

基于深度学习 LDAMP 网络的量子状态估计

林文瑞¹ 丛爽¹

摘要 设计出一种基于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, 将其应用于量子状态的估计. 该网络将去噪卷积神经网络与基于去噪的近似消息传递算法相结合, 利用量子系统输出的测量值作为网络输入, 通过设计出的带有去噪卷积神经网络的 LDAMP 网络重构出原始密度矩阵, 从大量的训练样本中提取各种不同类型密度矩阵的结构特征, 来实现对量子本征态、叠加态以及混合态的估计. 在对 4 个量子位的量子态估计的具体实例中, 分别在无和有测量噪声干扰情况下, 对基于 LDAMP 网络的量子态估计进行了仿真实验性能研究, 并与基于压缩感知的交替方向乘法法和三维块匹配近似消息传递等算法进行估计性能对比研究. 数值仿真实验结果表明, 所设计的 LDAMP 网络可以在较少的测量的采样率下, 同时完成对 4 种量子态的更高精度估计.

关键词 量子状态估计, 近似消息传递法, 压缩感知, 密度矩阵, 深度学习

引用格式 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

DOI 10.16383/j.aas.c210156

Quantum State Estimation Based on Deep Learning LDAMP Networks

LIN Wen-Rui¹ CONG Shuang¹

Abstract A learned denoising-based approximate message passing (LDAMP) deep learning network is proposed and trained in this paper, which is applied to the estimation of quantum states. This network combines denoising convolutional neural network with denoising-based approximate message passing algorithm. Using the measured output of the quantum system as the network input, the original density matrix was reconstructed by the designed LDAMP network with denoising convolutional neural network, and the structural features of various density matrices were extracted from a large number of training samples to realize the estimation of superposition and mixed states of quantum eigenstates. In the specific examples of quantum state estimation of four qubits, we study the performance of the quantum state estimation based on the LDAMP networks in the absence and presence of measurement interference, respectively, and compare the estimation performance with other algorithms which is based on compressed sensing such as alternating direction multiplier method and block matching 3D AMP. The numerical simulation results show that the LDAMP network can simultaneously estimate the four quantum states with higher accuracy in a small number of measurements.

Key words Quantum state estimation, approximate message passing, compressed sensing, density matrix, deep learning

Citation Lin Wen-Rui, Cong Shuang. Quantum state estimation based on deep learning LDAMP networks. *Acta Automatica Sinica*, 2023, 49(1): 79–90

在量子力学中, 量子状态密度矩阵包含了量子系统的全部信息, 是量子信息处理和量子控制研究的基础. 根据海森堡不确定性原理^[1], 对量子系统的每一次测量都会破坏其原有的状态, 这使得人们无法通过直接对一个量子系统进行测量来获得量子系统的状态, 只能测量到量子系统的状态在某个投影

方向上的投影 (坍缩) 结果, 再通过多次测量的统计结果, 计算出各投影分量的概率对真实量子状态进行估计^[2]. 为了解决量子态的估计问题, 1957 年 Fano^[3] 提出了量子层析的方法, 它是通过对一个量子状态进行多次投影测量, 根据测量值与量子态直接的关系式, 联立方程求解出量子状态. 该方法需要制备量子状态的大量全副本, 并进行重复地测量来进行量子态估计. 随着量子理论的逐步完善, 1969 年 Cahill 和 Glauber^[2] 提出利用重复测量获取的量子概率统计数据重构出量子态的密度矩阵来恢复量子状态信息, 奠定了量子态层析的基础. 一个 n 比特量子系统状态的密度矩阵 $\rho \in \mathbf{C}^{d \times d}$ 的维数为 $d \times d$, 并同时具有半正定, 单位迹和共轭对称的性

收稿日期 2021-02-21 录用日期 2021-06-25
Manuscript received February 21, 2021; accepted June 25, 2021
国家自然科学基金 (61973290, 61720106009) 资助
Supported by National Natural Science Foundation of China (61973290, 61720106009)
本文责任编辑 朱纪洪
Recommended by Associate Editor ZHU Ji-Hong
1. 中国科学技术大学自动化系 合肥 230027
1. Department of Automation, University of Science and technology of China, Hefei 230027

质, 其中 $d = 2^n$. 由于待估计的 ρ 的未知元素数目为 $d \times d = 4^n$ 个, 所以该系统的完备测量次数为 4^n , 是随着量子维数呈指数增加^[4]. 为了解决测量次数随量子位指数增长的问题, 人们根据压缩感知 (Compressive sensing, CS) 理论^[5] 将量子态密度矩阵通过高维测量矩阵投影到维数很低的压缩矩阵上, 只要测量矩阵满足限制等距特性 (Restricted isometry property, RIP)^[6], 就可以利用压缩后的低维压缩矩阵通过求解一个优化问题精确重构出待估计的密度矩阵^[7-9].

在实际量子测量过程中, 测量噪声的存在是不可避免的, 并且可以被假定为高斯噪声. 在量子状态估计算法研究中, Smith 等^[10] 基于最小二乘法从含有噪声的数据中实现了 4 比特量子状态估计. Li 等^[11] 基于压缩感知将交替方向乘法 (Alternating direction multiplier method, ADMM) 运用到量子态估计中^[12], 给出了算法迭代优化形式, 并在 5 比特的量子状态密度矩阵估计中获得了较高的估计精度. ADMM 算法是一种求解分布式凸优化问题的计算框架, 在图像处理及恢复、统计学习和量子态估计等领域有着广泛应用. Smolin 等^[13] 提出了一种从加性高斯噪声测量中计算最大似然量子态的有效方法, 其最坏时间复杂度为 $O(d^4)$. 但是这些优化算法在量子态重建过程中都会需要较多的计算内存和计算时间. 对于高维量子系统, 重建过程中需要测量和处理的数据会很多, 这样会导致所需的计算资源和计算时间增长很快. Metzler 等^[14] 提出 DAMP 算法, 它是一种基于压缩感知的迭代稀疏信号恢复算法, 通过在算法中添加 Onsager 校正项^[14-15], 修正每次迭代时有效噪声的偏差, 改进了基于去噪的迭代阈值 (Denoising-based iterative threshold, DIT) 算法收敛速度慢的问题. Liu 等^[16] 采用卷积神经网络的降噪自动编码器, 从带有噪声的密度矩阵中重建出高精度的密度矩阵.

本文基于压缩感知理论, 结合深度学习网络, 提出一种采用深度学习网络来估计量子状态密度矩阵的方法. 将密度矩阵视为二维自然图像信号, 采用基于学习降噪的近似消息传递 (Learned denoising-based approximate message passing, LDAMP) 神经网络^[17], 将去噪卷积神经网络 (Denoising convolutional neural network, DnCNN) 的图像降噪器^[18] 融合到基于去噪的近似消息传递 (Denoising-based approximate message passing, DAMP) 算法^[19] 中. 此外, 还提供关于 LDAMP 在量子状态估计中的渐近收敛性能的分析框架—状态演化 (State

evolution, SE) 方程^[17] 来预测 LDAMP 网络每一层的均方误差 (Mean square error, MSE) 性能^[20]. 在数值仿真实验中, 在固定 DnCNN 神经网络和 LDAMP 网络的层数, 选择高斯测量矩阵, 在无和有外部测量噪声两种情况下, 分别对 4 个量子位的量子本征态、叠加态、对角混合态和一般混合态的密度矩阵参数估计进行结果和性能分析, 并与其他典型的重建算法对密度矩阵参数估计的性能进行了对比研究.

本文结构安排如下: 第 1 节概述了基于压缩感知的量子状态估计作为一个信号恢复问题, 第 2 节介绍了应用于量子状态估计的 DAMP 信号恢复算法, 第 3 节具体描述了 LDAMP 网络的实现和 DnCNN 神经网络的结构及其训练过程, 第 4 节给出了对 DnCNN 训练后模型的测试结果和量子态估计数值模拟实验结果及分析, 第 5 节是结论.

1 基于压缩感知的量子态密度矩阵估计

压缩感知是近 10 多年来在信号处理领域诞生的一种新的信号恢复理论. 它主要是利用信号 $\mathbf{s} \in \mathbf{C}^N$ 的 M (远远小于完备测量需要的 N) 个随机线性测量值 \mathbf{y} , 建立欠定线性方程组 $\mathbf{y} = \mathbf{A}\mathbf{s}$ ($\mathbf{y} \in \mathbf{C}^M$), 并通过优化算法来高精度重构求解出原始信号 \mathbf{s} , 其中 $\mathbf{A} \in \mathbf{C}^{M \times N}$, $M < N$ 为测量矩阵. 在量子态估计中, 原始信号为密度矩阵 $\rho \in \mathbf{C}^{d \times d}$, 它与量子系统波函数之间的关系为:

$$\rho = \sum_{i=1}^d |\psi_i\rangle p_i \langle \psi_i| \quad (1)$$

式中, $|\psi_i\rangle$ 为系统的波函数, p_i 表示波函数的概率, d 为密度矩阵的维数. 将密度矩阵写成向量表达的形式是 $\rho = \boldsymbol{\psi} \cdot \boldsymbol{\psi}^*$, 其中 $\boldsymbol{\psi} = [\sqrt{p_1} |\psi_1\rangle, \sqrt{p_2} |\psi_2\rangle, \dots, \sqrt{p_d} |\psi_d\rangle]$. 当量子系统的量子位数为 n 时, 密度矩阵的维数 $d = 2^n$. 由于神经网络无法直接处理复数信号的去噪, 在本文中仅对实数密度矩阵进行估计. 在基于压缩感知的量子态估计中, 测量矩阵 \mathbf{A} 需要满足限制等距特性才可以唯一确定 d^2 个待估计的密度矩阵元素, 人们通常采用的是泡利测量矩阵^[21]. 由于随机高斯矩阵参数均为实数并且具有较好的重构效果^[22], 本文中在 RIP 条件下采用随机高斯矩阵来构造测量矩阵, 其设计方法为: 构造一个 $M \times N$ 大小的矩阵 \mathbf{A} , 使 \mathbf{A} 中的每一个元素独立的服从均值为 0, 方差为 $1/M$ 的高斯分布, 即:

$$\mathbf{A}(i, j) \sim \mathbf{N}(0, 1/M) \quad (2)$$

式中, \mathbf{A} 的行数与列数的比值定义为采样率 $\eta =$

M/N , 并且有 $N = d^2$.

