

On the Quadratic Convergence of the Cubic Regularization Method under a Local Error Bound Condition

Man-Chung Yue *

Zirui Zhou[†]

Anthony Man-Cho So [‡]

Abstract

In this paper we consider the cubic regularization (CR) method, a regularized version of the classical Newton’s method, for minimizing a twice continuously differentiable function. While it is well-known that the CR method is globally convergent and enjoys a superior global iteration complexity, existing results on its local quadratic convergence require a stringent non-degeneracy condition. We prove that under a local error bound (EB) condition, which is much weaker a requirement than the existing non-degeneracy condition, the sequence of iterates generated by the CR method converges at least Q-quadratically to a second-order critical point. This indicates that adding cubic regularization not only equips Newton’s method with remarkable global convergence properties but also enables it to converge quadratically even in the presence of degenerate solutions. As a byproduct, we show that without assuming convexity, the proposed EB condition is equivalent to a quadratic growth condition, which could be of independent interest. To demonstrate the usefulness and relevance of our convergence analysis, we focus on two concrete nonconvex optimization problems that arise in phase retrieval and low-rank matrix recovery and prove that with overwhelming probability, the sequence of iterates generated by the CR method for solving these two problems converges at least Q-quadratically to a global minimizer. We also present numerical results of the CR method when applied to solve these two problems to support and complement our theoretical development.

Keywords: cubic regularization, local quadratic convergence, error bound, second-order critical points, non-isolated solutions, phase retrieval, low-rank matrix recovery

AMS subject classifications: 90C26, 90C30, 65K05, 49M15

1 Introduction

Consider the unconstrained minimization problem

$$\min_{x \in \mathbb{R}^n} f(x), \tag{1}$$

where $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is assumed to be twice continuously differentiable. Newton’s method is widely regarded as an efficient local method for solving problem (1). The cubic regularization (CR) method is a regularized version of Newton’s method that is globally convergent. Roughly speaking, given

*Imperial College Business School, Imperial College London, United Kingdom. E-mail: m.yue@imperial.ac.uk

[†]Department of Mathematics, Simon Fraser University, Canada. E-mail: ziruiz@sfu.ca. This author is supported by an NSERC Discovery Grant and the SFU Alan Mekler postdoctoral fellowship.

[‡]Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, N. T., Hong Kong. E-mail: mancho@se.cuhk.edu.hk

the current iterate x^k , the CR method determines the next one by minimizing a cubic regularized quadratic model of f at x^k ; *i.e.*,

$$x^{k+1} \in \underset{x \in \mathbb{R}^n}{\text{Argmin}} \left\{ f(x^k) + \nabla f(x^k)^T(x - x^k) + \frac{1}{2}(x - x^k)^T \nabla^2 f(x^k)(x - x^k) + \frac{\sigma_k}{6} \|x - x^k\|^3 \right\}, \quad (2)$$

where the regularization parameter $\sigma_k > 0$ is chosen such that $f(x^{k+1}) \leq f(x^k)$. The idea of using cubic regularization first appeared in Griewank [17], where he proved that any accumulation point of $\{x^k\}_{k \geq 0}$ generated by the CR method is a second-order critical point of f ; *i.e.*, an $x \in \mathbb{R}^n$ satisfying $\nabla f(x) = 0$ and $\nabla^2 f(x) \succeq 0$. Later, Nesterov and Polyak [27] presented the remarkable result that the CR method has a better global iteration complexity bound than that for the steepest descent method. Elaborating on these results, Cartis *et al.* [11, 12] proposed an adaptive CR method for solving problem (1), where $\{\sigma_k\}_{k \geq 0}$ are determined dynamically and subproblems (2) are solved inexactly. They showed that the proposed method can still preserve the good global complexity bound established in [27]. Based on these pioneering works, the CR method has been attracting increasing attention over the past decade; see, *e.g.*, [10, 32, 35] and references therein.

In addition to these global convergence properties, the CR method, as a modified Newton's method, is also expected to attain a fast local convergence rate. It is known that if any accumulation point \bar{x} of the sequence $\{x^k\}_{k \geq 0}$ generated by the CR method satisfies

$$\nabla f(\bar{x}) = 0 \quad \text{and} \quad \nabla^2 f(\bar{x}) \succ 0, \quad (3)$$

then the whole sequence $\{x^k\}_{k \geq 0}$ converges at least Q-quadratically¹ to \bar{x} ; see [17, Theorem 4.1] or [27, Theorem 3]. Nevertheless, the non-degeneracy condition (3) implies that \bar{x} is an isolated local minimizer of f and hence does not hold for many nonconvex functions in real-world applications. For example, consider the problem of recovering a positive semidefinite matrix $X^* \in \mathbb{R}^{n \times n}$ with rank $r \ll n$, given a linear operator $\mathcal{A} : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^m$ and a measurement vector $b = \mathcal{A}(X^*)$. A practically efficient approach for recovering X^* is to solve the following nonconvex minimization problem (see, *e.g.*, [4]):

$$\min_{U \in \mathbb{R}^{n \times r}} f(U) := \frac{1}{4m} \|\mathcal{A}(UU^T) - b\|^2.$$

Noticing that $f(U) = f(UR)$ for any $U \in \mathbb{R}^{n \times r}$ and any orthogonal matrix $R \in \mathbb{R}^{r \times r}$, it is not hard to see that there is no isolated local minimizer of f when $r \geq 2$, which implies that there is no $U \in \mathbb{R}^{n \times r}$ such that $\nabla f(U) = 0$ and $\nabla^2 f(U) \succ 0$ when $r \geq 2$. Similar degeneracy features can also be found in various nonconvex optimization formulations used in phase retrieval [31] and deep learning [37]. In view of this, it is natural to study the local convergence properties of the CR method for solving problems with non-isolated minimizers. Moreover, the non-degeneracy condition (3) seems too stringent for the purpose of ensuring quadratic convergence of the CR method. Indeed, one can observe from (2) that due to the cubic regularization, the CR method is well defined even when the Hessian at hand has non-positive eigenvalues. In addition, the CR method belongs to the class of regularized Newton-type methods, many of which have been shown to attain a superlinear or quadratic convergence rate even in the presence of non-isolated solutions. For instance, Li *et al.* [20] considered a regularized Newton's method for solving the convex case of problem (1). They proved that if f satisfies a local error bound condition, which is a weaker requirement than (3), then the whole sequence $\{x^k\}_{k \geq 0}$ converges superlinearly or quadratically to an optimal solution. Yue *et al.* [36] extended such result to a regularized proximal Newton's method for solving a class of nonsmooth convex minimization problems. Other regularized Newton-type methods that have been shown to

¹A sequence of vectors $\{w^k\}_{k \geq 0}$ in \mathbb{R}^n is said to converge *Q-superlinearly* (*Q-quadratically*) of order $q \in (1, 2)$ ($q = 2$) to a vector w^∞ if there exists a positive constant M such that $\|w^k - w^\infty\|/\|w^{k-1} - w^\infty\|^q \leq M$ for all sufficiently large k ; see, *e.g.*, [28, Appendix A.2]. It is worth noting that Q-superlinear (Q-quadratic) convergence is a local notion in the sense that it concerns the convergence properties of $\{w^k\}_{k \geq 0}$ for all sufficiently large k .

attain superlinear or quadratic convergence for problems with non-isolated solutions include, among others, the classic Levenberg-Marquardt (LM) method [34, 15] for nonlinear equations, Newton-type methods for complementarity problems [33], and regularized Gauss-Newton methods for nonlinear least-squares [2].

In this paper we establish the quadratic convergence of the CR method under the assumption of the following local error bound condition.

Definition 1 (EB Condition). *We say that f satisfies the local error bound (EB) condition if there exist scalars $\kappa, \rho > 0$ such that*

$$\text{dist}(x, \mathcal{X}) \leq \kappa \|\nabla f(x)\| \quad \text{whenever} \quad \text{dist}(x, \mathcal{X}) \leq \rho,$$

where $\mathcal{X} = \{x \in \mathbb{R}^n : \nabla f(x) = 0, \nabla^2 f(x) \succeq 0\}$ is the set of second-order critical points of f and $\text{dist}(x, \mathcal{X})$ denotes the distance of x to \mathcal{X} .

As we shall see in Section 3, the above local EB condition is a weaker requirement than the non-degeneracy condition (3). We prove that if f satisfies the above local EB condition, then the whole sequence $\{x^k\}_{k \geq 0}$ generated by the CR method converges at least Q-quadratically to a second-order critical point of f . This, together with the pioneering works [17, 27, 11], indicates that adding cubic regularization not only equips Newton's method with superior global convergence properties but also enables it to converge quadratically even in the presence of degenerate solutions. We remark that our proof of quadratic convergence is not a direct extension of those from the aforementioned works on regularized Newton-type methods. In particular, a major difficulty in our proof is that the descent direction $d^k = x^{k+1} - x^k$ of the CR method is obtained by minimizing a nonconvex function, as one can see from (2). By contrast, the descent directions of the regularized Newton-type methods in [20, 36, 34, 15, 2] are all obtained by minimizing a strongly convex function. For instance, the LM method for solving the nonlinear equation $F(x) = 0$ computes its descent direction by solving the strongly convex optimization problem

$$d^k = \underset{d \in \mathbb{R}^n}{\operatorname{argmin}} \left\{ \|F(x^k) + F'(x^k)d\|^2 + \mu_k \|d\|^2 \right\},$$

where F' is the Jacobian of F and $\mu_k > 0$ is the regularization parameter; see [19, 25]. Consequently, we cannot utilize the nice properties of strongly convex functions in our proof. Instead, we shall exploit the fact that any accumulation point of the sequence $\{x^k\}_{k \geq 0}$ generated by the CR method is a second-order critical point of f in our analysis. It is also worth noting that our convergence analysis unifies and sharpens those in [27] for the so-called globally non-degenerate star-convex functions and gradient-dominated functions (see Section 2 for the definitions). In particular, we show that when applied to these two classes of functions, the CR method converges quadratically, which improves upon the sub-quadratic convergence rates established in [27].

Besides our convergence analysis of the CR method, the proposed local EB condition could also be of independent interest. A notable feature of the EB condition in Definition 1 is that its target set \mathcal{X} is the set of second-order critical points of f . This contrasts with other EB conditions in the literature, where \mathcal{X} is typically the set of first-order critical points (see, *e.g.*, [24]) or the set of optimal solutions (see, *e.g.*, [16, 38]). Such feature makes our EB condition especially useful for analyzing local convergence of iterative algorithms that are guaranteed to cluster at second-order critical points. Moreover, we prove that under some mild assumptions, our local EB condition is equivalent to a quadratic growth condition (see Theorem 1 (ii) for the definition). Prior to this work, the equivalence between these two regularity conditions was established when f is convex [1] or when f is nonconvex but satisfies a certain quadratic lower-bound condition [14]. Our result indicates that if the target set \mathcal{X} is the set of second-order critical points, the quadratic lower-bound condition in [14] holds automatically and hence a complete equivalence of the two regularity conditions can be established.

To demonstrate the usefulness and relevance of our convergence analysis, we apply it to study the local convergence behavior of the CR method when applied to minimize two concrete nonconvex functions that arise in phase retrieval and low-rank matrix recovery. A common feature of these nonconvex functions is that they do not have isolated local minimizers. Motivated by recent advances in probabilistic analysis of the global geometry of these nonconvex functions [31, 4], we show that with overwhelming probability, (i) the set of second-order critical points equals the set of global minimizers and (ii) the local EB condition in Definition 1 holds. As a result, our analysis implies that with overwhelming probability, the sequence of iterates generated by the CR method for solving these nonconvex problems converges at least Q-quadratically to a global minimizer. Numerical results of the CR method for solving these two nonconvex problems are also presented, which corroborate our theoretical findings.

The rest of this paper is organized as follows. In Section 2, we review existing results on the convergence properties of the CR method. In Section 3, we study the local EB condition and prove its equivalence to a quadratic growth condition. In Section 4, we prove the quadratic convergence of the CR method under the local EB condition. In Section 5, we study the CR method for solving two concrete nonconvex minimization problems that arise in phase retrieval and low-rank matrix recovery. In Section 6, we present numerical results of the CR method for solving these two nonconvex problems. Finally, we close with some concluding remarks in Section 7.

1.1 Notations

We adopt the following notations throughout the paper. Let \mathbb{R}^n be the n -dimensional Euclidean space and $\langle \cdot, \cdot \rangle$ be its standard inner product. For any vector $x \in \mathbb{R}^n$, we denote by $\|x\| = \sqrt{\langle x, x \rangle}$ its Euclidean norm. Given any $\bar{x} \in \mathbb{R}^n$ and $\rho > 0$, we denote by $\mathbb{B}(\bar{x}; \rho)$ the Euclidean ball with center \bar{x} and radius ρ ; *i.e.*, $\mathbb{B}(\bar{x}; \rho) := \{x \in \mathbb{R}^n : \|x - \bar{x}\| \leq \rho\}$. For any matrix $X \in \mathbb{R}^{m \times n}$, we denote by $\|X\|$ and $\|X\|_F$ its operator norm and Frobenius norm, respectively. If in addition X is symmetric, we write $\lambda_1(X) \geq \dots \geq \lambda_n(X)$ as the eigenvalues of X in decreasing order. Moreover, we write $X \succeq 0$ if X is positive semidefinite. We denote by \mathcal{O}^r the set of $r \times r$ orthogonal matrices; *i.e.*, $Q^T Q = Q Q^T = I_r$ for any $Q \in \mathcal{O}^r$, where I_r is the $r \times r$ identity matrix. For any complex vector $z \in \mathbb{C}^n$, we denote by $\Re(z)$ and $\Im(z)$ its real and imaginary parts, respectively. Moreover, we let \bar{z} be the conjugate of z , $z^H = \bar{z}^T$ be the Hermitian transpose of z , and $\|z\| = \sqrt{z^H z}$ be the norm of z . For any closed subset $C \subseteq \mathbb{R}^n$, we denote by $\text{dist}(x, C)$ the distance of $x \in \mathbb{R}^n$ to C . In addition, we use $\mathcal{N}(C; \rho)$ with some $\rho > 0$ to denote the neighborhood $\mathcal{N}(C; \rho) := \{x \in \mathbb{R}^n : \text{dist}(x, C) \leq \rho\}$ of C .

For any $x \in \mathbb{R}^n$, we define $\mathcal{L}(f(x)) := \{y \in \mathbb{R}^n : f(y) \leq f(x)\}$. We say that $x \in \mathbb{R}^n$ is a *second-order critical point* of f if it satisfies the second-order necessary condition for f ; *i.e.*, $\nabla f(x) = 0$ and $\nabla^2 f(x) \succeq 0$. Unless otherwise stated, we use \mathcal{X} to denote the set of second-order critical points of f and \mathcal{X}^* to denote the set of global minimizers of f . It is clear that $\mathcal{X}^* \subseteq \mathcal{X}$. Moreover, since f is twice continuously differentiable, both \mathcal{X} and \mathcal{X}^* are closed subsets of \mathbb{R}^n . We assume throughout the paper that \mathcal{X}^* is non-empty and the minimum value of f on \mathbb{R}^n , denoted by f^* , is lower bounded.

2 The Cubic Regularization Method

In this section, we review the cubic regularization (CR) method for solving problem (1) and some existing results on its convergence properties.

Given a vector $x \in \mathbb{R}^n$, we define the cubic regularized quadratic approximation of f at x as

$$f_\sigma(p; x) = f(x) + \nabla f(x)^T(p - x) + \frac{1}{2}(p - x)^T \nabla^2 f(x)(p - x) + \frac{\sigma}{6}\|p - x\|^3, \quad (4)$$

where $\sigma > 0$ is the regularization parameter. In addition, we define

$$\bar{f}_\sigma(x) := \min_{p \in \mathbb{R}^n} f_\sigma(p; x) \quad \text{and} \quad p_\sigma(x) \in \text{Argmin}_{p \in \mathbb{R}^n} f_\sigma(p; x). \quad (5)$$

In principle, starting with an initial point $x^0 \in \mathbb{R}^n$, the CR method generates a sequence of iterates $\{x^k\}_{k \geq 0}$ by letting $x^{k+1} = p_{\sigma_k}(x^k)$ for some $\sigma_k > 0$ such that

$$f(p_{\sigma_k}(x^k)) \leq \bar{f}_{\sigma_k}(x^k). \quad (6)$$

Notice that this requires the computation of $p_\sigma(x)$, which is a global minimizer of $f_\sigma(\cdot; x)$. Although $f_\sigma(\cdot; x)$ is in general nonconvex, it has been shown in [27] that $p_\sigma(x)$ can be obtained from the solution of a one-dimensional convex optimization problem, thus making the computation of $p_\sigma(x)$ viable from the theoretical point of view. Moreover, the optimality condition for the global minimizers of $f_\sigma(\cdot; x)$ is very similar to that of a standard trust-region subproblem [11, Theorem 3.1]. Such observation has led to the development of various efficient algorithms for finding $p_\sigma(x)$ in [11]. More recently, it is shown in [9] that the gradient descent method can also be applied to find $p_\sigma(x)$.

