arXiv:2602.02165v1 [quant-ph] 2 Feb 2026

AQER：面向数字量子计算机的可扩展高效数据加载器

Kaining Zhang[1], **Xinbiao Wang**[1],**Yuxuan Du**[1, 2*]、**Min-Hsiu Hsieh**[3*]、**Dacheng Tao**[1*]南洋理工大学生成式人工智能实验室、计算与数据科学学院，新加坡 639798[2]南洋理工大学物理与数学科学学院，新加坡 639798[3]鸿海（富士康）研究院，台北，台湾
yuxuan.du@ntu.edu.sg,min-hsiu.hsieh@foxconn.com,dacheng.tao@ntu.edu.sg

2026年2月3日

摘要

数字量子计算有望提供超越经典系统能力范围的计算能力，但其发展常受量子资源匮乏的制约。在此背景下，如何高效地将经典数据或量子数据加载到量子电路中，是一个关键瓶颈。近似量子加载器（AQLs）通过平衡保真度与电路复杂度，为这一问题提供了可行的解决方案。然而，现有的大多数 AQL 方法要么依赖启发式方法，要么仅针对特定输入类型提供保证，仍缺乏通用的理论框架。为弥补这一空白，我们在此将大多数 AQL 方法重新表述为一个统一框架，并建立其近似误差的信息论界。我们的分析表明，当加载电路作用于目标态时，制备态与目标态之间可达到的非保真度会随着各子系统间总纠缠熵的增加而线性增长。基于这一发现，我们提出 AQER，这是一种可扩展的 AQL 方法，通过系统性地减少目标态中的纠缠来构造加载电路。我们使用合成数据集、经典图像和语言数据集，以及最多包含 50 个量子比特的量子多体态数据集，开展系统实验以评估 AQER 的有效性。结果表明，AQER 在准确度和门效率方面均持续优于现有方法。我们的工作为可扩展的量子数据处理和实际量子计算应用铺平了道路。

1 引言

数字量子计算机 [1]有望在众多领域展现出超越经典系统的计算能力 [2–7]。经过四十年的探索 [8–16], , 我们已取得了重大进展，超导平台 [17]和中性原子平台 [18] 已在合成基准测试中展现出量子优势。然而，在可预见的未来，可用的量子资源仍将受到严重限制，例如高质量量子比特的数量和相干时间 [19]。这一局限凸显了持续提升量子计算三大基础模块效率的必要性：量子态制备 [20, 21], 、量子处理 [22, 23], 和读出 [24, 25]。最大限度地利用这些阶段中的资源，对于提升量子计算机的实际效用至关重要。在此背景下，高效的量子态制备——即构建用于将经典或量子输入编码为量子态的量子门序列——成为一项关键的先决条件

*通讯作者。

。这一环节一直以难度极高著称，因为理论结果表明，在最坏情况下，要在可证明的误差容限内制备任意量子态，可能需要指数级数量的量子门或辅助量子比特 [27, 28]。

近年来，一种称为 近似量子加载器（AQL） [29] 的新概念应运而生，为高效量子态制备开辟了一个颇具前景的方向。与以实现可证明的准确度保证为目标的早期方法不同，AQL 接受制备保真度与电路复杂度之间的权衡。这一范式的核心洞见是，许多量子算法能够容忍输入态存在一定的不精确性，因此可以大幅减少量子门数量和资源开销。例如，在量子机器学习中，输入特征的微小扰动通常对分类准确率影响甚微 [30] ，也不会损害量子优势的展现，尤其是在样本复杂度方面 [31]。受这一观察启发，人们投入了大量精力，为各种数据类型设计高效的 AQL 方案。这些方法大致分为两类：基于张量网络（TN）的方法 [32, 33], ，采用矩阵乘积态（MPS）及其他 TN 表示；以及基于电路的方法 [34, 35], ，直接优化量子门序列（一般框架见图 1，详情见第 3.1 节）。尽管取得了这些进展，现有的大多数技术要么依赖启发式方法，要么仅在限制性情形下提供理论保证。因此，对于 *AQL* 所能达到的近似误差的基本极限，目前仍缺乏系统认识。解决这一问题将为 AQL 算法的规范化开发以及数字量子计算机的资源高效利用提供关键见解。

在此，我们通过首先将多种 AQL 方法重新表述为一个统一框架，弥补上述知识空白。图 1 展示了其中的直观思路：AQL 构造实质上是寻找一系列（可调节的）量子门，使该序列演化得到的量子态与目标量子态之间的距离最小。基于这一框架，我们推导出 AQL 近似误差的**信息论下界和上界**（定理 3.1），且这些界不依赖于具体的 AQL 策略。具体而言，我们证明，近似误差随新提出的纠缠度量线性下降；该度量等于目标量子态经过 AQL 门序列的逆序列作用后，各单量子比特纠缠熵之和。这一结果提供了关键

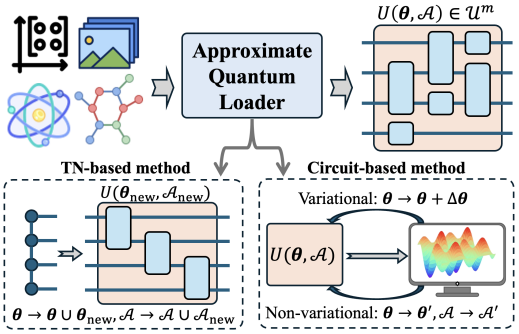

图 1：AQL 的一般框架。典型的 AQLs 分为两类：基于张量网络的方法和基于电路的方法，两者都旨在从给定的门集 $\mathcal{U}$ 构造量子电路，以近似制备目标态。

这一发现表明，AQL 的性能从根本上取决于 AQL 过程中可实现的纠缠约化程度；纠缠约化越多，加载就越准确。

受纠缠在决定 AQL 性能方面的理论重要性启发，我们提出了 AQER，这是一种可扩展且高效的 AQL 方法，它以最大纠缠约化为原则来构建门序列。具体而言，AQER 利用这一原则，通过逐步添加参数化的单量子比特门和双量子比特门来降低目标态的纠缠，并采用显式构造的单量子比特门序列进一步降低近似误差。与现有的 AQL 方法相比，AQER 具有**两大优势**。首先，AQER 灵活且通用，支持高效地近似加载经典数据和未知量子态。其次，AQER 稳健且易于优化。在纠缠约化优化的指导下，它不仅能实现较低的近似误差，还能缓解参数训练 [36], 期间的梯度消失问题，从而确保其可扩展至大规模量子比特系统。为验证 AQER 的有效性，我们开展了系统实验，对其在不同量子态加载任务中的表现进行基准测试，任务涵盖合成量子态、现实世界图像和语言数据集，以及多达 50 个量子比特的量子多体系统。实验结果

表明 AQER 的表现始终优于现有方法。我们的工作为可扩展的量子数据处理，以及量子计算在现实任务中的实际应用铺平了道路。

总而言之，我们的贡献主要有三点。(i) 我们为范围广泛的 AQL 方法提出了一个统一框架，并推导出两个关于 AQL 近似误差与纠缠度量之间关系的信息论界。据我们所知，这是首项从信息论角度确立 AQL 理论极限的研究。(ii) 受这些理论结果的启发，我们开发了 AQER，这是一种可扩展高效的 AQL 方法，通过有理论依据的纠缠缩减优化，高效利用量子门资源。(iii) 我们针对多种数据集开展了大规模数值模拟，数据集规模最高达 50 个量子比特，并将 AQER 与参考 AQL 方法进行比较。结果验证了 AQER 的有效性，表明其性能更优，近似误差更低，量子门数量更少。相应代码可在 GitHub 获取，以便复现和基准测试。

2 预备知识

本节介绍量子计算、量子纠缠和变分量子算法的基础知识，随后回顾典型的 AQL 方法。更多详情请参阅附录A。

量子计算基础。量子比特的纯态可写为 $|\psi\rangle = a|0\rangle + b|1\rangle$，其中 a、$b \in \mathbb{C}$ 满足 $|a|^2 + |b|^2 = 1$，且 $|0\rangle = (1, 0)^T$、$|1\rangle = (0, 1)^T$。N量子比特空间由 N 个单量子比特空间的张量积构成。对于纯态 $|\psi\rangle$，相应的密度矩阵定义为 $\rho = |\psi\rangle\langle\psi|$，其中 $\langle\psi| = (|\psi\rangle)^\dagger$。混合态由如下形式的密度矩阵表示：$\rho = \sum_k c_k|\psi_k\rangle\langle\psi_k|$，其中 $c_k \geq 0$ 且 $\sum_k c_k = 1$。对于 N量子比特态 $|\psi\rangle$，子系统 $A \subseteq [N]$ 的约化密度矩阵可通过偏迹 $\rho_A = \mathrm{Tr}_{[N]\setminus A}[|\psi\rangle\langle\psi|]$求得。量子门由作用于量子态的酉矩阵表示，并可在电路模型中按照量子电路表示法绘出。典型的量子操作包括固定量子门，例如 $CZ := \mathrm{diag}(1, 1, 1, -1)$，以及可调的单量子比特旋转门 $R_\sigma(\theta) = e^{-i\theta\sigma/2}$，其中 $\sigma \in \{X, Y, Z\}$ 和 $X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ $Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ $Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$分别为泡利X、Y、Z算子。旋转门可推广到双量子比特情形，例如 R_{ZZ}，其中 $ZZ = Z \otimes Z$。量子测量是从量子态中提取经典信息的过程。数学上，它由一个称为可观测量的厄米矩阵 H 描述。使用可观测量 H 测量量子态 $|\psi\rangle$ 或 ρ，会得到一个随机变量，其期望值分别为 $\langle\psi|H|\psi\rangle$ 或 $\mathrm{Tr}[H\rho]$。为了量化两个量子态之间的相似程度，我们使用保真度，定义为 $F = |\langle\psi_1|\psi_2\rangle|^2$ 或 $F = \mathrm{Tr}[\rho_1\rho_2]$；相应的非保真度为 $1 - F$ [37]。这两者都可以通过将 $|\psi_1\rangle$ 或 $|\psi_2\rangle$ 的密度矩阵作为可观测量来计算。

量子纠缠。量子纠缠是量子系统的一种独特性质，使其区别于经典系统，也是实现量子计算优势的关键量子资源。具体而言，若复合系统的 N 量子比特量子态 $|\psi\rangle$ 无法表示为两个子系统的量子态的张量积，即对于各自子系统中的任意量子态 $|\psi_A\rangle$，$|\psi_B\rangle$ 均不成立，则称该量子态在两个子系统 A，$B \subset [N]$ 之间纠缠。人们已提出多种度量来量化量子态的纠缠程度。常用的一种度量是 Rényi-2 熵，定义为 $\mathcal{S}_A(|\psi\rangle) = -\log_2 \mathrm{Tr}[\rho_A^2]$，其中 ρ_A 是 $|\psi\rangle$ 在子系统 A上的约化密度矩阵。当 $\mathcal{S}_A(|\psi\rangle) = 0$时，量子态 $|\psi\rangle$ 是可分的，并可写成乘积态 $|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$。

变分量子算法（VQA）。变分量子算法 [38] 是一种量子-经典混合范式，其中参数化量子电路（PQC）由经典优化器进行训练。它已成为量子机器学习的核心框架 [39, 40]，，应用涵盖判别式学习、生成学习 [41–43], 以及强化学习 [44–46]。具体而言，参数化量子电路可写为$V(\boldsymbol{\theta}) = V_m(\boldsymbol{\theta}_m) \cdots V_1(\boldsymbol{\theta}_1)$，参数 $\boldsymbol{\theta}$ 被优化，以最小化相对于输入态 ρ_{in} 和可观测量 O定义的代价函数，即，$f(\boldsymbol{\theta}) = \mathrm{Tr}\left[O\,V(\boldsymbol{\theta})\rho_{\mathrm{in}}V(\boldsymbol{\theta})^\dagger\right]$。参数由梯度下降或Adam等经典优化器更新，这需要获取（精确或估计的）梯度。当变分电路在经典计算机上模拟时，梯度可通过自动微分直接计算 [47]。当在量子设备上执行时，梯度通常由参数移位规则估计 [48]：$\frac{\partial f}{\partial \theta_j} = \frac{1}{2}f(\boldsymbol{\theta}_+) - \frac{1}{2}f(\boldsymbol{\theta}_-)$，其中 $\boldsymbol{\theta}_\pm$ 与 $\boldsymbol{\theta}$ 在第 j个参数上相差 $\pm\pi/2$。

2.1 相关工作

Exi面向AQL的sting方法可大致分为基于TN的方法和基于电路的方法 .

基于张量网络的方法。 基于张量网络的方法利用张量网络高效表示并制备低纠缠量子态。例如，键维数为 k的矩阵乘积态（MPS）可用 $\mathcal{O}(Nk^2)$ 个双量子比特门 [49], 精确制备，而对于具有紧凑MPS表示的量子态，近似编码也是可行的 [50]。有鉴于此，基于张量网络的策略已被用于为**经典数据** [29, 32, 33, 51]构建近似量子态编码。这些方法为低纠缠输入提供了具有可控近似误差的原则性框架，但其对量子数据或高纠缠经典数据的效用有限。

基于电路的方法。 基于电路的方法直接优化生成目标态的量子门，通常不依赖于显式的低纠缠表示。这些方法可大致分为变分策略与非变分策略。变分方法训练参数化电路，以最小化相对于目标态的非保真度 [34, 52, 53], ，而非变分方法则沿量子电路迭代优化局部双量子比特门，在无需显式参数训练的情况下逐步提高准确度 [35, 54]。基于电路的方法灵活且适用范围广，但缺乏严格的理论保证，并会遭遇贫瘠高原。

3 近似量子加载器

本节首先在统一优化框架下重新表述 AQL 的主要方法，并在第 3.1 节推导 AQLs 近似误差的两个信息论界。随后，我们将在第 3.2 节介绍 AQER 的实现细节。

3.1 AQL 的统一框架与理论分析

现有的 AQL 方法大致可分为基于 *TN* 的方法和基于量子电路的方法。为清晰起见，我们首先阐明 AQL 的统一框架，并说明如何在此框架下表述不同方法。基于这一框架，我们随后对 AQL 可实现的近似误差进行信息论分析。

AQL 的统一框架。我们首先回顾近似量子加载器（AQL）的基本定义。AQL 是指通过给定的门集 $\mathcal{U}$ 生成量子门序列，从而以可控的准确度制备目标态，如图 1 所示。具体而言，令 $U(\boldsymbol{\theta};\ \mathcal{A}) \in \mathcal{U}^m$ 表示由 m 个量子门组成的量子电路，其中 $\boldsymbol{\theta}$ 和 $\mathcal{A}$ 分别表示可调参数和电路架构。为确定最优的 $\boldsymbol{\theta}$ 和 $\mathcal{A}$，AQL 可重新表述为用于求解以下优化问题的统一框架

$$\arg\min_{\boldsymbol{\theta},\mathcal{A}} \left[1 - |\langle \boldsymbol{v}_{\text{target}}|U(\boldsymbol{\theta};\mathcal{A})|\psi_{\text{product}}\rangle|^2\right], \tag{1}$$

其中，$|\boldsymbol{v}_{\text{target}}\rangle$ 是目标量子态，$|\psi_{\text{product}}\rangle$是易于制备的乘积态。在这一框架下，各种 AQL 方法在构造量子电路 $U(\boldsymbol{\theta};\ \mathcal{A})$ 时有所不同，区别在于它们是更新 $\boldsymbol{\theta}$、设计 $\mathcal{A}$，还是两者兼顾。下文简要概述基于张量网络的方法和基于电路的方法如何对应于式（1），更多细节见附录A.3。

基于张量网络的方法。 对于这些方法， [50], 量子电路 $U(\boldsymbol{\theta},\ \mathcal{A})$ 是通过从张量网络表示中提取局部幺正算子并依次追加，逐步构造而成的。随后，每个酉算子都会通过 KAK 分解 [55], 进一步分解为硬件可用门，从而得到最终的量子电路。对于式（1），优化过程通过扩展量子电路 $U(\boldsymbol{\theta};\ \mathcal{A}) \rightarrow U(\boldsymbol{\theta}\cup\boldsymbol{\theta}_{\text{new}};\ \mathcal{A}\cup\mathcal{A}_{\text{new}})$ 来进行，同时保持先前的参数和架构不变。基于电路的方法。 基于电路的AQL方法可分为变分方法和非变分方法。在变分方法中，[34, 52, 53], 通过优化带有代价函数 $\ell(\boldsymbol{\theta}) = 1-|\langle \boldsymbol{v}_{\text{target}}|U(\boldsymbol{\theta};\ \mathcal{A})\ |\psi_{\text{product}}\rangle\ |^2$的变分量子电路，构造量子电路 $U(\boldsymbol{\theta};\ \mathcal{A})$，其中 $\boldsymbol{\theta}$ 是可调的，而 $\mathcal{A}$ 是固定的。相比之下，非变分方法 [35, 54] 按照预设的电路逐步更新所有双量子比特酉算子，并遵循之字形调度。这需要依次调整电路 $U(\boldsymbol{\theta};\ \mathcal{A})$的参数 $\boldsymbol{\theta}$和架构 $\mathcal{A}$，因此属于式（1）中的通用优化框架。

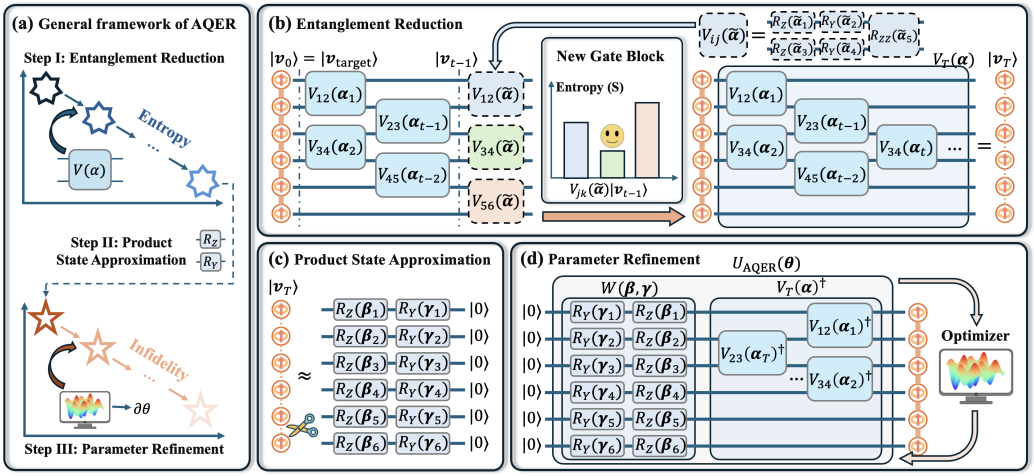

图 2：AQER 算法的工作流程。(a) AQER 概览，由三个步骤组成。(b) 步骤 *I*：纠缠约化。此步骤通过迭代添加双量子比特门块，逐步降低输入 $|\boldsymbol{v}_{\text{target}}\rangle$ 与量子电路 $V_T(\boldsymbol{\alpha})$的纠缠。(c) 步骤 *II*：乘积态近似。此步骤通过对初始态 $|0\rangle^{\otimes N}$施加单量子比特旋转 $\{R_Z(\boldsymbol{\beta}_n)\}_{n=1}^N$和$\{R_Y(\boldsymbol{\gamma}_n)\}_{n=1}^N$，近似低纠缠量子态 $|\boldsymbol{v}_T\rangle$。(d) 步骤 *III*：参数精化。此步骤对 $\boldsymbol{\theta}=(\boldsymbol{\alpha},\ \boldsymbol{\beta},\ \boldsymbol{\gamma})$中的所有电路参数进行微调，以最小化非保真度并获得最终的 AQL $U_{\text{AQER}}(\boldsymbol{\theta}^*)$。

信息论分析。 式（1）中的统一框架有一个优点：它能够对AQL进行与算法无关的理论分析。在此，我们利用一种纠缠度量，建立了AQL可达到的近似误差的两个信息论界，具体如下述定理所述，证明见附录 B.2。

定理 *3.1*。设 N量子比特量子态 $|\psi\rangle$ 的纠缠度量为 $\mathcal{S}(|\psi\rangle) = \sum_{i=1}^N \mathcal{S}_{\{i\}}(|\psi\rangle)$。那么，对于量子态 $|\boldsymbol{v}_{\text{target}}\rangle$ 以及具有 $\mathcal{S}(U^\dagger|\boldsymbol{v}_{\text{target}}\rangle) = S$的量子电路 U，$|\boldsymbol{v}_{\text{target}}\rangle$ 与由 U 作用于任意乘积态 $|\psi_{\text{product}}\rangle$ 所生成的量子态之间的非保真度下界为 $1-|\langle\boldsymbol{v}_{\text{target}}|U|\psi_{\text{product}}\rangle|^2 \geq f_1(S) := \frac{1}{2}\left(1-\sqrt{2^{1-S/N}-1}\right)$。此外，在能够访问 ρ的情况下，我们可以构造乘积态 $|\psi'_{\text{product}}\rangle$，使非保真度的上界为 $1-|\langle\boldsymbol{v}_{\text{target}}|U|\psi'_{\text{product}}\rangle|^2 \leq f_2(S) := \frac{1}{2}\left(1-\sqrt{2^{1-S+\lfloor S\rfloor}}-1+\lfloor S\rfloor\right)$。当 $S\to 0$ 时，$f_1(S) \to \ln_{2^N}^2 S + \mathcal{O}(S^3)$，且 $f_2(S) \to \ln_2^2 S + \mathcal{O}(S^3)$。

上述结果表明，AQL 的近似误差从根本上取决于演化态 $U^\dagger|\boldsymbol{v}_{\text{target}}\rangle$的纠缠度量 $\mathcal{S}$。具体而言，目标态 $|\boldsymbol{v}_{\text{target}}\rangle$ 与制备态 $U|\psi_{\text{product}}\rangle$ 之间的非保真度随纠缠度量值 S线性增长。因此，较小的 S 能够保证较低的非保真度和较小的近似误差。由于 S同时取决于 $|\boldsymbol{v}_{\text{target}}\rangle$ 和量子电路 U，通过 AQL 中的参数优化和架构优化来降低非保真度，等同于最小化纠缠度量 $\mathcal{S}$。

备注。 定理 3.1 可推广至附录 C 中所述的噪声信道情形。

3.2 AQER：高效且可扩展的 AQL

纠缠度量 $\mathcal{S}$ 在决定近似误差方面的理论作用表明，它可以作为评估 AQL 质量的实用指标。基于这一认识，我们提出 AQER，这是一种高效且可扩展的 AQL，它通过遵循最大纠缠减少原则构造门序列，从而求解式 (1) 中的优化问题。在本小节余下部分，我们将介绍 AQER 的实现。

AQER 概述： 如图 2(a) 所示，AQER 由三个关键组成部分构成：(I) 通过逐步添加量子门来抑制输入数据中的纠缠度量 $\mathcal{S}$；(II) 应用单量子比特旋转门以校正主导误差；(III) 优化电路参数以进一步细化。接下来，我们将分别说明这些步骤。

步骤 I：纠缠约化。 本步骤的目标是构建一个门序列，通过监测纠缠度量 $\mathcal{S}$ 来降低目标输入的纠缠。如图 2(b) 所示，这是通过**迭代地**追加双量子比特门块 $V_{\mathcal{I}_t}(\boldsymbol{\alpha}_t)$实现的，其中可调变量包括带有 $1 \le j_t \ne k_t \le N$ 的作用量子比特对 $\mathcal{I}_t = (j_t, k_t)$ 以及门参数$\boldsymbol{\alpha}_t$。具体而言，这些门块 $\{V_{\mathcal{I}_t}(\boldsymbol{\alpha}_t)\}$ 具有相同的结构 $R_{ZZ}R_YR_Z$，即对两个量子比特都施加单量子比特旋转门。令 $V_{t-1}(\boldsymbol{\alpha}_{1:t-1}) = V_{\mathcal{I}_{t-1}}(\boldsymbol{\alpha}_{t-1}) \cdots V_{\mathcal{I}_1}(\boldsymbol{\alpha}_1)$ 为经过 $(t-1)$ 次迭代后生成的门序列。在第 t 次迭代中，门块 $V_{\mathcal{I}_t}(\boldsymbol{\alpha}_t)$的作用量子比特对 $\mathcal{I}_t$ 和门参数 $\boldsymbol{\alpha}_t$ 通过求解以下优化问题得到

$$\mathcal{I}_t, \boldsymbol{\alpha_t} = \arg\min_{\tilde{\mathcal{I}},\tilde{\boldsymbol{\alpha}}} \mathcal{S}\left(V_{\tilde{\mathcal{I}}}(\tilde{\boldsymbol{\alpha}})|\boldsymbol{v}_{t-1}\rangle\right), \tag{2}$$

其中，$\mathcal{S}$ 是定理 3.1 中定义的纠缠度量，且 $|\boldsymbol{v}_{t-1}\rangle := V_{t-1}(\boldsymbol{\alpha}_{1:t-1})|\boldsymbol{v}_{\text{target}}\rangle$。重复这一迭代过程 T 次，使量子态 $|\boldsymbol{v}_T\rangle = V_T(\boldsymbol{\alpha})|\boldsymbol{v}_{\text{target}}\rangle$ 的纠缠足够小，其中 $\boldsymbol{\alpha} := \boldsymbol{\alpha}_{1:T} = (\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_T)$。

步骤 II：乘积态近似。在步骤 I 之后，AQER 旨在通过应用单量子比特旋转门进一步降低近似误差，如图 2(c) 所示。此步骤的动机在于，正如定理 3.1 所指出的，低纠缠的量子态 $|\boldsymbol{v}_T\rangle$ 可以用乘积态很好地近似。为了从标准初始态 $|0\rangle^{\otimes N}$制备这样的乘积态，我们应用额外的单量子比特旋转 $W(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \otimes_{i=1}^{N}(R_Z(\boldsymbol{\beta}_i)R_Y(\boldsymbol{\gamma}_i))$。请注意，最优参数 $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_N)$ 和 $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1, \cdots, \boldsymbol{\gamma}_N)$ 无需数值优化即可**显式推导**，如下方推论所示。显式形式及其推导请参见附录 B.1。

推论 3.2 （非正式）。 在能够持续获取纠缠较少的量子态 $|\boldsymbol{v}_T\rangle$ 的情况下，无需优化即可推导出 $(\boldsymbol{\beta}, \boldsymbol{\gamma})$ 中每个参数的显式形式。

