$$

定义

$$C_{\mathcal{M}} := \frac{\|\mathsf{K}\|_{\text{op}}}{2}, \qquad C_\Delta := 2d\,\|\mathsf{K}\|_{\text{op}}. \tag{C.14}$$

于是（4.3）–（4.6）成立。

记 $\mathsf{K} = \begin{pmatrix} \mathsf{K}_{qq} & \mathsf{K}_{pq} \\ \mathsf{K}_{pq} & \mathsf{K}_{pp} \end{pmatrix}$，则 $\nabla_p \mathcal{M}(q,\ p) = \mathsf{K}_{pq} q + \mathsf{K}_{pp} p$，因此

$$|\nabla_p \mathcal{M}(q,p)| \le (\|\mathsf{K}_{pq}\|_{\text{op}} + \|\mathsf{K}_{pp}\|_{\text{op}})\,(|q|+|p|) \le (\|\mathsf{K}_{pq}\|_{\text{op}} + \|\mathsf{K}_{pp}\|_{\text{op}})\,(1+|z|). \tag{C.15}$$

接下来，利用 $\mathcal{A}_0 = \mathcal{A}_\infty + \mathcal{A}_{\text{pert}}$ 和 $\mathcal{A}'\mathcal{V}_0 = Q + \mathcal{R}$，我们可以计算得出

$$\begin{aligned} \mathcal{A}_0\mathcal{M}(z) + \mathcal{A}'\mathcal{V}_0(z) &= \mathcal{A}_\infty\mathcal{M}(z) + \mathcal{A}_{\text{pert}}\mathcal{M}(z) + Q(z) + \mathcal{R}(z) \\ &= B_1(z) + \mathcal{A}_{\text{pert}}\mathcal{M}(z) + Q(z) + \mathcal{R}(z) \qquad \text{by (C.12)} \\ &= -\Xi(z) + \mathcal{A}_{\text{pert}}\mathcal{M}(z) + \mathcal{R}(z), \end{aligned} \tag{C.16}$$

其中

$$\Xi(z) = \Xi(q,p) := \frac{1}{2}|p|^2 + \frac{\gamma}{2}\langle q,p\rangle + \frac{1}{2}\left\langle \left(Q_\infty + \frac{\gamma^2}{2}(1-\lambda)I_d\right) q, q\right\rangle.$$

对于 $|q| \geq C_{\text{linear}}$，利用 (C.15) 和 $|r(q)| \leq \rho_\nabla(|q|)\,|q|$，我们可以计算得出

$$\begin{aligned} |\mathcal{A}_{\text{pert}}\mathcal{M}(z)| &= |\langle r(q), \nabla_p \mathcal{M}(z)\rangle| \\ &\leq \rho_\nabla(|q|)\,|q| \cdot (\|\mathsf{K}_{pq}\|_{\text{op}} + \|\mathsf{K}_{pp}\|_{\text{op}})\,(1+|z|) \\ &\leq 2\,(\|\mathsf{K}_{pq}\|_{\text{op}} + \|\mathsf{K}_{pp}\|_{\text{op}})\,\rho_\nabla(|q|)\,(1+|z|^2). \end{aligned} \tag{C.17}$$

结合 (C.9) 和 (C.17) 可得：对于所有 $|q| \geq \max\{1, C_{\text{linear}}\}$，

$$\mathcal{A}_0\mathcal{M}(z) + \mathcal{A}'\mathcal{V}_0(z) \leq -\Xi(z) + [2(\|\mathsf{K}_{pq}\|_{\text{op}} + \|\mathsf{K}_{pp}\|_{\text{op}})\rho_\nabla(|q|) + \rho_1(|q|)]\,(1+|z|^2). \tag{C.18}$$

现在我们推导 Ξ 的强制性界。下界将用于吸收 (C.18) 中的尾部扰动项，而上界稍后将用于以显式常数建立 Ξ 与 $\mathcal{V}_0$ 之间的关系。令

$$a_{\min} := \lambda_{\min}(Q_\infty) + \frac{\gamma^2}{2}(1-\lambda), \qquad a_{\max} := \lambda_{\max}(Q_\infty) + \frac{\gamma^2}{2}(1-\lambda).$$

因此，我们得到全局界

$$\Xi(z) \geq \underline{a}\,|z|^2, \qquad \Xi(z) \leq \overline{a}\,|z|^2, \tag{C.19}$$

with

$$\underline{a} := \frac{1}{4}\left(a_{\min} + 1 - \sqrt{(a_{\min}-1)^2 + \gamma^2}\right), \qquad \overline{a} := \frac{1}{4}\left(a_{\max} + 1 + \sqrt{(a_{\max}-1)^2 + \gamma^2}\right). \tag{C.20}$$

步骤 *4.5*：确保吸收的闭式截断值。为获得可计算的截断值，我们回顾 (4.12) 中定义的尾部模量：$\rho_\nabla(R) := \sup_{|q|\geq R} \frac{|r(q)|}{|q|}$。根据定义，对于任意 $|q| \geq R$，有$|r(q)| \leq \rho_\nabla(|q|)|q|$。利用 (ρ_1) 在 (C.10) 中的表达式，吸收扰动项的条件为

$$\sup_{r\geq R}\left[\left(2\,(\|\mathsf{K}_{pq}\|_{\text{op}} + \|\mathsf{K}_{pp}\|_{\text{op}}) + 4\lambda_{\max}(Q_\infty) + \gamma^2|1-\lambda| + \gamma\right)\rho_\nabla(r) + (\rho_\nabla(r))^2\right] \leq \frac{5}{16}\underline{a}. \tag{C.21}$$

令 A 为线性项的系数：

$$A := 2\,(\|\mathsf{K}_{pq}\|_{\text{op}} + \|\mathsf{K}_{pp}\|_{\text{op}}) + 4\lambda_{\max}(Q_\infty) + \gamma^2|1-\lambda| + \gamma.$$

使（C.21）成立的一个充分条件是 $A\rho_\nabla$（R）$+$（ρ_∇（R））$^2 \leq \frac{5}{16}\underline{a}$。考虑二次方程 $x^2 + Ax - \frac{5}{16}\underline{a} = 0$。其正根为

$$\rho_\star := \frac{-A + \sqrt{A^2 + \frac{5}{4}\underline{a}}}{2} > 0.$$

现在我们设定

$$R_0 := \inf\{R \geq \max\{1, C_{\text{linear}}\} :\ \rho_\nabla(R) \leq \rho_\star\}. \tag{C.22}$$

于是 $R_0 < \infty$ （由于 $\rho_\nabla(R) \to 0$）。根据我们对 R_0的选择，对于所有满足 $z = (q, p)$ 且 $|q| \geq R_0$ 的情形，（C.21）中的括号项有界于$\frac{5}{16}a$。利用$\Xi(z) \geq a|z|^2$以及$1+|z|^2 \leq 2|z|^2$（由于 $R_0 \geq 1$），我们可以明确地得到：

$$\left[2\left(\|\mathsf{K}_{pq}\|_{\mathrm{op}} + \|\mathsf{K}_{pp}\|_{\mathrm{op}}\right)\rho_\nabla(|q|) + \rho_1(|q|)\right](1+|z|^2) \leq \frac{5}{8}\,\Xi(z). \tag{C.23}$$

将（C.23）代入（C.18），可得 $|q| \geq R_0$：

$$\mathcal{A}_0\mathcal{M}(z) + \mathcal{A}'\mathcal{V}_0(z) \leq -\frac{3}{8}\Xi(z). \tag{C.24}$$

现在，我们将（C.24）转化为如下形式的漂移改进 $-c_{\mathrm{imp}}\mathcal{V}_0 + C$，其显式速率如（4.9）所述。注意，根据 $\mathcal{V}_0$ 和 Ξ的定义，我们有如下精确恒等式：

$$\mathcal{V}_0(z) - \Xi(z) = U(q) - \frac{1}{2}\langle Q_\infty q, q\rangle.$$

对于 $|q| \geq R_0 \geq 1$，根据 $\delta_U(R_0)$ 的定义以及事实 $1+|q|^2 \leq 2|q|^2 \leq 2|z|^2$，可得

$$U(q) - \frac{1}{2}\langle Q_\infty q, q\rangle \leq \delta_U(R_0)(1+|q|^2) \leq 2\delta_U(R_0)|z|^2.$$

利用来自（C.19）的上界 $\Xi(z) \leq a|z|^2$，我们得到

$$\mathcal{V}_0(z) \leq \Xi(z) + 2\delta_U(R_0)|z|^2 \leq \left(\bar{a} + 2\delta_U(R_0)\right)|z|^2.$$

最后，使用下界 $|z|^2 \leq \frac{1}{\underline{a}}\Xi(z)$，由（C.19）可得显式控制

$$\mathcal{V}_0(z) \leq \frac{\bar{a} + 2\delta_U(R_0)}{\underline{a}}\,\Xi(z), \qquad |q| \geq R_0. \tag{C.25}$$

结合（C.24）和（C.25）可得

$$\mathcal{A}_0\mathcal{M}(z) + \mathcal{A}'\mathcal{V}_0(z) \leq -\frac{3}{8}\cdot\frac{\underline{a}}{\bar{a} + 2\delta_U(R_0)}\,\mathcal{V}_0(z), \qquad |q| \geq R_0.$$

回顾（C.20）可知 $4\underline{a} = a_{\min} + 1 - \sqrt{(a_{\min}-1)^2 + \gamma^2}$以及 $4\bar{a} = a_{\max} + 1 + \sqrt{(a_{\max}-1)^2 + \gamma^2}$。将该系数的分子和分母同乘 4，即可准确得到在（$c_{\mathrm{imp}}$ 4.9）中定义的常数：

$$\frac{3}{8}\cdot\frac{4\underline{a}}{4\bar{a} + 8\delta_U(R_0)} = \underline{c}_{\mathrm{imp}}.$$

步骤 5：控制区域 $|q| \leq R_0$ 及全局延拓。

固定 R_0 如（C.22）所示。由于 ∇U 连续，函数 $r(q) = \nabla U(q) - Q_\infty q$ 连续。因此

$$B_0 := \sup_{|q| \leq R_0} |r(q)| < \infty.$$

对于 $|q| \leq R_0$，我们利用（C.15）对扰动项进行界定：

$$|\mathcal{A}_{\mathrm{pert}}\mathcal{M}(q,p)| = |\langle r(q), \mathsf{K}_{pq}q + \mathsf{K}_{pp}p\rangle| \leq B_0\|\mathsf{K}_{pq}\|_{\mathrm{op}}R_0 + B_0\|\mathsf{K}_{pp}\|_{\mathrm{op}}|p|.$$

此外，由于 $|q| \le R_0$ 和 $|r(q)| \le B_0$，余项（C.7）满足

$$|\mathcal{R}(q,p)| \le 2|Q_\infty q|\,|r(q)| + |r(q)|^2 + \frac{\gamma^2}{2}|1-\lambda|\,|r(q)|\,|q| + \frac{\gamma}{2}|r(q)|\,|p| \le C_{R_0}^{(1)} + C_{R_0}^{(2)}|p|,$$

其中我们可以取

$$C_{R_0}^{(1)} := 2\,\lambda_{\max}(Q_\infty)\,R_0\,B_0 + B_0^2 + \frac{\gamma^2}{2}|1-\lambda|\,R_0\,B_0, \qquad C_{R_0}^{(2)} := \frac{\gamma}{2}B_0.$$

将此与

$$|\mathcal{A}_{\mathrm{pert}}\mathcal{M}(q,p)| \le B_0\|\mathsf{K}_{pq}\|_{\mathrm{op}}R_0 + B_0\|\mathsf{K}_{pp}\|_{\mathrm{op}}|p|,$$

我们得到，对于 $|q| \le R_0$，

$$|\mathcal{A}_{\mathrm{pert}}\mathcal{M}(q,p)| + |\mathcal{R}(q,p)| \le A_0 + L_0\,|p|,$$

with

$$A_0 := B_0\|\mathsf{K}_{pq}\|_{\mathrm{op}}R_0 + C_{R_0}^{(1)}, \qquad L_0 := B_0\|\mathsf{K}_{pp}\|_{\mathrm{op}} + C_{R_0}^{(2)}.$$

