Applied Mathematics and Computation 495 (2025) 129313

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

A modulus-based framework for weighted horizontal linear complementarity problems

Francesco Mezzadri

Department of Engineering "Enzo Ferrari", University of Modena and Reggio Emilia, via P. Vivarelli 10/1, building 26, I-41125, Modena, Italy

ARTICLE INFO

Keywords:
Weighted linear complementarity problem
Modulus-based methods
Non-smooth Newton's method
Matrix splitting methods

ABSTRACT

We develop a modulus-based framework to solve weighted horizontal linear complementarity problems (WHLCPs). First, we reformulate the WHLCP as a modulus-based system whose solution, in general, is not unique. We characterize the solutions by discussing their sign pattern and how they are linked to one another. After this analysis, we exploit the modulus-based formulation to develop new solution methods. In particular, we present a non-smooth Newton iteration and a matrix splitting method for solving WHLCPs. We prove the local convergence of both methods under some assumptions. Finally, we solve numerical experiments involving symmetric and non-symmetric matrices. In this context, we compare our approaches with a recently proposed smoothing Newton's method. The experiments include problems taken from the literature. We also provide numerical insights on relevant parts of the algorithms, such as convergence, attraction basin, and starting iterate.

1. Introduction

1.1. Formulation of the problem

We focus on the weighted horizontal linear complementarity problem (WHLCP), which consists in determining two vectors $\boldsymbol{\xi}, \boldsymbol{\eta} \in \mathbb{R}^{n\times n}$ such that

$$A\boldsymbol{\xi} - B\boldsymbol{\eta} = \boldsymbol{q}, \quad \boldsymbol{\xi} \geq \mathbf{0}, \quad \boldsymbol{\eta} \geq \mathbf{0}, \quad \boldsymbol{\xi} \circ \boldsymbol{\eta} = \boldsymbol{w}, \tag{1}$$

where $A, B \in \mathbb{R}^{n\times n}$, $\boldsymbol{q} \in \mathbb{R}^n$ is a known vector, and $\boldsymbol{w} \geq \mathbf{0}$ is a given vector of non-negative weights. In the above equation, $\circ$ denotes the Hadamard product. It is easy to notice that the problem (1) generalizes the "standard" horizontal linear complementarity problem (HLCP), which is obtained if $\boldsymbol{w} = \mathbf{0}$. Similarly, if A or B are the identity matrix, the problem (1) reduces to the weighted linear complementarity problem (WLCP). If, in addition, $\boldsymbol{w} = \mathbf{0}$, the WLCP reduces to a "standard" linear complementarity problem (LCP).

Weighted complementarity problems (WCPs) were introduced in [1] with the motivation that they significantly extend the standard linear complementarity formulation. For instance, using the terminology of [2] which we adopt in this paper, [1] showed that the Fisher market equilibrium problem can be formulated as a weighted mixed horizontal linear complementarity problem instead of a nonlinear complementarity problem. Similarly, [1] showed that the quadratic generalization of the linear programming and weighted centering problem (LPWC, [3]) can be formulated as a weighted mixed horizontal linear complementarity problem.

E-mail address: francesco.mezzadri@unimore.it.

https://doi.org/10.1016/j.amc.2025.129313
Received 3 August 2024; Received in revised form 19 December 2024; Accepted 18 January 2025

Available online 23 January 2025

Successive developments focused on the introduction of new solution methods and on the formulation of existence conditions for WCPs. In particular, [1] introduced two interior-point methods to solve monotone weighted mixed horizontal linear complementarity problems. Solvability conditions and an extension of the methods in [1] were subsequently provided in [4]. Many other Newton-based methods were then formulated, with several recent contributions focusing on full-Newton step interior point methods [5–7] and on variants of smoothing Newton methods [8–11]. As regards the existence and uniqueness of the solution to WHLCPs, a thorough analysis in the setting of Euclidean Jordan algebra was recently provided in [2]. Finally, it is worth mentioning that the concept of weighted complementarity appeared even in [12,13].

1.2. Aim and novelty

In this paper, we aim at introducing a modulus-based framework to solve the WHLCP (1). This is motivated by the success of the modulus-based formulation, where the complementarity problem is first reformulated as a nonlinear system which contains the absolute value of an auxiliary variable $\boldsymbol{x}$. Such equivalent system can then be solved by a numerical method. The first discussions on a modulus-based formulation of the LCP can be found in [14,15]. This idea was then developed into an actual solution method in [16] and especially in [17], where modulus-based matrix splitting methods for LCPs were introduced. Since then, modulus-based matrix splitting methods were extended to many generalizations of the LCP, such as HLCPs [18], (horizontal) nonlinear complementarity problems [19–21], implicit and quasi complementarity problems [22,23], and vertical linear and nonlinear complementarity problems [24,25]. Multisplittings were also used to achieve a parallel implementation of the algorithms [26–32]. More convoluted splitting techniques and non-splitting modulus-based methods were also introduced and studied for LCPs and their generalizations.

However, no modulus-based reformulation has yet been provided for the WLCP, nor for the WHLCP. We address this gap in the literature. On the one hand, developing such reformulation is relatively straightforward. On the other hand, the WHLCP modulus-based system is radically different from the one that is obtained from other complementarity problems. Indeed, we show that such modulus-based system contains fractional terms and that it has multiple solutions even when the solution to the WHLCP is unique. Notwithstanding this, we show that all the solutions of the modulus-based system map to the same, unique solution of the WHLCP. In this context, we characterize the solutions of the modulus-based system by discussing their number, their sign pattern, and how they are linked together. The above considerations can directly be applied to the WLCP, which is just a special instance of WHLCP.

This analysis, which is performed in Section 2, provides a modulus-based framework which can be leveraged to develop new solution methods. Thus, a further novelty of the paper consists in introducing a modulus-based non-smooth Newton's method and a modulus-based matrix splitting method for solving WHLCPs and WLCPs. We do this in Sections 3 and 4, respectively. We analyze the convergence of both methods, proving that they converge locally to a solution of the modulus-based system under some assumptions. The convergence analysis significantly differs from other modulus-based methods. For instance, the convergence analysis of modulus-based matrix splitting methods for WHLCPs and WLCPs is based on proving that there exists an attracting fixed point of the splitting iteration among all solutions of the modulus-based system. This requires further characterizing such solutions by several auxiliary lemmas. This analysis and the convergence results provide some guidance to the choice of the starting iterates and the parameters of the methods in order to more easily fall within the convergence basin of the proposed solution methods.

We then provide many numerical experiments based on test problems with symmetric and non-symmetric matrices. This is done in Section 5, where we detail a further technical novelty with respect to existing state-of-the-art approaches. Indeed, we show that our methods significantly outperform a recently proposed smoothing Newton's method [11], especially at large dimensions. Indeed, the method in [11] needs to solve a nonlinear system of order $(2n+1)$, while the modulus-based formulation uses a nonlinear system of order n. We also provide insights on the convergence and on other aspects of the procedures, which confirm the theoretical results provided in Sections 3 and 4. A summary of the conclusions is finally provided in Section 6.

1.3. Definitions and notation

Inequalities and absolute values involving vectors and matrices must be intended to apply component-wise. By $\langle A\rangle$ we denote the comparison matrix of A, which is the matrix of elements $|a_{ii}|$ on the main diagonal and $-|a_{ij}|$ off-diagonal, for $i,j=1,\dots,n$ with $i\neq j$. Moreover, we say that A is a Z-matrix if its off-diagonal entries are non-positive; if, in addition, $A^{-1}\geq 0$, then A is an M-matrix [33]. The matrix A is an H-matrix if its comparison matrix $\langle A\rangle$ is an M-matrix [33]. Finally, A is an H_+-matrix if it is an H-matrix with positive diagonal entries [34]. Through the paper, I denotes the identity matrix of suitable order.

Next, we recall some existing results on the existence and uniqueness of the solution to the WHLCP (1). In this regard, we first recall the following definitions.

Definition 1. [35] Let $\mathcal{M}$ be the set

$$\mathcal{M}=\{M_0,M_1,\dots,M_k\}, \tag{2}$$

where $M_0,\dots,M_k\in\mathbb{R}^{n\times n}$. A matrix $R\in\mathbb{R}^{n\times n}$ is *column representative* of $\mathcal{M}$ if

$$R_{\cdot j}\in\{(M_0)_{\cdot j},(M_1)_{\cdot j},\dots,(M_k)_{\cdot j}\},\quad j=1,\dots,n, \tag{3}$$

where $R_{\cdot j}$ and $(M_i)_{\cdot j}$ denote the j-th column of R and of M_i, $i=1,\dots,k$, respectively.

Definition 2. [35] The set $\mathcal{M}$ defined as in (2) has the column $\mathcal{W}$-property if the determinants of the column representative matrices of $\mathcal{M}$ are all positive or all negative.

It is known that, if the set $\{A, B\}$ has the column $\mathcal{W}$-property, then the HLCP of matrices A, B has a unique solution for any known term [35]. This has recently been generalized in [2]. In particular, [2, Theorem 3] connects the $\mathcal{W}$-property to the concept of P-pairs in the $\mathbb{R}^n$ setting. Thus, it is thereby proved that, if the set $\{A, B\}$ has the column $\mathcal{W}$-property, then the WHLCP (1) has a unique solution for any $\boldsymbol{q} \in \mathbb{R}^n$. This has the interesting implication that the solution to the WHLCP (1) is unique in many special interesting cases, such as when A, B form a column H_+-set or a column M-set (i.e., sets where all column-representative matrices are H_+-matrices or M-matrices, respectively [36]). Based on these results, in the following we assume that the set $\{A, B\}$ has the column $\mathcal{W}$-property and, thus, that the solution to (1) is unique.

We are also going to use the Bouligand and the Clarke sub-differentials. For the definition of such sub-differentials, the reader is referred, for instance, to [37,38]. In our analysis, we are particularly interested in the sub-differential of the absolute value function. By $\partial_B|\boldsymbol{x}|$ we denote the Bouligand's sub-differential of $|\boldsymbol{x}|$, which is the set of diagonal matrices of elements $\partial_B|\boldsymbol{x}|_{ii} = 1$ if $x_i > 0$, $\partial_B|\boldsymbol{x}|_{ii} = -1$ if $x_i < 0$, and $\partial_B|\boldsymbol{x}|_{ii} \in \{-1, 1\}$ if $x_i = 0$, $i = 1, \dots, n$. By $\partial|\boldsymbol{x}|$ we denote the Clarke's generalized Jacobian of $|\boldsymbol{x}|$, which is the convex hull of the Bouligand sub-differential. Thus, $\partial|\boldsymbol{x}|$ is the set of diagonal matrices of elements $\partial|\boldsymbol{x}|_{ii} = 1$ if $x_i > 0$, $\partial|\boldsymbol{x}|_{ii} = -1$ if $x_i < 0$, and $\partial_B|\boldsymbol{x}|_{ii} \in [-1, 1]$ if $x_i = 0$, $i = 1, \dots, n$. In this context, we will also use the concept of semi-smooth function [39].

Finally, we will sometimes refer to fixed points, i.e., points $\boldsymbol{x}^* \in \mathbb{R}^n$ that satisfy $\boldsymbol{x}^* = \boldsymbol{g}(\boldsymbol{x}^*)$ for a given mapping $\boldsymbol{g}(\boldsymbol{x}) : \mathbb{R}^n \to \mathbb{R}^n$. We recall that a fixed point $\boldsymbol{x}^* \in \mathbb{R}^n$ is *attracting* (or *a point of attraction*) for the iteration $\boldsymbol{x}^{(k+1)} = \boldsymbol{g}(\boldsymbol{x}^{(k)})$, with $k = 0, 1, \dots$, if there exists an open neighborhood $\mathcal{B}$ of $\boldsymbol{x}^*$ such that the iteration is well defined and converges to $\boldsymbol{x}^*$ for any starting iterate belonging to $\mathcal{B}$ [40, p. 144].

2. The modulus-based formulation of the WHLCP

In this section, we show that the WHLCP (1) can be written in a modulus-based form. First, let us deduce the following lemma.

Lemma 1. *If ξ, η are two real numbers such that*

$$\xi \geq 0, \quad \eta \geq 0, \quad \xi\eta = 0, \tag{4}$$

then they can be written as

$$\xi = \gamma(|x| + x) \quad \eta = \theta(|x| - x) \tag{5}$$

for some $x \in \mathbb{R}$ and with $\gamma, \theta \in \mathbb{R}^+$ arbitrary, real positive parameters.

If ξ, η are real numbers such that

$$\xi > 0, \quad \eta > 0, \quad \xi\eta = w \tag{6}$$

with $w \in \mathbb{R}^+$ arbitrary, real positive constant, then they can be written as

$$\xi = \gamma\Big(|x| + x + \frac{|x| - x}{4x^2} w\Big) \quad \eta = \theta\Big(|x| - x + \frac{|x| + x}{4x^2} w\Big) \tag{7}$$

for some $x \in \mathbb{R}\backslash\{0\}$ and with $\gamma, \theta \in \mathbb{R}^+$ arbitrary, real positive parameters that satisfy $\gamma\theta = 1$.

Proof. The first claim coincides with the standard modulus-based formulation. E.g., see [17].

As regards the second claim, let us analyze it for all x where (7) is defined. In this regard, let us consider separately the positive and the negative values of x. If $x > 0$, then (7) reduces to

$$\xi = \gamma\Big(|x| + x\Big) = 2\gamma x \quad \eta = \theta\Big(\frac{|x| + x}{4x^2} w\Big) = \frac{\theta w}{2x}. \tag{8}$$

Evidently, if we take $x = \frac{\xi}{2\gamma}$ with $\gamma\theta = 1$, we find an $x > 0$ that satisfies both the equations in (8) and $\xi\eta = w$. The above equations for $x > 0$ also satisfy $\xi > 0, \eta > 0$, as required by (6) (indeed, $w > 0$ implies that both ξ and η must be strictly positive).

