Signal Processing 139 (2017) 136–142

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Efficient reconstruction of density matrices for high dimensional quantum state tomography

Jiaojiao Zhang[a], Kezhi Li[b], Shuang Cong[a,*], Haitao Wang[c]

[a] Department of Automation, University of Science and Technology of China, Hefei 230027, China
[b] Imperial College London, London W12 0NN, United Kingdom
[c] Beijing Institute of Satellite Information Engineering, Beijing 100086, China

ARTICLE INFO

Article history:
Received 16 February 2017
Accepted 7 April 2017
Available online 13 April 2017

Keywords:
Compressive sensing
Quantum state tomography
ADMM
ISTA

ABSTRACT

The conventional quantum state tomography (QST) needs large number of measurements to reconstruct the quantum state. Thanks to the compressive sensing (CS) theory, one can recover a pure or nearly pure quantum state with an acceptable accuracy given much fewer number of measurements. However, most existing algorithms for CS based QST are rather slow and difficult to be implemented in practice. To fill the gap between the CS theory and practical QST, this paper firstly applies an improved Alternating Direction Multiplier Method (ADMM) combining with the Iterative Shrinkage-Thresholding Algorithm (ISTA), IST–ADMM for short, aiming at improving the efficiency of QST problem in particular with much lower number of measurements. The IST–ADMM avoids computing the inverse of large-scale matrix, reduces the computational time and required memory space. The computation complexity is reduced from $\mathcal{O}(d^6)$ for least square (widely used in QST), and $\mathcal{O}(md^4)$ for Fixed Point-ADMM in our previous work, to IST–ADMM's $\mathcal{O}(md^2)$. The proposed algorithm makes it practical to reconstruct high dimensional quantum states provided fewer number of measurements. The simulations verify the superiority of the proposed algorithm, where it takes 3.13 minutes to reconstruct an 8-qubit density matrix with 96.17% accuracy, which is faster than many existing and our previous work.

1. Introduction

Quantum state tomography (QST) is a fundamental technique of quantum information processing, and its results determine the accuracy of the system afterwards such as in quantum computing and quantum communication [1]. The process of QST can be described as the process to recover a density matrix ρ of the quantum state given a series of measurements obtained from physical experiments. An n qubits quantum state can be fully represented by its density matrix ρ in a d dimensional Hilbert space [2], where $d = 2^n$ and ρ is a Hermitian matrix. $\mathcal{O}(d^2)$ measurements are usually required to fully recover the density matrix ρ due to the complexity of the system [3].

QST has its unique characteristics that distinguishe it from other optimization problems. Unlike classical systems, quantum mechanical measurements capture collapsed states in probabilities due to the Heisenberg uncertainty principle. Density matrices are supposed to hold its properties, and quantum measurements are costly. The new signal processing approach, compressed sensing (CS), gives us a better alternative to estimate a quantum state with fewer measurements. Proposed by Candes and Donaho [4,5], CS is a novel data acquisition and sampling theory that can process structured signals (sparse, low-rank, etc.) more efficiently [6,7]. It implies that only few essential data is required to exactly reconstruct the original signal by solving a CS-based optimization problem. Fortunately, people are usually interested in pure or nearly pure states practically in quantum systems [8], which means the corresponding density matrix ρ is low-rank and its singular values mostly vanish. In this case, the low-rank density matrix ρ can be reconstructed accurately with fewer measurements. Candes et al. [9] and Cai et al. [10] gave sufficient conditions to precisely reconstruct a k-sparse signal (the signal vector has k non-zero elements). Gross demonstrated that if the sensing matrix A satisfies the rank Restricted Isometry Property (RIP), $\mathcal{O}(d \cdot r \ln d)$ measurements are enough to estimate nearly pure quantum states[1]. However, these works usually provide theoretical bounds with orders, many constants and parameters in the expression, which might confuse researchers with quantum physics backgrounds. Meanwhile, few existing works studied how many measurements are really required in practice, and what the sampling rate is needed for CS based QST

* Corresponding author.
E-mail address: scong@ustc.edu.cn (S. Cong).

http://dx.doi.org/10.1016/j.sigpro.2017.04.007
0165-1684/

if we want to recover the density matrix with certain level of accuracy.

From the algorithmic perspective, the algorithms for convex optimization problems via CS have been studied widely. These problems mainly focus on minimizing the ℓ_1 norm or nuclear norm. For minimizing ℓ_1 norm, the algorithms include Interior Point (IP) [11], Gradient Projection (GP) [12] and Iterative Shrinkage-Thresholding Algorithm (ISTA). ISTA was proposed to solve Wavelet image convolution recovery problems [13,14] and later it was widely adopted to solve linear inverse problems with sparse constraints [15]. Daubechies et al. [16] proved the convergence of ISTA. For minimizing the nuclear norm, Yang and Yuan [17] introduced Augmented Lagrangian Method (ALM) and ADMM. Li and Cong [8] firstly applied ADMM to compressive quantum state reconstruction problems with both ℓ_1 norm and nuclear norm. The solution has a good accuracy but the complex computation caused by high dimensional matrix inversion limits its applications. For instance, it takes almost three hours with 92.43% accuracy to reconstruct a 7-qubit density matrix in MATLAB on the computer with 2 cores of 2.4 GHz Intel Xeon E5-2407 CPUs. Hence we hope to improve the algorithm further in terms of efficiency and accuracy.

In this paper, we make several contributions in CS based QST as follows. Firstly, we propose a novel algorithm for density matrix reconstruction given fewer number of measurements, with less computational time and better accuracy. The proposed algorithm combined an adaptive ADMM framework with a nuclear norm ISTA algorithm. Specifically, in each iteration, we use ISTA to obtain the solutions of sub-problems and then use the ADMM to update variables alternately. The IST–ADMM avoids computing the inverse of large-scale matrix, is capable of handling outlier errors, and approaches the true density matrix more rapidly comparing to several prevailing methods and our previous work. In particular, the computation has been reduced from $\mathcal{O}(d^6)$ for least square, $\mathcal{O}(md^4)$ for FP-ADMM in [18], to $\mathcal{O}(md^2)$ for IST–ADMM in this paper. In addition, we estimate the constants and parameters in the theoretical bounds of number of measurements for different qubits derived by previous works. These proper values match our simulation bounds and can provide a guidance to the sampling rates for researchers who intend to implement CS based QST in practice. Moreover, we study the time and accuracy required to reconstruct density matrices of qubits $n = 5, 6, 7, 8$ respectively. Compared with the least square (LS) and previous ADMM method for QST, the proposed algorithm demonstrates its superiority in time and much higher estimation accuracy, especially for high dimensional quantum systems.

