arXiv:2303.00209v1 [quant-ph] 1 Mar 2023

使用经典-量子混合内存进行学习的内存-样本下界

刘启鹏 * Ran Raz † Wei Zhan ‡

摘要

Raz 的一项工作（《ACM期刊》和 FOCS 16）证明，对于在 n 比特上进行奇偶学习的任何算法，都需要 $\Omega(n^2)$ 比特的经典内存，或数量呈 n 指数增长的随机样本。近期的一系列工作延续了这一研究方向，并表明，对于大量经典学习任务，要么需要超线性的经典内存，要么需要超多项式数量的样本。所有这些工作都将学习算法视为经典分支程序，即在有界内存中执行经典计算。

然而，这些结果并未涵盖所有物理计算模型，尤其是量子计算机以及量子内存的使用。这留下了一种可能性：一小部分量子内存就可能显著减少对经典内存或样本的需求，从而彻底改变经典学习任务的性质。尽管近期已有研究探讨了阴影层析和纯度测试等本质量子学习问题对量子内存的必要性，但量子内存在经典学习任务中的作用仍不明确。

在本研究中，我们探讨了存在量子存储器时的经典学习任务。我们证明，对于在 n 比特上进行奇偶学习的任何同时使用经典内存和量子存储器的量子算法，要么需要 $\Omega(n^2)$ 比特的经典内存，要么需要 $\Omega(n)$ 比特的量子存储器，要么需要指数数量的样本。换言之，即使学习算法除了使用经典内存外，还能使用大小为 cn 的量子存储器（其中 $c > 0$ 为某个常数），奇偶学习的内存-样本下界在定性上仍然相同。

我们的结果更为一般，适用于许多其他经典学习任务。沿用以往工作的做法，我们用矩阵 M：$A \times X \to \{-1, 1\}$ 表示以下学习任务。一个未知的 x 从概念类 X中均匀随机抽取，学习算法通过流式接收随机样本 $(a_i, b_i = M(a_i, x))$，尝试揭示 x；对于每个 i，$a_i \in A$ 均为均匀随机选取。假设 k、ℓ、r 为整数，并且 M 的任意一个至少有 $2^{-k} \cdot |A|$ 行、至少有 $2^{-\ell} \cdot |X|$ 列的子矩阵，其偏差至多为 2^{-r}。我们证明，对于与 M 对应的学习问题，任何使用经典与量子混合内存的算法若要达到至少为 $2^{-O(r)}$ 的成功概率，就必须满足以下条件之一：（1）需要 $\Omega(k \cdot \ell)$ 比特的经典内存；（2）需要 $\Omega(r)$ 量子比特的量子存储器；或（3）需要 $2^{\Omega(r)}$ 个随机样本。

我们的结果排除了这样一种可能性：少量量子存储器能够显著减少在这些问题上进行高效学习所需的经典内存大小。我们的结果还意味着，在有界存储模型中，若干现有密码协议的安全性得到提升（这些协议基于对 n 比特的奇偶学习）；我们证明，即使存在一个至多拥有 cn^2 比特经典内存和 cn 比特量子存储器的量子对手，安全性仍然成立（其中 $c > 0$ 为某个常数）。

*西蒙斯计算理论研究所。电子邮箱：qipengliu0@gmail.com。研究部分获西蒙斯计算理论研究所量子博士后奖学金、美国国防高级研究计划局 SIEVE-VESPA 资助（编号：HR00112020023）以及美国国家科学基金会量子信息科学与工程研究所计划资助（项目编号：OMA-2016245）。

†普林斯顿大学。电子邮箱：ranr@cs.princeton.edu。研究获西蒙斯研究员奖以及美国国家科学基金会项目资助，项目编号为 CCF-1714779 和 CCF-2007462。

‡普林斯顿大学。电子邮箱：weizhan@cs.princeton.edu。研究获西蒙斯研究员奖以及美国国家科学基金会项目资助，项目编号为 CCF-1714779 和 CCF-2007462。

1 引言

内存在学习中发挥着重要作用。从 Shamir [Sha14]以及 Steinhardt、Valiant 和 Wager 的奠基性工作 [SVW16], 开始，一系列工作开启并深化了对记忆限制下学习下界的研究。Steinhardt、Valiant 和 Wager [SVW16]猜想，为了从随机子集奇偶性的样本中学习一个未知的 n 比特字符串，算法要么需要大小关于 n 为二次的内存，要么需要指数数量的随机样本（同样相对于 n）。该猜想后来由 Raz [Raz18], 证明，这首次表明，对于某些学习问题，高效学习需要超线性的内存大小。该结果随后被推广到一大类学习问题 [KRT17, Raz17, MM18, BGY18, GRT18, SSV19, GKR20, GKLR21]。

尽管我们已经全面理解了在特定计算资源（内存）受限的情况下进行学习的（不可）行性，但上述先前工作并未涵盖所有物理计算模型；其中最显著的是量子计算以及量子存储器的能力。许多研究人员相信，大规模量子计算机最终将成为现实。近期实验展示了量子优势，例如 [AAB+19]，并表明实现量子存储器和量子计算机可能不存在根本性障碍。关于量子存储器在学习中所起作用的问题，Wright 在一般态层析的背景下提出了这一问题 [Wri16]，Aaronson 则针对阴影层析提出了这一问题 [Aar18]。一系列工作 [HHJ+16、BCL20、HKP20、CCHL22、ACQ22、CLO22] 开创了这一思路，并展示了在有量子存储器与无量子存储器的情况下进行学习时的多项式分离或指数分离，但这些工作所处理的都是内在的量子学习任务，例如态层析、阴影层析和纯度测试。

鉴于上述情况，研究在量子存储器以及混合经典—量子记忆存在的条件下进行经典学习任务，是很有吸引力的。上述所有经典结果的直接推论只能给出平凡结论。由于 k 量子比特的内存总能被高效地模拟为 $\sim 2^k$ 经典比特，我们只能得出这样的结论（例如，对于奇偶学习）：要么需要 $\sim 2\log\, n$ 量子比特的量子存储器，要么需要指数多的样本。在我们的工作之前，即使只有非常小容量的量子存储器可用，也可能显著减少对经典内存的需求，并带来一种高效的学习算法。

在这项工作中，我们证明了在存在混合内存的情况下，许多经典学习问题的内存-样本下界。与 [Raz17, GRT18], 一样，我们用一个矩阵 M 来表示学习问题：$A \times X \to \{-1, 1\}$ 矩阵的列对应概念类 X 中的概念，行对应随机样本。在学习任务中，一个未知概念 $x \in X$ 被均匀随机地抽取，每个随机样本的形式为 $(a_i, b_i) = (a_i, M(a_i, x))$，其中 $a_i \in A$ 是均匀随机选取的。学习者的目标是找出 x。[GRT18], 证明了，当底层矩阵 M 是一个误差为 2^{-r} 的 (k, ℓ)-L_2 双源提取器[1] 时，要使该学习任务的成功概率至少达到$2^{-O(r)}$，学习算法需要 $\Omega(k \cdot \ell)$ 比特的内存，或者需要 $2^{\Omega(r)}$ 个样本。

1.1 我们的结果

在这项工作中，我们将量子学习算法建模为一个程序，其混合内存由 q 个量子比特的量子存储器和 m 个比特的经典内存组成。在每个阶段，算法都会获得一个随机样本（a_i，$b_i = M$（a_i，x））。量子学习算法会根据该随机样本，对混合内存施加任意量子信道。尽管该量子信道

[1]粗略地说，这意味着，M 的每个子矩阵 M'，只要其行数至少为 $2^{-k}|A|$、列数至少为 $2^{-\ell}|X|$，相对偏差就至多为 2^{-r}。

可以是任意的，但我们要求其输出为一个至多包含 q 个量子比特和 m 个比特的混合经典-量子态。我们强调，对量子信道的复杂度没有任何限制（而且这只会使我们的结果更有力，因为我们在这里证明的是此类算法的下界）。

基于上述模型，我们给出以下主定理。

定理 1（主定理，非正式表述）。设 $M\colon A\times X\to\{-1,1\}$ 是一个矩阵。若 M 是一个误差为 2^{-r} 的 (k,ℓ)-L_2双源提取器，则量子学习算法至少需要满足以下条件之一：

1. $\Omega(k\cdot\ell)$ 比特的经典内存；或者，
2. $\Omega(r)$ 量子比特的量子存储器；或者，
3. $2^{\Omega(r)}$个样本，

才能在相应的学习任务中以至少 $2^{-O(r)}$ 的概率成功 k.

我们的主定理表明，对于许多学习问题，即使可以使用大小至多为 $\Omega(r)$ 的量子存储器，也不会减少所需经典内存的大小或样本数量。由于相干量子存储器对于近期的中等规模量子计算机而言颇具挑战性，即便量子计算机未来得到广泛应用，它也可能依然成本高昂，因此量子内存对这些学习问题的影响更加有限。

为了更精确地阐述该定理，我们以奇偶学习为例。上述定理表明，要进行高效学习，量子学习算法要么需要 $\Omega(n^2)$ 比特的内存，要么需要 $\Omega(n)$ 个量子存储器量子比特；否则，就需要 $2^{\Omega(n)}$ 随机样本。乍看之下，对量子存储器的限制似乎微不足道：如果目标是学习一个 n 比特的未知秘密，那么似乎总是需要线性规模的内存来存储该秘密。然而，注意到我们的主定理适用于使用混合内存的量子学习算法，并排除了使用 $n^2/1000$ 比特和 $n/1000$ 个量子比特混合内存进行奇偶学习的算法，因此主定理给出了非平凡且有力的内存-样本下界。还要注意，即使目标仅仅是输出秘密中的一位，我们的结果（以及此前的结果）仍然成立。目前，我们不知道主定理是否是紧的。对于奇偶学习，我们尚未发现任何仅使用 $O(n)$ 个量子存储器量子比特的量子学习算法。我们将弥合这一差距留作一个引人入胜的开放问题。

主定理自然适用于 [GRT18], 中讨论的其他学习问题，包括稀疏奇偶学习、从稀疏线性方程中学习，以及许多其他问题。此处不一一列举，详情请参阅 [GRT18] 。

在此过程中，我们提出了一种证明经典存储器-样本下界的新方法，称为“坏度等级”方法。该方法在技术上等同于 [Raz17、GRT18] 中先前提出的方法，但概念上更易于使用，而且我们能够将其推广到量子情形。

我们指出，在没有经典内存的情况下，证明量子内存大小的线性下界要简单得多（但据我们所知，在我们的工作之前还没有出现过此类证明）。我们在附录C中给出了这样的证明。在附录C中，我们陈述并证明了定理3 ，该定理给出了量子存储大小（不含经典内存）线性下界的一个更简单的证明。在大多数情况下，定理3 在定性上弱于我们的主要结果，因为它只给出了仅有量子内存、但没有（可能达到二次规模的）经典内存的程序的下界；不过，定理3 在技术上与之不可比较，而且它是以量子提取器而非经典提取器的形式陈述的。此外，定理3的证明比我们主定理的证明简单得多。

对有界存储模型中密码学的启示。由于学习理论与密码学可以被视为一体两面，我们的定理也将有界存储模型中许多现有密码协议（基于奇偶学习的协议）的安全性提升到了量子环境。据我们所知，这是首次证明经典密码协议能够抵御空间受限量子攻击者。[2] 我们将在下文进一步阐述。

（经典）有界存储模型中的密码学最早由 Maurer 提出 [Mau92]。在这种模型中，无需任何计算假设。诚实方通过长比特流序列执行协议。窃听者的存储能力有限，无法存储完整对话，因此无法破解协议。一系列研究 [CM97、AR99、ADR02、DR02、DM02、Lu02、DM04、MSTS04、HN06、DQW21、DQW22、.....] 在该模型中构建了高效且安全的密钥协商、不经意传输、比特承诺和时间戳协议。

基于 n 比特的奇偶学习的内存-样本下界， [Raz18] 在有界存储模型中提出了一种加密方案。Guan 和 Zhandry [GZ19] 提出了密钥协商、不经意传输和比特承诺方案，在面对内存最多为 $O(n^2)$比特的攻击者时，实现了更少的轮数和更好的正确性。沿用类似思路，Liu 和 Vusirikala[LV21] 证明了面对内存最多为 $O(n^2)$比特的攻击者时，可以实现半诚实多方计算。最近，Dodis、Quach 和 Wichs [DQW22]研究了基于奇偶学习的有界存储模型中的消息认证。我们关于奇偶学习的结果直接推广了上述所有结果。当密码协议基于 n 比特的奇偶学习（通常被视为安全参数）时，我们的结果表明，即使存在经典内存最多为 $O(n^2)$比特、量子存储器最多为 $O(n)$量子比特的量子对手，安全性仍然成立。

尽管此前已有许多关于量子有界存储模型中密码学的研究 [DFR+07，DFSS07、Sch07、DFSS08、WW08、PMLA13、BY21], 但它们都依赖于流式处理量子态。我们的内存-样本下界首次给出了一大类经典密码学方案（密钥协商、不经意传输和比特承诺），并证明了它们能够抵御空间受限量子攻击者。

2 证明概述

2.1 经典下界的证明回顾

由于我们的证明建立在此前关于学习的经典存储器-样本下界研究之上，具体而言，建立在 [Raz17、GRT18], 的证明技术之上，我们将简要回顾这些证明，并以奇偶学习 [Raz18] 为例。在下文中， $M(a,\ x)$ 表示 a和 x 在 $\mathbb{F}_2$中的内积。

考虑一个经典分支程序，它试图通过样本 (a,b) 学习一个未知且均匀随机的$x \in \{0,1\}^n$，其中 $a \in \{0,1\}^n$是均匀随机的，而 $b = M(a,x)$。我们可以将分支程序的每个量子态 v 与 $\{0,1\}^n$上的一个分布 $P_{X|v}$ 关联起来，该分布表示以到达该量子态为条件时 x 的分布。在初始量子态，由于对 x一无所知，分布是均匀的（其 ℓ_2范数尽可能小）。沿着分支程序的一条计算路径，分布 $P_{X|v}$ 逐渐演化，并最终变得集中（具有较大的 ℓ_2范数），以便正确输出 x 。因此，在演化过程中，$P_{X|v}$ 应当在某个阶段具有略大的 ℓ_2范数（对于某个较小的常数 $2^{\varepsilon n}$ ，比均匀分布大

[2]另一方面，已有一些经典安全的有界存储协议示例，而使用数量少一个指数级的量子存储器即可将其攻破。[GKK+08]。

倍）。如果我们将这样的分布设为目标分布，那么仅凭随机样本很难达到该分布。分支程序只有以 $2^{-\Omega(n)}$ 的概率才能朝目标取得显著进展；而大多数时候，一个样本只是将分布（当前分布和目标分布）各自分成均等的两部分，这对于接近目标分布（具有较大的 ℓ_2 范数）帮助不大。

更严谨地说，我们考察内积的演化

$$\langle P_{X|v}, P\rangle = \sum_{x\in\{0,1\}^n} P_{X|v}(x)\cdot P(x)$$

即当前状态 v 上的分布 $P_{X|v}$ 与目标分布 P 之间的内积。接收到一个样本（a，b）意味着 $M(a,x)=b$，因此只有 $P_{X|v}$ 中支撑在此类 x 上的部分会继续传播。如果这部分接近 $1/2$ 概率，我们就称 a 将 $P_{X|v}$ 均匀地分割。将新的分布记为 $P^{(a,b)}_{X|v}$，经过适当归一化后，新的内积为

$$\langle P^{(a,b)}_{X|v}, P\rangle = \sum_{\substack{x\in\{0,1\}^n\\ M(a,x)=b}} P_{X|v}(x)\cdot P(x) \Big/ \sum_{\substack{x\in\{0,1\}^n\\ M(a,x)=b}} P_{X|v}(x). \tag{1}$$

理想情况下，$P_{X|v}$ 和逐点乘积向量 $P_{X|v}\cdot P$ 都应具有相对较小的 ℓ_2 范数。由于 M 的提取器性质，$a\in\{0,1\}^n$ 中的大多数应能将这两个向量均匀地分割，因此分母接近 $1/2$，而分子接近 $\frac{1}{2}\langle P_{X|v}, P\rangle$。这意味着，给定一个均匀随机的 a，我们在内积上取得的进展有限。另一方面，从具有均匀分布的 $\langle U, P\rangle = 2^{-n}$ 到 $\langle P, P\rangle = 2^{2\varepsilon n}\cdot 2^{-n}$，分支程序需要经过多步推进。因此，这种情况发生的概率极小。

为确保上述论述顺利进行，我们要求分支程序中的每个量子态 v 都满足以下性质：

- ℓ_2-范数 $\left\|P_{X|v}\right\|_2$ 较小。
- ℓ_2-范数 $\left\|P_{X|v}\cdot P\right\|_2$ 较小；当 ℓ_∞-范数 $\left\|P_{X|v}\right\|_\infty$ 较小时，这一条件便得到满足。
- 方程（1）中的分母对每个样本（a，b）都与 0 保持一定距离。

这些性质本身并不成立。相反，我们会在选择目标分布之前，对分支程序执行截断过程。具体而言，我们修改分支程序，使其在以下情况下停止：

- （ℓ_2 截断）：到达一个 $\left\|P_{X|v}\right\|_2$ 较大的量子态 v；
- （ℓ_∞ 截断）：当未知概念为 x 时，达到具有较大 $P_{X|v}(x)$ 的量子态 v；
- （样本截断）：或者，对于下一个样本（a，b），a 不能被 $P_{X|v}$ 整除。

事实证明，在 ℓ_2 截断之后，另外两个截断步骤会在分支程序的每个阶段额外引入 $2^{-\Omega(n)}$ 误差。因此，证明归结为证明一个关于到达具有较大 $\left\|P_{X|v}\right\|_2$ 的量子态的概率的 $2^{-\Omega(n^2)}$ 界；由此通过标准的并集界，我们可以证明奇偶学习的内存-样本下界：至少需要 $2^{\Omega(n)}$ 样本，或 $\Omega(n^2)$ 比特的内存。

2.2 坏度等级

如上所述，为了界定到达具有较大 ℓ_2-范数的量子态的概率，基本思路是将其分布固定为目标分布 P，并界定内积 $\langle P_{X|v}$ 的增量， $P\rangle$。Raz17 [Raz17 和 GRT18] 通过设计一个势函数来追踪 $\langle P_{X|v}$ 的平均值， $P\rangle^k$ 其中 $k = \Theta(n)$ 中的平均值取遍 v 分支程序同一阶段的量子态。在此，我们提出另一种使用坏度等级概念的方法。尽管在经典情形下，它在技术上等价于势函数方法，但更加灵活，也更容易适用于量子情形。我们将这一方法视为本工作的独立贡献。

我们首先将坏事件定义为一对 (v, a)，其中包含量子态 v 和样本 a的后续部分，满足 $\langle P_{X|v}$, $P\rangle \geq 2^{-n}$，并且对于两种可能结果之一 b,

$$\sum_{\substack{x\in\{0,1\}^n \\ M(a,x)=b}} P_{X|v}(x) \cdot P(x) \geq \left(\frac{1}{2} + 2^{-\delta n}\right) \cdot \langle P_{X|v}, P\rangle \tag{2}$$

其中某个较小的常数 δ。换言之，内积 $\langle P_{X|v}$， $P\rangle$ 足够大，但并未在 a之间均匀分配。根据式（1），我们知道，经过一次坏事件后，内积至多大约翻倍。相反，在良好情形下，内积要么仅乘以一个（$1 + 2^{-\delta n}$）因子，要么已经小于基准值 2^{-n}。此外， M 的提取器性质确保，对于每个量子态 v，在均匀随机选取 a时，发生坏事件的概率至多为 $2^{-\Omega(n)}$ 。

