Solving Absolute Value Equation using Complementarity and Smoothing Functions

Tangi Migot, Lina Abdallah, Mounir Haddou

Solving Absolute Value Equation using Complementarity and Smoothing Functions

T. Migot[*] L. Abdallah[†] M. Haddou[‡]

Abstract

In this paper, we reformulate the NP-hard problem of the absolute value equation (AVE) as a horizontal linear complementarity one and then solve it using a smoothing technique. This approach leads to a new class of methods that are valid for general absolute value equation. An asymptotic analysis proves the convergence of our schemes and provides some interesting error estimates. This kind of error bound or estimate had never been studied for other known methods. The corresponding algorithms were tested on randomly generated problems and applications. These experiments show that, in the general case, one observes a reduction of the number of failures.

Keywords : smoothing function ; concave minimization ; complementarity ; absolute value equation
AMS Subject Classification : 90C59 ; 90C30 ; 90C33 ; 65K05 ; 49M20

1 Introduction

We consider the absolute value equation, which consists in finding $x \in \mathbb{R}^N$ such that

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

where $A \in \mathbb{R}^{N\times N}$, $b \in \mathbb{R}^N$ and $|x| := (|x_1|, \dots, |x_N|)^T$. A slightly more general problem has been introduced in [23]

$$Ax + B|x| = b \ , \tag{AVP}$$

where $A, B \in \mathbb{R}^{M\times N}$, $b \in \mathbb{R}^M$ and unknown $x \in \mathbb{R}^N$. Here, we focus here on (AVE), which has been more popular in the literature. The recent interest in these problems can be explained by the fact that frequently occurring optimization problems such as linear complementarity problems and mixed integer

[*]tangi.migot@insa-rennes.fr,IRMAR-INSA, Rennes, France
[†]lina_abdallah@hotmail.fr, Université Libanaise, Laboratoire de Mathématiques et Applications (LaMA), Tripoli, Liban
[‡]mounir.haddou@insa-rennes.fr,IRMAR-INSA, Rennes, France

programming problems can be reformulated as an (AVE), see [21, 16]. The general NP-hard linear complementarity problem can be formulated as an (AVE), which implies that it is NP-hard in general. Moreover, it has been proved in [21] that checking if an (AVE) has one or an infinite number of solutions is NP-complete.

Theoretical criteria regarding existence of solutions and unique solvability of (AVE) have been studied in [21, 17, 27, 25, 9]. An important criterion among others is that (AVE) has a unique solution if all of the singular values of the matrix A exceed 1. In the special case where the problem is uniquely solvable, a family of Newton methods has been proposed first in [15], then completed with global and quadratic convergence in [2], an inexact version in [1] and other related methods [5, 20, 30]. Also, Picard-HSS iteration methods and nonlinear HSS-like methods have been considered for instance in [28, 22, 31]. It is of a great interest to consider methods that remain valid in the general case. Most of such methods which are valid in the general case are due to Mangasarian in [14, 11, 12] by considering a concave or a bilinear reformulation of (AVE) solved by a sequence of linear programs. A hybrid method mixing Newton approach of [15] and [12] can be found in [13]. A method based on interval matrix has been studied by Rohn in [24, 26].

The special case where (AVE) is not solvable also received some interests in the literature. Prokopyev shows numerical results using a mixed integer programming solver in [21]. Theoretical study in order to correct b and A to make (AVE) feasible can be found in [7, 8].

Our aim in this paper is to pursue the study of (AVE) without additional hypothesis such as unique solvability and propose a new method, which solves a sequence of linear programs. The motivation is to diminish the number of instances where classical methods can not solve the problem. We propose a new reformulation of (AVE) as a sequence of concave minimization problems using complementarity and a smoothing technique.

The paper is organized as follows. In Section 2, we present the new formulation of (AVE) as a sequence of concave minimization problems. In Section 3, we prove convergence to a solution of (AVE) and in Section 4, we establish error estimate. Finally, Section 5 provides numerical results on simple examples and randomly generated test problems.

2 AVE as a Sequence of Concave Minimization Programs

We consider a reformulation of (AVE) as a sequence of concave minimization problems. First, we use a classical decomposition of the absolute value to reformulate (AVE) as an horizontal linear complementarity problem. Set $x = x^+ - x^-$, where $x^+ \geq 0$, $x^- \geq 0$ and $x^+ \perp x^-$, so that $x^+ = \max(x, 0)$

and $x^- = \max(-x, 0)$. This decomposition guarantees that $|x| = x^+ + x^-$. So (AVE) is equivalent to the following complementarity problem

$$\begin{cases} A(x^+ - x^-) - (x^+ + x^-) = b \\ x^+ \geq 0,\ x^- \geq 0 \\ x^+ \perp x^- \end{cases} . \tag{1}$$

Now, we reformulate this problem as a sequence of concave optimization problems using a smoothing technique. This technique has been first studied in [19, 4] and uses a family of non-decreasing continuous smooth concave functions $\theta : \mathbb{R} \to]-\infty, 1[$ that satisfies

$$\theta(t) < 0 \text{ if } t < 0,\ \theta(0) = 0 \text{ and } \lim_{t \to +\infty} \theta(t) = 1 \ .$$

One generic way to build such functions is to consider non-increasing probability density functions $f : \mathbb{R}_+ \to \mathbb{R}_+$ and then take the corresponding cumulative distribution functions

$$\forall t \geq 0, \qquad \theta(t) = \int_0^t f(x)dx \ .$$

By definition of f

$$\lim_{t \to +\infty} \theta(t) = \int_0^{+\infty} f(x)dx = 1 \text{ and } \theta(0) = \int_0^0 f(x)dx = 0 \ .$$

The hypothesis on f gives the concavity of θ on $\mathbb{R}_+$. We then extend the functions θ for negative values in a differentiable way, for instance taking $\theta(t < 0) = tf(0)$.

Two interesting examples of this family are $\theta^1(t) = t/(t+1)$ if $t \geq 0$ and $\theta^1(t) = t$ if $t < 0$, $\theta^2(t) = 1 - e^{-t}$ with $t \in \mathbb{R}$. In particular, θ^1 will play a special role in our analysis.

