arXiv:2601.01589v1 [quant-ph] 4 Jan 2026

量子游走与欠阻尼朗之万动力学之间的学习关系

Yazhen Wang 威斯康星大学麦迪逊分校

2026年1月6日

摘要

人们始终需要快速的计算算法，而量子加速和经典加速等进展推动了这类算法的发展。本文旨在研究量子计算中基于量子游走的搜索算法，以及经典计算中基于朗之万动力学的采样算法。在量子领域，基于量子游走的搜索算法相较于其经典对应算法可实现二次加速。在经典计算领域，大量研究聚焦于梯度加速，其中，从欠阻尼朗之万动力学推导出的梯度调整算法，相较于传统朗之万算法可实现二次加速。

由于搜索算法和采样算法都旨在处理学习任务，我们研究了硬币量子游走与欠阻尼朗之万动力学之间的学习关系。具体而言，我们证明，就 Le Cam亏损距离而言，经过随机化的量子游走与欠阻尼朗之万动力学渐近等价，而未经随机化的量子游走由于其高频振荡行为，并不渐近等价。我们进一步讨论了这些等价性和非等价性结果对于机器学习任务中相关算法的计算性质和推断性质的影响。我们的研究结果为理解量子游走与欠阻尼朗之万动力学之间的关系，以及量子加速和经典梯度加速背后的内在机制，提供了新的见解。

关键词：量子算法、经典算法、朗之万扩散、Le Cam亏损距离、经典随机游走、量子游走、量子加速、经典加速。

1 引言

当前机器学习对计算能力的需求正逼近计算机技术的极限，因此人们投入大量精力，致力于开发涵盖芯片、软件和系统的快速计算技术。值得注意的是，如何在经典计算和量子计算中开发更快的算法，正日益受到关注，也吸引了大量研究。本文聚焦于这类研究中两个彼此独立的方向：经典方法和量子方法。经典方法是加速的基于梯度的优化方法，量子方法则是基于量子游走的搜索算法。这两种方法都能比各自对应的标准方法实现二次加速。关于量子计算，参见 [19] [20] [41][43] [54] ；关于经典计算，参见 [34] [39] [40] [49] [55] [56] 。本文从学习的角度研究这些以往互不相关的经典与量子算法之间的关系，并探明将这些算法应用于机器学习中的推断任务时，它们之间的学习等价性与非等价性。

在量子计算中，量子游走搜索是一种量子算法，旨在通过量子游走在图中搜索标记节点。量子游走的概念源自经典随机游走，在经典随机游走中，行走者会在图或格中随机移动。在经典随机游走中，行走者的位置由图中各节点上的概率分布描述。相比之下，量子游走将行走者表示为量子态，该量子态可以同时处于多个位置的叠加态中。这些算法的计算效率和成功概率会受到搜索空间结构的强烈影响。通常，与其经典对应算法相比，量子游走搜索算法可实现平方级加速。例如，Grover 算法——一种用于无结构搜索的知名量子搜索算法——可以以 $O(\sqrt{N})$ 阶复杂度找到解，其中 N 表示搜索问题的规模。在量子计算中，解决同一问题需要 $O(N)$ 阶复杂度。此外， $O(N)$ 阶和 $O(\sqrt{N})$ 阶的计算复杂度结果代表了解决该

搜索问题的理论极限。具体而言，任何量子算法至少必须执行 $O(\sqrt{N})$ 次操作，而经典算法无法以少于 $O(N)$ 次操作解决该问题。因此，Grover 算法在渐近意义下是最优的，相较于最佳经典搜索算法可提供平方级加速。详情请参见 [19] [20] [35] [43] [45] [46]。

在经典计算的背景下，Nesterov 于 1983年提出的 Nesterov 加速梯度下降法（见 [40]）是一种旨在最小化凸目标函数的一阶算法。该方法设计了一种巧妙的算法方案，引入一个称为动量的额外变量，从而加速传统的梯度下降。具体而言，对于凸函数，Nesterov 方法的收敛速度为 $O(k^{-2})$，其中k 表示迭代次数；相比之下，传统梯度下降的收敛速度为 $O(k^{-1})$。加速的概念已成为基于梯度的优化方法的核心。令步长趋近于零，Nesterov 加速梯度法便导出一个二阶常微分方程（ODE），称为 Nesterov 加速梯度流或 Nesterov 常微分方程。详见 [39] [40] [49] [55] [56] 。

Jordan、Kinderlehrer 和 Otto 于 1998年发表的一项奠基性工作（见 [24]）确立了这样一个结果：朗之万动力学的福克–普朗克方程对应于概率测度空间中相对熵泛函的梯度流，该空间配备了 Wasserstein-2 度量；这一梯度流称为 Wasserstein 梯度流。这一发现通过在 Wasserstein-2 空间中研究优化问题，建立了采样与优化领域之间的联系。近期研究将 Wasserstein 梯度流推广至 Nesterov 加速梯度流，探究了朗之万动力学采样中的加速。朗之万动力学的一种加速版本称为欠阻尼朗之万动力学，它与 Nesterov 加速梯度流具有重要的相似之处。欠阻尼朗之万动力学结合了一个确定性分量（与 Nesterov 加速梯度流直接对应）和一个随机分量。这种方法使加速朗之万动力学的采样速度比传统朗之万动力学快一个二次量级。详见 [3] [10] [11][24] [34] [42] [61] 。

计算任务和推断任务是机器学习的两个关键组成部分。在计算方面，欠阻尼朗之万动力学和量子游走都能相较于各自对应的标准方法实现二次加速。本文使用 Le Cam 缺陷距离，从推断学习的视角探究二者之间的关系。具体而言，

我们将基于量子游走和欠阻尼朗之万动力学的学习算法（或方法）置于 Le Cam 范式中，并根据 Le Cam 缺陷距离建立它们的渐近等价性和非等价性。这里的等价性是指，与量子游走和欠阻尼朗之万动力学相关的两个学习模型中，一个模型的每个学习过程都有另一个模型中性能相同的对应学习过程，而渐近等价性则表示在渐近意义下具有相同性能。详情请见 [32] [52] [53] 。

我们对学习关系的研究或许能为经典计算中的加速现象和量子计算中的量子加速带来新的启示，并促使我们探索经典加速与量子加速之间是否存在内在联系。一维量子游走有对应的一维经典随机游走；然而，量子游走的混合时间比其经典对应算法快一个二次方。我们可以将经典随机游走提升为二维经典随机游走，使其在混合时间方面与量子游走相匹配。

通过这一比较，我们发现了若干相似之处。首先，与各自的经典对应算法——即经典随机游走和标准（过阻尼）朗之万动力学——相比，量子游走和欠阻尼朗之万动力学的混合时间都实现了二次加速。具体而言，量子游走的混合时间可能比相应经典随机游走快一个二次方，而基于欠阻尼朗之万动力学的算法，其混合时间也可能比基于标准朗之万动力学的算法快一个二次方。

其次，提升机制是两种情形中观察到二次加速的基础。在量子领域，量子游走可被解释为对应于一类经典提升马尔可夫链，它能加快混合速度，从而降低计算复杂度。在经典领域，由于 Nesterov 梯度流由二阶常微分方程（ODE）支配，欠阻尼朗之万动力学可视为二阶随机微分方程（SDE）。它将标准朗之万动力学从一维空间提升到二维过程，该过程由原有的位置变量和一个额外的动量变量组成，从而得到一个二维随机微分方程。

我们希望这一研究方向能推动人们进一步探究量子加速与经典加速之间的联系，并深入探索两者共有的机制和基本原理。

本文其余部分安排如下。第2节首先简要回顾量子游走和朗之万动力学，随后介绍量子游走与欠阻尼朗之万动力学渐近等价和非等价的主要结果，并讨论其意义。第3节给出总结性论述，附录则汇集了所有技术性证明。

2 量子游走与朗之万动力学的主要结果

2.1 量子游走

由于本文聚焦于实数轴上的带硬币离散时间量子游走，我们在此简要回顾其框架。这类量子游走由一个行走者、一个硬币、分别作用于行走者和硬币的演化算子，以及一组可观测量构成。

行走者是一个量子系统，存在于一个维数无限但可数的希尔伯特空间 $\mathcal{H}_p$ 中。通常使用 $\mathcal{H}_p$的标准计算基来表示行走者的位置，其基向量对应于离散的位置点。我们用$|\text{position}\rangle \in \mathcal{H}_p$ 表示行走者的状态，并注意到标准基态 $\{|i\rangle_p\}$ 张成 $\mathcal{H}_p$。因此，任何满足 $\sum_i |\alpha_i|^2 = 1$的 $\sum_i \alpha_i |i\rangle_p$形式的叠加态，都是行走者的合法状态。行走者通常初始化于原点，即 $|\text{position}\rangle_{\text{initial}} = |0\rangle_p$。

硬币是一个存在于二维希尔伯特空间$\mathcal{H}_c$中的量子系统。它可以取两个标准基态 $|0\rangle_c$ 和 $|1\rangle_c$中的任意一个，也可以是这些基态的任意叠加态。因此，硬币态 $|\text{coin}\rangle \in \mathcal{H}_c$，一般的归一化硬币态可写为 $|\text{coin}\rangle = a_0|0\rangle_c + a_1|1\rangle_c$，其中 $|a_0|^2 + |a_1|^2 = 1$。

量子游走的总态存在于 $\mathcal{H} = \mathcal{H}_p \otimes \mathcal{H}_c$中。通常选择 $\mathcal{H}$ 中的乘积态作为初始态；也就是说，$|\psi\rangle_{\text{initial}} = |\text{position}\rangle_{\text{initial}} \otimes |\text{coin}\rangle_{\text{initial}}$。

量子游走的演化由作用于硬币和行走者的两个算子决定。每一步中，首先将硬币算子作用于硬币态，随后将条件位移算子作用于整个复合系统。硬币算子的作用是变换硬币态，而随机性则通过对整个

系统反复应用这两个演化算子后进行测量而引入。

在通常记作 $\mathbf{C}$ 的硬币算子中，$\mathbf{H}$Hadamard 算子被广泛使用：

$$\mathbf{H} = \frac{1}{\sqrt{2}} \left(|0\rangle_c\langle 0| + |0\rangle_c\langle 1| + |1\rangle_c\langle 0| - |1\rangle_c\langle 1| \right). \tag{1}$$

对于条件位移算子，采用一个酉算子：当对应的硬币态是其中一个基态（例如 $|0\rangle$）时，该算子使行走者向前移动一步；当硬币态是另一个基态（例如 $|1\rangle$）时，则向后移动一步。一个合适的条件位移算子为

$$\mathbf{S} = |0\rangle_c\langle 0| \otimes \sum_j |j+1\rangle_p\langle j| + |1\rangle_c\langle 1| \otimes \sum_j |j-1\rangle_p\langle j|. \tag{2}$$

因此，作用于总希尔伯特空间的演化算子表示为 $\mathbf{U} = \mathbf{S} \cdot (\mathbf{C} \otimes \mathbf{I}_p)$，其中 $\mathbf{I}_p$ 表示 $\mathcal{H}_p$ 上的恒等算子。于是，量子游走在经过 t 步后的状态可简洁地表示为

$$|\psi\rangle_t = \mathbf{U}^t |\psi\rangle_{\text{initial}},$$

其中 $|\psi\rangle_{\text{初始}} = |\text{位置}\rangle_{\text{初始}} \otimes |\text{硬币}\rangle_{\text{初始}}$。如果采用的两个演化算子分别是上述 Hadamard 算子 $\mathbf{H}$ 和条件位移算子 $\mathbf{S}$，则该量子行走者称为 Hadamard 行走。

要进行测量并获得量子游走的结果，需要使用可观测量。我们根据用于描述 coin 和行走者的基态定义一组可观测量。从复合量子系统中提取信息有多种方法。例如，可以先使用可观测量对 coin 进行测量

$$\mathbf{M}_c = a_0^w |0\rangle_c\langle 0| + a_1^w |1\rangle_c\langle 1|.$$

随后，使用算符对行走者的位置态进行测量

$$\mathbf{M}_p = \sum_i \alpha_i^w |i\rangle_p\langle i|.$$

我们可以将状态 $|\psi\rangle_t$ 表示为

$$|\psi\rangle_t = \sum_k \left[b_{kt}|0\rangle_c + \beta_{kt}|1\rangle_c \right] |k\rangle_p,$$

其中， $|0\rangle_c$ 和 $|1\rangle_c$ 表示硬币态分量， $|k\rangle_p$ 表示位置态分量。例如，假设量子游走初始化为状态

$$|\psi\rangle_0 = |0\rangle_c \otimes |0\rangle_p,$$

其中，方程(1)和(2)分别定义了硬币算子和移位算子。在这种情况下，可以推导出 b_{kt} 和 β_{kt}的解析表达式，以及在第 t步进行测量时，游走者位于位置 k 的概率。

令 $\mathbb{C}$ 表示所有复数的集合，并令 $\varphi = (a_0, a_1)' \in \mathbb{C}^2$，其中 $|a_0|^2 + |a_1|^2 = 1$，表示一维带硬币量子游走的初始硬币态。令 $\mathbf{X}_t^q$ 表示时刻 t的一维带硬币量子游走，该游走从初始硬币态 φ 出发，并在由 2×2 酉矩阵 $\mathbf{U}$描述的硬币算子 $\mathbf{C}$ 作用下演化。

该量子游走的标准差呈线性增长，阶为 $O(t)$，混合时间阶为 $O(t)$；这与相应的经典随机游走形成鲜明对比，后者的标准差阶为 $O(\sqrt{t})$，混合时间阶为 $O(t^2)$。更多详情请参见 [1] [4] [26] [27] [35] [43] [45] [46] [51] [57]。

2.2 朗之万动力学

欠阻尼朗之万动力学由以下随机微分方程定义 ,

$$\begin{cases} dY_t^c = V_t^c dt, & (Y_0^c, V_0^c) \sim \rho_0, \\ dV_t^c = -V_t^c dt - \nabla\Psi(Y_t^c)dt + \sqrt{2}dB_t, \end{cases} \tag{3}$$

其中， B_t 表示标准的一维布朗运动， $\Psi(\cdot)$是势函数， ∇ 表示梯度算子，而 ρ_0 是初始概率密度函数（pdf）。该随机微分方程的解（Y_t^c， V_t^c）是一个扩散过程，其唯一的不变分布正比于 $\exp(-\Psi(x) - v^2/2)$。因此， Y_t^c的边缘不变分布正比于 $\exp(-\Psi(x))$。基于欠阻尼朗之万方程（3）的离散化而得的算法，与从标准（或过阻尼）朗之万动力学推导出的传统朗之万算法相比，能以二次加速实现从不变分布中采样。更多详情见 [7] [10] [15] [42] [61] 。

2.3 学习等价性与非等价性

通过对标准朗之万动力学进行离散化，可以得到经典随机游走；而标准朗之万动力学本身也可视为经典随机游走的连续类比。事实上，就数学近似和统计学习而言，经典随机游走与标准朗之万动力学渐近等价。量子游走和欠阻尼朗之万动力学的混合时间分别比其对应方法——经典随机游走和标准朗之万动力学——快一个二次因子；基于它们的算法，相较于基于经典随机游走和标准朗之万动力学的算法，也能实现二次加速。

此外，量子游走和欠阻尼朗之万动力学都可以通过适当的提升方案与各自的经典对应算法联系起来。具体而言，量子游走对应于提升后的经典随机游走，而欠阻尼朗之万动力学则可视为标准朗之万动力学的提升版本。这自然引出一个问题：从学习或算法的角度来看，量子游走与欠阻尼朗之万动力学之间是否存在更深层的联系?

