arXiv:2304.04167v1 [quant-ph] 9 Apr 2023

利用有限数据集进行神经网络辅助的量子态与量子过程层析

Akshay Gaikwad[#], [1], [*] Omkar Bihani[#], [1], [†] Arvind, [1, 2], [‡] 和 Kavita Dorai[1], [§]

[1] *Department of Physical Sciences, Indian Institute of Science Education & Research Mohali, Sector 81 SAS Nagar, Manauli PO 140306 Punjab India.* [2] *Vice Chancellor, Punjabi University Patiala, 147002, Punjab, India*

本研究采用前馈人工神经网络（FFNN）架构，对含噪实验数据中的量子态和量子过程进行层析。为评估 FFNN 的性能，我们使用大幅缩减的数据集，并表明能够以高保真度重建未知量子态和量子过程的密度矩阵与过程矩阵。我们使用 FFNN 模型对在核磁共振（NMR）量子处理器上实验生成的 100 个双量子比特态和 128 个三量子比特态进行层析。此外，FFNN 模型还用于表征不同的量子过程，包括双量子比特纠缠门、整形脉冲场梯度、核磁共振系统中存在的本征退相干过程，以及各种双量子比特噪声信道（关联比特翻转、关联相位翻转和比特与相位联合翻转）。我们将 FFNN 模型得到的结果与标准量子态和量子过程层析方法进行比较，计算所得的保真度表明，在所有情况下，FFNN 模型的层析表现均优于标准方法。

I. 引言

量子态层析（QST）和量子过程层析（QPT）分别是表征未知量子态和过程以及评估量子器件质量的关键技术 [1–3]。人们已经设计出许多计算和实验效率高的量子态层析和量子过程层析算法，例如自引导层析[4]，自适应层析 [5],基于压缩感知的量子态层析和量子过程层析协议，这些协议使用大幅减少的数据集 [6, 7], 选择性量子过程层析 [8–10], 以及利用弱测量进行直接量子态层析/量子过程层析 [11]。

最近，机器学习（ML）技术已被用于提高层析协议的效率 [12–14]。研究人员使用基于受限玻尔兹曼机的人工神经网络（ANN）模型，对纠缠量子态进行了量子态层析（QST） [15] ，并在光学系统上进行了实验实现 [16]。研究人员还进行了基于机器学习的自适应量子态层析，该方法能够适应实验并提出合适的后续测量方案 [17]。基于注意力机制的生成网络量子态层析已在 IBMQ 量子计算机上通过实验实现[18]。在重建光子量子数据集中的量子态时，通过将态制备和测量误差降至最低，研究人员实现了人工神经网络增强的量子态层析 [19]。研究人员采用卷积人工神经网络模型，在存在模拟噪声的情况下，利用层析测量重建量子态 [20]。基于人工神经网络的局域测量量子态层析已在核磁共振（NMR）上通过实验展示[21]。机器学习被用于检测实验中的多体纠缠结构针对核磁共振纠缠态 [22]。人工神经网络用于进行量子态层析，同时考虑测量不完美性 [23]，并经过训练，可在无需任何关于量子态的先验信息的情况下唯一重建量子态 [24]。人工神经网络用于以高保真度重建编码在光子空间自由度中的量子态 [25]。机器学习方法用于直接估计制备的量子态的保真度 [26]。人工神经网络用于在存在各种类型噪声的情况下重建量子态 [27]。通过条件生成对抗网络对中等规模量子器件中的量子态层析进行了研究 [28]。

在本研究中，我们采用前馈神经网络（FFNN）架构来执行量子态以及过程层析。我们在计算生-成的量子态/过程上训练并测试了该模型，随后在核磁共振量子处理器上生成的含噪实验数据上对其进行了验证。此外，我们还在大幅缩减的数据集上测试了前馈神经网络模型的效能，其中使用了总数据集的一个随机子集。即使使用这一大幅缩减的数据集，前馈神经网络模型仍能够以高保真度重-构真实的量子态和量子过程。

本文的组织结构如下：第II 节简要介绍前馈神经网络模型在量子态层析和量子过程层析中的基本框架；第IIA 节介绍前馈神经网络架构，第IIB 节详细说明如何构建用于执行量子态层析和量子过程层析的前馈神经网络训练数据集。第III 节和IV节分别给出使用前馈神经网络对实验核磁共振数据执行量子态层析和量子过程层析的结果。第V 节总结了几点结论。

[*] ph16010@iisermohali.ac.in
[†] omkarbihani10@gmail.com
[‡] arvind@iisermohali.ac.in[§] kavita@iisermohali.ac.in[#] 这些作者对本研究贡献相同。

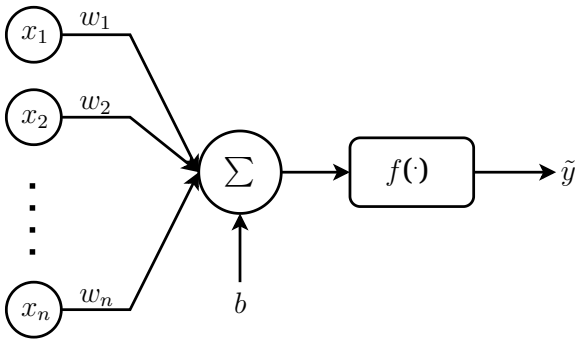

图1。（在线彩色版）人工神经网络模型的基本单元，其中 x_i 是输入， w_i 是权重， b 是偏置， $\sum$ 是求和函数， f 是激活函数， $\tilde{y}$ 是人工神经网络的输出。

II. 基于前馈神经网络的量子态层析和量子过程层析

A. 基本前馈神经网络架构

首先，我们介绍多层感知机模型，也称为前馈神经网络（FFNN），本文将其用于表征量子态和量子过程。人工神经网络是一种数学计算模型，其灵感来源于生物神经系统。生物神经系统由称为神经元的自适应单元构成，这些单元通过权重与其他神经元相连。当神经元的值大于称为偏置的“阈值”时，该神经元便会被激活。图 1 展示了人工神经网络的示意图，其中包含输入 n x_1、x_2、$\cdots$、x_n，这些输入通过权重 w_1、w_2、$\cdots$、w_n与一个神经元相连；这些输入的加权和与偏置 b 进行比较，并经由激活函数 f处理，得到输出$\tilde{y} = f(\sum_{i=1}^{n} w_i x_i - b)$。

多层前馈神经网络架构由三层组成：输入层、隐藏层和输出层。数据被馈送到输入层，随后传递到隐藏层，最后从最后一个隐藏层到达输出层。图 2 描绘了一个原型前馈神经网络模型的示意图，该模型包含一个输入层、两个隐藏层和一个输出层，并已被用作示例，使用大幅精简的数据集对实验双量子比特核磁共振量子态进行量子态层析。

