Superlinearly convergent smoothing Newton continuation algorithms for variational inequalities over definable sets

CHEK BENG CHUA AND LE THI KHANH HIEN

Abstract. In this paper, we use the concept of barrier-based smoothing approximations introduced by Chua and Li [7] to extend various smoothing Newton continuation algorithms to variational inequalities over general closed convex sets X. We prove that when the underlying barrier has a gradient map that is definable in some o-minimal structure, the iterates generated converge superlinearly to a solution of the variational inequality. We further prove that if X is proper and definable in the o-minimal structure $\mathbb{R}_{\text{an}}^{\mathbb{R}_{\text{alg}}}$, then the gradient map of its universal barrier is definable in the o-minimal expansion $\mathbb{R}_{\text{an,exp}}$.

1. Introduction

Let $\mathbb{E}$ denote a finite dimensional real vector space equipped with inner product $\langle\cdot,\cdot\rangle$, let X denote a closed convex subset of $\mathbb{E}$ with nonempty interior $int(X)$, let Ω denote a subset of $\mathbb{E}$ that contains X, and let $F:\Omega\to\mathbb{E}$ denote a continuous map that is differentiable in the interior $int(\Omega)$ of its domain Ω. The variational inequality $VI(X,F)$ is the problem of finding $x\in X$ such that

$$\langle F(x), y-x\rangle \geq 0 \text{ for all } y\in X. \tag{1}$$

The variational inequality (1) can be solved directly by interior point methods (see, e.g., [16, 34]), or indirectly via a reformulation (see, e.g., [5, 15, 17, 18]). For our approach, we use two of the most general nonsmooth reformulation equations of (1):

(1) the *natural map equation*

$$\begin{pmatrix} x-\Pi_X(x-y) \\ F(x)-y \end{pmatrix} = 0;$$

and

(2) the *normal map equation*

$$F(\Pi_X(z)) + z - \Pi_X(z) = 0.$$

Here, Π_X denotes the Euclidean projector onto X; i.e.,

$$\Pi_X(z) = \arg\min_{x\in X} \frac{1}{2}\|x-z\|^2,$$

where $\|\cdot\|$ is the norm induced by the inner product $\langle\cdot,\cdot\rangle$. We denote by G^{nat} the *natural map*

$$(x,y)\in\Omega\times\mathbb{E}\mapsto (x-\Pi_X(x-y), F(x)-y),$$

Date: February 13, 2014.

2010 *Mathematics Subject Classification.* 90C33, 65K05, 47J20 .

Key words and phrases. variational inequalities, smoothing Newton continuation, superlinear convergence, barrier-based smoothing approximation.

Le Thi Khanh Hien was supported by a SINGA scholarship.

and by G^{nor} the *normal map*

$$z \in \mathbb{E} \mapsto F(\Pi_X(z)) + z - \Pi_X(z).$$

Note that the domain of the natural map involves the domain Ω of F. This typically restricts the iterates in solution algorithms, unless the domain Ω is whole space $\mathbb{E}$. *To avoid this restriction, we shall assume that $\Omega = \mathbb{E}$ when the natural map is used.*

Since the Euclidean projector Π_X is generally nonsmooth, typical Newton-based methods do not apply to the natural map nor the normal map equations. One way to overcome this is to consider smoothing approximations. The study of smoothing approximations of Euclidean projectors is mostly limited to specific classes of convex sets, such as nonnegative orthants, second-order cones, positive semidefinite cones, symmetric cones and box-constrained sets.

To the best of the authors' knowledge, there are only two known approaches to developing smoothing approximations of Euclidean projectors onto general convex sets; Qi and Sun [22] developed convolution-based smoothing approximations, while Chua and Li [7] developed barrier-based smoothing approximations. Although the former can be used to smooth any nonsmooth map, it is generally computed as a multivariate integral, whence uncomputable in practice. In contrast, the latter applies only to the Euclidean projector onto any convex cone with nonempty interior, and can be computed via proximal mappings of smooth maximal monotone maps, which is generally no more difficult than computing the projector itself.

In a subsequent work [26], Qi and Sun studied the use of convolution-based smoothing approximations in a smoothing merit function algorithm to solve the natural map equation; and proved that the algorithm converges globally under the assumption of a certain P-type property on F and the boundedness of the solution set of $VI(X, F)$. They further proved that the algorithm converges superlinearly to a solution x^* under nonsingularity of the Clarke generalized Jacobian of the smoothing approximation at $(x^*, 0)$, and the semismoothness of F at x^*, and of the projector at $x^* - F(x^*)$. Besides being applicable to all variational inequalities, the approach of Qi and Sun also avoid the need for the Jacobian consistency of the smoothing approximation (i.e., the convergence of the distance between Jacobians of the smoothing approximation and the Clarke generalized Jacobian to 0, when the smoothing parameter converges to 0), which plays a crucial role in almost all other superlinearly convergent smoothing algorithms.

In this paper, the barrier-based smoothing approach is extended to convex sets, and the barrier-based smoothing approximation is shown to be semismooth whenever the barrier used to define the smoothing approximation has a gradient map that is definable in some o-minimal structure. This result allows us to deduce the local superlinear convergence of various existing smoothing Newton continuation algorithms—including the algorithm of Qi and Sun in [26]—for certain definable convex sets under assumptions of uniform nonsingularity (i.e., boundedness of the least singular value of the Jacobian away from 0) of the Newton system, and semismoothness of F. Just as in the work of Qi and Sun, the Jacobian consistency of the smoothing approximation is not required in establishing superlinear convergence.

In [22], Qi and Sun also considered a different smoothing approximation of the Euclidean projector onto a convex set finitely generated by twice-differentiable convex functions. This smoothing approximation is in fact a barrier-based smoothing approximation. Qi and Sun proved that when this smoothing approximation is used in a smoothing Newton continuation algorithm to solve the normal map equation, the algorithm converges superlinearly to a solution x^* under the linear independence constraint qualification at

x^*. Our result with the barrier-based smoothing approximation will show that when the generating convex functions are definable in some o-minimal structure, the assumption of linear independence constraint qualification can be dropped.

The paper is organized as follows. In the next section, some basic definitions and results on smoothing approximations, semismoothness and o-minimal structures are given. In Section 3, the barrier-based smoothing approximation is defined, and its semismoothness under the definability of the gradient of the barrier is established. We then present a common technique in proving the superlinear convergence of various existing smoothing Newton continuation algorithm in Section 4. Finally, in Section 5, we demonstrate the definability of the gradient of the universal barrier when the set X is definable in the o-minimal structure $\mathbb{R}_{\text{an}}^{\mathbb{R}_{\text{alg}}}$.

1.1. **Notations.** We shall use $\mathbb{R}^m_+$ (respectively, $\mathbb{R}^m_{++}$) to denote the cone of nonnegative (respectively, positive) vectors in $\mathbb{R}^m$. For two vectors $x, y \in \mathbb{R}^m$, the notation $x \geq y$ (respectively, $x > y$) means $x - y \in \mathbb{R}^m_+$ (respectively, $x - y \in \mathbb{R}^m_{++}$). For any function $f : \mathbb{E} \to \mathbb{R}$ with $\lim_{x\to 0} f(x) = 0$ and $f(x) \neq 0$ near 0, we use '$g(x) = o(f(x))$ as $x \to 0$' to mean that g is a function with domain containing a neighborhood of 0 and satisfying

$$\lim_{x\to 0} \frac{g(x)}{f(x)} = 0.$$

For any function $f : \mathbb{E} \to \mathbb{R}$, we use '$h(x) = O(f(x))$ as $x \to 0$' to mean that h is a function with domain containing a neighborhood of 0, and there exists $C > 0$ such that

$$|h(x)| \leq C|f(x)|$$

for all x near 0. For any sequence $\{x_k\}$ of real numbers with $\lim_{k\to\infty} x_k = 0$ and $x_k \neq 0$ for all k, we use $y_k = o(x_k)$ to mean that $\{y_k\}$ is a sequence of real numbers satisfying

$$\lim_{k\to\infty} \frac{y_k}{x_k} = 0.$$

For any sequence $\{x_k\}$ of real numbers, we use $z_k = O(x_k)$ to mean that $\{z_k\}$ is a sequence of real numbers and there exists $C > 0$ such that

$$|z_k| \leq C|x_k|$$

for all k. For a Fréchet-differentiable function $f : \Omega \subseteq \mathbb{E} \to \mathbb{R}$, we use ∇f to denote the gradient of f. For a Fréchet-differentiable map $F : \Omega \subseteq \mathbb{E} \to \mathbb{E}'$, we use $\mathbf{J}F$ to denote the derivative of F. For a Fréchet-differentiable map

$$F : \Omega \times \Omega' \subseteq \mathbb{E} \times \mathbb{E}' \to \mathbb{E}'' : (x, y) \mapsto F(x, y),$$

we use $\mathbf{J}_x F$ and $\mathbf{J}_y F$ to denote the partial derivatives of F with respect to x and y, respectively.

2. Background

This section gives various basic definitions on smoothing approximations, semismoothness and o-minimal structures, and establishes several basic results required in this paper.

2.1. **Smoothing approximations.** A *smoothing approximation* of a continuous map $G : \mathbb{E} \to \mathbb{E}'$ is a continuous map $H : \mathbb{E} \times \mathbb{R}^m_+ \to \mathbb{E}'$ such that $H(\cdot, 0) = G$, and for each $\mu \in \mathbb{R}^m_{++}$, $H(\cdot, \mu)$ is differentiable. The variable μ is called the (m-tuple of) *smoothing parameters*. When $H(\cdot, \mu)$ converges uniformly to G as $\mu \to 0$, we say that H is a *uniform smoothing approximation*. When $H(\cdot, \mu)$ is Lipschitz in the smoothing parameters (i.e., there exists $L > 0$ such that $\|H(x, \mu) - H(x, \nu)\| \leq L \|\mu - \nu\|$ for all $x \in \mathbb{E}$ and all $\mu, \nu \in \mathbb{R}^m_+$), we say that H is a *Lipschitzian smoothing approximation*.

