Journal of Computational Mathematics, Vol.26, No.2, 2008, 123–148.

A MONOTONE COMPACT IMPLICIT SCHEME FOR NONLINEAR REACTION-DIFFUSION EQUATIONS*

Yuanming Wang
Department of Mathematics, East China Normal University, Shanghai 200241, China;
Division of Computational Science, E-Institute of Shanghai Universities, Shanghai Normal University, Shanghai 200234, China
Email: ymwang@math.ecnu.edu.cn
Benyu Guo
Department of Mathematics, Shanghai Normal University, Shanghai 200234, China;
Division of Computational Science, E-Institute of Shanghai Universities, Shanghai, China
Email: byguo@shnu.edu.cn

Abstract

A monotone compact implicit finite difference scheme with fourth-order accuracy in space and second-order in time is proposed for solving nonlinear reaction-diffusion equations. An accelerated monotone iterative method for the resulting discrete problem is presented. The sequence of iteration converges monotonically to the unique solution of the discrete problem, and the convergence rate is either quadratic or nearly quadratic, depending on the property of the nonlinear reaction. The numerical results illustrate the high accuracy of the proposed scheme and the rapid convergence rate of the iteration.

Mathematics subject classification: 65M06, 65M12.
Key words: Nonlinear reaction-diffusion equation, Monotone compact implicit scheme, High accuracy, Monotone iteration, Rapid convergence rate.

1. Introduction

Many phenomena in physics, chemistry, biology and engineering are described by nonlinear reaction-diffusion equations. Much work has been done for the qualitative analysis of the equations (see [19] and references therein), as well as their numerical simulations (see, e.g., [7,10,13,17,18,20,21,23,24,28]). In this paper, we provide a new numerical treatment for a class of nonlinear reaction-diffusion equations. It includes the construction and analysis of a monotone compact implicit finite difference scheme with high accuracy, and an accelerated monotone iterative method with rapid convergence rate for solving the resulting discrete problem. The equation under consideration is of the form:

$$\begin{cases} \partial u/\partial t + \mathcal{L}u = f(x,t,u), \qquad 0 < x < 1, \quad 0 < t \le T, \\ u(0,t) = g_0(t), \quad u(1,t) = g_1(t), \qquad 0 < t \le T, \\ u(x,0) = u_0(x), \qquad 0 \le x \le 1, \end{cases} \tag{1.1}$$

where $g_0(t)$, $g_1(t)$ and $u_0(x)$ are given continuous functions satisfying the compatibility conditions $u_0(0) = g_0(0)$ and $u_0(1) = g_1(0)$. The operator $\mathcal{L}u$ in (1.1) is given by

$$\mathcal{L}u = -\frac{\partial}{\partial x}\left(k(x)\frac{\partial u}{\partial x}\right), \tag{1.2}$$

* Received November 21, 2006 / Revised version received April 5, 2007 / Accepted April 16, 2007 /

where the coefficient $k(x) \in C^1(0,1)$ and for certain constants α_0 and α_1,

$$0 < \alpha_0 \le k(x) \le \alpha_1, \qquad x \in (0,1). \tag{1.3}$$

The function $f(x,t,u)$ in (1.1) is continuous in its domain, and the function $f(\cdot,u)$, which is in general nonlinear in u, is continuously differentiable in u.

Various numerical methods have been developed for solving problem (1.1). In the usual finite difference methods, one approximates the term $\partial u/\partial t$ by Euler backward method and the differential operator $\mathcal{L}u$ by the central difference quotient (see, e.g., [7,10,13,17,18,20,21]). In this case, the resulting discrete system is tridiagonal, and so it does not need any fictitious points for implementing the scheme. However, such scheme has only the accuracy of $\mathcal{O}(\tau + h^2)$ where τ and h are the mesh sizes in time and in space, respectively (e.g., [15,17,18,20,21]). In other words, we must take small mesh sizes in order to obtain the desirable accuracy, and thus much computational work is involved.

As is well known, by using the Crank-Nicolson technique or the three-level Lees technique in the time discretization, the accuracy in time can be improved to second-order (see [4,15,25]). But if Lees technique is used, one has to evaluate the solution at the first time level by other method (see [4,15,25]). Another trick for improving accuracy in time is to use extrapolation technique (see [25]). For improvement of the accuracy in space, a conventional approach is to approximate $\mathcal{L}u$ by using more points in the space discretization (see [4]). However, this not only significantly increases the computational complexity but also causes difficulty in handling boundary conditions since fictitious points near boundaries are needed (see [4]).

An alternative approach of improving the accuracy in space is the so-called compact implicit method which has been developed and generalized by several investigators under the name *Operator Compact Implicit* (OCI) method (see, in particular, [2–4]). This method achieves the fourth-order accuracy while retaining the tridiagonal feature of a second-order method and not requiring additional fictitious points at the boundary (see [2–4,14]). Assume that the function $u(x)$ is independent of t. The main idea of the OCI method is to look for an approximation representation of $\mathcal{L}u$ by establishing the following relationship between $\mathcal{L}u$ and the function u on the three adjacent points of a uniform mesh $x_i = ih$ $(h = 1/L, i = 0, 1, \cdots, L)$:

$$r_i^- u_{i-1} + r_i^c u_i + r_i^+ u_{i+1} = q_i^- (\mathcal{L}u)_{i-1} + q_i^c (\mathcal{L}u)_i + q_i^+ (\mathcal{L}u)_{i+1}, \qquad 1 \le i \le L-1, \tag{1.4}$$

or

$$\mathcal{R}_i u_i = \mathcal{Q}_i (\mathcal{L}u)_i, \qquad 1 \le i \le L-1,$$

where u_i and $(\mathcal{L}u)_i$ are the approximations to u and $\mathcal{L}u$ at x_i, respectively, and the operators $\mathcal{R}_i$ and $\mathcal{Q}_i$ are tridiagonal operators:

$$\mathcal{R}_i u_i = r_i^- u_{i-1} + r_i^c u_i + r_i^+ u_{i+1}, \quad \mathcal{Q}_i u_i = q_i^- u_{i-1} + q_i^c u_i + q_i^+ u_{i+1}, \qquad 1 \le i \le L-1. \tag{1.5}$$

This approximation representation for $\mathcal{L}u$ is *explicit* if $q_i^- = q_i^+ = 0$, and *implicit* otherwise. Without loss of generality, throughout this paper, (1.4) is assumed normalized so that

$$\lim_{h \to 0} q_i^c = \text{a positive constant}, \qquad 1 \le i \le L-1. \tag{1.6}$$

Following the terminology of [3,4], a scheme of the form (1.4) will be referred to as an *Operator Compact Implicit* (OCI) scheme if it is a fourth-order accurate approximation to $\mathcal{L}u$, i.e., if its truncation error is $\mathcal{O}(h^4)$ after normalization. Note that the fourth-order accuracy is the highest that can be obtained by a scheme of the form (1.4) (see, e.g., [3,14]).

We now discretize (1.1). Let $t_n = n\tau$ ($\tau = T/N, n = 0, 1, \cdots, N$) be the mesh size in time. By combining the Crank-Nicolson discretization in time with the technique used in space approximation as in scheme (1.4), we obtain the following difference equation for (1.1):

$$\Big(\mathcal{Q}_i + \frac{\tau}{2}\mathcal{R}_i\Big) u_{i,n} = \Big(\mathcal{Q}_i - \frac{\tau}{2}\mathcal{R}_i\Big) u_{i,n-1} + \frac{\tau}{2}\mathcal{Q}_i\Big(f(x_i, t_n, u_{i,n}) + f(x_i, t_{n-1}, u_{i,n-1})\Big), \qquad (1.7)$$

where $u_{i,n}$ is the approximation to $u(x_i, t_n)$. Define

$$\begin{aligned}
&\Lambda = \{(i,n) : i = 1, \cdots, L-1;\ n = 1, \cdots, N\},\\
&\partial\Lambda = \{(i,n) : i = 0, L; n = 0, 1, \cdots, N\} \cup \{(i,0) : i = 0, 1, \cdots, L\},\\
&\overline{\Lambda} = \Lambda \cup \partial\Lambda.
\end{aligned}$$

Since a fundamental property of the problem (1.1) is its maximum principle (see [19]), it is reasonable to require that (1.7) also possesses an analogous discrete maximum principle as follows:

Discrete maximum principle. The inequality

$$\Big(\mathcal{Q}_i + \frac{\tau}{2}\mathcal{R}_i\Big) u_{i,n} \le \Big(\mathcal{Q}_i - \frac{\tau}{2}\mathcal{R}_i\Big) u_{i,n-1}, \qquad (i,n) \in \Lambda \qquad (1.8)$$

implies

$$\max_{(i,n)\in\Lambda} u_{i,n} \le \max_{(i,n)\in\partial\Lambda} u_{i,n}. \qquad (1.9)$$

The scheme satisfying the discrete maximum principle is often known as a *monotone* scheme. For the monotonicity of (1.7) we have the following result.

Theorem 1.1. *The scheme* (1.7) *is monotone if the following condition is satisfied:*

$$\text{(C)} \begin{cases} q_i^- \ge 0, \quad q_i^c > 0, \quad q_i^+ \ge 0, \qquad r_i^- + r_i^c + r_i^+ = 0, \\ q_i^- + \frac{\tau}{2} r_i^- < 0, \quad q_i^c - \frac{\tau}{2} r_i^c > 0, \quad q_i^+ + \frac{\tau}{2} r_i^+ < 0, \end{cases} \qquad 1 \le i \le L-1.$$

Proof. It suffices to prove that the inequality (1.8) implies (1.9). Define $a_i^{-,c,+}$ (i.e., a_i^-, a_i^c, a_i^+) and $b_i^{-,c,+}$ as follows:

$$a_i^{-,c,+} = q_i^{-,c,+} + \frac{\tau}{2} r_i^{-,c,+}, \qquad b_i^{-,c,+} = q_i^{-,c,+} - \frac{\tau}{2} r_i^{-,c,+}.$$

Then by the condition (C), we have

$$a_i^c > 0, \quad -a_i^{-,+} > 0, \quad b_i^{-,c,+} > 0, \quad a_i^c = -a_i^- - a_i^+ + b_i^- + b_i^c + b_i^+. \qquad (1.10)$$

Moreover, by the definitions of $\mathcal{Q}_i$ and $\mathcal{R}_i$, the inequality (1.8) reads

$$a_i^- u_{i-1,n} + a_i^c u_{i,n} + a_i^+ u_{i+1,n} \le b_i^- u_{i-1,n-1} + b_i^c u_{i,n-1} + b_i^+ u_{i+1,n-1}. \qquad (1.11)$$

Now, let

$$u_{i_0,n_0} = M_{\overline{\Lambda}} = \max_{(i,n)\in\overline{\Lambda}} u_{i,n}$$

for some $(i_0, n_0) \in \overline{\Lambda}$. If $(i_0, n_0) \in \partial\Lambda$, then (1.9) follows immediately. Otherwise, we have from (1.10) and (1.11) that

$$\begin{aligned} M_{\overline{\Lambda}} = u_{i_0,n_0} \le \frac{1}{a_{i_0}^c}(&-a_{i_0}^- u_{i_0-1,n_0} - a_{i_0}^+ u_{i_0+1,n_0} + b_{i_0}^- u_{i_0-1,n_0-1} \\ &+ b_{i_0}^c u_{i_0,n_0-1} + b_{i_0}^+ u_{i_0+1,n_0-1}) \le M_{\overline{\Lambda}}. \end{aligned} \qquad (1.12)$$

Using (1.10) again, we assert that all the values $u_{i,n}$ involved in (1.12) are equal to $M_{\overline{\Lambda}}$. Hence $M_{\overline{\Lambda}}$ is attained also at all the points which are the connected neighbors of (i_0, n_0). The same argument is valid at each of these points. Finally, all $u_{i,n}$ take the same value $M_{\overline{\Lambda}}$. Therefore, (1.9) is valid.

Since it is not easy to check the monotonicity of (1.7) directly, we turn to pay attention to the condition (C). A monotone scheme of the form (1.7) will be called a *Monotone Compact Implicit* (MCI) scheme in this paper if the corresponding steady scheme (1.4) is an OCI scheme.

Motivated by Hermite's generalization of the Taylor series, Collatz proposed an OCI scheme in [5,6]. Independently, Swartz [26] also derived a similar OCI scheme from approximation theory considerations. Later a broad family of OCI schemes was obtained in [4] using a straightforward Taylor expansion, which generalizes Swartz's work. For an overview of OCI schemes, see the monograph [14]. However, when the above technique is applied to the operator $\mathcal{L}u$ given by (1.2), the relationship (1.4) (or more precisely, the definitions of the operators $\mathcal{Q}_i$ and $\mathcal{R}_i$) becomes complicated since it depends not only on the function $k(x)$ itself but also on the derivative of it. In addition, we require some severe conditions on $k(x)$ as well as on $k'(x)$ (for example, $k'(x) \geq 0$) (see [14]) for the monotone condition (C). This feature limits the application of this method. In this paper, we propose a new approach for $\mathcal{L}u$ in its original form to obtain an OCI scheme and then establish a MCI scheme in the form (1.7) with the accuracy of $\mathcal{O}(\tau^2 + h^4)$. Such a scheme depends only on the function $k(x)$ itself, and the monotone condition (C) can also be easily verified with some conditions on the function $k(x)$ only. It is noted that our scheme was partially motivated by the work in [8] or [9] where a relationship between $\mathcal{L}u$ and u was established by using the local Green's function.

Due to the nonlinearity of the problem (1.7), some kind of iteration process is required. The frequently used iteration processes are either the Picard type or the Newton type [16–18,21,22,30]. The Picard's method yields a sequence converging to a unique solution monotonically. However, the convergence rate is only linear. In the treatment of a chemical reactor model in [17], for example, the convergence rate of the Picard iteration is very slow for certain physical parameters. Although the Newton's method is quadratically convergent, the corresponding sequence of iteration does not possess, in general, any monotone property. Moreover, the Newton's method has a strict requirement on the initial data for its quadratic convergence (see [16,30]). To increase the convergence rate while maintaining the monotone property of the iteration, we propose an accelerated monotone iteration for the problem (1.7) by the method of upper and lower solutions. It is shown, by using upper and lower solutions as a pair of coupled initial iterations, that the iteration yields two monotone sequences which converge monotonically from above and below, respectively, to a unique solution. Its convergence rate is either quadratic or nearly quadratic, with the usual differentiability requirement only on the function $f(\cdot, u)$. On the other hand, since the initial iteration in the monotone iteration is either an upper solution or a lower solution, which can be constructed directly from the problem (1.7) without any knowledge of the solution, our method avoids the search for the initial iteration as is often needed in the Newton's method. Indeed, this is another advantage of our method.

