This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2019.2934755, IEEE Transactions on Automatic Control

Quantum State Filter with Disturbance and Noise

Jiaojiao Zhang, Shuang Cong, *Senior Member, IEEE*, Qing Ling, *Senior Member, IEEE*, Kezhi Li, and Herschel Rabitz

***Abstract*—A quantum state filter (QSF) is proposed in this paper to estimate a low-rank quantum density matrix from informationally incomplete and contaminated measurements. There exist sparse disturbances on the quantum density matrix and Gaussian noise in the measurements. A proximal Jacobian variant of the alternating direction method of multipliers (PJ-ADMM) is proposed to design the QSF. The closed-form solutions to three resulting subproblems are given and the iterative QSF is developed. The proposed QSF is proved to be convergent and its superiority is demonstrated in the numerical illustrations compared with different state-of-the-art methods.**

***Index Terms*—Quantum state filter (QSF), proximal Jacobian, alternating direction method of multipliers (ADMM), convergence.**

I. Introduction

QUANTUM state estimation (QSE), which is also called quantum state tomography in quantum information, is a fundamental problem in quantum feedback control, quantum state preparation and quantum computing [1]–[4]. The task of QSE is to estimate the quantum state by means of experimental measurement results which are obtained by repeatedly projecting the quantum state onto different operators. An n-qubit state can be fully described by a density matrix ρ in a $d = 2^n$ dimensional Hilbert space [5]. This matrix is Hermitian, positive semidefinite and unit-trace. Since ρ has $\mathcal{O}(d^2)$ unknown elements, its recovery usually requires $\mathcal{O}(d^2)$ measurements [6]. Obtaining quantum measurements of ρ, due to the complexity of the quantum mechanics, is costly. In quantum information science, the states of interest are pure or nearly pure, which means that the corresponding density matrix ρ is low-rank [7]–[10]. Compressed sensing (CS), was introduced into low-rank matrix recovery with its efficiency and lower required number of measurements [11], [12]. According to the compressed sensing theory, $\mathcal{O}(rd \log d)$ measurements, instead of $\mathcal{O}(d^2)$, suffice to recover any low-rank matrix with overwhelming probability using proper projector bases, where r is the rank of the objective matrix. This feature has motivated researchers to estimate the density matrix ρ subject to quantum state constraints using CS [13]–[16].

This work was supported by the National Natural Science Foundation of China under Grants 61573330 and 61720106009. Herschel Rabitz acknowledges support from the US National Science Foundation.

J. J. Zhang and S. Cong with the Department of Automation, University of Science and Technology of China, Hefei, Anhui 230027, China (email: scong@ustc.edu.cn).

Q. Ling is with the School of Data and Computer Science, Sun Yat-Sen University, Guangzhou, Guangdong 510006, China.

K. Z. Li is with the Department of Electronic and Electrical Engineering, Imperial College London, London SW5 7AZ, UK.

H. Rabitz is with the Department of Chemistry, Princeton University, Princeton, NJ 08544, USA.

As in classical measurements, for quantum measurements, one also needs to consider the effect of disturbances to the quantum system [17]–[20]. The disturbances can appear both in the quantum measurements and in the quantum state. In the former case, the measurement disturbance is often assumed to be Gaussian noise and can be eliminated through least-squares techniques [21]. Liu *et al.* [22] referred to the estimation of a low-rank density matrix as a matrix Dantzig selector with quantum constraints or matrix LASSO with quantum constraints. For the latter case, sparse outliers were introduced on the elements of ρ as the state disturbances, which made accurate quantum state reconstruction difficult. Li *et al.* [23] applied the alternating direction method of multipliers (ADMM) and projected the solution to the feasible set. Recently, Zheng *et al.* [24] and Zhang *et al.* [25] considered the special case of $r = 1$ (i.e., ρ is unit rank), which significantly simplified the quantum state constraints. As a result they solved the subproblems in ADMM more efficiently. Specifically, Zheng *et al.* combined the fixed-point technique with ADMM while Zhang *et al.* applied the iterative shrinkage-thresholding operator in solving the subproblems of ADMM.

In this paper, we propose a quantum state filter with both sparse disturbance in the state and Gaussian noise in the measurement. For sparse disturbance in the state, the existing algorithms are ADMM [23], FP-ADMM [24] and IST-ADMM [25]. For Gaussian noise in the measurement, there are least-squares [21] and matrix Lasso [22]. However, these algorithms are unable to jointly handle both sparse state disturbance and Gaussian measurement noise. In addition, the existing algorithms either fail to satisfy the quantum state constraint, or have high computational complexity. This motivates us to develop a computationally-effective and convergent quantum state filter, which explicitly considers the quantum state constraint and is efficient to exactly recover the density matrix from the measurements contaminated by sparse state disturbance and Gaussian measurement noise. To be more specific, we focus on the problem of recovering a $d \times d$ low-rank density matrix ρ contaminated by a sparse $d \times d$ matrix S from m linear measurements contaminated by Gaussian noise e, where $m = \mathcal{O}(rd \log d) \ll d^2$. The task can be formulated as minimizing the nuclear norm of ρ, the ℓ_1 norm of S and the ℓ_2 norm of e, subject to linear measurement constraints and quantum constraints. This is a novel model in QSE where both state disturbances and measurement noise are considered. The two types of disturbances and noise inevitably exist in practical cases of quantum state estimation, and if either of them is ignored, then a large bias can lead the estimation being far from the optimal result.

We are inspired in this work by filters in classical con-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2019.2934755, IEEE Transactions on Automatic Control

trol theory which can eliminate both state and measurement disturbances. Thus, we propose a QSF which can estimate the quantum state while filtering out disturbances in the state and measurements. However, recovering the quantum state from contaminated measurements is a challenging task. Three variables, namely, the low-rank ρ, the sparse S, and the Gaussian noise e, are coupled in the measurements b. In addition to being contaminated, b is also informationally incomplete because only $m = \mathcal{O}(rd\log d) \ll d^2$ measurements are obtained instead of d^2. Furthermore, the density matrix ρ must satisfy its quantum constraints. We introduce the proximal Jacobian ADMM (PJ-ADMM) [26] to solve the QSF problem. The original problem is decomposed into three subproblems by adding the proximal terms to each of the subproblems to correct for the errors. Then the subproblems are solved in parallel at each iteration. An iterative QSF is developed, and the global convergence is proved. The fast convergence, efficient sampling rate and the computational superiority of the QSF are demonstrated by the numerical experiments compared with IST-ADMM and ADMM methods.

The remainder of the paper is organized as follows. Section II states the QSF problem. Section III proposes a QSF based on Proximal Jacobian ADMM and Section IV proves its global convergence. Numerical experiments are given in Section V and Section VI summarizes the paper.

Notation. For a vector, $\|\cdot\|_2$ denotes the ℓ_2 norm. For a matrix, $\|\cdot\|_*$, $\|\cdot\|_1$, $\|\cdot\|_{\mathcal{F}}$ and $tr(\cdot)$ denote its nuclear, ℓ_1 norm, Frobenius norm and trace, respectively. Superscript "H" denotes the conjugate transpose of a complex matrix. For a linear operator $\mathcal{A}:\mathbb{C}^{d\times d}\to\mathbb{C}^m$, $\mathcal{A}^H:\mathbb{C}^m\to\mathbb{C}^{d\times d}$ denotes its Hermitian transpose. $\langle\cdot,\cdot\rangle$ denotes the inner product such that $\langle Z,Z'\rangle = tr(Z^HZ')$; $G\succeq 0$ means that G is positive semidefinite. We define $\langle Z,Z'\rangle_G = \langle Z,GZ'\rangle$ and the induced matrix norm $\|Z\|_G = \sqrt{\langle Z,Z\rangle_G}$. $vec(\cdot)$ vectorizes a matrix to a column vector. ∂f denotes the subdifferential of a function f. The Kronecker product is denoted by $\otimes$.

II. Problem Statement

Consider a quantum system of n qubits and let $d = 2^n$. The quantum state can be described by a density matrix to be estimated $\rho\in\mathbb{C}^{d\times d}$, which is complex, positive semidefinite and unit-trace Hermitian; i.e., $\rho^H=\rho$, $\rho\succeq 0$, and $tr(\rho)=1$. In addition, for a pure or nearly pure quantum system, ρ is low-rank [7]–[9], meaning that its rank $r\ll d$. Mathematically, measuring the density matrix ρ is through Pauli bases. A Pauli basis of an n-qubit quantum system is generated by randomly selecting n Pauli matrices and calculating their normalized Kronecker product. Specifically, a Pauli basis is in the form of $W_i = (\sigma_1\otimes\cdots\otimes\sigma_n)/\sqrt{d}$, where $\{\sigma_j, j=1,\cdots,n\}$ are randomly selected from the set of Pauli matrices $\{\sigma_x,\sigma_y,\sigma_z,I\}$, with

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

and I denotes the 2×2 identity matrix. According to the definition, there are d^2 Pauli bases and everyone is Hermitian. Given a Pauli basis W_i, the measurement of the density matrix ρ is

$$b_i = tr(W_i^H\rho) = vec(W_i)^H vec(\rho). \tag{1}$$

Note that, in practice, such a measurement is obtained by preparing a large number of identical copies of the quantum state and calculating the mean value of the observations [5].

The fact that the quantum measurement is costly and the density matrix to be estimated ρ is low-rank motivates us to apply CS to quantum state tomography. Randomly choosing $m = \mathcal{O}(rd\log d)\ll d^2$ Hermitian bases out of the total d^2 and defining the sampling rate as

$$\eta = m/d^2, \tag{2}$$

we obtain m measurements according to (1). For simplicity, we write the measurement function in matrix form $b=\mathcal{A}(\rho)$, where $b\in\mathbb{R}^m$ is the measurement vector and $\mathcal{A}:\mathbb{C}^{d\times d}\to\mathbb{C}^m$ is a linear operator. To be more specific, define a matrix $A\in\mathbb{C}^{m\times d^2}$ whose i-th row corresponds to a Pauli basis W_i and is given by $vec(W_i)^H$. Thus, $\mathcal{A}(\rho) := Avec(\rho)$.