步骤 III：参数精化。 AQER 的最后一步通过微调步骤 I 和步骤 II 中构造的门序列参数，进一步提升准确度，如图 2(d) 所示。为清晰起见，我们将组合量子电路记为 $U_{\text{AQER}}(\boldsymbol{\theta}) = V_T(\boldsymbol{\alpha})^\dagger W(\boldsymbol{\beta}, \boldsymbol{\gamma})$，其中 $\boldsymbol{\theta} = (\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma})$ 汇集了所有可调参数。因此，参数精化即为最小化目标态 $|\boldsymbol{v}_{\text{target}}\rangle$ 与量子态 $U_{\text{AQER}}(\boldsymbol{\theta})|0\rangle^{\otimes N}$ 之间的非保真度，即

$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}} \left(1 - |\langle \boldsymbol{v}_{\text{target}}|U_{\text{AQER}}(\boldsymbol{\theta})|0\rangle^{\otimes N}|^2\right). \tag{3}$$

对公式 (3) 进行优化后，我们得到 AQL：$|\boldsymbol{v}_{\text{load}}\rangle = e^{-ig}U_{\text{AQER}}(\boldsymbol{\theta}^*)|0\rangle^{\otimes N}$，其中 g 是不会影响任何测量结果的全局相位。

注记。 (i) AQER 的一个重要特征是使用纠缠度量 $\mathcal{S}$ 作为近似误差的代理。对于量子数据，评估和优化 $\mathcal{S}$ 是高效的，因为它仅涉及局部测量。对于经典数据，可以经典地模拟 AQER 以构造 U_{AQER}。(ii) 通过首先抑制纠缠度量 $\mathcal{S}$，AQER 不仅降低了近似误差，还通过缓解贫瘠高原问题而有别于先前的基于电路的方法，从而增强了可训练性和可扩展性。有关更多讨论以及 AQER 的时间复杂度分析，请参见附录 D 和 G。(iii) 一般而言，AQER 是一种启发式算法。然而，对于某些结构化向量或状态族，AQER 具有高效的性能保证。特别地，如附录 H 所示，AQER 可证明地以多项式资源开销为 IQP 态生成最优加载电路。

4 实验

我们进行了大量数值模拟，以评估 AQER 在加载经典与量子数据集方面的性能。进一步的实现细节和更多结果分别见附录 E 和 F。

4.1 AQL 的数据集构建

我们简要介绍本文所用的数据集，更多详情见附录 E.1。

经典数据。 我们使用三个标准经典数据集：MNIST [56], CIFAR-10 [57], 和 SST-2 [58]。这些数据集广泛用于在视觉与语言任务中，针对多种数据类型对 AQL 方法进行基准测试。具体而言，MNIST 包含 28×28 灰度手写数字；

表 1： 不同 AQL 方法在 MNIST、CIFAR-10、SST-2、S-RQC 和 GS-TFIM 数据集上的非保真度 (↓)。我们将 AQER 与 $G \in \{20, 40, 80\}$ 进行比较，并以参考方法为对照；由于附录 E.2 中详述的可行性限制，后者使用相同或略大的 G。报告值为 M 个样本的均值（及标准差）。最佳 和 次佳 结果分别以蓝色 和 橙色标出。

	MNIST			CIFAR-10			SST-2			S-RQC			GS-TFIM		
MPS	G	36	54	90	30	60	90	36	54	90	27	54	81	36	72	90
		0.330 (0.101)	0.287 (0.089)	0.237 (0.076)	0.068 (0.038)	0.056 (0.031)	0.049 (0.028)	0.901 (0.022)	0.870 (0.022)	0.814 (0.022)	0.746 (0.100)	0.663 (0.118)	0.605 (0.127)	0.056 (0.004)	0.039 (0.003)	0.041 (0.002)
HEC	G	20	40	80	22	44	88	20	40	80	20	40	80	20	40	80
		0.430 (0.089)	0.234 (0.079)	0.103 (0.042)	0.081 (0.039)	0.050 (0.026)	0.030 (0.016)	0.892 (0.016)	0.759 (0.012)	0.546 (0.007)	0.731 (0.097)	0.484 (0.133)	0.367 (0.161)	0.168 (0.090)	0.020 (0.013)	0.007 (0.002)
AQCE	G	20	40	80	30	45	90	20	40	80	30	45	90	20	40	80
		0.296 (0.083)	0.145 (0.053)	0.051 (0.023)	0.068 (0.036)	0.048 (0.026)	0.024 (0.014)	0.891 (0.017)	0.761 (0.018)	0.518 (0.013)	0.534 (0.149)	0.363 (0.156)	0.267 (0.110)	0.108 (0.026)	0.068 (0.024)	0.056 (0.015)
AQER (我们的方法)	G	20	40	80	20	40	80	20	40	80	20	40	80	20	40	80
		0.195 (0.060)	0.090 (0.034)	0.034 (0.015)	0.043 (0.023)	0.029 (0.016)	0.018 (0.010)	0.819 (0.017)	0.652 (0.013)	0.406 (0.008)	0.285 (0.152)	0.128 (0.106)	0.067 (0.069)	0.028 (0.011)	0.009 (0.006)	0.003 (0.001)

CIFAR-10 包含 32×32 自然物体的 RGB图像，例如猫和汽车；SST-2 是一个情感分类数据集，我们使用预训练的 Sentence-BERT 模型 [59] 获取 1024 维句子嵌入。每个数据集都被预处理为 $M = 50$ 归一化向量，然后使用振幅编码 $|\boldsymbol{v}\rangle = \sum_{j=1}^{2^N} \boldsymbol{v}_j |j\rangle$或紧凑编码 [60] $|\boldsymbol{v}\rangle = \sum_{j=1}^{2^N} (\boldsymbol{v}_j + \imath \boldsymbol{v}_{j+2^N}) |j\rangle$ 编码为目标态，使用 $N \in \{10, 11\}$ 个量子比特。

量子数据。 我们构建了两种类型的量子数据集：(i) 由随机量子电路（RQC）生成的合成态；以及 (ii) 一维横场伊辛模型（1D TFIM）的基态 [61], , 分别记为 S-RQC 和 GS-TFIM。这些数据集代表了典型的量子电路演化和多体量子系统所产生的量子态示例。S-RQC 包含 $M = 50$个通过对量子态$|0\rangle^{\otimes N}$施加不同 RQC 生成的量子态。每个 RQC 都从集合 *RandomShuffle* $\left(\{CZ_{p_k,q_k}\}_{k=1}^{W} \cup \{R_{p_k}\}_{k=W+1}^{4W}\right)$中采样，该集合包含 W CZ 个量子门和 $3W$ 个在随机选定量子比特上执行的单量子比特旋转 $\{p_k, q_k\}$。这里，我们设定 $N = 10$和 $N_{\mathrm{RQC}} = 40$。GS-TFIM 包含 1D TFIM 哈密顿量$H = -\sum_{i=1}^{N-1} JZ_iZ_{i+1} - \sum_{i=1}^{N} gX_i$的基态。我们考虑系数 $g = 1$ 和 $J \in \{0.8, 0.9, 1, 1.1, 1.2\}$，为每个 $N \in \{10, 20, 30, 40, 50\}$构建大小为 $M = 5$ 的数据集。

4.2 实验设置

参考 AQL 方法。 我们选取三种典型的参考 AQL 方法，以便与 AQER 进行全面比较。这三种参考方法分别是：(i) 基于一维矩阵乘积态 [29],的张量网络方法，代表了在低纠缠数据上具有理论保证的方法；(ii) 基于硬件高效线路（HEC） [52], 的方法，该线路通常用于变分量子算法；(iii) 自动量子电路编码（AQCE） [35], , 展示了非变分AQL 的最新进展。
评估指标。 为了量化 AQL 的准确度，我们使用非保真度 $1 - |\langle \boldsymbol{v}_{\mathrm{load}} | \boldsymbol{v}_{\mathrm{target}} \rangle|^2$ 作为近似误差的度量。非保真度的取值范围为 0 至 1，其中，数值越小表示近似误差越低、准确度越高。为了评估 AQL 的效率，我们考察通过编码电路制备量子态所需的量子资源。具体而言，我们记录编码电路中所使用的双量子比特门数量（例如， CZ 和 R_{ZZ}）。该数量记为 G，作为量子资源消耗的度量，因为它决定了电路运行时间 [62]。 G 越小，表示 AQL 的效率越高。

AQER 的超参数设置。 AQER 步骤 I 的迭代次数默认设置为 $T \in \{5, 10, 20, 40, 60, 80, 100\}$。对于量子比特数较多的 GS-TFIM（$N \geq 20$），我们使用$T \in \{20, 40, 60, 80, 100, 120, 160, 200\}$，因为较大的系统通常需要更多量子门来捕获量子关联。在 AQER 中，迭代次数 T控制双量子比特门数量 G，每次迭代引入一个双量子比特门。在步骤 I 的第 t 次迭代中，公式 (2) 中的参数 $\boldsymbol{\alpha}_t$ 首先初始化为零，然后使用 Nelder–Mead方法，以 10^{-4} 的收敛容差进行优化。通过选择使其最小化的量子比特对来优化量子比特索引集 $\mathcal{I}_t$

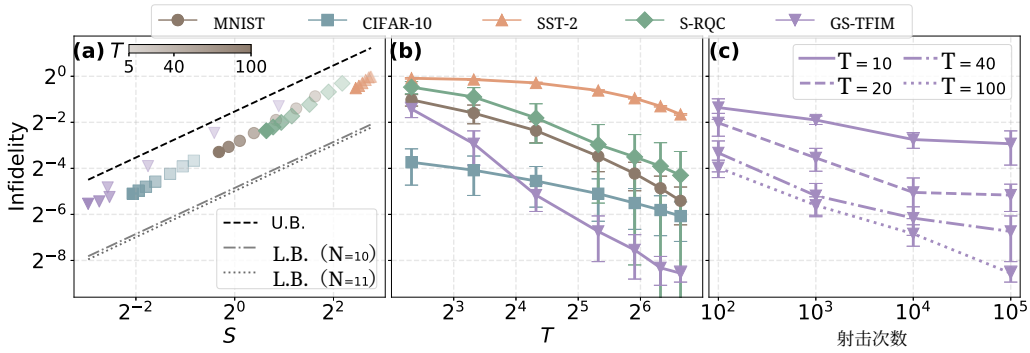

图3：AQER 在 MNIST、CIFAR-10、SST-2、S-RQC 和 GS-TFIM 数据集上的性能，以不同颜色和标记区分。(a) AQER 步骤 II 后，非保真度与纠缠度量值 S 的关系，数值为对 M 个样本取平均所得。颜色条由浅至深表示 T 递增。虚线表示定理 3.1 中线性化的上界（U.B.）和下界（L.B.），其中忽略了高阶项。(b) AQER 步骤 III 后，所有数据集上的非保真度与不同 T 值的关系。(c) GS-TFIM 数据集上，非保真度与不同测量次数的关系，其中 $T \in \{10, 20, 40, 100\}$ 不同。

S 通过调整 $\boldsymbol{\alpha}_t$。步骤 III 使用 Adam 优化器，以 10^{-2} 的学习率进行 $T_3 = 2000$ 次迭代的优化。对于量子数据集，S 等量以及梯度默认根据 10^5 次模拟测量次数进行估计。

4.3 实验结果

我们在经典与量子数据集上评估 AQER，以验证其准确度和效率，以及其在大型系统上的可训练性和可扩展性，并将其性能与现有 AQL 方法进行比较。更多数值结果见附录 F。

AQER 在经典数据和量子数据上均优于所有作为基准的 AQL 方法。我们首先通过测量不同双量子比特门数量 G 下多个数据集的近似误差（非保真度），比较各种 AQL 方法的性能。表 1 列出了 MNIST、CIFAR-10、SST-2、S-RQC 和 GS-TFIM（其中 $N = 10$）上的结果。AQER 的结果以 $G \in \{20, 40, 80\}$ 显示。对于作为基准的 AQL 方法，G 设为相同或略大的值，具体取决于附录 E.2 中说明的可行性约束。可以看出，AQER 始终优于现有 AQL 方法，在使用相同甚至更少的 G时实现最低非保真度。在 S-RQC 上的改进最为显著：对于 $G \in \{40, 80\}$，AQER 的非保真度比排名第二的方法（AQCE）降低了 60% 以上；此外，AQER 使用的双量子比特门数量比其他方法少 50%，同时实现了更低的非保真度。这些结果验证了 AQER 的优势：与现有 AQL 方法相比，它使用相同甚至更少的双量子比特门，便能实现更低的非保真度。

AQER 通过降低纠缠度量，持续减少非保真度 $\mathcal{S}$。接下来，我们考察步骤 I 中纠缠熵降低在减少 AQER 近似误差方面的有效性。为此，我们记录步骤 I 中在不同迭代次数 $T \in \{5, 10, 20, 40, 60, 80, 100\}$设置下的纠缠度量 S，并同时记录步骤 II 后对应的近似误差。这些结果如图 3(a)所示。对于所有数据集，增加 T 通常会使数据点向更低的 S 和非保真度方向移动，而这些数据点始终处于定理 3.1 给出的理论上下界之内，表明 AQER 会逐步扩展量子电路，以降低纠缠并提高 AQL 的准确度。例如，对于 MNIST 数据集，将 T 从 5 增加到 100，会使 S 降低约四倍，相应的非保真度也以类似幅度下降，从而验证了纠缠约化对降低非保真度的作用。

双量子比特门数量和射击次数对 AQER 性能的影响。我们进一步研究双量子比特门数量和测量次数对 AQER 的影响。具体而言，我们在 $\{10^2, 10^3, 10^4, 10^5\}$ 中针对不同的 $T \in \{5, 10, 20, 40, 60, 80, 100\}$ 和射击次数进行了实验。图 3(b) 展示了不同数据集在 AQER 的步骤 III 之后，非保真度随 T 变化的情况；图 3(c) 则说明了不同测量次数对 GS-TFIM 数据集的影响。较大的 T 值会显著降低非保真度。例如，将 T 从 5 增加到 100，可使

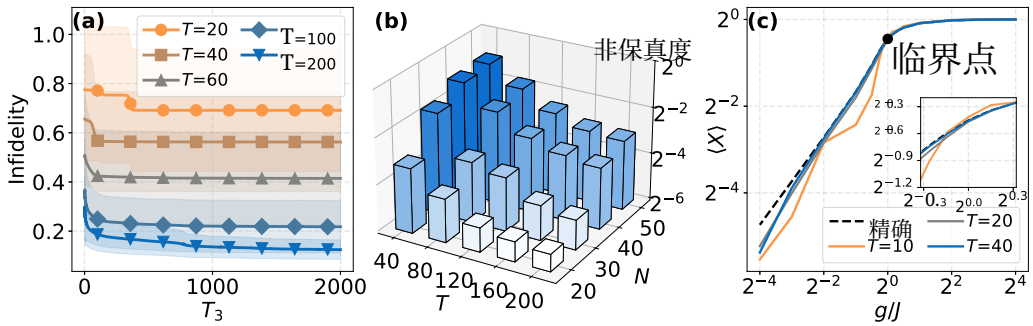

图 4：AQER 在 GS-TFIM 数据集上的性能。(a) 在 $N = 50$ 量子比特下，步骤 III 优化过程中不同 T取值的非保真度。(b) 不同量子比特数 N 以及步骤 I 迭代次数T下的非保真度。(c) 在 $N = 10$条件下，针对不同 g/J取值，在 AQER 加载的 GS-TFIM 量子态上测得的平均磁化强度 $\langle X\rangle$ 的期望值。每条曲线对应一个不同的 $T \in \{10, 20, 40\}$。

GS-TFIM 数据集的非保真度从高于 2^{-2} 降至低于 2^{-8}。同样，增加射击次数通常能通过抑制电路生成和优化中的统计噪声来降低非保真度。这种效果在较大的 T 下更为明显：对于 $T = 10$，降幅小于 4；而对于 $T = 100$，降幅大于 16。这些结果定量表明，较大的电路结合充足的测量次数能够有效提升 AQER 的性能。

AQER 的可训练性。 随后，我们通过在包含 $N = 50$ 量子比特的 GS-TFIM 数据集上的实验，展示了 AQER 在大规模系统上的可训练性。图 4(a) 给出了 AQER 步骤 III 中针对不同 $T \in \{20, 40, 60, 100, 200\}$ 的参数优化。优化曲线并未出现贫瘠高原，否则该过程会被困在接近 1 的高非保真度处。初始非保真度已远小于 1，这与定理 3.1 一致。例如，在 $T = 200$条件下，非保真度从约 0.3 开始，并有效下降至约 0.1。这些结果表明，AQER 中的纠缠降低机制成功缓解了步骤 III 中的贫瘠高原效应，从而保证了可训练性，并为其可扩展性提供了前提。

AQER 的可扩展性。 为验证 AQER 的可扩展性，我们在 GS-TFIM 数据集上针对不同的量子比特数 $N \in \{20, 30, 40, 50\}$ 和 $T \in \{40, 80, 120, 160, 200\}$进行了实验。如图 4(b) 所示，对于所有系统规模，非保真度均随 T 增加而持续降低。特别是，当 T 随 N线性增长（具体遵循 $T = 4N - 40$）时，AQER 在不同的 N 下均能保持近乎恒定的非保真度，凸显了其在量子比特数和双量子比特门数量方面良好的可扩展性。

AQER 的下游性能。 最后，我们在 Quantum 和经典数据集的下游任务中评估 AQER 的性能。我们首先检测横场伊辛模型中的量子相变，该相变通过基态的平均磁化强度 $\langle X\rangle$（作为序参量）进行测量。我们记录了 AQER加载态 在不同 T值下的 $\langle X\rangle$ 值随 g/J 变化的情况，如图 4(c) 所示。这些结果表明，AQER加载态能够捕捉从具有 $g/J < 1$ 的铁磁相到具有$g/J > 1$的顺磁相的转变，并且在 $g/J = 1$处的临界点附近发生快速变化。增加 T 通常能更准确地逼近精确值，表明更大的电路可以编码更多与序参量相关的关联。值得注意的是，即便 $T = 10$适中的量子态也能捕捉总体趋势，这表明AQER加载态能够有效捕捉量子相

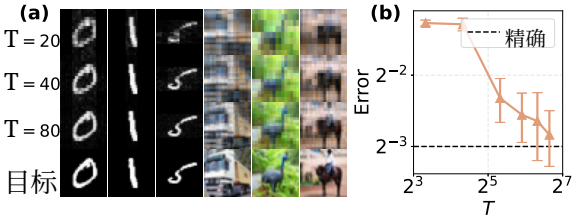

图 5：AQER 在经典数据上的下游性能。(a) 使用 AQER 和 $T \in \{20, 40, 80\}$ 重建的 MNIST 和 CIFAR-10 图像。(b) 使用$T \in \{10, 20, 40, 60, 80, 100\}$对应的 AQER加载态，在 SST-2 数据集上的分类错误率，并与精确加载（黑色虚线）进行比较。

变，且仅需有限的量子资源。接下来，对于经典数据，我们使用AQER在不同的 T 取值下重构MNIST和CIFAR-10图像。如图5(a)所示，随着 T 增大，重构结果逐渐接近目标。我们还使用量子核方法在SST-2上进行二分类评估，细节见附录 E.3。如图5(b)所示，误差随 T增大而减小，并在 $T = 100$处接近精确加载误差 2^{-3} 。这些结果表明，更大的 T可提升下游性能。

5 结论

在本研究中，我们为 AQLs 引入了一个统一框架，并推导出信息论界，表明非保真度从根本上受量子态纠缠的控制。基于这一洞见，我们提出了 AQER，这是一种可扩展且高效的方法，通过系统性地减少纠缠，在实现高效门使用量的同时达到较低的非保真度。在经典与量子数据集上的广泛基准测试表明，AQER 在准确度和电路效率方面均持续优于现有方法。这些结果既提供了理论保证，也提出了一种高效量子数据加载的实用方法，从而支持数据依赖型量子算法的更广泛应用。

致谢

本项目得到新加坡国家研究基金会的支持，资助项目为 NRF 教授奖学金，编号 NRF-P2024-001。

参考文献

[1] Richard P Feynman. 用计算机模拟物理。《国际理论物理学杂志》，21(6/7)，1982年。[2] Peter W Shor. 量子计算算法：离散对数与因数分解。载于《第35届计算机科学基础年会论文集》，第124–134页。IEEE，1994年。[3] Aram W. Harrow、Avinatan Hassidim 和 Seth Lloyd. 求解线性方程组的量子算法。《物理评论快报》，103:150502，2009年10月。[4] Yudong Cao、Jonathan Romero、Jonathan P Olson、Matthias Degroote、Peter D Johnson、Mária Kieferová、Ian D Kivlichan、Tim Menke、Borja Peropadre、Nicolas PD Sawaya 等. 量子计算时代的量子化学。《化学评论》，119(19):10856–10915，2019年。[5] Jin-Peng Liu、Herman Oie Kolden、Hari K. Krovi、Nuno F. Loureiro、Konstantina Trivisa 和 Andrew M. Childs. 求解耗散性非线性微分方程的高效量子算法。《美国国家科学院院刊》，118(35):e2026805118，2021年。[6] Hsin-Yuan Huang、Michael Broughton、Jordan Cotler、Sitan Chen、Jerry Li、Masoud Mohseni、Hartmut Neven、Ryan Babbush、Richard Kueng、John Preskill 和 Jarrod R. McClean. 从实验中学习的量子优势。《科学》，376(6598):1182–1186，2022年。[7] Junyu Liu、Minzhao Liu、Jin-Peng Liu、Ziyu Ye、Yunfei Wang、Yuri Alexeev、Jens Eisert 和 Liang Jiang. 迈向可证明高效的大规模机器学习模型量子算法。《自然·通讯》，15(1):434，2024年1月。[8] John M. Martinis、Michel H. Devoret 和 John Clarke. 电流偏置约瑟夫森结零电压态中的能级量子化。《物理评论快报》，55:1543–1546，1985年10月。[9] D. Jaksch、J. I. Cirac、P. Zoller、S. L. Rolston、R. Côté 和 M. D. Lukin. 中性原子的快速量子门。《物理评论快报》，85:2208–2211，2000年9月。

[10] Yuriy Makhlin、Gerd Schön 和 Alexander Shnirman。利用约瑟夫森结器件进行量子态工程。《现代物理评论》，73：357–400，2001年5月。

[11] John Clarke 和 Frank K. Wilhelm。超导量子比特。《自然》，453(7198)：1031–1042，2008年6月。[12] L. DiCarlo、J. M. Chow、J. M. Gambetta、Lev S. Bishop、B. R. Johnson、D. I. Schuster、J. Majer、A. Blais、L. Frunzio、S. M. Girvin 和 R. J. Schoelkopf。利用超导量子处理器演示双量子比特算法。《自然》，460(7252)：240–244，2009年7月。[13] L. Isenhower、E. Urban、X. L. Zhang、A. T. Gill、T. Henage、T. A. Johnson、T. G. Walker 和 M. Saffman。中性原子受控非门的演示。《物理评论快报》，104：010503，2010年1月。[14] Frank Arute、Kunal Arya、Ryan Babbush 等。利用可编程超导处理器实现量子霸权。《自然》，574(7779)：505–510，2019年10月。[15] Yulin Wu、Wan-Su Bao、Sirui Cao 等。利用超导量子处理器实现强量子计算优势。《物理评论快报》，127：180501，2021年10月。[16] T. M. Graham、Y. Song、J. Scott、C. Poole、L. Phuttitarn、K. Jooya、P. Eichler、X. Jiang、A. Marra、B. Grinkemeyer、M. Kwon、M. Ebert、J. Cherek、M. T. Lichtman、M. Gillette、J. Gilbert、D. Bowman、T. Ballance、C. Campbell、E. D. Dahl、O. Crawford、N. S. Blunt、B. Rogers、T. Noel 和 M. Saffman。在中性原子量子计算机上实现多量子比特纠缠与算法。《自然》，604(7906)：457–462，2022年4月。[17] Rajeev Acharya、Dmitry A. Abanin、Laleh Aghababaie-Beni、Igor Aleiner、Trond I. Andersen 等。低于表面码阈值的量子纠错。《自然》，638(8052)：920–926，2025年2月。[18] Muqing Xu、Lev Haldar Kendrick、Anant Kale、Youqi Gang、Chunhan Feng、Shiwei Zhang、Aaron W Young、Martin Lebrat 和 Markus Greiner。低温环境下的中性原子哈伯德量子模拟器。《自然》，第1–7页，2025年。[19] Yao-Yao Jiang、Chun-Qing Deng、Heng Fan、Bing-Yang Li、Lu-Yan Sun、Xin-Sheng Tan、Wei-Ting Wang、Guang-Ming Xue、Fei Yan、Hai-Feng Yu 等。超导量子计算的进展。《国家科学评论》，第nwaf246页，2025年。[20] Davide Girolami。制备一个量子态有多难？《物理评论快报》，122：010505，2019年1月。[21] Javier Gonzalez-Conde、Thomas W. Watts、Pablo Rodriguez-Grasa 和 Mikel Sanz。多项式函数的高效量子振幅编码。Quantum，8：1297，2024年3月。[22] Kishor Bharti、Alba Cervera-Lierta、Thi Ha Kyaw、Tobias Haug、Sumner Alperin-Lea、Abhinav Anand、Matthias Degroote、Hermanni Heimonen、Jakob S Kottmann、Tim Menke 等。含噪中尺度量子算法。《现代物理评论》，94(1)：015004，2022年。[23] F. Javier Cardama、Jorge Vázquez-Pérez、Tomás F. Pena、Juan C. Pichel 和 Andrés Gómez。量子编译流程：综述。载于 Silvina Caino-Lores、Demetris Zeinalipour、Thaleia Dimitra Doudali、David E. Singh、Gracia Ester Martín Garzón、Leonel Sousa、Diego Andrade、Tommaso Cucinotta、Donato D’Ambrosio、Patrick Diehl、Manuel F. Dolz、Admela Jukan、Raffaele Montella、Matteo Nardelli、Marta Garcia-Gasulla 和 Sarah Neuwirth 编，《Euro-Par 2024：并行处理研讨会》，第100–112页，沙姆，2025年。施普林格·自然瑞士。[24] Kaining Zhang、Min-Hsiu Hsieh、Liu Liu 和 Dacheng Tao。量子格拉姆－施密特过程及其在量子算法高效态读出中的应用。《物理评论研究》，3：043095，2021年11月。[25] Franz J. Schreiber、Jens Eisert 和 Johannes Jakob Meyer。参数化量子态的层析成像。《PRX 量子》，6：020346，2025年6月。[26] Deepak Ranga、Aryan Rana、Sunil Prajapat、Pankaj Kumar、Kranti Kumar 和 Athanasios V Vasilakos。量子机器学习：探索数据编码技术的作用、挑战与未来方向。《数学》，12(21)：3318，2024年。