如第 2.1 节所述，$\mathbf{X}_t^q$ 表示时刻t 的带硬币量子游走，其初始硬币态为 φ ，演化算子由酉矩阵 $\mathbf{U}$给出，如第 2.1 节所述，其中 $\theta = (\varphi, \mathbf{U})$,

$$\varphi = \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} \in \mathbb{C}^2 \text{ and } \mathbf{U} = \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} \in \mathbb{C}^{2\times 2}$$

满足 $|a_0|^2+|a_1|^2 = 1$，$\mathbf{U}^\dagger\mathbf{U} = \mathbf{I}_2$, 和 $u_{11}u_{12}u_{21}u_{22} \neq 0$。这里 $*$ 和 $\dagger$ 分别表示共轭和共轭转置，$|a|^2 = a^*a$，而 $\mathbf{I}_2$ 表示二维单位矩阵。令 X_t^q 表示 $\mathbf{X}_t^q$ 的位置测量结果。对于任意小的 $\eta > 0$，我们定义 $X_t^q(\eta)$为一个取值为 $X_t^q + j,\ j = 0, \pm 1, \cdots, \pm(r-1)$ r的随机变量，其概率为 $1/(2r)$，其中 $r = [\eta t]$。记 $\mathbb{P}_{t,\theta}^q$ 和 $\mathbb{P}_{t,\theta,\eta}^q$ 分别为 X_t^q 和 $X_t^q(\eta)$的概率分布。我们将 $X_t^q(\eta)$称为 X_t^q 的一种随机化。随机化是一种广泛应用于各领域的统计技术，尤其用于研究和科学应用，以增强统计学和机器学习方法的有效性。

记$\mathbb{P}_{t,\theta,\epsilon}^c$为 $X_t^c(\epsilon) = t|u_{11}|e^{-|Y_t^c|}sign(Y_t^c)$的概率分布，其中经典过程$Y_t^c$由带势函数 $\Psi_\epsilon(y) = -\log \pi_\epsilon(y;\theta)$的欠阻尼朗之万动力学（3）所支配；$\epsilon$ 是任意小的正常数，$\pi_\epsilon(y;\theta) = \pi(y;\theta)$，对于 $|y| \geq -\log(1-\epsilon)$

且 $\pi_\epsilon(y;\,\theta) = \varkappa_\epsilon$ 对于 $|y| < -\log(1-\epsilon)$,

$$\begin{aligned}\varkappa_\epsilon &= \frac{1}{-2\log(1-\epsilon)}\left[1 - \int_{|y|\geq -\log(1-\epsilon)} \pi_\epsilon(y;\theta)dy\right] \\ &= \frac{1}{-2\log(1-\epsilon)}\int_{\log(1-\epsilon)}^{-\log(1-\epsilon)} \pi(y;\theta)dy, \end{aligned} \tag{4}$$

且 $\pi(y;\theta)$ 是按如下方式为 $y \neq 0$ 定义的概率密度函数,

$$\begin{aligned}&\pi(y;\theta) = \frac{e^{-|y|}\sqrt{1-|u_{11}|^2}}{\pi(1-|u_{11}|^2e^{-2|y|})\sqrt{1-e^{-2|y|}}} \\ &\left\{1 - \left(|a_0|^2 - |a_1|^2 + \frac{a_0a_1^*u_{11}u_{12}^* + a_0^*a_1u_{11}^*u_{12}}{|u_{11}|^2}\right)|u_{11}|e^{-|y|}\mathrm{sign}(y)\right\}. \end{aligned} \tag{5}$$

这里，sign(y) 表示符号函数。

设 $(\Omega_q, \mathcal{F}_q)$ 和 $(\Omega_c, \mathcal{F}_c)$ 分别为 $\mathbb{P}_{t,\theta,\eta}^q$ 和 $\mathbb{P}_{t,\theta,\epsilon}^c$ 所在的可测空间。我们定义一个马尔可夫核 $\mathcal{K}(\cdot,\cdot)$，将 $(\Omega_c, \mathcal{F}_c)$ 上的 $\mathbb{P}_{t,\theta,\epsilon}^c$ 映射为 $(\Omega_q, \mathcal{F}_q)$ 上的另一个概率分布，以便将映射后的 $\mathbb{P}_{t,\theta,\epsilon}^c$ 与 $\mathbb{P}_{t,\theta,\eta}^q$ 进行比较。具体而言，对于给定的 $\omega \in \Omega_c$，$\mathcal{K}(\omega,\cdot)$ 是定义在 σ 域 $\mathcal{F}_q$ 上的概率测度；对于固定的 $A \in \mathcal{F}_q$，$\mathcal{K}(\cdot, A)$ 是 Ω_c 上的可测函数。马尔可夫核 $\mathcal{K}(\omega, A)$ 对 $\omega \in \Omega_c$ 和 $A \in \mathcal{F}_q$ 有定义，并将 $\mathbb{P}_{t,\theta,\epsilon}^c$ 映射为一个概率测度，记作 $(\Omega_q, \mathcal{F}_q)$ 上的 $\mathcal{K}(\mathbb{P}_{t,\theta,\epsilon}^c)$，如下所示：对于任意 $A \in \mathcal{F}_q$，我们定义

$$[\mathcal{K}(\mathbb{P}_{t,\theta,\epsilon}^c)](A) = \int_{\Omega_c} \mathcal{K}(\omega, A)\mathbb{P}_{t,\theta,\epsilon}^c(d\omega).$$

$X_t^c(\epsilon) \sim \mathbb{P}_{t,\theta,\epsilon}^c$ 相对于 $X_t^q(\eta) \sim \mathbb{P}_{t,\theta,\eta}^q$ 的亏损定义为

$$\beth(X_t^c(\epsilon), X_t^q(\eta)) = \inf_{\mathcal{K}} \sup_{\theta\in\Theta} \|\mathbb{P}_{t,\theta,\eta}^q - \mathcal{K}(\mathbb{P}_{t,\theta,\epsilon}^c)\|_{TV},$$

其中，$\|\cdot\|_{TV}$ 是全变差范数，下确界取遍所有马尔可夫核，上确界取遍某个参数空间 Θ。类似地，可以定义 $X_t^q(\eta)$ 相对于 $X_t^c(\epsilon)$ 的亏损 $\beth(X_t^q(\eta)\ X_t^c(\epsilon))$。我们将 $X_t^q(\eta)$ 与 $X_t^c(\epsilon)$ 之间的 Le Cam亏损距离定义为

$$\mho(X_t^q(\eta), X_t^c(\epsilon)) = \max\{\beth(X_t^q(\eta), X_t^c(\epsilon)), \beth(X_t^c(\epsilon), X_t^q(\eta))\}.$$

此外，利用 $\mathbb{P}_{t,\theta}^q$ X_t^q 的分布，我们可以定义 $\beth(X_t^c(\epsilon)\ X_t^q)$、$\beth(X_t^q,\ X_t^c(\epsilon))$ 和 $\mho(X_t^q,\ X_t^c(\epsilon))$（见 [32] [52] [53]）。

我们需要固定一些记号，并为理论分析定义参数空间 Θ 。我们将 φ 和 $\mathbf{U}$重写为

$$\varphi = \begin{pmatrix} e^{i\varsigma_1}\cos\vartheta \\ e^{i\varsigma_2}\sin\vartheta \end{pmatrix}, \mathbf{U} = \begin{pmatrix} e^{i\gamma_1}\cos\phi & e^{i\gamma_2}\sin\phi \\ e^{-i\gamma_2}\sin\phi & -e^{-i\gamma_1}\cos\phi \end{pmatrix}, \tag{6}$$

其中 ς_1， ς_2， γ_1， $\gamma_2 \in [-\pi/2, \pi/2]$ 和 ϑ， $\phi \in [0, \pi/2]$。定义

$$\Theta = \{\theta = (\varphi, \mathbf{U}) : \varphi \text{ and } \mathbf{U} \text{ are given in (6) with } \phi \in [\kappa, \pi/2 - \kappa]\}, \tag{7}$$

其中 κ 是一个较小的正常数。

我们有如下关于量子行走与欠阻尼朗之万动力学之间勒卡姆距离的定理。

定理1. 在条件 *(7)* 下，我们有

$$\limsup_{t\to\infty} \mho(X_t^q(\eta), X_t^c(\epsilon)) = O(\epsilon^{1/2} + \eta^{1/2}), \tag{8}$$

当 ϵ 和 $\eta \to 0$ 时，该值趋于零。此外，我们可以证明

$$\liminf_{t\to\infty} \mho(X_t^q, X_t^c(\epsilon)) \geq c_0 - O(\epsilon^{1/2}), \tag{9}$$

当 $\epsilon \to 0$ 时，该值以下界 c_0 为界，其中 $c_0 > 0$ 是一个一般常数。

考虑由基于量子游走或欠阻尼朗之万动力学的学习算法或程序所解决的推断学习任务。与标准机器学习框架一样，我们指定一个损失函数，并通过学习过程或算法所对应的风险值来评估其性能。基于量子游走和欠阻尼朗之万动力学的学习过程之间的性能关系，可以用 Le Cam 亏损距离来刻画。

具体而言，对于两个数据生成过程，如果它们的勒卡姆距离为零，那么基于其中一个过程的每个学习算法，都存在一个基于另一个过程且性能完全相同的对应算法，反之亦然。如果它们的勒卡姆距离被任意小的常数 δ从上界定，那么基于其中一个过程的每个学习过程或算法，都存在一个基于另一个过程的相应过程，其性能差异至多为δ。此外，如果勒卡姆距离随着时间趋于无穷而收敛于零，那么对于基于这两个过程之一的任意学习算法，

都存在一个基于另一个过程的相应算法，其性能差异收敛于零。更多详情请参见 [32] [52] [53] 。

如定理1所示，渐近结果（8）意味着，当 t 足够大且 η 和 ϵ 足够小时，$\mho(X_t^q(\eta)\ X_t^c(\epsilon))$ 可以任意小。这一结果表明，任何基于 $X_t^q(\eta)$ 的学习算法，都有一个基于 $X_t^c(\epsilon)$ 的对应算法，其性能差异可以任意小。更具体地说，对于任意 $\delta > 0$，定理1保证，对于某个足够大的 t 以及适当选择的较小 η 和 ϵ，$\mho(X_t^q(\eta)\ X_t^c(\epsilon)) \leq \delta$。因此，对于任何基于 $X_t^q(\eta)$ 或 $X_t^c(\epsilon)$ 的学习过程或算法，都存在一个基于另一方的对应过程，其性能差距不超过 δ。此外，这一结果可以轻松推广到多个样本。

尽管量子游走和欠阻尼朗之万动力学是独立发展的，但这一渐近等价性在这两个此前无关的框架之间建立了严格联系，并为基于量子游走和欠阻尼朗之万动力学的学习算法或过程带来了新的见解。量子游走和欠阻尼朗之万动力学分别在量子计算中快速量子算法以及经典计算中加速梯度算法的发展过程中发挥着核心作用，因此，我们的结果或将推动对量子加速和经典加速的进一步研究，并促进对二者内在联系和底层机制的深入探索。下文第2.4节将介绍与这些主题相关的其他重点和细节。

定理1中建立的结果（9）表明，随机化对于在量子游走与欠阻尼朗之万动力学之间建立任何学习等价性都是必不可少的。这一现象可以启发式地解释如下。对于量子游走而言，测量 X_t^q 呈现出一种渐近分布行为，其特征是在缓慢变化的全局分量之上叠加高频局部振荡。通过适当的卷积核，可以用量子游走逼近欠阻尼朗之万动力学：量子游走中缓慢变化的分量可以由欠阻尼朗之万动力学刻画，而振荡分量的频率与 $1/t$成正比，难以由任何连续时间扩散过程在渐近意义下加以逼近。事实上，（8）意味着

$$\limsup_{t\to\infty} \beth(X_t^q, X_t^c(\epsilon)) = O(\epsilon^{1/2}). \tag{10}$$

而（9）的证明实际上已经表明

$$\liminf_{t\to\infty} \mho(X_t^q, X_t^c(\epsilon)) = \liminf_{t\to\infty} \beth(X_t^c(\epsilon), X_t^q) \geq c_0 - O(\epsilon^{1/2}). \tag{11}$$

