# Efficient and Fast Optimization Algorithms for Quantum State Filtering and Estimation

Kun Zhang
*Department of Automation*
*University of Science and Technology of China*
Hefei, China
kkzhang@mail.ustc.edu.cn

Shuang Cong*
*Department of Automation*
*University of Science and Technology of China*
Hefei, China
*Corresponding Author:
scong@mail.ustc.edu.cn

Jiao Ding
*Department of Automation*
*University of Science and Technology of China*
Hefei, China
dingjiao@mail.ustc.edu.cn

Jiaojiao Zhang
*Department of Systems Engineering and Engineering Management*
*The Chinese University of Hong Kong, Shatin, N.T.*
Hong Kong
jjzhang@se.cuhk.edu.hk

Kezhi Li
*Department of Electronic and Electrical Engineering*
*Imperial College London*
London, England, SW7 2AZ, UK
kezhi.li@imperial.ac.uk

***Abstract*—In this paper, based on Alternating Direction Multiplier Method (ADMM) and Compressed Sensing (CS), we develop three types of novel convex optimization algorithms for the quantum state estimation and filtering. Considering sparse state disturbance and measurement noise simultaneously, we propose a quantum state filtering algorithm. At the same time, the quantum state estimation algorithms for either sparse state disturbance or measurement noise are proposed, respectively. Contrast with other algorithms in literature, simulation experiments verify that all three algorithms have low computational complexity, fast convergence speed and high estimation accuracy at lower measurement rates.**

***Keywords*—quantum state estimation and filtering, Alternating Direction Multiplier Method, convex optimization algorithm**

## I. Introduction

Quantum state estimation is the foundation of quantum information research [1]. Quantum state tomography (QST) is a method used commonly for quantum state estimation, which was first proposed by Stokes. Cahill and Glauber proposed to restore quantum state information by reconstructing the density matrix of quantum states in 1969, which laid the basis of quantum tomography [2]. An *n*-qubit quantum system can be fully described by a density matrix $\rho \in \mathbb{C}^{d \times d}$ ($d = 2^n$) with $d^2$ parameters, which is a positive semi-definite and unit-trace Hermitian. For fully reconstructing the density matrix, people usually require $O(d^2)$ measurements [3]. Quantum measurements increase exponentially with the number of qubits. Obviously, the cost of experiment complexity is tremendous. Fortunately, we are usually interested in pure or nearly pure quantum states, which means $\rho$ is low rank and sparse. ($r = rank(\rho) = d$ and most singular values of $\rho$ are 0). Because of this characteristic, we can use the Compressed Sensing (CS) to greatly reduce the number of measurements required. The CS theory is a novel method of data compression and recovery [4]. If the high-dimensional signal or after some transformation is sparse, it can be compressed to a low-dimensional space without the loss by a measurement matrix. Research results show that as long as the measurement matrix satisfies the Restricted Isometry Property (RIP)[5], the original signal can be accurately recovered by solving an optimization problem. In practice, people generally use the Pauli matrix to construct the measurement matrix in QST. Furthermore, CS indicates that the minimum number of measurements which can accurately reconstruct the density matrix is $m = O(rd \log_2 d) = d^2$, and we define the measurement rate as $\eta = m / d^2$. The interference is inevitable in the actual quantum measurements. Interference can be divided into the disturbance in the quantum state and the noise in the measurement process. The disturbance of the state itself introduces sparse outliers in elements of certain locations in the density matrix. When considering disturbance and measurement noise simultaneous, it is a problem of quantum state filter. When only one of the interferences is considered, it is a quantum state estimation problem. What we stand in need of is how to fast and efficient reconstruct quantum state density matrix with high accuracy.

Alternating Direction Multiplier Method (ADMM), firstly proposed in 1970s by Gabay and Mercier, is a promising and effective optimization framework to deal with separable objective functions. In recent years, ADMM has been paid much attentions due to its success in a wide variety of applications such as data-distributed machine learning, compressed sensing, semi-definite programming and statistics etc. [6], and some scholars have applied ADMM to quantum state estimation. However, they are only aimed at sparse state disturbance. In 2014, Li et al. first used ADMM to solve the problem of ignoring quantum state constraints [7], and projected the obtained solution onto the quantum state constrained feasible set and the computational complexity is $O(d^6)$. In 2016, Zheng et al. solved the special problem where the rank of $\rho$ equals one and proposed an ADMM algorithm based on fixed point equation (FP-ADMM) to reduce the computational complexity to $\mathrm{O}(md^4)$ [8]. In 2017, Zhang et al. proposed an ADMM algorithm combined with iterative threshold shrinkage method (IST-ADMM) [9]. In contrast with previous works, we remark that our algorithms consider more comprehensively and generally.