If $x < 0$, then (7) reduces to

$$\xi = \gamma\Big(\frac{|x| - x}{4x^2} w\Big) = -\frac{\gamma w}{2x} \quad \eta = \theta\Big(|x| - x\Big) = -2\theta x \tag{9}$$

Proceeding as in the previous case, if we take $x = -\frac{\eta}{2\theta}$ with $\gamma\theta = 1$, we find an $x < 0$ that, replaced into (7), satisfies all the conditions on ξ and η. Thus, there exists one positive and one negative value of x by which we can write (6) in the form (7). $\square$

Remark 1. We remark that the value of $x \in \mathbb{R}$ that satisfies (5) is unique, but there are two values of $x \in \mathbb{R}\backslash\{0\}$ that satisfy (6)–(7). By the equations (8) and (9), these two values, which we denote by $x_+ > 0$ and $x_- < 0$, can be determined by

$$
\begin{aligned}
&\xi = 2\gamma x_+ \quad \eta = \frac{\theta w}{2x_+}, \text{ with } x_+ > 0 \\
&\xi = -\frac{\gamma w}{2x_-} \quad \eta = -2\theta x_-, \text{ with } x_- < 0,
\end{aligned}
\tag{10}
$$

and map to the same (ξ, η). Equation (10) also implies that we can always compute x_+ from x_- (and viceversa) according to

$$
x_+ = -\frac{w}{4x_-}; \quad x_- = -\frac{w}{4x_+}. \tag{11}
$$

Remark 2. Instead of (7), we could use other modulus-based formulations, such as

$$
\xi = \gamma|x|, \quad \eta = \frac{\theta w}{|x|} \tag{12}
$$

or, analogously,

$$
\xi = \frac{\gamma w}{|x|}, \quad \eta = \theta|x|. \tag{13}
$$

Any couple ξ, η that satisfies (6) can be written as in (12) by two values of x, which are equal in absolute value. Such absolute value is the same as x_+ of (10) up to a scaling of 2. Similar considerations apply to (13), which allows to write any couple ξ, η that satisfies (6) by two values of x (equal in absolute value, which is the same as $|x_-|$ up to a scaling of 2). Thus, (12) – (13) are equivalent to (7). In the following, we will use (7) because it leads to formulations that are easier to treat in the formalism of modulus-based methods.

Based on Lemma 1, we can obtain a modulus-based reformulation of the WHLCP. In this regard, consider the WHLCP (1) and assume that the set $\{A, B\}$ has the column $\mathcal{W}$-property. Next, consider the modulus-based system

$$
(A\Gamma + B\Theta)\boldsymbol{x} = (B\Theta - A\Gamma)|\boldsymbol{x}| + (A\Gamma + B\Theta)\tilde{\boldsymbol{x}} + (B\Theta - A\Gamma)|\tilde{\boldsymbol{x}}| + \boldsymbol{q}, \tag{14}
$$

where $\Gamma, \Theta \in \mathbb{R}^{n\times n}$ are positive diagonal matrices of parameters that are defined in the following. Moreover, for compactness, we have set $\tilde{\boldsymbol{x}}$ as the n-dimensional vector of components

$$
\tilde{x}_i = 0 \text{ for } i = 1, \dots, p; \qquad \tilde{x}_i = \frac{w_i}{4x_i} \text{ for } i = p+1, \dots, n \tag{15}
$$

for some integer $p \in [0, n]$. Evidently, the system (14) is defined for any n-dimensional vector $\boldsymbol{x}$ of components $x_i \in \mathbb{R}$ for $i = 1, \dots, p$ and $x_i \in \mathbb{R}\backslash\{0\}$ for $i = p+1, \dots, n$. Within this range of definition, we can seek for solutions to (14). The next theorem proves that all such solutions map to the solution to the WHLCP (1) when p denotes the number of zero components of the weight.

Theorem 1. *Let $p \in [0, n]$ denote the number of zero components of the weight $\boldsymbol{w}$ of (1) and assume, for simplicity and with no loss of generality,[1] that all zero weights are in the first p entries of $\boldsymbol{w}$. Furthermore, let $\Gamma, \Theta \in \mathbb{R}^{n\times n}$ be two positive diagonal matrices of arbitrary diagonal elements $\gamma_{ii} > 0$ and $\theta_{ii} > 0$, respectively, $i = 1, \dots, n$, satisfying $\gamma_{ii}\theta_{ii} = 1$ for $i = p+1, \dots, n$. Then,*

(I) *Let $\boldsymbol{x}$ be an n-dimensional vector of components $x_i \in \mathbb{R}$ for $i = 1, \dots, p$ and $x_i \in \mathbb{R}\backslash\{0\}$ for $i = p+1, \dots, n$ that solves (14) with $\tilde{\boldsymbol{x}}$ as in (15). Then the couple*

$$
\boldsymbol{\xi} = \Gamma(|\boldsymbol{x}| + \boldsymbol{x} + |\tilde{\boldsymbol{x}}| - \tilde{\boldsymbol{x}}); \quad \boldsymbol{\eta} = \Theta(|\boldsymbol{x}| - \boldsymbol{x} + |\tilde{\boldsymbol{x}}| + \tilde{\boldsymbol{x}}) \tag{16}
$$

solves the WHLCP (1).

(II) *if $(\boldsymbol{\xi}, \boldsymbol{\eta})$ solves the WHLCP (1), every vector $\boldsymbol{x}$ of components*

$$
\begin{aligned}
&x_i = \frac{1}{2}(\gamma_{ii}\xi_i - \theta_{ii}\eta_i) \quad \text{for } i = 1, \dots, p \\
&x_i = \frac{\xi_i}{2\gamma_{ii}} \text{ or } x_i = -\frac{\eta_i}{2\theta_{ii}} \quad \text{for } i = p+1, \dots, n
\end{aligned}
$$

satisfies (14) with $\tilde{\boldsymbol{x}}$ as in (15).

Proof. As regards the first claim, by rearranging terms we find that any solution to (14) will satisfy

$$
A\Gamma(|\boldsymbol{x}| + \boldsymbol{x} + |\tilde{\boldsymbol{x}}| - \tilde{\boldsymbol{x}}) - B\Theta(|\boldsymbol{x}| - \boldsymbol{x} + |\tilde{\boldsymbol{x}}| + \tilde{\boldsymbol{x}}) = \boldsymbol{q}. \tag{17}
$$

Next, notice that $\Gamma(|\boldsymbol{x}| + \boldsymbol{x} + |\tilde{\boldsymbol{x}}| - \tilde{\boldsymbol{x}})$ and $\Theta(|\boldsymbol{x}| - \boldsymbol{x} + |\tilde{\boldsymbol{x}}| + \tilde{\boldsymbol{x}})$ of the previous equation are two vectors $\boldsymbol{\xi}, \boldsymbol{\eta}$ which satisfy $\boldsymbol{\xi} \geq \boldsymbol{0}, \boldsymbol{\eta} \geq \boldsymbol{0}, \boldsymbol{\xi} \circ \boldsymbol{\eta} = \boldsymbol{w}$, with $w_i = 0$ for $i = 1, \dots, p$ and $w_i > 0$ for $i = p+1, \dots, n$. This comes directly from Lemma 1: indeed, by the definition of

[1] If this is not the case, we can simply reorder the equations of the problem (1).

$\tilde{\boldsymbol{x}}$ in (15), the components of the above vectors are defined as in (5) and (7). Therefore, it follows that any solution to (17) satisfies (1) with $\boldsymbol{\xi}, \boldsymbol{\eta}$ as in (16).

As regards the second claim, it is sufficient to repeat the same passages backwards. $\square$

Remark 3. The previous theorem uses $\boldsymbol{w}$, which is a known term of the WHLCP much like $\boldsymbol{q}$. Everything that is inferred in Lemma 1 (i.e., that ξ_i, η_i cannot be zero if $w_i = 0$) is given solely by $\boldsymbol{w}$. Hence, p and all the used definitions do not impose any "external" requirement on $\boldsymbol{\xi}$, $\boldsymbol{\eta}$, nor they require the knowledge of any information on the solution. Therefore, the overall situation is analogous to other modulus-based settings: we can solve the nonlinear system (14) – (15) in its range of definition in place of the WHLCP (1).

We conclude this section by an example which helps visualizing the results of Lemma 1 – Theorem 1.

Example 1. Consider the WHLCP (1) with

$$A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \quad B = \begin{pmatrix} 4 & -3 & 0 \\ -2 & 4 & -1 \\ 0 & -1 & 2 \end{pmatrix} \quad \boldsymbol{q} = \begin{pmatrix} 5 \\ -3 \\ 1 \end{pmatrix} \quad \boldsymbol{w} = \begin{pmatrix} 0 \\ 8 \\ 2.5 \end{pmatrix}.$$

The set $\{A, B\}$ has the column $\mathcal{W}$-property, hence the solution to this problem is unique. In particular, for the chosen $\boldsymbol{q}$ and $\boldsymbol{w}$, the solution is $\boldsymbol{\xi}^* = (1.5, 4, 2.5)^T$ and $\boldsymbol{\eta}^* = (0, 2, 1)^T$.

It is easy to verify that the corresponding modulus-based system (14) (where, for simplicity, we consider $\Gamma = \Theta = \boldsymbol{I}$) has the following solutions:

$$\boldsymbol{x}_1^* = \begin{pmatrix} 0.75 \\ 2 \\ 1.25 \end{pmatrix} \quad \boldsymbol{x}_2^* = \begin{pmatrix} 0.75 \\ -1 \\ 1.25 \end{pmatrix} \quad \boldsymbol{x}_3^* = \begin{pmatrix} 0.75 \\ 2 \\ -0.5 \end{pmatrix} \quad \boldsymbol{x}_4^* = \begin{pmatrix} 0.75 \\ -1 \\ -0.5 \end{pmatrix}.$$

These solutions are in agreement with the previous analysis. Indeed, as $w_1 = 0$, the first component of $\boldsymbol{x}_1^*, \boldsymbol{x}_2^*, \boldsymbol{x}_3^*$, and $\boldsymbol{x}_4^*$ is always the same and satisfies (5) with $\gamma = \theta = 1$. On the other hand, for each of the other components we find two values x_{i+}^*, x_{i-}^*, $i = 2, 3$, which satisfy the relations (11) and are in agreement with (7) with $\gamma = \theta = 1$. Such values are then combined to form 2^{n-p} solution vectors, so that all sign combinations appear once in $\{\boldsymbol{x}_1^*, \dots, \boldsymbol{x}_4^*\}$.

Remark 4. As highlighted by the previous example, it can be noticed that the equivalence established by Lemma 1 and Theorem 1 implies that, if the WHLCP (1) has a unique solution, then its modulus-based form (14) has 2^{n-p} solutions (all mapping to the same WHLCP solution via (16)).

3. Modulus-based non-smooth Newton's method

The modulus-based framework allows to formulate new solution methods. For instance, set, for simplicity, $\Gamma = \Theta = \boldsymbol{I}$. By Theorem 1, we can solve the WHLCP (1) by finding a solution to the nonlinear system

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0} \tag{18}$$

where

$$\boldsymbol{f}(\boldsymbol{x}) := (A + B)\boldsymbol{x} + (A - B)|\boldsymbol{x}| - (A + B)\tilde{\boldsymbol{x}} + (A - B)|\tilde{\boldsymbol{x}}| - \boldsymbol{q}. \tag{19}$$

By the definition of $\tilde{\boldsymbol{x}}$ in (15), the mapping $\boldsymbol{f}(\boldsymbol{x})$ has components $f_i(\boldsymbol{x}) : \mathbb{R} \to \mathbb{R}$ for $i = 1, \dots, p$ and $f_i(\boldsymbol{x}) : \mathbb{R}\backslash\{0\} \to \mathbb{R}$ for $i = p + 1, \dots, n$.

We may think of solving (18) by a Newton's method. A non-smooth variant is however required, because of the non-differentiability of the absolute value function at 0. Such methods were analyzed for LCPs [41] and for other complementarity problems (see, e.g., [42,43]) but not for WHLCPs. Hence, we define the Clarke's generalized Jacobian of $|\boldsymbol{x}|$, which is the set of matrices

$$\partial|\boldsymbol{x}| = diag(\partial|x_i|) \text{ with } \partial|x_i| = \begin{cases} -1 \text{ for } x_i < 0 \\ \beta \text{ for } x_i = 0, \text{ with } \beta \in [-1, 1] \\ 1 \text{ for } x_i > 0. \end{cases} \tag{20}$$

Recall that $f_i(\boldsymbol{x}) : \mathbb{R}\backslash\{0\} \to \mathbb{R}$ for $i = p + 1, \dots, n$. Hence, by (15), $\tilde{\boldsymbol{x}}$ and $|\tilde{\boldsymbol{x}}|$ are always differentiable where (18) is defined. Thus, we define the following Jacobian matrices of $\tilde{\boldsymbol{x}}$ and of $|\tilde{\boldsymbol{x}}|$, respectively, which are defined at all possible solutions to (18):

$$J(\tilde{\boldsymbol{x}}) = diag(\partial\tilde{x}_i) \text{ with } \partial\tilde{x}_i = \begin{cases} 0 \text{ for } i = 1,\dots,p \\ -\frac{w_i}{4x_i^2} \text{ for } x_i \neq 0, \text{ for } i = p+1,\dots,n \end{cases}$$

$$J(|\tilde{\boldsymbol{x}}|) = diag(\partial|\tilde{x}_i|) \text{ with } \partial|\tilde{x}_i| = \begin{cases} 0 \text{ for } i = 1,\dots,p \\ \frac{w_i}{4x_i^2} \text{ for } i = p+1,\dots,n \text{ and } x_i < 0 \\ -\frac{w_i}{4x_i^2} \text{ for } i = p+1,\dots,n \text{ and } x_i > 0. \end{cases} \tag{21}$$

At this point, we can define the generalized Jacobian of (18) as the set of matrices

$$\begin{aligned} \partial F(\boldsymbol{x}) &= (A+B) + (A-B)\partial|\boldsymbol{x}| - (A+B)J(\tilde{\boldsymbol{x}}) + (A-B)J(|\tilde{\boldsymbol{x}}|) \\ &= A\Big(I + \partial|\boldsymbol{x}| - J(\tilde{\boldsymbol{x}}) + J(|\tilde{\boldsymbol{x}}|)\Big) + B\Big(I - \partial|\boldsymbol{x}| - J(\tilde{\boldsymbol{x}}) - J(|\tilde{\boldsymbol{x}}|)\Big). \end{aligned} \tag{22}$$

Such Jacobian is defined for $x_i \neq 0$, $i = p+1,\dots,n$.

We can finally set up the following non-smooth Newton's iteration, which is well-defined in the neighborhood of each solution to (18).

Method 1. *Let $\boldsymbol{x}^{(0)} \in \mathbb{R}^n$ with $x_i^{(0)} \neq 0$ for $i = p+1,\dots,n$. The non-smooth Newton's method computes the generic $(k+1)$-th iterate as*

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \Delta\boldsymbol{x}^{(k)}, \tag{23}$$

$k = 0,1,\dots$, where $\Delta\boldsymbol{x}^{(k)}$ is computed as the solution of

$$V_k\Delta\boldsymbol{x} = -\boldsymbol{f}(\boldsymbol{x}^{(k)}), \tag{24}$$

with $V_k \in \partial F(\boldsymbol{x}^{(k)})$.