This paper is organized as follows. In Section 2, we give the theoretical lower bounds of measurement rates for different qubits using some relative formulas. In Section 3, we propose the IST–ADMM algorithm in detail. Numerical experiments and results analysis are given in Section 4. Finally the conclusion is summarized in Section 5.

2. Lower bounds of measurement rates provided by CS

The state of an n-qubit quantum system can be described by a $d \times d$ density matrix ρ. The task of quantum state tomography is to reconstruct the density matrix ρ given certain measurements. The CS theory provides lower bounds of the number of measurements required to estimate the quantum state. Let $W_1, \cdots, W_{d^2}$ be an orthogonal basis for $\mathbb{C}^{d\times d}$, with respect to the inner product $(W_i, \rho) = tr(W_i^H \rho)$, W_i^H is the Hermitian transpose of W_i. We choose m bases out of d^2 elements, $\omega_1, \cdots, \omega_m$ at random from $\{W_1, \cdots, W_{d^2}\}$. Let the expectation of measurements $b_i \in \mathbb{R}^m$, and measuring operator $\mathcal{A} : \mathbb{C}^{d\times d\to m}$, then $b_i = (\mathcal{A}(\rho))_i + e_i = c \cdot tr(\omega_i^H \rho) + e_i, i = 1, \cdots, m$ or

$$b = A\mathrm{vec}(\rho) + e \tag{1}$$

where vec(·) represents the transformation from a matrix to a vector by stacking the matrix columns; $A \in \mathbb{C}^{m\times d^2}$ is the normalized measurement operator in a matrix form whose ith row consists of the concatenation of ω_i's rows; $e \in \mathbb{R}^m$ represents the noise caused by the system or measuring process; c is a normalized constant. If we set $E(\mathcal{A}^H\mathcal{A}) = \mathcal{I}$ where E represents the expectation overall $\mathcal{A}$, c would be $c = d/\sqrt{m}$ [8], m is the number of measurements.

Since the degree of the freedom of ρ is $d \times d$, usually people need $\mathcal{O}(d^2)$ measurements to give a unique solution by solving a system of linear equations. Yet if the quantum system is known as pure or nearly pure, the density matrix ρ has a low rank [1]. The density matrix reconstruction can be converted to a convex optimization problem by minimizing its nuclear norm [8]:

$$\begin{aligned} &\text{minimize } \|\rho\|_* \\ &\text{s.t. } \|A\mathrm{vec}(\rho) - b\|_2^2 \le \varepsilon, \rho^H = \rho, \rho \succeq 0 \end{aligned} \tag{2}$$

where $\rho \in \mathbb{C}^{d\times d}$; $\|\cdot\|_*$ represents the nuclear norm, which is equal to the sum of singular values, $m \ll d^2$; $\|\cdot\|_2$ represents the ℓ_2 norm, ρ^H is the Hermitian transpose of ρ; $\rho \succeq 0$ means ρ is a positive semi-definite matrix.

To make sure that the fewer measurements b contain the essential information of ρ, A is supposed to satisfy the rank RIP [19]. The lower bound of measurements has been derived according to RIP [20]. The lower bound implies that how many measurements do one need to exactly reconstruct an unknown low-rank matrix ρ given A's property. In general, there are several types of matrices that can be used as a sensing matrix A [21], such as Gaussian random matrix and Bernouli random matrix etc. Here we choose Pauli matrices in this paper as the bases of A, because they are widely used and easy to be implemented in practical quantum measurement.

The Pauli bases, $\omega_i = \otimes_1^n \sigma_k$, $k \in 1, \cdots, 4$, for an n-qubit quantum system are the Kronecker product of a series of complex and unitary elemental 2×2 Pauli matrices σ_i chosen from the four possibilities randomly:

$$\sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_4 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Thus there are $(2^n)^2 = d^2$ Pauli bases totally. We choose m Pauli bases randomly and record them as $\omega_{a_1}, \cdots, \omega_{a_m}$, where $a_1, \cdots, a_m \in [1, d^2]$, then the sensing matrix A is:

$$A = \begin{pmatrix} \mathrm{vec}(\omega(a_1))^T \\ \vdots \\ \mathrm{vec}(\omega(a_m))^T \end{pmatrix}. \tag{3}$$

In the quantum state tomography, we define the measurement rate η as

$$\eta = m/d^2. \tag{4}$$

The lower bound of η has been studied when Pauli bases are used as the sensing matrix and $\mathrm{rank}(\rho) = r$. If η satisfies [20]

$$\eta \ge C(1+\beta) r \ln d/d, \tag{5}$$

the solution ρ^* to the optimization problem (2) is unique and equals ρ with the overwhelming probability

$$P_s \ge 1 - e^{-\beta}, \tag{6}$$

where $\beta > 0$ is a parameter that balances the lower bound of η and the probability P_s. Some conclusions can be drawn from the

formulas (5) and (6). For the same number of qubits, when β decreases, the lower bound of η decreases and the correct probability P_s also decreases. The lower bound is not constant but can be adjusted by β according to the required correct probability P_s. For example, considering a pure state with $n = 5$ and $rank(\rho) = 1$, when β varies from 2.3 to 4.6, the lower bound varies from 0.089 to 0.151 and the correct probability P_s varies from 90% to 99%. It means that the measurement rate η decreases at the cost of the reduce of the correct probability P_s. The above theoretical lower bounds of η can provide instructions for us with regard to setting the right measurement rates. For pure state of quantum system with $n = 5, 6, 7, 8$, we set $r = 1$, $C = 0.25$ and $\beta = 4.60$, so $P_s \geq 99\%$, the lower bounds of η are 0.151, 0.091, 0.053 and 0.030, respectively. With suitable measurement rates, the reconstruction accuracy can be determined by the performance of the reconstruction algorithm.