In this paper, the research is mainly divided into three parts. Firstly, the problem of quantum state filtering for both interferences is studied. We use Prox-Jacobian ADMM method (PJ-ADMM) [10][11] to deal with the issue of the quantum state filtering, and propose a Quantum State Filtering ADMM (QSF-ADMM) algorithm. Secondly, we propose an Inexact ADMM (I-ADMM) algorithm for only having sparse disturbance. Thirdly, we develop an Improved-ADMM algorithm for estimating the quantum state with only measurement noise.

This paper is organized as follows. Three algorithms are proposed in detail in Sec.II. Experiments are carried out in Sec.III. Finally, Sec.IV is the conclusion.

## II. Quantum State Filtering and Estimation Algorithms

### A. *Quantum State Filtering Algorithm*

Sparse disturbance and measurement noise exist simultaneously. Assume the measurement noise is Gaussian noise. Enlightened by the filter in the control theory, we design a quantum state filtering algorithm to estimate quantum states, state disturbances and measurement noise simultaneously and reconstruct the density matrix accurately. When the objective variables are not less than 3, ADMM is difficult to guarantee the convergence. PJ-ADMM, which guarantees convergence by adding proximal term to each sub-problem, provides a new thought to solve convex optimization problems with 3 objective variables. It is noteworthy that both ADMM and PJ-ADMM are computational frameworks, which decompose the global problem into smaller and easier sub-problems, and obtain the global solution by coordinating the solutions of sub-problems. How to effectively solve each sub- problem is needed to determine the specific form of the problem.

The quantum state filtering problem can be described as reconstructing a low rank density matrix $\rho \in \mathbb{C}^{d\times d}$ from linear measurements $b \in \mathbb{R}^m$ with Gaussian noise $e \in \mathbb{R}^m$ and the density matrix with sparse disturbance $S \in \mathbb{R}^{d\times d}$. Let the measurement operator as $\mathcal{A}: \mathbb{C}^{d\times d} \to \mathbb{R}^m$, then the measurement formula is $b = \mathcal{A}(\rho + S) + e$. We learn that the low rank density matrix, sparse disturbance and Gaussian noise are all coupled in the measurement results. Meanwhile, only $m$ linear measurements are obtained, the information is not complete. In addition, the constraints of density matrix that is $\rho \succeq 0$, $\mathrm{tr}(\rho) = 1$, $\rho^H = \rho$ should be guaranteed.. The quantum state filtering problem can be formulated as:

$$\min \ \|\rho\|_* + \gamma\|S\|_1 + (\theta/2)\|e\|_2^2 \ \text{ s.t. } \ \mathcal{A}(\rho+S)+e=b, \ \rho \in \mathrm{C}, \tag{1}$$

in which the $\|\cdot\|_*$ is the nuclear norm, $\|\rho\|_* = \sum s_i$, $s_i$ are the singular values of $\rho$, $\|\cdot\|_1$ is the $l_1$ norm, $\|\cdot\|_2$ is the $l_2$ norm and $\rho \in \mathrm{C}$ indicates that the quantum state constraint $C := \{\rho \in \mathbb{C}^{d\times d} \mid \rho \succeq 0, \mathrm{tr}(\rho) = 1, \rho^H = \rho\}$ is satisfied. By means of minimizing $\|\rho\|_*$ and $\|S\|_1$, we force the density matrix to be low-rank while the state disturbance with sparseness, minimizing $\|e\|_2^2$ as a common method of filtering Gaussian noise.

The partial augmented Lagrangian of (1) is:

$$L := \|\rho\|_* + \gamma\|S\|_1 + (\theta/2)\|e\|_2^2 - \langle y, \mathcal{A}(\rho+S)+e-b\rangle + (\alpha/2)\|\mathcal{A}(\rho+S)+e-b\|_2^2,$$

where $\alpha > 0$ is the penalty parameter and $y \in \mathbb{R}^m$ is the Lagrange multiplier. The penalty term can relax condition of convergence. Thus, the constrained problem (1) can be transformed into an convex optimization problem :

$$\min \ \|\rho\|_* + \gamma\|S\|_1 + (\theta/2)\|e\|_2^2 + (\alpha/2)\|\mathcal{A}(\rho+S)+e-b-y/\alpha\|_2^2, \rho \in \mathrm{C}. \tag{2}$$

By substituting (2) into the PJ-ADMM frame [10], we can get the quantum state filtering iterates as (3).

