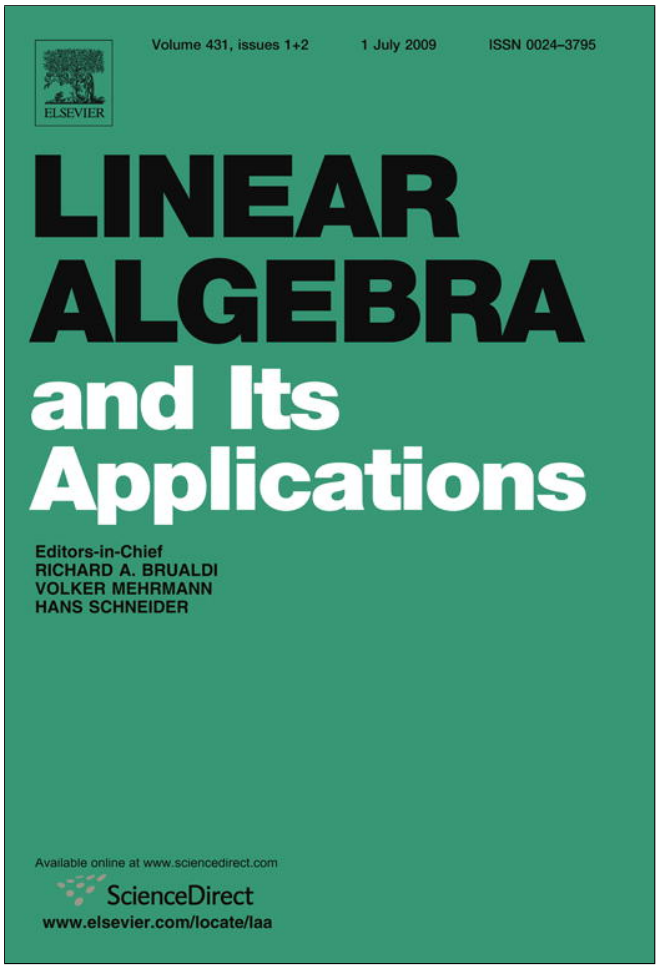

Linear Algebra and its Applications 431 (2009) 197–210

Contents lists available at ScienceDirect

Linear Algebra and its Applications

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

Nonstationary Extrapolated Modulus Algorithms for the solution of the Linear Complementarity Problem

A. Hadjidimos [a,*], M. Tzoumas [b]

[a] *Department of Computer and Communication Engineering, University of Thessaly, 10 Iasonos Street, GR-383 33 Volos, Greece*

[b] *Department of Mathematics, University of Ioannina, GR-451 10 Ioannina, Greece*

ARTICLE INFO

Article history:
Received 5 August 2008
Accepted 11 February 2009
Available online 27 March 2009

Submitted by V. Mehrmann

AMS classification:
Primary 65F10

Keywords:
LCP
P-matrices
Real symmetric positive definite matrices
Iterative schemes
Extrapolation
(Block) Modulus Algorithm

ABSTRACT

The *Linear Complementarity Problem* (LCP) has many applications as, e.g., in the solution of Linear and Convex Quadratic Programming, in Free Boundary Value problems of Fluid Mechanics, etc. In the present work we assume that the matrix coefficient $M \in \mathbb{R}^{n,n}$ of the LCP is *symmetric positive definite* and we introduce the *(optimal) nonstationary extrapolation* to improve the convergence rates of the well-known *Modulus Algorithm* and *Block Modulus Algorithm* for its solution. Two illustrative numerical examples show that the *(Optimal) Nonstationary Extrapolated Block Modulus Algorithm* is far better than all the previous similar Algorithms.

1. Introduction and preliminaries

The Linear Complementarity Problem (LCP) is met in many practical applications. For example, in linear and convex quadratic programming, in a problem of the theory of games [14,6], in problems in fluid mechanics [8], in problems in economics [19,13], etc. For more applications see, e.g., [16,7,5,17].

To state the LCP we need some notation. So, for a matrix $A \in \mathbb{R}^{m,n}$ we write $A \geqslant 0$ ($A > 0$) if each element of A is nonnegative (positive). The inequality $A \leqslant 0$ ($A < 0$) is defined in an obvious way. Also, $A \geqslant B$ ($A > B$) means $A - B \geqslant 0$ ($A - B > 0$). Finally, $|A|$ denotes the matrix whose elements are the moduli of the corresponding ones of A.

* Corresponding author.
E-mail addresses: hadjidim@inf.uth.gr (A. Hadjidimos), mtzoumas@cc.uoi.gr (M. Tzoumas).

0024-3795/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2009.02.024

The LCP is defined as follows (see, e.g., [16,7,5] or [17]):
Problem: Determine $x \in \mathbb{R}^{n,n}$, if it exists, satisfying the following conditions

$$r := Mx + q \geqslant 0, \quad x \geqslant 0, \quad r^T x = 0 \quad \text{with } M \in \mathbb{R}^{n,n}, \ q \in \mathbb{R}^n \ (q \not\geqslant 0). \tag{1.1}$$

Note: In (1.1) we set $q \not\geqslant 0$ since otherwise we have the trivial solution $x = 0$, $r = q \geqslant 0$.

A sufficient and necessary condition for LCP (1.1) to possess a unique solution, for all $q \in \mathbb{R}^n$, is that M is a *P-matrix*, that is all its principal minors are positive. The corresponding proof seems to go back to Samelson et al. [20]. Subclasses of P-matrices are the *real positive definite matrices*, the *M-matrices*, the *real H-matrices* with positive diagonals, etc. In this work we focus on *real symmetric positive definite matrices*.

To solve (1.1) we consider iterative methods, the first of which is attributed to Cryer [8]. Since then many researchers have proposed other iterative methods, e.g., Mangasarian [15], Ahn [1] and Pang [18]. Recently, a growing interest has been shown in them (see, e.g., [4,2,3,13,26,9], etc).

In the present work we are mainly concerned with the well-known *Modulus Algorithm* introduced by van Bokhoven [23] and extended by Kappel and Watson [12] to the *Block Modulus Algorithm*. In these Algorithms the LCP is transformed into a fixed-point problem, where a new "unknown" z is introduced so that

$$x = |z| + z \quad \text{and} \quad r = |z| - z, \tag{1.2}$$

see, e.g., [17]. Then, using (1.2) and replacing x and r in (1.1) it is readily obtained that

$$z = f(z) := D|z| + b, \tag{1.3}$$

$$z \in \mathbb{R}^n, \quad D = (I + M)^{-1}(I - M), \quad b = -(I + M)^{-1}q. \tag{1.4}$$

Note that the iteration matrix D is nothing but the *Cayley Transform* of M [10] or [11].