3. An improved ADMM with ISTA for quantum state reconstruction

3.1. ADMM and ISTA

To ensure the completeness of the paper and the consistency of pre-defined symbols, we introduce ADMM and ISTA briefly in advance. ADMM is an optimization method to solve problems with two objective functions [22]. Consider an optimization problem:

$$\begin{aligned} &\text{minimize } f(x) + g(z) \\ &\text{s.t. } Ax + Bz = c \end{aligned} \tag{7}$$

for some variable $x, z \in \mathbb{R}^n$, where $f, g : \mathbb{R}^n \to \mathbb{R}$ are two convex functions. The variables appear separately in the objectives and are coupled only in the constraint. We choose the augmented Lagrangian of (7): $L(x, z, y) = f(x) + g(z) + y^T(Ax + Bz - c) + \lambda/2\|Ax + Bz - c\|_2^2$, where $y \in \mathbb{R}^m$ is the Lagrangian multiplier, $\lambda > 0$ is the penalty parameter. Alternatively we combine the linear and quadratic terms in the augmented Lagrangian with $u = y/\lambda$, then it becomes:

$$L(x, z, y) = f(x) + g(z) + \lambda/2\|Ax + Bz - c + u\|_2^2 \tag{8}$$

In such a way, the study of the optimization problem (7) is converted into the minimization of augmented Lagrangian (8). The basic idea of ADMM framework is to utilize the separated structure in (8) and to update the two variables alternatively [23]. After just one sweep of alternating minimization with respect to x and z, the multiplier λ is updated immediately. The ADMM alterations can be written as follows:

$$\begin{cases} x^{k+1} = \arg\min\limits_x \left\{ f(x) + \lambda/2 \parallel Ax + Bz^k - c + u^k \|_2^2 \right\} & \text{(9a)} \\ z^{k+1} = \arg\min\limits_z \left\{ g(z) + \lambda/2 \parallel Ax^{k+1} + Bz - c + u^k \|_2^2 \right\} & \text{(9b)} \\ u^{k+1} = u^k + Ax^{k+1} + Bz^{k+1} - c & \text{(9c)} \end{cases}$$

in which ADMM splits the whole optimization problem (7) into two sub-problems (9a) and (9b) that are easier to be solved. Then ADMM coordinates the solvers of the sub-problems to get the global solution.

We should note that ADMM only provides a frame for the iteration. How to solve the sub-problems depends on the specific forms of $f(x)$ and $g(z)$. When ADMM is used in the quantum state estimation, the objectives are the ℓ_1 norm and the nuclear norm, respectively. (9a) and (9b) are both unconstrained optimization problems with a smooth term and a non-smooth one. We will introduce an adaptive extended Iterative Shrinkage-Thresholding Algorithm (ISTA) to solve these two kinds of sub-problems (9a) and (9b).

The ISTA can be viewed as an extension of the classical gradient descent method. This class of algorithm works on solving the unconstrained optimization problems with a smooth objective and a non-smooth one. The ISTA is attractive due to its simplicity and thus it is adequate for solving large-scale problems arising from quantum state estimation. Consider the problem of minimizing function $f(x)$ which is assumed to be smooth and convex with smooth Lipschitz gradient. That is, there exists a constant $L(f)$ for $\|\nabla f(x) - \nabla f(y)\| \leq L(f)\|x - y\|, \forall x, y \in \mathbb{R}^n$. The solution of minimizing $f(x)$ can be given by the gradient descent method as:

$$x^k = x^{k-1} - t^k \nabla f(x^{k-1}), x^0 \in \mathbb{R}^n, \tag{10}$$

where $t^k > 0$ is a scalar as step size in the k-th iteration. The gradient descent method makes x^k move in the negative gradient direction of $f(x^k)$, so $f(x^{k+1}) \leq f(x^k)$ can be always got as long as t^k is suitable. Further, we replace $f(x)$ with its second approximation by Taylor's expansion [24]. One can easily verify that the iteration x^k specified by (10) is also the solution to the following problem:

$$x^k = \arg\min_x \left\{ f(x^{k-1}) + \langle x - x^{k-1}, \nabla f(x^{k-1}) \rangle + 1/2t^k \|x - x^{k-1}\|_2^2 \right\}. \tag{11}$$

By neglecting constant terms, (11) can be written as

$$x^k = \arg\min_x \left\{ 1/2t^k \|x - (x^{k-1} - t^k \nabla f(x^{k-1}))\|_2^2 \right\} \tag{12}$$

From (11) and (12), each iteration in the gradient descent method is to minimize a second approximation of $f(x)$. By extending this gradient updating mechanism to the problem whose objectives contain a non-smooth function, i.e.

$$\min \{F(x) \equiv f(x) + g(x) : x \in \mathbb{R}^n\} \tag{13}$$

where $g(x) : \mathbb{R}^n \to \mathbb{R}$ is a continuous convex function which is non-smooth, and $f(x) : \mathbb{R}^n \to \mathbb{R}$ is a smooth convex function with gradient which is Lipschitz continuous, we obtain

$$x^k = \arg\min_x \left\{ 1/2t^k \|x - (x^{k-1} - t^k \nabla f(x^{k-1}))\|_2^2 + g(x) \right\} \tag{14}$$

If $g(x)$ is separable, Proximal Gradient [25] can be used to solve (14). Then the general step of ISTA is of the form

$$x^k = prox_{t^k}(g)(x^{k-1} - t^k \nabla f(x^{k-1})) \tag{15}$$

where the *prox* operation is defined by

$$prox_t(g)(x) := \arg\min_v \left\{ g(v) + 1/2t \|v - x\|_2^2 \right\} \tag{16}$$

Therefore, the simplicity of ISTA depends on the ability to compute the *prox*. If $g(x) = \gamma \|X\|_1$, the problem is reduced to solve a one dimensional minimization problem about x_i. In this case, the specific step of ISTA is:

$$x^k = \mathcal{S}_{\gamma t^k}(x^{k-1} - t^k \nabla f(x^{k-1})) \tag{17}$$

where $\mathcal{S}_{\gamma t^k} : \mathbb{R}^n \to \mathbb{R}^n$ is the shrinkage operator defined by $\mathcal{S}_\tau(x) = \text{sgn}[x] \max(|x| - \tau \mathbf{1}, \mathbf{0})$, where $\mathbf{1}$ is a vector with all elements 1. Let $g(x) = \gamma \|X\|_*$, where $\|X\|_*$ represents the nuclear norm which equals to the sum of singular values. If X is a Hermitian matrix with a low rank, its singular values are real and sparse. By doing the similar shrinkage operation to the ℓ_1 norm we can get the specific step of ISTA as

$$x^k = \mathcal{D}_{\gamma t^k}(x^{k-1} - t^k \nabla f(x^{k-1})) \tag{18}$$

where $\mathcal{D}_\tau$ is the singular value shrinkage operator defined as $\mathcal{D}_\tau(X) = U\mathcal{S}_\tau(X)V^T$, USV^T is the singular value decomposition of X.