To study the convergence of the method, we must analyze the non-singularity of the generalized Jacobian at the solutions to (18). In this regard, we prove the following theorem.

Theorem 2. *Let $\mathcal{M} = \{A, B\}$ satisfy the column $\mathcal{W}$-property and let $\boldsymbol{x}^*$ be any solution to (18). Then, any matrix $V^* \in \partial F(\boldsymbol{x}^*)$ is non-singular.*

Proof. First, let us define the following sets of positive, negative and zero elements of $\boldsymbol{x}^*$

$$\begin{aligned} &\mathcal{P}_1^* = \{i : x_i^* > 0, i = 1,\dots,p\}; \\ &\mathcal{P}_2^* = \{i : x_i^* > 0, i = p+1,\dots,n\}; \\ &\mathcal{N}_1^* = \{i : x_i^* < 0, i = 1,\dots,p\}; \\ &\mathcal{N}_2^* = \{i : x_i^* < 0, i = p+1,\dots,n\}; \\ &\mathcal{Z}^* = \{i : x_i^* = 0, i = 1,\dots,p\}. \end{aligned} \tag{25}$$

We recall that $x_i^* \neq 0$, $i = p+1,\dots,n$. Thus, we do not need to define the set of zero elements for $i = p+1,\dots,n$ and any generalized Jacobian (22) is defined at $\boldsymbol{x}^*$. We denote the cardinality of each set in (25) by r_1, r_2, s_1, s_2, t for $\mathcal{P}_1^*, \mathcal{P}_2^*, \mathcal{N}_1^*, \mathcal{N}_2^*, \mathcal{Z}^*$ respectively. In the most general case, r_1, r_2, s_1, s_2, t will all be nonzero. We focus on this situation, as all other special cases can be studied simply by removing the sets that are void from the following analysis.

Thus, there exists a permutation matrix P such that the rows of V^* corresponding to positive, negative, and zero values of x_i^* are grouped together, for $i = 1,\dots,p$ and for $i = p+1,\dots,n$. I.e., considering $\partial|\boldsymbol{x}^*|$, the permutation acts in the following way.

$$P\partial|\boldsymbol{x}^*|P^T = \begin{pmatrix} I_{r_1} & 0 & 0 & 0 & 0 \\ 0 & -I_{s_1} & 0 & 0 & 0 \\ 0 & 0 & \tilde{D} & 0 & 0 \\ 0 & 0 & 0 & I_{r_2} & 0 \\ 0 & 0 & 0 & 0 & -I_{s_2} \end{pmatrix}, \tag{26}$$

where $\tilde{D}$ is a diagonal matrix of order t and of elements $\beta \in [-1,1]$, and $I_{r_1}, I_{s_1}, I_{r_2}, I_{s_2}$ are identity matrices of order r_1, s_1, r_2, s_2, respectively. Similarly, we have

$$PJ(\tilde{\boldsymbol{x}}^*)P^T = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -D_{r_2} & 0 \\ 0 & 0 & 0 & 0 & -D_{s_2} \end{pmatrix} \qquad PJ(|\tilde{\boldsymbol{x}}^*|)P^T = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -D_{r_2} & 0 \\ 0 & 0 & 0 & 0 & D_{s_2} \end{pmatrix}, \tag{27}$$

where D_{r_2} and D_{s_2} are positive diagonal matrices of order r_2 and s_2, respectively. In particular, their entries are $\frac{w_i}{4(x_i^*)^2}$, $i = p+1, \dots, n$, as governed by the permutation. Hence, by (22), we can write

$$\begin{aligned} P\partial F(\boldsymbol{x}^*)P^T =& PAP^T\left(I + P\partial|\boldsymbol{x}^*|P^T - PJ(\tilde{\boldsymbol{x}}^*)P^T + PJ(|\tilde{\boldsymbol{x}}^*|)P^T\right) \\ &+ PBP^T\left(I - P\partial|\boldsymbol{x}^*|P^T - PJ(\tilde{\boldsymbol{x}}^*)P^T - PJ(|\tilde{\boldsymbol{x}}^*|)P^T\right) \\ =& PAP^T\tilde{D}_A + PBP^T\tilde{D}_B \end{aligned} \tag{28}$$

with

$$\tilde{D}_A = \begin{pmatrix} 2I_{r_1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I_t + \tilde{D} & 0 & 0 \\ 0 & 0 & 0 & 2I_{r_2} & 0 \\ 0 & 0 & 0 & 0 & 2D_{s_2} \end{pmatrix} \qquad \tilde{D}_B = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 2I_{s_1} & 0 & 0 & 0 \\ 0 & 0 & I_t - \tilde{D} & 0 & 0 \\ 0 & 0 & 0 & 2D_{r_2} & 0 \\ 0 & 0 & 0 & 0 & 2I_{s_2} \end{pmatrix}$$

In the previous expression, I_t denotes the identity matrix of order t. Evidently, $\tilde{D}_A$ and $\tilde{D}_B$ are non-negative matrices for $\beta \in [-1, 1]$ (which is, for any $V^* \in \partial F(\boldsymbol{x}^*)$) and $\tilde{D}_A + \tilde{D}_B$ is a positive diagonal matrix. Furthermore, by [43, Lemma 1], if $\mathcal{M}$ has the column $\mathcal{W}$-property, any permuted set $\{PAP^T, PBP^T\}$ has the column $\mathcal{W}$-property. Therefore, $PAP^T\tilde{D}_A + PBP^T\tilde{D}_B$ is non-singular for any $V^* \in \partial F(\boldsymbol{x}^*)$ by [35, Theorem 2]. □

Theorem 3. *Under the assumptions of Theorem 2, the non-smooth Newton's method 1 is locally convergent in a neighborhood of any solution* $\boldsymbol{x}^*$ *to* (18).

Proof. Via the column $\mathcal{W}$-property, Theorem 2 ensures that any matrix $V^* \in \partial F(\boldsymbol{x}^*)$ is not singular at any solution $\boldsymbol{x}^*$ of (18) whenever the solution to the corresponding WHLCP (1) is unique. In addition, the fact that $x_i^* \neq 0$, $i = p+1, \dots, n$, ensures that $\boldsymbol{f}(\boldsymbol{x})$ is (locally) Lipschitz continuous in a neighborhood of any solution. Indeed, $\boldsymbol{x}$ and $|\boldsymbol{x}|$ are globally Lipschitz in all components, while terms of the form $\frac{1}{x_i}$ are Lipschitz continuous provided that they are taken in an interval that excludes zero. Indeed, for $x_i \in (c, \infty)$ or $x_i \in (-\infty, -c)$ with $c > 0$, $\frac{1}{x_i}$ is Lipschitz continuous with constant c^{-2}. Similar considerations apply to $\frac{1}{|x_i|}$. Furthermore, $\boldsymbol{f}(\boldsymbol{x}^*)$ is semi-smooth. More precisely, the first p components of $\boldsymbol{f}(\boldsymbol{x})$ are semi-smooth for any $\boldsymbol{x}$, while the last $n - p$ components are smooth for $x_i \neq 0$, which is satisfied for any $\boldsymbol{x}^*$.

Hence, the modulus-based non-smooth Newton method for WHLCPs is locally convergent by [39, Theorem 3.2], which requires the non-singularity of all $V^* \in \partial F(\boldsymbol{x}^*)$ and that $\boldsymbol{f}(\boldsymbol{x}^*)$ is (locally) Lipschitz and semi-smooth. □

Remark 5. Take any solution $\boldsymbol{x}^*$ to (18). As the system (18) is not defined for $x_i = 0$, $i = p+1, \dots, n$, the neighborhood where the method is convergent does not include zero. Thus, one may argue that we need to correctly guess the sign of a solution for $i = p+1, \dots, n$ in order to achieve convergence in an actual implementation of the method. However, we have previously highlighted that (when $\mathcal{M} = \{A, B\}$ satisfies the column $\mathcal{W}$-property) the modulus-based system (18) has 2^{n-p} solutions with univocal sign pattern. Indeed, there exist two values $x_{i+}^* > 0$, $x_{i-}^* < 0$ corresponding to each i-th component of the unique solution of the WHLCP, $i = p+1, \dots, n$. This means that, for an arbitrary starting iterate of the non-smooth Newton method, we will "guess" the sign from the $(p+1)$-th to the n-th component of exactly one solution. Thus, we can achieve convergence to that solution without needing to change the sign with respect to the starting iterate.

4. Modulus-based matrix splitting methods

4.1. Formulation of the method

Another way to find a solution to the system (14) is to formulate a matrix splitting method. In this regard, for simplicity and consistently with most modulus-based splitting methods, let us set $\Gamma = \frac{1}{\gamma}I$ with $\gamma > 0$. Scaling the equation with respect to γ, the modulus-based system (14) can equivalently be written as

$$(A + B\Theta)\boldsymbol{x} = (B\Theta - A)|\boldsymbol{x}| + (A + B\Theta)\tilde{\boldsymbol{x}} + (B\Theta - A)|\tilde{\boldsymbol{x}}| + \gamma\boldsymbol{q} \tag{29}$$

where Θ must have components $\theta_{ii} = \gamma^2$, $i = p+1, \dots, n$, to satisfy the requirements of Theorem 1. In this setting, we could formulate, for instance, the following solution strategy.

Method 2. *Let $A = M_A - N_A$ and $B = M_B - N_B$ be two splittings of the matrices $A, B \in \mathbb{R}^{n\times n}$. Furthermore, let $\gamma > 0$ and let $\Theta \in \mathbb{R}^{n\times n}$ be a positive diagonal matrix of arbitrary components $\theta_{ii} > 0$ for $i = 1,\dots,p$ and $\theta = \gamma^2$ for $i = p+1,\dots,n$.*

Starting from an initial guess $\boldsymbol{x}^{(0)} \in \mathbb{R}^n$ and until a stopping condition is met, the modulus-based matrix splitting method for WHLCPs computes the generic $(k+1)$-th iterate as the solution to

$$(M_A + M_B\Theta)\boldsymbol{x} = (N_A + N_B\Theta)\boldsymbol{x}^{(k)} + (B\Theta - A)|\boldsymbol{x}^{(k)}| + (A + B\Theta)\tilde{\boldsymbol{x}}^{(k)} + (B\Theta - A)|\tilde{\boldsymbol{x}}^{(k)}| + \gamma\boldsymbol{q}, \tag{30}$$

for $k = 0,1,\dots$ and with $\tilde{\boldsymbol{x}}^{(k)}$ defined as in (15) with $x_i = x_i^{(k)}$. At each iteration, the vectors $\boldsymbol{\xi}^{(k)}, \boldsymbol{\eta}^{(k)} \in \mathbb{R}^n$ that satisfy the weighted complementarity condition have components

$$\begin{aligned} &\xi_i^{(k)} = \frac{1}{\gamma}(|x_i^{(k)}| + x_i^{(k)}); \quad \eta_i^{(k)} = \frac{\theta_{ii}}{\gamma}(|x_i^{(k)}| - x_i^{(k)}) \quad \text{for } i = 1,\dots,p \\ &\xi_i^{(k)} = \frac{1}{\gamma}\left(|x_i^{(k)}| + x_i^{(k)} + \frac{|x_i^{(k)}| - x_i^{(k)}}{4x_i^{(k)2}}w_i\right); \quad \eta_i^{(k)} = \gamma\left(|x_i^{(k)}| - x_i^{(k)} + \frac{|x_i^{(k)}| + x_i^{(k)}}{4x_i^{(k)2}}w_i\right) \quad \text{for } i = p+1,\dots,n. \end{aligned} \tag{31}$$

Naturally, an effective solution strategy requires that the sub-problems (30) be easy to solve. In this regard, we especially focus on well-known accelerated over-relaxation (AOR) splittings. Thus, let us define D_A, D_B as the diagonal matrices of diagonal elements a_{ii} and b_{ii}, respectively. Similarly, let $-L_A, -L_B$ be the lower-triangular parts of A and B, respectively, and let $-U_A, -U_B$ be their upper-triangular parts. Then, the modulus-based AOR method is obtained by using the following splittings in (30).

$$\begin{aligned} &M_A = \frac{D_A - \beta L_A}{\alpha}; \quad N_A = \frac{(1-\alpha)D_A + (\alpha-\beta)L_A + \alpha U_A}{\alpha} \\ &M_B = \frac{D_B - \beta L_B}{\alpha}; \quad N_B = \frac{(1-\alpha)D_B + (\alpha-\beta)L_B + \alpha U_B}{\alpha} \end{aligned} \tag{32}$$

We will particularly focus on the choices $\alpha = 1, \beta = 0$ (which corresponds to the Jacobi splitting), $\alpha = \beta = 1$ (corresponding to the Gauss-Seidel splitting), and $\alpha = \beta$ (corresponding to the SOR splitting).

4.2. Convergence analysis

To analyze the convergence of the method, let us view (30) as the fixed-point iteration that sets

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{g}(\boldsymbol{x}^{(k)}), \quad k = 0,1,\dots, \tag{33}$$

with

$$\boldsymbol{g}(\boldsymbol{x}) := (M_A + M_B\Theta)^{-1}\Big[(N_A + N_B\Theta)\boldsymbol{x} + (B\Theta - A)|\boldsymbol{x}| + (A + B\Theta)\tilde{\boldsymbol{x}} + (B\Theta - A)|\tilde{\boldsymbol{x}}| + \gamma\boldsymbol{q}\Big]. \tag{34}$$

Just as (29), the above mapping is defined for any n-dimensional $\boldsymbol{x}$ of components $x_i \in \mathbb{R}$ for $i = 1,\dots,p$ and $x_i \in \mathbb{R}\backslash\{0\}$ for $i = p+1,\dots,n$. Hence, (33) is well-defined in the neighborhood of any solution to (29).

4.2.1. Auxiliary lemmas

Clearly, the solutions that satisfy $\boldsymbol{x}^* = \boldsymbol{g}(\boldsymbol{x}^*)$ correspond to the solutions of the problem (29). Hence, 2^{n-p} fixed points exist. However, we do not know whether they are points of attraction for the fixed point iteration. Before directly studying the convergence properties of the fixed points, we need to further characterize the modulus-based solutions via the following auxiliary lemmas.

