



# Modified Newton-AGSOR method for solving nonlinear systems with block two-by-two complex symmetric Jacobian matrices

Xin Qi<sup>1</sup> · Hui-Ting Wu<sup>1</sup> · Xiao-Yong Xiao<sup>1</sup>

Received: 20 November 2019 / Revised: 26 February 2020 / Accepted: 9 March 2020 /

Published online: 14 March 2020

© Istituto di Informatica e Telematica (IIT) 2020

## Abstract

In this paper, we modify the accelerated generalized successive overrelaxation (AGSOR) method for block two-by-two complex linear systems, and use the AGSOR method as an inner iteration for the modified Newton equations to solve the nonlinear system whose Jacobian matrix is a block two-by-two complex symmetric matrix. Our new method is named modified Newton AGSOR (MN-AGSOR) method. Because generalized successive overrelaxation (GSOR) method is a special form of the AGSOR method, the modified Newton GSOR (MN-GSOR) method is also analyzed in the discussion. Next, we use the Hölder continuous condition instead of the Lipschitz assumption to analyze and prove the local convergence properties of the MN-AGSOR method. At last, numerical experiments verify the efficiency of the MN-AGSOR method, and it can be seen from the comparison of various aspects that the MN-AGSOR method is superior to some other recently proposed methods.

**Keywords** Complex systems of nonlinear equations · Convergence analysis · Modified Newton-DPMHSS · Modified Newton-AGSOR · Modified Newton-GSOR · Modified Newton-MDPMHSS

**Mathematics Subject Classification** 65F10 · 65F50

---

✉ Xin Qi  
15797699680@163.com

✉ Xiao-Yong Xiao  
xiaoxy31@gmail.com

<sup>1</sup> Department of Mathematics, School of Science, Nanchang University, 999 Xuefu Road, Nanchang 330031, People's Republic of China

### 1 Introduction

In this paper, we mainly consider the problem of solving a nonlinear system

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

where  $F : \mathbb{D} \subset \mathbb{C}^{2n} \rightarrow \mathbb{C}^{2n}$  is nonlinear and continuously differentiable. As studied in [6, 19], suppose that the Jacobian matrix of the above nonlinear system is a block two-by-two complex symmetric matrix in the form of

$$F'(x) = \begin{pmatrix} W(x) & iT(x) \\ iT(x) & W(x) \end{pmatrix}.$$

Here the matrices  $W(x), T(x) \in \mathbb{R}^{n \times n}$  are symmetric, and at least one of them is positive definite. The symbol  $i = \sqrt{-1}$  represents the imaginary unit, throughout the paper.

In general, the Newton method is widely used for solving the nonlinear system (1). By applying the Newton method, the solution of the nonlinear system is equivalently obtained by solving the corresponding Newton equation

$$F'(x_k)s_k = -F(x_k) \quad k = 0, 1, \dots, \quad x_{k+1} = s_k + x_k,$$

with  $s_k$  being the unknown vector. In order to improve the convergence speed of the Newton method, the modified Newton method was proposed by Darvishi and Barati [7] in 2007

$$\begin{cases} F'(x_k)d_k = -F(x_k) & k = 0, 1, \dots, \quad \hat{y}_k = d_k + x_k; \\ F'(x_k)h_k = -F(\hat{y}_k) & k = 0, 1, \dots, \quad x_{k+1} = h_k + \hat{y}_k. \end{cases} \tag{2}$$

Unlike the Newton method, the modified Newton method needs only one more evaluation of  $F$  per step, but it has the order of convergence three at least. For the nonlinear system (1) mentioned above, the corresponding Newton equation can be written in the following form:

$$\begin{pmatrix} W(x) & iT(x) \\ iT(x) & W(x) \end{pmatrix} \begin{pmatrix} d_{x1} \\ d_{x2} \end{pmatrix} = \begin{pmatrix} P(x) \\ Q(x) \end{pmatrix}, \tag{3}$$

where  $d_{x1}$  and  $d_{x2}$  are two unknown vectors with dimension  $n$ . As can be seen from the previous discussion, the main task of solving the nonlinear system (1) is to solve the linear system (3), so we pay our attention to solving the block two-by-two complex linear systems

$$\begin{pmatrix} W & iT \\ iT & W \end{pmatrix} \begin{pmatrix} z \\ y \end{pmatrix} = \begin{pmatrix} p \\ q \end{pmatrix}, \tag{4}$$

where the matrices  $W, T \in \mathbb{R}^{n \times n}$  are symmetric, and  $W$  is positive definite. This system is very common in many fields, especially in the finite element method discretization of elliptic PDE problems [1, 12, 14, 16, 17].

Since Bai and Golub proposed the Hermitian/skew-Hermitian splitting (HSS) method [4] in 2003, many novel and effective methods have been proposed for solving the linear system (4), such as the modified HSS (MHSS) [2], preconditioned MHSS (PMHSS)

[3] and generalized PMHSS (GPMHSS) [8] methods. These methods are based on the Hermitian/skew-Hermitian splitting of the coefficient matrix. They can efficiently solve non-Hermitian positive definite system of linear equations. Recently, some single-step iterative methods based on the HSS method have been extensively studied, and they are also quite effective for complex symmetric linear systems [13, 20, 22].

In 2015, the generalized successive overrelaxation (GSOR) [18] method was proposed by Salkuyeh et al. to solve the linear system of equations. Its performance is quite excellent in terms of the convergence speed and accuracy of the solution. In the same year, Edalatpour et al. proposed an accelerated generalized successive overrelaxation (AGSOR) [9] method which is an extension of the GSOR method. Subsequently, the preconditioned GSOR [11] method and the shifted GSOR [10] method were successively proposed. Inspired by their ideas, we propose to apply the AGSOR method to the block two-by-two complex linear systems. By applying the AGSOR method as the inner iteration, and using the modified Newton method as the outer iteration, we present the modified Newton-AGSOR method for solving nonlinear systems with block two-by-two complex Jacobian matrices.

Finally, we give a brief introduction to the structure of this paper. In Sect. 2, we will make convergence analysis of the AGSOR method for block two-by-two complex linear systems. We use the AGSOR method introduced in Sect. 2 as the inner iteration of our new method, and elaborate on the modified Newton-AGSOR (MN-AGSOR) method, including its algorithm and iterative formula in Sect. 3. In Sect. 4, the local convergence of the MN-AGSOR method will be analyzed and proved under the Hölder continuous condition. The numerical results of the MN-AGSOR method are shown and comparisons with several recently proposed methods are made to confirm its effectiveness in Sect. 5. At the end of the paper, we summarize the results of the entire article.

## 2 The AGSOR method for block two-by-two complex linear systems

In this section, we apply the AGSOR method proposed by Edalatpour et al. in [9] to solve block two-by-two complex linear systems, although it is slightly different from the standard AGSOR.

Consider the linear system of equations whose coefficient matrix has complex block two-by-two form, i.e.

$$\begin{pmatrix} W & iT \\ iT & W \end{pmatrix} \begin{pmatrix} z \\ y \end{pmatrix} = \begin{pmatrix} p \\ q \end{pmatrix}, \quad (5)$$

where  $W, T \in \mathbb{R}^{n \times n}$  are symmetric positive definite, and symmetric matrices, respectively. The difference between this system of linear equations and the one solved by AGSOR method in [9] is that the matrix on the skew-diagonal is complex symmetric.

Inspired by the AGSOR method, we can easily establish the iteration algorithm for solving the system (5).

---

**Algorithm 1.** The AGSOR method for solving linear systems (5)

---

1. Given an initial approximation  $x_0 = (z_0^T, y_0^T)^T$ , and positive constants  $\alpha$  and  $\beta$  ;
2. For  $k = 0, 1, 2, \dots$  until  $x_k = (z_k^T, y_k^T)^T$  converges, do

$$\begin{cases} Wz_{k+1} = (1 - \alpha)Wz_k - \alpha iTy_k + \alpha p, \\ Wy_{k+1} = -\beta iTz_{k+1} + (1 - \beta)Wy_k + \beta q. \end{cases} \tag{6}$$

3. End
- 

In fact, the iterative equation (6) can be changed to the following form:

$$\begin{pmatrix} z_{k+1} \\ y_{k+1} \end{pmatrix} = \mathcal{G}_{\alpha,\beta} \begin{pmatrix} z_k \\ y_k \end{pmatrix} + \mathcal{R}_\alpha \begin{pmatrix} p \\ q \end{pmatrix}, \tag{7}$$

where

$$\mathcal{R}_\alpha = \begin{pmatrix} W & 0 \\ \beta iT & W \end{pmatrix}^{-1} \begin{pmatrix} \alpha I & 0 \\ 0 & \beta I \end{pmatrix},$$

and

$$\begin{aligned} \mathcal{G}_{\alpha,\beta} &= \begin{pmatrix} W & 0 \\ \beta iT & W \end{pmatrix}^{-1} \begin{pmatrix} (1 - \alpha)W & -\alpha iT \\ 0 & (1 - \beta)W \end{pmatrix} \\ &= \begin{pmatrix} I & 0 \\ \beta iS & I \end{pmatrix}^{-1} \begin{pmatrix} (1 - \alpha)I & -\alpha iS \\ 0 & (1 - \beta)I \end{pmatrix}, \end{aligned}$$

with  $S = W^{-1}T$  and  $0 \neq \alpha\beta \in \mathbb{R}$ .

