

A barrier-based smoothing proximal point algorithm for NCPs over closed convex cones

CHEK BENG CHUA AND ZHEN LI

ABSTRACT. We present a new barrier-based method of constructing smoothing approximations for the Euclidean projector onto closed convex cones. These smoothing approximations are used in a smoothing proximal point algorithm to solve monotone nonlinear complementarity problems (NCPs) over a convex cones via the normal map equation. The smoothing approximations allow for the solution of the smoothed normal map equations with Newton's method, and do not require additional analytical properties of the Euclidean projector. The use of proximal terms in the algorithm adds stability to the solution of the smoothed normal map equation, and avoids numerical issues due to ill-conditioning at iterates near the boundary of the cones. We prove a sufficient condition on the barrier used that guarantees the convergence of the algorithm to a solution of the NCP. The sufficient condition is satisfied by all logarithmically homogeneous barriers. Preliminary numerical tests on semidefinite programming problems (SDPs) shows that our algorithm is comparable with the Newton-CG augmented Lagrangian algorithm (SDPNAL) proposed in [X. Y. Zhao, D. Sun, and K.-C. Toh, SIAM J. Optim. **20** (2010), 1737–1765].

1. INTRODUCTION

We consider the *nonlinear complementarity problem (NCP)*.

Given a continuous nonlinear map $F : \mathbb{E} \rightarrow \mathbb{E}$ on a finite dimensional Euclidean space \mathbb{E} with inner product $\langle \cdot, \cdot \rangle$, and a closed convex cone $K \subset \mathbb{E}$ with nonempty interior, find $x \in \mathbb{E}$ that satisfies

$$x \in K, \quad F(x) \in K^\sharp, \quad \text{and} \quad \langle x, F(x) \rangle = 0,$$

where K^\sharp denotes the dual cone of K .

We denote this problem by $\text{NCP}_K(F)$. When F is affine, we call the problem a *linear complementarity problem*.

In general, the KKT conditions for constrained nonlinear programs can be formulated as nonlinear complementarity problems [16]. Nonlinear complementarity problems also provide a general setting for various equilibrium problems, such as the computation of equilibria in finite games [41], the solution of spatial price equilibrium problems [42], and problems in option pricing [28]. Recently there has been much research in studying symmetric cone complementarity problems; see, e.g., [21], [22], [23], [34], [36], [45], [56] and [57].

While some solution methods, such as interior-point methods, solve the NCP directly (see, e.g., [31], [55], [59], [62], [64] and [65]), many other methods solve equivalent non-smooth equation reformulations of the problem. For instance, [32] uses a Hadamard product reformulation, [18] and [39] considers various C-functions in their reformulations,

Date: First draft, 1st May 2012; revised, 1st December 2012.

2010 Mathematics Subject Classification. 90C33, 90C55, 65K05 .

Key words and phrases. nonlinear complementarity problem, smoothing approximation, proximal point algorithm, normal map equation.

[4], [10], [26] and [29] uses a class of path-following algorithms known as non-interior continuation methods, and [17], [30], [38], [60] and [63] employs merit functions in solving reformulations of nonlinear complementarity problems.

The most general nonsmooth equation reformulations for the nonlinear linear complementarity problem are the normal map equation [43] and the natural map equation [14].

In this exposition, we consider the *normal map equation*

$$F(\text{Proj}_K(z)) + z - \text{Proj}_K(z) = 0, \quad (1)$$

where Proj_K denotes the Euclidean projection onto K ; i.e.,

$$\text{Proj}_K : z \in \mathbb{E} \mapsto \arg \min_{x \in K} \{\|z - x\|\},$$

where $\|\cdot\|$ denotes the norm induced by $\langle \cdot, \cdot \rangle$. Every solution z of the normal map equation gives a solution $x = \text{Proj}_K(z)$ of $\text{NCP}_K(F)$; see, e.g., [16, Section 1.5.2].

The normal map equation is typically solved with nonsmooth Newton's methods (see, e.g., [20] and [24]), semismooth Newton's methods (see, e.g., [13] and [17]), and smoothing Newton's methods (see, e.g., [6], [7], [8], [9], [11], [12], [19], [25], [27], [33], [44], [46], [47] and [48]). Existing nonsmooth and semismooth Newton's methods depend on the Bouligand-differentiability or the semismoothness of the nonsmooth reformulations, and is thus restricted to cones whose Euclidean projector possess the same analytic properties. On the other hand, solution methods that employs smoothing techniques avoid this constraint. However, a pre-requisite for employing smoothing techniques is the knowledge of a smoothing approximation of the Euclidean projector with computable derivatives.

For semidefinite programming, it was observed by Wang, Sun and Toh [61] that the addition of a log-determinant term to the objective function serves as a smoothing of the Euclidean projector. This connection between logarithmic barriers and smoothing approximations is the foundation of non-interior continuation methods for linear and semidefinite complementarity problems [4, 10, 26, 29]. In this paper, we generalize this idea and present a novel barrier-based method to construct smoothing approximations $p(\cdot, \mu)$, $\mu \geq 0$, of the Euclidean projector Proj_K (see Section 3 for its definition) of an arbitrary closed convex cone K , whose derivatives depend on those of the barrier used. This smoothing approximation allows us to write the normal map equation as the following auxiliary equation

$$\begin{pmatrix} F(p(z, \mu_+)) + z - p(z, \mu_+) \\ \mu \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad (2)$$

where μ_+ denotes the positive part $\max\{0, \mu\}$. The first component of this system is called the *smoothed normal map equation*.

For stability, we choose to employ the proximal point algorithm when solving the auxiliary equation (2). With appropriate control on the parameter c_k (which will be briefly described), we are able to avoid numerical issues due to ill-conditioning at primal and dual iterates near the boundary of their respective cones. The proximal point method was introduced by Martinet [40], and generalized by Rockafellar [51] to maps over Hilbert spaces. The algorithm can be used to solve problems such as the minimization of lower semicontinuous proper convex functions, and minimax problems [52]. It has resulted in many new algorithms in recent papers on several classes of optimization problems; see, e.g., [35], [61] and [67].

Briefly the proximal point algorithm finds a zero of a set-valued map $T : \mathbb{H} \rightrightarrows \mathbb{H}$ on a Hilbert space \mathbb{H} with a sequence of approximate zeros x_k of the map

$$x \mapsto T(x) + \frac{c_k^{-1}}{2}(x - x_{k-1}),$$

where $\{c_k\}$ is a sequence of positive real numbers. When T has a zero, Rockafellar [51, Theorem 1] proved that the sequence of approximate zeros converges to a zero of T under the assumptions that T is maximal monotone and $\{c_k^{-1}\}$ is bounded. When $\{c_k^{-1}\}$ is further bounded away from 0, we note that the above map is strongly monotone with modulus at least $\inf\{c_k^{-1}\}$, thus providing stability to the algorithm. As the proximal point algorithm can only be used to find zeros of maximal monotone maps, this approach can only solve the auxiliary equation of a limited class of nonlinear complementarity problems: it can only be applied directly when the *normal map*

$$F^{\text{nor}} : z \in \mathbb{E} \mapsto F(\text{Proj}_K(z)) + z - \text{Proj}_K(z)$$

is maximal monotone. In [66], Zhao and Li showed that when F is *co-coercive* with some constant $\alpha > 0$, that is,

$$\langle F(x) - F(y), x - y \rangle \geq \alpha \|F(x) - F(y)\|^2, \quad \forall x, y \in \mathbb{E},$$

the normal map is monotone, and subsequently maximal monotone when F is further continuous. We remark that being co-coercive is strictly stronger than being monotone.

In order to solve a more general class of nonlinear complementarity problem, we generalize the proximal point algorithm by using a nonlinear map $R : \mathbb{H} \rightarrow \mathbb{H}$ in the proximal term to get

$$z \mapsto T(z) + c_k^{-1}(R(z) - R(z_{k-1})).$$

With a simple change of variable $z \mapsto R^{-1}(z)$, the convergence of the proximal point algorithm translates to a similar conclusion for this variant: the sequence $\{R(z_k)\}$ converges to a zero of TR^{-1} under the assumptions that TR^{-1} is maximal monotone and $\{c_k^{-1}\}$ is bounded.

The idea of replacing the proximal term with nonlinear terms is not new. For example, Bregman functions [2, 15, 58] and logarithmic-quadratic terms [1] were studied. The main difference between these approaches and our approach is the assumption on T : while we require TR^{-1} to be maximal monotone, these other approaches require T to be maximal monotone. We remark that unless $R(\cdot) - R(0)$ is a multiple of the identity map, these two assumptions are incomparable¹; i.e., neither maximal monotonicity assumption implies the other.

For the purpose of solving the auxiliary equation, we use the nonlinear map $R : (z, \mu) \mapsto (p(z, \mu_+), \mu_+)$ in the proximal term to find zero of the *auxiliary function*

$$T : (z, \mu) \mapsto (F(p(z, \mu_+)) + z - p(z, \mu_+), \mu).$$

Under suitable conditions (see Proposition 4.2) on the barrier used to construct the smoothing approximation, we show that the maximal monotonicity of the composition

¹If $R(\cdot) - R(0)$ is a not multiple of the identity map, then there exists $z \in \mathbb{E}$ with $\langle w_z - w_0, z \rangle < \|w_z - w_0\| \|z\|$ for some $(z, w_z), (0, w_0) \in \mathcal{G}(R)$. Therefore, for the maximal monotone map $T : x \mapsto x + \alpha \langle v, x \rangle v$, where $\alpha \geq 0$ is arbitrary and $v = w_z - w_0 - \frac{\|w_z - w_0\|}{\|z\|} z$, we have $(z, Tw_z), (0, Tw_0) \in \mathcal{G}(TR)$ and

$$\langle z - 0, Tw_z - Tw_0 \rangle = \langle z, w_z - w_0 \rangle + \alpha \langle v, w_z - w_0 \rangle \langle z, v \rangle$$

is negative for all α sufficiently large, since

$$\langle v, w_z - w_0 \rangle = \left\langle w_z - w_0 - \frac{\|w_z - w_0\|}{\|z\|} z, w_z - w_0 \right\rangle = \|w_z - w_0\|^2 - \frac{\|w_z - w_0\|}{\|z\|} \langle z, w_z - w_0 \rangle > 0$$

and

$$\langle z, v \rangle = \left\langle z, w_z - w_0 - \frac{\|w_z - w_0\|}{\|z\|} z \right\rangle = \langle z, w_z - w_0 \rangle - \|w_z - w_0\| \|z\| < 0.$$

Thus TR is not monotone while T is maximal monotone. Similarly, there is a maximal monotone map T' with $T'R^{-1}$ not monotone; this gives a nonmonotone $T'R^{-1}$ with $(T'R^{-1})R = T'$ maximal monotone.

TR^{-1} is implied by the monotonicity of F .² Thus our approach only requires the monotonicity of F , and not that of the normal map F^{nor} .

We further modify the proximal point algorithm slightly (see Algorithm 4.1) to allow for different values of c_k between the two components of the auxiliary function (or equivalently, between two components of the nonlinear proximal term R). Subsequently we adapt the proof of Rockafellar [51, Theorem 1] to show that the algorithm generates a sequence of iterates $\{(z_k, \mu_k)\}$ with $\{p(z_k, \mu_k)\}$ converging to a solution of $\text{NCP}_K(F)$ under the assumption of monotonicity of F ; see Theorem 2.3.

In the setting of symmetric cone complementarity problems, where K is a *symmetric cone*, we show when a subclass of smoothing functions proposed by Chen and Mangasarian [7] for the nonnegative real line \mathbb{R}_+ (namely those generated by symmetric density functions with infinite support) is extended to smoothing approximations for symmetric cones, these Chen-Mangasarian-type smoothing approximations are special cases of our barrier-based smoothing approximations³. This subclass of smoothing approximations was successfully used in many smoothing methods for solving both linear and nonlinear complementarity problems, and more generally variational inequalities; see, e.g., [3], [5], [6], [12] and [48]. In particular, [6] and [12] employ the Chen-Mangasarian smoothing functions to solve the nonlinear complementarity problems via the normal map equation.

The paper is organized as follows. In Section 2, we briefly introduce the proximal point algorithm and relevant results on maximal monotone operators, and a variant of the proximal point algorithm used in this paper. This is followed with a newly proposed barrier-based method of constructing smoothing approximations of Euclidean projectors over closed convex cones in Section 3. In Section 4, we combine the smoothing technique and our variant of the proximal point algorithm to solve nonlinear complementarity problems over closed convex cones. We further prove the convergence of the algorithm under the assumption that F is monotone and a sufficient condition on the barrier used in the construction of the smoothing approximation. In Section 5, we demonstrate that the class of Chen-Mangasarian smoothing approximations fits into the framework of our barrier-based smoothing approximations. In Section 6, we report the results of preliminary numerical comparisons between an implementation our algorithm and the *Newton-CG augmented Lagrangian method (SDPNAL)* proposed by Zhao, Sun and Toh [67], by applying them to several classes of SDPs in SDPLIB. Since SDPNAL is also based on the proximal point algorithm, we find the comparison appropriate.

