



# A cubic regularization of Newton's method with finite difference Hessian approximations

G. N. Grapiglia<sup>1</sup> · M. L. N. Gonçalves<sup>2</sup> · G. N. Silva<sup>3</sup>

Received: 12 May 2021 / Accepted: 11 September 2021 / Published online: 13 October 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

## Abstract

In this paper, we present a version of the cubic regularization of Newton's method for unconstrained nonconvex optimization, in which the Hessian matrices are approximated by forward finite difference Hessians. The regularization parameter of the cubic models and the accuracy of the Hessian approximations are jointly adjusted using a nonmonotone line search criterion. Assuming that the Hessian of the objective function is globally Lipschitz continuous, we show that the proposed method needs at most  $\mathcal{O}(n\epsilon^{-3/2})$  function and gradient evaluations to generate an  $\epsilon$ -approximate stationary point, where  $n$  is the dimension of the domain of the objective function. Preliminary numerical results corroborate our theoretical findings.

**Keywords** Nonconvex optimization · Second-order methods · Finite differences · Worst-case complexity

## 1 Introduction

The cubic regularization of Newton's method (CNM) is a globally convergent variant of Newton's method for unconstrained minimization of twice continuously

---

✉ G. N. Grapiglia  
grapiglia@ufpr.br

M. L. N. Gonçalves  
maxlng@ufg.br

G. N. Silva  
gilson.silva@ufob.edu.br

<sup>1</sup> Departamento de Matemática, Universidade Federal do Paraná, Centro Politécnico, Cx. postal 19.081, 81531-980, Curitiba, PR, Brazil

<sup>2</sup> IME, Universidade Federal de Goiás, Campus II- Cx. postal 131, 74001-970, Goiânia, GO, Brazil

<sup>3</sup> Colegiado de Matemática, Universidade Federal do Oeste da Bahia, 47808-021, Barreiras, BA, Brazil

differentiable functions [16, 22]. At the  $t$ -th iteration of CNM, the next iterate  $x_{t+1}$  is obtained by minimizing a *cubic model* that consists of a third-order regularization of the second-order Taylor approximation of the objective function  $f(\cdot)$  around  $x_t$ . For a given tolerance  $\epsilon > 0$ , Nesterov and Polyak [22] proved that the CNM takes at most  $\mathcal{O}(\epsilon^{-3/2})$  iterations to generate an  $\epsilon$ -approximate stationary point of the objective function (i.e., an iterate  $x_t$  such that  $\|\nabla f(x_t)\|_2 \leq \epsilon$ ), when  $f(\cdot)$  is a nonconvex function with Lipschitz continuous Hessian. Remarkably, Cartis, Gould and Toint [7] proved that in the same problem class the standard Newton's method (without regularization) may need a number of iterations arbitrarily close to  $\mathcal{O}(\epsilon^{-2})$  to generate an  $\epsilon$ -approximate stationary point of the objective function. In view of these complexity results, several second-order methods inspired by the CNM have been proposed for nonconvex optimization in the last decade (see, e.g., [2, 4, 9, 11, 13, 14, 19]). More recently, Carmon et al. [6] showed that the worst-case complexity bound of  $\mathcal{O}(\epsilon^{-3/2})$  is the best that second-order methods can achieve when applied to functions with Lipschitz continuous Hessians, establishing the optimality of the CNM in this problem class.

As aforementioned, the iterates in the CNM are computed by minimizing cubic models, which involve Hessian matrices of the objective function. However, in several applications the computation of Hessian matrices can be computationally very expensive (see, e.g., [20, 24]). For the case in which only function values and first-order derivatives are provided by the user, Cartis, Gould and Toint [10] analyzed a variant of the CNM with Hessian matrices  $\nabla^2 f(x_t)$  approximated by forward finite difference Hessians  $B_t$ . Assuming that the gradient and the Hessian of the objective function are Lipschitz continuous on the path of iterates, and that the gradient is bounded over the iterates, they showed that the referred method (called ARC-FDH) needs at most  $\mathcal{O}(n\epsilon^{-3/2} + n|\log(\epsilon)|)$  calls of the oracle<sup>1</sup> to generate an  $\epsilon$ -approximate stationary point, where  $n$  is the dimension of the domain of the objective function. Regarding the Hessian approximations, the key condition required in [10] is that

$$\|\nabla^2 f(x_t) - B_t\| \leq \kappa_B \|x_{t+1} - x_t\|, \quad (1)$$

where  $\kappa_B \geq 0$ . Since  $x_{t+1}$  is unknown during the computation of  $B_t$ , the ARC-FDH is endowed with an adaptive procedure in which the stepsize that defines the finite difference approximation  $B_t$  is reduced until a sufficient condition for (1) is satisfied. This is the source of the additional logarithmic term in the complexity bound for the number of calls of the oracle (Lemma 3.2 in [10]).

In this paper, we present a new first-order version of the CNM with Hessian matrices approximated by forward finite difference Hessians. Different from [10], the stepsize that defines the finite difference approximation  $B_t$  is adjusted aiming the condition

$$\|\nabla^2 f(x_t) - B_t\| \leq \kappa_B \|x_t - x_{t-1}\|. \quad (2)$$

The use of (2) instead of (1) was first suggested by Kohler and Lucchi [18] and further investigated (theoretically and numerically) by Wang et al. [26, 27], and by Bellavia, Gurioli and Morini [1] in the context of sub-sampled variants of the CNM

<sup>1</sup>Throughout this paper, by *call of the oracle* we mean one function evaluation or one gradient evaluation.

for finite-sum minimization<sup>2</sup>. Based on condition (2), in our method, the regularization parameter of the cubic models and the accuracy of the Hessian approximations are jointly adjusted using a nonmonotone line search criterion. Assuming that the Hessian of the objective function is globally Lipschitz continuous, we show that the proposed method needs at most  $\mathcal{O}(n\epsilon^{-3/2})$  calls of the oracle to generate an  $\epsilon$ -approximate stationary point. Moreover, we also show that the method needs at most  $\mathcal{O}\left(n \max\left\{\epsilon_g^{-3/2}, \epsilon_H^{-3}\right\}\right)$  calls of the oracle to generate an  $(\epsilon_g, \epsilon_H)$ -approximate second-order stationary point, i.e., an iterate  $x_t$  such that

$$\|\nabla f(x_t)\| \leq \epsilon_g \quad \text{and} \quad \lambda_{\min}\left(\nabla^2 f(x_t)\right) \geq -\epsilon_H.$$

The paper is organized as follows. In Section 2, we formulate the problem and establish the crucial auxiliary results. In Section 3, we present our first-order CNM variant and analyze its worst-case evaluation complexity. Finally, in Section 4, we report preliminary numerical results that corroborate our theoretical findings.

**Notation** The symbol  $\|\cdot\|$  denotes the 2-norm for vectors or matrices (depending on the context), while  $\|\cdot\|_F$  denotes the Frobenius norm for matrices. The Euclidean inner product of  $x, y \in \mathbb{R}^n$  is denoted by  $\langle x, y \rangle$ . For  $j = 1, \dots, n$ ,  $e_j \in \mathbb{R}^n$  is the  $j$ -th orthonormal vector of the canonical basis for  $\mathbb{R}^n$ . We denote the identity matrix of  $\mathbb{R}^n \times \mathbb{R}^n$  by  $I$ , and for any symmetric matrix  $A \in \mathbb{R}^{n \times n}$ ,  $\lambda_{\min}(A)$  is the smallest eigenvalue of  $A$ . Given two square matrices  $A$  and  $B$ , the inequality  $A \succeq B$  means that the matrix  $A - B$  is positive semidefinite.

## 2 Problem formulation and auxiliary results

We consider the unconstrained minimization problem

$$\min_{x \in \mathbb{R}^n} f(x), \tag{3}$$

where  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is a twice continuously differentiable function, potentially nonconvex. Our analysis will be carried out under the following assumptions:

**A1** The Hessian of  $f$  is  $L$ -Lipschitz continuous on the whole  $\mathbb{R}^n$ , i.e.,

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

**A2** There exists  $f_{low} \in \mathbb{R}$  such that  $f(x) \geq f_{low}$  for all  $x \in \mathbb{R}^n$ .

From A1, it can be shown that, for all  $x, y \in \mathbb{R}^n$ ,

$$f(y) \leq f(x) + \langle \nabla f(x), y - x \rangle + \frac{1}{2} \langle \nabla^2 f(x)(y - x), y - x \rangle + \frac{L}{6} \|y - x\|^3 \tag{4}$$

---

<sup>2</sup>The numerical experiments reported in [1] show that in certain variants of CNM with inexact Hessians, the difference between  $\|x_{t+1} - x_t\|$  and  $\|x_t - x_{t-1}\|$  may reach different orders of magnitude. Thus, in practice, inequalities (1) and (2) induce very different error bounds.

and

$$\|\nabla f(y) - \nabla f(x) - \nabla^2 f(x)(y - x)\| \leq \frac{L}{2} \|y - x\|^2. \quad (5)$$

In view of inequality (4), we will consider the following cubic models for  $f(y)$ :

$$\Omega_{x,\sigma}(y) := f(x) + \langle \nabla f(x), y - x \rangle + \frac{1}{2} \langle \nabla^2 f(x)(y - x), y - x \rangle + \frac{\sigma}{6} \|y - x\|^3 \quad (6)$$

and

$$M_{x,\sigma}(y) := f(x) + \langle \nabla f(x), y - x \rangle + \frac{1}{2} \langle B(y - x), y - x \rangle + \frac{\sigma}{6} \|y - x\|^3, \quad (7)$$

where  $\sigma > 0$  and  $B \in \mathbb{R}^{n \times n}$  is an approximation to  $\nabla^2 f(x)$ .