So far, we have completed the construction of the AGSOR method adapted to solve the block two-by-two complex linear systems.

Obviously, the GSOR method introduced in [18] is a special form of the AGSOR method [9], with the condition  $\alpha = \beta$ .

**Lemma 1** *Let  $\alpha$  and  $\beta$  be two real numbers with  $\alpha\beta \neq 0$  and  $\mathcal{G}_{\alpha,\beta}$  be the iteration matrix of the AGSOR method adapted to solve the block two-by-two complex linear systems (5). Then, for every eigenvalue  $\lambda$  of  $\mathcal{G}_{\alpha,\beta}$  there is an eigenvalue  $\gamma$  of  $S = W^{-1}T$  which satisfies*

$$(1 - \alpha - \lambda)(1 - \beta - \lambda) = -\lambda\alpha\beta\gamma^2.$$

**Proof** Let  $\alpha \neq \beta$  and  $\lambda$  be an eigenvalue of  $\mathcal{G}_{\alpha,\beta}$  with  $x = (z^T, y^T)^T$  being the eigenvector of  $\lambda$ . Then, we have

$$\begin{pmatrix} I & 0 \\ \beta iS & I \end{pmatrix}^{-1} \begin{pmatrix} (1 - \alpha)I & -\alpha iS \\ 0 & (1 - \beta)I \end{pmatrix} \begin{pmatrix} z \\ y \end{pmatrix} = \lambda \begin{pmatrix} z \\ y \end{pmatrix},$$

which is equivalent to

$$\begin{pmatrix} (1 - \alpha)I & -\alpha iS \\ 0 & (1 - \beta)I \end{pmatrix} \begin{pmatrix} z \\ y \end{pmatrix} = \lambda \begin{pmatrix} I & 0 \\ \beta iS & I \end{pmatrix} \begin{pmatrix} z \\ y \end{pmatrix}.$$

Therefore, we can obtain the following equations:

$$\begin{aligned} (1 - \alpha - \lambda)z &= \alpha iSy, \\ (1 - \beta - \lambda)y &= \beta \lambda iSz. \end{aligned} \tag{8}$$

First we show  $y \neq 0$  if  $\lambda \neq 1 - \alpha$ . In fact, when  $\lambda \neq 1 - \alpha$ , we assume  $y = 0$ , then according to the first equation of (8), we have  $z = y = 0$ , which contradicts with the fact that  $x = (z^T, y^T)^T$  is an eigenvector.

Now, if  $\lambda \neq 1 - \alpha$ , then

$$(1 - \alpha - \lambda)(1 - \beta - \lambda)y = \lambda \alpha \beta (iS)^2 y = -\lambda \alpha \beta S^2 y,$$

which means that for every eigenvalue  $\lambda$  of  $\mathcal{G}_{\alpha, \beta}$ , there is an eigenvalue  $\gamma$  of  $S$  with its corresponding eigenvector  $y$ , i.e.

$$Sy = \gamma y,$$

thus

$$(1 - \alpha - \lambda)(1 - \beta - \lambda) = -\lambda \alpha \beta \gamma^2.$$

If  $\lambda = 1 - \alpha$ , since  $\alpha \neq \beta$  which implies  $\lambda \neq 1 - \beta$ , the conclusion is still correct. In addition, if  $\alpha = \beta$ , by the above process, we can get a simpler equation:

$$(1 - \alpha - \lambda)^2 = -\lambda \alpha^2 \gamma^2. \tag{9}$$

Now we prove the equation (9) only for  $\lambda = 1 - \alpha = 1 - \beta$ , which means that  $Sy = 0$  and  $\lambda Sz = 0$  according to (8). Therefore,  $S$  has an eigenvalue  $\gamma = 0$  or  $\lambda = 0$ , and then (9) holds true. □

**Remark 1** We can see that the above Lemma 1 is consistent with the Lemma 2 in [9] and the conditions of the matrices  $W$  and  $T$  are the same as those given in [9]. Therefore, the AGOSR method for solving linear systems (5) has the same convergence condition as the AGOSR method for solving block two-by-two real linear system, i.e.

$$0 < \alpha\beta < \alpha + \beta < \alpha\beta \frac{1 - \lambda_{max}^2}{2} + 2.$$

Its proof can be found in Theorem 1 in [9]. And the selection of the optimal parameters  $\alpha_*$  and  $\beta_*$  of the AGSOR method is also given in Theorem 2 of [9]. We directly show the results as follows:

$$\alpha_\star = \frac{b_\star + \sqrt{b_\star^2 - 4c_\star}}{2}, \quad \beta_\star = \frac{b_\star - \sqrt{b_\star^2 - 4c_\star}}{2},$$

where

$$b_\star = 4 \frac{1 + \sqrt{(1 + \lambda_{\min}^2)(1 + \lambda_{\max}^2)}}{\sqrt{1 + \lambda_{\min}^2} + \sqrt{1 + \lambda_{\max}^2}}, \quad c_\star = 4 \frac{1}{\sqrt{1 + \lambda_{\min}^2} + \sqrt{1 + \lambda_{\max}^2}}.$$

Here  $\lambda_{\max}$  and  $\lambda_{\min}$  are the largest and smallest eigenvalues of the modulus of the matrix  $S = W^{-1}T$ , respectively. Moreover, it has been mentioned in [9] that the optimal values of parameters  $\alpha_\star$  and  $\beta_\star$  of the AGSOR method are the optimal values of  $\alpha$  and  $\beta$  in the set

$$\Omega_{b,c} = \left\{ \begin{array}{l} 0 < c < b < c \frac{1 - \lambda_{\max}^2}{2} + 2, \\ b, c \in \mathbb{R} \mid b = -c\lambda_{\max}^2 + 2\sqrt{c(1 + \lambda_{\max}^2)}, \\ (\lambda_{\max}^2 + \lambda_{\min}^2)c + 2b - 4 \geq 0 \end{array} \right\}$$

where  $b = \alpha + \beta$  and  $c = \alpha\beta$ .

### 3 The modified Newton-AGSOR method

Based on the extension of the AGSOR method introduced in the previous section, next we will use the AGSOR method as an inner iteration method for the modified Newton method to solve a class of large sparse nonlinear systems with block two-by-two complex symmetric Jacobian matrices.

Now, we focus on the nonlinear system described as

$$F(x) = 0, \tag{10}$$

where  $F : \mathbb{D} \subset \mathbb{C}^{2n} \rightarrow \mathbb{C}^{2n}$  is continuously differentiable and its Jacobian matrix is large sparse and complex symmetric with the following form

$$F'(x) = \begin{pmatrix} W(x) & iT(x) \\ iT(x) & W(x) \end{pmatrix}.$$

In the above expression,  $W(x) \in \mathbb{R}^{n \times n}$  and  $T(x) \in \mathbb{R}^{n \times n}$  are symmetric matrices and  $W(x)$  satisfies the positive definite condition. It should be noted that the condition for  $T(x)$  is weaker than the modified Newton DPMHSS (MN-DPMHSS)[19] and the modified Newton-MDPMHSS (MN-MDPMHSS)[6] methods.

For simplicity of later discussion, we disassemble  $F(x)$  as follows:

$$F(x) = \begin{pmatrix} P(x) \\ Q(x) \end{pmatrix},$$

where  $P(x), Q(x) \in \mathbb{C}^n$ .

When we use AGSOR method as the inner iteration for the modified Newton method, it is equivalent to using AGSOR method to solve the two Newton equations with the following form:

$$\begin{aligned} F'(x_k)d_k &= -F(x_k), & x_{k+\frac{1}{2}} &= d_k + x_k; \\ F'(x_k)h_k &= -F(x_{k+\frac{1}{2}}), & x_{k+1} &= h_k + x_{k+\frac{1}{2}}. \end{aligned} \tag{11}$$

Now, we have a preliminary understanding of the basic structure and specific form of the modified Newton-AGSOR (MN-AGSOR) method. Afterwards, we can apply the MN-AGSOR method for solving nonlinear system (10).

