

最简结构神经网络的量子态估计及其性能对比

丛爽[†], 李友志

(中国科学技术大学 自动化系, 安徽 合肥 230027)

摘要: 本文提出并设计两种具有最简结构的前向神经网络, 来高精度实现对量子态密度矩阵的估计. 训练出具有函数逼近功能的反向传播(BP)神经网络和径向基函数(RBF)网络进行量子密度矩阵估计的应用. 根据量子态密度矩阵与量子系统实验装置的输出测量值之间的关系, 建立并构造出训练神经网络的输入/输出样本对; 通过对网络的归一化处理, 获得满足量子密度矩阵条件的网络输出. 分别对2量子位的本征态、叠加态和混合态的估计设计和训练出不同网络, 并在给定的性能指标下, 与采用深度学习算法的具有两个隐含层的宽度网络(WNN)的量子密度矩阵估计性能进行对比分析. 在此基础上, 采用RBF神经网络对高量子位密度矩阵进行估计实验. 分别在最少隐含层节点数、最少训练样本数、最短训练时间, 以及对非样本输入数据的泛化能力方面, 通过仿真实验对所设计网络的量子密度矩阵估计的优越性能进行对比研究.

关键词: 神经网络; 量子态估计; 结构优化

引用格式: 丛爽, 李友志. 最简结构神经网络的量子态估计及其性能对比. 控制理论与应用, 2024, 41(12): 2401 – 2407

DOI: 10.7641/CTA.2023.21008

Quantum state estimation with minimal structured neural networks and its performance comparison

CONG Shuang[†], LI You-zhi

(Department of Automation, University of Science and Technology of China, Hefei Anhui 230027, China)

Abstract: In this paper, two feedforward neural networks with the minimal structure are proposed and designed to estimate the quantum state density matrix with high accuracy. The back propagation (BP) neural network and the radial basis function (RBF) with function approximation function are designed and trained for the application of quantum density matrix estimation. According to the relationship between the quantum state density matrix and the output measurement value of the quantum system experimental device of the quantum system, the input/output sample pairs for training the neural networks' weights are established and constructed. The network output satisfying the condition of quantum density matrix is obtained by normalizing the networks. Different networks are designed and trained for 2-qubit eigenstate, superposition state and mixed state, and the performances of different networks are compared with the quantum density matrix estimation results of width neural network (WNN) with two hidden layers using deep learning algorithm under the same given performance index. On this basis, the RBF neural network is used to estimate the high qubit density matrix. The superior performances of the quantum density matrix estimation of the designed networks in the minimum number of hidden layer nodes, the minimum training samples, the minimum training time and the generalization ability of non-sample input data are compared by simulation experiments.

Key words: neural networks; quantum state estimation; structural optimization

Citation: CONG Shuang, LI Youzhi. Quantum state estimation with minimal structured neural networks and its performance comparison. *Control Theory & Applications*, 2024, 41(12): 2401 – 2407

1 引言

量子系统采用希尔伯特空间中的矢量来描述密度矩阵状态, 对一个 n 比特量子系统状态的密度矩阵 $\rho \in \mathbb{C}^{d \times d}$, $d = 2^n$ 的估计, 一般是通过测量获取量子

系统的输出值^[1], 然后根据输出值与量子密度矩阵之间的关系, 联立方程求解. 为了求出密度矩阵中的 4^n 参数, 一般需要相同的测量次数, 即完备测量的次数, 同时, 密度矩阵还必须满足半正定、单位迹和共轭对

收稿日期: 2022-11-14; 录用日期: 2023-10-23.

[†]通信作者. E-mail: scong@ustc.edu.cn.

本文责任编辑: 席在荣.

国家自然科学基金项目(62473354)资助.

Supported by the National Science Foundation of China (62473354).

称的性质条件^[2-3]. 量子层析技术是最早的量子态估计技术^[4]. 随后, Cahill和Glauber^[5]提出通过对量子系统大量重复测量获得的统计结果来估计量子态的密度矩阵, 为量子态估计的实现奠定了基础. 近些年, 多种优化算法被应用到量子态估计中. 2017年杨靖北等^[6]提出一种基于两步测量的量子态估计方法, 对任意量子纯态的估计提供了最少的观测次数. 2020年, 一种改进的迭代收缩阈值算法被提出, 得到了量子态估计的更小误差^[7]. 2021年针对连续弱测量中存在高斯噪声的情况, 一种带有自适应学习速率的矩阵指数梯度量子态估计算法的提出获得了更高的估计精度^[8].

