可分量子过程中的内禀信息与测量信息

David Gier* 和 James P. Crutchfield†加利福尼亚大学戴维斯分校
复杂性科学中心与物理与天文学系，*One Shields Avenue*，戴维斯，*CA 95616*（日期：2023年3月2日）

平稳量子信息源会发出相关量子位序列，即具有结构的量子随机过程。若观察者对量子位序列进行相同的测量，测量结果便构成一个经典随机过程的实现。我们提出适用于可分离量子位序列的量子信息论性质，这些性质可作为后续测量过程的经典信息性质的界限。对于由隐马尔可夫动力学驱动的信息源，我们阐述观察者如何利用特定的正算子值测度或自适应测量协议，暂时或永久地与信源的内部状态同步。我们提出一种方法，通过层析重构，使用独立同分布模型、马尔可夫模型或具有更大记忆容量的模型来近似信息源。我们依据量子信息性质，确定了可分过程的广泛类别，并考察了同步这些过程并准确重构它们所需的测量复杂度。

arXiv:2303.00162v1 [quant-ph] 1 Mar 2023

目录

I. 引言 2
A. 量子与经典随机性 2 B. 具有记忆的源 2
2

II. 随机过程 3
A. 经典过程 3 B. 表示 4 C. 量子过程 5 1. 无记忆 5 2. 有记忆 5 D. 量子过程的表示 7 E. 测量过程 7 F. 自适应测量协议 8 G. 讨论 9

III. 量子过程中的信息 9
A. 冯·诺依曼熵 10 B. 量子块熵 11 C. 冯·诺依曼熵率 12 D. 量子冗余 13 E. 量子熵增益 14 F. 量子可预测性增益 15 G. 总量子可预测性 15 H. 量子超额熵 16 I. 量子瞬态信息 18 J. 量子马尔可夫阶 19

IV. 量子位过程 19
A. 独立同分布过程 19 B. 经典过程的量子表征 20 C. 周期过程 20 D. 量子黄金均值过程 21 E. 三符号量子黄金均值 21 F. 单一分支与非单一分支量子比特源 22 G. 单一分支量子三能级系统源 23 H. 讨论 24

V. 与量子源同步 24
A. 知识状态 24 B. 平均态不确定性与同步信息 25 C. 与经典过程的量子表征同步 26 D. 与周期过程同步 26 E. 使用投影值测度同步 27 F. 通过自适应测量维持同步 28 G. 与三能级量子源同步 29 H. 讨论 31

VI. 量子过程系统辨识 32
A. 经典系统辨识 32 B. 量子多能级系统层析 32 C. 量子多能级系统过程层析 33 D. 独立同分布的成本 33 E. 信息性质的有限长度估计 34 F. 使用已知量子字母表的层析 34 1. $\ell = 1$ 35 2. $\ell = 2$ 36 3. $\ell \geq 3$ 37

* dgier@加州大学戴维斯分校.edu
† chaos@ucdavis.edu

G. 源重构 37 1.
$\ell = 1$ 372.
$\ell = 2$ 383.
$\ell \geq 3$ 38H. 讨论
38

VII. 结论
39 致谢 39 A. 经典过程中的信息 40 1. 香农熵
40 2. 块熵 40 3. 香农熵率 40 4. 冗余 41 5. 块熵导数与
积分 41 6. 熵增益 41 7. 可预测性增益 42 8. 总可预测
性 42 9. 超额熵 42 10. 瞬态信息 43 11. 马尔可夫阶数
43 B. 制备与测量的量子信道 43 参考文献 44

I. 引言

确定量子系统的状态，需要应对多种不确定性来源，其中有些并不存在于经典物理学中。特别是，测量存在不可消除的限制；自海森堡在 1927 [1]提出位置-动量不确定性原理以来，这一直是量子物理学的标志。对于一般的纯量子态，也存在类似的不相容测量 [2]。

对于一个二能级量子系统，或量子比特，不可能同时测量自旋在 x、y 和 z 方向上的取值。（用数学语言表述，泡利矩阵 σ_x、σ_y 和 σ_z 不对易。）此外，在每个基上只测量一次是不够的。必须在每个基上测量许多副本，才能确定结果分布。因此，要通过量子态层析确定未知的量子比特状态，就需要用一组互不相容基 [3] 或一个信息完备的正算子值测度 （POVM） [4]，对大量相同副本组成的系综进行测量。

这些不确定性来源在量子物理学中广为人知。将其与观察者接收到一串相关量子比特时的情形作比较：逐个测量这些量子比特，他们会看到什么？进而，他们能推断出生成这些量子比特串所需的资源吗？下文将通过剖析测量量子过程中表观随机性和相关性的来源，来回答这些问题。

A. 量子和经典随机性

同样是在1927年，冯·诺依曼用统计系综表述量子力学，并量化了这些混合量子态的熵。在此过程中，他将吉布斯关于统计系综和经典热力学熵的研究拓展到了量子领域 [5]。混合量子态 ρ 具有熵$S(\rho) = -tr(\rho \log \rho)$，现在称为冯·诺依曼熵。（$tr(\cdot)$是迹算子。） $S(\rho) = 0$ 当且仅当 ρ 是纯（非混合）量子态时成立。一方面，冯·诺依曼熵是理解量子系统的关键，尤其是那些具有纠缠子系统并表现出非经典关联的系统。另一方面，不确定性 $S(\rho)$量化了统计系综的一个普遍特征。它并不对应于任何特定的量子力学效应。

这两种不确定性——分别源于系综和量子不确定性——在量子信息论框架内得以结合；量子信息论将经典信息论推广到了量子可观测量 [6]。一个显著例子是无噪声编码。香农量化了无噪声经典独立同分布（i.i.d.）信息源产生的信息量——这种信息源在每个时间步[7]都发出一个从同一分布中抽取的状态。舒马赫量子无噪声编码定理将这一理论推广到了量子信息源。这为冯·诺依曼熵赋予了新的物理解释：对于发出状态 ρ的独立同分布量子源，$S(\rho)$是可靠压缩方案 [8]所需的量子比特数。

B. 具有记忆的源

非独立同分布平稳信息源会引入额外形式的不确定性。例如，源可能具有内部记忆，从而导致连续量子比特之间以及测量结果之间产生相关性。这种相关性可能纯粹是经典的，也可能具有独特的量子性质。正如我们将要展示的，若实验者错误地假设此类源

是独立同分布的，随后应用现有的层析方法，就无法检测到这些相关性，因此会高估源的随机性，并低估其可压缩性。

经典有记忆信息源可在计算力学框架内描述，其中稳态动力系统作为信息源，具有自身的内部状态和动态 [9]。对动力系统进行序贯有限精度测量会形成离散时间随机过程。由此产生的过程统计特征可用于构建信息源模型，并计算其渐近熵率、内部内存需求及其他与物理相关的性质 [10, 11]。重要的是，能够与过程内部状态同步的观察者，可以降低与经典信息源的序贯测量相关的不确定性，有时甚至能大幅降低 [12, 13]。

对开放量子系统进行序贯量子位探针测量，带来了类似但更为一般的挑战，因为观察者通过测量从单个量子位中获取的信息量有限。近期研究结果表明，对量子比特施加特定测量仪器会使测量序列呈现复杂行为 [14]。本文通过研究量子态本身的性质，以及特定的测量结果序列，对这些结果加以拓展。

以下介绍可分离——即非纠缠——量子位序列的一些新颖量子信息理论性质。我们以此前关于平稳量子信息源的熵率、压缩极限和最优编码策略的研究成果 [15–17]，为基础，同时借鉴了针对独立同分布假设的特定实验动机偏离所开展的研究 [18]。我们采用的方法与近期关于量子随机过程的研究有所不同，但又相辅相成；在那些研究中，观察者直接测量量子系统。由于量子系统会与不可访问环境相互作用，并由此在连续测量结果中引入记忆效应 [19–23]，这类过程分析起来较为复杂。

第二节介绍经典过程、可分离量子比特过程，以及通过经典-量子通道和测量通道在二者之间进行转换的方法。接着，第三节结合附录 A，分别定义与量子过程和经典过程相关的熵。借鉴参考文献 [11] 的熵层级结构，我们采用离散时间导数和积分，得到一系列不同的量化指标，用以衡量量子过程随机性和相关性。我们证明，对于投影测量或信息完备测量，测量结果序列形成的经典过程，其信息属性受到被测量子过程信息属性的约束。

接着，第四节将综述结构日益复杂的可分离量子比特和量子三能级系统过程的实例。第五节讨论观察者如何通过连续测量与具有记忆的源同步——即确定其内部状态。第六节利用由此得到的各种可能过程行为目录，回答尝试对量子过程进行层析的观察者所面临的实际问题。最后，第七节总结经验并提出未来的研究方向和应用，尤其是将研究成果扩展到可通过实验实现的任意纠缠量子位态生成 [24, 25] ，以及利用相关性作为资源来执行热力量子信息处理 [26, 27]。

II. 随机过程

我们将信息源的输出视为离散时间的平稳随机过程。如果源的输出是经典随机变量——X_t 每个时间步 t——我们便可以直接应用计算力学 [11]的方法。我们的目标是扩展这些方法，以描述可分离的量子位序列，其中每个量子位都由 d维希尔伯特空间中的一个纯态表示：$|\psi_t\rangle \in \mathcal{H}^d$ 在每个时间步 t。对于这样的量子位序列，可以进行重复的相同测量，使测量结果构成一个经典随机过程。由于可以选择在许多不同的基中测量量子位状态，经典测量过程的性质由相关量子位的状态和测量选择共同决定。因此，量子过程与经典测量过程之间是一对多的关系。图 1 展示了这一设置。

A. 经典过程

经典随机过程 由定义在随机变量链上的概率测度 $\Pr(\overleftrightarrow{X})$ 给出：

$$\overleftrightarrow{X} \equiv \ldots X_{-2}X_{-1}X_0X_1X_2\ldots \ ,$$

其中每个 X_t 取值均来自有限字母表 $\mathcal{X}$。一段由 ℓ 个连续随机变量组成的序列记作 $X_{0:\ell} = X_0X_1\cdots X_{\ell-1}$。索引左端包含、右端不包含。一个特定的双无限过程实现 记作 $\overleftrightarrow{x} \equiv \ldots x_{-2}x_{-1}x_0x_1x_2\ldots$ ，其中事件 x_t 取值于离散集合 $\mathcal{X}$。实现

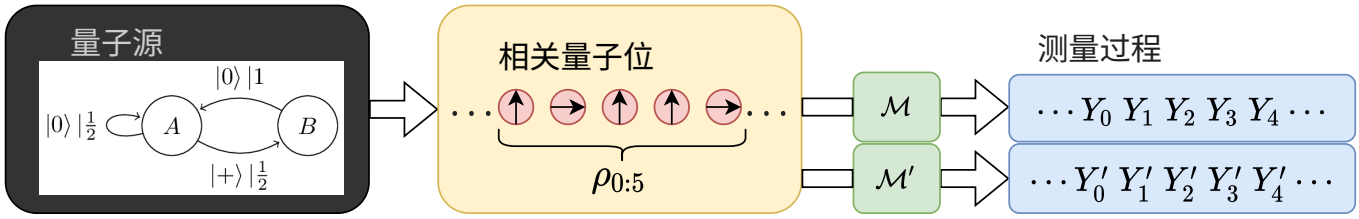


图 1. 平稳量子信息源会发出量子位，这些量子位由于源的内部记忆而彼此相关。实验者以不同方式测量这些量子位（$\mathcal{M}$ 或 $\mathcal{M}'$），从而得到一族经典随机过程。

长度为 ℓ 的一段序列的实现称为词，记作$x_{0:\ell}$。所有长度为 ℓ 的词构成的集合为 $\mathcal{X}^\ell$。

我们考虑平稳过程，即词概率 $\Pr(X_{0:\ell})$ 不随时间变化：

$$\Pr(X_{0:\ell}) = \Pr(X_{t:t+\ell}) \ ,$$

对所有 $t \in \mathbb{Z}$ 和 $\ell \in \mathbb{Z}^+$ 都成立。平稳过程的统计特征完全由长度为 ℓ 的词分布 $\Pr(X_{0:\ell})$ 描述。长度为 ℓ 的区块至多有 $|\mathcal{X}|^\ell$ 种可能的实现（词）。

一个重要的过程子类是独立同分布（i.i.d.）过程。独立同分布过程的联合块概率具有如下形式：

$$\Pr(X_{0:\ell}) \equiv \Pr(X_0)\Pr(X_1)\cdots\Pr(X_{\ell-1}) \ , \qquad (1)$$

对所有 $\ell \in \mathbb{Z}^+$ 均成立。块概率的这种因式分解意味着任意随机变量之间都不存在统计相关性。由于平稳性，对于独立同分布过程，对所有 t，$\Pr(X_t) = \Pr(X_0$。

另一个常见的研究对象子类是马尔可夫过程，其中每个 X_t 的分布仅取决于紧邻其前的随机变量 X_{t-1}。对于马尔可夫过程，有限长度块的联合概率可分解为：

$$\Pr(X_{0:\ell}) \equiv \Pr(X_0)\Pr(X_1|X_0)\cdots\Pr(X_{\ell-1}|X_{\ell-2}) \ , \quad (2)$$

其中 $\Pr(X|Y)$ 是在随机变量 Y 条件下随机变量 X 的概率分布。

最后，还有一个规模明显更大的子类，即隐马尔可夫过程，其内部马尔可夫动力学无法被直接观测。尽管联合概率不会像公式（2）中那样进行因子分解，但内部马尔可夫动力学会限制过程统计特征，具体如下文所述。

B. 表示

过程的一个表示是一种模型，由一组内部状态以及这些状态之间的转移动力学组成，二者共同准确地重现过程的统计特征。给定的过程可能有许多种表示。我们关注的是那些通过状态转移图（有向图）描绘的表示，它们以自然的方式生成平稳的离散时间随机过程。

马尔可夫链 是由二元组（$\mathcal{X}$ T）定义的过程表示：

- 由 m 个符号 $x \in \mathcal{X}$ 组成的有限字母表$\mathcal{X}$，以及
- 一个 $m \times m$ 转移矩阵 T。也就是说，如果源发射符号 x_i，其概率为 T_{ij} =$\Pr(x_j|x_i)$，那么下一个发射的符号是 x_j。

马尔可夫链的平稳分布记为 π，它是$\mathcal{S}$ 上因果状态的分布，并满足 $\pi = \pi T$。对于马尔可夫链，内部状态的集合恰好就是发射符号的集合，因为下一个符号的概率分布完全由前一个符号决定。我们将每个状态表示为图中的一个节点，并将每次转移表示为节点之间的一条有向边，其上标注相应的概率。

马尔可夫链足以表示马尔可夫过程，但通过允许存在无法直接观测的内部状态，我们可以描述更一般的隐马尔可夫过程。这些过程由隐马尔可夫链（HMC）生成，其定义为三元组（$\mathcal{S}$，$\mathcal{X}$，$\mathcal{T}$）：

- 有限个$\mathcal{S}$ $=\{\sigma_1, \ldots, \sigma_n\}$ 内部状态，
- 一个有限字母表$\mathcal{X}$，由 m 符号 $x \in \mathcal{X}$组成，并且
- 一组$\mathcal{T}$ $=\{T^x:\ x \in \mathcal{X}\}$ 个 m $n \times n$符号标记的转移矩阵。也就是说，如果源处于状态 σ_i，则它以概率 T^x_{ij} =$\Pr(x,\ \sigma_j|\sigma_i)$发射符号 x，同时转移到状态σ_j。

我们将状态之间的每一种可能转移表示为连接相应节点的一条边，并在边上标注发射符号和转移概率。隐马尔可夫链的平稳分布 π 唯一满足 $\pi = \pi \sum_x T^x$。

任何能够精确再现某一过程统计特征的隐马尔可夫链，都是一种生成式隐马尔可夫链。这一区分十分重要，因为其中只有一部分还属于限制更严格的预测式隐马尔可夫链。如果其在时间 $t+1$ 的状态完全由时间 t 的状态和发射符号决定，那么该隐马尔可夫链就是预测式的。这一性质称为单一性。

此时我们必须强调一个过程与该过程的特定表示之间的区别。在将过程和模型界定为“经典”或“量子”时，这一区分至关重要。离散时间经典随机过程之所以是经典的，是因为它由一连串经典随机变量组成。马尔可夫链和隐马尔可夫链是经典模型，因为它们的内部状态及其动力学都是经典的。相反，也可以用一组量子状态来构造经典过程的一种表示，模型通过某种量子动力学在这些状态之间进行转移。观察者可以通过对系统本身，或对每个时间步与系统相互作用的辅助量子比特进行序贯测量，恢复经典过程的统计特征。利用量子资源模拟经典随机过程，是量子计算力学的目标。在该领域中，一类量子模型（量子模拟器）在内存需求方面优于经证明达到最小的经典预测模型（ϵ-机器） [28–32]。同样，量子过程的不同表示可能具有经典或量子两种底层动力学。下面我们转向量子过程及其表示。

C. 量子过程

离散时间经典随机过程由每个时间步对应的一个经典随机变量组成。同样，离散时间量子随机过程由每个时间步对应的一个量子态 $|\psi_t\rangle \in \mathcal{H}^d$ 组成。我们首先描述独立同分布的量子信息源，然后将其推广到具有记忆的源。

1. 无记忆

一个离散时间量子信息源在每个时间步发射一个 d 能级量子系统或量子位。一个源所发射的无限量子比特序列的统计混合构成一个量子过程。与经典情形一样，不同类别的量子过程由其时间相关性加以区分。然而，对于量子源，我们必须使用量子信息论来同时描述经典相关性和量子相关性。

首先，考虑独立同分布（无记忆）量子信息源的输出。设 $\mathcal{H}$ 是一个 d 维希尔伯特空间，其纯态为 $|\psi_x\rangle \in \mathcal{H}$。一个 d 级独立同分布量子信息源由一组纯量子比特态 $\mathcal{Q}$ 以及定义在这些状态上的概率分布组成，并满足对于所有 $|\psi_x\rangle \in \mathcal{Q}$，$\Pr(|\psi_x\rangle) > 0$。我们将 $\mathcal{Q}$ 称为量子字母表，并且只考虑包含有限个纯态的量子字母表。

在每个离散时间步 t，源以概率 $\Pr(|\psi_x\rangle)$ 发出状态 $|\psi_x\rangle$。所得系综由如下 $d \times d$ 密度矩阵描述：

$$\rho_{iid} = \sum_{|\psi_x\rangle \in \mathcal{Q}} \Pr(|\psi_x\rangle)\, |\psi_x\rangle \langle\psi_x| \ . \tag{3}$$

ρ_{iid} 的这种特定纯态分解并非唯一。此外，观察者无法通过观测确定$\mathcal{Q}$——除非$\mathcal{Q}$ 仅包含一种状态——因为许多纯态系综都对应同一个密度矩阵。

如果一个独立同分布源在每个时间步都发出 ρ_{iid}，那么该源生成的量子过程就是无限张量积态：

$$|\Psi_{iid}\rangle = \cdots \otimes \rho_{iid} \otimes \rho_{iid} \otimes \rho_{iid} \otimes \cdots \ . \tag{4}$$

2. 有记忆

我们无法用$\mathcal{Q}$ 上的单一概率分布来描述非独立同分布源，而必须引入一个定义在状态序列上的概率分布，这些状态取自$\mathcal{Q}$。为此，我们将$\mathcal{Q}$ 中的每个元素与底层经典随机过程 $\overleftrightarrow{X}$ 的符号字母表$\mathcal{X}$ 中的一个元素相关联。无限量子位序列随后继承自 $\overleftrightarrow{X}$ 的概率。这一构造产生的量子位序列是可分离的，即不纠缠。

我们通过无记忆经典-量子通道$\mathcal{E}$ 来表示符号与纯量子态之间的关系：$\mathcal{X} \to \mathcal{H}$，其中 $x \to |\psi_x\rangle$。这也称为制备信道（或编码器）；更多内容请参见附录 B。

制备信道与后文所述的测量信道互为对偶；测量信道将量子态映射为经典概率分布，并通过采样映射为特定符号。

对于经典过程 $\overleftrightarrow{X}$，其实现由符号 $x \in \mathcal{X}$ 组成，相应的量子字母表 $\mathcal{Q}$ 为

通过 $\mathcal{E}$ 传递 $\mathcal{X}$ 中的每个元素而构造，使得 $\mathcal{Q} = \{\mathcal{E}(x)\}$。因此，$\mathcal{Q}$ 完全由$\mathcal{X}$ 和$\mathcal{E}$确定。例如，在独立同分布的情况下，公式 (3) 也可以写成 $\rho_{iid} = \mathcal{E}(X)$，而 ρ_{iid} 的每一种可能的纯态分解现在都可以解释为经典随机变量 X 与制备信道 $\mathcal{E}$的一种不同组合。

稍微滥用一下记号，我们写作 $|\Psi_{-\infty:\infty}\rangle = \mathcal{E}(\overleftrightarrow{X})$，表示量子过程 $|\Psi_{-\infty:\infty}\rangle$ 是通过将 $\overleftrightarrow{X}$ 中的每个随机变量传过经典-量子信道 $\mathcal{E}$而形成的。

请注意，一个无限量子位序列（可分离或纠缠的）可以看作由 $t \in \mathbb{Z}$索引的量子位一维晶格。这些可能相互纠缠的状态可以使用算子代数方法进行完整描述。我们在这一数学框架中构建量子过程的形式化定义。（参考文献 [17]对 $\mathbb{Z}$上的纠缠量子位序列的可观测量代数进行了更为详细的介绍。）

令 $\mathcal{B}_t$ 表示描述晶格点 t上所有可能观测量的 d维矩阵代数。（对于 $d = 2$，观测量空间由单位算子和 2×2 个复 Pauli（厄米、幺正）矩阵张成。）晶格点 t 处量子位的状态可以由作用于$\mathcal{H}_t$ 上的密度矩阵 ρ_t 描述，其中$\mathcal{H}_t$ 的维数为 d。对于由 ℓ 个连续量子位组成的区块，所有观测量都可以由晶格上 ℓ 个晶格点的联合代数描述：$\mathcal{B}_{0:\ell} = \bigotimes_{t=0}^{\ell-1} \mathcal{B}_t$，而该区块的状态为 $\rho_{0:\ell}$，作用于 $\mathcal{H}_{0:\ell} = \bigotimes_{t=0}^{\ell-1} \mathcal{H}_t$上；这是一个维数为$d^\ell$的希尔伯特空间。结合所有局域代数，即可在无限晶格上定义代数 $\mathcal{B}$ 。量子过程是无限晶格上的一个特定状态，可以写作 $|\Psi_{-\infty:\infty}\rangle$。

作为必要的第一步，为了更方便地将经典过程中的信息论工具加以运用，我们回到限制更多的情况：从有限字母表$\mathcal{Q}$ 中选取纯量子位态所构成的可分量子位序列。给定一个经典词 $w = x_0 x_1 \ldots x_{\ell-1} \in \mathcal{X}^\ell$和一个制备信道 $\mathcal{E} : \mathcal{X} \to \mathcal{Q}$，量子位序列的形式为：

$$\begin{aligned} |\psi_w\rangle &= \mathcal{E}(w) \\ &= |\psi_{x_0}\rangle \otimes |\psi_{x_1}\rangle \otimes \cdots \otimes |\psi_{x_{\ell-1}}\rangle \ , \end{aligned} \tag{5}$$

其中 ℓ 是序列的长度，且 $|\psi_w\rangle \in \mathcal{H}_{0:\ell}$。注意，$\dim(\mathcal{H}_{0:\ell}) = d^\ell$，$\mathcal{Q} = \{\mathcal{E}(x)\}$，长度为 ℓ 的可能量子位序列数为 $|\mathcal{Q}|^\ell$ ；或者，假设所有 $|\psi_x\rangle$ 均可区分，则为$|\mathcal{X}|^\ell$。

因此，可分量子过程 定义为$|\Psi_{-\infty:\infty}\rangle = \mathcal{E}(\overleftrightarrow{X})$。不同的制备——也就是说，$\overleftrightarrow{X}$ 与$\mathcal{E}$的不同组合——可能产生相同的量子过程。

量子过程的长度为 ℓ 的块密度矩阵集合为：

$$\begin{aligned} \rho_{0:\ell} &= \mathcal{E}(X_{0:\ell}) \\ &= \sum_{w \in \mathcal{X}^\ell} \Pr(|\psi_w\rangle)\, |\psi_w\rangle \langle\psi_w| \ , \end{aligned} \tag{6}$$

其中 $|\psi_w\rangle$ 是公式 (5) 中给出的可分向量。值得注意的是，它们的概率由底层经典随机过程 $\Pr(\overleftrightarrow{X})$ 的概率决定：$\Pr(|\psi_w\rangle) = \Pr(w)$。每个 $\rho_{0:\ell}$ 都是无限格点上纯量子态 $|\Psi_{-\infty:\infty}\rangle$ 的一个有限子系统。密度矩阵也采用左闭右开索引。

对于给定的 $\rho_{0:\ell}$，还可以在有限维希尔伯特空间 [33]中得到一种纯化。需要注意的是，由于 $\rho_{0:\ell}$ 没有唯一的纯态分解，因此通常无法由此重构概率 $\Pr(|\psi_w\rangle)$。相反，$\rho_{0:\ell}$ 仅包含观察者能够获取的信息。而且，如果 $\mathcal{Q}$ 包含非正交量子位态（对于某个$|\psi_x\rangle$，有$\langle\psi_x|\psi_{x'}\rangle \neq 0$ 、$|\psi_{x'}\rangle \in \mathcal{Q}$），那么观察者就无法明确地区分这些状态。

除了可分性之外，我们还关注平稳量子过程，其定义如下：

$$\rho_{0:\ell} = \rho_{t:\ell+t} \ , \tag{7}$$

对所有 $\ell \in \mathbb{Z}^+$ 和 $t \in \mathbb{Z}$均成立。如果 $\overleftrightarrow{X}$是平稳的，那么根据构造，$|\Psi_{-\infty:\infty}\rangle = \mathcal{E}(\overleftrightarrow{X})$也将是平稳的。
对于独立同分布量子过程，$\overleftrightarrow{X}$ 的联合概率按式（1）分解，因此该量子过程具有式（4）的形式。长度为ℓ的块密度矩阵可表示为乘积态：

$$\rho_{0:\ell} = \bigotimes_{t=0}^{\ell-1} \rho_{iid} \ , \tag{8}$$

其中 ρ_{iid} 采用式（3）的形式。
对于一个马尔可夫的底层经典过程 $\overleftrightarrow{X}$，还存在一些额外的微妙之处。$\overleftrightarrow{X}$的联合概率按式（2）进行因式分解，因此，$|\Psi_{-\infty:\infty}\rangle = \mathcal{E}(\overleftrightarrow{X})$ 的联合概率也会进行因式分解，使得：

$$\begin{aligned} \Pr(|\psi_w\rangle) = \Pr(|\psi_{x_0}\rangle) \Pr(|\psi_{x_1}\rangle \,|\, |\psi_{x_0}\rangle) \\ \cdots \Pr(|\psi_{x_{\ell-1}}\rangle \,|\, |\psi_{x_{\ell-2}}\rangle) \ . \end{aligned} \tag{9}$$

然而，观察者在测量量子过程时，无法可靠地区分不同的状态 $|\psi_x\rangle$ ，而底层的马尔可夫动力学也无法被观察到。因此，有记忆量子过程的一般情形属于隐马尔可夫过程。这些

过程最好通过能够直接表示过程结构的具体模型来引入。

D. 量子过程的表示

量子过程的表示 是一种模型，它包含内部状态以及状态之间的转移动态，并发射纯量子态，而非经典符号。与经典过程的表示一样，我们用状态转移图来描绘它们。当 $\overleftrightarrow{X}$是马尔可夫过程或隐马尔可夫过程时， $|\Psi_{-\infty:\infty}\rangle = \mathcal{E}(\overleftrightarrow{X})$ 可以通过扩展文献 [14]中的经典控制量子比特源（cCQS）来表示，如下所示。

一个隐马尔可夫链量子源（HMCQS）是一个三元组（*S, Q, T*），由以下部分组成：

- 一个由有限个内部状态组成的集合$\mathcal{S} = \{\sigma_1, \ldots, \sigma_n\}$，
- 一个有限字母表$\mathcal{Q} = \{|\psi_0\rangle \ldots, |\psi_{m-1}\rangle\}$，其中包含纯量子比特态，且每个 $|\psi_x\rangle \in \mathcal{H}^d$，以及
- 一个$\mathcal{T} = \{T^x : |\psi_x\rangle \in \mathcal{Q}\}$ 组成的转移矩阵集合。也就是说，如果源处于状态 σ_i，它以概率 $T^x_{ij} = \Pr(|\psi_x\rangle \sigma_j|\sigma_i)$发射量子位 $|\psi_x\rangle$，同时转移至内部状态σ_j。

与 HMC 类似，HM-CQS 的平稳分布满足 $\pi = \pi \sum_x T^x$。

任何能够精确再现量子过程的 HMCQS 都是该过程的生成式HMCQS，或称为该过程的生成器。尽管量子模型无法像经典模型那样进行预测，但我们可以定义经典单值性的类似概念。若对于每个状态 $\sigma \in \mathcal{S}$，都存在一种可能的测量，使得当时间 t 的状态为 σ 时，时间 $t+1$ 的内部状态完全确定，则称 HMCQS 是量子单值的。我们将在后文讨论量子单值性的若干含义。

我们称 HMCQS 为量子过程的经典控制器，因为其内部状态或转移动态都没有量子特性。这与相关的量子模型类别形成对比；后者在每个时间步都根据一个量子操作（通过一组 Kraus 算子定义）演化有限维量子系统。这些模型包括量子马尔可夫链(QMC) [34] 和隐量子马尔可夫模型(HQMM) [35–37]。HM-CQS 发出可分离量子态，而 QMC 和 HQMM 则生成测量结果序列（每个结果都对应一个特定的 Kraus 算子），这些序列构成经典随机过程。

考虑到未来的工作量，我们认为此时有必要提出一些关于相邻量子比特之间纠缠的观察。纠缠意味着有限长度的量子比特序列并非可分离的，因而无法用公式 (5) 描述。此外，无法直接参照某个基础经典随机过程来定义这些序列的概率。

尽管如此，仍有系统的方法来定义平稳的 $|\Psi_{-\infty:\infty}\rangle$，使得边缘分布集合 $\rho_{0:\ell}$ 能够描述对由 ℓ 个量子比特组成的块所进行的所有测量。例如，如果源的内部结构由一个 D 维量子系统构成，并且该系统在每个时间步与一个量子比特进行酉相互作用，那么它会生成键维数最大为 D [38]的矩阵乘积态（MPS）。如果源以随机方式（而非酉方式）运行，则可能以不同的概率发射许多不同的 MPS。此时，这一集合由矩阵乘积密度算符 （MPDO） [39]描述。我们将这些称为纠缠量子 ***dits*** 过程。其动力学与信息分析留待其他工作探讨。本文旨在为这些研究奠定基础。