Each iteration of ISTA is comprised of a gradient step of the smooth term $f(x)$ followed by a shrinkage operation of the non-smooth term $g(x)$. The step size t^k is important to the convergence and the speed of the algorithms. Beck and Teboulle [26] proved that: Let $\{x^k\}$ be the sequential solution for the problem $\min \{F(x) \equiv f(x) + g(x) : x \in \mathbb{R}^n\}$ generated by ISTA with a constant step size $t^k = 1/L(f)$. Then for any $k > 1$, $F(x^k) - F(x^*) \leq L(f) \|x^0 - x^*\|^2/2k$, $\forall x^* \in X^*$. Specially, for $f(x) = \|Ax - b\|_2^2$ we choose $L(f) = 2\lambda_{\max}(A^H A)$, where $\lambda_{\max}$ is the greatest eigenvalue of $A^H A$. In this case, the function values $F(x^k)$ for ISTA converge sublinearly, i.e. with a rate of $\mathcal{O}(1/k)$.

3.2. IST–ADMM proposed for quantum state reconstruction

During the measurement process, sometimes there are abnormal circumstances perturbations might occur leading to noises in ρ. If the perturbation is Gaussian random, it can be handled by least square constraints and regular optimization process. However, the outlier noises are difficult to process. Here, the corresponding perturbation can be formulated as a sparse matrix S which is referenced as outlier noises. Taking S into consideration optimization problem (2) becomes

$$\begin{aligned} &\text{minimize } \|\rho\|_* + I_C(\rho) + \gamma\|S\|_1 \\ &\text{s.t. } \|A\text{vec}(\rho + S) - b\|_2^2 \le \varepsilon \end{aligned} \tag{19}$$

where $\gamma > 0$ is a compromise factor, $\|\cdot\|_1$ is the ℓ_1 norm, $\|X\|_1 = \sum_{i=1}^{m}\sum_{j=1}^{n}|x_{ij}|$, $I_C(\rho)$ is the indictor function on a convex set C, $I_C(\rho) = \begin{cases}\infty & \text{if } \rho^H = \rho,\ \rho \succeq 0 \\ 0 & \text{otherwise}\end{cases}$. The function of $I_C(\rho)$ projects ρ into a positive semi-definite Hermitian matrix. The augmented Lagrangian for (19) is

$$\begin{aligned} &L(\rho, S, y, \lambda) \\ &= \|\rho\|_* + I_C(\rho) + \gamma\|S\|_1 + \langle y, A\text{vec}(\rho + S) - b\rangle \\ &\quad + \lambda/2\|A\text{vec}(\rho + S) - b\|_2^2 \\ &= \|\rho\|_* + I_C(\rho) + \gamma\|S\|_1 + \lambda/2\|A\text{vec}(\rho + S) - b + u\|_2^2 \end{aligned} \tag{20}$$

where $y \in \mathbb{R}^m$ is the Lagrangian multiplier, $\lambda > 0$ is the penalty parameter, $u = y/\lambda$. By substituting (20) into (19) we can get the ADMM iterations for quantum state reconstruction:

$$\begin{cases} \rho^{k+1} = \arg\min\limits_{\rho} \{\|\rho\|_* + I_C(\rho) \\ \qquad + \lambda/2\|A\text{vec}\left(\rho + S^k\right) - b + u^k\|_2^2\} & \text{(21a)} \\ S^{k+1} = \arg\min\limits_{S} \{\gamma\|S\|_1 \\ \qquad + \lambda/2\|A\text{vec}\left(\rho^{k+1} + S\right) - b + u^k\|_2^2\} & \text{(21b)} \\ u^{k+1} = u^k + A\text{vec}\left(\rho^{k+1} + S^{k+1}\right) - b & \text{(21c)} \end{cases}$$

For (21a) and (21b), $\|\cdot\|_*$ and $\|\cdot\|_1$ are both non-smooth functions while $\|\cdot\|_2^2$ is a smooth function with Lipschitz continuous gradient. So we introduce ISTA to get an improved ADMM. For (21a), we transform $\|\rho\|_* + I_C(\rho) + \lambda/2\|A\text{vec}(\rho + S^k) - b + u^k\|_2^2$ by multiplying $2/\lambda$, then let $f(x) = \|A\text{vec}(\rho + S^k) - b + u^k\|_2^2$ and $g(x) = 2/\lambda(\|\rho\|_* + I_C(\rho))$. The solution to (21a) will not change after the transformation. According to the ISTA steps, firstly we carry out a gradient step of the smooth term $f(x)$ and denote the result as c_1^{k+1}

$$c_1^{k+1} = \rho^k - 2t^k\text{mat}(A^H(A\text{vec}(\rho^k + S^k) - b + u^k)) \tag{22}$$

To satisfy $\rho^H = \rho$, c_1^{k+1} is projected to a Hermitian matrix

$$c_{k+1} = 1/2(c_1^{k+1} + c_1^{k+1^H}).$$

To satisfy $\rho \succeq 0$, we can change the definition of the shrinkage operator to

$$\mathcal{S}'_{\tau}(x) = \begin{cases}\max(|x| - \tau\mathbf{1}, \mathbf{0}), & \text{sgn}[x] > 0 \\ \delta\max(|x| - \tau\mathbf{1}, \mathbf{0}), & \text{sgn}[x] < 0\end{cases}.$$

Here $0 < \delta < 1$ is a shrinkage parameter to reduce the negative singular value components (In practice we set $\delta = 0.9$ in simulations), $\mathcal{D}'_{\tau}(X) = U\mathcal{S}'_{\tau}(X)V^T$.

Then by substituting c^{k+1} to (18) we get the solution to the sub-problem (21a):

$$\rho^{k+1} = \mathcal{D}'_{2t^k/\lambda}(c^{k+1}) \tag{23}$$

For (21b), we do the similar transformation as (21a) and let $f(x) = \|A\text{vec}(\rho^{k+1} + S) - b + u^k\|_2^2$ and $g(x) = 2\gamma/\lambda\|S\|_1$. Combining (17) we can get the solution to the sub-problem (21b):

$$S^{k+1} = \mathcal{S}'_{2\gamma t^k/\lambda}(S^k - 2t^k\text{mat}(A^H(A\text{vec}(\rho^{k+1} + S^k) - b + u^k)) \tag{24}$$

Finally, we get

$$\begin{cases} \rho^{k+1} = \mathcal{D}'_{2t^k/\lambda}(c^{k+1}) & \text{(25a)} \\ S^{k+1} = \mathcal{S}'_{2\gamma t^k/\lambda}(S^k - 2t^k\text{mat}(A^H(A\text{vec}(\rho^{k+1} \\ \qquad + S^k) - b + u^k))) & \text{(25b)} \\ u^{k+1} = u^k + A\text{vec}(\rho^{k+1} + S^{k+1}) - b & \text{(25c)} \end{cases}$$

The iterations (25) is called IST–ADMM for quantum state reconstruction. There are three parameters: the step size for gradient descent method t^k; the compromise factor γ; the penalty parameter λ of ADMM.

