arXiv:2304.00549v2 [quant-ph] 9 Nov 2023

变分量子本征求解器的变分去噪

Quoc Hoan Tran, [1], * Shinji Kikuchi, [1] 和 Hirotaka Oshima[1†] *Quantum Laboratory, Fujitsu Research, Fujitsu Limited, Kawasaki, Kanagawa 211-8588, Japan*(日期：2023年11月10日)

变分量子本征求解器（VQE）是一种混合算法，有望在实际化学问题中实现量子优势，而这些问题目前尚无法由经典计算机有效求解。VQE 使用经典优化器训练参数化量子电路，以近似给定哈密顿量的本征值和本征态。然而，VQE 面临任务特定设计和机器特定架构方面的挑战，尤其是在含噪量子设备上运行时。这会对其可训练性、准确性和效率产生负面影响，导致产生含噪量子数据。我们提出了变分去噪，这是一种无监督学习方法，利用参数化量子神经网络从含噪的变分量子本征求解器输出中学习，从而改进 VQE 的求解结果。与含噪输入数据相比，我们的方法能够显著降低 H_2, LiH 和 BeH_2 分子哈密顿量的能量估计误差，提高与基态的保真度，并适用于横场伊辛模型。令人意外的是，训练只需要含噪数据。变分去噪可以集成到量子硬件中，从而增强其作为端到端量子数据处理方法的通用性。

引言。— “量子优势”一词源于量子计算机有望比最强大的经典计算机更快地解决实际问题。尽管人们预计容错量子计算机能够带来量子优势，但相关量子硬件中的噪声效应意味着，要实现有用且可靠的计算，需要大量量子比特。尽管如此，含噪中等规模量子计算（NISQ）设备 [1] 仍因有望在某些数学任务 [2–5]中超越经典计算机而持续激发研究界的热情。鉴于经典神经网络（NN）在各种实际机器学习问题中展现出具备抗噪声能力的可训练性，近期人们重新燃起了乐观情绪，认为含噪中等规模量子计算机上的量子神经网络有望带来量子优势。尽管在使用人工和人工构造的数据集 [6], 的非实用监督学习任务中已展示出量子加速，人们仍在努力寻找这样的复杂分布：经典神经网络在这些分布上本质上效率低下，而量子神经网络却能表现出色 [7]。

使用变分量子算法（VQA）训练参数化量子电路（PQC），是含噪中等规模量子设备中的一个重要因素 [8]。在 VQA 中，会构建一个优化问题，其损失函数根据 PQC 中可观测量的期望值计算得出。经典计算机通过调整量子电路的参数来最小化该损失函数。然而，复杂损失景观中随机优化器的不稳定性、量子误差以及实验设置中的系统误差等因素，都可能阻碍 PQC 的可训练性。在某些条件下，例如量子电路较深、定义损失的可观测量具有非局域性 [10], , 以及量子纠缠过强 [11], 就可能出现著名的贫瘠高原现象 [9], , 此时损失景观实际上是平坦的。即使使用浅层局域量子模型和局域代价函数来避免贫瘠高原，较差的局部最优解若远离全局最优解，也会导致模型无法训练 [12, 13]。此外，量子噪声、任务特定电路设计以及超参数设置等因素，都会对训练性能产生显著影响。

VQA 中两种具有代表性的方法是量子近似优化算法（量子近似优化算法） [14] 和变分量子本征求解器（变分量子本征求解器） [15]。变分量子本征求解器旨在应用于量子化学和材料科学，以确定哈密顿量的最低能级和基态。由于在经典计算机上寻找此类问题的解，其计算成本会随系统规模呈指数级增长，因此变分量子本征求解器有望展现量子优势 [16]。尽管已成功应用于小规模问题 [17], , 变分量子本征求解器的大规模实现仍面临挑战 [18]。

变分量子本征求解器通过对某个任意的初始状态 $|\psi\rangle$ 施加参数为 $\boldsymbol{\zeta}$ 的酉算子 $V(\boldsymbol{\zeta})$，生成量子态 $|\psi(\boldsymbol{\zeta})\rangle = V(\boldsymbol{\zeta})|\psi\rangle$ 。给定一个厄米算子 H，变分量子本征求解器可估计其基态 $|\psi_{\min}\rangle$ ，该基态对应于一个未知的最小本征值（最低能量） $E_{\min}$ ，其对应的算子为 H。变分量子本征求解器的能量估计 $E_{\boldsymbol{\zeta}}$ 满足界限 $E_{\min} \leq E_{\boldsymbol{\zeta}} = \langle\psi(\boldsymbol{\zeta})| H |\psi(\boldsymbol{\zeta})\rangle$，随后由经典优化器进行优化，以获得最小值。然而，对 $E_{\boldsymbol{\zeta}}$ 的估计可能成为实现近期实用量子优势的障碍，因为所需的量子测量次数可能多到难以承受 [19]。这可能会使变分量子本征求解器无法进行大量迭代 [12]。此外，优化器的随机性以及量子设备中的噪声，可能导致不同的最终 $|\psi(\boldsymbol{\zeta})\rangle$ ，例如 $|\psi_1\rangle$ ， . . . ， $|\psi_N\rangle$。这些状态被视为含噪数据，只能提供有关目标基态的部分信息。

在此，我们提出一种从含噪变分量子本征求解器数据中学习目标态的方法。这个想法可以通过自编码器框架的量子扩展来实现 [20], 在

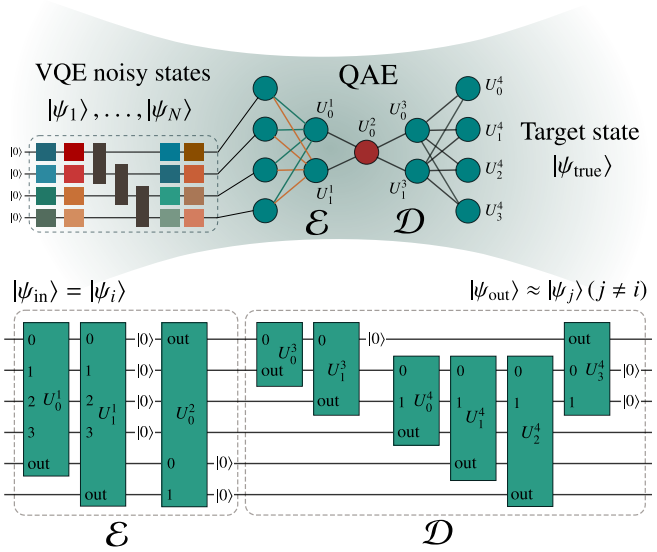

图 1. 通过量子自编码器进行变分去噪。由变分量子本征求解器电路制备的量子态被视为“含噪”态。量子自编码器训练编码和解码参数化量子电路，将每个含噪态压缩到一个子系统中，使得从约化态恢复出的状态尽可能接近其他含噪态。量子自编码器是从经典前馈神经网络推广而来的，其中每个神经元对应一个量子比特，而酉电路则连接后续耦合层中的神经元（上图）。隐藏层中的神经元数量减少，以压缩输入态。下图展示了量子自编码器的一种基于电路的实现示例[4,2,1,2,4]。

其中，前馈神经网络经过训练，可在输出层重现输入。在自编码器中，隐藏层的节点数通常少于输入层和输出层的节点数，这表明重构时保留了某些关键信息。文献 [21] 介绍了量子自编码器（QAE），并将其应用于压缩和去噪特定量子态 [22–25]。我们的方案通过硬件高效方法从含噪量子数据中学习，从而增强了量子自编码器的实际应用能力。它不仅可作为针对特定态的去噪方法，还致力于提升变分量子本征求解器的鲁棒性和效率。

量子自编码器（*QAE*）。— 量子自编码器生成一个编码器 $\mathcal{E}$，将 n量子比特的输入态映射为 k量子比特的（$k < n$）约化态，并生成一个解码器 $\mathcal{D}$，从约化态重建输入态。 $\mathcal{E}$ 和 $\mathcal{D}$ 分别使用参数为 $\boldsymbol{\theta}_e$ 和 $\boldsymbol{\theta}_d$的参数化量子电路实现。训练它们的目标是最大化输入数据分布 $\mathcal{C}$ 上输入态 ρ^{in} 与量子自编码器输出之间的期望保真度 $\mathcal{F}$

$$\max_{\boldsymbol{\theta}_e,\boldsymbol{\theta}_c} \boldsymbol{E}_{\rho^{\text{in}}\sim\mathcal{C}} \left[\mathcal{F}(\rho^{\text{in}}, \mathcal{D}(\mathcal{E}(\rho^{\text{in}}))\right]. \tag{1}$$

实际上，当将纯态$|\psi_1\rangle$、……、$|\psi_N\rangle$ 视为从 $\mathcal{C}$中抽样得到的训练数据时，式（1）可简化为最大化经验代价

$$\mathcal{L}(\boldsymbol{\theta}_e, \boldsymbol{\theta}_c) = \frac{1}{N}\sum_{i=1}^{N} \mathcal{F}(|\psi_i\rangle, \mathcal{D}(\mathcal{E}(|\psi_i\rangle))). \tag{2}$$

这里，纯态 $|\psi\rangle$ 与密度矩阵 ρ之间的保真度定义为 $\mathcal{F}(|\psi\rangle, \rho) = \langle\psi|\rho|\psi\rangle$。该量可通过 SWAP 测试实现，我们在训练电路中的 [26] 展示了这一测试。

图 1 上图展示了一个基于耗散量子神经网络的量子自编码器示例 [27, 28]。这里，耗散量子神经网络将经典前馈神经网络推广为一种模型，其中每个神经元对应一个量子比特，酉电路则连接相邻耦合层中的神经元。在下一层完成前向过程后，每层中的量子比特都会发生耗散，从而能够为量子自编码器构造约化空间。考虑结构为 $[m_1, m_2, \ldots, m_M]\,(M \geq 2)$ 的量子自编码器，对于第 j 个神经元（位于第 $i+1$层），我们将 U_j^i 记作作用于该神经元自身量子比特及前一层神经元的参数化酉算子。第 i 层与第 $i+1$ 层之间的酉算子记作 $U^i = \prod_j U_j^i$。我们将 β_i 记作第 i层含有 m_i 个神经元时的状态。因此，β_i 与 β_{i+1} 之间的变换可通过通道表示

$$\beta_{i+1} = \mathcal{Q}_i(\beta_i) = \text{Tr}_i\left[U^i(\beta_i \otimes (|0\rangle\langle 0|)^{\otimes m_{i+1}})(U^i)^\dagger\right]. \tag{3}$$

在这里， $\mathcal{Q}_i$ 将一个 m_i量子比特态映射为一个（m_i+m_{i+1}）量子比特态，然后通过对第 m_{i+1}层中的量子比特进行偏迹，将其缩减为一个 i量子比特态。例如，在图[4,2,1,2,4]中的量子自编码器1中，编码器 $\mathcal{E} = \mathcal{Q}_1\mathcal{Q}_2$ 包含两个通道：$\mathcal{Q}_1$ 将 4 量子比特态（第 1 层）扩展为 6 量子比特态，然后缩减为 2 量子比特态（第 2 层）；以及 $\mathcal{Q}_2$将 2 量子比特态扩展为 3 量子比特态，然后缩减为 1 量子比特态（第 3 层）。解码器 $\mathcal{D}$ =$\mathcal{Q}_3\mathcal{Q}_4$ 对这个 1 量子比特态进行逆向处理，以重构最终的 4 量子比特态。量子自编码器可在量子电路中实现，其量子比特数量为 $\max\{(m_i+m_{i+1})\}$ （图1下方面板）。

参考文献 [22, 23, 25] 中的 QAE 实现是在酉算子空间内进行优化，这限制了实验规模。我们提出了一种基于试探态的 QAE 实现，该实现硬件高效，并能随量子比特数量增加而良好扩展

$$U_j^i(\boldsymbol{\theta}) = \prod_{l=1}^{L}\left[R_{\text{loc}}^{(l)}(\boldsymbol{\theta}_l)V_{\text{ent}}\right] R_{\text{loc}}^{(L+1)}(\boldsymbol{\theta}_{L+1}). \tag{4}$$

在此，每个酉算子 U_j^i 由 L 个模块以及作用于每个量子比特的末端单量子比特旋转组成。每个模块包含局部单量子比特旋转 $R^{(l)}{}_{\text{loc}}(\boldsymbol{\theta}_l) = \otimes_k R(\theta_{k,l})$，以及双量子比特纠缠门。在以下实验中，我们使用 RY CZ 试探态，其中单量子比特旋转门 R 是 Pauli Y 矩阵的参数化旋转门 $R_Y(\theta) = e^{-iY\theta/2}$，而纠缠门 V_{ent} 由

受控相位门 $CZ(k,k+1)$组成，并按环形排列，作用于索引为 k 的量子比特。参数数量仅随量子比特数和试探态模块数线性增长。与参考文献 [22, 23], 相比，这种方法在训练资源方面显著高效，但在扩展规模时仍需进一步改进。

去噪量子自编码器。由于量子自编码器会从输入中去除无关信息，并将其压缩为较小的状态，因此可用于量子数据去噪。这一理念源自经典去噪自编码器 [29], , 后者经过训练，能够从受损输入中重建出干净的输入。在文献 [24], 中，研究者使用无噪声的逻辑态作为目标态，训练量子自编码器去除预定义码空间中各状态的噪声。然而，由于不可克隆定理，在量子训练中制备目标态的副本作为参考态往往颇具挑战。我们可以通过使用来自同一含噪数据源的输入态和目标态对来克服这一问题，例如带有比特翻转和退相干错误的量子态，以及小幅随机酉变换 [22]。不过，目前量子自编码器的去噪应用仅限于特定的量子态系列，例如 GHZ 态、W 态、Dicke 态和簇态 [22]。我们将展示如何将量子自编码器应用于变分量子本征求解器：量子噪声和优化器的随机行为可能导致最终态含噪。

考虑在特定设置下（例如初始化、优化器的迭代次数和量子噪声条件）运行 VQE 后得到的最终量子态 ρ。我们可以假设， ρ 是根据目标态 ρ^*从分布 $\mathcal{C}_{\rho^*}$ 中生成的。例如， ρ^* 是基态，而 $\mathcal{C}_{\rho^*}$ 是在随机优化器中经过固定次数的迭代后得到最终 VQE 解的过程。训练过程旨在最大化经验损失

$$\mathcal{L}(\boldsymbol{\theta}_e,\boldsymbol{\theta}_c)=\frac{1}{N}\sum_{i=1}^{N}\mathcal{F}(\rho_i^A,\mathcal{D}(\mathcal{E}(\rho_i^B))), \tag{5}$$

给定 N成对训练态（ρ_i^A， ρ_i^B），其中 ρ_i^A， ρ_i^B从 $\mathcal{C}_{\rho^*}$中抽取。我们将 $\mathcal{C}_{\rho^*}$ 视为含噪数据源，它可以生成这些量子态的不同样本。如果这些量子态与 ρ^*共享关键特征，训练量子自编码器（QAE）就能输出与目标态具有较高重叠度的量子态。在我们的优化过程中，我们将同时扰动随机逼近（SPSA）方法 [30] 与 AMSGrad 方法 [31] 结合使用。SPSA 的效率较高，因为测量次数与训练参数总数 [26]无关。