数据分为两部分：训练数据集用于训练模型；在量子过程中，网络参数（权重和偏置）会根据结果进行更新；测试数据集则用于评估网络性能。考虑“m”个训练元素 $\{(\vec{x}^{(1)},\ \vec{y}^{(1)}),\ (\vec{x}^{(2)},\ \vec{y}^{(2)})\cdots,\ (\vec{x}^{(p)},\ \vec{y}^{(p)})\}$ 其中 $\vec{x}^{(i)}$ 是输入， $\vec{y}^{(i)}$ 是对应的输出。将这些输入馈入网络会产生输出 $[\tilde{\vec{y}}^{(1)}, \tilde{\vec{y}}^{(2)}, ..., \tilde{\vec{y}}^{(p)}]$。由于网络参数是随机初始化的，预测输出并不等于预期输出。训练该网络可以通过最小化均方误差代价函数实现。

，相对于网络参数，使用随机梯度下降法和反向传播算法 [29]:

$$w_{ij} \to w'_{ij} = w_{ij} - \frac{\eta}{p'} \sum_{i=1}^{p'} \frac{\partial}{\partial w_{ij}} \mathcal{L}(\vec{x}^{(i)}) \tag{1}$$

$$b_i \to b'_i = b_i - \frac{\eta}{p'} \sum_{i=1}^{p'} \frac{\partial}{\partial b_i} \mathcal{L}(\vec{x}^{(i)}) \tag{2}$$

其中 $\mathcal{L}(x^{(i)}) = ||\vec{y}^{(i)} - \tilde{\vec{y}}^{(i)}||^2$ 是随机选择的 m' 训练输入 $x^{(i)}$， η 是学习率，而 w'_{ij} 和 b'_i 分别是更新后的权重和偏置。

B. 用于 QST 和 QPT 的前馈神经网络训练数据集

一个 n-量子比特密度算符 ρ 可以在乘积基中表示为矩阵:

$$\rho = \sum_{i=0}^{3} \sum_{j=0}^{3} ... \sum_{n=0}^{3} a_{ij...n} \sigma_i \otimes \sigma_j \otimes ... \sigma_n \tag{3}$$

其中 $a_{00...0} = 1/2^n$、 σ_0表示 2×2单位矩阵， $\sigma_i, i = 1,2,3$ 表示单量子比特泡利矩阵。

量子态层析的目标是根据一组层析测量重构 ρ 。量子态层析的标准步骤包括求解形如 [30]的线性方程组:

$$\mathcal{A}\mathcal{X} = \mathcal{B} \tag{4}$$

其中， $\mathcal{A}$ 是一个固定的系数矩阵，仅取决于所选的测量设置； $\mathcal{X}$ 是一个列矩阵，其中包含需要重构的密度矩阵的元素，而输入向量 $\mathcal{B}$ 包含实际的实验数据。

前馈神经网络模型在一个包含随机生成的纯态和混合态的数据集上进行训练。为了生成这些系综，考虑一个均值为零、方差为单位的正态分布 $\mathcal{N}(\mu = 0, \sigma^2 = 1)$。计算基中的一个 n量子比特纯随机量子态表示为一个 2^n 列向量 C，其 i 个条目 c_i 按如下方式从随机分布中生成:

$$c_i = \frac{1}{N}(\mathfrak{d}_i + i\,\mathfrak{e}_i) \tag{5}$$

其中， $\mathfrak{d}_i, \mathfrak{e}_i$ 从分布 $\mathcal{N}$ 中随机选取， N 是一个归一化因子，用于确保 C 表示一个单位向量。

对于混合态:

$$R = \mathcal{D}_i + i\,\mathcal{E}_i \tag{6}$$

其中 R是一个 $2^n \times 2^n$矩阵，其元素 $\mathcal{D}_i, \mathcal{E}_i$从正态分布 $\mathcal{N}$ 中随- 机采样得到。使用

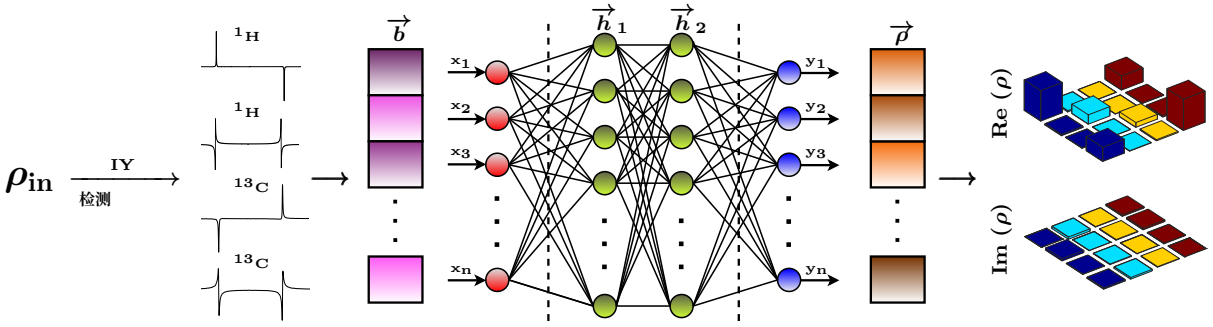

图2。（在线彩色）展示用于对在核磁共振量子系统上生成的双量子比特量子态执行量子态层析的前馈神经网络模型的流程图；左侧的 ρ_{in} 表示待进行层析的量子态；IY表示一种层析操作，随后进行信号检测，所示的一组核磁共振谱是在层析测量后获得的。接下来展示的是具有两个隐藏层的前馈神经网络模型，该模型随后使用缩减的数据集重构右侧所示的最终实验层析结果。

矩阵 R，构造相应的混合量子态密度矩阵 ρ_{mix}，如下所示：$\rho_{\mathrm{mix}} = \frac{RR^\dagger}{\mathrm{Tr}(RR)}$。

前馈神经网络在纯态和混合态上进行训练，并生成相应的密度矩阵。生成密度矩阵后 $\mathcal{X}_i$，使用方程（ $\mathcal{B}_i$ 计算相应的4）。随后使用训练元素 $\{\mathcal{B}_i, \mathcal{X}_i\}$ 训练图 2所示的前馈神经网络模型，其中 $\mathcal{B}_i$ 是前馈神经网络的输入，而 $\mathcal{X}_i$ 是相应的标记输出。

