

Inexact monotone methods for solving nonlinear elliptic problems

I. Boglaev¹

(Received 19 February 2015; revised 13 October 2015)

Abstract

We numerically solving semilinear elliptic problems with the method of upper and lower solutions. Inexact monotone iterative methods are constructed, where monotone linear systems are solved by the Jacobi or Gauss–Seidel methods only approximately. The inexact monotone methods combine the quadratic monotone iterative method at outer iterations and the Jacobi or Gauss–Seidel methods at inner iterations, and possess global monotone convergence. Results of numerical experiments are presented.

<http://journal.austms.org.au/ojs/index.php/ANZIAMJ/article/view/9317>

gives this article, © Austral. Mathematical Soc. 2015. Published October 27, 2015, as part of the Proceedings of the 17th Biennial Computational Techniques and Applications Conference. ISSN 1446-8735. (Print two pages per sheet of paper.) Copies of this article must not be made otherwise available on the internet; instead link directly to this URL for this article.

Contents

1	Introduction	C69
2	The monotone iterative methods	C70
2.1	The monotone iterative method	C71
2.2	The monotone inexact methods	C72
3	Numerical experiments	C77
	References	C80

1 Introduction

We consider the semilinear elliptic problem

$$\begin{aligned}
 & - \sum_{\nu=1}^{\kappa} (D_{\nu} u_{x_{\nu}})_{x_{\nu}} + \sum_{\nu=1}^{\kappa} \nu_{\nu}(x) u_{x_{\nu}} + f(x, u) = 0, \quad x \in \omega, \quad (1) \\
 & u(x) = g(x), \quad x \in \partial\omega, \quad u_{x_{\nu}} = \partial u / \partial x_{\nu},
 \end{aligned}$$

where ω is a connected bounded domain in \mathbb{R}^{κ} for $\kappa = 1, 2, \dots$ with boundary $\partial\omega$, and $D_{\nu}(x) > 0$ with $x \in \bar{\omega} = \omega \cup \partial\omega$ for $\nu = 1, \dots, \kappa$. We assume that for some constant $c_* > 0$,

$$f_u \geq c_*, \quad (x, u) \in \bar{\omega} \times (-\infty, \infty), \quad f_u \equiv \partial f / \partial u. \quad (2)$$

Section 2 introduces a nonlinear difference scheme for the numerical solution of (1).

These types of elliptic problems arise in various fields of applied sciences [8]. In the study of numerical solutions of nonlinear elliptic problems, the corresponding discrete problem is usually formulated as a system of nonlinear algebraic equations [7]. A fruitful method for solving the nonlinear difference scheme is the method of upper and lower solutions and its associated

monotone iterations. Boglaev [2] and Pao [9] gave some accelerated monotone iterative methods with a quadratic rate for solving discrete elliptic boundary value problems. However, solving a monotone linear system at each iterative step can be expensive if the number of unknowns is large [2, 9]. It seems reasonable to solve the monotone linear system at each iterative step (outer iteration) only approximately with a fixed number of iterative steps (inner iterations). One advantage of using an iterative method to solve monotone linear systems approximately is that the systems do not need to be solved too accurately in the early stages of the monotone iterative method. Inner iterations should keep global monotonicity of outer iterations. We call this combined monotone method an inexact monotone method (cf. inexact Newton methods [5]). Inexact monotone methods offer a trade-off between the accuracy with which the monotone linear systems (outer iterations) are solved and the amount of work per iteration (inner iterations).

Boglaev [4] constructed and investigated inexact monotone methods for solving semilinear parabolic problems. The aim of this article is to construct inexact monotone methods for solving semilinear elliptic problems, where the outer iterations are based on the monotone method discussed by Boglaev [2] and the inner iterations are based on the Jacobi and Gauss–Seidel iterative methods. The numerical results show that the inexact monotone methods converge with a fast linear rate of convergence (superlinear rate). The inexact monotone iterative methods are presented in Section 3.

2 The monotone iterative methods

On $\bar{\omega}$, we introduce a connected mesh $\bar{\Omega}$. For a mesh function $\mathbf{U}(\mathbf{p})$, $\mathbf{p} \in \bar{\Omega}$, consider a nonlinear difference scheme of the form

$$\begin{aligned} \mathcal{L}\mathbf{U}(\mathbf{p}) + f(\mathbf{p}, \mathbf{U}) &= 0, & \mathbf{p} \in \Omega, & \quad \mathbf{U}(\mathbf{p}) = g(\mathbf{p}), & \mathbf{p} \in \partial\Omega, & \quad (3) \\ \mathcal{L}\mathbf{U}(\mathbf{p}) &= d(\mathbf{p})\mathbf{U}(\mathbf{p}) - \sum_{\mathbf{p}' \in \sigma'(\mathbf{p})} \alpha(\mathbf{p}, \mathbf{p}')\mathbf{U}(\mathbf{p}'), \end{aligned}$$

where $\sigma'(\mathbf{p}) = \sigma(\mathbf{p}) \setminus \{\mathbf{p}\}$, $\sigma(\mathbf{p})$ is a stencil of the scheme at an interior mesh point $\mathbf{p} \in \Omega$ and $\partial\Omega$ is the boundary of $\overline{\Omega}$. We make the following assumptions on the coefficients of the difference operator \mathcal{L} :

$$d(\mathbf{p}) > 0, \quad a(\mathbf{p}, \mathbf{p}') \geq 0, \quad d(\mathbf{p}) - \sum_{\mathbf{p}' \in \sigma'(\mathbf{p})} a(\mathbf{p}, \mathbf{p}') \geq 0, \quad \mathbf{p} \in \Omega. \quad (4)$$

Remark 1. These assumptions are satisfied by the standard difference approximation of the differential operator in (1).

We now formulate the maximum principle for the difference operator \mathcal{L} from (3).

