Numer Algor (2013) 64:659–683
DOI 10.1007/s11075-012-9684-5

ORIGINAL PAPER

Convergence analysis of modified Newton-HSS method for solving systems of nonlinear equations

Qingbiao Wu · Minhong Chen

Received: 11 January 2012 / Accepted: 10 December 2012 / Published online: 18 January 2013

Abstract Hermitian and skew-Hermitian splitting(HSS) method has been proved quite successfully in solving large sparse non-Hermitian positive definite systems of linear equations. Recently, by making use of HSS method as inner iteration, Newton-HSS method for solving the systems of nonlinear equations with non-Hermitian positive definite Jacobian matrices has been proposed by Bai and Guo. It has shown that the Newton-HSS method outperforms the Newton-USOR and the Newton-GMRES iteration methods. In this paper, a class of modified Newton-HSS methods for solving large systems of nonlinear equations is discussed. In our method, the modified Newton method with R-order of convergence three at least is used to solve the nonlinear equations, and the HSS method is applied to approximately solve the Newton equations. For this class of inexact Newton methods, local and semilocal convergence theorems are proved under suitable conditions. Moreover, a globally convergent modified Newton-HSS method is introduced and a basic global convergence theorem is proved. Numerical results are given to confirm the effectiveness of our method.

Keywords Hermitian and Skew-Hermitian splitting · Newton-HSS method · Large sparse systems · Nonlinear equations · Positive-definite Jacobian matrices · Convergence analysis

Q. Wu · M. Chen (✉)
Department of Mathematics, Zhejiang University, Hangzhou 310027, Zhejiang, People's Republic of China
e-mail: mhchen@zju.edu.cn

Q. Wu
e-mail: qbwu@zju.edu.cn

1 Introduction

In this paper, we consider an effective and robust algorithm for solving large sparse systems of nonlinear equations. In particular, we focus on the type of the systems in which the Jacobian matrices are large, sparse, non-Hermitian and positive-definite. Consider the given nonlinear system

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

where $F : \mathbb{D} \subset \mathbb{C}^n \to \mathbb{C}^n$ is nonlinear and continuously differentiable, and $F = (F_1, \cdots, F_n)^T$ with $F_i = F_i(x)$, $i = 1, 2, \cdots, n$ and $x = (x_1, \cdots, x_n)^T$. The numerical solution for system (1) is often required in many scientific and engineering computing areas such as the discretizations of nonlinear differential and integral equations, see [1, 2]. Inexact Newton methods [3–6] are commonly used for solving such systems, especially when the problems are large and sparse. The inexact Newton methods can be briefly described as follows:

Algorithm 1.1 IN (inexact Newton method)

1. Let x_0 be given.
2. For $k = 0$ step 1 until "convergence" do:
   Find some $\eta_k \in [0, 1)$ and s_k that satisfy
$$\|F(x_k) + F'(x_k)s_k\| \le \eta_k \|F(x_k)\|. \tag{2}$$
3. Set $x_{k+1} = x_k + s_k$.

Here $F'(x_k)$ is the Jacobian matrix and $\eta_k \in [0, 1)$ is commonly called *forcing term* which is used to control the level of accuracy. Obviously, these methods are variants of Newton's method in which the Newton equation

$$F'(x_k)s_k = -F(x_k), \quad k \ge 0, \tag{3}$$

is solved approximately at each iteration. Local convergence analysis for inexact Newton methods [3] shows that if x_0 is sufficiently close to a solution x_* of nonlinear system (1) and the $\eta_k' s$ are uniformly bounded away from one, then the sequence $\{x_k\}$ converges to x_*.

Usually, when the dimension n of the problem is large, linear iterative methods, such as the classical splitting method and the modern Krylov subspace method [7], are applied to approximately solve the Newton equation (3). Thus, the inexact Newton methods are inner-outer iterative methods. We refer to the linear iteration, such as the Krylov subspace method, as an *inner iteration*, and the nonlinear iteration that generates the sequence $\{x_k\}$ as an *outer iteration*. Newton-CG and Newton-GMRES methods, using CG and GMRES methods as inner iterations, respectively, have been successfully used and widely studied, see [8–10].

Recently, based on the use of Hermitian and skew-Hermitian splitting, Bai et al. [11] have proposed the Hermitian and skew-Hermitian splitting(HSS) method for non-Hermitian positive-definite linear systems. The HSS method converges unconditionally to the exact solution of the system of linear equations. Moreover, it has the same upper bound for the convergence rate as that of the CG method when the optimal parameters are used. Because of the effectiveness and robustness of the HSS method, it is extensively studied, see [12–15], and the method also succeeds in solving problems such as stokes problems and distributed control problems and so on, see [16–18].

Using the HSS method as the inner iteration, Bai and Guo [19] have proposed the Newton-HSS method for solving the system of nonlinear equations with non-Hermitian positive-definite Jacobian matrices. The Newton-HSS method is summarized in the following.

Algorithm 1.2 NHSS (Newton-HSS method)

1. Given an initial guess x_0, positive constants α and tol, and a positive integer sequence $\{\ell_k\}_{k=0}^{\infty}$.
2. For $k = 0, 1, \cdots$ until $\|F(x_k)\| \leq tol\|F(x_0)\|$ do:

   2.1. Set $d_{k,0} := 0$.

   2.2. For $\ell = 0, 1, \cdots, \ell_k - 1$, apply Algorithm HSS:

   $$\begin{cases} (\alpha I + H(x_k))d_{k,\ell+\frac{1}{2}} = (\alpha I - S(x_k))d_{k,\ell} - F(x_k), \\ (\alpha I + S(x_k))d_{k,\ell+1} = (\alpha I - H(x_k))d_{k,\ell+\frac{1}{2}} - F(x_k), \end{cases}$$

   and obtain d_{k,ℓ_k} such that

   $$\|F(x_k) + F'(x_k)d_{k,\ell_k}\| \leq \eta_k\|F(x_k)\| \quad \text{for some} \quad \eta_k \in [0, 1),$$

   where $H(x_k) = \frac{1}{2}(F'(x_k) + F'(x_k)^*)$ and $S(x_k) = \frac{1}{2}(F'(x_k) - F'(x_k)^*)$.

   2.3. Set

   $$x_{k+1} = x_k + d_{k,\ell_k}.$$

Numerical results on two-dimensional nonlinear convection-diffusion equations have shown that, the Newton-HSS method costs fewer iteration steps and less CPU time than those in the Newton-USOR, the Newton-GMRES and the Newton-CG methods.

In this paper, we consider the modified Newton method, obtained by Darvish and Barati [20], as the outer iteration method instead of the Newton method. The method reads as follows:

$$\begin{cases} y_k = x_k - F'(x_k)^{-1}F(x_k), \\ x_{k+1} = y_k - F'(x_k)^{-1}F(y_k), \quad k = 0, 1, 2, \cdots. \end{cases} \tag{4}$$

Observe that the modified Newton method (4) requires only one more evaluation of F per step than the Newton method, and also it has R-order of convergence three at least (see [20]) while Newton method converges quadratically. And at each outer iteration step, the following two linear systems

$$\begin{cases} F'(x_k)d_k = -F(x_k), \quad y_k = x_k + d_k, \\ F'(x_k)h_k = -F(y_k), \quad x_{k+1} = y_k + h_k, \end{cases} \tag{5}$$

are needed to be solved. By making use of the HSS method to approximately solve the linear systems (5), a modified Newton-HSS (MN-HSS) method is obtained. Local and semilocal convergence theorems of our method are proved under suitable conditions. In the local convergence theorem, the convergence ball is determined based on the information around the solution x_* and also the convergence rate is obtained. In the semilocal convergence theorem, the convergence criterion is provided based on the information around the initial point x_0. Also a globally convergent modified Newton-HSS method is introduced in our paper and a basic global convergence theorem of the global modified Newton-HSS method is proved.

The rest of the paper is organized as follows. In Section 2, we introduce the modified Newton-HSS (MN-HSS) method. In Section 3, we show that the MN-HSS method is locally convergent. Semilocal and global convergence analyses of the method MN-HSS are presented in Sections 4 and 5. In Section 6, numerical results are given to confirm the effectiveness and robustness of our method. Finally, in Section 7, some brief conclusions are given.

2 The modified Newton-HSS methods

In this section, we describe a nonlinear iteration method for solving systems of nonlinear equations with non-Hermitian positive-definite Jacobian matrices.

First, let us review the HSS method [11].

The HSS iteration method Split linear matrix A into its Hermitian part H and skew-Hermitian part S,

$$A = H + S,$$

where

$$H = \frac{1}{2}(A + A^*) \quad \text{and} \quad S = \frac{1}{2}(A - A^*).$$

Given an initial guess $x_0 \in \mathbb{C}^n$, compute x_{k+1} for $k = 0, 1, 2, \cdots$ using the following iteration scheme until $\{x_k\}$ satisfies the stopping criterion:

$$\begin{cases} (\alpha I + H)x_{k+\frac{1}{2}} = (\alpha I - S)x_k + b, \\ (\alpha I + S)x_{k+1} = (\alpha I - H)x_{k+\frac{1}{2}} + b, \end{cases} \tag{6}$$

where α is a given positive constant and I denotes the identity matrix.