[27] Xiao-Ming Zhang、Tongyang Li和Xiao Yuan。最优电路深度的量子态制备：实现与应用。《物理评论快报》，129:230504，2022年11月。[28] Kaiwen Gui、Alexander M. Dalzell、Alessandro Achille、Martin Suchara和Frederic T. Chong。时空高效低深度量子态制备及其应用。Quantum，8:1257，2024年2月。[29] Jason Iaconis和Sonika Johri。基于张量网络的高效图像量子数据加载。arXiv预印本 arXiv:2310.05897，2023年。[30] Nam H Nguyen、Elizabeth C Behrman和James E Steck。含噪声与退相干的量子学习：一种鲁棒量子神经网络。Quantum Machine Intelligence，2(1):1，2020年。[31] Hsin-Yuan Huang、Yu Tong、Di Fang和Yuan Su。以海森堡极限标度学习多体哈密顿量。《物理评论快报》，130:200403，2023年5月。[32] Bernhard Jobst、Kevin Shen、Carlos A. Riofrío、Elvira Shishenina和Frank Pollmann。由经典图像数据的傅里叶模态得到的高效MPS表示与量子电路。Quantum，8:1544，2024年12月。[33] Jason Iaconis、Sonika Johri和Elton Yechao Zhu。利用矩阵乘积态进行正态分布的量子态制备。npj Quantum Information，10(1):15，2024年1月。[34] Naoki Mitsuda、Tatsuhiro Ichimura、Kouhei Nakaji、Yohichi Suzuki、Tomoki Tanaka、Rudy Raymond、Hiroyuki Tezuka、Tamiya Onodera和Naoki Yamamoto。近似复振幅编码算法及其在数据分类问题中的应用。《物理评论A》，109:052423，2024年5月。[35] Tomonori Shirakawa、Hiroshi Ueda和Seiji Yunoki。给定任意量子态的自动量子电路编码。《物理评论研究》，6:043008，2024年10月。[36] Martin Larocca、Supanut Thanasilp、Samson Wang、Kunal Sharma、Jacob Biamonte、Patrick J Coles、Lukasz Cincio、Jarrod R McClean、Zoë Holmes和Marco Cerezo。变分量子计算中的贫瘠高原。Nature Reviews Physics，第1–16页，2025年。[37] Michael A Nielsen和Isaac L Chuang。量子计算与量子信息。剑桥大学出版社，2010年。[38] Marco Cerezo、Andrew Arrasmith、Ryan Babbush、Simon C Benjamin、Suguru Endo、Keisuke Fujii、Jarrod R McClean、Kosuke Mitarai、Xiao Yuan、Lukasz Cincio等。变分量子算法。Nature Reviews Physics，3(9):625–644，2021年。[39] Maria Schuld和Nathan Killoran。特征希尔伯特空间中的量子机器学习。《物理评论快报》，122:040504，2019年2月。[40] Yuxuan Du、Xinbiao Wang、Naixu Guo、Zhan Yu、Yang Qian、Kaining Zhang、Min-Hsiu Hsieh、Patrick Rebentrost和Dacheng Tao。量子机器学习简明导论。Artificial Intelligence (R0)。Springer Nature Singapore，2025年。[41] Seth Lloyd和Christian Weedbrook。量子生成对抗学习。《物理评论快报》，121:040502，2018年7月。[42] X. Gao、Z.-Y. Zhang和L.-M. Duan。基于生成模型的量子机器学习算法。Science Advances，4(12):eaat9004，2018年。[43] Jinkai Tian、Xiaoyu Sun、Yuxuan Du、Shanshan Zhao、Qing Liu、Kaining Zhang、Wei Yi、Wanrong Huang、Chaoyue Wang、Xingyao Wu、Min-Hsiu Hsieh、Tongliang Liu、Wenjing Yang和Dacheng Tao。量子神经网络在生成学习中的最新进展。IEEE模式分析与机器智能汇刊，45(10):12321–12340，2023年。[44] Samuel Yen-Chi Chen、Chao-Han Huck Yang、Jun Qi、Pin-Yu Chen、Xiaoli Ma和Hsi-Sheng Goan。用于深度强化学习的变分量子电路。IEEE Access，8:141007–141024

，2020年。

[45] Zhihao Cheng、Kaining Zhang、Li Shen 和 Dacheng Tao。保守式离线量子强化学习。载于 AAAI人工智能会议论文集，第37卷，第7148–7156页，2023年。[46] Zhihao Cheng、Kaining Zhang、Li Shen 和 Dacheng Tao。量子模仿学习。《IEEE神经网络与学习系统汇刊》，35(10):14190–14204，2024年。[47] Ville Bergholm、Josh Izaac、Maria Schuld、Christian Gogolin、Shahnawaz Ahmed、Vishnu Ajith、M Sohaib Alam、Guillermo Alonso-Linaje、B AkashNarayanan、Ali Asadi 等。PennyLane：量子-经典混合计算的自动微分。arXiv预印本 arXiv:1811.04968，2018年。[48] David Wierichs、Josh Izaac、Cody Wang 和 Cedric Yen-Yu Lin。量子梯度的通用参数移位规则。Quantum，6:677，2022年。[49] C. Schön、E. Solano、F. Verstraete、J. I. Cirac 和 M. M. Wolf。纠缠多量子比特态的序贯生成。《物理评论快报》，95:110503，2005年9月。[50] Shi-Ju Ran。将矩阵乘积态编码到由单量子比特门和双量子比特门构成的量子电路中。《物理评论A》，101:032310，2020年3月。[51] Adam Holmes 和 A. Y. Matsuura。用于精确制备光滑可微函数态的高效量子电路。载于2020年IEEE量子计算与工程国际会议（QCE），第169–179页，2020年。[52] Kouhei Nakaji、Shumpei Uno、Yohichi Suzuki、Rudy Raymond、Tamiya Onodera、Tomoki Tanaka、Hiroyuki Tezuka、Naoki Mitsuda 和 Naoki Yamamoto。浅层参数化量子电路中的近似振幅编码及其在金融市场指标中的应用。《物理评论研究》，4:023136，2022年5月。[53] Manuel S Rudolph、Jacob Miller、Danial Motlagh、Jing Chen、Atithi Acharya 和 Alejandro Perdomo-Ortiz。通过张量网络对参数化量子电路进行协同预训练。《自然·通讯》，14(1):8367，2023年。[54] Manuel S Rudolph、Jing Chen、Jacob Miller、Atithi Acharya 和 Alejandro Perdomo-Ortiz。将矩阵乘积态分解为浅层量子电路。《量子科学与技术》，9(1):015012，2023年11月。[55] Robert R. Tucci。面向量子计算程序员的 Cartan KAK 分解简介。arXiv预印本 quant-ph/0507171，2005年。[56] Y. Lecun、L. Bottou、Y. Bengio 和 P. Haffner。应用于文档识别的基于梯度的学习。《IEEE会刊》，86(11):2278–2324，1998年。[57] Alex Krizhevsky、Geoffrey Hinton 等。学习微小图像中的多层特征。2009年。[58] Richard Socher、Alex Perelygin、Jean Wu、Jason Chuang、Christopher D Manning、Andrew Ng 和 Christopher Potts。情感树库上的语义组合性递归深度模型。载于2013年自然语言处理实证方法会议论文集，第1631–1642页，2013年。[59] Nils Reimers 和 Iryna Gurevych。Sentence-BERT：使用孪生 BERT 网络生成句子嵌入。载于2019年自然语言处理实证方法会议及第九届自然语言处理国际联合会议（EMNLP-IJCNLP）论文集，第3982页。计算语言学协会，2019年。[60] Carsten Blank、Adenilton J Da Silva、Lucas P de Albuquerque、Francesco Petruccione 和 Daniel K Park。紧凑型基于量子核的二元分类器。《量子科学与技术》，7(4):045007，2022年。[61] Pierre Pfeuty。横场一维伊辛模型。《物理学年鉴》，57(1):79–90，1970年。

[62] Ning Ma 和 Heng Li。理解并估算量子程序的执行时间。arXiv 预印本 arXiv:2411.15631，2024年。[63] Lov K. Grover。通过量子计算合成量子叠加态。《物理评论快报》，85:1334–1337，2000年8月。[64] Phillip Kaye 和 Michele Mosca。用于生成任意量子态的量子网络。载于光纤通信会议和国际量子信息会议，第 PB28 页。Optica 出版集团，2001年。[65] Gui-Lu Long 和 Yang Sun。以高效方案将量子寄存器初始化为任意叠加态。《物理评论A》，64:014303，2001年6月。[66] Ville Bergholm、Juha J. Vartiainen、Mikko Möttönen 和 Martti M. Salomaa。具有均匀受控单量子比特门的量子电路。《物理评论A》，71:052330，2005年5月。[67] Martin Plesch 和 Caslav Brukner。通过通用量子门分解进行量子态制备。《物理评论A》，83:032302，2011年3月。[68] Xiaoming Sun、Guojing Tian、Shuai Yang、Pei Yuan 和 Shengyu Zhang。量子态制备与一般酉合成的渐近最优电路深度。IEEE集成电路与系统计算机辅助设计汇刊，42(10):3301–3314，2023年。[69] Niels Gleinig 和 Torsten Hoefler。稀疏量子态制备的高效算法。载于2021年第58届ACM/IEEE设计自动化会议（DAC），第433–438页，2021年。[70] Ulrich Schollwöck。矩阵乘积态时代的密度矩阵重整化群。《物理学年鉴》，326(1):96–192，2011年。2011年1月特刊。[71] Xinbiao Wang、Yuxuan Du、Kecheng Liu、Yong Luo、Bo Du 和 Dacheng Tao。通过无免费午餐定理的视角探究经典与量子学习协议的可分能力。arXiv 预印本 arXiv:2405.07226，2024年。[72] Xuchen You 和 Xiaodi Wu。量子神经网络中的指数多个局部极小值。载于国际机器学习会议，第12144–12155页。PMLR，2021年。[73] Xinbiao Wang、Junyu Liu、Tongliang Liu、Yong Luo、Yuxuan Du 和 Dacheng Tao。量子神经网络中的对称剪枝。载于第十一届国际学习表征会议，2023年。[74] Xuchen You、Shouvanik Chakrabarti 和 Xiaodi Wu。过参数化变分量子本征求解器的收敛理论。arXiv 预印本 arXiv:2205.12481，2022年。[75] Junyu Liu、Khadijeh Najafi、Kunal Sharma、Francesco Tacchino、Liang Jiang 和 Antonio Mezzacapo。宽量子神经网络动力学的解析理论。《物理评论快报》，130(15):150601，2023年。[76] Xuchen You、Shouvanik Chakrabarti、Boyang Chen 和 Xiaodi Wu。分析量子神经网络中的收敛：偏离神经正切核。载于国际机器学习会议，第40199–40224页。PMLR，2023年。[77] Arthur Pesah、M. Cerezo、Samson Wang、Tyler Volkoff、Andrew T. Sornborger 和 Patrick J. Coles。量子卷积神经网络中不存在贫瘠高原。物理评论X，11:041011，2021年10月。[78] Kaining Zhang、Min-Hsiu Hsieh、Liu Liu 和 Dacheng Tao。迈向深度量子神经网络的可训练性。arXiv 预印本 arXiv:2112.15002，2022年。[79] Edward Grant、Leonard Wossnig、Mateusz Ostaszewski 和 Marcello Benedetti。用于解决参数化量子电路中贫瘠高原问题的初始化策略。Quantum，3:214，2019年12月。

[80] Kaining Zhang、Liu Liu、Min-Hsiu Hsieh 和 Dacheng Tao。通过高斯初始化摆脱深度变分量子电路中的贫瘠高原。载于 S. Koyejo、S. Mohamed、A. Agarwal、D. Belgrave、K. Cho 和 A. Oh 编，《神经信息处理系统进展》，第35卷，第18612–18627页。Curran Associates, Inc.，2022年。[81] Yabo Wang、Bo Qi、Chris Ferrie 和 Daoyi Dong。通过缩减域参数初始化提升参数化量子电路的可训练性。《物理评论应用》，22:054005，2024年11月。[82] Yifeng Peng、Xinyi Li、Samuel Yen-Chi Chen、Kaining Zhang、Zhiding Liang、Ying Wang 和 Yuxuan Du。TITAN：一种用于大规模变分量子本征求解器自适应参数冻结的轨迹感知技术。载于第三十九届神经信息处理系统年会，2025年。[83] Marco Cerezo、Akira Sone、Tyler Volkoff、Lukasz Cincio 和 Patrick J Coles。浅层参数化量子电路中依赖于代价函数的贫瘠高原。《自然·通讯》，12(1):1791，2021年。[84] Xinbiao Wang、Yuxuan Du、Zhuozhuo Tu、Yong Luo、Xiao Yuan 和 Dacheng Tao。纠缠数据在量子机器学习中的转变作用。《自然·通讯》，15(1):3716，2024年5月。[85] Kaining Zhang、Junyu Liu、Liu Liu、Liang Jiang、Min-Hsiu Hsieh 和 Dacheng Tao。随机量子数据的诅咒。arXiv 预印本 arXiv:2408.09937，2024年。[86] Giacomo Torlai 和 Matthew Fishman。PastaQ：用于量子计算机模拟、层析成像与分析的软件包，2020年。[87] Farrokh Vatan 和 Colin Williams。一般双量子比特门的最优量子电路。《物理评论A》，69:032315，2004年3月。[88] Patrick Rebentrost、Masoud Mohseni 和 Seth Lloyd。用于大数据分类的量子支持向量机。《物理评论快报》，113:130503，2014年9月。[89] Xinbiao Wang、Yuxuan Du、Yong Luo 和 Dacheng Tao。迈向理解含噪中等规模量子时代量子核的能力。Quantum，5:531，2021年。

A 初步介绍

A.1 记号

本文统一说明全文所使用的记号。我们用 $[N]$ 表示集合 $\{1,\cdots,N\}$。符号 $\boldsymbol{a}_j$ 表示向量 $\boldsymbol{a}$的第 j 个分量。张量积运算记作“$\otimes$”。矩阵 A 的共轭转置记作 $A^\dagger$。矩阵 A 的迹记作 $\mathrm{Tr}[A]$。记号 $\lfloor x \rfloor$ 表示小于或等于 x的最大整数。我们使用 $\mathcal{O}$描述复杂度概念。复数值 x 的相位记作 $\arg(x)$。

对于向量，记号 $\|\cdot\|_2$ 表示 ℓ_2 范数。对于算子，我们使用 Schatten 范数：迹范数 $\|X\|_1 := \mathrm{Tr}\sqrt{X^\dagger X}$，以及算子范数 $\|X\|_\infty := \sup_{\|\psi\|_2=1}\|X\psi\|_2$，其等于 X的最大奇异值。特别地，任意密度算子 ρ 都满足 $\|\rho\|_1 = 1$ 和 $\|\rho\|_\infty \le 1$。

对于作用于算符的线性映射（量子信道）Φ，我们用 $\|\Phi\|_\diamond$ 表示其菱形范数：$\|\Phi\|_\diamond := \sup_{d\ge 1}\sup_{X\neq 0}\frac{\|(\Phi\otimes \mathrm{id}_d)(X)\|_1}{\|X\|_1}$，~~其中 id_d 表~~示作用于 d 维辅助系统上的恒等信道。我们用 I（或 I_d）表示单位矩阵，并用 id 表示底层系统上的恒等信道。量子信道 $\mathcal{E}$ 是完全正且保迹（CPTP）的。

A.2 相关工作

量子数据加载问题已在多个研究方向得到广泛研究。例如，早期研究已证明，对任意 N 量子比特纯态进行精确振幅加载需要指数级的量子资源。具体而言，研究表明，若不假设存在先验结构，则制备一般量子态需要 $\mathcal{O}(2^N)$ 个单量子比特门和双量子比特门 [63–67]。这种尺度与一般量子态的内在复杂度相符：一个 N 量子比特纯态对应于 $\mathbb{C}^{2^N}$ 中具有 $2^{N+1}-1$ 个自由度的归一化向量。即使引入辅助量子比特以降低量子电路深度 [68], ，量子门总数仍呈指数增长。对于具有结构的实例，可以采用更高效的构造方法：例如，一个 S 稀疏向量可以通过 $\mathcal{O}(SN)$ 个 CNOT 门精确加载 [69]。然而，源自经典数据集的大多数量子态（例如图像） [56, 57], 既非稀疏，也非无结构，因此需要采用近似方法。

基于张量网络（TN）的近似加载。张量网络方法利用低纠缠结构，高效地表示量子态并构造相应的加载电路。用于量子数据加载的最广泛应用的张量网络家族是矩阵乘积态（MPS），它通过键维数为 k 的形式，对 N 个量子比特的量子态进行序列化表示。给定一个 MPS，可以构造包含 $\mathcal{O}(Nk^2)$ 个双量子比特门的电路，从而精确制备量子态 [49], ；而截断奇异值分解（SVD）则可对更一般的目标态 [50, 70] 进行可控近似。近期研究已将基于 MPS 的方法应用于合成分布和真实世界的经典数据集。例如，文献 [29] 提出了一种将图像的 MPS 表示转换为量子电路的方法，其电路规模随像素数量呈对数增长，并在囚禁离子设备上通过实验展示了复杂图像的振幅编码。文献 [32] 利用经典图像衰减的傅里叶谱构造近似 MPS编码，从而降低量子电路复杂度；此外，还可以使用受 MPS 结构启发的简单序列电路进一步优化。其他研究则致力于利用源自 MPS 的线性深度电路，高效制备实值平滑概率分布 [33, 51]。这些过程将经典预处理（例如傅里叶特征压缩）与基于 MPS 的电路综合相结合，从而实现高保真度的量子态加载，适用于近期量子硬件。尽管基于张量网络的方法为低纠缠态提供了原理清晰且资源高效的构造方式，但当目标态具有高纠缠或强长程关联时，其适用性会受到限制，因为所需的键维数（对应于电路深度和双量子比特门数量）可能迅速增长。

基于电路的近似加载。基于电路的方法直接构建量子电路，以近似目标态。这些方法大致可分为变分方法和非变分方法。

变分方法。变分方法采用参数化量子电路（参数化量子电路），例如硬件高效型或序贯拟设，并优化其参数，以最小化与目标态之间的非保真度 [34, 52, 53, 71]。对于实值数据向量，使用经典程序以及在计算基中进行的测量；必要时，也会在经哈达玛变换的

基中进行测量，以获取振幅和符号信息 [52]。对于复值数据向量，可直接使用保真度作为代价函数，并采用经典阴影技术，在优化过程中高效估计保真度及其梯度 [34]。此外，有研究提出混合策略，利用基于张量网络的初始化：首先使用张量网络表示（例如 MPS）近似目标态，将其映射为参数化电路，然后进行变分微调 [53]。这种方法保留了张量网络的结构优势，同时将适用范围扩展到严格低纠缠态之外。

非变分方法。非变分方法通过迭代细化局部门来近似目标态。Shirakawa 等人 [35] 提出了 AQCE 算法，该算法通过依次更新双量子比特酉算子来提高保真度。Rudolph 等人 [54] 通过使用 MPS 近似而非平凡恒等算子来初始化 AQCE，改进了这一框架，并展示了更快的收敛速度和更高的准确度。从基于 MPS 的电路出发并进行局部迭代细化的混合工作流，将张量网络的可控近似与电路级更新的自适应灵活性结合起来，从而为中度纠缠的数据集提供高效且准确的加载方式。

尽管这些方法灵活且适用范围广，但仍可能遭遇贫瘠高原和优化陷阱；对于量子比特数较多或纠缠程度较高的目标态，其扩展性也不理想 [72–76]。此外，基于电路的方法仍属于启发式方法。通常缺乏关于保真度和资源成本的严格理论保证，这限制了实用量子数据加载的系统化设计和可预测性。

A.3 典型近似量子加载方法概述

本节总结几种具有代表性的近似量子加载（AQL）方法，并展示如何将它们统一到正文第 3.1 节介绍的通用近似量子加载框架中。回顾一下，在该框架中，近似量子加载过程构造一个量子电路 $U(\boldsymbol{\theta};\,\mathcal{A})$，将目标态 $|\boldsymbol{v}_{\text{target}}\rangle$ 从初始乘积态 $|\psi_{\text{product}}\rangle$加载出来，并通过调整参数 $\boldsymbol{\theta}$、架构 $\mathcal{A}$，或同时调整二者，对其进行优化。

A.3.1 基于张量网络的方法

这里介绍一种基于矩阵乘积态（MPS） [50], 的近似量子加载方法，该矩阵乘积态构成一维张量网络。MPS 方法通过迭代构造当前量子态的键维数为 2 的 MPS 近似。从这一近似中，可以得到一系列 $N-1$ 双量子比特酉算子；将这些算子取逆后应用于当前量子态，以生成更新后的量子态。具体而言，在第（$i+1$）次迭代中，MPS 方法考虑当前量子态

$$|\psi^{(i)}_{\text{MPS}}\rangle = G^{\dagger}_{(N-1)i}\cdots G^{\dagger}_{1}|\boldsymbol{v}_{\text{target}}\rangle, \tag{4}$$

从中得到一组新的双量子比特酉算子

$$G_{(N-1)(i+1)},\ldots,G_{(N-1)i+1} \tag{5}$$

被提取并附加到现有量子电路中。通过充分多次重复该过程，式 (4) 中的量子态趋近于乘积态 $|0\rangle^{\otimes N}$。因此，集合 $\{G_i\}$ 定义了 AQL 的编码电路。我们注意到，每个双量子比特酉算子 G_i 会通过 KAK 分解 [55], 进一步分解为硬件原生量子门，对于实矩阵得到 2 个 CNOT（或 CZ门）量子门，对于一般复矩阵得到 3 个量子门。
在我们的框架中，基于 MPS 的 AQL 对应于依次更新电路架构和参数：

$$U(\boldsymbol{\theta};\mathcal{A}) \to U(\boldsymbol{\theta}\cup\boldsymbol{\theta}_{\text{new}};\mathcal{A}\cup\mathcal{A}_{\text{new}}), \tag{6}$$

而先前添加的量子门保持固定。增加迭代次数会系统性地降低相对于目标态的非保真度。

A.3.2 基于电路的方法

变分： 这里给出一个变分方法的示例，即硬件高效（HE）电路 [52], ，这类电路通常用作 AQL 电路。
HE 电路 $U_{\text{HE}}(\boldsymbol{\theta})$由参数化单量子比特旋转和作用于相邻量子比特的 CNOT 门所构成的重复层组成。每一层

i 先施加一个旋转层 R_Y，接着施加 R_Z，然后施加一个 CNOT 层。CNOT 排列模式在偶数—奇数和奇数—偶数量子比特对之间交替：在第 i 层，CNOT 门施加于量子比特对 $(2n, (2n+1)\%N)$（当 $i \bmod 2 = 0$ 时），以及量子比特对 $(2n+1, (2n+2)\%N)$（当 $i \bmod 2 = 1$ 时），其中 $0 \le 2n \le N-1$ 或 $0 \le 2n \le N-2$（视情况而定）。参数 $\boldsymbol{\theta}$ 通常使用基于梯度的优化器进行更新，以最小化损失函数

$$\ell(\boldsymbol{\theta}) = 1 - |\langle \boldsymbol{v}_{\text{target}} | U_{\text{HE}}(\boldsymbol{\theta}) | \psi_0 \rangle|^2. \tag{7}$$

在我们的通用框架中，HE 电路表示一种变分方法，其中仅优化参数 $\boldsymbol{\theta}$，而电路架构 $\mathcal{A}$ 保持固定。

非变分方法：自动量子电路编码（AQCE）方法 [35] 是一种非变分方法，它以正向和反向交替的方式，迭代更新预先指定电路中的双量子比特酉算子。假设当前的编码酉算子列表为 $\{G_1, \ldots, G_M\}$。第 m个双量子比特门的局部最优更新如下。对于给定的双量子比特子系统 $\mathcal{Q}_m$，定义

$$F_m = \operatorname*{Tr}_{[N]\setminus\mathcal{Q}_m} \left[G_{m+1}^\dagger \cdots G_M^\dagger |\boldsymbol{v}_{\text{target}}\rangle\langle\psi_0| G_1^\dagger \cdots G_{m-1}^\dagger \right], \tag{8}$$

其中 $|\psi_0\rangle$ 是初始乘积态。进行奇异值分解（SVD）

$$F_m = XDY, \tag{9}$$

更新后的双量子比特酉算子为

$$G_m^{\text{new}} = XY. \tag{10}$$

实际上，该算法会考虑双量子比特子系统 $\mathcal{Q}_m$ 的所有可能选择，并选出使 $|\mathrm{Tr}[D]|$ 最大的那个。这一选择可以这样理解：与更新后的酉算子对应的编码态和目标态之间的保真度为

$$\begin{aligned}
F &= \left| \langle\psi_0| G_1^\dagger \cdots G_{m-1}^\dagger G_m^{\text{new}\dagger} G_{m+1}^\dagger \cdots G_M^\dagger |\boldsymbol{v}_{\text{target}}\rangle \right|^2 \\
&= \left| \operatorname{Tr}\left[G_m^{\text{new}\dagger} G_{m+1}^\dagger \cdots G_M^\dagger |\boldsymbol{v}_{\text{target}}\rangle\langle\psi_0| G_1^\dagger \cdots G_{m-1}^\dagger \right] \right|^2 \\
&= \left| \operatorname{Tr}_{\mathcal{Q}_m}\left[G_m^{\text{new}\dagger} \operatorname{Tr}_{[N]\setminus\mathcal{Q}_m}\left[G_{m+1}^\dagger \cdots G_M^\dagger |\boldsymbol{v}_{\text{target}}\rangle\langle\psi_0| G_1^\dagger \cdots G_{m-1}^\dagger \right] \right] \right|^2 \\
&= \left| \operatorname{Tr}_{\mathcal{Q}_m}\left[G_m^{\text{new}\dagger} F_m \right] \right|^2 \\
&= \left| \operatorname{Tr}\left[Y^\dagger X^\dagger X D Y \right] \right|^2 \\
&= \left| \operatorname{Tr}\left[D \right] \right|^2 .
\end{aligned}$$

因此，最大化 $|\mathrm{Tr}[D]|$ 相当于最大化保真度 F。

经过若干次前向-后向扫描后，按照相同的流程添加新的量子门：将 $G_{M+1} = I$视为.....。随后，通过 KAK 分解 [55], 进一步将最终的双量子比特酉算子序列 $\{G_i\}$ 分解为适合硬件实现的量子门：对于实矩阵，需要 2 个 CNOT（或 CZ门）量子门；对于一般复矩阵，则需要 3 个量子门。

