arXiv:2611.28468v1 [math.PR] 21 Nov 2026

平滑适应 Wasserstein 距离的极限分布与自助法

Xiaozhen Wang*

2026年11月15日

摘要

本文针对高斯平滑适应 Wasserstein 距离给出极限定理。设 $\widehat{\mu}_n$ 为概率律 μ 在 $(\mathbb{R}^d)^T$ 上的独立路径的经验测度。令$P = \mu * \gamma_\sigma$和 $P_n = \widehat{\mu}_n * \gamma_\sigma$，其中 $\gamma_\sigma = \mathcal{N}(0, \sigma^2 I_{dT})$ 是均值为零、协方差矩阵为 $\sigma^2 I_{dT}$ 的高斯测度。我们首先假设 μ 具有紧支集。对于任意 $1 < p < \infty$ 和固定的 d, T和 $\sigma > 0$，我们证明

$$\sqrt{n}\,\mathcal{AW}_p(P_n, P) \implies Q_{P,p}(G)^{1/p},$$

其中 G是密度涨落的高斯极限，而$Q_{P,p}$ 是线性化连续性方程的最小适应p-能量。我们证明全路径自助法一致性，以及一阶矩和 p 阶矩的收敛。对于满足适当尾部条件和结构条件的所有 $1 < p < \infty$，该极限结果和自助法也适用于非紧支撑概率律。对于强对数凹分布的双侧有界密度扰动，当 $1 < p \le 2$ 时，带宽不受限制。对于 $p > 2$，我们给出充分的尾部条件，并提供一个高斯反例，其中小带宽下根-n 紧性不成立。我们的证明基于双因果传输的双端点局部展开。作为应用，我们利用自助法分位数，给出平滑模型下最优停止值的渐近误差界。

关键词 ： 适应Wasserstein距离；高斯平滑；极限分布；自助法 ；切空间.

1 引言

在多期决策问题中，比较两个随机过程概率律可能比仅仅在路径空间 $(\mathbb{R}^d)^T$上用普通 Wasserstein度量衡量它们的接近程度复杂得多。这是因为在时刻 t 所作出的决策只能使用截至该时刻可获得的信息。因此，就普通Wasserstein度量 W_p 而言，两个路径律可能非常接近，但在给定过去的情况下，它们未来状态的条件分布仍可能存在显著差异。这反过来又会改变最优停止或对冲问题的取值。为了沿着各自的滤过比较多阶段模型，Pflug [24] 以及随后 Pflug 和 Pichler [25] 提出了嵌套分布和嵌套距离的概念。最近，Backhoff-Veraguas、Bartl、Beiglböck 和 Eder 证明了各种适应弱拓扑 [6] 之间的等价性，并使用适应Wasserstein距离 $\mathcal{AW}_p$ （通过双因果耦合定义）证明了数理金融中的 Lipschitz 稳定性结果 [5].

当我们从总体稳定性转向根据数据估计分布时，对条件律的敏感性就成了一个问题。例如，假设我们观测到 n 条独立轨迹 $A^{(1)}, \dots, A^{(n)}$，它们来自一个 μ第一边缘分布无原子的概率律。那么，以概率

*CMAP，巴黎综合理工学院，巴黎理工学院，帕莱索，法国。

1，起点 $A_1^{(1)},\dots,A_1^{(n)}$ 彼此均不相同。在原始经验测度 $\widehat{\mu}_n$下，观测第一坐标便能识别抽样路径，并揭示其整个未来。因此， $\widehat{\mu}_n$ 不一定会在 μ 中收敛到 $\mathcal{AW}_p$ ，当 $n\to\infty$。Backhoff、Bartl、Beiglböck 和 Wiesel [4] 通过对状态空间进行划分来构造适应经验测度，并推导出定量收敛速率，从而解决了这一难题。

一种自然的替代方法是通过卷积平滑进行正则化：在计算传输距离之前，先将两个测度分别与高斯核 $\gamma_\sigma=\mathcal{N}(0,\sigma^2 I_{dT})$进行卷积。对于 Wasserstein 距离，这种平滑方法由 Goldfeld 和 Greenewald [16]提出，并在 [22], 中推广到一般情形 p ；相应的中心极限定理和自助法理论则在 [17]中建立。在多期设定中，拥有光滑的联合密度只是问题的一半。还必须控制条件密度随过去轨迹变化的方式。近期研究从多个方面探讨了这一问题：平滑经验测度在$\mathcal{AW}_p$ [19]中的依赖带宽的行为、比较适应性传输度量与经典传输度量的定量界 [8], ，以及$\mathcal{AW}_p$ 投影在经验定价中恢复鞅约束的作用[9]。近期，Larsson、Park 和 Wiesel [20, 定理 4] 证明，在固定 $\sigma>0$ 且满足依赖带宽的二次指数矩条件时， $\mathbb{E}[\mathcal{AW}_p(\widehat{\mu}_n*\gamma_\sigma,\mu*\gamma_\sigma)]$ 以速率 $n^{-1/2}$衰减。缩放后的误差是否具有极限分布，在 [20,第 1 节3]中仍是一个开放问题。

我们来探讨这个问题。固定 $\sigma>0$、 $1<p<\infty$和整数 d， $T\ge 1$，并令 $P=\mu*\gamma_\sigma$和 $P_n=\widehat{\mu}_n*\gamma_\sigma$ 在 $\mathbb{R}^{dT}$ 上具有密度 ρ 和 ρ_n。我们研究$\sqrt{n}\,\mathcal{AW}_p(P_n,\ P)$ 的渐近行为，并证明在下文给出的关于 μ 的假设下，

$$\sqrt{n}\,\mathcal{AW}_p(P_n,P)\implies N_{P,p}(G)\qquad \text{as } n\to\infty.$$

在此极限中， G 是由经验密度涨落 $\sqrt{n}(\rho_n-\rho)$ 产生的中心高斯场，而 $N_{P,\ p}=Q_P^{1/p}{}_{,\ p}$ 是定义在适当的零质量有符号密度空间上的适应切向范数。这里，$Q_{P,\ p}(h)$ 是满足线性化连续性方程 $-\sum_{t=1}^{T}\nabla_t\cdot(\rho v_t)=h$、同时遵循因果性的速度场 $v=(v_1,\ \dots,\ v_T)$ 的最小 $L^p(P)$ 动能；因果性意味着每个 v_t 仅是 $x_{1:t}$ 的函数。在 $L^p(P)$–$L^{p'}(P)$ 对偶变分问题中，若 $p'=p/(p-1)$，这一限制会将静态光滑 Wasserstein 极限 [17] 中的测试梯度 $\nabla_t\phi(X)$ 替换为其条件期望 $\mathbb{E}_P[\nabla_t\phi(X)\mid\mathcal{F}_t]$[1, 2]。

由于 $N_{P,\ p}$ 以及 G 的协方差都取决于未知概率律 μ，因此要直接计算极限分布，就必须估计这些量。我们证明，只要对完整轨迹进行重抽样，而不是分别抽取各个时间坐标，普通自助法就能一致地估计 $N_{P,\ p}(G)$ 的概率律；全路径重抽样保留了重抽样路径分布中不同日期之间的依赖关系。在第 3节中，我们结合这些自助法分位数与最优停止的 Lipschitz 稳定性，得到平滑模型 P下停止值的渐近误差界，并将其与 P和 μ之间的确定性平滑偏差分别处理。由于自助法距离 $\mathcal{AW}_p(P_n^*,\ P_n)$ 同时涉及 P的两个随机扰动，证明依赖于一个确定性的双端点展开

$$|\varepsilon|^{-1}\mathcal{AW}_p(P+\varepsilon h,P+\varepsilon j)\longrightarrow N_{P,p}(j-h)\qquad \text{as } \varepsilon\to 0,$$

只要 h 和 j 是 μ的有界零均值倾斜的高斯平滑，就能在我们的假设下成立。在静态最优传输中，沿线性密度插值 ρ_s 的上界通常通过将速度场按 ρ/ρ_s缩放来证明；但在适应性设定下，这可能引入对未来坐标的依赖，从而破坏适应性。相反，我们沿光滑的块三角形常微分方程流输运 $P+\varepsilon h$ ，并通过条件$L^{p'}$-Poincaré 不等式控制到 $P+\varepsilon j$ 的剩余误差。

当 μ 具有紧支集时，对于所有 $1<p<\infty$ 和所有 $\sigma>0$，极限分布、全路径自助法一致性以及一阶矩和 p阶矩的收敛性均成立。在 $p=2$处，最优

速度关于 h呈线性，因此Q_P,2(G)是独立卡方随机变量的加权级数；对于 $p \neq 2$，切能量是非二次的，附录 A.1 中的一个双期例子可以说明这一点。没有紧支集时， $1 < p \leq 2$ 和 $p > 2$ 两种情形表现迥异。对于 $1 < p \leq 2$，当 $d = 1$ 且 μ 相对于 σ满足二次指数矩条件时，极限分布和自助法成立；在任意维度下，当 $\mu = f\mu_0$是强对数凹分布 $\mu_0(da) \propto e^{-V_0}(a)da$的双侧有界密度扰动，且 $\nabla^2 V_0 \succeq \kappa I_{dT}$时， $d \geq 1$ 也成立；对于这类扰动，任意固定的 $\sigma > 0$ 均适用，一阶矩和 p阶矩也收敛。对于 $p > 2$，施加更强的依赖带宽的指数矩条件后，这两类非紧类仍具有相同的极限分布和自助法；然而，仅有对数凹性在带宽较小时并不足够：对于 $d = T = 1$ 和 $\mu = \mathrm{N}(0, s^2)$，只要 $\sigma^2 < (p-2)s^2$， $\sqrt{n}\, W_p(P_n, P)$就会依概率发散至 $+\infty$，因为切能量中的因子 ρ^{1-p} 在尾部增长过快。第 2.1 节和第 2.2 节陈述紧支集和非紧支集情形的定理，第 3 节则在给出标量能量公式和数值实验的同时，推导最优停止界。证明见第 4 节和第 5 节，附录 A 则包含谱公式、更多例子和数值细节。

2 主要结果

2.1 紧支集路径概率律

2.1.1 设定与适应切范数

固定 $1 < p < \infty$、状态维数 $d \geq 1$、时间范围 $T \geq 1$和平滑尺度 $\sigma > 0$，并令 $N = dT$。当 $n \to \infty$时，这些参数保持不变。$(\mathbb{R}^d)$ 中的轨迹^T 记为$x = (x_1, \ldots\ldots, x_T)$，其前缀为 $x_{1:t} = (x_1, \ldots\ldots, x_t)$，其中 $1 \leq t \leq T$和 $x_{1:0} = \varnothing$。我们将 $|\cdot|$ 记作欧几里得范数，并令$|x|_{\ell^p} = (\sum_{t=1}^{T} |x_t|^p)^{1/p}$。

设 μ 是 $(\mathbb{R}^d)^T$ 上的博雷尔概率测度，并令 $A^{(1)}, \ldots\ldots, A^{(n)}$ 是随机路径 $A = (A_1, \ldots\ldots, A_T) \sim \mu$ 的独立副本。记 $\gamma_\sigma = \mathcal{N}(0, \sigma^2 I_N)$，其密度及低维边缘分布也使用相同记号。由经验测度 $\widehat{\mu}_n = n^{-1} \sum_{i=1}^{n} \delta_{A^{(i)}}$，我们构造平滑分布

$$P = \mu * \gamma_\sigma, \qquad P_n = \widehat{\mu}_n * \gamma_\sigma, \qquad \rho = \frac{dP}{dx}, \qquad \rho_n = \frac{dP_n}{dx}.$$

记 $e_a(x) = \gamma_\sigma(x - a) = (2\pi\sigma^2)^{-N/2} e^{-|x-a|^2/(2\sigma^2)}$为 $a, x \in \mathbb{R}^N$，则有 $\rho(x) = \mathbb{E}_\mu[e_A(x)]$以及 $\rho_n(x) = n^{-1} \sum_{i=1}^{n} e_{A^{(i)}}(x)$。对于同时涉及 A 和 X的期望，我们使用耦合

$$X = A + \sigma Z \sim P, \qquad Z \sim \mathcal{N}(0, I_N), \quad Z \perp A,$$

其他以下标标示的潜在分布也遵循相同约定。

令 $\mathcal{F}_t = \sigma(X_1, \ldots, X_t)$ 为典范滤子，其中 $\mathcal{F}_0$是平凡的。对于定义在 $(\mathbb{R}^d)^T$ 上的两个路径律 P和Q，令 $\Pi(P, Q)$ 表示它们的耦合集合，其坐标过程为 (X, Y)。集合$\Pi_{\mathrm{bc}}(P, Q)$ 中的双因果耦合 由所有满足以下条件的 $\pi \in \Pi(P, Q)$ 构成：对每个 $1 \leq t \leq T$，

$$\begin{aligned} &\mathcal{L}_\pi(Y_{1:t} \mid X_{1:T}) = \mathcal{L}_\pi(Y_{1:t} \mid X_{1:t}), \\ &\mathcal{L}_\pi(X_{1:t} \mid Y_{1:T}) = \mathcal{L}_\pi(X_{1:t} \mid Y_{1:t}), \end{aligned} \qquad \pi\text{-a.s.}$$

适应的 p-Wasserstein 距离定义为

$$\mathcal{AW}_p(P, Q)^p = \inf_{\pi \in \Pi_{\mathrm{bc}}(P,Q)} \int \sum_{t=1}^{T} |x_t - y_t|^p \, d\pi(x, y). \tag{2.1}$$

对于未平滑律 ν， $\widetilde{\nu}$，我们还定义 $\mathcal{AW}_p^{(\sigma)}(\nu, \widetilde{\nu}) := \mathcal{AW}_p(\nu * \gamma_\sigma, \widetilde{\nu} * \gamma_\sigma)$。当 $T = 1$时，$\mathcal{AW}_p$退化为 p-Wasserstein 距离 W_p 在 $\mathbb{R}^d$上的通常形式。

假设 2.1（紧支集）。 μ 的支集包含于 $[-R,R]^N$ 中，其中$R<\infty$。

对于 $1\le t\le T$，令 $P_{1:t}$ 表示 $X_{1:t}$ 在 P 下的边缘分布，其密度为 $\rho_{1:t}(x_{1:t}) = \int \rho(x_{1:t},x_{t+1:T})\,dx_{t+1:T}$。令 $\rho_{1:0}=1$，并令 $k_t^P(x_{1:t-1},x_t) = \rho_{1:t}(x_{1:t})/\rho_{1:t-1}(x_{1:t-1})$ 表示一步条件密度。为简洁起见，明确时省略 P，并将测度 $k_t^P(z,y)\,dy$ 也记作 $k_t^P(z)$。适应的 L^p 速度场 $v=(v_1,\dots,v_T)$，满足 $v_t = v_t(x_{1:t})$，构成空间

$$\mathcal{A}_{P,p} = \bigoplus_{t=1}^{T} L^p(\rho_{1:t}(x_{1:t})\,dx_{1:t};\mathbb{R}^d), \qquad \|v\|_{P,p}^p = \sum_{t=1}^{T}\int_{\mathbb{R}^N} |v_t(x_{1:t})|^p \rho(x)\,dx.$$

记 ∇_t 和 $\nabla_{t\cdot}$ 为第 t 个分块 $x_t\in\mathbb{R}^d$ 中的梯度和散度。对于 $h\in L^1(\mathbb{R}^N)$，满足 $\int h(x)\,dx=0$，定义

$$\begin{aligned} Q_{P,p}(h) &= \inf_{\substack{v\in\mathcal{A}_{P,p}\\ -\sum_{t=1}^{T}\nabla_t\cdot(\rho v_t)=h}} \sum_{t=1}^{T}\int_{\mathbb{R}^N}|v_t(x_{1:t})|^p\rho(x)\,dx, \\ N_{P,p}(h) &= Q_{P,p}(h)^{1/p}. \end{aligned} \tag{2.2}$$

式 (2.2) 中的约束按分布意义理解：$\int \phi h = \sum_{t=1}^{T}\int \nabla_t\phi\cdot v_t\rho$ 对每个 $\phi\in C_c^\infty(\mathbb{R}^N)$ 都成立。令 $\inf\varnothing=+\infty$。对于 $Q_{P,p}(h)<\infty$，可行集是闭仿射集，因此 P 的自反性和严格凸性$\mathcal{A}_{P,p}$ 保证了唯一极小化元。在有限能量扰动上，$N_{P,p}$ 是由商空间所诱导的范数：将 $\mathcal{A}_{P,p}$ 除以闭散度核 $\{w\in\mathcal{A}_{P,p}:\sum_{t=1}^{T}\nabla_t\cdot(\rho w_t)=0\}$。

泛函（2.2）也有对偶表示。令 $p'=p/(p-1)$，并将 $\mathcal{A}_{P,p}$ 的对偶空间认作 $\mathcal{A}_{P,p'} = \bigoplus_{t=1}^{T} L^{p'}(P_{1:t};\mathbb{R}^d)$。由于 $v_t(X_{1:t})$ 是 $\mathcal{F}_t$-可测的，每个 $\phi\in C_c^\infty(\mathbb{R}^N)$ 和 $v\in\mathcal{A}_{P,p}$ 都满足

$$\sum_{t=1}^{T}\int_{\mathbb{R}^N}\nabla_t\phi\cdot v_t\rho = \sum_{t=1}^{T}\mathbb{E}_P\big[\mathbb{E}_P[\nabla_t\phi(X)\mid\mathcal{F}_t]\cdot v_t(X_{1:t})\big].$$

因此，散度核是这些配对的共同零空间。由自反性可知，它在 $\mathcal{A}_{P,p'}$ 中的湮没子是条件梯度 $\{\mathbb{E}_P[\nabla_t\phi\mid\mathcal{F}_t]\}_{t=1}^{T},\phi\in C_c^\infty(\mathbb{R}^N)$ 的范数闭包。设

$$\mathcal{E}_{P,p'}(\phi) = \sum_{t=1}^{T}\mathbb{E}_P\,|\mathbb{E}_P[\nabla_t\phi\mid\mathcal{F}_t]|^{p'}, \qquad \phi\in C_c^\infty(\mathbb{R}^N),$$

商对偶性和杨氏不等式给出

$$\begin{aligned} N_{P,p}(h) &= \sup_{\substack{\phi\in C_c^\infty(\mathbb{R}^N)\\ \mathcal{E}_{P,p'}(\phi)\le 1}} \int \phi h, \\ Q_{P,p}(h) &= \sup_{\phi\in C_c^\infty(\mathbb{R}^N)} \left\{ p\int\phi h - (p-1)\mathcal{E}_{P,p'}(\phi)\right\}. \end{aligned}$$

2.1.2 极限分布与自助法

在下述极限定理的假设下，第4.3 节指定了一个可分希尔伯特空间 H，其上的带符号密度使 $h\mapsto\int h$ 连续，$\mathbb{E}_\mu\|e_A\|_{\mathcal{H}}^2<\infty$，且 $N_{P,\ p}$ 在$\mathcal{H}_0=\{h\in\mathcal{H}:\ \int h=0\}$上连续。令 G 为取值于$\mathcal{H}_0$的中心化高斯元素，其协方差算子为

$$C_G u = \mathbb{E}_\mu\big[\langle e_A-\rho,u\rangle_{\mathcal{H}}(e_A-\rho)\big], \qquad u\in\mathcal{H}.$$

的概率律仅取决于协方差核 $\mathrm{Cov}^{\mu}(e_A(x)\ e_A(y))$ 和（2.2），而不取决于$\mathcal{H}$的选择。

在给定 $\mathcal{D}_n = \sigma(A^{(1)},\ \ldots,\ A^{(n)})$ 的条件下，令 $I_1,\ \ldots,\ I_n$ 为在 $\{1,\ \ldots,\ n\}$ 上服从均匀分布的独立同分布随机变量，并构造全路径自助法分布

$$\widehat{\mu}_n^* = n^{-1}\sum_{j=1}^{n}\delta_{A^{(I_j)}}, \qquad P_n^* = \widehat{\mu}_n^* * \gamma_\sigma, \qquad \rho_n^* = \frac{dP_n^*}{dx}.$$