Lemma 2. *Let the coefficients of the difference operator \mathcal{L} from (3) satisfy (4) and let the mesh $\overline{\Omega}$ be connected. If a mesh function $W(\mathbf{p})$ satisfies the conditions*

$$(\mathcal{L} + c)W(\mathbf{p}) \geq 0 (\leq 0), \quad \mathbf{p} \in \Omega, \quad W(\mathbf{p}) \geq 0 (\leq 0), \quad \mathbf{p} \in \partial\Omega,$$

where $c(\mathbf{p}) \geq 0$, $\mathbf{p} \in \overline{\Omega}$, then $W(\mathbf{p}) \geq 0 (\leq 0)$, $\mathbf{p} \in \overline{\Omega}$.

Abraham and Plemmons [1] proved this lemma.

2.1 The monotone iterative method

We say that $\mathbf{U}_1(\mathbf{p})$ is an upper solution if it satisfies

$$\mathcal{R}(\mathbf{p}, \mathbf{U}_1) \equiv \mathcal{L}\mathbf{U}_1(\mathbf{p}) + f(\mathbf{p}, \mathbf{U}_1) \geq 0, \quad \mathbf{p} \in \Omega, \quad \mathbf{U}_1(\mathbf{p}) \geq g(\mathbf{p}), \quad \mathbf{p} \in \partial\Omega,$$

where $\mathcal{R}(\mathbf{p}, \mathbf{U}_1)$ is the residual of the difference scheme (3) on $\mathbf{U}_1(\mathbf{p})$. Similarly, $\mathbf{U}_{-1}(\mathbf{p})$ is a lower solution if it satisfies the reversed inequalities. Initial upper and lower solutions $\mathbf{U}_\alpha^{(0)}(\mathbf{p})$ ($\alpha = 1$ and $\alpha = -1$ correspond to the upper and lower cases, respectively) are calculated by solving the linear problems

$$\begin{aligned} (\mathcal{L} + c_*)\mathbf{Z}_\alpha^{(0)}(\mathbf{p}) &= \alpha|\mathcal{R}(\mathbf{p}, \mathbf{S})|, \quad \mathbf{p} \in \Omega, \quad \mathbf{Z}_\alpha^{(0)}(\mathbf{p}) = 0, \quad \mathbf{p} \in \partial\Omega, \quad (5) \\ \mathbf{Z}_\alpha^{(0)}(\mathbf{p}) &= \mathbf{U}_\alpha^{(0)}(\mathbf{p}) - \mathbf{S}(\mathbf{p}), \quad \mathcal{R}(\mathbf{p}, \mathbf{S}) \equiv \mathcal{L}\mathbf{S}(\mathbf{p}) + f(\mathbf{p}, \mathbf{S}), \quad \mathbf{p} \in \overline{\Omega}, \end{aligned}$$

where \mathbf{c}_* is from (2), and $\mathcal{S}(\mathbf{p})$ is defined on $\overline{\Omega}$ and satisfies the boundary condition $\mathcal{S}(\mathbf{p}) = \mathbf{g}(\mathbf{p})$ on $\partial\Omega$. For $n \geq 1$ we calculate upper and lower solutions by using the recurrence formulae

$$\begin{aligned} [\mathcal{L} + \mathbf{c}^{(n-1)}(\mathbf{p})]Z_\alpha^{(n)}(\mathbf{p}) &= -\mathcal{R}(\mathbf{p}, \mathbf{U}_\alpha^{(n-1)}), \quad \mathbf{p} \in \Omega, \\ Z_\alpha^{(n)}(\mathbf{p}) &= \mathbf{U}_\alpha^{(n)}(\mathbf{p}) - \mathbf{U}_\alpha^{(n-1)}(\mathbf{p}), \quad \mathbf{p} \in \overline{\Omega}, \quad Z_\alpha^{(n)}(\mathbf{p}) = 0, \quad \mathbf{p} \in \partial\Omega, \\ \mathcal{R}(\mathbf{p}, \mathbf{U}_\alpha^{(n-1)}) &\equiv \mathcal{L}\mathbf{U}_\alpha^{(n-1)}(\mathbf{p}) + \mathbf{f}(\mathbf{p}, \mathbf{U}_\alpha^{(n-1)}). \end{aligned} \tag{6}$$

The mesh function $\mathbf{c}^{(n-1)}$ is

$$\mathbf{c}^{(n-1)}(\mathbf{p}) = \max_{\mathbf{u}} \left[\mathbf{f}_u(\mathbf{p}, \mathbf{u}), \mathbf{U}_{-1}^{(n-1)}(\mathbf{p}) \leq \mathbf{u} \leq \mathbf{U}_1^{(n-1)}(\mathbf{p}) \right], \quad \mathbf{p} \in \overline{\Omega}. \tag{7}$$

Boglaev [2] proved that the sequences $\{\mathbf{U}_\alpha^{(n)}\}$ for $\alpha = \pm 1$, generated by (5)–(7), converge monotonically and quadratically to a unique solution \mathbf{U}^* of (3).