We introduce $\theta_r(t) := \theta\left(\frac{t}{r}\right)$ for $r > 0$. This definition is similar to the perspective functions in convex analysis. These functions satisfy

$$\theta_r(0) = 0\ \forall r > 0 \text{ and } \forall t > 0,\ \lim_{r \searrow 0} \theta_r(t) = 1 \ .$$

Previous examples lead to $\theta^1_r(t) = t/(t+r)$ if $t \geq 0$ and $\theta^1_r(t) = t/r$ if $t < 0$, $\theta^2_r(t) = 1 - e^{-t/r}$ with $t \in \mathbb{R}$.

The following lemma shows the link between this family of functions and the complementarity.

Lemma 2.1. *Given $s, t \in \mathbb{R}_+$, we have*

$$s \perp t \iff \lim_{r \searrow 0} \theta_r(s) + \theta_r(t) \leq 1 \ .$$

Proof. Prove by contradiction that

$$\lim_{r \searrow 0} \theta_r(s) + \theta_r(t) \leq 1 \Longrightarrow s \perp t \ .$$

Suppose $s, t > 0$, then

$$\lim_{r \searrow 0} (\theta_r(s) + \theta_r(t)) = \lim_{r \searrow 0} \theta_r(s) + \lim_{r \searrow 0} \theta_r(t) = 2 \ .$$

This leads to a contradiction and therefore $s \perp t$. Conversely it is clear that $s \perp t$ implies $s = 0$ or $t = 0$ and the result follows. $\square$

In the case of the function θ_r^1, it holds that

$$\theta_r^1(s) + \theta_r^1(t) = 1 \Longleftrightarrow st = r^2 \ .$$

Using the previous lemma, problem (1) can be replaced by a sequence of concave optimization problems for $r > 0$:

$$\begin{aligned} &\min_{x^+, x^- \in \mathbb{R}^N} \sum_{i=1}^N \theta_r(x_i^+) + \theta_r(x_i^-) - 1 \\ &A(x^+ - x^-) - (x^+ + x^-) = b \\ &x^+ \geq 0, \ x^- \geq 0 \end{aligned} \quad . \tag{2}$$

In order to avoid compensation phenomenon and generate strictly feasible iterates we consider a relaxed version of (2):

$$\begin{cases} \displaystyle\min_{x^+, x^- \in \mathbb{R}^N} \sum_{i=1}^N \theta_r(x_i^+) + \theta_r(x_i^-) - 1 \\ -g(r)|A|e - g(r)e \leq A(x^+ - x^-) - (x^+ + x^-) - b \leq g(r)|A|e + g(r)e, \\ x^+ + x^- \geq g(r)e \\ 0 \leq x^+ \leq M, \ 0 \leq \ x^- \leq M \end{cases} \tag{P_r}$$

where e is the unit vector, $|A|$ denotes the matrix where each element is the absolute value of the corresponding element in A, M is some positive constant to be specified later and $g : \mathbb{R}_+^* \to \mathbb{R}_+^*$ is a function satisfuing

$$\lim_{r \searrow 0} \frac{r}{g(r)} = 0 \text{ and } \lim_{r \searrow 0} g(r) = 0 \ .$$

For instance, we can choose $g(r) = r^\alpha$ with $0 < \alpha < 1$.

3 Convergence

From now on, we assume that the set of solutions of (AVE) denoted $S^*_{(AVE)}$ is non-empty and denote $S^*_{(P_r)}$ the optimal set of (P_r). In order to simplify the notation, we denote $x \in S^*_{(P_r)}$ when $(x^+, x^-) \in S^*_{(P_r)}$ with $x = x^+ - x^-$ and $x^+ = \max(x, 0)$, $x^- = \max(-x, 0)$. Let M be a positive constant such that

$$M \geq ||x^*||_\infty \ ,$$

where x^* is some solution of $S^*_{(AVE)}$. The following theorem shows that for $r > 0$, the set of solutions $S^*_{(P_r)}$ is non-empty.

Theorem 3.1. *(P_r) has at least one solution for any $r > 0$.*

Proof. Since $S^*_{(AVE)} \neq \emptyset$, there exists a point $\bar{x} \in S^*_{(AVE)}$. We can write $\bar{x} = \bar{x}^+ - \bar{x}^-$ with $\bar{x}^+ \perp \bar{x}^-$. It follows that $(y^{r+} := \bar{x}^+ + g(r), y^{r-} := \bar{x}^-)$ is a feasible point of (P_r). Furthermore, we minimize a continuous function over a non-empty compact set so the objective function attains its minimum. □

The following three lemmas will be used to prove the main convergence Theorem 3.5.

Lemma 3.2. *For $r > 0$, functions θ_r and g defined above and $x^+, x^- \in \mathbb{R}^N_+$, such that $x^+ + x^- \geq g(r)e$. It holds that*

$$\forall i \in \{1, ..., N\}, \quad \theta_r(x_i^+) + \theta_r(x_i^-) - 1 \geq \theta_r(g(r)) - 1 \ .$$

Proof. θ_r is concave and $\theta_r(0) = 0$ so θ_r is subadditive on $\mathbb{R}_+$. Thus, for all $i \in \{1, ..., N\}$ it follows that

$$\theta_r(x_i^+) + \theta_r(x_i^-) - 1 \geq \theta_r(x_i^+ + x_i^-) - 1 \ .$$

Furthermore θ_r is non-decreasing and $x^+ + x^- \geq g(r)e$ therefore

$$\theta_r(x_i^+) + \theta_r(x_i^-) - 1 \geq \theta_r(g(r)) - 1 \ .$$

□

Lemma 3.3. *Given functions θ_r and g defined above we have*

$$\lim_{r \searrow 0} \ \theta_r(g(r)) - 1 = 0 \ .$$

Proof. Since $\theta_r(g(r)) = \theta_{\frac{r}{g(r)}}(1)$ and $\lim_{r \searrow 0} r/g(r) = 0$, it follows that

$$\lim_{r \searrow 0} \ \theta_r(g(r)) = \lim_{r \searrow 0} \ \theta_{\frac{r}{g(r)}}(1) = 1 \ . \tag{3}$$

□

In the special case, where many solutions of (AVE) have at least a zero component, it can be difficult to find a feasible point (x^+, x^-) of (P_r) such that $x^+ + x^- \geq g(r)e$. The following lemma explains how to build such point in this case.