给定量子过程的量子过程层析通常使用克劳斯算符表示来执行，其中，对于固定的算符基组 $\{E_i\}$，作用于输入态 ρ_{in} 的量子映射 Λ 可写为 [31]:

$$\Lambda(\rho_{in}) = \sum_{m,n} \chi_{mn} E_m \rho_{in} E_n^\dagger \tag{7}$$

其中，χ_{mn} 是表征量子映射 Λ的量子过程矩阵 χ 的元素。通过制备一组完整的线性无关输入态、估计量子映射作用后的输出态，最后利用形如[32]的线性方程，根据这些实验估计的输出态计算 χ_{mn} 的元素，即可通过实验确定 χ 矩阵:

$$\beta\vec{\chi} = \vec{\lambda} \tag{8}$$

其中 β 是系数矩阵，$\vec{\chi}$包含待确定的元素$\{\chi_{mn}\}$，而 $\vec{\lambda}$ 是表示实验数据的向量。

用于使用前馈神经网络模型执行量子过程层析的训练数据集，是通过随机生成一组幺正算符构建的。生成的幺正算符作用于输入态$\rho_{in} = \{|0\rangle,\ |1\rangle,\ \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\ \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)\}^{\otimes n}$，以获得$\rho_{out} = U\rho_{in}U^\dagger$。所有输出态 ρ_{out}堆叠起来以形成 $\vec{\lambda}$。最后，$\vec{\chi}$ 使用公式（8）计算。训练元素 $\{\vec{\lambda}_i, \vec{\chi}_i\}$ 随后将用于训练前馈神经网络模型，其中 λ_i 作为前馈神经网络的输入，而$\vec{\chi}_i$ 是相应的标记输出。

表 I. 在针对3000 个测试双量子比特态（$M_{\mathrm{data}} = 20$）执行 QST 的前馈神经网络模型训练完成后获得的平均态保真度；对于不同大小的训练数据集，每个数据集的轮次数量在50 至 150 之间变化（一个轮次指对完整训练数据集进行一次迭代）。

数据集大小	保真度		
	轮次(50)	轮次(100)	轮次(150)
500	0.8290	0.9176	0.9224
1000	0.9244	0.9287	0.9298
5000	0.9344	0.9379	0.9389
10000	0.9378	0.9390	0.9400
20000	0.9394	0.9414	0.9409
80000	0.9426	0.9429	0.9422

表 II. 在规模不同的训练数据集上训练前馈神经网络模型，以对 3000 个测试三量子比特态执行量子态层析（ $M_{\mathrm{data}} = 120$ ）后获得的平均态保真度；对于每个数据集，轮次从50 到 150 不等，其中一个轮次指完整训练数据集的一次迭代。

数据集大小	保真度		
	轮次(50)	轮次(100)	轮次(150)
500	0.6944	0.8507	0.8716
1000	0.8793	0.8994	0.9025
5000	0.9231	0.9262	0.9285
10000	0.9278	0.9321	0.9332
20000	0.9333	0.9362	0.9393
80000	0.9413	0.9433	0.9432

要对给定的量子态或量子过程进行层析，必须进行一系列实验。随后构建输入向量，其中的条目是各项实验的输出/读数。对于标准的量子态层析或量子过程层析，需要进行一组层析完备的实验；与这组层析完备实验对应的输入向量称为完整数据集。要进行量子态层析和

表 III. 训练前馈神经网络模型以对 3000 个测试三量子比特态 ($M_{\text{data}} = 200$) 执行量子过程层析后得到的平均量子过程保真度。对于不同大小的训练数据集，轮次数在 50 到 150 之间变化（一个轮次指对完整训练数据集进行一次迭代）。

数据集大小	保真度		
	轮次（50）	轮次（100）	轮次（150）
500	0.4904	0.5421	0.5512
2000	0.6872	0.7090	0.7202
15000	0.7947	0.8128	0.8218
20000	0.8047	0.8203	0.8295
50000	0.8305	0.8482	0.8598
80000	0.8441	0.8617	0.8691

通过前馈神经网络，利用大小为 m 的大幅缩减数据集进行量子过程层析时，通过从输入向量中随机选取 m 个元素，构建元素更少的缩减输入向量 $\vec{b}_m$ （相应缩减后的 $\vec{\lambda}_m$），其余元素设为 0（零填充）；这些缩减后的输入向量及其对应的标记输出向量用于训练前馈神经网络。

FFNN 使用Keras Python库进行训练和实现 [33]使用 Tensor-Flow 后端，在配备 48GB RAM、CPU 基础频率为 3.90GHz 的 Intel Xeon 处理器上运行。为了执行 QST 和 QPT，LeakyReLU（$\alpha = 0.5$）激活函数被用于 FFNN 的输入层和隐藏层：

$$\begin{aligned}\text{LeakyReLU(x)} =& x\,;\, x > 0 \\ =& \alpha x\,;\, x < 0\end{aligned} \tag{9}$$

输出层使用线性激活函数。验证时使用余弦相似度损失函数 $\mathcal{L} = \arccos\left(\frac{\vec{y}.\vec{y}}{||\vec{y}||.||\vec{y}||}\right)$，训练网络时使用学习率为 η 的 *adagrad* ($\eta = 0.5$) 优化器。*adagrad* 优化器会根据参数在训练过程中被更新的频率调整学习率。

前馈神经网络用于对 3000 个双量子比特和三量子比特测试量子态进行量子态层析，并对 3000 个双量子比特测试量子过程进行量子过程层析，训练数据集的大小各不相同。每个数据集的轮次数在 50 到 150 之间变化，其中一轮次是指前馈神经网络训练过程中对训练数据集进行的一次迭代。对 3000 个双量子比特和三量子比特测试量子态以及 3000 个双量子比特测试量子过程计算得到的平均保真度分别如表I、II 和III所示；M_{data} 表示缩减后的数据集大小。在比较训练数据集大小、M_{data} 和轮次数的影响后，选择 80000 作为训练数据集的最大规模，并将最大轮次数设为 150，以执行量子态层析和量子过程层析。

训练 150 轮次后，验证损失函数保持不变。

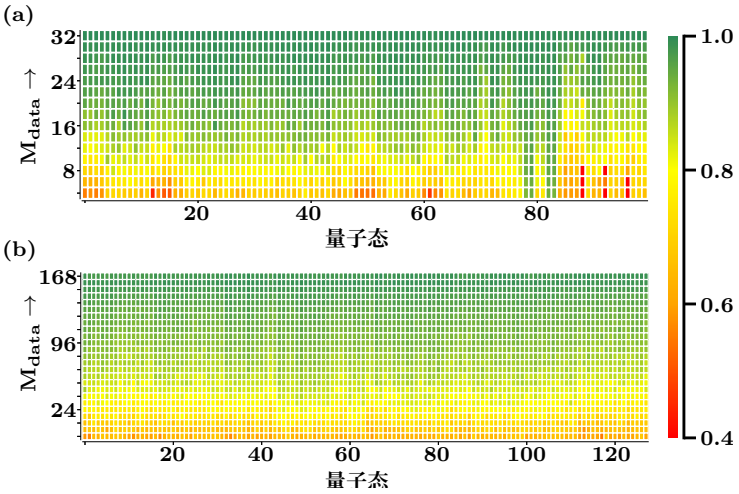

图 3.（网络版为彩色）针对 (a) 100 个双量子比特态和 (b) 128 个三量子比特态进行量子态层析时，前馈神经网络模型与标准线性反演方法之间的保真度 ($\bar{\mathcal{F}}$) 随大幅缩减的数据集大小 (M_{data}) 的变化。量子态按 x 轴编号，右侧的颜色编码条表示保真度的数值。