现在，量子态 β（v）的坏度等级 v 记录了计算路径在到达 v之前经过坏事件的次数。[3] 根据上述关于坏事件的观察，可以得到（省略较小的因子）：

- 对于每个量子态 v, $\langle P_{X|v}$, $P\rangle$ 都被 $2^{\beta(v)} \cdot 2^{-n}$所界定；
- 进入下一阶段时， $\beta(v)$以 $2^{-\Omega(n)}$的概率增加 1。

因此，在每个阶段，坏度等级为 β 的量子态总权重至多为 $2^{-\Omega(\beta n)}$。所以，任何满足 $\langle P_{X|v}$, $P\rangle \geq 2^{2\varepsilon n} \cdot 2^{-n}$ 的量子态都必须具有 $2^{-\Omega(n^2)}$ 概率。

2.3 证明量子下界的障碍

在本节中，我们尝试针对纯量子情形证明相同的 $2^{\Omega(n)}$-样本或 $\Omega(n^2)$-量子存储下界。在此过程中，我们指出了证明量子学习算法的内存-样本下界时面临的一些障碍；下一节中，我们将展示如何克服这些障碍，并证明混合学习算法的下界，其中经典内存大小为二次级，量子内存大小为线性级。

沿用上文经典情形证明的相同框架，我们首先需要将所有概念迁移到量子算法中：

- 状态 v 是量子存储器希尔伯特空间中的量子态；
- 分布 $P_{X|v}$ 仍然是良定义的：它是当测量量子存储器得到 x 时的分布 v （参见第3.4 节和式（3））；

[3] 目前我们将 $\beta(v)$视为自然数。在实际证明中， $\beta(v)$是自然数上的一个分布，因为到达同一状态的不同计算路径上，坏事件的数量可能不同。

- 我们仍然能够实施 ℓ_2 截断：如果 $P_{X|v}$ 的 ℓ_2范数很大，就将整个系统投影到 $v^\perp$ 的 v 正交子空间，并重复此过程，直到不存在这样的量子态 v （详情参见第4.1 节）。

- 我们同样能够以类似于 ℓ_2 截断的方式实现样本截断。由于这里的判据取决于 a，我们针对每个 a分别创建当前系统的副本；当每个副本中的 $P_{X|v}$ 不能被 a 整除时，使用投影对 v 量子态进行截断，然后再将它们合并回来。我们证明了这种截断引入的误差很小。

这里出现了第一个主要障碍： ℓ_∞ 截断。在经典情形中， ℓ_∞ 截断是针对每个单独的 x实施的；与之相反，在 ℓ_2 截断中，量子态会被整体移除。利用已知分布的 ℓ_2 范数很小这一事实，并使用马尔可夫不等式，可以证明 ℓ_∞ 截断引入的误差很小。

然而，当我们尝试用量子截断来模拟经典情形下 ℓ_∞ 截断的实现方式时，也就是说，仅在特定 x 对应的 $P_{X|v}(x)$较大时，才将系统投影到 $v^\perp$ ，而不是针对每个 x都进行投影，这可能会导致 $P_{X|u}$ 在与 v不正交的 u上的分布发生巨大变化。下面的例子说明了这种情形：

例子。考虑一个量子学习算法，并假设在计算的某个阶段，对每个 $x \in \{0,1\}^n$，量子存储器都处于某个纯态 $v(x)$。我们在维数为 $d \approx 2^{n/2}$ 的希尔伯特空间中均匀随机地选取每个 $v(x)$，并考虑 $v(x)$ 的典型配置。现在，对于每个量子态 v， ℓ_2-范数都有界：最坏的情况发生在某个 x 下，$v = v(x)$，此时 $\|P_{X|v(x)}\|_2$ 通常约为 $d \cdot 2^{-n}$，接近均匀分布的 ℓ_2-范数。然而，这些最坏分布的 ℓ_∞-范数也接近 $d \cdot 2^{-n}$，远大于均匀分布的 ℓ_∞-范数，因此需要截断。但对于 x ，截去 $v(x)$ 意味着 x 被完全抹去，最终我们会把所有内容都删掉。

接下来，我们固定一个目标态 v ，其目标分布为 $P_{X|v}$ ，且超过 ℓ_2-范数阈值；我们的目标仍然是证明 $2^{-\Omega(n^2)}$ 关于 v 的振幅界。坏事件仍应定义为满足公式（2）的二元组 (v, a)，其中 v 现在是一个量子态。随后我们遇到了第二个主要障碍：目前还不清楚该如何定义坏度等级。

如果我们通过考察历史状态上的坏事件，分别为每个状态 v 定义坏度等级 $\beta(v)$，那么就不清楚该如何衡量坏度等级 β的总权重。在经典情形中，我们只需将坏度等级为 β的状态的总概率定义为总权重。但在这里的量子情形中，这样的定义要么取决于基的选择，而基在每个阶段都可能发生较大变化；要么完全无法推出目标态所需的振幅界。

另一种选择是采用更具操作性的坏度等级定义。确实，一个很诱人的做法是将 β 定义为另一个寄存器，其更新由量子存储器控制。这类定义的问题在于，坏事件（式（2））关于 v并非线性的。因此，将坏度等级定义为线性算子的操作性定义，必然会引入误差，而误差会随着阶段数迅速累积。

2.4 小型量子存储器下的混合存储器下界

上一节中提到的障碍是针对证明二次量子存储器下界的。我们指出，证明线性量子存储器下界（不含经典内存）并不难：

证明完全可以基于信息论：当内存极其有限时，例如只有 $\frac{1}{2}n$ 个量子比特时，尽管内存是量子的，从每个样本中获得的信息仍然小到呈指数级。我们在附录C中给出了这样的证明。

我们在此证明的下界针对混合内存：要使用经典内存和量子存储器学习奇偶校验，算法至少需要 $2^{\Omega(n)}$ 个样本，或 $\Omega(n^2)$ 的经典内存，或 $\Omega(n)$ 的量子存储器（定理 1）。现在我们将介绍如何克服前面提到的障碍。

ℓ_∞ 截断。当只有少量量子存储器且没有经典内存时，对 ℓ_∞ 截断的处理很直接。我们通过将系统投影到正交子空间 $v^\perp$，移除所有具有大 ℓ_∞-范数分布的量子态 v，这与ℓ_2 截断的过程相同。由于 x 上的整体分布是均匀的，任何满足 $\|P_{X|v}\|_\infty \geq 2^{\delta n} \cdot 2^{-n}$的量子态 v 的权重都至多为 $2^{-\delta n}$。因此，只要希尔伯特空间的维度远小于 δn，此次截断引入的误差就很小。[4]

存在经典内存时，实际的 ℓ_∞ 截断步骤（见第 4.2 节，步骤 2）会更加复杂。我们首先对经典内存 ℓ_∞ 应用原始的经典截断 W。现在，对于每个经典内存状态 $\|P_{X|w}\|_\infty$， w都有界，因此我们可以如上所述，通过投影移除 v 具有较大$\|P_{X|v\text{的量子态，}w}\|_\infty$ 。由于经典截断 ℓ_∞ 取决于 x，它可能会改变分布 $P_{X|v\text{。不过，}w}$与经典情形一样， $P_{X|w}$ 不会有太大变化。因此，凡是$P_{X|v\text{发生}w}$ 剧烈变化的地方，其权重必定很小，也可以通过投影移除。这种移除对应于第 G_t 4.2 节中的 截断。

坏度等级。 有趣的是，我们可以完全避免在量子存储器上定义坏度等级所带来的问题，只需将其保留为经典内存的一种属性。为此，我们需要修改坏事件的定义：现在，坏事件是一个二元组（w， a），由经典内存状态 w和样本 a组成，满足 存在某个量子存储器状态 v，使得 $P_{X|v,\ w}$ 并满足式（2）。

对于每个固定的经典内存状态 w，我们仍需确保坏事件发生的概率很小。我们通过证明这一点（引理5.2）：如果存在许多不同的样本 a，且每个样本都对应某个满足式（2）的量子态 v_a，那么就存在某个量子态 v，它能同时满足式（2），并适用于其中大多数 a （这与提取器性质矛盾）。归根结底，这是因为式（2）具有连续性：经过适当的合同变换后，式（2）会变成关于 v上的二次型的一个简单阈值不等式。现在，如果某个 v_a满足该不等式，那么当阈值参数 δ显著减小时，大多数 v 也会满足该不等式，因此 v必然存在一个同时满足条件的量子态。[5] 在这一论证中，我们使用引理3.1，该引理由高斯二次型的反集中界推得，并且关键依赖于维数至多为 $2^{\varepsilon n}$ 某个较小的 ε 这一事实。

另一个技术问题是，为了使用提取器性质，我们需要确保$\langle P_{X|v\text{、}\ w\text{、}}\ P\rangle \geq 2^{-n}$ 对于同时满足 v。因此，我们在引理5.2 中所做的是，首先从概念上移除$\langle P_{X|v\text{、}\ w}$以及 $P\rangle$ 过小的部分，所使用的投影方式类似于截断步骤。移除之后，我们得到一个子空间 $\mathcal{V}'$，其中$\langle P_{X|v\text{、}\ w}$以及 $P\rangle$ 始终具有下界；我们证明，对于每个满足式（ v 的量子态，投影2）后，该不等式仍然近似成立，其中将 v 投影到 $\mathcal{V}'$上。因此，我们仍然可以应用上述论证，并在该子空间内找到一个同时满足 v 的。

[4]上一节中说明无法将 ℓ_∞ 截断同样处理的例子 ℓ_2 在这里不适用，因为它需要 $n/2$ 量子比特内存，而这里的内存规模更小。[5]我们注意到，样本截断的误差界（引理 4.12）也是用这一论证证明的。

3 预备知识

3.1 向量与矩阵

对于向量 $v \in \mathbb{C}^d$ 和 $p \in [1,\infty]$，我们将 v 的 ℓ_p 范数定义为

$$\|v\|_p = \left(\sum_{i=1}^{d} |v_i|^p\right)^{1/p}.$$

对于两个向量 u、$v \in \mathbb{C}^d$，定义它们的内积为 $\langle u, v\rangle = u^\dagger v = \sum_{i=1}^{d} \overline{u_i} v_i$。因此，$\|v\|_2^2 = \langle v, v\rangle$。我们还将集合 $\mathcal{X}$ 上的每个分布 P 视为满足 $\|P\|_1 = 1$的非负实向量。

我们特别使用狄拉克记号表示单位向量，$|v\rangle \in \mathbb{C}^d$ 意味着 $\||v\rangle\|_2 = 1$。对于非零向量 $u \in \mathbb{C}^d$，令 $|v\rangle \sim u$ 为 u的归一化，即 $|v\rangle = u/\|u\|_2$。

对于每个向量 $v \in \mathbb{C}^d$，令 $\mathsf{Diag}\, v \in \mathbb{C}^{d\times d}$ 为对角矩阵，其对角线元素表示 v。反之，对于每个方阵 M，令 $\mathsf{diag}\ M$为由 M的对角线元素组成的向量。对于矩阵（或一般的线性算子） M，我们分别使用 $\|M\|_{\mathrm{Tr}}$和 $\|M\|_2$表示其迹范数和谱范数，即

$$\big\|M\big\|_{\mathrm{Tr}} = \mathrm{Tr}\left[\sqrt{MM^\dagger}\right], \quad \big\|M\big\|_2 = \max_{v\neq 0} \|Mv\|_2/\|v\|_2.$$

对于厄米 $M \in \mathbb{C}^{d\times d}$，如果对每个 $v \in \mathbb{C}^d$,$v^\dagger M v \geq 0$，则称其为半正定算子。密度算子（或部分密度算子）是迹为 1（或分别至多为 1）的半正定算子。

将学习问题视为矩阵

设 $M\colon \mathcal{A}\times\mathcal{X} \to \{-1,1\}$ 是一个矩阵。矩阵 M 对应于以下学习问题。存在一个未知元素 $x \in \mathcal{X}$，它是从均匀随机分布中选取的。学习者试图通过样本（a,b）学习 x，其中 $a \in \mathcal{A}$是从均匀随机分布中选取的，并且 $b = M(a,x)$。也就是说，学习算法会获得一个样本流：（a_1,b_1）、（a_2,b_2）.....其中每个 a_t均服从均匀分布，并且对于每个 t，$b_t = M(a_t,x)$。对于每个 $a \in \mathcal{A}$，我们使用 $M_a\colon \mathcal{X}\to\{-1,1\}$来表示对应于 a行的向量 M。

提取器

矩阵 $M\colon \mathcal{A}\times\mathcal{X}\to\{-1,1\}$具有 $n = \log_2|\mathcal{X}|$，是一个误差为 2^{-r} 的 (k,ℓ)-L_2提取器；如果对于每个在 $\mathcal{X}$ 上满足 $\|P\|_2 \leq 2^\ell\cdot 2^{-n/2}$ 的分布 P，至多有 $2^{-k}\cdot|\mathcal{A}|$ 行 $a \in \mathcal{A}$ 满足

$$|\langle M_a, P\rangle| \geq 2^{-r}.$$

3.2 单位向量上的二次型反集中界

引理 3.1. 存在一个绝对常数 c，使得以下结论成立。设 σ 是希尔伯特空间 $\mathcal{V} = \mathbb{C}^d$ 上的厄米算子，且 v 是 $\mathcal{V}$ 中的均匀随机单位向量。那么对于每个 $\varepsilon > 0$，都有

$$\Pr\left[|v^\dagger\sigma v| \leq \frac{\varepsilon\|\sigma\|_2}{d}\right] \leq c\sqrt{\varepsilon} + e^{-d}.$$

证明。令 $g = (g_1,\ldots,g_d) \sim \mathcal{N}(0,1)^d$为标准高斯随机变量。注意到 $\|g\|_2^2$服从 χ^2_d-分布，且 $|g^\dagger\sigma g|/\|g\|_2^2$ 与 $|v^\dagger\sigma v|$同分布。因此，根据并集界，有

$$\Pr_v\left[|v^\dagger\sigma v| \leq \frac{\varepsilon\|\sigma\|_2}{d}\right] = \Pr_g\left[|g^\dagger\sigma g| \leq \varepsilon\|\sigma\|_2\cdot\frac{\|g\|_2^2}{d}\right] \leq \Pr_g\left[|g^\dagger\sigma g| \leq 5\varepsilon\|\sigma\|_2\right] + \Pr_g\left[\|g\|_2^2 \geq 5d\right].$$

对于第一项，注意到 $\mathrm{Var}[g^\dagger \sigma g] = 2\mathrm{Tr}[\sigma^2]$ （例如参见 [RS08，第 5]章），其不小于 $2\|\sigma\|_2^2$。因此，根据 Carbery–Wright 不等式 [CW01], ，存在一个绝对常数 c使得

$$\Pr\left[|g^\dagger \sigma g| \le 5\varepsilon\|\sigma\|_2\right] \le \Pr\left[|g^\dagger \sigma g| \le 4\varepsilon \mathrm{Var}[g^\dagger \sigma g]^{1/2}\right] \le c\sqrt{\varepsilon}.$$

对于第二项，关于 χ^2-分布的标准 Laurent-Massart 界 [LM00] 为：

$$\Pr\left[\|g\|_2^2 \ge 5d\right] \le e^{-d}. \qquad \square$$

3.3 多体量子系统

q 量子比特的量子态可表示在希尔伯特空间 $\mathcal{V} = (\mathbb{C}^2)^{\otimes q} = \mathbb{C}^{2^q}$中。在希尔伯特空间 $\mathcal{V}_{[m]} = \mathcal{V}_1 \otimes \cdots \otimes \mathcal{V}_m$的张量积中，多体部分系统 $V_1, \ldots, V_m$ 由部分密度算子 $\rho_{V_{[m]}}$表示。对于指标子集 $I \subseteq [m]$， $\{V_i\}_{i\in I}$ 上的子系统（简称 V_I）定义为对 $j \notin I$取偏迹，即

$$\rho_{V_I} = \mathrm{Tr}_{V_{j\notin I}}[\rho_{V_{[m]}}].$$

现在，对于任意两个不相交的子集 $I, J \subset [m]$，给定某个 $|v_J\rangle \in \mathcal{V}_J = \bigotimes_{j\in J} \mathcal{V}_j$， V_I 上的条件系统定义为

$$\rho_{V_I|v_J} = \left(\mathbb{I}_{V_I} \otimes \langle v_J|\right) \rho_{V_{I\cup J}} \left(\mathbb{I}_{V_I} \otimes |v_J\rangle\right),$$

这是一个定义在 V_I上的部分密度算子。注意，迹

$$\mathrm{Tr}\left[\rho_{V_I|v_J}\right] = \langle v_J|\rho_{V_J}|v_J\rangle$$

only 取决于系统 ρ 和 $|v_J\rangle$，但与 I的选择无关 .