Combining the two equations of (6) into the form

$$x_{k+1} = T(\alpha)x_k + G(\alpha)b \tag{7}$$

leads to

$$x_{k+1} = T(\alpha)^{k+1}x_0 + \sum_{j=0}^{k} T(\alpha)^j G(\alpha)b, \quad k = 0, 1, 2, \cdots, \tag{8}$$

where

$$T(\alpha) = (\alpha I + S)^{-1}(\alpha I - H)(\alpha I + H)^{-1}(\alpha I - S)$$

and

$$G(\alpha) = 2\alpha(\alpha I + S)^{-1}(\alpha I + H)^{-1}.$$

Here, $T(\alpha)$ is the iteration matrix of the HSS method. In fact, splitting A into the form

$$A = B(\alpha) - C(\alpha)$$

with

$$B(\alpha) = \frac{1}{2\alpha}(\alpha I + H)(\alpha I + S),$$
$$C(\alpha) = \frac{1}{2\alpha}(\alpha I - H)(\alpha I - S),$$

also results in the (8), and

$$T(\alpha) = B(\alpha)^{-1}C(\alpha) \text{ and } G(\alpha) = B(\alpha)^{-1}.$$

Now, we employ the modified Newton method (4) as the outer iteration and the HSS method as the inner iteration. In other words, we apply the HSS iteration method to the following linear systems:

$$F'(x_k)d_k = -F(x_k), \quad y_k = x_k + d_k, \tag{9}$$
$$F'(x_k)h_k = -F(y_k), \quad x_{k+1} = y_k + h_k. \tag{10}$$

Then, the modified Newton-HSS method for solving nonlinear system (1) is obtained.

The modified Newton-HSS iteration method Let $F : \mathbb{D} \subset \mathbb{C}^n \to \mathbb{C}^n$ be a continuously differentiable function with the positive-definite Jacobian matrix $F'(x)$ at any $x \in \mathbb{D}$, and let

$$H(x) = \frac{1}{2}(F'(x) + F'(x)^*) \text{ and } S(x) = \frac{1}{2}(F'(x) - F'(x)^*)$$

be its Hermitian and skew-Hermitian parts, respectively. Given an initial guess $x_0 \in \mathbb{D}$, a positive constant α and two sequences $\{l_k\}_{k=0}^{\infty}$ and $\{m_k\}_{k=0}^{\infty}$ of positive integers, compute x_{k+1} for $k = 0, 1, 2, \cdots$ until $\{x_k\}$ converges. The algorithm can be concluded as follows :

Algorithm 2.1 MN-HSS (modified Newton-HSS method)

1. Given an initial guess x_0, positive constants α and tol, and two positive integer sequences $\{l_k\}_{k=0}^{\infty}$, $\{m_k\}_{k=0}^{\infty}$.
2. For $k = 0, 1, \cdots$ until $\|F(x_k)\| \leq tol\|F(x_0)\|$ do:

   2.1. Set $d_{k,0} = h_{k,0} := 0$.

   2.2. For $l = 0, 1, \cdots, l_k - 1$, apply Algorithm HSS to the linear system (9):

   $$\begin{cases} (\alpha I + H(x_k))d_{k,l+\frac{1}{2}} = (\alpha I - S(x_k))d_{k,l} - F(x_k), \\ (\alpha I + S(x_k))d_{k,l+1} = (\alpha I - H(x_k))d_{k,l+\frac{1}{2}} - F(x_k), \end{cases}$$

   and obtain d_{k,l_k} such that

   $$\|F(x_k) + F'(x_k)d_{k,l_k}\| \leq \eta_k\|F(x_k)\| \text{ for some } \eta_k \in [0, 1), \tag{11}$$

   where $H(x_k) = \frac{1}{2}(F'(x_k) + F'(x_k)^*)$ and $S(x_k) = \frac{1}{2}(F'(x_k) - F'(x_k)^*)$.

2.3. Set

$$y_k = x_k + d_{k,l_k}.$$

2.4. Compute $F(y_k)$.

2.5. For $m = 0, 1, \cdots, m_k - 1$, apply Algorithm HSS to the linear system (10):

$$\begin{cases} (\alpha I + H(x_k))h_{k,m+\frac{1}{2}} = (\alpha I - S(x_k))h_{k,m} - F(y_k), \\ (\alpha I + S(x_k))h_{k,m+1} = (\alpha I - H(x_k))h_{k,m+\frac{1}{2}} - F(y_k), \end{cases}$$

and obtain h_{k,m_k} such that

$$\|F(y_k) + F'(x_k)h_{k,m_k}\| \le \tilde{\eta}_k \|F(y_k)\| \quad \text{for some} \quad \tilde{\eta}_k \in [0, 1). \quad (12)$$

2.6. Set

$$x_{k+1} = y_k + h_{k,m_k}.$$

Based on the use of the (8), after straightforward operations we can obtain the following uniform expressions for d_{k,l_k} and h_{k,m_k},

$$d_{k,l_k} = -\sum_{j=0}^{l_k-1} T(\alpha; x_k)^j G(\alpha; x_k) F(x_k),$$

$$h_{k,m_k} = -\sum_{j=0}^{m_k-1} T(\alpha; x_k)^j G(\alpha; x_k) F(y_k),$$

where

$$T(\alpha; x) = (\alpha I + S(x))^{-1}(\alpha I - H(x))(\alpha I + H(x))^{-1}(\alpha I - S(x)),$$

and

$$G(\alpha; x) = 2\alpha(\alpha I + S(x))^{-1}(\alpha I + H(x))^{-1}.$$

Then the modified Newton-HSS method can be rewritten as

$$\begin{cases} y_k = x_k - \sum_{j=0}^{l_k-1} T(\alpha; x_k)^j G(\alpha; x_k) F(x_k), \\ x_{k+1} = y_k - \sum_{j=0}^{m_k-1} T(\alpha; x_k)^j G(\alpha; x_k) F(y_k), \quad k = 0, 1, 2, \cdots. \end{cases}$$

Define matrices $B(\alpha; x)$ and $C(\alpha; x)$ by

$$B(\alpha; x) = \frac{1}{2\alpha}(\alpha I + H(x))(\alpha I + S(x)),$$

$$C(\alpha; x) = \frac{1}{2\alpha}(\alpha I - H(x))(\alpha I - S(x)).$$

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

$$F'(x) = B(\alpha; x) - C(\alpha; x)$$

and

$$T(\alpha; x) = B(\alpha; x)^{-1}C(\alpha; x), \quad B(\alpha; x) = G(\alpha; x)^{-1},$$
$$F'(x)^{-1} = (I - T(\alpha; x))^{-1}G(\alpha; x).$$

Therefore, the modified Newton-HSS method can be equivalently expressed as

$$\begin{cases} y_k = x_k - (I - T(\alpha; x_k)^{l_k})F'(x_k)^{-1}F(x_k), \\ x_{k+1} = y_k - (I - T(\alpha; x_k)^{m_k})F'(x_k)^{-1}F(y_k), \quad k = 0, 1, 2, \cdots . \end{cases} \tag{13}$$

3 Local convergence theorem of the modified Newton-HSS method

In this section we prove that the modified Newton-HSS method has the similar local convergence properties as the Newton-HSS method under the same conditions.

Let $F : \mathbb{D} \subset \mathbb{C}^n \to \mathbb{C}^n$ be G-differentiable on an open neighborhood $\mathbb{N}_0 \subset \mathbb{D}$ of a point $x_* \in \mathbb{D}$ at which $F'(x)$ is continuous, positive definite, and $F(x_*) = 0$. Suppose $F'(x) = H(x) + S(x)$, where $H(x) = \frac{1}{2}(F'(x) + F'(x)^*)$ and $S(x) = \frac{1}{2}(F'(x) - F'(x)^*)$ are the Hermitian and the skew-Hermitian parts of the Jacobian matrix $F'(x)$, respectively. Denote with $\mathbb{N}(x_*, r)$ an open ball centered at x_* with radius $r > 0$.