---

**Algorithm 2.** The Modified Newton-AGSOR method

---

1. Given an initial approximation  $x_0$ , two positive integer sequences  $\{l_k\}_{k=0}^\infty$ ,  $\{m_k\}_{k=0}^\infty$  and positive constants  $\alpha, \beta$  and  $tol$ ;
  2. For  $k = 0, 1, 2, \dots$  until  $\|F(x_k)\| \leq tol\|F(x_0)\|$  do;
    - 2.1. Set  $d_{k,0} = (z_{k,0}^T, y_{k,0}^T)^T$ , where  $z_{k,0} = 0, y_{k,0} = 0$ ;
    - 2.2. For  $l = 0, 1, 2, \dots, l_k - 1$ , employ **Algorithm 1** to solve the subsystem (6)
 
$$\begin{cases} W(x_k)z_{k,l+1} = (1 - \alpha)W(x_k)z_{k,l} - \alpha iT(x_k)y_{k,l} + \alpha P(x_k), \\ W(x_k)y_{k,l+1} = -\beta iT(x_k)z_{k,l+1} + (1 - \beta)W(x_k)y_{k,l} + \beta Q(x_k), \end{cases}$$
 to obtain the  $d_{k,l_k} = (z_{k,l_k}^T, y_{k,l_k}^T)^T$  such that  $\|F(x_k) + F'(x_k)d_{k,l_k}\| \leq \delta_k\|F(x_k)\|$  for some  $\delta_k \in [0, 1)$
    - 2.3. Set  $x_{k+\frac{1}{2}} = x_k + d_{k,l_k}$
    - 2.4. Compute  $F(x_{k+\frac{1}{2}})$
    - 2.5. Set  $h_{k,0} = (z_{k,0}^T, y_{k,0}^T)^T$ , where  $z_{k,0} = 0, y_{k,0} = 0$ ;
    - 2.6. For  $m = 0, 1, 2, \dots, m_k - 1$ , employ **Algorithm 1** to solve the subsystem (6)
 
$$\begin{cases} W(x_k)z_{k,l+1} = (1 - \alpha)W(x_k)z_{k,l} - \alpha iT(x_k)y_{k,l} + \alpha P(x_{k+\frac{1}{2}}), \\ W(x_k)y_{k,l+1} = -\beta iT(x_k)z_{k,l+1} + (1 - \beta)W(x_k)y_{k,l} + \beta Q(x_{k+\frac{1}{2}}), \end{cases}$$
 to obtain the  $h_{k,m_k} = (z_{k,l_k}^T, y_{k,l_k}^T)^T$  such that  $\|F(x_{k+\frac{1}{2}}) + F'(x_k)h_{k,m_k}\| \leq \tilde{\delta}_k\|F(x_{k+\frac{1}{2}})\|$  for some  $\tilde{\delta}_k \in [0, 1)$
    - 2.7  $x_{k+1} = x_{k+\frac{1}{2}} + x_{k,m_k}$
  3. End
- 

From the above iteration table, the MN-AGSOR method can be easily rewritten as an equivalent form as follows:

$$\begin{cases} x_{k+\frac{1}{2}} = x_k - \sum_{n=0}^{l_k-1} \mathcal{G}_{\alpha,\beta}^n(x_k) \mathcal{R}_\alpha(x_k) F(x_k), \\ x_{k+1} = x_{k+\frac{1}{2}} - \sum_{n=0}^{m_k-1} \mathcal{G}_{\alpha,\beta}^n(x_k) \mathcal{R}_\alpha(x_k) F\left(x_{k+\frac{1}{2}}\right), \end{cases} \tag{12}$$

where

$$\begin{aligned} \mathcal{G}_{\alpha,\beta}(x) &= \begin{pmatrix} W(x) & 0 \\ \beta iT(x) & W(x) \end{pmatrix}^{-1} \begin{pmatrix} (1-\alpha)W(x) & -\alpha iT(x) \\ 0 & (1-\beta)W(x) \end{pmatrix}, \\ &= \begin{pmatrix} I & 0 \\ \beta iS(x) & I \end{pmatrix}^{-1} \begin{pmatrix} (1-\alpha)I & -\alpha iS(x) \\ 0 & (1-\beta)I \end{pmatrix}, \end{aligned}$$

and

$$\mathcal{R}_{\alpha,\beta}(x) = \begin{pmatrix} W(x) & 0 \\ \beta T(x) & W(x) \end{pmatrix}^{-1} \begin{pmatrix} \alpha I & 0 \\ 0 & \beta I \end{pmatrix},$$

with  $S(x) = W(x)^{-1}T(x)$  and  $0 \neq \alpha\beta \in \mathbb{R}$ .

Notice that the matrix  $W(x)$  is positive definite and, in general, we can use the Cholesky decomposition method or the conjugate gradient method to solve the two linear subsystems in (6).

In addition, we define matrices  $\mathcal{M}_{\alpha,\beta}(x)$  and  $\mathcal{N}_{\alpha,\beta}(x)$  by

$$\begin{aligned} \mathcal{M}_{\alpha,\beta}(x) &= \begin{pmatrix} \frac{1}{\alpha}I & 0 \\ 0 & \frac{1}{\beta}I \end{pmatrix} \begin{pmatrix} W(x) & 0 \\ \beta T(x) & W(x) \end{pmatrix}, \\ \mathcal{N}_{\alpha,\beta}(x) &= \begin{pmatrix} \frac{1}{\alpha}I & 0 \\ 0 & \frac{1}{\beta}I \end{pmatrix} \begin{pmatrix} (1-\alpha)W(x) & -\alpha T(x) \\ 0 & (1-\beta)W(x) \end{pmatrix}. \end{aligned}$$

Then the Jacobian matrix  $F'(x)$  can be split as

$$F'(x) = \mathcal{M}_{\alpha,\beta}(x) - \mathcal{N}_{\alpha,\beta}(x),$$

and the following formulas hold true

$$\begin{aligned} \mathcal{G}_{\alpha,\beta}(x) &= \mathcal{M}_{\alpha,\beta}^{-1}(x)\mathcal{N}_{\alpha,\beta}(x), & \mathcal{M}_{\alpha,\beta}(x) &= \mathcal{R}_{\alpha,\beta}(x)^{-1}, \\ F'(x)^{-1} &= (I - \mathcal{G}_{\alpha,\beta}(x))^{-1}\mathcal{R}_{\alpha,\beta}(x). \end{aligned} \tag{13}$$

Using the above equations, we can convert the expression (12) equivalently to

$$\begin{cases} x_{k+\frac{1}{2}} = x_k - (I - \mathcal{G}_{\alpha,\beta}(x_k)^{l_k})F'(x_k)^{-1}F(x_k), \\ x_{k+1} = x_{k+\frac{1}{2}} - (I - \mathcal{G}_{\alpha,\beta}(x_k)^{m_k})F'(x_k)^{-1}F\left(x_{k+\frac{1}{2}}\right). \end{cases} \tag{14}$$



### 4 Convergence analysis

The main content of this section is to analyze and prove the local convergence properties of the MN-AGSOR iteration method. We will complete our work under the Hölder continuous condition, similar to [5], which is weaker than Lipschitz continuous assumption [21].

Assume that the mapping  $F : \mathbb{D} \subset \mathbb{C}^{2n} \rightarrow \mathbb{C}^{2n}$  is G-differentiable in an open domain  $\mathbb{D}_0 \subset \mathbb{D}$ ,  $F'(x)$  is symmetric and continuous, and there exists a point  $x_* \in \mathbb{D}_0$  satisfying  $F(x_*) = 0$ . For analyzing the convergence properties of the MN-AGSOR method and facilitating the symbolic operation in the proof process, we need some assumptions and appointments.

*Appointment 1* In the article,  $\|z\|$  represents the 2-norm of a matrix or a vector  $z$ . We use the symbols  $\Delta = \min \{ \alpha, \beta \}$  and  $\Lambda = \max \{ |1 - \alpha|, |1 - \beta| \}$ .

**Assumption 1** Suppose  $x_* \in \mathbb{D}_0$  is a solution to  $F(x) = 0$ , and there exists a positive constant  $r$ , for any  $u \in \mathbb{N}(x_*, r) \subset \mathbb{D}_0$ , the following conditions hold.

(A1) (The Bounded Condition) There exist positive constants  $\eta$  and  $\gamma$  satisfying

$$\max \{ \|T(x_*)\|, \|W(x_*)\| \} \leq \eta \quad \text{and} \quad \|F'(x_*)^{-1}\| \leq \gamma.$$

(A2) (The Hölder Condition) For some  $p \in (0, 1]$ , there exist nonnegative constants  $H_t$  and  $H_\omega$  satisfying

$$\begin{aligned} \|T(x_*) - T(x)\| &\leq H_t \|x_* - x\|^p, \\ \|W(x_*) - W(x)\| &\leq H_\omega \|x_* - x\|^p. \end{aligned}$$

The following perturbation lemma is useful for later discussion and analysis of convergence; see Lemma 2.3.2 in [15].

**Lemma 2** Assume that  $M, N \in \mathbb{C}^{n \times n}$ , with  $M$  being nonsingular and  $\|M^{-1}\| \leq \xi$ . If  $\|M - N\| \leq \zeta$  and  $\zeta \xi < 1$ , then  $N^{-1}$  exists and its norm satisfies

$$\|N^{-1}\| \leq \frac{\xi}{1 - \zeta \xi}.$$

**Lemma 3** If  $r \in (0, 1/(\gamma H)^\frac{1}{p})$  with  $H = H_\omega + H_t$  and Assumption 1 holds, then for any  $x, v \in \mathbb{N}(x_*, r) \subset \mathbb{D}_0$ , the matrix  $F'(x)$  is nonsingular and the following four inequalities are true when  $p \in (0, 1]$ , for any  $x, v \in \mathbb{N}(x_*, r)$ :

- (1)  $\|F'(x_*) - F'(x)\| \leq H \|x_* - x\|^p,$
- (2)  $\|F'(x)^{-1}\| \leq \frac{\gamma}{1 - \gamma H \|x_* - x\|^p},$
- (3)  $\|F(v)\| \leq \frac{H}{H_t} \|v - x_*\|^{p+1} + 2\eta \|v - x_*\|,$
- (4)  $S(v) \leq \frac{p H_t}{1 - \gamma H \|x - x_*\|^p} (\frac{1}{p+1} \|v - x_*\|^p + \|x - x_*\|^p) \|v - x_*\|,$

where  $S(v) = \|v - x_* - F'(x)^{-1}F(v)\|$ .