E. 测量过程

观察量子过程的主体可以通过多种方式对其进行测量。设 M 表示对处于状态 ρ的量子位进行的测量。一般而言， M 是一个正算子值测度 （POVM），由希尔伯特空间$\mathcal{H}$ 上的一组半正定厄米算子$\{E_y\}$ 描述，该空间的维数为 d。每个 E_y 对应一个可能的测量结果 y，而 POVM 元素之和必须等于单位算子：

$$\sum_y E_y = \mathbb{I} \, .$$

投影值测度(投影值测量)是 POVM 的一个重要子类，并具有额外约束：算子 E_y 必须是正交投影算子。投影值测量至多包含 d 个元素。仅由作用于 $\mathcal{H}^d$ 上的秩一投影算子组成的投影值测量称为冯·诺依曼测量，并且恰好包含 d 个元素 [5]。

对一组 ℓ 量子比特进行的测量由希尔伯特空间$\mathcal{H}_{0:\ell}$ 上的某个块正算子值测度$\mathcal{M}_{0:\ell}$ 描述，其元素为$\{E_{y_{0:\ell}}\}$，该空间的维数为 d^ℓ。$\mathcal{M}_{0:\ell}$可以包括在多个量子比特的联合基中进行的测量——这对于完全刻画纠缠过程至关重要。

对于可分过程，我们关注“局部”测量——作用于单个量子位的算子。因此，对 ℓ 个量子位组成的区块，其测量算子具有张量积结构：

$$\mathcal{M}_{0:\ell} = \bigotimes_{t=0}^{\ell-1} M_t \, , \tag{10}$$

其中，每个 M_t 都是作用于 $\mathcal{H}_t$ 上的 POVM。

如果我们对每个量子比特应用相同的局部 POVM M，那么

$$\mathcal{M}_{0:\ell} = \bigotimes_{t=0}^{\ell-1} M \ . \tag{11}$$

我们将其称为重复*POVM*测量。

观察者还可以拥有多个 POVM，并根据某种测量协议在不同时间步进行不同的测量。我们将在稍后更详细地描述测量协议。

为简洁起见，下面忽略 $\rho_{0:\ell}$ 的测量后状态，只考虑测量结果 $y_{0:\ell}$。因此，我们将 $\mathcal{M}_{0:\ell}$ 视为一个随机映射 $\rho_{0:\ell} \to Y_{0:\ell}$，其中随机变量表示测量结果。

将形如 Eq. (11) 的测量应用于由 ℓ 个量子比特组成的有限块时，测量结果分解为一组 ℓ 经典随机变量：

$$\begin{aligned} Y_{0:\ell} &= Y_0 Y_1 \cdots Y_{\ell-1} \\ &= \mathcal{M}_{0:\ell}(\rho_{0:\ell}) \ , \end{aligned}$$

其中，每个 Y_t 的可能取值是 POVM 测量结果 $y \in \mathcal{Y}$。$Y_{0:\ell}$ 有 $|\mathcal{Y}|^\ell$ 种可能的实现。我们将长度为 ℓ 的一个实现（词）写作 $y_{0:\ell}$。

处于状态 $\rho_{0:\ell}$ 的一组 ℓ 量子位中，任意特定测量结果的概率为：

$$\Pr(y_{0:\ell}) = tr(E_{y_{0:\ell}} \rho_{0:\ell}) \ .$$

对于具有式 (6) 可分形式的 $\rho_{0:\ell}$，并且在每个量子位上都进行如式 (11) 所示的相同 POVM 测量，我们可以将 $E_{y_{0:\ell}}$ 分解为 ℓ 局部算子 E_{y_t}：

$$\begin{aligned} \Pr(y_{0:\ell}) &= tr\left(E_{y_{0:\ell}} \sum_{w \in \mathcal{X}^\ell} \Pr(w) \left|\psi_w\right\rangle \left\langle\psi_w\right| \right) \\ &= tr\left(\sum_{w \in \mathcal{X}^\ell} \Pr(w) \prod_{t=0}^{\ell-1} E_{y_t} \left|\psi_{x_t}\right\rangle \left\langle\psi_{x_t}\right| \right) . \end{aligned} \tag{12}$$

对于可分量子位过程，局部测量结果序列 $y_{0:\ell}$ 也可以解释为将随机变量 $X_{0:\ell}$ 从 $\overleftrightarrow{X}$ 经由同一个无记忆噪声信道 $\mathcal{C}: \mathcal{X} \to \mathcal{Y}$ 发送后得到的结果。$\mathcal{C}$ 可分解为确定性制备 $\mathcal{E}$ 和随机测量 $\mathcal{M}$：

$$\mathcal{C} = \mathcal{M} \circ \mathcal{E} \ .$$

（附录B对经典-量子信道 $\mathcal{E}$ 和 $\mathcal{M}$ 作了更详细的说明。）

这一构造清楚地表明，测量结果对应于经典随机变量 Y_t，其取值为 $y \in \mathcal{Y}$，并构成一个经典过程 $\overleftrightarrow{Y}$，其概率由式 (12) 定义。为表示量子过程与测量所得经典过程之间的关系，我们写作：

$$\overleftrightarrow{Y} = \mathcal{M}(|\Psi_{-\infty:\infty}\rangle) \ ,$$

其中 $\mathcal{M}$ 是一个重复施行的局部 POVM。如果量子位过程是可分离的，我们还可以写作：

$$\overleftrightarrow{Y} = \mathcal{C}(\overleftrightarrow{X}) \ .$$

F. 自适应测量协议

观察者无需在每个量子位上重复进行相同的测量，而可以根据某种算法，在不同时间步应用不同的 POVM。如果行动者利用过去的测量结果来决定选择哪种 POVM，我们就称其采用了自适应测量协议。

下面讨论以确定性有限自动机（DFA）作为底层控制器的测量协议。结合量子测量与 DFA 的类似构造已出现在量子文法的语境中 [40–42]。

一个确定性量子测量协议（DQMP）定义为五元组 $(\mathcal{S}, S_0, M, Y, \delta)$：

- 内部状态的有限集合 $\mathcal{S} = \{\sigma_1, \ldots\ldots, \sigma_n\}$，
- 唯一的初始状态 $S_0 \in \mathcal{S}$，
- 一组 POVM $M = \{M_s\}_{\sigma_s \in \mathcal{S}}$，每个内部状态对应一个 POVM，
- 一个字母表 $\mathcal{Y}$，其中包含与不同测量结果对应的 m 符号，以及
- 一个确定性转移映射 $\delta: \mathcal{S} \times \mathcal{Y} \to \mathcal{S}$。

如果 $\mathcal{S}$ 仅由 S_0 组成，那么该 DQMP 就是对 POVM M_{S_0} 进行的重复POVM测量。当 $\mathcal{S}$ 具有多个内部状态时，与不同状态对应的正算子值测度可能包含相同或不同数量的元素。同样，与其测量结果对应的符号集也可能彼此不相交，或者包含重复符号。

我们可以对集合 $\mathcal{Y}$ 的大小给出以下界：$m \leq \sum_s |\{E_{s,\ y}\}|$，其中 $\{E_{s,\ y}\}$ 是与 POVM M_s 对应的算子集合，而 $m \geq \max_s |\{E_{s,\ y}\}|$ 是元素最多的 POVM 的大小。

对于 DQMP $\mathcal{M}$ 和量子比特过程 $|\Psi_{-\infty:\infty}\rangle$，获得测量过程 $\overleftrightarrow{Y} = \mathcal{M}(|\Psi_{-\infty:\infty}\rangle)$，通常比重复进行 POVM 测量更困难。当观察者在 $t = 0$ 开始使用协议$\mathcal{M}$时，会经历两种截然不同的运行状态：首先是暂态动力学，然后是循环动力学。我们将简要概述这一过程，并在第五节讨论同步这一与暂态动力学密切相关的任务时，再回到这个主题。

$\mathcal{M}$ 从状态 S_0 于 $t = 0$ 开始。对于给定的（平稳、遍历的）输入 $|\Psi_{-\infty:\infty}\rangle$，随着 $t \to \infty$ DQMP 逐渐趋近于其部分内部状态上的平稳分布 $\pi = \{\Pr(\sigma_i) > 0,$ 对所有 $\sigma_i \in \mathcal{S}_r\}$，其中$\mathcal{S}_r \subseteq \mathcal{S}$ 是循环状态集。该分布（甚至哪些状态属于$\mathcal{S}_r$）取决于 $|\Psi_{-\infty:\infty}\rangle$。再现动力学由该平稳分布以及$\mathcal{S}_r$中状态之间的转移概率决定。不属于$\mathcal{S}_r$ 的任何状态都属于暂态状态集$\mathcal{S}_t$。

暂态动力学描述了$\mathcal{M}$ 如何从 S_0 于 $t = 0$ 转变为其在$\mathcal{S}_r$ 上的再现动力学，这可能在有限时间 $t = t_{sync}$ 内发生，也可能仅在渐近意义上发生，如 $t \to \infty$。通常，两种动力学会产生两个不同的测量过程 $\overleftrightarrow{Y}_r$ 和 $Y_{0:t_{sync}}$：前者按构造是平稳且遍历的，后者则不是。最终的测量过程包含两个部分，即 $\overleftrightarrow{Y} = \{\overleftrightarrow{Y}_r, Y_{0:t_{sync}}\}$。

G. 讨论

这些层层叠加的复杂性表明，我们应通过一个具体示例来梳理问题，并重申总体目标。

设想观察者对量子源发出的状态 ρ_t中的单个量子比特进行测量，所用的是计算基 $M_{01} = \{|0\rangle\langle 0|, |1\rangle\langle 1|\}$上的投影测量。可能的测量结果$y_0 = 0$ 和 $y_0 = 1$ 出现的概率分别为：

$$\Pr(y_0 = 0) = tr(|0\rangle\langle 0|\,\rho_t)$$
$$\Pr(y_0 = 1) = tr(|1\rangle\langle 1|\,\rho_t)\ ,$$

分别为这两个值确定了随机变量 Y_0 的分布；并且，对 $\rho_{0:\ell}$应用相同的投影测量后，我们便完全确定了测量块的统计特征 $\Pr(Y_{0:\ell})$。对 $\ell \to \infty$ 继续这一过程，即可定义测量过程 $\overleftrightarrow{Y}$。

当然，观察者也可以选择在另一个基底中进行测量；例如，$M_{\pm} = \{|+\rangle\langle +|, |-\rangle\langle -|\}$，其中 $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$。这通常会产生一个具有截然不同统计特征的测量过程 $\overleftrightarrow{Y}'$。

最后，观察者可以使用自适应测量协议 $\mathcal{M}$。观察者从状态 $s = S_0$ 开始，并使用M_{01}进行测量。如果 $y_0 = 0$，则停留在 S_0 并继续使用 M_{01}。如果 $y_0 = 1$，则转移到新的内部状态，并在下一个量子比特上使用 $M_{\pm}$。无论 $M_{\pm}$ 的测量结果如何，观察者都会返回 S_0，并使用 M_{01} 测量下一个量子比特。测量过程 $\overleftrightarrow{Y}''$ 将不同于 $\overleftrightarrow{Y}$和 $\overleftrightarrow{Y}'$，并且可能同时包含暂态分量和再现分量。

明确了这一设定后，我们现在可以更准确地陈述下文将要回答的问题：

1. 给定描述 $\rho_{0:\ell}$ 可分离量子位序列的密度矩阵，一般而言，测量结果序列有哪些性质？第三节将聚焦于此。在那里，对于某些类别的测量，$\rho_{0:\ell}$的量子信息性质为测量序列 $Y_{0:\ell}$ 的经典信息性质提供了界限。

2. 对于隐马尔可夫链量子源，掌握源信息的观察者何时能够确定其内部状态（即实现同步）？观察者能否在之后的时刻保持同步？第五节将探讨这些问题。

3. 如果观察者遇到一个未知的量子位源，那么通过有限资源的层析，观察者能够多准确地估计所发射过程的信息性质？如果观察者针对某个有限的 ℓ重建 $\rho_{0:\ell}$，又该如何构建源的近似模型？这是第六节的主题。

此外，第四节通过量子位过程的具体示例，阐释这些一般性结果及所需的分析方法。

III. 量子过程中的信息

我们希望对这样一类量子过程进行信息论分析：其观测到的序列取决于观察者选择的测量方式。（其中许多内容与附录 A 中回顾的经典信息度量相平行。）这需要采用一种更一般的方法，使用密度矩阵 $\rho_{0:\ell}$，其中包含描述对 ℓ 个量子位块进行任意测量所得测量结果所需的全部信息。我们运用量子信息论研究 $\rho_{0:\ell}$ 的集合所具有的性质，然后将这些性质与附录 A 中描述的测量

序列的经典性质联系起来。首先，我们简要回顾量子信息论中的几个基本量。参考文献 [6]和 [33] 对这一主题有更全面的介绍。

A. 冯·诺依曼熵

在量子信息论中，冯·诺依曼熵所起的作用类似于香农熵在经典信息论中的作用。给定一个混合量子态 ρ，冯·诺依曼熵为：

$$\begin{aligned} S(\rho) &= -tr(\rho \log_2 \rho) \\ &= -\sum_i \lambda_i \log_2 \lambda_i \ , \end{aligned} \tag{13}$$

其中， λ_i 是密度矩阵 ρ 的特征值。$S(\rho) = 0$ 当且仅当 ρ 是纯态。我们使用以 2 为底的对数 $\log_2(\cdot)$，因此冯·诺依曼熵的单位为比特。

由式（13）可知，冯·诺依曼熵是密度矩阵ρ的特征值分布的香农熵。因此：

$$S(\rho) = \min_M \mathrm{H}\left[M(\rho)\right] \ , \tag{14}$$

其中 $\mathrm{H}[\cdot]$ 是香农熵，最小值取遍所有秩一正算子值测度的集合。该最小值总能由其投影算子构成 ρ 的特征基 [6] 的一个投影值测量达到。我们使用方括号来表示 $M(\rho)$ 是关于测量结果的经典概率分布。

为监测两个量子系统之间的相关性，我们使用量子相对熵：

$$S(\rho\|\sigma) \equiv tr(\rho \log_2 \rho) - tr(\rho \log_2 \sigma) \ , \tag{15}$$

其中， ρ 和 σ 分别是两个系统的密度算子。量子相对熵非负：

$$S(\rho\|\sigma) \geq 0 \ , \tag{16}$$

当且仅当 $\rho = \sigma$时等号成立，这一结果称为克莱因不等式 [33]。

双体系统 AB 的状态 ρ^{AB} 的联合量子熵为：

$$\begin{aligned} S(A,B) &= S(\rho^{AB}) \\ &= -tr(\rho^{AB} \log_2 \rho^{AB}) \ , \end{aligned}$$

我们可以进一步定义条件量子熵：

系统 A 在给定系统 B 条件下的熵：

$$S(A|B) = S(A,B) - S(B) \ , \tag{17}$$

其中 $S(B) = S(\rho^B)$。注意， $S(A|B) \neq S(B|A)$。

与经典情形不同，条件量子熵可能为负值——超密集编码协议 [6]正是利用了这一现象。等价地，条件量子熵可以用量子相对熵写为：

$$\begin{aligned} S(A|B) &= -S(\rho^{AB}\|\mathbb{I}^A \otimes \rho^B) \\ &= \log_2(d_A) - S\left(\rho^{AB}\|\frac{\mathbb{I}^A}{d_A} \otimes \rho^B\right) \ , \end{aligned} \tag{18}$$

其中 $\mathbb{I}^A$ 是希尔伯特空间$\mathcal{H}^A$上的恒等算子，其维数为 d_A。

系统 A 和 B 之间的**量子互信息** 定义为：

$$\begin{aligned} S(A{:}B) &= S(A) - S(A|B) \\ &= S(A) + S(B) - S(A,B) \ . \end{aligned} \tag{19}$$

量子互信息是对称且非负的。如果联合系统 AB 处于纯态，那么 $S(A, B)$ 将为零，并且 $S(A) = S(B)$。它也可以表示为量子相对熵：

$$S(A{:}B) = S\left(\rho^{AB}\|\rho^A \otimes \rho^B\right) \ . \tag{20}$$

冯·诺依曼熵的一些其他众所周知的性质有助于推导后续结果。首先，对于可分量子位序列，每个 $\rho_{0:\ell}$ 都是由某个底层经典过程的长度为 ℓ 的词所形成的状态的有限混合，因此以下结果将很有用：

引理 1. 设随机变量 X 取值于集合 $x \in \{0、1、\ldots\ldots、n\}$，对应的概率为 $\{p_0、p_1、\ldots p_n\}$。给定一组密度矩阵 $\{\rho_0、\rho_1、\ldots\ldots、\rho_n\}$，以下不等式成立 *[33]*：

$$S\left(\sum_{x=0}^{n} p_x \rho_x\right) \leq \mathrm{H}\left[X\right] + \sum_{x=0}^{n} p_x S(\rho_x) \ ,$$

当且仅当所有 ρ_x 的支撑都位于正交子空间上时，等号成立。

其次，由于我们使用量子信道来制备和测量量子比特过程，因此我们利用了量子相对熵的单调性这一事实 [6]：

$$S(\rho\|\sigma) \geq S(\mathcal{E}(\rho)\|\mathcal{E}(\sigma)) \ , \tag{21}$$

其中 $\mathcal{E}$ 是任意量子信道。当且仅当存在一个恢复

映射 $\mathcal{R}$，使得 $\mathcal{R}(\mathcal{E}(\rho)) = \rho$ 且 $\mathcal{R}(\mathcal{E}(\sigma)) = \sigma$ [43]时，这个不等式取等号。

B. 量子块熵

由于平稳量子位过程在时间上存在相关性，我们探讨量子位块的冯·诺依曼熵如何随块大小变化。以下结果给出了从量子信息源中可能观测到的测量序列的界限。冯·诺依曼熵是香农熵的推广，因此这里的结果（以及许多证明）自然推广了附录 A 中的结果。我们还指出，对于较大的 ℓ，精确确定 $\rho_{0:\ell}$ 在实践中变得不可行。因此，第六节将讨论在仅限于测量有限长度块时，如何近似量子位过程的性质和模型。

对于量子位过程，我们将量子块熵定义为由 ℓ 个连续量子位组成的块的冯·诺依曼熵：

$$S(\ell) \equiv -tr(\rho_{0:\ell} \log_2 \rho_{0:\ell}) \ .$$

如果 $\rho_{0:\ell}$ 是纯态，则 $S(\ell) = 0$。根据经典情形中的相同逻辑，$S(0) \equiv 0$。
经典块熵的许多性质对于$S(\ell)$也成立。

命题 1。对于平稳量子位过程 $S(\ell)$ 是 ℓ的非递减函数。

证明。根据冯·诺依曼熵 *[33]*的强次可加性：

$$S(\rho^A) + S(\rho^C) \leq S(\rho^{AB}) + S(\rho^{BC}) \ . \tag{22}$$

设 $\rho^{ABC} = \rho_{0:2\ell+1}$，其中 $\rho^A = \rho_{0:\ell}$、$\rho^B = \rho_\ell$和 $\rho^C = \rho_{\ell+1:2\ell+1}$。
结合量子位过程的平稳性，我们将式（*22*）改写为：

$$\begin{aligned} S(\rho_{0:\ell}) + S(\rho_{\ell+1:2\ell+1}) &\leq S(\rho_{0:\ell+1}) + S(\rho_{\ell:2\ell+1}) \\ S(\ell) + S(\ell) &\leq S(\ell+1) + S(\ell+1) \\ S(\ell) &\leq S(\ell+1) \ , \end{aligned} \tag{23}$$

其中，式（*23*）由平稳性得出。
因此，对于所有 $\ell \geq 0$，$S(\ell) \leq S(\ell+1)$，且 $S(\ell)$ 是 ℓ 的非递减函数。

命题 2. 对于平稳量子位过程 $S(\ell)$ 是凹的。

证明。冯·诺依曼熵具有强次可加性*[33]*, 这意味着：

$$S(\rho^{ABC}) + S(\rho^B) \leq S(\rho^{AB}) + S(\rho^{BC}) \ . \tag{24}$$

对于 $\ell \geq 3$，令 $\rho^{ABC} = \rho_{0:\ell}$，其中 $\rho^A = \rho_0$、$\rho^B = \rho_{1:\ell-1}$和 $\rho^C = \rho_{\ell-1}$
利用量子比特过程的平稳性，我们可以改写公式 *(24)*：

$$\begin{aligned} S(\rho_{0:\ell}) + S(\rho_{1:\ell-1}) &\leq S(\rho_{0:\ell-1}) + S(\rho_{1:\ell}) \\ S(\ell) + S(\ell-2) &\leq S(\ell-1) + S(\ell-1) \\ S(\ell) - 2S(\ell-1) + S(\ell-2) &\leq 0 \ , \end{aligned} \tag{25}$$

其中公式 *(25)* 由平稳性推出。
因此，$S(\ell)$ 是凹的。

对于由将经典过程通过经典-量子通道传输而形成的可分量子位过程，它们的块熵之间满足以下关系：

命题 3。令 $|\Psi_{-\infty:\infty}\rangle = E(\overleftrightarrow{X})$。$|\Psi_{-\infty:\infty}\rangle$ 和 $\overleftrightarrow{X}$ 的块熵满足：

$$S(\rho_{0:\ell}) \leq \mathrm{H}\left[X_{0:\ell}\right] \ ,$$

对于所有 ℓ，当且仅当 Q 由 H 中维数为 $d \geq |X|$ 的 $|X|$正交纯态组成时，等号成立。

证明。由公式 *(6)* 可知，对于可分量子位过程：

$$\rho_{0:\ell} = \sum_{w \in \mathcal{X}^\ell} \Pr(w) \left|\psi_w\right\rangle \left\langle\psi_w\right| \ ,$$

其中每个 $|\psi_w\rangle$ 取公式 *(5)* 的可分形式，并且$w = x_{0:\ell}$。

我们首先注意到，要使 X 中的所有符号都与正交量子位态相关联，希尔伯特空间的最小维数为 $d_{min} = |X|$。

将 $\rho_{0:\ell}$ 写成可分量子位字的混合态后，我们应用引理 *1* 得到：

$$\begin{aligned} S(\rho_{0:\ell}) &\leq \mathrm{H}\left[X_{0:\ell}\right] + \sum_{w \in \mathcal{X}^\ell} \Pr(w) S(|\psi_w\rangle) \\ &\leq \mathrm{H}\left[X_{0:\ell}\right] \ , \end{aligned}$$

其中 $w = x_{0:\ell}$。由于每个 $|\psi_w\rangle$ 都是纯态，第二项在最后一行中计算为零。也就是说，$S(|\psi_w\rangle) = 0$ 对于所有 $w \in X^\ell$。

当且仅当状态 $|\psi_w\rangle$ 支撑在正交子空间上时等号成立，这要求 $d \geq |X|$ 以及 Q 的所有元素均相互正交。

对于一般的 POVM 测量，我们无法使用 $S(\rho_{0:\ell})$ 来界定测量过程 $\overleftrightarrow{Y}$的块熵。然而，当测量 $\mathcal{M}_{0:\ell}$ 仅由秩一 POVM（包括所有投影值测量）构成时，以下结论成立：

$$S(\rho_{0:\ell}) \leq \mathrm{H}\left[\mathcal{M}_{0:\ell}(\rho_{0:\ell})\right] \ , \tag{26}$$

当且仅当测量是在 $\rho_{0:\ell}$ 的最小熵（特征）基中进行时，等号成立。这可直接由公式 (14) 得出。

命题 4. 设 $\overleftrightarrow{Y} = M(|\Psi_{-\infty:\infty}i)$，其中 M 是重复秩一 *POVM* 测量。 $|\Psi_{-\infty:\infty}i$ 和 $\overleftrightarrow{Y}$的块熵满足：

$$S(\rho_{0:\ell}) \leq \mathrm{H}\left[Y_{0:\ell}\right] \ ,$$

对所有 ℓ 均成立；当且仅当 $|\Psi_{-\infty:\infty}i$ 是一个具有正交字母表 Q 的可分过程，且 M 使用的 *POVM* 算子中包含 Q 中每个元素对应的一个投影算子时，等号成立。

证明. 该界可直接由公式 *(26)* 得出，因为重复秩一 *POVM* 测量是更一般的测量序列 $M_{0:\ell}$的一个子类。

等号成立的条件也可由公式 *(26)* 得出，但需要进一步说明。首先，考虑使用 *POVM* M 测量单量子比特边缘态 ρ_0 。若等号成立，则 ρ_0 *({*$|e_i i$*})*的特征基中的每个元素都必须在 M 中有一个对应的算子，且该算子是作用于相应本征子空间的投影算子（$E_i = |e_i i\, h e_i|$）。

我们可以写作 $\rho_0 = \sum_i p_i\, |e_i i h e_i|$。由于我们对每个量子位都应用相同的 *POVM*，所有长度为 ℓ 的块都必须具有形式为 $\bigotimes_{t=0}^{\ell-1} |e_t$ 的本征态——也就是说，它们必须采用公式 *(5)* 的可分形式——因此 $|\Psi_{-\infty:\infty}$ 是一个具有由正交态组成的量子字母表 Q 的可分过程。由 Q 中各状态上的投影算子构成的 M ，便是对块 $\rho_{0:\ell}$ 进行测量时熵最小的测量。请注意，*POVM* 中可能还有其他元素并非投影算子，因为将它们应用于 $\rho_{0:\ell}$ 时，相应测量结果的概率为 0 。（如果该过程未使用希尔伯特空间的那部分，那么如何测量它并不重要。）

总之，对于可分离量子位过程，$S(\ell)$ 的上限由其底层经典过程的块熵 $\mathrm{H}\left[X_{0:\ell}\right]$ 给出。对于采用秩一 POVM 的重复测量， $S(\ell)$ 是所有经典测量过程块熵 $\mathrm{H}\left[Y_{0:\ell}\right]$ 的下限。$\mathrm{H}\left[X_{0:\ell}\right]$ 与 $\mathrm{H}\left[Y_{0:\ell}\right]$ 之间没有直接关系，而是取决于$\mathcal{E}$ 和 $\mathcal{M}$的具体情况。

C. 冯·诺依曼熵率

量子位过程的冯·诺依曼熵率为：

$$s = \lim_{\ell\to\infty} \frac{S(\ell)}{\ell} \ . \tag{27}$$

s 的单位是 每个时间步的比特数。这一量等同于 平均熵，最初是在量子统计力学的背景下引入的 [44]。对于所有平稳过程 [45]，该极限都存在。从操作层面来看，s 是平稳量子源 [17] 的最优编码率。

命题 5. 对于平稳量子位过程，我们可以将冯·诺依曼熵率等价地写为：

$$s = \lim_{\ell\to\infty} S(\rho_0|\rho_{-\ell:0}) \ . \tag{28}$$

证明。 此证明与参考文献 *[46]* 中经典熵率的证明十分相似。我们先证明$\lim_{\ell\to\infty} S(\rho_0|\rho_{-\ell:0})$ 存在，然后证明它等价于公式（*27*）中的极限。当 $S(\rho_0|\rho_{-\ell:0})$是关于 ℓ 的非负递减函数时，该极限存在：

$$\begin{aligned} S(\rho_0|\rho_{-\ell:0}) &= S(\rho_{-\ell:1}) - S(\rho_{-\ell:0}) \\ &= S(\ell+1) - S(\ell) \qquad (29) \\ &\geq 0 \ , \qquad (30) \end{aligned}$$

其中，公式（*29*）由平稳性推出，公式（*30*）则由 $S(\ell)$ 的非递减性质推出。结合 $S(\ell)$ 为凹函数这一事实，可知$\lim_{\ell\to\infty} S(\rho_0|\rho_{-\ell:0})$ 存在。

现在，我们证明$\lim_{\ell\to\infty} S(\rho_0|\rho_{-\ell:0})$ $=\lim_{\ell\to\infty} S(\ell)/\ell$。对长度为 ℓ 的区块反复应用公式（*17*），可得冯·诺依曼熵的以下链式法则：

$$S(\rho_{0:\ell}) = \sum_{j=0}^{\ell-1} S(\rho_j|\rho_{j-1}...\rho_0) \ . \tag{31}$$

我们可以调整指标（由于平稳性），并将两边除以 ℓ ，得到：

$$\frac{S(\rho_{0:\ell})}{\ell} = \frac{1}{\ell}\sum_{i=1}^{\ell} S(\rho_i|\rho_{i-1}...\rho_1) \ . \tag{32}$$

最后几步需要用到以下结果，称为切萨罗平均 *[46]*：

引理 2. 如果 $a_n \to a$ 和 $b_n = \frac{1}{n}\sum_{i=1}^{n} a_i$，那么 $b_n \to a$。

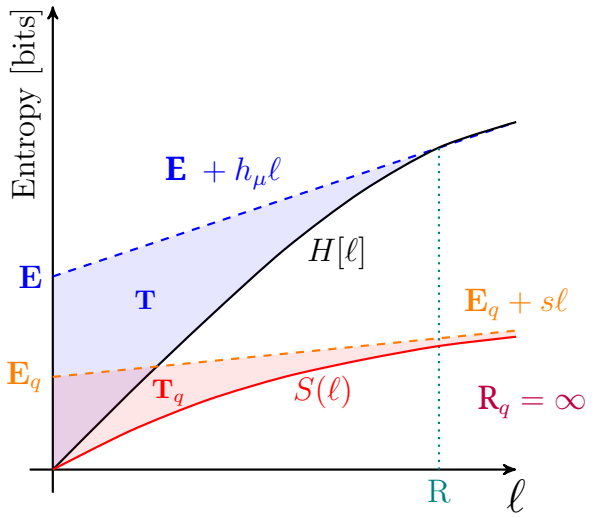


图 2. 块熵收敛到其线性渐近线。$\mathbf{H}[\ell]$ 是具有马尔可夫阶数 R（见附录 A）的有限经典过程的块熵，而 $S(\ell)$ 是具有无限量子马尔可夫阶 R_q 的有限量子过程的量子块熵。对于经典过程，$\mathbf{E}$ 是超额熵，而 h_μ 是其香农熵率。类似地，对于量子过程，$\mathbf{E}_q$ 是量子超额熵，而 s 是其冯·诺依曼熵率。蓝色阴影区域的面积是经典瞬态信息 $\mathbf{T}$，红色阴影区域的面积是量子瞬态信息 $\mathbf{T}_q$。

