NUMERICAL LINEAR ALGEBRA WITH APPLICATIONS
Numer. Linear Algebra Appl. (2010)
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/nla.713

Semilocal and global convergence of the Newton-HSS method for systems of nonlinear equations

Xue-Ping Guo[1,2,*,†] and Iain S. Duff[2,3]

[1]*Department of Mathematics, East China Normal University, Shanghai 200241, People's Republic of China*
[2]*CERFACS, 42 av. Gaspard Coriolis, 31057 Toulouse, Cedex 1, France*
[3]*RAL, Oxfordshire, England*

SUMMARY

Newton-HSS methods, which are variants of inexact Newton methods different from the Newton–Krylov methods, have been shown to be competitive methods for solving large sparse systems of nonlinear equations with positive-definite Jacobian matrices (*J. Comp. Math.* 2010; **28**:235–260). In that paper, only local convergence was proved. In this paper, we prove a Kantorovich-type semilocal convergence. Then we introduce Newton-HSS methods with a backtracking strategy and analyse their global convergence. Finally, these globally convergent Newton-HSS methods are shown to work well on several typical examples using different forcing terms to stop the inner iterations. Copyright © 2010 John Wiley & Sons, Ltd.

Received 6 January 2009; Revised 9 January 2010; Accepted 28 January 2010

KEY WORDS: systems of nonlinear equations; semilocal convergence; inexact Newton methods; the Newton-HSS method; globally convergent Newton-HSS method

1. INTRODUCTION

Consider solving the system of nonlinear equations with n equations in n variables:

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

where $F:\mathscr{D}\subset\mathscr{C}^n\rightarrow\mathscr{C}^n$ is a nonlinear continuously differentiable operator mapping from an open convex subset $\mathscr{D}$ of the n-dimensional complex linear space $\mathscr{C}^n$ into $\mathscr{C}^n$, and the Jacobian matrix $F'(x)$ is sparse, nonsymmetric and positive definite. This is satisfied in many practical cases [1–3]. The Newton method is the most common iterative method for solving (1.1) (see [4, 5], for

*Correspondence to: Xue-Ping Guo, Department of Mathematics, East China Normal University, Shanghai 200241, People's Republic of China.
†E-mail: xpguo@math.ecnu.edu.cn

Contract/grant sponsor: NSFC; contract/grant number: 10971070
Contract/grant sponsor: EPSRC; contract/grant number: EP/E053351/1

example). It has the following form:

$$x_{k+1} = x_k - F'(x_k)^{-1} F(x_k), \quad k = 0, 1, \ldots. \tag{2}$$

Hence, it is necessary to solve the Newton equation

$$F'(x_k) s_k = -F(x_k), \tag{3}$$

to obtain the $(k+1)$th iteration $x_{k+1} = x_k + s_k$. Equation (3) is a system of linear equations that we denote generally by

$$Ax = b. \tag{4}$$

In general, there are two types of iterative methods for solving (4) [6]. One comprises nonstationary iterative methods such as the Krylov methods. If Krylov subspace methods are used to solve the Newton equation, then we get Newton–Krylov subspace methods. We call the linear iteration, for example the Krylov subspace iteration, an *inner iteration*, whereas the nonlinear iteration that generates the sequence $\{x_k\}$ is an *outer iteration*. Newton-CG and Newton-GMRES iterations, using CG and GMRES as an inner iteration, respectively, are widely studied [6–10]. The second type of iterative methods that include methods such as Jacobi, Gauss–Seidel and successive overrelaxation (SOR) are classical stationary iterative methods. These methods do not depend on the history of their iterations. They are based on splittings of A. When splitting the coefficient matrix A of the linear equation into B and C, $A = B - C$, splitting methods to solve (4) of the form

$$Bx_\ell = Cx_{\ell-1} + b, \quad \ell = 0, 1, \ldots \tag{5}$$

are obtained. Hence, if these methods are regarded as inner iterations (and we assume that as is common the initial iterate is 0), we obtain the inner/outer iteration [3, 4, 11–17]

$$\begin{aligned} &x_0 \text{ given},\\ &x_{k+1} = x_k - (T_k^{\ell_k - 1} + \cdots + T_k + I) B_k^{-1} F(x_k),\\ &T_k = B_k^{-1} C_k,\\ &F'(x_k) = B_k - C_k, \quad k = 0, 1, \ldots, \end{aligned} \tag{6}$$

where ℓ_k is the number of inner iteration steps.

Bai *et al.* [3] have proposed the Hermitian/skew-Hermitian splitting (HSS) method for non-Hermitian positive-definite linear systems based on the Hermitian and skew-Hermitian splittings. They have proved that this method converges unconditionally to the unique solution of the system of linear equations and, when the optimal parameters are used, it has the same upper bound for the convergence rate as that of the CG method.

In [18], Bai and Guo use the HSS method as the inner iteration and obtain the Newton-HSS method to solve the system of nonlinear equations with non-Hermitian positive-definite Jacobian matrices. Numerical results on two-dimensional nonlinear convection–diffusion equations have shown that the Newton-HSS method considerably outperforms the Newton-USOR, the Newton-GMRES and the Newton-GCG methods in the sense of number of iterations and CPU time.

There are three fundamental problems concerning the convergence of the iteration [4]. The first is local convergence that assumes a particular solution x_*. The second type of convergence, called semilocal, does not require knowledge of the existence of a solution, but imposes all the conditions on the initial vectors. Finally, global convergence, the third and most elegant type of convergence

DOI: 10.1002/nla

result, states that beginning from an arbitrary point in $\mathscr{C}^n$, or at least in a large part of it, the iterates will converge to a solution. In [18], we gave two types of local convergence theorems. In this paper, we will first present semilocal convergence theorems for the Newton-HSS method. Then, to obtain the globally convergent result, we define a Newton-HSS method with backtracking and prove its global convergence. Finally, computational results are demonstrated.

2. PRELIMINARIES

Throughout this paper, the norm is the Euclidean norm. We denote by $\mathscr{B}(x,r) \equiv \{y \mid \|y-x\| < r\}$ an open ball centred at x with radius $r > 0$, whereas $\overline{\mathscr{B}(x,r)}$ is its closed ball. A^* represents the conjugate transpose of A. We also use $x_{k,\ell}$ with subscripts k as the step of the outer iteration and ℓ as the step of the inner iteration, respectively.

Inexact Newton methods [19] compute an approximate solution of the Newton equation as follows:

Algorithm IN [*19*]

1. Given x_0 and a positive constant tol.
2. For $k=0,1,2,\ldots$ until $\|F(x_k)\| \leqslant \text{tol}\|F(x_0)\|$ do:

   2.1. For a given $\eta_k \in [0,1)$ find s_k that satisfies
   $$\|F(x_k)+F'(x_k)s_k\| < \eta_k \|F(x_k)\|.$$

   2.2. Set $x_{k+1} = x_k + s_k$.

For large sparse non-Hermitian and positive-definite systems of linear equations (4), the HSS iteration method [3, 20] can be written as

Algorithm HSS [*3*]

1. Given an initial guess x_0, and positive constants α and tol.
2. Split A into its Hermitian part H and its skew-Hermitian part S
$$H = \tfrac{1}{2}(A+A^*) \quad \text{and} \quad S = \tfrac{1}{2}(A-A^*).$$
3. For $\ell = 0,1,2,\ldots$ until $\|b-Ax_\ell\| \leqslant \text{tol}\|b-Ax_0\|$, compute $x_{\ell+1}$ by
$$\begin{aligned}(\alpha I+H)x_{\ell+1/2} &= (\alpha I-S)x_\ell + b,\\ (\alpha I+S)x_{\ell+1} &= (\alpha I-H)x_{\ell+1/2}+b.\end{aligned}$$

3. SEMILOCAL CONVERGENCE OF THE NEWTON-HSS METHOD

Now we present a Newton-HSS algorithm to solve large systems of nonlinear equations with a positive-definite Jacobian matrix (1):

Algorithm NHSS (*the Newton-HSS method* [*18*])

1. Given an initial guess x_0, positive constants α and tol, and a positive integer sequence $\{\ell_k\}_{k=0}^{\infty}$.

