J. Fixed Point Theory Appl. (2023) 25:39
https://doi.org/10.1007/s11784-022-01026-8

Journal of Fixed Point Theory and Applications

Method of alternating projections for the general absolute value equation

Jan Harold Alcantara, Jein-Shan Chen and Matthew K. Tam

Abstract. A novel approach for solving the general absolute value equation $Ax + B|x| = c$ where $A, B \in \mathbb{R}^{m\times n}$ and $c \in \mathbb{R}^m$ is presented. We reformulate the equation as a nonconvex feasibility problem which we solve via the method of alternating projections (MAP). The fixed points set of the alternating projections map is characterized under nondegeneracy conditions on A and B. Furthermore, we prove local linear convergence of the algorithm. Unlike most of the existing approaches in the literature, the algorithm presented here is capable of handling problems with $m \neq n$, both theoretically and numerically.

Mathematics Subject Classification. Primary 90-08, 65K10.

Keywords. Absolute value equation, alternating projections, fixed point sets.

1. Introduction

In this paper, we consider the *absolute value equation (AVE)* given by

$$Ax + B|x| = c, \tag{1.1}$$

where $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{m\times n}$, $c \in \mathbb{R}^m$ and $|x|$ denotes the componentwise absolute value of $x \in \mathbb{R}^n$. Equation (1.1) with $m = n$ was first introduced by Rohn in [1] as a generalization of the equation

$$Ax - |x| = c, \tag{1.2}$$

which has been the subject of numerous research works for almost two decades; see [2–11]. Interest in Eq. (1.2) is primarily motivated by its equivalence with the *linear complementarity problem (LCP)*, which encompasses several mathematical programming problems [9,12–15]. In addition, absolute value equations of the form (1.2) are also intimately related with mixed integer programming [15] and interval linear equations [16]. On the other hand, the (possibly non-square) absolute value equation (1.1) was first investigated by Mangasarian in [14] to provide a more general framework for studying the

Birkhäuser

Published online: 03 January 2023

traditional AVE (1.2), which is our motivation for considering the general case (1.1).

As shown in [14, Proposition 2], we note that solving (1.1) is an NP-hard problem. Meanwhile, conditions for existence, non-existence and uniqueness of solutions of the AVE are reported in [5,9,10,17]. On the numerical side, there are already many algorithms aimed at solving the special case (1.2). These algorithms can be roughly classified into four categories:

(a) *Newton methods.* Most of the algorithms for solving AVE in the literature are based on modifications of the Newton method. For instance, a semismooth Newton method is proposed in [7] to directly handle the nonsmooth equation (1.1) using the limiting subdifferential of $|x|$. Variants of the semismooth Newton method were also proposed, which include the inexact semismooth Newton method [3] and the generalized Traub's method [4]. Another approach followed by several works involves replacing the absolute value function by its smooth approximation, which then permits the use of the classical Newton method. This technique, known as the smoothing Newton method, was employed in several works such as in [2,18]. A combination of both the semismooth and smoothing Newton method is also described in [11].

(b) *Picard iteration methods.* The Newton methods described above involve solving (exactly or approximately) linear systems of equations with different coefficient matrices at each iteration, which may be computationally expensive. On the other hand, in the Picard iteration method proposed in [10], a linear system with a fixed coefficient matrix A is solved in each iteration (see also Eq. (4.3)), and thus may be more efficient than Newton methods. However, this approach is limited to the case when A is invertible. A variant of this algorithm, known as the Picard-HSS iteration, is proposed in [19] for handling the case that A is non-Hermitian positive definite. The Douglas–Rachford splitting method recently proposed in [20] may also be viewed as an extension of the Picard iterations (4.3).

(c) *Matrix splitting iteration method.* Under this category are two algorithms, namely the SOR (successive over-relaxation)-like iteration [21] and the Gauss–Seidel iteration method [22]. We note the observation that the Picard iteration method [10] is a special case of the SOR-like iteration method, although the latter is derived from a matrix splitting approach.

(d) *Concave minimization approach.* Mangasarian pioneered this approach by reformulating the AVE as a concave minimization problem and then using the successive linearization algorithm to solve the resulting reformulated problem [6,8]. In another recent work [23], the AVE is reformulated as a complementarity problem, which was smoothly approximated by a concave minimization problem.

Meanwhile, to the best of our knowledge, the only method which can handle AVE (1.1) when $B \neq -I_n$ and $m \neq n$ is the successive linearization algorithm via concave minimization proposed in [14].

1.1. Our approach

Due to limited work on the general AVE, we propose a simple approach that can solve (1.1) which, like [14], does not require $B = -I_n$ or $m = n$. With its generality, our proposed algorithm can handle as well the conventional AVE (1.2), and in turn useful for mathematical programming problems that can be reformulated as (1.2). Moreover, for the traditional case when $B = -I_n$ and $m = n$, our new approach does not fall in any of the four categories (a)–(d) described above.

We reformulate the AVE as a feasibility problem and then use the method of alternating projections (MAP) to solve the resulting problem. By introducing an auxiliary variable $y \in \mathbb{R}^n$, we have that $x \in \mathbb{R}^n$ solves (1.1) if and only if the pair $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ solves

$$Ax + By = c \qquad \text{and} \qquad y = |x|.$$

The above system of equations suggests the reformulation of (1.1) as a *feasibility problem* given by

$$\text{find } (x, y) \in S_1 \cap S_2 \subseteq \mathbb{R}^n \times \mathbb{R}^n, \tag{1.3}$$

where the constraint sets, S_1 and S_2, are given by

$$\begin{aligned} S_1 &:= \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n : Ax + By = c\} \qquad \text{and} \\ S_2 &:= \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n : y = |x|\}. \end{aligned} \tag{1.4}$$

Observe that S_1 is an affine set, and is, thus, convex. On the other hand, S_2 is a nonconvex set but is expressible as a finite union of convex sets.

A simple algorithm to solve (1.3) is the *method of alternating projections (MAP)*: Given an initial point $z^0 = (x^0, y^0) \in \mathbb{R}^n \times \mathbb{R}^n$, MAP generates a sequence of iterates according to the rule:

$$z^{k+1} = (x^{k+1}, y^{k+1}) \in (P_{S_1} \circ P_{S_2})(z^k) \qquad \forall k \in \mathbb{N}, \tag{1.5}$$

where P_S is the possibly multivalued metric projector onto the set S given by

$$P_S(z) := \{s \in S : \|s - z\| \leq \|t - z\| \ \forall t \in S\}.$$

When S is nonempty and closed, the image of P_S at each point is nonempty. If in addition, S is convex, the function P_S is single-valued everywhere. Whenever $P_S(z)$ is single-valued, say $P_S(z) = \{s\}$, we simply write $s = P_S(z)$.

1.2. Contributions

Our primary contribution includes the proposal of a new algorithm, namely MAP (1.5), for solving the general absolute value equation (1.1). The nonconvexity of the set S_2 complicates the convergence analysis of the proposed algorithm. Nonetheless, we outline our important theoretical contributions as follows.

(a) A significant portion of our analyses involve the characterization of *fixed points* of the alternating projections mapping $P_{S_1} \circ P_{S_2}$, that is, the set

$$\text{Fix}(P_{S_1} \circ P_{S_2}) = \{z \in \mathbb{R}^n \times \mathbb{R}^n : z \in (P_{S_1} \circ P_{S_2})(z)\}, \tag{1.6}$$

which necessarily contains $S_1 \cap S_2$, but not conversely. Hence, if a sequence generated by MAP (1.5) is convergent, then its limit, which belongs to $\mathrm{Fix}(P_{S_1} \circ P_{S_2})$, need only be a candidate solution to the feasibility problem (1.3). We determine what conditions on A and B are sufficient to guarantee that a fixed point of $P_{S_1} \circ P_{S_2}$ is indeed a point in $S_1 \cap S_2$ (see Theorems 2.7, 2.13, 2.17 and 2.19).

(b) We establish in Theorem 3.1 the local convergence of the algorithm (1.5) using the theory developed by Dao and Tam (2019) in [24], which uses ideas originally developed in [25,26]. We also present a new complementarity function in Sect. 3.2 which we use to provide an alternative convergence analysis of the MAP iterates. Local linear convergence of MAP (1.5) to a point in $S_1 \cap S_2$ is proved in Theorems 3.17 and 3.19 under the same conditions used to characterize the fixed points of $P_{S_1} \circ P_{S_2}$.

(c) Despite the difficulty of proving the global convergence of (1.5) due to the nonconvexity of S_2, we prove in Proposition 3.16 a weaker result implying the impossibility of the iterates to be trapped in some particular region not containing a point in $S_1 \cap S_2$. Moreover, by utilizing the convergence theory of Attouch, Bolte and Svaiter (2013) for semi-algebraic and tame problems [27], we prove in Theorem 3.20 the global convergence to stationary points of a relaxed version of the iterations (1.5) given by

$$w^{k+1} \in (1-\gamma)P_{C_2}(w^k) + \gamma(P_{C_1} \circ P_{C_2})(w^k), \qquad \gamma \in (0,1). \tag{1.7}$$

That is, we take the convex combination of the iterates (1.5) with the mapping P_{C_2}. Although MAP iterations (1.5) are not covered by the above relaxation, we note that the former is the limiting case of (1.7) when $\gamma = 1$.

On the other hand, the numerical contributions of our work include the superior performance of our proposed algorithm MAP in solving randomly generated AVEs of the form (1.2) as compared with other methods from the four categories (a)–(d) of algorithms described above. We also derive from the MAP iterations a related algorithm called MAP-LS, which can serve as an alternative method for instances when a given AVE is difficult to solve by MAP. For the general AVE (1.1), we illustrate the dominant performance of MAP over the successive linearization algorithm in [14], which is the only algorithm we can compare our method with. Hence, our proposed algorithms have several merits from a numerical perspective, and are indeed an important contribution to the growing literature of AVE.

1.3. Outline

The structure of this paper is as follows: In Sect. 2, we characterize the fixed point sets of the alternating projections mapping. Next, we present the convergence analysis of the algorithms in Sect. 3. Finally, we illustrate the applicability of our approach through numerical experiments in Sect. 4.

2. Fixed points of the alternating projections map

This section is devoted to characterizing the set of fixed points of the alternating projections map. More precisely, we provide conditions on the matrices A and B which will allow us to determine which fixed points of $P_{S_1} \circ P_{S_2}$ belong to $S_1 \cap S_2$.

2.1. Change of variables

Instead of directly dealing with the sets S_1 and S_2 given by (1.4), we consider a change of variables which will reveal the close connection of the absolute value equation (1.1) and complementarity. More precisely, we obtain a feasibility problem equivalent to (1.3) involving two transformed sets C_1 and C_2, where C_1 is an affine set (see (2.3)) and C_2 is a complementarity set (see (2.4)). As we shall see in Sect. 2.4, such a change of variables will also help in our goal of characterizing fixed points in the case that $m = n$. Moreover, the relationship with the linear complementarity problem will also become more apparent with such a transformation (see Remark 2.18).

In general, we may consider any linear transformation $z = Rw$ where $z = (x, y)$, $w = (u, v)$ and $R \in \mathbb{R}^{2n\times 2n}$ is a unitary matrix. Letting $R = \begin{bmatrix} R_1 & R_2 \\ R_3 & R_4 \end{bmatrix}$

with $R_i \in \mathbb{R}^{n\times n}$, we see that $z \in S_i$ if and only if $w \in C_i$ for $i = 1, 2$, where

$$C_1 := \{w = (u, v) \in \mathbb{R}^n \times \mathbb{R}^n : \begin{bmatrix} AR_1 + BR_3 \ AR_2 + BR_4 \end{bmatrix} w = c\},$$

and

$$C_2 := \{w = (u, v) \in \mathbb{R}^n \times \mathbb{R}^n : |R_1 u + R_2 v| = R_3 u + R_4 v\}.$$

With these, AVE (1.1) is also equivalent to the feasibility problem

$$\text{find } w = (u, v) \in C_1 \cap C_2. \tag{FP}$$

Accordingly, we consider the MAP iterates given by

$$w^{k+1} \in (P_{C_1} \circ P_{C_2})(w^k). \tag{MAP}$$

Since we have chosen R to be a unitary matrix, it follows that for all $z = (x, y) \in \mathbb{R}^n \times \mathbb{R}^n$,

$$P_{S_i}(z) = RP_{C_i}(R^\mathsf{T} z), \quad i = 1, 2.$$

Consequently,

$$(P_{S_1} \circ P_{S_2})(z) = R\left((P_{C_1} \circ P_{C_2})(R^\mathsf{T} z)\right). \tag{2.1}$$

Thus, if z^0 is the initial point for the original MAP iterates (1.5) and we set $w^0 = R^\mathsf{T} z^0$ for the iterations (MAP), then $z^{k+1} = Rw^{k+1}$ for all $k \geq 0$. Moreover, we also have from (2.1) that

$$\operatorname{Fix}(P_{S_1} \circ P_{S_2}) = \{Rw : w \in \operatorname{Fix}(P_{C_1} \circ P_{C_2})\} = R \operatorname{Fix}(P_{C_1} \circ P_{C_2}). \tag{2.2}$$

In the sequel, all our analyses and results are based on the constraint sets induced by $R = \frac{1}{\sqrt{2}} \begin{bmatrix} I_n & -I_n \\ I_n & I_n \end{bmatrix}$.

Defining

$$T := \begin{bmatrix} A+B & -A+B \end{bmatrix} \in \mathbb{R}^{m\times 2n}, \tag{T}$$

it can be verified that C_1 is given by

$$C_1 = \{w \in \mathbb{R}^n \times \mathbb{R}^n \ : \ Tw = \sqrt{2}c\}. \tag{2.3}$$

On the other hand, $(u,v) \in C_2$ if and only if $|u-v| = u+v$. Using the fact that $t + |t| = 2t_+$ where $t_+ := \max(0,t)$ (with the maximum understood in the pointwise sense), we see that $|u-v| = u+v$ if and only if $u-(u-v)_+ = 0$. Then, C_2 reduces to

$$C_2 = \{w = (u,v) \in \mathbb{R}^n \times \mathbb{R}^n \ : \ u \geq 0, \ v \geq 0, \text{ and } \langle u, v\rangle = 0\}. \tag{2.4}$$

It follows that if x solves (1.1), then (u,v) with $u = \frac{1}{\sqrt{2}}x_+$ and $v = \frac{1}{\sqrt{2}}(-x)_+$ solves the feasibility problem (FP) with C_1 and C_2 given by (2.3) and (2.4), respectively. Conversely, if (u,v) solves (FP), then $x = \frac{1}{\sqrt{2}}(u-v)$ solves the AVE (1.1).

2.2. Projection formulas

Important for our subsequent analysis and numerical simulations are the exact formulas for the projections involved in the iterations (MAP). The projection onto the affine set C_1 is well known, which we recall in the following proposition.

Proposition 2.1. [28, Lemma 4.1] *Suppose that $c \in \mathbb{R}^m$ is in the range of T given by* (T)*. Then for any $w \in \mathbb{R}^n$, we have*

$$P_{C_1}(w) = w - T^\dagger(Tw - \sqrt{2}c),$$

where $T^\dagger$ is the Moore–Penrose inverse of T.

While P_{C_1} is a single-valued operator, the projection onto C_2 is not due to the nonconvexity of C_2.