利用杨氏不等式 $L_0|p| \le \eta|p|^2 + \frac{L_0^2}{4\eta}$，对于任意 $\eta \in (0,1)$，我们得到

$$|\mathcal{A}_{\mathrm{pert}}\mathcal{M}(q,p)| + |\mathcal{R}(q,p)| \le \eta|p|^2 + C_{R_0,\eta}, \qquad |q| \le R_0,$$

其中，$C_{R_0,\eta}$ 的一种显式取法为

$$C_{R_0,\eta} := A_0 + \frac{L_0^2}{4\eta}.$$

回顾式（C.16），我们推得对于 $|q| \le R_0$,

$$\mathcal{A}_0\mathcal{M}(z) + \mathcal{A}'\mathcal{V}_0(z) \le -\Xi(z) + \eta|p|^2 + C_{R_0,\eta}.$$

利用 $\Xi(z) \ge a(|q|^2 + |p|^2)$（根据 C.19），并选择 $\eta := a/2$，得到

$$\mathcal{A}_0\mathcal{M}(z) + \mathcal{A}'\mathcal{V}_0(z) \le -\frac{1}{2}\,\Xi(z) + C_{R_0}, \qquad |q| \le R_0,$$

其中 $C_{R_0} := C_{R_0,\,a/2}$。

结合（C.24）（在 $|q| \ge R_0$ 上成立），我们得到全局界

$$\mathcal{A}_0\mathcal{M}(z) + \mathcal{A}'\mathcal{V}_0(z) \le -\frac{3}{8}\Xi(z) + C_{R_0}, \qquad z \in \mathbb{R}^{2d}.$$

最后，我们按照（C.25）在 $|q| \ge R_0$ 上将 Ξ 转换为 $\mathcal{V}_0$；而在 $|q| \le R_0$ 上，我们利用恒等式 $\mathcal{V}_0 - \Xi = U(q) - \frac{1}{2}\langle Q_\infty q, q\rangle$ 以及界 $|U(q) - \frac{1}{2}\langle Q_\infty q, q\rangle| \preceq \delta_U(1)(1+R_0^2)$（参见（C.3））， 得出 $\mathcal{V}_0(z) \le \Xi(z) + C'_{R_0}$ 在 $|q| \le R_0$ 上成立。由此可得，对所有 z，式（4.8）均成立，其中 c_{imp} 与（4.9）中的相同，且 C_{imp} 按引理陈述中的定义给出。C_{imp} 的有限性成立，因为对每个固定的 q，映射

$$p \mapsto \mathcal{A}_0\mathcal{M}(q,p) + \mathcal{A}'\mathcal{V}_0(q,p) + c_{\mathrm{imp}}\mathcal{V}_0(q,p)$$

是关于 p 的凹二次多项式。事实上，对于固定的 q，$\mathcal{A}_0\mathcal{M} + \mathcal{A}'\mathcal{V}_0 + c_{\mathrm{imp}}\mathcal{V}_0$ 关于 p 的二次系数矩阵是负定的（且对 $|q| \le R_0$ 一致），因此对 p 取上确界是有限的。由此完成证明。

□

C.3 引理的证明4.5

证明。 回顾式（4.16）可知 $\mathcal{V}_\alpha = \mathcal{V}_0 + \alpha\mathcal{M}$。由式（2.10）和 $U \geq 0$, 我们有 e

$$1 + \mathcal{V}_0(q,p) \;\geq\; c_1\left(1 + U(q) + |q|^2 + |p|^2\right) \;\geq\; c_1\left(1 + |q|^2 + |p|^2\right),$$

因此

$$1 + |q|^2 + |p|^2 \;\leq\; \frac{1}{c_1}\left(1 + \mathcal{V}_0(q,p)\right). \tag{C.26}$$

根据引理4.3（增长界（4.3））,

$$|\mathcal{M}(q,p)| \;\leq\; C_{\mathcal{M}}\left(1 + |q|^2 + |p|^2\right).$$

与（C.26）结合可得

$$|\mathcal{M}(q,p)| \;\leq\; \frac{C_{\mathcal{M}}}{c_1}\left(1 + \mathcal{V}_0(q,p)\right). \tag{C.27}$$

因此，对于任意 $\alpha \geq 0$,

$$|\alpha\mathcal{M}(q,p)| \;\leq\; \alpha\frac{C_{\mathcal{M}}}{c_1}\left(1 + \mathcal{V}_0(q,p)\right).$$

设 $\alpha_* := \frac{c_1}{2C_{\mathcal{M}}}$。则对于所有 $\alpha \in [0, \alpha_*]$，我们有 $\alpha\frac{C_{\mathcal{M}}}{c_1} \leq \frac{1}{2}$，从而

$$\begin{aligned}
1 + \mathcal{V}_\alpha(q,p) &= 1 + \mathcal{V}_0(q,p) + \alpha\mathcal{M}(q,p) \\
&\geq 1 + \mathcal{V}_0(q,p) - |\alpha\mathcal{M}(q,p)| \\
&\geq \left(1 - \alpha\frac{C_{\mathcal{M}}}{c_1}\right)\left(1 + \mathcal{V}_0(q,p)\right) \;\geq\; \frac{1}{2}\left(1 + \mathcal{V}_0(q,p)\right),
\end{aligned}$$

类似地,

$$\begin{aligned}
1 + \mathcal{V}_\alpha(q,p) &= 1 + \mathcal{V}_0(q,p) + \alpha\mathcal{M}(q,p) \\
&\leq 1 + \mathcal{V}_0(q,p) + |\alpha\mathcal{M}(q,p)| \\
&\leq \left(1 + \alpha\frac{C_{\mathcal{M}}}{c_1}\right)\left(1 + \mathcal{V}_0(q,p)\right) \;\leq\; \frac{3}{2}\left(1 + \mathcal{V}_0(q,p)\right).
\end{aligned}$$

这就证明了（4.19）。 □

C.4 引理的证明4.6

证明。 我们使用 $\mathcal{L}_\alpha = \mathcal{A}_0 + \alpha\mathcal{A}' + \alpha\Delta_q + \gamma\Delta_p$ 来自（2.4）以及 $\mathcal{V}_\alpha = \mathcal{V}_0 + \alpha\mathcal{M}$ 来自（4.16）来写出

$$\mathcal{L}_\alpha\mathcal{V}_\alpha = \mathcal{L}_0\mathcal{V}_0 + \alpha\left(\mathcal{L}_0\mathcal{M} + \mathcal{A}'\mathcal{V}_0 + \Delta_q\mathcal{V}_0\right) + \alpha^2\left(\mathcal{A}'\mathcal{M} + \Delta_q\mathcal{M}\right), \tag{C.28}$$

其中 $\mathcal{L}_0 = \mathcal{A}_0 + \gamma\Delta_p$。由（2.12），可得

$$\mathcal{L}_0\mathcal{V}_0 \leq \gamma(d + A) - \lambda\mathcal{V}_0. \tag{C.29}$$

式（ α）中的项 (C.28)。 由于 $\mathcal{L}_0\mathcal{M} = \mathcal{A}_0\mathcal{M} + \gamma\Delta_p\mathcal{M}$，我们有

$$\mathcal{L}_0\mathcal{M} + \mathcal{A}'\mathcal{V}_0 = \mathcal{A}_0\mathcal{M} + \mathcal{A}'\mathcal{V}_0 + \gamma\Delta_p\mathcal{M} \le C_{\mathrm{imp}} - \underline{c}_{\mathrm{imp}}\mathcal{V}_0 + \gamma\Delta_p\mathcal{M}$$

由（4.8）。由于 $\mathcal{M}(z) \coloneqq \frac{1}{2}z^\top \mathsf{K} z$ 是二次项，因此有 $\nabla^2_{pp}\mathcal{M} = \mathsf{K}_{pp}$，从而

$$\Delta_p\mathcal{M} = \mathrm{tr}(\mathsf{K}_{pp}).$$

特别地， $\Delta_p\mathcal{M}$ 与（q， p）无关，因此项 $\gamma\Delta_p\mathcal{M}$ 可以吸收到常数中。因此，

$$\mathcal{L}_0\mathcal{M} + \mathcal{A}'\mathcal{V}_0 \le (C_{\mathrm{imp}} + \gamma\,\mathrm{tr}(\mathsf{K}_{pp})) - \underline{c}_{\mathrm{imp}}\mathcal{V}_0.$$

此外，根据假设 2.1， ∇U 是利普希茨连续的，因此 $\nabla^2 U$ 几乎处处存在，且 $\|\nabla^2 U(q)\|_{\mathrm{op}} \le L$ 几乎处处成立。因此， $|\Delta U(q)| \le dL$ 几乎处处成立。因此，

$$|\Delta_q\mathcal{V}_0(q,p)| = \left|\Delta U(q) + \frac{\gamma^2}{2}d(1-\lambda)\right| \le dL + \frac{\gamma^2}{2}d|1-\lambda| = K_\Delta \quad \text{a.e.}$$

合并这些界可得

$$\mathcal{L}_0\mathcal{M} + \mathcal{A}'\mathcal{V}_0 + \Delta_q\mathcal{V}_0 \le (C_{\mathrm{imp}} + \gamma\,\mathrm{tr}(\mathsf{K}_{pp}) + K_\Delta) - \underline{c}_{\mathrm{imp}}\mathcal{V}_0. \tag{C.30}$$

（ α^2项，见 (C.28)。 由于 $\mathcal{A}' = -\nabla U(q)\cdot\nabla_q$，利用 $|\nabla U(q)| \le L|q| + |\nabla U(0)|$和（4.5），我们得到

$$|\mathcal{A}'\mathcal{M}| \le |\nabla U(q)|\,|\nabla_q\mathcal{M}(q,p)| \le (L|q| + |\nabla U(0)|)\,C_{\mathcal{M}}(1+|q|+|p|).$$

利用（$1+|q|+|p|$）$^2 \le 3$（$1+|q|^2+|p|^2$）以及 $L|q| + |\nabla U(0)| \le$（$L + |\nabla U(0)|$）（$1+|q|$）可得

$$|\mathcal{A}'\mathcal{M}| \le 3\,C_{\mathcal{M}}(L + |\nabla U(0)|)\,(1+|q|^2+|p|^2).$$

根据（2.10）和 $U \ge 0$， $1+|q|^2+|p|^2 \le c_1^{-1}$（$1+\mathcal{V}_0$）。因此

$$|\mathcal{A}'\mathcal{M}| \le 3\,C_{\mathcal{M}}\frac{(L+|\nabla U(0)|)}{c_1}\,(1+\mathcal{V}_0).$$

最后，由于 $\Delta_q\mathcal{M} = \mathrm{tr}(\mathsf{K}_{qq})$ 是一个常数，且 $1+\mathcal{V}_0 \ge 1$，

$$|\Delta_q\mathcal{M}| = |\mathrm{tr}(\mathsf{K}_{qq})| \le |\mathrm{tr}(\mathsf{K}_{qq})|\,(1+\mathcal{V}_0).$$

因此，

$$\mathcal{A}'\mathcal{M} + \Delta_q\mathcal{M} \le \left(|\mathrm{tr}(\mathsf{K}_{qq})| + 3\,C_{\mathcal{M}}\frac{(L+|\nabla U(0)|)}{c_1}\right)(1+\mathcal{V}_0). \tag{C.31}$$

将 $\mathcal{L}_0\mathcal{V}_0$ 项的界（C.29）、 α项的界（C.30）和 α^2项的界（C.31）代入（C.28），得到（4.20），其中 C_1， C_2 由（4.21）给出。 □