$$\begin{cases} \rho^{k+1} = \arg\min\limits_{\rho\in C}\{\|\rho\|_* + \frac{\alpha}{2}\|\mathcal{A}(\rho+S^k)+e^k-b-y^k/\alpha\|_2^2 + \frac{1}{2}\|\rho-\rho^k\|_{P_1}^2\}, \\ S^{k+1} = \arg\min\limits_{S}\{\gamma\|S\|_1 + \frac{\alpha}{2}\|\mathcal{A}(\rho^k+S)+e^k-b-y^k/\alpha\|_2^2 + \frac{1}{2}\|S-S^k\|_{P_2}^2\}, \\ e^{k+1} = \arg\min\limits_{e}\{\frac{\theta}{2}\|e\|_2^2 + \frac{\alpha}{2}\|\mathcal{A}(\rho^k+S^k)+e-b-y^k/\alpha\|_2^2 + \frac{1}{2}\|e-e^k\|_{P_3}^2\}, \\ y^{k+1} = y^k - \kappa\alpha(\mathcal{A}(\rho^{k+1}+S^{k+1})+e^{k+1}-b), \end{cases} \tag{3}$$

where $\kappa > 0$ is the parameter for adjusting the Lagrange multiplier $y$ update step size; $(1/2)\|\rho-\rho^k\|_{P_1}^2$, $(1/2)\|S-S^k\|_{P_2}^2$ and $(1/2)\|e-e^k\|_{P_3}^2$ are the proximal terms, which added after each sub-problem to correct the error in time. $\|\cdot\|_{P_i}^2$ ($i$=1,2,3) is defined as $\|x_i\|_{P_i}^2 = x_i^T P_i x_i$, where $P_i \succeq 0$ is the parameter of the proximal terms and can be some symmetric and positive semi-definite matrix.

In the $(k+1)$ th iteration in (3), for a specific original variable, by fixing other original variables, we minimize the sum of the partial Lagrangian function and the proximal term, and update original variables $\rho$, $S$ and $e$ respectively.Finally, the Lagrange multiplier $y$ is updated immediately.

When we choose appropriate proximal term parameters, the solving of the sub-problems can be simplified. For the sub-problems of $\rho$ and $S$, non-differentiable norms and $\mathcal{A}$ involved in the quadratic penalty term make the direct solution complicated. However, through selecting $P_i = \tau_i I - \alpha\mathcal{A}^H\mathcal{A}$ ($i$=1,2), we can linearize the quadratic term of augmented Lagrangian, which can cancel the term $(\alpha/2)\rho^H\mathcal{A}^H\mathcal{A}\rho$ and add $(\tau_i/2)\rho^H\rho$ ; Because the sub-problem of $e$ has explicit solution, we have $P_3 = \tau_3 I$ ; $\tau_1 > 0$, $\tau_2 > 0$, $\tau_3 > 0$ are proximal step size. Therefore, we can solve the smooth terms of each sub-problem approximately using prox-linear approach and defining intermediate variables $\tilde{\rho}^k$, $\tilde{S}^k$ and $\tilde{e}^k$ as:

$$\tilde{\rho}^k := \rho^k - (\alpha/\tau_1)\mathcal{A}^H(\mathcal{A}(\rho^k + S^k) + e^k - b - y^k/\alpha), \qquad (4)$$

$$\tilde{S}^k := S^k - (\alpha/\tau_2)\mathcal{A}^H(\mathcal{A}(\rho^k + S^k) + e^k - b - y^k/\alpha), \qquad (5)$$

$$\tilde{e}^k := (\tau_3 e^k - \alpha(\mathcal{A}(\rho^k + S^k) - b - y^k/\alpha))/(\theta + \alpha + \tau_3). \qquad (6)$$

Consequently, we rewrite the algorithm (3) as follows:

$$\begin{cases} \rho^{k+1} = \arg\min_{\rho \in C}\{\|\rho\|_* + \frac{\tau_1}{2}\|\rho - \tilde{\rho}^k\|_F^2\}, \\ S^{k+1} = \arg\min_{S}\{\gamma\|S\|_1 + \frac{\tau_2}{2}\|S - \tilde{S}^k\|_F^2\}, \\ e^{k+1} = \arg\min_{e}\{\frac{\theta+\alpha+\tau_3}{2}\|e - \tilde{e}^k\|_2^2\}, \\ y^{k+1} = y^k - \kappa\alpha(\mathcal{A}(\rho^{k+1} + S^{k+1}) + e^{k+1} - b), \end{cases} \qquad (7)$$

in which $\|\cdot\|_F$ is the Frobenius norm.

Next we discuss the solution of three sub-problems $\rho$, $S$ and $e$ respectively.

*1) For solving the sub-problem of density matrix*

For the constraints of quantum state $\rho \succeq 0$, $\mathrm{tr}(\rho) = 1$, $\rho^H = \rho$, we know that $\|\rho\|_* = \mathrm{tr}(\rho) = 1$. Thus, the sub-problem of $\rho^{k+1}$ can be simplified to solve the following semi-definite programming (SDP) problem

$$\min \ \|\rho - (\tilde{\rho}^k + (\tilde{\rho}^k)^H)/2\|_F^2, \quad \text{s.t. } \rho \succeq 0, \mathrm{tr}(\rho)=1. \qquad (8)$$

