Lagged diffusivity fixed point iteration for solving steady-state reaction diffusion problems

Emanuele Galligani*

Dipartimento di Matematica Pura e Applicata ‘G. Vitali’, Università degli Studi di Modena e Reggio Emilia, Via Campi 213/b, I-41125, Modena, Italy

(Received 28 September 2011; revised version received 5 January 2012; second revision received 15 February 2012; accepted 24 February 2012)

The paper concerns with the computational algorithms for a steady-state reaction diffusion problem. A lagged diffusivity iterative algorithm is proposed for solving resulting system of quasilinear equations from a finite difference discretization. The convergence of the algorithm is discussed and the numerical results show the efficiency of this algorithm.

Keywords: nonlinear problems; lagging diffusivity; Newton method; arithmetic mean method

2010 AMS Subject Classifications: 65H10; 65N06; 65N22

1. Introduction

We consider the problem of finding a solution of the system of nonlinear equations

\[ F(u) = 0, \]

where \( F : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^n \) is a continuously differentiable mapping.

We are interested in large-scale systems for which the Jacobian of \( F \) is not available or is difficult to compute.

Particularly, we consider the solution of the system of quasilinear equations

\[ F(u) \equiv A(u)u + G(u) - s = 0, \]

where \( A(u) \) is a real matrix of order \( n \) and \( G : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^n \) is a continuously differentiable mapping.

The systems of the form (1) appear in many problems of practical interest. For example, in problems related to the study of reaction and diffusion processes that can be described by nonlinear partial differential equations of elliptic type (e.g. [1,11,14]).

*Email: emanuele.galligani@unimore.it

ISSN 0020-7160 print/ISSN 1029-0265 online
© 2012 Taylor & Francis
http://dx.doi.org/10.1080/00207160.2012.671478
http://www.tandfonline.com
2. Statement of the problem

We are concerned with the numerical solution of the steady-state reaction diffusion problem by the finite difference method.

In the rectangular domain $R$ with boundary $\partial R$ of the $x,y$ plane, we consider the reaction diffusion equation

$$-\text{div} (\sigma (x, y, u) \nabla u(x,y)) + g(x, y, u) = s(x, y), \quad (x,y) \in R,$$

supplemented by the Dirichlet boundary condition

$$u(x,y) = 0, \quad (x,y) \in \partial R. \quad \text{(3)}$$

The functions $\sigma(x, y, u)$, $g(x, y, u)$ and $s(x, y)$ are assumed to satisfy the following Smoothness Conditions for $x$ and $y$ in $R \cup \partial R$ and $u$ in a neighbourhood of a solution of problem (2)–(3). We assume that this model problem has a solution.

(i) The functions $\sigma(x, y, u)$ and $g(x, y, u)$ are piecewise continuous in $x$ and $y$ and continuous in $u$; the source term $s(x, y)$ is piecewise continuous in $x, y$.

(ii) There exist two positive constants $\sigma_{\text{min}}$ and $\sigma_{\text{max}}$ such that

$$0 < \sigma_{\text{min}} \leq \sigma(x, y, u) \leq \sigma_{\text{max}},$$

uniformly in $x, y$ and $u$.

(iii) For fixed $x$ and $y$, the function $\sigma(x, y, u)$ is locally Lipschitz continuous at $u$ (uniformly in $x$ and $y$), with constant $\Lambda > 0$.

(iv) For fixed $x$ and $y$, the function $g(x, y, u)$ is uniformly monotone\(^1\) at $u$ (uniformly in $x$ and $y$), with constant $c > 0$ and is continuously differentiable at $u$.

There exist various techniques for discretizing the problem (2)–(3). The use of the truncated Taylor series to represent the derivative in Equation (2) provides more insight into the nature of the truncation error that arises when the continuous model (2) is replaced by a discrete set of finite difference equations.

Using the Taylor series approach, Equation (2) will be solved with the following standard finite difference scheme.

If the domain $R$ is the unit square, we superimpose on $R$ a uniform grid of points $R_h$. More precisely, $R_h$ is defined in this manner. In the plane $Oxy$, we draw a system of straight lines parallel to the coordinate axes: $x = x_i$ and $y = y_j$ with mesh spacings uniform in each coordinate direction. We denote the intersection of these lines as mesh points or grid points. The coordinates of the interior mesh points $(x_i,y_j)$ of $R_h$ are given by $x_i = i \Delta x$ and $y_j = j \Delta y$ with $i = 1, \ldots, N_x$ and $j = 1, \ldots, N_y$ and $\Delta x = 1/(N_x + 1)$, $\Delta y = 1/(N_y + 1)$.

For simplicity, we assume $\Delta x = \Delta y = h$ and consequently $N_x = N_y = N$.