考虑从有外部测量干扰的线性度量中恢复密度矩阵 $\rho \in \mathbf{R}^{d \times d}$ 的问题:

$$\mathbf{y} = \mathbf{A} \cdot \text{vec}(\rho) + w \quad (3)$$

式中, $\text{vec}(\cdot)$ 表示将一个矩阵按列的方向组合成一个列向量, $w \in \mathbf{R}^m$ 为测量噪声并且通常被假设为独立且同分布的高斯噪声. 定义测量噪声 $w \sim \mathbf{N}(\mathbf{0}, \gamma_w^2 \mathbf{I})$ 为加性高斯白噪声 (Additive white Gaussian noise, AWGN), 其中 γ_w 为噪声标准差. 在本文的模拟实验中, 采用含测量噪声信号 \mathbf{y} 和理想测量信号 $\mathbf{y}_0 = \mathbf{A} \cdot \text{vec}(\rho)$ 之间的信噪比 (Signal noise ratio, SNR) 作为噪声强度, 计算公式为:

$$\text{SNR} = 10 \lg \left\{ \frac{\sum_{i=1}^M \sum_{j=1}^N \mathbf{y}^2(i, j)}{\sum_{i=1}^M \sum_{j=1}^N [\mathbf{y}(i, j) - \mathbf{y}_0(i, j)]^2} \right\} \quad (4)$$

式中, $\mathbf{y}(i, j)$ 和 $\mathbf{y}_0(i, j)$ 分别为含测量噪声信号和理想测量信号在向量中 (i, j) 点处的值.

利用正则化二次损失最小化方法解决该线性逆问题, 可以通过求解一个凸优化问题计算得到 $\text{vec}(\rho)$ 的估计值^[23-24]:

$$\text{vec}(\hat{\rho}) = \arg \min_{\text{vec}(\rho) \in \mathbf{R}^{d^2}} \left\{ \frac{1}{2} \|\mathbf{y} - \mathbf{A} \cdot \text{vec}(\rho)\|_2^2 + \lambda \|\text{vec}(\rho)\|_1 \right\} \quad (5)$$

式中, $\lambda > 0$ 为正则化参数. 然而这个极小化问题是一个凸优化问题, 在大规模问题中, 由于数据维度较大, 常规求解算法中, 如内点法等算法复杂度较大, 求解起来非常耗时. 这使得众多研究人员通过简单的基于梯度的方法来求解式 (5).

2 DIT 和 DAMP 算法

在众多基于梯度的算法中, 最受关注的一类方法是迭代阈值 (Iterative threshold, IT) 算法. 为了方便表示, 定义 $\mathbf{s} = \text{vec}(\rho)$. 当待估计的密度矩阵为 ρ_0 时, 称 $\mathbf{s}_0 = \text{vec}(\rho_0)$ 为待估计的信号. 从 $\mathbf{s}^0 = \mathbf{0}$ 和 $\mathbf{z}^0 = \mathbf{0}$ 开始, 迭代公式为:

$$\mathbf{s}^{l+1} = \eta(\mathbf{s}^l + \mathbf{A}^H \mathbf{z}^l; \lambda^l) \quad (6a)$$

$$\mathbf{z}^l = \mathbf{y} - \mathbf{A} \mathbf{s}^l \quad (6b)$$

式中, \mathbf{s}^l 是 l 次迭代密度矩阵的估计值. \mathbf{z}^l 是 l 次迭代的残差, 如果密度矩阵恢复过程是无噪声的, 并且算法性能执行得很好, 它将收敛于 $\mathbf{0}$. $\eta(\cdot)$ 称为阈

值函数, 它是一个被应用到矢量的分量元素上的标量非线性函数. λ^l 是 l 次迭代阈值函数使用的参数.

$$\eta_{\text{soft}}(\mathbf{x}; \lambda) = \text{sign}(x_i)(|x_i| - \lambda)_+ \quad (7)$$

在信号恢复中, 阈值函数的选择非常重要, 常见的有软阈值函数 (7) 等. 其中符号 $(w)_+$ 表示 $w \geq 0$ 时结果为 w , 否则结果为 0. 另外一种实现阈值函数的方法是使用降噪器 $D_{\hat{\sigma}^l}$, 此时阈值参数 λ^l 的含义为输入降噪器的信号的所含高斯噪声的标准差估计 $\hat{\sigma}^l$, 并且计算公式为 $\hat{\sigma}^l = \|\mathbf{z}^l\|_2 / \sqrt{M}$, 它随着迭代次数的变化而变化. 利用降噪器实现阈值函数的迭代阈值算法称为基于去噪的迭代阈值 (Denoising-based iterative threshold, DIT) 算法^[17], 迭代公式为:

$$\mathbf{s}^{l+1} = D_{\hat{\sigma}^l}(\mathbf{s}^l + \mathbf{A}^H \mathbf{z}^l) \quad (8a)$$

$$\mathbf{z}^l = \mathbf{y} - \mathbf{A} \mathbf{s}^l \quad (8b)$$

为了解释降噪器在 DIT 算法中所发挥的作用, 定义 \mathbf{v}^l 为有效噪声, 它表示降噪器的输入信号 $(\mathbf{s}^l + \mathbf{A}^H \mathbf{z}^l)$ 和待估计信号 \mathbf{s}_0 之间的差: $\mathbf{v}^l = \mathbf{s}^l + \mathbf{A}^H \mathbf{z}^l - \mathbf{s}_0$. 通过迭代式 (8a) 和 $\mathbf{s}^l + \mathbf{A}^H \mathbf{z}^l = \mathbf{s}_0 + \mathbf{v}^l$ 可得 $\mathbf{s}^{l+1} = D_{\hat{\sigma}^l}(\mathbf{s}^l + \mathbf{A}^H \mathbf{z}^l) = D_{\hat{\sigma}^l}(\mathbf{s}_0 + \mathbf{v}^l)$, 相当于对在含有有效噪声 \mathbf{v}^l 的 $(\mathbf{s}_0 + \mathbf{v}^l)$ 的信号中, 求得真实信号 \mathbf{s}_0 的估计信号 \mathbf{s}^{l+1} . 在 DIT 算法中, 大多数降噪器都默认 \mathbf{v}^l 作为加性高斯白噪声. 不过, 随着迭代次数的增加, 输入降噪器的信号会产生偏差, 会导致 \mathbf{v}^l 不再满足 AWGN, 使得迭代算法性能降低并且估计值收敛到真实值的速度变的很慢^[15, 19, 25]. 为了解决这个问题, 在信号恢复中, 采用基于去噪的近似消息传递 DAMP 算法对 DIT 算法进行改进. DAMP 算法通过增加一个 Onsager 校正项 \mathbf{b}^l ($l = 0, 1, 2, \dots$), 来修正 DIT 算法 l 次迭代中有效噪声 \mathbf{v}^l 与 AWGN 的偏差, 从而提高算法恢复性能和解决算法收敛慢的问题. DAMP 算法如下所示.

算法 1. DAMP 算法

要求. 高斯测量矩阵 $\mathbf{A} \in \mathbf{R}^{M \times N}$, 测量值向量 $\mathbf{y} \in \mathbf{R}^M$, 待估计信号 $\mathbf{s}_0 = \text{vec}(\rho_0)$.

1. 初始化变量 $\mathbf{s}^0 = \mathbf{0}$, 残差 $\mathbf{z}^0 = \mathbf{y}$, $\hat{\sigma}^0 = \|\mathbf{z}^0\|_2 / \sqrt{M}$, 迭代次数为 L ;
2. for $l = 0, 1, 2, \dots, (L - 1)$ do;
3. $\mathbf{s}^{l+1} = D_{\hat{\sigma}^l}(\mathbf{s}^l + \mathbf{A}^H \mathbf{z}^l)$;
4. $\mathbf{b}^{l+1} = \mathbf{z}^l \text{div} D_{\hat{\sigma}^l}(\mathbf{s}^l + \mathbf{A}^H \mathbf{z}^l) / M$;
5. $\mathbf{z}^{l+1} = \mathbf{y} - \mathbf{A} \mathbf{s}^{l+1} + \mathbf{b}^{l+1}$;
6. $\hat{\sigma}^{l+1} = \|\mathbf{z}^{l+1}\|_2 / \sqrt{M}$;
7. end for;

8. 返回 L 次迭代后的估计信号 s^L , 从而得到 ρ_0 的估计值.

