A framework for convex-constrained monotone nonlinear equations and its special cases

Max L. N. Gonçalves [*] Tiago C. Menezes[*]

June 30, 2021

Abstract

This work refers to methods for solving convex-constrained monotone nonlinear equations. We first propose a framework, which is obtained by combining a safeguard strategy on the search directions with a notion of approximate projections. The global convergence of the framework is established under appropriate assumptions and some examples of methods which fall into this framework are presented. In particular, inexact versions of steepest descent-based, spectral gradient-like, Newton-like and limited memory BFGS methods are discussed. Numerical experiments illustrating the practical behavior of the algorithms are discussed and comparisons with existing methods are also presented.

Keywords: Monotone nonlinear equations; approximate projection; global convergence; steepest descent-based method; spectral gradient-like method; Newton-like method.

1 Introduction

We consider the convex-constrained monotone nonlinear equations problem: finding $x_* \in C$ such that

$$F(x_*) = 0, \tag{1}$$

where C is a nonempty closed convex set and $F : \mathbb{R}^n \to \mathbb{R}^n$ is a continuous and monotone nonlinear function, not necessarily differentiable. The monotonicity of $F : \mathbb{R}^n \to \mathbb{R}^n$ here means $\langle F(x) - F(y), x - y \rangle \geq 0$, for all $x, y \in \mathbb{R}^n$. We assume that the solution set of (1), denoted by C^*, is nonempty. Problems of this nature appear in many applications such as power engineering, chemical equilibrium systems and economic equilibrium problems; see, e.g., [7, 10, 18, 26].

Recently, many methods, which are extensions of Newton, spectral gradient and conjugate gradient methods for solving the unconstrained monotone nonlinear equations, have been proposed for solving (1);

[*]IME, Universidade Federal de Goiás, Goiânia, GO 74001-970, Brazil. (E-mails: maxlng@ufg.br and tiagomenezes@discente.ufg.br). The work of these authors was supported in part by CAPES and CNPq Grant 408123/2018-4.

see, for example, [14, 15, 21, 24, 25, 27]. In particular, these constrained methods have to compute the exact orthogonal projection of a point onto the constrained set C. However, it is well-known that, depending on the geometry of C, the projection onto it neither has a closed-form nor can it be easily computed.

This work refers to some methods for solving (1) in which approximate projections are allowed. Toward this goal, we first present a framework for solving convex-constrained monotone nonlinear equations. More precisely, at each iteration, the framework imposes a safeguard strategy on the search directions. A suitable linesearch procedure based on [22] is considered, which, in particular, provides a hyperplane that strictly separates the current iteration from zeroes of the system of equations. Then, we compute an approximate projection of a point, which belongs to the aforementioned hyperplane, onto the intersection between C and the hyperplane (or onto the constrained set C). Under mild assumptions, we prove that the sequence generated by the proposed framework converges to a solution of (1). Some examples of methods which fall into this framework are presented. In particular, inexact versions of steepest descent-based, spectral gradient-like, Newton-like and limited memory BFGS methods are discussed.

In order to illustrate the robustness and effectiveness among the instances of the framework, we report some preliminary numerical experiments on a set of monotone nonlinear equations with polyhedral constraints problems. Moreover, we also applied the methods for solving the constrained absolute value equation and compare their performances with the inexact Newton method with feasible inexact projections [20].

The remainder of this paper is organized as follows. In Section 2, we list some notations and introduce a concept of an approximate solution for a specific quadratic problem with some of its properties. In Section 3, a framework, which combines a safeguard strategy on the search directions with approximate projections, is proposed and its global convergence is established. In Section 4, we present some instances of the framework by means of some examples of search directions d_k that satisfy the safeguard conditions. Some preliminary numerical experiments are reported in Section 5 and final remarks are given in Section 6.

2 Notation and preliminary results

This section presents some definitions and notations used in this paper. A concept of approximate solution for a specific quadratic problem is introduced and some of its useful properties are discussed.

Let $\mathbb{B}$ be the set of $n \times n$ symmetric positive definite matrices such that

$$\|B\| \leq L \qquad \text{and} \qquad \|B^{-1}\| \leq L, \tag{2}$$

where $L > 1$ and $\|.\|$ is a sub-multiplicative matrix norm. Note that $\mathbb{B}$ is a compact set of $\mathbb{R}^{n\times n}$. Consider an inner product on $\mathbb{R}^n$ by setting $\langle x, z\rangle_B = \langle x, Bz\rangle$, where $B \in \mathbb{B}$ and $\langle ., .\rangle$ denotes the usual inner product. Hence, the corresponding induced norm $\|.\|_B$ is equivalent to the usual norm on $\mathbb{R}^n$, since the following inequalities hold

$$\frac{1}{\|B^{-1}\|}\|x\|^2 \leq \|x\|_B^2 \leq \|B\|\|x\|^2. \tag{3}$$

We next present the definition of an approximate solution for a specific quadratic problem.

Definition 1. *Given $B \in \mathbb{B}$, $w \in \mathbb{R}^n$, $\varepsilon \geq 0$ and a nonempty closed convex set $\Omega \subset \mathbb{R}^n$, we say that $\tilde{y}^B_\Omega(w)$ is an ε–approximate solution for the problem*

$$\min_{y \in \Omega} \frac{1}{2}\langle By, y\rangle - \langle w, y\rangle, \tag{4}$$

iff

$$\tilde{y}^B_\Omega(w) \in \Omega \qquad \text{and} \qquad \langle B\tilde{y}^B_\Omega(w) - w, y - \tilde{y}^B_\Omega(w)\rangle \geq -\varepsilon, \quad \forall\, y \in \Omega. \tag{5}$$

Remark 1. *Since in* (4) *we are minimizing a strictly convex quadratic function over a convex set,* (5) *is a natural condition for an approximate solution. Indeed, the optimality condition for* (4) *is*

$$\langle B\tilde{y}^B_\Omega(w) - w, y - \tilde{y}^B_\Omega(w)\rangle \geq 0, \quad \forall\, y \in \Omega,$$

which can be obtained from (5) *by setting $\varepsilon = 0$. Note that, if $\tilde{y}^B_\Omega(w)$ is a zero–approximate solution, then $\tilde{y}^B_\Omega(w)$ is the unique exact solution of* (4)*, which we will denote by $y^B_\Omega(w)$.*

Note that if $w := Bx - F(z)$ with $x, z \in \mathbb{R}^n$, then problem (4) can be rewritten, ignoring constant terms, as

$$\min_{y \in \Omega} \frac{1}{2}\|y - (x - B^{-1}F(z))\|^2_B. \tag{6}$$

and (5) is equivalent to

$$\langle x - B^{-1}F(z) - \tilde{y}^B_\Omega(w), y - \tilde{y}^B_\Omega(w)\rangle_B \leq \varepsilon, \quad \forall\, y \in \Omega. \tag{7}$$

In this case, we can say that $\tilde{y}^B_\Omega(Bx - F(z))$ is an approximate projection (in the norm $\|\cdot\|_B$) of the direction $x - B^{-1}F(z)$.

Our condition (5) can be checked when, for example, Ω is bounded and the conditional gradient (CondG) method, also known as Frank-Wolfe method [9, 11], is used to solve (4). If we apply the CondG method to (4), we generate a sequence $\{z_j\} \subset \Omega$, where $z_j = z_{j-1} + \alpha_j(\bar{z}_j - z_{j-1})$, with $\alpha_j \in (0,1)$, and $\bar{z}_j$ is a solution of the subproblem

$$\begin{array}{ll} \min & \langle Bz_{j-1} - w, z - z_{j-1}\rangle, \\ \text{s.t.} & z \in \Omega \end{array} \tag{8}$$

which we assume is given exactly by a "linear oracle". If the CondG iterations are stopped when

$$\langle Bz_{j-1} - w, \bar{z}_j - z_{j-1}\rangle \geq -\varepsilon, \tag{9}$$

then condition (5) holds with $\tilde{y}^B_\Omega(w) = z_{j-1}$.