2. PROXIMAL POINT ALGORITHM

The proximal point algorithm [51] is an iterative algorithm for finding a zero of a *maximal monotone* set-valued map $T : \mathbb{H} \rightrightarrows \mathbb{H}$ over a Hilbert space \mathbb{H} . Each iteration of the proximal point algorithm finds an approximate value of a *proximal mapping* associated with T . When the approximates are sufficiently good (see [51, Theorem 1]), it can be proved that they converge to a zero of the maximal monotone operator T . In this section, we give basic definitions and results on the proximal point algorithm that are necessary for this paper. Interested readers are referred to Rockafellar's classic papers [51] and [52] for a more comprehensive introduction.

²When F is continuous with domain K , one can easily check that the monotonicity of F is also a necessary condition for the maximal monotonicity of the composition TR^{-1} .

³In the special case of $K = \mathbb{R}_+$, our smoothing approximation coincides with this subclass.

Definition 2.1. A set-valued map $T : \mathbb{H} \rightrightarrows \mathbb{H}$ over a Hilbert space \mathbb{H} with inner product $\langle \cdot, \cdot \rangle_{\mathbb{H}}$ is said to be a *monotone operator* if

$$\forall (z, w), (z', w') \in \mathcal{G}(T), \quad \langle z - z', w - w' \rangle_{\mathbb{H}} \geq 0,$$

where $\mathcal{G}(T) := \{(z, w) \in \mathbb{H}^2 : w \in T(z)\}$ denotes the *graph of T* . A monotone operator is *maximal monotone* if its graph is not contained in the graph of another monotone operator. Equivalently, a monotone operator is maximal monotone if

$$\forall (z, w) \in \mathbb{H}^2, \quad \left(\inf_{(z', w') \in \mathcal{G}(T)} \langle z - z', w - w' \rangle_{\mathbb{H}} \geq 0 \implies (z, w) \in \mathcal{G}(T) \right).$$

Observe that (maximal) monotonicity is preserved by positive scalings; i.e., T is (maximal) monotone if and only if cT is (maximal) monotone, where $c > 0$.

In 1962, Minty gave the following characterization of maximal monotonicity, upon which Rockafellar built the proximal point algorithm.

Theorem 2.1 (Minty's characterization of maximal monotonicity; see [43]). *For each $c > 0$, a monotone operator $T : \mathbb{H} \rightrightarrows \mathbb{H}$ is maximal monotone if and only if the range of $cT + I$ is \mathbb{H} ; or equivalently, $\mathbf{dom}((cT + I)^{-1}) = \mathbb{H}$, where $\mathbf{dom}((cT + I)^{-1})$ denotes the domain $\{x \in \mathbb{H} : (cT + I)^{-1}(x) \neq \emptyset\}$ of $(cT + I)^{-1}$.*

The set-valued map $(cT + I)^{-1}$ is called the *proximal mapping* associated with cT . It plays a key role in the proof of the convergence of the proximal point algorithm. When cT is monotone, $cT + I$ is strongly monotone with modulus 1; i.e., for all $(x, z), (x', z') \in \mathcal{G}(cT + I)$,

$$\langle z - z', x - x' \rangle_{\mathbb{H}} \geq \|x - x'\|_{\mathbb{H}}^2,$$

where $\|\cdot\|_{\mathbb{H}}$ is the norm induced by $\langle \cdot, \cdot \rangle_{\mathbb{H}}$. This leads to the nonexpansiveness of proximal mappings of monotone operators. Hence proximal mappings of monotone operators are single-valued.

Theorem 2.2 (Theorem 12.12 of [53]). *If $T : \mathbb{H} \rightrightarrows \mathbb{H}$ is a monotone operator, then the proximal mapping $P = (cT + I)^{-1}$ is monotone and nonexpansive; i.e., for all $(x, z), (x', z') \in \mathcal{G}(P)$, $\|z - z'\|_{\mathbb{H}} \leq \|x - x'\|_{\mathbb{H}}$.*

We now introduce a slight variant of the proximal point algorithm to find a zero of a set-valued map $T : \mathbb{H} \rightrightarrows \mathbb{H}$ on a Hilbert space \mathbb{H} , and demonstrate its convergence.

Algorithm 2.1 (Proximal point algorithm, a variant).

Choose a map $R : \mathbb{H} \rightarrow \mathbb{H}$ and a sequence of bijective linear maps $\{C_k : \mathbb{H} \rightarrow \mathbb{H}\}_{k=0}^{\infty}$. Pick $x_0 \in \mathbb{H}$. For $k = 0, 1, \dots$, find an approximate zero x_{k+1} of the set-valued map

$$\varphi_k : x \in \mathbb{H} \mapsto T(x) + C_k^{-1}(R(x) - R(x_k)) \quad (3)$$

until $0 \in T(x_k)$.

Remark 2.1. In Rockafellar's proximal point algorithm, R is the identity map and the C_k 's are positive multiples of the identity map.

The following convergence theorem is the analogue to that of Rockafellar's proximal point algorithm, and its proof follows that of [51, Theorem 1]. For the sake of completeness, we give the proof in Appendix A.

Theorem 2.3. *If $T : \mathbb{H} \rightrightarrows \mathbb{H}$ has at least one zero, and in Algorithm 2.1, $C_k T R^{-1}$ is a maximal monotone operator for each k , $\{C_k^{-1}\}$ is bounded, and the infinite sequence $\{x_k\}$ generated satisfies*

$$\sum_{k=0}^{\infty} \|C_k \varphi_k(x_{k+1})\|_{\mathbb{H}} < \infty, \quad (4)$$

then the sequence $\{R(x_k)\}$ converges weakly to a zero of TR^{-1} .

3. SMOOTHING APPROXIMATION

A *smoothing approximation* of the Euclidean projector Proj_K is a continuously differentiable map $p : \mathbb{E} \times \mathbb{R}_{++} \rightarrow \mathbb{E}$, parameterized by $\mu \in \mathbb{R}_{++}$, such that p converges point-wise to Proj_K as $\mu \rightarrow 0^+$; i.e.,

$$\forall z \in \mathbb{E}, \lim_{\mu \rightarrow 0^+} p(z, \mu) = \text{Proj}_K(z).$$

If the convergence is uniform, we say that p is a *uniform smoothing approximation*. In this section, we demonstrate a general method of constructing smoothing approximations of Euclidean projectors of closed convex cones.

Consider a convex twice-differentiable barrier $f : \mathbf{int}(K) \rightarrow \mathbb{R}$ with positive definite Hessians $\nabla^2 f(x)$. Being a barrier means that $f(x_k) \rightarrow \infty$ for any sequence $\{x_k\} \subset \mathbf{int}(K)$ converging to a point on the boundary of K . Together with its convexity, we deduce that for any sequence $\{x_k\} \subset \mathbf{int}(K)$ converging to a point on the boundary of K ,

$$\|\nabla f(x_k)\| \|x_k - x_1\| \geq \langle \nabla f(x_k), x_k - x_1 \rangle \geq f(x_k) - f(x_1) \rightarrow \infty,$$

whence $\|\nabla f(x_k)\| \rightarrow \infty$.

We now show that for each $\mu > 0$, the map $\varrho_\mu : x \mapsto x + \mu \nabla f(x/\mu)$ is a bijection between $\mathbf{int}(K)$ and \mathbb{E} , and subsequently deduce sufficient conditions for the inverse maps to produce a smoothing approximation $(z, \mu) \in \mathbb{E} \times \mathbb{R}_{++} \mapsto \varrho_\mu^{-1}(z)$ of the Euclidean projector Proj_K .

We first establish the maximal monotonicity of ϱ_μ via Löhne's characterization.

Theorem 3.1 (Löhne's characterization of maximal monotonicity; see [37]). *A set-valued map $T : \mathbb{H} \rightrightarrows \mathbb{H}$ is maximal monotone if and only if the following are satisfied.*

- (i) T is monotone.
- (ii) T has a nearly convex domain (i.e., the closure $\mathbf{cl}(\mathbf{dom}(T))$ is convex).
- (iii) The values of T are convex.
- (iv) The recession cone of $T(x)$ equals the normal cone to $\mathbf{cl}(\mathbf{dom}(T))$ at every $x \in \mathbf{dom}(T)$.
- (v) The graph of T is closed.

Proposition 3.1. *For each $\mu > 0$, and each continuous monotone map $F : \mathbb{H} \rightarrow \mathbb{H}$ with $\mathbf{dom}(F) \supseteq K$, the set-valued map $x \mapsto F(x) + \mu \nabla f(x/\mu)$ is maximal monotone.*

Proof. We shall use Löhne's characterization to check that the map $H_\mu : x \mapsto F(x) + \mu \nabla f(x/\mu)$ is maximal monotone for each $\mu > 0$. Indeed,

- (i) H_μ is the sum of the monotone map F and the derivative map of the convex map $x \mapsto \mu^2 f(x/\mu)$, whence monotone;
- (ii) H_μ has the convex domain $\mathbf{int}(K)$;
- (iii) when $x \in \mathbf{dom}(F) \cap \mathbf{dom}(f) = \mathbf{int}(K)$, $H_\mu(x) = \{F(x) + \mu \nabla f(x/\mu)\}$ is a singleton, but otherwise $H_\mu(x)$ is the empty set, whence the values of H_μ are convex;
- (iv) the recession cone of H_μ at each x in its domain $\mathbf{int}(K)$ is the trivial subspace $\{0\}$, which agrees with the normal cone to $\mathbf{cl}(\mathbf{int}(K)) = K$ at x ;
- (v) the graph of H_μ is closed as F is continuous on K and $x \mapsto \nabla f(x/\mu)$ is continuous on $\mathbf{int}(K)$ with $\|\nabla f(x_k/\mu)\| \rightarrow \infty$ for any sequence $\{x_k\} \subset \mathbf{int}(K)$ converging to a point on the boundary of K .

□

As a corollary to the maximal monotonicity of $x \mapsto \mu \nabla f(x/\mu)$, we deduce the bijectiveness of ϱ_μ .

Proposition 3.2. *For each $\mu > 0$, the map ϱ_μ is a bijection between $\mathbf{int}(K)$ and \mathbb{E} .*

Proof. For each $\mu > 0$, ϱ_μ is strongly monotone, whence injective. From Proposition 3.1 with $F \equiv 0$, and Minty's characterization, we deduce that ϱ_μ is surjective. \square

Henceforth, for each $\mu > 0$, we denote inverse map of ϱ_μ by p_μ , and remark that p_μ is the proximal mapping associated with the monotone operator $x \mapsto \mu \nabla f(x/\mu)$. We further define

$$p : \mathbb{E} \times \mathbb{R}_{++} \rightarrow \mathbb{E} : (z, \mu) \mapsto p_\mu(z),$$

and remark that as $J\varrho_\mu(x) = I + \nabla^2 f(x/\mu)$ is nonsingular for each $x \in \mathbf{int}(K)$, we can deduce the continuous differentiability of p follows from the Implicit Function Theorem.

Sufficient and necessary conditions for p to be a (uniform) smoothing approximation of the Euclidean projector Proj_K are given in the following proposition, where we adopt the notation

$$\begin{aligned} \limsup_{\mu \rightarrow 0^+, y \rightarrow x} \mu \nabla f(y/\mu) &:= \bigcup_{\mu_k \downarrow 0, y_k \rightarrow x} \lim_{k \rightarrow \infty} \mu_k \nabla f(y_k/\mu_k) \\ &= \{d : \exists \mu_k \downarrow 0, \exists y_k \rightarrow x, \mu_k \nabla f(y_k/\mu_k) \rightarrow d\}. \end{aligned}$$

Proposition 3.3. *For the following statements, we have (a) \implies (b) \implies (c) \implies (d).*

- (a) *The function p is a uniform smoothing approximation of the Euclidean projector Proj_K .*
- (b) *For every $x \in \mathbb{E}$,*

$$\limsup_{\mu \rightarrow 0^+, y \rightarrow x} \mu \nabla f(y/\mu) \subseteq N_K(x).$$

- (c) *The function p is a smoothing approximation of the Euclidean projector Proj_K .*
- (d) *For every $x \in \mathbb{E}$,*

$$N_K(x) \subseteq \limsup_{\mu \rightarrow 0^+, y \rightarrow x} \mu \nabla f(y/\mu).$$

Proof. “(a) \implies (b)”: For arbitrary sequences $\mu_k \downarrow 0$ and $\mathbf{int}(K) \ni y_k \rightarrow x$ with $\mu_k \nabla f(y_k/\mu_k)$ convergent, say to limit d , consider the sequence $\{z_k := \varrho_{\mu_k}(y_k) = y_k + \mu_k \nabla f(y_k/\mu_k)\}$, which converges to $x + d$. By the continuity of $p(\cdot, \mu)$ for $\mu > 0$ and the hypothesis of uniform convergence of $p(\cdot, \mu)$ to Proj_K as $\mu \rightarrow 0^+$, we deduce that $p(z_k, \mu_k) \rightarrow \text{Proj}_K(x + d)$. Hence

$$x = \lim_{k \rightarrow \infty} y_k = \lim_{k \rightarrow \infty} p(z_k, \mu_k) = \text{Proj}_K(x + d)$$

shows that $d \in N_K(x)$.