Assumption 3.1 *For all $x \in \mathbb{N}(x_*, r) \subset \mathbb{N}_0$, assume the following conditions given in* [19] *hold.*

(A1) *(THE BOUNDED CONDITION) there exist positive constants β and γ such that*

$$\max\{\|H(x_*)\|, \ \|S(x_*)\|\} \leq \beta \ \text{ and } \ \|F'(x_*)^{-1}\| \leq \gamma.$$

(A2) *(THE LIPSCHITZ CONDITION) there exist nonnegative constants L_h and L_s such that*

$$\|H(x) - H(x_*)\| \leq L_h\|x - x_*\|,$$
$$\|S(x) - S(x_*)\| \leq L_s\|x - x_*\|.$$

Lemma 3.1 *If $r \in \left(0, \frac{1}{\gamma L}\right)$ and Assumption* 3.1 *holds, then $F'(x)^{-1}$ exists for any $x \in \mathbb{N}(x_*, r) \subset \mathbb{N}_0$. And the following inequalities hold with $L := L_h + L_s$ for all $x, y \in \mathbb{N}(x_*, r)$:*

$$\|F'(x) - F'(x_*)\| \leq L\|x - x_*\|,$$
$$\|F'(x)^{-1}\| \leq \frac{\gamma}{1 - \gamma L\|x - x_*\|},$$
$$\|F(y)\| \leq \frac{L}{2}\|y - x_*\|^2 + 2\beta\|y - x_*\|,$$
$$\|y - x_* - F'(x)^{-1}F(y)\| \leq \frac{\gamma}{1 - \gamma L\|x - x_*\|}\left(\frac{L}{2}\|y - x_*\| + L\|x - x_*\|\right)\|y - x_*\|.$$

Proof The Lipschitz condition directly implies

$$\begin{aligned}\|F'(x)-F'(x_*)\| &= \|H(x)+S(x)-H(x_*)-S(x_*)\| \leq \|H(x)-H(x_*)\| \\ &\quad + \|S(x)-S(x_*)\| \\ &\leq (L_h+L_s)\|x-x_*\| = L\|x-x_*\|.\end{aligned}$$

Hence

$$\|F'(x_*)^{-1}(F'(x_*)-F'(x))\| \leq \|F'(x_*)^{-1}\|\,\|F'(x_*)-F'(x)\| \leq \gamma L\|x-x_*\| < 1.$$

By making use of Banach Lemma, $F'(x)^{-1}$ exists, and

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

Since

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

the bounded condition leads to

$$\|F'(x_*)\| = \|H(x_*)+S(x_*)\| \leq \|H(x_*)\|+\|S(x_*)\| \leq 2\beta$$

and

$$\begin{aligned}\|F(y)\| &\leq \left\|\int_0^1 (F'(x_*+t(y-x_*))-F'(x_*))dt(y-x_*)\right\| + \|F'(x_*)(y-x_*)\| \\ &\leq \frac{L}{2}\|y-x_*\|^2+2\beta\|y-x_*\|.\end{aligned}$$

Clearly, it holds that

$$\begin{aligned}y-x_*-F'(x)^{-1}F(y) &= -F'(x)^{-1}(F(y)-F(x_*)-F'(x)(y-x_*)) \\ &= -F'(x)^{-1}(F(y)-F(x_*)-F'(x_*)(y-x_*)) \\ &\quad + F'(x)^{-1}(F'(x)-F'(x_*))(y-x_*) \\ &= -F'(x)^{-1}\int_0^1 (F'(x_*+t(y-x_*))-F'(x_*))dt(y-x_*) \\ &\quad + F'(x)^{-1}(F'(x)-F'(x_*))(y-x_*).\end{aligned}$$

Therefore

$$\begin{aligned}\|y-x_*-F'(x)^{-1}F(y)\| &= \|-F'(x)^{-1}\int_0^1\big(F'(x_*+t(y-x_*))-F'(x_*)\big)\,dt\,(y-x_*)\\ &\quad+F'(x)^{-1}(F'(x)-F'(x_*))(y-x_*)\|\\ &\le \|-F'(x)^{-1}\|\bigg(\int_0^1\|F'(x_*+t(y-x_*))-F'(x_*)\|dt\\ &\qquad+\|F'(x)-F'(x_*)\|\bigg)\|y-x_*\|\\ &\le \frac{\gamma}{1-\gamma L\|x-x_*\|}\left(\frac{L}{2}\|y-x_*\|+L\|x-x_*\|\right)\|y-x_*\|.\end{aligned}$$

This completes the proof of Lemma 3.1. □

Lemma 3.2 *Under the assumptions of Lemma* 3.1*, suppose* $r \in (0, r_0)$ *and define* $r_0 := \min_{1\le j\le 2}\{r_+^{(j)}\}$*, where*

$$\begin{aligned} r_+^{(1)} &= \frac{\alpha+\beta}{L}\left(\sqrt{\frac{2\tau\alpha\theta}{\gamma(2+\tau\theta)(\alpha+\beta)^2}+1}-1\right),\\ r_+^{(2)} &= \frac{1-2\beta\gamma[(\tau+1)\theta]^u}{3\gamma L},\end{aligned}$$

with $u = \min\{l_*, m_*\}$, $l_* = \liminf_{k\to\infty} l_k$, $m_* = \liminf_{k\to\infty} m_k$*, and the constant* u *satisfies*

$$u > \left\lfloor -\frac{\ln(2\beta\gamma)}{\ln((\tau+1)\theta)}\right\rfloor,$$

where the symbol $\lfloor\cdot\rfloor$ *is used to denote the smallest integer no less than the corresponding real number,* $\tau \in \left(0, \frac{1-\theta}{\theta}\right)$ *a prescribed positive constant and*

$$\theta \equiv \theta(\alpha; x_*) = \|T(\alpha; x_*)\| \le \max_{\lambda\in\sigma(H(x_*))}\frac{|\alpha-\lambda|}{|\alpha+\lambda|} \equiv \sigma(\alpha; x_*).$$

Then, for any $x \in \mathbb{N}(x_*, r) \subset \mathbb{N}_0$, $t \in (0, r)$ *and* $v > u$*, it holds that*

$$\begin{aligned}&\|T(\alpha; x)\| \le (\tau+1)\theta < 1,\\ &g(t; v) = \frac{2\gamma}{1-\gamma Lt}(Lt+\beta[(\tau+1)\theta]^v) < g(r_0; u) < 1.\end{aligned}$$

Proof Details see Theorem 3.2 in [19]. □

Theorem 3.1 *Under the assumptions of Lemmas* 3.1 *and* 3.2*, then for any* $x_0 \in \mathbb{N}(x_*, r)$ *and any sequences* $\{l_k\}_{k=0}^\infty$, $\{m_k\}_{k=0}^\infty$ *of positive integers, the iteration sequence* $\{x_k\}_{k=0}^\infty$ *generated by the modified Newton-HSS method is well-defined and converges to* x_**. Moreover, it holds that*

$$\limsup_{k\to\infty}\|x_k-x_*\|^{\frac{1}{k}} \le g(r_0; u)^2,$$

with $u = \min\{l_*, m_*\}$, $l_* = \liminf_{k\to\infty} l_k$, $m_* = \liminf_{k\to\infty} m_k$.

Proof From Lemmas 3.1 and 3.2, we have

$$\begin{aligned}
\|y_k - x_*\| &= \|x_k - x_* - (I - T(\alpha; x_k)^{l_k})F'(x_k)^{-1}F(x_k)\| \\
&\leq \|x_k - x_* - F'(x_k)^{-1}F(x_k)\| + \|T(\alpha; x_k)^{l_k}\|\,\|F'(x_k)^{-1}F(x_k)\| \\
&\leq \frac{\gamma}{1-\gamma L\|x_k - x_*\|}\frac{3L}{2}\|x_k - x_*\|^2 + [(\tau+1)\theta]^{l_k} \\
&\quad \times \frac{\gamma}{1-\gamma L\|x_k - x_*\|}\left(\frac{L}{2}\|x_k - x_*\|^2 + 2\beta\|x_k - x_*\|\right) \\
&= \frac{(3 + [(\tau+1)\theta]^{l_k})\gamma L}{2(1-\gamma L\|x_k - x_*\|)}\|x_k - x_*\|^2 + \frac{2\beta\gamma[(\tau+1)\theta]^{l_k}}{1-\gamma L\|x_k - x_*\|}\|x_k - x_*\| \\
&\leq \frac{2\gamma}{1-\gamma L\|x_k - x_*\|}(L\|x_k - x_*\| + \beta[(\tau+1)\theta]^{l_k})\|x_k - x_*\| \\
&= g(\|x_k - x_*\|; l_k)\|x_k - x_*\| < g(r_0; u)\|x_k - x_*\| < \|x_k - x_*\|
\end{aligned}$$

