

# A New Inexact Proximal Linear Algorithm with Adaptive Stopping Criteria for Robust Phase Retrieval

Zhong Zheng, Shiqian Ma, and Lingzhou Xue \*

## Abstract

This paper considers the robust phase retrieval problem, which can be cast as a nonsmooth and nonconvex optimization problem. We propose a new inexact proximal linear algorithm with the subproblem being solved inexactly. Our contributions are two adaptive stopping criteria for the subproblem. The convergence behavior of the proposed methods is analyzed. Through experiments on both synthetic and real datasets, we demonstrate that our methods are much more efficient than existing methods, such as the original proximal linear algorithm and the subgradient method.

**Keywords**— Robust Phase Retrieval, Nonconvex and Nonsmooth Optimization, Proximal Linear Algorithm, Complexity, Sharpness.

## 1 Introduction

Phase retrieval aims to recover a signal from intensity-based or magnitude-based measurements. It finds various applications in different fields, including X-ray crystallography [1], optics [2], array and high-power coherent diffractive imaging [3], astronomy [4] and microscopy [5]. Mathematically, phase retrieval tries to find the true signal vectors  $x_*$  or  $-x_*$  in  $\mathbb{R}^n$  from a set of magnitude measurements:

$$b_i = (a_i^\top x_*)^2, \text{ for } i = 1, 2, \dots, m, \quad (1)$$

where  $a_i \in \mathbb{R}^n$  and  $b_i \geq 0, i = 1, 2, \dots, m$ . Directly solving the equations leads to an NP-hard problem [6], and nonconvex algorithms based on different designs of objective functions have been well studied in the literature, including Wirtinger flow [7], truncated Wirtinger flow [8], truncated amplitude flow [9], reshaped Wirtinger flow [10], etc.

In this paper, we focus on the robust phase retrieval (RPR) problem [11], which considers the case where  $b_i$  contains noise due to measurement errors of equipment. That is,

$$b_i = \begin{cases} (a_i^\top x_*)^2, & i \in \mathbb{I}_1, \\ \xi_i, & i \in \mathbb{I}_2, \end{cases} \quad (2)$$

in which  $\mathbb{I}_1 \cup \mathbb{I}_2 = \{1, 2, \dots, m\}$ ,  $\mathbb{I}_1 \cap \mathbb{I}_2 = \emptyset$ , and  $\xi_i$  denotes a random noise. [11] proposed to formulate RPR as the following optimization problem:

$$\min_{x \in \mathbb{R}^n} F(x) := \frac{1}{m} \sum_{i=1}^m |(a_i^\top x)^2 - b_i|. \quad (3)$$

---

\*Z. Zheng and L. Xue are with the Department of Statistics at Penn State University, University Park, PA 16802. Research is partly supported by NSF grants DMS-1953189, CCF-2007823, and DMS-2210775. Email: {zvv5337, lzxue}@psu.edu. S. Ma is with the Department of Computational Applied Mathematics and Operations Research at Rice University, Houston, TX 77005. Research is supported in part by NSF grants DMS-2243650, CCF-2308597, CCF-2311275, and ECCS-2326591, and a startup fund from Rice University. Email: sqma@rice.edu. S. Ma and L. Xue are co-corresponding authors.

It is demonstrated in [11] that using (3) for RPR possesses better recoverability compared to the median truncated Wirtinger flow algorithm [12] based on the  $\ell_2$ -loss.

Solving (3) is challenging because it is a nonconvex and nonsmooth optimization problem. In [13], the authors proposed the subgradient method to solve it. This method requires geometrically decaying step size, and it is unclear how to schedule this kind of step size in practice. [11] proposed to use the proximal linear (PL) algorithm to solve (3). For ease of presentation, we rewrite (3) as

$$\min_{x \in \mathbb{R}^n} F(x) = h(c(x)) = \frac{1}{m} \left\| |Ax|^2 - b \right\|_1, \quad (4)$$

where  $Ax = [\langle a_1, x \rangle, \dots, \langle a_m, x \rangle]^\top$ ,  $b = [b_1, \dots, b_m]^\top$ ,  $h(z) := \frac{1}{m} \|z\|_1$ , and  $c(x) := |Ax|^2 - b$  is a smooth map in which  $|\cdot|^2$  is element-wise square. One typical iteration of the PL algorithm is

$$x^{k+1} \approx \arg \min_{x \in \mathbb{R}^n} F_t(x; x^k), \quad (5)$$

where  $t > 0$  is the step size,

$$F(z; y) := h(c(y) + \nabla c(y)(z - y)), \quad (6)$$

$$F_t(z; y) := F(z; y) + \frac{1}{2t} \|z - y\|_2^2, \quad (7)$$

$\nabla c$  denotes the Jacobian of  $c$ , and “ $\approx$ ” means that the subproblem is solved inexactly. The subproblem (5) is convex and can be solved by various methods such as the proximal operator graph splitting (POGS) algorithm used in [11]. The PL method has drawn lots of attention recently. It has been studied by [14–16] and applied to solving many important applications such as RPR [11], robust matrix recovery [17, 18], and sparse spectral clustering [19]. The subproblem (5) is usually solved inexactly for practical concerns. As pointed out in [11], the PL implemented in [11] is much slower than the median truncated Wirtinger flow method. We found that this is mainly due to their stopping criterion for solving the subproblem (5), which unnecessarily solves (5) to very high accuracy in the early stage of the algorithm. Moreover, we found that the POGS algorithm used in [11] is ineffective in solving the subproblem (5). In this paper, we propose adaptive stopping criteria for inexactly solving (5) with the fast iterative shrinkage-thresholding algorithm (FISTA) [20–22]. We found that our new inexact PL (IPL) with the adaptive stopping criteria greatly outperforms existing implementations of PL methods [11] for solving RPR (4).

**Our Contributions.** In this paper, we propose two new adaptive stopping criteria for inexactly solving (5). The first one ensures that (5) is solved to a relatively low accuracy:

$$\begin{aligned} \text{(LACC)} \quad & F_t(x^{k+1}; x^k) - \min_{x \in \mathbb{R}^n} F_t(x; x^k) \\ & \leq \rho_l (F_t(x^k; x^k) - F_t(x^{k+1}; x^k)), \quad \rho_l > 0, \end{aligned} \quad (8)$$

and the second one ensures that (5) is solved to a relatively high accuracy:

$$\begin{aligned} \text{(HACC)} \quad & F_t(x^{k+1}; x^k) - \min_{x \in \mathbb{R}^n} F_t(x; x^k) \\ & \leq \frac{\rho_h}{2t} \|x^{k+1} - x^k\|_2^2, \quad 0 < \rho_h < 1/4. \end{aligned} \quad (9)$$

Here,  $\rho_l$  and  $\rho_h$  are given constants. Similar to the proximal bundle method [23] for nonsmooth convex problems, (LACC) and (HACC) are designed to ensure the sufficient decrease of the objective function for the nonsmooth and nonconvex RPR problem. Note that both (LACC) and (HACC) are only used theoretically because  $\min_{x \in \mathbb{R}^n} F_t(x; x^k)$  is not available. Later we will propose more practical stopping criteria that can guarantee (LACC) and (HACC). The connections of our approach to existing work are listed below.

- (a) Our (LACC) condition coincides with the inexact stopping criterion proposed in [24–26] for the proximal gradient method. In these papers, the authors focus on a different optimization problem

$$\min_{x \in \mathbb{R}^n} f_0(x) := f_1(x) + f_2(x),$$

in which  $f_1$  is a smooth function, and  $f_2$  is a proper, convex, and lower semi-continuous function. One typical iteration of their algorithms can be written as

$$y^{k+1} \approx \min_{x \in \mathbb{R}^n} f_{0k}(x) = f_1(x^k) + (x - x^k)^\top \nabla f_1(x^k) + f_2(x) + \frac{1}{2}(x - x^k)^\top H_k(x - x^k), \quad (10a)$$

$$x^{k+1} = x^k + \lambda_k(y^{k+1} - x^k), \quad (10b)$$

where  $H_k \in \mathbb{R}^{n \times n}$  is a positive semi-definite matrix and  $\lambda_k \in [0, 1]$  is a step size. The stopping criterion for inexactly solving (10a) proposed in [24–26] is

$$f_{0k}(y^{k+1}) - \tilde{f}_{0k} \leq \eta \left( f_{0k}(x^k) - \tilde{f}_{0k} \right), \quad (11)$$

where  $\tilde{f}_{0k} = \min_{x \in \mathbb{R}^n} f_{0k}(x)$  and  $\eta \in (0, 1)$ . We note that this is the same as our (LACC). Therefore, our (LACC) is essentially an extension of (11) from the proximal gradient method to the proximal linear method.

- (b) To the best of our knowledge, our (HACC) criterion is new and serves as a good alternative to (LACC). From our numerical experiments, we found that (HACC) works comparably with (LACC), and we believe that it can be useful for other applications.
- (c) We analyze the overall complexity and the local convergence of our IPL algorithm for solving RPR under the sharpness condition. To the best of our knowledge, this is the first time such results have been obtained under the sharpness condition.
- (d) We propose to solve (5) inexactly using FISTA [20–22], which uses easily verifiable stopping conditions that can guarantee (LACC) and (HACC). Through extensive numerical experiments, we demonstrate that our IPL with the new stopping criteria significantly outperforms existing algorithms for solving RPR.

**Organization.** The rest of this paper is organized as follows. In Section 2, we propose the main framework of our inexact proximal linear algorithm with two new adaptive stopping criteria for the subproblem. We establish its iteration complexity for obtaining an  $\epsilon$ -stationary point and its local convergence under the sharpness condition. Connections with some existing methods are also discussed. In Section 3, we discuss how to adapt the FISTA to solve the subproblem inexactly. We also establish the overall complexity of FISTA – the total number of iterations of the FISTA – in order to obtain an  $\epsilon$ -optimal solution under the sharpness condition. In Section 4, we show the numerical results on both synthetic and real datasets to demonstrate the advantage of the proposed methods over some existing methods. The proofs for all the theorems and lemmas are given in Section 5. Finally, we include some concluding remarks in Section 6.

## 2 IPL and Its Convergence Analysis

In this section, we introduce our IPL algorithm for solving the RPR (4) with the inexact stopping criteria (LACC) and (HACC) for the subproblem (5) and analyze its convergence. We will discuss the FISTA for solving (5) that guarantees (LACC) and (HACC) in the next section.

We first follow [15] to introduce some notation. Let

$$S_t(y) := \arg \min_{x \in \mathbb{R}^n} F_t(x; y),$$

$$\varepsilon(x; y) := F_t(x; y) - F_t(S_t(y); y),$$

where  $F_t(z; y)$  is defined in (7). We will also use the notation

$$L = \frac{2}{m} \|A\|_2^2 = \frac{2}{m} \left\| \sum_{i=1}^m a_i a_i^\top \right\|_2.$$

A Meta Algorithm of our IPL is summarized in Algorithm 1. We again emphasize that Algorithm 1 cannot be implemented because  $\min_{x \in \mathbb{R}^n} F_t(x; x^k)$  is not available, and we will discuss practical versions of it in the next section.

