Global complexity bound analysis of the Levenberg-Marquardt method for nonsmooth equations and its application to the nonlinear complementarity problem

Kenji Ueda* and Nobuo Yamashita†

Abstract

We investigate a global complexity bound of the Levenberg-Marquardt Method (LMM) for nonsmooth equations $F(x) = 0$. The global complexity bound is an upper bound to the number of iterations required to get an approximate solution such that $\|\nabla f(x)\| \le \epsilon$, where f is a least square merit function and ϵ is a given positive constant. We show that the bound of the LMM is $O(\epsilon^{-2})$. We also show that it is reduced to $O(\log \epsilon^{-1})$ under some regularity assumption on the generalized Jacobian of F. Furthermore, by applying these results to nonsmooth equations equivalent to the nonlinear complementarity problem (NCP), we get global complexity bounds for the NCP. In particular, we show that the bound is $O(\log \epsilon^{-1})$ when the mapping involved in the NCP is a uniformly P-function.

Keywords Levenberg-Marquardt methods, Global complexity bound, Nonlinear complementarity problems

Mathematics Subject Classification (2000) 90C56, 90C33, 49M15

1 Introduction

We consider a system of nonsmooth equations

$$F(x) = 0, \tag{1}$$

where $F : \mathbb{R}^n \to \mathbb{R}^m$ is a locally Lipschitz continuous mapping. When the system (1) has a solution, it is equivalent to the following nonlinear least squares problem.

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \;\; f(x) := \frac{1}{2}\|F(x)\|^2. \tag{2}$$

In this paper, we assume that the least squares merit function f is continuously differentiable, though F is nonsmooth. The system (1) satisfying these assumptions includes important applications such as the nonlinear complementarity problem (NCP) and the Karush-Kuhn-Tacker (KKT) system [1]. For the system (1) or the NCP, the Levenberg-Marquardt method (LMM) is known to be an efficient solution method [2, 3, 4, 5, 6, 7, 8].

A global complexity bound is one of the important factors for choosing an appropriate solution method [9, 10, 11, 12, 13, 14, 15, 16, 17]. When we solve an unconstrained minimization problem of a nonconvex function ϕ by some iterative methods, the global complexity bound is defined as an upper bound of the number of iterations required to get an approximate stationary point x such that $\|\nabla \phi(x)\| \le \epsilon$, where ϵ is a given positive constant. Since it corresponds to the worst computational time, it is useful when we want to estimate in advance the time for solving a large-scale problem. Recently,

*Corresponding author. Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan, E-mail: kueda@amp.i.kyoto-u.ac.jp

†Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan, E-mail: nobuo@i.kyoto-u.ac.jp

the bounds of some general iterative methods for the unconstrained minimization problem, such as the steepest descent method and the Newton-type methods, have been actively investigated. Thus, if we apply these results to the least squares problem (2), we can estimate the bound for (2). However, since these methods are not specialized to the problem (2), they are not efficient. In fact, the steepest descent method converges slow in general. Moreover, the Newton-type methods [9, 10, 11, 13, 14, 15, 16] require the twice continuous differentiability of F. Recently, Ueda and Yamashita [17] investigated the bound of the LMM, which is a special method for (2). Under the assumption that F is continuously differentiable, they showed that it is $O(\epsilon^{-2})$ without any regularity assumption on F. However, we cannot directly apply this result to a system of nonsmooth equations.

In this paper, we consider an LMM for the nonsmooth equations that uses the generalized Jacobian of F. We show that it has the same bound $O(\epsilon^{-2})$ as [17]. Moreover, under some regularity assumption of the generalized Jacobian of F, we also show that an upper bound of the number of iterations required to get an approximate solution x such that $\|F(x)\| \leq \epsilon$ is $O(\log \epsilon^{-1})$. By applying these results to the NCP, we get the global complexity bounds for the NCP. In particular, we can get the bound $O(\log \epsilon^{-1})$ when the mapping involved in the NCP is a uniformly P-function.

This paper is organized as follows. In the next section, we give some definitions related to the generalized Jacobian. In Section 3, we introduce the LMM for nonsmooth equations. In Section 4, we give the global complexity bounds of the LMM. In Section 5, we apply the results on the bounds to the NCP. Finally, Section 6 concludes the paper.

Throughout the paper, we use the following notations. For a vector $x \in \mathbb{R}^n$, $\|x\|$ denotes the Euclidean norm defined by $\|x\| := \sqrt{x^T x}$. For a symmetric matrix $M \in \mathbb{R}^{n\times n}$, we denote the maximum eigenvalue and the minimum eigenvalue of M as $\lambda_{\max}(M)$ and $\lambda_{\min}(M)$, respectively. For a matrix $M \in \mathbb{R}^{n\times m}$, $\|M\|$ denotes the ℓ_2 norm of M defined by $\|M\| := \sqrt{\lambda_{\max}(M^T M)}$. If M is symmetric positive semidefinite matrix, then $\|M\| = \lambda_{\max}(M)$. $B(x, r)$ denotes the closed sphere with center x and radius r, i.e., $B(x, r) := \{y \in \mathbb{R}^n \mid \|y - x\| \leq r\}$. For sets $S_1 \subseteq \mathbb{R}^n$ and $S_2 \subseteq \mathbb{R}^n$, $S_1 + S_2$ denotes the sum of S_1 and S_2 defined by $S_1 + S_2 := \{x + y \in \mathbb{R}^n \mid x \in S_1, y \in S_2\}$. For a set S, $\mathcal{P}(S)$ denotes the set consisting of all the subsets of S.

2 Preliminaries

In this section, we give some definitions that will be used in the subsequent sections.

When a vector mapping F is nonsmooth, we cannot necessarily use the Jacobian of F. Nevertheless, we can define the generalized Jacobian of F if F is locally Lipschitz continuous [18, 19].