At the mesh points of $R_h = R_h \cup \partial R_h$, the function $u(x_i,y_j)$ is to be approximated by a grid function $u_{ij}$. In order to approximate partial derivatives, we shall make use of various difference quotients of grid functions. The forward, backward and centred difference quotients with respect to $x$ and $y$ of the grid function $\{u_{ij}\}$ at the mesh point $(x_i,y_j)$ are, respectively,

$$\Delta_x u_{ij} = \frac{u_{i+1,j} - u_{ij}}{h}, \quad \Delta_y u_{ij} = \frac{u_{ij+1} - u_{ij}}{h},$$

$$\nabla_x u_{ij} = \frac{u_{ij} - u_{i-1,j}}{h}, \quad \nabla_y u_{ij} = \frac{u_{ij} - u_{ij-1}}{h},$$

$$\delta_x u_{ij} = \frac{1}{2}(\Delta_x u_{ij} + \nabla_x u_{ij}), \quad \delta_y u_{ij} = \frac{1}{2}(\Delta_y u_{ij} + \nabla_y u_{ij}).$$

This convenient notation was introduced by Courant et al. [3].
A particular important case is the centred second difference quotients which can be written as

$$\delta^2_x u_{ij} = \frac{1}{h^2} (u_{i-1,j} - 2u_{ij} + u_{i+1,j}) = \nabla_x \Delta_x u_{ij} = \Delta_x \nabla_x u_{ij},$$

$$\delta^2_y u_{ij} = \frac{1}{h^2} (u_{ij-1} - 2u_{ij} + u_{ij+1}) = \nabla_y \Delta_y u_{ij} = \Delta_y \nabla_y u_{ij}.$$ 

A centred finite difference approximation of Equation (2) at the mesh point \((x_i, y_j)\) may be written as

$$- \frac{1}{h} \left[ \sigma_{i+j}(u_{i+j}) \frac{u_{i+j} - u_{ij}}{h} - \sigma_j(u_{ij}) \frac{u_{ij} - u_{i-1,j}}{h} \right]$$

$$- \frac{1}{h} \left[ \sigma_{i+j}(u_{i+j}) \frac{u_{ij} - u_{ij-1}}{h} - \sigma_j(u_{ij}) \frac{u_{ij} - u_{i,j+1}}{h} \right] + g_{ij}(u_{ij}) = s(x_i, y_j), \quad (4)$$

for \(i, j = 1, \ldots, N\).

The Dirichlet boundary condition (3) produces the following conditions:

$$u_{i0} = u_{iN+1} = 0, \quad i = 1, \ldots, N,$$

$$u_{0j} = u_{N+1,j} = 0, \quad j = 1, \ldots, N. \quad (5)$$

The total number of difference equations is \(n = N \times N\).

Here we denote

$$\sigma_{ip}(u_{ij}) \equiv \sigma(x_i, y_p, u_{ij}), \quad g_{ip}(u_{ij}) \equiv g(x_i, y_p, u_{ij}).$$

These difference equations may be written in a well-known matrix form.

The above five point discretization formula which approximates the partial derivatives transforms (2) and (3) into a system of nonlinear \(n\) algebraic equations as (1) where the vector \(\mathbf{u}\) is the approximation of the restriction of \(u(x, y)\) on \(R_h\). For natural ordering by horizontal lines of the mesh points, the irreducible \(n \times n\) matrix \(A(\mathbf{u})\) can be partitioned in the form of a block tridiagonal matrix of order \(n\), where each square block on the diagonal of \(A(\mathbf{u})\) is a diagonally dominant tridiagonal matrix of order \(N\) and each block on the subdiagonal and on the superdiagonal of \(A(\mathbf{u})\) is a diagonal matrix of order \(N\). That is, we can express the matrix \(A(\mathbf{u})\) as

$$A(\mathbf{u}) = \begin{pmatrix} A_{11}(\mathbf{u}) & A_{12}(\mathbf{u}) & & & \\ A_{21}(\mathbf{u}) & A_{22}(\mathbf{u}) & & & \\ & \ddots & \ddots & \ddots & \\ & & A_{N-2N-1}(\mathbf{u}) & A_{N-1N-1}(\mathbf{u}) & A_{NN-1}(\mathbf{u}) \\ & & & A_{NN-1}(\mathbf{u}) & A_{NN}(\mathbf{u}) \end{pmatrix}. $$