III. 基于实验数据的前馈神经网络量子态层析

我们使用核磁共振量子处理器作为实验测试平台，为前馈神经网络模型生成数据。我们应用前馈神经网络，利用在核磁共振量子处理器上生成的大幅精简噪声数据集，对双量子比特和三量子比特量子态进行量子态层析。通过计算平均态保真度或平均过程保真度来评估前馈神经网络的性能。态保真度定义为 [34]:

$$\mathcal{F} = \frac{\left|\text{Tr}\left[\rho_{\text{FFNN}}\rho_{\text{STD}}^{\dagger}\right]\right|}{\sqrt{\text{Tr}\left[\rho_{\text{FFNN}}^{\dagger}\rho_{\text{FFNN}}\right]\text{Tr}\left[\rho_{\text{STD}}^{\dagger}\rho_{\text{STD}}\right]}} \tag{10}$$

其中 ρ_{FFNN} 和 ρ_{STD} 分别是通过前馈神经网络和标准线性反演法获得的密度矩阵。将公式（ ρ 10）中的 χ替换为 χ_{FFNN}，其中 χ_{STD}分别是通过前馈神经网络和标准线性反演法获得的过程矩阵。

双量子比特核磁共振系统的量子态层析通常使用一组四个幺正旋转来完成：$\{II, IX, IY, XX\}$其中，I 表示恒等操作，而 $X(Y)$表示对指定量子比特进行的 $90°$ x 旋转。输入向量 $\vec{b}$ (式（4）)由施加层析脉冲后进行测量而构建；记录到的时域信号随后经傅里叶变换，得到核磁共振谱。对于两个量子比特，核磁共振谱中有四个谱峰，每次测量可得到向量 $\vec{b}$的八个元素；输入向量 $\vec{b}$ 的维数为 33×1 (32

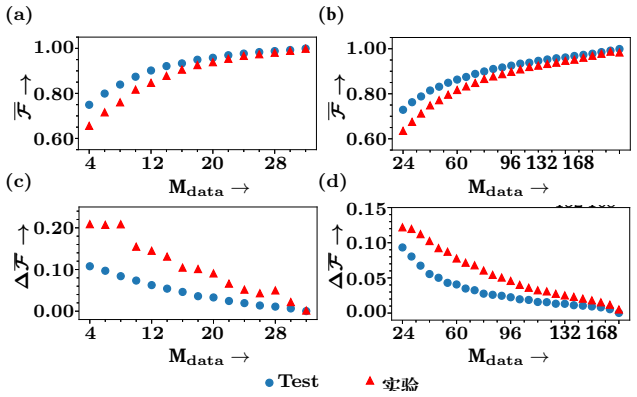

图4.（在线彩色）平均保真度（$\bar{\mathcal{F}}$）和标准偏差 $\Delta\mathcal{F}$ 随大幅缩减的数据集（M_{data}）大小的变化曲线，分别展示了基于前馈神经网络的量子态层析在两个量子比特（(a)和(c)）及三个量子比特（(b)和(d)）上的计算结果。测试数据集的平均保真度（蓝点）根据3000个量子态计算；对于实验数据集，则从完整数据集中随机选取缩减数据集 M_{data} 元素，分别用于100个双量子比特量子态和128个三量子比特量子态，然后重复该过程50次，以计算平均保真度。

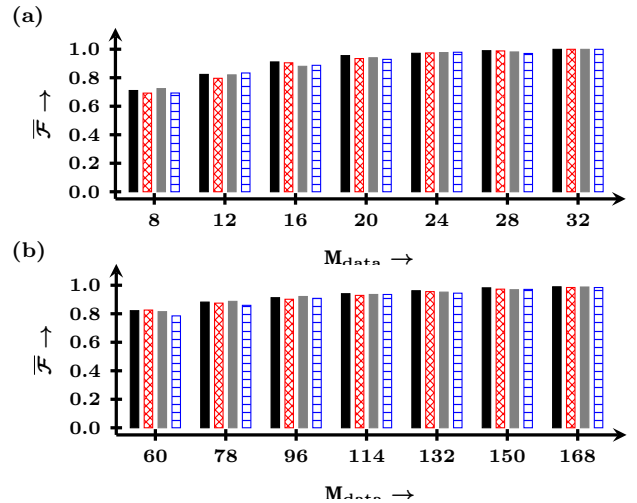

图5.（在线彩色）基于前馈神经网络的量子态层析计算得到的保真度（$\bar{\mathcal{F}}$）随大幅缩减的数据集（$M_{\rm data}$）大小的变化：(a) 双量子比特贝尔态，不同柱条对应四种不同的贝尔态；(b) 三量子比特GHZ态（黑色和红色交叉斜线柱条）以及双可分态（灰色和蓝色水平条纹柱条）。

来自层析脉冲，另有 1 个来自单位迹条件）。类似地，三量子比特核磁共振系统的量子态层析通常使用一组七个幺正旋转来完成：$\{III, IIY, IYY, YII, XYX, XXY, XXX\}$。每次测量会在核磁共振谱中产生 12 个共振谱峰（每个量子比特 4 个）；输入向量 b 的维数为 169×1。为了评估前馈神经网络模型对双量子比特态和三量子比特态进行完整量子态层析的性能，我们通过实验制备了 100 个双量子比特态

以及使用不同制备设置制备的 128 个三量子比特态，并计算了前馈神经网络模型预测的密度矩阵与通过量子态层析的标准线性反演方法获得的密度矩阵之间的平均保真度。我们还使用大幅缩减的数据集，对最大纠缠双量子比特贝尔态以及三量子比特 GHZ 态和双可分态进行了完整的基于前馈神经网络的量子态层析。