Notably, (8) has an analytic solution. Solve the singular values $a_i (i = 1 \ldots d)$ of $(\tilde{\rho}^k + (\tilde{\rho}^k)^H)/2$ and decompose it into $V\, diag\{a_i\} V^H$, in which $V \in \mathbb{C}^{d \times d}$ is a unitary matrix, and the singular values $a_i$ are arranged in descending order as $a_1 \ge a_2 \ge \ldots \ge a_d$. The optimal solution of (8) is

$$\rho^{k+1} = V\, diag\{x_i\} V^H, \qquad (9)$$

where $\{x_i, i = 1, \ldots, d\}$ are the singular values of $\rho^{k+1}$ and calculated from $\min\ (1/2)\sum_{i=1}^d (x_i - a_i)^2$, s.t. $\sum_{i=1}^d x_i = 1, x_i > 0, \forall i$. The Lagrangian as $L(\{x_i\}, \beta) := (1/2)\sum_{i=1}^d (x_i - a_i)^2 + \sum_{i=1}^d (x_i - 1)$, $x_i \ge 0$, $\forall i$, where $\beta \in \mathbb{R}$ is the Lagrange multiplier. Base on the convex optimization theory, if $\beta$ is the optimal Lagrange multiplier, we can minimize $L(\{x_i\}, \beta)$ over $\{x_i\}$ by $\min\ (1/2)\sum_{i=1}^d (x_i - a_i + \beta)^2$, s.t. $x_i \ge 0$, $\forall i$, and get the optimal primal solution as $x_i = \max\{a_i - \beta, 0\}$, $\forall i$.

Due to restriction of $\sum_{i=1}^d x_i = 1, x_i > 0, \forall i$, the equivalent expression is $\sum_{i=1}^d \max\{a_i - \beta, 0\} = 1$. Thus, we let $\beta = a_i$, $i = 1, \ldots, d$ determine the interval in which the optimal $\beta$ belongs. Assuming that $\beta$ is in $[a_{t+1}, a_t]$, the optimal $\beta$ can be calculated by $\sum_{i=1}^t (a_i - \beta) = 1$ as:

$$\beta = (\sum_{i=1}^t a_i - 1)/t. \qquad (10)$$

Finally we can obtain $\{x_i\}$ as:

$$\begin{cases} x_i = a_i - \beta, \ \forall i \le t, \\ x_i = 0, \qquad \forall i \ge t+1. \end{cases} \qquad (11)$$

*2) For the solution of the sparse matrix sub-problem*

Because the regularization of $S$ is $l_1$ norm, the corresponding proximal operation is the well-known soft-threshold operation. Thus, the sub-problem exists closed-form solution:

$$S^{k+1} = shrink_{\gamma/\tau_2}(S^k - \tilde{S}^k), \qquad (12)$$

in which $shrink_{\gamma/\tau_2}$ is a soft threshold contraction operator . For any scalar $s$ there is

$$shrink_{\gamma/\tau_2}(s) := \max\{|s - \gamma/\tau_2|, 0\} sign(s - \gamma/\tau_2).$$

*3) For the solution of Gaussian noise sub-problem*

The first-order optimality condition of $e$ is given by:

$$e^{k+1} = \tilde{e}^k. \qquad (13)$$

To ensure the convergence of the quantum state filtering algorithm, the gradient step size parameters in (7) must satisfy

$$\tau_1 > 3\alpha/(2-\kappa), \tau_2 > 3\alpha/(2-\kappa), \tau_3 > \alpha(3/(2-\kappa) - 1). \qquad (14)$$

In short, the QSF-ADMM algorithm is as follows:

1) Obtain the output signal $b$ of the experimental system and construct the measurement matrix $\mathcal{A}$;
2) Initialize variables to $\rho^0=0$, $S^0=0$, $e^0=0$, $y^0=0$; Set algorithm parameters $\kappa > 0$, $\gamma > 0$, $\alpha > 0$, $\theta > 0$ and $\tau_1, \tau_2, \tau_3$ satisfying (14);
3) For $k = 1, 2, \ldots N$
4) Calculate $\tilde{\rho}^k, \tilde{S}^k, \tilde{e}^k$ according to (4), (5), (6) respectively;
5) Calculate singular values of $(\tilde{\rho}^k + (\tilde{\rho}^k)^H)/2$;
6) Calculate singular values $\{x_i, i = 1, \ldots, d\}$ of $\rho^{k+1}$ according to (10) and (11);
7) Update $\rho^{k+1}$ according to (9);
8) Update $S^{k+1}$ according to (12);
9) Update $e^{k+1}$ according to (13);
10) Update $y^{k+1}$ according to (7);
11) End For

## *B. Quantum State Estimation Algorithms*

If we only consider one of the interferences, the problem is reduced to a two-objective convex optimization problem. In either case, we propose the I-ADMM algorithm for the sparse disturbance and the Improved-ADMM algorithm for the Gaussian noise, respectively, which can realize quantum state estimation fast and efficiently.