For convenience, we shall extend the domain of a smoothing approximation H to include negative smoothing parameters, by defining

$$H(\cdot, \mu) := H(\cdot, \Pi_{\mathbb{R}^m_+}(\mu)) \text{ for any } \mu \in \mathbb{R}^m \setminus \mathbb{R}^m_+,$$

and call it an *extended smoothing approximation*. We note that the extended smoothing approximation remains continuous, and retain any uniform convergence or Lipschitzian property.

A (extended) smoothing approximation $H : \mathbb{E} \times \mathbb{R}^m \to \mathbb{E}'$ of $G : \mathbb{E} \to \mathbb{E}'$ is said to *approximate superlinearly at* $\bar{x} \in \mathbb{E}$ if for any $(x, \mu) \to (\bar{x}, 0)$,

$$\|H(x, \mu) - G(\bar{x}) - \mathbf{J}_x H(x, \mu)(x - \bar{x})\| = o(\|x - \bar{x}\|) + O(\|\mu\|);$$

see [22]. For any $\gamma > 0$, it is said to *approximate with order* $(1 + \gamma)$ *at* $\bar{x}$ if $o(\|x - \bar{x}\|)$ is replaced by $O(\|x - \bar{x}\|^{1+\gamma})$ in the above equation. An order-2 approximation is also called a *quadratic approximation*.

2.2. **Semismoothness.** A locally Lipschitz continuous map $F : \mathbb{E} \mapsto \mathbb{E}'$ is said to be *semismooth at* x if the limit

$$\lim_{\substack{V \in \partial F(x+th') \\ h' \to h,\ t \downarrow 0}} V h'$$

exists for every $h \in \mathbb{E}$, where $\partial F(x + th')$ denotes the generalized Jacobian of F at $x + th'$ as defined by Clarke [8, §2.6]; see [23]. If a locally Lipschitz continuous map $F : \mathbb{E} \mapsto \mathbb{E}'$ is semismooth at x, then the directional derivative

$$F'(x; h) := \lim_{t \downarrow 0} \frac{F(x + th) - F(x)}{t}$$

exists, and equals the limit

$$\lim_{\substack{V \in \partial F(x+th') \\ h' \to h,\ t \downarrow 0}} V h';$$

see [23, Proposition 2.1].

It follows from [23, Theorem 2.3] and [10, Proposition 3.1.3] that a locally Lipschitz continuous map $F : \mathbb{E} \mapsto \mathbb{E}'$ is semismooth at x if and only if it is directionally differentiable at x and

$$\lim_{\substack{x+h \in D_F \\ h \to 0}} \frac{\|F(x + h) - F(x) - F'(x + h; h)\|}{\|h\|} = 0.$$

where D_F denotes the set of points at which F is Fréchet-differentiable. This observation leads to the following notion of higher-order semismoothness: a locally Lipschitz continuous map $F : \mathbb{E} \mapsto \mathbb{E}'$ is said to be γ*-order semismooth at* x for some $\gamma \in (0, 1]$ if it is directionally differentiable at x and

$$\limsup_{\substack{x+h \in D_F \\ h \to 0}} \frac{\|F(x + h) - F(x) - F'(x + h; h)\|}{\|h\|^{1+\gamma}} < \infty$$

see [27, Theorem 3.7]. A 1-order semismooth map is also said to be *strongly semismooth*. Finally, we note that a map is (γ-order) semismooth if its component functions are (γ-order) semismooth, and that compositions of (γ-order) semismooth maps are (γ-order) semismooth. These follow from [23, Corollary 2.4], [19, Theorem 5] and [11, Theorem 19].

Theorem 2.1. *If a locally Lipschitz continuous smoothing approximation $H : \mathbb{E} \times \mathbb{R}^m \to \mathbb{E}'$ of $G : \mathbb{E} \to \mathbb{E}'$ is semismooth (respectively, γ-order semismooth) at $(\bar{x}, 0)$, then it approximates superlinearly (respectively, with order $1 + \gamma$) at $\bar{x}$.*

Proof. Since H is locally Lipschitz continuous at $(\bar{x}, 0)$, the Jacobian $\mathbf{J}_\mu H$ is locally bounded near $(\bar{x}, 0)$. Therefore for any $(x, \mu) \to (\bar{x}, 0)$,

$$\begin{aligned}&\|H(x,\mu) - G(\bar{x}) - \mathbf{J}_x H(x,\mu)(x-\bar{x})\|\\ &\leq \|H(x,\mu) - H(\bar{x},0) - H'(x,\mu; x-\bar{x},\mu)\| + \|\mathbf{J}_\mu H(x,\mu)\mu\|\\ &= o(\|(x-\bar{x},\mu)\|) + O(\|\mu\|)\end{aligned}$$

(respectively, $O(\|(x-\bar{x},\mu)\|^{1+\gamma}) + O(\|\mu\|)$). Finally, $\|(x-\bar{x},\mu)\| = O(\max\{\|x-\bar{x}\|, \|\mu\|\})$. □

2.3. **O-minimal structures.** An *o-minimal structure on the real ordered field* $\mathbb{R}$ is a sequence of Boolean algebras $\mathcal{O} = \{\mathcal{O}_n\}_{n=1}^\infty$ of subsets of $\mathbb{R}^n$ such that for each $n \geq 1$,

1. if $A \in \mathcal{O}_n$, then both $A \times \mathbb{R}$ and $\mathbb{R} \times A$ belong to $\mathcal{O}_{n+1}$;
2. $\mathcal{O}_n$ contains every algebraic subsets of $\mathbb{R}^n$;
3. if $A \in \mathcal{O}_{n+1}$, then $\pi(A) \in \mathcal{O}_n$, where $\pi : \mathbb{R}^{n+1} \to \mathbb{R}^n : (x_1, \ldots, x_n, x_{n+1}) \mapsto (x_1, \ldots, x_n)$ is the projector on the first n coordinates; and
4. the sets in $\mathcal{O}_1$ are exactly the finite unions of intervals and points.

The sets in each $\mathcal{O}_n$ are said to be *definable in* $\mathcal{O}$. A map $F : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is said to be *definable in* $\mathcal{O}$ if its graph is a definable set in $\mathcal{O}$; i.e., $\{(x, y) \in A \times \mathbb{R}^m : y = F(x)\} \in \mathcal{O}_{n+m}$.

Example 2.1 (Semialgebraic sets). The smallest o-minimal structure on $\mathbb{R}$ is the class $\mathcal{SA}$ of all semialgebraic sets. A set is semialgebraic if it can be written as a finite union of sets of the form

$$\{x \in \mathbb{R}^n : p_1(x) = \cdots = p_k(x) = 0, q_1(x) > 0, \ldots, q_l(x) > 0\},$$

where $p_1, \ldots, p_k, q_1, \ldots, q_l \in \mathbb{R}[X]$. This example appeals to the fact that the projection of a semialgebraic set is semialgebraic by the Tarski-Seidenberg principle; see, e.g., [2].

Example 2.2 (Globally subanalytic sets). A set is said to be globally subanalytic if its image under the map

$$(x_1, \ldots, x_n) \in \mathbb{R}^n \mapsto \left(\frac{x_1}{\sqrt{1+x_1^2}}, \ldots, \frac{x_n}{\sqrt{1+x_n^2}}\right)$$

is a subanalytic set; see, e.g., [28]. The collection of all globally subanalytic sets is an o-minimal structure $\mathbb{R}_{\text{an}}$ on $\mathbb{R}$. It is the smallest o-minimal structure that contains sets of the form

$$\{(x,t) \in [-1,1]^n \times \mathbb{R} : f(x) = t\},$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a restricted analytic function; i.e., a function such that $f|_{[-1,1]^n}$ is analytic and vanishes identically off $[-1,1]^n$; see [32].

Example 2.3 ($\mathbb{R}_{\text{an,exp}}$). The smallest o-minimal structure that contains $\mathbb{R}_{\text{an}}$ and the set $\{(x, e^x) : x \in \mathbb{R}\}$ is denoted by $\mathbb{R}_{\text{an,exp}}$. We say that $\mathbb{R}_{\text{an,exp}}$ is the o-minimal expansion of $\mathbb{R}_{\text{an}}$ by the exponential function; see, e.g., [30, 31].

Example 2.4 ($\mathbb{R}^{\mathbb{R}}_{\text{an}}$ and $\mathbb{R}^{\mathbb{R}_{\text{alg}}}_{\text{an}}$). We denote by $\mathbb{R}^{\mathbb{R}}_{\text{an}}$ the o-minimal expansion of $\mathbb{R}_{\text{an}}$ by all power functions

$$x \mapsto \begin{cases} x^r & \text{if } x > 0, \\ 0 & \text{if } x \leq 0; \end{cases}$$

see, e.g., [20]. When we restrict the powers to real algebraic numbers, we get a smaller o-minimal expansion of $\mathbb{R}_{\text{an}}$, which we denote by $\mathbb{R}^{\mathbb{R}_{\text{alg}}}_{\text{an}}$.

The o-minimal structures in these examples satisfy the following strict inclusions [32, Part 2.5]

$$\mathcal{SA} \subsetneqq \mathbb{R}_{\text{an}} \subsetneqq \mathbb{R}^{\mathbb{R}_{\text{alg}}}_{\text{an}} \subsetneqq \mathbb{R}^{\mathbb{R}}_{\text{an}} \subsetneqq \mathbb{R}_{\text{an,exp}}.$$

From the definition of an o-minimal structure, especially closure under projection, one can establish many stability results for definable sets and functions. We list some of these results here as they will be used in this paper. We refer the readers to [9, Theorem 1.13], [3], and [29] for their proofs.