另一个之后会反复用到的简单事实是，对于V_J的正交基 $\mathcal{B}$ ，我们有

$$\rho_{V_I} = \mathrm{Tr}_{V_J}[\rho_{V_{I\cup J}}] = \sum_{|v_J\rangle\in\mathcal{B}} \rho_{V_I|v_J}.$$

3.4 经典-量子系统

在多部系统的底层空间 $\mathcal{V}_1 \otimes \cdots \otimes \mathcal{V}_m$ 中，如果存在一个固定的正交基 $\mathcal{B}_i$ ，使得对于每个多部系统 $\rho_{V_{[m]}}$、每一对不同的$|v_i\rangle \ne |v_i'\rangle \in \mathcal{B}_i$ 以及任意两个量子态 $|v\rangle$、 $|v'\rangle \in \bigotimes_{j\ne i} \mathcal{V}_j$，都有

$$\langle v_i, v|\rho_{V_{[m]}}|v_i', v'\rangle = 0.$$

不失一般性，在本文余下部分中，我们始终假设 $\mathcal{B}_i$ 是计算基态的集合。我们也将 $\mathcal{V}_i$ 与离散集合 $\mathcal{B}_i$ 等同起来，并且在讨论 $\mathcal{V}_i$ 中的经典元素时省略狄拉克括号。在这种情况下，每个多部系统 $\rho_{V_{[m]}}$ 都可以写成直和

$$\rho_{V_{[m]}} = \bigoplus_{v_i\in\mathcal{V}_i} \rho_{V_{[m]\setminus\{i\}}|v_i}.$$

读者可能会发现，在某些后续情形中，这种直和视角更容易处理。

当 V_I 为经典时，在给定任意满足 J 与 I不相交的 $|v_J\rangle \in \mathcal{V}_J$ 的条件下，系统 $\rho_{V_I|v_J}$在 V_I上表示为一个对角矩阵。若 $\mathrm{Tr}[\rho_{V_I|v_J}] > 0$，则它会在 V_I的计算基态上诱导出一个分布，其定义为

$$P^{\rho}_{V_I|v_J} = \mathrm{diag}\,\rho_{V_I|v_J}/\mathrm{Tr}[\rho_{V_I|v_J}]. \tag{3}$$

在本文的其余部分，每当我们使用此记号 $P^{\rho}_{V_I|v_J}$时，都默认假设 $\mathbf{Tr}[\rho_{V_I|v_J}] > 0$，且该分布存在。

在本文中，我们通常考虑以下情形：存在一个量子存储寄存器V，其取值于希尔伯特空间 $\mathcal{V}$；以及一个经典存储寄存器 W，其取值于存储状态集合 $\mathcal{W}$。此外，还有一些与 V 和 W相关的经典信息 $X \in \mathcal{X}$（在本文后文中，它就是要学习的概念）。我们将使用以下事实：

命题 3.2. 设 ρ_{XVW}是定义在经典变量 X、W和量子变量 V上的经典-量子系统。对于每个 $w \in \mathcal{W}$，$P^{\rho}_{X|w}$都是某些 $\{|v\rangle\} \subseteq \mathcal{V}$对应的$P^{\rho}_{X|v,\ w}$的凸组合。

证明. 设 $\mathcal{B}$ 是 $\mathcal{V}$的一组正交基，因此我们有（由上一节末尾可知）

$$\rho_{X|w} = \sum_{|v\rangle \in \mathcal{B}} \rho_{X|v,w}.$$

因此，$P^{\rho}_{X|w}$ 是 $P^{\rho}_{X|v,\ w}$ 的线性组合，其中 $|v\rangle \in \mathcal{B}$的系数非负。由于它们都是分布，因此必为凸组合。 □

经典-量子混合系统上的算子刻画。现在，我们刻画经典-量子混合存储空间 $\mathcal{V} \otimes \mathcal{W}$ 上所有可能的算子。在假设 W 是经典的之前，我们将作用于该系统的量子信道视为作用于底层空间 $\mathcal{V} \otimes \mathbb{C}^{|\mathcal{W}|}$ 上的量子信道。现在，记 $\mathscr{T}_{\mathcal{V}\otimes\mathcal{W}}$ 为满足以下条件的所有此类量子信道 Φ 的集合：对于 $\mathcal{V} \otimes \mathcal{W}$ 中的每个经典-量子系统 ρ_{VW}，W 在 $\Phi(\rho_{VW})$ 中仍是经典的。也就是说，对于任意两个量子态 $|v\rangle$、$|v'\rangle \in \mathcal{V}$ 以及任意一对不同的 w、$w' \in \mathcal{W}$，我们有

$$\langle v, w|\Phi(\rho_{VW})|v', w'\rangle = 0.$$

请注意，$\mathscr{T}_{\mathcal{V}\otimes\mathcal{W}}$ 中并非所有量子信道都能在物理上实现。例如，对于只有一位经典内存且没有量子存储器的情况，该量子信道

$$\begin{pmatrix} a & c \\ \overline{c} & b \end{pmatrix} \mapsto \begin{pmatrix} a & ic \\ -i\overline{c} & b \end{pmatrix}$$

不是经典算子。然而，由于我们受限于经典-量子系统，该量子信道实际上等价于作用在一位经典内存上的恒等信道。一般而言，$\mathscr{T}_{\mathcal{V}\otimes\mathcal{W}}$ 中的每个量子信道都等价于一个由 $\mathcal{W}$ 控制、将 $\mathcal{V}$ 映射到 $\mathcal{V} \otimes \mathcal{W}$ 的量子信道。下面，我们将证明这一观察结果，并用它来说明以下论断：

命题 3.3。设 ρ_{XVW} 是定义在经典 X、W 和量子 V 上的经典-量子系统。设 $\Phi \in \mathscr{T}_{\mathcal{V}\otimes\mathcal{W}}$，并用 $\Phi(\rho)$ 表示对 VW 应用 Φ，并对 X 应用恒等变换后得到的系统。那么对于每个 $|v\rangle \in \mathcal{V}$ 和 $w \in \mathcal{W}$，$P^{\Phi(\rho)}_{X|v}{}_{,\ w}$ 都是某些 $\{|v'\rangle\} \subseteq \mathcal{V}$ 和 $\{w'\} \subseteq \mathcal{W}$ 的 $P^{\rho}_{X|v',\ w'}$ 的凸组合。

命题 3.2 与命题 3.3 的一个区别是，在命题 3.3 中，不总是能将 $P^{\Phi(\rho)}_{X|v}{}_{,\ w}$表示为以下各项的凸组合：$P^{\rho}_{X|v',\ w'}$，其中 $|v'\rangle$ 取自 $\mathcal{V}$的一组正交基。但在命题 3.2中总是可以做到这一点。尽管这一区别对本文并不重要，我们仍在此提及，以便说明清楚。

证明。由于 $\Phi \in \mathscr{T}_{\mathcal{V}\otimes\mathcal{W}}$，对于经典-量子系统，以下量子信道在功能上等价于 Φ：

$$\Phi' : \rho \to \sum_{w\in\mathcal{W}} \Phi(\rho_{V|w} \otimes |w\rangle\langle w|).$$

Φ' 的物理含义是在计算基下测量 W （这不应改变我们所关心的功能），并应用 Φ。

通过定义量子信道 $\Phi_w(\cdot) := \Phi(\cdot \otimes |w\rangle\langle w|)$，上述内容也可以写为：

$$\Phi' : \rho \to \sum_{w \in \mathcal{W}} \Phi_w(\rho_{V|w}).$$

现在考虑每个 Φ_w 的 Kraus表示，即一组有限的线性算子$E_{w,\ k} \colon \mathcal{V} \to \mathcal{V} \otimes \mathcal{W}$，使得

$$\Phi_w(\rho_{V|w}) = \sum_k E_{w,k} \rho_{V|w} E_{w,k}^\dagger, \qquad \sum_k E_{w,k}^\dagger E_{w,k} = \mathbb{I}_V.$$

我们可以写成

$$\begin{aligned}
\Phi(\rho)_{X|v,w} = \Phi'(\rho)_{X|v,w} &= (\mathbb{I}_X \otimes \langle v, w|)\Phi'(\rho)(\mathbb{I}_X \otimes |v, w\rangle) \\
&= \sum_{w' \in \mathcal{W}} \sum_k (\mathbb{I}_X \otimes \langle v, w|E_{w',k})\rho_{XV|w'}(\mathbb{I}_X \otimes E_{w',k}^\dagger|v, w\rangle) \\
&= \sum_{w' \in \mathcal{W}} \sum_k \left\| E_{w',k}^\dagger |v, w\rangle \right\|_2 \cdot \rho_{X|v',w'}
\end{aligned}$$

其中，在求和的每一项中，$|v'\rangle \sim E_{w',\ k}^\dagger|v,\ w\rangle$。与命题 3.2，$P^{\Phi(\rho)\, X|v,\ w}$是 $P^{\rho}_{X|v',\ w'}$的凸组合。

□

3.5 混合内存分支程序

对于对应于矩阵 M 的学习问题，具有 m 位经典内存、 q 量子比特量子存储器且长度为 T 的混合内存分支程序定义如下。

在每个阶段 $0 \le t \le T$，分支程序的内存状态都表示为一个经典-量子系统 $\rho_{VW}^{(t)}$，其量子存储空间为 $\mathcal{V} = (\mathbb{C}^2)^{\otimes q}$，经典内存空间为 $\mathcal{W} = \{0,1\}^m$。内存状态会根据分支程序接收到的样本而演化，因此取决于未知元素 $x \in_R \mathcal{X}$。于是，我们可以将整体系统解释为作用于 XVW 上的系统，其中 X 由一个未知概念 x组成，从而得到一个经典-量子系统 $\rho_{XVW}^{(t)}$。 x 的分布始终是均匀的，即

$$\rho_X^{(t)} = \mathrm{Tr}_{VW}[\rho_{XVW}^{(t)}] = \frac{1}{2^n}\mathbb{I}_X.$$

最初，内存 VW 独立于 X，且可以任意初始化。我们假设

$$\rho_{XVW}^{(0)} = \frac{1}{2^n}\mathbb{I}_X \otimes \frac{1}{2^q}\mathbb{I}_V \otimes \frac{1}{2^m}\mathbb{I}_W.$$

在每个阶段 $0 \le t < T$，分支程序都会接收一个样本 (a, b)，其中 $a \in_R \mathcal{A}$ 且$b = M(a, x)$，并对其内存状态施加一个操作 $\Phi_{t,\ a,\ b} \in \mathscr{T}_{\mathcal{V}\otimes\mathcal{W}}$。因此，整个系统的演化可以写为

$$\rho_{XVW}^{(t+1)} = \mathop{\mathbf{E}}_{a \in_R \mathcal{A}} \left[\sum_{x \in \mathcal{X}} |x\rangle\langle x| \otimes \Phi_{t,a,M(a,x)}\left(\rho_{VW|x}^{(t)}\right) \right].$$

最后，在阶段 $t = T$，对 $\rho_{VW}^{(T)}$ 进行计算基测量，分支程序根据测量结果 $(v, w) \in \{0,1\}^{q+m}$ 输出一个元素 $\widetilde{x} \in X$。程序的成功概率是 $\widetilde{x} = x$ 的概率，可表述为

$$\sum_{\substack{x \in \mathcal{X}, v \in \{0,1\}^q, w \in \mathcal{W} \\ \widetilde{x}(v,w) = x}} \langle x, v, w|\rho_{XVW}^{(T)}|x, v, w\rangle.$$

4 主要结果

定理2. 设 $\mathcal{X}, \mathcal{A}$ 是两个有限集合，满足 $n = \log_2 |\mathcal{X}|$。设 $M\colon \mathcal{A} \times \mathcal{X} \to \{-1, 1\}$ 是一个矩阵，它是一个 (k', ℓ')-L_2 提取器，误差为 $2^{-r'}$，其中 k', ℓ' 和 r' 足够大，且 $\ell' \leq n$。令

$$r = \min\left\{\frac{1}{4}r', \frac{1}{26}\ell' + \frac{1}{6}, \frac{1}{2}(k'-1)\right\}.$$

设 ρ 是一个分支程序，用于求解与 M 对应的学习问题，该程序由经典-量子系统 $\rho^{(t)}_{XVW}$ 描述，具有 q 量子比特的量子存储器 V、m 比特的经典内存 W，长度为 T。若 $m \leq \frac{1}{44}(k'-1)\ell'$、$q \leq r - 7$ 且 $T \leq 2^{r-2}$，则 ρ 的成功概率至多为 $O(2^{q-r})$。

从现在起，我们令 $k = k' - 1$ 和 $\ell = \frac{1}{5}(\ell' - 13r - 2)$。于是得到以下不等式，后文将会用到：

$$\begin{aligned} q + r + 1 - r' &\leq -2r. && (4)\\ 2\ell + 9r - n &\leq -r. && (5)\\ (k-r)\ell &\geq 2m + 4r + 1. && (6) \end{aligned}$$

上述不等式的详细计算过程见附录A。

4.1 截断经典-量子系统

这里我们描述如何根据 $\rho_{X|v,\ w}$ 上所需的某个性质 $G(v, w)$，对部分经典-量子系统 ρ_{XVW} 进行截断。目标是移除 ρ_{XVW} 中不满足 G 的部分。我们执行以下步骤：

1. 将部分系统 ρ'_{XVW}初始化为 ρ_{XVW}，并将每个 $w \in \mathcal{W}$对应的子空间 $\mathcal{V}_w \subseteq \mathcal{V}$初始化为 $\mathcal{V}$。

2. 选择 $w \in \mathcal{W}$ 和 $|v\rangle \in \mathcal{V}_w$，使得 $\mathbf{Tr}[\rho'_{X|v,w}] > 0$ 和 $G(v, w)$ 不成立。

3. 通过投影将部分系统 ρ'_{XVW} 变为以下系统 ：

$$\left(\mathbb{I}_X \otimes (\mathbb{I}_{VW} - |v, w\rangle\langle v, w|)\right) \rho'_{XVW} \left(\mathbb{I}_X \otimes (\mathbb{I}_{VW} - |v, w\rangle\langle v, w|)\right),$$

并将 $\mathcal{V}_w$ 变为与 $|v\rangle$正交的子空间，即

$$\{|v'\rangle \in \mathcal{V}_w \mid \langle v|v'\rangle = 0\}.$$

4. 重复 从步骤 2 开始，直到不存在这样的 w 和 $|v\rangle$。将最终系统记为 $\rho^{|G}_{XVW}$.

在步骤 2 中，我们任意选取 w 和 $|v\rangle$，只要满足要求即可；不过，我们也可以始终将其视为遍历 $w \in \mathcal{W}$，并分别处理每个 $\rho_{XV|w}$。对每个 w所作的$|v\rangle$ 选择确实会影响最终系统 $\rho^{|G}_{XVW}$；但正如我们稍后将看到的，这些选择与我们的证明无关。

下面，我们给出关于截断系统的两个有用引理。

引理 4.1。对于每个满足 $\mathbf{Tr}[\rho^{|G}_{X|v,w}] > 0$ 的 $|v\rangle \in \mathcal{V}$ 和 $w \in \mathcal{W}$，剩余子空间 $\mathcal{V}_w$ 中都存在 $|v'\rangle$，使得

$$P^{\rho^{|G}}_{X|v,w} = P^{\rho}_{X|v',w} = P^{\rho^{|G}}_{X|v',w}.$$

证明。只需证明执行一轮截断过程时该引理成立。假设在步骤 2 中选取了 $|v_1\rangle$、$w_1\rangle$，从而得到部分系统

$$\rho'_{XVW} = \big(\mathbb{I}_X \otimes (\mathbb{I}_{VW} - |v_1, w_1\rangle\langle v_1, w_1|)\big)\rho_{XVW}\big(\mathbb{I}_X \otimes (\mathbb{I}_{VW} - |v_1, w_1\rangle\langle v_1, w_1|)\big).$$

我们可以写成

$$\begin{aligned}\rho'_{X|v,w} &= \big(\mathbb{I}_X \otimes \langle v, w|\big)\rho'_{XVW}\big(\mathbb{I}_X \otimes |v, w\rangle\big) \\ &= \big(\mathbb{I}_X \otimes (\langle v, w| - \langle v, w|v_1, w_1\rangle\langle v_1, w_1|)\big)\rho_{XVW}\big(\mathbb{I}_X \otimes (|v, w\rangle - |v_1, w_1\rangle\langle v_1, w_1|v, w\rangle)\big).\end{aligned}$$

- 如果 $w \neq w_1$，那么

  $$\rho'_{X|v,w} = \big(\mathbb{I}_X \otimes \langle v, w|\big)\rho_{XVW}\big(\mathbb{I}_X \otimes |v, w\rangle\big) = \rho_{X|v,w}.$$

  选择 $|v'\rangle = |v\rangle$ 即可直接得到该引理。

- 如果 $w = w_1$，那么对于 $\langle v_1$、$w_1|v$、$w\rangle = \langle v_1|v\rangle = \lambda$，我们有

  $$\rho'_{X|v,w} = \big(\mathbb{I}_X \otimes (\langle v| - \overline{\lambda}\langle v_1|)\langle w|\big)\rho_{XVW}\big(\mathbb{I}_X \otimes (|v\rangle - \lambda|v_1\rangle)|w\rangle\big).$$

  根据 $\mathbf{Tr}[\rho^{|G}_{X|v,w}] > 0$ 这一事实，我们必有 $|v\rangle \neq |v_1\rangle$。因此，若令 $|v'\rangle \sim |v\rangle - \lambda|v_1\rangle$ 为 $|v\rangle$ 在 $|v_1\rangle$ 的正交子空间上的归一化投影，则上述等式意味着 $P^{\rho'}_{X|v',\ w} = P^{\rho}_{X|v',\ w}$。同时，由于 $\langle v_1|v'\rangle = 0$，我们有 $\rho'_{X|v',\ w} = \rho_{X|v',\ w}$，证毕。

□

上述引理的一个直接推论是：如果 $G(v, w)$ 仅取决于分布 $P^{\rho}_{X|v,\ w}$，那么对于截断系统 $\rho^{|G}_{XVW}$ 中的任意 $|v\rangle \in \mathcal{V}$ 和 $w \in \mathcal{W}$，$G(v, w)$ 均成立，即使 $|v\rangle$ 不属于剩余子空间 $\mathcal{V}_w$。

我们的第二个引理基于以下事实：该事实给出了部分系统与其投影之间的迹距离上界，其证明见附录B。

命题 4.2. 对于任意部分系统 ρ和作用于 ρ上的投影算子 Π，有

$$\big\|\rho - \Pi\rho\Pi\big\|^2_{\mathrm{Tr}} \leq 4\mathrm{Tr}[\rho]^2 - 4\mathrm{Tr}[\Pi\rho]^2.$$

引理 4.3. F　　对于每个 $w \in W$，令 $|v_1\rangle$、…、$|v_d\rangle$ 为在 $\mathcal{V}_w$中第 2 步选取的量子态。则

$$\big\|\rho_{XV|w} - \rho^{|G}_{XV|w}\big\|_{\mathrm{Tr}} \leq 3\sum_{i=1}^{d}\sqrt{\mathrm{Tr}[\rho_{X|v_i,w}]\mathrm{Tr}[\rho_{XV|w}]}.$$

证明。由命题 4.2，取 ρ 为 $\rho_{XV|w}$，并取 Π 为

$$\mathbb{I}_X \otimes \prod_{i=1}^{d}(\mathbb{I}_V - |v_i\rangle\langle v_i|) = \mathbb{I}_X \otimes \left(\mathbb{I}_V - \sum_{i=1}^{d}|v_i\rangle\langle v_i|\right).$$

Then $\Pi\rho\Pi = \rho^{|G}_{XV|w}$ 且 $\mathbf{Tr}[\Pi\rho] = \mathbf{Tr}[\rho_{XV|w}] - \sum_{i=1}^{d}\mathbf{Tr}[\rho_{X|v_i,w}]$。因此我们有　　e

$$\begin{aligned}\big\|\rho_{XV|w} - \rho^{|G}_{XV|w}\big\|_{\mathrm{Tr}} &\leq \sqrt{4\mathrm{Tr}[\rho]^2 - 4\mathrm{Tr}[\Pi\rho]^2} \\ &\leq \sqrt{8(\mathrm{Tr}[\rho] - \mathrm{Tr}[\Pi\rho])\mathrm{Tr}[\rho]} \\ &= \sqrt{8\sum_{i=1}^{d}\mathrm{Tr}[\rho_{X|v_i,w}]\mathrm{Tr}[\rho_{XV|w}]} \\ &\leq 3\sum_{i=1}^{d}\sqrt{\mathrm{Tr}[\rho_{X|v_i,w}]\mathrm{Tr}[\rho_{XV|w}]}.\end{aligned}$$

□

由于 $\mathrm{Tr}[\rho_{XV|w}] \le 1$ 始终成立，对所有 $w \in \mathcal{W}$ 求和可得以下推论：推论 4.4。设 $|v_1, w_1\rangle, \ldots, |v_d, w_d\rangle$ 为步骤 2 中选取的所有存储状态。那么

$$\left\|\rho_{XVW} - \rho^{|G}_{XVW}\right\|_{\mathrm{Tr}} \le 3 \sum_{i=1}^{d} \sqrt{\mathrm{Tr}[\rho_{X|v_i,w_i}]}.$$

4.2 截断分支程序

The 我们对部分系统 ρ_{XVW} 所期望的性质由三部分组成：

- 小 L_2 范数：令 $G_2(v,w)$ 表示如下性质：

$$\left\|P^{\rho}_{X|v,w}\right\|_2 \le 2^{\ell} \cdot 2^{-n/2}.$$

- 小 L_∞ 范数：令 $G_\infty(v,w)$ 表示如下性质：

$$\left\|P^{\rho}_{X|v,w}\right\|_\infty \le 2^{2\ell+9r} \cdot 2^{-n}.$$

- 均匀划分：对于每个 $a \in \mathcal{A}$，令 $G_a(v,w)$ 表示如下性质：

$$|\langle M_a, P^{\rho}_{X|v,w}\rangle| \le 2^{-r}.$$

现在，我们通过指定每个阶段 t 的截断部分经典-量子系统 $\tau^{(t)}_{XVW}$来定义截断分支程序。初始时，令 $\tau^{(0)}_{XVW} = \rho^{(0)}_{XVW}$。对于每个阶段 $0 \le t \le T$，截断由三个要素组成（为方便起见，以下忽略 P 上的上标）：