Definition 2.1. *Let $D_F \subseteq \mathbb{R}^n$ be the set where F is differentiable.*

(a) *The B-subdifferential of F at x is defined by*

$$\partial_B F(x) = \{J \in \mathbb{R}^{n\times m} \mid J = \lim_{k\to\infty} \nabla F(x^k),\ \lim_{k\to\infty} x^k = x,\ \{x^k\} \subseteq D_F\}.$$

(b) *The Clarke generalized Jacobian of F at x is defined by*

$$\partial F(x) = \mathrm{co}\, \partial_B F(x),$$

where co *denotes the convex hull of the set.*

Remark 2.1. *Note that since F is assumed to be locally Lipschitz continuous in this paper, we can use the above subdifferentials. Note also that $\partial_B F(x)$ and $\partial F(x)$ are nonempty and compact set for each x [18]. Moreover, if a least squares merit function $f(x) = \frac{1}{2}\|F(x)\|$ is continuously differentiable, we have $\nabla f(x) = J^T F(x)$, $\forall J \in \partial F(x)$ by using the standard calculus rules [18].*

Remark 2.2. *The Fischer-Burmeister function defined by $\psi(a, b) = \sqrt{a^2 + b^2} - a - b$ is not differentiable at $(0, 0)$, but it is locally Lipschitz continuous. Thus, the generalized Jacobian $\partial\psi$ is well-defined at $(0, 0)$.*

Remark 2.3. *To solve the nonsmooth equations* (1), *the Newton-type methods with the generalized Jacobian are often used [3, 20]. For example, the generalized Newton method updates the k-th iterative point as $x^{k+1} = x^k + d^k$, where d^k is a search direction such that $J_k d^k = -F(x^k)$, $J_k \in \partial F(x^k)$.*

From Definition 2.1, the generalized Jacobian ∂F is a point-to-set mapping from $\mathbb{R}^n$ into $\mathcal{P}(\mathbb{R}^{n\times m})$. Next, we introduce the upper semi-continuity of a point-to-set mapping [21].

Definition 2.2. *Let X be a subset of $\mathbb{R}^n$, Y be a subset of $\mathbb{R}^{n\times m}$, and Θ be a point-to-set mapping from X into $\mathcal{P}(Y)$.*

(a) *Θ is uniformly compact near $\bar{x} \in X$ if there exists a neighborhood N of $\bar{x}$ such that the closure of $\cup_{x\in N}\Theta(x)$ is compact.*

(b) *Θ is closed at $\bar{x}$ if $x^k \to \bar{x}, y^k \in \Theta(x^k)$ and $y^k \to \bar{y}$ imply $\bar{y} \in \Theta(\bar{x})$.*

(c) *Θ is upper semi-continuous at $\bar{x}$ if Θ is uniformly compact near $\bar{x}$ and closed at $\bar{x}$.*

It is well-known that ∂F is upper semi-continuous [18]. Thus, for each x, $\max_{J\in\partial F(x)}\|J\|$ is bounded above.

3 The Levenberg-Marquardt method

In this section, we explain the LMM for the system of nonsmooth equations (1). In what follows, let x^k be the k-th iterative point, F_k be $F(x^k)$, and $J_k \in \partial F(x^k)$. Throughout the paper, we need the following assumptions.

Assumption 3.1.

(a) *The vector mapping F is locally Lipschitz continuous.*

(b) *The least squares merit function f is continuously differentiable.*

As mentioned in Remark 2.1, we can use the generalized Jacobian under Assumption 3.1 (a). Moreover, the system (1) satisfying Assumption 3.1 includes important applications such as the nonlinear complementarity problem (NCP) and the Karush-Kuhn-Tacker (KKT) system.

For the current iterative point x^k, an LMM adopts a search direction $d^k(\mu_k)$ defined by

$$d^k(\mu_k) = -(J_k^T J_k + \mu_k I)^{-1} J_k^T F_k,$$

where μ_k is a positive parameter. In order to guarantee global convergence property, μ_k is updated based on the idea of the trust-region method [22, 23]. Note that a search direction $d^k(\mu_k)$ is given as a solution of a trust-region subproblem of (2), that is,

$$\underset{d\in\mathbb{R}^n}{\text{minimize}} \ \frac{1}{2}\|F(x^k) + J_k d\|^2 \quad \text{subject to } \|d\|^2 \leq \Delta_k^2,$$

and μ_k corresponds to the Lagrange multiplier of the Karush-Kuhn-Tucker conditions of the subproblem. Since the trust-region method controls the trust-region radius Δ_k for global convergence, it requires to solve the subproblem at each iteration [22]. On the other hand, Osborne [23] proposed to update μ_k directly instead of Δ_k. Then, $d^k(\mu_k)$ is given as a solution of the linear equations which is much easier to solve than the trust-region subproblem. Therefore, we adopt his updating rule with the following little modification. We set μ_k as

$$\mu_k = \nu_k \|F_k\|^\delta,$$

and we control a positive parameter ν_k instead of μ_k. Here, δ is a given constant such that $\delta \geq 0$. In what follows, we denote the search direction as $d^k(\nu_k)$ instead of $d^k(\mu_k)$.

We control ν_k as follows. Let $f_k : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ be a model function of f at x^k defined by

$$f_k(d,\nu) := \frac{1}{2}\|F_k + J_k d\|^2 + \frac{1}{2}\nu\|F_k\|^\delta\|d\|^2.$$

Let $\rho_k : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ be the ratio of the reduction of the merit function value to that of the model function value, i.e.,

$$\rho_k(d, \nu) := \frac{f(x^k) - f(x^k + d)}{f(x^k) - f_k(d, \nu)}.$$

If $\rho_k(d^k(\nu_k), \nu_k)$ is large, then the LMM adopts $d^k(\nu_k)$ and decreases the parameter ν_k. On the other hand, if $\rho_k(d^k(\nu_k), \nu_k)$ is small, then the LMM increases ν_k and computes $d^k(\nu_k)$ once again.

We describe the precise description of the LMM as follows.

The Levenberg-Marquardt Method

Step 0 : Choose parameters $\epsilon, \nu_0, \delta, \gamma_1, \gamma_2, \eta_1, \eta_2$ such that

$$0 < \epsilon < 1,\ \nu_0 > 0,\ \delta \geq 0,\ \gamma_1 < 1 < \gamma_2,\ 0 < \eta_1 \leq \eta_2 \leq 1.$$

Choose a starting point x^0. Set $k := 0$.

Step 1 : Choose $J_k \in \partial F(x^k)$. If $\|J_k^T F_k\| \leq \epsilon$, then terminate. Otherwise, go to Step 2.

Step 2 : Step 2.0 : Set $l_k := 1$ and $\bar{\nu}_{l_k} = \nu_k$.

Step 2.1 : Compute

$$d^k(\bar{\nu}_{l_k}) = -(J_k^T J_k + \bar{\nu}_{l_k} \|F_k\|^\delta I)^{-1} J_k^T F_k.$$

Step 2.2 : Compute

$$\rho_k(d^k(\bar{\nu}_{l_k}), \bar{\nu}_{l_k}) = \frac{f(x^k) - f(x^k + d^k(\bar{\nu}_{l_k}))}{f(x^k) - f_k(d^k(\bar{\nu}_{l_k}), \bar{\nu}_{l_k})}.$$

If $\rho_k(d^k(\bar{\nu}_{l_k}), \bar{\nu}_{l_k}) < \eta_1$, then update $\bar{\nu}_{l_k+1} := \gamma_2 \bar{\nu}_{l_k}$, set $l_k := l_k + 1$, and go to Step 2.1. Otherwise, go to Step 3.

