

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

丛 爽, 汪 涛[†], 张 坤

(中国科学技术大学 自动化系, 安徽 合肥 230027)

摘要: 针对连续弱测量中存在高斯噪声的情况, 提出一种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. 将量子态估计问题转化为含量子约束的凸优化问题, 通过引入 Von Neumann 散度(量子相对熵), 在一阶最优条件下, 推导出指数型的量子态迭代公式, 保证量子态密度矩阵的半正定性. 再通过对迭代结果进行迹为 1 的投影, 得到最终的量子态估计值. 同时在迭代公式中设计自适应的学习速率, 进一步加快算法的收敛速度. 将所提出的算法分别在 1, 2, 3 和 4 个量子位系统上, 进行了在线量子态估计的数值仿真实验, 并与现有的在线量子态估计算法进行了性能对比. 实验结果表明, 所提出的算法具有更好的快速收敛性, 以及更高的状态估计精度.

关键词: 量子态在线估计; 连续弱测量; 矩阵指数梯度算法; 自适应学习速率

引用格式: 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

DOI: 10.7641/CTA.2020.00425

Online quantum state estimation optimization algorithm with adaptive learning rate matrix exponentiated gradient

CONG Shuang, WANG Tao[†], ZHANG Kun

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

Abstract: For the presence of Gaussian noise in the continuous weak measurement process, an online quantum state estimation optimization algorithm with adaptive learning rate matrix exponentiated gradient (ALR-MEG) is proposed. The quantum state online estimation problem is transformed into convex optimization problem with quantum state constraints, and Von Neumann divergence (quantum relative entropy) is introduced into the problem. Under the first-order optimal condition, an exponential quantum state iteration formula is derived, which ensures the positive semi-definite of the quantum state density matrix, and then by projecting the iteration result with one trace, the final quantum state estimate is obtained. At the same time, the adaptive learning rate is designed in the iterative formula to accelerate the convergence speed of the algorithm. The proposed algorithm is applied on the 1, 2, 3 and 4 qubit systems, and numerical simulation experiments of online quantum state estimation are carried out. The performance is compared with the existing online quantum state estimation algorithm. The results show that the proposed algorithm has better rapid convergence, higher state estimation accuracy.

Key words: online quantum state estimation; continuous weak measurement; matrix exponentiated gradient algorithm; adaptive learning rate

Citation: CONG Shuang, WANG Tao, ZHANG Kun. Online quantum state estimation optimization algorithm with adaptive learning rate matrix exponentiated gradient. *Control Theory & Applications*, 2021, 38(8): 1188 – 1196

1 引言

量子状态的估计(quantum state estimation, QSE)是实现高精度量子系统状态反馈控制的重要前提^[1], 也是量子通信和量子计算的基础. 量子状态可以用一个密度矩阵表示. 一个 n 比特量子系统中的状态可以采用希尔伯特空间中的一个密度矩阵 $\rho \in \mathbb{C}^{d \times d}$ 来描

述, 其中, $d = 2^n$ ^[2], 同时满足半正定, 单位迹的厄米矩阵约束要求. 量子层析术(quantum state tomography, QST)^[3]是人们最早用来进行量子状态估计的技术, 此技术通过测量未知量子态的大量全同副本的输出值, 计算出测量的平均值, 再重构出量子状态. 自 2006 年压缩感知理论^[4]的提出, 使得人们可以通过较少的

收稿日期: 2020-07-08; 录用日期: 2020-12-31.

[†]通信作者. E-mail: scong@ustc.edu.cn; Tel.: +86 551-63600710.

本文责任编辑: 席在荣.

国家自然科学基金项目(61973290)资助.

Supported by the National Natural Science Foundation of China (61973290).

测量次数,高精度地重构出量子状态密度矩阵. Heinosaari等人基于压缩感知的量子层析技术,研究了先验条件对于重构所需最少测量值的影响^[5]. Zhang等人基于压缩感知与交替方向乘子法(alternating direction multiplier method, ADMM),解决了含有稀疏干扰与高斯噪声的量子态估计问题^[6-7]. 不过,他们都是采用离线的数据处理,对一个固定的量子态进行估计. 在实际应用中,尤其是量子系统的状态反馈控制中,需人们实时在线地对动态变化的量子状态进行估计.

Silberfarb提出的连续弱测量(continuous weak measurement, CWM)^[8],可以在对量子系统影响较小的情况下,通过计算系统均值来重构出量子状态^[9],为在线量子态估计提供了理论基础. Silberfarb等人首先采用极大似然法(maximum likelihood method, ML)开发出在线ML估计器^[8-9]. Ralph等人推导出完整的贝叶斯估计算法,用于单量子位系统的频率跟踪和参数估计^[10]. 他们是在海森堡绘景中,先估计出初始量子态,然后通过系统模型的演化,获得动态密度矩阵. Yang等人提出一种在薛定谔绘景下直接估计当前时刻量子态的方法:将每个时刻的状态估计问题转换为一个含有约束最小二乘问题,并采用MATLAB-CVX凸优化工具箱对其进行求解^[11]. 不过,他们所采用的方法实际上是将离线算法应用到在线求解过程中,在每一次在线估计过程中,进行一个双循环:在外部循环中在线估计量子状态,而在内部循环的每一次状态估计中,执行多次迭代,计算量非常大. Zhang等人提出一种基于在线邻近梯度下降交替方向乘子法(proximal gradient descent for online alternating direction multiplier method, OPG-ADMM)^[12]的量子态在线估计算法^[13]. 这种方法利用一阶随机梯度信息对量子态估计值进行更新,实现了量子态的在线估计. 该算法对输出结果进行半定规划(semi-definite programming, SDP)以保证量子态估计值满足半正定性与迹为1.

本文基于矩阵指数梯度的在线学习算法(matrix exponentiated gradient algorithm, MEG),设计自适应的学习速率,提出一种带有自适应学习速率的矩阵指数梯度(adaptive learning rate matrix exponentiated gradient, ALR-MEG)在线量子态估计算法. 通过建立 n 比特开放量子系统的离散演化模型,构造连续弱测量下的输出值序列. 将量子态估计问题转化为含量子状态约束的凸优化问题,并引入表示真实状态与估计状态距离的、对数型Von Neumann散度(量子相对熵),保证了量子态密度矩阵的半正定性,再通过对迭代结果进行迹为1投影,得到最终的量子态估计值. 由于所提算法对对数形式的性能指标进行求导,得到指数型迭代公式,指数与对数运算的对称性保证了输出结果的厄米特性. 因此避免了现有的在线量子态估计算法在每次估计中都中需要求解一个SDP问题,减少了计

算量. 另外,在迭代公式中采用自适应学习速率,提高了收敛速度. 将所提算法应用于1, 2, 3和4个量子位状态的估计,并与现有的在线量子态估计算法OPG-ADMM进行性能对比.