C.5 命题的证明 4.7

证明。我们从引理 4.6出发：对所有 $\alpha \in (0,1]$ 以及所有 $z=(q,\ p) \in \mathbb{R}^{2d}$,

$$\begin{aligned}\mathcal{L}_\alpha \mathcal{V}_\alpha(z) &\le \gamma(d+A) - \lambda \mathcal{V}_0(z) + \alpha\left(C_1 - \underline{c}_{\mathrm{imp}} \mathcal{V}_0(z)\right) + C_2\alpha^2(1+\mathcal{V}_0(z)) \\ &= \left[\gamma(d+A) + \alpha C_1 + C_2\alpha^2\right] - \left[\lambda + \alpha \underline{c}_{\mathrm{imp}} - C_2\alpha^2\right] \mathcal{V}_0(z).\end{aligned} \tag{C.32}$$

Set

$$\varsigma(\alpha) := \lambda + \alpha \underline{c}_{\mathrm{imp}} - C_2\alpha^2.$$

步骤 *1*：比较 $\mathcal{V}_0$和 $\mathcal{V}_\alpha$。由（4.3），$|\mathcal{M}| \le C_{\mathcal{M}}\ (1+|q|^2+|p|^2)$。由（2.10），$1+\mathcal{V}_0 \ge c_1$ $(1+|q|^2+|p|^2)$，因此 $1+|q|^2+|p|^2 \le c_1^{-1}\ (1+\mathcal{V}_0)$。因此

$$|\mathcal{M}| \le \widetilde{C}_{\mathcal{M}}\,(1+\mathcal{V}_0), \qquad \widetilde{C}_{\mathcal{M}} := \frac{C_{\mathcal{M}}}{c_1}.$$

因此，对于 $\alpha \in (0,1]$,

$$\mathcal{V}_\alpha = \mathcal{V}_0 + \alpha\mathcal{M} \le \mathcal{V}_0 + \alpha|\mathcal{M}| \le \mathcal{V}_0 + \alpha\widetilde{C}_{\mathcal{M}}(1+\mathcal{V}_0) = (1+\alpha\widetilde{C}_{\mathcal{M}})\mathcal{V}_0 + \alpha\widetilde{C}_{\mathcal{M}}.$$

重新整理后得到下界

$$\mathcal{V}_0 \ge \frac{\mathcal{V}_\alpha - \alpha\widetilde{C}_{\mathcal{M}}}{1+\alpha\widetilde{C}_{\mathcal{M}}}. \tag{C.33}$$

步骤 *2*：用 $\mathcal{V}_\alpha$表示的漂移界。将（C.33）代入（C.32）得到

$$\begin{aligned}\mathcal{L}_\alpha \mathcal{V}_\alpha &\le \left[\gamma(d+A) + \alpha C_1 + C_2\alpha^2\right] - \varsigma(\alpha)\left(\frac{\mathcal{V}_\alpha - \alpha\widetilde{C}_{\mathcal{M}}}{1+\alpha\widetilde{C}_{\mathcal{M}}}\right) \\ &= \left[\gamma(d+A) + \alpha C_1 + C_2\alpha^2 + \frac{\alpha\widetilde{C}_{\mathcal{M}}\varsigma(\alpha)}{1+\alpha\widetilde{C}_{\mathcal{M}}}\right] - \frac{\varsigma(\alpha)}{1+\alpha\widetilde{C}_{\mathcal{M}}}\,\mathcal{V}_\alpha.\end{aligned}$$

定义

$$\lambda_\alpha := \frac{\varsigma(\alpha)}{1+\alpha\widetilde{C}_{\mathcal{M}}}.$$

常数项和可容许性。使用 $\alpha \le 1,\ C_2\alpha^2 \le \alpha$（$\alpha C_2$），并且

$$\frac{\alpha\widetilde{C}_{\mathcal{M}}\varsigma(\alpha)}{1+\alpha\widetilde{C}_{\mathcal{M}}} \le \alpha\widetilde{C}_{\mathcal{M}}\varsigma(\alpha) \le \alpha\widetilde{C}_{\mathcal{M}}(\lambda + \underline{c}_{\mathrm{imp}}),$$

我们得到

$$\begin{aligned}&\gamma(d+A) + \alpha C_1 + C_2\alpha^2 + \frac{\alpha\widetilde{C}_{\mathcal{M}}\varsigma(\alpha)}{1+\alpha\widetilde{C}_{\mathcal{M}}} \\ &\le \gamma(d+A) + \alpha\left[C_1 + \alpha C_2 + \widetilde{C}_{\mathcal{M}}(\lambda + \underline{c}_{\mathrm{imp}})\right] = \gamma\left(d + A'_\alpha\right),\end{aligned}$$

其中 A'_α 由（4.25）给出。因此

$$\mathcal{L}_\alpha \mathcal{V}_\alpha \le \gamma\left(d + A'_\alpha - \lambda_\alpha \mathcal{V}_\alpha\right),$$

即， $\mathcal{V}_\alpha$ 对于 $\mathcal{L}_\alpha$ 是（λ_α， A'_α）-可容许的。

速率展开。由于 $\frac{1}{1+x} \geq 1 - x$ 对任意 $x \geq 0$,

$$\lambda_\alpha = \frac{\varsigma(\alpha)}{1+\alpha\widetilde{C}_{\mathcal{M}}} \geq \varsigma(\alpha)\left(1-\alpha\widetilde{C}_{\mathcal{M}}\right) = \left(\lambda + \alpha\underline{c}_{\text{imp}} - C_2\alpha^2\right)\left(1-\alpha\widetilde{C}_{\mathcal{M}}\right).$$

展开乘积可得

$$\lambda_\alpha \geq \lambda + \alpha\left(\underline{c}_{\text{imp}} - \lambda\widetilde{C}_{\mathcal{M}}\right) - \alpha^2\left(C_2 + \underline{c}_{\text{imp}}\widetilde{C}_{\mathcal{M}}\right) + \alpha^3 C_2\widetilde{C}_{\mathcal{M}}.$$

去掉非负的三次项可得

$$\lambda_\alpha \geq \lambda + \alpha\left(\underline{c}_{\text{imp}} - \lambda\widetilde{C}_{\mathcal{M}}\right) - \alpha^2\left(C_2 + \underline{c}_{\text{imp}}\widetilde{C}_{\mathcal{M}}\right) = \lambda + \delta\,\alpha - C_\lambda\,\alpha^2,$$

其中 δ 和 C_λ 定义于 (4.23)–(4.24)。在 (4.27) 中对 α_1 的选择确保辅助约束（来自命题 2.2 以及关于 λ_α的正性要求）同时成立，因此结论得证。

□

C.6 定理 4.8 的证明

证明。 取 $\alpha \in (0, \alpha_1]$ 并回顾 (3.4) 中 $L_{\text{eff}}(\alpha) = (1+\alpha\gamma)L$。我们使用度量参数$\Lambda_\alpha(\lambda)$，其定义见 (4.30)，即

$$\Lambda_\alpha(\lambda) = J_2\,\frac{(1+\alpha\gamma)L}{\lambda}, \qquad \Lambda_0 := \Lambda_0(\lambda) = J_2\frac{L}{\lambda}.$$

另设 $\mathcal{S}_h := 1 - \frac{1}{2\Lambda_0} > 0$ 和

$$\Delta_\Lambda := J_2 L\frac{\delta - \gamma\lambda}{\lambda^2} > 0.$$

根据命题 4.7，存在 $\alpha_1 > 0$ 和 $C_\lambda \geq 0$，使得对于所有 $\alpha \in (0, \alpha_1]$ 都可以选取$\lambda_\alpha > 0$ 满足

$$\lambda_\alpha \;\geq\; \underline{\lambda}_\alpha := \lambda + \delta\alpha - C_\lambda\alpha^2.$$

定义代理量

$$\bar{\Lambda}_\alpha := J_2\,\frac{(1+\alpha\gamma)L}{\underline{\lambda}_\alpha}.$$

由于 $\lambda_\alpha \geq \underline{\lambda}_\alpha$ 和 $\Lambda_\alpha(\lambda)$ 随 λ递减，

$$\Lambda_\alpha(\lambda_\alpha) = J_2\,\frac{(1+\alpha\gamma)L}{\lambda_\alpha} \leq J_2\,\frac{(1+\alpha\gamma)L}{\underline{\lambda}_\alpha} = \bar{\Lambda}_\alpha. \tag{C.34}$$

设 $D := \delta - \gamma\lambda > 0$ 并定义

$$g(\alpha) := \frac{1+\alpha\gamma}{\underline{\lambda}_\alpha} = \frac{1+\alpha\gamma}{\lambda + \delta\alpha - C_\lambda\alpha^2}.$$

直接计算可得

$$g(\alpha) - g(0) = \frac{1+\alpha\gamma}{\underline{\lambda}_\alpha} - \frac{1}{\lambda} = \frac{-D\alpha + C_\lambda\alpha^2}{\lambda\,\underline{\lambda}_\alpha}. \tag{C.35}$$

选择

$$\alpha_{3,a} := \min\left\{\alpha_1,\ 1,\ \frac{D}{4C_\lambda}\ (\text{if } C_\lambda > 0),\ \sqrt{\frac{\lambda}{2C_\lambda}}\ (\text{if } C_\lambda > 0)\right\}, \tag{C.36}$$

约定当 $C_\lambda = 0$ 时，删除涉及 C_λ 的项。于是，对所有 $\alpha \in (0, \alpha_{3,a}]$，有

$$-D\alpha + C_\lambda \alpha^2 \le -\frac{3D}{4}\alpha, \qquad \underline{\lambda}_\alpha \ge \frac{\lambda}{2}, \tag{C.37}$$

因此，由 (C.35) 和 (C.37) 可得

$$g(\alpha) - g(0) \le -\frac{3D}{4}\alpha \cdot \frac{2}{\lambda^2} \le -\frac{D}{8\lambda^2}\alpha.$$

两边乘以 $J_2 L$，并利用 $g(0) = 1/\lambda$，得到

$$\bar{\Lambda}_\alpha = J_2 L\, g(\alpha) \le J_2 L \left(\frac{1}{\lambda} - \frac{D}{8\lambda^2}\alpha\right) = \Lambda_0 - \frac{\Delta_\Lambda}{8}\alpha.$$

结合 (C.34) 即可证明 (i)，其中显式取 $c_\Lambda := \Delta_\Lambda/8$。

回顾 (4.31)。由于 $\Lambda_0 > \frac{1}{2}$，我们有 $h'(\Lambda_0) < 0$ 并且

$$h'(\Lambda_0) = -h(\Lambda_0)\left(1 - \frac{1}{2\Lambda_0}\right) = -h(\Lambda_0)\mathcal{S}_h.$$

此外

$$h''(\Lambda) = \frac{\Lambda^2 - \Lambda - \frac{1}{4}}{\Lambda^{3/2}}\, e^{-\Lambda}.$$

定义有限常数

$$M_h := \sup_{\Lambda \in [\Lambda_0/2, \Lambda_0]} |h''(\Lambda)|.$$

令 $t := c_\Lambda \alpha$。为确保 $\Lambda_0 - t \in [\Lambda_0/2, \Lambda_0]$，并使二次余项受线性项支配，设

$$\alpha_{3,b} := \min\left\{\frac{\Lambda_0}{2c_\Lambda},\ \frac{h(\Lambda_0)\mathcal{S}_h}{M_h\, c_\Lambda}\right\}. \tag{C.38}$$

最后定义

$$\alpha_{\text{metric,acc}} := \min\{\alpha_{3,a}, \alpha_{3,b}\}. \tag{C.39}$$

对于 $\alpha \in (0,\ \alpha_{\text{metric,\ acc}}]$，第 (i) 部分表明 $\Lambda_\alpha(\lambda_\alpha) \le \Lambda_0 - t$，其中 $t = c_\Lambda \alpha$。由于 h 在 $[\Lambda_0/2, \Lambda_0]$ 上递减，因此得到

$$h(\Lambda_\alpha(\lambda_\alpha)) \ge h(\Lambda_0 - t).$$

对 h 在 Λ_0 处作二阶泰勒展开，并使用拉格朗日余项，可得

$$h(\Lambda_0 - t) \ge h(\Lambda_0) + |h'(\Lambda_0)|\, t - \frac{1}{2}M_h t^2 \ge h(\Lambda_0) + \frac{1}{2}h(\Lambda_0)\mathcal{S}_h\, t,$$