2. Extrapolating LCP

For the iterative solution of (1.3) the simplest iterative scheme is the following

$$z^{(m+1)} = D|z^{(m)}| + b, \quad m = 0, 1, 2, \ldots, \ \text{with any } z^{(0)} \geqslant 0. \tag{2.1}$$

For the convergence of (2.1) to the (unique) solution of (1.3) there must hold $||D|| < 1$, where $||\cdot||$ denotes the *absolute matrix norm* induced by the *absolute vector norm* $||\cdot||$ as follows: For a given $A \in \mathbb{R}^{n,n}$, $||A|| := \sup_{\forall y \in \mathbb{R}^n \setminus \{0\}} \frac{||Ay||}{||y||}$. The *absolute vector norm*, in addition to the three well-known conditions for a *vector norm*, satisfies the following two:

$$(i) \ ||\,|x|\,|| = ||x||\,, \quad \forall x \in \mathbb{R}^n \quad \text{and} \quad (ii) \ |x| \leqslant |y| \Longrightarrow ||x|| \leqslant ||y||\,, \quad \forall x, y \in \mathbb{R}^n. \tag{2.2}$$

For the proof see [23] or [12] or Theorem 9.4 of [17]. Note that all vector norms defined by

$$||y||_p = \left(\sum_1^n |y_i|^p \right)^{\frac{1}{p}}, \quad \forall p \geqslant 1, \tag{2.3}$$

also satisfy (2.2), with the most common ones being those for $p = 1, 2, \infty$. Restricting to *symmetric positive definite matrices* M, D in (1.4) is *(real) symmetric*. Let λ_i (> 0), $i = 1(1)n$, be the eigenvalues of M, then those of D are $\frac{1-\lambda_i}{1+\lambda_i}$, $i = 1(1)n$. Consequently, the *absolute spectral norm* for D is $||D||_2 = \rho(D) = \max_{\lambda_i \in \sigma(M)} |\frac{1-\lambda_i}{1+\lambda_i}| < 1$, and so scheme (2.1) always converges. Therefore $z^{(m)}$ tends to the solution z of (1.3) as $k \to \infty$ from which x and r are recovered using (1.2).

To accelerate the convergence of (2.1) we apply *extrapolation* to (1.1). So, we multiply through by ω (> 0), the *extrapolation parameter*, in which case (1.1) becomes

$$(\omega r) := (\omega M)x + (\omega q) \geqslant 0, \quad x \geqslant 0, \quad (\omega r)^T x = 0. \tag{2.4}$$

Due to the positivity of ω, relations (1.1) imply (2.4) and vice versa; also, the matrix properties of M are inherited by ωM and $\omega q \in \mathbb{R}^n \backslash \{0\}$ ($\omega q \not\geqslant 0$).

The *extrapolated* iterative scheme based on (2.1) is constructed from (2.4) in the same way as (2.1) is constructed from (1.3). Hence

$$z^{(m+1)} = D_\omega |z^{(m)}| + b_\omega, \quad \text{with any } z^{(0)} \geqslant 0, \tag{2.5}$$

where

$$D_\omega = (I + \omega M)^{-1}(I - \omega M), \quad b_\omega = -(I + \omega M)^{-1}\omega q, \tag{2.6}$$

with D_ω being the *Extrapolated Caley Transform of M* (see [11]).

Obviously, iterative scheme (2.5) converges for any $\omega \in (0, +\infty)$ because

$$||D_\omega||_2 = \rho(D_\omega) = \max_{\omega > 0, \lambda_i \in \sigma(M)} \left| \frac{1 - \omega\lambda_i}{1 + \omega\lambda_i} \right| < 1. \tag{2.7}$$

The problem of minimization of $\rho(D_\omega)$ in (2.7) was solved in a more general form in [11] from which we borrow the following:

Theorem 2.1 (**Formulas** (**4.3**) **of** [**11**]). *Let $\lambda_{\min}$ and $\lambda_{\max}$ be the smallest and the largest eigenvalues of the real symmetric positive definite matrix M. Then, the optimal extrapolation parameter ω in (2.5) and the corresponding spectral radius of D_ω in (2.6) are given by*

$$\omega^* = \frac{1}{\sqrt{\lambda_{\min}\lambda_{\max}}}, \quad \rho(D_{\omega^*}) = \frac{\sqrt{\lambda_{\max}} - \sqrt{\lambda_{\min}}}{\sqrt{\lambda_{\max}} + \sqrt{\lambda_{\min}}}. \tag{2.8}$$

Corollary 2.1. *Under the assumptions of Theorem 2.1, $\rho(D_{\omega^*})$ is a strictly increasing function of the spectral condition number $\kappa_2 := \kappa_2(M) = \frac{\lambda_{\max}}{\lambda_{\min}}$.*

Proof. By dividing both terms of the fraction giving $\rho(D_{\omega^*})$ in (2.8) by $\sqrt{\lambda_{\min}}$ and differentiating with respect to (*wrt*) the ratio $\frac{\lambda_{\max}}{\lambda_{\min}}$ the conclusion immediately follows. □

3. Nonstationary Extrapolated Block Modulus Algorithm (*NSEBMA*)

We begin this section with the discussion of the two Modulus Algorithms:
van Bokhoven's *Modulus Algorithm (MA)*: The following lemma is taken from [12].

Lemma 3.1. *Under the notation and the assumptions made so far, if we apply van Bokhoven's MA to scheme (2.1), with $z^{(0)} = 0 \in \mathbb{R}^n$, then after N iterations,*

$$N = \left\lceil \frac{\ln\left(\frac{1-\rho(D)}{1+\rho(D)}\right) - \ln(1 + \sqrt{n})}{\ln(\rho(D))} \right\rceil, \tag{3.1}$$

one component of $z^{(N)}$ will become positive (negative), say that corresponding to the index l

$$|z_l^{(N)}| = \max_{i=1(1)n} |z_i^{(N)}|, \tag{3.2}$$

and will remain positive (negative), thereafter.

Proof. For the proof see Theorem 3 of [12] and the note(s) immediately after it. □

By Lemma 3.1 and (1.2), if $z_l^{(N)} < 0$, then $x_l^{(N)} = 0$. If $z_l^{(N)} > 0$, then $x_l^{(N)} > 0$, forcing $r_l^{(N)} = 0$. In the former case we delete the lth equation of $r = Mx + q$ and the lth column of M. In the latter we do the

same after pivoting about m_{ll}. So, the new LCP is reduced in size by one. If we assign the subscript 1 to the original M, r, D, b, and 2 to the corresponding ones of the new LCP, we will find $N_2 \leqslant N_1$, since for M_2, $\rho(D_2) < \rho(D_1)$ in general (see Theorems 3.1, 4.2 and 4.4). Hence, the total number of iterations to solve our LCP will be

$$N_1 + N_2 + \cdots + N_{n-1}, \quad \text{where } N_1 \geqslant N_2 \geqslant N_3 \geqslant \cdots \geqslant N_{n-2} \geqslant N_{n-1}. \tag{3.3}$$

Theorem 3.1. *Under the assumptions of Lemma 3.1, N is an increasing function of $\rho(D)$.*

Proof. Let $\widehat{N}$ be the quantity in the *ceiling* function in (3.1), namely

$$\widehat{N} := \frac{\ln\left(\frac{1-\rho}{1+\rho}\right) - \ln(1+\sqrt{n})}{\ln \rho}, \tag{3.4}$$

with $\rho = \rho(D)$ (< 1). Differentiating $\widehat{N}$ *wrt* ρ we obtain

$$\frac{d\widehat{N}}{d\rho} = -\frac{1}{\ln \rho}\left[\frac{2}{1-\rho^2} + \frac{\widehat{N}}{\rho}\right] > 0. \tag{3.5}$$

Therefore $\widehat{N}$ strictly increases and hence N is an increasing function of ρ. □

As is obvious, we can apply to *van Bokhoven's MA* a *nonstationary extrapolation* with ω^* being recalculated in the beginning of each cycle. If $\lambda_{\min}\lambda_{\max} = 1$, whence $\omega^* = 1$ by (2.8), $\rho(D_{\omega^*}) = \rho(D)$, otherwise $\rho(D_{\omega^*}) < \rho(D)$. Therefore, it will be expected that the total number of iterations and CPU time to solve the LCP at hand will be drastically reduced despite the recalculation of ω_i^*'s, $i = 1(1)n-1$. To realize how the Nonstationary Extrapolated Modulus Algorithm (*NSEMA*) is related to (1.1) we will express the process in matrix form.