and

$$\begin{aligned}
\|x_{k+1} - x_*\| &= \|y_k - x_* - (I - T(\alpha; x_k)^{m_k})F'(x_k)^{-1}F(y_k)\| \\
&\leq \|y_k - x_* - F'(x_k)^{-1}F(y_k)\| + \|T(\alpha; x_k)^{m_k}\|\,\|F'(x_k)^{-1}F(y_k)\| \\
&\leq \frac{\gamma}{1-\gamma L\|x_k - x_*\|}\left(\frac{L}{2}\|y_k - x_*\| + L\|x_k - x_*\|\right)\|y_k - x_*\| \\
&\quad + \frac{\gamma[(\tau+1)\theta]^{m_k}}{1-\gamma L\|x_k - x_*\|}\left(\frac{L}{2}\|y_k - x_*\|^2 + 2\beta\|y_k - x_*\|\right) \\
&\leq \left(\frac{\gamma L}{1-\gamma L\|x_k - x_*\|}\left(\frac{1 + [(\tau+1)\theta]^{m_k}}{2}\|y_k - x_*\| + \|x_k - x_*\|\right)\right. \\
&\quad \left. + \frac{2\beta\gamma[(\tau+1)\theta]^{m_k}}{1-\gamma L\|x_k - x_*\|}\right)\|y_k - x_*\| \\
&\leq \frac{2\gamma g(\|x_k - x_*\|; l_k)}{1-\gamma L\|x_k - x_*\|}\left(\frac{1 + g(\|x_k - x_*\|; l_k)}{2}\right. \\
&\qquad \left. \times L\|x_k - x_*\| + \beta[(\tau+1)\theta]^{m_k}\right)\|x_k - x_*\| \\
&< \frac{2\gamma g(\|x_k - x_*\|; l_k)}{1-\gamma L\|x_k - x_*\|}\left(L\|x_k - x_*\| + \beta[(\tau+1)\theta]^{m_k}\right)\|x_k - x_*\| \\
&= g(\|x_k - x_*\|; l_k)g(\|x_k - x_*\|; m_k)\|x_k - x_*\| \\
&\leq g(\|x_k - x_*\|; u)^2\|x_k - x_*\| < g(r_0; u)^2\|x_k - x_*\| < \|x_k - x_*\|.
\end{aligned}$$

We can further prove that $\{x_k\}_{k=0}^{\infty} \subset \mathbb{N}(x_*, r)$ by induction. When $k = 0$, we can get $\|x_0 - x_*\| < r < r_0$ and

$$\|x_1 - x_*\| < g(\|x_0 - x_*\|; u)^2\|x_0 - x_*\| < \|x_0 - x_*\| < r,$$

since $x_0 \in \mathbb{N}(x_*, r)$. Hence $x_1 \in \mathbb{N}(x_*, r)$. Now, when $k = n$, suppose that $x_n \in \mathbb{N}(x_*, r)$, and then we can straightforwardly deduce the estimate

$$\begin{aligned}
\|x_{n+1} - x_*\| &< g(\|x_n - x_*\|; u)^2\|x_n - x_*\| \\
&< g(r_0; u)^2\|x_n - x_*\| < g(r_0; u)^{2(n+1)}\|x_0 - x_*\| < r,
\end{aligned}$$

which shows that $x_{n+1} \in \mathbb{N}(x_*, r)$ for $k = n + 1$. Moreover, as $n \to \infty$, $x_{n+1} \to x_*$. This completes the proof of Theorem 3.3. □

4 Semilocal convergence of the modified Newton-HSS method

In this section, under the conditions given in [21], we prove the semilocal convergence for the modified Newton-HSS method by using the major function. That is, if the initial point x_0 satisfies some conditions, the existence of the solution of the nonlinear system (1) can be ascertained directly from the iterative process.

Let $F : \mathbb{D} \subset \mathbb{C}^n \to \mathbb{C}^n$ be G-differentiable on an open neighborhood $\mathbb{N}_0 \subset \mathbb{D}$ of a point $x_0 \in \mathbb{D}$ at which $F'(x)$ is continuous and positive definite. Suppose $F'(x) = H(x) + S(x)$, where $H(x) = \frac{1}{2}(F'(x) + F'(x)^*)$ and $S(x) = \frac{1}{2}(F'(x) - F'(x)^*)$ are the Hermitian and skew-Hermitian parts of the Jacobian matrix $F'(x)$, respectively. Denote with $\mathbb{N}(x_0, r)$ an open ball centered at x_0 with radius $r > 0$.

Assumption 4.1 *Let $x_0 \in \mathbb{C}^n$ and assume the following conditions given in* [21] *hold.*

(A1) *(THE BOUNDED CONDITION) there exist positive constants β and γ such that*

$$\max\{\|H(x_0)\|, \|S(x_0)\|\} \le \beta, \quad \|F'(x_0)^{-1}\| \le \gamma, \quad \|F(x_0)\| \le \delta. \tag{14}$$

(A2) *(THE LIPSCHITZ CONDITION) there exist nonnegative constants L_h and L_s such that for all $x, y \in \mathbb{N}(x_0, r) \subset \mathbb{N}_0$,*

$$\|H(x) - H(y)\| \le L_h\|x - y\|, \tag{15}$$

$$\|S(x) - S(y)\| \le L_s\|x - y\|. \tag{16}$$

From Assumption 4.1, the integral mean-value theorem, Banach's Lemma and $F'(x) = H(x) + S(x)$, we can easily get Lemma 4.1 with $L := L_h + L_s$.

Lemma 4.1 *Under Assumption* 4.1*, for all $x, y \in \mathbb{N}(x_0, r)$, if $r \in \left(0, \frac{1}{\gamma L}\right)$, then $F'(x)^{-1}$ exists. And we have the following inequalities:*

$$\begin{aligned}
&\|F'(x) - F'(y)\| \le L\|x - y\|,\\
&\|F'(x)\| \le L\|x - x_0\| + 2\beta,\\
&\|F(x) - F(y) - F'(y)(x - y)\| \le \frac{L}{2}\|x - y\|^2,\\
&\|F'(x)^{-1}\| \le \frac{\gamma}{1 - \gamma L\|x - x_0\|}.
\end{aligned}$$

Define

$$\begin{aligned}
g(t) &= \frac{1}{2}at^2 - bt + c,\\
h(t) &= at - 1,
\end{aligned}$$

with the constants satisfying $a = L\gamma(1+\eta)$, $b = 1-\eta$ and $c = 2\gamma\delta$, where $\eta = \max_k\{\max\{\eta_k, \tilde{\eta}_k\}\} < 1$.

Let $t_0 = 0$, and the sequences $\{t_k\}$, $\{s_k\}$ are generated by the following formula

$$\begin{cases} s_k = t_k - \dfrac{g(t_k)}{h(t_k)}, \\ t_{k+1} = s_k - \dfrac{g(s_k)}{h(t_k)}. \end{cases} \tag{17}$$

Some properties of the functions $g(t)$, $h(t)$ and the sequences $\{t_k\}$, $\{s_k\}$ are given by the following lemmas.

Lemma 4.2 *Assume that the constants satisfy*

$$\gamma^2\delta L \le \frac{(1-\eta)^2}{4(1+\eta)}. \tag{18}$$

Denote $t_* = \frac{b-\sqrt{b^2-2ac}}{a}$, *and then when* $t \in [0, t_*]$, *the following inequalities hold that*

$$\begin{aligned} &g(t) \ge 0,\ \ g'(t) < 0,\ \ g''(t) > 0, \\ &h(t) < g'(t) < 0. \end{aligned}$$

The proof is omitted since it is straightforward.

Lemma 4.3 *Suppose the sequences* $\{t_k\}$, $\{s_k\}$ *are generated by the formula* (17). *Under the assumption of Lemma* 4.2, *then the sequences* $\{t_k\}$, $\{s_k\}$ *increase and converge to* t_*. *Moreover,*

$$\begin{aligned} &0 \le t_k < s_k < t_{k+1} < t_*, \\ &t_{k+1} - s_k < s_k - t_k. \end{aligned}$$

Proof Denote

$$U(x) = x - \frac{g(x)}{h(x)} \ \text{ and } \ \mathrm{V(x) = U(x) - \frac{g(U(x))}{h(x)}}.$$

Then

$$\begin{aligned} U'(x) &= \frac{(h(x) - g'(x))h(x) + g(x)h'(x)}{h(x)^2}, \\ V'(x) &= \frac{\big(h(x) - g'(U(x))\big)h(x)U'(x) + g(U(x))h'(x)}{h(x)^2}. \end{aligned}$$

From Lemma 4.2, we have $U'(x) > 0$ for $x \in [0, t_*]$. So $U(x)$ increases strictly on $[0, t_*]$, and then $x < U(x) < U(t_*) = t_*$ for $x \in [0, t_*]$. It follows that

$h(x) < g'(x) < g'(U(x))$. Hence, it is easy to prove that $V(x)$ increases strictly on $[0, t_*]$, and $U(x) < V(x) < V(t_*) = t_*$. Moreover,

$$s_k = U(t_k), \quad t_{k+1} = V(t_k), \quad t_0 = 0 < t_*.$$

Then we can easily prove the inequality $0 \leq t_k < s_k < t_{k+1} < t_*$ by induction. As $g'(t) < 0$ for $t \in [0, t_*)$, so $g(t_k) > g(s_k)$. Hence

$$t_{k+1} - s_k = -\frac{g(s_k)}{h(t_k)} < -\frac{g(t_k)}{h(t_k)} = s_k - t_k.$$

The proof of Lemma 4.3 completes. □

Theorem 4.1 *Under the assumptions of Lemmas* 4.1 *and* 4.2*, define* $r := \min(r_1, r_2)$ *with*

$$r_1 = \frac{\alpha+\beta}{L}\left(\sqrt{1 + \frac{2\alpha\tau\theta}{(2\gamma+\gamma\tau\theta)(\alpha+\beta)^2}} - 1\right),$$
$$r_2 = \frac{b - \sqrt{b^2 - 2ac}}{a},$$

and define $u = \min\{m_*, l_*\}$ *with* $m_* = \liminf_{k\to\infty} m_k$, $l_* = \liminf_{k\to\infty} l_k$*, and the constant u satisfies* $u > \lfloor \ln\eta / \ln((\tau+1)\theta) \rfloor$*, where the symbol* $\lfloor\cdot\rfloor$ *is used to denote the smallest integer no less than the corresponding real number,* $\tau \in \left(0, \frac{1-\theta}{\theta}\right)$ *and*

$$\theta \equiv \theta(\alpha; x_0) = \|T(\alpha; x_0)\| < 1.$$

Then the iteration sequence $\{x_k\}_{k=0}^{\infty}$ *generated by the modified Newton-HSS method is well-defined and converges to* x_**, which satisfies* $F(x_*) = 0$.