Lemma 3.4. *Let $\bar{x} \in S^*_{(AVE)}$ and $r > 0$ be such that $g(r) < \min_{\bar{x}_i \neq 0} |\bar{x}_i|$. Then $y^r := \bar{x} + g(r)$ is a solution of the following equation*

$$Ax - |x| = b + g(r)Ae - g(r)\delta(\bar{x}) \ , \qquad ((\text{AVE})_r)$$

where $\delta(x) \in \mathbb{R}^N$ is such that $\delta_i(x) := \begin{cases} 1 \ if \ x_i \geq 0 \\ -1 \ if \ x_i < 0 \end{cases}$.

Proof. $\bar{x}$ is a solution of (AVE), that is

$$A\bar{x} - |\bar{x}| = b \ .$$

Therefore it holds that

$$A\bar{x} + g(r)Ae - |\bar{x}| - g(r)\delta(\bar{x}) = b + g(r)Ae - g(r)\delta(\bar{x}),$$

and so

$$A(\bar{x} + g(r)e) - |\bar{x} + g(r)e| = b + g(r)Ae - g(r)\delta(\bar{x}).$$

Thus, $y^r = \bar{x} + g(r)$ is a solution of $((\text{AVE})_r)$. □

We now prove the convergence of $\{x^r\}_{r>0}$, where $x^r := x^{r+} - x^{r-}$ with $(x^{r+}, x^{r-}) \in S^*_{(P_r)}$, to an element of $S^*_{(AVE)}$. It is to be noted that $S^*_{(P_r)}$ is not necessarily a singleton.

Theorem 3.5. *Every limit point of the sequence $\{x^r\}$ such that $x^r \in S_{(P_r)}$ for $r > 0$ is a solution of* (AVE).

Proof. By Theorem 3.1 there exists at least one solution of (P_r). According to Lemma 3.4, we can build a sequence $\{y^r\}_{r>0}$ where $y^r = y^{r+} - y^{r-}$ with $y^{r+} \perp y^{r-}$ that are solutions of $(\text{AVE})_r$. Furthermore, for r sufficiently small (y^{r+}, y^{r-}) is a feasible point of (P_r). Let $x^r = (x^{r+}, x^{r-})$ with $\{x^r\}_{r>0}$ be a sequence of optimal solutions of (P_r), then

$$\sum_{i=1}^{N} (\theta_r(x_i^{r+}) + \theta_r(x_i^{r-}) - 1) \leq \sum_{i=1}^{N} (\theta_r(y_i^{r+}) + \theta_r(y_i^{r-}) - 1) \leq 0 \ .$$

For all $i \in \{1, ..., N\}$, it holds that

$$\theta_r(x_i^{r+}) + \theta_r(x_i^{r+}) - 1 \leq - \sum_{j=1; j \neq i}^{N} (\theta_r(x_j^{r+}) + \theta_r(x_j^{r+}) - 1) \ .$$

By Lemma 3.2, we obtain for all $i \in \{1, ..., N\}$, that

$$\theta_r(x_i^{r+}) + \theta_r(x_i^{r-}) \leq 1 + (N-1)(1 - \theta_r(g(r))) \ .$$

For any limit point $\bar{x} = (\bar{x}^+, \bar{x}^-)$ of the sequence $\{x^r\}_r$, where $\bar{x}^+ = \lim_{r \searrow 0} x^{r+}$ and $\bar{x}^- = \lim_{r \searrow 0} x^{r-}$, using Lemma (3.3) ($\lim_{r \searrow 0} 1 - \theta_r(g(r)) = 0$) and passing to the limit, it follows that

$$\lim_{r \searrow 0} \ \theta_r(\bar{x}_i^+) + \theta_r(\bar{x}_i^-) \leq 1 \ .$$

Thus, $\bar{x}^+ \perp \bar{x}^-$ by the previous inequality and Lemma 2.1.

Now, we verify that $\bar{x}$ is a solution of (AVE). Let $r > 0$ and x^r be a solution of (P_r), we have

$$b - g(r)|A|e - g(r)e \leq A(x^{r+} - x^{r-}) - (x^{r+} + x^{r-}) \leq b + g(r)|A|e + g(r)e.$$

Passing to the limit when $r \searrow 0$, we obtain

$$A(\bar{x}^+ - \bar{x}^-) - (\bar{x}^+ + \bar{x}^-) = b.$$

So, $\bar{x} = \bar{x}^+ - \bar{x}^-$ is a solution of (AVE). □

We now continue the discussion on convergence by a characterization of the limit points in the case where $S^*_{(AVE)} = \emptyset$.

Theorem 3.6. *Assume that $S^*_{(AVE)}$ is empty and that* (2) *admits a feasible point whose infinite norm is bounded by M. Then, every limit point of $\{x^r\}$ such that $x^r \in S_{(P_r)}$ for $r > 0$, does not satisfy the complementarity constraint.*

Proof. A straightforward adaptation of Theorem 3.1 to the new assumption proves that there exists a sequence $\{x^r\}$ such that $x^r \in S_{(P_r)}$ for $r > 0$. Therefore, we can extract up to a subsequence a limit point of this sequence since by assumption it is bounded.

By contradiction, if the limit point satisfies the complementarity constraint. Then, this point will be a solution of (AVE). However, this is a contradiction with the assumption that $S^*_{(AVE)}$ is empty. □

In the case where no feasible point of (2) exists. It clearly holds that for r sufficiently small the problem (P_r) becomes infeasible.

4 Error Estimate

In this section we study, the asymptotic behaviour of the sequence $\{x^r\}_{r>0}$ for small values of r. We remind the definition of the Landau notation O often used in the context of asymptotic comparison. Given two functions f and h, one writes

$$f(x) = O_{x \to a}(h(x))$$

if $\exists C > 0,\ \exists d > 0$ such that $|f(x)| \leq C|h(x)|$ when $|x - a| \leq d$. This notation becomes $O(h(x))$ when a is 0.

The following simple lemma will be useful for the rest of this section.