Lemma 2. *For $i = p+1,\dots,n$, let x_{i+}^*, x_{i-}^* denote the positive and the negative i-th component of the solutions to the system (29), respectively. If $x_{i+}^* = -x_{i-}^*$, then*

$$\frac{w_i}{4(x_{i+}^*)^2} = \frac{w_i}{4(x_{i-}^*)^2} = 1. \tag{35}$$

If, instead, $x_{i+}^* \neq -x_{i-}^*$, then either

$$\frac{w_i}{4(x_{i+}^*)^2} < 1; \quad \frac{w_i}{4(x_{i-}^*)^2} > 1 \tag{36}$$

or

$$\frac{w_i}{4(x_{i+}^*)^2} > 1; \quad \frac{w_i}{4(x_{i-}^*)^2} < 1. \tag{37}$$

Proof. Let us assume $\frac{w_i}{4(x^*_{i+})^2} = K < 1$. Then, replacing x_{i+} according to (11), we have

$$\frac{4(x^*_{i-})^2}{w_i} = K,$$

hence

$$\frac{w_i}{4(x^*_{i-})^2} = \frac{1}{K} > 1.$$

If $\frac{w_i}{4(x^*_{i+})^2} = K > 1$, we proceed analogously to show that $\frac{w_i}{4(x^*_{i-})^2} = \frac{1}{K} < 1$. Finally, if $\frac{w_i}{4(x^*_{i+})^2} = K = 1$, proceeding as above we obtain (35), which implies

$$x^*_{i+} = -x^*_{i-} = \sqrt{\frac{w_i}{4}}. \quad \square \tag{38}$$

As an example, one may again consider Example 1. Indeed, it is easy to verify that the solutions that were found in Example 1 satisfy the lemma. Furthermore, Lemma 2 together with the considerations in Remark 1 on the pattern of the signs of the solutions, immediately leads to the following corollary.

Corollary 1. *If $x^*_{i+} \neq -x^*_{i-}$ for $i = p+1, \dots, n$, there exists one and only one solution $\boldsymbol{x}^*$ to (29) that satisfies*

$$\frac{w_i}{4(x^*_i)^2} < 1 \quad \text{for } i = p+1, \dots, n. \tag{39}$$

It is now useful to analyze how the solution changes as γ in the system (29) is varied.

Lemma 3. *Let $\gamma_1, \gamma_2 \in \mathbb{R}$ and let $\boldsymbol{x}^*_1$ be any one of the solutions to (29) with $\gamma = \gamma_1$. Then, the vector $\boldsymbol{x}^*_2$ with the same sign pattern as $\boldsymbol{x}^*_1$ that solves (29) with $\gamma = \gamma_2$ has components*

$$\begin{aligned} x^*_{2_i} &= \frac{\gamma_2}{\gamma_1} x^*_{1_i} \quad &\text{if } x^*_{1_i} > 0 \\ x^*_{2_i} &= \frac{\gamma_1}{\gamma_2} x^*_{1_i} \quad &\text{if } x^*_{1_i} < 0, \end{aligned} \tag{40}$$

for $i = p+1, \dots, n$.

Proof. While γ can scale and move the solutions of (29), both $\boldsymbol{x}^*_1$ and $\boldsymbol{x}^*_2$ map to the same $\boldsymbol{\xi}^*, \boldsymbol{\eta}^*$ by the equivalence between the WHLCP and the modulus-based formulation. Thus, the expressions in (31) at the solution must give the same values for any $x^*_{1_i}$ and $x^*_{2_i}$, which gives the equalities

$$\begin{aligned} &\frac{1}{\gamma_1}\left(|x^*_{1_i}| + x^*_{1_i} + \frac{|x^*_{1_i}| - x^*_{1_i}}{4(x^*_{1_i})^2} w_i\right) = \frac{1}{\gamma_2}\left(|x^*_{2_i}| + x^*_{2_i} + \frac{|x^*_{2_i}| - x^*_{2_i}}{4(x^*_{2_i})^2} w_i\right); \\ &\gamma_1\left(|x^*_{1_i}| - x^*_{1_i} + \frac{|x^*_{1_i}| + x^*_{1_i}}{4(x^*_{1_i})^2} w_i\right) = \gamma_2\left(|x^*_{2_i}| - x^*_{2_i} + \frac{|x^*_{2_i}| + x^*_{2_i}}{4(x^*_{2_i})^2} w_i\right) \end{aligned}$$

for $i = p+1, \dots, n$. If $x^*_{1_i}, x^*_{2_i} > 0$, the first equality immediately reduces to

$$\frac{2x^*_{1_i}}{\gamma_1} = \frac{2x^*_{2_i}}{\gamma_2},$$

while, if $x^*_{1_i}, x^*_{2_i} < 0$, the second equality reduces to

$$-2x^*_{1_i}\gamma_1 = -2x^*_{2_i}\gamma_2. \tag{41}$$

These relations evidently imply (40). $\square$

Lemma 3 evidently implies that any positive component of a solution increases as γ in the system (29) increases, while any negative component increases as γ approaches zero.

4.2.2. Convergence for $p = 0$

Assume $p = 0$. Then, at any solution $\boldsymbol{x}^*$, we have $x^*_i \neq 0$, for $i = 1, \dots, n$. Hence, $\boldsymbol{g}(\boldsymbol{x})$ as in (34) is differentiable at any solution $\boldsymbol{x}^*$ and its Jacobian matrix is

$$G(\boldsymbol{x}) = (M_A + M_B\Theta)^{-1}\left[N_A + N_B\Theta + (B\Theta - A)D_s(\boldsymbol{x}) - (A + B\Theta)D_w(\boldsymbol{x}) - (B\Theta - A)D_w(\boldsymbol{x})D_s(\boldsymbol{x})\right], \tag{42}$$

which is defined for any n-dimensional $\boldsymbol{x}$ with nonzero components and where

- $D_s(\boldsymbol{x})$ is the diagonal matrix of elements $d_{s_{ii}} = sign(x_i)$, $i = 1, \dots, n$. Notice that, by the definition of $\tilde{\boldsymbol{x}}$ in (15), when $p = 0$, $d_{s_{ii}}$ is also equal to $sign(\tilde{x}_i)$. We have used this fact to write $G(\boldsymbol{x})$ as in (42);
- $D_w(\boldsymbol{x})$ denotes the positive diagonal matrix of elements $d_{w_{ii}} = \frac{w_i}{4x_i^2}$, $i = 1, \dots, n$.

For simplicity, we first prove the convergence of the modulus-based Jacobi method. We then extend the results to a general SOR iteration by proceeding in the same way. In the following, we define $\boldsymbol{x}_+^* > \boldsymbol{0}$ as the solution of positive components x_{i+}^*, $i = 1, \dots, n$. Similarly, $\boldsymbol{x}_-^* < \boldsymbol{0}$ denotes the solution of negative components x_{i-}^*, $i = 1, \dots, n$. For compactness of notation, we also define $D_w^+ := D_w(\boldsymbol{x}_+^*)$ and $D_w^- := D_w(\boldsymbol{x}_-^*)$. We further remark that both matrices are positive diagonal.

Theorem 4. *Let $\boldsymbol{x}_+^* > \boldsymbol{0}$ be the solution to (29) with positive components. Furthermore, let A, B have positive diagonal entries and assume that γ is chosen sufficiently large to have $I - 2D_w^+$ positive diagonal matrix and $D_B\Theta(I - 2D_w^+) > D_A$.*

If $\langle A\rangle + \langle B\rangle\Theta D_w^+$ is an M-matrix, then $\boldsymbol{x}_+^*$ is an attracting fixed point for the iteration $\boldsymbol{x}^{(k+1)} = \boldsymbol{g}(\boldsymbol{x}^{(k)})$ with the Jacobi splitting.

Proof. First, we notice that there exists a sufficiently large γ such that $I - 2D_w^+$ is a positive diagonal matrix and $D_B\Theta(I - 2D_w^+) > D_A$. Indeed, D_w^+ is the diagonal matrix of components $d_{w_{ii}}^+ = \frac{w_i}{4(x_{i+}^*)^2}$, $i = 1, \dots, n$, where we know that any x_{i+}^* increases with γ by Lemma 3. Hence, there must exist a $\hat{\gamma}$ such that all components of the positive solution $\boldsymbol{x}_+^*$ are sufficiently large to ensure that $I - 2D_w^+$ is positive diagonal for any $\gamma \geq \hat{\gamma}$. Recalling $\Theta = \gamma^2 I$, the condition $D_B\Theta(I - 2D_w^+) > D_A$ can also be satisfied by a sufficiently large γ.

Next, consider the Jacobian matrix (42) and the Jacobi splittings $A = D_A - N_A$, $B = D_B - N_B$. At $\boldsymbol{x}_+^*$, we know that the Jacobian is non-singular and that $D_s(\boldsymbol{x}_+^*) = I$. Hence,

$$|G(\boldsymbol{x}_+^*)| = \left|(D_A + D_B\Theta)^{-1}\left[N_A + N_B\Theta + (B\Theta - A) - (A + B\Theta)D_w^+ - (B\Theta - A)D_w^+\right]\right|.$$

Collecting terms and by the triangle inequality, we can further evaluate

$$\begin{aligned}
|G(\boldsymbol{x}_+^*)| =& \left|(D_A + D_B\Theta)^{-1}\left[D_B\Theta - D_A + 2N_A - 2B\Theta D_w^+\right]\right| \\
=& (D_A + D_B\Theta)^{-1}\left[|D_B\Theta - D_A - 2D_B\Theta D_w^+| + |2N_A + 2N_B\Theta D_w^+|\right] \\
\leq& (D_A + D_B\Theta)^{-1}\left[|D_B\Theta(I - 2D_w^+) - D_A| + 2|N_A| + 2|N_B|\Theta D_w^+\right] \\
:=& \mathcal{M}^{-1}\mathcal{N},
\end{aligned}$$

with $\mathcal{M} = (D_A + D_B\Theta)$ and $\mathcal{N} = |D_B\Theta(I - 2D_w^+) - D_A| + 2|N_A| + 2|N_B|\Theta D_w^+$. By the assumptions $I - 2D_w^+$ positive diagonal matrix and $D_B\Theta(I - 2D_w^+) > D_A$, we can easily evaluate

$$\begin{aligned}
\mathcal{M} - \mathcal{N} =& D_A + D_B\Theta - D_B\Theta(I - 2D_w^+) + D_A - 2|N_A| - 2|N_B|\Theta D_w^+ \\
=& 2D_A + 2D_B\Theta D_w^+ - 2|N_A| - 2|N_B|\Theta D_w^+ \\
=& 2\langle A\rangle + 2\langle B\rangle\Theta D_w^+.
\end{aligned}$$

Therefore, under the assumption that $\langle A\rangle + \langle B\rangle\Theta D_w^+$ is an M-matrix, $\mathcal{M} - \mathcal{N}$ is an M-splitting with $\rho(\mathcal{M}^{-1}\mathcal{N}) < 1$ [44, Theorem 3.4]. As $\rho(|G(\boldsymbol{x}_+^*)|) \leq \rho(\mathcal{M}^{-1}\mathcal{N})$ by $|G(\boldsymbol{x}_+^*)| \leq \mathcal{M}^{-1}\mathcal{N}$ (see, e.g., [33, Corollary 1.5]) and as $\rho(T) \leq \rho(|T|)$ for any $T \in \mathbb{R}^{n\times n}$ (see, e.g., [45, §6.2.P4]), we finally get $\rho(G(\boldsymbol{x}_+^*)) < 1$. Hence, $\boldsymbol{x}_+^*$ is attracting by [40, p. 145]. □

Remark 6. Although the values of γ that satisfy $I - 2D_w^+$ positive diagonal and $D_B\Theta(I - 2D_w^+) > D_A$ depend on the unknown solution $\boldsymbol{x}^*$, we have established in the theorem that there exists a sufficiently large $\hat{\gamma}$ that satisfies these conditions. Furthermore, notice that θ_{ii} grows quadratically with γ, and the entries of D_w^+ decrease with the square of the entries of $\boldsymbol{x}_+^*$ (which increase linearly with γ by Lemma 3). Hence, reasonably small values of γ will generally be sufficient to satisfy the assumptions on $I - 2D_w^+$ and $D_B\Theta(I - 2D_w^+)$.

Similarly, the assumption that $\langle A\rangle + \langle B\rangle\Theta D_w^+$ is an M-matrix is satisfied by many commonly used matrices. For instance, assume that A, B are strictly or irreducibly diagonally dominant by columns. Then, $\langle A\rangle + \langle B\rangle\Theta D_w^+$ is an M-matrix for any positive diagonal matrices Θ, D_w^+.

Remark 7. It is interesting to notice that Theorem 4 does not put any upper bound to the choice of γ for the Jacobi splitting. Indeed, D_w^+ will become arbitrarily close to zero as γ increases. At the same time, Θ will become arbitrarily large. However, for excessively large values of γ, the Jacobian matrix will resemble the spectral properties of the identity matrix. Hence, the spectral radius of $G(\boldsymbol{x}_+^*)$ will tend to 1, slowing down the convergence.

Similarly to Theorem 4, we can easily establish a convergence result for $\boldsymbol{x}_-^*$ with a sufficiently small γ.

Theorem 5. *Let $\mathbf{x}_-^* < \mathbf{0}$ be the solution to (29) with negative components. Furthermore, let A, B have positive diagonal elements and assume that $\gamma > 0$ is chosen sufficiently small to have $I - 2D_w^-$ positive diagonal matrix and $D_A(I - 2D_w^-) > D_B\Theta$.*

If $\langle A\rangle D_w^- + \langle B\rangle\Theta$ is an M-matrix, then $\mathbf{x}_-^*$ is an attracting fixed point for the iteration $\mathbf{x}^{(k+1)} = \mathbf{g}(\mathbf{x}^{(k)})$ with the Jacobi splitting.

Proof. Similarly to the proof to Theorem 4, there exists a γ such that $I - 2D_w^-$ is a positive diagonal matrix and $D_A(I - 2D_w^-) > D_B\Theta$. Indeed, Lemma 3 implies that the components of $\mathbf{x}_-^*$ become arbitrarily large when γ approaches zero. Consequently, the components of D_w^- become arbitrarily close to zero. At the same time, $\Theta = \gamma^2 I$ becomes smaller as $\gamma \to 0$. It follows that there exists a $\hat{\gamma}$ such that both conditions are satisfied for any $\gamma \in (0, \hat{\gamma})$.