This paper is organized as follows. In Section 2, we design a new OCI scheme of the form (1.4) and then combine the Crank-Nicolson discretization in time with the OCI scheme to propose a MCI scheme of the form (1.7). This MCI scheme is treated in Section 3 by the method of upper and lower solutions. The existence and uniqueness of the solution is discussed. In Section 4, we prove that the MCI scheme has the accuracy of $\mathcal{O}(\tau^2+h^4)$. Section 5 is devoted to an accelerated monotone iteration for resolving the resulting discrete system. This iteration is reduced to the Newton's method if the reaction function $f(\cdot,u)$ possesses a concavity or convexity property between upper and lower solutions. The quadratic convergence rate of the iteration is analyzed. Some explicit estimates for the convergence rate are given. In Section 6, we give some numerical results which demonstrate the monotonicity of the iteration and the high accuracy of the numerical solution. The comparison of the convergence rate of the accelerated monotone iteration with that of the Picard iteration is also given. The final section gives some concluding remarks.

2. The Derivation of the MCI Scheme

In this section we derive a MCI scheme and discuss its basic properties.

2.1. An OCI scheme

Consider a uniform mesh $x_i = ih$ $(i = 0, 1, \cdots, L)$ where $h = 1/L$ is the mesh size. The local Green's function $G_i(x,s)$ of the operator $\mathcal{L}$ given by (1.2) is defined by

$$\begin{cases} \mathcal{L}G_i(x,s) = \delta(x,s), & (x,s) \in (x_{i-1},x_i) \times [x_{i-1},x_i], \qquad 1 \le i \le L, \\ G_i(x_{i-1},s) = G_i(x_i,s) = 0, & s \in [x_{i-1},x_i], \qquad 1 \le i \le L. \end{cases}$$

It can be verified that (see [8,9])

$$G_i(x,s) = \begin{cases} g_{1,i}(s)g_{2,i}(x)/J_i, & x \le s, \\ g_{1,i}(x)g_{2,i}(s)/J_i, & x > s, \end{cases} \tag{2.1a}$$

where

$$J_i = \left(\int_{x_{i-1}}^{x_i} \frac{1}{k(s)}\,\mathrm{d}s \right)^{-1}, \quad g_{1,i}(x) = J_i \int_x^{x_i} \frac{1}{k(s)}\,\mathrm{d}s, \quad g_{2,i}(x) = J_i \int_{x_{i-1}}^{x} \frac{1}{k(s)}\,\mathrm{d}s. \tag{2.1b}$$

By the above local Green's function, we have the following relationship between $\mathcal{L}u(x)$ and $u(x)$ (see [8,9]):

$$-J_i u(x_{i-1}) + (J_i + J_{i+1})u(x_i) - J_{i+1}u(x_{i+1}) = \Psi_i(\mathcal{L}u), \tag{2.2}$$

where

$$\Psi_i(\mathcal{L}u) = \int_{x_i}^{x_{i+1}} \mathcal{L}u(x) g_{1,i+1}(x)\,\mathrm{d}x + \int_{x_{i-1}}^{x_i} \mathcal{L}u(x) g_{2,i}(x)\,\mathrm{d}x.$$

Define $v(x) = \mathcal{L}u(x)$, and let $H_i(x)$ be the Hermite interpolant of $v(x)$ on $[x_{i-1}, x_{i+1}]$ such that

$$H_i'(x_i) = v'(x_i), \quad H_i(x_j) = v(x_j), \qquad j = i-1, i, i+1.$$

Then by approximation theory we have

$$H_i(x) = \phi_{i,1}(x)v(x_{i-1}) + \phi_{i,2}(x)v(x_i) + \phi_{i,3}(x)v(x_{i+1}) + p_i(x),$$

where

$$\phi_{i,1}(x) = -\frac{x-x_i}{2h} + \frac{(x-x_i)^2}{2h^2}, \quad \phi_{i,2}(x) = 1 - \frac{(x-x_i)^2}{h^2}, \quad \phi_{i,3}(x) = \frac{x-x_i}{2h} + \frac{(x-x_i)^2}{2h^2},$$

and

$$p_i(x) = -\frac{1}{h^2}v'(x_i)(x-x_{i-1})(x-x_i)(x-x_{i+1})$$
$$+\frac{1}{2h^3}(v(x_{i+1})-v(x_{i-1}))(x-x_{i-1})(x-x_i)(x-x_{i+1}).$$

Moreover for all $x \in [x_{i-1}, x_{i+1}]$,

$$v(x) = H_i(x) + \frac{1}{24}\frac{\mathrm{d}^4 v}{\mathrm{d}x^4}(\xi_i)(x-x_{i-1})(x-x_i)^2(x-x_{i+1}),$$

where $\xi_i \in [x_{i-1}, x_{i+1}]$. On the other hand,

$$v'(x_i) = \frac{1}{2h}\left(v(x_{i+1})-v(x_{i-1})\right) - \frac{h^2}{6}\frac{\mathrm{d}^3 v}{\mathrm{d}x^3}(\eta_i),$$

where $\eta_i \in [x_{i-1}, x_{i+1}]$ and depends only on x_j $(j = i-1, i, i+1)$. Consequently, for all $x \in [x_{i-1}, x_{i+1}]$,

$$v(x) = \phi_{i,1}(x)v(x_{i-1}) + \phi_{i,2}(x)v(x_i) + \phi_{i,3}(x)v(x_{i+1}) + \varepsilon_i(x),$$

or equivalently

$$\mathcal{L}u(x) = \phi_{i,1}(x)\mathcal{L}u(x_{i-1}) + \phi_{i,2}(x)\mathcal{L}u(x_i) + \phi_{i,3}(x)\mathcal{L}u(x_{i+1}) + \varepsilon_i(x), \tag{2.3}$$

where

$$\varepsilon_i(x) = \frac{1}{6}\frac{\mathrm{d}^3 v}{\mathrm{d}x^3}(\eta_i)(x-x_{i-1})(x-x_i)(x-x_{i+1})$$
$$+\frac{1}{24}\frac{\mathrm{d}^4 v}{\mathrm{d}x^4}(\xi_i)(x-x_{i-1})(x-x_i)^2(x-x_{i+1}).$$

By substituting (2.3) into the expression of $\Psi_i(\mathcal{L}u)$, we deduce from (2.2) that

$$\begin{aligned} &-J_i u(x_{i-1}) + (J_i + J_{i+1})u(x_i) - J_{i+1}u(x_{i+1}) \\ &= E_i\mathcal{L}u(x_{i-1}) + F_i\mathcal{L}u(x_i) + G_i\mathcal{L}u(x_{i+1}) + \widetilde{\varepsilon}_i, \end{aligned} \tag{2.4}$$

where

$$E_i = \int_{x_i}^{x_{i+1}} \phi_{i,1}(x)g_{1,i+1}(x)\mathrm{d}x + \int_{x_{i-1}}^{x_i} \phi_{i,1}(x)g_{2,i}(x)\mathrm{d}x,$$
$$F_i = \int_{x_i}^{x_{i+1}} \phi_{i,2}(x)g_{1,i+1}(x)\mathrm{d}x + \int_{x_{i-1}}^{x_i} \phi_{i,2}(x)g_{2,i}(x)\mathrm{d}x,$$
$$G_i = \int_{x_i}^{x_{i+1}} \phi_{i,3}(x)g_{1,i+1}(x)\mathrm{d}x + \int_{x_{i-1}}^{x_i} \phi_{i,3}(x)g_{2,i}(x)\mathrm{d}x,$$
$$\widetilde{\varepsilon}_i = \int_{x_i}^{x_{i+1}} \varepsilon_i(x)g_{1,i+1}(x)\mathrm{d}x + \int_{x_{i-1}}^{x_i} \varepsilon_i(x)g_{2,i}(x)\mathrm{d}x.$$

Here, $g_{1,i+1}(x)$ and $g_{2,i}(x)$ are defined by (2.1b). Neglecting the term $\widetilde{\varepsilon}_i$ in (2.4) we obtain the following relationship between u_i and $(\mathcal{L}u)_i$:

$$\begin{aligned} &-J_i u_{i-1} + (J_i + J_{i+1})u_i - J_{i+1}u_{i+1} \\ &= E_i(\mathcal{L}u)_{i-1} + F_i(\mathcal{L}u)_i + G_i(\mathcal{L}u)_{i+1}, \end{aligned} \tag{2.5}$$

where u_i and $(\mathcal{L}u)_i$ are the approximations to $u(x_i)$ and $\mathcal{L}u(x_i)$, respectively. A similar scheme is obtained in [12] for interface problems. Clearly, the relation (2.5) has the form (1.4) with the coefficients

$$\begin{aligned} r_i^- &= -J_i, \quad r_i^c = J_i + J_{i+1}, \quad r_i^+ = -J_{i+1}, \\ q_i^- &= E_i, \quad q_i^c = F_i, \quad q_i^+ = G_i. \end{aligned} \tag{2.6}$$

Theorem 2.1. *The scheme* (2.5) *is an OCI scheme.*

Proof. Since η_i in the expression of $\varepsilon_i(x)$ depends only on x_j $(j = i-1, i, i+1)$ we have by the Taylor expansion of $\sigma(x) = 1/k(x)$ at x_i that

$$\widetilde{\varepsilon}_i = \frac{1}{6}\frac{\mathrm{d}^3 v}{\mathrm{d}x^3}(\eta_i)\frac{J_{i+1}J_i}{k^2(x_i)}\left(\widetilde{\varepsilon}_{i,1} + \widetilde{\varepsilon}_{i,2}\right)h + \mathcal{O}(h^5),$$

where

$$\begin{aligned} \widetilde{\varepsilon}_{i,1} &= \int_{x_i}^{x_{i+1}}\int_{x_i}^{x}(s - x_{i-1})(s - x_i)(s - x_{i+1})\mathrm{d}s\mathrm{d}x, \\ \widetilde{\varepsilon}_{i,2} &= \int_{x_{i-1}}^{x_i}\int_{x}^{x_i}(s - x_{i-1})(s - x_i)(s - x_{i+1})\mathrm{d}s\mathrm{d}x. \end{aligned}$$

A direct calculation yields $\widetilde{\varepsilon}_{i,1} + \widetilde{\varepsilon}_{i,2} = 0$. This implies $\widetilde{\varepsilon}_i = \mathcal{O}(h^5)$. It is easy to see from condition (1.3) that $F_i = \mathcal{O}(h)$ and $F_i > 0$ for each i. These statements imply that the truncation error of (2.5) is $\mathcal{O}(h^4)$ after normalization according to (1.6), and thus (2.5) is an OCI scheme. □

Remark 2.1. The OCI scheme (2.5) depends only on the function $k(x)$. Although some integrals are involved in this scheme, they can be easily calculated if $k(x)$ is an elementary function. It will be seen in the next subsection that for this scheme the monotone condition (C) can also be easily checked with some conditions on $k(x)$ and the mesh sizes.

2.2. A MCI scheme

Let the operators $\mathcal{R}_i$ and $\mathcal{Q}_i$ be given in (1.5), where the coefficients $r_i^{-,c,+}$ (i.e., r_i^-, r_i^c, r_i^+) and $q_i^{-,c,+}$ are defined by (2.6). Combining the Crank-Nicolson time discretization with the trick used in the OCI scheme (2.5) we obtain the following finite difference equation of the form (1.7) for (1.1),

$$\begin{cases} \left(\mathcal{Q}_i + \dfrac{\tau}{2}\mathcal{R}_i\right)u_{i,n} = \left(\mathcal{Q}_i - \dfrac{\tau}{2}\mathcal{R}_i\right)u_{i,n-1} + \dfrac{\tau}{2}\mathcal{Q}_i\left(f(x_i, t_n, u_{i,n})\right. \\ \qquad\qquad \left.+ f(x_i, t_{n-1}, u_{i,n-1})\right), \qquad 1 \le i \le L-1; \quad 1 \le n \le N, \\ u_{0,n} = g_0(t_n), \quad u_{L,n} = g_1(t_n), \qquad 1 \le n \le N, \\ u_{i,0} = u_0(x_i), \qquad 0 \le i \le L. \end{cases} \tag{2.7}$$

Obviously, the scheme (2.7) is a generalization of the Crandall scheme (see [11]) or the Douglas scheme (see [25]) for the equation $u_t = u_{xx}$.

To check the monotone property of the scheme (2.7) we establish some lemmas.

Lemma 2.1. *There exists a positive constant $h^* \le 1$ such that for all $h \le h^*$,*

$$E_i \ge 0, \quad F_i > 0, \quad G_i \ge 0, \qquad 1 \le i \le L-1.$$

Proof. By the condition (1.3), it is clear that $F_i > 0$ for all $1 \le i \le L-1$. Furthermore by the Taylor expansion of $\sigma(x) = 1/k(x)$ at x_i, we have that for all $1 \le i \le L-1$,

$$\begin{aligned} E_i &= J_i J_{i+1} \left[\left(\sigma(x_i)h + \mathcal{O}(h^2)\right) \left(-\frac{1}{24}\sigma(x_i)h^2 + \mathcal{O}(h^3) \right) \right. \\ &\quad \left. + \left(\sigma(x_i)h + \mathcal{O}(h^2)\right) \left(\frac{1}{8}\sigma(x_i)h^2 + \mathcal{O}(h^3) \right) \right] \\ &= J_i J_{i+1} \left(\frac{1}{12}\sigma^2(x_i)h^3 + \mathcal{O}(h^4) \right). \end{aligned}$$

Similarly,

$$G_i = J_i J_{i+1} \left(\frac{1}{12}\sigma^2(x_i)h^3 + \mathcal{O}(h^4) \right), \qquad 1 \le i \le L-1.$$

Since $J_i > 0$ and $J_{i+1} > 0$ for all $h > 0$, there exists a positive constant $h^* \le 1$ such that for all $h \le h^*$, both E_i and G_i are nonnegative. This completes the proof. $\square$

Remark 2.2. In some specific problems, we may obtain the precise value of h^* through a direct calculation of E_i and G_i. For example, a simple calculation leads to that

$$E_i \ge \left(\frac{\alpha_0}{8\alpha_1} - \frac{\alpha_1}{24\alpha_0} \right) h, \qquad G_i \ge \left(\frac{\alpha_0}{8\alpha_1} - \frac{\alpha_1}{24\alpha_0} \right) h.$$

Thus if $\alpha_1 \le \sqrt{3}\alpha_0$, then $E_i \ge 0$ and $G_i \ge 0$ for all $h \ge 0$.

