Journal of Optimization Theory and Applications
https://doi.org/10.1007/s10957-022-02152-6

An Accelerated Smoothing Newton Method with Cubic Convergence for Weighted Complementarity Problems

Jingyong Tang[1] · Jinchuan Zhou[2] · Hongchao Zhang[3]

Received: 6 September 2022 / Accepted: 16 December 2022

Abstract
Smoothing Newton methods, which usually inherit local quadratic convergence rate, have been successfully applied to solve various mathematical programming problems. In this paper, we propose an accelerated smoothing Newton method (ASNM) for solving the weighted complementarity problem (wCP) by reformulating it as a system of nonlinear equations using a smoothing function. In spirit, when the iterates are close to the solution set of the nonlinear system, an additional approximate Newton step is computed by solving one of two possible linear systems formed by using previously calculated Jacobian information. When a Lipschitz continuous condition holds on the gradient of the smoothing function at two checking points, this additional approximate Newton step can be obtained with a much reduced computational cost. Hence, ASNM enjoys local cubic convergence rate but with computational cost only comparable to standard Newton's method at most iterations. Furthermore, a second-order nonmonotone line search is designed in ASNM to ensure global convergence. Our numerical experiments verify the local cubic convergence rate of ASNM and

Communicated by Dr. Qianchuan Zhao.

This work was supported by the National Natural Science Foundation of China (11771255), the Shandong Province Natural Science Foundation (ZR2021MA066, 2019KJI013), the Henan Province Natural Science Foundation (222300420520), the Key Scientific Research Projects of Higher Education of Henan Province (22A110020) and the USA National Science Foundation (1819161, 2110722).

✉ Hongchao Zhang
hozhang@math.lsu.edu

Jingyong Tang
tangjy@xynu.edu.cn

Jinchuan Zhou
jinchuanzhou@163.com

1 School of Mathematics and Statistics, Xinyang Normal University, Xinyang 464000, China

2 School of Mathematics and Statistics, Shandong University of Technology, Zibo 255049, China

3 Department of Mathematics, Louisiana State University, 303 Lockett Hall, Baton Rouge, LA 70803-4918, USA

Published online: 30 December 2022

show that the acceleration techniques employed in ASNM can significantly improve the computational efficiency compared with some well-known benchmark smoothing Newton method.

Keywords Nonlinear programming · Weighted complementarity problem · Accelerated smoothing Newton method · Nonmonotone line search · Cubic convergence

1 Introduction

In 2012, Potra [23] introduced the notion of a weighted complementarity problem (wCP), which consists in finding a pair of vectors (x, s) belonging to the intersection of a manifold with a cone, such that their product in a certain algebra, $x \circ s$, equals a given weight vector w. The significance of studying the wCP lies in the fact that a large variety of equilibrium problems in economics can be formulated in a natural way as wCP [23]. Moreover, those formulations could lead to the development of highly efficient algorithms for solving the corresponding equilibrium problems [23]. More details on the applications of wCP can be found in the survey paper [25].

After the introduction of wCP by Potra [23], many researchers studied the weighted linear complementarity problem (wLCP) over the nonnegative orthant of $\mathbf{R}^n$, which finds $(x, s, y) \in \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^m$ such that

$$\text{(wLCP)} \quad x \geq 0,\ s \geq 0,\ Px + Qs + Ry = a,\ xs = w. \tag{1}$$

Here, $P \in \mathbf{R}^{(n+m)\times n}$, $Q \in \mathbf{R}^{(n+m)\times n}$, $R \in \mathbf{R}^{(n+m)\times m}$ and $a \in \mathbf{R}^{n+m}$ are given matrices and vector with R assuming to have full column rank, $w \geq 0$ is a given weight vector and xs denotes the component-wise product. Various types of interior-point methods [1, 23, 24], smoothing Newton-type methods [32, 40] and damped Gauss–Newton methods [35] have been studied for solving wLCP. More recently, infeasible interior-point methods as well as their computational complexities were proposed in [5, 6] for solving the special wLCP:

$$\text{(special wLCP)} \quad x \geq 0,\ s \geq 0,\ Ax = b,\ A^T y + s = c,\ \ xs = w, \tag{2}$$

which often appears in the iterations of interior-point methods for linear programming. In 2021, Tang and Zhou [36] proposed a nonmonotone Levenberg–Marquardt type method, which has local quadratic convergence under some local error bound condition, to solve the weighted nonlinear complementarity problem (wNCP):

$$\text{(wNCP)} \quad x \geq 0,\ s \geq 0,\ F(x, s, y) = 0,\ xs = w, \tag{3}$$

where $F : \mathbf{R}^{2n+m} \to \mathbf{R}^{n+m}$ is a continuously differentiable nonlinear function.

In this paper, we aim to study the wCP in a general Jordan algebra [8] setting since this general framework reveals the essential geometric features of a wCP. Let $(\mathbf{V}, \langle\cdot,\cdot\rangle, \circ)$ be a Euclidean Jordan algebra (see the definition in Sect. 2) and $\mathbf{K} =$

$\{x \circ x : x \in \mathbf{V}\}$ be the symmetric cone formed by the squares of its elements. Given a vector $w \in \mathbf{K}$, we consider the wCP which finds $(x, s, y) \in \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m$ such that

$$\text{(wCP)} \quad x \in \mathbf{K},\ s \in \mathbf{K},\ F(x, s, y) = 0,\ x \circ s = w, \tag{4}$$

where $F : \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m \to \mathbf{V} \times \mathbf{R}^m$ is a continuously differentiable map. When w is the zero vector, wCP would reduce to the complementarity problem (CP) over symmetric cones, for example, studied in [39], which finds $(x, s, y) \in \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m$ such that

$$\text{(CP)} \quad x \in \mathbf{K},\ s \in \mathbf{K},\ F(x, s, y) = 0,\ x \circ s = 0. \tag{5}$$

The wCP in general Jordan algebra setting has been also studied in the literature. Chi et. al. [4] considered the weighted horizontal LCP in the setting of Euclidean Jordan algebras and established some existence and uniqueness results. Gowda [9] studied wLCPs and interior-point systems for copositive linear transformations on Euclidean Jordan algebras and showed that both problems have solutions under suitable conditions. Tang and Zhang [33] proposed a nonmonotone smoothing Newton algorithm for solving the general wCP (4) and established its local quadratic convergence under proper assumptions weaker than the usual nonsingularity assumption of the considered nonlinear system.

Since Smale [29] initiated the study on smoothing (non-interior continuation) Newton-type methods for solving linear programming problems and LCPs, there has been much interests in studying smoothing Newton-type methods. The main idea of this class of methods is to use a smoothing function to reformulate the problem concerned as a system of smooth nonlinear equations and then solve it approximately by Newton method. When driving the smoothing parameter to zero, one can expect to find a solution of the original problem. Note that the convergence rates of many smoothing Newton-type methods (e.g., [7, 31]) strongly depend on the strict complementarity condition. In 2000, Qi, Sun and Zhou [28] proposed a class of new smoothing Newton methods for solving NCPs and box constrained variational inequality problems. The Qi–Sun–Zhou (QSZ) method treats the smoothing parameter as a free variable and solves one system of linear equations (the Newton step) at each iteration. Based on the strong semismoothness of the smoothing function, QSZ method possess local quadratic convergence under the nonsingularity assumption. Due to its encouraging convergent properties and numerical performances, the QSZ algorithmic framework has been extensively studied to deal with various optimization problems (e.g., [2, 3, 11–21, 26, 37, 40, 41]).

It is well known that the convergence rate of the classical two-step Newton method for solving nonlinear system is cubic if its Jacobian is Lipschitz continuous and nonsingular at the solution [22]. Motivated from this observation and the QSZ method, in this paper, we propose to design an accelerated smoothing Newton method (ASNM) for solving wCP (4). Similar to the QSZ method, ASNM computes one approximate Newton step at the beginning iterations. But when the iteration points are close to a solution, an additional approximate Newton step is calculated by solving one of two possible linear systems by using previously calculated Jacobian information. A Lipschitz continuous condition on the gradient of the smoothing function at two checking points is used to determine which linear system should be solved. This technique ensures ASNM has local cubic convergence rate but with computational cost comparable to standard Newton's method at most iterations. In addition, a second-order nonmonotone line search is proposed in ASNM to ensure global convergence. Based on the strong semismoothness of the smoothing function, we show that the convergence rate of ASNM is cubic under certain nonsingularity assumption and the assumption that gradient of F is locally Lipschitz continuous, which obviously holds when F is a linear map. Moreover, we give brief discussions on how to solve the resulted approximate Newton system efficiently by equivalently translating it to a smaller dimensional problem. Our numerical experiments verify the local cubic convergence rate as well as show that ASNM is much more efficient comparing with the well-established QSZ method. To the best of our knowledge, this is the first smoothing Newton-type method for solving wCPs in mathematical programming which achieves local cubic convergence rate.

This paper is organized as follows. We reformulate wCP as a system of smooth nonlinear equations in Sect. 2 and our ASNM is proposed for solving wCP in Sect. 3. The global convergence of ASNM is studied in Sect. 4, while the local cubic convergence properties of ASNM are analyzed in Sect. 5. We further give some discussions on solving the (approximate) Newton equations in Sect. 6 and show some numerical experiments in Sect. 7. Finally, some conclusions are drawn in the last section.

2 Equivalent Reformulation of the wCP

As it is well known, the theory of Euclidean Jordan algebras is a basic tool for analyzing optimization problems over symmetric cones. A *Euclidean Jordan algebra* is a triple $(\mathbf{V}, \langle\cdot,\cdot\rangle, \circ)$, where $(\mathbf{V}, \langle\cdot,\cdot\rangle)$ is a finite dimensional inner product space over $\mathbf{R}$ and $(x, s) \mapsto x \circ s : \mathbf{V} \times \mathbf{V} \to \mathbf{V}$ is a bilinear mapping satisfying the following conditions:

(i) $x \circ s = s \circ x$ for all $x, s \in \mathbf{V}$;

(ii) $x \circ (x^2 \circ s) = x^2 \circ (x \circ s)$ for all $x, s \in \mathbf{V}$, where $x^2 := x \circ x$;

(iii) $\langle x \circ s, z\rangle = \langle x, s \circ z\rangle$ for all $x, s, z \in \mathbf{V}$.

We call $x \circ s$ the *Jordan product* of x and s. For a comprehensive discussion of Jordan algebras, one can refer to [8]. Some basic results and properties of Euclidean Jordan algebras can be found in [13, 15–20, 33].