其中，最后一个不等式用到了 $t \le h(\Lambda_0)\mathcal{S}_h/M_h$（根据 $\alpha_{3,b}$ 的定义）。因此

$$h(\Lambda_\alpha(\lambda_\alpha)) \ge h(\Lambda_0)\left(1 + \frac{\mathcal{S}_h}{2}\, c_\Lambda\, \alpha\right).$$

因此，对于任意 $c_3 < c_3^* := \frac{s_h}{2} e_\Lambda$，我们有

$$\widetilde{\Lambda}_{3,\alpha}(\lambda_\alpha) = \kappa_{\text{adjust}}\, h(\Lambda_\alpha(\lambda_\alpha)) \geq \kappa_{\text{adjust}}\, h(\Lambda_0)\,(1 + c_3\alpha) = \widetilde{\Lambda}_{3,0}(\lambda)\,(1 + c_3\alpha).$$

最后，

$$\begin{aligned}\widetilde{\Lambda}_{2,\alpha}(\lambda_\alpha) = h(\Lambda_\alpha(\lambda_\alpha))\,\frac{L_{\text{eff}}(\alpha)}{\gamma^2} &= \frac{L}{\gamma^2}(1+\alpha\gamma)\,h(\Lambda_\alpha(\lambda_\alpha)) \\ &\geq \widetilde{\Lambda}_{2,0}(\lambda)\,(1+\alpha\gamma)(1+c_3\alpha) \geq \widetilde{\Lambda}_{2,0}(\lambda)\,(1+(\gamma+c_3)\alpha)\,,\end{aligned}$$

因此，该界以 $c_2 := \gamma + c_3$ 成立。至此证明完毕。 □

C.7 引理的证明4.10

证明。由 $\alpha = 0$ 处的严格活跃性，我们有 $\Delta(0) > 0$。由 Δ 的连续性，存在 $\varepsilon > 0$，使得对所有 $\alpha \in [0, \varepsilon]$，$\Delta(\alpha) > 0$。因此，$\alpha_{\text{branch}} \geq \varepsilon > 0$，并且根据 α_{branch} 的定义，对所有 $\alpha \in (0, \alpha_{\text{branch}}]$，我们有

$$\widetilde{\Lambda}_{1,\alpha}(\lambda_\alpha) \leq \min\left\{\widetilde{\Lambda}_{2,\alpha}(\lambda_\alpha),\, \widetilde{\Lambda}_{3,\alpha}(\lambda_\alpha)\right\}.$$

回顾定理 3.8(见（3.32））可知

$$c(\lambda_\alpha) = \frac{\gamma}{384}\min\left\{\widetilde{\Lambda}_{1,\alpha}(\lambda_\alpha), \widetilde{\Lambda}_{2,\alpha}(\lambda_\alpha), \widetilde{\Lambda}_{3,\alpha}(\lambda_\alpha)\right\},$$

t ˜上述不等式表明，最小值在 $\Lambda_{1,\alpha}(\lambda_\alpha)$ 处取得，因此 s

$$c_\alpha = \frac{\gamma}{384}\widetilde{\Lambda}_{1,\alpha}(\lambda_\alpha) \qquad \text{for all } \alpha \in (0, \alpha_{\text{branch}}].$$

□

C.8 定理4.12 的证明

证明。由 引理4.10，对所有 $\alpha \in (0, \alpha_{\text{branch}}]$，李雅普诺夫分支保持激活。Hen ce

$$c_\alpha = \frac{\gamma}{384}\,\widetilde{\Lambda}_{1,\alpha}(\lambda_\alpha) = \frac{\gamma}{384}\,\frac{\lambda_\alpha L_{\text{eff}}(\alpha)}{\gamma^2},$$

其中 $L_{\text{eff}}(\alpha) = (1+\gamma\alpha)L$。在 $\alpha = 0$，

$$c_0 = \frac{\gamma}{384}\,\widetilde{\Lambda}_{1,0}(\lambda) = \frac{\gamma}{384}\,\frac{\lambda L}{\gamma^2} = \frac{L}{384\,\gamma}\,\lambda.$$

根据命题4.7，对所有 $\alpha \in (0, \alpha_1]$，

$$\lambda_\alpha \;\geq\; \lambda + \delta\alpha - C_\lambda\alpha^2.$$

因此，对于 $\alpha \in (0, \alpha_{\text{branch}}]$，

$$\begin{aligned}c_\alpha = \frac{L}{384\,\gamma}\,\lambda_\alpha(1+\gamma\alpha) &\geq \frac{L}{384\,\gamma}\,(\lambda + \delta\alpha - C_\lambda\alpha^2)(1+\gamma\alpha) \\ &= \frac{L}{384\,\gamma}\left[\lambda + (\delta+\gamma\lambda)\alpha + (\gamma\delta - C_\lambda)\alpha^2 - \gamma C_\lambda\alpha^3\right].\end{aligned}$$

舍去可能为正的项 $(\gamma\delta)\alpha^2$，并使用 $\alpha \le 1$ 来界定 $-\gamma C_\lambda \alpha^3 \ge -\gamma C_\lambda \alpha^2$，我们得到

$$c_\alpha \ge \frac{L}{384\,\gamma}\left[\lambda + (\delta+\gamma\lambda)\alpha - (1+\gamma)C_\lambda\alpha^2\right] = c_0 + \tilde{\kappa}\,\alpha - C'\alpha^2,$$

与 $\tilde{\kappa} = \frac{L}{(\delta+\gamma\lambda)\ 384\gamma}$ 和

$$C' = \frac{L}{384\gamma}(1+\gamma)C_\lambda = \frac{L}{384\gamma}(1+\gamma)\left(C_2 + \widetilde{C}_{\mathcal{M}}\,\underline{c}_{\mathrm{imp}}\right).$$

选择 $\alpha_{\mathrm{branch,acc}} := \text{最小值}\{\alpha_{\mathrm{branch}}, 1, \tilde{\kappa}/(2C')\}$。那么对于所有 $\alpha \in (0, \alpha_{\mathrm{branch,acc}}]$，我们有 $C'\alpha \le \tilde{\kappa}/2$。因此，令 $\kappa := \tilde{\kappa}/2$ 可得 (4.36)：

$$c_\alpha \ge c_0 + \frac{\tilde{\kappa}}{2}\alpha = c_0 + \kappa\alpha.$$

□

C.9 推论 4.13 的证明

证明。回顾定理 3.32 中的式 (3.8，可知

$$c(\lambda) = \frac{\gamma}{384}\min\left\{\widetilde{\Lambda}_{1,\alpha}(\lambda), \widetilde{\Lambda}_{2,\alpha}(\lambda), \widetilde{\Lambda}_{3,\alpha}(\lambda)\right\}.$$

对于 HFHR 动力学，令

$$f_1(\alpha) := \frac{\gamma}{384}\widetilde{\Lambda}_{1,\alpha}(\lambda_\alpha), \qquad f_2(\alpha) := \frac{\gamma}{384}\widetilde{\Lambda}_{2,\alpha}(\lambda_\alpha), \qquad f_3(\alpha) := \frac{\gamma}{384}\widetilde{\Lambda}_{3,\alpha}(\lambda_\alpha),$$

于是 $c_\alpha = \min\{f_1(\alpha)\ f_2(\alpha)\ f_3(\alpha)\}$ 且 $c_0 = \min\{f_1(0)\ f_2(0)\ f_3(0)\}$。特别地，对于每个 i，$f_i(0) \ge c_0$。

步骤 *1*：李雅普诺夫分支下界。根据定理 4.12，对于所有 $\alpha \in (0, \alpha_{\mathrm{branch,\ acc}}]$，

$$f_1(\alpha) \ \ge\ c_0 + \kappa\,\alpha. \tag{C.40}$$

步骤 *2*：度量分支下界。根据定理 4.8，对于所有 $\alpha \in (0, \alpha_{\mathrm{metric,\ acc}}]$，

$$f_i(\alpha) \ge f_i(0)(1 + c_i\alpha) = f_i(0) + c_i f_i(0)\alpha, \qquad i = 2, 3,$$

其中 c_2，$c_3 > 0$。利用 $f_i(0) \ge c_0$ 可得

$$f_i(\alpha) \ge c_0 + c_0 c_i\,\alpha, \qquad i = 2, 3. \tag{C.41}$$

步骤 *3*：取最小值。对于所有 $\alpha \in (0,\ \alpha_{\mathrm{global}}]$，(C.40) 和 (C.41) 均成立。因此

$$c_\alpha = \min_{i=1,2,3} f_i(\alpha) \ \ge\ \min\left\{c_0 + \kappa\alpha,\ c_0 + c_0c_2\alpha,\ c_0 + c_0c_3\alpha\right\} \ =\ c_0 + \kappa_{\mathrm{global}}\alpha,$$

其中 $\kappa_{\mathrm{global}} := \min\{\kappa, c_0c_2, c_0c_3\} > 0$。 □

C.10 推论的证明4.14

证明。取 $\alpha \in (0,\ \alpha_{\mathrm{W2}}]$ 并选择 $\lambda_\alpha = \lambda_{\alpha\circ}$ 根据 $\alpha \le \alpha_{\mathrm{pos}}$ 和（4.28），有 $\lambda_\alpha \ge \lambda_- = \lambda/2$。另外，由于 $\lambda_\alpha = \lambda + \delta\alpha - C_\lambda \alpha^2 \le \lambda + \delta\alpha \le \lambda_+$，有 $\lambda_\alpha \in I_{\lambda\circ}$

应用引理3.11，其中 $\mathcal{V} = \mathcal{V}_\alpha$ 和 $\varepsilon = \varepsilon_{\alpha\circ}$ 利用一致界 k_1^-，$R_1(\lambda)^+$，g_*^-，c_r^-，c_0^-以及C_V^{unif}，并根据 ε^-的定义，引理（3.41）中的显式常数3.11 给出

$$\mathcal{W}_2^2(\mu,\nu) \le C_\rho^{\mathrm{unif}}\, \mathcal{W}_{\rho\mathcal{V}_\alpha}(\mu,\nu),$$

因此我们得到：

$$\mathcal{W}_2(\mu P_t^\alpha, \nu P_t^\alpha) \le \left(C_\rho^{\mathrm{unif}}\right)^{1/2} \left(\mathcal{W}_{\rho\mathcal{V}_\alpha}(\mu P_t^\alpha, \nu P_t^\alpha)\right)^{1/2}.$$

根据推论4.13,

$$\mathcal{W}_{\rho\mathcal{V}_\alpha}(\mu P_t^\alpha, \nu P_t^\alpha) \le e^{-c_\alpha t}\, \mathcal{W}_{\rho\mathcal{V}_\alpha}(\mu,\nu),$$

因此

$$\mathcal{W}_2(\mu P_t^\alpha, \nu P_t^\alpha) \le \left(C_\rho^{\mathrm{unif}}\right)^{1/2} e^{-\frac{1}{2}c_\alpha t} \left(\mathcal{W}_{\rho\mathcal{V}_\alpha}(\mu,\nu)\right)^{1/2}.$$

这证明了（4.39），其中 $c_\alpha^{(2)} = \frac{1}{2}c_{\alpha\circ}$

最后，由于 $\alpha_{\mathrm{W2}} \le \alpha_{\mathrm{global}}$，推论4.13 给出 $c_\alpha \ge c_0 + \kappa_{\mathrm{global}}\alpha$，因此

$$c_\alpha^{(2)} = \frac{1}{2}c_\alpha \ge \frac{1}{2}\left(c_0 + \kappa_{\mathrm{global}}\alpha\right) = c_0^{(2)} + \kappa^{(2)}\alpha,$$