数值上，我们估计 $c_0 \approx 0.3$，尽管证明中推导出的理论下界要小得多。(10)表明，欠阻尼朗之万动力学可以通过由马尔可夫核诱导的变换来近似为量子游走，而(11)则表明反向近似是不可能的。(11)的证明进一步揭示，这一限制源于量子游走的高频振荡分量。因此，通过 $X_t^q(\eta)$ 引入随机化，对于建立量子游走与欠阻尼朗之万动力学之间的学习等价性似乎至关重要。此外，由于随机化过程 $X_t^q(\eta)$ 在 $\eta \to 0$ 时趋近于 X_t^q，我们可以允许 η 依赖于 t，前提是 $\eta t \to \infty$。也就是说，为确保随机化渐近地持续存在，η 必须以远慢于 $1/t$ 的速率趋近于零。否则，若 η 下降得足够快，使得 $[\eta t] < 1$，那么 $X_t^q(\eta)$ 就与 X_t^q 重合，不会发生随机化。

在我们的表述中引入 ϵ，是出于内在的技术考量。具体而言，式（5）中的概率密度函数 $\pi(y;\theta)$ 在 y 趋近原点时会发散；这一奇异性源于量子游走 X_t^q 的渐近行为。这种发散令人担忧：相应的欠阻尼朗之万动力学（3）是否定义良好。由于$\epsilon \to 0$，正则化概率密度函数 $\pi_\epsilon(y;\theta)$ 收敛于 $\pi(y;\theta)$，因此或许可以将 X_t^c 定义为 $X_t^c(\epsilon)$ 的一个适当极限。例如，可以按如下方式定义一个分布 $\mathbb{P}^c_{t,\,\theta}$：对于实数轴上的任意 x，

$$\mathbb{P}^c_{t,\theta}((-\infty, x]) = \limsup_{\epsilon\to 0} \mathbb{P}^c_{t,\theta,\epsilon}((-\infty, x]),$$

并将 X_t^c 定义为 $\epsilon \to 0$ 时 $X_t^c(\epsilon)$ 的适当极限，具体而言，对于实数轴上的任意博雷尔集 A，都有$Pr(X_t^c \in A) = \mathbb{P}^c_{t,\,\theta}(A)$。根据此定义，$X_t^c$ 可视为势函数为 $\Psi(y) = -\log \pi(y;\theta)$ 的欠阻尼朗之万动力学 (3)。如果 $\mathbb{P}^c_{t,\,\theta}$ 和 X_t^c均有定义，我们可以证明

$$\limsup_{t\to\infty} \mho(X_t^q(\eta), X_t^c) = O(\eta^{1/2}), \tag{12}$$

and

$$\limsup_{t\to\infty} \beth(X_t^q, X_t^c) = 0, \text{ but } \liminf_{t\to\infty} \beth(X_t^c, X_t^q) \geq c_0 > 0, \tag{13}$$

其中 c_0 与 (9) 中的一般常数相同。然而，在定义 $\mathbb{P}^c_{t,\ \theta}$ 和 X^c_t 时所用的收敛未必是弱收敛。因此，值得探究 X^c_t 是否为 $\epsilon \to 0$ 时 $X^c_t(\epsilon)$ 的弱极限。由于 $\pi(y;\theta)$ 中存在奇异性，要证明$X^c_t(\epsilon)$ 的随机等度连续性颇具挑战；我们猜想，当 $\epsilon \to 0$ 时， $X^c_t(\epsilon)$ 不存在弱极限。

2.4 等价性结果的含义

定理1中建立的渐近等价性，可能对经典计算中的加速现象和量子计算中的量子加速具有重要意义。量子游走和欠阻尼朗之万动力学都能够分别比其经典对应算法实现快一个二次方的混合时间。在这两种情况下，这种二次提升都与某种潜在的提升机制相关，甚至正是由该机制实现的。

在量子方面，相对于其经典模拟——马尔可夫链，量子游走可以将混合时间实现二次加速。此外，还可以对经典马尔可夫链进行提升，使所得的提升后经典马尔可夫链在混合时间方面与量子游走相匹配。提升是加速马尔可夫链蒙特卡罗算法的一种成熟框架。具体而言，经典马尔可夫链 $\mathcal{M}$ 的提升，是通过将原链 $\mathcal{M}$ 的每个状态替换为一组状态，并定义转移概率以保留 $\mathcal{M}$ 的结构，从而构造出的更大的经典马尔可夫链 $\mathcal{M}$ 。特别地，对提升后链 $\mathcal{M}$ 中适当的状态进行合并（或投影）后，可以恢复原链 $\mathcal{M}$，且 $\mathcal{M}$ 的不变分布恰好投影为 $\mathcal{M}$ 的不变分布。这类提升方案能够使混合时间实现二次加速，从而显著降低马尔可夫链蒙特卡罗算法的计算成本。值得注意的是，提升后链的构造引入了动量这一概念，以实现这种加速。在混合时间方面与提升后链相对应的量子游走，也能相对于原链实现二次加速。因此，探索量子二次提升与提升机制之间的联系颇具意义。有关提升的更多细节，请参见 [5] [8] [13] [14] [21] [22] [25]。

在经典领域，加速的概念已成为基于梯度的优化方法的核心。作为一个二阶常微分方程，Nesterov 梯度流在离散优化方法与连续时间动力学之间架起了概念上的桥梁。作为 Nes-

terov 梯度流的随机类比，欠阻尼朗之万动力学由一个二阶随机微分方程描述。通过将标准（或过阻尼）朗之万动力学从一维空间提升为二维过程，可以自然地构造出欠阻尼朗之万动力学。标准朗之万动力学由以下随机微分方程描述：

$$dY_t = -\nabla\Psi(Y_t)dt + \sqrt{2}\,dB_t,$$

其中，$(\nabla, \Psi(\cdot)\ B_t)$ 的定义与式 (3) 中相同。欠阻尼朗之万动力学通过在传统的位置变量之外引入一个额外的动量变量，对这一表述进行了扩展，从而得到式 (3) 所示的二维随机微分方程。更多详情请参阅 [16] [17] [33] [37] [42] [47] [50] [58][59] [60]。

基于朗之万动力学的采样算法可视为在概率测度空间上求解优化问题的一种梯度下降形式。在这一框架下，基于梯度的朗之万算法处于采样与优化的交界处：它们旨在从不变分布中抽取样本，同时在 Wasserstein-2 空间上求解优化问题，其中 Kullback–Leibler 散度充当目标泛函。这种优化视角自然促使人们研究由朗之万动力学衍生的算法中的加速。

加速这一概念在基于梯度的优化中发挥着关键作用，Nesterov 加速梯度下降便是一个典型例子。在朗之万动力学中，加速算法指的是对欠阻尼朗之万动力学 (3) 进行离散化，其收敛速度较传统朗之万算法实现了二次提升；传统朗之万算法对应于对标准（过阻尼）朗之万扩散的直接离散化。事实上，若从 (3) 中去掉布朗运动项，所得常微分方程便对应于 Nesterov 加速梯度流。作为朗之万动力学的一种加速变体，欠阻尼朗之万动力学与 Nesterov 加速梯度流具有关键的结构相似性，并且与标准朗之万动力学相比，能够实现快一个二次量级的采样速度。更多详情请参见 [3] [6] [7] [9] [10][11] [12] [15] [23] [24] [31] [34] [42] [61] 。

我们已经证明，量子游走与欠阻尼朗之万动力学在结构上具有许多相似之处。正如标准朗之万动力学可以被视为马尔可夫链的连续类比一样，二者

量子游走和欠阻尼朗之万动力学分别都可以与提升方案联系起来；就混合时间和算法运行时间而言，这些方案相较于马尔可夫链和标准朗之万动力学分别实现了二次提升。量子游走与欠阻尼朗之万动力学之间所建立的渐近等价，预计将推动人们进一步研究量子加速和经典加速，并深入探索这些现象之间的内在联系，以及促成它们的共同底层机制。

3 结语

基于量子游走的搜索算法能够相较于其经典对应算法实现二次加速，而源自欠阻尼朗之万动力学的梯度调整采样算法能够使传统朗之万算法实现二次加速。由于搜索和采样算法都旨在解决学习任务，我们研究了带硬币量子游走与欠阻尼朗之万动力学之间的学习关系。

我们根据 Le Cam 缺失距离，建立了欠阻尼朗之万动力学与带随机化的量子游走之间的渐近等价性。然而，由于量子游走的位置测量在分布行为中存在高频振荡，我们证明了不带随机化的量子游走与欠阻尼朗之万动力学并不渐近等价。我们进一步讨论了这些等价性与非等价性结果对机器学习任务中相关算法的计算和推断方面的影响。

总体而言，我们的研究结果为量子游走与欠阻尼朗之万动力学之间的关系，以及量子加速和经典梯度加速背后的内在机制，提供了新的见解。我们希望这些结果能推动量子计算和经典计算领域的算法开发研究进一步发展。

致谢

Yazhen Wang 的研究部分得到了美国国家科学基金会（NSF）DMS-2514240 项目的资助。

参考文献

[1] A. Ambainis. 量子搜索算法。*SIGACT News*，35(2):22?35，2004年6月。[2] A. Ambainis、E. Bach、A. Nayak、A. Vishwanath 和 J. Watrous。一维量子游走。载于第三十三届 *ACM*计算理论研讨会论文集，STOC ’01，第37–49页，纽约，美国，2001年。计算机协会。[3] L. Ambrosio、N. Gigli 和 G. Savaré。梯度流：度量空间与概率测度空间中的梯度流。Springer Science & Business Media，2008年。[4] S. Apers、A. Sarlette 和 F. Ticozzi。量子游走模拟以及利用经典过程实现快速混合。*Physical Review A*，98:032115，2018年9月。[5] S. Apers、F. Ticozzi 和 A. Sarlette。提升马尔可夫链以加快混合：局限与机遇。*arXiv*预印本 *arXiv:1705.08253v1*，2017年。[6] S. Calogero。动理学 Fokker-Planck 方程向平衡态的指数收敛。*Communications in Partial Differential Equations*，37:1357–1390，2010年。[7] Y. Cao、J. Lu 和 L. Wang。欠阻尼朗之万动力学显式 L^2-收敛率估计。*Archive for Rational Mechanicsand Analysis*，247(5):90，2023年。[8] F. Chen、L. Lovász 和 I. Pak。提升马尔可夫链以加快混合。载于第三十一届 *ACM* 计算理论研讨会论文集，STOC ’99，第275–281页，纽约，美国，1999年。计算机协会。

[9] X. Cheng 和 P. Bartlett。Langevin 马尔可夫链蒙特卡罗在 KL 散度下的收敛性。载于 F. Janoos、M. Mohri 和 K. Sridharan 编，*算法学习理论论文集*，*机器学习研究论文集*第83卷，第186–211页。PMLR，2018年4月7–9日。[10] X. Cheng、N. S. Chatterji、P. L. Bartlett 和 M. I. Jordan。欠阻尼 Langevin 马尔可夫链蒙特卡罗：非渐近分析。载于 S. Bubeck、V. Perchet 和 P. Rigollet 编，第*31*届学习理论会议论文集，*机器学习研究论文集*第75卷，第300–323页。PMLR，2018年7月6–9日。[11] X. Cheng、D. Yin、P. Bartlett 和 M. Jordan。随机梯度与 Langevin 过程。载于第*37*届国际机器学习会议论文集，ICML'20。JMLR.org，2020年。[12] S. Chewi、M. A. Erdogdu、M. B. Li、R. Shen 和 M. Zhang。从 Poincaré 到对数 Sobolev 的 Langevin 蒙特卡罗分析。*arXiv*预印本 *arXiv:2112.12662*，2021年。[13] D. Dervovic。对每个量子游走，都存在一个混合时间更短的（经典）提升马尔可夫链。*arXiv*预印本 *arXiv:1712.02318v2*，2018年。[14] P. Diaconis、S. Holmes 和 R. M. Neal。不可逆马尔可夫链采样器分析。*应用概率年刊*，10(3):726–752，2000年。[15]A. Eberle、A. Guillin 和 R. Zimmer。朗之万动力学的耦合与定量收缩率。*The Annals of Probability*，47(4):1982–2010，2019年。[16] J. Fan、Q. Zhang、A. Taghvaei 和 Y. Chen。变分 Wasserstein 梯度流。载于 K. Chaudhuri、S. Jegelka、L. Song、C. Szepesvari、G. Niu 和 S. Sabato 编，第*39*届国际机器学习会议论文集，*机器学习研究论文集*第162卷，第6185–6215页。PMLR，2022年7月。[17]C. W. Gardiner 等。随机方法手册，第3卷。Springer Berlin，1985年。[18] G. Grimmett、S. Janson 和 P. F. Scudo。量子随机游走的弱极限。*Physical Review E*，69:026119，2004年2月。

[19] L. K. Grover。《style id='2'>第二十八届 ACM计算理论研讨会年会论文集：一种用于数据库搜索的快速量子力学算法》，第212–219页。ACM，1996年。[20] L. K. Grover。《量子力学有助于大海捞针式搜索》。《style id='5'>物理评论快报，79(2)：325，1997年。[21] T. P. Hayes 和 A. Sinclair。《树结构马尔可夫链的提升》。载于 M. Serna、R. Shaltiel、K. Jansen 和 J. Rolim 编，《style id='8'>近似、随机化与组合优化：算法与技术》，第602–616页，柏林、海德堡，2010年。Springer Berlin Heidelberg。[22]M. Hildebrand。《具有 V 形平稳概率的 Diaconis-Holmes-Neal 马尔可夫链采样器的收敛速率》。《style id='11'>Markov Processes And Related Fields，10(4)：687–704，2004年。[23] R. Jordan 和 D. Kinderlehrer。《扩展变分原理》。载于 G. Talenti 编，《style id='14'>偏微分方程及其应用：纪念 Carlo Pucci 的论文集》，第187–200页。劳特利奇，2017年。[24] R. Jordan、D. Kinderlehrer 和 F. Otto。《福克–普朗克方程的变分形式》。《style id='17'>《SIAM 数学分析杂志》，29(1)：1–17，1998年。[25] K. Jung、D. Shah 和 J. Shin。《快速且精简的提升马尔可夫链》。载于《style id='20'>ALLERTON 通信、控制与计算会议。Curran Associates Inc.，2007年。[26] K. Kadian、S. Garhwal 和 A. Kumar。《量子游走及其应用领域：系统综述》。《style id='23'>《计算机科学评论》，41：100419，2021年。[27] J. Kempe。《量子随机游走：入门概述》。《style id='26'>当代物理学，44(4)：307–327，2003年。[28] N. Konno。《一维量子随机游走》。《style id='29'>量子信息处理，1(5)：345–354，2002年。[29]N. Konno。《一维量子随机游走的一类新极限定理》。《style id='32'>《日本数学会杂志》，57(4)：1179–1195，2005年。