其中, $\text{div}D_{\hat{\sigma}^l}(\cdot)$ 是散度运算, \mathbf{b}^l 为 Onsager 修正项, 用来修正中间解 z^l 的偏差, 使得有效噪声 \mathbf{v}^l 符合去噪所期望的 AWGN 模型.

散度 $\text{div}D_{\hat{\sigma}^l}(\cdot)$ 是一个不易计算的项, 人们一般采用蒙特卡洛近似来计算散度^[19, 26]. 对于一个训练好的降噪器 $D_{\hat{\sigma}^l}(\cdot)$, 利用一个独立的同分布随机向量 $\mathbf{d} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$, 来近似计算散度:

$$\begin{aligned} \text{div}D_{\hat{\sigma}^l}(s^l + \mathbf{A}^H z^l) &= \\ \lim_{\epsilon \rightarrow 0} \mathbb{E}_{\mathbf{d}} \left\{ \mathbf{d}^T \left(\frac{D_{\hat{\sigma}^l}(s^l + \epsilon \mathbf{d}) - D_{\hat{\sigma}^l}(s^l)}{\epsilon} \right) \right\} &\approx \\ \frac{1}{\epsilon} \mathbf{d}^T (D_{\hat{\sigma}^l}(s^l + \epsilon \mathbf{d}) - D_{\hat{\sigma}^l}(s^l)) &\quad (9) \end{aligned}$$

式中, ϵ 是一个很小的数, 设置 $\epsilon = \|\mathbf{s}^l\|_{\infty}/1000$.

基于不同的降噪器, DAMP 包含许多算法^[27], 包括高斯-AMP (Gauss-AMP)、双边-AMP (Bilateral-AMP)、非局部平均-AMP (Non-local means AMP, NLM-AMP) 和三维块匹配-AMP (Block matching 3D AMP, BM3D-AMP)^[28] 等. 本文采用的降噪器 $D_{\hat{\sigma}^l}$ 和散度 $\text{div}D_{\hat{\sigma}^l}(\cdot)$ 计算都是通过设计并训练一个深度卷积神经网络 DnCNN 来实现. 同时通过设计一个学习的基于去噪的近似消息传递 LDAMP 网络, 来实现算法 1 的 DAMP 算法, 并通过多个相同结构的级联, 来获得高精度的信号 \mathbf{s}_0 即 $\text{vec}(\rho_0)$ 的估计, 得到密度矩阵 ρ_0 的估计.

本文通过 LDAMP 网络, 使用量子系统输出的测量值, 来实现对量子态密度矩阵估计, 具体流程如下:

步骤 1. 根据 $\rho = \varphi_r \varphi_r^H / \text{tr}(\varphi_r \varphi_r^H)$ 生成 DnCNN 神经网络的训练集, 验证集和测试集, 具体过程在第 3.3 节;

步骤 2. 在 9 个不同噪声标准差范围下, 训练出相应的 DnCNN 神经网络, 训练完成后利用测试集对网络进行测试;

步骤 3. 对于待估计的密度矩阵 ρ_0 , 在不同采样率下利用高斯测量算符获取的输出测量值 \mathbf{y}_0 . 对于含噪声情况下测量值, 通过 $\mathbf{y} = \text{awgn}(\mathbf{y}_0, \text{SNR})$ 添加固定 SNR 数值的高斯噪声获得好噪声测量值 \mathbf{y} ;

步骤 4. 设计一个包含训练好的 DnCNN 网络和实现 DAMP 算法的 10 层 LDAMP 网络, 网络各层中的 DnCNN 是根据当前层的噪声标准差估计 $\hat{\sigma}$ 落入的 9 个范围来选定. LDAMP 网络的输入为测量值 \mathbf{y}_0 或 \mathbf{y} ;

步骤 5. LDAMP 网络最后的输出为根据输入测量值 \mathbf{y}_0 或 \mathbf{y} 得到的估计的密度矩阵 $\hat{\rho}_0$.

3 LDAMP 和 DnCNN 网络的结构设计及其训练

第 3.1 节对 LDAMP 网络的结构进行设计; 第 3.2 节进行 DnCNN 降噪器及其中的卷积神经网络的结构设计; 第 3.3 节进行 DnCNN 降噪器中神经网络的训练过程.

3.1 LDAMP 网络的结构设计

LDAMP 网络由 L 层相同的 AMP 网络结构级联而成, 其中, LDAMP 网络单层的结构图如图 1 所示. 每层网络包含两个在固定噪声标准差区间下训练好的 DnCNN 降噪器, 一个用于计算每层 AMP 网络的输出 s^{l+1} , 另一个用于实现散度 $\text{div}D_{\hat{\sigma}^l}(\cdot)$. 图 1 中的 $x^l = s^l + \mathbf{A}^H z^l$ 为 LDAMP 网络的第 $(l+1)$ 层中 DnCNN 降噪器的输入信号, $\hat{\mathbf{v}}^l$ 为降噪器中 DnCNN 卷积神经网络的输出, 它表示对有效噪声 \mathbf{v}^l 的估计.

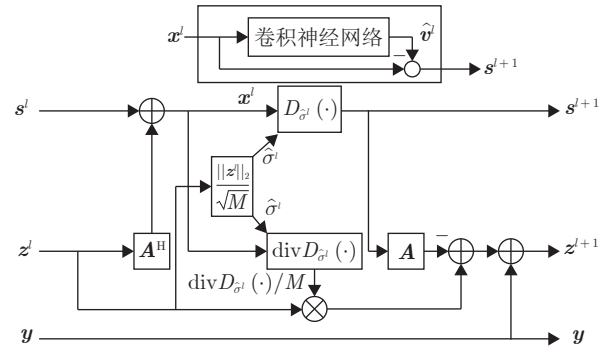


图 1 LDAMP 中第 l 级网络结构图

Fig.1 Structure of the l -level network in the LDAMP

需要指出的是, LDAMP 网络中只有 DnCNN 降噪器是由神经网络设计而成, 而网络的其他部分只是对 DAMP 算法的实现.

对于 L 层的 LDAMP 网络, 可以使用不同的方法来训练, 包括端到端训练, 逐层训练和逐降噪器训练, 其中, 端到端训练是指将 L 层的 LDAMP 网络视作一个整体, 同时训练网络的所有权值. 这是训练神经网络的标准方法. 逐层训练是指每次训练 LDAMP 网络中的一层 AMP 网络来恢复信号, 固定这些权值, 然后添加一层 AMP 网络, 训练产生的 2 层网络的第 2 层来恢复信号, 再固定这些权值, 重复操作直到已经训练一个 L 层 LDAMP 网络. 逐降噪器训练是指将 DnCNN 降噪器从网络中

单独出来,并在不同噪声标准差范围区间下训练每个降噪器的 AWGN 去噪问题,从而得到多个不同去噪水平的降噪器,以供 LDAMP 网络中的各层选择,即每层 AMP 网络,根据当前层的噪声标准差的估计,来选择使用哪一个降噪器。

本文采用的逐降噪器训练方法是最小均方误差 (Minimum mean squared error, MMSE)^[17] 最优的. 具体训练过程为: 首先,对噪声水平的预期范围进行离散化,将噪声标准差 $[0, 2]$ 范围内的噪声分成多个小范围;然后,在这些小范围噪声水平下训练 DnCNN 降噪器,得到多个训练好的降噪器模型;最后,LDAMP 网络实现时,每层网络根据本层的噪声标准差估计 $\hat{\sigma}$ 落在某个区间内,来选择在该噪声标准差区间下训练好的降噪器模型. LDAMP 网络易于训练,可以应用于各种不同的测量矩阵,并带有一个可以准确预测其性能的状态演化框架。

状态演化 (State evolution, SE) 是一个可以预测 LDAMP 网络各级的输出与真实信号之间 MSE 性能的分析框架. 它由一系列 SE 方程组成:

$$\theta^{l+1}(\mathbf{s}_0, \eta, \gamma_w^2) = \frac{1}{N} \mathbb{E}_w \left\| D_{\sigma^l}(\mathbf{s}_0 + \sigma^l w) - \mathbf{s}_0 \right\|_2^2 \quad (10)$$

$$(\sigma^l)^2 = \frac{1}{\eta} \theta^l(\mathbf{s}_0, \eta, \gamma_w^2) + \gamma_w^2 \quad (11)$$

式中, \mathbf{s}_0 是信号 \mathbf{s} 的确定性实现, $\eta = M/N$ 表示采样比率, 标量 γ_{mw}^2 是测量噪声 w 的方差, 符号 $\theta^l(\mathbf{s}_0, \eta, \gamma_w^2)$ 为 LDAMP 网络的第 l 层去噪输出的平均 MSE, 且 $\theta^0 = \|\mathbf{s}_0\|_2^2/N$. 在一个大系统 ($N \rightarrow \infty$) 中, SE 方程可以从 LDAMP 网络本身来解释. 这种解释主要依赖于降噪器输入 $\mathbf{x}^l = \mathbf{s}^l + \mathbf{A}^H \mathbf{z}^l = \mathbf{s}_0 + \mathbf{v}^l$ 中等效的 AWGN 模型, 其中降噪器输出的平均 MSE 是通过式 (9) 中的蒙特卡罗方法计算的. 式 (10) 中的期望只涉及到分布 $w \sim \mathcal{N}(\mathbf{0}, \gamma_w^2 \mathbf{I})$, 有效噪声 \mathbf{v}^l 的方差 $(\sigma^l)^2$ 由式 (11) 计算得到. 通过递推