对于所有 $\alpha \in (0,\ \alpha_{\mathrm{W2}}]$。这就完成了证明。 □

D 第5节结果的证明

D.1 命题5.1的证明

证明。我们验证 t he properties sequentially based on the 可分结构 $U(q) = \sum_{i=1}^d v(q_i)$。

(a) 正则性。首先，我们验证假设2.1(i)-(ii)。一维势函数 $v(s)$ 具有连续导数 $v'(s)$，满足 $|v'(s) - v'(t)| \le |s-t|$ 对所有 $s,\ t \in \mathbb{R}$ (因为 v' 是斜率为 ± 1或 0 的分段线性函数)。对于 d维势函数，我们对各分量的平方求和：

$$|\nabla U(q) - \nabla U(q')|^2 = \sum_{i=1}^d |v'(q_i) - v'(q_i')|^2 \le \sum_{i=1}^d |q_i - q_i'|^2 = |q-q'|^2.$$

因此，∇U 是全局 Lipschitz 的，其常数为 $L = 1$，且与 d 无关。

(b) 耗散性。我们验证假设 2.1(iii)。令

$$\bar\lambda := \frac{1}{4+\gamma^2} \in \left(0, \frac{1}{4}\right],$$

并且对于 $s \in \mathbb{R}$ 定义

$$\Delta_{\bar\lambda}(s) := \bar\lambda\left(v(s) + \frac{\gamma^2}{4}s^2\right) - \frac{1}{2}\, s\, v'(s).$$

我们声称，$\sup_{s\in\mathbb{R}} \Delta_{\bar\lambda}(s) \le A_1(\gamma)$，其中

$$A_1(\gamma) := \frac{\gamma^4 + 6\gamma^2 + 16}{4(\gamma^4 + 10\gamma^2 + 24)}.$$

情况 1： $|s| \le \frac{1}{2}$。这里 $v(s) = -\frac{1}{4} - \frac{1}{2}s^2$ 和 $v'(s) = -s$，因此

$$\Delta_{\bar\lambda}(s) = \frac{\bar\lambda}{4} + \left(\frac{1}{2} + \bar\lambda\frac{\gamma^2 - 2}{4}\right)s^2.$$

由于 s^2 的系数为正，$\Delta_{\bar\lambda}$ 在 $|s| = \frac{1}{2}$ 处取最大值：

$$\sup_{|s|\le 1/2} \Delta_{\bar\lambda}(s) = \Delta_{\bar\lambda}\left(\frac{1}{2}\right) = \frac{3\gamma^2 + 10}{16(\gamma^2 + 4)}.$$

情况 2： $|s| > \frac{1}{2}$。这里 $v(s) = \frac{1}{2}(|s| - 1)^2 = \frac{1}{2}(s^2 - 2|s| + 1)$ 和 $sv'(s) = s^2 - |s|$。因此

$$\Delta_{\bar\lambda}(s) = -a|s|^2 + b|s| + \frac{\bar\lambda}{2},$$

其中

$$a := \frac{1}{2} - \bar\lambda\left(\frac{1}{2} + \frac{\gamma^2}{4}\right) = \frac{\gamma^2 + 6}{4(\gamma^2 + 4)}, \qquad b := \frac{1}{2} - \bar\lambda = \frac{\gamma^2 + 2}{2(\gamma^2 + 4)}.$$

凹二次函数 $-ax^2 + bx$ 在 $x_* = \frac{b}{2a}$ 处取得最大值，因此

$$\sup_{|s|>1/2} \Delta_{\bar\lambda}(s) \le \frac{b^2}{4a} + \frac{\bar\lambda}{2} = \frac{\gamma^4 + 6\gamma^2 + 16}{4(\gamma^4 + 10\gamma^2 + 24)} = A_1(\gamma).$$

此外，

$$A_1(\gamma) - \Delta_{\bar\lambda}\left(\frac{1}{2}\right) = \frac{(\gamma^2 - 2)^2}{16(\gamma^2 + 4)(\gamma^2 + 6)} \ge 0,$$

因此，$A_1(\gamma)$ 也支配情况 1 中的最大值。因此，对于所有 $s \in \mathbb{R}$，$\Delta_{\bar\lambda}(s) \le A_1(\gamma)$，即

$$\frac{1}{2}\, s\, v'(s) \ge \bar\lambda\left(v(s) + \frac{\gamma^2}{4}s^2\right) - A_1(\gamma), \qquad s \in \mathbb{R}.$$

应用这个不等式，其中 $s = q_i$ 对每个坐标 $i = 1, \ldots, d$ 都成立，并对 i 求和，可得

$$\frac{1}{2} q \cdot \nabla U(q) = \sum_{i=1}^{d} \frac{1}{2} q_i v'(q_i) \ge \bar\lambda\left(\sum_{i=1}^{d} v(q_i) + \frac{\gamma^2}{4}\sum_{i=1}^{d} q_i^2\right) - d\, A_1(\gamma),$$

由此得到 (2.6)，其中 $\lambda = \bar\lambda$ 且 $A = dA_1(\gamma)$。

扩展至所有更小的 λ。令 $\lambda \in (0, \bar\lambda]$ 为任意值。由于对所有 s，$v(s) \pm \frac{\gamma^2}{4}s^2 \ge 0$，因此逐点有 $\lambda\left(v(s) + \frac{\gamma^2}{4}s^2\right) \le \bar\lambda\left(v(s) + \frac{\gamma^2}{4}s^2\right)$，从而

$$\Delta_\lambda(s) := \lambda\left(v(s) + \frac{\gamma^2}{4}s^2\right) - \frac{1}{2}\, s\, v'(s) \le \Delta_{\bar\lambda}(s) \le A_1(\gamma), \qquad s \in \mathbb{R}.$$

等价地，

$$\frac{1}{2}\, s\, v'(s) \geq \lambda \left(v(s) + \frac{\gamma^2}{4} s^2 \right) - A_1(\gamma), \qquad s \in \mathbb{R},$$

并对各坐标求和，可得 (2.6)，这对每个 $\lambda \in (0, 1/(4+\gamma^2)]$ 均成立，且具有相同的常数$A = dA_1(\gamma)$。

(c) 梯度的渐近线性增长。最后，我们验证假设4.1。取$Q_\infty := I_d$。由于 $|v'(s) - s| \leq 1$ 对所有 s均成立，因此有 $|\nabla U(q) - q| \leq \sqrt{d}$。对于 $|q| \geq \sqrt{d}$，定义$\varrho(r) := \sqrt{d}/r \leq 1$。于是

$$|\nabla U(q) - q| \leq \sqrt{d} = \varrho(|q|)\, |q|, \qquad |q| \geq \sqrt{d},$$

并且显然，当 $r \to \infty$ 时（固定 d），$\varrho(r) \to 0$。至此，证明完毕。 □

D.2 命题5.2的证明

证明。我们证明 (i)–(iii)。

步骤 *0*：矩阵形式及预备常数。写出

$$a := \frac{2+\gamma^2}{4\gamma}, \qquad b := \frac{1}{2\gamma}, \qquad \mathcal{M}(q,p) = a|q|^2 + b|p|^2.$$

因此 $\mathcal{M}(z) = \frac{1}{2} z^\top \mathsf{K} z$，其中 $z = (q,\ p)$，并且

$$\mathsf{K} = \begin{pmatrix} 2a\, I_d & 0 \\ 0 & 2b\, I_d \end{pmatrix} = \begin{pmatrix} \frac{2+\gamma^2}{2\gamma} I_d & 0 \\ 0 & \frac{1}{\gamma} I_d \end{pmatrix}, \qquad \|\mathsf{K}\|_{\mathrm{op}} = \frac{2+\gamma^2}{2\gamma}.$$

对于多阱势，我们有 $U \geq 0$。利用 (2.7) 并展开，

$$\mathcal{V}_0(q,p) = U(q) + \frac{\gamma^2}{4}(1-\lambda)|q|^2 + \frac{1}{2}|p|^2 + \frac{\gamma}{2}\langle q, p\rangle.$$

舍去 $U(q)$，并将剩余的二次型写为

$$\frac{\gamma^2}{4}(1-\lambda)|q|^2 + \frac{1}{2}|p|^2 + \frac{\gamma}{2}\langle q,p\rangle = \frac{1}{4}\,(q,p)\cdot A\,(q,p), \qquad A := \begin{pmatrix} \gamma^2(1-\lambda) I_d & \gamma I_d \\ \gamma I_d & 2 I_d \end{pmatrix},$$

我们得到 (5.6)，其中 $c_1^{\mathrm{MW}} = \frac{1}{4}\lambda_{\min}(A_1)$，且 $A_1 = \begin{pmatrix} \gamma^2(1-\lambda & \gamma \\ \gamma & 2 \end{pmatrix}$。的特征值可显式求得；因此 (5.5) 得证。

最后，由于 $|\mathcal{M}(q,\ p)| \leq a|q|^2 + b|p|^2 \leq \frac{\|\mathsf{K}\|_{\mathrm{op}}}{2}(|q|^2+|p|^2)$，结合式 (5.6) 得到式 (5.8)，其中 $\widetilde{C}_{\mathcal{M}}^{\mathrm{MW}} = \frac{\|\mathsf{K}\|_{\mathrm{op}}}{2c_1^{\mathrm{MW}}} = \frac{2+\gamma^2}{4\gamma c_1^{\mathrm{MW}}}$，从而证明了 (ii)。

步骤 *1*：证明 *(i)*。对于这个例子，我们可以写成 $\nabla U(q) = q + r(q)$，其中 $|r(q)| \leq \sqrt{d}$ 对所有 q(参见命题 5.1，其中 $Q_\infty = I_d$)。直接计算可得

$$\mathcal{A}_0 \mathcal{M}(q,p) + \mathcal{A}' \mathcal{V}_0(q,p) = -B|q|^2 - |p|^2 + \mathcal{R}_{\mathrm{total}}(q,p), \tag{D.1}$$

其中 $B = 1 + \frac{\gamma^2}{2}(1-\lambda)$，并且

$$\mathcal{R}_{\text{total}}(q,p) = -|r|^2 - C_q\langle q, r\rangle - C_p\langle p, r\rangle, \qquad C_q := 2 + \frac{\gamma^2}{2}(1-\lambda), \quad C_p := \frac{1}{\gamma} + \frac{\gamma}{2}.$$

我们使用 $|ab| \le \frac{\varepsilon}{2}a^2 + \frac{1}{2\varepsilon}b^2$ 对交叉项进行界定。对于 p 项，取 $a = \sqrt{C_p}|p|$、$b = \sqrt{C_p}|r|$ 和 $\varepsilon = 1/C_p$：

$$C_p|\langle p, r\rangle| \le \frac{1}{2}|p|^2 + \frac{C_p^2}{2}|r|^2.$$

对于 q 项，取 $a = \sqrt{C_q}|q|$、$b = \sqrt{C_q}|r|$ 和 $\varepsilon = B/C_q$：

$$C_q|\langle q, r\rangle| \le \frac{B}{2}|q|^2 + \frac{C_q^2}{2B}|r|^2.$$

舍去该项 $-|r|^2 \le 0$ 并在 $|r|^2 \le d$ 中使用（D.1），得到

$$\mathcal{A}_0\mathcal{M} + \mathcal{A}'\mathcal{V}_0 \le -\left(\frac{B}{2}|q|^2 + \frac{1}{2}|p|^2\right) + \left(\frac{C_q^2}{2B} + \frac{C_p^2}{2}\right) d.$$

定义

$$Q(q,p) := \frac{B}{2}|q|^2 + \frac{1}{2}|p|^2, \qquad C(\gamma,\lambda) := \frac{C_q^2}{2B} + \frac{C_p^2}{2}.$$

Thus

$$\mathcal{A}_0\mathcal{M} + \mathcal{A}'\mathcal{V}_0 \le -Q(q,p) + C(\gamma,\lambda)\, d. \tag{D.2}$$