“(b) \implies (c)”: Fix an $z \in \mathbb{E}$ and an arbitrary $e \in \mathbf{int}(K)$ and consider the bounded sequence $\{\varrho_{\mu_k}(\mu_k e) = \mu_k e + \mu_k \nabla f(e)\}$ with $\mu_k \rightarrow 0$. The proximal mapping p_{μ_k} is nonexpansive, hence

$$\|p(z, \mu_k)\| \leq \|p_{\mu_k}(z) - p_{\mu_k}(\varrho_{\mu_k}(\mu_k e))\| + \|\mu_k e\| \leq \|z - \varrho_{\mu_k}(\mu_k e)\| + \|\mu_k e\|.$$

This shows that $\{p(z, \mu_k)\}$ is bounded for any $\mu_k \rightarrow 0$. Therefore, it suffices to show that convergent $\{p(z, \mu_k)\}$ with $\mu_k \rightarrow 0$ has limit $\text{Proj}_K(z)$. Indeed, if $p(z, \mu_k) \rightarrow x$, then, with x_{μ_k} denoting $p(z, \mu_k)$ so that $z = x_{\mu_k} + \mu_k \nabla f(x_{\mu_k}/\mu_k)$, we have

$$z \in \limsup_{k \rightarrow \infty} (x_{\mu_k} + \mu_k \nabla f(x_{\mu_k}/\mu_k)) \subseteq x + N_K(x)$$

under the hypothesis (b), and consequently $x = \text{Proj}_K(z)$.

“(c) \implies (d)”: For arbitrary $x \in K$ and arbitrary $d \in N_K(x)$, consider the sequence $\{\mu_k := 1/k\}$, and the sequence $\{y_k := p(x + d, \mu_k)\}$. Then

$$\mu_k \nabla f(y_k/\mu_k) = x + d - y_k \xrightarrow{(c)} x + d - \text{Proj}_K(x + d) = d$$

shows that $d \in \limsup_{\mu \rightarrow 0^+, y \rightarrow x} \mu \nabla f(y/\mu)$. \square

When p is a smoothing approximation of Proj_K , we call p the *smoothing approximation derived from the barrier f* . The above sufficient condition for p to be a smoothing approximation is satisfied by ϑ -logarithmically homogeneous barriers; i.e., barriers f satisfying

$$\forall x \in \mathbf{int}(K), \forall t \in \mathbb{R}_{++}, f(tx) = f(x) - \vartheta \log t. \quad (5)$$

Corollary 3.1. *If f is a ϑ -logarithmically homogeneous barrier, then p is a smoothing approximation of the Euclidean projector Proj_K .*

Proof. We prove the corollary by establishing statement (b) of Proposition 3.3.

If f is a ϑ -logarithmically homogeneous barrier, then for any $x \in \mathbf{int}(K)$,

$$\langle x, -\nabla f(x) \rangle = \lim_{t \rightarrow 0^+} \frac{f(x) - f(x + tx)}{t} = \lim_{t \rightarrow 0^+} \frac{\vartheta \log(1 + t)}{t} = \vartheta, \quad (6)$$

and, in addition, for any $h \in \mathbf{int}(K)$, the convexity of f results in

$$\langle \nabla f(x), h \rangle \leq \lim_{t \rightarrow \infty} \frac{f(x + th) - f(x)}{t} = \lim_{t \rightarrow \infty} \frac{f(x/t + h) - \vartheta \log t - f(x)}{t} = 0. \quad (7)$$

Thus, for any $x \in K$, any sequences $\mu_k \rightarrow 0$ and $y_k \rightarrow x$ with $\{d_k := -\mu_k \nabla f(y_k/\mu_k)\}$ convergent, say with limit d , we have

$$d = \lim_{k \rightarrow \infty} -\mu_k \nabla f(y_k/\mu_k) \stackrel{(7)}{\in} K^\sharp,$$

and

$$0 \leq \langle x, d \rangle = \lim_{k \rightarrow \infty} \langle y_k, -\mu_k \nabla f(y_k/\mu_k) \rangle \stackrel{(6)}{=} \lim_{k \rightarrow \infty} \mu_k^2 \vartheta = 0,$$

whence $d \in N_K(x)$. \square

Example 3.1. *A 1-logarithmically homogeneous barrier for the positive real line \mathbb{R}_{++} is*

$$f : x \mapsto -\log x.$$

The smoothing approximation derived from this barrier is

$$p : (z, \mu) \mapsto \frac{z + \sqrt{z^2 + 4\mu^2}}{2},$$

which is precisely the CHKS map (see Example 5.2).

Example 3.2. *An n -logarithmically homogeneous barrier for the positive definite cone \mathbb{S}_{++}^n is*

$$f : X \mapsto -\log \det X.$$

Its gradient map is $\nabla f : X \mapsto -X^{-1}$ and the smoothing approximation derived from this barrier

$$p : (Z, \mu) \mapsto \frac{Z + \sqrt{Z^2 + 4\mu^2 I}}{2},$$

where I is the n -by- n identity matrix and \sqrt{X} denotes the unique $Y \in \mathbb{S}_{++}^n$ satisfying $Y^2 = X \in \mathbb{S}_{++}^n$, is an extension of the CHKS map.

Example 3.3. An r -logarithmically homogeneous barrier for a symmetric cone Ω of rank r is

$$f : x \mapsto -\log \det x,$$

where \det is the determinant function in a Euclidean Jordan algebra \mathfrak{J} whose cone of squares coincide with $K = \mathbf{cl}(\Omega)$. Its gradient map is $\nabla f : x \mapsto -x^{-1}$ and the smoothing approximation derived from this barrier

$$p : (z, \mu) \mapsto \frac{z + \sqrt{z^2 + 4\mu^2 e}}{2},$$

where e is the unit of \mathfrak{J} and \sqrt{x} denotes the unique $y \in \Omega$ satisfying $y^2 = x \in \Omega$, is another extension of the CHKS map.

4. SMOOTHING PROXIMAL POINT ALGORITHM

When p is a smoothing approximation of the Euclidean projector Proj_K , the normal map equation (1) can be approximated by the *smoothed normal map equation*

$$F(p(z, \mu)) + z - p(z, \mu) = 0, \quad (8)$$

so that for any positive sequence $\{\mu_k\}$ converging to 0, and any sequence of solutions $\{(z_k, \mu_k)\}$ of the smoothed normal map equation, we have that all limit points of $\{z_k\}$ solves the normal map equation. For convenience, we denote the *smoothed normal map* by

$$F_\mu^{\text{nor}} : z \in \mathbb{E} \mapsto F(p(z, \mu)) + z - p(z, \mu). \quad (9)$$

Moreover, using the convention $p_0 \equiv p(\cdot, 0) \equiv \text{Proj}_K$, the smoothed normal map F_μ^{nor} becomes the normal map F^{nor} when $\mu = 0$. For subsequent use, we define $p_\mu(z)$ to be the empty set when $\mu < 0$.

Therefore, we consider the application of the proximal point algorithm (Algorithm 2.1) to find a zero of the *auxiliary function*

$$T : (z, \mu) \in \mathbb{E} \times \mathbb{R} \mapsto (F_{\mu_+}^{\text{nor}}(z), \mu).$$

For convergence, we start from $(z_0, \mu_0) \in \mathbb{E} \times \mathbb{R}_{++}$, and pick linear maps C_k of the form $(z, \mu) \mapsto (c_k z, \gamma_k^{-1} \mu)$ for some positive real numbers c_k and γ_k , and the nonlinear map $R : (z, \mu) \mapsto (p(z, \mu_+), \mu_+)$. With these choices, the algorithm simplifies to the following.

Algorithm 4.1 (Smoothing proximal point algorithm).

Choose positive sequences $\{c_k\}_{k=0}^\infty$, $\{\gamma_k\}_{k=0}^\infty$. Pick $z_0 \in \mathbb{E}$ and $\mu_0 > 0$. For $k = 0, 1, \dots$, set

$$\mu_{k+1} = \frac{\gamma_k}{1 + \gamma_k} \mu_k,$$

and find an approximate zero z_{k+1} of the single-valued map

$$\varphi_k : z \in \mathbb{E} \mapsto F(p(z, \mu_{k+1})) + z - p(z, \mu_{k+1}) + c_k^{-1}(p(z, \mu_{k+1}) - p(z_k, \mu_k)).$$

To establish the convergence of the algorithm using Theorem 2.3, we would require the maximal monotonicity of

$$(x, \mu) \mapsto C_k T R^{-1}(x, \mu) = \begin{cases} (c_k(F(x) + p_\mu^{-1}(x) - x), \gamma_k^{-1} \mu) & \text{if } \mu > 0, \\ \{c_k(F(x) + \text{Proj}_K^{-1}(x) - x)\} \times \mathbb{R}_- & \text{if } \mu = 0 \end{cases}$$

for each k . For convenience, we define the set-valued maps

$$\varrho : \mathbb{E} \times \mathbb{R} \rightrightarrows \mathbb{E} : (x, \mu) \mapsto p_\mu^{-1}(x) = \begin{cases} x + \mu \nabla f(x/\mu) & \text{if } x \in \mathbf{int}(K) \text{ and } \mu > 0, \\ \text{Proj}_K^{-1}(x) & \text{if } x \in K \text{ and } \mu = 0, \\ \emptyset & \text{otherwise,} \end{cases}$$

and

$$\begin{aligned} H : \mathbb{E} \times \mathbb{R} \rightrightarrows \mathbb{E} : (x, \mu) &\mapsto F(x) + \varrho(x, \mu) - x \\ &= \begin{cases} F(x) + \mu \nabla f(x/\mu) & \text{if } x \in \mathbf{int}(K) \text{ and } \mu > 0, \\ F(x) + \text{Proj}_K^{-1}(x) - x & \text{if } x \in K \text{ and } \mu = 0, \\ \emptyset & \text{otherwise.} \end{cases} \end{aligned}$$

Proposition 4.1. *If $F : \mathbb{E} \rightarrow \mathbb{E}$ is continuous and monotone with $\mathbf{dom}(F) \supseteq K$, then for each $\mu \geq 0$, the set-valued map $x \mapsto H(x, \mu)$ is maximal monotone.*

Proof. When $\mu > 0$, we have proved in Proposition 3.1 that the map $H(\cdot, \mu)$ is maximal monotone under the hypotheses of the proposition. In the case $\mu = 0$, $\text{Proj}_K^{-1}(x) - x = \{z - x : \text{Proj}_K(z) = x\}$ is the normal cone of K at x , and it is known that the normal cone map (also called *normality map*) is monotone [49, 50]. Hence, $H(\cdot, 0)$ is the sum of two monotone operators, whence monotone, under hypotheses of the proposition. From here on, we can follow the proof of Proposition 3.1 to establish the maximal monotonicity of $H(\cdot, 0)$ via Löhne's characterization. Alternatively, we can use Minty's characterization since the monotonicity of F leads to the strong monotonicity of $F + I$, which in turn implies the global unique solvability of $\text{NCP}_K(F + I)$; i.e., $x \mapsto H(x, 0) + x$ is a bijection between K and \mathbb{E} (see, e.g., [16, Proposition 2.3.3]). \square

Proposition 4.2. *If $F : \mathbb{E} \rightarrow \mathbb{E}$ is continuous and monotone with $\mathbf{dom}(F) \supseteq K$, p is the smoothing approximation of the Euclidean projector Proj_K derived from the barrier f , and*

$$c\gamma \sup_{x \in \mathbf{int}(K)} \langle \nabla f(x) - \nabla^2 f(x)x, (\nabla^2 f(x))^{-1} \nabla f(x) - x \rangle \leq 4\omega, \quad (10)$$

for some $\omega, c, \gamma > 0$, then the set-valued map S defined by

$$(x, \mu) \in \mathbb{H}_\omega \mapsto \begin{cases} (cH(x, \mu), \gamma^{-1}\mu) & \text{if } \mu > 0, \\ \{cH(x, 0)\} \times \mathbb{R}_- & \text{if } \mu = 0 \end{cases} \quad (11)$$

is maximal monotone, where \mathbb{H}_ω is the Hilbert space $\mathbb{E} \times \mathbb{R}$ with inner product $\langle \cdot, \cdot \rangle_\omega : ((x, \mu), (x', \mu')) \mapsto \langle x, x' \rangle + \omega\mu\mu'$.

Proof. When F is continuous and monotone, the map $H(\cdot, \mu)$ is maximal monotone for each $\mu \geq 0$ by Proposition 4.1. Thus, by Minty's characterization, the map $H(\cdot, \mu) + c^{-1}I$ is surjective onto \mathbb{E} for each $\mu \geq 0$ and each $c > 0$. Hence, when $c > 0$ and $\gamma > 0$,

$$(S + I)^{-1}(z, \nu) = \left(H \left(\cdot, \frac{\gamma}{1 + \gamma} \nu_+ \right) + c^{-1}I \right)^{-1} (c^{-1}z) \times \left\{ \frac{\gamma}{1 + \gamma} \nu_+ \right\} \neq \emptyset$$

for each $(z, \nu) \in \mathbb{H}_\omega$. Hence $S + I$ is surjective onto \mathbb{H}_ω . Therefore, by Minty's characterization, S is maximal monotone if and only if it is monotone. It remains to prove monotonicity.