Proposition 2.1 (Stability results)**.** *Let $\mathcal{O}$ be an o-minimal structure on $\mathbb{R}$.*

(1) *If $A \subseteq \mathbb{R}^m$, and $B \subseteq \mathbb{R}^{n+m}$ are definable in $\mathcal{O}$, then the sets $\{x : \forall y \in A, (x, y) \in B\}$ and $\{x : \exists y \in A, (x, y) \in B\}$ are definable in $\mathcal{O}$.*
(2) *The closure, interior and product of definable sets in $\mathcal{O}$ are definable in $\mathcal{O}$.*
(3) *If a map $G : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is definable in $\mathcal{O}$, then its derivative $\boldsymbol{J}G$ and its partial derivatives $\boldsymbol{J}_{x_i}G$ (if they exist) are definable in $\mathcal{O}$. If, in addition, G is injective, then its inverse map is definable in $\mathcal{O}$.*
(4) *If the maps $G : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and $F : V \supseteq G(U) \to \mathbb{R}^p$ are definable in $\mathcal{O}$, then the composition $F \circ G : U \to \mathbb{R}^p$ is definable in $\mathcal{O}$.*
(5) *A vector-valued map is definable in $\mathcal{O}$ if and only if each of its component function is definable in $\mathcal{O}$.*

3. Barrier-based smoothing approximations of closed convex sets

For each differentiable barrier $f : int(X) \to \mathbb{R}$ on a closed convex set $X \in \mathbb{E}$ with nonempty interior (i.e., $f(x_k) \to \infty$ for any convergent sequence $\{x_k\}$ in $int(X)$ with limit in the boundary of X), we define the map $p : \mathbb{E} \times \mathbb{R} \to \mathbb{E}$ via

$$\begin{cases} p(z, \mu) + \mu^2 \nabla f(p(z, \mu)) = z & \text{when } \mu > 0, \\ p(\cdot, \mu) = \Pi_X & \text{when } \mu \leq 0. \end{cases} \tag{2}$$

We note that for each $\mu > 0$, the map $p(\cdot, \mu)$ is the proximal mapping of $x \mapsto \mu^2 \nabla f(x)$, which is maximal monotone by Löhne's characterization (cf. [7, Proposition 3.1]), whence is a bijection between $int(X)$ and $\mathbb{E}$ by Minty's criterion. Thus p is well-defined.

Theorem 3.1 (Barrier-based smoothing approximation)**.** *If f is a twice continuously differentiable barrier on $X \subset \mathbb{E}$, then the map $p : \mathbb{E} \times \mathbb{R} \to \mathbb{E}$ defined via (2) is an extended smoothing approximation of the Euclidean projector Π_X.*

Proof. By the Implicit Function Theorem, the continuous differentiability of ∇f implies the differentiability of p over $\mathbb{E} \times \mathbb{R}_{++}$.

It remains to show that for each fixed $z \in \mathbb{E}$, $\{p(w_k, \mu_k)\}$ converges to $\Pi_X(z)$ for any sequence $\{(w_k, \mu_k)\}$ in $\mathbb{E} \times \mathbb{R}_{++}$ converging to $(z, 0)$. We first show that the sequence $\{p(w_k, \mu_k)\}$ is bounded. Pick an arbitrary but fixed $e \in int(X)$. The sequence $\{e_k := e + \mu_k^2 \nabla f(e)\}$ is bounded. Moreover, $p(e_k, \mu_k) = e$ by definition. For each k, the nonexpansiveness of the proximal mapping $p(\cdot, \mu_k)$ implies

$$\|p(w_k, \mu_k)\| \le \|p(w_k, \mu_k) - p(e_k, \mu_k)\| + \|p(e_k, \mu_k)\| \le \|w_k - e_k\| + \|e\|\,;$$

thus $\{p(w_k, \mu_k)\}$ is bounded. Finally, we note that since $\mu_k > 0$, it follows that $p(w_k, \mu_k)$ is the unique minimizer to the barrier problem $\min\{\frac{1}{2}\|x - w_k\|^2 + \mu_k^2 f(x)\}$, and every limit point of these minimizers must be the unique minimizer $\Pi_X(z)$ of the convex optimization problem $\min\{\frac{1}{2}\|x - z\|^2 : x \in X\}$.[1] □

Definition 3.1 (Barrier-based smoothing approximation)**.** Given a twice continuously differentiable barrier f on $S \subset \mathbb{E}$, the *(extended barrier-based) smoothing approximation defined by* f is the map $p : \mathbb{E} \times \mathbb{R} \to \mathbb{E}$ that satisfies (2). It is a smoothing approximation of the Euclidean projector Π_X.

Definition 3.2 (ϑ-barrier)**.** The *barrier parameter* of a twice continuously differentiable barrier f on $X \subset \mathbb{E}$ is

$$\inf\left\{\vartheta \ge 0 : \inf_{x \in int(X),\ h \in \mathbb{E}} \vartheta \left\langle h, \nabla^2 f(x) h\right\rangle - \langle \nabla f(x), h\rangle^2 \ge 0\right\}.$$

A *ϑ-barrier* is a twice continuously differentiable barrier with a finite barrier parameter ϑ.

Theorem 3.2 (Barrier-based uniform smoothing approximation)**.** *The smoothing approximation p defined by a ϑ-barrier $f : int(X) \to \mathbb{R}$ is $\sqrt{\vartheta}$-Lipschitzian; i.e., Lipschitz continuous with modulus $\sqrt{\vartheta}$ in the smoothing parameter.*

Consequently, in this case, p is a uniform smoothing approximation of the Euclidean projector Π_X.

Proof. By appealing to the continuity of the smoothing approximation, it suffices to only consider positive smoothing parameters. Since

$$\begin{aligned}
&\vartheta(\mu - \nu)^2 - \|p(z, \mu) - p(z, \nu)\|^2 \\
&= \vartheta(\mu - \nu)^2 - \langle p(z, \mu) - p(z, \nu), p(z, \mu) - p(z, \nu)\rangle \\
&= \vartheta(\mu - \nu)^2 + \left\langle p(z, \mu) - p(z, \nu), \mu^2 \nabla f(p(z, \mu)) - \nu^2 \nabla f(p(z, \nu))\right\rangle \\
&= \left\langle (p(z, \mu) - p(z, \nu), \mu - \nu), (\mu^2 \nabla f(p(z, \mu)) - \nu^2 \nabla f(p(z, \nu)), \vartheta\mu - \vartheta\nu)\right\rangle,
\end{aligned}$$

it suffices to show that the map $(x, \mu) \in int(X) \times \mathbb{R}_{++} \mapsto (\mu^2 \nabla f(x), \vartheta\mu)$ is monotone. To this end, we check that its Jacobian

$$(h, \tau) \in \mathbb{E} \times \mathbb{R} \mapsto (\mu^2 \nabla^2 f(x) h + 2\mu \nabla f(x)\tau, \vartheta\tau)$$

is a monotone linear map at each $(x, \mu) \in int(X) \times \mathbb{R}_{++}$. Indeed, its symmetric part

$$(h, \tau) \in \mathbb{E} \times \mathbb{R} \mapsto (\mu^2 \nabla^2 f(x) h + \mu \nabla f(x)\tau, \mu \langle \nabla f(x), h\rangle + \vartheta\tau)$$

is positive semidefinite if and only if the Schur complement

$$h \in \mathbb{E} \mapsto \frac{\mu^2}{\vartheta}(\vartheta \nabla^2 f(x) h - \langle \nabla f(x), h\rangle \nabla f(x))$$

is positive semidefinite. □

[1] See, e.g., [1, Proposition 4.1.1].

Henceforth, we assume that the barrier f is a ϑ-barrier.

With the uniform smoothing approximation p defined by f, we can now define a smoothing approximation

$$H^{\text{nat}} : (x, y, \mu, \varepsilon) \in \mathbb{E}^2 \times \mathbb{R}^2 \mapsto \begin{pmatrix} x - p(x - y, \mu) \\ F(x) + \Pi_{\mathbb{R}_+}(\varepsilon)x - y \end{pmatrix} \tag{3}$$

of the natural map, which incorporates a regularization of the map F; and a smoothing approximation

$$H^{\text{nor}} : (z, \mu) \in \mathbb{E} \times \mathbb{R} \mapsto F(p(z, \mu)) + z - p(z, \mu) \tag{4}$$

of the normal map.

We note that the local Lipschitz continuity of F carries over to the smoothing approximations H^{nat} and H^{nor} under the $\sqrt{\vartheta}$-Lipschitz continuity of p in the smoothing parameter.

Proposition 3.1. *If f is a ϑ-barrier, and F is locally Lipschitz continuous, then the smoothing approximations H^{nat} and H^{nor} are locally Lipschitz continuous.*

3.1. **Definability and superlinear approximations.** As shown by Bolte et al [3], a locally Lipschitz definable function (more generally a locally Lipschitz tame function) is semismooth. Moreover, a locally Lipschitz function that is definable in a polynomially bounded o-minimal structure is γ-order semismooth for some $\gamma > 0$; see Remarks 3 and 4 of [3]. Examples of polynomially bounded o-minimal structures are substructures of $\mathbb{R}_{\text{an}}^{\mathbb{R}}$; see, e.g., [25, p 184]. Thus if the smoothing approximation p is definable in some o-minimal structure (respectively, polynomially bounded o-minimal structure), then it is semismooth (respectively, γ-order semismooth), and hence we may apply Theorem 2.1 to deduce that it approximates superlinearly (respectively, with order $1+\gamma$). The following proposition gives a necessary and sufficient condition for p to be definable.

Proposition 3.2. *The smoothing approximation p defined by f is definable in an o-minimal structure $\mathcal{O}$ if and only if ∇f is definable in $\mathcal{O}$.*

Consequently, p is semismooth when f is definable in $\mathcal{O}$. Moreover, p is γ-order semismooth for some $\gamma > 0$ when f is definable in a polynomially bounded o-minimal structure $\mathcal{O}$.

