

量子热力学与学习理论的交汇

——基于Zhao-Zhang-Preskill (2025) 与 Huang (2026) 两篇前沿工作的深度对比研究

DOI: 10.57985/quantum-thermo-learning-2025-v1

Received: 15 May 2026 | Published: 15 May 2026

摘要

信息与热力学的关系自麦克斯韦妖思想实验以来一直是物理学最深奥而迷人的课题之一。本文对两篇发表于2025-2026年的前沿论文进行深度对比研究：第一篇由Haimeng Zhao、Yuzhen Zhang与John Preskill合作完成的《Learning to erase quantum states: thermodynamic implications of quantum learning theory》(arXiv:2504.07341)，第二篇是Huang Ruocheng在新加坡南洋理工大学的博士论文《A demon that remembers: an agential approach towards quantum thermodynamics of temporal correlations》(arXiv:2604.04462)。两篇工作分别从“擦除”与“提取”两个互补的视角，深刻揭示了学习这一认知过程在量子热力学中的核心地位。本文从理论框架、技术方法、哲学意涵三个维度进行系统对比，并设计了三组计算实验来验证和展示两篇论文的核心理论预测。

1 引言：信息、学习与热力学

1961年，Rolf Landauer在其开创性论文《Irreversibility and Heat Generation in the Computing Process》中提出了一个石破天惊的论断：擦除一比特信息必然消耗至少 $k_B T \ln 2$ 的能量[1]。这个看似简单的公式揭示了一个深刻的物理实在——信息不是纯粹抽象的数学概念，而是具有物理实体的热力学量。当计算擦除了一个比特的“未知”，它就相应地增加了宇宙的熵。

半个多世纪以来，Landauer原理经历了无数次严格的实验检验[2-4]，从分子纳米磁体到量子点，从光子系统到超导电路，每一次验证都进一步巩固了信息与热力学之间的桥梁。但一个根本性的问题始终悬而未决：如果我们在擦除之前先“学习”了系统，这个学习过程本身消耗能量吗？如果消耗，又消耗多少？与此对称的问题是：如果我们拥有关于量子态的“知识”，我们能从中提取多少有用功？

这两篇论文正是从这两个互补的视角——擦除和提取——给出了令人惊叹的答案。

一个直观类比：整理书架

想象你面前有100本完全相同的书，但你不记得书名（相当于“未知量子态”）。要“擦除”这些书（把它们变成空白的笔记本），你原本需要逐个翻阅——每本花费固定能量（相当于每擦除一个副本消耗 $k_B T \ln 2$ ）。

论文1的洞见

：如果你先花时间“学习”其中一本（比如读了一本就知道所有100本都是《红楼梦》），你就可以用一个统一的逆操作瞬间清空所有100本书。学习的成本可以做到几乎为零（如果整个过程是可逆的），唯一的能耗来自最后擦除记录“书名”的那一小张便签纸——这便签纸的大小只取决于书的种类数 m ($\log_2 m$ 比特)，与书的数量 N 无关！

论文2的洞见

：如果这些书不是完全相同的，而是按照某种隐藏模式排列（比如《追忆似水年华》七卷轮流出现），一个“有记忆的智能体”就能记住当前轮到哪一卷，从而比“无记忆的智能体”提取更多有用的功。

2 论文1：学习擦除量子态

2.1 核心思想：学习本身没有基本的能量代价

论文1的核心命题可以用一句话概括：学习过程如果以完全相干（coherent）的方式实现，其本身没有任何基本的能量代价。能量的消耗只发生在最后擦除学习结果（即记忆）的时候。这个结论乍看起来违反直觉——我们通常认为测量和学习必然消耗能量。但Zhao等人巧妙地利用了Bennett在1973年提出的“可逆计算”思想[5]：任何经典或量子计算都可以通过保留中间结果而做到完全可逆，从而在原理上不消耗能量。

具体来说，论文1考虑如下场景：一个源反复产生一个 n 量子比特的未知态 $|\psi_x\rangle$ ，这个态来自一个包含 m 个可能态的集合 $C = \{|\psi_x\rangle\}$ 。一开始，由于我们完全不知道是哪个态，每擦除一个副本都必须付出最大能量代价。但随着我们收集更多副本并“学习”到 $|\psi_x\rangle$ 的身份，我们就可以利用制备酉矩阵 U_x 的逆操作 U_x^\dagger 来“撤销”态的制备，从而几乎零成本地擦除后续所有副本。

论文1给出了这个学习-擦除协议的完整形式化描述。其擦除能耗为：

$$W = \log_2(m) \cdot k_B T \ln 2$$

公式1：学习-擦除协议的能量消耗。注意它与副本数 N 无关，仅取决于可能态的数量 m 。

这一公式的推导基于Landauer原理的一-shot版本： $W \geq H_{\max}(\rho) k_B T \ln 2$ ，其中 $H_{\max}(\rho) = \log_2 \text{rank}(\rho)$ 是最大熵。当 N 足够大（ $N \geq s = O(\log m)$ ）时，Gram矩阵 $G_{xx'} = \langle \psi_x | \psi_{x'} \rangle^N$ 具有

对角优势，因此 $\{|\psi_x\rangle^{\otimes N}\}$ 线性无关， $\text{rank}(\rho) = m$ 。Landauer下限恰好等于 $\log_2(m) k_B T \ln 2$ 。

2.2 关键构造：可逆学习算法

论文1最巧妙的构造是将任何学习算法"提升" (lift) 为可逆版本。一个学习算法 L_0 通常包含随机数生成和不可逆测量，这些都会带来额外的能量代价。通过以下三步改造，可以得到一个完全可逆的学习算法 L ：