We next establish a useful relationship between exact and inexact solutions of (4).

Lemma 1. *For every $w, \hat{w} \in \mathbb{R}^n$ and $\varepsilon \geq 0$, we have*

$$\|\tilde{y}^I_\Omega(w) - y^I_\Omega(\hat{w})\|^2 \leq \|w - \hat{w}\|^2 + 2\varepsilon.$$

Proof. Since $\tilde{y}^I_\Omega(w) \in \Omega$ and $y^I_\Omega(\hat{w}) \in \Omega$, it follows from Definition 1 that

$$\langle \tilde{y}^I_\Omega(w) - w, \tilde{y}^I_\Omega(w) - y^I_\Omega(\hat{w}) \rangle \leq \varepsilon, \qquad \langle \hat{w} - y^I_\Omega(\hat{w}), \tilde{y}^I_\Omega(w) - y^I_\Omega(\hat{w}) \rangle \leq 0. \tag{10}$$

On the other hand, after some simple algebraic manipulations we have

$$\begin{aligned} \|w - \hat{w}\|^2 = {} & \|\tilde{y}^I_\Omega(w) - y^I_\Omega(\hat{w})\|^2 + 2\langle w - \tilde{y}^I_\Omega(w) - (\hat{w} - y^I_\Omega(\hat{w})),\ \tilde{y}^I_\Omega(w) - y^I_\Omega(\hat{w}) \rangle \\ & + \|(w - \tilde{y}^I_\Omega(w)) - (\hat{w} - y^I_\Omega(\hat{w}))\|^2, \end{aligned}$$

which implies that

$$\begin{aligned} \|\tilde{y}^I_\Omega(w) - y^I_\Omega(\hat{w})\|^2 \leq {} & \|w - \hat{w}\|^2 + 2\langle \tilde{y}^I_\Omega(w) - w,\ \tilde{y}^I_\Omega(w) - y^I_\Omega(\hat{w}) \rangle \\ & + 2\langle \hat{w} - y^I_\Omega(\hat{w}),\ \tilde{y}^I_\Omega(w) - y^I_\Omega(\hat{w}) \rangle. \end{aligned}$$

By the last inequality, (10) and (2), it yields

$$\|\tilde{y}^I_\Omega(w) - y^I_\Omega(\hat{w})\|^2 \leq \|w - \hat{w}\|^2 + 2\varepsilon,$$

which is equivalent to the desired inequality. □

3 A framework for solving convex-constrained monotone equations

This section describes a framework for solving (1) and presents its global convergence analysis. The framework imposes a safeguard strategy on the search directions which, combined with a suitable line-search procedure, turns it a globalized scheme. Approximate projections are also considered in order to deal with the case where projecting exactly onto C is expensive. Formally, the framework is described as follows.

Framework 1. A framework for solving (1)

Step 0. Let $x_0 \in C$, $\eta_1, \eta_2 > 0$, $\gamma, \sigma \in (0,1)$, $\bar{\mu} \in [0,1)$ and $\{\mu_k\} \subset [0, \bar{\mu}]$ be given, and set $k = 0$.

Step 1. If $\|F(x_k)\| = 0$, then stop.

Step 2. Compute the direction d_k in $\mathbb{R}^n$ such that

$$F(x_k)^T d_k \leq -\eta_1 \|F(x_k)\|^2, \tag{11}$$

$$\|d_k\| \leq \eta_2 \|F(x_k)\|. \tag{12}$$

Step 3. Find $z_k = x_k + \alpha_k d_k$, where $\alpha_k = \gamma^{m_k}$ with m_k being the smallest nonnegative integer m such that

$$-\langle F(x_k + \gamma^m d_k),\ d_k\rangle \geq \sigma\gamma^m \|d_k\|^2. \tag{13}$$

Step 4. Define $\xi_k := (\langle F(z_k),\ x_k - z_k\rangle)/\|F(z_k)\|^2$, $w_k := x_k - \xi_k F(z_k)$ and $\varepsilon_k := \mu_k^2\|\xi_k F(z_k)\|^2$. Set

$$x_{k+1} := \tilde{y}^I_{C\cap H_k}(w_k), \tag{14}$$

where $H_k := \{x \in \mathbb{R}^n;\ \langle F(z_k), x - z_k\rangle \leq 0\}$.

Step 5. Set $k \leftarrow k+1$ and go to Step 1.

end

Remark 2. *i) If F is the gradient of some function $f : \mathbb{R}^n \to \mathbb{R}$, then condition* (11) *implies that d_k is a sufficient descent direction for f at x_k. In its turn, condition* (12) *essentially says that the length of $d(x_k)$ should be proportional to the length of $F(x_k)$. The way to obtain d_k satisfying* (11) *and* (12) *will depend on the particular instance of the framework; see Section 4 for some examples. (ii) Note that condition* (11) *implies that there exists a nonnegative number m_k satisfying* (13)*, for all $k \geq 1$. Indeed, suppose that there exists $k_0 \geq 1$ such that* (13) *is not satisfied for any nonnegative integer m, i.e.,*

$$-\langle F(x_{k_0} + \gamma^m d_{k_0}),\ d_{k_0}\rangle < \sigma\gamma^m \|d_{k_0}\|^2,\ \forall\, m \geq 1.$$

Let $m \to \infty$ and by continuity of F, we have

$$-\langle F(x_{k_0}),\ d_{k_0}\rangle \leq 0. \tag{15}$$

On the other hand, by (11)*, we obtain*

$$-\langle F(x_{k_0}),\ d_{k_0}\rangle \geq \eta_1 \|F(x_{k_0})\|^2 > 0,$$

which contradicts (15)*. Therefore, the linesearch procedure in Step 3 is well defined. iii) In Step 4, note that w_k is the projection of x_k in H_k (which has a closed-form) and x_{k+1} is an ε_k–approximate solution of the problem* (4) *with $B := I$, $w := w_k$ and $\Omega := C \cap H_k$. Another choice of x_{k+1} in* (14) *would be*

$$x_{k+1} := \tilde{y}^I_C(w_k). \tag{16}$$

For this choice, we mention that Lemma 2 and Theorem 3 also holds. iv) It will follow form (18) *and* (19) *that the hyperplane H_k strictly separates the current iteration from the elements of the solution set C^*.*

In order to investigate the global convergence of Framework 1, the following properties of the sequences $\{x_k\}$ and $\{z_k\}$ will be needed.

Lemma 2. *The sequences $\{x_k\}$ and $\{z_k\}$ generated by Framework 1 are both bounded. Furthermore, it holds that*

$$\lim_{k\to\infty} \|x_k - z_k\| = 0. \tag{17}$$

Proof. From Step 3, we have

$$\langle F(z_k),\, x_k - z_k\rangle = -\alpha_k \langle F(z_k),\, d_k\rangle \geq \sigma\alpha_k^2 \|d_k\|^2 = \sigma\|x_k - z_k\|^2. \tag{18}$$