To simplify matters, assume that l of $z_l^{(\sum_{i=1}^{p} N_i)}$, $p = 1(1)n-1$, in (3.2), is found in the natural order $(1, 2, 3, \ldots, n-1)$ and that none of the $z_l^{(\sum_{i=1}^{p} N_i)}$'s is zero. (*Note:* If $z_l^{(\sum_{i=1}^{p} N_i)} = 0$, $p < n-1$, then all the remaining components of x and r are zero.) Hence *NSEMA* terminates after $n-1$ cycles. Beginning the first cycle, (1.1) is multiplied through by ω_1^* to obtain (2.4). In (2.4), r, M, q are multiplied by ω_1^* while x remains unchanged. Note that the properties of $\omega_1^* r$, $\omega_1^* M$, $\omega_1^* q$ do not differ from those of r, M, q. After the first cycle, if $x_1^{(N_1)} = 0$ then $\omega_1^* r_1^{(N_1)} > 0$. So, the first equation and the first column of $\omega_1^* M$ are deleted. If $x_1^{(N_1)} > 0$, then $\omega_1^* r_1^{(N_1)} = 0$. Then, the pivoting follows, with pivot $\omega_1^* m_{11}^{(1)}$, where the upper index denotes cycle. (*Notes:* (i) All the multipliers in the pivoting process are those that they should have been if no extrapolation had been applied. and (ii) By Theorem 2.1 and Corollary 2.1, the ratios of the extreme eigenvalues of M and of $\omega_1^* M$ as well as those of the corresponding principal submatrices remain unchanged.) Then, a deletion, such as before, follows. To return to the original LCP in (1.1) we can follow one of three alternatives: (i) Divide all n equations, including the first one, by ω_1^* to recover (1.1). Then, the first cycle of the *NSEMA* is completed and the second cycle follows. At the end of the $n-1$ cycles the actual values for x and r are obtained. (ii) Begin the second cycle by multiplying the $n-1$ equations from the second to the last by $\frac{\omega_2^*}{\omega_1^*}$ and so on. In this alternative, setting

$$\Omega_1^* = \operatorname{diag}\left(\omega_1^*, \omega_2^*, \omega_3^*, \ldots, \omega_{n-1}^*, 1\right), \tag{3.6}$$

the Algorithm we use solves the following *Nonstationary Extrapolated* LCP

$$(\Omega_1^* r) = (\Omega_1^* M)x + (\Omega_1^* q) \geqslant 0, \quad (\Omega_1^* r)^T x = 0. \tag{3.7}$$

Since x has remained unchanged, only $\Omega_1^* r$ has to be premultiplied by Ω_1^{*-1} to recover r.
(iii) Multiply the last $n-1$ equations by ω_2^*, noting by (2.8) that the present ω_2^* differs from the previous one by the factor ω_1^*, and go on with the second cycle. Setting

$$\Omega_2^* = \operatorname{diag}\left(\omega_1^*, \omega_1^*\omega_2^*, \omega_1^*\omega_2^*\omega_3^*, \ldots, \prod_{i=1}^{n-1}\omega_i^*, \prod_{i=1}^{n-1}\omega_i^*\right), \tag{3.8}$$

the Algorithm used solves the following *Nonstationary Extrapolated* LCP

$$(\Omega_2^* r) \;=\; (\Omega_2^* M)x + (\Omega_2^* q) \;\geqslant\; 0, \quad (\Omega_2^* r)^T x \;=\; 0. \tag{3.9}$$

Obviously, x remains unchanged and a premultiplication of $\Omega_2^* r$ by Ω_2^{*-1} recovers r.

Two points have to be clarified. (i) From the second cycle onwards $\Omega_1^* M$ and $\Omega_2^* M$ are not symmetric. This is true, but we should recall that the submatrix used in each cycle is a positive multiple of the original one. Therefore all the properties of the latter are inherited by the one used. (ii) In a real situation the ordering of l's in all three alternatives would not be the natural one and so the components of x appear in a permuted order. Let P be the corresponding permutation matrix. Then, the problem we solve, say in alternative (iii), is

$$(\Omega_2^* Pr) = (\Omega_2^* PMP^T)(Px) + (\Omega_2^* Pq) \geqslant 0, \quad (\Omega_2^* Pr)^T (Px) = 0. \tag{3.10}$$

Obviously, we have to keep track of the ordering of l's, as in the *Gauss elimination*. Then, x and r are recovered in an obvious way.

Kappel and Watson's *Block Modulus Algorithm (BMA)*: Lemma 3.2 below is from [12].

Lemma 3.2. *Under the notation and the assumptions made so far, if we apply Kappel and Watson's Block Modulus Algorithm (BMA) to iterative scheme* (2.1), *with* $z^{(0)} = 0 \in \mathbb{R}^n$, *then after N iterations, where N is given by* (3.1), *not only the absolutely largest component of* $z^{(N)}$ *will preserve its sign thereafter, but also all other components of it satisfying*

$$|z_l^{(N)}| \geqslant T := \frac{1}{\sqrt{n}} \left(\frac{1}{1+\rho(D)} - \frac{\rho^N(D)}{1-\rho(D)} \right) \|b\|_2. \tag{3.11}$$

Proof. For the proof see Theorem 4 of [12] and the notes following it. □

In general, there may be more than one component of $z^{(N)}$ that will allow to determine the corresponding $x_l^{(N)}$ and $r_l^{(N)}$. In such a case, more that one equation (and corresponding columns of M) will be deleted and the next LCP will be drastically reduced in size. It is then expected that the *Kappel and Watson's Algorithm* will produce the solution sought in fewer iterations in each cycle, and maybe in fewer cycles, than that of van Bokhoven's.

In what follows we state and prove a theorem which seems to be a negative result.

Theorem 3.2. *Under the assumptions of Lemma* 3.2, T *strictly decreases with* $\rho(D)$ *increasing*.