Step 3 : If $\eta_2 > \rho_k(d^k(\bar{\nu}_{l_k}), \bar{\nu}_{l_k}) \geq \eta_1$, then update $\nu_{k+1} := \bar{\nu}_{l_k}$.
If $\rho_k(d^k(\bar{\nu}_{l_k}), \bar{\nu}_{l_k}) \geq \eta_2$, then update $\nu_{k+1} := \gamma_1 \bar{\nu}_{l_k}$.
Update $x^{k+1} = x^k + d^k(\bar{\nu}_{l_k})$. Set $k := k + 1$, and go to Step 1.

In what follows, for simplicity, we denote l_k and $\bar{\mu}_{l_k}$ at the last iteration of the inner loops of Steps 2.0–2.2 for each k as l_k^* and μ_k^*, respectively.

In the remainder of this section, we show that the LMM is well-defined when $\|J_k^T F_k\| \neq 0$. First, we give a lower bound of the reduction of the model function.

Lemma 3.1. *Suppose that Assumption 3.1 holds. Then,*

$$f(x^k) - f_k(d^k(\nu), \nu) = -\frac{1}{2} F_k^T J_k d^k(\nu) \geq \frac{\|J_k^T F_k\|^2}{2(\|J_k\|^2 + \nu \|F_k\|^\delta)}.$$

Proof. By the definitions of $f(x^k)$, $f_k(d^k(\nu),\nu)$ and $d^k(\nu)$, we have

$$\begin{aligned}
f(x^k) - f_k(d^k(\nu),\nu) &= \frac{1}{2}\|F_k\|^2 - \left(\frac{1}{2}\|F_k + J_k d^k(\nu)\|^2 + \frac{1}{2}\nu\|F_k\|^\delta \|d^k(\nu)\|^2\right) \\
&= -F_k^T J_k d^k(\nu) - \frac{1}{2} d^k(\nu)^T (J_k^T J_k + \nu\|F_k\|^\delta I) d^k(\nu) \\
&= -\frac{1}{2} F_k^T J_k d^k(\nu) \\
&= \frac{1}{2} F_k^T J_k (J_k^T J_k + \nu\|F_k\|^\delta I)^{-1} J_k^T F_k \\
&\geq \frac{\lambda_{\min}\left((J_k^T J_k + \nu\|F_k\|^\delta I)^{-1}\right)}{2} \|J_k^T F_k\|^2 \\
&= \frac{\|J_k^T F_k\|^2}{2\lambda_{\max}(J_k^T J_k + \nu\|F_k\|^\delta I)} \\
&= \frac{\|J_k^T F_k\|^2}{2(\|J_k\|^2 + \nu\|F_k\|^\delta)}.
\end{aligned}$$

This completes the proof. □

Next, we give an upper bound of $\|d^k(\nu)\|$.

Lemma 3.2. *Suppose that Assumption 3.1 holds. Then,*

$$\|d^k(\nu)\| \leq \frac{\|J_k^T F_k\|}{\nu\|F_k\|^\delta}.$$

Proof. By the definition of $d^k(\nu)$, we have

$$\begin{aligned}
\|d^k(\nu)\| &= \|(J_k^T J_k + \nu\|F_k\|^\delta I)^{-1} J_k^T F_k\| \\
&\leq \|(J_k^T J_k + \nu\|F_k\|^\delta I)^{-1}\| \cdot \|J_k^T F_k\| \\
&= \lambda_{\max}\left((J_k^T J_k + \nu\|F_k\|^\delta I)^{-1}\right) \|J_k^T F_k\| \\
&= \frac{1}{\lambda_{\min}(J_k^T J_k + \nu\|F_k\|^\delta I)} \|J_k^T F_k\| \\
&\leq \frac{\|J_k^T F_k\|}{\nu\|F_k\|^\delta},
\end{aligned}$$

where the last inequality follows from the positive semidefiniteness of $J_k^T J_k$. □

From Lemmas 3.1 and 3.2, we give an upper bound of $f(x^k + d^k(\nu))$.

Lemma 3.3. *Suppose that Assumption 3.1 holds. Then,*

$$f(x^k + d^k(\nu)) \leq f_k(d^k(\nu),\nu) - \frac{\|J_k^T F_k\|^2}{2(\|J_k\|^2 + \nu\|F_k\|^\delta)} + \frac{\|J_k^T F_k\|}{\nu\|F_k\|^\delta} \int_0^1 \|(\nabla f(x^k + \tau d^k(\nu)) - \nabla f(x^k))\| \mathrm{d}\tau.$$

Proof. Since f is continuously differentiable, we have

$$\begin{aligned}
&f(x^k + d^k(\nu)) \\
&= f(x^k) + \int_0^1 \nabla f(x^k + \tau d^k(\nu))^T d^k(\nu) \mathrm{d}\tau \\
&= f(x^k) + \int_0^1 \nabla f(x^k + \tau d^k(\nu))^T d^k(\nu) \mathrm{d}\tau + f_k(d^k(\nu),\nu) - f_k(d^k(\nu),\nu) + F_k^T J_k d^k(\nu) - F_k^T J_k d^k(\nu) \\
&= f_k(d^k(\nu),\nu) + (f(x^k) - f_k(d^k(\nu),\nu) + F_k^T J_k d^k(\nu)) + \int_0^1 (\nabla f(x^k + \tau d^k(\nu)) - J_k^T F_k)^T d^k(\nu) \mathrm{d}\tau.
\end{aligned}$$

It then follows from Lemma 3.1 and $\nabla f(x^k) = J_k^T F_k$ that

$$
\begin{aligned}
& f(x^k + d^k(\nu)) \\
&= f_k(d^k(\nu), \nu) - (f(x^k) - f_k(d^k(\nu), \nu)) + \int_0^1 (\nabla f(x^k + \tau d^k(\nu)) - \nabla f(x^k))^T d^k(\nu) \mathrm{d}\tau \\
&\le f_k(d^k(\nu), \nu) - (f(x^k) - f_k(d^k(\nu), \nu)) + \|d^k(\nu)\| \int_0^1 \|(\nabla f(x^k + \tau d^k(\nu)) - \nabla f(x^k))\| \mathrm{d}\tau \\
&\le f_k(d^k(\nu), \nu) - \frac{\|J_k^T F_k\|^2}{2(\|J_k\|^2 + \nu \|F_k\|^\delta)} + \frac{\|J_k^T F_k\|}{\nu \|F_k\|^\delta} \int_0^1 \|(\nabla f(x^k + \tau d^k(\nu)) - \nabla f(x^k))\| \mathrm{d}\tau,
\end{aligned}
$$

where the last inequality follows from Lemmas 3.1 and 3.2. □

Next, we give the following key lemma for the well-definedness.