Our first auxiliary result gives upper bounds for  $\|\nabla f(x^+)\|$  and also for  $-\lambda_{\min}(\nabla^2 f(x^+))$  under suitable conditions about  $x^+$  and the matrix  $B$ .

**Lemma 1** *Suppose that A1 holds and assume that  $x^+ \in \mathbb{R}^n$  satisfies*

$$\|\nabla M_{x,\sigma}(x^+)\| \leq \theta \|x^+ - x\|^2 \quad (8)$$

for some  $x \in \mathbb{R}^n$ ,  $\sigma > 0$  and  $\theta \geq 0$ . If for some  $\kappa_B \geq 0$ ,  $\gamma > 0$  and  $\hat{x} \in \mathbb{R}^n$ , we have

$$\|B - \nabla^2 f(x)\| \leq \kappa_B \min \{\|x - \hat{x}\|, \gamma \|\nabla f(x)\|\}, \quad (9)$$

then

$$\|\nabla f(x^+)\| \leq \frac{\sigma + L + 2(\theta + \kappa_B)}{2} \max \{\|x^+ - x\|, \min \{\|x - \hat{x}\|, \gamma \|\nabla f(x)\|\}\}^2. \quad (10)$$

If additionally

$$B + \frac{\sigma}{2} \|x^+ - x\| I \geq -\theta \|x - \hat{x}\| I, \quad (11)$$

then

$$-\lambda_{\min}(\nabla^2 f(x^+)) \leq \frac{\sigma + 2(\theta + \kappa_B + L)}{2} \max \{\|x^+ - x\|, \|x - \hat{x}\|\}. \quad (12)$$

*Proof* Given  $y \in \mathbb{R}^n$ , denote

$$\Phi_x(y) := f(x) + \langle \nabla f(x), y - x \rangle + \frac{1}{2} \langle \nabla^2 f(x)(y - x), y - x \rangle.$$

Then, by (5)–(8), we get

$$\begin{aligned} \|\nabla f(x^+)\| &\leq \|\nabla f(x^+) - \nabla \Phi_x(x^+)\| + \|\nabla \Phi_x(x^+) - \nabla M_{x,\sigma}(x^+)\| + \|\nabla M_{x,\sigma}(x^+)\| \\ &\leq \frac{L}{2} \|x^+ - x\|^2 + \left\| \left( \nabla^2 f(x) - B \right) (x^+ - x) - \frac{\sigma}{2} \|x^+ - x\| (x^+ - x) \right\| \\ &\quad + \|\nabla M_{x,\sigma}(x^+)\| \\ &\leq \left( \frac{L + \sigma}{2} + \theta \right) \|x^+ - x\|^2 + \|\nabla^2 f(x) - B\| \|x^+ - x\|. \end{aligned}$$

Hence, it follows from (9) that

$$\begin{aligned} \|\nabla f(x^+)\| &\leq \left(\frac{L + \sigma}{2} + \theta\right) \|x^+ - x\|^2 + \kappa_B \min \{\|x - \hat{x}\|, \gamma \|\nabla f(x)\|\} \|x^+ - x\| \\ &\leq \left(\frac{L + \sigma}{2} + \theta + \kappa_B\right) \max \{\|x^+ - x\|, \min \{\|x - \hat{x}\|, \gamma \|\nabla f(x)\|\}\}^2, \end{aligned}$$

that is, (10) is true.

On the other hand, by A1 and (9), for any  $d \in \mathbb{R}^n$ , we have

$$\begin{aligned} \langle (B - \nabla^2 f(x^+))d, d \rangle &= \langle (B - \nabla^2 f(x))d, d \rangle + \langle (\nabla^2 f(x) - \nabla^2 f(x^+))d, d \rangle \\ &\leq \|B - \nabla^2 f(x)\| \|d\|^2 + \|\nabla^2 f(x) - \nabla^2 f(x^+)\| \|d\|^2 \\ &\leq \kappa_B \|x - \hat{x}\| \|d\|^2 + L \|x^+ - x\| \|d\|^2 \\ &\leq (\kappa_B + L) \max \{\|x^+ - x\|, \|x - \hat{x}\|\} \|d\|^2 \\ &= \langle (\kappa_B + L) \max \{\|x^+ - x\|, \|x - \hat{x}\|\} Id, d \rangle. \end{aligned}$$

Since the inequality above holds for all  $d \in \mathbb{R}^n$ , it follows that

$$\begin{aligned} B - \nabla^2 f(x^+) &\preceq (\kappa_B + L) \max \{\|x^+ - x\|, \|x - \hat{x}\|\} I \\ \implies B &\preceq \nabla^2 f(x^+) + (\kappa_B + L) \max \{\|x^+ - x\|, \|x - \hat{x}\|\} I \end{aligned}$$

Then, using Weyl’s inequality [12, 28], we get

$$\lambda_{\min}(B) \leq \lambda_{\min}(\nabla^2 f(x^+)) + (\kappa_B + L) \max \{\|x^+ - x\|, \|x - \hat{x}\|\}. \tag{13}$$

Now, assuming that (11) is true, we also have

$$\begin{aligned} \lambda_{\min}(B) + \frac{\sigma}{2} \|x^+ - x\| &\geq -\theta \|x - \hat{x}\| \\ \implies \lambda_{\min}(B) + \frac{\sigma}{2} \|x^+ - x\| + \theta \|x - \hat{x}\| &\geq 0 \\ \implies \lambda_{\min}(B) + \left(\frac{\sigma}{2} + \theta\right) \max \{\|x^+ - x\|, \|x - \hat{x}\|\} &\geq 0, \end{aligned}$$

which gives

$$\lambda_{\min}(B) \geq -\left(\frac{\sigma}{2} + \theta\right) \max \{\|x^+ - x\|, \|x - \hat{x}\|\}. \tag{14}$$

Finally, combining (13) and (14), we obtain (12). □

The next lemma provides a lower bound on  $f(x) - f(x^+)$  when  $M_{x,\sigma}(x^+) \leq M_{x,\sigma}(x) = f(x)$  and  $\sigma$  is sufficiently large.

**Lemma 2** *Suppose that A1 holds and assume that  $x^+$  satisfies*

$$M_{x,\sigma}(x^+) \leq f(x) \tag{15}$$

*for some  $x \in \mathbb{R}^n$  and  $\sigma > 0$ . Moreover, suppose that for some  $\kappa_B \geq 0$  and  $\hat{x} \in \mathbb{R}^n$ , inequality (9) hold. If*

$$\sigma \geq 2(L + 3\kappa_B) \tag{16}$$

then

$$f(x) - f(x^+) \geq \frac{\sigma}{12} \|x^+ - x\|^3 - \frac{\kappa_B}{2} \|x - \hat{x}\|^3. \quad (17)$$

*Proof* By (4), (7), (9) and (15), we have

$$\begin{aligned} f(x^+) &\leq f(x) + \langle \nabla f(x), x^+ - x \rangle + \frac{1}{2} \langle \nabla^2 f(x)(x^+ - x), x^+ - x \rangle + \frac{L}{6} \|x^+ - x\|^3 \\ &= f(x) + \langle \nabla f(x), x^+ - x \rangle + \frac{1}{2} \langle B(x^+ - x), x^+ - x \rangle + \frac{\sigma}{6} \|x^+ - x\|^3 \\ &\quad + \frac{1}{2} \langle (\nabla^2 f(x) - B)(x^+ - x), x^+ - x \rangle + \frac{(L - \sigma)}{6} \|x^+ - x\|^3 \\ &= M_{x,\sigma}(x^+) + \frac{1}{2} \langle (\nabla^2 f(x) - B)(x^+ - x), x^+ - x \rangle + \frac{(L - \sigma)}{6} \|x^+ - x\|^3 \\ &\leq f(x) + \frac{\kappa_B}{2} \|x - \hat{x}\| \|x^+ - x\|^2 + \frac{(L - \sigma)}{6} \|x^+ - x\|^3. \end{aligned} \quad (18)$$

Note that

$$\|x - \hat{x}\| \|x^+ - x\|^2 \leq \|x - \hat{x}\|^3 + \|x^+ - x\|^3. \quad (19)$$

Combining (18) and (19), we get

$$f(x^+) \leq f(x) + \frac{\kappa_B}{2} \|x - \hat{x}\|^3 + \frac{L + 3\kappa_B - \sigma}{6} \|x^+ - x\|^3,$$

and so

$$f(x) - f(x^+) \geq \frac{\sigma - L - 3\kappa_B}{6} \|x^+ - x\|^3 - \frac{\kappa_B}{2} \|x - \hat{x}\|^3.$$

Finally, using (16), it follows that (17) is true.  $\square$

The third auxiliary result gives sufficient conditions under which the error bound (9) is satisfied by a suitable finite difference approximation  $B$ . This is a classical result. For the reader's convenience, we give the proof here.