在我们的框架中，AQCE 自然属于一种非变分方法，它会依次更新电路参数 $\boldsymbol{\theta}$（由更新后的双量子比特酉算子的 KAK 分解得到）和电路架构 $\mathcal{A}$（由更新后的酉算子作用于哪些量子比特对来决定）。

B 定理证明

B.1 技术引理

在证明主定理之前，我们先给出一个技术引理，该引理给出了任意单量子比特纯态对混合单量子比特态进行近似时所能达到的最大保真度。

引理 B.1。 记 ρ为一个单量子比特混合态。则

$$\max_{|\phi\rangle} \mathrm{Tr}[|\phi\rangle\langle\phi|\rho] = \frac{1 + \sqrt{2^{1-\mathcal{S}(\rho)} - 1}}{2}, \tag{11}$$

其中 $\mathcal{S}(\rho)$ 表示量子态 ρ的 *Rényi-2* 熵。

证明。为方便起见，我们将 ρ的谱分解记为 $\rho=\sum_{j=1}^{2}\lambda_j|\psi_j\rangle\langle\psi_j|$，其中$\lambda_{1,2}\in\mathbb{R}$，且根据密度矩阵的性质，$\lambda_1+\lambda_2=\mathrm{Tr}[\rho]=1$。不失一般性，假设 $\lambda_1\geq\lambda_2$。因此，当 $|\phi\rangle=|\psi_1\rangle$时，式（11）左侧取得最大值，即

$$\max_{|\phi\rangle}\mathrm{Tr}[|\phi\rangle\langle\phi|\rho]=\mathrm{Tr}[|\psi_1\rangle\langle\psi_1|\rho]=\lambda_1. \tag{12}$$

此外，谱分解也可用于Rényi熵

$$\begin{aligned}\mathcal{S}(\rho)=&-\log_2\mathrm{Tr}[\rho^2]\\ =&-\log_2\mathrm{Tr}\left[\sum_{j=1}^{2}\sum_{k=1}^{2}\lambda_j\lambda_k|\psi_j\rangle\langle\psi_j|\psi_k\rangle\langle\psi_k|\right]\\ =&-\log_2\mathrm{Tr}\left[\sum_{j=1}^{2}\lambda_j^2|\psi_j\rangle\langle\psi_j|\right]\\ =&-\log_2\left(\lambda_1^2+\lambda_2^2\right).\end{aligned} \tag{13}$$

将式 (13) 与 $\lambda_1+\lambda_2=1$结合，可得到 λ_1 关于Rényi熵 $\mathcal{S}(\rho)$的表达式，如下：

$$\lambda_1=\frac{1+\sqrt{2^{1-\mathcal{S}(\rho)}-1}}{2}. \tag{14}$$

比较式 (12) 和 (14)，可得

$$\max_{|\phi\rangle}\mathrm{Tr}[|\phi\rangle\langle\phi|\rho]=\frac{1+\sqrt{2^{1-\mathcal{S}(\rho)}-1}}{2}.$$

因此，引理 B.1 得证。

□

引理 B.2。用 ρ_{ij} 表示 ρ的密度矩阵形式中的(i,j)项。则引理 *B.1* 中的最大保真度可由 $|\phi\rangle=R_Z(\beta)R_Y(\gamma)|0\rangle$达到，其中 $\beta=\arg(\rho_{10})$ 且$\gamma=\frac{\pi}{2}-\arcsin\frac{\rho_{00}-\rho_{11}}{\sqrt{4|\rho_{10}|^2+(\rho_{00}-\rho_{11})^2}}$。

证明。首先，我们将量子态 $|\phi\rangle$展开为 β 和 γ的形式：

$$|\phi\rangle=e^{-\beta Z/2}e^{-\gamma Y/2}|0\rangle=\begin{bmatrix}e^{-i\frac{\beta}{2}}&0\\0&e^{i\frac{\beta}{2}}\end{bmatrix}\begin{bmatrix}\cos\frac{\gamma}{2}&-\sin\frac{\gamma}{2}\\\sin\frac{\gamma}{2}&\cos\frac{\gamma}{2}\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}e^{-i\frac{\beta}{2}}\cos\frac{\gamma}{2}\\e^{i\frac{\beta}{2}}\sin\frac{\gamma}{2}\end{bmatrix}. \tag{15}$$

那么，ρ 和 $|\phi\rangle$ 之间的保真度为

$$\begin{aligned}\mathrm{Tr}[|\phi\rangle\langle\phi|\rho]=&\mathrm{Tr}\left[\begin{bmatrix}\frac{1}{2}+\frac{1}{2}\cos\gamma&\frac{1}{2}e^{-i\beta}\sin\gamma\\\frac{1}{2}e^{i\beta}\sin\gamma&\frac{1}{2}-\frac{1}{2}\cos\gamma\end{bmatrix}\begin{bmatrix}\rho_{00}&\rho_{01}\\\rho_{10}&\rho_{11}\end{bmatrix}\right]\\ =&\left(\frac{1}{2}+\frac{1}{2}\cos\gamma\right)\rho_{00}+\frac{1}{2}e^{-i\beta}\sin\gamma\rho_{10}+\frac{1}{2}e^{i\beta}\sin\gamma\rho_{01}+\left(\frac{1}{2}-\frac{1}{2}\cos\gamma\right)\rho_{11}\\ =&\frac{1}{2}+\frac{\rho_{00}-\rho_{11}}{2}\cos\gamma+\mathrm{Re}\left[e^{-i\beta}\rho_{10}\right]\sin\gamma,\end{aligned} \tag{16}$$

其中，式（16）由密度矩阵的性质 $\mathrm{Tr}[\rho]=1$ 和 $\rho=\rho^\dagger$ 得出。随后，考虑 β 和 γ 关于 ρ 的表达式，可得

$$\begin{aligned}\mathrm{Tr}[|\phi\rangle\langle\phi|\rho]=&\frac{1}{2}+\frac{\rho_{00}-\rho_{11}}{2}\cos\gamma+|\rho_{10}|\sin\gamma\\ =&\frac{1}{2}+\sqrt{\left(\frac{\rho_{00}-\rho_{11}}{2}\right)^2+|\rho_{10}|^2}.\end{aligned} \tag{17}$$

此外，Rényi熵 $\mathcal{S}(\rho)$ 可用 ρ 表示如下

$$\begin{aligned}\mathcal{S}(\rho) =& -\log_2 \mathrm{Tr}[\rho^2] \\ =& -\log_2\left(\rho_{00}^2+\rho_{11}^2+2|\rho_{10}|^2\right).\end{aligned}$$

因此，我们有

$$\begin{aligned}2^{1-S(\rho)}-1 &= 2\left(\rho_{00}^2+\rho_{11}^2+2|\rho_{10}|^2\right)-1 \\ &= 2\rho_{00}^2+2\rho_{11}^2-\left(\rho_{00}+\rho_{11}\right)^2+4|\rho_{10}|^2 \\ &= \left(\rho_{00}-\rho_{11}\right)^2+4|\rho_{10}|^2. \end{aligned} \tag{18}$$

因此，结合式（17）、（18）和引理 B.1 中的最大保真度值，即可证明引理 B.2。

□

引理 B.3。 设 $\mathcal{N}$ 是一个作用于 N 个量子比特的 *CPTP* 映射，并设 $\mathcal{U}_\ell(\rho):=U_\ell\rho U_\ell^\dagger$ 表示与 N 量子比特酉算子 U_ℓ 相关联的酉信道，其中每个 $\ell\in[L]$ 均如此。给定一个初始态 ρ_0，将理想输出态和含噪输出态定义为

$$\rho := \big(\mathcal{U}_L\circ\cdots\circ\mathcal{U}_1\big)(\rho_0), \tag{19}$$

$$\hat{\rho} := \big(\mathcal{N}\circ\mathcal{U}_L\circ\mathcal{N}\circ\cdots\circ\mathcal{N}\circ\mathcal{U}_1\circ\mathcal{N}\big)(\rho_0), \tag{20}$$

其中，$\hat{}\rho$ 是通过插入 $L+1$ 层噪声 $\mathcal{N}$，并使其与 L 个酉算子交替而得到的。于是，

$$\|\hat{\rho}-\rho\|_1 \le (L+1)\,\|\mathcal{N}-\mathrm{id}\|_\diamond, \tag{21}$$

其中，id 表示恒等信道，$\|\cdot\|_\diamond$ 表示菱形范数。

证明。 为方便起见，我们将无噪声信道记为

$$\mathcal{E}_0^{(L)} := \mathcal{U}_L\circ\cdots\circ\mathcal{U}_1 \tag{22}$$

并将完全含噪声信道记为

$$\mathcal{E}_{\mathcal{N}}^{(L)} := \mathcal{N}\circ\mathcal{U}_L\circ\mathcal{N}\circ\cdots\circ\mathcal{N}\circ\mathcal{U}_1\circ\mathcal{N}. \tag{23}$$

根据定义，

$$\rho=\mathcal{E}_0^{(L)}(\rho_0), \qquad \hat{\rho}=\mathcal{E}_{\mathcal{N}}^{(L)}(\rho_0). \tag{24}$$

因此，

$$\|\hat{\rho}-\rho\|_1 = \left\|\mathcal{E}_{\mathcal{N}}^{(L)}(\rho_0)-\mathcal{E}_0^{(L)}(\rho_0)\right\|_1 \le \left\|\mathcal{E}_{\mathcal{N}}^{(L)}-\mathcal{E}_0^{(L)}\right\|_\diamond, \tag{25}$$

其中，不等式由菱形范数的定义可得；菱形范数对任意输入态下输出的迹范数差给出了上界。

接下来，我们推导 $\left\|\mathcal{E}_{\mathcal{N}}^{(L)}-\mathcal{E}_0^{(L)}\right\|_\diamond$ 的一个上界。噪声信道 $\mathcal{E}_{\mathcal{N}}^{(L)}$ 中有 $L+1$ 个噪声位置：U_1 之前一个、每对连续的 U_j 和 U_{j+1} 之间各一个，以及 U_L 之后一个。我们构造一列中间信道 $\{\Phi^{(k)}\}_{k=0}^{L+1}$，使得

$$\Phi^{(0)}=\mathcal{E}_0^{(L)}, \qquad \Phi^{(L+1)}=\mathcal{E}_{\mathcal{N}}^{(L)}, \tag{26}$$

并且对于每个 $k\in\{1,\dots,L\}$，信道 $\Phi^{(k)}$ 是通过应用前 k 个噪声层得到的。因此，我们有如下望远镜分解

$$\mathcal{E}_{\mathcal{N}}^{(L)}-\mathcal{E}_0^{(L)} = \Phi^{(L+1)}-\Phi^{(0)} = \sum_{k=1}^{L+1}\left(\Phi^{(k)}-\Phi^{(k-1)}\right). \tag{27}$$

每个 $\Phi^{(k)}-\Phi^{(k-1)}$ 都对应于将一个恒等信道替换为 $\mathcal{N}$，而其他所有位置保持不变。因此，存在 CPTP 映射 $\Lambda^{(k)}{}_{\mathrm{L}}$ 和 $\Lambda^{(k)}{}_{\mathrm{R}}$（由酉信道 $\mathcal{U}_\ell$ 与噪声/恒等信道复合而成），使得

$$\Phi^{(k)}-\Phi^{(k-1)} = \Lambda_{\mathrm{R}}^{(k)}\circ(\mathcal{N}-\mathrm{id})\circ\Lambda_{\mathrm{L}}^{(k)}. \tag{28}$$

由于 $\Lambda^{(k)}{}_{\mathrm{L}}$ 和 $\Lambda^{(k)}{}_{\mathrm{R}}$ 都是 CPTP 映射，它们的菱形范数满足

$$\|\Lambda_{\mathrm{L}}^{(k)}\|_\diamond = \|\Lambda_{\mathrm{R}}^{(k)}\|_\diamond = 1. \tag{29}$$

因此，我们有

$$\left\|\Phi^{(k)} - \Phi^{(k-1)}\right\|_\diamond \le \|\Lambda_{\mathrm{R}}^{(k)}\|_\diamond \, \|\mathcal{N} - \mathrm{id}\|_\diamond \, \|\Lambda_{\mathrm{L}}^{(k)}\|_\diamond = \|\mathcal{N} - \mathrm{id}\|_\diamond. \tag{30}$$

利用三角不等式，并对所有 $(L+1)$ 个噪声位置求和，可得

$$\begin{aligned}\left\|\mathcal{E}_{\mathcal{N}}^{(L)} - \mathcal{E}_0^{(L)}\right\|_\diamond &\le \sum_{k=1}^{L+1} \left\|\Phi^{(k)} - \Phi^{(k-1)}\right\|_\diamond \\ &\le (L+1)\,\|\mathcal{N} - \mathrm{id}\|_\diamond.\end{aligned} \tag{31}$$

结合式 (25) 和式 (31)，可得

$$\|\hat{\rho} - \rho\|_1 \le \left\|\mathcal{E}_{\mathcal{N}}^{(L)} - \mathcal{E}_0^{(L)}\right\|_\diamond \le (L+1)\,\|\mathcal{N} - \mathrm{id}\|_\diamond. \tag{32}$$

因此，我们证明了引理 B.3。 □

B.2 主定理的证明

这里给出正文中定理的完整表述及其证明。

定理 B.4. 将 N 量子比特量子态 $|\psi\rangle$ 的纠缠度量记为 $\mathcal{S}(|\psi\rangle) = \sum_{i=1}^{N} \mathcal{S}_{\{i\}}(|\psi\rangle)$。那么，对于量子态 $|\boldsymbol{v}_{\mathrm{target}}\rangle$ 和量子电路 U，若满足 $\mathcal{S}(U^\dagger|\boldsymbol{v}_{\mathrm{target}}\rangle) = S$，则 $|\boldsymbol{v}_{\mathrm{target}}\rangle$ 与从任意乘积态 $|\psi_{\mathrm{product}}\rangle$ 经 U 生成的量子态之间的非保真度有如下下界：

$$1 - |\langle \boldsymbol{v}_{\mathrm{target}}|U|\psi_{\mathrm{product}}\rangle|^2 \ge f_1(S) := \frac{1 - \sqrt{2^{1-\frac{1}{N}S} - 1}}{2}. \tag{33}$$

此外，在能够访问 $U^\dagger|\boldsymbol{v}_{\mathrm{target}}\rangle$ 的情况下，我们可以构造一个乘积态 $|\psi'_{\mathrm{product}}\rangle$，使得非保真度的上界如下：

$$1 - \left|\langle \boldsymbol{v}_{\mathrm{target}}|U|\psi'_{\mathrm{product}}\rangle\right|^2 \le f_2(S) := \frac{1}{2}\left(1 - \sqrt{2^{1-S+\lfloor S\rfloor} - 1} + \lfloor S\rfloor\right), \tag{34}$$

其中，$\lfloor\cdot\rfloor$ 表示向下取整函数。通过计算泰勒展开，我们注意到 $f_1(S) = \frac{\ln 2}{2N} S + \mathcal{O}(S^3)$ 且 $f_2(S) = \frac{\ln 2}{2} S + \mathcal{O}(S^3)$。

证明。为方便起见，我们用 ρ 表示量子态 $U^\dagger|\boldsymbol{v}_{\mathrm{target}}\rangle$的密度矩阵形式。我们用 $\rho_n := \mathrm{Tr}_{[N]/\{n\}}[\rho]$ 表示 n个量子比特子系统中的第 n个单量子比特子系统的密度矩阵，并将 Rényi熵 $S_n := \mathcal{S}(\rho_n)$ 简记为 $S_n := \mathcal{S}(\rho_n)$。由于 ρ_n 是单量子比特密度矩阵，因此 $S_n \in [0,1]$。我们构造一组一般的单量子比特投影算符 $\{|\psi_n\rangle\}_{n=1}^N$，满足 $|\psi\rangle = \otimes_{n=1}^N |\psi_n\rangle$。相应地，我们定义一组随机变量 $\{y_n\}_{n=1}^N$，其中，若对第 n 个量子比特的测量得到 $|\psi_n\rangle$，则 $y_n = 1$；否则 $y_n = 0$。因此，我们可以将 $|\boldsymbol{v}_{\mathrm{target}}\rangle$ 与$U|\psi\rangle$ 之间的非保真度改写为

$$1 - |\langle \boldsymbol{v}_{\mathrm{target}}|U|\psi\rangle|^2 = 1 - \mathrm{Tr}\left[\rho\left(\otimes_{n=1}^N |\psi_n\rangle\langle\psi_n|\right)\right] = 1 - \Pr\big(y_n = 1,\ \forall n \in [N]\big). \tag{35}$$

首先，我们推导出式（35）中非保真度的下界如下。我们有

$$\begin{aligned}1 - \Pr\big(y_n = 1,\ \forall n \in [N]\big) &\ge 1 - \min_{n\in[N]} \Pr\big(y_n = 1\big) \\ &= 1 - \min_{n\in[N]} \mathrm{Tr}\left[\rho\left(I^{\otimes(n-1)} \otimes |\psi_n\rangle\langle\psi_n| \otimes I^{\otimes(N-n)}\right)\right] \\ &= 1 - \min_{n\in[N]} \mathrm{Tr}\left[\rho_n |\psi_n\rangle\langle\psi_n|\right] \\ &\ge 1 - \min_{n\in[N]} \frac{1 + \sqrt{2^{1-S_n} - 1}}{2},\end{aligned} \tag{36}$$

其中，最后一个不等式由引理 B.1 得出。我们进一步处理式 36 中的各项。具体而言，函数 $f(x) = \sqrt{2^{1-x}-1}$ 对于 $x \in [0,1]$ 递减。因此，

$$\min_{n\in[N]} \sqrt{2^{1-S_n}-1} = \min_{n\in[N]} f(S_n) \le f\left(\frac{1}{N}\sum_{n=1}^{N} S_n\right) = f\left(\frac{S}{N}\right) = \sqrt{2^{1-\frac{S}{N}}-1}. \tag{37}$$

结合式 (36) 和 (37)，我们得到

$$1 - |\langle \boldsymbol{v}_{\text{target}}|U|\psi\rangle|^2 \ge \frac{1-\sqrt{2^{1-\frac{S}{N}}-1}}{2}. \tag{38}$$

对于 $S \to 0$ 的情况，通过计算泰勒展开式，式 (38) 趋于 $\frac{\ln 2}{2N}^2 S + \mathcal{O}(S^3)$ n.

接下来，我们推导非保真度的上界。我们考虑一组特定的投影算子 $\{|\phi_n\rangle\}_{n=1}^N$，其中每个 $|\phi_n\rangle$ 都是 $\rho_{\{n\}}$ 的纯态近似，并且根据引理 B.1 和 B.2 具有最大的保真度。类似地，我们定义一组随机变量 $\{x_n\}_{n=1}^N$，其中，若测量 ρ 的第 n 个量子比特得到量子态 $|\phi_n\rangle$，则 $x_n = 1$；否则为 $x_n = 0$。根据引理 B.1，我们有

$$\Pr\big(x_n=1\big) = \frac{1+\sqrt{2^{1-S_n}-1}}{2},\ \Pr\big(x_n=0\big) = \frac{1-\sqrt{2^{1-S_n}-1}}{2}. \tag{39}$$

因此，量子态 $|\boldsymbol{v}_{\text{target}}\rangle$ 与 $U|\phi\rangle = U(\otimes_{n=1}^N |\phi_n\rangle)$ 之间的非保真度可以改写为

$$1 - |\langle \boldsymbol{v}_{\text{target}}|U|\phi\rangle|^2 = 1 - \text{Tr}\left[\rho\left(\otimes_{n=1}^N |\phi_n\rangle\langle\phi_n|\right)\right] = 1 - \Pr\big(x_n=1,\ \forall n\in[N]\big). \tag{40}$$

下面，我们从式 (40) 中的概率出发进行推导，以获得非保真度的上界。我们有

$$\begin{aligned}
\Pr\big(x_n=1,\ \forall n\in[N]\big) &= 1 - \Pr\big(\exists\, n\in[N],\ \text{s.t. } x_n=0\big) \\
&\ge 1 - \sum_{n=1}^{N} \Pr\big(x_n=0\big) \\
&= 1 - \sum_{n=1}^{N} \frac{1-\sqrt{2^{1-S_n}-1}}{2},
\end{aligned} \tag{41}$$

其中，式 (41) 由式 (39) 推得。函数 $f(x) = \sqrt{2^{1-x}-1}$ 对于 $x \in [0,1]$ 是凹函数，因为其二阶导数为 $f''(x) = \frac{(\ln 2)^2\, 2^{1-x}(2^{1-x}-2)}{4(2^{1-x}-1)^{3/2}} \le 0$ 对于 $x \in [0,1]$。因此，我们有

$$f(a)+f(b) \ge \begin{cases} f(0)+f(a+b) & \text{when } a,b,a+b\in[0,1], \\ f(a+b-1)+f(1) & \text{when } a,b\in[0,1] \text{ and } a+b\in[1,2]. \end{cases} \tag{42}$$

由于凹函数的性质，可以通过选择量子比特对 (i,j) 使得 $S_i, S_j \notin \{0,1\}$，并将新值 S_i', S_j' 推向 $[0,1]$ 的边界，将式 (42) 应用于式 (41)。最多执行上述过程 $N-1$ 次后，我们可以将概率界定如下

$$\begin{aligned}
\Pr\big(x_n=1,\ \forall n\in[N]\big) &\ge 1 - \frac{N}{2} + \frac{1}{2}\Big(\lfloor S_{\boldsymbol{\theta}}\rfloor f(1) + f(S-\lfloor S\rfloor) + (N-\lfloor S\rfloor - 1)\, f(0)\Big) \\
&= \frac{1}{2}\left(\sqrt{2^{1-S+\lfloor S\rfloor}-1} - \lfloor S\rfloor + 1\right),
\end{aligned} \tag{43}$$

其中，式 (43) 通过计算函数 f 的值给出。因此，我们得到非保真度的上界：

$$1 - |\langle \boldsymbol{v}_{\text{target}}|U|\phi\rangle|^2 = 1 - \Pr\big(x_n=1,\ \forall n\in[N]\big) \le \frac{1}{2}\left(1 - \sqrt{2^{1-S+\lfloor S\rfloor}-1} + \lfloor S\rfloor\right). \tag{44}$$

对于 $S \to 0$ 的情况，通过计算泰勒展开，式 (44) 趋于 $\frac{\ln 2}{2}^2 S + \mathcal{O}(S^3)$。因此，我们证明了定理 B.4。

□

C 噪声情形的理论结果

虽然 AQL 是为无噪声和有噪声两种情形设计的，但其在有噪声情形下的表现对于近期应用尤为关键。在实际的含噪中等规模量子（NISQ）设置中，硬件噪声会扰动量子电路的演化，因此需要分析其对误差的影响是否超出无噪声界限。这里，我们将定理 B.4 的结果推广到一般的完全正迹保持通道（CPTP 通道）情形。

定理 C.1. 记 N 量子比特（纯态或混态）量子态 ρ 的纠缠度量为 $\mathcal{S}(\rho)=\sum_{n=1}^{N}\mathcal{S}_{\{n\}}(\rho)$。设 $\mathcal{S}(\mathcal{E}(\rho_{\text{target}}))=S$，其中 $\mathcal{E}$ 是一个 *CPTP* 映射。那么，对于任意乘积态$|\psi_{\text{product}}\rangle$，$\mathcal{E}(\rho_{\text{target}})$ 与 $|\psi_{\text{product}}\rangle$ 之间的非保真度下界为

$$1-\langle\psi_{\text{product}}|\mathcal{E}(\rho_{\text{target}})|\psi_{\text{product}}\rangle\geq f_1(S):=\frac{1-\sqrt{2^{1-\frac{1}{N}S}-1}}{2}. \tag{45}$$

此外，若能访问 $\mathcal{E}(\rho_{\text{target}})$，我们可以构造一个乘积态 $|\psi'_{\text{product}}\rangle$，使得非保真度的上界为

$$1-\langle\psi'_{\text{product}}|\mathcal{E}(\rho_{\text{target}})|\psi'_{\text{product}}\rangle\leq f_2(S):=\frac{1}{2}\left(1-\sqrt{2^{1-S+\lfloor S\rfloor}-1}+\lfloor S\rfloor\right), \tag{46}$$

其中 $\lfloor\cdot\rfloor$ 表示向下取整函数。我们注意到，通过计算泰勒展开，$f_1(S)=\frac{\ln 2}{2N}S+\mathcal{O}(S^3)$ 且 $f_2(S)=\frac{\ln 2}{2}S+\mathcal{O}(S^3)$。

接下来，我们考虑分层含噪电路模型，并推导与噪声相关的界，如定理 C.2 所述。

定理 *C.2.* 记 N 量子比特量子态 ρ 的纠缠度量为 $\mathcal{S}(\rho)=\sum_{n=1}^{N}\mathcal{S}_{\{n\}}(\rho)$。设 $|\boldsymbol{v}_{\text{target}}\rangle$ 为一个 N 量子比特纯目标态，并令 $\rho_{\text{target}}=|\boldsymbol{v}_{\text{target}}\rangle\langle\boldsymbol{v}_{\text{target}}|$。设 $\mathcal{M}$、$\mathcal{N}$ 为两个 *CPTP* 噪声信道，且对于酉算子 U_ℓ，有 $\mathcal{U}_\ell(\rho):=U_\ell\rho U_\ell^\dagger$。考虑两个 L 层含噪电路$\mathcal{E}^{(L)}:=\mathcal{M}\circ\mathcal{U}_L\circ\mathcal{M}\circ\cdots\circ\mathcal{M}\circ\mathcal{U}_1\circ\mathcal{M}$ 和 $\mathcal{F}^{(L)}:=\mathcal{N}\circ\mathcal{U}_1^\dagger\circ\mathcal{N}\circ\cdots\circ\mathcal{N}\circ\mathcal{U}_L^\dagger\circ\mathcal{N}$。假设$\mathcal{S}(\mathcal{F}^{(L)}(\rho_{\text{target}}))=S$。那么，对于任意乘积态 $|\psi_{\text{product}}\rangle$，目标态 ρ_{target} 与量子态 $\mathcal{E}^{(L)}(|\psi_{\text{product}}\rangle\langle\psi_{\text{product}}|)$ 之间的非保真度有如下下界

$$1-\mathrm{Tr}[\rho_{\text{target}}\mathcal{E}^{(L)}(|\psi_{\text{product}}\rangle\langle\psi_{\text{product}}|)]\geq f_1(S)-(L+1)\left(\|\mathcal{M}-\mathrm{id}\|_\diamond+\|\mathcal{N}-\mathrm{id}\|_\diamond\right). \tag{47}$$