Lemma 3.4. *Suppose that Assumption 3.1 holds. Suppose also that* $\|J_k^T F_k\| \neq 0$*. Then,*

$$\rho_k(d^k(\nu), \nu) \ge 1$$

for ν *sufficiently large.*

Proof. Since $\|J_k^T F_k\| \neq 0$, we have $\|F_k\| \neq 0$. Thus, if ν is sufficiently large, $\nu \|F_k\|^\delta \ge \|J_k\|^2$ holds. In what follows, we suppose that $\nu \|F_k\|^\delta \ge \|J_k\|^2$ holds without loss of generality. It then follows from Lemma 3.3 that

$$
\begin{aligned}
f(x^k + d^k(\nu)) &\le f_k(d^k(\nu), \nu) - \frac{\|J_k^T F_k\|^2}{4\nu \|F_k\|^\delta} + \frac{\|J_k^T F_k\|}{\nu \|F_k\|^\delta} \int_0^1 \|(\nabla f(x^k + \tau d^k(\nu)) - \nabla f(x^k))\| \mathrm{d}\tau \\
&\le f_k(d^k(\nu), \nu) + \frac{\|J_k^T F_k\|}{4\nu \|F_k\|^\delta} \left(-\|J_k^T F_k\| + 4 \int_0^1 \|(\nabla f(x^k + \tau d^k(\nu)) - \nabla f(x^k))\| \mathrm{d}\tau \right). \quad (3)
\end{aligned}
$$

Taking $\nu \to \infty$, we have $\lim_{\nu \to \infty} \|d^k(\nu)\| = 0$ from the definition of $d^k(\nu)$, and hence

$$\lim_{\nu \to \infty} \int_0^1 \|(\nabla f(x^k + \tau d^k(\nu)) - \nabla f(x^k))\| \mathrm{d}\tau = 0.$$

Thus, since $\|J_k^T F_k\| \neq 0$, the following inequality holds for sufficiently large ν.

$$4 \int_0^1 \|(\nabla f(x^k + \tau d^k(\nu)) - \nabla f(x^k))\| \mathrm{d}\tau \le \|J_k^T F_k\|.$$

It then follows from (3) that

$$f(x^k + d^k(\nu)) \le f_k(d^k(\nu), \nu).$$

Therefore, by the definition of $\rho_k(d^k(\nu), \nu)$, we have

$$\rho_k(d^k(\nu), \nu) = \frac{f(x^k) - f(x^k + d^k(\nu))}{f(x^k) - f_k(d^k(\nu), \nu)} \ge 1,$$

which is the desired inequality. □

Now, we show the well-definedness of the LMM.

Theorem 3.1. *Suppose that Assumption 3.1 holds. Suppose also that* $\|J_k^T F_k\| \neq 0$*. Then, the LMM is well-defined, i.e., the number* l_k *of inner iteration is finite.*

Proof. From the updating rule of $\bar{\nu}_{l_k}$, we have $\bar{\nu}_{l_k} \to \infty$ as $l_k \to \infty$. Thus, when l_k is sufficiently large, we have from Lemma 3.4 that

$$\rho_k(d^k(\bar{\nu}_{l_k}), \bar{\nu}_{l_k}) = \frac{f(x^k) - f(x^k + d^k(\bar{\nu}_{l_k}))}{f(x^k) - f_k(d^k(\bar{\nu}_{l_k}), \bar{\nu}_{l_k})} \ge 1 \ge \eta_1.$$

Therefore, the LMM is well-defined. □

4 Global complexity bound

In this section, we estimate the global complexity bound of the LMM. Let K_{outer} be the total number of outer iterations when the algorithm terminates. If there does not exist such K_{outer}, we define $K_{\mathrm{outer}} := \infty$. Moreover, let K_{total} be the total number of inner iterations, i.e.,

$$K_{\mathrm{total}} := \sum_{k=0}^{K_{\mathrm{outer}}-1} l_k^*.$$

Note that K_{total} means the total number of solving linear equations.

In order to investigate K_{total}, we firstly make the following assumption.

Assumption 4.1.

(a) $\delta \leq 1$.

(b) *The level set of f at the initial point x^0 is compact, i.e., $\Omega := \{x \in \mathbb{R}^n \mid f(x) \leq f(x^0)\}$ is compact.*

Since $\{f(x^k)\}$ is monotonically decreasing, the sequence $\{x^k\}$ is included in the compact set Ω. Moreover, since the generalized Jacobian ∂F is upper semi-continuous as mentioned in section 2, there exist positive constants U_F and U_J such that

$$\|F(x)\| \leq U_F,\ \max(\|J\|, \|J^T\|) \leq U_J,\ \forall J \in \partial F(x),\ \forall x \in \Omega. \tag{4}$$

Now, we show that $\|d^k(\nu)\|$ is bounded from above when $\nu \in [\nu_0, \infty)$.

Lemma 4.1. *Suppose that Assumptions 3.1 and 4.1 hold. Then, for any $\nu \in [\nu_0, \infty)$,*

$$\|d^k(\nu)\| \leq U_d,$$

where $U_d = \frac{U_J U_F^{1-\delta}}{\nu_0}$.

Proof. It follows from Lemma 3.2 that

$$\|d^k(\nu)\| \leq \frac{\|J_k^T F_k\|}{\nu \|F_k\|^\delta} \leq \frac{\|J_k^T\| \cdot \|F_k\|}{\nu \|F_k\|^\delta} \leq \frac{U_J U_F^{1-\delta}}{\nu_0},$$

where the last inequality follows from (4) and $\nu \geq \nu_0$. □

When F is continuously differentiable, Ueda and Yamashita [17] assumed that the Jacobian of F is Lipschitz continuous to investigate the global complexity bound of the LMM. However, since F is nonsmooth in this paper, the assumption does not hold in general. Instead, we assume that the gradient of the merit function f is Lipschitz continuous.

Assumption 4.2. *Let $U_d = U_J U_F^{1-\delta}/\nu_0$. ∇f is Lipschitz continuous on $\Omega + B(0, U_d)$, i.e., there exists a positive constant L such that*

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|, \quad \forall x, y \in \Omega + B(0, U_d).$$

By using the assumption, we show that $\rho_k(d^k(\nu), \nu) \geq 1$ if ν is greater than a specific value depending on F_k.