Proof. Since n and $\|b\|_2$ are positive constants it is obvious that $\frac{dT}{d\rho}$ and $\frac{d\widehat{T}}{d\rho}$, with

$$\widehat{T} := \frac{1}{1+\rho} - \frac{\rho^{\widehat{N}}}{1-\rho} \tag{3.12}$$

and $\rho = \rho(D)$, are of the same sign. Differentiating we have

$$\frac{d\widehat{T}}{d\rho} = -\frac{1}{(1+\rho)^2} - \frac{(1-\rho)\frac{d\rho^{\widehat{N}}}{d\rho} + \rho^{\widehat{N}}}{(1-\rho)^2}. \tag{3.13}$$

To find $\frac{d\rho^{\widehat{N}}}{d\rho}$, we put $y = \rho^{\widehat{N}}$, take logarithms, and differentiate *wrt* ρ to obtain

$$\frac{1}{y}\frac{dy}{d\rho} = \frac{d\widehat{N}}{d\rho} \ln\rho + \widehat{N}\frac{1}{\rho}. \tag{3.14}$$

Substituting $\frac{d\widehat{N}}{d\rho}$ and $\widehat{N}$, from (3.5) and (3.4), respectively, as well as $y = \rho^{\widehat{N}}$ into (3.14), we can obtain after some simple manipulations that $\frac{d\rho^{\widehat{N}}}{d\rho} = -\frac{2\rho^{\widehat{N}}}{1-\rho^2}$. Substituting the last expression into (3.13) and using (3.12) we finally obtain that

$$\frac{d\widehat{T}}{d\rho} = -\frac{1}{1+\rho}\left(\frac{1}{1+\rho} - \frac{\rho^{\widehat{N}}}{1-\rho}\right) = -\frac{\widehat{T}}{1+\rho} < 0. \tag{3.15}$$

Consequently, $\widehat{T}$ and T are strictly decreasing functions of ρ. □

Remark 3.1. The above surprising result states that ρ should increase rather than decrease to get a smaller T and so increase the possibility to have more than one components of $z^{(N)}$ satisfying (3.11). However, we should bear in mind that the new feature of the *BMA* is the exploitation of the fact that $|z_l^{(N)}| \geqslant T$ may be satisfied by more than one l.

In corroboration to the above remark it should be mentioned that in a plethora of examples we have run, in none of them the simple *MA* has beaten the *BMA*. Also, a partial answer as to what actually happens is given theoretically by the following statement.

Theorem 3.3. *As $\rho = \rho(D)$ decreases in the interval $(0,1)$, the number $\widehat{N}$ in (3.4) decreases faster than what $\widehat{T}$ in (3.12) increases. More specifically*

$$\frac{d(\widehat{N}\widehat{T})}{d\rho} > 0. \tag{3.16}$$

Proof. Considering the derivative in (3.16) and using (3.5) and (3.15) we successively obtain

$$\begin{aligned}\frac{d(\widehat{N}\widehat{T})}{d\rho} &= \frac{d\widehat{N}}{d\rho}\widehat{T} + \widehat{N}\frac{d\widehat{T}}{d\rho} = -\frac{1}{\ln\rho}\left[\frac{2}{1-\rho^2} + \frac{\widehat{N}}{\rho}\right]\widehat{T} + \widehat{N}\left(-\frac{\widehat{T}}{1+\rho}\right)\\ &= \frac{\widehat{T}}{(1+\rho)\ln\rho}\left[\frac{2}{1-\rho} + \frac{(1+\rho+\rho\ln\rho)}{\rho}\widehat{N}\right].\end{aligned} \tag{3.17}$$

For the coefficient of $\widehat{N}$ in the second term in the brackets above, it is found that $\frac{d\left(\frac{1+\rho+\rho\ln\rho}{\rho}\right)}{d\rho} = \frac{\rho-1}{\rho^2} < 0$, and so

$$\inf_{\rho\in(0,1)} \frac{(1+\rho+\rho\ln\rho)}{\rho} = \frac{(1+\rho+\rho\ln\rho)}{\rho}|_{\rho=1} = 2,$$

meaning that the coefficient in question is always positive. Hence the right side of the equalities in (3.17) is positive proving our claim in (3.16). □

It is realized that the *nonstationary extrapolation*, with the three alternatives for the *MA*, can also be applied to the *BMA*. Then, one should expect to obtain the solution in fewer iterations than those required for the simple *BMA*. So, the *Nonstationary Extrapolated Block Modulus Algorithm (NSEBMA)* is expected to give optimal results in terms of iterations and CPU time for a specific LCP. It is understood that one has to deal with blocks instead of with points. For example, let p ($\leqslant n$) be the total number of cycles required to solve the *NSEBMA*, let n_i, with $\sum_1^p n_i = n$, be the number of components in each block and ω_i^*, $i = 1(1)p$, be the optimal extrapolation parameters. Then, the analogous to (3.8) extrapolation matrix and that to (3.10) Nonstationary Extrapolated LCP, which is solved, are

$$\begin{aligned}&\Omega_2^{*(b)} = \mathrm{diag}\left(\omega_1^* I_{n_1}, \omega_1^*\omega_2^* I_{n_2}, \omega_1^*\omega_2^*\omega_3^* I_{n_3}, \ldots, \prod_{i=1}^{p-1}\omega_i^* I_{n_{p-1}}, \prod_{i=1}^{p-1}\omega_i^* I_{n_p}\right),\\ &(\Omega_2^{*(b)}Pr) = (\Omega_2^{*(b)}PMP^T)(Px) + (\Omega_2^{*(b)}Pq) \geqslant 0, \quad (\Omega_2^{*(b)}Pr)^T(Px) = 0.\end{aligned} \tag{3.18}$$

4. Further theoretical background

In this section we prove a number of statements that apply to either of the *Nonstationary Extrapolated Modulus Algorithms*. Bearing in mind the two *Notes* in the discussion preceding the three alternatives for the *NSEBA* were presented, our analysis can put aside the extrapolation parameters ω_i^*'s.

First we investigate the case of *NSEMA* and then the results obtained are generalized to cover the *NSEBMA*.

Note that going from one cycle of iterations, say the very first one, to the next of *MA* we do fewer operations per iteration due to the reduced size of the new LCP. Besides, the extrapolation applied to the new LCP will be faster than that applied to the old problem. To prove this, in view of Theorem 2.1 and Corollary 2.1 we have to compare the ratios of the largest to the smallest eigenvalue of the coefficient matrices in the two LCPs. To make such a comparison we distinguish two cases depending on the sign of $z_l^{(N)}$ in (3.2). If $x_l^{(N)}$ is to be zero, then the lth equation of the LCP and the lth column of M are deleted. If $r_l^{(N)}$ is to be zero, a Gauss elimination takes place with pivot m_{ll} before the LCP is reduced in size by one as before. The following statements describe what happens in each case.

Theorem 4.1. *Let $M \in \mathbb{R}^{n,n}$ be symmetric and positive definite. The submatrix M_{22} obtained by deleting the l*th *row and column of M is also symmetric and positive definite.*

Proof. It is well known that any principal submatrix of a real symmetric positive definite matrix is also symmetric and positive definite (see, e.g., [24,25] or [5]). □

Theorem 4.2. *Let $M \in \mathbb{R}^{n,n}$ be symmetric and positive definite with $\lambda_{\min}$, $\lambda_{\max}$ being its smallest and largest (positive) eigenvalues. Let $\widehat{\lambda}_{\min}$ and $\widehat{\lambda}_{\max}$ be the corresponding eigenvalues of the submatrix M_{22} of Theorem 4.1. Then,*

$$\lambda_{\min} \leqslant \widehat{\lambda}_{\min} \leqslant \widehat{\lambda}_{\max} \leqslant \lambda_{\max}. \tag{4.1}$$

Proof. As is known (see, e.g., Theorem 2.2.2 in [25]), for any $w \in \mathbb{R}^n \backslash \{0\}$ there hold

$$\lambda_{\min} \leqslant \frac{w^T M w}{w^T w} \leqslant \lambda_{\max}, \tag{4.2}$$

where equality holds at the left (resp. right) end with w being the eigenvector associated with $\lambda_{\min}$ (resp. $\lambda_{\max}$). For simplicity, let $l = 1$ and M be partitioned as follows

$$M = \left[\begin{array}{c|c} m_{11} & y^T \\ \hline y & M_{22} \end{array}\right] \quad \text{with } y = [m_{21}\ m_{31}\ \ldots\ m_{n1}]^T. \tag{4.3}$$