**Lemma 3** *Suppose that A1 holds. Given  $x \in \mathbb{R}^n$  and  $h > 0$ , let  $A \in \mathbb{R}^{n \times n}$  be defined by*

$$A := \left[ \frac{\nabla f(x + he_1) - \nabla f(x)}{h}, \dots, \frac{\nabla f(x + he_n) - \nabla f(x)}{h} \right]. \quad (20)$$

then the matrix

$$B := \frac{1}{2} (A + A^T) \quad (21)$$

satisfies

$$\|B - \nabla^2 f(x)\| \leq \frac{\sqrt{n}L}{2} h. \quad (22)$$

*Proof* It follows from (5) with  $y = x + he_i$  that

$$\begin{aligned} \|\nabla f(x + he_i) - \nabla f(x) - h\nabla^2 f(x)e_i\| &\leq \frac{L}{2}h^2 \\ \implies \left\| \left( \frac{\nabla f(x + he_i) - \nabla f(x)}{h} \right) - \nabla^2 f(x)e_i \right\| &\leq \frac{L}{2}h \\ \implies \|(A - \nabla^2 f(x))e_i\| &\leq \frac{L}{2}h, \end{aligned}$$

and so

$$\|A - \nabla^2 f(x)\|^2 \leq \|A - \nabla^2 f(x)\|_F^2 = \sum_{i=1}^n \|(A - \nabla^2 f(x))e_i\|_2^2 \leq n \left(\frac{L}{2}\right)^2 h^2,$$

which gives

$$\|A - \nabla^2 f(x)\| \leq \frac{\sqrt{n}L}{2}h. \tag{23}$$

Finally, combining (21) and (23), we get

$$\|B - \nabla^2 f(x)\| \leq \|A - \nabla^2 f(x)\| \leq \frac{\sqrt{n}L}{2}h.$$

□

Combining the last two results, we have the following theorem, which is the basis for the nonmonotone line search criterion used in our method.

**Theorem 1** *Suppose that A1 holds and assume that  $x^+$  satisfies (8) and (15) for some  $x \in \mathbb{R}^n$  and  $\sigma > 0$ . Moreover, suppose that the matrix  $B$  in  $M_{x,\sigma}(\cdot)$  is defined as*

$$B := \frac{1}{2} (A + A^T),$$

where  $A$  is given by (20) with

$$0 < h \leq \frac{2\kappa_B}{\sqrt{n}\sigma} \min \{ \|x - \hat{x}\|, \gamma \|\nabla f(x)\| \} \tag{24}$$

for some  $\kappa_B > 0$ ,  $\gamma > 0$  and  $\hat{x} \in \mathbb{R}^n$ . If

$$\sigma \geq 2(L + \theta + 3\kappa_B), \tag{25}$$

it holds

$$f(x) - f(x^+) \geq \frac{\sigma}{12} \|x^+ - x\|^3 - \frac{\kappa_B}{2} \|x - \hat{x}\|^3 \quad (26)$$

and

$$\|\nabla f(x^+)\| \leq \sigma \max \{ \|x^+ - x\|, \min \{ \|x - \hat{x}\|, \hat{\gamma} \|\nabla f(x)\| \} \}^2, \quad (27)$$

where  $\hat{\gamma} = \max \{1, \gamma\}$ .

*Proof* By (25) and (24), we have

$$0 < h < \frac{2\kappa_B}{\sqrt{n}L} \min \{ \|x - \hat{x}\|, \gamma \|\nabla f(x)\| \}.$$

Then, it follows from Lemma 3 that

$$\|B - \nabla^2 f(x)\| \leq \kappa_B \min \{ \|x - \hat{x}\|, \gamma \|\nabla f(x)\| \}. \quad (28)$$

Finally, in view of (8), (15), (25) and (28), by Lemmas 1 and 2, we conclude that (26) and (27) hold.  $\square$

**Lemma 4** Given  $\tau, \lambda > 0$  and a set  $\{z_j\}_{j=1}^k$  of nonnegative real numbers, with  $k \geq 2$ , let

$$m(k) := \operatorname{argmin}_{j \in \{1, \dots, k-1\}} \left( z_j^\tau + z_{j+1}^\tau \right). \quad (29)$$

If

$$\sum_{j=1}^k z_j^\tau \leq \lambda, \quad (30)$$

then

$$\max \{ z_{m(k)}, z_{m(k)+1} \} \leq \left( \frac{2\lambda}{k-1} \right)^{\frac{1}{\tau}}. \quad (31)$$

*Proof* It follows from (29) and (30) that

$$z_{m(k)}^\tau + z_{m(k)+1}^\tau = \min_{j \in \{1, \dots, k-1\}} \left( z_j^\tau + z_{j+1}^\tau \right) \leq \frac{1}{k-1} \sum_{j=1}^{k-1} \left( z_j^\tau + z_{j+1}^\tau \right)$$

Therefore,  $z_{m(k)} \leq [2\lambda/(k-1)]^{1/\tau}$  and  $z_{m(k)+1} \leq [2\lambda/(k-1)]^{1/\tau}$ , which implies (31).  $\square$

### 3 CNM with finite difference Hessian approximations

We are now in position to present our cubic regularization of Newton’s method with finite difference Hessian approximations to problem (3).

---

**Algorithm 1** CNM with finite-difference hessian approximations.

---

**Step 0.** Given  $x_0, x_1 \in \mathbb{R}^n$  ( $x_0 \neq x_1$ ),  $\sigma_1, \gamma > 0$ ,  $\theta \geq 0$ , set  $\kappa_B = \sigma_1/6$ ,  $\hat{\gamma} = \max\{1, \gamma\}$ , and  $t := 1$ .

**Step 1.** Find the smallest integer  $i \geq 0$  such that  $2^i \sigma_t \geq 2\sigma_1$ .

**Step 1.1.** For

$$h_i = \frac{2\kappa_B \min\{\|x_t - x_{t-1}\|, \gamma \|\nabla f(x_t)\|\}}{\sqrt{n} (2^i \sigma_t)}, \tag{32}$$

compute

$$A_{t,i} = \left[ \frac{\nabla f(x_t + h_i e_1) - \nabla f(x_t)}{h_i}, \dots, \frac{\nabla f(x_t + h_i e_n) - \nabla f(x_t)}{h_i} \right] \in \mathbb{R}^{n \times n}, \tag{33}$$

and define

$$B_{t,i} := \frac{1}{2} (A_{t,i} + A_{t,i}^T). \tag{34}$$

**Step 1.2.** Consider the model

$$M_{x_t, 2^i \sigma_t}(y) := f(x_t) + \langle \nabla f(x_t), y - x_t \rangle + \frac{1}{2} \langle B_{t,i}(y - x_t), y - x_t \rangle + \frac{2^i \sigma_t}{6} \|y - x_t\|^3,$$

and compute an approximate solution  $x_{t,i}^+$  to the subproblem

$$\min_{y \in \mathbb{R}^n} M_{x_t, 2^i \sigma_t}(y), \tag{35}$$

such that

$$M_{x_t, 2^i \sigma_t}(x_{t,i}^+) \leq f(x_t) \quad \text{and} \quad \|\nabla M_{x_t, 2^i \sigma_t}(x_{t,i}^+)\| \leq \theta \min\{\|x_{t,i}^+ - x_t\|^2, \|\nabla f(x_t)\|\}. \tag{36}$$

**Step 1.3.** If

$$f(x_t) - f(x_{t,i}^+) \geq \frac{2^i \sigma_t}{12} \|x_{t,i}^+ - x_t\|^3 - \frac{\sigma_1}{12} \|x_t - x_{t-1}\|^3 \tag{37}$$

and

$$\|\nabla f(x_{t,i}^+)\| \leq \left(2^i \sigma_t\right) \max\left\{\|x_{t,i}^+ - x_t\|, \min\{\|x_t - x_{t-1}\|, \hat{\gamma} \|\nabla f(x_t)\|\}\right\}^2 \tag{38}$$

hold, set  $i_t = i$  and go to Step 2. Otherwise, set  $i := i + 1$  and go to Step 1.1.

**Step 2.** Set  $x_{t+1} = x_{t,i_t}^+$ ,  $\sigma_{t+1} = 2^{i_t-1} \sigma_t$ ,  $t := t + 1$ , and go to Step 1.

---

*Remark 1* Conditions in (36), which are similar to those proposed in [5], only require a decrease of the cubic regularized model and an approximate first-order stationary point.

*Remark 2* Notice that condition (37) allows the acceptance of a trial step  $x_{t,i}^+$  for which

$$f(x_t) < f(x_{t,i}^+) \leq f(x_t) + \frac{\sigma_1}{12} \|x_t - x_{t-1}\|^3.$$

Consequently, the corresponding sequence  $\{f(x_t)\}_{t \geq 0}$  may be nonmonotone. Thus, Algorithm 1 is a nonmonotone CNM. Different nonmonotone variants of CNM have been proposed in [3, 23, 30]. Specifically, the method presented in [3] is inspired in the nonmonotone line search of Grippo, Lampariello and Lucidi [17], while the methods presented in [23, 30] are inspired in the nonmonotone line search of Zhang and Hager [29].