对等式 *(32)* 两边取极限并应用引理 *2*，我们得到：

$$\lim_{\ell\to\infty}\frac{S(\rho_{0:\ell})}{\ell} = \lim_{\ell\to\infty} S(\rho_\ell|\rho_{\ell-1}\ldots\rho_1)\ . \tag{33}$$

利用平稳性，$\lim_{\ell\to\infty} S(\rho_\ell|\rho_{\ell-1}\ldots\rho_1)\lim_{\ell\to\infty} S(\rho_0|\rho_{-\ell:0})$结合方程，这证明了我们对 s 的两种定义是等价的。 *(33)*

为了说明下文的一些结果，重要的是要认识到，仅仅因为一个过程具有由公式（27）给出的冯·诺依曼熵率，并不意味着观察者能够对任意量子比特进行测量，使得该次测量的不确定性为 s。相反，获得 s 对应于对整个量子比特链进行测量所用的测量基，在该测量基下，结果分布的香农熵最小。对于许多非平凡的例子，这个基高度非局域，或者在实验上无法实现。与经典情形一样，s 在图形上表现为 $S(\ell)$的斜率，其中 $\ell\to\infty$，如图 2 所示。

对于可分离量子比特过程，我们还可以将 s 与底层过程 $\overleftrightarrow{X}$的经典熵率联系起来，如下所示：

命题 6。设 $|\Psi_{-\infty:\infty}\rangle\!\rangle = E(\overleftrightarrow{X})$。那么，$|\Psi_{-\infty:\infty}\rangle\!\rangle$ 的冯·诺依曼熵率 s 满足以下界：

$$s \le h_\mu^X\ ,$$

其中 h_μ^X 是 $\overleftrightarrow{X}$的香农熵率。当且仅当 Q 由 H^d 中 $d \ge |X|$维空间里的 $|X|$ 个正交纯态组成时，等号成立。

证明。将命题 *3* 两边除以 ℓ，再令 $\ell\to\infty$，并对两边取极限，得到：

$$\lim_{\ell\to\infty}\frac{S(\rho_{0:\ell})}{\ell} \le \lim_{\ell\to\infty}\frac{\mathrm{H}[X_{0:\ell}]}{\ell}\ .$$

根据公式（*27*），左侧为 s；根据式（*A1*），右侧为 h_μ^X。等号成立的条件继承自命题 *3*，证明完毕。

再次将范围限定为使用秩一 POVM 的重复测量，我们可以为测量过程证明以下界：

命题 7. 设 $\overleftrightarrow{Y} = M(|\Psi_{-\infty:\infty}\rangle\!\rangle)$，且 M 为重复秩一 *POVM*。测量熵率h_μ^Y 满足：

$$s \le h_\mu^Y\ ,$$

其中，s 是 $|\Psi_{-\infty:\infty}\rangle\!\rangle$ 的冯·诺依曼熵率；等号成立当且仅当 $|\Psi_{-\infty:\infty}\rangle\!\rangle$ 是具有正交字母表 Q 的可分离过程，且 M 使用的 *POVM* 算子中包含 Q 中每个元素对应的一个投影算子。

证明。将命题 *4* 两边同时除以 ℓ，并令 $\ell\to\infty$ 趋于极限，可得：

$$\lim_{\ell\to\infty}\frac{S(\rho_{0:\ell})}{\ell} \le \lim_{\ell\to\infty}\frac{\mathrm{H}[Y_{0:\ell}]}{\ell}\ .$$

左侧为 s，右侧为 h_μ^Y。等号成立的条件继承自命题 *4*，证明完毕。

D. 量子冗余

不同于经典过程，量子比特过程的最大熵率取决于希尔伯特空间的大小，而非字母表 $\mathcal{Q}$ 的大小。对于维数为 d 的希尔伯特空间，s 的最大可能值为 $\log_2(d)$，这对应于一个独立同分布序列，其中每个量子位都处于最大混合态 $\rho_{iid} = \mathbb{I}/d$。

量子比特过程可以压缩到其冯·诺依曼熵率 s。它可以被压缩的量是

压缩量即为量子冗余：

$$\mathbf{R}_q \equiv \log_2(d) - s \ . \tag{34}$$

单个量子位中的统计偏差以及它们之间的时间相关性，为压缩提供了机会。$\mathbf{R}_q$ 则包含了这两种效应。

对于可分离量子位过程，我们可以利用底层经典过程的性质来界定量子冗余：

命题 8. 设 $\overleftrightarrow{X}$ 是一个冗余为 $\mathbf{R}^X$、符号字母表为 X、熵率为 h_μ^X 的经典过程；设 $|\Psi_{-\infty:\infty}\rangle = E(\overleftrightarrow{X})$ 是一个冗余为 $\mathbf{R}_q$、希尔伯特空间维度为 d、熵率为 s 的量子位过程。

对于 $d \geq |X|$:

$$\mathbf{R}_q \geq \mathbf{R}^X \ ,$$

当且仅当 $d = |X|$ 且 Q 由 $|X|$ 个正交纯态组成时，等号成立。
对于 $d < |X|$,

$$\mathbf{R}_q < \mathbf{R}^X + (h_\mu^X - s) \ ,$$

其中，$(h_\mu^X - s)$ 这一项根据命题（6）始终为正。

证明。 首先，考虑 $d = |X|$:

$$\begin{aligned} \mathbf{R}_q &= \log_2(|\mathcal{X}|) - s \\ &= \mathbf{R}^X + h_\mu^X - s \\ &\geq \mathbf{R}^X \ , \end{aligned}$$

最后一行来自命题 6，等号成立的条件也是如此。

对于 $d > |X|$:

$$\begin{aligned} \mathbf{R}_q &= \log_2(d) - s \\ &> \log_2(|\mathcal{X}|) - s \\ &> \mathbf{R}^X \ . \end{aligned}$$

不存在等号成立的可能。在这种情况下，Q 不会张成 H，自然会导致更多冗余。
最后，对于 $d < |X|$:

$$\begin{aligned} \mathbf{R}_q &< \log_2(d) - s \\ &< \log_2(|\mathcal{X}|) - s \\ &< \mathbf{R}^X + (h_\mu^X - s) \ , \end{aligned}$$

至此证明完毕。

我们还可以比较对某个测量过程（通过重复使用秩为一的 POVM 获得）的经典冗余与被测量的量子比特过程的量子冗余。

命题 9。 设 $\overleftrightarrow{Y}$ 是一个测量过程，使得 $\overleftrightarrow{Y} = M(|\Psi_{-\infty:\infty}\rangle)$，且 M 是一个重复秩一 POVM。设 $\overleftrightarrow{Y}$ 的冗余为 $\mathbf{R}^Y$，$|\Psi_{-\infty:\infty}\rangle$ 的量子冗余为 $\mathbf{R}_q$。则：

$$\mathbf{R}_q \leq \mathbf{R}^Y \ , \tag{35}$$

等号成立当且仅当 $d = |Y|$，$|\Psi_{-\infty:\infty}\rangle$ 是一个具有正交字母表 Q 的可分过程，且 M 使用的 POVM 的算子中，对 Q 中每个元素都包含一个投影算子。

证明。 H 上的秩一正算子值测度至少包含 d 个元素，因此 $|Y| \geq d$:

$$\begin{aligned} \mathbf{R}_q &= \log_2(d) - s \\ &\leq \log_2(|\mathcal{Y}|) - s \\ &\leq \mathbf{R}^Y + h_\mu^Y - s \\ &\leq \mathbf{R}^Y \ . \end{aligned}$$

从第一行到第二行可得到等号成立的一个条件：$d = |Y|$。最后一行用到了命题 7，并给出了等号成立的其他条件。

E. 量子熵增益

我们可以对 $S(\ell)$ 求离散时间导数，正如在 [11] 中对 $\mathrm{H}[\ell]$ 所做的那样。附录 A 总结了这一过程。我们将 $S(\ell)$ 的一阶导数称为量子熵增益：

$$\Delta S(\ell) \equiv S(\ell) - S(\ell - 1) \ , \tag{36}$$

对于 $\ell > 0$。量子熵增益的单位是每个时间步的比特数，并设边界条件为 $\Delta S(0) = \log_2(d)$。由于 $S(\ell)$ 单调递增，$\Delta S(\ell) \geq 0$。

量子熵增益是将 ℓ^{th} 量子位纳入一个区块后引入的额外不确定性，其不确定性以冯·诺依曼熵来量化。

结合式 (36) 和式 (17)，我们可以将 $\Delta S(\ell)$ 写为：

$$\Delta S(\ell) = S(\rho_0 | \rho_{-\ell:0}) \ .$$

由此可以将量子熵增益与冯·诺依曼熵率联系起来，如下所示：

$$s = \lim_{\ell \to \infty} \Delta S(\ell) \ . \tag{37}$$

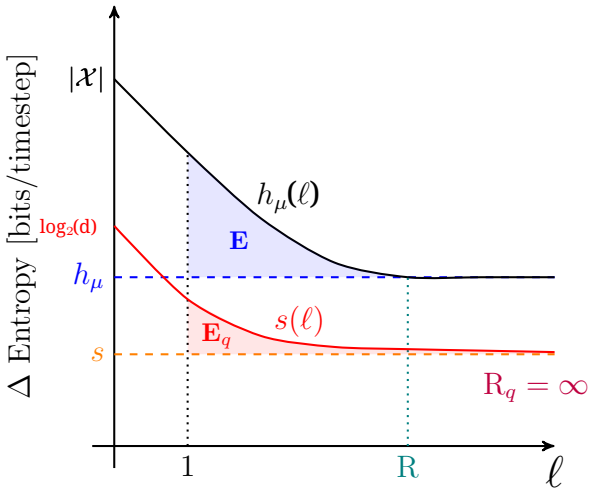


图 3. $\Delta H(\ell)$（对于马尔可夫阶数为 R的有限型经典过程）和 $\Delta S(\ell)$（对于量子马尔可夫阶为无限的有限型量子过程 R_q）向过程熵率 h_μ 和 s收敛的情况。阴影区域分别表示经典（蓝色）和量子（红色）超额熵。

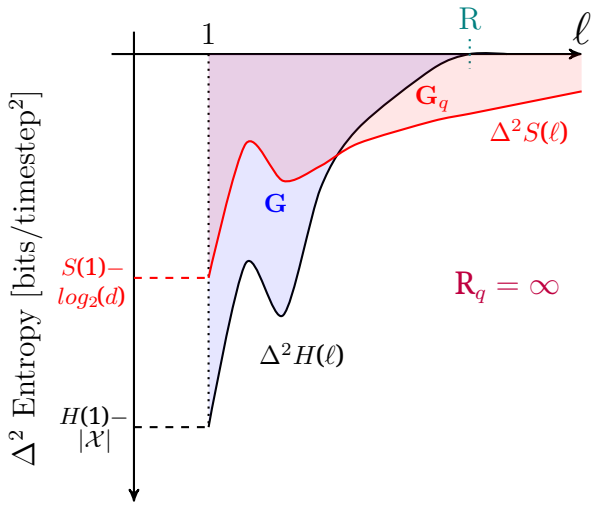


图 4. 对于马尔可夫阶数为 R 的有限经典过程，以及量子马尔可夫阶为无穷的有限量子过程，可预测性（$\Delta^2 H(\ell)$ 和$\Delta^2 S(\ell)$）收敛至 0。请注意，可预测性并非单调变化（不同于块熵及其一阶导数）。重叠的阴影区域表示经典（蓝色）和量子（红色）可预测性 $\mathbf{G}$ 和 $\mathbf{G}_q$ 的大小，按照约定它们为负值。

因此，与经典情形类似，量子熵增益可作为冯·诺依曼熵率的有限-ℓ 近似：

$$s(\ell) \equiv \Delta S(\ell) \ . \tag{38}$$

$s(\ell)$是观察者对量子位过程熵率所能作出的最佳估计，而该观察者只能获取长度为ℓ 的量子位块的测量统计。

熵率估计的收敛方式及其与量子位过程其他信息性质之间的关系总结于图 3 中。

F. 量子可预测性增益

我们将 $S(\ell)$ 的二阶导数称为<style id='3'>量子可预测性增益，其定义为：

$$\begin{aligned}\Delta^2 S(\ell) &\equiv \Delta s(\ell) \\ &= s(\ell) - s(\ell-1) \ ,\end{aligned}$$

对于 $\ell > 0$而言， $\Delta^2 S(\ell)$的单位是 每个时间步的比特数²。由于 $S(\ell)$是凹函数，因此 $\Delta^2 S(\ell) \leq 0$。 $|\Delta^2 S(\ell)|$ 量化了这样一种程度：如果观察者将其观察范围从 $\ell - 1$ 量子比特的块扩大到 ℓ 量子比特的块，其对冯·诺依曼熵率的估计 s 会得到改善。 $\Delta^2 S(\ell)$的一般收敛行为如图 4所示。

对于 $S(\ell)$的所有高阶离散导数（与经典块熵一样）：

$$\lim_{\ell\to\infty} \Delta^n S(\ell) = 0, n \geq 2 \ .$$

这直接源于定态量子态的公式（27）中的极限存在。

G. 总量子可预测性

截至目前，通过对冯诺依曼块熵取离散时间导数来引入可分量子过程的新信息论特征。类似地，我们也可以对函数 $\Delta^n S(\ell)$ 进行积分，就像经典情形中使用公式（A3）所做的那样。尽管这一过程起初相当直接，但会产生一些有趣的新信息量。

量子比特过程的这些性质具有如下的一般形式：

$$\mathcal{I}_n^q \equiv \sum_{\ell=\ell_0}^{\infty} [\Delta^n S(\ell) - \lim_{\ell\to\infty} \Delta^n S(\ell)] \ ,$$

其中， ℓ_0 是使 $\Delta^n S(\ell)$有定义的 ℓ 的第一个值。 $\mathcal{I}_2^q$ ，前者监测量子可预测性增益 $\Delta^2 S(\ell)$ 收敛到其极限 0 的过程，适用于

$\ell \to \infty$。我们使用 $\ell_0 = 1$ 来得到总量子可预测性$\mathbf{G}_q$:

$$\begin{aligned}\mathbf{G}_q &\equiv \mathcal{I}_2^q \\ &= \sum_{\ell=1}^{\infty} \Delta^2 S(\ell) \ . \end{aligned} \tag{39}$$

$\mathbf{G}_q$ 的单位是每个时间步的比特数。请注意，对于所有 ℓ, $\mathbf{G}_q \leq 0$，因为 $\Delta^2 S(\ell) \leq 0$ 。
通过将$\mathbf{G}_q$ 与量子比特过程一个先前已确立的性质——量子冗余——联系起来，可以对其进行诠释。

命题 10。 对于平稳量子位过程:

$$\mathbf{G}_q = -\mathbf{R}_q \ .$$

证明。 将式 *(A3)* 应用于式 *(39)*，可得:

$$\begin{aligned}\mathbf{G}_q &= \lim_{\ell\to\infty} \Delta S(\ell) - \Delta S(0) \\ &= s - \log_2(d) \\ &= -\mathbf{R}_q \ . \end{aligned}$$

这里，第二行由式 *(37)* 推出，第三行由式 *(34)* 推出。

因此， $|\mathbf{G}_q|$ 是量子比特过程每个时间步中可预测信息的总量。
很容易看出，单个量子位中的信息量可分解为:

$$\log_2(d) = |\mathbf{G}_q| + s \ .$$

$|\mathbf{G}_q|$ 是一个量子位中可预测的量子信息量，而 s 是不可约随机的信息量。

关系式 $|\mathbf{G}_q| = \mathbf{R}_q$ 可与命题 8 和 9 结合起来，证明两个推论。
推论 1. 设 $\overleftrightarrow{X}$ 是一个经典过程，其总可预测性为$\mathbf{G}^X$，符号字母表为 $\boldsymbol{X}$，熵率为h_μ^X；设 $|\Psi_{-\infty:\infty}\rangle = E(\overleftrightarrow{X})$ 是一个量子比特过程，其总量子可预测性为$\mathbf{G}_q$，希尔伯特空间的维数为d ，熵率为 s。

对于 $d \geq |\boldsymbol{X}|$:

$$|\mathbf{G}_q| \geq |\mathbf{G}^X| \ ,$$

当且仅当 $d = |\boldsymbol{X}|$ 且 Q 由 $|\boldsymbol{X}|$个正交纯态构成时，等号成立。

对于 $d < |\boldsymbol{X}|$,

$$|\mathbf{G}_q| < |\mathbf{G}^X| + (h_\mu^X - s) \ ,$$

其中，项 $(h_\mu^X - s)$ 根据命题（*6*）始终为正。

证明。 这可由结合命题*8*和命题*10*立即得出。

推论2。 设 $\overleftrightarrow{Y}$是一个测量过程，使得 $\overleftrightarrow{Y} = M(|\Psi_{-\infty:\infty}\rangle)$，且 M 是重复秩一*POVM*。设 $\overleftrightarrow{Y}$的冗余为 $\mathbf{G}^Y$，且 $|\Psi_{-\infty:\infty}\rangle$ 的量子冗余为 $\mathbf{G}_q$。则:

$$|\mathbf{G}_q| \leq |\mathbf{G}^Y| \ ,$$

当且仅当 $d = |Y|$, $|\Psi_{-\infty:\infty}\rangle$ 是具有正交字母表 Q 的可分过程，且 M 使用的 *POVM* 其算子包含 Q 中每个元素对应的一个投影算子时，等号成立。

证明。 这可由结合命题9和10立即得出。

从图形上看，总量子可预测性就是可预测性增益曲线与其为 0 的线性渐近线之间的面积，如图 4 所示。冯·诺依曼熵率和总可预测性有助于理解平稳源的压缩极限。然而，它们并未说明这种压缩是否可实现，也没有说明其可实现性是源于单个量子位状态中的偏差，还是量子位之间的相关性。为此，我们必须沿熵层次结构继续向上追溯。

H. 量子超额熵

$s(\ell)$ 收敛至真实冯·诺依曼熵率 s 的情况，可用量子超额熵来量化:

$$\begin{aligned}\mathbf{E}_q \equiv \mathcal{I}_1^q &= \sum_{\ell=1}^{\infty} \left[\Delta S(\ell) - \lim_{\ell\to\infty} \Delta S(\ell)\right] \\ &= \sum_{\ell=1}^{\infty} \left[s(\ell) - s\right] \ . \end{aligned} \tag{40}$$

$\mathbf{E}_q$ 的单位是比特。与经典情形类似，我们将具有有限$\mathbf{E}_q$ 的任何量子位过程称为有限型，将具有无限$\mathbf{E}_q$ 的过程称为无限型。
我们还可以根据 $S(\ell)$ 的渐近行为来表示$\mathbf{E}_q$ 。

命题11。 量子超额熵可以写为:

$$\mathbf{E}_q = \lim_{\ell\to\infty} \left[S(\ell) - s\ell\right] \ . \tag{41}$$

证明。 我们利用部分和，通过公式 *(A3)* 计算公式 *(40)* 中的离散积分：

$$\begin{aligned}\mathbf{E}_q &= \lim_{\ell\to\infty}\sum_{m=1}^{\ell}\left[\Delta s(m) - s\right] \\ &= \lim_{\ell\to\infty}\left[S(\ell) - S(0) - s\ell\right] \\ &= \lim_{\ell\to\infty}\left[S(\ell) - s\ell\right] \ ,\end{aligned}$$

根据定义，自 $S(0) = 0$ 。

对于有限量子过程，$\mathbf{E}_q$ 是熵增曲线与其渐近线之间的面积 s ，如图 3 所示。它还出现在图 2 中，表示线性渐近线相对于 $S(\ell)$ 曲线的垂直偏移。

这使得量子块熵可以自然地按以下方式进行标度：

$$S(\ell) \sim \mathbf{E}_q + s\ell \ ,$$

当 $\ell \to \infty$ 时。

以下命题对 $\mathbf{E}_q$ 作为量子互信息给出了更清晰的解释：

命题12。 量子超额熵可以写为：

$$\mathbf{E}_q = \lim_{\ell\to\infty} S(\rho_{-\ell:0}{:}\rho_{0:\ell}) \ ,$$

其中， $\rho_{-\ell:0}$ 和 $\rho_{0:\ell}$ 是两个由 ℓ 个连续量子位组成且具有共享边界的块。

证明。 根据式（*19*），两个相邻的、各含 ℓ 个量子位的块之间的量子互信息可表示为：

$$\begin{aligned}S(\rho_{0:\ell}{:}\rho_{-\ell:0}) &= S(\rho_{0:\ell}) - S(\rho_{0:\ell}|\rho_{-\ell:0}) \\ &= S(\ell) - \sum_{t=0}^{\ell-1} S(\rho_t|\rho_{-\ell:t-1}) \ ,\end{aligned}$$

其中，最后一行是通过反复应用式（*17*）得到的。

取 $\ell \to \infty$：

$$\begin{aligned}\lim_{\ell\to\infty} S(\rho_{0:\ell}{:}\rho_{-\ell:0}) &= \lim_{\ell\to\infty}\left[S(\ell) - \sum_{t=0}^{\ell-1} S(\rho_t|\rho_{-\ell:t-1})\right] \\ &= \lim_{\ell\to\infty}\left[S(\ell) - s\ell\right] \ ,\end{aligned}$$

其中，最后一行由式（*32*）和稳态性得出。

这个最终表达式等价于命题*11*中推导出的$\mathbf{E}_q$ 形式，从而完成证明。

$\mathbf{E}_q$ 因此是无限量子比特序列两半之间所有相关性的度量。$\mathbf{E}_q = 0$，当且仅当源是独立同分布的（且 $S(\ell) = s\ell$）。

我们可以将可分量子位过程的$\mathbf{E}_q$ 与其底层经典过程的 $\mathbf{E}^X$ 联系起来：

命题 13。 设 $\overleftrightarrow{X}$ 是一个字母表为 X、超额熵为$\mathbf{E}^X$的经典过程，并设 $|\Psi_{-\infty:\infty}\rangle = E(\overleftrightarrow{X})$ 是一个字母表为 Q、量子超额熵为$\mathbf{E}_q$的量子位过程。

那么：

$$\mathbf{E}_q \leq \mathbf{E}^X \ ,$$

当且仅当 Q 由 $|X|$ 个正交态组成时，等号成立。

证明。 考虑经典过程 $\overleftrightarrow{X}$中长度为 2ℓ 的块 $X_{-\ell:\ell}$。我们可以将 $X_{-\ell:\ell}$ 的实现写入经典寄存器，从而得到状态：

$$\rho^C_{-\ell:\ell} = \sum_{x_{-\ell:\ell}} \Pr(x_{-\ell:\ell}) \left|x_{-\ell:\ell}\right\rangle \left\langle x_{-\ell:\ell}\right| \ , \tag{42}$$

接下来，我们让每个符号通过制备信道 E，得到量子位过程 $\rho_{-\ell:\ell} = E^{2\ell}(\rho^C_{-\ell:\ell})$的块，其中所有 $|x_{-\ell:\ell}\rangle$ 都彼此正交。这里 $E^{2\ell} = \bigotimes_{i=0}^{2\ell} E$。

我们可以将量子互信息表示为量子相对熵（式 *(20)*），从而得到以下关系：

$$\begin{aligned}\mathrm{I}\left[X_{-\ell:0}{:}X_{0:\ell}\right] &= S(\rho^C_{-\ell:0}{:}\rho^C_{0:\ell}) \\ &\geq S(\mathcal{E}^\ell(\rho^C_{-\ell:0}){:}\mathcal{E}^\ell(\rho^C_{0:\ell})) \\ &\geq S(\rho_{-\ell:0}{:}\rho_{0:\ell}) \ ,\end{aligned} \tag{43}$$

其中，式 *(43)* 来自式 *(21)* 中量子相对熵的单调性。等号成立的条件也来自式 *(21)*。恢复映射必须存在的状态集合为 Q，而这只有在所有 $|\psi_x\rangle$ 都可区分——即彼此正交——时才有可能。

利用式 *(A9)*，将 $\overleftrightarrow{X}$ 的超额熵写成极限，我们得到：

$$\begin{aligned}\lim_{\ell\to\infty}\mathrm{I}\left[X_{-\ell:0}{:}X_{0:\ell}\right] &\geq \lim_{\ell\to\infty} S(\rho_{-\ell:0}{:}\rho_{0:\ell}) \\ \mathbf{E}^X &\geq \mathbf{E}_q \ .\end{aligned}$$

当我们对量子过程进行重复POVM测量以得到经典过程时，也会出现类似的界。

命题 14。 设 $\overleftrightarrow{Y}$为一个测量过程，使得 $\overleftrightarrow{Y} = M(|\Psi_{-\infty:\infty}\rangle)$，且 M 是重复秩一

POVM。设 $\overleftrightarrow{Y}$的超额熵为 $\mathbf{E}^Y$，量子超额熵为$\mathbf{E}_q$。则：

$$\mathbf{E}_q \geq \mathbf{E}^Y \ ,$$

当且仅当 $|\Psi_{-\infty:\infty}\rangle$ 是具有正交字母表 Q 的可分过程，且 M 使用的 *POVM* 算子中包含 Q 中每个元素对应的投影算子时，等号成立。

证明。 考虑 $\rho_{-\ell:\ell}$——量子过程 $|\Psi_{-\infty:\infty}\rangle$ 中长度为 2ℓ 的一个块——并令 M 为对秩一 *POVM* M 的重复测量，其元素为 $\{E_y\}$，使得$M(\rho_t) = \sum_y \Pr(y)\ |y\rangle\langle y|$, $\Pr(y) = tr(E_y\rho_t)$，且所有 $|y\rangle$ 都彼此正交。于是，在长度为 ℓ 的块上进行重复测量，得到 $M_{0:\ell} = \bigotimes_{i=0}^{\ell-1} M$。

我们将量子互信息表示为量子相对熵（使用式 *(20)*），并应用式 *(21)*，得到：

$$\begin{aligned} S(\rho_{-\ell:0}{:}\rho_{0:\ell}) &\geq S\left(\mathcal{M}_{0:\ell}(\rho_{-\ell:0}){:}\mathcal{M}_{0:\ell}(\rho_{0:\ell})\right) \\ &\geq S\Bigg(\sum_{y_{-\ell:0}} \Pr(y_{-\ell:0}) \bigotimes_{t=-\ell}^{0} |y_t\rangle\langle y_t| : \\ &\qquad \sum_{y_{0:\ell}} \Pr(y_{0:\ell}) \bigotimes_{t=0}^{\ell-1} |y_t\rangle\langle y_t| \Bigg) \\ &\geq \mathrm{I}\left[Y_{-\ell:0} : Y_{0:\ell}\right] \ . \end{aligned} \tag{44}$$

等号成立的条件同样由式 *(21)* 得出。恢复映射必须存在所针对的状态集合为 $\{|y\rangle\}$；这只有在所有 $|\psi_x\rangle \in Q$ 彼此正交，且 M 中包含每个 $|\psi_x\rangle$ 对应的投影算子时才可能成立。根据命题 *4* 中相同的论证，$|\Psi_{-\infty:\infty}\rangle$ 必须是可分过程，而 Q 必须由正交态构成。

取极限 $\ell \to \infty$，我们看到：

$$\lim_{\ell\to\infty} S(\rho_{-\ell:0}{:}\rho_{0:\ell}) \geq \lim_{\ell\to\infty} \mathrm{I}\left[Y_{-\ell:0} : Y_{0:\ell}\right]$$
$$\mathbf{E}_q \geq \mathbf{E}^Y .$$

结合上述证明，我们得到以下推论，它关联了底层经典过程 $\overleftrightarrow{X}$ 与测量过程 $\overleftrightarrow{Y}$的超额熵：

推论3。 设 $\overleftrightarrow{X}$是一个超额熵为 $\mathbf{E}^X$、字母表为 X 的经典过程；设 $|\Psi_{-\infty:\infty}\rangle = E(\overleftrightarrow{X})$ 是一个字母表为 Q 的可分量子位过程；并设 $\overleftrightarrow{Y}$是一个超额熵为 $\mathbf{E}^Y$的测量过程，使得 $\overleftrightarrow{Y} = M(|\Psi_{-\infty:\infty}\rangle)$，其中 M 是一个重复秩一*POVM*。

于是：

$$\mathbf{E}^X \geq \mathbf{E}^Y \ ,$$

当且仅当 Q 由 $|X|$ 正交态构成，且 M 使用的 *POVM* 中包含 Q 中每个元素对应的一个投影算子时，等号成立。

证明。 这可由命题*13*和*14*立即得出。

要得到$\mathbf{E}_q$ 的精确值，通常需要刻画无限长度的量子位序列。不过，我们可以利用公式 **(41)** 写出$\mathbf{E}_q$ 的有限-ℓ 估计：

$$\mathbf{E}_q(\ell) = S(\ell) - \ell s(\ell) \ , \tag{45}$$

而这通常会低估$\mathbf{E}_q$的真实值。

I. 量子瞬态信息

现在我们来考察量子块熵曲线如何收敛至其线性渐近线 $\mathbf{E}_q + s\ell$。我们将量子瞬态信息定义为：

$$\mathbf{T}_q \equiv -\mathcal{I}_0^q = \sum_{\ell=1}^{\infty} [\mathbf{E}_q + s\ell - S(\ell)] \ . \tag{46}$$

$\mathbf{T}_q$ 的单位为比特×时间步。

$\mathbf{T}_q$ 在图形上表示为$S(\ell)$曲线与其线性渐近线之间的面积，对 $\ell \to \infty$而言，如图 2 所示。我们将看到，$\mathbf{T}_q$ 能够区分某些周期性量子位过程，而先前的信息量（例如$\mathbf{E}_q$ 和 s）无法区分这些过程。

命题15. 量子瞬态信息$\mathbf{T}_q$ 可写为：

$$\mathbf{T}_q = \sum_{\ell=1}^{\infty} \ell\left[s(\ell) - s\right] \ .$$

证明。 该证明归结为参考文献 *[11]* 中针对经典随机过程瞬态信息的直接证明。它仅依赖于式 *(46)* 和式 *(A3)*；在量子情形下，这两个式子的形式与经典情形相同。

该表达式使我们能够估算给定量子过程的$\mathbf{T}_q$：

$$\mathbf{T}_q(\ell) = \sum_{m=1}^{\ell-1} m\left[s(m) - s(\ell)\right] \ , \tag{47}$$

