IMA Journal of Numerical Analysis (2023) **43**, 1036–1060
https://doi.org/10.1093/imanum/drab105
Advance Access publication on 11 February 2022

Exact and inexact Douglas–Rachford splitting methods for solving large-scale sparse absolute value equations

Cairong Chen
School of Mathematics and Statistics, FJKLMAA and Center for Applied Mathematics of Fujian Province, Fujian Normal University, Fuzhou 350007, P.R. China

Dongmei Yu
Institute for Optimization and Decision Analytics, Liaoning Technical University, Fuxin 123000, P.R. China

and

Deren Han*
LMIB of the Ministry Education, School of Mathematical Sciences, Beihang University, Beijing 100191, P.R. China
*Corresponding author: handr@buaa.edu.cn

[Received on 30 March 2020; revised on 18 December 2021]

Exact and inexact Douglas–Rachford splitting methods are developed to solve the large-scale sparse absolute value equation (AVE) $Ax - |x| = b$, where $A \in \mathbb{R}^{n\times n}$ and $b \in \mathbb{R}^n$. The inexact method adopts a relative error tolerance and, therefore, in the inner iterative processes, the LSQR method is employed to find a qualified approximate solution of each subproblem, resulting in a lower cost for each iteration. When $\|A^{-1}\| \leq 1$ and the solution set of the AVE is nonempty, the algorithms are globally and linearly convergent. When $\|A^{-1}\| = 1$ and the solution set of the AVE is empty, the sequence generated by the exact algorithm diverges to infinity on a trivial example. Numerical examples are presented to demonstrate the viability and robustness of the proposed methods.

Keywords: absolute value equation; Douglas–Rachford splitting method; exact and inexact; global and linear convergence; LSQR.

1. Introduction

Consider the problem of solving the absolute value equation (AVE) of the form

$$Ax - |x| - b = 0, \tag{1.1}$$

where $A \in \mathbb{R}^{n\times n}$ and $b \in \mathbb{R}^n$ are given and $|x| = (|x_1|, \ldots, |x_n|)^{\mathrm{T}} \in \mathbb{R}^n$ denotes the componentwise absolute value of the unknown vector x. AVE (1.1) is a particular case of the generalized absolute value equation (GAVE)

$$Ax + B|x| - b = 0 \tag{1.2}$$

with $A, B \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. Clearly, if $m = n$ and B is nonsingular, then GAVE (1.2) can be transformed into AVE (1.1). GAVE (1.2) was formally introduced by Rohn (2004) and further investigated in Mangasarian (2007a), Prokopyev (2009) and the references therein.

© The Author(s) 2022. Published by Oxford University Press on behalf of the Institute of Mathematics and its Applications. All rights reserved.

Downloaded from https://academic.oup.com/imajna/article/43/2/1036/6526317 by Zhongshan University user on 25 March 2025

AVE (1.1) and GAVE (1.2) are intimately related to many significant optimization problems, such as the linear complementarity problem (LCP) and mixed-integer programming; see, e.g., Mangasarian (2007a), Mangasarian & Meyer (2006), Prokopyev (2009). The LCP subsumes many mathematical programming problems (Cottle *et al.*, 1992), and mixed-integer programming has various real applications (Bixby *et al.*, 2004). In addition, AVE (1.1) and GAVE (1.2) also arise from a system of linear interval equations (Rohn, 1989), which is one of the critical problems encountered in scientific and engineering computing (Xia & Li, 2020). Both AVE (1.1) and GAVE (1.2) have received considerable attention since their appearance, and abundant theoretical results and numerical strategies for the equations have been established.

Owing to the combinatorial character introduced by the absolute value operator, solving GAVE (1.2) is NP-hard (Mangasarian, 2007a, Proposition 2). Moreover, if GAVE (1.2) is solvable, checking whether GAVE (1.2) has a unique solution or multiple solutions is NP-complete (Prokopyev, 2009, Proposition 2.1). It should be emphasized that even if AVE (1.1) can be converted into an LCP (Hu & Huang, 2010) and LCP possesses a polynomial-time algorithm if the underlying matrix is positive definite (Kojima, 1989), solving AVE (1.1) is still NP-hard because the resulting matrix is generally not positive definite. For some classes of AVEs (1.1) that are polynomial-time solvable one is referred to Zamani & Hladík (2021). Research has thus focused on constructing conditions under which AVE (1.1) has a unique solution for any $b \in \mathbb{R}^n$. In the past two decades some sufficient conditions have been given, for instance[1] $\|A^{-1}\| < 1$ in Mangasarian & Meyer (2006, Proposition 3), $\varrho(|A^{-1}|) < 1$ in Rohn *et al.* (2014), and the regularity of the interval matrix $[A-I, A+I]$ in Zhang & Wei (2009). The sufficient condition $\|A^{-1}\| < 1$ is not necessary for the unique solvability of AVE (1.1), and a counterexample can be found in Zhang & Wei (2009, Example 2.1) (see Section 3 for another example). In addition, the sufficient conditions $\|A^{-1}\| < 1$ and $\varrho(|A^{-1}|) < 1$ are independent of each other (Rohn *et al.*, 2014). Moreover, the interval matrix $[A - I, A + I]$ is regular if and only if AVE (1.1) has a unique solution for any $b \in \mathbb{R}^n$ (Wu & Li, 2018, Theorem 3.2 or Theorem 3.3). Other sufficient and necessary conditions can be obtained from Mezzadri (2020) and Wu & Shen (2021).

Some complementarity reformulations of AVE (1.1) have been explored in the literature. For instance AVE (1.1) is equivalent to the LCP of finding a vector $u \in \mathbb{R}^l$ such that

$$u \geq 0, \quad Mu + q \geq 0 \quad \text{and} \quad \langle u, Mu + q\rangle = 0. \tag{1.3}$$

One such reformulation is given in Prokopyev (2009, Proposition 3.1), where AVE (1.1) is reformulated to (1.3) with

$$M = \begin{bmatrix} -I & 2I & 0 \\ A & -(I+A) & 0 \\ -A & A+I & 0 \end{bmatrix}. \tag{1.4}$$

Note that here $l = 3n$, i.e., the number of the variables in LCP (1.3) with (1.4) is three times the number of variables in the original problem (1.1). The alternative LCP reformulation

$$u \geq 0, \quad (A+I)(A-I)^{-1}u + q \geq 0 \quad \text{and} \quad \langle u, (A+I)(A-I)^{-1}u + q\rangle = 0 \tag{1.5}$$

[1] Here and in the sequel, $\|\cdot\|$ denotes the 2-norm of a matrix or a vector, $\varrho(\cdot)$ denotes the spectral radius of a matrix and I is the identity matrix with suitable dimension.

Downloaded from https://academic.oup.com/imajna/article/43/2/1036/6526317 by Zhongshan University user on 25 March 2025

is proposed in Mangasarian (2007a) with

$$q = (A - I)^{-1}b \quad \text{and} \quad u = \frac{1}{2}[(A - I)x - b],$$

and in Mangasarian & Meyer (2006) with

$$q = \left[(A + I)(A - I)^{-1} - I\right]b \quad \text{and} \quad u = (A - I)x - b.$$

Though the reformulation (1.5) does not increase the number of variables it is unavailable to apply in the case that 1 is an eigenvalue of A. Whereafter, without giving any condition on A, AVE (1.1) is equivalently reformulated in the form of LCP (1.3) with $l = n$ (Hu & Huang, 2010). Nevertheless, the reformulation developed in Hu & Huang (2010) involves the inverse of the matrix $AD - I$, where D is a diagonal matrix determined by an index set. In Abdallah *et al.* (2018) it was shown that AVE (1.1) is equivalent to a horizontal LCP. In Mangasarian & Meyer (2006) it was proved that AVE (1.1) is equivalent to the problem of finding an $x \in \mathbb{R}^n$ such that

$$Q(x) \doteq Ax + x - b \geq 0, \quad F(x) \doteq Ax - x - b \geq 0 \quad \text{and} \quad \langle Q(x), F(x)\rangle = 0. \tag{1.6}$$

Problem (1.6) is a generalized linear complementarity problem (GLCP) (Facchinei & Pang, 2003). The aforementioned complementarity problems are ubiquitous in mathematical programming and such equivalent reformulations have attracted more and more interest in the optimization community.

In the literature a series of numerical methods have been developed to solve AVE (1.1). Mangasarian (2007a,b) suggested solving AVE (1.1) via concave minimization and proposed solving the concave minimization problem by a finite succession of linear programs. Mangasarian's method stops in a finite number of steps at a point satisfying the necessary optimality condition for the concave minimization problem (Mangasarian, 2007a,b). However, this point is not necessarily a solution of AVE (1.1); see Zamani & Hladík (2021, Example 1). Recently, Zamani & Hladík (2021) proposed a new concave minimization algorithm, which terminated in a finite number of steps at a solution of AVE (1.1) under mild conditions. Since $|x|$ is nonlinear and nonsmooth, AVE (1.1) can be viewed as a special case of the nonlinear equation

$$\mathscr{F}(x) = 0,$$

where $\mathscr{F}$ is a nonsmooth mapping from $\mathbb{R}^n$ to $\mathbb{R}^n$. From this point of view numerical algorithms for solving nonsmooth equations can be adopted to find a solution of AVE (1.1). In Mangasarian (2009) a generalized Newton method (also known as the exact semismooth Newton method) was directly utilized to solve AVE (1.1). If $\|A^{-1}\| < 1/4$ (it can be relaxed to $\|A^{-1}\| < 1/3$; Lian *et al.*, 2018) and $\|\mathcal{D}(x^k)\| \neq 0$[2] then the generalized Newton method is well defined and the generated sequence converges linearly to the unique solution x^* of AVE (1.1), and the reported numerical results in Mangasarian (2009) are promising. Shortly after, a smoothing Newton method was proposed in Caccetta *et al.* (2011). Under the condition that $\|A^{-1}\| < 1$, the smoothing Newton method is globally convergent and the convergence

[2] Throughout this paper $\mathcal{D}(x) \doteq \operatorname{diag}(\operatorname{sign}(x))$ with $\operatorname{sign}(x)$ denoting a vector with components equal to -1, 0 or 1, depending on whether the corresponding element in the vector x is negative, zero or positive respectively. For $x \in \mathbb{R}^n$, $\operatorname{diag}(x)$ represents a diagonal matrix with x_i as its diagonal entries for every $i = 1, 2, \ldots, n$.