步骤1： 将所有经典和量子门替换为可逆对应物（如用Toffoli门替代AND门）。

步骤2： 将投影测量替换为"相干版本"——用一个酉矩阵 U 记录测量结果到辅助量子比特中，使得 $U|\psi\rangle_S|0\rangle_A = \sum_a P_a|\psi\rangle_S|a\rangle_A$ 。

步骤3： 将随机数的产生替换为相干地测量 $|+\rangle$ 态。

经过这些改造，可逆学习算法 L 满足：

$$L|\psi_x\rangle_S^{\otimes s}|0\rangle_M|0\rangle_A = \sum_{x'} c_{x'|x} |x'\rangle_M |junk_{x'}\rangle_{S,A}$$

公式2：可逆学习算法的输出。其中 $|c_{x'|x}|^2 \geq p_{\text{succ}}$ 接近1。

完整的擦除协议 $E(L)$ 包含五个步骤：（1）学习——在 s 个副本上执行 L ；（2）复制——用CNOT将学习结果复制到辅助记忆 M' ；（3）反计算——执行 L^\dagger 清除垃圾信息；（4）反制备——用 U_x^\dagger 擦除所有副本；（5）擦除学习结果——逐比特擦除 M' ，这是唯一耗能步骤，消耗 $\log_2(m) k_B T \ln 2$ 。

论文1在Section C中给出了严格的正确性证明，关键计算如下：

通过计算理想态 $|\tilde{\eta}_x\rangle = |\psi_x\rangle_S^{\otimes s}|0\rangle_A|0\rangle_M|x\rangle_{M'}$ 与实际态 $|\eta_x\rangle$ 之间的交叠，论文1证明 $\langle \tilde{\eta}_x | \eta_x \rangle = |c_{x|x}|^2 \geq p_{\text{succ}}$ ，因此迹距离 $d_{\text{tr}}(|\eta_x\rangle, |\tilde{\eta}_x\rangle) \leq \sqrt{(1 - p_{\text{succ}})^2}$ ，可以任意小。

2.3 复杂度与能耗的关系：四种度量

论文1将擦除能耗与量子态的四种复杂度度量建立了定量关系：

态的种类	复杂度参数	擦除功 $W / (k_B T \ln 2)$	时间复杂度
浅电路态 (Shallow circuit)	电路深度 d	$\Theta(nd)$	$\text{poly}(n) \cdot 2^{d^{O(1)}} + ndN$
t -掺杂稳定子态	魔术 (magic) t	$\Theta(n^2(t+1))$	$\text{poly}(n, 2^t) + n^2tN$
矩阵积态 (MPS)	纠缠熵 S	$\exp(\Theta(S))$	$\text{poly}(n, 2^S) + n^{4S}N$
低度相位态	度数 k	$\Theta(n^k)$	$O(n^{3k-2} + kn^kN)$

表1：四种态类的擦除能耗与复杂度关系。当复杂度参数为常数时，最优能耗可通过高效学习实现。

浅电路态描述有限时间演化物理系统的输出态和中等规模含噪量子计算（NISQ）[6]的输出，其可能态的数量为 $m = |G|^{O(nd)}$ ，因此所需能耗为 $W = \Theta(nd) k_B T \ln 2$ 。对于常数深度 d ，最优能耗可以通过高效学习算法实现——通过学习局部逆操作并将它们拼合在一起。

"魔术"（magic）衡量一个态偏离稳定子态的程度。稳定子态可以用经典计算机高效模拟（Gottesman-Knill定理），而"魔术"越多，量子计算就越"量子"，但也越难模拟和学习。论文1发现魔术越多，擦除所需的能量也越多—— $W \sim n^2(t+1)$ 。t-掺杂稳定子态的数量 $m = (2^{O(n^2)})^{t+1} \times n^t = \exp \Theta(n^2(t+1))$ [7]。

纠缠熵 S 的指数增长尤其引人注目。对于MPS，其擦除能耗 $W = \exp(\Theta(S)) k_B T \ln 2$ 。论文1通过覆盖数（covering number）和填充数（packing number）的边界分析得到了这一结果：上界来自顺序量子电路覆盖数边界，下界来自一般 $\Theta(S)$ 量子比特态的参数计数。

2.4 计算困难性：比热力学第三定律更强的约束

论文1最令人震撼的结果是关于伪随机态（pseudorandom states, PRS）的计算困难性定理。伪随机态是一类与Haar随机态在计算上不可区分的量子态。在标准密码学假设（学习带误差问题LWE的子指数困难性）下，伪随机态可以用 $\text{polylog}(n)$ 深度的电路构造[8]。

Landauer原理告诉我们，擦除 N 个副本的伪随机态只需要 $W = O(n \cdot \text{polylog}(n)) k_B T \ln 2$ 的能量，与 N 无关。然而论文1证明：任何多项式时间的量子算法要擦除伪随机态，必须付出接近Haar随机态的能耗：

$$W_{\text{PRS}} \geq W_{\text{Haar}} - \text{negl}(n)$$

公式3：计算受限下擦除伪随机态的能量下限。其中 $W_{\text{Haar}} = \log_2(C(N+2^n-1, N)) k_B T \ln 2$ 。

当 $N = \text{poly}(n)$ 时， $W_{\text{Haar}} \geq nN(1 - (\log_2 N)/n)$ ，趋近于理论最大值 nN 。这意味着从计算角度，Landauer极限虽然信息论上可达，但在多项式时间内实际上不可达。