Next, notice that we have $D_s = -I$ at $\mathbf{x}_-^*$. Hence, proceeding similarly to the proof to Theorem 4,

$$
\begin{aligned}
|G(\mathbf{x}_-^*)| &= \left|(D_A + D_B\Theta)^{-1}\left[N_A + N_B\Theta - (B\Theta - A) - (A + B\Theta)D_w^- + (B\Theta - A)D_w^-\right]\right| \\
&= \left|(D_A + D_B\Theta)^{-1}\left[D_A - D_B\Theta + 2N_B\Theta - 2AD_w^-\right]\right| \\
&\le (D_A + D_B\Theta)^{-1}[D_A(I - 2D_w^-) - D_B\Theta + 2|N_B|\Theta + 2|N_A|D_w^-] \\
&:= \mathcal{M}^{-1}\mathcal{N},
\end{aligned}
$$

with $\mathcal{M} = (D_A + D_B\Theta)$ and $\mathcal{N} = D_A(I - 2D_w^-) - D_B\Theta + 2|N_B|\Theta + 2|N_A|D_w^- \ge 0$. It is then easy to verify that $\mathcal{M} - \mathcal{N}$ is an M-splitting if $\langle A\rangle D_w^- + \langle B\rangle\Theta$ is an M-matrix, thus proving the theorem. □

Considerations analogous to the Remarks 6 and 7 can be applied to this theorem. Indeed, D_w^- will decrease arbitrarily as γ approaches zero, as discussed in the proof. Furthermore, studying when other solutions are attracting fixed-points can be done by considering column-representative matrices of A, B. Indeed, Θ and D_w multiply A and B on the right. Thus, a column-representative "mixing" of the evaluations of the Theorems 4 and 5 can, in principle, provide conditions for solutions where plus and minus signs are mixed. In this regard, Corollary 1 implies that at most one solution is going to satisfy the assumption of the theorems. Indeed, Corollary 1 implies that, if γ is such that x_{i+}^* satisfies $1 - d_{w_{ii}}^+ > 0$, then we necessarily have $1 - d_{w_{ii}}^- < 0$ (and vice-versa). Repeating the same considerations for $i = 1, \dots, n$, there is, at most, one sign pattern that can satisfy $I - 2D_w$ positive diagonal matrix. This motivates why we are focusing on the special cases where $\mathbf{x}_-^*$ or $\mathbf{x}_+^*$ are attracting. Indeed, we have established that $\mathbf{x}_+^*$ will be attracting for sufficiently large γ, while $\mathbf{x}_-^*$ will be attracting for γ sufficiently close to zero (if also the other assumptions of the theorems are satisfied). As γ is an arbitrary (positive) parameter of the method, we can predict, to some extent, whether $\mathbf{x}_+^*$ or $\mathbf{x}_-^*$ will be attracting based on the choice of γ. In turn, this is useful to suitably choose the starting iterate of the modulus-based splitting methods. This is further discussed numerically in Section 5.3.

We now extend the analysis to the more general SOR splitting. The outline of the proof follows Theorem 4, with additional conditions on off-diagonal elements that potentially add an upper bound to the choice of γ.

Theorem 6. *Let $\mathbf{x}_+^* > \mathbf{0}$ be the solution to (29) with positive components and let $\alpha > 0$. Furthermore, let A, B have positive diagonal elements and assume that γ is chosen sufficiently large to have $I - 2\alpha D_w^+$ positive diagonal matrix and $D_B\Theta(I - 2\alpha D_w^+) > -(1 - 2\alpha)D_A$. Finally, let $M_A + M_B\Theta$ be an H_+-matrix and let $|L_B|\Theta \le |L_A|$. If $\langle A\rangle + \langle B\rangle\Theta D_w^+$ is an M-matrix, then $\mathbf{x}_+^*$ is an attracting fixed point for the iteration $\mathbf{x}^{(k+1)} = \mathbf{g}(\mathbf{x}^{(k)})$ with the SOR splitting.*

Proof. Considerations similar to those in the proof to Theorem 4 can be made with regard to the existence of γ such that $I - 2\alpha D_w^+$ is a positive diagonal matrix and $D_B\Theta(I - 2\alpha D_w^+) > -(1 - 2\alpha)D_A$. Next, consider the Jacobian matrix (42). We can proceed similarly to the proof of Theorem 4 and evaluate

$$
\begin{aligned}
|G(\mathbf{x}_+^*)| &= \left|(M_A + M_B\Theta)^{-1}\left[N_A + N_B\Theta + (B\Theta - A)D_s - (A + B\Theta)D_w^+ - (B\Theta - A)D_w^+ D_s\right]\right| \\
&\le \langle M_A + M_B\Theta\rangle^{-1}\left|M_B\Theta - M_A + 2N_A - 2B\Theta D_w^+\right|,
\end{aligned}
$$

where we have used the fact that $|(M_A + M_B\Theta)^{-1}| \le \langle M_A + M_B\Theta\rangle^{-1}$ if $M_A + M_B\Theta$ is an H_+-matrix [44]. Using the SOR splitting, we can replace the splitting matrices by their definitions in (32) with $\alpha = \beta$. Hence, by simple passages where we collect diagonal and off-diagonal terms, and using the triangle inequality, we get

$$
\begin{aligned}
|G(\mathbf{x}_+^*)| \le& \left\langle \frac{1}{\alpha}(D_A + D_B\Theta) - (L_A + L_B\Theta)\right\rangle^{-1}\left[\left|\frac{1 - 2\alpha}{\alpha}D_A + D_B\Theta\left(\frac{1}{\alpha}I - 2D_w^+\right)\right| + \right. \\
&\left. + |L_A - L_B\Theta + 2U_A + 2L_B\Theta D_w^+ + 2U_B\Theta D_w^+|\right] \\
\le& \left\langle \frac{1}{\alpha}(D_A + D_B\Theta) - (L_A + L_B\Theta)\right\rangle^{-1}\left[\frac{1 - 2\alpha}{\alpha}D_A + D_B\Theta\left(\frac{1}{\alpha}I - 2D_w^+\right) + \right. \\
&\left. + |L_A - L_B\Theta| + 2|U_A| + 2|L_B|\Theta D_w^+ + 2|U_B|\Theta D_w^+\right] \\
:=& \mathcal{M}^{-1}\mathcal{N}
\end{aligned}
$$

where $\mathcal{M}$ and $\mathcal{N}$ are redefined accordingly to the matrices in the last passage of the equation. In the last passage, we have also used the assumptions $I - 2\alpha D_w^+$ positive diagonal matrix and $D_B\Theta(I - 2\alpha D_w^+) > -(1-2\alpha)D_A$. Finally, by the assumption $|L_B|\Theta \le |L_A|$, we have

$$\begin{aligned}\mathcal{M} - \mathcal{N} =&\frac{1}{\alpha}(D_A + D_B\Theta) - |L_A + L_B\Theta| - \frac{1-2\alpha}{\alpha}D_A - D_B\Theta\Big(\frac{1}{\alpha}I - 2D_w^+\Big)\\ &- |L_B\Theta - L_A| - 2|U_A| - 2|L_B|\Theta D_w^+ - 2|U_B|\Theta D_w^+\\ =&2D_A + 2D_B\Theta D_w^+ - 2|L_A| - 2|U_A| - 2|L_B|\Theta D_w^+ - 2|U_B|\Theta D_w^+\\ =&2\langle A\rangle + 2\langle B\rangle\Theta D_w^+.\end{aligned}$$

This is the same expression that was found at the end of Theorem 4. Hence, by repeating the same considerations, $\boldsymbol{x}_+^*$ is an attracting fixed point for the SOR iteration. □

Remark 8. Contrarily to the Theorems 4 and 5, γ can no longer be taken arbitrarily large, in order not to violate the assumption $|L_B|\Theta \le |L_A|$ of Theorem 6. In case one needs to relax this condition, the proof to Theorem 6 immediately suggests to replace the assumption $|L_B|\Theta \le |L_A|$ by requiring that $\langle A\rangle + \langle B\rangle\Theta D_w^+ - |L_B|\Theta$ be an M-matrix. Such condition is harder to check for general A, B, but can be satisfied even when $|L_B|\Theta \ge |L_A|$.

Finally, we can prove the following theorem for $\boldsymbol{x}_-^*$ with sufficiently small γ.

Theorem 7. *Let $\boldsymbol{x}_-^* < \boldsymbol{0}$ be the solution to (29) with negative components and let $\alpha > 0$. Furthermore, let A, B have positive diagonal elements and assume that $\gamma > 0$ is chosen sufficiently small to have $I - 2\alpha D_w^-$ positive diagonal matrix and $D_A(I - 2\alpha D_w^-) > -(1-2\alpha)D_B\Theta$. Finally, let $M_A + M_B\Theta$ be an H_+-matrix and let $|L_B|\Theta \ge |L_A|$. If $\langle A\rangle D_w^- + \langle B\rangle\Theta$ is an M-matrix, then $\boldsymbol{x}_-^*$ is an attracting fixed point for the iteration $\boldsymbol{x}^{(k+1)} = \boldsymbol{g}(\boldsymbol{x}^{(k)})$ with the SOR splitting.*

Proof. The proof runs exactly as in the Theorems 5 – 6. In particular, repeating exactly the same passages of the proof of Theorem 6 with $D_s = -I$, we find $|G(\boldsymbol{x}_-^*)| \le \mathcal{M}^{-1}\mathcal{N}$ where $\mathcal{M}$ is an M-matrix, $\mathcal{N}$ is non-negative, and $\mathcal{M} - \mathcal{N} = 2\langle A\rangle D_w^- + 2\langle B\rangle\Theta$. □

Analogous considerations to Remark 8 can be applied to Theorem 7 by requiring that $\langle A\rangle D_w^- + \langle B\rangle\Theta - |L_A|$ be an M-matrix in place of requiring $|L_B|\Theta \ge |L_A|$.

Finally, it is worth noticing that the above convergence analyses can be adapted to study the convergence of other modulus-based matrix splitting methods. As an example, consider the AOR splitting (32). Proceeding as in Theorem 6, we get

$$\begin{aligned}|G(\boldsymbol{x}_+^*)| \le&\Big\langle \frac{1}{\alpha}(D_A + D_B\Theta) - \frac{\beta}{\alpha}(L_A + L_B\Theta)\Big\rangle^{-1}\Big[\Big|\frac{1-2\alpha}{\alpha}D_A + D_B\Theta\Big(\frac{1}{\alpha}I - 2D_w^+\Big)\Big| +\\ &+ \frac{\beta}{\alpha}|L_A - L_B\Theta| + 2\frac{|\alpha-\beta|}{\alpha}|L_A| + 2|L_B|\Theta D_w^+ + 2|U_A| + 2|U_B|\Theta D_w^+|\Big]\\ :=&\mathcal{M}^{-1}\mathcal{N}\end{aligned}$$

If $\beta \le \alpha$, under the same assumptions of Theorem 6 we have

$$\begin{aligned}\mathcal{M} - \mathcal{N} =&\frac{1}{\alpha}(D_A + D_B\Theta) - \frac{\beta}{\alpha}|L_A + L_B\Theta| - \frac{1-2\alpha}{\alpha}D_A - D_B\Theta\Big(\frac{1}{\alpha}I - 2D_w^+\Big)\\ &- \frac{\beta}{\alpha}|L_B\Theta - L_A| - 2\frac{\alpha-\beta}{\alpha}|L_A| - 2|L_B|\Theta D_w^+ - 2|U_A| - 2|U_B|\Theta D_w^+\\ =&2D_A + 2D_B\Theta D_w^+ - 2\frac{\beta}{\alpha}|L_A| - 2|L_A| + 2\frac{\beta}{\alpha}|L_A| - 2|L_B|\Theta D_w^+ - 2|U_A| - 2|U_B|\Theta D_w^+\\ =&2\langle A\rangle + 2\langle B\rangle\Theta D_w^+.\end{aligned}$$

This result is the same as in the proof to Theorem 6, which implies that $\boldsymbol{x}_+^*$ is an attracting fixed point for the AOR method with $\beta \le \alpha$ under the same assumptions of Theorem 6.

If $\beta > \alpha$, proceeding analogously we find

$$\mathcal{M} - \mathcal{N} = 2\langle A\rangle + 2\langle B\rangle\Theta D_w^+ + 4\Big(1 - \frac{\beta}{\alpha}\Big)|L_A|.$$

Hence, $\boldsymbol{x}_+^*$ is an attracting fixed point for the AOR method with $\beta > \alpha$ under the assumptions of Theorem 6 and if $\langle A\rangle + \langle B\rangle\Theta D_w^+ + 2\Big(1 - \frac{\beta}{\alpha}\Big)|L_A|$ is an M-matrix.

The convergence for $\boldsymbol{x}_-^*$ and, possibly, for solutions with different sign patterns can be studied in a similar way.

4.2.3. *Convergence for $p \ne 0$*

When $p \ne 0$, the first p components of the solution are uniquely determined and the Jacobian might not exist at such components. In particular, the fixed point iteration (33) might contain terms $|x_i^*|$ with $x_i^* = 0$, at some $i = 1, \dots, p$. In this case, $\boldsymbol{g}(\boldsymbol{x}^*)$ would not be differentiable at any solution. However, for the remaining $n-p$ components, the considerations performed in the previous subsection still hold. Furthermore, $\boldsymbol{g}(\boldsymbol{x}^*)$ is still well-defined, as the first p components of $\tilde{\boldsymbol{x}}^*$ are zero by definition (15). Hence, $\boldsymbol{g}(\boldsymbol{x})$ still exists at all solutions to (29).

Thus, $\boldsymbol{g}(\boldsymbol{x})$ will still have 2^{n-p} fixed-points, while the issues are limited to the Jacobian matrix. Therefore, we may still think to evaluate whether any fixed point is attracting by referring to a generalized Jacobian. In this regard, for compactness, we first define the following matrices.

- $\partial_B|\boldsymbol{x}|$ denotes the Bouligand sub-differential of $|\boldsymbol{x}|$;
- $\tilde{D}_s(\boldsymbol{x})$ is the diagonal matrix of elements $\tilde{d}_{s_{ii}}=0$, $i=1,\dots,p$ and $\tilde{d}_{s_{ii}}=sign(x_i)$, $i=p+1,\dots,n$;
- $\tilde{D}_w(\boldsymbol{x})$ denotes the non-negative diagonal matrix of elements $\tilde{d}_{w_{ii}}=0$, $i=1,\dots,p$ and $\tilde{d}_{w_{ii}}=\frac{w_i}{4x_i^2}$, $i=p+1,\dots,n$.