Downloaded from https://academic.oup.com/imajna/article/43/2/1036/6526317 by Zhongshan University user on 25 March 2025

rate is quadratic. To alleviate the burden of solving the system of linear equations during the Newton iteration, the inexact semismooth Newton method was presented in Cruz *et al.* (2016). As stated in Cruz *et al.* (2016), if $\|A^{-1}\| < 1/3$ and the residual relative error tolerance θ satisfies

$$0 \leq \theta < \frac{1 - 3\|A^{-1}\|}{\|A^{-1}\|(\|A\| + 3)}, \tag{1.7}$$

then the inexact semismooth Newton method is well defined and the generated sequence globally converges to the unique solution x^* of AVE (1.1), and its rate of convergence is Q-linear.[3] By using the matrix splitting technique a generalized Gauss–Seidel iterative method was developed to solve AVE (1.1) in Edalatpour *et al.* (2017). By introducing an auxiliary variable and using the idea of the classical successive overrelaxation (SOR) iterative method (Saad, 2003), an SOR-like iterative method was proposed in Ke & Ma (2017) for solving AVE (1.1). If $\|A^{-1}\| < 1$ the sequence generated by the SOR-like iterative method will converge whenever

$$\frac{1+\sqrt{5}}{3+\sqrt{5}} < \omega < \min\left\{\frac{5+\sqrt{5}}{3+\sqrt{5}}, \sqrt{\frac{2}{(3+\sqrt{5})\|A^{-1}\|}}\right\}$$

(Ke & Ma, 2017), where ω is the iterative parameter. For more numerical algorithms we refer the interested reader to, e.g., Gao & Wang (2014), Iqbal *et al.* (2015), Gu *et al.* (2017), Mansoori *et al.* (2017), Miao *et al.* (2017), Wang *et al.* (2017), Abdallah *et al.* (2018), Mansoori & Erfanian (2018), Saheya *et al.* (2018), Ke (2020); Chen *et al.* (2021) and the references therein.

The main purpose of this study is to develop numerical algorithms for solving the large-scale sparse AVE (1.1). Our methods have two main features: (i) the new algorithms can be more efficient than some existing methods for the case where $\|A^{-1}\| < 1$; (ii) when the solution set of AVE (1.1) is nonempty the new algorithms can deal with the case where $\|A^{-1}\| = 1$. Note that for the general case discussed in this paper, i.e., $\|A^{-1}\| \leq 1$, the existing algorithms to solve (1.1) via solving (1.5) are infeasible due to the possible singularity of $A - I$. In addition, if AVE (1.1) is reformulated to LCP (1.3) with (1.4), the scale will increase.

The new methods are based on reformulation (1.6), which, among the equivalent forms, is the most economical one in the sense that it is inverse-free and it does not increase the number of variables. Our methods are inspired by the operator splitting methods for solving the generalized equation

$$0 \in T(x),$$

where $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a set-valued mapping that has the form $T := \Phi + \Psi$, with two maximal monotone mappings Φ and Ψ. We focus on the recently resurrected Douglas–Rachford splitting (DRs) method (Douglas & Rachford, 1956), whose recursion is very simple: at the kth iteration the next iteration x^{k+1} is generated by solving the linear system

$$2Ax = 2Ax^k - \gamma\rho(x^k)G^{-1}(Ax^k - |x^k| - b), \tag{1.8}$$

where G is a symmetric positive definite (SPD) matrix functioning as a preconditioner and $\rho(x^k)$ is a parameter that can be evaluated easily. To reduce the computational cost per iteration we allow equation (1.8) to be solved approximately with a constructive accuracy criterion. Since $\|A^{-1}\| \leq 1$ is concerned,

[3] A sequence $\{x^k\} \subset \mathbb{R}^n$ converging to x^* is said to converge at a Q-linear rate if $\|x^{k+1} - x^*\| \leq \kappa\|x^k - x^*\|$ with $0 < \kappa < 1$.

Downloaded from https://academic.oup.com/imajna/article/43/2/1036/6526317 by Zhongshan University user on 25 March 2025

two possible cases need to be considered, i.e., the solution set of AVE (1.1) is nonempty, and AVE (1.1) does not possess a solution. For the first case we show that the new algorithms converge at a global linear rate to a solution of AVE (1.1). For the second case a trivial example is given to investigate the behavior of our methods. As a by-product of this paper we also derive a new sufficient condition ensuring the unique solvability of AVE (1.1). Though it is stronger than the popular one, $\|A^{-1}\| < 1$, the new condition provides some useful insights into the problem and inspires a new proof of the sufficient condition $\|A^{-1}\| < 1$.

Before ending this section we summarize the algorithm frameworks of the exact semismooth Newton method, the inexact semismooth Newton method and the SOR-like iterative method because they will be numerically compared with our methods.

ALGORITHM 1.1 [Mangasarian, 2009; Exact semismooth Newton method] Let $x^0 \in \mathbb{R}^n$ be an initial guess. For $k = 0, 1, \ldots$, until convergence, compute x^{k+1} according to

$$\left[A - \mathcal{D}(x^k)\right] x^{k+1} = b. \tag{1.9}$$

ALGORITHM 1.2 [Cruz *et al.*, 2016; Inexact semismooth Newton method] Let $x^0 \in \mathbb{R}^n$ be an initial guess. For $k = 0, 1, \ldots$, until convergence, find an x^{k+1} such that

$$\left[A - \mathcal{D}(x^k)\right] x^{k+1} = b + r_k, \quad \text{where} \quad \|r_k\| \le \theta \|Ax^k - |x^k| - b\|,$$

and θ is defined as in (1.7).

ALGORITHM 1.3 [Ke & Ma, 2017; SOR-like iterative method] Let $x^0 \in \mathbb{R}^n$ and $y^0 \in \mathbb{R}^n$ be initial guesses. For $k = 0, 1, \ldots$, until convergence, compute

$$\begin{cases} x^{k+1} = (1-\omega)x^k + \omega A^{-1}(y^k + b), \\ y^{k+1} = (1-\omega)y^k + \omega |x^{k+1}|, \end{cases} \tag{1.10}$$

where ω is the iterative parameter.

The paper is organized as follows. In Section 2 we present some classical definitions and preliminary results relevant to our later developments. Section 3 gives a new sufficient condition for AVE (1.1) to be uniquely solvable for any $b \in \mathbb{R}^n$. Section 4 describes the exact and inexact DRs methods and proves their contractive properties and convergence. In Section 5, five numerical examples are given to demonstrate our claims. Some concluding remarks are given in Section 6. Finally, in Appendix A, we give a new proof of the sufficient condition $\|A^{-1}\| < 1$.

Notation. We denote by $\mathbb{R}^{n\times n}$ the set of all $n \times n$ real matrices, $\mathbb{R}^n = \mathbb{R}^{n\times 1}$ and $\mathbb{R} = \mathbb{R}^1$, and by $|\cdot|$ the absolute value for real scalar. For a complex scalar λ, $\mathrm{Re}(\lambda)$ denotes the real part of λ. The largest eigenvalue and the smallest eigenvalue of A are denoted by $\lambda_{\max}(A)$ and $\lambda_{\min}(A)$, respectively, and $\sigma_{\min}(A)$ denotes the smallest singular value of A. The transpose of a matrix or a vector is denoted by $\cdot^{\mathrm{T}}$. The inner product of two vectors in $\mathbb{R}^n$ is defined as $\langle x, y\rangle \doteq x^{\mathrm{T}}y = \sum_{i=1}^n x_i y_i$ and it holds that $\|x\| \doteq \sqrt{\langle x, x\rangle}$. By $\mathrm{tridiag}(a, b, c)$ we denote a matrix that has a, b, c as the subdiagonal, main diagonal, superdiagonal entries in the matrix, respectively. The projection mapping from $\mathbb{R}^n$ onto Ω, denoted by P_Ω, is defined as $P_\Omega[x] = \arg\min\{\|x - y\| : y \in \Omega\}$.

Downloaded from https://academic.oup.com/imajna/article/43/2/1036/6526317 by Zhongshan University user on 25 March 2025

2. Preliminaries

In this section we collect some mathematical concepts and facts required in our later technical arguments.

A matrix $A \in \mathbb{R}^{n\times n}$ is said to be positive definite if $\langle Au, u\rangle > 0$ for all $0 \neq u \in \mathbb{R}^n$; see, e.g., Saad (2003, p. 29). If, in addition, a positive definite matrix $A \in \mathbb{R}^{n\times n}$ is symmetric then A is said to be SPD. It is well known that $A \in \mathbb{R}^{n\times n}$ is positive definite if and only if the symmetric part of A is SPD. Furthermore, if $A \in \mathbb{R}^{n\times n}$ is positive definite, then $\mathrm{Re}(\lambda) > 0$ for any eigenvalue λ of A; see, e.g., Saad (2003, pp. 30–31). When a matrix $A \in \mathbb{R}^{n\times n}$ is SPD, the mapping $x, y \to \langle x, y\rangle_A \doteq x^{\mathrm{T}}Ay$ from $\mathbb{R}^n \times \mathbb{R}^n$ to $\mathbb{R}$ is an inner product on $\mathbb{R}^n$ and we denote $\|x\|_A = \sqrt{\langle x, x\rangle_A}$. A matrix $A \in \mathbb{R}^{n\times n}$ is called a P-matrix if all its principal minors are positive (Cottle *et al.*, 1992, p. 147). A real symmetric matrix is positive definite if and only if it is a P-matrix.

We suppose temporarily that F and Q are two continuous linear mappings from $\mathbb{R}^n$ to $\mathbb{R}^n$. Then the GLCP involves finding an $x \in \mathbb{R}^n$ such that

$$F(x) \geq 0, \quad Q(x) \geq 0 \quad \text{and} \quad \langle F(x), Q(x)\rangle = 0. \tag{2.1}$$

GLCP (2.1) is monotone if

$$[Q(u) - Q(v)]^{\mathrm{T}}[F(u) - F(v)] \geq 0 \ \forall\, u, v \in \mathbb{R}^n,$$

and it is strongly monotone if there exists a $\mu > 0$ such that

$$[Q(u) - Q(v)]^{\mathrm{T}}[F(u) - F(v)] \geq \mu\|u - v\|^2 \ \forall\, u, v \in \mathbb{R}^n.$$

The linear operator $Q + F$ defined as in (2.1) is nonsingular (as defined by He, 1999) if there is a $\nu > 0$ such that

$$\|(Q + F)(u - v)\| \geq \nu\|u - v\| \ \forall\, u, v \in \mathbb{R}^n.$$

GLCP (2.1) is a special case of the following generalized variational inequality (GVI) problem:

$$\text{find an } x^* \in \mathbb{R}^n, \text{ such that } Q(x^*) \in \Omega, \ \left\langle v - Q(x^*), F(x^*)\right\rangle \geq 0 \ \forall\, v \in \Omega, \tag{2.2}$$

where Ω is a nonempty closed convex subset of $\mathbb{R}^n$. In fact, GLCP (2.1) corresponds to the particular case where $\Omega = \{x \in \mathbb{R}^n : x \geq 0\}$. When $Q(x) \equiv x$, the GVI problem (2.2) is reduced to the classical variational inequality (VI) problem. The VI problem has important applications in many fields, including economics, operations research and nonlinear analysis. A survey of methods for solving the VI problem can be found in Harker & Pang (1990), Cottle *et al.* (1992) and Facchinei & Pang (2003).