Proof Firstly, we can get the estimate about the iterative matrix $T(\alpha; x)$ of the linear solver. Whenever $x \in \mathbb{N}(x_0, r)$,

$$\|T(\alpha; x)\| \leq (\tau + 1)\theta < 1.$$

As for the proof of the above inequality, see Theorem 3.2 in [21].

Now, we prove by induction

$$\begin{cases} \|x_k - x_0\| \leq t_k - t_0, \\ \|F(x_k)\| \leq \dfrac{1-\gamma L t_k}{(1+\eta)\gamma}(s_k - t_k), \\ \|y_k - x_k\| \leq s_k - t_k, \\ \|F(y_k)\| \leq \dfrac{1-\gamma L t_k}{(1+\eta)\gamma}(t_{k+1} - s_k), \\ \|x_{k+1} - y_k\| \leq t_{k+1} - s_k. \end{cases} \tag{19}$$

Since

$$\begin{aligned}
\|x_0 - x_0\| &= 0 \le t_0 - t_0,\\
\|F(x_0)\| &\le \delta \le \frac{2\gamma\delta}{\gamma(1+\eta)} = \frac{1-\gamma L t_0}{(1+\eta)\gamma}(s_0 - t_0),\\
\|y_0 - x_0\| &= \|I - T(\alpha; x_0)^{l_0}\| \, \|F'(x_0)^{-1}F(x_0)\| \le (1+\theta^{l_0})\gamma\delta < 2\gamma\delta = s_0,\\
\|F(y_0)\| &\le \|F(y_0) - F(x_0) - F'(x_0)(y_0 - x_0)\| + \|F(x_0) + F'(x_0)(y_0 - x_0)\|\\
&\le \frac{L}{2}\|y_0 - x_0\|^2 + \eta\|F(x_0)\| \le \frac{L}{2}s_0^2 + \eta\delta \le \frac{1-\gamma L t_0}{(1+\eta)\gamma}(t_1 - s_0),\\
\|x_1 - y_0\| &\le \|I - T(\alpha; x_0)^{m_0}\| \, \|F'(x_0)^{-1}F(y_0)\|\\
&\le (1+\theta^{m_0})\|F'(x_0)^{-1}\| \, \|F(y_0)\| < (1+\eta)\gamma\|F(y_0)\| \le t_1 - s_0,
\end{aligned}$$

(19) is satisfied for $k = 0$. Suppose that (19) holds for all nonnegative integers less than k. We prove that it holds for k. For the first inequality in (19), we have

$$\|x_k - x_0\| \le \|x_k - y_{k-1}\| + \|y_{k-1} - x_{k-1}\| + \|x_{k-1} - x_0\| \le t_k - t_0 < t_* < r.$$

Since $x_{k-1}, \ y_{k-1} \in \mathbb{N}(x_0, r)$ and by the inequality (11), we have

$$\begin{aligned}
(1+\eta)\gamma\|F(x_k)\| &\le (1+\eta)\gamma\|F(x_k) - F(y_{k-1}) - F'(x_{k-1})(x_k - y_{k-1})\|\\
&\quad + (1+\eta)\gamma\|F(y_{k-1}) + F'(x_{k-1})(x_k - y_{k-1})\|\\
&\le \frac{(1+\eta)\gamma L}{2}\|x_k - y_{k-1}\|^2 + \eta(1+\eta)\gamma\|F(y_{k-1})\|\\
&\le \frac{(1+\eta)\gamma L}{2}(t_k - s_{k-1})^2 + \eta(1-\gamma L t_{k-1})(t_k - s_{k-1})\\
&= g(t_k) - g(s_{k-1}) + b(t_k - s_{k-1}) - a s_{k-1}(t_k - s_{k-1})\\
&\quad + \eta(1-\gamma L t_{k-1})(t_k - s_{k-1})\\
&= g(t_k) - g(s_{k-1}) + (1 - \gamma L(1+\eta)s_{k-1} - \eta\gamma L t_{k-1})\frac{g(s_{k-1})}{-h(t_{k-1})}\\
&= g(t_k) + \frac{(1+\eta)\gamma L s_{k-1} - \gamma L t_{k-1}}{h(t_{k-1})}g(s_{k-1})\\
&< g(t_k) = -h(t_k)(s_k - t_k)\\
&< (1 - \gamma L t_k)(s_k - t_k).
\end{aligned}$$

It follows that

$$\|F(x_k)\| \le \frac{(1-\gamma L t_k)}{(1+\eta)\gamma}(s_k - t_k),$$

and then

$$\begin{aligned}
\|y_k - x_k\| &\le \|I - T(\alpha; x_k)^{l_k}\| \, \|F'(x_k)^{-1}F(x_k)\|\\
&\le \left(1 + ((1+\tau)\theta)^{l_k}\right)\|F'(x_k)^{-1}\| \, \|F(x_k)\|\\
&\le (1+\eta)\frac{\gamma}{1-\gamma L t_k}\|F(x_k)\|\\
&\le s_k - t_k.
\end{aligned}$$

Similarly, we can prove that

$$\|F(y_k)\| \le \frac{(1-\gamma L t_k)}{(1+\eta)\gamma}(t_{k+1}-s_k),$$

and

$$\|x_{k+1}-y_k\| \le t_{k+1}-s_k.$$

Therefore, (19) is true for all k. Since the sequences $\{t_k\}$, $\{s_k\}$ converge to t_* and

$$\begin{aligned}\|x_{k+1}-x_0\| &\le \|x_{k+1}-y_k\| + \|y_k-x_k\| + \|x_k-x_0\|\\ &\le t_{k+1}-t_0 < t_* < r,\end{aligned}$$

the sequence $\{x_k\}$ also converges, to say x_*. Since $\|T(\alpha; x_*)\| < 1$, we have

$$F(x_*) = 0$$

from the iteration (13).

It completes the proof of this theorem. □

5 Global convergence theorem of the modified Newton-HSS method

In the above two sections, we have proved two convergence theorems of the modified Newton-HSS method. One is the local convergence theorem, in which the convergence ball is determined based on the information around the solution x_*, and also the convergence rate is obtained. The other is the semilocal convergence theorem, in which the convergence criterion is provided based on the information around the initial point x_0.

In this section, we discuss the stronger type of convergence, our purpose is to introduce and analyze globally convergent modified Newton-HSS method which is designed to improve convergence from an arbitrary starting point. A basic global convergence result is established to effect that, under reasonable assumptions, that is, if the sequence generated by the iterates has a limit point at which F' is invertible, then that limit point is a solution and the sequence converges to it [22]. Now, we introduce the algorithm of global modified Newton-HSS, Algorithm GMN-HSS.

Algorithm 5.1 GMN-HSS (global modified Newton-HSS method)

1. Given an initial guess x_0 and a positive constant $t \in (0, 1)$.
2. For $k = 0$ step 1 until ∞ do:

    2.1. Find some $\eta_k \in [0, 1)$ and d_k that satisfy

    $$\|F(x_k) + F'(x_k)d_k\| \le \eta_k \|F(x_k)\| \tag{20}$$

    and

    $$\|F(x_k + d_k)\| \le [1 - t(1-\eta_k)]\|F(x_k)\|. \tag{21}$$

    2.2. Set $y_k = x_k + d_k$.

    2.3. Compute $F(y_k)$.

    2.4. Find some $\tilde{\eta}_k \in [0, 1)$ and h_k that satisfy

    $$\|F(y_k) + F'(x_k)h_k\| \le \tilde{\eta}_k \|F(y_k)\| \tag{22}$$

and

$$\|F(y_k + h_k)\| \le [1 - t(1 - \tilde{\eta}_k)]\|F(y_k)\|. \tag{23}$$

2.5. Set $x_{k+1} = y_k + h_k$.

Denote with $\mathbb{N}_\delta(x) = \{y \mid \|y - x\| < \delta\}$, for $\delta > 0$.