DOI: 10.1002/nla

2. For $k=0,1,\ldots$ until $\|F(x_k)\| \leqslant \text{tol}\|F(x_0)\|$ do:
   2.1. Set $d_{k,0}:=0$.
   2.2. For $\ell=0,1,2,\ldots,\ell_k-1$, apply Algorithm HSS:

$$\begin{aligned}(\alpha I+H(x_k))d_{k,\ell+1/2} &= (\alpha I-S(x_k))d_{k,\ell}-F(x_k),\\ (\alpha I+S(x_k))d_{k,\ell+1} &= (\alpha I-H(x_k))d_{k,\ell+1/2}-F(x_k),\end{aligned} \tag{7}$$

   and obtain d_{k,ℓ_k} such that

$$\|F(x_k)+F'(x_k)d_{k,\ell_k}\| < \eta_k\|F(x_k)\| \quad \text{for some } \eta_k\in[0,1), \tag{8}$$

   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

$$x_{k+1}=x_k+d_{k,\ell_k}. \tag{9}$$

In [18], η_k, for all k, is set equal to a positive constant $\tilde{\eta}$ less than 1 that is pre-specified before implementing the algorithm. However, the globally convergent algorithm in the next section is based on different η_k [21] per step.

If the last direction d_{k,ℓ_k} at the kth step is given in terms of the first direction $d_{k,0}$ in (7) (here the value is 0), we get

$$d_{k,\ell_k}=(I-T_k^{\ell_k})(I-T_k)^{-1}B_k^{-1}F(x_k), \tag{10}$$

where $T_k:=T(\alpha;x_k)$, $B_k:=B(\alpha;x_k)$ and

$$\begin{aligned}T(\alpha;x) &= B(\alpha;x)^{-1}C(\alpha;x),\\ B(\alpha;x) &= \frac{1}{2\alpha}(\alpha I+H(x))(\alpha I+S(x)),\\ C(\alpha;x) &= \frac{1}{2\alpha}(\alpha I-H(x))(\alpha I-S(x)).\end{aligned} \tag{11}$$

Then, from the expressions for T_k in (11) and d_{k,ℓ_k} in (10), the Newton-HSS iteration in (9) can be written as

$$x_{k+1}=x_k-(I-T_k^{\ell_k})F'(x_k)^{-1}F(x_k). \tag{12}$$

In order to get a Kantorovich-type semilocal convergence theorem for the above Newton-HSS method, we need the following assumption.

Assumption 3.1
Let $x_0\in\mathscr{C}^n$, and $F:\mathscr{D}\subset\mathscr{C}^n\longrightarrow\mathscr{C}^n$ be G-differentiable on an open neighbourhood $\mathscr{N}_0\subset\mathscr{D}$ on 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 the skew-Hermitian parts of the Jacobian matrix $F'(x)$, respectively. In addition, assume the following conditions hold.

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

$$\max\{\|H(x_0)\|,\|S(x_0)\|\}\leqslant\beta,\quad \|F'(x_0)^{-1}\|\leqslant\gamma,\quad \|F(x_0)\|\leqslant\delta, \tag{13}$$

DOI: 10.1002/nla

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

$$\begin{aligned}\|H(x)-H(y)\| &\leqslant L_h\|x-y\|,\\ \|S(x)-S(y)\| &\leqslant L_s\|x-y\|.\end{aligned} \tag{14}$$

From Assumption 3.1, $F'(x)=H(x)+S(x)$, $L=L_h+L_s$ and Banach's theorem (see Theorem V.4.3 in [22], or the Perturbation Lemma 2.3.2 in [4]), Lemma 3.1 easily follows.

Lemma 3.1
Under Assumption 3.1, we have

(1) $\|F'(x)-F'(y)\|\leqslant L\|x-y\|$;
(2) $\|F'(x)\|\leqslant L\|x-x_0\|+2\beta$;
(3) If $r\leqslant 1/(\gamma L)$, then $F'(x)$ is nonsingular and satisfies

$$\|F'(x)^{-1}\|\leqslant\frac{\gamma}{1-\gamma L\|x-x_0\|}.$$

Then we can give the following semilocal convergence theorem.

Theorem 3.2
Assume that Assumption 3.1 holds with the constants satisfying

$$\delta\gamma^2 L\leqslant\frac{1-\eta}{2(1+\eta^2)}, \tag{15}$$

where $\eta=\max_k\{\eta_k\}<1$, $r=\min(r_1,r_2)$ with

$$\begin{aligned} r_1 &= \frac{\alpha+\beta}{L}\left(\sqrt{1+\frac{2\alpha\tau\theta}{(2\gamma+\gamma\tau\theta)(\alpha+\beta)^2}}-1\right),\\ r_2 &= \frac{b-\sqrt{b^2-2ac}}{a},\\ a &= \frac{\gamma L(1+\eta)}{1+2\gamma^2\delta L\eta},\quad b=1-\eta,\quad c=2\gamma\delta,\end{aligned} \tag{16}$$

and with $\ell_*=\liminf_{k\to\infty}\ell_k$ satisfying[‡] $\ell_*>\lfloor\ln\eta/\ln((\tau+1)\theta)\rfloor$, $\tau\in(0,(1-\theta)/\theta)$ and

$$\theta\equiv\theta(\alpha;x_0)=\|T(\alpha;x_0)\|<1. \tag{17}$$

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

Proof
First of all, we will show the following estimate about the iterative matrix $T(\alpha;x)$ of the linear solver: if $x\in\mathscr{B}(x_0,r)$, then

$$\|T(\alpha;x)\|\leqslant(\tau+1)\theta<1. \tag{18}$$

[‡]Here, the 'floor' symbol $\lfloor\cdot\rfloor$ represents the largest integer less than or equal to the corresponding real number.

DOI: 10.1002/nla

In fact, from the definition of $B(\alpha;x)$ in (11) and Assumption (A2), we can obtain

$$\begin{aligned}
&\|B(\alpha;x)-B(\alpha;x_0)\|\\
&\leqslant\frac{1}{2}\|(H(x)-H(x_0))+(S(x)-S(x_0))\|+\tfrac{1}{2\alpha}\|H(x)S(x)-H(x_0)S(x_0)\|\\
&\leqslant\frac{L}{2}\|x-x_0\|+\tfrac{1}{2\alpha}(\|(H(x)-H(x_0))+H(x_0)\|\|S(x)-S(x_0)\|+\|H(x)-H(x_0)\|\|S(x_0)\|)\\
&\leqslant\frac{L}{2}\|x-x_0\|+\tfrac{1}{2\alpha}((L_h\|x-x_0\|+\beta)L_s\|x-x_0\|+\beta L_h\|x-x_0\|)\\
&\leqslant\frac{L}{2}\|x-x_0\|+\tfrac{1}{2\alpha}\left(\frac{L^2}{2}\|x-x_0\|^2+\beta L\|x-x_0\|\right)\\
&=\frac{L^2}{4\alpha}\|x-x_0\|^2+\frac{(\alpha+\beta)L}{2\alpha}\|x-x_0\|. \qquad (19)
\end{aligned}$$

Similarly, we have

$$\|C(\alpha;x)-C(\alpha;x_0)\|\leqslant\frac{L^2}{4\alpha}\|x-x_0\|^2+\frac{(\alpha+\beta)L}{2\alpha}\|x-x_0\|. \qquad (20)$$

Then because $F'(x)=B(\alpha;x)-C(\alpha;x)$ and the definition of $T(\alpha;x)$ in (11), it follows that

$$\begin{aligned}
\|B(\alpha;x_0)^{-1}\|&=\|(I-T(\alpha;x_0))F'(x_0)^{-1}\|\\
&<(1+\theta)\gamma\\
&<2\gamma. \qquad (21)
\end{aligned}$$

Meanwhile, $r\leqslant r_1$ implies that

$$L^2r^2+2(\alpha+\beta)Lr<\frac{2\alpha\tau\theta}{2\gamma+\gamma\tau\theta}<\frac{2\alpha}{\gamma}. \qquad (22)$$