Practical experiments and measurements are inevitably subject to disturbances. In this paper, we simultaneously consider two kinds of disturbances, *sparse disturbance* of the quantum state ρ and *Gaussian noise* in the measurements b, both of which exist in practical physical measurement processes. Define a sparse disturbance matrix $S\in\mathbb{R}^{d\times d}$ whose nonzero elements are much less than d^2 and a Gaussian noise vector $e\in\mathbb{R}^m$. The contaminated measurement can be written as

$$b = \mathcal{A}(\rho+S)+e. \tag{3}$$

Given the linear operator $\mathcal{A}$ and the measurement vector b contaminated by sparse disturbance S and Gaussian noise e, the goal of the QSF is to estimate the density matrix ρ. With prior knowledge that ρ is low-rank and S is sparse, we propose to solve the following convex optimization problem

$$\begin{aligned}\min\ & \|\rho\|_* + \gamma\|S\|_1 + \frac{\theta}{2}\|e\|_2^2,\\ s.t.\ & \mathcal{A}(\rho+S)+e=b,\\ & \rho\in C,\end{aligned} \tag{4}$$

where γ and θ are nonnegative regularization parameters, $C := \{\rho\in\mathbb{C}^{d\times d}|\rho^H=\rho,\rho\succeq 0, tr(\rho)=1\}$ is a convex set, and $\rho\in C$ is the quantum state constraint. Minimizing $\|\rho\|_*$ and $\|S\|_1$ forces ρ to be low-rank and S to be sparse, respectively, while minimizing $\|e\|_2^2$ is common in filtering Gaussian noise e. If without S being present, (4) reduces to the quantum state estimation model that only considers Gaussian measurement noise [21], [22]; without observation error e, (4) becomes the model that only considers sparse state disturbances [23]–[25]. It should be noted that the item $\|\rho\|_*$ is equal to $tr(\rho)$ under the quantum constraints $\rho\in C$, because the singular values of a Hermitian positive semipositive matrix are equal to their eigenvalues. This fact will be used in the derivation and analysis of the proposed algorithm.

Solving (4) is a challenging task since the three variables (i.e., the low-rank ρ, the sparse S and the Gaussian e), are coupled in the objective function $\|\rho\|_*+\gamma\|S\|_1+(\theta/2)\|e\|_2^2$ and the measurement constraint $\mathcal{A}(\rho+S)+e=b$. In addition,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2019.2934755, IEEE Transactions on Automatic Control

the density matrix ρ must satisfy the quantum constraints, which makes the problem more complicated. In the Section III, we shall develop a novel and provably convergent quantum state filter that iteratively estimates the density matrix ρ, while filtering out the sparse disturbance S and the Gaussian noise e.

III. Quantum State Filter With Disturbance And Noise

We first overview the proposed quantum state filter in III-A and then derive the detailed algorithm in III-B.

A. Overview of the Filter

The flowchart of the proposed iterative QSF is depicted in Fig. 1. Here we define k as the iteration index and q^{-1} as a one-step delay.

The top of Fig. 1 refers to the measurement process (see also Section II). The true density matrix ρ is disturbed by a sparse signal S and goes through the measurement block to obtain $\mathcal{A}(\rho+S)$. Then, Gaussian noise e is superimposed into the result, yielding the measurement vector $b=\mathcal{A}(\rho+S)+e$. The QSF estimates the quantum state ρ and filters out the disturbances S and e from contaminated measurement b, respectively.

At the $(k+1)$-th iteration of the proposed quantum state filter, the outputs from the previous iteration, ρ^k, S^k and e^k, are imported into the ρ-, S- and e-linear operators so as to obtain intermediate values $\tilde{\rho}^k$, $\tilde{S}^k$ and $\tilde{e}^k$, respectively (the forms of these linear operators is given in Section III-B). Then, $\tilde{\rho}^k$ is imported into the projection operator, where $\tilde{\rho}^k$ is projected onto the quantum constraint set C, such that the output ρ^{k+1} is positive semidefinite and unit-trace Hermitian. Meanwhile, $\tilde{S}^k$ is imported into the soft thresholding operator to obtain a sparse output S^{k+1}. Here $\tilde{e}^k$ is equal to e^{k+1}, denoting the new estimate of the Gaussian noise.

The proposed algorithm works as a type of a filter. When the number of iterations evolves, the QSF gradually removes the Gaussian noise e from the measurement vector b, separates the sparse signal S from the density matrix ρ and obtains an estimate of the quantum state.

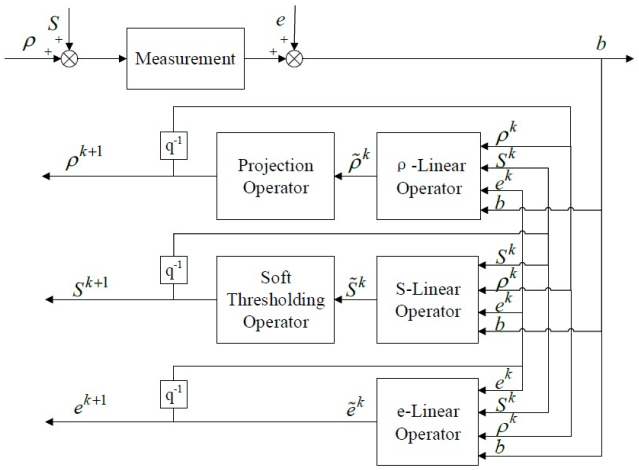


Fig. 1: Flowchart of the proposed quantum state filter (QSF).

B. Development of the Filter

Here we introduce the proximal Jacobian variant of the alternating direction method of multipliers (PJ-ADMM) [26] to develop the QSF.

The partial augmented Lagrangian of (4) can be defined as

$$\begin{aligned} L_\alpha :=& \|\rho\|_* + \gamma\|S\|_1 + \frac{\theta}{2}\|e\|_2^2 - \langle y, \mathcal{A}(\rho+S)+e-b\rangle \\ &+ \frac{\alpha}{2}\|\mathcal{A}(\rho+S)+e-b\|_2^2,\ \rho\in C, \end{aligned} \tag{5}$$

where, $y\in\mathbb{R}^m$ is a Lagrange multiplier and a real vector, which is associated with the equality constraint $\mathcal{A}(\rho+S)+e=b$; α is a positive penalty parameter. The quantum constraint $\rho\in C$ remains explicit. During the optimization process, we will guarantee that every iteration of ρ stays in the quantum constraint set C such that $\rho^H=\rho$. Thus, we know that $\mathcal{A}(\rho)$ is real, and so is $\mathcal{A}(\rho+S)-b$.

At the $(k+1)$-th iteration, we use the PJ-ADMM updates the primal variables ρ, S and e in parallel, and then updates the dual variable y. To be specific, for every primal variable, we fix the rest and minimize the partial augmented Lagrangian plus a proximal term. After updating all the primal variables, the dual variable is calculated through dual gradient ascent. The resultant PJ-ADMM for the QSF is

$$\begin{cases} \rho^{k+1} = \arg\min\limits_{\rho\in C}\Big\{\|\rho\|_* + \langle y^k, \mathcal{A}(\rho+S^k)+e^k-b\rangle \\ \qquad + \frac{\alpha}{2}\|\mathcal{A}(\rho+S^k)+e^k-b\|_2^2 + \frac{1}{2}\|\rho-\rho^k\|_{P_1}^2\Big\}, & \text{(6a)} \\ S^{k+1} = \arg\min\limits_{S}\Big\{\gamma\|S\|_1 + \langle y^k, \mathcal{A}(\rho^k+S)+e^k-b\rangle \\ \qquad + \frac{\alpha}{2}\|\mathcal{A}(\rho^k+S)+e^k-b\|_2^2 + \frac{1}{2}\|S-S^k\|_{P_2}^2\Big\}, & \text{(6b)} \\ e^{k+1} = \arg\min\limits_{e}\Big\{\frac{\theta}{2}\|e\|_2^2 + \langle y^k, \mathcal{A}(\rho^k+S^k)+e-b\rangle \\ \qquad + \frac{\alpha}{2}\|\mathcal{A}(\rho^k+S^k)+e-b\|_2^2 + \frac{1}{2}\|e-e^k\|_{P_3}^2\Big\}, & \text{(6c)} \\ y^{k+1} = y^k - \kappa\alpha(\mathcal{A}(\rho^{k+1}+S^{k+1})+e^{k+1}-b). & \text{(6d)} \end{cases}$$

where, $(1/2)\|\rho-\rho^k\|_{P_1}^2$, $(1/2)\|S-S^k\|_{P_2}^2$ and $(1/2)\|e-e^k\|_{P_3}^2$ are proximal terms attached to the primal subproblems and κ is a damping parameter for the dual update.

Through adjusting κ, we are able to choose a proper dual stepsize so as to achieve fast convergence, as we will show in the later Sections of the paper. For the proximal terms, the squared pseudo norms are defined as $\|z_i\|_{P_i}^2 := \langle z_i, P_i z_i\rangle$, $i=1,2,3$. The baseline is that $P_1+\alpha\mathcal{A}^H\mathcal{A}\succ 0$, $P_2+\alpha\mathcal{A}^H\mathcal{A}\succ 0$ and $P_3+\alpha I\succ 0$ so that the subproblems (6a), (6b) and (6c) are all convex.

To implement a provably convergent and computationally efficient QSF, we shall specifically choose the matrices P_i so that: (i) the algorithm converges to the global optimal solution of (4); (ii) the subproblems in (6) have computationally efficient explicit solutions. Below are the choices of the matrices P_i.