We first consider the restriction of S to $\mathbb{E} \times \mathbb{R}_{++} \subset \mathbb{H}_\omega$. This restriction of S is the sum of the monotone map $(x, \mu) \in H_\omega \mapsto (cF(x), 0)$ and the map

$$(x, \mu) \in H_\omega \mapsto \left(c(\varrho(x, \mu) - x), \frac{\mu}{\gamma} \right).$$

It suffices to show that the latter is monotone on $\mathbf{int}(K) \times \mathbb{R}_{++}$, or equivalently, that

$$(x, \mu) \in H_1 \mapsto \left(c(\varrho(x, \mu) - x), \omega \frac{\mu}{\gamma} \right)$$

is monotone on $\mathbf{int}(K) \times \mathbb{R}_{++}$. Since this map is single-valued and differentiable on $\mathbf{int}(K) \times \mathbb{R}_{++}$, its monotonicity is equivalent to the symmetric part of its Jacobian

$$\begin{aligned} & \begin{pmatrix} c(J_x \varrho(x, \mu) - I) & \frac{1}{2} c J_\mu \varrho(x, \mu) \\ \frac{1}{2} c J_\mu \varrho(x, \mu)^T & \frac{\omega}{\gamma} \end{pmatrix} \\ &= \begin{pmatrix} c \nabla^2 f(x') & \frac{1}{2} c (\nabla f(x') - \nabla^2 f(x') x') \\ \frac{1}{2} c (\nabla f(x') - \nabla^2 f(x') x')^T & \frac{\omega}{\gamma} \end{pmatrix} \end{aligned} \quad (12)$$

being positive semidefinite on $\mathbf{int}(K) \times \mathbb{R}_{++}$, where x' denotes x/μ . The Schur complement of $\frac{\omega}{\gamma}$ in the symmetric part of the Jacobian (12) is

$$\begin{aligned} & \frac{\omega}{\gamma} - \frac{c}{4} \langle \nabla f(x') - \nabla^2 f(x') x', (\nabla^2 f(x'))^{-1} (\nabla f(x') - \nabla^2 f(x') x') \rangle \\ &= \frac{\omega}{\gamma} - \frac{c}{4} \langle \nabla f(x') - \nabla^2 f(x') x', (\nabla^2 f(x'))^{-1} \nabla f(x') - x' \rangle. \end{aligned}$$

Thus the symmetric part of the Jacobian (12) is positive semidefinite under the hypotheses of the proposition.

We now consider S on the whole space \mathbb{H}_ω . We first establish that for each $x \in K$, and each $\tilde{x} \in cH(x, 0)$, there is a sequence $\mathbb{E} \times \mathbb{R}_{++} \ni (x_k, \mu_k) \rightarrow (x, 0)$ such that $cH(x_k, \mu_k) \rightarrow \tilde{x}$. Indeed, by statement (d) of Proposition 3.3, for the direction $c^{-1} \tilde{x} - F(x) \in N_K(x)$, we can find a sequence $\{(x_k, \mu_k)\} \subset \mathbf{int}(K) \times \mathbb{R}_{++}$ such that $(x_k, \mu_k) \rightarrow (x, 0)$ and $\mu_k \nabla f(x_k/\mu_k) \rightarrow c^{-1} \tilde{x} - F(x)$; hence, under the continuity of F ,

$$cH(x_k, \mu_k) = c(F(x_k) + \mu_k \nabla f(x_k/\mu_k)) \rightarrow c(F(x) + c^{-1} \tilde{x} - F(x)) = \tilde{x}.$$

Next, for each $(x, \mu), (x', \mu') \in \mathbf{dom}(S)$, each $(z, \nu) \in S(x, \mu)$ and each $(z', \nu') \in S(x', \mu')$, by taking sequences $\{(x_k, \mu_k)\}, \{(x'_k, \mu'_k)\} \subset \mathbf{int}(K) \times \mathbb{R}_{++}$ satisfying

$$\begin{aligned} (x_k, \mu_k) &\rightarrow (x, \mu), & (x'_k, \mu'_k) &\rightarrow (x', \mu'), \\ S(x_k, \mu_k) &\rightarrow (z, \nu_+), & S(x'_k, \mu'_k) &\rightarrow (z', \nu'_+), \end{aligned}$$

we can deduce from the monotonicity of S over $\mathbf{int}(K) \times \mathbb{R}_{++}$ that

$$\begin{aligned} 0 &\leq \lim_{k \rightarrow \infty} \langle (x_k, \mu_k) - (x'_k, \mu'_k), S(x_k, \mu_k) - S(x'_k, \mu'_k) \rangle_\omega \\ &= \langle (x, \mu) - (x', \mu'), (z, \nu_+) - (z', \nu'_+) \rangle_\omega \\ &= \langle (x, \mu) - (x', \mu'), (z, \nu) - (z', \nu') \rangle_\omega + \omega(\mu - \mu')(\nu_+ - \nu - \nu'_+ + \nu'). \end{aligned}$$

Since $\mu, \mu', \nu_+ - \nu, \nu'_+ - \nu' \geq 0$ with $\nu_+ - \nu > 0$ (resp., $\nu'_+ - \nu' > 0$) only when $\mu = 0$ (resp., $\mu' = 0$), it follows that

$$(\mu - \mu')(\nu_+ - \nu - \nu'_+ + \nu') = -\mu(\nu'_+ - \nu') - \mu'(\nu_+ - \nu) \leq 0.$$

Thus $\langle (x, \mu) - (x', \mu'), (z, \nu) - (z', \nu') \rangle_\omega \geq 0$, showing that S is monotone on \mathbb{H}_ω . \square

Remark 4.1. The inequality in the proposition is satisfied by a ϑ -logarithmically homogeneous barrier whenever $c\gamma\vartheta \leq \omega$. Indeed, if f is a ϑ -logarithmically homogeneous barrier, then differentiating the definition (5) of logarithmic homogeneity in x followed by in t results in $\nabla^2 f(x)x = -\nabla f(x)$ at $t = 1$, which, together with $\langle x, -\nabla f(x) \rangle = \vartheta$ as derived in the proof of Corollary 3.1, shows that

$$\langle \nabla f(x) - \nabla^2 f(x)x, (\nabla^2 f(x))^{-1} \nabla f(x) - x \rangle = 4\vartheta$$

for any $x \in \text{int}(K)$.

As a corollary of this proposition, we deduce the Lipschitz continuity of $p(z, \cdot)$ for each $z \in \mathbb{E}$.

Corollary 4.1. *If*

$$\sup_{x \in \text{int}(K)} \langle \nabla f(x) - \nabla^2 f(x)x, (\nabla^2 f(x))^{-1} \nabla f(x) - x \rangle < \infty,$$

then for any $z \in \mathbb{E}$, the function $\mu \in \mathbb{R}_{++} \mapsto p(z, \mu)$ is Lipschitz continuous. In this case, the Lipschitz constant is bounded by the above supremum, and $p(\cdot, \mu)$ converges uniformly as $\mu \rightarrow 0^+$.

Proof. We take $F \equiv 0$, $c = \gamma = 1$ and

$$\omega = \frac{1}{4} \sup_{x \in \text{int}(K)} \langle \nabla f(x) - \nabla^2 f(x)x, (\nabla^2 f(x))^{-1} \nabla f(x) - x \rangle$$

in Proposition 4.2. In the proof of Proposition 4.2, the monotonicity of S on $\mathbb{E} \times \mathbb{R}_{++}$ is established without the hypothesis that $\{p(\cdot, \mu_k)\}$ converges point-wise to the Euclidean projector Proj_K when $\mu_k \rightarrow 0$. Thus, for any $z \in \mathbb{E}$, any $\mu, \mu' > 0$, we have

$$\langle (x, \mu) - (x', \mu'), (z - x, \mu) - (z - x', \mu') \rangle_\omega \geq 0,$$

where $x = p(z, \mu)$ and $x' = p(z, \mu')$, so that $(z - x, \mu) = S(x, \mu)$ and $(z - x', \mu) = S(x', \mu')$. This reduces to

$$-\|x - x'\|^2 + \omega(\mu - \mu')^2 \geq 0,$$

or equivalently, $\|p(z, \mu) - p(z, \mu')\| \leq \sqrt{\omega}|\mu - \mu'|$. \square

As a corollary of the convergence of the proximal point algorithm, we have the convergence of Algorithm 4.1.

Corollary 4.2. *If $F : \mathbb{E} \rightarrow \mathbb{E}$ is continuous and monotone and $\text{NCP}_K(F)$ has at least one solution, the barrier f satisfies*

$$\sup_{x \in \text{int}(K)} \langle \nabla f(x) - \nabla^2 f(x)x, (\nabla^2 f(x))^{-1} \nabla f(x) - x \rangle < \infty,$$

p is the smoothing approximation of the Euclidean projector Proj_K derived from f , and in Algorithm 4.1, the sequences $\{c_k^{-1}\}$, $\{\gamma_k\}$ and $\{c_k \gamma_k\}$ are bounded, and the infinite sequence $\{(z_k, \mu_k)\}$ generated satisfies

$$\sum_{k=0}^{\infty} c_k \|\varphi_k(z_{k+1})\| < \infty,$$

then the sequence $\{x_k := p(z_k, \mu_k)\}$ converges to a solution of $\text{NCP}_K(F)$.

Proof. Define the set-valued map

$$T : (z, \mu) \in \mathbb{E} \times \mathbb{R} \mapsto (F_{\mu_+}^{\text{nor}}(z), \mu),$$

whose zeros $(z, 0)$ give solutions $\text{Proj}_K(z)$ of $\text{NCP}_K(F)$, and denote the linear map $(x, \mu) \mapsto (c_k x, \gamma_k^{-1} \mu)$ by C_k , so that the map φ_k in Algorithm 4.1, together with $\mu \mapsto \mu + \gamma_k(\mu - \mu_k)$, is precisely the same map in Algorithm 2.1 when we associate $x \equiv (z, \mu)$ and take $R : (z, \mu) \mapsto (p(z, \mu_+), \mu_+)$. Pick any ω satisfying the inequality in Proposition 4.2 with $(c, \gamma) = (c_k, \gamma_k)$ for every k , so that $C_k T R^{-1} = C_k(H(x, \mu), \mu)$ is maximal monotone

on the Hilbert space $H_\omega = \mathbb{E} \times \mathbb{R}$ with inner product $\langle \cdot, \cdot \rangle_\omega : ((x, \mu), (y, \nu)) \mapsto \langle x, y \rangle + \omega \mu \nu$. This is possible under the assumptions that

$$\sup_{x \in \text{int}(K)} \langle \nabla f(x) - \nabla^2 f(x)x, (\nabla^2 f(x))^{-1} \nabla f(x) - x \rangle < \infty$$

and $\{c_k \gamma_k\}$ is bounded. Algorithm 4.1 is then a realization of Algorithm 2.1 where μ_{k+1} is computed exactly. Finally, the hypotheses of Theorem 2.3 are satisfied under the hypotheses of this proposition. \square

5. SYMMETRIC CONE COMPLEMENTARITY PROBLEM

In [7], Chen and Mangasarian proposed uniform smoothing approximations p of the positive part $z_+ = \max\{0, z\}$, which can be defined by twice integrating probability density functions d as follows.

$$\begin{aligned} p : (z, \mu) \in \mathbb{R} \times \mathbb{R}_{++} &\mapsto \int_{-\infty}^z \int_{-\infty}^t \frac{1}{\mu} d\left(\frac{x}{\mu}\right) dx dt = \int_{-\infty}^z \int_{-\infty}^{\frac{t}{\mu}} d(x) dx dt \\ &= \mu \int_{-\infty}^{\frac{z}{\mu}} \int_{-\infty}^t d(x) dx dt. \end{aligned} \quad (13)$$

From its definition, it is immediate that p is positively homogeneous of degree one. When the density function satisfies the conditions

- (A1) $d(t)$ is an even and piecewise continuous map with finite number of pieces, and
- (A2) $\int_{-\infty}^{\infty} |t|d(t) dt < \infty$,

it can be easily shown that

$$p(z, \mu) = \int_{-\infty}^z \frac{1}{\mu} (z - t) d\left(\frac{t}{\mu}\right) dt = z \int_{-\infty}^{\frac{z}{\mu}} d(t) dt - \mu \int_{-\infty}^{\frac{z}{\mu}} t d(t) dt \quad \forall z, \forall \mu > 0. \quad (14)$$

Moreover, the following properties of $p(z, \mu)$ hold under the above assumptions, and the following assumption.

- (A3) $d(t)$ has infinite support.

Proposition 5.1. *If d satisfies Assumptions (A1), (A2) and (A3), then $p(z, \mu)$ has the following properties.*

- (1) For each $\mu > 0$, $z \mapsto p(z, \mu)$ is convex, strictly increasing and continuously differentiable, $0 < p(z, \mu) - z_+ \leq D\mu$ for all z , where D is a constant depending only on d , and $\lim_{z \rightarrow -\infty} p(z, \mu) = 0$.
- (2) For each $\mu > 0$, $0 < \frac{\partial}{\partial z} p(z, \mu) = \int_{-\infty}^{\frac{z}{\mu}} d(t) dt < 1$, and $\frac{\partial}{\partial z} p(-z, \mu) = 1 - \frac{\partial}{\partial z} p(z, \mu)$.
- (3) For each $\mu > 0$ and each $b > 0$, $p(z, \mu) = b$ has a unique solution.

Proof. See proof of Proposition 1 of [6]. \square

Here are some examples of probability density functions satisfying Assumptions (A1)–(A3), and their associated approximations of the plus map.