To reformulate wCP (4) as a system of smooth nonlinear equations, we apply the following smoothing function

$$\psi(\mu, x, s) = x + s - \sqrt{(x - s)^2 + 4w + 4\mu^2 e}, \quad \forall(\mu, x, s) \in \mathbf{R} \times \mathbf{V} \times \mathbf{V}, \tag{6}$$

where w is the weight vector given in the wCP and e is the identity element in $\mathbf{V}$ satisfying $e \circ x = x \circ e = x$ for all $x \in \mathbf{V}$. The function ψ is a special case of the weighted smoothing function introduced in [33]. By [33, Theorem 3], we have that ψ is continuously differentiable on $\mathbf{R}_{++} \times \mathbf{V} \times \mathbf{V}$ and it satisfies

$$\psi(0, x, s) = 0 \Longleftrightarrow x \in \mathbf{K},\ s \in \mathbf{K},\ x \circ s = w. \tag{7}$$

In the following, let us denote $z = (\mu, x, s, y) \in \mathbf{R} \times \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m$ and define the function $\mathrm{H}(z)$ as

$$\mathrm{H}(z) = \begin{pmatrix} \mu \\ F(x, s, y) \\ \psi(\mu, x, s) \end{pmatrix}, \tag{8}$$

where the smoothing function ψ is given in (6). Then, from (7) it holds that

$$\mathrm{H}(z) = 0 \iff \mu = 0 \ \text{ and } (x, s, y) \text{ is a solution of the wCP (4).} \tag{9}$$

Moreover, $\mathrm{H}(z)$ is continuously differentiable at any $z \in \mathbf{R}_{++} \times \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m$ and its Jacobian is

$$\mathrm{H}'(z) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & F'_x(x, s, y) & F'_s(x, s, y) & F'_y(x, s, y) \\ \psi'_\mu(\mu, x, s) & \psi'_x(\mu, x, s) & \psi'_s(\mu, x, s) & 0 \end{bmatrix}. \tag{10}$$

Here, $\psi'_\mu(\mu, x, s)$, $\psi'_x(\mu, x, s)$ and $\psi'_s(\mu, x, s)$ are the derivatives of ψ with respect to μ, x and s, respectively. In particular, given $c = \sqrt{(x - s)^2 + 4w + 4\mu^2 e}$, defining the *Lyapunov transformation* $\mathcal{L}_c x := x \circ c,\ \forall x \in \mathbf{V}$, then for any $(\mu, x, s) \in \mathbf{R}_{++} \times \mathbf{V} \times \mathbf{V}$ and $(h, u, v) \in \mathbf{R} \times \mathbf{V} \times \mathbf{V}$, we will have

$$\psi'_\mu(\mu, x, s)h = -\sum_{i=1}^{r} \frac{4\mu h}{\sqrt{\lambda_i + 4\mu^2}} c_i,$$

$$\psi'_x(\mu, x, s)u = u - \mathcal{L}_c^{-1}[(x - s) \circ u] \text{ and } \psi'_s(\mu, x, s)v = v + \mathcal{L}_c^{-1}[(x - s) \circ v],$$

where (λ_i, c_i), $i, \ldots, r$, are given by the *spectral decomposition* of $(x - s)^2 = \sum_{i=1}^{r} \lambda_i c_i$ and $\mathcal{L}_c^{-1}$ is the inverse of $\mathcal{L}_c$. One could refer to [33, Theorem 3] for more detail explanations.

To discuss convergence properties of our accelerated smoothing Newton method, we define

$$\mathrm{J}_{\mathrm{H}}(z, \hat{z}) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & F'_x(x, s, y) & F'_s(x, s, y) & F'_y(x, s, y) \\ \psi'_\mu(\hat{\mu}, \hat{x}, \hat{s}) & \psi'_x(\hat{\mu}, \hat{x}, \hat{s}) & \psi'_s(\hat{\mu}, \hat{x}, \hat{s}) & 0 \end{bmatrix}, \tag{11}$$

where $z = (\mu, x, s, y) \in \mathbf{R}_{++} \times \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m$ and $\hat{z} = (\hat{\mu}, \hat{x}, \hat{s}, \hat{y}) \in \mathbf{R}_{++} \times \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m$. Obviously, $\mathrm{H}'(z) = \mathrm{J}_{\mathrm{H}}(z, \hat{z})$ when $z = \hat{z}$. In addition, we assume that $F'(x, s, y)$ has the following rank and monotone property.

Assumption 1 rank $F'_y(x, s, y) = m$ and for any $(u, v, \Lambda) \in \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m$,

$$F'(x, s, y)(u, v, \Lambda) = 0 \Longrightarrow \langle u, v \rangle \geq 0.$$

Assumption 1 is very standard and has been extensively used to analyze smoothing Newton-type methods (e.g., [2, 21, 33, 34]). For wLCP (1), Assumption 1 in fact reduces to require that the matrix R has full column rank and the wLCP is monotone, that is,

$$Pu + Qv + R\Lambda = 0 \quad \text{implies} \quad u^T v \geq 0,$$

which was originally introduced by Potra [23] and has been widely used to analyze interior-point methods [1, 23] and smoothing Newton methods [32, 40]. We now show the nonsingularity property of $\mathrm{J_H}(z, \hat{z})$ under Assumption 1, which plays a key role for our ASNM.

Theorem 1 *If Assumption* 1 *holds, then* $\mathrm{J_H}(z, \hat{z})$ *defined by (11) is nonsingular at any* $z = (\mu, x, s, y) \in \mathbf{R}_{++} \times \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m$ *and* $\hat{z} = (\hat{\mu}, \hat{x}, \hat{s}, \hat{y}) \in \mathbf{R}_{++} \times \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m$.

Proof Let int$\mathbf{K}$ denote the interior of $\mathbf{K}$ and we write $u \succ_{\mathbf{K}} v$ if $u - v \in \mathrm{int}\mathbf{K}$ for any $u, v \in \mathbf{V}$. For any $z = (\mu, x, s, y) \in \mathbf{R}_{++} \times \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m$ and $\hat{z} = (\hat{\mu}, \hat{x}, \hat{s}, \hat{y}) \in \mathbf{R}_{++} \times \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m$, it suffices to prove that the following system

$$F'_x(x, s, y)\Delta x + F'_s(x, s, y)\Delta s + F'_y(x, s, y)\Delta y = 0, \tag{12}$$

$$\psi'_x(\hat{\mu}, \hat{x}, \hat{s})\Delta x + \psi'_s(\hat{\mu}, \hat{x}, \hat{s})\Delta s = 0, \tag{13}$$

has only zero solution. Let $\hat{c} = \sqrt{(\hat{x} - \hat{s})^2 + 4w + 4\hat{\mu}^2 e}$. Since $w \in \mathbf{K}$ and $\hat{\mu} > 0$, we have $\hat{c} \succ_{\mathbf{K}} 0$ and $\hat{c}^2 \succ_{\mathbf{K}} (\hat{x} - \hat{s})^2 = (\hat{s} - \hat{x})^2$, which together with [10, Proposition 8] gives

$$\hat{c} - (\hat{x} - \hat{s}) \succ_{\mathbf{K}} 0, \quad \hat{c} - (\hat{s} - \hat{x}) \succ_{\mathbf{K}} 0. \tag{14}$$

In addition, by $w \in \mathbf{K}$ and $\hat{\mu} > 0$, it holds

$$[\hat{c} - (\hat{x} - \hat{s})] \circ [\hat{c} - (\hat{s} - \hat{x})] = 4w + 4\hat{\mu}^2 e \succ_{\mathbf{K}} 0. \tag{15}$$

Moreover, Assumption 1 and (12) imply

$$\langle \Delta x, \Delta s \rangle \geq 0. \tag{16}$$

By [33, Theorem 3], we have

$$\psi'_x(\hat{\mu}, \hat{x}, \hat{s})\Delta x = \Delta x - \mathcal{L}_{\hat{c}}^{-1}[(\hat{x} - \hat{s}) \circ \Delta x],$$

$$\psi'_s(\hat{\mu}, \hat{x}, \hat{s})\Delta s = \Delta s + \mathcal{L}_{\hat{c}}^{-1}[(\hat{x} - \hat{s}) \circ \Delta s].$$

This and (13) give

$$\mathcal{L}_{\hat{c}}(\Delta x + \Delta s) - [(\hat{x} - \hat{s}) \circ (\Delta x - \Delta s)] = 0,$$

i.e.,

$$\hat{c} \circ (\Delta x + \Delta s) - (\hat{x} - \hat{s}) \circ (\Delta x - \Delta s) = 0.$$

It follows that

$$[\hat{c} - (\hat{x} - \hat{s})] \circ \Delta x + [\hat{c} - (\hat{s} - \hat{x})] \circ \Delta s = 0. \tag{17}$$

By (14)–(17), we can obtain from [39, Lemma 2.6][1] that $\Delta x = 0$ and $\Delta s = 0$. So, by (12) we have $F_y'(x, s, y)\Delta y = 0$, which and the assumption on $\operatorname{rank} F_y'(x, s, y) = m$ give $\Delta y = 0$. We complete the proof. □

Since $\mathrm{H}'(z) = \mathrm{J}_{\mathrm{H}}(z, z)$, we immediately have the following corollary.

Corollary 1 *If Assumption 1 holds, then $\mathrm{H}'(z)$ defined by (10) is nonsingular at any $z = (\mu, x, s, y) \in \mathbf{R}_{++} \times \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m$.*

3 An Accelerated Smoothing Newton Method

Let $\mathrm{H}(z)$ be defined by (8) and the merit function $f : \mathbf{R} \times \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m \to \mathbf{R}_+$ be defined as

$$f(z) = \frac{1}{2}\|\mathrm{H}(z)\|^2. \tag{18}$$

Then, our accelerated smoothing Newton method is described as the following.