此外，若可访问 $\mathcal{F}^{(L)}(\rho_{\text{target}})$，我们可以构造一个乘积态 $|\psi'_{\text{product}}\rangle$，使其非保真度的上界为

$$1-\mathrm{Tr}[\rho_{\text{target}}\mathcal{E}^{(L)}(|\psi'_{\text{product}}\rangle\langle\psi'_{\text{product}}|)]\leq f_2(S)+(L+1)\left(\|\mathcal{M}-\mathrm{id}\|_\diamond+\|\mathcal{N}-\mathrm{id}\|_\diamond\right). \tag{48}$$

函数 f_1 和 f_2 遵循定理 *C.1* 中的定义。

以上结果表明，定理 B.4 中由纠缠度量决定的项在含噪情形下依然成立，此外还多出一个随深度 L 和噪声强度线性增长的噪声项。更具体地说，噪声强度由累积量 $(L+1)(\|\mathcal{M}-\mathrm{id}\|_\diamond+\|\mathcal{N}-\mathrm{id}\|_\diamond)$刻画。对于具体的 CPTP 噪声模型，这一修正很容易理解。例如，错误率为 p 的去极化信道 $\mathcal{D}_p(\rho)=(1-p)\rho+pI/d$ 满足 $\|\mathcal{D}_p-\mathrm{id}\|_\diamond=\mathcal{O}(p)$，因此带来的修正量级为 $(L+1)p$。退相干和振幅阻尼通道等其他常见噪声模型也具有类似的线性标度。当纠缠度量和累积噪声都适中时，AQL 可在含噪中等规模量子硬件上实现较小的近似误差。

我们指出，对于噪声量子信道，纠缠度量 $\mathcal{S}$ 通常高于相应的无噪声情形。这里我们在定理 C.3 中给出了去极化信道的一个示例。因此，由于定理 B.4 和定理 C.1 中的界具有完全相同的形式，含噪电路达到的非保真度会高于无噪声线路。

定理 C.3. 设 N 量子比特（纯态或混合态）量子态 ρ 的纠缠度量为 $\mathcal{S}(\rho)=\sum_{n=1}^{N}\mathcal{S}_{\{n\}}(\rho)$。令 $\mathcal{D}_p\rho=(1-p)\rho+pI/d$ 为错误率为 p 的去极化信道。那么

$$\left(1-\frac{p}{\ln 4}\right)\mathcal{S}(\rho)+\frac{Np}{\ln 4}\leq\mathcal{S}(\mathcal{D}_p(\rho))\leq\mathcal{S}(\rho)+N\log_2\frac{2}{1+(1-p)^2}. \tag{49}$$

C.1 定理 C.1 的证明。

证明. 为方便起见，我们将 ρ 量子比特量子态 $\mathcal{E}(\rho_{\text{target}})$ 记为 N，将 $\rho_n := \text{Tr}_{[N]/\{n\}}[\rho]$ 记为量子态 ρ 的第 n 个单量子比特子系统的密度矩阵。为简便起见，我们将 Rényi 熵 $S_n := \mathcal{S}(\rho_n)$ 记为 ρ_n。由于 ρ_n 是单量子比特密度矩阵，$S_n \in [0,1]$。我们构造一组一般的单量子比特投影算符 $\{|\psi_n\rangle\}_{n=1}^N$，满足 $|\psi_{\text{product}}\rangle = \otimes_{n=1}^N |\psi_n\rangle$。相应地，我们定义一组随机变量 $\{y_n\}_{n=1}^N$：若对第 n 个量子比特进行测量得到 $|\psi_n\rangle$，则 $y_n = 1$；否则 $y_n = 0$。因此，我们可以将 ρ 与 $|\psi_{\text{product}}\rangle$ 之间的非保真度改写为

$$\begin{aligned} 1 - \langle\psi_{\text{product}}|\rho|\psi_{\text{product}}\rangle &= 1 - \text{Tr}\left[\rho\left(\otimes_{n=1}^N |\psi_n\rangle\langle\psi_n|\right)\right] \\ &= 1 - \Pr\big(y_n = 1,\ \forall n \in [N]\big). \end{aligned} \tag{50}$$

首先，我们推导式 (50) 中非保真度的下界如下。我们有

$$\begin{aligned} 1 - \Pr\big(y_n = 1,\ \forall n \in [N]\big) &\geq 1 - \min_{n\in[N]} \Pr\big(y_n = 1\big) \\ &= 1 - \min_{n\in[N]} \text{Tr}\left[\rho\left(I^{\otimes(n-1)} \otimes |\psi_n\rangle\langle\psi_n| \otimes I^{\otimes(N-n)}\right)\right] \\ &= 1 - \min_{n\in[N]} \text{Tr}\left[\rho_n |\psi_n\rangle\langle\psi_n|\right] \\ &\geq 1 - \min_{n\in[N]} \frac{1 + \sqrt{2^{1-S_n} - 1}}{2}, \end{aligned} \tag{51}$$

其中，最后一个不等式由引理 B.1 得出，该引理根据 Rényi 熵给出了单量子比特密度矩阵与纯态之间的最大保真度。

接下来，我们继续处理式 (51) 中的各项。特别地，函数 $f(x) = \sqrt{2^{1-x} - 1}$ 在 $x \in [0,1]$ 时递减。因此，

$$\min_{n\in[N]} \sqrt{2^{1-S_n} - 1} = \min_{n\in[N]} f(S_n) \leq f\left(\frac{1}{N}\sum_{n=1}^N S_n\right) = f\left(\frac{S}{N}\right) = \sqrt{2^{1-\frac{S}{N}} - 1}. \tag{52}$$

结合式 (51) 和式 (52)，我们得到

$$1 - \langle\psi_{\text{product}}|\rho|\psi_{\text{product}}\rangle \geq \frac{1 - \sqrt{2^{1-\frac{S}{N}} - 1}}{2} =: f_1(S). \tag{53}$$

对于 $S \to 0$ 的情况，通过计算泰勒展开，式 (53) 趋于 $\frac{\ln 2}{2N^2} S + \mathcal{O}(S^3)$.

接下来，我们推导非保真度的上界。我们考虑一组特定的投影算子$\{|\phi_n\rangle\}_{n=1}^N$，其中每个 $|\phi_n\rangle$ 都是 $\rho_{\{n\}}$ 的纯态近似，且根据引理 B.1 和 B.2，其保真度最大。也就是说，选择 $|\phi_n\rangle$ 以最大化 $\text{Tr}[\rho_n|\phi_n\rangle\langle\phi_n|]$。类似地，我们定义一组随机变量 $\{x_n\}_{n=1}^N$，其中，若测量 ρ 的第 n 个量子比特后得到量子态 $|\phi_n\rangle$，则 $x_n = 1$；否则为 $x_n = 0$。根据引理 B.1，我们有

$$\Pr\big(x_n = 1\big) = \frac{1 + \sqrt{2^{1-S_n} - 1}}{2}, \quad \Pr\big(x_n = 0\big) = \frac{1 - \sqrt{2^{1-S_n} - 1}}{2}. \tag{54}$$

因此，ρ 与乘积态 $|\phi\rangle = \otimes_{n=1}^N |\phi_n\rangle$ 之间的非保真度可以重新表述为

$$\begin{aligned} 1 - \langle\phi|\rho|\phi\rangle &= 1 - \text{Tr}\left[\rho\left(\otimes_{n=1}^N |\phi_n\rangle\langle\phi_n|\right)\right] \\ &= 1 - \Pr\big(x_n = 1,\ \forall n \in [N]\big). \end{aligned} \tag{55}$$

接下来，我们将从式（55）中的概率出发进行推导，以得到非保真度的上界。我们有

$$\begin{aligned} \Pr\big(x_n = 1,\ \forall n \in [N]\big) &= 1 - \Pr\big(\exists\, n \in [N],\ \text{s.t. } x_n = 0\big) \\ &\geq 1 - \sum_{n=1}^N \Pr\big(x_n = 0\big) \\ &= 1 - \sum_{n=1}^N \frac{1 - \sqrt{2^{1-S_n} - 1}}{2}, \end{aligned} \tag{56}$$

其中，式（56）由式（54）和并集界得出。

函数 $f(x)=\sqrt{2^{1-x}-1}$ 对于 $x\in[0,1]$ 是凹函数，因为其二阶导数

$$f''(x)=\frac{(\ln 2)^2\,2^{1-x}\,(2^{1-x}-2)}{4\,(2^{1-x}-1)^{3/2}}\le 0 \tag{57}$$

对于 $x\in[0,1]$。因此，我们有

$$f(a)+f(b)\ge\begin{cases} f(0)+f(a+b) & \text{when } a,b,a+b\in[0,1],\\ f(a+b-1)+f(1) & \text{when } a,b\in[0,1] \text{ and } a+b\in[1,2],\end{cases} \tag{58}$$

根据凹函数的性质。通过选取量子比特对 $(i,\ j)$，使得 $S_i,\ S_j\notin\{0,1\}$，并将新值 $S_i',\ S_j'$ 推向 $[0,1]$的边界，同时保持 $S_i'+S_j'=S_i+S_j$，即可将式（58）应用于式（56）。至多执行上述过程 $N-1$ 次后，我们可以得到如下概率界：

$$\begin{aligned}\Pr\big(x_n=1,\ \forall n\in[N]\big)&\ge 1-\frac{N}{2}+\frac{1}{2}\Big(\lfloor S\rfloor f(1)+f(S-\lfloor S\rfloor)+(N-\lfloor S\rfloor-1)\,f(0)\Big)\\&=\frac{1}{2}\left(\sqrt{2^{1-S+\lfloor S\rfloor}-1}-\lfloor S\rfloor+1\right),\end{aligned} \tag{59}$$

其中，式（59）是通过计算函数 f 在 0 和 1 处的值得出的。因此，我们得到非保真度的上界：

$$\begin{aligned}1-\langle\phi|\rho|\phi\rangle&=1-\Pr\big(x_n=1,\ \forall n\in[N]\big)\\&\le\frac{1}{2}\left(1-\sqrt{2^{1-S+\lfloor S\rfloor}-1}+\lfloor S\rfloor\right)=:f_2(S).\end{aligned} \tag{60}$$

For 在 $S\to 0$ 的情况下，通过计算泰勒展开，式 (60) 趋于 $\log_2{}^2 S+\mathcal{O}(S^3)$。Thus, we 已证明定理 C.1。 □

C.2 定理 C.2 的证明。

证明。 为方便起见，我们将通过 ρ_{target} 在有噪声通道和无噪声通道的情况下分别得到的量子态记为 $\hat{\rho}:=\mathcal{F}^{(L)}(\rho_{\text{target}})$ 和 $\rho:=\mathcal{U}_1^\dagger\circ\cdots\circ\mathcal{U}_L^\dagger(\rho_{\text{target}})$。接下来，我们关注式 (47) 中的下界，而式 (48) 中的上界可用类似方法推导。根据定理 C.1，我们有

$$1-\langle\psi_{\text{product}}|\hat{\rho}|\psi_{\text{product}}\rangle=1-\operatorname{Tr}[\hat{\rho}\sigma_{\text{product}}]\ge f_1(S) \tag{61}$$

对于所有乘积态 $|\psi_{\text{product}},\rangle$ 其中 $\sigma_{\text{product}}:=\ \psi_{\text{product}}\ \rangle\!\langle\psi_{\text{product}}\circ|$ 因此，目标态 ρ_{target} 与由 $\mathcal{E}^{(L)}$ 恢复的量子态之间的非保真度为

$$\begin{aligned}&1-\operatorname{Tr}[\rho_{\text{target}}\mathcal{E}^{(L)}(\sigma_{\text{product}})]\\&=1-\text{迹}[\hat{\rho}\sigma_{\text{product}}]+\text{迹}[\hat{\rho}\sigma_{\text{product}}]-\text{迹}[\rho\sigma_{\text{product}}]+\text{迹}[\rho\sigma_{\text{product}}]-\text{迹}[\rho_{\text{target}}\mathcal{E}^{(L)}(\sigma_{\text{product}})]\\&\ge f_1(S)-\left|\operatorname{Tr}[\hat{\rho}\sigma_{\text{product}}]-\operatorname{Tr}[\rho\sigma_{\text{product}}]\right|-\left|\operatorname{Tr}[\rho\sigma_{\text{product}}]-\operatorname{Tr}[\rho_{\text{target}}\mathcal{E}^{(L)}(\sigma_{\text{product}})]\right|.\end{aligned} \tag{62}$$

式 (62) 中的第二项可以界定为

$$|\operatorname{Tr}[\hat{\rho}\sigma_{\text{product}}]-\operatorname{Tr}[\rho\sigma_{\text{product}}]|\le\|\hat{\rho}-\rho\|_1\|\sigma_{\text{product}}\|_\infty\le(L+1)\|\mathcal{N}-\mathrm{id}\|_\diamond, \tag{63}$$

其中，最后一个不等式来自引理 B.3，且算子范数 $\|\sigma\|_\infty\le 1$。

式 (62) 中的第三项可用类似的方法界定

$$\begin{aligned}&\left|\operatorname{Tr}[\rho\sigma_{\text{product}}]-\operatorname{Tr}[\rho_{\text{target}}\mathcal{E}^{(L)}(\sigma_{\text{product}})]\right|\\&=\left|\operatorname{Tr}[\mathcal{U}_1^\dagger\circ\cdots\circ\mathcal{U}_L^\dagger(\rho_{\text{target}})\sigma_{\text{product}}]-\operatorname{Tr}[\rho_{\text{target}}\mathcal{E}^{(L)}(\sigma_{\text{product}})]\right|\\&=\left|\operatorname{Tr}[\rho_{\text{target}}\mathcal{U}_L\circ\cdots\circ\mathcal{U}_1(\sigma_{\text{product}})]-\operatorname{Tr}[\rho_{\text{target}}\mathcal{E}^{(L)}(\sigma_{\text{product}})]\right|\\&\le\left\|\mathcal{U}_L\circ\cdots\circ\mathcal{U}_1(\sigma_{\text{product}})-\mathcal{E}^{(L)}(\sigma_{\text{product}})\right\|_1\|\rho_{\text{target}}\|_\infty\\&\le(L+1)\|\mathcal{M}-\mathrm{id}\|_\diamond.\end{aligned} \tag{64}$$

结合式 (63) 和 (64) 与式 (62)，我们得到

$$1-\operatorname{Tr}[\rho_{\text{target}}\mathcal{E}^{(L)}(\sigma_{\text{product}})]\ge f_1(S)-(L+1)\left(\|\mathcal{M}-\mathrm{id}\|_\diamond+\|\mathcal{N}-\mathrm{id}\|_\diamond\right),$$

w这就是式 (47)。式 (48) 可以概括如下。因此，我们证明了定理 C.2。 □

C.3 定理 C.3 的证明

证明。为方便起见，我们分别用 ρ_n 和 σ_n 表示 ρ 和 $\sigma = \mathcal{D}_p(\rho)$ 在第 n 个量子比特上的约化密度矩阵。于是，我们有

$$\sigma_n = \mathrm{Tr}_{[N]/\{n\}}\left[\mathcal{D}_p(\rho)\right] = \mathrm{Tr}_{[N]/\{n\}}\left[(1-p)\rho + pI/2^N\right] = (1-p)\rho_n + \frac{p}{2}I. \tag{65}$$

因此，量子态 σ_n 的Rényi熵可表示为

$$\begin{aligned}
\mathcal{S}(\sigma_n) =& -\log_2 \mathrm{Tr}[\sigma_n^2] = -\log_2 \mathrm{Tr}\left[((1-p)\rho_n + \frac{p}{2}I)^2\right] \\
=& -\log_2\left\{(1-p)^2\mathrm{Tr}[\rho_n^2] + p(1-p)\mathrm{Tr}[\rho_n] + \frac{p^2}{4}\mathrm{Tr}[I]\right\} \\
=& -\log_2\left\{(1-p)^2\mathrm{Tr}[\rho_n^2] + p\left(1-\frac{p}{2}\right)\right\}.
\end{aligned} \tag{66}$$

接下来，我们基于式 (66) 推导 $\mathcal{S}(\sigma_n)$ 的不等式。我们有

$$\begin{aligned}
\mathcal{S}(\sigma_n) =& -\log_2\left\{(1-p)^2\mathrm{Tr}[\rho_n^2] + p\left(1-\frac{p}{2}\right)\right\} \\
\leq& -\log_2\left\{(1-p)^2\mathrm{Tr}[\rho_n^2] + p\left(1-\frac{p}{2}\right)\mathrm{Tr}[\rho_n^2]\right\} \\
=& \mathcal{S}(\rho_n) - \log_2\left(1-p+\frac{1}{2}p^2\right) \\
=& \mathcal{S}(\rho_n) + \log_2\frac{2}{1+(1-p)^2},
\end{aligned} \tag{67}$$

其中，不等式由 $\mathrm{Tr}[\rho_n^2] \leq 1$ 得出。

另一方面，

$$\begin{aligned}
\mathcal{S}(\sigma_n) =& -\log_2\left\{(1-p)^2\mathrm{Tr}[\rho_n^2] + p\left(1-\frac{p}{2}\right)\right\} \\
=& -\log_2\mathrm{Tr}[\rho_n^2] - \log_2\left\{(1-p)^2 + \frac{p\left(1-\frac{p}{2}\right)}{\mathrm{Tr}[\rho_n^2]}\right\} \\
=& \mathcal{S}(\rho_n) - \log_2\left\{1 - p\left(1-\frac{p}{2}\right)\left(2-\frac{1}{\mathrm{Tr}[\rho_n^2]}\right)\right\} \\
\geq& \mathcal{S}(\rho_n) + \frac{1}{\ln 2}p\left(1-\frac{p}{2}\right)\left(2-\frac{1}{\mathrm{Tr}[\rho_n^2]}\right) \\
\geq& \mathcal{S}(\rho_n) + \frac{p}{\ln 4}\left(2-\frac{1}{\mathrm{Tr}[\rho_n^2]}\right) \\
=& \mathcal{S}(\rho_n) + \frac{p}{\ln 4}\left(2-2^{\mathcal{S}(\rho_n)}\right) \\
\geq& \mathcal{S}(\rho_n) + \frac{p}{\ln 4}\left(2-(1+\mathcal{S}(\rho_n))\right) \\
=& \left(1-\frac{p}{\ln 4}\right)\mathcal{S}(\rho_n) + \frac{p}{\ln 4},
\end{aligned} \tag{68}$$

其中，第一个不等式由 $\log_2(1-x) \leq -\frac{x}{\ln 2}$ 得出，第二个不等式由 $p \in [0,1]$ 得出，第三个不等式对 $x \in [0,1]$ 由 $2^x \leq 1+x$ 得出。分别对 $n \in [N]$ 求和式 (66) 和 (68)，即可得到式 (49)。

□

D AQER 算法的更多实现细节

本节将详细讨论 AQER 的技术细节，包括纠缠度量 $\mathcal{S}$ 的计算与优化、乘积态近似的显式构造，以及 AQER 中变分量子电路优化的可训练性。我们还将讨论为何 $\mathcal{S}$ 可作为全局近似误差的高效代理指标，以及它如何助力在大型量子系统上实现可扩展训练。最后，我们将说明如何将 AQER 用于经典数据。

D.1 步骤 I：纠缠度量的计算与优化

在 AQER 的步骤 I 中，纠缠度量 $\mathcal{S}$ 的计算与优化只需有限的量子资源。由于 $\mathcal{S}$ 定义为单量子比特雷尼熵之和，仅由局部项构成，因此可通过局部测量高效评估。具体而言，任何常数规模子系统的 Rényi 纠缠熵都可以通过对其约化密度矩阵进行量子态层析来估计；要获得精度为 ϵ 的估计，只需进行 $\mathcal{O}(1/\epsilon^2)$ 次测量。此外，在步骤 I 的每次迭代中，对于每个候选双量子比特门块，只有涉及的两个量子比特的单量子比特雷尼熵会受到影响。因此，添加一个新门块所导致的 $\mathcal{S}$ 变化，等于这两个单量子比特熵之和的变化。由此，在步骤 I 的每次迭代中，都会选取能使相关单量子比特熵降幅最大的门块，作为新的结构。利用 Nelder-Mead 算法等经典方法优化这些门块中的参数。通过这一过程，仅凭局部信息即可系统地降低纠缠。这种降低使其能够高效扩展至大型量子系统。

D.2 步骤 II：参数计算

在 AQER 的步骤 II 中，单量子比特旋转的参数可通过单量子比特态层析成像显式且高效地计算。根据定理 3.1，经过步骤 I 后的量子态 $|\boldsymbol{v}_T\rangle$ 纠缠度降低，可由乘积态良好近似；该乘积态由 $|0\rangle^{\otimes N}$通过单量子比特旋转制备。在步骤 II 中，单量子比特旋转$R_Z(\boldsymbol{\beta}_i)R_Y(\boldsymbol{\gamma}_i)$的参数（$\boldsymbol{\beta}_i$，$\boldsymbol{\gamma}_i$）可根据层析成像得到的单量子比特约化密度矩阵的矩阵元显式计算。具体而言，根据引理 B.2，每个 β_i 和 γ_i 都可直接由相应约化密度矩阵的矩阵元 ρ_{00}、ρ_{11}和 ρ_{10} 计算得到。如步骤 I 的讨论所述，对每个量子比特进行层析成像需要 $\mathcal{O}(1/\epsilon^2)$ 次测量，才能获得精度为 ϵ的估计，且所需测量次数与总系统规模 N无关。得到这些参数后，即可显式构造 $|\boldsymbol{v}_T\rangle$ 的乘积态近似，无需任何迭代优化，从而为步骤 III 提供高保真初始化。

D.3 步骤 III：参数优化的可训练性

通过在步骤 I 中减少纠缠，并在步骤 II 中显式构造单量子比特门，AQER 减轻了贫瘠高原对步骤 III 后续参数优化的影响。通常，在变分量子算法的训练中，当损失函数初始化在其平均值附近且梯度呈指数级微小时，就会出现贫瘠高原现象。这使参数优化极具挑战性。对于定义为相对于一个 N 量子比特目标态的非保真度的损失函数，随机初始化的参数化电路的平均损失随 $1-\mathcal{O}(1/2^N)$ 缩放。相比之下，在 AQER 的步骤 III 中，在完整优化之前会抑制纠缠度量 $\mathcal{S}$。根据定理 3.1，较小的 $\mathcal{S}$ 值对应非保真度损失较小的量子态。因此，变分电路的初始点实际上位于贫瘠高原区域之外。这确保后续优化从条件良好的区域开始，此处的梯度具有意义，从而减轻贫瘠高原的影响并提升可训练性。这一机制不同于以往避免或减轻贫瘠高原的策略；后者主要通过重塑优化景观或放大梯度信号来实现（例如，通过架构约束 [77, 78], 巧妙的参数初始化策略 [79–82], 或局部代价函数 [83]），但不一定能保证以低损失状态作为起点。此外，AQER 主要缓解参数空间贫瘠高原，但无法解决数据诱导的贫瘠高原；后者仍可能导致严重的可训练性和泛化问题 [84, 85]。

D.4 AQER：作为近似误差代理的纠缠度量

使用纠缠度量 $\mathcal{S}$ 作为全局近似误差的代理指标有多项优势。直接优化不保真度损失需要进行全局测量，这在计算上开销很大，而且容易受到统计方差的影响，在大型系统中尤为如此。相比之下，$\mathcal{S}$ 只需通过局部测量即可高效评估并降低。通过先降低 $\mathcal{S}$，AQER 不仅能减小近似误差，还能使量子电路处于有利状态，从而进一步进行全局优化。这一策略提升了 AQER 在大规模量子比特系统上的评估效率，以及整体可训练性和可扩展性。

D.5 使用经典数据的 AQER

对于经典数据，AQER 通过经典模拟实现，同时保留相同的三步结构。具体而言，对应于经典数据的量子态以经典向量或张量网络的形式表示。

经典模拟： 在步骤 I 中，通过经典地计算约化密度矩阵及其 Rényi 熵，得到纠缠度量 $\mathcal{S}$。门优化通过对候选模块进行经典模拟来完成，并使用基于梯度的方法，根据精确的梯度信息优化参数。在步骤 II 中，根据引理 B.2，直接从计算得到的约化密度矩阵中提取单量子比特旋转参数（β_i，γ_i）。步骤 III 利用经典模拟进行精确的损失评估和梯度计算，不受统计噪声影响。

E 实验设置

本节首先介绍经典与量子数据集的构建和预处理。随后，我们将介绍实验中所用参考量子数据加载器的超参数配置。

数值模拟设置。 本文中的所有实验均使用经典模拟器完成。对于涉及少于 20 个量子比特的计算任务，我们使用了基于 Python 的软件包 PennyLane [47],。对于含有 20 个或更多量子比特的较大型量子系统，我们使用了基于 Julia 的软件包 PastaQ [86],，该软件包采用张量网络技术，以支持对大规模量子系统进行高效模拟。

E.1 经典数据集和量子数据集的构建

经典数据。 对于 MNIST，每张图像都补零至 32×32，展平为一个 1024 维向量 $\boldsymbol{v}$，并进行归一化，使其 ℓ_2 范数为 1。对应的目标态由振幅编码 $|\boldsymbol{v}\rangle=\sum_{j=1}^{2^N}\boldsymbol{v}_j|j\rangle$ 定义，使用 $N=10$ 个量子比特。对于 CIFAR-10，每张 32×32 RGB 图像都通过拼接三个颜色信道展平为一个 3072 维向量，补零至 4096 维，然后进行归一化。随后将向量的后半部分视为虚部，并与前半部分组合，构成紧凑编码 [60]:$|\boldsymbol{v}\rangle=\sum_{j=1}^{2^N}(\boldsymbol{v}_j+i\,\boldsymbol{v}_{j+2048})|j\rangle$，使用 $N=11$ 个量子比特；该编码作为目标，用于评估 AQER 在不同编码方案下的表现。对于 SST-2，每个句子首先由预训练的 Sentence-BERT 模型表示为一个 1024 维嵌入向量。随后对嵌入向量进行归一化，并通过使用 $N=10$ 个量子比特的振幅编码定义对应的目标态。