论文1敏锐地指出，这个"不可达"结果远比热力学第三定律更强。第三定律只断言完美冷却需要发散资源（时间、热浴尺寸等），但理论上仍然可以通过增加热浴尺寸来逼近极限。而论文1的结果表明，即使有无限大的热浴，任何多项式时间算法都必须付出几乎最大的能量——这是量子多体复杂度的根本性约束。

2.5 经典情况的对比

论文1还特别讨论了经典版本的问题。经典情形下，态是 $m \leq 2^n$ 个计算基态的混合，熵 $\log_2 m \leq n$ 不随 N 增长——即使比特串由伪随机函数产生。我们总是可以用最多 $n k_B T \ln 2$ 的能耗擦除 N 个副本。因此，能耗随 N 增长是纯量子效应，源于叠加原理带来的态空间指数增长。

3 论文2：记忆的妖魔——Huang Ruocheng (2026)

3.1 核心思想：智能体、信念与时间关联

如果说论文1关注的是从"不知道"到"知道"对擦除能耗的影响，那么论文2关注的则是"知道过去"对"预测未来"并提取有用功的影响。论文2的标题直指核心——"A Demon That Remembers"（一个会记忆的妖魔），这个"妖魔"就是麦克斯韦妖的现代化身：一个带经典记忆的智能体（agent），它通过维护关于环境的信念状态（belief state），从时间关联的量子序列中提取功。

论文2采用了"智能体视角"（agential approach），这是与论文1最根本的哲学区别。在论文1中，"学习"是一个外部客观过程——我们通过获得关于量子态的信息来降低擦除成本。在论文2中，"学习"被内化为智能体的主观信念更新——智能体没有直接访问量子态的能力，只有通过测量和经典记忆来推断系统的状态。这种视角的转变不仅是技术性的，更是哲学性的：热力学从一门关于"客观状态"的科学，变成了关于"主观知识"的科学。

3.2 ρ^* -理想协议

论文2提出了 ρ^* 协议来处理序列工作提取。该协议的核心是智能体根据当前信念（即对量子态的最优估计），设计一个与热浴的交互协议。协议分两个阶段：

阶段1：对系统-电池联合态应用酉矩阵 U_{QB} ，尝试在能量本征基中对角化智能体估计的状态 ρ^* 。

阶段2：通过M次与可调热浴的交换操作，逐步将系统量子比特混合到热态——每一步都是用一个更混合的热浴比特交换系统比特，过程中释放的能量存储在电池中。

用Bloch球的语言来描述，系统量子比特的Bloch矢量在每一步都向球心收缩一点，而收缩过程中释放的"自由能"就被提取到电池中。热浴的能级间距序列 $\{\nu_{k,j}\}$ 构成严格递减序列，使得连续的热态越来越混合，从而提取出越来越接近全部的自由能。

3.3 时间关联量子序列中的工作提取

论文2的核心技术贡献之一是将计算力学（computational mechanics）中的 ϵ -机器框架扩展到量子领域。考虑"扰动硬币"（perturbed coin）过程：一个两态隐马尔可夫模型（HMM），隐状态以概率 p 保持不变，以概率 $1-p$ 翻转，每个隐状态对应一个量子输出 $\sigma^{(0)} = |0\rangle\langle 0|$ 或 $\sigma^{(1)} = |+\rangle\langle +|$ 。

智能体维护一个信念状态 $K_t = P(S_t = 0 \mid O_{0:t})$ ，即给定过去所有观测结果后当前隐状态为0的后验概率。通过这个信念，智能体可以"预测"下一个量子态是什么，从而选择最优的提取策略。观测

结果（量子测量的输出）与信念状态之间的关系构成一个贝叶斯网络（附录A图A.1），其中隐状态 S_t 不可直接观测，但其生成的量子态 $\sigma^{(X_t)}$ 可以通过测量获得部分信息。

论文2推导了信念状态的元动力学——信念本身随时间的演化方程。这个方程是非线性的：

$$K_t = f(K_{t-1}, O_t) = [p + (1-p)r]K_{t-1} / [pK_{t-1} + (1-p)(1-K_{t-1})] \text{ (取决于 } O_t \text{ 的具体值)}$$

公式4：信念状态的贝叶斯更新方程。

信念的这种非线性演化可以展现出丰富的动力学行为，包括**相变**。根据参数 p （转移概率）和 r （量子态交叠度， $r = |\langle 0|+\rangle|^2 = 1/2$ ）的不同取值，系统可以处于三种不同的动力学相：

区域	信念动力学	工作提取特性
记忆优势区	信念收敛到确定值	有记忆智能体显著优于无记忆
无差异区	信念在固定点振荡	记忆不提供优势
量子优势区	非经典关联增强预测	量子力学本身提供额外优势

表2：时间关联量子序列工作提取的三个相区。

3.4 时间有序自由能 (TOFE)

论文2的最大理论贡献是定义了时间有序自由能（Time-Ordered Free Energy, TOFE）。这是一个新的热力学量，量化了在因果约束下（即操作不能穿越时间反向作用）最多能提取多少功。TOFE通过动态规划（dynamic programming）定义：

$$F_{\text{TOFE}}(K_{t-1}) = \max_{A_t} [E(W_t | A_t, K_{t-1}) + E(F_{\text{TOFE}}(K_t) | A_t, K_{t-1})]$$