Note that $\|x_k - z_k\| > 0$, for all $k \geq 0$. Otherwise, since (11) and the Cauchy-Schwartz inequality imply that $\eta_1\|F(x_k)\| \leq \|d_k\|$, we would have $F(x_k) = 0$. Let $x_* \in C^*$ be given. By the monotonicity of F and the fact that $F(x_*) = 0$, we obtain

$$\langle F(z_k),\, x_* - z_k\rangle \leq 0. \tag{19}$$

Hence, $x_* \in H_k$ (see the definition of H_k in Step 4). Since $x_{k+1} = \tilde{y}^I_{C\cap H_k}(w_k)$, it follows from the fact that $x^* \in C \cap H_k$ and Lemma 1 with $B = I$, $x = w_k$ and $\hat{x} = x_*$ that

$$\begin{aligned} \|x_{k+1} - x_*\|^2 &= \|\tilde{y}^I_{C\cap H_k}(w_k) - y^I_{C\cap H_k}(x_*)\|^2 \leq \|w_k - x_*\|^2 + 2\varepsilon_k \\ &= \|x_k - x_*\|^2 - 2\xi_k\langle F(z_k),\, x_k - x_*\rangle + \xi_k^2\|F(z_k)\|^2 + \mu_k^2\xi_k^2\|F(z_k)\|^2. \end{aligned} \tag{20}$$

where we used that $\varepsilon_k^2 = (\mu_k^2\|\xi_k F(z_k)\|^2)/2$ in the last equality. It is easy to see that (20) also holds when $x_{k+1} = \tilde{y}^I_C(w_k)$. By the monotonicity of the mapping F and the fact that $x_* \in C^*$, we get

$$\begin{aligned} \langle F(z_k),\, x_k - z_k\rangle &= \langle F(x_*),\, z_k - x_*\rangle + \langle F(z_k),\, x_k - z_k\rangle \\ &\leq \langle F(z_k),\, z_k - x_*\rangle + \langle F(z_k),\, x_k - z_k\rangle \\ &= \langle F(z_k),\, x_k - x_*\rangle. \end{aligned} \tag{21}$$

By combining (20) and (21), we find

$$\begin{aligned} \|x_{k+1} - x_*\|^2 &\leq \|x_k - x_*\|^2 - 2\xi_k\langle F(z_k),\, x_k - z_k\rangle + \xi_k^2\|F(z_k)\|^2 + \mu_k^2\xi_k^2\|F(z_k)\|^2 \\ &\leq \|x_k - x_*\|^2 + (\mu_k^2 - 1)\frac{\langle F(z_k),\, x_k - z_k\rangle^2}{\|F(z_k)\|^2} \\ &\leq \|x_k - x_*\|^2 + (\bar{\mu}^2 - 1)\sigma^2\frac{\|x_k - z_k\|^4}{\|F(z_k)\|^2}, \end{aligned} \tag{22}$$

where the second inequality follows from the definition of ξ_k and the last inequality is due to the fact that $\mu_k \leq \bar{\mu}$ and (18). By (22) and the fact that $\bar{\mu} < 1$, we have

$$\|x_{k+1} - x_*\|^2 \leq \|x_k - x_*\|^2, \quad k \geq 0, \tag{23}$$

which implies that the sequence $\{x_k\}$ is bounded. It follows from the Cauchy-Schwartz inequality, the monotonicity of F and (18) that

$$\|F(x_k)\| \geq \frac{\langle F(x_k),\ x_k - z_k\rangle}{\|x_k - z_k\|} \geq \frac{\langle F(z_k),\ x_k - z_k\rangle}{\|x_k - z_k\|} \geq \sigma\|x_k - z_k\|.$$

Therefore, by the continuity of F and the boundedness of $\{x_k\}$, we have that $\{z_k\}$ is also bounded. Since $\{z_k\}$ is bounded and F is continuous on $\mathbb{R}^n$, there exists a constant $M > 0$ such that $\|F(z_k)\| \leq M$ for all $k \geq 0$, which, combined with (22), yields

$$\frac{(1-\bar{\mu}^2)\sigma^2}{M^2} \sum_{k=0}^{\infty} \|x_k - z_k\|^4 \leq \sum_{k=0}^{\infty} \left(\|x_k - x_*\|^2 - \|x_{k+1} - x_*\|^2\right) < \infty,$$

which implies $\lim_{k\to\infty} \|x_k - z_k\| = 0$. □

We are now ready to establish the global convergence of Framework 1.

Theorem 3. *The sequence $\{x_k\}$ generated by Framework 1 converges to a solution of* (1).

Proof. Since $z_k = x_k + \alpha_k d_k$, from Lemma 2, it holds that

$$\lim_{k\to\infty} \alpha_k \|d_k\| = \lim_{k\to\infty} \|x_k - z_k\| = 0. \tag{24}$$

We also have, from Lemma 2, that $\{x_k\}$ is bounded and therefore $\{F(x_k)\}$ is bounded as well. Thus, it follows from the second inequality in (12) that $\{d_k\}$ is bounded. Consider now two different cases: (i) $\liminf_{k\to\infty} \|d_k\| = 0$ or (ii) $\liminf_{k\to\infty} \|d_k\| > 0$.

Case (i). Note that (11) and the Cauchy-Schwartz inequality imply that $\eta_1\|F(x_k)\| \leq \|d_k\|$. Hence, since $\liminf_{k\to\infty} \|d_k\| = 0$, it follows that $\liminf_{k\to\infty} \|F(x_k)\| = 0$. Since F is continuous, we have that the sequence $\{x_k\}$ has some cluster point $\bar{x}$ such that $F(\bar{x}) = 0$. Replacing x_* by $\bar{x}$ in (23), we obtain

$$\|x_{k+1} - \bar{x}\|^2 \leq \|x_k - \bar{x}\|^2,$$

which implies that $\{\|x_k - \bar{x}\|\}$ converges. Therefore, we can conclude that the whole sequence $\{x_k\}$ converges to $\bar{x}$, a solution of (1).

Case (ii). Since $\liminf_{k\to\infty} \|d_k\| > 0$, it follows from (24) that there exists a subsequence of indices $K \subset \mathbb{N}$ such that $\lim_{k\to\infty} \alpha_k = 0$, where $k \in K$. By (13), we have

$$-\langle F(x_k + \gamma^{m_k - 1} d_k),\ d_k\rangle < \sigma\gamma^{m_k-1}\|d_k\|^2.$$

Since $\{x_k\}$ and $\{d_k\}$ are bounded, we can choose a subsequence $K_1 \subset K$ such that $\{(x_k, d_k)\} \xrightarrow{K_1} (\bar{x}, \bar{d})$. Hence, using the continuity of F and taking the limit in the last inequality as $k \to \infty$ with $k \in K_1$, we have

$$-\langle F(\bar{x}),\ \bar{d}\rangle \leq 0. \tag{25}$$

On the other hand, by taking the limit in (11) as $k \to \infty$ with $k \in K_1$, we obtain

$$-\langle F(\bar{x}),\ \bar{d}\rangle \geq \delta \|F(\bar{x})\|^2 > 0,$$

where the last inequality is due to the inequality in (12) and the fact that $\liminf_{k\to\infty} \|d_k\| > 0$. Thus, the last inequality contradicts (25). Hence, $\liminf_{k\to\infty} \|d_k\| = 0$. Therefore, using a similar argument as in the first case, we conclude that the whole sequence $\{x_k\}$ converges to a solution of (1). This completes the proof. $\square$

4 Some instances of Framework 1

This section presents some examples of search directions d_k that satisfy the safeguard conditions (11) and (12) and, as a consequence, some instances of Framework 1. These instances of methods allow approximate projections onto $C \cap H_k$, which can be advantageous when the exact projections are difficult (where the projection cannot be easily performed).

Let us begin by presenting inexact versions of two well-known methods.

1) Steepest descent-based method with approximate projections (SDM-AP). This method corresponds to Framework 1 with the direction d_k in the Step 2 defined by $d_k = -F(x_k)$, for every $k \geq 0$. It is easy to see that this choice of d_k satisfies the conditions (11) and (12) with $\eta_1 = 1$ and $\eta_2 \geq 1$. Therefore, from Theorem 3, it holds that the sequence $\{x_k\}$ generated by SDM-AP converges to a solution of (1).