For the global convergence of the CR method, we need the following assumption.

Assumption 1. *The Hessian of the function f is Lipschitz continuous on a closed convex set \mathcal{F} with $\mathcal{L}(f(x^0)) \subseteq \text{int}(\mathcal{F})$; i.e., there exists a constant $L > 0$ such that*

$$\|\nabla^2 f(x) - \nabla^2 f(y)\| \leq L\|x - y\|, \quad \forall x, y \in \mathcal{F}. \quad (7)$$

A direct consequence of Assumption 1 is that for any $x \in \mathcal{F}$, it holds that $f(p_\sigma(x)) \leq \bar{f}_\sigma(x)$ whenever $\sigma \geq L$ (see [27, Lemma 4]). This further implies that for all $k \geq 0$, we can find a $\sigma_k \leq 2L$ such that (6) holds. Indeed, if the Lipschitz constant L is known, we can let $\sigma_k = L$. If not, by using a line search strategy that starts from $\sigma_k = \bar{\sigma}$ for some $\bar{\sigma} \in (0, L]$ and doubles σ_k after each trial [27, Section 5.2], we can find a $\sigma_k \in [\bar{\sigma}, 2L]$ such that (6) holds. We now state the details of the CR method as follows.

Algorithm 1 (The Cubic Regularization Method).

0. Input an initial point $x^0 \in \mathbb{R}^n$, a scalar $\bar{\sigma} \in (0, L]$, and set $k = 0$.

1. Find $\sigma_k \in [\bar{\sigma}, 2L]$ such that

$$f(p_{\sigma_k}(x^k)) \leq \bar{f}_{\sigma_k}(x^k). \quad (8)$$

2. Set $x^{k+1} = p_{\sigma_k}(x^k)$ and $k = k + 1$, and go to Step 1.

End.

The following result, which can be found in [17, Theorem 4.1] and [27, Theorem 2], shows that any accumulation point of the sequence $\{x^k\}_{k \geq 0}$ generated by the CR method is a second-order critical point of f .

Fact 1. *Suppose that Assumption 1 holds. Let $\{x^k\}_{k \geq 0}$ be the sequence of iterates generated by the CR method. If $\mathcal{L}(f(x^k))$ is bounded for some $k \geq 0$, then the following statements hold.*

(i) $v := \lim_{k \rightarrow \infty} f(x^k)$ exists.

(ii) $\lim_{k \rightarrow \infty} \|x^k - x^{k-1}\| = 0$.

(iii) *The sequence $\{x^k\}_{k \geq 0}$ has at least one accumulation point. Moreover, every accumulation point \bar{x} of $\{x^k\}_{k \geq 0}$ satisfies*

$$f(\bar{x}) = v, \quad \nabla f(\bar{x}) = 0, \quad \nabla^2 f(\bar{x}) \succeq 0.$$

We next review some existing results on the local convergence rate of the CR method. We start with the following result, which can be found in [17, Theorem 4.1].

Fact 2. Suppose that Assumption 1 holds. Let $\{x^k\}_{k \geq 0}$ be the sequence generated by Algorithm 1 for solving problem (1). If an accumulation point \bar{x} of $\{x^k\}_{k \geq 0}$ satisfies

$$\nabla f(\bar{x}) = 0, \quad \nabla^2 f(\bar{x}) \succ 0, \quad (9)$$

then the whole sequence $\{x_k\}_{k \geq 0}$ converges at least Q -quadratically to \bar{x} .

As discussed in the Introduction, the non-degeneracy condition (9) implies that \bar{x} is an isolated local minimizer of f , which does not hold in many applications. In an attempt to overcome such a limitation, Nesterov and Polyak [27] considered two classes of functions for which there can be non-isolated second-order critical points and showed that Algorithm 1 converges superlinearly locally when applied to these functions. The first class is the so-called globally non-degenerate star-convex functions.

Definition 2. We say that f is star-convex if for any $x^* \in \mathcal{X}^*$,

$$f(\alpha x^* + (1 - \alpha)x) \leq \alpha f^* + (1 - \alpha)f(x), \quad \forall x \in \mathbb{R}^n, \forall \alpha \in [0, 1]. \quad (10)$$

Definition 3. We say that the optimal solution set \mathcal{X}^* of f is globally non-degenerate if there exists a scalar $\alpha > 0$ such that

$$f(x) - f^* \geq \frac{\alpha}{2} \cdot \text{dist}^2(x, \mathcal{X}^*), \quad \forall x \in \mathbb{R}^n. \quad (11)$$

Fact 3 ([27, Theorem 5]). Suppose that Assumption 1 holds, f is star-convex, and \mathcal{X}^* is globally non-degenerate. Then, there exist a scalar $\gamma > 0$ and an integer $k_0 \geq 0$ such that

$$f(x^{k+1}) - f^* \leq \gamma \left(f(x^k) - f^* \right)^{\frac{3}{2}}, \quad \forall k \geq k_0.$$

The second class of functions studied in [27] is the so-called gradient-dominated functions.

Definition 4. We say that f is gradient-dominated of degree 2 if there exists a scalar $\tau_f > 0$ such that

$$f(x) - f^* \leq \tau_f \|\nabla f(x)\|^2, \quad \forall x \in \mathbb{R}^n. \quad (12)$$

It is worth mentioning that the inequality (12) is an instance of the Łojasiewicz inequality, which has featured prominently in the convergence analysis of iterative methods; see, e.g., [21] and the references therein. Indeed, recall that f is said to satisfy the Łojasiewicz inequality with exponent $\theta \in [\frac{1}{2}, 1)$ at $\bar{x} \in \mathbb{R}^n$ if there exist a scalar $c > 0$ and a neighborhood \mathcal{U} of \bar{x} such that

$$|f(x) - f(\bar{x})|^\theta \leq c \|\nabla f(x)\|, \quad \forall x \in \mathcal{U}.$$

Hence, the inequality (12) is simply the Łojasiewicz inequality at any global minimizer of f with $\theta = \frac{1}{2}$ and $\mathcal{U} = \mathbb{R}^n$.

Fact 4 ([27, Theorem 7]). Suppose that Assumption 1 holds and f is gradient-dominated of degree 2. Then, there exist a scalar $\gamma > 0$ and an integer $k_0 \geq 0$ such that

$$f(x^{k+1}) - f^* \leq \gamma \left(f(x^k) - f^* \right)^{\frac{4}{3}}, \quad \forall k \geq k_0.$$

From the definitions, it is not hard to see that both globally non-degenerate star-convex functions and gradient-dominated functions can be nonconvex and can have non-isolated second-order critical points. Nevertheless, the convergence rates obtained in Facts 3 and 4 are weaker than that in Fact 2: only superlinear rates of order $\frac{3}{2}$ and $\frac{4}{3}$ are established for these two classes respectively, while a quadratic rate is achieved in Fact 2. As we shall see in Section 4, using our analysis approach, the superlinear convergence rates in Facts 3 and 4 can be improved to quadratic.

3 Error Bound for the Set of Second-Order Critical Points

Recall that \mathcal{X} is the set of second-order critical points of f , which is a closed subset of \mathbb{R}^n and assumed to be non-empty. In this section, we are interested in the local EB condition for \mathcal{X} , which we repeat here for the convenience of the readers.

Assumption 2 (EB Condition). *There exist scalars $\kappa, \rho > 0$ such that*

$$\text{dist}(x, \mathcal{X}) \leq \kappa \|\nabla f(x)\|, \quad \forall x \in \mathcal{N}(\mathcal{X}; \rho).$$

Assumption 2 is much weaker than the non-degeneracy assumption (9). Indeed, if $\bar{x} \in \mathcal{X}$ satisfies condition (9), then it is routine to show that \bar{x} is an isolated second-order critical point and there exist scalars $\kappa, \rho > 0$ such that $\text{dist}(x, \mathcal{X}) \leq \kappa \|\nabla f(x)\|$ whenever $\|x - \bar{x}\| \leq \rho$. On the other hand, Assumption 2 can still be satisfied when f has no isolated second-order critical points. For instance, it is not hard to verify that $f(x) = (\|x\|^2 - 1)^2$, whose set of second-order critical points is $\mathcal{X} = \{x : \|x\| = 1\}$, satisfies Assumption 2. Furthermore, at the end of this section we shall show that both the globally non-degenerate star-convex functions and the gradient-dominated functions considered in Facts 3 and 4 satisfy Assumption 2. In Section 5 we shall show that certain nonconvex functions that arise in phase retrieval and low-rank matrix recovery satisfy Assumption 2 with overwhelming probability.

In what follows, we prove that under some mild assumptions, the EB condition is equivalent to a quadratic growth condition. For any $x \in \mathbb{R}^n$, we denote by \hat{x} a projection² of x onto \mathcal{X} ; i.e., $\hat{x} \in \text{Argmin}_{z \in \mathcal{X}} \|x - z\|$.

Theorem 1. *Suppose that $\nabla^2 f(x)$ is uniformly continuous on $\mathcal{N}(\mathcal{X}; \gamma)$ for some $\gamma > 0$. Also, suppose that f satisfies the following separation property: there exists an $\epsilon > 0$ such that $\|x - y\| \geq \epsilon$ for any $x, y \in \mathcal{X}$ with $f(x) \neq f(y)$. Then, the following statements are equivalent.*

(i) *There exist scalars $\kappa, \rho > 0$ such that*

$$\text{dist}(x, \mathcal{X}) \leq \kappa \|\nabla f(x)\|, \quad \forall x \in \mathcal{N}(\mathcal{X}; \rho). \quad (13)$$

(ii) *There exist scalars $\alpha, \beta > 0$ such that*

$$f(x) \geq f(\hat{x}) + \frac{\alpha}{2} \cdot \text{dist}^2(x, \mathcal{X}), \quad \forall x \in \mathcal{N}(\mathcal{X}; \beta). \quad (14)$$

Before presenting the proof, some remarks on the assumptions in Theorem 1 are in order. First, the uniform continuity of $\nabla^2 f(x)$ on $\mathcal{N}(\mathcal{X}; \gamma)$ for some $\gamma > 0$ holds if \mathcal{X} is a compact set. Second, the separation property in Theorem 1 has appeared in [24], in which it was referred to as *proper separation of isocost surfaces*, and has long played a role in the study of error bounds. It holds for many nonconvex functions in applications and holds trivially if f is convex.

Proof of Theorem 1. We first prove (i) \Rightarrow (ii). Suppose that (13) holds with some $\kappa, \rho > 0$. Since $\nabla^2 f(x)$ is uniformly continuous on $\mathcal{N}(\mathcal{X}; \gamma)$, there exists a scalar $\beta_0 > 0$ such that

$$\|\nabla^2 f(x) - \nabla^2 f(y)\| \leq \frac{1}{4\kappa}, \quad \forall x, y \in \mathcal{N}(\mathcal{X}; \gamma) \text{ with } \|x - y\| \leq \beta_0. \quad (15)$$

Let $\beta_1 := \min\{\beta_0, \rho, \gamma\} > 0$, $x \in \mathcal{N}(\mathcal{X}; \beta_1)$ be arbitrarily chosen, and $x(t) = \hat{x} + t(x - \hat{x})$ for $t \in [0, 1]$. Thus, $\|x(t) - \hat{x}\| \leq \|x - \hat{x}\| \leq \beta_1$ for any $t \in [0, 1]$. By (15), we have

$$\|\nabla^2 f(x(t)) - \nabla^2 f(\hat{x})\| \leq \frac{1}{4\kappa}, \quad \forall t \in [0, 1].$$

²Since \mathcal{X} is in general nonconvex, the projection onto \mathcal{X} is not necessarily unique.

This, together with the inequality $|\lambda_{\min}(A) - \lambda_{\min}(B)| \leq \|A - B\|$ for any real symmetric matrices A and B (see, *e.g.*, [3, Corollary III.2.6]), yields

$$\lambda_{\min}[\nabla^2 f(x(t))] \geq \lambda_{\min}[\nabla^2 f(\hat{x})] - \frac{1}{4\kappa} \geq -\frac{1}{4\kappa}, \quad \forall t \in [0, 1], \quad (16)$$

where the last inequality is due to $\nabla^2 f(\hat{x}) \succeq 0$. By the integral form of Taylor's series, we have

$$f(x) - f(\hat{x}) = \langle \nabla f(\hat{x}), x - \hat{x} \rangle + \int_0^1 (1-t)(x - \hat{x})^T \nabla^2 f(x(t))(x - \hat{x}) dt.$$

This, together with (16), $\nabla f(\hat{x}) = 0$, and $\|x - \hat{x}\| = \text{dist}(x, \mathcal{X})$, yields

$$f(x) - f(\hat{x}) \geq -\frac{1}{8\kappa} \cdot \text{dist}^2(x, \mathcal{X}), \quad \forall x \in \mathcal{N}(\mathcal{X}; \beta_1). \quad (17)$$

Our next goal is to prove that there exists a scalar $\beta > 0$ such that

$$f(x) \geq f(\hat{x}) + \frac{1}{16\kappa} \cdot \text{dist}^2(x, \mathcal{X}), \quad \forall x \in \mathcal{N}(\mathcal{X}; \beta). \quad (18)$$

This would then imply that statement (ii) holds. Suppose that (18) does not hold for any $\beta > 0$. Then, there exist a sequence $\{x^k\}_{k \geq 0}$ and a sequence of positive scalars $\{t_k\}_{k \geq 0}$ such that $\lim_{k \rightarrow \infty} \text{dist}(x^k, \mathcal{X}) = 0$ and

$$f(x^k) \leq f(\hat{x}^k) + \frac{1}{16\kappa} \cdot \text{dist}^2(x^k, \mathcal{X}) - t_k, \quad \forall k \geq 0. \quad (19)$$

Without loss of generality, we assume that $x^k \in \mathcal{N}(\mathcal{X}; \beta_1)$ for all $k \geq 0$. By (19), we have $x^k \notin \mathcal{X}$ for all $k \geq 0$. Let $\lambda_k := \frac{1}{2} \cdot \text{dist}(x^k, \mathcal{X})$. Hence, $\lim_{k \rightarrow \infty} \lambda_k = 0$ and $\lambda_k > 0$ for all $k \geq 0$. Given any $k \geq 0$, consider the problem

$$\begin{aligned} v_k := \min \left\{ f(x) + \frac{1}{8\kappa} \cdot \text{dist}^2(x, \mathcal{X}) \right\} \\ \text{s.t. } x \in \mathcal{N}(\mathcal{X}; \beta_1) \cap \mathbb{B}\left(\hat{x}^k; \frac{\epsilon}{3}\right). \end{aligned} \quad (20)$$