更新式 (10) 和式 (11), 将根据 SE 方程所得到的理论性能, 与本文提出的 LDAMP 网络的平均 MSE 性能进行对比。

3.2 DnCNN 降噪器的结构设计

在 DnCNN 神经网络结构设计上, DnCNN 网络源于对 VGG 网络的改进, 去除了池化层和全连接层, 在卷积层 (Convolution, Conv) 和修正线性单元 (Rectified linear unit, ReLU)^[29] 之间加入了批归一化 (Batch normalization, BN)^[30]. 这种改变可以减轻内部协变量移位^[18] 带来的影响, 从而提高网络的学习收敛速度. 在模型学习中, DnCNN 借鉴了 ResNet 的残差学习方法^[31-33], 来提高训练速度和去噪性能。

图 2 为 DnCNN 降噪器的网络结构图, 图 3 是 DnCNN 降噪器输入变量的尺寸变换过程图. 输入矢量的维数为 $d \times d$, 其中 $d = 2^n$, n 为量子态的位数. 深度为 D 的 DnCNN 神经网络, 在本文应用中, D 取值为 20. 本文设计的 DnCNN 有三种不同的网络层, 具体排列顺序为: 第 1 层为卷积层 + 修正线性单元 (Conv + ReLU): 使用 64 个大小为 $3 \times 3 \times 1$ 的卷积核 (或称为滤波器), 生成 64 个特征图. 然后, 使用修正线性单元进行非线性处理. 第 2 ~ ($D-1$) 层为卷积层 + 批归一化 + 修正线性单元 (Conv + BN + ReLU), 使用 64 个大小为 $3 \times 3 \times 64$ 的卷积核, 在卷积层和 ReLU 层之间加入 BN 层. 最后一层为卷积层 (Conv), 使用 1 个大小为 $3 \times 3 \times 64$ 的卷积核重构输出。

DnCNN 神经网络的输入是有噪声的信号矩阵 $\mathbf{x} \in \mathbf{R}^{d \times d}$, 所以 LDAMP 网络各层把数据输入到 DnCNN 神经网络之前, 需要将列向量变成原始信号大小的矩阵形式, 变换过程如图 3 所示. 含噪声输入的矩阵 \mathbf{x} 和向量 \mathbf{x}^l 分别为:

$$\mathbf{x} = \boldsymbol{\rho}_0 + \mathbf{v} \quad (12)$$

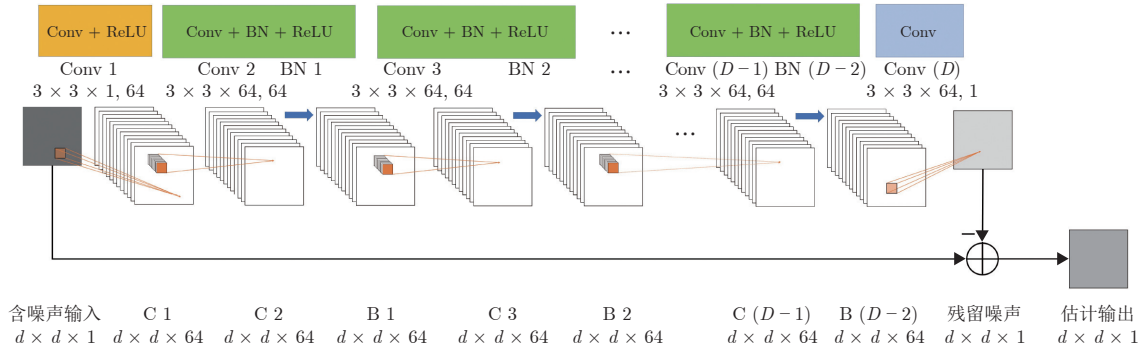


图 2 DnCNN 降噪器的网络结构图

Fig.2 Network structure of the DnCNN denoiser

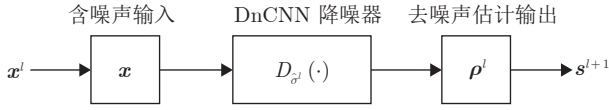


图 3 DnCNN 降噪器输入变量的尺寸变换过程

Fig. 3 Size transformation process of input variable of the DnCNN denoiser

$$\mathbf{x}^l = \mathbf{s}_0 + \mathbf{v}^l \quad (13)$$

式中, $\mathbf{x} \in \mathbf{R}^{d \times d}$ 是输入向量 $\mathbf{x}^l \in \mathbf{R}^{d^2}$ 转化成的矩阵形式, $\boldsymbol{\rho}_0 \in \mathbf{R}^{d \times d}$ 是真实信号 $\mathbf{s}_0 \in \mathbf{R}^{d^2}$ 的矩阵形式, $\mathbf{v} \in \mathbf{R}^{d \times d}$ 是有效噪声 $\mathbf{v}^l \in \mathbf{R}^{d^2}$ 转化成的矩阵形式. 不同于其他去噪方法, 如多层感知器 (Multi-layer perceptron, MLP) 和收缩域级联 (Cascade of shrinkage fields, CSF) 算法^[34] 是学习一个映射函数 $\mathcal{F}(\mathbf{x}) = \mathbf{s}$ 来估计真实的信号 \mathbf{s}_0 , DnCNN 神经网络采用残差学习规则训练映射 $\mathcal{R}(\mathbf{x}) = \hat{\mathbf{v}}$ 来估计真实噪声信号 \mathbf{v} , 然后得到的信号估计为 $\mathbf{s} = \mathbf{x} - \mathcal{R}(\mathbf{x})$. DnCNN 神经网络的损失函数表示为:

$$l(\Theta) = \frac{1}{2Q} \sum_{i=1}^N \|\mathcal{R}(x_i; \Theta) - (x_i - s_i)\|_{\text{F}}^2 \quad (14)$$

它是估计噪声信号 $\mathcal{R}(\mathbf{x}; \Theta)$ 与真实噪声信号 $\mathbf{v} = \mathbf{x} - \mathbf{s}_0$ 之间的均方误差. 其中, Q 是批大小 (Batch Size), Θ 是 DnCNN 神经网络中可以调整的参数, 包括网络的权值和偏置, 训练次数和目标误差等.

在许多信号恢复算法中, 通常要求输出信号的尺寸与输入信号的尺寸一致, 这可能会导致恢复的矩阵信号中边界元素值受到较大影响, 也就是产生边界伪影. 为了保证隐含层的每个特征信号与输入信号的尺寸大小相同, 在 DnCNN 神经网络的每个卷积层中进行卷积运算之前, 对输入卷积层的信号采用简单的零填充策略, 不过这种零填充策略不会产生任何边界伪影^[18].

3.3 DnCNN 神经网络训练过程

在对 DnCNN 神经网络进行训练时, 采用的样本数据的尺寸大小为 $2^4 \times 2^4 = 16 \times 16$, 也就是 4 个量子位的密度矩阵. 对样本数据加入不同标准差范围下的高斯噪声作为训练输入, 原始样本数据作为训练输出目标, 在未知高斯噪声水平下训练 DnCNN 神经网络. 为了训练更小噪声范围下的 DnCNN 网络, 将噪声标准差范围 $[0, 2]$ 划分成 9 个小范围, 具体分段结点为 0、0.005、0.01、0.03、0.05、0.1、0.3、0.5、1 和 2, 并在这 9 个小范围噪声水平下分别训练了 DnCNN 网络. 然后 LDAMP 网络根据各级的噪声标准差估计 $\hat{\sigma}$ 的大小选择对应的

DnCNN 模型. 训练和测试过程中, 样本数据包含满足和近似满足量子态密度矩阵性质的矩阵, 其中满足性质的两种量子态的密度矩阵生成公式为:

$$\boldsymbol{\rho} = \frac{\boldsymbol{\varphi}_r \boldsymbol{\varphi}_r^{\text{H}}}{\text{tr}(\boldsymbol{\varphi}_r \boldsymbol{\varphi}_r^{\text{H}})} \quad (15)$$

式中, $\boldsymbol{\varphi}_r$ 为 $d \times r$ 的 Wishart 矩阵, 并且矩阵内各元素满足高斯分布^[35]. $r = 1$ 时, 为叠加态密度矩阵, $r > 1$ 时为一般混合态密度矩阵.

本文研究的另外两种量子态还包括本征态和对角混合态, 其中本征态密度矩阵为仅有对角线某一位置有元素为 1, 其他位置元素均为 0 的矩阵, 对角混合态密度矩阵为仅在对角线上所有位置有非 0 元素, 且对角线所有元素之和为 1, 其他位置元素均为 0 的矩阵. 样本数据中近似满足性质的矩阵是指在构造样本数据时, 改变部分满足性质的密度矩阵中的元素, 加入一些噪声干扰, 使矩阵的迹在 $[0, 1]$ 之间且近似共轭对称的半正定矩阵. 在构造过程中, 采用 $[0, 1]$ 之间的随机数乘以原始密度矩阵或在原始矩阵生成过程中给矩阵 $\boldsymbol{\varphi}_r$ 的共轭转置矩阵 $\boldsymbol{\varphi}_r^{\text{H}}$ 添加噪声干扰. 值得注意的是, 训练集和测试集中近似满足性质的密度矩阵所占的比例在 30% ~ 50% 可以取得较好的估计效果, 另外 $[0, 0.5]$ 之间的随机数应多于 $(0.5, 1]$ 之间的随机数. 这种构造方法的原因是因为 LDAMP 网络具有数层相同的结构, 网络的初始输入 $\mathbf{s}^0 = \text{vec}(\boldsymbol{\rho}^0)$ 为零向量, 网络的最终输出 $\hat{\boldsymbol{\rho}}^L$ 为估计的密度矩阵, 所以网络进行密度矩阵估计是各层网络渐近逼近真实密度矩阵的过程, 因此需要大量的中间估计结果和少量的零矩阵包含在训练样本集中.