A fundamental property of the projection from $\mathbb{R}^n$ onto Ω is the following. Given $u \in \mathbb{R}^n$ and a nonempty closed convex subset Ω of $\mathbb{R}^n$, w is the projection of u onto Ω, i.e., $w = P_\Omega[u]$ if and only if

$$\langle u - w, v - w\rangle \leq 0 \ \forall\, v \in \Omega. \tag{2.3}$$

Downloaded from https://academic.oup.com/imajna/article/43/2/1036/6526317 by Zhongshan University user on 25 March 2025

It follows from this property that x^* is a solution of GVI (2.2) if and only if it solves the nonsmooth projection equation

$$Q(x) = P_{\Omega}\left[Q(x) - F(x)\right]. \tag{2.4}$$

Now we consider GLCP (1.6) arising from AVE (1.1). In such a case we have

$$F(x) = Ax - x - b \quad \text{and} \quad Q(x) = Ax + x - b.$$

Unless stated otherwise, throughout the rest of this paper, F and Q are reserved for the ones here. Consequently, for any $u, v \in \mathbb{R}^n$, we have

$$\begin{aligned}[Q(u) - Q(v)]^{\mathrm{T}}[F(u) - F(v)] &= [A(u-v) + (u-v)]^{\mathrm{T}}[A(u-v) - (u-v)] \\ &= (u-v)^{\mathrm{T}}A^{\mathrm{T}}A(u-v) - (u-v)^{\mathrm{T}}A^{\mathrm{T}}(u-v) \\ &\quad + (u-v)^{\mathrm{T}}A(u-v) - (u-v)^{\mathrm{T}}(u-v) \\ &= (u-v)^{\mathrm{T}}(A^{\mathrm{T}}A - I)(u-v),\end{aligned}$$

which implies that GLCP (1.6) is monotone if and only if $\sigma_{\min}(A) \geq 1$ (or, equivalently, $\|A^{-1}\| \leq 1$), and it is strongly monotone if and only if $\sigma_{\min}(A) > 1$ (or, equivalently, $\|A^{-1}\| < 1$). Thus, according to Mangasarian & Meyer (2006, Proposition 3), AVE (1.1) has a unique solution for any $b \in \mathbb{R}^n$ if the corresponding GLCP (1.6) is strongly monotone. On the other hand,

$$\begin{aligned}\|(Q+F)(u-v)\|^2 &= \|(Q+F)(u) - (Q+F)(v)\|^2 \\ &= \|2Au - 2b - (2Av - 2b)\|^2 \\ &= 4(u-v)^{\mathrm{T}}A^{\mathrm{T}}A(u-v) \\ &\geq 4\sigma_{\min}^2(A)\|u-v\|^2.\end{aligned}$$

Hence, a necessary and sufficient condition for the linear operator $Q + F$ defined as in (1.6) being nonsingular is that A is nonsingular.

Let

$$e(x) \doteq Q(x) - P_{\Omega}\left[Q(x) - F(x)\right] \tag{2.5}$$

be the residual of equation (2.4); the following theorem reveals the relationship between the solutions of AVE (1.1) and the zeros of $e(x)$.

THEOREM 2.1 For (2.5) we have

$$e(x) = Ax - |x| - b. \tag{2.6}$$

Thus, x^* is a solution of AVE (1.1) if and only if $e(x^*) = 0$.

Downloaded from https://academic.oup.com/imajna/article/43/2/1036/6526317 by Zhongshan University user on 25 March 2025

Proof. For the first assertion note that

$$P_{\Omega}[x] = \max\{x, 0\} = (|x| + x)/2$$

for $\Omega = \{x \in \mathbb{R}^n : x \geq 0\}$. Hence, it turns out from the definitions of Q and F in (1.6) that

$$\begin{aligned} e(x) &= Q(x) - P_{\Omega}\left[Q(x) - F(x)\right] \\ &= Ax + x - b - P_{\Omega}\left[Ax + x - b - (Ax - x - b)\right] \\ &= Ax + x - b - P_{\Omega}\left[2x\right] \\ &= Ax + x - b - (|x| + x) \\ &= Ax - |x| - b. \end{aligned}$$

Then the proof of the second part of this theorem is trivial. □

The proof of the next theorem is inspired by that of He (1999), Theorem 2.

THEOREM 2.2 If x^* is a solution of AVE (1.1) then

$$\left\langle A(x - x^*), e(x)\right\rangle \geq \frac{1}{2}\left\{\|e(x)\|^2 + \left\langle x - x^*, (A^{\mathrm{T}}A - I)(x - x^*)\right\rangle\right\} \ \forall\, x \in \mathbb{R}^n. \tag{2.7}$$

Proof. Recall the equivalence between AVE (1.1) and GLCP (1.6). Then x^* is also a solution of GLCP (1.6), which implies that $Q(x^*) \in \Omega$ and

$$\langle v - Q(x^*), F(x^*)\rangle \geq 0 \ \forall\, v \in \Omega,$$

where $\Omega = \{x \in \mathbb{R}^n : x \geq 0\} \subseteq \mathbb{R}^n$. Let $x \in \mathbb{R}^n$ be an arbitrary vector. We have, upon setting $v := P_{\Omega}[Q(x) - F(x)]$,

$$\langle P_{\Omega}[Q(x) - F(x)] - Q(x^*), F(x^*)\rangle \geq 0 \ \forall\, x \in \mathbb{R}^n. \tag{2.8}$$

On the other hand, setting $u := Q(x) - F(x)$ and $v := Q(x^*)$, then it follows from (2.3) that

$$\langle Q(x) - F(x) - P_{\Omega}[Q(x) - F(x)], P_{\Omega}[Q(x) - F(x)] - Q(x^*)\rangle \geq 0 \ \forall\, x \in \mathbb{R}^n. \tag{2.9}$$

Adding (2.8) and (2.9), and using (2.5), we obtain

$$\langle Q(x) - Q(x^*) - e(x), e(x) - F(x) + F(x^*)\rangle \geq 0 \ \forall\, x \in \mathbb{R}^n.$$

Rearranging the above inequality we get

$$\begin{aligned} &\left\langle Q(x) - Q(x^*) + F(x) - F(x^*), e(x)\right\rangle \\ &\qquad \geq \|e(x)\|^2 + \langle Q(x) - Q(x^*), F(x) - F(x^*)\rangle. \end{aligned} \tag{2.10}$$

Downloaded from https://academic.oup.com/imajna/article/43/2/1036/6526317 by Zhongshan University user on 25 March 2025

In addition, it follows from the definitions of Q and F in (1.6) that

$$Q(x) + F(x) = 2(Ax - b) \tag{2.11}$$

and

$$\langle Q(x) - Q(x^*), F(x) - F(x^*)\rangle = \langle x - x^*, (A^{\mathrm{T}}A - I)(x - x^*)\rangle. \tag{2.12}$$

The assertion then follows from (2.10)–(2.12). □

In what follows we turn to the *Fejér monotonicity* (Bauschke & Combettes, 2011, Definition 5.1), which is useful in our convergence analysis. The Fejér monotonicity of a sequence $\{v^k\}$ with respect to a closed convex set S says that for two consecutive points v^{k+1} and v^k, we have

$$\operatorname{dist}(v^{k+1}, S) \leq \operatorname{dist}(v^k, S),$$

where

$$\operatorname{dist}(v, S) := \inf_{u \in S} \|u - v\|$$

denotes the distance between a point v and the set S. A Fejér monotonicity technique (see, for instance, Bot *et al.*, 2019, Lemma 1.2) is summarized in the following lemma.

LEMMA 2.3 Let $\{\xi_n\}$ be a sequence of real numbers and $\{\omega_n\}$ be a sequence of real non-negative numbers. Assume that $\{\xi_n\}$ is bounded from below and that, for any $n \geq 1$,

$$\xi_{n+1} \leq \xi_n - \omega_n.$$

Then the following statements hold:

(i) the sequence $\{\omega_n\}$ is summable, i.e., $\sum_{n=1}^{\infty} \omega_n < \infty$;

(ii) the sequence $\{\xi_n\}$ is monotonically decreasing and convergent.

Our proofs of the convergence of the proposed algorithms are based on Theorem 2.2 and the Fejér monotonicity of the iterations with respect to the solution set.

3. A new sufficient condition

In this section we will develop a new sufficient condition for the unique solvability of AVE (1.1).

Let

$$\mathcal{F}(x) = x - \nu e(x)$$

with a real number $\nu \neq 0$; then Theorem 2.1 implies that x^* is a solution of AVE (1.1) if and only if $\mathcal{F}(x^*) = x^*$, that is, x^* is a fixed point of $\mathcal{F}$. This provides a way to derive a sufficient condition

Downloaded from https://academic.oup.com/imajna/article/43/2/1036/6526317 by Zhongshan University user on 25 March 2025

for AVE (1.1) to have a unique solution. As a matter of fact it follows from the triangle inequality and $\||x|-|y|\| \le \|x-y\|$ that

$$\begin{aligned}\|\mathcal{F}(x)-\mathcal{F}(y)\| &= \|(I-\nu A)(x-y)+\nu(|x|-|y|)\| \\ &\le \|I-\nu A\|\cdot\|x-y\|+\nu\|x-y\| \\ &= (\|I-\nu A\|+\nu)\|x-y\| \quad \forall x,\ y\in\mathbb{R}^n.\end{aligned}$$

Therefore, the mapping $\mathcal{F}$ is contractive if $\|I-\nu A\|+\nu \in (0,1)$. According to the Banach fixed-point theorem (see, e.g., Facchinei & Pang, 2003, p. 144), if there exists a $\nu\in(0,1)$ such that $\|I-\nu A\|<1-\nu$, then the mapping $\mathcal{F}$ has a unique fixed point in $\mathbb{R}^n$, which means that AVE (1.1) has a unique solution for any $b\in\mathbb{R}^n$. Apparently, the above discussion gives rise to the following theorem.