Since \hat{x}^k is feasible for (20) and $\hat{x}^k \in \mathcal{X}$, we have $v_k \leq f(\hat{x}^k)$. Let x be an arbitrary feasible point of (20). Then, it follows from (17) that $f(x) + \frac{1}{8\kappa} \cdot \text{dist}^2(x, \mathcal{X}) \geq f(\hat{x})$. In addition, since $x \in \mathbb{B}\left(\hat{x}^k; \frac{\epsilon}{3}\right)$, we have $\|\hat{x} - \hat{x}^k\| \leq \|x - \hat{x}\| + \|x - \hat{x}^k\| \leq 2\|x - \hat{x}^k\| \leq \frac{2}{3}\epsilon < \epsilon$. This, together with the fact that $\hat{x}, \hat{x}^k \in \mathcal{X}$ and our assumption in Theorem 1, implies that $f(\hat{x}) = f(\hat{x}^k)$. Hence, every feasible point x of (20) satisfies $f(x) + \frac{1}{8\kappa} \cdot \text{dist}^2(x, \mathcal{X}) \geq f(\hat{x}^k)$, which implies that $v_k \geq f(\hat{x}^k)$. Thus, we can conclude that $v_k = f(\hat{x}^k)$. Combining this with (19), we obtain

$$f(x^k) + \frac{1}{8\kappa} \cdot \text{dist}^2(x^k, \mathcal{X}) \leq v_k + \tau_k, \quad \forall k \geq 0, \quad (21)$$

where $\tau_k = \frac{3}{16\kappa} \cdot \text{dist}^2(x^k, \mathcal{X}) - t_k$. Since $\lim_{k \rightarrow \infty} \text{dist}(x^k, \mathcal{X}) = 0$, there exists a $k_0 \geq 0$ such that x^k is feasible for (20) for any $k \geq k_0$. By this, (20), (21), and Ekeland's variational principle (see, *e.g.*, [26, Theorem 2.26]), there exists a sequence $\{z^k\}_{k \geq k_0}$ such that for all $k \geq k_0$, $\|x^k - z^k\| \leq \lambda_k$ and

$$\begin{aligned} z^k = \operatorname{argmin} \left\{ f(x) + \frac{1}{8\kappa} \cdot \text{dist}^2(x, \mathcal{X}) + \frac{\tau_k}{\lambda_k} \|x - z^k\| \right\} \\ \text{s.t. } x \in \mathcal{N}(\mathcal{X}; \beta_1) \cap \mathbb{B}\left(\hat{x}^k; \frac{\epsilon}{3}\right). \end{aligned} \quad (22)$$

Since $\lim_{k \rightarrow \infty} \lambda_k = 0$, we have $\lim_{k \rightarrow \infty} \|x^k - z^k\| = 0$. In addition, noticing that

$$\text{dist}(z^k, \mathcal{X}) \leq \|z^k - \hat{x}^k\| \leq \|z^k - x^k\| + \|x^k - \hat{x}^k\| = \|z^k - x^k\| + \text{dist}(x^k, \mathcal{X}),$$

we obtain $\lim_{k \rightarrow \infty} \|z^k - \hat{x}^k\| = \lim_{k \rightarrow \infty} \text{dist}(z^k, \mathcal{X}) = 0$. Hence, there exists a $k_1 \geq k_0$ such that z^k is in the interior of the feasible set of (22) for all $k \geq k_1$. Consequently, by the generalized Fermat's rule (see, *e.g.*, [29, Theorem 10.1]), we have

$$0 \in \partial \left(f(\cdot) + \frac{1}{8\kappa} \cdot \text{dist}^2(\cdot, \mathcal{X}) + \frac{\tau_k}{\lambda_k} \|\cdot - z^k\| \right) (z^k),^3 \quad \forall k \geq k_1. \quad (23)$$

Since f is continuously differentiable, we obtain from [26, Corollary 1.82] that $\partial f(z^k) = \{\nabla f(z^k)\}$. In addition, we have

$$\partial (\text{dist}^2(\cdot, \mathcal{X})) (z^k) = 2 \cdot \text{dist}(z^k, \mathcal{X}) \cdot \partial (\text{dist}(\cdot, \mathcal{X})) (z^k) \subseteq 2 \cdot \text{dist}(z^k, \mathcal{X}) \cdot \mathbb{B}(0; 1),$$

where the equality follows from [26, Corollary 1.111(i)] and the inclusion is due to [29, Example 8.53]. Also, we have $\partial (\|\cdot - z^k\|) (z^k) = \mathbb{B}(0; 1)$. These, together with (23), yield

$$\begin{aligned} 0 &\in \partial \left(f(\cdot) + \frac{1}{8\kappa} \cdot \text{dist}^2(\cdot, \mathcal{X}) + \frac{\tau_k}{\lambda_k} \|\cdot - z^k\| \right) (z^k) \\ &= \nabla f(z^k) + \partial \left(\frac{1}{8\kappa} \cdot \text{dist}^2(\cdot, \mathcal{X}) + \frac{\tau_k}{\lambda_k} \|\cdot - z^k\| \right) (z^k) \end{aligned} \quad (24)$$

$$\subseteq \nabla f(z^k) + \partial \left(\frac{1}{8\kappa} \cdot \text{dist}^2(\cdot, \mathcal{X}) \right) (z^k) + \partial \left(\frac{\tau_k}{\lambda_k} \|\cdot - z^k\| \right) (z^k) \quad (25)$$

$$\subseteq \nabla f(z^k) + \left(\frac{1}{4\kappa} \cdot \text{dist}(z^k, \mathcal{X}) + \frac{\tau_k}{\lambda_k} \right) \mathbb{B}(0; 1), \quad \forall k \geq k_1, \quad (26)$$

where (24) and (25) are due to [29, Exercise 10.10]. By (26), we have

$$\|\nabla f(z^k)\| \leq \frac{1}{4\kappa} \cdot \text{dist}(z^k, \mathcal{X}) + \frac{\tau_k}{\lambda_k}, \quad \forall k \geq k_1. \quad (27)$$

Moreover, we have $z^k \in \mathcal{N}(\mathcal{X}; \beta_1)$ for all $k \geq k_0$ from (22). This, together with $\beta_1 \leq \rho$ and (13), yields $\text{dist}(z^k, \mathcal{X}) \leq \kappa \|\nabla f(z^k)\|$ for all $k \geq k_0$. By this, $k_1 \geq k_0$, and (27), we have

$$\text{dist}(z^k, \mathcal{X}) \leq \kappa \|\nabla f(z^k)\| \leq \frac{1}{4} \cdot \text{dist}(z^k, \mathcal{X}) + \frac{\kappa \tau_k}{\lambda_k}, \quad \forall k \geq k_1,$$

which results in $\text{dist}(z^k, \mathcal{X}) \leq \frac{4\kappa \tau_k}{3\lambda_k}$ for all $k \geq k_1$. This further leads to

$$\text{dist}(x^k, \mathcal{X}) = \|x^k - \hat{x}^k\| \leq \|x^k - \hat{z}^k\| \leq \|x^k - z^k\| + \text{dist}(z^k, \mathcal{X}) \leq \lambda_k + \frac{4\kappa \tau_k}{3\lambda_k}, \quad \forall k \geq k_1.$$

By the definitions of τ_k and λ_k , the above yields

$$\text{dist}^2(x^k, \mathcal{X}) \leq \text{dist}^2(x^k, \mathcal{X}) - \frac{4\kappa t_k}{3}, \quad \forall k \geq k_1,$$

which is a contradiction since $\kappa > 0$ and $t_k > 0$ for all $k \geq 0$. Therefore, there exists a scalar $\beta > 0$ such that (18) holds, which implies that statement (ii) holds.

We next prove (ii) \Rightarrow (i). Suppose that (14) holds with some $\alpha, \beta > 0$. Since $\nabla^2 f(x)$ is uniformly continuous on $\mathcal{N}(\mathcal{X}; \gamma)$, there exists a scalar $\rho_0 > 0$ such that

$$\|\nabla^2 f(x) - \nabla^2 f(y)\| \leq \frac{\alpha}{2}, \quad \forall x, y \in \mathcal{N}(\mathcal{X}; \gamma) \text{ with } \|x - y\| \leq \rho_0. \quad (28)$$

³Given an extended-real-valued function $h : \mathbb{R}^n \rightarrow (-\infty, +\infty]$ and an $x \in \text{dom}(h) := \{z \in \mathbb{R}^n : h(z) < \infty\}$, we denote by $\partial h(x)$ the *limiting subdifferential* (known also as the general or Mordukhovich subdifferential) of h at x ; see, *e.g.*, [26, Definition 1.77].

Let $\rho_1 = \min\{\rho_0, \beta, \gamma\} > 0$, $x \in \mathcal{N}(\mathcal{X}; \rho_1)$ be arbitrarily chosen, and $\tilde{x}(t) = x + t(\hat{x} - x)$ for $t \in [0, 1]$. Using the same arguments as those for (16), one has

$$\lambda_{\min}[\nabla^2 f(\tilde{x}(t))] \geq -\frac{\alpha}{2}, \quad \forall t \in [0, 1]. \quad (29)$$

By (14), (29), and the integral form of Taylor's series, we obtain

$$\begin{aligned} \langle \nabla f(x), x - \hat{x} \rangle &= f(x) - f(\hat{x}) + \int_0^1 (1-t)(\hat{x} - x)^T \nabla^2 f(\tilde{x}(t))(\hat{x} - x) dt \\ &\geq \frac{\alpha}{2} \|x - \hat{x}\|^2 - \frac{\alpha}{4} \|x - \hat{x}\|^2 = \frac{\alpha}{4} \|x - \hat{x}\|^2. \end{aligned}$$

Applying the Cauchy-Schwarz inequality and using $\text{dist}(x, \mathcal{X}) = \|x - \hat{x}\|$, the above yields

$$\text{dist}(x, \mathcal{X}) \leq \frac{4}{\alpha} \|\nabla f(x)\|, \quad \forall x \in \mathcal{N}(\mathcal{X}; \rho_1).$$

Therefore, statement (i) holds as well. \square

Remark. When f is convex, \mathcal{X} reduces to the set of optimal solutions to f and it is known that the EB condition (13) is equivalent to the quadratic growth condition (14); see, *e.g.*, [1]. When f is nonconvex, Drusvyatskiy *et al.* [14] studied these two regularity conditions for the set of first-order critical points (replacing \mathcal{X} in both (13) and (14) by the set of first-order critical points) and proved that they are equivalent under a quadratic lower-bound condition; see [14, Theorem 3.1]. Our Theorem 1 is motivated by [14, Theorem 3.1] and shows that for the set of second-order critical points of a twice continuously differentiable function, the EB condition (13) and the quadratic growth condition (14) are equivalent, as the quadratic lower-bound condition in [14] holds automatically.

Corollary 1. *Suppose that Assumption 2 and the premise of Theorem 1 hold. Then, any second-order critical point of f is a local minimizer.*

Proof. Let \bar{x} be an arbitrary second-order critical point of f . By Theorem 1 and Assumption 2, the quadratic growth condition (14) holds for some $\alpha, \beta > 0$. Let $\delta = \min\{\beta, \frac{\epsilon}{3}\}$ and x be an arbitrary point in $\mathcal{N}(\bar{x}; \delta)$. It then follows from (14) that $f(x) \geq f(\hat{x})$. Moreover, it holds that $\|\hat{x} - \bar{x}\| \leq \|x - \hat{x}\| + \|x - \bar{x}\| \leq 2\|x - \bar{x}\| \leq \frac{2}{3}\epsilon < \epsilon$. By this and the separation property in Theorem 1, we have $f(\hat{x}) = f(\bar{x})$. Hence, we obtain $f(x) \geq f(\bar{x})$ for all $x \in \mathcal{N}(\bar{x}; \delta)$, which implies that \bar{x} is a local minimizer of f . \square

For the rest of this section, we show that the classes of functions considered in Facts 3 and 4 satisfy Assumption 2.

Proposition 1. *Suppose that f is star-convex, \mathcal{X}^* is globally non-degenerate, and $\nabla^2 f(x)$ is uniformly continuous on $\mathcal{N}(\mathcal{X}^*; \gamma)$ for some $\gamma > 0$. Then, f satisfies Assumption 2.*

Proof. We first show that for star-convex functions, the set of second-order critical points equals the set of optimal solutions; *i.e.*, $\mathcal{X} = \mathcal{X}^*$. Since it is clear that $\mathcal{X}^* \subseteq \mathcal{X}$, it suffices to show that $\mathcal{X} \subseteq \mathcal{X}^*$. Suppose on the contrary that $x \notin \mathcal{X}^*$ for some $x \in \mathcal{X}$. Hence, $\nabla f(x) = 0$ and $f(x) > f(x^*)$ for any $x^* \in \mathcal{X}^*$. By this and (10), we have that for any $x^* \in \mathcal{X}^*$,

$$\begin{aligned} \langle \nabla f(x), x^* - x \rangle &= \lim_{\alpha \downarrow 0} \frac{f(x + \alpha(x^* - x)) - f(x)}{\alpha} \\ &\leq \lim_{\alpha \downarrow 0} \frac{\alpha f(x^*) + (1 - \alpha)f(x) - f(x)}{\alpha} = f(x^*) - f(x) < 0, \end{aligned}$$

which contradicts with $\nabla f(x) = 0$. Hence, we obtain $\mathcal{X} = \mathcal{X}^*$. This, together with our assumption in Proposition 1, implies that $\nabla^2 f(x)$ is uniformly continuous on $\mathcal{N}(\mathcal{X}; \gamma)$ for some $\gamma > 0$. Also, since

$\mathcal{X} = \mathcal{X}^*$, we have $f(x) = f(y) = f^*$ for any $x, y \in \mathcal{X}$, which implies that the separation property in Theorem 1 holds. Moreover, by $\mathcal{X} = \mathcal{X}^*$ and the assumption that \mathcal{X}^* is globally non-degenerate, statement (ii) of Theorem 1 holds. Hence, statement (i) of Theorem 1 holds as well, which implies that f satisfies Assumption 2. \square

Proposition 2. *Suppose that f is gradient-dominated of degree 2 and $\nabla^2 f(x)$ is uniformly continuous on $\mathcal{N}(\mathcal{X}^*; \gamma)$ for some $\gamma > 0$. Then, f satisfies Assumption 2.*

Proof. Due to (12), one can see that for any $x \notin \mathcal{X}^*$, we have $\nabla f(x) \neq 0$, which immediately implies that $\mathcal{X} \subseteq \mathcal{X}^*$. This, together with $\mathcal{X}^* \subseteq \mathcal{X}$, yields $\mathcal{X} = \mathcal{X}^*$. It then follows from the same arguments as those in the proof of Proposition 1 that the premise of Theorem 1 holds. Our next goal is to prove

$$f(x) - f^* \geq \frac{1}{4\tau_f} \cdot \text{dist}^2(x, \mathcal{X}^*), \quad \forall x \in \mathbb{R}^n. \quad (30)$$

Notice that (30) holds trivially for $x \in \mathcal{X}^*$. Let $\tilde{x} \in \mathbb{R}^n \setminus \mathcal{X}^*$ be arbitrarily chosen. Consider the differential equation

$$\begin{cases} u(0) = \tilde{x}, \\ \dot{u}(t) = -\nabla f(u(t)), \quad \forall t > 0. \end{cases} \quad (31)$$

Since ∇f is continuously differentiable on \mathbb{R}^n , it is Lipschitz continuous on any compact subset of \mathbb{R}^n . It then follows from the Picard-Lindelöf Theorem (see, e.g., [18, Theorem II.1.1]) that there exists a $\delta > 0$ such that (31) has a unique solution $u_{\tilde{x}}(t)$ for $t \in [0, \delta]$. Let $[0, \nu)$ be the maximal interval of existence for $u_{\tilde{x}}(t)$, where $\nu \leq \infty$.⁴ Define $H(t) := f(u_{\tilde{x}}(t)) - f^*$ for $t \in [0, \nu)$. Then, we have

$$\dot{H}(t) = \langle \nabla f(u_{\tilde{x}}(t)), \dot{u}_{\tilde{x}}(t) \rangle = -\|\nabla f(u_{\tilde{x}}(t))\| \cdot \|\dot{u}_{\tilde{x}}(t)\|, \quad \forall t \in [0, \nu), \quad (32)$$

where the second equality is due to (31). Using (12) and the definition of H , we get

$$\dot{H}(t) \leq -\frac{H(t)^{\frac{1}{2}}}{\sqrt{\tau_f}} \|\dot{u}_{\tilde{x}}(t)\|, \quad \forall t \in [0, \nu).$$

Recall that $\nabla f(x) = 0$ for any $x \in \mathcal{X}^*$. This implies that there does not exist a $\bar{t} \in [0, \nu)$ such that $u_{\tilde{x}}(\bar{t}) \in \mathcal{X}^*$, for otherwise $u_{\tilde{x}} \equiv u_{\tilde{x}}(\bar{t})$ is the unique solution to (31), which contradicts with $u_{\tilde{x}}(0) = \tilde{x} \notin \mathcal{X}^*$. Hence, $H(t) > 0$ for all $t \in [0, \nu)$ and

$$\|\dot{u}_{\tilde{x}}(t)\| \leq -\sqrt{\tau_f} \frac{\dot{H}(t)}{H(t)^{\frac{1}{2}}} = -2\sqrt{\tau_f} \left[H(t)^{\frac{1}{2}} \right]'. \quad (33)$$

Then, for any $0 \leq s_1 < s_2 < \nu$, we have

$$\begin{aligned} \|u_{\tilde{x}}(s_2) - u_{\tilde{x}}(s_1)\| &= \left\| \int_{s_1}^{s_2} \dot{u}_{\tilde{x}}(t) dt \right\| \leq \int_{s_1}^{s_2} \|\dot{u}_{\tilde{x}}(t)\| dt \\ &\leq \int_{s_1}^{s_2} -2\sqrt{\tau_f} \left[H(t)^{\frac{1}{2}} \right]' dt \\ &= 2\sqrt{\tau_f} \left[H(s_1)^{\frac{1}{2}} - H(s_2)^{\frac{1}{2}} \right]. \end{aligned} \quad (34)$$

Substituting $s_1 = 0$ in (34) and using $u_{\tilde{x}}(0) = \tilde{x}$ and $H(t) > 0$ for any $t \in [0, \nu)$, we obtain

$$\|u_{\tilde{x}}(s_2) - \tilde{x}\| \leq 2\sqrt{\tau_f} H(0)^{\frac{1}{2}} = 2\sqrt{\tau_f} (f(\tilde{x}) - f^*)^{\frac{1}{2}}, \quad \forall s_2 \in [0, \nu). \quad (35)$$

⁴An interval $[0, \nu)$ is called a *maximal interval* of existence for $u_{\tilde{x}}(t)$ if there does not exist an extension $\tilde{u}_{\tilde{x}}(t)$ of $u_{\tilde{x}}(t)$ over an interval $[0, \tilde{\nu})$ such that $\tilde{u}_{\tilde{x}}(t)$ remains a solution to (31) and $\tilde{\nu} > \nu$; see, e.g., [18, p. 12].