结果。我们将量子自编码器（QAE）应用于对变分量子本征求解算法生成的状态进行去噪。在接下来的实验中，我们在一个控制良好的平台上考察 QAE 电路，不考虑量子噪声的影响，但会在 [26] 中讨论这一影响。含噪数据可视为由施加于 VQE 电路的含噪量子信道生成的状态（第一个实验），也可视为 VQE 优化过程中提前停止所得的结果（第二和第三个实验）。在 [26], 中，我们还考虑由量子噪声与提前停止优化共同产生的含噪数据，或由 VQE 电路的不同拟设产生的含噪数据。我们使用 Qiskit [32, 33] 模拟对 H_2和 LiH 分子进行数值实验，其中以斯莱特型轨道（STO-3G）基组表示的分子哈密顿量分别映射为含 2 个量子比特（ H_2 ）和 4 个量子比特（ LiH ）的量子比特哈密顿量 [34, 35]。对于 VQE 电路，我们构造硬件高效电路，其中包含由 SU(2) 张成的单量子比特操作和双量子比特受控-X 纠缠门 [36]。这里，我们使用泡利 Y 和 Z 的旋转算符作为单量子比特门 [26]。

在第一个实验中，我们将过程 $\mathcal{C}_{\rho^*}$ 视为比特翻转或退极化噪声信道，并在对 H_2 进行 250 次 SPSA优化、对 LiH 进行 1000 次 SPSA优化后，将其应用于 VQE 电路末端。在比特翻转信道中，每个量子比特以 p_b 的概率施加泡利X门。在退极化信道中，每个量子比特以相应的概率$\{1-\frac{3p_d}{4},\frac{p_d}{4},\frac{p_d}{4},\frac{p_d}{4}\}$施加门 {I、泡利X、泡利Y、泡利Z}。此处，我们设定 $p_b=p_d=0.2$ （改变 p_b 和 p_d时的结果见 [26]）。

在第二个实验中，我们将 $\mathcal{C}_{\rho^*}$ 视为 VQE 的提前停止过程，分别考虑 H_2 和 LiH 在经过 10 次和 100 次 SPSA 迭代后的 VQE 电路输出。含噪变分量子本征求解器数据由不同的 SPSA 随机种子生成。对于这两个实验，我们分别使用 $L=1$ 个试探态模块，为 H_2 构建量子自编码器[2,1,2] ，为 LiH 构建量子自编码器[4,1,4] 。在每个键长下，我们使用 200 对含噪变分量子本征求解器数据训练量子自编码器。我们生成 1000 个含噪样本，并用能量最高的 200 个样本进行测试。

图2 和 3 展示了不同键长下含噪数据和去噪数据的能量。插图描绘了 $|\Delta E|$ 原始能量和去噪能量相对于基态能量的能量误差。有趣的是，即使使用含噪数据进行训练，我们的方法平均而言仍能以较小的波动改善基态能量估计。

接下来，我们展示量子自编码器（QAE）也可用于复杂度更高的非平凡态。我们将量子自编码器应用于含有 N 个自旋的一维横场伊辛模型（TFIM）的VQE数据。哈密顿量定义为 $H=-\sum_i J_i\sigma_i^z\sigma_{i+1}^z-\sum_i g_i\sigma_i^x$，其中 σ_i^x 和 σ_i^z是作用于第 j个自旋的泡利X和Z矩阵， J_i 为交换耦合， g_i 是沿 x 轴的横向磁场。我们考虑具有开放边界条件的均匀相互作用$J_i=1$ 和均匀格点势 $g_i=g>0$。该模型存在两个相：有序相（$g<1.0$）和无序相（$g>1.0$），具体取决于基态是否破坏或保持$\prod_j\sigma_j^x$ 自旋翻转对称性。含噪数据由提前停止优化时的VQE电路生成，其中使用同时扰动随机逼近（SPSA）进行$4N$次迭代。我们构建了含 $L=3$ 个试探态模块的量子自编码器$[N,1,N]$。我们使用100

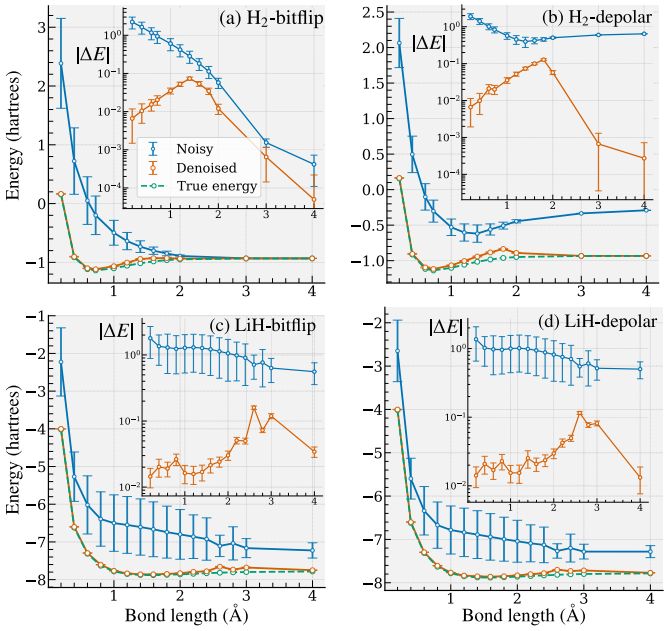

图 2. 在不同键长和不同噪声类型（比特翻转、退极化）下，H_2 和 LiH 分子的含噪变分量子本征求解器数据与去噪数据的能量。在比特翻转信道中，每个量子比特以 0.2 的概率发生翻转。在退极化信道中，对每个量子比特以 0.2 的概率施加单量子比特退极化操作。插图展示了这些能量与基态能量之间的能量差 $|\Delta E|$ 。实线和误差条分别表示 200 个测试样本的平均值和标准差。

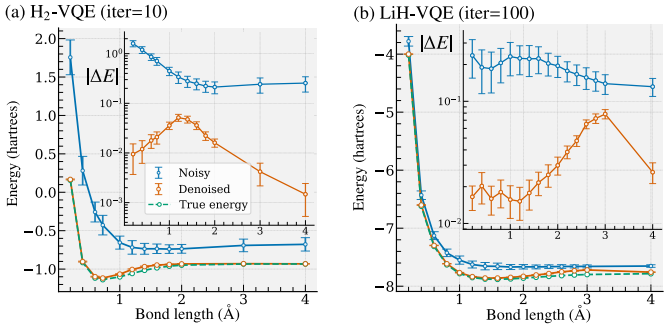

图 3. 变分量子本征求解器优化早期步骤中含噪数据的能量，以及（a）H_2 （10 次迭代）和（b）LiH 分子（100 次迭代）的去噪数据能量。插图展示了这些能量与基态能量之间的能量差 $|\Delta E|$ 。实线和误差条分别表示 200 个测试样本的平均值和标准差。

对含噪样本进行训练，并使用1000个含噪样本进行测试。

图 4 展示了一维横场伊辛模型在 $|\Delta E|$ 、6、8 个自旋以及 $N = 4$，1，0（基态具有较大纠缠熵时）的 1000 个测试样本的能量误差分布。虽然增大 N 会使问题更难，但我们的方法可以将误差降低近一个数量级 [26]。

最后，为了理解去噪机制

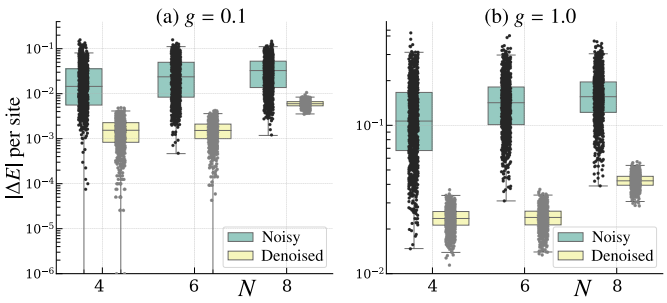

图 4. 箱线图和抖动图展示了一维横场伊辛模型在 $N = 4$、6、8 个自旋条件下 1000 个测试样本的能量误差 $|\Delta E|$ 分布。箱体表示四分位距，须线标示第 5 和第 95 百分位数。箱体中的横线表示中位数。

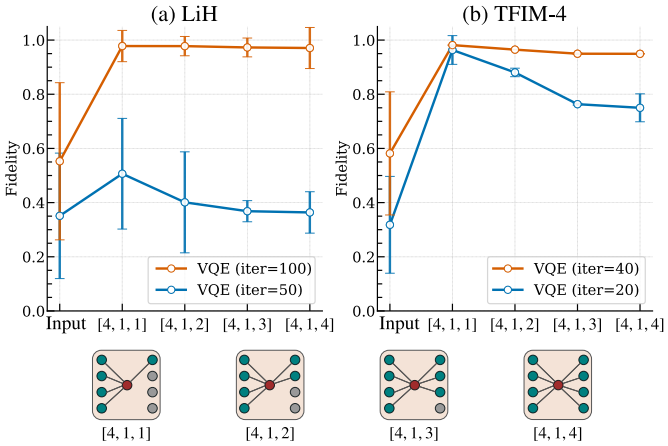

图 5. QAE[4,1,4] 各子系统输入与输出之间的保真度，以及它们与 (a) LiH 和 (b) TFIM-4 基态中对应部分之间的保真度。含噪数据由经过不同 SPSA 迭代的 VQE 电路生成。灰色节点表示非活跃神经元，其中子系统 [4,1,1],[4,1,2], 和 [4,1,3] 分别对应基态中的第一个、第一个和第二个、以及第一个、第二个和第三个量子比特组成的子系统。实线和误差线分别表示 1000 个测试样本的平均值和标准差。

量子自编码器中，我们研究数据向前传递时，量子自编码器各子系统的输出与目标态对应部分之间的保真度。我们考察了 LiH 在键长为 1 时，经过 50 次（高噪声）和 100 次（低噪声）同时扰动随机逼近迭代的含噪变分量子本征求解器数据。此外，还考察了横场伊辛模型在 $N = 4$ （TFIM-4）下，经过 $g = 1.0$ 20 次（高噪声）和 40 次（低噪声）同时扰动随机逼近迭代的数据。量子自编码器[4,1,4]采用 $L = 1$ 试探态模块，并按照先前实验中的相同条件进行训练，再用 1000 个含噪数据样本进行测试。如图 5所示，当输入态经过量子自编码器传递时，目标态的一部分会在子系统的输出中得到重构。各子系统中的变分量子本征求解器输出态 [4,1,1], [4,1,2], [4,1,3],和 [4,1,4] 分别对应于重构目标态的第一个量子比特、前两个量子比特、前三个量子比特，最后是全部量子比特。在使用低

噪声水平数据（橙色曲线）进行成功训练时，量子自编码器的耗散特性使其能够在瓶颈神经元处保留目标态的部分信息，并依次重构目标态的其他部分。然而，如果使用高噪声水平数据训练量子自编码器，此时变分量子本征求解器输出态与基态的重叠度较低，无关信息便会泄漏至瓶颈处，从而阻碍解码器依次重构基态（蓝色曲线）。

结论.—我们提出了一种从 VQE 含噪数据中学习并分离噪声的方法。该方法能够舍弃含噪且无关的信息，同时将相关信息压缩到约化空间中，从而提高 VQE 的输出与真实目标态之间的重叠度。由于量子自编码器的训练独立于 VQE 问题中的哈密顿量，因此它可以在优化过程中通过细化中间态来补充 VQE。我们的方法尤其适用于含噪中等规模量子设备，因为这类设备中含噪量子数据很常见，而且获取无噪声数据受到限制。它可以为进一步研究量子模型在学习量子数据集方面的优势铺平道路 [37, 38]。另一个值得关注的方面是，我们的方法有望用于连接混合量子系统。对于同一问题，如果不同平台上的多种算法得到不同的输出态，我们的方法可以在稳健的平台上产生稳定的结果。