2.2 The monotone inexact methods

For solving the linear systems (5), we employ the Jacobi or Gauss–Seidel iterative methods and show that inexact iterative methods still maintain the monotone convergence of the iterative sequences. We write the linear problems (6) in the equivalent form for $\alpha = \pm 1$,

$$\begin{aligned} \mathcal{L}^{(n)}\mathbf{U}_\alpha^{(n)}(\mathbf{p}) + \Phi(\mathbf{p}, \mathbf{U}_\alpha^{(n-1)}) &= 0, \quad \mathbf{p} \in \Omega, \\ \mathcal{L}^{(n)} &\equiv \mathcal{L} + \mathbf{c}^{(n-1)}(\mathbf{p}), \\ \Phi(\mathbf{p}, \mathbf{U}_\alpha^{(n-1)}) &= -\mathbf{c}^{(n-1)}(\mathbf{p})\mathbf{U}_\alpha^{(n-1)}(\mathbf{p}) + \mathbf{f}(\mathbf{p}, \mathbf{U}_\alpha^{(n-1)}), \\ \mathbf{U}_\alpha^{(n)}(\mathbf{p}) &= \mathbf{g}(\mathbf{p}), \quad \mathbf{p} \in \partial\Omega. \end{aligned} \tag{8}$$

The iterative sequences $\{\mathbf{U}_\alpha^{(n,i)}(\mathbf{p})\}$ for $i = 0, 1, \dots$, generated by the Jacobi

and Gauss–Seidel methods, are defined by the recurrence formulae

$$\begin{aligned} \mathcal{L}_*^{(n)} Z_\alpha^{(n,i)}(\mathbf{p}) &= -\mathcal{H}(\mathbf{p}, \mathbf{U}_\alpha^{(n,i-1)}), \quad \mathbf{p} \in \Omega, \quad Z_\alpha^{(n,i)}(\mathbf{p}) = 0, \quad \mathbf{p} \in \partial\Omega, \quad (9) \\ Z_\alpha^{(n,i)}(\mathbf{p}) &= \mathbf{U}_\alpha^{(n,i)}(\mathbf{p}) - \mathbf{U}_\alpha^{(n,i-1)}(\mathbf{p}), \quad \mathbf{p} \in \bar{\Omega}, \quad i \geq 1, \\ \mathbf{U}_\alpha^{(n,0)}(\mathbf{p}) &= \mathbf{U}_\alpha^{(n-1)}(\mathbf{p}), \quad \mathbf{p} \in \bar{\Omega}, \\ \mathcal{H}(\mathbf{p}, \mathbf{U}_\alpha^{(n,i-1)}) &\equiv \mathcal{L}^{(n)} \mathbf{U}_\alpha^{(n,i-1)}(\mathbf{p}) + \Phi_\alpha^{(n)}(\mathbf{p}), \end{aligned}$$

where $\Phi_\alpha^{(n)}(\mathbf{p}) \equiv \Phi(\mathbf{p}, \mathbf{U}_\alpha^{(n-1)})$, and $\mathcal{H}(\mathbf{p}, \mathbf{U}_\alpha^{(n,i-1)})$ are the residuals of the difference equations (8) on $\mathbf{U}_\alpha^{(n,i-1)}(\mathbf{p})$. For the Jacobi and Gauss–Seidel methods, $\mathcal{L}_*^{(n)}$ is defined by

$$\mathcal{L}_*^{(n)} Z_\alpha^{(n,i)}(\mathbf{p}) = \begin{cases} \mathcal{L}_{\text{JAC}}^{(n)} Z_\alpha^{(n,i)}(\mathbf{p}) = [\mathbf{d}(\mathbf{p}) + \mathbf{c}^{(n-1)}(\mathbf{p})] Z_\alpha^{(n,i)}(\mathbf{p}), \\ \mathcal{L}_{\text{GS}}^{(n)} Z_\alpha^{(n,i)}(\mathbf{p}) = \mathcal{L}_{\text{JAC}}^{(n)} Z_\alpha^{(n,i)}(\mathbf{p}) - \sum_{\mathbf{p}' \in \sigma'_\Gamma(\mathbf{p})} \mathbf{a}(\mathbf{p}, \mathbf{p}') Z_\alpha^{(n,i)}(\mathbf{p}'), \end{cases} \quad (10)$$

where $\mathcal{L}_{\text{JAC}}^{(n)}$ and $\mathcal{L}_{\text{GS}}^{(n)}$ denote the difference operators in the Jacobi and Gauss–Seidel methods, respectively, and $\sigma'_\Gamma(\mathbf{p})$ is a set of stencil points corresponding to a strictly lower triangular part of $\sigma(\mathbf{p})$.


Remark 3. The sequences $\{\mathbf{U}_\alpha^{(n)}\}$ with $\alpha = \pm 1$ and $n \geq 0$ are called outer iterations, and for n fixed, the sequences $\{\mathbf{U}_\alpha^{(n,i)}\}$ with $\alpha = \pm 1$ and $i \geq 0$ are called inner iterations.

Lemma 4. *Let the coefficients $\mathbf{d}(\mathbf{p})$ and $\mathbf{a}(\mathbf{p}, \mathbf{p}')$ in the difference operator \mathcal{L} from (3) satisfy (4) and the mesh $\bar{\Omega}$ be connected. Then for the difference operators $\mathcal{L}_{\text{JAC}}^{(n)}$ and $\mathcal{L}_{\text{GS}}^{(n)}$ from (10), the maximum principle in Lemma 2 holds.*

Proof: In the case of $\mathcal{L}_{\text{JAC}}^{(n)}$, $\mathbf{d}(\mathbf{p}) + \mathbf{c}^{(n-1)}(\mathbf{p}) > 0$ is a diagonal entry and $\mathbf{a}(\mathbf{p}, \mathbf{p}') = 0$ for $\mathbf{p}' \in \sigma'(\mathbf{p})$. In the case of $\mathcal{L}_{\text{GS}}^{(n)}$, $\mathbf{d}(\mathbf{p}) + \mathbf{c}^{(n-1)}(\mathbf{p}) > 0$ is a diagonal entry and $\mathbf{a}(\mathbf{p}, \mathbf{p}') \geq 0$ for $\mathbf{p}' \in \sigma'_\Gamma(\mathbf{p})$. Thus, the coefficients of the difference operators $\mathcal{L}_{\text{JAC}}^{(n)}$ and $\mathcal{L}_{\text{GS}}^{(n)}$ satisfy the conditions (4). From here

and (7) it follows that

$$d(\mathbf{p}) + c^{(n-1)}(\mathbf{p}) - \sum_{\mathbf{p}' \in \sigma'_i(\mathbf{p})} a(\mathbf{p}, \mathbf{p}') \geq d(\mathbf{p}) - \sum_{\mathbf{p}' \in \sigma'(\mathbf{p})} a(\mathbf{p}, \mathbf{p}') \geq 0.$$

Thus, from Lemma 2, we conclude the maximum principle for $\mathcal{L}_{\text{JAC}}^{(n)}$ and $\mathcal{L}_{\text{GS}}^{(n)}$. 