本文安排如下:第2节为连续弱测量下的开放量子系统演化模型与测量序列构造,第3节提出在线量子态矩阵指数梯度算法,第4节为仿真性能对比实验,第5节为全文总结.

2 连续弱测量下的开放量子系统

一个 n 比特量子系统可以由薛定谔绘景下的连续随机开放量子系统主方程描述:

$$\begin{aligned} \rho(t + \Delta t) - \rho(t) = & \sum [L\rho(t)L^H - \frac{1}{2}(L^H L\rho(t) + \rho(t)L^H L)]\Delta t - \\ & \frac{i}{\hbar}[H, \rho(t)]\Delta t + \sqrt{\eta} \sum [L\rho(t) + \rho(t)L^H]dW, \quad (1) \end{aligned}$$

其中: $\rho(t) \in \mathbb{C}^{d \times d}$ 为量子状态密度矩阵; $H \in \mathbb{C}^{d \times d}$ 为系统总哈密顿量; L 为Lindblad算符; L^H 表示 L 的共轭转置; Δt 为相互作用时间; $\hbar$ 通常设置为1; η 为测量效率; dW 是零差测量时产生的随机噪声,可满足 $E(dW) = 0$ 的Wiener过程.

量子状态在线估计过程如图1所示,它是由连续弱测量过程和量子状态在线估计器两部分组成. 在单量子位系统的弱测量过程中,通过引入一个二能级探测系统 P ,并将其与被测系统 S 组成联合系统,联合系统状态演化方程为

$$|\Psi(\Delta t)\rangle = U(\Delta t)(|\phi\rangle \otimes |\varphi\rangle),$$

其中: $|\phi\rangle$ 和 $|\varphi\rangle$ 分别为系统 P 和 S 的初始状态, $\otimes$ 表示 Kronecker积, $U(\Delta t) := \exp(\frac{-i\xi H \Delta t}{\hbar})$ 为联合演化算符, $\xi > 0$ 为系统相互作用强度. $|O\rangle$ 表示投影测量后,探测系统 P 的本征态. 联立联合演化算符 $U(\Delta t)$ 以及弱测量前后联合系统状态 $|\Psi(\Delta t)\rangle$ 变化关系式,推导出通过对一个二能级探测系统 P 的直接投影测量,间接作用在被测系统 S 上的弱测量算符: $m_0(\Delta t)$ 和 $m_1(\Delta t)$ ^[11]. 基于二能级量子系统的弱测量算符,可以利用张量积计算多能级量子系统的弱测量算符 M ^[14].

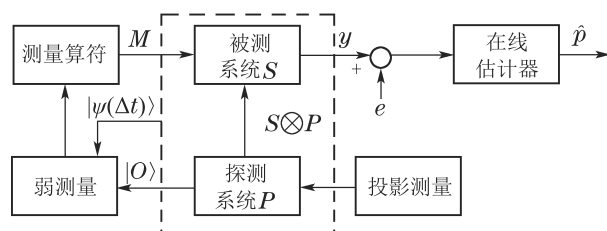


图1 基于连续弱测量的量子状态在线估计过程

Fig. 1 Online quantum state estimation process based on continuous weak measurement

由连续弱测量过程可知, 在每个采样时刻, 可以获得一个含有测量噪声的测量值. 在线估计器期望能够即时精确重构当前采样时刻的量子状态密度矩阵 $\hat{\rho}$, 以达到高效的在线估计随时间变化的量子系统的状态.

2.1 n -比特开放量子系统离散演化模型

二能级量子系统的弱测量算符为

$$\begin{aligned} m_0(\Delta t) &:= I - \left(\frac{L^H L}{2} + iH \right) \Delta t, \\ m_1(\Delta t) &:= L \sqrt{\Delta t}^{[6]}, \end{aligned}$$

其中: $H = H_0 + u_x H_x$ (H_0 为系统自由哈密顿量, H_x 为外加控制量强度, u_x 为外加控制哈密顿量); $L = \xi \sigma$, 其中 σ 为Pauli算符:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

$I_{2 \times 2}$ 为单位矩阵. n -比特量子系统的弱测量算符可以通过 $m_0(\Delta t)$ 和 $m_1(\Delta t)$ 的张量积计算获得

$$\begin{cases} M_1(\Delta t) = \underbrace{m_0(\Delta t) \otimes \cdots \otimes m_0(\Delta t)}_n \otimes m_0(\Delta t), \\ M_2(\Delta t) = \underbrace{m_0(\Delta t) \otimes \cdots \otimes m_0(\Delta t)}_n \otimes m_1(\Delta t), \\ \vdots \\ M_{2^n}(\Delta t) = \underbrace{m_1(\Delta t) \otimes \cdots \otimes m_1(\Delta t)}_n \otimes m_1(\Delta t). \end{cases} \quad (2)$$

考虑随机开放量子系统(1)中的随机噪声和弱测量效率, 二能级量子系统的演化算符可以定义为

$$\begin{aligned} a_0(\Delta t) &= m_0(\Delta t) + \sqrt{\eta} L \cdot dW, \\ a_1(\Delta t) &= m_1(\Delta t) + \sqrt{\eta} L \cdot dW^{[6]}. \end{aligned}$$

n -比特量子系统的演化算符可以通过 $a_0(\Delta t)$ 和 $a_1(\Delta t)$ 的张量积计算为

$$\begin{cases} A_1(\Delta t) = \underbrace{a_0(\Delta t) \otimes \cdots \otimes a_0(\Delta t)}_n \otimes a_0(\Delta t), \\ A_2(\Delta t) = \underbrace{a_0(\Delta t) \otimes \cdots \otimes a_0(\Delta t)}_n \otimes a_1(\Delta t), \\ \vdots \\ A_{2^n}(\Delta t) = \underbrace{a_1(\Delta t) \otimes \cdots \otimes a_1(\Delta t)}_n \otimes a_1(\Delta t). \end{cases} \quad (3)$$

根据式(1)的随机开放量子系统主方程以及式(3)的演化算符, 通过令 $t = \Delta t \cdot k$, 被测量子系统 S 的动态离散演化模型为

$$\rho_{k+1} = \sum_{i=1}^{2^n} A_i(\Delta t) \rho_k A_i(\Delta t)^H, \quad (4)$$

其中 $k = 1, 2, \dots, N$ 表示采样时间.