**Proof** Using Assumption (A2), we obtain

$$\begin{aligned} \|F'(x) - F'(x_*)\| &= \left\| \begin{pmatrix} W(x) - W(x_*) & iT(x) - iT(x_*) \\ iT(x) - iT(x_*) & W(x) - W(x_*) \end{pmatrix} \right\| \\ &\leq \left\| \begin{pmatrix} W(x) - W(x_*) & 0 \\ 0 & W(x) - W(x_*) \end{pmatrix} \right\| \\ &\quad + \left\| \begin{pmatrix} 0 & iT(x) - iT(x_*) \\ iT(x) - iT(x_*) & 0 \end{pmatrix} \right\| \\ &= \|W(x) - W(x_*)\| + \|T(x) - T(x_*)\| \\ &\leq H_\omega \|x - x_*\|^p + H_t \|x - x_*\|^p \\ &= H \|x - x_*\|^p. \end{aligned}$$

Then the first formula of Lemma 3 is proved.

Moreover, the condition  $r \in (0, 1/(\gamma H)^{\bar{p}})$  implies  $\gamma H \|x - x_*\|^p < 1$ . Hence by using  $\|F'(x_*)^{-1}\| \leq \gamma$  and Lemma 2, we know that  $F'(x)^{-1}$  exists and satisfies

$$\|F'(x)^{-1}\| \leq \frac{\|F'(x_*)^{-1}\|}{1 - \|F'(x_*)^{-1}\| \|F'(x) - F'(x_*)\|} \leq \frac{\gamma}{1 - \gamma H \|x - x_*\|^p}.$$

Thus the second formula of Lemma 3 is true.

Since

$$\begin{aligned} F(v) &= F(v) - F(x_*) - F'(x_*)(v - x_*) + F'(x_*)(v - x_*) \\ &= \int_0^1 [F'(x_* + t(v - x_*)) - F'(x_*)](v - x_*) dt + F'(x_*)(v - x_*) \end{aligned}$$

and

$$\begin{aligned} \|F'(x_*)\| &\leq \left\| \begin{pmatrix} W(x_*) & 0 \\ 0 & W(x_*) \end{pmatrix} \right\| + \left\| \begin{pmatrix} 0 & iT(x_*) \\ iT(x_*) & 0 \end{pmatrix} \right\| \\ &= \|W(x_*)\| + \|T(x_*)\| \leq 2\eta, \end{aligned}$$

then it holds that

$$\begin{aligned} \|F(v)\| &\leq \int_0^1 \|F'(x_* + t(v - x_*)) - F'(x_*)\| \|v - x_*\| dt + \|F'(x_*)\| \|v - x_*\| \\ &\leq \frac{H}{p+1} \|v - x_*\|^{p+1} + 2\eta \|v - x_*\|. \end{aligned}$$

So the third formula of Lemma 3 is correct.

Next, because

$$\begin{aligned} v - x_* - F'(x)^{-1}F(v) &= -F'(x)^{-1}[F(v) - F(x_*) - F'(x_*)(v - x_*) + (F'(x_*) - F'(x))(v - x_*)] \\ &= -F'(x)^{-1} \left[ \int_0^1 (F'(x_* + t(v - x_*)) - F'(x_*))dt \right. \\ &\quad \left. + (F'(x_*) - F'(x)) \right] (v - x_*), \end{aligned}$$

we have

$$\begin{aligned} &||v - x_* - F'(x)^{-1}F(v)|| \\ &\leq ||F'(x)^{-1}|| \left[ \int_0^1 ||F'(x_* + t(v - x_*)) - F'(x_*)|| dt \right. \\ &\quad \left. + ||F'(x_*) - F'(x)|| \right] ||v - x_*|| \\ &\leq \frac{\gamma}{1 - \gamma H ||x - x_*||^p} \left( \frac{H}{p + 1} ||v - x_*||^p + H ||x - x_*||^p \right) ||v - x_*||. \end{aligned}$$

Now proof of Lemma 3 is complete. □

**Theorem 1** Let  $\Delta = \min \{ \alpha, \beta \}$  and  $\Lambda = \max \{ |1 - \alpha|, |1 - \beta| \}$ . Under the conditions of Lemma 3, assume  $r \in (0, r_0)$ , with  $r_0 = \min \{ r_1, r_2, r_3 \}$ ,

$$\begin{aligned} r_1 &= \left( \frac{1}{2\gamma \left( \frac{1}{\Delta} H_\omega + \frac{\beta}{\Delta} H_t \right)} \right)^{\frac{1}{p}}, \\ r_2 &= \left( \frac{\sigma\theta}{2\gamma \left[ \frac{1}{\Delta} (1 + \Lambda + \sigma\theta) H_\omega + \frac{1}{\Delta} (\sigma\theta\beta + \alpha + \beta) H_t \right]} \right)^{\frac{1}{p}}, \\ r_3 &= \left( \frac{1 - 2\beta\gamma[(\sigma + 1)\theta]^{\mu_*}}{4\gamma H} \right)^{\frac{1}{p}}, \end{aligned}$$

where  $\mu_* = \min \{ m_*, l_* \}$ ,  $m_* = \liminf_{k \rightarrow \infty} m_k$ ,  $l_* = \liminf_{k \rightarrow +\infty} l_k$ , and the constant  $\mu_*$  satisfies

$$\mu_* > \left\lceil -\frac{\ln(2\eta\gamma)}{\ln((1 + \sigma)\theta)} \right\rceil. \tag{15}$$

Here the symbol  $\lceil x \rceil$  is an upper bound function, representing the smallest integer no less than the corresponding real number  $x$ , the number  $\sigma \in (0, \sigma_0)$  is a prescribed positive constant with  $\sigma_0 = \frac{1-\theta}{\theta}$ , and

$$\theta \equiv \theta(\alpha, \beta; x_*) = \|\mathcal{G}_{\alpha, \beta}(x_*)\| < 1 \tag{16}$$

with  $\alpha, \beta$  satisfying  $0 < \alpha\beta < \alpha + \beta < \alpha\beta^{\frac{1-\rho(W(x_*)^{-1}T(x_*))}{2}} + 2$ . Then for all  $x_0 \in \mathbb{N}(x_*, r)$ , any positive integer sequences  $\{l_k\}_{k=0}^\infty$  and  $\{m_k\}_{k=0}^\infty$ , the iteration solution sequence  $\{x_k\}_{k=0}^\infty$  generated by the MN-AGSOR method is well-defined and converges to the solution  $x_*$ . Moreover, it holds that

$$\limsup_{k \rightarrow \infty} \|x_k - x_*\|^{\frac{1}{k}} \leq g(r_0, x_*)^2, \tag{17}$$

where

$$g(s, \lambda) = \frac{\gamma}{1 - \gamma H s^p} [3Hs^p + 2\eta((1 + \sigma)\theta)^\lambda], \text{ for } s \in (0, r) \text{ and } \lambda > \mu_*.$$

**Proof** The idea to prove this theorem is to find some  $r$ , such that for any vector  $x \in \mathbb{N}(x_*, r)$ , it holds that

$$\|\mathcal{G}_{\alpha, \beta}(x)\| \leq (\sigma + 1)\theta < 1,$$

which is satisfied if we can show  $\|\mathcal{G}_{\alpha, \beta}(x_*) - \mathcal{G}_{\alpha, \beta}(x)\| < \sigma\theta$  since

$$\|\mathcal{G}_{\alpha, \beta}(x)\| \leq \|\mathcal{G}_{\alpha, \beta}(x) - \mathcal{G}_{\alpha, \beta}(x_*)\| + \|\mathcal{G}_{\alpha, \beta}(x_*)\|.$$

By using (16) and Assumption (A1), we obtain

$$\begin{aligned} \|\mathcal{M}_{\alpha, \beta}(x_*)^{-1}\| &= \|(I - \mathcal{G}_{\alpha, \beta}(x_*))F'(x_*)^{-1}\| \\ &\leq (1 + \|\mathcal{G}_{\alpha, \beta}(x_*)\|)\|F'(x_*)^{-1}\| < 2\gamma. \end{aligned} \tag{18}$$