记 $\mathcal{L}^*$、$\mathbb{P}^*$、$\mathbb{E}^*$，Var^* 分别表示给定 $\mathcal{D}_n$ 后的条件分布、概率、期望和方差。在度量空间 $(E,\ d_E)$ 上，令

$$d_{\mathrm{BL},E}(\nu,\widetilde{\nu}) = \sup_{\|f\|_\infty\le 1,\,\mathrm{Lip}_{d_E}(f)\le 1}\left|\int f\,d\nu - \int f\,d\widetilde{\nu}\right|,$$

其中 $d_{\mathrm{BL}} := d_{\mathrm{BL},\ \mathbb{R}}$；在 $\mathbb{R}$ 上，柯尔莫哥洛夫距离为 $\sup_{x\in\mathbb{R}}|\nu((-\infty,\ x]) - \widetilde{\nu}((-\infty,\ x])|$。

定理 2.2 (紧支撑极限与自助法)。 在假设 *2.1* 下，我们有$Q_{P,\ p}(G) < \infty$ 几乎必然成立，并且当 $n\to\infty$ 时，

$$\begin{aligned}\sqrt{n}\,\mathcal{AW}_p(P_n,P) &\Longrightarrow N_{P,p}(G),\\ n^{p/2}\mathcal{AW}_p(P_n,P)^p &\Longrightarrow Q_{P,p}(G).\end{aligned} \tag{2.3}$$

全路径自助法满足

$$d_{\mathrm{BL}}\big(\mathcal{L}^*(\sqrt{n}\,\mathcal{AW}_p(P_n^*,P_n)),\ \mathcal{L}(N_{P,p}(G))\big)\longrightarrow 0 \quad \textit{in probability.} \tag{2.4}$$

如果 μ 不是点质量，$N_{P,\ p}(G)$ 具有连续分布函数，并且 (2.4) 在柯尔莫哥洛夫距离下也成立；否则，$\mathcal{AW}_p(P_n,\ P)$ 和 $\mathcal{AW}_p(P_n^*,\ P_n)$ 均恒等于零。此外，对于 $r\in\{1,\ p\}$，

$$\mathbb{E}\big[(\sqrt{n}\,\mathcal{AW}_p(P_n,P))^r\big]\longrightarrow\mathbb{E}\big[N_{P,p}(G)^r\big]. \tag{2.5}$$

注 2.3（二次极限分布）。 对于 $p=2$，(2.2) 中的最小能量映射 $\mathcal{V}_P\colon\mathcal{H}_0\to\mathcal{A}_{P,\ 2}$有界且为线性映射。设 $\lambda_1\ge\lambda_2\ge\cdots\ge 0$ 是迹类算子 $\mathcal{V}_P C_G\mathcal{V}_P^*$ 在 $\mathcal{A}_{P,\ 2}$ 上的特征值（按重数计数；若为有限秩，则用零补齐）。对于独立同分布的 $Z_j\sim\mathcal{N}(0,\ 1)$，

$$n\,\mathcal{AW}_2(P_n,P)^2\Longrightarrow Q_{P,2}(G)\overset{d}{=}\sum_{j=1}^{\infty}\lambda_j Z_j^2,$$

其中 $\sum_{j=1}^{\infty}\lambda_j<\infty$；证明和迹恒等式见附录 A.1 。对于 $p\neq 2$，切能量$Q_{P,\ p}$ 不是二次的。

2.2 无界支撑的概率律

当 μ 在 $(\mathbb{R}^d)^T$ 上具有无界支集时，我们考虑两类：具有二次指数矩的标量路径分布（$d=1$），以及相对于强对数凹参考分布 μ_0 的密度在正无穷之间有界的向量值分布（$d,\ T\ge 1$）。

假设 2.4 （标量指数矩）。 对于 $d=1$，我们假设存在某个 $\beta>0$，使得 $\mathbb{E}_\mu[e^{\beta|A|^2}]<\infty$。

假设 2.5 （强对数凹参考分布）。 存在 $\kappa>0$ 以及势函数$V_0\in C^2(\mathbb{R}^N;\ \mathbb{R})$，使得对于每个 $a\in\mathbb{R}^N$，$\nabla^2 V_0(a)\succeq\kappa I_N$。定义参考概率分布为 $\mu_0(da) = \mathcal{Z}_0^{-1}e^{-V_0(a)}\,da$，其中 $\mathcal{Z}_0 = \int_{\mathbb{R}^N}e^{-V_0(a)}\,da\in(0,\infty)$。我们假设 μ 相对于 μ_0 绝对连续，且其密度 $f=d\mu/d\mu_0$ 满足

$$0<c_f\le f\le C_f<\infty \qquad \mu_0\text{-almost everywhere}, \tag{2.6}$$

对于某些常数 $c_f,\ C_f$。

对于整数 $m \geq 1$，定义 θ_m，并且当 $p > 2$ 时，定义 $\Theta_{p,\,m}$，其定义为

$$\begin{aligned} \theta_m &= \frac{2m - 1 + \sqrt{4m^2 + 12m - 7}}{4}, \qquad m \leq \theta_m < m + \tfrac{1}{2}, \\ \Theta_{p,m} &= \frac{p(m+1) - 2 + \sqrt{[p(m+1) - 2]^2 + 8pm}}{4} < \frac{p(m+1)}{2}. \end{aligned} \tag{2.7}$$

这些阈值源自第 5.1 节和5.2节中的加权估计。

定理 *2.6*（具有指数矩的标量律）。假设假设 *2.4* 在$\beta\sigma^2 > \theta_T$（若 $1 < p \leq 2$）或 $\beta\sigma^2 > \Theta_{p、\,T}$（若 $p > 2$）的条件下成立。则（*2.3*）–（*2.4*）成立（在$p = 2$处采用注 *2.3*）。如果 μ 不是点质量，则 $N_{P、\,p}(G)$ 具有连续分布函数，并且（*2.4*）在柯尔莫哥洛夫距离下也成立；否则，AW_p（P_n、P）和 AW_p（P_n^*、P_n）均恒等于零。

根据（2.7）中的上界，定理 2.6 的假设在以下条件下成立

$$\mathbb{E}_\mu \exp\left(\frac{T + \frac{1}{2}}{\sigma^2}|A|^2\right) < \infty \quad (1 < p \leq 2), \qquad \mathbb{E}_\mu \exp\left(\frac{p(T+1)}{2\sigma^2}|A|^2\right) < \infty \quad (p > 2).$$

定理 2.7（具有强对数凹参考分布的概率律）。假设 *2.5* 成立。对于 $1 < p \leq 2$，令 $\sigma > 0$ 任意。对于 $p > 2$，固定 $\sigma > 0$，并假设对于某个满足 $\beta\sigma^2 > \Theta_p$的 $\beta > 0$，$\mathbb{E}_\mu[e^{\beta|A|^2}] < \infty$，$_{dT}$。于是，当 $n \to \infty$ 时，

$$\sqrt{n}\,\mathcal{AW}_p(P_n, P) \Longrightarrow N_{P,p}(G).$$

极限具有连续分布函数，且条件全路径自助法律$\mathcal{L}^*(\sqrt{n}\,\mathcal{AW}_p(P_n^*, P_n))$ 依概率收敛到 $\mathcal{L}(N_{P,p}(G))$，且收敛所依据的是柯尔莫哥洛夫距离。对于 $p = 2$，极限由注释 *2.3* 描述。此外，若 $1 < p \leq 2$，则

$$\mathbb{E}\Big[\big(\sqrt{n}\,\mathcal{AW}_p(P_n, P)\big)^r\Big] \longrightarrow \mathbb{E}[N_{P,p}(G)^r], \qquad r \in \{1, p\}.$$

下面的高斯例子表明，对于 $p > 2$，仅有强对数凹性并不能保证在每个带宽下的根-n 紧性。

命题 2.8 (高斯阻碍)。设 $d = T = 1$, $p > 2$，且 $\mu = N(0, s^2)$ 满足 $s > 0$。对于满足 $0 < \sigma^2 < (p-2)s^2$ 的任意固定带宽，我们有

$$\sqrt{n}\,W_p(\widehat{\mu}_n * \gamma_\sigma, \mu * \gamma_\sigma) \longrightarrow +\infty \quad \textit{in probability.} \tag{2.8}$$

特别地，序列 $\{\sqrt{n}\,W_p(P_n, P)\}_{n \geq 1}$ 不是紧的。

命题 2.8 在第4.3节中证明。

2.3 双端点局部展开

我们的证明基于有界潜在倾斜下$\mathcal{AW}_p$ 的局部展开：

$$\mathcal{D} = \{h_g = (g\mu) * \gamma_\sigma : g \in L^\infty(\mu),\ \mathbb{E}_\mu g = 0\}, \qquad P_\varepsilon^h = (\rho + \varepsilon h)\,dx. \tag{2.9}$$

由于 $|h_g| \leq \|g\|_\infty \rho$，$P_\varepsilon^{h_g}$是 $|\varepsilon|\|g\|_\infty \leq 1$的概率律。

命题 2.9（双端点局部展开）。假设以下条件之一成立：

(i) 假设*2.1*。*(ii)* 假设 *2.4*，当 $\beta\sigma^2 > \theta_T$ *if* $1 < p \leq 2$，或 $\beta\sigma^2 > \Theta_{p,\,T}$ *if* $p > 2$。

(iii) 假设 2.5，并附加要求：当 $p>2$ 时，$\mathbb{E}_{\mu}[e^{\beta|A|^2}]<\infty$ *for some* $\beta>0$ 满足 $\beta\sigma^2>\Theta_{p,\,dT}$。

此处 θ_m 以及 $\Theta_{p,\,m}$ 均在(2.7)中定义。于是

$$\lim_{\varepsilon\to 0}\frac{\mathcal{AW}_p(P_\varepsilon^h,P_\varepsilon^j)}{|\varepsilon|}=N_{P,p}(j-h),\qquad h,j\in\mathcal{D}. \tag{2.10}$$

双端点形式处理自助法中的随机对 $(P_n^*,\ P_n)$；第4.3节通过希尔伯特范数比较，将2.10从 $\mathcal{D}$ 推广到经验扰动。

3 应用与数值实验

下面的标量公式使极限能量可以直接计算。随后，我们将自助法应用于最优停止问题，并在两期金融模型中检验所得校准。谱计算及更多示例汇集于附录A。

3.1 可计算的切向能量

在维数为 $d=1$ 的情形下，只要存在一个属于 $\mathrm{A}_{P,p}$ 的解，连续性方程就能唯一确定适应速度场，因此 (2.2) 中的下确界便不再需要。为写出 $T=2$ 对应的能量，令 $\mathbb{R}^2$ 上的坐标为 (x,z)，并设 $\rho_1(x)=\rho_{1:1}(x)$，$k_x(z)=\rho(x,z)/\rho_1(x)$，以及 $K(x,z)=\int_{-\infty}^{z}k_x(s)\,ds$。给定一个零质量扰动 h，定义

$$h_1(x)=\int_{\mathbb{R}}h(x,z)\,dz,\qquad H_h(x)=\int_{-\infty}^{x}h_1(s)\,ds,\qquad A_h(x,z)=\int_{-\infty}^{z}h(x,s)\,ds.$$

令 $v=(v_1,\ v_2)\in\mathcal{A}_{P,\,p}$ 满足 $-\partial_x(\rho v_1)-\partial_z(\rho v_2)=h$。由于两个通量均可积，在 $z\in\mathbb{R}$ 上积分可得 $-\partial_x(\rho_1 v_1)=h_1$，从而 $v_1(x)=-H_h(x)/\rho_1(x)$。将 $\rho v_1=-H_h k_x$ 代回方程，并在 z 上对 $(-\infty,\ z]$ 进行积分，得到 $\rho v_2=-A_h+Kh_1+H_h\partial_x K$。因此，切能量为

$$Q_{P,p}(h)=\int_{\mathbb{R}}\frac{|H_h(x)|^p}{\rho_1(x)^{p-1}}\,dx+\iint_{\mathbb{R}^2}\frac{|A_h-Kh_1-H_h\partial_x K|^p}{\rho^{p-1}}\,dx\,dz. \tag{3.1}$$

这里的项 $-H_h\partial_x K=\rho_1 v_1\partial_x K$ 源自条件累积分布函数 $K(x,\ \cdot)$在 x中的变分。对 $T\ge 3$ 个标量周期重复这一消元过程，得到递归速度（5.3）。一旦 $d\ge 2$，适应散度便不再唯一确定 v：转到 dT 标量坐标仍会在（2.2）中产生一个可行场，但通常不是块能量的最小化场。在多参数条件分布下，局部二次能量最小化速度场的相容性将在 [27]中研究。

3.2 最优停止与自助法校准

设 Q 是定义在 $(\mathbb{R}^d)^T$ 上的路径律，并配备典范滤子 $\mathcal{F}_t$。设 $g_t:\mathbb{R}^d\to\mathbb{R}(1\le t\le T)$ 是可测的贴现收益，它们具有共同的 Lipschitz 常数 $L<\infty$，且在 Q 下期望有限。记 $\mathcal{T}(Q)$ 为取值于 $\{1,\dots,T\}$ 的 $\mathcal{F}_t$-停止时刻集合，我们考虑最优停止值

$$\mathsf{V}(Q)=\sup_{\tau\in\mathcal{T}(Q)}\mathbb{E}_Q[g_\tau(X_\tau)].$$

对于第一坐标 $x^{(1)}$ 上行权价为 $K>0$、每期利率为 $r\ge 0$ 的百慕大看跌期权，其贴现收益

$$g_t(x)=e^{-rt}(K-e^{x^{(1)}})_+,\qquad u_+=\max\{u,0\},$$

是 L-Lipschitz 的，常数为 $L=K$。

命题 3.1 (停止值的稳定性)。 对于定义在 $(\mathbb{R}^d)^T$ 上、具有有限 p 阶矩且满足 $1 \le p < \infty$ 的任意路径律 P、Q，

$$|\mathsf{V}(P) - \mathsf{V}(Q)| \le L\, \mathcal{AW}_p(P, Q),$$

按照(2.1)中的成本约定。

证明。 设 $\pi \in \Pi_{\mathrm{bc}}(P, Q)$ 是 (X, Y) 的一个双因果耦合，并令 $\tau \in \mathcal{T}(P)$。由于 π 从 Y 到 X 是因果的，给定 Y 时 τ 的条件分布定义了 Y 的一个随机停止规则。其在 Q 下的期望收益是普通停止时刻收益的混合，因此不超过 $\mathsf{V}(Q)$。根据霍尔德不等式，

$$\mathbb{E}_P[g_\tau(X_\tau)] - \mathsf{V}(Q) \le \mathbb{E}_\pi\big[g_\tau(X_\tau) - g_\tau(Y_\tau)\big] \le L\Big(\mathbb{E}_\pi \sum_{t=1}^{T} |X_t - Y_t|^p\Big)^{1/p}.$$

对 $\tau \in \mathcal{T}(P)$ 取上确界，再对 $\pi \in \Pi_{\mathrm{bc}}(P, Q)$ 取下确界，得到 $\mathsf{V}(P) - \mathsf{V}(Q) \le L\mathcal{AW}_p(P, Q)$。交换 P 和 Q 可得反向界。 □

固定显著性水平 $\alpha \in (0,1)$，并定义全路径自助法统计量及其条件 $(1-\alpha)$ 分位数如下

$$T_n^* = \sqrt{n}\, \mathcal{AW}_p(P_n^*, P_n), \qquad q_{n,1-\alpha}^* = \inf\{z \in \mathbb{R} : \mathbb{P}^*(T_n^* \le z) \ge 1 - \alpha\}.$$

令 $F_\infty(z) = \mathbb{P}(N_{P,p}(G) \le z)$。当对每个 $\varepsilon > 0$ 都有 $F_\infty(q_{1-\alpha} - \varepsilon) < 1 - \alpha < F_\infty(q_{1-\alpha} + \varepsilon)$ 时，其 $(1-\alpha)$ 分位数 $q_{1-\alpha} = \inf\{z : F_\infty(z) \ge 1 - \alpha\}$ 唯一。

推论 3.2 （停止值的自助法误差半径）。 假设定理 *2.2*、定理 *2.6* 或定理 *2.7* 的条件成立，并假设$N_{P,\,p}(G)$是非退化的，且具有唯一的（$1-\alpha$）分位数。那么

$$\liminf_{n\to\infty} \mathbb{P}\left(|\mathsf{V}(P_n) - \mathsf{V}(P)| \le \frac{L\, q_{n,1-\alpha}^*}{\sqrt{n}}\right) \ge 1 - \alpha.$$

证明。 F_∞的柯尔莫哥洛夫收敛于$\mathcal{L}^*(T_n^*)$，加上 $q_{1-\alpha}$的唯一性，推出$q^*_{n,\,1-\alpha} \to q_{1-\alpha}$ 依概率成立，因此

$$\mathbb{P}(\sqrt{n}\, \mathcal{AW}_p(P_n, P) \le q_{n,1-\alpha}^*) \longrightarrow 1 - \alpha.$$

根据命题3.1，事件 $\sqrt{n}\,\mathcal{AW}_p(P_n, \bar{P}) \le q^*_{n,\,1-\alpha}$意味着 $|\mathsf{V}(P_n) - \mathsf{V}(P)| \le n^{-1/2} L\, q^*_{n,\,1-\alpha}$，由此得出结论。 □

在计算中，$q^*_{n,\,1-\alpha}$ 由$\widehat{q}^*_{n,\,1-\alpha}$ 这一经验分位数替代；该分位数基于 B_n 次条件独立的自助抽样得到。给定数据后，经验自助法累积分布函数的逐点方差至多为 $1/(4B_n)$。当 $B_n \to \infty$ 时，有限网格论证结合 F_∞ 的连续性和（2.4），可得累积分布函数在联合概率意义下一致收敛。由 $q_{1-\alpha}$ 的唯一性，进而推出$\widehat{q}^*_{n,\,1-\alpha} \to q_{1-\alpha}$ 依概率成立，因此推论3.2 在以$\widehat{q}^*_{n,\,1-\alpha}$ 替代 $q^*_{n,\,1-\alpha}$ 时仍成立（有关实验中采用的重抽样约定，请参见附录A.3 ）。

固定带宽目标。 记 $P_\sigma = \mu * \gamma_\sigma$和 $P_{n,\,\sigma} = \widehat{\mu}_n * \gamma_\sigma$。推论 3.2 在指定的平滑尺度 $\sigma > 0$下校准停止价值。如果观测状态为 $Y = A + \sigma Z$，决策依据 Y的历史，且行权收益为 $g_t(Y_t)$，那么 $\mathsf{V}(P_\sigma)$ 就是该观测模型的直接价值；如果 σ 被选作建模容差，则 $\mathsf{V}(P_\sigma)$ 可作为正则化基准，其抽样误差由自助法校准。

若要对未平滑值 $\mathsf{V}(\mu)$进行推断，令 $b_\sigma = |\mathsf{V}(P_\sigma) - \mathsf{V}(\mu)|$。任意上界$b_\sigma \ge b_\sigma$ 都会直接加到统计半径上：

$$|\mathsf{V}(P_{n,\sigma}) - \mathsf{V}(\mu)| \le |\mathsf{V}(P_{n,\sigma}) - \mathsf{V}(P_\sigma)| + \bar{b}_\sigma.$$

在对 μ采用一致 W_1-Lipschitz 单步条件核的条件下， [19,定理 5.1] 中的带宽界与 $p = 1$ 处的命题 3.1 一起给出 $b_\sigma \le L\mathrm{AW}_1(P_\sigma, \mu) = O(\sigma)$。另一方面，对于 $p > 2$ ，存在直接阻碍带宽 $\sigma_n \downarrow 0$缩小的因素：在命题 2.8 的高斯情形下，固定任意满足 $0 < \sigma_0^2 < (p-2)s^2$的 σ_0。一旦 $\sigma_n \le \sigma_0$，在共同高斯卷积下 W_p 的收缩性给出

$$W_p(\widehat{\mu}_n * \gamma_{\sigma_0}, \mu * \gamma_{\sigma_0}) \le W_p(\widehat{\mu}_n * \gamma_{\sigma_n}, \mu * \gamma_{\sigma_n}),$$