Example 5.1. *Neural network smoothing map [7].*

$$p_{\text{NN}}(z, \mu) = z + \mu \log(1 + e^{-\frac{z}{\mu}}),$$

where $d(t) = e^{-t}/(1 + e^{-t})^2$.

Example 5.2. *Chen-Harker-Kanzow-Smale (CHKS) map [4, 29, 54].*

$$p_{\text{CHKS}}(z, \mu) = (z + \sqrt{z^2 + 4\mu^2})/2,$$

where $d(t) = 2/(t^2 + 4)^{\frac{3}{2}}$.

For each $\mu > 0$, the preceding proposition implies that $z \mapsto p(z, \mu)$ is a bijection between \mathbb{R} and \mathbb{R}_{++} . Moreover, by the Inverse Function Theorem, its inverse map $\varrho_\mu : \mathbb{R}_{++} \rightarrow \mathbb{R}$ is continuously differentiable with $\varrho'_\mu(x) = (\frac{\partial}{\partial z} p(\varrho_\mu(x), \mu))^{-1}$ since $\frac{\partial}{\partial z} p(z, \mu) \neq 0$ for all z . Furthermore, ϱ_μ is strictly increasing and satisfies $\lim_{x \rightarrow 0} \varrho_\mu(x) = -\infty$, and $\varrho : (x, \mu) \mapsto \varrho_\mu(x)$ is positively homogeneous of degree one. Define the function

$$f : \mathbb{R}_{++} \rightarrow \mathbb{R} : x \mapsto \int_1^x \varrho_1(t) - t \, dt,$$

so that for $(x, \mu) \in \mathbf{int}(K) \times \mathbb{R}_{++}$,

$$\varrho(x, \mu) = \mu \varrho(x/\mu, 1) = \mu \varrho_1(x/\mu) = x + \mu f'(x/\mu).$$

Then f is a twice differentiable barrier for \mathbb{R}_{++} with second derivative $\varrho'(t) - 1 = (\frac{\partial}{\partial z} p(\varrho(x), 1))^{-1} - 1 > 0$. Hence p is the smoothing approximation derived from the barrier f .

We now turn our attention to the setting where K is the symmetric cone associated with a Euclidean Jordan algebra \mathfrak{J} of rank r . The Euclidean projector is the Löwner operator

$$\text{Proj}_K : z \in \mathfrak{J} \mapsto \sum_{i=1}^r \lambda_i(z) c_i,$$

where $\sum_{i=1}^r \lambda_i(z) c_i = z$ is a spectral decomposition; see, e.g., [56]. The smoothing approximation p then extends to the smoothing approximation

$$p_{\mathfrak{J}} : (z, \mu) \in \mathfrak{J} \times \mathbb{R}_{++} \mapsto \sum_{i=1}^r p(\lambda_i(z), \mu) c_i$$

of Proj_K . By defining the barrier

$$f_{\mathfrak{J}} : z \in \mathbf{int}(K) \mapsto \sum_{i=1}^r f(\lambda_i(z)),$$

we see that $p_{\mathfrak{J}}$ is the smoothing approximation derived from the barrier $f_{\mathfrak{J}}$. Thus, when the smoothing proximal point algorithm (Algorithm 4.1) is applied to solve $\text{NCP}_K(F)$ for nonlinear monotone maps $F : \mathfrak{J} \rightarrow \mathfrak{J}$, we may deduce convergence when

$$\sup_{x \in \mathbf{int}(K)} \langle \nabla f_{\mathfrak{J}}(x) - \nabla^2 f_{\mathfrak{J}}(x)x, (\nabla^2 f_{\mathfrak{J}}(x))^{-1} \nabla f_{\mathfrak{J}}(x) - x \rangle < \infty.$$

We now check that this inequality holds under the following additional assumption.

(A4) $\lim_{x \rightarrow \infty} x^3 d(x) < \infty$ exists.

Remark 5.1. Both the neural network smoothing map and the CHKS map satisfy the above assumption; see, e.g., [6].

Proposition 5.2. *Under Assumptions (A1), (A2), (A3) and (A4),*

$$\sup_{x \in \mathbf{int}(K)} \langle \nabla f_{\mathfrak{J}}(x) - \nabla^2 f_{\mathfrak{J}}(x)x, (\nabla^2 f_{\mathfrak{J}}(x))^{-1} \nabla f_{\mathfrak{J}}(x) - x \rangle < \infty.$$

Proof. The map $(x, \mu) \in \mathbf{int}(K) \times \mathbb{R}_{++} \mapsto (x + \mu \nabla f_{\mathfrak{J}}(x/\mu), \mu)$ is the inverse map of $(z, \mu) \in \mathfrak{J} \times \mathbb{R}_{++} \mapsto (p_{\mathfrak{J}}(z, \mu), \mu)$. Thus its Jacobian is

$$\begin{aligned} & \begin{pmatrix} I + \nabla^2 f_{\mathfrak{J}}(x/\mu) & \nabla f_{\mathfrak{J}}(x/\mu) - \frac{1}{\mu} \nabla^2 f_{\mathfrak{J}}(x/\mu)x \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} J_z p_{\mathfrak{J}}(z, \mu) & J_\mu p_{\mathfrak{J}}(z, \mu) \\ 0 & 1 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} (J_z p_{\mathfrak{J}}(z, \mu))^{-1} & -(J_z p_{\mathfrak{J}}(z, \mu))^{-1} J_\mu p_{\mathfrak{J}}(z, \mu) \\ 0 & 1 \end{pmatrix}, \end{aligned}$$

where $z = x + \mu \nabla f_{\mathfrak{J}}(x/\mu)$. Since

$$J_\mu p_{\mathfrak{J}}(z, 1) = \sum_{i=1}^r \frac{\partial}{\partial \mu} p(\lambda_i(z), 1) c_i$$

and

$$J_z p_{\mathfrak{J}}(z, 1) : v \in \mathfrak{J} \mapsto \sum_{i=1}^r [\lambda_i(z), \lambda_i(z)]_{p(\cdot, 1)} v_i c_i + \sum_{1 \leq i < j \leq r} [\lambda_i(z), \lambda_j(z)]_{p(\cdot, 1)} v_{ij},$$

where $\sum_{i=1}^r \lambda_i(z) c_i = z$ is a spectral decomposition, $\sum_{i=1}^r v_i c_i + \sum_{1 \leq i < j \leq r} v_{ij}$ is a Peirce decomposition, and $[\alpha, \beta]_{p(\cdot, 1)}$ denotes the first divided difference

$$[\alpha, \beta]_{p(\cdot, 1)} = \begin{cases} \frac{\partial}{\partial z} p(\alpha, 1) & \text{if } \alpha = \beta, \\ \frac{p(\alpha, 1) - p(\beta, 1)}{\alpha - \beta} & \text{if } \alpha \neq \beta, \end{cases}$$

this means

$$\nabla f_{\mathfrak{J}}(x) - \nabla^2 f_{\mathfrak{J}}(x)x = -(J_z p_{\mathfrak{J}}(z, 1))^{-1} J_\mu p_{\mathfrak{J}}(z, 1) = - \sum_{i=1}^r \frac{\frac{\partial}{\partial \mu} p(\lambda_i(z), 1)}{\frac{\partial}{\partial z} p(\lambda_i(z), 1)} c_i$$

and

$$\begin{aligned} & (\nabla^2 f_{\mathfrak{J}}(x))^{-1} \nabla f_{\mathfrak{J}}(x) - x (\nabla^2 f_{\mathfrak{J}}(x))^{-1} (\nabla f_{\mathfrak{J}}(x) - \nabla^2 f_{\mathfrak{J}}(x)x) \\ &= -((J_z p_{\mathfrak{J}}(z, 1))^{-1} - I)^{-1} (J_z p_{\mathfrak{J}}(z, 1))^{-1} J_\mu p_{\mathfrak{J}}(z, 1) \\ &= -(I - J_z p_{\mathfrak{J}}(z, 1))^{-1} J_\mu p_{\mathfrak{J}}(z, 1) \\ &= - \sum_{i=1}^r \frac{\frac{\partial}{\partial \mu} p(\lambda_i(z), 1)}{1 - \frac{\partial}{\partial z} p(\lambda_i(z), 1)} c_i. \end{aligned}$$

Thus it suffices to show that

$$\sup_{z \in \mathbb{R}} \frac{(\frac{\partial}{\partial \mu} p(z, 1))^2}{(1 - \frac{\partial}{\partial z} p(z, 1)) \frac{\partial}{\partial z} p(z, 1)} < \infty.$$

By differentiating the last expression of $p(z, \mu)$ in (14) with respect to μ , we deduce that

$$\frac{\partial}{\partial \mu} p(z, \mu) = \int_{-\infty}^{\frac{z}{\mu}} t d(t) dt.$$

Moreover, noting that $\int_{-z/\mu}^{z/\mu} td(t) dt = 0$ under Assumption (A1), we conclude that $\frac{\partial}{\partial \mu} p(\cdot, \mu)$ is even. Under Assumption (A4), we have

$$\begin{aligned} \lim_{z \rightarrow \infty} \frac{\left(\frac{\partial}{\partial \mu} p(z, 1)\right)^2}{\left(1 - \frac{\partial}{\partial z} p(z, 1)\right) \frac{\partial}{\partial z} p(z, 1)} &= \lim_{z \rightarrow \infty} \frac{\left(\frac{\partial}{\partial \mu} p(z, 1)\right)^2}{\frac{\partial}{\partial z} p(-z, 1) \frac{\partial}{\partial z} p(z, 1)} \\ &= \lim_{z \rightarrow \infty} \frac{\left(\int_{-\infty}^z td(t) dt\right)^2}{\int_{-\infty}^{-z} d(t) dt \int_{-\infty}^z d(t) dt} \\ &= \lim_{z \rightarrow \infty} \frac{1}{\int_{-\infty}^z d(t) dt} \times \lim_{z \rightarrow \infty} \frac{2zd(z) \int_{-\infty}^z td(t) dt}{-d(-z)} \\ &= -2 \lim_{z \rightarrow \infty} \frac{\int_{-\infty}^z td(t) dt}{1/z} = \lim_{z \rightarrow \infty} 2z^3 d(z) < \infty. \end{aligned}$$

This establishes the proposition since $\frac{\left(\frac{\partial}{\partial \mu} p(z, 1)\right)^2}{\left(1 - \frac{\partial}{\partial z} p(z, 1)\right) \frac{\partial}{\partial z} p(z, 1)}$ is even and continuous in z . \square

Corollary 5.1. *If $F : \mathfrak{J} \rightarrow \mathfrak{J}$ is monotone, $\text{NCP}_K(F)$ has at least one solution, the density function d satisfies Assumptions (A1), (A2), (A3) and (A4), and in Algorithm 4.1, we use the smoothing approximation defined by (13), the sequences $\{c_k^{-1}\}$, $\{\gamma_k\}$ and $\{c_k \gamma_k\}$ are bounded, and the infinite sequence $\{(z_k, \mu_k)\}$ generated satisfies*

$$\sum_{k=0}^{\infty} c_k \|\varphi_k(z_{k+1})\| < \infty,$$

then the sequence $\{x_k := p(z_k, \mu_k)\}$ converges to a solution of $\text{NCP}_K(F)$.

6. NUMERICAL EXPERIMENTS

In this section, we report results of preliminary numerical experiments on semidefinite programming problems from SDPLIB (<http://euler.nmt.edu/~brian/sdplib/sdplib.html>). We compare our smoothing proximal point algorithm (SPPA) with the Newton-CG augmented Lagrangian algorithm (SDPNAL) proposed by Zhao, Sun and Toh [67].