接下来，我们将 Q 与 $\mathcal{V}_0$ 联系起来。利用 $v(s) \le \frac{1}{2}s^2 + \frac{1}{4}$ 以及可分性，

$$U(q) = \sum_{i=1}^{d} v(q_i) \le \frac{1}{2}|q|^2 + \frac{d}{4}.$$

因此，再次利用 $\mathcal{V}_0$ 的显式展开（参见（2.7））

$$\mathcal{V}_0(q,p) \le \widetilde{\mathcal{V}}_0(q,p) + \frac{d}{4}, \qquad \widetilde{\mathcal{V}}_0(q,p) := \frac{B}{2}|q|^2 + \frac{1}{2}|p|^2 + \frac{\gamma}{2}\langle q, p\rangle.$$

作为二次型，Q 和 $\mathcal{V}_0$ 分解为相同的 2×2 块。因此，只需考虑

$$Q_{\text{mat}} := \begin{pmatrix} B/2 & 0 \\ 0 & 1/2 \end{pmatrix}, \qquad P_{\text{mat}} := \begin{pmatrix} B/2 & \gamma/4 \\ \gamma/4 & 1/2 \end{pmatrix}.$$

由于 $\det(P_{\text{mat}}) = \frac{4B-\gamma^2}{16} = \frac{4+\gamma^2(1-2\lambda)}{16} > 0$（对于 $\lambda \le 1/4$），我们有 $P_{\text{mat}} \succ 0$。因此，$Q \ge c\,\mathcal{V}_0$ 中的最佳常数 c 是最小的广义特征值，即

$$c_{\text{imp}} = \inf_{z\neq 0} \frac{z^\top Q_{\text{mat}} z}{z^\top P_{\text{mat}} z} = \frac{2\sqrt{B}}{2\sqrt{B} + \gamma}.$$

将其与（D.2）结合可得

$$\mathcal{A}_0\mathcal{M} + \mathcal{A}'\mathcal{V}_0 \le -c_{\text{imp}}\widetilde{\mathcal{V}}_0(q,p) + C(\gamma,\lambda)\, d.$$

最后，由于 $\mathcal{V}_0 \le \widetilde{\mathcal{V}}_0 + \frac{d}{4}$，我们有 $-\widetilde{\mathcal{V}}_0 \le -\mathcal{V}_0 + \frac{d}{4}$，因此

$$\mathcal{A}_0\mathcal{M} + \mathcal{A}'\mathcal{V}_0 \le -c_{\mathrm{imp}}\mathcal{V}_0(q,p) + \left(C(\gamma,\lambda) + \frac{c_{\mathrm{imp}}}{4}\right) d.$$

这就证明了（5.4），其中

$$C_{\mathrm{imp}}^{(d)} := \left(C(\gamma,\lambda) + \frac{c_{\mathrm{imp}}}{4}\right) d.$$

步骤 *2*：证明（*iii*）。根据（5.9）中的定义，

$$\mathrm{Err}^{(d)}(q,p) = |\mathcal{A}'\mathcal{M}(q,p)| + |\Delta_q\mathcal{M}(q,p)|.$$

由于 $\mathcal{M}(q,\ p) = a|q|^2 + b|p|^2$，

$$\Delta_q\mathcal{M}(q,p) = 2ad = \frac{2+\gamma^2}{2\gamma}\, d.$$

此外，$\mathcal{A}' = -\nabla U(q)\cdot\nabla_q$ 且 $\nabla_q\mathcal{M}(q,\ p) = 2a\,q$。利用 $|\nabla U(q)| \le L|q| + |\nabla U(0)| = |q|$ 和杨氏不等式，

$$|\mathcal{A}'\mathcal{M}(q,p)| = |\nabla U(q)\cdot\nabla_q\mathcal{M}(q,p)| \le |q|\cdot 2a|q| = 2a|q|^2 \le \frac{2a}{c_1^{\mathrm{MW}}}\,\mathcal{V}_0(q,p).$$

因此（5.10）成立，其中

$$C_2^{\mathrm{MW}} := \frac{2a}{c_1^{\mathrm{MW}}} = 2\widetilde{C}_{\mathcal{M}}^{\mathrm{MW}} \quad \text{and} \quad C_2^{(d),\mathrm{MW}} := \frac{2+\gamma^2}{2\gamma}\, d.$$

最后，取 $c_{\mathrm{imp}} = c_{\mathrm{imp}}$（来自（i））和 $\widetilde{C}_{\mathcal{M}} = \widetilde{C}_{\mathcal{M}}^{\mathrm{MW}}$（来自（ii）），更新后的命题4.7 给出了漂移率展开（5.12），其中 δ_{MW} 以及$C_{\lambda,\ \mathrm{MW}}$ 如（5.11）所示。这就完成了证明。 □

D.3 引理 5.3 的证明

证明。步骤 *1*：耗 散性（多阱情形）。对于 $|s| \le 1/2$，我们有 $v'(s) = -s$ 且 $s^2 \le 1/4$。因此

$$s\,v'(s) = -s^2 \ge \lambda s^2 - \frac{1+\lambda}{4}. \tag{D.3}$$

对于 $|s| \ge 1/2$，我们有 $v'(s) = s - \mathrm{sign}(s)$。因此，$s\,v'(s) = s^2 - |s|$。令 $x := |s| \ge 1/2$。那么对于任意 $\lambda \in (0,1)$，

$$x^2 - x = (1-\lambda)x^2 - x + \lambda x^2 \ge -\frac{1}{4(1-\lambda)} + \lambda x^2,$$

因为 $\inf_{x\ge 0}\{(1-\lambda)x^2 - x\} = -\frac{1}{4(1-\lambda)}$。因此，

$$s\,v'(s) \ge \lambda s^2 - \frac{1}{4(1-\lambda)}, \qquad |s| \ge \frac{1}{2}. \tag{D.4}$$

结合两种情形（D.3）和（D.4），可得一维耗散性界

$$s\,v'(s) \ge \lambda s^2 - D_0(\lambda), \qquad D_0(\lambda) := \max\left\{\frac{1+\lambda}{4},\ \frac{1}{4(1-\lambda)}\right\}.$$

特别地，这对每个 $\lambda \in (0, 1/4]$ 都成立（加性常数 $D_0(\lambda)$ 有限）。对于 d 维可分势函数 $U(q) = \sum_{i=1}^{d} v(q_i)$，对各坐标求和可得 $\langle q, \nabla U(q)\rangle \geq \lambda |q|^2 - d\, D_0(\lambda)$。

步骤 *2*：当 λ较小时 $\delta_{\mathrm{MW}} > \gamma\lambda$ 的可行性。回顾 $\delta_{\mathrm{MW}} = c_{\mathrm{imp}} - \lambda \widetilde{C}_{\mathcal{M}}^{\mathrm{MW}}$。定义 e

$$F(\lambda) := \delta_{\mathrm{MW}} - \gamma\lambda = c_{\mathrm{imp}}(\lambda) - \left(\gamma + \widetilde{C}_{\mathcal{M}}^{\mathrm{MW}}(\lambda)\right) \lambda,$$

其中 $c_{\mathrm{imp}}(\lambda)$ 和 $\widetilde{C}_{\mathcal{M}}^{\mathrm{MW}}$($\lambda$在命题中有明确给出5.2。二者关于 $\lambda \in [0, 1/4]$ 均连续，且在 $\lambda = 0$处有限。此外，

$$F(0) = c_{\mathrm{imp}}(0) = \frac{2\sqrt{1 + \gamma^2/2}}{2\sqrt{1 + \gamma^2/2} + \gamma} > 0.$$

因此，根据连续性，存在 $\lambda_\star(\gamma) \in (0,\ 1/4]$，使得对于所有 $\lambda \in (0,\ \lambda_\star(\gamma)],$，都有 $F(\lambda) > 0$，即 $\delta_{\mathrm{MW}} > \gamma\lambda$。至此，证明完毕。 □

D.4 定理 5.4证明

证明。 我们将推论4.13 应用于一维的多阱模型。通过在引理 5.3中选择 λ，可确保条件 $\delta_{\mathrm{MW}} > \gamma\lambda$。随后，通过对代价函数 $\rho_{\alpha,\ d} = \sum_{i=1}^{d} \rho_{\alpha,\ 1}$ 进行张量化，并利用 U的乘积结构，即可得到 d 维的结论。

步骤 *1*：具有显式常数的一维加速收缩。首先考虑 $d = 1$。根据命题 5.1，一维势函数 v 满足假设 2.1。此外，命题 5.2 给出了一个显式的二次校正项 $\mathcal{M}$，使得一阶改进条件在$c_{\mathrm{imp}} = c_{\mathrm{imp}}$下成立；该命题还给出了 $\widetilde{C}_{\mathcal{M}}^{\mathrm{MW}}$和 C_2^{MW}的显式取值，用于控制扰动项。因此，命题 4.7 适用于维度为 1 的情形，并给出改进漂移

$$\lambda_\alpha \ \geq\ \lambda + \delta_{\mathrm{MW}}\alpha - C_{\lambda,\mathrm{MW}}\alpha^2, \qquad \delta_{\mathrm{MW}} := c_{\mathrm{imp}} - \lambda\, \widetilde{C}_{\mathcal{M}}^{\mathrm{MW}}, \quad C_{\lambda,\mathrm{MW}} := C_2^{\mathrm{MW}} + \widetilde{C}_{\mathcal{M}}^{\mathrm{MW}}\, c_{\mathrm{imp}}.$$

特别是，若 $\delta_{\mathrm{MW}} > 0$，则 $\lambda_\alpha > \lambda$对所有充分小的 α均成立。将推论 4.13 应用于维度为 1 的情形，可得常数 $\alpha_{\mathrm{MW}} > 0$ 和 $\kappa_{\mathrm{MW}} > 0$ （仅取决于一维模型和 γ），使得对于所有 $\alpha \in (0, \alpha_{\mathrm{MW}}]$ 以及所有定义在 μ, ν上的概率测度 $\mathbb{R}^2$,

$$\mathcal{W}_{\rho_{\alpha,1}}\left(\mu P_t^{\alpha,(1)}, \nu P_t^{\alpha,(1)}\right) \leq e^{-(c_0 + \kappa_{\mathrm{MW}}\alpha)t}\, \mathcal{W}_{\rho_{\alpha,1}}(\mu, \nu), \qquad t \geq 0. \tag{D.5}$$

这里， c_0 是 $\alpha = 0$处的一维收缩率。

步骤 *2*：张量化。由于 $U(q) = \sum_{i=1}^{d} v(q_i)$ 是可分离的，且驱动布朗运动在各坐标上相互独立， d 维 HFHR 动力学可分解为 d 个相互独立的一维 HFHR 动力学副本，因此 $P^{\alpha,}{}_{t}{}^{(d)} = \bigotimes_{i=1}^{d} P_t^{\alpha,\ (1)}$。

固定任意一个 π与 μ和 ν的耦合，并令 $(Z, Z') \sim \pi$，其中 $Z = (Z_1, \ldots, Z_d)$ 和 $Z' = (Z'_1, \ldots, Z'_d)$，且 $Z_i, Z'_i \in \mathbb{R}^2$。在给定 (Z, Z') 的条件下，对每个坐标分别运行相互独立的（近乎最优）一维耦合，并将时刻 t 所得的耦合对记为 (Z_t, Z'_t)。由代价函数的可加性 $\rho_{\alpha,d}(z, z') = \sum_{i=1}^{d} \rho_{\alpha,1}(z_i, z'_i)$，以及独立性，

$$\mathbb{E}\left[\rho_{\alpha,d}(Z_t, Z'_t) \,|\, Z, Z'\right] = \sum_{i=1}^{d} \mathbb{E}\left[\rho_{\alpha,1}(Z_{t,i}, Z'_{t,i}) \,|\, Z_i, Z'_i\right].$$

对每个坐标应用一维收缩（D.5）可得

$$\mathbb{E}\big[\rho_{\alpha,d}(Z_t, Z_t') \,|\, Z, Z'\big] \le e^{-(c_0+\kappa_{\mathrm{MW}}\alpha)t} \sum_{i=1}^{d} \rho_{\alpha,1}(Z_i, Z_i') = e^{-(c_0+\kappa_{\mathrm{MW}}\alpha)t}\, \rho_{\alpha,d}(Z, Z').$$