With these definitions, we write the Bouligand sub-differential of $\boldsymbol{g}(\boldsymbol{x})$ as the set of matrices

$$\partial_B G(\boldsymbol{x})=(M_A+M_B\Theta)^{-1}\Big[N_A+N_B\Theta+(B\Theta-A)\partial_B|\boldsymbol{x}|-(A+B\Theta)\tilde{D}_w(\boldsymbol{x})-(B\Theta-A)\tilde{D}_w(\boldsymbol{x})\tilde{D}_s(\boldsymbol{x})\Big]. \tag{43}$$

For compactness of notation, in the following we will simply denote $\tilde{D}_w^+:=\tilde{D}_w(\boldsymbol{x}_+^*)$ and $\tilde{D}_w^-:=\tilde{D}_w(\boldsymbol{x}_-^*)$. Furthermore, $\boldsymbol{x}_+^*>\boldsymbol{0}$ now denotes the solution of positive components x_{i+}^* for $i=p+1,\dots,n$, and $\boldsymbol{x}_-^*<\boldsymbol{0}$ denotes the solution of negative components x_{i-}^*, $i=p+1,\dots,n$. The first p components are instead shared by all solutions, and can be positive, negative, or zeros.

Theorem 8. *Let $\boldsymbol{x}_+^*>\boldsymbol{0}$ be the solution to (29) with positive components, for $i=p+1,\dots,n$. Furthermore, let A,B have positive diagonal elements and assume that γ is chosen sufficiently large to have $I-2\alpha\tilde{D}_w^+$ positive diagonal matrix and $D_B\Theta(I-2\alpha\tilde{D}_w^+)>-(1-2\alpha)D_A$. For $i=1,\dots,p$, let $[\langle M_B\rangle\Theta]_{\cdot i}\ge\langle M_A\rangle_{\cdot i}$ and $|N_B\Theta|_{\cdot i}\le|N_A|_{\cdot i}$, where θ_{ii} is an arbitrary positive parameter. For $i=p+1,\dots,n$, instead, let $\theta_{ii}=\gamma^2$ and $[|L_B|\Theta]_{\cdot i}\le|L_A|_{\cdot i}$. Finally, let $M_A+M_B\Theta$ be an H_+-matrix and let C be the matrix of columns $C_{\cdot i}=\langle A\rangle_{\cdot i}$ for $i=1,\dots,p$ and $C_{\cdot i}=[\langle A\rangle+\langle B\rangle\Theta D_w^+]_{\cdot i}$ for $i=p+1,\dots,n$. If C is an M-matrix, then $\boldsymbol{x}_+^*$ is an attracting fixed point for the iteration $\boldsymbol{x}^{(k+1)}=\boldsymbol{g}(\boldsymbol{x}^{(k)})$ with the SOR splitting with $0<\alpha<2$.*

More in general, if $\{\langle A\rangle,\langle A\rangle+\langle B\rangle\Theta D_w^+\}$ is a column M-set [36], then $\boldsymbol{x}_+^*$ is an attracting fixed point for the iteration $\boldsymbol{x}^{(k+1)}=\boldsymbol{g}(\boldsymbol{x}^{(k)})$ with the SOR splitting with $0<\alpha<2$ for any $p=0,1,\dots,n$.

Proof. Assume that $p=n$. In this case, the WHLCP is a standard HLCP and its solution is unique. For consistency of notation, we still name such solution $\boldsymbol{x}_+^*$, but, in general, it will have positive, negative and zero components, as the positivity requirement is present only for $i\ge p+1$. The generalized Jacobian (43) at such solution becomes

$$\begin{aligned}|\partial_B G(\boldsymbol{x}_+^*)|=&\Big|(M_A+M_B\Theta)^{-1}\Big[N_A+N_B\Theta+(B\Theta-A)\partial_B|\boldsymbol{x}_+^*|\Big]\Big|\\ \le&\langle M_A+M_B\Theta\rangle^{-1}\Big|N_A+N_B\Theta+(M_B\Theta-M_A)\partial_B|\boldsymbol{x}_+^*|+(N_A-N_B\Theta)\partial_B|\boldsymbol{x}_+^*|\Big|,\end{aligned}$$

By the definition of the Bouligand sub-differential, any matrix $\hat{J}\in\partial_B|\boldsymbol{x}_+^*|$ satisfies $|\hat{J}|=I$. By this fact and by applying the triangle inequality to the previous expression, any matrix $V^*\in\partial_B G(\boldsymbol{x}_+^*)$ satisfies

$$|V^*|\le\langle M_A+M_B\Theta\rangle^{-1}\Big[|N_A+N_B\Theta|+|M_B\Theta-M_A|+|N_A-N_B\Theta|\Big]$$

Naming $\mathcal{M}:=\langle M_A+M_B\Theta\rangle$ and $\mathcal{N}:=|N_A+N_B\Theta|+|M_B\Theta-M_A|+|N_A-N_B\Theta|$, it can be shown that $\mathcal{M}-\mathcal{N}=2\langle A\rangle$ by proceeding as in [46, Theorem 2.4] (as all the assumptions of such theorem are satisfied for SOR with $0<\alpha<2$). Therefore, $\rho(V^*)<1$ for any $V^*\in\partial_B G(\boldsymbol{x}_+^*)$. In addition, notice that the absolute value function is semi-smooth. Hence, $\boldsymbol{x}_+^*$ is attracting by [47, Proposition 4.1].

In the opposite extreme case $p=0$, we have already found $\mathcal{M}-\mathcal{N}=2\langle A\rangle+2\langle B\rangle\Theta D_w^+$, with $\mathcal{M},\mathcal{N}$ defined as in the proof of Theorem 6.

In all other cases, we will have a mix of the two above extreme situations. In particular, we have that any matrix $V^*\in\partial_B G(\boldsymbol{x}_+^*)$ satisfies

$$|V^*|\le\langle M_A+M_B\Theta\rangle^{-1}\mathcal{N}, \tag{44}$$

where $\mathcal{N}$ is a matrix made of the following columns:

$$\mathcal{N}_{\cdot i}=\Big[|N_A+N_B\Theta|+|M_B\Theta-M_A|+|N_A-N_B\Theta|\Big]_{\cdot i}\ \text{for } i=1,\dots,p;$$

$$\mathcal{N}_{\cdot i}=\Big[\frac{1-2\alpha}{\alpha}D_A+D_B\Theta\Big(\frac{1}{\alpha}I-2D_w^+\Big)+|L_A-L_B\Theta|+2|U_A|+2|L_B|\Theta D_w^++2|U_B|\Theta D_w^+\Big]_{\cdot i}\ \text{for } i=p+1,\dots,n.$$

It follows that the i-th column of $\langle M_A+M_B\Theta\rangle-\mathcal{N}$ is equal to the i-th column of $2\langle A\rangle$ for $i=1,\dots,p$, while it is equal to the i-th column of $2\langle A\rangle+2\langle B\rangle\Theta D_w^+$ for $i=p+1,\dots,n$. Thus, if $\{\langle A\rangle,\langle A\rangle+2\langle B\rangle\Theta D_w^+\}$ is a column M-set, $\rho(V^*)<1$ for any $V^*\in\partial_B G(\boldsymbol{x}_+^*)$ and $\boldsymbol{x}_+^*$ is an attracting fixed point for any $p=0,\dots,n$. □

A similar result can be obtained with regard to $\boldsymbol{x}_-^*$ requiring that γ is sufficiently small. Furthermore, the same considerations can be applied also to other splittings. For instance, the considerations made with regard to the AOR splitting with $p = 0$ can be extended to the case with $p \neq 0$ following the line of reasoning of the proof to Theorem 8.

4.2.4. An important special case: the WLCP

If either A or B is the identity matrix, the WHLCP reduces to a WLCP. This case is contained in the previous analysis. Nonetheless, the convergence theory is significantly simplified. Indeed, several assumptions of the previous theorems are immediately satisfied. With no loss of generality, let us then consider $B = I$. Theorem 6 becomes as follows. Most notably, the requirement $|L_B|\Theta \le |L_A|$ is necessarily satisfied. Hence, the local convergence to $\boldsymbol{x}_+^*$ is ensured by a sufficiently large γ (provided that A satisfies some other assumptions).

Theorem 9. *Let $\boldsymbol{x}_+^* > \boldsymbol{0}$ be the solution to (29) with positive components and let $\alpha > 0$. Furthermore, let A have positive diagonal elements and assume that γ is sufficiently large to have $I - 2\alpha D_w^+$ positive diagonal matrix and $\Theta(I - 2\alpha D_w^+) > -(1-2\alpha)D_A$. Finally, let $M_A + \frac{1}{\alpha}\Theta$ be an H_+-matrix. If $\langle A\rangle + \Theta D_w^+$ is an M-matrix, then $\boldsymbol{x}_+^*$ is an attracting fixed point for the iteration $\boldsymbol{x}^{(k+1)} = \boldsymbol{g}(\boldsymbol{x}^{(k)})$ with the SOR splitting.*

Analogous results can easily be obtained for all the other convergence theorems. In this regard, it is sufficient to remove the assumptions that are automatically satisfied when $B = I$. Clearly, such theorems can be further simplified if $B = I$ is not split. This would be the most common setting for modulus-based methods for WLCPs, and is contained in the previous analysis.

5. Numerical experiments

5.1. Setting of the experiments

We solve several test problems by the modulus-based non-smooth Newton's (MNN), modulus-based Jacobi (MJ), modulus-based Gauss-Seidel (MGS), and modulus-based SOR (MSOR) methods. We compare these methods with [11], which studied a smoothing Newton method based on a special instance of the weighted smoothing function introduced in [10]. In the following, we denote the smoothing Newton method [11] by SN. To make a fair and meaningful comparison, we solve the numerical experiments used in [11] (among others).

In particular, we consider numerical experiments defined by the following matrices.

- **Problem 1** uses the same matrices as [11, Example 6.1] (which were earlier used also in [18] for non-weighted HLCPs). In particular, let $n = m^2$. We define $S \in \mathbb{R}^{m\times m}$ as the tridiagonal matrix of elements $S = tridiag\{-1, 4, -1\}$. Then, we define the matrices of the WHLCP (1) as

  $A^{n\times n}$ block-tridiagonal matrix of blocks $\{-I, S, -I\}$;

  $B^{n\times n}$ block-diagonal matrix of blocks $\{S\}$.
- **Problem 2** uses the same matrices as [11, Example 6.2]. Let n be an integer divisible by 4 and set $n_i = n/4$ for $i = 1, \dots, n$. Let M_i be a random matrix of order n_i, $i = 1, \dots, n$. Problem 2 is a WHLCP as (1) where A is the block diagonal matrix of diagonal blocks $\frac{M_1 M_1^T}{\|M_1 M_1^T\|}, \dots, \frac{M_4 M_4^T}{\|M_4 M_4^T\|}$, while $B = I$. Notice that we are here considering the special case of a weighted (non-horizontal) LCP.
- **Problem 3** uses the same matrices as [11, Example 6.3]. Let $U, V \in \mathbb{R}^{n\times n}$ be random matrices. We set $A = \frac{UU^T}{\|UU^T\|}$ and $B = I + \frac{VV^T}{\|VV^T\|}$.
- **Problem 4** is a WHLCP of non-symmetric matrices of order $n = m^2$. We introduce this numerical experiment because all test problems of [11] used symmetric matrices. In particular, define $\hat{S} \in \mathbb{R}^{m\times m}$ as the tridiagonal matrix of elements $\hat{S} = tridiag\{-1.5, 4, -0.5\}$. Then, we define the matrices of the WHLCP (1) as

  $A^{n\times n}$ block-tridiagonal matrix of blocks $\{-1.5I, \hat{S}, -0.5I\}$;

  $B^{n\times n}$ block-diagonal matrix of blocks $\{\hat{S}\}$.

  Notice that Problem 4 can be viewed as a non-symmetric variant of Problem 1 and is based on the matrices of [18].

All experiments are performed on an M1 MacBook Pro (16", 2021) with 32 GBs of RAM, using numerical codes implemented in Matlab R2023b. When not otherwise specified, we choose $\gamma = 1.5$ and $\Theta = \gamma^2 I$. In the computation of the generalized Jacobian in MNN, we consider any value of $x_i^{(k)}$ in $[-10^{-10}, 10^{-10}]$ to be numerically zero. Other parameters such as $\boldsymbol{w}, \boldsymbol{x}^{(0)}, \dots$ are specified in the analysis of each experiment.

The MNN, MJ, MGS, and MSOR methods are stopped when the Euclidean norm of the residual of the modulus-based system (14) is smaller than a tolerance $tol = 10^{-8}$. As SN is not modulus-based, we cannot use the same stopping criterion as in the modulus-based methods. Hence, we use the same stopping criterion as in [11] and require that the Euclidean norm of the smoothing function used

Table 1
Results of the numerical solution of Problem 1 (left) and Problem 4 (right).

m	method	it	t	$conv$	e	m	method	it	t	$conv$	e
	MNN	6	0.01	4.7E-14	1.2E-14		MNN	6	0.02	2.8E-13	5.2E-14
	MJ	263	0.03	9.3E-9	1.1E-8		MJ	163	0.03	9.4E-9	5.6E-9
20	MGS	155	0.02	9.0E-9	1.1E-8	20	MGS	94	0.02	1.0E-8	7.9E-9
	MSOR	124	0.02	9.9E-9	1.2E-8		MSOR	78	0.02	7.6E-9	6.2E-9
	SN	5	0.01	5.9E-12	2.7E-11		SN	5	0.01	3.2E-12	1.2E-11
	MNN	5	0.04	7.2E-9	1.3E-9		MNN	6	0.06	3.7E-13	1.1E-13
	MJ	277	0.18	9.6E-9	1.3E-8		MJ	266	0.20	9.9E-9	1.2E-8
75	MGS	171	0.15	9.6E-9	1.3E-8	75	MGS	132	0.12	8.9E-9	1.1E-8
	MSOR	160	0.13	9.5E-9	1.2E-8		MSOR	106	0.11	9.7E-9	1.2E-8
	SN	5	0.54	5.5E-11	2.7E-10		SN	5	0.53	5.1E-11	2.5E-10
	MNN	6	0.27	3.8E-13	1.2E-13		MNN	6	0.27	4.9E-13	1.4E-13
	MJ	311	0.90	9.5E-9	1.3E-8		MJ	306	0.98	9.5E-9	1.2E-8
150	MGS	187	0.61	9.8E-9	1.3E-8	150	MGS	140	0.54	8.7E-9	1.1E-8
	MSOR	152	0.51	9.5E-9	1.3E-8		MSOR	113	0.40	8.2E-9	1.1E-8
	SN	5	12.1	1.2E-10	6.1E-10		SN	5	12.2	1.2E-10	5.8E-10

by [11] is smaller than tol. We then account for the possible differences in the used stopping criteria in the later discussion. All the other parameters of SN are as in [11, Sec. 6.2].