- For ρ-subproblem (6a), the main computational difficulty comes from the quadratic term $(\alpha/2)\rho^H\mathcal{A}^H\mathcal{A}\rho$. We let $P_1=\tau_1 I-\alpha\mathcal{A}^H\mathcal{A}$ with $\tau_1>0$, which cancels the quadratic term $(\alpha/2)\rho^H\mathcal{A}^H\mathcal{A}\rho$ and adds $(\tau_1/2)\rho^H\rho$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2019.2934755, IEEE Transactions on Automatic Control

- For S-subproblem (6b), similar to what we have done for ρ-subproblem (6a), we let $P_2 = \tau_2 I - \alpha \mathcal{A}^H \mathcal{A}$ with $\tau_2 > 0$.
- For e-subproblem (6c), since it has already had an explicit solution, we let $P_3 = \tau_3 I$ with $\tau_3 > 0$.

By substituting P_1, P_2 and P_3 defined above, (6) with matrix norm $\|\cdot\|_{P_i}$ can be simplified as the following algorithm with Frobenius norm $\|\cdot\|_{\mathcal{F}}$ and ℓ_2 norm $\|\cdot\|_{\ell_2}$

$$\begin{cases} \rho^{k+1} = \arg\min\limits_{\rho \in C} \left\{ \|\rho\|_* + \frac{\tau_1}{2} \left\|\rho - \tilde{\rho}^k\right\|_{\mathcal{F}}^2 \right\}, & \text{(7a)} \\ S^{k+1} = \arg\min\limits_{S} \left\{ \gamma \|S\|_1 + \frac{\tau_2}{2} \left\|S - \tilde{S}^k\right\|_{\mathcal{F}}^2 \right\}, & \text{(7b)} \\ e^{k+1} = \arg\min\limits_{e} \left\{ \frac{\theta + \alpha + \tau_3}{2} \|e - \tilde{e}^k\|_2^2 \right\}, & \text{(7c)} \\ y^{k+1} = y^k - \kappa\alpha(\mathcal{A}(\rho^{k+1} + S^{k+1}) + e^{k+1} - b). & \text{(7d)} \end{cases}$$

Here we define intermediate variables

$$\tilde{\rho}^k := \rho^k - \frac{\alpha}{\tau_1} \mathcal{A}^H (\mathcal{A}(\rho^k + S^k) + e^k - b - \frac{y^k}{\alpha}), \quad (8)$$

$$\tilde{S}^k := S^k - \frac{\alpha}{\tau_2} \mathcal{A}^H (\mathcal{A}(\rho^k + S^k) + e^k - b - \frac{y^k}{\alpha}), \quad (9)$$

and

$$\tilde{e}^k := \frac{\tau_3 e^k - \alpha(\mathcal{A}(\rho^k + S^k) - b - \frac{y^k}{\alpha})}{\theta + \alpha + \tau_3}. \quad (10)$$

Below we proceed to show that the subproblems in (7) have computationally efficient explicit solutions.

ρ-subproblem: For (7a), considering the quantum state constraints $\rho = \rho^H$, $\rho \succeq 0$ and $tr(\rho) = 1$, we have that $\|\rho\|_* = tr(\rho) = 1$. Thus, (7a) reduces to solving the following semidefinite program (SDP) in the form of

$$\begin{aligned} \min \ & \left\|\rho - \tilde{\rho}^k\right\|_{\mathcal{F}}^2, \\ s.t. \ & \rho = \rho^H, \rho \succeq 0, tr(\rho) = 1, \end{aligned} \quad (11)$$

or equivalently, solving

$$\begin{aligned} \min \ & \left\| \rho - \frac{\tilde{\rho}^k + (\tilde{\rho}^k)^H}{2} \right\|_{\mathcal{F}}^2, \\ s.t. \ & \rho \succeq 0, tr(\rho) = 1. \end{aligned} \quad (12)$$

Note that the SDP (12) has a closed-form solution [27]. Run the eigenvalue decomposition to rewrite $(\tilde{\rho}^k + (\tilde{\rho}^k)^H)/2$ as $V diag\{a_i\} V^H$, where $V \in \mathbb{C}^{d \times d}$ is a unitary matrix and $\{a_i, i = 1, \cdots, d\}$ are eigenvalues of $(\tilde{\rho}^k + (\tilde{\rho}^k)^H)/2$. Here we sort the eigenvalues as $a_1 \geq a_2 \geq \cdots \geq a_d$. The optimal solution to (12) is

$$\rho^{k+1} = V diag\{x_i\} V^H, \quad (13)$$

where $\{x_i, i = 1, \cdots, d\}$ are eigenvalues of ρ^{k+1} and solved from

$$\begin{aligned} \min \ & \frac{1}{2} \sum_{i=1}^{d} (x_i - a_i)^2, \\ s.t. \ & \sum_{i=1}^{d} x_i = 1, x_i \geq 0, \forall i. \end{aligned} \quad (14)$$

Write the partial Lagrangian for (14) as

$$\begin{aligned} L(\{x_i\}, \beta) := & \frac{1}{2} \sum_{i=1}^{d} (x_i - a_i)^2 + \beta(\sum_{i=1}^{d} x_i - 1), \\ & x_i \geq 0, \forall i, \end{aligned} \quad (15)$$

where $\beta \in \mathbb{R}$ is the Lagrange multiplier.

According to convex optimization theory, if β is the optimal Lagrange multiplier, minimizing $L(\{x_i\}, \beta)$ over $\{x_i\}$ from

$$\begin{aligned} \min \ & \frac{1}{2} \sum_{i=1}^{d} (x_i - a_i + \beta)^2, \\ s.t. \ & x_i \geq 0, \forall i, \end{aligned} \quad (16)$$

yields the optimal primal solution

$$x_i = \max\{a_i - \beta, 0\}, \forall i. \quad (17)$$

For the optimal β, we must have $\sum_{i=1}^{d} x_i = 1$. Using the relation (17), it is equivalent to $\sum_{i=1}^{d} \max\{a_i - \beta, 0\} = 1$. This equality suggests trying $\beta = a_i$ at every $i = 1, \cdots, d$ and then deducing which interval the optimal β must be within. Suppose that β is within $[a_{t+1}, a_t]$, we know that $x_i = a_i - \beta, \forall i \leq t$ and $x_i = 0, \forall i \geq t+1$. Thus, the optimal β is solved from $\sum_{i=1}^{t} (a_i - \beta) = 1$; namely, $\beta = (\sum_{i=1}^{t} a_i - 1)/t$. Using this optimal β, we calculate $\{x_i\}$ from (17) and consequently ρ^{k+1} from (13).

S-subproblem: For (7b), it has a well-known soft thresholding solution

$$S^{k+1} = shrink_{\gamma/\tau_2}(S^k - \tilde{S}^k), \quad (18)$$

where $shrink_{\gamma/\tau_2}$ is an element-wise soft thresholding (or shrinkage) operator. Given any scalar s, $shrink_{\gamma/\tau_2}(s) := \max\{|s - \gamma/\tau_2|, 0\} sign(s - \gamma/\tau_2)$.

e-subproblem: For (7c), applying its optimal condition, we have

$$e^{k+1} = \tilde{e}^k. \quad (19)$$

The algorithm used in the proposed QSF is outlined in Algorithm 1, which corresponds to the flowchart in Fig. 1. Note that the ρ, S and e-subproblems are solved in parallel. The main computational complexity comes from the calculations of $\tilde{\rho}^k$ and $\tilde{S}^k$, both of which scale as $\mathcal{O}(md^2)$. The eigenvalue decomposition in the ρ-subproblem has a complexity order of $\mathcal{O}(d^3)$, which is lower than $\mathcal{O}(md^2)$ since $m = \eta d^2$.

Remark. *Note that in the quantum state filter* (6)*, if we choose $P_1 = 0$, $P_2 = 0$ and $P_3 = 0$, then the algorithm reduces to the traditional ADMM with Jacobian primal update [28], [29]. However, it has been shown in [30] that when the number of primal blocks is larger than 2, the traditional ADMM does not have global convergence in general. In the Section IV, we show that when P_1, P_2 and P_3 are properly chosen, the proposed QSF* (6) *converges to the global optimal solution of* (4)*.*

IV. Convergence Proof and Analysis

In this Section, we shall give the convergence of the QSF (6) in Theorem 1 and prove the convergence of the proposed QSF to the global optimal solution of (4). To simplify the notation,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2019.2934755, IEEE Transactions on Automatic Control

Algorithm 1 PJ-ADMM for Quantum State Filter

Require: Initialize variables to $\rho^0 = 0, S^0 = 0, e^0 = 0, y^0 = 0$ and set algorithm parameters $\tau_1, \tau_2, \tau_3, \alpha, \kappa > 0$.

For $k = 1, 2, \cdots$ **Do**
  Calculate $\tilde{\rho}^k = \rho^k - \frac{\alpha}{\tau_1}\mathcal{A}^H(\mathcal{A}(\rho^k + S^k) + e^k - b - \frac{y^k}{\alpha})$.
  Run eigenvalue decomposition on $(\tilde{\rho}^k + (\tilde{\rho}^k)^H)/2$ to obtain $Vdiag\{a_i\}V^H$, where $a_1 \geq a_2 \geq \cdots \geq a_d$.
  Set β as $a_1, a_2, \cdots$ to find t such that $\sum_{i=1}^{t} \max\{a_i - \beta, 0\} < 1$ and $\sum_{i=1}^{t+1} \max\{a_i - \beta, 0\} > 1$.
  Find optimal $\beta = \frac{1}{t}\sum_{i=1}^{t} a_i$ and $x_i = \max\{a_i - \beta, 0\}, \forall i$.
  Update $\rho^{k+1} = Vdiag\{x_i\}V^H$.
  Calculate $\tilde{S}^k = S^k - \frac{\alpha}{\tau_2}\mathcal{A}^H(\mathcal{A}(\rho^k + S^k) + e^k - b - \frac{y^k}{\alpha})$.
  Update $S^{k+1} = shrink_{\gamma/\tau_2}(S^k - \tilde{S}^k)$.
  Calculate $\tilde{e}^k := \frac{\tau_3 e^k - \alpha(\mathcal{A}(\rho^k + S^k) - b - \frac{y^k}{\alpha})}{\theta + \alpha + \tau_3}$.
  Update $e^{k+1} = \tilde{e}^k$.
  Update $y^{k+1} = y^k - \kappa\alpha(\mathcal{A}(\rho^{k+1} + S^{k+1}) + e^{k+1} - b)$.
End For

define a triple $x^k := (\rho^k, S^k, e^k)$ of the primal variables and a quadruple $u^k := (\rho^k, S^k, e^k, y^k)$ of the primal and dual variables. The main idea behind the proof is to establish the convergence of u^k to $u^* := (\rho^*, S^*, e^*, y^*)$, where ρ^*, S^*, e^* and y^* are the optimal primal and dual solutions of (4).