THEOREM 3.1 AVE (1.1) is uniquely solvable for any $b\in\mathbb{R}^n$ if there exists a $\nu\in(0,1)$ such that $\|I-\nu A\|<1-\nu$.

Now we are in position to give a new sufficient condition for AVE (1.1) to be uniquely solvable for any $b\in\mathbb{R}^n$. Note that for $\nu\in(0,1)$,

$$\begin{aligned}\|I-\nu A\|<1-\nu &\iff \|I-\nu A\|^2<(1-\nu)^2 \\ &\iff \lambda_{\max}\left((I-\nu A)^{\mathrm{T}}(I-\nu A)\right)<(1-\nu)^2 \\ &\iff \max_{0\neq x\in\mathbb{R}^n}\frac{x^{\mathrm{T}}(I-\nu A)^{\mathrm{T}}(I-\nu A)x}{x^{\mathrm{T}}x}<(1-\nu)^2 \\ &\iff \left[x^{\mathrm{T}}(A^{\mathrm{T}}A-I)x\right]\cdot\nu^2-\left[2x^{\mathrm{T}}(A-I)x\right]\cdot\nu<0\ \forall\ 0\neq x\in\mathbb{R}^n.\end{aligned}$$

Then there exists a $\nu\in(0,1)$ such that $\|I-\nu A\|<1-\nu$ if one of the following conditions holds:

(a) $x^{\mathrm{T}}(A^{\mathrm{T}}A-I)x>0$ and $x^{\mathrm{T}}(A-I)x>0$ for all $0\neq x\in\mathbb{R}^n$;

(b) $x^{\mathrm{T}}(A^{\mathrm{T}}A-I)x<0$ and $x^{\mathrm{T}}(A-I)x>0$ for all $0\neq x\in\mathbb{R}^n$.

However, we argue that condition (b) is invalid. Indeed, on the one hand, $x^{\mathrm{T}}(A^{\mathrm{T}}A-I)x<0$ for all $0\neq x\in\mathbb{R}^n$ implies that $\|A\|<1$. Hence, $\varrho(A)\le\|A\|<1$. On the other hand, $x^{\mathrm{T}}(A-I)x>0$ for all $0\neq x\in\mathbb{R}^n$ implies that $\mathrm{Re}(\lambda-1)>0$ with λ being any eigenvalue of A, from which we have $\varrho(A)>1$, a contradiction with $\varrho(A)<1$. In conclusion, if $\sigma_{\min}(A)>1$ and $A-I$ is positive definite, then there exists a $\nu\in(0,1)$ such that $\|I-\nu A\|<1-\nu$. In particular, if A is SPD, then $\sigma_{\min}(A)>1$ implies that $A-I$ is positive definite. However, in general, the positive definiteness of $A-I$ is required. As an example let $A=\mathrm{diag}((1.8,-2)^{\mathrm{T}})$. It can be verified that $\sigma_{\min}(A)>1$, $A-I$ is not positive definite and $\|I-\nu A\|>1-\nu$ for all $\nu\in(0,1)$. Obviously, the new sufficient condition, $\sigma_{\min}(A)>1$ and $A-I$ is positive definite, is stronger than the sufficient condition $\|A^{-1}\|<1$ proposed in Mangasarian & Meyer (2006). Nevertheless, the new sufficient condition guarantees the convergence of the fixed-point contraction algorithm $x^{k+1}=\mathcal{F}(x^k)$ $(k=0,1,2,\ldots)$, which is inverse-free. We will not study the property of this fixed-point iteration since it is not the main subject of this paper.

Downloaded from https://academic.oup.com/imajna/article/43/2/1036/6526317 by Zhongshan University user on 25 March 2025

The above discussion also inspires us to obtain a new proof of the sufficient condition $\|A^{-1}\| < 1$ from the point of view of the Banach fixed-point theorem, which differs from that of Mangasarian & Meyer (2006, Proposition 3). The detail of the new proof is given in Section Appendix A.

Earlier, it was mentioned that $\|A^{-1}\| < 1$ is not necessary for AVE (1.1) to be uniquely solvable. The reason may be that a positive definite matrix is usually not equivalent to a P-matrix. For example, let $A = \begin{bmatrix} 1 & 2 \\ -\frac{2}{3} & 1 \end{bmatrix}$; then $\|A^{-1}\| = 1$ and AVE (1.1) is uniquely solvable for any $b \in \mathbb{R}^n$ because $M = (A+I)(A-I)^{-1} = \begin{bmatrix} 1 & -3 \\ 1 & 1 \end{bmatrix}$ is a P-matrix. However, this M is not positive definite since $u^{\mathrm{T}}Mu = 0$ for any $u = (u_1, u_2)^{\mathrm{T}}$ with $u_1 = u_2$.

4. The methods and convergence

In this section we describe our methods and prove their convergence. We will use the following assumption in the convergence analysis.

ASSUMPTION 4.1 Assume that

$\|A^{-1}\| \leq 1$ and the solution set S^* of AVE (1.1) is nonempty.

Assumption 4.1 includes $\|A^{-1}\| < 1$ as a special case, in which AVE (1.1) has a unique solution for any $b \in \mathbb{R}^n$ (Mangasarian & Meyer, 2006). Furthermore, as shown in Section 2, it is worth mentioning that the condition $\|A^{-1}\| \leq 1$ is equivalent to saying that the corresponding GLCP (1.6) is monotone.

As stated in Section 2, solving AVE (1.1) is equivalent to finding a zero of the residual function $e(x)$, as defined in (2.6). Given an invertible matrix $G \in \mathbb{R}^{n\times n}$ and a real number $\gamma \neq 0$, $e(x) = 0$ is equivalent to

$$Q(x) + F(x) = Q(x) + F(x) - \gamma\rho(x)G^{-1}e(x)$$

whenever $\rho(x) \neq 0$. If $Q + F$ is nonsingular then we can construct the fixed-point iteration

$$(Q+F)(x^{k+1}) = (Q+F)(x^k) - \gamma\rho(x^k)G^{-1}e(x^k)$$

to solve AVE (1.1). When $\gamma \in (0, 2)$, G is SPD and

$$\rho(x) = \frac{\|e(x)\|^2}{e(x)^{\mathrm{T}}G^{-1}e(x)}, \tag{4.1}$$

the aforementioned fixed-point iteration becomes the implicit method proposed in He (1999), which is used to solve the monotone GVI problem (2.2). Essentially, it is developed from the perspective of operator splitting, and when $\gamma = 1$ and $G = I$, it is reduced to the Douglas–Rachford algorithm (Douglas & Rachford, 1956). To solve the GLCP (1.6), an equivalent reformulation of AVE (1.1), our DRs algorithm is summarized in Algorithm 4.2, where (2.11) has been used to simplify the representation.

Downloaded from https://academic.oup.com/imajna/article/43/2/1036/6526317 by Zhongshan University user on 25 March 2025

ALGORITHM 4.2 [The DRs method for AVE (1.1)] Given $x^0 \in \mathbb{R}^n$, $\gamma \in (0,2)$ and an SPD matrix $G \in \mathbb{R}^{n\times n}$, for $k = 0, 1, \ldots$, if $x^k \notin S^*$ then

$$x^{k+1} \quad \text{exactly solves} \quad \Theta_k(x) = 0, \tag{4.2}$$

where

$$\Theta_k(x) = 2Ax - 2Ax^k + \gamma\rho(x^k)G^{-1}(Ax^k - |x^k| - b) \tag{4.3}$$

and $\rho(x)$ is defined as in (4.1).

From Algorithm 4.2 each step of the DRs iteration requires the exact solution of subproblem (4.2). Indeed, it follows from (4.3) that

$$x^{k+1} = x^k - \frac{1}{2}\gamma\rho(x^k)A^{-1}G^{-1}(Ax^k - |x^k| - b). \tag{4.4}$$

Then each step of the DRs iteration requires the exact solution of the linear system with the coefficient matrix GA. In particular, let $G = I$; then $\rho(x) \equiv 1$ and (4.4) is reduced to

$$x^{k+1} = \left(1 - \frac{1}{2}\gamma\right)x^k + \frac{1}{2}\gamma A^{-1}(|x^k| + b). \tag{4.5}$$

Each step of (4.5) requires the exact solution of the system of linear equations with the coefficient matrix A. The iterative scheme (4.5) is similar to the first stage of the SOR-like iterative method (1.10); however, (4.5) is not dependent on the iterative sequence $\{y^k\}$. Comparing with the SOR-like iterative scheme, to some extent the iterative scheme (4.5) is preferred, especially for large-scale problems (see Section 5.1 for an example).

The iterative sequence $\{x^k\}$ generated by Algorithm 4.2 has the following property.

THEOREM 4.3 Assume that Assumption 4.1 holds and let $x^* \in S^*$ be an arbitrary solution of AVE (1.1). Then the sequence $\{x^k\}$ generated by Algorithm 4.2 satisfies

$$\left\|A(x^{k+1} - x^*)\right\|_G^2 \le \left\|A(x^k - x^*)\right\|_G^2 - \frac{\gamma(2-\gamma)\rho(x^k)}{4}\left\|e(x^k)\right\|^2. \tag{4.6}$$

Proof. Note that $\Theta_k(x^{k+1}) = 0$ means

$$2Ax^{k+1} = 2Ax^k - \gamma\rho(x^k)G^{-1}e(x^k).$$

In addition, the condition $\|A^{-1}\| \le 1$ implies

$$x^{\mathrm{T}}(A^{\mathrm{T}}A - I)x \ge 0 \quad \forall x \in \mathbb{R}^n. \tag{4.7}$$

Downloaded from https://academic.oup.com/imajna/article/43/2/1036/6526317 by Zhongshan University user on 25 March 2025

Hence,

$$\begin{aligned}
&\left\|2A(x^{k+1}-x^*)\right\|_G^2 \\
&\quad= \left\|2A(x^k-x^*)-\gamma\rho(x^k)G^{-1}e(x^k)\right\|_G^2 \\
&\quad= \left\|2A(x^k-x^*)\right\|_G^2 - 4\gamma\rho(x^k)(x^k-x^*)^{\mathrm{T}}A^{\mathrm{T}}e(x^k) + \left(\gamma\rho(x^k)\right)^2 e(x^k)^{\mathrm{T}}G^{-1}e(x^k) \\
&\quad\le \left\|2A(x^k-x^*)\right\|_G^2 - 2\gamma\rho(x^k)\left\|e(x^k)\right\|^2 + \left(\gamma\rho(x^k)\right)^2 e(x^k)^{\mathrm{T}}G^{-1}e(x^k) \\
&\quad= \left\|2A(x^k-x^*)\right\|_G^2 - \gamma(2-\gamma)\rho(x^k)\left\|e(x^k)\right\|^2,
\end{aligned}$$

where the inequality follows from (2.7) and (4.7), and the last equality from (4.1). □

Since G is SPD,

$$e(x)^{\mathrm{T}}G^{-1}e(x) \le \lambda_{\max}(G^{-1})\|e(x)\|^2.$$

Combining the above inequality and (4.1) we have

$$\rho(x) \ge \frac{1}{\lambda_{\max}(G^{-1})} = \lambda_{\min}(G) > 0,$$

provided that x is not a solution of AVE (1.1). As a consequence,

$$\left\|A(x^{k+1}-x^*)\right\|_G^2 \le \left\|A(x^k-x^*)\right\|_G^2 - \frac{\gamma(2-\gamma)\lambda_{\min}(G)}{4}\left\|e(x^k)\right\|^2,$$

which indicates that the sequence $\{x^k\}$ is Fejér monotone with respect to the solution set S^* of AVE (1.1), measured with the norm $\|\cdot\|_{A^{\mathrm{T}}GA}$. Hence, due to the fact that A is nonsingular and G is SPD, the global convergence of $\{x^k\}$ to a solution can be easily established, with the help of Lemma 2.3.