实际上数值优化算法只能对一个固定的量子状态进行估计, 而且需要对每一个量子态都进行反复的迭代优化. 由于神经网络具有对非训练样本的泛化能力, 近年来在量子态估计的应用中得到研究者的关注. 在对典型或一定数量的输入/输出数据进行训练后, 神经网络的最大特点是具有对未训练样本的泛化能力, 一旦训练好一个神经网络, 网络就能够对任意同类的测量值进行密度矩阵逼近估计, 这是数值优化算法所不具有的优势. Torlai等^[9]将玻尔兹曼机用于一维和二维的高度纠缠态的量子状态估计, 重建了纠缠熵. Xu等^[10]提出了一种用于全量子状态估计的神经网络估计模型, 实验结果证明了模型的准确性. Xin等^[11]利用全连接神经网络进行了非完备测量下的量子态估计, 对于一个4比特量子位系统有较高的精确度. Lohani等^[12]利用卷积神经网络实现了量子态估计, 并获得了较好的性能. 林文瑞和丛爽^[13]设计了一种基于深度学习去噪的近似消息传递(learned denoising-based approximate message passing, LDAMP)的深度学习网络, 在较少的测量的采样率下, 高精度地完成了对4种量子态同时进行估计. 本文研究组通过设计含有两个隐含层, 且隐含层节点数为300的宽度神经网络, 采用自适应训练算法, 实现对量子态的估计, 并达到了和深度神经网络同样高的准确性和精度^[14]. 通过研究发现, 虽然深度神经网络对庞大而又复杂的数据有较好的处理能力, 在量子态估计中获得了较高的准确性和精度, 然而, 深度神经网络结构复杂, 训练过程繁琐, 不利于将估计出的量子状态用于状态反馈控制系统中. 实际上, 采用前向网络对量子态估计可能并不需要使用很复杂的深度学习网络. 现在的问题是: 能否训练出一个结构简单, 同时对量子态的估计性能高以及泛化能力强的神经网络? 本文想回答这个问题.

本文的研究重点放在如何设计仅具有一个隐含层的网络结构, 并通过研究最少的隐含层节点数、最少的训练样本的情况下, 获得更少训练时间, 以及高精度估计量子密度矩阵的典型神经网络. 本文通过设计两种含有一个隐含层的前向神经网络实现对量子态

密度矩阵的估计, 分别通过对反向传播(back-propagation, BP)网络采用Levenberg-Marquardt(L-M)算法, 以及对径向基函数(radial basis function, RBF)网络的隐含层节点自动寻优算法, 训练出两种具有函数逼近功能的前向神经网络, 并通过对网络输出的归一化处理, 获得满足密度矩阵条件的网络输出; 根据量子态密度矩阵与量子系统实验装置的输出测量值之间的关系, 构造出训练前向神经网络的输入/输出样本对, 分别设计、训练和验证不同网络对本征态、叠加态和混合态的逼近估计性能. 鉴于RBF展现出的优良估计性能, 进而采用RBF网络对2量子位的任意量子态进行估计, 以及对3-7高量子位的本征态进行估计的研究, 并分别对无及含有噪声情况下的估计性能进行了性能对比分析.

本文的安排如下: 第2节为量子态估计与神经网络输入/输出关系以及训练用数据的生成; 第3节为神经网络结构设计及其性能指标; 第4节为神经网络训练及性能对比分析; 第5节是结论.

2 量子态估计与神经网络输入/输出关系及网络训练用数据的生成

量子态密度矩阵 ρ 与量子系统输出测量值 y 之间的关系为

$$y = \text{tr}(M\rho), \quad (1)$$

其中: M 为所选择的测量算符, ρ 为量子态密度矩阵.

采用神经网络进行量子密度矩阵的估计任务, 就是选择合适的测量算符 M , 通过完备的 $d \times d$, $d = 2^n$ 次测量, 重复地作用于量子系统的待估计的状态 ρ 上, 将所获得的与测量算符相对应的每一个测量值 y 组成的数据, 作为神经网络的一组输入, 通过训练网络的权值, 使得网络的输出 $\hat{\rho}$ 逼近待估计的状态 ρ , 换句话说, 使得网络的输入/输出关系逼近式(1).