To characterize the convergence, define multi-tuples $G_x := (P_1 + \alpha\mathcal{A}^H\mathcal{A}, P_2 + \alpha\mathcal{A}^H\mathcal{A}, P_3 + \alpha I)$ and $G := (P_1 + \alpha\mathcal{A}^H\mathcal{A}, P_2 + \alpha\mathcal{A}^H\mathcal{A}, P_3 + \alpha I, I/\alpha\kappa)$ of matrices, where I denotes the $d \times d$ identity matrix. Observe that in G_x and G, all the matrices are positive definite if we choose parameters such that $\alpha > 0$, $\kappa > 0$, $P_1 + \alpha\mathcal{A}^H\mathcal{A} \succ 0$, $P_2 + \alpha\mathcal{A}^H\mathcal{A} \succ 0$ and $P_3 + \alpha I \succ 0$, as suggested in Section III-B. For quadruples $G = (H_1, H_2, H_3, H_4)$ of matrices as well as (ρ, S, e, y) and (ρ', S', e', y') of variables, define the product

$$G(\rho, S, e, y) = (H_1\rho, H_2 S, H_3 e, H_4 y),$$

the inner product

$$\begin{aligned}&\langle(\rho, S, e, y), (\rho', S', e', y')\rangle_G \\ &= \langle\rho, \rho'\rangle_{H_1} + \langle S, S'\rangle_{H_2} + \langle e, e'\rangle_{H_3} + \langle y, y'\rangle_{H_4},\end{aligned}$$

and the norm

$$\|(\rho, S, e, y)\|_G = \sqrt{\|\rho\|_{H_1}^2 + \|S\|_{H_2}^2 + \|e\|_{H_3}^2 + \|y\|_{H_4}^2}.$$

Similar definitions apply to triples. Since all the matrices in G_x and G are positive definite, the norms are well defined.

In the proof, we first give a lower bound for $\|u^k - u^*\|_G^2$, the distance from the primal-dual pair to the optimal one in Lemma 1. Then we prove that $\|u^k - u^*\|_G^2 - \|u^{k+1} - u^*\|_G^2$, the decrease of the distance, is sufficiently large when the parameters are properly chosen in Lemma 2. Finally, we reach the main result that the limit point of the primal-dual pair is unique and optimal in Theorem 1.

Lemma 1. *For the quantum state filter* (6)*, if* $\alpha > 0$*,* $\kappa > 0$*,* $P_1 + \alpha\mathcal{A}^H\mathcal{A} \succ 0$*,* $P_2 + \alpha\mathcal{A}^H\mathcal{A} \succ 0$ *and* $P_3 + \alpha I \succ 0$*, then we have* $\forall k \geq 1$ *that*

$$\|u^k - u^*\|_G^2 - \|u^{k+1} - u^k\|_G^2 \geq h(u^k, u^{k+1}), \tag{20}$$

where

$$\begin{aligned}h(u^k, u^{k+1}) := &\frac{2 - \kappa}{\kappa^2\alpha}\|y^k - y^{k+1}\|_2^2 + \|x^k - x^{k+1}\|_{G_x}^2 \\ &+ \frac{2}{\kappa}(y^k - y^{k+1})^H(\mathcal{A}, \mathcal{A}, I)(x^k - x^{k+1}).\end{aligned} \tag{21}$$

Proof. The Lagrangian of (4) is

$$L = \|\rho\|_* + \gamma\|S\|_1 + \frac{\theta}{2}\|e\|_2^2 - \langle y, \mathcal{A}(\rho + S) + e - b\rangle, \quad \rho \in C. \tag{22}$$

Let (ρ^*, S^*, e^*, y^*) be any optimal primal-dual solution of (4). According to convex optimization theory, the KKT conditions of (4) are: (i) primal feasibility

$$\mathcal{A}(\rho^* + S^*) + e^* = b, \tag{23}$$

and (ii) stationarity

$$(\rho^*, S^*, e^*) = \arg\min_{\rho, S, e} L(\rho, S, e, y^*). \tag{24}$$

Observe that in (24), the optimization variables ρ, S and e are separable. Thus, for ρ we have

$$\rho^* = \arg\min_{\rho \in C} \|\rho\|_* - \langle y^*, \mathcal{A}(\rho)\rangle. \tag{25}$$

Similar to the discussion in Section III, the quantum state constraint $\rho \in C$ means that $\rho = \rho^H$, $\rho \succeq 0$ and $tr(\rho) = 1$, and hence leads to $\|\rho\|_* = tr(\rho) = 1$. Therefore, we can remove the nuclear norm term from (25) to obtain

$$\rho^* = \arg\min_{\rho \in C} \; - \langle y^*, \mathcal{A}(\rho)\rangle,$$

and consequently write its optimality condition as

$$\langle\rho^* - \rho, \mathcal{A}^H(y^*)\rangle \geq 0, \ \forall \rho \in C, \text{ and } \rho^* \in C. \tag{26}$$

For S we have

$$S^* = \arg\min_{S} \gamma\|S\|_1 - \langle y^*, \mathcal{A}(S)\rangle,$$

and consequently its optimality condition

$$\mathcal{A}^H(y^*) \in \gamma\partial\|S^*\|_1. \tag{27}$$

For e we have

$$e^* = \arg\min_{e} \frac{\theta}{2}\|e\|_2^2 - \langle y^*, e\rangle,$$

and the optimal condition is

$$\theta e^* = y^*. \tag{28}$$

In summary, the KKT conditions of (4) are (23), (26), (27) and (28).

We proceed to consider the quantum state filter in (6). For convenience, let $\hat{y} := y^k - \alpha(\mathcal{A}(\rho^{k+1} + S^{k+1}) + e^{k+1} - b)$ so that the dual update is $y^{k+1} = y^k - \kappa(y^k - \hat{y})$.

Using the fact of $\|\rho\|_* = tr(\rho) = 1$, (6a) is equivalent to

$$\begin{aligned}\rho^{k+1} = \arg\min_{\rho} \Big\{ &\frac{\alpha}{2}\|\mathcal{A}(\rho + S^k) + e^k - b - \frac{y^k}{\alpha}\|_2^2 \\ &+ \frac{1}{2}\|\rho - \rho^k\|_{P_1}^2, \ \rho \in C\Big\}.\end{aligned} \tag{29}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2019.2934755, IEEE Transactions on Automatic Control

Note that the optimality condition of (29) is given by

$$\langle \rho-\rho^{k+1},\ \alpha\mathcal{A}^H(\mathcal{A}(\rho^{k+1})+\mathcal{A}(S^k)+e^k-b-\frac{y^k}{\alpha}) + P_1(\rho^{k+1}-\rho^k)\rangle \geq 0,\ \forall \rho\in C. \tag{30}$$

Because $\rho^*\in C$, from (30) we have

$$\langle \rho^*-\rho^{k+1},\ \alpha\mathcal{A}^H(\mathcal{A}(\rho^{k+1})+\mathcal{A}(S^k)+e^k-b-\frac{y^k}{\alpha}) + P_1(\rho^{k+1}-\rho^k)\rangle \geq 0. \tag{31}$$

From (26) and $\rho^{k+1}\in C$, we also have

$$\langle \rho^*-\rho^{k+1}, \mathcal{A}^H(y^*)\rangle \geq 0. \tag{32}$$

Adding (31) and (32) and then substituting $\hat{y} := y^k - \alpha(\mathcal{A}(\rho^{k+1}+S^{k+1})+e^{k+1}-b)$, we have

$$\langle \mathcal{A}(\rho^{k+1}-\rho^*), \hat{y}-y^*-\alpha(\mathcal{A}(S^k-S^{k+1})+e^k-e^{k+1})\rangle + (\rho^*-\rho^{k+1})^H P_1(\rho^k-\rho^{k+1})\geq 0. \tag{33}$$

Similarly, the optimality condition for (6b) is given by

$$-\alpha\mathcal{A}^H(\mathcal{A}(\rho^k)-\mathcal{A}(S^{k+1})+e^k-b-\frac{y^k}{\alpha})-P_2(S^{k+1}-S^k) \in \gamma\partial\|S^{k+1}\|_1. \tag{34}$$

Combining (34) with (27), substituting $\hat{y}:=y^k-\alpha(\mathcal{A}(\rho^{k+1}+S^{k+1})+e^{k+1}-b)$ and using the fact that $\partial\|\cdot\|_1$ is a monotone operator, we get

$$\langle \mathcal{A}(S^{k+1}-S^*), \hat{y}-y^*-\alpha(\mathcal{A}(\rho^k-\rho^{k+1})+e^k-e^{k+1})\rangle + (S^*-S^{k+1})^H P_2(S^k-S^{k+1})\geq 0. \tag{35}$$

For (6c), its optimal condition is

$$\theta e^{k+1} = y^k-\alpha(\mathcal{A}(\rho^k)+\mathcal{A}(S^k)+e^{k+1}-b) + P_3(e^k-e^{k+1}). \tag{36}$$

Subtracting (36) by (28) and substituting the definition $\hat{y}:=y^k-\alpha(\mathcal{A}(\rho^{k+1}+S^{k+1})+e^{k+1}-b)$, we get

$$\theta(e^{k+1}-e^*) = \hat{y}-y^*-\alpha(\mathcal{A}(\rho^k-\rho^{k+1}) + \mathcal{A}(S^k-S^{k+1})+P_3(e^k-e^{k+1}). \tag{37}$$