Next, we claim that $\nu = \infty$. Suppose to the contrary that $\nu < \infty$. Then, it follows from [18, Corollary II.3.2] that $\|u_{\tilde{x}}(t)\| \rightarrow \infty$ as $t \nearrow \nu$. However, the above inequality implies that

$$\|u_{\tilde{x}}(t)\| \leq \|\tilde{x}\| + \|u_{\tilde{x}}(t) - \tilde{x}\| \leq \|\tilde{x}\| + 2\sqrt{\tau_f}(f(\tilde{x}) - f^*)^{\frac{1}{2}}, \quad \forall t \in [0, \nu),$$

which yields a contradiction. Hence, the claim $\nu = \infty$ is true. In addition, we have $\dot{H}(t) \leq 0$ for all $t \in [0, \infty)$ from (32), which implies that $H(t)$ is non-increasing on $[0, \infty)$. This, together with $H(t) > 0$ for all $t \in [0, \infty)$, implies that $\lim_{t \rightarrow \infty} H(t)$ exists. It then follows from this and (34) that $u_{\tilde{x}}(t)$ has the Cauchy property and hence $\lim_{t \rightarrow \infty} u_{\tilde{x}}(t)$ exists. Let $u_{\tilde{x}}(\infty) := \lim_{t \rightarrow \infty} u_{\tilde{x}}(t)$. We claim that $\nabla f(u_{\tilde{x}}(\infty)) = 0$. Indeed, if $\nabla f(u_{\tilde{x}}(\infty)) \neq 0$, then by (31), (32), and the continuity of ∇f , we have

$$\lim_{t \rightarrow \infty} \dot{H}(t) = - \lim_{t \rightarrow \infty} \|\nabla f(u_{\tilde{x}}(t))\| \cdot \|\dot{u}_{\tilde{x}}(t)\| = - \lim_{t \rightarrow \infty} \|\nabla f(u_{\tilde{x}}(t))\|^2 = -\|\nabla f(u_{\tilde{x}}(\infty))\|^2 < 0,$$

which contradicts with the fact that $\lim_{t \rightarrow \infty} H(t)$ exists. Hence, we have $\nabla f(u_{\tilde{x}}(\infty)) = 0$. This, together with (12), yields $f(u_{\tilde{x}}(\infty)) = f^*$ and hence $u_{\tilde{x}}(\infty) \in \mathcal{X}^*$. By (35), this gives

$$\text{dist}(\tilde{x}, \mathcal{X}^*) \leq \|u_{\tilde{x}}(\infty) - \tilde{x}\| = \lim_{t \rightarrow \infty} \|u_{\tilde{x}}(t) - \tilde{x}\| \leq 2\sqrt{\tau_f}(f(\tilde{x}) - f^*)^{\frac{1}{2}},$$

which implies that (30) holds for $x = \tilde{x}$. Since $\tilde{x} \in \mathbb{R}^n \setminus \mathcal{X}^*$ is arbitrary, we conclude that (30) holds for all $x \in \mathbb{R}^n$. By (30) and the fact that $\mathcal{X} = \mathcal{X}^*$, statement (ii) of Theorem 1 holds. Hence, statement (i) of Theorem 1 holds as well, which implies that f satisfies Assumption 2. \square

It is worth noting that the argument on the gradient flow (31) in the proof of Proposition 2 follows that in [22]. In addition, a similar argument on subgradient flows has been applied in the convex setting [5]. Also, as pointed out by a referee of this paper, Proposition 2 follows from [13, Theorem 3.7], where a more general setting was considered and whose proof was based on the use of Ekeland's variational principle.

4 Quadratic Convergence of the CR Method

In this section, we establish the quadratic rate of convergence of the CR method under the local EB condition proposed in Section 3. To proceed, we start with the following consequence of Assumption 1.

Fact 5 ([27, Lemma 1]). *Suppose that Assumption 1 holds. Then, for any $x, y \in \mathcal{F}$,*

$$\|\nabla f(y) - \nabla f(x) - \nabla^2 f(x)(y - x)\| \leq \frac{L}{2}\|y - x\|^2. \quad (36)$$

We next prove the following intermediate lemma.

Lemma 1. *Suppose that Assumption 1 holds. Let $x \in \mathcal{F}$ and \hat{x} be a projection point of x to \mathcal{X} . If $\hat{x} \in \mathcal{F}$, then for any $\sigma > 0$, we have*

$$\|p_\sigma(x) - x\| \leq \left(1 + \frac{L}{\sigma} + \sqrt{\left(1 + \frac{L}{\sigma}\right)^2 + \frac{L}{\sigma}}\right) \cdot \text{dist}(x, \mathcal{X}). \quad (37)$$

Proof. For simplicity, we denote $x^+ := p_\sigma(x)$. By (5) and the first-order optimality condition of (4), one has

$$0 = \nabla f(x) + \nabla^2 f(x)(x^+ - x) + \frac{\sigma}{2}\|x^+ - x\|(x^+ - x). \quad (38)$$

Since $\hat{x} \in \mathcal{X}$, we have $\nabla f(\hat{x}) = 0$ and $\nabla^2 f(\hat{x}) \succeq 0$. By (38) and $\nabla f(\hat{x}) = 0$, it is not hard to verify that

$$\begin{aligned} \left(\nabla^2 f(\hat{x}) + \frac{\sigma \|x^+ - x\|}{2} I_n \right) (x^+ - \hat{x}) &= \nabla f(\hat{x}) - \nabla f(x) - \nabla^2 f(\hat{x})(\hat{x} - x) \\ &\quad - \frac{\sigma}{2} \|x^+ - x\|(\hat{x} - x) - (\nabla^2 f(x) - \nabla^2 f(\hat{x}))(x^+ - x). \end{aligned}$$

Since $\nabla^2 f(\hat{x}) \succeq 0$, we have

$$\left\| \left(\nabla^2 f(\hat{x}) + \frac{\sigma \|x^+ - x\|}{2} I_n \right) (x^+ - \hat{x}) \right\| \geq \frac{\sigma}{2} \|x^+ - x\| \cdot \|x^+ - \hat{x}\|.$$

This, together with the above equality, yields

$$\begin{aligned} \frac{\sigma}{2} \|x^+ - x\| \cdot \|x^+ - \hat{x}\| &\leq \|\nabla f(\hat{x}) - \nabla f(x) - \nabla^2 f(\hat{x})(\hat{x} - x)\| + \frac{\sigma}{2} \|x^+ - x\| \cdot \|x - \hat{x}\| \\ &\quad + \|\nabla^2 f(x) - \nabla^2 f(\hat{x})\| \cdot \|x^+ - x\| \\ &\leq \frac{L}{2} \|x - \hat{x}\|^2 + \left(\frac{\sigma}{2} + L \right) \|x^+ - x\| \cdot \|x - \hat{x}\|, \end{aligned}$$

where the second inequality is due to Fact 5 and the assumption that $\hat{x} \in \mathcal{F}$. Using the triangle inequality $\|x^+ - \hat{x}\| \geq \|x^+ - x\| - \|x - \hat{x}\|$, we further obtain

$$\frac{\sigma}{2} \|x^+ - x\|^2 \leq \frac{L}{2} \|x - \hat{x}\|^2 + (\sigma + L) \|x^+ - x\| \cdot \|x - \hat{x}\|.$$

By solving the above quadratic inequality, one has

$$\|x^+ - x\| \leq \left(1 + \frac{L}{\sigma} + \sqrt{\left(1 + \frac{L}{\sigma} \right)^2 + \frac{L}{\sigma}} \right) \cdot \|x - \hat{x}\|.$$

Noticing that $x^+ = p_\sigma(x)$ and $\text{dist}(x, \mathcal{X}) = \|x - \hat{x}\|$, we obtain the desired inequality (37). \square

Remark. As we shall see in the sequel, Lemma 1 implies that there exists a $c_1 > 0$ such that

$$\|x^{k+1} - x^k\| \leq c_1 \cdot \text{dist}(x^k, \mathcal{X}) \quad (39)$$

for all sufficiently large k , where $\{x^k\}_{k \geq 0}$ is the sequence generated by Algorithm 1. It is known that establishing (39) is an important step for analyzing local convergence of Newton-type methods with non-isolated solutions. However, our proof of Lemma 1 is novel. Indeed, in most cases (39) is obtained based on the property that x^{k+1} is the minimizer of a strongly convex function (see, *e.g.*, [20, 36, 34, 15, 2]), which does not apply to the CR method. Moreover, in our proof of Lemma 1, the fact that $\nabla^2 f(x) \succeq 0$ for any $x \in \mathcal{X}$ plays a crucial role.

Now we are ready to present the main result of this section.

Theorem 2. *Suppose that Assumptions 1 and 2 hold. Let $\{x^k\}_{k \geq 0}$ be the sequence of iterates generated by Algorithm 1. If $\mathcal{L}(f(x^k))$ is bounded for some $k \geq 0$, then the whole sequence $\{x^k\}_{k \geq 0}$ converges at least Q -quadratically to a point $x^* \in \mathcal{X}$.*

Proof. Let \hat{x}^k be a projection point of x^k to \mathcal{X} ; *i.e.*, $\hat{x}^k \in \text{Argmin}_{z \in \mathcal{X}} \|z - x^k\|$. Let $\bar{\mathcal{X}}$ be the set of accumulation points of $\{x^k\}_{k \geq 0}$. By (8), we have $f(x^{k+1}) \leq f_{\sigma_k}(x^k) \leq f(x^k)$ for all k . This, together with the boundedness of $\mathcal{L}(f(x^k))$ for some $k \geq 0$, implies the boundedness of $\{x^k\}_{k \geq 0}$. Hence, $\bar{\mathcal{X}}$ is non-empty and bounded, and we have $\lim_{k \rightarrow \infty} \text{dist}(x^k, \bar{\mathcal{X}}) = 0$. By Fact 1 (iii), we have $\bar{\mathcal{X}} \subseteq \mathcal{X}$.

Thus, $\lim_{k \rightarrow \infty} \text{dist}(x^k, \mathcal{X}) = \lim_{k \rightarrow \infty} \|x^k - \hat{x}^k\| = 0$. It then follows from Assumption 2 that there exists a $k_1 \geq 0$ such that

$$\text{dist}(x^k, \mathcal{X}) \leq \kappa \|\nabla f(x^k)\|, \quad \forall k \geq k_1.$$

In addition, since $x^k \in \mathcal{L}(f(x^0)) \subseteq \text{int}(\mathcal{F})$ for all $k \geq 0$ and $\{x^k\}_{k \geq 0}$ is bounded, there exists a compact set $\mathcal{M} \subseteq \text{int}(\mathcal{F})$ such that $\{x^k\}_{k \geq 0} \subseteq \mathcal{M}$. Also, it follows from $\{x^k\}_{k \geq 0} \subseteq \mathcal{M}$ and $\lim_{k \rightarrow \infty} \|x^k - \hat{x}^k\| = 0$ that $\lim_{k \rightarrow \infty} \text{dist}(\hat{x}^k, \mathcal{M}) = 0$. This, together with $\mathcal{M} \subseteq \text{int}(\mathcal{F})$ and the compactness of \mathcal{M} , implies that $\hat{x}^k \in \text{int}(\mathcal{F})$ for all sufficiently large k . Hence, there exists a $k_2 \geq 0$ such that

$$x^k \in \mathcal{F}, \quad \hat{x}^k \in \mathcal{F}, \quad \forall k \geq k_2. \quad (40)$$

Hence, for any $k \geq \bar{k} := \max\{k_1, k_2\}$, we have

$$\begin{aligned} \text{dist}(x^{k+1}, \mathcal{X}) &\leq \kappa \|\nabla f(x^{k+1})\| \\ &= \kappa \left\| \nabla f(x^{k+1}) - \nabla f(x^k) - \nabla^2 f(x^k)(x^{k+1} - x^k) - \frac{\sigma_k}{2} \|x^{k+1} - x^k\| (x^{k+1} - x^k) \right\| \\ &\leq \kappa \left\| \nabla f(x^{k+1}) - \nabla f(x^k) - \nabla^2 f(x^k)(x^{k+1} - x^k) \right\| + \frac{\kappa \sigma_k}{2} \|x^{k+1} - x^k\|^2 \\ &\leq \frac{\kappa(L + \sigma_k)}{2} \|x^{k+1} - x^k\|^2 \leq \frac{3}{2} \kappa L \|x^{k+1} - x^k\|^2, \end{aligned} \quad (41)$$

where the equality is due to the first-order optimality condition of (4), the third inequality is by (40) and Fact 5, and the last inequality is by $\sigma_k \leq 2L$ for all $k \geq 0$. Using (40), Lemma 1, and $\sigma_k \geq \bar{\sigma} > 0$ for all k , we have

$$\|x^{k+1} - x^k\| \leq c_1 \cdot \text{dist}(x^k, \mathcal{X}), \quad \forall k \geq \bar{k}, \quad (42)$$

where $c_1 = \left(1 + \frac{L}{\bar{\sigma}} + \sqrt{\left(1 + \frac{L}{\bar{\sigma}}\right)^2 + \frac{L}{\bar{\sigma}}}\right)$. Combining (41) and (42), we obtain

$$\text{dist}(x^{k+1}, \mathcal{X}) \leq c_2 \cdot \text{dist}^2(x^k, \mathcal{X}), \quad \forall k \geq \bar{k}, \quad (43)$$

where $c_2 = \frac{3}{2} \kappa c_1^2 L$. We next show that the whole sequence $\{x^k\}_{k \geq 0}$ is convergent. Let $\eta > 0$ be arbitrary. Since $\lim_{k \rightarrow \infty} \text{dist}(x^k, \mathcal{X}) = 0$, there exists a $k_3 \geq 0$ such that

$$\text{dist}(x^k, \mathcal{X}) \leq \min \left\{ \frac{1}{2c_2}, \frac{\eta}{2c_1} \right\}, \quad \forall k \geq k_3.$$

It then follows from (43) that

$$\text{dist}(x^{k+1}, \mathcal{X}) \leq c_2 \cdot \text{dist}^2(x^k, \mathcal{X}) \leq \frac{1}{2} \text{dist}(x^k, \mathcal{X}), \quad \forall k \geq \max\{k_3, \bar{k}\}.$$

This, together with (42), implies that for any $k \geq \max\{k_3, \bar{k}\}$ and any $j \geq 0$, we have

$$\begin{aligned} \|x^{k+j} - x^k\| &\leq \sum_{i=k}^{\infty} \|x^{i+1} - x^i\| \leq \sum_{i=k}^{\infty} c_1 \cdot \text{dist}(x^i, \mathcal{X}) \\ &\leq c_1 \cdot \text{dist}(x^k, \mathcal{X}) \cdot \sum_{i=0}^{\infty} \frac{1}{2^i} \leq 2c_1 \cdot \text{dist}(x^k, \mathcal{X}) \leq \eta, \end{aligned} \quad (44)$$

which implies that $\{x^k\}_{k \geq \max\{k_3, \bar{k}\}}$ is a Cauchy sequence. Therefore, the whole sequence $\{x^k\}_{k \geq 0}$ is convergent. Finally, we study the convergence rate of $\{x^k\}_{k \geq 0}$. Let $x^* := \lim_{k \rightarrow \infty} x^k$. By Fact 1, we have $x^* \in \mathcal{X}$. It follows from (43) and (44) that for any $k \geq \max\{k_3, \bar{k}\}$,

$$\|x^* - x^{k+1}\| = \lim_{j \rightarrow \infty} \|x^{k+1+j} - x^{k+1}\| \leq 2c_1 \cdot \text{dist}(x^{k+1}, \mathcal{X}) \leq 2c_1 c_2 \cdot \text{dist}^2(x^k, \mathcal{X}).$$

Combining this with $\text{dist}(x^k, \mathcal{X}) \leq \|x^k - x^*\|$, we obtain

$$\frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|^2} \leq 2c_1c_2, \quad \forall k \geq \max\{k_3, \bar{k}\}.$$

Therefore, $\{x^k\}_{k \geq 0}$ converges at least Q-quadratically to an element x^* in \mathcal{X} . \square

From the proof of Theorem 2, we immediately have the following result that explicitly characterizes the region where local quadratic convergence occurs.