Defining the vector w

$$w = [0\, w_{n-1}^T]^T \in \mathbb{R}^n, \quad w_{n-1} \in \mathbb{R}^{n-1} \backslash \{0\} \Longrightarrow w^T w = w_{n-1}^T w_{n-1}, \tag{4.4}$$

we will have

$$w^T M w = [0 | w_{n-1}^T] \left[\begin{array}{c|c} m_{11} & y^T \\ \hline y & M_{22} \end{array}\right] [0 | w_{n-1}^T]^T = w_{n-1}^T M_{22} w_{n-1}. \tag{4.5}$$

Taking w_{n-1} to be the eigenvector of M_{22} associated with $\widehat{\lambda}_{\min}$ we have

$$\widehat{\lambda}_{\min} = \frac{w_{n-1}^T M_{22} w_{n-1}}{w_{n-1}^T w_{n-1}} = \frac{w^T M w}{w^T w} \geqslant \lambda_{\min}.$$

Hence, by virtue of (4.2), the left inequality in (4.1) is proved. Similarly, taking w_{n-1} to be the eigenvector of M_{22} associated with $\widehat{\lambda}_{\max}$ the right inequality in (4.1) is also proved. □

Theorem 4.3. *Under the assumptions and notation of Theorems 4.1 and 4.2 and in view of Corollary 2.1 the extrapolation applied to the reduced LCP will make it converge at least as fast as the extrapolation applied to the original one.*

Proof. In view of (4.1) and Corollary 2.1 the proof is immediate. □

Remark 4.1. Note that we have identical rates of convergence in the old and the new LCPs, namely $\widehat{\lambda}_{\min} = \lambda_{\min}$ and $\widehat{\lambda}_{\max} = \lambda_{\max}$ simultaneously hold, if and only if (*iff*) the eigenvectors w_{n-1_m} and w_{n-1_M} associated with $\widehat{\lambda}_{\min}$ and $\widehat{\lambda}_{\max}$ of M_{22} are orthogonal to the vector y in (4.3) and, also, $w_m = [0\, w_{n-1_m}^T]^T$ and $w_M = [0\, w_{n-1_M}^T]^T$ are the eigenvectors of M associated with $\lambda_{\min}$ and $\lambda_{\max}$, respectively.

Now, we come to the case where a pivoting process takes place.

Theorem 4.4. *Let $M \in \mathbb{R}^{n,n}$ be symmetric and positive definite. Applying Gauss elimination to it with pivot any diagonal element m_{ll}, $l = 1(1)n$, the submatrix $\widehat{M}$ obtained by deleting the lth row and column of the resulting matrix is also symmetric and positive definite.*

Proof. For simplicity we assume that m_{11} is taken as pivot in the Gauss elimination. If we also assume that M is partitioned as in (4.3), then Gauss elimination results to

$$\left[\begin{array}{c|c} 1 & 0_{n-1}^T \\ \hline -\frac{1}{m_{11}}y & I_{n-1} \end{array}\right]\left[\begin{array}{c|c} m_{11} & y^T \\ \hline y & M_{22} \end{array}\right] = \left[\begin{array}{c|c} m_{11} & y^T \\ \hline 0_{n-1} & M_{22} - \frac{1}{m_{11}}yy^T \end{array}\right]. \tag{4.6}$$

Since M_{22} is symmetric so is the matrix

$$\widehat{M} = M_{22} - \frac{1}{m_{11}}yy^T. \tag{4.7}$$

To prove that $\widehat{M}$ is also positive definite we consider any vector

$$w = [w_1\, w_{n-1}^T]^T \in \mathbb{R}^n \quad \text{with } w_1 = -\frac{1}{m_{11}}(w_{n-1}^T y) \in \mathbb{R},\ w_{n-1} \in \mathbb{R}^{n-1}\backslash\{0\}. \tag{4.8}$$

Then, we successively have

$$\begin{aligned} 0 < w^T M w &= \left[-\tfrac{1}{m_{11}}(w_{n-1}^T y)|w_{n-1}^T\right]\left[\begin{array}{c|c} m_{11} & y^T \\ \hline y & M_{22} \end{array}\right]\left[-\tfrac{1}{m_{11}}(w_{n-1}^T y)|w_{n-1}^T\right]^T \\ &= w_{n-1}^T(M_{22} - \tfrac{1}{m_{11}}yy^T)w_{n-1} = w_{n-1}^T\widehat{M}w_{n-1}, \end{aligned} \tag{4.9}$$

proving our assertion. □

Theorem 4.5. *Let M be the matrix of Theorem 4.4 and $\lambda_{\min}$, $\lambda_{\max}$ be its smallest and largest eigenvalues. Let the smallest and largest eigenvalues of $\widehat{M}$ in (4.7) of Theorem 4.4 be $\widehat{\lambda}_{\min}$, $\widehat{\lambda}_{\max}$, respectively. Then, there will hold*

$$\lambda_{\min} \leqslant \widehat{\lambda}_{\min} \leqslant \widehat{\lambda}_{\max} \leqslant \lambda_{\max}. \tag{4.10}$$

Proof. Let w be the vector

$$w = [w_1\, w_{n-1}^T]^T \in \mathbb{R}^n \quad \text{with } w_1 \in \mathbb{R},\ w_{n-1} \in \mathbb{R}^{n-1}\backslash\{0\}, \tag{4.11}$$

then we have that

$$\frac{w^T M w}{w^T w} = \frac{m_{11}\left(w_1 + \frac{1}{m_{11}}(w_{n-1}^T y)\right)^2 + w_{n-1}^T\widehat{M}w_{n-1}}{w_1^2 + w_{n-1}^T w_{n-1}}. \tag{4.12}$$

Taking as w_{n-1} the eigenvector of $\widehat{M}$ associated with its smallest eigenvalue $\widehat{\lambda}_{\min}$ and $w_1 = -\frac{1}{m_{11}}(w_{n-1}^T y)$, then the vector w has the form in (4.8) and we successively obtain

$$\lambda_{\min} \leqslant \frac{w^T M w}{w^T w} = \frac{w_{n-1}^T\widehat{M}w_{n-1}}{\frac{1}{m_{11}^2}(w_{n-1}^T y)^2 + w_{n-1}^T w_{n-1}} \leqslant \frac{w_{n-1}^T\widehat{M}w_{n-1}}{w_{n-1}^T w_{n-1}} = \widehat{\lambda}_{\min} \tag{4.13}$$

proving the left inequality in (4.10). Taking w_{n-1} to be the eigenvector of $\widehat{M}$ associated with $\widehat{\lambda}_{\max}$ and $w_1 = 0$, so that the vector w has the form of (4.4), we have

$$\lambda_{\max} \geqslant \frac{w^T M w}{w^T w} = \frac{\frac{1}{m_{11}}(w_{n-1}^T y)^2 + w_{n-1}^T \widehat{M} w_{n-1}}{w_{n-1}^T w_{n-1}} \geqslant \frac{w_{n-1}^T \widehat{M} w_{n-1}}{w_{n-1}^T w_{n-1}} = \widehat{\lambda}_{\max}, \tag{4.14}$$

proving the right inequality in (4.10). □

Remark 4.2. It is similar to Remark 4.1. Namely, the equalities $\widehat{\lambda}_{\min} = \lambda_{\min}$ and $\widehat{\lambda}_{\max} = \lambda_{\max}$ simultaneously hold *iff* the eigenvectors $w_{n-1} = w_{n-1_m}$ and $w_{n-1} = w_{n-1_M}$ of $\widehat{M}$ are orthogonal to y, and $[0\, w_{n-1_m}^T]^T$ and $[0\, w_{n-1_M}^T]^T$ are the eigenvectors of M associated with $\lambda_{\min}$ and $\lambda_{\max}$, respectively.