Multiplying both sides of (37) with $e^{k+1}-e^*$ yields

$$\langle e^{k+1}-e^*, \hat{y}-y^*-\alpha(\mathcal{A}(\rho^k-\rho^{k+1})+\mathcal{A}(S^k-S^{k+1}) + P_3(e^k-e^{k+1})\rangle = \theta\|e^{k+1}-e^*\|_2^2\geq 0. \tag{38}$$

Adding the inequalities (33), (35) and (38), we have

$$\begin{aligned}&\langle \mathcal{A}(\rho^{k+1}-\rho^*)+\mathcal{A}(S^{k+1}-S^*)+(e^{k+1}-e^*), \hat{y}-y^*\rangle\\ &+(\rho^{k+1}-\rho^*)^H(P_1+\alpha\mathcal{A}^H\mathcal{A})(\rho^k-\rho^{k+1})\\ &+(S^{k+1}-S^*)^H(P_2+\alpha\mathcal{A}^H\mathcal{A})(S^k-S^{k+1})\\ &+(e^{k+1}-e^*)^H(P_3+\alpha I)(e^k-e^{k+1})\\ \geq&\alpha\langle(\mathcal{A},\mathcal{A},I)(x^{k+1}-x^*),(\mathcal{A},\mathcal{A},I)(x^k-x^{k+1})\rangle.\end{aligned} \tag{39}$$

Consider the equality

$$\begin{aligned}&(\mathcal{A},\mathcal{A},I)(x^{k+1}-x^*)\\ =&\mathcal{A}(\rho^{k+1}-\rho^*)+\mathcal{A}(S^{k+1}-S^*)+(e^{k+1}-e^*)\\ =&\frac{1}{\kappa\alpha}(y^k-y^{k+1}),\end{aligned} \tag{40}$$

which comes from the update of (6d) that

$$\mathcal{A}\rho^{k+1}+\mathcal{A}S^{k+1}+e^{k+1}-b=\frac{1}{\kappa\alpha}(y^k-y^{k+1}),$$

and comes from the feasibility of the optimal solution of (4) that

$$\mathcal{A}\rho^*+\mathcal{A}S^*+e^*-b=0.$$

Also note that

$$\begin{aligned}\hat{y}-y^* &= (\hat{y}-y^{k+1})+(y^{k+1}-y^*)\\ &= \frac{\kappa-1}{\kappa}(y^k-y^{k+1})+(y^{k+1}-y^*).\end{aligned} \tag{41}$$

With the above two equations (40) and (41), the inequality (39) can be rewritten as

$$\begin{aligned}&\langle\frac{1}{\kappa\alpha}(y^k-y^{k+1}), y^{k+1}-y^*\rangle\\ &+(\rho^{k+1}-\rho^*)^H(P_1+\alpha\mathcal{A}^H\mathcal{A})(\rho^k-\rho^{k+1})\\ &+(S^{k+1}-S^*)^H(P_2+\alpha\mathcal{A}^H\mathcal{A})(S^k-S^{k+1})\\ &+(e^{k+1}-e^*)^H(P_3+\alpha I)(e^k-e^{k+1})\\ \geq&\frac{1-\kappa}{\alpha\kappa^2}\|y^k-y^{k+1}\|_2^2+\frac{1}{\kappa}(y^k-y^{k+1})^H(\mathcal{A},\mathcal{A},I)(x^k-x^{k+1}).\end{aligned} \tag{42}$$

By the definitions of u^{k+1}, u^k, u^* and G, (42) is equivalent to

$$\begin{aligned}&(u^k-u^{k+1})^H G(u^{k+1}-u^*)\\ \geq&\frac{1-\kappa}{\alpha\kappa^2}\|y^k-y^{k+1}\|_2^2+\frac{1}{\kappa}(y^k-y^{k+1})^H(\mathcal{A},\mathcal{A},I)(x^k-x^{k+1}).\end{aligned} \tag{43}$$

Substituting the equality $\|u^k-u^*\|_G^2-\|u^{k+1}-u^*\|_G^2-\|u^k-u^{k+1}\|_G^2=2(u^k-u^{k+1})^HG(u^{k+1}-u^*)$ into (43) yields (20) and completes the proof. □

Lemma 2. *For the quantum state filter* (6)*, if $\alpha>0$, $\kappa>0$, and the matrices P_1, P_2 and P_3 are chosen such that*

$$\begin{cases}P_1\succ\alpha(\frac{1}{\xi_1}-1)\mathcal{A}^HA,\\ P_2\succ\alpha(\frac{1}{\xi_2}-1)\mathcal{A}^HA,\\ P_3\succ\alpha(\frac{1}{\xi_3}-1)I,\\ \xi_1+\xi_2+\xi_3<2-\kappa \text{ where } \xi_i>0, i=1,2,3,\end{cases} \tag{44}$$

then there exists a constant $\eta>0$ such that

$$h(u^k,u^{k+1})\geq\eta\|u^k-u^{k+1}\|_G^2, \tag{45}$$

and consequently

$$\|u^k-u^*\|_G^2-\|u^{k+1}-u^*\|_G^2\geq\eta\|u^k-u^{k+1}\|_G^2. \tag{46}$$

Proof. By the Cauchy-Schwartz inequality, it holds

$$\begin{aligned}&\frac{2}{\kappa}(y^k-y^{k+1})^H(\mathcal{A},\mathcal{A},I)(x^k-x^{k+1})\\ =&\frac{2}{\kappa}(y^k-y^{k+1})^H\times\\ &\big(\mathcal{A}(\rho^k-\rho^{k+1})+\mathcal{A}(S^k-S^{k+1})+(e^k-e^{k+1})\big)\\ \geq&-\frac{\xi_1+\xi_2+\xi_3}{\kappa^2\alpha}\|y^k-y^{k+1}\|_2^2-\frac{\alpha}{\xi_1}\|\mathcal{A}(\rho^k-\rho^{k+1})\|_2^2\\ &-\frac{\alpha}{\xi_2}\|\mathcal{A}(S^k-S^{k+1})\|_2^2-\frac{\alpha}{\xi_3}\|e^k-e^{k+1}\|_2^2.\end{aligned} \tag{47}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2019.2934755, IEEE Transactions on Automatic Control

Substituting (47) into the definition of $h(\cdot)$ in (21) yields

$$\begin{aligned} h(u^k-u^{k+1}) \geq & \frac{2-\kappa-\xi_1-\xi_2-\xi_3}{\kappa^2\alpha}\|y^k-y^{k+1}\|_2^2 \\ & +(\rho^k-\rho^{k+1})^H(P_1+\alpha(1-\frac{1}{\xi_1})\mathcal{A}^H\mathcal{A})(\rho^k-\rho^{k+1}) \\ & +(S^k-S^{k+1})^H(P_2+\alpha(1-\frac{1}{\xi_2})\mathcal{A}^H\mathcal{A})(S^k-S^{k+1}) \\ & +(e^k-e^{k+1})^H(P_3+\alpha(1-\frac{1}{\xi_3})I)(e^k-e^{k+1}). \end{aligned} \tag{48}$$

Conditions in (44) guarantee that $P_1+\alpha(1-(1/\xi_1))\mathcal{A}^H\mathcal{A}\succ 0$, $P_2+\alpha(1-(1/\xi_2))\mathcal{A}^H\mathcal{A}\succ 0$, $P_3+\alpha(1-(1/\xi_3))I\succ 0$ and $2-\kappa-\xi_1-\xi_2-\xi_3>0$. Therefore, from (48) we can find some $\eta>0$ such that (45) holds. By (20) in Lemma 1, (46) follows immediately. □

Lemma 2 provides guidelines for how to set the parameters. If specifically choosing the matrices $P_1=\tau_1 I-\alpha\mathcal{A}^H\mathcal{A}$, $P_2=\tau_2 I-\alpha\mathcal{A}^H\mathcal{A}$ and $P_3=\tau_3 I$ in the quantum state filter (7), from (44) we must have $\tau_1\succ(\alpha/\xi_1)\mathcal{A}^H\mathcal{A}$, $\tau_2\succ(\alpha/\xi_2)\mathcal{A}^H\mathcal{A}$ and $\tau_1>\alpha/\xi_1$, where $\xi_1+\xi_2+\xi_3<2-\kappa$.

Given Lemma 1 and Lemma 2, we are ready to obtain the global convergence of the QSF (6) in the following Theorem 1:

Theorem 1. *For the quantum state filter* (6)*, if $\alpha>0$, $\kappa>0$, and the matrices P_1, P_2 and P_3 are chosen to satisfy* (44)*, then the sequence $\{u^k\}$ converges to an optimal solution u^* to* (4)*.*

Proof. Due to the non-negativity of the right-hand side of (46), we have: (i) $\|u^k-u^*\|_G^2$ is monotonically decreasing and thus converges to a nonnegative constant; (ii) $\|u^k-u^{k+1}\|_G\to 0$; (iii) $\{u^k\}$ lies in a compact region.

It follows from (ii) that $\rho^k-\rho^{k+1}\to 0$, $S^k-S^{k+1}\to 0$, $e^k-e^{k+1}\to 0$ and $y^k-y^{k+1}\to 0$ as $k\to\infty$. Then by the update in (6d), $y^k=y^{k-1}-\kappa\alpha(\mathcal{A}(\rho^k+S^k)+e^k-b)$ implies that $\mathcal{A}(\rho^k+S^k)+e^k-b\to 0$. From (iii), u^k has a subsequence $\{u^{k_j}\}$ that converges to $\bar{u}=(\bar{\rho},\bar{S},\bar{e},\bar{y})$. Therefore, $\bar{u}=(\bar{\rho},\bar{S},\bar{e},\bar{y})$ is a limit point of $\{u^k=(\rho^k,S^k,e^k,y^k)\}$ where $\mathcal{A}(\bar{\rho}+\bar{S})+\bar{e}-b=0$ and $\bar{\rho}\in C$ (because ρ^k stays in C throughout the optimization process). Below we prove that any limit point is an optimal solution to (4).