---

**Algorithm 1** IPL – A Meta Algorithm

---

**Input:** Initial point  $x^0$ , step size  $t = 1/L$ , parameters  $\rho_l > 0$  and  $\rho_h \in (0, 1/4)$

**for**  $k = 0, 1, \dots$ , **do**

Obtain  $x^{k+1}$  by inexactly solving (5) with one of the following stopping criteria:

**Option 1:** (LACC), i.e., (8)

**Option 2:** (HACC), i.e., (9)

**end for**

---

## 2.1 Convergence under General Settings

In this subsection, we analyze the convergence rate of IPL (Algorithm 1) for obtaining an  $\epsilon$ -stationary point of (4) under the general settings when the sharpness condition may not hold. We use the definition of  $\epsilon$ -stationary point as introduced in [15].

**Definition 1.** We call  $\tilde{x}$  an  $\epsilon$ -stationary point of (4) if the following inequality holds:

$$\|\mathcal{G}_t(\tilde{x})\|_2 \leq \epsilon, \quad (12)$$

where  $\mathcal{G}_t(x)$  is the proximal gradient which is defined as:

$$\mathcal{G}_t(x) = t^{-1}(x - S_t(x)). \quad (13)$$

Our convergence rate result of Algorithm 1 is given in Theorem 1, and the proof is given in Section 5.

**Theorem 1.** Denote  $F^* = \inf_{x \in \mathbb{R}^n} F(x)$ . For Algorithm 1 with  $t = 1/L$ , the following conclusion holds.

(a) When (LACC) holds with  $\rho_l > 0$  for any  $k \in \mathbb{N}$ , we can find an  $\epsilon$ -stationary point in

$$\left\lceil \frac{2(1 + \rho_l)(F(x^0) - F^*)}{t\epsilon^2} \right\rceil$$

iterations for any  $\epsilon > 0$ .

(b) When (HACC) holds with  $0 < \rho_h < 1/4$  for any  $k \in \mathbb{N}$ , we can find an  $\epsilon$ -stationary point in

$$\left\lceil \frac{2(1 - \sqrt{\rho_h})^2(F(x^0) - F^*)}{(1 - 2\sqrt{\rho_h})t\epsilon^2} \right\rceil$$

iterations for any  $\epsilon > 0$ .

Theorem 1 shows that IPL finds an  $\epsilon$ -stationary point in  $O(1/\epsilon^2)$  main iterations with the adaptive IPL stopping conditions. Moreover, Theorem 1 achieves the best known convergence rate for PL in [15]. We should point out that we use two adaptive stopping criteria for the subproblem, but [15] requires solving the subproblem (5) exactly (see their Proposition 3) or using their pre-determined subproblem accuracy conditions (see their Theorem 5.2).

## 2.2 Local Convergence under Sharpness Assumption

In this subsection, we analyze the local convergence of IPL (Algorithm 1) to the global optimal solution under the sharpness condition.

**Assumption 1** (Sharpness). There exists a constant  $\lambda_s > 0$  such that the following inequality holds for any  $x \in \mathbb{R}^n$ :

$$F(x) - F(x_*) \geq \lambda_s \Delta(x), \quad (14)$$

where  $\Delta(x) := \min\{\|x - x_*\|_2, \|x + x_*\|_2\}$ .

[11] proved that the sharpness condition (Assumption 1) is satisfied by the RPR (4) with high probability under certain mild conditions.

Another assumption is about the closeness between the initial point and the optimal solution, which can be guaranteed by the modified spectral initialization (see Algorithm 3 in [11]) with high probability under some mild conditions.

**Assumption 2.** *Under Assumption 1, we assume that the initial point  $x^0$  in Algorithm 1 satisfies the following inequalities.*

(a) *If (LACC) is chosen in Algorithm 1, then we assume  $x^0$  satisfies*

$$F(x^0) - F(x_*) \leq \lambda_s^2 / (2L). \quad (15)$$

(b) *If (HACC) is chosen in Algorithm 1, then we assume  $x^0$  satisfies*

$$\Delta(x^0) \leq \frac{\lambda_s(1 - 4\rho_h)}{2(1 - 3\rho_h)L}. \quad (16)$$

We now define the  $\epsilon$ -optimal solution to RPR (4).

**Definition 2.** *We call  $\bar{x}$  an  $\epsilon$ -optimal solution to RPR (4), if  $\Delta(\bar{x}) \leq \epsilon$ .*

Now we are ready to show in Theorem 2 that, in terms of main iteration number, (LACC) leads to local linear convergence and (HACC) leads to local quadratic convergence.

**Theorem 2.** *Let  $t = \frac{1}{L}$  and suppose that Assumption 1 holds. For the sequence  $\{x^k\}_{k=0}^\infty$  generated by Algorithm 1, we have the following conclusions.*

(a) *(Low Accuracy) When (15) holds and (8) holds with  $\rho_l > 0$  for any  $k \in \mathbb{N}$ , we have*

$$\Delta(x^k) \leq \frac{F(x^0) - F(x_*)}{\lambda_s} \left( \frac{1 + 4\rho_l}{2 + 4\rho_l} \right)^k, \quad \forall k \in \mathbb{N}.$$

(b) *(High Accuracy) When (16) holds and (9) holds with  $0 < \rho_h < 1/4$  for any  $k \in \mathbb{N}$ , we have*

$$\Delta(x^k) \leq \frac{\lambda_s(1 - 4\rho_h)}{L(1 - 3\rho_h)} \zeta^{2^k}, \quad \forall k \in \mathbb{N},$$

$$\text{where } \zeta := \frac{L\Delta(x^0)(1 - 3\rho_h)}{\lambda_s(1 - 4\rho_h)}.$$

Theorem 2 shows that, with a good initialization, using (LACC) finds an  $\epsilon$ -optimal solution to (4) within  $O(\log \frac{1}{\epsilon})$  iterations, which is a linear rate, and using (HACC) finds an  $\epsilon$ -optimal solution to (4) within  $O(\log \log \frac{1}{\epsilon})$  iterations, which is a quadratic rate.

## 2.3 Related Work

There are two closely related works that need to be discussed here. [11] studied the PL algorithm for solving RPR (4), and established its local quadratic convergence under the sharpness condition. But their theoretical analysis requires the subproblem (5) to be solved exactly. In practice, [11] proposed to use POGS [27], which is a variant of the alternating direction method of multipliers (ADMM), to solve (5) inexactly. However, they did not provide any convergence analysis for the algorithm when the subproblem (5) is solved inexactly by POGS. [15] also considered solving (4) for obtaining an  $\epsilon$ -stationary point as defined in Definition 1.<sup>1</sup> Indeed, several algorithms were proposed and analyzed in [15]. In particular, a practical algorithm proposed by [15] uses FISTA [20–22] to inexactly solve

<sup>1</sup>The authors of [15] actually considered solving a more general problem  $\min_x g(x) + h(c(x))$ . Here, for simplicity, we assume that  $g = 0$  and this does not affect the discussion.

$$\begin{aligned} \min_{\lambda \in \mathbb{R}^m} \phi_{k,\nu}(\lambda) &= \frac{t}{2} \|x^k/t - \nabla c(x^k)^\top \lambda\|_2^2 \\ &\quad - \lambda^\top (c(x^k) - \nabla c(x^k)x^k) + (h_\nu)^*(\lambda), \end{aligned}$$

which is the dual problem of a smoothed version of (5). Here  $(h_\nu)^*(\cdot)$  is the Fenchel conjugate of  $h_\nu$ , and  $h_\nu$  is the Moreau envelope of  $h$ , which is defined as

$$h_\nu(\lambda) = \inf_{\lambda' \in \mathbb{R}^m} h(\lambda') + \frac{1}{2\nu} \|\lambda' - \lambda\|_2^2.$$

[15] proposed to terminate the FISTA when  $\text{dist}(0; \partial\phi_{k,\nu}(\lambda^{k+1})) \leq \frac{1}{L_h(k+1)^2}$ , and then update  $x^k$  by  $x^{k+1} = x^k - t\nabla c(x^k)^\top \lambda^{k+1}$ . The authors established the overall complexity of this algorithm for suitably chosen parameters  $t$  and  $\nu$ . Compared to [15], we use adaptive stopping criteria and provide a better convergence rate based on Assumption 1.

### 3 FISTA for Solving the Subproblem Inexactly

In this section, we propose to use the FISTA to inexactly solve (5) with more practical stopping criteria that guarantee (LACC) and (HACC). Therefore, the convergence results (Theorems 1 and 2) in Section 2 still apply here.

For simplicity, we let  $t = 1/L$  throughout this section and rewrite (5) as follows.

$$\min_{z \in \mathbb{R}^n} H_k(z) = \frac{1}{2t} \|z\|_2^2 + \|B_k z - d_k\|_1, \quad (17)$$

where we denote  $z = x - x^k$ ,  $B_k = \frac{2}{m} \text{diag}(Ax^k)A$ , and  $d_k = \frac{1}{m} (b - (Ax^k)^2)$ . As a result, (LACC) and (HACC) can be rewritten respectively as

$$H_k(z_k) - \min_{z \in \mathbb{R}^n} H_k(z) \leq \rho_l (H_k(0) - H_k(z_k)), \rho_l \geq 0, \quad (18)$$

and

$$H_k(z_k) - \min_{z \in \mathbb{R}^n} H_k(z) \leq \frac{\rho_h}{2t} \|z_k\|_2^2, 0 \leq \rho_h < 1/4. \quad (19)$$

In IPL, we set  $x^{k+1} = x^k + z_k$ , where  $z_k$  satisfies either (18) or (19). The dual problem of (17) is

$$\max_{\lambda \in \mathbb{R}^m, \|\lambda\|_\infty \leq 1} D_k(\lambda) = -\frac{t}{2} \|B_k^\top \lambda\|_2^2 - \lambda^\top d_k. \quad (20)$$

From weak duality, we know that

$$D_k(\lambda) \leq H_k(z), \forall z \in \mathbb{R}^n, \text{ and } \|\lambda\|_\infty \leq 1. \quad (21)$$

Therefore,  $D_k(\lambda)$  can serve as a lower bound for  $\min_z H_k(z)$ , and we can obtain verifiable stopping criteria that are sufficient conditions for (18) and (19). This leads to our inexact FISTA for solving (17), which is summarized in Algorithm 2. Here we define  $z_k(\lambda) = -tB_k^\top \lambda$ .