In addition,  $r \in (0, 1/(\gamma H)^{\frac{1}{p}})$  implies  $\gamma H \|x - x_*\|^p < 1$ , then by Assumption (A2), we have

$$\begin{aligned}
 & ||\mathcal{M}_{\alpha,\beta}(x_*) - \mathcal{M}_{\alpha,\beta}(x)|| \\
 &= \left\| \begin{pmatrix} \frac{1}{\alpha}I & 0 \\ 0 & \frac{1}{\beta}I \end{pmatrix} \begin{pmatrix} W(x_*) & 0 \\ \beta iT(x_*) & W(x_*) \end{pmatrix} - \begin{pmatrix} \frac{1}{\alpha}I & 0 \\ 0 & \frac{1}{\beta}I \end{pmatrix} \begin{pmatrix} W(x) & 0 \\ \beta iT(x) & W(x) \end{pmatrix} \right\| \\
 &= \left\| \begin{pmatrix} \frac{1}{\alpha}I & 0 \\ 0 & \frac{1}{\beta}I \end{pmatrix} \begin{pmatrix} W(x_*) - W(x) & 0 \\ \beta iT(x_*) - \beta iT(x) & W(x_*) - W(x) \end{pmatrix} \right\| \\
 &\leq \left\| \begin{pmatrix} \frac{1}{\alpha}I & 0 \\ 0 & \frac{1}{\beta}I \end{pmatrix} \right\| \left[ \left\| \begin{pmatrix} W(x_*) - W(x) & 0 \\ 0 & W(x_*) - W(x) \end{pmatrix} \right\| \right. \\
 &\quad \left. + \left\| \begin{pmatrix} 0 & 0 \\ \beta iT(x_*) - \beta iT(x) & 0 \end{pmatrix} \right\| \right] \\
 &\leq \max \left\{ \frac{1}{\alpha}, \frac{1}{\beta} \right\} (||W(x_*) - W(x)|| + \beta ||T(x_*) - T(x)||) \\
 &\leq \frac{1}{\Delta} H_\omega ||x - x_*||^p + \frac{\beta}{\Delta} H_t ||x - x_*||^p \\
 &= \left( \frac{1}{\Delta} H_\omega + \frac{\beta}{\Delta} H_t \right) ||x - x_*||^p,
 \end{aligned}$$

and

$$\begin{aligned}
 & ||\mathcal{N}_{\alpha,\beta}(x_*) - \mathcal{N}_{\alpha,\beta}(x)|| \\
 &= \left\| \begin{pmatrix} \frac{1}{\alpha}I & 0 \\ 0 & \frac{1}{\beta}I \end{pmatrix} \begin{pmatrix} (1 - \alpha)(W(x_*) - W(x)) & -\alpha(T(x_*) - T(x)) \\ 0 & (1 - \beta)(W(x_*) - W(x)) \end{pmatrix} \right\| \\
 &\leq \left\| \begin{pmatrix} \frac{1}{\alpha}I & 0 \\ 0 & \frac{1}{\beta}I \end{pmatrix} \right\| \left\| \begin{pmatrix} (1 - \alpha)(W(x_*) - W(x)) & 0 \\ 0 & (1 - \beta)(W(x_*) - W(x)) \end{pmatrix} \right\| \\
 &\quad + \left\| \begin{pmatrix} \frac{1}{\alpha}I & 0 \\ 0 & \frac{1}{\beta}I \end{pmatrix} \right\| \left\| \begin{pmatrix} 0 & -\alpha(T(x_*) - T(x)) \\ 0 & 0 \end{pmatrix} \right\| \\
 &\leq \max \left\{ \frac{1}{\alpha}, \frac{1}{\beta} \right\} \left[ \max \{ |1 - \alpha|, |1 - \beta| \} ||W(x_*) - W(x)|| + \alpha ||T(x_*) - T(x)|| \right] \\
 &\leq \frac{1}{\Delta} [\Lambda H_\omega ||x_* - x||^p + \alpha H_t ||x_* - x||^p] \\
 &= \left[ \frac{\Lambda}{\Delta} H_\omega + \frac{\alpha}{\Delta} H_t \right] ||x - x_*||^p.
 \end{aligned}$$

Thus, from (18) and Lemma 2, for any  $x \in \mathbb{N}(x_*, r)$ , we obtain

$$\begin{aligned} \|\mathcal{M}_{\alpha,\beta}(x)^{-1}\| &\leq \frac{\|\mathcal{M}_{\alpha,\beta}(x_*)^{-1}\|}{1 - \|\mathcal{M}_{\alpha,\beta}(x_*)^{-1}\| \|\mathcal{M}_{\alpha,\beta}(x_*) - \mathcal{M}_{\alpha,\beta}(x)\|} \\ &\leq \frac{2\gamma}{1 - 2\gamma\left(\frac{1}{\Delta}H_\omega + \frac{\beta}{\Delta}H_t\right)\|x - x_*\|^p}, \end{aligned} \tag{19}$$

given that

$$\|x - x_*\|^p < \frac{1}{2\gamma\left(\frac{1}{\Delta}H_\omega + \frac{\beta}{\Delta}H_t\right)},$$

which is satisfied since  $r < r_1$ .

On the one hand, by direct calculations, we have

$$\begin{aligned} &\|\mathcal{G}_{\alpha,\beta}(x) - \mathcal{G}_{\alpha,\beta}(x_*)\| \\ &= \|\mathcal{M}_{\alpha,\beta}(x)^{-1}\mathcal{N}_{\alpha,\beta}(x) - \mathcal{M}_{\alpha,\beta}(x_*)^{-1}\mathcal{N}_{\alpha,\beta}(x_*)\| \\ &\leq \|\mathcal{M}_{\alpha,\beta}(x)^{-1}\| \|\mathcal{N}_{\alpha,\beta}(x) - \mathcal{N}_{\alpha,\beta}(x_*)\| \\ &\quad + \|\mathcal{M}_{\alpha,\beta}(x) - \mathcal{M}_{\alpha,\beta}(x_*)\| \|\mathcal{G}_{\alpha,\beta}(x_*)\| \\ &\leq \frac{2\gamma\left[\left(\frac{1}{\Delta}H_\omega + \frac{\beta}{\Delta}H_t\right)\|x - x_*\|^p + \left[\frac{\Lambda}{\Delta}H_\omega + \frac{\alpha}{\Delta}H_t\right]\|x - x_*\|^p\right]}{1 - 2\gamma\left(\frac{1}{\Delta}H_\omega + \frac{\beta}{\Delta}H_t\right)\|x - x_*\|^p}. \end{aligned} \tag{20}$$

On the other hand,  $r < r_1$  implies

$$1 - 2\gamma\left(\frac{1}{\Delta}H_\omega + \frac{\beta}{\Delta}H_t\right)\|x - x_*\|^p > 0.$$

Furthermore, in order to make  $\|\mathcal{G}_{\alpha,\beta}(x) - \mathcal{G}_{\alpha,\beta}(x_*)\| < \sigma\theta$ , we only need to show

$$\frac{2\gamma\left[\left(\frac{1}{\Delta}H_\omega + \frac{\beta}{\Delta}H_t\right)\|x - x_*\|^p + \left[\frac{\Lambda}{\Delta}H_\omega + \frac{\alpha}{\Delta}H_t\right]\|x - x_*\|^p\right]}{1 - 2\gamma\left(\frac{1}{\Delta}H_\omega + \frac{\beta}{\Delta}H_t\right)\|x - x_*\|^p} < \sigma\theta,$$

which is equivalent to

$$\|x - x_*\|^p < \frac{\sigma\theta}{2\gamma\left[\left(\frac{1}{\Delta} + \frac{\Lambda}{\Delta} + \frac{1}{\Delta}\sigma\theta\right)H_\omega + \left(\sigma\theta\frac{\beta}{\Delta} + \frac{\alpha+\beta}{\Delta}\right)H_t\right]},$$

and it is true since  $r < r_2$ . Therefore, when  $r < r_1$  and  $r < r_2$ , we have

$$\|\mathcal{G}_{\alpha,\beta}(x) - \mathcal{G}_{\alpha,\beta}(x_*)\| < \sigma\theta.$$

Hence, for any  $u \in \mathbb{N}(x_*, r)$ , with  $r < \min\{r_1, r_2\}$ , we obtain

$$\begin{aligned} \rho(\mathcal{G}_{\alpha,\beta}(x)) &\leq \|\mathcal{G}_{\alpha,\beta}(x)\| \\ &\leq \|\mathcal{G}_{\alpha,\beta}(x_*) - \mathcal{G}_{\alpha,\beta}(x)\| + \|\mathcal{G}_{\alpha,\beta}(x_*)\| \\ &< (1 + \sigma)\theta < 1, \end{aligned} \tag{21}$$

since  $\sigma < \sigma_0 = \frac{1-\theta}{\theta}$ .