Lemma 4.2. *Suppose that Assumptions 3.1, 4.1 and 4.2 hold. Suppose also that*

$$\nu \geq \frac{\max(U_J^2, \nu_0 U_F^\delta, 4L)}{\|F_k\|^\delta}.$$

Then,

$$\rho_k(d^k(\nu), \nu) \geq 1.$$

Proof. From (4) and the assumption on ν, we have the following three inequalities.

$$\nu\|F_k^\delta\| \geq U_J^2 \geq \|J_k\|^2, \tag{5}$$

$$\nu \geq \frac{\nu_0 U_F^\delta}{\|F_k\|^\delta} \geq \nu_0, \tag{6}$$

$$\nu\|F_k^\delta\| \geq 4L. \tag{7}$$

By using (5) and Lemma 3.3, we have

$$\begin{aligned} f(x^k + d^k(\nu)) &\leq f_k(d^k(\nu), \nu) - \frac{\|J_k^T F_k\|^2}{4\nu\|F_k\|^\delta} + \frac{\|J_k^T F_k\|}{\nu\|F_k\|^\delta}\int_0^1 \|(\nabla f(x^k + \tau d^k(\nu)) - \nabla f(x^k))\|\mathrm{d}\tau \\ &\leq f_k(d^k(\nu), \nu) + \frac{\|J_k^T F_k\|}{4\nu\|F_k\|^\delta}\left(-\|J_k^T F_k\| + 4\int_0^1 \|(\nabla f(x^k + \tau d^k(\nu)) - \nabla f(x^k))\|\mathrm{d}\tau\right). \end{aligned} \tag{8}$$

On the other hand, by using (6) and Lemma 4.1, we have $x^k + \tau d^k(\nu) \in \Omega + B(0, U_d)$ for any $\tau \in [0, 1]$. It then follows from Assumption 4.2 that

$$\begin{aligned} 4\int_0^1 \|(\nabla f(x^k + \tau d^k(\nu)) - \nabla f(x^k))\|\mathrm{d}\tau &\leq 4\int_0^1 L\|\tau d^k(\nu)\|\mathrm{d}\tau \\ &\leq 4L\|d^k(\nu)\| \\ &\leq \frac{4L\|J_k^T F_k\|}{\nu\|F_k\|^\delta} \\ &\leq \|J_k^T F_k\|, \end{aligned}$$

where the second inequality follows from $\tau \in [0, 1]$, the third inequality follows from Lemma 3.2, and the last inequality follows from (7). It then follows from (8) that

$$f(x^k + d^k(\nu)) \leq f_k(d^k(\nu), \nu).$$

Therefore, by the definition of $\rho_k(d^k(\nu), \nu)$, we have

$$\rho_k(d^k(\nu), \nu) = \frac{f(x^k) - f(x^k + d^k(\nu))}{f(x^k) - f_k(d^k(\nu), \nu)} \geq 1,$$

which is the desired inequality. □

From Lemma 4.2, we can show that $\nu_k^*\|F_k\|^\delta$ is bounded from above.

Lemma 4.3. *Suppose that Assumptions 3.1, 4.1 and 4.2 hold. Then,*

$$\nu_k^*\|F_k\|^\delta \leq U_{\nu F},$$

where $U_{\nu F} = \gamma_2 \max(U_J^2, \nu_0 U_F^\delta, 4L)$.

Proof. From Lemma 4.2, if $\bar{\nu}_{l_k}\|F_k\|^\delta \geq \max(U_J^2, \nu_0 U_F^\delta, 4L)$, then $\rho_k(d^k(\bar{\nu}_{l_k}), \bar{\nu}_{l_k}) \geq 1$, and hence the inner loops of Step 2 must terminate. Therefore, if $\bar{\nu}_1\|F_k\|^\delta \geq \max(U_J^2, \nu_0 U_F^\delta, 4L)$ at the k-th iteration, then $\nu_k^*\|F_k\|^\delta = \bar{\nu}_1\|F_k\|^\delta$. On the other hand, if $\bar{\nu}_1\|F_k\|^\delta < \max(U_J^2, \nu_0 U_F^\delta, 4L)$, then $\nu_k^*\|F_k\|^\delta$ must satisfy $\nu_k^*\|F_k\|^\delta \leq \gamma_2 \max(\nu_0 U_F^\delta, 4L)$. Otherwise, $\bar{\nu}_{l_k^*-1}\|F_k\|^\delta > \max(U_J^2, \nu_0 U_F^\delta, 4L)$, which contradicts $\rho_k(d^k(\bar{\nu}_{l_k^*-1}), \bar{\nu}_{l_k^*-1}) < \eta_1 \leq 1$. Consequently, we have

$$\begin{aligned} \nu_k^*\|F_k\|^\delta &\leq \max(\bar{\nu}_1\|F_k\|^\delta, \gamma_2 U_J^2, \gamma_2\nu_0 U_F^\delta, \gamma_2 4L) \\ &= \max(\nu_{k-1}^*\|F_k\|^\delta, \gamma_2 U_J^2, \gamma_2\nu_0 U_F^\delta, \gamma_2 4L) \\ &\leq \max(\nu_{k-1}^*\|F_{k-1}\|^\delta, \gamma_2 U_J^2, \gamma_2\nu_0 U_F^\delta, \gamma_2 4L) \\ &\leq \cdots \leq \max(\nu_0\|F_0\|^\delta, \gamma_2 U_J^2, \gamma_2\nu_0 U_F^\delta, \gamma_2 4L) \\ &= \gamma_2 \max(\gamma_2 U_J^2, \nu_0 U_F^\delta, 4L) \end{aligned}$$

from the updating rule of ν. □

By using the above lemma, we give a lower bound of the reduction of the merit function when $k < K_{\text{outer}}$.

Lemma 4.4. *Suppose that Assumptions 3.1, 4.1 and 4.2 hold. Then, for all k such that* $k < K_{\text{outer}}$,

$$f(x^k) - f(x^{k+1}) > p\epsilon^2.$$

where $p = \frac{\eta_1}{2(U_J^2 + U_{\nu F})}$.

Proof. Since $\rho_k(d^k(\nu_k^*), \nu_k^*) \geq \eta_1$ from the definition of ν_k^*, we have

$$f(x^k) - f(x^{k+1}) \geq \eta_1(f(x^k) - f_k(d^k(\nu_k^*), \nu_k^*)) \geq \frac{\eta_1 \|J_k^T F_k\|^2}{2(\|J_k\|^2 + \nu_k^* \|F_k\|^\delta)}, \tag{9}$$

where the last inequality follows from Lemma 3.1. On the other hand, we have $\|J_k^T F_k\| > \epsilon$, $\forall k < K_{\text{outer}}$ from the definition of K_{outer}. It then follows from Lemma 4.3, (4) and (9) that

$$f(x^k) - f(x^{k+1}) \geq \frac{\eta_1 \|J_k^T F_k\|^2}{2(\|J_k\|^2 + \nu_k^* \|F_k\|^\delta)} \geq \frac{\eta_1}{2(U_J^2 + U_{\nu F})}\epsilon^2,$$

which is the desired inequality. □

Now, we give an upper bound of K_{outer}.