Using only one index for numbering the mesh points \(P_k, k = 1, \ldots, n\), where \(k = (j-1)N + i\), the elements \(a_{kl}(\mathbf{u})\) of the \(k\)th row, \(k = 1, \ldots, n\), of the matrix \(A(\mathbf{u})\) are given by

$$a_{kk-1}(\mathbf{u}) = -\frac{1}{h^2} \sigma_{ij}(u_{ij}), \quad a_{kk+1}(\mathbf{u}) = -\frac{1}{h^2} \sigma_{i+1,j}(u_{i+1,j}),$$

$$a_{kk-N}(\mathbf{u}) = -\frac{1}{h^2} \sigma_{ij}(u_{ij}), \quad a_{kk+N}(\mathbf{u}) = -\frac{1}{h^2} \sigma_{ij}(u_{ij}),$$

$$a_{kk}(\mathbf{u}) = (2\sigma_{ij}(u_{ij}) + \sigma_{ij+1}(u_{ij+1}) + \sigma_{ij+1}(u_{ij+1}))/h^2,$$

$$a_{ij}(\mathbf{u}) = 0 \quad \text{for} \ l \neq k - N, k - 1, k + 1, k + N; \ l = 1, \ldots, n.$$ 

Now we can define the vector \(\mathbf{s} \in \mathbb{R}^n\) whose component \(s_k, k = 1, \ldots, n\) is the value of the source term \(s(x, y)\) at the mesh point \(P_k\) of \(R_h\). Analogously we define the nonlinear mapping \(G(\mathbf{u})\) whose
kth component $G_k(u)$ is a function only of the coordinates of the mesh point $P_k$ of $R_h$ and of the approximation of $u(x, y)$ at $P_k$, $k = 1, \ldots, n$.

Thus, we can write Equations (4) and (5) in the matrix form (1)

$$ F(u) \equiv A(u)u + G(u) - s = 0. $$

This system of nonlinear equations is characterized by the following Properties (see, e.g. [12]).

(I) The system (1) has at least one solution and all the solutions belong to a well-defined closed ball in $\mathbb{R}^n$

$$ \Omega = \{ u \| u \|_h \leq \rho \}, $$

where $\rho$ is independent of $h$ and of $\|s\|_h$.

Here, $\|u\|_h$ is the discrete $l_2(R_h)$ norm of grid functions $\{u_{ij}\}$:

$$ \|u\|_h = \left( \sum_{i=1}^{N} \sum_{j=1}^{N} h^2 |u_{ij}|^2 \right)^{1/2}. \quad (6) $$

(II) At each mesh point $(x_i, y_j)$ of $\bar{R}_h$ the backward difference quotients with respect to $x$ and to $y$ of a grid function $\{u_{ij}\}$ belonging to $\Omega$ are bounded. The bound is independent of $h$, but depends on $\|s\|_h$.

It is possible to prove the following propositions (see, e.g. [12, pp. 37–43]).

**Proposition 1** No solution of the matrix Equation (1) generated by the discretization of the model problem (2)–(3) can lie outside the ball $\Omega$, where

$$ \rho = \frac{(\|G(0)\|_h + \|s\|_h)}{(\tilde{c} + c)}, $$

with $\tilde{c}$ and $c$ positive constants independent of $h$ for which

$$ \langle A(w)u, u \rangle \geq \tilde{c}\|u\|_h^2, $$

$$ \langle G(u) - G(v), u - v \rangle \geq c\|u - v\|_h^2, $$

for all $u, v, w \in \Omega$.

**Proposition 2** Let $u$ be a solution of the system of 1-D finite difference Equation (1). Then

$$ \frac{|u_i - u_{i-1}|}{h} \leq \frac{2c_0 + \sqrt{c_0}(\sigma_{\text{max}}\|u\|_h)^{1/2}}{\sigma_{\text{min}}} , \quad i = 1, \ldots, N + 1, $$

where

$$ c_0 = \|G(u) + s\|_h. $$

**Definition 1** We say that the grid functions $\{u_{ij}\}$ defined on $\bar{R}_h$ and vanishing on $\partial R_h$, i.e., the vectors $\{u\}$ representing in the interior mesh points $R_h$ these grid functions, satisfy Property $S$ on $\Omega$ if they are uniformly bounded in $\Omega$ and have uniformly bounded backward difference quotients $\nabla_x u_{ij}$ and $\nabla_y u_{ij}$ at each mesh point $(x_i, y_j)$ of $\bar{R}_h$. 
3. Monotonicity of the mapping $F$

We begin with two useful lemmas which establish the discrete analogs of two continuous $L_2(R)$ inner products used frequently in the study of elliptic partial differential equations.