Quantum 数据。我们考虑两类量子数据集：由随机量子电路（RQC）生成的合成态，以及一维横场伊辛模型（1D TFIM）的基态。这些数据集是典型示例，代表了由量子电路演化或量子多体系统产生的、具有物理意义的量子态。具体而言，第一个数据集由在量子态 $|0\rangle$ 上通过不同 RQC 生成的 $M=50$ 个量子态组成。每个 RQC 均从集合 $RandomShuffle\left(\{CZ_{p_k,q_k}\}_{k=1}^{W}\cup\{R_{\sigma_k,p_k}(\theta_k)\}_{k=W+1}^{4W}\right)$ 中抽样，其中 $1\le p_k\neq q_k\le N$ 从 $[N]$ 中随机抽取。每个 $R_{\sigma,\ p_k}$ 都是作用于量子比特 p_k 上的单量子比特旋转，其参数分别服从 $\sigma_k\sim$ 均匀分布和 $\theta_k\sim$ 均匀分布。当双量子比特门的数量为 $W\ge\mathcal{O}(N)$ 时，这些量子电路会生成高度纠缠的量子态，并作为具有代表性的合成量子数据。实际使用中，我们采用 $N=10$ 和 $W=40$。第二个数据集由一维横场伊辛模型的基态组成，其哈密顿量定义为 $H_{\mathrm{TFIM}}=-\sum_{i=1}^{N-1}JZ_iZ_{i+1}-\sum_{i=1}^{N}gX_i$。我们考虑相变点附近的系数，即 $g=1$ 和 $J\in\{0.8,0.9,1,1.1,1.2\}$，以构建规模为 $M=5$ 的数据集，其中 $N\in\{10,20,30,40,50\}$。

E.2 参考量子数据加载器的超参数设置

这里给出了实验中使用的参考量子数据加载器的超参数设置。这些方法的描述见 A.3 节。

表 2：不同编码方法所用 CNOT/CZ 门的可行数量。术语 N 表示量子比特的数量，术语 $k \in \mathbb{N}$可以是任意正整数。

方法	双量子比特门数量 G
AQCE（$\mathbb{C}$）	$G = 15k$
AQCE（$\mathbb{R}$）	$G = 10k$
MPS ($\mathbb{C}$)	$G = 3(N-1)k$
MPS ($\mathbb{R}$)	$G = 2(N-1)k$
HEC	$G = \lceil \frac{1}{2} Nk \rceil$
AQER（我们的）	$G = k$

自动量子电路编码 (AQCE)。AQCE 方法会以先前向后向的方式依次更新双量子比特酉算子数次，然后再向量子电路中添加新的量子门。在实验中，每个加门步骤都会添加 5 个新的酉算子，随后进行 200 轮前向和后向的量子门更新。

硬件高效电路（HEC）。我们采用分层的硬件高效（HE）架构，其中每一层由一个 R_Y 旋转层、一个 R_Z 旋转层以及一个作用于相邻量子比特对的 CNOT 层组成。具体而言，对于第 i层，当 i %$2 = 0$时，CNOT 门施加于量子比特对（$2n$，（$2n+1$）%N） $0 \le 2n \le N-1$；当 i%$2 = 1$ 时，施加于（$2n+1$，（$2n+2$）%N） $0 \le 2n \le N-2$。HEC 中的参数从$[0, 2\pi]$上的均匀分布随机初始化。训练使用 Adam 优化器进行 2000 次迭代，与 AQER 中的设置一致。

我们指出，参考文献中的 AQL 方法对硬件可用双量子比特门的数量施加了不同的限制，例如 CNOT 和 CZ 门。我们在表 2 中总结了这些方法下可行的双量子比特门数量。具体而言，在实数情形和一般复数情形下，任意双量子比特酉算子的分解分别需要两个和三个 CNOT/CZ门 [87]。

E.3 带交叉验证的量子核方法

我们采用量子核方法（QKM） [88, 89] 进行二分类。QKM 使用量子特征映射将经典输入向量 $\boldsymbol{x}^{(i)}$ 嵌入量子态 $|\psi(\boldsymbol{x}^{(i)})\rangle$，并据此构建核矩阵 $K_{ij} = |\langle\psi(\boldsymbol{x}^{(i)})|\psi(\boldsymbol{x}^{(j)})\rangle|^2$，以量化数据点之间的两两相似度。设 $y_i \in \{-1, 1\}$ 为数据 $\boldsymbol{x}^{(i)}$的标签。那么，支持向量机（SVM）的训练对应于求解对偶优化问题

$$\max_{\boldsymbol{\alpha}} \left[\sum_i \boldsymbol{\alpha}_i - \frac{1}{2} \sum_{ij} \boldsymbol{\alpha}_i \boldsymbol{\alpha}_j y_i y_j K_{ij} \right], \text{ s.t. } 0 \le \boldsymbol{\alpha}_i \le C, \sum_i \boldsymbol{\alpha}_i y_i = 0,$$

其中，C 是惩罚参数，默认设为 1。QKM 的参数 $\boldsymbol{\alpha}$ 训练完成后，使用决策函数计算新样本 $\boldsymbol{x}$ 的预测标签

$$\hat{y} = \text{sign} \left[\sum_{i=1}^{N} \boldsymbol{\alpha}_i y_i K(\boldsymbol{x}^{(i)}, \boldsymbol{x}) \right],$$

其中，$K(\boldsymbol{x}^{(i)}, \boldsymbol{x}) = |\langle\psi(\boldsymbol{x}^{(i)})|\psi(\boldsymbol{x})\rangle|^2$ 是训练样本 $\boldsymbol{x}^{(i)}$ 与新样本 $\boldsymbol{x}$之间的核。

在实验中，我们将 AQER 的近似编码与精确振幅编码进行比较。我们在大小为 200 的数据集上采用分层 k 折交叉验证（使用 $k = 5$）来评估分类性能。对于每一折，SVM 使用预先计算的核矩阵在训练子集上进行训练，并使用 K 的相应子矩阵对留出的验证子集进行预测。报告的误差为所有折的平均值。

F 其他实验结果

在本节中，我们展示 AQER 在经典数据集和量子数据集上的性能的更多数值结果。

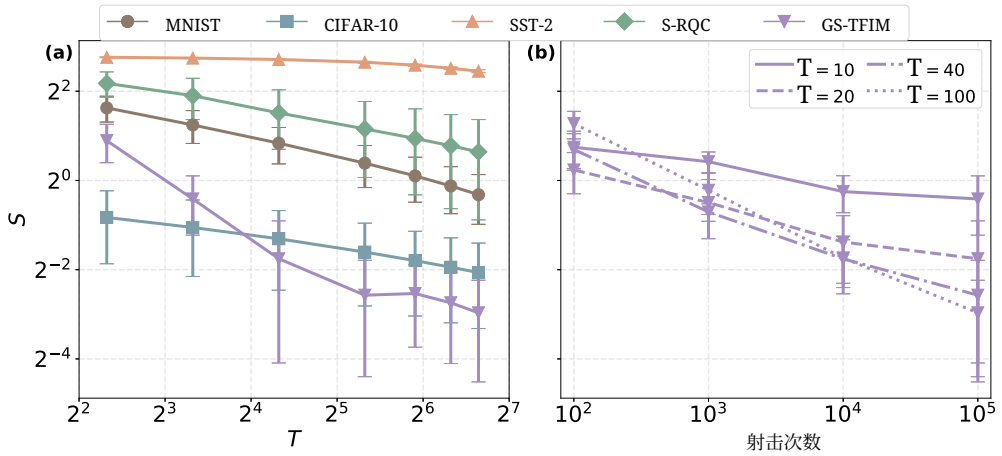

图 6：AQER 在 MNIST、CIFAR-10、SST-2、S-RQC 和 GS-TFIM 数据集上的纠缠抑制性能，以不同颜色和标记区分。(a) AQER 的步骤 III 之后，所有数据集上的纠缠度量值 S 与不同的 $T \in \{5, 10, 20, 40, 60, 80, 100\}$ 之间的关系。(b) GS-TFIM 数据集上的纠缠度量值 S 与不同测量次数之间的关系，其中使用了不同的 $T \in \{10, 20, 40, 100\}$。

AQER 的步骤 I 能够随着迭代次数增加有效降低纠缠。我们首先考察 AQER 如何逐步降低 MNIST、CIFAR-10、SST-2、S-RQC 和 GS-TFIM 数据集的纠缠度量值 S。图 6(a) 展示了所有数据集在不同 $T \in \{5, 10, 20, 40, 60, 80, 100\}$ 下经过 AQER 步骤 III 后的 S。对于大多数数据集，增加 T 会持续降低 S，这表明 AQER 通过逐步扩大量子电路，有效降低了目标态的纠缠。由于测量次数有限，GS-TFIM 数据集出现了轻微波动，这是统计噪声所致，但整体下降趋势依然明显。这些结果证实，随着 T 增大，AQER 能够逐步减轻纠缠。

射击次数对纠缠约化的影响。接下来，我们研究不同测量次数对纠缠约化的影响。图 6(b) 展示了 GS-TFIM 数据集的纠缠度量 S，其中 $T \in \{10, 20, 40, 100\}$。增加射击次数通常会降低 S，这与正文中观察到的非保真度变化趋势一致。对于较大的 T，这种降低更为显著。例如，将测量次数从 10^2 增加到 10^5，对于 $T = 10$，S 大约降低至原来的二分之一；而对于 $T = 100$，降幅则约为 16 倍。这些结果表明，充足的测量次数能够显著增强 AQER 抑制纠缠的能力。

随着迭代步数增加，AQER 的步骤 III 能有效降低非保真度。随后，我们研究步骤 III 的不同迭代次数 $T_3 \in \{40, 80, 160, 320, 640, 1280, 2000\}$ 对 AQER 性能的影响。图 7 展示了 MNIST、CIFAR-10、SST-2、S-RQC 和 GS-TFIM 数据集的非保真度，各子图分别对应 $T = 10$，20，40，和 100。对于所有情况，增加 T_3 都会持续降低非保真度，表明更长时间的步骤 III 优化能够有效提升近似质量。对于较大的 T，这种降低更为显著，说明 AQER 能够受益于将更大的初始量子电路与步骤 III 中更充分的优化相结合。

使用 AQER 进行二分类。我们进一步考察了使用 AQER加载态时 QKM 的分类性能。具体而言，我们使用具有不同 T 值的 AQER加载态，对 MNIST 中数字 0 和 1 进行二分类。结果如图 8 所示，其中将 $T \in \{5, 10, 20, 40, 60, 80, 100\}$ 的错误率与精确加载进行比较。我们观察到，随着 T 增加，分类误差稳步下降；从 $T \geq 20$ 开始，性能已达到与精确加载相当的错误率。这表明，AQER 只需相对较小的电路规模，就能在分类任务中实现近乎精确的下游性能。

AQER 与 AQL 基线方法的资源—精度权衡。 我们进一步比较了 S-RQC 数据集上不同 AQL 方法在准确度与资源之间的权衡。图 9 展示了 AQER、HEC、AQCE 和 MPS 的非保真度随四项资源指标变化的情况：量子电路深度、可训练参数数量、测量次数总数，以及经典训练时间（单位为

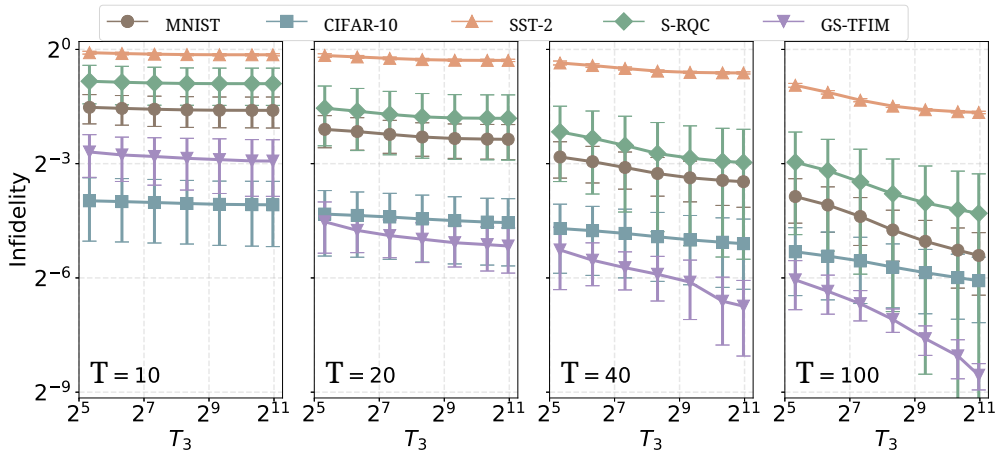

图 7：AQER 在步骤 III 中采用不同优化迭代次数（即 $T_3 \in \{40, 80, 160, 320, 640, 1280, 2000\}$）时的非保真度。每个子图分别对应 $T = 10, 20, 40,$ 和 100。

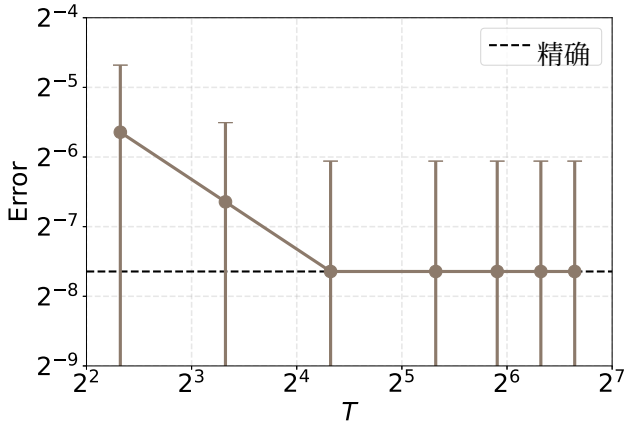

图 8：使用带有$T \in \{5, 10, 20, 40, 60, 80, 100\}$的 AQER加载态对 MNIST 数字 0 与 1 进行二分类的误差，与精确加载的结果（黑色虚线）进行比较。

秒）。具体而言，量子电路深度定义为编译量子电路中顺序门层的数量，其中不同量子比特上的对易门已完全并行化。我们仅统计至少包含一个双量子比特门的层，因为在实际应用中，双量子比特门通常比单量子比特门慢一个数量级。参数数量是量子电路中所有独立连续参数（例如旋转角）的数量。测量次数是整个 AQL 流程中使用的量子测量次数。训练时间是在配备 Apple M2 芯片和 8 GB 内存的笔记本电脑上测得的。总体而言，在量子电路深度和参数数量相当或更低的情况下，AQER 实现了比其他方法更低的非保真度，表明其对量子电路表达能力的利用更加高效。测量次数和训练时间方面的表现则更为复杂。由于算法较为简单，MPS 基线方法所需的测量次数更少、训练时间更短；但将其实现为量子电路时，量子电路深度、双量子比特门数量和参数数量都显著更高，且非保真度也明显更高。与基于电路的基线方法（AQCE 和 HEC）相比，AQER 始终能以相当或更低的测量开销和训练成本实现更低的非保真度。因此，在基于电路的 AQL 方法中，AQER 实现了较为理想的资源—精度权衡。

大量子比特数下 AQER 的梯度缩放。我们分析了 TFIM 数据集上 AQER 梯度范数在大量子比特数下的缩放行为，以证明步骤 III 的优化

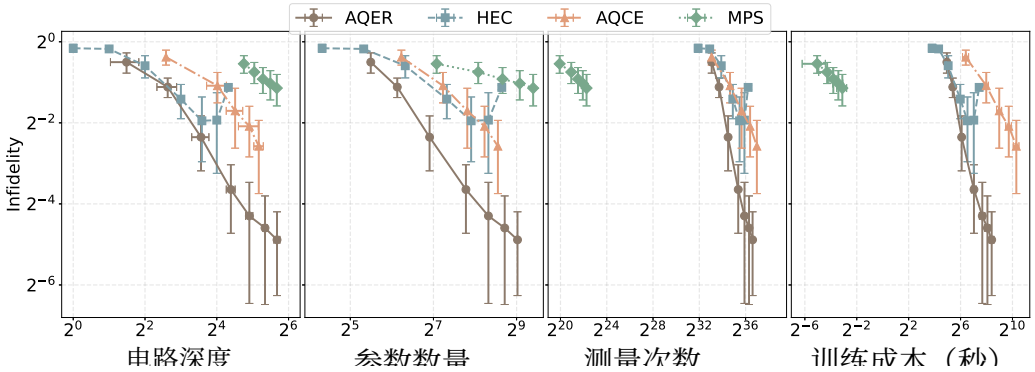

图 9：S-RQC 数据集上不同 AQL 方法的非保真度随各项指标的变化。我们考察了四项指标：电路深度、参数数量、测量次数以及以秒为单位的训练成本。每个数据点均为 5 个样本的平均值。

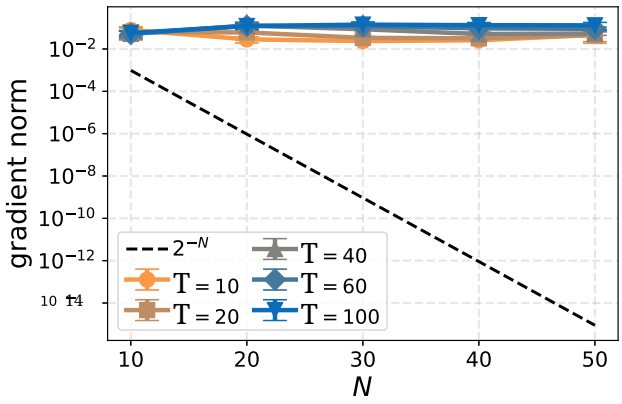

图 10：TFIM 数据集中 AQER 梯度范数随量子比特数量的变化。该图展示了第三阶段训练初始阶段的梯度范数随系统规模$N \in \{10, 20, 30, 40, 50\}$的变化，并对不同的电路规模 $T \in \{10, 20, 40, 60, 100\}$进行了比较。黑色虚线表示与 2^{-N} 成正比的指数衰减参考曲线。

在 AQER 中不存在贫瘠高原问题。对于系统规模 $N \in \{10, 20, 30, 40, 50\}$ 和电路规模 $T \in \{10, 20, 40, 60, 100\}$，我们在优化开始时计算了关于所有可训练参数的梯度范数。结果如图 10 所示。在所有系统规模和电路规模下，初始梯度范数保持在 10^{-2} 量级，且不随 N 呈现任何指数衰减，这与图中指数消失的参考曲线 2^{-N}形成鲜明对比。这些观测结果提供了数值证据，表明当系统规模从 $N = 10$ 增加到 $N = 50$时，AQER 的可训练性保持稳定。

AQER在噪声信道下的稳健性。 我们在 $N = 10$横场伊辛模型任务上进行含噪模拟，以评估AQER在现实的含噪中等规模量子（NISQ）时代中的性能。具体而言，在每一层单量子比特门和双量子比特门层之后，我们分别施加错误率为 $p_1 \in \{10^{-3}, 3 \times 10^{-4}, 10^{-4}\}$ 和 $p_2 = 10p_1$的去极化信道，这对应于当前量子设备所报告的代表性错误率范围。对于每种噪声设置和双量子比特门数量 $T \in \{5, 7, 10, 12, 15, 20, 24, 28, 32, 36, 40, 50, 60, 80, 100\}$，我们运行AQER并记录最终非保真度。结果如图 11所示。在全部三种噪声水平下，我们观察到依赖于噪声的最优电路规模 T^*：当 T 从很小的值开始增大时，非保真度下降，这表明在噪声累积占据主导之前，纠缠引导的电路增长提升了近似质量。超过 T^*之后，进一步增大 T 会导致非保真度逐渐上升，因为噪声的影响超过了表达能力提升。此外，当物理错误率从$(p_1, p_2) = (10^{-3}, 10^{-2})$降至$(10^{-4}, 10^{-3})$时，可达到的最佳非保真度从高于 2^{-3} 改善至

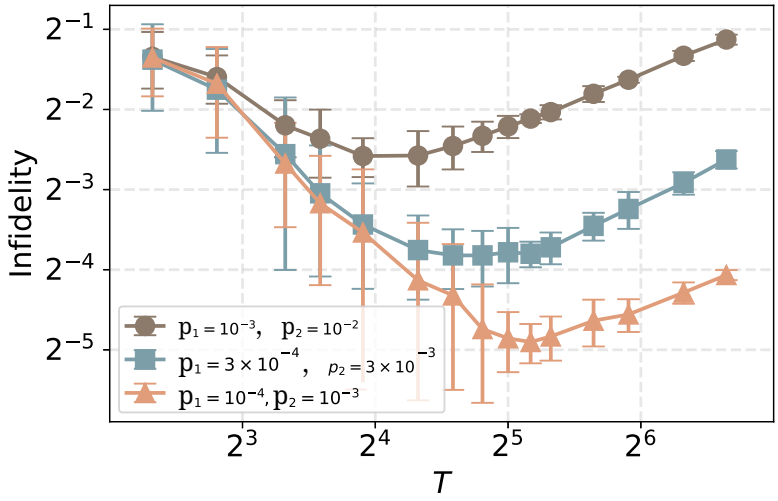

图 11：全局退极化噪声下 AQER 在 $N = 10$ 横场伊辛模型任务中的非保真度。该图展示了三种噪声水平下最终非保真度随双量子比特门数量 T 的变化，其中单量子比特门和双量子比特门错误率分别为 (p_1 p_2 $\in$ { 10^{-3} 10^{-2} , 3×10^{-4} $3 \times$,) (,) (,),(,)}。 10^{-3} 10^{-4} 10^{-3}

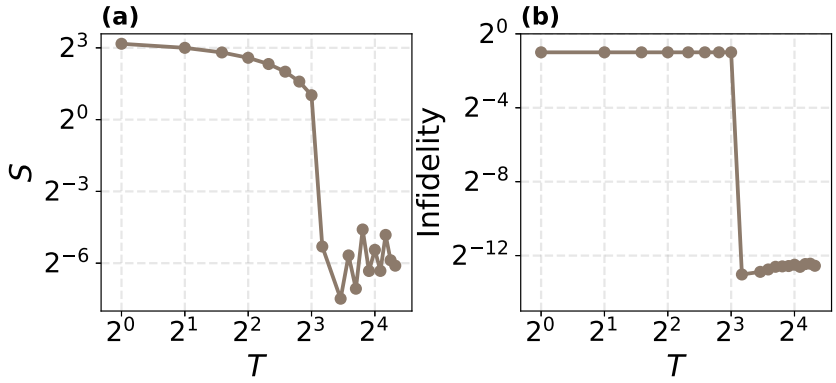

图 12：AQER 在 $N = 10$ GHZ态上的性能。（a）纠缠度量 $\mathcal{S}$ 随双量子比特门数量 T的变化。（b）非保真度随 T的变化。

约 2^{-5}，且最优 T^* 增大。总体而言，这些结果表明AQER在NISQ级噪声区间中仍然有效且稳健。

用于制备 GHZ态的 AQER。我们以 $N = 10$ GHZ态为案例，评估 AQER 在高度纠缠但结构简单的量子态上的性能。GHZ态具有高度纠缠性，这是因为每个单量子比特约化密度矩阵都是最大混合态，且初始纠缠度量满足 $\mathcal{S} = N$。尽管如此，AQER 仍能非常高效地降低这种纠缠。如图 12 所示，纠缠度量 $\mathcal{S}$ 会随双量子比特门数量 T 快速衰减，并在 $T = 9$时就降至 2^{-3} 以下。仅使用 9 个双量子比特门，相应的非保真度便降至 2^{-12} 以下。该示例表明，一个量子态可以具有高度纠缠性，同时仍保持结构简单；在这种情况下，尽管 $\mathcal{S}$的初始值很大，AQER 仍能高效找到可实现极低非保真度的低深度电路。

用于制备二维随机电路态的 AQER。我们考虑在矩形晶格上制备二维随机电路态，以评估 AQER 在具有实际物理意义的目标态上的表现。目标态由一个深度为4的量子电路生成，该电路由交替排列的单量子比特层和 CZ 层组成。在每个单量子比特层中，每个量子比特都会受到从 $\{R_X, R_Y, R_Z\}$ 中随机选取的旋转作用，旋转角度均匀采样。在每个 CZ 层中，CZ 门沿水平方向或垂直方向作用于最近邻

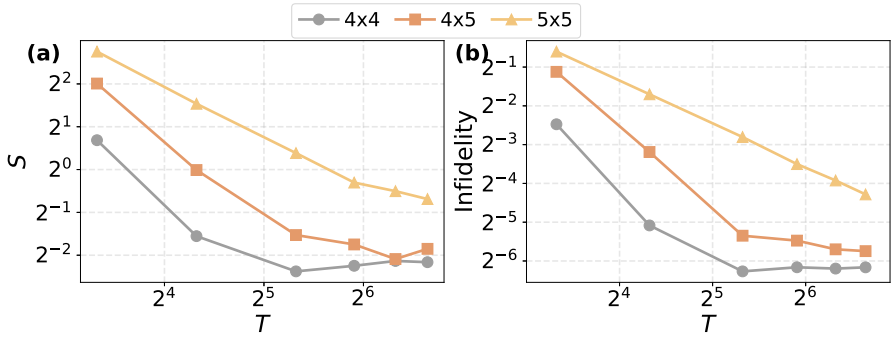

图 13：AQER 在二维随机电路态上的性能。目标态由作用于 4×4、4×5 和 5×5 晶格的深度为 4 的最近邻线路生成。(a) 纠缠度量 $\mathcal{S}$ 与双量子比特门数量 $T\in\{10,20,40,60,80,100\}$ 的关系。(b) 同一组二维随机电路态的非保真度与 T 的关系。

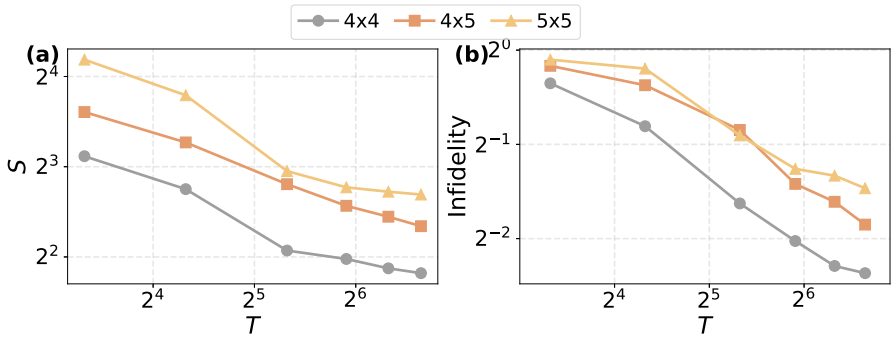

图 14：AQER 在临界点二维 XXZ 基态上的性能。(a) 对于 4×4、4×5 和 5×5 晶格，纠缠度量 $\mathcal{S}$ 随双量子比特门数量 $T\in\{10,20,40,60,80,100\}$ 的变化关系。(b) 非保真度随 T 的变化关系。