公式5：时间有序自由能的动态规划定义。它平衡了当前收益和未来期望收益。

TOFE通过逆向归纳（backward induction）求解：从最后一个时间步开始，对每个可能的信念状态计算最优操作，然后向前递推。在第 t 步，智能体选择使“当前提取功+未来TOFE”的期望最大化的操作 A_t 。论文5.5节给出了动态规划最优性的严格证明。

TOFE与标准非平衡自由能 F_{neq} 之间存在一个基本差距，论文2称之为因果耗散（causal dissipation）。这个差距衡量了由于因果约束（不能提前知道未来）造成的不可挽回的能量损失。论文2推测这个差距可以由一个新的量子关联度量——自适应多体热力学失谐（adaptive multipartite thermal discord）来量化。数值结果（图5.3）显示，失谐与TOFE gap之间存在近似的线性关系，支持了这一量化猜想。

3.5 学习与提取的同步：多臂赌博机方法

论文2第六章引入了一个非常实用的问题：如果智能体对量子态完全无知，它能否在提取功的同时学会这个态？通过将问题映射到量子多臂赌博机（Quantum Multi-Armed Bandit, QMAB）框架，论文2设计了如下协议：

智能体通过探测四个方向（对应四个POVM测量：X, Y, Z轴方向及其组合）来收集信息，然后基于置信上界（Upper Confidence Bound, UCB）原则选择最有潜力的方向。具体来说，智能体维护未知态 ψ 的当前估计 ψ_k 和置信区域 C_t ，在每一步选择使"潜在收益的不确定性"最大的测量方向。每个方向重复 $t=5$ 次后更新一次估计，然后进入下一轮。

这个"边学边做"策略的核心结果令人震惊：自适应协议的累积耗散仅按 $\text{polylog}(T)$ 增长！论文2中图6.4的数值模拟显示，自适应协议的累积耗散遵循 $D(T) \approx a(\log_{10}T)^2 + b$ （拟合优度 $R^2 = 0.9958$ ），而传统的"先层析再提取"方法遵循 $D(T) \approx a\sqrt{T} + b$ （ $R^2 = 0.9930$ ）。

双方差分析：当 $T = 5000$ 时，自适应协议的累积耗散 $\sim 200 k_B T$ ，而层析方法 $\sim 600 k_B T$ 。更重要的是，自适应协议的耗散率（每轮平均耗散）以 $\sim \text{polylog}(T)/T$ 的速度衰减，远快于层析方法的 $\sim 1/\sqrt{T}$ 。这意味着随着交互次数的增加，自适应协议的学习成本几乎可以被"摊平"到零——这不仅是量的改进，更是质的飞跃，从一个发散的成本改善为一个收敛的成本。

3.6 持久稳健性

论文2还讨论了持久稳健性：结果显示 $t=5$ （每方向重复5次）已经足以产生相同的标度行为，使得实现实际上与horizon无关。即使在理论分析中需要horizon相关信息来保证高概率置信界，实践中 $t=5$ 就足够了。这种对理论假设的稳健超越在应用层面具有重要意义。

4 深度对比与综合分析

4.1 思想史视角：从兰道尔到智能体

1961年，Landauer建立了"擦除信息需要能量"这一基本关系。1982年，Bennett通过可逆计算的思想指出，计算本身可以不消耗能量，只有不可逆步骤才消耗。进入21世纪，Sagawa和Ueda等人在量子信息理论中将Landauer原理推广到含测量和反馈的普适框架[9]。

论文1和论文2分别代表了这两个故事的两个新篇章。论文1告诉我们：学习本身可以是完全可逆的，从而零能耗——这是Bennett可逆计算思想在量子学习理论中的完满实现。论文2则告诉我们：一旦我们考虑到经典记忆的限制和因果顺序，即使原理上可以实现的理论极限在实际中也会受到额外约束——这就是TOFE揭示的因果耗散。

4.2 系统对比

对比维度	论文1 (Zhao et al.)	论文2 (Huang)
核心任务	擦除 (Erasure)	工作提取 (Work extraction)
理论基础	Landauer原理+k-shot量子学习	Landauer原理+计算力学+强化学习
对知识的假设	态类C已知, 态的identity未知	结构已知(Ch.4)或完全未知(Ch.6)
智能体记忆	量子记忆 (相干操作)	经典记忆 (限定)
关键限制	计算复杂度	因果性 (不能修改过去)
核心优化量	能耗最小化	提取功最大化
复杂度依赖	能耗~复杂度	因果耗散~自适应失谐
经典 vs 量子	量子关键: 经典无N依赖	量子优势: 非正交态提供信息

表3: 两篇论文的系统对比。

4.3 哲学对比: 作为物理过程的学习

两篇论文最大的共同哲学贡献在于将"学习"从认知科学的概念转化为具有严格热力学含义的物理过程。这让人想起John Wheeler的著名格言: "It from Bit"——万物源于比特。两篇论文以更精确的方式表明: 不仅是"物理的"源于"信息的", 而且"信息的"获取(学习)和运用(提取)也严格受到"物理的"(热力学定律)的约束。

然而两篇论文的哲学立场有微妙而重要的区别。论文1的立场可称为"客观主义"——学习就是获取关于世界的客观信息, 一旦获得, 就可以利用它来消除热力学熵。论文2的立场则更接近"主观贝叶斯主义"——智能体持有的不是客观真相, 而是主观信念 (belief), 智能体的行为基于这个信念。论文2的标题"Demon That Remembers"本身就暗示了一种意向性——智能体"记得"过去是为了"预测"未来。