*1) I-ADMM algorithm for sparse disturbance*

Quantum state measurements with only disturbance can be expressed as $b = \mathcal{A}(\rho + S)$, where $S \in \mathbb{R}^{d \times d}$ is the sparse disturbance matrix. The aim of the quantum state estimation is to reconstruct the density matrix with sparse disturbance. The original problem is decomposed as minimizing the nuclear norm of density matrix with quantum state constraints and the

$l_1$ norm of sparse disturbance. Thus, the problem can be transformed a convex optimization problem as:

$$\min \ \|\rho\|_* + \gamma \|S\|_1 + I_C(\rho), \quad \text{s.t.} \quad \mathcal{A}(\rho + S) = b, \tag{15}$$

where $\gamma > 0$ is the weight factor, $I_C(\rho)$ is the indicator function on a convex set $C$ to make $\rho$ in constraints,

$$I_C(\rho) = \begin{cases} 0 & \text{if } \rho \succeq 0, \mathrm{tr}(\rho) = 1, \rho^H = \rho, \\ \infty & \text{otherwise.} \end{cases} \tag{16}$$

Via the augmented Lagrangian function, we can transform (15) to an unconstrained problem:

$$\min \ \|\rho\|_* + I_C(\rho) + \gamma \|S\|_1 + (\alpha/2) \|\mathcal{A}(\rho + S) - b - y/\alpha\|_2^2,$$

where $\alpha > 0$ is the penalty parameter and $y \in \mathbb{R}^m$ is the Lagrange multiplier.

For the (15) with two-objective variables, the classical ADMM framework decomposes the original problem into two small sub-problems[6], and optimizes the two sub-problems alternatively. After one sweep of updating $\rho$ and $S$, the multiplier $y$ is updated by the gradient ascent method immediately. In brief, the ADMM algorithm is

$$\begin{cases} \rho^{k+1} = \arg\min_{\rho} \{ \|\rho\|_* + I_C(\rho) + \frac{\alpha}{2} \|\mathcal{A}(\rho + S^k) - b - y^k/\alpha\|_2^2 \}, \\ S^{k+1} = \arg\min_{S} \{ \gamma \|S\|_1 + \frac{\alpha}{2} \|\mathcal{A}(\rho^{k+1} + S) - b - y^k/\alpha\|_2^2 \}, \\ y^{k+1} = y^k - \kappa\alpha(\mathcal{A}(\rho^{k+1} + S^{k+1}) - b), \end{cases} \tag{17}$$

where $\kappa > 0$ is the parameter for adjusting the Lagrange multiplier step size. Normally $\kappa = 1$ and studies have shown that $\kappa$ can speed up the convergence when it is adjusted in $(0, (\sqrt{5}+1)/2)$.

Because nuclear norm, indicator function and $l_1$ norm are non-differentiable, the solutions of $\rho^{k+1}$ and $S^{k+1}$ in (17) are tricky and no closed-form solution. Therefore, we solve the two sub-problems inexactly to reduce the computational complexity by proximal gradient method. By gradient descent of least square terms in sub-problems $\rho^{k+1}$ and $S^{k+1}$ respectively, we get $\tilde{\rho}$ and $\tilde{S}$ as:

$$\tilde{\rho} := \rho^k - \tau_1 \mathcal{A}^H (\mathcal{A}(\rho^k + S^k) - b - y^k/\alpha), \tag{18}$$

$$\tilde{S} := S^k - \tau_2 \mathcal{A}^H (\mathcal{A}(\rho^{k+1} + S^k) - b - y^k/\alpha), \tag{19}$$

where $\tau_1 > 0$ and $\tau_2 > 0$ are the proximal step size.

For sub-problem of $\rho^{k+1}$, the $\|\rho\|_*$ is a constant when the quantum state constraints satisfied. Therefore, we can obtain the solution $V diag\{x_i\} V^H$ through solving the problem $\min_{\rho \in C} \|\rho - \tilde{\rho}\|_F^2$ whose solution process is the same as in (7). The sub-problem of $S^{k+1}$ has soft threshold solution.

In summary, we propose the I-ADMM algorithm as:

$$\begin{cases} \rho^{k+1} = V diag\{x_i\} V^H, \\ S^{k+1} = shrink_{\gamma\tau_2/\alpha}(S^k - \tilde{S}), \\ y^{k+1} = y^k - \kappa\alpha(\mathcal{A}(\rho^{k+1} + S^{k+1}) - b). \end{cases} \tag{20}$$

In the simulation experiment, the weight $\gamma$ is set to $1/\sqrt{d}$ and slightly larger $\kappa$ can accelerate convergence. In order to ensure the convergence of the algorithm, it is proved that $\tau_1$ and $\tau_2$ should satisfy:

$$\tau_1 \lambda_{\max} < 1, \ \tau_2 \lambda_{\max} + \kappa < 2, \tag{21}$$

where $\lambda_{\max}$ is the maximum eigenvalue of $\mathcal{A}^H \mathcal{A}$. Because $\mathcal{A}$ is generate by Pauli matrices, the $\lambda_{\max}$ of $\mathcal{A}^H \mathcal{A}$ equals 1.