In this paper, three parameters are set as: $t^k = 1/2\lambda_{\max}(A^HA)$ [26], while the sensing matrix A is made of Pauli matrices, $t^k = 1/2$; $\gamma = 1/\sqrt{d}$ [27]; the penalty parameter $\lambda = 1/2\|b\|_2$ [18]. In the updating processes, if $\|A\text{vec}(\rho^k + S^k) - b\|_2/\|b\|_2 < \varepsilon_1$ or the iterations $k > k_{\max}$, IST–ADMM stops updating where $\varepsilon_1 = 10^{-7}$. The reconstruction performances are evaluated by the normalized reconstruction error

$$\text{error} = \|\hat{\rho} - \rho\|_F^2/\|\rho\|_F^2 \tag{26}$$

where $\|\cdot\|_F$ represents Frobenius norm, ρ and $\hat{\rho}$ denote the true state, whose rank is one, and the estimate state respectively. If the error is larger than 1, we record it as 1. We call $(1 - \text{error})$ the reconstruction accuracy.

Remarks:

1. We must distinguish the reconstruction accuracy $(1 - \text{error})$ and the correct probability P_s in (6). P_s is the probability of the incident that the solution ρ^* to the optimization problem (2) is unique and equals ρ. P_s is affected by the lower bound of the measurement rate η. While $(1 - \text{error})$ represents the similarity between the estimate state $\hat{\rho}$ calculated by the IST–ADMM and the true state ρ. $(1 - \text{error})$ is affected by the parameter settings of our algorithm such as iterations etc. In this paper, if $(1 - \text{error})$ is more than 95%, we will claim the incident that " the solution ρ^* to the optimization problem (2) is unique and equals to ρ " happens. $(1 - \text{error})$ is an indication we defined to reflect the performance of the proposed algorithm. It is meaningful only in the simulations because we do not know the true state ρ in practice.
2. Compared with LS and Li's ADMM [8], IST–ADMM proposed in this paper is faster and more accurate. Li's ADMM needs to compute the pseudo-inverse $(A^HA)^{-1}A^H$ of $A \in m \times d^2$. Computing $(A^HA)^{-1}A^H$ involves the inverse of a $d^2 \times d^2$ matrix and the multiplication between a $d^2 \times d^2$ matrix and a $d^2 \times m$ one, whose computational complexity is $\mathcal{O}(d^6)$. Likewise, LS computational complexity is also $\mathcal{O}(d^6)$. However, the iterations (25) shows that the maximal computation of IST–ADMM is the multiplication between an $m \times d^2$ matrix and a $d^2 \times 1$ vector, whose computational complexity is $\mathcal{O}(md^2)$. In addition, IST–ADMM provides greater accuracy than ADMM. In IST–ADMM the solutions of the sub-problems are obtained on the whole with one step by ISTA, while in ADMM they are obtained in several segments. Hence, IST–ADMM can reduce the computing time and improve the accuracy.
3. Compared with FP-ADMM [18], the proposed IST–ADMM has a big improvement on reducing the computational complexity and time. With FP-ADMM, the computing sequence of ρ is $\rho^{k+1} = D_{2t^k/\lambda}(\text{mat}((I - 2t^kA^HA)\text{vec}(\rho^k) + 2t^kA^H(b - A\text{vec}(S^k) - u^k)))$, where $(I - 2t^kA^HA)$ costs $\mathcal{O}(md^4)$ computational complexity. Besides, massive memory space is

consumed because two $d^2 \times d^2$ matrices, I and $2t^k A^H A$, need to be saved at the same time. When $d = 2^8$, it needs 128G memory that general computers cannot cope with. However, with IST–ADMM, the computing sequence of ρ is $\rho^k - 2t^k \mathrm{mat}(A^H(A\mathrm{vec}(\rho^k + S^k) - b + u^k))$. $(A\mathrm{vec}(\rho^k + S^k) - b + u^k)$ enclosed in parentheses has higher precedence and it is computed first. The result is an $m \times 1$ vector and then the vector is multiplied by A^H. So the computational complexity is $\mathcal{O}(md^2)$. This improvement can greatly reduce the computational time. With higher number of qubits, the superiority of the proposed method is more significant. For example, using MATLAB on a computer with 16G memory and 2 cores of 2.4GHz Intel Xeon E5 – 2407 CPUs, $n = 5, 6, 7$, $\eta = 0.4$ and the number of iterations is 30, the computational time of FP-ADMM is 1.80, 30.60, 67.39 seconds, respectively, and the computational time of IST–ADMM is 0.15, 1.71, 2.74 seconds, respectively. Similarly, the computing sequence of (25b) also reduces the computations and the memory space compared with FP-ADMM. The proposed method extends the number of qubits to 8. It takes 3.13 minutes to reconstruct an 8-qubit density matrix with 96.17% accuracy. Though the convergence of IST–ADMM is not proved, from the simulations of high dimensional quantum state tomography we can see that results given by the proposed algorithms are satisfactory.

4. Numerical experiments and results analysis

To verify the superiority of the IST–ADMM, we carry out three experiments. In experiment 1, we use the IST–ADMM to get the lower bounds of the measurement rates with different number of qubits. Then we compare the experimental bounds with the theoretical values in (5) to verify the efficiency of the proposed IST–ADMM . In experiment 2, we study the reconstruction errors change with iterations after setting the suitable η according to experiment 1. This experiment aims at finding the fewer iterations to reduce the computing time. In experiment 3, with the outlier noise we compare IST–ADMM with LS and ADMM in aspect of accuracy and computation speed. When $n = 8$, the complete Pauli matrix is a $d^2 \times d^2 = 65536 \times 65536$ complex matrix. If $\eta = 0.2$, the sensing matrix A is an $m \times d^2 = 13107 \times 65536$ complex matrix, where $m = \eta \times d^2$. The complete Pauli matrix and the sensing matrix need 76.8G. Besides, some temporary variables also need to be stored. Hence, we use a supercomputer with 256G memory and 64 cores of 2.66 GHz Intel Xeon E7 – 8837 CPUs. In three experiments, we use the Pauli measurements.