**Remark 1.** Here we remark on the step size  $t_{k,j}$  in (22). It can be chosen as  $(tL_k^2)^{-1}$  for some  $L_k \geq \|B_k\|_2$  or chosen by the Armijo backtracking line search. More specifically, suppose that we have an initial step size  $t_{k(-1)} > 0$ . Given the step size  $t_{k(j-1)}, j \geq 0$ , denote

$$\begin{aligned} \lambda_{b,t_s}^{j+1} &= \arg \min_{\lambda \in \mathbb{R}^m, \|\lambda\|_\infty \leq 1} \frac{\gamma_j}{2t_s} \|\lambda - \lambda_c^{j+1}\|_2^2 \\ &\quad + (\lambda - \lambda_c^{j+1})^\top (tB_k B_k^\top \lambda_c^{j+1} + d_k), \end{aligned}$$

---

**Algorithm 2** FISTA for Solving (20)

---

**Input:**  $\lambda^0 \in \mathbb{R}^m$  satisfying  $\|\lambda^0\|_\infty \leq 1$ .  $\lambda_a^0 = \lambda_b^0 = \lambda_c^0 = \lambda^0$ ,  $\gamma_0 = 1$ ,  $\rho_l > 0$  and  $\rho_h \in (0, 1/4)$ .  
**for**  $j = 0, 1, 2 \dots$  **do**

$$\begin{aligned}
\lambda_c^{j+1} &= (1 - \gamma_j)\lambda_a^j + \gamma_j\lambda_b^j, \\
\lambda_b^{j+1} &= \arg \min_{\|\lambda\|_\infty \leq 1} \frac{\gamma_j}{2t_{kj}} \|\lambda - \lambda_c^{j+1}\|_2^2 + \\
&\quad (\lambda - \lambda_c^{j+1})^\top \left( tB_k B_k^\top \lambda_c^{j+1} + d_k \right), t_{kj} > 0, \\
\lambda_a^{j+1} &= (1 - \gamma_j)\lambda_a^j + \gamma_j\lambda_b^{j+1}, \\
\gamma_{j+1} &= 2 / \left( 1 + \sqrt{1 + 4/\gamma_j^2} \right),
\end{aligned} \tag{22}$$

Terminate if one of the following stopping criteria is satisfied:

$$\begin{aligned}
\text{(LACC-FISTA)} \quad & H_k(z_k(\lambda_a^{j+1})) - D_k(\lambda_a^{j+1}) \leq \\
& \rho_l (H_k(0) - H_k(z_k(\lambda_a^{j+1}))),
\end{aligned} \tag{23}$$

$$\begin{aligned}
\text{(HACC-FISTA)} \quad & H_k(z_k(\lambda_a^{j+1})) - D_k(\lambda_a^{j+1}) \leq \\
& \frac{\rho_h}{2t} \|z_k(\lambda_a^{j+1})\|_2^2.
\end{aligned} \tag{24}$$

**end for**

**Output:**  $\lambda_k = \lambda_a^{j+1}$ ,  $z_k = -tB_k^\top \lambda_k$ ,  $x^{k+1} = x^k + z_k$ .

---

and

$$\lambda_{a,t_s}^{j+1} = (1 - \gamma_j)\lambda_a^j + \gamma_j\lambda_{b,t_s}^{j+1}$$

$t_{kj}$  can be selected as

$$\begin{aligned}
t_{kj} &= \max\{t_s | t_s = 2^{-s}t_{k(j-1)}, s \in \mathbb{N}, \\
&\quad \frac{1}{2t_s} \|\lambda_{a,t_s}^{j+1} - \lambda_c^{j+1}\|_2^2 \geq \frac{t}{2} \|B_k^\top \lambda_{a,t_s}^{j+1} - B_k^\top \lambda_c^{j+1}\|_2^2\}.
\end{aligned}$$

We now discuss the overall complexity of the IPL (Algorithm 1) with the subproblem (5) solved inexactly by FISTA (Algorithm 2). For ease of presentation, we denote this algorithm as IPL+FISTA. We assume that IPL is terminated after  $K_\epsilon$  iterations, when an  $\epsilon$ -optimal solution is found, i.e.,  $\Delta(x^{K_\epsilon}) \leq \epsilon$ ,  $\Delta(x^{K_\epsilon-1}) > \epsilon$ . We use  $J_k, k \geq 0$  to denote the number of iterations of FISTA when it is called in the  $k$ -th iteration (getting  $x^{k+1}$  from  $x^k$ ) of IPL. The overall complexity of IPL+FISTA for obtaining an  $\epsilon$ -optimal solution is thus given by  $J(\epsilon) = \sum_{k=0}^{K_\epsilon-1} J_k$ , which equals the times that we call (22). Now we are ready to give the overall complexity of IPL+FISTA in Theorem 3.

**Theorem 3.** Let  $t = 1/L$ ,  $t_{kj} = (t\|B_k\|_2^2)^{-1}$  and suppose that Assumption 1 holds.

- (a) (Low Accuracy) For the overall complexity of Algorithm 1, when using Algorithm 2 with option (23) with  $\rho_l > 0$ , for any  $x^0 \in \mathbb{R}^n$  that satisfies  $\Delta(x^0) \leq E_1$ , we have

$$J(\epsilon) \leq E_2/\epsilon, \forall \epsilon > 0. \tag{25}$$

Here  $E_1, E_2$  are positive constants that only depend on  $\{a_i\}_{i=1}^m, \{b_i\}_{i=1}^m, x_\star, \lambda_s, L$  and  $\rho_l$ , and they will be specified later in the proof.

- (b) (*High Accuracy*) For the overall complexity of Algorithm 1, when using Algorithm 2 with option (24) with  $\rho_h \in (0, 1/4)$ , there exist positive constants  $E_3, E_4$ , and  $\{\epsilon_i\}_{i=1}^\infty$  with  $\epsilon_i > \epsilon_{i+1}, \forall i \in \mathbb{N}_+$  and  $\lim_{i \rightarrow \infty} \epsilon_i = 0$  such that if  $\Delta(x^0) \leq E_3$ , then

$$J(\epsilon_i) \leq E_4/\epsilon_i, \forall i \in \mathbb{N}_+. \quad (26)$$

Here  $E_3, E_4$  only depend on  $\{a_i\}_{i=1}^m, \{b_i\}_{i=1}^m, x_*, \lambda_s, L$  and  $\rho_h$ ,  $\{\epsilon_i\}_{i=1}^\infty$  depends on  $\{\Delta(x^i)\}_{i=1}^\infty$  and they will be specified later in the proof. Note that (26) implies that the worst-case overall complexity might be higher than  $O(1/\epsilon)$  – see the explanation below for more details.

Theorem 3 shows that under the (LACC-FISTA) stopping criterion, we need  $O(1/\epsilon)$  iterations to find an  $\epsilon$ -optimal solution, and under the (HACC-FISTA) stopping criterion, we have the same rate with regard to a countable positive sequence that decreases to zero. Theorem 3 provides better theoretical rates compared to  $O(1/\epsilon^3)$  in [15]. Moreover, the results in Theorem 3 are about the convergence to  $\epsilon$ -optimal solution, while the results in [15] are for convergence to  $\epsilon$ -stationary point. Our results require the sharpness condition, which was not assumed in [15].

For (b) in Theorem 3, we can only find a countable sequence of diminishing  $\epsilon_i$ 's to show the  $O(1/\epsilon_i)$  rate. We cannot show the  $O(1/\epsilon)$  rate for any fixed  $\epsilon > 0$ . This is because of the local quadratic convergence under (HACC) shown in (b) of Theorem 3. For instance, if our initial point  $x^0$  satisfies  $\Delta(x^0) = 2\epsilon > \epsilon$ , under the (HACC), the quadratic convergence result in Theorem 3 (b) implies that  $\Delta(x^1) \leq C\epsilon^2 < \epsilon$  for some constant  $C > 0$  when  $\epsilon$  is sufficiently small. Therefore, IPL finds an  $\epsilon$ -optimal stationary point with only one main iteration. However, Lemma 15 (b) indicates that  $J(\epsilon) = J_0 = O(1/\epsilon^2)$ . Therefore, it is possible that the overall complexity becomes  $O(1/\epsilon^2)$ , which is higher than  $O(1/\epsilon)$ .

## 4 Numerical Experiments

In this section, we conduct numerical experiments to compare our IPL method with existing methods for solving the RPR problem (4). Readers can find the code and datasets to replicate the experiments in this section via <https://github.com/zhengzhongpku/IPL-code-share>. The algorithms that we test include the following ones.

- (i) **PL**: The original proximal linear algorithm proposed by [11] where the subproblem (5) is solved by POGS [27]. POGS terminates when both the primal residual and the dual residual are small enough. In their code, the authors [11] implemented a two-stage trick that uses a relatively larger tolerance in early iterations and a smaller tolerance in later iterations to terminate the POGS. In our comparison, we use all the default parameters set by the authors in their code<sup>2</sup>.
- (ii) **Subgradient method**. The subgradient method with geometrically decaying step sizes was proposed by [13], and they used this algorithm to solve the RPR (4). One typical iteration of this algorithm is

$$x^{k+1} = x^k - \lambda_0 q^k \xi_k / \|\xi_k\|_2, k \geq 0, \quad (27)$$

in which  $\lambda_0 > 0, q \in (0, 1)$  are hyper-parameters and  $\xi_k = \frac{1}{n} \sum_{i=1}^m 2a_i^\top x^k \text{sign}((a_i^\top x^k)^2 - b_i)$ .

- (iii) **IPL-FISTA-Low, IPL-FISTA-High**: our IPL+FISTA algorithm with stopping criteria (LACC-FISTA) and (HACC-FISTA) in Algorithm 2, respectively, and we also used the Armijo backtracking line search discussed in Remark 1.

The initial point for all the tested algorithms is generated by the spectral initialization given in Algorithm 3 in [11]. All the code is run on a server with Intel Xeon E5-2650v4 (2.2GHz). Each task is limited to a single core – no multi-threading is used.

---

<sup>2</sup>The code of [11] can be downloaded from <https://web.stanford.edu/~jduchi/projects/phase-retrieval-code.tgz>



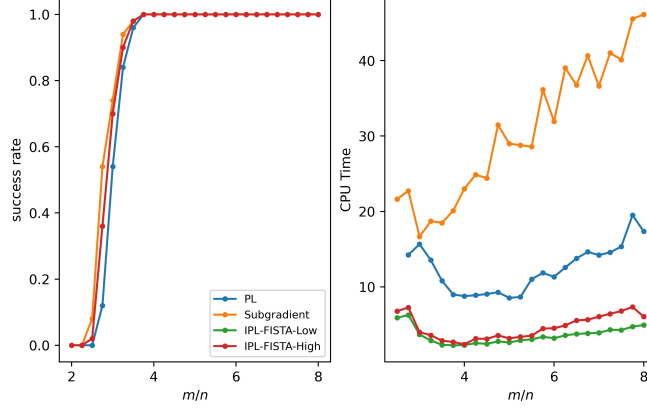


Figure 1: The comparison of success rates and CPU time on synthetic datasets with  $p_{\text{fail}} = 0.05$  and  $n = 500$ .

## 4.1 Synthetic Data

We generate synthetic data following the same manner as [11]. Specifically,  $a_i$ 's are drawn randomly from the normal distribution  $\mathcal{N}(0, I_n)$ . The entries of  $x_\star \in \{-1, 1\}^n$  are drawn randomly from discrete Bernoulli distribution. We denote  $p_{\text{fail}} = |\mathbb{I}_2|/m$ , where  $\mathbb{I}_2$  is generated by random sampling without replacement from  $\{1, 2, \dots, m\}$ .  $\xi_i$ 's for these samples in (2) are independently drawn from Cauchy distribution, which means that

$$b_i = \xi_i = \tilde{M} \tan\left(\frac{\pi}{2} U_i\right), U_i \sim U(0, 1), \forall i \in \mathbb{I}_2,$$

where  $\tilde{M}$  is the sample median of  $\{(a_i^\top x_\star)^2\}_{i=1}^m$ . For a given threshold  $\epsilon > 0$ , we call an algorithm successful if it returns an  $x$  such that the relative error

$$\Delta(x)/\|x_\star\|_2 \leq \epsilon. \quad (28)$$

For each combination of  $n, k = m/n$  and  $p_{\text{fail}}$ , we randomly generate 50 instances according to the above procedure, and we report the success rate of the algorithm among the 50 instances. For IPL-FISTA-Low and IPL-FISTA-High, we set  $\rho_l = \rho_h = 0.24$ . For the subgradient method, we set  $q = 0.998$  which is one of the suggested choices of  $q$  in [13]. Moreover, [13] did not specify how to choose  $\lambda_0$ , and we set  $\lambda_0 = 0.1\|x^0\|_2$  as we found that this choice gave good performance. Since we found that in most cases, the relative error given by PL is in the level of  $[10^{-5}, 10^{-3}]$ , we set  $\epsilon = 10^{-3}$  in (28) for PL. In our comparison, we first run PL using the default settings of the code provided by the authors of [11]. If the returned  $x$  satisfies (28) with  $\epsilon = 10^{-3}$ , then we claim that PL is successful, and we terminate IPL and Subgradient method once they found an iterate with a smaller objective value than the one given by PL. If the iterate returned by PL does not satisfy (28) with  $\epsilon = 10^{-3}$ , then we claim that PL failed, and we then terminate IPL and Subgradient method when  $F(x^k) - F(x_\star) \leq 10^{-7}$ . The CPU time is only reported for the successful cases for PL. The cost of computing the spectral norm  $\|A\|_2$  to obtain  $L$  is included in the total CPU time of PL and IPL.