这通常会低估$\mathbf{T}_q$的真实值。

$\mathbf{T}_q$ 与观察者同步到 HMCQS 所需的最少信息量有关。我们

将能够确定信源的内部状态的观察者称为已同步。如果 $S(\ell)$ 在有限的 ℓ 下收敛到其线性渐近线，那么存在一个作用于 $\rho_{0:\ell}$ （在 $\rho_{0:\ell}$ 的特征基中）的最优 POVM，能够准确确定 HMCQS 的内部状态。请注意，这不一定是重复的 POVM，甚至也不一定由局部 POVM 构成。该测量中有助于同步的信息量由 $\mathbf{T}_q$ 量化。相反，如果 $S(\ell)$ 在有限的 ℓ 下不收敛，那么对于任何有限块的量子比特，都不存在这样的 POVM；观察者至多只能渐近地收敛到信源的内部状态。我们将在第五节看到，当我们将测量限制为局域测量时，对于大多数信源而言，即使这种情况也不可能实现。

J. 量子马尔可夫阶

量子马尔可夫阶对应于下一量子位有条件依赖的先前量子位数量。若 R_q 是满足以下性质的最小值，则过程的量子马尔可夫阶为 R_q ：

$$S(\rho_0|\rho_{-R_q:0}) = S(\rho_0|\rho_{-\infty:0}).$$

图形解释是，当块大小达到 R_q时， $S(\ell)$趋于平稳并具有恒定斜率——即 s，如图 2 所示。因此， R_q 是使 $\Delta S(\ell)$和 $\Delta^2 S(\ell)$分别收敛到其渐近值的 ℓ 值，如图 3 和图 4 所示。

请注意，若 $R = 1$，则称经典过程为“马尔可夫”。如果一个可分量子位过程所依据的经典过程是马尔可夫过程，那么它满足公式（9）中的性质，但通常不具有 $R_q = 1$。

考虑一个可分量子过程 $|\Psi_{-\infty:\infty}\rangle = \mathcal{E}(\overleftrightarrow{X})$，其中 $\overleftrightarrow{X}$的马尔可夫阶数为 R^X ，而 $|\Psi_{-\infty:\infty}\rangle$ 的量子马尔可夫阶为 R_q。R_q 可以等于、小于或大于 R^X。下面我们将分别给出一个简单示例：

- $R_q = R^X$ 根据命题 3，若$\mathcal{Q}$ 由正交态组成，则结论显然成立，因为它们具有相同的块熵曲线。

- $R_q < R^X$ 若 $R^X > 0$ 且$\mathcal{X}$ 中的所有符号都映射到同一个纯态 $|\psi_x\rangle$。在这种情况下$R_q = 0$，所得过程是独立同分布的。

- $R_q > R^X$ 当 $R^X > 0$、 $|\mathcal{X}| = |\mathcal{Q}|$ 且 $\mathcal{Q}$ 由非正交态组成时。通常， $R_q = \infty$ 因为无法通过有限 POVM 可靠地区分任意的非正交态序列。

比较 R_q 与测量过程 $\overleftrightarrow{Y} = \mathcal{M}(|\Psi_{-\infty:\infty}\rangle)$的马尔可夫阶数 R^Y 时，也适用类似规则：

- $R_q = R^Y$ 若$|\Psi_{-\infty:\infty}\rangle$ 是一个具有正交字母表$\mathcal{Q}$的可分过程，且$\mathcal{M}$ 使用的 POVM 中包含针对$\mathcal{Q}$中每个元素的投影算子，则根据命题 4，结论成立。

- $R_q < R_Y$ 若$\mathcal{Q}$ 由正交态组成， $R_q = 1$且$\mathcal{M}$ 是一个重复秩一POVM，其中不包含投影到$\mathcal{Q}$中各状态上的投影算子。

- $R_q > R_Y$ 若 $R_q > 0$ 且$\mathcal{M}$ 是使用单元素 POVM 的重复POVM测量， $\mathbb{I}$。请注意，这不是秩一 POVM。

现在，我们已经掌握了一套量子信息性质工具，接下来一节将计算（或估算）这些性质在典型量子位过程中的取值。

IV. 量子位过程

现在，我们介绍多种可分离量子位过程，并大致按照结构复杂度递增的顺序加以组织，同时展示第三节的信息度量能够在多大程度上量化它们的行为。待掌握这些过程的信息度量后，我们将在文末附近的表 I 中总结其性质。

A. 独立同分布过程

回顾一下，独立同分布（独立且同分布）的量子位过程在任意量子位之间都不存在经典相关性或量子关联，其长度为 ℓ 的密度矩阵处于乘积态，即：

$$\rho_{0:\ell} = \bigotimes_{i=0}^{\ell-1} \rho_{iid} \ ,$$

其中， $\rho_{iid} \in \mathcal{H}$ 是单个时间步的密度矩阵。量子块熵的形式为 $S(\ell) = \ell S(\rho_{iid})$，因此冯·诺依曼熵率为 $s = S(\rho_{iid})$。这意味着存在一种重复的投影测量，使得测量过程 $\overleftrightarrow{Y}$具有经典熵率 $h_\mu = S(\rho_{iid})$。实现这一界限的测量由正交投影算子P_y构成，每个投影算子均由ρ_{iid}的一个特征向量构造而成。在 ρ_{iid} 为最大混合态的特殊情形下，$s = \log_2(d)$，而对块 $\rho_{0:\ell}$ 上的任意一组正交投影算子进行测量，都会使所有测量结果 $y_{0:\ell}$服从均匀分布。由于量子位之间不存在相关性$\mathbf{E}_q = 0$、$\mathbf{T}_q = 0$，以及 $R_q = 0$

这一点显然成立。具有非相关噪声的单量子位源的输出可视为一个独立同分布量子位过程。

B. 经典过程的量子表示

任何字母表为$\mathcal{X}$的经典过程，都可以用具有正交字母表 $\mathcal{Q}$ 的量子位过程来表示，其中每个符号 $x \in \mathcal{X}$ 对应一个纯态 $|\psi_x\rangle \in \mathcal{Q}$。此时，编码 $\mathcal{E}$ 是平凡的，并且可以通过使用正交投影算子$\{P_x = |\psi_x\rangle\langle\psi_x|\}$进行重复测量，重新获得底层过程 $\overleftrightarrow{X}$。这要求 $d \geq |\mathcal{X}|$。

底层过程、量子位过程和测量结果的信息度量满足：

$$\begin{aligned} \mathrm{H}\left[X_{0:\ell}\right] &= S(\rho_{0:\ell}) \\ &\leq \mathrm{H}\left[Y_{0:\ell}\right] \ , \end{aligned}$$

对于使用秩一正算子值测度进行的重复测量，若其测量元素包含投影算子集合$\{P_x\}$，则等号成立。
从这一关系可以看出，许多量子信息量，例如 s、$\mathbf{E}_q$、$\mathbf{T}_q$ 和 R_q，都等于 $\overleftrightarrow{X}$ 的经典性质。例外包括依赖于 d 与 $|\mathcal{X}|$ 之间关系的量，例如量子冗余。

由于量子-经典信道 $\mathcal{E}$ 是平凡的，输出过程 $\overleftrightarrow{Y}$是将 $\overleftrightarrow{X}$通过有噪经典信道后得到的结果，其中噪声水平取决于具体的测量方案 $\mathcal{M}$。

对于重复秩一POVM $\mathrm{H}\left[X_{0:\ell}\right] \leq \mathrm{H}\left[Y_{0:\ell}\right]$，对所有 ℓ，由命题 (3) 和 (4) 可得。类似的不等式明确关联了其他经典过程性质，例如熵率 ($h_\mu^X \leq h_\mu^Y$)、可预测性 ($|\mathbf{G}^X| \geq |\mathbf{G}^Y|$) 和超额熵（见推论 3）。

C. 周期过程

经典随机过程 $\overleftrightarrow{X}$ 若由模板序列的重复构成——即长度为 p 的符号块——则称其具有周期p。周期可分离量子位过程 $|\Psi_{-\infty:\infty}\rangle = \mathcal{E}(\overleftrightarrow{X})$ 是指其底层经典过程 $\overleftrightarrow{X}$具有周期性。

对于经典周期性过程，块熵曲线在其马尔可夫阶数 $R = p$ 处达到最大值，之后随着 ℓ ($h_\mu = 0$)的增加保持不变。该最大值就是超额熵，它完全由周期决定，公式为$\mathbf{E} = \log_2(p)$。最后，试图与不同长度的p 模板同步的观察者可能会遇到

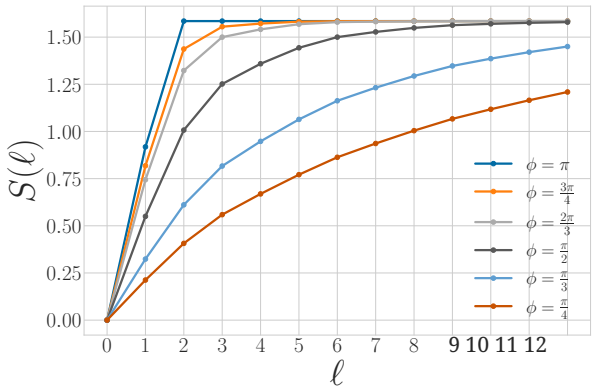


图 5. 周期过程发射状态 $|\psi_{00\phi}\rangle$ 时，量子块熵 $S(\ell)$ 随长度 ℓ 的变化，其中 ϕ取不同值。每条曲线都趋近于最大值 $\mathbf{E}_q = \log_2 3$。 ϕ 值越大，对应的字母表态越可区分，且 $\mathbf{T}_q$值越低。

或多或少的不确定性，具体取决于模板本身；这一特征由瞬态信息$\mathbf{T}$ [11]所体现。

周期性量子比特过程具有许多这类性质（正如我们通过第三节的结果所展示的），但也表现出更丰富的行为，因为它们可以由非正交态组成。图 5 展示了由重复量子字 $|\psi_{00\phi}\rangle = |0\rangle \otimes |0\rangle \otimes |\psi(\phi)\rangle$ 构成的周期 3 过程的量子块熵，其中$|\psi(\phi)\rangle = \cos\frac{\phi}{2}|0\rangle + \sin\frac{\phi}{2}|1\rangle$。当 $\phi = \pi$ 时，我们得到经典周期为 3 的字“001”的块熵。随着 ϕ 减小，$|\psi(\phi)\rangle$ 与 $|0\rangle$的可区分性也会降低。

由命题 6 和 13 可知，周期为 p 的周期性量子比特过程满足 $s = 0$，且 $\mathbf{E}_q \leq \log_2(p)$。（并且，$\mathbf{E}_q = \log_2(p)$，除非经典符号 $\mathcal{X}$ 中的两个符号被映射到 $\mathcal{Q}$ 中的同一纯态，从而使量子比特过程的有效周期降至小于 p。）然而，如果 $\mathcal{Q}$ 包含非正交态，那么量子块熵曲线不一定会在 $\ell = p$ 时达到其最大值 $\mathbf{E}_q$。此时，$R_q \to \infty$ 因为无论采用何种有限测量，观察者都无法明确分辨一个长度为 p的块从何处开始、到何处结束。

尽管所有周期为p 的量子位过程都具有相同的冯·诺依曼熵率和量子超额熵，但仍可通过两种不同方式下的量子瞬态信息值加以区分。

首先，不同的量子字母表 $\mathcal{Q}$ 会给出不同的$\mathbf{T}_q$值。图 5 表明，对于一个包含 2 个状态的量子比特字母表，随着这些状态变得越来越难以区分，量子瞬态信息会增加。对于$\phi = \pi$（正交字母表），$\mathbf{T}_q \approx 2.33$ 比特×时间步；而对于 $\phi = \frac{\pi}{2}$，$\mathbf{T}_q \approx 4.22$ 比特×时间步。这些

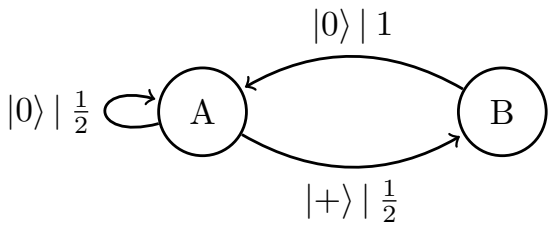


图 6. |0⟩-|+⟩ 量子黄金均值过程生成器。

$\mathbf{T}_q$ 值（以及本节中的其他值）均使用带有 $\ell = 12$ 的公式 (47) 进行数值近似。

其次，$\mathbf{T}_q$ 可以区分长度不同的p词。参考文献 [11] 表明，$\mathbf{T}$能够区分不同的周期为 5 的经典词（‘00001’、‘00011’和‘00101’），而$\mathbf{T}_q$ 将这一行为推广了。虽然在全局比特交换和移位下，所有周期为 3 的词都等价于‘001’，但周期为 5 的词并非如此。第五节将更详细地讨论同步到周期为 5 的量子位源，并将这一任务与$\mathbf{T}_q$的值联系起来。

D. 量子黄金均值过程

经典黄金均值过程由所有不含连续‘1’的二进制字符串组成。它是一个马尔可夫过程（$R = 1$），因为块的联合概率 $\Pr(X_{0:\ell})$ 按照式（2）分解，其中 $\Pr(0|0) = \frac{1}{2}$、$\Pr(1|0) = \frac{1}{2}$、$\Pr(0|1) = 1$，且 $\Pr(1|1) = 0$。对于经典黄金均值，每个符号含 $h_\mu = \frac{2}{3}$ 比特，且 $\mathbf{E} \approx 0.2516$ 比特。

将经典符号字母表替换为量子字母表 $\mathcal{Q} = \{|0\rangle \ |+\rangle\}$，其中 $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$，得到 |0⟩-|+⟩ 量子黄金均值过程，该过程在文献 [14] 中提出。图 6 展示了它的生成器。

我们还可以将这一过程进一步推广为 |0⟩-$|\psi(\phi)\rangle$量子黄金均值过程，其量子字母表为$\{|0\rangle \ |\psi(\phi)\rangle = \cos\frac{\phi}{2}|0\rangle + \sin\frac{\phi}{2}|1\rangle\}$。图 7 展示了该过程在不同 ϕ 值下的量子熵率。当 $\phi = \pi$, $|\psi(\phi)\rangle = |1\rangle$ 时，我们恢复经典黄金均值过程。随着 ϕ 减小至 0，$\mathcal{Q}$ 中的状态变得更难区分，且 s 减小，这与命题 6 的预期一致。

此外，在图7中，我们还看到了对量子黄金均值过程 |0⟩-$|\psi(\phi)\rangle$ 施加重复投影值测量后所得测量过程的熵率 h_μ^Y。该投影值测量由以角度 θ参数化的投影算子组成，其中 $M_\theta = \{|\psi(\theta)\rangle\langle\psi(\theta)|, |\psi(\theta+\pi)\rangle\langle\psi(\theta+\pi)|\}$。根据命题 7 的规定，对于所有 ϕ 和 θ，$h_\mu^Y \geq s$。

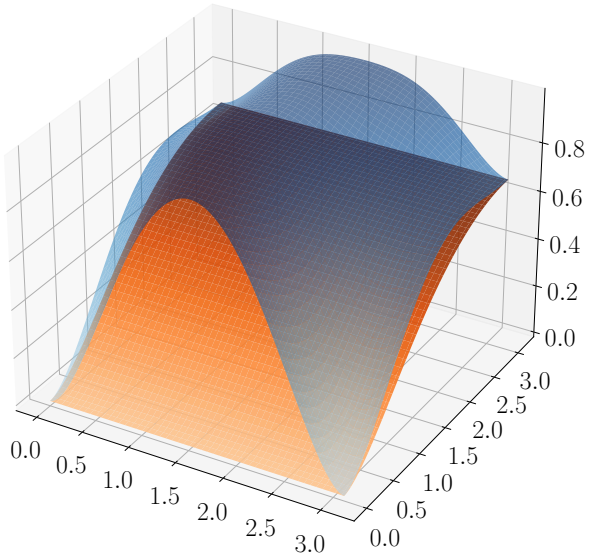


图 7. 对|0⟩-$|\phi\rangle$ 量子黄金均值过程使用重复投影值测量 M_θ 时的冯·诺依曼熵率 s（下方橙色曲面）和测量熵率 h_μ^Y（上方蓝色曲面）。随着 ϕ 增大，s 也随之增大，字母表变得更容易区分。当 $\phi = \pi$ 且 $\theta = 0, \pi$ 时，我们恢复经典黄金均值过程。对于 $(\theta - \phi) = \frac{\pi}{2}$，作用于 $|\phi\rangle$ 的 M_θ 是最大熵投影值测量（测量结果的分布为 50-50）。h_μ^Y 的极大值出现在此区域。

图8展示了第三节中的另一项量子信息性质——量子超额熵$\mathbf{E}_q$——与经典测量过程的超额熵$\mathbf{E}^Y$之间的关系。对于所有 ϕ 和 θ，命题14中的界（$\mathbf{E}_q \geq \mathbf{E}^Y$）成立；$\mathbf{E}^Y$ 在 $\phi = \pi$ 以及$\theta = 0, \pi$处达到最大值，此时 $\mathcal{Q} = \{|0\rangle \ |1\rangle\}$和 $M = M_{01}$。这些量是使用 $\ell = 10$ 估算得到的。

支撑 |0⟩-$|\psi(\phi)\rangle$ 量子黄金均值的是经典黄金均值过程，它是马尔可夫过程。因此，尽管对于大多数 ϕ 值，它的量子马尔可夫阶数为无限，它仍满足公式（9）中的量子马尔可夫性质。这对观察者同步到量子黄金均值源的能力有着重要影响，第五节将对此进行探讨。

E. 3 符号量子黄金均值

图9展示了三符号量子黄金均值过程的生成器，其字母表为$\mathcal{Q} = \{|0\rangle \ |1\rangle \ |+\rangle\}$。尽管其生成器与 |0⟩-|+⟩ 量子黄金均值具有相同的内部状态和转移概率，但三符号量子黄金均值的量子字母表$\mathcal{Q}$ 与生成器状态$\mathcal{S}$ （$|0\rangle \to A$ 和

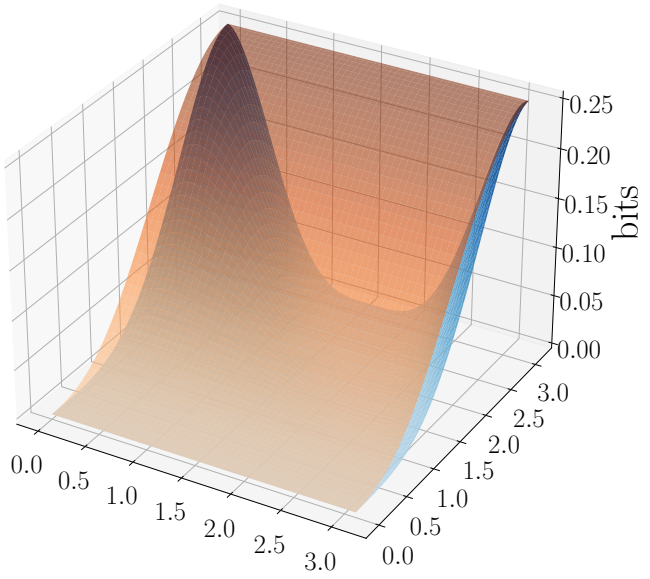


图 8. $|0\rangle$-$|\phi\rangle$ 量子黄金均值过程经重复投影值测量 M_θ测量得到的量子超额熵$\mathbf{E}_q$ （较高的橙色曲面）和测得的超额熵$\mathbf{E}^Y$ （较低的蓝色曲面）。随着 ϕ 增大，$\mathbf{E}_q$ 也随之增大，因为 $|0\rangle$ 和 $|\phi\rangle$ 变得更加可区分。$\mathbf{E}_Y$在 $(\theta - \phi) = 0, \pi$ 时达到最大，因为此时 M_θ 最能确定是否发出了 $|\phi\rangle$。

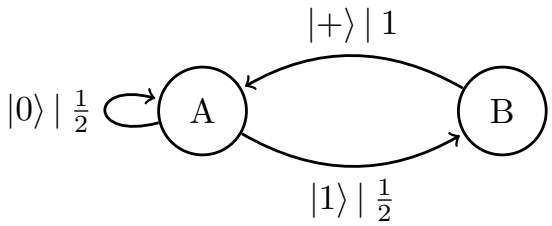


图 9. 三符号量子黄金均值过程生成器。

$|+\rangle \to B$）之间并非一一对应。相反，$|\mathcal{Q}| = 3$，这三个状态不可能两两正交，且 $d = 2$。

然而，与 $|0\rangle$-$|+\rangle$ 量子黄金均值不同，我们可以直接根据生成器计算量子熵率，因为（1）它具有量子单值性这一性质，并且（2）可以与之同步。第五节将对此进行详细讨论。目前，当且仅当对于每个 $\sigma \in \mathcal{S}$，都存在某个 POVM M_σ，使得知道σ 和 M_σ 的测量结果的观察者能够唯一确定 HMCQS 转移到的内部状态时，HMCQS 才是量子单值 的。图 9 中的生成器满足这一条件，其中 $M_A = M_{01}$。M_B 可以是任意 POVM。

对于经典单值隐马尔可夫链，h_μ 可以计算如下：

$$h_\mu = \sum_{\sigma_i, \sigma_j \in \mathcal{S}} \pi_i \sum_{x \in \mathcal{X}} T_{ij}^x \log_2 T_{ij}^x \ , \tag{48}$$

其中，π 为平稳态分布。该结果可追溯至信息论的奠基时期 [7]。

类似地，对于一个能够同步的量子单值 HMCQS，我们可以写为：

$$s = \sum_{\sigma_i, \sigma_j \in \mathcal{S}} \pi_i \sum_{x \in \mathcal{X}} T_{ij}^x \log_2 T_{ij}^x \ . \tag{49}$$

我们来逐步梳理三符号量子黄金均值中的这一逻辑。在状态 A ($\pi_A = \frac{2}{3}$) 下，生成器以相同的概率1 2发出一个处于状态 $|0\rangle$ 或 $|1\rangle$的量子比特。描述该量子比特的密度矩阵是一个最大混合态，对其进行的任何测量都包含 1 比特的不可约随机性。若处于状态B ($\pi_B = \frac{1}{3}$)，则它确定性地发出一个处于状态 $|+\rangle$ 的量子比特。因此，对各状态取平均后，冯·诺依曼熵率为 $s = \frac{2}{3}$ 比特/时间步。

尽管如此，如果仅限于重复使用秩一 POVM 进行测量，所观察到的经典过程的熵率无法达到2 3比特/符号的下界，因为状态为 A 时的最小熵基是 M_{01}，而状态为 B 时的最小熵基是 $M_\pm$。然而，使用图 10 中自适应测量协议的实验者，在同步后观察到的符号序列熵率将达到2 3 比特/符号。

他们首先在 $M_{01} = \{|0\rangle \ |1\rangle\}$ 基中进行测量，并根据初始测量的测量结果选择一个新基。如果测量结果是“0”，他们就能够与过程生成器同步，而该生成器必然处于状态 A。他们继续处于状态 A，使用 POVM M_{01} 进行测量，直到观察到“1”，此时他们转移到 B，并使用 $M_\pm$进行测量；这是一种零熵测量。他们观察到“+”，并返回状态 A。

这样，使用自适应测量方案时，测量序列的循环部分 $\overleftrightarrow{Y}_r$ 的熵率可能低于 h_μ^Y 对任何重复秩一 POVM 测量所能达到的熵率，即使像本例中那样，相关的 DQMP 仅使用秩一 PVM。这些想法将在下一节中加以形式化和拓展。

F. 单一分支与非单一分支量子比特源

接下来要讨论的两类量子比特过程，分别由图 11 和图 12 所示的单叶量子比特源和非单一化量子比特源生成

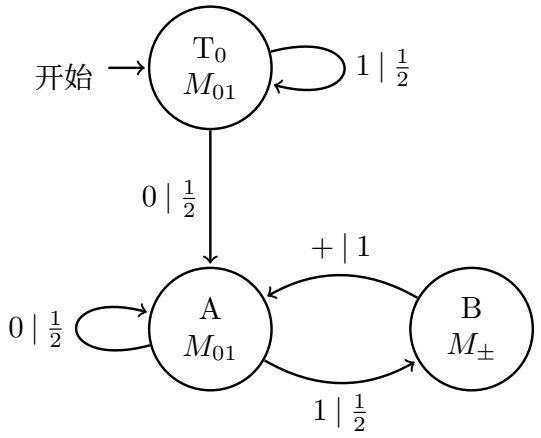


图 10. 三符号量子黄金均值过程的自适应测量协议（以 DQMP 的形式）。为实现同步，观察者从 T_0 （一个暂态）开始，并使用 M_{01}进行测量。恰好观测到 n 个“1”的概率为 $\frac{1}{2^n}$。观测到“0”后，观察者即实现同步。状态 A 和 B 分别与图 9 中源的内部状态 A 和 B 完全对应。“–”转移未显示，因为其概率为 0。该源是量子单值的，因此观察者在未来时刻仍将保持同步。

。这两个生成器都由两个内部状态（$\mathcal{S} = \{A, B\}$）组成，并发射四种可能的纯量子比特态（$\mathcal{Q} = \{|0\rangle \ |1\rangle \ |+\rangle \ |-\rangle\}$），这些态构成两组正交对。然而，对于单叶量子比特源的每个内部状态，两个可能发射的状态彼此正交，因此可以明确确定下一个内部状态。换言之，它具有量子单值性的性质。非单一化量子比特源则不具备这一性质。下一节将说明，这种差异如何显著影响观察者进行同步的能力。

通过改变参数 p，可以在此前已经分析过的若干个较简单过程之间进行插值。从图 11 中的单叶量子比特源开始，当$p = 0$时，生成器变为一个周期二源，发射词 $|\psi_{0+}\rangle$。当 $p = \frac{1}{2}$ 时，我们得到一个生成独立同分布最大混合过程的双状态模型。而当 $p \to 1$ 时，生成器会发射更长的字符串：根据源处于 A 还是 B，字符串中的量子比特要么全部为$|1\rangle$，要么全部为 $|-\rangle$。当 $p = 1$ 时，该过程变为非遍历的（此处表现为内部状态之间的断开），源将只以确定性的方式发射 $|1\rangle$ 或 $|-\rangle$。

类似地，图 12 中的非单一化量子比特源在某些 p下会简化。对于 $p = 0$，我们得到一个周期二源，这次其序列为 $|\psi_{01}\rangle$；而对于 $p = \frac{1}{2}$，它也会生成独立同分布最大混合过程。如–

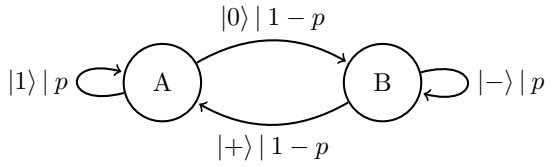


图 11. 单叶量子比特源。每个内部状态都会发射两个正交态之一，随后发生转移——例如， A 发射 $|0\rangle$ 或 $|1\rangle$，二者可通过测量 M_{01}加以区分——因此该源具有量子单值性。p 是一个取值范围为 0 到 1 的参数。其他过程对应特定的 p 值，例如，非正交周期-2 过程（$p = 0$）、最大混合独立同分布过程（$p = \frac{1}{2}$），以及由 $|1\rangle$ 或 $|-\rangle$构成的确定性序列（$p = 1$）。

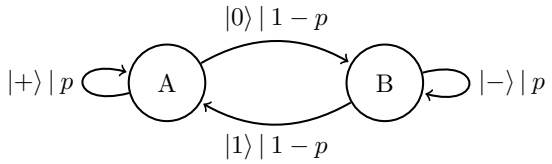


图 12. 非单一化量子比特源。每个内部状态都会发射两个非正交态之一，随后发生转移。观察者无法通过任何 POVM 确定源转移到了哪个状态。 p 的取值范围为 0 到 1。其他过程对应特定的 p 值，例如，正交周期 2 过程（$p = 0$）、最大混合独立同分布过程（$p = \frac{1}{2}$），以及由 $|1\rangle$ 或 $|-\rangle$ 构成的确定性序列（$p = 1$）。

$p \to 1$，它会发出全为 $|+\rangle$ 或全为 $|-\rangle$ 的长序列，并在 $p = 1$下变为非遍历。

G. 单一分支的三能级量子比特源

在超越量子比特过程的范围后，我们考虑一个三能级量子比特（$d = 3$）过程的单一示例，其生成器如图 13 所示，并采用五量子比特字母表 $\mathcal{Q} = \{|0\rangle \ |1\rangle \ |2\rangle \ |+\rangle \ |-\rangle\}$。使用更高维的希尔伯特空间，可为观察者提供更多用于同步的测量。因此，该示例过程展现出仅凭量子比特过程不可能实现的行为：希尔伯特空间的一个子空间（由 $|2\rangle$ 占据）始终可与 $\mathcal{Q}$ 中的所有其他状态区分开来，并可保留用于同步。另一个子空间（包括 $|0\rangle$、$|1\rangle$、$|+\rangle$和 $|-\rangle$）由无法可靠区分的状态组成。注意，该过程是量子单值的，但是

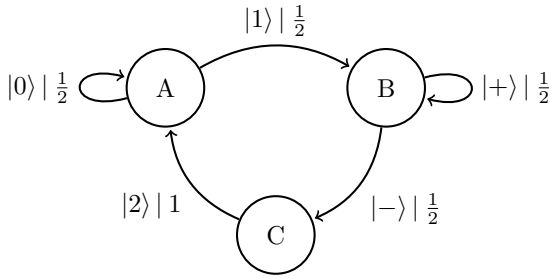


图 13. 单一分支量子三能级系统源。当处于内部状态A 和 B 时，它会发射一个位于希尔伯特空间中由 $|0\rangle$ 和 $|1\rangle$张成的子空间内的三能级量子比特。当处于 C 时，它会发射 $|2\rangle$，而且始终可以将其与$\mathcal{Q}$中的所有其他状态区分开来。这展示了在更高维希尔伯特空间中实现同步的其他可能性。

通过在单一基底中进行测量，无法保持同步。下一节将讨论多个能够实现这一点的自适应测量协议。

H. 讨论

表 I 总结了上述示例的信息性质，其中包括解析结果或数值估计。这些示例共同展示了可分离量子位过程的多种不同特征。它们说明，第三节中定义和刻画的信息性质，既能反映量子信息源的底层结构特征，也会受到过程中量子字母表内各状态可区分性的显著影响。因此，我们对量子块熵收敛到其线性渐近线的分析，为量化可分离量子过程中的随机性和相关性提供了有意义且易于解释的方法。

接下来，我们将讨论观察者面对一个发射可分量子过程的量子信息源时可能想要完成的两项任务。首先，如果观察者预先了解该源的内部结构，可能会想要确定它在给定时间步所处的内部状态。这项任务称为同步，我们将在下一节讨论。其次，如果观察者对该源一无所知，可能会想要测量它产生的过程，以推断其内部结构。这项任务称为系统辨识，我们将在第六节展示观察者如何通过层析协议来完成这项任务。