Lemma 2.2. ([27, Corollary 1, p.85]) *If $A = (a_{i,j})$ is a real, irreducibly diagonally dominant $n \times n$ matrix with $a_{i,j} \le 0$ for all $i \ne j$, and $a_{i,i} > 0$ for all $1 \le i \le n$, then A^{-1} exists and is nonnegative.*

Throughout the paper, $R = (R_{i,j})$ and $Q = (Q_{i,j})$ denote the $(L-1) \times (L-1)$ tridiagonal matrices with the respective elements

$$\begin{aligned} R_{i,j} &= -\delta_{i,j+1}J_{j+1} + \delta_{i,j}(J_j + J_{j+1}) - \delta_{i,j-1}J_j, \\ Q_{i,j} &= \delta_{i,j+1}E_{j+1} + \delta_{i,j}F_j + \delta_{i,j-1}G_{j-1}, \end{aligned}$$

where $\delta_{i,j} = 1$ if $i = j$, and $\delta_{i,j} = 0$ otherwise. We have the following result.

Lemma 2.3. *Let $M_{i,n}$ $(i = 0, 1, \cdots, L; n = 0, 1, \cdots, N)$ be some given constants, and define $D_n = \mathrm{diag}(M_{1,n}, \cdots, M_{L-1,n})$. Assume that $3\alpha_1 \le 10\alpha_0$ and the mesh sizes h and τ satisfy that*

$$\begin{cases} h \le h^*, \qquad -2 < \tau M_{i,n} < \dfrac{20\alpha_0 - 6\alpha_1}{10\alpha_0 + 3\alpha_1}, \\ \dfrac{2 + \tau M_{i,n}}{8\alpha_0} < \dfrac{\tau}{h^2} < \dfrac{10 - 5\tau M_{i,n}}{12\alpha_1}, \end{cases} \qquad 0 \le i \le L;\ 0 \le n \le N, \tag{2.8}$$

where h^* is the constant in Lemma 2.1. Then for all $1 \le i \le L-1$ and $0 \le n \le N$,

(i) $E_i \ge 0$, $F_i > 0$, *and* $G_i \ge 0$;

(ii) $\left(1 + \frac{\tau}{2}M_{i-1,n}\right) E_i - \frac{\tau}{2}J_i < 0$ *and* $\left(1 + \frac{\tau}{2}M_{i+1,n}\right) G_i - \frac{\tau}{2}J_{i+1} < 0$;

(iii) $\left(1 - \frac{\tau}{2}M_{i,n}\right) F_i - \frac{\tau}{2}(J_i + J_{i+1}) > 0$;

(iv) $\left(Q + \frac{\tau}{2}QD_n + \frac{\tau}{2}R\right)^{-1} \ge 0$.

Proof. (i) The result (i) follows from Lemma 2.1.

(ii) By the condition (1.3), we have

$$E_i \le \frac{h^2}{8\alpha_0} J_i, \quad G_i \le \frac{h^2}{8\alpha_0} J_{i+1}, \qquad 1 \le i \le L-1.$$

This relation and the condition (2.8) lead to the result (ii).

(iii) Due to the condition (1.3) we can easily verify that

$$F_i \ge \frac{5}{12\alpha_1} h^2 \left(J_i + J_{i+1}\right).$$

Then the result (iii) follows from the above estimate and the condition (2.8).

(iv) The results in (i) and (ii) imply that the tridiagonal matrix $Q + \frac{\tau}{2} Q D_n + \frac{\tau}{2} R$ satisfies the conditions in Lemma 2.2, and thus its inverse exists and is nonnegative. The conclusion (iv) is proved. □

By Lemma 2.3, we see immediately that the monotone condition (C) is valid for (2.7) if the condition (2.8) holds with $M_{i,n} \equiv 0$. This leads to the following conclusion.

Theorem 2.2. *Assume that $3\alpha_1 \le 10\alpha_0$, and that the mesh sizes h and τ fulfill the condition* (2.8) *with $M_{i,n} \equiv 0$. Then the scheme* (2.7) *is a MCI scheme.*

More generally, we have the following positivity lemma.

Lemma 2.4. *Let the conditions in Lemma* 2.3 *be satisfied. If $u_{i,n}$ is given by*

$$\begin{cases} \left(\mathcal{Q}_i + \frac{\tau}{2}\mathcal{R}_i\right) u_{i,n} + \frac{\tau}{2}\mathcal{Q}_i \left(M_{i,n} u_{i,n}\right) \ge \left(\mathcal{Q}_i - \frac{\tau}{2}\mathcal{R}_i\right) u_{i,n-1} - \frac{\tau}{2}\mathcal{Q}_i \left(M_{i,n-1} u_{i,n-1}\right), \\ \qquad\qquad 1 \le i \le L-1; \quad 1 \le n \le N, \\ u_{0,n} \ge 0, \quad u_{L,n} \ge 0, \qquad 1 \le n \le N, \\ u_{i,0} \ge 0, \qquad 0 \le i \le L, \end{cases} \tag{2.9}$$

then $u_{i,n} \ge 0$ for all $0 \le i \le L$ and $0 \le n \le N$.

Proof. By Lemma 2.3,

$$Q - \frac{\tau}{2} Q D_n - \frac{\tau}{2} R \ge 0, \quad \left(Q + \frac{\tau}{2} Q D_n + \frac{\tau}{2} R\right)^{-1} \ge 0.$$

Writing (2.9) in the matrix form, we obtain the desired result by an induction on n. □

Remark 2.3. In the special case $k(x) \equiv D$, where D is a positive constant, the condition (2.8) can be weakened as

$$\begin{cases} -2 < \tau M_{i,n} < \frac{4}{3}, \\ \frac{2 + \tau M_{i,n}}{12D} < \frac{\tau}{h^2} < \frac{10 - 5\tau M_{i,n}}{12D}, \end{cases} \qquad 0 \le i \le L; \ 0 \le n \le N. \tag{2.10}$$

3. The Existence and Uniqueness of the Solution

To study the existence and uniqueness of solution of (2.7) we need a pair of ordered upper and lower solutions defined as follows.

Definition 3.1. *A function* $\overline{u}_{i,n}$ $(i = 0, 1, \cdots, L; n = 0, 1, \cdots, N)$ *is called an upper solution of* (2.7) *if*

$$\begin{cases} \left(\mathcal{Q}_i + \frac{\tau}{2}\mathcal{R}_i\right)\overline{u}_{i,n} \geq \left(\mathcal{Q}_i - \frac{\tau}{2}\mathcal{R}_i\right)\overline{u}_{i,n-1} + \frac{\tau}{2}\mathcal{Q}_i\left(f(x_i, t_n, \overline{u}_{i,n}) \right. \\ \qquad\qquad \left. + f(x_i, t_{n-1}, \overline{u}_{i,n-1})\right), \qquad 1 \leq i \leq L-1; \quad 1 \leq n \leq N, \\ \overline{u}_{0,n} \geq g_0(t_n), \quad \overline{u}_{L,n} \geq g_1(t_n), \qquad 1 \leq n \leq N, \\ \overline{u}_{i,0} \geq u_0(x_i), \qquad 0 \leq i \leq L. \end{cases} \tag{3.1}$$

Similarly, $\underline{u}_{i,n}$ *is called a lower solution if it satisfies the above inequalities in reversed order. The pair* $\overline{u}_{i,n}, \underline{u}_{i,n}$ *are said to be ordered if* $\overline{u}_{i,n} \geq \underline{u}_{i,n}$ *for all* $0 \leq i \leq L$ *and* $0 \leq n \leq N$*.*

It is clear that every solution of (2.7) is an upper solution as well as a lower solution. For any pair of ordered upper and lower solutions $\overline{u}_{i,n}$, $\underline{u}_{i,n}$ we define the sector

$$\langle \underline{u}_{i,n}, \overline{u}_{i,n} \rangle = \{u_{i,n} \in \mathbf{R} : \underline{u}_{i,n} \leq u_{i,n} \leq \overline{u}_{i,n}\}. \tag{3.2}$$

Theorem 3.1. *Let* $\overline{u}_{i,n}$, $\underline{u}_{i,n}$ *be a pair of ordered upper and lower solutions of* (2.7)*, and assume that there exist constants* $\underline{M}_{i,n}$ *such that*

$$f_u(x_i, t_n, \xi_{i,n}) \geq -\underline{M}_{i,n}, \qquad \xi_{i,n} \in \langle \underline{u}_{i,n}, \overline{u}_{i,n} \rangle, \quad 0 \leq i \leq L; \ 0 \leq n \leq N, \tag{3.3}$$

where $f_u = \partial f/\partial u$*. If* $3\alpha_1 \leq 10\alpha_0$ *and the condition* (2.8) *holds with* $M_{i,n} = \underline{M}_{i,n}$*, then* (2.7) *has a maximal solution* $\overline{u}^*_{i,n}$ *and a minimal solution* $\underline{u}^*_{i,n}$ *in* $\langle \underline{u}_{i,n}, \overline{u}_{i,n} \rangle$*. Here, the maximal property of* $\overline{u}^*_{i,n}$ *means that for any solution* $u_{i,n}$ *of* (2.7) *in* $\langle \underline{u}_{i,n}, \overline{u}_{i,n} \rangle$*, we have* $u_{i,n} \leq \overline{u}^*_{i,n}$ *for all* $0 \leq i \leq L$ *and* $0 \leq n \leq N$*. The minimal property of* $\underline{u}^*_{i,n}$ *is understood similarly.*

Proof. The proof is constructive. Consider the following iteration:

$$\begin{cases} \left(\mathcal{Q}_i + \frac{\tau}{2}\mathcal{R}_i\right)u^{(m)}_{i,n} + \frac{\tau}{2}\mathcal{Q}_i\left(\underline{M}_{i,n}u^{(m)}_{i,n}\right) = \left(\mathcal{Q}_i - \frac{\tau}{2}\mathcal{R}_i\right)u^{(m)}_{i,n-1} + \frac{\tau}{2}\mathcal{Q}_i\left(\underline{M}_{i,n}u^{(m-1)}_{i,n}\right) \\ \quad + \frac{\tau}{2}\mathcal{Q}_i\left(f(x_i, t_n, u^{(m-1)}_{i,n}) + f(x_i, t_{n-1}, u^{(m)}_{i,n-1})\right), \qquad 1 \leq i \leq L-1; \quad 1 \leq n \leq N, \\ u^{(m)}_{0,n} = g_0(t_n), \quad u^{(m)}_{L,n} = g_1(t_n), \qquad 1 \leq n \leq N, \\ u^{(m)}_{i,0} = u_0(x_i), \qquad 0 \leq i \leq L. \end{cases} \tag{3.4}$$

By Lemma 2.3, the above iteration is well defined under the condition (2.8) as long as the initial iteration $u^{(0)}_{i,n}$ is given. Denote the sequence by $\{\overline{u}^{(m)}_{i,n}\}$ if $u^{(0)}_{i,n} = \overline{u}_{i,n}$, and by $\{\underline{u}^{(m)}_{i,n}\}$ if $u^{(0)}_{i,n} = \underline{u}_{i,n}$. We shall first prove that for all $m \geq 0$,

$$\underline{u}^{(m)}_{i,n} \leq \underline{u}^{(m+1)}_{i,n} \leq \overline{u}^{(m+1)}_{i,n} \leq \overline{u}^{(m)}_{i,n}, \qquad 0 \leq i \leq L; \ 0 \leq n \leq N. \tag{3.5}$$

Let

$$\overline{w}^{(0)}_{i,n} = \overline{u}^{(0)}_{i,n} - \overline{u}^{(1)}_{i,n}, \quad \underline{w}^{(0)}_{i,n} = \underline{u}^{(1)}_{i,n} - \underline{u}^{(0)}_{i,n}, \quad w^{(1)}_{i,n} = \overline{u}^{(1)}_{i,n} - \underline{u}^{(1)}_{i,n}$$

for all $0 \le i \le L$ and $0 \le n \le N$. Then by (3.1), (3.3) and (3.4),

$$\begin{aligned}
&\left(\mathcal{Q}_i + \frac{\tau}{2}\mathcal{R}_i\right) \overline{w}_{i,n}^{(0)} + \frac{\tau}{2}\mathcal{Q}_i \left(\underline{M}_{i,n}\overline{w}_{i,n}^{(0)}\right) \ge \left(\mathcal{Q}_i - \frac{\tau}{2}\mathcal{R}_i\right) \overline{w}_{i,n-1}^{(0)} - \frac{\tau}{2}\mathcal{Q}_i \left(\underline{M}_{i,n-1}\overline{w}_{i,n-1}^{(0)}\right) \\
&\quad + \frac{\tau}{2}\mathcal{Q}_i \left(\underline{M}_{i,n-1}\left(\overline{u}_{i,n-1}^{(0)} - \overline{u}_{i,n-1}^{(1)}\right) + f(x_i, t_{n-1}, \overline{u}_{i,n-1}^{(0)}) - f(x_i, t_{n-1}, \overline{u}_{i,n-1}^{(1)})\right), \\
&\left(\mathcal{Q}_i + \frac{\tau}{2}\mathcal{R}_i\right) \underline{w}_{i,n}^{(0)} + \frac{\tau}{2}\mathcal{Q}_i \left(\underline{M}_{i,n}\underline{w}_{i,n}^{(0)}\right) \ge \left(\mathcal{Q}_i - \frac{\tau}{2}\mathcal{R}_i\right) \underline{w}_{i,n-1}^{(0)} - \frac{\tau}{2}\mathcal{Q}_i \left(\underline{M}_{i,n-1}\underline{w}_{i,n-1}^{(0)}\right) \\
&\quad + \frac{\tau}{2}\mathcal{Q}_i \left(\underline{M}_{i,n-1}\left(\underline{u}_{i,n-1}^{(1)} - \underline{u}_{i,n-1}^{(0)}\right) + f(x_i, t_{n-1}, \underline{u}_{i,n-1}^{(1)}) - f(x_i, t_{n-1}, \underline{u}_{i,n-1}^{(0)})\right), \\
&\left(\mathcal{Q}_i + \frac{\tau}{2}\mathcal{R}_i\right) w_{i,n}^{(1)} + \frac{\tau}{2}\mathcal{Q}_i \left(\underline{M}_{i,n}w_{i,n}^{(1)}\right) \ge \left(\mathcal{Q}_i - \frac{\tau}{2}\mathcal{R}_i\right) w_{i,n-1}^{(1)} - \frac{\tau}{2}\mathcal{Q}_i \left(\underline{M}_{i,n-1}w_{i,n-1}^{(1)}\right) \\
&\quad + \frac{\tau}{2}\mathcal{Q}_i \left(\underline{M}_{i,n-1}\left(\overline{u}_{i,n-1}^{(1)} - \underline{u}_{i,n-1}^{(1)}\right) + f(x_i, t_{n-1}, \overline{u}_{i,n-1}^{(1)}) - f(x_i, t_{n-1}, \underline{u}_{i,n-1}^{(1)})\right).
\end{aligned}$$

Since $\overline{u}_{i,0}^{(0)} \ge \overline{u}_{i,0}^{(1)} = \underline{u}_{i,0}^{(1)} \ge \underline{u}_{i,0}^{(0)}$, it follows from (3.3) and Lemma 2.3 that

$$\overline{w}_{i,1}^{(0)} \ge 0, \quad \underline{w}_{i,1}^{(0)} \ge 0, \quad w_{i,1}^{(1)} \ge 0, \qquad 0 \le i \le L.$$

Consequently,

$$\underline{u}_{i,1}^{(0)} \le \underline{u}_{i,1}^{(1)} \le \overline{u}_{i,1}^{(1)} \le \overline{u}_{i,1}^{(0)}, \qquad 0 \le i \le L.$$

By an induction on n, we know that (3.5) holds for $m = 0$. Finally, the desired result (3.5) follows from an induction argument about m.