The simulation results corresponding to  $p_{\text{fail}} = 0.05$  and  $p_{\text{fail}} = 0.15$  are shown in Figures 1 and 2, where the  $x$ -axis corresponds to different values of  $m$  since  $n = 500$  is fixed. From both Figures 1 and 2, we can see that the four algorithms have similar success rates, but the total CPU time of IPL-FISTA-Low and IPL-FISTA-High that includes the cost of computing the spectral norm  $\|A\|_2$  is significantly less than that of others.

## 4.2 Image Recovery

In this section, we compare the four candidate algorithms on images in a similar manner as [11]. In particular, suppose we have an RGB image array  $X_\star \in \mathbb{R}^{n_1 \times n_2 \times 3}$ , we construct the signal as  $x_\star = [\text{vec}(X_\star); 0] \in \mathbb{R}^n$ , in which  $n = \min\{2^s \mid s \in N, 2^s \geq 3n_1n_2\}$ . Let  $H_n \in \frac{1}{\sqrt{n}}\{-1, 1\}^{n \times n}$  be the Hadamard matrix and  $S_j \in \text{diag}(\{-1, 1\}^n), j = 1, 2, \dots, k$  is a random diagonal matrix, and its diagonal elements are independently

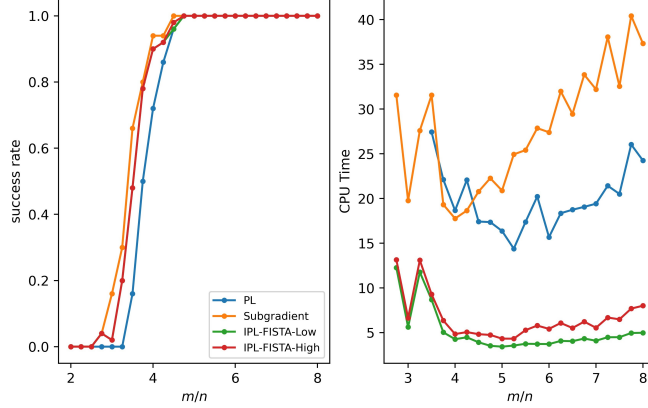


Figure 2: The comparison of success rates and CPU time on synthetic datasets with  $p_{\text{fail}} = 0.15$  and  $n = 500$ .

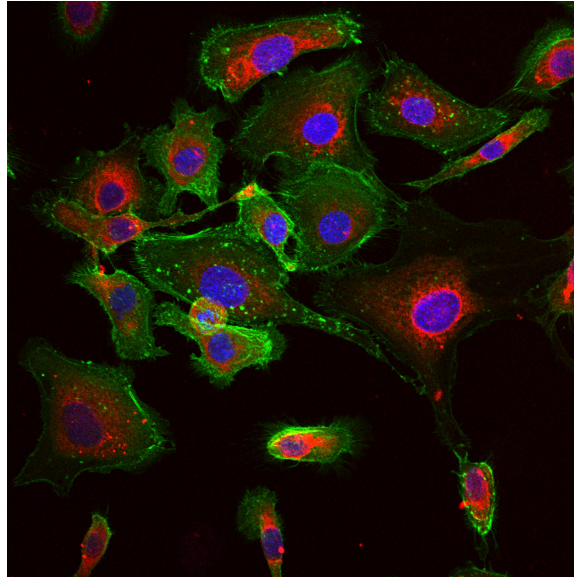


Figure 3: A real RNA nanoparticles image.

distributed as discrete uniform distribution. We then let  $A = \frac{\sqrt{m}}{\sqrt{k}} [H_n S_1; H_n S_2 \dots H_n S_k]$  and we know  $L = 2$  in this case. The advantage of such a mapping is that it mimics the fast Fourier transform, and calculating  $Ay$  is only of time complexity  $O(m \log m)$ . We first examine the numerical comparisons on a real RNA nanoparticles image<sup>3</sup> as shown in Figure 3, and we follow the code of [11] for the experiments on a sub-image with  $n = 2^{18}$ . We also take  $p_{\text{fail}} \in \{0.05, 0.1, 0.15, 0.2\}$ ,  $k \in \{3, 6\}$  and set the noise in the same way as the synthetic datasets. For each combination of dataset parameters, we run 50 replicates by generating 50 different  $A$  and test all the candidate algorithms. We use the same way as the synthetic datasets to define success. For IPL,  $\rho_l = \rho_h = 0.24$ . For the Subgradient method,  $\lambda_0 = 0.1 \|x^0\|_2$ ,  $q = 0.998$ . For a replicate, if PL succeeds, the CPU time is the time needed to reach (28).

Table 1 reports the median CPU time (in seconds) of the candidate algorithms for  $p_{\text{fail}} = 0.1$  and  $m/n = 6$  based on two tolerances  $\epsilon = 10^{-1}$  and  $\epsilon = 10^{-7}$ . We only show the results for this combination of  $p_{\text{fail}}$  and  $m/n$  because other choices give similar results. It is noted from Table 1 that PL can only reach a relative error that takes value in  $[10^{-2}, 10^{-1}]$ , and IPL-FISTA is much more efficient than PL and

<sup>3</sup><https://visualsonline.cancer.gov/details.cfm?imageid=11167>

Table 1: The comparison of the median CPU time in seconds for RNA image recovery.

	$\epsilon = 0.1$	$\epsilon = 10^{-7}$
PL	3166.93	NA
Subgradient	88.14	659.64
IPL-FISTA-Low	6.01	218.14
IPL-FISTA-High	48.56	175.60

Table 2: The comparison of the median CPU time in hours with the interquartile range (IQR) in the parentheses for multiple image recovery tasks.

	<b>RNA</b> ( $n = 2^{22}$ )	<b>Hubble</b> <sup>4</sup> ( $n = 2^{22}$ )
<b>PL</b>	> 10	> 10
<b>Subgradient</b>	4.39 (0.24)	4.72 (0.27)
<b>IPL-FISTA-Low</b>	1.65 (0.41)	2.07 (0.14)
<b>IPL-FISTA-High</b>	1.30 (0.08)	1.26 (0.06)
	<b>James Web</b> <sup>5</sup> ( $n = 2^{21}$ )	<b>Penn State</b> <sup>6</sup> ( $n = 2^{22}$ )
<b>PL</b>	> 10	> 10
<b>Subgradient</b>	1.86 (0.14)	4.84 (0.66)
<b>IPL-FISTA-Low</b>	0.90 (0.21)	2.23 (0.16)
<b>IPL-FISTA-High</b>	0.56 (0.09)	1.41 (0.21)

Subgradient methods.

Additional experimental results are reported in Table 2 for comparing CPU time for the four candidate algorithms. We provide four images with  $n$  being at most  $2^{22}$ . The experiments use the same  $m/n$  and  $p_{\text{fail}}$ , and the CPU time based on ten replications is reported in the form of “median (Interquartile Range)”. Subgradient method, IPL-FISTA-Low and IPL-FISTA-High are terminated with tolerance  $\epsilon = 10^{-7}$ . We can see that PL cannot terminate within 10 hours, and IPL-FISTA still enjoys the best efficiency.

## 5 Proofs

### 5.1 Proof of Theorem 1

Before proceeding, we first present some lemmas.

**Lemma 1** (Weak Convexity). (*Discussion of Condition C2 in [11]*) *The following inequalities hold for any  $x, y \in \mathbb{R}^n$ , and  $L = \frac{2}{m}\|A\|_2^2$ .*

$$|F(x) - F(x; y)| \leq \frac{L}{2}\|x - y\|_2^2, \quad (29)$$

$$F(x) \leq F_t(x; y), \quad \forall t \leq \frac{1}{L}. \quad (30)$$

**Lemma 2** (see, e.g., equation (5.2) in [15]). *The following inequality holds for any  $0 < t \leq 1/L$  and  $x, y \in \mathbb{R}^n$ .*

$$F(x) - F(y) + \varepsilon(y; x) \geq \frac{1}{2t}\|x - S_t(x)\|_2^2. \quad (31)$$

The following lemma studies one iteration of our IPL algorithm (as summarized in Algorithm 1).

<sup>4</sup><https://www.nasa.gov/image-feature/goddard/2017/hubble-hones-in-on-a-hypergiants-home>

<sup>5</sup><https://www.nasa.gov/webbfirstimages>

<sup>6</sup><https://www.britannica.com/topic/Pennsylvania-State-University>

**Lemma 3.** *If  $0 < t \leq \frac{1}{L}$ , we have the following inequalities.*

(a) *When the low accuracy condition (8) is satisfied, we have*

$$F(x^k) - F(x^{k+1}) \geq \frac{1}{2(1 + \rho_l)t} \|x^k - S_t(x^k)\|_2^2. \quad (32)$$

(b) *When the high accuracy condition (9) is satisfied, we have*

$$F(x^k) - F(x^{k+1}) \geq \frac{1 - 2\sqrt{\rho_h}}{2t(1 - \sqrt{\rho_h})^2} \|x^k - S_t(x^k)\|_2^2. \quad (33)$$

*Proof of Lemma 3.* We first prove part (a) of Lemma 3 and then prove part (b) of Lemma 3.

(a). Letting  $x = x^k$  and  $y = x^{k+1}$  in (31), we have

$$\begin{aligned} F(x^k) - F(x^{k+1}) &\geq \frac{1}{2t} \|x^k - S_t(x^k)\|_2^2 - \varepsilon(x^{k+1}; x^k) \\ &\geq \frac{1}{2t} \|x^k - S_t(x^k)\|_2^2 - \rho_l(F(x^k) - F(x^{k+1})), \end{aligned}$$

where the second inequality follows from (8) and (30). This proves (32) in Lemma 3(a).