式(4)可以通过线性化变换为

$$\text{vec}(\rho_{k+1}) = \left(\sum_{i=1}^{2^n} A_i(\Delta t) \otimes (A_i(\Delta t)^H)^T \right) \text{vec}(\rho_k), \quad (5)$$

其中: $\text{vec}(X)$ 表示将矩阵 X 的所有列组合串联为一个列向量, X^T 表示矩阵 X 的转置.

连续弱测量过程中作用在被测系统状态上的测量算符的动态离散演化模型为

$$M_{k+1} = \sum_{i=1}^{2^n} M_i(\Delta t)^H M_k M_i(\Delta t). \quad (6)$$

2.2 连续弱测量输出序列的构造

在薛定谔绘景下, 每一采样时刻人们在系统 S 输出端获得的测量输出为 y_i . 设初始测量算符为 M_1 , 离散化的量子系统状态为 $\{\rho_i\}_{i=1}^k$, 那么, 每一采样时刻测量输出的计算公式为

$$y_i = \text{tr}(M_1^H \rho_i) = \text{vec}(M_1)^H \text{vec}(\rho_i), \quad (7)$$

其中 $\text{tr}(X)$ 表示矩阵 X 的主对角线元素之和.

实现在线算法的思路是: 基于连续弱测量, 在连续不断的每一个时间序列里, 获得一个测量值, 通过利用前面已经获得了的测量值, 结合当前的获得的输出值, 构造出一组对当前时刻量子态估计的数据集和相应的采样矩阵, 然后再结合测量值与密度矩阵之间的关系, 重构出密度矩阵. 所以, 通过在线测量实现对一个量子态估计的关键, 在于构造出合适的测量值数据序列. 将测量值序列用 b_k 表示, 在每一个采样时刻获取一个测量值, 从第2个时刻起, 所获得的测量值序列由当前时刻获得的值与以前所有时刻获得测量值构成, 即 k 时刻的测量值序列 $b_k = [y_1 \ y_2 \ \cdots \ y_k]$, 其中:

$$y_i = \text{tr}(M_{k-i+1}^H \rho_k) = \text{vec}(M_{k-i+1})^H \text{vec}(\rho_k). \quad (8)$$

根据式(8)测量数据序列的构造方法^[15], 测量值序列的长度会随着采样次数 ρ_k 的增加而增大, 不能一味地增加测量值序列的长度, 来增加计算机的运算负担. 基于压缩传感理论, 在一定的测量次数下, 就可以获得较高的估计精度. 考虑到估计精度以及在线处理的计算代价, 在量子态估计中, 需要限制并且固定一个测量值序列的长度, 从第1次测量获取数据开始, 直到达到给定的测量长度之后, 每一次新获得测量数据, 将更新替代掉已有数据序列中的一个数据, 从而始终保持估计数据序列限制在给定的长度里, 本文称此为滑动窗口. 由此可得实际量子态估计中所使用的、带有滑动窗口的测量值序列为

$$b_k = \begin{cases} (y_1 \ \cdots \ y_{k-1} \ y_k)^T, & k < l, \\ (y_{k-l+1} \ \cdots \ y_{k-1} \ y_k)^T, & k \geq l, \end{cases} \quad (9)$$

其中 l 为测量值序列的滑动窗口长度.

此时, b_k 为一个具有长度 l 的动态滑动窗口的数据

集, 滑动窗口的更新策略为先进先出. 根据式(8), 可以构造与式(9)对应的采样矩阵为

$$\mathcal{A}_k = \begin{cases} (\text{vec}(M_k) \cdots \text{vec}(M_2) \text{vec}(M_1))^H, & k < l, \\ (\text{vec}(M_l) \cdots \text{vec}(M_2) \text{vec}(M_1))^H, & k \geq l. \end{cases} \quad (10)$$

考虑到弱测量过程中不可避免的存在测量噪声, 并结合式(9)–(10), 本文将测量值序列 b_k 重写为

$$b_k = \mathcal{A}_k \text{vec}(\rho_k) + e_k, \quad (11)$$

其中 $e_k \in \mathbb{R}^k (l < k)$ 或 $e_k \in \mathbb{R}^l (k \geq l)$ 为测量噪声, 并被假设为均值为0的高斯噪声.

到此, 本文完成对连续弱测量过程中的量子系统离散状态演化模型式(5)的建立, 以及测量值序列式(11)的构造.

3 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

传统的矩阵指数梯度(MEG)学习算法, 是应用于对称正定矩阵的矩阵核学习算法. 它通过对数运算, 误差修正, 再指数运算的方式输出结果, 确保了计算结果的正定性与自共轭性. 该运算过程与量子态密度矩阵的物理约束条件相吻合. 在对量子态在线估计中, 需要同时考虑满足密度矩阵的物理约束条件, 以及对动态变化的量子密度矩阵估计的实时性, 算法的设计中需要考虑对计算状态空间距离的Bregman散度表达式的选择, 以及学习速率的设置. 本节中, 通过将Bregman散度转换为Von Neumann散度, 重新定义量子态的测量更新中的损失函数, 以及设计带有自适应的学习速率, 来推导出在线量子态密度估计中的ALR-MEG算法.

3.1 Von Neumann散度

在量子态估计中, 当 F 在定义域内是一个严格的凸函数, 且 $f(\rho) = \nabla_\rho F(\rho)$, 那么, 关于量子密度矩阵 ρ 与 $\hat{\rho}$ 的Bregman散度定义为

$$D_B(\hat{\rho}, \rho) = F(\hat{\rho}) - F(\rho) + \text{tr}((\hat{\rho} - \rho)f(\rho)^T). \quad (12)$$

当取 $F(\rho) = \text{tr}(\rho \log \rho - \rho)$, Bregman散度成为Von Neumann散度(量子相对熵):

$$D_V(\hat{\rho}, \rho) = \text{tr}(\hat{\rho} \log \hat{\rho} - \hat{\rho} \log \rho - \hat{\rho} + \rho). \quad (13)$$

在所研究问题中, 量子密度矩阵必须满足条件为 $\text{tr}(\hat{\rho}) = \text{tr}(\rho) = 1$. 将其代入式(13)中, 可以得到简化后的量子相对熵表达式为

$$D_V(\hat{\rho}, \rho) = \text{tr}(\hat{\rho} \log \hat{\rho} - \hat{\rho} \log \rho). \quad (14)$$

3.2 量子态估计值的迭代更新

根据测量值序列的表达式(11), 定义 k 时刻测量过程的损失函数为

$$L_k(\hat{\rho}) := \|\mathcal{A}_k \text{vec}(\hat{\rho}) - b_k\|_2^2, \quad (15)$$