However, in many cases, exactly solving a system of linear equations is either expensive or impossible. On the other hand, there seems to be little justification of the effort required to calculate an accurate solution of $\Theta_k(x) = 0$ in each iteration. Hence, a new algorithm adopting approximate solutions of subproblems is desirable. Before describing the inexact algorithm we first recall the following error bound theorem for piecewise linear multifunctions, which was established in Zheng & Ng (2014, Theorem 3.3).

THEOREM 4.4 Let $\mathcal{G}$ be a piecewise linear multifunction. For any $\kappa > 0$ there exists a $\mu > 0$ such that

$$\operatorname{dist}(x, \mathcal{G}^{-1}(0)) \le \mu \operatorname{dist}(0, \mathcal{G}(x)) \ \forall \|x\| < \kappa.$$

Taking Theorem 2.1 into consideration again, solving AVE (1.1) is equivalent to finding zeros of the residual function $e(x)$, as defined in (2.6), which is piecewise linear. Hence, for any $\kappa > 0$, there is a $\mu > 0$ such that

$$\operatorname{dist}(x, S^*) \le \mu\|e(x)\| \ \forall \|x\| < \kappa. \tag{4.8}$$

Downloaded from https://academic.oup.com/imajna/article/43/2/1036/6526317 by Zhongshan University user on 25 March 2025

In the following we develop an inexact method, in which we solve $\Theta_k(x) = 0$ approximately.

ALGORITHM 4.5 [The inexact DRs method for AVE (1.1)] Given $x^0 \in \mathbb{R}^n$, $\gamma \in (0,2)$, $\delta \in (0,1)$ and an SPD matrix $G \in \mathbb{R}^{n\times n}$, for $k = 0, 1, \ldots$, if $x^k \notin S^*$ then find x^{k+1} such that

$$\left\|\Theta_k(x^{k+1})\right\| \le \alpha_k \left\|e(x^k)\right\|, \tag{4.9}$$

where $\Theta_k(x)$ is defined as in (4.3) and $\rho(x)$ is defined as in (4.1), and

$$0 \le \alpha_k \le \frac{(1-\delta)\gamma(2-\gamma)\rho(x^k)}{4\mu\|A^{\mathrm{T}}G\| + 2\gamma\rho(x^k) + \lambda_{\max}(G)} < 1. \tag{4.10}$$

Now we can prove the following contractive property between the sequence $\{x^k\}$ generated by Algorithm 4.5 and the solution set S^*.

THEOREM 4.6 Assume that Assumption 4.1 holds and let $x^* \in S^*$ be an arbitrary solution of AVE (1.1). Then the sequence $\{x^k\}$ generated by Algorithm 4.5 satisfies

$$\left\|A(x^{k+1} - x^*)\right\|_G^2 \le \left\|A(x^k - x^*)\right\|_G^2 - \frac{\delta\gamma(2-\gamma)\rho(x^k)}{4}\left\|e(x^k)\right\|^2 \ \forall\, k \ge 0. \tag{4.11}$$

Proof. From (4.3) we have

$$2A(x^{k+1} - x^k) = -\gamma\rho(x^k)G^{-1}e(x^k) + \Theta_k(x^{k+1}).$$

Hence,

$$\begin{aligned}
&\left\|2A(x^{k+1} - x^*)\right\|_G^2 \\
&\quad = \left\|2A(x^k - x^*) - \left[\gamma\rho(x^k)G^{-1}e(x^k) - \Theta_k(x^{k+1})\right]\right\|_G^2 \\
&\quad = \left\|2A(x^k - x^*)\right\|_G^2 + \left\|\gamma\rho(x^k)G^{-1}e(x^k) - \Theta_k(x^{k+1})\right\|_G^2 \\
&\qquad - 4\gamma\rho(x^k)(x^k - x^*)^{\mathrm{T}}A^{\mathrm{T}}e(x^k) + 4(x^k - x^*)^{\mathrm{T}}A^{\mathrm{T}}G\Theta_k(x^{k+1}) \\
&\quad \le \left\|2A(x^k - x^*)\right\|_G^2 - 2\gamma\rho(x^k)\left\|e(x^k)\right\|^2 \\
&\qquad + \left\|\gamma\rho(x^k)G^{-1}e(x^k) - \Theta_k(x^{k+1})\right\|_G^2 + 4(x^k - x^*)^{\mathrm{T}}A^{\mathrm{T}}G\Theta_k(x^{k+1}),
\end{aligned} \tag{4.12}$$

where the inequality follows from (2.7) and (4.7). We now deal with the last two terms in (4.12) one at a time.

Downloaded from https://academic.oup.com/imajna/article/43/2/1036/6526317 by Zhongshan University user on 25 March 2025

For the first term we have

$$\begin{aligned}
&\left\|\gamma\rho(x^k)G^{-1}e(x^k) - \Theta_k(x^{k+1})\right\|_G^2 \\
&\quad = \left\|\gamma\rho(x^k)G^{-1}e(x^k)\right\|_G^2 + \left\|\Theta_k(x^{k+1})\right\|_G^2 - 2\gamma\rho(x^k)\Theta_k(x^{k+1})^{\mathrm{T}}e(x^k) \\
&\quad \le \left\|\gamma\rho(x^k)G^{-1}e(x^k)\right\|_G^2 + \left\|\Theta_k(x^{k+1})\right\|_G^2 + 2\gamma\rho(x^k)\left\|\Theta_k(x^{k+1})\right\|\left\|e(x^k)\right\| \\
&\quad \le \gamma^2\rho(x^k)\left\|e(x^k)\right\|^2 + \lambda_{\max}(G)\alpha_k^2\left\|e(x^k)\right\|^2 + 2\gamma\rho(x^k)\alpha_k\left\|e(x^k)\right\|^2, \qquad (4.13)
\end{aligned}$$

where the first inequality follows from the Cauchy–Schwarz inequality and the last one from the definition of ρ (see (4.1)) and the approximate criterion (4.9).

We now turn to the second term. Using the Cauchy–Schwarz inequality, (4.8) and (4.9), we have

$$4(x^k - x^*)^{\mathrm{T}}A^{\mathrm{T}}G\Theta_k(x^{k+1}) \le 4\mu\alpha_k\left\|A^{\mathrm{T}}G\right\|\left\|e(x^k)\right\|^2. \qquad (4.14)$$

By substituting (4.13) and (4.14) into inequality (4.12) we obtain

$$\begin{aligned}
&\left\|2A(x^{k+1} - x^*)\right\|_G^2 \\
&\quad \le \left\|2A(x^k - x^*)\right\|_G^2 - \gamma(2-\gamma)\rho(x^k)\left\|e(x^k)\right\|^2 \\
&\qquad + \left(\lambda_{\max}(G)\alpha_k^2 + 2\gamma\rho(x^k)\alpha_k + 4\mu\alpha_k\left\|A^{\mathrm{T}}G\right\|\right)\left\|e(x^k)\right\|^2 \\
&\quad \le \left\|2A(x^k - x^*)\right\|_G^2 - \delta\gamma(2-\gamma)\rho(x^k)\left\|e(x^k)\right\|^2,
\end{aligned}$$

where the last inequality follows from (4.10). □

REMARK 4.7 Comparing the contractive properties of the sequences generated by Algorithms 4.2 and 4.5, i.e., the inequalities (4.6) and (4.11), we find that the only difference between them is the parameter δ in (4.11). Specially, if $\delta = 1$, then (4.11) is reduced to (4.6). In such a case $\alpha_k \equiv 0$, which implies that the inexact algorithm is returned to the exact one. Hence, in what follows, we only analyze the convergence behavior of Algorithm 4.5, and that for Algorithm 4.2 is similar.

We are now ready to present the convergence of Algorithm 4.5.

THEOREM 4.8 Assume that Assumption 4.1 holds. Then the sequence $\{x^k\}$ generated by Algorithm 4.5 globally converges to a solution of AVE (1.1). Furthermore, the rate of convergence is linear.