作者感谢 Koki Chinzei、Yuichi Ka-mata、Mikio Morita 和Yasuhiro Endo 富有成效的讨论。

* tran.quochoan@fujitsu.com

[1] J. Preskill, Q NISQ 时代及以后量子计算，Quantum **2**，79（2018年）。[2] C. Neill、P. Roushan、K. Kechedzhi、S. Boixo、S. V. Isakov、V. Smelyanskiy、A. Megrant、B. Chiaro、A. Dunsworth、K. Arya、R. Barends、B. Burkett、Y. Chen、Z. Chen、A. Fowler、B. Foxen、M. Giustina、R. Graff、E. Jeffrey、T. Huang、J. Kelly、P. Klimov、E. Lucero、J. Mutus、M. Neeley、C. Quintana、D. Sank、A. Vainsencher、J. Wenner、T. C. White、H. Neven 和 J. M. Martinis，使用超导量子比特展示量子霸权的蓝图，《科学》**360**，195（2018年）。[3] F. Arute、K. Arya、R. Babbush、D. Bacon、J. C. Bardin、R. Barends、R. Biswas、S. Boixo、F. G. S. L. Brandao、D. A. Buell、B. Burkett、Y. Chen、Z. Chen、B. Chiaro、R. Collins、W. Courtney、A. Dunsworth、E. Farhi、B. Foxen、A. Fowler、C. Gidney、M. Giustina、R. Graff、K. Guerin、S. Habegger、M. P. Harrigan、M. J. Hartmann、A. Ho、M. Hoffmann、T. Huang、T. S. Humble、S. V. Isakov、E. Jeffrey、Z. Jiang、D. Kafri、K. Kechedzhi、J. Kelly、P. V. Klimov、S. Knysh、A. Korotkov、F. Kostritsa、D. Landhuis、M. Lind- mark、E. Lucero、D. Lyakh、S. Mandr`a、J. R. Mc- Clean、M. McEwen、A. Megrant、X. Mi、K. Michielsen、M. Mohseni、J. Mutus、O. Naaman、M. Neeley、C. Neill、M. Y. Niu、E. Ostby、A. Petukhov、J. C. Platt、C. Quin- tana、E. G. Rieffel、P. Roushan、N. C. Rubin、D. Sank、K. J. Satzinger、V. Smelyanskiy、K. J. Sung、M. D. Tre- vithick、A. Vainsencher、B. Villalonga、T. White、Z. J. Yao、P. Yeh、A. Zalcman、H. Neven 和 J. M. Marti- nis，《使用可编程超导处理器实现量子霸权》，《自然》574，505（2019）。[4] H.-S. Zhong、H. Wang、Y.-H. Deng、M.-C. Chen、L.-C. Peng、Y.-H. Luo、J. Qin、D. Wu、X. Ding、Y. Hu、P. Hu、X.-Y. Yang、W.-J. Zhang、H. Li、Y. Li、X. Jiang、L. Gan、G. Yang、L. You、Z. Wang、L. Li、N.-L. Liu、C.-Y. Lu 和 J.-W. Pan，《利用光子实现量子计算优势》，《科学》（纽约州纽约市）370，1460（2020）。[5] H.-S. Zhong、Y.-H. Deng、J. Qin、H. Wang、M.-C. Chen、L.-C. Peng、Y.-H. Luo、D. Wu、S.-Q. Gong、H. Su、Y. Hu、P. Hu、X.-Y. Yang、W.-J. Zhang、H. Li、Y. Li、X. Jiang、L. Gan、G. Yang、L. You、Z. Wang、L. Li、N.-L. Liu、J. J. Renema、C.-Y. Lu 和 J.-W. Pan，《利用受激压缩光进行相位可编程高斯玻色采样》，《物理评论快报》127，180502（2021）。[6] Y. Liu、S. Arunachalam 和 K. Temme，《监督式机器学习中严谨且稳健的量子加速》，《自然·物理》（2021）。[7] X. Gao、E. R. Anschuetz、S.-T. Wang、J. I. Cirac 和 M. D. Lukin，《利用量子关联增强生成模型》，《物理评论X》12，021037（2022）。[8] M. Cerezo、A. Arrasmith、R. Babbush、S. C. Benjamin、S. Endo、K. Fujii、J. R. McClean、K. Mitarai、X. Yuan、L. Cincio 和 P. J. Coles，《变分量子算法》，《自然综述：物理》，1（2021）。[9] J. R. McClean、S. Boixo、V. N. Smelyanskiy、R. Babbush 和 H. Neven，《量子神经网络训练景观中的贫瘠高原》，Nat. Commun. 9，4812（2018）。[10] M. Cerezo、A. Sone、T. Volkoff、L. Cincio 和 P. J. Coles，《浅层参数化量子电路中依赖代价函数的贫瘠高原》，Nat. Commun. 12，1791（2021）。[11] C. Ortiz Marrero、M. Kieferov´a 和 N. Wiebe，《纠缠诱导的贫瘠高原》，《PRX量子》2，040316（2021）。[12] E. R. Anschuetz 和 B. T. Kiani，《量子变分算法深陷陷阱》，Nat. Commun. 13，7760（2022）。[13] L. Bittel 和 M. Kliesch，《训练变分量子算法是 NP难问题》，《物理评论快报》127，120502（2021）。[14] E. Farhi、J. Goldstone 和 S. Gutmann，《量子近似优化算法》，arXiv 预印本：1411.4028（2014）。[15] A. Peruzzo、J. McClean、P. Shadbolt、M.-H. Yung、X.-Q. Zhou、P. J. Love、A. Aspuru-Guzik 和 J. L. O'Brien，《光子量子处理器上的变分本征值求解器》，Nat. Commun. 5，4213（2014）。[16] I. O. Sokolov、P. K. Barkoutsos、P. J. Ollitrault、D. Greenberg、J. Rice、M. Pistoia 和 I. Taver- nelli，《强关联区域中的量子轨道优化酉耦合簇方法：量子算法能否胜过其经典对应方法？》，《化学物理杂志》152，124107（2020）。[17] A. Kandala、A. Mezzacapo、K. Temme、M. Takita、M. Brink、J. M. Chow 和 J. M. Gambetta，《适用于小分子和量子磁体的硬件高效变分量子本征求解器》，《自然》549，242

(2017年)。
[18] JTilly、H. Chen、S. Cao、D. Picozzi、K. Setia、Y. Li、E. Grant、L. Wossnig、I. Rungger、G. H. Booth 和 J. Tennyson，《变分量子本征求解器：方法与最佳实践综述》，Physics Reports 986，1（2022年），《变分量子本征求解器：方法与最佳实践综述》。[19] J. F. Gonthier、M. D. Radin、C. Buda、E. J. Doskocil、C. M. Abuan 和 J. Romero，《测量是近期实现化学领域实用量子优势的障碍：资源分析》，Phys. Rev. Res. 4，033154（2022年）。[20] I. J. Goodfellow、Y. Bengio 和 A. Courville，*DeepLearning*（MIT出版社，美国马萨诸塞州剑桥，2016年）http://www.deeplearningbook.org。[21] J. Romero、J. P. Olson 和 A. Aspuru-Guzik，《用于高效压缩量子数据的量子自编码器》，Quantum Science and Technology 2，045001（2017年）。[22] D. Bondarenko 和 P. Feldmann，《利用量子自编码器对量子数据进行去噪》，《物理评论快报》124，130502（2020年）。[23] T. Achache、L. Horesh 和 J. Smolin，《利用量子自编码器对量子态进行去噪——理论与应用》，arXiv预印本：2012.14714（2020年）。[24] D. F. Locher、L. Cardarelli 和 M. M¨uller，《利用量子自编码器进行量子纠错》，arXiv预印本：2202.00555（2022年）。[25] J. Pazem 和 M. H. Ansari，《利用brainbox量子自编码器缓解纠缠态中的误差》，arXiv预印本：2303.01134（2023年）。[26]有关量子自编码器的训练、变分量子本征求解器（VQE）的实现、更多去噪结果，以及探究如何在量子自编码器（QAE）训练中针对贫瘠高原设计试探态的详情，请参阅补充材料。[27] K. Beer、D. Bondarenko、T. Farrelly、T. J. Osborne、R. Salzmann、D. Scheiermann 和 R. Wolf，《深度量子神经网络的训练》，Nat. Commun. 11（2020年）。[28] K. Sharma、M. Cerezo、L. Cincio 和 P. J. Coles，《训练基于耗散感知机的量子神经网络的能力，《物理评论快报》128，180505（2022年）。[29] P. Vincent、H. Larochelle、Y. Bengio 和 P.-A. Man- zagol，《利用去噪自编码器提取并组合鲁棒特征》（计算机协会，美国纽约州纽约市，2008年），第1096–1103页。[30] J. C. Spall，《随机优化与同时扰动法》，载于 *Proceedings of the 31st Conference on Winter Simulation: Simulation—a Bridge to the Future - Volume 1*，WSC '99（计算机协会，美国纽约州纽约市，1999年），第101–109页。[31] S. J. Reddi、S. Kale 和 S. Kumar，《论Adam及其后续方法的收敛性》，载于 *Proceedings of the 6th International Conference on Learning Representations ICLR 2018* 2018 ()。[32] Qiskit：用于量子计算的开源框架（2021年）。[33]IBM 量子挑战赛 2021，https://github.com/qiskit-community/ibm-quantum-challenge-2021（2021年），访问日期：2022年12月20日。[34] S. B. Bravyi 和 A. Y. Kitaev，《费米子量子计算》，《物理学年鉴》298，210（2002年）。[35] J. T. Seeley、M. J. Richard 和 P. J. Love，《用于量子计算电子结构的 Bravyi-Kitaev 变换》，《化学物理学杂志》137，224109 2012 1208 5986（），..[36] P. K. Barkoutsos、J. F. Gonthier、I. Sokolov、N. Moll、G. Salis、A. Fuhrer、M. Ganzhorn、D. J. Egger、M. Troyer、A. Mezzacapo、S. Filipp 和 I. Tavernelli，《电子结构计算的量子算法：粒子-空穴哈密顿量与优化波函数展开》，《物理评论A辑》98，022322（2018年）。[37] L. Schatzki、A. Arrasmith、P. J. Coles 和 M. Cerezo，《量子机器学习的纠缠数据集》，arXiv预印本：2109.03400（2021年）。[38] A. Nakayama、K. Mitarai、L. Placidi、T. Sugimoto 和 K. Fujii，《用于机器学习的VQE生成量子电路数据集》，arXiv预印本：2302.09751（2023年）。

《变分量子本征求解器的变分去噪》补充材料

Quoc Hoan Tran, [1, *] Shinji Kikuchi, [1] 和 Hirotaka Oshima[1†] *Quantum Laboratory, Fujitsu Research, Fujitsu Limited, Kawasaki, Kanagawa 211-8588, Japan*（日期：2023年11月10日）

本补充材料详细介绍了计算过程、正文中展示的实验以及补充图表。本节中的公式、图和表的编号均以 S 为前缀（例如，公式 (S1)、图 S1、表 S1）；不带此前缀的编号（例如，公式 (1)、图 1、表 1）则指正文中的相应内容。

arXiv:2304.00549v2 [quant-ph] 9 Nov 2023

目录

I. 量子自编码器训练 A. 电路设计 1
B. 训练算法 1
3

II. 变分量子本征求解器的实现 5

III. 分子哈密顿量的变分量子本征求解器去噪 5
A. 去噪演示 B. 量子自编码器的噪声鲁棒性 C. 数据中不同 5
噪声水平下量子自编码器的性能 6
8

IV. 横场伊辛模型变分量子本征求解器的去噪 9

V. 为量子自编码器设计试探态 10

VI. 量子自编码器训练中的贫瘠高原 14

VII. 去噪机制 14

参考文献 16

I. 训练量子自编码器

本节介绍量子自编码器（QAE）的量子电路实现，以及用于训练的完整电路。本研究采用的 QAE 概念参考了文献 [1]，其量子电路实现基于文献 [3]中的 Qiskit 库 [2] 。然而，先前的实现未提供可在真实含噪中等规模量子计算机上运行的硬件高效拟设。在该实现中，训练所需的测量次数随参数总数增长。此外，参数数量会随系统中所用量子比特数量呈指数增长。在我们的实现中，我们为 QAE 提供了一种硬件高效拟设电路，可显著减少训练参数的数量。我们还采用同时扰动随机逼近（SPSA） [4]优化来训练这些量子线路，其测量次数与训练参数总数无关。

A. 电路设计

图S1 展示了用于训练量子自编码器的电路设计。由于本研究中的量子自编码器将输入态 ρ 作为含噪态，并训练参数化电路，使输出态尽可能接近

* tran.quochoan@fujitsu.com

表 I. 参考文献 [1, 3] 中早期实现所用的参数数量，以及我们针对 QAE 的实现。此处L 表示 QAE 的每个酉算子中使用的试探态模块数量（见图S1右下角）。

结构	参考文献 [3]中的实现	我们的实现（L=1）	我们的实现（L=2）	我们的实现（L=3）
[2, 1, 2]	96	14	21	28
[3, 1, 3]	304	20	30	40
[4, 1, 4]	1088	26	39	52
[4, 2, 1, 2, 4]	2400	58	87	116

另一含噪态 $\tilde{\rho}$，因此整个训练电路由四部分组成：（I）用于制备输入态 ρ 和目标态 $\tilde{\rho}$ 的电路（在我们的实验中，输入态和参考态均为通过 VQE 电路制备的纯态）；（II）量子自编码器电路；（III）用于计算量子自编码器输出与参考态 $\tilde{\rho}$之间保真度的电路；以及（IV）测量部分。量子自编码器输出的密度矩阵与参考纯态之间的保真度计算为 $2p_0 - 1$，其中 p_0 是测量得到 0 的概率。

图 S1 展示了具有 $[4, 2, 1, 2, 4]$ 层节点的量子自编码器（QAE）结构的电路实现，用于对 4 量子比特态进行去噪（右下图）。每个神经元对应一个量子比特，酉电路连接相邻耦合层中的神经元。每层中的量子比特在下一层的前向过程之后会发生耗散，因此可以为 QAE 构建约化空间。考虑结构为 $[m_1, \ldots, m_M]\,(M \geq 2)$ 的 QAE。对于第 j 个神经元和第 $i+1$ 层，我们将 U_j^i 记为作用于该神经元及其前一层神经元的参数化酉算子。这些酉矩阵需要进行优化。第 i 层与第 $i+1$ 层之间的酉算子可表示为 $U^i = \prod_j U_j^i$。我们将 β_i 记为第 i 层中包含 m_i 个神经元的状态。因此，β_i 与 β_{i+1} 之间的变换可通过量子信道表示

$$\beta_{i+1} = \mathcal{Q}_i(\beta_i) = \mathrm{Tr}_i \left[U^i (\beta_i \otimes (|0\rangle\langle 0|)^{\otimes m_{i+1}}) (U^i)^\dagger \right]. \tag{S1}$$

这里，$\mathcal{Q}_i$ 将一个 m_i量子比特态扩展为一个（$m_i + m_{i+1}$）量子比特态，然后通过对第 i层的量子比特做偏迹，将其缩减为一个 m_{i+1}量子比特态。这样，采用拓扑结构 [4,2,1,2,4], 时，编码器 $\mathcal{E} = \mathcal{Q}_1\mathcal{Q}_2$ 由两个通道中的三个幺正变换组成：第一个通道 $\mathcal{Q}_1$ 将4量子比特态（第1层）扩展为6量子比特态，然后将其缩减为2量子比特态（第2层）；第二个通道 $\mathcal{Q}_2$ 进一步将2量子比特态扩展为3量子比特态，然后将其缩减为1量子比特态（第3层）。解码器 $\mathcal{D} = \mathcal{Q}_3\mathcal{Q}_4$ 对这个1量子比特态执行逆向操作，以通过六个幺正变换重构最终的4量子比特态。因此，这种量子自编码器的耗散式设计最多需要{（$m_i + m_{i+1}$）} 个量子比特，而训练电路最多需要{（$m_i + m_{i+1}$）} 个量子比特，其中 $m_1 + 1$ 个量子比特用于计算保真度。

参考文献 [1, 3] 中的作者考虑了量子电路部分 (II) 中每个酉算子 U_j^i（作用于第 j 个神经元和前一层的神经元）所有可能取值的优化。这里，U_j^i 表示为

$$U_j^i = e^{iK_j^i} \text{ where } K_j^i = \sum_{\sigma \in \{I,X,Y,Z\}^{\otimes(m_i+1)}} k_\sigma \sigma, \tag{S2}$$

其中，厄米矩阵 K_j^i 由泡利基 $\sigma \in \{I, X, Y, Z\}^{\otimes(m_i+1)}$ 展开中的 4^{m_i+1} 个系数 k_σ 唯一确定，该展开对应于 $2^{m_i+1} \times 2^{m_i+1}$的实向量空间。因此，对于结构为 $[m_1, m_2, \ldots, m_M]$的 QAE，训练参数总数为

$$P = \sum_{i=1}^{M-1} m_{i+1} 4^{m_i+1}, \tag{S3}$$

其随系统中使用的量子比特数量呈指数级增长。因此，我们需要探索大量训练参数，这使得训练阶段既耗时又充满挑战。此外，原则上，在真实的含噪中等规模量子设备上实现式（S2）中的表示效率低下。

在我们的研究中，我们为量子自编码器提供了一种基于试探态的实现，其中每个酉块 U_j^i 由一个参数化量子电路实现，该电路包含 L 个模块，并在 R 每个单量子比特上施加最终的参数化旋转门（图 S1右下方）。在这里，每个模块都包含作用于每个单量子比特的参数化旋转门和双量子比特纠缠门。在我们的实验中，我们使用泡利 Y 门的旋转算符作为 $R_Y(\theta) = e^{-iY\theta/2}$。非参数化的双量子比特纠缠门为受控 X 门或受控 Z 门（受控量子比特以灰色圆圈表示，作用酉算子以灰色方框表示），并按照量子比特的索引沿圆周排列。我们的实现具有硬件高效性，可以减少训练所需的参数数量（见表 I）。在相同的去噪量子自编码器深度下，参数数量仅随每个试探态模块所用的量子比特数和层数线性增长。