In view of (3.5), there exist the limits

$$\lim_{m\to\infty} \overline{u}_{i,n}^{(m)} = \overline{u}_{i,n}^{*}, \quad \lim_{m\to\infty} \underline{u}_{i,n}^{(m)} = \underline{u}_{i,n}^{*}, \qquad 0 \le i \le L; \ \ 0 \le n \le N. \tag{3.6}$$

Letting $m \to \infty$ in (3.4), we assert that both $\overline{u}_{i,n}^{*}$ and $\underline{u}_{i,n}^{*}$ are solutions of (2.7).

Now, if $u_{i,n}$ is a solution of (2.7) in $\langle \underline{u}_{i,n}, \overline{u}_{i,n}\rangle$, then $u_{i,n}$, $\underline{u}_{i,n}$ is also a pair of ordered upper and lower solutions of (2.7). The above arguments imply that $\underline{u}_{i,n}^{*} \le u_{i,n}$ for all $0 \le i \le L$ and $0 \le n \le N$. Similarly, we have $u_{i,n} \le \overline{u}_{i,n}^{*}$ for all $0 \le i \le L$ and $0 \le n \le N$. This implies that $\overline{u}_{i,n}^{*}$ and $\underline{u}_{i,n}^{*}$ are the maximal and minimal solutions of (2.7) in $\langle \underline{u}_{i,n}, \overline{u}_{i,n}\rangle$, respectively. The proof is complete. □

Theorem 3.2. *Let the condition in Theorem* 3.1 *hold. Assume, in addition, that there exist constants* $\overline{M}_{i,n}$ *such that* $\tau\overline{M}_{i,n} < 2$ *and*

$$f_u(x_i, t_n, \xi_{i,n}) \le \overline{M}_{i,n}, \qquad \xi_{i,n} \in \langle \underline{u}_{i,n}, \overline{u}_{i,n}\rangle, \quad 0 \le i \le L; \ 0 \le n \le N. \tag{3.7}$$

Then (2.7) *has a unique solution* $u_{i,n}^{*}$ *in* $\langle \underline{u}_{i,n}, \overline{u}_{i,n}\rangle$.

Proof. Let $\overline{u}_{i,n}^{*}$ and $\underline{u}_{i,n}^{*}$ be the limits in (3.6). It suffices to verify that $\overline{u}_{i,n}^{*} = \underline{u}_{i,n}^{*}$ for all $0 \le i \le L$ and $0 \le n \le N$. Let $w_{i,n}^{*} = \overline{u}_{i,n}^{*} - \underline{u}_{i,n}^{*}$. Then $w_{i,n}^{*} \ge 0$ and

$$\begin{aligned}
\left(\mathcal{Q}_i + \frac{\tau}{2}\mathcal{R}_i\right) w_{i,n}^{*} &= \left(\mathcal{Q}_i - \frac{\tau}{2}\mathcal{R}_i\right) w_{i,n-1}^{*} + \frac{\tau}{2}\mathcal{Q}_i\left(f(x_i, t_n, \overline{u}_{i,n}^{*}) - f(x_i, t_n, \underline{u}_{i,n}^{*})\right) \\
&\quad + \frac{\tau}{2}\mathcal{Q}_i\left(f(x_i, t_{n-1}, \overline{u}_{i,n-1}^{*}) - f(x_i, t_{n-1}, \underline{u}_{i,n-1}^{*})\right) \\
&\le \left(\mathcal{Q}_i - \frac{\tau}{2}\mathcal{R}_i\right) w_{i,n-1}^{*} + \frac{\tau}{2}\mathcal{Q}_i\left(\overline{M}_{i,n}w_{i,n}^{*}\right) + \frac{\tau}{2}\mathcal{Q}_i\left(\overline{M}_{i,n-1}w_{i,n-1}^{*}\right),
\end{aligned}$$

which gives

$$\left(\mathcal{Q}_i+\frac{\tau}{2}\mathcal{R}_i\right)w^*_{i,n}-\frac{\tau}{2}\mathcal{Q}_i\left(\overline{M}_{i,n}w^*_{i,n}\right)\le\left(\mathcal{Q}_i-\frac{\tau}{2}\mathcal{R}_i\right)w^*_{i,n-1}+\frac{\tau}{2}\mathcal{Q}_i\left(\overline{M}_{i,n-1}w^*_{i,n-1}\right).$$

Since condition (2.8) holds with $M_{i,n}=\underline{M}_{i,n}$ and $\underline{M}_{i,n}\ge-\overline{M}_{i,n}$, it also holds with $M_{i,n}=-\overline{M}_{i,n}$ under the condition $\tau\overline{M}_{i,n}<2$. By Lemma 2.4, $w^*_{i,n}\le 0$. Therefore, $w^*_{i,n}=0$ which implies $\overline{u}^*_{i,n}=\underline{u}^*_{i,n}$ for all $0\le i\le L$ and $0\le n\le N$. This completes the proof. □

Remark 3.1. It is seen from the proofs of Theorems 3.1 and 3.2 that the monotone iteration (3.4) not only leads to the existence and uniqueness of solution of (2.7) but also provides an algorithm for the numerical solution. However, its convergence rate is only linear because it is of Picard type. An accelerated monotone iteration will be developed in Section 5.

4. The Convergence of the MCI Scheme (2.7)

In this section, we deal with the convergence of the MCI scheme (2.7). Specifically, we prove that the solution of MCI scheme (2.7) converges to the solution of (1.1) with the accuracy of $\mathcal{O}(\tau^2+h^4)$ as $\tau\to 0$ and $h\to 0$. For this purpose we introduce the following lemma.

Lemma 4.1. *Let $\{\zeta_i\}$ be a sequence of real numbers such that for certain $0<\gamma<1$ and $\delta>0$,*

$$|\zeta_i|\le\gamma|\zeta_i|+(1+\gamma)|\zeta_{i-1}|+\delta,\qquad i=1,2,\cdots. \tag{4.1}$$

Then

$$|\zeta_i|\le \mathrm{e}^{\frac{2i\gamma}{1-\gamma}}|\zeta_0|+\frac{\delta}{2\gamma}\left(\mathrm{e}^{\frac{2i\gamma}{1-\gamma}}-1\right),\qquad i=0,1,\cdots. \tag{4.2}$$

Proof. We prove the lemma by induction. Clearly, (4.2) holds for $i=0$. Now, assume that (4.2) is true for some $i\ge 0$. Then

$$\begin{aligned}|\zeta_{i+1}|&\le\frac{1+\gamma}{1-\gamma}|\zeta_i|+\frac{\delta}{1-\gamma}\\&\le\left(1+\frac{2\gamma}{1-\gamma}\right)|\zeta_i|+\frac{\delta}{1-\gamma}\\&\le\left(1+\frac{2\gamma}{1-\gamma}\right)\mathrm{e}^{\frac{2i\gamma}{1-\gamma}}|\zeta_0|+\frac{\delta}{2\gamma}\left(1+\frac{2\gamma}{1-\gamma}\right)\left(\mathrm{e}^{\frac{2i\gamma}{1-\gamma}}-1\right)+\frac{\delta}{1-\gamma}.\end{aligned}$$

Since

$$1+\frac{2\gamma}{1-\gamma}\le\mathrm{e}^{\frac{2\gamma}{1-\gamma}},$$

we have

$$|\zeta_{i+1}|\le\mathrm{e}^{\frac{2(i+1)\gamma}{1-\gamma}}|\zeta_0|+\frac{\delta}{2\gamma}\left(\mathrm{e}^{\frac{2(i+1)\gamma}{1-\gamma}}-1\right).$$

The induction is complete. □

In the rest part of this paper, the letter C with subscript denotes a generic positive constant that is independent of τ and h and may not be the same at different occurrences.

The main result in this section is stated in the following theorem.

Theorem 4.1. *Let $u(x,t)$ and $u_{i,n}$ be the solutions of* (1.1) *and* (2.7)*, respectively. Assume that*

(i) there exists a positive constant $\overline{M}$ such that $\tau\overline{M} < 1$ and for $1 \le i \le L-1$ and $1 \le n \le N$,

$$|f_u(x_i, t_n, \xi_{i,n})| \le \overline{M}, \qquad \xi_{i,n} \in \langle \min\{u(x_i,t_n), u_{i,n}\}, \max\{u(x_i,t_n), u_{i,n}\}\rangle; \tag{4.3}$$

(ii) the functions $u_{ttt}(x,t)$ and $v_{xxxx}(x,t)$ are all continuous for all $x \in [0,1]$ and $t \in [0,T]$, where $v(x,t) = \mathcal{L}u(x,t)$;

(iii) $3\alpha_1 \le 10\alpha_0$ and the condition (2.8) *holds with $M_{i,n} \equiv 0$.*

Then

$$|u(x_i,t_n) - u_{i,n})| \le C_1\left(\tau^2 + h^4\right), \qquad 0 \le i \le L;\ 0 \le n \le N. \tag{4.4}$$

Proof. Combining the Crank-Nicolson time discretization with the relation (2.4) we have that for all $1 \le i \le L-1$ and $1 \le n \le N$,

$$\begin{aligned}\Big(\mathcal{Q}_i + \frac{\tau}{2}\mathcal{R}_i\Big) u(x_i,t_n) = \Big(\mathcal{Q}_i - \frac{\tau}{2}\mathcal{R}_i\Big) u(x_i,t_{n-1}) \\ +\frac{\tau}{2}\mathcal{Q}_i\Big(f(x_i,t_n,u(x_i,t_n)) + f(x_i,t_{n-1},u(x_i,t_{n-1}))\Big) + \widetilde{\varepsilon}_{i,n} + \widehat{\varepsilon}_{i,n},\end{aligned} \tag{4.5}$$

where $\widehat{\varepsilon}_{i,n}$ is the truncation error due to the Crank-Nicolson time discretization, which satisfies

$$|\widehat{\varepsilon}_{i,n}| \le C_2\tau^3 h, \qquad 1 \le i \le L-1;\ \ 1 \le n \le N. \tag{4.6}$$

On the other hand, we have from the proof of Theorem 2.1 that

$$|\widetilde{\varepsilon}_{i,n}| \le C_3\tau h^5, \qquad 1 \le i \le L-1;\ \ 1 \le n \le N. \tag{4.7}$$

Let $w_{i,n} = u(x_i,t_n) - u_{i,n}$. Then by (4.5) and (2.7),

$$\begin{cases} \Big(\mathcal{Q}_i + \frac{\tau}{2}\mathcal{R}_i\Big) w_{i,n} = \Big(\mathcal{Q}_i - \frac{\tau}{2}\mathcal{R}_i\Big) w_{i,n-1} + \frac{\tau}{2}\mathcal{Q}_i\Big(f(x_i,t_n,u(x_i,t_n)) - f(x_i,t_n,u_{i,n})\Big) \\ \qquad + \frac{\tau}{2}\mathcal{Q}_i\Big(f(x_i,t_{n-1},u(x_i,t_{n-1})) - f(x_i,t_{n-1},u_{i,n-1})\Big) + \widetilde{\varepsilon}_{i,n} + \widehat{\varepsilon}_{i,n}, \\ \qquad\qquad 1 \le i \le L-1;\ \ 1 \le n \le N, \\ w_{0,n} = 0,\ \ w_{L,n} = 0, \qquad 1 \le n \le N, \\ w_{i,0} = 0, \qquad 0 \le i \le L. \end{cases} \tag{4.8}$$

Now let the matrices Q and R be the same as before, and

$$W_n = \left(w_{1,n}, \cdots, w_{L-1,n}\right)^T, \ \ \mathcal{E}_n = \left(\widetilde{\varepsilon}_{1,n} + \widehat{\varepsilon}_{1,n}, \cdots, \widetilde{\varepsilon}_{L-1,n} + \widehat{\varepsilon}_{L-1,n}\right)^T.$$

Then the first equality in (4.8) can be written as the following matrix form:

$$\Big(Q + \frac{\tau}{2}R\Big) W_n = \Big(Q - \frac{\tau}{2}R\Big) W_{n-1} + \frac{\tau}{2}Q\left(V_n W_n + V_{n-1}W_{n-1}\right) + \mathcal{E}_n, \tag{4.9}$$

where

$$V_n = \operatorname{diag}\left(f_u(x_1,t_n,\theta_{1,n}), \cdots, f_u(x_{L-1},t_n,\theta_{L-1,n})\right)$$

and $\theta_{i,n}$ lies between $u(x_i,t_n)$ and $u_{i,n}$. Since the condition (2.8) holds with $M_{i,n} \equiv 0$, we have from Lemma 2.3 that

$$\Big(Q + \frac{\tau}{2}R\Big)^{-1} \ge 0, \qquad \Big(Q - \frac{\tau}{2}R\Big) \ge 0.$$