[30] N. Konno。《量子游走》。载于 U. Franz 和 M. Schüermann 编，《style id='2'>量子势理论，《数学讲义》，第309–452页。Springer，2008年。[31] M. Lambert、S. Chewi、F. Bach、S. Bonnabel 和 P. Rigollet。《通过 Wasserstein 梯度流进行变分推断》。《style id='5'>神经信息处理系统进展，35：14434–14447，2022年。[32] L. LeCam。《style id='8'>统计决策理论中的渐近方法》。Springer统计学丛书。Springer，美国纽约州纽约，1986年。[33] S. Lim、E. Yoon、T. Byun、T. Kang、S. Kim、K. Lee 和 S. Choi。《希尔伯特空间中通过随机演化方程实现的基于分数的生成建模》。载于《style id='11'>第37届神经信息处理系统国际会议论文集，NIPS ’23，第37799–37812页，美国纽约州雷德胡克，2023年。Curran Associates Inc。[34] Y.-A. Ma、N. S. Chatterji、X. Cheng、N. Flammarion、P. L. Bartlett 和 M. I. Jordan。《梯度型MCMC是否存在 Nesterov加速的类似物？》。《style id='14'>Bernoulli，27(3)：1942–1992，2021年。[35] F. Magniez、A. Nayak、J. Roland 和 M. Santha。《通过量子行走进行搜索》。载于《style id='17'>第三十九届 ACM计算理论研讨会年会论文集，STOC ’07，第575–584页，美国纽约州纽约，2007年。计算机协会。[36] S. P. Meyn 和 R. L. Tweedie。《马尔可夫过程的稳定性 II：连续时间过程与采样链》。《style id='20'>应用概率进展，25(3)：487–517，1993年。[37] P. Mokrov、A. Korotin、L. Li、A. Genevay、J. M. Solomon 和 E. Burnaev。《大规模Wasserstein梯度流》。《style id='23'>神经信息处理系统进展，34：15243–15256，2021年。[38] A. Nayak 和 A. Vishwanath。《直线上的量子行走》。《style id='26'>arXiv预印本 arXiv:quant-ph/0010117v1，2000年。[39]Y. Nesterov。《style id='29'>凸优化导论：基础课程》，第87卷。Springer Science & Business Media，2013年。

[40] Y. E. Nesterov. 一种以收敛速率 $O(1/k^2)$ 求解凸规划问题的方法。载于*Dokl. Akad. Nauk SSSR*，第269卷，第543–547页，1983年。[41] M. A. Nielsen 和 I. L. Chuang. 量子计算与量子信息。剑桥大学出版社，10周年纪念版，2010年。[42] G. A. Pavliotis. 随机过程及其应用：扩散过程、*Fokker–Planck*方程与朗之万方程。《应用数学文本》，第60卷。Springer，纽约，2014年。[43] R. Portugal. 量子游走与搜索算法。《量子科学与技术》。Springer，第二版，2018年。[44] G. O. Roberts 和 R. L. Tweedie. 朗之万分布及其离散近似的指数收敛。*Bernoulli*，2(4):341–363，1996年。[45] M. Santha. 基于量子游走的搜索算法。载于 M. Agrawal、D. Du、Z. Duan 和 A. Li 编，《计算模型理论与应用》，第31–46页，柏林、海德堡，2008年。施普林格·柏林·海德堡。[46] N. Shenvi、J. Kempe 和 K. B. Whaley. 量子随机游走搜索算法。《物理评论*A*》，67:052307，2003年5月。[47] Y. Song、J. Sohl-Dickstein、D. P. Kingma、A. Kumar、S. Ermon 和 B. Poole. 通过随机微分方程进行基于分数的生成建模。载于第九届国际学习表征会议，ICLR。OpenReview.net，2021年5月。[48] O. Stramer 和 R. L. Tweedie. 具有非光滑系数的随机微分方程弱解的存在性与稳定性。《中华统计学志》，7(3):577–593，1997年。[49] W. Su、S. Boyd 和 E. J. Candès. 用于建模 Nesterov 加速梯度法的微分方程：理论与见解。《机器学习研究杂志》，17(153):1–43，2016年。

[50] S. Vempala 和 A. Wibisono. 未校正朗之万算法的快速收敛：等周性已足够。神经信息处理系统进展，32，2019年。[51] S. E. Venegas-Andraca. 量子游走：全面综述。量子信息处理，11(5):1015–1106，2012年。[52] Y. Wang. GARCH模型与扩散过程的渐近非等价性。统计年鉴，30(3):754–783，2002年。[53] Y. Wang. 量子态层析与含噪矩阵补全的渐近等价性。统计年鉴，41(5):2462–2504，2013年。[54] Y. Wang. 当量子计算遇上数据科学：让数据科学量子化。《哈佛数据科学评论》，4(1)，2022年1月27日。https://hdsr.mitpress.mit.edu/pub/kpn45eyx。[55] Y. Wang 和 S. Wu. 通过随机微分方程，对统计与计算范式中的梯度下降算法进行渐近分析。《机器学习研究杂志》，21(1):8179–8281，2020年。[56] A. Wibisono、A. C. Wilson 和 M. I. Jordan. 优化中加速方法的变分视角。《美国国家科学院院刊》，113(47):E7351–E7358，2016年。[57] T. G. Wong. Szegedy量子游走与硬币量子游走的等价性。量子信息处理，16(9):215，2017年7月。[58] Y. Yan、K. Wang 和 P. Rigollet. 利用 Wasserstein-Fisher-Rao 梯度流学习高斯混合。*arXiv*预印本 *arXiv:2301.01766*，2023年。[59] B. Zhang、P. Xu、X. Chen 和 Q. Zhuang. 通过去噪扩散概率模型实现生成式量子机器学习。《物理评论快报》，132:100602，2024年3月。[60] R. Zhang、C. Chen、C. Li 和 L. Carin. 将策略优化视为 Wasserstein 梯度流。国际机器学习会议，第5737–5746页，2018年。

[61] X. Zuo、S. Osher 和 W. Li。《用于采样的梯度调整欠阻尼朗之万动力学》。*arXiv:2410.08987v2*，2024年10月。

附录：定理1及结果（10）—（13）的证明

A1. 记号、弱收敛与证明思路

记 $\mathbb{P}_{1t}$ 为 X_t^q/t的分布， $\mathbb{P}_{1t\eta}$ 为 $X^q(\eta)_t/t$的分布， $\mathbb{P}_{2t\epsilon}$ 为 $X_t^c(\epsilon)/t$的分布。也就是说， $\mathbb{P}_{1t}$、 $\mathbb{P}_{1t\eta}$和 $\mathbb{P}_{2t\epsilon}$ 分别由$\mathbb{P}_t^q$ 、θ、 $\mathbb{P}_t^q$ 、θ、η以及$\mathbb{P}_t^c$、θ、ϵ 经因子 t缩放得到。下文中，为简便起见，在不会引起混淆时省略下标 θ ，例如将 $\mho(X^q(\eta)_t/t, X_t^c(\epsilon)/t)$和 $\beth(X_t^q(\eta)/t, X_t^c(\epsilon)/t)$分别记作 $\mho(\mathbb{P}_{1t\eta}, \mathbb{P}_{2t\epsilon})$和 $\beth(\mathbb{P}_{1t\eta}, \mathbb{P}_{2t\epsilon})$。定义一一变换$\mathcal{T}$: $(-\infty, \infty) \to (-|u_{11}|, |u_{11}|)$为 $x = \mathcal{T}(y) = |u_{11}|e^{-|y|}\text{sign}(y)$。注意，全变差范数在任意一一变换（例如缩放和 $\mathcal{T}$）下保持不变。因此有 $\mho(X^q(\eta)_t, X_t^c(\epsilon)) = \mho(\mathbb{P}_{1t\eta}, \mathbb{P}_{2t\epsilon})$。由于 $\mho(\mathbb{P}_{1t\eta}, \mathbb{P}_{2t\epsilon}) = \max\{\beth(\mathbb{P}_{1t\eta}, \mathbb{P}_{2t\epsilon}) \ \beth(\mathbb{P}_{2t\epsilon}, \mathbb{P}_{1t\eta})\}$，下文将证明，当 $t \to \infty$时， $\beth(\mathbb{P}_{1t\eta}, \mathbb{P}_{2t\epsilon})$和 $\beth(\mathbb{P}_{2t\epsilon}, \mathbb{P}_{1t\eta})$的极限满足定理中给出的界。

令 $p(k, t; \theta)$ 为 X_t^q 的概率函数。根据定义，对于任意 $y \in \mathbf{R}$,

$$\mathbb{P}_{1t}((-\infty, y]) = \sum_{k \le ty} p(k, t; \theta). \tag{14}$$

X_t^q 在 $t \to \infty$ 时的极限分布已分别由 [28], 中的定理3、 [29], 中的定理1以及 [18] 中的定理1确立如下,

$$\frac{X_t^q}{t} \Rightarrow \mathbb{Q} \text{ as } t \to \infty, \tag{15}$$

其中 $\Rightarrow$表示弱收敛， $\mathbb{Q}$ 的概率密度函数为 $f(x; \theta)$，称为 Konno密度函数，其定义为

$$\begin{aligned} f(x;\theta) &= \frac{\sqrt{1-|u_{11}|^2}}{\pi(1-x^2)\sqrt{|u_{11}|^2-x^2}} \\ &\quad \left\{1 - \left(|a_0|^2 - |a_1|^2 + \frac{u_{11}a_0\overline{u_{12}a_1} + \overline{u_{11}a_0}u_{12}a_1}{|u_{11}|^2}\right)x\right\} \end{aligned} \tag{16}$$

对于 $x \in (-|u_{11}|, |u_{11}|)$ 和 $f(x;\theta) = 0$ 其他情况。式 (15) 中的弱收敛是通过证明 X_t^q/t 的特征函数收敛到 $\mathbb{Q}$ 的特征函数得出的，其中 $t \to \infty$ （见 [18] [27] [28] [29] [30] [51]）。请注意，首先， $\mathbb{Q}$ 是绝对连续的，其概率密度函数为 $f(x;\theta)$；其次，时刻 t 的量子游走最多可向起始位置的右侧或左侧移动 t 步，因此 X_t^q/t 有上下界。因此，弱收敛意味着

$$\lim_{t\to\infty} \sum_{k\le ty} p(k,t;\theta) = \mathbb{Q}((-\infty, y]) \text{ uniformly in } y \in \mathbf{R},$$

以及相应矩的收敛和矩生成函数的收敛。由于概率函数 $X_t^q(\eta)$ 是 X_t^q 对应概率函数的平均值，根据 Skorokhod表示定理，我们很容易证明，对于任意 $\eta > 0$,

$$\frac{X_t^q(\eta)}{t} \Rightarrow \mathbb{Q} \text{ as } t \to \infty. \tag{17}$$

容易看出，Konno密度 $f(x;\theta)$ 与式 (5) 中的概率密度函数 $\pi(y;\theta)$ 通过变换 $x = \mathcal{T}(y)$ 相关。定义 $f_\epsilon(x;\theta) = f(x;\theta)$，其中 $x \in [-|u_{11}|(1-\epsilon)\ |u_{11}|(1-\epsilon)]$， $f(x;\theta) = \varkappa_\epsilon$ 当 $x \in (-|u_{11}|, -|u_{11}|(1-\epsilon)) \cup (|u_{11}|(1-\epsilon)\ |u_{11}|)$ 时，以及 $f(x;\theta) = 0$ 在其他情况下，其中 $\varkappa_\epsilon$ 在式 (4) 中给出，其表达式如下:

$$\begin{aligned}
\varkappa_\epsilon &= \frac{1}{2\epsilon|u_{11}|}\left[1 - \int_{-|u_{11}|(1-\epsilon)}^{|u_{11}|(1-\epsilon)} f(x;\theta)dx\right] \\
&= \frac{\sqrt{1-|u_{11}|^2}}{\pi\epsilon}\int_{1-\epsilon}^{1} \frac{1}{(1-|u_{11}|^2z^2)\sqrt{1-z^2}}dz \\
&= \frac{\sqrt{2}}{\pi\sqrt{\epsilon}\sqrt{1-|u_{11}|^2}} + O(\sqrt{\epsilon}).
\end{aligned} \tag{18}$$

A2. Langevin 动力学的全变差范数结果

ics

用 $\rho_\epsilon(y,v,t)$ 表示由势函数 $\Psi_\epsilon(y)$ 决定、满足式（3）的欠阻尼Langevin过程 (Y_t^c, V_t^c) 的概率密度函数，并用 $\pi_\epsilon(y,v;\theta)$ 表示其不变分布。 $\rho_\epsilon(y,v,t)$ 满足如下形式的福克–普朗克方程

:

$$\frac{\partial \rho_\epsilon}{\partial t} = -v \cdot \nabla_y \rho_\epsilon + \nabla_y \Psi_\epsilon \cdot \nabla_v \rho_\epsilon + \nabla_v \cdot (v \rho_\epsilon) + \Delta_v \rho_\epsilon, \;\; \rho_\epsilon(y, v, 0) = \rho_{\epsilon,0}(y, v),$$