V. 与量子源同步

了解过程的量子生成器后，观察者如何确定其内部状态？当观察者确知其内部状态时，便与量子源同步 。接下来将探讨如何与量子过程同步——既包括观测如何帮助推断源的状态，也包括衡量部分同步和完全同步的定量指标。

与经典过程同步相比，量子测量使这项任务变得更为微妙。对于经典过程，若已知其最小单一化模型—— ϵ-机——便可完成同步，相关描述见参考文献 [12, 47]。

A. 已知状态

回顾一下，隐马尔可夫链量子源（HM-CQS）由一组内部状态（$\mathcal{S}$）、一个纯态字母表（$\mathcal{Q}$）和一组带标签转移矩阵（$\mathcal{T}$）构成。我们假设观察者完全了解生成某一过程的 HMCQS，但只能在时刻 $t = \ell$ 通过对区块进行测量 $\mathcal{M}_{0:\ell}$ 并观测到结果 $y_{0:\ell}$，来推断内部状态。他们无法接触在 $t = 0$之前发射的量子比特。

观察者根据不同观测序列对信源内部状态所作的最佳猜测，可以表示为信源内部状态上的分布，称为混合态（不要与由密度矩阵表示的混合量子态混淆）。经典地，观测到特定的长度为 ℓ 的词 $w = x_{0:\ell}$ 后，观察者处于混合态$\eta(w) = (p_A, p_B, \ldots)$。观测到下一个结果 X_ℓ 后，他们将根据测量结果转移到一组新的混合态之一： $\eta(w, x_\ell = 0)$、 $\eta(w, x_\ell = 1)$，依此类推。与新混合态对应的词，是 w 与新的观测结果 x_ℓ的串接。经典过程的混合态集合及其间的动态关系定义了过程的混合态表示（MSP），按构造它是单一分支的 [48]。

使用经典过程的混合态过程，而不是该过程的非单一分支生成器，具有许多计算优势。以下两点尤其值得关注：它使人们能够计算不具有有限单一分支表示的过程的熵率 [49] ，还能够计算观察者试图与过程的生成器同步时所经历的不确定性 [50]。

对于量子过程，不存在唯一的混合态过程，而是存在多个可能的混合态过程，每个都对应于一种不同的测量协议。对于给定的

	s	$\lvert\mathbf{G}_q\rvert$	$\mathbf{E}_q$	$\mathbf{T}_q$	R_q
	（比特/时间步）		（比特）	（比特×时间步）	（时间步）
独立同分布量子比特过程	$S(\rho_{iid})$	$1-S(\rho_{iid})$	0	0	0
周期 3 过程（$\phi=\pi$）	0	1	$\log_2(3)$	2.33	3
周期 3 过程 ($\phi=\pi/2$)	0	1	$\log_2(3)$	4.22	∞
量子黄金均值 ($\phi=\pi$)	2/3	1/3	0.2516	1/3	1
量子黄金均值 ($\phi=\pi/2$) 0.4495		0.5505	0.1092	0.5687	∞
三符号量子黄金均值 0.6667		0.3333	0.4652	0.8855	∞
单叶量子比特源 ($p=1/3$)	0.9184	0.0816	0.0808	0.1976	∞
非单一分支量子比特源（$p=1/3$）0.7306		0.2614	0.3217	0.4090	∞
单一分支量子三能级系统源	0.8002	0.7848	1.290	2.156	∞

表 I. 示例量子过程/源的信息性质。小数值是使用 $\ell=8$ 对单一分支量子三能级系统源、 $\ell=10$ 对单一分支与非单一分支量子比特源，以及 $\ell=12$ 对所有其他过程，通过方程 (38)、(45) 和 (47) 进行数值估算得到的。其他数值则通过解析计算得到。

源和给定的测量协议 $\mathcal{M}$ ，我们可以定义一组混合态，每个混合态都对应于可能观察到的测量序列。

考虑与长度为 ℓ 的观测序列相对应的混合态。我们将 $\mathcal{M}_{0:\ell}$ 限定为由局部 POVM 组成，并允许进行自适应测量。观察者施加测量 $\mathcal{M}_{0:\ell}$ 并得到测量结果 $y_{0:\ell}$ 后，其对生成器内部状态的最佳猜测由条件分布 $\eta(y_{0:\ell}|\mathcal{M}_{0:\ell}) = \{\Pr(\sigma|y_{0:\ell},\mathcal{M}_{0:\ell})$ 对所有 $\sigma\}$ 表示。

对于 $t=0$，观察者没有任何测量结果可用于预测信源的内部状态。不过，他们确实知道该模型的平稳态分布 π （因为可以直接从$\mathcal{T}$计算得到）。在没有任何测量的情况下，这可作为对信源内部状态的“最佳猜测”。如果初始状态分布$\mathcal{S}_0$ 不是 π 且观察者知晓这一事实，则该过程的混合态为η（$y_{0:\ell}|\mathcal{M}$，$\mathcal{S}_0$）。如果$\mathcal{S}_0=\pi$，则省略关于$\mathcal{S}_0$ 的条件。对于量子比特过程的大多数重复投影值测量而言，混合态的数量是不可数无穷的[14]。即使对于 $|\mathcal{S}|=2$，也会出现这种测量诱导复杂性 。本节中的示例暂时避开这种复杂性，转而关注那些信息足够充分的测量，使观察者能够通过有限的观测序列 $y_{0:\ell}$ 实现同步，并且仅产生有限个或（至多）可数无穷个循环混合态。

B. 平均态不确定性与同步信息

为了比较不同测量协议的同步行为，使用熵量会比直接处理混合态集合及其动力学更为方便。

具体而言，我们考察长度为-ℓ 的测量序列所对应的可能混合态的熵。

观察者在应用测量序列$\mathcal{M}_{0:\ell}$（根据某个协议$\mathcal{M}$）并观察到测量结果 $y_{0:\ell}$ 后，对信源的内部状态所具有的不确定性，由相应混合态分布的香农熵给出：

$$\begin{aligned}\mathrm{H}\left[\eta(y_{0:\ell}|\mathcal{M})\right] &= \mathrm{H}\left[\Pr(\sigma|y_{0:\ell},\mathcal{M})\right] \\ &= -\sum_{\sigma\in\mathcal{S}} \Pr(\sigma|y_{0:\ell},\mathcal{M}) \log_2 \Pr(\sigma|y_{0:\ell},\mathcal{M}) \ .\end{aligned}$$

对长度为 ℓ 个量子比特的区块上的所有测量结果取平均，即可得到平均态不确定性：

$$\begin{aligned}\mathcal{H}(\ell|\mathcal{M}) &\equiv \mathrm{H}\left[\eta(Y_{0:\ell}|\mathcal{M})\right] \\ &\equiv -\sum_{y_{0:\ell}} \Pr(y_{0:\ell}|\mathcal{M})\, \mathrm{H}\left[\eta(y_{0:\ell}|\mathcal{M})\right] \ .\end{aligned}$$

对于给定的测量协议$\mathcal{M}$，该量通常会在 $\ell\to\infty$ 极限下收敛到一个有限值。如果不具备遍历性，例如存在某种潜在的周期性，则该极限不一定存在。如果该极限存在，则渐近态不确定性为：

$$C_\infty(\mathcal{M}) = \lim_{\ell\to\infty} \mathcal{H}(\ell|\mathcal{M}) \ .$$

$\mathcal{H}(\ell|\mathcal{M})$ 和 $C_\infty(\mathcal{M})$ 均以比特为单位进行测量。
如果 $\mathrm{H}\left[\eta(y_{0:\ell}|\mathcal{M})\right]=0$，那么 $y_{0:\ell}$ 是一种同步观察，观察到测量结果 $y_{0:\ell}$ 的观察者便能准确识别信源的内部状态。如果 $\mathcal{H}(\ell|\mathcal{M})=0$，那么对 ℓ 个量子比特应用协议 $\mathcal{M}$ 时观察到的任何测量结果，都能让观察者与信源同步。对于经典和量子过程，马尔可夫

阶数是使 $\mathcal{H}(\ell)$ 变为零的 ℓ 的第一个取值（量子情形下，某个 $\mathcal{M}_{0:\ell}$ 也满足这一条件）。对于具有$\mathcal{Q}$中非正交态的 HMCQS，这通常不会发生，只有在 $\ell \to \infty$ 极限下例外，因为 R_q 通常是无限的。

测量 ℓ 个量子比特后实现同步，并不能保证观察者在测量量子比特 $\ell+1$后仍然保持同步。在会出现这种情况的经典过程中，持续同步要求观察者所使用的该过程的隐马尔可夫链满足单一性这一附加条件。在量子情形下，当且仅当底层 HMCQS 是量子单值的，并且当观察者处于内部状态 σ时，其测量协议确保执行相应测量，使得时刻 $t+1$的内部状态完全确定，同步才会持续。

这等价于以下表述：“存在某个协议$\mathcal{M}$，其测量为$\mathcal{M}_{0:\ell}$ 和 M_ℓ，使得：

$$\begin{aligned} &\mathrm{H}\left[\Pr(\sigma|y_{0:\ell},\mathcal{M}_{0:\ell})\right]=0 \\ &\quad\Longrightarrow \mathrm{H}\left[\Pr(\sigma|y_{0:\ell}y_\ell,\mathcal{M}_{0:\ell}\otimes M_\ell)\right]=0\ , \end{aligned} \tag{50}$$

对于所有 $y_\ell \in \mathcal{Y}$。”

当源处于一个内部状态（σ_i）时，维持同步所需的测量，不必与源处于另一个内部状态（σ_j）时维持同步所需的测量相同。当维持同步所需的测量取决于 HMCQS 的当前状态时，自适应测量协议即使在任何固定基测量都无法维持同步的情况下，也能够实现同步，下面的多个示例将对此进行说明。

最后，观察者使用给定的测量协议$\mathcal{M}$ 同步到某个过程时所经历的状态不确定性总量，称为同步信息，其定义如下：

$$\mathbf{S}(\mathcal{M})=\sum_{\ell=0}^{\infty}\mathcal{H}(\ell|\mathcal{M})\ . \tag{51}$$

请注意，如果渐近态不确定性 $C_\infty(\mathcal{M})$大于 0，则 $\mathbf{S}(\mathcal{M})$为无穷大。当 $C_\infty(\mathcal{M})=0$时，我们可以在某个有限的 ℓ处截断求和，以估算$\mathbf{S}(\mathcal{M})$。

对于经典过程，同步信息与瞬态信息$\mathbf{T}$密切相关。类似地，我们将$\mathbf{S}(\mathcal{M})$与量子瞬态信息$\mathbf{T}_q$ 联系起来，并将其用于比较不同测量协议下的同步。

本节其余部分将探讨与第 IV 节中的各种模型进行同步。我们将每种模型分别与多种测量方式搭配，包括重复的 POVM 和自适应协议，以展示这一选择如何影响$\mathcal{H}(\ell|\mathcal{M})$。

C. 同步到经典过程的量子表示

若 HMCQS 的量子比特字母表完全由正交态组成（对所有 $|\psi_x\rangle$ 和 $|\psi_{x'}\rangle\ \in\mathcal{Q}$，均有$\langle\psi_x|\psi_{x'}\rangle=0$），则它是底层经典过程 $\overleftrightarrow{X}$的量子表示。观察者可以使用正交投影算子 $P_x=|\psi_x\rangle\langle\psi_x|$ 进行测量（对所有 $|\psi_x\rangle\ \in\mathcal{Q}$ ），从而明确区分源发出的纯量子比特态。

同步到这类 HMCQS 的任务，等同于同步到一个发射经典符号的源。若它是量子单值的，那么其底层经典 HMC 就是单一分支的，且同步速度呈指数级（平均而言） [12, 47]。若不具备单一性，即使使用正交投影算子进行测量，观察者也可能无法实现渐近同步——即对于所有 $\mathcal{M}$，都有$C_\infty(\mathcal{M})>0$ 。

D. 同步到周期过程

所有周期性量子过程的冯·诺依曼熵率都为零——即 $s=0$——且每个状态都只有一个入射转移和一个出射转移。由于这些简化，要实现同步，观察者只需确定源的相位；也就是说，确定源在时间 t所处的 p 内部状态中的哪一个。对于任何 t，一旦确定了该相位，其他所有 t的相位也就确定了。初始相位不确定性等同于初始状态不确定性 $\mathrm{H}[\pi]=\log_2(p)$，因为 π 是均匀分布的。

任何能够区分$\mathcal{Q}$ 中状态的测量协议$\mathcal{M}$ 都能提供有关相位的信息，并使观察者得以渐近地与源同步。如果$\mathcal{Q}$ 中的两个状态无法通过$\mathcal{M}$区分，那么即使进行无限次测量，也可能无法实现同步。

对于经典周期性过程，观察者会在马尔可夫阶数 $\ell=p$ 处实现同步，且同步信息与瞬态信息相等：$\mathbf{S}=\mathbf{T}$[11]。相比之下，如果$\mathcal{Q}$中存在非正交态，周期性量子过程通常具有无限马尔可夫阶数，并且只能渐近同步。因此，$\mathbf{S}(\mathcal{M})\geq\mathbf{T}_q$。两者相等的条件是，$\mathcal{M}$ 是对该过程的最优测量。对于 $R_q=\infty$，这意味着$\mathcal{M}$ 必须是对双无限量子位链进行的全局测量。

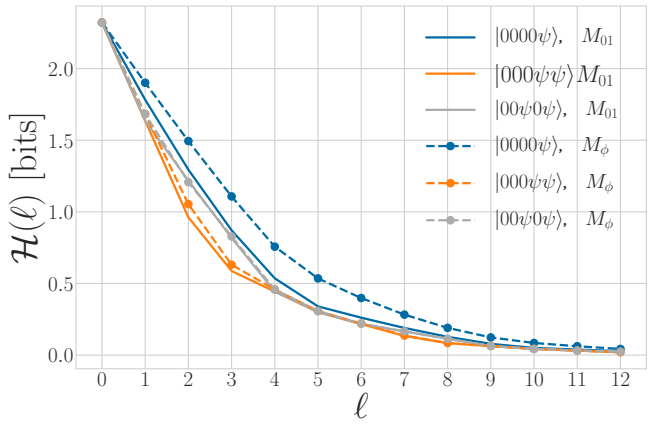


图14。周期为5的量子位过程中，平均态不确定性 $\mathcal{H}(\ell)$： ψ 表示状态 $|\psi(\phi)\rangle$。对应的投影值测量为 $M_\phi = \{|\psi(\phi)\rangle \langle\psi(\phi)|, |\psi(\phi+\pi)\rangle\langle\psi(\phi+\pi)|\}$。我们设定 $\phi = 3\pi/4$。每条曲线下的面积就是该过程和测量的同步信息**S** 。

图14展示了观察者测量三个不同的周期为5的量子比特过程时的平均态不确定性；这些过程由两个字母表态组成： $|0\rangle$ 和 $|\psi(\phi = 3\pi/4)\rangle$。该量子比特字母表中所有其他周期为5的序列，在平移和交换对称变换下都与这三种序列等价。正如预期的那样，对于序列和测量基底的所有组合，$\mathcal{H}(\ell)$ 会从 $\mathcal{H}(0) = \log_2 5$ 降至 $C_\infty = 0$ 。需要指出的是，同步发生的速率——以及观察者所见不确定性的总量，即同步信息——取决于具体的序列和对其反复施行的具体测量。

对于词为‘0000ψ’的周期为 5 的过程，使用式（47）和 $\ell = 12$，需要$\mathbf{T}_q \approx 6.32$比特×时间步。基底 M_{01} (M_ϕ) 测量时，我们发现$\mathbf{S}(M_{01}) \approx 7.92$比特（$\mathbf{S}(M_\phi) \approx 9.30$ 比特）。这些值也可通过计算至 $\ell = 12$ 来估算。类似地，对于其他周期为 5 的词，$\mathbf{S}(\mathcal{M})$ 与 $\mathbf{T}_q$ 之间的界也成立。对于词‘$000\psi\psi$’，$\mathbf{T}_q \approx 4.86$ 比特×时间步，$\mathbf{S}(M_{01}) \approx 6.84$ 比特，且$\mathbf{S}(M_\phi) \approx 7.05$ 比特。对于词‘$00\psi0\psi$’，$\mathbf{T}_q \approx 5.51$ 比特×时间步，$\mathbf{S}(M_{01}) \approx 7.39$ 比特，且$\mathbf{S}(M_\phi) \approx 7.47$ 比特。

此时我们希望强调，对于具有非正交字母表 $\mathcal{Q}$ 的一般周期-5过程， s、$\mathbf{E}_q$ 和 R_q 是相等的。尽管如此，$\mathbf{T}_q$ 和 $\mathbf{S}(\mathcal{M})$ 仍能揭示这些过程在同步任务中具有物理意义的差异。这些差异既内在于量子过程本身（$\mathbf{T}_q$），也体现在观察者获得的测量结果中（$\mathbf{S}(\mathcal{M})$）。

最后，我们对周期过程补充一点：尽管对于所有周期量子过程，$s = 0$，但许多测量协议（甚至是那些满足 $C_\infty = 0$ 的协议）都会产生

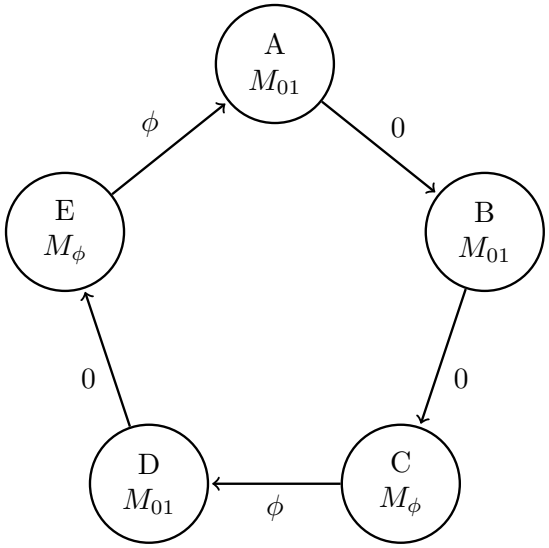


图 15. 针对周期-5序列 $|00\psi0\psi\rangle$ 的自适应测量协议。每个状态都标注了下一步要执行的测量： M_{01}，其可能的测量结果为‘0’和‘1’；或 M_ϕ，即由元素 $|\psi(\phi)\rangle\langle\psi(\phi)|$ 和 $|\psi(\phi+\pi)\rangle\langle\psi(\phi+\pi)|$ 构成的投影值测量（PVM），其可能的测量结果为 ϕ 和 $(\phi+\pi)$。图中所示的五个状态是该测量协议的循环状态，观察者只有在同步后才会遇到这些状态。同步后使用此协议的观察者会看到一个测得的周期-5过程，其词为‘$00\phi0\phi$’。

具有非零熵率h_μ^Y的测量所得经典过程。事实上，对周期过程进行固定基测量，只有在满足以下条件时，观测到的过程才会具有 $h_\mu^Y = 0$： **(i)** $\mathcal{Q}$ 中的所有状态彼此正交；**(ii)** 测量采用的正交基中，为$\mathcal{Q}$中的每个状态都包含一个投影算子。

相比之下，使用自适应测量协议的观察者在同步后可以轻易地使测量结果的不确定性为零。例如，图 15 展示了一个 DQMP 的循环状态；该 DQMP 会根据源的内部状态，在两种不同的测量——M_{01} 和 M_ϕ——之间切换。观察到的测量结果序列是确定性的，而循环测量过程是一个周期为 5 的过程，其词为‘$00\phi0\phi$’，且 $h_\mu^{Y_r} = 0$。

E. 使用投影值测量进行同步

量子黄金均值过程有多个同步观测；参见图 6 中的生成器。使用 M_{01} 进行测量并看到“1”，会使观察者同步到内部状态 B；使用 $M_\pm$ 进行测量并看到“–”，会使观察者同步到内部状态

A。让我们依次考虑这两种重复投影值测量所产生的混合态，从 M_{01}开始。

图16展示了通过反复应用 M_{01}得到的测量过程的混合态。观测到同步测量 y =‘1’意味着源刚刚转移到了状态 B，同时发射了一个 $|+\rangle$量子比特；也就是说，$\mathbf{H}[\eta(\text{‘1’})] = 0$。此外，处于状态 B 的源只能转移到状态 A，同时发射一个 $|0\rangle$量子比特，因此观察者将看到‘0’，且$\mathbf{H}[\eta(\text{‘10’})] = 0$。然而，一旦源处于状态 A，如果观察者又观测到一个‘0’，就很容易与源失去同步，因为这与源的两种转移都相符。随着观察者看到更多‘0’，他们会转移到图16中更靠右的混合态。总之，使用 M_{01} 进行测量利用了同步观测‘1’和‘10’，但其混合态过程有可数无穷多个循环状态，分别对应于由 n 个‘0’组成的序列。

这一测量过程是一个具有无限状态的经典更新过程，其信息性质可以使用参考文献 [51] 中的方法进行估计。在连续观测到 n 个“0”后，下一次观测为“1”的概率为：

$$\begin{aligned}\Pr(\text{‘1’}|\text{‘0’}^n) &= \tfrac{1}{4} * \Pr(A|\text{‘0’}^n) \\ &= \tfrac{3}{16}\left(1 - \left(-\tfrac{1}{3}\right)^n\right) \ .\end{aligned}$$

我们得到 $h_\mu^Y \approx 0.60$ 每符号比特数、$\mathbf{E}^Y \approx 0.053$ 比特，以及 $[1] \approx 0.65$ 比特的单符号熵 $\mathbf{H}$。$\mathbf{E}^Y$的较小值，以及 h_μ^Y 并未显著低于 $\mathbf{H}[\ell = 1]$ 这一事实表明，相较于使用 $\Pr(\text{“0”}) = 5/6$的偏置硬币，无限状态更新过程表示仅提供了很小的预测优势。$\Pr(B|\text{“0”}^n) = \frac{1}{4}\left(1 - (-1/3)^{n-1}\right)$ 以指数速度快速收敛至其渐近值 $\Pr(B|\text{“0”}^n) = 1/4$，这一点也进一步证明了上述结论。

使用 $M_\pm$ 进行测量时，符号“–”是一种同步观测，表明源处于状态 A。对于该测量过程，其递归混合态如图 17 所示，并且它们还构成一个经典更新过程。自上一个“–”以来观测到 n “+”时，生成器处于状态 A 的概率为 $\Pr(A|\text{“+”}^n) = \frac{3}{5}\left(1 - (-2/3)^{n+1}\right)$。使用 $M_\pm$ 获得的测量过程具有 $h_\mu^Y \approx 0.90$ 每符号比特数、$\mathbf{E}^Y \approx 0.020$ 比特和 $\mathbf{H}[\ell = 1] \approx 0.91$ 比特。同样，无限状态混合态过程相较于 $\Pr(\text{“+”}) = 2/3$ 的偏置硬币，仅提供很小的预测优势。此外，$\Pr(A|\text{“+”}^n)$ 会以指数速度快速收敛到 $\Pr(A|\text{“+”}^n) = 3/5$。

采用这两种固定基测量中的任一种，都会得到随 ℓ 单调降低的平均态不确定性，如图18所示。由于该源不是量子单值的，观察者在测量该过程时会反复同步和失去同步，而不论基如何，且当 $\ell \to \infty$ 时，$\mathcal{H}(\ell)$ 并不趋近于 0。我们发现，$C_\infty(M_{01}) \approx 0.62$ 比特，而 $C_\infty(M_\pm) \approx 0.54$比特。使用 $M_\pm$ 进行测量不仅在渐近意义下会产生更少的状态不确定性，而且对于 ℓ 的所有取值都会产生更低的平均不确定性。

图18还展示了另外两个相关投影值测量的平均态不确定性：M_θ 在 $\theta = \pi/4$ 和 $\theta = 3\pi/4$下的情况。回想一下，M_θ 是由投影算子构成的投影值测量，它们分别投影到正交态 $|\psi(\theta)\rangle$ 和 $|\psi(\theta + \pi)\rangle$上。作为比较，$M_{01}$ 对应 $\theta = 0$，而 $M_\pm$ 对应 $\theta = \pi/2$。

对于 $\theta = \pi/4$，符号态 $|0\rangle$ 和 $|+\rangle$ 向观察者提供完全相同的测量结果分布，因此观察者无法从平稳态分布之外获取有关源状态的任何信息。这也可从图19中渐近态不确定性的最大值看出。

对于 $\theta = 3\pi/4$——以及 θ的大多数取值——M_θ都没有同步观测。尽管如此，图18表明，对于所有 ℓ，$\mathcal{H}(\ell|M_{\theta=3\pi/4})$都低于能够实现精确同步的两个基。

对于一般的 θ，测量过程并不具有图16和图17所示的更新过程结构。相反，在缺乏同步观测的情况下，混合态的数量通常会随 ℓ呈指数增长。如果我们用长度为ℓ 的观测来近似混合态过程（MSP），那么将会有 $|\mathcal{Y}^\ell|$ 个混合态——每种可能的测量序列对应一个混合态——且每个混合态都对应信源内部状态上的一种不同分布。这些无限状态的混合态过程可用其统计复杂度维数 [52]来刻画。

尽管复杂度急剧增加，许多这类投影值测量所产生的平均状态不确定性仍低于 M_{01} 和 $M_\pm$。$\theta = 3\pi/4$ 是一个典型例子。应用 M_θ 时的渐近态不确定性如图19所示。请注意，如前所述，$C_\infty(\mathcal{M}_\theta) = \mathbf{H}[\pi]$ 在 $\theta = \pi/4$时达到最大值。这组投影值测量中的最小渐近态不确定性为 $C_\infty \approx 0.49$ 比特，在 $\theta \approx 2.01$时出现。观察者可以选择使用具有指数数量混合态的一组基，以降低对信源内部状态的不确定性，但代价是必须追踪更大数量混合态的概率。

F. 通过自适应测量维持同步

自适应测量协议也能够在任何固定基测量都无法实现同步时维持同步。为了理解这一点，请看图 11 中的单叶量子比特源。

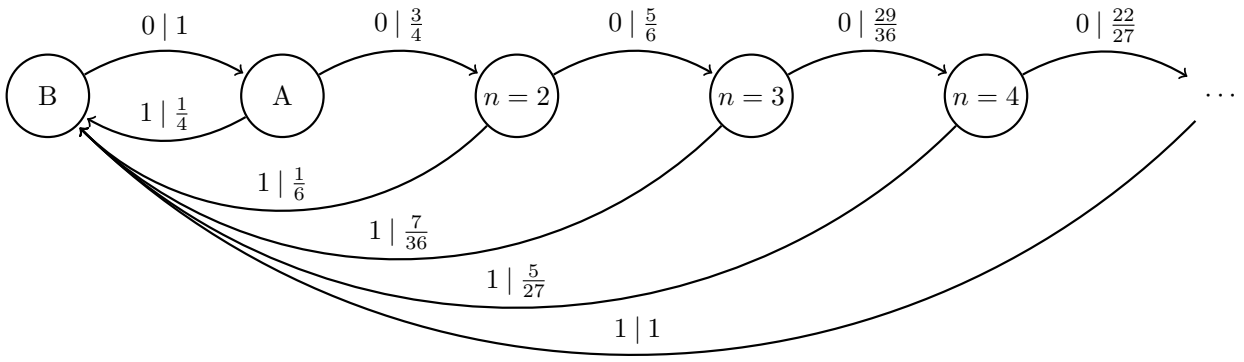


图 16. 使用 M_{01} 测量的 $|0\rangle$-$|+\rangle$ 量子黄金均值过程的混合态表示。 n 指自最近一次出现 '1' 以来连续出现的 '0' 的数量。

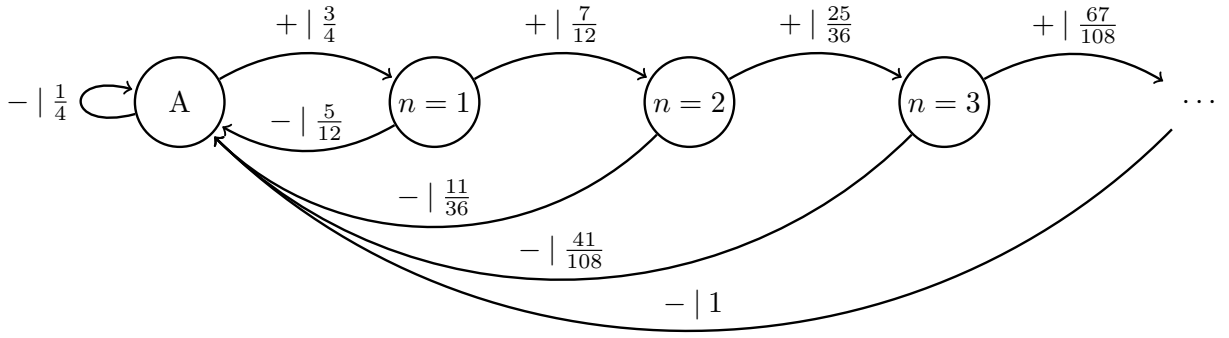


图 17. 使用 $M_{\pm}$ 测量的 $|0\rangle$-$|+\rangle$ 量子黄金均值过程的混合态表示。 n 指数量自最近一次出现 '−' 以来连续出现的 '+'。

对于 $0 < p < 1$，不存在同步观测；然而，几乎任何测量都能将平均态不确定性降至 $\mathrm{H}[\pi]$ 以下。此外，如果观测者通过其他方式得知源的状态（也许源在 $t = 0$ 时刻初始化于状态 A），那么他们就能够通过一个简单的自适应测量协议 $\mathcal{M}^{adaptive}$ 维持同步。

此协议仅针对循环状态定义，具体如下：若源处于状态 A，则施加测量M_{01}；若源处于状态 B，则施加测量 $M_{\pm}$。只有因为该源是量子单值的，才可能如此简单地定义协议并维持同步。

图 20 展示了观察者与以状态 A 初始化的源逐渐失去同步时$\mathcal{H}(\ell)$ 的变化。对于每次重复投影值测量，它最终会达到一个渐近值，不过$\mathcal{H}(\ell)$ 可能并非单调变化。

G. 与三能级量子比特源同步

量子比特源（$d = 2$）在无歧义状态判别和同步方面的机会有限。三能级量子比特允许更一般的行为，并已被建议用于量子通信网络，其中专门保留一种状态用于同步 [53]。

让我们将这一逻辑应用于图 13 中的单一分支量子三能级系统源。该过程的同步观测序列为"2"、"1+"和"1−"，它们分别能确定地使源处于状态 A、 B 和 C。同步之后