**Lemma 1** Let $\{u_{ij}\}$, $\{v_{ij}\}$ and $\{w_{ij}\}$ be three grid functions defined at the mesh points $(x_i, y_j)$ of $\bar{R}_h$, $i, j = 0, \ldots, N + 1$. Then,

$$h \sum_{i=1}^{N} \left( -\sigma_{i+1j}(w_{i+1}) \frac{u_{i+1j} - u_{ij}}{h^2} + \sigma_{ij}(w_{ij}) \frac{u_{ij} - u_{i-1j}}{h^2} \right) v_{ij}$$

$$= \sum_{i=1}^{N} \sigma_{ij}(w_{ij}) \nabla_x u_{ij} \nabla_x v_{ij} h + \sigma_{1j}(w_{1j}) \frac{u_{1j} - u_{0j}}{h} v_{0j} - \sigma_{N+1j}(w_{N+1j}) \frac{u_{N+1j} - u_{Nj}}{h} v_{Nj} \quad (7)$$

and

$$h \sum_{i=1}^{N} \left( -\sigma_{ij+1}(w_{ij+1}) \frac{u_{ij+1} - u_{ij}}{h^2} + \sigma_{ij}(w_{ij}) \frac{u_{ij} - u_{ij-1}}{h^2} \right) v_{ij}$$

$$= \sum_{i=1}^{N} \sigma_{ij}(w_{ij}) \nabla_y u_{ij} \nabla_y v_{ij} h + \sigma_{i1}(w_{i1}) \frac{u_{ij} - u_{i0}}{h} v_{i0} - \sigma_{iN+1}(w_{iN+1}) \frac{u_{iN+1} - u_{iN}}{h} v_{iN} \quad (8)$$

**Proof** We have

$$h \sum_{i=1}^{N} \left( -\sigma_{i+1j}(w_{i+1}) \frac{u_{i+1j} - u_{ij}}{h^2} + \sigma_{ij}(w_{ij}) \frac{u_{ij} - u_{i-1j}}{h^2} \right) v_{ij}$$

$$= \sigma_{ij}(w_{ij}) \frac{u_{ij} - u_{0j}}{h} v_{ij} + \sum_{i=2}^{N} \sigma_{ij}(w_{ij}) \frac{u_{ij} - u_{i-1j}}{h} v_{ij}$$

$$- \sum_{i=1}^{N-1} \sigma_{i+1j}(w_{i+1}) \frac{u_{i+1j} - u_{ij}}{h} v_{ij} - \sigma_{N+1j}(w_{N+1j}) \frac{u_{N+1j} - u_{Nj}}{h} v_{Nj}$$

$$= \sigma_{ij}(w_{ij}) \frac{u_{ij} - u_{0j}}{h} v_{ij} + \sum_{i=2}^{N} \sigma_{ij}(w_{ij}) \frac{u_{ij} - u_{i-1j}}{h} v_{ij}$$

$$- \sum_{i=2}^{N} \sigma_{ij}(w_{ij}) \frac{u_{ij} - u_{i-1j}}{h} v_{ij} - \sigma_{N+1j}(w_{N+1j}) \frac{u_{N+1j} - u_{Nj}}{h} v_{Nj}$$

$$= \sigma_{ij}(w_{ij}) \frac{u_{ij} - u_{0j}}{h} v_{ij} + \sum_{i=2}^{N} \sigma_{ij}(w_{ij}) \frac{u_{ij} - u_{i-1j}}{h} v_{ij}$$

$$- \sigma_{N+1j}(w_{N+1j}) \frac{u_{N+1j} - u_{Nj}}{h} v_{Nj},$$

then, we have the result (7). Similarly, we obtain formula (8).

**Lemma 2** Let $\{u_{ij}\}$, $\{v_{ij}\}$ and $\{w_{ij}\}$ be three grid functions defined at the mesh points $(x_i, y_j)$ of $\bar{R}_h$, $i, j = 0, \ldots, N + 1$, which are zero at the boundary $\partial R_h$ of $R_h$, i.e. satisfy the condition (5) on $\partial R_h$. 


Then, we have the following expression for the discrete $l_2(R_h)$ inner product of the vectors $A(w)u$ and $v$:

$$
\langle A(w)u, v \rangle = h^2 \sum_{j=1}^{N} \sum_{i=1}^{N} \sigma_{ij}(w_{ij}) \nabla_h u_{ij} \nabla_h v_{ij} + \sum_{j=1}^{N} \sigma_{N+1j}(w_{N+1j}) u_{Nj} v_{Nj} + \sum_{i=1}^{N} \sigma_{iN+1}(w_{iN+1}) u_{iN} v_{iN},
$$

(9)

where $\nabla_h u_{ij}$ denotes the gradient of the grid function $\{u_{ij}\}$ at the mesh point $(x_i, y_j)$ (see formulae (10) and (11)).