**Lemma 5** *Suppose that A1 holds. Then, the sequence of regularization parameters  $\{\sigma_t\}$  in Algorithm 1 satisfies*

$$\sigma_1 \leq \sigma_t \leq 2 \left( L + \theta + \frac{3\sigma_1}{2} \right) := \sigma_{max}, \quad (39)$$

for all  $t \geq 1$ . Moreover, the number  $O_T$  of calls of the oracle up to the  $T$ -th iteration is bounded as follows:

$$O_T \leq (n + 2) \left[ 2T + \log_2 \left( 2 \left( L + \theta + \frac{3\sigma_1}{2} \right) \right) - \log_2(\sigma_1) \right]. \quad (40)$$

*Proof* Clearly, (39) is true for  $t = 1$ . Suppose that (39) holds for some  $t \geq 1$ . If  $i_t = 0$ , then by Step 1 and the induction assumption, we have

$$\sigma_1 \leq \sigma_{t+1} = \frac{1}{2} \sigma_t \leq \sigma_t \leq 2 \left( L + \theta + \frac{3\sigma_1}{2} \right),$$

that is, (39) holds for  $t + 1$ . On the other hand, if  $i_t \geq 1$ , then we must have

$$2^{i_t-1} \sigma_t \leq 2 \left( L + \theta + \frac{3\sigma_1}{2} \right). \quad (41)$$

Indeed, assuming by contradiction that (41) is not true and using  $\kappa_B = \sigma_1/6$ , it follows that

$$2^{i_t-1} \sigma_t > 2 \left( L + \theta + \frac{3\sigma_1}{2} \right) = 2(\sigma_1 + L + \theta + 3\kappa_B).$$

In this case, by Theorem 1, inequalities (37) and (38) would have been satisfied for  $i = i_t - 1$ , contradicting the minimality of  $i_t$ . Thus, (41) is true. Consequently, by Step 1 and (41), we have

$$\sigma_1 \leq \sigma_{t+1} = \frac{1}{2} (2^{i_t} \sigma_t) \leq 2 \left( L + \theta + \frac{3\sigma_1}{2} \right),$$

that is, (39) also holds for  $t + 1$  in this case. This completes the induction argument.

Finally, note that at the  $t$ -th iteration of Algorithm 1 the number of calls of the oracle is bounded by  $(n + 2)(i_t + 1)$ . On the other hand,

$$\sigma_{t+1} = 2^{i_t-1}\sigma_t \implies (n + 2)(i_t + 1) = (n + 2) [2 + \log_2(\sigma_{t+1}) - \log_2(\sigma_t)].$$

Thus,

$$\begin{aligned} O_T &\leq \sum_{t=1}^T (n + 2)(i_t + 1) = (n + 2) [2T + \log_2(\sigma_{T+1}) - \log_2(\sigma_1)] \\ &\leq (n + 2) \left[ 2T + \log_2 \left( 2 \left( L + \theta + \frac{3\sigma_1}{2} \right) \right) - \log_2(\sigma_1) \right], \end{aligned}$$

where the last inequality is due to (39). □

*Remark 3* It follows from (40) that

$$\frac{1}{T} O_T \leq 2(n + 2) + \frac{(n + 2)}{T} \left[ \log_2 \left( 2 \left( L + \theta + \frac{3\sigma_1}{2} \right) \right) - \log_2(\sigma_1) \right].$$

Thus, in Algorithm 1, the average number of oracle calls per iteration, up to the  $T$ -th iteration, is asymptotically bounded by  $2(n + 2)$ .

The theorem below establishes an iteration-complexity bound of  $\mathcal{O}(\epsilon^{-3/2})$  for Algorithm 1.

**Theorem 2** *Suppose that A1 and A2 hold. Given  $\epsilon > 0$ , let  $\{x_t\}_{t=1}^T$  be generated by Algorithm 1, such that*

$$\|\nabla f(x_t)\| > \epsilon, \quad t = 1, \dots, T. \tag{42}$$

Then,

$$T < 3 + \left[ \frac{24(f(x_1) - f_{low})}{\sigma_1} + 2\|x_1 - x_0\|^3 \right] \left[ 4 \left( L + \theta + \frac{3\sigma_1}{2} \right) \right]^{\frac{3}{2}} \epsilon^{-\frac{3}{2}}. \tag{43}$$

*Proof* Notice that  $2^{i_t}\sigma_t = 2\sigma_{t+1}$  and, by Step 1,  $\sigma_t \geq \sigma_1$  for all  $t \geq 1$ . Then, it follows from (37) that

$$f(x_t) - f(x_{t+1}) \geq \frac{\sigma_{t+1}}{6} \|x_{t+1} - x_t\|^3 - \frac{\sigma_t}{12} \|x_t - x_{t-1}\|^3, \quad t = 1, \dots, T - 1.$$

As in [5], summing up these inequalities and using the lower bound on  $f(\cdot)$  and  $\sigma_t \geq \sigma_1$ , we get

$$\begin{aligned}
 f(x_1) - f_{low} &\geq f(x_1) - f(x_T) \\
 &= \sum_{t=1}^{T-1} f(x_t) - f(x_{t+1}) \\
 &\geq \sum_{t=1}^{T-1} \frac{\sigma_{t+1}}{6} \|x_{t+1} - x_t\|^3 - \sum_{t=1}^{T-1} \frac{\sigma_t}{12} \|x_t - x_{t-1}\|^3 \\
 &= \sum_{t=2}^T \frac{\sigma_t}{6} \|x_t - x_{t-1}\|^3 - \sum_{t=2}^{T-1} \frac{\sigma_t}{12} \|x_t - x_{t-1}\|^3 - \frac{\sigma_1}{12} \|x_1 - x_0\|^3 \\
 &\geq \sum_{t=2}^T \frac{\sigma_t}{6} \|x_t - x_{t-1}\|^3 - \sum_{t=2}^T \frac{\sigma_t}{12} \|x_t - x_{t-1}\|^3 - \frac{\sigma_1}{12} \|x_1 - x_0\|^3 \\
 &= \sum_{t=2}^T \frac{\sigma_t}{12} \|x_t - x_{t-1}\|^3 - \frac{\sigma_1}{12} \|x_1 - x_0\|^3 \\
 &\geq \frac{\sigma_1}{12} \sum_{t=1}^{T-1} \|x_{t+1} - x_t\|^3 - \frac{\sigma_1}{12} \|x_1 - x_0\|^3,
 \end{aligned}$$

and so

$$\sum_{t=1}^{T-1} \|x_{t+1} - x_t\|^3 \leq \frac{12(f(x_1) - f_{low})}{\sigma_1} + \|x_1 - x_0\|^3. \quad (44)$$

Let us denote  $s_t := x_{t+1} - x_t$ . In this way, we can rewrite (44) as

$$\sum_{t=1}^{T-1} \|s_t\|^3 \leq \frac{12(f(x_1) - f_{low})}{\sigma_1} + \|s_0\|^3. \quad (45)$$

Since (43) is clearly true for  $T \in \{1, 2\}$ , let us assume that  $T \geq 3$ . Define

$$t_* := \operatorname{argmin}_{j \in \{1, \dots, T-2\}} (\|s_j\|^3 + \|s_{j+1}\|^3).$$

Then, by Lemma 4 with  $z_j = \|s_j\|$ ,  $k = T - 1$  and  $\tau = 3$ , it follows from (45) that

$$\max \{ \|s_{t_*}\|, \|s_{t_*+1}\| \} \leq \left[ \frac{24(f(x_1) - f_{low})}{\sigma_1} + 2\|s_0\|^3 \right]^{\frac{1}{3}} \frac{1}{(T-2)^{\frac{1}{3}}}. \quad (46)$$

On the other hand, by (38) and (42), we have

$$\begin{aligned}
 2(\sigma_{t_*+1}) \max \{ \|s_{t_*+1}\|, \|s_{t_*}\| \}^2 &\geq \|\nabla f(x_{t_*+1})\| \geq \epsilon \\
 \implies \max \{ \|s_{t_*}\|, \|s_{t_*+1}\| \} &> \left( \frac{\epsilon}{2\sigma_{t_*+1}} \right)^{\frac{1}{2}}. \quad (47)
 \end{aligned}$$

Then, combining (46) and (47), it follows that

$$\begin{aligned} \left(\frac{\epsilon}{2\sigma_{t_*+1}}\right)^{\frac{1}{2}} &< \left[\frac{24(f(x_1) - f_{low})}{\sigma_1} + 2\|s_0\|^3\right]^{\frac{1}{3}} \frac{1}{(T-2)^{\frac{1}{3}}} \\ \implies \left(\frac{\epsilon}{2\sigma_{t_*+1}}\right)^{\frac{3}{2}} &< \left[\frac{24(f(x_1) - f_{low})}{\sigma_1} + 2\|s_0\|^3\right] \frac{1}{T-2} \\ \implies T-2 &< \left[\frac{24(f(x_1) - f_{low})}{\sigma_1} + 2\|s_0\|^3\right] \left(\frac{2\sigma_{t_*+1}}{\epsilon}\right)^{\frac{3}{2}}. \end{aligned}$$

Finally, using the upper bound on  $\sigma_{t_*+1}$  given by Lemma 5, we obtain (43). □

Combining (40) and (43), we obtain the following evaluation-complexity bound.