6.1. Problem description. Let \mathbb{S}^n denote the vector space of n -by- n real symmetric matrices, equipped with the trace inner product $\langle X, Y \rangle := \sum_{i,j=1}^n X_{ij} Y_{ij}$, and let \mathbb{S}_+^n denote the cone of positive semidefinite matrices, which is self-dual under the trace inner product. We use the notation $X \succeq 0$ to mean $X \in \mathbb{S}_+^n$. The primal semidefinite program is

$$\min\{\langle C, X \rangle : \mathcal{A}(X) = b, X \succeq 0\},$$

where $\mathcal{A} : \mathbb{S}^n \rightarrow \mathbb{R}^m$ is linear, $b \in \mathbb{R}^m$ and $C \in \mathbb{S}^n$. Its dual problem is

$$\max\{b^T y : \mathcal{A}^*(y) + S = C, S \succeq 0\},$$

where \mathcal{A}^* denotes the adjoint operator of \mathcal{A} . When strong duality holds, all primal and dual optimal solutions X^* and (y^*, S^*) satisfy

$$X^* \succeq 0, \quad S^* \succeq 0, \quad \mathcal{A}(X^*) = b, \quad \mathcal{A}^*(y^*) + S^* = C, \quad \langle X^*, S^* \rangle = 0.$$

These conditions are, of course, sufficient for optimality in general. These conditions are equivalent to the complementarity problem $\text{NCP}_K(F)$, where K is the closed convex cone $\mathbb{S}_+^n \times \mathbb{R}^m$, and F is the affine operator

$$(X, y) \in \mathbb{S}^n \times \mathbb{R}^m \mapsto (-\mathcal{A}^*(y) + C, \mathcal{A}(X) - b).$$

6.2. Numerical algorithm. In our numerical experiments, we use SPPA to solve the above complementarity problem, in which we employ the CHKS smoothing approximation (5.2) for the cone \mathbb{S}_+^n , and the identity map as the smoothing approximation of the space \mathbb{R}^m ; i.e., we use the smoothing approximation

$$p : (Z, y, \mu) \mapsto (p_Z(Z, \mu), y),$$

where $p_Z(Z, \mu) = (Z + \sqrt{Z^2 + 4\mu^2 I})/2$ is the smoothing approximation derived from the log-determinant barrier of \mathbb{S}_+^n . We remark that its Jacobian is

$$Jp_Z(Z, \mu) : V_Z \mapsto Q(\Theta \circ (Q^T V_Z Q))Q^T = \sum_{i,j=1}^n \Theta_{ij} (q_i^T V_W q_j) q_i q_j^T,$$

where $Z = Q\Lambda Q^T = \sum_{i=1}^n \lambda_i q_i q_i^T$ is a diagonalization, \circ is the Hadamard (i.e., entry-wise) product, and Θ is the matrix of first divided differences; i.e.,

$$\Theta_{ij} := [\lambda_i, \lambda_j]_{p_{\text{CHKS}(\cdot, \mu)}} = \frac{1}{2} \left(1 + \frac{\lambda_i + \lambda_j}{\sqrt{\lambda_i^2 + 4\mu^2} + \sqrt{\lambda_j^2 + 4\mu^2}} \right) \in (0, 1).$$

At each iteration, when estimating the zero of

$$\begin{aligned} \varphi_k : (Z, y) &\mapsto F_{\mu_{k+1}}^{\text{nor}}(Z, y) + c_k^{-1}(p(Z, y, \mu_{k+1}) - p(Z_k, y_k, \mu_k)) \\ &= \begin{bmatrix} -\mathcal{A}^*(y) + C + Z - p_Z(Z, \mu_{k+1}) + c_k^{-1}(p_Z(Z, \mu_{k+1}) - p_Z(Z_k, \mu_k)) \\ \mathcal{A}(p_Z(Z, \mu_{k+1})) - b + c_k^{-1}(y - y_k) \end{bmatrix} \end{aligned}$$

we first employ a change of variable

$$\begin{aligned} W &= p_Z(Z, \mu_{k+1})/\sqrt{c_k} + \sqrt{c_k}(Z - p_Z(Z, \mu_{k+1})) \\ &= p_Z(Z, \mu_{k+1})/\sqrt{c_k} - \sqrt{c_k}\mu_{k+1} \left(\frac{p_Z(Z, \mu_{k+1})}{\mu_{k+1}} \right)^{-1} \\ &= p_Z(Z, \mu_{k+1})/\sqrt{c_k} - \mu_{k+1} \left(\frac{p_Z(Z, \mu_{k+1})/\sqrt{c_k}}{\mu_{k+1}} \right)^{-1} \end{aligned}$$

so that $p_Z(W, \mu_{k+1}) = p_Z(Z, \mu_{k+1})/\sqrt{c_k}$ and $W - p_Z(W, \mu_{k+1}) = \sqrt{c_k}(Z - p_Z(Z, \mu_{k+1}))$. This reduces the equation $\varphi_k(Z, y) = 0$ to

$$\begin{bmatrix} -\mathcal{A}^*(y) + C + W/\sqrt{c_k} - c_k^{-1}p_Z(Z_k, \mu_k) \\ \sqrt{c_k}\mathcal{A}(p_Z(W, \mu_{k+1})) - b + c_k^{-1}(y - y_k) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

from which we further reduce to the Schur-complement equation

$$\sqrt{c_k}\mathcal{A}(p_Z(\sqrt{c_k}\mathcal{A}^*(y) - C_k, \mu_{k+1})) - b + c_k^{-1}(y - y_k) = 0$$

by substituting $W = \sqrt{c_k}\mathcal{A}^*(y) - C_k$ with $C_k = \sqrt{c_k}(C - c_k^{-1}p_Z(Z_k, \mu_k))$ from the first component into the second component. Since $p_Z(\cdot, \mu)$ is the gradient map of the convex function

$$q_\mu : Z \mapsto \frac{Z^2 + Z\sqrt{Z^2 + 4\mu^2}}{4} + \mu^2 \log(Z + \sqrt{Z^2 + 4\mu^2}),$$

solving the Schur-complement equation is equivalent to minimizing the strongly convex function

$$\theta_k : y \mapsto \mathcal{A}q_{\mu_{k+1}}(\sqrt{c_k}\mathcal{A}^*(y) - C_k) - b^T y + \frac{1}{2c_k} \|y - y_k\|^2. \quad (15)$$

We use Newton's method with line search for strong Wolfe conditions to minimize this strongly convex function. The Newton step is inexactly computed with MATLAB[®]'s preconditioned conjugate gradient (PCG) function `pcg` to predetermined relative and absolute accuracies.

The details of the algorithm are presented below.

Algorithm 6.1 (Numerical smoothing proximal point algorithm).

Initialization

- Choose positive sequences $\{c_k\}$, $\{\gamma_k\}$ and $\{\varepsilon_k\}$ such that $\{1/c_k\}$, $\{\gamma_k\}$ and $\{c_k\gamma_k\}$ are bounded from above, and $\sum_{k=0}^{\infty} \varepsilon_k < \infty$.
- Set accuracy parameters $\varepsilon_{\text{infeas}} > 0$, $\eta_{\text{rel}} > 0$ and $\eta_{\text{abs}} > 0$, and line search parameters $\sigma_1, \sigma_2 \in (0, 1)$.
- Choose initial values $\mu_0 \in \mathbb{R}_{++}$ and $(Z_0, y_0) \in \mathbb{S}^n \times \mathbb{R}^m$, and set iteration counter $k = 0$.

Main iteration: Repeat the following steps until

$$\max \left\{ \frac{\|\mathcal{A}X_k - b\|}{1 + \|b\|}, \frac{\|\mathcal{A}^*y_k + S_k - C\|}{1 + \|C\|} \right\} \leq \varepsilon_{\text{infeas}},$$

where $X_k := \text{Proj}_{\mathbb{S}^n_+}(Z_k)$ and $S_k := \text{Proj}_{\mathbb{S}^n_+}(Z_k) - Z_k = \text{Proj}_{\mathbb{S}^n_+}(-Z_k)$.

Step 1. Set $\mu_{k+1} = \frac{\gamma_k}{1+\gamma_k}\mu_k$ and $(\tilde{Z}, \tilde{y}) = (Z_k, y_k)$.

Step 2. Sub-iteration: Repeat the following steps until $\|\varphi_k(\tilde{Z}, \tilde{y})\| \leq c_k^{-1}\varepsilon_k$.

(a) Determine a search direction v_y that satisfies

$$\|\nabla^2\theta_k(\tilde{y})v_y + \nabla\theta_k(\tilde{y})\| \leq \min\{\eta_{\text{abs}}, \eta_{\text{rel}}\|\nabla\theta_k(\tilde{y})\|\},$$

where θ_k is defined in (15).

(b) Find $\alpha \in [0, 1]$ such that the following strong Wolfe conditions hold:

$$\theta_k(\tilde{y} + \alpha v_y) \leq \theta_k(\tilde{y}) + \sigma_1\alpha\nabla\theta_k(\tilde{y})^T v_y;$$

$$|\theta_k(\tilde{y} + \alpha v_y)^T v_y| \leq \sigma_2|\theta_k(\tilde{y})^T v_y|.$$

(c) Update $\tilde{y} \leftarrow \tilde{y} + \alpha v_y$.

Step 3. Compute $\tilde{W} = \sqrt{c_k}\mathcal{A}^*(\tilde{y}) - \sqrt{c_k}(C - c_k^{-1}p_Z(Z_k, \mu_k))$ and

$$\tilde{Z} = \sqrt{c_k}p_Z(\tilde{W}, \mu_{k+1}) + (\tilde{W} - p_Z(\tilde{W}, \mu_{k+1}))\sqrt{c_k}.$$

Step 4. Set $(Z_{k+1}, y_{k+1}) = (\tilde{Z}, \tilde{y})$ and update $k \leftarrow k + 1$.

In our numerical tests, we

- set an upper bound of 40 sub-iterations per main iteration, and an upper bound of 400 total sub-iterations;
- set an upper bound of 200 iterations, a relative accuracy of $\eta_{\text{rel}} = 10^{-2}$ and an absolute accuracy of $\eta_{\text{abs}} = 10^{-3}$ in every call to the `pcg` function;
- use the sequences $\{\gamma_k = 0.5\}$ and $\{\varepsilon_k = c_0 k^{-1.5}\}$;
- adaptively update c_k to $\min\{2c_{k-1}, 10^8\}$ if $\|\varphi_k(Z_k, y_k)\| < c_{k-1}^{-1}$, but keep $c_k = c_{k-1}$ otherwise, where we take $c_{-1} = 1$;
- set the accuracy parameter $\varepsilon_{\text{infeas}} = 10^{-6}$;
- use the line search parameters $\sigma_1 = 10^{-4}$ and $\sigma_2 = 0.1$; and
- start from the initial values $\mu_0 = 1$ and $(X_0, y_0) = (0, 0)$.

Remark 6.1. The adaptive choice of c_k is to follow the decrease of $\|\varphi_k(Z_k, y_k)\|$.

Remark 6.2. When computing $\nabla^2\theta_k(\tilde{y})v_y$ in each PCG step, we need to evaluate

$$\begin{aligned} Jp_Z(\tilde{W}, \mu_{k+1})V_W &= Q(\Theta \circ (Q^T V_W Q))Q^T \\ &= \sum_{i,j=1}^n \Theta_{ij}(q_i^T V_W q_j)q_i q_j^T = V_W - \sum_{i,j=1}^n (1 - \Theta_{ij})(q_i^T V_W q_j)q_i q_j^T \end{aligned}$$

where $V_W = \sqrt{c_k} \mathcal{A}^*(v_y) - C_k$ and $\tilde{W} = Q\Lambda Q^T = \sum_{i=1}^n \lambda_i q_i q_i^T$ is a diagonalization. We follow the technique in [67] to reduce computational cost by dropping the (i, j) 'th term when $\Theta_{ij} \approx 0$. Alternatively, we can use the second sum and drop those (i, j) 'th terms with $\Theta_{ij} \approx 1$, if this gives more computational savings.

6.3. Tests results. We tested the SPPA and SDPNAL on the SDP relaxations of *graph partition problems* and *max-cut problems*, and the computation of *Lovász theta functions* via SDP formulation. The problem instances are drawn from SDPLIB (<http://euler.nmt.edu/~brian/sdplib/sdplib.html>).

We execute the Matlab implementations of the algorithms in Matlab version 7.13.0.564 on a machine with Intel® Q6600 CPU @ 2.40GHz, 3GB of RAM, and running Windows 7 Enterprise (64-bit). The code for SDPNAL is obtained from the authors' website (<http://www.math.nus.edu.sg/~mattohkc/SDPNAL.html>).

It should be remarked that pre-processing in the implementation of SDPNAL scales the inputs \mathcal{A} , b and C , and generates initial primal-dual iterates. For a fairer comparison, we run our code on the same scaled inputs produced by the pre-processing in the SDPNAL code.

The results of the numerical tests are shown in Tables 1, 2 and 3, respectively. The column " $m|n$ " gives the number of equality constraints and the order of the matrices involved. The column "M|Sub" gives the number of main and sub- iterations. The two columns under "Relative infeas." give, respectively, the values

$$\frac{\|\mathcal{A}X_k - b\|}{1 + \|b\|} \quad \text{and} \quad \frac{\|\mathcal{A}^*y_k + S_k - C\|}{1 + \|C\|}$$

in the final iteration, and the column "Relative gap" gives the value

$$\frac{|\langle C, X_k \rangle - b^T y_k|}{1 + |\langle C, X_k \rangle| + |b^T y_k|}$$

in the final iteration. It can be observed from the tables that the SPPA is comparable with SDPNAL.