其中$\rho_{\epsilon,\ 0}$是初始概率密度函数。由 [48] 的定理 2.6 和命题 4.2(i) 可得，(Y_t^c，V_t^c) 关于勒贝格测度是不可约 T-过程，且具有不变测度$\pi_\epsilon(y,\ v;\ \theta)$，并且是正 Harris 常返的。因此，[36] 的定理 6.1 表明它是遍历的——也就是说，当 $t \to \infty$ 时，分布 $\rho_\epsilon(y,\ v,\ t)$ 在全变差范数下收敛到$\pi_\epsilon(y,\ v;\ \theta) \propto \exp[-\Psi_\epsilon(y) - v^2/2]$，其中条件验证如下：

$$\operatorname{rank}\left(\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0 & 1\\0 & 1\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}\right) = 2;$$

$$\begin{aligned}\nabla \Psi_\epsilon(y) &= \operatorname{sign}(y)\left[-1 + \frac{2}{1-|u_{11}|^2 e^{-2|y|}} + \frac{1}{1-e^{-2|y|}} - \frac{1}{1-\varpi e^{-|y|}}\right]\\ &\to \mp 1 \text{ as } y \to \pm\infty;\end{aligned}$$

and

$$\begin{aligned}&2\left(\begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}\begin{pmatrix} y\\ v\end{pmatrix}, \begin{pmatrix} v\\ -v - \nabla\Psi_\epsilon(y)\end{pmatrix}\right) + tr\left(\begin{pmatrix}0 & 0\\0 & 2\end{pmatrix}\begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}\right)\\ &= -2v^2 - 2v\nabla\Psi_\epsilon(y) + 2,\end{aligned}$$

当 $|y|$ 足够大时，它被 $-v^2$ 和 $|v|$所界（另见 [44]中定理 2.1 的证明）。这里 $\varpi = \left(|a_0|^2 - |a_1|^2 + \frac{u_{11}a_0u_{12}a_1 + u_{11}a_0u_{12}\bar{a}_1}{|u_{11}|^2}\right)\overline{|u_{11}|}$

用 $\rho_\epsilon(y,t)$ 表示 $\rho_\epsilon(y,v,t)$ 关于 y 的边缘概率密度函数。注意，式（5）中定义的 $\pi_\epsilon(y;\theta)$ [] 是 $\pi_\epsilon(y,v;\ \theta)$ 关于y 的边缘概率密度函数。因此，我们有 $\|\rho_\epsilon(y,t) - \pi_\epsilon(y;\ \theta)\|_{TV} \le \|\rho_\epsilon(y,v,t) - \pi_\epsilon(y,v;\ \theta)\|_{TV} \to 0$ 如$t \to \infty$。注意，从 Y_t^c 到 $X_t^c(\epsilon)$ 的变换是一一对应的，且分布 $\mathbb{P}_{2t\epsilon}$ 和概率密度函数 $\pi_\epsilon(y;\theta)$ 分别由概率密度函数 $\rho_\epsilon(y;t)$ 和 $f_\epsilon(x;\theta)$ 经由变换 $\mathcal{T}$ 及缩放因子 t 得到。由于全变差范数在变换下保持不变，$\|\rho_\epsilon(y,t) - \pi_\epsilon(y;\ \theta)\|_{TV} = \|\mathbb{P}_{2t\epsilon} - \mathbb{Q}_\epsilon\|_{TV}$，其中 $\mathbb{Q}_\epsilon$ 表示概率密度函数为 $f_\epsilon(x;\theta)$ 的分布。因此，对每个 ϵ,

$$\|\mathbb{P}_{2t\epsilon} - \mathbb{Q}_\epsilon\|_{TV} \to 0 \text{ as } t \to \infty.$$

概率密度函数分别为 $f(x;\theta)$ 的 $\mathbb{Q}$ 与概率密度函数为$f_\epsilon(x;\theta)$ 的 $\mathbb{Q}_\epsilon$ 之间的全变差距离满足以下界:

$$\begin{aligned}
&\|\mathbb{Q}-\mathbb{Q}_\epsilon\|_{TV}=\frac{1}{2}\int|f(x;\theta)-f_\epsilon(x;\theta)|dx\\
&=\frac{1}{2}\int_{x\in(-|u_{11}|,-|u_{11}|(1-\epsilon))\cup(|u_{11}|(1-\epsilon),|u_{11}|)}|f(x;\theta)-\kappa_\epsilon|dx\\
&\leq\frac{1}{2}\int_{x\in(-|u_{11}|,-|u_{11}|(1-\epsilon))\cup(|u_{11}|(1-\epsilon),|u_{11}|)}[f(x;\theta)+\kappa_\epsilon]dx\\
&=2\epsilon|u_{11}|\varkappa_\epsilon=\frac{2\sqrt{2\epsilon}|u_{11}|}{\pi\sqrt{1-|u_{11}|^2}}+O(\epsilon^{3/2})\\
&\leq\frac{2\sqrt{2\epsilon}}{\pi\tan\kappa}+O(\epsilon^{3/2}).
\end{aligned}$$

将上述两个结果结合起来，我们得到

$$\begin{aligned}
&\limsup_{t\to\infty}\|\mathbb{P}_{2t\epsilon}-\mathbb{Q}\|_{TV}\leq\limsup_{t\to\infty}[\|\mathbb{P}_{2t\epsilon}-\mathbb{Q}_\epsilon\|_{TV}+\|\mathbb{Q}_\epsilon-\mathbb{Q}\|_{TV}]\\
&\leq\frac{2\sqrt{2\epsilon}}{\pi\tan\kappa}+O(\epsilon^{3/2}).
\end{aligned}\tag{19}$$

A3. 量子游走的全变差范数结果

回顾 $r=[\eta t]$。$\mathbb{P}^q_{t,\theta,\eta}$ 具有一个概率函数

$$\begin{aligned}
&\sum_{j=-r+1}^{r}Pr(X^q_t+j=y)\frac{1}{2r}=\frac{1}{2r}\sum_{j=-r+1}^{r}p(y-j,t;\theta)\\
&=\frac{1}{2r}\sum_{k=y-r}^{y+r-1}p(k,t;\theta).
\end{aligned}$$

在定义 $X^q_t(\eta)$ 的随机化方案中，我们添加一个取值于 $(-1/2,1/2]$ 到 $X^q_t+j,\ j=0,\pm1,\cdots,\pm(r-1)$ r 上的独立均匀随机变量 U_j，并将所得随机变量的分布记为 $\mathbb{P}^q_{*t,\theta,\eta}$。也就是说，$\mathbb{P}^q_{*t,\theta,\eta}$是 $\mathbb{P}^q_{t,\theta,\eta}$ 与取值于 $(-1/2,1/2]$ 上的独立均匀分布的卷积分布。于是 $\mathbb{P}^q_{*t,\theta,\eta}$是绝对连续的

分布，其累积分布函数和概率密度函数如下：$\mathbb{P}^q_{*t,\theta,\eta}$ 的累积分布函数为

$$\sum_{j=-r+1}^{r} Pr(X_t^q + j + U_j \le y)\frac{1}{2r} = \frac{1}{2r}\sum_{j=-r+1}^{r} Pr(U_j \le y - j - X_t^q)$$
$$= \frac{1}{2r}\sum_{j=-r+1}^{r}\sum_{\ell} p(\ell, t; \theta) Pr(U_j \le y - j - \ell)$$
$$= \frac{1}{2r}\sum_{j=-r+1}^{r} p([y] - j, t; \theta) Pr(U_j \le y - [y])$$
$$= \frac{1}{2r}\sum_{k=[y]-r}^{[y]+r-1} p(k, t; \theta) Pr(U_j \le y - [y]),$$

其中，第三个等号是因为 U_j 在（$-1/2$，$1/2$]，$y - j - \ell \in$（$-1/2$，$1/2$] 上服从均匀分布，因此 $0 = [y - k - \ell] = [y] - k - \ell$；所以其概率密度函数等于

$$\frac{1}{2r}\sum_{k=[y]-r}^{[y]+r-1} p(k, t; \theta).$$

将通过变换后再除以 t 得到的 $\mathbb{P}^q_{*t,\ \theta,\ \eta}$ 的分布记为 $\mathbb{P}^*_{1t\eta}$。也就是说，在上述添加均匀随机变量的过程中，我们分别用（X_t^q/t，（$X_t^q + j$）$/t$，U_j/t）替换（X_t^q，$X_t^q + j$，U_j）。注意，$\mathbb{P}_{1t\eta}$ 是将 $\mathbb{P}_{t,\ \theta,\ \eta}$ 通过相同的变换后再除以 t 得到的分布。于是 $\mathbb{P}_{1t\eta}$ 的概率函数为

$$p_{1t\eta}(x) = \frac{1}{2[\eta t]}\sum_{k=[tx]-r}^{[tx]+r-1} p(k, t; \theta),$$

且 $\mathbb{P}^*_{1t\eta}$ 具有概率密度函数

$$p_{*1t\eta}(x) = \frac{t}{2[\eta t]}\sum_{k} p(k, t; \theta) 1\{-\eta < ([tx] - k)/t \le \eta\}$$
$$= \frac{\eta t}{[\eta t]}\sum_{k} p(k, t; \theta)\frac{1}{2\eta} 1\{-\eta < ([tx] - k)/t \le \eta\}$$
$$= \sum_{k} p(k, t; \theta)\frac{1}{2\eta} 1\{-\eta < ([tx] - k)/t \le \eta\} + O((\eta^2 t)^{-1}),$$

其中最后一个等号是由于

$$\frac{t}{2[\eta t]} = \frac{1}{2\eta} + O((\eta^2 t)^{-1}).$$

或者，我们在（$-\eta$， η]到 X_t^q/t）上加入一个独立的均匀随机变量，并将所得随机变量的分布记为 $\check{}\,\mathbb{P}^*_{1t\eta}$。也就是说， $\mathbb{P}^*_{1t\eta}$ 是 $\mathbb{P}_{1t}$ 与（$-\eta$， η]）上的独立均匀分布的卷积分布。$\check{}\,\mathbb{P}^*_{1t\eta}$ 是绝对连续的，其概率密度函数为

$$\check{p}_{*1t\eta}(x) = \sum_k p(k,t;\theta)\frac{1}{2\eta}1\{-\eta < ([tx]-k)/t \le \eta\}. \tag{20}$$

注意

$$p_{*1t\eta}(x) = \frac{\eta t}{[\eta t]}\check{p}_{*1t\eta}(x),$$

and

$$\begin{aligned}\|\check{\mathbb{P}}^*_{1t\eta} - \mathbb{P}^*_{1t\eta}\|_{TV} &= \int |\check{p}_{*1t\eta}(x) - p_{*1t\eta}(x)|dx \\ &= \left|\frac{\eta t}{[\eta t]} - 1\right| \le \frac{1}{[\eta t]} = O((\eta t)^{-1}).\end{aligned} \tag{21}$$

定义海灵格距离 $H(\check{}\,\mathbb{P}^*_{1t\eta}$， $\mathbb{Q})$ 为

$$H^2(\check{\mathbb{P}}^*_{1t\eta},\mathbb{Q}) = \int_{\mathbf{R}}\left|\sqrt{\frac{d\check{\mathbb{P}}^*_{1t\eta}}{dx}} - \sqrt{\frac{d\mathbb{Q}}{dx}}\right|dx = 2 - 2E_{\check{\mathbb{P}}^*_{1t\eta}}\left[\sqrt{\frac{d\mathbb{Q}}{d\check{\mathbb{P}}^*_{1t\eta}}}\right],$$

其中 dx 表示勒贝格测度。简单的概率分析可得

$$\|\check{\mathbb{P}}^*_{1t\eta} - \mathbb{Q}\|^2_{TV} \le H^2(\check{\mathbb{P}}^*_{1t\eta},\mathbb{Q}) \le 2\check{\mathbb{P}}^*_{1t\eta}(A^c) + E_{\check{\mathbb{P}}^*_{1t\eta}}\left[1_A \log\frac{d\check{\mathbb{P}}^*_{1t\eta}}{d\mathbb{Q}}\right], \tag{22}$$

其中 A 是任意固定的博雷尔集，而 A^c 表示其补集（参见 [53]中的 [32] 及公式 (42)-(43)）。

对于酉矩阵**U** 由式（6）给出的量子游走， $\phi = \pi/4$ 对应一般无偏行走，而 ϕ 的其他取值表示有偏

行走。一般无偏行走的概率行为与 Hadamard 行走完全相同。参数 ϕ 标记了一族行走，其中每个元素在相同初始条件下都会产生相同的概率分布。有偏行走的概率行为与 Hadamard 行走相似，只需根据偏置参数 ϕ 进行一些调整（例如使用有偏因子 $\cos\phi$ 而非 $1/\sqrt{2}$）（见 [2] [18] [38]）。为简便起见，我们以 Hadamard 行走为例进行论述，其初始状态为（a_0， a_1）=（1，0），酉矩阵U 则取式（1）给出的 Hadamard 算子。

令 ζ 为一个稍后确定的小正数。令 $A_\zeta = (-1/\sqrt{2}+\zeta,\ 1/\sqrt{2}-\zeta)$。给定 $x \in A_\zeta$，对于任意 t，我们选择一个整数 $\ell_x = [xt]$，其中$[\cdot]$表示最近整数值。注意，$x - 1/(2t) < \ell_x/t \le x + 1/(2t)$，并且当 $t \to \infty$时，$\ell_x/t \to x$。首先，在不等式（22）中，我们取 A 为 A_ζ。其次，我们有

$$\frac{d\check{\mathbb{P}}^*_{1t\eta}}{d\mathbb{Q}} = \frac{\check{p}_{*1t\eta}(x)}{f(x;\theta)}, \tag{23}$$