取期望，然后对所有耦合取下确界 π ，得到

$$\mathcal{W}_{\rho_{\alpha,d}}\Big(\mu P_t^{\alpha,(d)}, \nu P_t^{\alpha,(d)}\Big) \le e^{-(c_0+\kappa_{\mathrm{MW}}\alpha)t}\, \mathcal{W}_{\rho_{\alpha,d}}(\mu,\nu), \qquad t \ge 0,$$

对所有 $\alpha \in (0,\ \alpha_{\mathrm{MW}}]$。

步骤 *3*：维数无关性。常数 α_{MW} 和 κ_{MW} 完全来自一维估计（D.5），因此不依赖于 d。具体而言，我们取 $\alpha_{\mathrm{MW}} := \alpha_{\mathrm{global}}$如推论4.13 中针对 $d = 1$ （其中 $L = 1$）的情形所定义的值，它是第4节中导出的分支加速阈值与度量加速阈值中的较小者。关于 $c_\alpha^{(d)}$的界立即可得。 □

D.5 命题5.6的证明

证明。(a) 由于 $\varepsilon > 0$，每个映射 $q_j \mapsto (q_j^2 + \varepsilon^2)^{p/2}$ 在 $\mathbb{R}$上光滑；因此 $g \in C^\infty$ ，从而$U \in C^\infty$。于是，假设2.1(i) 成立。直接计算可得

$$\nabla^2 U(q) = \frac{1}{\sigma^2} X^\top X + \nabla^2 g(q), \qquad \nabla^2 g(q) = \iota\, \mathrm{diag}\left(\psi''(q_j)\right)_{j=1}^d,$$

其中 $\psi(t) := (t^2 + \varepsilon^2)^{p/2}$ 且

$$\psi''(t) = p\left(t^2+\varepsilon^2\right)^{\frac{p}{2}-2}\left(\varepsilon^2 + (p-1)t^2\right) \ge 0.$$

对于 $\psi''(t)$ 的上界，可以使用 $\varepsilon^2 + (p-1)t^2 \le \varepsilon^2 + t^2$ 得到

$$\psi''(t) \le p\left(t^2+\varepsilon^2\right)^{\frac{p}{2}-2}(t^2+\varepsilon^2) = p(t^2+\varepsilon^2)^{\frac{p}{2}-1} \le p(\varepsilon^2)^{\frac{p}{2}-1} = p\varepsilon^{p-2}.$$

因此 $\|\nabla^2 g(q)\|_{\mathrm{op}} \le \iota p \varepsilon^{p-2}$，并且

$$\|\nabla^2 U(q)\|_{\mathrm{op}} \le \frac{\|X^\top X\|_{\mathrm{op}}}{\sigma^2} + \iota p \varepsilon^{p-2} = \frac{M}{\sigma^2} + \iota p \varepsilon^{p-2},$$

这给出了 ∇U的所需全局 Lipschitz 常数。因此，假设2.1(ii)成立。

(b) 我们可以计算得

$$\nabla U(q) = \frac{1}{\sigma^2} X^\top X q - \frac{1}{\sigma^2} X^\top y + \nabla g(q), \qquad \text{with} \quad \nabla g(q) = \iota p \left(q_j (q_j^2+\varepsilon^2)^{\frac{p}{2}-1}\right)_{j=1}^d.$$

注意，$\langle \nabla g(q)\ q\rangle = \iota p \sum_{j=1}^d q_j^2 \left(q_j^2+\varepsilon^2\right)^{\frac{p}{2}-1} \ge 0$。利用 $X^\top X \succeq m I_d$ 和柯西–施瓦茨不等式，

$$\langle \nabla U(q), q\rangle \ge \frac{1}{\sigma^2}\left\langle X^\top X q, q\right\rangle - \frac{1}{\sigma^2}\left\langle X^\top y, q\right\rangle \ge \frac{m}{\sigma^2}|q|^2 - \frac{|X^\top y|}{\sigma^2}|q|.$$

配方可得，对于所有 q,

$$\frac{m}{\sigma^2}|q|^2 - \frac{|X^\top y|}{\sigma^2}|q| \ge \frac{m}{2\sigma^2}|q|^2 - \frac{|X^\top y|^2}{2m\sigma^2},$$

这证明了耗散性。因此，假设2.1（iii）成立。

（c）设定 $Q_\infty = \sigma^{-2} X^\top X$。于是

$$\nabla U(q) - Q_\infty q = -\frac{1}{\sigma^2} X^\top y + \iota p\, v(q), \qquad v_j(q) := q_j (q_j^2 + \varepsilon^2)^{\frac{p}{2}-1}.$$

对于每个坐标，都有以下初等界（例如将 $|q_j| \geq \varepsilon$ 和 $|q_j| < \varepsilon$ 分开）：

$$|v_j(q)| \leq |q_j|^{p-1} + \varepsilon^{p-1}.$$

因此，

$$|v(q)| \leq \left(\sum\nolimits_{j=1}^{d} |q_j|^{2p-2} \right)^{1/2} + \sqrt{d}\, \varepsilon^{p-1}.$$

回顾标准向量 r 范数的记号： $\|q\|_r := (\sum_{j=1}^d |q_j|^r)^{1/r}$ ，其中 $r > 0$。利用范数关系 $\|q\|_r \leq d^{\frac{1}{r}-\frac{1}{2}} \|q\|_2$（对于 $0 < r < 2$）（此处取 $r = 2p-2$ 应用该关系，并注意到 $1 < p < 2$ 蕴含 $0 < 2p-2 < 2$），我们有

$$\left(\sum\nolimits_{j=1}^{d} |q_j|^{2p-2} \right)^{1/2} = \|q\|_{2p-2}^{p-1} \leq \left(d^{\frac{1}{2p-2}-\frac{1}{2}} \|q\|_2 \right)^{p-1} = d^{\frac{2-p}{2}} |q|^{p-1}.$$

因此，对于所有 q，

$$|\nabla U(q) - Q_\infty q| \leq \frac{|X^\top y|}{\sigma^2} + \iota p \left(d^{\frac{2-p}{2}} |q|^{p-1} + \sqrt{d}\, \varepsilon^{p-1} \right) = c_0^{\mathrm{LR}} + c_1^{\mathrm{LR}} |q|^{p-1},$$

其中 c_0^{LR}， c_1^{LR}如（5.15）所述。两边除以 $|q|$ （当 $|q| \geq 1$ 时）可得

$$|\nabla U(q) - Q_\infty q| \leq \left(\frac{c_0^{\mathrm{LR}}}{|q|} + c_1^{\mathrm{LR}} |q|^{p-2} \right) |q| = \varrho(|q|)|q|.$$

由于 $p - 2 < 0$，两项 c_0^{LR}/r 和 $c_1^{\mathrm{LR}} r^{p-2}$ 均随 r递减，并且当 $\varrho(r) \to 0$ 时， $r \to \infty$。这验证了假设4.1。证明完毕。 □

D.6 命题 5.7的证明

证明。 将引理 4.3 直接应用于（5.13）中定义的贝叶斯线性回归模型，并利用命题 5.6中推导出的显式性质和界，即可直接得到这些结果。具体而言，将 U 和 ∇U 的具体形式代入一般框架，即可得到谱界、尾模量和校正项构造。

□

D.7 引理 5.8的证明

证明。步骤 *1*：*d* 贝叶斯线性回归的耗散性。 根据命题 5.6（b），对于所有 $q \in \mathbb{R}^d$:

$$\langle \nabla U(q), q \rangle \geq \frac{m}{2\sigma^2} |q|^2 - \frac{|X^\top y|^2}{2m\sigma^2}. \tag{D.6}$$

任取 $\lambda \in (0, \bar{\lambda}]$，其中 $\bar{\lambda} \leq m/(2\sigma^2)$ 。弱化（D.6）中的二次系数可得

$$\langle \nabla U(q), q \rangle \geq \lambda |q|^2 - \frac{|X^\top y|^2}{2m\sigma^2},$$

这恰好就是假设 2.1(iii)（相差一个加性常数），从而证明了 (i)。

步骤 2：关于$c_{\mathrm{imp}}(\lambda)$的显式一致下界。固定 $\lambda \in (0, \lambda]$。在引理4.3中，一阶改进常数$c_{\mathrm{imp}}(\lambda)$可按 (4.9) 选取，其中

$$a_{\min}(\lambda) = \lambda_{\min}(Q_\infty) + \frac{\gamma^2}{2}(1-\lambda), \qquad a_{\max}(\lambda) = \lambda_{\max}(Q_\infty) + \frac{\gamma^2}{2}(1-\lambda).$$

由于 $\lambda \le \lambda$，我们有如下确定性界

$$a_{\min}(\lambda) \ge m_\infty + \frac{\gamma^2}{2}(1-\bar{\lambda}) = a_{\min}^-, \qquad a_{\max}(\lambda) \le M_\infty + \frac{\gamma^2}{2} = a_{\max}^+.$$

此外，由命题5.7(ii)，对所有 $R' \ge 1$,

$$\delta_U(R') \le \frac{|X^\top y|}{\sigma^2 R'} + \iota\, d^{1-\frac{p}{2}}\,(R')^{p-2} + \frac{\iota\, d\, \varepsilon^p + \frac{1}{2\sigma^2}|y|^2}{(R')^2}.$$

在引理4.3 中，截断半径满足 $R_0(\lambda) \ge \max\{1, C_{\mathrm{linear}}\} = R$，且由于 $\delta_U(\cdot)$ 关于其自变量非增，

$$\delta_U(R_0(\lambda)) \le \delta_U(R) \le \delta_U^+,$$

其中 δ_U^+ 由 (5.22) 给出。将上述三个界代入 (4.9) 得到

$$c_{\mathrm{imp}}(\lambda) \ge c_{\mathrm{imp}}^-,$$

其中c_{imp}^- 定义见 (5.27)。

步骤 3：关于 $\widetilde{C}_{\mathcal{M}}^{\mathrm{LR}}(\lambda)$ 的显式一致上界。根据引理 4.3，校正子可以选为一个二次函数：$\mathcal{M}(z) = \frac{1}{2} z^\top \mathsf{K}(\lambda) z$，其中

$$\mathsf{K}(\lambda) = \int_0^\infty e^{tB^\top}\, C_{B_1}(\lambda)\, e^{tB}\, dt.$$

取算子范数可得

$$\|\mathsf{K}(\lambda)\|_{\mathrm{op}} \le \left(\int_0^\infty \|e^{tB}\|_{\mathrm{op}}^2\, dt\right) \|C_{B_1}(\lambda)\|_{\mathrm{op}}.$$

由于 $C_{B_1}(\lambda)$ 是一个显式对称矩阵，且仅通过系数 λ 依赖于 $Q_\infty + \frac{\gamma^2}{2}(1-\lambda) I_d$（见引理 4.3），一个粗略但显式的界给出

$$\|C_{B_1}(\lambda)\|_{\mathrm{op}} \le 2\left(1 + \gamma + \|Q_\infty\|_{\mathrm{op}} + \frac{\gamma^2}{2}\right) = C_{B_1}^+,$$

其中 $C_{B_1}^+$ 定义见（5.26）。

接下来我们对 $\int_0^\infty \|e^{tB}\|_{\mathrm{op}}^2\, dt$ 进行界定。将 $Q_\infty = S^\top \mathrm{diag}(\nu_1, \dots, \nu_d) S$ 对角化，其中 S 正交且$\nu_i \in [m_\infty, M_\infty]$。于是，$B$ 与一个分块对角矩阵正交相似，该矩阵有 2×2 个块 $B_\nu := \left(\begin{smallmatrix} 0 & 1 \\ -\nu & -\gamma \end{smallmatrix}\right)$。直接计算 e^{tB_ν}（等价地，$x'' + \gamma x' + \nu x = 0$ 的基矩阵）可得一致界

$$\|e^{tB}\|_{\mathrm{op}} = \max_{1 \le i \le d} \|e^{tB_{\nu_i}}\|_{\mathrm{op}} \le C_B e^{-\eta t},$$