Theorem 4.6. *Under the assumptions of Theorems 4.4 and 4.5, if any of the four inequalities in (4.13) and (4.14) does* **not** *hold, then, one of the two extreme inequalities in (4.10) of Theorem 4.5 will be a* **strict** *one. Furthermore, the optimal spectral radius in (2.8) corresponding to the matrix $D(\widehat{M})$ will be* **strictly less** *than that corresponding to $D(M)$, with the matrix of the form $D(\cdot)$ being defined in (1.4) in terms of $\widehat{M}$ and M, respectively.*

Proof. The first part comes directly from the implied strict inclusion $[\widehat{\lambda}_{\min}, \widehat{\lambda}_{\max}] \subset [\lambda_{\min}, \lambda_{\max}]$ as a consequence of which we have $\frac{\widehat{\lambda}_{\max}}{\widehat{\lambda}_{\min}} < \frac{\lambda_{\max}}{\lambda_{\min}}$. The second part comes from the previous strict inequality and Corollary 2.1. □

Coming now to the case of the *NSEBMA* it is clear that, in general, we have to deal with a repeated application of Theorems 4.2 and 4.5 since more than one components of $z^{(N)}$ may satisfy (3.11). Of course, one can use blocks to prove the analogous propositions to Theorems 4.1–4.6. To see what the difference is, we outline below a block analogue of a combination of Theorems 4.4–4.5 and Remark 4.2.

Theorem 4.7. *Let $M \in \mathbb{R}^{n,n}$ be symmetric and positive definite and that the first p $r_i^{(N)}$'s are to become zeros $(2 \leqslant p < n)$. Let M be of the block form*

$$M = \left[\begin{array}{c|c} M_{11} & Y^T \\ \hline Y & M_{22} \end{array}\right] \quad \text{with } M_{11} \in \mathbb{R}^{p,p},\ M_{22} \in \mathbb{R}^{n-p,n-p},\ Y \in \mathbb{R}^{n-p,p}. \tag{4.15}$$

Then: (i) *Applying a "block" Gauss elimination, where all p columns below the diagonal of M_{11} are eliminated, and deleting the first block row and column of the resulting matrix, the submatrix $\widehat{M}$ obtained is symmetric and positive definite.*
(ii) *Let $\lambda_{\min}$ and $\lambda_{\max}$ be the smallest and the largest eigenvalues of M and $\widehat{\lambda}_{\min}$ and $\widehat{\lambda}_{\max}$ be the corresponding ones of $\widehat{M}$. Then, there will hold*

$$\lambda_{\min} \leqslant \widehat{\lambda}_{\min} \leqslant \widehat{\lambda}_{\max} \leqslant \lambda_{\max}. \tag{4.16}$$

(iii) *Equalities in (4.16) hold at both ends iff the pair of eigenvectors $w_{n-p} = w_{n-p_m}$ and $w_{n-p} = w_{n-1_M}$ associated with $\widehat{\lambda}_{\min}$ and $\widehat{\lambda}_{\max}$ of $\widehat{M}$ are orthogonal to the columns of Y, and $[-w_{n-p}^T Y M_{11}^{-1} | w_{n-p_m}^T]^T$ and $[0 | w_{n-p_M}^T]^T$ are the eigenvectors of M associated with $\lambda_{\min}$ and $\lambda_{\max}$, respectively.*

Proof. (i) Recall that the matrices M_{11} and M_{22} are symmetric positive definite. Hence M_{11} admits a Cholesky decomposition which can be written as $L_{11}U_{11}$, where L_{11} is lower triangular with $\mathrm{diag}(L_{11}) = I_p$ and U_{11} upper triangular that can be written as $\mathrm{diag}(U_{11})L^T$ with $\mathrm{diag}(U_{11})$ positive diagonal. So, the "block" pivoting process will be as follows:

$$\left[\begin{array}{c|c} L_{11}^{-1} & 0_{p,n-p} \\ \hline -YM_{11}^{-1} & I_{n-p} \end{array}\right] \left[\begin{array}{c|c} M_{11} & Y^T \\ \hline Y & M_{22} \end{array}\right] = \left[\begin{array}{c|c} U_{11} & L_{11}^{-1}Y^T \\ \hline 0_{n-p,p} & \widehat{M} \end{array}\right],$$

$$\widehat{M} = M_{22} - YM_{11}^{-1}Y^T. \tag{4.17}$$

Hence $\widehat{M}$ is symmetric and positive definite because M_{11} and therefore $M_{11}^{-1} \in \mathbb{R}^{p,p}$ possess both these properties. Letting

$$w = [w_p^T\, w_{n-p}^T]^T \in \mathbb{R}^n\backslash\{0\} \quad \text{with } w_p = -M_{11}^{-1}Y^T w_{n-p} \in \mathbb{R}^p,\ w_{n-p} \in \mathbb{R}^{n-p}\backslash\{0\} \tag{4.18}$$

it is obtained that

$$\begin{aligned} 0 < w^T M w &= \left[-w_{n-p}^T Y M_{11}^{-1} | w_{n-p}^T\right] \left[\begin{array}{c|c} M_{11} & Y^T \\ \hline Y & M_{22} \end{array}\right] \left[-w_{n-p}^T Y M_{11}^{-1} | w_{n-p}^T\right]^T \\ &= w_{n-p}^T (M_{22} - Y M_{11}^{-1} Y^T) w_{n-p} = w_{n-p}^T \widehat{M} w_{n-p}, \end{aligned} \tag{4.19}$$

which proves that $\widehat{M}$ is also positive definite.
(ii) Let w be the vector

$$w = [w_p^T\, w_{n-p}^T]^T \in \mathbb{R}^n\backslash\{0\}, \quad w_p \in \mathbb{R}^p,\ w_{n-p} \in \mathbb{R}^{n-p}\backslash\{0\}. \tag{4.20}$$