Algorithm 1: An accelerated smoothing Newton method (ASNM)

Step 0: Choose $\delta, \tau \in (0, 1)$, $\lambda > 0$, $\mu_0 > 0$ and $L > 0$. Choose $(x^0, s^0, y^0) \in \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m$, set $z^0 = (\mu_0, x^0, s^0, y^0)$ and $C_0 = f(z^0) + c$, where $c > 0$ is a constant. Choose $\gamma \in (0, 1)$ such that $\gamma < \min\left\{\frac{\mu_0}{C_0^{3/2}}, \frac{2}{\mu_0 C_0^{1/2}}\right\}$. Let $h = (1, 0, 0, 0) \in \mathbf{R} \times \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m$. Set $k = 0$.
Step 1: If $\|\mathrm{H}(z^k)\| = 0$, then stop.
Step 2: Compute $\Delta \bar{z}^k = (\Delta \bar{\mu}_k, \Delta \bar{x}^k, \Delta \bar{s}^k, \Delta \bar{y}^k) \in \mathbf{R} \times \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m$ by solving

$$\mathrm{H}'(z^k)\Delta \bar{z}^k = -\mathrm{H}(z^k) + \gamma C_k^{\frac{3}{2}} h. \tag{19}$$

Set $\hat{z}^k = z^k + \Delta \bar{z}^k$. If $\|\mathrm{H}(\hat{z}^k)\| = 0$, then stop.
Step 3: If

$$\|\mathrm{H}(\hat{z}^k)\| > \lambda \min\{1, \|\mathrm{H}(z^k)\|\}, \tag{20}$$

then set $\Delta \hat{z}^k = 0$ and go to Step 5.
Step 4: If

$$\|\psi'(\mu^k, x^k, s^k) - \psi'(\hat{\mu}^k, \hat{x}^k, \hat{s}^k)\| \leq L\|(\mu^k, x^k, s^k) - (\hat{\mu}^k, \hat{x}^k, \hat{s}^k)\|, \tag{21}$$

[1] [39, Lemma 2.6]: Let $a, b \in \mathbf{V}$ with $a \succ_{\mathbf{K}} 0$, $b \succ_{\mathbf{K}} 0$ and $a \circ b \succ_{\mathbf{K}} 0$. Then for all $u, v \in \mathbf{V}$ satisfying $\langle u, v\rangle \geq 0$ and $a \circ u + b \circ v = 0$, we have $u = v = 0$.

then compute $\Delta\hat{z}^k = (\Delta\hat{\mu}_k, \Delta\hat{x}^k, \Delta\hat{s}^k, \Delta\hat{y}^k) \in \mathbf{R} \times \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m$ by solving

$$\mathrm{H}'(z^k)\Delta\hat{z}^k = -\mathrm{H}(\hat{z}^k) + \gamma C_k^{\frac{3}{2}} h; \tag{22}$$

Else, compute $\Delta\hat{z}^k = (\Delta\hat{\mu}_k, \Delta\hat{x}^k, \Delta\hat{s}^k, \Delta\hat{y}^k) \in \mathbf{R} \times \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m$ by solving

$$\mathrm{J_H}(z^k, \hat{z}^k)\Delta\hat{z}^k = -\mathrm{H}(\hat{z}^k) + \gamma C_k^{\frac{3}{2}} h. \tag{23}$$

Step 5: Set $\alpha_k = \delta^{l_k}$, where l_k is the smallest nonnegative integer l satisfying

$$f(z^k + \delta^l \Delta\bar{z}^k + (\delta^l)^2 \Delta\hat{z}^k) \leq C_k - \tau(\delta^l f(z^k))^2. \tag{24}$$

Step 6: Set $z^{k+1} = z^k + \alpha_k \Delta\bar{z}^k + \alpha_k^2 \Delta\hat{z}^k$. Set

$$C_{k+1} = \frac{(C_k + 1) f(z^{k+1})}{f(z^{k+1}) + 1}. \tag{25}$$

Set $k := k + 1$ and go to Step 1.

For ASNM, we have the following remarks.

(I) Different from existing smoothing Newton methods, a new perturbation term $\gamma C_k^{3/2}$ is used into the approximated Newton equations (19), (22) and (23) in Step 2 and Step 4 of ASNM. This particularly designed perturbation term is crucial for establishing the local cubic convergence of ASNM.

(II) As QSZ method [28], ASNM only solves one system of linear equations when the condition (20) holds. Otherwise, ASNM solves an additional system of linear equations. However, different from the classical two-step Newton methods, ASNM includes two possible approximated Newton steps (22) and (23) by using previously obtained Jacobian information. The Lipschitz continuous condition (21) is used to determine which one should be solved. This technique ensures ASNM has cubic convergence rate but with computational cost only comparable to standard Newton's method at most iterations.

(III) Note that $\Delta\bar{z}^k + \Delta\hat{z}^k$ calculated by ASNM may not be a decent direction of the merit function f at z^k. Hence, to ensure the global convergence of ASNM, we adopt a second-order nonmonotone line search in Step 5. It is worth pointing out that the technique of updating C_k in (25) had been originally introduced in [33].

The following theorem shows ASNM is well defined and gives some of its basic properties.

Theorem 2 *If Assumption* 1 *holds, then* ASNM *is well defined and its generated sequence* $\{z^k = (\mu_k, x^k, s^k, y^k)\}$ *satisfies for all* $k \geq 0$,

(i) $z^k \in \mathbf{R}_{++} \times \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m$; (ii) $C_k > f(z^k)$; (iii) $\mu_k > \gamma C_k^{\frac{3}{2}}$.

Proof We will prove the theorem by induction. Suppose $z^k = (\mu_k, x^k, s^k, y^k) \in \mathbf{R}_{++} \times \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m$, $C_k > f(z^k)$ and $\mu_k > \gamma C_k^{\frac{3}{2}}$ for some k. Then, $\mathrm{H}'(z^k)$ is nonsingular by Corollary 1. Moreover, by the first equation of (19), we have $\Delta\bar{\mu}_k = -\mu_k + \gamma C_k^{\frac{3}{2}}$, which yields

$$\hat{z}^k = (\gamma C_k^{\frac{3}{2}}, x^k + \Delta\bar{x}^k, s^k + \Delta\bar{s}^k, y^k + \Delta\bar{y}^k). \tag{26}$$

This indicates $\hat{z}^k \in \mathbf{R}_{++} \times \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m$. Then, it follows from Theorem 1 that $\mathrm{J_H}(z^k, \hat{z}^k)$ is nonsingular. Thus, Eqs. (19), (22) and (23) are all solvable. Moreover, by (19) we have

$$\begin{aligned}\nabla f(z^k)^T \Delta\bar{z}^k &= \mathrm{H}(z^k)^T \mathrm{H}'(z^k)\Delta\bar{z}^k \\ &= -\|\mathrm{H}(z^k)\|^2 + \mu_k \gamma C_k^{\frac{3}{2}} \\ &= -\mu_k(\mu_k - \gamma C_k^{\frac{3}{2}}) - \|F(x^k, s^k, y^k)\|^2 - \|\psi(\mu_k, x^k, s^k)\|^2,\end{aligned}$$

which together with $\mu_k > 0$ and $\mu_k > \gamma C_k^{\frac{3}{2}}$ yields

$$\nabla f(z^k)^T \Delta\bar{z}^k < 0. \tag{27}$$

Notice that if for any nonnegative integer l,

$$f(z^k + \delta^l \Delta\bar{z}^k + (\delta^l)^2 \Delta\hat{z}^k) > C_k - \tau(\delta^l f(z^k))^2,$$

then by $C_k > f(z^k)$ we have

$$\frac{f(z^k + \delta^l(\Delta\bar{z}^k + \delta^l \Delta\hat{z}^k)) - f(z^k)}{\delta^l} > -\tau\delta^l f(z^k)^2. \tag{28}$$

By letting $l \to \infty$ in both sides of (28), we have $\nabla f(z^k)^T \Delta\bar{z}^k \geq 0$ which contradicts (27). Thus, we can find a step size $\alpha_k \in (0, 1]$ satisfying (24) in Step 5 and get the $(k+1)$th iteration $z^{k+1} = z^k + \alpha_k \Delta\bar{z}^k + \alpha_k^2 \Delta\hat{z}^k$ in Step 6.

Now we show $z^{k+1} = (\mu_{k+1}, x^{k+1}, s^{k+1}, y^{k+1}) \in \mathbf{R}_{++} \times \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m$, $C_{k+1} > f(z^{k+1})$ and $\mu_{k+1} > \gamma C_{k+1}^{\frac{3}{2}}$. In fact, by (26) and the first equations of (22) and (23), we have $\Delta\hat{\mu}_k = -\gamma C_k^{\frac{3}{2}} + \gamma C_k^{\frac{3}{2}} = 0$. It follows that

$$\mu_{k+1} = \mu_k + \alpha_k \Delta\bar{\mu}_k + \alpha_k^2 \Delta\hat{\mu}_k = (1 - \alpha_k)\mu_k + \alpha_k \gamma C_k^{\frac{3}{2}}, \tag{29}$$

which gives $\mu_{k+1} > 0$, i.e., $z^{k+1} \in \mathbf{R}_{++} \times \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m$. Moreover, since $f(z^k) > 0$, by (24) we have $C_k > f(z^{k+1})$. This together with (25) and $f(z^{k+1}) > 0$ gives

$$C_{k+1} = \frac{(C_k + 1) f(z^{k+1})}{f(z^{k+1}) + 1} > f(z^{k+1}).$$

Furthermore, by $C_k > f(z^{k+1})$ and (25), we have

$$C_{k+1} = \frac{C_k f(z^{k+1}) + f(z^{k+1})}{f(z^{k+1}) + 1} < \frac{C_k f(z^{k+1}) + C_k}{f(z^{k+1}) + 1} = C_k. \tag{30}$$

By (29), (30) and $\mu_k > \gamma C_k^{\frac{3}{2}}$, we have

$$\mu_{k+1} \geq (1 - \alpha_k)\gamma C_k^{\frac{3}{2}} + \alpha_k \gamma C_k^{\frac{3}{2}} = \gamma C_k^{\frac{3}{2}} > \gamma C_{k+1}^{\frac{3}{2}}.$$

Therefore, we can conclude that if $z^k \in \mathbf{R}_{++} \times \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m$, $C_k > f(z^k)$ and $\mu_k > \gamma C_k^{\frac{3}{2}}$ for some k, then z^{k+1} generated by Algorithm ASNM satisfies $z^{k+1} \in \mathbf{R}_{++} \times \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m$, $C_{k+1} > f(z^{k+1})$ and $\mu_{k+1} > \gamma C_{k+1}^{\frac{3}{2}}$. Since $z^0 \in \mathbf{R}_{++} \times \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m$, $C_0 > f(z^0)$ and $\mu_0 > \gamma C_0^{\frac{3}{2}}$, by induction on k, we prove the theorem. □

4 Global Convergence

In this section, we discuss global convergence of ASNM. For this purpose, we first give the following lemmas.

Lemma 1 *Let Assumption* 1 *hold and* $\{z^k = (\mu_k, x^k, s^k, y^k)\}$ *be the iteration sequence generated by* ASNM. *Then,* $C_k > C_{k+1}$ *and* $\mu_k > \mu_{k+1}$ *for all* $k \geq 0$.