其中 η 且 C_B 如（5.25）中所示。因此，

$$\int_0^\infty \|e^{tB}\|_{\mathrm{op}}^2\,dt \le \int_0^\infty C_B^2 e^{-2\eta t}\,dt = \frac{C_B^2}{2\eta}.$$

将积分的界和一致界 $\|C_{B_1}(\lambda)\|_{\mathrm{op}} \le C_{B_1}^+$ 代入关于 $\|\mathsf{K}(\lambda)\|_{\mathrm{op}}$ 的不等式，可得一致估计

$$\|\mathsf{K}(\lambda)\|_{\mathrm{op}} \le \frac{C_B^2}{2\eta}\, C_{B_1}^+.$$

此外，由于 $\lambda \le \lambda$， $\mathcal{V}_0$的二次下界常数满足 $c_1(\gamma,\ \lambda) \ge c_1(\gamma,\ \lambda) = c_1^-$（因为 $1-\lambda \ge 1-\lambda$ 增大了用于定义该界的 2×2 块）。因此，

$$\widetilde{C}_{\mathcal{M}}^{\mathrm{LR}}(\lambda) = \frac{\|\mathsf{K}(\lambda)\|_{\mathrm{op}}}{2c_1(\gamma,\lambda)} \le \frac{1}{2c_1}\cdot\frac{C_B^2}{2\eta}\, C_{B_1}^+ = \widetilde{C}_{\mathcal{M}}^{+},$$

其中 $\widetilde{C}_{\mathcal{M}}^{+}$ 定义见（5.26）。

步骤 *4*：在一个显式区间上得出 $\delta_{\mathrm{LR}} > \gamma\lambda$。对于任意 $\lambda \in (0,\ \lambda]$，

$$\delta_{\mathrm{LR}} - \gamma\lambda = \underline{c}_{\mathrm{imp}}(\lambda) - \left(\gamma + \widetilde{C}_{\mathcal{M}}^{\mathrm{LR}}(\lambda)\right)\lambda \ge \underline{c}_{\mathrm{imp}}^{-} - \left(\gamma + \widetilde{C}_{\mathcal{M}}^{+}\right)\lambda.$$

因此，若 $\gamma + \widetilde{C}_{\mathcal{M}}^{+}$，则 $\delta_{\mathrm{LR}} > \gamma\lambda$。结合 $\lambda \le \bar{\lambda}$ 可得显式选择（5.28），从而证明（ii）。证明完毕。

□

D.8 定理5.9的证明

证明。该结果可直接由推论4.13得出。命题5.6证明了假设2.1 和4.1 成立。命题5.7 给出了二次修正项 $\mathcal{M}$ 的显式构造，并证明了一阶漂移改进，其显式常数为 δ_{LR} 和$C_{\lambda,\ \mathrm{LR}}$。引理5.8 保证，选择 $\lambda \le \lambda_\star(\gamma)$ 后，加速条件 $\delta_{\mathrm{LR}} > \gamma\lambda$ 得到满足。因此，推论4.13的所有条件均已满足，这意味着加速常数 α_{LR} 和 κ_{LR}存在。证明完毕。

□

D.9 命题5.11的证明

证明。(a) 由于 $h,\ \varphi \in C^2$ 且 $q \mapsto \langle q,\ x_i\rangle$ 是线性的，因此每个求和项 $q \mapsto \varphi(y_i - h(\langle q,\ x_i\rangle))$ 都是 C^2，因而 U也是如此。此外， $\varphi \ge 0$和 $\frac{\iota}{2}|q|^2 \ge 0$推出 $U \ge 0$。因此，假设2.1(i) 成立。

通过对 U 求导，我们可以计算得出

$$\nabla U(q) = -\frac{1}{n}\sum_{i=1}^{n} \varphi'(y_i - h(\langle q, x_i\rangle))\ h'(\langle q, x_i\rangle)\, x_i + \iota q.$$

令 $s_i(q) := \langle q, x_i\rangle$ 和 $t_i(q) := y_i - h(s_i(q))$。对 ∇U 求导，得到

$$\nabla^2 U(q) = \iota I_d + \frac{1}{n}\sum_{i=1}^{n} a_i(q)\, x_i x_i^\top,$$

with

$$a_i(q) = \varphi''(t_i(q))\,\big(h'(s_i(q))\big)^2 - \varphi'(t_i(q))\,h''(s_i(q)).$$

利用 $|\varphi'|$、$|\varphi''|$、$|h'|$ 和 $|h''|$ 的一致界，对于所有 q，我们得到：

$$|a_i(q)| \le \Phi_2 H_1^2 + \Phi_1 H_2.$$

因此，

$$\|\nabla^2 U(q)\|_{\mathrm{op}} \le \iota + \frac{1}{n}\sum_{i=1}^{n} |a_i(q)|\,\|x_i x_i^\top\|_{\mathrm{op}} \le \iota + (\Phi_2 H_1^2 + \Phi_1 H_2)\,B_x^2,$$

由于 $\|x_i x_i^\top\|_{\mathrm{op}} = |x_i|^2 \le B_x^2$。因此，假设 2.1（ii）成立。

（b）根据梯度表达式和柯西–施瓦茨不等式，

$$|\nabla U(q) - \iota q| \le \frac{1}{n}\sum_{i=1}^{n} |\varphi'(t_i(q))|\,|h'(s_i(q))|\,|x_i| \le \Phi_1 H_1 B_x =: C_0.$$

因此

$$\langle \nabla U(q), q\rangle = \iota|q|^2 + \langle \nabla U(q) - \iota q,\, q\rangle \ge \iota|q|^2 - C_0|q| \ge \frac{\iota}{2}|q|^2 - \frac{C_0^2}{2\iota},$$

其中，我们配方得到了最后一个不等式。这给出了所需的耗散不等式。因此，假设2.1（iii）成立。

（c）根据 $Q_\infty = \iota I_d$ 和界 $|\nabla U(q) - \iota q| \le C_0$，对于 $|q| \ge 1$，我们有

$$|\nabla U(q) - Q_\infty q| \le C_0 = \frac{C_0}{|q|}\,|q| = \varrho(|q|)\,|q|.$$

由于 $\varrho(r) = C_0/r$ 是递减的，且在无穷远处趋于零，因此假设4.1 成立。证明完毕。

□

D.10 命题5.12的证明

证明。将引理4.3直接应用于式（5.29）定义的贝叶斯二元分类模型即可得到结果。谱界、尾模量和校正项构造，可通过将命题5.11中推导出的具体势函数性质代入一般框架得到。

□

D.11 引理5.13的证明

证明。步骤 *1*：对任意 $\lambda \le \iota/2$，耗散性。由命题5.11(b)（其中岭系数记为 ι），对所有 $q \in \mathbb{R}^d$，$\langle \nabla U(q)\ q\rangle \ge \frac{\iota}{2}|q|^2 - \frac{C_0^2}{2\iota}$。

固定任意 $\lambda \in (0,\ \bar{\lambda}]$；由于 $\bar{\lambda} \le \iota/2$，我们得到

$$\langle \nabla U(q), q\rangle \ge \lambda|q|^2 - \frac{C_0^2}{2\iota},$$

即为假设 2.1（iii）（相差一个加性常数）。

¯步骤 2：$c_{\mathrm{imp}}(\lambda)$ 的一致下界。对于 $\lambda \in (0, \lambda]$，根据引理 4.3，我们有 $a_{\min} = a_{\max} = a(\lambda) = \iota + \frac{\gamma^2}{2}(1 - \lambda) \geq a^-$。此外，截断半径满足 $R_0(\lambda) \geq R$，并且由 (5.32) 和单调性可得，$\delta_U(R_0(\lambda)) \leq \delta_U(R) \leq \delta_U^+$。将这些界代入 (4.9) 得到

$$\underline{c}_{\mathrm{imp}}(\lambda) \geq \underline{c}_{\mathrm{imp}}^-,$$

其中 $\underline{c}_{\mathrm{imp}}^-$ 定义于（5.41）（使用 a^-，其取决于 ι）。

步骤 *3*：给出$\widetilde{C}_{\mathcal{M}}^{\mathrm{BC}}(\lambda)$的一致上界。引理4.3 给出一个二次校正项 $\mathcal{M}(z) = \frac{1}{2} z^\top \mathsf{K}(\lambda) z$ 满足 $\mathsf{K}(\lambda) = \int_0^\infty e^{tB^\top} C_{B_1}(\lambda) e^{tB}\, dt$。取算子范数，

$$\|\mathsf{K}(\lambda)\|_{\mathrm{op}} \leq \left(\int_0^\infty \|e^{tB}\|_{\mathrm{op}}^2\, dt \right) \|C_{B_1}(\lambda)\|_{\mathrm{op}}. \tag{D.7}$$

由于 $Q_\infty = \iota I_d$，可得粗略上界 $\|C_{B_1}(\lambda)\|_{\mathrm{op}} \leq C_{B_1}^+$，其中 $C_{B_1}^+$ 按引理陈述中的定义（使用 ι）。此外，e^{tB} 的分块常微分方程表示（等价于阻尼振子 $x'' + \gamma x' + \iota x = 0$）意味着 $\|e^{tB}\|_{\mathrm{op}} \leq C_B e^{-\eta t}$，其中 η，C_B 使用 ι 定义。因此

$$\int_0^\infty \|e^{tB}\|_{\mathrm{op}}^2\, dt \leq \frac{C_B^2}{2\eta}. \tag{D.8}$$

因此，由（D.7）和（D.8）可得

$$\|\mathsf{K}(\lambda)\|_{\mathrm{op}} \leq \frac{C_B^2}{2\eta} C_{B_1}^+. \tag{D.9}$$

¯最后，对于 $\lambda \leq \lambda$，基准二次下界常数满足 $c_1(\gamma, \lambda) \geq c_1(\gamma, \lambda) = c_1$。因此，由 (D.9) 可得

$$\widetilde{C}_{\mathcal{M}}^{\mathrm{BC}}(\lambda) = \frac{\|\mathsf{K}(\lambda)\|_{\mathrm{op}}}{2c_1(\gamma, \lambda)} \leq \frac{1}{2c_1} \cdot \frac{C_B^2}{2\eta} C_{B_1}^+ = \widetilde{C}_{\mathcal{M}}^+.$$

¯步骤 *4*：在一个显式区间上得出结论 $\delta_{\mathrm{BC}} > \gamma\lambda$。对于 $\lambda \in (0, \lambda]$，

$$\delta_{\mathrm{BC}} - \gamma\lambda = \underline{c}_{\mathrm{imp}}(\lambda) - \left(\gamma + \widetilde{C}_{\mathcal{M}}^{\mathrm{BC}}(\lambda)\right)\lambda \geq \underline{c}_{\mathrm{imp}}^- - \left(\gamma + \widetilde{C}_{\mathcal{M}}^+\right)\lambda.$$

因此 $\delta_{\mathrm{BC}} > \gamma\lambda$ 当 $\lambda \leq c_{\mathrm{imp}}^-/(\gamma + \widetilde{C}_{\mathcal{M}}^+)$ 时成立。结合 $\lambda \leq \bar{\lambda}$ 可得 (5.42)。至此，证明完毕。

□

D.12 定理 5.14 的证明

证明。结论可直接由推论 4.13 得出。命题 5.11证明假设 2.1和 4.1 成立。命题 5.12 给出了二次修正项的显式构造 $\mathcal{M}$，并以显式常数 δ_{BC} 和 $C_{\lambda,\ \mathrm{BC}}$保证一阶漂移改进。引理 5.13 保证，选择 $\lambda \leq \lambda_\star(\gamma)$ 后，加速条件 $\delta_{\mathrm{BC}} > \gamma\lambda$ 得以满足。因此，推论 4.13 的所有条件均已满足，这意味着加速常数 α_{BC} 和 κ_{BC}存在。证明完毕。

□