Next, let $Y=(1,1,\cdots,1)^T$ and define

$$S^{(1)}=\left(Q+\frac{\tau}{2}R\right)Y,\quad S^{(2)}=\left(Q+\frac{\tau}{2}R\right)^{-1}\left(Q-\frac{\tau}{2}R\right)Y,\quad S^{(3)}=\left(Q+\frac{\tau}{2}R\right)^{-1}QY.$$

Denoting by $S_i^{(l)}$ the i-th component of $S^{(l)}$, $l=1,2,3$, we have

$$S_i^{(1)}\geq\frac{\alpha_0 h}{\alpha_1},\qquad 1\leq i\leq L-1,$$

which implies

$$\left\|\left(Q+\frac{\tau}{2}R\right)^{-1}\right\|_\infty\leq\frac{\alpha_1}{\alpha_0 h}.$$

Since

$$\left(Q+\frac{\tau}{2}R\right)(S^{(2)}-Y)=-\tau RY\leq 0,$$

we obtain from the nonnegativity of $\left(Q+\frac{\tau}{2}R\right)^{-1}$ that $S^{(2)}\leq Y$. This leads to

$$\left\|\left(Q+\frac{\tau}{2}R\right)^{-1}\left(Q-\frac{\tau}{2}R\right)\right\|_\infty=\max_i S_i^{(2)}\leq 1.$$

Similarly,

$$\left\|\left(Q+\frac{\tau}{2}R\right)^{-1}Q\right\|_\infty\leq 1.$$

Also, it can be shown that

$$\|V_n\|_\infty\leq\overline{M},\quad \|V_{n-1}\|_\infty\leq\overline{M},\quad \|\mathcal{E}_n\|_\infty\leq C_4(\tau^3h+\tau h^5),\qquad 1\leq n\leq N.$$

By the above estimates and (4.9), we arrive at

$$\|W_n\|_\infty\leq\frac{\tau}{2}\overline{M}\|W_n\|_\infty+\left(1+\frac{\tau}{2}\overline{M}\right)\|W_{n-1}\|_\infty+\frac{\alpha_1C_4}{\alpha_0}\tau(\tau^2+h^4).$$

Since $\|W_0\|_\infty=0$ and $\tau\overline{M}<1$, an application of Lemma 4.1 gives

$$\begin{aligned}\|W_n\|_\infty&\leq\left(e^{\frac{T\overline{M}}{1-\tau\overline{M}/2}}-1\right)\frac{\alpha_1C_4}{\alpha_0\overline{M}}\left(\tau^2+h^4\right)\\&\leq\left(e^{2T\overline{M}}-1\right)\frac{\alpha_1C_4}{\alpha_0\overline{M}}\left(\tau^2+h^4\right).\end{aligned}$$

Then the desired result (4.4) follows. □

5. An Accelerated Monotone Iterative Scheme

The Picard iteration (3.4) gives an algorithm for resolving the solution of (2.7). However, its convergence rate is only of linear order as pointed out before. To raise the convergence rate while maintaining the monotone property, we extend the accelerated monotone iteration in [20] for (2.7). Its convergence rate is either quadratic or nearly quadratic with the usual differentiability requirement only on the function $f(\cdot,u)$. If the reaction function possesses a concavity or convexity property between upper and lower solutions, then this iteration is reduced to the Newton's method.

5.1. Monotone iteration

Let $\overline{u}_{i,n}$ and $\underline{u}_{i,n}$ be a pair of ordered upper and lower solutions of (2.7). It follows from Theorem 3.2 that (2.7) has a unique solution $u^*_{i,n}$ in $\langle \underline{u}_{i,n}, \overline{u}_{i,n}\rangle$ under the conditions of the theorem. To compute the solution $u^*_{i,n}$ we use the following iteration:

$$\begin{cases} \left(\mathcal{Q}_i + \frac{\tau}{2}\mathcal{R}_i\right) u^{(m)}_{i,n} + \frac{\tau}{2}\mathcal{Q}_i\left(M^{(m-1)}_{i,n} u^{(m)}_{i,n}\right) = \left(\mathcal{Q}_i - \frac{\tau}{2}\mathcal{R}_i\right) u^*_{i,n-1} \\ \quad + \frac{\tau}{2}\mathcal{Q}_i\left(M^{(m-1)}_{i,n} u^{(m-1)}_{i,n}\right) + \frac{\tau}{2}\mathcal{Q}_i\left(f(x_i,t_n,u^{(m-1)}_{i,n}) + f(x_i,t_{n-1},u^*_{i,n-1})\right), \\ \qquad\qquad 1 \le i \le L-1;\quad 1 \le n \le N, \\ u^{(m)}_{0,n} = g_0(t_n),\quad u^{(m)}_{L,n} = g_1(t_n), \qquad 1 \le n \le N, \\ u^{(m)}_{i,0} = u_0(x_i), \qquad 0 \le i \le L, \end{cases} \tag{5.1}$$

where $u^{(0)}_{i,n}$ is either $\overline{u}_{i,n}$ or $\underline{u}_{i,n}$, and

$$M^{(m)}_{i,n} = \max\left\{-f_u(x_i,t_n,\xi_{i,n}) : \xi_{i,n} \in \langle \underline{u}^{(m)}_{i,n}, \overline{u}^{(m)}_{i,n}\rangle\right\}. \tag{5.2}$$

The functions $\overline{u}^{(m)}_{i,n}$, $\underline{u}^{(m)}_{i,n}$ in the definition of $M^{(m)}_{i,n}$ are obtained from (5.1) with $u^{(0)}_{i,n} = \overline{u}_{i,n}$ and $u^{(0)}_{i,n} = \underline{u}_{i,n}$, respectively. By the definition of $M^{(m)}_{i,n}$, it is clear that if $f(\cdot,u)$ is a C^2-function then

$$M^{(m)}_{i,n} = \begin{cases} -f_u(x_i,t_n,\underline{u}^{(m)}_{i,n}), & \text{if } f_{uu}(x_i,t_n,\xi_{i,n}) \ge 0,\ \ \forall\, \xi_{i,n} \in \langle \underline{u}^{(m)}_{i,n}, \overline{u}^{(m)}_{i,n}\rangle, \\ -f_u(x_i,t_n,\overline{u}^{(m)}_{i,n}), & \text{if } f_{uu}(x_i,t_n,\xi_{i,n}) \le 0,\ \ \forall\, \xi_{i,n} \in \langle \underline{u}^{(m)}_{i,n}, \overline{u}^{(m)}_{i,n}\rangle. \end{cases} \tag{5.3}$$

Hence if $f_u(\cdot,u)$ is either monotone nondecreasing or monotone nonincreasing in u then the iteration (5.1) is reduced to the Newton's iteration

$$\begin{cases} \left(\mathcal{Q}_i + \frac{\tau}{2}\mathcal{R}_i\right) u^{(m)}_{i,n} - \frac{\tau}{2}\mathcal{Q}_i\left(f_u(x_i,t_n,u^{(m-1)}_{i,n}) u^{(m)}_{i,n}\right) = \left(\mathcal{Q}_i - \frac{\tau}{2}\mathcal{R}_i\right) u^*_{i,n-1} \\ \quad - \frac{\tau}{2}\mathcal{Q}_i\left(f_u(x_i,t_n,u^{(m-1)}_{i,n}) u^{(m-1)}_{i,n}\right) + \frac{\tau}{2}\mathcal{Q}_i\left(f(x_i,t_n,u^{(m-1)}_{i,n})\right. \\ \quad \left. + f(x_i,t_{n-1},u^*_{i,n-1})\right), \qquad 1 \le i \le L-1;\quad 1 \le n \le N, \\ u^{(m)}_{0,n} = g_0(t_n),\quad u^{(m)}_{L,n} = g_1(t_n), \qquad 1 \le n \le N, \\ u^{(m)}_{i,0} = u_0(x_i), \qquad 0 \le i \le L. \end{cases} \tag{5.4}$$

To show that the sequence given by (5.1) is well-defined, it is crucial that the sequences $\{\overline{u}^{(m)}_{i,n}\}$ and $\{\underline{u}^{(m)}_{i,n}\}$ possess the property $\overline{u}^{(m)}_{i,n} \ge \underline{u}^{(m)}_{i,n}$ for every m. This is given in the following lemma.

Lemma 5.1. *Let $\overline{u}_{i,n}$, $\underline{u}_{i,n}$ be a pair of ordered upper and lower solutions of* (2.7)*, and let $\overline{M}_{i,n}$ be some constants such that $\tau\overline{M}_{i,n} < 2$ and*

$$f_u(x_i,t_n,\xi_{i,n}) \le \overline{M}_{i,n}, \qquad \xi_{i,n} \in \langle \underline{u}_{i,n}, \overline{u}_{i,n}\rangle,\ \ 0 \le i \le L;\ 0 \le n \le N.$$

Assume that $3\alpha_1 \le 10\alpha_0$ and the condition (2.8) *holds with $M_{i,n} = M^{(0)}_{i,n}$, where $M^{(0)}_{i,n}$ is defined by* (5.2) *with $\underline{u}^{(0)}_{i,n} = \underline{u}_{i,n}$ and $\overline{u}^{(0)}_{i,n} = \overline{u}_{i,n}$. Then the sequences $\{\overline{u}^{(m)}_{i,n}\}$, $\{\underline{u}^{(m)}_{i,n}\}$ and $\{M^{(m)}_{i,n}\}$ given by* (5.1) *and* (5.2) *are all well-defined and possess the monotone property*

$$\underline{u}_{i,n} \le \underline{u}^{(m)}_{i,n} \le \underline{u}^{(m+1)}_{i,n} \le \overline{u}^{(m+1)}_{i,n} \le \overline{u}^{(m)}_{i,n} \le \overline{u}_{i,n}, \quad 0 \le i \le L;\ 0 \le n \le N;\ m \ge 0. \tag{5.5}$$

Proof. Since the condition (2.8) holds with $M_{i,n} = M_{i,n}^{(0)}$, we have from Lemma 2.3 that $\overline{u}_{i,n}^{(1)}$ and $\underline{u}_{i,n}^{(1)}$ exist and satisfy

$$\left(\mathcal{Q}_i + \frac{\tau}{2}\mathcal{R}_i\right)\left(\overline{u}_{i,n}^{(1)} - \underline{u}_{i,n}^{(1)}\right) + \frac{\tau}{2}\mathcal{Q}_i\left(M_{i,n}^{(0)}\left(\overline{u}_{i,n}^{(1)} - \underline{u}_{i,n}^{(1)}\right)\right) \geq 0.$$

Writing the above relation in the matrix form and using Lemma 2.3, we obtain $\overline{u}_{i,n}^{(1)} \geq \underline{u}_{i,n}^{(1)}$ for all $0 \leq i \leq L$ and $0 \leq n \leq N$. Therefore, $M_{i,n}^{(1)}$ is well-defined. Also we have

$$\begin{aligned}&\left(\mathcal{Q}_i + \frac{\tau}{2}\mathcal{R}_i\right)\left(\overline{u}_{i,n} - \overline{u}_{i,n}^{(1)}\right) + \frac{\tau}{2}\mathcal{Q}_i\left(M_{i,n}^{(0)}\left(\overline{u}_{i,n} - \overline{u}_{i,n}^{(1)}\right)\right)\\ \geq &\left(\mathcal{Q}_i - \frac{\tau}{2}\mathcal{R}_i\right)\left(\overline{u}_{i,n-1} - u_{i,n-1}^{*}\right) - \frac{\tau}{2}\mathcal{Q}_i\left(M_{i,n-1}^{(0)}\left(\overline{u}_{i,n-1} - u_{i,n-1}^{*}\right)\right).\end{aligned}$$

Again by Lemma 2.3, $\overline{u}_{i,n} \geq \overline{u}_{i,n}^{(1)}$ for all $0 \leq i \leq L$ and $0 \leq n \leq N$. Similarly, $\underline{u}_{i,n} \leq \underline{u}_{i,n}^{(1)}$ for all $0 \leq i \leq L$ and $0 \leq n \leq N$. The above results imply that (5.5) holds for $m = 0$. Now, assume that for some $m \geq 1$ and all $1 \leq k \leq m$, $\overline{u}_{i,n}^{(k)}$, $\underline{u}_{i,n}^{(k)}$ and $M_{i,n}^{(k)}$ are all well-defined and

$$\underline{u}_{i,n}^{(k-1)} \leq \underline{u}_{i,n}^{(k)} \leq \overline{u}_{i,n}^{(k)} \leq \overline{u}_{i,n}^{(k-1)}, \qquad 0 \leq i \leq L;\ 0 \leq n \leq N;\ 1 \leq k \leq m. \tag{5.6}$$

Clearly, by the definition of $M_{i,n}^{(m)}$ and (5.6),

$$-\overline{M}_{i,n} \leq M_{i,n}^{(m)} \leq M_{i,n}^{(0)}.$$

This implies that the condition (2.8) holds for $M_{i,n} = M_{i,n}^{(m)}$. Therefore by Lemma 2.3, $\overline{u}_{i,n}^{(m+1)}$ and $\underline{u}_{i,n}^{(m+1)}$ are well-defined and

$$\left(\mathcal{Q}_i + \frac{\tau}{2}\mathcal{R}_i\right)\left(\overline{u}_{i,n}^{(m+1)} - \underline{u}_{i,n}^{(m+1)}\right) + \frac{\tau}{2}\mathcal{Q}_i\left(M_{i,n}^{(m)}\left(\overline{u}_{i,n}^{(m+1)} - \underline{u}_{i,n}^{(m+1)}\right)\right) \geq 0.$$

It follows from Lemma 2.3 that $\overline{u}_{i,n}^{(m+1)} \geq \underline{u}_{i,n}^{(m+1)}$ for all $0 \leq i \leq L$ and $0 \leq n \leq N$. This ensures also that $M_{i,n}^{(m+1)}$ is well-defined. By the iteration (5.1),

$$\begin{aligned}&\left(\mathcal{Q}_i + \frac{\tau}{2}\mathcal{R}_i\right)\left(\overline{u}_{i,n}^{(m)} - \overline{u}_{i,n}^{(m+1)}\right) + \frac{\tau}{2}\mathcal{Q}_i\left(M_{i,n}^{(m)}\left(\overline{u}_{i,n}^{(m)} - \overline{u}_{i,n}^{(m+1)}\right)\right)\\ = &\frac{\tau}{2}\mathcal{Q}_i\left(M_{i,n}^{(m-1)}\left(\overline{u}_{i,n}^{(m-1)} - \overline{u}_{i,n}^{(m)}\right) + f(x_i, t_n, \overline{u}_{i,n}^{(m-1)}) - f(x_i, t_n, \overline{u}_{i,n}^{(m)})\right).\end{aligned}$$