Corollary 2. *Suppose that Assumptions 1 and 2 hold and $\mathcal{X} \subseteq \mathcal{F}$. Let $\{x^k\}_{k \geq 0}$ be the sequence of iterates generated by Algorithm 1 with $\bar{\sigma} = L$. Then, we have*

$$\text{dist}(x^{k+1}, \mathcal{X}) \leq 27\kappa L \cdot \text{dist}^2(x^k, \mathcal{X}) \quad \text{whenever} \quad \text{dist}(x^k, \mathcal{X}) \leq \rho. \quad (45)$$

Proof. Suppose that $\text{dist}(x^k, \mathcal{X}) \leq \rho$ for some k . By Assumption 1 and $f(x^{k+1}) \leq f(x^k)$ for all k , we have $x^k \in \mathcal{F}$. In addition, since $\mathcal{X} \subseteq \mathcal{F}$, we obtain $\hat{x}^k \in \mathcal{F}$. Also, it follows from Assumption 2 that $\text{dist}(x^k, \mathcal{X}) \leq \kappa \|\nabla f(x^k)\|$. By these and similar arguments as those for deriving (43), one has

$$\text{dist}(x^{k+1}, \mathcal{X}) \leq \frac{3}{2}\kappa c_1^2 L \cdot \text{dist}^2(x^k, \mathcal{X}),$$

where $c_1 = \left(1 + \frac{L}{\bar{\sigma}} + \sqrt{\left(1 + \frac{L}{\bar{\sigma}}\right)^2 + \frac{L}{\bar{\sigma}}}\right)$. This, together with $\bar{\sigma} = L$, implies (45). \square

Equipped with Theorem 2 and Propositions 1 and 2, we can improve the results in [27] on the local convergence rate of the CR method when applied to globally non-degenerate star-convex functions and gradient-dominated functions.

Corollary 3. *Suppose that Assumption 1 holds, f is star-convex, and \mathcal{X}^* is globally non-degenerate. If $\mathcal{L}(f(x^k))$ is bounded for some $k \geq 0$, then the whole sequence $\{x^k\}_{k \geq 0}$ converges at least Q-quadratically to a point $x^* \in \mathcal{X}^*$.*

Proof. Since $\mathcal{L}(f(x^k))$ is bounded for some $k \geq 0$, we have that \mathcal{X}^* is bounded. This, together with Assumption 1, implies that $\nabla^2 f(x)$ is uniformly continuous on $\mathcal{N}(\mathcal{X}^*; \gamma)$ for some $\gamma > 0$. The premise of Proposition 1 then holds. Hence, by Proposition 1 and its proof, we obtain that Assumption 2 holds and $\mathcal{X} = \mathcal{X}^*$. The conclusion of Corollary 3 then follows from Theorem 2. \square

Corollary 4. *Suppose that Assumption 1 holds and f is gradient-dominated with degree 2. If $\mathcal{L}(f(x^k))$ is bounded for some $k \geq 0$, then the whole sequence $\{x^k\}_{k \geq 0}$ converges at least Q-quadratically to a point $x^* \in \mathcal{X}^*$.*

Proof. Using the same arguments as those in the proof of Corollary 3, we have that the premise of Proposition 2 holds. Hence, by Proposition 2 and its proof, we obtain that Assumption 2 holds and $\mathcal{X} = \mathcal{X}^*$. The conclusion of Corollary 4 then follows from Theorem 2. \square

5 Applications to Structured Nonconvex Optimization Problems

In this section, we study the CR method for solving two concrete nonconvex minimization problems that arise in phase retrieval and low-rank matrix recovery.

5.1 Phase Retrieval

In this subsection, we consider the application of the CR method for solving (noiseless) phase retrieval problems. Specifically, the problem of interest is to recover an unknown complex signal $z^* = x^* + iy^* \in \mathbb{C}^n$ from the measurements

$$b_j = |a_j^H z^*|, \quad j = 1, \dots, m, \quad (46)$$

where $\{a_j\}_{j=1}^m \subseteq \mathbb{C}^n$ are assumed to be independently sampled from the standard complex Gaussian distribution $\mathcal{CN}(0, I_n)$. For non-triviality, we assume that $z^* \neq 0$. Since for any $\phi \in [0, 2\pi)$, $z^* e^{i\phi}$ provides exactly the same measurements, we can only expect to recover z^* up to this ambiguity. Such problem has broad applications in science and engineering, including optics, signal processing, computer vision, and quantum mechanics. For more discussions on its applications and recent developments, the readers are invited to the survey papers [30, 23].

Given the form (46), the following optimization formulation arises naturally:

$$\min_{z \in \mathbb{C}^n} f_c(z) := \frac{1}{2m} \sum_{j=1}^m (|a_j^H z|^2 - b_j^2)^2. \quad (47)$$

Let $\mathcal{Z}^* := \{z^* e^{i\phi} : \phi \in [0, 2\pi)\}$ be the set of target signals. Observe that $f_c(z) \geq 0$ for any $z \in \mathbb{C}^n$ and $f_c(z) = 0$ for any $z \in \mathcal{Z}^*$. Hence, any $z \in \mathcal{Z}^*$ is a global minimizer of (47). By letting $f(x, y) := f_c(x + iy)$, the corresponding real-variable problem of (47) is given by

$$\min_{x, y \in \mathbb{R}^n} f(x, y) = \frac{1}{2m} \sum_{j=1}^m \left(\left\| \begin{pmatrix} \Re(a_j) & -\Im(a_j) \\ \Im(a_j) & \Re(a_j) \end{pmatrix}^T \begin{pmatrix} x \\ y \end{pmatrix} \right\|^2 - b_j^2 \right)^2. \quad (48)$$

Let $\mathcal{X}^* := \{(x^* \cos \phi - y^* \sin \phi, x^* \sin \phi + y^* \cos \phi) : \phi \in [0, 2\pi)\}$. Using similar arguments, one can verify that any $(x, y) \in \mathcal{X}^*$ is a global minimizer of (48). Also, it holds that $(x, y) \in \mathcal{X}^*$ if and only if $x + iy \in \mathcal{Z}^*$. Hence, as long as we obtain an element $(x, y) \in \mathcal{X}^*$, the phase retrieval problem is solved by letting $z = x + iy$.

We now state the main result of this subsection.

Theorem 3. *There exist constants $c_0, c_1 > 0$ such that when $m \geq c_0 n \log^3 n$, it holds with probability at least $1 - c_1 m^{-1}$ that with any arbitrary initialization, the sequence of iterates $\{(x^k, y^k)\}_{k \geq 0}$ generated by Algorithm 1 for solving (48) converges at least Q -quadratically to an element in \mathcal{X}^* .*

The rest of this subsection is devoted to proving the above theorem. Before we proceed, let us lay out the concepts of Wirtinger calculus that are necessary for our developments on complex-variable functions. Let $h_c : \mathbb{C}^n \rightarrow \mathbb{R}$ be a real-valued function on \mathbb{C}^n and $h : \mathbb{R}^{2n} \rightarrow \mathbb{R}$ be defined as $h(x, y) = h_c(x + iy)$ for any $x, y \in \mathbb{R}^n$. We define

$$\frac{\partial}{\partial z} := \frac{1}{2} \left(\frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right) \quad \text{and} \quad \frac{\partial}{\partial \bar{z}} := \frac{1}{2} \left(\frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right),$$

which can be understood as operators acting on real-valued functions of (x, y) . Then, the *Wirtinger gradient* $\nabla_w h_c$ and *Wirtinger Hessian* $\nabla_w^2 h_c$ of h_c are defined, respectively, as

$$\nabla_w h_c := \left[\frac{\partial h}{\partial z}, \frac{\partial h}{\partial \bar{z}} \right]^H \quad \text{and} \quad \nabla_w^2 h_c := \begin{pmatrix} \frac{\partial}{\partial z} \left(\frac{\partial h}{\partial z} \right)^H & \frac{\partial}{\partial \bar{z}} \left(\frac{\partial h}{\partial z} \right)^H \\ \frac{\partial}{\partial z} \left(\frac{\partial h}{\partial \bar{z}} \right)^H & \frac{\partial}{\partial \bar{z}} \left(\frac{\partial h}{\partial \bar{z}} \right)^H \end{pmatrix}.$$

Define the matrix

$$J = \frac{1}{2} \begin{pmatrix} I_n & iI_n \\ I_n & -iI_n \end{pmatrix},$$

which satisfies $2JJ^H = 2J^HJ = I_{2n}$ and

$$\nabla_w h_c(x + iy) = J\nabla h(x, y), \quad \nabla_w^2 h_c(x + iy) = J\nabla^2 h(x, y)J^H.$$

In particular, for the function f_c in (47), we have

$$\nabla_w f_c(z) = \frac{1}{m} \sum_{j=1}^m \begin{pmatrix} (|a_j^H z|^2 - b_j^2) (a_j a_j^H) z \\ (|a_j^H z|^2 - b_j^2) (a_j a_j^H)^T \bar{z} \end{pmatrix} \quad (49)$$

and

$$\nabla_w^2 f_c(z) = \frac{1}{m} \sum_{j=1}^m \begin{pmatrix} (2|a_j^H z|^2 - b_j^2) a_j a_j^H & (a_j^H z)^2 a_j a_j^T \\ (a_j^T \bar{z})^2 \bar{a}_j a_j^H & (2|a_j^H z|^2 - b_j^2) \bar{a}_j a_j^T \end{pmatrix} \quad (50)$$

for any $z \in \mathbb{C}^n$; see, e.g., [7, Section 7.2].

5.1.1 Second-Order Critical Points and Local EB Condition

We first show that with high probability, the set of second-order critical points of f is \mathcal{X}^* . Moreover, we show that in a neighborhood of \mathcal{X}^* , the local EB condition (13) holds. For this purpose, we need the following result, which is directly implied by [31, Theorem 2].

Fact 6. *Let \mathcal{U}_c be a neighborhood of \mathcal{Z}^* defined as $\mathcal{U}_c := \{z \in \mathbb{C}^n : \text{dist}(z, \mathcal{Z}^*) \leq \frac{1}{\sqrt{7}} \|z^*\|\}$. There exist constants $c_2, c_3 > 0$ such that when $m \geq c_2 n \log^3 n$, the following statements hold with probability at least $1 - c_3 m^{-1}$.*

(i) *For any $z \notin \mathcal{U}_c$, if $\nabla_w f_c(z) = 0$, then*

$$\begin{pmatrix} \hat{z} \\ \bar{\hat{z}} \end{pmatrix}^H \nabla_w^2 f_c(z) \begin{pmatrix} \hat{z} \\ \bar{\hat{z}} \end{pmatrix} \leq -\frac{\|z^*\|^4}{100}, \quad (51)$$

where \hat{z} is defined as the unique projection of z to \mathcal{Z}^* ; i.e.,

$$\hat{z} = z^* e^{i\phi(z)}, \quad \text{with } \phi(z) = \underset{\phi \in [0, 2\pi)}{\operatorname{argmin}} \|z - z^* e^{i\phi}\|.$$

(ii) *For any $z \in \mathcal{U}_c$, it holds that*

$$\left(\frac{g(z)}{g(z)} \right)^H \nabla_w^2 f_c(z) \left(\frac{g(z)}{g(z)} \right) \geq \frac{\|z^*\|^2}{4} \cdot \|g(z)\|^2, \quad (52)$$

where $g(z) := z - \hat{z}$.

Proposition 3. *There exist constants $c_2, c_3 > 0$ such that when $m \geq c_2 n \log^3 n$, the following statements on f hold with probability at least $1 - c_3 m^{-1}$.*

(i) *\mathcal{X}^* equals the set of second-order critical points of f .*

(ii) *The following error bound holds:*

$$\text{dist}((x, y), \mathcal{X}^*) \leq \frac{4}{\|z^*\|^2} \|\nabla f(x, y)\| \quad \text{whenever } \text{dist}((x, y), \mathcal{X}^*) \leq \frac{1}{\sqrt{7}} \|z^*\|. \quad (53)$$

Proof. It suffices to prove that statements (i) and (ii) of Fact 6 lead to the statements (i) and (ii) herein. We first prove (ii). Let (x, y) be an arbitrary point satisfying $\text{dist}((x, y), \mathcal{X}^*) \leq \frac{1}{\sqrt{7}}\|z^*\|$ and (\hat{x}, \hat{y}) be the projection of (x, y) to \mathcal{X}^* . By definition, one can easily verify that $x + iy \in \mathcal{U}_c$ and the projection of $x + iy$ to \mathcal{Z}^* is $\hat{x} + i\hat{y}$. We assume that $(x, y) \notin \mathcal{X}^*$ since (53) holds trivially otherwise. Thus, $x + iy \notin \mathcal{Z}^*$ and we have $g(x + iy) = x - \hat{x} + i(y - \hat{y})$. This, together with the identity $\nabla_w^2 f_c(z) = J\nabla^2 f(x, y)J^H$ and (52), yields

$$\begin{pmatrix} x - \hat{x} \\ y - \hat{y} \end{pmatrix}^T \nabla^2 f(x, y) \begin{pmatrix} x - \hat{x} \\ y - \hat{y} \end{pmatrix} \geq \frac{\|z^*\|^2}{4} \cdot \|(x, y) - (\hat{x}, \hat{y})\|^2. \quad (54)$$

Let $(x(t), y(t)) = t \cdot (x, y) + (1 - t) \cdot (\hat{x}, \hat{y})$ for $t \in [0, 1]$. Note that the projection of $(x(t), y(t))$ to \mathcal{X}^* is (\hat{x}, \hat{y}) for any $t \in [0, 1]$. Using the same arguments, (54) holds if we substitute (x, y) by $(x(t), y(t))$ for any $t \in [0, 1]$. Hence, by the integral form of Taylor's series, we obtain

$$\begin{aligned} f(x, y) &= f(\hat{x}, \hat{y}) + \nabla f(\hat{x}, \hat{y})^T \begin{pmatrix} x - \hat{x} \\ y - \hat{y} \end{pmatrix} + \int_0^1 (1 - t) \cdot \begin{pmatrix} x - \hat{x} \\ y - \hat{y} \end{pmatrix}^T \nabla^2 f(x(t), y(t)) \begin{pmatrix} x - \hat{x} \\ y - \hat{y} \end{pmatrix} dt \\ &\geq f(\hat{x}, \hat{y}) + \nabla f(\hat{x}, \hat{y})^T \begin{pmatrix} x - \hat{x} \\ y - \hat{y} \end{pmatrix} + \frac{\|z^*\|^2}{8} \cdot \|(x, y) - (\hat{x}, \hat{y})\|^2, \end{aligned}$$

and similarly,

$$f(\hat{x}, \hat{y}) \geq f(x, y) - \nabla f(x, y)^T \begin{pmatrix} x - \hat{x} \\ y - \hat{y} \end{pmatrix} + \frac{\|z^*\|^2}{8} \cdot \|(x, y) - (\hat{x}, \hat{y})\|^2.$$

Noticing that $f(\hat{x}, \hat{y}) = 0$ and $\nabla f(\hat{x}, \hat{y}) = 0$ (by the global optimality of (\hat{x}, \hat{y})), we obtain (53) by summing up the above two inequalities.