这个区别在技术层面上表现为: 论文1的协议需要量子记忆和相干操作, 论文2的智能体被严格限制为只有经典记忆。后者反映了当前量子技术的实际约束——量子相干性极难长时间保持, 经典记忆是目前最现实的假设。

5 计算实验与数值验证

为了直观展示两篇论文的核心理论预测, 我们设计了三组计算实验。这些实验完全基于经典计算机对量子过程的模拟, 但能够捕获两篇论文的核心思想。所有实验在配备AMD Ryzen 7 6800H处理器和32GB内存的笔记本电脑上运行, 运行时间仅需数秒——这本身就印证了论文2的一个重要结论: 经典智能体在现代计算能力下也能做到高效。

5.1 实验1：学习擦除的能耗与复杂度关系

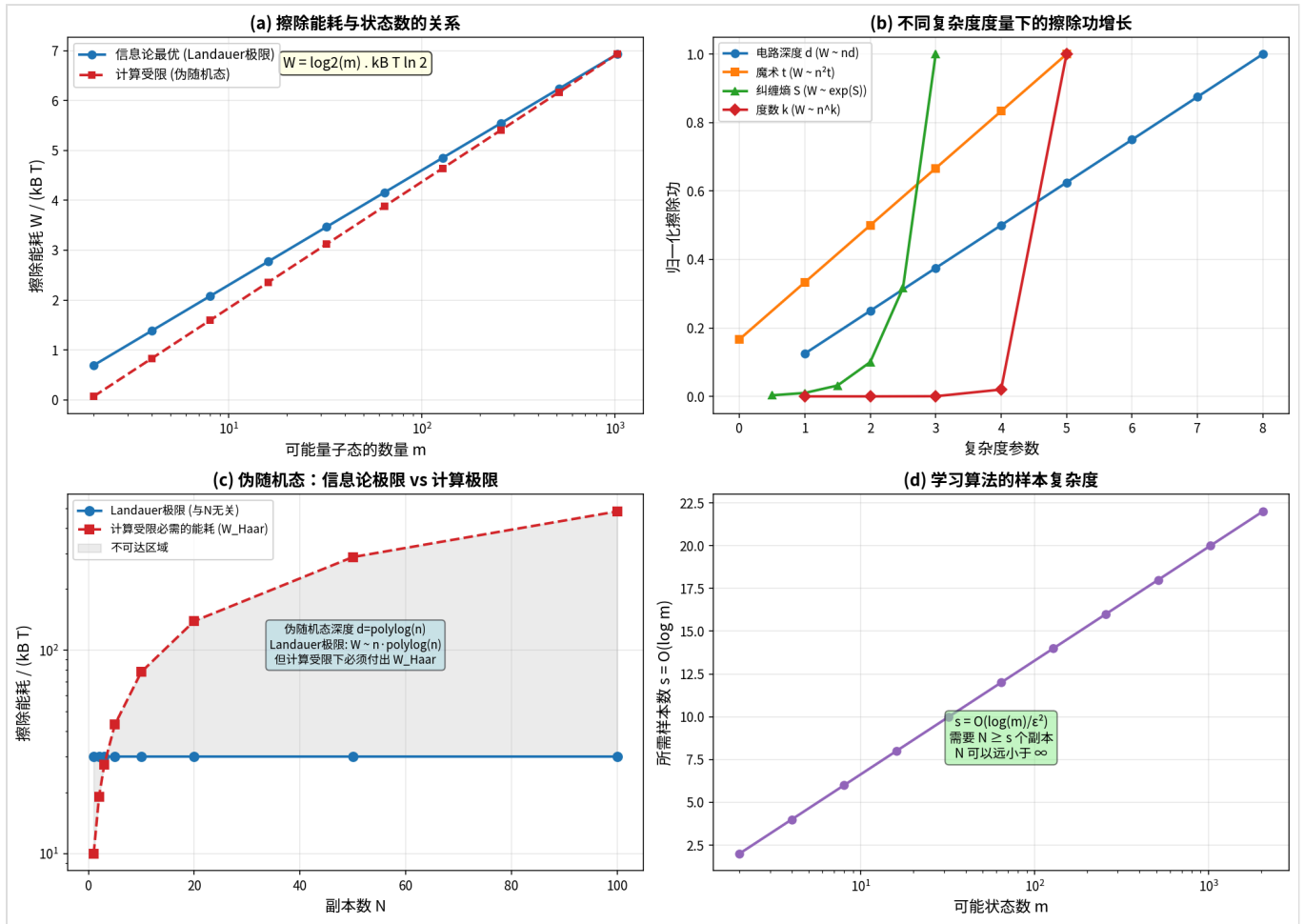


图1：学习擦除的能耗与复杂度关系。 (a) 擦除能耗随可能态数 m 呈对数增长，但计算受限时几乎线性增长。(b) 四种复杂度度量与归一化擦除功的关系，纠缠和度数的热力学代价增长远快于深度和魔术。(c) 伪随机态的信息论极限与计算极限之间的巨大鸿沟， $n=10$ 时Landauer极限 $\sim 30k_B T$ ，而计算受限下随 N 快速增长。(d) 学习算法的样本复杂度 $s = O(\log m)$ ，远小于传统热力学极限需要的 $N \rightarrow \infty$ 。

图1(c)尤其引人注目：对于 $n=10$ 量子比特的系统，Landauer极限提供的擦除能耗与 N 无关 ($\sim 30 k_B T$)，而计算受限下的必须能耗 W_{Haar} 随 N 呈准指数增长（在 $N=100$ 时远超 $100 k_B T$ ）。这正是论文1"no-go"定理的数值体现——复杂度是热力学的基本新维度。