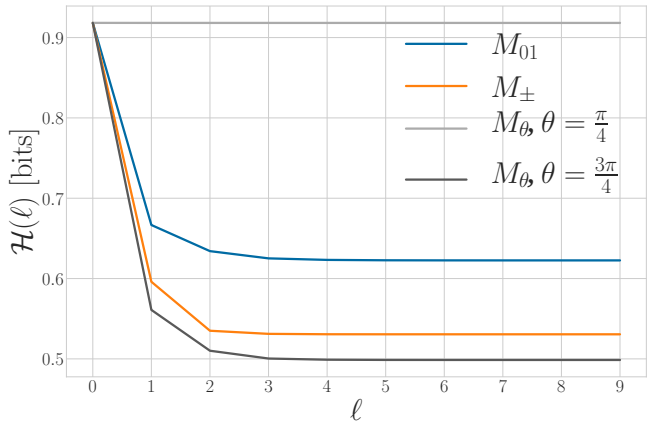


图18. 量子黄金均值生成器经过 ℓ 次测量后的 $|0\rangle$-$|+\rangle$ 平均态不确定性 $\mathcal{H}(\ell)$。

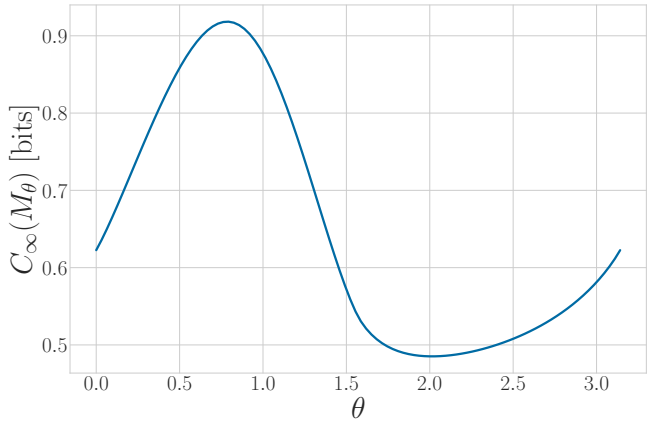


图19. 对$|0\rangle$-$|+\rangle$ 量子黄金均值应用投影值测量 M_θ 时的渐近态不确定性 $|0\rangle$。

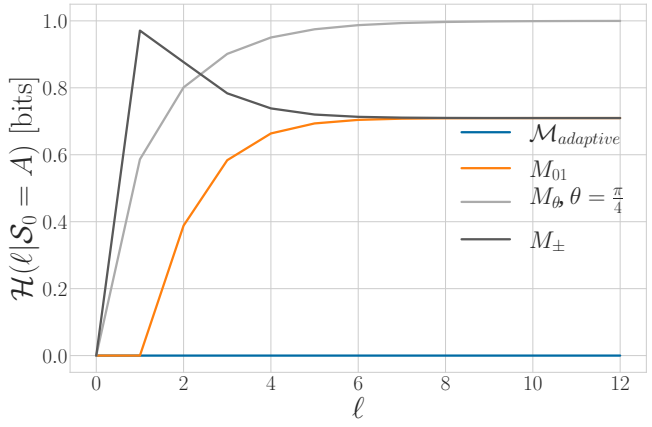


图20. 单一分支的量子比特源（$p = 0.6$）在状态 A 初始化时的平均态不确定性 $\mathcal{H}(\ell)$。只有自适应测量协议（文中所述）能够维持同步。

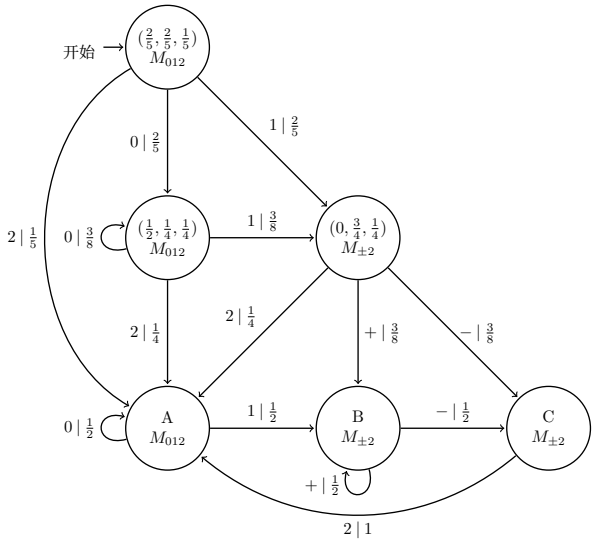


图21. 为图 13 中的量子三能级过程生成器定义的自适应测量协议。三个暂态混合态以内部源状态概率（p_A，p_B，p_C）标记，三个循环状态则与这些状态完全对应。由于该源是量子单值的，这个自适应协议能够永久与源同步。省略概率为零的转移。

观察者可以在状态 A 时通过测量$M_{012} = \{|0\rangle\langle 0|,\ |1\rangle\langle 1|,\ |2\rangle\langle 2|\}$保持同步，在状态 B 时通过测量 $M_{\pm 2} = \{|+\rangle\langle +|,\ |-\rangle\langle -|,\ |2\rangle\langle 2|\}$ 保持同步，而在状态 C 时则可采用上述任一种测量方式。

我们使用五种不同的测量协议探索如何与该源同步，并在图22中进行比较。前两种是在 M_{012} 基和 $M_{\pm 2}$ 基上重复投影值测量。其余三种是自适应测量协议，它们共享相同的再现动力学，但具有不同的暂态状态。考虑在 M_{012} ($M_{\pm 2}$) 基上测量，直到观测到 2，从而与源状态 A 同步。然后，当源处于状态 A 时使用 M_{012}，当源处于状态 B 或 C 时使用 $M_{\pm 2}$。我们将此协议称为 $\mathcal{M}_{012,\ sync}(\mathcal{M}_{\pm 2,\ sync})$。第五种协议 $\mathcal{M}_{adaptive}$ 由图21中的 DQMP 定义，它在刚定义的循环状态协议之外，还对三个暂态混合态采用自适应协议。

观察者实施这五种测量协议时的平均状态不确定性如图22所示。固定基测量无法实现持续同步，且渐近态不确定性非零（$C_\infty(M_{012}) \approx 0.40$ 比特和$C_\infty(M_{\pm 2}) \approx 0.72$ 比特）。如果先在固定基下测量，直到实现同步（观察到一个 2），然后利用该源的量子单值性保持同步，那么渐近态不确定性便会消

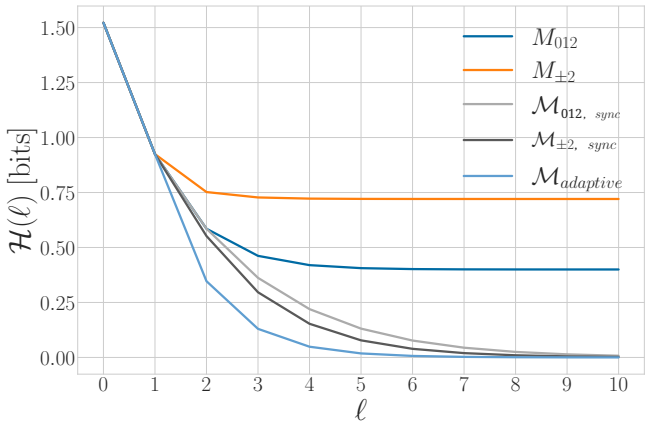


图22。测量由单一分支的三能级量子系统源生成的过程时的平均态不确定性 $\mathcal{H}(\ell)$。 M_{012} 和 $M_{\pm 2}$ 是固定基测量。$\mathcal{M}_{012,\ sync}$ 和 $\mathcal{M}_{\pm 2,\ sync}$在固定基上进行测量，直到观测到 2，此后始终保持同步。$\mathcal{M}^{adaptive}$ 指的是图 21 中的协议。

失。对于$\mathcal{M}_{012,\ sync}$和$\mathcal{M}_{\pm 2,\ sync}$，情况正是如此；它们的同步信息分别为$\mathbf{S}(\mathcal{M}_{012,\ sync})$ ≈3.91 比特和 $\mathbf{S}(\mathcal{M}_{012,\ sync}) \approx 3.60$ 比特。测量协议$\mathcal{M}^{adaptive}$ 的额外复杂性使其能够使用更多同步词（“1+”和“1−”），并获得低于“等待观察到一个‘2’”这一简单策略的同步信息（$\mathbf{S}(\mathcal{M}^{adaptive}) \approx 3.00$比特）。请注意，这三个同步信息值都高于我们对该过程的量子瞬态信息的估计值（$\mathbf{T}_q \approx 2.16$ 比特 × 个符号）。这些$\mathbf{S}(\mathcal{M})$值是使用 $\ell = 10$估算的。

H. 讨论

本节详细介绍了如何与已知的量子多能级系统源同步。与经典随机变量源不同，观察者可以尝试使用各种测量协议与量子多能级系统源同步。我们可以使用上文引入的两个信息量来比较不同协议：渐近态不确定性 $C_\infty(\mathcal{M})$和同步信息 $\mathbf{S}(\mathcal{M})$。

对于并非量子单值的源，没有任何协议能够与源保持同步。然而，对于不同的测量协议 $\mathcal{M}_0$ 和$\mathcal{M}_1$，我们可以比较 $C_\infty(\mathcal{M}_0)$ 和 $C_\infty(\mathcal{M}_1)$。数值较低的协议在同步任务中表现更好，因为观察者对信源内部状态的不确定性平均而言会更低。

这自然引出了一个问题：用于与源同步的最佳测量协议是什么？为回答这个问题，我们引入量子位过程的一种与协议无关的新性质——最小渐近状态不确定性：

$$C_{min} = \min_{\mathcal{M}} C_\infty(\mathcal{M}) \ ,$$

其中的最小值取遍所有通过 DQMP 定义的可能测量协议。实际上，对于一个量子过程，确定 C_{min} 需要证明不存在任何测量协议能够达到更低的值。全面探索测量协议的空间超出了本文的范围。尽管如此，我们在上述示例中提出了 C_{min} 的候选方案，并通过数值方法找到了受限类别的重复投影值测量的最小渐近态不确定性；回顾图19。

对于属于量子单值的源，我们探索了几种能够实现持续同步的协议。对于此类过程，$C_{min} = 0$比特。我们可以通过同步信息 $\mathcal{M}_0$ 和 $\mathcal{M}_1$ 来比较不同的同步测量协议 $\mathbf{S}(\mathcal{M}_0)$ 和 $\mathbf{S}(\mathcal{M}_1)$。数值越低，意味着观察者在同步时经历的状态不确定性越小。观察者所能经历的最小状态不确定性是多少？我们将量子过程的最小同步信息定义为：

$$\mathbf{S}_{min} = \min_{\mathcal{M}} \mathbf{S}(\mathcal{M}) \ .$$

这个最小值也取遍所有 DQMP。要证明某个协议是最小的并非易事。

这项工作引出了一个尚未解决的问题：“是否总能与量子单值源同步？”或者等价地，“对于量子单值源生成的所有过程，$C_{min} = 0$ 是否成立？”回答这个问题也需要更深入地理解 DQMP 的空间。相关进展可能包括证明存在或不存在一种能够与图 11 中的单叶量子比特源同步的协议。

最后，我们列举若干原因，说明同步不仅是确定信源内部状态的重要任务，也有助于改进对未来测量结果的预测。一般而言，HMCQS 具有内在的随机性。周期源则是例外。通过同步，我们可以大幅降低测量结果的不确定性。在极端情况下，若信源的内部状态只有一种可能的转移，这种不确定性便会消失，我们可以使用投影值测量测量下一个量子位，并得到确定性的结果。

例如，如果我们知道一个 $|0\rangle$-$|+\rangle$ 量子黄金均值生成器处于状态 B，那么下一个量子位就是 $|0\rangle$；如果我们应用测量 M_{01}，则总会看到“0”。因此，

我们将同步视为一种动态推断形式：观察者利用对源的内部结构以及一系列测量结果的了解，从而更准确地预测未来的测量结果。每种混合态都对应一种不同的预测。未来符号预测的状态，也可以针对具有一定数量状态的机器拓扑进行推断——这称为贝叶斯结构推断 [54]——以确定该源最可能的 ϵ 机器。

VI. 量子过程系统辨识

同步是观察者拥有量子信息源的准确模型（此处为 HMCQS）后执行的一项任务。接下来自然会问：观察者如何为一个未知的 量子信息源建立准确模型？本节将着手回答这个问题。首先简要回顾如何从数据推断经典信息源，然后讨论如何通过量子态层析辨识单个量子位的状态。将这两种思路结合起来，我们便得到一种针对平稳量子比特序列的推断方法。

A. 经典系统辨识

未知平稳经典过程的观察者只能利用现有数据——符号字母表$\mathcal{X}$上各词的分布——来推断源的结构。我们将长度为ℓ 的词的分布记作$\mathcal{P}(\ell) = \{w = x_0 \cdots x_{\ell-1} | w \in \mathcal{X}^\ell\}$。

对于 $\ell = 1$ ，我们得到$\mathcal{X}$上各符号的分布。利用$\mathcal{P}(1)$，我们可以重建一个无记忆（独立同分布）源模型，该模型只有一个内部状态，且每个 $\Pr(x \in \mathcal{X})$ 都可直接从$\mathcal{P}(1)$得到。

对于 $\ell = 2$ ，我们使用$\mathcal{P}(2)$来创建下一个符号在给定前一个符号条件下的条件概率分布；即，对于所有 x_0，$x_1 \in \mathcal{X}$，$\Pr(x_1|x_0)$。根据这些条件分布，我们可以构造该源的马尔可夫近似，其中包含 $|\mathcal{X}|$ 个状态，每个符号对应一个状态。这些条件分布确定了该源马尔可夫近似的转移矩阵：即，$\Pr(x_1|x_0) = T_{x_0,\ x_1}$。

同样，对于 $\ell > 2$ ，我们通过以长度为 $\ell - 1$ 的词为条件，得到该源的长度为 ℓ 的 HMM 近似；每个词都对应一个不同的内部状态。在这一简化图景中，模型的内部状态数量随 ℓ 呈指数增长，因为长度为 ℓ 的词共有 $|\mathcal{X}|^\ell$ 个。这会在推断具有长时间周期相关性的过程时引发数值问题。尽管如此，仍有一些方法可以合并反映相同

B. 量子位的层析

我们不能直接将根据词分布推断经典源动力学的程序应用于量子位过程，因为观测取决于所使用的测量基/协议。要全面表征平稳量子源，我们必须改为在不同基下进行多次测量，以重构量子位密度矩阵。这项任务称为量子态层析。我们首先推断单个量子位状态 ρ_0，然后介绍一种使用可分量子位序列进行量子系统辨识的通用方法。

通过测量对单个未知量子位密度矩阵 ρ_0 进行层析重建，主要面临两个挑战：

1. ρ_0 的完整描述需要的参数数量会随状态的希尔伯特空间维数呈指数增长。2. 量子测量具有概率性，因此必须制备并测量 ρ_0 的许多副本，才能估计单个参数。

特定的测量组合对这项任务尤其有用。例如，可以通过一组互为无偏基（MUB） [3] 进行测量，或使用单个信息完备 POVM（IC-POVM） [4]来推断 ρ_0 。对于量子比特，一种可能的互补无偏基由 x、y 和 z 基组成。分别在这三个基中测量许多量子比特副本，可以得到三个概率——$\Pr(+x)$、$\Pr(+y)$ 和 $\Pr(+z)$——它们能够唯一确定密度矩阵 ρ_0。对于一般的量子位层析，我们不讨论将这些参数确定到所需容差内所需的测量次数；这是一个已有充分研究的问题[33, 55]。

量子比特的一种信息完备 POVM 示例由投影到以下状态的投影算子构成：

$$
\begin{aligned}
|\phi_1\rangle &= |0\rangle\,, \\
|\phi_2\rangle &= \frac{1}{\sqrt{3}}\,|0\rangle + \sqrt{\frac{2}{3}}\,|1\rangle\,, \\
|\phi_3\rangle &= \frac{1}{\sqrt{3}}\,|0\rangle + \sqrt{\frac{2}{3}} e^{i2\pi/3}\,|1\rangle\,, \quad \text{and} \\
|\phi_4\rangle &= \frac{1}{\sqrt{3}}\,|0\rangle + \sqrt{\frac{2}{3}} e^{i4\pi/3}\,|1\rangle\ . \qquad (52)
\end{aligned}
$$

这也是一种对称信息完备 POVM，即 SIC-POVM，因为任意两个投影算子的组合都具有相同的内积。

通过使用相同的信息完备 POVM 测量 ρ_0 的许多相同副本，可以得到涵盖d^2 种可能测量结果的概率分布。由此可获得 $d^2 - 1$个参数（由于归一化），这些参数可以唯一确定密度矩阵 ρ_0。

高维希尔伯特空间中 SIC-POVM 的存在及其性质，仍是一个活跃的研究领域[56]。

C. 量子比特过程的层析

既然我们已经知道如何估计单个量子位的密度矩阵，就可以开始分析长度为ℓ 的密度矩阵 $\rho_{0:\ell}$。我们将依次测量每个量子位，而不是对整个长度为ℓ 的量子位组块进行联合测量。我们之所以能够这样做，是因为我们关注的是可分量子 dit 序列——纠缠量子 dits 过程的一般情形则需要在非局域基中进行测量。

对于 $\ell = 1$，我们按照上述方法重构 ρ_0，并根据式（4）得到一个关于每个时间步发射量子位 ρ_0 的源的无记忆（独立同分布）估计。除非 ρ_0是纯态，否则会有许多单态 HMCQS 生成这一过程，因为许多不同的纯态系综对应于同一个密度矩阵。对 ρ_0 进行对角化，并令源以相应本征值 λ_i为概率发射每个纯本征态 $|\psi_i\rangle$，即可得到该源唯一的无记忆模型。

对于 $\ell = 2$，我们必须重构双量子位密度矩阵 $\rho_{0:2}$。（回想一下，我们的索引左端包含、右端不包含，因此 $\rho_{0:2}$ 是 $t = 0$，1 个量子位的联合状态。）由于平稳性，$\rho_{0:2}$ （两个量子位的联合状态）必须与单量子位边缘分布一致，即 $tr_0(\rho_{0:2}) = \rho_1 = \rho_0 = tr_1(\rho_{0:2})$。

我们将详细介绍重构 $\rho_{0:\ell}$ 量子比特 的迭代过程。当 $d = 2$, $\rho_{0:2}$ 有 15 个实参数需要通过层析确定。对单量子比特边缘分布进行层析可确定 3 个参数，而平稳性条件可确定另外 3 个参数。通过考虑一组互为无偏测量的组合，可以完整重构该状态。对于两个量子比特，这意味着泡利矩阵的 16 种组合（$\sigma_{I_0} \otimes \sigma_{x_1}$, $\sigma_{x_0} \otimes \sigma_{x_1}$, $\sigma_{x_0} \otimes \sigma_{y_1}$，等等。）[57]。要完整表征 $\rho_{0:2}$ 9，必须确定这些值——不涉及恒等算符 σ_{I_0} 和 σ_{I_1} 的值由单量子比特边缘分布确定。对于 $d > 2$，可以对这一过程进行修改

在更高维的希尔伯特空间中使用一组互为无偏基即可。

确定 $\rho_{0:2}$ 后，我们可以继续确定 $\rho_{0:3}$ （对于量子比特而言有 63 个实参数）。这些参数中有许多已由先前对单量子比特边缘分布（3 个参数）、双量子比特边缘态（15 个参数）的层析及其平稳性条件（分别为 6 个和 15 个参数）确定。三个单点泡利矩阵的组合足以完成完整重构。对于更大的ℓ，也可以继续这一过程，通常直到测量数量在实验上变得不可行。

D. 独立同分布的代价

人们通常假设量子信息源是独立同分布的 [6, 33]。如果进行量子态层析的观察者假设未知量子信息源是独立同分布的，那么他们就不会进一步探究，而只会确定单量子位边缘态 ρ_0。如果量子位之间存在相关性，就会导致对源的随机性估计过高。他们会错误地认为每个量子位都处于状态 ρ_0，并且熵率为 S（ρ_0）每个时间步的比特数。后者对真实熵率的估计高出 S（1）倍。观察者可以按照上述步骤，对长度为 ℓ 的量子位块进行层析重构，从而更准确地估计熵率和其他信息量。

为了说明这一假设在多大程度上可能误导观察者，请考虑图 12 中非单一化量子比特源生成的过程。假设源以其平稳态分布开始：($p_A = 1/2$, $p_B = 1/2$)。对于任意满足 $0 \leq p < 1$ 的 p， ρ_0 都是最大混合态，而假设过程为独立同分布的观察者会估计每个时间步有 $s = 1$ 比特。这仅在 $p = 1/2$ 时才是对该源的准确描述。然而，对于许多 p 值，该源中后续量子比特之间存在显著相关性。

让我们仔细看看 p 的极端值。对于$p = 0$，该过程的周期为 2，每个时间步有 $s = 0$ 比特，并且 $\mathbf{E}_q = 1$ 比特。这种模式很容易通过在 $\{|0\rangle \ |1\rangle\}$ 基底中进行测量来检测：测得 0（1）会立即使观察者同步到状态 B（A）。随着 $p \to 1$，这两个源状态变得越来越互不相连， $s \to 0$, 且 $\mathbf{E}_q \to 1$。在 $\{|+\rangle \ |-\rangle\}$ 基底中测量的观察者发现，测得一个 $+$（$-$）后，很可能会再测得一个 $+$ （$-$）。当只考虑单量子比特密度矩阵 ρ_0 时，这个源在 p 取不同值时丰富多样的行为将完全无法得到体现。

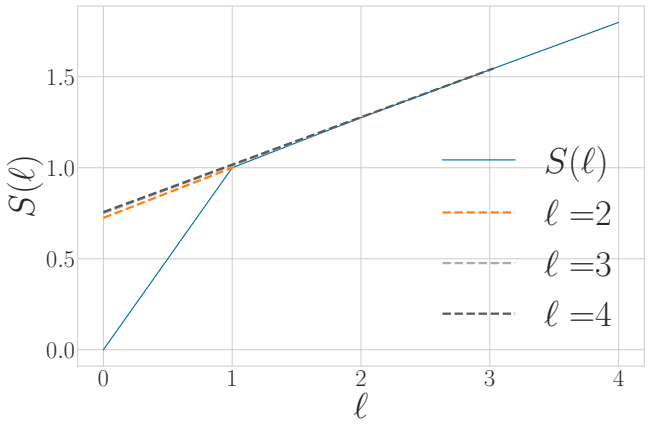


图 23. 图 12 中具有 $p = 0.05$ 的非单一化量子比特源所生成过程的信息性质的块熵和长度为 ℓ 的估计。线性估计的斜率和 y 截距分别为 s 和 $\mathbf{E}_q$。请注意，当 $\ell > 2$ 时，估计结果并未显著改善，这表明双量子比特关联对于确定过程的信息性质最为重要。

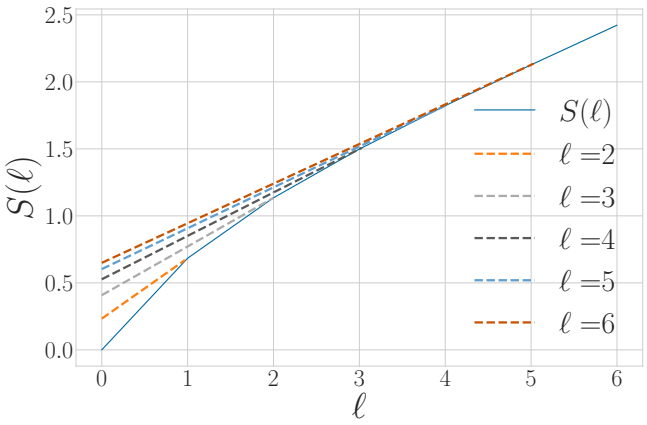


图 24. 对于图 11 中具有 $p = 0.05$ 的单一分支量子比特源所生成过程的信息性质，块熵和长度-ℓ 估计值。线性估计的斜率和 y-截距分别为 s 和 $\mathbf{E}_q$。请注意，随着 ℓ 增大，估计值持续改善，这表明存在长程关联。

E. 信息性质的有限长度估计

我们刚刚看到了一个例子：如果观察者假设一个源是独立同分布的，他们通常会低估量子过程的结构和相关性，并高估其熵率。如果他们只对有限长度、最长为 ℓ 的量子位块进行层析，情况也一样。现在考虑一位实验者，他通过层析重构密度矩阵 $\rho_{0:\ell}$，随后假设量子比特过程中不存在其他（更长程的）相关性。

其对量子熵率、量子超额熵和量子瞬态信息的估计分别由式（38）、（45）和（47）给出。

信息性质的长度为 ℓ 的估计值与其真实值之间的差异，取决于过程的内部结构和量子字母表。图 23 和图 24 中非单一化与单一分支量子比特源的估计分别代表了这方面的两个极端。如前所述，非单一化量子比特源的单量子比特密度矩阵为最大混合态。若观察者改为重构 $\rho_{0:2}$，则会显著提高对 s、$\mathbf{E}_q$ 和 $\mathbf{T}_q$ 的估计精度。然而，对于 $\ell > 2$，这些估计并不会大幅改善。对过程进行层析重构所需的实验次数与依据该重构所得估计的准确度之间始终存在权衡。对于此源，$\ell = 2$ 在这两种资源之间取得了平衡。

相比之下，对于单叶量子比特源，以下量的估计值 s、$\mathbf{E}_q$和$\mathbf{T}_q$ 会随着 ℓ 的增加而稳步改善。通过增加重构密度矩阵 $\rho_{0:\ell}$的长度，可以捕获更长程的相关性。在观测该源时，或许值得为实验上可行的最大ℓ 寻找 $\rho_{0:\ell}$。

面对未知量子源时，应如何选择合适的 ℓ 值？一种策略是不断增大 ℓ，直到将 $\ell+1$ 引入的修正应用于相关信息量后，其幅度低于某个阈值。平稳性（以及由此产生的量子块熵的凹性）确保未来的修正也将低于该阈值。

F. 使用已知量子字母表的层析

如果观察者额外知晓某个 HMCQS 可能发射哪些纯态（即量子字母表$\mathcal{Q}$），便可利用这一知识，通过推断底层经典过程的词概率 $\overleftrightarrow{X}$，而非进行完整的层析重构，来简化系统辨识任务。对于量子比特，我们可以借助布洛赫球从几何角度表示这种简化；见图 25。布洛赫球表面上的每一点都表示 $\mathcal{H}^2$ 中的一个纯量子比特态——例如，位于 z 轴两极的 $|0\rangle$ 和 $|1\rangle$ ——而内部的每一点则表示一个可能的量子比特密度矩阵。以下假设$\mathcal{Q}$ 的每个元素都是唯一的。

1. $\ell = 1$

长度为 1 的密度矩阵 ρ_0 必须满足以下方程

$$\rho_0 = \sum_{|\psi_x\rangle \in \mathcal{Q}} \Pr(|\psi_x\rangle)\, |\psi_x\rangle\langle\psi_x| \ .$$

通过层析找到 ρ_0 后，可以重排此方程，以推断在 $\Pr(X = x) = \Pr(|\psi_x\rangle)$ 的条件下，$\overleftrightarrow{X}$ 的长度为 1 的词分布。这项任务能否完成取决于 $|\mathcal{Q}|$ 与 d 之间的关系。对于量子比特情形（$d = 2$），若 $|\mathcal{Q}| \leq 3$，则可以唯一推断 $\Pr(X_0)$；假设 $\mathcal{Q}$ 中不存在简并态。

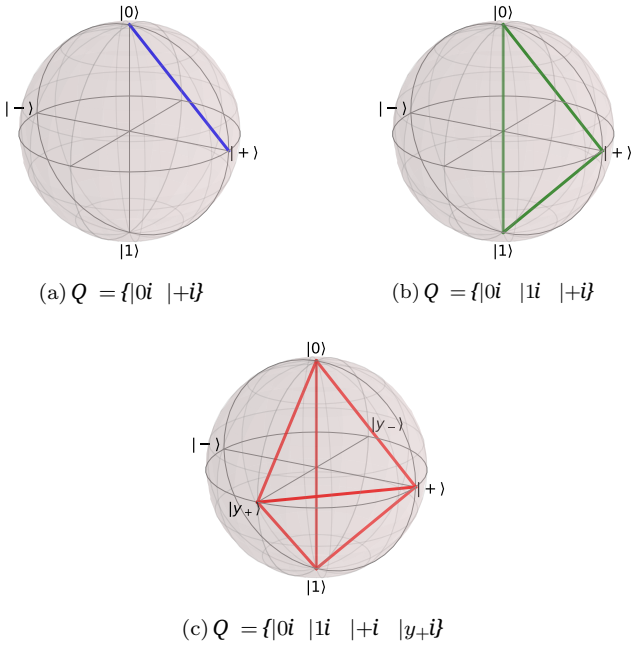
已知 $\mathcal{Q}$ 时，可能无需完整重构 ρ_0。我们先介绍几个简单示例，然后给出一种通用算法，用于求得 $\rho_{0:\ell}$，而无需对 $\rho_{0:\ell}$ 完整地进行量子态层析。

对于 $d = 2$ 和 $|\mathcal{Q}| = 2$，ρ_0 的可能取值被限制在布洛赫球内的一条弦上，该弦由 $\rho_0 = p\,|\psi_0\rangle\langle\psi_0| + (1-p)\,|\psi_1\rangle\langle\psi_1|$ 确定，其中 $0 < p < 1$。只需确定参数 p，而无需完整重构 ρ_0。观察者也可以选择一种信息量唯一确定 p 的测量。为此，最优的投影值测量是这样一种测量：其一对对跖投影算子可以由布洛赫球上一条直径连接，该直径平行于 ρ_0 可能取值所在的直线。例如，若 $\mathcal{Q} = \{|0\rangle, |+\rangle\}$，则可能的密度矩阵集合位于图 25a 中的线段上。此时，最佳投影值测量为 M_θ，其中 $\theta = \frac{3\pi}{4}$；测量结果 y_0（对应于对纯态 $|\psi(\theta = \frac{3\pi}{4})\rangle$ 的投影）和 y_1（对应于正交投影算子）。在这种情况下，很容易证明 $p = \frac{\sqrt{2}+1}{2} - \sqrt{2}\Pr(y_0|\rho_0)$，且 $\cos^2(\frac{3\pi}{8}) \leq \Pr(y_0|\rho_0) \leq \sin^2(\frac{3\pi}{8})$。$M_{01}$ 和 $M_\pm$ 也能确定 p，但需要更多样本，才能将 p 确定在所需的容差范围内。