Lemma 4.1. *Let θ_r be such that $\theta_r \geq \theta_r^1$. For $x^r \in S^*_{(P_r)}$ and r sufficiently small, we have*

$$\forall i \in \{1, ..., N\},\quad x_i^{r+} x_i^{r-} = O(rg(r))\ .$$

Proof. Set $i \in \{1, ..., N\}$. Thanks to the convergence proof of the Theorem 3.5 for r sufficiently small it holds that

$$\theta_r(x_i^{r+}) + \theta_r(x_i^{r-}) - 1 \leq (N-1)(1 - \theta_r(g(r))) \ .$$

Thus, it also holds, in particular for $\theta = \theta^1$, that

$$\theta_r^1(x_i^{r+}) + \theta_r^1(x_i^{r-}) - 1 \leq (N-1)(1 - \theta_r^1(g(r))) \ .$$

Therefore for r sufficiently small (i.e. such that $g(r) \geq r$ and $(1-(N-1)g(r)) > 0$) it holds that

$$\begin{aligned}
&\frac{x_i^{r+}}{x_i^{r+} + r} + \frac{x_i^{r-}}{x_i^{r-} + r} - 1 \leq (N-1)(1 - \frac{g(r)}{g(r) + r}) \leq (N-1)g(r), \\
&\frac{2x_i^{r+}x_i^{r-} + rx_i^{r+} + rx_i^{r-} - x_i^{r+}x_i^{r-} - rx_i^{r+} - rx_i^{r-} - r^2}{(x_i^{r+} + r)(x_i^{r-} + r)} \leq (N-1)g(r), \\
&\frac{x_i^{r+}x_i^{r-} - r^2}{(x_i^{r+} + r)(x_i^{r-} + r)} \leq (N-1)g(r), \\
&x_i^{r+}x_i^{r-} - r^2 \leq (N-1)g(r)(x_i^{r+}x_i^{r-} + rx_i^{r+} + rx_i^{r-} + r^2), \\
&x_i^{r+}x_i^{r-} \leq r^2 g(r)\frac{1 + (N-1)}{1 - (N-1)g(r)} + rg(r)\frac{(N-1)(x_i^{r+} + x_i^{r-})}{1 - (N-1)g(r)},
\end{aligned}$$

and the results follows. □

The following proposition gives a first error estimate that concerns only the components of (x^{r+}, x^{r-}) converging to zero.

Proposition 4.2. *Let θ_r be such that $\theta_r \geq \theta_r^1$. Let $(\bar{x}^+, \bar{x}^-)$ be a limit point of the sequence $\{x^{r+}, x^{r-}\}_r$ of optimal solutions of (P_r). The convergence of the components of the variable x^{r+} or x^{r-} to the possibly zero part of the accumulation point is done in $O(r)$.*

Proof. Set $i \in \{1, ..., N\}$. We work with one component. Assume that $\bar{x}_i^+ = 0$. The opposite case is completely similar. By assumption on x_i^{r+}, for r sufficiently small it holds that $\max(x_i^{r+}, x_i^{r-}) = x_i^{r-}$. Thus, for r sufficiently small we have

$$g(r) \leq x_i^{r+} + x_i^{r-} \leq 2x_i^{r-} \ .$$

Additionally, by Lemma 4.1 it follows

$$\begin{aligned}
&x_i^{r+}x_i^{r-} = O(rg(r)), \\
&\Longrightarrow x_i^{r+} = \frac{O(rg(r))}{x_i^{r-}}, \\
&\Longrightarrow x_i^{r+} = \frac{O(rg(r))}{g(r)}, \\
&\Longrightarrow |x_i^{r+} - \bar{x}_i^+| = O(r).
\end{aligned}$$

□

In the next theorem we provide an error estimate for the non-zero part of the solution in the couple (x^+, x^-).

To establish this result, we use the classical Hoffman's lemma on linear inequalities.

Lemma 4.3 (Hoffman's lemma,[6]). *Given a convex polyhedron P such that*

$$P = \{x \in \mathbb{R}^n \mid Ax \leq b\} \ .$$

We set $d_P(x)$ the distance from x to P, by choosing a norm $||.||$, where $d_P(x) = \inf_{y \in P} \ ||y - x||$. There exists a constant K that only depends on A, such that

$$\forall x \in \mathbb{R}^n \ : \ d_P(x) \leq K||(Ax - b)^+|| \ .$$

It is to be noted that if the constraints are given by $Ax = b$ with A a square full-rank matrix instead of $Ax \leq b$ then the polyhedron is reduced to a singleton and one can choose $K = ||A^{-1}||$.

Theorem 4.4. *Given $(\bar{x}^+, \bar{x}^-)$ a limit point of the sequence $\{x^{r+}, x^{r-}\}$, if we denote $\bar{x} = \bar{x}^+ - \bar{x}^-$ and $x^r = x^{r+} - x^{r-}$. Then, for r sufficiently small*

$$d_{S^*_{(AVE)}}(x^r) = O(g(r)) \ . \tag{4}$$

We remind here that the function g can be chosen such that $g(r) = r^\alpha$ with $\alpha \in (0, 1)$. This means that (4) is almost a linear bound.

Proof. We split the proof in two cases, either $\min_{i \in \{1,...,N\}} |\bar{x}_i| \neq 0$, either $\exists i \in \{1, ..., N\}, \ \bar{x}_i = 0$ respectively denoted as a) and b).

a) First, assume that there is no zero component in $\bar{x}$. Let V be a neighbourhood of $\bar{x}$ defined as

$$V = B_\infty(\bar{x}, \alpha) = \{x \mid \max_{1 \leq i \leq N} \ |x_i - \bar{x}_i| \leq \alpha\} \ ,$$

where $\alpha = \min_{i \in \{1,...,N\}} |\bar{x}_i|/2$. For all $x \in V$, $\bar{x}$ and x have the same signs component-wise. Denote $D = \ diag(\delta(\bar{x}))$, where $\delta(x) \in \mathbb{R}^N$ with $\delta_i(x) = \begin{cases} 1 \ if \ x_i \geq 0 \\ -1 \ if \ x_i < 0 \end{cases}$.