**Proof** Since the discrete $l_2(R_h)$ inner product of two grid functions $\{u_{ij}\}$ and $\{v_{ij}\}$ which are zero on the boundary $\partial R$ is defined by

$$
\langle u, v \rangle = h^2 \sum_{j=1}^{N} \sum_{i=1}^{N} u_{ij} v_{ij},
$$

(10)

we have

$$
\langle A(w)u, v \rangle = h \sum_{j=1}^{N} \left( h \sum_{i=1}^{N} \left( -\sigma_{i+1j}(w_{i+1j}) \frac{u_{i+1j} - u_{ij}}{h^2} + \sigma_{ij}(w_{ij}) \frac{u_{ij} - u_{i-1j}}{h^2} \right) v_{ij} \right) + h \sum_{j=1}^{N} \left( h \sum_{j=1}^{N} \left( -\sigma_{ij+1}(w_{ij+1}) \frac{u_{ij+1} - u_{ij}}{h^2} + \sigma_{ij}(w_{ij}) \frac{u_{ij} - u_{ij-1}}{h^2} \right) v_{ij} \right) = h \sum_{j=1}^{N} \left( \sum_{i=1}^{N} \sigma_{ij}(w_{ij}) \nabla_x u_{ij} \nabla_x v_{ij} h + \frac{1}{h} \sigma_{N+1j}(w_{N+1j}) u_{Nj} v_{Nj} \right) + h \sum_{i=1}^{N} \left( \sum_{j=1}^{N} \sigma_{ij}(w_{ij}) \nabla_y u_{ij} \nabla_y v_{ij} h + \frac{1}{h} \sigma_{iN+1}(w_{iN+1}) u_{iN} v_{iN} \right).
$$

We have used formulae (7) and (8) and the hypothesis that the grid functions $\{u_{ij}\}$ and $\{v_{ij}\}$ are zero at the boundary $\partial R$, i.e. satisfy the condition (5) on $\partial R$.

We define the gradient of the grid function $\{u_{ij}\}$ at $(i, j)$ entry to be

$$
\nabla_h u_{ij} \equiv (\nabla_x u_{ij}, \nabla_y u_{ij}), \quad i, j = 1, \ldots, N.
$$

(11)

Thus, we can write

$$
\nabla_h u_{ij} \cdot \nabla_h v_{ij} = \nabla_x u_{ij} \nabla_x v_{ij} + \nabla_y u_{ij} \nabla_y v_{ij}.
$$

Therefore, we obtain the result (9).

**Remark 1** While the grid function $\{u_{ij}\}$ is defined on the whole mesh region $\bar{R}_h$, the vector $u \in \mathbb{R}^n$ represents the grid function $\{u_{ij}\}$ defined only on the interior mesh points $R_h$.

Now we want to prove that under suitable conditions the mapping (1) is uniformly monotone on the closed ball $\Omega$ (see, Property (I)).

Let $\{u_{ij}\}$ and $\{v_{ij}\}$ be two grid functions belonging to $\Omega$, i.e. two solutions of the nonlinear system (1). For the condition (5), they are represented by the vectors $u$ and $v$ of $n$ components.
As consequence of the identity

\[ A(u)v - A(v)w = A(u)(v - w) + (A(u) - A(v))w, \quad (12) \]

we have, by Equation (1),

\[ F(u) - F(v) = A(u)(u - v) + (A(u) - A(v))v + G(u) - G(v). \quad (13) \]

Thus, from Equation (13), the discrete \( L^2(R_h) \) inner product

\[ \langle F(u) - F(v), u - v \rangle = \langle A(u)(u - v), u - v \rangle + \langle (A(u) - A(v))v, u - v \rangle + \langle G(u) - G(v), u - v \rangle. \quad (14) \]

Using Equation (9), we have

\[
\langle A(u)(u - v), u - v \rangle = h^2 \sum_{j=1}^{N} \sum_{i=1}^{N} \sigma_{ij}(u_{ij})|\nabla_h(u_{ij} - v_{ij})|^2 \\
+ \sum_{j=1}^{N} \sigma_{N+1,j}(u_{N+1,j})|u_{Nj} - v_{Nj}|^2 + \sum_{i=1}^{N} \sigma_{iN+1}(u_{iN+1})|u_{IN} - v_{IN}|^2. \quad (15)
\]

Assumption (iv) implies that, for all grid functions \( u \) and \( v \) belonging to \( \Omega \), there exists a positive constant \( c \) such that

\[
(g_{ij}(u_{ij}) - g_{ij}(v_{ij}))(u_{ij} - v_{ij}) \geq c(u_{ij} - v_{ij})^2,
\]

for all \( i, j = 1, \ldots, N \) (see [16, p. 141]). The constant \( c \) is independent of \( h \).