此时, 可将量子态在线估计问题转化为一个包含Bregman散度以及以损失函数为修正项, 并同时满足量子状态约束的优化问题^[16]:

$$\begin{aligned} \min_{\hat{\rho}_k} & D_B(\hat{\rho}_k, \hat{\rho}_{k-1}) + \eta_k L_k(\hat{\rho}_k), \\ \text{s.t. } & \hat{\rho}_k \succeq 0, \text{tr}(\hat{\rho}_k) = 1, \hat{\rho}_k^H = \hat{\rho}_k, \end{aligned} \quad (16)$$

其中: $\hat{\rho}_k$ 为待估计 k 时刻的密度矩阵; $\hat{\rho}_{k-1}$ 表示 $k-1$ 时刻密度矩阵估计值; $D_B(\hat{\rho}_k, \hat{\rho}_{k-1})$ 表示Bregman散度, 其作用是保留历史估计值的影响; $L_k(\hat{\rho}_k)$ 表示 k 时刻的损失函数, 其作用是更新当前时刻的估计值, 使损失函数尽可能的小; η_k 表示 k 时刻的学习速率, 用于平衡 $D_B(\hat{\rho}_k, \hat{\rho}_{k-1})$ 与 $L_k(\hat{\rho}_k)$ 的作用强度.

首先考虑优化问题(16)中无约束情况下的求解问题, 对于这样一个具有平方损失函数的优化问题, 在每一步迭代过程中, 要解决的问题是

$$\hat{\rho}_k = \underset{\rho}{\text{argmin}} D_B(\rho, \hat{\rho}_{k-1}) + \eta_k L_k(\rho). \quad (17)$$

问题(17)是一个机器学习中经典的迭代问题, 只要保证该问题是一个凸优化问题, 便能通过令其导函数为零的方法直接求解. 损失函数(15)是一个凸二次函数, Von Neumann散度是Bregman散度的一种特殊形式, 且是严格的凸函数^[17]. 因而, 将式(14)代入式(17), 可以得到

$$\hat{\rho}_k = \underset{\rho}{\text{argmin}} \text{tr}(\rho \log \rho - \rho \log \hat{\rho}_{k-1}) + \eta_k L_k(\rho). \quad (18)$$

问题(18)为一个凸优化问题. 对式(18)求导并令其导函数为零, 可以获得方程

$$\hat{\rho}_k = \exp(\log(\hat{\rho}_{k-1}) - \eta_k \nabla L_k(\hat{\rho}_{k-1})). \quad (19)$$

方程(19)是一个指数型函数, 为非线性的, 且等式两边均有待估计状态, 无法直接得到 $\hat{\rho}_k$ 的闭式解. 常用的解决方法是将 $\nabla L_k \hat{\rho}_k$ 近似为 $\nabla L_k \hat{\rho}_{k-1}$ ^[18], 从而得到 $\hat{\rho}_k$ 的解析解为

$$\hat{\rho}_k \approx \exp(\log(\hat{\rho}_{k-1}) - \eta_k \nabla L_k(\hat{\rho}_{k-1})), \quad (20)$$

其中, $\log$ 运算将对称的半正定矩阵映射到一个对称矩阵, 加上一个对称的误差梯度修正, 通过 $\exp$ 运算, 将对称矩阵映射为一个对称的半正定矩阵, 由此保证了运算过程中, 只要 $k-1$ 时刻估计值 $\hat{\rho}_{k-1}$ 是对称的半正定阵, 那么迭代产生的下一个 k 时刻的估计值 $\hat{\rho}_k$ 一定是对称的半正定阵. 通过选择Von Neumann散度, 不需要经过任何其他处理, 便可通过式(20)获得估计值, 同时还满足问题(16)中的约束条件: $\hat{\rho}_k \succeq 0, \hat{\rho}_k^H = \hat{\rho}_k$.

对于式(20)中的学习速率, 设计自适应学习速率: 根据误差的变化情况, 自动的调整学习速率的大小, 来进一步加快算法的收敛速度.

$$\eta_k = \begin{cases} 1.04\eta_{k-1}, & \text{SSE}_k < \text{SSE}_{k-1}, \\ 0.88\eta_{k-1}, & \text{SSE}_k > 1.055\text{SSE}_{k-1}, \\ \eta_{k-1}, & \text{其他}, \end{cases} \quad (21)$$

其中: η_k 表示 k 时刻的学习速率, SSE_k 表示 k 时刻的平方误差 $L_k(\hat{\rho}_{k-1})$.

最后, 考虑问题(16)中迹为1的约束, 对式(20)的结果进行投影. 得到最终的估计状态的迭代式为

$$\hat{\rho}_k = \frac{\exp(\log(\hat{\rho}_{k-1}) - \eta_k \nabla L_k(\hat{\rho}_{k-1}))}{\text{tr}(\exp(\log(\hat{\rho}_{k-1}) - \eta_k \nabla L_k(\hat{\rho}_{k-1})))}. \quad (22)$$

式(22)就是本文推导出的带有自适应学习速率的矩阵指数梯度(ALR-MEG)在线量子态估计算法.

在算法的实际实现过程中, 为了减少迭代的计算量, 提高运行效率, 对(20)式两边取对数, 得到

$$\log(\hat{\rho}_k) \approx \log(\hat{\rho}_{k-1}) - \eta_k \nabla L_k(\hat{\rho}_{k-1}).$$

此时, 令 $G_k = \log \hat{\rho}_k$, 可以得到

$$G_k = G_{k-1} - \eta_k \nabla L_k(\hat{\rho}_{k-1}). \quad (23)$$

将式(23)代入式(22), 得到

$$\hat{\rho}_k = \frac{\exp G_k}{\text{tr}(\exp G_k)}. \quad (24)$$

算法在迭代过程最后一步对 G_k 进行指数运算, 并进行迹为1的投影, 得到 k 时刻密度矩阵的估计值. 这种方法, 只需在第一次迭代时, 对密度矩阵估计值计算 $\log \hat{\rho}_0$, 而后每次迭代, 对 G_k 指数运算即可, 使得迭代的计算量大大降低.

本文所提的带有自适应学习速率矩阵指数梯度算法(ALR-MEG)的在线量子态估计优化算法步骤总结在算法1中.

算法1 ALR-MEG.