(b). When (9) holds, since  $F_t(\cdot; x^k)$  is  $\frac{1}{t}$ -strongly convex, we have

$$\frac{\rho_h}{2t} \|x^k - x^{k+1}\|_2^2 \geq \varepsilon(x^{k+1}; x^k) \geq \frac{1}{2t} \|x^{k+1} - S_t(x^k)\|_2^2. \quad (34)$$

Let  $u = x^k - x^{k+1}$ ,  $v = x^{k+1} - S_t(x^k)$ . From (34) we have

$$\rho_h \|u\|_2^2 \geq \|v\|_2^2, \quad (35)$$

and

$$\begin{aligned} &\|u + v\|_2^2 - (1 - \sqrt{\rho_h})^2 \|u\|_2^2 \\ &= (2\sqrt{\rho_h} - \rho_h) \|u\|_2^2 + 2u^\top v + \|v\|_2^2 \\ &\geq \sqrt{\rho_h}(2 - \sqrt{\rho_h}) \|u\|_2^2 - 2\|u\|_2 \|v\|_2 + \|v\|_2^2 \\ &= (\sqrt{\rho_h} \|u\|_2 - \|v\|_2) ((2 - \sqrt{\rho_h}) \|u\|_2 - \|v\|_2) \\ &\geq 0, \end{aligned} \quad (36)$$

where the first inequality is from the Cauchy-Schwarz inequality, and the second inequality follows from (35) and the fact that  $\rho_h \in (0, 1/4)$ . Therefore, from (34) and (36) we have

$$\begin{aligned} \varepsilon(x^{k+1}; x^k) &\leq \frac{\rho_h}{2t} \|x^k - x^{k+1}\|_2^2 \\ &\leq \frac{\rho_h}{2t(1 - \sqrt{\rho_h})^2} \|x^k - S_t(x^k)\|_2^2, \end{aligned}$$

which, together with (31), yields

$$\begin{aligned} &F(x^k) - F(x^{k+1}) \\ &\geq \frac{1}{2t} \|x^k - S_t(x^k)\|_2^2 - \varepsilon(x^{k+1}; x^k) \\ &\geq \frac{1}{2t} \|x^k - S_t(x^k)\|_2^2 - \frac{\rho_h}{2t(1 - \sqrt{\rho_h})^2} \|x^k - S_t(x^k)\|_2^2 \\ &= \frac{1 - 2\sqrt{\rho_h}}{2t(1 - \sqrt{\rho_h})^2} \|x^k - S_t(x^k)\|_2^2. \end{aligned}$$

This proves (33) in Lemma 3(b).

Therefore, the proof of Lemma 3 is complete.  $\square$

Now we are ready to prove Theorem 1.

*Proof of Theorem 1.* We will prove (a) and (b) together. Both (32) and (33) indicate that  $\{x^k\}$  generated by Algorithm 1 satisfies

$$F(x^k) - F(x^{k+1}) \geq \beta t \|\mathcal{G}_t(x^k)\|_2^2,$$

in which  $\beta > 0$  is a constant that  $\beta = 1/(2(1 + \rho_l))$  for LACC and  $\beta = (1 - 2\sqrt{\rho_h})/(2(1 - \sqrt{\rho_h})^2)$  for HACC. Letting  $F^* = \inf_{x \in \mathbb{R}^n} F(x)$ , we have

$$\begin{aligned} F(x^0) - F^* &\geq \sum_{k=0}^{K_0-1} \beta t \|\mathcal{G}_t(x^k)\|_2^2 \\ &\geq \beta t K_0 \min_{0 \leq k \leq K_0-1} \|\mathcal{G}_t(x^k)\|_2^2. \end{aligned}$$

Then, our IPL (Algorithm 1) finds an  $\epsilon$ -stationary point to RPR (4) in

$$\left\lceil \frac{F(x^0) - F^*}{\beta t \epsilon^2} \right\rceil$$

iterations. Therefore, the proof of both (a) and (b) in Theorem 1 is complete.  $\square$

## 5.2 Proof of Theorem 2

To prove Theorem 2, we need the following lemmas.

**Lemma 4.** (Part of the Proof for Theorem 1 in [11]) Let  $t = 1/L$ . For any  $x^k, x^{k+1} \in \mathbb{R}^n$ , we have

$$\begin{aligned} F(x^{k+1}) - F(x_*) + \frac{L}{4} \|x^{k+1} - x_*\|_2^2 \\ \leq L \|x^k - x_*\|_2^2 + 2\varepsilon(x^{k+1}, x^k), \end{aligned} \quad (37)$$

which also holds if we replace  $x_*$  by  $-x_*$ .

*Proof of Lemma 4.* We have

$$\begin{aligned} &\frac{L}{2} \|x_* - S_t(x^k)\|_2^2 \\ &\geq \frac{L}{4} \|x_* - x^{k+1}\|_2^2 - \frac{L}{2} \|S_t(x^k) - x^{k+1}\|_2^2 \\ &\geq \frac{L}{4} \|x_* - x^{k+1}\|_2^2 - \varepsilon(x^{k+1}; x^k). \end{aligned} \quad (38)$$

where the first inequality follows from the Cauchy-Schwarz inequality and the second one is from the convexity of  $\|\cdot\|_2^2$ . We then have

$$\begin{aligned} &F(x^{k+1}) \\ &\leq F(x^{k+1}; x^k) + \frac{L}{2} \|x^k - x^{k+1}\|_2^2 \\ &= F_t(S_t(x^k); x^k) + \varepsilon(x^{k+1}; x^k) \\ &\leq F(x_*; x^k) + \frac{L}{2} \|x^k - x_*\|_2^2 - \frac{L}{2} \|x_* - S_t(x^k)\|_2^2 \\ &\quad + \varepsilon(x^{k+1}; x^k) \\ &\leq F(x_*) + L \|x^k - x_*\|_2^2 - \frac{L}{2} \|x_* - S_t(x^k)\|_2^2 + \varepsilon(x^{k+1}; x^k), \end{aligned}$$

where the first and the last inequalities are from (29), and the second inequality is from the strong convexity of  $F_t(\cdot; x^k)$ . Combining this inequality with (38) yields (37). It is easy to find that all the proofs still hold if we replace  $x_*$  with  $-x_*$ . Therefore, the proof of Lemma 4 is complete.  $\square$

**Lemma 5** (One-step progress). *Let  $t = 1/L$  and suppose that Assumption 1 holds.*

(a) *(Low accuracy condition) When (8) is satisfied for some  $\rho_l \geq 0$ , we have*

$$\begin{aligned} & (1 + 2\rho_l) (F(x^{k+1}) - F(x_*)) \\ & \leq 2\rho_l (F(x^k) - F(x_*)) + L(\Delta(x^k))^2. \end{aligned} \quad (39)$$

*If we also have  $F(x^k) - F(x_*) \leq \lambda_s^2/(2L)$ , then we have*

$$\begin{aligned} & (1 + 2\rho_l) (F(x^{k+1}) - F(x_*)) \\ & \leq \left(\frac{1}{2} + 2\rho_l\right) (F(x^k) - F(x_*)). \end{aligned} \quad (40)$$

(b) *(High accuracy condition) When (9) is satisfied, we have*

$$\lambda_s \Delta(x^{k+1}) \leq \frac{L(1 - 3\rho_h)}{(1 - 4\rho_h)} (\Delta(x^k))^2. \quad (41)$$

*Proof of Lemma 5.* We first prove part (a) of Lemma 5 and then prove part (b) of Lemma 5.

(a). When the low accuracy condition (8) holds, from (30) we have

$$\begin{aligned} \varepsilon(x^{k+1}; x^k) & \leq \rho_l (F_t(x^k; x^k) - F_t(x^{k+1}; x^k)) \\ & \leq \rho_l (F(x^k) - F(x^{k+1})), \end{aligned}$$

which, combining with (37), yields

$$\begin{aligned} & (1 + 2\rho_l) (F(x^{k+1}) - F(x_*)) + \frac{L}{4} \|x^{k+1} - x_*\|_2^2 \\ & \leq 2\rho_l (F(x^k) - F(x_*)) + L \|x^k - x_*\|_2^2. \end{aligned}$$

Discarding the term  $\frac{L}{4} \|x^{k+1} - x_*\|_2^2$ , we get

$$\begin{aligned} & (1 + 2\rho_l) (F(x^{k+1}) - F(x_*)) \\ & \leq 2\rho_l (F(x^k) - F(x_*)) + L \|x^k - x_*\|_2^2. \end{aligned}$$

Since (37) also holds when  $x_*$  is replaced by  $-x_*$ , we also have

$$\begin{aligned} & (1 + 2\rho_l) (F(x^{k+1}) - F(x_*)) \\ & \leq 2\rho_l (F(x^k) - F(x_*)) + L \|x^k + x_*\|_2^2. \end{aligned}$$

This proves (39). (40) holds because of Assumption 1. Hence, it proves part (a) of Lemma 5.

(b). When the high accuracy condition (9) holds, we have

$$\begin{aligned} & 2\varepsilon(x^{k+1}; x^k) \\ & \leq \rho_h L \|x^k - x^{k+1}\|_2^2 \\ & \leq \rho_h L \left( \frac{\|x^{k+1} - x_*\|_2^2}{4\rho_h} + \frac{\|x^k - x_*\|_2^2}{1 - 4\rho_h} \right), \end{aligned}$$

where the second inequality is from the Cauchy-Schwarz inequality. Combining with (37), we have

$$F(x^{k+1}) - F(x_*) \leq \frac{(1 - 3\rho_h)L}{1 - 4\rho_h} \|x^k - x_*\|_2^2.$$

Similarly, replacing  $x_*$  by  $-x_*$ , we have

$$F(x^{k+1}) - F(x_*) \leq \frac{(1 - 3\rho_h)L}{1 - 4\rho_h} \|x^k + x_*\|_2^2.$$

Combining the above two inequalities with Assumption 1 yields (41), which proves part (b) of Lemma 5.

Therefore, the proof of Lemma 5 is complete.  $\square$

Now we are ready to give the proof of Theorem 2.

*Proof of Theorem 2.* (a). We will prove that for any  $k \in \mathbb{N}$ ,

$$F(x^k) - F(x_*) \leq (F(x^0) - F(x_*)) \left( \frac{1 + 4\rho_l}{2 + 4\rho_l} \right)^k \quad (42)$$

by induction, which immediately leads to the conclusion of (a) by Assumption 1. First, (42) clearly holds for  $k = 0$ . Now we assume that it holds for  $k$ . For  $k + 1$ , since (15) holds, we have

$$\begin{aligned} & F(x^k) - F(x_*) \\ & \leq (F(x^0) - F(x_*)) \left( \frac{1 + 4\rho_l}{2 + 4\rho_l} \right)^k \\ & \leq F(x^0) - F(x_*) \\ & \leq \lambda_s^2 / (2L). \end{aligned}$$

Using (40) directly proves that (42) holds when  $k$  is replaced by  $k + 1$ . This proves part (a) of Theorem 2. (b). Note that (41) is equivalent to

$$\frac{L\Delta(x^{k+1})(1 - 3\rho_h)}{\lambda_s(1 - 4\rho_h)} \leq \left( \frac{L\Delta(x^k)(1 - 3\rho_h)}{\lambda_s(1 - 4\rho_h)} \right)^2. \quad (43)$$

Since (16) holds, from (43) we know that

$$\frac{L\Delta(x^1)(1 - 3\rho_h)}{\lambda_s(1 - 4\rho_h)} \leq \left( \frac{1}{2} \right)^2.$$

Plug this back into (43) and work recursively, we get

$$\frac{L\Delta(x^k)(1 - 3\rho_h)}{\lambda_s(1 - 4\rho_h)} \leq \left( \frac{1}{2} \right)^{2^k}.$$

Then,  $\Delta(x^k) \rightarrow 0$ , which together with (43) proves the quadratic convergence in part (b) of Theorem 2. Therefore, the proof of Theorem 2 is complete.  $\square$

### 5.3 Proof of Theorem 3

Throughout this subsection, we assume that the assumptions in Theorem 3 hold. That is, we assume  $t = 1/L$ ,  $t_{kj} = (t\|B_k\|_2^2)^{-1}$  in Algorithm 2, and Assumption 1 holds. To prove Theorem 3, we first present some lemmas.