*2) Improved-ADMM algorithm for Gaussian noise*

This sub-section focuses on the quantum state estimation problem with only the noise in measurement process. We introduce auxiliary variables $e \in \mathbb{C}^m$, and the measurement result is $b = \mathcal{A}(\rho) + e$. We decompose the original problem into two sub-problems, which are minimizing the nuclear norm of density matrix with quantum state constraints and minimizing the $l_2$ norm of Gaussian noise. The problem can be written as:

$$\min \ \|\rho\|_* + I_C(\rho) + (1/2\gamma) \|e\|_2^2, \quad \text{s.t.} \quad \mathcal{A}(\rho) + e = b. \tag{22}$$

in which $\gamma > 0$ is the weight factor, $I_C(\rho)$ is the indicator function as same as (16).

With the help of the augmented Lagrangian, the constrained problem (22) can be rewritten as:

$$\min \ \|\rho\|_* + I_C(\rho) + (1/2\gamma) \|e\|_2^2 + (\alpha/2) \|\mathcal{A}(\rho) + e - b - y/\alpha\|_2^2,$$

where $\alpha > 0$ is the penalty parameter and $y \in \mathbb{R}^m$ is the Lagrange multiplier. Using ADMM iterative framework [6], we have

$$\begin{cases} e^{k+1} = \arg\min_{e} \{ (1/2\gamma) \|e\|_2^2 + (\alpha/2) \|\mathcal{A}(\rho^k) + e - b - y^k/\alpha\|_2^2 \}, \\ \rho^{k+1} = \arg\min_{\rho} \{ \|\rho\|_* + I_C(\rho) + (\alpha/2) \|\mathcal{A}(\rho) + e^{k+1} - b - y^k/\alpha\|_2^2 \}, \\ y^{k+1} = y^k - \kappa\alpha(\mathcal{A}(\rho^{k+1}) + e^{k+1} - b), \end{cases} \tag{23}$$

where $\kappa > 0$ is the parameter for adjusting the Lagrange multiplier update step size. For calculating proximal step of the least square term of $\rho^{k+1}$ in (23), we get $\tilde{\rho}$ as:

$$\tilde{\rho} := \rho^k - \tau(\mathcal{A}^H(\mathcal{A}(\rho) + e^{k+1} - b - y^k/\alpha)). \tag{24}$$

where $\tau > 0$ is step size.

For sub-problem of $\rho^{k+1}$, the $\|\rho\|_*$ is a constant when the quantum state constraints satisfied. we can obtain the solution $V diag\{x_i\} V^H$ through solving the problem $\min_{\rho \in C} \|\rho - \tilde{\rho}\|_F^2$. The later solution process is the same as in (7). For the solution of sub-problem $e^{k+1}$, we can solve it by making the first derivative equal to 0.

Therefore, we propose the Improved-ADMM algorithm as:

$$\begin{cases} e^{k+1} = (\gamma\alpha/(1+\gamma\alpha))(y^k/\alpha - \mathcal{A}(\rho^k) + \mathrm{b}), \\ \rho^{k+1} = V diag\{x_i\} V^H, \\ y^{k+1} = y^k - \kappa\alpha(\mathcal{A}(\rho^{k+1}) + e^{k+1} - b). \end{cases} \tag{25}$$

In order to ensure the convergence of the algorithm, it is proved that the parameter $\tau$ must meet:

$$\tau\lambda_{\max} + \kappa < 2, \tag{26}$$

where $\lambda_{\max}$ =1 is the maximum eigenvalue of $\mathcal{A}^H\mathcal{A}$.

We remark that the core of three algorithms in this section is to decompose the problem without closed-form solution into a multi-step method with less computational complexity in each step, which greatly reduces the complexity of the algorithm. Compared with previous algorithms, the maximal computation of each algorithm in this paper is the multiplication between an $m \times d^2$ matrix and a $d^2$ vector, whose computation complexity is $O(md^2)$. We give the convergence conditions of the three algorithms explicitly. However, Li's ADMM algorithm cannot guarantee convergence because it considers the objective function and constraints separately. FP-ADMM and IST-ADMM algorithms can only deal with $rank(\rho) = 1$. (All three algorithms in this paper are also valid for $rank(\rho) > 1$.) Moreover, QSF-ADMM algorithm can be considered as the Jacobi type method while I-ADMM and Improved-ADMM algorithms as Gauss-Seidel type method. The difference is that the Gauss-Seidel type method utilizes the latest information of $k+1$ iteration step, such as the sub-problem of $S^{k+1}$ in (17) and the sub-problem of $\rho^{k+1}$ in (23). Thus, I-ADMM and Improved-ADMM algorithms are believed to converge more faster than QSF-ADMM algorithm method because the latest information is always used.

## III. Numerical Experiments

In the experiments, we firstly describe the parameters' settings and the algorithm performance index. Then, the simulation experiments consists of two parts: (1) comparison of the three algorithms in this paper with ADMM and IST-ADMM algorithms; (2) the reconstruction performance of the three algorithms with different measurement rates.