4.1. Comparisons of lower bounds of measurement rates with different qubits

In this experiment, we use the IST–ADMM in (25) to estimate quantum states without outlier noise. We record the lower bounds of η of different qubits when the reconstruction accuracy achieves above 99%. Then we compare these bounds with the theoretical values obtained in (5). We set qubits $n = 5, 6, 7, 8$, respectively; the measurement rate increases from $\eta = 0.01$ to $\eta = 0.2$ and the incremental step is $\Delta\eta = 0.01$. Under each measurement rate, the number of iterations is 100. Fig. 1 depicts the normalized estimation errors with different η, in which purple dot-dash cross, red dotted triangle, blue dashed asterisk and black solid circle lines represent $n = 5, 6, 7, 8$, respectively. From Fig. 1 one can see that to reconstruct the density matrix exactly, that is $(1 - \text{error}) \approx 100\%$, $n = 5, 6, 7, 8$ need the measurement rate η at least to be 0.15, 0.09, 0.05 and 0.03, respectively.

The comparison of the experimental lower bounds of measurement rate η with the theoretical values obtained in (5) is shown in Table 1, from which one can see that:

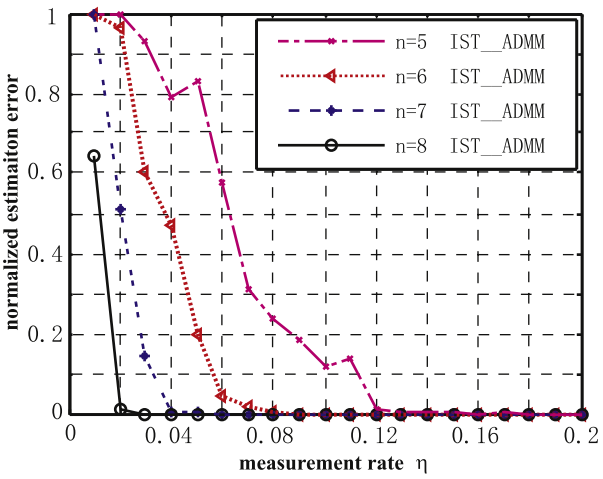

Fig. 1. Normalized estimation errors with different measurement rates.

Table 1
Comparison of theoretical and experimental lower bounds of η.

Pauli measurement	Lower bounds of η			
	$n = 5$	$n = 6$	$n = 7$	$n = 8$
Theoretical values	0.151	0.091	0.053	0.030
Experimental values	0.15	0.09	0.05	0.03

1. In the case of $n = 5, 6, 7, 8$, the theoretical lower bounds are 0.151, 0.091, 0.053 and 0.030 while the experimental values are 0.15, 0.09, 0.05 and 0.03, they are very close.
2. With the number of qubits increasing, the lower bounds of measurement rate η decrease, which verifies that the CS theory can sharply reduce the number of measurements.
3. The experimental values basically agree well with theoretical values, which indicates that IST–ADMM is efficient. Meanwhile, the results provide a guidance of η setting for us.

4.2. Effects of iterations on reconstruction errors with different qubits

In experiment 2, without outlier noise, given $n = 5, 6, 7, 8$, we set the measurement rate η as 0.15, 0.09, 0.05 and 0.03, respectively. The iterations increase from 10 to 100 and the incremental step is 10. Fig. 2 depicts the normalized estimation errors with

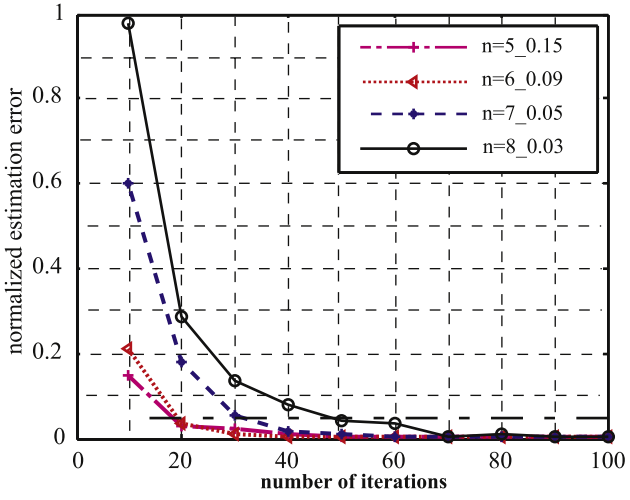

Fig. 2. Normalized estimation errors with different number of iterations.

Table 2
Reconstruction accuracy and time for $n = 5, 6, 7, 8$.

Qubit	$n = 5$	$n = 6$	$n = 7$	$n = 8$
Measurement rate η	0.15	0.09	0.05	0.03
Number of iterations	20	20	30	50
Reconstruction accuracy (%)	97.31	96.30	94.35	96.17
Reconstruction time (s)	0.06	0.51	1.25	188.22

different iterations, in which purple dot-dash cross, red dotted triangle, blue dashed asterisk and black solid circle lines represent the result of $n = 5, 6, 7, 8$, respectively. The black dash-dot-dot line means that the normalized estimation error is 0.05. From Fig. 2 one can see that when the number of qubits increases, the iterations required increase to obtain the same error. For $n = 5, 6, 7, 8$, the required iterations are 20, 20, 30 and 50 , respectively with about 95% accuracy.

Table 2 records the reconstruction accuracy and the time for $n = 5, 6, 7, 8$ in the experiment 2, where the measurement rates are 0.15, 0.09, 0.05 and 0.03, respectively, and the iterations are 20, 20, 30, and 50, respectively. Here, the reconstruction time means how long IST–ADMM needs to compute the solution within the set iterations given A and the expectation of measurements b. From Table 2 one can see that the reconstruction accuracy of IST–ADMM is 97.31%, 96.30%, 94.35% and 96.17%, respectively; and the consuming time is 0.06, 0.51, 1.25 and 188.22 seconds, respectively. As mentioned in Section 3.2, IST–ADMM avoids computing the inverse and multiplication of large scale matrices leading to faster computation.