**Lemma 6** (Local Lipschitz Constant of  $F(x)$ ). *It holds that*

$$\sup_{x, y \in \mathbb{R}^n, \Delta(x) \leq r, \Delta(y) \leq r, x \neq y} \frac{|F(x) - F(y)|}{\|x - y\|_2} \leq L(\|x_*\|_2 + r).$$

*Proof of Lemma 6.* Denote  $u = (x - y)/\|x - y\|_2$  and  $v = (x + y)/\|x + y\|_2$  when  $x + y \neq 0$ .  $v$  is set as 0

when  $x + y = 0$ .

$$\begin{aligned}
& |F(x) - F(y)| / \|x - y\|_2 \\
&= \frac{1}{m\|x - y\|_2} \left| \sum_{i=1}^m |(a_i^\top x)^2 - b_i| - \sum_{i=1}^m |(a_i^\top y)^2 - b_i| \right| \\
&\leq \frac{1}{m\|x - y\|_2} \sum_{i=1}^m |(a_i^\top x)^2 - (a_i^\top y)^2| \\
&= \frac{1}{m\|x - y\|_2} \sum_{i=1}^m |(a_i^\top (x - y))(a_i^\top (x + y))| \\
&= \|x + y\|_2 \frac{1}{m} \sum_{i=1}^m |a_i^\top u| \cdot |a_i^\top v|.
\end{aligned}$$

Recalling that  $L = \frac{2}{m} \|A\|_2^2 = \frac{2}{m} \|\sum_{i=1}^m a_i a_i^\top\|_2$  as defined in Algorithm 1 and noticing the fact that when  $\Delta(x) \leq r$  and  $\Delta(y) \leq r$ , we have  $\|x + y\|_2 \leq 2(\|x_\star\|_2 + r)$ , and hence we can claim that

$$\begin{aligned}
& \|x + y\|_2 \frac{1}{m} \sum_{i=1}^m |a_i^\top u| \times |a_i^\top v| \\
&\leq \|x + y\|_2 \left( u^\top \left( \frac{1}{2m} \sum_{i=1}^m a_i a_i^\top \right) u + v^\top \left( \frac{1}{2m} \sum_{i=1}^m a_i a_i^\top \right) v \right) \\
&\leq L(\|x_\star\|_2 + r),
\end{aligned}$$

which proves the desired result. Therefore, the proof of Lemma 6 is complete.  $\square$

**Lemma 7.** *If Assumption 1 holds, then for any  $r \geq 0$ , we have*

$$\begin{aligned}
& \{x \in \mathbb{R}^n : \Delta(x) \leq E(r)\} \\
&\subseteq \{x \in \mathbb{R}^n : F(x) - F(x_\star) \leq r\} \\
&\subseteq \{x \in \mathbb{R}^n : \Delta(x) \leq r/\lambda_s\},
\end{aligned}$$

in which

$$E(r) = \left( \sqrt{L^2 \|x_\star\|_2^2 + 4rL} - L\|x_\star\|_2 \right) / (2L). \quad (44)$$

This relationship also indicates that  $E(r) \leq r/\lambda_s$ .

*Proof of Lemma 7.* For  $x \in \mathbb{R}^n$ , if  $\Delta(x) \leq E(r)$  for some  $r \geq 0$ , without loss of generality, we assume that  $\Delta(x) = \|x - x_\star\|_2$ . From Lemma 6, we have

$$F(x) - F(x_\star) \leq LE(r)(\|x_\star\|_2 + E(r)) = r,$$

which proves the first inclusion. The second inclusion follows immediately from Assumption 1. Therefore, the proof of Lemma 7 is complete.  $\square$

**Lemma 8** (Bound of  $\|B_k\|_2$ ). *For any  $r \geq 0$ , if  $\sup_{k \in \mathbb{N}} \Delta(x^k) \leq r$ , then*

$$\sup_{k \in \mathbb{N}} \|B_k\|_2 \leq B(r) := \frac{2}{m} \|A\|_2 (\|x_\star\|_2 + r) \max_{i=1,2,\dots,m} \|a_i\|_2.$$

*Proof of Lemma 8.* Since  $B_k = \frac{2}{m} \text{diag}(Ax^k)A$ , we have  $\|B_k\|_2 \leq \frac{2}{m} \|Ax^k\|_\infty \|A\|_2$  and  $\|x^k\|_2 \leq \|x_\star\|_2 + r$ . The desired result follows by using  $\|Ax^k\|_\infty \leq (\|x_\star\|_2 + r) \max_{i=1,2,\dots,m} \|a_i\|_2$ . Therefore, the proof of Lemma 8 is complete.  $\square$



Next, we provide Lemmas 9 and 10 to show the convergence rate for solving the subproblem with Algorithm 2. These results can be used for both conditions for (a) and (b).

**Lemma 9** (see [20]). *In the  $j$ -th iteration of Algorithm 2, we have*

$$\begin{aligned} \max_{\|\lambda\|_\infty \leq 1} D_k(\lambda) - D_k(\lambda_a^j) &\leq \frac{tC\|B_k\|_2^2}{(j+1)^2} \|\lambda^0 - \lambda_{k+1}^*\|_2^2 \\ &\leq \frac{tCm\|B_k\|_2^2}{(j+1)^2}, \forall j \in \mathbb{N}, \end{aligned}$$

where  $\lambda_{k+1}^* \in \arg \max_{\|\lambda\|_\infty \leq 1} D_k(\lambda)$  and  $C > 0$  is universal constant.

**Lemma 10** (Theorem 4 in [28]). *For Algorithm 2, there exist universal positive constants  $C', C''$  such that, when  $j \geq C', j \in \mathbb{N}$ , it holds*

$$H_k(z_k(\lambda_a^j)) - D_k(\lambda_a^j) \leq \frac{C''tm\|B_k\|_2^2}{j+1}. \quad (45)$$

*Proof of Lemma 10.* Theorem 4 in [28] indicates that if

$$\begin{aligned} (j+1)^2 &\geq \max \left\{ \frac{2Ctm\|B_k\|_2^2/t}{m\|B_k\|_2^2}, \frac{2Ctm\|B_k\|_2^2\|B_k\|_2^2m}{(1/t)\epsilon^2} \right\} \\ &= \max \left\{ 2C, \frac{2Ct^2m^2\|B_k\|_2^4}{\epsilon^2} \right\} \end{aligned}$$

then we have

$$H_k(z_k(\lambda_a^j)) - D_k(\lambda_a^j) \leq \epsilon.$$

Here  $C$  is the constant used in Lemma 9. Specifically, by choosing  $\epsilon = \frac{tm\|B_k\|_2^2\sqrt{2C}}{j+1}$ , we know that if  $j \geq C' := \sqrt{2C} - 1$  holds, then (45) holds, where  $C'' = \sqrt{2C}$ . Therefore, the proof of Lemma 10 is complete.  $\square$

We now define some constants.

$$\begin{aligned} E_1 &= \min\{E(\lambda_s^2/(2L)), E(\lambda_s M_1 B^2(\lambda_s/(2L))/C')\}, \\ E_2 &= \frac{(2+4\rho_l)M_1 B^2(\lambda_s/(2L))L(\|x_\star\|_2 + \lambda_s/(2L))}{\lambda_s}, \\ E_3 &= \min\{E_0/2, B(E_0/2)\sqrt{M_2/C'}, E(\lambda_s^2/(2L))\}, \\ E_4 &= 4M_2 B^2(E_0/2)/(3E_0), \end{aligned}$$

and  $\{\epsilon_i\}_{i=1}^\infty = \{\Delta(x^i)\}_{i=1}^\infty$ , where  $M = \frac{\lambda_s}{2L} (\frac{\lambda_s}{2L} + \|x_\star\|_2)^{-1}$ ,  $M_1 = 2C''tm(\rho_l+1)/(\lambda_s\rho_l)$ ,  $M_2 = 4C''t^2m(\rho_h+1)/(\rho_h M^2)$ ,  $E_0 = \frac{\lambda_s(1-4\rho_h)}{(1-3\rho_h)L}$  and  $C', C''$  are universal positive constants mentioned in Lemma 10.

We now formally state the sufficiency of (23) and (24) for (8) and (9) in Lemma 11 so that we can use the linear and quadratic convergence rate for main iterations.

**Lemma 11.** *For Algorithm 2, (23) indicates (8) and (24) indicates (9).*

*Proof of Lemma 11.* Note that  $z_k(\lambda_a^{j+1}) = -tB_k^\top \lambda_a^{j+1}$ , and  $x^{k+1} = x^k + z_k(\lambda_a^{j+1})$ . The conclusion immediately holds because  $H_k(z_k(\lambda_a^{j+1})) = F_t(x^{k+1}; x^k)$  and

$$H_k(z_k(\lambda_a^{j+1})) - D_k(\lambda_a^{j+1}) \geq H_k(z_k(\lambda_a^{j+1})) - \min_{z \in \mathbb{R}^n} H_k(z)$$

which is from strong duality. Therefore, the proof of Lemma 11 is complete.  $\square$

Based on Lemma 11, we prove some properties induced by conditions of Theorem 3. In particular, Lemma 12 gives the ones induced by part (a) of Theorem 3, and Lemma 13 gives the ones induced by part (b) of Theorem 3.

**Lemma 12.** *Assume Assumption 1 and (8) hold for any  $k \in \mathbb{N}$  with some  $\rho_l \geq 0$ . If  $\Delta(x^0) \leq E_1$ , then for any  $k \in \mathbb{N}$ , we have*

$$F(x^k) - F(x_\star) \leq \lambda_s^2/(2L) \quad (46)$$

and

$$\Delta(x^k) \leq \min\{\lambda_s/(2L), M_1 B^2(\lambda_s/(2L))/C'\}. \quad (47)$$

*Proof of Lemma 12.* Note that  $E(\cdot)$  defined in (44) is monotonically increasing. Since  $\Delta(x^0) \leq E_1$ , from Lemma 7 we have

$$F(x^0) - F(x_\star) \leq \lambda_s \min\{\lambda_s/(2L), M_1 B^2(\lambda_s/(2L))/C'\}.$$

Therefore, we can apply (40) and it implies that

$$\begin{aligned} F(x^k) - F(x_\star) &\leq F(x^0) - F(x_\star) \\ &\leq \lambda_s \min\{\lambda_s/(2L), M_1 B^2(\lambda_s/(2L))/C'\}, \end{aligned} \quad (48)$$

which proves (46). From (14), we have

$$\lambda_s \Delta(x^k) \leq F(x^k) - F(x_\star),$$

which leads to (47). Therefore, the proof of Lemma 12 is complete.  $\square$

**Lemma 13.** *Assume Assumption 1 and (9) hold for any  $k \in \mathbb{N}$  with some  $0 \leq \rho_h < 1/4$ . If  $\Delta(x^0) \leq E_3$ , then for any  $k \in \mathbb{N}$ , (46) holds, and the following two inequalities hold*

$$\Delta(x^{k+1}) \leq \frac{1}{2} \Delta(x^k) \quad (49)$$

$$\Delta(x^k) \leq E_3. \quad (50)$$

*Proof of Lemma 13.* Based on (41), we have

$$E_0 \Delta(x^{k+1}) \leq \Delta^2(x^k), \forall k \geq 0. \quad (51)$$

We prove (49) by induction. When  $k = 0$ , noticing that  $\Delta(x^0) \leq E_0/2$ , by (51), we have  $E_0 \Delta(x^1) \leq \Delta(x^0)(E_0/2)$ , and therefore (49) holds for  $k = 0$ . We now assume that (49) holds for  $k < k_0$  for  $k_0 \in \mathbb{N}$ . We then have  $\Delta(x^{k_0}) \leq \Delta(x^0) \leq E_0/2$  and using (51), we have  $E_0 \Delta(x^{k_0+1}) \leq \Delta(x^{k_0})(E_0/2)$ , which implies that (49) holds for  $k = k_0$ . This completes the proof for (49), which immediately indicates (50) for any  $k \geq 0$ . This indicates that  $\Delta(x^k) \leq \Delta(x^0) \leq E_3 \leq E(\lambda_s^2/(2L))$ . By Lemma 7,  $F(x^k) - F(x_\star) \leq \lambda_s^2/(2L)$ , which indicates that (46) holds for any  $k \geq 0$ . Therefore, the proof of Lemma 13 is complete.  $\square$

We can notice that a common property under conditions for (a) or (b) is that (46) holds for any  $k \geq 0$ , based on which, we give Lemma 14 to show another common property.