We next prove (i). Let \mathcal{X} be the set of second-order critical points of f . Clearly, we have $\mathcal{X}^* \subseteq \mathcal{X}$ since any $(x, y) \in \mathcal{X}^*$ is a global minimizer of f . We now show that $\mathcal{X} \subseteq \mathcal{X}^*$. Let $(x, y) \in \mathcal{X}$ be arbitrary. By definition, $\nabla f(x, y) = 0$ and $\nabla^2 f(x, y) \succeq 0$. Using $\nabla f(x, y) = 0$ and the result in (i), we see that $(x, y) \in \mathcal{X}^*$ or $\text{dist}((x, y), \mathcal{X}^*) > \frac{1}{\sqrt{7}}\|z^*\|$. If the latter holds, we have by definition that $x + iy \notin \mathcal{U}_c$. In addition, it holds that $\nabla_w f_c(x + iy) = J\nabla f(x, y) = 0$. Hence, the inequality (51) holds for $x + iy$. This, together with the identity $\nabla_w^2 f_c(x + iy) = J\nabla^2 f(x, y)J^H$, implies that $\nabla^2 f(x, y) \not\succeq 0$, which contradicts with $(x, y) \in \mathcal{X}$. Therefore, we have $\mathcal{X} \subseteq \mathcal{X}^*$ and hence $\mathcal{X} = \mathcal{X}^*$. \square

5.1.2 Lipschitz Continuity of $\nabla^2 f$

Our next step is to verify the Lipschitz continuity of $\nabla^2 f$. Let $A = (a_1, \dots, a_m) \in \mathbb{C}^{n \times m}$ and $M = \max_j \|a_j\|$. We need the following result, which combines Lemma 23 and Lemma 28 of [31].

Fact 7. *There exist constants $c_4, c_5, c_6 > 0$ such that when $m \geq c_4 n$, it holds with probability at least $1 - c_5 \exp(-c_6 m)$ that*

$$\frac{m}{2} \leq \lambda_{\min}(AA^H) \leq \lambda_{\max}(AA^H) \leq 2m \quad (55)$$

and

$$\frac{1}{m} \sum_{j=1}^m |a_j^H w|^2 - |a_j^H w'|^2 \leq \frac{3}{2} \|w - w'\| (\|w\| + \|w'\|), \quad \forall w, w' \in \mathbb{C}^n. \quad (56)$$

Proposition 4. *Suppose that (55) and (56) hold. Then, for any $R > 0$, $\nabla^2 f$ is Lipschitz continuous on $\mathbb{B}(0; R) = \{(x, y) \in \mathbb{R}^{2n} : \|(x, y)\| \leq R\}$ with Lipschitz constant $L = 20M^2 R$.*

Proof. Let $w = x + iy$ and $w' = x' + iy'$ with $(x, y), (x', y') \in \mathbb{B}(0; R)$. Thus, $\|w\| \leq R$ and $\|w'\| \leq R$. By the identities $\nabla_w^2 f_c(x + iy) = J \nabla^2 f(x, y) J^H$ and $2J^H J = I_{2n}$, we have

$$\begin{aligned} \|\nabla^2 f(x, y) - \nabla^2 f(x', y')\| &= \sup_{\|(u, v)\|=1} \left| \begin{pmatrix} u \\ v \end{pmatrix}^T [\nabla^2 f(x, y) - \nabla^2 f(x', y')] \begin{pmatrix} u \\ v \end{pmatrix} \right| \\ &= \sup_{\|(u, v)\|=1} \left| 4 \begin{pmatrix} u \\ v \end{pmatrix}^T J^H [\nabla_w^2 f_c(w) - \nabla_w^2 f_c(w')] J \begin{pmatrix} u \\ v \end{pmatrix} \right| \\ &= \sup_{\|z\|=1} \left| \begin{pmatrix} z \\ \bar{z} \end{pmatrix}^H [\nabla_w^2 f_c(w) - \nabla_w^2 f_c(w')] \begin{pmatrix} z \\ \bar{z} \end{pmatrix} \right|. \end{aligned}$$

Using (50), (55), and (56), we further have

$$\begin{aligned} &\|\nabla^2 f(x, y) - \nabla^2 f(x', y')\| \\ &\leq \sup_{\|z\|=1} \left| \frac{4}{m} \sum_{j=1}^m (|a_j^H w|^2 - |a_j^H w'|^2) |a_j^H z|^2 \right| + \left| \frac{2}{m} \sum_{j=1}^m \Re \left([(a_j^H w)^2 - (a_j^H w')^2] (z^H a_j)^2 \right) \right| \\ &\leq \sup_{\|z\|=1} \frac{4}{m} \sum_{j=1}^m ||a_j^H w|^2 - |a_j^H w'|^2| |a_j^H z|^2 + \frac{2}{m} \sum_{j=1}^m |a_j^H w - a_j^H w'| |a_j^H w + a_j^H w'| |a_j^H z|^2 \\ &\leq 4M^2 \cdot \frac{1}{m} \sum_{j=1}^m ||a_j^H w|^2 - |a_j^H w'|^2| + 4M^2 R \|w - w'\| \cdot \left\| \frac{1}{m} \sum_{j=1}^m a_j a_j^H \right\| \\ &\leq 6M^2 \|w - w'\| (\|w\| + \|w'\|) + \frac{4}{m} M^2 R \|w - w'\| \cdot \lambda_{\max}(AA^H) \\ &\leq 20M^2 R \|(x, y) - (x', y')\|. \end{aligned}$$

The proof is then completed. \square

5.1.3 Proof of Theorem 3

In view of Theorem 2, it suffices to prove that with high probability, the following statements hold simultaneously: (i) Assumption 1 holds, (ii) Assumption 2 holds, (iii) $\mathcal{L}(f(x^k, y^k))$ is bounded for some $k \geq 0$, and (iv) \mathcal{X}^* equals the set of second-order critical points of f .

Let $c_0 := \max\{c_2, c_4\}$ and $m \geq c_0 n \log^3 n$. Suppose that $(x^0, y^0) \in \mathbb{R}^n \times \mathbb{R}^n$ is the initial point of Algorithm 1. We define

$$\begin{aligned} \bar{R} &:= \left(\frac{2\sqrt{2}}{\sqrt{m}} \sqrt{2mf(x^0, y^0) + \sum_{j=1}^m b_j^4} \right)^{\frac{1}{2}} > 0 \quad \text{and} \\ \mathcal{F} &:= \mathbb{B}(0; 2\bar{R}) = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n : \|(x, y)\| \leq 2\bar{R}\}. \end{aligned}$$

Suppose that (55) and (56) hold. Let (x, y) be an arbitrary point in $\mathcal{L}(f(x^0, y^0))$ and set $z := x + iy$. By definition, we have $f_c(z) = f(x, y) \leq f(x^0, y^0)$. It then follows from (55) and the definition of A that

$$\|(x, y)\|^2 = \|z\|^2 \leq \frac{1}{\lambda_{\min}(AA^H)} \cdot z^H AA^H z \leq \frac{2}{m} \sum_{j=1}^m |a_j^H z|^2.$$

Using the inequality $(\sum_{j=1}^m \alpha_j)^2 \leq m \sum_{j=1}^m \alpha_j^2$, which holds for any real numbers $\{\alpha_i\}_{i=1}^m$, we further

have

$$\begin{aligned}\|(x, y)\|^2 &\leq \frac{2}{\sqrt{m}} \cdot \sqrt{\sum_{j=1}^m |a_j^H z|^4} \leq \frac{2}{\sqrt{m}} \cdot \sqrt{\sum_{j=1}^m \left[2 \left(|a_j^H z|^2 - b_j^2 \right)^2 + 2b_j^4 \right]} \\ &= \frac{2\sqrt{2}}{\sqrt{m}} \sqrt{2mf_c(z) + \sum_{j=1}^m b_j^4} \leq \frac{2\sqrt{2}}{\sqrt{m}} \sqrt{2mf(x^0, y^0) + \sum_{j=1}^m b_j^4} = \bar{R}^2.\end{aligned}$$

By the definition of \mathcal{F} , we have $\mathcal{L}(f(x^0, y^0)) \subseteq \text{int}(\mathcal{F})$ and $\mathcal{L}(f(x^0, y^0))$ is bounded. In addition, by Proposition 4, $\nabla^2 f$ is Lipschitz continuous on \mathcal{F} with Lipschitz constant $L = 40M^2\bar{R}$. Hence, statements (i) and (iii) above hold with probability at least $1 - c_5 \exp(-c_6 m)$. Furthermore, by Proposition 3, statements (ii) and (iv) above hold with probability at least $1 - c_4 m^{-1}$. Therefore, there exists a $c_1 > 0$ such that all the statements (i)–(iv) hold with probability at least $1 - c_1 m^{-1}$. The proof is then completed. \square

5.2 Low-Rank Matrix Recovery

In this subsection, we consider the application of the CR method for solving low-rank matrix recovery problems. Specifically, the problem of interest is to recover an unknown low-rank matrix $X^* \in \mathbb{R}^{n_1 \times n_2}$ with $\text{rank}(X^*) = r \ll \min\{n_1, n_2\}$ from the measurements

$$\mathbb{R}^m \ni b = \mathcal{A}(X^*), \quad (57)$$

where the linear operator $\mathcal{A} : \mathbb{R}^{n_1 \times n_2} \rightarrow \mathbb{R}^m$ is given by $\mathcal{A}(X) = (\langle A_1, X \rangle, \dots, \langle A_m, X \rangle)$ for any $X \in \mathbb{R}^{n_1 \times n_2}$. For simplicity, we assume that $n_1 = n_2 = n$, A_i 's are symmetric, and the target matrix X^* is symmetric and positive semidefinite.

Since $X^* \succeq 0$ with $\text{rank}(X^*) = r$, we have $X^* = U^* U^{*T}$ for some $U^* \in \mathbb{R}^{n \times r}$. This motivates the following nonconvex formulation for recovering X^* :

$$\min_{U \in \mathbb{R}^{n \times r}} f(U) := \frac{1}{4m} \|\mathcal{A}(UU^T) - b\|^2. \quad (58)$$

By letting $\mathcal{U} := \{U^* Q : Q \in \mathcal{O}^r\}$, it holds that $\mathcal{U} = \{U \in \mathbb{R}^{n \times r} : X^* = UU^T\}$. Hence, we can recover the unknown matrix X^* as long as we find any $U \in \mathcal{U}$. Observe that f is non-negative and $f(U) = 0$ for any $U \in \mathcal{U}$. Hence, any $U \in \mathcal{U}$ is a global minimizer of (58).

We next introduce the so-called restricted isometry property (RIP) of the operator \mathcal{A} .

Definition 5. We say that the linear operator \mathcal{A} satisfies (r, δ_r) -RIP if for any matrix $X \in \mathbb{R}^{n \times n}$ with $\text{rank}(X) \leq r$,

$$(1 - \delta_r) \|X\|_F^2 \leq \frac{1}{m} \sum_{i=1}^m \langle A_i, X \rangle^2 \leq (1 + \delta_r) \|X\|_F^2.$$

The above definition has played an important role in the literature of low-rank matrix recovery. One well-known case where the RIP holds is when \mathcal{A} is a random measurement operator. For example, if $\{A_i\}_{i=1}^m$ are mutually independent random Gaussian matrices, then when $m \geq Dnr$, \mathcal{A} satisfies the RIP for some $\delta_r < 1$ with probability at least $1 - C \exp(-dm)$, where C, D, d are absolute positive scalars [8, Theorem 2.3].

We now state the main result of this subsection.

Theorem 4. Suppose that \mathcal{A} satisfies $(2r, \delta_{2r})$ -RIP with $\delta_{2r} < \frac{1}{10}$. Then, with any arbitrary initialization, the sequence of iterates $\{U^k\}_{k \geq 0}$ generated by Algorithm 1 for solving (58) converges at least Q -quadratically to an element in \mathcal{U} .

The rest of this subsection is devoted to proving Theorem 4. Before we proceed, let us introduce some notations and preliminaries. Since $X^\star \succeq 0$ and $\text{rank}(X^\star) = r$, we have $\lambda_1(X^\star) \geq \dots \geq \lambda_r(X^\star) > 0$. As a result, the singular values of any $U \in \mathcal{U}$ are $\{\sqrt{\lambda_i(X^\star)}\}_{i=1}^r$. Let $\nabla f(U) \in \mathbb{R}^{n \times r}$ and $\nabla^2 f(U) \in \mathbb{R}^{(nr) \times (nr)}$ be the gradient and Hessian of f at U , respectively. For problem (58), a routine calculation gives

$$\nabla f(U) = \frac{1}{m} \sum_{i=1}^m \langle A_i, UU^T - X^\star \rangle A_i U \quad (59)$$

and

$$\begin{aligned} \text{vec}(Z)^T \nabla^2 f(U) \text{vec}(Z) &= \left\langle Z, \lim_{\tau \rightarrow 0} \frac{\nabla f(U + \tau Z) - \nabla f(U)}{\tau} \right\rangle \\ &= \frac{1}{m} \sum_{i=1}^m 2 \langle A_i, UZ^T \rangle^2 + \langle A_i, UU^T - X^\star \rangle \langle A_i, ZZ^T \rangle, \end{aligned} \quad (60)$$

where $\text{vec}(Z) \in \mathbb{R}^{nr}$ is the vector obtained by stacking the columns of Z . The following result, which is stated in [4, Lemma 4.1] and is related to [6, Lemma 2.1], is crucial to our analysis.

Fact 8. *For any $X, Y \in \mathbb{R}^{n \times n}$ with $\text{rank}(X), \text{rank}(Y) \leq r$, if \mathcal{A} is $(2r, \delta_{2r})$ -RIP, then it holds that*

$$\left| \frac{1}{m} \sum_{i=1}^m \langle A_i, X \rangle \langle A_i, Y \rangle - \langle X, Y \rangle \right| \leq \delta_{2r} \|X\|_F \|Y\|_F.$$

5.2.1 Second-Order Critical Points and Local EB Condition

We first show that under the RIP, the set of second-order critical points of f is \mathcal{U} . Moreover, we show that in a neighborhood of \mathcal{U} , the local EB condition (13) holds. The following result is due to [4, Theorem 3.2].

Fact 9. *Suppose that \mathcal{A} satisfies $(2r, \delta_{2r})$ -RIP with $\delta_{2r} < \frac{1}{10}$. Then, for any $U \in \mathbb{R}^{n \times r}$ such that $\nabla f(U) = 0$ and $UU^T \neq X^\star$, it holds that*

$$\lambda_{\min}(\nabla^2 f(U)) \leq -\frac{\lambda_r(X^\star)}{5} < 0.$$

Proposition 5. *Suppose that \mathcal{A} satisfies $(2r, \delta_{2r})$ -RIP with $\delta_{2r} < \frac{1}{10}$. Then, the following statements hold.*

(i) \mathcal{U} equals the set of second-order critical points of f .