其中 $\check{p}_{*1t\eta}(x)$ 表示由（20）给出的 $\check{\mathbb{P}}^*_{1t\eta}$ 的概率密度函数。显然，$f(x;\theta)$ 在 A_ζ 上有下界 $f(0;\theta)$，上界为 $\max\{f(-1/\sqrt{2}+\zeta;\theta)\ f(1/\sqrt{2}-\zeta;\theta)\}$。利用 [2] 中的定理 2 和 [38], 中的方程（8），可以证明，当 t 足够大时，对于 $x \in A_\zeta$，$tp([tx],t;\theta)$ 介于 $2f(x;\theta)$ 与 $\zeta f(x;\theta)$ 之间。此外，如 [2] 和 [38], 所示，我们可以将 $tp([tx],t;\theta)$ 分解为慢分量和快分量，其中慢分量是一个缓慢变化且不振荡的函数，用以描述 $tp([tx],t;\theta)$ 的整体性质；快分量则是剩余的快速振荡函数。将慢分量视为 x 的函数时，它是一个概率密度函数，等于 $f(x,\theta)$。用 $g(x,t;\theta)$ 表示快分量。那么对于 $x \in A_\zeta$，

$$tp([tx],t;\theta) = f([tx]/t,\theta) + g([tx]/t,t;\theta), \tag{24}$$

其中

$$g(x,t;\theta) = f(x;\theta)H(x,t) + O(t^{-1}), \tag{25}$$

$$H(x,t) = \sqrt{2}x\sin(2[\omega_{\alpha_x} + x\alpha_x]t - \beta_x), \tag{26}$$

$$\omega_\alpha = \sin^{-1}\left(\frac{\sin\alpha}{\sqrt{2}}\right),\ \cos\alpha_x = \frac{x}{\sqrt{1-x^2}},\ \cos\beta_x = \frac{2x-1}{\sqrt{2}(1-x)}. \tag{27}$$

根据式 (24)–(27)，对于 $x \in A_\zeta$，我们将式 (20) 中的 $\check{p}_{*1t\eta}(x)$ 写成以下两项之和：

$I_1(x)$ 和 $I_2(x)$，其中

$$
\begin{aligned}
I_1(x) &= \sum_{u_k=k/t} \frac{1}{t} f(u_k;\theta) \frac{1}{2\eta} 1\{-\eta < x - u_k < \eta\} \\
&= \frac{1}{2\eta} \int_{x-\eta}^{x+\eta} f(u;\theta) du + O(t^{-1}) \\
&= f(x;\theta) + O(\eta) + O(t^{-1}), \\
I_2(x) &= \sum_{u_k=k/t} \frac{1}{t} g(u_k,t;\theta) \frac{1}{2\eta} 1\{-\eta < x - u_k < \eta\} \\
&= \sum_{u_k=k/t} \frac{1}{t} f(u_k;\theta) H(u_k,t) \frac{1}{2\eta} 1\{-\eta < x - u_k < \eta\} + O(\eta t^{-1}) \\
&= \frac{1}{2\eta} \int_{x-\eta}^{x+\eta} f(u;\theta) H(u,t) du + O(t^{-1}) + O(\eta t^{-1}) \\
&= O(\eta + (\eta t)^{-1}) + O(t^{-1}) + O(\eta t^{-1}).
\end{aligned}
$$

这里，对 $I_2(x)$ 的上述分析中，最后一个等式是基于以下事实：

$$
\begin{aligned}
&\frac{1}{2\eta} \int_{x-\eta}^{x+\eta} f(u;\theta) H(u,t) du = \frac{1}{2\eta} \int_{x-\eta}^{x+\eta} f(u;\theta) \sqrt{2} u \sin\left(2[\omega_{\alpha_u} + u\alpha_u]t - \beta_u\right) du \\
&= \frac{1}{\sqrt{2}\eta} \int_{x-\eta}^{x+\eta} f(x;\theta) x \sin\left(2[\omega_{\alpha_u} + u\alpha_u]t - \beta_x\right) du + O(\eta) \\
&= \frac{f(x;\theta)x}{\sqrt{2}\eta} \left(\frac{d[\omega_{\alpha_x} + x\alpha_x]}{dx} \right)^{-1} \frac{1}{2t} \{-\cos\left(2[\omega_{\alpha_u} + u\alpha_u]t - \beta_x\right)\}|_{u=x-\eta}^{u=x+\eta} + O(\eta) \\
&= O((\eta t)^{-1} + \eta).
\end{aligned}
$$

利用得到的 $I_1(x)$、$I_2(x)$ 和 $\check{p}_{*1t\eta}(x) = I_1(x) + I_2(x)$ 的表达式，我们可得，对于 $x \in A_\zeta$，

$$
\check{p}_{*1t\eta}(x) = f(x;\theta) + O(\eta + (\eta t)^{-1}). \tag{28}
$$

结合（23）和（28），我们得到，对于 $x \in A_\zeta$，

$$
\log \frac{d\check{\mathbb{P}}^*_{1t\eta}}{d\mathbb{Q}} = \log \left[\frac{f(x;\theta) + O(\eta + (\eta t)^{-1})}{f(x;\theta)} \right] \le \frac{O(\eta + (\eta t)^{-1})}{f(x;\theta)},
$$

因此，应用控制收敛定理，我们有

$$\limsup_{t\to\infty} E_{\check{\mathbb{P}}^*_{1t\eta}}\left[1_{A_\zeta}\log\frac{d\check{\mathbb{P}}^*_{1t\eta}}{d\mathbb{Q}}\right]=O(\eta). \tag{29}$$

事件 $X_t^q\in tA_\zeta=((-1/\sqrt{2}+\zeta)t,\ (1/\sqrt{2}-\zeta)t)$ 表示量子游走在时间 t 的测量结果落在区间 $((-1/\sqrt{2}+\zeta)t,\ (1/\sqrt{2}-\zeta)t)$ 内。由 [2] 中的定理1可知，当 $t\to\infty$ 时，对于 $x\in(-1,\ -1/\sqrt{2})\cup(1/\sqrt{2},\ 1)$，$p([xt],\ t;\theta)=O(\varrho^t)$，其中 $\varrho\in(0,\ 1)$ 是一个固定常数。利用 [2] 中定理 2 给出的 $p(k,\ t;\ \theta)$ 在 $t\to\infty$ 时的极限表达式，我们可以直接计算

$$\sum_{x=k/t,k\in[(-1/\sqrt{2}+\zeta)t,(1/\sqrt{2}-\zeta)t]} p(x,t;\theta)=1-\frac{2\zeta}{\pi}-O(t^{-1}),$$

and

$$\sum_{x=k/t,k\in[-t/\sqrt{2},(-1/\sqrt{2}+\zeta)t]\cup[(1/\sqrt{2}-\zeta)t,t/\sqrt{2}]} p(x,t;\theta)=\frac{2\zeta}{\pi}+O(t^{-1}).$$

事实上，上述结果也见于 [2] 中定理 2 之后的讨论、 [51] 中的定理4以及 [38] 的附录C。因此，我们有

$$\begin{aligned}\check{\mathbb{P}}^*_{1t\eta}(A_\zeta^c) \ &\le\ \mathrm{Prob}(X_t^q/t\in A_\zeta^c)=\mathbb{P}_{1t}(A_\zeta^c)=\sum_{x=k/t,k\in(-t,-t/\sqrt{2})\cup(t/\sqrt{2},t)} p(x,t;\theta)\\ &\quad+\sum_{x=k/t,k\in[-t/\sqrt{2},(-1/\sqrt{2}+\zeta)t]\cup[(1/\sqrt{2}-\zeta)t,t/\sqrt{2}]} p(x,t;\theta)\\ &=O(t\varrho^t)+\frac{2\zeta}{\pi}+O(t^{-1}).\end{aligned} \tag{30}$$

将（29）和（30）代入（22），我们得出

$$\begin{aligned}&\limsup_{t\to\infty}\|\check{\mathbb{P}}^*_{1t\eta}-\mathbb{Q}\|^2_{TV}\le 2\limsup_{t\to\infty}\left[O(t\varrho^t)+\frac{2\zeta}{\pi}+O(t^{-1})\right]+O(\eta)\\ &=\frac{4\zeta}{\pi}+O(\eta).\end{aligned} \tag{31}$$

由于 (31) 的左侧与 ζ 无关，我们令 ζ 趋于零，得到

$$\limsup_{t\to\infty}\|\check{\mathbb{P}}^*_{1t\eta}-\mathbb{Q}\|_{TV}=O(\eta^{1/2}). \tag{32}$$

A4. 定理1中朗之万动力学与量子游走之间勒卡姆距离的结果（8）

结合 (19)、(21) 和 (32)，我们得出

$$\begin{aligned}\limsup_{t\to\infty}\|\mathbb{P}^*_{1t\eta}-\mathbb{P}_{2t\epsilon}\|_{TV} &\le \limsup_{t\to\infty}[\|\mathbb{P}^*_{1t\eta}-\check{\mathbb{P}}^*_{1t\eta}\|_{TV}+\|\check{\mathbb{P}}^*_{1t\eta}-\mathbb{Q}\|_{TV}\\ &\quad +\|\mathbb{P}_{2t\epsilon}-\mathbb{Q}\|_{TV}]\\ &\le \frac{2\sqrt{2\epsilon}}{\pi\tan\kappa}+O(\epsilon^{3/2}+\eta^{1/2}),\end{aligned}\tag{33}$$

这对任意 $\theta\in\Theta$ 均成立。注意，φ 属于 $\mathbb{C}^2$ 中单位球面上的一个闭集，$\mathbf{U}$ 位于紧集 $\{|u_{11}|^2+|u_{12}|^2=|u_{21}|^2+|u_{22}|^2=1\}$ 中，且 Θ 是紧的。此外，$\|\mathbb{P}^*_{1t\eta}-\mathbb{P}_{2t\epsilon}\|_{TV}$ 关于 θ 连续，因为 $\mathbb{P}_{1t}$、$\mathbb{P}^*_{1t\eta}$ 和 $\mathbb{P}_{2t\epsilon}$ 都连续依赖于 θ。因此，

$$\limsup_{t\to\infty}\sup_{\theta\in\Theta}\|\mathbb{P}^*_{1t\eta}-\mathbb{P}_{2t\epsilon}\|_{TV}\le\frac{2\sqrt{2\epsilon}}{\pi\tan\kappa}+O(\epsilon^{3/2}+\eta^{1/2}).\tag{34}$$

最后，利用 $\mathbb{P}_{1t\eta}$ 与均匀分布的卷积是一种特殊的马尔可夫核这一事实，我们得到

$$\begin{aligned}\limsup_{t\to\infty}\beth(\mathbb{P}_{1t\eta},\mathbb{P}_{2t\epsilon}) &\le \limsup_{t\to\infty}\sup_{\theta\in\Theta}\|\mathbb{P}^*_{1t\eta}-\mathbb{P}_{2t\epsilon}\|_{TV}\\ &\le \frac{2\sqrt{2\epsilon}}{\pi\tan\kappa}+O(\epsilon^{3/2}+\eta^{1/2}).\end{aligned}\tag{35}$$

对于 $\beth(\mathbb{P}_{2t\epsilon},\ \mathbb{P}_{1t\eta})$ 的情况，我们通过离散化 $\mathbb{P}_{2t\epsilon}$ 来反转上述卷积方案，以便与 $\mathbb{P}_{1t\eta}$ 进行比较。具体而言，定义 $\mathbb{P}^*_{2t\epsilon}$ 为取值 k/t、$k=0,\ \pm1,\ \pm2,\ \cdots$ 的离散分布，其在 k/t 处的概率函数等于 $\mathbb{P}_{2t\epsilon}(\lfloor\frac{k}{t}-\frac{1}{2t},\ \frac{k}{t}+\frac{1}{2t}\rfloor)$。此外，我们将 $\mathbb{P}^*_{1t\eta}$ 离散化。令 $\mathbb{P}^{**}_{1t\eta}$ 为取值 k/t、$k=0,\ \pm1,\ \pm2,\ \ldots..$ 的离散分布，其在 k/t 处的概率函数等于 $\mathbb{P}^*_{1t\eta}(\lfloor\frac{k}{t}-\frac{1}{2t},\frac{k}{t}+\frac{1}{2t}\rfloor)$。注意，$r=[\eta t]$。根据 (20)，对于 $x=\ell_x/t$ 或 $\ell_x=xt$，

$$\begin{aligned}&\mathbb{P}^*_{1t\eta}\left(\left(\frac{k}{t}-\frac{1}{2t},\frac{k}{t}+\frac{1}{2t}\right]\right)=\int_{x-1/(2t)}^{x+1/(2t)}p^*_{1t\eta}(u)du\\ &=\sum_{k=\ell_x-r}^{\ell_x+r-1}p(k,t;\theta)\frac{1}{2[\eta t]}=\frac{1}{2r}\sum_{k=\ell_x-r}^{\ell_x+r-1}p(k,t;\theta)=p_{1t\eta}(x).\end{aligned}$$

因此， $\mathbb{P}_{1t\eta}^{**} = \mathbb{P}_{1t\eta}$，也就是说，对 $\mathbb{P}_{1t\eta}^{*}$ 进行离散化即可恢复 $\mathbb{P}_{1t\eta}$。

由于 $\mathbb{P}_{2t\epsilon}^{*}$ 和 $\mathbb{P}_{1t\eta}$ 分别是 $\mathbb{P}_{2t\epsilon}$ 和 $\mathbb{P}_{1t\eta}^{*}$的离散化版本，且离散化会减小它们对应的全变差距离，因此我们有 $\|\mathbb{P}_{1t\eta} - \mathbb{P}_{2t\epsilon}^{*}\|_{TV} \leq \|\mathbb{P}_{1t\eta}^{*} - \mathbb{P}_{2t\epsilon}\|_{TV}$。于是，由 (33) 可得

$$\limsup_{t\to\infty} \|\mathbb{P}_{1t\eta} - \mathbb{P}_{2t\epsilon}^{*}\|_{TV} \leq \frac{2\sqrt{2\epsilon}}{\pi \tan\kappa} + O(\epsilon^{3/2} + \eta^{1/2}). \tag{36}$$