，具体模式依据二维网格的四种不同铺排方式之一确定（两种水平模式和两种垂直模式）。这四种 CZ 模式在四层中循环使用，因此整个量子电路中的所有最近邻耦合都会被激活。我们考虑 4×4、4×5 和 5×5 晶格，并以双量子比特门数量 $T\in\{10,20,40,60,80,100\}$ 应用 AQER。所得的纠缠度量和非保真度如图 13 所示。对于所有晶格规模，随着 T 从 10 增至 100，纠缠度量 $\mathcal{S}$ 至少降低 8 倍，表明纠缠降低策略对二维随机电路态仍然有效。相应的非保真度也会随着 T 增大而降低。例如，在 4×4 晶格上，非保真度从 $T=10$ 时高于 2^{-3} 降至 $T=100$ 时约为 2^{-6}；在 5×5 晶格上，则从高于 2^{-1} 降至低于 2^{-4}。这些结果表明，随着电路规模 T 增大，AQER 能够显著且稳定地降低浅层二维随机电路态的纠缠度量和近似误差。

用于制备临界点处二维 XXZ 基态的 AQER。我们考虑在晶格上制备自旋-$\frac{1}{2}$ XXZ 模型的基态，以评估 AQER 在二维多体量子态上的表现。哈密顿量为

$$H_{\mathrm{XXZ}}=\sum_{\langle i,j\rangle}\Big[J_{xy}\big(\sigma_i^x\sigma_j^x+\sigma_i^y\sigma_j^y\big)+J_z\,\sigma_i^z\sigma_j^z\Big], \tag{69}$$

其中，$\langle i,\ j\rangle$ 遍历二维网格上的最近邻对，$\sigma^{x,\ y,\ z}$ 表示泡利矩阵。我们聚焦于临界点 $J_{xy}=J_z$，并考虑 4×4、4×5和 5×5 晶格。对于每个晶格，我们采用 AQER，双量子比特门数量为 $T\in\{10,20,40,60,80,100\}$。纠缠度量和保真度损失如图 14 所示。对于所有晶格规模，纠缠度量 $\mathcal{S}$ 都会随着 T 的增加而稳步下降；当 T 从 10 增加到 100 时，通常会降低约 2–3 倍（例如，在 4×4 晶格上从高于 2^3 降至低于 2^2，以及在 5×5 晶格上从高于 2^4 降至低于 2^3）。

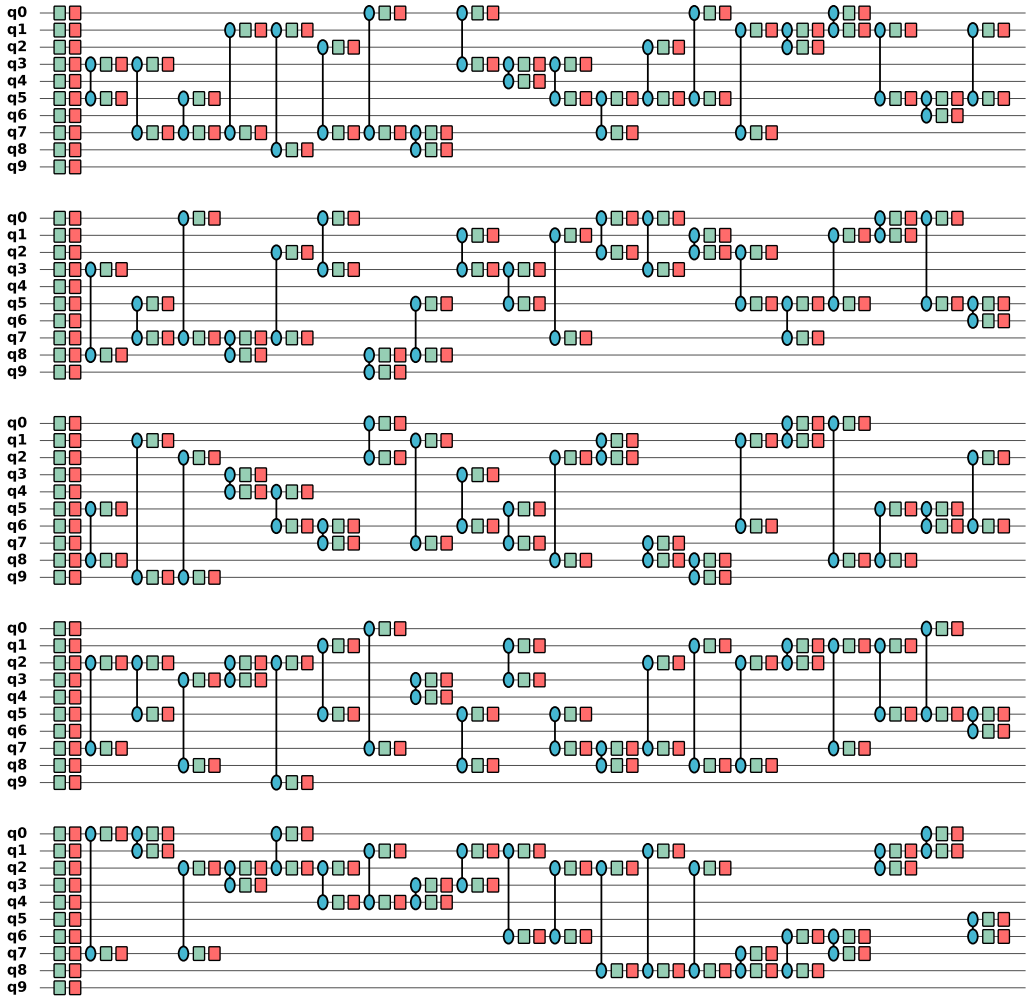

图 15：使用 $T = 20$ 为 MNIST 数据集生成的 AQER 加载电路。绿色和红色方块分别表示 R_Y 和 R_Z 量子门。由线连接的蓝色圆圈表示 R_{ZZ} 量子门。

相应的保真度损失也会随着 T 单调下降，降幅超过 2 倍。这些结果表明，AQER 能够可靠地压缩纠缠，并降低二维量子多体系统的近似误差；随着电路规模 T 的增加，$\mathcal{S}$ 和非保真度都会系统性地降低。

量子电路可视化。图 15–19 分别展示了使用 $T = 10$时，AQER 算法为 MNIST、CIFAR-10、SST-2、S-RQC 和 GS-TFIM 数据集生成的量子电路。每幅图展示了五个示例的量子电路布局，包括单量子比特旋转门 R_Y，R_Z 以及双量子比特 R_{ZZ} 门。

G AQER 的计算复杂度

下面我们给出 AQER 的计算复杂度。我们将分析分为经典数据设定和量子数据设定：在经典数据设定中，目标态在经典计算机上以态矢量的形式显式表示；在量子数据设定中，只能访问在量子设备上制备的目标量子态副本。下文中，我们分别用 N 表示量子比特的数量，用 $d = 2^N$ 表示经典数据向量的维度。AQER 步骤 I 的迭代次数记为 T，AQER 步骤 III 的迭代次数记为 T_3。

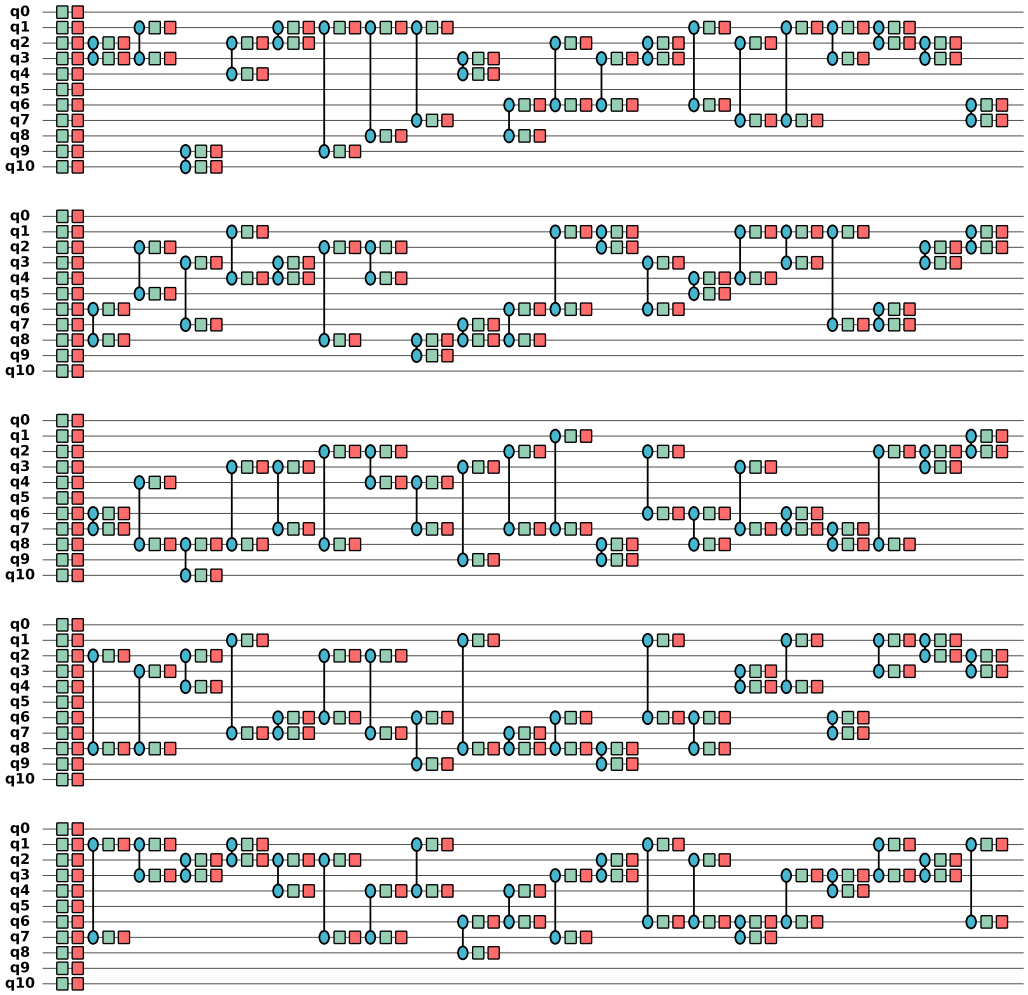

图16：由AQER使用 $T = 20$ 为CIFAR-10数据集生成的加载电路。绿色和红色方块分别表示 R_Y 和 R_Z 量子门。由线条连接的蓝色圆圈表示 R_{ZZ} 量子门。

G.1 经典数据

我们首先分析目标态在经典计算机上以显式态矢量 $\boldsymbol{v} \in \mathbb{C}^d$ 给出时 AQER 的成本。在这一设定下，AQER 所需的所有量（约化密度矩阵、熵、梯度）都可以根据 $\boldsymbol{v}$ 精确计算。

施加量子门的代价。AQER中使用的每个量子门要么是对角的（R_Z 和 R_{ZZ}），要么是二稀疏的（对于R_Y）。因此，将单量子比特门或双量子比特门 U 作用于态矢量 $\boldsymbol{v} \in \mathbb{C}^d$，可以实现为与对角矩阵或二稀疏矩阵的矩阵-向量乘法，因而需要 $\mathcal{O}(d)$ 次算术运算。随后，包含 $\mathcal{O}(T)$ 的量子电路的一次前向传播需要 $\mathcal{O}(dT)$。

计算约化密度矩阵与熵的代价。令计算基下的态矢量写为

$$\boldsymbol{v} = \sum_{i_1=0}^{1} \cdots \sum_{i_N=0}^{1} v_{i_1,\ldots,i_N} \boldsymbol{e}_{i_1,\ldots,i_N}. \tag{70}$$

那么，量子比特 (1，2) 的约化密度矩阵（RDM）为

$$\rho_{j_1,j_2;k_1,k_2} = \sum_{i_3,\ldots,i_N} v_{j_1,j_2,i_3,\ldots,i_N} v^*_{k_1,k_2,i_3,\ldots,i_N}, \qquad j_1, j_2, k_1, k_2 \in \{0,1\}. \tag{71}$$

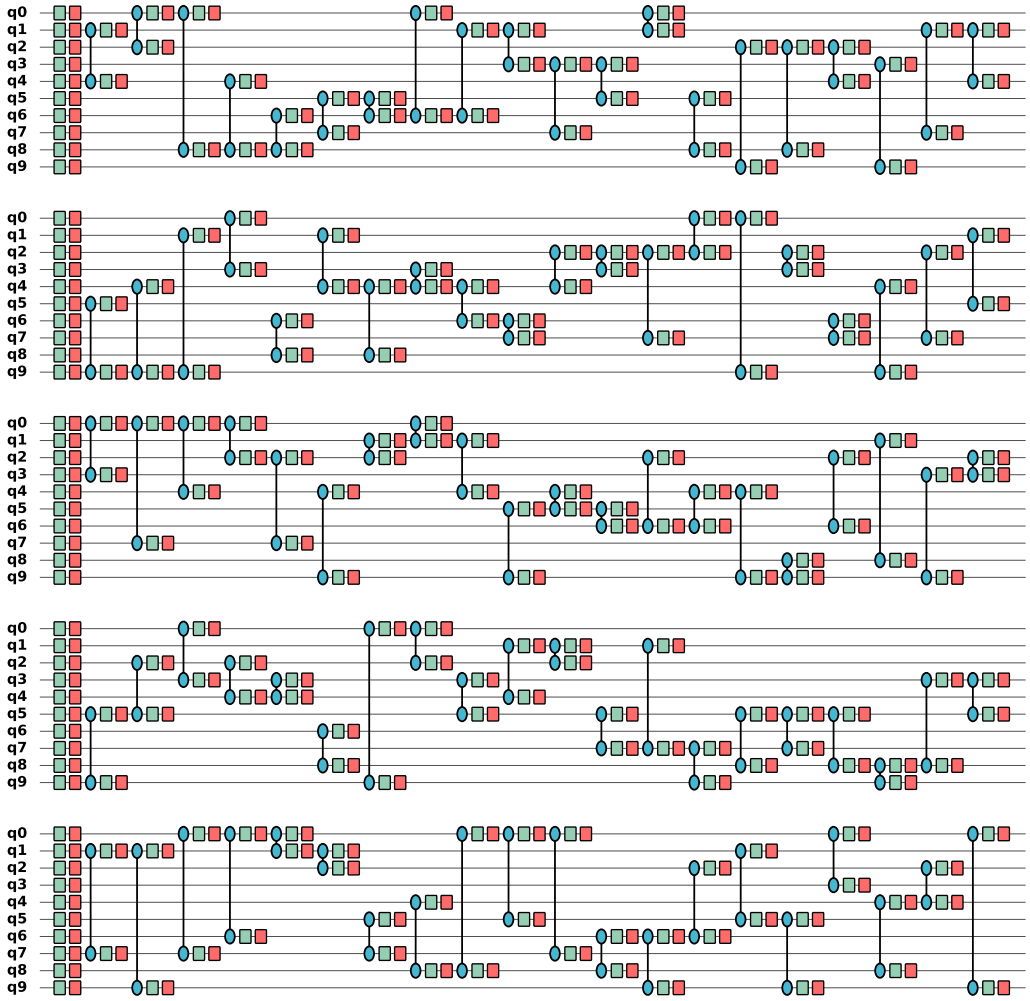

图 17：使用 $T = 20$ 为 SST-2 数据集生成的加载量子电路。绿色和红色方块分别表示 R_Y 和 R_Z 量子门。由线连接的蓝色圆圈表示 R_{ZZ} 量子门。

计算该约化密度矩阵需要对 (i_3、.....、i_N) 的所有 2^{N-2} 种赋值求和，因此需要 $\mathcal{O}(2^N) = \mathcal{O}(d)$ 次运算。对任意固定数量的量子比特以及任意量子比特索引选择，其约化密度矩阵的计算复杂度也相同。得到约化密度矩阵后，通过对常数阶矩阵进行对角化，即可计算 AQER 中使用的单量子比特熵和纠缠度量，额外开销可以忽略。因此，计算固定数量量子比特的约化密度矩阵及相应熵的成本为 $\mathcal{O}(d)$。

步骤 I：贪心纠缠约简。在步骤 I 中，AQER 通过迭代添加可降低纠缠度量的单量子比特门和双量子比特门，贪心地构建包含 $5T$ 量子门的量子电路。每次迭代中，算法最多评估 $\mathcal{O}\left(N^2\right)$ 个候选量子比特对，而评估一个候选项需要 $\mathcal{O}\left(d\right)$ 时间。因此，步骤 I 的总成本为

$$\mathcal{O}(d\,T\,N^2). \tag{72}$$

步骤 II：乘积态近似。在步骤 II 中，AQER 计算所有N 量子比特的单量子比特约化密度矩阵，并利用这些矩阵初始化单量子比特旋转的参数。计算N 单量子比特约化密度矩阵的代价为 $\mathcal{O}(dN)$，随后计算 $2N$ 门参数仅需 $\mathcal{O}(N)$ 次操作。因此，步骤 II 的复杂度为

$$\mathcal{O}(dN). \tag{73}$$

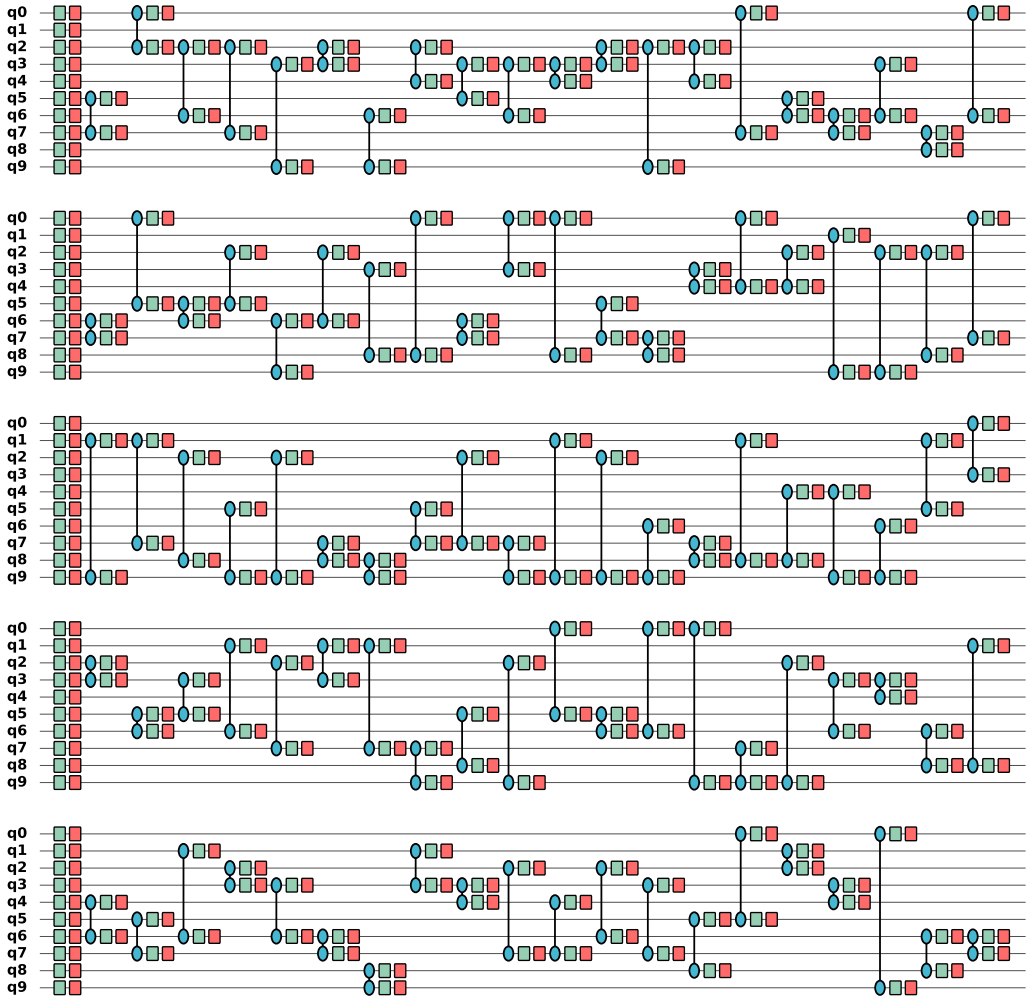

图18：使用 $T = 20$ 为 S-RQC 数据集生成的 AQER 加载电路。绿色和红色模块分别表示 R_Y 和 R_Z 量子门。由线连接的蓝色圆圈表示 R_{ZZ} 量子门。

步骤 III：基于梯度的参数微调。步骤 III 对 AQER 量子电路中的所有可训练参数执行 T_3 次基于梯度的优化迭代。参数总数为

$$P = 5T + 2N, \tag{74}$$

其中，$5T$ 来自步骤 I 中 T 次迭代里的单量子比特旋转和双量子比特旋转，而 $2N$ 来自步骤 II 中确定的额外单量子比特旋转。

为了估算步骤 III 中一次迭代的复杂度，我们考虑计算损失函数相对于所有 P 参数的梯度所需的成本。使用参数移位规则，或对状态向量模拟进行自动微分时，计算每个偏导数都需要对整个量子电路进行两次前向计算。每次前向计算的成本为 $\mathcal{O}(d(T+N))$，因此计算一个偏导数的成本为 $\mathcal{O}(d(T+N))$。由于共有 $P = \mathcal{O}(T+N)$ 个参数，完整的梯度计算成本为 $\mathcal{O}(d(T+N)^2)$。因此，在步骤 III 中执行 T_3 个梯度步的总成本为

$$\mathcal{O}(dT_3(T+N)^2). \tag{75}$$

Ove经典数据的总体复杂度。结合公式（72）、（73）和（75），总体计算 ation在经典数据场景下，AQER 的复杂度为 a-

$$\mathcal{O}\big(dTN^2 + dN + dT_3(T+N)^2\big), \tag{76}$$

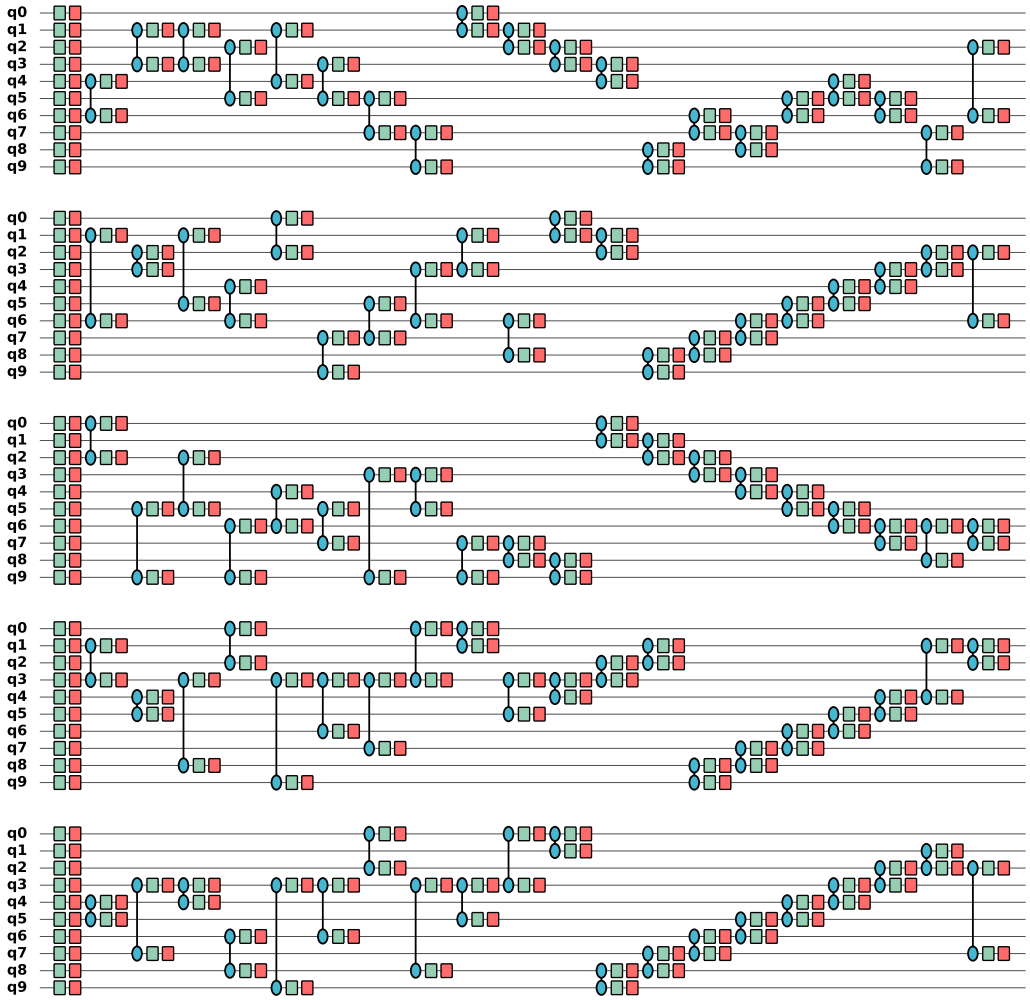

图 19：使用 $T = 20$ 为 GS-TFIM 数据集（$N = 10$）生成的 AQER 加载电路。绿色和红色方块分别表示 R_Y 和 R_Z 量子门。由线连接的蓝色圆圈表示 R_{ZZ} 量子门。

可概括为

$$\mathcal{O}\big(d\,T_3\,T^2\ \mathrm{polylog}(d)\big). \tag{77}$$

这种缩放与希尔伯特空间维数 d 呈线性关系（对于精确的态矢量模拟而言，这是不可避免的），与步骤 I 的迭代次数 T 及梯度步数 T_3呈多项式关系。

G.2 量子数据

现在我们来考虑量子数据设定：目标态 ρ 可在量子设备上获得，且不假设存在对 ρ 的显式经典描述。在这种情况下，AQER 所需的所有约化密度矩阵、熵和梯度，均通过有限次测量估算。为方便起见，我们用 M 表示用于估算每个期望值或约化密度矩阵中每个元素的射击次数。