Now, we can estimate the error of the iteration sequence  $\{x_k\}_{k=0}^\infty$  generated by MN-AGSOR method. Using (14) and Lemma 3, we get

$$\begin{aligned} \|x_{k+\frac{1}{2}} - x_*\| &= \|x_k - x_* - (I - \mathcal{G}_{\alpha,\beta}(x_k))^{l_k} F'(x_k)^{-1} F(x_k)\| \\ &\leq \|x_k - x_* - F'(x_k)^{-1} F(x_k)\| + \|\mathcal{G}_{\alpha,\beta}(x_k)\|^{l_k} \cdot \|F'(x_k)^{-1} F(x_k)\| \\ &\leq \frac{(p + 2 + ((\sigma + 1)\theta)^{l_k})\gamma H}{(p + 1)(1 - \gamma H\|x_k - x_*\|^p)} \|x_k - x_*\|^{p+1} \\ &\quad + \frac{2\gamma\eta((\sigma + 1)\theta)^{l_k}}{1 - \gamma H\|x_k - x_*\|^p} \|x_k - x_*\| \\ &\leq \frac{\gamma}{1 - \gamma H\|x_k - x_*\|^p} \left[ \frac{3+p}{1+p} H\|x_k - x_*\|^p + 2\eta((\sigma + 1)\theta)^{l_k} \right] \|x_k - x_*\| \\ &\leq \frac{\gamma}{1 - \gamma H\|x_k - x_*\|^p} [3H\|x_k - x_*\|^p + 2\eta((\sigma + 1)\theta)^{l_k}] \|x_k - x_*\| \\ &= g(\|x_k - x_*\|, l_k) \|x_k - x_*\|, \end{aligned}$$

where

$$g(\|x_k - x_*\|, l_k) = \frac{\gamma}{1 - \gamma H\|x_k - x_*\|^p} [3H\|x_k - x_*\|^p + 2\eta((\sigma + 1)\theta)^{l_k}]$$

and  $(1 + \sigma)\theta < 1$ .

Set  $\mu_* = \min\{m_*, l_*\}$ ,  $m_* = \liminf_{k \rightarrow \infty} m_k$ , and  $l_* = \liminf_{k \rightarrow \infty} l_k$ . It is clear that the function  $g(s, \lambda)$  is strictly monotone decreasing with respect to  $\lambda$ . Additionally, by direct calculations, we have

$$\frac{\partial g(s, \lambda)}{\partial s} = \frac{\gamma H p s^{p-1} [3 + 2\gamma\eta((\sigma + 1)\theta)^\lambda]}{(1 - \gamma H s^p)^2} > 0,$$

which implies that  $g(s, \lambda)$  is strictly monotone increasing with respect to  $s$ . Then, for  $x_k \in \mathbb{N}(x_*, r)$ , we get

$$g(\|x_k - x_*\|, l_k) < \frac{\gamma}{1 - \gamma H r^p} (3H r^p + 2\eta((1 + \sigma)\theta)^{\mu_*}) = g(r, \mu_*) < 1,$$

under the conditions that

$$2\eta((1 + \sigma)\theta)^{\mu_*} < 1, \quad \text{and} \quad r < \left( \frac{1 - 2\eta((1 + \sigma)\theta)^{\mu_*}}{4\gamma H} \right)^{\frac{1}{p}}.$$

Actually, the above two inequalities are correct with  $\mu_*$  satisfying (15) and  $r < r_3$ , thus, we obtain

$$\|x_{k+\frac{1}{2}} - x_*\| < \|x_k - x_*\|.$$

Similarly, it holds that

$$\begin{aligned} \|x_{k+1} - x_*\| &= \|x_{k+\frac{1}{2}} - x_* - (I - \mathcal{G}_{\alpha,\beta}(x_k)^{m_k})F'(x_k)^{-1}F\left(x_{k+\frac{1}{2}}\right)\| \\ &\leq \|x_{k+\frac{1}{2}} - x_* - F'(x_k)^{-1}F(x_k)\| + \|\mathcal{G}_{\alpha,\beta}(x_k)\|^{m_k} \cdot \|F'(x_k)^{-1}F\left(x_{k+\frac{1}{2}}\right)\| \\ &\leq \frac{\gamma}{1 - \gamma H\|x_k - x_*\|^p} \left( \frac{1 + ((\sigma + 1)\theta)^{m_k}}{p + 1} H\|x_{k+\frac{1}{2}} - x_*\|^p \right) \|x_{k+\frac{1}{2}} - x_*\| \\ &\quad + \frac{\gamma}{1 - \gamma H\|x_k - x_*\|^p} (H\|x_k - x_*\|^p + 2\eta((\sigma + 1)\theta)^{m_k}) \|x_{k+\frac{1}{2}} - x_*\| \\ &\leq \frac{\gamma g(\|x_k - x_*\|, l_k)}{1 - \gamma H\|x_k - x_*\|^p} \times \|x_k - x_*\| \\ &\quad \times \left( \frac{2g(\|x_k - x_*\|, l_k)^p + 1 + p}{p + 1} H\|x_k - x_*\|^p + 2\eta((\sigma + 1)\theta)^{m_k} \right) \\ &\leq \frac{\gamma g(\|x_k - x_*\|, l_k)}{1 - \gamma H\|x_k - x_*\|^p} \left( 3H\|x_k - x_*\|^p + 2\eta((\sigma + 1)\theta)^{m_k} \right) \|x_k - x_*\| \\ &= g(\|x_k - x_*\|, l_k)g(\|x_k - x_*\|, m_k)\|x_* - x_k\| \\ &< g(r, \mu_*)^2 \|x_* - x_k\| \\ &< \|x_* - x_k\|. \end{aligned}$$

Therefore, for any  $u_0 \in \mathbb{D}(x_*, r) \subset \mathbb{D}_0$ , since

$$0 \leq \dots < \|x_{k+1} - x_*\| < \|x_k - x_*\| < \dots < \|x_0 - x_*\| < r,$$

then we know that the iteration solution sequence  $\{x_k\}_{k=0}^\infty$  generated by the modified Newton-GSOR method is convergent to the solution  $x_*$  and well-defined. Moreover,  $\|x_{k+1} - x_*\| < g(r, x_*)^2 \|u_k - x_*\|$  directly leads to  $\|x_k - x_*\| < g(r_0, x_*)^{2k} \|x_0 - x_*\|$ , or equivalently,

$$\|x_k - x_*\|^{\frac{1}{k}} < g(r_0, x_*)^2 \|x_0 - x_*\|^{\frac{1}{k}},$$

then by letting  $k \rightarrow \infty$ , we establish (17), at this point, proof of this theorem is completed.  $\square$

## 5 Numerical examples

Next, we will make comparisons between the Modified Newton-AGSOR (MN-AGSOR) method, the Modified Newton-DPMHSS (MN-DPMHSS) method [19] and the Modified Newton-MDPMHSS (MN-MDPMHSS) method [6] by several numerical experiments to show the validity and superiority of the MN-AGSOR method. As mentioned earlier, the AGSOR method contains the GSOR method (as long as the two



parameters are the same), so MN-GSOR is a special form of MN-AGSOR, but in the numerical experiments, we use MN-GSOR as an independent method.

Consider the following nonlinear equations [6, 19]

$$\begin{cases} \begin{bmatrix} \tilde{u}_t \\ \tilde{u}_t \end{bmatrix} - \begin{bmatrix} \alpha_1 I_n & i\beta_1 I_n \\ i\beta_1 I_n & \alpha_1 I_n \end{bmatrix} \begin{bmatrix} \tilde{u}_{xx} \\ \tilde{u}_{yy} \end{bmatrix} + \zeta \begin{bmatrix} \tilde{u} \\ \tilde{u} \end{bmatrix} = - \begin{bmatrix} \alpha_2 I_n & i\beta_2 I_n \\ i\beta_2 I_n & \alpha_2 I_n \end{bmatrix} \begin{bmatrix} \tilde{u}^{\frac{4}{3}} \\ \tilde{u}^{\frac{4}{3}} \end{bmatrix}, & \text{in } \mathbb{D}, \\ \tilde{u}(0, x, y) = \tilde{u}_0(x, y), & \text{in } \Omega, \\ \tilde{u}(t, x, y) = 0, & \text{in } (0, 1] \times \partial\Omega, \end{cases} \tag{22}$$

where  $\mathbb{D} = (0, 1] \times \Omega$  with  $(x, y) \in \Omega = (0, 1) \times (0, 1)$ , and  $\partial\Omega$  being the boundary of  $\Omega$ . The constant  $\zeta > 0$  represents the magnitude of the reaction term. By discretizing the above problem on equidistant grids  $\Delta t = h = 1/(N + 1)$ , then at each temporal step of the implicit scheme, we should solve a system of nonlinear equations

$$F(u) = M(\hat{u}^T, \hat{u}^T)^T + \begin{pmatrix} \alpha_2 I_n & i\beta_2 I_n \\ i\beta_2 I_n & \alpha_2 I_n \end{pmatrix} h\Delta t \psi(\hat{u}) = 0, \tag{23}$$

where

$$M = h(1 + \zeta \Delta t)I_{2n} + \frac{\Delta t}{h} \begin{pmatrix} \alpha_1 I_n & i\beta_1 I_n \\ i\beta_1 I_n & \alpha_1 I_n \end{pmatrix} \otimes (A_N \otimes I_N + I_N \otimes A_N),$$

and

$$\psi(\hat{u}) = (\hat{u}_1^{\frac{4}{3}}, \hat{u}_2^{\frac{4}{3}}, \dots, \hat{u}_n^{\frac{4}{3}}, \hat{u}_1^{\frac{4}{3}}, \hat{u}_2^{\frac{4}{3}}, \dots, \hat{u}_n^{\frac{4}{3}})^T \text{ for vector } \hat{u} = (\hat{u}_1, \hat{u}_2, \dots, \hat{u}_n)^T,$$

with  $n = N \times N$  and  $A_N = tridiag(-1, 2, -1)$ . Here  $\otimes$  means the Kronecker product.