In the simulation experiments, the measured value vector is constructed by $b = \mathcal{A}(\tilde{\rho} + \tilde{S}) + \tilde{e}$, where $\tilde{\rho}$ is the actual density matrix to be restored, $\tilde{S}$ is the sparse interference matrix, and $\tilde{e}$ is the measurement noise. To insure that $\tilde{\rho}$ is in an arbitrary pure/superposition state, we set $rank(\tilde{\rho}) = 1$. The true density matrix $\tilde{\rho}$ is generated by $\tilde{\rho} = (\psi_r \psi_r^H) / tr(\psi_r \psi_r^H)$. $\psi_r$ is a Wishart matrix of complex domain $\mathbb{C}^{d \times r}$, and its elements obey a random Gaussian distribution. $\tilde{S} \in \mathbb{R}^{d \times d}$ contains $d^2/10$ non-zero elements, and the position is random, and the amplitude satisfies the Gaussian distribution $N(0, \|\rho\|_F/100)$. The signal-to-noise ratio (SNR) of Gaussian noise is 70dB. The simulation experiments are conducted in MATLAB R2018a, Inter Core i5-8400M CPU, clocked at 2.8GHz, memory 8GB.

We use $Error = \|\rho - \tilde{\rho}\|_F^2 / \|\tilde{\rho}\|_F^2$ to measure the reconstruction accuracy of estimated density matrix $\rho$.

### A. *Experimental Simulation*

#### *1) Performance comparison with existing algorithms*

Fixed $\eta$ = 50% and 1000 iterations for 5-qubit system, we compare the convergence performance of the three proposed algorithms with ADMM [7] algorithm, IST-ADMM [9] algorithm. In the simulation experiment, the parameters of the five algorithms are adjusted to the optimum. Fig. 1 depicts the variation of estimation errors with the number of iterations.

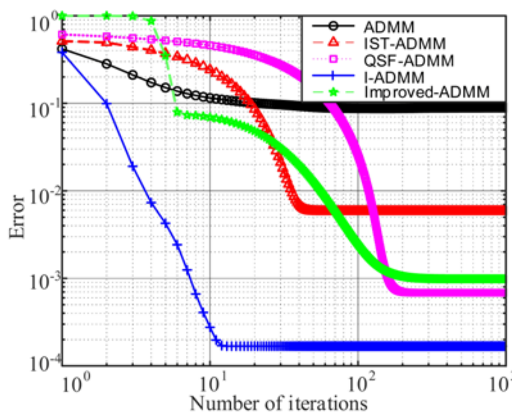


Fig. 1 Comparison of estimation errors of five algorithms with measurement rate of 50% for 5-qubit system.

We can see from the Fig. 1 that:

*1)* Under suitable parameters, the five algorithms can converge. For the number of iterations, I-ADMM algorithm achieves the estimation error of less than 0.001 with the least iterations. Compared with QSF-ADMM, I-ADMM and Improved-ADMM algorithm have faster convergence speed, which is consistent with the theoretical analysis.

*2)* ADMM, IST-ADMM, QSF-ADMM, I-ADMM and Improved-ADMM algorithm iterate 1000 times. The estimated error and time are 0.0819 (136.84s), 0.0058 (2.64s), 0.0007 (2.64s), 0.00017 (3.38s) and 0.0010 (2.80s), respectively. The time required by the three algorithms is not much different from IST-ADMM algorithm, and much lower than ADMM algorithm. Moreover, the three algorithms have lower estimation error bounds, which are obviously better than that of ADMM and IST-ADMM algorithm.

Therefore, we learn that the three algorithms in this paper have better estimation accuracy and convergence performance.

#### *2) Performance analysis of three algorithms*

For 3-qubit systems, we analyze the recovery accuracy and convergence performance of QSF-ADMM, I-ADMM and Improved-ADMM algorithms for quantum state density matrix at different measurement rates. And measurement rates $\eta$ are set as 37.5%, 75%, 100%, respectively. (the minimum of $\eta$ is 37.5% via compressed sensing theory.)

The adjustable parameters involved in all three algorithms include: 1) gradient descent step size $\tau_i$ ; 2) Lagrange multiplier update step size $\kappa$; 3) balance low rank and error term weight $\gamma$, $\theta$; 4) penalty parameter $\alpha$. According to the algorithm convergence requirements in Section II, the parameter used in the experiment are set as follows: $\gamma = 0.0001, \alpha = 1$; when $\eta$ = 37.5%, 75% and 100%, $\kappa$=1.451, 1.465 and 1.665 respectively; for QSF-ADMM algorithm $\tau_2$ = 0.28, 0.24 and 0.20 respectively, and $\tau_1$= 0.6, $\tau_3$= 0.6, $\theta$=1; for I-ADMM algorithm $\tau_2$ = 0.28, 0.24 and 0.20 respectively, and $\tau_1$ = 0.6; for Improved-ADMM algorithm $\tau$ = 0.6. The experimental results are shown in Fig. 2.