Hence, again using the Banach theorem, we get

$$\begin{aligned}
\|B(\alpha;x)^{-1}\|&\leqslant\frac{\|B(\alpha;x_0)^{-1}\|}{1-\|B(\alpha;x_0)^{-1}\|\|B(\alpha;x)-B(\alpha;x_0)\|}\\
&\leqslant\frac{8\alpha\gamma}{4\alpha-2\gamma(L^2\|x-x_0\|^2+2(\alpha+\beta)L\|x-x_0\|)}. \qquad (23)
\end{aligned}$$

Hence, as in Theorem 3.2 of [18], this together with (17), (19), (20) and (22), the estimate about the inner iteration matrix $T(\alpha;x)$ and $T(\alpha;x_0)$ is obtained as follows:

$$\begin{aligned}
&\|T(\alpha;x)-T(\alpha;x_0)\|\\
&\quad=\|B(\alpha;x)^{-1}(C(\alpha;x)-C(\alpha;x_0))-B(\alpha;x)^{-1}(B(\alpha;x)-B(\alpha;x_0))B(\alpha;x_0)^{-1}C(\alpha;x_0)\|
\end{aligned}$$

DOI: 10.1002/nla

$$\leqslant \frac{4\gamma(L^2\|x-x_0\|^2+2(\alpha+\beta)L\|x-x_0\|)}{4\alpha-2\gamma(L^2\|x-x_0\|^2+2(\alpha+\beta)L\|x-x_0\|)}$$
$$<\tau\theta. \tag{24}$$

Consequently,

$$\|T(\alpha;x)\|\leqslant\|T(\alpha;x)-T(\alpha;x_0)\|+\|T(\alpha;x_0)\|\leqslant(\tau+1)\theta<1. \tag{25}$$

Second, we claim that the following iterative sequence $\{t_k\}$ converges monotone increasingly to r_2:

$$\begin{aligned} t_0 &= 0, \\ t_{k+1} &= t_k-\frac{g(t_k)}{h(t_k)}, \quad k=0,1,2,\ldots, \end{aligned} \tag{26}$$

where

$$\begin{aligned} g(t) &= \tfrac{1}{2}at^2-bt+c, \\ h(t) &= at-1. \end{aligned} \tag{27}$$

In short, the following inequalities hold:

$$t_k<t_{k+1}<r_2 \quad \text{for } k=0,1,\ldots. \tag{28}$$

In fact, from inequality (15) we have

$$\delta\gamma^2 L\leqslant\frac{(1+2\gamma^2\delta L\eta)(1-\eta)}{2(1+\eta)},$$

that is,

$$c=t_1<\frac{b}{a}.$$

Based on $g(t_1)=\frac{1}{2}ac^2+\eta c>0=g(r_2)$, and $g(t)$ decreasing in $[0,b/a]$, it holds that

$$t_1<r_2.$$

Hence (28) is true for $k=0$.

Suppose that $t_{k-1}<t_k<r_2$. Then we consider

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

$g(t)$ decreasing and $g'(t)$ increasing in $[0,b/a]$ imply that $g(t_k)>0$ and $g'(t_k)\leqslant 0$, respectively. Furthermore, $h(t_k)\leqslant g'(t_k)$ implies $-h(t_k)\geqslant 0$. Hence,

$$t_{k+1}>t_k.$$

On the other hand, the function

$$t-\frac{g(t)}{g'(t)}$$

DOI: 10.1002/nla

increases in $[0, b/a]$. Combining with

$$-\frac{g(t_k)}{h(t_k)} \leqslant -\frac{g(t_k)}{g'(t_k)},$$

we have

$$t_{k+1} \leqslant r_2 - \frac{g(r_2)}{g'(r_2)} = r_2.$$

Hence, (28) is also true for k. Consequently, the claim (28) holds for all nonnegative integers. Furthermore, there exists t_* such that $\lim_k t_k = t_*$. Then we can assert $t_* = r_2$ [23].

Finally, we prove by induction

$$\begin{aligned} \|x_{k+1} - x_k\| &\leqslant t_{k+1} - t_k, \\ \|x_{k+1} - x_0\| &\leqslant t_{k+1} - t_0 \leqslant r_2, \\ \|F(x_k)\| &\leqslant \frac{1 - \gamma L t_k}{\gamma(1+\eta)}(t_{k+1} - t_k) \quad \text{for } k = 0, 1, \ldots. \end{aligned} \tag{29}$$

Since

$$\begin{aligned} \|x_1 - x_0\| &= \|F'(x_0)^{-1} F(x_0) + T_0^{\ell_*} F'(x_0)^{-1} F(x_0)\| \\ &\leqslant \gamma(1+\theta^{\ell_*})\delta \\ &\leqslant 2\gamma\delta \\ &= t_1 - t_0, \end{aligned}$$

and

$$\|F(x_0)\| \leqslant \delta \leqslant \frac{2\delta}{1+\eta} = \frac{1-\gamma L t_0}{\gamma(1+\eta)}(t_1 - t_0),$$

(29) is correct for $k = 0$. Suppose that (29) holds for all nonnegative integers less than k. We need to prove that it holds for k.

It follows from the integral mean-value theorem and Lemma 3.1 that when $x, y \in \mathscr{B}(x_0, r)$,

$$\begin{aligned} \|F(x) - F(y) - F'(y)(x-y)\| &= \left\| \int_0^1 F'(y + t(x-y))(x-y)\mathrm{d}t - F'(y)(x-y) \right\| \\ &\leqslant \int_0^1 \|F'(y + t(x-y)) - F'(y)\| \|x-y\| \,\mathrm{d}t \\ &\leqslant \int_0^1 Lt\|x-y\|^2 \mathrm{d}t \\ &= \frac{L}{2}\|x-y\|^2. \end{aligned} \tag{30}$$

DOI: 10.1002/nla

Therefore, because of (8) and $x_{k-1}, x_k \in \mathscr{B}(x_0, r_2)$,

$$\begin{aligned}\|F(x_k)\| &\leqslant \|F(x_k)-F(x_{k-1})-F'(x_{k-1})(x_k-x_{k-1})\|+\|F(x_{k-1})+F'(x_{k-1})(x_k-x_{k-1})\|\\ &\leqslant \frac{L}{2}\|x_k-x_{k-1}\|^2+\eta\|F(x_{k-1})\|,\end{aligned}$$

then using the inductive hypothesis, we have

$$\begin{aligned}\frac{\gamma(1+\eta)}{1-\gamma L t_k}\|F(x_k)\| &\leqslant \frac{\gamma(1+\eta)}{1-\gamma L t_k}\left(\frac{L}{2}(t_k-t_{k-1})^2+\eta\frac{1-\gamma L t_{k-1}}{\gamma(1+\eta)}(t_k-t_{k-1})\right)\\ &\leqslant \frac{1}{2}\frac{(1+\eta)\gamma L}{1-\gamma L t_k}(t_k-t_{k-1})^2+\frac{\eta}{1-\gamma L t_k}(t_k-t_{k-1})\\ &\leqslant \frac{1}{2}\frac{a}{-h(t_k)}(t_k-t_{k-1})^2+\frac{\eta}{-h(t_k)}(t_k-t_{k-1}).\end{aligned}$$

The above last inequality is true because (15) implies $\delta \leqslant 1/(2\gamma^2 L)$; therefore, $1-\gamma L t_k \geqslant -h(t_k)$, whereas $t_k > t_1 = 2\gamma\delta$ implies $(1+\eta)\gamma L/(1-\gamma L t_k) < a/(-h(t_k))$.