在训练 DnCNN 时, 本文使用了 $9\,900 \times 128 = 1\,267\,200$ 个样本用于训练, $1\,100 \times 128 = 140\,800$ 个样本用于验证和 $11 \times 128 = 1\,408$ 个样本矩阵用于测试. 训练集、验证集及测试集中 4 种量子态密度矩阵的比例为本征态: 对角混合态: 叠加态: 一般混合态 = 1:2:4:4. 所有的样本矩阵在生成时都是无噪声样本, 而在训练、验证和测试时, 会在对应的噪声水平范围内加入噪声. 为了获取足够的空间信息进行去噪, 将 DnCNN 神经网络的深度设置为 20 层. 采用了一种鲁棒的初始化方法 MSRA^[36] 来初始化 DnCNN 神经网络权值, 并采用权值衰减为 0.0001、动量为 0.9、小批量尺寸为 128 的带动量的随机梯度下降法 (Stochastic gradient descent with momentum, SGDM) 训练网络. 为每个小范围噪声水平下的神经网络设置的最大训练次数为 50, 最大允许无用训练次数 (Max bad epochs, MBE) 为 2, 初始无用训练次数 (Bad epochs, BE) 为 0. 在训练过

程中, 初始最小验证误差采用验证集验证未训练的网络得到, 后续过程中, 网络训练一次就验证一次得到新的验证误差, 并将新的验证误差与上一次的验证误差相比, 如果小于上一次的验证误差就将此次验证误差设置为最小验证误差并将 BE 清零, 否则最小验证误差不变并将 BE 加 1. 当 BE 大于等于 MBE 时, 网络就减小学习速率, 并将 BE 清零. 学习速率变化过程为初始值 0.001 衰减为 0.0001, 最后衰减至 0.00001. 当学习速率衰减为 0.00001 且 BE 大于等于 MBE 时, DnCNN 神经网络的训练结束. 在实验中, 使用 Matlab 中的 MatCovNet 工具箱^[37]来实现 DnCNN 神经网络. 所有实验均在 Intel(R) Core(TM) i7-8700K CPU 3.70 GHz、NVIDIA GeForce GTX 1080 GPU 的 PC 上运行的 Matlab (R2018b) 环境下进行. 所有 9 个小范围噪声水平下的 20 层 DnCNN 网络的 GPU 训练时间需要 59 小时左右.

4 量子态估计仿真实验及其结果分析

本节对本文提出的 LDAMP 网络的量子态估计性能, 在不同参数下, 进行了数值性能对比实验, 进行 3 个实验:

1) 不同噪声标准差区间下, 对所设计训练出的 20 层 DnCNN 神经网络降噪器去噪性能的测试实验;

2) 本文的 LDAMP 网络与其他 5 种不同方法对量子态密度矩阵估计的性能对比实验;

3) 不同测量噪声以及不同采样率分别对 LDAMP 网络估计量子态密度矩阵性能的影响实验.

实验中, 量子状态测量值 \mathbf{y} 是根据式 (3) 计算: $\mathbf{y} = \mathbf{A} \cdot \text{vec}(\rho_0) + \mathbf{w}$, 其中 \mathbf{A} 为随机高斯测量矩阵, $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \gamma_w^2 \mathbf{I})$ 表示测量过程中产生的高斯噪声, ρ_0 为真实密度矩阵, 可以从本征态、叠加态、对角混合态和一般混合态密度矩阵中选取. 定义了三个性能指标来衡量实验结果:

1) 估计出来的密度矩阵 $\hat{\rho}$ 与真实密度矩阵 ρ_0 之间的归一化距离 $D(\hat{\rho}, \rho_0)$, 定义为:

$$D(\hat{\rho}, \rho_0) = \|\hat{\rho} - \rho_0\|_F^2 / \|\rho_0\|_F^2 \quad (16)$$

当 $D(\hat{\rho}, \rho_0)$ 大于 1 时, 取 1, 范围是 $[0, 1]$; 并且当 2 个矩阵完全相同时, $D(\hat{\rho}, \rho_0)$ 为 0.

2) 估计的密度矩阵与真实密度矩阵之间的均方误差 MSE, 定义为:

$$MSE = E_w \left\{ \|\hat{\rho} - \rho_0\|_2^2 / d^2 \right\} \quad (17)$$

当 MSE 以 dB 作为单位时, 定义为:

$$MSE(\text{dB}) = 10 \lg(MSE) \quad (18)$$

3) 估计的密度矩阵与真实密度矩阵之间的保真度 (Fidelity)^[38]:

$$Fidelity = \text{tr} \left(\sqrt{\sqrt{\rho_0} \hat{\rho} \sqrt{\rho_0}} \right) \quad (19)$$

式中, 保真度的范围是 $[0, 1]$, 并且当 2 个矩阵完全相同时, 保真度为 1^[39].

4.1 DnCNN 去噪性能的测试实验

图 4 为使用 $11 \times 128 = 1408$ 个随机样本矩阵, 分别测试噪声水平分别从 $[0, 0.005]$, $(0.005, 0.03]$, \dots , $(1, 2]$ 共 9 个范围情况下, 对训练后得到的 20 层 DnCNN 进行去噪性能的测试实验的误差曲线图, 其中方形实线表示网络输入样本与对应的无噪声样本之间的平均均方误差 MSE (dB); 圆圈虚线为网络输出样本与对应的无噪声样本之间的平均均方误差 MSE (dB); 三角形实线是输出误差与输入误差之间的 MSE 比值变化曲线. 由图 4 可以看出, 网络输出误差始终小于输入误差, 并且随着噪声标准差的增加, 输出误差与输入误差之间的比值逐渐减小: 从 $[0, 0.005]$ 区间的 0.5956 减小至 $(1, 2]$ 区间的 0.0311. 这说明, 当输入噪声很小时, 输出误差仍然能够滤掉接近一半噪声, 随着噪声的增加, 去噪效果越明显. 表明经过训练后的 DnCNN 降噪器拥有良好的去噪效果.

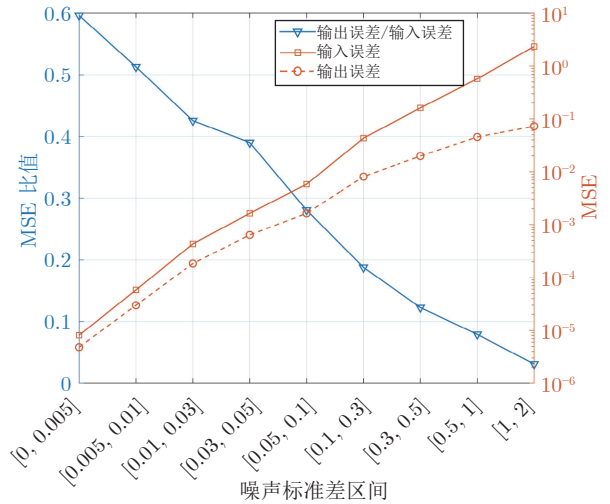


图 4 DnCNN 降噪器的 MSE 性能
Fig. 4 MSE performance of the DnCNN denoiser

4.2 LDAMP 网络与其他方法的对比实验

本节中将本文的 LDAMP 网络与 ADMM 算法, 以及其他 4 种 DAMP 算法 (BM3D-AMP、NLM-

AMP、Gauss-AMP 和 Bilateral-AMP) 对量子态估计的性能, 进行对比实验. 实验中采用的样本为对角混合态密度矩阵. 分别在采样率从 0.05 每隔 0.05 增加到 0.5, 测量噪声强度固定为 $\text{SNR} = 40 \text{ dB}$ 情况下, 不同量子态估计方法的归一化距离随采样率变化情况, 实验结果如图 5 所示. 固定采样率 $\eta = 0.1$, 不同量子态估计方法 MSE (dB) 性能随 SNR 变化情况如表 1 所示. 由图 5 可以看出, 在每一个采样率下, LDAMP 网络的性能是所有算法中最好的, 并且在采样率为 0.15 时, 就能够达到 10^{-3} 的性能. 由表 1 可以看出, 在所有不同噪声等级干扰和无噪声情况下, LDAMP 网络的估计性能均优于其他方法, 并且在无噪声干扰下的 MSE (dB) 为 -41.0905 dB , 估计误差达到 10^{-4} 量级.

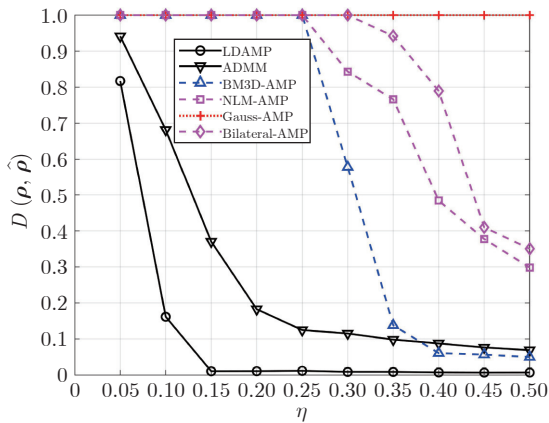


图 5 当 $\text{SNR} = 40 \text{ dB}$ 时, LDAMP 网络和其他方法的归一化距离性能对比

Fig.5 Comparison of normalized distance performance between LDAMP network and other methods with $\text{SNR} = 40 \text{ dB}$

4.3 LDAMP 网络的量子态估计实验

本节采用本文设计出的具有 20 层 DnCNN 以及 10 层算法迭代的 LDAMP 网络, 在无和有噪声的 2 种情况下, 分别对本征态、叠加态、对角混合态

和一般混合态的估计进行了性能对比实验. 影响状态估计性能的因素有采样率的多少和噪声的大小, 共进行了 4 个实验.