**Lemma 14.** *If (46) and Assumption 1 hold, then we have*

$$\frac{\lambda_s}{2L} \left( \frac{\lambda_s}{2L} + \|x_\star\|_2 \right)^{-1} \Delta(x^k) \leq \|x^k - S_t(x^k)\|_2, \quad (52)$$

$$\frac{1}{2} \lambda_s \Delta(x^k) \leq F_t(x^k; x^k) - F_t(S_t(x^k); x^k). \quad (53)$$

*Proof of Lemma 14.* For fixed  $k \geq 0$ , we define  $\tilde{x}^0 = x^k$  and  $\tilde{x}^1 = S_t(x^k)$ . Therefore  $\tilde{x}^0$  and  $\tilde{x}^1$  satisfy (8) (with  $k$  replaced by 0 and  $x$  replaced by  $\tilde{x}$ ) with  $\rho_l = 0$ . Hence we have  $\lambda_s \Delta(\tilde{x}^0) \leq F(\tilde{x}^0) - F(x_*) \leq \lambda_s^2/(2L)$  and

$$\begin{aligned}
& F(\tilde{x}^0) - F(\tilde{x}^1) \\
&= F(\tilde{x}^0) - F(x_*) - (F(\tilde{x}^1) - F(x_*)) \\
&\geq F(\tilde{x}^0) - F(x_*) - \frac{1}{2} (F(\tilde{x}^0) - F(x_*)) \\
&= \frac{1}{2} (F(\tilde{x}^0) - F(x_*))
\end{aligned} \tag{54}$$

where the inequality is from (40) with  $\rho_l = 0$ . Moreover, from Assumption 1 we have

$$\begin{aligned}
\lambda_s \Delta(\tilde{x}^1) &\leq F(\tilde{x}^1) - F(x_*) \leq (1/2) (F(\tilde{x}^0) - F(x_*)) \\
&\leq \lambda_s^2/(2L),
\end{aligned}$$

where the second inequality can be obtained by applying (40). Thus, we have  $\max\{\Delta(\tilde{x}^0), \Delta(\tilde{x}^1)\} \leq \lambda_s/(2L)$ . So, based on Lemma 6 and Assumption 1,

$$\begin{aligned}
\lambda_s \Delta(\tilde{x}^0) &\leq F(\tilde{x}^0) - F(x_*) \leq 2 (F(\tilde{x}^0) - F(\tilde{x}^1)) \\
&\leq 2L \left( \frac{\lambda_s}{2L} + \|x_*\|_2 \right) \|\tilde{x}^0 - \tilde{x}^1\|_2,
\end{aligned} \tag{55}$$

where the second inequality is from (54), and the last inequality is from Lemma 6. (55) implies that

$$\|\tilde{x}^0 - \tilde{x}^1\|_2 \geq \frac{\lambda_s}{2L} \left( \frac{\lambda_s}{2L} + \|x_*\|_2 \right)^{-1} \Delta(\tilde{x}^0).$$

This proves (52). To prove (53), without loss of generality, we assume that  $\Delta(\tilde{x}^0) = \|\tilde{x}^0 - x_*\|_2$ . Using (29) and the fact that  $F_t(\tilde{x}^0; \tilde{x}^0) = F(\tilde{x}^0)$  and noticing that  $\tilde{x}^1 = S_t(\tilde{x}^0)$  is the minimizer of  $F_t(\cdot; \tilde{x}^0)$ , we have

$$\begin{aligned}
F_t(\tilde{x}^0; \tilde{x}^0) - F_t(\tilde{x}^1; \tilde{x}^0) &\geq F(\tilde{x}^0) - F_t(x_*; \tilde{x}^0) \\
&= F(\tilde{x}^0) - F(x_*; \tilde{x}^0) - \frac{L}{2} \Delta^2(\tilde{x}^0) \\
&\geq F(\tilde{x}^0) - F(x_*) - L \Delta^2(\tilde{x}^0) \\
&\geq \lambda_s \Delta(\tilde{x}^0) - L \Delta^2(\tilde{x}^0),
\end{aligned}$$

where the second inequality is due to (29), and the last inequality is due to Assumption 1. Using the fact that  $\Delta(\tilde{x}^0) \leq (F(\tilde{x}^0) - F(x_*))/\lambda_s \leq \lambda_s/(2L)$  based on Assumption 1, we have

$$F_t(\tilde{x}^0; \tilde{x}^0) - F_t(\tilde{x}^1; \tilde{x}^0) \geq (\lambda_s - L \frac{\lambda_s}{2L}) \Delta(\tilde{x}^0) = \frac{\lambda_s}{2} \Delta(\tilde{x}^0),$$

which proves (53). Therefore, the proof of Lemma 14 is complete.  $\square$

Now, we are ready to give an upper bound for  $J_k$ .

**Lemma 15.** *Assume that Assumption 1 holds. For Algorithm 2 under either options (23) with  $\rho_l > 0$  or (24) with  $\rho_h \in (0, 1/4)$ , for any  $k \in \mathbb{N}$ , if (46) holds, the following statements hold.*

(a) (Low Accuracy) When using option (23), we have

$$J_k \leq \left\lceil \max \left\{ \frac{M_1 \|B_k\|_2^2}{\Delta(x^k)}, C' \right\} \right\rceil - 1. \tag{56}$$

(b) (High Accuracy) When using option (24), we have

$$J_k \leq \left\lceil \max \left\{ \frac{M_2 \|B_k\|_2^2}{\Delta^2(x^k)}, C' \right\} \right\rceil - 1. \tag{57}$$

Here,  $C'$  and  $C''$  are the constants in Lemma 10 and  $M = \frac{\lambda_s}{2L} \left( \frac{\lambda_s}{2L} + \|x_\star\|_2 \right)^{-1}$ ,  $M_1 = 2C''tm(\rho_l + 1)/(\lambda_s\rho_l)$  and  $M_2 = 4C''t^2m(\rho_h + 1)/(\rho_hM^2)$ .

*Proof of Lemma 15 . (a).* By choosing

$$j = \left\lceil \max \left\{ C', \left( \frac{2C''tm\|B_k\|_2^2(\rho_l + 1)}{\rho_l\lambda_s\Delta(x^k)} \right) \right\} \right\rceil - 1, \quad (58)$$

we have

$$\frac{C''tm\|B_k\|_2^2}{j+2} \leq \lambda_s\rho_l\Delta(x^k)/(2+2\rho_l), \quad (59)$$

which, together with (45), yields that

$$\begin{aligned} & H_k(z_k(\lambda_a^{j+1})) - D_k(\lambda_a^{j+1}) \\ & \leq \lambda_s\rho_l\Delta(x^k)/(2+2\rho_l) \\ & \leq \frac{\rho_l}{1+\rho_l} \left( H_k(0) - \min_{z \in \mathbb{R}^n} H_k(z) \right), \end{aligned} \quad (60)$$

where the last inequality is from (53). From (60) we have

$$\begin{aligned} & H_k(z_k(\lambda_a^{j+1})) - D_k(\lambda_a^{j+1}) \\ & \leq \rho_l \left( -H_k(z_k(\lambda_a^{j+1})) + D_k(\lambda_a^{j+1}) + H_k(0) - \min_{z \in \mathbb{R}^n} H_k(z) \right) \\ & \leq \rho_l (H_k(0) - H_k(z_k(\lambda_a^{j+1}))), \end{aligned}$$

where the last inequality is due to (21), and this gives (23). Therefore, (23) should have already been satisfied when (58) holds. This proves (56) in Lemma 15(a).

(b). By choosing

$$j = \left\lceil \max \left\{ C', \left( \frac{4C''t^2m\|B_k\|_2^2(\rho_h + 1)}{M^2\rho_h\Delta^2(x^k)} \right) \right\} \right\rceil - 1, \quad (61)$$

we have

$$\frac{C''tm\|B_k\|_2^2}{j+2} \leq M^2\rho_h\Delta^2(x^k)/(4t(1+\rho_h)), \quad (62)$$

which, together with (45), yields that

$$\begin{aligned} & H_k(z_k(\lambda_a^{j+1})) - D_k(\lambda_a^{j+1}) \\ & \leq M^2\rho_h\Delta^2(x^k)/(4t(1+\rho_h)) \\ & \leq \rho_h\|x^k - S_t(x^k)\|_2^2/(4t(1+\rho_h)), \end{aligned} \quad (63)$$

where the last inequality is from (52). Denote  $z_{k\star} = \arg \min_{z \in \mathbb{R}^n} H_k(z)$ , which indicates that  $\|z_{k\star}\|_2 = \|x^k - S_t(x^k)\|_2$ . We have

$$\frac{1}{2t}\|z_k(\lambda_a^{j+1}) - z_{k\star}\|_2^2 \leq H_k(z_k(\lambda_a^{j+1})) - D_k(\lambda_a^{j+1}),$$

which follows from the fact that  $H_k(\cdot)$  is  $\frac{1}{t}$ -strongly convex and (21). From (63), we have  $\|z_k(\lambda_a^{j+1}) - z_{k\star}\|_2^2 \leq \rho_h/(2(1+\rho_h))\|z_{k\star}\|_2^2$ . So, by the Cauchy-Schwarz inequality, we have

$$\begin{aligned} \frac{\rho_h}{2t}\|z_k(\lambda_a^{j+1})\|_2^2 & \geq \frac{\rho_h}{4t}\|z_{k\star}\|_2^2 - \frac{\rho_h}{2t}\|z_k(\lambda_a^{j+1}) - z_{k\star}\|_2^2 \\ & \geq \rho_h\|z_{k\star}\|_2^2/(4t(1+\rho_h)). \end{aligned}$$

Together with (63), (24) should have already been satisfied when (61) holds. This proves (57) in Lemma 15(b).

Therefore, the proof of Lemma 15 is complete.  $\square$

Next, we are ready to present the proof of Theorem 3.

*Proof of Theorem 3.* We first prove part (a) of Theorem 3 and then prove part (b) of Theorem 3 in what follows.