Hence, from

$$g(t_k)-g(t_{k-1})-h(t_k)(t_k-t_{k-1}) = \tfrac{1}{2}a(t_k-t_{k-1})^2+\eta(t_k-t_{k-1}),$$

we obtain

$$\begin{aligned}\frac{\gamma(1+\eta)}{1-\gamma L t_k}\|F(x_k)\| &\leqslant \frac{1}{-h(t_k)}(g(t_k)-g(t_{k-1})-h(t_k)(t_k-t_{k-1}))\\ &= t_{k+1}-t_k. \qquad (31)\end{aligned}$$

Consequently, from the iterative formula (12), (25) and Lemma 3.1, we get

$$\begin{aligned}\|x_{k+1}-x_k\| &\leqslant \|(I-T_k^{\ell_k})F'(x_k)^{-1}F(x_k)\|\\ &\leqslant (1+((\tau+1)\theta)^{\ell_*})\frac{\gamma}{1-\gamma L t_k}\|F(x_k)\|.\end{aligned}$$

Then, based on the choice of ℓ_*, it follows that

$$\|x_{k+1}-x_k\| \leqslant (1+\eta)\frac{\gamma}{1-\gamma L t_k}\|F(x_k)\|,$$

so that, from (31), the first inequality in (29) is also correct for k. The second one in (29) is easy to get from

$$\begin{aligned}\|x_{k+1}-x_0\| &\leqslant \|x_{k+1}-x_k\|+\|x_k-x_{k-1}\|+\cdots+\|x_1-x_0\|\\ &\leqslant t_{k+1}-t_k+t_k-t_{k-1}+\cdots+t_1-t_0\\ &\leqslant t_{k+1}-t_0\\ &\leqslant r_2.\end{aligned}$$

DOI: 10.1002/nla

Hence, (29) is true for all k. Since the sequence $\{t_k\}$ converges to t_*, the sequence $\{x_k\}$ also converges, to say x_*. Because $\|T(\alpha; x_*)\| < 1$ [3], from the iteration (12), we have

$$F(x_*) = 0.$$

Therefore, the conclusion of this theorem follows. □

Note: The semilocal convergence of inexact Newton methods was proved in [17] under the following assumptions:

$$\begin{aligned} \|F'(x_0)^{-1}F(x_0)\| &\leqslant \beta, \\ \|F'(x_0)^{-1}(F'(x) - F'(y))\| &\leqslant \gamma\|x-y\|, \\ \frac{\|F'(x_0)^{-1}s_n\|}{\|F'(x_0)^{-1}F(x_n)\|} &\leqslant \eta_n, \end{aligned} \tag{32}$$

and

$$\beta\gamma \leqslant g_1(\eta), \tag{33}$$

where

$$g_1(\eta) = \frac{\sqrt{(4\eta+5)^3} - (2\eta^2 + 14\eta + 11)}{(1+\eta)(1-\eta)^2}. \tag{34}$$

Later, Shen and Li [23] substituted $g_1(\eta)$ with $g_2(\eta)$, where

$$g_2(\eta) = \frac{(1-\eta)^2}{(1+\eta)(2(1+\eta) - \eta(1-\eta)^2}. \tag{35}$$

These could be used to obtain convergence results for the Newton-HSS method as a subclass of these techniques. However, our detailed proof in Theorem 3.2 is targeted specifically at the Newton-HSS method and so gives better bounds for the semilocal convergence result, see Figure 1. The corresponding bound in our theorem is

$$g_3(\eta) = \frac{1-\eta}{2(1+\eta^2)}. \tag{36}$$

An unconditional convergence theorem of the HSS iteration (see Theorem 2.2 in [3]) shows that the Euclidean norm of T satisfies

$$\|T(\alpha; x)\| \leqslant \max_{\lambda \in \sigma(H)} \frac{\alpha - \lambda}{\alpha + \lambda} < 1,$$

where $\lambda(\cdot)$ represents the spectrum of the corresponding matrix. Hence, Equation (17) holds and is not a part of the assumption. It serves only to define the scalar parameter θ.

4. GLOBAL CONVERGENCE OF THE NEWTON-HSS METHOD

In the previous section, we have answered an important question [4]. From a specific initial approximation x_0, the existence of solutions can be ascertained directly from the iterative process.

DOI: 10.1002/nla

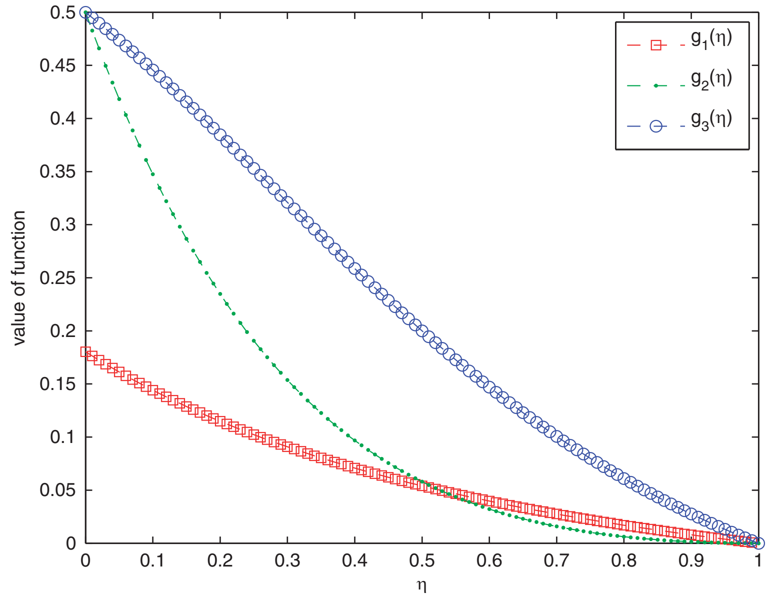

Figure 1. Graphs of $g_1(\eta)$, $g_2(\eta)$ and $g_3(\eta)$.

This point x_0 needs to satisfy some conditions. These may (at least in principle) be used to find solutions. In fact, in [24] we have utilized such criteria on x_0 to solve an integral equation. We now look at a stronger type of convergence. A globally convergent algorithm for solving (1) means an algorithm with the property that, for any initial iterate, the iteration either converges to a root of F or fails to do so in one of the small number of ways [6]. There are three main ways that such algorithms can be globalized: linear search methods, trust region methods and continuation/homotopy methods. Various inexact Newton methods with such global strategies are at present widely considered to be among the best approaches for solving nonlinear systems of equations, especially globally convergent Newton-GMRES subspace methods [6, 10, 25–28].

A general framework can be obtained by augmenting the inexact Newton condition with a sufficient decrease condition on $\|F\|$ [25].

Algorithm GIN (*global inexact Newton method* [25])
Let x_0 and $t \in (0,1)$ be given.

1. For $k=0$ step 1 until ∞ do:

   1.1. For a given $\eta_k \in [0,1)$, find an s_k that satisfies

   $$\|F(x_k)+F'(x_k)s_k\| \leqslant \eta_k \|F(x_k)\|$$

   and

   $$\|F(x_k+s_k)\| \leqslant [1-t(1-\eta_k)]\|F(x_k)\|.$$

   1.2. Set $x_{k+1}=x_k+s_k$.

Eisenstat and Walker [25] give a thorough demonstration that trust region methods are in some sense dual to line search methods and the inexact Newton method with practically implemented Goldstein–Armijo conditions can be regarded as a special case of Algorithm GIN.

Hence, in this section, we select a backtracking linear search paradigm to implement the Newton-HSS method. Eisenstat and Walker [25] offer the following inexact Newton backtracking method containing strong global convergence properties combined with potentially fast local convergence:

Algorithm INB (*inexact Newton backtracking method* [*21, 25*])
Let x_0, $\eta_{\max} \in [0, 1)$, $t \in (0, 1)$, and $0 < \theta_{\min} < \theta_{\max} < 1$ be given.

1. For $k = 0$ step 1 until ∞ do:
   1.1. Choose $\eta_k \in [0, \eta_{\max})$ and d_k such that
   $$\|F(x_k) + F'(x_k)d_k\| \leqslant \eta_k \|F(x_k)\|.$$
   1.2. Set $\overline{d}_k = d_k$ and $\overline{\eta}_k = \eta_k$.
   1.3. While $\|F(x_k + \overline{d}_k)\| > [1 - t(1 - \overline{\eta}_k)]\|F(x_k)\|$ do:
      1.3.1 Choose $\theta \in [\theta_{\min}, \theta_{\max}]$.
      1.3.2 Update $\overline{d}_k \longleftarrow \theta \overline{d}_k$ and $\overline{\eta}_k \longleftarrow 1 - \theta(1 - \overline{\eta}_k)$.
   1.4. Set $x_{k+1} = x_k + \overline{d}_k$.