By taking $S^* = \{x \in \mathbb{R}^n \mid Ax - Dx = b\} \cap V$ we obtain a convex polyhedron. This set is non-empty because $\bar{x} \in S^*$. In the neighbourhood V, solving $Ax - Dx = b$ gives a solution of (AVE). Using Hoffman lemma [6] for r sufficiently small such that $x_r \in V$ there exists a constant K such that

$$\begin{aligned} d_{S^*}(x^r) \leq K \left\| \begin{matrix} (A - D)x^r - b \\ (x^r - \alpha - \bar{x})^+ \\ (-x^r - \alpha + \bar{x})^+ \end{matrix} \right\| \quad &\leq \quad K(\|(A - D)x^r - b\| + \|(x^r - \alpha - \bar{x})^+\| \\ &+ \quad \|(-x^r - \alpha + \bar{x})^+\|), \\ &= \quad K\|(A - D)x^r - b\|, \\ &= \quad K\|Ax^r - |x^r| - b\|. \end{aligned}$$

Since x^r is feasible for (P_r), it holds that

$$\begin{aligned}||Ax^r - |x^r| - b|| &= ||g(r)Ae - g(r)\delta(x^r)||, \\ &= ||(Ae - \delta(x^r))g(r)||, \\ &\leq ||Ae - \delta(x^r)||\ |g(r)|, \\ &= ||Ae - \delta(x^r)||g(r) = O(g(r)).\end{aligned}$$

Combining both previous inequalities, we obtain

$$d_{S^*}(x^r) \leq K||Ae - \delta(x^r)||g(r) = O(g(r))\ .$$

b) Now we move to the case where $\exists i \in \{1, ..., N\},\ \bar{x}_i = 0$. We denote $\sigma(t) = \{i | t_i \neq 0\}$. Set $\alpha = \min_{i\in\sigma(\bar{x})} |\bar{x}_i|/2$ and a neighbourhood V of $\bar{x}$ defined as

$$V = B_\infty(\bar{x}, \alpha) = \{x \mid \max_{i\in\sigma(\bar{x})} |x_i - \bar{x}_i| \leq \alpha\}\ .$$

V is non-empty because $\bar{x} \in V$. For all $x \in V$, $\bar{x}$ and x have the same signs for the components $i \in \sigma(\bar{x})$. Furthermore, for r sufficiently small we have $x^r \in V$.

Taking $S^* = \{x \in \mathbb{R}^n \mid Ax - Dx = b\ ,\ Dx \geq 0\} \cap V$ with $D =\ diag(\delta(x^r))$ we obtain again a convex polyhedron. The choice of D depending on x^r is not restrictive as we can always take a subsequence of the sequence $\{x^r\}_{r>0}$, which converge to $\bar{x}$, with constant signs near $\bar{x}$. This set is non-empty because $\bar{x} \in S^*$. In the neighbourhood V, solving $Ax - Dx = b$ with the constraints $Dx \geq 0$ gives a solution of (AVE). Using Hoffman lemma [6] there exists a constant K such that

$$\begin{aligned}d_{S^*}(x^r) \leq K\left\| \begin{matrix} (A-D)x^r - b \\ (x^r - \alpha - \bar{x})^+ \\ (-x^r - \alpha + \bar{x})^+ \\ (-Dx)^+ \end{matrix} \right\| &\leq K(\|(A-D)x^r - b\| + \|(x^r - \alpha - \bar{x})^+\| \\ &+ \|(-x^r - \alpha + \bar{x})^+\| + \|(-Dx)^+\|), \\ &= K\|(A-D)x^r - b\|, \\ &= K\|Ax^r - |x^r| - b\|.\end{aligned}$$

As x^r is feasible for (P_r), we have

$$\begin{aligned}||Ax^r - |x^r| - b|| &\leq ||g(r)Ae - g(r)\delta(x^r)||, \\ &= ||(Ae - \delta(x^r))g(r)||, \\ &\leq ||Ae - \delta(x^r)||\ |g(r)|, \\ &= ||Ae - \delta(x^r)||g(r) = O(g(r)).\end{aligned}$$

Combining both previous inequalities, we obtain

$$d_{S^*}(x^r) \leq K||Ae - \delta(x^r)||g(r) = O(g(r)).$$

In both cases a) and b), the proof is complete since $S^* \subset S^*_{(AVE)}$. $\square$

Remark 1. We can be a bit more specific in the case where $(A-D)$ is invertible. In this case $S^* = \{\bar{x}\}$, so (4) becomes

$$||x^r - \bar{x}|| \leq ||(A-D)^{-1}|| \; ||Ae - \delta(x)||g(r) = O(g(r)) \; .$$

This case corresponds to the special cases where (AVE) has isolated solutions.

5 Algorithm

In the previous sections we present theoretical results about convergence and error estimate of an algorithm to compute a solution of (AVE). In this section, we focus on the algorithm and its implementation.

Consider the generic algorithm where C_k is the feasible set of (P_{r^k}) :

$$\begin{cases} \{r_k\}_{k\in\mathbb{N}}, \; r_0 > 0 \text{ and } \lim\limits_{k\to+\infty} r_k = 0 \\ \text{find } x^k : \; x^k \in \arg\min\limits_{x\in C_k} \sum_{i=1}^n \theta_{r_k}(x_i^+) + \theta_{r_k}(x_i^-) - 1 \end{cases} . \tag{TAVE}$$

In a practical implementation of (TAVE) one should probably more likely use the initial problem (2) with the constraint $x^+ + x^- \geq g(r)e$. The sequence of computed points will probably be infeasible but we believe that it leads to improved numerical behaviour. The constraint $x^+ + x^- \geq g(r)e$ prevents the sequence to possibly go to a local minimum with a zero component.

Algorithm TAVE requires an initial point. In a same way as in [12, 13], one can use the solution of the following linear program

$$\begin{cases} \min\limits_{x^+, x^- \in \mathbb{R}^N_+} \; (x^+ + x^-)^T e \\ A(x^+ - x^-) - (x^+ + x^-) = b \end{cases} .$$

Indeed, this program find an initial feasible point of (P_r) and the objective function may encourage this point to satisfy the complementarity condition.

In this study we put the variables in a compact set. Indeed, the functions θ_r are more efficient when their arguments live in $[0,1]$. Besides, we use one way to express complementarity with Lemma 2.1 another way, which will be used in the numerical study, is to consider the following

$$\theta_r(s) + \theta_r(t) - \theta_r(s+t) = 0. \tag{5}$$

In this case we don't necessarily need the constraint $x^+ + x^- \geq g(r)$, since it is a reformulation of the complementarity and no longer a relaxation.