因此，沿任意序列 $\sigma_n \downarrow 0$，$\sqrt{n}\, W_p(\widehat{\mu}_n * \gamma_{\sigma_n},\ \mu * \gamma_{\sigma_n}) \to +\infty$ 依概率成立。

3.3 半合成百慕大期权实验

我们在一个两时点考克斯–罗斯–鲁宾斯坦二项模型 [13]中展示极限分布和全路径自助法。以 12 月 31 日的芝加哥期权交易所波动率指数收盘值 14．95 作为年化波动率 $\nu = 0.1495$，我们设定 $S_0 = 100$、行权价 $K = 105$、半年步长 $\Delta = 1/2$，年利率为 3%。利用

$$a = \nu\sqrt{\Delta}, \qquad U = e^{a}, \qquad D = e^{-a}, \qquad \pi_+ = \frac{e^{0.03\Delta} - D}{U - D},$$

每个时期以风险中性概率 π_+向上变动，使贴现股票价格在潜在分布 μ下构成鞅。在对数相对坐标中，四条路径及其概率为

$$\begin{aligned} (a_i)_{i=1}^{4} &= \big((-a,-2a),(-a,0),(a,0),(a,2a)\big), \\ (w_i)_{i=1}^{4} &= \big((1-\pi_+)^2,(1-\pi_+)\pi_+,\pi_+(1-\pi_+),\pi_+^2\big). \end{aligned} \tag{3.2}$$

我们设定 $\sigma = 0.65a$ 并取贴现百慕大看跌期权收益 $g_t(x) = e^{-0.03t\Delta}(105 - 100e^{x})_+ (t = 1,\ 2)$，它们具有相同的 Lipschitz 常数 $L = 105$。由于 μ 具有有限支撑，定理 2.2适用于 $p = 2$ 和 $p = 4$。

对于（3.2）中四条路径上的任意概率权重，高斯平滑产生的第二期条件分布关于第一坐标随机递增。因此，共同均匀 Rosenblatt 耦合是最优的，从而我们可以通过一维分位数求积直接计算$\mathcal{AW}_p$ ；附录 A.3 给出了证明、求积细节和停止价值积分。

对于每个 $n \in \{50, 200, 800, 3200, 12800\}$，我们从（3.2）模拟生成 600 个独立数据集，并为每个数据集抽取 249 个全路径自助重抽样，对每次重抽样重新计算 $\mathrm{AW}_p(P_n^*, P_n)$。图1比较了 $\sqrt{n}\,\mathrm{AW}_p(P_n, P)$的抽样分布与 $\sqrt{n}\,\mathrm{AW}_p(P_n^*, P_n)$的合并自助分布，并使用根据（3.1）计算的极限 0.95 分位数 $q_{0.95}$进行归一化。图2比较了围绕 $\mathsf{V}(P)$ 和 $\mathsf{V}(\mu)$ 得到的停止价值误差与自助半径。

根据（3.1），渐近0.95分位数为 $q_{0.95} \approx 0.5666$ 当 $p = 2$ 时，以及 $q_{0.95} \approx 0.8947$ 当$p = 4$ 时。即使在 $n = 50$处，有限样本上尾低于其渐近极限（在 $p = 4$时尤为明显），自助法仍能紧密跟踪抽样分布。随着 n 增大，两种分布的标准化0.95分位数都收敛至1，且在全部五种样本量下，名义覆盖率为95% 的距离界的经验覆盖率介于94.2% 和96.3% 之间（表1）。

反向归纳得到 $\mathsf{V}(\mu) = 7.94165$，而在平滑模型下进行数值积分得到 $\mathsf{V}(P) = 8.96675$，从而产生 1.02511 的固定平滑偏差。在 $n = 12800$ 时，

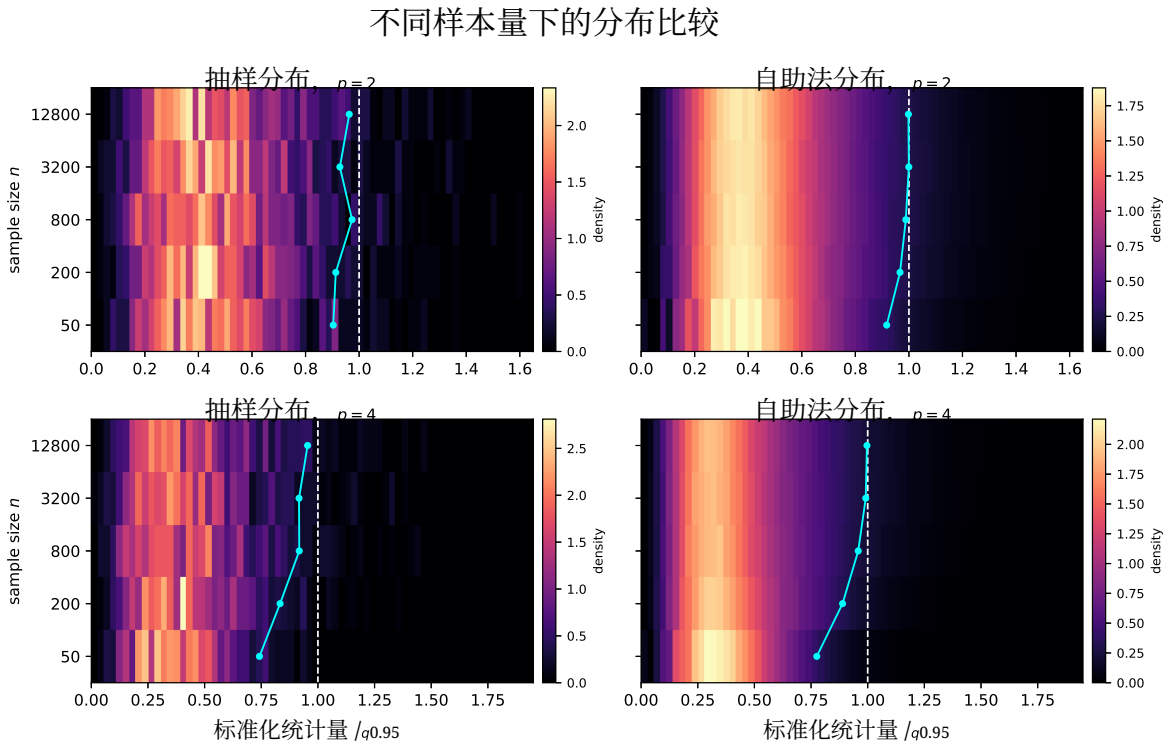

图1：模型（3.2）中 $n \in \{50、200、800、3200、12800\}$的抽样分布和合并的全路径自助法分布。横轴按极限0.95分位数 $q_{0.95}$（来自3.1）（竖直虚线）进行归一化；青色折线描绘了每个0.95分位数处的经验 n。

自助法值半径中位数分别为 0.5235（$p=2$）和 0.8261（$p=4$），而 $|\mathrm{V}(P_n)-\mathrm{V}(P)|$的经验 0.95 分位数为 0.1332。在所有样本量下，自助法区间在全部 600 次试验中都覆盖 $\mathrm{V}(P)$；加入平滑偏差 $|\mathrm{V}(P)-\mathrm{V}(\mu)|$后，得到围绕 $\mathrm{V}(\mu)$ 的相应误差界，如图2右面板所示。

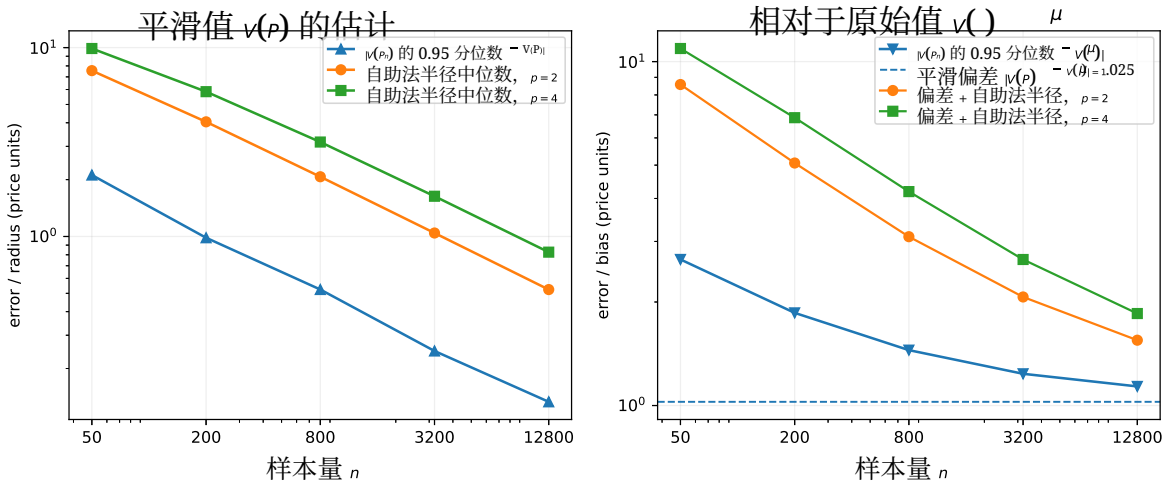

图2：百慕大期权估值误差与自助半径（以价格单位计）。左：平滑值 $\mathrm{V}(P)$ 附近的误差，显示 $|\mathrm{V}(P_n) - \mathrm{V}(P)|$ 的经验 0.95 分位数（蓝色），以及 $L\widehat{q}^{*}_{n,0.95}/\sqrt{n}$ 的自助法半径中位数，分别对应 $p=2$（橙色）和 $p=4$（绿色）。右：相对于 $\mathrm{V}(\mu)$ 的同一实验，显示 $|\mathrm{V}(P_n) - \mathrm{V}(\mu)|$ 的经验 0.95 分位数（蓝色）、固定平滑偏差 $|\mathrm{V}(P) - \mathrm{V}(\mu)|$（虚线），以及将该偏差加到相应自助半径后的结果（橙色和绿色）。

4 主要结果的证明

在本节中，我们将 $q = p'$ 和 $L^q(f) = L^q(f(x)\,dx)$ 记作一个正密度 f。映射或向量场在 $(\mathbb{R}^d)^T$ 上是 分块三角的，当且仅当其第 t 个分量仅依赖于 $x_{1:t}$。对于光滑映射 F， DF 表示其雅可比矩阵，$\|DF\|$ 表示其欧几里得算子范数。

4.1 传输比较与适应散度

对于具有正密度 ρ 的路径律 P，我们的比较界使用均匀条件L^q-Poincaré 不等式

$$\|\phi - \int \phi k_t(z,\cdot)\|_{L^q(k_t(z,\cdot))} \le C_q \|\nabla\phi\|_{L^q(k_t(z,\cdot))}, \qquad z \in \mathbb{R}^{d(t-1)}. \tag{4.1}$$

引理 4.1 (普通传输比较)。 设 f, g 是 $\mathbb{R}^d$ 上具有有限p阶矩的概率密度。若 f 满足常数为 C_q 的 L^q-*Poincaré* 不等式，则

$$W_p(f\,dx, g\,dx) \le p C_q \left\| \frac{g-f}{f} \right\|_{L^p(f\,dx)}.$$

证明。 对于 $\|(g-f)/f\|_{L^p(f)} < \infty$，由霍尔德不等式和 $L^q(f)$-Poincaré 不等式可得，对任意 $\phi \in C_c^\infty(\mathbb{R}^d)$,

$$\left| \int \phi(g-f) \right| \le C_q \left\| \frac{g-f}{f} \right\|_{L^p(f)} \|\nabla\phi\|_{L^q(f)}.$$

Hahn–Banach延拓和 $L^q(f;\mathbb{R}^d)$ 对偶性给出满足 $-\mathrm{div}(fv) = g - f$且具有$\|v\|_{L^p(f)} \le C_q\|(g-f)/f\|_{L^p(f)}$ 的 $v \in L^p(f;\mathbb{R}^d)$。沿着非线性插值

$$f_s = (1-\vartheta(s))f + \vartheta(s)g, \qquad \vartheta(s) = 1-(1-s)^p, \quad s \in [0,1],$$

t通量 $J_s = \vartheta'(s) f v$ 满足 $\partial_s f_s +$ 散度 $J_s = 0$，其中 $f_s \ge (1-\vartheta(s))f$ 对于 $s < 1$。由于 e

$$\int_0^1 \frac{\vartheta'(s)^p}{(1-\vartheta(s))^{p-1}}\,ds \int |v|^p f = p^p \int |v|^p f,$$

连续性方程界 W_p [3, 第 8]得出该结果。 □

引理 4.2 (单侧适应比较)。 设 P, Q 为具有密度 ρ, η、前缀密度 $\rho_{1:t}$, $\eta_{1:t}$且具有有限 p 阶矩的路径律。假设 P 满足 (4.1),

$$W_p(k_t(z,\cdot),k_t(z',\cdot)) \le L|z-z'|_{\ell^p},$$

且 $\eta \ge c\rho$ 对某个 $c>0$成立。那么

$$\mathcal{AW}_p(P,Q) \le C(p,T,C_q,L,c)\left\|\frac{\eta-\rho}{\rho}\right\|_{L^p(P)}.$$

证明。 令 $u=\eta/\rho$并令 $u_t=\mathbb{E}_P[u \mid X_{1:t}]\,(u_0=1)$。在任意共享历史 $z=x_{1:t-1}$处，我们有 $k_t^Q(z,\cdot)/k_t^P(z,\cdot)=u_t/u_{t-1}$。由于 $u_{t-1}\ge c$，在历史 z处应用引理 4.1 并使用条件詹森不等式可得

$$\begin{aligned}\int \eta_{1:t-1}(z)W_p(k_t^P(z,\cdot),k_t^Q(z,\cdot))^p\,dz &\le (pC_q)^p\mathbb{E}_P[u_{t-1}^{1-p}|u_t-u_{t-1}|^p]\\ &\le 2^p(pC_q)^p c^{1-p}\|u-1\|^p_{L^p(P)}.\end{aligned}$$

将耦合 $(X,\ Y)$ 通过 $t=1,\ \ldots\ldots,\ T$ 进行归纳构造，所用的是 $k_t^P(X_{1:t-1},\cdot)$ 与 $k_t^Q(Y_{1:t-1},\cdot)$ 的可测最优（或 ε-最优）单步耦合。将 $k_t^P(Y_{1:t-1},\cdot)$ 插入 W_p 三角不等式中，可得 $S_t=\sum_{i\le t}\mathbb{E}|X_i-Y_i|^p$ 满足

$$S_t \le (1+2^{p-1}L^p)S_{t-1}+2^{2p-1}(pC_q)^p c^{1-p}\|u-1\|^p_{L^p(P)}.$$

从 $t=1$ 到 T 迭代即可证明该引理。 □

当 $|\nabla_i \log k_t| \le B$对所有 $i<t$ 成立时，我们有 $|\log(k_t(z',y)/k_t(z,y))| \le B\sum_{i<t}|z_i-z_i'|$，因此将引理 4.1 应用于 $m\to\infty$ 线段的等分 $[z,z']$ 可得

$$W_p(k_t(z,\cdot),k_t(z',\cdot)) \le pC_qB\sum_{i<t}|z_i-z_i'| \le pC_qBT^{1/q}|z-z'|_{\ell^p}. \tag{4.2}$$

在$\mathcal{A}^{P,p}$上，将分布意义下的适应散度算子定义为

$$D_\rho v = -\rho^{-1}\sum_{t=1}^{T}\nabla_t\cdot(\rho v_t),$$

其图定义域为 $\{v\in\mathcal{A}^{P,p}\colon D_\rho v\in L^p(P)\}$并赋予图范数$\|v\|_{P,p}+\|D_\rho v\|_{L^p(P)}$。在此记号下，式（2.2）中的约束写作 $D_\rho v=h/\rho$。

引理 4.3 (适应散度的可解性)。 假设 ρ 及其前缀密度 $\rho_{1:t}$ 均为正且光滑，(4.1) 成立，并且历史得分 $\nabla_i\log k_t\,(i<t)$ 有界。那么，对于每个满足 $\mathbb{E}_P s=0$ 的 $s\in L^p(P)$，都存在唯一的最小范数速度场 $v\in\mathcal{A}^{P,p}$使其满足$D_\rho v=s$，并且

$$\|v\|_{P,p}\le C\|s\|_{L^p(P)},\qquad N_{P,p}(h)\le C\|h/\rho\|_{L^p(P)}\quad\Big(\int h=0\Big).$$

证明。 令 $s_t=\mathbb{E}_P[s\mid X_{1:t}]$，其中 $s_0=0$。给定 $v_1,\ \ldots\ldots,\ v_{t-1}$，考虑第 t 个前缀方程

$$-\rho_{1:t}^{-1}\nabla_t\cdot(\rho_{1:t}v_t)=s_t-s_{t-1}+\sum_{i<t}v_i\cdot\nabla_i\log k_t=:b_t.$$

对 $\int k_t(z,\ y)\,dy=1$ 在 $z_i\,(i<t)$ 求导可得，$\mathbb{E}_P[\nabla_i\log k_t\mid X_{1:t-1}]=0$，因此 $\mathbb{E}_P[b_t\mid X_{1:t-1}]=0$。由（4.1）和霍尔德不等式，每个 $\phi\in C_c^\infty(\mathbb{R}^{dt})$ 都满足

$$\left|\int b_t\phi\,dP_{1:t}\right|\le C_q\|b_t\|_{L^p(P_{1:t})}\|\nabla_t\phi\|_{L^q(P_{1:t})}.$$

Hahn–Banach延拓和 $L^q(P_{1:t};\mathbb{R}^d)$ 对偶性给出满足前缀方程的 $v_t \in L^p(P_{1:t};\mathbb{R}^d)$，其中 $\|v_t\|_{L^p(P_{1:t})} \le C_q\|b_t\|_{L^p(P_{1:t})}$。对 $t=1,\dots,T$ 进行归纳法可得到一个解 $v \in \mathcal{A}_{P,\,p}$ 它满足范数界；随后，$D_\rho^{-1}(\{s\})$ 中的闭仿射空间因 $\mathcal{A}_{P,\,p}$ 的自反性和严格凸性而存在唯一的最小范数元素。 □

引理 4.4 (适应性部分散度核心)。 若除引理 *4.3* 的假设外，未来评分 $\nabla_t \log(\rho/\rho_{1:t})$ 在 $1 \le t \le T$ 上有界，则每个满足 $D_\rho v = s \in L^p(P)$ 的 $v \in \mathcal{A}_{P,p}$ 都可以用光滑场 $v^m = (v_1^m,\dots,v_T^m)$ 逼近，使得 $v_t^m \in C_c^\infty(\mathbb{R}^{dt};\mathbb{R}^d)$ 满足

$$\|v^m - v\|_{P,p} \longrightarrow 0, \qquad \|D_\rho v^m - s\|_{L^p(P)} \longrightarrow 0.$$

证明。 由于 $\rho v_i \in L^1(\mathbb{R}^N)$，将 $D_\rho v = s$ 与 $\psi(x_{1:t})\chi_M(x_{t+1:T})$ 配对，并取一个满足 $|\nabla\chi_M| \le C/M$ 的光滑未来截断函数 χ_M，再令 $M \to \infty$，可得部分散度 $D_t v_t := -\rho_{1:t}^{-1}\nabla_t \cdot (\rho_{1:t} v_t)$ 满足

$$D_t v_t = s_t - s_{t-1} + \sum_{i<t} v_i \cdot \nabla_i \log k_t \in L^p(P_{1:t}), \qquad s_t = \mathbb{E}_P[s \mid X_{1:t}], \quad s_0 = 0.$$