Proof The first result holds by (30). By (iii) of Theorem 2, we have $\mu_k > \gamma C_k^{\frac{3}{2}}$ for all $k \geq 0$. So, we obtain from (29) that $\mu_{k+1} < (1-\alpha_k)\mu_k + \alpha_k \mu_k = \mu_k$ for all $k \geq 0$. □

Lemma 2 *Let Assumption* 1 *hold and* $\{z^k = (\mu_k, x^k, s^k, y^k)\}$ *be the iteration sequence generated by* ASNM. *Then, there exists a constant* $C^* > 0$ *such that*

$$\lim_{k\to\infty} f(z^k) = \lim_{k\to\infty} C_k = C^*. \tag{31}$$

Moreover, we have

$$\lim_{k\to\infty} \alpha_k f(z^k) = 0. \tag{32}$$

Proof By Lemma 1, $\{C_k\}$ is strictly monotonically decreasing. Hence, there exists a constant $C^* \geq 0$ such that $\lim_{k\to\infty} C_k = C^*$. Moreover, by (25) we have

$$\lim_{k\to\infty} f(z^k) = \lim_{k\to\infty} \left(\frac{C_k}{1 + C_{k-1} - C_k} \right) = C^*.$$

This proves (31). By Step 5 and Step 6, we have

$$\tau(\alpha_k f(z^k))^2 \leq C_k - f(z^{k+1}),$$

which together with (31) proves (32). ◻

Theorem 3 *Let Assumption* 1 *hold and* $\{z^k = (\mu_k, x^k, s^k, y^k)\}$ *be the iteration sequence generated by* ASNM. *Then, any accumulation point of* $\{z^k\}$ *is a solution of* $\mathrm{H}(z) = 0$.

Proof By Lemma 2, there exists a constant $C^* \geq 0$ such that $\lim_{k\to\infty} f(z^k) = \lim_{k\to\infty} C_k = C^*$. It follows from (18) that

$$\lim_{k\to\infty} \|\mathrm{H}(z^k)\| = \sqrt{2C^*}. \tag{33}$$

Let $z^* = (\mu^*, x^*, s^*, y^*)$ be an accumulation point of $\{z^k = (\mu_k, x^k, s^k, y^k)\}$, and without loss of generality, we assume $\lim_{(\mathcal{K}\ni)k\to\infty} z^k = z^*$ where $\mathcal{K}$ is a subset of $\{0, 1, ...\}$. Then, by the continuity of H and (33), we have

$$\lim_{(\mathcal{K}\ni)k\to\infty} \|\mathrm{H}(z^k)\| = \|\mathrm{H}(z^*)\| = \sqrt{2C^*}. \tag{34}$$

Now we assume $\|\mathrm{H}(z^*)\| > 0$, i.e., $C^* > 0$ and will derive a contradiction. Since $\mu_k > \gamma C_k^{\frac{3}{2}}$ for all $k \geq 0$, we have $\mu^* \geq \gamma(C^*)^{\frac{3}{2}} > 0$ which means $z^* \in \mathbf{R}_{++} \times \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m$. Then, $\mathrm{H}(z)$ is continuously differentiable at z^* and $\mathrm{H}'(z^*)$ is nonsingular by Corollary 1. So, from (19), we have

$$\lim_{(\mathcal{K}\ni)k\to\infty} \Delta\bar{z}^k = \mathrm{H}'(z^*)^{-1}[-\mathrm{H}(z^*) + \gamma(C^*)^{\frac{3}{2}}h] =: \Delta\bar{z}^*.$$

Denote $\Delta\bar{z}^* = (\Delta\bar{\mu}^*, \Delta\bar{x}^*, \Delta\bar{s}^*, \Delta\bar{y}^*)$. Then, by (26) we have

$$\lim_{(\mathcal{K}\ni)k\to\infty} \hat{z}^k = (\gamma(C^*)^{\frac{3}{2}}, x^* + \Delta\bar{x}^*, s^* + \Delta\bar{s}^*, y^* + \Delta\bar{y}^*) =: \hat{z}^*.$$

Obviously, $\hat{z}^* \in \mathbf{R}_{++} \times \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m$. Since ψ is continuously differentiable at any $(\mu, x, s) \in \mathbf{R}_{++} \times \mathbf{V} \times \mathbf{V}$, by the continuity, we have

$$\lim_{(\mathcal{K}\ni)k\to\infty} \mathrm{J_H}(z^k, \hat{z}^k) = \mathrm{J_H}(z^*, \hat{z}^*).$$

Since $z^*, \hat{z}^* \in \mathbf{R}_{++} \times \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m$, $\mathrm{H}'(z^*)$ and $\mathrm{J_H}(z^*, \hat{z}^*)$ are all nonsingular. So, there exists a constant $M > 0$ such that $\|\mathrm{H}'(z^k)^{-1}\| \leq M$ and $\|\mathrm{J_H}(z^k, \hat{z}^k)^{-1}\| \leq M$ for all $k \in \mathcal{K}$. Hence, for any $k \in \mathcal{K}$, if the condition (20) holds, then $\Delta\hat{z}^k = 0$. Otherwise, we have

$$\|\mathrm{H}(\hat{z}^k)\| \leq \lambda\|\mathrm{H}(z^k)\| = \lambda\sqrt{2f(z^k)} < \lambda\sqrt{2C_k} < \lambda\sqrt{2C_0},$$

which together with (22) and (23) yields

$$\|\Delta\hat{z}^k\| \leq M(\|\mathrm{H}(\hat{z}^k)\| + \gamma C_k^{\frac{3}{2}}) < M(\lambda\sqrt{2C_0} + \gamma C_0^{\frac{3}{2}}).$$

This indicates that $\{\Delta\hat{z}^k\}_{k\in\mathcal{K}}$ is bounded. Since $\lim_{(\mathcal{K}\ni)k\to\infty} f(z^k) = C^* > 0$, by (32) we have $\lim_{(\mathcal{K}\ni)k\to\infty} \alpha_k = 0$. Let $\tilde{\alpha}_k = \delta^{-1}\alpha_k$. Then,

$$\lim_{(\mathcal{K}\ni)k\to\infty} \tilde{\alpha}_k = 0, \quad \lim_{(\mathcal{K}\ni)k\to\infty} (\Delta\bar{z}^k + \tilde{\alpha}_k\Delta\hat{z}^k) = \Delta\bar{z}^*. \tag{35}$$

Moreover, by Step 5, for all sufficiently large $k \in \mathcal{K}$,

$$f(z^k + \tilde{\alpha}_k\Delta\bar{z}^k + \tilde{\alpha}_k^2\Delta\hat{z}^k) > C_k - \tau(\tilde{\alpha}_k f(z^k))^2,$$

which together with $C_k > f(z^k)$ gives

$$\frac{f(z^k + \tilde{\alpha}_k(\Delta\bar{z}^k + \tilde{\alpha}_k\Delta\hat{z}^k)) - f(z^k)}{\tilde{\alpha}_k} > -\tau\tilde{\alpha}_k f(z^k)^2. \tag{36}$$

Since f is continuously differentiable at $z^* \in \mathbf{R}_{++} \times \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m$, by letting $k \to \infty$ with $k \in \mathcal{K}$ in (36), we have from (35) that

$$\nabla f(z^*)^T \Delta\bar{z}^* \geq 0. \tag{37}$$

On the other hand, by (19) and (33), we have

$$\begin{aligned} \nabla f(z^*)^T \Delta\bar{z}^* &= \mathrm{H}(z^*)^T \mathrm{H}'(z^*)\Delta\bar{z}^* \\ &= -\|\mathrm{H}(z^*)\|^2 + \gamma\mu^*(C^*)^{\frac{3}{2}} \\ &= -(2 - \gamma\mu^*(C^*)^{\frac{1}{2}})C^*. \end{aligned} \tag{38}$$

Since the sequences $\{\mu_k\}$ and $\{C_k\}$ are all strictly monotonically decreasing, we have $\gamma\mu^*(C^*)^{\frac{1}{2}} < \gamma\mu_0 C_0^{\frac{1}{2}} < 2$. Thus, by (38) and $C^* > 0$, we have

$$\nabla f(z^*)^T \Delta\bar{z}^* < 0,$$

which is contrary to (37). We complete the proof. □

5 Local Cubic Convergence

In this section, we establish the local cubic convergence properties of ASNM by employing the strong semismoothness of the function H. The readers can refer to [30] for the definition of (strong) semismoothness. In addition, note that H is (strongly) semismooth if and only if its each component function is (strongly) semismooth, see [27, Corollary 2.4].

Lemma 3 *If* H *is semismooth at* z^* *and all* $V \in \partial \mathrm{H}(z^*)$ *are nonsingular where* $\partial \mathrm{H}(z^*)$ *is the Clarke's generalized Jacobian of* H *at* z^**, then there exist* $\epsilon >$ *and* $\xi > 0$ *such that*

$$\|z - z^*\| \leq \xi \|\mathrm{H}(z)\|, \ \ \forall\, z \in N(z^*, \epsilon) := \{z| \ \|z - z^*\| \leq \epsilon\}.$$

Proof The proof of the lemma is similar as that of [34, Lemma 7]. □

Theorem 4 *Let Assumption* 1 *hold and* z^* *be any accumulation point of the iteration sequence* $\{z^k\}$ *generated by* ASNM. *Suppose that* F' *is locally Lipschitz continuous and* H *is strongly semismooth at* z^*. *If all* $V \in \partial \mathrm{H}(z^*)$ *are nonsingular, then the whole sequence* $\{z^k\}$ *converges to* z^* *and for all* z^k *sufficiently close to* z^**,*

$$z^{k+1} = z^k + \Delta \bar{z}^k + \Delta \hat{z}^k, \tag{39}$$

and

$$\|z^{k+1} - z^*\| = O(\|z^k - z^*\|^3), \tag{40}$$

and

$$\|\mathrm{H}(z^{k+1})\| = O(\|\mathrm{H}(z^k)\|^3). \tag{41}$$

Proof By Theorem 3, we have $\mathrm{H}(z^*) = 0$. Since all $V \in \partial \mathrm{H}(z^*)$ are nonsingular, by [27, Proposition 3.1], there is a constant $M_1 > 0$ such that for all z^k sufficiently close to z^*,

$$\|\mathrm{H}'(z^k)^{-1}\| \leq M_1. \tag{42}$$

Since H is strongly semismooth at z^*, for all z^k sufficiently close to z^*,

$$\|\mathrm{H}(z^k) - \mathrm{H}(z^*) - \mathrm{H}'(z^k)(z^k - z^*)\| = O(\|z^k - z^*\|^2). \tag{43}$$

This together with (19) and (42) implies that for all z^k sufficiently close to z^*,

$$\begin{aligned}
\|\hat{z}^k - z^*\| &= \|z^k + \Delta \bar{z}^k - z^*\| \\
&= \|z^k + \mathrm{H}'(z^k)^{-1}[-\mathrm{H}(z^k) + \gamma C_k^{\frac{3}{2}} h] - z^*\| \\
&\leq \|\mathrm{H}'(z^k)^{-1}\| \Big[\|\mathrm{H}(z^k) - \mathrm{H}(z^*) - \mathrm{H}'(z^k)(z^k - z^*)\| + \gamma C_k^{\frac{3}{2}} \Big] \\
&= O(\|z^k - z^*\|^2) + O(C_k^{\frac{3}{2}}).
\end{aligned} \tag{44}$$