Regarding the choice of the parameters α, r_0 and the update parameter of r, it is to be noted that they are all used in the constraint $x^+ + x^- \geq g(r)e$ with $g(r) = r^\alpha$ and $0 < \alpha < 1$. Theorem 4.4 shows that the convergence to the zero part of the solution is a $O(g(r))$. So it is clear that α needs to be taken as big as possible, for instance $\alpha = 0.99$. Also there is a link between the value of

α and the update rule for r. We choose to select a sequence of values with an update constant T, so that $r_{k+1} = \frac{r_k}{T}$. The initial parameter r_0 can be chosen according the relation

$$\theta_r^1(s) + \theta_r^1(t) = 1 \Longleftrightarrow st = r^2 \ .$$

At each step in r, we solve a concave optimization problem to get the current point. The following heuristic can be rather useful to accelerate convergence and assure a good precision when we are close to the solution. After finding the current point x^k we solve if possible the linear system

$$(A - \operatorname{diag}(\delta(x^k)))z = b. \tag{6}$$

If x solves (AVE), then we solved (AVE) with the same precision as we solved the linear system. However, if x does not solve (AVE), we continue the iteration in r with x^k. This idea is similar to compute a Newton iteration.

6 Numerical Simulations

We present numerical results on two examples and some randomly generated problems. Those random problems can be divided into two different classes: problems with singular values exceeding one and problems that are completely general without any particular property. The latter class is the most interesting for us since numerical methods can fail to obtain a solution. This is the main motivation for our approach.

These simulations have been done using MATLAB, [18], with the linear programming solver GUROBI, [3]. We used the Successive Linearisation Algorithm (SLA) of [10] to solve concave minimization problems encountered at each iteration.

Proposition 6.1 (SLA for concave minimization)**.** *Given ϵ sufficiently small and r^k. Denote C the feasible set of (P_{r^k}). Given $x^k = x^{k+} - x^{k-}$, x^{k+1} is designed as a solution of the linear problem*

$$\min_{y^+, y^- \in C} \ (y^+)^T \nabla \theta_{r^k}(x^{k+}) + (y^-)^T \nabla \theta_{r^k}(x^{k-}),$$

with $x^0 = x^{0+} - x^{0-}$ a random point. We stop when

$$x^{k+1} \in C \text{ and } (x^{k+1} - x^k)^T \nabla \theta_{r^k}(x^k) \leq \epsilon.$$

This algorithm generates a finite sequence with strictly decreasing objective function values.

Proof. see [[10], Theorem 4.2]. □

Along these simulations we use the parameters detailed in Table 1 for TAVE. The maximum number of iterations in r for one instance is fixed to 20 and the

T	initial r : r_0	function θ_r	α
1.8	1	θ_r^2	0.99

Table 1: Parameters for the simulations

maximum number of linear programs for one SLA is fixed to 10. We measure the time in seconds, the number of linear programs solved and the number of linear systems solved respectively denoted by nb-LP-method and nb-lin-syst-method.

We consider two concrete examples and then two kinds of randomly generated problems. The first one is a second order ordinary differential equation with initial conditions and the second example is an obstacle problem. We remind that our main motivation is to consider general absolute value equations and this has been treated in subsection 6.4.

6.1 An Ordinary Differential Equation

We consider the ordinary differential equation

$$\ddot{x}(t) - |x(t)| = 0,\ x(0) = x_0,\ \dot{x}(0) = \gamma\ .$$

We get an (AVE) by using a finite difference scheme in order to discretize this equation. We use the following second-order backward difference to approximate the second derivative

$$\frac{x_{i-2} - 2x_{i-1} + x_i}{h^2} - |x_i| = 0\ .$$

This equation was derived with an equispaced grid $x_i = ih$, $i = 1, ...N$. Neumann boundary conditions were approximated using a centred finite difference scheme

$$\frac{x_{-1} - x_1}{2h} = \gamma\ .$$

We compare the obtained solution by TAVE to the one of the predefined Runge-Kutta ode45 function in MATLAB, [18]. The domain is $t \in [0, 4]$, initial conditions $x_0 = -1$, $\gamma = 1$ and $N = 100$. Results are presented in Figure 1. TAVE solves the problem and gives consistent results.

6.2 Obstacle Problem

In this simple obstacle problem, we try to compute a trajectory joining the boundary of a domain with an obstacle, g, and a minimal curvature, f. This can be formulated using the following equation and inequalities: find u such that

$$(\ddot{u}(x) - f(x))^T(u(x) - g(x)) = 0,\ \ddot{u}(x) - f(x) \geq 0,\ u(x) - g(x) \geq 0\ .$$

We approximate the second order derivative with a second-order centred finite difference to get a discrete version on an equispaced grid $x_i = ih$, $i = 1, ...N$.

$$(Du-f)^T(u-g) = 0,\ Du-f \geq 0,\ u-g \geq 0,\ \text{where } D = \begin{pmatrix} \frac{2}{h^2} & \frac{-1}{h^2} & & \\ \frac{-1}{h^2} & \ddots & \ddots & \\ & \ddots & \ddots & \frac{-1}{h^2} \\ & & \frac{-1}{h^2} & \frac{2}{h^2} \end{pmatrix},$$

$g_i = g(x_i)$ and $f_i = f(x_i)$. This can be written as a linear complementarity problem by setting $z = u - g$, $M = D$ and $q = Dg - f$, that is

$$(Mz + q)^T z = 0,\ Mz + q \geq 0,\ z \geq 0\ . \tag{7}$$

This equation is equivalent to an (AVE) whenever 1 is not an eigenvalue of M by proposition 2 of [14]:

$$(7) \Longleftrightarrow (M - I)^{-1}(M + I)x - |x| = (M - I)^{-1}q\ .$$

We present results for our method and for the LPM method from [12], with $g(x) = \max(0.8 - 20(x - 0.2)^2, \max(1 - 20(x - 0.75)^2, 1.2 - 30(x - 0.41)^2))$, $f(x) = 1$, $N = 50$ in Figure 2. Both methods give 20 points on the curve g and none below g over 50 points. Once again TAVE method gives consistent results.