In the following theorem, we prove the monotone property of the inexact iterative methods from (5)–(7), (9) and (10).

Theorem 5. *Under the assumptions of Lemma 4, the sequences $\{\mathbf{U}_\alpha^{(n)}\}$ and $\{\mathbf{U}_\alpha^{(n,i)}\}$ for $\alpha = \pm 1$, generated by (5)–(7), (9), (10) are, respectively, upper and lower solutions and converge monotonically:*

$$\begin{aligned} \mathbf{U}_{-1}^{(n-1)}(\mathbf{p}) &\leq \mathbf{U}_{-1}^{(n,i-1)}(\mathbf{p}) \leq \mathbf{U}_{-1}^{(n,i)}(\mathbf{p}) \leq \mathbf{U}_1^{(n,i)}(\mathbf{p}), & \mathbf{p} \in \overline{\Omega}, & (11) \\ \mathbf{U}_1^{(n,i)}(\mathbf{p}) &\leq \mathbf{U}_1^{(n,i-1)}(\mathbf{p}) \leq \mathbf{U}_1^{(n-1)}(\mathbf{p}), & \mathbf{p} \in \overline{\Omega}, & \end{aligned}$$

where $i \geq 1$, $\mathbf{U}_\alpha^{(n,0)}(\mathbf{p}) = \mathbf{U}_\alpha^{(n-1)}(\mathbf{p})$ and $\mathbf{U}_\alpha^{(n,i)}(\mathbf{p}) = \mathbf{U}_\alpha^{(n,i_n)}(\mathbf{p})$ for $\alpha = \pm 1$, and i_n is a number of iterative steps in the iterative method (9) for n fixed. The upper sequence $\{\mathbf{U}_1^{(n)}\}$ converges from above to the unique solution \mathbf{U}^* of (3) and the lower sequence $\{\mathbf{U}_{-1}^{(n)}\}$ converges from below to \mathbf{U}^* .

Proof: Boglaev [2] proved the existence of the unique solution \mathbf{U}^* of (3).

Taking into account that

$$\begin{aligned} \mathbf{U}_\alpha^{(1,0)}(\mathbf{p}) &= \mathbf{U}_\alpha^{(0)}(\mathbf{p}), & \mathcal{H}(\mathbf{p}, \mathbf{U}_\alpha^{(0)}) &= \mathcal{R}(\mathbf{p}, \mathbf{U}_\alpha^{(0)}), & \mathbf{p} \in \overline{\Omega}, & \alpha = \pm 1, \\ \mathcal{R}(\mathbf{p}, \mathbf{U}_1^{(0)}) &\geq 0, & \mathcal{R}(\mathbf{p}, \mathbf{U}_{-1}^{(0)}) &\leq 0, & \mathbf{p} \in \Omega, & \end{aligned}$$

from (9), by the maximum principle in Lemma 4, we conclude that

$$\mathbf{Z}_{-1}^{(1,1)}(\mathbf{p}) \geq 0, \quad \mathbf{Z}_1^{(1,1)}(\mathbf{p}) \leq 0, \quad \mathbf{p} \in \overline{\Omega}. \quad (12)$$

Thus, $\mathbf{U}_{-1}^{(0)}(\mathbf{p}) \leq \mathbf{U}_{-1}^{(1,1)}(\mathbf{p})$ and $\mathbf{U}_1^{(1,1)}(\mathbf{p}) \leq \mathbf{U}_1^{(0)}(\mathbf{p})$ for $\mathbf{p} \in \overline{\Omega}$. We now prove that

$$\mathbf{U}_{-1}^{(1,1)}(\mathbf{p}) \leq \mathbf{U}_1^{(1,1)}(\mathbf{p}), \quad \mathbf{p} \in \overline{\Omega}. \quad (13)$$

Introduce the notation

$$\mathbf{W}^{(n,i)}(\mathbf{p}) = \mathbf{U}_1^{(n,i)}(\mathbf{p}) - \mathbf{U}_{-1}^{(n,i)}(\mathbf{p}), \quad \mathbf{W}^{(n)}(\mathbf{p}) = \mathbf{U}_1^{(n)}(\mathbf{p}) - \mathbf{U}_{-1}^{(n)}(\mathbf{p}). \quad (14)$$

For $\mathbf{U}(\mathbf{p}) \geq \mathbf{V}(\mathbf{p})$ with $\mathbf{p} \in \overline{\Omega}$, we define the sector $\langle \mathbf{V}, \mathbf{U} \rangle = \{ \mathbf{V}(\mathbf{p}) \leq \mathbf{W}(\mathbf{p}) \leq \mathbf{U}(\mathbf{p}), \mathbf{p} \in \overline{\Omega} \}$. Using the mean-value theorem, from (9), we get the difference problem for $\mathbf{W}^{(1,1)}$:

$$\begin{aligned} \mathcal{L}_*^{(1)} \mathbf{W}^{(1,1)}(\mathbf{p}) &= (\mathcal{L}_*^{(1)} - \mathcal{L}^{(1)}) \mathbf{W}^{(0)}(\mathbf{p}) + (\mathbf{c}^{(0)}(\mathbf{p}) - \mathbf{f}_u(\mathbf{p}, \mathbf{Q}^{(0)})) \mathbf{W}^{(0)}(\mathbf{p}), \\ \mathbf{p} \in \Omega, \quad \mathbf{W}^{(1,1)}(\mathbf{p}) &= \mathbf{0}, \quad \mathbf{p} \in \partial\Omega, \end{aligned} \quad (15)$$