前馈神经网络模型使用 80,000 个量子态进行训练，以执行量子态层析。为了对双量子比特态和三量子比特态进行基于前馈神经网络的量子态层析，我们分别使用了包含 100、100 和 50 个神经元的三个隐藏层，以及包含 300、200 和 100 个神经元的三个隐藏层。训练后的前馈神经网络的性能如图所示 3。通过前馈神经网络和标准线性反演方法获得的密度矩阵之间的保真度，分别如图 3(a) 和 (b) 所示；这些密度矩阵对应于 100 个实验生成的双量子比特态和 128 个实验生成的三量子比特态。大小为 $M_{\rm data}$ 的缩减输入向量绘制在 y轴上，量子态沿 x轴编号。

FFNN 用于量子态层析的性能如图4所示，通过计算在一组测试态/实验态上的平均态保真度来进行评估。沿 M_{data} 轴绘制输入向量的缩减大小，该输入向量被馈入 FFNN。在 (a) 和 (c) 中，双量子比特态的平均保真度以及平均态保真度的标准差分别沿 y 轴绘制；在 (b) 和 (d) 中，则分别绘制三量子比特态的相应结果。对于给定的 M_{data} 值，通过从相应的完整输入向量中随机选择 M_{data} 个元素并重复 50 次，计算经 FFNN 预测的给定量子态的平均保真度。对于测试数据集（蓝色圆圈），通过计算 3000 个双量子比特态和三量子比特态上的平均保真度来评估 FFNN 的性能。对于实验数据集（红色三角形），通过分别计算 100 个双量子比特态和 128 个三量子比特态上的平均保真度来评估 FFNN 的性能。平均态保真度的标准差为：

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(\bar{\mathcal{F}}_i - \bar{\mathcal{F}})^2}{N-1}} \tag{11}$$

根据图4可推断，前馈神经网络模型能够以平均保真度 $\mathcal{F} \geq 0.8392 \pm 0.084$ 预测大小为$M_{data} \geq 8$的未知双量子比特测试量子态，并能够以平均保真度 $\mathcal{F} \geq 0.8630 \pm 0.0407$ 预测大小为 $M_{data} \geq 60$的未知三量子比特测试量子态。类似地，对于实验量子态，前馈神经网络模型能够以平均保真度 $\mathcal{F} \geq 0.8466 \pm 0.1450$ 预测大小为 $M_{data} \geq 12$的双量子比特量子态，而对于三量子比特实验态，前馈神经网络能够以平均保真度 $\mathcal{F} \geq 0.8327 \pm 0.0716$ 预测大小为$M_{data} \geq 60$的未知量子态。当考虑完整的输入向量 b 时，在3000个双量子比特和三量子比特

量子比特测试态结果分别为 $\mathcal{F} = 0.9993$ 和 $\mathcal{F} = 0.9989$。对于完整输入数据集，在100个双量子比特和128个三量子比特实验态上计算得到的平均保真度分别为 $\mathcal{F} = 0.9983$和 $\mathcal{F} = 0.9833$。

前馈神经网络模型被应用于对两个量子比特最大纠缠贝尔态以及三个量子比特GHZ 态和双可分态进行量子态层析。图5描绘了两个量子比特贝尔态以及三个量子比特GHZ态和双可分态的实验保真度 $\mathcal{F}(\rho_{\mathrm{FFNN}}, \rho_{\mathrm{STD}})$，该保真度是针对数据集大小为 M_{data}的精简数据集，通过比较前馈神经网络预测的密度矩阵与利用标准线性反演量子态层析获得的密度矩阵计算得到的。图5(a)中的黑色、交叉阴影红色、灰色和水平蓝色柱分别对应贝尔态 $|B_1\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$、$|B_2\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$、$|B_3\rangle = (|00\rangle - |11\rangle)/\sqrt{2}$ 和$|B_4\rangle = (|01\rangle + |10\rangle)/\sqrt{2}$。图 5(b) 中的黑色和红色交叉阴影柱分别对应三个量子比特 GHZ 态 $|\psi_1\rangle = (|000\rangle + |111\rangle)/\sqrt{2}$ 和 $|\psi_2\rangle = (|010\rangle + |101\rangle)/\sqrt{2}$，而灰色和水平蓝色柱分别对应三个量子比特双可分态 $|\psi_3\rangle = (|000\rangle + |001\rangle + |110\rangle + |111\rangle)/2$ 和$|\psi_4\rangle = (|000\rangle + |010\rangle + |101\rangle + |111\rangle)/2$。图5 中的柱状图清楚地表明，对于精简数据集，前馈神经网络模型能够以非常高的保真度预测两个和三个量子比特的纠缠量子态。

这里顺便指出，大幅缩减的数据集 M_{data} 的大小等同于用于执行量子态层析（量子过程层析）的实验读出数量，而基于线性反演的标准量子态层析（量子过程层析）方法始终使用完整数据集。因此，M_{data} 的最高值与完整数据集的大小相同；对于双量子比特量子态层析、三量子比特量子态层析和双量子比特量子过程层析，该大小分别为 32、168 和 256。

IV. 基于实验数据的前馈神经网络 QPT

我们使用前馈神经网络模型，对三组不同的实验核磁共振数据集进行双量子比特量子过程层析：(i) 幺正量子门；(ii) 非幺正过程，例如自然核磁共振退相干过程和脉冲场梯度；(iii) 在核磁共振量子处理器上使用对偶算法实验模拟的关联比特翻转、关联相位翻转和关联比特+相位翻转噪声通道。

A. 前馈神经网络对双量子比特幺正和非幺正过程的重构

前馈神经网络模型使用大幅精简的 80,000 个合成双量子比特量子过程进行训练。
数据集，包含三个隐藏层，分别有 600、400 和 300 个神经元。使用 3000 个测试样本和 10 个 ex- 对训练后的前馈神经网络的性能进行了评估。

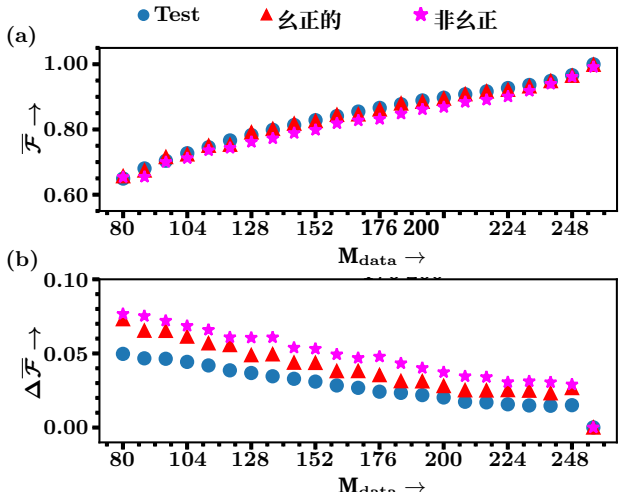

图 6.（彩色在线版）（a）使用前馈神经网络模型对双量子比特过程进行量子过程层析所得的平均过程保真度（$\bar{\mathcal{F}}$），以及（b）标准差 $\Delta\mathcal{F}$ 随数据集大小（M_{data}）的变化。对于测试数据集（蓝色圆点），平均保真度根据 3000 个过程计算；对于实验实现的幺正过程（红色三角形）和非酉过程（洋红色星形），则从四个幺正量子门和六个非酉过程的完整集合中随机选取缩减数据集 M_{data} 个元素来计算平均保真度，并重复该过程 300 次。