Input: 初始化 $\hat{\rho}_0$, $G_0 = \log \hat{\rho}_0$, 步长 η_0 ; 滑动窗口长度 $l \in \mathbb{Z}_+$.

for $k = 1, 2, \dots$, do

 获取测量输出 b_k ;

 根据式(15)计算 k 时刻的误差 $L_k(\hat{\rho}_{k-1})$;

 根据式(21)计算更新 k 时刻的学习速率 η_k ;

 根据式(23)计算 G_k ;

 根据式(24)计算 k 时刻的估计值 $\hat{\rho}_k$;

end

本文所提的ALR-MEG算法, 从算法推导的过程中就保证了输出结果的半正定性与自共轭性. 对比现有OPG-ADMM算法, 省去了对每次输出结果的半定规划(SDP)处理. 在OPG-ADMM算法中, 其学习速率 η_k 计算公式为^[15]

$$\eta_k = \frac{\tau}{\alpha \lambda_{\max} + c}, \quad (25)$$

其中: τ 与 c 均为正常量, $\lambda_{\max}$ 为 $\mathcal{A}_k^H \mathcal{A}_k$ 的最大特征值.

由式(25)可以看出, OPG-ADMM算法在更新学习速率的过程中, 需要获得 $\mathcal{A}_k^H \mathcal{A}_k$ 的最大特征值 $\lambda_{\max}$, 这需要对 $\mathcal{A}_k^H \mathcal{A}_k$ 进行奇异值分解. 根据式(10)可知, 测量矩阵 $\mathcal{A}_k$ 的维度随窗口长度 l 的增加而增大, 因而使

得OPG-ADMM算法在窗口长度较长的情况下, 耗时大大增加. 而本文所提的ALR-MEG算法中, 主要的复杂度体现在对于 $\hat{\rho}_k$ 的对数运算与指数运算上, $\mathcal{A}_k$ 仅仅参与损失函数(15)中的乘法运算. 且算法的学习速率是根据实际情况自适应变化的, 不受窗口长度 l 的影响, 因此在运行效率上优于OPG-ADMM算法.

4 量子态在线估计数值仿真实验及其结果分析

本节将ALR-MEG算法与OPG-ADMM算法进行性能对比的数值仿真实验, 比较不同参数在量子态在线重构中对性能的影响. 本文将进行3个实验: 1) 滑动窗口大小对性能影响的实验; 2) 固定窗口长度下的性能对比实验; 3) 两种算法耗时性能对比实验.

实验中, 量子系统的真实状态 ρ_k 由式(5)产生, 测量矩阵 $\mathcal{A}_k$ 和测量值序列 b_k 分别由式(10)–(11)构造. 对于 n -比特量子系统模型, 密度矩阵的初始值设为

$$\rho_1^n = \underbrace{\rho_1 \otimes \dots \otimes \rho_1}_n,$$

其中: $\rho_1 = [0.5, \frac{1-i}{\sqrt{8}}; \frac{1+i}{\sqrt{8}}, 0.5]$; 密度矩阵估计值的初值根据ALR-MEG算法的初值要求设置为 $\hat{\rho}_1^n = \frac{I_d}{d}$, $d = 2^n$. 初始测量矩阵取值为

$$M_1^n = \underbrace{\sigma_z \otimes \dots \otimes \sigma_z}_n;$$

实验中分别研究了1, 2, 3和4量子位下的算法性能, 并着重研究4量子位系统两种算法性能表现. 实验中设置系统相互作用强度为 $\xi = 0.5$; 外加控制量强度为 $u_x = 2$; 系统测量效率为 $\eta = 0.5$; 随机噪声为 dW 的幅值为0.01; 高斯噪声的信噪比(signal-to-noise ratio, SNR)为60 dB. 对于估计值精度的衡量, 本文采用采用了两种评价标准. 一是在纯态研究中常用的保真度 fidelity($\hat{\rho}_k, \rho_k$), 其定义为

$$\text{fidelity}(\hat{\rho}_k, \rho_k) = \text{tr}(\sqrt{\sqrt{\hat{\rho}_k} \rho_k \sqrt{\hat{\rho}_k}}). \quad (26)$$

另一种是误差归一化的距离 $D(\hat{\rho}_k, \rho_k)$, 其定义为

$$D(\hat{\rho}_k, \rho_k) = \frac{\|\hat{\rho}_k - \rho_k\|_F^2}{\|\rho_k\|_F^2}, \quad (27)$$

其中: $\hat{\rho}_k$ 表示 k 时刻密度矩阵的估计值, ρ_k 表示 k 时刻密度矩阵的真实值. 保真度作为纯态研究中重要的精度衡量指标, 其主要反应了密度矩阵估计值与真实值在对角线元素上的保真程度. 保真度的值在(0, 1)之间, 越接近1, 那么意味着两个状态间越相似. 误差的归一化距离 $D(\hat{\rho}_k, \rho_k)$ 体现了密度矩阵估计值与真实值之间所有元素的保真程度. 其值也在(0, 1)之间, $D(\hat{\rho}_k, \rho_k)$ 的值越接近于0, 代表估计值与真实值之间相似程度越高. 本文将采用两种精度标准, 全面评价算法的各项性能. 如引言中所说, 基于极大似然估计

法(ML)和凸优化工具箱(CVX-LS)的量子态在线优化算法,本质上是双重迭代的方式. 在每个状态下均需多次迭代才能获得估计值. 由于其运行效率较低,在此不对其进行比较. 本文将重点比较ALR-MEG算法与OPG-ADMM算法在各项性能上的优劣. 实验中ALR-MEG算法与OPG-ADMM算法的参数均已调至最优,运行环境为MATLAB 2016b, 2.20 GHz Inter Core i7 8750H CPU, 16 GB内存.

4.1 滑动窗口的大小对于算法影响的实验

在本实验中,针对1, 2, 3和4量子位,验证不同长度的滑动窗口对于算法的影响. 实验中,系统的外加控制量强度 $u_x = 2$,总采样次数设置为 $N = 500$,滑动窗口的长度取值 $l = 1, 2, \dots, 100$,性能指标设置为保真度 $\text{fidelity}(\hat{\rho}_k, \rho_k) \geq 0.90$. 在每一种窗口长度下,记录不同量子位,达到性能指标所需要的最小采样次数 $k_{\min}$. 实验结果如图2所示.