where $\mathbf{Q}^{(0)} \in \langle \mathbf{U}_{-1}^{(0)}, \mathbf{U}_1^{(0)} \rangle$. From here, (4), (8) and (10), we obtain

$$(\mathcal{L}_*^{(1)} - \mathcal{L}^{(1)}) \mathbf{W}^{(0)}(\mathbf{p}) = \sum_{\mathbf{p}' \in \sigma'_*(\mathbf{p})} \alpha(\mathbf{p}, \mathbf{p}') \mathbf{W}^{(0)}(\mathbf{p}'),$$

where $\sigma'_*(\mathbf{p}) = \sigma'(\mathbf{p})$ for $\mathcal{L}_{\text{JAC}}^{(n)}$ and $\sigma'_*(\mathbf{p}) = \sigma'_u(\mathbf{p})$ for $\mathcal{L}_{\text{GS}}^{(n)}$, where $\sigma'_u(\mathbf{p})$ is a set of stencil points corresponding to a strictly upper triangular part of $\sigma(\mathbf{p})$. From here, (4), (7) and taking into account that $\mathbf{U}_{-1}^{(n)}(\mathbf{p}) \leq \mathbf{U}_1^{(n)}(\mathbf{p})$, by Lemma 4 applied to (15), we conclude that $\mathbf{W}^{(1,1)}(\mathbf{p}) \geq \mathbf{0}$ for $\mathbf{p} \in \overline{\Omega}$, and prove (13). From (12) and (13), follows (11) for $n = 1$ and $i = 1$.

We now show that $\mathbf{U}_\alpha^{(1,1)}$ are upper ($\alpha = 1$) and lower ($\alpha = -1$) solutions to problems (8) and (3), that is,

$$\begin{aligned} \mathcal{H}(\mathbf{p}, \mathbf{U}_1^{(1,1)}) &\geq \mathbf{0}, \quad \mathcal{H}(\mathbf{p}, \mathbf{U}_{-1}^{(1,1)}) \leq \mathbf{0}, \quad \mathbf{p} \in \Omega, \\ \mathcal{R}(\mathbf{p}, \mathbf{U}_1^{(1,1)}) &\geq \mathbf{0}, \quad \mathcal{R}(\mathbf{p}, \mathbf{U}_{-1}^{(1,1)}) \leq \mathbf{0}, \quad \mathbf{p} \in \Omega. \end{aligned} \quad (16)$$

We represent $\mathcal{H}(\mathbf{p}, \mathbf{U}_\alpha^{(1,1)})$ with $\alpha = \pm 1$ in the form

$$\mathcal{H}(\mathbf{p}, \mathbf{U}_\alpha^{(1,1)}) = \mathcal{L}^{(1)} \mathbf{U}_\alpha^{(1,1)}(\mathbf{p}) + \Phi_\alpha^{(1)}(\mathbf{p}) = \mathcal{L}^{(1)} \mathbf{Z}_\alpha^{(1,1)}(\mathbf{p}) + \mathcal{H}(\mathbf{p}, \mathbf{U}_\alpha^{(1,0)}),$$

where $\mathcal{H}(\mathbf{p}, \mathbf{U}_\alpha^{(1,0)}) = \mathcal{H}(\mathbf{p}, \mathbf{U}_\alpha^{(0)})$. Using the above difference equations for $\mathbf{Z}_\alpha^{(1,1)}$ with $\alpha = \pm 1$, from (3), (9) and (10), we get

$$\mathcal{H}(\mathbf{p}, \mathbf{U}_\alpha^{(1,1)}) = (\mathcal{L}^{(1)} - \mathcal{L}_*^{(1)})\mathbf{Z}_\alpha^{(1,1)}(\mathbf{p}) = - \sum_{\mathbf{p}' \in \sigma'_*(\mathbf{p})} \mathbf{a}(\mathbf{p}, \mathbf{p}')\mathbf{Z}_\alpha^{(1,1)}(\mathbf{p}'), \quad \mathbf{p} \in \Omega, \tag{17}$$

where $\sigma'_*(\mathbf{p}) = \sigma'(\mathbf{p})$ for $\mathcal{L}_{\text{JAC}}^{(n)}$ and $\sigma'_*(\mathbf{p}) = \sigma'_\text{U}(\mathbf{p})$ for $\mathcal{L}_{\text{GS}}^{(n)}$. From here, (4) and (12), we prove (16) for $\mathcal{H}(\mathbf{p}, \mathbf{U}_\alpha^{(1,1)})$ for $\alpha = \pm 1$.

To prove (16) for $\mathcal{R}(\mathbf{p}, \mathbf{U}_\alpha^{(1,1)})$ with $\alpha = \pm 1$, from (9) and using the mean-value theorem, we represent $\mathcal{H}(\mathbf{p}, \mathbf{U}_\alpha^{(1,1)})$ in the form

$$\mathcal{H}(\mathbf{p}, \mathbf{U}_\alpha^{(1,1)}) = [\mathbf{c}^{(0)}(\mathbf{p}) - \mathbf{f}_\text{u}(\mathbf{p}, \mathbf{E}_\alpha^{(1,1)})] [\mathbf{U}_\alpha^{(1,1)}(\mathbf{p}) - \mathbf{U}_\alpha^{(0)}(\mathbf{p})] + \mathcal{R}(\mathbf{p}, \mathbf{U}_\alpha^{(1,1)}), \tag{18}$$

where $\mathbf{E}_1^{(1,1)} \in \langle \mathbf{U}_1^{(1,1)}, \mathbf{U}_1^{(0)} \rangle$ and $\mathbf{E}_{-1}^{(1,1)} \in \langle \mathbf{U}_{-1}^{(0)}, \mathbf{U}_{-1}^{(1,1)} \rangle$. From here, (7), (12) and (16), we conclude that

$$\mathcal{R}(\mathbf{p}, \mathbf{U}_1^{(1,1)}) \geq \mathcal{H}(\mathbf{p}, \mathbf{U}_1^{(1,1)}) \geq 0, \quad \mathcal{R}(\mathbf{p}, \mathbf{U}_{-1}^{(1,1)}) \leq \mathcal{H}(\mathbf{p}, \mathbf{U}_{-1}^{(1,1)}) \leq 0, \quad \mathbf{p} \in \Omega. \tag{19}$$