再次利用 $\mathbb{P}_{1t\eta}$ 和 $\mathbb{P}_{2t\epsilon}^{*}$ 关于 θ 的连续性以及 Θ的紧致性，我们得出 $\sup_{\theta\in\Theta|} \|\mathbb{P}_{1t\eta} - \mathbb{P}_{2t\epsilon}^{*}\|_{TV}$的极限具有与 (36) 中相同的上界，这又意味着 $\beth(\mathbb{P}_{2t\epsilon}, \mathbb{P}_{1t\eta})$的极限也具有相同的上界，因为离散化是一个马尔可夫核。也就是说，

$$\limsup_{t\to\infty} \beth(\mathbb{P}_{2t\epsilon}, \mathbb{P}_{1t\eta}) \leq \frac{4\sqrt{2\epsilon}}{\pi \tan\kappa} + O(\epsilon^{3/2} + \eta^{1/2}). \tag{37}$$

定理1的第一个结果 (8) 是 (35) 和 (37) 的推论。

A5. 朗之万动力学与量子游走之间的勒卡姆距离定理1中的结果（9）

在亏损 $\beth$ 和勒卡姆距离 $\mho$的定义中，我们需要对参数空间取上确界。因此，缩小参数空间会降低亏损 $\beth$ 和勒卡姆距离 $\mho$。为了给出（9）的一个正下界，我们可以将参数空间 Θ缩小为一个单点集合 $\Theta_0 = \{\theta_0\}$，该集合对应于（6）和（7）中的 $\gamma_1 = \gamma_2 = \varsigma_1 = 0$、 $\varsigma_2 = \pi/2$和 $\vartheta = \phi = \pi/4$。在定理证明的其余部分，我们考虑单点参数空间 Θ_0，处理$\beth$和 $\mho$时不再取上确界。

利用距离 $\mho$的三角不等式以及 $\mho$与 $\beth$之间的关系，我们有

$$\begin{aligned} \mho(\mathbb{P}_{2t\epsilon}, \mathbb{P}_{1t}) &\geq \mho(\mathbb{Q}^{*}, \mathbb{P}_{1t}) - \mho(\mathbb{Q}^{*}, \mathbb{P}_{2t\epsilon}) \\ &\geq \beth(\mathbb{Q}^{*}, \mathbb{P}_{1t}) - \mho(\mathbb{Q}^{*}, \mathbb{Q}) - \mho(\mathbb{Q}, \mathbb{P}_{2t\epsilon}), \end{aligned} \tag{38}$$

其中， $\mathbb{Q}^{*}$ 定义为取值为 k/t、 k =0， ± 1， ± 2， $\cdots$的离散分布，在 k/t 处的概率函数等于 $\mathbb{Q}((\frac{k}{t} - \frac{1}{2t}, \frac{k}{t} + \frac{1}{2t}])$。我们将分析（38）右侧的三项。

首先，很容易看出（19）意味着

$$\limsup_{t\to\infty} \mho(\mathbb{Q}, \mathbb{P}_{2t\epsilon}) = O(\epsilon^{1/2}). \tag{39}$$

其次，我们将证明

$$\limsup_{t\to\infty} \mho(\mathbb{Q}^*, \mathbb{Q}) = 0. \tag{40}$$

由于 $\mathbb{Q}^*$ 是 $\mathbb{Q}$ 的离散化，且离散化是一个马尔可夫核，因此我们有 $\beth(\mathbb{Q}, \mathbb{Q}^*) = 0$。为处理 $\beth(\mathbb{Q}^*, \mathbb{Q})$，我们定义 $\mathbb{Q}^{**}$ 为 $\mathbb{Q}^*$ 与在 $(\frac{k}{t} - \frac{1}{2t}, \frac{k}{t} + \frac{1}{2t}]$ 上的独立均匀分布的卷积。由于卷积是一个马尔可夫核，我们得到 $\beth(\mathbb{Q}^*, \mathbb{Q}) \le \|\mathbb{Q}^{**} - \mathbb{Q}\|_{TV}$。因此，只需证明 $\mathbb{Q}$ 与 $\mathbb{Q}^{**}$ 之间的全变差距离趋于零。记 $f^{**}(x; \theta_0)$ 为 $\mathbb{Q}^{**}$ 的概率密度函数。则对于 $x \in (-1/\sqrt{2}, 1/\sqrt{2})$ 和 $k = [tx]$，我们有

$$\begin{aligned} f^{**}(x, \theta_0) &= t\mathbb{Q}\left(\left(\frac{k}{t} - \frac{1}{2t}, \frac{k}{t} + \frac{1}{2t}\right]\right) \\ &= f(k/t; \theta_0) + O(t^{-1}) = f([tx]/t; \theta_0) + O(t^{-1}). \end{aligned} \tag{41}$$

这里我们通过以下两个结果推导出 (41) 中的第二个等式：

$$\mathbb{Q}\left(\left(\frac{k}{t} - \frac{1}{2t}, \frac{k}{t} + \frac{1}{2t}\right]\right) = \int_{k/t-1/(2t)}^{k/t+1/(2t)} f(u; \theta_0)du = \frac{1}{t} f(\check{x}; \theta_0), \tag{42}$$

其中，最后一个等式由中值定理得出，并且 $\check{x} \in [k/t - 1/(2t), k/t + 1/(2t)]$，这意味着 $x - 3/(2t) \le \check{x} \le x + 1/(2t)$，且

$$f(\check{x}; \theta_0) = f(x; \theta_0) + O(t^{-1}) = f([tx]/t; \theta_0) + O(t^{-1}). \tag{43}$$

回顾 $A_\zeta = (-1/\sqrt{2} + \zeta, 1/\sqrt{2} - \zeta)$ 对于 $\zeta > 0$。对于 $x \in A_\zeta$，我们有

$$\begin{aligned} \frac{d\mathbb{Q}}{d\mathbb{Q}^{**}} &= \frac{f(x; \theta_0)}{f^{**}(x; \theta_0)} \\ &= \frac{f(x; \theta_0)}{f([tx]/t; \theta_0) + O(t^{-1})} \\ &= 1 + O(t^{-1}), \end{aligned} \tag{44}$$

其中第二个等式由（41）得出，而最后一个等式来自

$$x - \frac{1}{2t} \le \frac{[tx]}{t} \le x + \frac{1}{2t}.$$

将 $\frac{d\mathbb{Q}}{d\mathbb{Q}^{**}}$ 的表达式代入（44），并且在 A_ζ 上，$\frac{d\mathbb{Q}}{d\mathbb{Q}^{**}}$ 有界。因此，根据控制收敛定理可知 (44)

$$\limsup_{t\to\infty} E_{\mathbb{Q}}\left[1_{A_\zeta}\log\frac{d\mathbb{Q}}{d\mathbb{Q}^{**}}\right]=0. \tag{45}$$

与 (18) 中相同的计算可得

$$\mathbb{Q}(A_\zeta^c)=\int_{A_\zeta^c} f(x;\theta_0)dx=\int_{|x|>1/\sqrt{2}-\zeta} f(x;\theta_0)dx=\frac{2^{3/4}\zeta^{1/2}}{\pi}+O(\zeta^{3/2}). \tag{46}$$

将不等式 (22) 应用于 $\mathbb{Q}$ 和 $\mathbb{Q}^{**}$，并利用 (45)-(46)，可得

$$\begin{aligned}&\limsup_{t\to\infty}\|\mathbb{Q}-\mathbb{Q}^{**}\|_{TV}^2\le\limsup_{t\to\infty}\left\{2\mathbb{Q}(A_\zeta^c)+E_{\mathbb{Q}}\left[1_{A_\zeta}\log\frac{d\mathbb{Q}}{d\mathbb{Q}^{**}}\right]\right\}\\&=\frac{2^{2+3/4}\zeta^{1/2}}{\pi}+O(\zeta^{3/2}).\end{aligned}$$

由于上述不等式的左端与 ζ无关，我们令 $\zeta\to 0$，得到

$$\limsup_{t\to\infty}\|\mathbb{Q}-\mathbb{Q}^{**}\|_{TV}=0,$$

由此完成了 (40) 的证明。

第三，我们推导出 $\mathfrak{I}(\mathbb{Q}^*,\mathbb{P}_{1t})$ 的一个下界。用 (λ_{jk}) 表示一个随机矩阵，它代表一个马尔可夫核，并满足对所有 (j,k) 都有 $\lambda_{jk}\ge 0$

$\sum_j \lambda_{jk} = 1$ 对所有 k 均成立。于是我们有

$$
\begin{aligned}
\beth(\mathbb{Q}^*, \mathbb{P}_{1t}) &= \inf_{(\lambda_{jk})} \sum_k \left| p(k,t;\theta_0) - \sum_j \lambda_{jk} \mathbb{Q}\left(\left(\frac{j}{t} - \frac{1}{2t}, \frac{j}{t} + \frac{1}{2t} \right] \right) \right| \\
&= \inf_{(\lambda_{jk})} \sum_k \frac{1}{t} \left| f(k/t;\theta_0)[1 + H(k/t,t)] - \sum_j \lambda_{jk} f(j/t;\theta_0) \right| + O(t^{-1}) \\
&\geq \inf_{\lambda(y,x)} \int_{-1/\sqrt{2}}^{1/\sqrt{2}} \left| f(x;\theta_0)[1 + H(x,t)] - \int_{-1/\sqrt{2}}^{1/\sqrt{2}} \lambda(y,x) f(y;\theta_0) dy \right| dx + O(t^{-1}) \\
&\geq \inf_{\lambda(y,x)} \int_{\Omega_t} \left| f(x;\theta_0)[1 + H(x,t)] - \int_{-1/\sqrt{2}}^{1/\sqrt{2}} \lambda(y,x) f(y;\theta_0) dy \right| dx + O(t^{-1}) \\
&\geq \int_{\Omega_t} f(x;\theta_0) |H(x,t)| dx + O(t^{-1}) \\
&= \frac{1}{2} \int_{-1/\sqrt{2}}^{1/\sqrt{2}} f(x;\theta_0) |H(x,t)| dx + O(t^{-1}) \\
&\geq \frac{1}{2\pi} \int_{-1/\sqrt{2}}^{1/\sqrt{2}} |H(x,t)| dx + O(t^{-1}), \qquad (47)
\end{aligned}
$$

其中，$H(x,t)$ 定义见 (26)，$\lambda(y,x)$ 表示一个满足以下条件的核函数：对所有 (y,x)，$\lambda(y,x) \geq 0$；对所有 x，$\Omega_t = \{x \in (-1/\sqrt{2}, 1/\sqrt{2}) : H(x,t) \leq 0\}$。第一个数组中的等号使用了参数空间为单点参数空间 Θ_0 时 $\beth(\mathbb{Q}^*, \mathbb{P}_{1t})$ 的定义；第二个数组中的等号由 (24) 和 (42)-(43) 得出；第三个数组中的不等式使用了积分近似，以及核函数 $\lambda(y,x)$ 覆盖了 λ_{jk} 的离散情形；第四个数组中的不等式是因为 $\Omega_t \subset (-1/\sqrt{2}, 1/\sqrt{2})$；第五个数组中的不等式源于 $f(x;\theta_0)$ 的凹性，这意味着

$$
f(x;\theta_0) - \int_{-1/\sqrt{2}}^{1/\sqrt{2}} \lambda(y,x) f(y;\theta_0) dy \leq 0,
$$

第六个数组中的等号可由以下结果得出:

$$
\begin{aligned}
1 &= \sum_k p(k,t;\theta_0) = \sum_k \{f(k/t;\theta_0)[1+H(k/t,t)]/t + O(t^{-2})\} \\
&= \int_{-1/\sqrt{2}}^{1/\sqrt{2}} f(x;\theta_0)[1+H(x,t)]dx + O(t^{-1}) \\
&= 1 + \int f(x;\theta_0)H(x,t)dx + O(t^{-1}) \\
&= 1 - \int_{\Omega_t} f(x;\theta_0)|H(x,t)|dx + \int_{\Omega_t^c} f(x;\theta_0)|H(x,t)|dx + O(t^{-1}), \\
&\int_{\Omega_t} f(x;\theta_0)|H(x,t)|dx = \int_{\Omega_t^c} f(x;\theta_0)|H(x,t)|dx + O(t^{-1}) \\
&= \int_{\Omega_t^c} f(x;\theta_0)|H(x,t)|dx + \frac{1}{2}\int f(x;\theta_0)H(x,t)dx + O(t^{-1}) \\
&= \frac{1}{2}\int f(x;\theta_0)|H(x,t)|dx + O(t^{-1}),
\end{aligned}
$$

最后一个不等式由.....得出

$$
f(x;\theta_0) = \frac{1}{\pi(1-x^2)\sqrt{1-2x^2}} \geq \frac{1}{\pi}.
$$

下面我们将证明

$$
\liminf_{t\to\infty} \int_{-1/\sqrt{2}}^{1/\sqrt{2}} |H(x,t)|dx > 0.
$$

定义 x_ℓ, $\ell = 0, \pm 1, \cdots, \pm L_t$, 为区间 $(-1/\sqrt{2}, 1/\sqrt{2})$ 内满足以下条件的所有点构成的序列

$$
2[\omega_{\alpha_{x_\ell}} + x_\ell \alpha_{x_\ell}]t - \beta_{x_\ell} = \pi\ell,
$$

并且我们选取一点 $\xi_\ell \in (x_\ell, x_{\ell+1})$, 使得

$$
2[\omega_{\alpha_{\xi_\ell}} + \xi_\ell \alpha_{\xi_\ell}]t - \beta_{\xi_\ell} \in [\pi\ell + \pi/4, \pi\ell + 3\pi/4].
$$