Forming $\frac{w^T M w}{w^T w}$, replacing w from (4.20), using for M the above block partitioned form and for M_{22} the expression from (4.15) in terms of $\widehat{M}$, after some manipulation, we obtain that

$$\frac{w^T M w}{w^T w} = \frac{||M_{11}^{\frac{1}{2}} \left(w_p + M_{11}^{-1} Y^T w_{n-p}\right)||_2^2 + w_{n-p}^T \widehat{M} w_{n-p}}{||M_{11}^{-1} Y^T w_{n-p}||_2^2 + w_{n-p}^T w_{n-p}}, \tag{4.21}$$

where $M_{11}^{\frac{1}{2}}$ is the unique real symmetric positive definite square root of M_{11} (see, e.g., Theorem 2.2.7 in [25]). Now we work in a similar way as before in Theorem 4.5. Namely, taking as w_{n-p} the eigenvector of $\widehat{M}$ associated with its smallest eigenvalue $\widehat{\lambda}_{\min}$ and $w_p = -M_{11}^{-1} Y^T w_{n-p}$, we can obtain

$$\lambda_{\min} \leqslant \frac{w^T M w}{w^T w} = \frac{w_{n-p}^T \widehat{M} w_{n-p}}{||M_{11}^{-1} Y^T w_{n-p}||_2^2 + w_{n-p}^T w_{n-p}} \leqslant \frac{w_{n-p}^T \widehat{M} w_{n-p}}{w_{n-p}^T w_{n-p}} = \widehat{\lambda}_{\min} \tag{4.22}$$

proving the left inequality in (4.16). Taking w_{n-p} to be the eigenvector of $\widehat{M}$ associated with the largest eigenvalue $\widehat{\lambda}_{\max}$ and $w_p = 0$, we have

$$\lambda_{\max} \geqslant \frac{w^T M w}{w^T w} = \frac{||M_{11}^{-\frac{1}{2}} Y^T w_{n-p}||_2^2 + w_{n-p}^T \widehat{M} w_{n-p}}{w_{n-p}^T w_{n-p}} \geqslant \frac{w_{n-p}^T \widehat{M} w_{n-p}}{w_{n-p}^T w_{n-p}} = \widehat{\lambda}_{\max}, \tag{4.23}$$

where $M_{11}^{-\frac{1}{2}}$ is the inverse of $M_{11}^{\frac{1}{2}}$, proving the right inequality in (4.16).
(iii) For the first part of our assertion to hold, the norms in (4.22) and (4.23) must be zero. Due to the invertibility of M_{11}^{-1} and $M_{11}^{-\frac{1}{2}}$ this holds *iff* the associated eigenvectors with $\widehat{\lambda}_{\min}$ and $\widehat{\lambda}_{\max}$ must be orthogonal to the columns of the submatrix Y. The second part of our assertion readily follows. □

5. Numerical examples

Before we present our specific examples we make a number of points.

(i) We have run numerous examples of various sizes from $n = 3$ to $n = 50$ using all six methods. Namely, iterative methods (2.1) and (2.5) of Section 2, the van Bokhoven's *MA*, its *nonstationary extrapolated* counterpart (*NSEMA*), and similarly, the Kappel and Watson's *BMA* and the nonstationary extrapolated one (*NSEBMA*). For *NSEMA* and *NSEBMA* of the three alternatives of Section 3 the one in (iii) was adopted.
(ii) For each n and for all six methods the vector $q \in \mathbb{R}^n$ was the same and was selected by using the Matlab command 10*(rand(n, 1)-0.5), so that each component q_i, $i = 1(1)n$, was chosen randomly in the interval $(-5, 5)$. It was observed that for the same matrix M but for different random vectors q the results were pretty much the same.

Table 1
Spectral condition numbers of the matrix coefficient $M = \text{tridiag}(-1, 2, -1) \in \mathbb{R}^{n,n}$.

n	10	20	30	40	50
$\kappa_2(tridiag(-1, 2, -1))$	48.3742	178.064	388.812	680.617	1053.48

Table 2
Number of iterations (iter) and CPU times in seconds.

n		*Iterative Methods for Example 1*					
		MA	*BMA*	$(n-1)N$[a]	*NSEMA*	*NSEBMA*	$(n-1)N_\omega$[b]
10	iter	111	28	225	66	14	108
	CPU	0.060	0.040		0.040	0.030	
20	iter	553	128	2356	183	34	551
	CPU	0.120	0.060		0.080	0.040	
30	iter	3733	319	9135	723	54	1392
	CPU	0.731	0.171		0.240	0.100	
40	iter	6224	629	23,712	1064	97	2691
	CPU	2.003	0.440		0.511	0.180	
50	iter	17,253	1015	49,539	2236	102	4459
	CPU	6.780	1.061		1.201	0.431	

[a] $(n-1)N$ is the possible maximum number of iterations for *MA* and *BMA*.
[b] $(n-1)N_\omega$ is the possible maximum number of iterations for *NSEMA* and *NSEBMA*.

(iii) If $z^{(0)} = 0$ in (2.1) and (2.5), then all three unextrapolated methods have identical the first N $z^{(k)}$'s, $k = 1(1)N$, with N of (3.1). The same holds for the three extrapolated methods.

(iv) Recall that all four (Block) Modulus Algorithms are exact that is if exact arithmetic were used the exact result would be obtained after at most $(n-1)N$ iterations followed by the solution of a linear system. In contrast with the (Block) Modulus Algorithms, the methods (2.1) and (2.5) are iterative. Hence it is not easy to have a fair stopping criterion. What we did was the following. After the solution was found by any of the four (Block) Modulus Algorithms exhausting all $K = \sum_{i=1}^{p} N_i$ iterations, provided $K \leqslant 10^6$, we determined the "worst" relative absolute error e for the last two iterations for *NSEMA* and NSEBMA, that is $e = \frac{||x^{(K)} - x^{(K-1)}||_2}{||x^{(K)}||_2}$. This was subsequently used as a stopping criterion for the two iterative methods, specifically

$$\frac{\|x^{(k+1)} - x^{(k)}\|_2}{\|x^{(k+1)}\|_2} = \frac{\| \,|z^{(k+1)}| + z^{(k+1)} - |z^{(k)}| - z^{(k)}\|_2}{\| \,|z^{(k+1)}| + z^{(k+1)}\|_2} \leqslant e, \; k = 1, 2, 3, \ldots,$$

and a check was made after each iteration.

(v) It was observed that in almost all the cases of the four (Block) Modulus Algorithms the number of iterations required for the solution of an LCP was much less than the theoretical computed one ($(n-1)N$).

(vi) In more than 98% of the examples we ran $e = 0$ to the Matlab accuracy something which could not happen with the iterative methods (2.1) and (2.5). So, what we would suggest is that if the obtained relative absolute error e is not very satisfactory then use the last z of *NSEMA* or *NSEBMA* and run a small number of iterations, say 5 to 10, using (2.1) as a "smoother" until an e of satisfactory accuracy is obtained.

Table 3
Spectral condition numbers for the *Hilbert* matrix $M = \mathcal{H} \in \mathbb{R}^{n,n}$.

n	3	4	5	6	7	8	9
$\kappa_2(\mathcal{H})$	$5.241 * 10^2$	$1.551 * 10^4$	$4.766 * 10^5$	$1.495 * 10^7$	$4.754 * 10^8$	$1.526 * 10^{10}$	$4.931 * 10^{11}$

Table 4
Number of iterations (iter) and CPU times in seconds.

n		Iterative methods for Example 2					
		MA	*BMA*	$(n-1)N$	*NSEMA*	*NSEBMA*	$(n-1)N_\omega$
3	iter	1650	1289	2578	61	48	96
	CPU	0.090	0.100		0.030	0.030	
4	iter	53,476	53,476	160,428	369	369	1107
	CPU	2.133	2.664		0.051	0.050	
5	iter	$>10^6$	$>10^6$	$8,394,108 > 10^6$	3184	2662	10,648
	CPU	–[a]	–		0.181	0.190	
6	iter	$>10^6$	$>10^6$		18,364	18,369	91,820
	CPU	–	–		0.841	1.052	
7	iter	$>10^6$	$>10^6$		123,006	123,856	738,036
	CPU	–	–		6.189	7.271	
8	iter	$>10^6$	$>10^6$		807,005	807,005	$5,649,035 > 10^6$
	CPU	–	–		44.684	54.949	
9	iter	$>10^6$	$>10^6$		$>10^6$	$>10^6$	
	CPU	–	–		–	–	

[a] A dash (–) means that no convergence has been achieved.