Theorem 4.1. *Suppose that Assumptions 3.1, 4.1 and 4.2 hold. Then,*

$$K_{\text{outer}} \leq \left\lceil \frac{f(x^0)}{p}\epsilon^{-2} + 1 \right\rceil.$$

Proof. Let K be $\lceil (f(x^0)\epsilon^{-2}/p) + 1 \rceil$. Suppose the contrary, i.e., $K_{\text{outer}} > K$. It then follows from Lemma 4.4 that

$$f(x^0) \geq f(x^0) - f(x^K) = \sum_{j=0}^{K-1} (f(x^j) - f(x^{j+1})) > \sum_{j=0}^{K-1} p\epsilon^2 = p\epsilon^2 K. \tag{10}$$

On the other hand, we have

$$p\epsilon^2 K = p\epsilon^2 \left\lceil \left(\frac{f(x^0)}{p\epsilon^2} \right) + 1 \right\rceil > f(x^0)$$

from the definition of K. This contradicts (10), and hence we obtain the theorem. □

From Theorem 4.1, the next theorem gives the global complexity bound K_{total} of the LMM.

Theorem 4.2. *Suppose that Assumptions 3.1, 4.1 and 4.2 hold. Then,*

$$K_{\text{total}} \leq \left\lceil \log_{\gamma_2} \left(\frac{U_{\nu F} U_J^\delta \gamma_2^{K_{\text{outer}}}}{\nu_0 \gamma_1^{K_{\text{outer}}}} \epsilon^{-\delta} \right) + 1 \right\rceil,$$

and hence $K_{\text{total}} = O(\epsilon^{-2})$.

Proof. Since $\epsilon < \|J_{K_{\text{outer}}-1}^T F_{K_{\text{outer}}-1}\| \leq U_J \|F_{K_{\text{outer}}-1}\|$ from (4), we have

$$\|F_{K_{\text{outer}}-1}\| > \frac{\epsilon}{U_J}.$$

Now we suppose the contrary of the theorem, i.e., $K_{\text{total}} > \lceil \log_{\gamma_2}(\epsilon^{-\delta} U_{\nu F} U_J^{\delta} \gamma_2^{K_{\text{outer}}} / \nu_0 \gamma_1^{K_{\text{outer}}}) + 1 \rceil$. The number of satisfying $\rho_k(d^k(\bar{\nu}_{l_k}), \bar{\nu}_{l_k}) < \eta_1$ is $\sum_{k=0}^{K_{\text{outer}}-1}(l_k^* - 1)$. Moreover, the number of satisfying $\rho_k(d^k(\bar{\nu}_{l_k}), \bar{\nu}_{l_k}) \geq \eta_2$ is at most K_{outer}. It then follows from the updating rule of ν_k that

$$\begin{aligned}
\nu_{K_{\text{outer}}-1}^* \|F_{K_{\text{outer}}-1}\|^{\delta} &> \nu_{K_{\text{outer}}-1}^* U_J^{-\delta} \epsilon^{\delta} \\
&= \nu_0 \gamma_2^{\sum_{k=0}^{K_{\text{outer}}-1}(l_k^*-1)} \gamma_1^{K_{\text{outer}}} U_J^{-\delta} \epsilon^{\delta} = \nu_0 \gamma_2^{K_{\text{total}}} \gamma_2^{-K_{\text{outer}}} \gamma_1^{K_{\text{outer}}} U_J^{-\delta} \epsilon^{\delta} \\
&> \nu_0 \gamma_2^{\log_{\gamma_2}\left(\frac{U_{\nu F} U_J^{\delta} \gamma_2^{K_{\text{outer}}}}{\nu_0 \gamma_1^{K_{\text{outer}}}} \epsilon^{-\delta}\right)} \gamma_2^{-K_{\text{outer}}} \gamma_1^{K_{\text{outer}}} \epsilon^{\delta} = U_{\nu F},
\end{aligned}$$

where the last inequality follows from the assumption that $K_{\text{total}} > \lceil \log_{\gamma_2}(U_{\nu F} U_J^{\delta} \gamma_2^{K_{\text{outer}}} / \nu_0 \gamma_1^{K_{\text{outer}}}) + 1 \rceil$. This contradicts Lemma 4.3. It then follows from Theorem 4.1 that

$$\begin{aligned}
K_{\text{total}} &\leq \left\lceil \log_{\gamma_2}\left(\frac{U_{\nu F} U_J^{\delta} \gamma_2^{K_{\text{outer}}}}{\nu_0 \gamma_1^{K_{\text{outer}}}} \epsilon^{-\delta}\right) + 1 \right\rceil \\
&= \lceil K_{\text{outer}}(1 - \log_{\gamma_2} \gamma_1) + \log_{\gamma_2} U_{\nu F} + \delta \log_{\gamma_2} U_J + \delta \log_{\gamma_2} \epsilon^{-1} - \log_{\gamma_2} \nu_0 + 1 \rceil \\
&\leq \left\lceil \left\lceil \frac{f(x^0)}{p} \epsilon^{-2} + 1 \right\rceil (1 - \log_{\gamma_2} \gamma_1) + \log_{\gamma_2} U_{\nu F} + \delta \log_{\gamma_2} U_J + \delta \log_{\gamma_2} \epsilon^{-1} - \log_{\gamma_2} \nu_0 + 1 \right\rceil,
\end{aligned}$$

and hence $K_{\text{total}} = O(\epsilon^{-2})$. □

Note that since $J_k^T F_k = 0$ does not imply $F_k = 0$, Theorem 4.2 does not provide a global complexity bound of $\|F_k\| \leq \hat{\epsilon}$ for some positive constant $\hat{\epsilon}$. To get the bound, we replace the termination criterion in Step 1 with $\|F_k\| \leq \hat{\epsilon}$ in the remainder of this section. We call the resulting method the modified LMM, and denote the total number of inner iterations of the modified LMM as $\hat{K}_{\text{total}}$. Note that since f is nonconvex, the modified LMM may not terminate. Thus, we further assume a regularity of the generalized Jacobian.

Assumption 4.3. *There exists a positive constant σ such that $\lambda_{\min}(J_k J_k^T) \geq \sigma$ for all $k \geq 0$.*

Under Assumption 4.3, we give the global complexity bound $\hat{K}_{\text{total}}$.