Thanks to (5.6) and the definition of $M_{i,n}^{(m-1)}$,

$$\left(\mathcal{Q}_i + \frac{\tau}{2}\mathcal{R}_i\right)\left(\overline{u}_{i,n}^{(m)} - \overline{u}_{i,n}^{(m+1)}\right) + \frac{\tau}{2}\mathcal{Q}_i\left(M_{i,n}^{(m)}\left(\overline{u}_{i,n}^{(m)} - \overline{u}_{i,n}^{(m+1)}\right)\right) \geq 0.$$

By Lemma 2.3, $\overline{u}_{i,n}^{(m)} \geq \overline{u}_{i,n}^{(m+1)}$ for all $0 \leq i \leq L$ and $0 \leq n \leq N$. Similarly, $\underline{u}_{i,n}^{(m+1)} \geq \underline{u}_{i,n}^{(m)}$ for all $0 \leq i \leq L$ and $0 \leq n \leq N$. So the conclusion of the lemma follows from the induction. □

In view of the monotone property (5.5), there exist limits

$$\lim_{m\to\infty} \overline{u}_{i,n}^{(m)} = \overline{u}_{i,n}^{*}, \qquad \lim_{m\to\infty} \underline{u}_{i,n}^{(m)} = \underline{u}_{i,n}^{*}. \tag{5.7}$$

By the monotone property of $M_{i,n}^{(m)}$ and

$$-\overline{M}_{i,n} \leq M_{i,n}^{(m)} \leq M_{i,n}^{(0)},$$

the sequences $\{M_{i,n}^{(m)}\}$ is also convergent as $m \to \infty$. Letting $m \to \infty$ in (5.1), we deduce that

$$\begin{cases} \Big(\mathcal{Q}_i + \frac{\tau}{2}\mathcal{R}_i\Big) v^*_{i,n} = \Big(\mathcal{Q}_i - \frac{\tau}{2}\mathcal{R}_i\Big) u^*_{i,n-1} + \frac{\tau}{2}\mathcal{Q}_i \left(f(x_i,t_n,v^*_{i,n}) + f(x_i,t_{n-1},u^*_{i,n-1}\right), \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad 1 \le i \le L-1;\ \ 1 \le n \le N, \\ v^*_{0,n} = g_0(t_n), \ \ v^*_{L,n} = g_1(t_n), \qquad 1 \le n \le N, \\ v^*_{i,0} = u_0(x_i), \qquad 0 \le i \le L, \end{cases} \tag{5.8}$$

where $v^*_{i,n} = \overline{u}^*_{i,n}$ or $\underline{u}^*_{i,n}$. This leads to the following conclusion.

Theorem 5.1. *Let the conditions in Lemma* 5.1 *hold. Then the sequences* $\{\overline{u}_{i,n}^{(m)}\}$ *and* $\{\underline{u}_{i,n}^{(m)}\}$ *given by* (5.1) *with* $\overline{u}_{i,n}^{(0)} = \overline{u}_{i,n}$ *and* $\underline{u}_{i,n}^{(0)} = \underline{u}_{i,n}$, *converge monotonically from above and below, respectively, to the unique solution* $u^*_{i,n}$ *of* (2.7) *in* $\langle \underline{u}_{i,n}, \overline{u}_{i,n}\rangle$, *and*

$$\underline{u}_{i,n} \le \underline{u}_{i,n}^{(m)} \le \underline{u}_{i,n}^{(m+1)} \le u^*_{i,n} \le \overline{u}_{i,n}^{(m+1)} \le \overline{u}_{i,n}^{(m)} \le \overline{u}_{i,n}, \\ 0 \le i \le L;\ \ 0 \le n \le N;\ \ m \ge 0. \tag{5.9}$$

Proof. Let $\overline{u}^*_{i,n}$ and $\underline{u}^*_{i,n}$ be the limits in (5.7). To complete the proof of the theorem we only need to show $\overline{u}^*_{i,n} = \underline{u}^*_{i,n} = u^*_{i,n}$ for all $0 \le i \le L$ and $0 \le n \le N$. Let $w^*_{i,n} = v^*_{i,n} - u^*_{i,n}$ where $v^*_{i,n}$ is either $\overline{u}^*_{i,n}$ or $\underline{u}^*_{i,n}$. Subtracting (2.7) from (5.8) and using the mean-value theorem lead to

$$\Big(\mathcal{Q}_i + \frac{\tau}{2}\mathcal{R}_i\Big) w^*_{i,n} - \frac{\tau}{2}\mathcal{Q}_i \left(f_u(x_i,t_n,\xi^*_{i,n}) w^*_{i,n}\right) = 0,$$

where $\xi^*_{i,n}$ is an intermediate value between $v^*_{i,n}$ and $u^*_{i,n}$. Since

$$-\overline{M}_{i,n} \le -f_u(x_i,t_n,\xi^*_{i,n}) \le M_{i,n}^{(0)},$$

the condition (2.8) holds for $M_{i,n} = -f_u(x_i,t_n,\xi^*_{i,n})$. Hence by Lemma 2.3, $w^*_{i,n} = 0$ which implies $\overline{u}^*_{i,n} = \underline{u}^*_{i,n} = u^*_{i,n}$ for all $0 \le i \le L$ and $0 \le n \le N$. $\square$

When $f_u(\cdot,u)$ is monotone nondecreasing or monotone nonincreasing in u the iteration (5.1) is reduced to the Newton's iteration (5.4). As a consequences of Theorem 5.1 we have the following result.

Corollary 5.1. *Let the conditions in Lemma* 5.1 *hold, and assume that* $f(\cdot,u)$ *is a* C^2*-function of* u. *Then the sequence* $\{\overline{u}_{i,n}^{(m)}\}$ *given by* (5.4) *with* $\overline{u}_{i,n}^{(0)} = \overline{u}_{i,n}$ *converges monotonically from above to the unique solution* $u^*_{i,n}$ *of* (2.7) *in* $\langle \underline{u}_{i,n}, \overline{u}_{i,n}\rangle$ *if* $f_{uu}(x_i,t_n,\xi_{i,n}) \le 0$ *for all* $\xi_{i,n} \in \langle \underline{u}_{i,n}, \overline{u}_{i,n}\rangle$. *Similarly, the sequence* $\{\underline{u}_{i,n}^{(m)}\}$ *given by* (5.4) *with* $\underline{u}_{i,n}^{(0)} = \underline{u}_{i,n}$ *converges monotonically from below to* $u^*_{i,n}$ *if* $f_{uu}(x_i,t_n,\xi_{i,n}) \ge 0$ *for all* $\xi_{i,n} \in \langle \underline{u}_{i,n}, \overline{u}_{i,n}\rangle$.

Remark 5.1. The procedure of creating the sequence $\{u_{i,n}^{(m)}\}$ by the iteration (5.1) is described as follows: Starting from $u_{i,1}^{(0)}$, which is either $\overline{u}_{i,1}$ or $\underline{u}_{i,1}$, we compute $u_{i,1}^{(1)}$ from (5.1) where $u^*_{i,0} = u_0(x_i)$ and $M_{i,1}^{(0)}$ is given by (5.2). Repeating this process leads to the sequence $\{u_{i,1}^{(m)}\}$. Using $u^*_{i,1} = v^*_{i,1}$ where $v^*_{i,1}$ is the limit of the sequence $\{u_{i,1}^{(m)}\}$ (this is ensured by Theorem 5.1), the same process gives the sequence $\{u_{i,2}^{(m)}\}$. Continuing this process yields the sequence $\{u_{i,n}^{(m)}\}$ for all $0 \le i \le L$, $0 \le n \le N$ and $m \ge 1$.

5.2. Convergence rate

In this subsection we show that the sequences given by (5.1) and (5.4) possess either a quadratic or a nearly quadratic convergence rate.

Theorem 5.2. *Let the conditions in Lemma* 5.1 *hold, and let* $\{\overline{u}_{i,n}^{(m)}\}$ *and* $\{\underline{u}_{i,n}^{(m)}\}$ *be the sequences given by* (5.1) *with* $\overline{u}_{i,n}^{(0)}=\overline{u}_{i,n}$ *and* $\underline{u}_{i,n}^{(0)}=\underline{u}_{i,n}$. *Let also* $u_{i,n}^*$ *be the unique solution of* (2.7) *in* $\langle\underline{u}_{i,n},\overline{u}_{i,n}\rangle$. *Then there exists a positive constant* ρ_n, *independent of* m, *such that for all* $0\le i\le L$, $0\le n\le N$ *and* $m\ge 2$,

$$\begin{aligned}\max_i\left|\overline{u}_{i,n}^{(m)}-u_{i,n}^*\right|&\le\rho_n\max_i\left|\overline{u}_{i,n}^{(m-1)}-\underline{u}_{i,n}^{(m-1)}\right|\cdot\max_i\left|\overline{u}_{i,n}^{(m-1)}-u_{i,n}^*\right|,\\ \max_i\left|\underline{u}_{i,n}^{(m)}-u_{i,n}^*\right|&\le\rho_n\max_i\left|\overline{u}_{i,n}^{(m-1)}-\underline{u}_{i,n}^{(m-1)}\right|\cdot\max_i\left|\underline{u}_{i,n}^{(m-1)}-u_{i,n}^*\right|,\end{aligned}\tag{5.10}$$

and

$$\begin{aligned}&\max_i\left|\overline{u}_{i,n}^{(m)}-u_{i,n}^*\right|+\max_i\left|\underline{u}_{i,n}^{(m)}-u_{i,n}^*\right|\\ &\le\rho_n\left(\max_i\left|\overline{u}_{i,n}^{(m-1)}-u_{i,n}^*\right|+\max_i\left|\underline{u}_{i,n}^{(m-1)}-u_{i,n}^*\right|\right)^2.\end{aligned}\tag{5.11}$$

Proof. Consider the sequence $\{\overline{u}_{i,n}^{(m)}\}$ and let $\overline{w}_{i,n}^{(m)}=\overline{u}_{i,n}^{(m)}-u_{i,n}^*$. We have $\overline{w}_{i,n}^{(m)}\ge 0$ for all $0\le i\le L$, $0\le n\le N$ and $m\ge 1$. Subtracting (2.7) from (5.1) gives

$$\begin{aligned}&\left(\mathcal{Q}_i+\frac{\tau}{2}\mathcal{R}_i\right)\overline{w}_{i,n}^{(m)}+\frac{\tau}{2}\mathcal{Q}_i\left(M_{i,n}^{(m-1)}\overline{w}_{i,n}^{(m)}\right)\\ &=\frac{\tau}{2}\mathcal{Q}_i\left(M_{i,n}^{(m-1)}\overline{w}_{i,n}^{(m-1)}\right)+\frac{\tau}{2}\mathcal{Q}_i\left(f(x_i,t_n,\overline{u}_{i,n}^{(m-1)})-f(x_i,t_n,u_{i,n}^*)\right).\end{aligned}\tag{5.12}$$

By the mean-value theorem, there exist $\xi_{i,n}^{(m-1)}$ in $\langle\underline{u}_{i,n}^{(m-1)},\overline{u}_{i,n}^{(m-1)}\rangle$ and $\eta_{i,n}^{(m-1)}$ in $\langle u_{i,n}^*,\overline{u}_{i,n}^{(m-1)}\rangle$ such that

$$\begin{aligned}&M_{i,n}^{(m-1)}=-f_u(x_i,t_n,\xi_{i,n}^{(m-1)}),\\ &f(x_i,t_n,\overline{u}_{i,n}^{(m-1)})-f(x_i,t_n,u_{i,n}^*)=f_u(x_i,t_n,\eta_{i,n}^{(m-1)})\overline{w}_{i,n}^{(m-1)}.\end{aligned}$$

Again by the mean-value theorem, there exists an intermediate value $\theta_{i,n}^{(m-1)}$ between $\eta_{i,n}^{(m-1)}$ and $\xi_{i,n}^{(m-1)}$ such that

$$f_u(x_i,t_n,\eta_{i,n}^{(m-1)})-f_u(x_i,t_n,\xi_{i,n}^{(m-1)})=f_{uu}(x_i,t_n,\theta_{i,n}^{(m-1)})(\eta_{i,n}^{(m-1)}-\xi_{i,n}^{(m-1)}).$$

Using the above relations we obtain from (5.12) that

$$\begin{aligned}&\left(\mathcal{Q}_i+\frac{\tau}{2}\mathcal{R}_i\right)\overline{w}_{i,n}^{(m)}+\frac{\tau}{2}\mathcal{Q}_i\left(M_{i,n}^{(m-1)}\overline{w}_{i,n}^{(m)}\right)\\ &=\frac{\tau}{2}\mathcal{Q}_i\left(f_{uu}(x_i,t_n,\theta_{i,n}^{(m-1)})(\eta_{i,n}^{(m-1)}-\xi_{i,n}^{(m-1)})\overline{w}_{i,n}^{(m-1)}\right).\end{aligned}\tag{5.13}$$

Let $M_n=\max_i\{|f_{uu}(x_i,t_n,\xi_{i,n})|:\xi_{i,n}\in\langle\underline{u}_{i,n},\overline{u}_{i,n}\rangle\}$. Then

$$\left|f_{uu}(x_i,t_n,\theta_{i,n}^{(m-1)})(\eta_{i,n}^{(m-1)}-\xi_{i,n}^{(m-1)})\right|\le M_n\max_i\left|\overline{u}_{i,n}^{(m-1)}-\underline{u}_{i,n}^{(m-1)}\right|.$$

It follows from (5.13) that

$$\begin{aligned}&\left(\mathcal{Q}_i+\frac{\tau}{2}\mathcal{R}_i\right)\overline{w}_{i,n}^{(m)}+\frac{\tau}{2}\mathcal{Q}_i\left(M_{i,n}^{(m-1)}\overline{w}_{i,n}^{(m)}\right)\\&\le\frac{\tau}{2}M_n\max_i\left|\overline{u}_{i,n}^{(m-1)}-\underline{u}_{i,n}^{(m-1)}\right|\mathcal{Q}_i\left(\overline{w}_{i,n}^{(m-1)}\right).\end{aligned}\tag{5.14}$$