Consider any limit point $\bar{u}=(\bar{\rho},\bar{S},\bar{e},\bar{y})$. Observing (30), replacing (ρ^k,S^k,e^k,y^k) and $(\rho^{k+1},S^{k+1},e^{k+1},y^{k+1})$ by $(\bar{\rho},\bar{S},\bar{e},\bar{y})$, as well as utilizing the fact of $\mathcal{A}(\bar{\rho}+\bar{S})+\bar{e}-b=0$, from (30) we have

$$\langle\mathcal{A}^H(\bar{y}),\bar{\rho}-\rho\rangle\geq 0,\ \forall\rho\in C. \tag{49}$$

Similarly, (34) implies that

$$\mathcal{A}^H(\bar{y})\in\gamma\partial\|\bar{S}\|_1, \tag{50}$$

and (37) implies that

$$\theta\bar{e}=\bar{y}. \tag{51}$$

Along with $\bar{\rho}\in C$ and $\mathcal{A}(\bar{\rho}+\bar{S})+\bar{e}-b=0$, (49), (50) and (51) imply that the quadruple $(\bar{\rho},\bar{S},\bar{e},\bar{y})$ satisfies the KKT conditions of (1) – see (23), (26), (27) and (28) – and is thus an optimal solution to (4).

To complete the proof, it remains to derive that the primal-dual sequence $\{u^k=(\rho^k,S^k,e^k,y^k)\}$ has a unique limit point. Let $\bar{u}_1=(\bar{\rho}_1,\bar{S}_1,\bar{e}_1,\bar{y}_1)$ and $\bar{u}_2=(\bar{\rho}_2,\bar{S}_2,\bar{e}_2,\bar{y}_2)$ be any two limit points. As we have shown, both $\bar{u}_1$ and $\bar{u}_2$ are optimal solutions to (4). Thus, u^* in (20) can be replaced by $\bar{u}_1$ and $\bar{u}_2$. This results in $\|u^{k+1}-\bar{u}_i\|_G^2\leq\|u^k-\bar{u}_i\|_G^2$, $i=1,2$, and consequently $\lim_{k\to\infty}\|u^k-\bar{u}_i\|_G^2=\nu_i$, $i=1,2$, where ν_1 and ν_2 are two finite constants. Using the equality

$$\begin{aligned} &\|u^k-\bar{u}_1\|_G^2-\|u^k-\bar{u}_2\|_G^2 \\ =&-2\langle u^k,\bar{u}_1-\bar{u}_2\rangle_G+\|\bar{u}_1\|_G^2-\|\bar{u}_2\|_G^2, \end{aligned}$$

and taking limit for $u^k\to\bar{u}_1$ and $u^k\to\bar{u}_2$, respectively, we have

$$\nu_1-\nu_2=-2\langle\bar{u}_1,\bar{u}_1-\bar{u}_2\rangle_G+\|\bar{u}_1\|_G^2-\|\bar{u}_2\|_G^2=-\|\bar{u}_1-\bar{u}_2\|_G^2,$$

and

$$\nu_1-\nu_2=-2\langle\bar{u}_2,\bar{u}_1-\bar{u}_2\rangle_G+\|\bar{u}_1\|_G^2-\|\bar{u}_2\|_G^2=\|\bar{u}_1-\bar{u}_2\|_G^2.$$

Therefore, we must have $\|\bar{u}_1-\bar{u}_2\|_G^2=0$, which indicates that the limit point of $\{u^k=(\rho^k,S^k,e^k,y^k)\}$ is unique and completes the proof. □

Remark. *Our theoretical proof is along the lines of that in [26]. However, when the PJ-ADMM is applied for the QSF, the proof must be adapted according to the special quantum features. The measurement matrix in the linear constraint is complex, which is different to the case in [26]. The quantum state to be estimated is also complex, and must satisfy the quantum constraints, namely, $\rho^H=\rho$, $\rho\succeq 0$, and $tr(\rho)=1$. Thus, the ρ-subproblem is a constrained one, which calls for constrained optimization techniques in our proof.*

V. Numerical comparison experiments

In this Section, we will provide three numerical comparison experiments to demonstrate the sampling rate efficiency in V-A and convergence properties in V-B of the proposed PJ-ADMM, compared with ADMM [23] and IST-ADMM [25]. In V-C, we reconstructed an 8-qubit experimentally-realized state using the proposed PJ-ADMM and compared our result with that of maximum likelihood estimation (MLE) [15], [34].

In the experiments, the measurement is generated from $b=\mathcal{A}(\ddot{\rho}+\ddot{S})+\ddot{e}$, where $\ddot{\rho}$ represents the true quantum state to recover, $\ddot{S}$ is the sparse disturbance and $\ddot{e}$ is the Gaussian noise. We generate $\ddot{\rho}$ by

$$\ddot{\rho}=\frac{\psi_r{\psi_r}^H}{tr(\psi_r{\psi_r}^H)}, \tag{52}$$

where ψ_r is a complex $d\times r$ Wishart matrix with i.i.d. complex random Gaussian entries [31]. We set $d=2^n$ with $n=5$ and $r=2$. The disturbance matrix $\ddot{S}\in\mathbb{R}^{d\times d}$ has $d^2/10$ nonzero entries located uniformly and randomly, with values satisfying the Gaussian distribution $\mathcal{N}(0,\|\rho\|_{\mathcal{F}}/100)$. Gaussian noise $\ddot{e}$ is generated by the MATLAB command randn$(n,1)$ multiplied by an appropriately chosen constant to obtain a desired signal-to-noise ratio (SNR). As is usually done, we measure the SNR

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2019.2934755, IEEE Transactions on Automatic Control

of b in terms of decibel (dB) and use the following definition of SNR of b:

$$\text{SNR}(b) = 20\log_{10}\left(\frac{\|b - \text{E}(b)\|_2}{\|\ddot{e}\|_2}\right), \tag{53}$$

where $\text{E}(b)$ represents the mean value of b. We set $\text{SNR}(b) = 60$ dB. The linear operator $\mathcal{A}$ is constructed from the Pauli bases as shown in Section II such that $\mathcal{A}\mathcal{A}^H = I$. As a result, we have $\lambda_{\max}(\mathcal{A}^H\mathcal{A}) = 1$, which is useful in setting the parameters for PJ-ADMM. In the optimization problem (4), the weight is $\gamma = 1/\sqrt{d}$ as suggested by [32] and we set $\theta = 1$.

The performance of an estimated ρ is evaluated by three criteria. The first one is the normalized distance $D(\rho, \ddot{\rho})$ between the estimated ρ and the true density matrix $\ddot{\rho}$, defined by

$$D(\rho, \ddot{\rho}) = \|\rho - \ddot{\rho}\|_{\mathcal{F}}^2 / \|\ddot{\rho}\|_{\mathcal{F}}^2. \tag{54}$$

The second one is the normalized distance $D(\rho, \rho^*)$ between the estimate ρ and the optimal solution ρ^*, defined by

$$D(\rho, \rho^*) = \|\rho - \rho^*\|_{\mathcal{F}}^2 / \|\rho^*\|_{\mathcal{F}}^2. \tag{55}$$

The third one is the fidelity defined by

$$\text{fidelity} = tr\left(\sqrt{\sqrt{\ddot{\rho}}\rho\sqrt{\ddot{\rho}}}\right). \tag{56}$$

Note that the fidelity lies within $[0, 1]$, and its value is 1 when the two states are identical [33].

We compare PJ-ADMM with two existing algorithms ADMM [23] and IST-ADMM [25]. As we have discussed in Section I, ADMM in [23] cannot guarantee the convergence to the optimal solution, while IST-ADMM in [25] is only able to handle the case of $r = 1$. In case of $r = 2$ as we set in the illustrations, IST-ADMM cannot guarantee the positive semidefiniteness of the solution. Note that another existing algorithm FP-ADMM in [24], which we do not compare in our illustrations, has the same solution accuracy as IST-ADMM but much higher computational complexity. All the illustrations were run in MATLAB R2016b on a laptop with 2.5GHz Intel Core 2 i5-3210M CPU and 6 GB memory.

According to the theoretical results given in Theorem 1, the parameters τ_1, τ_2, τ_3 and κ must satisfy $\tau_1 > \frac{3\alpha}{2-\kappa}\|\mathcal{A}\|^2$, $\tau_2 > \frac{3\alpha}{2-\kappa}\|\mathcal{A}\|^2$ and $\tau_3 > \alpha(\frac{3}{2-\kappa} - 1)$, which become $\tau_1 > \frac{3\alpha}{2-\kappa}$, $\tau_2 > \frac{3\alpha}{2-\kappa}$ and $\tau_3 > \alpha(\frac{3}{2-\kappa} - 1)$ since $\|\mathcal{A}\| = 1$. In the illustrations, the penalty parameter is $\alpha = 100$, while τ_1, τ_2, τ_3 and κ vary to demonstrate the impact on the performance of PJ-ADMM. The parameters of ADMM and IST-ADMM are hand-tuned to be the best. All the three algorithms are terminated if $k > k_{\max}$, and $k_{\max} = 1000$.

A. Impact of the Sampling Rate η in PJ-ADMM, ADMM and IST-ADMM

The sampling rate η defined in (2) is an important factor to quantify the measurement complexity in QSF. A low sampling rate η means that the number of measurements is much less than d^2, the size of ρ. However, the lower the sampling rate η, the more difficult to accurately recover the density matrix ρ.

We compare the proposed PJ-ADMM ($\kappa = 0.1$, $\tau_1 = \tau_2 = 158$ and $\tau_3 = 58$) with ADMM and IST-ADMM, and run them until $k > 1000$ under different sampling rates. The sampling rate increases from $\eta = 0.20$ to $\eta = 0.6$ with the incremental step 0.05.