Proposition 2.2. *Let $w = (u,v) \in \mathbb{R}^n \times \mathbb{R}^n$. Then $z \in P_{C_2}(w)$ if and only if for each $i = 1, \ldots, n$,*

$$(z_i, z_{n+i}) \in \begin{cases} \{(0,(v_i)_+)\} & u_i < v_i \\ \{((u_i)_+, 0)\} & u_i > v_i \ . \\ \{(0,(v_i)_+), ((u_i)_+, 0)\} & u_i = v_i \end{cases} \tag{2.5}$$

In particular, P_{C_2} is multivalued on $\{(u,v) : \exists i$ such that $u_i = v_i > 0\}$.

Proof. Fix $w = (u,v) \in \mathbb{R}^n \times \mathbb{R}^n$. To prove the result, we need to solve the minimization problem

$$\min_{\bar{w} \in C_2} \|\bar{w} - w\|^2.$$

Letting $\bar{w} = (\bar{u}, \bar{v}) \in \mathbb{R}^n \times \mathbb{R}^n$, we have

$$\|\bar{w} - w\|^2 = \sum_{i=1}^n (\bar{u}_i - u_i)^2 + \sum_{i=1}^n (\bar{v}_i - v_i)^2$$

$$= \sum_{i=1}^{n} \|(\bar{u}_i - u_i, \bar{v}_i - v_i)\|^2.$$

As the last expression is separable, we only need to consider the projection of an arbitrary point $(s,t) \in \mathbb{R}^2$ onto the set

$$M := \{(a,b) : a \geq 0,\ b \geq 0 \text{ and } ab = 0\}, \tag{2.6}$$

which can be easily calculated as

$$P_M(s,t) = \begin{cases} \{(0,t_+)\} & \text{if } s < t \\ \{(s_+,0)\} & \text{if } s > t\ . \\ \{(0,t_+),(s_+,0)\} & \text{if } s = t \end{cases} \tag{2.7}$$

This gives the formula (2.5). □

Example 2.3. Let $n = 3$ and $w = (u,v) \in \mathbb{R}^3 \times \mathbb{R}^3$ where $u = (2,-4,1)$ and $v = (1,3,1)$. Then $u_1 > v_1$, $u_2 < v_2$ and $u_3 = v_3$. By Eq. (2.5), we have $P_{C_2}(w) = \{(2,0,1,0,3,0),(2,0,0,0,3,1)\}$.

Note that because of the convexity of C_1, we know that the map P_{C_1} is firmly nonexpansive, i.e., $\|P_{C_1}(w) - P_{C_1}(w')\|^2 \leq \langle w - w', P_{C_1}(w) - P_{C_1}(w')\rangle$ for all $w, w' \in \mathbb{R}^n \times \mathbb{R}^n$. The same cannot be said for P_{C_2} due to the nonconvexity of C_2. However, P_{C_2} is firmly nonexpansive on some subsets of $\mathbb{R}^n \times \mathbb{R}^n$ as we will prove in Corollary 2.4. Before we present this result, we first introduce some notations which will be used for the remaining parts of this paper.

We denote by $\mathscr{T}$ the collection of all functions $\tau : \{1,2,\ldots,n\} \to \{1,2\}$, so that $|\mathscr{T}| = 2^n$. For each $\tau \in \mathscr{T}$, we let S_τ denote the set of all $w = (u,v) \in \mathbb{R}^n \times \mathbb{R}^n$ such that for each $i = 1,\ldots,n$, we have $(u_i,v_i) \in K_j$ if $\tau(i) = j$ for $j = 1,2$, where

$$K_1 := \{(a,b) \in \mathbb{R}^2 : a > b \text{ or } a = b \leq 0\} \quad \text{and} \tag{2.8}$$

$$K_2 := \{(a,b) \in \mathbb{R}^2 : a < b \text{ or } a = b \leq 0\}. \tag{2.9}$$

Observe that

$$\bigcup_{\tau \in \mathscr{T}} S_\tau = \mathbb{R}^n \times \mathbb{R}^n \backslash \{(u,v) : u_i = v_i > 0 \text{ for some } i\}.$$

For each $\tau \in \mathscr{T}$, we also let $R_\tau := S_\tau \cap C_2$ so that $C_2 = \bigcup_{\tau \in \mathscr{T}} R_\tau$. Thus, $w \in R_\tau$ if and only if for each $i = 1,\ldots,n$, we have $(u_i,v_i) \in M_j$ if $\tau(i) = j$ for $j = 1,2$ where

$$M_1 := \{(a,b) \in \mathbb{R}^2 : a \geq 0 \text{ and } b = 0\} \quad \text{and} \tag{2.10}$$

$$M_2 := \{(a,b) \in \mathbb{R}^2 : b \geq 0 \text{ and } a = 0\}. \tag{2.11}$$

Corollary 2.4. *P_{C_2} is firmly nonexpansive on S_τ for any $\tau \in \mathscr{T}$.*

Proof. First, we note that the restriction of P_M given by (2.7) to K_j is precisely the projection mapping P_{M_j}, where M_j is given by (2.10)–(2.11). Since M_j is convex, then P_{M_j} is firmly nonexpansive on K_j. It follows that P_M is firmly nonexpansive on K_j.

Given $\tau \in \mathscr{T}$, take two points $w = (u, v) \in S_\tau$ and $w' = (u', v') \in S_\tau$. Then the points (u_i, v_i) and (u'_i, v'_i) both lie on K_1 or K_2 for each $i = 1, \ldots, n$. Then by firm nonexpansiveness of P_M, we obtain

$$\begin{aligned} \|P_{C_2}(w) - P_{C_2}(w')\|^2 &= \sum_{i=1}^{n} \|P_M(u_i, v_i) - P_M(u'_i, v'_i)\|^2 \\ &\leq \sum_{i=1}^{n} \langle (u_i, v_i) - (u'_i, v'_i), P_M(u_i, v_i) - P_M(u'_i, v'_i)\rangle \\ &= \langle w - w', P_{C_2}(w) - P_{C_2}(w')\rangle. \end{aligned}$$

This proves the desired result. □

By invoking the fact that C_1 given by (2.3) is an affine set, the next proposition describes a property of the iterates (MAP) which is based on the following observation: When $n = 1$, the set C_1 defines a straight line in $\mathbb{R}^2$ provided that A and B are not both zero. Intuitively, one can see that if the line C_1 intersects C_2 but does not pass through the origin, then $P_{C_1}(w) \not\leq 0$ for any $w \in C_2$. For $n > 1$, we may conjecture that if $\bar{w} := P_{C_1}(w)$ with $w \in C_2$, we either have (i) $\bar{w} \notin \mathbb{R}^n_- \times \mathbb{R}^n_-$ or (ii) $(\bar{w}_i, \bar{w}_{n+i}) \notin \mathbb{R}^2_-$ for all $i = 1, \ldots, n$ for any $w \in C_2$, where $\mathbb{R}^n_-$ denotes the set of nonpositive vectors in $\mathbb{R}^n$. The following proposition indicates that (i) holds, and we illustrate in Example 2.6 that (ii) does not hold in general.

Proposition 2.5. *If $c \neq 0$, $C_1 \cap C_2 \neq \emptyset$ and $\{w^k\}_{k=0}^{\infty}$ is any sequence generated by (MAP), then $w^k \notin \mathbb{R}^n_- \times \mathbb{R}^n_-$ for all $k \geq 1$.*

Proof. It is enough to show that given a point $w \in C_2$, we have $\bar{w} := P_{C_1}(w) \notin \mathbb{R}^n_- \times \mathbb{R}^n_-$. Since C_1 is a convex set, we have

$$\langle w - \bar{w}, w' - \bar{w}\rangle \leq 0 \qquad \forall w' \in C_1.$$

In particular, we can take $w' = w^* \in C_1 \cap C_2$ and obtain

$$\langle w - \bar{w}, w^* - \bar{w}\rangle \leq 0. \tag{2.12}$$

Since $c \neq 0$ and $T\bar{w} = \sqrt{2}c$, then $\bar{w} \neq 0$. Thus, if $\bar{w} \in \mathbb{R}^n_- \times \mathbb{R}^n_-$, then there exists some $i \in \{1, 2 \ldots, 2n\}$ such that $\bar{w}_i < 0$. Since $w, w^* \geq 0$, we must have $w_i - \bar{w}_i > 0$ and $w^*_i - \bar{w}_i > 0$. Meanwhile, we also have that $w_j - \bar{w}_j \geq 0$ and $w^*_j - \bar{w}_j \geq 0$ for all j. In turn, we will obtain $\langle w - \bar{w}, w^* - \bar{w}\rangle > 0$ which contradicts (2.12). Hence, $\bar{w} \notin \mathbb{R}^n_- \times \mathbb{R}^n_-$ as desired. □

Example 2.6. Let $A = \begin{bmatrix} 3 & -8 \\ 3 & 0 \end{bmatrix}$, $B = -I_2$ and $c = (6, 9)/\sqrt{2}$. It can be verified that $C_1 \cap C_2 = \{(3/\sqrt{2}, 0, 0, 0)\}$. For $w = (0, 0, 1, 0) \in C_2$, one can check that $\bar{w} := P_{C_1}(w) \approx (1.8042, -0.5569, -0.7921, -0.6540)$. That is, $\bar{w} \notin \mathbb{R}^n_- \times \mathbb{R}^n_-$, illustrating Proposition 2.5. On the other hand, observe that $(\bar{w}_2, \bar{w}_4) \in \mathbb{R}^2_-$ indicating that statement (ii) in the discussion preceding Proposition 2.5 does not necessarily hold.

2.3. Characterization of fixed points for arbitrary m and n

We now provide a general condition which will allow us to distinguish which fixed points of $P_{C_1} \circ P_{C_2}$ belong to $C_1 \cap C_2$. In the following, we denote by $\mathrm{Ker}(T)$ and $\mathrm{Ran}(T)$ the kernel and range of T, respectively. Given any affine set $S \subseteq \mathbb{R}^n$, we denote its orthogonal complement by $S^\perp$. We also denote

$$\Omega := \{w = (u, v) \in \mathbb{R}^n \times \mathbb{R}^n : (u_i, v_i) \notin \mathbb{R}^2_{--} \ \forall i = 1, \ldots, n\}, \tag{2.13}$$

where $\mathbb{R}^n_{--}$ denotes the set of negative vectors in $\mathbb{R}^n$.

Theorem 2.7. (Characterization of fixed point sets for arbitrary m, n) *Let $T \in \mathbb{R}^{m \times 2n}$ be given by* (T) *and suppose that*

$$\mathrm{Ker}(T)^\perp \cap \hat{C}_2 = \{0\}, \tag{C}$$

where

$$\hat{C}_2 := \{w = (u, v) \in \mathbb{R}^n \times \mathbb{R}^n : u_i v_i = 0 \ \forall i = 1, \ldots, n\}. \tag{2.14}$$

Then for any $c \in \mathbb{R}^m$,

$$\mathrm{Fix}(P_{C_1} \circ P_{C_2}) \cap \Omega = C_1 \cap C_2.$$

Proof. We note first that if $c \notin \mathrm{Ran}(T)$, then $C_1 = \emptyset$. Since $\mathrm{Fix}(P_{C_1} \circ P_{C_2}) \subseteq C_1$, then $\mathrm{Fix}(P_{C_1} \circ P_{C_2}) = \emptyset$. Hence, the result necessarily holds. Suppose now that $c \in \mathrm{Ran}(T)$ so that $C_1 \neq \emptyset$. Since $C_1 \cap C_2 \subseteq \mathrm{Fix}(P_{C_1} \circ P_{C_2})$ and $C_2 \subseteq \Omega$, then $C_1 \cap C_2 \subseteq \mathrm{Fix}(P_{C_1} \circ P_{C_2}) \cap \Omega$. To prove the other inclusion, suppose that $w = (u, v) \in \mathrm{Fix}(P_{C_1} \circ P_{C_2}) \cap \Omega$. Since $w \in (P_{C_1} \circ P_{C_2})(w)$, then $w = P_{C_1}(w')$ for some $w' \in P_{C_2}(w)$. Since C_1 is an affine set, it follows that $w - w' \in \mathrm{Ker}(T)^\perp$.

We also have that $w \in \Omega$ so that we may partition its components using the following index sets:

$$\begin{aligned} I &:= \{i \in \{1, 2, \ldots, n\} : u_i > v_i \text{ and } u_i \geq 0\}, \\ J &:= \{i \in \{1, 2, \ldots, n\} : u_i = v_i \geq 0\}, \\ K &:= \{i \in \{1, 2, \ldots, n\} : u_i < v_i \text{ and } v_i \geq 0\}. \end{aligned}$$

By rearranging the columns of A and B if necessary, we may suppose that $u = (u_I, u_J, u_K) \in \mathbb{R}^n$ where u_Λ denotes the components of u indexed by $\Lambda \in \{I, J, K\}$. Accordingly, we let $v = (v_I, v_J, v_K) \in \mathbb{R}^n$. Consequently, we have from Proposition 2.2 that $w' = (u', v')$ where $u' = (u_I, u'_J, 0_{|K|})$ and $v' = (0_{|I|}, v'_J, v_K)$ with $(u'_j, v'_j) \in \{(u_j, 0), (0, v_j)\}$. Then $(w - w')_i (w - w')_{n+i} = 0$ for all $i = 1, \ldots, n$, that is, we have $w - w' \in \hat{C}_2$.

To summarize, we have shown that $w - w' \in \mathrm{Ker}(T)^\perp \cap \hat{C}_2$. By (C), it follows that $w = w'$. Hence, $w \in C_2$. This completes the proof. □

The geometric interpretation of condition (C) in terms of normal cones is discussed in Sect. 3.3. Note, however, that this condition is not easy to verify for the case $m \neq n$. In the next subsection where we discuss the case $m = n$, we use the notion of nondegenerate matrices to provide an easier-to-verify condition on A and B that will result to a map T that satisfies (C).

Moreover, we also note the following observation: From Proposition 2.5, we see that $\mathrm{Fix}(P_{C_1} \circ P_{C_2}) \subseteq (\mathbb{R}^n \times \mathbb{R}^n) \backslash (\mathbb{R}^n_{--} \times \mathbb{R}^n_{--})$ and note that

$$\Omega \subseteq (\mathbb{R}^n \times \mathbb{R}^n) \backslash (\mathbb{R}^n_{--} \times \mathbb{R}^n_{--}).$$

Hence, it is not necessary that $\mathrm{Fix}(P_{C_1} \circ P_{C_2}) \subseteq \Omega$. In Example 2.15, we illustrate the importance of intersecting the set of fixed points with the set Ω. For the case $m = n$, we also provide a sufficient condition so that the set of fixed points is necessarily contained in Ω (see Theorem 2.17), which in turn implies that $\mathrm{Fix}(P_{C_1} \circ P_{C_2})$ is equal to $C_1 \cap C_2$.

2.4. Characterization of fixed points for $m = n$

To prove our main result for the case $m = n$, we first establish some key lemmas which are prerequisites to obtaining conditions on A and B that will imply (C).

Lemma 2.8. *Let $T \in \mathbb{R}^{m \times 2n}$ be given by (T). Suppose that at least one among the matrices A, B, $A+B$ and $A-B$ is of full row rank. Then* $\mathrm{rank}(T) = m$. *In particular,* $\mathrm{rank}(T) = n$ *when $m = n$ and $B = -I_n$.*

Proof. If either $A+B$ or $A-B$ has rank m, then it is clear that $\mathrm{rank}(T) = m$. For the other cases when either A or B is of full row rank, we note that $\mathrm{rank}(T) = \mathrm{rank}(TT^\top) = \mathrm{rank}(2(AA^\top + BB^\top))$. The matrices $AA^\top$ and $BB^\top$ are positive semidefinite, and at least one of them is positive definite by our full rank assumption. Hence, the claim of the lemma follows. □

The next lemma precisely describes the elements of the set $\mathrm{Ker}(T)^\perp$.