Theorem 4.3. *Suppose that Assumptions 3.1, 4.1, 4.2 and 4.3 hold. Then, $\hat{K}_{\text{total}} = O(\log \hat{\epsilon}^{-1})$.*

Proof. Since $\rho_k(d^k(\nu_k^*), \nu_k^*) \geq \eta_1$ from the definition of ν_k^*,

$$f(x^k) - f(x^{k+1}) \geq \eta_1(f(x^k) - f_k(d^k(\nu_k^*), \nu_k^*)) \geq \frac{\eta_1 \|J_k^T F_k\|^2}{2(\|J_k\|^2 + \nu_k^* \|F_k\|^{\delta})}, \tag{11}$$

where the last inequality follows from Lemma 3.1. On the other hand, Assumption 4.3 implies that $\|J_k^T F_k\|^2 \geq \sigma \|F_k\|^2$. It then follows from (11) that

$$\begin{aligned}
f(x^k) - f(x^{k+1}) &\geq \frac{\eta_1 \|J_k^T F_k\|^2}{2(\|J_k\|^2 + \nu_k^* \|F_k\|^{\delta})} \\
&\geq \frac{\eta_1 \sigma}{2(\|J_k\|^2 + \nu_k^* \|F_k\|^{\delta})} \|F_k\|^2 \\
&\geq \frac{\eta_1 \sigma}{2(U_J^2 + U_{\nu F})} \|F_k\|^2 \\
&= \frac{\eta_1 \sigma}{U_J^2 + U_{\nu F}} f(x^k).
\end{aligned}$$

where the third inequality follows from (4) and Lemma 4.3, and the last equality follows from the definition of f. Therefore, we have

$$f(x^k) \leq \left(1 - \frac{\eta_1 \sigma}{U_J^2 + U_{\nu F}}\right) f(x^{k-1}) \leq \left(1 - \frac{\eta_1 \sigma}{U_J^2 + U_{\nu F}}\right)^k f(x^0),$$

and hence if

$$k \geq \frac{\log \frac{2f(x^0)}{\hat{\epsilon}^2}}{\log \left(1 - \frac{\eta_1 \sigma}{U_J^2 + U_{\nu F}}\right)^{-1}}$$

then $f(x^k) \leq \hat{\epsilon}^2/2$, i.e., $\|F_k\| \leq \hat{\epsilon}$. Thus, we have $\hat{K}_{\mathrm{outer}} = O(\log \hat{\epsilon}^{-1})$, where $\hat{K}_{\mathrm{outer}}$ is the total number of outer iterations of the modified LMM. It then directly follows from Theorem 4.2 that $\hat{K}_{\mathrm{total}} = O(\log \hat{\epsilon}^{-1})$. □

When F is continuously differentiable, Ueda and Yamashita [17] also give the global complexity bound $\hat{K}_{\mathrm{total}}$ under the same regularity assumption. However, their result is $\hat{K}_{\mathrm{total}} = O(\hat{\epsilon}^{-2})$. Therefore, the result in Theorem 4.3 is much better than that of [17].

5 Application to the nonlinear complementarity problem

We apply the results obtained in the previous section to the nonlinear complementarity problem (NCP(G)) [1]: Find $x \in \mathbb{R}^n$ such that

$$x \geq 0,\ G(x) \geq 0,\ x^T G(x) = 0,$$

where $G : \mathbb{R}^n \to \mathbb{R}^n$. In this section, we assume that the mapping G satisfy the following assumptions.

Assumption 5.1.

(a) *The vector mapping G is continuously differentiable.*

(b) ∇G *is locally Lipschitz continuous.*

By using the Fischer-Burmeister function, we can reformulate NCP(G) into the following nonsmooth equations [24].

$$F(x) = \begin{pmatrix} \psi(x_1, G_1(x)) \\ \vdots \\ \psi(x_n, G_n(x)) \end{pmatrix} = 0, \tag{12}$$

where $\psi : \mathbb{R}^2 \to \mathbb{R}$ is the Fischer-Burmeister function defined by

$$\psi(a, b) = \sqrt{a^2 + b^2} - a - b.$$

Note that ψ is not differentiable at $(0, 0)$. Therefore, if there exists i such that $x_i = G_i(x) = 0$, then F is not differentiable at x. Nevertheless, F is locally Lipschitz continuous under Assumption 5.1 [25]. Moreover, the least squares merit function $f(x) = \frac{1}{2}\|F(x)\|^2$ has the following properties [26].

Lemma 5.1. *Suppose that Assumption 5.1.*

(a) F *is locally Lipschitz continuous.*

(b) f *is continuously differentiable.*

(c) ∇f *is locally Lipschitz continuous.*

Lemma 5.1 (c) implies that ∇f is Lipschitz continuous on any compact set.

By using Lemma 5.1, we get the global complexity bound of the LMM for the equations (12) equivalent to the NCP as a direct application of Theorem 4.2.

Theorem 5.1. *Suppose that Assumption 5.1 holds. Suppose also that $\delta \leq 1$ and a sequence generated by the LMM is bounded. Then, the global complexity bound of the LMM for the NCP is $O(\epsilon^{-2})$.*

Remark 5.1. *A sequence generated by the LMM is bounded if the level set of f is compact. The level set of f is compact if G is a uniformly P-function [25] (see Assumption 5.2 for the definition). The level set of f is also compact if G is monotone, NCP(G) has a strictly feasible solution and the Fischer-Burmeister function is replaced with the penalized Fischer-Burmeister function $\psi_\tau(a,b) = \tau\psi(a,b) + (1-\tau)\max(0,a)\max(0,b)$, where $\tau \in (0,1)$ is an arbitrary but fixed constant [27].*

Remark 5.2. *Note that the bound in Theorem 5.1 is not for a solution of NCP(G) but for a stationary point of f. However, a stationary point of f is a solution of NCP(G) if G is P_0-function, i.e., there exists i such that $x_i \neq y_i$ and $(x_i - y_i)(G_i(x) - G_i(y)) \geq 0, \forall x, y \in \mathbb{R}^n$ [25].*

Next, as related to Assumption 4.3, we further make the following assumption on G.

Assumption 5.2. *G is a uniformly P-function, i.e., there exists a positive constant $\alpha > 0$ such that*

$$\max_{1\leq i\leq n}(x_i - y_i)(G_i(x) - G_i(y)) \geq \alpha\|x-y\|^2, \forall x, y \in \mathbb{R}^n.$$

When G is a uniformly P-function, it is well-known that the following properties hold [25, 28, 29, 30].

Lemma 5.2. *Suppose that Assumptions 5.1 and 5.2.*

(a) *The level set Ω of the merit function f is compact.*

(b) *For any $J \in \partial F(x)$ and $x \in \mathbb{R}^n$, J is nonsingular.*

(c) *The NCP(G) has a unique solution x^*.*

(d) *There exists a positive constant c such that $\|x - x^*\| \leq c\|F(x)\|$ for any $x \in \Omega$.*

From Lemma 5.2 (a), (b) and the upper semi-continuity of the generalized Jacobian ∂F, there exists a positive constant σ such that $\lambda_{\min}(JJ^T) \geq \sigma$ for any $J \in \partial F(x)$ and $x \in \Omega$. Therefore, Assumption 4.3 holds.