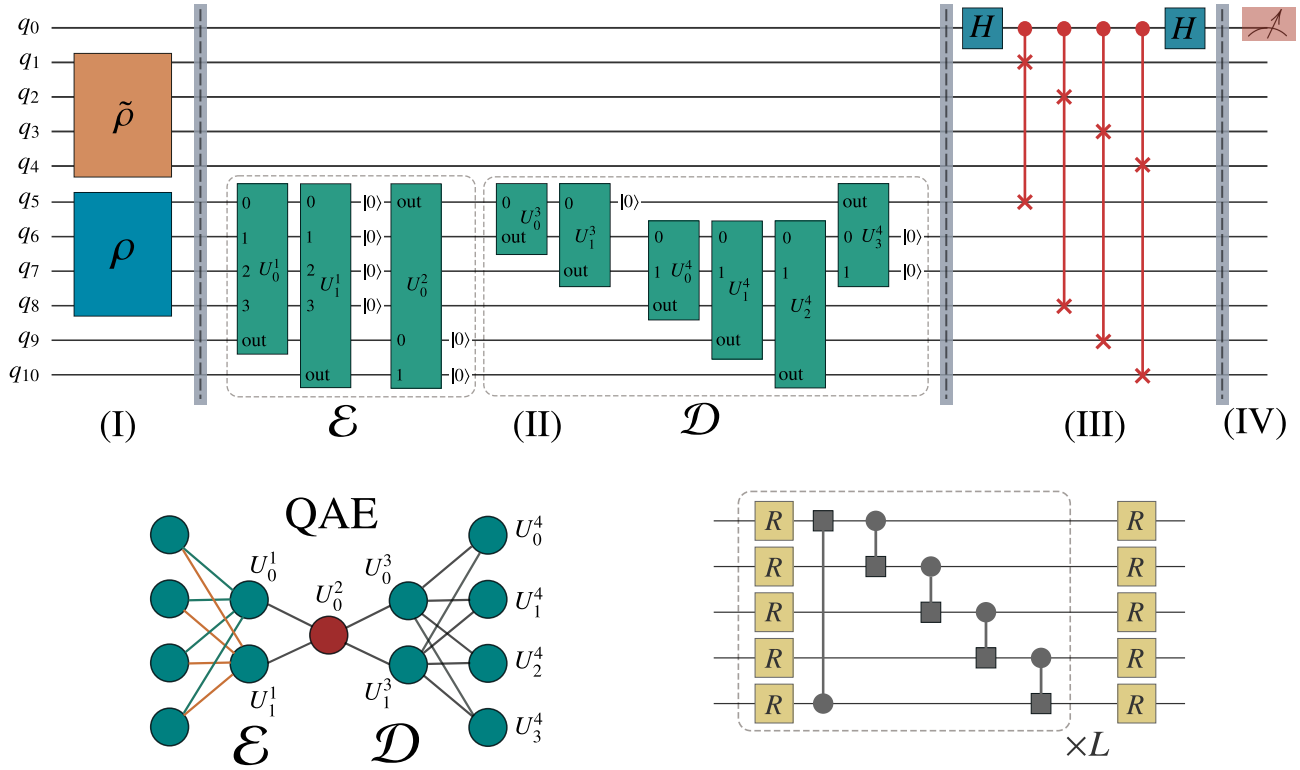

图 S1. 用于训练量子自编码器的电路实现。上图展示了使用含噪训练数据训练量子自编码器所用的量子电路。我们采用了参考文献 [1, 3]中的实现概念。该电路由四个部分组成，并由灰色竖线分隔：(I) 用于制备输入态 ρ 和参考态 $\tilde{\rho}$ 的电路（注意，这些态是从其他 VQE 电路中获得的含噪态）；(II) 用于实现量子自编码器抽象结构的电路，其中包含编码器 $\mathcal{E}$ 和解码器 $\mathcal{D}$（此处，该电路表示量子自编码器左下方的结构，其中每层有 $[4, 2, 1, 2, 4]$ 个节点）；(III) 由两个哈达玛门和四个交换门组成，用于计算量子自编码器输出与目标态 $\tilde{\rho}$之间保真度的电路；以及 (IV) 测量部分。在量子自编码器中，每个酉块 U_j^i 均由一个参数化试探电路实现，该电路由 L 个模块以及作用于每个单量子比特的末端参数化旋转门 R 组成，如右下图所示。此处，每个模块都包含作用于每个单量子比特的参数化旋转门和双量子比特纠缠门。在我们的实验中，我们采用泡利 Y 门的旋转算符作为 $R_Y(\theta) = e^{-iY\theta/2}$。纠缠门可以是非参数化受控门（受控 X 门或受控 Z 门），也可以是参数化受控旋转门（受控 RX 或受控 RZ），并按照量子比特的索引循环排列。它们的受控量子比特以灰色圆圈表示，所执行的酉算子以灰色方块表示。

B. 训练算法

去噪量子自编码器的训练可以表述为最大化经验代价

$$\mathcal{L}(\boldsymbol{\theta}) = \frac{1}{N}\sum_{i=1}^{N} \mathcal{F}(\rho_i^A, \mathcal{D}(\mathcal{E}(\rho_i^B))), \tag{S4}$$

给定 N成对训练态（ρ_i^A，ρ_i^B），其中 $\boldsymbol{\theta} = (\theta_1、\ldots\ldots、\theta_P)$ 是用于构造 $\mathcal{E}$ 和 $\mathcal{D}$ 的参数向量，包含 P 个参数。成对训练态（ρ_i^A，ρ_i^B）从含噪过程 $\mathcal{C}_{\rho^*}$ 中采样得到，而我们不知道真实的目标态 ρ^*。因此，这与经典自编码器的训练不同，后者仅考虑由相同数据组成的数据对进行训练。

最大化式（S4）中代价的一种典型方法是普通梯度下降（GD），该方法已用于训练参考文献 [3] 中的量子自编码器。这里，参数在迭代过程中更新，更新依据代价函数的梯度，以提高代价。在第 $(t+1)$ 次迭代中，梯度下降将参数 θ_p 更新为

$$\theta_p^{(t+1)} = \theta_p^{(t)} + \eta \frac{\partial \mathcal{L}}{\partial \theta_p}(\boldsymbol{\theta}^{(t)}), \tag{S5}$$

其中，η 是学习率。在每次迭代中，对每个索引 p计算梯度 $\frac{\partial \mathcal{L}}{\partial \theta_p}(\boldsymbol{\theta})$。考虑到

扰动步长 ϵ 足够小时，我们可以使用有限差分法近似计算梯度，如下所示：

$$\frac{\partial \mathcal{L}}{\partial \theta_p}(\boldsymbol{\theta}) \approx \frac{\mathcal{L}(\boldsymbol{\theta}+\boldsymbol{\epsilon}_p)-\mathcal{L}(\boldsymbol{\theta})}{\epsilon}, \tag{S6}$$

其中 $\boldsymbol{\epsilon}_p$ 是一个指示向量，在参数向量 $\boldsymbol{\theta}$的第 ϵ 个 p位置取值为 $\boldsymbol{\theta}$，其余位置均为 0。根据式（S5）和式（S6），我们可以用每次迭代所需的测量次数来估算复杂度：N（$P+1$）$\times S$，其中 N是训练数据的数量， P是训练参数的数量，而 S是运行量子电路以估计测量部分获得 0 的概率、进而计算保真度所需的测量次数。由于 N 和 P 通常是固定的，量子测量的次数很大程度上取决于模型中的参数数量，这正是先前方法受到主要限制的原因，参见文献 [3]。

为克服这一限制，我们采用 SPSA [4] 优化器，以随机方式取代逐维梯度估计；无论参数数量多少，该方法都只需对代价函数进行两次评估。SPSA 不会逐步提供无偏梯度估计，但在存在噪声的情况下，对于大规模优化问题仍然有效。因此，建议在含噪中等规模量子设备或含噪量子模拟器的优化流程中使用该方法。

在 SPSA 中，从定义在$\{-1,1\}^P$上的 P维离散均匀分布中采样一个 P维随机方向 $\boldsymbol{b}$。沿着这一采样方向，使用有限差分法通过两次代价函数评估来近似梯度分量，得到：

$$\boldsymbol{\nabla}_{\boldsymbol{b}}\mathcal{L}(\boldsymbol{\theta}) \equiv \boldsymbol{b}\cdot\boldsymbol{\nabla}\mathcal{L}(\boldsymbol{\theta}) \approx \frac{\mathcal{L}(\boldsymbol{\theta}+\epsilon\boldsymbol{b})-\mathcal{L}(\boldsymbol{\theta}-\epsilon\boldsymbol{b})}{2\epsilon}. \tag{S7}$$

因此，梯度的随机估计量构造为 $g_{\boldsymbol{b}}(\boldsymbol{\theta}) = (\boldsymbol{\nabla}_{\boldsymbol{b}}\mathcal{L}(\boldsymbol{\theta}))\,\boldsymbol{b}$，由此得到更新规则：

$$\boldsymbol{\theta}^{(t+1)} = \boldsymbol{\theta}^{(t)} + \eta g_{\boldsymbol{b}^{(t)}}(\boldsymbol{\theta}^{(t)}), \tag{S8}$$

其中， $\boldsymbol{b}^{(t)}$ 在第 t次迭代时采样。此外，为了降低随机估计量 $g_{\boldsymbol{b}}(\boldsymbol{\theta})$的波动，我们考虑在每次迭代中对 $\boldsymbol{b}$ 进行多次采样，然后取该估计量的平均值。在我们的实验中，每次迭代使用 $\boldsymbol{b}$ 的两个不同样本。因此，每次迭代的测量次数为 $4NS$；在所有模拟中，我们使用 $N=200$ 训练数据，测量次数为 $S=1000$ 。

为了降低波动并提高优化过程的收敛性，我们将 AMSGrad 优化器的概念引入 [5] SPSA 算法中。AMSGrad 优化器在更新参数时会考虑过去梯度平方的最大值。对公式 (S8) 进行变换后，得到以下方程：

$$m_t = \beta_1 m_{t-1} + (1-\beta_t) g_t, \tag{S9}$$
$$v_t = \beta_2 v_{t-1} + (1-\beta_2) g_t^2, \tag{S10}$$
$$\hat{v}_t = \max(\hat{v}_{t-1}, v_t), \tag{S11}$$
$$\boldsymbol{\theta}^{(t+1)} = \boldsymbol{\theta}^{(t)} + \frac{\eta}{\sqrt{\hat{v}_t}+\nu} m_t, \tag{S12}$$

其中， g_t， g_{t-1}、.....表示过去的随机估计量序列 $g_{\boldsymbol{b}}(\boldsymbol{\theta})$，而 ν 是引入的一个非零小扰动，用于避免除以零。这里，我们将该优化器定义为 AMSGradSPSA。

在我们的实现中，我们进一步引入小批量学习和学习率调度 μ ，以提高学习效率。这里，小批量学习是指使用较小的数据批次训练模型，而不是一次性使用整个数据集。在小批量学习中，数据被划分为较小的批次，模型依次在每个批次上进行训练。由于小批量学习能让模型在训练过程中接触到更多样的数据，因此可以增强模型的泛化能力。在训练过程中，我们在每个训练轮次都将数据集随机划分为每批 50 个样本。我们在每个批次上计算代价函数和梯度，以更新参数。因此，每个训练轮次需要更新参数四次。所有实验均进行 100 个训练轮次。此外，用于梯度估计的小扰动 ϵ 设为 $\epsilon=0.1$ 。第一个训练轮次的学习步长设为 $\eta=\eta_0=0.1$ ，之后每经过十个训练轮次便按 0.8 的比例缩小。AMSGradSPSA 的其他参数为 $\beta_1=0.9$， $\beta_2=0.999$ 和 $\nu=10^{-8}$。

II. 变分量子本征求解器的实现

在我们的实验中，我们针对量子化学中的电子结构问题，利用变分量子本征求解器生成量子自编码器的训练数据。首先，通过宇称变换，将二次量子化费米子哈密顿量表示的分子哈密顿量映射为量子比特哈密顿量。我们可以采用 Z_2 对称性和双量子比特约化等约化技术，减少所需的量子比特数量。在实验中，我们使用宇称映射和 STO-3G 基组模拟 H_2, LiH 和 BeH_2分子。氢原子的电子位于 1s 原子轨道中；锂原子有 3 个电子，铍原子有 4 个电子，分别位于 1s、2s 和 2p（含 x、y、z分量）原子轨道中。形成分子时，这些原子轨道会参与形成分子轨道。对于我们的分子体系，H=5+1LiH 和 BeH=5+2分别对应总计 2、6（=5+1）和 7（=5+2）个分子轨道。若按自旋轨道（计入自旋向上和自旋向下）计算，则 H_2, LiH 和 BeH_2, 分子哈密顿量中分别对应 4、12 和 14 个自旋轨道。因此，表示 H_2, LiH 和 BeH_2 的原始哈密顿量分别需要 4、12 和 14 个量子比特。在后续步骤中，我们遵循参考文献 [6] 及 IBM Quantum Challenge 中的说明 2021 [7]，将这些问题约化为更紧凑的形式，同时不显著影响目标基态能量。

对于 H_2, ，宇称映射会移除与系统自旋宇称相关的两个量子比特，从而将哈密顿量编码到一个双量子比特系统中。对于 BeH_2，我们选择由四个分子轨道组成的活性空间。随后，我们使用宇称映射对八个自旋轨道的哈密顿量进行映射。此过程会消去两个量子比特，将计算规模缩减为一个六量子比特问题。

对于 LiH，原始问题需要 12 个量子比特。我们首先冻结 Li 1s 轨道中的芯电子以进行约化。由于这些电子对化学的影响并不显著，这一步可以消去两个量子比特。在 LiH 的基态中，主要相互作用涉及 H 1s 和 Li 2s 轨道。不过，由于分子的取向，$2p_z$ 轨道也会参与部分相互作用。考虑到这一点，我们可以排除与 $2p_x$ 和 $2p_y$对应的分子轨道。这一处理可进一步将系统缩减两个量子比特。接下来，我们结合使用宇称映射和双量子比特约化，再缩减两个量子比特。因此，LiH 的哈密顿量现在只需用六个量子比特表示。进一步利用哈密顿量中的对称性，并注意到它与自旋向上数算符和自旋向下数算符均对易，我们可以针对每种对称性各移除一个量子比特，从而再消去两个量子比特。完成所有这些步骤后，LiH 问题最终可约化为一个四量子比特系统。

VQE电路的主要目标是通过一系列单量子比特旋转和纠缠操作，准确表示基态波函数。量子计算机的一项显著优势是能够高效地表示和存储精确波函数。对于包含多个原子以上的系统，这种表示方式在经典计算机上会带来难以承受的计算开销。对于化学系统，我们有两种主要的拟设可供选择：q-UCC拟设 [8] 以及注重硬件效率的启发式拟设，例如双局域拟设 [9]。q-UCC拟设借鉴了物理过程，旨在将电子激发映射到量子电路，但由于需要许多层和量子门，效率往往较低。在本研究中，我们重点研究用于VQE问题的启发式拟设。VQE量子态通过以下双局域变分拟设态 [9] 制备，其参数为 $\boldsymbol{\zeta}$:

$$|\psi(\boldsymbol{\zeta})\rangle = \prod_{d=1}^{D}\left(\prod_{q=1}^{Q}\left[U_R^{q,d}(\boldsymbol{\zeta})\right] \times U_{\mathrm{Ent}}\right) \times \prod_{q=1}^{Q}\left[U_R^{q,0}(\boldsymbol{\zeta})\right]|00\ldots0\rangle\,, \tag{S13}$$

适用于 Q 个量子比特，由 D 个纠缠门 U_{Ent} 与每个量子比特上的 Q（$D+1$）旋转门交替组成。在我们的案例中，我们从零态 $|0\rangle$ 开始，而非哈特里–福克态。在实验中，对于VQE的启发式拟设，我们使用 U_R（$\boldsymbol{\zeta}$） $= R_Y$（$\boldsymbol{\zeta}_1$）R_Z（$\boldsymbol{\zeta}_2$），并且 U_{Ent} 由按线性排列、作用于索引为（k， $k+1$）的量子比特上的CNOT门组成。该电路中的参数数量为 Q（$2D+1$）。在正文中，我们使用 $D=1$ 构造变分量子电路，并通过Qiskit对同时扰动随机逼近（SPSA）的实现来优化这些电路。

III. 分子哈密顿量的变分量子本征求解器去噪

本节 展示了对分子哈密顿量的变分量子本征求解器数据进行去噪的更多结果 .