Thus, for the discrete \( L^2(R_h) \) inner product (see definition (10))

\[
\langle G(u) - G(v), u - v \rangle = h^2 \sum_{j=1}^{N} \sum_{i=1}^{N} (g_{ij}(u_{ij}) - g_{ij}(v_{ij}))(u_{ij} - v_{ij}),
\]

and from Equation (16), we have the inequality (see, definition (6))

\[
\langle G(u) - G(v), u - v \rangle \geq c \|u - v\|^2_h.
\]

The next step is to find a bound for \( |\langle (A(u) - A(v))v, u - v \rangle| \). Using Equation (9), we have

\[
|\langle (A(u) - A(v))v, u - v \rangle| = \left| \sum_{j=1}^{N} \sum_{i=1}^{N} (\sigma_{ij}(u_{ij}) - \sigma_{ij}(v_{ij}))\nabla_x v_{ij} \\
\times \nabla_x (u_{ij} - v_{ij})h^2 + \sum_{j=1}^{N} \sum_{i=1}^{N} (\sigma_{ij}(u_{ij}) - \sigma_{ij}(v_{ij}))\nabla_y v_{ij} \nabla_y (u_{ij} - v_{ij})h^2 \\
+ \sum_{j=1}^{N} (\sigma_{N+1,j}(u_{N+1,j}) - \sigma_{N+1,j}(v_{N+1,j}))v_{Nj}(u_{Nj} - v_{Nj}) \\
+ \sum_{i=1}^{N} (\sigma_{iN+1}(u_{iN+1}) - \sigma_{iN+1}(v_{iN+1}))v_{IN}(u_{IN} - v_{IN}) \right|.
\]

Specifically, assumption (iii) says that for given grid functions \( u \) and \( v \) belonging to \( \Omega \), there exist some \( w \) on the line between \( u \) and \( v \) and a positive constant \( \Lambda(w_{ij}) \) such that

\[
|\sigma_{ij}(u_{ij}) - \sigma_{ij}(v_{ij})| \leq \Lambda(w_{ij})|u_{ij} - v_{ij}|
\]

for all \( i, j = 1, \ldots, N + 1 \).
We define

$$\Lambda = \sup_{w_{ij}} \Lambda(w_{ij});$$  \hspace{1cm} (20)

here, \(\Lambda\) is independent of \(h\) and \(\{w_{ij}\} \in \Omega\).

Property (II) assures that there exists a constant \(\beta > 0\) such that

$$|\nabla_x w_{ij}| \leq \beta \quad \text{and} \quad |\nabla_y w_{ij}| \leq \beta,$$

for all \(i,j = 1, \ldots, N + 1\) and all grid function \(\{w_{ij}\}\) belonging to \(\Omega\). \(\beta\) is independent of \(h\).

Now, we may apply the inequalities (19) and (21) and the definition (20) to estimate expression (18), then

$$|\langle (A(u) - A(v))v, u - v \rangle| \leq \Lambda \beta \sum_{j=1}^{N} \sum_{i=1}^{N} |u_{ij} - v_{ij}| h^2 + \Lambda \beta \sum_{j=1}^{N} \sum_{i=1}^{N} |u_{ij} - v_{ij}| |\nabla_x (u_{ij} - v_{ij})| h^2 + \Lambda \sum_{j=1}^{N} |u_{N+1j} - v_{N+1j}| |v_{Nj}| |u_{Nj} - v_{Nj}| + \Lambda \sum_{i=1}^{N} |u_{iN+1} - v_{iN+1}| |v_{iN}| |u_{iN} - v_{iN}|.$$

The last two terms are zero because the grid functions \(\{u_{ij}\}\) and \(\{v_{ij}\}\) belonging to \(\Omega\) are bounded and satisfy the boundary condition (5).

Using a well-known technical trick involving the arithmetic mean and geometric mean inequality of positive numbers \((\sqrt{ab} \leq a/2 + b/2)\), we obtain

$$|\langle (A(u) - A(v))v, u - v \rangle| \leq \frac{\Lambda \beta}{2} \sum_{j=1}^{N} \sum_{i=1}^{N} h^2 \left( \frac{|u_{ij} - v_{ij}|}{\sqrt{\alpha}} |\nabla_x (u_{ij} - v_{ij})| + \frac{|u_{ij} - v_{ij}|}{\sqrt{\alpha}} |\nabla_y (u_{ij} - v_{ij})| \right)$$

$$\leq \frac{\Lambda \beta}{2} \sum_{j=1}^{N} \sum_{i=1}^{N} h^2 \left( \frac{|u_{ij} - v_{ij}|^2}{\alpha} + \alpha |\nabla_x (u_{ij} - v_{ij})|^2 \right)$$

$$+ \frac{\Lambda \beta}{2} \sum_{j=1}^{N} \sum_{i=1}^{N} h^2 \left( \frac{|u_{ij} - v_{ij}|^2}{\alpha} + \alpha |\nabla_y (u_{ij} - v_{ij})|^2 \right),$$  \hspace{1cm} (22)

where \(\alpha\) is as yet an undetermined positive number.