Proof. The graph of p is

$$\left\{(z, \mu, x) : x \in int(K),\ \mu > 0,\ x + \mu^2 \nabla f(x) = z\right\}$$
$$\bigcup \left\{(z, \mu, x) : \mu \leq 0,\ x \in K,\ \langle x, z - x\rangle = 0 \text{ and } \forall w \in K, \langle w, z - x\rangle \leq 0\right\},$$

which is definable when ∇f and K are definable. Since K is the closure of the domain of ∇f, it is definable whenever ∇f is.

Conversely, the graph of ∇f is $\{(x, y) : x = p(x + y, 1)\}$, which is definable when p is definable.

The final statements then follow from Theorem 1, and Remarks 3 and 4 of [3]. □

As consequences of Propositions 3.1 and 3.2, and Theorem 2.1, we deduce sufficient conditions for the smoothing approximations of the natural and normal maps to approximate superlinearly (respectively, with order $1+\gamma$).

Proposition 3.3. *For any $x, y \in \mathbb{E}$, if f is a ϑ-barrier with a gradient map that is definable in an o-minimal structure (respectively, polynomially bounded o-minimal structure), and F is semismooth (respectively, γ'-order semismooth) at x, then the smoothing approximation H^{nat} defined in (3) is semismooth (respectively, γ-order semismooth for* *some $\gamma \in (0, \gamma']$) at $(x, y, 0)$. Consequently, it approximates superlinearly (respectively, with order $1 + \gamma$) at $(x, y, 0)$.*

Proposition 3.4. *For any $z \in \mathbb{E}$, if f is a ϑ-barrier with a gradient map that is definable in an o-minimal structure (respectively, polynomially bounded o-minimal structure), and F is semismooth (respectively, γ'-order semismooth) at $\Pi_X(z)$, then the smoothing approximation H^{nat} defined in (3) is semismooth (respectively, γ-order semismooth for some $\gamma \in (0, \gamma']$) at $(z, 0)$. Consequently, it approximates superlinearly (respectively, with order $1 + \gamma$) at $(z, 0)$.*

We now give a few examples of twice continuously differentiable barriers with finite barrier parameters and with gradients that are definable is some o-minimal structures.

Example 3.1 (Polyhedral sets). A barrier of the polyhedral set $\{x : a_i^T x - b_i \leq 0 \ i = 1, \ldots, m\}$, where $a_1, \ldots, a_m \in \mathbb{R}^n$ and $b_1, \ldots, b_m \in \mathbb{R}$, is $x \mapsto -\sum_{i=1}^m \log(b_i - a_i^T x)$. Its barrier parameter is m, and its gradient is definable in the o-minimal structure $\mathcal{SA}$ of semialgebraic sets.

Example 3.2 (Symmetric cones). A symmetric cone K is the cone of squares of some Euclidean Jordan algebra $\mathfrak{J}$, and is thus definable in the o-minimal structure $\mathcal{SA}$. Its standard log-determinant barrier is $x \in int(K) \mapsto -\log\det(x)$, where the determinant $\det(\cdot)$ is a polynomial in x. The gradient of the log-determinant barrier is thus definable in the o-minimal structure $\mathcal{SA}$. Its barrier parameter is the rank of the symmetric cone.

Example 3.3 (Homogeneous cones). A homogeneous cone K is the cone associated with some T-algebra $\mathfrak{A}$; see, e.g., [6, Theorem 1]. It is thus definable in the o-minimal structure $\mathcal{SA}$. The only known optimal self-concordant barrier of K [6, Section 3.1] has the form $x \mapsto -\sum_{i=1}^r \log \rho_i(u_x)^2$, where $x \mapsto \rho_i(u_x)^2$ are rational functions; see [33, §III.3]. Thus the gradient of this barrier is definable in the o-minimal structure $\mathcal{SA}$. Its barrier parameter is the rank of the homogeneous cone.

Example 3.4 (Hyperbolicity cones). By Proposition 18 and Theorem 20 of [24], the hyperbolicity cone defined by a hyperbolic polynomial q is definable in the o-minimal structure $\mathcal{SA}$. A barrier of this hyperbolicity cone is $x \mapsto -\log q(x)$. The barrier parameter of this barrier is the degree of the polynomial q. Since q is a polynomial, the gradient of this barrier is definable in the o-minimal structure $\mathcal{SA}$.

Example 3.5 (Power cones). A (high-dimensional) power cone is

$$\left\{(x_1, \ldots, x_n, z_1, \ldots, z_m) \in \mathbb{R}_+^n \times \mathbb{R}^m : \prod_{i=1}^n x_i^{\alpha_i} \geq \|z\|\right\}$$

where the exponents $\alpha_1, \ldots, \alpha_n \in (0, 1]$ sum to 1. A barrier of this cone is $-\log(\prod_{i=1}^n x_i^{2\alpha_i} - \|z\|^2)$, whose gradient is definable in the o-minimal structure $\mathbb{R}_{\text{an}}^{\mathbb{R}}$. It has barrier parameter 2.

Example 3.6 (Finitely generated convex sets). Consider the convex set $X = \{x : g_i(x) \leq 0 \text{ for } i = 1, \ldots, N\}$ generated by a finite number of twice continuously differentiable convex functions $g_1, \ldots, g_N : \mathbb{E} \to \mathbb{R}$. When we assume that for $i = 1, \ldots, N$, $g_i(x) < 0$ for all $x \in int(X)$, and that $int(X) \neq \emptyset$, the function $f : x \in int(X) \mapsto -\sum_{i=1}^N \log(-g_i(x))$ is a barrier of X. This barrier has gradient and Hessian

$$\nabla f : x \mapsto \sum_{i=1}^N \frac{-\nabla g_i(x)}{-g_i(x)}$$

and

$$\nabla^2 f(x) : h \mapsto \sum_{i=1}^{N} \frac{\nabla^2 g_i(x) h}{-g_i(x)} + \sum_{i=1}^{N} \frac{\langle \nabla g_i(x), h \rangle \nabla g_i(x)}{g_i(x)^2},$$

respectively. The gradient is definable in an o-minimal structure whenever the functions $g_1, \ldots, g_N$ are definable in the same o-minimal structure. Since $\nabla^2 g_i(x)$ is positive semidefinite for each i, we have that, for any $x \in int(X)$ and any $h \in \mathbb{E}$,

$$\begin{aligned}
\langle \nabla f(x), h \rangle^2 &= \left(\sum_{i=1}^{N} \frac{\langle -\nabla g_i(x), h \rangle}{-g_i(x)} \right)^2 \\
&\leq N \sum_{i=1}^{N} \left(\frac{\langle -\nabla g_i(x), h \rangle}{-g_i(x)} \right)^2 \\
&\leq N \left\langle h, \nabla^2 f(x) h \right\rangle;
\end{aligned}$$

i.e., f has a finite barrier parameter $\vartheta \leq N$. The uniform smoothing approximation defined by f coincide with the one defined in [22, Section 4].

4. Smoothing Newton continuation algorithms

Given $G : \mathbb{E} \to \mathbb{E}$ and a smoothing approximation $H : \mathbb{E} \times \mathbb{R}^m \to \mathbb{E}$ of G, consider an algorithm that generates an infinite sequence $\{(w_k, \mu_k)\}$ in $\mathbb{E} \times \mathbb{R}^m_+$ such that $\mu_k \to 0$ and $\|H(w_k, \mu_k)\| \to 0$. The continuity of H then implies that any accumulation point of $\{w_k\}$ is a zero of G. We shall now consider the local superlinear convergence of such algorithms.

Lemma 4.1. *Suppose $\{w_k\}$ is a convergent sequence in $\mathbb{E}$ converging to a zero w^* of G. If, in addition, H is Lipschitz continuous near $(w^*, 0)$ and approximates superlinearly (respectively, with order $1+\gamma$) at $(w^*, 0)$, and $\{\mu_k\}$ and $\{\nu_k\}$ are sequences in $\mathbb{R}^m_{++}$ such that*

(1) *$\{\boldsymbol{J}_w H(w_k, \mu_k)\}$ is uniformly nonsingular (i.e., each term is nonsingular with the sequence of inverses having uniformly bounded norms), and*
(2) *$\|\mu_k\|, \|\nu_k\| = o(\|H(w_k, 0)\|)$*
(respectively, $\|\mu_k\|, \|\nu_k\| = O(\|H(w_k, 0)\|^{1+\gamma})$),

then the sequence of solutions $\{d_k\}$ to $H(w_k, \nu_k) + \boldsymbol{J}_w H(w_k, \mu_k) d_k = 0$ satisfies

$$\|w_k + d_k - w^*\| = o(\|w_k - w^*\|)$$

(respectively, $\|w_k + d_k - w^*\| = O(\|w_k - w^*\|^{1+\gamma})$).