A. 去噪演示

首先，我们分别将 H_2,LiH 和 BeH_2，在 SPSA 优化器迭代 10、100 和 500 次时，变分量子本征求解器电路的输出作为含噪训练数据。含噪变分量子本征求解器数据是在 SPSA 优化器的不同随机种子下生成的。

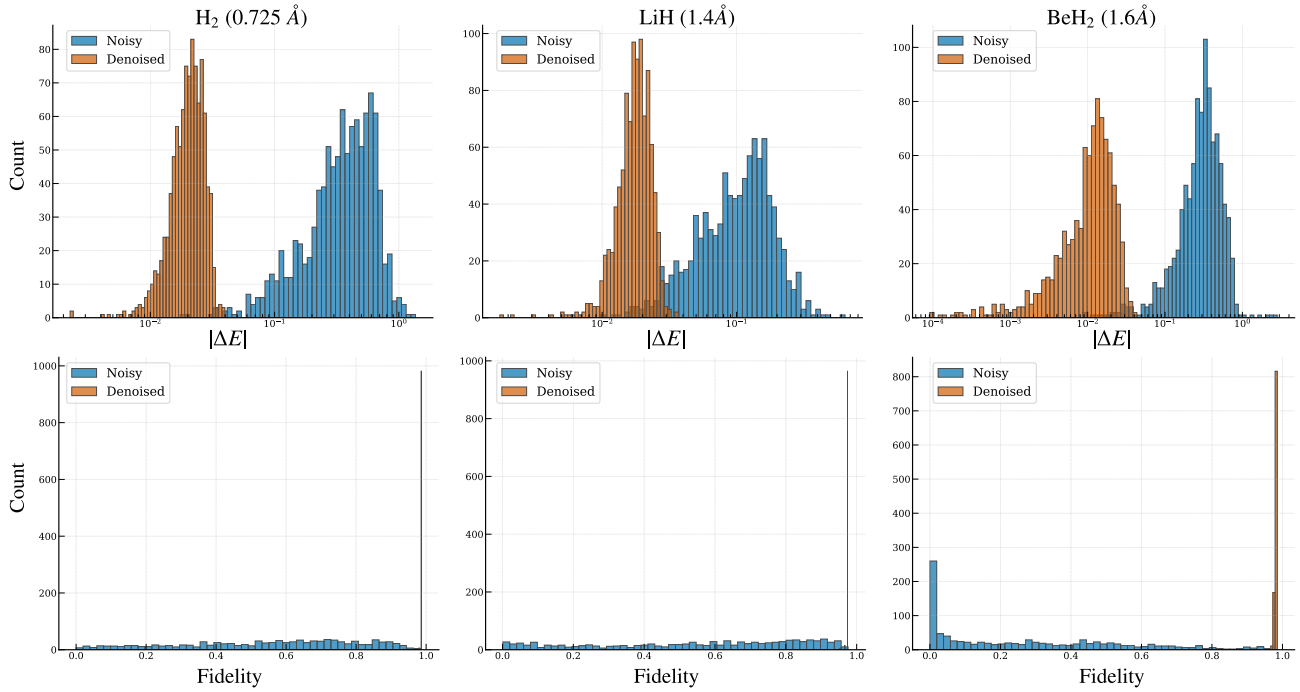

图 S2. H_2, LiH 和 BeH_2 分子在键长分别为 0.725$\mathring{A}$、1.4$\mathring{A}$ 和 1.6$\mathring{A}$ 时，与基态能量的能量差直方图，以及含噪和去噪 VQE数据与基态之间保真度的直方图。对于 H_2, LiH 和 BeH_2,，量子自编码器分别在提前停止于 10、100 和 500 次同时扰动随机逼近迭代时的 200 对训练对上进行训练，并使用 1000 个含噪 VQE样本进行测试。

我们使用 200 对含噪数据训练量子自编码器，并使用 1000 个含噪数据样本进行测试。这里，我们采用量子自编码器 [2,1,2] 用于 H_2,量子自编码器[4,1,4] 用于 LiH，以及量子自编码器[6,1,6] 用于 BeH_2 ，其包含 $L = 1$ 个试探态模块。在量子自编码器中，每个量子比特上的参数化旋转门 R 是绕泡利 Y 轴的旋转，纠缠门则是非参数化的受控 Z 门。图S2 展示了能量差的直方图 $|\Delta E|$ ，即估计能量与基态能量之差，以及含噪数据和去噪数据相对于基态的保真度直方图，分别对应 H_2, LiH 和 BeH_2 分子。在所有情况下，能量估计和与基态的重叠度均得到显著改善。

此外，我们分析了训练对数量对 LiH 的 VQE数据去噪性能的影响。图 S3 展示了 1000 个测试样本的去噪态保真度随训练对数量的变化。对于 LiH，我们使用量子自编码器[4,1,4] 和 $L = 3$ 个试探态模块。值得注意的是，即使训练对数量较少，我们的方法仍能实现较高的保真度，表明其具有高效学习的潜力。

接下来，我们考虑这样一项实验：在提前停止于 10 次和 100 次迭代时，对变分量子本征求解器电路的输出施加比特翻转和退极化噪声通道，以生成 H_2 和 LiH 的含噪数据。对于每个键长，我们使用在不同随机种子下生成的 200 对含噪变分量子本征求解器数据来训练量子自编码器。此外，我们还生成 1000 个含噪数据样本用于测试，以观察其在基态估计中的去噪效果。图 S4展示了 H_2 和 LiH 分子在各个键长下含噪数据与去噪数据的能量，以及它们与基态的保真度。这里，我们为 H[2,1,2] 构建量子自编码器$_2$ ，并为 LiH 构建量子自编码器[4,1,4] ，采用 $L = 1$ 个试探态模块；其中，量子自编码器中每个量子比特上的旋转门均为泡利 Y 门的参数化旋转门，纠缠门则为非参数化的受控 Z 门。

B. 量子自编码器中的噪声鲁棒性

在正文中，我们没有考虑量子自编码器中的噪声模型。这是因为我们对混合系统有浓厚兴趣：在这类系统中，可以利用控制良好的系统去除噪声量子平台中的噪声。

在本节中，我们进行了数值实验，以探究 QAE 在噪声下的鲁棒性。我们模拟了含有单量子比特退极化误差、双量子比特退极化误差或两者兼有的 QAE 电路。图S5展示了随着 QAE 量子电路各噪声模式中的退极化错误率（噪声水平）变化，使用 QAE（橙色曲线）求得的基态$^\circ$ $^\circ$能量 H_2 （键长为 0.725A）、LiH（键长为 1.595A）和 BeH_2 （键长为 1.6A ）的输出态能量误差如何变化。用于计算输出态能量的量子电路也在与 QAE 相同的噪声条件下进行模拟，

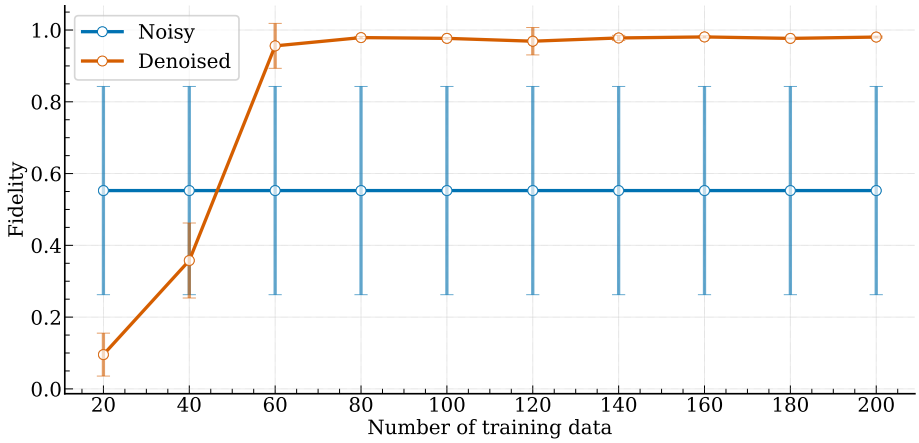

图 S3. 随训练数据数量变化，LiH 分子（键长为 1.4$\mathring{A}$）的含噪 VQE 态和去噪态与基态的保真度。量子自编码器使用 200 对数据进行训练，并使用 1000 个含噪样本进行测试。我们将经过 100 次同时扰动随机逼近迭代后提前停止的 VQE 态视为含噪数据。我们采用包含 $L = 3$ 个试探态模块的 QAE[4,1,4]，其中每个量子比特上的参数化旋转门为泡利 Y 旋转，纠缠门为非参数化受控 Z 门。

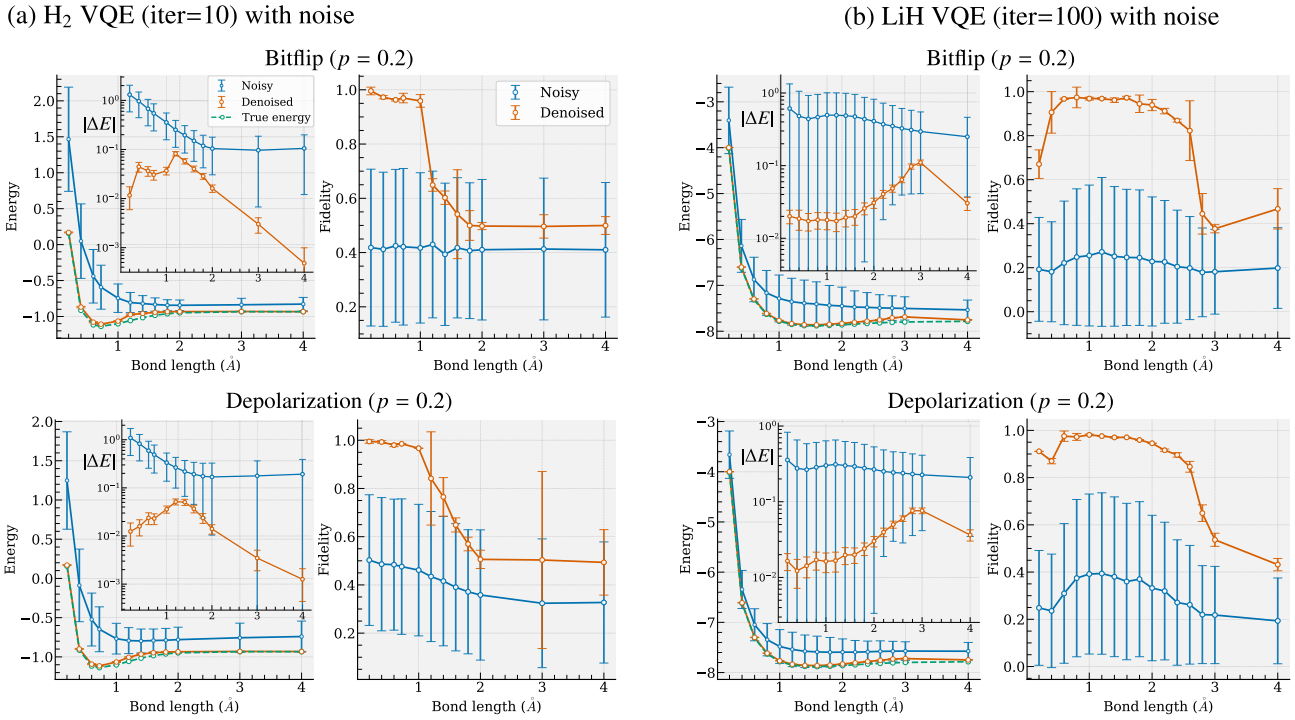

图 S4. H_2 和 LiH 在各个键长下，含噪 VQE 数据和去噪数据的能量、与基态能量的能量差（插图）以及与基态的保真度。这里，含噪 VQE 数据由**提前停止同时扰动随机逼近优化，并施加比特翻转或退极化噪声等其他噪声**生成。对变分量子本征求解器优化中经过 10 次（对于 H_2）和 100 次（对于 LiH）同时扰动随机逼近迭代的 VQE 数据，施加比特翻转或退极化信道，噪声强度为 0.2。对于 QAE[2,1,2]（对于 H_2）和 QAE[4,1,4]（对于 LiH），我们使用一个试探态模块，其中每个量子比特上的参数化旋转门为泡利 Y 门的旋转，纠缠门为非参数化 CZ 门。实线和误差条分别表示 1000 个测试样本的平均值和标准差。