在核磁共振平台上实验实现的量子过程。

图 6展示了前馈神经网络对各种双量子比特实验量子过程进行量子过程层析的结果。通过计算以下平均值来评估前馈神经网络的质量：过程保真度 $\mathcal{F}$，该值比较了前馈神经网络使用大小为χ_{FFNN}的缩减数据集所预测的过程矩阵（M_{data}）与使用完整数据集通过标准量子过程层析法获得的过程矩阵（χ_{STD}）。

图 6（a）描绘了在 3000 个双量子比特测试量子过程（蓝色圆圈）上对前馈神经网络进行评估的性能，其中 y轴表示平均保真度$\bar{\mathcal{F}} = \frac{1}{3000}\sum_{n=1}^{3000}\bar{\mathcal{F}}_n$，其中 $\bar{\mathcal{F}}_n = \frac{1}{300}\sum_{i=1}^{300}\bar{\mathcal{F}}_i$ 是第 n个测试量子过程的平均保真度，该值通过随机构造给定大小的输入向量并重复该过程 300 次计算得到。类似地，红色三角形和粉色星形分别对应于从实验数据中获得的四个幺正量子过程和六个非幺正量子过程。图 6（a）所示的图清楚地表明，前馈神经网络模型能够从含噪的实验简化数据集中，以较高准确度预测幺正量子过程以及非幺正量子过程。例如，对于 $M_{\mathrm{data}} = 160$，前馈神经网络能够预测保真度为 $\mathcal{F} = 0.8411 \pm 0.0284$的测试量子过程，而实验中的幺正过程和非幺正过程分别获得了 $\mathcal{F} = 0.8447 \pm 0.038$ 和 0.8187 ± 0.0493。因此，可以根据所需的准确度和精度，相应地设置 M_{data} 的值。平均保真度的标准差 $\mathcal{F}$ 通过对 3000 个量子过程使用公式（11）计算得到，并绘于图 6（b）中。从图 6（b）可以看出

观察到，与非酉过程相比，前馈神经网络模型在幺正过程的量子过程层析方面表现更好，因为相应的过程矩阵更加稀疏。

通过前馈神经网络获得的各个量子过程的实验保真度如图 8 所示，其中，针对给定的缩减数据集大小 M_{data}，计算一组量子过程的平均保真度。对于测试数据集， $\mathcal{F}$ 是基于 3000 个测试过程计算得出的；而对于实验数据集， $\mathcal{F}$ 是基于四个幺正过程和六个非幺正过程计算得出的。对于幺正量子门：恒等、CX180、受控非门和 CY90（分别对应“无操作”门、比特翻转门、绕 x 轴旋转 180° 的受控旋转门，以及绕 y 轴旋转 90° 的受控旋转门），使用大小为 160 的缩减数据集时，前馈神经网络能够预测相应的过程矩阵，其平均保真度分别为 $\mathcal{F} = 0.8767 \pm 0.0356$、$0.8216 \pm 0.0463$、$0.8314 \pm 0.0387$ 和 0.8489 ± 0.0315。待进行层析的六个非幺正过程包括两种不同时间的自由演化过程：$D1 = 0.05$ 秒和 $D2 = 0.5$ 秒、磁场梯度脉冲（MFGP），以及三个误差通道，即关联比特翻转（CBF）通道、关联相位翻转（CPF）通道和关联比特相位翻转（CBPF）通道。在自由演化时间 $D1$ 和 $D2$ 内，多个噪声通道会同时作用于所有量子比特，例如相位阻尼通道（对应于 T_2 核磁共振弛豫过程）和振幅阻尼通道（对应于 T_1 核磁共振弛豫过程）。在核磁共振硬件中，MFGP 过程通常通过梯度线圈实现，磁场梯度沿 z 轴方向。待进行层析的 MFGP 过程是一个持续时间为 1000μ 秒的正弦形脉冲，包含 100 个时间间隔 $= 100$，施加的梯度强度为 15%。对于固有的非幺正量子过程 D1、D2 和 MFGP，使用大小为 160 的缩减数据集时，前馈神经网络能够预测相应的过程矩阵，其平均保真度分别为 $\mathcal{F} = 0.8373 \pm 0.0381$、$0.7607 \pm 0.0690$ 和 0.7858 ± 0.0703。计算得到的保真度值表明，若过程矩阵是稀疏的，前馈神经网络的表现会更好。

尽管我们的主要目标是证明，即使对于大幅缩减的数据集，前馈神经网络也能够以高保真度重构量子态和量子过程，但我们还希望验证该网络应用于完整数据集时的有效性。通过前馈神经网络针对完整数据集获得的过程保真度值如表IV所示，其中显然可以看出，前馈神经网络能够以极高的保真度预测潜在的量子过程，并且即使对于非幺正量子过程也能准确运行。与其他量子过程相比，D2 过程的保真度略低，这可以归因于相应的过程矩阵较为稀疏。

表 IV. 实验保真度 $\mathcal{F}$ 计算于χ_{FFNN}，即使用完整数据集通过前馈神经网络预测的过程矩阵，与 χ_{STD}，即通过标准量子过程层析法获得的过程矩阵之间。

幺正的 量子过程	$\mathcal{F}$	非幺正过程 量子过程	$\mathcal{F}$
Test	0.9997	D1	0.9987
恒等	0.9943	D2	0.9635
CNOT	0.9996	Grad	0.9917
CX180	0.9996	CBF	0.9943
CY90	0.9996	CPF	0.9996
		CBPF	0.9996

B. 关联噪声信道的前馈神经网络重构

对偶性模拟算法（DSA）可用于模拟完全关联的双量子比特噪声通道，即 CBF、CPF 和 CBPF 通道 [35]。随后，采用前馈神经网络模型对这些通道进行完整表征。DSA 使我们能够在一次实验中模拟开放量子系统的任意动力学，其中辅助量子比特系统的维度等于表征给定量子通道的 Kraus算符总数。具有 d 个 Kraus算符的任意量子通道都可以通过 DSA，使用幺正操作 V、W 和控制操作 $U_c = \sum_{i=0}^{d-1} |i\rangle\langle i| \otimes U_i$ 进行模拟，使其满足以下条件：

$$E_k = \sum_{i=0}^{d-1} W_{ki} V_{i0} U_i \quad (k = 0, 1, 2, ..., d-1) \tag{12}$$