2) Newton method with approximate projections (NM-AP). Assume that F is continuously differentiable. By taking d_k in the Step 2 of Framework 1 as $d_k = -B(x_k)^{-1}F(x_k)$ for every $k \geq 0$, where $B(x_k)$ is a positive definite matrix, we obtain a variant of the Newton method proposed in [25] with approximate projections. Note that $B(x_k)$ may be the Jacobian of F at x_k or an approximation of it. Assuming that there exist constants $0 < a \leq b$ such that $aI \prec B(x_k) \prec bI$, for every k, then d_k satisfies (11) and (12) with $\eta_1 = 1/b$ and $\eta_2 = 1/a$. Indeed, since $B_k d_k = -F(x_k)$, we obtain

$$\langle d_k, F(x_k)\rangle = \langle -B_k^{-1}F(x_k), F(x_k)\rangle = -\|F(x_k)\|^2_{B_k^{-1}} \leq -\left(\frac{1}{b}\right)\|F(x_k)\|^2$$

and

$$a\|d_k\|^2 \leq \|d_k\|^2_{B_k} = \langle B_k d_k,\ d_k\rangle = -\langle F(x_k),\ d_k\rangle \leq \|F(x_k)\|\|d_k\|,$$

which proves the statement. Therefore, since this method can be seen as an instance of Framework 1, we trivially have, from Theorem 3, that the sequence $\{x_k\}$ generated by it converges to a solution of (1).

We next present two examples of methods, in the spirit of the one in example 2, for the nonsmooth case. Recall that $F : \mathbb{R}^n \to \mathbb{R}^n$ is said to be τ-strongly monotone if there is a constant $\tau > 0$ such that $\langle x - y, F(x) - F(y)\rangle \geq \tau\|x - y\|^2$, for all $x, y \in \mathbb{R}^n$, and the function $F : \mathbb{R}^n \to \mathbb{R}^n$ is said to be $\mathcal{L}$-Lipschitz continuous if there is a constant $\mathcal{L} > 0$ such that $\|F(x) - F(y)\| \leq \mathcal{L}\|x - y\|$, for all $x, y \in \mathbb{R}^n$.

3) Spectral gradient-like methods with approximate projections (SGM-AP). Consider $d_k = -\lambda_k F(x_k)$ for every $k \geq 0$, where λ_k is the spectral coefficient which is related to the Barzilai-Borwein choice of the step-size [2]. Firstly, let us discuss some existing choices of λ_k.

3.1) In [13], λ_k is defined by

$$\lambda_k = \frac{\langle s_k, s_k \rangle}{\langle s_k, u_k \rangle}, \tag{26}$$

where $s_k := x_k - x_{k-1}$ and $u_k := F(x_k) - F(x_{k-1})$. Under the assumption that F is τ-strongly monotone and $\mathcal{L}$-Lipschitz continuous, we have that $d_k = -\lambda_k F(x_k)$ satisfies (11) and (12) with $\eta_1 = 1/\mathcal{L}$ and $\eta_2 = 1/\tau$. Indeed, using that F is τ-strongly monotone, we have

$$\langle s_k, u_k \rangle = \langle x_k - x_{k-1}, F(x_k) - F(x_{k-1}) \rangle \geq \tau \langle x_k - x_{k-1}, x_k - x_{k-1} \rangle = \tau \langle s_k, s_k \rangle > 0,$$

for some $\tau > 0$, and therefore, $\lambda_k \leq 1/\tau$. Now, using the Cauchy-Schwarz inequality and that F is $\mathcal{L}$-Lipschitz continuous, we obtain

$$\langle s_k, u_k \rangle = \langle s_k, F(x_k) - F(x_{k-1}) \rangle \leq \|F(x_k) - F(x_{k-1})\| \|s_k\| \leq \mathcal{L} \langle s_k, s_k \rangle,$$

which implies $1/\mathcal{L} \leq \lambda_k$. Thus, $1/\mathcal{L} \leq \lambda_k \leq 1/\tau$ and, as a consequence, the statement trivially follows from the fact that $d_k = -\lambda_k F(x_k)$.

3.2) In [27, 28], the coefficient λ_k is as in (26) with $s_k := x_k - x_{k-1}$ and $u_k := F(x_k) - F(x_{k-1}) + r s_k$, where $r > 0$ is a given scalar. Using that F is monotone, we have

$$\begin{aligned} \langle s_k, u_k \rangle &= \langle s_k, F(x_k) - F(x_{k-1}) + r s_k \rangle \\ &= \langle x_k - x_{k-1}, F(x_k) - F(x_{k-1}) \rangle + r \langle s_k, s_k \rangle \\ &\geq r \langle s_k, s_k \rangle > 0, \end{aligned}$$

which implies that $\lambda_k \leq 1/r$. Now, by assuming that F is $\mathcal{L}$-Lipschitz continuous, we obtain

$$\begin{aligned} \langle s_k, u_k \rangle &= \langle s_k, F(x_k) - F(x_{k-1}) + r s_k \rangle \\ &= \langle x_k - x_{k-1}, F(x_k) - F(x_{k-1}) \rangle + r \langle s_k, s_k \rangle \\ &\leq (\mathcal{L} + r) \langle s_k, s_k \rangle, \end{aligned}$$

which yields $1/(\mathcal{L} + r) \leq \lambda_k$. Therefore, as $1/(\mathcal{L} + r) \leq \lambda_k \leq 1/r$, we can conclude, from the fact that $d_k = -\lambda_k F(x_k)$, that d_k satisfies the conditions (11) and (12) with $\eta_1 = 1/(\mathcal{L} + r)$ and $\eta_2 = 1/r$.

3.3) In the works [1, 19] the coefficient λ_k is a convex combination of the default spectral coefficient in [2] and the positive spectral coefficient in [6]. More specifically, λ_k is defined by

$$\lambda_k = (1 - t)\theta_k^* + t\theta_k^{**},$$

where $t \in [0,1]$, $\theta_k^* = \|s_k\|^2/\langle u_k, s_k\rangle$, $\theta_k^{**} = \|s_k\|/\|u_k\|$, $s_k := x_k - x_{k-1}$, $u_k := F(x_k) - F(x_{k-1}) + rs_k$ and $r > 0$. In [1, Lemma 2], it was shown that if F is $\mathcal{L}$-Lipschitz continuous, then $d_k = -\lambda_k F(x_k)$ satisfies (11) and (12) with $\eta_1 = \max\{1, 1/(\mathcal{L}+r)\}$ and $\eta_2 = \min\{1, 1/r\}$.

Since the search directions in examples 3.1, 3.2 and 3.3 satisfy (11) and (12) for specific values of η_1 and η_2, we can conclude, from Theorem 3, that the SGM-AP (i.e., Framework 1 with the above three choice of search directions) converges to a solution of (1).

4) Limited memory BFGS method with approximate projections (L-BFGS-AP). Consider the L-BFGS direction d_k proposed in [29], which is obtained by solving the system $B_k d_k = -F(x_k)$, where the sequence $\{B_k\}$ is given by $B_0 = I$ and B_{k+1} is computed by the following modified L-BFGS update process: let $m > 0$ be given and set $\tilde{m} = \min\{k+1, m\}$, $B_k^{(0)} = B_0 = I$. Choose a set of increasing integers $L_k = \{j_0, \ldots, j_{\tilde{m}-1}\} \subset \{0, \ldots, k\}$. Update B_{k+1} by using the pairs $\{y_{j_l}, s_{j_l}\}_{l=0}^{\tilde{m}-1}$, i.e., for $l = 0, \ldots, \tilde{m}-1$,

$$B_{k+1} := B_{k+1}^{(l+1)} = \begin{cases} B_k^{(l)} - \frac{B_k^{(l)} s_{j_l} s_{j_l}^T B_k^{(l)}}{s_{j_l}^T B_k^{(l)} s_{j_l}} + \frac{y_{j_l} y_{j_l}^T}{y_{j_l}^T s_{j_l}}, & if \ \frac{y_{j_l}^T s_{j_l}}{\|s_{j_l}\|^2} \geq \varepsilon, \\ B_k^{(l)}, & \text{otherwise}, \end{cases}$$