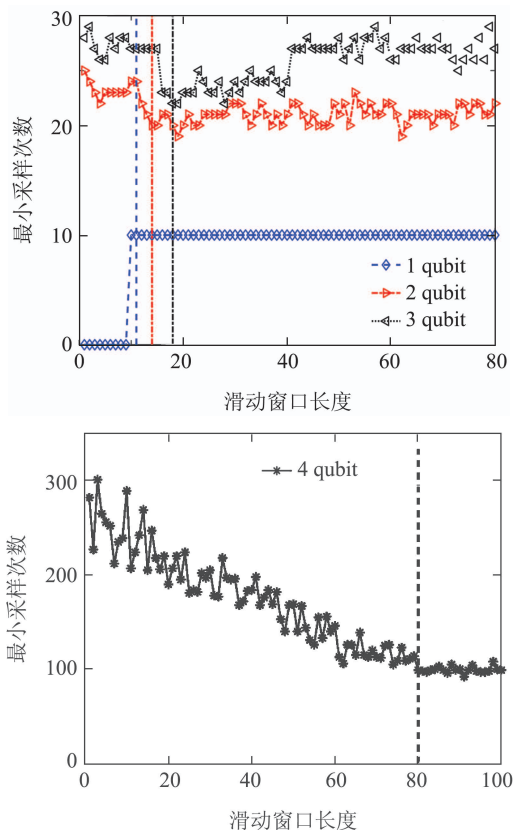


图2 滑动窗口大小对在线估计的影响

Fig. 2 Impact of the sliding window size on online estimation

从图2可以看到,随着滑动窗口尺寸的不断增大,对于不同的量子位,最小采样次数 $k_{\min}$ 均有不断减小并趋于平缓的过程. 这意味着,ALR-MEG算法对于不同的量子位有各自不同的最佳窗口长度. 在1, 2, 3, 4量子位下的最佳窗口长度分别为11, 13, 16和80,对应达到性能指标的最小采样次数分别为11, 20, 23和101.

为了更好的体现两种算法对于密度矩阵全部元素

的估计效果,对4量子位的状态估计进行了进一步研究,采用误差归一化距离 $D(\hat{\rho}_k, \rho_k) \leq 0.10$ 的性能指标,滑动窗口的长度取值 $l = 1, 2, \dots, 100$,在每个窗口长度下重复十次实验,记录每个窗口长度下达到性能指标的最少采样次数 $k_{\min}$ 十组,并计算每个窗口长度下最小采样次数 $k_{\min}$ 的标准差. 对比ALR-MEG与OPG-ADMM算法达到性能指标的最小采样次数随窗口长度的变化及标准差大小. 实验结果如图3所示.

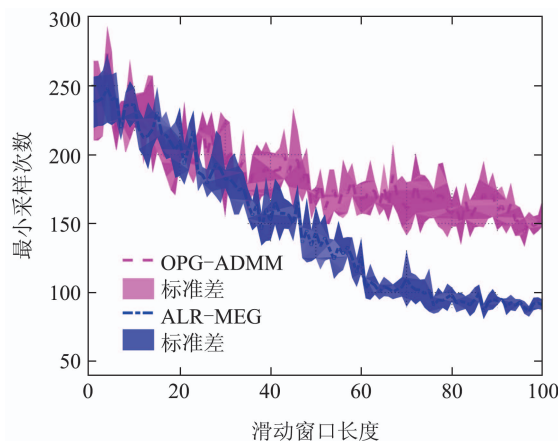


图3 可变窗口长度下算法性能对比

Fig. 3 Algorithm performance comparison under variable window length

从图3的实验结果中可以得到以下结论:

1) 4量子位下,随着窗口长度的增加,两种算法达到性能指标所需的采样次数逐渐递减,并在窗口长度达到一定值后趋于稳定. ALR-MEG和OPG-ADMM算法最佳窗口长度分别为83和81,但是两种算法达到稳定的采样次数分别为89和153. 这表明4量子位下,ALR-MEG比OPG-ADMM有更快的收敛速率.

2) 对于每一个窗口长度,标准差表示随机噪声对于算法稳定性的影响. 从图3可以看出,对两种算法而言,窗口长度的增加都有助于降低标准差,增强算法的抗干扰能力. 当 $k_{\min}$ 稳定,即窗口长度在 $[80, 100]$ 时,ALR-MEG算法的标准差范围在 $[0.41, 8.80]$,标准差的均值为5.07,而OPG-ADMM算法的标准差在 $[5.82, 24.03]$ 之间,标准差的均值为12.55. 这表明当 $k_{\min}$ 趋于稳定时,ALR-MEG算法表现出比OPG-ADMM更优异的抗干扰能力.

为了增强算法对噪声干扰的稳定性,选取略比最佳窗口长度稍大一些的窗口进行接下来的实验. 下面的所有实验中1, 2, 3和4量子位的窗口长度分别选取13, 16, 16, 88.

4.2 固定窗口长度下的性能对比实验

本节中固定窗口长度,采用保真度 $\text{fidelity}(\hat{\rho}_k, \rho_k) \geq 0.90$,和归一化距离 $D(\hat{\rho}_k, \rho_k) \leq 0.10$,两种性能指标. 对比ALR-MEG算法和OPG-ADMM算法在不同

量子位下达到目标精度所需的采样次数. 实验结果如图4所示, 横坐标为量子位数, 纵坐标为达到性能指标所花费的采样次数.

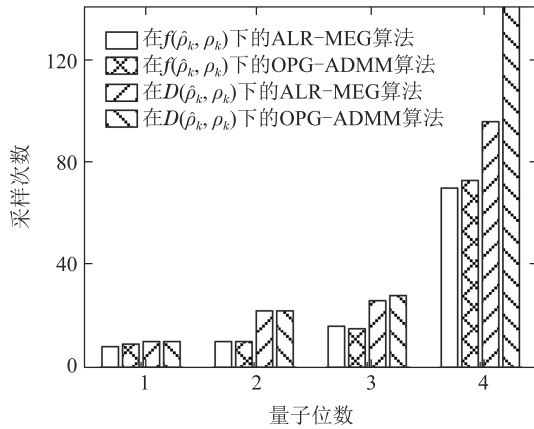


图4 固定窗口长度的性能对比

Fig. 4 Online estimation performance of two algorithms under fixed sliding window size

从图4的实验结果中可以得到如下结论:

1) 在最大采样次数 $N = 500$ 内, ALR-MEG 与 OPG-ADMM 算法均能达到目标精度. 且随着量子位数的增加, 达到目标精度所需的采样次数逐渐递增. 用保真度衡量两种算法的性能时, 两种算法所表现的性能差距不大.

2) 随着量子位数的增加, 达到 $D(\hat{\rho}_k, \rho_k) \leq 0.10$ 所需的采样次数不断增加. 两种算法在1, 2和3量子位下的性能相差不大, 但是在4量子位时, ALR-MEG 算法的采样次数小于 OPG-ADMM 算法.