Mention that all the numerical tests are finished in Matlab (R2016a) on an Intel quad-core processor (2.79GHz, 8GB RAM). We choose the initial guess  $x_0 = 1$  for all the considered iteration methods, and the program’s termination condition for the outer modified Newton iteration is

$$\frac{\|F(u_k)\|_2}{\|F(u_0)\|_2} \leq 10^{-6}.$$

We set the tolerance of inner iteration methods  $\delta_k = \tilde{\delta}_k = \delta$  for the considered four methods.

It is easy to see that the solution of (23) is  $\hat{u}_* = 0$ , and

$$F'(u) = M + \frac{4}{3}h\Delta t \begin{pmatrix} \alpha_2 I_n & i\beta_2 I_n \\ i\beta_2 I_n & \alpha_2 I_n \end{pmatrix} \times diag\left(\hat{u}_1^{\frac{1}{3}}, \hat{u}_2^{\frac{1}{3}}, \dots, \hat{u}_n^{\frac{1}{3}}\right),$$

thus  $F'(u_*) = M$ . In addition, we can get that

$$\|F'(u_*) - F'(u)\| \leq \frac{4}{3}h\Delta t \sqrt{\alpha_2^2 + \beta_2^2} \|u_* - u\|^{\frac{1}{3}},$$

for any vector  $u \in N(u_*, r)$ .

Hence, the solution sequence  $\{u_k\}$  generated by the MN-AGSOR method converges to the solution  $u_* = 0$  according to the previous theoretical analysis. We obtain the optimal experimental parameters of the MN-DPMHSS method and the MN-MDPMHSS method in different situations from the article [6, 19].

First we choose  $\alpha_1 = \beta_1 = 1$  and  $\alpha_2 = \beta_2 = 1$ . The optimal experimental parameters of the MN-AGSOR method and the MN-GSOR method are obtained by some experimental tests. The detailed data of the experimental optimal parameters are shown in Tables 1, 2, 3 and 4.

In Tables 5, 6, 7 and 8, we compare our MN-AGSOR method and MN-GSOR method with MN-DPMHSS method and MN-MDPMHSS method in the following four aspects: the step of inner iteration steps denoted as ‘‘In Step’’, the step

**Table 1** The experimental optimal parameters  $\alpha, \beta$  of MN-DPMHSS

N	$\zeta = 1$			$\zeta = 10$			$\zeta = 200$		
	$\delta = 0.1$	0.2	0.4	0.1	0.2	0.4	0.1	0.2	0.4
30	$\alpha = 0.90$	0.90	0.90	1.00	0.80	0.90	0.50	0.50	0.50
	$\beta = 1.10$	1.10	1.10	0.90	1.30	1.20	1.00	1.00	0.70
40	$\alpha = 0.70$	0.90	0.90	0.80	1.00	0.70	0.50	0.50	0.50
	$\beta = 1.00$	1.10	1.10	0.90	1.20	1.20	0.90	0.70	0.60
50	$\alpha = 1.00$	0.80	0.80	0.90	0.60	0.60	0.50	0.60	0.40
	$\beta = 1.10$	1.20	1.00	1.10	1.10	0.90	0.70	0.80	0.70

**Table 2** The experimental optimal parameters  $\alpha, \beta$  of MN-MDPMHSS

N	$\zeta = 1$			$\zeta = 10$			$\zeta = 200$		
	$\delta = 0.1$	0.2	0.4	0.1	0.2	0.4	0.1	0.2	0.4
30	$\alpha = 2.10$	2.10	2.00	2.20	2.10	2.10	1.30	1.70	1.90
	$\beta = 0.50$	0.40	0.80	0.40	0.80	0.40	0.60	0.60	0.70
40	$\alpha = 1.80$	1.60	1.90	2.20	2.10	1.30	1.90	1.90	1.40
	$\beta = 0.40$	0.80	0.80	0.50	1.00	0.60	0.90	0.70	0.60
50	$\alpha = 2.10$	2.20	2.00	1.90	2.10	1.90	2.20	1.20	1.70
	$\beta = 0.50$	1.00	0.40	0.40	1.00	0.40	1.00	0.60	0.70

**Table 3** The experimental optimal parameters  $\alpha, \beta$  of MN-AGSOR

N	$\zeta = 1$			$\zeta = 10$			$\zeta = 200$		
	$\delta = 0.1$	0.2	0.4	0.1	0.2	0.4	0.1	0.2	0.4
30	$\alpha = 0.75$	0.85	1.05	0.85	0.75	0.80	0.80	0.75	0.60
	$\beta = 0.85$	0.65	0.50	0.75	0.75	0.70	0.85	0.80	1.05
40	$\alpha = 0.80$	0.70	1.00	0.75	1.00	0.85	0.65	0.70	0.75
	$\beta = 0.80$	0.75	0.55	0.85	0.55	0.65	0.95	0.85	0.80
50	$\alpha = 0.70$	1.05	1.05	0.85	0.80	0.85	0.95	0.70	0.80
	$\beta = 0.90$	0.50	0.50	0.75	0.70	0.65	0.65	0.80	0.70

**Table 4** The experimental optimal parameters  $\alpha$  of MN-GSOR

N	$\zeta = 1$			$\zeta = 10$			$\zeta = 200$		
	$\delta = 0.1$	0.2	0.4	0.1	0.2	0.4	0.1	0.2	0.4
30	$\alpha = 0.80$	0.75	0.75	0.80	0.75	0.75	0.80	0.80	0.80
40	$\alpha = 0.80$	0.75	0.75	0.80	0.75	0.75	0.60	0.60	0.70
50	$\alpha = 0.80$	0.75	0.75	0.80	0.75	0.75	0.80	0.60	0.75

of outer iteration steps denoted as “Out Step”, the elapsed CPU time in seconds denoted as “CPU(s)”, and the error estimates denoted as “RES”.

Next we choose  $\alpha_1 = \alpha_2 = 1$ ,  $\beta_1 = 1/2$  and  $\beta_2 = -2$  for further comparison. At this time, the experimental optimal parameters of the four considered methods are listed in Table 9 for  $N = 50$ , and numerical results of the four methods are presented in Tables 10 and 11, respectively.

**Table 5**  $\delta = 0.1, N = 40$

$\zeta$	Method	RES	CPU(s)	Out Step	In Step
1	MN-MDPMHSS	1.23E-07	0.414	3	15
	MN-DPMHSS	8.01E-08	0.823	3	24
	MN-AGSOR	6.34E-08	0.174	2	12
	MN-GSOR	6.34E-08	0.175	2	12
10	MN-MDPMHSS	1.09E-07	0.421	3	16
	MN-DPMHSS	7.31E-08	0.563	3	24
	MN-AGSOR	5.60E-08	0.170	2	12
	MN-GSOR	6.90E-08	0.197	2	12
200	MN-MDPMHSS	1.57E-07	0.770	3	21
	MN-DPMHSS	1.24E-07	0.529	3	24
	MN-AGSOR	4.03E-07	0.225	3	13
	MN-GSOR	3.32E-07	0.242	3	18

**Table 6**  $\delta = 0.2, N = 40$

$\zeta$	Method	RES	CPU(s)	Out Step	In Step
1	MN-MDPMHSS	6.46E-07	0.476	3	18
	MN-DPMHSS	6.32E-08	0.608	4	24
	MN-AGSOR	8.46E-07	0.229	3	12
	MN-GSOR	4.13E-07	0.260	3	12
10	MN-MDPMHSS	7.56E-07	0.435	3	18
	MN-DPMHSS	7.82E-08	0.555	4	24
	MN-AGSOR	7.16E-07	0.225	3	12
	MN-GSOR	4.64E-07	0.228	3	12
200	MN-MDPMHSS	1.36E-07	0.560	4	21
	MN-DPMHSS	2.01E-07	0.557	4	24
	MN-AGSOR	7.06E-07	0.226	3	12
	MN-GSOR	3.32E-07	0.270	3	18

**Table 7**  $\delta = 0.4, N = 30$ 

$\zeta$	Method	RES	CPU(s)	Out Step	In Step
1	MN-MDPMHSS	4.62E-07	0.167	4	16
	MN-DPMHSS	9.81E-07	0.203	5	20
	MN-AGSOR	9.68E-07	0.077	3	12
	MN-GSOR	4.62E-07	0.088	3	12
10	MN-MDPMHSS	8.67E-07	0.170	5	13
	MN-DPMHSS	6.47E-08	0.231	6	24
	MN-AGSOR	5.64E-07	0.074	3	12
	MN-GSOR	4.57E-07	0.077	3	12
200	MN-MDPMHSS	7.58E-07	0.224	5	19
	MN-DPMHSS	1.68E-07	0.229	6	24
	MN-AGSOR	4.40E-07	0.095	4	15
	MN-GSOR	7.23E-08	0.100	4	16

**Table 8**  $\delta = 0.4, N = 50$ 

$\zeta$	Method	RES	CPU(s)	Out Step	In Step
1	MN-MDPMHSS	4.22E-07	1.039	5	14
	MN-DPMHSS	6.69E-08	1.478	6	24
	MN-AGSOR	8.98E-07	0.469	3	12
	MN-GSOR	4.42E-07	0.469	3	12
10	MN-MDPMHSS	7.45E-07	1.038	5	14
	MN-DPMHSS	1.16E-07	1.510	6	24
	MN-AGSOR	8.18E-07	0.473	3	12
	MN-GSOR	4.82E-07	0.497	3	12
200	MN-MDPMHSS	7.98E-07	1.240	5	18
	MN-DPMHSS	4.33E-07	1.562	6	24
	MN-AGSOR	3.80E-07	0.582	4	14
	MN-GSOR	2.62E-07	0.587	4	14

According to the results of the numerical tests in Tables 5, 6, 7 and 8 and Tables 10 and 11, we can see that the inner and outer iteration steps of the MN-AGSOR method are significantly smaller than the MN-DPMHSS and MN-MDPMHSS methods, and the CPU time of the MN-AGSOR method is significantly less which implies that the modified Newton-AGSOR method is more efficient and superior than the MN-DPMHSS and the MN-MDPMHSS methods. On the other hand, we can see that the performances of the MN-GSOR method and the MN-AGSOR method are similar, but the MN-GSOR method requires only one parameter, so when we handle the scientific and engineering problems, it is better to choose the MN-GSOR method instead of the MN-AGSOR method.