Eisenstat and Walker [21] have examined in great detail the choice of the parameter η_k (the so-called forcing term). This is used to reduce the effort required to obtain too accurate a solution of the Newton equation.

Choice 1: For $\overline{\rho} = (1 + \sqrt{5})/2$ and any $\eta_0 \in [0, 1)$, choose

$$\eta_k = \frac{|\|F(x_k)\| - \|F(x_{k-1}) + F'(x_{k-1})d_{k-1}\||}{\|F(x_{k-1})\|}, \quad k = 1, 2, \ldots,$$

with
Choice 1 safeguard: $\eta_k = \max\{\eta_k, \eta_{k-1}^{\overline{\rho}}\}$ wherever $\eta_{k-1}^{\overline{\rho}} > 0.1$.

Choice 2: Given $\lambda \in [0, 1]$ and $\rho \in (1, 2]$, select any $\eta_0 \in [0, 1)$ and choose

$$\eta_k = \lambda \left(\frac{\|F(x_k)\|}{\|F(x_{k-1})\|} \right)^{\rho}, \quad k = 1, 2, \ldots,$$

with
Choice 2 safeguard: $\eta_k = \max\{\eta_k, \lambda\eta_{k-1}^{\rho}\}$ wherever $\lambda\eta_{k-1}^{\rho} > 0.1$.

Then in both cases after applying the above two safeguards, it is necessary for Algorithm INB to use another additional safeguard [26]:

$$\eta_k \longleftarrow \min\{\eta_k, \eta_{\max}\}. \tag{37}$$

When $\eta_k \|F(x_k)\|$ is small enough, a larger η_k can actually be used for the outer iteration. Hence, Pernice and Walker [26] impose the following final safeguard:

$$\eta_k \longleftarrow 0.8\varepsilon / \|F(x_k)\| \quad \text{wherever } \eta_k \leqslant 2\varepsilon / \|F(x_k)\|. \tag{38}$$

Now we give the Newton-HSS method with backtracking as follows.

DOI: 10.1002/nla

Algorithm NHSSB (*the Newton-HSS method with backtracking*).
Let x_0, $\eta_{\max} \in [0, 1)$, $t \in (0, 1)$, $0 < \theta_{\min} < \theta_{\max} < 1$, tol$>0$ be given.

1. While $\|F(x_k)\| > \text{tol} \min\{\|F(x_0)\|, \sqrt{n}\}$ and $k < 1000$ do:
   1.1. Choose $\eta_k \in [0, \eta_{\max}]$, apply Algorithm HSS to the kth Newton equation to obtain d_k such that

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

   1.2. Perform the **Backtracking Loop (BL)**, i.e.
      1.2.1. Set $\overline{d}_k = d_k, \overline{\eta}_k = \eta_k$.
      1.2.2. While $\|F(x_k + \overline{d}_k)\| > [1 - t(1 - \overline{\eta}_k)]\|F(x_k)\|$do:
         1.2.2.1. Choose $\theta \in [\theta_{\min}, \theta_{\max}]$.
         1.2.2.2. Update $\overline{d}_k = \theta \overline{d}_k$ and $\overline{\eta}_k = 1 - \theta(1 - \overline{\eta}_k)$.
   1.3. Set $x_{k+1} = x_k + \overline{d}_k$.

Assumption 3.1 guarantees that Lemma 3.1 holds. That means that F' is Lipschitz continuous with Lipschitz constant L, and there exists a positive constant m_f such that $\|F'(x)^{-1}\| \leqslant m_f$ on the set

$$\Omega(\{x_n\}, r) = \bigcup_{n=0}^{\infty} \{x \mid \|x - x_n\| \leqslant r\}.$$

Hence, we have the following two global convergence theorems for NHSSB by Theorem 8.2.1 of Kelley [6], Theorems 2.2 and 2.3 of Pernice and Walker [26].

Theorem 4.1
Let $x_0 \in \mathscr{C}^n$ and $t \in (0, 1)$ be given. Assume that $\{x_k\}$ is given by Algorithm NHSSB, in which each η_k is given by Choice 1 followed by all 'the safeguards'. Furthermore, suppose that $\{x_k\}$ is bounded, and Assumption 3.1 holds. Then $\{x_k\}$ converges to a root $\{x_*\}$ of F. Moreover

(1) if $\eta_k \longrightarrow 0$, the convergence is q-superlinear, and
(2) if $\eta_k \leqslant \mathscr{K}_\eta \|F(x_k)\|^p$ for some $\mathscr{K}_\eta > 0$, the convergence is q-superlinear with q-order $1 + p$.

Theorem 4.2
Let $x_0 \in \mathscr{C}^n$ and $t \in (0, 1)$ be given. Assume that $\{x_k\}$ is given by Algorithm NHSSB, in which each η_k is given by Choice 2 followed by all 'the safeguards'. Furthermore, suppose that $\{x_k\}$ is bounded, and Assumption 3.1 holds. Then $\{x_k\}$ converges to a root $\{x_*\}$ of F. Moreover,

(1) if $\lambda < 1$, the convergence is of q-order ρ, and
(2) if $\lambda = 1$, the convergence is of r-order ρ and of q-order p for every $p \in [1, \rho)$.

Remark
In fact, it is not necessary here to use the first bound condition (A1) in Assumption 3.1 for proving these two theorems.

5. NUMERICAL TESTS

In this section, we illustrate Algorithm NHSSB with five kinds of forcing terms on nonlinear convection–diffusion equations. Since a comparison between the Newton-HSS method and other methods such as Newton-GMRES, Newton-USOR and Newton-GCG were shown in detail in [18], in this paper we just demonstrate the effectiveness of the Newton-HSS method with backtracking and the effect of the forcing terms. Also, comparison between the Newton-HSS method and the Newton-GMRES method with backtracking and forcing terms is shown.

We consider the two-dimensional nonlinear convection–diffusion equation

$$\begin{aligned} -(u_{xx}+u_{yy})+q(u_x+u_y) &= -\mathrm{e}^u \quad (x,y)\in\Omega, \\ u(x,y) &= 0 \quad (x,y)\in\partial\Omega, \end{aligned} \tag{39}$$

where $\Omega=(0,1)\times(0,1)$, with $\partial\Omega$ its boundary, and q is a positive constant used to control the magnitude of the convective terms [1, 2, 18]. Now we apply the five-point finite-difference scheme to the diffusive terms and the central difference scheme to the convective terms, respectively. Let $h=1/(N+1)$ and $Re=qh/2$ denote the equidistant step-size and the mesh Reynolds number, respectively. Then we get the system of nonlinear equations of the form

$$\overline{A}u+h^2\mathrm{e}^u=0,$$

$$u=(u_1,u_2,\ldots,u_N)^\top,\quad u_i=(u_{i1},u_{i2},\ldots,u_{iN})^\top,\quad i=1,2,\ldots,N,$$

where the coefficient matrix of the linear term is

$$\overline{A}=T_x\otimes I+I\otimes T_y.$$

Here, $\otimes$ means the Kronecker product, and T_x and T_y are the tridiagonal matrices

$$T_x=\mathrm{tridiag}(t_2,t_1,t_3),\quad T_y=\mathrm{tridiag}(t_2,0,t_3),$$

with

$$t_1=4,\quad t_2=-1-Re,\quad t_3=-1+Re.$$

Discretization is on a 100×100 uniform grid, so that the dimension $n=10\,000$. In the implementations of Algorithm NHSSB, $\alpha=qh/2$ is adopted [3]. Five types of forcing terms in the figures are represented as follows:

Choice 1 denotes Choice 1 with Choice 1 safeguard, additional safeguard and final safeguard;
Choice 2 denotes Choice 2 with Choice 2 safeguard, additional safeguard and final safeguard;
Choice 3 $\eta_k=0.1$, for all k;
Choice 4 $\eta_k=0.0001$, for all k;
Choice 5 $\eta_k=|\|F(x_k)\|-\|F(x_{k-1})+F'(x_{k-1})d_{k-1}\||/\|F(x_k)\|,\quad k=1,2,\ldots,$ with Choice 1 safeguard, additional safeguard and final safeguard.