Lemma 5.1 ([1, §2.3.3] [22, *Lemma* 1.1]) *Assume that $F'(x)$ is invertible. Then for any $\epsilon > 0$, there exists $\delta > 0$ such that $F'(y)$ is invertible and*

$$\|F'(y)^{-1} - F'(x)^{-1}\| < \epsilon$$

whenever $y \in \mathbb{N}_\delta(x)$.

Lemma 5.2 ([1, §3.2.10] [22, *Lemma* 1.1]) *For any x and $\epsilon > 0$, there exists $\delta > 0$ such that*

$$\|F(z) - F(y) - F'(y)(z - y)\| \le \epsilon\|z - y\|$$

whenever $y, z \in \mathbb{N}_\delta(x)$.

Theorem 5.1 *Assume that $\{x_k\}$, $\{y_k\}$ are two sequences such that $F(x_k) \to 0$ and, for each k,*

$$\|F(x_k) + F'(x_k)d_k\| \le \eta\|F(x_k)\| \ \text{ and } \ \|\mathrm{F(y_k)}\| \le \|\mathrm{F(x_k)}\|,$$
$$\|F(y_k) + F'(x_k)h_k\| \le \eta\|F(y_k)\| \ \text{ and } \ \|\mathrm{F(x_{k+1})}\| \le \|\mathrm{F(y_k)}\|,$$

where $d_k = y_k - x_k$, $h_k = x_{k+1} - y_k$ and η is independent of k. If x_* is a limit point of $\{x_k\}$ such that $F'(x_*)$ is invertible, then $F(x_*) = 0$ and $x_k \to x_*$.

Proof Obviously $F(x_*) = 0$. Set $K \equiv \|F'(x_*)^{-1}\|$. By Lemmas 5.1 and 5.2, for any $z \in \mathbb{N}_\delta(x_*)$, there exists a sufficiently small δ such that $\|F'(z)^{-1}\|$ exists and

$$\begin{aligned}
&\|F'(z)^{-1}\| \le 2K,\\
&\|F(z) - F(x_*) - F'(x_*)(y - x_*)\| \le \frac{1}{2K}\|z - x_*\|.
\end{aligned}$$

If $z \in \mathbb{N}_\delta(x_*)$, then

$$\begin{aligned}
\|F(z)\| &\ge \|F'(x_*)(z - x_*)\| - \|F(z) - F(x_*) - F'(x_*)(z - x_*)\|\\
&\ge \frac{1}{\|F'(x_*)^{-1}\|}\|z - x_*\| - \frac{1}{2K}\|z - x_*\| = \frac{1}{2K}\|z - x_*\|,
\end{aligned}$$

so that, whenever $z \in \mathbb{N}_\delta(x_*)$,

$$\|z - x_*\| \le 2K\|F(z)\|. \tag{24}$$

Let $\epsilon \in (0, \delta/6)$ be given. Since x_* is a limit point of $\{x_k\}$ and $F(x_*) = 0$, there is a sufficiently large k such that

$$x_k \in S_\epsilon \equiv \left\{z \middle| \|z - x_*\| < \delta/3 \ \text{ and } \ \|\mathrm{F(z)}\| < \epsilon/[\mathrm{K}(1 + \eta)]\right\}.$$

Then

$$\begin{aligned}\|d_k\| &= \|F'(x_k)^{-1}(-F(x_k) + (F(x_k) + F'(x_k)d_k))\| \\ &\leq \|F'(x_k)^{-1}\|\big(\|F(x_k)\| + \|F(x_k) + F'(x_k)d_k\|\big) \\ &\leq 2K(1+\eta)\|F(x_k)\| \\ &< 2\epsilon < \delta/3, \\ \|h_k\| &= \|F'(x_k)^{-1}(-F(y_k) + (F(y_k) + F'(x_k)h_k))\| \\ &\leq \|F'(x_k)^{-1}\|\big(\|F(y_k)\| + \|F(y_k) + F'(x_k)h_k\|\big) \\ &\leq 2K(1+\eta)\|F(y_k)\| \\ &\leq 2K(1+\eta)\|F(x_k)\| \\ &< 2\epsilon < \delta/3,\end{aligned}$$

and so

$$\|x_{k+1} - x_*\| \leq \|x_{k+1} - y_k\| + \|y_k - x_k\| + \|x_k - x_*\| < \delta.$$

Since

$$\|F(x_{k+1})\| \leq \|F(x_k)\| < \epsilon/[K(1+\eta)]$$

and, from inequality (24),

$$\|x_{k+1} - x_*\| \leq 2K\|F(x_{k+1})\| < 2K\epsilon/[K(1+\eta)] < \delta/3,$$

it follows that $x_{k+1} \in S_\epsilon$. We conclude that $x_k \in S_\epsilon \subset \mathbb{N}_\delta(x_*)$ for all sufficiently large k, and since $\|F(x_k)\| \to 0$, it follows from inequality (24) that $x_k \to x_*$. □

Now, we prove the basic global convergence theorem of the Algorithm GMN-HSS.

Theorem 5.2 *Assume that Algorithm GMN-HSS does not break down. If $\sum_{k\geq 0}(2 - \eta_k - \tilde{\eta}_k)$ is divergent, then $F(x_k) \to 0$. If, in addition, x_* is a limit point of $\{x_k\}$ such that $F'(x_*)$ is invertible, then $F(x_*) = 0$ and $x_k \to x_*$.*

Proof From the inequalities (21) and (23), we have

$$\|F(y_{k-1})\| \leq [1 - t(1 - \eta_{k-1})]\|F(x_{k-1})\|,$$

and

$$\begin{aligned}\|F(x_k)\| &\leq [1 - t(1 - \tilde{\eta}_{k-1})]\|F(y_{k-1})\| \\ &\leq [1 - t(1 - \eta_{k-1})][1 - t(1 - \tilde{\eta}_{k-1})]\|F(x_{k-1})\| \\ &\leq \|F(x_0)\|\Pi_{0\leq j<k}[1 - t(1 - \eta_j)]\Pi_{0\leq j<k}[1 - t(1 - \tilde{\eta}_j)] \\ &\leq \|F(x_0)\|\exp[-t\Sigma_{0\leq j<k}(1 - \eta_j)]\exp[-t\Sigma_{0\leq j<k}(1 - \tilde{\eta}_j)] \\ &= \|F(x_0)\|\exp[-t\Sigma_{0\leq j<k}(2 - \eta_j - \tilde{\eta}_j)].\end{aligned}$$

Since $t > 0$ and $2 - \eta_j - \tilde{\eta}_j \geq 0$, the divergence of $\sum_{k\geq 0}(2 - \eta_k - \tilde{\eta}_k)$ implies $\|F(x_k)\| \to 0$. The remainder of the theorem follows from Theorem 5.1 with $\eta = 1$.

This completes the proof of the theorem. □

To implement the Algorithm GMN-HSS, there are three main ways: linear search methods, trust region methods and continuation/homotopy methods. The backtracking linear method is used in our method because of its simplicity, and the algorithm is stated as follows.

Algorithm 5.2 Algorithm GMN-HSSB (global modified Newton-HSS method with backtracking)

1. Let $x_0, \eta_{\max} \in [0, 1),\ t \in (0, 1),\ 0 < \theta_{\min} < \theta_{\max} < 1$, and $tol > 0$ be given.
2. While $\|F(x_k)\| > tol\|F(x_0)\|$ and $k < 1000$ do:
    - 2.1. Choose $\eta_k \in [0, \eta_{\max}]$, apply HSS method to the linear system (9) to obtain d_k such that

$$\|F(x_k) + F'(x_k)d_k\| \leq \eta_k\|F(x_k)\|.$$

    - 2.2. Perform the Backtracking Loop(BL) i.e.
        - 2.2.1. Set $\bar{d}_k = d_k$ and $\bar{\eta}_k = \eta_k$.
        - 2.2.2. While $\|F(x_k + d_k)\| > [1 - t(1 - \bar{\eta}_k)]\|F(x_k)\|$ do:
            - 2.2.2.1. Choose $\theta \in [\theta_{\min}, \theta_{\max}]$.
            - 2.2.2.2. Update $\bar{d}_k = \theta\bar{d}_k$ and $\bar{\eta}_k = 1-\theta(1-\bar{\eta}_k)$.
    - 2.3. Set $y_k = x_k + \bar{d}_k$.
    - 2.4. Compute $F(y_k)$.
    - 2.5. Choose $\tilde{\eta}_k \in [0, \eta_{\max}]$, apply HSS method to the linear system (10) to obtain h_k such that

$$\|F(y_k) + F'(x_k)h_k\| \leq \tilde{\eta}_k\|F(y_k)\|.$$

    - 2.6. Perform the Backtracking Loop (BL) i.e.
        - 2.6.1. Set $\bar{h}_k = h_k$ and $\bar{\eta}_k = \tilde{\eta}_k$.
        - 2.6.2. While $\|F(y_k + h_k)\| > [1 - t(1 - \bar{\eta}_k)]\|F(y_k)\|$ do:
            - 2.6.1.1. Choose $\theta \in [\theta_{\min}, \theta_{\max}]$.
            - 2.6.1.2. Update $\bar{h}_k = \theta\bar{h}_k$ and $\bar{\eta}_k = 1-\theta(1-\bar{\eta}_k)$.
    - 2.7. Set $x_{k+1} = y_k + \bar{h}_k$.