由式(1)可知, 要想获得系统的测量值 y , 需要首先选择量子态的测量算符 M . 本文采用泡利矩阵来构造测量算符 M . 神经网络的输入样本为泡利矩阵构造的测量算符作用在待估计的密度矩阵上获得的测量值. 泡利矩阵是一组3个 2×2 的么正厄米复矩阵, 用 σ_x , σ_y , σ_z 表示, 将单位矩阵 I 视为第零号泡利矩阵, 有

$$\begin{cases} \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \\ \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{cases} \quad (2)$$

对于一个 n 比特的量子系统, 由泡利矩阵构成的测量算符 M_i 是从集合 $\{\sigma_x, \sigma_y, \sigma_z, I\}$ 中随机挑选 n 个矩阵, 然后计算这些矩阵的归一化张量积组成, 即

$$M_i = \frac{\sigma_1 \otimes \cdots \otimes \sigma_n}{\sqrt{d}}. \quad (3)$$

由式(3)可以得到所有完备测量算符. 对于 n 比特量子系统, 共有 $d \times d$ 个不同的测量算符. 根据式(3)中的测量算符 M_i , 代入式(1)可以获得对应的密度矩阵 ρ , 使得测量值 y_i 为

$$y_i = \text{tr}(M_i^\dagger \rho) = \text{vec}(M_i)^\dagger \text{vec}(\rho), \quad (4)$$

其中 $\text{vec}(\cdot)$ 表示将矩阵按列组成一个向量. 将所有在不同测量算符 M_i 作用下得到的测量值 y_i 排列成一个向量 b_k , 可得

$$b_k = [y_1 \ y_2 \ \cdots \ y_{d \times d}], \quad (5)$$

其中的下标 $d \times d$ 为完备测量次数.

考虑到实际测量过程会存在噪声, 对每一个待测量子密度矩阵 ρ 的测量向量 b_k 加上随机高斯噪声 e_k 可得

$$\begin{cases} b'_k = A_k \text{vec}(\rho) + e_k, \\ A_k = (\text{vec}(M_1), \text{vec}(M_2), \dots, \text{vec}(M_{d \times d}))^\dagger. \end{cases} \quad (6)$$

式(6)为构造的带有噪声的神经网络输入数据, 对应于在测量算符 M_i 对式(1)中密度矩阵作用下的完备测量数据 b_k , 组成训练神经网络权值的输入样本.

任意一个密度矩阵只有满足一定的条件才能成为量子本征态、叠加态或混合态. 所以对于训练出的神经网络的输出值, 本文还必须进行适当的处理才能成为量子密度矩阵: 1) 对本征态估计的网络输出, 所对应的真实本征态最多只有有限的 d 个在对角线某一位置的元素为 1, 其他位置元素均为 0 的矩阵; 2) 训练用的生成真实的叠加态和混合态密度矩阵网络输出数据公式必须满足归一化条件

$$\rho = \frac{\phi_r^\dagger \phi_r}{\text{tr}(\phi_r^\dagger \phi_r)}, \quad (7)$$

其中: ϕ_r 是 $d \times r$ 维的 Wishart 矩阵(波函数), 矩阵内各元素满足均值为 0, 标准差为 1 的正态分布^[15]; r 为密度矩阵的秩, 当秩 $r = 1$ 时, 为叠加态; 当秩 $r \geq 1$ 时, 为混合态.

3 神经网络结构设计及其性能指标

本节根据第 2 节中有关量子态密度矩阵与量子系统输入测量值之间的关系(1), 分别进行 BP 神经网络和 RBF 神经网络结构设计, 训练 BP 神经网络和 RBF 神经网络权值的算法以及在对网络权值进行训练和验证过程中所采用的性能指标的选择.

3.1 神经网络结构设计

本文所研究的反向传播(BP)神经网络和径向基函数(RBF)神经网络的结构如图 1 所示, 两种神经网络都只含有一个隐含层. 对于一个 n 比特的量子系统, 其密度矩阵的维数为 $d \times d = 2^n \times 2^n$, 采用对一个密度矩阵的完备的测量次数作为网络的输入节点数, 所以, 每一组输入数据, 就是一个密度矩阵的完备测量结果.