Lemma 2.9. *Suppose $m = n$ and $T \in \mathbb{R}^{n \times 2n}$ is given by (T), and suppose that $A - B$ is nonsingular. Then* $\mathrm{Ker}(T)^\perp = \mathrm{Ker}(\begin{bmatrix} I_n & Q \end{bmatrix})$, *where*

$$Q := (A^\top + B^\top)(A^\top - B^\top)^{-1}. \tag{2.15}$$

Proof. Let $w = (u, v) \in \mathrm{Ker}(T)^\perp = \mathrm{Ran}(T^\top)$. Then there exists $x \in \mathbb{R}^n$ such that $T^\top x = w$. It follows that $u = (A^\top + B^\top)x$ and $v = -(A^\top - B^\top)x$. By invertibility of $A - B$, we see that

$$u = (A^\top + B^\top)x = -(A^\top + B^\top)(A^\top - B^\top)^{-1}v = -Qv.$$

Hence, we have $\mathrm{Ker}(T)^\perp \subseteq \mathrm{Ker}(\begin{bmatrix} I_n & Q \end{bmatrix})$. Meanwhile, we also have that

$$\dim(\mathrm{Ker}(\begin{bmatrix} I_n & Q \end{bmatrix})) = 2n - \mathrm{rank}(\begin{bmatrix} I_n & Q \end{bmatrix}) = n$$

by the rank-nullity theorem, and

$$\dim(\mathrm{Ker}(T)^\perp) = \mathrm{rank}(T^\top) = \mathrm{rank}(T) = n,$$

by Lemma 2.8. Thus,

$$\dim(\mathrm{Ker}(\begin{bmatrix} I_n & Q \end{bmatrix})) = \dim(\mathrm{Ker}(T)^\perp).$$

With these, we conclude that $\mathrm{Ker}(T)^\perp = \mathrm{Ker}(\begin{bmatrix} I_n & Q \end{bmatrix})$. □

Note that we can derive a result similar to Lemma 2.9 if we rather assume that $A + B$ is nonsingular.

Now that we have described the set $\mathrm{Ker}(T)^{\perp}$, we next focus on finding conditions which will imply (C). Nondegenerate matrices, as defined below, will play a major role in our analysis.

Definition 2.10. [13] A matrix $Q \in \mathbb{R}^{n\times n}$ is nondegenerate if all its principal minors are nonzero, i.e., the principal submatrix $Q_{\Lambda\Lambda}$ is nonsingular for all $\Lambda \subseteq \{1,\ldots,n\}$. We call Q degenerate if it is not a nondegenerate matrix.

Lemma 2.11. *Let $A, Q \in \mathbb{R}^{n\times n}$ be nonsingular matrices where Q is a nondegenerate matrix, and let $B := AQ$. Let $\Lambda \subseteq \{1,\ldots,n\}$ and let A' be the $n \times n$ matrix obtained by replacing the columns of A indexed by Λ by those columns of B indexed by Λ. Then A' is nonsingular.*

Proof. Without loss of generality, assume that $\Lambda = \{1, 2, \ldots, k\}$ with $k \leq n$. Denote the columns of A by $\{v_1, v_2, \ldots, v_n\}$ and the columns of B by $\{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n\}$. To prove the claim, we only need to show that $\{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k, v_{k+1}, \ldots, v_n\}$ is linearly independent.

Suppose that $\sum_{j=1}^{k} a_j\bar{v}_j + \sum_{j=k+1}^{n} a_jv_j = 0$ for some constants $a_1, \ldots, a_n$. By the definition of B, we have $\bar{v}_j = \sum_{i=1}^{n} q_{ij}v_i$ for all j, where q_{ij} is the (i, j)-entry of Q. Direct computations lead us to

$$\left(\sum_{j=1}^{k} a_jq_{1j}\right) v_1 + \cdots + \left(\sum_{j=1}^{k} a_jq_{kj}\right) v_k + \left(a_{k+1} + \sum_{j=1}^{k} a_jq_{k+1,j}\right) v_{k+1}$$
$$+ \cdots + \left(a_n + \sum_{j=1}^{k} a_jq_{nj}\right) v_n = 0.$$

Since the v_is are linearly independent, all the coefficients above should be equal to zero. From the first k terms, we obtain that $Q_{\Lambda\Lambda}(a_1, \ldots, a_k)^{\mathsf{T}} = 0$. Since $Q_{\Lambda\Lambda}$ is nonsingular by nondegeneracy of Q, then $a_j = 0$ for all $j = 1, \ldots, k$ which consequently gives $a_j = 0$ for all $j > k$. □

Proposition 2.12. *Let $m = n$ and suppose that the matrix Q defined by (2.15) is nondegenerate. Let $\Lambda_1 \subseteq \{1, \ldots, n\}$ and $\Lambda_2 = \{n + i : i \notin \Lambda_1\}$. Then the columns of $[I_n\ Q]$ indexed by $\Lambda_1 \cup \Lambda_2$ are linearly independent. Consequently, condition (C) holds.*

Proof. Set $A = I_n$ and $\Lambda = \{1, \ldots, n\}\backslash\Lambda_1$. Then the columns of matrix A' described in Lemma 2.11 are precisely the columns of $D := [I_n\ Q]$ indexed by $\Lambda_1 \cup \Lambda_2$. Consequently, A' is nonsingular and so the first claim of the proposition holds.

If $w = (u, v) \in \mathrm{Ker}(T)^{\perp} \cap \hat{C}_2$, then since $\mathrm{Ker}(T)^{\perp} = \mathrm{Ker}(D)$ by Lemma 2.9, we have

$$0 = Dw = \sum_{i\in\Lambda_1} u_id_i + \sum_{i\in\Lambda_2'} v_id_i, \tag{2.16}$$

where $d_i \in \mathbb{R}^n$ is the ith column of D, $\Lambda_1 := \{i : u_i \neq 0\}$ and $\Lambda_2' := \{i : v_i \neq 0\}$. In other words, the right-hand side of (2.16) is a linear combination of the columns of D indexed by

$$\Lambda := \Lambda_1 \cup \{n + i : i \in \Lambda_2'\} \subseteq \Lambda_1 \cup \Lambda_2.$$

Thus, the columns indexed by Λ must be linearly independent, i.e., $\Lambda_1 = \Lambda_2' = \emptyset$ so that $w = 0$. □

As an immediate consequence of the above result and Theorem 2.7, we have the following.

Theorem 2.13. (Characterization of fixed point sets for $m = n$) *Let $m = n$. Suppose that Q given by (2.15) is a nondegenerate matrix and Ω is given by (2.13). Then for any $c \in \mathbb{R}^n$,*

$$\operatorname{Fix}(P_{C_1} \circ P_{C_2}) \cap \Omega = C_1 \cap C_2.$$

Proof. Since Q is nondegenerate, we have from Proposition 2.12 that condition (C) holds. Hence, the claim follows from Theorem 2.7. □

Remark 2.14. One can guarantee that the matrix Q given by (2.15) is nondegenerate if $\sigma_{\min}(A) > \sigma_{\max}(B)$, i.e. the smallest singular value of A is greater than the largest singular value of B. To see this, note that for all $x \in \mathbb{R}^n$,

$$\begin{aligned} x^\top(A^\top + B^\top)(A^\top - B^\top)^{-1}x &= y^\top(A - B)(A^\top + B^\top)y, \; y = (A^\top - B^\top)^{-1}x \\ &= y^\top(AA^\top - BA^\top + AB^\top - BB^\top)y \\ &= y^\top(AA^\top - BB^\top)y \\ &\geq (\lambda_{\min}(AA^\top) - \lambda_{\max}(BB^\top))\|y\|^2 \\ &= (\sigma_{\min}(A) - \sigma_{\max}(B))\|y\|^2, \end{aligned}$$

where the third equality follows from $y^\top BA^\top y = y^\top AB^\top y$. It follows that Q is a positive definite matrix, which is necessarily nondegenerate. By a similar computation, the condition $\sigma_{\max}(A) < \sigma_{\min}(B)$ implies nondegeneracy of Q. ■

The following example demonstrates the importance of nondegeneracy of Q as well as the significance of intersecting the set of fixed points with Ω.

Example 2.15.

1 Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $B = -I_2$ and $c = (-10, -19)/\sqrt{2}$. Then $Q = \begin{bmatrix} -1.5 & 1.5 \\ 1 & 0 \end{bmatrix}$, which is a degenerate matrix. Moreover, it can be verified that $w = (-0.9231, 4.7026, 9.0872, 0.6154) \in \operatorname{Fix}(P_{C_1} \circ P_{C_2}) \cap \Omega$. Clearly, however, $w \notin C_1 \cap C_2$. We note that the problem is feasible, i.e. $C_1 \cap C_2 \neq \emptyset$. For instance, both $(0, 0, 3, 2)/\sqrt{2}$ and $(2, 0, 0, 5)/\sqrt{2}$ are solutions of the feasibility problem.

2 Let $A = 1/2$, $B = 3/2$ and $c = -\sqrt{2}$. Then $Q = -2$, which is nondegenerate. Moreover, $C_1 \cap C_2 = \emptyset$ and $\operatorname{Fix}(P_{C_1} \circ P_{C_2}) = \{(-0.8, -0.4)\}$ so that $C_1 \cap C_2 \neq \operatorname{Fix}(P_{C_1} \circ P_{C_2})$. Nevertheless, we see that $\operatorname{Fix}(P_{C_1} \circ P_{C_2}) \cap \Omega = C_1 \cap C_2$.

It turns out that the signs of the principal minors of Q play an important role in characterizing the fixed points of the alternating projections map. In particular, the fixed points are necessarily contained in the set Ω given by (2.13) if all the principal minors of Q are positive. Such a matrix is called a P-matrix [13]. In contrast, we see from Example 2.15.2 that intersecting the set of fixed points with Ω is necessary if there exists a negative principal minor.

To characterize the set of fixed points for the P-matrix case, we need the following lemma.

Lemma 2.16. [13, Theorem 3.3.4] *$Q \in \mathbb{R}^{n\times n}$ is a P-matrix (i.e., all of its principal minors are positive) if and only if whenever $x_i(Qx)_i \leq 0$ for all $i = 1, \ldots, n$, we have $x = 0$.*

Theorem 2.17. *Let $m = n$. Suppose that Q given by (2.15) is a P-matrix. Then for any $c \in \mathbb{R}^n$, we have*

$$\mathrm{Fix}(P_{C_1} \circ P_{C_2}) = C_1 \cap C_2.$$

In particular, by Remark 2.14, the above equality holds if $\sigma_{\min}(A) > \sigma_{\max}(B)$.

Proof. Suppose that $w \in (P_{C_1} \circ P_{C_2})(w)$. As in the proof of Theorem 2.7, we have $w - w' \in \mathrm{Ker}(T)^{\perp}$ where $w' \in P_{C_2}(w)$. Since Q is nondegenerate, $A - B$ is necessarily nonsingular so that by Lemma 2.9, $w - w' \in \mathrm{Ker}([I_n\ Q])$, i.e.

$$u - u' + Q(v - v') = 0. \tag{2.17}$$

Observe that to prove the desired result, Theorem 2.13 implies that it is enough to prove that $w \in \Omega$, i.e. $(u_i, v_i) \notin \mathbb{R}^2_{--}$ for all i. Suppose to the contrary that there exists an index j such that $u_j, v_j < 0$. Then from Proposition 2.2, we know that $u'_j = v'_j = 0$ so that $v_j - v'_j < 0$. In particular, $v - v'$ is a nonzero vector. Consequently, by Lemma 2.16, there exists some l such that

$$(v_l - v'_l)(Q(v - v'))_l > 0. \tag{2.18}$$

We consider two cases:

(i) Suppose that $v_l - v'_l > 0$. From Proposition 2.2, this can only happen if $u_l \geq v_l$ and $(u'_l, v'_l) = (u_l, 0)$. Thus, we obtain that $u_l - u'_l = 0$. From Eq. (2.17), it follows that $(Q(v - v'))_l = 0$. This is a contradiction to (2.18).

(ii) Suppose that $v_l - v'_l < 0$. We conclude from Proposition 2.2 that $v_l < 0$ and $v'_l = 0$. Moreover, it also follows from the same proposition that $u_l - u'_l \leq 0$. From Eq. (2.17), it must be the case that $(Q(v - v'))_l \geq 0$. Hence, $(v_l - v'_l)(Q(v - v'))_l \leq 0$. However, this is a direct contradiction to (2.18).

Hence, it is impossible that there exists j such that $(u_j, v_j) \in \mathbb{R}^2_{--}$, i.e., $w \in \Omega$. This completes the proof. □

Remark 2.18. If $A - B$ is nonsingular, then the feasibility problem (FP) is equivalent to solving the system

$$u \geq 0, \quad F(u) := Q^{\mathsf{T}} u - \sqrt{2}(A - B)^{-1} c \geq 0, \quad \text{and} \quad \langle u, F(u)\rangle = 0,$$

known in the literature as a *linear complementarity problem* (LCP). The above LCP has a unique solution for all $c \in \mathbb{R}^n$ if and only if Q is a P-matrix [13]. Thus, Theorem 2.17 indicates that if Q is a P-matrix, then for any $c \in \mathbb{R}^n$, the set of fixed points of $P_{C_1} \circ P_{C_2}$ consists of a single point, which is precisely the solution of the feasibility problem (FP). ■

The next result, which is a very special case, provides another condition for the equality of the set of fixed points and the intersection of C_1 and C_2.

Theorem 2.19. *Suppose that $C_1 \cap R_\tau \neq \emptyset$ for all $\tau \in \mathscr{T}$. Then*

$$\mathrm{Fix}(P_{C_1} \circ P_{C_2}) = C_1 \cap C_2.$$

Proof. Suppose $w = P_{C_1}(w')$ where $w' \in P_{C_2}(w)$. Choose $\tau \in \mathscr{T}$ such that $w' \in R_\tau$, so that $w' = P_{R_\tau}(w)$. Taking $w^* \in C_1 \cap R_\tau$ and using the convexity of C_1 and R_τ, we obtain $\langle w' - w, w^* - w \rangle \leq 0$ and $\langle w - w', w^* - w' \rangle \leq 0$, respectively. Adding these two inequalities, we see that $\|w' - w\|^2 \leq 0$ so that $w = w'$ and therefore $w \in C_1 \cap C_2$. The other inclusion is trivial, and thus, the proof is complete. □

As a consequence, we state the following corollary whose hypothesis is the setting considered in [9].

Corollary 2.20. *Let $A \in \mathbb{R}^{n \times n}$, $B = -I_n$ and $c < 0$. If $\|A\|_\infty < \frac{\alpha}{2}$ where $\alpha = \frac{\min_i |c_i|}{\max_i |c_i|}$, then*

$$\mathrm{Fix}(P_{C_1} \circ P_{C_2}) = C_1 \cap C_2.$$

Proof. From [9, Proposition 6], we know that the AVE (1.1) has exactly 2^n distinct solutions, each of which has no zero components and has different sign pattern. Thus, each R_τ contains a point in $C_1 \cap C_2$ in its interior. The claim then follows from Theorem 2.19. □

3. Convergence analysis

In this section, we discuss the convergence issues related to the proposed method of alternating projections. In Sect. 3.1, we present some local convergence results which are direct consequences of the theory developed in [24]. We present an alternative local convergence analysis in Sect. 3.2 through the use of a new complementarity function. A by-product of this alternative analysis is the global convergence of MAP for homogeneous AVE. In addition, we also prove in Sect. 3.2 that under a nondegeneracy assumption, the MAP iterates cannot be trapped in some region S_τ (defined in Sect. 2.2) if S_τ does not contain a solution of the feasibility problem (FP). In Sect. 3.3, we establish the local linear rate of convergence of MAP. A globally convergent relaxation of MAP is presented in 3.4. Finally, another algorithm derived from the fixed point relation $w \in (P_{C_1} \circ P_{C_2})(w)$ is described in Sect. 3.5.