**Corollary 1** *Suppose that A1 and A2 hold. Then, given  $\epsilon > 0$ , Algorithm 1 needs at most  $\mathcal{O}\left(n\epsilon^{-\frac{3}{2}}\right)$  calls of the oracle to generate an iterate  $x_t$  such that  $\|\nabla f(x_t)\| \leq \epsilon$ .*

As a consequence of Corollary 1, we also have a liminf-type global convergence result for Algorithm 1.

**Corollary 2** *Suppose that A1 and A2 hold and let  $\{x_t\}_{t \geq 1}$  be a sequence generated by Algorithm 1. Then, either exists  $\hat{t}$  such that  $\nabla f(x_{\hat{t}}) = 0$  or*

$$\liminf_{t \rightarrow +\infty} \|\nabla f(x_t)\| = 0. \tag{48}$$

*Proof* Suppose that  $\nabla f(x_t) \neq 0$  for all  $t \geq 1$ . In this case, by Corollary 1, the sequence  $\{x_t\}_{t \geq 1}$  has a subsequence  $\{x_{t_j}\}_{j \geq 1}$  such that

$$\|\nabla f(x_{t_j})\| \leq \epsilon_j, \tag{49}$$

where

$$\epsilon_j = \begin{cases} 1, & \text{for } j = 1, \\ \min\left\{\frac{1}{j}, \min_{i=1, \dots, j-1} \|\nabla f(x_{t_i})\|\right\}, & \text{for } j \geq 2. \end{cases}$$

Since  $\lim_{j \rightarrow +\infty} \epsilon_j = 0$ , by (49), we have

$$\lim_{j \rightarrow +\infty} \|\nabla f(x_{t_j})\| = 0,$$

and so (48) is true. □

Let us now consider a variant of Algorithm 1 in which, besides (36), we require

$$B_{t,i} + \frac{2^i \sigma_t}{2} \|x_{t,i}^+ - x_t\| I \geq -\theta \|x_t - x_{t-1}\| I. \tag{50}$$

This means that we will assume that  $x_{t,i}^+$  approximately satisfies the first and the second-order optimality conditions for a local minimizer of the cubic model  $M_{x_t, 2^i \sigma_t}(\cdot)$ . In this scenario, we can prove second-order complexity and global convergence results for Algorithm 1.

**Theorem 3** Suppose that A1 and A2 hold and let  $\{x_t\}_{t=1}^T$  be generated by Algorithm 1 with (36) and (50) being satisfied for all  $t$  and  $i \leq i_t$ . Given  $\epsilon_g, \epsilon_H > 0$ , if

$$\|\nabla f(x_t)\| > \epsilon_g \quad \text{or} \quad \lambda_{\min}(\nabla^2 f(x_t)) < -\epsilon_H, \quad \text{for } t = 1, \dots, T, \quad (51)$$

then

$$T < 3 + \left[ \frac{24(f(x_1) - f_{low})}{\sigma_1} + 2\|x_1 - x_0\|^3 \right] \left( \max \{ \sqrt{2\sigma_{max}}, \sigma_{max} \} + \theta + \frac{\sigma_1}{6} + L \right)^3 \max \{ \epsilon_g^{-3/2}, \epsilon_H^{-3} \}. \quad (52)$$

where  $\sigma_{max}$  is defined in (39).

*Proof* As in the proof of Theorem 2, we have that

$$\max \{ \|s_{t^*}\|, \|s_{t^*+1}\| \} \leq \left[ \frac{24(f(x_1) - f_{low})}{\sigma_1} + 2\|s_0\|^3 \right]^{\frac{1}{3}} \frac{1}{(T-2)^{\frac{1}{3}}}, \quad (53)$$

where  $s_t = x_{t+1} - x_t$  and

$$t^* = \operatorname{argmin}_{j \in \{1, \dots, T-2\}} (\|s_j\|^3 + \|s_{j+1}\|^3).$$

If  $\|\nabla f(x_{t^*+1})\| > \epsilon_g$ , then, by (38), we have

$$\begin{aligned} 2(\sigma_{t^*+1}) \max \{ \|s_{t^*+1}\|, \|s_{t^*}\| \}^2 &> \epsilon_g \\ \implies \max \{ \|s_{t^*+1}\|, \|s_{t^*}\| \} &> \left( \frac{\epsilon_g}{2\sigma_{t^*+1}} \right)^{\frac{1}{2}} \end{aligned}$$

On the other hand, if  $\|\nabla f(x_{t^*+1})\| \leq \epsilon_g$  then, by (51), we must have

$$-\lambda_{\min}(\nabla^2 f(x_{t^*+1})) > \epsilon_H.$$

In this case, it follows from Lemma 1 that

$$\max \{ \|s_{t^*+1}\|, \|s_{t^*}\| \} > \frac{\epsilon_H}{\sigma_{t^*+1} + \theta + \kappa_B + L}.$$

Thus, in any case, we have

$$\max \{ \|s_{t^*+1}\|, \|s_{t^*}\| \} > \frac{1}{\max \{ \sqrt{2\sigma_{t^*+1}}, \sigma_{t^*+1} \} + \theta + \kappa_B + L} \min \{ \epsilon_g^{1/2}, \epsilon_H \}. \quad (54)$$

Combining (53) and (54), we get

$$T - 2 < \left[ \frac{24(f(x_1) - f_{low})}{\sigma_1} + 2\|s_0\|^3 \right] \left( \max \{ \sqrt{2\sigma_{t^*+1}}, \sigma_{t^*+1} \} + \theta + \kappa_B + L \right)^3 \max \{ \epsilon_g^{-3/2}, \epsilon_H^{-3} \}.$$

Finally, using the upper bound on  $\sigma_{t^*+1}$  given by Lemma 5, we obtain (52). □

**Corollary 3** Suppose that A1 and A2 hold and let  $\{x_t\}_{t \geq 1}$  be generated by Algorithm 1 with (36) and (50) being satisfied for all  $t$  and  $i \leq i_t$ . Then, given  $\epsilon_g, \epsilon_H \in (0, 1)$ ,

Algorithm 1 needs at most  $O\left(n \max\left\{\epsilon_g^{-3/2}, \epsilon_H^{-3}\right\}\right)$  calls of the oracle to generate  $x_t$  such that

$$\|\nabla f(x_t)\| \leq \epsilon_g \quad \text{and} \quad \lambda_{\min}\left(\nabla^2 f(x_t)\right) \geq -\epsilon_H.$$

Consequently, either there exists  $\hat{t}$  such that  $\nabla f(x_{\hat{t}}) = 0$  and  $\nabla^2 f(x_{\hat{t}}) \succeq 0$  or

$$\liminf_{t \rightarrow +\infty} \max\left\{\|\nabla f(x_t)\|, -\lambda_{\min}\left(\nabla^2 f(x_t)\right)\right\} = 0. \tag{55}$$

*Proof* The evaluation-complexity bound follows directly from Theorem 3 and Lemma 5, while (55) follows by the same argument used to prove Corollary 2.  $\square$

*Remark 4* From (55), we see that the addition of requirement (50) in Step 1.2 of Algorithm 1 allows the iterates to escape from nondegenerate saddle points.

Finally, we can establish a local convergence rate under the following additional assumption:

**A3** There exists  $\mu > 0$  such that  $\nabla^2 f(x) \succeq \mu I$  whenever

$$f(x) \leq 2f(x_1) - f_{low} + \frac{\sigma_1}{6} \|x_1 - x_0\|^3.$$

Note first that, by (37), we have

$$f(x_{t+1}) - f(x_t) \leq \frac{\sigma_1}{12} \|x_t - x_{t-1}\|^3, \quad \forall k \geq 1.$$

Consequently, for all  $t \geq 2$ , we have

$$\begin{aligned} f(x_t) - f(x_1) &= \sum_{j=1}^{t-1} f(x_{j+1}) - f(x_j) \\ &\leq \frac{\sigma_1}{12} \sum_{j=1}^{t-1} \|x_j - x_{j-1}\|^3 \\ &= \frac{\sigma_1}{12} \|x_1 - x_0\|^3 + \frac{\sigma_1}{12} \sum_{j=1}^{t-2} \|x_{j+1} - x_j\|^3 \\ &\leq \frac{\sigma_1}{6} \|x_1 - x_0\|^3 + f(x_1) - f_{low}, \end{aligned}$$

where the last inequality is due to (45). Thus, it follows from A3 that

$$\nabla^2 f(x_t) \succeq \mu I, \quad \forall t \geq 1. \tag{56}$$

Using this remark, we can prove the following local quadratic convergence rate for Algorithm 1.