where $s_k := x_{k+1} - x_k$ and $y_k := F(x_{k+1}) - F(x_k)$. If d_k in the Step 2 of Framework 1 is defined as above, we obtain an L-BFGS method with approximate projections. Under the assumption that F is $\mathcal{L}$-Lipschitz continuous, it was proven in [29] that B_k and B_k^{-1} are bounded for all $k \geq 0$, i.e., $\{B_k\} \subset \mathbb{B}$. Since $B_k d_k = -F(x_k)$ and using (3), we obtain

$$\langle d_k, F(x_k)\rangle = \langle -B_k^{-1} F(x_k), F(x_k)\rangle = -\|F(x_k)\|_{B_k^{-1}}^2 \leq -\left(\frac{1}{\|B_k\|}\right) \|F(x_k)\|^2,$$

which yields $\langle d_k, F(x_k)\rangle \leq -1/L\|F(x_k)\|^2$. Now, from the Cauchy-Schwarz inequality, we have

$$\|d_k\|_{B_k}^2 = \langle B_k d_k,\, d_k\rangle = -\langle F(x_k),\, d_k\rangle \leq \|F(x_k)\|\|d_k\|,$$

which, combined with (3) and $\|B_k^{-1}\| \leq L$, yields $(1/L)\|d_k\| \leq \|F(x_k)\|$. Thus, d_k satisfies (11) and (12) with $\eta_1 = 1/L$ and $\eta_2 = L$. Therefore, we conclude, from Theorem 3, that the sequence $\{x_k\}$ generated by L-BFGS-AP (i.e., Framework 1 with the above choice of search direction) converges to a solution of (1).

We end this section by proposing a new convergent method for solving (1), which is an instance of Framework 1. This method is inspired by [23, Algorithm 2.1] for solving variational inequalities. In the context that the projection operator is computationally expensive, the latter algorithm was devised in order to minimize the total number of performed projection operations. Let us now present our extension of [23, Algorithm 2.1] to the convex-constrained monotone nonlinear equations setting.

5) Modified Newton-like method with approximate projections (MNM-AP). Consider the direction d_k defined as follows: let $\eta > 0$, $\bar{\theta} \in [0, \eta)$ and $\{\theta_k\} \subset [0, \bar{\theta}]$ be given. Let $B_k \subset \mathbb{B}$, and set $w_k^1 := B_k x_k - F(x_k)$ and

$\varepsilon_k^1 := \theta_k^2 \|F(x_k)\|^2$. Compute s_k^1 in $\mathbb{R}^n$ such that

$$s_k^1 = \tilde{y}_C^{B_k}(w_k^1) - x_k, \tag{27}$$

where $\tilde{y}_C^{B_k}(w_k^1)$ is an ε_k^1–approximate solution of the problem (4). If $\eta \|F(x_k)\| \leq \|s_k^1\|_{B_k}$, then $d_k := s_k^1$. Otherwise, compute s_k^2 in $\mathbb{R}^n$ such that

$$F(x_k) + B_k s_k^2 = 0, \tag{28}$$

and set $d_k := s_k^2$. Note that the matrix B_k can be taken as those in the examples 3 and 4. We will now prove that the d_k described above satisfies (11) and (12), for all $k \geq 0$. If $\eta \|F(x_k)\| \leq \|s_k^1\|_{B_k}$, then $d_k = \tilde{y}_C^{B_k}(w_k^1) - x_k$. By (5) with $B = B_k$, $w = w_k^1$ and $y = x_k$, we have

$$\begin{aligned}\theta_k^2 \|F(x_k)\|^2 \geq & \langle B_k(x_k - \tilde{y}_C^{B_k}(w_k^1)) - F(x_k), x_k - \tilde{y}_C^{B_k}(w_k^1)\rangle \\ = & \|\tilde{y}_C^{B_k}(w_k^1) - x_k\|_{B_k}^2 - \langle F(x_k),\, x_k - \tilde{y}_C^{B_k}(w_k^1)\rangle,\ \forall\, k \geq 0,\end{aligned}$$

which, combined with the definition of d_k, yields

$$-\langle F(x_k), d_k\rangle + \theta_k^2 \|F(x_k)\|^2 \geq \|d_k\|_{B_k}^2 \geq \|F(x_k)\|^2 \eta^2,\ \forall\, k \geq 0, \tag{29}$$

or, equivalently,

$$-\langle F(x_k), d_k\rangle \geq \|F(x_k)\|^2 (\eta^2 - \theta_k^2),\ \forall\, k \geq 0.$$

Therefore, since $\theta_k \leq \bar{\theta}$ for all $k \geq 0$ and $\bar{\theta} \in [0, \eta)$, we have

$$\langle F(x_k), d_k\rangle \leq -(\eta^2 - \bar{\theta}^2)\|F(x_k)\|^2.$$

Hence, (11) holds with $\eta_1 = (\eta^2 - \bar{\theta}^2)$. From (29) and using the Cauchy-Schwarz inequality, we have

$$\|d_k\|_{B_k}^2 \leq \theta_k^2 \|F(x_k)\|^2 - \langle B_k^{-1} F(x_k), d_k\rangle_{B_k} \leq \theta_k^2 \|F(x_k)\|^2 + \|B_k^{-1} F(x_k)\|_{B_k} \|d_k\|_{B_k},$$

which, after some algebraic manipulations, yields

$$\|d_k\|_{B_k}^2 \leq \theta_k^2 \|F(x_k)\|^2 + \frac{\|B_k^{-1} F(x_k)\|_{B_k}^2}{2} + \frac{\|d_k\|_{B_k}^2}{2}.$$

Using the definition of scalar product $\langle \cdot, \cdot \rangle_B = \langle \cdot, B \cdot \rangle$ and (3), we obtain

$$\frac{\|d_k\|_{B_k}^2}{2} \leq \theta_k^2 \|F(x_k)\|^2 + \frac{\|F(x_k)\|_{B_k^{-1}}^2}{2} \leq \theta_k^2 \|F(x_k)\|^2 + \frac{\|F(x_k)\|^2 \|B_k^{-1}\|}{2} = \left(\theta_k^2 + \frac{\|B_k^{-1}\|}{2}\right) \|F(x_k)\|^2,$$

which implies that

$$\|d_k\|_{B_k}^2 \leq \left(2\theta_k^2 + \|B_k^{-1}\|\right) \|F(x_k)\|^2.$$

Therefore, by (3), $\|B_k^{-1}\| \leq L$ and $\theta_k \leq \bar{\theta}$ for all $k \geq 0$, we have

$$\|d_k\|^2 \leq L\left(2\bar{\theta}^2 + L\right)\|F(x_k)\|^2,$$

and hence (12) holds with $\eta_2 = \sqrt{L\left(2\bar{\theta}^2 + L\right)}$. On the other hand, if $d_k := s_k^2$, the proof is similar to the one in example 4. Therefore, we conclude, from Theorem 3, that the sequence $\{x_k\}$ generated by the MNM-AP (i.e., Framework 1 with the above choice of search direction) converges to a solution of (1).

5 Numerical experiments

This section summarizes the numerical experiments carried out to verify the efficiency of the instances of Framework 1. Numerical experiments are divided into two sections. In Section 5.1, the methods are tested for a group of convex-constrained monotone nonlinear equations, whereas in Section 5.2 they are tested for solving the system of constrained absolute value equations (CAVE). The computational results are obtained using MATLAB R2018a on a 2.4GHz Intel(R) i5 with 8GB of RAM and Windows 10 as ultimate system.

5.1 Monotone nonlinear equations with polyhedral constraints

In this subsection, our aim is to illustrate the behavior of the methods to solve 52 monotone nonlinear equations with polyhedral constraints; see Table 1. Some of these problems are originally unconstrained ones, for which constrains were added. In Pb11, the matrix $A \in \mathbb{R}^{10\times n}$ is randomly generated so that a solution of the problem 11 belongs to the feasible set.