Since H is strongly semismooth at z^*, H is locally Lipschitz continuous near z^*. Thus, for all z sufficiently close to z^*,

$$\|\mathrm{H}(z)\| = \|\mathrm{H}(z) - \mathrm{H}(z^*)\| = O(\|z - z^*\|). \tag{45}$$

For all $k \geq 1$, since $C_k < C_0$, by (18) and (25) we have

$$C_k = \frac{(C_{k-1}+1)f(z^k)}{f(z^k)+1} \leq \frac{C_0+1}{2}\|\mathrm{H}(z^k)\|^2,$$

which together with (45) gives for all z^k sufficiently close to z^*,

$$C_k^{\frac{3}{2}} = O(\|z^k - z^*\|^3). \tag{46}$$

By (44) and (46), for all z^k sufficiently close to z^*,

$$\|\hat{z}^k - z^*\| = O(\|z^k - z^*\|^2). \tag{47}$$

Thus, we have that $\hat{z}^k$ is sufficiently close to z^* when z^k is sufficiently close to z^*. Hence, by (45), (47) and Lemma 3, for all z^k sufficiently close to z^*,

$$\|\mathrm{H}(\hat{z}^k)\| = O(\|\hat{z}^k - z^*\|) = O(\|z^k - z^*\|^2) = O(\|\mathrm{H}(z^k)\|^2). \tag{48}$$

It follows that for all z^k sufficiently close to z^*,

$$\|\mathrm{H}(\hat{z}^k)\| \leq \lambda\|\mathrm{H}(z^k)\| = \lambda \min\{1, \|\mathrm{H}(z^k)\|\},$$

for any $\lambda > 0$. Hence, by Step 3 of ASNM, when z^k is sufficiently close to z^*, $\Delta\hat{z}^k$ is always computed in Step 4 of ASNM. Now, let us denote

$$\mathcal{J}_\mathrm{H}(z^*) := \mathrm{conv}\left\{ V \,|\, V = \lim_{\substack{z \to z^* \\ \hat{z} \to z^*}} \mathrm{J}_\mathrm{H}(z, \hat{z}),\ z, \hat{z} \in D_\mathrm{H} \right\},$$

where D_H is the set of points at which H is differentiable. Since F is continuously differentiable at z^*, we have $\partial \mathrm{H}(z^*) = \mathcal{J}_\mathrm{H}(z^*)$. So, by our assumption, all $V \in \mathcal{J}_\mathrm{H}(z^*)$ are nonsingular. Thus, there exists a constant $M_2 > 0$ such that for all z^k sufficiently close to z^*,

$$\|\mathrm{J}_\mathrm{H}(z^k, \hat{z}^k)^{-1}\| \leq M_2. \tag{49}$$

Thus, by (22), (23), (42), (46), (48) and (49), for all z^k sufficiently close to z^*, we have

$$\|\Delta\hat{z}^k\| \leq \max\{M_1, M_2\}(\|\mathrm{H}(\hat{z}^k)\| + \gamma C_k^{\frac{3}{2}}) = O(\|z^k - z^*\|^2). \tag{50}$$

So, by (47) and (50), for all z^k sufficiently close to z^*,

$$\begin{aligned} &\|z^k + \Delta\bar{z}^k + \Delta\hat{z}^k - z^*\| \\ &= \|\hat{z}^k - z^* + \Delta\hat{z}^k\| \leq \|\hat{z}^k - z^*\| + \|\Delta\hat{z}^k\| = O(\|z^k - z^*\|^2). \end{aligned} \tag{51}$$

This implies that $z^k + \Delta\bar{z}^k + \Delta\hat{z}^k$ is sufficiently close to z^* when z^k is sufficiently close to z^*. Hence, by (18), (45), (51) and Lemma 3, for all z^k sufficiently close to z^*,

$$\begin{aligned} f(z^k + \Delta\bar{z}^k + \Delta\hat{z}^k) &= \frac{1}{2}\|\mathrm{H}(z^k + \Delta\bar{z}^k + \Delta\hat{z}^k)\|^2 \\ &= O(\|z^k + \Delta\bar{z}^k + \Delta\hat{z}^k - z^*\|^2) = O(\|z^k - z^*\|^4) \\ &= O(\|\mathrm{H}(z^k)\|^4) = O(f(z^k)^2). \end{aligned} \tag{52}$$

So, for all z^k sufficiently close to z^*,

$$f(z^k + \Delta\bar{z}^k + \Delta\hat{z}^k) + \tau f(z^k)^2 \le f(z^k) < C_k,$$

where $\tau \in (0, 1)$ is the parameter in line search (24) of ASNM. This shows that, for all z^k sufficiently close to z^*, unit step size $\alpha_k = 1$ will be always accepted in Step 5 of ASNM. Hence, for all z^k sufficiently close to z^*, we have

$$z^{k+1} = z^k + \Delta\bar{z}^k + \Delta\hat{z}^k,$$

i.e., (39) holds. This together with (51) also shows $\lim_{k\to\infty} z^k = z^*$.

In the following, we show the local cubic convergence rate of z^k to z^*. By (47), we have $\lim_{k\to\infty} \hat{z}^k = z^*$. Since H is strongly semismooth at z^*, for all sufficiently large k,

$$\|\mathrm{H}(\hat{z}^k) - \mathrm{H}(z^*) - \mathrm{H}'(\hat{z}^k)(\hat{z}^k - z^*)\| = O(\|\hat{z}^k - z^*\|^2). \tag{53}$$

In addition, by (19), (42), (45) and (46), for all sufficiently large k,

$$\begin{aligned} \|\Delta\bar{z}^k\| &= \|\mathrm{H}'(z^k)^{-1}[-\mathrm{H}(z^k) + \gamma C_k^{\frac{3}{2}} h]\| \\ &\le M_1[\|\mathrm{H}(z^k)\| + \gamma C_k^{\frac{3}{2}}] \\ &= O(\|z^k - z^*\|). \end{aligned} \tag{54}$$

By (10), (11), (54) and the Lipschitz continuity of F', there exists a constant $\bar{L} > 0$ such that for all sufficiently large k,

$$\begin{aligned} \|\mathrm{H}'(\hat{z}^k) - \mathrm{J_H}(z^k, \hat{z}^k)\| &= \|F'(\hat{x}^k, \hat{s}^k, \hat{y}^k) - F'(x^k, s^k, y^k)\| \\ &\le \bar{L}\|(\hat{x}^k, \hat{s}^k, \hat{y}^k) - (x^k, s^k, y^k)\| \\ &\le \bar{L}\|\hat{z}^k - z^k\| = \bar{L}\|\Delta\bar{z}^k\| = O(\|z^k - z^*\|). \end{aligned} \tag{55}$$

We now consider how $\Delta\hat{z}^k$ is calculated according to the condition (21) in Step 4 of ASNM. If the condition (21) holds, then

$$\begin{aligned} \|\mathrm{J_H}(z^k, \hat{z}^k) - \mathrm{H}'(z^k)\| &= \|\psi'(\hat{\mu}^k, \hat{x}^k, \hat{s}^k) - \psi'(\mu^k, x^k, s^k)\| \\ &\le L\|(\hat{\mu}^k, \hat{x}^k, \hat{s}^k) - (\mu^k, x^k, s^k)\| \\ &\le L\|\hat{z}^k - z^k\| = L\|\Delta\bar{z}^k\| = O(\|z^k - z^*\|), \end{aligned}$$

which together (22), (39), (46), (47), (53) and (55) gives

$$
\begin{aligned}
\|z^{k+1}-z^*\| &= \|z^k+\Delta\bar{z}^k+\Delta\hat{z}^k-z^*\| \\
&= \|\hat{z}^k+\Delta\hat{z}^k-z^*\| = \|\hat{z}^k+\mathrm{H}'(z^k)^{-1}[-\mathrm{H}(\hat{z}^k)+\gamma C_k^{\frac{3}{2}}h]-z^*\| \\
&\le \|\mathrm{H}'(z^k)^{-1}\|\Big[\|\mathrm{H}(\hat{z}^k)-\mathrm{H}'(z^k)(\hat{z}^k-z^*)\|+\gamma C_k^{\frac{3}{2}}\Big] \\
&\le \|\mathrm{H}'(z^k)^{-1}\|\Big[\|\mathrm{H}(\hat{z}^k)-\mathrm{H}(z^*)-\mathrm{H}'(\hat{z}^k)(\hat{z}^k-z^*)\| \\
&\quad +\|(\mathrm{H}'(\hat{z}^k)-\mathrm{J}_{\mathrm{H}}(z^k,\hat{z}^k))(\hat{z}^k-z^*)\| \\
&\quad +\|(\mathrm{J}_{\mathrm{H}}(z^k,\hat{z}^k)-\mathrm{H}'(z^k))(\hat{z}^k-z^*)\|+\gamma C_k^{\frac{3}{2}}\Big] \\
&= O(\|\hat{z}^k-z^*\|^2)+O(\|z^k-z^*\|\|\hat{z}^k-z^*\|)+O(\|z^k-z^*\|^3) \\
&= O(\|z^k-z^*\|^3);
\end{aligned}
$$

otherwise, by (23), (39), (46), (47), (53) and (55) we have

$$
\begin{aligned}
\|z^{k+1}-z^*\| &= \|z^k+\Delta\bar{z}^k+\Delta\hat{z}^k-z^*\| \\
&= \|\hat{z}^k+\Delta\hat{z}^k-z^*\| = \|\hat{z}^k+\mathrm{J}_{\mathrm{H}}(z^k,\hat{z}^k)^{-1}[-\mathrm{H}(\hat{z}^k)+\gamma C_k^{\frac{3}{2}}h]-z^*\| \\
&\le \|\mathrm{J}_{\mathrm{H}}(z^k,\hat{z}^k)^{-1}\|\Big[\|\mathrm{H}(\hat{z}^k)-\mathrm{J}_{\mathrm{H}}(z^k,\hat{z}^k)(\hat{z}^k-z^*)\|+\gamma C_k^{\frac{3}{2}}\Big] \\
&\le \|\mathrm{J}_{\mathrm{H}}(z^k,\hat{z}^k)^{-1}\|\Big[\|\mathrm{H}(\hat{z}^k)-\mathrm{H}(z^*)-\mathrm{H}'(\hat{z}^k)(\hat{z}^k-z^*)\| \\
&\quad +\|(\mathrm{H}'(\hat{z}^k)-\mathrm{J}_{\mathrm{H}}(z^k,\hat{z}^k))(\hat{z}^k-z^*)\|+\gamma C_k^{\frac{3}{2}}\Big] \\
&= O(\|\hat{z}^k-z^*\|^2)+O(\|z^k-z^*\|\|\hat{z}^k-z^*\|)+O(\|z^k-z^*\|^3) \\
&= O(\|z^k-z^*\|^3).
\end{aligned}
$$