5.2. Verification of the methods on symmetric and non-symmetric matrices

We verify the methods by solving the problems 1 and 4 considering dimensions up to $n = m^2 = 150^2$. The terms $\boldsymbol{w}$ and $\boldsymbol{q}$ are obtained by imposing the solution $\boldsymbol{\xi}^* = 4\boldsymbol{e}$, $\boldsymbol{\eta}^* = \boldsymbol{e}$. Hence, $\boldsymbol{w} = 4\boldsymbol{e}$ and $p = 0$. The starting vector is chosen as

$$x_i^{(0)} = (-1)^{i+1}, \text{ for } i = 1, \dots, n.$$

This choice has alternating signs to avoid encouraging either $\boldsymbol{x}_+$ or $\boldsymbol{x}_-$ in the solution of the modulus-based system. For consistency, the starting iterate is the same for all methods. Notice, however, that for splitting methods it would make sense to use a positive or negative starting iterate, depending on whether γ is chosen to be large or close to zero, respectively. In SN, the starting vector is given by the vectors $\{\boldsymbol{\xi}^{(0)}, \boldsymbol{\eta}^{(0)}\}$ which are computed from $\boldsymbol{x}^{(0)}$ according to (16) with $\Gamma = \Theta = I$.

As regards the convergence of the methods, $\{A, B\}$ satisfy the column $\mathcal{W}$-property. Hence, the MNN and SN methods are locally convergent. As regards the modulus-based matrix splitting methods, first notice that A, B of Problem 1 are irreducibly diagonally dominant and have the same sign pattern. Hence, their sum is again irreducibly diagonally dominant. Consequently, it is easy to verify that $\langle A\rangle + \langle B\rangle \Theta D_w$ is an M-matrix at the solution. Furthermore, for $\gamma = 1.5$, the solution is $x_{i+}^* = 3$, $x_{i-}^* = -\frac{1}{3}$ for $i = 1, \dots, n$. Thus, D_w^+ has diagonal elements $d_{w_{ii}}^+ = \frac{1}{9}$, $i = 1, \dots, n$. Therefore, $(I - 2D_w^+) = \frac{7}{9} I$ and $D_B \Theta (I - 2D_w^+) = 7I > D_A = 4I$. Hence, we expect $\boldsymbol{x}_+^* = 3\boldsymbol{e}$ to be an attracting fixed point for the MJ method. As regards MGS and MSOR, similar considerations apply. Indeed, for $\alpha = 1.1$, we have $D_B \Theta (I - 2\alpha D_w^+) = 6.8I > (1 - 2\alpha) D_A = 4.4I$ and, using Remark 8, $\langle A\rangle + \langle B\rangle \Theta D_w^+ - |L_B|\Theta$ is a block-tridiagonal matrix whose diagonal blocks are tridiagonal matrices of elements $tridiag(-3.5, 5, -1.25)$. The lower- and upper-diagonal blocks are $-I$. Such matrix is an M-matrix, as can be verified in small dimensions. Hence, $\boldsymbol{x}^* = 3\boldsymbol{e}$ is an attracting fixed point for MGS and MSOR as well. Similar considerations can be applied to Problem 4.

The results for the problems 1 and 4 are reported in Table 1. In the tables, it denotes the number of iterations, t the computational time (in seconds), $conv$ the value of the stopping criterion at convergence, and e the Euclidean norm of the maximum between the error on $\boldsymbol{\xi}$ and on $\boldsymbol{\eta}$.

All methods successfully solve the problems. Nonetheless, significant differences are apparent. For instance, MNN converges in the least time in most of the problems. Modulus-based matrix splitting methods trigger significantly more iterations, which, however, are computationally less onerous than a Newton's iteration. Hence, computational times remain small even when the dimension is large. On top of this, modulus-based matrix splitting methods can easily exploit the sparsity of A, B and MJ could easily be parallelized. Hence, they remain an attractive alternative to MNN, especially if parallel computing or more convoluted splittings were used. Finally, SN takes the least iterations to converge, but the computational times are much higher than the other methods when the dimension of the problem is large. In some cases, the difference is almost two orders of magnitude. The main reason is that SN [11] actually solves a $(2n + 1) \times (2n + 1)$ system at every Newton iteration. All the modulus-based methods, on the other hand, solve systems of order n. Thus, performing an iteration of SN becomes increasingly onerous much quicker than in all the other considered methods. The difference is so significant that the effect of having used a slightly different stopping criterion is almost inconsequential. Indeed, if we were to perform just 4 iterations of SN, we would still be far from the solution (for instance, for Problem 1 the Euclidean norm of the error on $\boldsymbol{\xi}$ after 4 SN iterations is in the order of 10^{-4}, while at convergence of MJ such norm is in the order of 10^{-8}), but the computational time would still remain in the order of 10 seconds. Hence, the proposed methods lead to a significant reduction in computational cost and complexity with respect to recent, state-of-the art approaches.

5.3. Convergence of MNN and modulus-based matrix splittings

Let us further analyze the convergence of MNN and of modulus-based matrix splitting methods. We consider the problems in the previous subsection and focus, for simplicity, on $m = 75$ and on the MJ splitting. Similar conclusions apply to the other matrix splitting methods and to problems of different dimensions.

First, let us print the computed solutions to Problem 1, $m = 75$. With MNN, we get

$$\boldsymbol{x}_{MNN}^{(end)} = (1.999999999997050, -0.499999999924515, 1.999999999995934, -0.499999999992846, \dots)^T$$

while with MJ (and dividing by γ in accordance with Lemma 3 to obtain values readily comparable with MNN) we obtain

$$\frac{\boldsymbol{x}_{MJ}^{(end)}}{\gamma} = (1.999999999999382, 1.999999999998769, 1.999999999998163, 1.999999999997571, \dots)^T.$$

We notice that MNN tends to converge to the solution that has the same sign pattern as the starting iterate. This is in agreement with Theorem 3, by which the method is convergent to any solution of the modulus-based system. Thus, it makes sense to achieve convergence to the solution with the same sign pattern of the starting iterate (although this is not necessarily required by the convergence theory). On the other hand, the MJ method converges to the positive solution of the modulus-based system. This is in accordance with the convergence Theorem 4, according to which that is the attracting fixed point in the numerical setting. In this context, it is interesting to see what happens when we change γ. For instance, let us set $\gamma = 0.04$. As γ is now small, we expect the negative solution $\boldsymbol{x}_-^*$ to become attracting. Indeed, the conditions of Theorem 5 are satisfied: with $\gamma = 0.04$, we have $x_{i-}^* = -\frac{25}{2}$ and D_w^- has diagonal elements $d_{ii_w}^- = \frac{4}{625}$, $i = 1, \dots, n$. This implies $(I - 2D_w^-) = 0.9872I$ positive diagonal matrix and $D_A(I - 2D_w^-) = 3.9488I > D_B\Theta = 0.0064I$ (with, in addition, $\langle A\rangle D_w^- + \langle B\rangle\Theta$ M-matrix for any positive diagonal matrix D_w^-, Θ). If we solve the problem by MJ, we indeed find the following solution (which we have multiplied by γ in accordance with Lemma 3 to obtain values readily comparable with those computed by MNN)

$$\gamma \cdot \boldsymbol{x}_{MJ}^{(end)} = (-0.500000000017337, -0.500000000034363, -0.500000000050757, -0.500000000066231, \dots)^T.$$

As expected, the method converged to the negative solution. However, the convergence was slower, and we even had to increase the maximum number of iterations to achieve the same precision of the other experiments. Further experiments show that we still have convergence if we reduce γ, as detailed in the next paragraph.

In this context, some remarks on the convergence basin of the methods are in order. In particular, in accordance with Remark 5, the fact that MNN can converge to the solution with the same sign pattern as the starting guess further shows that we do not actually need to change signs during the MNN iterations (at least, not if we are in a neighborhood of the solution itself). This helps avoid iterates where $\tilde{\boldsymbol{x}}$ is not defined and can explain why the MNN method always converges in our experiments. Instead, in MJ we have just one attracting fixed point. If, as in the above analysis, we start from a starting guess with alternating signs and the convergence is expected to a positive (or negative) solution, some components will necessarily have to change sign at some iteration. In this case, numerical issues would occur if, at some iterate, we got $x_i^{(k)} \approx 0$ for some $i = p+1, \dots, n$. Nonetheless, it is worth noticing the following:

- it is reasonably simple to avoid these issues. Indeed, depending on how we choose γ, we know which solution tends to become attracting according to the Theorems 4, 5 and, more in general, to the Theorems 6, 7. Hence, it is sufficient to choose a starting iterate compatible with the performed choice of γ. This has significant effects. For instance, in the previous paragraph we mentioned some convergence issues and slow-downs of MJ when γ is small. However, it is sufficient to choose the starting iterate as, for instance, $\boldsymbol{x}^{(0)} = -\boldsymbol{e}$ to achieve convergence in a much larger basin. E.g., with $\gamma = 0.2$ we achieve convergence to $\boldsymbol{x}_-^*$ with a speed and general behavior comparable to what we observed in Section 5.2 with regard to $\boldsymbol{x}_+^*$. Our choice of alternating signs was earlier performed only for comparison with the other methods.
  It is also possible to implement a safeguard against zero iterates, based on the fact that the solution cannot have null components where the weight is non-zero;
- the convergence basin with a general starting iterate is not banal. Nonetheless, numerically the convergence basin appears to be large, at least for the considered experiments. Indeed, we did not observe convergence issue to the attracting fixed point.

Finally, we notice that the above remarks have a visible consequence for the convergence curve of the methods. In particular, we again use the sign-alternating starting iterate. Fig. 1 shows that the convergence of MNN is monotone, while the convergence of MJ is not. In the figure, this is highlighted by the logarithmic plot, which we use for visualization purposes. The non-monotonicity is consistent with the previous analysis and is produced by the competing effects of $\boldsymbol{x}^{(k)}$ and $\tilde{\boldsymbol{x}}^{(k)}$ when there is a change of sign (i.e., $x_i^{(k)}$ is small when it approaches zero, while $\tilde{x}_i^{(k)}$ increases). Such non-monotone convergence is particularly evident in MJ, while it seems to occur less often in MGS and MSOR. Nonetheless, it is still visible in both these methods for Problem 1 with $m = 75$.

5.4. Comparison with the literature on random problems

We here complete the comparison with the problems in the literature. In this regard, the solution of Problem 1 in the previous sections already presents a comparison with [11, Example 6.1]. Let us then pass to the problems 2 and 3, which are made of random

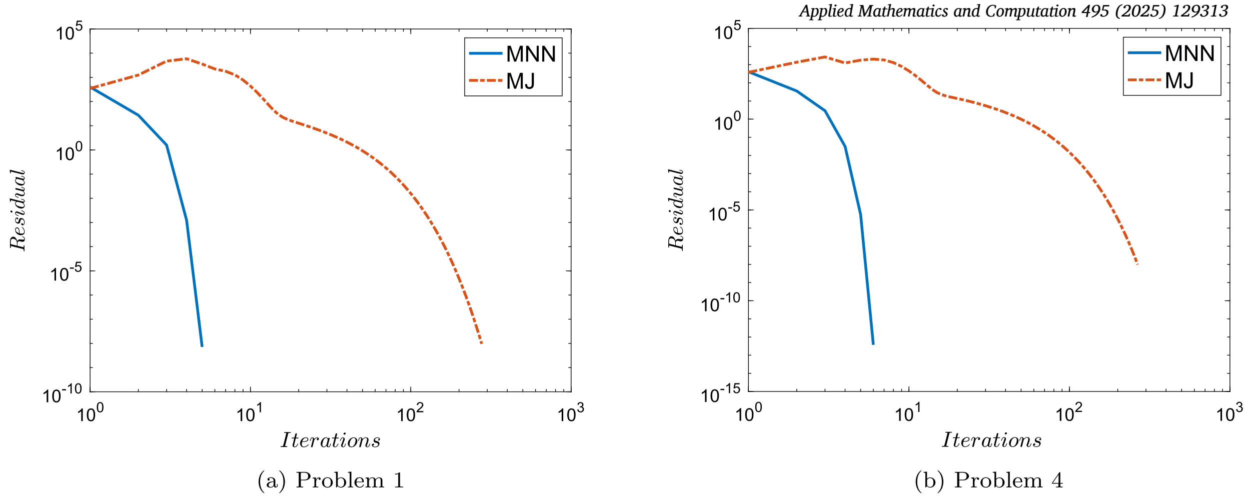

Fig. 1. Comparison between the convergence curves of MNN and of MJ, $m = 75$.

Table 2
Results from solving 10 instances of Problem 2 and Problem 3 with random matrices and vectors as in [11]. Size $n = 5000$.

Problem	method	$\overline{it}$	$\overline{t}$	$\overline{conv}$	$It_{\max}$	$t_{\max}$	$conv_{\max}$
2	MNN	18	11.1	2.7E-8	20	11.9	2.2E-7
	SN	6.8	20.2	1.8E-7	7	21.9	7.7E-7
3	MNN	17.7	10.9	1.3E-7	20	12.7	7.0E-7
	SN	16.8	50.8	1.6E-7	21	63.7	7.0E-7

matrices. The known vector $\boldsymbol{q}$ and the weight $\boldsymbol{w}$ are chosen randomly according to the examples 6.2 and 6.3 of [11]. Finally, the starting iterates are also chosen randomly, as in many sets of experiments of [11]. For consistency with [11], we also set the tolerance to $tol = 10^{-6}$.

Let us solve a set of 10 problems of this kind. The dimension is set to $n = 5000$, which is the maximum size that was considered in [11]. The results are reported in Table 2, where we compare SN and MNN. In this comparison, we do not include the splitting methods because their convergence conditions are not satisfied (namely, the M-matrix condition). In the table, $\overline{it}$, $\overline{t}$, and $\overline{conv}$ represent the mean number of iterations, computational times, and final values of the stopping criterion, respectively, over all the 100 random experiments. To verify whether the algorithms always converge, we also report $It_{\max}$, $t_{\max}$, and $conv_{\max}$, which denote the maximum number of iterations, the maximum computational time, and the maximum value of the stopping criterion, respectively.