Name of problem	$m n$	SPPA					SDPNAL				
		Iter	Relative infeas.		Relative gap	Time (m:s)	Iter	Relative infeas.		Relative gap	Time (m:s)
		M Sub	Primal	Dual			M Sub	Primal	Dual		
gpp100	101 100	16 44	8.7e-07	9.7e-10	4.0e-07	2.0	25 94	1.3e-07	6.0e-07	-2.1e-06	5.2
gpp124-1	125 124	17 58	8.0e-07	4.3e-09	8.0e-07	3.4	25 237	3.6e-07	5.8e-07	-1.3e-05	10.2
gpp124-2	125 124	14 44	9.4e-07	2.5e-07	5.1e-07	1.8	22 100	5.6e-07	4.5e-07	-2.8e-06	4.0
gpp124-3	125 124	14 40	4.3e-07	3.0e-07	3.8e-07	1.7	25 85	7.3e-07	1.0e-06	-2.2e-06	3.5
gpp124-4	125 124	16 43	8.9e-07	6.8e-10	3.5e-07	2.1	25 90	8.0e-07	5.1e-07	-1.9e-06	3.9
gpp250-1	251 250	15 49	2.2e-07	9.6e-07	1.8e-06	10.0	26 299	6.8e-07	4.0e-07	-1.5e-05	33.4
gpp250-2	251 250	14 45	2.8e-07	3.6e-07	4.3e-07	7.4	24 159	5.3e-07	7.7e-07	-4.3e-06	15.3
gpp250-3	251 250	14 44	2.2e-07	1.4e-07	1.3e-07	6.9	25 84	5.2e-07	5.9e-07	-1.9e-06	8.2
gpp250-4	251 250	13 40	8.7e-07	8.3e-07	5.8e-07	5.6	24 84	4.7e-07	4.9e-07	-1.7e-06	7.9
gpp500-1	501 500	15 56	1.2e-08	1.6e-07	1.4e-07	57.9	22 394	2.0e-06	1.1e-06	-3.6e-05	3:41.7
gpp500-2	501 500	13 47	6.1e-07	4.2e-07	1.5e-07	41.0	22 234	2.0e-07	7.8e-07	-7.0e-06	1:43.9
gpp500-3	501 500	13 45	1.1e-07	2.4e-08	3.7e-09	35.4	21 98	1.7e-07	7.2e-07	-2.7e-06	42.2
gpp500-4	501 500	12 41	6.0e-07	1.0e-07	1.2e-08	31.4	19 73	1.1e-07	7.5e-07	-2.2e-06	29.8
equalG11	801 801	21 66	7.7e-07	7.5e-07	5.6e-06	2:35.2	29 103	2.1e-07	6.5e-07	-1.0e-04	2:18.8
equalG51	1001 1001	20 65	5.7e-07	4.5e-07	2.9e-06	3:51.7	26 149	9.4e-07	7.5e-07	-2.1e-05	5:09.7

TABLE 1. Graph partitioning problems. For gpp250-1, SDPNAL stops after 26 iterations as it detects a lack of progress in the relative primal infeasibility. For gpp500-1, SDPNAL stops as the dual infeasibility is less than 8×10^{-7} .

Name of problem		SPPA					SDPNAL						
		Iter		Relative infeas.		Relative gap	Time (m:s)	Iter		Relative infeas.		Relative gap	Time (m:s)
		M Sub	Primal	Dual	M Sub			Primal	Dual				
mcp100	100 100	13 33	1.1e-07	1.6e-07	6.6e-08	1.4	22 48	1.2e-07	4.9e-07	-4.1e-07	1.4		
mcp124-1	124 124	15 42	4.1e-08	8.4e-07	3.1e-07	2.4	23 65	7.7e-07	5.5e-07	-3.7e-07	2.3		
mcp124-2	124 124	13 34	3.1e-07	2.6e-07	1.1e-07	1.8	21 51	1.0e-07	7.4e-07	-4.5e-07	1.8		
mcp124-3	124 124	12 31	7.1e-07	6.8e-07	2.4e-07	1.6	21 52	7.1e-07	7.1e-07	-3.8e-07	1.7		
mcp124-4	124 124	12 30	1.9e-07	1.7e-07	5.5e-08	1.6	21 45	8.4e-07	6.8e-07	-7.1e-07	1.7		
mcp250-1	250 250	15 45	2.7e-08	7.9e-07	6.1e-07	8.9	25 69	8.1e-07	7.9e-07	-9.6e-07	5.3		
mcp250-2	250 250	14 37	4.0e-07	3.3e-07	2.0e-07	6.4	23 63	1.7e-07	5.3e-07	-5.6e-07	4.4		
mcp250-3	250 250	14 41	1.3e-07	6.2e-08	2.8e-08	6.3	24 57	2.6e-07	6.7e-07	-5.7e-07	4.4		
mcp250-4	250 250	13 40	2.3e-08	8.5e-07	3.5e-07	6.3	24 58	2.7e-07	6.7e-07	-7.2e-07	4.3		
mcp500-1	500 500	16 50	2.0e-08	6.1e-07	7.4e-07	57.1	27 77	1.1e-07	6.2e-07	-1.2e-06	28.5		
mcp500-2	500 500	15 46	4.4e-08	3.9e-07	3.9e-07	42.8	24 66	2.9e-07	7.0e-07	-1.3e-06	19.0		
mcp500-3	500 500	14 43	3.3e-08	6.1e-07	4.3e-07	30.8	24 58	1.6e-07	9.4e-07	-1.4e-06	16.3		
mcp500-4	500 500	14 41	4.6e-07	2.4e-07	1.7e-07	29.8	23 55	2.8e-07	7.4e-07	-1.3e-06	15.8		
maxG11	800 800	20 62	7.9e-07	9.6e-07	2.6e-06	3:50.4	30 105	2.5e-07	8.0e-07	-3.9e-06	2:13.2		
maxG51	1000 1000	16 51	1.1e-08	1.7e-07	1.2e-07	3:36.1	23 70	3.1e-07	8.7e-07	-2.6e-06	1:29.9		
maxG32	2000 2000	21 67	2.0e-08	6.7e-07	2.2e-06	37:51.8	31 110	2.2e-07	5.3e-07	-3.8e-06	19:45.0		

TABLE 2. Max-cut problems.

Name of problem		SPPA					SDPNAL						
		Iter		Relative infeas.		Relative gap	Time (m:s)	Iter		Relative infeas.		Relative gap	Time (m:s)
		M Sub	Primal	Dual	M Sub			Primal	Dual				
theta1	104 50	14 43	1.7e-07	6.4e-09	1.0e-07	1.4	9 24	3.0e-07	2.7e-07	6.8e-08	0.5		
theta2	498 100	17 77	4.7e-07	4.0e-09	6.3e-08	9.2	22 26	1.5e-07	6.4e-07	-1.9e-07	1.3		
theta3	1106 150	17 66	9.1e-07	4.3e-09	2.1e-07	10.6	21 23	5.7e-08	5.6e-07	-6.7e-08	1.9		
theta4	1949 200	17 84	7.9e-07	8.9e-09	4.4e-07	24.3	20 24	1.1e-07	9.0e-07	-2.0e-07	2.4		
theta5	3028 250	17 74	9.8e-07	1.1e-08	2.4e-07	38.9	20 23	5.1e-08	4.8e-07	-1.1e-07	3.4		
theta6	4375 300	18 101	3.2e-07	3.5e-09	1.2e-07	1:25.5	21 29	3.7e-08	5.0e-07	-1.1e-07	4.6		
thetaG11	2401 801	33 193	7.1e-08	3.9e-14	1.5e-08	11:23.7	31 162	2.5e-07	1.9e-12	1.0e-07	3:32.8		
thetaG51	6910 1001	58 400	8.5e-02	1.2e-04	9.9e-01	90:41.8	36 493	3.3e-03	7.0e-07	1.9e-03	51:36.6		

TABLE 3. Lovász theta function. Both algorithms have difficulties solving thetaG51 to the desired accuracy. SDPNAL stops as the dual infeasibility is less than 8×10^{-7} ; SPPA reaches its maximal number of total iterations.

7. CONCLUSION

A new barrier-based method of constructing smoothing approximations of Euclidean projectors onto closed convex cones is introduced. With these smoothing approximations, we propose a smoothing proximal point algorithm to solve monotone nonlinear complementarity problems over closed convex cones via their normal map equations.

The combination of smoothing approximations and proximal terms allow for the stable application of inexact global Newton’s methods. We demonstrate a sufficient condition for the convergence of the algorithm when the nonlinear complementarity problem has a solution. This sufficient condition is a condition on the barrier used in the construction of the smoothing approximation, and is satisfied by all logarithmically homogeneous barriers. Thus our method and its convergence analysis applies to all monotone nonlinear complementarity problems over closed convex cones.

Preliminary numerical tests show that in practice, our method is comparable with the Newton-CG augmented Lagrangian algorithm when it comes to solving semidefinite programs.

8. ACKNOWLEDGEMENT

We thank the anonymous referees for the very helpful suggestions on improving the presentation of this paper.

REFERENCES

- [1] A. Auslender, M. Teboulle, and S. Ben-Tiba, *A logarithmic-quadratic proximal method for variational inequalities*, Comput. Optim. Appl. **12** (1999), 31–40.
- [2] R. S. Burachik and A. N. Iusem, *A generalized proximal point algorithm for the variational inequality problem in a Hilbert space*, SIAM J. Optim. **8** (1998), 197–216.
- [3] B. Chen and X. Chen, *A global and local superlinear continuation-smoothing method for P_0 and R_0 NCP or monotone NCP*, SIAM J. Optim. **9** (1999), 624–645.
- [4] B. Chen and P. T. Harker, *A noninterior-point continuation method for linear complementarity problems*, SIAM J. Matrix Anal. Appl. **14** (1993), 1168–1190.
- [5] ———, *Smooth approximation to nonlinear complementarity problems*, SIAM J. Optim. **7** (1997), 403–420.
- [6] B. Chen, P. T. Harker, and M. Ç. Pınar, *Continuation method for nonlinear complementarity problems via normal maps*, Eur. J. Oper. Res. **116** (1999), 591–606.
- [7] C. Chen and O. L. Mangasarian, *A class of smoothing functions for nonlinear and mixed complementarity problems*, Comput. Optim. Appl. **5** (1996), 97–138.
- [8] X. Chen, H. Qi, and P. Tseng, *Analysis of nonsmooth symmetric-matrix functions with applications to semidefinite complementarity problems*, SIAM J. Optim., **13** (2003), 960–985.
- [9] 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** (2003), 39–56.
- [10] X. Chen and P. Tseng, *Non-interior continuation methods for solving semidefinite complementarity problems*, Math. Program., **95** (2003), 431–474.
- [11] J.-S. Chen and P. Tseng, *An unconstrained smooth minimization reformulation of the second-order cone complementarity problem*, Math. Program., **104** (2005), 293–327.
- [12] C. B. Chua and P. Yi, *A continuation method for nonlinear complementarity problems over symmetric cones*, SIAM J. Optim. **20** (2010), 2560–2583.
- [13] T. De Luca, F. Facchinei, and C. Kanzow. *A semismooth equation approach to the solution of nonlinear complementarity problems*, Math. Program. **75** (2000) 407–439.
- [14] B. C. Eaves, *On the basic theorem of complementarity*, Math. Program. **1** (1971), 68–75.
- [15] J. Eckstein, *Nonlinear proximal point algorithms using Bregman functions, with applications to convex programming*, Math. Oper. Res. **18** (1993), 202–226.
- [16] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems Vol I, II*, Springer-Verlag, New York, NY, USA, 2003.
- [17] F. Facchinei, J. Soares. *A new merit function for nonlinear complementarity problems and a related algorithm*, SIAM J. Optim. **7** (1997), 225–247.
- [18] A. Fischer, *A Newton-type method for positive-semidefinite linear complementarity problems*, J. Optimization Theory Appl. **86** (1995), 585–608.
- [19] M. Fukushima, Z.-Q. Luo, and P. Tseng, *Smoothing functions for second-order-cone complementarity problems*, SIAM J. Optim., **12** (2001), 436–460.
- [20] S. A. Gabriel and J. S. Pang, *An inexact NE/SQP method for solving the nonlinear complementarity problem*, Comput. Optim. Appl. **1** (1992), 67–91.
- [21] M. S. Gowda and R. Sznajder, *Automorphism invariance of \mathbf{P} - and \mathbf{GUS} -properties of linear transformations on Euclidean Jordan algebras*, Math. Oper. Res., **31** (2006), 109–123.
- [22] ———, *Some global uniqueness and solvability results for linear complementarity problems over symmetric cones*, SIAM J. Optim., **18** (2007), 461–481.
- [23] M. S. Gowda, R. Sznajder, and J. Tao, *Some \mathbf{P} -properties for linear transformations on Euclidean Jordan algebras*, Linear Algebra Appl., **393** (2004), 203–232.
- [24] P. T. Harker and B. Xiao, *Newton’s method for the nonlinear complementarity problem: A B -differentiable equation approach*, Math. Program. **48** (1990), 339–358.
- [25] S. Hayashi, N. Yamashita, and M. Fukushima, *A combined smoothing and regularization method for monotone second-order cone complementarity problems*, SIAM J. Optim., **15** (2005), 593–615.