针对4量子位两种算法的性能差异, 进行了进一步

的实验. 固定相同的窗口长度为88, 记录误差归一化距离 $D(\hat{\rho}_k, \rho_k)$ 随着采样次数的变化. 实验结果如图5所示.

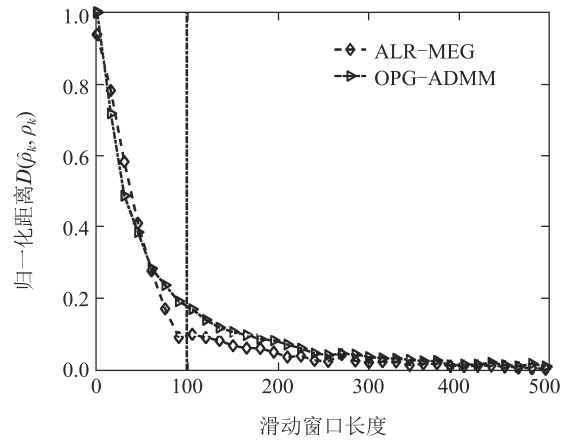


图5 4量子位归一化距离随采样次数变化

Fig. 5 Normalized distance changes with the number of sample for 4 qubit

从图5的实验结果中可以看到, 在4量子位下, ALR-MEG 与 OPG-ADMM 算法达到目标精度所需采样次数分别为106与166次, 并且ALR-MEG 比 OPG-ADMM 算法早60次采样达到目标精度. 在采样次数 $k = 100$ 时, ALR-MEG 与 OPG-ADMM 的归一化距离分别为(100, 0.08814)和(100, 0.1665).

当 $k = 100$ 时, 真实的量子状态密度矩阵幅值为 $|\rho_{100}|$, 如图6(a)所示, 采用ALR-MEG 与 OPG-ADMM 算法的估计出的状态密度矩阵幅值分别如图6(b)和6(c)所示.

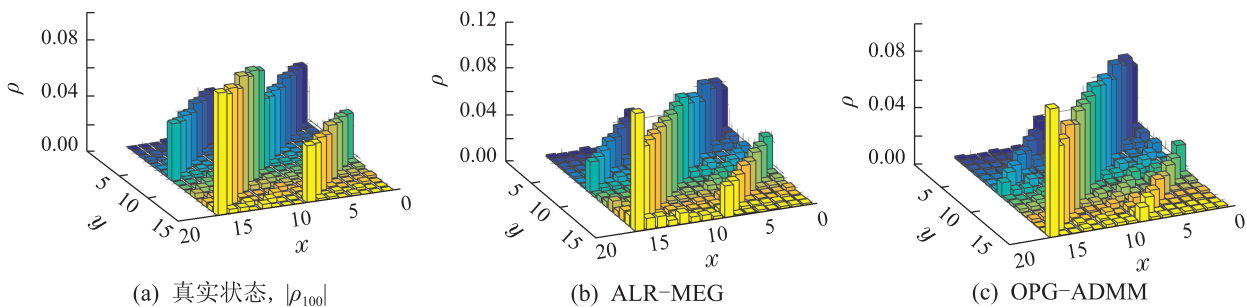


图6 $k = 100$ 时真实以及估计状态的密度矩阵幅值

Fig. 6 When $k = 100$ the density matrix amplitude of the real state and the estimated state

对比图6(b)与6(c), 可以发现, ALR-MEG 算法在 $k = 100$ 时, 不仅更好的反映了主对角线元素的特征, 而且对于非对角线元素, 相比 OPG-ADMM 算法也有更好的估计效果. 这是由于ALR-MEG 算法采用的是自适应的学习速率, 其步长会根据误差的变化而变化, 以更快的达到收敛. 而 OPG-ADMM

算法的学习速率是通过求测量矩阵的最大奇异值得, 当测量矩阵选定时, 其步长会按照固定的曲线递减. 4量子位的状态估计中, 两种算法的学习速率随采样次数的变化曲线如图7所示.

从图7的实验结果中可以得到, ALR-MEG 算法的学习速率是一个先增加再递减的过程, 而 OPG-

ADMM算法的学习速率是一个递减到定值的过程, 这与采样次数达到窗口长度后, 每次采样所使用的测量矩阵保持不变相对应. 由于两种算法的梯度计算方式不同, 学习速率的绝对值大小并没有意义, 本文只需关注学习速率的变化趋势即可. 在线估计问题中, 本文将 $k-1$ 时刻产生的估计值, 当作 k 时刻的真实值并计算误差来得到 k 时刻的估计值. 但是由于演化造成的误差并不时一个单纯的递减过程, 这使得OPG-ADMM算法的学习速率选择存在一定不足, 也让ALR-MEG算法表现出更好的跟踪效果.

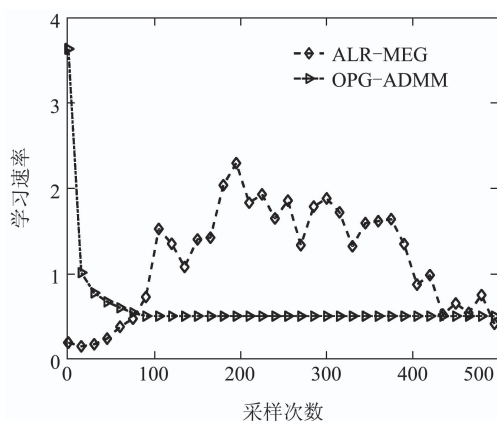


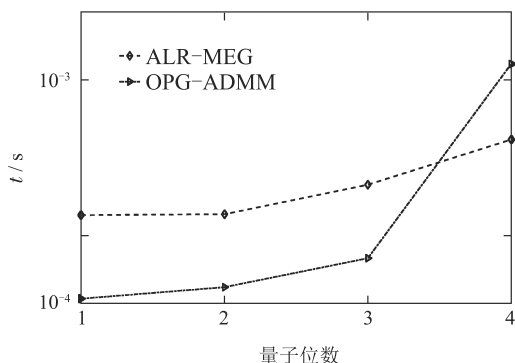
图7 两种算法的学习速率随采样次数的变化

Fig. 7 Learning rate of the two algorithms varies with the number of samples

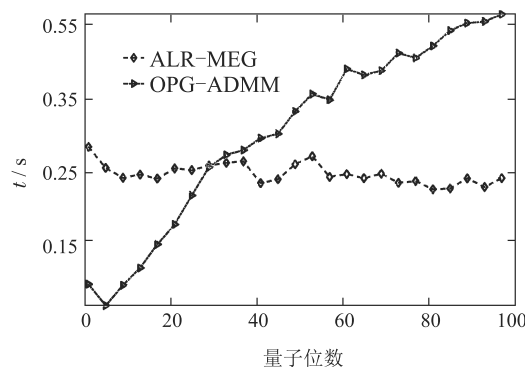
4.3 两种算法耗时性能对比实验

本节中, 对两种算法的计算时间进行了性能对比实验, 实验结果如图8所示. 其中, 图8(a)为量子位数分别在1, 2, 3和4下, 固定窗口长度分别为13, 16, 16, 88时, 达到窗口长度后每估计一次所花费的计算时间; 图8(b)为4量子位下, 采样500次, 且在不同窗口长度下所花费的计算时间.