(a). Lemma 8 and (47) indicate that  $\sup_{k \in \mathbb{N}} \|B_k\|_2 \leq B(\lambda_s/(2L))$ . Together with (46), (56), (47), we obtain

$$J_k \leq M_1 B^2(\lambda_s/(2L))/\Delta(x^k). \quad (64)$$

We now prove the desired result. If  $\epsilon \geq \Delta(x^0)$ , then  $J(\epsilon) = 0$  and (25) holds. If  $\epsilon < \Delta(x^0)$ , we denote  $K_\epsilon$  as the smallest index such that  $\Delta(x^k) > \epsilon, \forall k \leq K_\epsilon$  and  $\Delta(x^{K_\epsilon+1}) \leq \epsilon$ . From Lemma 6 we have

$$\begin{aligned} & L(\|x_\star\|_2 + \lambda_s/(2L))\Delta(x^k) \\ & \geq F(x^k) - F(x_\star) \\ & \geq (F(x^{K_\epsilon}) - F(x_\star)) \left( \frac{2 + 4\rho_l}{1 + 4\rho_l} \right)^{K_\epsilon - k} \\ & \geq \lambda_s \Delta(x^{K_\epsilon}) \left( \frac{2 + 4\rho_l}{1 + 4\rho_l} \right)^{K_\epsilon - k} \\ & \geq \epsilon \lambda_s \left( \frac{2 + 4\rho_l}{1 + 4\rho_l} \right)^{K_\epsilon - k}, \quad \forall 0 \leq k \leq K_\epsilon. \end{aligned} \quad (65)$$

Here we explain how these inequalities were obtained. The first inequality follows from Lemma 6 and (47). Since (46) holds, (40) holds for any  $k \in \mathbb{N}$  and it implies the second inequality in (65). The third inequality in (65) follows from Assumption 1 and the last inequality in (65) follows from the fact that  $\Delta(x^{K_\epsilon}) > \epsilon$ .

From (65) we know that  $\forall 0 \leq k \leq K_\epsilon$ , it holds that

$$\Delta(x^k) \geq \frac{\epsilon \lambda_s}{L(\|x_\star\|_2 + \lambda_s/(2L))} \left( \frac{2 + 4\rho_l}{1 + 4\rho_l} \right)^{K_\epsilon - k}. \quad (66)$$

Therefore, we have

$$\begin{aligned} J(\epsilon) &= \sum_{k=0}^{K_\epsilon} J_k \\ &\leq \sum_{k=0}^{K_\epsilon} M_1 B^2(\lambda_s/(2L))/\Delta(x^k) \\ &\leq \sum_{k=0}^{K_\epsilon} \frac{M_1 B^2(\frac{\lambda_s}{2L}) L(\|x_\star\|_2 + \lambda_s/(2L))}{\lambda_s \epsilon} \left( \frac{2 + 4\rho_l}{1 + 4\rho_l} \right)^{k - K_\epsilon} \\ &= \frac{E_2}{\epsilon(2 + 4\rho)} \sum_{k=0}^{K_\epsilon} \left( \frac{2 + 4\rho_l}{1 + 4\rho_l} \right)^{k - K_\epsilon} \leq \frac{E_2}{\epsilon}. \end{aligned}$$

where the first inequality follows from (64) and the second inequality follows from (66). This completes the proof of Theorem 3(a).

(b). Lemma 8 and (50) indicates that  $\sup_{k \in \mathbb{N}} \|B_k\|_2 \leq B(E_0/2)$ . Together with (46), (49), (57), ,

$$J_k \leq M_2 B^2(E_0/2)/\Delta^2(x^k), \quad \forall k \geq 0, \quad (67)$$

which follows from our choice of  $E_3 \leq B(E_0/2)\sqrt{M_2/C'}$  such that  $M_2 B^2(E_0/2)/\Delta^2(x^k) \geq M_2 B^2(E_0/2)/\Delta^2(x^0) \geq C'$ . Next, we prove the conclusion. We adopt the same definition of  $K_\epsilon$  as in (a) and pick  $\{\epsilon_i\}_{i=1}^\infty = \{\Delta(x^i)\}_{i=1}^\infty$  so that  $K_{\epsilon_i} = i - 1$ . From (51) we have

$$\Delta(x^{i-1}) \geq \sqrt{E_0 \epsilon_i}, \quad \forall i \in \mathbb{N}. \quad (68)$$

From (49) and (68) we have

$$\Delta(x^k) \geq 2^{i-1-k} \sqrt{E_0 \epsilon_i}, \forall 0 \leq k \leq i-1. \quad (69)$$

Finally, for any  $i \in \mathbb{N}$ , we have

$$\begin{aligned} J(\epsilon_i) &= \sum_{k=0}^{i-1} J_k \leq \sum_{k=0}^{i-1} \frac{M_2 B^2 (E_0/2)}{4^{i-1-k} E_0 \epsilon_i} \\ &\leq \frac{4M_2 B^2 (E_0/2)}{3E_0 \epsilon_i} = \frac{E_4}{\epsilon_i}. \end{aligned}$$

Here, the first inequality follows from (67) and (69), and the second inequality follows from the fact that

$$\sum_{k=0}^{i-1} \frac{1}{4^{i-1-k}} \leq \sum_{k=-\infty}^{i-1} \frac{1}{4^{i-1-k}} = \frac{4}{3}.$$

This completes the proof of Theorem 3(b).

Therefore, the proof of Theorem 3 is complete.  $\square$

## 6 Conclusion

In this paper, we proposed a new inexact proximal linear algorithm for solving the robust phase retrieval problem. Our contribution lies in the two adaptive stopping criteria for the subproblem in the proximal linear algorithm. We showed that the iteration complexity of our inexact proximal linear algorithm is in the same order as the exact proximal linear algorithm. Under the sharpness condition, we are able to prove the local convergence of the proposed method. Moreover, we discussed how to use the FISTA for solving the subproblem in our inexact proximal linear algorithm, and analyzed the total oracle complexity for obtaining an  $\epsilon$ -optimal solution under the sharpness condition. Numerical results demonstrated the superior performance of the proposed methods over some existing methods.

## References

- [1] J. Miao, P. Charalambous, J. Kirz, and D. Sayre, "Extending the methodology of x-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens," *Nature*, vol. 400, no. 6742, pp. 342–344, 1999.
- [2] R. P. Millane, "Phase retrieval in crystallography and optics," *Journal of the Optical Society of America A*, vol. 7, no. 3, pp. 394–411, 1990.
- [3] A. Chai, M. Moscoso, and G. Papanicolaou, "Array imaging using intensity-only measurements," *Inverse Problems*, vol. 27, no. 1, p. 015005, 2010.
- [4] C. Fienup and J. Dainty, "Phase retrieval and image reconstruction for astronomy," *Image Recovery: Theory and Application*, vol. 231, p. 275, 1987.
- [5] J. Miao, T. Ishikawa, Q. Shen, and T. Earnest, "Extending x-ray crystallography to allow the imaging of noncrystalline materials, cells, and single protein complexes," *Annual Review of Physical Chemistry*, vol. 59, pp. 387–410, 2008.
- [6] M. Fickus, D. G. Mixon, A. A. Nelson, and Y. Wang, "Phase retrieval from very few measurements," *Linear Algebra and its Applications*, vol. 449, pp. 475–499, 2014.
- [7] E. J. Candes, X. Li, and M. Soltanolkotabi, "Phase retrieval via wirtinger flow: Theory and algorithms," *IEEE Transactions on Information Theory*, vol. 61, no. 4, pp. 1985–2007, 2015.
- [8] Y. Chen and E. J. Candès, "Solving random quadratic systems of equations is nearly as easy as solving linear systems," *Communications on Pure and Applied Mathematics*, vol. 70, no. 5, pp. 822–883, 2017.

- [9] G. Wang, G. B. Giannakis, and Y. C. Eldar, “Solving systems of random quadratic equations via truncated amplitude flow,” *IEEE Transactions on Information Theory*, vol. 64, no. 2, pp. 773–794, 2017.
- [10] H. Zhang, Y. Zhou, Y. Liang, and Y. Chi, “A nonconvex approach for phase retrieval: Reshaped wirtinger flow and incremental algorithms,” *Journal of Machine Learning Research*, vol. 18, 2017.
- [11] J. C. Duchi and F. Ruan, “Solving (most) of a set of quadratic equalities: Composite optimization for robust phase retrieval,” *Information and Inference: A Journal of the IMA*, vol. 8, no. 3, pp. 471–529, 2019.
- [12] H. Zhang, Y. Chi, and Y. Liang, “Provable non-convex phase retrieval with outliers: Median truncated wirtinger flow,” in *International Conference on Machine Learning*. PMLR, 2016, pp. 1022–1031.
- [13] D. Davis, D. Drusvyatskiy, K. J. MacPhee, and C. Paquette, “Subgradient methods for sharp weakly convex functions,” *Journal of Optimization Theory and Applications*, vol. 179, no. 3, pp. 962–982, 2018.
- [14] A. S. Lewis and S. J. Wright, “A proximal method for composite minimization,” *Mathematical Programming*, vol. 158, no. 1, pp. 501–546, 2016.
- [15] D. Drusvyatskiy and C. Paquette, “Efficiency of minimizing compositions of convex functions and smooth maps,” *Mathematical Programming*, vol. 178, no. 1, pp. 503–558, 2019.
- [16] J. C. Duchi and F. Ruan, “Stochastic methods for composite and weakly convex optimization problems,” *SIAM Journal on Optimization*, vol. 28, no. 4, pp. 3229–3259, 2018.
- [17] V. Charisopoulos, Y. Chen, D. Davis, M. Díaz, L. Ding, and D. Drusvyatskiy, “Low-rank matrix recovery with composite optimization: good conditioning and rapid convergence,” *Foundations of Computational Mathematics*, pp. 1–89, 2021.
- [18] V. Charisopoulos, D. Davis, M. Díaz, and D. Drusvyatskiy, “Composite optimization for robust blind deconvolution,” *arXiv preprint arXiv:1901.01624*, 2019.
- [19] Z. Wang, B. Liu, S. Chen, S. Ma, L. Xue, and H. Zhao, “A manifold proximal linear method for sparse spectral clustering with application to single-cell RNA sequencing data analysis,” *INFORMS Journal on Optimization*, vol. 4, no. 2, pp. 200–214, 2022.
- [20] Y. Nesterov, “On an approach to the construction of optimal methods of minimization of smooth convex functions,” *Ekonomika i Mateaticheskie Metody*, vol. 24, no. 3, pp. 509–517, 1988.
- [21] A. Beck and M. Teboulle, “A fast iterative shrinkage-thresholding algorithm for linear inverse problems,” *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.
- [22] Y. Nesterov, “Gradient methods for minimizing composite functions,” *Mathematical Programming*, vol. 140, no. 1, pp. 125–161, 2013.
- [23] M. Díaz and B. Grimmer, “Optimal convergence rates for the proximal bundle method,” *SIAM Journal on Optimization*, vol. 33, no. 2, pp. 424–454, 2023.
- [24] S. Bonettini, I. Loris, F. Porta, and M. Prato, “Variable metric inexact line-search-based methods for nonsmooth optimization,” *SIAM Journal on Optimization*, vol. 26, no. 2, pp. 891–921, 2016.
- [25] C.-p. Lee and S. J. Wright, “Inexact successive quadratic approximation for regularized optimization,” *Computational Optimization and Applications*, vol. 72, no. 3, pp. 641–674, 2019.
- [26] —, “Inexact variable metric stochastic block-coordinate descent for regularized optimization,” *Journal of Optimization Theory and Applications*, vol. 185, no. 1, pp. 151–187, 2020.
- [27] N. Parikh and S. Boyd, “Block splitting for distributed optimization,” *Mathematical Programming Computation*, vol. 6, no. 1, pp. 77–102, 2014.
- [28] C. Dünnér, S. Forte, M. Takác, and M. Jaggi, “Primal-dual rates and certificates,” in *International Conference on Machine Learning*. PMLR, 2016, pp. 783–792.