其中， E_k 是Kraus算符， V_{i0} 和 W_{ki} 分别是 V和 W的元素。参考文献 [35], 给出了DSA的量子电路，其中系统的初始量子态被编码为 $|0\rangle_a \otimes |\psi\rangle_s$，随后依次作用于 $V \otimes I$、U_c 和 $W \otimes I$，最后对系统量子比特进行测量。

在本研究中，使用两个Kraus算符对双量子比特CBF、CPF和CBPF通道进行表征，如下所示：

$$\begin{aligned}
\text{CBF}: E_0 &= \sqrt{1-p} I^{\otimes 2}, \quad E_1 = \sqrt{p} \sigma_x^{\otimes 2} \\
\text{CPF}: E_0 &= \sqrt{1-p} I^{\otimes 2}, \quad E_1 = \sqrt{p} \sigma_z^{\otimes 2} \\
\text{CBPF}: E_0 &= \sqrt{1-p} I^{\otimes 2}, \quad E_1 = \sqrt{p} \sigma_y^{\otimes 2}
\end{aligned} \tag{13}$$

其中， p 表示噪声强度，也可以解释为系统量子态受到给定噪声通道影响的概率。对于 $p = 0$，系统的量子态不受影响；对于 $p = 1$，系统的量子态受到给定噪声通道的最大影响。由于本研究考察的三个噪声通道都只有两个Kraus算符，因此可以使用单个辅助量子比特对其进行模拟。因此，对于全部三个噪声

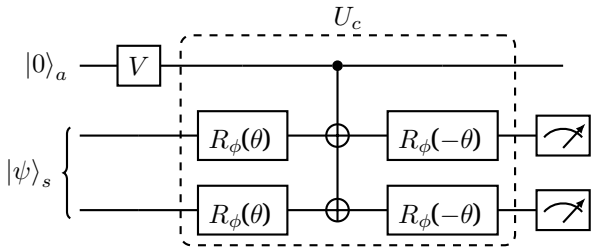

图 7. 用于模拟关联比特翻转、关联相位翻转和关联比特+相位翻转噪声通道作用的量子电路。 $|\psi\rangle_s$ 是一组线性无关的双量子比特输入态， $|0\rangle_a$ 表示辅助量子比特的状态， V 是单量子比特旋转门， U_c 表示一组控制操作，其值（θ, ϕ）随所模拟的噪声通道而变化。

通道，可以设置 $V = \begin{pmatrix} \sqrt{1-p} & -\sqrt{p} \\ \sqrt{p} & \sqrt{1-p} \end{pmatrix}$, $W = I$, 以及 $U_0 = I \otimes I$。不同的 U_1 对于 CBF、CPF 和 CBPF 通道分别设置为 $\sigma_x \otimes \sigma_x$, $\sigma_z \otimes \sigma_z$ 和 $\sigma_y \otimes \sigma_y$, 从而满足式（12）中给出的条件。注意， V 可解释为绕 y 轴旋转角度 θ，使得 $p = \sin^2(\frac{\theta}{2})$ 。

使用 DSA 模拟全部三种错误信道的广义量子电路如图7所示。对于 CBF 信道， U_c 结果是一个受控非门，其中 θ （图7）的值为零。对于 CPF 和 CBPF 信道， θ、 ϕ （旋转的角度和轴）分别为（$\frac{\pi}{2}$、 y）和（$\frac{\pi}{2}$、 z）。对系统量子比特进行层析测量得到的输出构成列向量 $\vec{\lambda}$。对于给定的 p值，通过将系统量子比特制备为一组完整的线性无关输入态，可以构造完整向量 $\vec{\lambda}$。

如图 8所示，对于 $\mathcal{F} = 0.8738 \pm 0.0366$、$0.8272 \pm 0.0403$ 和 0.8273 ± 0.0416，分别使用大小为160的缩减数据集对实验模拟的噪声通道CBF、CPF 和 CBPF 进行分析。由于这三个关联噪声通道均仅由两个Kraus算符刻画，相应的过程矩阵具有足够的稀疏性（过程矩阵中仅有两个非零元素）。因此，FFNN可用于对具有任意噪声强度的此类噪声通道进行准确的量子层析，同时仅需使用大幅缩减的数据集。

V. 结论

近期有大量研究致力于训练人工神经网络，以执行多种量子信息处理任务，包括层析、纠缠表征和量子门优化。我们设计并应用了一个前馈神经网络，利用实验核磁共振数据进行量子态层析和量子过程层析，从而重建用于表征该

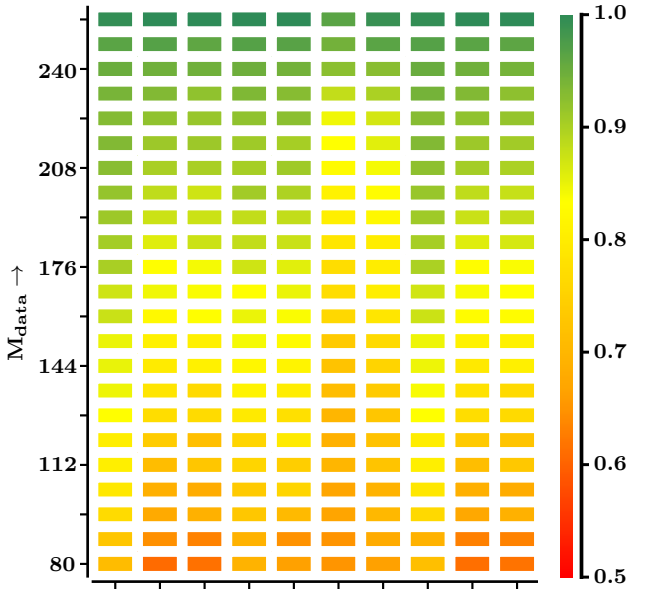

图 8.（网络版彩图）对于不同的幺正和非幺正量子过程，前馈神经网络模型与标准线性反演方法之间的过程保真度（$\mathcal{F}$）随大幅缩减的数据集大小（M_{data}）的变化。各种量子过程标注在 x 轴上，右侧的彩色条表示保真度的数值。