Fig. 2 depicts $D(\rho, \ddot{\rho})$, the normalized distance between the estimate ρ and the true density matrix $\ddot{\rho}$, with different η, from which one can see that when the sampling rate is $\eta = 0.4$, the value of $D(\rho, \ddot{\rho})$ of PJ-ADMM is 0.0030 (the fidelity defined by (56) is 98.42%), and those of IST-ADMM and ADMM are 0.0053 (97.00%) and 0.3144 (71.10%), respectively. When the sampling rate is $\eta = 0.6$, the value of $D(\rho, \ddot{\rho})$ of PJ-ADMM decreases to 8.845×10^{-4} and the fidelity is 99.17%. The value of $D(\rho, \ddot{\rho})$ (fidelity) is 0.0013 (98.60%) for IST-ADMM and 0.1217 (88.90%) for ADMM. The experimental results demonstrate that PJ-ADMM and IST-ADMM both need much smaller sampling rates to achieve a target accuracy than ADMM.

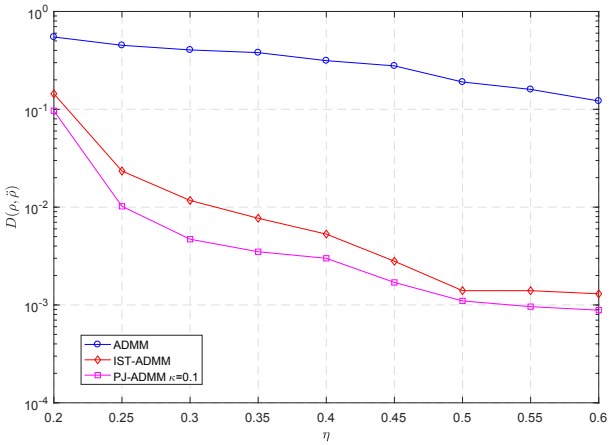


Fig. 2: Comparison of recovery accuracies of PJ-ADMM, ADMM and IST-ADMM that run for $k_{\max} = 1000$ iterations. The sampling ratio is η.

B. Convergence Properties of PJ-ADMM, IST-ADMM and ADMM

In order to verify the convergence performance, we compare the proposed PJ-ADMM with IST-ADMM and ADMM, in terms of the normalized distance $D(\rho, \rho^*)$ between the estimate ρ and the optimal solution ρ^*, as well as the normalized distance $D(\rho, \ddot{\rho})$ between the estimate ρ and the true density matrix $\ddot{\rho}$. We set the sampling rate $\eta = 0.4$ as in experiment A, and $\kappa = 0.1$ according to Theorem 1.

Fig. 3 depicts the evolution of $D(\rho, \rho^*)$ with respect to the number of iterations, from which one can see that when the algorithms run for 1000 iterations, the performance $D(\rho, \rho^*)$ of PJ-ADMM ($\kappa = 0.1$), IST-ADMM and ADMM are 2.31×10^{-10}, 1.58×10^{-3} and 0.145, respectively, which shows that the proposed PJ-ADMM has the best convergence performance among the three methods.

Fig. 4 further shows the evolution of $D(\rho, \ddot{\rho})$ with respect to the number of iterations. After 100 iterations, the values of $D(\rho, \ddot{\rho})$ for PJ-ADMM ($\kappa = 0.1$), IST-ADMM and ADMM

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2019.2934755, IEEE Transactions on Automatic Control

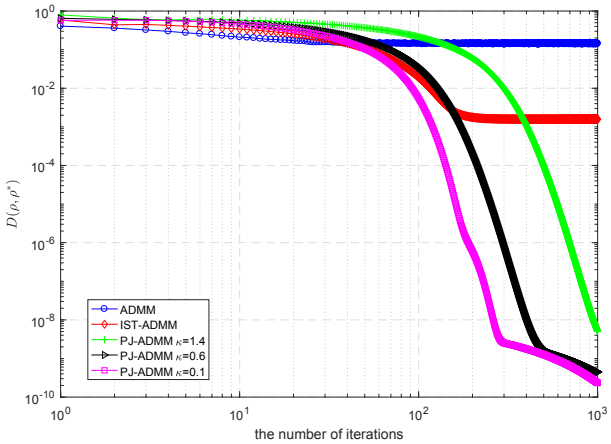


Fig. 3: Comparison of convergence rates of PJ-ADMM, ADMM and IST-ADMM that run for the sampling rate $\eta = 0.4$.

are 0.0113 (the fidelity is 96.45%), 0.0267 (92.44%) and 0.1517 (87.80%), respectively. At the end of 1000 iterations, the values of $D(\rho, \ddot{\rho})$ for PJ-ADMM ($\kappa = 0.1$), IST-ADMM and ADMM are 0.0022 (the fidelity is 98.65%), 0.0038 (97.40%) and 0.1523 (88.08%), respectively.

We also compare the run times of 200 iterations: 62.65s (ADMM), 1.46s (IST-ADMM), 1.36s (PJ-ADMM with $\kappa = 0.1$), 1.25s (PJ-ADMM with $\kappa = 0.6$) and 1.18s (PJ-ADMM with $\kappa = 1.4$). The iteration complexity of ADMM is significantly higher, while those of PJ-ADMM and IST-ADMM are similar.

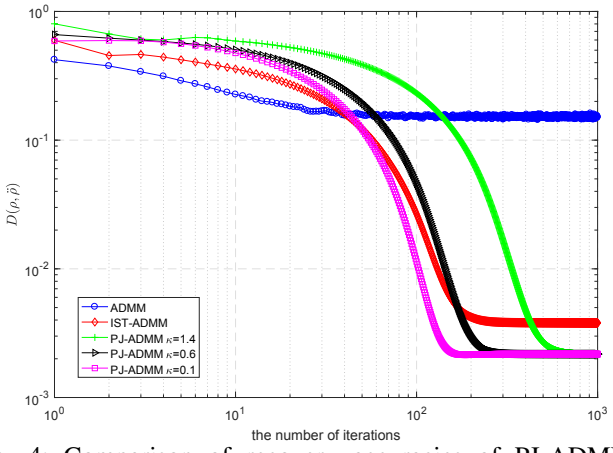


Fig. 4: Comparison of recovery accuracies of PJ-ADMM, ADMM and IST-ADMM that run for the sampling rate $\eta = 0.4$.

Among IST-ADMM and ADMM, the former has closer performance as PJ-ADMM. However, IST-ADMM has no convergence guarantee and the recovered density matrix may not be positive definite, which violates the quantum constraint. In addition, when the rank of the true density matrix $\ddot{\rho}$ increases from $r = 2$ to larger values, the disadvantage of IST-ADMM will be amplified.

Table I shows the values of $D(\rho, \ddot{\rho})$ for PJ-ADMM and IST-ADMM under different ranks, from which one can see that as the true rank increases from 2 to 4, the recovery accuracy of PJ-ADMM slightly increases from 0.0022 to 0.0067, while that of IST-ADMM sharply increases from 0.0038 to 0.0422.

rank of $\ddot{\rho}$	$D(\rho, \ddot{\rho})$ for IST-ADMM	$D(\rho, \ddot{\rho})$ for PJ-ADMM
2	0.0038	0.0022
3	0.0126	0.0033
4	0.0422	0.0067

TABLE I: Values of $D(\rho, \ddot{\rho})$ for PJ-ADMM and IST-ADMM under different ranks.

C. Reconstruction of the Experimentally-Realized States Using Proposed PJ-ADMM

We also apply the proposed PJ-ADMM to reconstruct the real 8-qubit quantum states provided in [34]. The states are no longer pure but entangled. We select a random subset of Pauli measurements with the sampling rate $\eta = 15\%$. We compare our reconstructed ρ to ρ_{ML} which is the reconstructed result obtained in the full tomography procedure using maximum likelihood estimation (MLE) [15], [34]. The absolute values of the reconstructed density matrices are shown in Fig. 5. After 1000 iterations, PJ-ADMM achieves a norm difference of $D(\rho_{ML}, \rho) = 0.0016$, with partial details shown in Fig. 5(c) and 5(d). This verifies the effectiveness of the algorithm under real quantum states, in addition to showing that it can achieve a reconstruction approaching MLE obtained from full tomography but with lower rate samples.

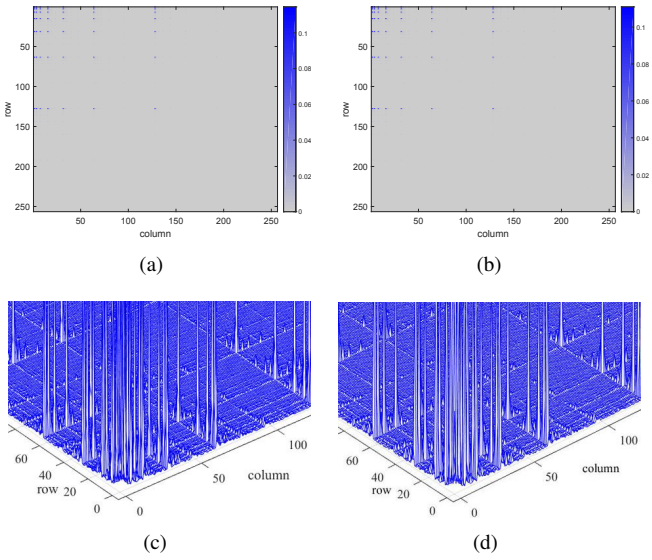


Fig. 5: Absolute value of corresponding reconstructed density matrix of the experimentally-realized state. (a) Maximum likelihood estimate of full quantum state tomography $|\rho_{ML}|$ [34]. (b) Reconstructed $|\rho|$ by PJ-ADMM with $\eta = 15\%$. (c) The magnitudes of $|\rho_{ML}|$. (d) The magnitudes of $|\rho|$.

VI. Conclusion

A rapidly convergent and efficient QSF is proposed in this paper. The proposed QSF can be used to estimate the density

matrix in the quantum system with sparse disturbances in the state and Gaussian noise in the measurements. The iterative QSF algorithm is appropriate to filter the quantum density matrix both off-line and on-line, which may be used as the feedback state in the quantum closed-loop control systems.

REFERENCES

[1] C. Zu, W. B. Wang, L. He, W. G. Zhang, C. Y. Dai, F. Wang, and L. M. Duan, "Experimental realization of universal geometric quantum gates with solid-state spins," *Nature*, vol. 514, no. 7520, pp. 72–75, 2014.