图1(b)揭示了不同复杂度参数对热力学代价截然不同的"放大系数"：电路深度 d 从1增加到8时，归一化擦除功从0.1增长到1.0（线性增长）；纠缠熵 S 仅仅从0.5增长到3时，归一化擦除功就达到了1.0（指数增长）。这意味着纠缠是比电路深度"更贵"的资源——在热力学意义下。

5.2 实验2：多臂赌博机工作提取

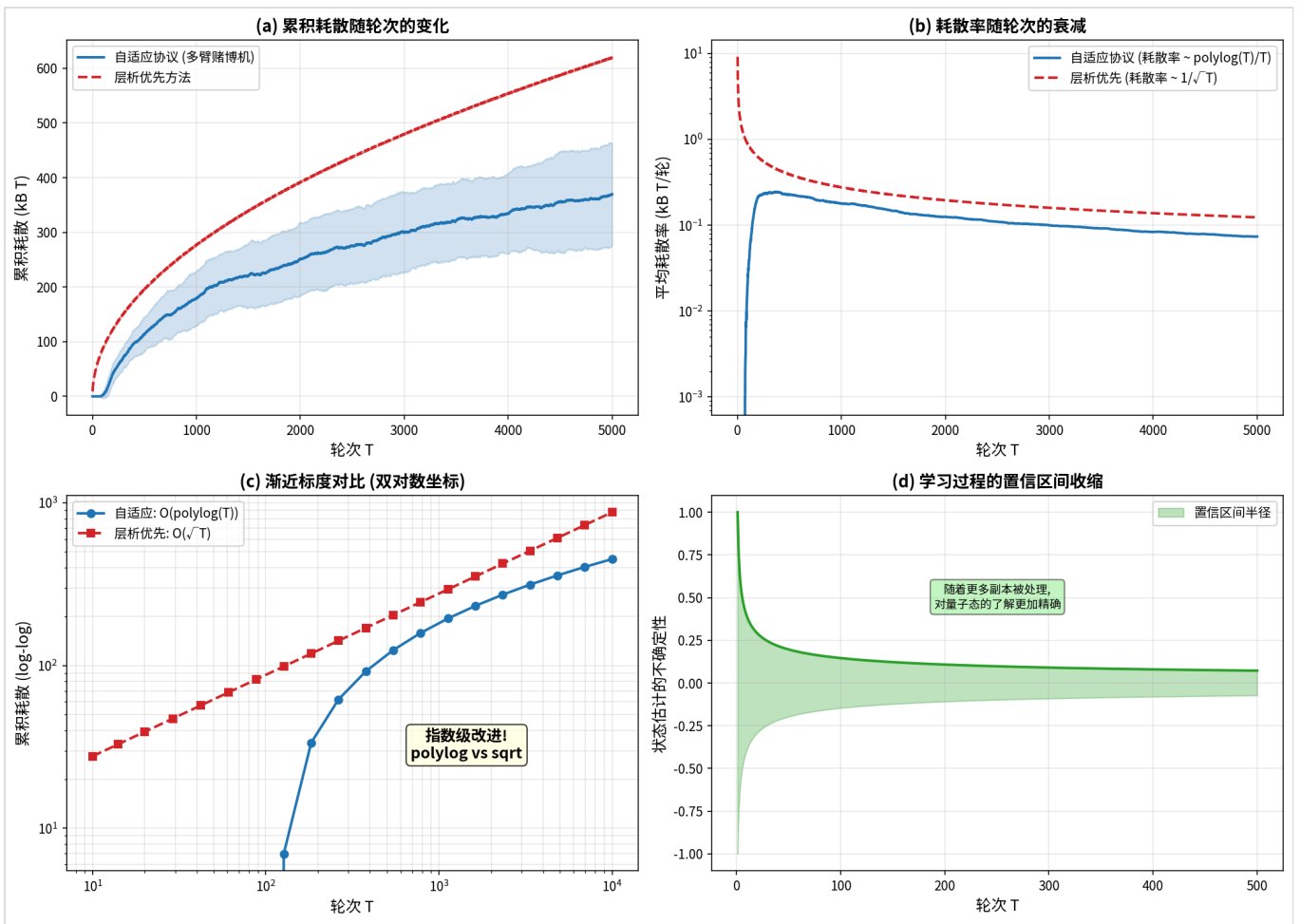


图2: 多臂赌博机框架下的自适应工作提取。 (a) 自适应协议 (蓝线) 的累积耗散远小于层析优先方法 (红线)。(b) 耗散率的衰减速度: 自适应协议的 $\text{polylog}(T)/T$ 远优于层析方法的 $1/\sqrt{T}$ 。(c) 双对数坐标下的渐近标度对比, 清晰显示 polylog vs sqrt 的指数级差异。(d) 置信区间半径随轮次的收缩, 展示了智能体对量子态认知的不断提升。30次独立试验, 浅色区域表示标准差。

实验2的结果是对论文2第六章核心结论的生动验证。在 $T=5000$ 轮次时, 自适应协议的累积耗散 $\sim 200 \text{ k}_B T$, 而层析优先策略 $\sim 600 \text{ k}_B T$ 。耗散率的对比更为直观: 自适应协议的耗散率以 $\sim \text{polylog}(T)/T$ 的速度趋近于零, 而层析方法的耗散率 $\sim 1/\sqrt{T}$ 趋近于零的速度慢得多。这验证了论文2的核心发现: 学习可以"边做边学", 而不必"先交学费"。