Moreover, for any sequences $\{\tilde{\mu}_k\}, \{\tilde{\nu}_k\}, \{\hat{\nu}_k\}$ in $\mathbb{R}^m$ such that $\|\tilde{\mu}_k\|, \|\tilde{\nu}_k\|, \|\hat{\nu}_k\| = o(\|H(w_k, 0)\|)$,

(i) *$\|H(w_k + d_k, \tilde{\mu}_k)\|^2 - \|H(w_k, \tilde{\mu}_k)\|^2 + \|H(w_k, \tilde{\nu}_k)\|^2 = o(\|H(w_k, \tilde{\nu}_k)\|^2)$ and*
(ii) *$\|H(w_k + d_k, \hat{\nu}_k)\| = o(\|H(w_k, \hat{\nu}_{k-1})\| + \|\hat{\nu}_k - \hat{\nu}_{k-1}\|)$.*

Proof. We first note that the Lipschitz continuity of H near $(w^*, 0)$ implies that

$$\begin{aligned}
o(\|H(w_k, 0)\|) = o(\|H(w_k, 0) - H(w^*, 0)\|) = o(\|w_k - w^*\|) \quad \text{and} \\
O(\|H(w_k, 0)\|^{1+\gamma}) = O(\|H(w_k, 0) - H(w^*, 0)\|^{1+\gamma}) = O(\|w_k - w^*\|^{1+\gamma}).
\end{aligned} \tag{5}$$

Under the hypothesis of the boundedness of $\{\mathbf{J}_w H(w_k,\mu_k)^{-1}\}$,

$$\begin{aligned}\|w_k+d_k-w^*\| &= \left\|w_k-\mathbf{J}_w H(w_k,\mu_k)^{-1}H(w_k,\nu_k)-w^*\right\| \\ &= O(\|\mathbf{J}_w H(w_k,\mu_k)(w_k-w^*)-H(w_k,\nu_k)\|) \\ &= O(\|\mathbf{J}_w H(w_k,\mu_k)(w_k-w^*)-H(w_k,\mu_k)+H(w^*,0)\|) \\ &\quad + O(\|H(w_k,\mu_k)-H(w_k,\nu_k)\|).\end{aligned}$$

The local Lipschitz continuity of H near $(w^*,0)$ bounds the second term by $O(\|\mu_k-\nu_k\|)$. The first term is $o(\|w_k-w^*\|)+O(\|\mu_k\|)$ (respectively, $O(\|w_k-w^*\|^{1+\gamma})+O(\|\mu_k\|)$) under the hypothesis that H is a superlinear (respectively, order-$(1+\gamma)$) approximation at $(w^*,0)$. Subsequently, we deduce from the hypothesis $\|\mu_k\|,\|\nu_k\|=o(\|H(w_k,0)\|)$ (respectively, $\|\mu_k\|,\|\nu_k\|=O(\|H(w_k,0)\|^{1+\gamma})$) and (5) that

$$\|w_k+d_k-w^*\|=o(\|w_k-w^*\|) \tag{6}$$

(respectively, $\|w_k+d_k-w^*\|=O(\|w_k-w^*\|^{1+\gamma})$).

(i) We now show that $\|H(w_k+d_k,\tilde{\mu}_k)\|^2-\|H(w_k,\tilde{\mu}_k)\|^2+\|H(w_k,\tilde{\nu}_k)\|^2=o(\|H(w_k,\tilde{\nu}_k)\|^2)$. The last two terms on the left can be bounded by

$$\begin{aligned}&\|H(w_k,\tilde{\nu}_k)\|^2-\|H(w_k,\tilde{\mu}_k)\|^2 \\ &=(\|H(w_k,\tilde{\nu}_k)\|-\|H(w_k,\tilde{\mu}_k)\|)(\|H(w_k,\tilde{\nu}_k)\|+\|H(w_k,\tilde{\mu}_k)\|) \\ &\le \|H(w_k,\tilde{\nu}_k)-H(w_k,\tilde{\mu}_k)\|\,(2\,\|H(w_k,\tilde{\nu}_k)\|+\|H(w_k,\tilde{\mu}_k)-H(w_k,\tilde{\nu}_k)\|) \\ &=O(\|\tilde{\mu}_k-\tilde{\nu}_k\|\,(\|H(w_k,\tilde{\nu}_k)\|+\|\tilde{\mu}_k-\tilde{\nu}_k\|)) \\ &=o(\|w_k-w^*\|\,(\|H(w_k,\tilde{\nu}_k)\|+\|w_k-w^*\|))\end{aligned}$$

via the Lipschitz continuity of H near $(w^*,0)$, (5), and the hypothesis $\|\tilde{\mu}_k\|,\|\tilde{\nu}_k\|=o(\|H(w_k,0)\|)$. To bound the first term on the left, we note from (6) that $w_k+d_k\to w^*$, whence we may apply the Lipschitz continuity of H near $(w^*,0)$ to deduce

$$\begin{aligned}\|H(w_k+d_k,\tilde{\mu}_k)\| &= \|H(w_k+d_k,\tilde{\mu}_k)-H(w^*,0)\| \\ &= O(\|w_k+d_k-w^*\|+\|\tilde{\mu}_k\|) \\ &= o(\|w_k-w^*\|),\end{aligned} \tag{7}$$

where we have used the hypothesis $\|\tilde{\mu}_k\|=o(\|H(w_k,0)\|)$, (5) and (6) in the last equality. It remains to show that $\|w_k-w^*\|=O(\|H(w_k,\tilde{\nu}_k)\|)$. This follows from

$$\begin{aligned}\|w_k-w^*\| &= \left\|w_k+d_k-w^*-\mathbf{J}_w H(w_k,\mu_k)^{-1}H(w_k,\nu_k)\right\| \\ &= o(\|w_k-w^*\|)+O(\|H(w_k,\nu_k)\|) \\ &= o(\|w_k-w^*\|)+O(\|H(w_k,\tilde{\nu}_k)\|+\|\nu_k-\tilde{\nu}_k\|) \\ &= o(\|w_k-w^*\|)+O(\|H(w_k,\tilde{\nu}_k)\|),\end{aligned} \tag{8}$$

where we have used (6), the boundedness of $\{\mathbf{J}_w H(w_k,\mu_k)^{-1}\}$, the Lipschitz continuity of H near $(w^*,0)$, the hypothesis $\|\nu_k\|,\|\tilde{\nu}_k\|=o(\|H(w_k,0)\|)$ and (5).

(ii) Since the equations in (7) and (8) holds when, respectively, $\tilde{\mu}_k$ and $\tilde{\nu}_k$ are replaced by $\hat{\nu}_k$, we can further deduce that

$$\begin{aligned}\|H(w_k+d_k,\hat{\nu}_k)\| &= o(\|w_k-w^*\|) \\ &= o(\|H(w_k,\hat{\nu}_k)\|) \\ &= o(\|H(w_k,\hat{\nu}_k)-H(w_k,\hat{\nu}_{k-1})\|+\|H(w_k,\hat{\nu}_{k-1})\|) \\ &= o(\|\hat{\nu}_k-\hat{\nu}_{k-1}\|+\|H(w_k,\hat{\nu}_{k-1})\|),\end{aligned}$$

where we have used the Lipschitz continuity of H near $(w^*,0)$.

□

We now consider three algorithms for solving either the natural map equation or the normal map equation, and prove their local superlinear convergence. The three algorithms, Algorithms 4.1, 4.2 and 4.3, were analyzed, respectively, in [13], [4] and [22] for their global and local superlinear convergence. These algorithms, although only stated in the context of specific classes of variational inequalities and smoothing approximations, have been shown in the respective works to be globally convergent for general variational inequalities and smoothing approximations, albeit with various additional assumptions on the problem instance and the sequence of iterates generated by the algorithms.

The first algorithm, Algorithm 4.1, solves the natural map equation. Regularization is incorporated to ensure the boundedness of the level sets of $H^{\text{nat}}(\cdot,\mu,\varepsilon)$ for all $\mu,\varepsilon>0$. We note the slight difference in the update of μ; the additional condition in the update formula in [13] was imposed to exploit the Jacobian consistency of the smoothing approximation, and has no effect on global convergence. We drop this additional condition here because we no longer assume the Jacobian consistency of the smoothing approximation.

Algorithm 4.1 (Algorithm 2 of [13]).

Inputs: Initial data $w_0=(x_0,y_0)\in\mathbb{E}\times\mathbb{E}$, and parameters $\beta>0$, $\alpha,\eta\in(0,1)$, $\bar{\eta}\in(0,\eta]$, $\sigma\in(0,1/2)$, and $\kappa>0$.

Set $k=0$ and $\mu_0=\varepsilon_0=\left\|G^{\text{nat}}(w_0)\right\|$, and repeat the following steps until $\left\|G^{\text{nat}}(w_k)\right\|=0$.

Step 1. Set $j=0$ and $v_{k0}=w_k$.

Step 1a. Find $d_{kj}\in\mathbb{E}^2$ such that

$$H^{\text{nat}}(v_{kj},\mu_k,\varepsilon_k)+\mathbf{J}_w H^{\text{nat}}(v_{kj},\mu_k,\varepsilon_k)d_{kj}=0.$$

Step 1b. If $\left\|H^{\text{nat}}(v_{kj}+d_{kj},\mu_k,\varepsilon_k)\right\|\le\beta\eta^k$, set $w_{k+1}=v_{kj}+d_{kj}$ and proceed to Step 2.

Step 1c. Otherwise, find the largest $\lambda_{kj}\in\{1,\alpha,\alpha^2,\dots\}$ such that

$$\begin{aligned}&\left\|H^{\text{nat}}(v_{kj}+\lambda_{kj}d_{kj},\mu_k,\varepsilon_k)\right\|^2-\left\|H^{\text{nat}}(v_{kj},\mu_k,\varepsilon_k)\right\|^2\\&\le-2\sigma\lambda_{kj}\left\|H^{\text{nat}}(v_{kj},\mu_k,\varepsilon_k)\right\|^2,\end{aligned}$$

and set $v_{k,j+1}=v_{kj}+\lambda_{kj}d_{kj}$.

Step 1d. If $\left\|H^{\text{nat}}(v_{k,j+1},\mu_k,\varepsilon_k)\right\|\le\beta\eta^k$, set $w_{k+1}=v_{k,j+1}$ and proceed to Step 2.

Otherwise, update $j=j+1$ and return to Step 1a.

Step 2. Set $\mu_{k+1}=\min\{\kappa\left\|G^{\text{nat}}(w_{k+1})\right\|^2,\mu_0\bar{\eta}^{k+1}\}$ and $\varepsilon_{k+1}=\min\{\kappa\left\|G^{\text{nat}}(w_{k+1})\right\|^2,\varepsilon_0\bar{\eta}^{k+1}\}$, and update $k=k+1$.