对于 $d = 2$ 和 $|\mathcal{Q}| = 3$，ρ_0 的可能取值被限制在布洛赫球内的一个单纯形中，该单纯形由 $\rho_0 = p_0\,|\psi_0\rangle\langle\psi_0| + p_1\,|\psi_1\rangle\langle\psi_1| + (1-p_0-p_1)\,|\psi_2\rangle\langle\psi_2|$ 定义，其中 $0 < p_0$，$p_1 < 1$ 和 $p_0 + p_1 < 1$。可以确定参数 p_0 和 p_1，而不必像通常表征量子比特混合态那样确定 3 个参数。为此有两种简单的测量选择：使用信息完备 POVM，或使用两个不同的投影值测量。

对于第一种情况，考虑使用具有元素 $E_y = \frac{1}{2}\,|\phi_y\rangle\langle\phi_y|$ 的 SIC- POVM 对 ρ_0 进行测量，其中每个 $|\phi_y\rangle$ 由方程 52 描述。观察到测量结果 y 的概率可写为：

$$\begin{aligned}\Pr(y|\rho_0) &= \sum_x \Pr(|\psi_x\rangle)\Pr(y||\psi_x\rangle) \\ &= \sum_x \frac{p_x}{2}|\langle\psi_x|\phi_y\rangle|^2 \ .\end{aligned}$$



(a) $\mathcal{Q} = \{|0\rangle, |+\rangle\}$　(b) $\mathcal{Q} = \{|0\rangle, |1\rangle, |+\rangle\}$

(c) $\mathcal{Q} = \{|0\rangle, |1\rangle, |+\rangle, |y_+\rangle\}$

图 25. 给定 $\mathcal{Q}$ 时，长度为 1 的密度矩阵集合（ρ_0）边界的布洛赫球表示。(a) 有效态集合是一条线段。观察者只需确定一个参数。(b) 有效态集合是一个三角形的内部。观察者必须确定两个参数。(c) 有效态集合是一个四面体的内部。观察者必须确定三个参数，并且分解为基态系综并不唯一。这里 $|y_+\rangle = \frac{1}{\sqrt{2}}(|0\rangle - i\,|1\rangle)$。

可以重新排列这组 4 个方程（每个 POVM 元素对应一个方程），从而得到一组唯一的 p_x。

或者，可以使用两个投影值测量，其投影算子能够通过与可能的 ρ_0 值单纯形平行的（理想情况下彼此正交的）布洛赫球直径相连接。这样将得到 2 个参数，从而唯一确定 ρ_0 单纯形上的一个点。图 25b 展示了一个 $\mathcal{Q} = \{|0\rangle, |1\rangle, |+\rangle\}$ 的例子，其中 ρ_0 的可能值被限制在布洛赫球内的一个三角形内部。通过使用正交投影值测量 M_{01} 和 $M_\pm$（以及许多其他组合）进行测量，可以确定 p_0 和 p_1；此时 $(1 - p_0 - p_1) = 2\Pr(\text{‘+’}|\rho_0, M_\pm) - 1$，且 $(p_0 - p_1) = 2\Pr(\text{‘0’}|\rho_0, M_{01}) - 1$。对于 $d = 2$ 和 $|\mathcal{Q}| = 4$，ρ_0 的可能值被限制在布洛赫球内的一个四面体中，其顶点是 $\mathcal{Q}$ 的元素；而且，仅凭完全重构的 ρ_0 无法唯一推断出经典符号分布。例如，如果 $\mathcal{Q} = \{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}$ 且 ρ_0 是最大混合态，那么它可能对应于形如 $\rho_0 = p\,|0\rangle\langle 0| + p\,|1\rangle\langle 1| + (1-p)\,|+\rangle\langle +| + (1-p)\,|-\rangle\langle -|$ with $0 < p < 1$ 的任意混合。因此，对于 $d = 2$ 和 $|\mathcal{Q}| \geq 3$，了解层析优势将

Q 的层析优势有所降低，但并未消失，因为观察者可以立即排除任何位于$\mathcal{Q}$的元素所定义的凸多面体之外的 ρ_0 值。图 25c 展示了这一点，其中 ρ_0 值的可能区域被限制在布洛赫球体积的不到$_{1\ 4}$。

当$\mathcal{Q}$已知时， $d=3$ （三能级量子比特）也可以进行类似的简化。对任意混合量子三能级态进行完整层析需要确定 8 个参数，而在给定$\mathcal{Q}$ 的情况下确定经典分布需要$|\mathcal{Q}|-1$个参数。当 $|\mathcal{Q}|\leq 8$时，这通常体现出一种优势。我们不明确构造能够实现这一优势的测量，因为要在 8 维空间$\mathcal{H}^3$ 上几何理解混合态，远比在描述$\mathcal{H}^2$上混合态的三维布洛赫球复杂得多。

对于一般的量子位，完整层析所需的参数数量为 d^2-1。因此，当 $|\mathcal{Q}|<d^2$时，我们预计已知$\mathcal{Q}$ 会带来明显的层析优势（需要确定的参数更少）。

现在，我们已经可以给出一个适用于$|\mathcal{Q}|=n$ 和任意 d 的一般协议。我们的目标是仅凭测量统计恢复字母表态的底层分布（$p_x=$**Pr(**$|\psi_x\rangle$**)**，$0\leq x<n$）。首先，我们构造一个包含 $n+1$ 个元素的 POVM：对于$\mathcal{Q}$中的每个 $|\psi_x\rangle$，取 $E_y=c_y\,|\psi_x\rangle\langle\psi_x|$，以及$E_n=\mathbb{I}-\sum_{y=0}^{n-1}E_y$。每个 c_y 都是一个参数，可以对其进行调整以确保 E_n 为半正定。将该 POVM 应用于 ρ_0 后，我们得到一个关于 $n+1$ 种可能测量的分布（**Pr(**$y|\rho_0$**)**）。前 n 个与我们所需的分布相关，其关系为：

$$\begin{aligned}\Pr(y|\rho_0)&=\sum_x p_x\Pr(y|\,|\psi_x\rangle)\\&=\sum_x p_x c_y|\langle\psi_x|\psi_y\rangle|^2\ ,\end{aligned}$$

每个 $y<n$ 对应一个方程。如果一组 p_x 满足这一线性方程组，那么它与观测到的测量结果相一致。对于 $|\mathcal{Q}|<d^2$，解是唯一的。

2. $\ell=2$

了解$\mathcal{Q}$，在考虑对多个量子多能级系统进行层析时会带来进一步的优势。长度为 ℓ 的经典字分布具有 $|\mathcal{Q}|^\ell-1$ 个参数，而对 ℓ 个量子位进行完整层析则需要确定 $d^{2\ell}-1$ 个参数。

对于 $\ell=2$:

$$\begin{aligned}\rho_{0:2}=\sum_{|\psi_{x_0}\rangle,|\psi_{x_1}\rangle\in\mathcal{Q}}\Big[&\Pr(|\psi_{x_0}\rangle\otimes|\psi_{x_1}\rangle)\\&\big(|\psi_{x_0}\rangle\otimes|\psi_{x_1}\rangle\big)\big(\langle\psi_{x_0}|\otimes\langle\psi_{x_1}|\big)\Big]\ .\end{aligned}\tag{53}$$

我们再次明确考虑 $d=2$ 和 $|\mathcal{Q}|=2$ 的情形。长度为 2 的经典字有 4 种词概率，但存在 3 个约束条件：(i) 归一化，(ii) 平稳性，以及 (iii) 与单量子比特边缘分布一致。因此，只需确定一个参数即可重构 $\rho_{0:2}$。

考虑这样一项任务：在已知$\mathcal{Q}=\{|0\rangle\ \ |+\rangle\}$的情况下，重构图 6 中的 $|0\rangle$-$|+\rangle$量子黄金均值生成器所产生的长度为 2 的密度矩阵。首先分析单量子比特密度矩阵，可得 **Pr(**$|0\rangle$**)** $=\rho_{0\,3}^{\ 2}$ 和 $\frac{2}{3}\,|0\ \ 0|+\frac{1}{3}\,|+\ \ +|$ $\rangle\langle\rangle\langle$。

词 概率 **Pr(**$|00\rangle$**)** 是求出 $\rho_{0:2}$ 所需的唯一额外信息。以下简单的测量协议可以确定 **Pr(**$|00\rangle$**)**：使用 $M_\pm$ 测量两个连续的量子比特。如果量子比特处于状态$|+\rangle$，则永远不会看到测量结果“–”。如果量子比特处于状态 $|0\rangle$，则以$_{1\ 2}$ 的概率看到测量结果“–”。因此，**Pr(**$|00\rangle$**)** $=4$ **Pr(**‘$--$’$|\rho_{0:2},M_\pm\otimes M_\pm$**)**。

该方法可以推广到任意的二元素字母表 $\mathcal{Q}=\{|\psi_0\rangle\ \ |\psi_1\rangle\}$。首先，使用一个 PVM $M_{\tilde{0}1}$ 测量两个连续的量子位，其中一个元素是投影到 $|\psi_1\rangle$ 上的投影算子，对应测量结果“1”；另一个正交投影算子对应测量结果“0”。其次：

$$\begin{aligned}&\Pr(\,|\psi_0\rangle\otimes|\psi_0\rangle)\\&\quad=(\Pr(\text{`}\tilde{0}\text{'}|\,|\psi_0\rangle\,,M_{\tilde{0}1}))^{-2}\Pr(\text{`}\tilde{0}\tilde{0}\text{'}|\rho_{0:2},M_{\tilde{0}1}\otimes M_{\tilde{0}1})\\&\quad=(1-|\langle\psi_0|\psi_1\rangle|)^{-2}\Pr(\text{`}\tilde{0}\tilde{0}\text{'}|\rho_{0:2},M_{\tilde{0}1}\otimes M_{\tilde{0}1})\ .\end{aligned}$$

与通常重构 $\rho_{0:2}$ 所需的 9 个参数相比，这显然更有优势，因为它考虑到了单量子比特边缘态和平稳性各自施加 3 个约束。

对于 $d=2$ 和 $|\mathcal{Q}|\geq 3$，我们必须确定底层经典分布在$\mathcal{Q}$上的其他参数。为此，我们反复应用元素为 $E_y=\frac{1}{2}\,|\phi_y\rangle\langle\phi_y|$的 SIC-POVM，其中每个 $|\phi_y\rangle$ 由公式 52 描述。

观测到长度为 2 的测量的概率

序列 y_0y_1 可以写为：

$$\begin{aligned}\Pr(y_0y_1|\rho_{0:2}) &= \sum_{x_0,x_1} p_{x_0,x_1} \Pr(y_0y_1|\,|\psi_{x_0}\rangle \otimes |\psi_{x_1}\rangle) \\ &= \sum_{x_0,x_1} \frac{p_{x_0,x_1}}{4} |\langle\psi_{x_0}|\phi_{y_0}\rangle|^2 |\langle\psi_{x_1}|\phi_{y_1}\rangle|^2 \ .\end{aligned}$$

共有 $|\mathcal{Q}|^2$ $p_{x_0,\ x_1}$ 个。如果一组 $p_{x_0,\ x_1}$ 是该方程组的解，则它与测量统计特征相符；对于$|\mathcal{Q}|=3$，该解是唯一的。

对于 $|\mathcal{Q}|=n$ 以及任意的 d，我们可以像对 $\ell=1$所做的那样，由 $\mathcal{Q}$ 构造一个 POVM。随后便可以根据测量统计推断出一组p_{x_0,x_1}，使其满足由此得到的方程组。对于长度为2的词，方程如下：

$$\begin{aligned}\Pr(y_0y_1|\rho_{0:2}) &= \sum_{i} p_{x_0,x_1} \Pr(y_0y_1|\,|\psi_{x_0}\rangle \otimes |\psi_{x_1}\rangle) \\ &= \sum_{x_0,x_1} p_{x_0,x_1} c_{y_0} c_{y_1} |\langle\psi_{x_0}|\psi_{y_0}\rangle|^2 |\langle\psi_{x_1}|\psi_{y_1}\rangle|^2 \ .\end{aligned}$$

共有 n^2 个方程，每个方程对应 POVM 元素$E_{y_0}\otimes E_{y_1}$所对应的一种长度为2的测量序列。由于归一化，计算其他可能的测量结果（对应于 E_n）是冗余的。

3. $\ell \geq 3$

将这一分析推广到长度为ℓ 的密度矩阵 $\rho_{0:\ell}$，其形式为：

$$\rho_{0:\ell} = \sum_{|\psi_w\rangle\in\mathcal{Q}^\ell} \left[\Pr(|\psi_w\rangle)\, |\psi_w\rangle\langle\psi_w| \right] ,$$

其中每个 $|\psi_w\rangle$ 的形式如公式 (5) 所示。对于一般情形 $|\mathcal{Q}|^\ell < d^{2\ell}$，可以唯一确定长度为$\ell$ 的词分布。

上述针对$\ell=2$ 探讨的各种测量策略可以推广到任意的 ℓ。我们将具体介绍两种策略：对 $d=2$ 使用 SIC-POVM，以及对任意的 d使用由 $\mathcal{Q}$ 构造的 POVM。

考虑反复使用由元素 $E_y=\frac{1}{2}|\phi_y\rangle\langle\phi_y|$ 构成的 SIC-POVM，每次 $|\phi_y\rangle$ 均由式（52）描述。进行 ℓ 次测量后，可观测到一个长度为 ℓ 的词 $y_{0:\ell}$。$\mathcal{Y}$ 是由测得符号组成的四元素字母表。

观测长度为ℓ 的测量序列 $y_{0:\ell}$ 可以写成：

$$\begin{aligned}\Pr(y_{0:\ell}|\rho_{0:\ell}) &= \sum_{|\psi_w\rangle} \Pr(|\psi_w\rangle) \Pr(y_{0:\ell}|\,|\psi_w\rangle) \\ &= \sum_{|\psi_w\rangle} \frac{\Pr(|\psi_w\rangle)}{2^\ell} |\langle\psi_w|\phi_{y_{0:\ell}}\rangle|^2 \ ,\end{aligned}$$

其中 $|\phi_{y_{0:\ell}}\rangle = \bigotimes_{t=0}^{\ell}|\phi_{y_t}\rangle$，因子 2^ℓ 来自 POVM 元素。每个 4^ℓ 序列都有一个可通过测量估计的概率。求解这个方程组，可以得到与测量结果一致的底层长度为 ℓ 的词概率。如果 $|\mathcal{Q}|\leq 3$，则该解唯一。

我们也可以使用直接由 $\mathcal{Q}$ 构造的 POVM 进行 ℓ 次测量。在这种情况下，得到的方程为：

$$\begin{aligned}\Pr(y_{0:\ell}|\rho_{0:\ell}) &= \sum_{|\psi_w\rangle} \Pr(|\psi_w\rangle) \Pr(y_{0:\ell})|\,|\psi_w\rangle) \\ &= \sum_{|\psi_w\rangle} p_w \left(\prod_{t=0}^{\ell-1} c_{y_t}\right) |\langle\psi_w|\psi_{y_{0:\ell}}\rangle|^2 \ .\end{aligned}$$

与前面一样，当 $|\mathcal{Q}|<d^2$ 时，可以唯一确定 $|\mathcal{Q}|^\ell$ 底层的词概率（p_w）。

G. 源重构

观察一个可分量子位过程并得到长度为ℓ 密度矩阵 $\rho_{0:\ell}$ 后，能否推断出生成该过程的隐马尔可夫链 HMCQS? 下面针对 ℓ的不同取值，重构生成未知过程 $\rho_{0:\ell}$ 的源。请注意，能够生成 $\rho_{0:\ell}$ 的源可能无法生成 $\rho_{0:\ell+1}$。

1. $\ell=1$

确定 ρ_0 后，观察者可以构造量子信息源的独立同分布近似。给定将 ρ_0 分解为纯态 $|\psi_x\rangle$的任意分解——如式 (3) 所示——相应的 HMCQS 由一个内部状态 $\mathcal{Q}=\{|\psi_x\rangle\}$ 组成，而每个 $|\psi_x\rangle$对应的唯一转移概率，就是它在 ρ_0 的分解中的对应概率 $\Pr(|\psi_x\rangle)$。通过对 ρ_0进行特征分解，可以得到唯一模型。在这种情况下，模型以等于相应本征值的概率发射每个本征态 $|\psi_i\rangle$ ： $\Pr(|\psi_i\rangle)=\lambda_i$。

2. $\ell = 2$

借助 $\rho_{0:2}$，观察者开始对产生量子比特间相关性的源进行建模。为此，需要找到符合式 (53) 形式的 $\rho_{0:2}$ 的可分解分解。若$\mathcal{Q}$ 已知，上文已介绍多种求取此类可分解分解的方法。对于一般的双量子位密度矩阵，判定其可分离还是纠缠通常是 NP 困难的 [58]。此外，还有许多关于可分性的充要条件，例如正部分转置（PPT）判据 [59]。

以下假设观察者拥有 $\rho_{0:2}$ 的可分解分解，其由字母表态 $\mathcal{Q}'$ 组成；该字母表可能与源的字母表 $\mathcal{Q}$ 相同，也可能不同。一般而言，二量子位密度矩阵的分解中，基态集合可能不同，但我们要求采用对称分解，使两个量子位的基态均为 $\mathcal{Q}'$。通过该分解，他们可以构造一个 HMCQS，其底层马尔可夫动力学由公式（9）描述。该 HMCQS 具有 $|\mathcal{Q}'|$ 个内部状态，转移概率为 $T_{\sigma_{x'}\,\sigma_{x''}} = \Pr(|\psi_{x''}\rangle\,||\psi_{x'}\rangle)$，可由 $\rho_{0:2}$ 计算得出。$\rho_{0:2}$ 的不同可分解分解会产生不同的 HMCQS，其较长序列的统计特征可能有所不同。要确定哪一种是对源更准确的模型，需要对 $\rho_{0:3}$ 进行层析，以改进该模型。

3. $\ell \geq 3$

随着量子位数量增加，求得可分离分解的计算开销也会增大[60]。不过，如果能得到 $\rho_{0:3}$ 的可分离分解，其中所有 3 个量子位的基态为$\mathcal{Q}'$，那么就可以构造该源的长度为 2 的马尔可夫近似。$\mathcal{Q}'$ 中的每个长度为 2 的量子位序列都对应一个不同的内部 HMCQS 状态。因此，它有 $|\mathcal{Q}'|^2$ 个内部状态，除非某些长度为 2 的序列不可能出现。转移概率的形式为$T_{\sigma_{x_0,x_1},\sigma_{x_1,x_2}} = \Pr(|\psi_{x_2}\rangle\,||\psi_{x_0}\rangle\,|\psi_{x_1}\rangle)$。只有符合拼接规则的转移概率才非零。

通过这种方式可以继续进行：在给定 $\rho_{0:\ell}$ 的可分解分解后对源进行近似，从而在某些序列被禁止的情况下，得到一个具有 $|\mathcal{Q}'|^{\ell-1}$ 个或更少状态的 HMCQS。每个状态对应于一个长度为 $\ell-1$ 的词，其中组成该词的纯态取自$\mathcal{Q}'$。模型中的内部状态数量随 ℓ呈指数增长，但并非所有这些状态都能产生唯一的未来预测。如果是这样，则可以在不损失预测能力的情况下将它们合并，这正是经典计算力学中所采用的方法。用于执行此操作的一般算法超出了本文的范围。

H. 讨论

前文详细介绍了如何通过对长度为ℓ 的密度矩阵 $\rho_{0:\ell}$进行层析重构，来识别未知量子过程的诸多方面。当量子比特之间存在相关性时，这相较于通常假设源是独立同分布的情况，能够提供预测优势。我们从重构经典过程的方法出发，针对 d 和 $|\mathcal{Q}|$ 以及 ℓ 的不同取值，开发了多种测量协议，以便在已知$\mathcal{Q}$ 时确定过程的统计特征。随后，我们介绍了完全由通过层析重构得到的密度矩阵 $\rho_{0:\ell}$ 的可分解分解推导出的量子源有效模型。

由于我们的目标是利用相关性来改进对未来测量结果的预测，因此值得探究：未来的测量结果能够被多准确地预测？为初步回答这个问题，下面将针对 $\ell = 2$ 这一特殊情形以及任意 ℓ，定义最大预测性测量。

对于相关量子位过程，已知长度为 2 的密度矩阵 $\rho_{0:2}$ 的观察者可以对第一个量子位进行测量，并根据测量结果进行条件化，从而降低测量第二个量子位时的不确定性。他们在第一个量子比特 ρ_0 上施加 M_0 （可能的测量结果为 y_i）。此时联合系统处于如下经典-量子态:

$$\rho_{0:2}^{M_0} = \sum_{y_i} \Pr(y_i)\,|y_i\rangle\langle y_i| \otimes \rho_1^{y_i}\ ,$$

其中， $\rho_1^{y_i}$ 是在测量结果 y_i 条件下的量子比特密度矩阵。

随后，第二个量子比特的条件冯·诺依曼熵为:

$$S(\rho_1|M_0(\rho_0)) = \sum_{y_i} \Pr(y_i) S(\rho_1^{y_i})\ .$$

测量结果不确定性最小的秩一测量是在 $\rho_1^{y_i}$ 的特征基中进行的。不同的 y_i 值通常对应于第二个量子比特上不同的最小熵测量。

对于两个量子比特，我们可以找到使第二个量子比特的条件冯·诺依曼熵最小化的投影值测量。这引出了该过程的一个基无关性质:

$$S_{min}(\rho_1|M_{min}(\rho_0)) = \min_{M_0} S(\rho_1|M_0(\rho_0))\ ,$$

其中，最小值是在 ρ_0 上的所有投影值测量中取到的。
如果实验者重构了 $\rho_{0:\ell}$，则对块中除最后一个量子位外的所有量子位进行测量，采用测量

协议$\mathcal{M}$，其可能的测量结果为 $y_{0:\ell-1}$，会使该块处于经典-量子态：

$$\rho_{0:\ell}^{\mathcal{M}} = \sum_{y_{0:\ell-1}} \Pr(y_{0:\ell-1}) \left|y_{0:\ell-1}\right\rangle \left\langle y_{0:\ell-1}\right| \otimes \rho_{\ell}^{y_{0:\ell-1}} \ .$$

为使第ℓ个量子比特的条件冯·诺依曼熵最小，我们计算：

$$S_{min}(\rho_{\ell}|\mathcal{M}_{min}(\rho_{0:\ell-1})) = \min_{\mathcal{M}} S(\rho_{\ell}|\mathcal{M}(\rho_0)) \ ,$$

其中，最小值是在 $\rho_{0:\ell-1}$上所有局域测量协议中取得的。

进一步的研究需要探索测量协议空间，以找出对于任意可分离量子位过程而言，相较于独立同分布模型具有最大预测优势的协议。

最后需要指出，我们的源重构程序需要的模型状态数量随 ℓ呈指数增长。未来，从密度矩阵中寻找最小模型的研究，也需要合并具有相同预测的状态，并对可能的模型拓扑结构进行推断，这与经典贝叶斯结构推断相同。

VII. 结论

受以往关于经典随机过程的信息论研究启发，我们提出了用于量化平稳量子信息源的结构和信息产生的方法。我们确定了与量子块熵 $S(\ell)$相关的性质，据此可以确定给定量子位过程中的随机性和结构的数量。我们给出了由此产生的经典测量过程的信息性质界限。特别是，我们证明，在任何 ℓ下，它们的熵率或块熵都不可能低于原始量子过程。

我们分析了若干隐马尔可夫链量子信源（HMCQS），说明观察者如何通过测量发射出的量子比特，与信源的内部状态同步。我们证明，如果信源允许同步观察，那么在固定基测量无法实现同步的情况下，自适应测量协议能够实现同步并维持同步。

后续研究将以多种方式扩展这些方法和结果。尽管本文为求简洁而聚焦于可分离量子序列，但也可以通过结合 HMCQS 和一个能够按序生成矩阵乘积态的 D 维量子系统，类似地研究纠缠量子位序列 [24]。这样做将开启对矩阵乘积算子熵收敛的研究 [61]。

关于经典随机过程，已有许多结果可以推广到量子过程。例如，不可对角化经典动力学的信息度量存在闭式表达式[50, 62–64]。将这些结果推广到量子动力学，有助于更准确地确定本文介绍的量子信息性质。同样，前文为可分量子过程的涨落定理和大偏差理论奠定了基础。最后，值得进一步发展适用于量子随机过程的因果等价关系，并通过推广经典结果[65]来发展量子 ϵ-机。

可分离的量子序列也可作为有限状态量子信息转换器进行信息处理的资源，这类转换器能够将一个量子过程转换为另一个量子过程。这些操作本身就值得关注，此外还会产生热力学后果：它们要么需要做功才能运行（例如克服兰道尔擦除 [66]所引发的耗散），要么充当量子版的信息驱动引擎，利用环境关联来完成有用功 [26, 67, 68]。某些量子过程已展现出这种行为 [69]。

最后，尽管自旋链在各种哈密顿量作用下的基态和热态具有空间延展性，它们仍是量子过程。因此，本文引入的量子信息度量可用于依据生成这些状态所需的源复杂度对其进行分类。

致谢

我们感谢 Fabio Anza、Sam Loomis、Alex Jurgens、Ariadna Venegas-Li，以及特柳赖德科学研究中心信息引擎研讨会的与会者所进行的有益讨论。JPC 感谢在访问期间，特柳赖德科学研究中心、圣菲研究所和加州理工学院的热情款待。本研究基于或部分基于以下资助项目的支持：基础问题研究所和费策·富兰克林基金（硅谷社区基金会的捐赠者建议基金）资助的项目编号 FQXi-RFP-IPW-1902 和 FQXi-RFP-1809，以及美国陆军研究实验室和美国陆军研究办公室资助的项目编号 W911NF-18-1-0028 和 W911NF-21-1-0048。

附录 A：经典过程中的信息

本节简要回顾文献 [11] 中提出的、用于量化经典随机过程的随机性和相关性的若干性质。详细内容和完整证明见该文献。本文正文将介绍适用于量子过程的相应版本。

1. 香农熵

我们用具有可能测量结果 $\{x_1, x_2, ...x_n\}$ 的离散随机变量 X 的香农熵来量化其不确定性：[7]:

$$\mathrm{H}[X] \equiv -\sum_{i=1}^{n} \Pr(x_i) \log_2 \Pr(x_i) \ .$$

我们使用 $\log_2$，，在这种情况下，香农熵的单位为比特。

为了研究多个随机变量之间的相关性，我们使用若干个与香农熵相关的其他信息量。首先，两个离散随机变量的联合熵——X 和 Y，其可能的测量结果分别为 $\{x_1, x_2, ...x_n\}$ 和 $\{y_1, y_2, ...y_m\}$——定义为：

$$\mathrm{H}[X,Y] \equiv -\sum_{i=1}^{n}\sum_{j=1}^{m} \Pr(x_i, y_j) \log_2 \Pr(x_i, y_j) \ .$$

当且仅当联合熵分解为 $\mathrm{H}[X,Y] = \mathrm{H}[X] + \mathrm{H}[Y]$ 时，X 和 Y 才是统计独立的。
其次，X 在给定 Y 条件下的条件熵为：

$$\mathrm{H}[X|Y] \equiv -\sum_{i=1}^{n}\sum_{j=1}^{m} \Pr(x_i, y_j) \log_2 \Pr(x_i|y_j) \ .$$

$\mathrm{H}[X|Y]$ 表示在已知 Y 的值后，X 取值的不确定性。请注意，$\mathrm{H}[X|Y] \geq 0$ 与 $\mathrm{H}[X|Y] \neq$ 并不对称。如果 X 和 Y 不相关，则 $\mathrm{H}[X|Y] =$ 与 $\mathrm{H}[X]$ 相等。

根据上述定义，可以推导出以下单位算子，它将条件熵与联合熵联系起来：

$$\mathrm{H}[X|Y] = \mathrm{H}[X,Y] - \mathrm{H}[Y] \ .$$

第三，最后，两个随机变量之间的互信息为：

$$\mathrm{I}[X:Y] \equiv \mathrm{H}[X] - \mathrm{H}[X|Y] \ .$$

完全了解 Y 后，我们可以获得关于 X 的信息量。互信息具有对称性，并且 $\mathrm{I}[X:Y] = 0$ 当且仅当 X 和 Y 在统计上相互独立时成立。

2. 块熵

对于经典随机过程，我们通过计算 ℓ 个连续随机变量构成的块的联合熵 $\mathrm{H}[X_{0:\ell}]$，来量化其中的不确定性。这就是块熵：

$$\begin{aligned}\mathrm{H}[\ell] &\equiv \mathrm{H}[X_{0:\ell}] \\ &= -\sum_{x_{0:\ell}} \Pr(x_{0:\ell}) \log_2 \Pr(x_{0:\ell}) \ ,\end{aligned}$$

其中，求和遍历所有长度为 ℓ 的词，且 $\mathrm{H}[0] \equiv 0$。

$\mathrm{H}[\ell]$ 单调递增且向下凹，其在 $\ell \to \infty$ 时的行为反映了过程的相关性和随机性 [11]。$\mathrm{H}[\ell]$ 的一般行为及其与我们将要定义的其他信息性质之间的关系见图 2。

3. 香农熵率

以下简要概述文献 [11] 的研究结果。有关多变量系统及块熵相关信息论量的更深入探讨，请参阅该文献。

其中首先介绍的是香农熵率 h_μ：

$$h_\mu = \lim_{\ell \to \infty} \frac{\mathrm{H}[\ell]}{\ell} \ . \tag{A1}$$

式（A1）中的极限保证对所有平稳过程 [70] 都存在。h_μ 是信息源产生的不可约随机性，其单位为每符号比特数。

香农熵率也可以用条件熵表示：

$$h_\mu = \lim_{\ell \to \infty} \mathrm{H}[X_0|X_{-\ell:0}] \ , \tag{A2}$$

因此，h_μ 也可以等价地理解为：已知此前所有符号时，下一个符号的平均不确定性。

为了更好地理解 h_μ，我们来看几个简单的例子。
对于独立同分布过程，块熵显然为 $\mathrm{H}[\ell] = \ell\mathrm{H}[X_0]$，因此，$h_\mu = \mathrm{H}[X_0]$。由于不存在

变量之间的相关性，了解过去的符号无法减少下一个符号的不确定性。
如果一个过程具有周期性（例如由交替出现的 0 和 1 组成），那么敏锐的观察者会注意到这一模式，并能够确定性地预测该过程未来的所有符号。在这种情况下 $h_\mu = 0$。

对于平稳马尔可夫过程和隐马尔可夫过程，观察者可以利用过去的观测来减少对后续符号的不确定性。 h_μ 将是他们在考虑了与过去符号的所有相关性之后，对下一个符号的平均不确定性。