4.3. Comparison of IST–ADMM, ADMM and LS with outlier noise

To validate the superiority of IST–ADMM algorithm in aspect of reconstruction accuracy and speed, we compare IST–ADMM, ADMM and LS in terms of reconstruction of ρ with the same measurement matrix A and the same expectation of measurements b. Meanwhile, to validate the robustness of IST–ADMM, the outlier noise is considered in the density matrix. The outlier noise S is set with $0.01d^2$ nonzero values located uniform randomly and the magnitudes satisfying Gaussian distribution $N(0, 0.1\|\rho\|_2)$. With $n = 5, 6, 7$, the measurement rate increases from $\eta = 0.05$ to $\eta = 0.5$ and the incremental step is $\Delta\eta = 0.05$. Under each η, the number of iterations is set as 30. Fig. 3 depicts the normalized estimation errors with different η, in which the solid, dashed and dot-dashed lines represent IST–ADMM, ADMM and LS, respectively. The purple cross, red triangle and blue asterisk represent $n = 5, 6, 7$, respectively. From Fig. 3 one can see that:

1. Even with noise, the IST–ADMM can still reconstruct the density matrix exactly when η is more than 0.15, which evaluates the robustness of IST–ADMM.
2. For the same number of qubits and η, IST–ADMM can significantly reduce the estimation error compared with ADMM and LS. For example, when $n = 5, 6, 7$ and $\eta = 0.15$, the reconstruction accuracy by IST–ADMM is 98.71%, 99.39%, 99.30%, respectively while by ADMM it is 34.10%, 47.76%, 59.56%, respectively and by LS it is 14.07%, 13.30%, 12.04%, respectively. Results show that IST–ADMM significantly improves the reconstruction accuracy of the quantum state estimation via CS.

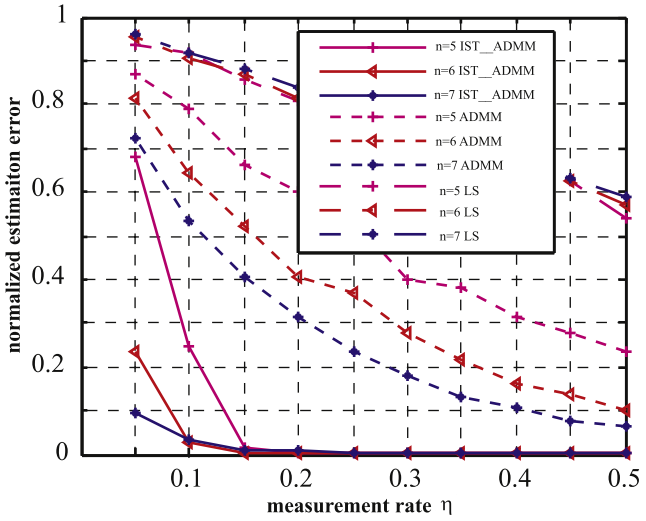

Fig. 3. Comparison of IST–ADMM, ADMM and LS with outlier noise.

Table 3 records the reconstruction time of IST–ADMM, ADMM and LS with $n = 5, 6, 7$ and $\eta = 0.1, 0.25, 0.4$. The reconstruction time means how long the algorithms need to run for 30 iterations. From Table 3 one can see that:

1. Under the same number of qubits and measurement rate η, the reconstruction time of IST–ADMM reduces significantly compared to ADMM and LS.
2. While η varies, the time of IST–ADMM is much less than that of ADMM and LS. The higher number of qubits is, the more time reduces. For example, when $\eta = 0.25$ and $n = 5$, the reconstruction time of ADMM is $16.36/0.13 = 126$ times longer than ADMM. It is $387/1.00 = 387$ times when $n = 6$, $9050/1.63 = 5552$ times when $n = 7$. IST–ADMM is much faster than ADMM in the quantum state estimation. With higher number of qubits, the superiority of the proposed method is more significant.

From three experiments above, one might notice that:

1. In experiment 1, for each measurement rate η, the true density matrix ρ is randomly generated every time. With the randomness, we verify that although the m rows of the measurement matrix A are chosen randomly, IST–ADMM is efficient for any pure density matrix ρ. However, in order to reduce the time, for each density matrix ρ and each measurement matrix A, IST–ADMM performs the number of iterations set only once. The feasibility of this method was studied in [28].
2. The rank of true density matrix ρ is set as 1, because we focus on the reconstruction of pure state whose rank is 1. However, IST–ADMM is suitable for all the low-rank density matrix reconstruction. Following numerical experiments address this point.

4.4. The experiment in the case of $rank(\rho) = 2$ and randomness

Without outlier noise, we set qubit $n = 6$ and $rank(\rho) = 2$. The measurement rate increases from $\eta = 0.02$ to $\eta = 0.4$ and the incremental step is $\Delta\eta = 0.02$. For each measurement rate η, the

Table 3
Comparison of reconstruction time about IST–ADMM, ADMM and LS for 30 iterations (s).

n	5			6			7		
η	0.1	0.25	0.4	0.1	0.25	0.4	0.1	0.25	0.4
IST–ADMM	0.12	0.13	0.14	0.47	1.00	1.5	1.36	1.63	1.93
ADMM	15.34	16.36	19.11	3.17e2	3.87e2	4.54e2	7.67e3	9.05e3	9.81e3
LS	4.72	5.66	5.85	1.31e2	1.47e2	1.60e2	3.29e3	3.32e3	3.41e3

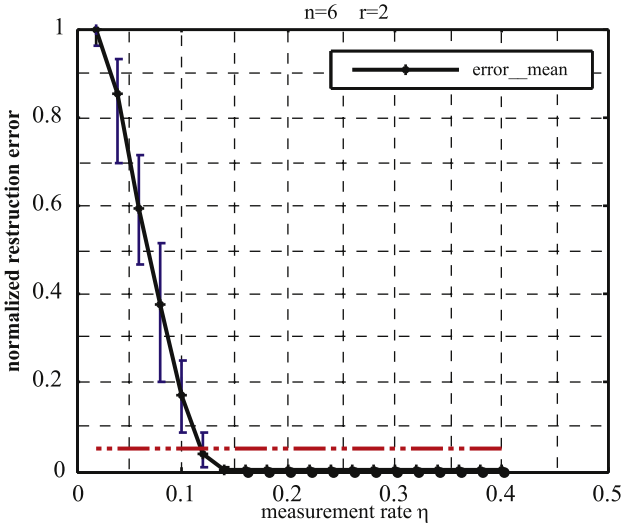

Fig. 4. The experiment in the case of $rank(\rho) = 2$ and randomness.

true density matrix ρ is randomly generated every time. For each reconstruction, the number of iterations is 100. Run the above experiments in each case for 20 times. In this way, for each measurement rate η, we obtain 20 different errors. Calculate the average, maximum and minimum of these 20 errors respectively. Fig. 4 depicts the average, maximum and minimum of the normalized estimation errors with different η, in which the black solid asterisk line represents the average, the upper and lower endpoints of the vertical blue segments represents the maximum and minimum of errors, respectively. The red dash-dot-dot line represents the normalized estimation error is 0.05, i.e. the reconstruction accuracy is 95%. From Fig. 4 one can see that:

1. When the measurement rate η is less than 0.14, the gaps of the average, maximum and minimum cases are obvious. But the three values are almost same when η is more than 0.14. This shows that we can ignore the influence of randomness if the density matrix is reconstructed accurately with η more than its lower bound. Because we only care about the case in which the density matrix is reconstruct accurately, the influence of randomness can be ignored. In our experiments, we set the measurement rate η according to its lower bound and did the experiment once for each reconstruction, which can make sure the expected accuracy.
2. When $rank(\rho) = 2$, i.e. $r = 2$ and the measurement rate $\eta = 0.14$, the average, maximum and minimum of the reconstruction accuracy are 99.57%, 99.99% and 99.21%, respectively. With $\eta \geq 0.16$, the accuracy is almost 100%. This shows that our IST–ADMM is also efficient for $\text{rank}(\rho) > 1$ and can reconstruct the density matrix accurately as long as η is suitable.

5. Conclusion

This paper proposed an improved ADMM combined with ISTA for quantum state tomography via compressive sensing. The algorithm utilized ISTA to solve the sub-problems and then used ADMM to update the density matrix and the sparse noises iteratively. For the simple of ISTA we can avoid computing the inverse and multiplication of large scaled matrices. Simulations showed that the proposed algorithm is much faster and achieves a better recovery accuracy than conventional ADMM. In the case of $n = 8$ and $\eta = 0.03$, IST–ADMM used 50 iterations and 188 seconds to reconstruct the density matrix with the accuracy of 96.17%. We can forecast that the measurement rates are at least 0.0167 and 0.0093 respectively when qubit $n = 9$ and $n = 10$.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under grant no. 61573330 and the Open Research Fund of State Key Laboratory of Space- Ground Integrated Information Technology under grant no. 2015_SGIIT_KFJJ_DH_04.

References

[1] D. Gross, Y.K. Liu, S.T. Flammia, S. Becker, J. Eisert, Quantum state tomography via compressed sensing, Phys. Rev. Lett. 105 (2010) 150401.
[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. 14 (28) (2012) 95022–95049.
[3] A.I. Lvovsky, Continuous-variable optical quantum state tomography, Rev. Mod. Phys. 81 (2005) 299–332.
[4] E.J. Candes, J. Romberg, T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, IEEE Trans. Inf. Theory 52 (2004) 489–509.
[5] D.L. Donoho, Compressed sensing, IEEE Trans. Inf. Theory 52 (2006) 1289–1306.
[6] E.J. Candes, T. Tao, Near-optimal signal recovery from random projections: universal encoding strategies, IEEE Trans. Inf. Theory 52 (2006) 5406–5425.
[7] S. Mallakpour, S. Soltanian, An automated signal reconstruction method based on analysis of compressive sensed signals in noisy environment, Signal Process. 104 (2014) 43–50.
[8] K. Li, S. Cong, A robust compressive quantum state tomography algorithm using ADMM, IFAC Proc. Vol. 47 (2014) 6878–6883.
[9] E.J. Candes, J.K. Romberg, T. Tao, Stable signal recovery from incomplete and inaccurate measurements, Commun. Pure Appl. Math. 59 (2006) 1207–1223.
[10] T.T. Cai, L. Wang, G. Xu, New bounds for restricted isometry constants, IEEE Trans. Inf. Theory 56 (2010) 4388–4394.
[11] K. Koh, S.J. Kim, S. Boyd, An interior-point method for large-scale ℓ_1-regularized logistic regression, J. Mach. Learn. Res. 8 (2007) 1519–1555.
[12] M.A.T. Figueiredo, R.D. Nowak, S.J. Wright, Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems, IEEE J. Selected Top. Signal Process. 1 (2007) 586–597.
[13] M.A.T. Figueiredo, R.D. Nowak, An EM algorithm for wavelet-based image restoration, IEEE Trans. Image Process. 12 (2003) 906–916.
[14] Y. Zhang, Z. Dong, P. Phillips, S. Wang, G. Ji, J. Yang, Exponential wavelet iterative shrinkage thresholding algorithm for compressed sensing magnetic resonance imaging, Inf. Sci. 322 (2015) 115–132.
[15] Y. Gu, N.A. Goodman, S. Hong, Y. Li, Robust adaptive beam forming based on interference covariance matrix sparse reconstruction, Signal Process. 96 (2014) 375–381.
[16] I. Daubechies, M. Defrise, C. De Mol, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint, Commun. Pure Appl. Math. 57 (2004) 1413–1457.
[17] J. Yang, X. Yuan, Linearized augmented lagrangian and alternating direction methods for nuclear norm minimization, Math. Comput. 82 (2011) 301–329.
[18] K. Zheng, K. Li, S. Cong, A reconstruction algorithm for compressive quantum tomography using various measurement sets, Sci. Rep. 6 (2016) 38497.
[19] Y. Tsaig, D.L. Donoho, Extensions of compressed sensing, Signal Process. 86 (2006) 549–571.
[20] D. Gross, Recovering low-rank matrices from few coefficients in any basis, IEEE Trans. Inf. Theory 57 (2009) 1548–1566.
[21] E.J. Candes, Y. Plan, Tight oracle inequalities for low-rank matrix recovery from a minimal number of noisy random measurements, IEEE Trans. Inf. Theory 57 (2011) 2342–2359.
[22] D. Gabay, B. Mercier, A dual algorithm for the solution of nonlinear variational problems via finite element approximation, Comput. Math. Appl. 2 (1976) 17–40.
[23] W. Shi, Q. Ling, K. Yuan, G. Wu, On the linear convergence of the ADMM in decentralized consensus optimization, IEEE Trans. Signal Process. 62 (2014) 1750–1761.
[24] D.P. Bertsekas, Nonlinear Programming, American Mathematical Society, 2003.
[25] P.L. Combettes, V.R. Wajs, Signal recovery by proximal forward-backward splitting, Multiscale Model. Simul. 4 (2005) 1168–1200.
[26] A. Beck, M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, SIAM J. Imag. Sci. 2 (2009) 183–202.
[27] E.J. Candes, X. Li, Y. Ma, J. Wright, Robust principal component analysis? J. ACM (JACM) 58 (2011) 11.
[28] J. Zhang, K. Zheng, K. Li, S. Cong, Further improved ADMM and its application in quantum state estimation, in: Proceedings of 17th Chinese Conference on System Simulation Technology, 2016, pp. 82–87. (in Chinese)