将 v_t 乘以一个满足 $|\nabla_t\chi_M| \le C/M$ 的光滑前缀截断函数 $\chi_M(x_{1:t})$，得到 $\chi_M v_t \to v_t$，并且当 $M \to \infty$ 时，$D_t(\chi_M v_t) = \chi_M D_t v_t - \nabla_t\chi_M \cdot v_t \to D_t v_t$ 在 $L^p(P_{1:t})$ 中成立。固定 M 后，$F = \rho_{1:t}\chi_M v_t$ 和 $\nabla_t \cdot F$ 均属于具有紧支集的 $L^p(\mathbb{R}^{dt})$；用定义在 $\mathbb{R}^{dt}$ 上的标准磨 mollifier ζ_δ 对 F 进行卷积，再除以 $\rho_{1:t}$，得到场 $v^{M,\,\delta}_{\,t} = \rho_{1:t}^{-1}(F * \zeta_\delta) \in C_c^\infty(\mathbb{R}^{dt};\mathbb{R}^d)$，它们在 $L^p(P_{1:t})$ 中及在 D_t-范数下收敛到 $\chi_M v_t$，当 $\delta \downarrow 0$ 时成立。由于分解中的未来评分 $\nabla_t \log(\rho/\rho_{1:t})$ 有界

$$D_\rho v = \sum_{t=1}^{T} \left(D_t v_t - v_t \cdot \nabla_t \log \frac{\rho}{\rho_{1:t}} \right),$$

在 $t = 1,\dots,T$ 上取 (M,δ) 的对角序列，即可得到所述结论。 □

注 4.5（高斯乘子）。 引理 4.3和4.4 的假设适用于以下形式的任意密度：$\widetilde{\rho} = m\gamma$ 在 $\mathbb{R}^N$上，其中 γ 为标准高斯密度（γ_k 在 $\mathbb{R}^k$上），$0 < a \le m \le b < \infty$，$\|\nabla m\|_\infty < \infty$，且 m 及其前缀平均

$$m_t(z_{1:t}) = \int m(z_{1:t}, z_{t+1:T})\gamma_{d(T-t)}(z_{t+1:T})\,dz_{t+1:T} \qquad (m_0 = 1,\ m_T = m)$$

均为光滑函数。事实上，条件分布 $k_t(z_{1:t}) = \gamma_d(z_t)m_t(z_{1:t})/m_{t-1}(z_{1:t-1})$ 继承了4.1）中的高斯 L^q-Poincaré 不等式，其中 $C_q = 2(b/a)^{1/q}(\pi/2(\mathbb{E}|Z|^q)^{1/q}(Z \sim \mathcal{N}(0,\ 1))$，而历史得分函数和未来评分

$$\nabla_i \log k_t = \nabla_i \log m_t - \nabla_i \log m_{t-1}, \qquad \nabla_t \log(\widetilde{\rho}/\widetilde{\rho}_{1:t}) = \nabla_t \log m - \nabla_t \log m_t$$

其绝对值均不超过 $2\|\nabla m\|_\infty/a$。

4.2 双端点切向展开

现在证明命题2.9，首先给出一般下界，然后在紧致情形、对数凹参考分布情形以及标量指数矩情形下构造相匹配的双因果上界。

引理 4.6 （局部下界）。 设 P 具有密度 $\rho > 0$，且 h、k 的积分为零，其中h/ρ、$k/\rho \in L^\infty$ 和 $N_{P,\,p}(k-h) < \infty$。对于 $P_\varepsilon^h = (\rho + \varepsilon h)\,dx$ 和 $P_\varepsilon^k = (\rho + \varepsilon k)\,dx$，

$$\liminf_{\varepsilon \to 0} \frac{\mathcal{AW}_p(P_\varepsilon^h, P_\varepsilon^k)}{|\varepsilon|} \ge N_{P,p}(k-h).$$

证明。 当下极限有限时，取一个达到该下极限的子序列，选取 (X,Y) 的双因果耦合$\pi_\varepsilon \in \Pi_{\mathrm{bc}}(P_\varepsilon^h, P_\varepsilon^k)$，使其代价至多比 $\mathcal{AW}_p(P_\varepsilon^h, P_\varepsilon^k)^p$大 $o(|\varepsilon|^p)$，并令

$$V_t^\varepsilon(x) = \mathbb{E}_{\pi_\varepsilon}[(Y_t - X_t)/\varepsilon \mid X = x].$$

π_ε 的因果性确保 $V_t^\varepsilon(x)$仅依赖于 $x_{1:t}$，而$(\rho + \varepsilon h)/\rho = 1 + O(|\varepsilon|)$和詹森不等式使 V^ε 在$\mathcal{A}_{P,p}$中保持有界。取一个子序列，我们得到$\mathcal{A}_{P,p}$中的弱极限$V^\varepsilon \rightharpoonup V$。对于任意 $\phi \in C_c^\infty(\mathbb{R}^N)$，由泰勒定理可得

$$|\phi(y) - \phi(x) - \nabla\phi(x)\cdot(y-x)| \le C_\phi \min\{|y-x|^2, |y-x|\}.$$

由于 $\int |y-x|^p \, d\pi_\varepsilon = O(|\varepsilon|^p)$，在 $1 < p < 2$ 时使用$\min\{r^2, r\} \le r^p$，在 $p \ge 2$ 时使用霍尔德不等式，可知

$$\frac{1}{|\varepsilon|}\int \min\{|y-x|^2, |y-x|\}\, d\pi_\varepsilon \le C \begin{cases} |\varepsilon|^{p-1}, & 1 < p < 2, \\ |\varepsilon|, & p \ge 2, \end{cases} \longrightarrow 0.$$

将 $\mathbb{E}_{\pi_\varepsilon}[\phi(Y) - \phi(X)]$ 除以 ε，并令 $\varepsilon \to 0$，因此得到

$$\int \phi(k-h) = \lim_{\varepsilon\to 0} \int \sum_{t=1}^T \nabla_t\phi \cdot V_t^\varepsilon(\rho + \varepsilon h) = \int \sum_{t=1}^T \nabla_t\phi \cdot V_t \rho.$$

因此 $-\sum_{t=1}^T \nabla_t \cdot (\rho V_t) = k - h$，而 $\|\cdot\|_{P,p}$ 的弱下半连续性给出下界。

□

紧支撑概率律。 我们首先验证具有紧支撑混合律的高斯混合的条件庞加莱不等式、得分函数和比较估计。

引理4.7 (一致混合庞加莱不等式)。 若 ν 是由位置参数 a 构成混合的 $N(a, \sigma^2 I_d)$，且其支集的直径至多为 D，那么对于 $1 \le q < \infty$ 和 $Z \sim N(0, 1)$，

$$\|\phi - \mathbb{E}_\nu\phi\|_{L^q(\nu)} \le \left(\frac{\pi}{2}(\mathbb{E}|Z|^q)^{1/q}\sigma + De^{D^2/(8q\sigma^2)}\right)\|\nabla\phi\|_{L^q(\nu)}. \tag{4.3}$$

证明。 令 B 独立地服从混合概率律，且 $Z_d \sim \mathcal{N}(0, I_d)$，并令 $m(a) = \mathbb{E}\phi(a + \sigma Z_d)$。对 $[0, \pi/2]$ 应用Pisier旋转插值，得到

$$\|\phi(a + \sigma Z_d) - m(a)\|_{L^q} \le \frac{\pi}{2}(\mathbb{E}|Z|^q)^{1/q}\sigma\|\nabla\phi(a + \sigma Z_d)\|_{L^q}.$$

令 B' 为 B的一个独立副本。对于任意两个位置 a, a'，满足 $|a - a'| \le D$ 且任意$s \in [0, 1]$，对高斯密度配方可得

$$\gamma_\sigma(x - [(1-s)a + sa']) \le e^{D^2/(8\sigma^2)}[(1-s)\gamma_\sigma(x-a) + s\gamma_\sigma(x-a')].$$

因此，沿$(1-s)B + sB'$ 对 ∇m 积分可得

$$\mathbb{E}|m(B) - m(B')|^q \le D^q e^{D^2/(8\sigma^2)} \int |\nabla\phi|^q \, d\nu,$$

以及在 $L^q(\nu)$ 中的三角不等式给出 (4.3)。 □

当 μ 的支集包含于紧集 $K \subset \mathbb{R}^{dT}$ 且 $P = \mu * \gamma_\sigma$时，引理4.7 给出一个条件庞加莱常数 C_q，使得在(4.1)中该常数关于 $x_{1:t-1}$一致。对积分号下的高斯核求导，还可得到条件得分恒等式

$$\begin{aligned} \nabla_i \log k_t(x_{1:t}) &= \sigma^{-2}(\mathbb{E}[A_i \mid X_{1:t}] - \mathbb{E}[A_i \mid X_{1:t-1}]), \quad i < t, \\ \nabla_t \log \frac{\rho(x)}{\rho_{1:t}(x_{1:t})} &= \sigma^{-2}(\mathbb{E}[A_t \mid X_{1:T}] - \mathbb{E}[A_t \mid X_{1:t}]), \end{aligned} \tag{4.4}$$

二者均以 L^∞ by $\sigma^{-2}\mathrm{diam}(K)$ 为界。

命题 4.8 (固定紧致潜在类上的比较)。 假设 $P = \nu * \gamma_\sigma$ 和$Q = \widetilde{\nu} * \gamma_\sigma$ 具有密度 ρ 和 η，其中 ν 和 $\widetilde{\nu}$ 的支集均包含于同一个紧集$K \subset \mathbb{R}^{dT}$。那么

$$\mathcal{AW}_p(P, Q) \le C(p, d, T, K, \sigma) \left\| \frac{\eta - \rho}{\rho} \right\|_{L^p(P)}.$$

证明。按照引理 4.2 设定 $u = \eta/\rho$ 和 $u_t = \mathbb{E}_P[u \mid X_{1:t}]$，并在日期 $t-1$ 处将历史空间划分为 $\{u_{t-1} \ge 1/2\}$ 和 $\{u_{t-1} < 1/2\}$。在 $\{u_{t-1} \ge 1/2\}$ 上，引理 4.1 将条件 W_p^p 积分界定为 $(pC_q)^p 2^{p-1} \mathbb{E}_P|u_t - u_{t-1}|^p \le C\|u-1\|_{L^p(P)}^p$。在 $\{u_{t-1} < 1/2\}$ 上，$k_t^P(z,\cdot)$ 和 $k_t^Q(z,\cdot)$ 都是 $\mathsf{N}(a, \sigma^2 I_d)$ 的混合分布，其中 a 位于第 t 个投影 K_t 中，因此 $W_p(k_t^P(z,\cdot)\ k_t^Q(z,\cdot)) \le D_t := \mathsf{diam}(K_t)$。由于 $u_{t-1} < 1/2$ 蕴含 $|u_{t-1} - 1| > 1/2$，我们得到

$$\begin{aligned} D_t^p \mathbb{E}_P[u_{t-1} \mathbf{1}_{\{u_{t-1} < 1/2\}}] &\le D_t^p P_{1:t-1}(u_{t-1} < 1/2) \\ &\le 2^p D_t^p \mathbb{E}_P|u_{t-1} - 1|^p \le 2^p D_t^p \|u-1\|_{L^p(P)}^p. \end{aligned}$$

将这两个界代入引理 4.2 的耦合递推，并结合（4.2）和（4.4），即可得到所述不等式。

□

命题 4.9 （紧支集条件下的双端点切向展开）。 在假设 *2.1* 下，对所有 h， $k \in \mathcal{D}$ 如 (2.9)所示，

$$\lim_{\varepsilon \to 0} \frac{\mathcal{AW}_p(P_\varepsilon^h, P_\varepsilon^k)}{|\varepsilon|} = N_{P,p}(k - h).$$

证明。设 $v \in \mathsf{A}_{P,p}$ 为引理 4.3 中 $D_\rho v = (k-h)/\rho$ 的最小范数解，并设$w = (w_1, \dots, w_T)$ 满足 $w_t \in C_c^\infty(\mathbb{R}^{dt}; \mathbb{R}^d)$，为引理 4.4 中的一个光滑近似函数。常微分方程流 $\partial_\varepsilon F_\varepsilon(x) = w(F_\varepsilon(x))$ 满足 $F_0(x) = x$，是一个全局块三角微分同胚，其逆映射 $F_{-\varepsilon}$ 为块三角形式，因此 $(\mathrm{Id}\ F_\varepsilon)_\# P_\varepsilon^h$ 是 P_ε^h 与 Q_ε^w 之间的双因果耦合：$= (F_\varepsilon)_\# P_\varepsilon^h$。由于 $(F_\varepsilon - \mathrm{Id})/\varepsilon \to w$ 一致成立，由控制收敛可得

$$\frac{1}{|\varepsilon|^p} \int \sum_{t=1}^{T} |F_{\varepsilon,t}(x) - x_t|^p (\rho + \varepsilon h)(x)\, dx \longrightarrow \|w\|_{P,p}^p.$$

令 $a = h/\rho$ 和 $b_w = D_\rho w$。由于 $\nabla_t \log \rho = \nabla_t \log \rho_{1:t} + \nabla_t \log(\rho/\rho_{1:t})$ 在 w_t 的紧支集上有界（由 4.4）， b_w 在 $\mathbb{R}^N$上有界且连续，并且沿 F_ε 应用变量替换公式可得

$$\frac{dQ_\varepsilon^w}{dP}(x) = [1 + \varepsilon a(F_{-\varepsilon}(x))] \exp\left(\int_0^\varepsilon b_w(F_{-s}(x))\, ds \right).$$

因此， $dQ_\varepsilon^w/dP = 1 + O(|\varepsilon|)$在 $\mathbb{R}^N$ 上一致成立，且

$$\frac{dQ_\varepsilon^w/dP - 1}{\varepsilon} \longrightarrow a + b_w \quad \text{in } L^p(P).$$

当 $|\varepsilon|$ 足够小时，有 $dQ_\varepsilon^w/dP_\varepsilon^k \ge 1/2$。利用引理 4.2，以 P_ε^k 作为参考概率律，并令 $\varepsilon \to 0$ 趋于极限，可得

$$\limsup_{\varepsilon \to 0} \frac{\mathcal{AW}_p(Q_\varepsilon^w, P_\varepsilon^k)}{|\varepsilon|} \le C \left\| D_\rho w - \frac{k-h}{\rho} \right\|_{L^p(P)},$$

与 C 无关， w。将其与 F_ε的传输代价结合起来，令 $w \to v$ 在引理4.4的图范数中，并与引理 4.6 匹配，即可得到该展开式。 □

推论 4.10（紧致潜在类上的一个共同希尔伯特控制）。 固定一个整数 $s > N/2$，并选择 $0 < a < 2\sigma^{-2} < \alpha$ $pa > p-1\ \alpha$ $_{a\,s} = h\ e^{a|\cdot|^2}h \in$（）满足（）。在希尔伯特空 $H^s\ \mathbb{R}^N$ $\|h\|_{\mathcal{H}_{a\,s}} = \|e^{a|\cdot|^2}h\|_{H^s\ \mathbb{R}^N}$ $_{a\,s\,0} = h \in\ _{a\,s}\ \int h = 0$ $P = \nu * \gamma_\sigma$ 间 H, {: （）}上，其范数为 .（），且 H, , {H, : }，所有满足 $\mathrm{supp}(\nu)$、$\mathrm{supp}(\widetilde{\nu}) \subset K$ 的和 $Q = \widetilde{\nu} * \gamma_\sigma$ 均为紧致的

$$\mathcal{AW}_p(P,Q) \le C\|\rho - \eta\|_{\mathcal{H}_{a,s}}, \qquad N_{P,p}(h) \le C\|h\|_{\mathcal{H}_{a,s}} \quad \Big(\int h = 0\Big).$$

证明。 由于 ν 的支集包含在紧集 K中，因此 $\rho(x) \ge c_\alpha e^{-\alpha|x|^2}$；而索伯列夫嵌入（$s > N/2$）给出 $|h(x)| \le C_s\|h\|_{\mathcal{H}_{a,\ s}} e^{-a|x|^2}$。因此

$$\int |h|^p \rho^{1-p} \le C\|h\|^p_{\mathcal{H}_{a,s}} \int e^{-[pa-(p-1)\alpha]|x|^2}\,dx \le C'\|h\|^p_{\mathcal{H}_{a,s}},$$

这两个界由命题4.8 和引理4.3得出。 □

强对数凹参考律。 当 $\mu = f\mu_0$ 满足假设 2.5时，我们构造一个全局双Lipschitz三角微分同胚 F，将标准高斯分布 γ 从 $\mathbb{R}^N$ 映射到$P_0 = \mu_0 * \gamma_\sigma$，将适应性连续性方程拉回到高斯图坐标，并通过 F共轭常微分方程恢复流。

记 $\rho_0 = dP_0/dx$，按字典序列出 N 标量坐标（t, i），并令$\bar\rho_{0,\ j}(x_{1:j})$ 为 P_0 ($\bar\rho_{0,\ 0} = 1$) 的第 j 个标量前缀密度。在 $A \sim \mu_0$和 $X = A + \sigma Z$的条件下，给定 $X_{1:j} = x \in \mathbb{R}^j$ 时 A 的条件分布具有势函数

$$W_{j,x}(a) = V_0(a) + \frac{|x - a_{1:j}|^2}{2\sigma^2} + \mathrm{const}, \qquad \nabla^2 W_{j,x} \succeq K_j := \kappa I_N + \sigma^{-2}\Pi_j,$$

其中 Π_j 投影到前 j 个标量坐标上。根据 Brascamp–Lieb方差不等式[11, 式（1.3）], Cov($A \mid X_{1:j} = x$) $\preceq K_j^{-1}$，因此 Cov($A_{1:j} \mid X_{1:j} = x$) $\preceq (\kappa + \sigma^{-2})^{-1} I_j$。 对 $\bar\rho_{0,j}(x)$ $= \mathbb{E}_{\mu_0}[\gamma_\sigma(x - A_{1:j})$] 求两次导数，得到 ∇^2(–log $\bar\rho_{0,j}$)(x) $= \sigma^{-2} I_j - \sigma^{-4}$ Cov($A_{1:j} \mid X_{1:j} = x$)，因此

$$mI_j \preceq \nabla^2(-\log \bar\rho_{0,j}) \preceq MI_j, \qquad m = \frac{\kappa}{1 + \kappa\sigma^2}, \quad M = \sigma^{-2}. \tag{4.5}$$

对于 $1 \le j \le N$和标量历史 $a \in \mathbb{R}^{j-1}$，条件密度 $k_a(z) = \bar\rho_{0,j}(a, z)/\bar\rho_{0,j-1}(a)$ 以及势函数 $U_a = -$ log k_a 满足

$$m \le U_a'' \le M, \qquad |U_a'(z) - U_b'(z)| \le M|a - b|.$$

在 $r = M|a - b|/m$下，不等式 $U_{\mathit{ℓ}b}(z) \ge U_{\mathit{ℓ}a}(z - r)$ 意味着 $k_b(z)/k_a(z - r)$ 关于 z单调不增，因此条件分位数 $q_a(u)$ $= \inf\{z : \int_{-\infty}^{z} k_a(y)\,dy \ge u\}$ 满足

$$|q_a(u) - q_b(u)| \le B|a - b|, \qquad B = M/m, \quad 0 < u < 1.$$

此外，将 Caffarelli 收缩定理 [10, 14] 应用于 $T_a(z) = q_a(\Phi(z))$ 和 T_a^{-1}可得$M^{-1/2} \le \partial_z T_a(z) \le m^{-1/2}$。定义 $F_1(z_1)$ $= T_\varnothing(z_1)$ 和 $F_j(z_{1:j}) = T_{F_{1:j-1}(z)}(z_j)$ 后，可得到一个光滑三角微分同胚 $F\colon \mathbb{R}^N \to \mathbb{R}^N$，满足 $F_\#\gamma = P_0$，其逆为 $(F^{-1})^{T^{-1}_{x_{1:j-1}}}\ x_j$ $\ell = m^{-1/2}$ ( ) ( )。记作 ，分

位数界给出

$$|F_j(z) - F_j(z')| \le B|F_{1:j-1}(z) - F_{1:j-1}(z')| + \ell|z_j - z_j'|,$$

由此可得 Lip(F)$^2 \le 2m^{-1}(1 + 2B^2)^{N-1}$、Lip($F^{-1}$)$^2 \le M\{N + (N-1)B^2\}$，从而得到 sup $_z \|DF(z)\| < \infty$ 和 sup$_x \|D(F^{-1})(x)\| < \infty$。