因此原始含噪数据的能量误差增大（蓝色曲线）。有趣的是，在试探态电路的噪声水平处于一定容限范围内时，VQE 可以表现出降噪效果。

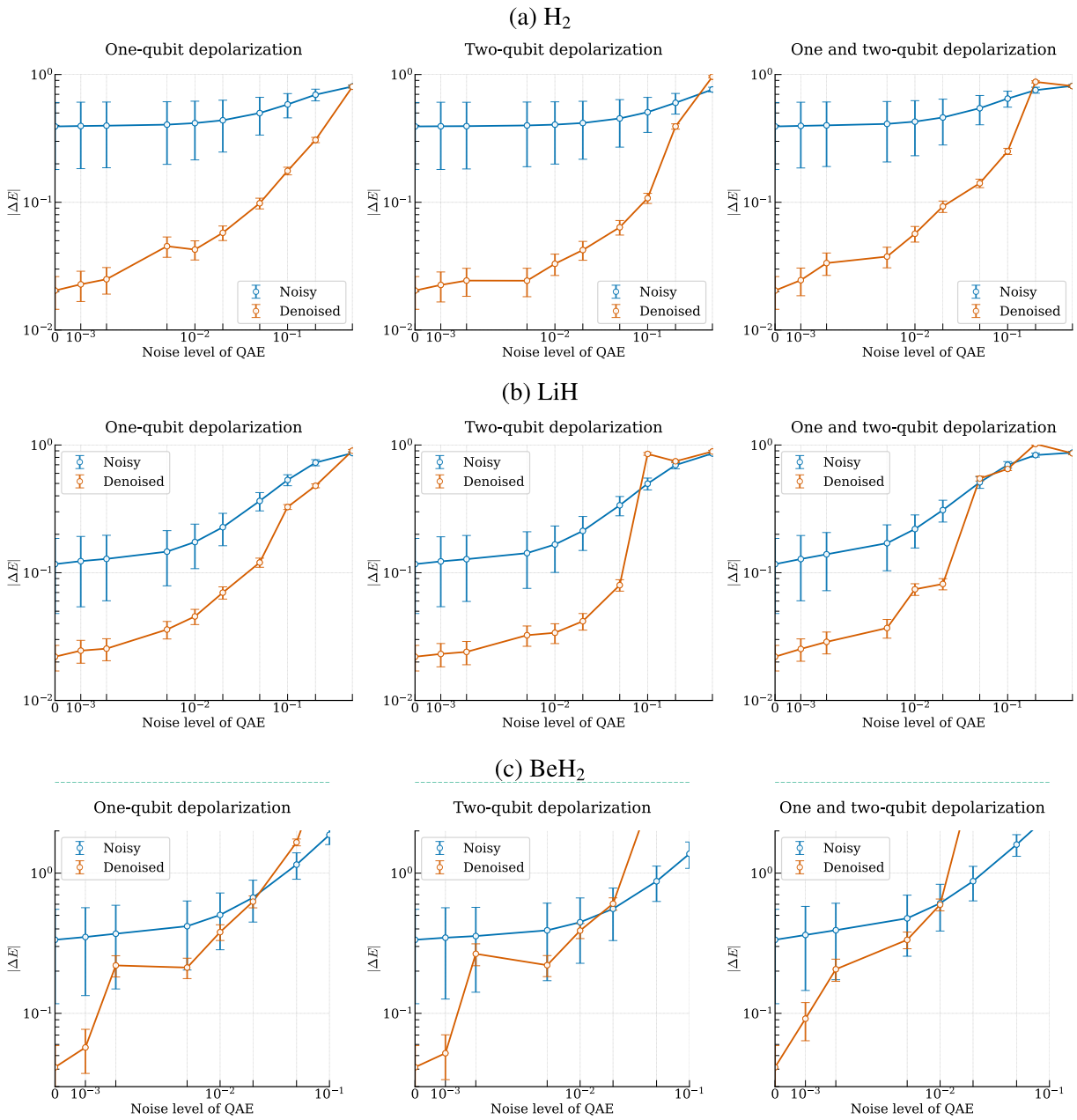

图 S5. 使用量子自编码器（橙色线）获得的输出态在寻找以下分子的基态能量时，其能量误差随量子自编码器量子电路中每种噪声模式的退极化错误率（噪声水平）变化的情况：(a) H_2（键长为 0.725$\mathring{A}$）；(b) LiH（键长为 1.595$\mathring{A}$）；以及 (c) BeH_2（键长为 1.6$\mathring{A}$）。用于评估输出态能量的量子电路在与量子自编码器相同的噪声条件下进行模拟，因此原始含噪数据中的能量误差也随之增大（蓝色线）。

C. 不同数据噪声水平下量子自编码器的性能

我们研究量子自编码器能否在误差较大时保持其性能。在第一个实验中，我们改变数据中的噪声强度，其中退极化噪声强度为 $p_d \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0\}$，比特翻转噪声强度为 $p_b \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$。此处，退极化信道通过以相应概率 $\{1-3p_d/4, p_d/4, p_d/4, p_d/4\}$ 对每个量子比特施加一个泡利矩阵 $\{I, X, Y, Z\}$ 来实现；比特翻转信道则通过以相应概率 $\{1-p_b, p_b\}$ 对每个量子比特施加 $\{I, X\}$ 来实现。

图 S6 展示了量子自编码器对不同噪声类型（退极化和比特翻转）及噪声强度的鲁棒性。实线表示平均能量误差（相对于基态能量）或保真度

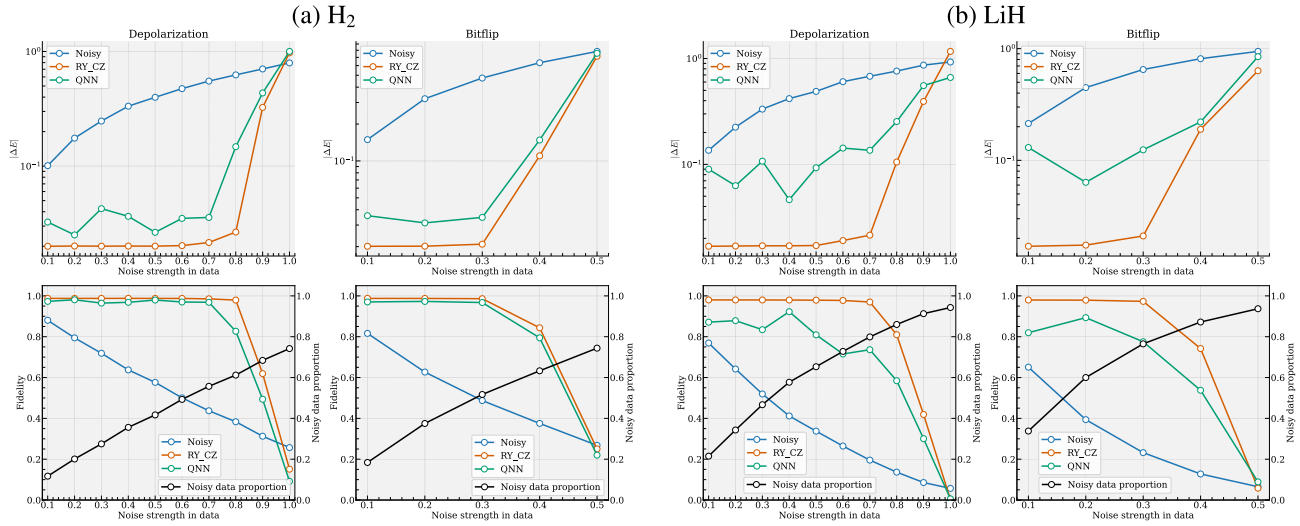

图 S6. 对于 (a) H_2（键长为 $0.725\mathring{A}$）和 (b) LiH（键长为 $1.4\mathring{A}$）的变分量子本征求解器最终结果，量子自编码器对不同噪声类型（退极化和比特翻转）及噪声强度的鲁棒性。实线表示平均能量误差或与基态的保真度。量子自编码器由一个试探态模块构成，其中旋转门 RY 作用于所有量子比特，而两个量子比特之间的受控 Z 门则按照量子比特的索引循环排列（记作 RY CZ）。标签 QNN 表示此前研究中量子自编码器的实现 [1, 3]。对于每次随机初始化，我们使用 200 对含噪数据训练两种结构 100 个训练轮次，并用 1000 个含噪数据样本进行测试。预测的能量误差和保真度均取这 1000 个样本的平均值，并对随机初始化重复十次后再求平均。在保真度图中，黑色实线表示训练数据集中“噪声极大的数据”的平均占比。此处，“噪声极大的数据”指与基态的保真度低于 0.2 的数据。

。 °与 H 的基态$_2$（键长为 $0.725A$）以及 LiH（键长为 $1.4A$）。在这里，我们分别为 H 和 LiH 构建了量子自编码器[2,1,2] 和量子自编码器[4,1,4]，采用 $L=1$ 试探态模块用于 H_2 和 LiH。每个试探态中的旋转门 R均为参数化泡利 Y 旋转门，双量子比特纠缠门则为受控 Z 门，并按照量子比特的索引以环形方式排列。我们在图中将此试探态称为 RY CZ。此外，我们还将量子自编码器的设计与先前研究中的量子自编码器实现进行了比较 [1, 3], ；后者对量子自编码器的量子神经网络表示中的每个酉算子的所有可能取值进行优化（见第 IA节）。我们将该实现称为量子神经网络，并采用相同的优化设置来优化 RY CZ。对于每次随机初始化，_我们都使用随机抽取的 200 对含噪数据训练这两种结构 100 个训练轮次，并使用 1000 个含噪数据样本进行测试。对这 1000 个样本的预测能量误差和预测保真度取平均值。此外，我们重复随机初始化十次，并取平均值在图中展示。在保真度图中，我们还绘制了训练数据集中“噪声极大数据”的平均占比（黑色实线）。这里，我们将“噪声极大数据”定义为与基态的保真度低于 0.2 的数据。

图 S6 表明，即使数据中的噪声水平相对较高（包括训练数据集中含噪数据占比较大），我们的方法依然表现良好。此外，该方法优于此前的量子神经网络（QNN）实现，减少了训练参数的数量，使优化更加容易。

IV. 横场伊辛模型的变分量子本征求解器去噪

我们展示了在变分量子本征求解算法中进行去噪的实验，以求解一维横场伊辛模型（TFIM）的基态。这里，一维横场伊辛模型的哈密顿量定义为

$$H = -\sum_i J_i \sigma_i^z \sigma_{i+1}^z - \sum_i g_i \sigma_i^x, \tag{S14}$$

其中，i 表示直线上的格点。该模型假设自旋为$\frac{1}{2}$ 的粒子，其泡利矩阵 σ_i^x 和 σ_i^z 作用于格点i。 J_i 是能量量纲的常数， g_i 是耦合参数，用于确定外部横场与最近邻相互作用之间的相对强度。

在数值模拟中，我们考虑均匀相互作用 $J_i = 1$ 和均匀格点势 $g_i = g > 0$，并采用开放边界条件（即对于 N个自旋，$H = -\sum_{i=1}^{N-1} \sigma_i^z \sigma_{i+1}^z - g\sum_{i=1}^{N} \sigma_i^x$）。该模型有两个相（有序

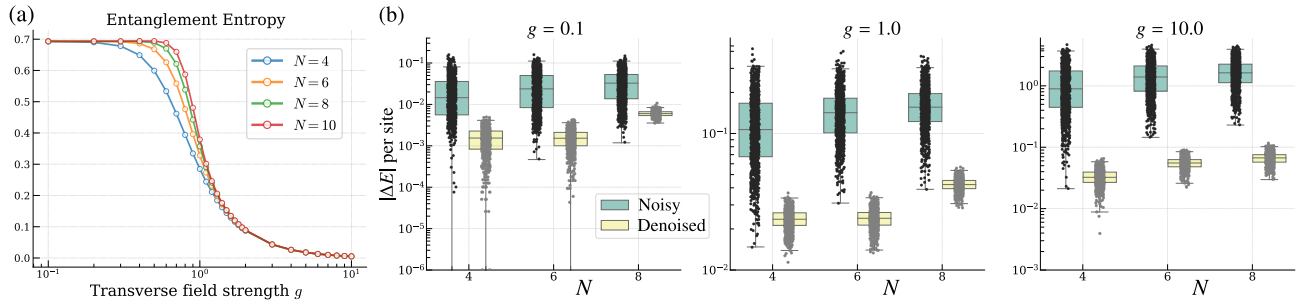

图 S7。（a）开放边界条件下，一维横场伊辛模型（TFIM）基态在不同横场强度 g 下的纠缠熵。（b）箱线图和抖动图展示了在 $N=4$、6、8 个自旋的一维横场伊辛模型中，1000 个测试数据的能量误差 $|\Delta E|$（估计能量与基态能量之差）的分布。箱体范围为四分位距，须线标示第 5 和第 95 百分位数。箱体上的横线表示中位数。

相和无序相），具体取决于基态是否破坏或保持 $\prod_j \sigma_j^x$ 自旋翻转对称性。当 $g=1$时，系统会发生量子相变。

为了理解一维横场伊辛模型的纠缠行为，我们计算了左侧包含 $N/2$ 个自旋的子系统与右侧其余 $N/2$ 个自旋之间的纠缠熵。图 S7（a）展示了这些纠缠熵随横场强度 g 的变化。为了证明我们的方法能够处理复杂度更高的非平凡态，我们使用量子自编码器对变分量子本征求解器的输出进行去噪，以寻找 $g=0.1$、1。0、和 10。0 处的基态。这里，含噪数据由提前停止优化时经过 $4\times N$ 次同时扰动随机逼近迭代的变分量子本征求解器电路生成。与其他实验类似，我们构建量子自编码器$[N,1,N]$，其中包含 $L=3$ 个试探态模块，旋转门为参数化泡利 Y 旋转门 $R_Y(\theta)$，$=e^{-iY\theta/2}$ 非参数化双量子比特纠缠门为受控 Z 门。我们仅使用 100 对含噪变分量子本征求解器数据训练量子自编码器，并使用 1000 个含噪变分量子本征求解器数据样本进行测试。

图 S7 (b) 展示了 $|\Delta E|$ 时，一维横场伊辛模型在 $N=4$、6 和 8 个自旋的 1000 个测试数据中能量误差（估计能量与基态能量之差）的分布。虽然增大 N会使问题更难，但对于所有 N 和 g的取值，我们的方法都能将误差降低约一个数量级。

我们进一步研究了 QAE 对不同噪声类型（退极化和相位翻转）及噪声强度的鲁棒性，这些噪声作用于横场长度为 $g=1.0$时的 (a) TFIM-4 和 (b) TFIM-6 的基态（图 S8）。退极化信道通过以相应概率 $\{I,X,Y,Z\}$ 对每个量子比特施加一个泡利矩阵 $\{1-3p_d/4,p_d/4,p_d/4,p_d/4\}$来实现，而相位翻转信道则通过以相应概率 $\{I,Z\}$ 对每个量子比特施加 $\{1-p_{\text{phase}},p_{\text{phase}}\}$来实现。我们改变数据中的噪声强度，噪声强度为 p_d、$p_{\text{phase}}\in\{0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,1.0\}$。实线表示平均能量误差（相对于基态能量）或与基态的保真度。这里，我们为 TFIM-4 和 TFIM-6 分别构造了具有 [4,1,4] 和 [6,1,6] 个 $L=1$ 试探态模块的 QAE R。在每个试探态中，旋转门 R 是参数化泡利 Y 旋转门，双量子比特纠缠门是按量子比特索引呈环形排列的受控 Z 门。我们在图中将此试探态记作 RY CZ。我们使用随机抽取的 200 对含噪数据训练这两种结构 100 轮，并使用 1000 个含噪数据样本进行测试。预测的能量误差和预测的保真度均取这 1000 个样本的平均值。在保真度图中，我们还绘制了训练数据集中“噪声极大数据”的平均占比（黑色实线）。这里，我们将“噪声极大数据”定义为与基态保真度低于 0.2 的数据。