Note: Choice 5 is a new option that we introduce that has a different denominator from Choice 1. Since the term $\|F(x_k)\|$ is closer to the current iterate than $\|F(x_{k-1})\|$, it is reasonable

DOI: 10.1002/nla

to adopt this value in our forcing term. From the proof of the local convergence order of Choice 1 (see Theorem 2.2 of [21]), the corresponding estimate can easily be obtained as

$$\|x_{k+1}-x_*\| \leqslant \beta(\|x_k-x_*\|^2+\|x_{k-1}-x_*\|^2).$$

It looks a bit weaker than q-superlinear and two-step q-quadratic convergence, but it is effective in numerical tests (see Figures 4 and 7).

We follow [21] in our choice of values for the other parameters. We set $\eta_0=0.5$ in Choices 1, 2 and 5 and use $\eta_{\max}=0.9$ for safeguards in Choices 1, 2 and 5. As for the backtracking search, we choose $t=10^{-4}$, $\theta_{\min}=\frac{1}{10}$, $\theta_{\max}=\frac{1}{2}$. We select θ in Step 1.2.2.1 of Algorithm NHSSB such that the merit function $g(\theta)\equiv\|F(x_k+\theta d_k)\|^2$ is minimized over $[\theta_{\min},\theta_{\max}]$. $\lambda=1$ and $\rho=(1+\sqrt{5})/2$ are chosen in Choice 2 as they have been seen to be most effective by [21].

We give the results in the following six figures. The horizontal axis indicates the total number of inner iterations (denoted as 'IT'), the total number of outer iterations (denoted as 'OT') and the total CPU time (denoted as 't'), respectively. The corresponding vertical axis is $\log\|F(x_k)\|$. In every figure, results for two values of q (200 and 600) are shown. We let $\mathbf{e}$ be the vector of all 1s. We use $x_0=\mathbf{e}$ in the first three figures and $x_0=16\mathbf{e}$ in the last three figures. The reason for choosing these two points is that the solution is near 0 and any points over $16\mathbf{e}$ can cause problems with the convergence for some choices of parameter values resulting in a large increase in run time. For example, when we let $x_0=17\mathbf{e}$ and $x_0=18\mathbf{e}$, we require more than our limit of 1000 inner iterations to get the direction d_k for the iteration with Choice 4.

One sees from Figure 2 that fewer inner iterations are needed in the case of Choice 5 than in the other cases. The poorest is Choice 4. Choices 1, 2 and 3 are almost the same. But a different situation is found with the number of outer iterations. That is, the least number of outer iterations are performed for Choice 4, then the order is Choices 1, 5, 2 and 3 (see Figure 3). This is not that surprising as, if the inner equation is solved more accurately (hence more inner iterations), one might expect that fewer outer iterations would be needed. Hence, because the initial point is close to the solution, it is not very astonishing that the CPU times have the same behaviour as the number of iterations in Figure 3. This is shown in Figure 4.

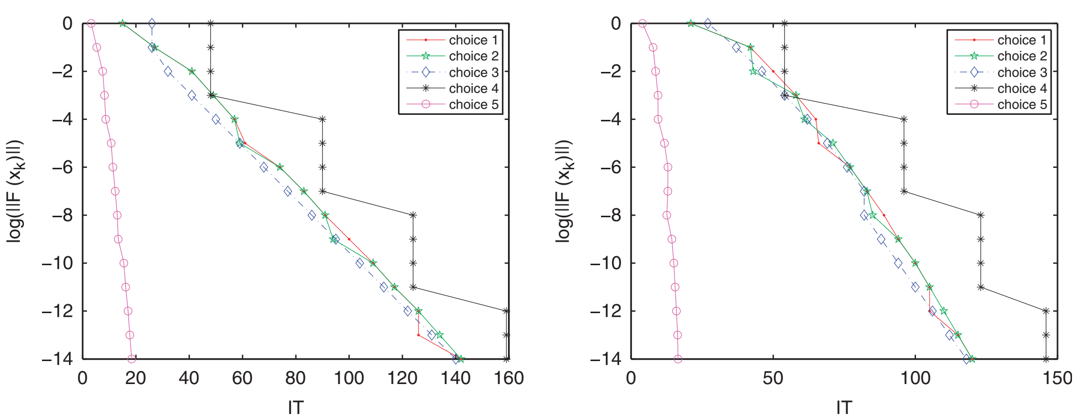

Figure 2. The total number of inner iterations versus the norms of the nonlinear function values when $q=200$ and $q=600$, respectively, with $x_0=\mathbf{e}$.

DOI: 10.1002/nla

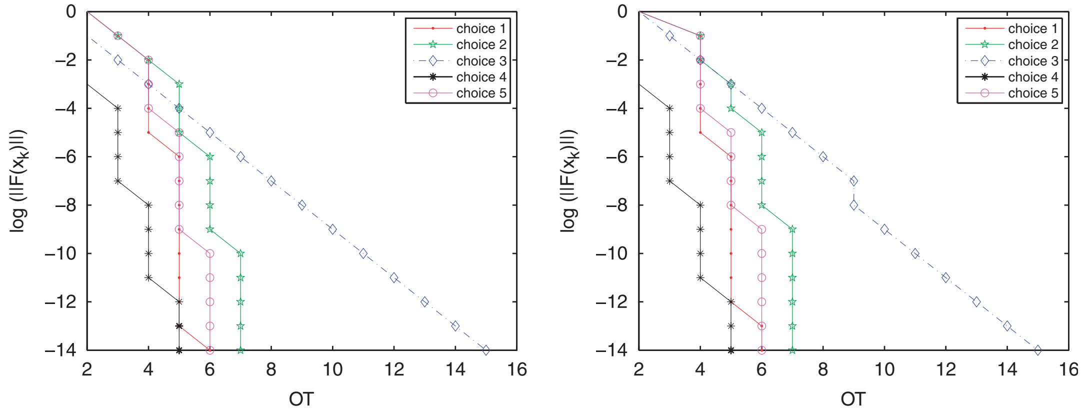

Figure 3. The number of outer iterations versus the norms of the nonlinear function values when $q=200$ and $q=600$, respectively, with $x_0=\mathbf{e}$.

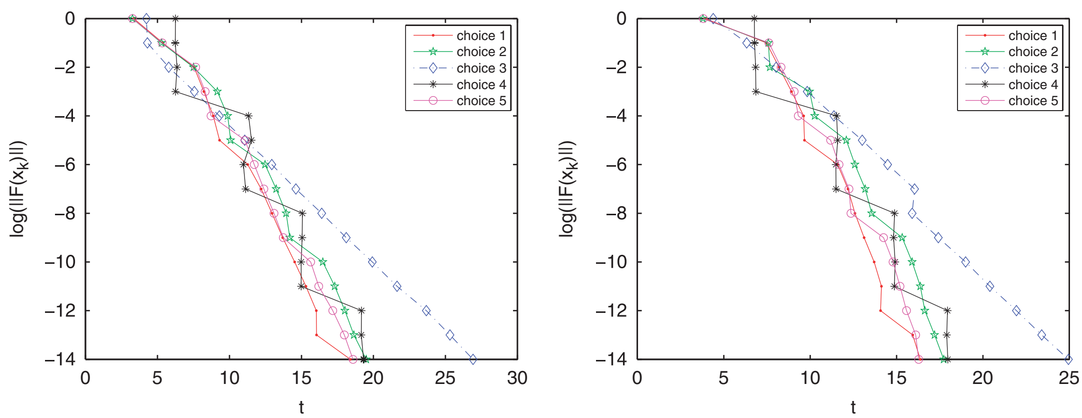

Figure 4. CPU times versus the norms of the nonlinear function values when $q=200$ and $q=600$, respectively, with $x_0=\mathbf{e}$.

When the starting point is far from the solution (for example, $x_0=16\mathbf{e}$ in our tests), more inner iterations are needed than in the case of $x_0=\mathbf{e}$. The number of inner iterations is still the main influence on the CPU time, see Figures 5, 6 and 7. Choice 5 is still a good choice.