1. 移除 $\left\|P_{X|v,\ w}\right\|_2$ 较大的部分。也就是说，令 $\tau^{(t,\star)}_{XVW} = \tau^{(t)|G_2}_{XVW}$。

2. 移除 $\left\|P_{X|v,\ w}\right\|_\infty$ 较大的部分。这分两步完成。

   - 首先，令 $g \in \{0,1\}^{\mathcal{X}\otimes\mathcal{W}}$ 为一个指示向量，使得当且仅当 $g(x,w) = 1$

   $$\mathrm{Tr}[\tau^{(t,\star)}_{X|w}] > 0 \text{ and } P^{\tau^{(t,\star)}}_{X|w}(x) \le 2^{2\ell+5r} \cdot 2^{-n}.$$

   令 $\tau^{(t,\circ)}_{XVW} = (gg^\dagger \otimes \mathbb{I}_V)\tau^{(t,\star)}_{XVW}(gg^\dagger \otimes \mathbb{I}_V)$，其中 $gg^\dagger$ 是作用于 $\mathcal{X}\otimes\mathcal{W}$ 上的投影算子。

   - 为确保投影 $gg^\dagger$ 后分布没有发生太大变化，对每个 $0 \le t < T$，令 $G_t(v,w)$ 表示如下性质

   $$\mathrm{Tr}[\tau^{(t,\circ)}_{X|v,w}] \ge (1-2^{-r})\mathrm{Tr}[\tau^{(t,\star)}_{X|v,w}].$$

   令 $\tau^{(t,\infty)}_{XVW} = \tau^{(t,\circ)|G_\infty \wedge G_t}_{XVW}$。

3. 对每个 $a \in \mathcal{A}$，移除（仅针对这个 a）满足 $P_{X|v,w}$ 不能被 a 整除的部分。也就是说，对每个 $a \in \mathcal{A}$，令 $\tau^{(t,a)}_{XVW} = \tau^{(t,\infty)|G_a}_{XVW}$。

然后，如果 $t < T$，对于每个 $a \in_R \mathcal{A}$，我们通过应用样本操作 $\Phi_{t,\ a,\ b}$，使系统像原始分支程序在 $\tau^{(t,a)}_{XVW}$上一样演化，从而得到

$$\tau^{(t+1)}_{XVW} = \mathop{\mathbf{E}}_{a\in_R\mathcal{A}}\left[\sum_{x\in\mathcal{X}} |x\rangle\langle x| \otimes \Phi_{t,a,M(a,x)}\left(\tau^{(t,a)}_{VW|x}\right)\right].$$

4.3 截断差异的界定

为了证明原始分支程序 $\rho^{(t)}$ 的成功概率较低，我们计划证明截断分支程序 $\tau^{(t)}$的成功概率上界，并界定这两个概率之间的差异。

这里，我们通过系统 $\rho^{(t)}_{XVW}$ 和 $\tau^{(t)}_{XVW}$ 之间的迹距离来界定差异。我们将证明，截断过程的每个组成部分对迹距离的贡献都很小，而且演化会保持迹距离。

4.3.1 按 G_2 截断

引理 4.5. 对于任意满足 $G_2(v,w)$ 被违反（即$\left\|P^{\tau^{(t)}}_{X|v,\ w}\right\|_2 > 2^{\ell}\cdot 2^{-n/2}$）的 $0\le t\le T$、$|v\rangle\in\mathcal{V}$ 和 $w\in\mathcal{W}$，必有 $\mathrm{Tr}[\tau^{(t)}_{X|v,w}] < 2^{-2m}\cdot 2^{-4r}$。

该引理指出，对于截断过程选取的任意方向 $|v、w\rangle$，其权重都很小，因此截断不会显著改变量子态。

证明。这是我们的主要技术引理，证明推迟至第 5 节。 □

由于在截断过程中至多选取了 2^{q+m} 个这样的方向，我们得到如下推论。

推论 4.6。对于每个 $0\le t\le T$，我们有 $\left\|\tau^{(t,\star)}_{XVW}-\tau^{(t)}_{XVW}\right\|_{\mathrm{Tr}}\le 3\cdot 2^{q-\quad 2r}$.

证明。回顾 $\tau^{(t,\star)}_{XVW}=\tau^{(t)|G_2}_{XVW}$。由于 $\dim(\mathcal{V}\otimes\mathcal{W})=2^{q+m}$，截断至多持续$2^{q+m}$轮。由于在每一轮中，所选取的 $|v,\ w\rangle$ 具有 $\mathrm{Tr}[\tau^{(t)}_{X|v,w}] < 2^{-2m}\cdot 2^{-4r}$，根据推论 4.4，我们有

$$\left\|\tau^{(t,\star)}_{XVW}-\tau^{(t)}_{XVW}\right\|_{\mathrm{Tr}}\le 3\cdot 2^{q+m}\cdot\sqrt{2^{-2m}\cdot 2^{-4r}}=3\cdot 2^{q-2r}.$$ □

4.3.2 通过 G_∞进行截断

引理 4.7。对于每个 $0\le t\le T$ 和 $w\in\mathcal{W}$，我们都有

$$\sum_{\substack{x\in\mathcal{X}\\ g(x,w)=0}} P^{\tau^{(t,\star)}}_{X|w}(x)\le 2^{-5r}.$$

证明。根据断言 3.2，$P^{\tau^{(t,\star)}}_{X|w}$ 是 $P^{\tau^{(t,\star)}}_{X|v\,w,}$ 的凸组合。由引理 4.1 可知，对于每个 $|v\rangle$ 和 w，$G_2($ $P^{\tau^{(t,\star)}}_{X|v\,w,}$ $)$ 成立，因此根据 ℓ_2-范数的凸性，我们也知道 $G_2(P^{\tau^{(t,\star)}}_{X|w})$ 成立。这意味着

$$\mathop{\mathbf{E}}_{x\sim P^{\tau^{(t,\star)}}_{X|w}}\left[P^{\tau^{(t,\star)}}_{X|w}(x)\right]=\left\|P^{\tau^{(t,\star)}}_{X|w}\right\|_2^2\le 2^{2\ell}\cdot 2^{-n}.$$

因此，根据马尔可夫不等式，我们有

$$\sum_{\substack{x\in\mathcal{X}\\ g(x,w)=0}} P^{\tau^{(t,\star)}}_{X|w}(x)=\Pr_{x\sim P^{\tau^{(t,\star)}}_{X|w}}\left[P^{\tau^{(t,\star)}}_{X|w}(x)>2^{2\ell+5r}\cdot 2^{-n}\right]\le 2^{-5r}.$$ □

Coro推论 4.8。对于每个 $0\le t\le T$ 以及每个 $w\in\mathcal{W}$，我们有 $\tau^{(t,\circ)}_{XV|w}\le\tau^{(t,\star)}_{XV|w}$，一个 d

$$\mathrm{Tr}[\tau^{(t,\circ)}_{XV|w}]\ge(1-2^{-5r})\cdot\mathrm{Tr}[\tau^{(t,\star)}_{XV|w}].$$

此外，我们有 $\left\|\tau^{(t,\circ)}_{XVW}-\tau^{(t,\star)}_{XVW}\right\|_{\mathrm{Tr}}\le 2^{-5r}$。

证明。由于 X和 W均为经典的，且 $\tau_{XVW}^{(t,\circ)}=(gg^\dagger\otimes\mathbb{I}_V)\tau_{XVW}^{(t,\star)}(gg^\dagger\otimes\mathbb{I}_V)$，因此我们有

$$\tau_{XV|w}^{(t,\star)}-\tau_{XV|w}^{(t,\circ)}=\sum_{\substack{x\in\mathcal{X}\\ g(x,w)=0}}|x\rangle\langle x|\otimes\tau_{V|x,w}^{(t,\star)},$$

它是半正定的。回顾一下（公式（3））

$$\mathrm{Tr}[\tau_{V|x,w}^{(t,\star)}]=\langle x,w|\tau_{XW}^{(t,\star)}|x,w\rangle=\mathrm{diag}\,\tau_{X|w}^{(t,\star)}(x)=P_{X|w}^{\tau^{(t,\star)}}(x)\mathrm{Tr}[\tau_{X|w}^{(t,\star)}],$$

we have

$$\mathrm{Tr}[\tau_{XV|w}^{(t,\star)}]-\mathrm{Tr}[\tau_{XV|w}^{(t,\circ)}]=\sum_{\substack{x\in\mathcal{X}\\ g(x,w)=0}}\mathrm{Tr}[\tau_{V|x,w}^{(t,\star)}]=\sum_{\substack{x\in\mathcal{X}\\ g(x,w)=0}}P_{X|w}^{\tau^{(t,\star)}}(x)\cdot\mathrm{Tr}[\tau_{X|w}^{(t,\star)}]\le 2^{-5r}\cdot\mathrm{Tr}[\tau_{X|w}^{(t,\star)}].$$

因此，由于 $\tau_{XVW}^{(t,\circ)}-\tau_{XVW}^{(t,\star)}$ 是半正定的，我们有

$$\left\|\tau_{XVW}^{(t,\circ)}-\tau_{XVW}^{(t,\star)}\right\|_{\mathrm{Tr}}=\sum_{w\in\mathcal{W}}\mathrm{Tr}[\tau_{XV|w}^{(t,\star)}]-\mathrm{Tr}[\tau_{XV|w}^{(t,\circ)}]\le 2^{-5r}\sum_{w\in\mathcal{W}}\mathrm{Tr}[\tau_{X|w}^{(t,\star)}]\le 2^{-5r}.$$ □

引理 4.9. 对于任意满足以下条件的 $0\le t\le T$、$|v\rangle\in\mathcal{V}$ 和 $w\in\mathcal{W}$：$G_\infty(v,w)$ 不成立（即 $\left\|P_{X|v\ \ ,w}^{\tau^{(t,\circ)}}\right\|_\infty>2^{2\ell+9r}\cdot 2^{-n}$），或者 $G_t(v,w)$ 不成立（即 $\mathbf{Tr}[\tau_{X|v,w}^{(t,\circ)}]<(1-2^{-r})\mathbf{Tr}[\tau_{X|v,w}^{(t,\star)}]$），必有 $\mathbf{Tr}[\tau_{X|v,w}^{(t,\circ)}]<2\cdot 2^{-4r}\cdot\mathbf{Tr}[\tau_{X|w}^{(t,\circ)}]$。

证明。若 $G_\infty(v,w)$ 不成立，令 $x\in\mathcal{X}$ 为满足 $P_{X|v\ w}^{\tau^{(t,\circ)}}(x)>2^{2\ell+9r}\cdot 2^{-n}$ 的那个。如果 $g(x,w)=0$，那么 $P_{X|w}^{\tau^{(t,\circ)}}(x)=0$；而如果 $g(x,w)=1$，则根据推论 4.8，

$$P_{X|w}^{\tau^{(t,\circ)}}(x)\le\frac{\mathrm{Tr}[\tau_{X|w}^{(t,\star)}]}{\mathrm{Tr}[\tau_{X|w}^{(t,\circ)}]}\cdot 2^{2\ell+5r}\cdot 2^{-n}\le(1-2^{-5r})^{-1}\cdot 2^{2\ell+5r}\cdot 2^{-n}.$$

因此我们总有

$$\mathrm{Tr}[\tau_{X|v,w}^{(t,\circ)}]\le\frac{P_{X|w}^{\tau^{(t,\circ)}}(x)}{P_{X|v,w}^{\tau^{(t,\circ)}}(x)}\cdot\mathrm{Tr}[\tau_{X|w}^{(t,\circ)}]\le 2\cdot 2^{-4r}\cdot\mathrm{Tr}[\tau_{X|w}^{(t,\circ)}],$$

wh这里第一个不等式来自以下事实：$\tau_{X|w}^{(t,\circ)}\ge\tau_{X|v}^{(t,\circ)}$，$w$ 以及方程（3 ）.

若 G_t（v，w）不成立，因为我们从推论4.8 知道

$$\begin{aligned}\left|\mathrm{Tr}[\tau_{X|v,w}^{(t,\circ)}]-\mathrm{Tr}[\tau_{X|v,w}^{(t,\star)}]\right|&\le\left\|\tau_{XV|w}^{(t,\circ)}-\tau_{XV|w}^{(t,\star)}\right\|_{\mathrm{Tr}}\le 2^{-5r}\cdot\mathrm{Tr}[\tau_{XV|w}^{(t,\star)}]\\&\le 2^{-5r}\cdot(1-2^{-5r})^{-1}\cdot\mathrm{Tr}[\tau_{XV|w}^{(t,\circ)}],\end{aligned}$$

因此，由 $\mathbf{Tr}[\tau_{(t,\circ)X|v,w}]<(1-2^{-r})\mathbf{Tr}[\tau_{X|v,w}^{(t,\star)}]$，我们推得

$$\begin{aligned}\mathrm{Tr}[\tau_{X|v,w}^{(t,\circ)}]&<(2^r-1)\cdot\left(\mathrm{Tr}[\tau_{X|v,w}^{(t,\star)}]-\mathrm{Tr}[\tau_{X|v,w}^{(t,\circ)}]\right)\\&\le(2^r-1)\cdot 2^{-5r}\cdot(1-2^{-5r})^{-1}\cdot\mathrm{Tr}[\tau_{XV|w}^{(t,\circ)}]\\&<2\cdot 2^{-4r}\cdot\mathrm{Tr}[\tau_{X|w}^{(t,\circ)}].\end{aligned}$$ □

推论 4.10。对于每个 $0\le t\le T$，都有 $\left\|\tau_{XVW}^{(t,\infty)}-\tau_{XVW}^{(t,\circ)}\right\|_{\mathrm{Tr}}\le 5\cdot 2^{q-2r}$。

证明。回顾 $\tau_{XVW}^{(t,\infty)}=\tau_{XVW}^{(t,\circ)|G_\infty\wedge G_t}$。对于每个 $w\in\mathcal{W}$，截断至多选取维数为 $\mathcal{V}=2^q$的量子态 $|v,\ w\rangle$，每个量子态满足 $\mathbf{Tr}[\tau_{(t,\circ)X|v,\ w}]<2\cdot 2^{-4r}\cdot\mathrm{Tr}[\tau_{X|w}^{(t,\circ)}]$。因此，对每个 $w\in\mathcal{W}$应用引理4.3，可得

$$\left\|\tau_{XVW}^{(t,\infty)}-\tau_{XVW}^{(t,\circ)}\right\|_{\mathrm{Tr}}\le 3\cdot\sum_{w\in\mathcal{W}}2^q\cdot\sqrt{2\cdot 2^{-4r}}\cdot\mathrm{Tr}[\tau_{X|w}^{(t,\circ)}]\le 5\cdot 2^{q-2r}.$$ □

4.3.3 按 G_a 截断

注意，在从 $\tau^{(t,\star)}$ 到 $\tau^{(t,\circ)}$ 的截断步骤中，分布 $P^{\tau^{(t,\star)}}_{X|v}$，$w$ 可能会发生变化，不再满足 G_2。然而，通过使用 G_t 进行截断，任何变化过大的此类分布都会被排除，因此我们得到以下保证。

引理 4.11. 对于任意 $0 \le t \le T$、$|v\rangle \in \mathcal{V}$ 和 $w \in \mathcal{W}$，我们有

$$\left\|P^{\tau^{(t,\infty)}}_{X|v,w}\right\|_2 \le (1-2^{-r})^{-1} \cdot 2^{\ell} \cdot 2^{-n/2}.$$

证明。由引理 4.1，存在 $|v'\rangle \in \mathcal{V}$ 使得 $P^{\tau^{(t,\infty)}}_{X|v},w = P^{\tau^{(t,\infty)}}_{X|v'},w = P^{\tau^{(t,\circ)}}_{X|v'},w$。按 G_t 截断可确保 $\mathbf{Tr}[\tau^{(t,\circ)}_{X|v',w}] \ge (1-2^{-r})\mathbf{Tr}[\tau^{(t,\star)}_{X|v',w}]$，因此

$$\left\|P^{\tau^{(t,\infty)}}_{X|v,w}\right\|_2 = \left\|P^{\tau^{(t,\circ)}}_{X|v',w}\right\|_2 = \frac{\left\|\operatorname{diag}\tau^{(t,\circ)}_{X|v',w}\right\|_2}{\mathrm{Tr}[\tau^{(t,\circ)}_{X|v',w}]} \le \frac{\left\|\operatorname{diag}\tau^{(t,\star)}_{X|v',w}\right\|_2}{(1-2^{-r})\mathrm{Tr}[\tau^{(t,\star)}_{X|v',w}]} \le (1-2^{-r})^{-1} \cdot 2^{\ell} \cdot 2^{-n/2}. \quad \square$$

引理 4.12. 对于定义在 $\mathcal{X} \otimes \mathcal{V}$ 上的任意部分经典-量子系统 τ_{XV}，若对每个 $|v\rangle \in \mathcal{V}$ 都有 $\left\|P^{\tau}_{X|v}\right\|_2 \le 2^{\ell'} \cdot 2^{-n/2}$ 成立，则我们有

$$\Pr_{a \in_R \mathcal{A}}\left[\exists |v\rangle \in \mathcal{V}, |\langle M_a, P^{\tau}_{X|v}\rangle| \ge 2^{-r}\right] \le 2^{-2r}.$$

证明。注意，我们可以将 $\tau_V = \mathbf{Tr}_X[\tau_{XV}]$ 视为 $\mathbb{I}_V$。这是因为我们可以首先假设 τ_V 是满秩的（否则将 $\mathcal{V}$ 改为其子空间，本引理的结论仍然成立）；如果对于某个非奇异的 Q，我们有对角化 $Q^{\dagger}\tau_V Q = \mathbb{I}_V$，那么考虑新的系统

$$\tau'_{XV} = (\mathbb{I}_X \otimes Q^{\dagger})\tau_{XV}(\mathbb{I}_X \otimes Q),$$

而在 $|v\rangle \in \mathcal{V}$ 上的分布集合 $\{P^{\tau}_{X|v}\}$ 和 $\{P^{\tau'}_{X|v}\}$ 是相同的，因为对于 $|v'\rangle \sim Q|v\rangle$ 有 $P^{\tau'}_{X|v} = P^{\tau}_{X|v'}$。由 $\tau_V = \mathbb{I}_V$ 可知，对于每个 $|v\rangle \in \mathcal{V}$ 都有 $\mathbf{Tr}[\tau_{X|v}] = 1$，因此 $P^{\tau}_{X|v} = \mathbf{diag}\,\tau_{X|v}$。

令 $\mathcal{A}' \subseteq \mathcal{A}$ 为满足存在 $|v\rangle \in \mathcal{V}$ 且有 $|\langle M_a, P^{\tau}_{X|v}\rangle| \ge 2^{-r}$ 的 $a \in \mathcal{A}$ 的集合。对于每个 $a \in \mathcal{A}'$，令

$$\sigma_a = \mathrm{Tr}_X[(\mathrm{Diag}\, M_a \otimes \mathbb{I}_V)\tau_{XV}]$$

它是作用于 $\mathcal{V}$ 上的厄米算子。存在 $|v\rangle \in \mathcal{V}$，使得

$$|\langle v|\sigma_a|v\rangle| = |\langle M_a, \operatorname{diag}\tau_{X|v}\rangle| = |\langle M_a, P^{\tau}_{X|v}\rangle| \ge 2^{-r},$$

这意味着 $\|\sigma_a\|_2 \ge 2^{-r}$。现在令 $|u\rangle$ 为 $\mathcal{V}$ 中的均匀随机单位向量，根据引理3.1 可知，存在某个绝对常数 c，

$$\Pr_{|u\rangle}\left[|\langle u|\sigma_a|u\rangle| \ge 2^{-r'}\right] \ge 1 - 2^{(q+r-r')/2}c - e^{-2^q} \ge 1 - 2^{-r}c - e^{-1} \ge 1/2.$$