[2] A. Soare, H. Ball, D. Hayes, J. Sastrawan, M. C. Jarratt, J. J. Mcloughlin, X. Zhen, T. J. Green, and M. J. Biercuk, "Experimental noise filtering by quantum control," *Nature Physics*, vol. 10, no. 11, pp. 825–829, 2014.

[3] C. Schwemmer, L. Knips, D. Richart, H. Weinfurter, T. Moroder, M. Kleinmann, and O. Gühne, "Systematic errors in current quantum state tomography tools," *Physical Review Letters*, vol. 114, no. 8, p. 080403, 2015.

[4] D. Dong and I. R. Petersen, "Sliding mode control of two-level quantum systems," *Automatica*, vol. 48, no. 12, pp. 3089–3097, 2010.

[5] D. Gross, Y. K. Liu, S. T. Flammia, S. Becker, and J. Eisert, "Quantum state tomography via compressed sensing," *Physical Review Letters*, vol. 105, no. 15, p. 150401, 2010.

[6] G. S. Thekkadath, L. Giner, Y. Chalich, M. J. Horton, J. Banker, and J. S. Lundeen, "Direct measurement of the density matrix of a quantum system," *Physical Review Letters*, vol. 117, no. 12, 2016.

[7] W. T. Liu, T. Zhang, J. Y. Liu, P. X. Chen, and J. M. Yuan, "Experimental quantum state tomography via compressed sampling," *Physical Review Letters*, vol. 108, no. 17, p. 170403, 2012.

[8] N. Yamamoto and L. Bouten, "Quantum risk-sensitive estimation and robustness," *IEEE Transactions on Automatic Control*, vol. 54, no. 1, pp. 92–107, 2009.

[9] C. Xiang, I. R. Petersen, and D. Dong, "Performance analysis and coherent guaranteed cost control for uncertain quantum systems using small gain and popov methods," *IEEE Transactions on Automatic Control*, vol. 62, no. 3, pp. 1524–1529, 2017.

[10] D. Suess, A. Eisfeld, and W. Strunz, "Hierarchy of stochastic pure states for open quantum system dynamics," *Physical Review Letters*, vol. 113, no. 15, p. 150403, 2014.

[11] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[12] Y. Shoukry and P. Tabuada, "Event-triggered state observers for sparse sensor noise/attacks," *IEEE Transactions on Automatic Control*, vol. 61, no. 8, pp. 2079–2091, 2016.

[13] D. Gross, Y. Liu, S. T. Flammia, S. Becker, and J. Eisert, "Quantum state tomography via compressed sensing," *Physical review letters*, vol. 105, no. 15, p. 150401, 2010.

[14] A. Shabani, R. L. Kosut, M. Mohseni, H. Rabitz, M. A. Broome, M. P. Almeida, A. Fedrizzi, and A. G. White, "Efficient measurement of quantum dynamics via compressive sensing." *Physical Review Letters*, vol. 106, no. 10, p. 100401, 2011.

[15] C. Riofrío, D. Gross, S. Flammia, T. Monz, D. Nigg, R. Blatt, and J. Eisert, "Experimental quantum compressed sensing for a seven-qubit system." *Nature Communications*, vol. 8, p. 15305, 2017.

[16] S. Lloyd, M. Mohseni, and P. Rebentrost, "Quantum principal component analysis," *Nature Physics*, vol. 10, no. 9, pp. 631–633, 2014.

[17] C. A. Fuchs and A. Peres, "Quantum-state disturbance versus information gain: Uncertainty relations for quantum information," *Physical Review A*, vol. 53, no. 4, p. 2038, 1996.

[18] D. Dong, I. R. Petersen, and H. Rabitz, "Sampled-data design for robust control of a single qubit," *IEEE Transactions on Automatic Control*, vol. 58, no. 10, pp. 2654–2659, 2013.

[19] R. Long, G. Riviello, and H. Rabitz, "The gradient flow for control of closed quantum systems," *IEEE Transactions on Automatic Control*, vol. 58, no. 10, pp. 2665–2669, 2013.

[20] B. Qi, "A two-step strategy for stabilizing control of quantum systems with uncertainties," *Automatica*, vol. 49, no. 3, pp. 834–839, 2013.

[21] A. Smith, C. Riofrío, B. Anderson, H. Sosa-Martinez, I. Deutsch, and P. Jessen, "Quantum state tomography by continuous measurement and compressed sensing," *Physical Review A*, vol. 87, no. 3, p. 030102, 2013.

[22] Y. K. Liu, "Universal low-rank matrix recovery from Pauli measurements," *Advances in Neural Information Processing Systems*, pp. 1638–1646, 2011.

[23] K. Z. Li and S. Cong, "A robust compressive quantum state tomography algorithm using ADMM," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 6878–6883, 2014.

[24] K. Zheng, K. Z. Li, and S. Cong, "A reconstruction algorithm for compressive quantum tomography using various measurement sets," *Scientific Reports*, vol. 6, p. 38497, 2016.

[25] J. J. Zhang, K. Z. Li, S. Cong, and H. T. Wang, "Efficient reconstruction of density matrices for high dimensional quantum state tomography," *Signal Processing*, vol. 139, pp. 136–142, 2017.

[26] W. Deng, M. Lai, Z. Peng, and W. Yin, "Parallel multi-block ADMM with O(1/k) convergence," *Journal of Scientific Computing*, vol. 71, no. 2, pp. 712–736, 2017.

[27] D. S. Gonçalves, M. A. Gomesruggiero, and C. Lavor, "A projected gradient method for optimization over density matrices," *Optimization Methods & Software*, vol. 31, no. 2, pp. 1–14, 2015.

[28] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite element approximation," *Computers & Mathematics with Applications*, vol. 2, no. 1, pp. 17–40, 1976.

[29] J. Eckstein and D. P. Bertsekas, "On the douglas —rachford splitting method and the proximal point algorithm for maximal monotone operators," *Mathematical Programming*, vol. 55, no. 1, pp. 293–318, 1992.

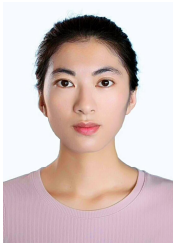
[30] C. Chen, B. He, Y. Ye, and X. Yuan, "The direct extension of admm for multi-block convex minimization problems is not necessarily convergent," *Mathematical Programming*, vol. 155, no. 1-2, pp. 57–79, 2016.

[31] K. Zyczkowski, K. A. Penson, I. Nechita, and B. Collins, "Generating random density matrices," *Journal of Mathematical Physics*, vol. 52, no. 6, p. 255, 2010.

[32] J. Wright, A. Ganesh, K. Min, and Y. Ma, "Compressive principal component pursuit," *IEEE International Symposium on Information Theory*, vol. 2, no. 1, pp. 1276–1280, 2012.

[33] S. T. Flammia and Y. K. Liu, "Direct fidelity estimation from few pauli measurements." *Physical Review Letters*, vol. 106, no. 106, p. 230501, 2011.

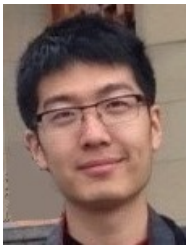
[34] H. Häffner, W. Hänsel, C. Roos, J. Benhelm, M. Chwalla, T. Körber, U. Rapol, M. Riebe, P. Schmidt and C. Becher, "Scalable multiparticle entanglement of trapped ions," *Nature*, vol. 438, no. 7068, p. 643, 2005.



Jiaojiao Zhang received the B.E. degree in automation from School of Automation, Harbin Engineering University, Harbin, China, in 2015. She is currently pursuing the master degree in control theory and control engineering from University of Science and Technology of China, Hefei, China. Her current research interests include convex optimization, quantum state estimation and quantum state filter.

Shuang Cong (M'07–SM'12) graduated from Beijing University of Aeronautics and Astronautics (BUAA), China, in 1982. She received the Ph. D. degree in System Engineering from the University of Rome La Sapienza, Italy, in 1995. She is currently a professor in the Department of Automation, University of Science and Technology of China (USTC). 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.

Qing Ling (M'07–SM'15) received the B.E. degree in automation and the Ph.D. degree in control theory and control engineering from University of Science and Technology of China, Hefei, Anhui, China, in 2001 and 2006, respectively. He was a Postdoctoral Research Fellow in the Department of Electrical and Computer Engineering, Michigan Technological University, Houghton, Michigan, USA, from 2006 to 2009, and an Associate Professor in the Department of Automation, University of Science and Technology of China, from 2009 to 2017. Since 2017, he has been a Professor in the School of Data and Computer Science, Sun Yat-Sen University, Guangzhou, Guangdong, China. His current research interests include decentralized network optimization and its applications. He is an Associate Editor of IEEE Transactions on Network and Service Management and IEEE Signal Processing Letters.



Kezhi Li obtained the Ph.D. degree at Imperial College London, and B. Eng. degree in Electronic Engineering from University of Science and Technology of China (USTC), Hefei, China, respectively. Prior to his current position, he was a research associate at University of Cambridge, a research fellow at Royal Institute of Technology (KTH) in Stockholm and a research assistant at Microsoft Research Asia (MSRA) and USTC. He is currently a research fellow at Department of Electrical and Electronic Engineering, Imperial College London. His research interests are quite broad in several inter-discipline areas, such as statistical signal processing and their applications in quantum tomography and bio-engineering.

Herschel Rabitz graduated from Harvard University in 1970, earning his Ph.D. degree in chemical physics. This was followed by post-doctoral work at the University of Wisconsin. In 1971 Professor Rabitz joined the faculty of the Department of Chemistry, Princeton University, and from July 1993 to July 1996 he was chair of the department. He is also an affiliated member of Princeton University 's Program in Applied and Computational Mathematics. Professor Rabitz 's research interests include the interface of chemistry, physics, and engineering, with principal areas of focus including: molecular dynamics, biophysical chemistry, chemical kinetics, and optical interactions with matter. An overriding theme throughout his research is the emphasis on molecular scale.