Now, we get the bound for the NCP(G) as a direct application of Theorem 4.3.

Theorem 5.2. *Suppose that Assumptions 5.1 and 5.2 hold. Suppose also that $\delta \leq 1$. Then, the global complexity bound of the modified LMM defined in Section 4 is $O(\log \hat{\epsilon}^{-1})$. Moreover, for an approximate solution $\hat{x}$ such that $\|F(\hat{x})\| \leq \hat{\epsilon}$, the distance $\|\hat{x} - x^*\| = O(\hat{\epsilon})$.*

Proof. The first part of the theorem directly follows from Theorem 4.3. The second part of the theorem follows from Lemma 5.2 (d) and the assumption on $\hat{x}$. □

6 Concluding remarks

In this paper, we have investigated the global complexity bound of the LMM for the nonsmooth equations. We have shown that the bound is $O(\epsilon^{-2})$ without any regularity or convex assumptions. We have also shown that the bound is $O(\log \epsilon^{-1})$ under the regularity assumption of the generalized Jacobian. Moreover, by applying these results to the NCP, we have obtained the same global complexity bounds of the LMM for the NCP. In this paper, we have assumed that the mapping G involved in the NCP is a uniformly P-function for the regularity assumption of the generalized Jacobian. By using other assumption such as the monotonicity of G, we may have a better global complexity bound. Furthermore, it would be worth estimating global complexity bounds of other solution methods for the NCP such as the generalized Newton's method [28, 31].

References

[1] Facchinei, F., Pang, J.-S.: Finite-Dimensional Variational Inequalities and Complementarity Problems, Springer-Verlag, New York (2003)

[2] Facchinei, F., Kanzow, C.: A nonsmooth inexact Newton method for the solution of large-scale nonlinear complementarity problems, Mathematical Programming, 76, 493–512 (1997)

[3] Jiang, H., Ralph, D.: Global and local superlinear convergence analysis of Newton-type methods for semismooth equations with smooth least squares, in Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods, In: Fukushima, M., Qi, L. (eds.): Kluwer Academic Publishers, Dordrecht, pp. 181–209 (1999)

[4] Kanzow, C., Yamashita, N., Fukushima, M.: Levenberg-Marquardt methods with strong local convergence properties for solving nonlinear equations with convex constraints, Journal of Computational and Applied Mathematics, 173, 321–343 (2005)

[5] Ma, C., Tang, J.: The quadratic convergence of a smoothing Levenberg-Marquardt method for nonlinear complementarity problem, Applied Mathematics and Computation, 197, 566–581 (2008)

[6] Ma, C., Tang, J., Chen, X.: A globally convergent Levenberg-Marquardt method for solving nonlinear complementarity problem, Applied Mathematics and Computation, 192, 370–381 (2007)

[7] Sabramanian, P. K.: Gauss-Newton methods for the complementarity problem, Journal of Optimization Theory and Applications, 77, 467–482 (1993)

[8] Zhang, J., Zhang, X.: A smoothing Levenberg-Marquardt method for NCP, Applied Mathematics and Computation, 178, 212–228 (2006)

[9] 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, Technical Report 09/14, Department of Mathematics, FUNDP - University of Namur (2009)

[10] ———: Adaptive cubic regularization methods for unconstrained optimization. PartII: worst-case function- and derivative-evaluation complexity, Mathematical Programming, doi: 10.1007/s10107-009-0337-y.

[11] Gratton, S., Sartenaer, A., Toint, P. L.: Recursive trust-region methods for multiscale nonlinear optimization, SIAM Journal on Optimization, 19, 414–444 (2008)

[12] Nesterov, Yu.: Introductory Lectures on Convex Optimization, Kluwer Academic Publishers, Dordrecht (2004)

[13] ———: Accelerating the cubic regularization of Newton's method on convex problems, Mathematical Programming, 112, 159–181 (2008)

[14] Nesterov, Yu., Polyak, B. T.: Cubic regularization of Newton method and its global performance, Mathematical Programming, 108, 177–205 (2006)

[15] Polyak, R. A.: Regularized Newton method for unconstrained convex optimization, Mathematical Programming, 120, 125–145 (2009)

[16] Ueda, K., Yamashita, N.: Convergence properties of the regularized Newton method for the unconstrained nonconvex optimization, Applied Mathematics and Optimization, 62, 27–46 (2010)

[17] ———: On a global complexity bound of the Levenberg-Marquardt method, Journal of Optimization Theory and Applications, 147, 443–453 (2010)

[18] Clarke, F. H.: Optimization and Nonsmooth Analysis, SIAM, Philadelphia (1983)

[19] Qi, L.: Convergence analysis of some algorithms for solving nonsmooth equations, Mathematics of Operations Research, 18, 227–244 (1993)

[20] Qi, L., Sun, J.: A nonsmooth version of Newton's method, Mathematical Programming, 58, 353–367 (1993)

[21] Hogan, W. W.: Point-to-set maps in mathematical programming, SIAM Review, 15, 591–603 (1973)

[22] Moré, J. J.: The Levenberg-Marquardt algorithm: Implementation and theory, Numerical Analysis, 630, 105–116 (1978)

[23] Osborne, M. R.: Nonlinear least squares — the Levenberg algorithm revisited, Journal of the Australian Mathematical Society, 19, 343–357 (1976)

[24] Fischer, A.: A special Newton-type optimization method, Optimization, 24, 269–284 (1992)

[25] Facchinei, F., Soares, J.: A new merit function for nonlinear complementarity problems and a related algorithm, SIAM Journal on Optimization, 7, 225–247 (1997)

[26] Chen, J.-S.: The semismooth-related properties of a merit function and a descent method for the nonlinear complementarity problem, Journal of Global Optimization, 36, 565–580 (2006)

[27] Chen, B., Chen, X., Kanzow, C.: A penalized Fischer-Burmeister NCP-function, Mathematical Programming, 88, 211–216 (2000)

[28] Jiang, H., Qi, L.: A new nonsmooth equations approach to nonlinear complementarity problems, SIAM Journal on Control and Optimization, 35, 178–193 (1997)

[29] Kanzow, C., Fukushima, M.: Equivalence of the generalized complementarity problem to differentiable unconstrained minimization, Journal of Optimization Theory and Applications, 90, 581–603 (1996)

[30] Moré, J. J.: Classes of functions and feasibility conditions in nonlinear complementarity problems, Mathematical Programming, 6, 327–338 (1974)

[31] Luca, T. D., Facchinei, F., Kanzow, C.: A semismooth equation approach to the solution of nonlinear complementarity problems, Mathematical Programming, 75, 407–439 (1996)