It now follows from Equation (14) that

$$\langle F(u) - F(v), u - v \rangle \geq \langle A(u)(u - v), u - v \rangle + \langle G(u) - G(v), u - v \rangle$$

$$- |\langle (A(u) - A(v))v, u - v \rangle|,$$

and from Equations (15), (17) and (22) and assumption (ii), formula (23) becomes

$$\langle F(u) - F(v), u - v \rangle \geq \sum_{j=1}^{N} \sum_{i=1}^{N} h^2 \sigma_{\min}(||\nabla_x (u_{ij} - v_{ij})||^2 + ||\nabla_y (u_{ij} - v_{ij})||^2)$$

$$+ \sigma_{\min} \left( \sum_{j=1}^{N} |u_{Nj} - v_{Nj}|^2 + \sum_{i=1}^{N} |u_{iN} - v_{iN}|^2 \right).$$
\[ + c\|u - v\|_h^2 - \Lambda \beta \sum_{j=1}^{N} \sum_{i=1}^{N} h^2 |u_{ij} - v_{ij}|^2 \]

\[- \frac{\Lambda \beta}{2} \sum_{j=1}^{N} \sum_{i=1}^{N} h^2 \alpha (|\nabla_x (u_{ij} - v_{ij})|^2 + |\nabla_y (u_{ij} - v_{ij})|^2) \]

\[= \sigma_{\min} \left( \sum_{j=1}^{N} |u_{Nj} - v_{Nj}|^2 + \sum_{i=1}^{N} |u_{iN} - v_{iN}|^2 \right) \]

\[+ \sum_{j=1}^{N} \sum_{i=1}^{N} h^2 \left( \sigma_{\min} - \frac{\Lambda \beta \alpha}{2} \right) (|\nabla_x (u_{ij} - v_{ij})|^2 + |\nabla_y (u_{ij} - v_{ij})|^2) \]

\[+ \frac{\Lambda \beta}{\sigma_{\min}} (\|u - v\|_h^2 - c) h. \tag{24} \]

If we set

\[\alpha = \frac{2\sigma_{\min}}{\Lambda \beta}, \]

then the inequality (24) can be written as

\[\langle F(u) - F(v), u - v \rangle \geq \left( c - \frac{\Lambda \beta}{\alpha} \right) \|u - v\|_h^2. \tag{25} \]

When

\[c - \frac{\Lambda^2 \beta^2}{2\sigma_{\min}} > 0 \]

or

\[\frac{\Lambda^2 \beta^2}{2\sigma_{\min}} < 1, \tag{26} \]

inequality (25) shows that the mapping \(F(u)\) is uniformly monotone on \(\Omega\).

When inequality (25) holds, there exists only one solution \(u^*\) of the system \(F(u) = 0\) in \(\Omega\).

Indeed, suppose that \(\tilde{u}^*\) is some other solution of \(F(u) = 0\) in \(\Omega\), \(\tilde{u}^* \neq u^*\). Then, since the mapping \(F(u)\) is uniformly monotone on \(\Omega\) and Equation (25) holds with \(c - (\Lambda^2 \beta^2)/(2\sigma_{\min}) > 0\), we have

\[0 = \langle F(u^*) - F(\tilde{u}^*), u^* - \tilde{u}^* \rangle \geq \left( c - \frac{\Lambda^2 \beta^2}{2\sigma_{\min}} \right) \|u^* - \tilde{u}^*\|_h^2 > 0, \]

which is a contradiction. Hence \(u^*\) is the only solution of \(F(u) = 0\) in \(\Omega\).

Inequality (26) says that the best chances for the applicability of the algorithm to be defined in next section, are assured when the local Lipschitz constant of \(\sigma(x, y, u)\) is small and the constant \(c\) (in condition (iv)) on the monotonicity of \(g(x, y, u)\) is large.

4. Lagged diffusivity fixed point iteration

We will investigate the solvability of the system of nonlinear difference equations (1) under the assumption that the computation of the Jacobian matrix of \(F(u)\) is not convenient.

For solving this system the easiest and may be the most common method is to lag part of the nonlinear term in Equation (1) (see [20]).
An approximate solution of the weakly nonlinear system (27) is computed by a Newton iterative method in such a way that its solution represents the discretized weakly nonlinear difference system which can be solved inexactly with a Newton iterative method.

That is, we will apply the so-called lagged diffusivity fixed point iteration procedure. Many numerical experiments have shown the following facts.