Thus, we prove (16) for $\mathcal{R}(\mathbf{p}, \mathbf{U}_\alpha^{(1,1)})$ with $\alpha = \pm 1$.

From (9) and (16), by the maximum principle in Lemma 4, we conclude that

$$\mathbf{Z}_{-1}^{(1,2)}(\mathbf{p}) \geq 0, \quad \mathbf{Z}_1^{(1,2)}(\mathbf{p}) \leq 0, \quad \mathbf{p} \in \overline{\Omega}. \tag{20}$$

Thus,

$$\mathbf{U}_{-1}^{(0)}(\mathbf{p}) \leq \mathbf{U}_{-1}^{(1,1)}(\mathbf{p}) \leq \mathbf{U}_{-1}^{(1,2)}(\mathbf{p}), \quad \mathbf{U}_1^{(1,2)}(\mathbf{p}) \leq \mathbf{U}_1^{(1,1)}(\mathbf{p}) \leq \mathbf{U}_1^{(0)}(\mathbf{p}), \quad \mathbf{p} \in \overline{\Omega}.$$

We now prove that

$$\mathbf{U}_{-1}^{(1,2)}(\mathbf{p}) \leq \mathbf{U}_1^{(1,2)}(\mathbf{p}), \quad \mathbf{p} \in \overline{\Omega}. \tag{21}$$

Similar to (15), using the mean-value theorem, from (9), we get the difference problem for $\mathbf{W}^{(1,2)}$:

$$\begin{aligned} \mathcal{L}_*^{(1)}\mathbf{W}^{(1,2)}(\mathbf{p}) &= [\mathcal{L}_*^{(1)} - \mathcal{L}^{(1)}] \mathbf{W}^{(1,1)}(\mathbf{p}) + [\mathbf{c}^{(0)}(\mathbf{p}) - \mathbf{f}_\text{u}(\mathbf{p}, \mathbf{Q}^{(0)})] \mathbf{W}^{(0)}(\mathbf{p}), \\ \mathbf{p} \in \Omega, \quad \mathbf{W}^{(1,2)}(\mathbf{p}) &= 0, \quad \mathbf{p} \in \partial\Omega, \end{aligned} \tag{22}$$

where $\mathbf{Q}^{(0)} \in \langle \mathbf{V}_{-1}^{(0)}, \mathbf{V}_1^{(0)} \rangle$. From here, (4), (8) and (10),

$$(\mathcal{L}_*^{(1)} - \mathcal{L}^{(1)})\mathbf{W}^{(1,1)}(\mathbf{p}) = \sum_{\mathbf{p}' \in \sigma'_*(\mathbf{p})} \alpha(\mathbf{p}, \mathbf{p}')\mathbf{W}^{(1,1)}(\mathbf{p}'), \quad (23)$$

where $\sigma'_*(\mathbf{p}) = \sigma'(\mathbf{p})$ for $\mathcal{L}_{\text{JAC}}^{(n)}$ and $\sigma'_*(\mathbf{p}) = \sigma'_u(\mathbf{p})$ for $\mathcal{L}_{\text{GS}}^{(n)}$. From here, (4), (7), (11) with $\mathbf{n} = 1$, $\mathbf{i} = 1$, and taking into account that $\mathbf{U}_{-1}^{(n)}(\mathbf{p}) \leq \mathbf{U}_1^{(n)}(\mathbf{p})$, by Lemma 4 applied to (22), we conclude that $\mathbf{W}^{(1,2)}(\mathbf{p}) \geq 0$ for $\mathbf{p} \in \overline{\Omega}$ and prove (21). From (20) and (21), follows (11) for $\mathbf{n} = 1$ and $\mathbf{i} = 2$.

This completes the proof that $\mathbf{U}_\alpha^{(1,2)}$ are upper ($\alpha = 1$) and lower ($\alpha = -1$) solutions to problems (8) and (3), which is equivalent to (16).

By induction on \mathbf{i} , we prove (11) for $\mathbf{n} = 1$, and prove that $\mathbf{U}_1^{(1)} = \mathbf{U}_1^{(1, \mathbf{i}_1)}$ and $\mathbf{U}_{-1}^{(1)} = \mathbf{U}_{-1}^{(1, \mathbf{i}_1)}$ are, respectively, upper and lower solutions to (3). Now, by induction on \mathbf{n} and \mathbf{i} , we prove (11).