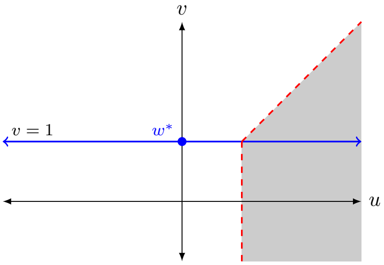

FIGURE 1. The method of alternating projections converges to a fixed point of $P_{C_1} \circ P_{C_2}$ when the initial point w^0 lies on the gray region. However, the convergence point is not the solution w^* of the feasibility problem. If w^0 lies on the red dashed line, the convergence to w^* depends on the selected element of $P_{C_2}(w^0)$

3.1. Convergence of MAP

The method of alternating projections and its generalization to more than two sets are globally convergent when the involved sets are convex [29]. For our problem (FP), the set C_1 is affine (hence, convex) while C_2 is a nonconvex set. Nevertheless, C_2 is a *union convex set*, i.e., it can be expressed as a finite union of closed convex sets [24]. In particular, we can write C_2 as $C_2 = \bigcup_{\tau \in \mathscr{T}} R_\tau$ where each R_τ is a closed convex set as defined in the preceding section. Thus, the local convergence of MAP is a direct consequence of [24, Corollary 6.2].

Theorem 3.1. (Local convergence of MAP) *Suppose $w^* \in C_1 \cap C_2$. Then there exists sufficiently small $\delta > 0$ such that for any w^0 with $\|w^0 - w^*\| < \delta$, any sequence generated by* (MAP) *converges to a solution of* (FP).

On the other hand, the global convergence of MAP to solutions of the feasibility problem (FP) is not always guaranteed.

Example 3.2. In Example 2.15.1, the iterates (MAP) may converge to a fixed point of $P_{C_1} \circ P_{C_2}$ that does not belong to $C_1 \cap C_2$. For instance, if we set $w^0 = (-1, 5, 9, 1)$, it can be verified that w^k converges to the point $w = (-0.9231, 4.8077, 9.1923, 0.6154) \in \mathrm{Fix}(P_{C_1} \circ P_{C_2}) \backslash (C_1 \cap C_2)$.

In fact, the following example shows that unique solvability does not imply global convergence.

Example 3.3. Let $A = 1$, $B = -1$ and $c = -2/\sqrt{2}$. Then, C_1 is the horizontal line $v = 1$ while C_2 is the union of the nonnegative u and v axes. In Fig. 1, we see that $w^* = (0, 1)$ is the unique solution to (FP). Meanwhile, MAP is not globally convergent to w^*.

In both of the examples above, we note that the matrix Q defined by (2.15) is degenerate. This suggests that for the case $m = n$, nondegeneracy of Q may be a necessary condition for global convergence to $C_1 \cap C_2$. We leave this as a conjecture which is worth further investigation. Note, however, that nondegeneracy is not sufficient for global convergence to solutions (for example, see Example 2.15.2).

We close this section by identifying two specific instances when the method of alternating projections is globally convergent.

Proposition 3.4. *Suppose $T \in \mathbb{R}^{m\times 2n}$ has full column rank. Then the feasibility problem* (FP) *has a solution if and only if $TT^{\dagger}c = c$ and $\sqrt{2}T^{\dagger}c \in C_2$. In particular, $\sqrt{2}T^{\dagger}c$ is the unique solution to* (FP) *whenever a solution exists. Moreover, any sequence generated by* (MAP) *converges finitely to $\sqrt{2}T^{\dagger}c$ (after one iteration).*

Proof. If $C_1 \cap C_2 \neq \emptyset$, then there exists $w^* \in C_1$ so that by Proposition 2.1, $w^* = w^* - T^{\dagger}(Tw^* - \sqrt{2}c)$. Since T has full column rank, then $T^{\dagger}T = I_{2n}$. Thus, $w^* = \sqrt{2}T^{\dagger}c$ is the unique point in C_1 and $\sqrt{2}c = Tw^* = \sqrt{2}TT^{\dagger}c$. Moreover, since $C_1 \cap C_2$ is nonempty, then w^* must be in C_2, i.e $\sqrt{2}T^{\dagger}c \in C_2$. Conversely, $TT^{\dagger}c = c$ and $\sqrt{2}T^{\dagger}c \in C_2$ implies that $\sqrt{2}T^{\dagger}c \in C_1 \cap C_2$. The convergence of any sequence generated by (MAP) is an immediate consequence of Proposition 2.1. In particular, $w^k = \sqrt{2}T^{\dagger}c$ for all $k \geq 1$ given any initial point $w^0 \in \mathbb{R}^n \times \mathbb{R}^n$. □

Another specific case when we obtain global convergence can be obtained when $0 \in C_1 \cap C_2$.

Proposition 3.5. *If $c = 0 \in \mathbb{R}^m$, then any sequence generated by* (MAP) *converges to a point in* $\mathrm{Fix}(P_{C_1} \circ P_{C_2})$.

Proof. This is a direct consequence of [24, Corollary 4.3]. □

A result stronger than the above proposition is derived in the next section (see Remark 3.13). In particular, we shall see that any sequence of MAP iterates generated by (MAP) will always converge to a point in $C_1 \cap C_2$ for any initial point w^0 whenever $c = 0$. That is, MAP is globally convergent to a solution of the feasibility problem (FP) for homogeneous AVEs.

3.2. Convergence analysis using a new C-function

We now provide an alternative convergence analysis for the method of alternating projections by introducing a C-function that is new to the literature. We recall first the notion of C-functions.

Definition 3.6. A function $\phi : \mathbb{R}^2 \to \mathbb{R}$ is called a *complementarity function* (or a *C-function*) if its zeros are precisely the points on the nonnegative axes, i.e.,

$$\phi(s,t) = 0 \quad \Longleftrightarrow \quad s \geq 0,\ t \geq 0,\ \text{and}\ st = 0.$$

There are several examples of C-functions [30,31], as well as methods to construct these functions [32]. Popular choices include the natural residual (NR) function and the Fischer–Burmeister (FB) function given, respectively, by

$$\phi_{\mathrm{NR}}(s,t) = \min(s,t) \qquad \text{and} \qquad \phi_{\mathrm{FB}}(s,t) = \sqrt{s^2+t^2} - (s+t).$$

Given any C-function ϕ, we define $\Phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ as

$$\Phi(u, v) := \begin{pmatrix} \phi(u_1, v_1) \\ \vdots \\ \phi(u_n, v_n) \end{pmatrix}.$$

It is then easy to see that

$$\begin{aligned} (u^*, v^*) \in C_2 \quad & \Longleftrightarrow \quad \Phi(u^*, v^*) = 0 \\ & \Longleftrightarrow \quad (u^*, v^*) \in \underset{(u,v)\in\mathbb{R}^n\times\mathbb{R}^n}{\arg\min} \; \Psi(w) := \frac{1}{2}\|\Phi(u, v)\|^2. \end{aligned}$$

Consequently, the feasibility problem (FP) can be equivalently reformulated as a constrained minimization problem

$$\min_{w\in C_1} \Psi(w), \tag{3.1}$$

provided that $C_1 \cap C_2 \neq \emptyset$. Note that if we define $\psi := \frac{1}{2}\phi^2$, then ψ is also a C-function and $\Psi(w) = \sum_{i=1}^n \psi(u_i, v_i)$.

Although different C-functions yield different formulations (3.1), a suitable choice of ϕ (or ψ) can facilitate the convergence analysis of MAP. Inspired by the equivalence of the method of alternating projections and the projected gradient method in the case of sparse affine feasibility problem as discussed in [33], we aim to choose a suitable C-function ψ such that the induced function Ψ satisfies

$$(P_{C_1} \circ P_{C_2})(w) = P_{C_1}(w - \nabla\Psi(w)),$$

(where Ψ should be differentiable to begin with). Unfortunately, P_{C_2} is multivalued as shown in Proposition 2.2 while the right-hand side of the above equation is single-valued. Thus, we instead find a C-function which induces a function Ψ satisfying

$$(P_{C_1} \circ P_{C_2})(w) \subseteq P_{C_1}(w - \partial\Psi(w)), \tag{3.2}$$

where $\partial\Psi(w)$ denotes the Clarke generalized gradient of $\Psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ at w.

Definition 3.7. [34] Let $F : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz continuous on $\mathbb{R}^n$.

(a) The *B-subdifferential* of F at x, denoted by $\partial_B F(x)$ is given by

$$\partial_B F(x) := \left\{ \lim_{x^k\to x} \nabla F(x^k) \;:\; F \text{ is differentiable at } x^k \in \mathbb{R}^n \right\}.$$

(b) The *Clarke generalized gradient* of F at a point $x \in \mathbb{R}^n$, denoted by $\partial F(x)$, is defined as the convex hull of $\partial_B F(x)$.

In the next example, we illustrate that the NR and FB functions do not satisfy condition (3.2).

Example 3.8. Let $n = 1$, $A = 1$, $B = 0$, $c = 0$ and consider the point $w = (-1, -1)$. Denote the function Ψ induced by the NR and FB functions by Ψ_{NR} and Ψ_{FB}, respectively. From Definition 3.7, one can verify that

$$\partial_B \Psi_{\text{NR}}(w) = \{(-1, 0), (0, -1)\} \quad \text{and} \quad \partial_B \Psi_{\text{FB}}(w) = \{(-3 - 2\sqrt{2}, -3 - 2\sqrt{2})\}.$$

Thus,

$$P_{C_1}(w - \partial\Psi_{\text{NR}}(w)) = \{(-0.5, -0.5)\},\ P_{C_1}(w - \partial\Psi_{\text{FB}}(w)) \\ = \{(2+2\sqrt{2}, 2+2\sqrt{2})\}.$$

Meanwhile, we have $(P_{C_1} \circ P_{C_2})(w) = \{(0,0)\}$.

In the following result, we propose a C-function that is new to the literature and gives the desired inclusion (3.2).

Proposition 3.9. *The function defined by*

$$\psi(s,t) = \begin{cases} \dfrac{s^2}{2} + \dfrac{(-t)_+^2}{2} & \text{if } s \le t, \\ \dfrac{t^2}{2} + \dfrac{(-s)_+^2}{2} & \text{if } s > t \end{cases} = \frac{\min(s,t)^2}{2} + \frac{\max(-\max(s,t),0)^2}{2} \tag{3.3}$$

is a nonnegative C-function. Moreover, ψ is differentiable on $K_1 \cup K_2$, where K_1 and K_2 are given by (2.8) and (2.9), respectively, and the B-subdifferential of ψ is given by

$$\partial_B\psi(s,t) = \begin{cases} \{(s, -(-t)_+)\} & \text{if } s < t \text{ or } s = t \le 0 \\ \{(-(-s)_+, t)\} & \text{if } s > t \\ \{(s,0),(0,t)\} & \text{if } s = t > 0. \end{cases} \tag{3.4}$$

Proof. Due to the symmetry of ψ (that is, $\psi(s,t) = \psi(t,s)$), we only need to verify the equivalence in Definition 3.6 for $s \le t$. In this case,

$$\psi(s,t) = 0 \quad \Longleftrightarrow \quad s = 0 \text{ and } (-t)_+ = 0 \quad \Longleftrightarrow \quad s = 0 \text{ and } t \ge 0.$$

This proves that ψ is a C-function. It can also be verified that ψ is locally Lipschitz continuous on $\mathbb{R}^2$ (see also [34, Lemma 4.6.1] or [35, Proposition 4.1.2]). Next, note that ψ is differentiable only on $K_1 \cup K_2$. The first two cases in formula (3.4) can be easily verified. If $s = t > 0$ and $\{(s^k, t^k)\}_{k=1}^\infty$ is a sequence in $K_1 \cup K_2$ converging to (s,t), then for sufficiently large k, the sequence lie in $\mathbb{R}^2_{++}$. Hence, the only subsequential limits of $\{\nabla\psi(s^k,t^k)\}_{k=1}^\infty$ are the limits of $\{(s^k,0)\}_{k=1}^\infty$ and $\{(0,t^k)\}_{k=1}^\infty$, which are $(s,0)$ and $(0,t)$, respectively. This completes the proof. □

We next show that the induced function $\Psi(w)$ of (3.3) indeed gives the desired inclusion (3.2). In fact, the following corollary shows that the MAP iterates (MAP) are the same as the "projected B-subdifferential" iterates.

Corollary 3.10. *If ψ is given by (3.3) and $\Psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ is given by $\Psi(w) := \sum_{i=1}^n \psi(u_i, v_i)$, then*

$$P_{C_2}(w) = w - \partial_B\Psi(w).$$

In particular, (3.2) holds.

Proof. Denote $w = (u,v) \in \mathbb{R}^n \times \mathbb{R}^n$. A direct verification shows that $\partial\Psi_B(w) = \sum_{i=1}^n D^i$, where the summation denotes the Minkowski sum of sets, and $D^i \subseteq \mathbb{R}^n \times \mathbb{R}^n$ denotes the set of all d^i such that

$$(d^i_j, d^i_{n+j}) \in \begin{cases} \{(0,0)\} & \text{if } j \neq i \\ \partial_B\psi(u_i, v_i) & \text{if } j = i \end{cases}, \qquad i = 1, 2, \ldots, n$$

with $\partial_B\psi(u_i,v_i)$ given by (3.4). To establish the result, we only need to show that $P_M(s,t) = (s,t) - \partial_B\psi(s,t)$, where P_M is given by (2.7). This equality can be directly verified by using the fact that $x + (-x)_+ = x_+$ for all $x \in \mathbb{R}$. □

We next establish one more important property of ψ as defined in (3.3) which will later be useful in proving our convergence result.

Lemma 3.11. *Let ψ be given by (3.3) and let $(a,b),(s,t) \in \mathbb{R}^2$. If $(a,b),(s,t) \in K_1$ or $(a,b),(s,t) \in K_2$, where K_1 and K_2 are given by (2.8) and (2.9), respectively, then*

$$\psi(s,t) - \psi(a,b) \geq \frac{1}{2}\langle \nabla\psi(a,b), (s-a, t-b)\rangle \\ -\frac{\min(a,b)^2}{8} - \frac{\max(a,b)^2}{8}\mathbb{1}_{\mathbb{R}^2_-}(\max(a,b), \max(s,t)), \tag{3.5}$$

where $\mathbb{1}_{\mathbb{R}^2_-}(c,d) = 1$ if $(c,d) \in \mathbb{R}^2_-$ and 0 otherwise. In particular, if $\psi(s,t) = 0$, then

$$2\psi(a,b) \leq \langle \nabla\psi(a,b), (a-s, b-t)\rangle. \tag{3.6}$$

Moreover,

$$2\psi(a,b) \leq \langle \psi'(a,b), (a,b)\rangle \qquad \forall (a,b) \in \mathbb{R}^2,\ \forall \psi'(a,b) \in \partial_B\psi(a,b). \tag{3.7}$$

Proof. By symmetry of ψ, it suffices to consider the case when $(a,b),(s,t) \in K_1$ to prove (3.5). By direct computation, we get from (3.3) and (3.4) that

$$\psi(s,t) - \psi(a,b) - \frac{1}{2}\langle \nabla\psi(a,b), (s-a,t-b)\rangle = \frac{t^2}{2} - \frac{bt}{2} + \frac{(-s)_+^2}{2} + \frac{(-a)_+ s}{2}.$$

Noting that $t^2 - bt \geq -b^2/4$ and $s^2 - as \geq -a^2/4$, we get the desired inequality. On the other hand, (3.6) directly follows from (3.6). Finally, in view of (3.6), we only need to verify inequality (3.7) for $a = b > 0$ which is a routine calculation. □

We now present our convergence result using the C-function (3.3).