Theorem 4.1. *If*

(1) $F:\mathbb{E}\to\mathbb{E}$ *is locally Lipschitz continuous and monotone, and*
(2) *the solution set of* $VI(X,F)$ *is nonempty and bounded,*

then Algorithm 4.1 generates a bounded sequence $\{w_k=(x_k,y_k)\}$ *with an accumulation point* $w^*=(x^*,y^*)$ *that is a zero of* G^{nat}.

If, in addition,

(1) H^{nat} *is based on the smoothing approximation defined by a* ϑ*-barrier with a gradient map that is definable in an o-minimal structure (respectively, polynomially bounded o-minimal structure),*
(2) F *is semismooth (respectively,* γ'*-order semismooth) at* x^**, and*

(3) *for any subsequence* $\{w_{k_l}\}$ *converging to* w^*, $\{\boldsymbol{J}_w H^{\mathrm{nat}}(w_{k_l}, \mu_{k_l}, \varepsilon_{k_l})\}$ *is uniformly nonsingular,*

then the full Newton step is eventually always taken (i.e., $w_{k+1} = w_k + d_{k0}$ *for all* k *sufficiently large), and the* x*-component of the sequence* $\{w_k\}$ *converges superlinearly (respectively, with order* $1+\gamma$ *for some* $\gamma \in (0, \gamma']$*) to the solution* x^* *of* $VI(X, F)$*.*

Proof. The global convergence of the algorithm is proved in Theorem 4.3 of [13].

To prove superlinear convergence, construct a subsequence $\{w_{k_l}\}$ by taking $k_0 = 0$, and recursively taking $k_{l+1} \geq k_l$ to be the least index satisfying $\|w_{k_{l+1}} - w^*\| \leq \frac{1}{2}\|w_{k_l} - w^*\|$. This subsequence is well-defined since w^* is an accumulation point of $\{w_k\}$. Since $\|(\mu_{k_l}, \varepsilon_{k_l})\| \leq \kappa\sqrt{2}\,\|H^{\mathrm{nat}}(w_{k_l}, 0)\|^2 = O(\|H^{\mathrm{nat}}(w_{k_l}, 0)\|^2) = o(\|H^{\mathrm{nat}}(w_{k_l}, 0)\|)$, we can apply Lemma 4.1, together with Proposition 3.3, on this subsequence by taking $\{\mu_k\}, \{\nu_k\}, \{\tilde{\mu}_k\}, \{\tilde{\nu}_k\}, \{\hat{\nu}_k\}$ in the lemma to be $\{(\mu_{k_l}, \varepsilon_{k_l})\}$ to get

(1) $\|w_{k_l} + d_{k_l 0} - w^*\| = o(\|w_{k_l} - w^*\|)$
(respectively, $\|w_{k_l} + d_{k_l 0} - w^*\| = O(\|w_{k_l} - w^*\|^{1+\gamma})$),
(2) $\|H^{\mathrm{nat}}(w_{k_l} + d_{k_l 0}, \mu_{k_l}, \varepsilon_{k_l})\|^2 = o(\|H^{\mathrm{nat}}(w_{k_l}, \mu_{k_l}, \varepsilon_{k_l})\|^2)$, and
(3) $H^{\mathrm{nat}}(w_{k_l} + d_{k_l 0}, \mu_{k_l}, \varepsilon_{k_l}) = o(\|H^{\mathrm{nat}}(w_{k_l}, \mu_{k_l - 1}, \varepsilon_{k_l - 1})\| + \|(\mu_{k_l} - \mu_{k_l - 1}, \varepsilon_{k_l} - \varepsilon_{k_l - 1})\|)$,

where $d_{k_l 0}$ is the search direction determined in Step 1a for $j = 0$. The third conclusion, together with $\|H^{\mathrm{nat}}(w_{k_l}, \mu_{k_l - 1}, \varepsilon_{k_l - 1})\| \leq \beta\eta^{k_l - 1}$ and

$$(|\mu_{k_l} - \mu_{k_l - 1}|, |\varepsilon_{k_l} - \varepsilon_{k_l - 1}|) \leq (\mu_{k_l - 1}, \varepsilon_{k_l - 1}) \leq \bar{\eta}^{k_l - 1}(\mu_0, \varepsilon_0) \leq \eta^{k_l - 1}(\mu_0, \varepsilon_0),$$

implies that $\|H^{\mathrm{nat}}(w_{k_l} + d_{k_l 0}, \mu_{k_l}, \varepsilon_{k_l})\| \leq \beta\eta^{k_l}$ eventually always hold. Thus the full Newton step is eventually always taken at Step 1b. It then follows from the first conclusion that for all sufficiently large l,

$$\|w_{k_l + 1} - w^*\| = \|w_{k_l} + d_{k_l 0} - w^*\| \leq \frac{1}{2}\|w_{k_l} - w^*\|,$$

whence $k_{l+1} = k_l + 1$. This means $w_{k+1} = w_k + d_{k0}$ for all k sufficiently large, and w_k converges superlinearly (respectively, with order $1+\gamma$) to w^*. □

Algorithm 4.2, on the other hand, solves the normal map equation. A significant difference between this algorithm and the other two is in the computation of the Newton direction; here, the value of the normal map G^{nor} is used, instead of that of the smoothing approximation H^{nor}. Such algorithms are often called Jacobian smoothing algorithm (since only the Jacobian of the smoothing approximation is used) or splitting algorithm (where the normal map G^{nor} is seen to be split into a smooth part $H^{\mathrm{nor}}(\cdot, \mu)$ and a nonsmooth part $G^{\mathrm{nor}} - H^{\mathrm{nor}}(\cdot, \mu)$). Once again, we drop, from the original algorithm, the condition that exploits the Jacobian consistency of the smoothing approximation, which has no effect on global convergence.

Theorem 4.2. *If* $F : \mathbb{E} \to \mathbb{E}$ *is continuously differentiable and strongly monotone, then Algorithm 4.2 generates a sequence* $\{z_k\}$ *converging to a zero* z^* *of* G^{nor}*.*

If, in addition,

(1) H^{nor} *is based on the smoothing approximation defined by a* ϑ*-barrier with a gradient map that is definable in an o-minimal structure (respectively, polynomially bounded o-minimal structure),*
(2) F *is semismooth (respectively,* γ'*-order semismooth) at* $\Pi_X(z^*)$*, and*
(3) *for any subsequence* $\{z_{k_l}\}$ *converging to* z^**,*
   (a) $\{\boldsymbol{J}_z H^{\mathrm{nor}}(z_{k_l}, \mu_{k_l})\}$ *is uniformly nonsingular,*
   (b) $\mu_{k_l} = o(\|H^{\mathrm{nor}}(z_{k_l}, 0)\|)$ *(respectively,* $\mu_{k_l} = O(\|H^{\mathrm{nor}}(z_{k_l}, 0)\|^{1+\gamma'})$*),*

Algorithm 4.2 (Algorithm 3.1 of [4]).

Inputs: Initial data $z_0 \in \mathbb{E}$, parameters $\alpha, \rho, \eta \in (0,1)$, $\sigma \in (0, \frac{1}{2}(1-\rho))$, and $L = \sup_{z\in\mathbb{E},\ \mu\in\mathbb{R}_{++}} \frac{1}{\mu} \|H^{\text{nor}}(z,\mu) - G^{\text{nor}}(z)\| \in (0,\infty)$.

Set $k=0$, $\delta_0 = \|G^{\text{nor}}(z_0)\|$ and $\mu_0 = \frac{\rho}{2L}\delta_0$, and repeat the following steps until $\|G^{\text{nor}}(z_k)\| = 0$.

Step 1. Set $j=0$ and $v_{kj} = z_k$.

Step 1a. Find $d_{kj} \in \mathbb{E}$ such that

$$G^{\text{nor}}(v_{kj}) + \mathbf{J}_z H^{\text{nor}}(v_{kj}, \mu_k) d_{kj} = 0.$$

Step 1b. Find the largest $\lambda_{kj} \in \{1, \alpha, \alpha^2, \dots\}$ such that

$$\|H^{\text{nor}}(v_{kj} + \lambda_{kj} d_{kj}, \mu_k)\|^2 - \|H^{\text{nor}}(v_{kj}, \mu_k)\|^2 \le -2\sigma\lambda_{kj} \|G^{\text{nor}}(v_{kj})\|^2,$$

and set $v_{k,j+1} = v_{kj} + \lambda_{kj} d_{kj}$.

Step 1c. If

$$\|G^{\text{nor}}(v_{k,j+1})\| \le \max\{\eta \|G^{\text{nor}}(z_k)\|, \rho^{-1} \|G^{\text{nor}}(v_{k,j+1}) - H^{\text{nor}}(v_{k,j+1}, \mu_k)\|\},$$

set $z_{k+1} = v_{k,j+1}$ and proceed to Step 2.
Otherwise, update $j = j+1$ and return to Step 1a.

Step 2. Choose $0 < \mu_{k+1} \le \min\{\frac{\rho}{2L} \|G^{\text{nor}}(z_{k+1})\|, \frac{1}{2}\mu_k\}$, and update $k = k+1$.

---

then the full Newton step is eventually always taken (i.e., $z_{k+1} = z_k + d_{k0}$ for all k sufficiently large), and the sequence $\{\Pi_X(z_k)\}$ converges superlinearly (respectively, with order $1+\gamma$ for some $\gamma \in (0, \gamma']$) to the solution $\Pi_X(z^*)$ of $VI(X,F)$.

Proof. The global convergence of the algorithm is proved in Corollary 4.1 of [4].