对于 $\mu = f\mu_0$，我们有 $\rho(x) = \rho_0(x)r(x)$，其中 $r(x) = \mathbb{E}_{\mu_0}[f(A) \mid A + \sigma Z = x] \in [c_f, C_f]$。对于任意 $g \in L^\infty(\mu_0)$ 和单位向量 $e \in \mathbb{R}^N$，上述方差界给出

$$|\partial_e \mathbb{E}_{\mu_0}[g(A) \mid X = x]| = \sigma^{-2} |\operatorname{Cov}_{\mu_0}(g(A), e \cdot A \mid X = x)| \le \frac{\|g\|_\infty}{\sigma^2 \sqrt{\kappa + \sigma^{-2}}}.$$

在坐标图 F 中，概率律 P 及其倾斜 $P_\varepsilon^h (h = (g\mu) * \gamma_\sigma \in \mathcal{D})$ 的形式为

$$P = F_\# \widetilde{P}, \qquad \widetilde{\rho}(z) = \gamma(z)\widetilde{r}(z), \qquad \widetilde{r} = r \circ F,$$

以及 $P_\varepsilon^h = F_\#(\widetilde{\rho}_\varepsilon^h \, dz)$，其中 $\widetilde{\rho}_\varepsilon^h = \gamma(\widetilde{r} + \varepsilon \widetilde{r}_h)$ 和 $\widetilde{r}_h(z) = \mathbb{E}_{\mu_0}[f(A)g(A) \mid X = F(z)]$。$\widetilde{r}$ 和 $\widetilde{r}_h$ 均有界，且其梯度也有界；它们的分块前缀平均同样光滑——对于 $\widetilde{r}$，第 t 个分块前缀平均等于 $(\rho_{1:t}/\rho_{1:t}) \circ F_{1:t}$，倾斜密度也满足同一恒等关系。因此，注4.5 对小 $|\varepsilon|$ 下的 $\widetilde{\rho}_\varepsilon^h$ 一致成立。

给定分块适配场 $v \in \mathcal{A}_{P,p}$，定义拉回速度 $u(z) = DF(z)^{-1} v(F(z))$ 和密度方向 $\widetilde{h}(z) = h(F(z))|\det DF(z)|$。由于 DF 和 DF^{-1} 是下分块三角的，u 是分块适配的，并且

$$h = -\operatorname{div}(\rho v) \quad \Longleftrightarrow \quad \widetilde{h} = -\operatorname{div}(\widetilde{\rho} u),$$

而物理动能变为

$$\|v\|_{P,p}^p = \int \sum_{t=1}^{T} |(DF(z)u(z))_t|^p \, \widetilde{\rho}(z) \, dz. \tag{4.6}$$

由 $\|DF\|$ 和 $\|DF^{-1}\|$的一致有界性，p次方根定义了 4.6中的一个等价的严格凸范数，该范数定义在 $\widetilde{P}$上。$_p$应用注记 4.5 以及引理 4.3–4.4 于 $\widetilde{\rho} = \gamma\widetilde{r}$，结合 $\int |\widetilde{h}/\widetilde{\rho}|^p \widetilde{\rho} = \int |h|^p \rho^{1-p}$，可得

$$N_{P,p}(h) \le C_{p,P} \|h/\rho\|_{L^p(P)}, \tag{4.7}$$

以及（4.6）的唯一极小化元 u 和光滑分块前缀紧支近似 $u^m \to u^{\sim}$ $^{\sim}$，它们属于 $\mathcal{A}_{\widetilde{P}.}$ $_p$并满足 $D_{\widetilde{\rho}} u^m \to h/\widetilde{\rho}$于 $L^p(P)$ 中。

现在令 h，$j \in \mathcal{D}$，令 u 为 $j - h$的拉回最小速度，令 Φ_ε^m为 u^m的常微分方程流，并令 $\Psi_\varepsilon^m = F \circ \Phi_\varepsilon^m \circ F^{-1}$。由于 F 和 Φ_ε^m 均为分块三角映射，且其逆映射为块三角逆映射，$(\mathrm{Id}\ \Psi_\varepsilon^m)_\# P_\varepsilon^h$ 是双因果的，由控制收敛可得

$$|\varepsilon|^{-p} \int \sum_{t=1}^{T} |\Psi_{\varepsilon,t}^m(x) - x_t|^p \, dP_\varepsilon^h(x) \longrightarrow \int \sum_{t=1}^{T} |(DF(z)u^m(z))_t|^p \, d\widetilde{P}(z).$$

在坐标图坐标中，输运密度 $\widetilde{q}_\varepsilon^m = d((\Phi_\varepsilon^m)_\# \widetilde{P}_\varepsilon^h)/dz$ 满足

$$\frac{\widetilde{q}_\varepsilon^m - \widetilde{\rho}}{\varepsilon} \longrightarrow \widetilde{h} + \widetilde{\rho} D_{\widetilde{\rho}} u^m \quad \text{in } L^p(\widetilde{\rho}^{1-p} \, dz).$$

$^{\sim}$应用引理 4.2，参考密度为 $\widetilde{\rho} + \varepsilon j$，并将所得的双因果耦合通过 F向前推送（其至多将 $\mathcal{A}\mathcal{W}_p$ 乘以 $L_p = T^{|1/p - 1/2|} \operatorname{Lip}(F)$ ）可得

$$\limsup_{\varepsilon \to 0} \frac{\mathcal{A}\mathcal{W}_p((\Psi_\varepsilon^m)_\# P_\varepsilon^h, P_\varepsilon^j)}{|\varepsilon|} \le C L_p \left\| D_{\widetilde{\rho}} u^m - \frac{\widetilde{j} - \widetilde{h}}{\widetilde{\rho}} \right\|_{L^p(\widetilde{P})}.$$

先令 $\varepsilon \to 0$，再令 $m \to \infty$，并结合引理 4.6，即可在假设 2.5下证明 (2.10)。

具有指数矩的标量路径。假设 $d=1$ 且定理 2.6 的条件成立，并令 $r=\max(p,2)$。根据第 5.1 节中的加权能量估计（当$1<p\le 2$、 $r=2$时为 (5.4)）以及第 5.2 节中的加权能量估计（当 $p>2$、 $r=p$时为 (5.8)），(5.3) 中的递归标量速度 v_h 满足 $\|v_h\|_{P,p}\le T^{(1/p-1/2)_+}\|v_h\|_{P,r}<\infty$。在一维情形下， v_h 是满足 $A_{P,p}$ 并解 $-\sum_{t=1}^{T}\partial_t(\rho v_t)=h$ 的唯一场：所有通量 ρv_t 都属于 $L^1(\mathbb{R}^T)$，因此，对未来坐标积分并对 $1\le j\le T$ 使用归纳法可知，任意两个解之差的通量都与 x_j无关，而根据 $L^1(\mathbb{R})$ 可积性，该通量必为零。因此，$Q_{P,p}(h)=\sum_{t=1}^{T}\int|v_{h,t}|^p\rho_{1:t}$。

对于 $\mathbb{R}^T$ 上的正光滑密度 f，令 $R^f\colon(0,1)^T\to\mathbb{R}^T$ 表示其逆 Rosenblatt 映射，该映射定义为 $R_1^f(u)=F_{f\,1}^{-1}(u_1)$，以及 $R_t^f(u)=F_{f\ t|R_1^f:t-1(u)}^{-1}(u_t)$。给定 $h=h_g\in\mathbf{D}$，设 $\rho_\varepsilon=\rho+\varepsilon h$ 和 $R_\varepsilon=R^{\rho_\varepsilon}$，并选择 $\varepsilon_0>0$，使得 $\varepsilon_0\|g\|_\infty\le 1/2$ 以及所有 $\rho_\varepsilon(|\varepsilon|\le\varepsilon_0)$ 都属于同一个矩类 $\mathbf{C}_{M_0}$（定义见第 5.1 节）。对条件分位数方程求导可得

$$w_{\varepsilon,t}(u):=\partial_\varepsilon R_{\varepsilon,t}(u)=v_{\rho_\varepsilon,h,t}(R_\varepsilon(u)_{1:t})\longrightarrow v_{\rho,h,t}(R_0(u)_{1:t})$$

在 $(0,1)^T$ 上逐点成立，其中 $\varepsilon\to 0$。

根据高斯卷积、前缀密度的正性以及递推式（5.3），每个速度 $v_{\rho_\varepsilon,\,h,\,t}(x_{1:t})$ 都关于 $(\varepsilon,x_{1:t})$ 联合连续，因此

$$M_R:=\max_{1\le t\le T}\ \sup_{\substack{|\varepsilon|\le\varepsilon_0\\ |x_{1:t}|\le R}}|v_{\rho_\varepsilon,h,t}(x_{1:t})|<\infty\qquad\text{for every } R<\infty.$$

对于 $M>M_R$，通过前缀 Rosenblatt 映射进行变量替换，并应用（5.4）（当 $r=2$时）或（5.8）（当 $r=p>2$时），得到

$$\begin{aligned}&\int_{(0,1)^T}|w_{\varepsilon,t}(u)|^r\mathbf{1}_{\{|w_{\varepsilon,t}(u)|>M\}}\,du\\&\quad\le\int_{|x_{1:t}|>R}|v_{\rho_\varepsilon,h,t}(x_{1:t})|^r(\rho_\varepsilon)_{1:t}(x_{1:t})\,dx_{1:t}\le Ce^{-\eta R^2}\|h\|_{\mathcal{H}}^r,\end{aligned}\tag{4.8}$$

其中常数关于 $|\varepsilon|\le\varepsilon_0$ 一致。先取 R 充分大，再取 $M>M_R$，可知$\{|w_{\varepsilon,\,t}|^r\}_{|\varepsilon|\le\varepsilon_0}$一致可积，因此由 Vitali 定理可得 $w_{\varepsilon,\,t}\to v_{\rho,\,h,\,t}\circ R_0$ 在 $L^r((0,\,1)^T)$ 中。对 ε 积分，得到

$$\frac{R_\varepsilon-R_0}{\varepsilon}=\int_0^1\partial_\delta R_\delta|_{\delta=s\varepsilon}\,ds\longrightarrow v_{\rho,h}\circ R_0\quad\text{in } L^r((0,1)^T;\ell^r),$$

因此在 $L^p((0,\,1)^T;\ \ell^p)$ 中也成立，因为 $p\le r$。将此展开应用于 h和 k，并使用双因果耦合$(R_\varepsilon^h(U)\ R_\varepsilon^k(U))$ 以及 $U\sim$均匀分布$((0,\,1)^T)$，可得（2.10）中的上界；引理4.6 给出相应的下界。

4.3 统计极限与自助法

我们在四种情形下分别指定辅助希尔伯特空间 $\mathcal{H}$，使得 $\mathbb{E}_\mu\|e_A\|_{\mathcal{H}}^2<\infty$,$N_{P,\,p}$ 在 $\mathcal{H}_0$，上连续，且 $\mathcal{AW}_p$ 受 $\|\cdot\|_{\mathcal{H}}$ 控制：

- 紧支集（假设 2.1， $1<p<\infty$）。固定一个整数 $s>N/2$，取 $0<a<(2\sigma^2)^{-1}<\alpha$，使得 $pa>(p-1)\alpha$，并令 $\mathbf{H}=\mathbf{H}_{a,s}$ 取自推论 4.10。于是 $\sup_{y\in[-R,R]^N}\|e_y\|_{\mathbf{H}_{a,s}}<\infty$，而推论 4.10 表明，对于支集包含于 $[-R,R]^N$ 的混合律，$\mathrm{AW}_p(\rho\,dx,\eta\,dx)\le C\|\eta-\rho\|_{\mathbf{H}}$ 且 $N_{P,p}(h)\le C\|h\|_{\mathbf{H}}$ 一致成立。

- 标量指数矩（假设 2.4， $1<p\le 2$， $\beta\sigma^2>\theta_T$）。取 $\mathcal{H}$ $=$ $_\omega=$ $L^2(\mathbb{R}^T,\ e^{\omega|x|^2}\,dx)$ 5 1 $\mathcal{H}(,\ )$第 节中的

$$\frac{1}{2\sigma^2}+\frac{T-1}{\sigma^2(1+2\beta\sigma^2)}<\omega<\frac{\beta}{1+\beta\sigma^2}.$$

直接积分可得

$$\|e_y\|_{\mathcal{H}_\omega}^2 = [4\pi\sigma^2(1-\omega\sigma^2)]^{-T/2}\exp\left(\frac{\omega|y|^2}{1-\omega\sigma^2}\right),$$

其中 $\omega/(1-\omega\sigma^2) < \beta$，因此 $\mathbb{E}_\mu\|e_A\|_{\mathcal{H}_\omega}^2 < \infty$。对于任意矩类 $\mathcal{C}_{M_0}$，命题 5.1结合霍尔德不等式

$$\mathcal{AW}_p \le T^{1/p-1/2}\mathcal{AW}_2, \qquad N_{P,p} \le T^{1/p-1/2}N_{P,2} \tag{4.9}$$

给出了度量比较以及$N_{P,\,p}$的连续性。

- 对数凹参考分布（假设 2.5，$1 < p \le 2$）。取$\mathcal{H} = \mathcal{H}_P = L^2(\mathbb{R}^N,\ \rho^{-1}\,dx)$，并且$\mathcal{H}_{P,\,0} = \{h \in \mathcal{H}_P : \int h = 0\}$。引理5.4 给出了每个 $\sigma > 0$处的$\mathbb{E}_\mu\|e_A\|^2_{\mathcal{H}_P} < \infty$，而（5.11）、（4.7）和（4.9）给出了度量比较以及$N_{P,\,p}$的连续性。
- 在 $p > 2$ 下的无界支撑（假设 2.4 或 2.5，$\beta\sigma^2 > \Theta_{p,\,N}$）。取 $\mathcal{H} = \mathcal{H}_{a,\,s}$ 其中$(\alpha,\ b,\ a)$ 满足（ 5.5 ）；特征矩和度量比较由（ 5.7 ）及命题 5.2 给出。

在所有四种情形下，中心化核特征 $\xi_i = e_{A^{(i)}} - \rho$ 是 $\mathcal{H}_0$,中独立同分布的平方可积随机元素，因此希尔伯特空间中心极限定理 [21]给出

$$G_n := \sqrt{n}(\rho_n - \rho) = \frac{1}{\sqrt{n}}\sum_{i=1}^n \xi_i \Longrightarrow G \quad \text{in } \mathcal{H}, \tag{4.10}$$

其中 G 是协方差为 $C_G u = \mathbb{E}[\langle \xi_1, u\rangle_{\mathcal{H}}\xi_1]$ 的中心高斯分布。

一致局部展开与弱极限。 定义 $\mathsf{T}:\ L_0^2(\mu) \to \mathcal{H}$，其定义域为 $L_0^2(\mu) = \{g \in L^2(\mu) : \int g\,d\mu = 0\}$，并令 T $g = \mathbb{E}[g(A)(e_A - \rho)]$。由于 $C_G = \mathsf{T}\mathsf{T}^*$，且有界零均值函数在 $L_0^2(\mu)$ 中稠密，G 在 $\mathcal{H}$ 中的拓扑支撑为

$$\mathcal{S} = \overline{\operatorname{Ran}\mathsf{T}}^{\mathcal{H}} = \overline{\mathcal{D}}^{\mathcal{H}},$$

且 $\xi_i \in \mathcal{S}$ 几乎必然成立。在指数矩情形下，度量比较适用于有界矩类 $\mathcal{C}_{M_0}$，令 $M_i = e^{\beta|A^{(i)}|^2}$、$m_\beta = \mathbb{E}M_1 < M_0$ 和 $M_i^K = \min(M_i, K)$。那么 $n^{-1}\sum_{i=1}^n M_i \to m_\beta$ 依概率收敛，且

$$\operatorname{Var}^*\left(\frac{1}{n}\sum_{j=1}^n M_{I_j}^K\right) \le \frac{K^2}{n}, \qquad \mathbb{E}\mathbb{E}^*\left[\frac{1}{n}\sum_{j=1}^n (M_{I_j} - M_{I_j}^K)\right] = \mathbb{E}(M_1 - M_1^K),$$

因此先发送 $n \to \infty$，然后发送 $K \to \infty$，即可得到 $\mathbb{P}(P_n \in \mathcal{C}_{M_0},\ P_n^* \in \mathcal{C}_{M_0}) \to 1$；此外，一旦 $|\varepsilon|$ 足够小，对于任意 $g \in L^\infty(\mu)$，都有 $m_\beta + \varepsilon\mathbb{E}_\mu[g(A)e^{\beta|A|^2}] < M_0$。

对于方向 $a,\ b \in \mathcal{H}_0$，若 $\rho + \varepsilon a$ 和 $\rho + \varepsilon b$ 是可行概率密度（属于紧类、$\mathcal{C}_{M_0}$ 或 $\mathcal{H}_P$），则定义 $F_\varepsilon(a,\ b) = \varepsilon^{-1}\mathcal{AW}_p((\rho + \varepsilon a)\,dx,\ (\rho + \varepsilon b)\,dx)$。我们声称，对于每个紧 $K \subset \mathcal{S}^2$,

$$\sup_{\substack{(a,b)\in K\\ (a,b)\text{ feasible at }\varepsilon}} |F_\varepsilon(a,b) - N_{P,p}(b-a)| \longrightarrow 0 \qquad (\varepsilon \downarrow 0). \tag{4.11}$$

事实上，固定 $\delta > 0$，并利用$\mathcal{D}$ 在$\mathcal{S}$中的密度，选取一个有限的 δ-网$\{(a_\ell, b_\ell)\}_{\ell=1}^M \subset \mathcal{D}^2$，用于 K。当$\varepsilon > 0$ 足够小时，所有 $\rho + \varepsilon a_\ell$ 和 $\rho + \varepsilon b_\ell$ 都是可行的，而上述度量比较给出

$$|F_\varepsilon(a,b) - F_\varepsilon(a_\ell, b_\ell)| \le L_{\mathrm{cmp}}(\|a - a_\ell\|_{\mathcal{H}} + \|b - b_\ell\|_{\mathcal{H}}),$$

且 L_{cmp} 与网及 ε无关：在紧情形和指数矩情形中，所有端点都属于同一个固定紧类或属于 $\mathcal{C}_{M_0}$；而在基准情形（$1 < p \le 2$）中，对网中的有限多个倾斜 $|\varepsilon|$取最大值$_g \, \|g\|_\infty \le 1/2$ 可使倾斜乘子保持在 $[c_f/2, 3C_f/2]$内，因此（5.11）给出一个一致的常数。令 $L_{\tan}$表示，p 在$\mathcal{H}_0$,上的 Lipschitz 常数，三角不等式给出

$$\begin{aligned}&\sup_{\substack{(a,b)\in K\\ \text{feasible at }\varepsilon}} \left|F_\varepsilon(a,b) - N_{P,p}(b-a)\right| \\ &\quad\le \max_{1\le \ell\le M} \left|F_\varepsilon(a_\ell,b_\ell) - N_{P,p}(b_\ell - a_\ell)\right| + (L_{\mathrm{cmp}} + L_{\tan})\delta.\end{aligned}$$

对于一个固定的网，当 $\varepsilon \downarrow 0$ 时，由命题 2.9，最大值趋于零；令 $\delta \downarrow 0$ 即可证明（4.11）。由于 $\{G_n\}$ 由（*4.10*）在S 中是紧的，且 P_n 以趋于 1 的概率是可行的，将（4.11）应用于 $(a,b) = (0, G_n)$，其中 $\varepsilon = n^{-1/2}$，得到

$$\sqrt{n}\,\mathcal{AW}_p(P_n, P) - N_{P,p}(G_n) \longrightarrow 0 \quad \text{in probability},$$

连续映射给出（2.3）。

自助法、连续性与矩收敛。令 $G_n^* = \sqrt{n}(\rho_n^* - \rho_n)$。我们首先检验在 $\overline{\mathcal{H}}$ 中，

$$d_{\mathrm{BL},\mathcal{H}}(\mathcal{L}^*(G_n^*), \mathcal{L}(G)) \longrightarrow 0 \quad \text{in probability}. \tag{4.12}$$