6.3 Random Easy Problems with Unique Solution

We consider the special case where (AVE) has a unique solution. Following [29], we generate data (A, b) by the following Matlab code for $n = 100, 200, 400, 800$

```
n=input('dimension of matrix A =');
rand('state',0);
R=rand(n,n);
b=rand(n,1);
A=R'*R+n*eye(n);
```

Then, the corresponding (AVE) has a unique solution and A is a symmetric matrix. The required precision for solving (AVE) is 10^{-6} and thanks to the heuristic from Section 5 in equation (6) we get in the worst case 10^{-10}. For each n we consider 100 instances. TAVE is compared to a Newton method from [15], which is denoted GN. Results are summarized in Table 2, which gives for TAVE the number of linear programs solved, the time required to solve all the instances and gives for GN the number of linear systems and the time required. Note that other Newton methods like [2, 5, 20] should give similar conclusions so we do not include them in our comparison. It is to be expected that the method GN is faster than the method TAVE in particular when the dimension grows since it solves only linear systems whereas TAVE solves linear programs.

n	nb-LP-TAVE	time TAVE	nb-lin-syst-GN	time GN
32	100	0.1841	217	0.0403
64	100	0.4702	224	0.0553
128	100	1.5880	219	0.1079
256	100	14.5161	226	0.3924
512	100	129.1686	214	2.2327

Table 2: TAVE and Newton method from [15], GN, for (AVE) in the case with singular values of A exceeds 1.

Our method solves all the problems, which once again valid our approach. We observe that the number of solved linear programs is very low. Indeed in each case, the initialization step has been sufficient to solve the problem. The method GN outperform in time TAVE as it was expected. Since some other methods to solve (AVE) have no problem solving these instances they are not our main focus here.

6.4 General Problems

Now, we consider general (AVE) without any assumption on A. This kind of problems correspond to the main goal of this paper. We generate for several n and several values of the parameters one hundred instances of the problem following [14]:

"Choose a random A from a uniform distribution on $[-10, 10]$, then choose a random x from a uniform distribution on $[-1, 1]$ and set $b = Ax - |x|$."

We compare 4 methods tailored for general (AVE):

- TAVE method from Algorithm TAVE;
- TAVE2 which is the same algorithm with the different objective based on (5)
$$\sum_{i=1}^{N} \theta_r(x_i^+) + \theta_r(x_i^-) - \theta_r(x_i^+ + x_i^-) \ ;$$
- concave minimization method CMM from [14];
- successive linear programming method LPM from [12].

In this situation GN method has not been applied since it has no guarantee of convergence.

In Tables 3–6,"nnztot" gives the number of violated expressions for all problems, "nnzx" gives the maximum violated expressions for one problem, "out-iter" gives the number of iteration in r and "in-iter" gives the number of linear programs solved for all the problems. We also provide the time in seconds and the number of failures.

In each case our methods manage to reduce the number of unsolved problems. This confirm the interest of the relaxation method presented here. Also one should note that an improved number of solved problems comes with a price, since it requires more time.

Table 4 shows promising results for TAVE2. It is a slightly different method, since it is not a relaxation but a reformulation of the complementarity. In every case it gives the smallest number of unsolved problem in a very reasonable time.

n	nnztot	nnzx	out-iter	in-iter	time	nb-failure
32	0	0	74	306	0.6234	0
64	3	1	156	491	2,8173	3
128	8	1	269	841	20,9447	8
256	8	1	324	1129	281,6190	8

Table 3: TAVE

n	nnztot	nnzx	out-iter	in-iter	time	nb-failure
32	0	0	33	**164**	**0.3137**	**0**
64	2	1	81	**221**	**1,1280**	**2**
128	4	1	136	**303**	**6,3953**	**4**
256	4	1	131	**292**	**56,4148**	**4**

Table 4: TAVE2

n	nnztot	nnzx	out-iter	in-iter	time	nb-failure
32	9	1	-	485	1.0823	9
64	8	1	-	458	2,9234	8
128	10	1	-	568	18,4404	10
256	11	1	-	595	124,5728	11

Table 5: CMM

Conclusion and Perspectives

In this paper, we propose a class of heuristic schemes to solve the NP-hard problem (AVE). A complete analysis is provided including convergence and error estimate.

Furthermore, a numerical study shows that our approach is full of interest. Indeed, our methods turn out to be consistent with real examples and problems

n	nnztot	nnzx	out-iter	in-iter	time	nb-failure
32	7	1	-	248	0.5546	7
64	19	4	-	342	2,5822	13
128	19	3	-	409	16,6830	13
256	29	5	-	439	143,0973	11

Table 6: LPM

with unique solution. We do not compare our method with those specially designed for these problems since they do not belong to the same class. Finally, the last set of generated problems concerns general (AVE). In comparison with some existing methods, our approach improve the number of failures. This was our main goal.

It is of interest to note that the methods presented here can also solve the linear complementarity problem using the same technique as for the obstacle problem example.

Further studies could improve the choice of parameters in order to reduce the computational time to solve the problems, especially for large instances. Promising results were shown by the modified algorithm TAVE2, which considers a different way to formulate the penalty. So we may wonder if it is possible to improve our algorithms in this case and if there exists other similar reformulations of the complementarity which can give even better results.

We are working on a hybrid algorithm which can benefit from both the minimization methods and Newton methods as in [13]. Indeed, this philosophy is fully applicable here and can lead to computational improvements.

Acknowledgements

The authors would like to thank anonymous referees for their helpful remarks and comments.

References

References

[1] J. Y. Bello Cruz, O. P. Ferreira, and L. F. Prudente. On the global convergence of the inexact semi-smooth Newton method for absolute value equation. *Computational Optimization and Applications*, 65(1):93–108, sep 2016.

[2] L. Caccetta, B. Qu, and G. Zhou. A globally and quadratically convergent method for absolute value equations. *Computational Optimization and Applications*, 48(1):45–58, 2011.

[3] Inc. Gurobi Optimization. Gurobi optimizer reference manual, 2015.

[4] M. Haddou and P. Maheux. Smoothing methods for nonlinear complementarity problems. *Journal of Optimization Theory and Applications*, 160(3):711–729, 2014.

[5] F. K. Haghani. On generalized traub's method for absolute value equations. *Journal of Optimization Theory and Applications*, pages 1–7, 2015.

[6] A. J. Hoffman. On approximate solutions of systems of linear inequalities. *Journal of Research of the National Bureau of Standard*, 49(2):pp 263–265, 1952.

[7] S. Ketabchi and H. Moosaei. An efficient method for optimal correcting of absolute value equations by minimal changes in the right hand side. *Computers & Mathematics with Applications*, 64(6):1882–1885, 2012.

[8] S. Ketabchi, H. Moosaei, and S. Fallahi. Optimal error correction of the absolute value equation using a genetic algorithm. *Mathematical and Computer Modelling*, 57(9):2339–2342, 2013.