6 Numerical examples

In this section, we compare our method with NHSS by the example given in [19], and the numerical results show that our method is more competitive than NHSS. We also solve a model problem BROYDN3D with our method.

Example 1 [19] Let us consider the two-dimensional nonlinear convection-diffusion equation

$$\begin{cases} -(u_{xx} + u_{yy}) + q_1 u_x + q_2 u_y = -e^u, & \text{for } (\mathrm{x}, \mathrm{y}) \in \Omega, \\ u(x, y) = 0, & \text{for } (\mathrm{x}, \mathrm{y}) \in \partial\Omega, \end{cases}$$

where $\Omega = (0, 1) \times (0, 1)$, with $\partial\Omega$ its boundary, and q_1, q_2 are positive constants used to measure magnitudes of the convective terms. By applying the centered finite difference scheme on the equidistant discretization grid with the stepsize $h = \frac{1}{N+1}$, the system of nonlinear equations (1) is obtained with following form

$$F(x) \equiv Mx + h^2\phi(x) = 0,$$

where N is a prescribed positive integer,

$$M = T_x \otimes I + I \otimes T_y,$$
$$\phi(x) = (e^{x_1}, e^{x_2}, \cdots, e^{x_n})^T,$$

with T_x and T_y being tridiagonal matrices given by

$$T_x = \text{tridiag}(-1 - \text{Re}_1, 2, -1 + \text{Re}_1) \text{ and } T_y = \text{tridiag}(-1 - \text{Re}_2, 2, -1 + \text{Re}_2).$$

Here, $\text{Re}_1 = \frac{1}{2}q_j h$, $j = 1, 2$, $\text{Re} = \max\{\text{Re}_1, \text{Re}_2\}$ is the mesh Reynolds number, $\otimes$ the Kronecker product symbol, and $n = N \times N$.

It has been shown by the authors in [19] that Newton-HSS method outperforms the Newton-USOR, the Newton-GMRES and the Newton-GCG methods. So in this paper, we just compare our method with Newton-HSS method. Here we choose the same parameters as those given in [19]. The positive constant $q_2 = \frac{1}{h}$, the initial guess $x_0 = 0$, the stopping criterion for the outer Newton iteration is set to be

$$\frac{\|F(x_k)\|_2}{\|F(x_0)\|_2} \leq 10^{-6},$$

and the prescribed tolerance η_k and $\tilde{\eta}_k$ for controlling the accuracy of the HSS iteration are both set to be η.

In the implementations, we use the optimal parameters α for the Newton-HSS method listed in [19] and we adopt experimentally optimal parameters α for the modified Newton-HSS method, see Tables 1 and 2.

Table 1 The optimal values α for Newton-HSS method

N	$q = 600$			$q = 800$			$q = 1000$		
	$\eta = 0.1$	$\eta = 0.2$	$\eta = 0.4$	$\eta = 0.1$	$\eta = 0.2$	$\eta = 0.4$	$\eta = 0.1$	$\eta = 0.2$	$\eta = 0.4$
30	3.0	2.7	2.9	1.1	1.2	1.1	1.1	1.1	1.4
40	1.3	1.2	1.3	1.2	1.1	1.3	1.4	1.2	1.2
50	1.6	1.5	1.8	1.2	1.5	1.2	1.2	1.2	1.3

Table 2 The optimal values α for modified Newton-HSS method

N	$q = 600$			$q = 800$			$q = 1000$		
	$\eta = 0.1$	$\eta = 0.2$	$\eta = 0.4$	$\eta = 0.1$	$\eta = 0.2$	$\eta = 0.4$	$\eta = 0.1$	$\eta = 0.2$	$\eta = 0.4$
30	3.1	2.9	3.6	3.5	3.3	3.5	3.7	3.8	4.4
40	3.1	2.8	3.4	3.4	3.3	3.5	4.1	3.6	3.7
50	2.0	2.1	2.8	3.0	2.6	2.9	3.2	3	2.8

Table 3 $\eta = 0.1, N = 30$

q	Method	Error estimates	CPU time(s)	Outer IT	Inner IT
600	NHSS	2.5317×10^{-7}	3.399424	6	47
	MN-HSS	3.6278×10^{-7}	2.787244	3	46
800	NHSS	5.7008×10^{-7}	5.049233	6	82
	MN-HSS	6.9747×10^{-7}	3.048042	3	51
1000	NHSS	2.3342×10^{-7}	6.103315	6	103
	MN-HSS	5.7774×10^{-7}	3.405098	3	58

Table 4 $\eta = 0.1, N = 40$

q	Method	Error estimates	CPU time(s)	Outer IT	Inner IT
600	NHSS	3.1709×10^{-7}	14.45626	6	64
	MN-HSS	3.6415×10^{-7}	9.094802	3	46
800	NHSS	3.6561×10^{-7}	15.72509	6	75
	MN-HSS	4.8204×10^{-7}	9.834077	3	51
1000	NHSS	5.2253×10^{-7}	17.04021	6	83
	MN-HSS	6.6359×10^{-7}	10.4798	3	56

Table 5 $\eta = 0.1, N = 50$

q	Method	Error estimates	CPU time(s)	Outer IT	Inner IT
600	NHSS	3.0139×10^{-7}	34.48944	6	53
	MN-HSS	5.6893×10^{-7}	23.76745	3	47
800	NHSS	2.0475×10^{-7}	42.55545	6	79
	MN-HSS	5.3148×10^{-7}	25.13673	3	51
1000	NHSS	4.6039×10^{-7}	45.81205	6	88
	MN-HSS	5.7395×10^{-7}	26.60133	3	56

Table 6 $\eta = 0.2, N = 30$

q	Method	Error estimates	CPU time(s)	Outer IT	Inner IT
600	NHSS	5.6144×10^{-7}	4.063045	8	47
	MN-HSS	5.3797×10^{-7}	2.904874	4	44
800	NHSS	1.8154×10^{-7}	5.722528	9	82
	MN-HSS	8.1726×10^{-7}	3.230078	4	50
1000	NHSS	8.6362×10^{-7}	6.06679	8	93
	MN-HSS	9.5968×10^{-7}	3.460101	4	56

Table 7 $\eta = 0.2, N = 40$

q	Method	Error estimates	CPU time(s)	Outer IT	Inner IT
600	NHSS	1.6504×10^{-7}	17.94886	9	68
	MN-HSS	6.1775×10^{-7}	9.873137	4	45
800	NHSS	8.8256×10^{-7}	16.64199	8	70
	MN-HSS	7.2855×10^{-7}	10.57403	4	50
1000	NHSS	8.6838×10^{-7}	18.11303	8	79
	MN-HSS	9.4418×10^{-7}	11.01499	4	54

Table 8 $\eta = 0.2, N = 50$

q	Method	Error estimates	CPU time(s)	Outer IT	Inner IT
600	NHSS	6.7809×10^{-7}	39.42675	8	52
	MN-HSS	5.7596×10^{-7}	25.86464	4	46
800	NHSS	5.6707×10^{-7}	42.24997	8	63
	MN-HSS	8.3840×10^{-7}	27.31638	4	49
1000	NHSS	2.2986×10^{-7}	55.61846	9	93
	MN-HSS	8.7973×10^{-7}	28.63291	4	54

Table 9 $\eta = 0.4, N = 30$

q	Method	Error estimates	CPU time(s)	Outer IT	Inner IT
600	NHSS	3.6490×10^{-7}	5.099263	13	46
	MN-HSS	4.7466×10^{-7}	3.670551	7	46
800	NHSS	5.7009×10^{-7}	6.756781	14	82
	MN-HSS	6.9747×10^{-7}	3.900176	7	51
1000	NHSS	6.7839×10^{-7}	6.752883	14	80
	MN-HSS	7.0287×10^{-7}	4.106306	7	57

Table 10 $\eta = 0.4, N = 40$

q	Method	Error estimates	CPU time(s)	Outer IT	Inner IT
600	NHSS	9.1867×10^{-7}	19.10262	12	59
	MN-HSS	4.1738×10^{-7}	12.38975	7	46
800	NHSS	3.6332×10^{-7}	22.35775	14	75
	MN-HSS	3.6695×10^{-7}	13.37436	7	52
1000	NHSS	4.3902×10^{-7}	23.67613	14	83
	MN-HSS	3.4567×10^{-7}	13.95786	7	57

The numerical results displayed in Tables 3, 4, 5, 6, 7, 8, 9, 10, 11 indicate that the modified Newton-HSS method outperforms the Newton-HSS in the sense of number of iterations and CPU time.