Theorem 3.12. *Let $\{w^k\}_{k=0}^\infty$ be any sequence generated by (MAP). Let $w^* = (u^*, v^*) \in C_1 \cap C_2$, and denote*

$$I_1^* := \{i : u_i^* > v_i^* = 0\}, \\ I_2^* := \{i : 0 = u_i^* < v_i^*\},$$

and let $\Gamma^* := \{w = (u,v) : (u_i, v_i) \in K_i \text{ if } i \in I_i^* \ (i = 1,2)\}$. If $w^k \in \Gamma^*$ for all sufficiently large k, then $\Psi(w^k) \to 0$ as $k \to \infty$. Moreover, there exists a point $\bar{w} \in C_1 \cap C_2$ such that $w^k \to \bar{w}$ as $k \to \infty$.

Proof. We have from Corollary 3.10 that $w^k - \Psi'(w^k) \in P_{C_2}(w^k)$ where $\Psi'(w^k) \in \partial_B \Psi(w^k)$. Thus,

$$\begin{aligned}
\|w^{k+1} - w^*\|^2 &= \|P_{C_1}(w^k - \Psi'(w^k)) - P_{C_1}(w^*)\|^2 \\
&\leq \|(w^k - w^*) - \Psi'(w^k)\|^2 \\
&= \|w^k - w^*\|^2 - 2\langle w^k - w^*, \Psi'(w^k)\rangle + \|\Psi'(w^k)\|^2, \quad (3.8)
\end{aligned}$$

where the inequality holds by nonexpansiveness of P_{C_1}. Meanwhile, since $w^k, w^* \in \Gamma^*$, then inequalities (3.6) and (3.7) yield

$$\begin{aligned}
\langle w^k - w^*, \Psi'(w^k)\rangle &= \langle (u^k - u^*, v^k - v^*), \Psi'(w^k)\rangle \\
&= \sum_{i \in I_1^* \cup I_2^*} \langle (u_i^k - u_i^*, v_i^k - v_i^*), \nabla\psi(u_i^k, v_i^k)\rangle \\
&\quad + \sum_{i \notin I_1^* \cup I_2^*} \langle (u_i^k, v_i^k), \psi'(u_i^k, v_i^k)\rangle \\
&\geq 2\sum_{i=1}^{n} \psi(u_i^k, v_i^k) \\
&= 2\Psi(w^k), \quad (3.9)
\end{aligned}$$

where $\psi'(u_i^k, v_i^k) \in \partial_B \psi(u_i^k, v_i^k)$. On the other hand, we have

$$\|\Psi'(w^k)\|^2 - 2\Psi(w^k) = \sum_{i=1}^{n} \left[\|\psi'(u_i^k, v_i^k)\|^2 - 2\psi(u_i^k, v_i^k)\right] = 0, \quad (3.10)$$

where the last equality can be verified directly from (3.3) and (3.4). Continuing from (3.8), we have

$$\begin{aligned}
\|w^{k+1} - w^*\|^2 &\leq \|w^k - w^*\|^2 - 4\Psi(w^k) + \|\Psi'(w^k)\|^2, \quad (3.11) \\
&= \|w^k - w^*\|^2 - 2\Psi(w^k), \quad (3.12)
\end{aligned}$$

where (3.11) and (3.12) follow from (3.9) and (3.10), respectively. From (3.12), we get

$$2\sum_{k=1}^{N} \Psi(w^k) = \|w^0 - w^*\|^2 - \|w^{N+1} - w^*\|^2 \leq \|w^0 - w^*\|^2 \qquad \forall N \in \mathbb{N}.$$

Thus, $\Psi(w^k) \to 0$ as $k \to \infty$. This proves the first claim.

Meanwhile, note that (3.12) implies that the sequence $\{w^k\}_{k=0}^{\infty}$ is bounded. Thus, the sequence has an accumulation point $\bar{w}$, i.e. there exists a subsequence $\{w^{k_j}\}_{j=1}^{\infty}$ such that $w^{k_j} \to \bar{w}$ as $j \to \infty$. Since $\{w^k\}_{k=0}^{\infty} \subseteq C_1$ and C_1 is closed, then $\bar{w} \in C_1$. Moreover, $\Psi(w^{k_j}) \to \Psi(\bar{w})$ as $j \to \infty$ since Ψ is continuous. Since the full sequence $\{\Psi(w^k)\}_{k=0}^{\infty}$ converges to zero, then $\Psi(\bar{w}) = 0$, that is, $\bar{w} \in C_2$. Hence, we obtain that $\bar{w} \in C_1 \cap C_2$ and since $\bar{w}$ must also be a point in the closure of Γ^*, then $\bar{w} \in \Gamma^*$. By applying the same argument for w^* as above, we obtain $\|w^{k+1} - \bar{w}\| \leq \|w^k - \bar{w}\|$ as in (3.12). Thus, $\{\|w^k - \bar{w}\|\}$ is a decreasing sequence of nonnegative numbers and must therefore be convergent. Since $\|w^{k_j} - \bar{w}\| \to 0$ as $j \to \infty$, then it follows that $\|w^k - \bar{w}\| \to 0$, i.e., $w^k \to \bar{w}$ as $k \to \infty$. This completes the proof. □

Remark 3.13. We observe that the above theorem implies the local convergence result given by Theorem 3.1. In addition, we obtain the global convergence to $C_1 \cap C_2$ if $c = 0$, which is stronger than the claim of Proposition 3.5. Hence, the above discussion provides an alternative proof for the aforementioned results in light of the new C-function ψ.

To see precisely how one gets the local convergence given by Theorem 3.1, let $B(w, \delta)$ denote the open ball centered at w with radius δ. For each $i \in I_1^*$, let $\delta_i > 0$ be such that $B((u_i^*, v_i^*), \delta_i) \subseteq \{(a, b) \in \mathbb{R}^2 : a > b\}$ and for each $i \in I_2^*$, let $\delta_i > 0$ so that $B((u_i^*, v_i^*), \delta_i) \subseteq \{(a, b) \in \mathbb{R}^2 : a < b\}$. Taking $\delta := \min\{\delta_i : i \in I_1^* \cup I_2^*\}$, then $B(w^*, \delta) \subseteq \Gamma^*$. Moreover, for any $w \in \Gamma^*$, note that

$$\begin{aligned}\|P_{C_2}(w) - P_{C_2}(w^*)\|^2 &= \sum_{i \in I_1^* \cup I_2^*} \|P_M(u_i, v_i) - P_M(u_i^*, v_i^*)\|^2 \\ &\quad + \sum_{i \notin I_1^* \cup I_2^*} \|P_M(u_i, v_i) - P_M(0, 0)\|^2 \\ &\leq \sum_{i \in I_1^* \cup I_2^*} \|(u_i, v_i) - (u_i^*, v_i^*)\|^2 + \sum_{i \notin I_1^* \cup I_2^*} \|P_M(u_i, v_i)\|^2 \\ &\leq \sum_{i \in I_1^* \cup I_2^*} \|(u_i, v_i) - (u_i^*, v_i^*)\|^2 + \sum_{i \notin I_1^* \cup I_2^*} \|(u_i, v_i)\|^2 \\ &= \|w - w^*\|^2, \end{aligned} \tag{3.13}$$

where M is as defined in the proof of Proposition 2.2. The first inequality above follows from the proof of Corollary 2.4, while the second inequality holds since $\|P_M(a, b)\| \leq \|(a, b)\|$ for all $(a, b) \in \mathbb{R}^2$. It follows from inequality (3.13) that if $w^0 \in B(w^*, \delta)$, then $w^k \in B(w^*, \delta)$ for all $k \geq 0$. Thus, $w^k \in \Gamma^*$ for all k and by Theorem 3.12, w^k converges to some $\bar{w} \in C_1 \cap C_2$. This is precisely the claim of Theorem 3.1.

If $c = 0$, we see that $w^* = 0 \in C_1 \cap C_2$ so that $I_1^* = I_2^* = \emptyset$. The above discussion reveals that the MAP iterates will converge to a point in $C_1 \cap C_2$ given any initial point w^0. ■

We provide a geometric interpretation of Theorem 3.12 for the case that $I_1^* \cup I_2^* = \{1, \ldots, n\}$. In this case, there exists $\tau^* \in \mathscr{T}$ such that w^* is contained in the interior of S_{τ^*}. Theorem 3.12 indicates that if the iterates will eventually be "trapped" in S_{τ^*}, then the iterates must converge to a solution of the feasibility problem (FP). Observe that the above remark implies that this could occur if we choose an initial point w^0 that is close enough to w^*. However, it is in general difficult to prove this when the initial point is arbitrarily set. Nevertheless, we will prove that for the case $m = n$, it is impossible for the iterates to be eventually trapped in some S_τ that does not contain a point in $C_1 \cap C_2$ if we assume nondegeneracy of Q as defined in (2.15). To this end, we need the following lemma.

Given any matrix $A \in \mathbb{R}^{m \times n}$, we denote by $\sigma_k(A)$ the kth largest singular value of A. Moreover, the norm of A is the largest singular value, i.e., $\|A\| = \sigma_1(A)$. If $k > \min\{m, n\}$, we set $\sigma_k(A) = 0$.

Lemma 3.14. [36, Corollary 3.1.3] *Let $A \in \mathbb{R}^{m\times n}$ and let A_r denote a submatrix of A obtained by deleting a total of r rows and/or columns of A. Then*

$$\sigma_k(A) \geq \sigma_k(A_r) \geq \sigma_{k+r}(A), \quad k = 1, \ldots, \min\{m, n\}.$$

Lemma 3.15. *Suppose that Q given by (2.15) is nondegenerate. Let $\Lambda := \Lambda_1 \cup \Lambda_2$ where $\Lambda_1 \subseteq \{1, \ldots, n\}$ and $\Lambda_2 = \{n + i : i \notin \Lambda_1\}$. If $L_{\cdot\Lambda}$ is the submatrix of $L := I_{2n} - T^\dagger T$ containing all its columns indexed by Λ and all of its $2n$ rows, then $\|L_{\cdot\Lambda}\| < 1$.*

Proof. Let $E_1 \in \mathbb{R}^{n\times n}$ such that the first $|\Lambda_1|$ columns of E_1 are the standard unit vectors $e_i \in \mathbb{R}^n$ with $i \in \Lambda_1$, while the other remaining columns are zeros. In addition, let $E_2 \in \mathbb{R}^{n\times n}$ be such that the first $|\Lambda_1|$ columns are zeros and the last $|\Lambda_2|$ columns are composed of e_i's where $i \notin \Lambda_1$. Further, let $E := \begin{bmatrix} E_1 & E_2 \\ E_2 & E_1 \end{bmatrix}$. We note that

$$E_1E_1^\top + E_2E_2^\top = I_n, \qquad E_iE_j^\top = 0\ (\forall i \neq j) \quad \text{and} \quad EE^\top = E^\top E = I_{2n}. \tag{3.14}$$

Then, the matrix $L_{\cdot\Lambda}$ is precisely the submatrix of $\tilde{L} := E^\top ME$ containing all its rows and its first n columns. Meanwhile, using the identities (3.14), it can be verified that the matrix $\tilde{L}$ is also equal to $I_{2n} - \tilde{T}^\dagger\tilde{T}$ where $\tilde{T} := TE = [\ U \ \ V\ ]$, $U := T\begin{bmatrix} E_1 \\ E_2 \end{bmatrix}$ and $V := T\begin{bmatrix} E_2 \\ E_1 \end{bmatrix}$. Calculating $\tilde{L}$ using this formula, we see that

$$L_{\cdot\Lambda} = \begin{bmatrix} I_n - U^\top WU \\ -V^\top WU \end{bmatrix},$$

where $W = (TT^\top)^{-1}$. Noting that $UU^\top + VV^\top = W^{-1}$, which can be derived from (3.14), we obtain

$$U^\top WVV^\top WU = U^\top WU - (U^\top WU)^2. \tag{3.15}$$

Then

$$\begin{aligned} \|L_{\cdot\Lambda}\|^2 &= \lambda_{\max}(L_{\cdot\Lambda}^\top L_{\cdot\Lambda}) \\ &= \lambda_{\max}\left((I_n - U^\top WU)^2 + U^\top WVV^\top WU\right) \\ &= \lambda_{\max}(I_n - U^\top WU), \end{aligned}$$

where the last equality follows from (3.15). Meanwhile, we know from the definition of U that it is composed of the columns of T which are indexed by Λ. By Lemma 2.12, these columns must be linearly independent so that U is nonsingular and 1 is not an eigenvalue of $I_n - U^\top WU$. We conclude that $\|L_{\cdot\Lambda}\| \neq 1$. But by Lemma 3.14, we know that $\|L_{\cdot\Lambda}\| \leq \|L\| = 1$. Hence, we arrive at the desired conclusion. □

Using the above lemma, we obtain the following proposition.

Proposition 3.16. *Let $\{w^k\}_{k=0}^\infty$ be any sequence generated by (MAP), and suppose there exists $\tau \in \mathscr{T}$ such that $S_\tau \cap (C_1 \cap C_2) = \emptyset$, i.e. S_τ does not contain a solution of (FP). Then there does not exist $N \in \mathbb{N}$ such that $\{w^k\}_{k=N}^\infty \subseteq S_\tau \cap \Omega$, where Ω is given by (2.13).*

Proof. Suppose to the contrary that there exists N such that $w^k \in S_\tau \cap \Omega$ for all $k \geq N$. To prove the result, we will show that w^k converges to some point $w^* \in S_\tau \cap (C_1 \cap C_2)$ which is a contradiction to our hypothesis. To this end, we apply a convenient change of variables based on τ. First, let $\Lambda_1 := \{i : \tau(i) = 1\}$ and $\Lambda_2 := \{n + i : \tau(i) = 2\}$. With these index sets, define the matrices E, $L_{\cdot\Lambda}$, $\tilde{L}$ and $\tilde{T}$ be as in the proof of Lemma 3.15.

We consider the transformation $w = E\tilde{w}$. Similar to the discussion in Sect. 2.1, we see that $w \in C_i$ if and only if $\tilde{w} \in \tilde{C}_i$ where $\tilde{C}_1 := \{\tilde{w} : \tilde{T}\tilde{w} = \sqrt{2}c\}$ and $\tilde{C}_2 = C_2$. Moreover,

$$w \in S_\tau \cap \Omega \Longleftrightarrow \tilde{u}_i > \tilde{v}_i \quad \text{and} \quad \tilde{u}_i \geq 0 \quad \forall i = 1, \ldots, n. \tag{3.16}$$

Analogous to Eqs. (2.1) and (2.2), we have

$$(P_{C_1} \circ P_{C_2})(w) = E(P_{\tilde{C}_1} \circ P_{\tilde{C}_2})(\tilde{w}),$$

and

$$\mathrm{Fix}(P_{C_1} \circ P_{C_2}) = E\left(\mathrm{Fix}(P_{\tilde{C}_1} \circ P_{\tilde{C}_2})\right) \tag{3.17}$$

since E is unitary.