The tolerance in the stopping criterion $\|F(x_k)\| < \varepsilon$ was set to $\varepsilon = 10^{-6}$. If the stopping criterion is not satisfied, the method stops when a maximum of 500 iterations has been performed. In this first group of test problems, we set $\sigma = 10^{-4}$, $\gamma = 1/2$ and $\mu_k = \bar{\mu} = 0.25$, for every k, in all algorithms. Moreover, the ε_k–approximate solution in (14) was computed by the conditional gradient method, which stopped when either the stopping criterion is satisfied or a maximum of 300 iterations is performed. In order to avoid an excessive number of inner iterations, input ε_k was replaced by $\max\{\mu_k^2\|\xi_k F(z_k)\|^2, 10^{-2}\}$. Linear optimization subproblems in the conditional gradient method (see (8)) were solved via the MATLAB command `linprog`. We denote by SGM-AP1, SGM-AP2 and SGM-AP3, the method SGM-AP, with the coefficient λ_k given in examples 3.1, 3.2 and 3.3, respectively. In SGM-AP2, we set $r = 0.01$, whereas, in SGM-AP3, we set $t = 1/(exp(k+1)^{k+1})$ and $r = 1/(k+1)^2$. In the L-BFGS-AP, we used $m = 1$. Finally, we set $\eta = 0.5$, $\theta_k = \bar{\theta} = 0.25$ in the MNM-AP.

We consider 4 different starting points (following the suggestions from the original sources of the problems) for each problem of Table 1: For problem 1, $x_1 = (0.1,\ldots,0.1)$, $x_2 = (1,\ldots,1)$, $x_3 = ((n-1)/n, 0.1,\ldots,0.1,(n-1)/n)$ and $x_4 = (-1,\ldots,-1)$. For problem 2, $x_1 = (0.1,\ldots,0.1)$, $x_2 = (1,\ldots,1)$, $x_3 = (0,\ldots,0)$ and $x_4 = (-1,\ldots,-1)$. For problem 3, $x_1 = (10,0,\ldots,0)$, $x_2 = (9,0,\ldots,0)$, $x_3 = (3,0,3,0,3)$ and $x_4 = (0,2,2,2,2)$. For problem 5, $x_1 = (0,\ldots,0)$, $x_2 = (3,0,0,0)$, $x_3 = (1,1,1,0)$ and $x_4 = (0,1,1,1)$. For problems 14 and 16, $x_1 = (-1,\ldots,-1)$, $x_2 = (-0.1,\ldots,-0.1)$, $x_3 = (-1/2,-1/2^2,\ldots,-1/2^n)$ and

$x_4 = (-1, -1/2, \ldots, -1/n)$. For problem 17, $x_1 = ((n-1)/n, 0.1, \ldots, 0.1, (n-1)/n)$, $x_2 = (0.1, \ldots, 0.1)$, $x_3 = (1/2, 1/2^2, \ldots, 1/2^n)$ and $x_4 = (1, 1/2, \ldots, 1/n)$. For problems 4, 6 to 13, 15 and 18, $x_1 = (1, \ldots, 1)$, $x_2 = (0.1, \ldots, 0.1)$, $x_3 = (1/2, 1/2^2, \ldots, 1/2^n)$ and $x_4 = (1, 1/2, \ldots, 1/n)$. Figures 1 and 2 report the numerical results of SDM-AP, SGM-AP1, SGM-AP2, SGM-AP3, L-BFGS-AP and MNM-AP for solving the 52 problems using performance profiles [8]. We adopted the CPU time as performance measurement. Recall that in the performance profile, efficiency and robustness can be accessed on the left and right extremes of the graphic, respectively. We consider that a method is the most efficient if its runtime does not exceed in 5% the CPU time of the fastest one.

From Figures 1 and 2, we can see that all the variations of the SGM-AP achieved a better performance (in terms of efficiency and robustness) compared to L-BFGS-AP and MNM-AP. In the group of SGM-AP variants, SGM-AP1 and SGM-AP2 were better than the others.

Problem	Ref.	n	Set C
Pb 1	[27, Problem 1]	1000/5000/10000	$[-1, n]$ and $\sum_{i=1}^n x_i \leq n$
Pb 2	[27, Problem 2]	1000/5000/10000	$[-1, n]$ and $\sum_{i=1}^n x_i \leq n$
Pb 3	[25, Problem 2]	5/5/5	$[0, n]$ and $\sum_{i=1}^n x_i \leq n$
Pb 4	[25, Problem 3]	10/10/10	$[0, n]$ and $\sum_{i=1}^n x_i \leq n$
Pb 5	[25, Problem 4]	4	$[-1, n]$ and $\sum_{i=1}^n x_i \leq 3$
Pb 6	[1, Problem 1]	1000/5000/10000	$[-1, 2]$ and $\sum_{i=1}^n x_i \leq n$
Pb 7	[1, Problem 2]	1000/5000/10000	$[-1, 2]$ and $\sum_{i=1}^n x_i \leq n$
Pb 8	[1, Problem 3]	1000/5000/10000	$[0, n]$ and $\sum_{i=1}^n x_i \leq n$
Pb 9	[1, Problem 5]	1000/5000/10000	$[-1, 7]$ and $\sum_{i=1}^n x_i \leq 1.1 \cdot n$
Pb 10	[1, Problem 6]	1000/5000/10000	$[0, e]$ and $\sum_{i=1}^n x_i \leq e \cdot n$
Pb 11	[12, Problem 1]	1000/5000/10000	$[-1, 2]$; $Ax \leq b$, where $A \in \mathbb{R}^{10 \times n}$ and $b = (n, \ldots, n) \in \mathbb{R}^{10}$
Pb 12	[12, Problem 4]	1000/5000/10000	$[-1, n]$ and $\sum_{i=1}^n x_i \leq 20 \cdot n$
Pb 13	[12, Problem 7]	1000/5000/10000	$[-1, n]$ and $\sum_{i=1}^n x_i \leq n$
Pb 14	[12, Problem 8]	1000/5000/10000	$[-n, 1]$ and $\sum_{i=1}^n x_i \leq n$
Pb 15	[12, Problem 9]	1000/5000/10000	$[-n, n]$ and $\sum_{i=1}^n x_i \leq n$
Pb 16	[14, Problem 2]	1000/5000/10000	$[-n, 1]$ and $\sum_{i=1}^n x_i \leq 1$
Pb 17	[14, Problem 3]	1000/5000/10000	$[-1, n]$ and $\sum_{i=1}^n x_i \leq n$
Pb 18	[14, Problem 7]	1000/5000/10000	$[-1, n]$ and $\sum_{i=1}^n x_i \leq n$

Table 1: Test problems

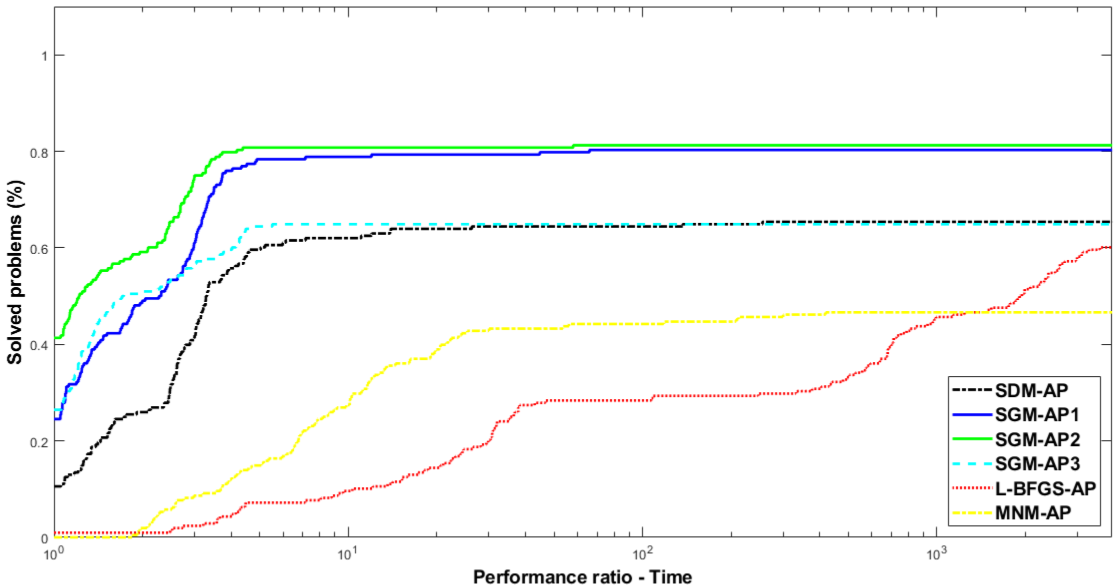

Figure 1: Performance of SDM-AP, SGM-AP1, SGM-AP2, SGM-AP3, L-BFGS-AP and MNM-AP with x_{k+1} as in (14)