1) 不同采样率下无和含固定测量噪声的量子状态估计实验

在采样率从 0.05 每隔 0.05 增加到 0.5, 分别在无和含固定测量噪声 $\text{SNR} = 40 \text{ dB}$ 两种情况下, 对 4 种不同算法的量子状态进行了网络训练设计, 网络的估计结果与真实状态之间的归一化距离变化过程如图 6 所示. 从图 6 可以看出, 随着采样率的增加, 所有状态估计的归一化距离呈下降趋势, 估计效果逐渐变好. 由于测量噪声的存在, 所有状态的估计结果均略差于无噪声情况下的状态估计. 其中, 本征态的估计最为准确, 归一化距离均可以在最低采样率 0.05 时达到 10^{-4} 量级, 其在无噪声情况下的保真度为 0.999993, 有噪声情况下的保真度为 0.989028. 对角混合态的估计次之, 其在 0.15 采样率时无噪声情况下的归一化距离 (保真度) 才达到 0.0017 (0.999898). 这 2 种量子状态具有较为简单的密度矩阵结构, 因此在估计时也达到了较好的估计效果. 在无噪声情况下, 叠加态在采样率 0.3 时的归一化距离 (保真度) 达到 0.0037 (0.980570); 一般混合态在采样率 0.5 时的归一化距离 (保真度), 达到 0.0198 (0.992958). 在采样率小于 0.3 时, 对角混合态的估计效果不如一般混合态; 当采样率大于 0.3 时, 对角混合态效果优于一般混合态. 对于这种现象, 经过采用多个不同密度矩阵的验证, 可以得出以下原因: 在采样率较小时, 两种状态的估计效果均不理想, 但是由于叠加态密度矩阵中元素数值相对一般混合态中的元素数值较大一些, 所以计算出的归一化距离的值会更大, 而采样率变大后, 两种状态的估计效果均逐渐变好, 由于叠加态相对于一般混合态的密度矩阵结构更为简单, 此时叠加态的估计效果明显优于一般混合态, 其归一化距离更低.

2) 三种采样率及噪声固定下不同层数的 LDAMP 网络的性能实验

表 1 当 $\eta = 0.1$ 时, LDAMP 网络和其他方法 MSE (dB) 性能比较

Table 1 Comparison of MSE (dB) performance between LDAMP network and other methods with $\eta = 0.1$

SNR	0	10	20	30	40	50	60	70	80	无干扰
LDAMP	-25.7984	-29.0129	-32.8600	-33.6109	-36.0653	-36.5386	-38.0527	-38.9837	-40.0702	-41.0905
ADMM	-14.6620	-23.0837	-30.2725	-33.0952	-34.5782	-35.3161	-35.4256	-35.5693	-35.9288	-36.1332
BM3D-AMP	-24.8475	-25.1181	-25.3271	-25.7576	-26.2451	-26.4625	-26.7987	-26.7337	-26.9364	-27.3626
NLM-AMP	-25.6457	-25.2723	-25.1595	-25.3828	-25.4296	-25.5618	-25.7596	-25.9587	-26.0826	-26.2108
Gauss-AMP	-25.6424	-25.6433	-25.6620	-25.6991	-25.6728	-25.6729	-25.6786	-25.6985	-25.6940	-25.7053
Bilateral-AMP	-25.1949	-25.1813	-25.1152	-25.1163	-25.1209	-25.1663	-25.2087	-25.2193	-25.2296	-25.2332

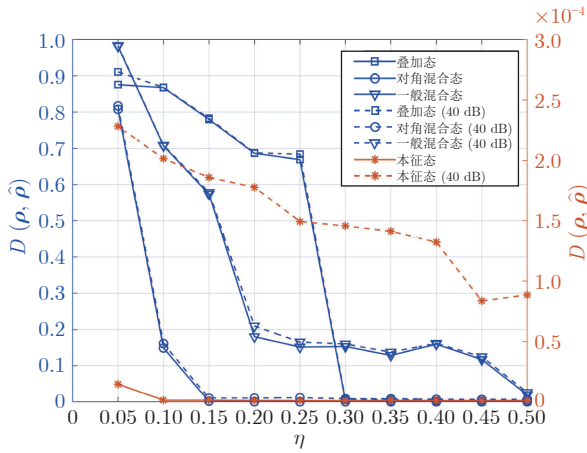


图 6 LDAMP 网络对不同量子态密度矩阵估计的归一化距离性能对比

Fig. 6 Comparison of normalized distance performance of LDAMP network for estimation of density matrices of different quantum states

图 7 为固定噪声 $\text{SNR} = 40 \text{ dB}$, 采样率分别取 0.05、0.1 和 0.2 三种情况下, 不同层数的 LDA-MP 网络估计效果与式 (10) 和式 (11) 所示的 SE

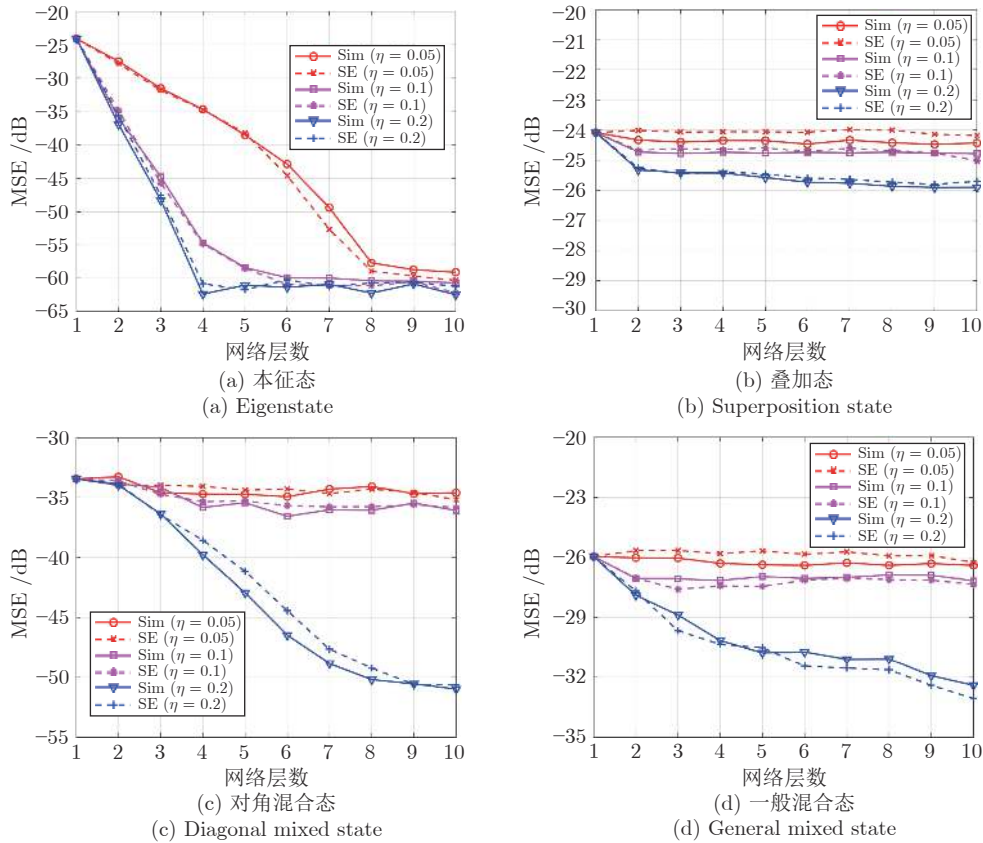


图 7 当 $\text{SNR} = 40 \text{ dB}$ 时, 不同采样率下 LDAMP 网络的仿真 (Sim) 状态演化方程的 MSE (dB) 性能对比

Fig. 7 Comparison of MSE (dB) performance between simulation (Sim) and SE analysis of LDAMP network for different sampling ratios with $\text{SNR} = 40 \text{ dB}$

方程的 MSE (dB) 性能对比的实验结果, 其中虚线为 SE 方程在三个不同采样率下的性能, 实线为本文设计的 LDAMP 网络所对应的性能 Sim. 由图 7 可以看出, 在采样率低于 0.2 时, LDAMP 网络在 10 层内收敛; 而采样率大于 0.2 时, 对于混合态和叠加态, LDAMP 网络则需要更多的层数达到收敛, 这也与密度矩阵的复杂度相关, 同时表明本文设计的 LDAMP 网络的高效性.

3) 固定采样率在无和有噪声下对角混合态的估计性能对比实验

图 8 为固定采样率 0.1, 在无和分别含 $\text{SNR} = 30 \text{ dB}$ 和 $\text{SNR} = 40 \text{ dB}$ 测量噪声的情况下, 对角混合态真实密度矩阵 ρ_3 测量值的柱状图. 图 8 中 3 个测量值之间的对比表现出不同大小测量噪声对真实测量值的影响效果. 由图 8 可以看出, 30 dB 明显比 40 dB 等级的噪声影响更大, 这种影响持续体现在对真实量子态的估计上.