倒数第二个不等式由等式（4）得出，而最后一个不等式是因为假设 r 足够大。

由于上述结论对每个 $a \in \mathcal{A}'$ 都成立，因此 $\mathbf{Pr}_{a\in\mathcal{A}',\ |u\rangle}[|\langle u|\sigma_a|u\rangle| \ge 2^{-r'}]$ 至少为1/2。这意味着存在某个 $|u\rangle \in \mathcal{V}$，使得对于至少一半的 $a \in \mathcal{A}'$，都有 $|\langle u|\sigma_a|u\rangle| \ge 2^{-r'}$。另一方面，由于 M 是一个误差为 $2^{-r'}$ 的（k'，ℓ'）- 提取器，且 $\left\|P^{\tau}_{X|u}\right\|_2 \le 2^{\ell'} \cdot 2^{-n/2}$，因此至多有 $2^{-k'}$ 比例的 $a \in \mathcal{A}$ 满足 $|\langle u|\sigma_a|u\rangle| = |\langle M_a,\ P^{\tau}_{X|u}\rangle| \ge 2^{-r'}$。这意味着

$$\Pr_{a \in_R \mathcal{A}}\left[a \in \mathcal{A}'\right] \le 2 \cdot 2^{-k'} \le 2^{-2r}.$$

根据 r 的定义，这里有 $k' - 1 \ge 2r$。 $\square$

推论4.13。对于每个 $0 \le t \le T$，我们有$\mathbf{E}_{a\in_R\mathcal{A}} \left\|\tau_{XVW}^{(t,a)} - \tau_{XVW}^{(t,\infty)}\right\|_{\mathrm{Tr}} \le 2^{-2r}$。

证明。 对于每个 $w \in \mathcal{W}$，部分系统 $\tau_{XV|w}^{(t,\infty)}$ 满足引理 4.12 的条件，因为对于每个 $|v\rangle \in \mathcal{V}$，

$$\left\|P_{X|v,w}^{\tau^{(t,\infty)}}\right\|_2 \le (1-2^{-r})^{-1} \cdot 2^{\ell} \cdot 2^{-n/2} \le 2^{\ell'} \cdot 2^{-n/2}.$$

注意，对于每个 $a \in \mathcal{A}$，如果不存在 $|v\rangle \in \mathcal{V}$使得 $|\langle M_a, P_{X|v}^{\tau^{(t,\infty)}}, {}_w\rangle| \ge 2^{-r}$ （即，当对于每个 $|v\rangle \in \mathcal{V}$均成立 $G_a(v,w)$ 时），子系统 $\tau_{XV|w}^{(t,\infty)}$在通过 G_a 进行的截断中不会被触及，并且我们有 $\tau_{XV|w}^{(t,a)} = \tau_{XV|w}^{(t,\infty)}$。因此

$$\begin{aligned}
\mathop{\mathbf{E}}_{a\in_R\mathcal{A}} \left\|\tau_{XVW}^{(t,a)} - \tau_{XVW}^{(t,\infty)}\right\|_{\mathrm{Tr}} &= \sum_{w\in\mathcal{W}} \mathop{\mathbf{E}}_{a\in_R\mathcal{A}} \left\|\tau_{XV|w}^{(t,a)} - \tau_{XV|w}^{(t,\infty)}\right\|_{\mathrm{Tr}} \\
&\le \sum_{w\in\mathcal{W}} \Pr_{a\in_R\mathcal{A}} \left[\exists |v\rangle \in \mathcal{V}, |\langle M_a, P_{X|v,w}^{\tau^{(t,\infty)}}\rangle| \ge 2^{-r}\right] \cdot \mathrm{Tr}[\tau_{XV|w}^{(t,\infty)}] \\
&\le 2^{-2r} \cdot \sum_{w\in\mathcal{W}} \mathrm{Tr}[\tau_{XV|w}^{(t,\infty)}] \le 2^{-2r}.
\end{aligned}$$

□

4.3.4 演化保持迹距离

引理 4.14。对于每个 $0 \le t < T$，我们有 $\left\|\tau_{XVW}^{(t+1)} - \rho_{XVW}^{(t+1)}\right\|_{\mathrm{Tr}} \le \mathbf{E}_{a\in_R\mathcal{A}} \left\|\tau_{XVW}^{(t,a)} - \rho_{XVW}^{(t)}\right\|_{\mathrm{Tr}}$。

证明。回顾

$$\rho_{XVW}^{(t+1)} = \mathop{\mathbf{E}}_{a\in_R\mathcal{A}} \left[\sum_{x\in\mathcal{X}} |x\rangle\langle x| \otimes \Phi_{t,a,M(a,x)}\left(\rho_{VW|x}^{(t)}\right)\right],$$

$$\tau_{XVW}^{(t+1)} = \mathop{\mathbf{E}}_{a\in_R\mathcal{A}} \left[\sum_{x\in\mathcal{X}} |x\rangle\langle x| \otimes \Phi_{t,a,M(a,x)}\left(\tau_{VW|x}^{(t,a)}\right)\right].$$

因此 根据三角不等式以及量子信道在迹范数下的收缩性 s,

$$\begin{aligned}
\left\|\tau_{XVW}^{(t+1)} - \rho_{XVW}^{(t+1)}\right\|_{\mathrm{Tr}} &\le \mathop{\mathbf{E}}_{a\in_R\mathcal{A}} \left\|\sum_{x\in\mathcal{X}} |x\rangle\langle x| \otimes \left(\Phi_{t,a,M(a,x)}\left(\tau_{VW|x}^{(t,a)}\right) - \Phi_{t,a,M(a,x)}\left(\rho_{VW|x}^{(t)}\right)\right)\right\|_{\mathrm{Tr}} \\
&= \mathop{\mathbf{E}}_{a\in_R\mathcal{A}} \sum_{x\in\mathcal{X}} \left\|\Phi_{t,a,M(a,x)}\left(\tau_{VW|x}^{(t,a)}\right) - \Phi_{t,a,M(a,x)}\left(\rho_{VW|x}^{(t)}\right)\right\|_{\mathrm{Tr}} \\
&\le \mathop{\mathbf{E}}_{a\in_R\mathcal{A}} \sum_{x\in\mathcal{X}} \left\|\tau_{VW|x}^{(t,a)} - \rho_{VW|x}^{(t)}\right\|_{\mathrm{Tr}} \\
&= \mathop{\mathbf{E}}_{a\in_R\mathcal{A}} \left\|\tau_{XVW}^{(t,a)} - \rho_{XVW}^{(t)}\right\|_{\mathrm{Tr}}.
\end{aligned}$$

□

4.4 定理 2的证明

证明。首先，结合推论 4.6、4.8、4.10 和4.13 以及引理 4.14 ，我们得到

$$\left\|\tau_{XVW}^{(t+1)} - \rho_{XVW}^{(t+1)}\right\|_{\mathrm{Tr}} \le \left\|\tau_{XVW}^{(t)} - \rho_{XVW}^{(t)}\right\|_{\mathrm{Tr}} + 8 \cdot 2^{q-2r} + 2^{-5r} + 2^{-2r}.$$

由于 $\tau_{XVW}^{(0)} = \rho_{XVW}^{(0)}$，根据三角不等式可知 $\left\|\tau_{XVW}^{(T)} - \rho_{XVW}^{(T)}\right\|_{\mathrm{Tr}} \le T \cdot 10 \cdot 2^{q-2r} \le 10 \cdot 2^{q-r}$，因此

$$\left\|\tau_{XVW}^{(T,\infty)} - \rho_{XVW}^{(T)}\right\|_{\mathrm{Tr}} \le 10 \cdot 2^{q-r} + 8 \cdot 2^{q-2r} + 2^{-5r}.$$

这给出了在任意测量下 $\tau_{(T,\infty)XVW}$ 与 $\rho^{(T)}_{XVW}$ 的测量概率之差的上界，具体来说，是分支程序 ρ 的成功概率与 τ 上以下值之间差异的上界：

$$\sum_{\substack{x\in\mathcal{X},v\in\{0,1\}^q,w\in\mathcal{W}\\ \widetilde{x}(v,w)=x}} \langle x,v,w|\tau^{(T,\infty)}_{XVW}|x,v,w\rangle = \sum_{v\in\{0,1\}^q,w\in\mathcal{W}} \mathrm{Tr}[\tau^{(T,\infty)}_{X|v,w}]\cdot P^{\tau^{(T,\infty)}}_{X|v,w}(\widetilde{x}(v,w)).$$

由于 $\left\|P^{\tau^{(T,\infty)}}_{X|v,\ w}\right\|_\infty \le 2^{2\ell+9r}\cdot 2^{-n}$ 且 $\mathbf{Tr}[\tau_{(T,\ \infty)XVW}]\le 1$，上述值至多为 $2^{2\ell+9r}\cdot 2^{-n}$。因此，分支程序 ρ 的成功概率至多为（回想一下，$2\ell+9r-n\le -r$）

$$10\cdot 2^{q-r}+8\cdot 2^{q-2r}+2^{-5r}+2^{2\ell+9r}\cdot 2^{-n}=O(2^{q-r}).$$ □

引理4.5 的证明

证明引理 4.5 的第一步是分析 $P^{\tau^{(t)}}_{X|v,\ w}$ 如何根据以下规则演化

$$\tau^{(t+1)}_{XVW}=\mathop{\mathbf{E}}_{a\in_R\mathcal{A}}\left[\sum_{x\in\mathcal{X}}|x\rangle\langle x|\otimes\Phi_{t,a,M(a,x)}\left(\tau^{(t,a)}_{VW|x}\right)\right].$$

我们引入以下记号。对于每个 $a\in\mathcal{A}$ 和 $b\in\{-1,1\}$，令

$$\mathbb{I}^{a,b}=\frac{1}{2}(\vec{1}+b\cdot M_a),$$

这是一个 0-1 向量，用于指示 $M(a,x)=b$。令

$$\tau^{(t,a,b)}_{XVW}=(\mathrm{Diag}_{\ \mathbb{I}^{a,b}}\otimes\mathbb{I}_{VW})\tau^{(t,a)}_{XVW}, \tag{7}$$

这样我们就可以写成

$$\tau^{(t+1)}_{XVW}=\mathop{\mathbf{E}}_{a\in_R\mathcal{A}}\left[(\mathbb{I}_X\otimes\Phi_{t,a,1})\left(\tau^{(t,a,1)}_{XVW}\right)+(\mathbb{I}_X\otimes\Phi_{t,a,-1})\left(\tau^{(t,a,-1)}_{XVW}\right)\right]. \tag{8}$$

因此，断言3.3 表明 是 $P^{\tau^{(t,a,b)}}_{X|v'\ w'}$ 的凸组合，其中 $P^{\tau^{(t+1)}}_{X|v\ w}$，，对某些 a、b、w' 和 $|v'\rangle$ 成立。

5.1 目标分布与坏度

在考察目标分布之前，我们先证明 $P^{\tau^{(t)}}_{X|v,\ w}$ 的 ℓ_2-范数不会太大：

引理 5.1. 对于任意 $0\le t\le T$、$|v\rangle\in\mathcal{V}$、$w\in\mathcal{W}$，有

$$\left\|P^{\tau^{(t)}}_{X|v,w}\right\|_2\le 4\cdot 2^{\ell}\cdot 2^{-n/2}.$$

证明。当 $t=0$ 时，该结论显然成立，因为 $P^{\tau^{(0)}}_{X|v,\ w}$ 始终是均匀的。

现在假设 $t>0$。由引理 4.1 和引理 4.11 可知

$$\left\|P^{\tau^{(t-1,a)}}_{X|v',w'}\right\|_2\le(1-2^{-r})^{-1}\cdot 2^{\ell}\cdot 2^{-n/2}$$

对于每个 $w'\in\mathcal{W},|v'\rangle\in\mathcal{V}$和 $a\in\mathcal{A}$，由于 $\tau^{(t-1,a)}_{XVW}$ 是从 $\tau^{(t-1,\infty)}_{XVW}$ 截断得到的，因此成立。由于 $G_a(P^{\tau^{(t-1,a)}}_{X|v'\ w'})$为真，这意味着分布由 a均匀划分，我们进一步得到

$$\left\|P^{\tau^{(t-1,a,b)}}_{X|v',w'}\right\|_2=\frac{\left\|_{\mathbb{I}^{a,b}}\cdot P^{\tau^{(t-1,a)}}_{X|v',w'}\right\|_2}{\left\|_{\mathbb{I}^{a,b}}\cdot P^{\tau^{(t-1,a)}}_{X|v',w'}\right\|_1}\le 2(1-2^{-r})^{-1}\cdot\left\|P^{\tau^{(t-1,a)}}_{X|v',w'}\right\|_2\le 4\cdot 2^{\ell}\cdot 2^{-n/2}.$$

由于 $P^{\tau^{(t)}}_{X|v,\ w}$ 是 $P^{\tau^{(t-1,a,b)}}_{X|v'\ w'}$ 的凸组合，根据凸性，其 ℓ_2-范数有界于$4\cdot 2^{\ell}\cdot 2^{-n/2}$。

□

从现在起，我们用 P 表示一个固定的目标分布（稍后我们会选择引理 4.5中的分布），使得

$$2^{\ell} \cdot 2^{-n/2} \le \|P\|_2 \le 4 \cdot 2^{\ell} \cdot 2^{-n/2}.$$

我们想要界定 $\langle P^{\tau^{(t)}}_{X|v},w, P\rangle$ 的进展，它起初为 2^{-n} at $t = 0$，并在 $P^{\tau^{(t)}}_{X|v},w = P$ 时变为至少 $2^{2\ell} \cdot 2^{-n}$。注意，根据柯西-施瓦茨不等式，我们总有

$$\langle P^{\tau^{(t)}}_{X|v,w}, P\rangle \le \left\|P^{\tau^{(t)}}_{X|v,w}\right\|_2 \|P\|_2 \le 16 \cdot 2^{2\ell} \cdot 2^{-n}. \tag{9}$$

为了界定进展，我们引入一些新的记号。对于部分系统上的任意上标（例如 (t, a)），我们用 σ_{XVW} 表示 τ_{XVW}(Diag $P \otimes \mathbb{I}_{VW}$)。注意

$$\mathrm{Tr}[\sigma_{X|v,w}] = \mathrm{Tr}[\tau_{X|v,w}\,\mathrm{Diag}\,P] = \mathrm{Tr}[\tau_{X|v,w}] \cdot \langle P^{\tau}_{X|v,w}, P\rangle.$$

类似地，$P^{\sigma}_{X|v,w}$ 可以通过以下方式从 $P^{\tau}_{X|v,w}$ 推导得出

$$P^{\sigma}_{X|v,w}(x) = \frac{\mathrm{Tr}[\tau_{X|v,w}]}{\mathrm{Tr}[\sigma_{X|v,w}]} \cdot P^{\tau}_{X|v,w}(x) \cdot P(x) = \frac{P^{\tau}_{X|v,w}(x) \cdot P(x)}{\langle P^{\tau}_{X|v,w}, P\rangle}. \tag{10}$$

因此，我们可以将 ℓ_2 范数的 $P^{\sigma}_{X|v,w}$ 界定为

$$\left\|P^{\sigma}_{X|v,w}\right\|_2 \le \frac{1}{\langle P^{\tau}_{X|v,w}, P\rangle} \cdot \left\|P^{\tau}_{X|v,w}\right\|_{\infty} \cdot \|P\|_2.$$

现在，我们可以找出 $\langle P^{\tau^{(t)}}_{X|v},w, P\rangle$ 大幅增加的位置；当内积没有被某个 $a \in \mathcal{A}$ 均匀划分时，就会发生这种情况（我们稍后会在分析中说明原因）。形式上，在阶段 $0 \le t < T$，如果

$$\exists |v\rangle \in \mathcal{V}, \text{ s.t. } |\langle M_a, P^{\sigma^{(t,a)}}_{X|v,w}\rangle| > 2^{-r} \text{ and } \langle P^{\tau^{(t,a)}}_{X|v,w}, P\rangle \ge \frac{1}{2} \cdot 2^{-n}. \tag{11}$$

引理 5.2。对于每个 $0 \le t < T$ 和 $w \in \mathcal{W}$，我们有

$$\Pr_{a \in_R \mathcal{A}}[(w, a) \textit{ is bad}] \le 2^{-k}.$$

证明。由于 $\tau_{(t,a)XVW}$是由 $\tau^{(t,\infty)}_{XVW}$截断得到的，引理 4.1表明，对于每个 $|v\rangle \in \mathcal{V}$, $w \in \mathcal{W}$和 $a \in \mathcal{A}$，存在 $|v'\rangle \in \mathcal{V}$ 使得

$$P^{\tau^{(t,a)}}_{X|v,w} = P^{\tau^{(t,\infty)}}_{X|v',w}$$

并且根据式（10），还可推出

$$P^{\sigma^{(t,a)}}_{X|v,w} = P^{\sigma^{(t,\infty)}}_{X|v',w}.$$

现在固定某个 $w \in \mathcal{W}$，并令 $\mathcal{A}' \subseteq \mathcal{A}$ 为满足以下条件的所有 $a \in \mathcal{A}$ 构成的集合：

$$\exists |v\rangle \in \mathcal{V}, \text{ s.t. } |\langle M_a, P^{\sigma^{(t,\infty)}}_{X|v,w}\rangle| > 2^{-r} \text{ and } \langle P^{\tau^{(t,\infty)}}_{X|v,w}, P\rangle \ge \frac{1}{2} \cdot 2^{-n}.$$

那么 $\mathcal{A}'$ 包含所有满足 a 的 $w, a)$ 均为坏的，而我们的目标是界定 $\mathcal{A}'$ 在 $\mathcal{A}$中所占的比例。在证明的其余部分，我们暂时省略上标，并将 $\tau^{(t,\infty)}$ 和 $\sigma^{(t,\infty)}$ 分别简写为 τ和 σ。出于与引理 4.12 相同的理由，我们可以假设 $\tau_{V|w} = \mathbb{I}_V$，因此

$$\langle v|\sigma_{V|w}|v\rangle = \mathrm{Tr}[\sigma_{X|v,w}] = \langle P^{\tau}_{X|v,w}, P\rangle, \text{ and } \mathrm{Tr}[\sigma_{XV|w}] = \langle P^{\tau}_{X|w}, P\rangle \le 16 \cdot 2^{2\ell} \cdot 2^{-n}.$$

其中最后　　不等式由引理 4.11 和柯西-施瓦茨不等式得出，推导方式与式（9　　）.