另一方面, 由于密度矩阵可能是复数, 其中含有实部和虚部, 通过将实部与虚部分开, 可以得到两组实数矩阵, 然后按列拉成两个向量, 将虚部拉成的向量放在实部下面组成一个合并向量, 成为神经网络的输出, 这样, 神经网络的输出节点数为 $2 \times d \times d$, 用 ρ 表示. 神经网络的输入向量为 $[y_1 \ y_2 \ \cdots \ y_{d \times d}]^T$, 输出向量为 $[\rho_{r11} \ \rho_{r21} \ \cdots \ \rho_{r(d-1)d} \ \rho_{rdd} \ \rho_{i11} \ \rho_{i21} \ \cdots \ \rho_{i(d-1)d} \ \rho_{idd}]^T$.

图 1 中的归一化处理层的目的是为了使得由网络估计出的密度矩阵同时满足量子密度矩阵的“厄密和迹为 1”条件. 对于本征态和叠加态, 神经网络在输出之后的归一化处理计算为

$$\hat{\rho}_t = \frac{\hat{\rho}^\dagger \hat{\rho}}{\text{tr}(\hat{\rho}^\dagger \hat{\rho})}. \quad (8)$$

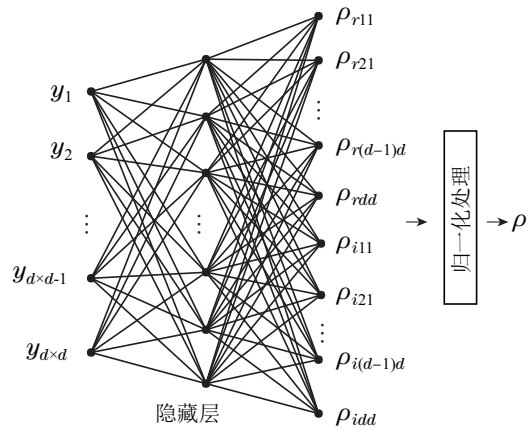


图 1 量子态估计的前向神经网络结构图

Fig. 1 Structure diagram of a forward neural network for quantum state estimation

对于 2 比特混合态, 神经网络输出之后的归一化处理计算为: 先将密度矩阵 ρ 分解为下三角矩阵 τ ^[16], τ 与 ρ 中各元素关系为

$$\tau = \begin{bmatrix} \sqrt{\frac{\text{Det}(\rho)}{m_{11}^{(1)}}} & 0 & 0 & 0 \\ \frac{m_{12}^{(1)}}{\sqrt{m_{11}^{(1)} m_{11,22}^{(2)}}} & \sqrt{\frac{m_{11}^{(1)}}{m_{11,22}^{(2)}}} & 0 & 0 \\ \frac{m_{12,23}^{(2)}}{\sqrt{\rho_{44}} \sqrt{m_{11,22}^{(2)}}} & \frac{m_{11,23}^{(2)}}{\sqrt{\rho_{44}} \sqrt{m_{11,22}^{(2)}}} & \sqrt{\frac{m_{11,22}^{(2)}}{\rho_{44}}} & 0 \\ \frac{\rho_{41}}{\sqrt{\rho_{44}}} & \frac{\rho_{42}}{\sqrt{\rho_{44}}} & \frac{\rho_{43}}{\sqrt{\rho_{44}}} & \sqrt{\rho_{44}} \end{bmatrix}, \quad (9)$$

其中: $m_{ij}^{(1)}$ 是 ρ 的一阶余子式, $m_{ij,kl}^{(2)}$ 是 ρ 的二阶余子式, 则估计的密度矩阵 $\hat{\rho}_t$ 为

$$\hat{\rho}_t = \frac{\hat{\tau}^\dagger \hat{\tau}}{\text{tr}(\hat{\tau}^\dagger \hat{\tau})}. \quad (10)$$

可以验证经过处理后的密度矩阵满足半正定性、单位迹及共轭对称性。

3.2 训练网络权值的算法

对于隐含层输出范围在 $(-1, 1)$ 的Sigmoid激活函数的BP神经网络, 采用MATLAB工具箱中的L-M算法来训练网络权值. 对于隐含层激活函数为径向基函数的RBF神经网络, 采用MATLAB中对隐含层节点进行自动寻优的newrb.m函数来训练RBF神经网络. 该方法当隐含层节点数等于样本输入组数时, 可以达到0误差.