令 Π_m 为有限秩正交投影，在$\mathcal{H}$ 中强收敛递增至恒等算子。对于每个固定的 m，$\mathbb{E}\|\xi_1\|_{\mathcal{H}}^2 < \infty$ 给出 Lindeberg 条件

$$\mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n \|\xi_i\|_{\mathcal{H}}^2 \mathbf{1}_{\{\|\xi_i\|_{\mathcal{H}} > \delta\sqrt{n}\}}\right] \longrightarrow 0,$$

因此有限维自助法中心极限定理 [26, 第 23] 章适用于 $\Pi_m G_n^*$。与此同时，投影尾项满足 $\mathbb{E}^*\|(I-\Pi_m)G_n^*\|_{\mathcal{H}}^2 \le n^{-1}\sum_{i=1}^n \|(I-\Pi_m)\xi_i\|_{\mathcal{H}}^2 \to \mathbb{E}\|(I-\Pi_m)\xi_1\|_{\mathcal{H}}^2 \to 0$当 $n\to\infty$ 时，继而 $m\to\infty$，从而得到（4.12）。

由于$(G_n, G_n + G_n^*)$ 在 $\mathcal{S}^2$ 中是紧的，并且 (P_n, P_n^*) 以趋于 1 的概率是可行的，(4.11) 蕴含

$$R_n := \sqrt{n}\,\mathcal{AW}_p(P_n^*, P_n) - N_{P,p}(G_n^*) \longrightarrow 0$$

在联合概率意义下。由于 $\mathbb{E}[\mathbb{P}^*(|R_n| > \delta)] = \mathbb{P}(|R_n| > \delta) \to 0$，我们有 $\mathbb{P}^*(|R_n| > \delta) \to 0$ 依概率对于每个 $\delta > 0$ 成立。对于每个 $f:\ \mathbb{R}\to\mathbb{R}$，其中 $\|f\|_\infty \le 1$，且 $\mathrm{Lip}(f) \le 1$,

$$\left|\mathbb{E}^* f(\sqrt{n}\,\mathcal{AW}_p(P_n^*, P_n)) - \mathbb{E}^* f(N_{P,p}(G_n^*))\right| \le \delta + 2\mathbb{P}^*(|R_n| > \delta),$$

将这一界与 (4.12) 以及 N_P 的连续性结合起来，p 证明了 (2.4)。

为了将 (2.4) 提升到柯尔莫哥洛夫距离，我们验证当 μ 不是点质量时，$N_{P,p}(G)$ 是无原子的。对于定义域中任意两个不同的 h, k，$Q_{P,p} = N_{P,p}^p$，它们的最小适应速度 v_h和 v_k 不同，因为 $D_\rho v_h \ne D_\rho v_k$。由于 $tv_h + (1-t)v_k$ 对 $th + (1-t)k$ 是可行的，$\|\cdot\|_P^p$ 的严格凸性给出 p

$$Q_{P,p}(th + (1-t)k) \le \|tv_h + (1-t)v_k\|_{P,p}^p < tQ_{P,p}(h) + (1-t)Q_{P,p}(k), \qquad 0 < t < 1.$$

当 μ 不是点质量时，$C_G \ne 0$。选取 C_G 的一个特征值为$\lambda > 0$的单位特征向量 $e \in \mathbf{H}_0$，并令 $h = \sqrt{\lambda}e$，我们将 $G = Zh + G^\perp$ 分解为 $Z \sim \mathrm{N}(0,1)$，其中后者与 $G^\perp$独立。对于每个固定的 $G^\perp$，函数 $z \mapsto Q_{P,p}(G^\perp + zh)$ 在 $\mathbb{R}$上严格凸，因此每个水平集至多包含两个点，从而其高斯测度为零。根据富比尼定理，$N_{P,p}(G)$ 具有连续分布函数，而波利亚定理将 (2.4) 提升为柯尔莫哥洛夫距离下的收敛。

剩下需要验证矩收敛 (2.5)。在假设 2.1 下，推论 4.10 给出 $\sqrt{n}\,\mathrm{AW}_p(P_n,P)\le C\|G_n\|_{\mathcal{H}}$ 逐样本成立。由于特征 ξ_i 在 H 中有界，拉ademacher 矩不等式 [21] 给出 $\sup_n \mathbb{E}\|G_n\|_{\mathcal{H}}^r<\infty$，对所有满足 $\max(p,2)$ 的 $r>$ 均成立，从而保证一致可积性以及 (2.5)。在假设 2.5 且满足 $1<p\le 2$时，由 (5.9) 和 (4.9) 可得逐样本界

$$\sqrt{n}\,\mathcal{AW}_p(P_n,P)\le T^{1/p-1/2}\sqrt{C_{\mathrm{A}}}\,\|G_n\|_{\mathcal{H}_P}.$$

令 $X_n=\|G_n\|_{\mathcal{H}_P}^2$ 和 $X=\|G\|_{\mathcal{H}_P}^2$。则 $X_n\Longrightarrow X$ 和 $\mathbb{E}X_n=\mathbb{E}\|\xi_1\|_{\mathcal{H}_P}^2=\mathbb{E}X<\infty$ 对每个 n，因此 $\mathbb{E}(X_n-K)_+\to\mathbb{E}(X-K)_+$对每个 $K>0$。利用 $X_n\mathbf{1}_{\{X_n>2K\}}\le 2(X_n-K)_+$可证明 $\{X_n\}$ 的一致可积性，并证明当 2.5对 $1<p\le 2$成立。

高斯阻碍的证明。

命题 2.8 的证明。设 $A\sim\mathrm{N}(0,s^2)$、$\tau^2=s^2+\sigma^2$，以及 $P=\mathrm{N}(0,\tau^2)$，其密度为 $\rho=\varphi_\tau$，累积分布函数为 $F(x)=\Phi(x/\tau)$，并记 $F_n(x)=\int_{-\infty}^x\rho_n(y)\,dy$。由于 $\mathbb{E}\int e_A^2/\rho=1+s^2/\sigma^2<\infty$，$\mathrm{H}_P$,0 中的希尔伯特空间中心极限定理蕴含 $\sqrt{n}(F_n-F)\Longrightarrow B_\sigma$在 $L^\infty(\mathbb{R})$ 中，其中 $B_\sigma(x)=\int_{-\infty}^x G(y)\,dy$的协方差为 $\mathrm{Cov}(B_\sigma(x)\ B_\sigma(y))=\mathrm{Cov}_\mu(\Phi_\sigma(x-A)\ \Phi_\sigma(y-A))$。记 $Y_i=A+\sigma Z_i(Z_1,Z_2\sim\mathrm{N}(0,1)$ 为独立同分布变量，且与 A 独立，则有 $V(x):=\mathrm{Var}(B_\sigma(x))=\mathbb{P}(Y_1>x,Y_2>x)-\mathbb{P}(Y_1>x)^2$。对大 $x>0$ 时的 $[x,x+x^{-1}]^2$ 上的 (Y_1,Y_2) 的二元高斯密度积分，得到 $\mathbb{P}(Y_1>x,Y_2>x)\ge cx^{-2}e^{-x^2/(2s^2+\sigma^2)}\gg\mathbb{P}(Y_1>x)^2,$，因此

$$V(x)\ge cx^{-2}e^{-x^2/(2s^2+\sigma^2)}.$$

对于 $J_R:=\int_{-R}^R|B_\sigma(x)|^p\rho(x)^{1-p}\,dx$，这个方差下界给出 $\mathbb{E}[|B_\sigma(x)|^p\rho(x)^{1-p}]\ge cx^{-p}e^{\zeta_p x^2}$，其中

$$\zeta_p=\frac{(p-2)s^2-\sigma^2}{2(s^2+\sigma^2)(2s^2+\sigma^2)}>0$$

由于 $\sigma^2<(p-2)s^2$。因此，$\mathbb{E}J_R\to\infty$因为 $R\to\infty$。对于 $Z\sim\mathcal{N}(0,1)$，取 $m_k=\mathbb{E}|Z|^k$，由柯西–施瓦茨不等式可得 $\mathbb{E}J_R^2\le(m_{2p}/m_p^2)(\mathbb{E}J_R)^2,$，而 Paley–Zygmund不等式推出 $\mathbb{P}(J_R>\frac12\mathbb{E}J_R)\ge m_p^2/(4m_{2p})>0$。令 $J_\infty=\lim_{R\to\infty}J_R$，我们得到对于任意 $M<\infty$，都有 $\mathbb{P}(J_\infty>M)\ge m_p^2/(4m_{2p})$，因此 $J_\infty=\infty$ 以正概率成立。

为将其提升为概率一，在 $\mathcal{H}_{P,0}$ 上定义积分算子 $(\mathcal{I}h)(x)=\int_{-\infty}^x h(y)\,dy$以及集合

$$\mathcal{B}_p=\left\{h\in\mathcal{H}_{P,0}:\int_{\mathbb{R}}|\mathcal{I}h(x)|^p\rho(x)^{1-p}\,dx<\infty\right\}.$$

由闵可夫斯基不等式，$\mathcal{B}_P$ 是一个线性子空间。由柯西–施瓦茨不等式，$\|\mathcal{I}h\|_\infty\le\|h\|_{\mathcal{H}_P}$，因此对每个整数 $k\ge1$，截断泛函 $\mathcal{J}_k(h)=\int_{-k}^k|\mathcal{I}h(x)|^p\rho(x)^{1-p}\,dx$ 在$\mathcal{H}_{P,0}$上连续。因此，

$$\mathcal{B}_p=\bigcup_{M=1}^{\infty}\bigcap_{k=1}^{\infty}\{h\in\mathcal{H}_{P,0}:\mathcal{J}_k(h)\le M\}$$

是 Borel 可测的。根据可测线性子空间的高斯零一律 [18, 推论 4.56],$\mathbb{P}(G\in\mathcal{B}_P)\in\{0,1\}$。由于 $B_\sigma=\mathcal{I}G$ 和 $J_\infty=\infty$ 具有正概率，$\mathbb{P}(G\in\mathcal{B}_P)=0$，即$J_\infty=\infty$ 几乎必然成立。

最后，在每个紧致分位数区间 $[F(-R)\ F(R)],$上，逆CDF Delta方法给出

$$n^{p/2}\int_{F(-R)}^{F(R)}|F_n^{-1}(u)-F^{-1}(u)|^p\,du\Longrightarrow J_R.$$

因此 下极限$_{n\to\infty}\mathbb{P}(n^{p/2}W_p(P_n,\ P)^p>M)\ge\mathbb{P}(J_R>M)\to1$当 $R\to\infty$对每个 $M<\infty$，从而证明（2.8）。 □

5 无界概率律的估计

本节将建立定理 2.6 和2.7的证明中所使用的加权传输界和特征矩估计。

5.1 二次加权估计

取 $p=2$ ($\mathsf{Q}_P = Q_P$,2)，并按时间顺序枚举这 $N=dT$ 个标量坐标 $(x_1^{(1)},\dots,x_T^{(d)})$。记 $\bar\rho_j(x_{1:j}) = \int \rho(x)\,dx_{j+1:N}$ 以及 $\bar h_j(x_{1:j}) = \int h(x)\,dx_{j+1:N}$，其中当 $\int h = 0$ 时，$\bar\rho_0 = 1$ 且 $\bar h_0 = 0$。固定 $\beta > 0$、$M_0 < \infty$ 和 $B = \beta\sigma^2 > \theta_N$（其中 θ_N 来自（2.7）），并令 C_{M_0} 表示满足 $\int e^{\beta|a|^2}\,\mu(da) \le M_0$ 的平滑密度 $\rho = d(\mu * \gamma_\sigma)/dx$ 类。

高斯包络和条件尾部。设 $\Delta = \alpha - b$ 并选择指数 α，b，η，$\omega > 0$ 满足

$$\begin{aligned} \alpha &> \frac{1}{2\sigma^2}, \qquad 0 < b < \frac{\beta}{1+2\beta\sigma^2}, \qquad \eta > 0, \\ \alpha + \eta + 2(N-1)\Delta &< \omega < \frac{\beta}{1+\beta\sigma^2}, \qquad \omega < 2b. \end{aligned} \tag{5.1}$$

条件 $B > \theta_N$ 等价于 $B^2 - (N-\frac{1}{2})B - (N-\frac{1}{2}) > 0$，或者

$$\frac{1}{2\sigma^2} + \frac{N-1}{\sigma^2(1+2B)} < \frac{\beta}{1+\beta\sigma^2},$$

因此，(5.1) 中的可行范围非空。令 $\mathcal{H} = \mathcal{H}_\omega = L^2(\mathbb{R}^N, e^{\omega|x|^2}\,dx)$，并令 $\mathcal{H}_0 = \{h \in \mathcal{H} : \int h = 0\}$。在承载 μ 一半质量的球上对 γ_σ 积分，针对 $e^{\beta|a|^2}$ 配方，并利用 $\mathbb{E}[e^{|x-A_{1:j}|^2/(4\sigma^2)} \mid X_{1:j} = x] \le Ce^{\alpha|x|^2}$，可得对 $\rho \in \mathcal{C}_{M_0}$ 和 $1 \le j \le N$ 一致成立的

$$\begin{aligned} ce^{-\alpha|x|^2} \le \bar\rho_j(x) &\le Ce^{-b|x|^2}, \\ |\nabla \log \bar\rho_j(x)| &\le C(1+|x|), \end{aligned} \tag{5.2}$$

以及对于任意 $b_+ \in (b, \beta/(1+2B))$，$|\nabla\bar\rho_j(x)| \le Ce^{-b_+|x|^2}$。

记 $u = x_{1:j-1}$ 和 $z = x_j$，其中一步条件密度 $k_j(u,z) = \bar\rho_j(u,z)/\bar\rho_{j-1}(u)$，条件累积分布函数 $\bar K_j(u,z) = \int_{-\infty}^z \bar k_j(u,r)\,dr$。令 $G_c(z) = \int_{|z|}^\infty e^{-cr^2}\,dr$ $(c>0)$，对 $z \ge 0$ 令 $I_z = (-\infty, z]$ 对 $z \le 0$ 和 $I_z = [z,\infty)$，并令 $\Lambda_j(u,z) = \int_{I_z} \bar\rho_j(u,r)\,dr$。于是 (5.2) 给出

$$\min(\bar K_j, 1-\bar K_j) \le \frac{\Lambda_j}{\bar\rho_{j-1}} \le Ce^{\Delta|u|^2} G_b(z).$$

对于 $i < j$，设定 $M_{ji} := \partial_{u_i}\Lambda_j - \Lambda_j \partial_{u_i} \log \bar\rho_{j-1}$，前缀归一化因子 $\bar\rho_{j-1}$ 从 $\partial_{u_i}(\Lambda_j/\bar\rho_{j-1})/k_j$ 中抵消，得到

$$C_{ji} := \frac{\partial_{u_i}\bar K_j}{\bar k_j} = \pm\frac{M_{ji}}{\bar\rho_j}.$$

结合 恒等算子与 b_+ 导数界以及式（5.2）中的得分估计可得

$$|M_{ji}(u,z)| \le Ce^{-b|u|^2} G_b(z) \le Ce^{-b|x_{1:j}|^2}, \qquad |C_{ji}(x_{1:j})| \le Ce^{\Delta|x_{1:j}|^2}.$$

三角右逆及其加权动能。对于 $h \in \mathcal{H}_0$，集合 $\mathcal{I}_j(u,z) = \int_{-\infty}^z h_j(u,r)\,dr$ 并且 $B_j = \mathcal{I}_j - K_j h_{j-1}$，并递归定义 $v = (v_1, \dots, v_N)$ 如下：

$$v_1 = -\frac{\mathcal{I}_1}{\bar\rho_1}, \qquad v_j = -\frac{B_j}{\bar\rho_j} - \sum_{i<j} v_i C_{ji}. \tag{5.3}$$

置于 $E_j(c) = \int \bar\rho_j(x_{1:j})|v_j(x_{1:j})|^2 e^{c|x_{1:j}|^2}\,dx_{1:j}$ 对于 $c \ge 0$。我们证明 $\sum_{j=1}^N E_j(\eta) \le C\|h\|_{\mathcal H}^2$.

在 I_z 上，上述条件尾界给出 $|B_j(u,z)| \le J_j(u,z) + Ce^{\Delta|u|^2} G_b(z)|\bar h_{j-1}(u)|$，其中 $J_j(u,z) := \int_{I_z} |\bar h_j(u,s)|\,ds$ 满足 $J_j(u,z)^2 \le G_\omega(z)\int_{I_z} e^{\omega s^2}|\bar h_j(u,s)|^2\,ds$，由柯西–施瓦茨不等式可得。由于 $\int_0^\infty e^{(\alpha+c)z^2} G_\omega(z)\,dz < \infty$ 对于 $\alpha+c<\omega$，直接项 $D_j(c) := \int (B_j^2/\bar\rho_j) e^{c|x_{1:j}|^2}\,dx_{1:j}$ 满足 $D_j(c) \le C\|h\|_{\mathcal H}^2$，只要

$$\alpha + c + 2\Delta < \omega, \qquad \alpha + c < 2b$$

（在 $j=1$ 处只需 $\alpha + c < \omega$，因为 $h_0 = 0$）。对于交叉项（$i<j$），使用 $C_{ji} = \pm M_{ji}/\bar\rho_j$ 可消除额外因子 $\bar\rho_{j-1}^{-1}$：

$$\begin{aligned}\int e^{c|x_{1:j}|^2}\bar\rho_j |v_i C_{ji}|^2\,dx_{1:j} &= \int e^{c|x_{1:j}|^2}|v_i|^2 \frac{|M_{ji}|^2}{\bar\rho_j}\,dx_{1:j} \\ &\le C\int |v_i(x_{1:i})|^2 e^{-(2b-\alpha-c)|x_{1:j}|^2}\,dx_{1:j}.\end{aligned}$$

当 $c < 2b-\alpha$ 时，对其余坐标 $x_{i+1:j}$ 积分并代入 $\bar\rho_i \ge c e^{-\alpha|x_{1:i}|^2}$，得到

$$\int e^{c|x_{1:j}|^2}\bar\rho_j |v_i C_{ji}|^2\,dx_{1:j} \le C E_i(c+2\Delta).$$

现在取 $c_j = \eta + 2\Delta(N-j)$。对于每个 $i<j$，都有 $c_i \ge c_j + 2\Delta$，因此第 j 层的每个交叉项都由 $E_i(c_i)$控制，而权重为 c_j 时的直接项条件则归结为 (5.1) 中的严格不等式。对 j 应用归纳法，便得到 $\sum_{j=1}^N E_j(\eta) \le C\|h\|_{\mathcal H}^2$。去掉权重或限制到 $\{|x_{1:j}|>K\}$ 则得到

$$\begin{aligned}\sum_{j=1}^N \int \bar\rho_j |v_j|^2 &\le C\|h\|_{\mathcal H}^2, \\ \int_{|x_{1:j}|>K} \bar\rho_j |v_j|^2 &\le C e^{-\eta K^2}\|h\|_{\mathcal H}^2.\end{aligned} \tag{5.4}$$

对于 $h \in C_c^\infty \cap \mathcal H_0$，写成 $F_i = \bar\rho_{j-1} v_i$ 可得 $\bar\rho_j v_j = -\mathcal I_j + K_j h_{j-1} - \sum_{i<j} F_i \partial_i K_j$；对关于 x_j 求导，利用 $\partial_j K_j = k_j$ 以及归纳恒等式 $\sum_{i<j}\partial_i F_i = -h_{j-1}$ 可得 $-\partial_j(\bar\rho_j v_j) - \sum_{i<j}\partial_i(\bar\rho_j v_i) = h_j$，根据 (5.4)，该结果可推广到所有 $h \in \mathcal H_0$。