Let the matrices Q and R be the same as before, and let $\overline{W}_n^{(m)}=(\overline{w}_{1,n}^{(m)},\cdots,\overline{w}_{L-1,n}^{(m)})^T$. Then for $m\ge 2$, (5.14) can be written in the matrix form

$$\left(Q+\frac{\tau}{2}QD_n^{(m-1)}+\frac{\tau}{2}R\right)\overline{W}_n^{(m)}\le\frac{\tau}{2}M_n\max_i\left|\overline{u}_{i,n}^{(m-1)}-\underline{u}_{i,n}^{(m-1)}\right|Q\overline{W}_n^{(m-1)},\tag{5.15}$$

where

$$D_n^{(m-1)}=\operatorname{diag}\left(M_{1,n}^{(m-1)},\cdots,M_{L-1,n}^{(m-1)}\right).$$

Since $M_{i,n}^{(m-1)}\ge-\overline{M}_{i,n}$, we have

$$Q+\frac{\tau}{2}QD_n^{(m-1)}+\frac{\tau}{2}R\ge Q+\frac{\tau}{2}Q\overline{D}_n+\frac{\tau}{2}R,$$

where

$$\overline{D}_n=\operatorname{diag}(-\overline{M}_{1,n},\cdots,-\overline{M}_{L-1,n}).$$

Therefore,

$$\left(Q+\frac{\tau}{2}Q\overline{D}_n+\frac{\tau}{2}R\right)\overline{W}_n^{(m)}\le\frac{\tau}{2}M_n\max_i\left|\overline{u}_{i,n}^{(m-1)}-\underline{u}_{i,n}^{(m-1)}\right|Q\overline{W}_n^{(m-1)}.\tag{5.16}$$

Since $-\overline{M}_{i,n}\le M_{i,n}^{(0)}$, the condition (2.8) holds with $M_{i,n}=-\overline{M}_{i,n}$ and so by Lemma 2.3,

$$\left(Q+\frac{\tau}{2}Q\overline{D}_n+\frac{\tau}{2}R\right)^{-1}\ge 0.$$

This implies that

$$\left\|\overline{W}_n^{(m)}\right\|_\infty\le\rho_n\max_i\left|\overline{u}_{i,n}^{(m-1)}-\underline{u}_{i,n}^{(m-1)}\right|\left\|\overline{W}_n^{(m-1)}\right\|_\infty,$$

where

$$\rho_n=\frac{\tau}{2}M_n\left\|\left(Q+\frac{\tau}{2}Q\overline{D}_n+\frac{\tau}{2}R\right)^{-1}\right\|_\infty\|Q\|_\infty.$$

Then the first relation in (5.10) is proved. The second relation in (5.10) can be proved similarly. Putting two results in (5.10) together, we deduce that

$$\begin{aligned}&\max_i\left|\overline{u}_{i,n}^{(m)}-u_{i,n}^*\right|+\max_i\left|\underline{u}_{i,n}^{(m)}-u_{i,n}^*\right|\\&\le\rho_n\max_i\left|\overline{u}_{i,n}^{(m-1)}-\underline{u}_{i,n}^{(m-1)}\right|\left(\max_i\left|\overline{u}_{i,n}^{(m-1)}-u_{i,n}^*\right|+\max_i\left|\underline{u}_{i,n}^{(m-1)}-u_{i,n}^*\right|\right).\end{aligned}\tag{5.17}$$

Obviously,

$$\max_i\left|\overline{u}_{i,n}^{(m-1)}-\underline{u}_{i,n}^{(m-1)}\right|\le\max_i\left|\overline{u}_{i,n}^{(m-1)}-u_{i,n}^*\right|+\max_i\left|\underline{u}_{i,n}^{(m-1)}-u_{i,n}^*\right|.\tag{5.18}$$

Finally, the relation (5.11) follows from (5.17) and (5.18).

Theorem 5.2 gives a nearly quadratic convergence for the sequences $\{\overline{u}_{i,n}^{(m)}\}$ and $\{\underline{u}_{i,n}^{(m)}\}$, and a quadratic convergence for the sum of these two sequences. Moreover, when $f(\cdot,u)$ possesses a concavity or convexity property in $\langle\underline{u}_{i,n},\overline{u}_{i,n}\rangle$, then one of the two sequences has the quadratic convergence. This is shown in the following theorem.

Theorem 5.3. *Let the conditions in Theorem* 5.2 *hold. Then there exists a positive constant* ρ_n, *independent of* m, *such that for all* $0 \le i \le L$, $0 \le n \le N$ *and* $m \ge 2$,

$$\max_i \left|\overline{u}_{i,n}^{(m)} - u_{i,n}^*\right| \le \rho_n \left(\max_i \left|\overline{u}_{i,n}^{(m-1)} - u_{i,n}^*\right|\right)^2, \tag{5.19}$$

if $f_{uu}(x_i, t_n, \xi_{i,n}) \le 0$ *for* $\xi_{i,n} \in \langle \underline{u}_{i,n}, \overline{u}_{i,n}\rangle$, *and*

$$\max_i \left|\underline{u}_{i,n}^{(m)} - u_{i,n}^*\right| \le \rho_n \left(\max_i \left|\underline{u}_{i,n}^{(m-1)} - u_{i,n}^*\right|\right)^2, \tag{5.20}$$

if $f_{uu}(x_i, t_n, \xi_{i,n}) \ge 0$ *for* $\xi_{i,n} \in \langle \underline{u}_{i,n}, \overline{u}_{i,n}\rangle$.

Proof. Consider the case $f_{uu}(x_i, t_n, \xi_{i,n}) \le 0$. By (5.3),

$$M_{i,n}^{(m-1)} = -f_u(x_i, t_n, \overline{u}_{i,n}^{(m-1)}).$$

This implies that $\xi_{i,n}^{(m-1)}$ appearing in (5.13) is given by $\xi_{i,n}^{(m-1)} = \overline{u}_{i,n}^{(m-1)}$. Since $\eta_{i,n}^{(m-1)}$ in (5.13) is in $\langle u_{i.n}^*, \overline{u}_{i,n}^{(m-1)}\rangle$ we have that

$$|\eta_{i,n}^{(m-1)} - \xi_{i,n}^{(m-1)}| \le |\overline{u}_{i,n}^{(m-1)} - u_{i,n}^*|.$$

In this case, the relation (5.14) becomes

$$\begin{aligned} &\left(\mathcal{Q}_i + \frac{\tau}{2}\mathcal{R}_i\right)\overline{w}_{i,n}^{(m)} + \frac{\tau}{2}\mathcal{Q}_i\left(M_{i,n}^{(m-1)}\overline{w}_{i,n}^{(m)}\right) \\ \le\ &\frac{\tau}{2}M_n \max_i \left|\overline{u}_{i,n}^{(m-1)} - u_{i,n}^*\right| \mathcal{Q}_i\left(\overline{w}_{i,n}^{(m-1)}\right). \end{aligned}$$

By the above and an argument as in the proof of Theorem 5.2, we arrive at (5.19). The proof of (5.20) is similar. □

Remark 5.2. Since the initial iteration in the accelerated monotone iteration (5.1) is either an upper or a lower solution, which can be constructed directly from the equation without any knowledge of the solution, this method avoids the search for the initial iteration as is often needed in the Newton's method. This feature is one of the great advantages of this approach.

Remark 5.3. Following the same line as in the derivation of (5.10) and (5.11), we can also prove the quadratic convergence of the sequence for the corresponding steady-state problems (cf. [29]).

6. Numerical Results

In this section, we present some numerical results. They demonstrate the monotonicity of the sequences given by (5.1) and the rapid convergence rate. They also indicate the high accuracy of the scheme (2.7).

Example 1: The chemical reactor model. We first consider a chemical reactor model from chemical engineering (see [1,19]). This model is given by

$$\begin{cases} \dfrac{\partial u}{\partial t} - \dfrac{\partial}{\partial x}\left(e^x \dfrac{\partial u}{\partial x}\right) = \sigma(1-u)e^{-\gamma/(1+u)}, & 0 < x < 1, \quad 0 < t \le T, \\ u(0,t) = u(1,t) = 0, \qquad 0 < t \le T, \\ u(x,0) = u_0(x), \qquad 0 \le x \le 1, \end{cases} \tag{6.1}$$

where σ and γ are given positive constants. For this example, the values J_i, E_i, G_i and F_i in the scheme (2.7) are, for $1 \le i \le L-1$,

$$J_i = \frac{e^{ih}}{e^h - 1}, \qquad E_i = \frac{1}{h} + \frac{1}{e^h - 1}\left(\frac{h}{12} + \frac{5h}{12}e^h - e^h\right),$$
$$F_i = -\frac{2}{h} + \frac{2h}{3} + \frac{e^h + 1}{e^h - 1}, \qquad G_i = \frac{1}{h} - \frac{1}{e^h - 1}\left(1 + \frac{5h}{12} + \frac{h}{12}e^h\right).$$

It is easy to see that $\overline{u}_{i,n} = 2$ and $\underline{u}_{i,n} = 0$ are a pair of ordered upper and lower solutions of (2.7) whenever $0 \le u_0(x) \le 2$ for all $0 \le x \le 1$. Let the initial function $u_0(x) = \sin(\pi x)$, and take the mesh sizes $h = \tau = 0.01$ and the physical parameters $\sigma = 5$, $\gamma = 1$. Using $\overline{u}^{(0)}_{i,n} = 2$ and $\underline{u}^{(0)}_{i,n} = 0$ in the iteration (5.1) we compute the corresponding sequences $\{\overline{u}^{(m)}_{i,n}\}$ and $\{\underline{u}^{(m)}_{i,n}\}$. In numerical computations, the basic feature of the monotone convergence of the sequences $\{\overline{u}^{(m)}_{i,n}\}$ and $\{\underline{u}^{(m)}_{i,n}\}$ were observed for all i and n, see Fig. 6.1, in which we plot the numerical results of these sequences for $n = 50$ and all i.

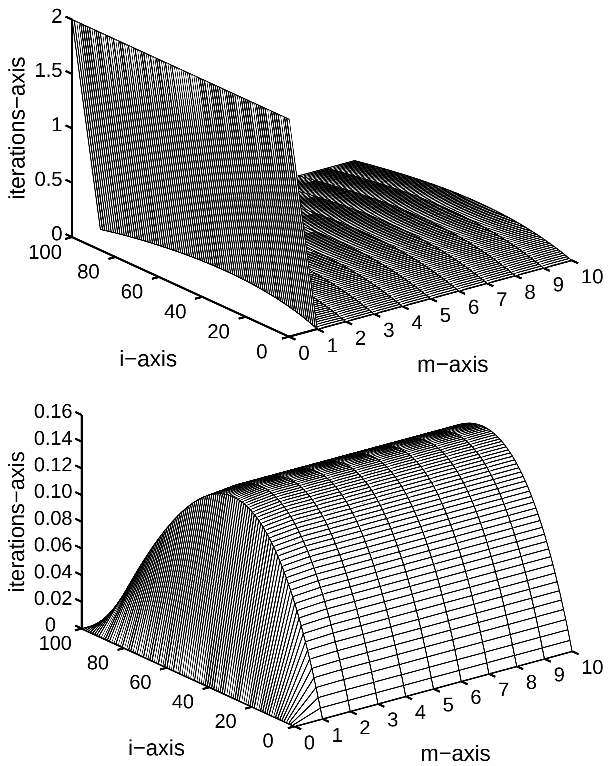

Fig. 6.1. The monotone properties of $\{\overline{u}^{(m)}_{i,50}\}$ (top) and $\{\underline{u}^{(m)}_{i,50}\}$ (bottom).

To demonstrate the monotone convergence of the sequences, the numerical results of these sequences at $(i,n) = (50,50)$ are sketched in Fig. 6.2, in which the solid line and dash-dot line represent the sequences $\{\overline{u}^{(m)}_{50,50}\}$ and $\{\underline{u}^{(m)}_{50,50}\}$, respectively. They coincide with the monotone convergence as described in Theorem 5.1.

We next compare the convergence rate of the accelerated iteration (5.1) with that of the Picard iteration (3.4) in terms of the number of iterations. To compute the number of iterations

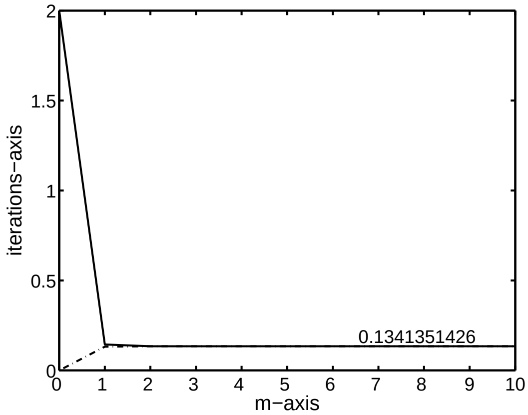

Fig. 6.2. The monotone convergence of $\{\overline{u}^{(m)}_{50,50}\}$ and $\{\underline{u}^{(m)}_{50,50}\}$.

we take the tolerance in the iterations as

$$\max_i |\overline{u}^{(m)}_{i,n} - \underline{u}^{(m)}_{i,n}| \leq 10^{-10}. \tag{6.2}$$

It is known from the chemical engineering literature that the value of σ ranges from 1 up to 10^7 (cf. [1]). With the initial function and the mesh sizes as before we compute the number of iterations for $\gamma = 4$ and various values of σ. In the iteration (3.4) we take $\underline{M}_{i,n} = \sigma$. Our numerical computations show that the number of iterations of the Picard iteration (3.4) is much larger than that of the accelerated iteration (5.1), especially for large σ. The number of iterations for the cases $\sigma = 10^2, 10^3$ and 10^4 are listed in Table 6.1.

Table 6.1: The number of iterations for various values of σ.