**Theorem 4** Suppose that A1–A3 hold and let  $\{x_t\}_{t \geq 1}$  be generated by Algorithm 1. If

$$\|\nabla f(x_1)\| \leq \min \left\{ \frac{6\mu}{L\gamma}, \frac{\mu^2}{2[2\hat{\gamma}^2\mu^2 + 8(\theta + 1)^2]\sigma_{max}} \right\}, \tag{57}$$

then

$$\|\nabla f(x_t)\| \leq \frac{\mu^2}{[2\hat{\gamma}^2\mu^2 + 8(\theta + 1)^2]\sigma_{max}} \left(\frac{1}{2}\right)^t, \quad \forall t \geq 2. \tag{58}$$

*Proof* First, we will show that

$$\|\nabla f(x_{t+1})\| \leq \frac{[2\hat{\gamma}^2\mu^2 + 8(\theta + 1)^2]\sigma_{max}}{\mu^2} \|\nabla f(x_t)\|^2 \tag{59}$$

for all  $t \geq 1$ . Assume that

$$\|\nabla f(x_t)\| \leq \min \left\{ \frac{6\mu}{L\gamma}, \frac{\mu^2}{2[2\hat{\gamma}^2\mu^2 + 8(\theta + 1)^2]\sigma_{max}} \right\} \tag{60}$$

for some  $t \geq 1$ . Then, by (32)–(34), Lemma 3, the facts that  $\kappa_B = \sigma_1/6$  and  $2^i\sigma_t \geq 2\sigma_1$ , and (60), we have

$$\begin{aligned} \|\nabla^2 f(x_t) - B_t\| &\leq \frac{\sqrt{n}L}{2} \frac{2\kappa_B\gamma}{\sqrt{n}2\sigma_{t+1}} \|\nabla f(x_t)\| \leq \frac{L\kappa_B\gamma}{2\sigma_1} \|\nabla f(x_t)\| = \frac{L\gamma}{12} \|\nabla f(x_t)\| \\ &\leq \frac{\mu}{2}. \end{aligned}$$

Thus, given  $v \neq 0$ , we have

$$\begin{aligned} v^T \left( \nabla^2 f(x_t) - B_t \right) v &\leq \|\nabla^2 f(x_t) - B_t\| \|v\|^2 \leq \frac{\mu}{2} \|v\|^2 = v^T \left( \frac{\mu}{2} I \right) v \\ &\implies v^T \nabla^2 f(x_t) v \leq v^T \left( B_t + \frac{\mu}{2} I \right) v. \end{aligned}$$

Thus,

$$\nabla^2 f(x_t) \leq B_t + \frac{\mu}{2} I.$$

and, by Weyl’s inequality, we get

$$\begin{aligned} \lambda_{min} \left( \nabla^2 f(x_t) \right) &\leq \lambda_{min} (B_t) + \frac{\mu}{2} \\ \implies \lambda_{min} (B_t) &\geq \lambda_{min} \left( \nabla^2 f(x_t) \right) - \frac{\mu}{2}. \end{aligned} \tag{61}$$

Combining (61) and (56), we get

$$\lambda_{min} (B_t) \geq \lambda_{min} \left( \nabla^2 f(x_t) \right) - \frac{1}{2} \lambda_{min} \left( \nabla^2 f(x_t) \right) = \frac{1}{2} \lambda_{min} \left( \nabla^2 f(x_t) \right). \tag{62}$$

On the other hand, it follows from the second inequality in (36) that

$$\|\nabla M_{x_t, 2^i\sigma_t}(x_{t+1})\| \leq \theta \|\nabla f(x_t)\|, \tag{63}$$

where

$$\nabla M_{x_t, 2^i \sigma_t}(x_{t+1}) = \nabla f(x_t) + B_t(x_{t+1} - x_t) + \frac{2^i \sigma_t}{2} \|x_{t+1} - x_t\| (x_{t+1} - x_t).$$

From the last equality, we get

$$\begin{aligned} \left( B_t + \frac{2^i \sigma_t}{2} \|x_{t+1} - x_t\| I \right) (x_{t+1} - x_t) &= \nabla M_{x_t, 2^i \sigma_t}(x_{t+1}) - \nabla f(x_t) \\ \implies x_{t+1} - x_t &= - \left( B_t + \frac{2^i \sigma_t}{2} \|x_{t+1} - x_t\| I \right)^{-1} (\nabla M_{x_t, 2^i \sigma_t}(x_{t+1}) - \nabla f(x_t)). \end{aligned} \tag{64}$$

Then, by (64), (62) and (63), we have

$$\begin{aligned} \|x_{t+1} - x_t\| &= \left\| \left( B_t + \frac{2^i \sigma_t}{2} \|x_{t+1} - x_t\| I \right)^{-1} (\nabla M_{x_t, 2^i \sigma_t}(x_{t+1}) - \nabla f(x_t)) \right\| \\ &\leq \frac{\|\nabla M_{x_t, 2^i \sigma_t}(x_{t+1}) - \nabla f(x_t)\|}{\lambda_{\min}(B_t)} \\ &\leq \frac{2\|\nabla M_{x_t, 2^i \sigma_t}(x_{t+1}) - \nabla f(x_t)\|}{\lambda_{\min}(\nabla^2 f(x_t))} \\ &\leq \frac{2(\|\nabla M_{x_t, 2^i \sigma_t}(x_{t+1})\| + \|\nabla f(x_t)\|)}{\lambda_{\min}(\nabla^2 f(x_t))} \\ &\leq \frac{2(\theta + 1)\|\nabla f(x_t)\|}{\lambda_{\min}(\nabla^2 f(x_t))}. \end{aligned} \tag{65}$$

Now, combining (38), (39) and (65), and recalling that  $2^i \sigma_t = 2\sigma_{t+1}$ , it follows that

$$\begin{aligned} \|\nabla f(x_{t+1})\| &\leq 2\sigma_{\max} \max \{ \|x_{t+1} - x_t\|, \min \{ \|x_t - x_{t-1}\|, \hat{\gamma} \|\nabla f(x_t)\| \} \}^2 \\ &\leq 2\sigma_{\max} \{ \|x_{t+1} - x_t\|, \hat{\gamma} \|\nabla f(x_t)\| \}^2 \\ &\leq 2\sigma_{\max} \max \left\{ \frac{2(\theta + 1)\|\nabla f(x_t)\|}{\lambda_{\min}(\nabla^2 f(x_t))}, \hat{\gamma} \|\nabla f(x_t)\| \right\}^2 \\ &\leq 2\sigma_{\max} \max \left\{ \frac{2(\theta + 1)}{\mu}, \hat{\gamma} \right\}^2 \|\nabla f(x_t)\|^2 \\ &\leq \left( 2\hat{\gamma}^2 + \frac{8(\theta + 1)^2}{\mu^2} \right) \sigma_{\max} \|\nabla f(x_t)\|^2 \\ &\leq \frac{[2\hat{\gamma}^2 \mu^2 + 8(\theta + 1)^2] \sigma_{\max}}{\mu^2} \|\nabla f(x_t)\|^2. \end{aligned}$$

Moreover, by (60), we also have

$$\begin{aligned} \|\nabla f(x_{t+1})\| &\leq \frac{[2\hat{\gamma}^2 \mu^2 + 8(\theta + 1)^2] \sigma_{\max}}{\mu^2} \frac{\mu^2}{2[2\hat{\gamma}^2 \mu^2 + 8(\theta + 1)^2] \sigma_{\max}} \|\nabla f(x_t)\| = \frac{1}{2} \|\nabla f(x_t)\| \\ &< \min \left\{ \frac{6\mu}{L\gamma}, \frac{\mu^2}{2[2\hat{\gamma}^2 \mu^2 + 8(\theta + 1)^2] \sigma_{\max}} \right\} \end{aligned} \tag{66}$$

Thus, by induction, (59) holds for all  $t \geq 1$ .

Denoting

$$\delta_t := \frac{[2\hat{\gamma}^2\mu^2 + 8(\theta + 1)^2]\sigma_{\max}}{\mu^2} \|\nabla f(x_t)\|,$$

it follows from (59) that

$$\delta_{t+1} \leq \delta_t^2 \quad \forall t \geq 1.$$

Moreover, by (57), we also have

$$\delta_1 = \frac{[2\hat{\gamma}^2\mu^2 + 8(\theta + 1)^2]\sigma_{\max}}{\mu^2} \|\nabla f(x_1)\| \leq \frac{1}{2}.$$

Therefore, for all  $t \geq 2$ ,

$$\begin{aligned} \|\nabla f(x_t)\| &= \frac{\mu^2}{[2\hat{\gamma}^2\mu^2 + 8(\theta + 1)^2]\sigma_{\max}} \delta_t \leq \frac{\mu^2}{[2\hat{\gamma}^2\mu^2 + 8(\theta + 1)^2]\sigma_{\max}} \delta_1^{2^t} \\ &\leq \frac{\mu^2}{[2\hat{\gamma}^2\mu^2 + 8(\theta + 1)^2]\sigma_{\max}} \left(\frac{1}{2}\right)^{2^t}, \end{aligned}$$

and the proof is complete.  $\square$

## 4 Illustrative numerical experiments

In this section, we present preliminary numerical results obtained by an Octave implementation of Algorithm 1. Our code corresponds to Algorithm 1 with the additional condition (50). Regarding the parameters, we used  $\sigma_1 = 1$ ,  $\theta = 10$ ,  $\gamma = 6\|\nabla f(x_1)\|^{-1}$  and  $\|x_1 - x_0\| = 6$ , resulting in  $h_1 = 1/\sqrt{n}$ . Each cubic subproblem (35) is approximately solved by a monotone BFGS line search method using as initial point the approximate solution generated by 10 iterations of the method described in Section 6.1 of [8]<sup>3</sup>. In our first experiment, we applied the referred code to a set of 10 nonconvex test problems formed with functions from [21] (each problem with two choices for the dimension  $n$ ), using the following stopping criterion:

$$\|\nabla f(x_t)\| \leq \epsilon. \quad (67)$$

The results are shown in Table 1, where  $T(\epsilon)$  represents the number of iterations required by the code to satisfy the stopping criterion (67),  $O(\epsilon)$  represents the corresponding number of calls of the oracle (function evaluations plus gradients evaluations), and  $D(\epsilon)$  is defined as

$$D(\epsilon) := \frac{O(\epsilon)}{T(\epsilon)(n + 2)}. \quad (68)$$

From Table 1, we can see that all problems were solved in the sense of condition (67). Moreover, except for Problem 10,  $D(\epsilon)$  is approximately bounded by 2, which is in accordance with Remark 3, made about Lemma 5.