- [26] Z.-H. Huang and J. Han, *Non-interior continuation method for solving the monotone semidefinite complementarity problem*, Appl. Math. Optim., **47** (2003), 195–211.
- [27] Z.-H. Huang and T. Ni, *Smoothing algorithms for complementarity problems over symmetric cones*, Comput. Optim. Appl., (2008). <http://dx.doi.org/10.1007/s10589-008-9180-y>.
- [28] J. Huang and J.-S. Pang, *Option pricing and linear complementarity*, J. Comput. Finance **2** (1998), 31–60.
- [29] C. Kanzow, *Some noninterior continuation methods for linear complementarity problems*, SIAM J. Matrix Anal. Appl. **17** (1996), 851–868.
- [30] C. Kanzow, N. Yamashita, and M. Fukushima, *New NCP-functions and their properties*, J. Optimization Theory Appl. **94**, (1997), 115–135.
- [31] M. Kojima, Y. Kurita, and S. Mizuno, *Large-step interior point algorithms for linear complementarity problems*, SIAM J. Optim. **3** (1993), 398–412.
- [32] M. Kojima, N. Mejidlo, and T. Noma, *Homotopy continuation method for nonlinear complementarity problems*, Math. Oper. Res. **16** (1991), 754–774.
- [33] L. Kong, J. Sun, and N. Xiu, *A regularized smoothing Newton method for symmetric cone complementarity problems*, SIAM J. Optim., **19** (2008), 1028–1047.
- [34] L. Kong and N. Xiu, *New smooth C-functions for symmetric cone complementarity problems*, Optim. Lett., **1** (2007), 391–400.
- [35] Y. J. Liu, D. Sun, and K. C. Toh, *An implementable proximal point algorithmic framework for nuclear norm minimization*, Research report, Department of Mathematics, National University of Singapore, Singapore, October 2010.
- [36] Y.-J. Liu, L.-W. Zhang, and Y.-H. Wang, *Some properties of a class of merit functions for symmetric cone complementarity problems*, Asia-Pac. J. Oper. Res., **23** (2006), 473–495.
- [37] A. Löhne, *A characterization of maximal monotone operators*, Set-Valued Anal. **16** (2008), 693–700.
- [38] Z. Q. Luo and P. Tseng, *A new class of merit functions for the nonlinear complementarity problem*, Complementarity and Variational Problems: State of the Art (M. C. Ferris and J.-S. Pang, eds.), SIAM Publications, Philadelphia, 1997, pp. 204–225.
- [39] O. L. Mangasarian, *Equivalence of the complementarity problem to a system of nonlinear equations*, SIAM J. Appl. Math. **31** (1976), 89–92.
- [40] B. Martinet, *Régularisation d'inéquations variationnelles par approximations successives*, Rev. Française Informat. Techerche Opérationnelle **4** (1970), 154–158.
- [41] R. D. McKelvey and A. McLennan, *Computation of equilibria in finite games*, Handbook of Computational Economics vol 1 (H. M. Amman, D. A. Kendrick, and J. Rust., eds.), Elsevier, North Holland, 1996.
- [42] T. C. Miller, T. L. Friesz, and R. L. Tobin, *Equilibrium Facility Location on Networks*, Springer-Verlag, New York, NY, USA, 1996.
- [43] G. J. Minty, *Monotone (nonlinear) operators in Hilbert space*, Duke Math J. **29** (1962), 341–346.
- [44] T. Ni and W.-Z. Gu, *Smoothing Newton algorithm for symmetric cone complementarity problems based on a one-parametric class of smoothing functions*, J. Appl. Math. Comput., (2009). <http://dx.doi.org/10.1007/s12190-009-0341-7>.
- [45] S. Pan and J.-S. Chen, *A one-parametric class of merit functions for the symmetric cone complementarity problem*, J. Math. Anal. Appl., **355** (2009), 195–215.
- [46] L. Qi and X. Chen, *A parametrized Newton method and a quasi-Newton method for nonsmooth equations*, Comput. Optim. Appl. **3** (1994), 157–179.
- [47] L. Qi and X. Chen, *A globally convergent successive approximation method for severely nonsmooth equations*, SIAM J. Control Optim. **33** (1995), 402–418.
- [48] L. Qi, D. Sun, and J. Sun, *A squared smoothing Newton method for nonsmooth matrix equations and its applications in semidefinite optimization problems*, SIAM J. Optim. **14** (2004), 783–806.
- [49] R. T. Rockafellar, *Characterization of the subdifferentials of convex functions*, Pacific J. Math. **17** (1966), 497–510.
- [50] ———, *On the maximal monotonicity of subdifferential mappings*, Pacific J. Math. **33** (1970), 209–216.
- [51] ———, *Monotone operators and the proximal point algorithm*, SIAM J. Control Optim. **14** (1976), 877–898.
- [52] ———, *Augmented Lagrangians and applications of the proximal point algorithm in convex programming*, Math. Oper. Res. **1** (1976), 97–116.
- [53] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer, Berlin, 1998.

- [54] S. Smale, *Algorithms for solving equations*, Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Berkeley, Calif., 1986), Amer. Math. Soc., Providence, RI, 1987, 172–195.
- [55] M. Shida and S. Shindoh, *Monotone semidefinite complementarity problems*, Research Report 312, Tokyo Institute of Technology, Tokyo, Japan, 1996.
- [56] D. Sun and J. Sun, *Löwner’s operator and spectral functions in Euclidean Jordan algebras*, Math. Oper. Res. **33** (2008), 421–445.
- [57] J. Tao and M. S. Gowda, *Some \mathbf{P} -properties for nonlinear transformations on Euclidean Jordan algebras*, Math. Oper. Res., **30** (2005), 985–1004.
- [58] M. Teboulle, *Convergence of proximal-like algorithms*, SIAM J. Optim. **7** (1997), 1069–1083.
- [59] P. Tseng, *An infeasible path-following method for monotone complementarity problems*, SIAM J. Optim. **7** (1997), 386–402.
- [60] P. Tseng, *Merit functions for semi-definite complementarity problems*, Math. Program., **83** (1998), 159–185.
- [61] C. J. Wang, D. Sun, and K. C. Toh, *Solving Log-Determinant Optimization Problems by a Newton-CG Primal Proximal Point Algorithm*, SIAM J. Optim. **20** (2010), 2994–3013.
- [62] S. J. Wright, *An infeasible-interior-point method for linear complementarity problems*, Math. Program., **67** (1994), 29–51.
- [63] N. Yamashita and M. Fukushima, *A new merit function and a descent method for semidefinite complementarity problems*, Appl. Optim., **22** (1999), 405–420.
- [64] Y. Ye and K. Anstreicher, *On quadratic and $O(\sqrt{n}L)$ convergence of a predictor-corrector algorithm for LCP*, Math. Program. **59** (1993), 537–551.
- [65] A. Yoshise, *Interior point trajectories and a homogeneous model for nonlinear complementarity problems over symmetric cones*, SIAM J. Optim., **17** (2006), 1129–1153.
- [66] Y. B. Zhao and D. Li, *Monotonicity of fixed point and normal mappings associated with variational inequality and its application*, SIAM J. Optim. **11** (2001), 962–973.
- [67] X. Y. Zhao, D. Sun, and K. C. Toh, *A Newton-CG augmented Lagrangian method for semidefinite programming*, SIAM J. Optim. **20** (2010), 1737–1765.

APPENDIX A. PROOF OF THEOREM 2.3

For the simplicity of notation, we drop the subscripts in $\langle \cdot, \cdot \rangle_{\mathbb{H}}$ and $\|\cdot\|_{\mathbb{H}}$.

Proposition A.1 (Proposition 1 of [51]). *If $T : \mathbb{H} \rightrightarrows \mathbb{H}$ is a maximal monotone operator and $P = (I + T)^{-1}$ is the proximal mapping for T , then for $Q = I - P$, and all $x, x' \in \mathbb{E}$,*

- (1) $Q(x) \in T(P(x))$,
- (2) $\langle P(x) - P(x'), Q(x) - Q(x') \rangle \geq 0$,
- (3) $\|P(x) - P(x')\|^2 + \|Q(x) - Q(x')\|^2 \leq \|x - x'\|^2$.

Proposition A.2 (cf. Proposition 3 of [51]). *For any maximal monotone operator $T : \mathbb{H} \rightrightarrows \mathbb{H}$, any bijective linear map $C : \mathbb{H} \rightarrow \mathbb{H}$, and any map $R : \mathbb{H} \rightarrow \mathbb{H}$, if CTR^{-1} is maximal monotone, and $P = (I + CTR^{-1})^{-1}$ is the proximal mapping for CTR^{-1} , then*

$$\|R(x') - P(R(x))\| \leq \min\{\|Cw\| : w \in T(x') + C^{-1}(R(x') - R(x))\}.$$

Proof. For any $w \in T(x') + C^{-1}(R(x') - R(x))$, we have $Cw + R(x) \in (I + CTR^{-1})(R(x'))$, and hence $R(x') = P(Cw + R(x))$. Therefore

$$\|R(x') - P(R(x))\| = \|P(Cw + R(x)) - P(R(x))\| \leq \|Cw\|$$

by Theorem 2.2. □

Proof of Theorem 2.3. Pick an arbitrary zero x_* of T . For each k , let P_k denote the proximal mapping $(I + C_k TR^{-1})^{-1}$ for $C_k TR^{-1}$, and let Q_k denote $I - P_k$. The image $R(x_*)$ is a fixed point of P_k for each k .

Denoting by ε_k each term $\|C_k T(x_{k+1}) + R(x_{k+1}) - R(x_k)\|$ in (4), we deduce from Theorem 2.2 that for each k ,

$$\begin{aligned} & \|R(x_{k+1}) - R(x_*)\| \\ & \leq \|R(x_{k+1}) - P_k(R(x_k))\| + \|P_k(R(x_k)) - R(x_*)\| \\ & = \|P_k(C_k T(x_{k+1}) + R(x_{k+1})) - P_k(R(x_k))\| + \|P_k(R(x_k)) - P_k(R(x_*))\| \\ & \leq \|C_k T(x_{k+1}) + R(x_{k+1}) - R(x_k)\| + \|R(x_k) - R(x_*)\| \\ & = \varepsilon_k + \|R(x_k) - R(x_*)\|; \end{aligned}$$

and subsequently for each $l < k$, $\|R(x_{k+1}) - R(x_*)\| - \|R(x_l) - R(x_*)\| \leq \sum_{i=l}^k \varepsilon_i$. This proves that

$$\lim_{k \rightarrow \infty} \|R(x_k) - R(x_*)\| = \inf_{k=0,1,\dots} \|R(x_k) - R(x_*)\|,$$

and thus $\{R(x_k)\}$ is bounded and has a weak limit point. It remains to show that every weak limit point of $\{R(x_k)\}$ is a zero of T , and that the sequence has at most one weak limit point.

Pick an arbitrary weak limit point z_∞ of this sequence. Since, by Proposition A.2, $0 \leq \|R(x_{k+1}) - P_k(R(x_k))\| \leq \varepsilon_k \rightarrow 0$, we have $\lim_{k \rightarrow \infty} \|R(x_{k+1}) - P_k(R(x_k))\| = 0$, whence z_∞ is a weak limit point of $\{P_k(R(x_k))\}$. By Proposition A.1, $C_1 C_k^{-1} Q_k(R(x_k)) \in C_1 TR^{-1}(P_k(R(x_k)))$ for each k . Hence by the monotonicity of $C_1 TR^{-1}$, we deduce that for any $(z, \tilde{z}) \in \mathcal{G}(C_1 TR^{-1})$ and any k ,

$$\langle z - P_k(R(x_k)), \tilde{z} - C_1 C_k^{-1} Q_k(R(x_k)) \rangle \geq 0. \tag{A.1}$$

Since for each k , $R(x_*)$ is a fixed point of P_k , and hence a zero of Q_k , we deduce from Proposition A.1 that

$$\begin{aligned} & \|P_k(R(x_k)) - R(x_*)\|^2 + \|Q_k(R(x_k))\|^2 \\ &= \|P_k(R(x_k)) - P_k(R(x_*))\|^2 + \|Q_k(R(x_k)) - Q_k(R(x_*))\|^2 \\ &\leq \|R(x_k) - R(x_*)\|^2; \end{aligned}$$

and subsequently

$$\begin{aligned} & \|R(x_{k+1}) - R(x_*)\|^2 + \|Q_k(R(x_k))\|^2 - \|R(x_k) - R(x_*)\|^2 \\ &\leq \|R(x_{k+1}) - R(x_*)\|^2 - \|P_k(R(x_k)) - R(x_*)\|^2 \\ &= \langle R(x_{k+1}) - P_k(R(x_k)), R(x_{k+1}) - R(x_*) + P_k(R(x_k)) - R(x_*) \rangle \\ &\leq \|R(x_{k+1}) - P_k(R(x_k))\| (\|R(x_{k+1}) - R(x_*)\| + \|P_k(R(x_k)) - P_k(R(x_*))\|) \\ &\leq \|R(x_{k+1}) - P_k(R(x_k))\| (\|R(x_{k+1}) - R(x_*)\| + \|R(x_k) - R(x_*)\|) \\ &\leq \varepsilon_k (\|R(x_{k+1}) - R(x_*)\| + \|R(x_k) - R(x_*)\|). \end{aligned}$$

In the limit, we see that $Q_k(R(x_k)) \rightarrow 0$ strongly, whence $C_k^{-1}Q_k(R(x_k)) \rightarrow 0$ strongly as $\{C_k^{-1}\}$ is assumed to be bounded. Thus, taking limit on a suitable subsequence of (A.1) gives $\langle z - z_\infty, \tilde{z} \rangle \geq 0$. Since $(z, \tilde{z}) \in \mathcal{G}(C_1TR^{-1})$ is arbitrary, the maximal monotonicity of C_1TR^{-1} implies that $0 \in C_1TR^{-1}(z_\infty)$. Hence, z_∞ is a zero of TR^{-1} .

Finally, using the same argument as in the proof of Theorem 1 of [51], we conclude that $\{R(x_k)\}$ has a unique weak limit point. \square

(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

(Zhen Li) DIVISION OF MATHEMATICAL SCIENCES, SCHOOL OF PHYSICAL & MATHEMATICAL SCIENCES, NANYANG TECHNOLOGICAL UNIVERSITY, SINGAPORE 637371, SINGAPORE

E-mail address: lizhen@ntu.edu.sg