This proves (40). Moreover, by (40), (45) and Lemma 3, for all sufficiently large k,

$$
\|\mathrm{H}(z^{k+1})\| = O(\|z^{k+1}-z^*\|) = O(\|z^k-z^*\|^3) = O(\|\mathrm{H}(z^k)\|^3).
$$

We complete the proof. □

To establish the local cubic convergence of ASNM, we need to assume that F' is locally Lipschitz continuous and H is strongly semismooth at z^*. It is worth pointing out that these assumptions are essential but by no means restrictive. For example, for wLCP (1), $F(x,s,y) = Px+Qs+Ry-a$ is linear and obviously Lipschitz continuously differentiable. In addition, the smoothing function ψ for wLCP becomes

$$
\psi_c(\mu,a,b) = a+b-\sqrt{(a-b)^2+4c+4\mu^2},\ \forall\,(\mu,a,b)\in\mathbf{R}^3, \tag{56}
$$

where $c \geq 0$ is a fixed constant. Then, the function H defined in (8) becomes

$$\mathrm{H}(z) = \begin{pmatrix} \mu \\ Px + Qs + Ry - a \\ \mathrm{fl}(\mu, x, s) \end{pmatrix}, \text{ where } \mathrm{fl}(\mu, x, s) = \begin{pmatrix} \psi_{w_1}(\mu, x_1, s_1) \\ \vdots \\ \psi_{w_n}(\mu, x_n, s_n) \end{pmatrix} \quad (57)$$

and $w = (w_1, ..., w_n)^T$ is the weight vector in wLCP. It can be verified that the function ψ_c given in (56) is strongly semismooth on $\mathbf{R}^3$. So, H defined by (57) is strongly semismooth since its each component function is strongly semismooth. As another example, consider the KKT optimality conditions of liner programming over symmetric cones $\mathbf{K}$ (e.g., [15, 18]), which is defined by

$$x \in \mathbf{K},\ s \in \mathbf{K},\ Ax = b,\ A^* y + s = c,\ x \circ s = 0, \quad (58)$$

where A is a linear operator and A^* is its ajoint. Note that the system (58) is equivalent to the CP (5) with

$$F(x, s, y) = \begin{pmatrix} Ax - b \\ A^* y + s - c \end{pmatrix}.$$

Hence, F is again Lipschitz continuously differentiable. For CP (5), the smoothing function ψ given in (6) is

$$\psi(\mu, x, s) = x + s - \sqrt{(x - s)^2 + 4\mu^2 e},\ \ \forall\ (\mu, x, s) \in \mathbf{R}_+ \times \mathbf{V} \times \mathbf{V},$$

which is strongly semismooth at any $(0, x, s) \in \mathbf{R} \times \mathbf{V} \times \mathbf{V}$ by [30, Proposition 3.4]. Hence, its associated function H is also strongly semismooth at any $z = (0, x, s, y) \in \mathbf{R} \times \mathbf{V} \times \mathbf{V} \times \mathbf{R}^m$.

6 Discussions on Solving the (Approximate) Newton Equations

In this section, we discuss an efficient practical way to solve the (approximate) Newton equations (19), (22) and (23) by taking wLCP (1) as an example. The general wCP (4) can be treated by the same way. For wLCP, the Jacobian of $\mathrm{H}(z)$ given in (57) at any $z \in \mathbf{R}_{++} \times \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^m$ is

$$\mathrm{H}'(z) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & P & Q & R \\ \mathrm{fl}'_\mu(\mu, x, s) & \mathrm{fl}'_x(\mu, x, s) & \mathrm{fl}'_s(\mu, x, s) & 0 \end{bmatrix}, \quad (59)$$

in which

$$\mathrm{fl}'_\mu(\mu, x, s) = \left(-\frac{4\mu}{\sqrt{(x_1 - s_1)^2 + 4w_1 + 4\mu^2}}, \ldots, -\frac{4\mu}{\sqrt{(x_n - s_n)^2 + 4w_n + 4\mu^2}} \right)^T,$$
$$\mathrm{fl}'_x(\mu, x, s) = I - D(\mu, x, s) \quad \text{and} \quad \mathrm{fl}'_s(\mu, x, s) = I + D(\mu, x, s),$$

where

$$D(\mu, x, s) = \mathbf{diag}\left(\frac{x_i - s_i}{\sqrt{(x_i - s_i)^2 + 4w_i + 4\mu^2}} \right).$$

For any $k \geq 0$, let $D_k = D(\mu^k, x^k, s^k)$. Then, by the structure of $\mathrm{H}(z)$ and $\mathrm{H}'(z)$ in (57) and (59), solving the Newton equation (19) is equivalent to setting $\Delta\bar{\mu}_k = -\mu_k + \gamma C_k^{\frac{3}{2}}$ and computing $(\Delta\bar{x}^k, \Delta\bar{s}^k, \Delta\bar{y}^k)$ satisfying

$$\begin{aligned} P\Delta\bar{x}^k + Q\Delta\bar{s}^k + R\Delta\bar{y}^k &= r_1^k \\ (I - D_k)\Delta\bar{x}^k + (I + D_k)\Delta\bar{s}^k &= r_2^k, \end{aligned} \tag{60}$$

where $r_1^k = -(Px^k + Qs^k + Ry^k - a)$ and $r_2^k = -\mathrm{fl}(\mu_k, x^k, s^k) - \mathrm{fl}'_\mu(\mu_k, x^k, s^k)\Delta\bar{\mu}_k$. The linear system (60) can be further simplified by changing of variables. Let $u_k = \Delta\bar{x}^k + \Delta\bar{s}^k$ and $v_k = \Delta\bar{x}^k - \Delta\bar{s}^k$. Then, the equations in (60) become

$$\begin{aligned} &\frac{P+Q}{2}u_k + \frac{P-Q}{2}v_k + R\Delta\bar{y}^k = r_1^k \\ &u_k = D_k v_k + r_2^k, \end{aligned}$$

for which, by denoting $\hat{P} = \frac{P+Q}{2}$ and $\hat{Q} = \frac{P-Q}{2}$, $(v_k, \Delta\bar{y}^k)$ can be solved by

$$(\hat{P}D_k + \hat{Q})v_k + R\Delta\bar{y}^k = r_1^k - \hat{P}r_2^k. \tag{61}$$

Then, u_k and $(\Delta\bar{x}^k, \Delta\bar{s}^k)$ can be finally obtained by back substitutions. Note that the coefficient matrices of linear systems (60) and (61) are

$$\begin{bmatrix} P & Q & R \\ I - D_k & I + D_k & 0 \end{bmatrix}_{(2n+m)\times(2n+m)} \quad \text{and} \quad \begin{bmatrix} \hat{P}D_k + \hat{Q} & R \end{bmatrix}_{(n+m)\times(n+m)}$$

respectively. Hence, the linear system (61) is in a much smaller compact format than the original linear system (60), which could significantly reduce the computational cost when solving large-scale wLCPs.

By the same way, when solving the linear system (22), we can simply set $\Delta\hat{\mu}_k = 0$ and solve

$$(\hat{P}D_k + \hat{Q})\hat{v}_k + R\Delta\hat{y}^k = \hat{r}_1^k - \hat{P}\hat{r}_2^k \tag{62}$$

to obtain $\hat{v}_k$ and $\Delta\hat{y}^k$, where $\hat{r}_1^k = -(P\hat{x}^k + Q\hat{s}^k + R\hat{y}^k - a)$ and $\hat{r}_2^k = -\mathrm{fl}(\hat{\mu}_k, \hat{x}^k, \hat{s}^k)$. Then, we can set $\hat{u}_k = D_k\hat{v}_k + \hat{r}_2^k$ and obtain $\Delta\hat{x}^k = \frac{\hat{u}_k+\hat{v}_k}{2}$ and $\Delta\hat{s}^k = \frac{\hat{u}_k-\hat{v}_k}{2}$. Moreover, noticing that the linear systems (61) and (62) have identical coefficient matrices and different right hand sides. Hence, significant computational cost can be saved when some direct methods are applied to solve the linear systems (61) and (62). For instance, the LU factorization of the coefficient matrix only needs to be computed once for solving these two linear systems.

Similarly, when solving the linear system (23), we set $\Delta\hat{\mu}_k = 0$ and solve

$$(\hat{P}\hat{D}_k + \hat{Q})\hat{v}_k + R\Delta\hat{y}^k = \hat{r}_1^k - \hat{P}\hat{r}_2^k \tag{63}$$

to obtain $\hat{v}_k$ and $\Delta\hat{y}^k$, where $\hat{D}_k = D(\hat{\mu}^k, \hat{x}^k, \hat{s}^k)$. Then, we can set $\hat{u}_k = \hat{D}_k\hat{v}_k + \hat{r}_2^k$ and obtain $\Delta\hat{x}^k = \frac{\hat{u}_k+\hat{v}_k}{2}$ and $\Delta\hat{s}^k = \frac{\hat{u}_k-\hat{v}_k}{2}$. In this case, note that there is only a low rank submatrix difference between the coefficient matrices of linear system (61) and (63). This property could be also exploited in practice to reduce the computational cost.

7 Numerical Results

We would like to perform some numerical experiments in this section. All experiments are carried on a PC with CPU of Inter(R) Core(TM)i7-7700 CPU @ 3.60 GHz and RAM of 8.00GB. The program codes are written in MATLAB and run in MATLAB R2018a environment. To show the efficiency of ASNM, we would compare it with the benchmark well-known Qi–Sun–Zhou smoothing Newton method [28], denoted by QSZ-SNM. In all the experiments, the parameters in ASNM are set as $\delta = 0.5,\ \mu_0 = 10^{-4},\ \tau = 10^{-7}, \lambda = 1, L = 10$. Moreover, we let $C_0 = f(z^0) + 1$ and take $\gamma = \frac{\mu_0}{C_0^{3/2}+1}$. We use $\min\{\|\mathrm{H}(z^k)\|,\ \|\mathrm{H}(\hat{z}^k)\|\} \leq 10^{-8}$ as the stopping criterion for ASNM. For QSZ-SNM, we take same parameters as those used in [28, Preliminary numerical results] and use $\|\mathrm{H}(z^k)\| \leq 10^{-8}$ as the stopping criterion.