To prove superlinear convergence, construct a subsequence $\{z_{k_l}\}$ as in the proof of Theorem 4.1. Applying Lemma 4.1, together with Proposition 3.4, on this subsequence by taking $\{\nu_k\}, \{\tilde\nu_k\}, \{\hat\nu_k\}$ in the lemma to be the zero sequence, and by taking $\{\mu_k\}, \{\tilde\mu_k\}$ in the lemma to be $\{\mu_{k_l}\}$, gives

(1) $\|z_{k_l} + d_{k_l 0} - z^*\| = o(\|z_{k_l} - z^*\|)$
(respectively, $\|z_{k_l} + d_{k_l 0} - z^*\| = O(\|z_{k_l} - z^*\|^{1+\gamma})$),
(2) $\|H^{\text{nor}}(z_{k_l} + d_{k_l 0}, \mu_{k_l})\|^2 - \|H^{\text{nor}}(z_{k_l}, \mu_{k_l})\|^2 + \|G^{\text{nor}}(z_{k_l})\|^2 = o(\|G^{\text{nor}}(z_{k_l})\|^2)$, and
(3) $G^{\text{nor}}(z_{k_l} + d_{k_l 0}) = o(\|G^{\text{nor}}(z_{k_l})\|)$.

The second conclusion implies that the step size $\lambda_{k_l 0}$ is eventually always 1, while the third conclusion implies that $\|G^{\text{nor}}(z_{k_l} + d_{k_l 0})\| \le \eta \|G^{\text{nor}}(z_{k_l})\|$ eventually always hold. Thus the full Newton step is eventually always taken, and we can follow the same argument as in the proof of Theorem 4.1 to deduce superlinear (respectively, order-$(1+\gamma)$) convergence. □

The last algorithm, Algorithm 4.3, also solves the normal map equation. This algorithm was designed to be as generally applicable as possible, by introducing a fallback search direction when the Jacobian of the smoothing approximation fails to be nonsingular, or when the Newton direction is almost orthogonal with the steepest descent direction; cf [12]. For this algorithm, the smoothing approximation is not assumed to be Jacobian consistent.

Theorem 4.3. *If*

(1) *there exists $\bar\mu > 0$ such that the set*

$$\{z : \|H^{\text{nor}}(z,\mu)\| \le \mu\beta \textit{ for some } 0 < \mu \le \bar\mu\}$$

Algorithm 4.3 (Algorithm 5.1 of [22]).

Inputs: Initial data $z_0 \in \mathbb{E}$ and $\mu_0 > 0$, and parameters $\beta > 0$, $\alpha \in (0,1)$, $\sigma \in (0, 1/2)$, $\rho_1 > 0$, and $\rho_2 > 2$.

Set $k = 0$ and repeat the following steps until $\|G^{\rm nor}(z_k)\| = 0$.

Step 1. Set $j = 0$ and $v_{kj} = z_k$.

Step 1a. If there exists $h_{kj} \in \mathbb{E}$ satisfying

$$H^{\rm nor}(v_{kj}, \mu_k) + \mathbf{J}_z H^{\rm nor}(v_{kj}, \mu_k) h_{kj} = 0$$

and

$$\langle H^{\rm nor}(v_{kj}, \mu_k), \mathbf{J}_z H^{\rm nor}(v_{kj}, \mu_k) h_{kj} \rangle \le -\rho_1 \|h_{kj}\|^{\rho_2},$$

then set $d_{kj} = h_{kj}$.

Otherwise set $d_{kj} = -\mathbf{J}_z H^{\rm nor}(v_{kj}, \mu_k) H^{\rm nor}(v_{kj}, \mu_k)$.

Step 1b. Find the largest $\lambda_{kj} \in \{1, \alpha, \alpha^2, \dots\}$ such that

$$\|H^{\rm nor}(v_{kj} + \lambda_{kj} d_{kj}, \mu_k)\|^2 - \|H^{\rm nor}(v_{kj}, \mu_k)\|^2 \le 2\sigma\lambda_{kj} \langle H^{\rm nor}(v_{kj}, \mu_k), \mathbf{J}_z H^{\rm nor}(v_{kj}, \mu_k) d_{kj} \rangle,$$

and set $v_{k,j+1} = v_{kj} + \lambda_{kj} d_{kj}$.

Step 1c. If $\|H^{\rm nor}(v_{k,j+1}, \mu_k)\| \le \mu_k \beta$ or $\|G^{\rm nor}(v_{k,j+1})\| \le \frac{1}{2}\|G^{\rm nor}(z_k)\|$, set $z_{k+1} = v_{k,j+1}$ and proceed to Step 2.

Otherwise, update $j = j + 1$ and return to Step 1a.

Step 2. Choose $0 < \mu_{k+1} \le \min\{\frac{1}{2}\mu_k, \frac{1}{2}\|G^{\rm nor}(z_{k+1})\|^2\}$, and update $k = k + 1$.

---

is bounded,

(2) *for any $\mu > 0$, the function $\theta_\mu : z \mapsto \|H^{\rm nor}(z, \mu)\|^2$ has bounded level sets, and all stationary points of θ_μ are zeros of θ_μ, and*

(3) *the function $\|G^{\rm nor}(\cdot)\|$ has a bounded level set,*

then Algorithm 4.3 generates a bounded sequence $\{z_k\}$ with an accumulation point z^* that is a zero of $G^{\rm nor}$.

If, in addition,

(1) *$H^{\rm nor}$ is based on the smoothing approximation defined by a ϑ-barrier with a gradient map that is definable in an o-minimal structure (respectively, polynomially bounded o-minimal structure),*

(2) *F is semismooth (respectively, γ'-order semismooth) at $\Pi_X(z^*)$, and*

(3) *for any subsequence $\{z_{k_l}\}$ converging to z^*, $\{\mathbf{J}_z H^{\rm nor}(z_{k_l}, \mu_{k_l})\}$ is eventually uniformly nonsingular,*

then the full Newton step is eventually always taken (i.e., $z_{k+1} = z_k + d_{k0}$ for all k sufficiently large), and the sequence $\{\Pi_X(z_k)\}$ converges superlinearly (respectively, with order $1 + \gamma$ for some $\gamma \in (0, \gamma']$) to the solution $\Pi_X(z^*)$ of $VI(X, F)$.

Proof. The global convergence of the algorithm is proved in Theorem 6.1 of [22].

To prove superlinear convergence, construct a subsequence $\{z_{k_l}\}$ as in the proof of Theorem 4.1. Since $\mu_{k_l} \le \frac{1}{2}\|H^{\rm nor}(z_{k_l}, 0)\|^2 = O(\|H^{\rm nor}(z_{k_l}, 0)\|^2) = o(\|H^{\rm nor}(z_{k_l}, 0)\|)$, we can apply Lemma 4.1, together with Proposition 3.4, on this subsequence by taking $\{\mu_k\}, \{\nu_k\}, \{\tilde{\nu}_k\}, \{\tilde{\mu}_k\}$ in the lemma to be $\{\mu_{k_l}\}$, and by taking $\{\hat{\nu}_k\}$ in the lemma to be the zero sequence, to get

(1) $\|z_{k_l} + h_{k_l 0} - z^*\| = o(\|z_{k_l} - z^*\|)$
(respectively, $\|z_{k_l} + h_{k_l 0} - z^*\| = O(\|z_{k_l} - z^*\|^{1+\gamma})$),

(2) $\|H^{\rm nor}(z_{k_l} + h_{k_l 0}, \mu_{k_l})\|^2 = o(\|H^{\rm nor}(z_{k_l}, \mu_{k_l})\|^2)$, and

(3) $G^{\rm nor}(z_{k_l}+h_{k_l0}) = o(\|G^{\rm nor}(z_{k_l})\|)$

whenever $\mathbf{J}_z H^{\rm nor}(z_{k_l},\mu_{k_l})$ is nonsingular. Assuming that $\{\mathbf{J}_z H^{\rm nor}(z_{k_l},\mu_{k_l})\}$ is eventually uniformly nonsingular, it follows from the local Lipschitz continuity of $H^{\rm nor}$ at $(x^*,0)$ and the first condition that $\|H^{\rm nor}(z_{k_l},\mu_{k_l})\| = \|\mathbf{J}_z H^{\rm nor}(z_{k_l},\mu_{k_l})h_{k_l0}\| = O(\|h_{k_l0}\|)$ converges to 0; whence

$$\begin{aligned}\rho_1 \|h_{k_l0}\|^{\rho_2} &= \rho_1 \left\|\mathbf{J}_z H^{\rm nor}(z_{k_l},\mu_{k_l})^{-1} H^{\rm nor}(z_{k_l},\mu_{k_l})\right\|^{\rho_2}\\ &= O(\|H^{\rm nor}(z_{k_l},\mu_{k_l})\|^{\rho_2})\\ &= o(\|H^{\rm nor}(z_{k_l},\mu_{k_l})\|^2)\\ &= o(-\langle H^{\rm nor}(z_{k_l},\mu_{k_l}), \mathbf{J}_z H^{\rm nor}(z_{k_l},\mu_{k_l})h_{k_l0}\rangle)\end{aligned}$$

means that $d_{k_l0} = h_{k_l0}$ for all sufficiently large l. The second conclusion then implies that the step size λ_{k_l0} is eventually always 1, while the third conclusion implies that $\|G^{\rm nor}(z_{k_l}+d_{k_l0})\| \le \frac{1}{2}\|G^{\rm nor}(z_{k_l})\|$ eventually always hold. Thus the full Newton step is eventually always taken, and we can follow the same argument as in the proof of Theorem 4.1 to deduce superlinear (respectively, order-$(1+\gamma)$) convergence. □

5. Barriers with definable gradients

A twice-differentiable barrier $f : int(X) \to \mathbb{R}$ with a finite barrier parameter and a definable gradient map is necessary to employ Theorems 4.1, 4.2 and 4.3 to deduce the superlinear convergence of, respectively, Algorithms 4.1, 4.2 and 4.3 when solving either the natural map or normal map equation. For every closed convex set X that is proper (i.e., has nonempty interior and does not contain any affine subspace), it is well-known that the universal barrier is a twice-differentiable barrier with a finite barrier parameter; see [21, Theorem 2.5.1]. The *universal barrier* of X is

$$f : x \in int(X) \mapsto \log(vol(X^\sharp(x))),$$

where

$$X^\sharp(x) = \{y \in \mathbb{E} : \forall z \in X\ \langle y, z-x\rangle \le 1\}$$

is the polar set of X at x, and $vol(\cdot)$ denotes the Lebesgue measure on $\mathbb{E}$.