We now look at the transformed iterates $\tilde{w}^k = E^\top w^k = (\tilde{u}_k, \tilde{v}_k)$. By (3.16), we have $\tilde{u}_k > \tilde{v}_k$ and $\tilde{u}_k \geq 0$ for all $k \geq N$ so that

$$\begin{aligned} \tilde{w}^{k+1} &= (P_{\tilde{C}_1} \circ P_{\tilde{C}_2})(\tilde{w}^k) && (3.18)\\ &= P_{\tilde{C}_1}\left((\tilde{u}^k, 0)\right) && (3.19)\\ &= \tilde{L}(\tilde{u}^k, 0)^\top + \sqrt{2}\tilde{T}^\dagger c && (3.20)\\ &\\ &= L_{\cdot\Lambda}\tilde{u}^k + \sqrt{2}\tilde{T}^\dagger c \qquad \forall k \geq N, \end{aligned}$$

where (3.19) and (3.20) follow from Propositions 2.2 and 2.1, respectively. Letting $L_{\cdot\Lambda} = \begin{bmatrix}(L_{\cdot\Lambda})_1\\(L_{\cdot\Lambda})_2\end{bmatrix}$ and $\sqrt{2}\tilde{T}^\dagger c = \begin{bmatrix}d_1\\d_2\end{bmatrix}$,

where $(L_{\cdot\Lambda})_1, (L_{\cdot\Lambda})_2 \in \mathbb{R}^{n\times n}$ and $d_1, d_2 \in \mathbb{R}^n$, then we see that $\tilde{u}^{k+1} = (L_{\cdot\Lambda})_1\tilde{u}^k + d_1$ and $\tilde{v}^{k+1} = (L_{\cdot\Lambda})_2\tilde{u}^k + d_2$. Since $\|L_{\cdot\Lambda}\| < 1$ by Lemma 3.15, we also have by Lemma 3.14 that $\|(L_{\cdot\Lambda})_1\| < 1$ so that $\{\tilde{u}^k\}_{k=0}^\infty$ is convergent. Consequently, $\{\tilde{v}^k\}_{k=0}^\infty$ is also convergent.

Therefore, there exists $\tilde{w}^*$ such that $\tilde{w}^k \to \tilde{w}^*$ as $k \to \infty$. Moreover, we have from (3.18) that $\tilde{w}^* = (P_{\tilde{C}_1} \circ P_{\tilde{C}_2})(\tilde{w}^*)$, i.e. $\tilde{w}^* \in \mathrm{Fix}(P_{\tilde{C}_1} \circ P_{\tilde{C}_2})$. Since $w^k = E\tilde{w}^k$, it also follows that $w^k \to w^* := E\tilde{w}^*$ and from (3.17), $w^* \in \mathrm{Fix}(P_{C_1} \circ P_{C_2})$. Since Ω is closed, it also follows that $w^* \in \Omega$. From Theorem 2.13, we must have $w^* \in C_1 \cap C_2$. However, since w^* must belong to the closure of S_τ, the fact that it is in $C_1 \cap C_2$ implies that $w^* \in S_\tau$. Hence, $w^* \in S_\tau \cap (C_1 \cap C_2)$. This is a contradiction. □

3.3. Rate of Convergence

An immediate consequence of Lemma 3.15 is the local linear rate of convergence of the iterates (MAP).

Theorem 3.17. *Let $m = n$ and suppose that $w^* \in C_1 \cap C_2$ such that $(u_i^*, v_i^*) \neq (0,0)$ for all $i = 1, \ldots, n$. If Q given by (2.15) is nondegenerate, then there exists sufficiently small $\delta > 0$ such that for any w^0 with $\|w^0 - w^*\| < \delta$, the sequence $\{w^k\}_{k=0}^{\infty}$ generated by (MAP) converges linearly to w^*.*

In the following, we denote by $\mathrm{supp}(w) := \{i : w_i \neq 0\}$ the support of a vector w.

Proof. Let $\tau^* \in \mathscr{T}$ such that $w^* \in S_{\tau^*}$. Observe that since $(u_i^*, v_i^*) \neq (0,0)$ for all $i = 1, \ldots, n$, then Γ^* defined in Theorem 3.12 is precisely the set S_{τ^*}. Choose $\delta > 0$ sufficiently small so that the closure of $B(w^*, \delta)$ is contained in the interior of $\Gamma^* = S_{\tau^*}$. From the discussion in Remark 3.13, we have $\{w^k\}_{k=0}^{\infty} \subseteq B(w^*, \delta)$ whenever $w^0 \in B(w^*, \delta)$. Moreover, there exists $\bar{w} \in C_1 \cap C_2$ in the interior of S_{τ^*} such that $w^k \to \bar{w}$ as $k \to \infty$ for any $w^0 \in B(w^*, \delta)$. Thus, $(\bar{u}_i, \bar{v}_i) \neq (0,0)$ for all $i = 1, \ldots, n$. Meanwhile, in view of the equivalence of AVE and the LCP described in Remark 2.18 together with [13, Theorem 3.6.3], the nondegeneracy assumption on Q implies that w^* is an isolated solution of the feasibility problem (FP). Thus, by choosing a smaller δ (if necessary), we have that $\bar{w} = w^*$. That is, $w^k \to w^*$ for all $w^0 \in B(w^*, \delta)$.

Denote by Λ the support of w^*. Since $\{w^k\}_{k=0}^{\infty}$ is contained in the interior of S_{τ^*} and $w^k \to w^*$, then we have by Proposition 2.2 that P_{C_2} is single-valued at w^k and $\mathrm{supp}(P_{C_2}(w^k)) = \Lambda$ for all $k \geq 0$. Thus,

$$\|(P_{C_2}(w^k) - P_{C_2}(w^*))_\Lambda\| = \|P_{C_2}(w^k) - P_{C_2}(w^*)\| \leq \|w^k - w^*\|,$$

where the inequality holds by nonexpansiveness of P_{C_2} on S_{τ^*} (Corollary 2.4). Then if $L_{\cdot\Lambda}$ denotes the submatrix of $L := I_{2n} - T^\dagger T$ containing all of its $2n$ rows and all its columns indexed by Λ, we have from Proposition 2.1 that

$$\begin{aligned}
\|w^{k+1} - w^*\| &= \|P_{C_1}(P_{C_2}(w^k)) - P_{C_1}(P_{C_2}(\bar{w}))\| \\
&= \|L(P_{C_2}(w^k)) - L(P_{C_2}(w^*))\| \\
&= \|L_{\cdot\Lambda}(P_{C_2}(w^k) - P_{C_2}(w^*))_\Lambda\| \\
&\leq \|L_{\cdot\Lambda}\| \cdot \|(P_{C_2}(w^k) - P_{C_2}(w^*))_\Lambda\| \\
&\leq \|L_{\cdot\Lambda}\| \cdot \|w^k - w^*\|.
\end{aligned}$$

Since $\|L_{\cdot\Lambda}\| < 1$ by Lemma 3.15, the conclusion of this theorem follows. □

The rate of convergence asserted by the above result can also be obtained using [37, Theorem 5.16] and Proposition 2.12. In fact, it can be extended to the general case when m is not necessarily equal to n using the notions of "super-regularity" and "linearly regular intersection". We recall from [37] that a closed set C is *super-regular* at w^* if, for all $\varepsilon > 0$, any two points z_1, z_2 sufficiently close to w^* with $z_2 \in C$, and any point $y \in P_C(z_1)$, satisfy $\langle z_1 - y, z_2 - y \rangle \leq \varepsilon \|z_1 - y\| \cdot \|z_2 - y\|$. In particular, a convex set is super-regular at each of its points. We refer the reader to [38] for more details on how super-regularity is related with other pre-existing notions.

To define the concept involving sets with linearly regular intersection, we first recall that the *limiting normal cone* to a closed set C at $w^* \in C$ is given by

$$N_C(w^*) = \left\{ \lim_{k\to\infty} t_k(w^k - z^k) : t_k \geq 0,\ w^k \to w^*,\ z^k \in P_C(w^k) \right\}.$$

We say that two closed sets C_1 and C_2 have a *linearly regular intersection* at $w^* \in C_1 \cap C_2$ if

$$N_{C_1}(w^*) \cap (-N_{C_2}(w^*)) = \{0\}. \tag{3.21}$$

With these definitions, we state the following convergence result from [37].

Lemma 3.18. [37, Theorem 5.16] *If C_1 and C_2 are closed sets which have a linearly regular intersection at $w^* \in C_1 \cap C_2$ and if either C_1 or C_2 is super-regular at w^*, then any alternating projection sequence with initial point sufficiently close to w^* converges linearly to a point in $C_1 \cap C_2$.*

Using the above result, we obtain the local linear convergence of the MAP iterates.

Theorem 3.19. *Suppose that $w^* \in C_1 \cap C_2$ such that $(u_i^*, v_i^*) \neq (0,0)$ for all $i = 1, \ldots, n$. If condition (C) holds, then any sequence generated by (MAP) with initial point sufficiently close to w^* converges linearly to a point in $C_1 \cap C_2$.*

Proof. We note that since C_1 is convex, then it is super-regular at each of its points. Thus, by Lemma 3.18, it suffices to show that C_1 and C_2 have a linearly regular intersection at $w^* \in C_1 \cap C_2$ where $(u_i^*, v_i^*) \neq (0,0)$ for all $i = 1, \ldots, n$. Directly from the definition, the limiting normal cones to C_1 and C_2 are given, respectively, by

$$N_{C_1}(w) = \operatorname{Ker}(T)^{\perp} \qquad \forall w \in C_1$$

and

$$N_{C_2}(w) = \{w' = (u', v') \in \mathbb{R}^n \times \mathbb{R}^n : (u_i', v_i') \in N_M(u_i, v_i)\} \qquad \forall w \in C_2, \tag{3.22}$$

where M is given by (2.6) and

$$N_M(s,t) = \begin{cases} \{(0,\lambda) : \lambda \in \mathbb{R}\} & \text{if } s > t = 0 \\ \{(\lambda,0) : \lambda \in \mathbb{R}\} & \text{if } 0 = s < t\ . \\ \mathbb{R}^2_- \cup M & \text{if } s = t = 0 \end{cases}$$

We note that the normal cone to C_2 can also be obtained using [39, Theorem 3.4]. Since $(u_i^*, v_i^*) \neq (0,0)$, it follows that $N_{C_2}(w^*) \subseteq \hat{C}_2$, where $\hat{C}_2$ is given by (2.14). By condition (C), we see that (3.21) holds, i.e., the intersection at w^* is linearly regular. This completes the proof. □

Notice that the above theorem guarantees local linear convergence of (MAP) to a point in $C_1 \cap C_2$, which may not be the same as the point w^*. On the other hand, Theorem 3.17 shows that local linear convergence to w^* is achieved.

3.4. Globally convergent relaxation of MAP

In the preceding sections, our analysis was focused on the iterates given by (MAP). To obtain a global result, we now focus on a relaxed version of the iterations (MAP) given by

$$w^{k+1} \in (1-\gamma)P_{C_2}(w^k) + \gamma(P_{C_1} \circ P_{C_2})(w^k), \tag{3.23}$$

where $\gamma \in (0,1)$ is fixed and the initial point is $w^0 = (1-\gamma)\bar{w}^0 + \gamma P_{C_1}(\bar{w}^0)$ with $\bar{w}^0 \in C_2$.

Theorem 3.20. *Let $\{w^k\}_{k=0}^{\infty}$ be a sequence generated by (3.23). If $\{w^k\}_{k=0}^{\infty}$ is bounded, then there exists $\bar{w}^* \in C_2$ such that $w^k \to w^*$ where $w^* = (1-\gamma)\bar{w}^* + \gamma P_{C_1}(\bar{w}^*)$. Moreover, if condition (C) holds and $(\bar{u}_i^*, \bar{v}_i^*) \neq (0,0)$ for all $i = 1, \ldots, n$, then $w^* \in C_1 \cap C_2$, that is, the sequence $\{w^k\}_{k=0}^{\infty}$ is globally convergent to a solution of the feasibility problem (FP).*

Proof. To prove the convergence of $\{w^k\}_{k=0}^{\infty}$, denote $\bar{w}^k \in P_{C_2}(w^k)$ for all $k \geq 0$. By (3.23), we have

$$w^{k+1} = (1-\gamma)\bar{w}^k + \gamma P_{C_1}(\bar{w}^k) \tag{3.24}$$

and so

$$\bar{w}^{k+1} \in P_{C_2}(w^{k+1}) = P_{C_2}((1-\gamma)\bar{w}^k + \gamma P_{C_1}(\bar{w}^k)). \tag{3.25}$$

Let $h(w) := \frac{1}{2}\|w - P_{C_1}(w)\|^2$. Then h is a Lipschitz continuous function with Lipschitz constant 1 and $\nabla h(w) = w - P_{C_1}(w)$. Thus, $w - \gamma \nabla h(w) = (1-\gamma)w + P_{C_1}(w)$. In turn, (3.25) reduces to $\bar{w}^{k+1} \in P_{C_2}(\bar{w}^k - \gamma\nabla h(\bar{w}^k))$. From [27, Theorem 5.3], we conclude that there exists a point $\bar{w}^* \in C_2$ such that $\bar{w}^k \to \bar{w}^*$ and

$$0 \in \nabla h(\bar{w}^*) + N_{C_2}(\bar{w}^*). \tag{3.26}$$

By continuity of P_{C_1} and using Eq. (3.24), we see that $w^k \to w^* = (1-\gamma)\bar{w}^* + \gamma P_{C_1}(\bar{w}^*)$.

To prove global convergence to a solution, note that from (3.26), there exists $z^* \in N_{C_2}(\bar{w}^*)$ such that $z^* = P_{C_1}(\bar{w}^*) - \bar{w}^*$. The latter equation implies that $z^* \in \operatorname{Ker}(T)^\perp$. Since $(\bar{u}_i^*, \bar{v}_i^*) \neq (0,0)$ for all $i = 1, \ldots, n$, it follows from (3.22) that $N_{C_2}(\bar{w}^*) \subset \hat{C}_2$. By condition (C), we conclude that $z^* = 0$ and therefore $\bar{w}^* = P_{C_1}(\bar{w}^*)$, i.e. $\bar{w}^* \in C_1$. Hence, $w^* = \bar{w}^* \in C_1 \cap C_2$. □

3.5. A related fixed point algorithm

Another algorithm can also be derived from the method of alternating projections. To describe this algorithm, we denote by D the multivalued mapping from $\mathbb{R}^n \times \mathbb{R}^n$ to $\mathbb{R}^{2n \times 2n}$ such that for each $w = (u,v) \in \mathbb{R}^n \times \mathbb{R}^n$, $D(w)$

is a set containing diagonal matrices D_w such that

$$((D_w)_{ii}, (D_w)_{n+i,n+i}) \in \begin{cases} \{(1,0)\} & \text{if } u_i > v_i,\ u_i \geq 0 \\ \{(0,1)\} & \text{if } u_i < v_i,\ v_i \geq 0 \\ \{(0,1),(1,0)\} & \text{if } u_i = v_i > 0 \\ \{(0,0)\} & \text{if } u_i = v_i \leq 0 \end{cases},$$

for all $i = 1, \dots, n$. Then, the projection onto C_2 can equivalently written as

$$P_{C_2}(w) = D(w)w = \{D_w w : D_w \in D(w)\}.$$

Suppose now that w is a fixed point of $P_{C_1} \circ P_{C_2}$, i.e. $w \in (P_{C_1} \circ P_{C_2})(w)$. Recalling that $P_{C_1}(w) = Lw + \sqrt{2}T^\dagger c$, where $L = I_{2n} - T^\dagger T$, we have

$$w = LD_w w + \sqrt{2}T^\dagger c, \qquad \text{where } D_w \in D(w).$$

That is, $w \in \operatorname{Fix}(P_{C_1} \circ P_{C_2})$ if and only if there exists $D_w \in D(w)$ such that

$$(I_{2n} - LD_w)w = \sqrt{2}T^\dagger c.$$

This motivates the iterations

$$w^{k+1} = \sqrt{2}(I_{2n} - LD_{w^k})^{-1}T^\dagger c, \tag{3.27}$$

where $D_{w^k} \in D(w^k)$. This algorithm is well-defined if 1 is not an eigenvalue of $LD(w^k)$ for all k. A particular case is described in the following proposition.