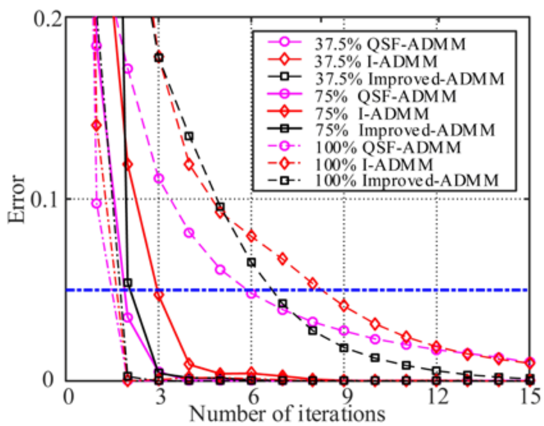


Fig. 2 Errors with different number of iterations for 3-qubit system. Dotted line, solid line and dashed dot line represent differences in measurement rates of 37.5%, 75%, 100%, respectively. Circle, diamond, and square represent different algorithms at the same measurement rate. The blue dashed line is the error of 0.05.

We can see from Fig. 2 that:

*1)* All three algorithms have fast convergence speed and can achieve high quantum state reconstruction accuracy. As the number of iterations increases, the estimation error decreases continuously. Under the minimum measurement rate of 37.5% for 3-qubit, the QSF-ADMM, Improved-ADMM, and I-ADMM algorithms achieve the estimation accuracy of over 95%, requiring 6, 7, 9 iterations, respectively and consuming 0.0067, 0.0436, 0.0548 seconds, respectively.

*2)* The same algorithm reduces the number of iterations as the measurement rate increases. For example, for the Improved-ADMM algorithm with sampling rates of 37.5%, 75%, and 100%, it achieves an estimation accuracy of more than 99% for 11, 3, and 2 iterations, respectively and consuming 0.0454, 0.0311, 0.0304 seconds, respectively.

*3)* The fixed measurement rate is 37.5% and the number of algorithm iterations is 15 times. The estimated errors of QSF-ADMM, Improved-ADMM, and I-ADMM algorithms are 98.83%, 99.05%, 99.02%, respectively and consuming 0.0123, 0.0461, 0.0606 seconds, respectively. We conclude that all three algorithms can quickly achieve a high estimation accuracy and QSF-ADMM is the shortest time-consuming algorithm. Furthermore, the proposed algorithm can restore the density matrix at a lower sampling rate and adopt the method of sparse storage to greatly reduce storage space. For example, when the number of quantum bits is 11 and the measurement operator $\mathcal{A}$ is constructed with $\eta$ =0.6%, 1.19G storage space is required, but full storage need 1573G.

## IV. Conclusion

We developed three fast and effective optimization algorithms to simultaneously filter the disturbance and/or Gaussian noise of quantum state in this paper. Simulation results showed that all three algorithms proposed can have fast convergence speed and high estimation accuracy at lower measurement rate, which embodies the superiorities of the algorithms in quantum state density matrix reconstruction .

## Acknowledgment

This work was supported by the National Natural Science Foundation of China under grants no. 61573330 and 61720106009

## References

[1] D. Gross, Y. K. Liu, S.T. Flammia, S. Becker, J. Eisert, " Quantum state tomography via compressed sensing,". Phys. Rev. Lett, vol. 105, pp. 150401, 2010.

[2] S.T. Flammia, D. Gross, Y.K. Liu, J. Eisert, "Quantum Tomography via Compressed Sensing: Error Bounds, Sample Complexity, and Efficient Estimators," New J. Phys,vol. 14(28), pp. 95022-95049, 2012.

[3] J. Yang, S. Cong, X. Liu, Z. Li, and K. Li, "Effective quantum state reconstruction using compressed sensing in NMR quantum computing," Phys. Rev. A, vol. 96, 052101, 2017.

[4] C.A. Riofrío, D. Gross, S.T. Flammia , T. Monz, D. Nigg , R. Blatt, & J. Eisert, "Experimental quantum compressed sensing for a seven-qubit system," Nature Communications, 15305, 2017.

[5] D. Gross, "Recovering low-rank matrices from few coefficients in any basis," IEEE Trans. Inf. Theory, vol. 57(3), pp.1548-1566, 2011.

[6] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, " Distributed optimization and statistical learning via the alternating direction method of multipliers," Foundations and Trends in Machine Learning,vol. 3(1), pp. 1-122, 2011.

[7] 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.

[8] K. Zheng, K. Z. Li, S. Cong, "A reconstruction algorithm for compressive quantum tomography using various measurement sets," Scientific Reports, 6:38497, 2016.

[9] 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.

[10] W. Deng, M.J. Lai, Z. Peng, et al. "Parallel Multi-Block ADMM with o (1/k) Convergence," Journal of Scientific Computing,vol. 71, pp.712-736, 2017.