3.3 训练和验证用性能指标

为了能够直接采用MATLAB工具箱中的程序训练所设计的神经网络, 网络在训练过程中不加归一化处理, 这样可以采用均值平方差(mean squared error, MSE)函数作为训练神经网络的性能指标, 通过使MSE达到期望的性能指标来分别训练本征态、叠加态和混合态神经网络, 其计算公式为

$$\text{MSE}(\rho, \hat{\rho}) = \frac{\sum_{i=1}^{2^n} \sum_{j=1}^{2^n} (\rho_{i,j} - \hat{\rho}_{i,j})^2}{4^n}, \quad (11)$$

其中: ρ 为真实的密度矩阵; $\hat{\rho}$ 为训练过程中神经网络的输出, 即密度矩阵的估计值; n 为量子比特数.

本文采用两个性能指标对已经训练好的经过归一化处理的神经网络的性能进行验证, 第1个是真实的密度矩阵和归一化处理过后的估计的密度矩阵的均方根(root-MSE, RMSE)

$$R(\rho, \hat{\rho}_t) = \sqrt{\frac{\sum_{i=1}^{2^n} \sum_{j=1}^{2^n} (\rho_{i,j} - \hat{\rho}_{i,j})^2}{4^n}}. \quad (12)$$

第2个是真实密度矩阵与归一化处理过后的估计密度矩阵 $\hat{\rho}_t$ 之间的保真度(fidelity)

$$F(\rho, \hat{\rho}_t) = \text{tr}(\sqrt{\sqrt{\rho} \hat{\rho}_t \sqrt{\rho}}). \quad (13)$$

保真度 $F(\rho, \hat{\rho}_t)$ 的范围是 $[0, 1]$, 真实密度矩阵 ρ 与估计密度矩阵 $\hat{\rho}_t$ 的差别越小, 保真度越接近于1.

4 不同类型及不同量子位的神经网络训练及性能对比分析

本节首先采用BP神经网络和RBF神经网络分别对本征态、叠加态和混合态3种量子态下 $n = 2$ 比特量子密度矩阵进行训练, 并将所设计的BP和RBF神经网络的实验结果与采用两个隐含层、隐含层的节点数均为300、权值训练算法采用动量法和RMSprop算法的结合的Adam梯度下降优化算法的宽度神经网络^[14]的结果进行性能对比.

本节对本征态训练时, 训练样本对包含所有可能的本征态, 验证时, 验证数据为所有可能的本征态的测量值, 验证RMSE和保真度为平均值. 对叠加态和混合态训练时, 训练样本对为随机生成的数据, 验证时, 验证数据为随机100个密度矩阵的测量值, 验证RMSE和保真度给出了最小值 R_{Min} , 最大值 R_{Max} 和平均值 R_{Mean} . 验证保真度的最大值为 F_{Max} 、最小值为 F_{Min} , 平均值为 F_{Mean} . 本节所给出的训练时间均为10次训练的平均值.

仿真实验运行环境为: MATLAB R2019a 1.60 Ghz Intel Core i5-8265U CPU, 内存8 GB; 宽度神经网络的仿真实验运行环境为PyCharm, 2.40 GHz Intel Core i5-1135G7 CPU, 内存16 GB. 实验中, e_k 取均值为0、标准差为0.001的高斯噪声序列, 第4.4节中 e_k 取均值为0、标准差为0.01的高斯噪声序列.

4.1 2量子比特本征态的估计

本实验是对2量子比特的本征态进行网络训练, 考察在给定的性能指标 10^{-7} 下, 达到对任意真实本征态密度矩阵的逼近所需要的最少网络隐含层数、最少训练样本数, 以及所对应的训练时间. 3种神经网络本征态训练实验结果的性能对比由表1所示, 从中可以看出:

1) RBF神经网络、BP神经网络和宽度神经网络的隐含层节点数分别为3, 4和300/300, 因为2量子比特的本征态只有 $d = 2^n = 4$ 个, 所以4组以上的输入就一定有重复, 所以从理论上最多只要4组应当能够完成逼近和估计, RBF的隐含层节点数最多也应当在与训练组数相等时, 达到零误差. 表1中的实验结果证实了这一点. 换句话说, 对于量子本征态的估计, 完全不需要采用具有深度或宽度神经网络来实现.