<sup>3</sup>The choice of performing 10 iterations was done based on a few preliminary numerical tests. Running the method in [8] with this number of iterations often provided a very good initial point for the BFGS method.

**Table 1** Numerical results of an implementation of Algorithm 1

PROBLEM ( <i>n</i> )	$\epsilon = 10^{-2}$			$\epsilon = 10^{-5}$		
	$T(\epsilon)$	$O(\epsilon)$	$D(\epsilon)$	$T(\epsilon)$	$O(\epsilon)$	$D(\epsilon)$
1. Extended Rosenbrock (8)	42	882	2.1000	45	942	2.0933
2. Extended Rosenbrock (16)	44	1640	2.0707	47	1748	2.0662
3. Extended Powell Singular (8)	14	252	1.8000	49	952	1.9429
4. Extended Powell Singular (16)	23	884	2.1353	67	2468	2.0464
5. Penalty I (8)	13	252	1.9385	172	3462	2.0128
6. Penalty I (16)	16	578	2.0069	196	7112	2.0159
7. Penalty II (8)	8	192	2.4000	71	1462	2.0592
8. Penalty II (16)	17	722	2.3595	212	7724	2.0241
9. Variably Dimensioned (8)	14	372	2.6571	16	392	2.4500
10. Variably Dimensioned (16)	18	902	2.7840	23	1496	3.6135
11. Trigonometric (8)	5	82	1.6400	8	122	1.5250
12. Trigonometric (16)	6	200	1.8519	8	236	1.6389
13. Discrete Boundary Value (8)	1	12	1.2000	8	82	1.0250
14. Discrete Boundary Value (16)	1	20	1.1111	23	416	1.0048
15. Discrete Integral Equation (8)	2	22	1.1000	3	32	1.0667
16. Discrete Integral Equation (16)	2	38	1.0556	3	56	1.0370
17. Broyden Tridiagonal (8)	4	42	1.0500	5	52	1.0400
18. Broyden Tridiagonal (16)	4	74	1.0278	4	74	1.0278
19. Broyden Banded (8)	6	132	2.2000	7	142	2.0286
20. Broyden Banded (16)	7	272	2.1587	8	290	2.0139

According with Theorem 2, the number of iterations  $T(\epsilon)$  satisfies  $T(\epsilon) \leq C_{f,A}\epsilon^{-3/2}$ , where the constant  $C_{f,A} > 0$  depends on the problem and the parameters used in Algorithm 1. As pointed in [15], by assuming that

$$T(\epsilon) = C_{f,A}\epsilon^{-p}, \quad \epsilon > 0,$$

we can estimate  $p$  numerically using the formula

$$p = \frac{1}{\log(\tau)} \log \left( \frac{T(\epsilon/\tau)}{T(\epsilon)} \right), \tag{69}$$

where  $\tau > 1$ . We estimated the complexity power  $p$  for the problems in Table 1 considering  $T(\epsilon)$  and  $T(\epsilon/\tau)$  with  $\epsilon = 10^{-2}$  and  $\tau = 10^3$ . The results are given in Table 2.

As we can see, the estimated power  $p$  in all problems is much smaller than  $3/2$ , which agrees with Theorem 2. In particular, the largest power obtained was  $p = 0.4539$ . This result illustrate the very pessimistic aspect of the worst-case complexity bounds in the nonconvex setting.

**Table 2** Numerical estimation of the complexity order  $p$  in  $T(\epsilon) = C_{f,A}\epsilon^{-p}$ . The largest power obtained was  $p = 0.4539$  for Problem 14

PROBLEM ( $n$ )	$T(\epsilon)$	$T(\epsilon/\tau)$	$p$
1. Extended Rosenbrock (8)	42	45	0.0100
2. Extended Rosenbrock (16)	44	47	0.0095
3. Extended Powell Singular (8)	14	49	0.1814
4. Extended Powell Singular (16)	23	67	0.1548
5. Penalty I (8)	13	172	0.3739
6. Penalty I (16)	16	196	0.3627
7. Penalty II (8)	8	71	0.3161
8. Penalty II (16)	17	212	0.3653
9. Variably Dimensioned (8)	14	16	0.0193
10. Variably Dimensioned (16)	18	23	0.0355
11. Trigonometric (8)	5	8	0.0680
12. Trigonometric (16)	6	8	0.0416
13. Discrete Boundary Value (8)	1	8	0.3010
14. Discrete Boundary Value (16)	1	23	0.4539
15. Discrete Integral Equation (8)	2	3	0.0587
16. Discrete Integral Equation (16)	2	3	0.0587
17. Broyden Tridiagonal (8)	4	5	0.0323
18. Broyden Tridiagonal (16)	4	4	0.0000
19. Broyden Banded (8)	6	7	0.0223
20. Broyden Banded (16)	7	8	0.0193

In our second experiment, we applied our implementation of Algorithm 1 for the two-dimensional problem [19]:

$$\min_{x \in \mathbb{R}^2} f(x) \equiv (1/4)x_1^4 + (1/4)x_2^4 - (5/3)x_1^3 - (5/3)x_2^3. \quad (70)$$

It is worth mentioning that the objective  $f(\cdot)$  in (70) has a global minimum at  $x^* = [5 \ 5]^T$ , a degenerate saddle point at  $[0 \ 0]^T$  and nondegenerate saddle points at  $[5 \ 0]^T$  and  $[0 \ 5]^T$ . We considered seven starting points close to the saddle points of  $f(\cdot)$ . The results are shown in Table 3, where  $x_F$  denotes the final point returned by our code. As we can see in Table 3, for all starting points, Algorithm 1 returned a point  $x_F$  very close to the global minimizer  $x^*$ , escaping from the saddle points. With respect to the nondegenerate saddle points, this result is in accordance with Remark 4, made about Corollary 3.

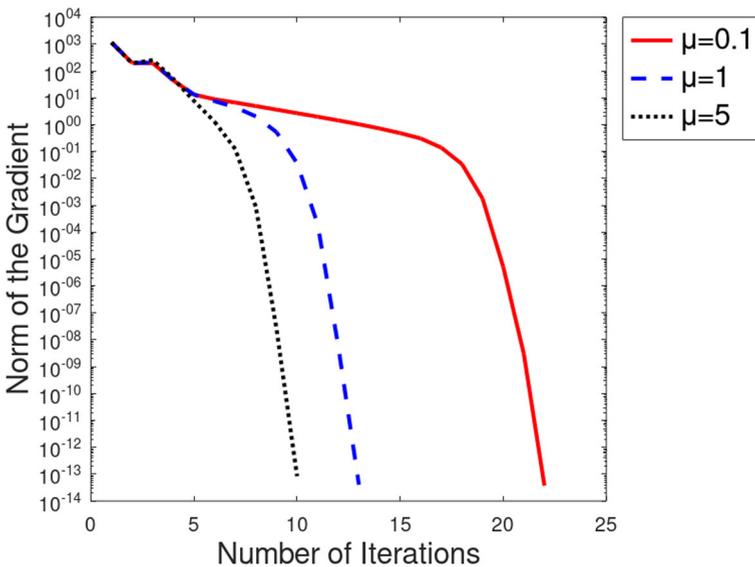
In our final experiment, we applied our code to  $l_2$ -regularized logistic problems of the form

$$\min_{x \in \mathbb{R}^n} f_\mu(x) := - \sum_{i=1}^m \left[ b^{(i)} \log(c_x(a^{(i)})) + (1 - b^{(i)}) \log(1 - c_x(a^{(i)})) \right] + \frac{\mu}{2} \|x\|_2^2, \quad (71)$$

**Table 3** Numerical results for Algorithm 1 applied to problem (70) with  $\epsilon = 10^{-5}$

$x_1$	$\ x_F - x^*\ $	$T(\epsilon)$	$O(\epsilon)$	$D(\epsilon)$
$\begin{bmatrix} 4.9 & -0.1 \end{bmatrix}^T$	4.9934E-11	6	26	1.0833
$\begin{bmatrix} 5.1 & -0.01 \end{bmatrix}^T$	2.3653E-08	6	30	1.2500
$\begin{bmatrix} 4.99 & 0.01 \end{bmatrix}^T$	1.3076E-09	6	30	1.2500
$\begin{bmatrix} -0.002 & 5.1 \end{bmatrix}^T$	4.7686E-09	6	30	1.2500
$\begin{bmatrix} 0.001 & 5 \end{bmatrix}^T$	7.3187E-09	5	26	1.3000
$\begin{bmatrix} 0.001 & 0.1 \end{bmatrix}^T$	9.5072E-10	11	70	1.5909
$\begin{bmatrix} 0.001 & -0.001 \end{bmatrix}^T$	3.3821E-09	11	70	1.5909

where  $\{(a^{(i)}, b^{(i)})\}_{i=1}^m \subset \mathbb{R}^n \times \{0, 1\}$  is the dataset,  $c_x(a) := 1/(1 + e^{-\langle a, x \rangle})$  is the logistic model, and  $\mu > 0$  is the regularization parameter. The objective function  $f_\mu(\cdot)$  in (71) is  $\mu$ -strongly convex and has Lipschitz continuous Hessian. Therefore,  $f_\mu(\cdot)$  satisfies assumptions A1–A3. We considered the Breast Cancer Wisconsin



**Fig. 1** Behavior of  $\|\nabla f_\mu(x_t)\|$  as a function of the iteration counter  $t$

dataset [25]<sup>4</sup> with  $n = 10$ ,  $m = 683$  and  $a_1^{(i)} = 1$  for  $i = 1, \dots, m$ . As starting point we used  $x_0 = [0 \dots 0]^T \in \mathbb{R}^{10}$ . Figure 1 shows the behavior of  $\{\|\nabla f_\mu(x_t)\|\}$  for  $\mu \in \{0.1, 1, 5\}$ . The curves confirm the quadratic rate of convergence established in Theorem 4.