(ii) The following error bound holds:

$$\text{dist}(U, \mathcal{U}) \leq \frac{2}{\lambda_r(X^\star)} \|\nabla f(U)\| \quad \text{whenever} \quad \text{dist}(U, \mathcal{U}) \leq \frac{1}{3} \sqrt{\lambda_r(X^\star)}. \quad (61)$$

Proof. By the global optimality of \mathcal{U} , any $U \in \mathcal{U}$ is a second-order critical point of f . On the other hand, due to Fact 9, any $U \notin \mathcal{U}$ cannot be a second-order critical point of f if \mathcal{A} satisfies $(2r, \delta_{2r})$ -RIP with $\delta_{2r} < \frac{1}{10}$. Therefore, the result in (i) holds.

We next prove (ii). Let \hat{U} be the projection of U to \mathcal{U} and $\Delta := U - \hat{U}$. Clearly, $\text{dist}(U, \mathcal{U}) = \|\Delta\|_F$. By the definition of \mathcal{U} , we have $\hat{U} = U^\star \bar{Q}$, where $\bar{Q} = \arg\min_{Q \in \mathcal{O}^r} \|U - U^\star Q\|^2$. Let $U^{\star T} U = P \Sigma R^T$ be the singular value decomposition of $U^{\star T} U$, where $\Sigma \in \mathbb{R}^{r \times r}$ is diagonal and $P, R \in \mathcal{O}^r$. Then, one can verify that $\bar{Q} = PR^T$. Hence, it follows that

$$\Delta^T \hat{U} = (U - U^\star \bar{Q})^T U^\star \bar{Q} = R \Sigma R^T - \hat{U}^T \hat{U} = \hat{U}^T \Delta.$$

Using $\mathcal{A}(X^*) = \mathcal{A}(\hat{U}\hat{U}^T) = b$ and (59), one has $\nabla f(U) = \frac{1}{m} \sum_{i=1}^m \langle A_i, UU^T - \hat{U}\hat{U}^T \rangle A_i U$. It then follows from Fact 8 that

$$\begin{aligned} \langle \nabla f(U), \Delta \rangle &= \frac{1}{m} \sum_{i=1}^m \langle A_i, UU^T - \hat{U}\hat{U}^T \rangle \langle A_i, \Delta U^T \rangle \\ &\geq \langle UU^T - \hat{U}\hat{U}^T, \Delta U^T \rangle - \delta_{2r} \|UU^T - \hat{U}\hat{U}^T\|_F \|\Delta U^T\|_F. \end{aligned} \quad (62)$$

Using $U = \Delta + \hat{U}$ and $\Delta^T \hat{U} = \hat{U}^T \Delta$, we obtain

$$\begin{aligned} \langle UU^T - \hat{U}\hat{U}^T, \Delta U^T \rangle &= \langle \Delta\Delta^T + \hat{U}\Delta^T + \Delta\hat{U}^T, \Delta\Delta^T + \Delta\hat{U}^T \rangle \\ &= \langle \Delta\Delta^T, \Delta\Delta^T \rangle + 3\langle \hat{U}\Delta^T, \Delta\Delta^T \rangle + 2\|\hat{U}\Delta^T\|_F^2 \\ &\geq \|\Delta\Delta^T\|_F^2 - 3\|\hat{U}\Delta^T\|_F \|\Delta\Delta^T\|_F + 2\|\hat{U}\Delta^T\|_F^2. \end{aligned}$$

Also, we have

$$\begin{aligned} \|UU^T - \hat{U}\hat{U}^T\|_F \|\Delta U^T\|_F &= \|\Delta\Delta^T + \hat{U}\Delta^T + \Delta\hat{U}^T\|_F \|\Delta\Delta^T + \Delta\hat{U}^T\|_F \\ &\leq \|\Delta\Delta^T\|_F^2 + 3\|\hat{U}\Delta^T\|_F \|\Delta\Delta^T\|_F + 2\|\hat{U}\Delta^T\|_F^2. \end{aligned}$$

Hence, it follows from (62) that

$$\begin{aligned} \langle \nabla f(U), \Delta \rangle &\geq (1 - \delta_{2r}) \|\Delta\Delta^T\|_F^2 - 3(1 + \delta_{2r}) \|\hat{U}\Delta^T\|_F \|\Delta\Delta^T\|_F + 2(1 - \delta_{2r}) \|\hat{U}\Delta^T\|_F^2 \\ &\geq \|\hat{U}\Delta^T\|_F \left(2(1 - \delta_{2r}) \|\hat{U}\Delta^T\|_F - 3(1 + \delta_{2r}) \|\Delta\Delta^T\|_F \right), \end{aligned} \quad (63)$$

where the second inequality uses $\delta_{2r} < 1$. Since the smallest singular value of \hat{U} is $\sqrt{\lambda_r(X^*)}$, it holds that $\|\hat{U}\Delta^T\|_F \geq \sqrt{\lambda_r(X^*)} \|\Delta\|_F$. This, together with $\|\Delta\|_F \leq \frac{1}{3} \sqrt{\lambda_r(X^*)}$ and $\delta_{2r} < \frac{1}{10}$, gives

$$\begin{aligned} 2(1 - \delta_{2r}) \|\hat{U}\Delta^T\|_F - 3(1 + \delta_{2r}) \|\Delta\Delta^T\|_F &\geq 2(1 - \delta_{2r}) \sqrt{\lambda_r(X^*)} \|\Delta\|_F - (1 + \delta_{2r}) \sqrt{\lambda_r(X^*)} \|\Delta\|_F \\ &\geq \frac{\sqrt{\lambda_r(X^*)}}{2} \|\Delta\|_F. \end{aligned}$$

Substituting this into (63) and using $\|\hat{U}\Delta^T\|_F \geq \sqrt{\lambda_r(X^*)} \|\Delta\|_F$, we obtain

$$\langle \nabla f(U), \Delta \rangle \geq \frac{\lambda_r(X^*)}{2} \|\Delta\|_F^2,$$

which, together with the Cauchy-Schwarz inequality, implies the required error bound (61). \square

5.2.2 Lipschitz Continuity of $\nabla^2 f$

We next verify the Lipschitz continuity of $\nabla^2 f$.

Proposition 6. *Suppose that \mathcal{A} satisfies $(2r, \delta_{2r})$ -RIP with $\delta_{2r} < \frac{1}{10}$. Then, for any $R > 0$, $\nabla^2 f$ is Lipschitz continuous on $\mathbb{B}(0; R) = \{U \in \mathbb{R}^{n \times r} : \|U\|_F \leq R\}$ with Lipschitz constant $L = 5R$.*

Proof. Let $U, U' \in \mathbb{B}(0; R)$. Hence, $\|U\|_F \leq R$ and $\|U'\|_F \leq R$. By (60) and Fact 8, we obtain

$$\begin{aligned} \|\nabla^2 f(U) - \nabla^2 f(U')\| &= \max_{\|Z\|_F=1} |\text{vec}(Z)^T (\nabla^2 f(U) - \nabla^2 f(U')) \text{vec}(Z)| \\ &= \max_{\|Z\|_F=1} \left| \frac{1}{m} \sum_{i=1}^m 2\langle A_i, UZ^T \rangle^2 - 2\langle A_i, U'Z^T \rangle^2 + \langle A_i, UU^T - U'U'^T \rangle \langle A_i, ZZ^T \rangle \right| \\ &\leq (1 + \delta_{2r}) \cdot \max_{\|Z\|_F=1} \left[\|(U + U')Z^T\|_F \|(U - U')Z^T\|_F + \|UU^T - U'U'^T\|_F \|ZZ^T\|_F \right] \\ &\leq (1 + \delta_{2r}) \cdot [(\|U\|_F + \|U'\|_F) \|(U - U')\|_F + \|(U - U')U^T + U'(U - U')^T\|_F] \\ &\leq (1 + \delta_{2r}) \cdot [(\|U\|_F + \|U'\|_F) \|(U - U')\|_F + \|U - U'\|_F \|U\|_F + \|U'\|_F \|U - U'\|_F] \\ &\leq 4(1 + \delta_{2r})R \cdot \|U - U'\|_F \leq 5R \cdot \|U - U'\|_F, \end{aligned}$$

where we use $\delta_{2r} < \frac{1}{10}$ in the last inequality. The proof is then completed. \square

5.3 Proof of Theorem 4

In view of Fact 2, it suffices to prove that under the RIP assumption in Theorem 4, the following statements hold: (i) Assumption 1 holds, (ii) Assumption 2 holds, (iii) $\mathcal{L}(f(U^k))$ is bounded for some $k \geq 0$, and (iv) \mathcal{U} equals the set of second-order critical points of f .

Suppose that $U \in \mathbb{R}^{n \times r}$ is the initial point of Algorithm 1. Define

$$\bar{R} := \left(10r f(U^0) + \frac{3r}{m} \|b\|^2 \right)^{\frac{1}{4}} > 0 \quad \text{and} \\ \mathcal{F} := \mathbb{B}(0; 2\bar{R}) = \{U \in \mathbb{R}^{n \times r} : \|U\|_F \leq 2\bar{R}\}.$$

Let $U \in \mathbb{R}^{n \times r}$ be an arbitrary point in $\mathcal{L}(f(U^0))$. Let $\lambda_1 \geq \dots \geq \lambda_r \geq 0$ be the eigenvalues of $U^T U$. Then, it holds that $\|UU^T\|_F^2 = \|U^T U\|_F^2 = \lambda_1^2 + \dots + \lambda_r^2$ and $\|U\|_F^2 = \lambda_1 + \dots + \lambda_r$. Hence, we obtain $\|U\|_F^4 \leq r \|UU^T\|_F^2$. This, together with Definition 5 and $f(U) \leq f(U^0)$, yields

$$\begin{aligned} \|U\|_F^4 &\leq r \|UU^T\|_F^2 \leq r \cdot \frac{1}{(1 - \delta_{2r})m} \|\mathcal{A}(UU^T)\|^2 \\ &\leq \frac{2r}{(1 - \delta_{2r})m} (\|\mathcal{A}(UU^T) - b\|^2 + \|b\|^2) \\ &\leq \frac{8r}{1 - \delta_{2r}} f(U^0) + \frac{2r}{(1 - \delta_{2r})m} \|b\|^2 \leq \bar{R}^4, \end{aligned}$$

where we use $\delta_{2r} < \frac{1}{10}$ in the last inequality. By the definition of \mathcal{F} , we have $\mathcal{L}(f(U^0)) \subseteq \text{int}(\mathcal{F})$ and $\mathcal{L}(f(U^0))$ is bounded. In addition, by Proposition 6, $\nabla^2 f$ is Lipschitz continuous on \mathcal{F} with Lipschitz constant $L = 10\bar{R}$. Hence, statements (i) and (iii) above hold. Note that Proposition 5 implies that statements (ii) and (iv) hold. The proof is then completed.

6 Numerical Experiments

In this section, we apply the CR method to solve nonconvex minimization problems considered in Section 5. Our primary goal is to validate Theorems 3 and 4, which concern the global convergence of Algorithm 1 to the target signals and its local quadratic convergence rate. All experiments are coded in Matlab and run on a Dell desktop with a 3.40-GHz Intel Core i7-3770 processor and 16 GB of RAM. The code to reproduce all the figures and numerical results in this section can be found online: https://github.com/ZiruiZhou/cubicreg_app.git.

6.1 Phase Retrieval

Our setup of the experiments for phase retrieval is as follows. We first generate a complex signal $z^* \in \mathbb{C}^n$ from the standard n -dimensional complex Gaussian distribution $\mathbb{CN}(0, I_n)$, which is considered to be our target signal. Next, we generate the measurement vectors $\{a_j\}_{j=1}^m$ independently and identically from $\mathbb{CN}(0, I_n)$ and compute the measurements $\{b_j\}_{j=1}^m$ by assigning $b_j = |a_j^H z^*|$ for each j . Here m is chosen to be $m = \lceil 3n \log^3(n) \rceil$, which empirically guarantees that the event in Theorem 3 holds with overwhelming probability. In addition, the set of target signals is given by $\mathcal{X}^* = \{(x^* \cos \phi - y^* \sin \phi, x^* \sin \phi + y^* \cos \phi) : \phi \in [0, 2\pi)\}$, where x^* and y^* are the real and imaginary parts of z^* , respectively.

We then apply Algorithm 1 to solve the resulting optimization problem (48). For the initial point (x^0, y^0) , we draw the entries of x^0 and y^0 independently and identically from the uniform distribution on the interval $[-5, 5]$. In the k -th iteration of Algorithm 1, we compute the relative error (RE) of the iterate (x^k, y^k) , which is defined as

$$\text{RE}(k) = \frac{\text{dist}((x^k, y^k), \mathcal{X}^*)}{\|(x^k, y^k)\|}.$$

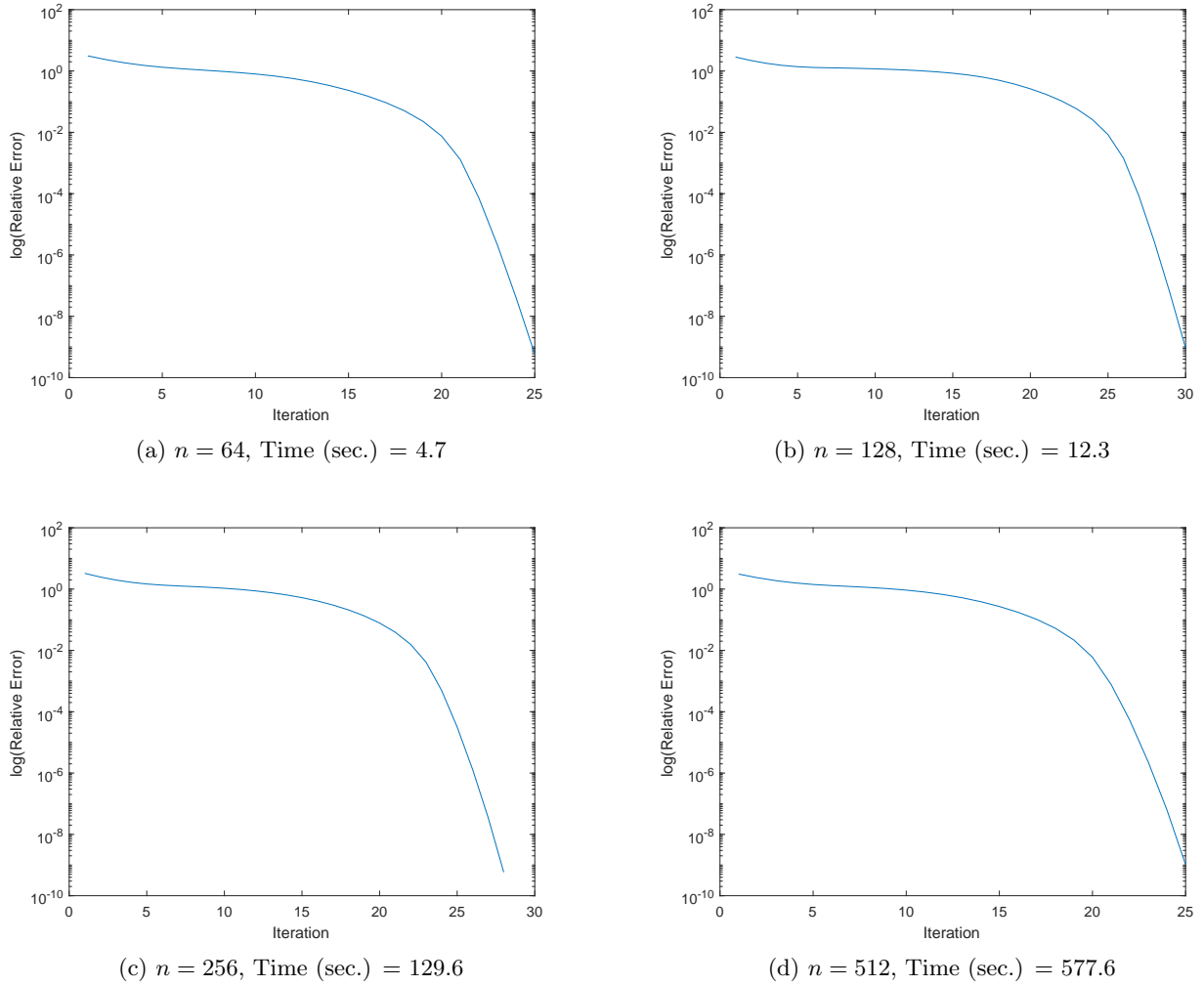


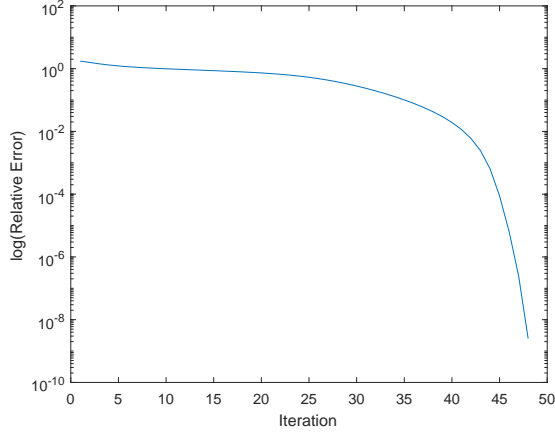
Figure 1: Algorithm 1 for the phase retrieval problem.