图 9 为相同情况下, 对角混合态真实密度矩阵 ρ_3 与密度矩阵估计 $\hat{\rho}_3$ 、 $\hat{\rho}_3$ ($w = 30 \text{ dB}$) 和 $\hat{\rho}_3$ ($w = 40 \text{ dB}$) 的模值元素分布图. 估计出来的 3 个密度矩

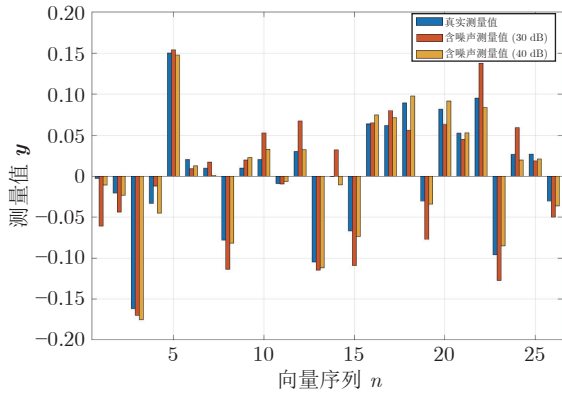


图 8 当 $\eta = 0.1$ 时, 对角混合态密度矩阵 ρ_3 的真实测量值与含噪声测量值对比

Fig.8 Comparison between the real measured values of diagonal mixed state density matrix ρ_3 and the measured values with noise in $\eta = 0.1$

阵与 ρ_3 之间的归一化距离分别为 0.1492、0.3197 和 0.1614.

4) 3 种采样率下 4 种不同量子态在不同噪声的估计性能实验

为了研究测量噪声和采样率对量子态估计的具

体影响, 还进行了采样率分别为 0.05、0.1 和 0.2 时, 4 种不同量子态估计后的 MSE (dB) 随测量噪声大小的变化的实验, 实验结果如图 10 所示. 由图 10 可以看出, 在采样率和噪声水平分别固定时, 所有量子态估计的 MSE 均随另一个未固定量的增加而减小, 其中 SNR 增加时表示噪声水平降低. LDAMP 网络的性能可以随着采样率的增加而提高. 不同量子态的表现也有差异, 对于本征态, LDAMP 网络在采样率 $\eta = 0.1$ 的情况下, 就能够获得优良的性能. 而对于其他 3 种量子态, 则需要增加采样率来达到期望的估计效果.

5 结束语

本文训练出一个具有 20 层的降噪卷积神经网络 (DnCNN), 与基于去噪的近似消息传递 (DAMP) 算法相结合, 设计出一个 10 层的基于学习去噪的近似消息传递 (LDAMP) 网络, 并将其应用于量子状态估计中. 本文提出的 LDAMP 网络在量子态估计中的性能, 优于基于压缩感知理论最先进的三维块匹配近似消息传递 BM3D-AMP 信号恢复算法和 ADMM 算法, 为采用深度学习网络进行量

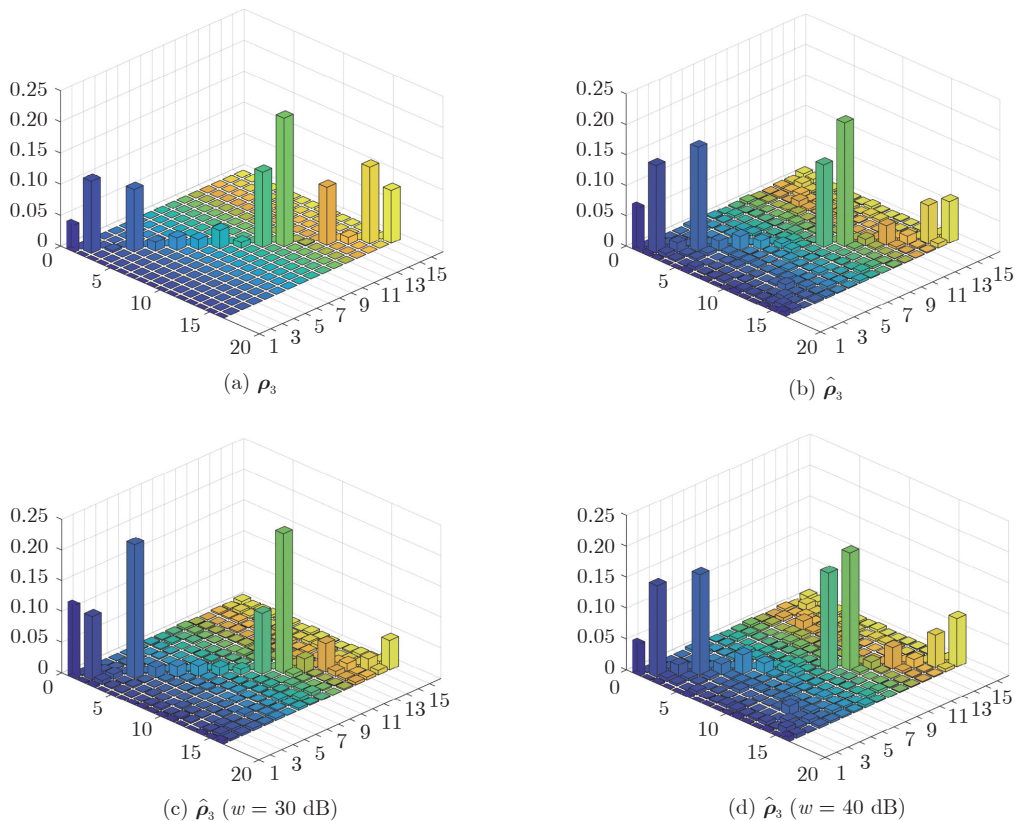


图 9 当 $\eta = 0.1$ 时, 对角混合态密度矩阵 ρ_3 及其在无和含测量噪声下估计矩阵的模值分布
Fig.9 Diagonal mixed state density matrix ρ_3 and its modulus distribution of estimation matrix without and including measurement noise with $\eta = 0.1$

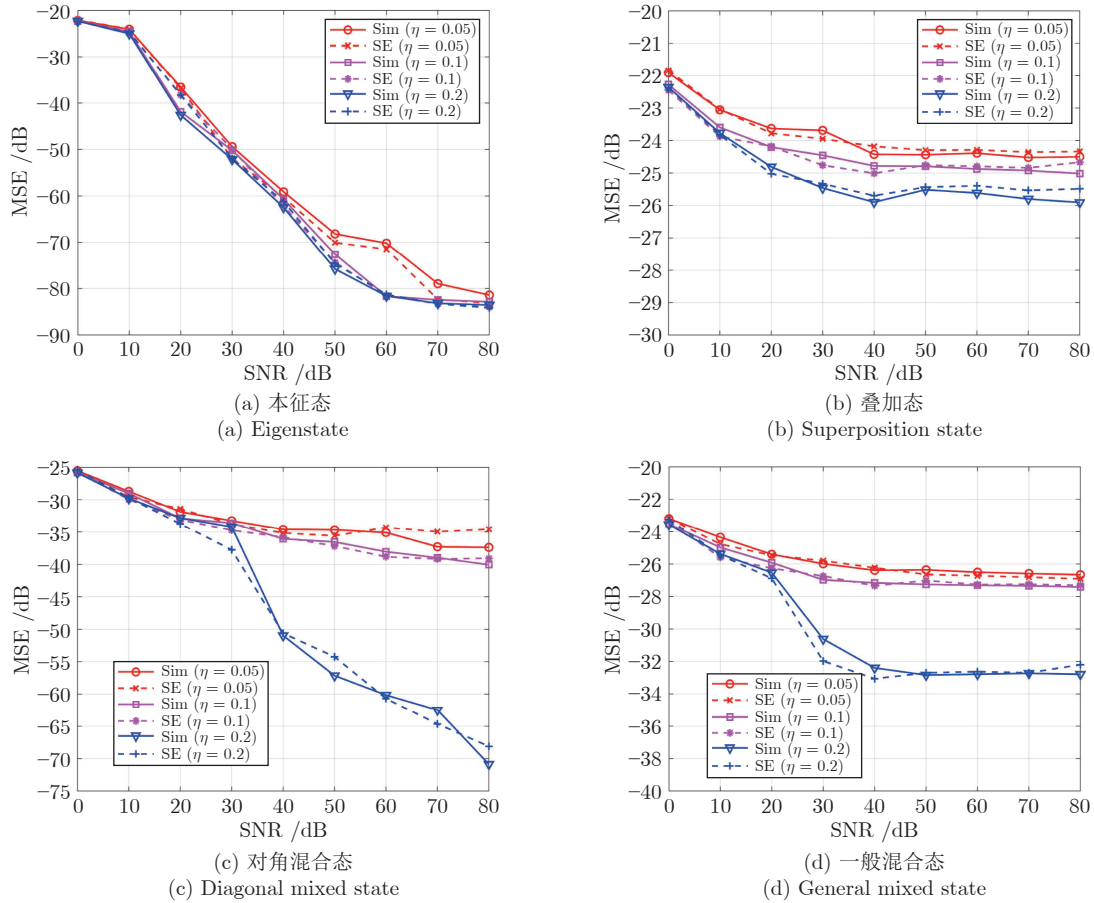


图 10 不同采样率下 LDAMP 网络对 4 种量子态密度矩阵估计时的 MSE 性能

Fig. 10 MSE performance of the LDAMP network for the estimation of four quantum state density matrices with different sampling ratios

子态估计开辟了一条新的路径. 更加深入的研究是带有虚部的深度学习网络的量子态估计, 以及实现在线量子态估计^[40]的深度学习网络的实现.