**Table 9** The experimental optimal parameters of the methods for  $N = 50$

Method	$\zeta = 1$		$\zeta = 10$		$\zeta = 200$	
	$\delta = 0.1$	0.2	0.1	0.2	0.1	0.2
MN-MDPMHSS	$\alpha = 1.70$	1.60	1.60	1.50	1.40	1.00
	$\beta = 0.80$	0.90	0.80	0.90	0.70	0.40
MN-DPMHSS	$\alpha = 0.30$	0.70	0.60	0.40	0.30	0.60
	$\beta = 1.00$	1.30	1.00	0.70	0.90	1.20
MN-AGSOR	$\alpha = 0.90$	0.90	0.85	1.05	0.95	1.00
	$\beta = 0.90$	0.90	0.95	0.75	0.85	0.85
MN-GSOR	$\alpha = 0.90$	0.90	0.90	0.90	0.90	0.90

**Table 10**  $\delta = 0.1, N = 50$

$\zeta$	Method	RES	CPU(s)	Out Step	In Step
1	MN-MDPMHSS	2.14E-07	0.819	3	18
	MN-DPMHSS	1.68E-07	0.827	3	18
	MN-AGSOR	1.82E-07	0.260	2	8
	MN-GSOR	1.82E-07	0.283	2	8
10	MN-MDPMHSS	3.07E-07	0.813	3	18
	MN-DPMHSS	8.39E-07	0.83	3	18
	MN-AGSOR	2.25E-07	0.263	2	8
	MN-GSOR	1.94E-07	0.283	2	8
200	MN-MDPMHSS	1.42E-07	0.917	3	21
	MN-DPMHSS	1.71E-07	0.83	3	18
	MN-AGSOR	4.17E-07	0.258	2	8
	MN-GSOR	3.44E-07	0.261	2	8

**Table 11**  $\delta = 0.2, N = 50$

$\zeta$	Method	RES	CPU(s)	Out Step	In Step
1	MN-MDPMHSS	5.70E-07	0.819	3	18
	MN-DPMHSS	8.72E-07	0.823	3	18
	MN-AGSOR	1.18E-07	0.265	2	8
	MN-GSOR	1.82E-07	0.251	2	8
10	MN-MDPMHSS	7.44E-07	0.828	3	18
	MN-DPMHSS	9.67E-07	0.826	3	18
	MN-AGSOR	3.74E-07	0.257	2	8
	MN-GSOR	1.94E-07	0.252	2	8
200	MN-MDPMHSS	2.32E-07	1.033	4	22
	MN-DPMHSS	5.68E-07	0.826	3	18
	MN-AGSOR	2.21E-07	0.256	2	8
	MN-GSOR	3.44E-07	0.260	2	8

## 6 Conclusions

For solving the nonlinear systems with block two-by-two complex symmetric Jacobian matrices, we have introduced a modified Newton-AGSOR method (MN-AGSOR) method based on the AGSOR algorithm. In the theoretical analysis, the local convergence properties of the MN-AGSOR have been discussed under the Hölder continuous condition instead of the stronger Lipschitz assumption. The numerical results confirm that the MN-AGSOR method has the advantage over the modified Newton-DPMHSS and the modified Newton-MDPMHSS methods in both CPU time and iteration steps. Because the performance of the MN-AGSOR method is very close to that of the MN-AGSOR method, we prefer the MN-AGSOR method in applications. Furthermore, the MN-AGSOR method needs less conditions for the Jacobian splitting matrix than the MN-DPMHSS and MN-MDPMHSS methods.

## References

1. Bai, Z.Z.: Block preconditioners for elliptic PDE-constrained optimization problems. *Computing* **91**(4), 379–395 (2011)
2. Bai, Z.Z., Benzi, M., Chen, F.: Modified HSS iteration methods for a class of complex symmetric linear systems. *Computing* **87**(3–4), 93–111 (2010)
3. Bai, Z.Z., Benzi, M., Chen, F.: On preconditioned MHSS iteration methods for complex symmetric linear systems. *Numer. Algorithms* **56**(2), 297–317 (2011)
4. Bai, Z.Z., Golub, G.H., Ng, M.K.: Hermitian and skew-Hermitian splitting methods for non-Hermitian positive definite linear systems. *SIAM J. Matrix Anal. Appl.* **24**(3), 603–626 (2003)
5. Chen, M., Lin, R., Wu, Q.: Convergence analysis of the modified Newton-HSS method under the Hölder continuous condition. *J. Comput. Appl. Math.* **264**, 115–130 (2014)
6. Chen, M.H., Wu, Q.B.: Modified Newton-MDPMHSS method for solving nonlinear systems with block two-by-two complex symmetric Jacobian matrices. *Numer. Algorithms* **80**(2), 355–375 (2019)
7. Darvishi, M., Barati, A.: A third-order Newton-type method to solve systems of nonlinear equations. *Appl. Math. Comput.* **187**(2), 630–635 (2007)
8. Dehghan, M., Dehghani-Madiseh, M., Hajararian, M.: A generalized preconditioned MHSS method for a class of complex symmetric linear systems. *Math. Model. Anal.* **18**(4), 561–576 (2013)
9. Edalatpour, V., Hezari, D., Khojasteh Salkuyeh, D.: Accelerated generalized SOR method for a class of complex systems of linear equations. *Math. Commun.* **20**(1), 37–52 (2015)
10. Edalatpour, V., Hezari, D., Salkuyeh, D.: Two efficient inexact algorithms for a class of large sparse complex linear systems. *Mediterr. J. Math.* **13**(4), 2301–2318 (2016)
11. Hezari, D., Edalatpour, V., Salkuyeh, D.: Preconditioned GSOR iterative method for a class of complex symmetric system of linear equations. *Numer. Linear Algebra Appl.* **22**(4), 761–776 (2015)
12. Karlsson, H.O.: The quasi-minimal residual algorithm applied to complex symmetric linear systems in quantum reactive scattering. *J. Chem. Phys.* **103**(12), 4914–4919 (1995)
13. Li, C.X., Wu, S.L.: A single-step HSS method for non-Hermitian positive definite linear systems. *Appl. Math. Lett.* **44**, 26–29 (2015)
14. Lions, J.L.: *Optimal control of systems governed by partial differential equations problèmes aux limites*. Springer, Berlin (1971)
15. Ortega, J.M., Rheinboldt, W.C.: *Iterative Solution of Nonlinear Equations in Several Variables*. Academic Press, New York (1970)
16. Papp, D., Vizvari, B.: Effective solution of linear Diophantine equation systems with an application in Chemistry. *J. Math. Chem.* **39**(1), 15–31 (2006)
17. Rees, T., Dollar, H.S., Wathen, A.J.: Optimal solvers for PDE-constrained optimization. *SIAM J. Sci. Comput.* **32**(1), 271–298 (2010)

18. Salkuyeh, D.K., Hezari, D., Edalatpour, V.: Generalized successive overrelaxation iterative method for a class of complex symmetric linear system of equations. *Int. J. Comput. Math.* **92**(4), 802–815 (2015)
19. Wang, J., Guo, X.P., Zhong, H.X.: MN-DPMHSS iteration method for systems of nonlinear equations with block two-by-two complex Jacobian matrices. *Numer. Algorithms* **77**(1), 167–184 (2018)
20. Wang, X., Xiao, X.Y., Zheng, Q.Q.: A single-step iteration method for non-Hermitian positive definite linear systems. *J. Comput. Appl. Math.* **346**, 471–482 (2019)
21. Wu, Q., Chen, M.: Convergence analysis of modified Newton-HSS method for solving systems of nonlinear equations. *Numer. Algorithms* **64**(4), 659–683 (2013)
22. Xiao, X.Y., Wang, X., Yin, H.W.: Efficient single-step preconditioned HSS iteration methods for complex symmetric linear systems. *Comput. Math. Appl.* **74**(10), 2269–2280 (2017)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.