与图 S6、图 S8 表明，即使数据中的噪声水平相对较高（包括训练数据集中含噪数据占比较大），我们的方法依然能够取得良好表现。

V. 量子自编码器试探态的设计

我们算法中采用的试探态结构依赖于硬件高效拟设。

$$U_j^i(\boldsymbol{\theta})=\prod_{l=1}^{L}\left[R_{\text{loc}}^{(l)}(\boldsymbol{\theta}_l)V_{\text{ent}}\right]R_{\text{loc}}^{(L+1)}(\boldsymbol{\theta}_{L+1}) \tag{S15}$$

这里，每个酉算子 U_j^i 由 L 个模块以及作用于每个量子比特的最终单量子比特旋转组成。每个模块都包含局部单量子比特旋转 $R^{(l)}{}_{\text{loc}}(\boldsymbol{\theta}_l)=\otimes_k R(\theta_{k,l})$，以及双量子比特纠缠门。在我们的实验中，我们使用

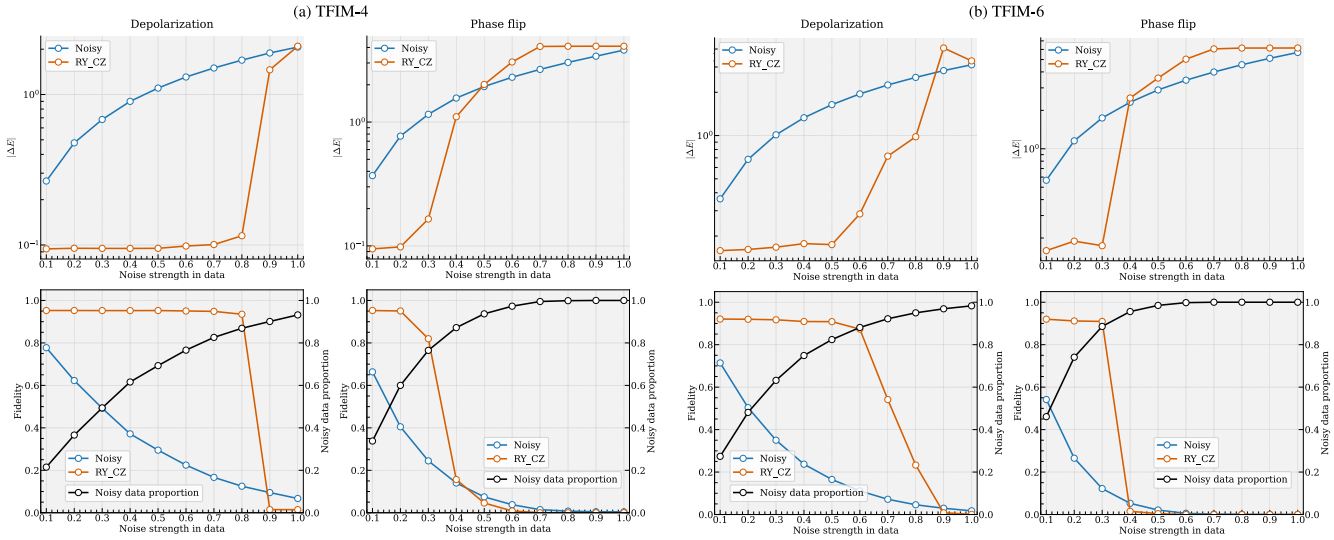

图 S8. 在横场长度为 $g = 1.0$ 时，量子自编码器对施加于 (a) TFIM-4 和 (b) TFIM-6 基态的不同噪声类型（退极化和相位翻转）及不同噪声强度的鲁棒性。实线表示平均能量误差或与基态的保真度。量子自编码器[N, 1, N] (N=4,6)由一个试探态模块构成，其中旋转门 RY 作用于所有量子比特，两个量子比特的受控 Z 门按量子比特索引循环排列（记为 RY CZ）。我们使用 200 对随机含噪数据训练这两种结构，共训练 100 轮，并使用 1000 个含噪数据样本进行测试。预测的能量误差和保真度均取这 1000 个样本的平均值。在保真度图中，黑色实线表示训练数据集中“高噪声数据”的平均占比。此处，“高噪声数据”定义为与基态的保真度低于 0.2 的数据。

试探态 RY CZ，其中单量子比特旋转门 R 是参数化旋转门 $R_Y(\theta) = e^{-iY\theta/2}$，即泡利 Y矩阵；纠缠门 V_{ent} 由受控相位门 $CZ(k, k+1)$ 组成，这些门按量子比特索引k 以环形方式排列。

我们可以提供一种选择试探态的方法。例如，在本研究中，我们考察了多种试探态结构，用少量数据对它们进行训练，并选取性能最佳的试探态。我们对几种常见的试探态类型进行了数值研究，以支持在本研究中选择 RY CZ 试探态。在数值实验中，我们考察了六种试探态模式：RY CZ、RY CX、RYZ CZ、RY CAN、RYZ CAN 和 RZYY CAN。此处，参数化旋转门由局部单量子比特旋转中的泡利 Y 和 Pauli Z矩阵构成，而纠缠门 V_{ent} 则为 CX、CZ 或双量子比特规范（CAN）门。双量子比特规范门定义为三个双量子比特门的复合

$$\text{CAN} = \text{RXX}(\pi)\text{RYY}(\pi)\text{RZZ}(\pi) = e^{-i\frac{\pi}{2}X\otimes X}e^{-i\frac{\pi}{2}Y\otimes Y}e^{-i\frac{\pi}{2}Z\otimes Z}. \tag{S16}$$

值得注意的是，最后一种模式 RZYY CAN 可用于实现任意双量子比特幺正操作。

我们将这些试探态模式用于 LiH 的去噪任务（键长为 1。595A 且进行 100 次 VQE 迭代）、BeH_2 （键长为 1。6A 且进行 500 次 VQE 迭代），以及 $g = 1.0$ 且 N $= 4$、6 和 8 个量子比特的一维横场伊辛模型，进行 $10 \times N$ 次 VQE 迭代。图 S9 展示了 1000 个测试数据点的能量误差 $|\Delta E|$ 和基态保真度分布，以证明 RY CZ 是最适合我们试探态的选择。此外，对于 LiH 和 BeH 分子，$_2$, 虽然 VQE试探态中使用了 CX 门，但它在去噪任务中无法取得良好结果。因此，量子自编码器的试探态不必与 VQE 或通用试探态相同。

我们进一步研究了增加 L 式（S15）所述试探态模块数量时的去噪性能。图S10 展示了 LiH 和 BeH_2 分子在不同试探态模块数量 L 以及原始训练数据中不同 VQE 迭代次数下的性能，其中较少的迭代次数对应噪声更大的数据。我们还可以研究采用不同量子自编码器配置时的去噪性能。图S11 展示了采用不同量子自编码器拓扑结构，以及在 VQE 流程中采用不同数量的同时扰动随机逼近迭代来生成原始训练数据时的性能，其中较少的迭代次数对应噪声更大的数据。增加量子自编码器拓扑结构的复杂度和试探态模块的数量，可以提升量子自编码器的表达能力，但参数数量过多会使训练更加困难。在我们的任务中，我们将量子自编码器的拓扑结构设为 $[N, 1, N]$，并设置一个瓶颈节点； $L = 1$ 或 $L = 3$ 已足以观察去噪性能的影响。不过，如果要处理数量更多的参数 L 或更复杂的量子自编码器拓扑结构，我们就必须改进优化器或谨慎选择初始化方式。

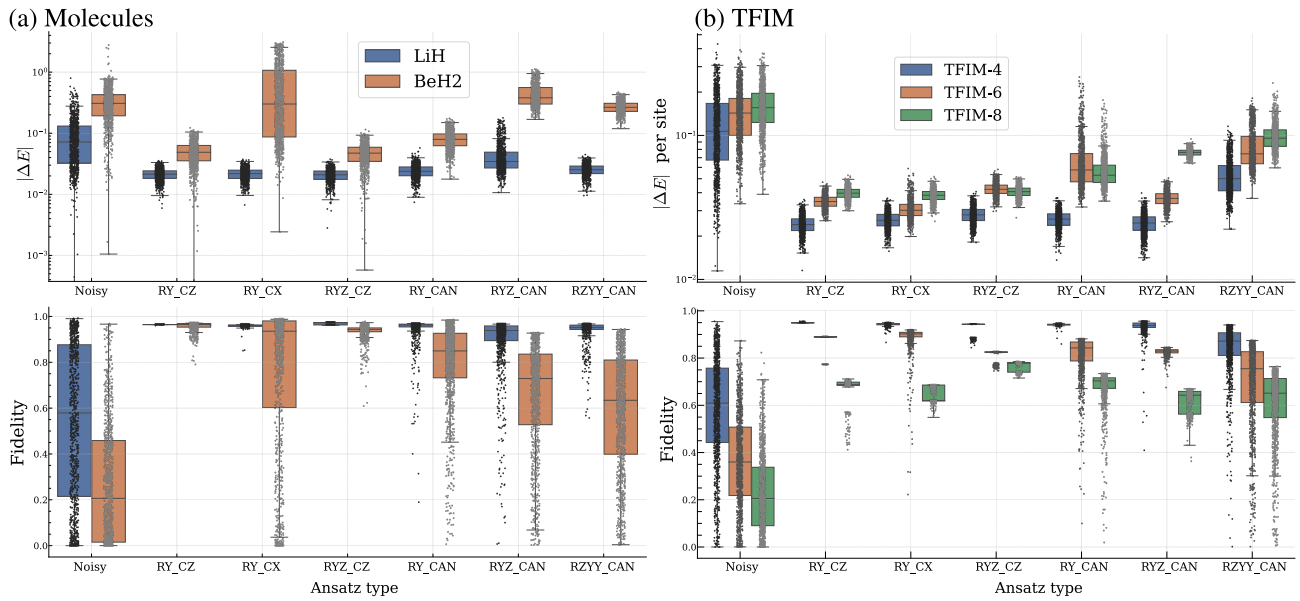

图 S9. 箱线图和抖动图展示了 (a) LiH（键长为 1.595$\mathring{A}$，VQE 流程中进行了 100 次同时扰动随机逼近迭代）和 BeH_2（键长为 1.6$\mathring{A}$，VQE 流程中进行了 500 次同时扰动随机逼近迭代）分子，以及 (b) $g = 1.0$、N=4、6、8 个量子比特的一维横场伊辛模型在 VQE 流程中进行 $10 \times N$ 次同时扰动随机逼近迭代时，采用不同量子自编码器试探态类型所得的 1000 个测试数据样本的能量误差 $|\Delta E|$ 和基态保真度分布。四分位距位于箱体内，第 5 和第 95 百分位数由须线标示。箱体上的横线表示中位数。

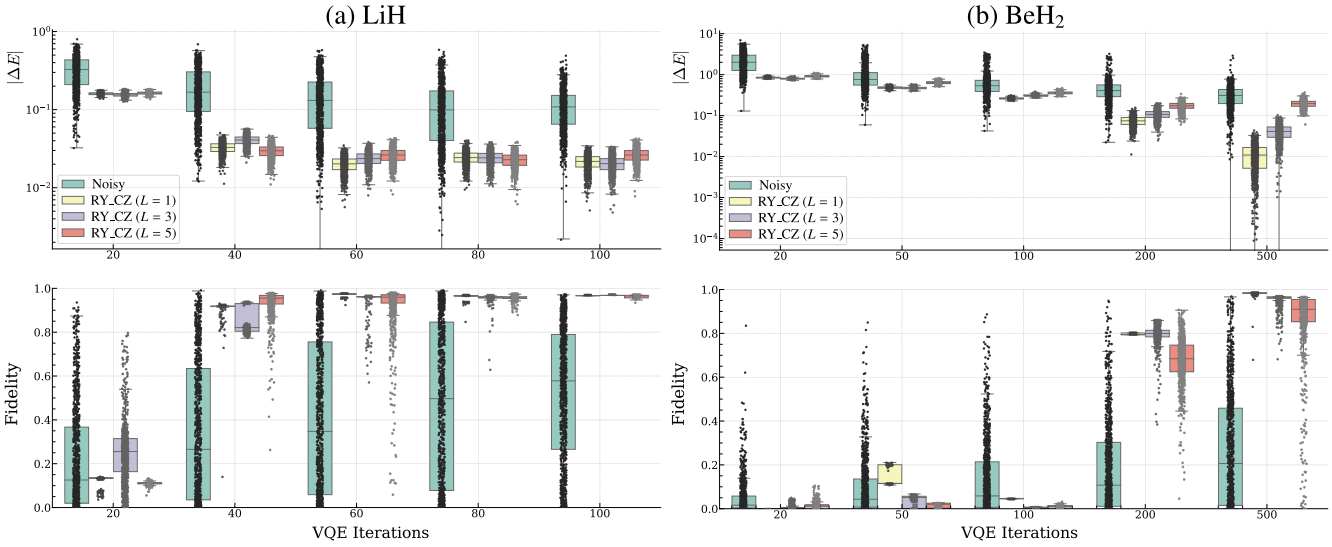

图 S10. 箱线图和抖动图展示了 (a) LiH（键长为 1.595$\mathring{A}$）和 (b) BeH_2（键长为 1.6$\mathring{A}$）分子的 1000 个测试数据样本的能量误差 $|\Delta E|$ 和基态保真度分布。这些图展示了在创建原始训练数据时，采用不同试探态模块数量 L 和不同数量的 VQE 流程中同时扰动随机逼近迭代时的性能；迭代次数较少对应的数据噪声较大。四分位距位于箱体内，第 5 和第 95 百分位数由须线标示。箱体上的横线表示中位数。