(vii) In all experiments the theory of Sections 2–4 was confirmed. Namely: (a) Regarding execution (CPU) times, all three Extrapolated schemes are better than the unextrapolated ones. (b) Both Block Modulus Algorithms are better than the corresponding simple Modulus Algorithms. (c) Going from one experiment to another of the same size the CPU time required for each method becomes larger as the condition number $\kappa_2(D)$ or $\kappa_2(D_\omega)$ increases.

(viii) In case the condition number is moderately large (see *Example 1*) all four (Block) Modulus Algorithms work exceptionally well. For extremely large condition numbers (see *Example 2*) all methods work only for very small numbers of n and this is due to the tremendous number of iterations required. For those n for which *NSEMA* and *NSEBMA* work the results are very satisfactory.

Example 1. M is the classical tridiagonal matrix $M = \text{tridiag}(-1, 2, -1) \in \mathbb{R}^{n,n}$, with $n = 10(10)50$. The corresponding spectral condition numbers for M are given in Table 1.

In all five cases of the present example the results are very good despite the relatively large condition numbers. This, in our opinion, is mainly due to the sparsity of the matrix and also to its irreducible diagonal dominance property. As is seen *NSEBMA* is the best method. There are two extra columns

under $(n-1)N$ and $(n-1)N_\omega$ which indicate the possible maximum number of iterations for *MA*, *BMA* and *NSEMA*, *NSEBMA*, respectively.

Example 2. M is the *Hilbert* matrix $\mathcal{H} \in \mathbb{R}^{n,n} := \{h_{i,j} = \frac{1}{i+j-1},\ i,j = 1(1)n\}$, with $n = 3(1)9$. The spectral condition numbers for M are illustrated in Table 3.

In the cases of this example, a "nightmare" case when solving (or pivoting) a linear system, the large condition numbers are disastrous even for rather small values of n. In our opinion, despite the irreducible diagonal dominance property of the coefficient matrix, the "poor" results may be due to the dense character of it. It is noted that this is the only example out of those run that *NSEMA* beats *NSEBMA*. Table 4 is similar to Table 2.

6. Concluding remarks

Before we conclude our work we would like to make a number of points:

(i) The theory developed in the present work is fully confirmed by the numerical experiments.
(ii) The principle of *extrapolation* as was introduced in Sections 2–4 increases the convergence rates for all three known methods, namely the iterative method (2.1), the *MA* and the *BMA*.
(iii) Kappel and Watson [12] introduced a kind of *nonstationary extrapolation* but it is very difficult, if not impossible, in practice to find the appropriate positive diagonal matrix Γ defined there. Our work gives a partial answer for *symmetric positive definite* matrices.
(iv) An extension of the theory of the present paper seems to work also in cases where the matrix M is an M-matrix or a (real) H-matrix with positive diagonal elements. It is well-known that these two classes of matrices are P-matrices and the LCP has a unique solution that can also be found by other iterative methods (see, e.g., [15,1,18,7,16,17,21,22]). In this direction we have been working with encouraging preliminary results.

References

[1] B.H. Ahn, Solution of nonsymmetric linear complementarity problems by iterative methods, J. Optim. Theory Appl. 33 (1981) 175–185.
[2] Z.-Z. Bai, On the monotone convergence of the matrix multisplitting relaxation methods for linear complementarity problem, IMA J. Numer. Anal. 18 (1998) 509–518.
[3] Z.-Z. Bai, On the convergence of the multisplitting methods for linear complementarity problem, SIAM J. Matrix Anal. Appl. 21 (1999) 67–78.
[4] Z.-Z. Bai, D.J. Evans, Matrix multisplitting relaxation methods for linear complementarity problems, Int. J. Comput. Math. 63 (1997) 309–326.
[5] A. Berman, R.J. Plemmons, Classics in Applied Mathematics, SIAM, Philadelphia, 1994
[6] R.W. Cottle, G.B. Dantzig, Complementarity pivot theory of mathematical programming, Linear Algebra Appl. 1 (1968) 103–125.
[7] R.W. Cottle, J.-S. Pang, R.E. Stone, The Linear Complementarity Problem, Academic Press, New York, 1992
[8] C.W. Cryer, The solution of a quadratic programming problem using systematic over-relaxation, SIAM J. Control 9 (1971) 385–392.
[9] Ll. Cvetković, S. Rapajić, How to improve MAOR method convergence area for the linear complementarity problems, Appl. Math. Comput. 162 (2005) 577–584.
[10] S.M. Fallatt, M.J. Tsatsomeros, On the Cayley transform of positivity classes of matrices, Electron. J. Linear Algebra 9 (2002) 190–196.
[11] A. Hadjidimos, M. Tzoumas, On the principle of extrapolation and the Cayley transform, Linear Algebra Appl. 428 (2008) 2761–2777.
[12] N.W. Kappel, L.T. Watson, Iterative algorithms for the linear complementarity problems, Int. J. Comput. Math. 19 (1986) 273–297.
[13] M.D. Koulisianis, T.S. Papatheodorou, Improving projected successive overrelaxation method for linear complementarity problems, Appl. Numer. Math. 45 (2003) 29–40.
[14] C.E. Lemke, Bimatrix equilibrium points and mathematical programming, Management Sci. 11 (1965) 681–689.
[15] O.L. Mangasarian, Solution of symmetric linear complementarity problems by iterative methods, J. Optim. Theory Appl. 22 (1977) 465–485.

[16] O.L. Mangasarian, Nonlinear Programming, McGraw Hill, New York, 1969 (Reprint: SIAM Classics in Applied Mathematics 10, Philadelphia, 1994).
[17] K.G. Murty, Linear Complementarity, Linear and Nolinear Programming. Internet ed., 1997.
[18] J.S. Pang, Necessary and sufficient conditions for the convergence of iterative methods for the linear complementarity problem, J. Optim. Theory Appl. 42 (1984) 1–17.
[19] K. Pantazopoulos, Numerical Methods and Software for the Pricing of American Financial Derivatives, Ph.D. Thesis, Department of Computer Sciences, Purdue University, West Lafayette, IN, 1998.
[20] H. Samelson, R.M. Thrall, O. Wesler, A partitioning theorem for Euclidean n-space, Proc. Amer. Math. Soc. 9 (1958) 805–807.
[21] U. Schäfer, A linear complementarity problem with a P-matrix, SIAM Rev. 46 (2004) 189–201.
[22] U. Schäfer, On the modulus algorithm for the linear complementarity problem, Oper. Res. Lett. 32 (2004) 350–354.
[23] W.M.G. van Bokhoven, Piecewise-linear Modelling and Analysis, Proeschrift, Eindhoven, 1981.
[24] R.S. Varga, Matrix Iterative Analysis. Prentice-Hall, Englewood Cliffs, NJ, 1962 (also: second ed., Revised and Expanded, Springer, Berlin, 2000).
[25] D.M. Young, Iterative Solution of Large Linear Systems, Academic Press, New York, 1971.
[26] D. Yuan, Y. Song, Modified AOR methods for linear complementarity problem, Appl. Math. Comput. 140 (2003) 53–67.