Proof. From (4.1) we can conclude that $\rho(x^k) \ge 1/\lambda_{\max}(G^{-1})$ for all k. Hence, it follows from the nonsingularity of A that $\{x^k\}$ is bounded, which implies that it does have at least one cluster point,

Downloaded from https://academic.oup.com/imajna/article/43/2/1036/6526317 by Zhongshan University user on 25 March 2025

denoted by $\hat{x}$. Moreover, (4.11) also implies

$$\lim_{k\to\infty} \|e(x^k)\| = 0,$$

which together with the continuity of the mapping $e(\cdot)$, implies that $e(\hat{x}) = 0$ and hence $\hat{x}$ is a solution of AVE (1.1). Since x^* is an arbitrary solution in (4.11) we can set $x^* := \hat{x}$ and obtain

$$\|A(x^{k+1} - \hat{x})\|_G^2 \le \|A(x^k - \hat{x})\|_G^2 - \frac{\delta\gamma(2-\gamma)}{4\lambda_{\max}(G^{-1})}\|e(x^k)\|^2. \tag{4.15}$$

It follows from (4.8) and (4.15) that

$$\|A(x^{k+1} - \hat{x})\|_G^2 \le \lambda_{\max}(A^{\mathrm{T}}GA)\|x^k - \hat{x}\|^2 \le \mu^2\lambda_{\max}(A^{\mathrm{T}}GA)\|e(x^k)\|^2.$$

By substituting the last inequality into (4.15) and then rearranging terms we get

$$\|A(x^{k+1} - \hat{x})\|_G^2 \le \zeta\|A(x^k - \hat{x})\|_G^2, \tag{4.16}$$

where

$$\zeta := \frac{4\mu^2\lambda_{\max}(G^{-1})\lambda_{\max}(A^{\mathrm{T}}GA)}{\delta\gamma(2-\gamma) + 4\mu^2\lambda_{\max}(G^{-1})\lambda_{\max}(A^{\mathrm{T}}GA)} \in (0,1).$$

Inequality (4.16) and the nonsingularity of A then imply $\{x^k\}$ converges globally and in a linear manner to a solution of AVE (1.1). □

Up to now we have proved the global linear convergence of the proposed algorithms under Assumption 4.1. We recall that $\|A^{-1}\| = 1$ does not guarantee the solvability of AVE (1.1). Hence, we borrow a trivial example from Mangasarian (2009) to illustrate the behavior of the proposed algorithms when AVE (1.1) does not have a solution. For the sake of simplicity we use the exact DRs method with $G = I$. Consider AVE (1.1) in R^1: $x - |x| - 1 = 0$. Then $\|A^{-1}\| = 1$ and AVE (1.1) has no solutions. Without loss of generality let $x^0 \ge 0$ and, according to (4.5), we have $x^{k+1} = x^k + \frac{1}{2}\gamma(|x^k| - x^k) + \frac{1}{2}\gamma = \cdots = x^0 + \frac{1}{2}\gamma(|x^0| - x^0) + \frac{k+1}{2}\gamma \to +\infty$ as $k \to +\infty$. Namely, for this example, the sequence $\{x^k\}$ generated by Algorithm 4.2 diverges to infinity, i.e., $\|x^k\| \to \infty$ as $k \to \infty$. For the general AVE (1.1), however, the following problem remains open.

Problem 4.9 Assume that $\|A^{-1}\| = 1$ and AVE (1.1) does not possess a solution. Let the sequence $\{x^k\}$ be generated by Algorithm 4.2 or 4.5. Does $\{x^k\}$ diverge to infinity?

5. Numerical results

In this section we present five numerical examples to illustrate the superior performance of our algorithms. The purpose here is to show that, on the one hand, the new algorithms have the most extensive range of convergence, i.e., besides the condition that $\|A^{-1}\| \le 1$ and the solution set is

Downloaded from https://academic.oup.com/imajna/article/43/2/1036/6526317 by Zhongshan University user on 25 March 2025

nonempty, there are no other requirements; on the other hand and more importantly, for large-scale sparse problems, the new algorithms can be much more efficient than three existing algorithms.

The following five algorithms will be tested:

1. DRs: the exact DRs method, i.e., Algorithm 4.2.
2. InexactDRs: the inexact DRs method, i.e., Algorithm 4.5. The inexact criterion (4.9) is used with $\alpha_k = \min\left\{0.9, \frac{1}{\max\{1,k-k_{\max}\}}\right\}$ and $k_{\max} = 10$. For the DRs and InexactDRs methods, we set $G = I$ and $\gamma = 1.99$.
3. Newton: the exact semismooth Newton method (Mangasarian, 2009), i.e., Algorithm 1.1.
4. InexactNewton: the inexact semismooth Newton method proposed in Cruz *et al.* (2016), i.e., Algorithm 1.2. As defined in Cruz *et al.* (2016),
$$\theta = 0.9999 \cdot \frac{1 - 3\|A^{-1}\|}{\|A^{-1}\|(\|A\| + 3)}$$
is used.
5. SOR-like: the exact SOR-like iterative method, which was proposed in Ke & Ma (2017) and further studied in Chen *et al.* (2020), i.e., Algorithm 1.3. We use the optimal iterative parameter proposed in Chen *et al.* (2020) and let $y^0 = x^0$.

Note that in all these algorithms, the main task of each iteration is solving a system of linear equations. The main difference between the Newton-type methods and the SOR-like algorithm, as well as the algorithms proposed in this paper, is the coefficient matrices: in Newton-type methods the matrix is $A - \mathcal{D}(x^k)$ (see, e.g., (1.9)), which may be varied along with the iteration because $\mathcal{D}(x^k)$ may vary with x^k, while in the SOR-like method and Algorithms 4.2 and 4.5, the matrices A and GA, respectively, are fixed. According to (4.5) and (1.10) each step of the tested DRs iteration and the SOR-like iteration requires the exact solution of the linear system with the constant coefficient matrix A. Solving systems $Ax = d^k$ for different k only requires one decomposition of A, which in the worst case costs $\mathcal{O}(n^3)$ flops; for each iteration the cost is reduced to $\mathcal{O}(n^2)$ flops (matrix-vector products). Hence, if there are $K(K \ll n)$ iteration steps, the cost is $\mathcal{O}(n^3) + K\mathcal{O}(n^2)$, which is less than $K\mathcal{O}(n^3)$ (the typical cost of Newton-type methods). For such a system of linear equations with the fixed coefficient matrix A we can adopt the intrinsic properties of A, especially when A is large scale and sparse.

In our implementations, for the sake of efficiency, $dA = \text{decomposition}(A, \text{'lu'})$ is used to precompute the LU decomposition of A for the corresponding exact methods, where decomposition is the routine in MATLAB. Furthermore, as suggested in Cruz *et al.* (2016), the LSQR algorithm (Paige & Saunders, 1982) is used as the inner iterative method for inexact methods. All methods are terminated if $\|e(x^k)\| \leq \epsilon$ with $\epsilon = 10^{-6}$. Unless otherwise stated we set $x^0 = -100 + 200 \times \text{rand}(n, 1)$ in the examples, where rand is the MATLAB function that returns a pseudorandom scalar drawn from the standard uniform distribution on the open interval $(0, 1)$. All computations are done in MATLAB R2017b with a machine precision 2.22×10^{-16} on a personal computer with a 2.60 GHz central processing unit (Intel Core i7), 16 GB memory and MacOS operating system.

For comparison, in Examples 5.2 and 5.3, the performance profile graphic (Dolan & Moré, 2002) is used, and the performance measurement is the running CPU time; that is, we use the performance ratio

$$r_{p,s} = \frac{t_{p,s}}{\min\{t_{p,s} : s \in \mathcal{S}\}},$$

Downloaded from https://academic.oup.com/imajna/article/43/2/1036/6526317 by Zhongshan University user on 25 March 2025

TABLE 1 *Numerical results for Example 5.1*

Method		n				
		24000	28000	32000	36000	40000
SOR-like	IT	13	13	13	13	13
	CPU	0.0428	0.0500	0.0576	0.0730	0.0740
	RES	1.58×10^{-7}	1.70×10^{-7}	1.83×10^{-7}	1.93×10^{-7}	2.04×10^{-7}
DRs	IT	13	13	13	13	13
	CPU	0.0417	0.0491	0.0537	0.0699	0.0693
	RES	2.55×10^{-7}	2.73×10^{-7}	2.93×10^{-7}	3.09×10^{-7}	3.27×10^{-7}

where $t_{p,s}$ is defined as the mean of the CPU time by using method s to solve problem p five times, and $\mathcal{S}$ is the set of corresponding solvers. We set $r_{p,s} = r_M$ if and only if solver s does not successfully solve problem p within 50 iterations, where $r_M = \text{ceil}(r_{\max}) + 10$ with $r_{\max} = \max\{r_{p,s} : p \in \mathcal{P}, s \in \mathcal{S}\}$ ($\mathcal{P}$ is the set of tested problems). In this way the performance profile graphic allows us to compare the efficiency and robustness of the tested methods (Cruz *et al.*, 2016). Indeed, in a performance profile graphic, efficiency and robustness can be accessed on the extreme left (when $\tau = 1$ in our plots) and right (when $\tau = r_M$ in our plots) of the graphic, respectively.

5.1 *Computational results for* $\|A^{-1}\| < \frac{1}{3}$

We first give an example to compare the numerical performance of the DRs method and the SOR-like method based on the number of iterations (denoted by 'IT'), elapsed CPU time in seconds (denoted by 'CPU') and residual error defined by $\|e(x^k)\|$ (denoted by 'RES').

EXAMPLE 5.1 [Guo *et al.*, 2019] Consider AVE (1.1) with $A = \text{tridiag}(-1, 8, -1) \in \mathbb{R}^{n\times n}$ and $b = Ax^* - |x^*|$, where $x^* = (-1, 1, -1, 1, \ldots, -1, 1)^{\mathrm{T}} \in \mathbb{R}^n$.

For this example $\|A^{-1}\| = 0.1667 < 0.25$ and the numerical results are reported in Table 1. From Table 1 we find that the total numbers of iterations for the SOR-like method and the DRs method are equal. However, the DRs method performs slightly better than the SOR-like method in terms of CPU time, which is consistent with the fact that each step of the DRs iteration is less expensive than that of the SOR-like iteration.

In the following we present the computational results for all tested methods in the case that $\|A^{-1}\| < \frac{1}{3}$ and the development here parallels the recent work by Cruz *et al.* (2016) on the InexactNewton method for solving AVE (1.1).

EXAMPLE 5.2 Consider AVE (1.1) generated by the method of Cruz *et al.* (2016). Here, $n = 10000$ is used and two sets of AVEs (1.1) are generated with the mean density of A being approximately equal to 0.003. For each set, 100 AVEs (1.1) are generated and the average condition number of A roughly equals 291.5162 (the smallest value and the largest value are 104.9930 and 999.7781, respectively) for the first set and 1.9007×10^5 (the smallest value is 1.0266×10^5, and the largest value is 8.1424×10^5) for the second set. Let $x^* = -100 + 200 \times \text{rand}(n, 1)$ and $b = Ax^* - |x^*|$.