Proposition 3.21. *The iterations (3.27) are well-defined for any initial point $w^0 \in \mathbb{R}^n \times \mathbb{R}^n$ if Q given by (2.15) is nondegenerate.*

Proof. We show that for all $w \in \mathbb{R}^n \times \mathbb{R}^n$, the matrix $I_{2n} - LD_w$ is nonsingular for any $D_w \in D(w)$. To this end, let $\Lambda = \{i \in \{1, \dots, 2n\} : (D_w)_{ii} = 1\}$. Then

$$\|LD_w\| = \|L_{\cdot\Lambda}\| \leq \|L\|,$$

where the inequality follows from Lemma 3.14. Since Q is nondegenerate, $\|L\| < 1$ by Lemma 3.9. Thus, $\|LD_w\| < 1$ and therefore, $I_{2n} - LD_w$ is nonsingular, as desired. □

Unlike the iterates (MAP), it is not difficult to show that any sequence generated via (3.27) is bounded.

Proposition 3.22. *Let $\{w^k\}_{k=0}^\infty$ be any sequence generated by (3.27). Then $\{w^k\}_{k=0}^\infty$ is a bounded sequence. Any accumulation point w^* of $\{w^k\}_{k=0}^\infty$ satisfies $w^* = \sqrt{2}(I_{2n} - LD^*)^{-1}T^\dagger c$, where $D^* \in \mathbb{R}^{2n\times 2n}$ is a diagonal matrix with diagonal elements of 1 or 0 and satisfy $D^*_{ii}D^*_{n+i,n+i} = 0$ for all $i = 1, \dots, n$.*

Proof. Note that the range of the multivalued mapping D is a finite set. In particular, the set $\{D_w : D_w \in D(w) \text{ and } w \in \mathbb{R}^n \times \mathbb{R}^n\}$ has 3^n elements. Thus, there exists a constant $\kappa \in (0, \infty)$ such that $\|(I_{2n} - LD_w)^{-1}\| \leq \kappa$ for all $w \in \mathbb{R}^n \times \mathbb{R}^n$ and $D_w \in D(w)$. Thus, $\|w^{k+1}\| \leq \sqrt{2}\|(I_{2n} - LD_{w^k})^{-1}\| \cdot \|T^\dagger c\| \leq \sqrt{2}\kappa\|T^\dagger c\|$ for all k. Hence, $\{w^k\}_{k=0}^\infty$ is a bounded sequence. To prove the last claim, let w^* be an arbitrary accumulation point of $\{w^k\}_{k=0}^\infty$, and let

$\{w^{k_j}\}_{j=1}^{\infty}$ be a subsequence that converges to w^*. Denote by d^{k_j-1} the diagonal entries of $D_{w^{k_j-1}}$. Then the sequence $\{(w^{k_j}, d^{k_j-1})\}_{j=1}^{\infty}$ is bounded and must have subsequence that converges to some point (w^*, d^*). Without loss of generality, we may assume that $\{(w^{k_j}, d^{k_j-1})\}_{j=1}^{\infty}$ converges to (w^*, d^*). It follows that $D_{w^{k_j-1}} \to D^*$ as $j \to \infty$, where D^* is the diagonal matrix with diagonal entries equal to d^*. Setting $k = k_j$ in (3.27) and letting $j \to \infty$, we get the desired conclusion. □

Both the MAP algorithm (MAP) and the iterations (3.27) are aimed at finding a fixed point of $P_{C_1} \circ P_{C_2}$. However, the iterations (3.27) require more computational effort than MAP since the former involves solving a linear system involving $2n$ equations in $2n$ unknowns for each iteration. Nevertheless, we may consider a hybrid algorithm where we generate first a sequence of MAP iterates, then use (3.27) for the succeeding iterations. We call this approach the MAP-LS algorithm (where LS denotes linear system involved in computing the iterations given by (3.27)) which is described in Algorithm 1. Whenever convergent, the limit of the sequence generated by MAP-LS algorithm is necessarily a fixed point of $P_{C_1} \circ P_{C_2}$.

Algorithm 1: MAP-LS algorithm

```
(H) Choose a termination parameter $\varepsilon$ and set $w^0 = T^\dagger c$. Let $N$ be
a positive integer and $\delta > 0$. Set $k = 0$.
Step 1. Let
    $w^{k+1} \in \begin{cases} (P_{C_1} \circ P_{C_2})(w^k) & \text{if } k \le N \text{ and } \|w^{k+1} - w^k\| > \delta \\ \sqrt{2}(I_{2n} - LD_{w^k})^{-1} T^\dagger c & \text{if } k > N \text{ or } \|w^{k+1} - w^k\| \le \delta. \end{cases}$.
Step 2. Set $x^{k+1} = \frac{1}{\sqrt{2}}(u^{k+1} - v^{k+1})$.
Step 3. Stop if $\|Ax^{k+1} + B|x^{k+1}| - c\| \le \varepsilon$. Otherwise, set $k = k + 1$
    and go to Step 1.
```

4. Numerical simulations

In this section, we demonstrate the applicability of MAP and MAP-LS in solving randomly generated absolute value equations (1.1). We first note some remarks on the implementation of our algorithms.

4.1. Implementation of MAP and MAP-LS

If $T = [\, A + B \; -A + B \,] \in \mathbb{R}^{m \times 2n}$ is of full row rank, then its Moore–Penrose inverse of T is well-known and is given by

$$T^\dagger = T^\intercal (TT^\intercal)^{-1}.$$

In view of Proposition 2.1, we calculate the projection onto C_1 of a point $w \in \mathbb{R}^n \times \mathbb{R}^n$ by first solving for z in

$$TT^\intercal z = Tw - \sqrt{2}c, \tag{4.1}$$

then setting $P_{C_1}(w) = w - T^{\mathsf{T}} z$.

Notice that since T is of full row rank, the coefficient matrix TT^{T} of the linear system (4.1) is a symmetric positive definite matrix, so we can use its Cholesky decomposition. In particular, we use the Matlab function `dS = decomposition(S,'chol')` where $S := TT^{\mathsf{T}} = 2(AA^{\mathsf{T}} + BB^{\mathsf{T}})$ and solve for z in (4.1) using the backslash operator, i.e., `z = dS\b`, where $b := Tw - \sqrt{2}c$.

In particular, by virtue of Lemma 2.8, the above procedure can be applied when dealing with the traditional AVE (1.1) with $A \in \mathbb{R}^{n\times n}$ and $B = -I_n$. Furthermore, in this case, the matrix-vector multiplication $T^{\mathsf{T}} z$ can be calculated more efficiently by computing first $z' := A^{\mathsf{T}} z$ so that $T^{\mathsf{T}} z = (z' - z, -z' - z)$.

On the other hand, the inversion of $2n \times 2n$ matrix in Eq. (3.27) may be computationally intensive. However, since $I_{2n} - LD_{w^k}$ can be partitioned into four $n \times n$ blocks, then its inverse can be calculated in terms of the inverses of two $n \times n$ matrices. Particularly, if we let $L = \begin{bmatrix} L_1 & L_2 \\ L_2^{\mathsf{T}} & L_3 \end{bmatrix}$ and $D_{w^k} = \begin{bmatrix} D_1^k & 0 \\ 0 & D_2^k \end{bmatrix}$, then

$$I_{2n} - LD_{w^k} = \begin{bmatrix} I_n - L_1 D_1^k & -L_2 D_2^k \\ -L_2^T D_1^k & I_n - L_3 D_2^k \end{bmatrix}.$$

Thus, the inverse of $I_{2n} - LD_{w^k}$ can be calculated in terms of the inverse of $I_n - L_1 D_1^k$ and the inverse of its Schur complement (or the inverse of $I_n - L_3 D_2^k$ and the inverse of its Schur complement). These inverses exist, in particular, if $\|LD_{w^k}\| < 1$ (such as when Q is nondegenerate) in which case $\|LD_1^k\| < 1$ and $\|LD_2^k\| < 1$ by Lemma 3.14. In general, such approach is more efficient than dealing directly with the inverse of $I_{2n} - LD_{w^k}$. Hence, we take this approach when using the MAP-LS algorithm.

4.2. Numerical results

We compare MAP and MAP-LS to four other algorithms in the literature, each of which is a representative of the four classifications described in the introduction. We only choose those algorithms which, like MAP and MAP-LS, do not require parameters which need to be tuned carefully. From the class of algorithms based on Newton methods, we choose the generalized Newton method (GNM) [7] as the other variants of the Newton method involve parameters that may be problem dependent or are difficult to tune. From the second group, we choose the Picard iteration method (PIM) in [10]. The variant of this method presented in [19] is only applicable for positive definite matrices and involves a problem-dependent parameter. On the other hand, the iterates of the Douglas–Rachford splitting method [20] are simply convex combinations of the PIM iterates and the current iterate (similar to the MAP relaxation (3.23)). In fact, if we use the prescribed parameters in [20], the Douglas–Rachford iterates approximate the PIM iterates. From matrix splitting iteration methods, we choose the Gauss–Seidel iteration [22]. The SOR-like iteration method [21] also requires a parameter, and from the numerical results presented in [21], we see that the SOR-like iteration also generates iterates which are approximately the same as the PIM iterations for optimally chosen parameters. Finally, we note that the concave minimization approach involves solving a linear program at each iteration, which may be inefficient for large scale problems. We omit comparisons with this approach for the case $B = -I_n$ as the current algorithms in the literature [6,8] are not competitive enough with the other methods. However, we use the successive linearization algorithm (SLA) in [14] for the general AVE (1.1), which is the only existing algorithm in the literature that can solve such problems.

We briefly describe the algorithms we have chosen for our numerical comparisons:

(a) *Generalized Newton method (GNM)* [7]
This algorithm is aimed at solving the AVE (1.1) with $m = n$ and $B = -I_n$, and the iterations are given by

$$x^{k+1} = (A - D^k)^{-1} c, \tag{4.2}$$

where $D^k = \operatorname{diag}(\operatorname{sgn}(x_1^k), \ldots, \operatorname{sgn}(x_n^k))$. The iterations are derived by applying the semismooth Newton method in solving the equation $Ax - |x| - c = 0$. As in [7], we use the Matlab's backslash operator "\" to obtain the iterates. The maximum iterations for this algorithm is set to 2000.

(b) *Picard iteration method (PIM)* [10]
This method is applicable whenever $m = n$ and A is invertible. The algorithm consists of the fixed point iterations for the equation $x = A^{-1}(-B|x| + c)$, that is,

$$x^{k+1} = A^{-1}(-B|x^k| + c). \tag{4.3}$$

From the above formula, we only need to compute A^{-1} once. For the sake of efficiency, we pre-compute the LU decomposition of A using the `decomposition` function of Matlab. We set the maximum number of iterations to 2000.

(c) *Gauss–Seidel iteration method (GSM)* [22]
Similar to the generalized Newton method, this Gauss–Seidel algorithm solves the AVE $Ax - |x| = c$ by decomposing A as $A = D - E - F$ where D, E and F are diagonal, strictly lower triangular and strictly upper triangular matrices. Using this decomposition, the Gauss–Seidel iterations are given by

$$(D - E)x^{k+1} - |x^{k+1}| = Fx^k + c.$$

Though the above system is nonlinear, the next iterate x^{k+1} can be easily solved since $D - E$ is lower triangular. In particular, having computed x_1^{k+1}, we inductively compute x_i^{k+1} using the previously obtained coordinates $x_1^{k+1}, x_2^{k+1}, \ldots, x_{i-1}^{k+1}$. We set the maximum iterations to 20000.

(d) *Successive linearization algorithm (SLA)* [14]
This is the only algorithm in the existing literature which can handle the general AVE (1.1). Given an initial point $(x^0, t^0, s^0) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m$,

we solve the linear programming problem

$$\begin{aligned} \min\ & \epsilon \sum_{i=1}^{n} (\operatorname{sgn}(x_i^k) x_i + t_i) + \sum_{j=1}^{m} s_i \\ \text{s.t.}\ & -s \leq Ax + Bt - c \leq s \\ & -t \leq x \leq t, \end{aligned}$$

and call its solution $(x^{k+1}, t^{k+1}, s^{k+1})$. To solve this linear program, we use the Matlab function `linprog`. We set the maximum number of iterations to 1000.

All simulations were carried out in Matlab R2020a on a desktop machine with an Intel Core i7-8700 3.20 GHz and 32GB of memory. We use the zero vector as the initial point for all the algorithms, and the stopping criterion is

$$\|Ax^k + B|x^k| - c\| \leq \varepsilon \quad \text{with} \quad \varepsilon = 10^{-6}. \tag{4.4}$$

For the case $m = n$ and $B = -I_n$, we compare our algorithms with GNM, PIM and GSM, since SLA takes a lot of computing time in solving the these problems. For the general case, we can only compare our algorithms with the SLA as the other solvers can only handle the case $m = n$.

Example 4.1. We generate a matrix A as in [7]. First, we generate a matrix $A' \in \mathbb{R}^{n \times n}$ whose entries are from the uniform distribution on $[-10, 10]$. Then, we let $A = A'/(t\sigma_{\min}(A'))$, where t is a uniform random number in $[0, 1]$. We then randomly generate a vector $x^* \in \mathbb{R}^n$ such that $x_i^* = r \cdot 10^{\alpha s}$ where $\alpha \in \{0, 1, 2, 3\}$, while r and s are generated from the uniform distribution on $[-1, 1]$ and $[0, 1]$, respectively. Finally, we set $c = Ax + B|x|$, where $B = -I_n$. We note that the case $\alpha = 0$ is precisely the test problem considered in [7].

In this example, $\sigma_{\min}(A) > \sigma_{\max}(B)$ so that the AVE (1.1) has a unique solution (see Remark 2.18). For our experiments, we let $n = 5000$ and generate 100 random AVEs as described above. We report in Table 1 the success rates and averages of CPU time and number of iterations (of successful simulations) of MAP, GNM, PIM and GSM. First, note that PIM has the best average CPU time in solving the AVEs, followed by GNM and our MAP algorithm. However, in terms of reaching a solution with residual given by (4.4), both GNM and PIM have relatively lower success rates compared to MAP. Moreover, GNM and PIM failed to solve several test problems as α increases. In particular, both of these algorithms failed to solve 100 randomly generated AVEs when $\alpha = 3$. On the other hand, our algorithm is still able to solve more than 60% of the problems when $\alpha = 3$. Finally, notice that Gauss–Seidel method failed to solve any of the problems. For this algorithm, each component of the iterate x^{k+1} is obtained by solving a nonlinear equation of the form $ax - |x| = b$. This equation might not have a solution for $b \neq 0$ if $b/(a-1) < 0$ and $b/(a+1) > 0$, which is the reason why GSM failed in solving the generated AVEs. In fact, this problem was encountered by GSM during the first iteration for all of the test problems considered.

Table 1. Numerical results for Example 4.1.

Method	α			
	0	1	2	3
MAP				
Success rate	1	0.99	0.87	0.62
Ave. Time	2.58	3.03	3.13	10.42
Ave. Iter	40.85	52.51	55.44	250.39
GNM				
Success rate	0.76	0.55	0	0
Ave. Time	2.23	2.29	–	–
Ave. Iter	3.93	4.00	–	–
PIM				
Success rate	0.75	0.54	0.01	0
Ave. Time	0.57	0.59	0.84	–
Ave. Iter	4.99	5.65	22.00	–
GSM				
Success rate	0	0	0	0
Ave. Time	–	–	–	–
Ave. Iter	–	–	–	–

We next demonstrate the use of MAP-LS algorithm in the following two examples.