7.1 wLCP Over Nonnegative Orthant

In this subsection, we consider to solve the following wLCP:

$$x \geq 0,\ s \geq 0,\ Px + Qs + Ry = a,\ xs = w, \tag{64}$$

where

$$P = \begin{pmatrix} A \\ M \end{pmatrix},\quad Q = \begin{pmatrix} 0 \\ -I \end{pmatrix},\quad R = \begin{pmatrix} 0 \\ -A^T \end{pmatrix},\quad a = \begin{pmatrix} b \\ -f \end{pmatrix}, \tag{65}$$

in which M is an $n \times n$ symmetric positive semidefinite matrix, $A \in \mathbf{R}^{m\times n}$ is a full row rank matrix with $m < n$ and $f \in \mathbf{R}^n$ and $b \in \mathbf{R}^m$ are given vectors. wLCP of the form (64) with (65) arises from the optimality conditions of the Quadratic Programming

Table 1 The value of $\|\mathrm{H}(z^k)\|$ at the k-th iteration for solving a wLCP

	ASNM	QSZ-SNM
$k = 1$	6.1371e+00	5.7391e+00
$k = 2$	2.6077e−01	6.6797e−01
$k = 3$	2.7683e−04	2.6777e−02
$k = 4$	1.2603e−12	2.6674e−03
$k = 5$	0	3.1728e−07
$k = 6$	0	6.9176e−13

Table 2 Comparison of **ASNM** and **QSZ-SNM** for solving wLCPs

n	ASNM			QSZ-SNM		
	AIT	ACPU	AHK	AIT	ACPU	AHK
1000	4.1	0.33	1.3237e−09	6.0	0.43	8.3745e−10
2000	4.2	1.70	7.0614e−10	7.0	2.55	1.7456e−12
3000	4.3	4.51	2.4453e−09	7.0	7.15	3.1857e−12
4000	4.3	9.77	3.5265e−09	7.0	15.34	4.8653e−12
5000	4.2	17.09	2.6718e−09	7.0	25.84	6.8262e−12
6000	4.3	27.85	2.7384e−09	7.0	44.82	8.9658e−12
7000	4.2	53.15	3.2102e−09	7.0	87.85	1.1353e−11
8000	4.3	108.95	4.0344e−09	7.0	229.14	1.3677e−11

and Weighted Centering problem [23, Theorem 2.1]. This wLCP is monotone [23] and has been used in numerical experiments for smoothing Newton methods [32, 33, 40]. In the experiments, we generate a random matrix $A \in \mathbf{R}^{m\times n}$ with full row rank and set $M = B^T B/\|B^T B\|$ with $B = \mathbf{rand}(n, n)$. Then, we choose $\hat{x} = \mathbf{rand}(n, 1)$, $f = \mathbf{rand}(n, 1)$ and set $b = A\hat{x}$, $\hat{s} = M\hat{x} + f$ and $w = \hat{x}\hat{s}$.

As a typical example of observing local convergence behaviors of ASNM and QSZ-SNM, we show the results of solving one test problem with $n = 500$ and $m = 250$, $x^0 = s^0 = (1, 0, ..., 0)^T$ and $y^0 = (0, ..., 0)^T$ as the starting point. Table 1 gives the value of $\|\mathrm{H}(z^k)\|$ at each iteration. We can clearly see that ASNM has local cubic convergence rate and converges faster than QSZ-SNM, which usually maintains local quadratic convergence rate.

Next, for each problem with different sizes n and $m = n/2$, we randomly generate 10 instances and test them by using the same starting points as before. Numerical results are provided in Table 2, where **AIT** and **ACPU** denote the average number of iterations and the average CPU time in seconds, respectively, and **AHK** denotes the final average value of $\|\mathrm{H}(z^k)\|$. From Table 2, we can see that ASNM always takes fewer iterations and often uses much less CPU time than QSZ-SNM to reach the stopping tolerance. Moreover, we could observe that the larger the problem size is the more CPU time ASNM can save compared with QSZ-SNM. This is because ASNM has local cubic convergence rate and the computational techniques discussed in Sect. 6 can significantly reduce the computational cost when the second approximate Newton system need to be solved.

7.2 wNCP Over Nonnegative Orthant

In this subsection, we consider to solve the following wNCP:

$$x \geq 0,\ s \geq 0,\ F(x, s, y) = 0,\ xs = w, \tag{66}$$

where

$$F(x, s, y) = \begin{pmatrix} Ax - b \\ \nabla f(x) - s - A^T y \end{pmatrix}, \tag{67}$$

in which $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$ and $f : \mathbf{R}^n \to \mathbf{R}$ is a twice continuously differentiable function. This wNCP is a perturbed Karush–Kuhn–Tucker (KKT) system of the nonlinear programming:

$$\min\ f(x),\quad \text{s.t. } Ax = b,\ x \geq 0.$$

Note that if $f(x)$ is linear, i.e., $f(x) = c^T x$, then the wNCP (66) reduces to the special wLCP (2).

In this experiment, we test the following nonlinear function:

$$f(x) = \frac{1}{2} x^T M x + (q + P(x))^T x,$$

where $M \in \mathbf{R}^{n \times n}$ is a positive semidefinite matrix, $q \in \mathbf{R}^n$ and $P(x) = (P_1(x), ..., P_n(x)) : \mathbf{R}^n \to \mathbf{R}^n$ is a twice continuously differentiable map. We randomly generate a matrix $A \in \mathbf{R}^{m \times n}$ with full row rank, choose $\hat{x} = \mathbf{rand}(n, 1)$ and set $b = A\hat{x}$. Moreover, we choose $M = \frac{n}{4} \frac{N^T N}{\|N^T N\|}$ with $N = \mathbf{randn}(n, n)$, $q = \mathbf{rand}(n, 1)$ and take $w = \mathbf{rand}(n, 1)$. The nonlinear function $P(x)$ is, respectively, set by

(a) $P_i(x) = d_i \cdot \arctan x_i,\ $ where $d_i = 4 * \mathbf{rand}(1, 1),\quad i = 1, ..., n;$

(b) $P_i(x) = x_i^2 + \sin x_i + \cos x_i + 1,\quad i = 1, ..., n;$

(c) $P_i(x) = \ln(x_i + 1) - \dfrac{x_i}{n},\quad i = 1, ..., n.$

Then, for each problem with different sizes n and $m = n/2$, we generate 10 instances and test them by using the starting points $x^0 = s^0 = (1, 0, ..., 0)^T$ and $y^0 = (0, ..., 0)^T$. Numerical results are given in Table 3, which again show that ASNM is much more efficient than QSZ-SNM in both the number of iterations and CPU time.

7.3 wCP Over the Second-order Cone

The *second-order cone* (SOC) in $\mathbf{R}^n (n \geq 1)$ defined by

$$\mathbf{L}^n := \{(x_1, x_{2:n}) \in \mathbf{R} \times \mathbf{R}^{n-1} : x_1 \geq \|x_{2:n}\|\}$$

Table 3 Comparison of **ASNM** and **QSZ-SNM** for solving wNCPs

$P(x)$	n	ASNM			QSZ-SNM		
		AIT	ACPU	AHK	AIT	ACPU	AHK
(a)	1000	10.0	0.84	2.2785e−09	13.6	1.11	2.9914e−09
	2000	11.0	4.65	1.9422e−09	15.3	5.93	2.9373e−09
	3000	12.2	12.66	1.9482e−09	18.9	19.33	2.7376e−09
	4000	12.2	26.41	2.7206e−09	19.2	42.26	2.5145e−09
	5000	12.5	46.91	1.1588e−09	20.6	76.93	2.1689e−09
	6000	13.0	79.72	5.4644e−10	21.9	138.53	2.3165e−09
(b)	1000	10.2	0.98	2.1380e−09	13.8	1.06	3.1705e−09
	2000	11.0	4.48	2.5691e−09	15.6	5.80	1.8689e−09
	3000	12.0	12.58	2.5474e−09	17.6	17.71	2.2413e−09
	4000	12.3	27.25	1.0696e−09	19.5	41.30	1.6055e−09
	5000	12.6	48.27	2.6266e−09	21.3	79.24	2.9997e−09
	6000	13.2	82.93	1.1333e−09	22.0	133.50	6.6072e−10
(c)	2000	11.2	10.89	2.0117e−09	15.3	15.02	2.4610e−09
	2500	11.4	20.25	3.1721e−09	16.4	28.27	3.7709e−09
	3000	11.8	33.89	3.4694e−09	17.3	48.79	3.0284e−09
	3500	12.1	51.29	1.7574e−09	17.8	76.55	2.0459e−09
	4500	12.5	102.42	2.0450e−09	19.5	164.40	1.5862e−09
	5000	12.7	141.12	7.2064e−10	20.5	235.72	2.0797e−09

is also called the Lorentz cone, where $\|\cdot\|$ denotes the Euclidean norm. In recent years, optimization problems with second-order cone constraints have received considerable attention because of its wide range of applications in many fields. One may see the survey paper [2] and references therein.

We now consider to solve the wCP (66) over the SOC $\mathbf{L}^n$ (wCP-SOC), which finds $(x, s, y) \in \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^m$ such that

$$x \in \mathbf{L}^n,\ \ s \in \mathbf{L}^n,\ \ F(x, s, y) = 0,\ \ x \circ s = w, \tag{68}$$

with

$$F(x, s, y) = \begin{pmatrix} Ax - b \\ Mx + q - s - A^T y \end{pmatrix}, \tag{69}$$

where the Jordan product " $\circ$ " is defined by $x \circ s = (x^T s, x_1 s_{2:n} + s_1 x_{2:n})$, $w \in \mathbf{L}^n$, $A \in \mathbf{R}^{m\times n}$, $b \in \mathbf{R}^m$ and $q \in \mathbf{R}^n$ are given vectors and $M \in \mathbf{R}^{n\times n}$ is a given positive semidefinite matrix. The system (68)–(69) often appears as a perturbed KKT system of the convex quadratic optimization over the SOC $\mathbf{L}^n$ [38]:

$$\min\ \frac{1}{2}x^T Mx + q^T x,\ \ \text{s.t. } Ax = b,\ \ x \in \mathbf{L}^n.$$

Table 4 Comparison of **ASNM** and **QSZ-SNM** for solving wCP-SOC

SP	n	ASNM			QSZ-SNM		
		AIT	ACPU	AHK	AIT	ACPU	AHK
(SP1)	1000	8.0	1.45	1.9697e−12	9.4	1.47	1.7276e−09
	2000	8.0	7.98	1.2333e−09	10.6	9.43	1.5471e−09
	3000	9.0	25.88	4.6535e−10	13.0	34.53	2.3388e−09
	4000	9.0	60.40	9.6934e−10	14.1	86.46	2.5829e−09
	5000	9.0	115.35	2.5780e−09	16.5	191.61	6.3682e−10
	6000	10.0	215.48	1.1055e−10	19.1	381.84	1.2667e−10
(SP2)	1000	7.8	1.42	1.2917e−09	9.2	1.45	1.6058e−09
	2000	8.1	7.99	1.4709e−09	10.8	9.65	2.2457e−09
	3000	9.0	25.79	1.4007e−09	13.2	34.30	2.2457e−09
	4000	9.0	61.48	5.8138e−11	14.2	88.10	1.1690e−09
	5000	9.3	119.58	8.4791e−10	16.9	199.83	1.0470e−09
	6000	10.0	209.91	1.1765e−10	19.2	379.28	1.3902e−10