[9] T. Lotfi and H. Veiseh. A note on unique solvability of the absolute value equation. 2013.

[10] O. L. Mangasarian. Machine learning via polyhedral concave minimization. In *Applied Mathematics and Parallel Computing*, pages 175–188. Springer, 1996.

[11] O. L. Mangasarian. Absolute value equation solution via dual complementarity. *Optimization Letters*, 7(4):625–630, 2013.

[12] O. L. Mangasarian. Absolute value equation solution via linear programming. *Journal of Optimization Theory and Applications*, 161(3):870–876, 2014.

[13] O. L. Mangasarian. A hybrid algorithm for solving the absolute value equation. *Optimization Letters*, pages 1–6, 2015.

[14] O.L. Mangasarian. Absolute value equation solution via concave minimization. *Optimization Letters*, 1(1):3–8, 2007.

[15] O.L. Mangasarian. A generalized newton method for absolute value equations. *Optimization Letters*, 3(1):101–108, 2009.

[16] O.L. Mangasarian. Linear complementarity as absolute value equation solution. *Optimization Letters*, 8(4):1529–1534, 2014.

[17] O.L. Mangasarian and R.R. Meyer. Absolute value equations. *Linear Algebra and Its Applications*, 419(2):359–367, 2006.

[18] MATLAB. *version R2013*. The MathWorks Inc., Natick, Massachusetts, 2014.

[19] T. Migot and M. Haddou. A smoothing method for sparse optimization over polyhedral sets. In *Modelling, Computation and Optimization in Information Systems and Management Sciences*, pages 369–379. Springer, 2015.

[20] H. Moosaei, S. Ketabchi, M.A. Noor, J. Iqbal, and V. Hooshyarbakhsh. Some techniques for solving absolute value equations. *Applied Mathematics and Computation*, 268:696–705, 2015.

[21] O. Prokopyev. On equivalent reformulations for absolute value equations. *Computational Optimization and Applications*, 44(3):363–372, 2009.

[22] Y.-E. Qi and M.-Z. Zhu. The nonlinear HSS-like iteration method for absolute value equations. *Applied Mathematics and Computation*, 265(11271174):266–274, mar 2014.

[23] J. Rohn. A theorem of the alternatives for the equation $ax + b|x| = b$. *Linear and Multilinear Algebra*, 52(6):421–426, 2004.

[24] J. Rohn. An algorithm for solving the absolute value equation. *Electronic Journal of Linear Algebra*, 18(589-599):5, 2009.

[25] J. Rohn. On unique solvability of the absolute value equation. *Optimization Letters*, 3(4):603–606, 2009.

[26] J. Rohn. An algorithm for solving the absolute value equation: An improvement. Technical report, 2010.

[27] J. Rohn, V. Hooshyarbakhsh, and R. Farhadsefat. An iterative method for solving absolute value equations and sufficient conditions for unique solvability. *Optimization Letters*, 8(1):35–44, 2014.

[28] D. K. Salkuyeh. The Picard–HSS iteration method for absolute value equations. *Optimization Letters*, 8(8):2191–2202, 2014.

[29] L. Yong. Particle swarm optimization for absolute value equations. *Journal of Computational Information Systems*, 6(7):2359–2366, 2010.

[30] L. Yong. A Smoothing Newton Method for Absolute Value Equation. *International Journal of Control and Automation*, 9(2):119–132, feb 2016.

[31] J.-J. Zhang. The relaxed nonlinear PHSS-like iteration method for absolute value equations. *Applied Mathematics and Computation*, 265:266–274, 2015.

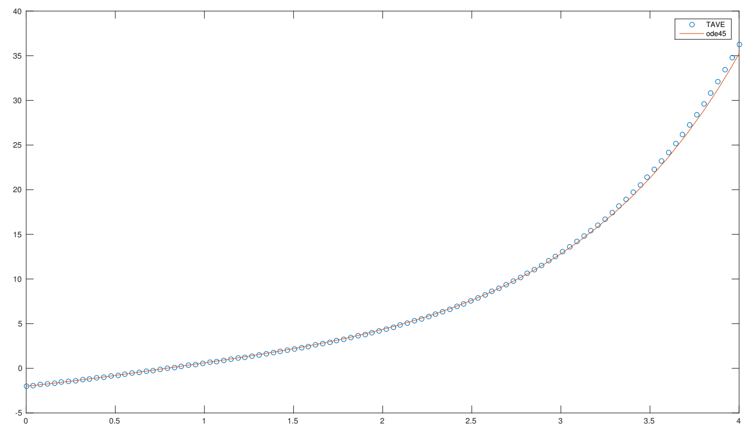

Figure 1: Numerical solution of (6.1) with ode45 and ThetaAVE.

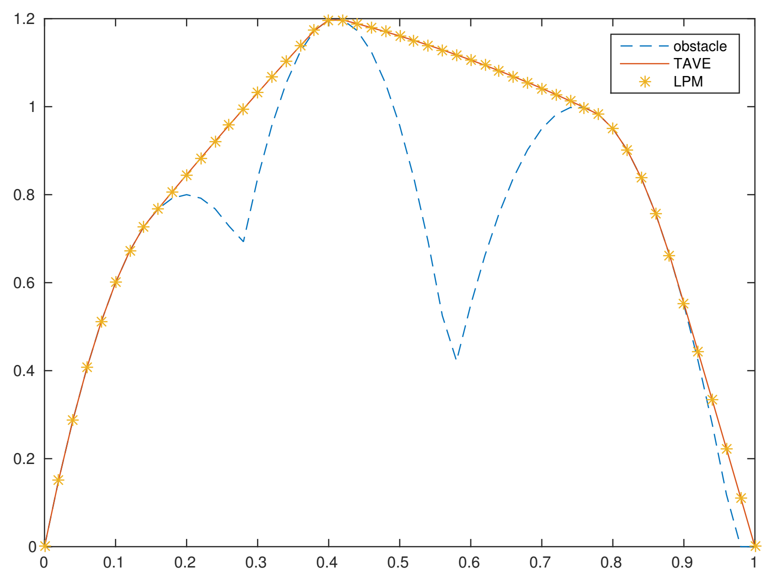

Figure 2: Numerical solution of the obstacle problem (6.2) with ThetaAVE and the method from [14]