表1 3种神经网络本征态训练实验结果对比

Table 1 Experimental results of eigenstate training of three neural networks

性能	RBF网络	BP网络	宽度网络
隐含层节点数	3	4	300
训练误差MSE	0	9.61×10^{-8}	8.00×10^{-8}
训练样本数	4	15	9000
训练次数	1	12	10
训练时间	0.48''	0.80''	1.97''
R_{Max}	5.53×10^{-4}	1.73×10^{-6}	无
R_{Min}	7.81×10^{-5}	3.27×10^{-8}	无
R_{Mean}	2.02×10^{-4}	8.16×10^{-7}	2.70×10^{-5}
F_{Max}	0.9999	1.0000	无
F_{Min}	0.9991	1.0000	无
F_{Mean}	0.9997	1.0000	0.9999

2) RBF神经网络只花费0.48 s就完成了训练, BP神经网络的训练时间为0.80 s, 宽度神经网络的训练时

间为1.97 s, 几乎是RBF神经网络用时的4倍. 实际上, RBF神经网络只需要对所求输入数据训练1次, 即可达到训练性能; BP神经网络训练了12次, 宽度神经网络训练了10次, RBF神经网络在2量子比特本征态估计中的多项性能中都表现最好.

4.2 2量子比特叠加态的估计

与本征态的数目只有 d 个不同, 叠加态有无数多个, 此时需要最少多少训练数据? 神经网络的结构是否会变得复杂呢? 这是笔者通过实验想弄清楚的问题. 训练误差指标选为 10^{-5} .

3种神经网络叠加态训练实验结果由表2所示, 从中可以看出:

1) RBF神经网络和BP神经网络都只需要45个训练数据, 凭借其泛化能力, 就能够对验证用的100个非训练数据获得大于98%的平均保真度, 并不需要像采用深度学习的宽度网络那样采用90000个训练样本.

2) RB的训练次数同样只需要1次, BP网络需要25次的迭代训练, 不过BP神经网络的隐含层节点数只需要16个, 比RBF的44要少28个, 所以BP网络的训练时间也是最短的, 只花费了1.58 s, 因此, 在叠加态实验中, BP神经网络所花费的计算时间最短.

表 2 3种神经网络叠加态训练实验结果

Table 2 Experimental results of superposition state training of three neural networks

性能	RBF 网络	BP 神经网络	宽度神经网络
隐含层节点数	44	16	300/300
训练误差MSE	0	9.76×10^{-6}	6.20×10^{-7}
训练样本数	45	45	90000
训练次数	1	25	1000
训练时间	2.80''	1.58''	14'2''
R_{Max}	0.0918	0.0745	无
R_{Min}	0.0100	0.0036	无
R_{Mean}	0.0322	0.0178	0.0076
F_{Max}	0.9992	0.9998	无
F_{Min}	0.9265	0.9529	无
F_{Mean}	0.9838	0.9932	0.9999

4.3 2量子比特混合态的估计

与叠加态相比, 混合态更加复杂, 本节研究单独设计混合态估计的神经网络的性能, 实验中网络输出归一化处理采用式(10), 其中秩取 $r = 4$. 3种神经网络混合态训练实验结果由表3所示, 从中可以看出:

1) RBF神经网络所需要的隐含层节点数最少, 与叠加态情况相比, 不但没有增加, 反而从44个减少到24个; BP神经网络保持16个;

2) RBF神经网络所需要的训练样本数据从45减少到25, 而BP神经网络从45增加到60. 由此可见RBF神经网络对于复杂的混合态表现出更加优越的性能.

表 3 3种神经网络混合态训练实验结果

Table 3 Experimental results of mixed state training of three neural networks

性能	RBF 网络	BP 神经网络	宽度神经网络
隐含层节点数	24	16	300/300
训练误差MSE	0	9.90×10^{-7}	无
训练样本数	25	60	90000
训练次数	1	60	无
训练时间	1.70''	2.77''	无
R_{Max}	0.0517	0.0798	无
R_{Min}	0.0092	0.0038	无
R_{Mean}	0.0252	0.0185	0.0019
F_{Max}	0.9985	0.9998	无
F_{Min}	0.9550	0.9497	无
F_{Mean}	0.9886	0.9920	0.9999