In this experiment, we take $M = \frac{n}{4}\frac{N^T N}{\|N^T N\|}$ with $N = \mathbf{randn}(n, n)$ and $q = \mathbf{rand}(n, 1)$. Moreover, we take $w = (w_1, w_{2:n}^T)^T \in \mathbf{L}^n$ with $w_{2:n} = \mathbf{rand}(n-1, 1)$ and $w_1 = \|w_{2:n}\| + \mathbf{rand}(1, 1)$. In addition, we generate a random matrix $A \in \mathbf{R}^{m\times n}$ with full row rank and set $b = A\hat{x}$ with $\hat{x} \in \mathbf{L}^n$ being generated by the same way as w. We again generate 10 instances with different sizes n and $m = n/2$, and solve them by using the following two starting points, respectively:

(SP1) $x^0 = s^0 = (1, 0, ..., 0)^T$, $y^0 = (0, ..., 0)^T$;

(SP2) $x^0 = \mathbf{rand}(n, 1)/n$, $s^0 = \mathbf{rand}(n, 1)/n$, $y^0 = \mathbf{rand}(m, 1)/m$,

where x^0 and s^0 in **(SP1)** are in strict interior of $\mathbf{L}^n$ while x^0 and s^0 in **(SP2)** in general do not belong to $\mathbf{L}^n$. Table 4 provides the numerical results of solving the 10 instances for each case. These numerical results again confirm the superior numerical performances of ASNM compared with QSZ-SNM. And we can see again that ASNM saves more CPU time as the problem dimension increases. Furthermore, the computational results show the flexibility of choosing the starting point of ASNM. Unlike interior-point methods, strictly interior starting point in $\mathbf{L}^n$ is not required by ASNM and the starting point will not affect the computational results significantly.

8 Conclusions

In this paper we propose an accelerated smoothing Newton method (ASNM) for solving the general wCP (4) by reformulating it as equivalent system of nonlinear equations. ASNM is designed to compute an additional approximate Newton step when the iterates are close to the solution set of wCP. A Lipschitz condition on the gradients of the smoothing function is used to determine which additional approximate Newton step will be computed to accelerate the convergence rate. Since previous Jacobian information is used, the cost of solving the additional approximate Newton system is much reduced at most iterations. A second-order nonmonotone line search technique is introduced in ASNM to ensure the global convergence. Under proper nonsingularity and strong semismooth assumptions, ASNM is shown to have local cubic convergence rate. Moreover, brief discussions are made on how to solve the resulted approximate Newton systems efficiently. Our numerical experiments show that ASNM significantly improves the computational efficiency compared with some well-known benchmark smoothing Newton method for solving various wCPs arising in mathematical programming.

References

1. Asadi, S., Darvay, Z., Lesaja, G., Mahdavi-Amiri, N., Potra, F.A.: A full-Newton step interior-point method for monotone weighted linear complementarity problems. J. Optim. Theory Appl. **186**, 864–878 (2020)
2. Chen, J.-S., Pan, S.H.: A survey on SOC complementarity functions and solution methods for SOCPs and SOCCPs. Pac. J. Optim. **8**, 33–74 (2012)
3. Chi, X.N., Wei, H.J., Wan, Z.P., Zhu, Z.B.: Smoothing Newton algorithm for the circular cone programming with a nonmonotone line search. Acta Math. Scientia **37B**(5), 1262–1280 (2017)
4. Chi, X.N., Gowda, M.S., Tao, J.: The weighted horizontal linear complementarity problem on a Euclidean Jordan algebra. J. Glob. Optim. **73**, 153–169 (2019)
5. Chi, X.N., Wan, Z.P., Hao, Z.J.: A full-modified-Newton step $O(n)$ infeasible interior-point method for the special weighted linear complementarity problem. J. Ind. Manag. Optim. **18**(4), 2579–2598 (2022)
6. Chi, X.N., Wang, G.Q.: A full-Newton step infeasible interior-point method for the special weighted linear complementarity problem. J. Optim. Theory Appl. **190**, 108–129 (2021)
7. Engelke, S., Kanzow, C.: Improved smoothing-type methods for the solution of linear programming. Numer. Math. **90**, 487–507 (2002)
8. Faraut, J., Korányi, A.: Analysis on Symmetric Cones. Oxford Mathematical Monographs, Oxford University Press, New York (1994)
9. Gowda, M.S.: Weighted LCPs and interior point systems for copositive linear transformations on Euclidean Jordan algebras. J. Global Optim. **74**, 285–295 (2019)
10. Gowda, M.S., Sznajder, R., Tao, J.: Some P-properties for linear transformations on Euclidean Jordan algebras. Linear Algebra Appl. **393**, 203–232 (2004)
11. Huang, Z.H., Gu, W.Z.: A smoothing-type algorithm for solving linear complementarity problems with strong convergence properties. Appl. Math. Optim. **57**, 17–29 (2008)
12. Huang, Z.H., Zhang, Y., Wu, W.: A smoothing-type algorithm for solving system of inequalities. J. Comput. Appl. Math. **220**, 355–363 (2008)
13. Huang, Z.H., Ni, T.: Smoothing algorithms for complementarity problems over symmetric cones. Comput. Optim. Appl. **45**, 557–579 (2010)
14. Jiang, X., Zhang, Y.: A smoothing-type algorithm for absolute value equations. J. Ind. Manag. Optim. **9**, 789–798 (2013)
15. Kong, L.: Quadratic convergence of a smoothing Newton method for symmetric cone programming without strict complementarity. Positivity **16**, 297–319 (2012)
16. Liu, L.X., Liu, S.Y., Wang, C.F.: Smoothing Newton methods for symmetric cone linear complementarity problem with the Cartesia P/P_0-property. J. Ind. Manag. Optim. **7**(1), 53–66 (2011)
17. Liu, L.X., Liu, S.Y.: A new nonmonotone smoothing Newton method for the symmetric cone complementarity problem with the Cartesian P_0-property. Math. Meth. Oper. Res. **92**, 229–247 (2020)
18. Liu, X.H., Huang, Z.H.: A smoothing Newton algorithm based on a one-parametric class of smoothing functions for linear programming over symmetric cones. Math. Math. Oper. Res. **70**, 385–404 (2009)
19. Liu, X.H., Gu, W.Z.: Smoothing Newton algorithm based on a regularized one-parametric class of smoothing functions for generalized complementarity problems over symmetric cones. J. Ind. Manag. Optim. **6**(2), 363–380 (2010)
20. Lu, N., Huang, Z.H.: A smoothing Newton algorithm for a class of non-monotonic symmetric cone linear complementarity problems. J. Optim. Theory Appl. **161**, 446–464 (2014)

21. Narushima, Y., Sagara, N., Ogasawara, H.: A smoothing Newton method with Fischer-Burmeister function for second-order cone complementarity problems. J. Optim. Theory Appl. **149**, 79–101 (2011)
22. Ortega, J.M., Rheinboldt, W.C.: Iterative Solution of Nonlinear Equations in Several Variables. Academic Press, New York (1970)
23. Potra, F.A.: Weighted complementarity problems-a new paradigm for computing equilibria. SIAM J. Optim. **22**, 1634–1654 (2012)
24. Potra, F.A.: Sufficient weighted complementarity problems. Comput. Optim. Appl. **64**(2), 467–488 (2016)
25. Potra, F.A.: Equilibria and weighted complementarity problems. In: Al-Baali M., Grandinetti L., Purnama A. (eds) Numerical Analysis and Optimization. NAO 2017. Springer Proceedings in Mathematics and Statistics, vol 235. Springer, Cham. https://doi.org/10.1007/978-3-319-90026-1_-12 (2018)
26. Qi, H.D.: A regularized smoothing Newton method for box constrained variational inequality problems with P_0-functions. SIAM J. Optim. **10**, 315–330 (2000)
27. Qi, L., Sun, J.: A nonsmooth version of Newton's method. Math. Program. **58**, 353–367 (1993)
28. Qi, L., Sun, D., Zhou, G.: A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequalities. Math. Program. **87**, 1–35 (2000)
29. Smale, S.: Algorithms for solving equations. In: Gleason, A.M. (ed.) Proceeding of International Congress of Mathematicians, pp. 172–195. American Mathematics Society, Providence, Rhode Island (1987)
30. Sun, D., Sun, J.: Löwner's operator and spectral functions in Euclidean Jordan algebras. Math. Oper. Res. **33**, 421–445 (2008)
31. Tseng, P.: Error bounds and superlinear convergence analysis of some Newton-type methods in optimization. In: Di Pillo, G., Giannessi, F. (eds.) Nonlinear Optimization and Related Topics, pp. 445–462. Kluwer Academic Publishers, Boston (2000)
32. Tang, J.Y.: A variant nonmonotone smoothing algorithm with improved numerical results for large-scale LWCPs. Comput. Appl. Math. **37**, 3927–3936 (2018)
33. Tang, J.Y., Zhang, H.C.: A nonmonotone smoothing Newton algorithm for weighted complementarity problems. J. Optim. Theory Appl. **189**, 679–715 (2021)
34. Tang, J.Y., Zhou, J.C.: Smoothing inexact Newton method based on a new derivative-free nonmonotone line search for the NCP over circular cones. Ann. Oper. Res. **295**, 787–808 (2020)
35. Tang, J.Y., Zhou, J.C.: A modified damped Gauss-Newton method for non-monotone weighted linear complementarity problems. Optim. Methods Softw. **37**, 1145–1164 (2022)
36. Tang, J.Y., Zhou, J.C.: Quadratic convergence analysis of a nonmonotone Levenberg-Marquardt type method for the weighted nonlinear complementarity problem. Comput. Optim. Appl. **80**, 213–244 (2021)
37. Tang, J., Ma, C.: A smoothing Newton method for symmetric cone complementarity problems. Optim. Lett. **9**, 225–244 (2015)
38. Wang, G.Q., Zhang, Z.H., Zhu, D.T.: On extending primal-dual interior-point method for linear optimization to convex quadratic symmetric cone optimization. Numer. Funct. Anal. Optim. **34**(5), 576–603 (2012)
39. Yoshise, A.: Interior point trajectories and a homogeneous model for nonlinear complementarity problems over symmetric cones. SIAM J. Optim. **17**, 1129–1153 (2006)
40. Zhang, J.: A smoothing Newton algorithm for weighted linear complementarity problem. Optim. Lett. **10**, 499–509 (2016)
41. Zhu, J.G., Liu, H.W., Li, X.L.: A regularized smoothing-type algorithm for solving a system of inequalities with a P_0-function. J. Comput. Appl. Math. **233**, 2611–2619 (2010)