1) 从图8(a)中可以看出: OPG-ADMM算法在1, 2, 3量子位时, 所花费的计算时间比ALR-MEG算法少, 但4量子位时, OPG-ADMM算法所需时间明显增加, 并且超过ALR-MEG算法的用时.



(a) 不同量子位下估计一次的计算时间



(b) 4量子位下计算时间随窗口长度变化

图8 两种算法的计算时间对比

Fig. 8 Comparison of the calculation time of two algorithms

2) 从图8(b)中可以看出: 在4量子位下, ALR-MEG算法估计所花费的计算时间并不随窗口长度的变化而发生太大变化, 而OPG-ADMM算法的计算时间随着窗口长度的增加而增加, 该算法学习速率的计算需要对 $\mathcal{A}_k^H \mathcal{A}_k$ 进行奇异值分解以求得其最大特征值 $\lambda_{\max}$, 随着估计量子态位数的增加, 窗口长度也增加, 使得奇异值分解计算量也随之增加. 这使得OPG-ADMM算法的耗时也不断增加, 如图8(b)中实线所示, 而ALR-MEG算法在更高维的量子态在线估计上, 因其估计计算时间不随窗口长度的增加而增长, 如图8(b)中虚线所示, 而表现出更好的性能.

5 结论

对于连续弱测量下的动态量子系统状态估计问题, 本文提出了一种带有自适应学习速率的矩阵指数梯度量子态在线估计算法ALR-MEG. 所提算法通过使用量子相对熵作为Bregman散度参与迭代, 省去了对输出结果的半定规划处理. 设计自适应学习速率加快算法的收敛速度, 并设置输出序列的滑动窗口进一步提高了估计效率. 数值实验表明, 所提方法在高量子位下, 能以更少的采样次数与耗时达到高精度的量子态在线估计, 显示出ALR-MEG算法在高量子位状态估计问题中的优越性.

参考文献:

- [1] SAYRIN C, DOTSENKO I, ZHOU X, et al. Real-time quantum feedback prepares and stabilizes photon number states. *Nature*, 2011, 477(7362): 73–77.
- [2] GYONGYOSI L. Quantum imaging of high-dimensional Hilbert spaces with Radon transform. *International Journal of Circuit Theory & Applications*, 2017, 45(7): 1029–1046.
- [3] GROSS D, LIU Y K, FLAMMIA S T, et al. Quantum state tomography via compressed sensing. *Physical Review Letters*, 2010, 105: 150401.

- [4] CANDÈS E J, ROMBERG J. Quantitative robust uncertainty principles and optimally sparse decompositions. *Foundations of Computational Mathematics*, 2006, 6(2): 529 – 531.
- [5] HEINOSAARI T, MAZZARELLA L, WOLF M M. Quantum tomography under prior information. *Communications in Mathematical Physics*, 2013, 318(2): 355 – 374.
- [6] ZHANG J, CONG S, LING Q, et al. An efficient and fast quantum state estimator with sparse disturbance. *IEEE Transactions on Cybernetics*, 2018, 49(7): 2546 – 2555.
- [7] ZHANG J, CONG S, LING Q, et al. Quantum state filter with disturbance and noise. *IEEE Transactions on Automatic Control*, 2020, 65(7): 2856 – 2866.
- [8] SILBERFARB A, JESSEN P S, DEUTSCH I H. Quantum state reconstruction via continuous measurement. *Physical Review Letters*, 2005, 95: 030402.
- [9] SMITH G A, SILBERFARB A, DEUTSCH I H, et al. Efficient quantum-state estimation by continuous weak measurement and dynamical control. *Physical Review Letters*, 2006, 97: 180403.
- [10] RALPH J F, JACOBS K, HILL C D. Frequency tracking and parameter estimation for robust quantum state estimation. *Physical Review A*, 2011, 84(5): 052119.
- [11] CONG S, TANG Y, HARRZ S, et al. On-line quantum state estimation using continuous weak measurement and compressed sensing. *Science China Information Sciences*, 2021, 64(8): 189202.
- [12] SUZUKI T. Dual averaging and proximal gradient descent for online alternating direction multiplier method. *Proceedings of the 30th International Conference on Machine Learning*. Atlanta, Georgia, USA: ICML, 2013: 392 – 400.
- [13] ZHANG K, CONG S. Online quantum state estimation algorithm based on ADMM. *Proceedings of the 20th Annual Conference on System Simulation Technology and its Application in China*. Urumqi, China: CCSSTA, 2019: 87 – 91.
- [14] HARRZ S, CONG S. State transfer via on-line state estimation and Lyapunov-based feedback control for an n -qubit system. *Entropy*, 2019, 21(8): 751.
- [15] YOUSSEY A, FERRIE C, TOMAMICHEL M. Efficient online quantum state estimation using a matrix exponentiated gradient method. *New Journal of Physics*, 2019, 21: 033006.
- [16] MATTINGELY J, BOYD S. Real-time convex optimization in signal processing. *Signal Processing Magazine*, 2010, 27(3): 50 – 61.
- [17] DE DOMENICO M, NICOSIA V, ARENAS A, et al. Structural reducibility of multilayer networks. *Nature Communications*, 2015, 6: 6864.
- [18] KHANNA S, MURTHY C R. Sparse recovery from multiple measurement vectors using exponentiated gradient updates. *IEEE Signal Processing Letters*, 2018, 25(10): 1485 – 1489.

作者简介:

丛 爽 教授, 博士生导师, 主要研究方向为量子系统控制、神经模糊系统、运动控制、机器人控制等, E-mail: scong@ustc.edu.cn;

汪 涛 硕士研究生, 研究方向为量子态估计在线优化算法, E-mail: wangt97@mail.ustc.edu.cn;

张 坤 博士研究生, 研究方向为量子态估计在线优化算法、跨模态社交媒体计算等, E-mail: kkzhang@mail.ustc.edu.cn.