Numerical results for this example are reported in Table 2 and Fig. 1. It follows from Table 2, for the case that the average condition number of A is relatively small ($\mathcal{O}(10^2)$), that the InexactDRs method performs best, and the probability of the InexactDRs method winning on a given problem is approximately equal to 0.62. In addition, the probabilities of the InexactNewton method, DRs method

Downloaded from https://academic.oup.com/imajna/article/43/2/1036/6526317 by Zhongshan University user on 25 March 2025

TABLE 2 *Numerical results for Example 5.2*

Method		Condition number	
		$\mathcal{O}(10^2)$	$\mathcal{O}(10^5)$
Newton	Efficiency (%)	0	0
	Robustness (%)	100	47
InexactNewton	Efficiency (%)	18	0
	Robustness (%)	100	33
DRs	Efficiency (%)	12	43
	Robustness (%)	100	75
InexactDRs	Efficiency (%)	62	0
	Robustness (%)	100	74
SOR-like	Efficiency (%)	8	32
	Robustness (%)	100	48

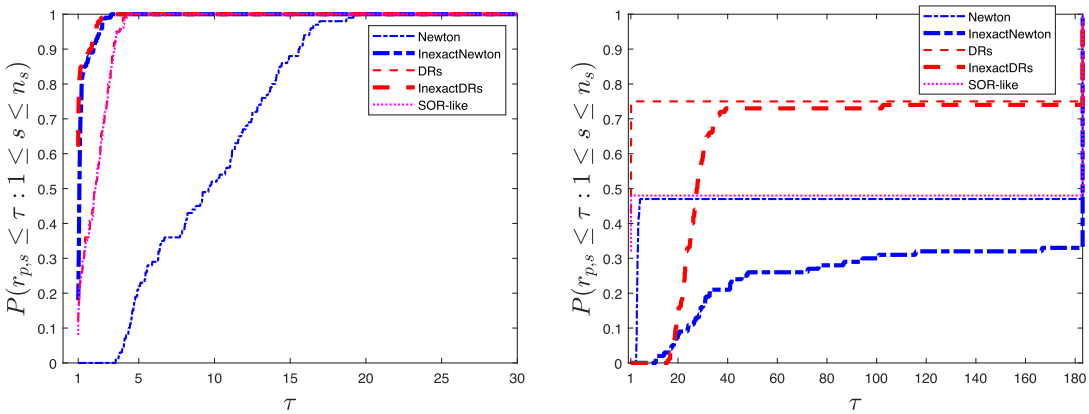

FIG. 1. Performance profile graphics for Example 5.2. Left: the average condition number of A is $\mathcal{O}(10^2)$. Right: the average condition number of A is $\mathcal{O}(10^5)$.

and SOR-like method to win are 0.18, 0.12 and 0.08, respectively. All methods have high robustness of 100%. However, for the case that the average condition number of A is relatively large ($\mathcal{O}(10^5)$), the DRs method has a probability of 0.43 to win, while the SOR-like method has a probability of 0.32 to win. Robustness rates are 75%, 74%, 48%, 47% and 33% for DRs, InexactDRs, SOR-like, Newton and InexactNewton methods, respectively. The conclusions are further supported by the performance profile graphics depicted in Fig. 1. It can be seen from the left plot of Fig. 1 that the InexactNewton method is also better than the exact methods. Both DRs and SOR-like methods are more efficient than the Newton method, and the DRs method is slightly more efficient than the SOR-like method. In addition, the performance of the InexactNewton method, the DRs method and the SOR-like method becomes much more competitive if we extend the τ of interest to 4. The numerical results also show that the inexact methods are more sensitive to the increase of the condition numbers of A than the exact methods, and the same phenomenon is reported in Cruz *et al.* (2016). Unfortunately, its theoretical basis remains to be further studied.

Downloaded from https://academic.oup.com/imajna/article/43/2/1036/6526317 by Zhongshan University user on 25 March 2025

TABLE 3 *Numerical results for Example 5.3*

Method	Newton	DRs	InexactDRs	SOR-like
Efficiency (%)	0	0	100	0
Robustness (%)	100	100	100	100

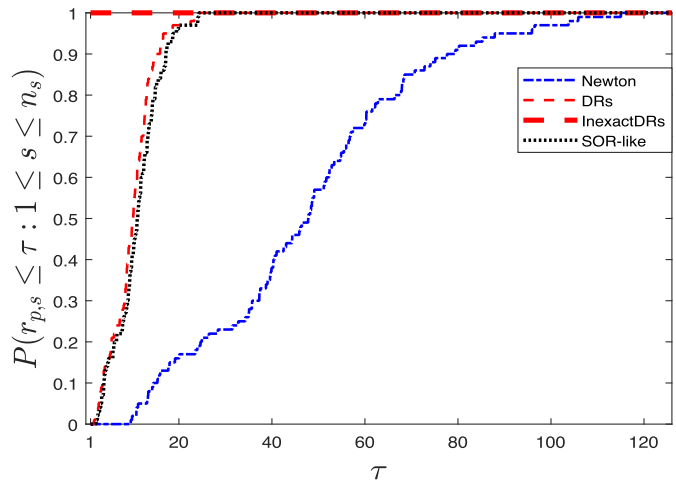

FIG. 2. The performance profile graphic for Example 5.3.

5.2 *Computational results for* $\frac{1}{3} \leq \|A^{-1}\| < 1$

In this subsection AVE (1.1) is generated in the same way as Example 5.2, except the singular value vector singval of A. Specifically, for a given singval, let mins = min(singval) and the singval is randomly updated by singval = (3/mins) ∗ singval or singval = (1/mins + (2/mins) ∗ rand)*singval such that $\frac{1}{3} \leq \|A^{-1}\| < 1$. In this case the InexactNewton method is not tested since $1 - 3\|A^{-1}\| \leq 0$ and the residual relative error tolerance θ is infeasible (see (1.7)), and the same reason applies to Section 5.3. Though there are no theoretical guarantees, the Newton method is still tested, as it often converges in this case. In this subsection we deal only with relatively 'well-conditioned' matrices for greater robustness.

EXAMPLE 5.3 Randomly generate 100 AVEs (1.1) with $n = 10000$. The mean density of A is approximately equal to 0.003 and the average condition number of A equals roughly 49.5642 (the smallest and the largest values are 11.3059 and 235.0726, respectively). Let $x^* = -100 + 200 \times \text{rand}(n, 1)$ and $b = Ax^* - |x^*|$.

Numerical results for this example are reported in Table 3 and Fig. 2. From Table 3 we see that all methods have robustness 100%. Efficiencies of the Newton method, the DRs method, the InexactDRs method and the SOR-like method are 0%, 0%, 100% and 0%, respectively. As expected the InexactDRs method is more efficient than the other methods. Analyzing Fig. 2 both DRs and SOR-like methods are more efficient than the Newton method, and the DRs method is superior to the SOR-like method in terms of CPU time. In addition, the performance of the DRs method and the SOR-like method becomes much more competitive if we extend the τ of interest to 20.

Downloaded from https://academic.oup.com/imajna/article/43/2/1036/6526317 by Zhongshan University user on 25 March 2025

TABLE 4 *Numerical results for Example 5.4*

Method	x^0	[0;0]	[1;1]	$[-1;1]$	$[1;-1]$	$[-1;-1]$
Newton	x^*	[0;0]	—	[0;0]	$[1;-1]$	—
DRs	x^*	[0;0]	$[1;-0.99]$	$[0.99;-0.99]$	$[1;-1]$	$[0.99;-1]$
InexactDRs	x^*	[0;0]	$[1;-0.99]$	$[0.99;-0.99]$	$[1;-1]$	$[0.99;-1]$
SOR-like	x^*	[0;0]	$[1;-0.2361]$	$[0.0557;-0.4427]$	$[1;-1]$	$[0.0983;-0.0132]$

5.3 *Computational results for* $\|A^{-1}\| \geq 1$

In this subsection we will present some toy examples for the case that $\|A^{-1}\| \geq 1$. Our main goal here is to further verify the theoretical results proposed in Section 4.

We first give an AVE (1.1) with $\|A^{-1}\| = 1$, which has infinitely many solutions.

EXAMPLE 5.4 [Chen *et al.*, 2021] Consider AVE (1.1) with $A = \left[\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right]$ and $b = \left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right]$; then it is reduced to

$$\begin{cases} x_1 - |x_1| = 0, \\ -x_2 - |x_2| = 0. \end{cases} \tag{5.1}$$

Note that $\|A^{-1}\| = 1$ and the AVE (5.1) has infinitely many solutions. Essentially, any $x = (x_1, x_2)^{\mathrm{T}}$ with $x_1 \geq 0$ and $x_2 \leq 0$ is a solution of AVE (5.1). Numerical results for this example are reported in Table 4. It follows from Table 4 that DRs, InexactDRs and SOR-like methods find a solution of the AVE from different initial points. In addition, the obtained solution is initial-point dependent. The Newton method fails to find a solution with initial points $x^0 = (1,1)^{\mathrm{T}}$ and $x^0 = (-1,-1)^{\mathrm{T}}$ (the results are denoted by '—'). It is easy to check that $A - \mathcal{D}(x^0)$ is singular in both cases, and the Newton iteration is not well defined.

Finally, we consider an AVE (1.1) with $\|A^{-1}\| > 1$.

EXAMPLE 5.5 [Cruz *et al.*, 2016] Consider AVE (1.1) with $A = \left[\begin{smallmatrix} 1 & -1 \\ 3 & -1 \end{smallmatrix}\right]$ and $b = \left[\begin{smallmatrix} -1 \\ -3 \end{smallmatrix}\right]$. Note that $\|A^{-1}\| \approx 1.7071 > 1$, and the matrix $A - \mathcal{D}(x)$ is invertible for any $x \in \mathbb{R}^2$. Moreover, the unique solution of the AVE (1.1) is $x^* = -(1,1)^{\mathrm{T}}$. Let $x^0 = (1,1)^{\mathrm{T}}$; as shown in Cruz *et al.* (2016), the sequence generated by the Newton method will oscillate between the points $x^1 = \left(-\frac{1}{3}, 1\right)^{\mathrm{T}}$ and $x^2 = (1,3)^{\mathrm{T}}$. However, under the same setup, all of the DRs, InexactDRs and SOR-like methods can find the solution of AVE (1.1) (the numbers of iterations are 30, 36 and 67 for the DRs, InexactDRs and SOR-like methods, respectively).

Examples 5.4 and 5.5 show that, for the general convergence of the Newton method, an additional condition on A must be assumed (Cruz *et al.*, 2016). Meanwhile, Examples 5.4 and 5.5 also show that $\|A^{-1}\| < 1$ is not necessary for the convergence of the SOR-like method. Example 5.5 implies that $\|A^{-1}\| \leq 1$ is not necessary for the convergence of the DRs method and the InexactDRs method. Up to now we are not aware whether it is possible to prove them theoretically.

Downloaded from https://academic.oup.com/imajna/article/43/2/1036/6526317 by Zhongshan University user on 25 March 2025

6. Conclusions

Regarding AVE (1.1) as a special GLCP, solving AVE (1.1) is equivalent to finding a zero of the residual function of a nonsmooth projection equation. Utilizing the special structure of the induced GLCP we design exact and inexact DRs methods for solving AVE (1.1). The most impressive feature of our methods is that we consider the case where $\|A^{-1}\| \leq 1$ rather than $\|A^{-1}\| < 1$, which is only theoretically applied by most of the existing works in the literature; some methods, such as the Newton type, can only be applied to the case with more restrictions. If the solution set of AVE (1.1) is nonempty then our exact and inexact methods converge globally to a solution, and the rate of convergence is linear. If the solution set of AVE (1.1) is empty, the sequence generated by our exact method diverges to infinity on an example.