量子数据的步骤 I 和步骤 II。与经典情况一样，步骤 I 会添加 $\mathcal{O}(T)$ 个量子门，以降低纠缠度量。每次迭代中，候选门的数量至多为 N^2。每个候选门的成本为 $\mathcal{O}(MT)$，因此，在量子数据场景下，步骤 I 的总成本随以下因素增长：

$$\mathcal{O}(MT^2N^2). \tag{78}$$

在步骤 II 中，AQER 会估计 N 个单量子比特约化密度矩阵。这需要 $\mathcal{O}(MN)$ 次射击和 $\mathcal{O}(N)$ 次经典运算，因此步骤 II 的复杂度为

$$\mathcal{O}(MN). \tag{79}$$

步骤 III：基于梯度的参数微调。 对于量子数据的基于梯度的优化，我们可以使用参数移位规则来计算梯度。对于每个参数，估计一个偏导数需要 $\mathcal{O}(M(T+N))$ 次运算。与之前一样，参数总数为$P = 5T + 2N = \mathcal{O}(T+N)$。因此，步骤 III 中的一次完整梯度计算需要

$$\mathcal{O}(M(T+N)^2) \tag{80}$$

次运算。执行 T_3 个梯度步后，步骤 III 的总复杂度为：

$$\mathcal{O}(MT_3(T+N)^2). \tag{81}$$

量子数据的总体复杂度。 结合式 (78)、(79) 和 (81)，AQER 处理量子数据的总体计算复杂度为

$$\mathcal{O}(MT_3T^2\mathrm{poly}(N)). \tag{82}$$

H 利用 AQER 高效加载 IQP 量子电路态

本节分析 AQER 在瞬时量子多项式（IQP）电路产生的经典向量或量子态上的性能。我们首先在定理 H.1 中证明，当 IQP 角度位于已知离散网格上，且我们可以经典访问完整态矢量时，AQER 至多经过 $|E|$ 次步骤 I 迭代即可精确重构一个加载电路。随后，我们放宽离散网格假设，并在定理 H.2 中证明，对于角度为独立同分布连续变量的一般 IQP 态，AQER 以高概率找到一个加载电路，使纠缠度量$\mathcal{S}$ 降至至多 ε，且 AQER 中步骤 I 的迭代次数不超过 $|E|$。最后，定理 H.3 针对角度固定为 $\pi/8$ 且相互作用图未知的离散 IQP 态族，建立了类似的精确加载保证；在此情形下，我们只能量子访问 IQP 态。设 D 为图 E的度。在这种情况下，AQER 仍能以至少 $1-\delta$的概率恢复一个精确加载电路，步骤 I 的迭代次数不超过 $|E|$，且量子访问总次数关于 N、$\log(1/\delta)$、2^D 和 $|E|$ 为多项式。

定理 H.1（离散 IQP 态矢量的精确加载）。 设 $N \geq 2$ 为量子比特数，并设 $K \in \mathbb{N}$ 为已知的网格大小。我们考虑一个全连接 *IQP* 电路态

$$|v(\boldsymbol{\omega})\rangle = H^{\otimes N}\left(\prod_{\{i,j\}\in E} e^{-i\omega_{ij}Z_iZ_j/2}\right)H^{\otimes N}|0^N\rangle\,, \tag{83}$$

其中 E 是 *IQP* 电路的未知相互作用图，且每个参数均未知，并被限制在离散网格上

$$\omega_{ij} = \frac{k_{ij}\pi}{2K+1}, \qquad k_{ij} \in \{-2K, -2K+1, \ldots, 2K-1, 2K\}. \tag{84}$$

假设我们对 $|v(\boldsymbol{\omega})\rangle$ 的态矢量具有经典访问。那么，在相差一个整体全局相位的意义下，*AQER* 算法可通过 T 次步骤 *I* 的迭代，为该 *IQP* 态重构一个精确加载电路，其中 $T \leq |E|$。

证明。 对于已知目标态为 IQP 态的情况，我们对 AQER 的步骤 I 中的优化作如下限制。为方便起见，令 $\mathcal{I}_t \subseteq [N]$ 表示在步骤 I 的第 t 次迭代中新增的双量子比特门块所涉及的量子比特集合，并定义累积集合 $\mathcal{J}_t := \bigcup_{t'=1}^{t}\mathcal{I}_{t'}$。在第 $t+1$ 次迭代中，考虑作用于量子比特对 $\{i,j\}$ 的候选双量子比特门块。对于每个量子比特$q \in \{i,j\}$，若 $q \in \mathcal{J}_t$，则我们通过$R_Z(0)R_Y(0) = I$将该门块中相应的单量子比特部分限制为恒等算符。若 $q \notin \mathcal{J}_t$，则将单量子比特部分设为 $R_Z(\pi)\,R_Y(-\pi/2) = -iH$，该操作在全局相位意义下实现哈达玛门。对于作用于 $\{i,j\}$上的纠缠门 R_{ZZ}，我们将其旋转角限制在网格 $\alpha_{ij} = a_{ij}\pi/(2K+1)$ 上，其中 $a_{ij} \in \{-2K, -2K+1, \ldots, 2K-1, 2K\}$。因此，纠缠度量的优化通过网格搜索完成，搜索对象为量子比特对 $\{i,j\}$ 以及 α_{ij}的离散取值。

由于纠缠度量 $\mathcal{S}$ 定义为单量子比特 Rényi 熵之和，因此它在任意单量子比特酉算子下保持不变。在上述步骤 I 的优化规则下，对目标 IQP 态的 $\mathcal{S}$ 进行优化，等价于对剩余IQP态 $|v'(\boldsymbol{\omega})\rangle$的 $\mathcal{S}$ 进行优化，后者定义如下：

$$|v'(\boldsymbol{\omega})\rangle := \left(\prod_{\{i,j\}\in E} e^{-i\omega_{ij}Z_iZ_j/2}\right) H^{\otimes N}|0^N\rangle = \left(\prod_{\{i,j\}\in E} e^{-i\omega_{ij}Z_iZ_j/2}\right)|+\rangle^{\otimes N}, \tag{85}$$

其中 $|+\rangle = H|0\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$。

接下来，我们聚焦于步骤 I 的第一次迭代。由于 $|E|>0$，量子比特 $\{p,q\}\in E$ 上存在一个候选门块，优化问题为

$$\begin{aligned}
&\arg\min_{\alpha_{pq}} \mathcal{S}\left(e^{-i\alpha_{pq}Z_pZ_q/2}|v'(\boldsymbol{\omega})\rangle\right)\\
=\ &\arg\min_{\alpha_{pq}} \mathcal{S}\left(e^{-i\alpha_{pq}Z_pZ_q/2}\left(\prod_{\{i,j\}\in E} e^{-i\omega_{ij}Z_iZ_j/2}\right)|+\rangle^{\otimes N}\right)\\
=\ &\arg\min_{\alpha_{pq}} \mathcal{S}\left(e^{-i\alpha_{pq}Z_pZ_q/2-i\sum_{\{i,j\}\in E}\omega_{ij}Z_iZ_j/2}|+\rangle^{\otimes N}\right)\\
=\ &\arg\min_{\alpha_{pq}} \mathcal{S}\left(e^{-i\sum_{\{i,j\}\in E}\beta_{ij}(\alpha_{pq})Z_iZ_j/2}|+\rangle^{\otimes N}\right),
\end{aligned} \tag{86}$$

其中我们定义有效角度

$$\beta_{ij}(\alpha_{pq}) := \begin{cases} \omega_{ij}, & \{i,j\}\neq\{p,q\},\\ \omega_{ij}+\alpha_{pq}, & \{i,j\}=\{p,q\}. \end{cases} \tag{87}$$

式 (86) 中的纠缠度量由单量子比特雷尼-2熵 $\mathcal{S}(\rho) = \sum_{n=1}^{N}\mathcal{S}_n(\rho_n)$ 定义，这些熵仅取决于单量子比特约化密度矩阵（RDM）：$\mathcal{S}_n(\rho_n) = -\log_2 \mathrm{Tr}[\rho_n^2]$，其中 ρ_n 是一个 n-量子比特量子态 ρ 在第 n-个量子比特上的 RDM。因此，只需计算式 (86) 中量子态的单量子比特约化密度矩阵。值得注意的是，任何单量子比特约化密度矩阵都可以作布洛赫分解 $\rho_n = \frac{1}{2}(I + x_nX_n + y_nY_n + z_nZ_n)$，其中 $x_n = \mathrm{Tr}(\rho_nX_n)$、$y_n = \mathrm{Tr}(\rho_nY_n)$ 和 $z_n = \mathrm{Tr}(\rho_nZ_n)$ 分别是相应泡利算符在量子比特 n 上的期望值。因此，计算单量子比特约化密度矩阵可归结为求出 x_n、y_n 和 z_n。特别地，对于式 (86) 中的量子态，我们有

$$\begin{aligned}
x_n(\alpha_{pq}) &= \langle+|^{\otimes N} e^{i\sum_{\{i,j\}\in E}\beta_{ij}(\alpha_{pq})Z_iZ_j/2} X_n e^{-i\sum_{\{i,j\}\in E}\beta_{ij}(\alpha_{pq})Z_iZ_j/2}|+\rangle^{\otimes N}\\
&= \langle+|^{\otimes N} e^{i\sum_{\{i,j\}\in \mathcal{N}(n)}\beta_{ij}(\alpha_{pq})Z_iZ_j/2} X_n e^{-i\sum_{\{i,j\}\in \mathcal{N}(n)}\beta_{ij}(\alpha_{pq})Z_iZ_j/2}|+\rangle^{\otimes N} &(88)\\
&= \prod_{\{i,j\}\in\mathcal{N}(n)} \cos\big(\beta_{ij}(\alpha_{pq})\big)\langle+|^{\otimes N}X_n|+\rangle^{\otimes N} &(89)\\
&= \prod_{\{i,j\}\in\mathcal{N}(n)} \cos\big(\beta_{ij}(\alpha_{pq})\big). &(90)
\end{aligned}$$

在式 (88) 中，我们将 $\mathcal{N}(n)\subseteq E$ 定义为与 n 相连的边的集合，即每个 $\{i,j\}\in\mathcal{N}(n)$ 都满足 $n\in\{i,j\}$；并利用这样一个事实：对于 $\{i,j\}\notin\mathcal{N}(n)$，算符 Z_iZ_j 与 X_n 对易，因此在共轭变换中会相互抵消，且不影响期望值。式 (89) 随后由单条相连边的海森堡绘景恒等式推出，例如 $e^{i\beta_{mn}Z_mZ_n/2}X_ne^{-i\beta_{mn}Z_mZ_n/2} = X_n\cos\beta_{mn} - Z_mY_n\sin\beta_{mn}$，再结合任何包含 Z_m 的项在 $|+\rangle^{\otimes N}$ 上的期望值均为零这一事实，这是因为 $\langle+|Z|+\rangle = 0$。式 (90) 随后由 $\langle+|I|+\rangle = \langle+|X|+\rangle = 1$ 得出。对 Y_n 和 Z_n 进行类似分析，可得 $y_n = z_n = 0$。

因此，纠缠度量是 α_{pq} 的函数，其形式为

$$\mathcal{S}(\alpha_{pq}) = -\sum_{n=1}^{N}\log_2\left(\mathrm{Tr}[\rho_n(\alpha_{pq})^2]\right) = -\sum_{n=1}^{N}\log_2\left(\frac{1+x_n(\alpha_{pq})^2}{2}\right), \tag{91}$$

其中， $x_n(\alpha_{pq})$ 由式 (90) 给出。根据构造，式 (87) 中的有效角度 $\beta_{ij}(\alpha_{pq})$ 仅在 $\{i,j\}=\{p,q\}$ 时依赖于 α_{pq}。因此，只有 $x_p(\alpha_{pq})$ 和 $x_q(\alpha_{pq})$ 依赖于 α_{pq}:

$$x_p(\alpha_{pq}) = \cos\big(\beta_{pq}(\alpha_{pq})\big)\, C_p, \qquad C_p := \prod_{\{p,m\}\in\mathcal{N}(p)\setminus\{\{p,q\}\}} \cos\beta_{pm},$$
$$x_q(\alpha_{pq}) = \cos\big(\beta_{pq}(\alpha_{pq})\big)\, C_q, \qquad C_q := \prod_{\{q,m\}\in\mathcal{N}(q)\setminus\{\{p,q\}\}} \cos\beta_{qm},$$

其中，对于 α_{pq} 和 C_p而言， C_p 和 C_q 是常数， $C_q\neq 0$ 则由角度网格的选择决定。将这些表达式代入式（91）并合并常数项，可得

$$\mathcal{S}(\alpha_{pq}) = \text{const} - \log_2\left(\frac{1+C_p^2\cos^2\beta_{pq}(\alpha_{pq})}{2}\right) - \log_2\left(\frac{1+C_q^2\cos^2\beta_{pq}(\alpha_{pq})}{2}\right),$$

因此，在网格 $\alpha_{pq}\in\{a_{pq}\pi/(2K+1)\}$ 上最小化 $\mathcal{S}(\alpha_{pq})$，等价于最小化函数

$$g(\alpha_{pq}) := -\log_2\left(\frac{1+C_p^2\cos^2\beta_{pq}(\alpha_{pq})}{2}\right) - \log_2\left(\frac{1+C_q^2\cos^2\beta_{pq}(\alpha_{pq})}{2}\right). \tag{92}$$

容易验证， $\alpha^*_{pq}=-\omega_{pq}$ ，即 $\beta_{pq}(\alpha^*_{pq})=\omega_{pq}+\alpha^*_{pq}=0$ 是上述目标函数的唯一最优值，并且可以通过在 $\alpha_{ij}=a_{ij}\pi/(2K+1)$上进行网格搜索找到，且$a_{ij}\in\{-2K,-2K+1,\ldots,2K-1,2K\}$。

对于固定的边 $\{p,q\}\in E$，我们已确定唯一的最优选择 $\alpha^*_{pq}=-\omega_{pq}$ ，使得$\beta_{pq}(\alpha^*_{pq})=\omega_{pq}+\alpha^*_{pq}=0$。现在，我们迭代这一构造。在步骤 I 的第 t次迭代中，令 $\beta^{(t)}_{ij}$ 表示 IQP 边上的有效角度，并定义残余边集

$$F^{(t)} := \{\{i,j\}\in E : \beta^{(t)}_{ij}\neq 0\}. \tag{93}$$

根据上述讨论，如果 $F^{(t)}$ 非空，我们可以任意选取 $\{p,q\}\in F_{(t)}$，并在角度网格 α_{pq} 上使 $\mathcal{S}$ 最小化。最小值的唯一性意味着，最优选择 $\alpha^{(t)}_{pq}$ 满足

$$\beta^{(t+1)}_{pq} = \beta^{(t)}_{pq} + \alpha^{(t)}_{pq} = 0, \tag{94}$$

而其他所有角度保持不变，即对于所有 $\{i,j\}\neq\{p,q\}$， $\beta^{(t+1)}_{ij}=\beta^{(t)}_{ij}$ 。因此，残余边集在每次迭代中都会严格缩小，

$$|F^{(t+1)}| = |F^{(t)}| - 1. \tag{95}$$

由于 $|F_{(0)}|\leq|E|$，经过 $T=|F_{(0)}|\leq|E|$ 次迭代后，我们得到一个满足 $F_{(T)}=\varnothing$ 的构型，即对于所有 $\{i,j\}\in E$， $\beta^{(T)}_{ij}=0$。在这种情况下，所有双量子比特纠缠旋转都会被抵消，所得量子态的纠缠度量 $\mathcal{S}=0$ 为零。根据 $\mathcal{S}$ 定义为单量子比特雷尼熵之和，这意味着经过 T 次步骤 I 迭代后，由 $|v(\boldsymbol{\omega})\rangle$ 得到的量子态是一个乘积态，我们将其记为 $|v_{\text{prod}}\rangle$。

最后，根据引理 B.2，我们可以构造一个由 R_Z 和 R_Y 旋转组成的单量子比特电路，使其将 $|v_{\text{prod}}\rangle$ 精确映射为计算基态 $|0^N\rangle$ ，但相差一个全局相位。令 U_{I} 为通过上述步骤得到的所有步骤 I 门块的乘积，令 U_{II} 为引理 B.2 中的单量子比特电路。于是

$$U_{\text{II}}U_{\text{I}}\,|v(\boldsymbol{\omega})\rangle = |0^N\rangle \tag{96}$$

但相差一个全局相位。取伴随算子并反转量子门的顺序，我们便得到一个显式的加载电路

$$U_{\text{load}} = (U_{\text{II}}U_{\text{I}})^\dagger = U_{\text{I}}^\dagger U_{\text{II}}^\dagger \tag{97}$$

使得 $U_{\text{load}}\,|0^N\rangle = |v(\boldsymbol{\omega})\rangle$ 仅相差一个全局相位。这就完成了定理 H.1 的证明。

□

定理 H.2（一般 IQP 状态向量的近似加载）。 我们沿用定理 *H.1* 中的记号，并考虑 *IQP* 量子电路的量子态

$$|v(\boldsymbol{\omega})\rangle = H^{\otimes N}\left(\prod_{\{i,j\}\in E} e^{-i\omega_{ij}Z_iZ_j/2}\right)H^{\otimes N}|0^N\rangle\,, \tag{98}$$

其中，角度 $\{\omega_{ij}\}_{\{i,j\}\in E}$ 是从 $[-\pi,\pi]$ 中均匀采样得到的独立同分布随机变量。令 D 为 E的最大度。我们假设可以经典访问 $|v(\boldsymbol{\omega})\rangle$的态矢量。那么，对于任意 $\delta,\epsilon>0$，*AQER* 通过搜索大小为 $K=\left\lceil \frac{\pi}{2}\sqrt{\frac{DN}{\epsilon}}\right\rceil$的网格，至多使用 $T\le|E|$ 次步骤 *I* 的迭代，即可生成满足 $\mathcal{S}(U_{\mathrm{I}}|v(\boldsymbol{\omega})\rangle)\le\varepsilon$ 的量子电路 U_{I}，其概率至少为 $1-\delta$。

证明。 首先，我们证明以高概率所有参数 ω_{ij} 都满足余弦值的下界。由于每个 ω_{ij} 都是从 $[-\pi,\pi]$ 中均匀采样得到的，因此有

$$\Pr\left[|\cos\omega_{ij}|\ \ge\ \frac{\delta}{|E|}\right]=1-\frac{2}{\pi}\arcsin\left(\frac{\delta}{|E|}\right)\ \ge\ 1-\frac{\delta}{|E|}. \tag{99}$$

对所有 $\{i,j\}\in E$应用并集界可知，以至少 $1-\frac{\delta}{|E|}\cdot|E|=1-\delta$的概率有 —

$$|\cos\omega_{ij}|\ \ge\ \frac{\delta}{|E|}\quad\text{for all }\{i,j\}\in E. \tag{100}$$

以下我们假设这一高概率事件成立。

Next，我们按照定理 H.1 的证明中步骤 I 的相同优化过程进行。具体来说 y, the 作用于 $\{i,j\}$ 的双量子比特门的角度为 α_{ij}，且受限于网格

$$\mathcal{G}_K:=\left\{\frac{a\pi}{2K+1}:a\in\{-2K,-2K+1,\dots,2K-1,2K\}\right\},$$

其中 $K=\left\lceil\frac{\pi}{2}\sqrt{\frac{DN}{\epsilon}}\right\rceil$。我们还规定，每对 $\{i,j\}\subseteq[N]$ 至多添加一个双量子比特门块。在此限制下，步骤 I 在每条边上至多更新一次，因此迭代总数满足 $T\le|E|$。对于每个 $\{p,q\}\in E$，式（92）优化问题中的网格搜索会找到一个值 $\alpha^\star_{pq}$，使其与 $-\omega_{pq}$ 的距离至多为 $\pi/(2K+1)$。因此，经过 $T=|E|$ 次迭代后，剩余的残余态可写为

$$|v(\boldsymbol{\beta})\rangle=e^{-i\sum_{\{i,j\}\in E}\beta_{ij}Z_iZ_j/2}|+\rangle^{\otimes N}\,, \tag{101}$$

其中，每个有效角均满足 $|\beta_{ij}|\le\pi/(2K+1)$。

量子态 $|v(\boldsymbol{\beta})\rangle$ 的纠缠度量可以通过类似于公式 (88)–(91) 的计算得到。具体而言，

$$\begin{aligned}
\mathcal{S}(|v(\boldsymbol{\beta})\rangle) =&\ -\sum_{n=1}^{N}\log_2\left(\frac{1+x_n^2}{2}\right)\\
&\le\sum_{n=1}^{N}(1-x_n^2) &(102)\\
&=N-\sum_{n=1}^{N}\prod_{\{i,j\}\in\mathcal{N}(n)}\cos^2\beta_{ij} &(103)\\
&\le N-N\cos^{2D}\left(\frac{\pi}{2K+1}\right) &(104)\\
&\le N-N\left[1-\left(\frac{\pi}{2K+1}\right)^2\right]^D &(105)\\
&\le N-N\left[1-D\left(\frac{\pi}{2K+1}\right)^2\right] &(106)
\end{aligned}$$

$$= DN\left(\frac{\pi}{2K+1}\right)^2 \leq \epsilon. \tag{107}$$

这里，公式 (102) 通过对 $x \in [0,1]$ 令 $x = 1 - x_n^2$，由不等式 $-\log_2 \frac{2-x}{2} \leq x$ 得出。公式 (103) 由公式 (90) 得出，该公式给出 $x_n^2 = \prod_{\{i,j\}\in\mathcal{N}(n)} \cos^2 \beta_{ij}$。公式 (104) 利用$|\beta_{ij}| \leq \pi/(2K+1)$ 和 $|\mathcal{N}(n)| \leq D$ 推导得出，因此每个乘积均以下界 $\cos^{2D}(\pi/(2K+1))$ 为界。公式 (105) 对 $|x| \leq \pi/2$ 使用 $\cos^2 x = 1 - \sin^2 x \geq 1 - x^2$。公式 (106) 由对 $x \in [0,1]$ 的伯努利不等式 $(1-x)^D \geq 1 - Dx$ 得出。最后，公式 (107) 由选择$K = \left\lceil \frac{\pi}{2}\sqrt{\frac{DN}{\epsilon}} \right\rceil$ 得出，该选择确保 $DN\left(\frac{\pi}{2K+1}\right)^2 \leq \epsilon$。定理 H.2 的证明至此完成。

□

定理 H.3（离散 IQP 态的精确加载）。设 $N \geq 2$ 为量子比特的数量。我们考虑一个全连接 *IQP* 电路态

$$|v(E)\rangle = H^{\otimes N}\left(\prod_{\{i,j\}\in E} e^{-i\pi Z_i Z_j/8}\right) H^{\otimes N}|0^N\rangle, \tag{108}$$

其中，E 是 *IQP* 电路的未知相互作用图，最大度为 D。假设我们可以量子访问该量子态 $|v(\boldsymbol{\omega})\rangle$。那么，除一个整体全局相位外，*AQER* 算法可以使用 T 次步骤 *I* 迭代，重构出 *IQP* 态的精确加载电路，其中 $T \leq |E|$，其概率至少为 $1-\delta$；通过对给定 *IQP* 态的调用次数不超过 $\mathcal{O}\left(|E|N^2 2^D \log\frac{N^2|E|}{\delta}\right)$ 次。

证明。我们遵循定理 H.1 中的步骤 I 优化方法，但将每个量子比特对 $\{p,q\}$ 的网格搜索限制为两个角度 $\alpha_{pq} \in \{0, -\pi/4\}$。需要指出的是，现在我们只有对 IQP 态的量子访问，而不是经典向量访问。我们的目标是保证在每次迭代中，网格搜索以至少 $1-\delta/|E|$ 的概率选中 E 中的一个量子比特对，其中失败概率完全来自有限的测量次数。因此，在经过 $T = |E|$ 次步骤 I 迭代后，E 中的所有量子比特对以至少 $1-\delta$ 的概率被消去。

接下来，我们关注第一次迭代。与定理 H.1 的证明类似，残余态可写为

$$|v'(E)\rangle = e^{-i\sum_{\{i,j\}\in E}\beta_{ij}(\alpha_{pq})Z_iZ_j/2}|+\rangle^{\otimes N}. \tag{109}$$

对于任意候选对 $\{p,q\}$ 以及任意 α_{pq}的取值，单量子比特布洛赫系数满足

$$x_n(\alpha_{pq}) = \prod_{\{i,j\}\in\mathcal{N}(n)} \cos\big(\beta_{ij}(\alpha_{pq})\big), \qquad y_n(\alpha_{pq}) = 0, \qquad z_n(\alpha_{pq}) = 0.$$

在定理 H.1 中，我们证明了步骤 I 对纠缠度量进行优化

$$\mathcal{S}(\alpha_{pq}) = -\sum_{n=1}^{N}\log_2\left(\frac{1+x_n(\alpha_{pq})^2}{2}\right), \tag{110}$$

并且，通过增加 $|x_n(\alpha_{pq})|$可以使 $\mathcal{S}$ 减小。由于 $|\mathcal{N}(n)| \leq D$ 且 $\beta_{ij}(\alpha_{pq}) \in\{0, \pm\pi/4\}$，因此我们有

$$x_n(\alpha_{pq}) \in \left\{\left(\tfrac{\sqrt{2}}{2}\right)^d : 0 \leq d \leq D\right\} \subseteq [2^{-D/2}, 1].$$

因此，对于每个 $\{p,q\} \subseteq [N]$，通过使用

$$\mathcal{O}\left(2^D \log\frac{N^2|E|}{\delta}\right)$$

通过量子测量次数，我们可以估计相关的 $x_p(\alpha_{pq})$ 和 $x_q(\alpha_{pq})$，其加性误差小于 $2^{-D/2}$ 的某个常数倍，从而以至少 $1-\delta/(N^2|E|)$ 的概率确定：当 α_{pq} 从 0 切换为 $-\pi/4$ 时，$|x_p|$ 和 $|x_q|$ 是否都会增加（如果 $\{p,q\} \in E$）。遍历所有 $\{p,q\} \subseteq [N]$ 后，我们便可以以至少 $1-\delta/|E|$ 的概率识别出至少一个 $\{p,q\} \in E$。

对所有 $T = |E|$ 次迭代重复这一过程，并集界表明，总体成功概率至少为 $1 - \delta$。量子测量次数总数为

$$\mathcal{O}\left(2^D \log \frac{N^2|E|}{\delta}\right) \cdot N^2 \cdot |E| = \mathcal{O}\left(|E|N^2 2^D \log \tfrac{N^2|E|}{\delta}\right).$$

至此，定理 H.3 的证明完成。

□

I 大语言模型的使用

在本研究中，大语言模型（LLM）仅作为通用工具，用于进行少量语言润色和表达优化。LLM 未对任何科学思想、实验或结果作出贡献，其协助生成的所有内容均由作者仔细审阅和编辑。作者对论文的全部内容承担全部责任，包括受 LLM 辅助文本影响的内容。