假设我们有对角化 $\sigma_{V|w} = U^\dagger D U$，其中 U是酉的， D是对角且非负的。令 $\mathcal{V}' \subseteq \mathcal{V}$ 为由计算基向量$|e\rangle \in \mathcal{V}$中满足 $\langle e|D|e\rangle \geq 2^{-4r} \cdot 2^{-2\ell} \cdot 2^{-n}$的那些向量张成的子空间 $U^\dagger|e\rangle$ 。因此，对于每个 $|v\rangle \in \mathcal{V}'$，我们都有

$$\langle P^\tau_{X|v,w}, P\rangle = \mathrm{Tr}[\sigma_{X|v,w}] \geq 2^{-4r} \cdot 2^{-2\ell} \cdot 2^{-n}.$$

我们声称，对于任意 $a \in \mathcal{A}'$，都存在 $|v\rangle \in \mathcal{V}'$使得 $|\langle M_a,\ P^\sigma_{X|v,\ w}\rangle| > \frac{1}{2} \cdot 2^{-r}$。为证明这一断言，令 Π 为从 $\mathcal{V}$到 $\mathcal{V}'$的投影算子，则 $(\mathbb{I}_X \otimes \Pi)\ ^{\sigma_{XV|w}}\ (\mathbb{I}_X \otimes \Pi)$ 在概念上可视为一个截断的部分系统 $\sigma^{|G}_{XV|w}$其中 G（v，w）在固定的 $[\sigma_{X|v,w}] \geq 2^{-4r-2\ell} \cdot 2^{-n}$ 下，当 $\mathbf{Tr}\ w$时成立。由引理 4.3 可得

$$\left\|\sigma^{|G}_{XV|w} - \sigma_{XV|w}\right\|_{\mathrm{Tr}} \leq 3 \cdot 2^q \cdot \sqrt{2^{-4r-2\ell-n} \cdot \mathrm{Tr}[\sigma_{XV|w}]} \leq 12 \cdot 2^{q-2r} \cdot 2^{-n}.$$

由于 $a \in \mathcal{A}'$，假设对于 $|u\rangle \in \mathcal{V}$，我们有 $|\langle M_a,\ P^\sigma_{X|u,\ w}\rangle| > 2^{-r}$ 且 $\mathbf{Tr}[\sigma_{X|u,w}] = \langle P^\tau_{X|u,\ w}, P\rangle \geq \frac{1}{2} \cdot 2^{-n}$。令 $|v\rangle \sim \Pi|u\rangle$，则有

$$\begin{aligned}
\left\|P^\sigma_{X|u,w} - P^\sigma_{X|v,w}\right\|_1 = \left\|P^\sigma_{X|u,w} - P^{\sigma^{|G}}_{X|u,w}\right\|_1 &\leq \left\|\frac{\sigma_{X|u,w}}{\mathrm{Tr}[\sigma_{X|u,w}]} - \frac{\sigma^{|G}_{X|u,w}}{\mathrm{Tr}[\sigma^{|G}_{X|u,w}]}\right\|_{\mathrm{Tr}} \\
&\leq \left\|\frac{\sigma_{X|u,w}}{\mathrm{Tr}[\sigma_{X|u,w}]} - \frac{\sigma^{|G}_{X|u,w}}{\mathrm{Tr}[\sigma_{X|u,w}]}\right\|_{\mathrm{Tr}} + \left\|\frac{\sigma^{|G}_{X|u,w}}{\mathrm{Tr}[\sigma_{X|u,w}]} - \frac{\sigma^{|G}_{X|u,w}}{\mathrm{Tr}[\sigma^{|G}_{X|u,w}]}\right\|_{\mathrm{Tr}} \\
&= \left\|\frac{\sigma_{X|u,w}}{\mathrm{Tr}[\sigma_{X|u,w}]} - \frac{\sigma^{|G}_{X|u,w}}{\mathrm{Tr}[\sigma_{X|u,w}]}\right\|_{\mathrm{Tr}} + \left|\frac{1}{\mathrm{Tr}[\sigma_{X|u,w}]} - \frac{1}{\mathrm{Tr}[\sigma^{|G}_{X|u,w}]}\right| \cdot \mathrm{Tr}[\sigma^{|G}_{X|u,w}] \\
&= \frac{\left\|\sigma_{X|u,w} - \sigma^{|G}_{X|u,w}\right\|_{\mathrm{Tr}}}{\mathrm{Tr}[\sigma_{X|u,w}]} + \frac{\left|\mathrm{Tr}[\sigma^{|G}_{X|u,w}] - \mathrm{Tr}[\sigma_{X|u,w}]\right|}{\mathrm{Tr}[\sigma_{X|u,w}]} \\
&\leq \frac{2\left\|\sigma_{X|u,w} - \sigma^{|G}_{X|u,w}\right\|_{\mathrm{Tr}}}{\mathrm{Tr}[\sigma_{X|u,w}]} \leq \frac{2\left\|\sigma_{XV|w} - \sigma^{|G}_{XV|w}\right\|_{\mathrm{Tr}}}{\mathrm{Tr}[\sigma_{X|u,w}]} \leq 48 \cdot 2^{q-2r} \leq \frac{1}{2} \cdot 2^{-r},
\end{aligned}$$

其中最后一步由 $q \leq r - 7$ 得出。因此

$$|\langle M_a, P^\sigma_{X|v,w}\rangle| \geq |\langle M_a, P^\sigma_{X|u,w}\rangle| - \left\|P^\sigma_{X|u,w} - P^\sigma_{X|v,w}\right\|_1 > \frac{1}{2} \cdot 2^{-r}.$$

与引理 4.12 的证明类似，对于每个 $a \in \mathcal{A}'$ 令

$$\pi_a = \mathrm{Tr}_X[(\mathrm{Diag}\, M_a \otimes U^\dagger D^{-1/2} U) \cdot \sigma_{XV|w} \cdot (\mathbb{I}_X \otimes U^\dagger D^{-1/2} U)]$$

它是作用于 $\mathcal{V}$ 上的厄米算子。对于每个 $|v\rangle \in \mathcal{V}$，令 $|v'\rangle \sim U^\dagger D^{1/2} U|v\rangle$。回顾一下，$\sigma_{V|w} = U^\dagger D U$，因此

$$\begin{aligned}
P^\sigma_{X|v,w} &= \frac{\mathrm{diag}\ (\mathbb{I}_X \otimes \langle v|)\sigma_{XV|w}(\mathbb{I}_X \otimes |v\rangle)}{\langle v|\sigma_{V|w}|v\rangle} \\
&= \frac{\mathrm{diag}\ (\mathbb{I}_X \otimes \langle v'|U^\dagger D^{-1/2} U)\sigma_{XV|w}(\mathbb{I}_X \otimes U^\dagger D^{-1/2} U|v'\rangle)}{\langle v'|U^\dagger D^{-1/2} U \sigma_{V|w} U^\dagger D^{-1/2} U|v'\rangle} \\
&= \mathrm{diag}\ (\mathbb{I}_X \otimes \langle v'|U^\dagger D^{-1/2} U)\sigma_{XV|w}(\mathbb{I}_X \otimes U^\dagger D^{-1/2} U|v'\rangle).
\end{aligned}$$

这意味着

$$\langle v'|\pi_a|v'\rangle = \left\langle M_a, \operatorname{diag}\left(\mathbb{I}_X \otimes \langle v'|U^\dagger D^{-1/2}U\right)\sigma_{XV|w}\left(\mathbb{I}_X \otimes U^\dagger D^{-1/2}U|v'\rangle\right)\right\rangle = \langle M_a, P^\sigma_{X|v,w}\rangle.$$

我们在上文证明了存在 $|v\rangle \in \mathcal{V}'$，因此也存在 $|v'\rangle \in \mathcal{V}'$，使得

$$|\langle v'|\pi_a|v'\rangle| = \left|\langle M_a, P^\sigma_{X|v,w}\rangle\right| \geq \frac{1}{2}\cdot 2^{-r},$$

这意味着，对于 $\Pi\pi_a\Pi$， π_a 在 $\mathcal{V}'$上的限制满足 $\|\Pi\pi_a\Pi\|_2 \geq \frac{1}{2}\cdot 2^{-r}$。现在考虑一个均匀随机单位向量 $|v'\rangle$，它属于 $\mathcal{V}'$。根据引理 3.1 ，我们知道存在某个绝对常数 c，使得

$$\Pr_{|v\rangle'}\left[|\langle v'|\sigma_a|v'\rangle| \geq 2^{-r'}\right] \geq 1 - 2^{(q+r+1-r')/2}c - e^{-2^q} \geq 1 - 2^{-r}c - e^{-1} \geq \frac{1}{2}.$$

因此，对于随机向量 $|v\rangle \sim U^\dagger D^{-1/2}U|v'\rangle$，其中 $|v'\rangle$在 $\mathcal{V}'$中均匀分布，我们得出

$$\Pr_{|v\rangle}\left[|\langle M_a, P^\sigma_{X|v,w}\rangle| \geq 2^{-r'}\right] \geq \frac{1}{2}.$$

另一方面，由于 $|v'\rangle \in \mathcal{V}'$，也有 $|v\rangle \in \mathcal{V}'$，因此$\langle P^\tau_{X|v,\ w},\ P\rangle \geq 2^{-4r}\cdot 2^{-2\ell}\cdot 2^{-n}$始终成立。因此，存在一个 $|v\rangle \in \mathcal{V}$，它同时满足

$$\langle P^\tau_{X|v,w}, P\rangle \geq 2^{-4r}\cdot 2^{-2\ell}\cdot 2^{-n} \quad \text{and} \quad |\langle M_a, P^\sigma_{X|v,w}\rangle| \geq 2^{-r'}$$

对于至少 $1/2$ 比例的 $a \in \mathcal{A}'$。由于

$$\left\|P^\sigma_{X|v,w}\right\|_2 \leq \frac{1}{\langle P^\tau_{X|v,w}, P\rangle}\cdot\left\|P^\tau_{X|v,w}\right\|_\infty\cdot\left\|P\right\|_2 \leq 4\cdot 2^{5\ell+13r}\cdot 2^{-n/2} = 2^{\ell'}\cdot 2^{-n/2},$$

并且 M是一个误差为 $2^{-r'}$的（k'，ℓ'）-提取器，因此至多有 $2^{-k'}$比例的 $a \in \mathcal{A}$满足$|\langle M_a, P^\sigma_{X|v',\ w}\rangle| \geq 2^{-r'}$，这意味着

$$\Pr_{a\in_R\mathcal{A}}[(w,a) \text{ is bad}] \leq \Pr_{a\in_R\mathcal{A}}[a \in \mathcal{A}'] \leq 2\cdot 2^{-k'} = 2^{-k}.$$ □

5.2 坏度等级

在阶段 t，对于每个经典内存状态 $w \in \mathcal{W}$，我们统计通向它的路径变坏的次数。这是一个取决于 $a \in \mathcal{A}$先前随机选择的随机变量。我们将其存储在另一个经典寄存器 B中，称为坏度等级，其取值为 $\beta \in \{0,\ldots,T\}$。初始值设为 0，即令

$$\tau^{(0)}_{XVWB} = \tau^{(0)}_{XVW}\otimes|0\rangle\langle 0|_B.$$

我们通过对每个阶段 $0 \leq t < T$的联合系统 τ_{XVWB}应用以下更新规则，确保 B 的分布始终只取决于 W，并且在给定 W的条件下独立于 X 和 V：

- 真 ncation 步骤独立于 B 执行。因此，对于每个 $a \in \mathcal{A}$，我们令 l et

$$\tau^{(t,a)}_{XVWB} = \sum_{w\in\mathcal{W}}\tau^{(t,a)}_{XV|w}\otimes|w\rangle\langle w|\otimes\operatorname{Diag}P^{\tau^{(t)}}_{B|w}. \tag{12}$$

- B 的值在演化步骤之前更新；对于每个 $a \in \mathcal{A}$ 和 $b \in \{-1, 1\}$，我们令

$$\tau_{XVWB}^{(t,a,b)} = (\mathrm{Diag}_{\mathbb{I}^{a,b}} \otimes \mathbb{I}_V \otimes U_a)\tau_{XVWB}^{(t,a)}(\mathbb{I}_{XV} \otimes U_a^\dagger).$$

这里，U_a 是一个置换算子，它依赖于 $\tau_{XVW}^{(t,a)}$，并作用于 $\mathcal{W} \otimes \{0, \ldots, T\}$，使得

$$U_a|w\rangle|\beta\rangle = \begin{cases} |w\rangle|(\beta+1) \bmod (T+1)\rangle & \text{if } (w,a) \text{ is bad,} \\ |w\rangle|\beta\rangle & \text{otherwise.} \end{cases}$$

- 对于演化步骤，我们将量子信道 $\Phi_{t,\ a,\ b}$ 作用于存储器 W 和 V，得到

$$\tau_{XVWB}^{(t+1)} = \mathop{\mathbf{E}}_{a \in_R \mathcal{A}} \left[(\mathbb{I}_X \otimes \Phi_{t,a,1} \otimes \mathbb{I}_B)\big(\tau_{XVWB}^{(t,a,1)}\big) + (\mathbb{I}_X \otimes \Phi_{t,a,-1} \otimes \mathbb{I}_B)\big(\tau_{XVWB}^{(t,a,-1)}\big) \right].$$

请注意，演化步骤可能会在 X，V 和 B 之间引入依赖关系。然而，由于我们处理截断步骤的方式，这些依赖关系之后会被消除（12），因此不会影响我们的证明。

我们可以验证，上文定义的组合部分系统 $\tau_{XVWB}^{(t)}$ 与前几节讨论的部分系统 $\tau^{(t)}{}_{XVW}$ 一致，即始终满足 $\mathrm{Tr}_B[\tau_{XVWB}^{(t)}] = \tau^{(t)}{}_{XVW}$：

- 对于截断步骤，可以直接验证

$$\mathrm{Tr}_B[\tau_{XVWB}^{(t,a)}] = \sum_{w \in \mathcal{W}} \tau_{XV|w}^{(t,a)} \otimes |w\rangle\langle w| = \tau_{XVW}^{(t,a)}.$$

- 置换算子 U_a 作用于 $\mathcal{W}$ 时等同于恒等算子，因为

$$\mathrm{Tr}_B \left[U_a |w, \beta\rangle\langle w, \beta| U_a^\dagger \right] = |w\rangle\langle w|.$$

回想式（7），$\tau^{(t,\ a,\ b)}{}_{XVW} = (\mathrm{Diag}\ \mathbb{I}^{a,\ b} \otimes \mathbb{I}_V)\tau_{XVW}^{(t,a)}$，可得 $\mathrm{Tr}_B[\tau_{XVWB}^{(t,a,b)}] = \tau_{XVW}^{(t,a,b)}$。

- Th演化步骤可以直接从公式中验证，无需 B (式 (8))：

$$\tau_{XVW}^{(t+1)} = \mathop{\mathbf{E}}_{a \in_R \mathcal{A}} \left[(\mathbb{I}_X \otimes \Phi_{t,a,1})\big(\tau_{XVW}^{(t,a,1)}\big) + (\mathbb{I}_X \otimes \Phi_{t,a,-1})\big(\tau_{XVW}^{(t,a,-1)}\big) \right].$$

因此，先前关于 $\tau_{XVW}^{(t)}$ 证明的所有性质都得以保留。此外，我们还证明了关于坏度等级的以下两个性质。

引理 5.3。对于任意 $0 \le t \le T$、$|v\rangle \in \mathcal{V}$ 和 $w \in \mathcal{W}$，有

$$\langle P_{X|v,w}^{\tau^{(t)}}, P\rangle \le \sum_{\beta=0}^{T} P_{B|w}^{\tau^{(t)}}(\beta) \cdot 2^{\beta} \cdot 2^{-n} \cdot (1 - 2^{-r})^{-3t}.$$

证明。我们对 t 使用归纳法来证明。对于 $t = 0$，该引理成立，因为 $\langle P_{X|v,\ w}^{\tau^{(t)}}、P\rangle = 2^{-n}$ 和 $P_{B|w}^{\tau^{(t)}}(0) = 1$。

假设引理对某个 $t < T$ 成立。通过与引理4.11 类似的论证，并多次应用引理4.1，我们知道，对于每个 $|v\rangle \in \mathcal{V}$，$w \in \mathcal{W}$ 和 $a \in \mathcal{A}$，都存在 $|v'\rangle$ 和 $|v''\rangle \in \mathcal{V}$，使得

$$\langle P_{X|v,w}^{\tau^{(t,a)}}, P\rangle = \langle P_{X|v',w}^{\tau^{(t,\circ)}}, P\rangle \le (1 - 2^{-r})^{-1} \langle P_{X|v',w}^{\tau^{(t,\star)}}, P\rangle = (1 - 2^{-r})^{-1} \langle P_{X|v'',w}^{\tau^{(t)}}, P\rangle,$$

因此

$$\langle P^{\tau(t,a)}_{X|v,w}, P\rangle \le \sum_{\beta=0}^{T} P^{\tau(t)}_{B|w}(\beta)\cdot 2^{\beta}\cdot 2^{-n}\cdot(1-2^{-r})^{-3t-1}. \tag{13}$$

此外，由 G_a 得到的截断步骤意味着 $|\langle M_a, P^{\tau(t,a)}_{X|v,w}\rangle| \le 2^{-r}$。也就是说，对于两个 $b\in\{-1,1\}$，

$$1-2^{-r} \le 2\left\| \mathbb{I}^{a,b}\cdot P^{\tau(t,a)}_{X|v,w}\right\|_1 \le 1+2^{-r}.$$

因此，无条件地有

$$\langle P^{\tau(t,a,b)}_{X|v,w}, P\rangle = \frac{\langle \mathbb{I}^{a,b}\cdot P^{\tau(t,a)}_{X|v,w}, P\rangle}{\left\| \mathbb{I}^{a,b}\cdot P^{\tau(t,a)}_{X|v,w}\right\|_1} \le 2(1-2^{-r})^{-1}\cdot\langle P^{\tau(t,a)}_{X|v,w}, P\rangle. \tag{14}$$

当内积被均匀划分时，即 $|\langle M_a, P^{\sigma(t,a)}_{X|v,w}\rangle| \le 2^{-r}$，我们还可以得到

$$\langle \mathbb{I}^{a,b}\cdot P^{\tau(t,a)}_{X|v,w}, P\rangle \le \frac{1}{2}(1+2^{-r})\langle P^{\tau(t,a)}_{X|v,w}, P\rangle \le \frac{1}{2}(1-2^{-r})^{-1}\langle P^{\tau(t,a)}_{X|v,w}, P\rangle,$$

这意味着

$$\langle P^{\tau(t,a,b)}_{X|v,w}, P\rangle = \frac{\langle \mathbb{I}^{a,b}\cdot P^{\tau(t,a)}_{X|v,w}, P\rangle}{\left\| \mathbb{I}^{a,b}\cdot P^{\tau(t,a)}_{X|v,w}\right\|_1} \le (1-2^{-r})^{-2}\cdot\langle P^{\tau(t,a)}_{X|v,w}, P\rangle. \tag{15}$$

现在有三种情况需要讨论：

- 若 (w,a) 是坏的，则对于每个 $\beta>0$，有 $P^{\tau(t,a,b)}_{B|w}(\beta) = P^{\tau(t)}_{B|w}(\beta-1)$。注意，$P^{\tau(t)}_{B|w}(T)=0$当 $t<T$ 时，因此式 (13) 和式 (14) 意味着

$$\begin{aligned}\langle P^{\tau(t,a,b)}_{X|v,w}, P\rangle &\le \sum_{\beta=0}^{T-1} P^{\tau(t)}_{B|w}(\beta)\cdot 2^{\beta+1}\cdot 2^{-n}\cdot(1-2^{-r})^{-3t-2}\\ &\le \sum_{\beta=0}^{T} P^{\tau(t,a,b)}_{B|w}(\beta)\cdot 2^{\beta}\cdot 2^{-n}\cdot(1-2^{-r})^{-3(t+1)}.\end{aligned}$$

- 如果（w、a）不是坏的，并且 $|\langle M_a、\ P^{\sigma}\ (_t、\ _a)\ _{X|v、\ w}\rangle| \le 2^{-r}$，那么对于每个 $\beta\ge 0$，都有 $P^{\tau}\ (_t、\ _a、\ _b)\ _{B|w}\ (\beta)\ = P^{\tau}\ (_t)\ _{B|w}\ (\beta)$。因此，式（13）和式（15）意味着

$$\begin{aligned}\langle P^{\tau(t,a,b)}_{X|v,w}, P\rangle &\le \sum_{\beta=0}^{T} P^{\tau(t)}_{B|w}(\beta)\cdot 2^{\beta}\cdot 2^{-n}\cdot(1-2^{-r})^{-3t-3}\\ &= \sum_{\beta=0}^{T} P^{\tau(t,a,b)}_{B|w}(\beta)\cdot 2^{\beta}\cdot 2^{-n}\cdot(1-2^{-r})^{-3(t+1)}.\end{aligned}$$