The upper sequence $\{\mathbf{U}_1^{(n)}\}$ is a monotone decreasing sequence, which is bounded by $\mathbf{U}_{-1}^{(0)}$. From here it follows that $\lim_{n \rightarrow \infty} \mathbf{Z}_1^{(n)}(\mathbf{p}) = 0$ for $\mathbf{p} \in \overline{\Omega}$. From here and (6) we conclude that $\lim_{n \rightarrow \infty} \mathbf{R}(\mathbf{p}, \mathbf{U}_1^{(n)}) = 0$ for $\mathbf{p} \in \Omega$. Thus, $\lim_{n \rightarrow \infty} \mathbf{U}_1^{(n)}(\mathbf{p}) = \mathbf{U}^*(\mathbf{p})$ for $\mathbf{p} \in \overline{\Omega}$. Similarly, we can prove this result for the lower sequence $\{\mathbf{U}_{-1}^{(n)}\}$. ♠

3 Numerical experiments

In this section, we compare numerically convergence properties of the inexact monotone Jacobi and Gauss–Seidel methods and the corresponding monotone iterative methods of Boglaev [3]. The monotone iterative methods of Boglaev [3] are constructed using the assumption that

$$0 < c_* \leq f_u \leq c^*, \quad c_*, c^* = \text{const} > 0. \quad (24)$$

These methods utilize \mathbf{c}^* in (6) instead of $\mathbf{c}^{(n-1)}(\mathbf{p})$. Denoting the difference operators \mathcal{L}_* , \mathcal{L}_{JAC} and \mathcal{L}_{GS} from (10) with $\mathbf{c}^{(n-1)}(\mathbf{p}) = \mathbf{c}^*$ by $\tilde{\mathcal{L}}_*$, $\tilde{\mathcal{L}}_{\text{JAC}}$ and $\tilde{\mathcal{L}}_{\text{GS}}$, respectively, we present the methods of Boglaev [3] in the form of (5) and the recurrence formulae

$$\begin{aligned} \tilde{\mathcal{L}}_* \tilde{\mathbf{Z}}_\alpha^{(n)}(\mathbf{p}) &= -\mathcal{R}(\mathbf{p}, \tilde{\mathbf{U}}_\alpha^{(n-1)}), \quad \mathbf{p} \in \bar{\Omega}, \\ \tilde{\mathbf{Z}}_\alpha^{(n)}(\mathbf{p}) &= \tilde{\mathbf{U}}_\alpha^{(n)}(\mathbf{p}) - \tilde{\mathbf{U}}_\alpha^{(n-1)}(\mathbf{p}), \quad \mathbf{p} \in \bar{\Omega}, \quad \tilde{\mathbf{Z}}_\alpha^{(n)}(\mathbf{p}) = \mathbf{0}, \quad \mathbf{p} \in \partial\Omega, \\ \mathcal{R}(\mathbf{p}, \tilde{\mathbf{U}}_\alpha^{(n-1)}) &\equiv \mathcal{L} \tilde{\mathbf{U}}_\alpha^{(n-1)}(\mathbf{p}) + \mathbf{f}(\mathbf{p}, \tilde{\mathbf{U}}_\alpha^{(n-1)}), \quad \alpha = \pm 1, \end{aligned} \tag{25}$$

where

$$\tilde{\mathcal{L}}_* = \begin{cases} \tilde{\mathcal{L}}_{\text{JAC}} \tilde{\mathbf{U}}_\alpha^{(n)}(\mathbf{p}) = [\mathbf{d}(\mathbf{p}) + \mathbf{c}^*] \tilde{\mathbf{U}}_\alpha^{(n)}(\mathbf{p}), \\ \tilde{\mathcal{L}}_{\text{GS}} \tilde{\mathbf{U}}_\alpha^{(n)}(\mathbf{p}) = \tilde{\mathcal{L}}_{\text{JAC}} \tilde{\mathbf{U}}_\alpha^{(n)}(\mathbf{p}) - \sum_{\mathbf{p}' \in \sigma'_L(\mathbf{p})} \mathbf{a}(\mathbf{p}, \mathbf{p}') \tilde{\mathbf{U}}_\alpha^{(n)}(\mathbf{p}'). \end{cases}$$

Boglaev [3] proved that the monotone iterative methods (5) and (25) converge linearly to the solution of the nonlinear difference scheme (3).

Our implementation of the inexact monotone Jacobi and Gauss–Seidel iterative methods is based on the framework of an inexact Newton method by Dembo et al. [5]. Dembo et al. [5] defined an inexact Newton method as a generalization of Newton’s method for solving the system of nonlinear equations $\mathbf{F}(\mathbf{x}) = \mathbf{0}$ for $\mathbf{F} : \mathbb{R}^k \rightarrow \mathbb{R}^k$. At the m th iteration, the step $\mathbf{s}_m = \mathbf{x}_{m+1} - \mathbf{x}_m$ is defined by

$$\mathbf{F}'(\mathbf{x}_m) \mathbf{s}_m = -\mathbf{F}(\mathbf{x}_m) + \mathbf{r}_m,$$

and satisfies

$$\frac{\|\mathbf{r}_m\|}{\|\mathbf{F}(\mathbf{x}_m)\|} \leq \eta_m,$$

where $\eta_m \in [0, 1)$ is the forcing term. One advantage of using an iterative method to solve the Newton linear system $\mathbf{F}'(\mathbf{x}_m) \mathbf{s}_m = -\mathbf{F}(\mathbf{x}_m)$ approximately is that the system does not need to be solved too accurately in the early stages of the Newton iteration. Thus, the inexact Newton method offers a trade-off between the accuracy with which the Newton system is solved and the amount of work per iteration. In typical applications, the choice of the forcing terms

is critical to the efficiency of the method and can affect robustness [6]. For the nonlinear system $F(x) = 0$, our choice of forcing term η_m is

$$\eta_m = \begin{cases} \eta_{\max}, & m = 0, \\ \min(\eta_{\max}, \|r_m\|/\|F(x_m)\|), & m > 0, \end{cases}$$

where $\eta_{\max} \in (0, 1)$ [6].