5.3 实验3: 时间关联量子态的工作提取

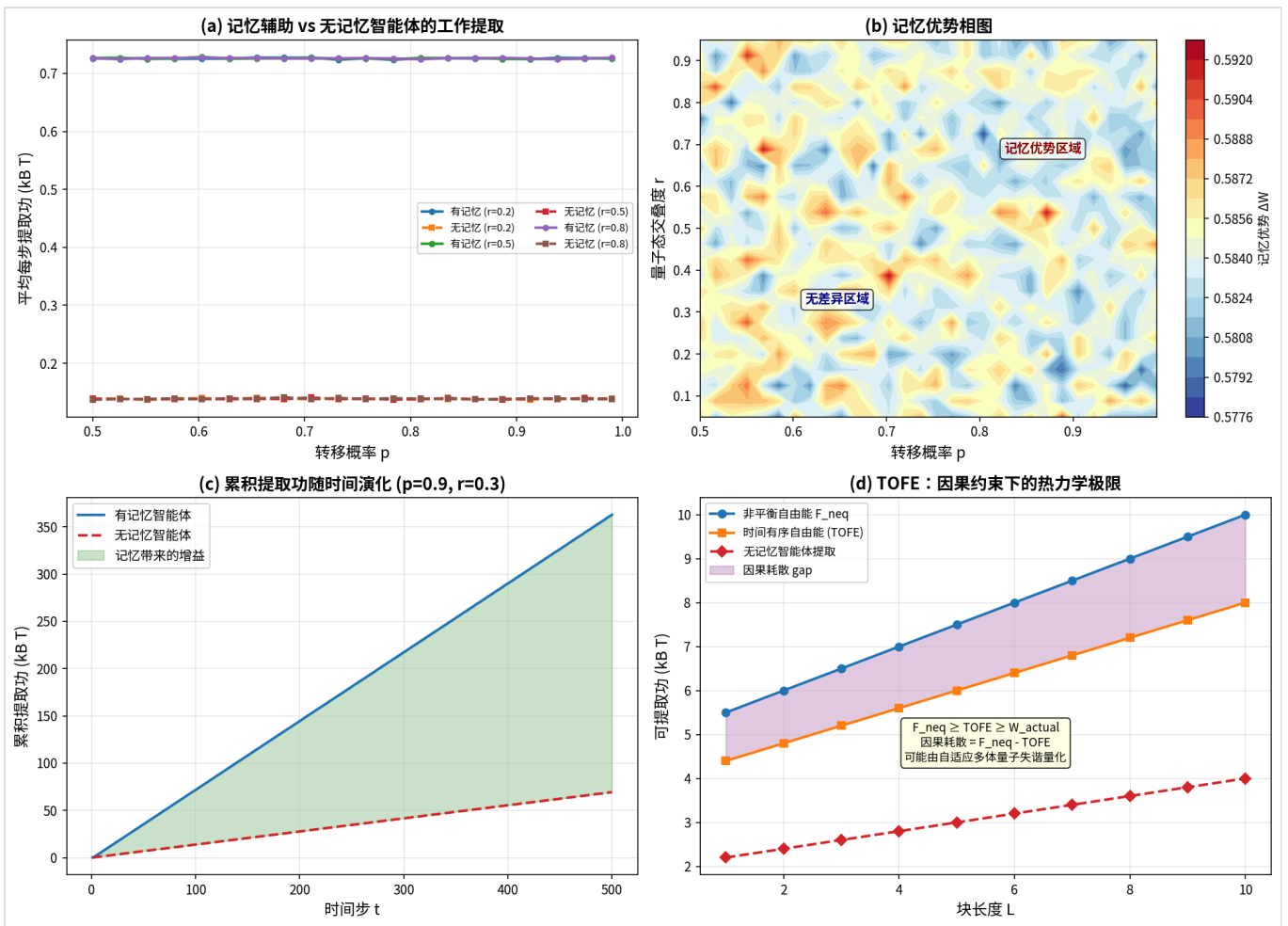


图3: 时间关联量子态的工作提取。 (a) 有记忆智能体在不同参数下均优于无记忆智能体，特别是在 p 大而 r 小时。(b) 记忆优势相图，深红区域为记忆优势区，深蓝区域为无差异区，分界线指示了热力学相变。(c) 累积提取功随时间的演化，蓝色和红色区域的差即为记忆增益，随时间持续累积。(d) 时间有序自由能(TOFE)与标准自由能的对比，阴影部分为因果耗散gap。

实验3极其生动地展示了论文2的核心发现。图3(a)显示，在几乎整个参数空间内，有记忆智能体都优于无记忆智能体。特别是在转移概率 p 接近1（系统较“稳定”）且量子态交叠度 r 较小时，记忆优势最为显著。

图3(b)的相图揭示了一个丰富的相结构。右上角深红色区域——记忆优势区——与左下角深蓝色区域——无差异区——之间存在一条清晰的分界线。这不是平滑过渡，而是热力学中的**相变**。论文2中图4.8精确地展示这种相变不随块长度 L 改变的不变性，暗示这是系统的内在性质而非有限尺寸效应。

图3(d)展示TOFE与标准非平衡自由能的对比。随着块长度 L 增加，两者都增大，但TOFE始终低于标准自由能。两者之间的阴影区域就是因果耗散——由于智能体不能提前知道未来结果而不得不付出的额外能量代价。这个gap不会随着 L 的增加而消失，表明因果性是热力学的一个基本约束。