5.2 Absolute value equations with polyhedral constraints

In this subsection, we consider the problem of finding a solution of the CAVE problem:

$$\text{find } x \in C \text{ such that } Ax - |x| = b, \tag{30}$$

where $C := \{x \in \mathbb{R}^n; \sum_{i=1}^n x_i \leq d, x_i \geq -1, i = 1, \ldots, n\}$, $A \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^n \equiv \mathbb{R}^{n \times 1}$, and $|x|$ denotes the vector whose i-th component is equal to $|x_i|$. The problem (30) draws attention for its simple formulation when compared to its equivalent linear complementarity problem (LCP) (see [3, 4, 17]) which in turn includes linear programs, quadratic programs, bimatrix games and other problems. Hence, interesting algorithms related to Newton-type methods to solve (30) have been developed; see, for example, [5, 16] and [20] for the unconstrained and constrained case, respectively.

Under the assumption that $\|A^{-1}\| \leq 1$, it was proven in [17, Proposition 4] that the problem (30), with $C = \mathbb{R}^n$, is uniquely solvable for any b. Now, if A is symmetric positive definite, then $F(x) = Ax - |x| - b$ is

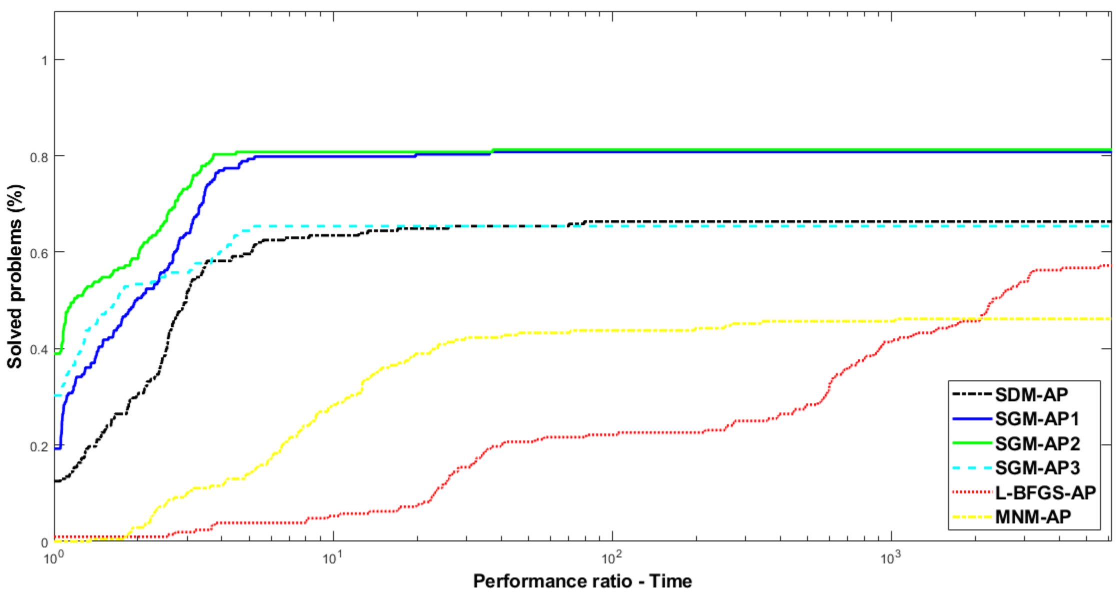

Figure 2: Performance of SDM-AP, SGM-AP1, SGM-AP2, SGM-AP3, L-BFGS-AP and MNM-AP with x_{k+1} as in (16)

monotone. In fact, for all $x, y \in \mathbb{R}^n$, we have

$$\begin{aligned}\langle F(x)-F(y), x-y\rangle &= \langle Ax-|x|-Ay+|y|, x-y\rangle = \|x-y\|_A^2+\langle |y|-|x|, x-y\rangle \\ &\geq \|x-y\|^2\frac{1}{\|A^{-1}\|}+\langle |y|-|x|, x-y\rangle \geq \|x-y\|^2+\langle |y|-|x|, x-y\rangle.\end{aligned} \tag{31}$$

where in the second equality we use that $\langle\cdot,\cdot\rangle_B = \langle\cdot, B\cdot\rangle$, (3) and $\|A^{-1}\| \leq 1$. Note that $|x|$ can be written as $|x| = P_{\mathbb{R}^n_+}(x)+P_{\mathbb{R}^n_+}(-x)$. So, from (31), the Cauchy-Schwarz inequality and the fact that $P_{\mathbb{R}^n_+}(\cdot)$ is monotone and nonexpansive, we obtain

$$\begin{aligned}\langle F(x)-F(y), x-y\rangle \geq &\|x-y\|^2+\langle P_{\mathbb{R}^n_+}(y)+P_{\mathbb{R}^n_+}(-y)-P_{\mathbb{R}^n_+}(x)-P_{\mathbb{R}^n_+}(-x), x-y\rangle \\ =&\|x-y\|^2-\langle P_{\mathbb{R}^n_+}(x)-P_{\mathbb{R}^n_+}(y), x-y\rangle+\langle P_{\mathbb{R}^n_+}(-y)-P_{\mathbb{R}^n_+}(-x), x-y)\rangle \\ \geq&\|x-y\|^2-\|P_{\mathbb{R}^n_+}(x)-P_{\mathbb{R}^n_+}(y)\|\|x-y\| \\ \geq&\|x-y\|^2-\|x-y\|^2 = 0,\end{aligned}$$

which proves the statement. In our implementation, we used the Matlab routine `sprandsym` to construct matrix A randomly, which generates a symmetric positive definite sparse matrix with predefined dimension, density and singular values. For this process, the density of matrix A was set to 0.003 and the vector of singular values was randomly generated from a uniform distribution on $(0,1)$. In this case, as the vector of singular values (rc) is a vector of length n, then A has eigenvalues rc. Thus, if rc is a positive (nonnegative) vector, then A is a positive (nonnegative) definite matrix. We chose a random solution x_* from a uniform distribution on $(0.1,10)$ and computed $b = Ax_* - |x_*|$ and $d = \sum_{i=1}^{n}(x_*)_i$, where $(x_*)_i$ denotes the i-th component of the vector x_*. The initial points were defined as $x_0 = (0,\ldots,0,d/2,0,\ldots,0,d/2,0,\ldots,0) \in \mathbb{R}^n$, where the two positions of $d/2$ were generated randomly on the set $\{1,2,\ldots,n\}$.