由于 $x_\ell, \xi_\ell \in (-1/\sqrt{2}, 1/\sqrt{2})$，我们得出，对于较大的 t、$L_t \propto t$，以及

$$
\begin{aligned}
&2[\omega_{\alpha_{x_{\ell+1}}} + x_{\ell+1}\alpha_{x_{\ell+1}}]t - \beta_{x_{\ell+1}} - 2[\omega_{\alpha_{x_\ell}} + x_\ell\alpha_{x_\ell}]t - \beta_{x_\ell} = \pi,\\
&2[\omega_{\alpha_{\xi_\ell}} + \xi_\ell\alpha_{\xi_\ell}]t - \beta_{\xi_\ell} - 2[\omega_{\alpha_{x_\ell}} + x_\ell\alpha_{x_\ell}]t - \beta_{x_\ell} \in [\pi/4, 3\pi/4],\\
&\{2[\dot\omega_{\alpha_{x_\ell}} + x_\ell\dot\alpha_{x_\ell} + \alpha_{x_\ell}]t - \dot\beta_{x_\ell}\}(x_{\ell+1} - x_\ell) = \pi + o(1),\\
&\{2[\dot\omega_{\alpha_{x_\ell}} + x_\ell\dot\alpha_{x_\ell} + \alpha_{x_\ell}]t - \dot\beta_{x_\ell}\}(\xi_\ell - x_\ell) \in [\pi/4, 3\pi/4] + o(1),\\
&x_{\ell+1} - x_\ell = \frac{\pi}{|2[\dot\omega_{\alpha_{x_\ell}} + x_\ell\dot\alpha_{x_\ell} + \alpha_{x_\ell}]t - \dot\beta_{x_\ell}|} + o(t^{-1}),\\
&\xi_\ell - x_\ell \in \frac{1}{|2[\dot\omega_{\alpha_{x_\ell}} + x_\ell\dot\alpha_{x_\ell} + \alpha_{x_\ell}]t - \dot\beta_{x_\ell}|}[\pi/4, 3\pi/4] + o(t^{-1}).
\end{aligned}
$$

注意到，$H(x,\ t)$中的正弦函数在 x_ℓ 处等于0，在 ξ_ℓ 处则大于 $1/2$ 或小于 $-1/2$（取决于其值为正还是负）；其在 x_ℓ 与 $x_{\ell+1}$ 之间的面积小于由三个点（x_ℓ，0）、（$x_{\ell+1}$，0）和（ξ_ℓ，$\pm 1/2$）构成的三角形面积（根据水平轴上方或下方的面积而定）。因此，我们得到

$$
\begin{aligned}
&\int_{-1/\sqrt{2}}^{1/\sqrt{2}} |H(x,t)|dx \geq \sum_{\ell=-L_t}^{L_t} \frac{|\xi_\ell|}{4}(x_{\ell+1} - x_\ell)\\
&= \sum_{\ell=-L_t}^{L_t} \frac{|x_\ell|}{4} \frac{\pi}{|2[\dot\omega_{\alpha_{x_\ell}} + x_\ell\dot\alpha_{x_\ell} + \alpha_{x_\ell}]t - \dot\beta_{x_\ell}|} + o(1)\\
&= \int_{-1/\sqrt{2}}^{1/\sqrt{2}} \frac{|x|}{4} \frac{\pi}{|2[\dot\omega_{\alpha_x} + x\dot\alpha_x + \alpha_x]t - \dot\beta_x|} \frac{2L_t + 1}{\sqrt{2}} dx + o(1). \qquad (48)
\end{aligned}
$$

综合这些结果，我们得出

$$
\begin{aligned}
&c_0 = \liminf_{t\to\infty} \beth(\mathbb{Q}^*, \mathbb{P}_{1t}) \geq \frac{1}{2\pi} \liminf_{t\to\infty} \int_{-1/\sqrt{2}}^{1/\sqrt{2}} |H(x,t)|dx\\
&\geq \frac{1}{2\pi} \int_{-1/\sqrt{2}}^{1/\sqrt{2}} \frac{|x|}{4} \liminf_{t\to\infty} \left[\frac{\pi}{|2[\dot\omega_{\alpha_x} + x\dot\alpha_x + \alpha_x]t - \dot\beta_x|} \frac{2L_t + 1}{\sqrt{2}} \right] dx\\
&\geq \frac{1}{8\pi} \min_{-1\leq\sqrt{2}x\leq 1} \{|\dot\omega_{\alpha_x} + x\dot\alpha_x + \alpha_x|\} \int_{-1/\sqrt{2}}^{1/\sqrt{2}} \left| \frac{x}{\dot\omega_{\alpha_x} + x\dot\alpha_x + \alpha_x} \right| dx > 0, \qquad (49)
\end{aligned}
$$

w这里第一个不等式来自（47），第二个不等式源于 (48)

法图引理，最后一个不等式使用了以下结果：

$$\liminf_{t\to\infty}\left[\frac{\pi t}{|2[\dot{\omega}_{\alpha_x}+x\dot{\alpha}_x+\alpha_x]t-\dot{\beta}_x|}\right]=\frac{\pi}{|2[\dot{\omega}_{\alpha_x}+x\dot{\alpha}_x+\alpha_x]|},$$

$$2L_t+1\geq\frac{\sqrt{2}}{\pi}\min\{|2[\dot{\omega}_{\alpha_{x_\ell}}+x_\ell\dot{\alpha}_{x_\ell}+\alpha_{x_\ell}]t-\dot{\beta}_{x_\ell}|:-L_t\leq\ell\leq L_t\}$$

$$\geq\frac{\sqrt{2}}{\pi}\min\left\{|2[\dot{\omega}_{\alpha_x}+x\dot{\alpha}_x+\alpha_x]t-\dot{\beta}_x|:x\in\left(-1/\sqrt{2},1/\sqrt{2}\right)\right\},$$

$$\liminf_{t\to\infty}\frac{L_t}{t}\geq\frac{1}{\sqrt{2}\pi}\liminf_{t\to\infty}\frac{\min\{|2[\dot{\omega}_{\alpha_x}+x\dot{\alpha}_x+\alpha_x]t-\dot{\beta}_x|:-1\leq\sqrt{2}x\leq1\}}{t}$$

$$\geq\frac{\sqrt{2}}{\pi}\min\left\{|\dot{\omega}_{\alpha_x}+x\dot{\alpha}_x+\alpha_x|:x\in\left(-1/\sqrt{2},1/\sqrt{2}\right)\right\}.$$

另一种下界：我们可以不用（49）中的下界，而是如下得到 c_0 的一个显式下界：

$$c_0\geq\frac{1}{2\pi}\liminf_{t\to\infty}\int_{-1/\sqrt{2}}^{1/\sqrt{2}}|H(x,t)|dx\geq\frac{1}{8\sqrt{2}\pi}.$$

事实上，利用（48）中的第一个不等式，我们有

$$\int_{-1/\sqrt{2}}^{1/\sqrt{2}}|H(x,t)|dx\geq\sum_{\ell=-L_t}^{L_t}\frac{|\xi_\ell|}{4}(x_{\ell+1}-x_\ell)\geq\frac{1}{4\sqrt{2}}+O(t^{-1}),$$

其中，最后一个不等式源于以下结果：

$$\xi_\ell-(x_\ell+x_{\ell+1})/2=\xi_\ell-x_\ell-(x_{\ell+1}-x_\ell)/2=O(t^{-1}),$$

$$\sum_{x_\ell\geq0}\frac{|\xi_\ell|}{4}(x_{\ell+1}-x_\ell)=\frac{1}{8}\sum_{x_\ell\geq0}(x_\ell+x_{\ell+1})(x_{\ell+1}-x_\ell)+O(t^{-1})$$

$$=\frac{1}{8}\sum_{x_\ell\geq0}(x_{\ell+1}^2-x_\ell^2)+O(t^{-1})=\frac{1}{8\sqrt{2}}+O(t^{-1}),$$

$$\sum_{x_{\ell+1}\leq0}\frac{|\xi_\ell|}{4}(x_{\ell+1}-x_\ell)=\frac{1}{8}\sum_{x_{\ell+1}\leq0}-(x_\ell+x_{\ell+1})(x_{\ell+1}-x_\ell)+O(t^{-1})$$

$$=\frac{1}{8}\sum_{x_{\ell+1}\geq0}(x_\ell^2-x_{\ell+1}^2)+O(t^{-1})=\frac{1}{8\sqrt{2}}+O(t^{-1}).$$

最后，将 (39)-(40) 和 (49) 代入 (38)，即可证明 (9)。

A6. (10) 和 (11) 的证明

$\mathbb{P}_{1t\eta}$ $X_t^q(\eta)$ 的分布是 X_t^q 的分布 $\mathbb{P}_{1t}$ 的平均值，因此， $\mathbb{P}_{1t\eta}$ 是 $\mathbb{P}_{1t}$ 与离散均匀分布的卷积。由于卷积是马尔可夫核，我们有 $\beth(X_t^q,\ X_t^c(\epsilon)) \leq \beth(X_t^q(\eta)\ X_t^c(\epsilon))$，其上界为 $\mho(X_t^q(\eta)\ X_t^c(\epsilon))$。因此，定理1中的结果 (8) 可立即推出 (10)。

$\beth$ 和 $\mho$ 的定义表明

$$\mho(X_t^q, X_t^c(\epsilon)) = \max\{\beth(X_t^c(\epsilon), X_t^q), \beth(X_t^q, X_t^c(\epsilon))\},$$

因此，(9) 和 (10) 共同推出 (11)。

A7. (12) 的证明

只需证明 $\beth(\mathbb{P}_{1t\eta},\ \mathbb{P}_{2t})$ 和 $\beth(\mathbb{P}_{2t},\ \mathbb{P}_{1t\eta})$ 的阶均为$\eta^{1/2}$，其中 $t \to \infty$。根据 $\mathbb{P}_{2t}$ 的定义，我们有

$$\begin{aligned}
&\limsup_{t\to\infty} \|\mathbb{P}_{1t\eta}^* - \mathbb{P}_{2t}\|_{TV} \leq \limsup_{t\to\infty} \limsup_{\epsilon\to 0} \|\mathbb{P}_{1t\eta}^* - \mathbb{P}_{2t\epsilon}\|_{TV} \\
&\leq \limsup_{\epsilon\to 0} \limsup_{t\to\infty} \|\mathbb{P}_{1t\eta}^* - \mathbb{P}_{2t\epsilon}\|_{TV} \\
&= O(\eta^{1/2}),
\end{aligned}$$

其中，最后一个等式由 (33) 得出。类似地，我们有

$$\begin{aligned}
&\limsup_{t\to\infty} \|\mathbb{P}_{1t\eta} - \mathbb{P}_{2t}^*\|_{TV} \leq \limsup_{t\to\infty} \limsup_{\epsilon\to 0} \|\mathbb{P}_{1t\eta} - \mathbb{P}_{2t\epsilon}^*\|_{TV} \\
&\leq \limsup_{\epsilon\to 0} \limsup_{t\to\infty} \|\mathbb{P}_{1t\eta} - \mathbb{P}_{2t\epsilon}^*\|_{TV} \\
&= O(\eta^{1/2}),
\end{aligned}$$

其中，最后一个等式由 (36) 得出。

利用上述两个结果，我们采用与从 (33) 和 (36) 推导 (35) 和 (37) 时相同的论证，得到 $t \to \infty$, $\beth(\mathbb{P}_{1t\eta},\ \mathbb{P}_{2t})$ 和$\beth(\mathbb{P}_{2t},\ \mathbb{P}_{1t\eta})$ 的阶为 $\eta^{1/2}$。

A8. (13) 的证明

式 (12) 的证明表明

$$\limsup_{t\to\infty} \|\mathbb{P}_{1t\eta}^* - \mathbb{P}_{2t}\|_{TV} = O(\eta^{1/2}),$$

再次将 $\mathbb{P}^*_{1t\eta}$ 表示为 $\mathbb{P}_{1t}$ 与均匀分布的卷积，我们得到

$$\limsup_{t\to\infty} \beth(\mathbb{P}_{1t}, \mathbb{P}_{2t}) = O(\eta^{1/2}).$$

由于上述等式的左端与 η 无关，我们令 $\eta \to 0$ ，以证明 (13) 的第一个结论。对于第二个结论，我们有

$$\beth(X_t^c, X_t^q) \geq \beth(X_t^c(\epsilon), X_t^q) - \mho(X_t^c, X_t^c(\epsilon)).$$

因此，我们如下证明 (13) 的第二个结论：

$$\begin{aligned}
&\liminf_{t\to\infty} \beth(X_t^c, X_t^q) \geq \liminf_{t\to\infty} \beth(X_t^c(\epsilon), X_t^q) - \limsup_{t\to\infty} \mho(X_t^c, X_t^c(\epsilon)).\\
&\geq c_0 - O(\epsilon^{1/2}) \to c_0 \text{ as } \epsilon \to 0,
\end{aligned}$$

其中我们注意到左端与 ϵ 无关，第二个不等式由 (11) 以及.....得出

$$\limsup_{t\to\infty} \mho(X_t^c, X_t^c(\epsilon)) = O(\epsilon^{1/2}).$$

要证明上述结论，只需证明

$$\limsup_{t\to\infty} \|\mathbb{P}_{2t} - \mathbb{P}_{2t\epsilon}\|_{TV} = O(\epsilon^{1/2}).$$

事实上，我们有

$$\begin{aligned}
&\limsup_{t\to\infty} \|\mathbb{P}_{2t} - \mathbb{P}_{2t\epsilon}\|_{TV} \leq \limsup_{t\to\infty} \limsup_{\epsilon'\to 0} \|\mathbb{P}_{2t\epsilon'} - \mathbb{P}_{2t\epsilon}\|_{TV}\\
&\leq \limsup_{\epsilon'\to 0} \limsup_{t\to\infty} \|\mathbb{P}_{2t\epsilon'} - \mathbb{P}_{2t\epsilon}\|_{TV}\\
&= O(\epsilon^{1/2}),
\end{aligned}$$

其中第一个不等式由 $\mathbb{P}_{2t}$ 的定义得出，最后一个等号则利用了 (19) 表明.....这一事实

$$\limsup_{t\to\infty} \|\mathbb{P}_{2t\epsilon'} - \mathbb{P}_{2t\epsilon}\|_{TV} = O(\epsilon^{1/2} + (\epsilon')^{1/2}).$$