Example 4.2. We set $B = -I_n$ and let $A = A'(A')^{\mathsf{T}}$ where $A' \in \mathbb{R}^{n\times n}$ is sampled from the standard normal distribution. We also randomly generate a vector x^* from the standard normal distribution, and set $c = Ax^* + B|x^*|$. For each $n \in \{50, 100, 500, 1000, 2000, 3000\}$, we generate 100 random AVEs as described. For this experiment, we note that the convergence of MAP can possibly be extremely slow (for instance, see Fig. 2). Hence, we solve the randomly generated problems using MAP-LS algorithm only, and we present comparisons with GNM, PIM and GSM. The summary of the results is reported in Table 2. For the MAP-LS algorithm, we set $N = 100$ and $\delta = 10^{-3}$ in Algorithm 1. In Table 2, we also report two averages of iteration numbers for MAP-LS: (i) "Ave. Iter (MAP)" indicates the average number of MAP iterations (MAP) of successful instances, and (ii) "Ave. Iter (LS)" indicates the average number of the linear system iterations (3.27) of successful simulations.

We see from Table 2 that MAP-LS used 100 iterations of the alternating projections (MAP) for all the test problems, before using the iterations (3.27). Moreover, the average number of iterations via (3.27) increases as the dimension n increases. Despite this, it is evident that the average CPU time required by MAP-LS to solve the AVEs is significantly shorter than the time required by GNM. In fact, the gap in CPU times spent by MAP-LS and GNM becomes more apparent as the dimension of the problem increases. This is due to the fact that GNM took much more iterations than MAP-LS.

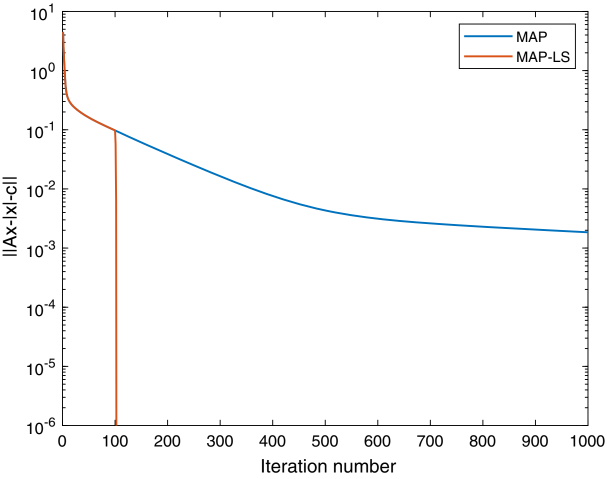

FIGURE 2. Convergence of MAP and MAP-LS for a test problem generated as in Example 4.2 with $n = 100$

Recall that using the implementation described in Sect. 4.1, each iteration of MAP-LS requires two $n \times n$ matrix inversions, while from (4.2), we see that GNM only needs to invert a single $n \times n$ matrix at each iteration. However, as GNM took significantly more iterations than MAP-LS, the latter significantly outperforms the former.

In addition, MAP-LS achieved at least 75% success rate in solving the AVEs of different dimensions n, while the success rate of GNM decreases dramatically as n increases. In particular, for $n = 3000$, GNM only achieved less than 25% success rate. Finally, PIM failed to solve all the generated test problems after reaching the maximum number of iterations set for all n considered, while GSM achieved low success rates for $n \in \{50, 100\}$ and failed to solve the problems when $n \in \{500, 1000, 2000, 3000\}$.

Example 4.3. As mentioned in the introduction, the linear complementarity problem (LCP), that is, the problem of finding $x \in \mathbb{R}^n$ satisfying

$$x \geq 0, \quad Mx + q \geq 0 \quad \text{and} \quad \langle x, Mx + Q \rangle = 0, \tag{4.5}$$

where $M \in \mathbb{R}^{n \times n}$ and $q \in \mathbb{R}^n$ is equivalent to the absolute value equation. In particular, we have from [9, Proposition 2] that (4.5) is equivalent to (1.2) with $A = (M - I)^{-1}(M + I)$ and $c = (M - I)^{-1}q$ provided that 1 is not an eigenvalue of M. We now use this equivalence to solve standard test problems for LCP given in [40, Example 7.4], where M is given by $M = C^{\mathsf{T}}C + D + \mathrm{diag}(\eta)$, $q \in \mathbb{R}^n$ is randomly generated from $(-500, 0)$, and $n \in \{50, 100, 150, 200\}$. Here, $C, D \in \mathbb{R}^{n \times n}$ and $\eta \in \mathbb{R}^n$ are randomly generated such that $c_{ij}, d_{ij} \in (-5, 5)$, $\eta_i \in (0, 0.3)$ and D is skew-symmetric. Similar to the preceding example, we compare only MAP-LS with the other three algorithms, and the results are summarized in Table 3. All of the methods considered were able to solve all the test problems generated, with GNM

Table 2. Numerical results for Example 4.2

Method	n					
	50	100	500	1000	2000	3000
MAP-LS						
Success(%)	0.89	0.84	0.78	0.81	0.76	0.76
Ave. Time	0.00242	0.0056	0.21	1.40	10.81	36.67
Ave. Iter (MAP)	98.47	99.48	100	100	100	100
Ave. Iter (LS)	3.15	4.94	10.38	15.14	19.41	23.38
GNM						
Success(%)	0.68	0.71	0.84	0.84	0.63	0.22
Ave. Time	0.00048	0.0039	0.51	6.28	74.91	252.69
Ave. Iter	11.03	24.44	118.55	335.90	803.70	1067.32
PIM						
Success(%)	0	0	0	0	0	0
Ave. Time	–	–	–	–	–	–
Ave. Iter	–	–	–	–	–	–
GSM						
Success(%)	0.24	0.17	0	0	0	0
Ave. Time	0.18671	1.6259	–	–	–	–
Ave. Iter	3937.29	15413.41	–	–	–	–

achieving the best running time among all. On the other hand, MAP-LS, PIM and GSM have almost the same running time.

Example 4.4. We sample the entries of $A, B \in \mathbb{R}^{m\times n}$ and $x^* \in \mathbb{R}^n$ from the standard normal distribution, and we set $c = Ax^* + B|x^*|$. We let $n = 500$ and for each $m = rn$ with $r \in \{0.25, 0.5, 0.75, 1.5, 2, 3\}$, we generate 100 random AVEs and solve these problems using MAP and SLA. The results are summarized in Table 4. Observe that both algorithms were able to solve all the randomly generated problems. However, it is noticeable that the difference in the average CPU time spent in solving the test problems is quite significant. More specifically, the ratios of the average CPU time of SLA to the average CPU time of MAP for the six values of r considered are 516.60, 712.23, 247.15, 211.84, 1604.34 and 470.73, respectively. This shows the substantial difference in performance of the two algorithms.

Acknowledgements

J. H. Alcantara and J.-S. Chen's research is supported by Ministry of Science and Technology, Taiwan. M. K. Tam is supported in part by DE200100063 from the Australian Research Council. This work was conducted while J. H. Alcantara was a postdoctoral fellow at National Taiwan Normal University.

Table 3. Numerical results for Example 4.3

Method	n			
	50	100	150	200
MAP-LS				
Success(%)	1	1	1	1
Ave. Time	0.0044	0.010	0.019	0.029
Ave. Iter (MAP)	100.00	100.00	100.00	100.00
Ave. Iter (LS)	9.66	10.86	11.93	12.23
GNM				
Success(%)	1	1	1	1
Ave. Time	0.0003	0.001	0.002	0.003
Ave. Iter	7.62	8.09	8.75	8.84
PIM				
Success(%)	1	1	1	1
Ave. Time	0.0053	0.012	0.028	0.041
Ave. Iter	1110.21	1482.51	1805.79	2009.94
GSM				
Success(%)	1	1	1	1
Ave. Time	0.0071	0.014	0.027	0.036
Ave. Iter	137.27	125.72	119.02	113.2

Table 4. Numerical results for Example 4.4

Method	r					
	0.25	0.5	0.75	1.5	2	3
MAP						
Success rate	1	1	1	1	1	1
Ave. Time	0.01	0.03	0.26	0.12	0.02	0.19
Ave. Iter	104.19	296.34	2162.84	227.16	1	1
SLA						
Success rate	1	1	1	1	1	1
Ave. Time	4.21	19.69	63.60	26.11	31.33	90.31
Ave. Iter	2.38	3.64	6.11	1	1	1

The authors would like to thank the anonymous referee for the valuable feedback and comments.

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

References

[1] Rohn, J.: A theorem of alternatives for the equation $Ax + B|x| = b$. Linear Multilinear Algebra **52**, 421–426 (2004)

[2] Caccetta, L., Qu, B., Zhou, G.: A globally and quadratically convergent method for absolute value equations. Comput. Optim. Appl. **48**, 45–58 (2011)

[3] Cruz, J.Y.B., Ferreira, O.P., Prudente, L.F.: On the global convergence of the inexact semi-smooth Newton method for absolute value equation. Comput. Optim. Appl. **65**, 93–108 (2016)

[4] Haghani, F.K.: On generalized Traub's method for absolute value equations. J. Optim. Theory Appl. **166**, 619–625 (2015)

[5] Hu, S.-L., Huang, Z.-H.: A note on absolute value equations. Optim. Lett. **4**, 417–424 (2010)

[6] Mangasarian, O.L.: Absolute value equation solution via concave minimization. Optim. Lett. **1**, 3–8 (2007)

[7] Mangasarian, O.L.: A generalized Newton method for absolute value equation. Optim. Lett. **3**, 101–108 (2008)

[8] Mangasarian, O.L.: A hybrid algorithm for solving the absolute value equation. Optim. Lett. **9**, 1469–1474 (2015)

[9] Mangasarian, O.L., Meyer, R.R.: Absolute value equations. Linear Algebra Appl. **419**, 359–367 (2006)

[10] Rohn, J., Hooshyarbakhsh, V., Farhadsefat, R.: An iterative method for solving absolute value equations and sufficient conditions for unique solvability. Optim. Lett. **8**, 35–44 (2014)

[11] Zhang, C., Wei, Q.J.: Global and finite convergence of a generalized Newton method for absolute value equations. J. Optim. Theory Appl. **143**, 391–403 (2009)

[12] Cottle, R.W., Dantzig, G.: Complementary pivot theory of mathematical programming. Linear Algebra Appl. **1**, 103–125 (1968)

[13] Cottle, R.W., Pang, J.-S., Stone, R.-E.: The Linear Complementarity Problem. Academic Press, New York (1992)

[14] Mangasarian, O.L.: Absolute value programming. Comput. Optim. Appl. **36**, 43–53 (2007)

[15] Prokopyev, O.: On equivalent reformulations for absolute value equations. Comput. Optim. Appl. **44**, 363–372 (2009)

[16] Rohn, J.: Systems of linear interval equations. Linear Algebra Appl. **126**, 39–78 (1989)

[17] Wu, S.-L., Li, C.-X.: A note on unique solvability of the absolute value equation. Optim. Lett. **14**, 1957–1960 (2020)

[18] Saheya, B., Yu, C.-H., Chen, J.-S.: Numerical comparisons based on four smoothing functions for absolute value equation. J. Appl. Math. Comput. **56**, 131–149 (2018)

[19] Salkuyeh, D.K.: The Picard-HSS iteration method for absolute value equation. Optim. Lett. **8**, 2191–2202 (2014)

[20] Chen, C., Yu, D., Han, D.: An inexact Douglas–Rachford splitting method for solving absolute value equations. arXiv:2103.09398 [math.OC] (2021)

[21] Ke, Y.-F., Ma, C.-F.: SOR-like iteration method for solving absolute value equations. Appl. Math. Comput. **311**, 195–202 (2017)

[22] Edalatpour, V., Hezari, D., Salkuyeh, D.K.: A generalization of the Gauss-Seidel iteration method for solving absolute value equations. Appl. Math. Comput. **293**, 156–167 (2017)

[23] Abdallah, L., Haddou, M., Migot, T.: Solving absolute value equation using complementarity and smoothing functions. J. Comput. Appl. Math. **327**, 196–207 (2018)

[24] Dao, M.N., Tam, M.K.: Union averaged operators with applications to proximal algorithms for min-convex functions. J. Optim. Theory Appl. **181**, 61–94 (2019)

[25] Bauschke, H.H., Noll, D.: On the local convergence of the Douglas–Rachford algorithm. Arch. Math. **102**, 589–600 (2014)

[26] Tam, M.K.: Algorithms based on unions of nonexpansive maps. Optim. Lett. **12**, 1019–1027 (2018)

[27] Attouch, H., Bolte, J., Svaiter, B.F.: Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward–backward splitting, and regularized Gauss-Seidel methods. Math. Program. **137**, 91–129 (2013)

[28] Bauschke, H.H., Kruk, S.G.: Reflection-projection method for convex feasibility problems with an obtuse cone. J. Optim. Theory Appl. **120**, 503–531 (2004)

[29] Bregman, L.M.: The method of successive projections for finding a common point of convex sets. Sov. Math. Dokl. **162**, 688–692 (1965)

[30] Alcantara, J.H., Chen, J.-S.: A novel generalization of the natural residual function and a neural network approach for the NCP. Neurocomputing **413**, 368–382 (2020)

[31] Galantai, A.: Properties and construction of NCP functions. Comput. Optim. Appl. **52**, 805–824 (2012)

[32] Alcantara, J.H., Lee, C.-H., Nguyen, C.T., Chang, Y.-L., Chen, J.-S.: On construction of new NCP functions. Oper. Res. Lett. **48**, 115–121 (2020)

[33] Hesse, R., Luke, D.R., Neumann, P.: Alternating projections and Douglas–Rachford for sparse affine feasibility. IEEE Trans. Signal Process. **62**, 4868–4881 (2014)

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

[35] Scholtes, S.: Introduction to Piecewise Differentiable Functions. Springer, Berlin (2012)

[36] Horn, R.A., Johnson, C.R.: Topics in Matrix Analysis. Cambridge University Press, Cambridge (1991)

[37] Lewis, A.S., Luke, D.R., Malick, J.: Local linear convergence for alternating and averaged nonconvex projections. Found. Comput. Math. **9**, 485–513 (2009)

[38] Danillidis, A., Luke, D.R., Tam, M.K.: Characterizations of super-regularity and its variants. In: Bauschke, H.H., Burachik, R.S., Luke, D.R. (eds.) Splitting Algorithms, Modern Operator Theory, and Applications, pp. 137–152. Springer, Berlin (2019)

[39] Tam, M.K.: Regularity properties of non-negative sparsity sets. J. Math. Anal. Appl. **447**, 758–777 (2017)

[40] Kanzow, C.: Some noninterior continuation methods for linear complementarity problems. SIAM J. Matrix Anal. Appl. **17**, 851–868 (1996)

Jan Harold Alcantara
Institute of Statistical Sciences
Academia Sinica Taipei11529
Taiwan
e-mail: janharold@stat.sinica.edu.tw

Jein-Shan Chen
Department of Mathematics
National Taiwan Normal University
Taipei 11677
Taiwan
e-mail: jschen@math.ntnu.edu.tw

Matthew K. Tam
School of Mathematics and Statistics
The University of Melbourne
Parkville VIC3010
Australia
e-mail: matthew.tam@unimelb.edu.au

Accepted: February 10, 2022.