Moreover, we terminate Algorithm 1 when $\text{RE}(k) < 10^{-8}$. It follows from (43) and the proof of Theorem 3 that with overwhelming probability, $\{\text{RE}(k)\}_{k \geq 0}$ converges to 0 and the local convergence rate is at least quadratic.⁵ To validate such a result, we present the logarithm of $\{\text{RE}(k)\}_{k \geq 0}$ against the number of iterations in Figure 1. Also, the time for reaching a required solution is recorded. It is clear from Figure 1 that $\{\text{RE}(k)\}_{k \geq 0}$ converges to 0 and in the final stages of the algorithm, the convergence rate is at least superlinear.

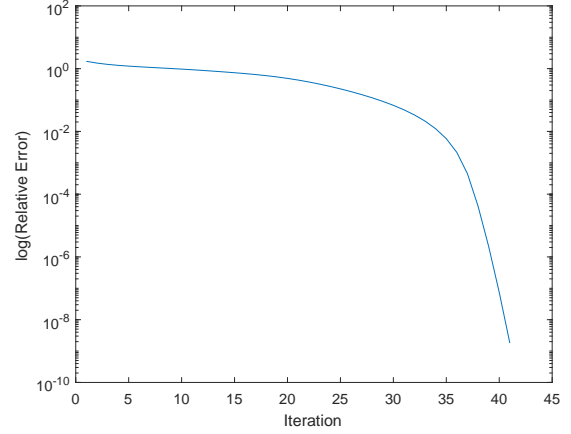
6.2 Low-Rank Matrix Recovery

Our setup of the experiments for low-rank matrix recovery is as follows. First, we generate a positive semidefinite matrix $X^* \in \mathbb{R}^{n \times n}$ with $\text{rank}(X^*) = r$. In particular, we generate a matrix $U^* \in \mathbb{R}^{n \times r}$ with its entries drawn independently and identically from the standard Gaussian distribution $\mathcal{N}(0, 1)$ and set $X^* = U^* U^{*T}$. Second, we generate the matrices $\{A_j\}_{j=1}^m \subseteq \mathbb{R}^{n \times n}$ that form the linear operator $\mathcal{A} : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^m$. For all $i = 1, \dots, m$, entries of A_i are drawn independently and identically from $\mathcal{N}(0, 1)$. Here m is chosen to be $m = 3nr$, which empirically guarantees that the event in Theorem 4 holds with overwhelming probability. Finally, we compute the measurements $\{b_j\}_{j=1}^m$ by assigning

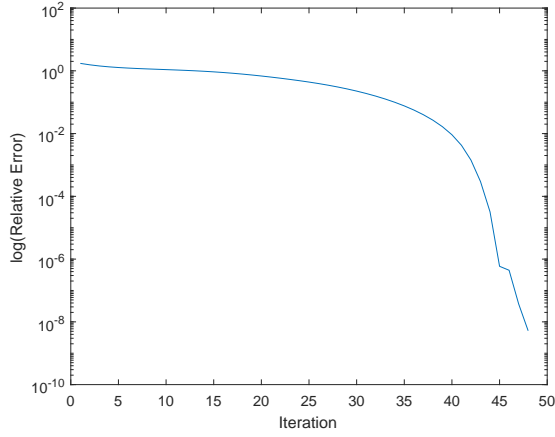
⁵Since the limit point of $\{(x^k, y^k)\}_{k \geq 0}$ is not known a priori, we use $\{\text{RE}(k)\}_{k \geq 0}$ to verify the local convergence rate of Algorithm 1.



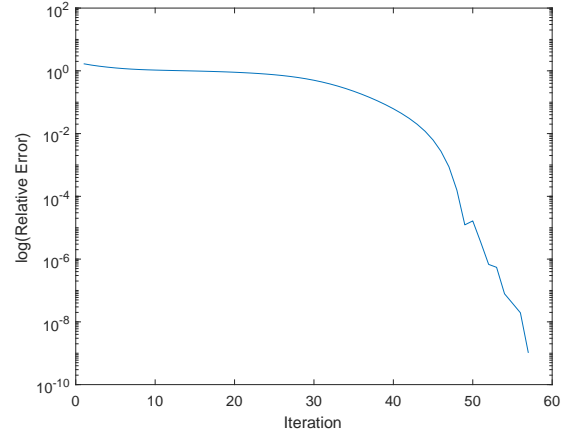
(a) $n = 32, r = 6$, Time (sec.) = 9.9



(b) $n = 64, r = 4$, Time (sec.) = 11.9



(c) $n = 128, r = 6$, Time (sec.) = 455.0



(d) $n = 256, r = 8$, Time (sec.) = 5017.6

Figure 2: Algorithm 1 for the low-rank matrix recovery problem.

$b_j = \langle A_j, X^* \rangle$ for all j . In addition, the set of target matrices is given by $\mathcal{U} = \{U^*Q : Q \in \mathcal{O}^r\}$.

We then apply Algorithm 1 to solve the resulting optimization problem (58). We use a random matrix $U_0 \in \mathbb{R}^{n \times r}$, whose entries are drawn independently and identically from the uniform distribution on the interval $[-5, 5]$, as the initial point. In the k -th iteration of Algorithm 1, we compute the relative error (RE) of the iterate U^k , which is defined as

$$\text{RE}(k) = \frac{\text{dist}(U^k, \mathcal{U})}{\|U^*\|_F}.$$

Moreover, we terminate Algorithm 1 when $\text{RE}(k) < 10^{-8}$. It follows from (43) and the proof of Theorem 4 that with overwhelming probability, $\{\text{RE}(k)\}_{k \geq 0}$ converges to 0 and the local convergence rate is at least quadratic.⁶ To validate such a result, we present the logarithm of $\{\text{RE}(k)\}_{k \geq 0}$ against the number of iterations in Figure 2. Also, the time for reaching a required solution is recorded. It is clear from Figure 2 that $\{\text{RE}(k)\}_{k \geq 0}$ converges to 0 and in the final stages of the algorithm, the convergence rate is at least superlinear.

⁶Since the limit point of $\{U^k\}_{k \geq 0}$ is not known a priori, we use $\{\text{RE}(k)\}_{k \geq 0}$ to verify the local convergence rate of Algorithm 1.

7 Conclusions

In this paper we established the quadratic convergence of the CR method under a local EB condition, which is much weaker a requirement than the non-degeneracy condition used in previous works. This indicates that adding cubic regularization not only equips Newton's method with remarkable global convergence properties but also enables it to converge quadratically even in the presence of degenerate solutions. As a byproduct, we showed that without assuming convexity, the proposed EB condition is equivalent to a quadratic growth condition, which could be of independent interest. In addition, we studied the CR method for solving two concrete nonconvex optimization problems that arise in phase retrieval and low-rank matrix recovery. We proved that with overwhelming probability, the sequence of iterates generated by the CR method for solving these two problems converges at least Q-quadratically to a global minimizer. Numerical results of the CR method for solving these two problems corroborated our theoretical findings.

Our proof of the quadratic convergence of the CR method is not a direct extension of those for other regularized Newton-type methods. The fact that any accumulation point of the sequence generated by the CR method is a second-order critical point plays a key role in our analysis. We believe that similar approaches could be employed for analyzing the local convergence of other iterative algorithms that cluster at second-order critical points, such as trust-region methods.

References

- [1] F. J. ARAGÓN ARTACHO AND M. H. GEOFFROY, *Characterization of Metric Regularity of Subdifferentials*, J. Convex Anal., 15 (2008), pp. 365–380.
- [2] S. BELLAVIA AND B. MORINI, *Strong Local Convergence Properties of Adaptive Regularized Methods for Nonlinear Least Squares*, IMA J. Numer. Anal., 35 (2014), pp. 947–968.
- [3] R. BHATIA, *Matrix Analysis*, vol. 169 of Graduate Texts in Mathematics, Springer–Verlag New York, Inc., New York, 1997.
- [4] S. BHOJANAPALLI, B. NEYSHABUR, AND N. SREBRO, *Global Optimality of Local Search for Low Rank Matrix Recovery*, in Advances in Neural Information Processing Systems (NIPS), 2016, pp. 3873–3881.
- [5] J. BOLTE, T. P. NGUYEN, J. PEYPOUQUET, AND B. W. SUTER, *From Error Bounds to the Complexity of First-Order Descent Methods for Convex Functions*, Math. Program., 165 (2017), pp. 471–507.
- [6] E. J. CANDÈS, *The Restricted Isometry Property and Its Implications for Compressed Sensing*, C. R. Acad. Sci. Paris, 346 (2008), pp. 589–592.
- [7] E. J. CANDÈS, X. LI, AND M. SOLTANOLKOTABI, *Phase Retrieval via Wirtinger Flow: Theory and Algorithms*, IEEE Trans. Inform. Theory, 61 (2015), pp. 1985–2007.
- [8] E. J. CANDÈS AND Y. PLAN, *Tight Oracle Inequalities for Low-Rank Matrix Recovery from a Minimal Number of Noisy Random Measurements*, IEEE Trans. Inform. Theory, 57 (2011), pp. 2342–2359.
- [9] Y. CARMON AND J. C. DUCHI, *Gradient Descent Efficiently Finds the Cubic-Regularized Non-Convex Newton Step*, arXiv preprint arXiv:1612.00547, (2016).
- [10] C. CARTIS, N. I. M. GOULD, AND PH. L. TOINT, *On the Complexity of Steepest Descent, Newton's and Regularized Newton's Methods for Nonconvex Unconstrained Optimization Problems*, SIAM J. Optim., 20 (2010), pp. 2833–2852.

- [11] C. CARTIS, N. I. M. GOULD, AND PH. L. TOINT, *Adaptive Cubic Regularisation Methods for Unconstrained Optimization. Part I: Motivation, Convergence and Numerical Results*, Math. Program., 127 (2011), pp. 245–295.
- [12] C. CARTIS, N. I. M. GOULD, AND PH. L. TOINT, *Adaptive Cubic Regularisation Methods for Unconstrained Optimization. Part II: Worst-Case Function-and Derivative-Evaluation Complexity*, Math. Program., 130 (2011), pp. 295–319.
- [13] D. DRUSVYATSKIY, A. D. IOFFE, AND A. S. LEWIS, *Nonsmooth Optimization Using Taylor-Like Models: Error Bounds, Convergence, and Termination Criteria*, arXiv preprint arXiv:1610.03446, (2016).
- [14] D. DRUSVYATSKIY, B. S. MORDUKHOVICH, AND T. T. NGHIA, *Second-Order Growth, Tilt Stability, and Metric Regularity of the Subdifferential*, J. Convex Anal., 21 (2014), pp. 1165–1192.
- [15] J.-Y. FAN AND Y.-X. YUAN, *On the Quadratic Convergence of the Levenberg-Marquardt Method without Nonsingularity Assumption*, Computing, 74 (2005), pp. 23–39.
- [16] A. FISCHER, *Local Behavior of an Iterative Framework for Generalized Equations with Nonisolated Solutions*, Math. Program., 94 (2002), pp. 91–124.
- [17] A. GRIEWANK, *The Modification of Newton’s method for Unconstrained Optimization by Bounding Cubic Terms*, tech. report, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, 1981.
- [18] P. HARTMAN, *Ordinary Differential Equations*, Society for Industrial and Applied Mathematics, 2002.
- [19] K. LEVENBERG, *A Method for the Solution of Certain Non-Linear Problems in Least Squares*, Quart. Appl. Math., 2 (1944), pp. 164–168.
- [20] D.-H. LI, M. FUKUSHIMA, L. QI, AND N. YAMASHITA, *Regularized Newton Methods for Convex Minimization Problems with Singular Solutions*, Comput. Optim. Appl., 28 (2004), pp. 131–147.
- [21] G. LI AND T. K. PONG, *Calculus of the Exponent of Kurdyka–Łojasiewicz Inequality and Its Applications to Linear Convergence of First-Order Methods*. Accepted for publication in *Found. Comput. Math.*, 2017.
- [22] S. ŁOJASIEWICZ, *Sur les trajectoires du gradient d’une fonction analytique*, Seminari di Geometria, 1983 (1982), pp. 115–117.
- [23] D. R. LUKE, *Phase Retrieval, What’s New*, SIAG/OPT Views and News, 25 (2017), pp. 1–5.
- [24] Z.-Q. LUO AND P. TSENG, *Error Bounds and Convergence Analysis of Feasible Descent Methods: A General Approach*, Ann. Oper. Res., 46 (1993), pp. 157–178.
- [25] D. W. MARQUARDT, *An Algorithm for Least-Squares Estimation of Nonlinear Parameters*, J. Soc. Ind. Appl. Math., 11 (1963), pp. 431–441.
- [26] B. S. MORDUKHOVICH, *Variational Analysis and Generalized Differentiation I: Basic Theory*, vol. 330, Springer Science & Business Media, 2006.
- [27] YU. NESTEROV AND B. T. POLYAK, *Cubic Regularization of Newton Method and its Global Performance*, Math. Program., 108 (2006), pp. 177–205.
- [28] J. NOCEDAL AND S. J. WRIGHT, *Numerical Optimization*, Springer, 2006.

- [29] R. T. ROCKAFELLAR AND R. J.-B. WETS, *Variational Analysis*, vol. 317, Springer-Verlag Berlin Heidelberg, 1998.
- [30] Y. SHECHTMAN, Y. C. ELDAR, O. COHEN, H. N. CHAPMAN, J. MIAO, AND M. SEGEV, *Phase Retrieval with Application to Optical Imaging: A Contemporary Overview*, IEEE Sig. Process. Mag., 32 (2015), pp. 87–109.
- [31] J. SUN, Q. QU, AND J. WRIGHT, *A Geometric Analysis of Phase Retrieval*. Accepted for publication in *Found. Comput. Math.*, 2017.
- [32] PH. L. TOINT, *Nonlinear Step-size Control, Trust Regions and Regularizations for Unconstrained Optimization*, Optim. Methods Softw., 28 (2013), pp. 82–95.
- [33] P. TSENG, *Error Bounds and Superlinear Convergence Analysis of Some Newton-Type Methods in Optimization*, in Nonlinear Optimization and Related Topics, G. D. Pillo and F. Giannessi, eds., Springer US, 2000, pp. 445–462.
- [34] N. YAMASHITA AND M. FUKUSHIMA, *On the Rate of Convergence of the Levenberg-Marquardt Method*, in Topics in Numerical Analysis, G. Alefeld and X. Chen, eds., vol. 15, Springer, Vienna, 2001, pp. 239–249.
- [35] Y.-X. YUAN, *Recent Advances in Trust Region Algorithms*, Math. Program., 151 (2015), pp. 249–281.
- [36] M.-C. YUE, Z. ZHOU, AND A. M.-C. SO, *A Family of Inexact SQA Methods for Non-Smooth Convex Minimization with Provable Convergence Guarantees Based on the Luo-Tseng Error Bound Property*, to appear in Math. Program., (2018).
- [37] Y. ZHOU AND Y. LIANG, *Characterization of Gradient Dominance and Regularity Conditions for Neural Networks*, arXiv preprint arXiv:1710.06910, (2017).
- [38] Z. ZHOU AND A. M.-C. SO, *A Unified Approach to Error Bounds for Structured Convex Optimization Problems*, Math. Program., 165 (2017), pp. 689–728.