## 5 Conclusion

In this paper, we presented a new variant of the cubic regularization of Newton's method (CNM) with Hessian matrices approximated by forward finite difference Hessians. The method is designed for the unconstrained minimization of twice differentiable functions with globally Lipschitz continuous Hessians. The stepsizes that define the finite difference approximations are adjusted jointly with the regularization parameters of the cubic models. Specifically, at the  $t$ -th iteration, our method approximates  $\nabla^2 f(x_t)$  by a matrix  $B_t$  such that

$$\|\nabla^2 f(x_t) - B_t\| \leq \mathcal{O}(\min\{\|x_t - x_{t-1}\|, \|\nabla f(x_t)\|\}).$$

A similar approximation has been considered in [27] in the context of a CNM variant with inexact Hessians. However, the method analyzed in [27] requires the knowledge of the Lipschitz constant (used in the regularization parameter) and also the exact solution of the subproblems. In contrast, our method uses a nonmonotone line search procedure to update the regularization parameters and allows the inexact solution of the subproblems. We proved that the proposed method needs at most  $\mathcal{O}(n\epsilon_g^{-3/2})$  calls of the oracle to generate an  $\epsilon_g$ -approximate first-order stationary point of the objective function. Moreover, we showed that the method needs at most  $\mathcal{O}(n \max\{\epsilon_g^{-3/2}, \epsilon_H^{-3}\})$  calls of the oracle to generate an  $(\epsilon_g, \epsilon_H)$ -approximate second-order stationary point. We also proved a quadratic convergence result for the proposed method. Finally, we presented illustrative numerical results, which confirmed our theoretical findings.

**Acknowledgments** We are very grateful to the two anonymous referees, whose comments helped to improve the paper."

**Funding** G.N. Grapiglia was partially supported by CNPq - Brazil grant 312777/2020-5. M.L.N. Gonçalves was partially supported by CNPq - Brazil grant 408123/2018-4.

## References

1. Bellavia, S., Gurioli, G., Morini, B.: Adaptive cubic regularization methods with dynamic inexact Hessian information and applications to finite-sum minimization. *IMA J. Numer. Anal.* **41**, 764–799 (2021)

<sup>4</sup>This dataset is freely available in the UCI Machine Learning Repository (<https://archive.ics.uci.edu/ml>).

2. Bergou, E., Diouane, Y., Gratton, S.: A line-search algorithm inspired by the adaptive cubic regularization framework with a worst-case complexity  $o(\epsilon^{-3/2})$ . *Optimization Online* (2017)
3. Bianconcini, T., Sciandrone, M.: A cubic regularization algorithm for unconstrained optimization using line search and nonmonotone techniques. *Optim. Methods Softw.* **31**, 1008–1035 (2016)
4. Birgin, E.G., Martínez, J.M.: The use of quadratic regularization with cubic descent condition for unconstrained optimization. *SIAM J. Optim.* **27**, 1049–1074 (2017)
5. Birgin, E.G., Gardenghi, J.L., Martínez, J.M., Santos, S.A., Toint, P.L.: Worst-case evaluation complexity for unconstrained nonlinear optimization using high-order regularized models. *Math. Program.* **163**, 359–368 (2017)
6. Carmon, Y., Duchi, J.C., Hinder, O., Sidford, A.: Lower bounds for finding stationary points I. *Math. Program.* **184**, 71–120 (2020)
7. Cartis, C., Gould, N.I.M., Toint, P.L.: On the complexity of steepest descent, Newton’s and regularized Newton’s methods for nonconvex unconstrained optimization problems. *SIAM J. Optim.* **20**, 2833–2852 (2010)
8. Cartis, C., Gould, N.I.M., Toint, P.L.: Adaptive cubic regularization methods for unconstrained optimization. Part I: motivation, convergence and numerical results. *Math. Program.* **127**, 245–295 (2011)
9. Cartis, C., Gould, N.I.M., Toint, Ph.L.: Adaptive cubic regularization methods for unconstrained optimization. Part II: worst-case function - and derivative - evaluation complexity. *Math. Program.* **130**, 295–319 (2011)
10. Cartis, C., Gould, N.I.M., Toint, P.L.: On the oracle complexity of first-order and derivative-free algorithms for smooth nonconvex minimization. *SIAM J. Optim.* **22**, 66–86 (2012)
11. Curtis, F.E., Robinson, D.P., Samadi, M.: A trust-region algorithm with a worst-case complexity of  $\mathcal{O}(\epsilon^{-3/2})$  for nonconvex optimization. *Math. Program.* **162**, 1–32 (2017)
12. Demmel, J.: *Applied numerical linear algebra*. SIAM, Philadelphia (1997)
13. Dussault, J.-P.: ARCq: a new adaptive regularization by cubics variant. *Optim. Methods Softw.* <https://doi.org/10.1080/10556788.2017.1322080> (2017)
14. Grapiglia, G.N., Yu, N.: Regularized Newton Methods for minimizing functions with hölder continuous Hessians. *SIAM J. Optim.* **27**, 478–506 (2017)
15. Grapiglia, G.N., Sachs, E.W.: On the worst-case evaluation complexity of non-monotone line search algorithms. *Comput. Optim. Appl.* **68**, 555–577 (2017)
16. Griewank, A.: *The Modification of Newton’s Method for Unconstrained Optimization by Bounding Cubic Terms*. Technical Report NA/12, Department of Applied Mathematics and Theoretical Physics. University of Cambridge, Cambridge (1981)
17. Grippo, L., Lampariello, F., Lucidi, S.: A nonmonotone line search technique for Newton’s method. *SIAM J. Numer. Anal.* **23**, 707–716 (1986)
18. Kohler, J.M., Lucchi, A.: Subsampled cubic regularization for non-convex optimization. *34th Int Conf Machine Learn ICML 2017(4)*, 2988–3011 (2017)
19. Martínez, J.M., Raydan, M.: Cubic-regularization counterpart of a variable-norm trust-region method for unconstrained minimization. *J. Glob. Optim.* **68**, 367–385 (2017)
20. Melicher, V., Haber, T., Vanroose, W.: Fast derivatives of likelihood functionals for ODE based models using adjoint-state method. *Comput. Stat.* **32**, 1621–1643 (2017)
21. Moré, J.J., Garbow, B.S., Hillstom, K.E.: Testing unconstrained optimization software. *ACM Trans. Math. Softw.* **7**, 17–41 (1981)
22. Nesterov, Y., Polyak, B.: Cubic regularization of Newton method and its global performance. *Math. Program.* **108**, 177–205 (2006)
23. Niri, T.D., Heydari, M., Hosseini, M.M.: An improvement of adaptive cubic regularization method for unconstrained optimization problems. *Int. J. Comput. Math.* **98**, 271–287 (2021)
24. Stapor, P., Fröhlich, F., Hasenauer, J.: Optimization and profile calculation of ODE models using second order adjoint sensitivity analysis. *Bioinformatics* **34**, i151–i159 (2018)
25. Street, W.N., Wolberg, W.H., Mangasarian, O.L.: Nuclear Feature Extraction for Breast Tumor Diagnosis. *IS&T/SPIE 1993 International Symposium on Electronic Imaging: Science and Technology*, vol. 1905, pp.861–870. San Jose, CA (1993)
26. Wang, Z., Zhou, Y., Liang, Y., Lan, G.: Stochastic Variance-Reduced Cubic Regularization for Nonconvex Optimization. *arXiv:1802.07372* (2018)
27. Wang, Z., Zhou, Y., Liang, Y., Lan, G.: A note on inexact gradient and Hessian conditions for cubic regularized Newton’s method. *Oper. Res. Lett.* **47**, 146–149 (2019)

28. Weyl, H.: Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen. *Math. Ann.* **71**, 441–479 (1912)
29. Zhang, H.C., Hager, W.W.: A nonmonotone line search technique for unconstrained optimization. *SIAM J Optim* **14**, 1043–1056 (2004)
30. Zheng, Y., Zheng, B.: A modified adaptive cubic regularization method for large-scale unconstrained optimization problem. arXiv:[1904.07440](https://arxiv.org/abs/1904.07440) [math.OC] (2019)

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.