4.4 2量子比特任意量子态的估计

由于RBF在已经做的3个实验中的性能表现突出, 本节将设计训练一个RBF神经网络来估计同时含有本征态、叠加态和混合态的任意量子态, 并考察其估计性能. 本文研究的重点同样放在寻求最少隐含层节点数, 以及所需要的最少训练样本上. 通过多次不同情况下实验及其结果分析, 最终选择的最少80个训练样本的组合为: 4个本征态、任意40个叠加态和任意36个混合态; 另外100个验证样本采用4个本征态、任意48个叠加态和任意48个混合态, 采用式(8)进行归一化处理, 其中混合态的秩 $r = 2$. 此外, 本文还考虑了无噪声和有噪声两种情况下, 采用RBF神经网络对任意量子态进行估计的训练, 实验结果由表4所示, 训练误差指标定为 10^{-6} . 从中可以看出: 无噪声情况下验证保真度比有噪声更接近于1, 精度更高. 在对100个非训练样本的任意态的估计中, 所设计的网络在有噪声情况下的平均保真度也高于97%.

表 4 RBF神经网络任意量子态训练实验结果

Table 4 Experimental results of arbitrary quantum state training of RBF neural networks

性能	无噪声	有噪声
隐含层节点数	77	78
训练误差MSE	9.27×10^{-7}	6.14×10^{-7}
训练样本数	80	80
训练次数	1	1
训练时间	6.54''	6.84''
R_{Max}	0.0752	0.0975
R_{Min}	3.65×10^{-5}	0.0055
R_{Mean}	0.0367	0.0442
F_{Max}	1.0000	0.9995
F_{Min}	0.9553	0.8754
F_{Mean}	0.9836	0.9739

4.5 多量子比特本征态的估计

本节中进一步研究不同神经网络对多量子比特密度矩阵估计逼近的性能。RBF神经网络多比特本征态训练实验结果如表5所示, 由于实验中所有情况下的RMSE测试性能均为0, 保真度值均为1, 本文仅给出一个平均值。从表5中可以看出:

1) RBF神经网络对2-7比特本征态估计的规律为: 样本数只需要所有可能的本征态 2^n , 隐含层节点数比样本数少1个, 并且训练次数都只需要1次。2-6比特时训练所花费时间都在5 s内。

2) 在无噪声的情况下, 以隐含层节点数最多等于

样本数来实现RBF网络的设计, 2-7量子位的实现结果验证了理论的正确性。

BP神经网络2-4比特本征态训练实验结果如表6所示, 从中可以看出, BP和RBF神经网络的训练样本数和隐含层节点数基本相同, 并且达到了保真度为1的验证精度。不过对于2和3量子比特, 训练BP神经网络所花费的时间比RBF神经网络更短, 但是当量子比特为4时, BP神经网络花费3 min 2 s, 而RBF只花费1.15 s, 由此可见, 在量子比特较少时, 网络结构简单的量子态估计实验中, BP神经网络有一定的优势, 而当量子比特数大于4后, RBF神经网络更有优势。

表5 RBF神经网络多比特本征态训练实验结果

Table 5 Experimental results of multiqubit eigenstate training of the RBF neural network

性能	2-量子位	3-量子位	4-量子位	5-量子位	6-量子位	7-量子位
隐含层节点数	3	7	15	31	63	127
训练误差MSE	0	0	0	0	0	0
训练样本数	4	8	16	32	64	128
训练次数	1	1	1	1	1	1
训练时间	0.57''	0.77''	1.15''	2.10''	5.43''	1'14''
R_{Mean}	0	0	0	0	0	0
F_{Mean}	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

表6 BP神经网络多比特本征态训练实验结果

Table 6 Experimental results of multiqubit eigenstate training of the BP neural network

性能	2-量子位	3-量子位	4-量子位
隐含层节点数	4	8	16
训练误差MSE	1.21×10^{-9}	3.17×10^{-7}	1.00×10^{-8}
训练样本数	4	8	16
训练次数	8	13	20
训练时间	0.11''	0.48''	3'2''
R_{Max}	3.79×10^{-5}	4.10×10^{-5}	6.77×10^{-5}
R_{Min}	1.10×10^{-6}	2.69×10^{-6}	2.30×10^{-6}
R_{Mean}	1.48×10^{-5}	1.25×10^{-5}	3.00×10^{-5}
F_{Mean}	1.0000	1.0000	1.0000

5 结论

本文设计了含有一个隐含层的BP和RBF两种前向神经网络来实现对量子态密度矩阵的估计。重点研究了2量子比特本征态、叠加态和混合态3种量子状态的前向神经网络估计方面的多种性能, 同时着重研究了RBF神经网络在任意态以及高量子位本征态密度矩阵估计中的优势。

参考文献:

[1] CAHILL K E, GLAUBER R J. Density operators and quasi-probability distributions. *Physical Review*, 1969, 177(5): 1882 – 1902.