分别为真实量子态和量子过程。前馈神经网络能够以极高的保真度预测真实量子态和量子过程；即使实验数据集大幅缩减，其表现也依然出色。压缩感知是另一种利用缩减后的数据集对量子态和量子过程进行层析的方法。然而，这种方法需要系统噪声等先验知识，还要求对目标态（量子过程）进行层析所依据的基底足够稀疏。相比之下，前馈神经网络不需要任何此类先验知识，并且适用于所有类型的量子态和量子过程。此外，使用大幅缩减的数据集有助于显著降低实验复杂度，因为随着系统规模增大，执行完备层析实验所需的实验数量会呈指数级增长。只需进行极少量实验，并将这些最精简的实验数据集作为输入馈入前馈神经网络，后者便可重建真实的密度矩阵或过程矩阵。因此，我们的结果表明，前馈神经网络架构是对大型量子比特寄存器进行量子态层析和量子过程层析的有前景的方法，也是标准方法的一种颇具吸引力的替代方案，因为它所需资源大幅减少。

致谢

所有实验均在 IISER Mohali 核磁共振研究设施的 Bruker Avance-III 600 MHz 傅里叶变换核磁共振波谱仪上完成。Arvind 感谢科学部提供的资助——科学与技术部（DST，印度）资助，项目编号为 DST/ICPS/QuST/Theme-1/2019/Q-68。K.D. 感谢印度科学与技术部（DST）的资助，项目编号为 DST/ICPS/QuST/Theme-2/2019/Q-74。

[1] M. A.Nielsen 和I.L. Chuang, 《 *Quantum Computation and Quantum Information* 》（剑桥大学出版社，英国剑桥，2010）。[2] A. M. Childs、I. L.Chuang 和 D. W. Leung, 《 Phys.Rev.A》64，012314（2001）。[3] J. L. O′Brien、G. J. Pryde、A. Gilchrist、D.F. V. James、N. K. Langford、T. C. Ralph和 A. G. White, 《Phys.Rev. Lett.》93，080502（2004）。[4] R. J. Chapman、C. Ferrie 和 A. Peruzzo, 《Phys.Rev.Lett.》117，040402（2016）。[5] I.A.Pogorelov、G. I. Struchalin、S.S. Straupe、I.V.Radchenko、K.S. Kravtsov 和 S. P. Kulik, 《Phys.Rev. A》95，012302（2017）。[6] C. A.Riofrao、D. Gross、S. T. Flammia、T. Monz、D. Nigg、R.Blatt 和J. Eisert, 《Nat. Commun.》8，15305（2017）。[7] A. Gaikwad、Arvind 和 K. Dorai, 《Quant. Inf. Proc.》20，19（2021）。[8] A. Bendersky、F. Pastawski 和J. P. Paz, 《Phys. Rev. Lett.》100，190403（2008）。[9] A. Gaikwad、D. Rehal、A. Singh、Arvind 和 K. Dorai, 《Phys. Rev. A》97，022311（2018）。[10] A. Gaikwad、K. Shende、Arvind 和K.Dorai, 《Sci.Rep.》12，3688（2022）。[11] Y. Kim、Y.-S.Kim、S.-Y. Lee、S.-W. Han、S.Moon、Y.-H.Kim和 Y.-W.Cho, 《Nat. Commun.》9，192（2018）。[12] G.Carleo 和 M. Troyer, 《Science》355，602（2017），https://www.science.org/doi/pdf/10.1126/science.aag2302

[13] G. Carleo、I. Cirac、K. Cranmer、L. Daudet、M. Schuld、N. Tishby、L. Vogt-Maranto 和 L.Zdeborov´a, Rev. Mod. Phys. **91**，045002 (2019)。[14] G. Torlai 和 R. G. Melko, Annu. Rev. Condens. Matter Phys. **11**，325 (2020), https://doi.org/10.1146/annurev- conmatphys-031119-050651。[15] G. Torlai、G. Mazzola、J. Carrasquilla、M. Troyer、R. Melko 和 G. Carleo, Nat. Phys. **14**，447 (2018)。[16] M. Neugebauer、L. Fischer、A. J¨ager、S. Czischek、S. Jochim、M. Weidem¨uller 和 M. G¨arttner, Phys. Rev. A **102**，042604 (2020)。[17] Y. Quek、S. Fort 和 H. K. Ng, Npj Quantum Inf. **7**，105 2021( )。

[18] P. Cha, P. Ginsparg, F.Wu, J. Carrasquilla,P. L. McMa-hon,and E.-A. Kim, 机器学习：科学与技术 3, 01 01 2021LT ( ).[19] A. M. Palmieri, E. Kovlakov, F. Bianchi, D. Yudin, S. Straupe, J.D. Biamonte, and S.Kulik, npj 量子信息 6, 20 (2020). [20] S.Lohani, B.T. Kirby,M. Brodsky, O.Danaci,and R. T. Glasser,机器学习：科学与技术1, 035007 (2020). [21] T.Xin, S. Lu, N. Cao, G.Anikeeva, D.Lu,J.Li,G. Long, and B.Zeng, npj量子信息 5,109 (2019).[22] Y. Tian, L.Che,X.Long, C. Ren, and D.Lu,高级量子技术 5, 2200025(2022).[23] C. Pan andJ. Zhang, 理论物理国际期刊 61, 227 (2022). [24] Y. S. Teo,S. Shin, H. Jeong, Y. Kim, Y.-H. Kim, G. I. Struchalin, E. V. Kovlakov,S. S. Straupe,S.P. Ku- lik, G. Leuchs, and L. L. Sanchez-Soto,新物理学杂志 23, 103021 (2021).[25] A. M. Palmieri, E. Kovlakov, F. Bianchi, D. Yudin, S. Straupe, J.D. Biamonte, and S.Kulik, npj量子信息 6,20(2020).[26] X. Zhang, M. Luo, Z. Wen,Q. Feng,S.Pang,W. Luo, and X. Zhou,物理评论快报 127, 130503 (2021). [27] D. Koutny, L. Motka, Z. Hradil, R. Jaroslav, and L. L. Sanchez-Soto,物理评论A 106,012409(2022). [28] S. Ahmed, C. S´anchez Mu˜noz,F. Nori, and A. F.Kockum,物理评论快报 127,140502 (2021). [29] S.Ruder,“梯度下降优化算法概述,” (2017), arXiv:1609.04747 [cs.LG]. [30] G. L.Long,H. Y.Yan, and Y. Sun, 光学杂志B：量子与半经典3, 376 (2001). [31] K. Kraus, A. Bohm, J. Dollard, and W. Wootters, *States, Effects and Operations: Fundamental Notions of Quantum Theory* 1983 (Springer-VerlagBerlin Heidelberg,). [32] I.L.Chuang and M. A. Nielsen, 现代光学杂志 44,2455 (1997). [33] F. Chollet, “ Keras,” https://github.com/keras-team/ keras (2015). [34] J. Zhang, A. M. Souza, F.D. Brandao, and D. Suter,物理评论快报 112,050502 (2014). [35] T. Xin, S.-J. Wei, J. S.Pedernales,E. Solano, and G.-L. Long, 物理评论A 96, 062303 (2017).