	$\sigma = 10^2$		$\sigma = 10^3$		$\sigma = 10^4$	
n	Acceler.	Picard	Acceler.	Picard	Acceler.	Picard
1	4	24	6	99	8	165
2	4	24	6	97	9	254
3	4	25	6	97	8	176
4	4	25	6	97	9	218
5	4	26	6	97	8	183
6	4	26	6	96	8	204
7	4	26	6	97	8	187
8	4	27	6	97	8	198
9	4	27	6	96	8	190
10	4	26	6	97	8	195
11	4	27	6	97	8	191
12	4	27	6	96	8	193
13	4	27	6	97	8	191
14	4	27	6	97	8	192
15	4	27	6	96	8	191
≥ 16	4	27 or 28	6	97 or 96	8	192

Example 2: The logistic equation. To show the high accuracy of numerical solution we consider the following logistic equation (see [18]):

$$\begin{cases} \dfrac{\partial u}{\partial t} - D\dfrac{\partial^2 u}{\partial x^2} = \sigma u(1-u) + q(x,t), \qquad 0 < x < 1, \quad 0 < t \leq T, \\ u(0,t) = u(1,t) = 0, \qquad 0 < t \leq T, \\ u(x,0) = u_0(x), \qquad 0 \leq x \leq 1, \end{cases} \tag{6.3}$$

where D and σ are positive constants, and $q(x,t)$ is a given function. The introduction of a source term q in (6.3) is to construct an analytical solution which is used to compare with numerical solution. For this example, the values J_i, E_i, G_i and F_i in the scheme (2.7) are given by

$$J_i = \frac{D}{h}, \quad E_i = \frac{h}{12}, \quad F_i = \frac{5h}{6}, \quad G_i = \frac{h}{12}, \qquad 1 \le i \le L-1.$$

Take the physical parameters $D = 5$ and $\sigma = 1$. Let the initial function be $u_0(x) = 0$. It is easy to check that when

$$q(x,t) = (5\pi^2 - 2)w(x,t) + w^2(x,t) + \sin(\pi x), \qquad w(x,t) = (1 - \mathrm{e}^{-t})\sin(\pi x),$$

the solution of (6.3) is given by $u(x,t) = w(x,t)$. Since $0 \le q(x,t) \le 5\pi^2$, the pair $\overline{u}_{i,n} = 8$ and $\underline{u}_{i,n} = 0$ form a pair of ordered upper and lower solutions of (2.7). Taking $\overline{u}_{i,n}^{(0)} = 8$ and $\underline{u}_{i,n}^{(0)} = 0$ in the iteration (5.1) we compute the corresponding sequences $\{\overline{u}_{i,n}^{(m)}\}$ and $\{\underline{u}_{i,n}^{(m)}\}$. Numerical experiments show that these sequences possess the same monotone convergence as described in Theorem 5.1. To demonstrate the accuracy of scheme (2.7), we calculate the maximum numerical error $\mathsf{error}_h(t_n)$ and the convergence order $\mathsf{order}_h(t_n)$ of the computed solution $u_{i,n}$, which are defined by

$$\mathsf{error}_h(t_n) = \max_i |u(x_i,t_n) - u_{i,n}|, \qquad \mathsf{order}_h(t_n) = \log_2\left(\frac{\mathsf{error}_h(t_n)}{\mathsf{error}_{h/2}(t_n)}\right), \tag{6.4}$$

where $u(x_i,t_n)$ denotes the value of the analytical solution at (x_i,t_n). In Tables 6.2 and 6.3, we list $\mathsf{error}_h(t_n)$ and $\mathsf{order}_h(t_n)$ for $t_n = 0.125$ and $t_n = 0.25$ respectively, where the computed solution $u_{i,n}$ is given by the iteration (5.1) with the tolerance $\varepsilon = 10^{-12}$ and the calculation is carried out with $\tau = h^2$. We see that the computed solution has the fourth-order accuracy. This is in good agreement with the theoretical prediction.

Table 6.2: The comparison between the schemes (2.7) and (6.5) at $t_n = 0.125$.

	Scheme (2.7)			Scheme (6.5)		
h	$\mathsf{error}_h(0.125)$	$\mathsf{order}_h(0.125)$	CPU time (s)	$\mathsf{error}_h(0.125)$	$\mathsf{order}_h(0.125)$	CPU time (s)
1/8	1.041e-05	4.0062	0.1400	1.155e-03	2.0063	0.0320
1/16	6.479e-07	4.0016	0.2970	2.875e-04	2.0012	0.1720
1/32	4.045e-08	4.0004	1.8590	7.182e-05	2.0003	1.2660
1/64	2.527e-09	4.0001	18.0780	1.795e-05	2.0000	12.4690
1/128	1.579e-10	4.0000	263.1100	4.488e-06	2.0000	161.8130
1/256	9.871e-12		5965.9060	1.122e-06		2965.7970

Table 6.3: The comparison between the schemes (2.7) and (6.5) at $t_n = 0.25$.

	Scheme (2.7)			Scheme (6.5)		
h	$\mathsf{error}_h(0.25)$	$\mathsf{order}_h(0.25)$	CPU time (s)	$\mathsf{error}_h(0.25)$	$\mathsf{order}_h(0.25)$	CPU time (s)
1/8	2.103e-05	4.0065	0.1720	2.554e-03	2.0059	0.0790
1/16	1.308e-06	4.0016	0.6880	6.359e-04	2.0015	0.4210
1/32	8.167e-08	4.0004	4.5630	1.588e-04	2.0004	2.7500
1/64	5.103e-09	4.0001	49.6410	3.969e-05	2.0001	24.6410
1/128	3.189e-10	4.0000	740.9540	9.922e-06	2.0000	364.6090
1/256	1.993e-11		19172.8430	2.480e-06		7107.2030

For comparison, we also use the following standard finite difference scheme to solve (6.3) (see [17,18,20,21]):

$$\begin{cases} -\lambda D u_{i-1,n} + (1+2\lambda D) u_{i,n} - \lambda D u_{i+1,n} \\ \quad = u_{i,n-1} + \tau\left(\sigma u_{i,n}(1-u_{i,n}) + q(x_i,t_n)\right), \qquad 1 \le i \le L-1, \quad 1 \le n \le N, \\ u_{0,n} = u_{1,n} = 0, \qquad 1 \le n \le N, \\ u_{i,0} = u_0(x_i), \qquad 0 \le i \le L, \end{cases} \tag{6.5}$$

where $\lambda = \tau/h^2$. Let all the parameters be the same as before, and the solution $u_{i,n}$ of (6.5) is computed by a similar iteration as that in (5.1). The corresponding maximum numerical error $\mathsf{error}_h(t_n)$, the convergence order $\mathsf{order}_h(t_n)$ and CPU times for $t_n = 0.125$ and $t_n = 0.25$ are given in Tables 6.2 and 6.3, respectively. We see that the standard method possesses only the second-order accuracy.

To compare time consumption, CPU times for scheme (2.7) and (6.5) at $t_n = 0.125$ and $t_n = 0.25$ are also listed in Tables 6.2 and 6.3. We see that with the same mesh size, scheme (2.7) costs more computational time than scheme (6.5). This is reasonable, since more arithmetic operations are involved in scheme (2.7). However, we see from Table 6.2 that for obtaining numerical solution of scheme (6.5) at $t_n = 0.125$, with the maximum numerical error around 1.122×10^{-6}, we need to take $h = 1/256$, which costs 2965.7970 CPU seconds. On the other hand, a more accurate numerical solution is provided by scheme (2.7) with $h = 1/16$. In this case, the maximum numerical error is 6.479×10^{-7}. But the corresponding cost is only 0.2970 CPU seconds. Similar comparison results at $t_n = 0.25$ are also observed from Table 6.3.

The above comparisons clearly indicate that the present scheme (2.7) is much more efficient than the standard finite difference method.

7. Concluding Remarks

In this paper, a monotone compact implicit (MCI) finite difference scheme is introduced for a class of nonlinear reaction-diffusion equations, and an accelerated monotone iteration is proposed for solving the resulting discrete problem. This new approach has superiority over the usual approaches. This is demonstrated by the numerical evidence. For simplicity, the coefficient $k(x)$ in (1.2) is independent of time t in our discussions. In the general case, the coefficient k may depend on both x and t. Accordingly, the operators $\mathcal{R}_i$ and $\mathcal{Q}_i$ in (2.7) becomes

$$\begin{aligned} \mathcal{R}_{i,n} u_{i,n} &= -J_{i,n} u_{i-1,n} + (J_{i,n} + J_{i+1,n}) u_{i,n} - J_{i+1,n} u_{i+1,n}, \\ \mathcal{Q}_{i,n} u_{i,n} &= E_{i,n} u_{i-1,n} + F_{i,n} u_{i,n} + G_{i,n} u_{i+1,n}, \end{aligned}$$

where

$$E_{i,n} = J_{i+1,n} \int_{x_i}^{x_{i+1}} \phi_{i,1}(x) \Psi_{i+1,n}(x) \mathrm{d}x - J_{i,n} \int_{x_{i-1}}^{x_i} \phi_{i,1}(x) \Psi_{i-1,n}(x) \mathrm{d}x,$$

$$J_{i,n} = \left(\int_{x_{i-1}}^{x_i} \frac{1}{k(s,t_n)} \mathrm{d}s \right)^{-1}, \quad \Psi_{i,n}(x) = \int_x^{x_i} \frac{1}{k(s,t_n)} \mathrm{d}s,$$

and $F_{i,n}$ and $G_{i,n}$ can be evaluated in the same manner. The analysis is similar to that in this paper.

Acknowledgments. The work of the first author was supported in part by NSF of China No. 10571059, E-Institutes of Shanghai Municipal Education Commission No.E03004, Shanghai Priority Academic Discipline, and the Scientific Research Foundation for the Returned

Overseas Chinese Scholars of the State Education Ministry. The work of the second author was supported in part by NSF of China No. 10471095, SF of Shanghai No. 04JC14062, the fund of Chinese Education Ministry No. 20040270002, the Shanghai Leading Academic Discipline Project No. T0401, the fund for E-Institutes of Shanghai Municipal Education Commission No. E03004 and the fund No. 04DB15 of Shanghai Municipal Education Commission.

References

[1] R. Aris, The Mathematical Theory of Diffusion and Reaction in Permeable Catalysis I and II, Oxford University Press (Clarendon), London, 1975.

[2] A.E. Berger, J.M. Solomon and M. Ciment, An analysis of a uniformly accurate difference method for a singular perturbation problem, *Math. Comp.*, **37** (1981), 79-94.

[3] A.E. Berger, J.M. Solomon, M. Ciment, S.H. Leventhal and B.C. Weinberg, Generalized OCI schemes for boundary layer problems, *Math. Comp.*, **35** (1980), 695-731.

[4] M. Ciment, S.H. Leventhal and B.C. Weinberg, The operator compact implicit method for parabolic equations, *J. Comput. Phys.*, **28** (1978), 135-166.

[5] L. Collatz, The Numerical Treatment of Differential Equations, 3rd Edition, Springer-Verlag, Berlin, 1960.

[6] L. Collatz, Hermitean methods for initial value problems in partial differential equations, In *Topics in Numerical Analysis*, Ed. by J.J.H. Miller, Academic Press, London-New York, 1973, 41-61.

[7] Ben-yu Guo, Difference Methods for Partial Differential Equations, Science Press, Beijing, 1988.

[8] Ben-yu Guo and J.J.H. Miller, Iterative and Petrov-Galerkin methods for a nonlinear two-point boundary value problem with application to semiconductor devices, in *Numerical Methods for Partial Differential Equations*, Ed. by Lungan Ying and Benyu Guo, World Scientific, Singapore, 1991, 23-34.

[9] Ben-yu Guo and J.J.H. Miller, Iterative and Petrov-Galerkin methods for solving a system of one-dimensional nonlinear elliptic equations, *Math. Comp.*, **58** (1992), 531-547.

[10] D. Hoff, Stability and convergence of finite difference methods for systems of nonlinear reaction-diffusion equations, *SIAM J. Numer. Anal.*, **15** (1978), 1161-1177.

[11] M.K. Jain, S.R.K. Iyengar and R.K. Jain, Numerical Methods for Scientific and Engineering Computation, John Wiley & Sons, New York, 1985.

[12] Z.L. Li and K. Ito, The Immersed Interface Method: Numerical Solutions of PDEs Involving Interfaces and Irregular Domains, SIAM Frontiers in Applied Mathematics 33, 2006.

[13] X. Lu and W. Feng, Dynamics and numerical simulation of food-chain populations, *Appl. Math. Comp.*, **65** (1994), 335-344.

[14] K.W. Morton, Numerical Solution of Convection-Diffusion Problems, Chapman & Hall, London, UK, 1996.

[15] K.W. Morton and D.F. Mayers, Numerical Solutions of Partial Differential Equations, Cambridge University Press, Cambridge, UK, 1994.

[16] J.M. Ortega and W.C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York, 1970.

[17] C.V. Pao, Monotone iterative methods for finite difference system of reaction-diffusion equations, *Numer. Math.*, **46** (1985), 571-586.

[18] C.V. Pao, Numerical methods of semilinear parabolic equations, *SIAM J. Numer. Anal.*, **24** (1987), 24-35.

[19] C.V. Pao, Nonlinear Parabolic and Elliptic Equations, Plenum Press, New York, 1992.

[20] C.V. Pao, Accelerated monotone iterative methods for finite difference equations of reaction-diffusion, *Numer. Math.*, **79** (1998), 261-281.

[21] C.V. Pao, Numerical analysis of coupled systems of nonlinear parabolic equations, *SIAM J. Numer. Anal.*, **36** (1999), 393-416.

[22] Z.-H. Qiao, Z.-L. Li and T. Tang, A finite difference scheme for solving the nonlinear Poisson-Boltzmann equation modeling charged spheres, *J. Comp. Math.*, 24 (2006), 252-264.

[23] J.I. Ramos, Adaptive methods of lines for one-dimensional reaction-diffusion equations, *Inter. J. Numer. Methods Fluids*, **16** (1993), 697-723.

[24] J.I. Ramos, Implicit, compact, linearized θ-methods with factorization for multidimensional reaction-diffusion equations, *Appl. Math. Comp.*, **94** (1998), 17-43.

[25] G.D. Smith, Numerical Solutions of Partial Differential Equations: Finite Difference Methods, 3rd Edition, Clarendon Press, Oxford, 1985.

[26] B.K. Swartz, The construction of finite difference analogs of some finite element schemes, In *Mathematical Aspects of Finite Elements in Partial Differential Equations*, Ed. by C. de Boor, Academic Press, New York, 1974, 279-312.

[27] V.S. Varga, Matrix Iterative Analysis, Prentice-Hall, Englewood Cliffs, NJ, 1962.

[28] Y.-M. Wang, Petrov-Galerkin methods for nonlinear reaction-diffusion equations, *Inter. J. Computer Math.*, **69** (1998), 123-145.

[29] Y.-M. Wang, On accelerated monotone iterations for numerical solutions of semilinear elliptic boundary value problems, *Appl. Math. Letters*, **18** (2005), 739-755.

[30] D.M. Young, Iterative Solution of Large Linear Systems, Academic Press, New York, 1971.