[2] GROSS D, LIU Y K, FLAMMIA S T, et al. Quantum state tomography via compressed sensing. *Physical Review Letters*, 2010, 105(15): 150401.

[3] GYONGYOSI L. Quantum imaging of high-dimensional hilbert spaces with radon transform. *International Journal of Circuit Theory and Applications*, 2017, 45(7): 1029 – 1046.

[4] STOKES G G. On the composition and resolution of streams of polarized light from different sources. *Transactions of the Cambridge Philosophical Society*, 1951, 9: 399 – 416.

[5] CAHILL K E, GLAUBER R J. Ordered expansions in boson amplitude operators. *Physical Review*, 1969, 177(5): 1857 – 1881.

[6] YANG Jingbei, CONG Shuang, CHEN Ding. Estimation of arbitrary quantum pure states based on a two-step measurement method and its minimum number of observations. *Control Theory & Applications*, 2017, 34(11): 1514 – 1521.
(杨靖北, 丛爽, 陈鼎. 基于两步测量方法及其最少观测次数的任意量子纯态估计. 控制理论与应用, 2017, 34(11): 1514 – 1521.)

[7] CONG Shuang, DING Jiao, ZHANG Kun. Improved iterative shrinkage threshold algorithm and its application to quantum state estimation. *Control Theory & Applications*, 2020, 37(7): 1667 – 1672.
(丛爽, 丁娇, 张坤. 改进的迭代收缩阈值算法及其在量子状态估计中的应用. 控制理论与应用, 2020, 37(7): 1667 – 1672.)

[8] CONG Shuang, WANG Tao, ZHANG Kun. Online quantum state estimation algorithm for matrix exponential gradient with adaptive learning rate. *Control Theory & Applications*, 2021, 38(8): 1188 – 1196.
(丛爽, 汪涛, 张坤. 带有自适应学习速率的矩阵指数梯度在线量子态估计算法. 控制理论与应用, 2021, 38(8): 1188 – 1196.)

[9] TORLAI G, MAZZOLA G, CARRASQUILLA J, et al. Neural-network quantum state tomography. *Nature Physics*, 2018, 14(5): 447 – 450.

[10] XU Q, XU S. Neural network state estimation for full quantum state tomography. *ArXiv Preprint*, 2018, arXiv: 1811.06654.

- [11] XIN T, LU D, KLASSEN J, et al. Quantum state tomography via reduced density matrices. *Physical Review Letters*, 2017, 118(2): 020401.
- [12] LOHANI S, KIRBY B T, BRODSKY M, et al. Machine learning assisted quantum state estimation. *Machine Learning Science and Technology*, 2020, 1(3): 035007.
- [13] LIN Wenrui, CONG Shuang. Quantum state estimation based on deep learning LDAMP network. *Acta Automatica Sinica*, 2023, 49(1): 79–90.
(林文瑞, 丛爽. 基于深度学习LDAMP网络的量子状态估计. 自动化学报, 2023, 49(1): 79–90.)
- [14] PAN Chengwei, CONG Shuang. Estimation of 2-bit quantum states based on width neural networks. *Proceedings of the 22nd China Simulation and Convergence Technology Academic Conference (CCSSTA 2021)*. Chongqing, China: 2021: 80–84.
(潘成威, 丛爽. 基于宽度神经网络的2比特量子态的估计. 第22届中国系统仿真技术及其应用学术年会(CCSSTA22nd 2021)论文集. 重庆, 中国: 2021: 80–84.)
- [15] JAMES D F V, KWIAT P G, MUNRO W J, et al. Measurement of qubits. *Physical Review A*, 2001, 64(5): 052312.
- [16] ZHANG Jiaojiao. *Research on quantum state estimation and filtering algorithm based on compressed sensing and its convergence*. Hefei: University of Science and Technology of China, 2018.
(张娇娇. 基于压缩感知的量子状态估计与滤波算法及其收敛性研究. 合肥: 中国科学技术大学, 2018.)

作者简介:

丛爽 教授, 博士生导师, 从事量子系统状态调控与实现的研究, E-mail: scong@ustc.edu.cn;

李友志 硕士, 主要研究方向基于神经网络的量子态估计, E-mail: lyouthz@mail.ustc.edu.cn.