从图形上看， h_μ 对应于块熵曲线在 $\ell \to \infty$ 处的斜率，如图 2 所示。

4. 冗余

对于字母表为 $\mathcal{X}$ 的随机过程，最大熵率为 $\log_2|\mathcal{X}|$，这对应于独立同分布随机变量 X_t，其在所有测量结果 x_t 上服从均匀分布。任何其他平稳过程都可以压缩 到其熵率 $h_\mu < \log_2|\mathcal{X}|$。

特定源可以被压缩的程度称为其冗余，定义如下：

$$\mathbf{R} \equiv \log_2|\mathcal{X}| - h_\mu \ .$$

$\mathbf{R}$ 包含两种截然不同的影响：单个随机变量内部的偏差，以及不同随机变量之间的相关性。要确定这两个因素的相对重要性，需要更仔细地考察 $\mathrm{H}[\ell]$。

5. 块熵的导数与积分

由于式（A1）中的极限存在，$\mathrm{H}[\ell]$ 至多呈线性增长。我们关注 $\mathrm{H}[\ell]$ 如何收敛到其线性渐近线，并将看到，对 $\mathrm{H}[\ell]$ 取离散导数（以及对这些导数取积分）可为过程分类提供有用的量。

考虑将离散导数算子 Δ 应用于函数 $F\colon \mathbb{Z} \to \mathbb{R}$：

$$\Delta F(\ell) = F(\ell) - F(\ell-1) \ .$$

我们可以多次对 F 应用 Δ，以获得更高阶导数：

$$\Delta^n F(\ell) = (\Delta \circ \Delta^{n-1})F(\ell) \ .$$

对 $\mathrm{H}[\ell]$ 求离散导数，可得到一组函数 $\Delta^n\mathrm{H}[\ell]$。接下来，我们研究这些离散导数本身如何收敛到其渐近值。为此，我们按以下方式对离散函数 $\Delta F(\ell)$ 求“积分”：

$$\sum_{\ell=A}^{B} \Delta F(\ell) = F(B) - F(A-1) \ . \tag{A3}$$

为了研究每个 $\Delta^n\mathrm{H}[\ell]$的收敛性质，我们将其在每个 ℓ 处与其渐近值 $\lim_{\ell\to\infty}\Delta^n\mathrm{H}[\ell]$进行比较。我们使用以下一般积分形式：

$$\mathcal{I}_n \equiv \sum_{\ell=\ell_0}^{\infty} [\Delta^n H(\ell) - \lim_{\ell\to\infty}\Delta^n\mathrm{H}[\ell]] \ , \tag{A4}$$

其中， ℓ_0 是 ℓ 的第一个取值，使得 $\Delta^n\mathrm{H}[\ell]$ 有定义。

6. 熵增益

$\mathrm{H}[\ell]$ 的一阶导数称为熵增益。其定义如下：

$$\Delta\mathrm{H}[\ell] \equiv \mathrm{H}[\ell] - \mathrm{H}[\ell-1] \ ,$$

对于 $\ell > 0$。我们令 $\Delta\mathrm{H}[0] \equiv \log_2(\mathcal{X})$。熵增益表示块大小增加一个随机变量时引入的额外不确定性，其单位为每符号比特数。请注意，由于 $\mathrm{H}[\ell]$ 单调递增且为凹函数，因此对于所有 ℓ， $\Delta\mathrm{H}[\ell] \geq \Delta\mathrm{H}[\ell+1] \geq 0$。图 3 展示了这一特性。

熵增益可以改写为条件熵：

$$\Delta\mathrm{H}[\ell] = \mathrm{H}[X_0|X_{-\ell:0}] \ , \tag{A5}$$

此时，它与熵率之间的关系便显而易见。利用公式 (A2)，我们得到：

$$h_\mu = \lim_{\ell\to\infty}\Delta\mathrm{H}[\ell] \ .$$

对于对信息源没有先验知识的观察者，通常需要利用有限数据序列来估计源的熵率。在这种情况下，熵增益可作为真实熵率的有限-ℓ 近似：

$$\begin{aligned} h_\mu(\ell) &\equiv \Delta\mathrm{H}[\ell] \\ &\equiv \mathrm{H}[\ell] - \mathrm{H}[\ell-1] \ . \end{aligned} \tag{A6}$$

7. 可预测性增益

$\mathrm{H}[\ell]$ 的二阶导数即为可预测性增益，定义如下：

$$\begin{aligned}\Delta^2 \mathrm{H}[\ell] &\equiv \Delta h_\mu(\ell) \\ &= h_\mu(\ell) - h_\mu(\ell-1) \ ,\end{aligned}$$

其中 $\ell > 0$。注意， $\Delta^2\mathrm{H}[\ell] \leq 0$。

$|\Delta^2 \mathrm{H}[\ell]|$ 是观察者将观测从长度为 $\ell-1$ 的块扩展到长度为ℓ的块时，获得的额外预测信息的平均量，其单位为每符号比特数[2]。$|\Delta^2 \mathrm{H}[\ell]|$ 的值较大，意味着第 ℓ 次测量对观察者而言特别有信息量，因此能大幅改善其对式 (A6) 所示熵率的估计。

还可以计算 $\mathrm{H}[\ell]$ 的高阶离散导数。就我们的目的而言，这并非必要；只需指出，对于平稳过程：

$$\lim_{\ell\to\infty} \Delta^n \mathrm{H}[\ell] = 0, n \geq 2 \ .$$

对于 $n = 2$，这可由 $h_\mu(\ell)$ 的收敛性得出；对于所有 $n > 2$，论证也类似。

8. 总可预测性

现在我们可以对函数 $\Delta^n\mathrm{H}[\ell]$进行积分。一般而言，这些积分的绝对值越大，过程中的相关性或统计偏差就越强。

我们首先研究可预测性增益 $\Delta^n\mathrm{H}[\ell]$ 如何收敛至其渐近值 $\lim_{\ell\to\infty}\Delta^2\mathrm{H}[\ell] = 0$。由于 $\Delta^2\mathrm{H}[0]$ 未定义，我们将使用式 (A4) 并令 $\ell_0 = 1$ 进行积分，从而得到总可预测性 $\mathbf{G}$：

$$\mathbf{G} \equiv \mathcal{I}_2 = \sum_{\ell=1}^{\infty} \Delta^2 H(\ell) \ . \tag{A7}$$

由于对所有 ℓ， $\Delta^2\mathrm{H}[\ell] < 0$，因此$\mathbf{G} < 0$ 也成立。$\mathbf{G}$ 的单位是每符号比特数。从图形上看，它是可预测性增益曲线与其趋近于 0 的线性渐近线之间的面积，如图 4 所示。

为了解释$\mathbf{G}$的值，我们应用式 (A3) 得到：

$$\begin{aligned}\mathbf{G} &= -\Delta \mathrm{H}[0] + \lim_{\ell\to\infty} \Delta \mathrm{H}[\ell] \\ &= -\log_2(|\mathcal{X}|) + h_\mu \\ &= -\mathbf{R} \ .\end{aligned}$$

过程的总可预测性在数值上等于其冗余，因此我们可以将 $|\mathbf{G}|$ 解释为一个过程每个符号所包含的可预测信息量。这里，我们再次强调，对于字母表为$\mathcal{X}$的给定过程，任何随机变量的最大熵都是 $\log_2(|\mathcal{X}|)$，其中包含两种信息：h_μ，即不可约随机性；以及 $|\mathbf{G}|$，即观察者可能对其进行预测的信息量。

因此，对于给定过程，总可预测性是熵率的函数，并且具有 h_μ同样的缺陷：它无法识别随机变量之间的统计相关性。较大的 $|\mathbf{G}|$ 可能是由随机变量严重偏倚的独立同分布过程造成的，也可能是由后续变量之间的强相关性造成的。幸运的是，我们接下来要介绍的量能够区分这两种情况。

9. 超额熵

研究熵增益收敛到渐近熵率 h_μ 的过程，会引出一个经过广泛研究的过程性质：超额熵：

$$\begin{aligned}\mathbf{E} \equiv \mathcal{I}_1 &= \sum_{\ell=1}^{\infty} \left[\Delta \mathrm{H}[\ell] - \lim_{\ell\to\infty} \Delta \mathrm{H}[\ell]\right] \\ &= \sum_{\ell=1}^{\infty} \left[\Delta \mathrm{H}[\ell] - h_\mu\right] \ .\end{aligned}$$

由于 $\Delta\mathrm{H}[\ell] \geq h_\mu$， $\mathbf{E} \geq 0$，其单位为比特。
将式 (A3) 应用于$\mathbf{E}$ 后，我们对它的理解会更加清晰：

$$\mathbf{E} = \lim_{\ell\to\infty} \left[\mathrm{H}[\ell] - h_\mu \ell\right] \ . \tag{A8}$$

虽然 h_μ 决定了 $\mathrm{H}[\ell]$的线性渐近行为，$\mathbf{E}$ 却涵盖了所有次线性效应。我们主要讨论$\mathbf{E}$ 有限的过程，这类过程称为有限型过程。若某个过程的$\mathbf{E}$ 为无穷（例如 $\mathrm{H}[\ell]$ 随 ℓ 呈对数增长），则称为非有限型 过程。从图形上看，$\mathbf{E}$ 对应于块熵曲线线性渐近线的 $\ell = 0$截距（如图 2 所示），也对应于熵增曲线与其渐近线之间的面积，如图 3 所示 h_μ。

过剩熵也可以表示为过程的随机变量链两半之间的互信息：

$$\mathbf{E} = \lim_{\ell\to\infty} \mathrm{I}\left[X_{-\ell:0} : X_{0:\ell}\right] \ . \tag{A9}$$

这表明，它是过程过去中可用于预测未来的信息总量。因此，它被视为过程记忆量的指标。更具结构性的方法表明，$\mathbf{E}$ 是预测随机过程 [65]所需实际记忆量的下界。

重要的是，$\mathbf{E}$可以轻易地区分独立同分布过程（此时$\mathbf{E} = 0$）与随机变量之间存在相关性的过程（此时$\mathbf{E} > 0$）。对于周期为 p 的所有过程，$h_\mu = 0$ 且 $\mathbf{H}[\ell]$ 在 $\ell = p$ 时达到最大值 $\log_2 p$ 。因此，所有周期为 p 的过程都有$\mathbf{E} = \log_2 p$。

与 h_μ一样，观察者通常可以只利用有限长度为 ℓ 的符号序列，来近似真实的超额熵。使用公式 (A8)，我们可以估计$\mathbf{E}$：

$$\mathbf{E}(\ell) \equiv \mathrm{H}[\ell] - \ell h_\mu(\ell) \ . \tag{A10}$$

10. 瞬态信息

基于对超额熵的分析，我们现在可以说，当 $\ell \to \infty$ 时，块熵曲线具有线性渐近线：

$$\mathrm{H}[\ell] \sim \mathbf{E} + h_\mu \ell \ . \tag{A11}$$

我们通过再进行一次离散积分来刻画 $\mathrm{H}[\ell]$ 收敛到该渐近线的方式，从而得到瞬态信息：

$$\mathbf{T} \equiv -\mathcal{I}_0 = \sum_{\ell=0}^{\infty} \left[\mathbf{E} + h_\mu \ell - \mathrm{H}[\ell]\right] \ .$$

$\mathbf{T}$ 的单位是比特×符号。$\mathbf{T}$ 在图 2 中以图形方式表示为 $\mathrm{H}[\ell]$ 曲线与其线性渐近线之间的面积，其中 $\ell \to \infty$。

瞬态信息可以改写为：

$$\mathbf{T} = \sum_{\ell=1}^{\infty} \ell \left[h_\mu(\ell) - h_\mu\right] \ ,$$

这表明$\mathbf{T}$ 衡量了同步到一个过程的难易程度。当观察者能够准确推断源当前所处的内部状态时，就认为观察者已经“同步”。如果他们保持同步，就能够最优地预测未来的测量结果。也就是说，他们对源的熵率 $h_\mu(\ell)$的估计等于其实际熵率 h_μ。

$\mathbf{T}$ 是一个值得注意的量，因为与$\mathbf{E}$不同，它能够区分不同周期-p 过程。

11. 马尔可夫阶数

我们为经典随机过程介绍的最后一个性质是马尔可夫阶数。若 R if R 是满足以下条件的最小值，则该过程的马尔可夫阶数为 R if R：

$$\Pr(X_0|X_{R:0}) = \Pr(X_0|X_{-\infty:0}) \ .$$

由式（2）可知，马尔可夫过程满足 $R \leq 1$。更一般地，对于马尔可夫阶数为 R 的过程，X_t 的分布仅取决于前面的 R 个符号。等价地，R 是使块熵达到其线性渐近线的 ℓ值：

$$\mathrm{H}[R] = \mathbf{E} + h_\mu R \ .$$

这一事实以图示方式呈现在图 2 中。

绝大多数有限状态隐马尔可夫过程都具有无限马尔可夫阶，这意味着对于所有 ℓ [71]，$\mathbf{H}[\ell] < \mathbf{E} + h_\mu \ell$。然而，观察者仍然可以利用有限的$\ell$ 估计来获取此处定义的信息性质。事实上，这些估计通常会随 ℓ [72]以指数速度收敛。

附录B：制备与测量的量子通道

正文声称，可分量子位过程是将经典过程 $\overleftrightarrow{X}$的实现（其符号字母表为$\mathcal{X}$）通过一个经典-量子通道所得，该通道将 $x \in \mathcal{X} \to |\psi_x\rangle \in \mathcal{H}$。正文还声称，测量操作可描述为一个以 $\rho_{0:\ell} \to y_{0:\ell}$为输入的量子-经典信道。以下采用参考文献 [6]中的形式体系，对这两项论断作进一步阐述。

我们首先说明如何将经典过程 $\overleftrightarrow{X}$的实现通过一个条件量子编码器传递。考虑一个经典-量子态，它由一个维数为 $|\mathcal{X}|$ 的经典寄存器，以及一个处于$\mathcal{H}$ 中并初始化为状态 $|0\rangle$的量子位组成。此外，设该经典寄存器存储某个经典随机变量 X_t 的测量结果，其结果取值于 $x \in \mathcal{X}$ ，使得：

$$\rho_{XA} = \sum_x \Pr(x) \left|x\right\rangle \left\langle x\right|_X \otimes \left|0\right\rangle \left\langle 0\right|_A \ . \tag{B1}$$

该状态将作为条件量子编码器 $\mathcal{E}_{XA\to B}$ 的输入，该编码器由一组$\{\mathcal{E}^x_{A\to B}\}$ $|\mathcal{X}|$完全正迹保持（CPTP）映射构成。我们构造每个映射，使其将初始量子态 $|0\rangle$ 转换为所需的纯量子比特态 $|\psi_x\rangle$；

即:

$$\mathcal{E}^x_{A\to B}(|0\rangle\langle 0|_A) = |\psi_x\rangle\langle\psi_x|_B \ , \tag{B2}$$

其中 $\mathcal{E}^x_{A\to B}$ 为酉操作。（对于 $d=2$，我们可以考虑布洛赫球上的旋转。）

编码为:

$$\begin{aligned}\rho_B &= \mathcal{E}_{XA\to B}(\rho_{XA}) \\ &= tr_X\left(\sum_x \Pr(x)\,|x\rangle\langle x|\otimes\mathcal{E}^x_{A\to B}\left(|0\rangle\langle 0|\right)\right) \\ &= \sum_x \Pr(x)\,|\psi_x\rangle\langle\psi_x|_B \ .\end{aligned}$$

现在，取由经典过程 $\overleftrightarrow{X}$的长度为 ℓ 的词 $x_{0:\ell}$ 组成的 ℓ 经典寄存器，我们可以使用 $\mathcal{E}_{XA\to B}$对 $X_{0:\ell}$ 进行编码并得到:

$$\begin{aligned}\rho_{0:\ell} &= tr_{X_{0:\ell}}\left(\sum_{x_{0:\ell}}\Pr(x_{0:\ell})\prod_{t=0}^{\ell-1}|x_t\rangle\langle x_t|\otimes\mathcal{E}^{x_t}_{A\to B}\left(|0\rangle\langle 0|\right)\right) \\ &= \sum_w \Pr(w)\,|\psi_w\rangle\langle\psi_w| \ ,\end{aligned}$$

其中 $|\psi_w\rangle$ 取公式 (5)的可分形式，因此 $\rho_{0:\ell}$与公式 (6)一致。

量子位的测量也以类似方式描述。设 M 是一个 POVM，其元素为 $\{E_y\}$，作用于量子比特态 ρ 其以这种方式将每个测量结果记录在经典寄存器 Y 中。Y 中各个值的分布由下式决定:

$$\begin{aligned}Y &= M(\rho) \\ &= \sum_y tr(E_y\rho)\,|y\rangle\langle y| \ .\end{aligned}$$

同样地，令 $\mathcal{M}_{0:\ell}$ 为由长度为 ℓ 的测量组成的序列，其可能的测量结果序列$y_{0:\ell}$ 根据某个协议 $\mathcal{M}$ 确定。当将 $\mathcal{M}_{0:\ell}$ 应用于联合状态$\rho_{0:\ell}$ 中的 ℓ 个连续量子比特时，我们假设 $\mathcal{M}_{0:\ell}$ 由公式 (10)形式的局域测量组成。在这种情况下，我们将对应于特定测量结果序列的 POVM 元素分解为: $E_{y_{0:\ell}} = \bigotimes_{t=0}^{\ell-1} E_{y_t}$。长度为 ℓ 的测量结果可以存储在 $Y_{0:\ell}$中，该集合包含 ℓ 个经典寄存器，其对应的概率为 $tr(E_{y_{0:\ell}}\rho_{0:\ell})$，因此:

$$\begin{aligned}Y_{0:\ell} &= \mathcal{M}_{0:\ell}(\rho_{0:\ell}) \\ &= \sum_{y_{0:\ell}} tr(E_{y_{0:\ell}}\rho_{0:\ell})\,|y_{0:\ell}\rangle\langle y_{0:\ell}| \\ &= \sum_{y_{0:\ell}} tr(\bigotimes_{t=0}^{\ell-1} E_{y_t}\rho_{0:\ell})\,|y_{0:\ell}\rangle\langle y_{0:\ell}| \ ,\end{aligned}$$

其中最后一行假设采用局域测量。此时，每个$|y_{0:\ell}\rangle$ 都是可分离状态，且所有 $|y_{0:\ell}\rangle$ 彼此正交。

[1] W. Heisenberg. 关于量子理论运动学和力学直观内容的探讨。*Z. Phys.*, 43(3-4):172–198, 1927年3月。2[2] A. Peres. Kochen-Specker 定理的两个简单证明。*J. Phys. A*, 24(4), 1991。2[3]W. K. Wootters and B. D. Fields. 通过互为无偏测量进行最优状态判定。*Ann.Phys. (N.Y.)*, 191(2):363–381, 1989。2, 32[4] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves. 对称信息完备量子测量。*J.Math. Phys.*, 45(6):2171–2180, 2004。2, 32[5] J. von Neumann. 量子力学的数学基础：新版。普林斯顿大学出版社，2018。2, 7[6] M. Wilde. 量子信息论。剑桥大学出版社，第2版，2017年12月。2, 10, 33, 43[7] C. E. Shannon. 通信的数学理论。*Bell Sys. Tech. J.*, 27:379–423, 623–656, 1948。2, 22, 40[8] B. Schumacher. 量子编码。《物理评论*A*》，51(4):2738–2747, 1995。 2[9]J. P. Crutchfield。秩序与混沌之间。《自然物理》，8:17–24，2012。 3[10] J. P. Crutchfield 和 K. Young。推断统计复杂性。《物理评论快报》，63:105–108，1989。 3[11]J. P. Crutchfield 和 D. P. Feldman。未被观察到的规律性，所观测到的随机性：熵收敛的层级。《混沌》，13(1):25–54，2003。 3, 14, 18, 20, 21, 26, 40[12] N. F. Travers 和 J. P. Crutchfield。有限状态源的精确同步。《统计物理学杂志》，145:1181–1201，2011。 3, 24, 26[13] N. F. Travers 和 J. P. Crutchfield。有限状态源的渐近同步。《统计物理学杂志》，145:1202–1223，2011。 3 [14] A. E. Venegas-Li、A. M. Jurgens 和 J. P. Crutchfield。受控量子比特过程中测量诱导的随机性与结构。《物理评论*E*》，102:040102，2020年10月。 3, 7, 21, 25[15] A. S. Holevo。量子编码定理。《数学科学进展》，53(6):193–230，1998。 3

[16] N. Datta 和 Y. Suhov。相互作用量子比特信息源的数据压缩极限。《量子信息处理》，1(4):257–281，2002。[17]D. Petz 和 M. Mosonyi。平稳量子源编码。《数学物理杂志》，42(10):4857–4864，2001。 3, 6, 12[18]Y. Nagamatsu、A. Mizutani、R. Ikuta、T. Yamamoto、N. Imoto 和 K. Tamaki。使用非独立同分布光源的量子密钥分发安全性。《物理评论*A*》，93 042325，2016 4 3 ( ): 。[19] F. A. Pollock、C. Rodriguez-Rosario、T. Frauenheim、M. Paternostro 和 K. Modi。量子过程的操作性马尔可夫条件。《物理评论快报》，120:040405，2018。 3[20] F. A. Pollock、C. Rodriguez-Rosario、T. Frauenheim、M. Paternostro 和 K. Modi。非马尔可夫量子过程：完整框架与高效表征。《物理评论*A*》，97:012127，2018。[21] P. Taranto、F. A. Pollock、S. Milz、M. Tomamichel 和 K. Modi。量子马尔可夫阶数。《物理评论快报》，122:14041，2019。[22]P. Taranto、F. A. Pollock 和 K. Modi K. 非马尔可夫记忆强度限制量子过程的可恢复性。《*npj*量子信息》，7(1):1–8，2021。[23] P. Taranto 、 S. Milz 、 F. A. Pollock 和 K. Modi。有限马尔可夫阶数的量子随机过程结构。《物理评论*A*》，99:042108，2019年4月。 3[24] C. Schon、E. Solano、F. Verstraete、J. I. Cirac 和 M. M. Wolf。纠缠多量子比特状态的序列生成。《物理评论快报》，95:110503，2005。 3, 39[25] C. Schon、K. Hammerer、M. M. Wolf、J. O. Cirac 和 E. Solano。腔量子电动力学中矩阵乘积态的序列生成。《物理评论*A*》，

55 3 032311, 2007 3 ( ): 。[26]
A. B. Boyd 、 D. Mandal 和 J. P. Crutchfield。关联驱动信息引擎与自修正热力学。《物理评论*E*》，95(1):012152，2017。3, 39[27] A. Chapman 和 A. Miyake。自主量子麦克斯韦妖如何利用关联信息。《物理评论*E*》，92(6):1–12，2015。 3[28] M. Gu、K. Wiesner、E. Rieper 和 V. Vedral。奥卡姆量子剃刀：量子力学如何降低经典模型的复杂性。*Nat. Commun.*，3，2012。 5[29] J. R. Mahoney、C. Aghamohammadi 和 J. P. Crutchfield。奥卡姆量子磨刀石：通过量子信道同步并压缩经典隐秘过程。《科学报告》，6:20495，2016。[30] W. Y. Suen、T. J. Elliot、J. Thompson、A. J. Garner、J. R. Mahoney、V. Vedral 和 M. Gu。量子多体系统中的结构复杂性综述。《统计物理学杂志》，187，2022年04月。[31] F. C. Binder、J. Thompson 和 M. Gu。用于非马尔可夫复杂过程的实用幺正模拟器。《物理评论快报》，120:240502，2018年6月。[32] S. P. Loomis 和 J. P. Crutchfield。随机过程经典与量子模型中的强优化与弱优化。《统计物理学杂志》，176(6):1317–1342，2019。 5[33] M. A. Nielsen 和 I. L. Chuang。《量子计算与量子信息：十周年纪念版》。剑桥大学出版社，美国，第10版，2011年。6、10、11、32、33[34]S. Gudder。量子马尔可夫链。《*J. Math. Phys.*》，49(7)，2008年。7[35] A. Monras、A. Beige 和 K. Wiesner。隐量子马尔可夫模型与多体态的非自适应读出。2010年。7[36]K. Wiesner 和 J. P. Crutchfield。有限型随机过程和量子过程中的计算。《*Physica D*》，237(9):1173–1195，2008年。[37] S. Srinivasan、G. Gordon 和 B. Boots。隐量子马尔可夫模型学习。《*AISTATS 2018*》，第1979–1987页，2018年。7[38]D. Perez-Garcia、F. Verstraete、M. M. Wolf 和 J. I. Cirac。矩阵乘积态表示。《***arXiv***预印本 ***quant-ph/***

0608197, 2006 7 》。[39]
F. Verstraete、J. J. García-Ripoll 和 J. I. Cirac。矩阵乘积密度算子：有限温度系统和耗散系统的模拟。《*Phys. Rev. Lett.*》，

93 20 1 5, 2004 7 ( ): – 。[40]
C. Moore 和 J. P. Crutchfield。量子自动机与量子文法。《*Theoret. Comp. Sci.*》，237:1-2:275–306，2000年。8[41] D. Qiu、L. Li、P. Mateus 和 J. Gruska。量子有限自动机。载于《*Handbook of Finite State Based Models andApplications*》，第113–144页。10 2012年。[42] S. Zheng、D. Qiu、L. Li 和 J. Gruska。具有量子态和经典状态的单向有限自动机。载于《*LectureNotes in Computer Science*》，第7300卷 LNAI，第273–290页。2012年。8[43]M. Junge、R. Renner、D. Sutter、M. M. Wilde 和 A. Winter。通用恢复映射与量子相对熵的近似充分性。《*Ann. HenriPoincaré*》，

19 10 2955 2978, 2018 11 ( ): – 。[44]
O. E. Lanford 和 D. W. Robinson。量子统计力学中状态的平均熵。《*J. Math. Phys.*》，9(7):1120–1125，1968年。12[45] M. Ohya 和 D. Petz。《量子熵及其应用》。1993年。12[46] T. M. Cover 和 J. A. Thomas。《信息论要素》。Wiley-Interscience，纽约，1991年。12[47]N. F. Travers 和 J. P. Crutchfield。历史与生成器 ϵ-机器的等价性。arxiv.org:1111.4500 [ ] 24, 26 math.PR。[48]
C. J. Ellison、J. R. Mahoney 和 J. P. Crutchfield。预测、回溯预测以及当前存储的信息量。《*J. Stat. Phys.*》，136 1034，2009年 6 1005 24 ( ): – 。[49] A. M. Jurgens 和 J. P. Crutchfield。隐马尔可夫过程的香农熵率。《*Journal of StatisticalPhysics*》，183 2 1 18, 2021 24 ( ): – 。[50]
J. P. Crutchfield 、 P. Riechers 和 C. J. Ellison。精确复杂度：内在计算的谱分解。《*Phys. Lett. A*》，

380 9 10 998 1002, 2016 24, 39 (-): – 。[51]
S. Marzen 和 J. P. Crutchfield。离散时间更新过程的信息与因果架构。《*Entropy*》，

17(7):4891–4917，2015年。28[52]A. Jurgens 和 J. P. Crutchfield。发散的预测状态：平稳遍历隐马尔可夫过程的统计复杂度维数。《混沌》，31(8):0050460，2021年。28[53] Y. Fujiwara。量子比特序列的解析。《*IEEE Trans.Inf. Theory*》，59(10):6796–6806，2013年。29[54] C. C. Strelioff 和 J. P. Crutchfield。隐过程的贝叶斯结构推断。《《物理评论 *E*》》，89(4):042119，2014年。32[55]Rob T Thew、Kae Nemoto、Andrew G White 和 William J Munro。量子高维比特量子态层析。《《物理评论*A*》》，

66 1 012303, 2002 32 ( ): 。I.

[56]
Bengtsson 和 K. Życzkowski。《量子态的几何：量子纠缠导论》。剑桥大学出版社，2017年。33[57] J. Lawrence、Č. Brukner 和 A. Zeilinger。n量子比特上的两两互 unbiased 的二元可观测量集合。《《物理评论*A*》》，65:032320，2002年2月。33[58]Sevag Gharibian。量子可分性问题的强 NP难性。《*arXiv*预印本 *arXiv:0810.4507*》，2008年。38[59] Pawel Horodecki。具有正部分转置的可分性判据与不可分混合态。《*Phys.Lett. A*》，

232 5 333 339, 1997 38 ( ): –

。M. Horodecki 、 P. Horodecki 和 R. Horodecki。n 粒子混合态的可分性：以线性映射表示的必要且充分条件。《*Phys, LettersA*》，

[60]
283 1 2 1 7, 2001 38 (-): – 。B. Pirvu、V. Murg、J. I. Cirac 和 F. Verstraete。矩阵乘积算子表示。《*New J. Phys.*》，

[61]
12 2 025012, 2010 39 ( ): 。P. Riechers 和 J. P. Crutchfield。表观复杂性的谱简洁性，第一部分：预测的不可对角化元动力学。《混沌》，28:033115，2018年。39

[62]
[63] P. M. Riechers 和 J. P. Crutchfield. 超越谱定理：分解不可对角化算子的任意函数。《*AIP*进展》，8:065305，2018。[64]P. Riechers 和 J. P. Crutchfield. 表观复杂性的谱简洁性，第二部分：精确复杂性与复杂性谱。《混沌》，28:033116，2018。39[65] C. R. Shalizi 和 J. P. Crutchfield. 计算力学：模式与预测、结构与简洁性。《统计物理学杂志》，

104 3 4 817 879, 2001 39, 43 (-): – 。

[66]
R. Landauer. 计算过程中的不可逆性与热产生。*IBM J. Res. Develop.*，5 191，1961 3 183 39 ( ): – 。[67] A. B. Boyd、D. Mandal 和 J. P. Crutchfield. 利用环境关联：必要多样性的热力学。《统计物理学杂志》，167(6):1555–1585，2016。39[68]A. M. Jurgens 和 J. P. Crutchfield. 麦克斯韦棘轮的功能热力学：模式的构建与解构、行为的随机化与去随机化。*Phys. Rev. Res.*，

2 3 033334, 2020 39 ( ):

[69]
。R. Huang 、 P. Riechers 、 M. Gu 和 V. Narasimhachar. 从有记忆量子随机过程中提取预测性功的引擎。2022年07月。39[70] T. M. Cover 和 J. A. Thomas. 信息论要素。Wiley-Interscience，纽约，第二版，2006。40[71]R. G. James、J. R. Mahoney、C. J. Ellison 和 J. P. Crutchfield. 通往同步的多条路径：自然时间尺度及其算法。《物理评论 *E*》，89:042135，2014。43[72]N. F. Travers. 满足路径可合并条件的隐马尔可夫模型中熵率近似收敛的指数界。*Stochastic Proc.Appln.*，124(12):4149–4170，2014。43