Though Newton-HSS and Newton-GMRES have been compared in [18], the comparison using a globalization strategy and choosing forcing terms dynamically was not carried out. Such a comparison is given in Figures 8 and 9, using GMRES(20). One can notice that Newton-GMRES needs more inner iterations because some inner GMRES iterations still do not converge after the maximum number of inner iteration steps (here, it is 1000). Tests with other different forcing terms give similar results to Figures 8 and 9.

Note: The Newton-HSS method with a back-tracking global strategy and choices of forcing terms performs better than the Newton-GMRES method with same global strategy and choices,

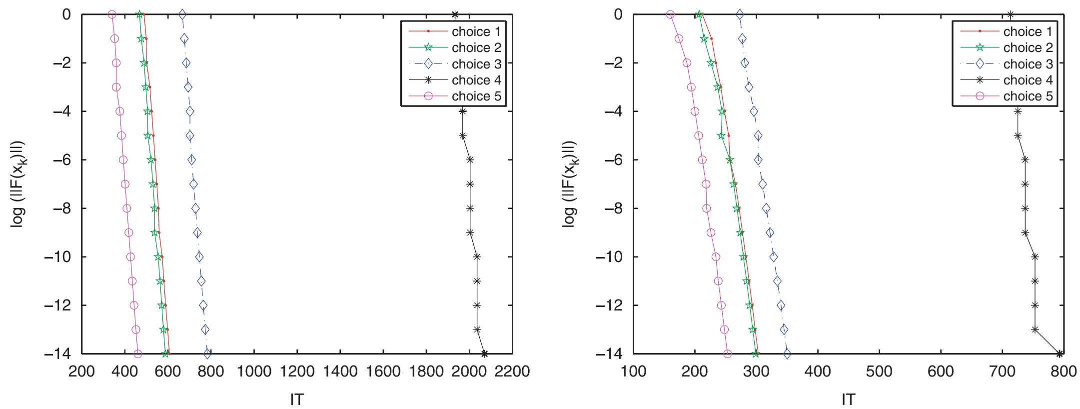

Figure 5. The total number of inner iterations versus the norms of the nonlinear function values when $q=200$ and $q=600$, respectively, with $x_0=16\mathbf{e}$.

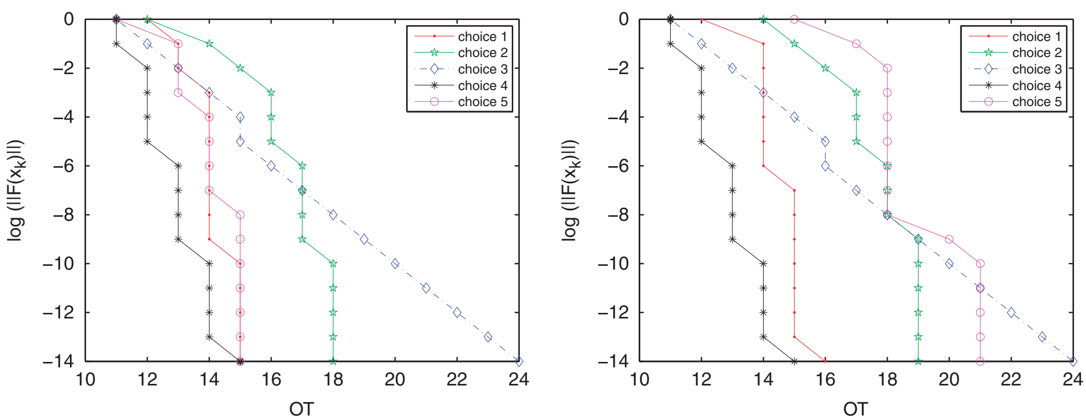

Figure 6. The number of outer iterations versus the norms of the nonlinear function values when $q=200$ and $q=600$, respectively, with $x_0=16\mathbf{e}$.

especially on the problem (39) and for big Reynolds numbers. But we are restricted in our choice of test problems because the Jacobian matrix $F'(x)$ needs to be positive definite. On the other hand, although Newton-HSS cannot be implemented in the Jacobian-free way as Newton-GMRES, it is less important for Newton-HSS. GMRES needs to compute and store $r, Ar, A^2r, \ldots$ (r is the residual), but HSS is a stationary iterative and so avoids this. For solving systems of nonlinear equations (1), one can use

$$F'(x)d \approx \frac{F(x+\varepsilon d)-F(x)}{\varepsilon}, \tag{40}$$

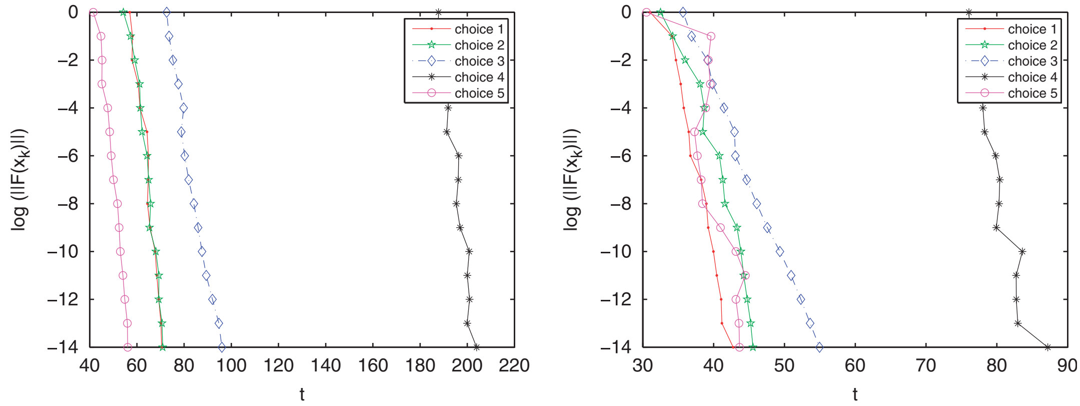

Figure 7. CPU times versus the norms of the nonlinear function values when $q=200$ and $q=600$, respectively, with $x_0=16\mathbf{e}$.

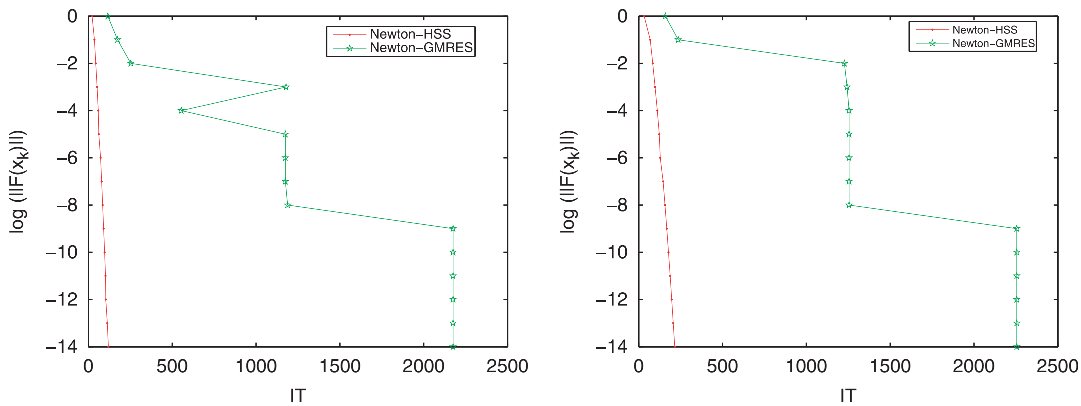

Figure 8. The total number of inner iterations versus the norms of the nonlinear function values when $q=600$ and $q=2000$, respectively, with $x_0=\mathbf{e}$ and Choice 1 for the Newton-HSS method and the Newton-GMRES method.

where ε is a small perturbation (say, see [7]), to carry out Jacobian-free Newton-GMRES. While, in Newton-HSS, $(\nabla_h F)(x)$ [6] can be introduced, where

$$(\nabla_h F)(x)_j = \begin{cases} \dfrac{F(x+h\|x\|e_j)-F(x)}{h\|x\|}, & x\neq 0, \\ \dfrac{F(he_j)-F(x)}{h}, & x=0, \end{cases} \tag{41}$$

to implement derivative-free Newton-HSS. Furthermore, it is very easy to compute Jacobian matrices for some problems such as (39).