度量比较与切空间范数的连续性。

命题 5.1（一致度量比较）。对于所有 ρ，$q \in \mathcal C_{M_0}$ 以及 ω 满足（5.1），

$$\mathcal{AW}_2(\rho\,dx, q\,dx) \le C\|\rho - q\|_{\mathcal H}.$$

证明。考虑线性路径 $\rho_s = (1-s)\rho + sq \in \mathsf{C}_{M_0}$ ($s\in[0,1]$)，其中 $h = q-\rho \in \mathrm{H}_0$，并令 $R_s : (0,1)^N \to \mathbb{R}^N$ 为其逆 Rosenblatt 映射。对于 $U \sim$ 均匀分布$((0,1)^N)$，耦合 $(R_0(U),\ R_1(U))$ 是双因果的，并满足 $\partial_s R_{s,j}(U) = v_{s,h,j}(R_s(U)_{1:j})$。根据（5.4），

$$\mathcal{AW}_2(\rho\,dx, q\,dx)^2 \le \mathbb{E}|R_1(U) - R_0(U)|^2 \le \int_0^1 \sum_{j=1}^N \int \bar\rho_{s,j}|v_{s,h,j}|^2\,ds \le C\|h\|_{\mathcal H}^2. \qquad \square$$

根据 (5.4)，对于所有 $h \in \mathrm{H}_0$，我们还有 $\mathsf{Q}_P(h)^{1/2} \le C\|h\|_{\mathcal H}$。令 $\ell = \frac12 + \frac{N-1}{1+2B}$、$u = \frac{B}{1+B}$ 和 $g = u - \ell > 0$，满足 (5.1) 的一组显式指数选择为 $\omega = (\ell+u)/(2\sigma^2)$、$\alpha = (1/2 + g/(16N))/\sigma^2$、$b = (B/(1+2B) - g/(16N))/\sigma^2$ 和 $\eta = g/(8\sigma^2)$。

5.2 二次指数以上的加权估计

现在取 $p>2$，保持 $\beta>0$、$M_0<\infty$、$B=\beta\sigma^2$ 和 $\mathcal{C}_{M_0}$，它们来自第 5.1 节，并去掉 ρ_j、h_j、k_j、K_j（$\rho_0=1$）上的上横线。假设 $B>\Theta_{p,N}$，其中 $\Theta_{p,N}$ 定义见（2.7）。

参数、密度包络和特征矩。对于 $B>\Theta_{p,N}$，存在满足以下条件的指数 α、b、a：

$$\alpha>\frac{1}{2\sigma^2},\qquad 0<b<\frac{\beta}{1+2B},\qquad 0<a<\frac{\beta}{2(1+B)},\qquad a<b,$$
$$p\{\alpha-a+(N-1)(\alpha-b)\}<b. \tag{5.5}$$

事实上，在 α、b、a的边界值处，最后一个不等式变为 $2B^2+\{2-p(N+1)\}B-pN>0$，其正根为 $\Theta_{p,N}$ 将该二次式在 $B=p(N+1)/2$ 处取值可得 $p>0$，从而确认 $\Theta_{p,N}<p(N+1)/2$ 属于（2.7）。

固定一个整数 $s>N/2$，并回顾推论 4.10中的希尔伯特空间：

$$\mathcal{H}_{a,s}=\{h:e^{a|x|^2}h\in H^s(\mathbb{R}^N)\},\qquad \|h\|_{\mathcal{H}_{a,s}}=\|e^{a|x|^2}h\|_{H^s},\qquad \mathcal{H}_{a,s,0}=\left\{h\in\mathcal{H}_{a,s}:\int h=0\right\}. \tag{5.6}$$

索伯列夫嵌入（$s>N/2$）给出 $|h(x)|\le C\|h\|_{\mathcal{H}_{a,s}}e^{-a|x|^2}$，对 $x_{j+1:N}$ 积分掉后得到$|h_j(x_{1:j})|\le C\|h\|_{\mathcal{H}_{a,s}}e^{-a|x_{1:j}|^2}$而每个 $\rho\in\mathcal{C}_{M_0}$ 都满足（5.2）。将 $e^{a|x|^2}e_y(x)$ 微分至 s 阶并配方可得，对于某个整数 $L<\infty$,

$$\|e_y\|^2_{\mathcal{H}_{a,s}}\le C(1+|y|)^L\exp\left(\frac{2a}{1-2a\sigma^2}|y|^2\right).$$

由于 $2a/(1-2a\sigma^2)<\beta$ 根据（5.5），我们得到

$$\mathbb{E}_\mu\|e_A\|^2_{\mathcal{H}_{a,s}}<\infty. \tag{5.7}$$

具有一致能量尾部的标量右逆。按照第 5.1 节中的 $u=x_{1:j-1}$, $z=x_j$, $\Delta=\alpha-b>0$ 和 Λ_j, M_{ji}, C_{ji}，我们再次有 $|M_{ji}(u,z)|\le Ce^{-b|u|^2}G_b(z)$ 和 $|C_{ji}(x_{1:j})|\le Ce^{\Delta|x_{1:j}|^2}$。对于 $h\in\mathrm{H}_{a,s,0}$，设 $h_0=0$, $\mathrm{I}_j(u,z)=\int_{-\infty}^z h_j(u,y)\,dy$，且 $B_j=\mathrm{I}_j-K_jh_{j-1}$。关于 I_z 的边缘界给出

$$|B_j(u,z)|\le C\|h\|_{\mathcal{H}_{a,s}}\{e^{-a|u|^2}G_a(z)+e^{-(a-\Delta)|u|^2}G_b(z)\}\le C\|h\|_{\mathcal{H}_{a,s}}e^{-(a-\Delta)|x_{1:j}|^2}$$

（对于 $j\ge2$，且在 $j=1$ 处没有第二项）。由 $v_1=-\mathcal{I}_1/\rho_1$ 和 $v_j=-B_j/\rho_j-\sum_{i<j}v_iC_{ji}$ 定义的递归速度 $v=(v_1,\dots,v_N)$，通过对 $1\le j\le N$ 作归纳法，满足

$$|v_j(x_{1:j})|\le C\|h\|_{\mathcal{H}_{a,s}}e^{[\alpha-a+(j-1)\Delta]|x_{1:j}|^2},$$

因为每个新坐标最多会使前述速度项的指数增加 Δ。因此，

$$\rho_j(x_{1:j})|v_j(x_{1:j})|^p\le C\|h\|^p_{\mathcal{H}_{a,s}}e^{-\{b-p[\alpha-a+(j-1)\Delta]\}|x_{1:j}|^2}.$$

由 $p\{\alpha-a+(N-1)\Delta\}<b$中的严格不等式（5.5），每个阶段都保留了正的高斯衰减率。将 $\eta>0$ 选为小于这些衰减率中的最小值，即可得到

$$\sum_{j=1}^N\int|v_j|^p\rho_j\le C\|h\|^p_{\mathcal{H}_{a,s}},$$
$$\int_{|x_{1:j}|>R}|v_j|^p\rho_j\le Ce^{-\eta R^2}\|h\|^p_{\mathcal{H}_{a,s}}, \tag{5.8}$$

且 $C_c^\infty\cap\mathcal{H}_{a,s,0}$ 在 $\mathcal{H}_{a,s,0}$ 中的密度给出 $h_j=-\sum_{i\le j}\partial_i(\rho_jv_i)$。

度量比较与分位数映射的微分。

命题 5.2 (一致加权比较)。 对于所有 $\rho,\ q \in \mathcal{C}^{M_0}$ 满足 (5.5)，N 阶标量适应距离 $\mathcal{AW}_{p,\,\mathrm{sc}}$ 以及 T 块距离 $\mathcal{AW}_p$ 满足

$$\mathcal{AW}_{p,\mathrm{sc}}(\rho\,dx, q\,dx) \le C_{M_0}\|q-\rho\|_{\mathcal{H}_{a,s}}, \qquad \mathcal{AW}_p(\rho\,dx, q\,dx) \le d^{(1/2-1/p)_+} C_{M_0}\|q-\rho\|_{\mathcal{H}_{a,s}},$$

且 $N_{P,\,p}$ 在 $\boldsymbol{H}_{a,\,s,\,0}$ 上关于 $\rho \in \mathcal{C}^{M_0}$ 连续。

证明。沿着 $\rho_\tau = (1-\tau)\rho + \tau q \in \mathcal{C}^{M_0}(h = q-\rho,\ \tau \in [0,1])$，在 $(0,1)^N$ 上对 $\partial_\tau R_{\tau,j}(u) = v_{\tau,h,j}(R_\tau(u)_{1:j})$ 积分，并应用闵可夫斯基不等式和 (5.8)，得到

$$\|R_1 - R_0\|_{L^p(du;\ell^p)} \le \int_0^1 \|v_{\tau,h}\circ R_\tau\|_{L^p(du;\ell^p)}\,d\tau \le C_{M_0}\|h\|_{\mathcal{H}_{a,s}},$$

这证明了 标量界。分块界由 $|z|_2 \le d^{(1/2-1/p)_+}|z|_p$ 在 $\mathbb{R}^d$ 上推出。 □

当 $d=1$ 且 $h \in \mathcal{D}$ 时，上述逐点速度界和（5.8）验证了（4.8），其中$r=p$。第 4.2 节中的 Vitali 论证继而给出 $\varepsilon^{-1}(R_\varepsilon - R_0) \to v_{\rho,\,h}\circ R_0$ 在 $L^p((0,1)^T;\ \ell^p)$，从而完成了关于 2.9 的命题 $p>2$。

5.3 强对数凹参考律

条件输运–熵比较。记 $\mathrm{KL}(Q \mid P) = \int \log(dQ/dP)\,dQ$ 若$Q \ll P$，否则 $+\infty$。

引理 5.3 （由条件输运–熵界得到适应输运）。 假设 P 具有有限二阶矩且条件密度 $k_t^P(x_{1:t-1})$ 为正，并且对所有 $1 \le t \le T$、$x_{1:t-1}$、$y_{1:t-1} \in \mathbb{R}^{d(t-1)}$ 和 $\nu \in \boldsymbol{P}(\mathbb{R}^d)$，

$$W_2(\nu, k_t^P(x_{1:t-1}))^2 \le 2C_{\mathrm{LS}}\,\mathrm{KL}(\nu \mid k_t^P(x_{1:t-1})),$$
$$W_2(k_t^P(x_{1:t-1}), k_t^P(y_{1:t-1})) \le L|x_{1:t-1} - y_{1:t-1}|.$$

那么，每个具有有限二阶矩的路径律 Q 都满足

$$\mathcal{AW}_2(P,Q)^2 \le C_{\mathrm{A}}\,\mathrm{KL}(Q \mid P), \qquad C_{\mathrm{A}} = 4C_{\mathrm{LS}}(1+2L^2)^{T-1}.$$

特别地，若 P 和 Q 的密度分别为 $\rho>0$ 和 q，则

$$\mathcal{AW}_2(P,Q)^2 \le C_{\mathrm{A}}\int \frac{(q-\rho)^2}{\rho}. \tag{5.9}$$

证明。对于 $\mathrm{KL}(Q \mid P) < \infty$，通过 $k_t^P(X_{1:t-1})$ 与 $k_t^Q(Y_{1:t-1})$ 的可测最优（或 ε-最优）单步 W_2 耦合，逐期耦合 (X,Y)。设 $S_t = \sum_{i\le t}\mathbb{E}|X_i - Y_i|^2$ 且 $H_t = \int Q_{1:t-1}(dy)\mathrm{KL}(k_t^Q(y) \mid k_t^P(y))$，通过 $k_t^P(Y_{1:t-1})$ 应用 W_2 三角不等式可得

$$S_t \le (1+2L^2)S_{t-1} + 4C_{\mathrm{LS}}H_t, \qquad S_0 = 0.$$

由 $\sum_{t=1}^T H_t = \mathrm{KL}(Q \mid P)$ 根据相对熵链式法则，从 $t=1$ 迭代至 T可得熵界；再利用 $\log(q/\rho) \le q/\rho - 1$ 即可得到（5.9）。 □

参考类的条件常数。 在假设 2.5下，回顾 $m = \kappa/(1+\kappa\sigma^2)$ 以及 $M = \sigma^{-2}$ 在（4.5）中的结果，并设

$$R_f = \frac{C_f}{c_f}, \qquad D_0 = \frac{C_f - c_f}{2\sigma^2\sqrt{\kappa + \sigma^{-2}}}.$$

P 的前缀密度分解为 $\rho_{1:t} = \rho_0 r_{1:t}$，其中 $r_t(x_{1:t}) = \mathbb{E}_{\mu_0}[f(A) \mid X_{1:t} = x_{1:t}] \in [c_f, C_f]$ 满足 $\|\nabla r_t\|_\infty \le D_0$。对于任意历史 $a = x_{1:t-1}$，在 P_0 和 P 下，条件分布 $k_{0,a}$ 和 k_a 由式 (4.5) 满足 $\nabla_z^2(-\log k_{0,a}) \succeq m I_d$，因此 $k_{0,a}$ 满足常数为 $1/m$ [10, 14] 的对数 Sobolev 不等式。由于 $dk_a/dk_{0,a}(z) = r_t(a,z)/r_{t-1}(a)$ 的变化幅度至多为因子 R_f，Holley–Stroock 扰动和 Otto–Villani 定理 [23, 15] 给出 k_a 的条件 T_2不等式，其常数为 $C_{\mathrm{LS}} = R_f/m$。

为界定历史 Lipschitz 常数 L，取一个单位历史向量 e并将其分解为

$$\partial_e \log \rho_{1:t}(a,z) = F_e(z) + G_e(z), \qquad F_e = \partial_e \log \rho_{0,1:t}, \quad G_e = \partial_e \log r_t.$$

由 $mI \preceq \nabla^2(-\text{对数}\,\rho_{0,1:t}) \preceq MI$，减去 $\frac{1}{2}(M+m)I$可知，$\nabla^2(-\text{对数}\,\rho_{0,1:t})$ 在 a 与 z 之间的非对角块的算子范数至多为 $M_\times = (M-m)/2$。因此$|\nabla_z F_e| \le M_\times$，于是对 $k_{0,a}$ 应用 Poincaré 不等式可得

$$\mathrm{Var}_{k_a}(F_e) \le R_f \,\mathrm{Var}_{k_{0,a}}(F_e) \le R_f M_\times^2/m, \qquad |G_e| \le D_0/c_f.$$

由于 ∂_e 对数 k_a 是 $F_e + G_e$的以 k_a 为中心的版本，因此可得

$$\int |\partial_e \log k_a|^2 \, dk_a \le S^2 := \frac{2R_f M_\times^2}{m} + 2\left(\frac{D_0}{c_f}\right)^2.$$

沿着 $a_s = (1-s)a + sb$ 进行插值并求解 $-\nabla_z \cdot (k_{a_s}\nabla u_s) = \partial_s k_{a_s}$，对 k_{a_s} 应用 Poincaré 不等式可得 $\int |\nabla u_s|^2 \, dk_{a_s} \le (R_f/m)S^2|a-b|^2$，而 Benamou–Brenier 公式给出 $W_2(k_a, k_b) \le L|a-b|$，其中 $L^2 = (R_f/m)S^2$。将 $m, M_\times, D_0$用 $z_0 = \kappa\sigma^2$表示并代入可得

$$L^2 = \frac{R_f^2}{2z_0^2} + \frac{R_f(R_f-1)^2}{2z_0}, \qquad C_{\mathrm{A}} = \frac{4R_f}{m}(1+2L^2)^{T-1}. \tag{5.10}$$

此外，对于任意有界零均值倾斜 g满足 $|\varepsilon|\|g\|_\infty \le 1/2$，我们有 $c_f/2 \le f(1+\varepsilon g) \le 3C_f/2$且 $\rho + \varepsilon h_g \ge \rho/2$；将 5.9) 中的 $P_\varepsilon^{h_g}$作为参考概率律应用，可得

$$\mathcal{AW}_2(Q, P_\varepsilon^{h_g}) \le C_* \|q - (\rho + \varepsilon h_g)\|_{\mathcal{H}_P} \tag{5.11}$$

对于每个在有限 $\mathcal{H}_P$-距离下的概率密度 $q = dQ/dx$。

内禀特征矩。

引理 5.4 (内禀二阶特征矩)。 在假设 *2.5* 下，对于每个 $\sigma > 0$,

$$\mathbb{E}_\mu \|e_A\|_{\mathcal{H}_P}^2 \le \frac{C_f}{c_f}\left(1 + \frac{2}{\kappa\sigma^2}\right)^N < \infty. \tag{5.12}$$

证明。 令 $q = \mu_0 * \gamma_{\sigma/\sqrt{2}}$, 于是 $\rho_0 = \mu_0 * \gamma_\sigma = q * \gamma_{\sigma/\sqrt{2}}$。在带宽 4.5 处应用 $\sigma/\sqrt{2}$ 可知 $U = -\log q$满足 $m_* I \preceq \nabla^2 U \preceq 2\sigma^{-2} I$其中 $m_* = \kappa/(1+\kappa\sigma^2/2)$。对于$Z \sim \mathcal{N}(0, \sigma^2 I/2)$，对 $U(x-Z)$ 作泰勒展开，得到

$$\frac{\rho_0(x)}{q(x)} = \mathbb{E}[e^{U(x)-U(x-Z)}] \ge \mathbb{E}\exp\left(-\nabla U(x)\cdot Z - \frac{|Z|^2}{\sigma^2}\right) = 2^{-N/2}\exp\left(\frac{\sigma^2|\nabla U(x)|^2}{8}\right).$$

由于 $|\nabla U(x)| \ge m_*|x - x_*|$围绕 U 的唯一极小化元 x_*，积分可得 $\int (q(x)/\rho_0(x))\, dx \le 2^{N/2}(8\pi/[\sigma^2 m_*^2])^{N/2}$。利用 $\mathbb{E}_{\mu_0}[e_A(x)^2] = (4\pi\sigma^2)^{-N/2} q(x)$ 和托内利定理，我们得到

$$\mathbb{E}_{\mu_0} \int \frac{e_A(x)^2}{\rho_0(x)}\, dx \le \left(\frac{2}{\sigma^2 m_*}\right)^N = \left(1 + \frac{2}{\kappa\sigma^2}\right)^N.$$

密度界（2.6）给出 $\mu \le C_f \mu_0$ 和 $\rho \ge c_f \rho_0$，这意味着（5.12）。 □

A 谱公式和数值细节

A.1 二次谱和四点分布

对于 $p=2$，记作 $\mathcal{Q}_P = Q_{P,\,2}$ 以及 $\mathcal{A}_P = \mathcal{A}_{P,\,2}$；到 $\ker(D_\rho)^\perp \subset \mathcal{A}_P$ 上的正交投影定义了一个有界线性最小能量算子 $\mathcal{V}_P : \mathcal{H}_0 \to \mathcal{A}_P$。

推论 A.1（迹恒等式）。 根据定理 *2.2*、 *2.6*或 *2.7* 在 $p=2$处，特征值 $\{\lambda_j\}$在备注*2.3* 中满足

$$\mathbb{E}\mathcal{Q}_P(G) = \mathbb{E}_\mu \mathcal{Q}_P(e_A - \rho) = \sum_j \lambda_j.$$

证明。由于 $\mathrm{V}_P : \mathrm{H}_0 \to \mathrm{A}_P$ 有界且线性，$\mathrm{V}_P G$ 是一个以 A_P为值域、均值为零的高斯元素，其协方差为迹类协方差 $\mathrm{V}_P C_G \mathrm{V}_P^*$。令 $\{u_j\}$ 为 A_P 中 $\mathrm{V}_P C_G \mathrm{V}_P^*$ 的标准正交特征基，则 Karhunen–Loève 展开 $\mathrm{V}_P G \overset{d}{=} \sum_j \sqrt{\lambda_j} Z_j u_j$（$Z_j \sim \mathrm{N}(0,1)$ 独立同分布）给出备注 2 3 $\sum_j \lambda_j = \mathbb{E}\| \ {}_P G\|^2_{\mathcal{A}_P} =$。对其取平方范数并取 $\mathbb{E}_\mu\| \ {}_P \ e_A - \rho \ \|^2_{\mathcal{A}_P}$ 期望可得 V V ( )。 □