References

- Mackey G W. Mathematical foundations of quantum mechanics. *Physics Today*, 1964, **17**(4): 68
- Cahill K E, Glauber R J. Density operators and quasiprobability distributions. *Physical Review*, 1969, **177**(5): 1882–1902
- Fano U. Description of states in quantum mechanics by density matrix and operator techniques. *Reviews of Modern Physics*, 1957, **29**(1): 74–93
- D'Alessandro D. On quantum state observability and measurement. *Journal of Physics A General Physics*, 2003, **36**(37): 9721–9735
- Baraniuk R G. Compressive sensing. *IEEE Signal Processing Magazine*, 2007, **24**(4): 118–121
- Candès E J. The restricted isometry property and its implications for compressed sensing. *Comptes Rendus Mathématique*, 2008, **346**(9–10): 589–592
- Flammia S T, Gross D, Liu Y K, Eisert J. Quantum tomography via compressed sensing: Error bounds, sample complexity and efficient estimators. *New Journal of Physics*, 2012, **14**(9): 95022–95049
- Miosso C J, Von Borries R, Arguez M, Velazquez L, Quintero C, Potes C M. Compressive sensing reconstruction with prior information by iteratively reweighted least-squares. *IEEE Transactions on Signal Processing*, 2009, **57**(6): 2424–2431
- Lian Qiu-Sheng, Fu Li-Peng, Chen Shu-Zhen, Shi Bao-Shun. A compressed sensing algorithm based on multi-scale residual reconstruction network. *Acta Automatica Sinica*, 2019, **45**(11): 2082–2091
(练秋生, 富利鹏, 陈书贞, 石保顺. 基于多尺度残差网络的压缩感知重构算法. *自动化学报*, 2019, **45**(11): 2082–2091)
- Smith A, Riofrío C A, Anderson B E, Sosa-Martinez H, Deutsch I H, Jessen P S. Quantum state tomography by continuous measurement and compressed sensing. *Physical Review A*, 2012, **87**(3): 184–191
- Li K Z, Cong S. A robust compressive quantumstate tomography algorithm using ADMM. *IFAC Proceedings Volumes*, 2014, **47**(3): 6878–6883
- Boyd S, Parikh N, Chu E, Peleato B, Eckstein J. *Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers*. Boston: Now publishers Incorporated, 2011.
- Smolin J A, Gambetta J M, Smith G. Efficient method for computing the maximum-likelihood quantum state from measurements with additive Gaussian noise. *Physical Review Letters*, 2012, **108**(7): 070502
- Mézard M, Montanari A. *Information, Physics, Computation: Probabilistic approaches*. Cambridge: Cambridge University Press, 2008.
- Donoho D L, Maleki A, Montanari A. Message-passing algo-

- rithms for compressed sensing. *Proceedings of the National Academy of Sciences of the United States of America*, 2009, **106**(45): 18914–18919
- 16 Liu X M, Lu S C, Wu R B. Machine learning for data Reduction in quantum state tomography. In: Proceedings of the 37th Chinese Control Conference. Wuhan, China: 2018. 227–231
- 17 Metzler C A, Mousavi A, Baraniuk R. Learned DAMP: A principled CNN-based compressive image recovery algorithm. arXiv preprint, 2017, arXiv: 1704.06625
- 18 Zhang K, Zuo W M, Chen Y J, Meng D Y, Zhang L. Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. *IEEE Transactions on Image Processing*, 2016, **26**(7): 3142–3155
- 19 Metzler C A, Maleki A, Baraniuk R G. From denoising to compressed sensing. *IEEE Transactions on Information Theory*, 2016, **62**(9): 5117–5144
- 20 Borgerding M, Schniter P, Rangan S. AMP-inspired deep networks for sparse linear inverse problems. *IEEE Transactions on Signal Processing*, 2017, **65**: 4293–4308
- 21 Zheng Kai, Yang Jing-Bei, Cong Shuang. A construction method of optimal measurement configuration set for quantum eigenstate estimation based on pauli measurement. In: Proceedings of the 18th China Conference on System Simulation Technology and Its Applications. Lanzhou, China: 2017. 216–219 (郑凯, 杨靖北, 丛爽. 基于泡利测量的量子本征态估计最优测量配置集的构造方法. 第18届中国系统仿真技术及其应用学术年会. 兰州, 中国: 2017. 216–219)
- 22 Zheng K, Li K Z, Cong S. A reconstruction algorithm for compressive quantum tomography using various measurement sets. *Scientific Reports*, 2016, **6**: 38497
- 23 Tibshirani R. Regression shrinkage selection via the LASSO. *Journal of the Royal Statistical Society Series B (Statistical Methodology)*, 2011, **73**(3): 273–282
- 24 Chen S S, Saunders D. Atomic decomposition by basis pursuit. *SIAM Review*, 2001, **43**(1): 129–159
- 25 Borgerding M, Schniter P. Onsager-corrected deep learning for sparse linear inverse problems. In: Proceedings of IEEE Global Conference on Signal and Information Processing. Washington DC, USA: IEEE, 2016. 227–231
- 26 Ramani S, Blu T, Unser M. Monte-Carlo sure: A black-box optimization of regularization parameters for general denoising algorithms. *IEEE Transactions on Image Processing*, 2008, **17**(9): 1540–1554
- 27 He H T, Wen C K, Jin S, Li G Y. Deep learning-based channel estimation for beamspace mmwave massive MIMO systems. *IEEE Wireless Communications Letters*, 2018, **7**(5): 852–855
- 28 Dabov K, Foi A, Katkovnik V, Egiazarian K. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 2007, **16**(8): 2080–2095
- 29 Krizhevsky A, Sutskever I, Hinton G. ImageNet classification with deep convolutional neural networks. In: Proceedings of the 25th International Conference on Neural Information Processing Systems. Lake Tahoe, USA: 2012. 1097–1105
- 30 Ioffe S, Szegedy C. Batch Normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint, 2015, arXiv: 1502.03167
- 31 He K M, Zhang X Y, Ren S Q, Sun J. Deep residual learning for image recognition. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016. 770–778.
- 32 Gai Shan, Bao Zhong-Yun. High noise image denoising algorithm based on deep learning. *Acta Automatica Sinica*, 2020, **46**(12): 2672–2680 (盖杉, 鲍中运. 基于深度学习的高噪声图像去噪算法. 自动化学报, 2020, **46**(12): 2672–2680)
- 33 Zhou Deng-Wen, Zhao Li-Juan, Duan Ran, Chai Xiao-Liang. Image super-resolution based on recursive residual networks. *Acta Automatica Sinica*, 2019, **45**(6): 1157–1165 (周登文, 赵丽娟, 段然, 柴晓亮. 基于递归残差网络的图像超分辨率重建. 自动化学报, 2019, **45**(6): 1157–1165)
- 34 Schmidt U, Roth S. Shrinkage fields for effective image restoration. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA: 2014. 2774–2781
- 35 Zyczkowski K, Penson K A, Nechita I, Collins B. Generating random density matrices. *Journal of Mathematical Physics*, 2010, **52**(6): 062201
- 36 He K M, Zhang X Y, Ren S Q, Sun J. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: Proceedings of the IEEE International Conference on Computer Vision. Santiago, Chile: 2015. 1026–1034
- 37 Vedaldi A, Lenc K. MatConvNet: Convolutional neural networks for Matlab. arXiv preprint, 2014, arXiv: 1412.4564
- 38 Michael A N, Isaac L C. *Quantum Computation and Quantum Information*. Cambridge: Cambridge University Press, 2002.
- 39 Flammia S T, Liu Y K. Direct fidelity estimation from few Pauli measurements. *Physical Review Letters*, 2011, **106**(23): 230501
- 40 Tang Ya-Ru, Cong Shuang, Yang Jing-Bei. On-line state estimation of one-qubit system. *Acta Automatica Sinica*, 2020, **46**(8): 1592–1599 (唐雅茹, 丛爽, 杨靖北. 单量子比特系统状态的在线估计. 自动化学报, 2020, **46**(8): 1592–1599)



林文瑞 中国科学技术大学自动化系硕士研究生。2019年获中国科学技术大学学士学位。主要研究方向为基于深度学习网络的量子状态估计。

E-mail: lwryjj@mail.ustc.edu.cn

(**LIN Wen-Rui** Master student in the Department of Automation,

University of Science and technology of China. He received his bachelor degree from University of Science and technology of China in 2019. His main research interest is quantum state estimation based on deep learning networks.)



丛爽 中国科学技术大学自动化系教授。1995年获意大利罗马大学系统工程博士学位。主要研究方向为运动控制中的先进控制策略, 模糊逻辑控制, 神经网络设计与应用, 机器人协调控制, 量子系统控制。本文通信作者。E-mail: scong@ustc.edu.cn

(**CONG Shuang** Professor in the Department of Automation, University of Science and Technology of China. She received her Ph.D. degree in system engineering from University of Rome, Italy, in 1995. Her research interest covers advanced control strategies for motion control, fuzzy logic control, neural networks design and applications, robotic coordination control, and quantum systems control. Corresponding author of this paper.)