For the CAVE problem, we consider only the SGM-AP2 since it was the best method in our first class of experiments described in Section 5.1. For a comparative purpose, we also run the inexact Newton method with feasible inexact projections (INM-InexP) of [20]. INM-InexP is an algorithm designed for solving smooth and nonsmooth equations subject to a set of constraints. We rescale the vector of singular values to ensure that the condition $\|A^{-1}\| \leq 1/3 < 1$ is fulfilled and, consequently, ensure the good definition of INM-InexP (see [16, Theorem 2] for more details). In INM-InexP, we set $\theta = \bar{\theta} = \bar{\mu} = 0.25$ and the other parameters were set as in [20]. For both algorithms, a failure was declared if the number of iterations was greater than 500. The procedure to obtain inexact projections used in the implementation of INM-InexP was also the CondG method and the procedure stopped when either the condition as in [20, Algorithm 1] was satisfied or a maximum of 10 iterations were performed. For our algorithms, the procedure stopped when either the stopping criterion was satisfied or a maximum of 10 iterations were performed.

As in Subsection 5.1, Figure 3 reports numerical results of algorithms using performance profiles. We generated 50 CAVEs of dimensions 1000, 5000 and 10000 and for each of them we test the algorithm for 5 different initial points. We see, from Figure 3, that the SGM-AP2-C (with x_{k+1} as in (16)) was the most robust one whereas INM-InexP was more efficient in terms of time saving than SGM-AP2-C and SGM-AP2-CH (with x_{k+1} as in (14)).

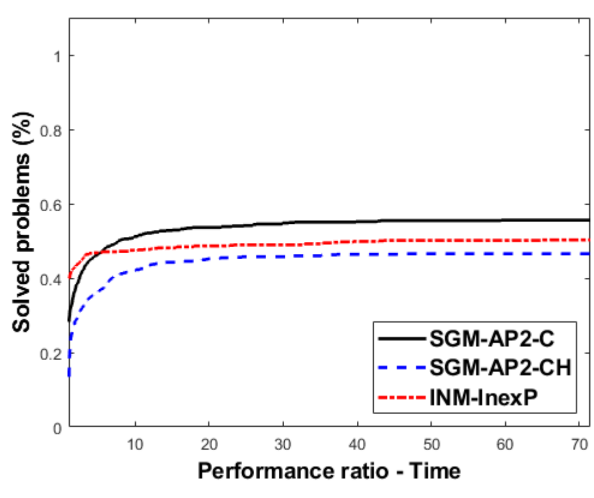

Figure 3: Performance of SGM-AP2-C, SGM-AP2-CH and INM-InexP

6 Final remarks

In this paper, we proposed and analyzed a framework with approximate projections for solving constrained monotone equations. Under mild assumptions, we proved that the sequence generated by the proposed framework converges to a solution of (1). Some examples of methods which fall into this framework were presented. Preliminary numerical experiments showed that the methods which fall into the framework performed well to solve constrained monotone nonlinear equations, and they are competitive in terms of robustness with the Inexact Newton method with feasible inexact projections in [20] for solving absolute value equations with polyhedral constraints.

References

[1] A. B. Abubakar, P. Kumam, and H. Mohammad. A note on the spectral gradient projection method for nonlinear monotone equations with applications. *Comp. Appl. Math.*, 39:1 – 35, 2020.

[2] J. Barzilai and J. M. Borwein. Two-point step size gradient methods. *IMA J. Numer. Anal.*, 8(1):141–148, 1988.

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

[4] R. W. Cottle, J.-S. Pang, and R. E. Stone. *The Linear Complementarity Problem*. Society for Industrial and Applied Mathematics, 2009.

[5] J. Y. B. Cruz, O. P. Ferreira, and L. F. Prudente. On the global convergence of the inexact semi-smooth Newton method for absolute value equation. *Comput. Optim. Appl.*, 65(1):93 – 108, 2016.

[6] Y.-H. Dai, M. Al-Baali, and X. Yang. A positive Barzilai-Borwein-like stepsize and an extension for symmetric linear systems. In *Numer. Anal. Optim.*, volume 134, pages 59 – 75. Springer, 2015.

[7] S. P. Dirkse and M. C. Ferris. MCPLIB: a collection of nonlinear mixed complementarity problems. *Optim. Meth. Softw.*, 5(4):319 – 345, 1995.

[8] E. D. Dolan and J. J. Moré. Benchmarking optimization software with performance profiles. *Math. Program.*, 91(2):201–213, 2002.

[9] J. C. Dunn. Convergence rates for conditional gradient sequences generated by implicit step length rules. *SIAM J. Control. Optim.*, 18(5):473–487, 1980.

[10] M. E. El-Hawary. *Optimal Power Flow: Solution Techniques, Requirements and Challenges: IEEE Tutorial Course*. Piscataway: IEEE, 1996.

[11] M. Frank and P. Wolfe. An algorithm for quadratic programming. *Nav. Res. Logist. Q.*, 3(1-2):95–110, 1956.

[12] W. La Cruz. A spectral algorithm for large-scale systems of nonlinear monotone equations. *Numer. Algor.*, 76(4):1109 – 1130, 2017.

[13] W. La Cruz and M. Raydan. Nonmonotone spectral methods for large-scale nonlinear systems. *Optim. Meth. Softw.*, 18:583 – 599, 2003.

[14] J. Liu and Y. Feng. A derivative-free iterative method for nonlinear monotone equations with convex constraints. *Numer. Algor.*, 82(1):245 – 262, 2019.

[15] J. Liu and S. Li. A projection method for convex constrained monotone nonlinear equations with applications. *Comput. Math. Appl.*, 70(10):2442 – 2453, 2015.

[16] O. L. Mangasarian. A generalized Newton method for absolute value equations. *Optim. Lett.*, 3(1):101 – 108, 2009.

[17] O. L. Mangasarian and R. R. Meyer. Absolute value equations. *Linear Algebra Appl.*, 419(2-3):359 – 367, 2006.

[18] K. Meintjes and A. P. Morgan. A methodology for solving chemical equilibrium systems. *Appl. Math. Comput.*, 22(4):333 – 361, 1987.

[19] H. Mohammad. A positive spectral gradient-like method for large-scale nonlinear monotone equations. *Bull. Comput. Appl. Math.*, 5(1):99 – 115, 2017.

[20] F. R. Oliveira and O. P. Ferreira. Inexact Newton method with feasible inexact projections for solving constrained smooth and nonsmooth equations. *Appl. Numer. Math.*, 156:63 – 76, 2020.

[21] Y. Ou and Y. Liu. Supermemory gradient methods for monotone nonlinear equations with convex constraints. *Comp. Appl. Math.*, 36(1):259 – 279, 2017.

[22] M. V. Solodov and B. F. Svaiter. *A Globally Convergent Inexact Newton Method for Systems of Monotone Equations*, pages 355 – 369. Springer US, Boston, MA, 1999.

[23] M. V. Solodov and B. F. Svaiter. A new projection method for variational inequality problems. *SIAM J. Control Optim.*, 37(3):765 – 776, 1999.

[24] C. Wang and Y. Wang. A superlinearly convergent projection method for constrained systems of nonlinear equations. *J. Glob. Optim.*, 44(2):283 – 296, 2009.

[25] W. Y. Wang, C. and C. Xu. A projection method for a system of nonlinear monotone equations with convex constraints. *Math. Meth. Oper. Res.*, 66(1):33 – 46, 2007.

[26] A. J. Wood and B. F. Wollenberg. Power generation operation and control, published by John Wiley and Sons. *New York, January*, 1996.

[27] Z. Yu, J. Lin, J. Sun, Y. Xiao, L. Liu, and Z. Li. Spectral gradient projection method for monotone nonlinear equations with convex constraints. *Appl. Numer. Math.*, 59(10):2416 – 2423, 2009.

[28] L. Zhang and W. Zhou. Spectral gradient projection method for solving nonlinear monotone equations. *J. Comput. Appl. Math.*, 196(2):478 – 484, 2006.

[29] W. Zhou and D. Li. Limited memory BFGS method for nonlinear monotone equations. *J. Comput. Math.*, 25(1):89 – 96, 2007.