In Tables 3–11, the number of outer iterations and the total number of inner iterations are denoted with Outer IT and Inner IT, respectively. From the numerical results displayed in Tables 3–11, we observe that the CPU time and the number of outer iterations in NHSS method are greater. The outer iterations in MN-HSS method are about half of those in NHSS method, and the CPU time for the NHSS method is about 1.5 times in average of that for the MN-HSS method. In the case that $N = 50, q = 1000, \eta = 0.4$, the CPU time for the NHSS method is 1.94 times of that for the MN-HSS method.

Example 2 [23] We consider the nonlinear problem BROYDN3D. It contains of n coupled quadratic equations. Let $X = (x_1, \cdots, x_n)$. The nonlinear system $F(X) = 0$ is given by $F = (F_1, \cdots, F_n)$ with

$$F_i(X) = (3 - 2x_i)x_i - x_{i-1} - 2x_{i+1} + 1,$$

where $x_0 = x_{n+1} = 0$ by convention. The initial guess is $X_0 = (-1, \cdots, -1)$.

Table 11 $\eta = 0.4, N = 50$

q	Method	Error estimates	CPU time(s)	Outer IT	Inner IT
600	NHSS	5.3810×10^{-7}	50.87866	13	49
	MN-HSS	3.1395×10^{-7}	34.38023	7	47
800	NHSS	3.0626×10^{-7}	65.59628	14	78
	MN-HSS	7.2375×10^{-7}	36.08129	7	50
1000	NHSS	3.4524×10^{-7}	68.68571	14	87
	MN-HSS	3.2225×10^{-7}	37.03921	7	56

Table 12 The optimal values α for NHSS and MN-HSS method

	100	500	1000	2000
NHSS	4.1	4.2	4.1	4
MN-HSS	4.4	4.3	4.3	4.2

Then

$$F'(X) = \begin{pmatrix} 3-4x_1 & -2 & \cdots & 0 & 0 \\ -1 & 3-4x_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 3-4x_{n-1} & -2 \\ 0 & 0 & \cdots & -1 & 3-4x_n \end{pmatrix}.$$

It is known $F'(X)$ is sparse and positive definite. Now we solve the nonlinear problem by the Newton method, the Newton-HSS method and the modified Newton-HSS Method. They are compared in error estimates, CPU time and the number of iterations. We use experimentally optimal parameters α for the Newton-HSS method and the modified Newton-HSS method corresponding to the problem dimension $n =$ 100, 500, 1000, 2000, see Table 12. The numerical results are displayed in Table 13.

From Table 13, we observe that the MN-HSS method outperforms the NHSS method in the sense of CPU time and the number of iterations, and the CPU time for the NHSS method is about 1.5 times of that for the MN-HSS method.

Table 13 Numerical Results for BROYDN3D

n	Method	Error estimates	CPU time(s)	Outer IT
100	Newton	1.0084×10^{-10}	0.001232	4
	NHSS	8.0089×10^{-7}	0.005718	5
	MN-HSS	6.6696×10^{-7}	0.003833	3
500	Newton	4.7042×10^{-11}	0.071939	4
	NHSS	7.7244×10^{-7}	0.3937667	5
	MN-HSS	4.1772×10^{-7}	0.2647427	3
1000	Newton	3.3482×10^{-11}	0.29026	4
	NHSS	7.4397×10^{-7}	1.982441	5
	MN-HSS	3.3558×10^{-7}	1.363745	3
2000	Newton	2.3793×10^{-11}	1.133462	9
	NHSS	7.1528×10^{-7}	11.299509	5
	MN-HSS	2.7542×10^{-7}	7.53185	3

7 Conclusions

Newton-HSS method is a considerable method for solving large sparse nonlinear systems with non-Hermitian positive definite Jacobian matrices. Based on the Newton-HSS method, we propose a modified Newton-HSS method, in which a modified Newton method which has R-order of convergence three at least is used for solving the nonlinear equations and the HSS method is applied to approximately solve the Newton equations. We have proved the local and semilocal convergence theorems of our method. Moreover, we have introduced a global modified Newton-HSS method and established a basic global convergence theorem. We also choose backtracking linear method to implement the Algorithm GMN-HSS and get the algorithm GMN-HSSB. However, we does not give the choice of the forcing terms in the Algorithm GMN-HSS in our paper, and it is an interesting topic in our future study. Finally, both the two-dimensional nonlinear convection-diffusion system and the BROYDN3D model problem show that the modified Newton-HSS outperforms the Newton-HSS method in the sense of CPU time and number of iterations.

Acknowledgements This work is supported by the National Basic Research 973 Program of China (No. 2011JB105001) and Important National Science & Technology Specific Projects (No. 2009ZX07424-001).

References

1. Ortega, J.M., Rheinboldt, W.C.: Iterative Solution of Nonlinear Equations in Several Variables. Academic Press, New York (1970)
2. Rheinboldt, W.C.: Methods of Solving Systems of Nonlinear Equations, The 2nd Edn. SIAM, Philadelphia (1998)
3. Dembo, R.S., Eisenstat, S.C., Steihaug, T.: Inexact Newton mehtods. SIAM J. Numer. Anal. **19**, 400–408 (1982)
4. Guo, X.-P.: On semilocal convergence of inexact Newton methods. J. Comput. Math. **25**, 231–242 (2007)
5. Eisenstat, S.C., Walker, H.F.: Globally convergent inexact Newton methods. SIAM J. Optim. **4**, 393–422 (1994)
6. Kelley, C.T.: Iterative Methods for Linear and Nonlinear Equations. SIAM, Philadelphia (1995)
7. Saad, Y.: Iterative Methods for Sparse Linear Systems, The 2nd edn. SIAM, Philadelphia (2003)
8. Brown, P.N., Saad, Y.: Convergence theory of nonlinear Newton–Krylov algorithms. SIAM J. Optim. **4**, 297–330 (1994)
9. Brown, P.N., Saad, Y.: Hybrid Krylov methods for nonlinear systems of equations. SIAM J. Sci. Statist. Comput. **11**, 450–481 (1990)
10. An, H.-B., Bai, Z.-Z.: A globally convergent Newton-GMRES method for large sparse systems of nonlinear equations. Appl. Numer. Math. **57**, 235–252 (2007)
11. 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**, 603–626 (2003)
12. Bai, Z.-Z., Golub, G.H., Pan, J.-Y.: Preconditioned Hermitian and skew-Hermitian splitting methods for non-Hermitian positive semidefinite linear systems. Numer. Math. **98**, 1–32 (2004)
13. Bai, Z.-Z., Golub, G.H., Li, C.-K.: Convergence properties of preconditioned Hermitian and skew-Hermitian splitting methods for non-Hermitian positive semidefinite matrices. Math. Comput. **76**, 287–298 (2007)
14. Bai, Z.-Z., Golub, G.H., Lu, L.-Z., Yin, J.-F.: Block triangular and skew-Hermitian splitting methods for positive-definite linear systems. SIAM J. Sci. Comput. **26**, 844–863 (2005)

15. Bai, Z.-Z., Golub, G.H., Li, C.-K.: Optimal parameter in Hermitian and skew-Hermitian splitting method for certain two-by-two block matrices. SIAM J. Sci. Comput. **28**, 583–603 (2006)
16. Bai, Z.-Z., Golub, G.H.: Accelerated Hermitian and skew-Hermitian splitting iteration methods for saddle-point problems. IMA J. Numer. Anal. **27**, 1–23 (2007)
17. Benzi, M., Gander, M.J., Golub, G.H.: Optimization of the Hermitian and skew-Hermitian splitting iteration for saddle-point problems. BIT Numer. Math. **43**, 881–900 (2003)
18. Bai, Z.-Z., Benzi, M., Chen, F., Wang, Z.-Q.: Preconditioned MHSS iteration methods for a class of block two-by-two linear systems with applications to distributed control problems. Technical Report 2011-001, Math/CS Department, Emory University (January 2011). To appear in IMA J. Numer. Anal.
19. Bai, Z.-Z., Guo, X.-P.: The Newton-HSS methods for systems of nonlinear equations with positive-definite Jacobian matrices. J. Comput. Math. **28**, 235–260 (2010)
20. Darvishi, M.T., Barati, A.: A third-order Newton-type method to solve systems of nonlinear equations. Appl. Math. Comput. **187**, 630–635 (2007)
21. Guo, X.-P., Duff, I.S.: Semilocal and global convergence of the Newton-HSS method for systems of nonlinear equations. Numer. Linear Algebra Appl. **18**, 299–315 (2011)
22. Eisenstat, S.C., Walker, H.F.: Globally convergent inexact Newton methods. SIAM J. Optim. **4**, 393–422 (1994)
23. Bargiacchi-Soula, S., Fehrenbach, J., Masmoudi, M.: From linear to nonlinear large scale systems. SIAM J. Matrix. Anal. Appl. **31**, 1552–1569 (2010)