- 若（w，a）不是坏的，且 $|\langle M_a,\ P^{\sigma}\ _{(t,\ a)\ X|v,\ w}\rangle| > 2^{-r}$根据坏性的定义（11），我们必有 $\langle P^{\tau}\ _{(t,\ a)\ X|v,\ w},\ P\rangle < \frac{1}{2}\cdot 2^{-n}$。因此，由式（14），

$$\langle P^{\tau(t,a,b)}_{X|v,w}, P\rangle < (1-2^{-r})^{-1}\cdot 2^{-n} \le \sum_{\beta=0}^{T} P^{\tau(t,a,b)}_{B|w}(\beta)\cdot 2^{\beta}\cdot 2^{-n}\cdot(1-2^{-r})^{-3(t+1)}.$$

最后一个不等式由 $\sum_{\beta=0}^{T} P_{B|w}^{\tau^{(t,a,b)}}(\beta) \cdot 2^{\beta} \cdot 2^{-n} \cdot (1-2^{-r})^{-3(t+1)} \geq 2^{-n}(1-2^{-r})^{-3(t+1)}$ 得出。因此，我们从这三种情形中都得到相同的结论。

对于演化步骤，由于 B 是经典的，我们可以将 X 和 B 视为一个整体，并在 $P_{XB|v}^{\tau^{(t+1)}},w$ 上应用命题3.3，该命题断言 $P_{XB|v}^{\tau^{(t+1)}},w$ 是某些 a、b、w' 和 $|v'\rangle$ 对应的 $P_{XB|v'}^{\tau^{(t,a,b)}},w'$ 的凸组合。然后利用线性性，我们得出 [6]

$$\langle P_{X|v,w}^{\tau^{(t+1)}}, P\rangle \leq \sum_{\beta=0}^{T} P_{B|w}^{\tau^{(t+1)}}(\beta) \cdot 2^{\beta} \cdot 2^{-n} \cdot (1-2^{-r})^{-3(t+1)}. \qquad \square$$

引理 5.4. 对于每个 $0 \leq \beta \leq t \leq T$，我们有

$$\langle \beta | \tau_B^{(t)} | \beta \rangle \leq 2^{-k\beta} \binom{t}{\beta}.$$

证明。我们对 t 使用归纳法证明。对于 $t=0$，引理成立，因为 $\tau^{(0)}{}_B = |0\rangle\langle 0|_B$。还要注意，当 $\beta = 0$ 时，对于每个 t，该引理显然成立。

现在假设该引理对某个 t 成立。根据定义，我们有

$$\tau_B^{(t+1)} = \mathop{\mathbf{E}}_{a \in_R \mathcal{A}} [\tau_B^{(t,a,1)} + \tau_B^{(t,a,-1)}] = \mathop{\mathbf{E}}_{a \in_R \mathcal{A}} \mathrm{Tr}_W [U_a \tau_{WB}^{(t,a)} U_a^{\dagger}].$$

因此

$$\langle \beta | \tau_B^{(t+1)} | \beta \rangle = \sum_{w \in \mathcal{W}} \mathop{\mathbf{E}}_{a \in_R \mathcal{A}} \left[\langle w, \beta | U_a \tau_{WB}^{(t,a)} U_a^{\dagger} | w, \beta \rangle \right].$$

根据 5.2 引理，我们知道对于每个 $w \in \mathcal{W}$，(w,a) 对 $a \in_R \mathcal{A}$ 为坏的概率至多为 2^{-k}。换言之，对于每个 $\beta > 0$,

$$U_a^{\dagger} |w, \beta\rangle = \begin{cases} |w,\beta\rangle, & \text{w.p. } \geq 1-2^{-k} \\ |w,\beta-1\rangle, & \text{w.p. } \leq 2^{-k} \end{cases}$$

其中，概率是对 a 的随机选择取的。这意味着

$$\begin{aligned} \langle \beta | \tau_B^{(t+1)} | \beta \rangle &\leq \sum_{w \in \mathcal{W}} \langle w, \beta | \tau_{WB}^{(t,a)} | w, \beta \rangle + 2^{-k} \sum_{w \in \mathcal{W}} \langle w, \beta-1 | \tau_{WB}^{(t,a)} | w, \beta-1 \rangle \\ &= \langle \beta | \tau_B^{(t,a)} | \beta \rangle + 2^{-k} \cdot \langle \beta-1 | \tau_B^{(t,a)} | \beta-1 \rangle. \end{aligned}$$

注意

$$\tau_B^{(t,a)} = \sum_{w \in \mathcal{W}} \mathrm{Tr}[\tau_{XV|w}^{(t,a)}] \cdot \mathrm{Diag}\, P_{B|w}^{\tau^{(t)}} \leq \sum_{w \in \mathcal{W}} \mathrm{Tr}[\tau_{XV|w}^{(t)}] \cdot \mathrm{Diag}\, P_{B|w}^{\tau^{(t)}} = \tau_B^{(t)},$$

因此我们得出

$$\begin{aligned} \langle \beta | \tau_B^{(t+1)} | \beta \rangle &\leq \langle \beta | \tau_B^{(t)} | \beta \rangle + 2^{-k} \cdot \langle \beta-1 | \tau_B^{(t)} | \beta-1 \rangle \\ &\leq 2^{-k\beta} \binom{t}{\beta} + 2^{-k} \cdot 2^{-k(\beta-1)} \binom{t}{\beta-1} = 2^{-k\beta} \binom{t+1}{\beta}. \end{aligned} \qquad \square$$

有了上述引理，我们终于可以证明引理 4.5。

[6] 需要注意的是，在 $\tau^{(t+1)}$ 中，X 和 B 并不独立。（在 $\tau^{(t,\,a,\,b)}$ 中，它们是独立的（在给定 v', w' 的条件下）。）尽管如此，这里并不需要或使用 X, B（在 $\tau^{(t+1)}$ 中）的独立性；通过线性性，对所有不等式取相应的凸组合，即可得出最终的不等式。

引理 4.5 的证明。对于目标分布 $P = P^{\tau^{(t)}}_{X|v}, w$ 我们有 $\langle P^{\tau^{(t)}}_{X|v}, w, P\rangle > 2^{2\ell} \cdot 2^{-n}$，因此根据引理 5.3，

$$\sum_{\beta=0}^{T} P^{\tau^{(t)}}_{B|w}(\beta) \cdot 2^{\beta} \cdot (1-2^{-r})^{-3t} > 2^{2\ell}.$$

由于 $t \le T \le 2^{r-2}$，我们有 $(1-2^{-r})^{-3t} \le 2$，因此

$$\sum_{\beta=\ell}^{T} P^{\tau^{(t)}}_{B|w}(\beta) \cdot 2^{\beta} > \frac{1}{2}\left(2^{2\ell} - 2 \cdot \sum_{\beta=0}^{\ell-1} 2^{\beta}\right) > 2^{\ell}.$$

另一方面，对于任意 $\beta \ge \ell$，由引理 5.4，

$$\mathrm{Tr}[\tau^{(t)}_{B|w}] \cdot P^{\tau^{(t)}}_{B|w}(\beta) \le \langle\beta|\tau^{(t)}_{B}|\beta\rangle \le (2^{-k}t)^{\beta} < 2^{-(k-r)\beta},$$

因此由式（6），

$$\mathrm{Tr}[\tau^{(t)}_{X|v,w}] \le \mathrm{Tr}[\tau^{(t)}_{B|w}] < 2^{-\ell}\sum_{\beta=\ell}^{T} 2^{-(k-r)\beta} \cdot 2^{\beta} \le 2 \cdot 2^{-(k-r)\ell} \le 2^{-2m} \cdot 2^{-4r}. \qquad \square$$

致谢

我们感谢 Uma Girish 在本文初稿方面提出的许多重要讨论和建议，也感谢匿名审稿人提供的宝贵意见。

参考文献

[AAB+19] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando GSL Brandao, David A Buell 等。使用可编程超导处理器实现量子霸权。《自然》，574(7779):505–510，2019年。2[Aar18] Scott Aaronson。量子态的阴影层析。载于第50届 ACM SIGACT 计算理论研讨会论文集，第325–338页，2018年。2[ACQ22] Dorit Aharonov、Jordan Cotler 和 Xiao-Liang Qi。量子算法测量。《自然·通讯》，13(1):1–9，2022年。2[ADR02] Yonatan Aumann、Yan Zong Ding 和 Michael O Rabin。有界存储模型中的永久安全性。IEEE 信息论汇刊，48(6):1668–1680，2002年。4[AR99] Yonatan Aumann 和 Michael O Rabin。有限存储空间模型中的信息论安全通信。载于国际密码学年会论文集，第65–79页。施普林格，1999年。4[Aud07] Koenraad MR Audenaert。冯·诺伊曼熵的尖锐连续性估计。《物理学杂志A：数学与理论》，40(28):8127，2007年。34[BCL20] Sebastien Bubeck、Sitan Chen 和 Jerry Li。最优量子性质测试离不开纠缠。载于2020年 IEEE 第61届计算机科学基础年会（FOCS）论文集，第692–703页。IEEE，2020年。2

[BGY18] Paul Beame、Shayan Oveis Gharan 和 Xin Yang。基于随机求值学习有限函数的时间-空间权衡及其在多项式中的应用。载于 COLT，机器学习研究论文集第75卷，第843–856页。PMLR，2018年。2[BY21] Anne Broadbent 和 Peter Yuen。基于有界量子存储模型和计算假设的设备无关不经意传输。arXiv 预印本 arXiv:2111.08595，2021年。4[CCHL22] Sitan Chen、Jordan Cotler、Hsin-Yuan Huang 和 Jerry Li。使用量子存储器与不使用量子存储器进行学习之间的指数分离。载于2021年 IEEE 第62届计算机科学基础年会（FOCS）论文集，第574–585页。IEEE，2022年。2[CLO22] Sitan Chen、Jerry Li 和 Ryan O'Donnell。迈向使用非相干测量的实例最优态认证。载于学习理论会议论文集，第2541–2596页。PMLR，2022年。2[CM97] Christian Cachin 和 Ueli Maurer。针对内存受限对手的无条件安全性。载于国际密码学年会论文集，第292–306页。施普林格，1997年。4[CW01] Anthony Carbery 和 James Wright。凸体上多项式的分布不等式与 L^q 范数不等式，见 $\mathbb{R}^n$。《数学研究快报》，8(3):233–248，2001年。10[DFR+07] Ivan B Damg˚ard、Serge Fehr、Renato Renner、Louis Salvail 和 Christian Schaffner。一种具有应用的紧致高阶熵量子不确定性关系。载于国际密码学年会论文集，第360–378页。施普林格，2007年。4[DFSS07] Ivan B Damg˚ard、Serge Fehr、Louis Salvail 和 Christian Schaffner。有界量子存储模型中的安全识别与量子密钥分发。载于国际密码学年会论文集，第342–359页。施普林格，2007年。4[DFSS08] Ivan B Damg˚ard、Serge Fehr、Louis Salvail 和 Christian Schaffner。有界量子存储模型中的密码学。《SIAM 计算杂志》，37(6):1865–1890，2008年。4[DM02] Stefan Dziembowski 和 Ueli Maurer。有界存储模型的紧致安全性证明。载于第34届 ACM 计算理论研讨会论文集，第341–350页，2002年。4[DM04] Stefan Dziembowski 和 Ueli Maurer。关于如何在有界存储模型中生成初始密钥。载于密码技术理论与应用国际会议论文集，第126–137页。施普林格，2004年。4[DQW21] Yevgeniy Dodis、Willy Quach 和 Daniel Wichs。多说少记：重新审视有界存储模型中的密码学。密码学电子印本档案，2021年。4[DQW22] Yevgeniy Dodis、Willy Quach 和 Daniel Wichs。有界存储模型中的身份认证。载于密码技术理论与应用国际会议论文集，第737–766页。施普林格，2022年。4

[DR02] Yan Zong Ding 和 Michael O Rabin。超加密与永久安全性。载于《计算机科学理论方面年度研讨会》，第1–26页。施普林格，2002年。4[FvdG99] Christopher A Fuchs 和 Jeroen van de Graaf。量子力学态的密码学可区分性度量。IEEE 信息论汇刊，45(4):1216–1227，1999年。31[GKK+08] Dmitry Gavinsky、Julia Kempe、Iordanis Kerenidis、Ran Raz 和 Ronald de Wolf。单向量子通信复杂度的指数分离及其在密码学中的应用。SIAM 计算杂志，38(5):1695–1708，2008年。4[GKLR21] Sumegha Garg、Pravesh K. Kothari、Pengda Liu 和 Ran Raz。带噪声奇偶学习的内存-样本下界。载于 APPROX-RANDOM，第207卷 LIPIcs，第60:1–60:19页。Schloss Dagstuhl - Leibniz-Zentrum für Informatik，2021年。2[GKR20] Sumegha Garg、Pravesh K. Kothari 和 Ran Raz。区分分布的时间-空间权衡及其在 Goldreich 伪随机生成器安全性中的应用。载于 APPROX-RANDOM，第176卷 LIPIcs，第21:1–21:18页。Schloss Dagstuhl - Leibniz-Zentrum für Informatik，2020年。2[GRT18] Sumegha Garg、Ran Raz 和 Avishay Tal。基于提取器的学习时间-空间下界。载于第50届 ACM SIGACT 计算理论研讨会论文集（STOC 2018），第990–1002页，美国纽约，2018年。计算机协会。2, 3, 4, 6[GZ19] Jiaxin Guan 和 Mark Zhandary。有界存储模型中的简单方案。载于国际密码技术理论与应用会议，第500–524页。施普林格，2019年。4[HHJ+16] Jeongwan Haah、Aram W Harrow、Zhengfeng Ji、Xiaodi Wu 和 Nengkun Yu。量子态的样本最优层析。载于第48届 ACM 计算理论研讨会论文集，第913–925页，2016年。2[HKP20] Hsin-Yuan Huang、Richard Kueng 和 John Preskill。通过极少量测量预测量子系统的多种性质。《自然·物理》，16(10):1050–1057，2020年。2[HN06] Danny Harnik 和 Moni Naor。混合有界存储模型中的永久安全性。载于国际自动机、语言与编程研讨会，第192–203页。施普林格，2006年。4[KK12] Roy Kasher 和 Julia Kempe。抵御量子对手的双源提取器。《计算理论》，8(1):461–486，2012年。33[KRT17] Gillat Kol、Ran Raz 和 Avishay Tal。稀疏奇偶学习的时间-空间困难性。载于第49届 ACM SIGACT 计算理论研讨会论文集，第1067–1080页，2017年。2[LM00] Beatrice Laurent 和 Pascal Massart。通过模型选择自适应估计二次泛函。《统计年刊》，28(5)，2000年10月。10

[Lu02] Chi-Jen Lu。基于在线强提取器、抵御空间受限对手的超加密。载于国际密码学年会，第257–271页。施普林格，2002年。4[LV21] Jiahui Liu 和 Satyanarayana Vusirikala。有界存储模型中的安全多方计算。载于 IMA国际密码学与编码会议，第289–325页。施普林格，2021年。4[Mau92] Ueli M Maurer。条件完美保密与可证明安全的随机密码。《密码学杂志》，5(1):53–66，1992年。4[MM18] Dana Moshkovitz 和 Michal Moshkovitz。空间受限学习的熵采样器与强一般下界。载于 ITCS，第94卷 LIPIcs，第28:1–28:20页。Schloss Dagstuhl - Leibniz-Zentrum für Informatik，2018年。2[MSTS04] Tal Moran、Ronen Shaltiel 和 Amnon Ta-Shma。有界存储模型中的非交互式时间戳。载于国际密码学年会，第460–476页。施普林格，2004年。4[OP04] Masanori Ohya 和 D´enes Petz。量子熵及其应用。施普林格科学与商业媒体，2004年。34[PMLA13] Stefano Pironio、Ll Masanes、Anthony Leverrier 和 Antonio Ac´ın。有界量子存储模型中设备无关量子密钥分发的安全性。《物理评论 X》，3(3):031007，2013年。4[Raz17] Ran Raz。一大类学习问题的时间-空间下界。载于2017年IEEE第58届计算机科学基础年会（FOCS），第732–742页，2017年。2, 3, 4, 6[Raz18] Ran Raz。快速学习需要良好记忆：奇偶学习的时间-空间下界。《ACM期刊》，66(1)，2018年12月。2, 4[RS08] Alvin C Rencher 和 G Bruce Schaalje。统计学中的线性模型。约翰·威立父子公司，2008年。10[Sch07] Christian Schaffner。有界量子存储模型中的密码学。arXiv 预印本 arXiv:0709.0289，2007年。4[Sha14] Ohad Shamir。统计学习与估计的在线和分布式算法的基本极限。神经信息处理系统进展，27，2014年。2[SSV19] Vatsal Sharan、Aaron Sidford 和 Gregory Valiant。小误差线性回归的内存-样本权衡。载于第51届 ACM SIGACT 计算理论研讨会论文集，第890–901页，2019年。2[SVW16] Jacob Steinhardt、Gregory Valiant 和 Stefan Wager。内存、通信与统计查询。载于学习理论会议，第1490–1516页。PMLR，2016年。2[Uhl76] Armin Uhlmann。*-代数状态空间中的“跃迁概率”。《数学物理报告》，9(2):273–279，1976年。32

[Wri16] John Wright。 如何学习量子态。博士论文，卡内基梅隆大学，2016年。 2

[WW08] Stephanie Wehner 和 J¨urg Wullschleger。在有界量子存储模型中的可组合安全性。载于国际自动机、语言与编程研讨会，第604–615页。施普林格，2008年。 4

定理2的参数上界

证明。对于公式（4），由于 $r \le r'/4$ 根据 r 和 $q \le r - 7$ 的定义，我们有 :

$$q + r + 1 - r' \le q + r + 1 - 4r \le -2r - 6 \le -2r.$$

对于公式（5），我们利用 $\ell = \frac{1}{5}$（$\ell' - 13r - 2$），因此

$$\begin{aligned} 2\ell + 9r - n =& \frac{2}{5}(\ell' - 13r - 2) + 9r - n \\ =& \frac{2}{5}\ell' + \frac{19}{5}r - \frac{4}{5} - n. \end{aligned}$$

求解不等式 $\frac{2}{5}\ell' + \frac{19}{5}r - \frac{4}{5} - n \le -r$（结合 $\ell' \le n$ 这一事实）可得 $r \le \frac{1}{8}\ell' + \frac{1}{6}$，这是因为根据定义，$r \le \frac{1}{26}\ell' + \frac{1}{6}$。

对于方程（6），当 $r \le \frac{1}{26}\ell' + \frac{1}{6}$ 时， $r \le \frac{1}{2}$（$k' - 1$）且 $k = k' - 1$,

$$\begin{aligned} k - r \ge k - (k' - 1)/2 = (k' - 1)/2, \\ \ell = \frac{1}{5}(\ell' - 13r - 2) \ge \frac{1}{5}\left(\ell' - \frac{\ell'}{2} - \frac{13}{6} - 2\right) > \frac{1}{10}\ell' - 1. \end{aligned}$$

当 ℓ' 足够大时，我们有 $\frac{1}{10}\ell' - 1 > \frac{1}{11}\ell' + 5$。因此，利用 $m \le (k' - 1)\ell'/44$ 这一事实:

$$\begin{aligned} (k - r)\ell &\ge \frac{1}{2}(k' - 1)\left(\frac{1}{11}\ell' + 5\right) \\ &= \frac{1}{22}(k' - 1)\ell' + \frac{5}{2}(k' - 1) \\ &\ge 2m + 5r \\ &\ge 2m + 4r + 1. \end{aligned}$$