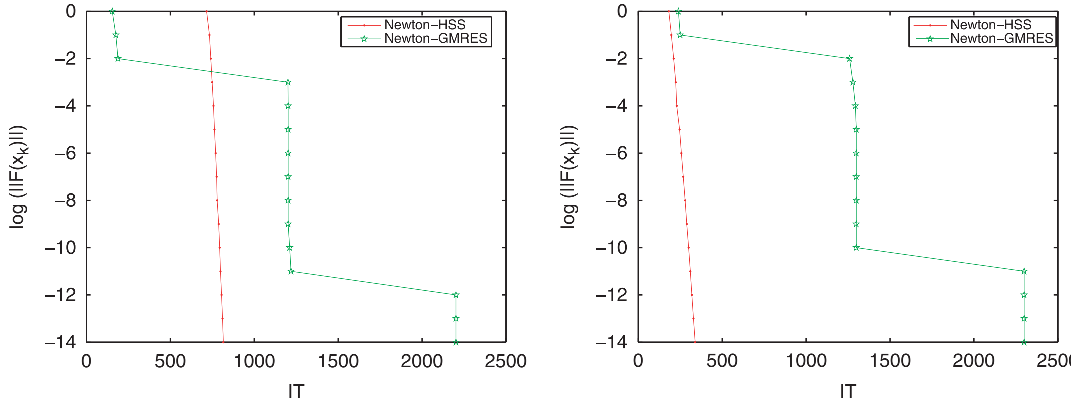

Figure 9. The total number of inner iterations versus the norms of the nonlinear function values when $q=600$ and $q=2000$, respectively, with $x_0=16\mathbf{e}$ and Choice 1 for the Newton-HSS method and the Newton-GMRES method.

6. CONCLUSIONS

We have proved the semilocal convergence for the Newton-HSS method, which ensures that the sequence $\{x_k\}$ produced by Algorithm NHSS converges to the solution of the system of nonlinear Equations (1) under some reasonable assumptions. This means that in principle any point can be tested to be an effective initial point or not by checking this semilocal convergence theorem. But in order to consider the convergence of iterates starting from an arbitrary point, we present Algorithm NHSSB combining the Newton-HSS method with a backtracking strategy and prove global convergence with two typical forcing terms. Finally, numerical tests are shown on convection–diffusion equations. We compare five choices for the stopping criteria of the inner iterations. Among them, from the results in this paper, Choice 5 needs the least run time.

ACKNOWLEDGEMENTS

We are indebted to Prof. Zhong-Zhi Bai and Prof. Serge Gratton for their helpful suggestions related to this paper, as well as to Doctor Azzam Haidar for his help with Matlab codes. Also, we gratefully acknowledge the many excellent suggestions from the anonymous referees and thank Prof. Chong-Li for sending us his paper. The first author appreciated very much Professor Iain Duff and other parallel algo colleagues for their pleasant reception and warm help when visiting CERFACS during the 2008–2009 year.

The work of the first author was supported in part by the NSFC grant 10971070. The work of the second author was supported EPSRC Grant EP/E053351/1.

REFERENCES

1. Axelsson O, Carey GF. On the numerical solution of two-point singularly perturbed boundary value problems. *Computer Methods in Applied Mechanics and Engineering* 1985; **50**:217–229.
2. Axelsson O, Nikolova M. Avoiding slave points in an adaptive refinement procedure for convection–diffusion problems in 2D. *Computing* 1998; **61**:331–357.
3. Bai Z-Z, Golub GH, Ng MK. Hermitian and skew-Hermitian splitting methods for non-Hermitian positive definite linear systems. *SIAM Journal on Matrix Analysis and Applications* 2003; **24**:603–626.

4. Ortega JM, Rheinboldt WC. *Iterative Solution of Nonlinear Equations in Several Variables*. Academic Press: New York, 1970.
5. Rheinboldt WC. *Methods of Solving Systems of Nonlinear Equations* (2nd edn). SIAM: Philadelphia, 1998.
6. Kelley CT. *Iterative Methods for Linear and Nonlinear Equations*. SIAM: Philadelphia, 1995.
7. Brown PN, Saad Y. Hybrid Krylov methods for nonlinear systems of equations. *SIAM Journal on Scientific and Statistical Computing* 1990; **11**:450–481.
8. Brown PN, Saad Y. Convergence theory of nonlinear Newton–Krylov algorithms. *SIAM Journal on Optimization* 1994; **4**:297–330.
9. Bellavia S, Macconi M, Morini B. In *A Hybrid Newton-GMRES Method for Solving Nonlinear Equations*, Vulkov L, Waśniewski J, Yalamov P (eds). Lecture Notes in Computer Science, vol. 1988. Springer: Berlin, Heidelberg, 2001; 68–75.
10. Bellavia S, Morini B. A globally convergent Newton-GMRES subspace method for systems of nonlinear equations. *SIAM Journal on Scientific Computing* 2001; **23**:940–960.
11. Sherman AH. On Newton-iterative methods for the solution of systems of nonlinear equations. *SIAM Journal on Numerical Analysis* 1978; **15**:755–771.
12. Ypma TJ. Convergence of Newton-like-iterative methods. *Numerische Mathematik* 1984; **45**:241–251.
13. Bai Z-Z. A class of two-stage iterative methods for systems of weakly nonlinear equations. *Numerical Algorithms* 1997; **14**:295–319.
14. Bai Z-Z, Migallón V, Penadés J, Szyld DB. Block and asynchronous two-stage methods for mildly nonlinear systems. *Numerische Mathematik* 1999; **82**:1–20.
15. Argyros IK. Local convergence of inexact Newton-like-iterative methods and applications. *Computers and Mathematics with Applications* 2000; **39**:69–75.
16. Bai Z-Z, Zhang S-L. A regularized conjugate gradient method for symmetric positive definite system of linear equations. *Journal of Computational Mathematics* 2002; **20**:437–448.
17. Guo X-P. On semilocal convergence of inexact Newton methods. *Journal of Computational Mathematics* 2007; **25**:231–242.
18. Bai Z-Z, Guo X-P. The Newton-HSS methods for systems of nonlinear equations with positive-definite Jacobian matrices. *Journal of Computational Mathematics* 2010; **28**:235–260.
19. Dembo RS, Eisenstat SC, Steihaug T. Inexact Newton methods. *SIAM Journal on Numerical Analysis* 1982; **19**:400–408.
20. Bai Z-Z, Golub GH, Pan J-Y. Preconditioned Hermitian and skew-Hermitian splitting methods for non-Hermitian positive semidefinite linear systems. *Numerische Mathematik* 2004; **98**:1–32.
21. Eisenstat SC, Walker HF. Choosing the forcing terms in an inexact Newton method. *SIAM Journal on Scientific Computing* 1996; **17**:16–32.
22. Kantorovich LV, Akilov GP. *Functional Analysis*. Pergamon Press: Oxford, 1982.
23. Shen W-P, Li C. Kantorovich-type convergence criterion for inexact Newton methods. *Applied Numerical Mathematics* 2009; **59**:1599–1611.
24. Guo X-P. On the convergence of Newton's method in Banach space. *Journal of Zhejiang University* (*Science Edition*) 2000; **27**:484–492.
25. Eisenstat SC, Walker HF. Globally convergent inexact Newton methods. *SIAM Journal on Optimization* 1994; **4**:393–422.
26. Pernice M, Walker HF. NITSOL: a Newton iterative solver for nonlinear systems. *SIAM Journal on Scientific Computing* 1998; **19**:302–318.
27. An H-B, Bai Z-Z. A globally convergent Newton-GMRES method for large sparse systems of nonlinear equations. *Applied Numerical Mathematics* 2007; **57**:235–252.
28. Gomes-Ruggiero MA, Lopes VLR, Toledo-Benavides JV. A globally convergent inexact Newton method with a new choice for the forcing term. *Annals of Operations Research* 2008; **157**:193–205.

DOI: 10.1002/nla