Both methods can solve all the numerical experiments. This shows the capabilities of MNN in a wide range of random problems and further demonstrates that it is competitive with solution methods in the literature. Furthermore, we again notice that MNN converges faster than SN. As previously discussed, this can be explained by the fact that MNN solves a smaller system than SN. Finally, it is worth noticing that Problem 2 is a WLCP, which is contained in our analysis as a special case of WHLCP. Hence, the above experiments show that we can effectively solve this kind of problems as well. Similarly, Problem 3 shows the effectiveness of MNN when full matrices are used.

5.5. Weight with zero values

Finally, we verify that the analyzed methods can solve problems with $p \neq 0$. In this regard, we consider the problems 1 and 4 within the same numerical setting of Section 5.2. However, we slightly modify the test problems by setting $p = 100$ and imposing $\eta_i^* = 0$ for $i = 1, \dots, 100$ in one case, and $\xi_i^* = \eta_i^* = 0$ for $i = 1, \dots, 100$ in another. Notice that the first case corresponds to $w_i = 0$ for $i = 1, \dots, 100$, but the solution with $\gamma = 1.5$ is $x_i^* = 0.75 \neq 0$ for $i = 1, \dots, 100$. In the second case, instead, we have both $w_i = 0$ and $x_i^* = 0$ for $i = 1, \dots, 100$. Hence, these two experiments cover different cases that may occur. The results for $m = 150$ are reported in Table 3.

As expected, Table 3 shows that the proposed methods converge even when the weight and/or the solution are zero in some components. Interestingly, even the number of iterations is similar to the case with $p = 0$ studied in Table 1. Hence, it seems that having a set of zero weights does not deeply change the behavior of the methods for these problems.

Table 3
Results when $\eta_i^* = 0$ (left) and $\xi_i^* = \eta_i^* = 0$ (right), $i = 1, \ldots, 100$, in the Problems 1 and 4 with $n = 150^2$.

Problem	method	*it*	*t*	*conv*	e
1	MNN	6	0.28	3.8E-13	1.2E-13
	MJ	311	0.99	9.6E-9	1.3E-8
	MGS	187	0.63	9.9E-9	1.3E-8
	MSOR	152	0.55	9.6E-9	1.3E-8
	SN	6	14.0	3.3E-11	4.4E-11
4	MNN	6	0.28	4.0E-13	1.4E-13
	MJ	306	0.91	9.6E-9	1.2E-8
	MGS	140	0.50	8.8E-9	1.1E-8
	MSOR	113	0.43	8.2E-9	1.1E-8
	SN	6	15.4	3.3E-11	4.8E-11

Problem	method	*it*	*t*	*conv*	e
1	MNN	6	0.27	3.8E-13	1.2E-13
	MJ	314	0.97	9.5E-9	1.3E-8
	MGS	187	0.55	9.8E-9	1.3E-8
	MSOR	152	0.49	9.5E-9	1.3E-8
	SN	5	11.9	1.2E-10	6.0E-10
4	MNN	6	0.29	4.1E-13	1.4E-13
	MJ	307	0.87	9.8E-9	1.3E-8
	MGS	140	0.48	8.7E-9	1.1E-8
	MSOR	113	0.46	8.1E-9	1.1E-8
	SN	5	12.1	1.2E-10	5.8E-10

6. Conclusions

We have reformulated the WHLCP as a nonlinear system which contains absolute value functions. When the solution to the WHLCP is unique, such modulus-based system has 2^{n-p} solutions, where p denotes the number of zero components of the given weight vector. Nonetheless, all modulus-based solutions map to the same solution of the WHLCP. We have further characterized the modulus-based solutions, studying especially their sign pattern. In particular, we have noticed that, for $i = p+1, \ldots, n$, there exist exactly one positive and one negative i-th component of the solution to the modulus-based system. We have then exploited the modulus-based formulation to introduce new solution methods that solve the modulus-based system in place of the WHLCP. In particular, we have introduced a non-smooth Newton's method and matrix splitting methods, and we have analyzed their convergence. We have found that the modulus-based non-smooth Newton's method is convergent in a neighborhood of every solution to the modulus-based system. On the other hand, only one modulus-based solution is an attracting fixed-point for the modulus-based matrix splitting methods. Nonetheless, we can guide which solution is attracting by tuning an arbitrary positive parameter, γ. Indeed, we have proved that a large γ makes the positive solution $\boldsymbol{x}_+^*$ attracting; γ close to zero makes the negative solution $\boldsymbol{x}_-^*$ attracting. The actual values that ensure such "attractingness" depend on the solution itself, but our analysis provides insights on the selection of γ. Finally, we have solved several numerical experiments with symmetric and non-symmetric matrices. In this context, we have shown the computational advantage with respect to state-of-the-art methods. In particular, while methods such as the one in [11] solve sub-problems of order $2n+1$, the modulus-based system is of order n. This is reflected by the computational times that we have registered. In the numerical experiments, we have also provided insights on the convergence and on the attraction basin of the methods. These results have confirmed the theoretical convergence analysis. Notably, we have observed that the modulus-based non-smooth Newton's method tends to preserve the sign pattern of the starting iterate, while for modulus-based matrix splitting methods only one solution is attracting. We have also confirmed the theoretical predictions on which solution is attracting depending on the choice of the parameter γ. Based on this, we have argued that it makes sense to choose as the starting iterate of the modulus-based matrix splitting methods a vector with the same sign pattern as the solution that is expected to be attracting.

Several future developments are possible. For instance, we have solved the modulus-based system by relatively simple methods. More advanced strategies could be used, for instance by introducing controls on the Newton's step or by using more convoluted splittings and multi-splittings. Safeguards against the risk of dividing by zero in numerical methods could also be implemented. Our analysis of the modulus-based framework should enable a ready adaptation of the convergence analysis to all such methods. Parallelization (either directly in the MJ method or using multi-splittings) is another possible future development. Lastly, the convergence of the modulus-based methods could be further analyzed and expanded.

Funding

The author acknowledges the support of the research fund FAR 2023 DIP - Fondo di Ateneo per la Ricerca per il finanziamento di progetti di ricerca dipartimentali (Department of Engineering "Enzo Ferrari", University of Modena and Reggio Emilia).

Data availability

Data will be made available on request.

References

[1] F.A. Potra, Weighted complementarity problems - a new paradigm for computing equilibria, SIAM J. Optim. 22 (4) (2012) 1634–1654.
[2] X. Chi, M.S. Gowda, J. Tao, The weighted horizontal linear complementarity problem on a Euclidean Jordan algebra, J. Glob. Optim. 73 (2019) 153–169.
[3] K.M. Anstreicher, Interior-point algorithms for a generalization of a linear programming and weighted centering, Tech. Rep., Department of Management Sciences, University of Iowa, 2011.
[4] F.A. Potra, Sufficient weighted complementarity problems, Comput. Optim. Appl. 64 (2016) 467–488.
[5] S. Asadi, Z. Darvay, G. Lesaja, N. Mahdavi-Amiri, F.A. Potra, A full-Newton step interior-point method for monotone weighted linear complementarity problems, J. Optim. Theory Appl. 186 (2020) 864–878.

[6] X. Chi, G. Wang, A full-Newton step infeasible interior-point method for the special weighted linear complementarity problem, J. Optim. Theory Appl. 190 (2021) 108–129.
[7] B. Kheirfam, Complexity analysis of a full-Newton step interior-point method for monotone weighted linear complementarity problems, J. Optim. Theory Appl. 202 (2024) 133–145.
[8] Z. Jian, A smoothing Newton algorithm for weighted linear complementarity problem, Optim. Lett. 10 (2016) 499–509.
[9] J. Tang, A variant nonmonotone smoothing algorithm with improved numerical results for large-scale LWCPs, Comput. Appl. Math. 37 (2018) 3927–3936.
[10] J. Tang, H. Zhang, A nonmonotone smoothing Newton algorithm for weighted complementarity problem, J. Optim. Theory Appl. 189 (2021) 679–715.
[11] J. Tang, J. Zhou, The solvability of weighted complementarity problems and a smoothing Newton algorithm under the local error bound, Optimization (2023), https://doi.org/10.1080/02331934.2023.2269943, in press.
[12] M. Kojima, S. Mizuno, T. Noma, Limiting behavior of trajectories generated by a continuation method for monotone complementarity problems, Math. Oper. Res. 15 (4) (1990) 662–675.
[13] A. Yoshise, Interior point trajectories and a homogeneous model for nonlinear complementarity problems over symmetric cones, SIAM J. Optim. 17 (2006) 1129–1153.
[14] W. van Bokhoven, Piecewise-Linear Modelling and Analysis, Proefschrift, Eindhoven, 1981.
[15] K. Murty, Linear Complementarity, Linear and Nonlinear Programming, Heldermann, Berlin, 1988.
[16] J.-L. Dong, M.-Q. Jiang, A modified modulus method for symmetric positive-definite linear complementarity problems, Numer. Linear Algebra Appl. 16 (2009) 129–143.
[17] Z.-Z. Bai, Modulus-based matrix splitting iteration methods for linear complementarity problems, Numer. Linear Algebra Appl. 17 (2010) 917–933.
[18] F. Mezzadri, E. Galligani, Modulus-based matrix splitting methods for horizontal linear complementarity problems, Numer. Algorithms 83 (1) (2020) 201–219.
[19] Z. Xia, C.-L. Li, Modulus-based matrix splitting iteration methods for a class of nonlinear complementarity problem, Appl. Math. Comput. 271 (2015) 34–42.
[20] N. Huang, C.-F. Ma, The modulus-based matrix splitting algorithms for a class of weakly nonlinear complementarity problems, Numer. Linear Algebra Appl. 23 (3) (2016) 558–569.
[21] F. Mezzadri, E. Galligani, Modulus-based matrix splitting methods for a class of horizontal nonlinear complementarity problems, Numer. Algorithms 87 (1) (2021) 667–687.
[22] J.-T. Hong, C.-L. Li, Modulus-based matrix splitting iteration methods for a class of implicit complementarity problems, Numer. Linear Algebra Appl. 23 (2016) 629–641.
[23] S.-L. Wu, P. Guo, Modulus-based matrix splitting algorithms for the quasi-complementarity problems, Appl. Numer. Math. 132 (2018) 127–137.
[24] F. Mezzadri, A modulus-based formulation for the vertical linear complementarity problem, Numer. Algorithms 90 (2022) 1547–1568.
[25] S. Xie, Z.-P. Yang, H. Xu, A modulus-based matrix splitting method for the vertical nonlinear complementarity problem, J. Appl. Math. Comput. 69 (2023) 2987–3003.
[26] Z.-Z. Bai, L.-L. Zhang, Modulus-based synchronous multisplitting iteration methods for linear complementarity problems, Numer. Linear Algebra Appl. 20 (2013) 425–439.
[27] F. Mezzadri, Modulus-based synchronous multisplitting methods for solving horizontal linear complementarity problems on parallel computers, Numer. Linear Algebra Appl. 27 (e2319) (2020) 1–15.
[28] C.-L. Li, J.-T. Hong, Modulus-based synchronous multisplitting iteration methods for an implicit complementarity problem, East Asian J. Appl. Math. 7 (2017) 363–375.
[29] C. Liu, C.-L. Li, Synchronous and asynchronous multisplitting iteration schemes for solving mixed linear complementarity problems with H-matrices, J. Optim. Theory Appl. 171 (2016) 169–185.
[30] H.-R. Xu, R. Chen, S.-L. Xie, L. Wu, Modulus-based multisplitting iteration methods for a class of nonlinear complementarity problems, East Asian J. Appl. Math. 8 (2018) 519–530.
[31] H. Zheng, Y. Zhang, X. Lu, S. Vong, Modulus-based synchronous multisplitting iteration methods for large sparse vertical linear complementarity problems, Numer. Algorithms 93 (2023) 711–729.
[32] Y. Zhang, H. Zheng, X. Lu, S. Vong, Modulus-based synchronous multisplitting iteration methods without auxiliary variable for solving vertical linear complementarity problems, Appl. Math. Comput. 458 (2023) 128248.
[33] A. Berman, R.J. Plemmons, Nonnegative Matrices in the Mathematical Sciences, Classics in Applied Mathematics, vol. 9, SIAM, Philadelphia, 1994.
[34] Z.-Z. Bai, On the convergence of the multisplitting methods for the linear complementarity problem, SIAM J. Matrix Anal. Appl. 21 (1999) 67–78.
[35] R. Sznajder, M.S. Gowda, Generalizations of P_0- and P-properties; extended vertical and horizontal linear complementarity problems, Linear Algebra Appl. 223–224 (1995) 695–715.
[36] F. Mezzadri, E. Galligani, A generalization of irreducibility and diagonal dominance with applications to horizontal and vertical linear complementarity problems, Linear Algebra Appl. 621 (2021) 214–234.
[37] Y. Cui, J.-S. Pang, Modern Nonconvex Nondifferentiable Optimization, SIAM, Philadelphia, 2021.
[38] F.H. Clarke, Optimization and Nonsmooth Analysis, John Wiley and Sons, New York, 1983.
[39] L. Qi, J. Sun, A nonsmooth version of Newton's method, Math. Program. 58 (1993) 353–367.
[40] J.M. Ortega, Numerical Analysis: A Second Course, Computer Science and Applied Mathematics, Academic Press, New York and London, 1972.
[41] H. Zheng, W. Li, The modulus-based nonsmooth Newton's method for solving linear complementarity problems, J. Comput. Appl. Math. 288 (2015) 116–126.
[42] H. Zheng, S. Vong, The modulus-based nonsmooth Newton's method for solving a class of nonlinear complementarity problems of P-matrices, Calcolo 55 (37) (2018) 1–17.
[43] F. Mezzadri, E. Galligani, A modulus-based nonsmooth Newton's method for solving horizontal linear complementarity problems, Optim. Lett. 15 (2021) 1785–1798.
[44] A. Frommer, D.B. Szyld, H-splittings and two-stage iterative methods, Numer. Math. 63 (1992) 345–356.
[45] R. Horn, C. Johnson, Matrix Analysis, 2nd edition, Cambridge University Press, 2013.
[46] H. Zheng, S. Vong, On convergence of the modulus-based matrix splitting iteration method for horizontal linear complementarity problems of H_+-matrices, Appl. Math. Comput. 369 (2020) 1–6.
[47] J.-J. Zhang, The relaxed nonlinear PHSS-like iteration method for absolute value equations, Appl. Math. Comput. 265 (2015) 266–274.