□

命题 4.2 的证明

我们首先陈述 Fuchs-van de Graaf 不等式 [FvdG99] 关于保真度的一个变体，该保真度定义于两个部分密度算子 ρ 和 σ 上，如下所示

$$F(\rho, \sigma) = \mathrm{Tr}\left[\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right].$$

Le引理 B.1. 设 ρ、σ 为两个半正定算子。假设 $\mathrm{Tr}[\rho] \ge \mathrm{Tr}[\sigma]$。 n

$$\frac{1}{2}\|\rho - \sigma\|_{\mathrm{Tr}} \le \sqrt{\frac{1}{4}(\mathrm{Tr}[\rho] + \mathrm{Tr}[\sigma])^2 - F(\rho, \sigma)} \le \sqrt{\mathrm{Tr}[\rho]^2 - F(\rho, \sigma)}.$$

注意，当 $\mathrm{Tr}[\rho] = \mathrm{Tr}[\sigma] = 1$ 时，上述不等式就是原始的 Fuchs-van de Graaf 不等式。

证明。设 u 和 v 是 ρ 和 σ的纯化，即 $u, v \in \mathcal{H} \otimes \mathcal{H}_A$ 满足 $\rho = \mathrm{Tr}_A[uu^\dagger]$ 且$\sigma = \mathrm{Tr}_A[vv^\dagger]$，其中 $\mathcal{H}$ 是 ρ 和 σ的环境空间，而 $\mathcal{H}_A$ 是某个有限维希尔伯特空间。

令 U 为将 $uu^\dagger - vv^\dagger$对角化的酉矩阵，即存在一个对角矩阵 $\Lambda \in \mathbb{C}^{d\times d}$使得 $uu^\dagger - vv^\dagger = U\Lambda U^\dagger$。令 p和 $q \in \mathbb{R}^d_{\geq 0}$分别为 $U^\dagger uu^\dagger U$和 $U^\dagger vv^\dagger U$的对角线。我们有

$$\begin{aligned}
u^\dagger u = \mathrm{Tr}[U^\dagger uu^\dagger U] &= \sum_{x\in[d]} p(x), \\
v^\dagger v = \mathrm{Tr}[U^\dagger vv^\dagger U] &= \sum_{x\in[d]} q(x), \\
\left\|uu^\dagger - vv^\dagger\right\|_{\mathrm{Tr}} = \left\|\Lambda\right\|_{\mathrm{Tr}} &= \sum_{x\in[d]} |p(x) - q(x)|, \\
|\langle u, v\rangle| = |\langle U^\dagger u, U^\dagger v\rangle| &\leq \sum_{x\in[d]} \sqrt{p(x)q(x)}.
\end{aligned}$$

因此，由柯西-施瓦茨不等式，

$$\begin{aligned}
\left\|uu^\dagger - vv^\dagger\right\|_{\mathrm{Tr}}^2 &= \left(\sum_{x\in[d]} |p(x) - q(x)|\right)^2 \\
&= \left(\sum_{x\in[d]} \left|\sqrt{p(x)} - \sqrt{q(x)}\right| \cdot \left|\sqrt{p(x)} + \sqrt{q(x)}\right|\right)^2 \\
&\leq \left(\sum_{x\in[d]} \left|\sqrt{p(x)} - \sqrt{q(x)}\right|^2\right)\left(\sum_{x\in[d]} \left|\sqrt{p(x)} + \sqrt{q(x)}\right|^2\right) \\
&= \left(\sum_{x\in[d]} p(x) + \sum_{x\in[d]} q(x)\right)^2 - 4\left(\sum_{x\in[d]} \sqrt{p(x)q(x)}\right)^2 \\
&\leq \left(u^\dagger u + v^\dagger v\right)^2 - 4|\langle u, v\rangle|^2.
\end{aligned}$$

注意到$\left\|\rho - \sigma\right\|_{\mathrm{Tr}} \leq \left\|uu^\dagger - vv^\dagger\right\|_{\mathrm{Tr}}$，$\mathrm{Tr}[\rho] = u^\dagger u$且 $\mathrm{Tr}[\sigma] = v^\dagger v$。由乌尔曼定理[Uhl76],，我们还可以选取 u和 v使得 $F(\rho, \sigma) = |\langle u, v\rangle|^2$。将它们代入上述不等式，即可完成证明。

□

现在我们可以证明命题 4.2

命题的证明*4.2*。根据 Fuchs-van de Graaf 不等式，只需证明保真度满足以下界：

$$F(\rho, \Pi\rho\Pi) \geq \mathrm{Tr}[\Pi\rho]^2.$$

设 u 是 ρ_{XV}的一个纯化，也就是说，对于某个希尔伯特空间 $\mathcal{H}_A$，有 $\rho = \mathrm{Tr}_A[uu^\dagger]$。那么，$(\Pi \otimes \mathbb{I}_A)u$是 $\Pi\rho\Pi$的一个纯化。根据乌尔曼定理，我们有

$$F(\rho, \Pi\rho\Pi) \geq \left|u^\dagger(\Pi \otimes \mathbb{I}_A)u\right|^2 = \mathrm{Tr}\left[(\Pi \otimes \mathbb{I}_A)uu^\dagger\right]^2 = \mathrm{Tr}\left[\Pi \cdot \mathrm{Tr}_A[uu^\dagger]\right]^2 = \mathrm{Tr}[\Pi\rho]^2.$$

□

C 线性量子存储下界

在本附录中，我们证明定理 3 ，该定理给出了一种更简单的线性量子存储下界（不含经典内存）证明。尽管在大多数情况下，定理 3 在定性上弱于我们的主要结果，因为它只给出了针对使用线性大小量子存储器、但不含（可能为二次大小）经典内存的程序的下界，但定理 3 在技术上与主要结果不可比，因为它是以量子提取器的形式表述的，而且量子存储大小的界取决于提取器的另一个参数。此外，定理 3 的证明比我们的主定理证明简单得多。

我们首先定义所需的量子提取器性质，这是 [KK12] 中所考虑性质的一个简化版本。给定一个矩阵 $M\colon \mathcal{A}\times\mathcal{X}\to\{-1,1\}$，考虑两个独立源 A 和 X，它们分别在 $\mathcal{A}$ 和 $\mathcal{X}$ 上均匀分布。假设存在某个量子寄存器 V，其量子态取决于 A 和 X，并且它们共同构成一个经典-量子系统

$$\rho_{AXV}=\bigoplus_{a\in\mathcal{A},x\in\mathcal{X}}\rho_{V|a,x},$$

其中，$\rho_{V|a,\ x}$ 是 V在 $A=a$ 和 $X=x$时的量子态。对于定义在 $A\times X$上的任意函数 f，如果对于任意 a、$a'\in\mathcal{A}$以及 x、$x'\in\mathcal{X}$，只要 $f(a,x)=f\ a'(x')$，就有$\rho_{V|a,\ x}=\rho_{V|a',\ x'}$，则称$V$仅取决于 $f(A,X)$。特别地， V仅取决于 A等价于 V与 X独立，或 $\rho_{XV}=\rho_X\otimes\rho_V$。

如果对于上述每个经典-量子系统ρ_{AXV}，其中量子子系统 V 包含 q个量子比特且仅依赖于 A，都满足以下条件，则称 M 是一个 X-强（q,r）量子提取器：

$$\left\|\rho_{M(A,X)XV}-U\otimes\rho_X\otimes\rho_V\right\|_{\mathrm{Tr}}\le 2^{-r}.$$

这里，$U=\begin{pmatrix}1/2 & 0\\ 0 & 1/2\end{pmatrix}$ 是作用于一个比特的均匀算子，而$\rho_{M(A,X)XV}$是通过添加一个存储 $M(A,X)$值的新经典寄存器，并对 A取偏迹而构造的经典-量子系统。换句话说，

$$\rho_{M(A,X)XV}=\bigoplus_{b\in\{-1,1\},x\in\mathcal{X}}\ \sum_{\substack{a\in\mathcal{A}\\ M(a,x)=b}}\rho_{V|a,x}.$$

注意，如果我们选择 V 为平凡情形，上述不等式便立即推出$|\mathbb{E}[M(A,X)]|\le 2^{-r}$。

例如， [KK12] 的结果表明，作用于 n 比特的内积函数，其中$\mathcal{A}=\mathcal{X}=\mathbb{F}_2^n$ 并且

$$M(a,x)=(-1)^{a\cdot x},$$

对于每个 $2\le k\le n$，都是一个 X-强（$k,n-k$）量子提取器。

本节证明以下定理：

定理 3. 设 $\mathcal{X}$、$\mathcal{A}$ 是两个有限集合，且 $n=\log_2|\mathcal{X}|$。设 $M\colon \mathcal{A}\times\mathcal{X}\to\{-1,1\}$ 是一个 X-强 (q, r)-量子提取器。设 ρ 是对应于 M 的学习问题的一个分支程序，由经典-量子系统 $\rho_{XV}^{(t)}$ 描述，具有 $q/2$ 量子比特的量子存储器V，长度为 T，且不含经典内存。那么， ρ 的成功概率至多为

$$2^{-n}+8T\sqrt{n+q}\cdot 2^{-r/4}.$$

为了证明该定理，我们首先需要定义以下依赖性度量：

定义：设 ρ_{XV} 是定义在经典 X 和量子 V 上的经典-量子系统。ρ_{XV} 中 V 对 X 的依赖性定义为

$$\xi^{\rho}(X;V)=\min_{\tau_V}\left\|\rho_{XV}-\rho_X\otimes\tau_V\right\|_{\mathrm{Tr}}$$

其中，τ_V 取遍 V 上的所有密度算子。请注意，在此定义中，取 $\tau_V=\rho_V$ 几乎是最优的，因为我们有

$$\left\|\rho_{XV}-\rho_X\otimes\rho_V\right\|_{\mathrm{Tr}}\leq\left\|\rho_{XV}-\rho_X\otimes\tau_V\right\|_{\mathrm{Tr}}+\left\|\rho_V-\tau_V\right\|_{\mathrm{Tr}}\leq 2\left\|\rho_{XV}-\rho_X\otimes\tau_V\right\|_{\mathrm{Tr}}. \tag{16}$$

我们还考虑了 ρ_{XV} 中 X 与 V 之间的量子互信息，其定义为

$$\mathbf{I}_{\rho}(X;V)=\mathbf{S}(\rho_X)+\mathbf{S}(\rho_V)-\mathbf{S}(\rho_{XV})=\mathbf{S}\left(\rho_{XV}\parallel\rho_X\otimes\rho_V\right),$$

其中，S(·)表示冯·诺伊曼熵，S(· ∥ ·)表示量子相对熵。当 V 由 q 个量子比特组成时，我们的依赖度量与量子互信息之间存在如下关系：

引理 C.1. $\frac{1}{2}\xi^{\rho}(X;V)^2\leq\mathrm{I}_{\rho}(X;V)\leq q\cdot\xi^{\rho}(X;V)+2\sqrt{\xi^{\rho}(X;V)}.$

证明。一方面，利用量子相对熵与迹距离之间的不等式（例如，参见[OP04，定理 1。15]），我们有

$$\mathbf{I}_{\rho}(X;V)=\mathbf{S}\left(\rho_{XV}\parallel\rho_X\otimes\rho_V\right)\geq\frac{1}{2}\left\|\rho_{XV}-\rho_X\otimes\rho_V\right\|_{\mathrm{Tr}}^2\geq\frac{1}{2}\xi^{\rho}(X;V)^2.$$

另一方面，Fannes-Audenaert 不等式 [Aud07] 告诉我们，对于任意 $x\in\mathcal{X}$，任意两个量子态 ρ 和 τ 在 V 上的冯·诺依曼熵之差有如下界

$$|\mathbf{S}(\rho)-\mathbf{S}(\tau)|\leq q\cdot\frac{1}{2}\left\|\rho-\tau\right\|_{\mathrm{Tr}}+h\left(\frac{1}{2}\left\|\rho-\tau\right\|_{\mathrm{Tr}}\right)$$

其中 $h(\epsilon)=-\epsilon\log_2\epsilon-(1-\epsilon)\log_2(1-\epsilon)$ 是二元熵函数。由于在给定 $X=x$ 的条件下，V 的量子态为 $\rho_{V|x}/\mathbf{Pr}[X=x]=2^n\rho_{V|x}$，因此我们有

$$\begin{aligned}\mathbf{I}_{\rho}(X;V)&=\mathop{\mathbf{E}}_{x\sim X}\left[\mathbf{S}(\rho_V)-\mathbf{S}(2^n\rho_{V|x})\right]\\&\leq\frac{1}{2}q\cdot\mathop{\mathbf{E}}_{x\sim X}\left\|\rho_V-2^n\rho_{V|x}\right\|_{\mathrm{Tr}}+\mathop{\mathbf{E}}_{x\sim X}h\left(\frac{1}{2}\left\|\rho_V-2^n\rho_{V|x}\right\|_{\mathrm{Tr}}\right)\\&\leq\frac{1}{2}q\cdot\left\|\rho_{XV}-\rho_X\otimes\rho_V\right\|_{\mathrm{Tr}}+h\left(\frac{1}{2}\left\|\rho_{XV}-\rho_X\otimes\rho_V\right\|_{\mathrm{Tr}}\right)\\&\leq\frac{1}{2}q\cdot\left\|\rho_{XV}-\rho_X\otimes\rho_V\right\|_{\mathrm{Tr}}+\sqrt{2\left\|\rho_{XV}-\rho_X\otimes\rho_V\right\|_{\mathrm{Tr}}},\end{aligned}$$

由于 h 是凹函数，且 $h(\epsilon)\leq 2\sqrt{\epsilon}$。现在令 τ_V 为 $\xi^{\rho}(X;V)$ 的定义中最优的密度算子。将其代入式（16），我们得到

$$\mathbf{I}_{\rho}(X;V)\leq q\cdot\xi^{\rho}(X;V)+2\sqrt{\xi^{\rho}(X;V)}.$$ □

引理 C.2. 对于每个经典-量子系统 ρ_{AXV}，若其 q 量子比特量子子系统 V 仅依赖于 A，则有

$$\mathbf{I}_{\rho}(X;M(A,X),V)\leq 2(n+q)\cdot 2^{-r/2}.$$

证明。由于 $\mathrm{I}_\rho(X;\ V)\ =\ 0$，只需界定 $\mathrm{I}_\rho(X;\ M(A,X)\ |\ V)\ \leq \mathrm{I}_\rho(M(A,X);\ X,V)$。为了界定后者，我们首先注意到，由于 M 是一个强 (q,r)-量子提取器，

$$\begin{aligned}\xi^\rho(M(A,X);X,V) &\leq \left\|\rho_{M(A,X)XV}-\rho_{M(A,X)}\otimes\rho_X\otimes\rho_V\right\|_{\mathrm{Tr}}\\ &\leq \left\|\rho_{M(A,X)XV}-U\otimes\rho_X\otimes\rho_V\right\|_{\mathrm{Tr}}+|\mathbf{E}[M(A,X)]|\\ &\leq 2\cdot 2^{-r}.\end{aligned}$$

由于 X 和 V 的总维数为 2^{n+q}，由引理C.1 可得

$$\begin{aligned}\mathbf{I}_\rho(X;M(A,X),V) &\leq \mathbf{I}_\rho(M(A,X);X,V)\\ &\leq (n+q)\cdot\xi^\rho(M(A,X);X,V)+2\sqrt{\xi^\rho(M(A,X);X,V)}\\ &\leq 5(n+q)\cdot 2^{-r/2}.\end{aligned}$$

□

引理 C.3. 对于每个经典-量子系统 ρ_{AXV}，其包含 $q/2$ 个量子比特的量子子系统 V 仅取决于 A 和 M（A，X），我们有

$$\xi^\rho(X;V)\leq 4\sqrt{n+q}\cdot 2^{-r/4}.$$

证明。令 $W=\rho_{a,\ 0}\otimes\rho_{a,\ 1}$，其中 $\rho_{a,\ b}$ 是当 $A=a$ 且 M（A，X）$=b$ 时 V 的密度矩阵。那么 W 是一个仅取决于 A 的 q 位量子系统。由于可以根据 M（A，X）和 W 判断 V，我们有

$$\xi^\rho(X;V)^2\leq 2\mathbf{I}_\rho(X;V)\leq 2\mathbf{I}_\rho(X;M(A,X)),W)\leq 10(n+q)\cdot 2^{-r/2}.$$

□

现在我们准备证明定理 3。令 $\Phi_{t,\ a,\ b}$ 表示在阶段 t 对 V 施加的量子信道，其样本为（a，b）；回忆系统 $\rho_{XV}^{(t)}$ 的演化可表示为

$$\rho_{XV}^{(t+1)}=\mathop{\mathbf{E}}_{a\sim A}\left[\sum_{x\in\mathcal{X}}|x\rangle\langle x|\otimes\Phi_{t,a,M(a,x)}\left(\rho_{V|x}^{(t)}\right)\right].$$

定理 3 的证明。我们将对 ξ_t的增量进行界定，它是 $\xi^{\rho^{(t)}}(X;\ V)$ 的简写。现在先考虑某个阶段 t，并令 τ 为使$\xi_t=\left\|\rho_{XV}^{(t)}-\rho_X^{(t)}\otimes\tau\right\|_{\mathrm{Tr}}$最小化的密度算子。注意，对于每个 $\rho_X^{(t)}=\rho_X=2^{-n}\mathbb{I}_X$，$t$。

由于 τ 是一个固定的量子态，我们可以制备 τ 并在 τ 上应用 $\Phi_{A,M(A,X)}$，从而得到一个新的量子寄存器 V'，它仅取决于 A 和 M（A,X）。注意，

$$\rho_{XV'}=\mathop{\mathbf{E}}_{a\sim A}\left[\sum_{x\in\mathcal{X}}|x\rangle\langle x|\otimes\Phi_{t,a,M(a,x)}(\tau)\right],$$

因此，类似于引理4.14 （演化不会增加迹距离），我们可以证明

$$\begin{aligned}\left\|\rho_{XV}^{(t+1)}-\rho_{XV'}\right\|_{\mathrm{Tr}} &\leq \mathop{\mathbf{E}}_{a\sim A}\sum_{x\in\mathcal{X}}\left\|\Phi_{t,a,M(a,x)}\left(\rho_{V|x}^{(t)}\right)-\Phi_{t,a,M(a,x)}(\tau)\right\|_{\mathrm{Tr}}\\ &\leq \sum_{x\in\mathcal{X}}\left\|\rho_{V|x}^{(t)}-\tau\right\|_{\mathrm{Tr}}\leq\left\|\rho_{XV}^{(t)}-\rho_X\otimes\tau\right\|_{\mathrm{Tr}}=\xi_t.\end{aligned}$$

因此我们有

$$\begin{aligned}\xi_{t+1} &\leq \left\|\rho_{XV}^{(t+1)} - \rho_X \otimes \rho_{V'}\right\|_{\mathrm{Tr}} \\ &\leq \left\|\rho_{XV}^{(t+1)} - \rho_{XV'}\right\|_{\mathrm{Tr}} + \left\|\rho_{XV'} - \rho_X \otimes \rho_{V'}\right\|_{\mathrm{Tr}} \\ &\leq \xi_t + 2\xi^{\rho}(X;V') \\ &\leq \xi_t + 8\sqrt{n+q}\cdot 2^{-r/4}.\end{aligned}$$

由于 $\xi_0 = 0$，我们得出

$$\xi_T \leq 8T\sqrt{n+q}\cdot 2^{-r/4}.$$

这个值界定了 ρ 的成功概率与一个内存独立于 X 的量子分支程序的成功概率之差。后者显然至多为 2^{-n}，证毕。

□