For the inexact monotone Jacobi and Gauss–Seidel iterative methods, the forcing term is

$$\eta_\alpha^{(n-1)} = \begin{cases} \eta_{\max,\alpha}, & n = 1, \\ \min\left(\eta_{\max,\alpha}, \frac{\|\mathcal{H}(\cdot, \mathbf{u}_\alpha^{(n,i)})\|_\Omega}{\|\mathcal{R}(\cdot, \mathbf{u}_\alpha^{(n-1)})\|_\Omega}\right), & n > 1, \end{cases}$$

where $\eta_{\max,\alpha} \in (0, 1)$. The stopping test for outer iterations is chosen to be

$$\|\mathcal{R}(\cdot, \mathbf{u}_\alpha^{(n)})\|_\Omega \leq \delta, \quad \alpha = \pm 1, \tag{26}$$

where δ is a prescribed accuracy.

As the test problem for (1), we consider the diffusion-reaction problem

$$\begin{aligned} -(\mathbf{u}_{xx} + \mathbf{u}_{yy}) + (e^{-1} - e^{-\mathbf{u}}) &= 0, \quad (x, y) \in \omega, \\ \omega &= \{0 < x < 1\} \times \{0 < y < 1\}, \quad \mathbf{u}(x, y) = 0, \quad (x, y) \in \partial\omega. \end{aligned}$$

From $f_u = e^{-u} > 0$ and $0 \leq u \leq 1$, condition (2) is satisfied with $c_* = e^{-1}$. From here and $u \geq 0$ we conclude that $c^* = 1$ in (24). Since $f_{uu} = -e^{-u} < 0$, in (7) $c^{(n-1)}(p) = f_u(p, \mathbf{u}_{-1}^{(n-1)})$ [2]. For the approximation of the differential operator, we use the central difference approximations on the five-point square stencil with $N_1 = N_2 = N$, which satisfies (4). A lower solution is defined by $\mathbf{u}_{-1}^{(0)}(\Omega) = 0$ and $\mathbf{u}_{-1}^{(0)}(\partial\Omega) = 1$. We choose the convergence tolerance $\delta = 10^{-5}$ in (26).

In Table 1, for various values of N , we give convergence iteration counts. Here I_1 and I_2 are the results, corresponding the monotone iterative methods of

Table 1: Comparison of iteration counts.

N	16	32	64	128
Jacobi method				
I_1	557	2233	8934	35740
I_2	400(5)	1321(5)	4521(5)	17979(5)
η_{\max}	0.095	0.097	0.098	0.098
I_1/I_2	1.392	1.690	1.976	1.999
Gauss–Seidel method				
I_1	281	1118	4469	17872
I_2	224(5)	707(5)	2744(5)	10324(6)
η_{\max}	0.096	0.097	0.098	0.098
I_1/I_2	1.232	1.458	1.629	1.731

Boglaev [3] and the inexact monotone iterative methods (5)–(7), (9), (10), respectively. Also, a number in (·) indicates the number of outer iterations. Optimal values of η_{\max} are also provided.

From the numerical data it follows that for all values of N the inexact monotone methods converge faster than the corresponding monotone iterative methods of Boglaev [3]. The iterative methods of Boglaev [3] utilize the constant \mathbf{c}^* in (6), which is independent of \mathbf{n} . Although the mesh function $\mathbf{c}^{(n-1)}$ in (6) is computed at each outer iteration, the overall amount of computational work in the inexact monotone methods is much less than in the case of the iterative methods of Boglaev [3].

References

- [1] B. Abraham and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*. Academic Press, New York, 1979.
doi:[10.1137/1.9781611971262](https://doi.org/10.1137/1.9781611971262) C71

- [2] I. Boglaev, Uniform quadratic convergence of monotone iterates for semilinear singularly perturbed elliptic problems. *Lecture Notes in Computational Science and Engineering* 81:37–46, 2011. doi:[10.1007/978-3-642-19665-2_5](https://doi.org/10.1007/978-3-642-19665-2_5) C70, C72, C74, C79
- [3] I. Boglaev, Monotone relaxation iterates and applications to semilinear singularly perturbed problems. *Int. J. Numer. Anal. Mod.(B)* 2:402–414, 2011. <http://www.math.ualberta.ca/ijnamb/Volume-2-2011/No-4-11/2011-04-08.pdf> C77, C78, C80
- [4] I. Boglaev, An inexact monotone method for solving semilinear parabolic problems. *Appl. Math. Comput.* 219:3253–3263, 2012. doi:[10.1016/j.amc.2012.09.067](https://doi.org/10.1016/j.amc.2012.09.067) C70
- [5] R. S. Dembo, S. C. Eisenstat and T. Steihaug, Inexact Newton methods. *SIAM J. Numer. Anal.* 19:400–408, 1982 doi:[10.1137/0719025](https://doi.org/10.1137/0719025) C70, C78
- [6] S. C. Eisenstat and H. F. Walker, Choosing the forcing terms in an inexact Newton method. *SIAM J. Sci. Comput.* 17:16–32, 1996. doi:[10.1137/0917003](https://doi.org/10.1137/0917003) C79
- [7] P. Knabner and L. Angerman, *Numerical Methods for Elliptic and Parabolic Partial Differential Equations*. Springer, New York, 2003. doi:[10.1007/b97419](https://doi.org/10.1007/b97419) C69
- [8] C. V. Pao, *Nonlinear Parabolic and Elliptic Equations*. Springer, New York, 1992. doi:[10.1007/978-1-4615-3034-3](https://doi.org/10.1007/978-1-4615-3034-3) C69
- [9] C. V. Pao, Accelerated monotone iterations for numerical solutions of nonlinear elliptic boundary value problems. *Computers Math. Applic.*, 46:1535–1544, 2003. doi:[10.1016/S0898-1221\(03\)00381-X](https://doi.org/10.1016/S0898-1221(03)00381-X) C70

Author address

1. **I. Boglaev**, Institute of Fundamental Sciences, Massey University, Palmerston North, New Zealand.
<mailto:i.boglaev@massey.ac.nz>