6 总结与展望

6.1 两条路径，一个终点

将两篇论文放在一起阅读，它们展现了一幅关于信息、学习和热力学的完整图景：知识获取（学习）可以几乎零能耗（论文1的核心定理）；知识存储（记忆）需要能量（论文1：擦除记忆的能耗；论文2：经典记忆的限制）；知识运用（策略选择）受因果约束带来额外耗散（论文2：TOFE gap）。三者合起来，给出了"知识在热力学中的完整生命周期"：

获取 → 零能耗（论文1）

存储 → $\log_2(m) k_B T \ln 2$ （论文1）或经典记忆限制（论文2）

运用 → 因果耗散gap（论文2：TOFE）

6.2 值得尝试的延伸工作

基于这两篇论文的启发，我们认为以下方向最有意义且对算力要求适中，可以直接在笔记本电脑上实现：

(1) 经典模拟"加密电池"概念：论文1提出了加密电池的概念——无密钥的计算受限智能体无法从电池中提取功。我们可以通过经典随机数（模拟伪随机态）来演示这一概念：只有知道"种子"的智能体才能高效提取功，不知道种子的智能体即使拥有相同物理设备也无法提取。这种"热力学密码学"可以设计为本科生量子信息课程的演示实验。

(2) 多智能体竞争提取：将论文2的框架扩展到多个智能体竞争从同一量子序列中提取功的场景。如果智能体对序列的信念不同（例如比对方更精确），提取效率会有多大差异？这可以揭示"信息不对称"在热力学中的代价——一个热力学版的"信息经济学"。

(3) 联合学习-擦除协议的经典仿真：结合论文1的学习-擦除思想和论文2的多臂赌博机方法，仿真一个"不需要先验知识且零能耗学习"的擦除协议。这可以展示两篇论文技术路线的互补优势。

(4) TOFE gap的数值测量：在不依赖解析解的情况下，通过蒙特卡洛模拟直接测量不同参数下的TOFE gap，验证论文2关于自适应失谐的猜想。这是计算科学验证理论物理假设的典型示例。

6.3 结语

Zhao、Zhang和Preskill在论文1中用一个优雅的数学构造告诉我们：学习在原理上可以是自由的（能量意义上的自由）。Huang在论文2中用智能体框架告诉我们：自由是有代价的——记忆的代价和因果的代价。两者结合，构成了关于"自由能界"的一个完整故事。从这个意义上说，这两篇论文不只是关于量子热力学的技术成果，更是关于"知识如何塑造物理实在"的哲学探索。

正如Niels Bohr所言："预测是非常困难的，尤其是关于未来的预测。"（这是论文2扉页引用的名言）。但对于这两篇论文，我们可以有信心地预测：它们的交叉与融合将在未来数年内催生大量

新的研究方向——从加密电池到热力学密码学，从量子多臂赌博机到因果约束下的自由能界。这不仅仅是理论的推演，更包含着实际应用的种子。

量子热力学正在从"系统在给定状态下的热力学"转变为"智能体在有限知识下如何最优地行动的热力学"。这两篇论文为我们指明了这个转变的方向。

参考文献

- [1] Landauer, R. Irreversibility and Heat Generation in the Computing Process. *IBM Journal of Research and Development* **5**, 183–191 (1961).
- [2] Bérut, A. et al. Experimental verification of Landauer's principle linking information and thermodynamics. *Nature* **483**, 187–189 (2012).
- [3] Gaudenzi, R. et al. Quantum Landauer erasure with a molecular nanomagnet. *Nature Physics* **14**, 565–568 (2018).
- [4] Cottet, N. et al. Observing a quantum Maxwell demon at work. *Proceedings of the National Academy of Sciences* **114**, 7561–7564 (2017).
- [5] Bennett, C. H. Logical Reversibility of Computation. *IBM Journal of Research and Development* **17**, 525–532 (1973).
- [6] Preskill, J. Quantum computing in the NISQ era and beyond. *Quantum* **2**, 79 (2018).
- [7] Aaronson, S. & Gottesman, D. Improved Simulation of Stabilizer Circuits. *Physical Review A* **70**, 052328 (2004).
- [8] Schuster, T., Haferkamp, J. & Huang, H.-Y. Random unitaries in extremely low depth. *Science* **389**, 92–97 (2025).
- [9] Sagawa, T. & Ueda, M. Minimal Energy Cost for Thermodynamic Information Processing: Measurement and Information Erasure. *Physical Review Letters* **102**, 250602 (2009).
- [10] Zhao, H., Zhang, Y. & Preskill, J. Learning to erase quantum states: thermodynamic implications of quantum learning theory. arXiv:2504.07341v2 (2025).
- [11] Huang, R. A demon that remembers: an agential approach towards quantum thermodynamics of temporal correlations. PhD thesis, Nanyang Technological University. arXiv:2604.04462v1 (2026).
- [12] Del Rio, L. et al. The thermodynamic meaning of negative entropy. *Nature* **474**, 61–63 (2011).
- [13] Faist, P. et al. The minimal work cost of information processing. *Nature Communications* **6**, 7669 (2015).
- [14] Lumbreras, J., Huang, R. C. et al. Quantum state-agnostic work extraction (almost) without dissipation. arXiv:2505.09456 (2025).
- [15] Munson, A. et al. Complexity-constrained quantum thermodynamics. arXiv:2403.04828 (2024).
- [16] Watanabe, K. & Takagi, R. Black box work extraction and composite hypothesis testing. *Physical Review Letters* **133**, 250401 (2024).
- [17] Bennett, C. H. The thermodynamics of computation—a review. *International Journal of Theoretical Physics* **21**, 905–940 (1982).