非零特征值$\{\lambda_j\}$ 也与定义在 $L^2(\mu)$ 上的积分算子$\mathcal{K}$ 的特征值相同，其中

$$(\mathcal{K}g)(a) = \int K(a,b) g(b)\, \mu(db), \qquad K(a,b) = \langle \mathcal{V}_P(e_a - \rho), \mathcal{V}_P(e_b - \rho) \rangle_{\mathcal{A}_P}.$$

事实上，定义 $S\colon L^2(\mu) \to \mathcal{A}_P$，使得 $Sg = \mathbb{E}_\mu[g(A)\mathcal{V}_P(e_A - \rho)]$，可得 $\mathcal{V}_P C_G \mathcal{V}_P^* = SS^*$，且$\mathcal{K} = S^*S$。

当 $p \neq 2$ 时，考虑四点分布 $\mu =$均匀分布$\{-a,a\} \otimes$均匀分布$\{-b,b\}$ 在 $\mathbb{R}^2 (a,b>0)$ 上。记 φ_σ 为密度，$\Phi_\sigma(x) = \Phi(x/\sigma)$ 为 $\mathcal{N}(0,\sigma^2)$ ($\Phi = \Phi_1$) 的累积分布函数，并令 $c>0$ 及整数 $k \geq 1$,

$$f_c(x) = \frac{\varphi_\sigma(x-c) + \varphi_\sigma(x+c)}{2}, \qquad D_c(x) = \frac{\Phi_\sigma(x-c) - \Phi_\sigma(x+c)}{2},$$

$$s_c(x) = \tanh(cx/\sigma^2), \qquad C_{c,p} = \int \frac{|D_c(x)|^p}{f_c(x)^{p-1}}\, dx, \qquad \kappa_{c,k} = \int s_c(x)^k f_c(x)\, dx.$$

这里 $\rho(x,z) = f_a(x) f_b(z)$。对于任意扰动 $h = \rho[u s_a(x) + v s_b(z) + w s_a(x) s_b(z)]$ ($u,v,w \in \mathbb{R}$)，公式 (3.1) 得到

$$v_1(x) = -u D_a(x)/f_a(x), \qquad v_2(x,z) = -(v + w s_a(x)) D_b(z)/f_b(z),$$

因此

$$Q_{P,p}(h) = C_{a,p}|u|^p + C_{b,p} \int |v + w s_a(x)|^p f_a(x)\, dx.$$

对于独立同分布的路径 $A^{(i)} = (a\varepsilon_i,\ b\zeta_i)$，其带有独立的 Rademacher 符号 $\varepsilon_i,\ \zeta_i \in \{\pm 1\}$ 以及样本均值

$$u_n = \frac{1}{n}\sum_{i=1}^n \varepsilon_i, \qquad v_n = \frac{1}{n}\sum_{i=1}^n \zeta_i, \qquad w_n = \frac{1}{n}\sum_{i=1}^n \varepsilon_i \zeta_i,$$

恒等算子 $\varphi_\sigma(x \mp a) = f_a(x)[1 \pm s_a(x)]$ gives

$$\rho_n(x,z) = f_a(x) f_b(z)[1 + u_n s_a(x) + v_n s_b(z) + w_n s_a(x) s_b(z)].$$

由于 $\sqrt{n}(u_n, v_n, w_n) \Longrightarrow (Z_1, Z_2, Z_3) \sim \mathcal{N}(0, I_3)$，定理 2.2 给出

$$n^{p/2} \mathcal{AW}_p(P_n, P)^p \Longrightarrow C_{a,p}|Z_1|^p + C_{b,p} \int |Z_2 + Z_3 s_a(x)|^p f_a(x)\, dx.$$

根据对称性，$\int s_a(x)^{2m-1} f_a(x)\,dx = 0$ 对所有 $m \geq 1$均成立。在 $p = 2$ 处，极限简化为 $C_{a,}$ $_2Z_1^2 + C_{b,}\ _2Z_2^2 + \kappa_{a,}\ _2C_{b,}\ _2Z_3^2$，而在 $p = 4$ 处，它等于

$$S_4 = C_{a,4} Z_1^4 + C_{b,4} \left(Z_2^4 + 6\kappa_{a,2} Z_2^2 Z_3^2 + \kappa_{a,4} Z_3^4 \right),$$

因此 $n^2 \mathcal{AW}_4(P_n,\ P)^4 \Longrightarrow S_4$ 且 $\sqrt{n}\,\mathcal{AW}_4(P_n,\ P) \Longrightarrow S_4^{1/4}$。若将 $\widehat{\mu}_n$ 替换为其经验边缘分布的乘积，则会将 w_n 替换为 $u_n v_n = o_P(n^{-1/2})$，从而使交互项 Z_3 在极限中消失。

A.2 无界支撑的例子

首先，假设 $d = 1$ 且 $\mu = \mathcal{N}(m_0,\ \Sigma)$ 定义在 $\mathbb{R}^T$ 上，协方差为 $\Sigma \succeq 0$（可能是奇异的）。记 $\lambda_{\max}(\Sigma)$ 为 Σ的最大特征值，则 $\mathbb{E}_\mu[e^{\beta|A|^2}] < \infty$ 对每个$\beta < (2\lambda_{\max}(\Sigma))^{-1}$都成立，因此定理 2.6 在

$$\lambda_{\max}(\Sigma) < \begin{cases} \sigma^2/(2\theta_T), & 1 < p \leq 2, \\ \sigma^2/(2\Theta_{p,T}), & p > 2. \end{cases}$$

相反，若 $\Sigma \succ 0$ 在任意维度 $d \geq 1$上均为非退化的，则 $\mathrm{N}(m_0, \Sigma)$——以及任何满足 $0 < c_f \leq f \leq C_f < \infty$的扰动$\mu = f\mathrm{N}(m_0, \Sigma)$——都满足假设 2.5，且 $\kappa = \lambda_{\max}(\Sigma)^{-1}$。于是，对每个 $1 < p \leq 2$ 及每个固定的 $\sigma > 0$，定理 2.7 均成立，且对$\lambda_{\max}(\Sigma)/\sigma^2$没有任何约束。若参考势函数在无穷远处的增长快于二次函数，例如 $\mu_0(da) \propto e^{-|a|^4 - \kappa|a|^2/2}\,da$，则对所有 $\beta > 0$均有 $\mathbb{E}_\mu[e^{\beta|A|^2}] < \infty$，并且定理 2.7 对所有$1 < p < \infty$ 和所有 $\sigma > 0$均成立。

在假设 Σ/σ^2 2 5, $\mu = m_0\ \Sigma$ 5 4 $I_\mu = \mathbb{E}_\mu\|e_A\|_{\mathcal{H}_P}^2$ $_2\ P_n\ P\ ^2 \leq C_{\mathrm{A}}\|\rho_n - \rho\|_{\mathcal{H}_P}^2$ 5 12 $I_\mu =$ 5 9 $I_N +$.引理 .界：由 (.)（当 $\det()$ 时，取 $\mathcal{N}(,)$）。将 (.$\mathcal{AW}(,)$中的结果与 (.中的 $\mathbb{E}\|\rho_n - \rho\|_{\mathcal{H}_P}^2 = n^{-1}(I_\mu - 1)$ 结合，可得有限样本界

$$\mathbb{E}\mathcal{AW}_2(P_n, P)^2 \leq \frac{C_{\mathrm{A}}}{n} \left[R_f \left(1 + \frac{2}{\kappa\sigma^2} \right)^{dT} - 1 \right],$$

其中 $R_f = C_f/c_f$ 和 C_{A} 由 (5.10) 给出。

A.3 数值实现

对于 (w 中的四条路径上的任意权重向量 3.2)，记 $P_w = \sum_{i=1}^4 w_i \mathcal{N}(a_i, \sigma^2 I_2)$，并令 R_w为其逆 Rosenblatt 映射。给定第一坐标 x，第一期上分支的后验赔率正比于$\exp(2ax/\sigma^2)$，因此随 x 增大。给定下分支，第二坐标服从均值为 $-2a$、0的高斯混合；而在上分支中，其均值为 0、$2a$；对于任意非负权重，后者在一阶随机序上均支配前者。因此，P_w 的第二期条件分位数随 x 增大，对于任意 w均成立，包括经验权重和自助法权重。

对于两个权重向量 w、$\widetilde{w}$，令 r(x、 u) 和 s(y、 u) 分别表示 P_w和 $P_{\widetilde{w}}$ 的第二期条件分位数函数。第二期传输成本为 $\int_0^1 |r(x、\ u) - s(y、\ u)|^p\,du$。由于 $r\ \ s$ $\int_0^1 |r\ x\ u\ - s\ y\ u\ |^p\,du$ $|r-s|^p$ $x\ y$ $|x-y|^p +$和 在其第一个参数上递增，且 为次模，因此总和 (、) (、) 关于 (、) 是次模的。因此，第一期单调耦合求解了动态规划问题（即 [7, 命题 3.5] 的两期情形），从而得到

$$\mathcal{AW}_p(P_w, P_{\widetilde{w}})^p = \int_{(0,1)^2} \sum_{t=1}^{2} |R_{w,t}(u_{1:t}) - R_{\widetilde{w},t}(u_{1:t})|^p\,du_1 du_2. \tag{A.1}$$

每个条件累积分布函数都是至多三个高斯累积分布函数的混合，我们通过带括区间的牛顿迭代将其反演，使累积分布函数残差 10^{-13} 或区间宽度 10^{-12}达到要求，并通过张量积高斯–勒让德求积计算（A.1），每个坐标使用48 个节点。

要根据 (3.1) 计算极限分位数，令 $\psi_i = e_{a_i} - e_{a_4} (i = 1, 2, 3)$。重缩放经验权重向量 $\sqrt{n}(\widehat{w}_{1:3} - w_{1:3})$ 的协方差为 $\Sigma_w = \mathrm{diag}(w_{1:3}) - w_{1:3}w_{1:3}^{\mathsf{T}}$，因此 $G = \sum_{i=1}^{3} Z_i\psi_i$，其中 $Z \sim \mathrm{N}(0, \Sigma_w)$。在每个坐标上使用 192 个 Gauss–Legendre 节点计算 (3.1) 中的三个标量速度场，可在 $p = 2$处得到一个 3×3能量矩阵（其谱见注 2.3），并在 $p = 4$处得到关于 Z 的齐次四次多项式；参考分位数由 500,000 次独立高斯抽样估计。

对于百慕大看跌期权，写作 $c_i = a_{i,1}$， $b_i = a_{i,2}$， $k = \log(K/S_0)$，以及

$$H_i = K\Phi\left(\frac{k - b_i}{\sigma}\right) - S_0 e^{b_i + \sigma^2/2}\Phi\left(\frac{k - b_i - \sigma^2}{\sigma}\right),$$

$(K - S_0e^{X_2})_+$ 在分量 i中的高斯期望。令 $\rho_{w,1}(x) = \sum_{i=1}^{4} w_i\varphi_\sigma(x - c_i)$，通过反向归纳可得

$$\mathrm{V}(P_w) = \int_{\mathbb{R}} \max\left\{ e^{-0.03\Delta}(K - S_0e^x)_+\rho_{w,1}(x),\, e^{-0.06\Delta}\sum_{i=1}^{4} w_i\varphi_\sigma(x - c_i)H_i \right\} dx,$$

采用自适应求积计算，绝对和相对误差容限为 10^{-9} （在 $[-a-10\sigma, a+10\sigma]$ 上，省略价格单位下小于 10^{-19} 的高斯尾部概率质量）。未平滑值 $\mathrm{V}(\mu)$ 通过对二叉树进行反向归纳计算。

每个数据集由按概率 w 抽取的多项路径计数组成（随机种子 20261116），每次自助抽样则根据经验权重对多项计数进行重抽样。每个数据集使用 $B_n = 249$ 次抽样，$\widehat{q}^{*}_{n,\ 0.95}$ 取为缩放后自助距离的第 $\lceil 0.95 \times 249 \rceil = 237$ 个次序统计量。

求积检验与覆盖率。为验证求积稳定性，我们在每个样本量下，使用每个坐标96 个节点，重新计算了前10 个数据集距离及其前20个自助法距离（总计1100 对）；最大绝对变化为 1.2×10^{-6}，非零距离中的最大相对变化为 2.2×10^{-5}。将极限计算的节点数从192增加到384 后，三个二次特征值——0.0122484、0.0139312和0.0757441——的相对变化均小于 3.6×10^{-7}。小倾斜下的有限差分检验结果也与切向公式（3.1）一致。表600报告了1次重复试验的经验覆盖率；每个数据集均使用其自身的自助法半径进行评估。

	距离界		平滑值		未平滑目标	
n	$p = 2$	$p = 4$	$p = 2$	$p = 4$	$p = 2$	$p = 4$
50	94.2	94.7	100.0	100.0	100.0	100.0
200	96.0	96.3	100.0	100.0	100.0	100.0
800	94.7	95.5	100.0	100.0	100.0	100.0
3200	95.8	96.2	100.0	100.0	57.3	100.0
12800	95.2	95.2	100.0	100.0	0.0	0.3

表 1：每个 n下 600 个独立数据集的经验覆盖率百分比，使用 249 次自助重抽样和 95% 的名义水平。各列报告 $\mathcal{AW}_p(P_n,\ P) \le \widehat{q}^{*}_{n,\ 0.95}/\sqrt{n}$, $|\mathrm{V}(P_n) - \mathrm{V}(P)| \le L\widehat{q}^{*}_{n,\ 0.95}/\sqrt{n}$ 的覆盖率，以及将相同半径应用于未平滑值 $\mathrm{V}(\mu)$ 周围时的覆盖率（覆盖率为 95% 时的蒙特卡洛标准误为 $\approx 0.9\%$）。

人工智能使用声明

在准备本文期间，使用了生成式人工智能工具辅助证明探索、中间推导、语言编辑以及数值代码和图表的开发。作者已独立检查并验证所有数学论证、参考文献、代码和数值结果，并对论文内容负全部责任。

参考文献

[1] B. Acciaio、D. Bartl、A. Grass、S. Hou 和 G. Pammer。《适应性 Bures–Wasserstein 空间的几何》，2026年。网址： https://arxiv.org/abs/2602.00623, arXiv:2602.00623。[2] B. Acciaio、D. Kršek、G. Pammer 和 M. Rodrigues。《随机过程的绝对连续曲线》，2025年。网址： https://arxiv.org/abs/2506.13634, arXiv:2506.13634。[3] L. Ambrosio、N. Gigli 和 G. Savaré。《度量空间与概率测度空间中的梯度流》。Birkhäuser，巴塞尔，2005年。 doi:10.1007/b137080。[4] J. Backhoff、D. Bartl、M. Beiglböck 和 J. Wiesel。《在适应Wasserstein距离下估计随机过程》。《应用概率年刊》，32(1)：529–550，2022年。网址： https://arxiv.org/abs/2002.07261, doi:10.1214/21-AAP1687。[5] J. Backhoff-Veraguas、D. Bartl、M. Beiglböck 和 M. Eder。《适应Wasserstein距离与数理金融中的稳定性》。《金融与随机性》，24(3)：601–632，2020年。doi:10.1007/s00780-020-00426-3。[6] J. Backhoff-Veraguas、D. Bartl、M. Beiglböck 和 M. Eder。《所有适应拓扑都相同》。《概率论及其相关领域》，178：1125–1172，2020年。 doi:10.1007/s00440-020-00993-8。[7] J. Backhoff-Veraguas、S. Källblad 和 B. A. Robinson。《SDEs 的概率律之间的适应Wasserstein距离》。《随机过程及其应用》，189：104689，2025年。网址： https://arxiv.org/abs/2209.03243, doi:10.1016/j.spa.2025.104689。[8] J. Blanchet、M. Larsson、J. Park 和 J. Wiesel。《适应Wasserstein度量的界》。《亨利·庞加莱研究所年刊·概率与统计》，2026年。即将刊出；arXiv:2407.21492v3，2026年8月9日。网址： https://arxiv.org/abs/2407.21492。[9] J. Blanchet、J. Wiesel、E. Zhang 和 Z. Zhang。《通过适应Wasserstein距离进行经验鞅投影》。《应用概率年刊》，36(1)：547–606，2026年。网址： https://arxiv.org/abs/2401.12197, doi:10.1214/25-AAP2239。[10] L. A. Caffarelli。《最优传输的单调性及 FKG 不等式和相关不等式的性质》。《数学物理通讯》，214(3)：547–563，2000年。[11] E. A. Carlen、D. Cordero-Erausquin 和 E. H. Lieb。《Brascamp–Lieb型非对称协方差估计及对数凹测度的相关不等式》。《亨利·庞加莱研究所年刊·概率与统计》，49(1)：1–12，2013年。网址： https://www.numdam.org/item/10.1214/11-AIHP462.pdf。[12] Cboe全球市场。《CBOE波动率指数：VIX [VIXCLS], 2025》。2025年12月31日每日收盘值：14.95。数据取自 FRED，圣路易斯联邦储备银行，2026年11月16日。网址： https://fred.stlouisfed.org/series/VIXCLS。

[13] J. C. Cox、S. A. Ross 和 M. Rubinstein。《期权定价：一种简化方法》。《金融经济学杂志》，7(3)：229–263，1979年。 doi:10.1016/0304-405X(79)90015-1。[14] M. Fathi、N. Gozlan 和 M. Prodhomme。《通过熵正则化证明 Caffarelli 收缩定理》，2019年。网址：https://arxiv.org/abs/1904.06053, arXiv:1904.06053。[15] N. Gigli 和 M. Ledoux。《从对数 Sobolev 到 Talagrand：一个快速证明》。《离散与连续动力系统》，33(5)：1927–1935，2013年。网址： https://www.math.univ-toulouse.fr/~ledoux/OttoVillani.pdf。[16] Z. Goldfeld 和 K. Greenewald。《高斯平滑最优传输：度量结构与统计效率》。载于 AISTATS 论文集，《机器学习研究论文集》第108卷，第3327–3337页，2020年。网址： https://proceedings.mlr.press/v108/goldfeld20a.html。[17] Z. Goldfeld、K. Kato、S. Nietert 和 G. Rioux。《平滑p-Wasserstein距离的极限分布理论》。《应用概率年刊》，34(2)：2447–2487，2024年。网址：https://arxiv.org/abs/2203.00159。[18] M. Hairer。《随机偏微分方程导论》，2025年。讲义，2025年11月10日版本，推论4.56。网址： https://hairer.org/notes/SPDEs.pdf。[19] S. Hou。《适应性平滑经验测度的收敛性》。《随机过程及其应用》，191：104775，2026年。网址： https://arxiv.org/abs/2401.14883,doi:10.1016/j.spa.2025.104775。[20] M. Larsson、J. Park 和 J. Wiesel。《平滑适应性 Wasserstein 距离的快速收敛率》，2026年。arXiv:2503.10827v2，2026年3月22日。网址： https://arxiv.org/abs/2503.10827。[21] M. Ledoux 和 M. Talagrand。《Banach 空间中的概率：等周不等式与随机过程》。施普林格，柏林，1991年。[22]S. Nietert、Z. Goldfeld 和 K. Kato。《平滑 p-Wasserstein 距离：结构、经验逼近与统计应用》。载于 ICML 论文集，《机器学习研究论文集》第139卷，第8172–8183页，2021年。网址：https://proceedings.mlr.press/v139/nietert21a.html。[23] F. Otto 和 C. Villani。《Talagrand 不等式的推广及其与对数 Sobolev 不等式的联系》。《泛函分析杂志》，173(2)：361–400，2000年。doi:10.1006/jfan.1999.3557。[24] G. Ch. Pflug。《多阶段随机优化中的版本无关性与嵌套分布》。《SIAM 优化杂志》，20(3)：1406–1420，2009年。doi:10.1137/080718401。[25] G. Ch. Pflug 和 A. Pichler。《多阶段随机优化模型的距离》。《SIAM 优化杂志》，22(1)：1–23，2012年。 doi:10.1137/110825054。[26] A. W. van der Vaart。《渐近统计学》。剑桥大学出版社，剑桥，1998年。[27] X. Wang、Z. Zhang 和 X. Zhou。《条件最优传输中相干性代价的曲率界》，2026年。arXiv:2610.09738。网址：https://arxiv.org/abs/2610.09738。