The proposed methods also have their advantages from a numerical viewpoint:

(i) Comparing with the Newton-type methods, the coefficient matrix during the iteration is fixed while that of Newton-type methods varies. This special intrinsic property allows us to manage the most expensive implementation, the inverse computation, in a 'once-for-all' manner. Hence, the cost per iteration reduces to the matrix-vector production cost, which is much lower than the inverse computation.

(ii) Comparing with the SOR-like method, in SOR-like methods, an auxiliary variable y is introduced; hence, the dimension of the system is enlarged to twice that of AVE (1.1). Although the updating scheme of variable x is basically consistent with the methods proposed in this paper, the updating of auxiliary variable y cannot be ignored, especially if A has some structure that makes the matrix-vector product for updating x low cost.

Numerical results indicate that our methods are competitive.

A by-product during our analysis is a new sufficient condition for the unique solvability of AVE (1.1). Although the new condition is stronger than the existing $\|A^{-1}\| < 1$, it inspires a new proof of $\|A^{-1}\| < 1$. Finally, in this paper we set $G = I$. Better choices for G may improve the performance of the proposed methods and how to choose G deserves further investigation.

Acknowledgements

The authors are grateful to the three anonymous reviewers for their helpful comments and suggestions, and in particular for their pointing out an error in our earlier citation of the Brouwer fixed-point theorem, leading us to delete the corresponding results. The authors also thank the editors for their comments, which have helped to improve the paper.

Funding

National Natural Science Foundation of China (11901024 to C.C., 12131004 and 11625105 to D.H.); Natural Science Foundation of Fujian Province (2021J01661 to C.C.); Ministry of Education in China of Humanities and Social Science Project (21YJCZH204 to D.Y.); Natural Science Foundation of Liaoning Province (2022lslwtkt-069, 2020-MS-301 to D.Y.).

References

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

Downloaded from https://academic.oup.com/imajna/article/43/2/1036/6526317 by Zhongshan University user on 25 March 2025

Bauschke, H. H. & Combettes, P. L. (2011) *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. New York: Springer.

Bixby, R., Fenelon, M., Gu, Z., Rothberg, E. & Wunderling, R. (2004) Mixed-integer programming: a progress report. *The Sharpest Cut: The Impact of Manfred Padberg and His Work*. MPS-SIAM Series on Optimization. Philadelphia: SIAM, pp. 309–325.

Bot, R. I., Csetnek, E. R. & Nguyen, D. K. (2019) A proximal minimization algorithm for structured nonconvex and nonsmooth problems. *SIAM J. Optim.*, **29**, 1300–1328.

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

Chen, C.-R., Yang, Y.-N., Yu, D.-M. & Han, D.-R. (2021) An inverse-free dynamical system for solving the absolute value equations. *Appl. Numer. Math.*, **168**, 170–181.

Chen, C.-R., Yu, D.-M. & Han, D.-R. (2020) Optimal parameter for the SOR-like iteration method for solving the system of absolute value equations. arXiv:2001.05781.

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

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

Dolan, E. D. & Moré, J. J. (2002) Benchmarking optimization software with performance profiles. *Math. Program.*, **91**, 201–213.

Douglas, J. & Rachford, H. H. (1956) On the numerical solution of heat conduction problems in two and three space variables. *Trans. Am. Math. Soc.*, **82**, 421–439.

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

Facchinei, F. & Pang, J.-S. (2003) *Finite-Dimensional Variational Inequalities and Complementarity Problems* vol. I.. New York: Springer.

Gao, X.-B. & Wang, J. (2014) Analysis and application of a one-layer neural network for solving horizontal linear complementarity problems. *Int. J. Comput. Int. Sys.*, **7**, 724–732.

Gu, X.-M., Huang, T.-Z., Li, H.-B., Wang, S.-F. & Li, L. (2017) Two CSCS-based iteration methods for solving absolute value equations. *J. Appl. Anal. Comput.*, **7**, 1336–1356.

Guo, P., Wu, S.-L. & Li, C.-X. (2019) On the SOR-like iteration method for solving absolute value equations. *Appl. Math. Lett.*, **97**, 107–113.

Harker, P. T. & Pang, J.-S. (1990) Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications. *Math. Program.*, **48**, 161–220.

He, B.-S. (1999) Inexact implicit methods for monotone general variational inequalities. *Math. Program.*, **86**, 199–217.

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

Iqbal, J., Iqbal, A. & Arif, M. (2015) Levenberg–Marquardt method for solving systems of absolute value equations. *J. Comput. Appl. Math.*, **282**, 134–138.

Ke, Y.-F. (2020) The new iteration algorithm for absolute value equation. *Appl. Math. Lett.*, **99**, 105990.

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

Kojima, M. (1989) A polynomial-time algorithm for a class of linear complementarity problems. *Math. Program.*, **44**, 1–26.

Lian, Y.-Y., Li, C.-X. & Wu, S.-L. (2018) Weaker convergent results of the generalized Newton method for the generalized absolute value equations. *J. Comput. Appl. Math.*, **338**, 221–226.

Mangasarian, O. L. (2007a) Absolute value programming. *Comput. Optim. Appl.*, **36**, 43–53.

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

Mangasarian, O. L. (2009) A generalized Newton method for absolute value equations. *Optim. Lett.*, **3**, 101–108.

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

Downloaded from https://academic.oup.com/imajna/article/43/2/1036/6526317 by Zhongshan University user on 25 March 2025

Mansoori, A. & Erfanian, M. (2018) A dynamic model to solve the absolute value equations. *J. Comput. Appl. Math.*, **333**, 28–35.
Mansoori, A., Eshaghnezhad, M. & Effati, S. (2017) An efficient neural network model for solving the absolute value equations. *IEEE T. Circuits-II*, **65**, 391–395.
Mezzadri, F. (2020) On the solution of general absolute value equations. *Appl. Math. Lett.*, **107**, 106462.
Miao, X.-H., Yang, J.-T., Saheya, B. & Chen, J.-S. (2017) A smoothing Newton method for absolute value equation associated with second-order cone. *Appl. Numer. Math.*, **120**, 82–96.
Paige, C. C. & Saunders, M. A. (1982) LSQR: an algorithm for sparse linear equations and sparse least squares. *ACM Trans. Math. Softw. (TOMS)*, **8**, 43–71.
Prokopyev, O. (2009) On equivalent reformulations for absolute value equations. *Comput. Optim. Appl.*, **44**, 363–372.
Rohn, J. (1989) Systems of linear interval equations. *Linear Algebra Appl.*, **126**, 39–78.
Rohn, J. (2004) A theorem of the alternatives for the equation $Ax + B\,|x| = b$. *Linear Multilinear Algebra*, **52**, 421–426.
Rohn, J., Hooshyarbakhsh, V. & Farhadsefat, R. (2014) An iterative method for solving absolute value equations and sufficient conditions for unique solvability. *Optim. Lett.*, **8**, 35–44.
Saad, Y. (2003) *Iterative Methods for Sparse Linear Systems*. Philadelphia: SIAM.
Saheya, B., Yu, C.-H. & Chen, J.-S. (2018) Numerical comparisons based on four smoothing functions for absolute value equation. *J. Appl. Math. Comput.*, **56**, 131–149.
Wang, H.-J., Cao, D.-X., Liu, H. & Qiu, L. (2017) Numerical validation for systems of absolute value equations. *Calcolo*, **54**, 669–683.
Wu, S.-L. & Li, C.-X. (2018) The unique solution of the absolute value equations. *Appl. Math. Lett.*, **76**, 195–200.
Wu, S.-L. & Shen, S.-Q. (2021) On the unique solution of the generalized absolute value equation. *Optim. Lett.*, **15**, 2017–2024.
Xia, M.-X. & Li, H.-H. (2020) New Farkas-type inequalities of mixed interval systems for AE solvability. *J. Phys. Conf. Ser.*, **1592**, 012082.
Zamani, M. & Hladík, M. (2021) A new concave minimization algorithm for the absolute value equation solution. *Optim. Lett.*, **15**, 2241–2254.
Zhang, C. & Wei, Q.-J. (2009) Global and finite convergence of a generalized Newton method for absolute value equations. *J. Optim. Theory Appl.*, **143**, 391–403.
Zheng, X.-Y. & Ng, K. F. (2014) Metric subregularity of piecewise linear multifunctions and applications to piecewise linear multiobjective optimization. *SIAM J. Optim.*, **24**, 154–174.

Appendix A. A new proof of the sufficient condition $\|A^{-1}\| < 1$

In this section we will give a new proof of the sufficient condition $\|A^{-1}\| < 1$, which is different from that of Mangasarian & Meyer (2006, Proposition 3). The main result in this section is the following lemma.

Lemma A.1 AVE (1.1) is uniquely solvable for any $b \in \mathbb{R}^n$ if $\|A^{-1}\| < 1$.

Proof. According to Theorem 2.1, x^* is a solution of AVE (1.1) if and only if $\tilde{\mathcal{F}}(x^*) = x^*$ with

$$\tilde{\mathcal{F}}(x) = x - \nu A^{-1} e(x), \quad \nu \in (0, 1). \tag{A.1}$$

Downloaded from https://academic.oup.com/imajna/article/43/2/1036/6526317 by Zhongshan University user on 25 March 2025

It follows from (A.1) that

$$\begin{aligned}\left\|\tilde{\mathcal{F}}(x)-\tilde{\mathcal{F}}(y)\right\| &= \|(1-\nu)(x-y)+\nu A^{-1}(|x|-|y|)\| \\ &\leq (1-\nu)\cdot\|x-y\|+\nu\|A^{-1}\|\|x-y\| \\ &= (1-\nu+\nu\|A^{-1}\|)\|x-y\|,\end{aligned}$$

which implies that $\tilde{\mathcal{F}}$ is contractive if $\|A^{-1}\| < 1$. The proof is completed by using the Banach fixed-point theorem. □

As a by-product the fixed-point iteration $x^{k+1}=\tilde{\mathcal{F}}(x^k)$ $(k=0,1,2,\ldots)$ is derived from (A.1), which is convergent if $\|A^{-1}\| < 1$. In addition, this fixed-point iteration is reduced to the DRs scheme (4.5) when $\nu=\frac{\gamma}{2}$.

Downloaded from https://academic.oup.com/imajna/article/43/2/1036/6526317 by Zhongshan University user on 25 March 2025