It remains to check the definability of the gradient of the universal barrier. We shall use a recent result of Kaiser [14]—which states that the Lebesgue measure is $\mathbb{R}_{\rm an}^{\mathbb{R}_{\rm alg}}$-compatible—to prove the definability of ∇f when f is the universal barrier of a proper closed convex set X definable in $\mathbb{R}_{\rm an}^{\mathbb{R}_{\rm alg}}$.

Given an o-minimal structure $\mathcal{O}$, a Borel measure λ on $\mathbb{R}^n$ is said to be *$\mathcal{O}$-compatible*, if there exists an o-minimal expansion $\mathcal{O}^*$ of $\mathcal{O}$ such that for every $A \subseteq \mathbb{R}^m \times \mathbb{R}^n$ definable in $\mathcal{O}$, the set

$$\{(x, \lambda(A_x)) : \lambda(A_x) < \infty\}$$

is definable in $\mathcal{O}^*$, where A_x denotes the set $\{y \in \mathbb{R}^n : (x,y) \in A\}$. We call $\mathcal{O}^*$ an *$\mathcal{O}$-measuring o-minimal structure of* λ. The measure λ is said to be *strongly $\mathcal{O}$-compatible* if the projection of the above set on the x-component is definable in $\mathcal{O}$.

Theorem 5.1 (Theorem 1.9 of [14])**.** *The Lebesgue measure vol on $\mathbb{R}^n$ is strongly $\mathbb{R}_{\rm an}^{\mathbb{R}_{\rm alg}}$-compatible, and $\mathbb{R}_{\rm an,exp}$ is a $\mathbb{R}_{\rm an}^{\mathbb{R}_{\rm alg}}$-measuring o-minimal structure of vol.*

Theorem 5.2. *The universal barrier f of a proper closed convex set $X \in \mathbb{R}^m$ that is definable in the o-minimal structure $\mathbb{R}_{\rm an}^{\mathbb{R}_{\rm alg}}$ has a gradient map ∇f that is definable in the o-minimal expansion $\mathbb{R}_{\rm an,exp}$.*

Proof. Let A be the set $\{(x,y) : x \in int(X),\ y \in X^\sharp(x)\} \cup \{(x,y) : x \notin int(X)\}$. Since X is definable in $\mathbb{R}_{\text{an}}^{\mathbb{R}_{\text{alg}}}$, so are $X^\sharp(x)$ and A. Moreover, for every $x \in \mathbb{R}^m$,

$$\{y \in \mathbb{R}^n : (x,y) \in A\} = \begin{cases} X^\sharp(x) & \text{if } x \in int(X), \\ \mathbb{R}^n & \text{otherwise.} \end{cases}$$

Therefore, the function $g : x \in int(X) \mapsto vol(X^\sharp(x))$, whose graph is

$$\{(x, vol(A_x)) : vol(A_x) < \infty\},$$

is definable in $\mathbb{R}_{\text{an,exp}}$ by Theorem 5.1. Thus the gradient map $\nabla f = \frac{\nabla g}{g}$ of the universal barrier of X is definable in $\mathbb{R}_{\text{an,exp}}$. □

We conclude this paper by mentioning a conjecture of Kaiser [14]:

> The Lebesgue measure in any arity is $\mathcal{O}$-compatible for every o-minimal structure $\mathcal{O}$.

If this conjecture is true, then we can trivially extend the above proof to deduce that the universal barrier of any definable proper closed convex set always have a gradient map definable in some o-minimal expansion. This will, of course, allows us to apply the main results in this paper to variational inequalities on any definable proper closed convex set.

References

[1] D. P. Bertsekas. *Nonlinear Programming.* Athena Scientific, Belmont, Massachusetts, 1999.

[2] J. Bochnak, M. Coste, and M.-F. Roy. *Real Algebraic Geometry.* Springer, Berlin Heidelberg, 1998.

[3] J. Bolte, A. Daniilidis, and A. Lewis. Tame functions are semismooth. *Math. Program.*, 117(1-2):5–19, 2009.

[4] X. Chen, L. Qi, and D. Sun. Global and superlinear convergence of the smoothing Newton method and its application to general box constrained variational inequalities. *Math. Comp.*, 67:519–540, 1998.

[5] X. Chen, D. Sun, and J. Sun. Complementarity functions and numerical experiments on some smoothing Newton methods for second-order-cone complementarity problems. *Comput. Optim. Appl.*, 25(1-3):39–56, 2003.

[6] C. B. Chua. A T-algebraic approach to primal-dual interior-point algorithms. *SIAM J. Optim.*, 20(1):503–523, 2009.

[7] C. B. Chua and Z. Li. A barrier-based smoothing proximal point algorithm for NCPs over closed convex cones. *SIAM J. Optim.*, 23(2):745–769, 2013.

[8] F. H. Clarke. *Optimization and Nonsmooth Analysis.* Wiley, New York, 1983.

[9] M. Costé. An introduction to o-minimal geometry. RAAG Notes, Institut de Recherche Mathématiques de Rennes, 1999.

[10] F. Facchinei and J.-S. Pang. *Finite-Dimensional Variational Inequalities and Complementarity Problems, Volume I.* Springer-Verlag, New York Berlin Heidelberg, 2003.

[11] A. Fischer. Solution of monotone complementarity problems with locally Lipschitzian functions. *Math. Program.*, 76(3):513–532, 1997.

[12] A. A. Goldstein and J. F. Price. An effective algorithm for minimization. *Numer. Math.*, 10:184–189, 1967.

[13] S. Hayashi, N. Yamashita, and M. Fukushima. A combined smoothing and regularization method for monotone second-order cone complementarity problems. *SIAM J. Optim.*, 15(2):593–615, 2005.

[14] T. Kaiser. First order tameness of measures. *Ann. Pure Appl. Logic*, 163(12):1903–1927, 2012.

[15] C. Kanzow and H. Pieper. Jacobian smoothing methods for nonlinear complementarity problems. *SIAM J. Optim.*, 9(2):342–373, 1999.

[16] M. Kojima, Y. Kurita, and S. Mizuno. Large-step interior point algorithms for linear complementarity problems. *SIAM J. Optim.*, 3(2):398–412, 1993.

[17] Z. Q. Lou and P. Tseng. A new class of merit functions for the nonlinear complementarity problem. In M. C. Ferris and J.-S. Pang, editors, *Complementarity and Variational Problems: State of the Art*, pages 204–225. SIAM Publications, Philadelphia, 1997.

[18] O. L. Mangasarian. Equivalence of the complementarity problem to a system of nonlinear equations. *SIAM J. Appl. Math.*, 31(1):89–92, 1976.
[19] R. Mifflin. Semismooth and semiconvex functions in constrained optimization. *SIAM J. Control Optim.*, 15(6):959–972, 1977.
[20] C. Miller. Expansions of the real field with power functions. *Ann. Pure Appl. Logic*, 68(1):79–94, 1994.
[21] Yu. E. Nesterov and A. S. Nemirovski. *Interior Point Polynomial Algorithms in Convex Programming*. SIAM Stud. Appl. Math. SIAM Publication, Philadelphia, PA, USA, 1994.
[22] L. Qi and D. Sun. Smoothing functions and smoothing Newton method for complementarity and variational inequality problems. *J. Optim. Theory Appl.*, 113(1):121–147, 2002.
[23] L. Qi and J. Sun. A nonsmooth version of Newton's method. *Math. Program.*, 58(1-3):353–367, 1993.
[24] J. Renegar. Hyperbolic programs, and their derivative relaxations. *Found. Comput. Math.*, 6(1):59–79, 2006.
[25] C. Steinhorn. Tame topology and o-minimal structures. In *Computational Aspects of General Equilibrium Theory*, volume 604 of *Lecture Notes in Economics and Mathematical Systems*, pages 165–191. Springer Berlin Heidelberg, 2008.
[26] D. Sun and L. Qi. Solving variational inequality problems via smoothing-nonsmooth reformulations. *J. Comput. Appl. Math.*, 129(1-2):37–62, 2001.
[27] D. Sun and J. Sun. Semismooth matrix-valued functions. *Math. Oper. Res.*, 27(1):150–169, 2002.
[28] L. van den Dries. A generalization of the Tarski-Seidenberg theorem, and some nondefinability results. *Bull. Amer. Math. Soc. (N.S.)*, 15(2):189–193, 1986.
[29] L. van den Dries. *Tame Topology and O-minimal Structures*. Cambridge University Press, May 1998.
[30] L. van den Dries, A. Macintyre, and D. Marker. The elementary theory of restricted analytic functions with exponentiation. *Ann. of Math.*, 140(1):183–205, 1994.
[31] L. van den Dries and C. Miller. On the real exponential field with restricted analytic functions. *Israel J. Math.*, 85(1-3):19–56, 1994.
[32] L. van den Dries and C. Miller. Geometric categories and o-minimal structures. *Duke Math. J.*, 84(2):497–540, 1996.
[33] È. B. Vinberg. The theory of convex homogeneous cones. *Trans. Moscow Math. Soc.*, 12:340–403, 1963.
[34] S. Wright. An infeasible-interior-point algorithm for linear complementarity problems. *Math. Program.*, 67(1-3):29–51, 1994.

(Chek Beng Chua) (Corresponding author) Division of Mathematical Sciences, School of Physical & Mathematical Sciences, Nanyang Technological University, Singapore 637371, Singapore
E-mail address: cbchua@ntu.edu.sg

(Le Thi Khanh Hien) Division of Mathematical Sciences, School of Physical & Mathematical Sciences, Nanyang Technological University, Singapore 637371, Singapore
E-mail address: TKLE1@e.ntu.edu.sg