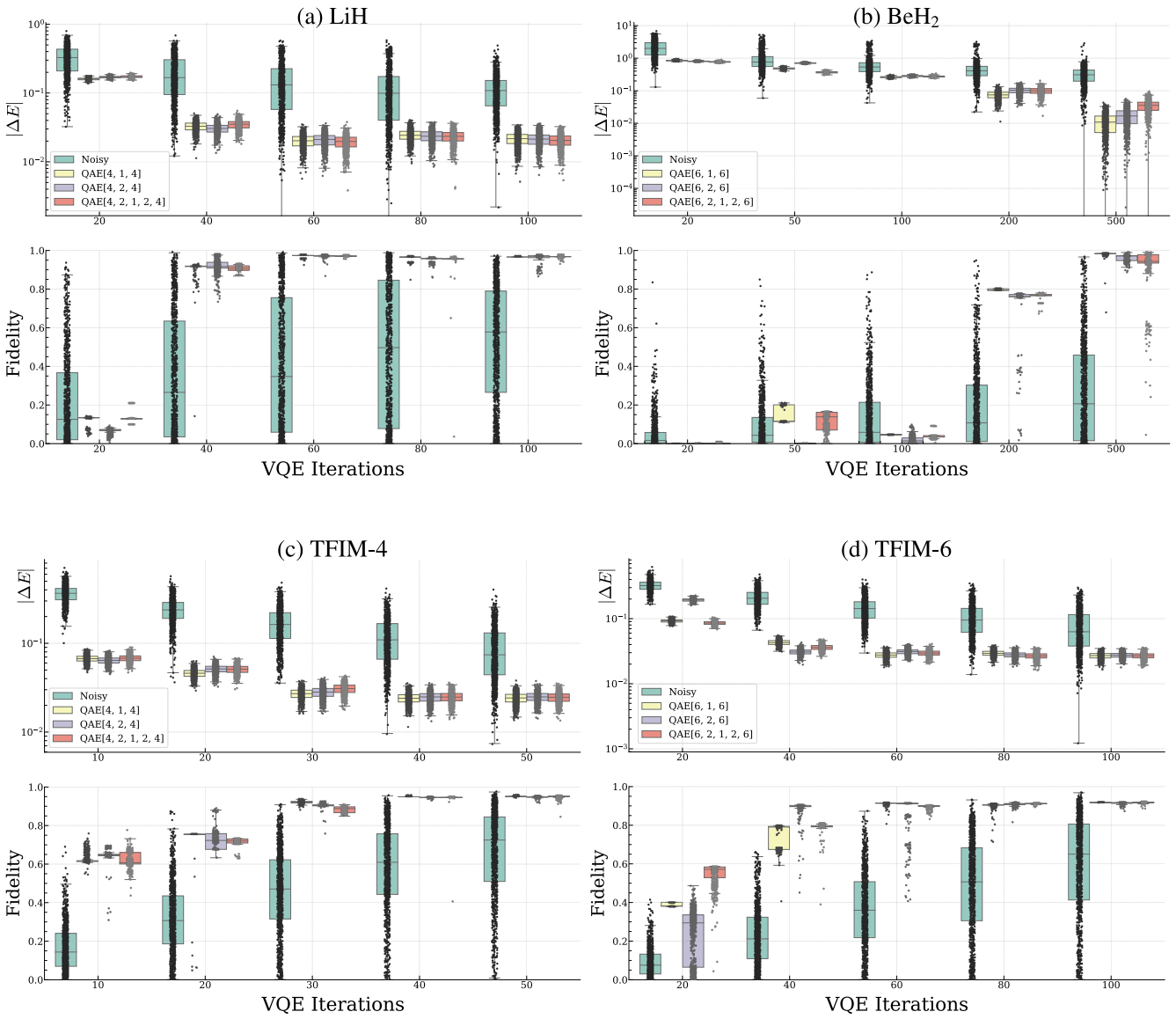

图 S11。箱线图和抖动图展示了以下对象在 1000 个测试数据样本上的能量误差 $|\Delta E|$ 和基态保真度分布：(a) LiH（键长为 1.595$\mathring{A}$）和 (b) BeH_2（键长为 1.6$\mathring{A}$）分子，以及在 $g = 1.0$ 下的一维横场伊辛模型：(c) N=4 量子比特（TFIM-4）和 (d) N=6 量子比特（TFIM-6）。这些图展示了在 VQE 流程中使用不同量子近似演化（QAE）拓扑结构和同时扰动随机逼近迭代来生成原始训练数据时的性能，其中较短的迭代对应噪声更大的数据。箱体包含四分位距，须线标示第 5 和第 95 百分位数。中位数由横贯箱体的线表示。

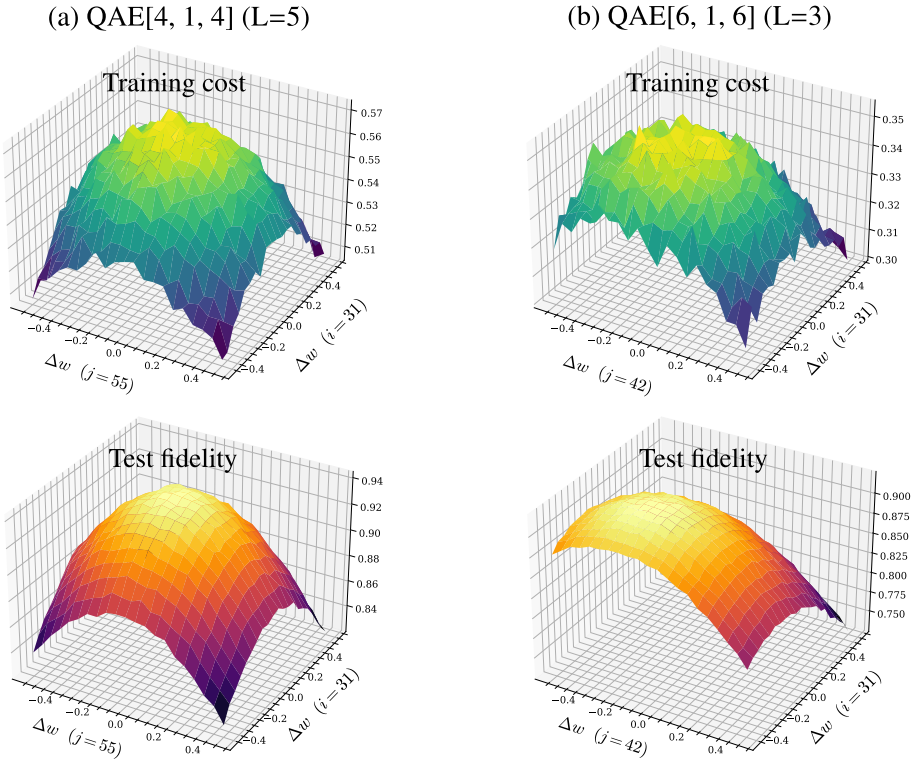

图 S12. 分别采用 $L=5$ 个和 $L=3$ 个试探态模块时，QAE$[4,1,4]$ 和 QAE$[6,1,6]$ 的训练代价与测试保真度地形图。我们固定其他参数，同时改变 $\theta_i \in \Omega_i$ 和 $\theta_j \in \Omega_j$，以绘制训练代价与测试保真度。这里，$\Omega_k = [\theta_k^t - \Delta w, \theta_k^t + \Delta w]$（其中 $\Delta w = 0.5$）表示参数 θ_k 在训练值 θ_k^t 附近的区间。

VI. 训练量子自编码器中的贫瘠高原

原则上，本研究采用的量子电路（RY CZ）和全局函数无法完全消除贫瘠高原，尤其是在扩大问题规模时。不过，我们的研究发现了缓解贫瘠高原问题的实用技术。首先，我们使用一小组训练数据训练量子自编码器，为参数建立一个相对较好的初始值。随后，在训练过程中，我们在小批量训练中结合使用同时扰动随机逼近（SPSA）与 AMSGrad，并引入随机参数更新，以增强近似梯度的变化。

我们预计，使用含噪数据进行训练能够缓解系统扩展时遇到的困难。例如，在图 S12中，我们展示了量子自编码器$[4,1,4]$ 和量子自编码器$[6,1,6]$ 在耗散型量子自编码器的每个酉算子中分别含有 $L=5$和 $L=3$个试探态模块时的训练代价和测试保真度景观。这里，我们首先使用横场伊辛模型训练量子自编码器，然后选取每个参数 θ_k，使其在训练后的取值 θ_k^t 附近变化，记为 $\Omega_k = [\theta_k^t - \Delta w, \theta_k^t + \Delta w](\Delta w = 0.5)$，同时固定其他参数。我们选择两个索引 i 和 j，使得训练代价在 Ω_i 和 Ω_j上的变化最大。最后，我们固定其他参数，仅改变 $\theta_i \in \Omega_i$ 和 $\theta_j \in \Omega_j$，以绘制图 S12中的训练代价和测试保真度。即使参数数量相对较多（78 和 76），我们仍能观察到，在当前规模下，训练景观并未出现贫瘠高原。不过，我们承认，这并不意味着更大规模系统中不会出现贫瘠高原。未来研究可以探索结合训练机制与数据特性来设计量子自编码器，从而规避贫瘠高原，这是一个值得关注的方向。

VII. 去噪机制

在正文中，我们讨论了对量子自编码器去噪机制的理解。本节将进一步展示更多结果。

我们研究量子自编码器各子系统的输出与目标基态中对应部分之间的保真度。为说明这一点，我们以量子自编码器[6,1,6]为例。首先，对整个量子自编码器结构 [6,1,6]（而不仅仅是

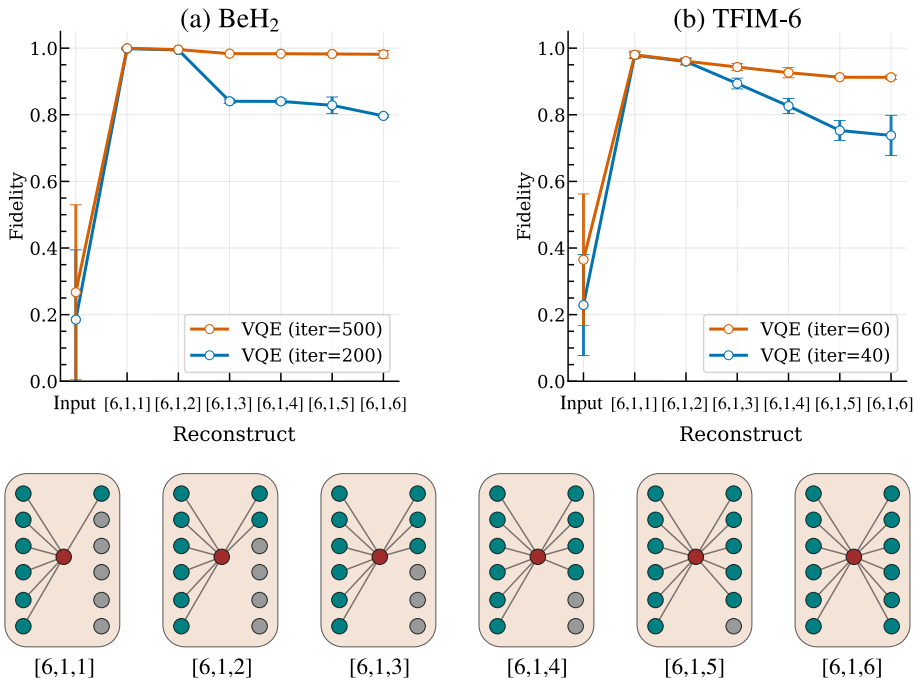

图 S13. QAE[6,1,6] 的子系统输出与 BeH_2 和 TFIM-6 基态中对应部分之间的保真度。这里，含噪数据由不同 SPSA 迭代下的 VQE 电路生成（迭代次数越少，噪声水平越高）。灰色节点表示非活跃神经元，其中子系统 [6,1,1], [6,1,2], [6,1,3], [6,1,4], 和 [6,1,5] 分别对应基态中的第一个、第一个和第二个、第一个至第三个、第一个至第四个，以及第一个至第五个量子比特构成的子系统。随着噪声水平升高，含噪成分从瓶颈层泄漏出来，导致解码器无法重构基态。实线和误差条表示基于 1000 个测试样本计算的平均值和标准差。

一个子空间）进行训练。然后，我们使用含噪数据测试训练好的量子自编码器，即在 1000 个含噪数据样本上计算量子自编码器[6,1,6]的输出与真实基态之间的保真度。

在这一测试阶段，我们采用了一种独特的方法，即使用非活跃神经元。非活跃神经元是指在训练好的 QAE 中连接已被移除的神经元。例如，如果我们将最后一层中的第 6 个神经元设为非活跃，就相当于得到一个 QAE[6,1,5]，其参数是从 QAE[6,1,6]的已训练参数中选取的一个子集。该 QAE 的输出包含五个量子比特，随后将其与目标基态中前五个量子比特的状态进行比较。当我们将最后一层中的第 5 个和第 6 个神经元设为非活跃时，也采用类似的流程，从而得到 QAE[6,1,4]（并通过停用相应的神经元组，继续得到 QAE[6,1,3], QAE[6,1,2], 和 QAE[6,1,1]），然后将该 QAE 的输出与目标基态中的前四个量子比特（或前 3 个、前 2 个及第 1 个量子比特）进行比较。

本研究揭示了量子自编码器如何处理不同噪声水平下的含噪数据。例如，在图 S13(a) 中，我们考察了由 BeH_2 的变分量子本征求解器电路生成的含噪数据（键长为 $1.6A$），这些数据是在 500 次（低噪声）和 200 次（高噪声）同时扰动随机逼近迭代后生成的。当输入态经过量子自编码器处理后，目标态的一部分会在子系统的输出中重构出来。

在噪声水平较低时，例如，VQE 输出态在经过 500 次 SPSA 迭代后生成的情况下，对于子系统 [6,1,1], [6,1,2], [6,1,3], [6,1,4], 和 [6,1,5], QAE 能够以较高保真度重构目标态的第一个量子比特、前两个量子比特、前三个量子比特、前四个量子比特、前五个量子比特，最终重构全部量子比特（由橙色线表示）。然而，如果我们使用噪声水平较高的 VQE 数据训练 QAE，而此时 VQE 输出态（经过 200 次 SPSA 迭代后）与基态的重叠度较低，那么与基态相关的无关信息就会在瓶颈处丢失，从而阻碍解码器重构基态（由蓝色线表示）。如图S13（b）所示，在 TFIM-6 实验中也进一步证实了类似的观察结果。我们考察了由 TFIM-6 的 VQE 电路（在 $g = 1.0$）生成的含噪数据，分别对应 50 次（低噪声）和 40 次（高噪声）SPSA 迭代。

[1] D. Bondarenko 和 P. Feldmann, 《量子自编码器对量子数据进行去噪》, 《物理评论快报》124, 130502（2020年）。[2] Qiskit：用于量子计算的开源框架（2021年）。[3] T. Achache、L. Horesh 和 J. Smolin, 《使用量子自编码器对量子态去噪——理论与应用》, arXiv 预印本：2012.14714（2020年）。[4] J. C. Spall, 《随机优化与同时扰动法》, 载于 *Proceedings of the 31st Conference onWinter Simulation: Simulation—a Bridge to the Future - Volume 1*, WSC '99（计算机协会, 美国纽约州纽约市, 1999年）, 第 101–109 页。[5] S. J. Reddi、S. Kale 和 S. Kumar, 《论 Adam 及其后续方法的收敛性》, 载于 *Proceedings of the 6th International Conference on Learning Representations ICLR 2018* 2018 ( )。[6] A. Kandala、A. Mezzacapo、K. Temme、M. Takita、M. Brink、J. M. Chow 和 J. M. Gambetta, 《面向小分子和量子磁体的硬件高效变分量子本征求解器》, 《自然》549, 242（2017年）。[7] IBM 量子挑战赛 2021, https://github.com/qiskit-community/ibm-quantum-challenge-2021（2021年）, 访问日期：2022年12月20日。[8] I. O. Sokolov、P. K. Barkoutsos、P. J. Ollitrault、D. Greenberg、J. Rice、M. Pistoia 和 I. Tavernelli, 《强关联区域中的量子轨道优化酉耦合簇方法：量子算法能否胜过其经典对应方法？》, 《化学物理杂志》152, 124107（2020年）。[9] P. K. Barkoutsos、J. F. Gonthier、I. Sokolov、N. Moll、G. Salis、A. Fuhrer、M. Ganzhorn、D. J. Egger、M. Troyer、A. Mezzacapo、S. Filipp 和 I. Tavernelli, 《电子结构计算的量子算法：粒子-空穴哈密顿量与优化波函数展开》, 《物理评论A辑》98, 022322（2018年）。