1. In the special case in which the diffusion term in Equation (2) dominates the reaction term, the lagged diffusivity procedure is an efficient and robust method, even if only linearly convergent, for solving problems with a highly nonlinear differential operator $-\text{div}(\sigma \nabla u)$. (For example, this may happen for nonlinear diffusion models in image denoising.)

When this operator is discretized, the matrix $A(u)$ has highly varying coefficients and Newton’s method for solving the nonlinear system (1) does not work satisfactorily, in the sense that its domain of convergence is very small.

2. In the general case, it may be noted that the steeper the slope of $g$ (in Equation (2)) and the flatter that of $\sigma$ relative to that of $g$, the better are the chances for the applicability of the lagged diffusivity procedure.

Specifically, let $u^{(0)} \in \Omega$ be an initial approximation to the solution of $u^*$ of the system (1). The mapping

$$ F_0(u) \equiv A(u^{(0)})u + G(u) - s = 0, $$

represents the discretized weakly nonlinear difference system which can be solved inexactly with a Newton iterative method. This solution $u^{(0)}$ produces a residual $F_0(u^{(0)})$ with bound $\|F_0(u^{(0)})\| \leq \varepsilon_1$. Here $\| \cdot \|$ is the Euclidean norm.

Thus, $u^{(0)}$ is a new approximation to $u^*$. Now, $u^{(1)}$ becomes the current iterate to generate the new iterate $u^{(2)}$, and so on.

In general, if $u^{(v)}$ is an estimate of the solution $u^*$ of Equation (1), we will determine a new estimate of $u^*$ by solving the weakly nonlinear system

$$ F_v(u) \equiv A(u^{(v)})u + G(u) - s = 0. \quad (27) $$

An approximate solution of the weakly nonlinear system (27) is computed by a Newton iterative method in such a way that its solution $u^{(v+1)}$ will be accepted if the residual $F_v(u^{(v+1)})$ satisfies the condition

$$ \|F_v(u^{(v+1)})\| \leq \varepsilon_{v+1}, \quad (28) $$

where $\varepsilon_{v+1}$ is a given tolerance such that $\varepsilon_{v+1} \to 0$ for $v \to \infty$.

If such suitable solution $u^{(v+1)}$ is found, we say that the algorithm does not break down.

The special form of Equation (1) and the large dimension of the system (1) suggest to use an iterative method for solving at each outer iteration $v + 1$ the system (27). (See, e.g. [16, Sections 7.4, 10.3], [4,5,8].)

Now, it is important to note that the iterate $u^{(v+1)}$ is the solution of a weakly nonlinear reaction diffusion equation (whose diffusivity $\sigma$ depends on the previous iterate) with inhomogenous term $-s - F_v(u^{(v+1)})$.

We assume that all the iterates $u^{(v)}$, $v = 0, 1, \ldots$, belong to $\Omega$ and satisfy Property S in $\Omega$ (see Definition 1).

Since the bound of the backward difference quotients depends on the inhomogeneous term (see Proposition 2), we have that there exist two constants $\beta$ and $\tilde{\beta}$ such that

$$ |\nabla_x u_i^{(v)}| \leq \beta + \varepsilon_i \tilde{\beta} \quad \text{and} \quad |\nabla_y u_i^{(v)}| \leq \beta + \varepsilon_i \tilde{\beta}, \quad (29) $$

instead of Equation (21), $i, j = 1, \ldots, N + 1$. 

We will show that under weak restrictions (i)-(iv) imposed on the functions $\sigma(x, y, u)$ and $g(x, y, u)$ the problem (2)–(3) can be solved via a sequence of systems of weakly nonlinear difference equations where only $g$ but not $\sigma$ depends on the approximate solution $u$ of $u(x, y)$.
Thus, we have

$$E. Galligani$$

$$\beta(\text{see formula (iv)},$$

arising from the discretization of the problem (2)–(3) subject to the Smoothness Conditions (i)–(iv), and characterized by the Properties (I) and (II), which define the closed ball $\Omega$ and the bound $\beta$ (see formula (21)).

When condition (26) holds, there exists a unique solution $u^* \in \Omega$ of $F(u) = 0$.

Suppose that $\{\varepsilon_v\}$ is a sequence of positive numbers such that $\varepsilon_v \to 0$ for $v \to \infty$.

Let $u^{(0)} \in \Omega$ be arbitrary and let $u^{(v+1)}$ be the solution of $F_v(u) = 0$ satisfying the condition (28) with $F_v(u)$ as in Equation (27).

If all the vectors $\{u^{(v)}\}$ belong to $\Omega$ and satisfy the Property S in $\Omega$ with the bound (29) instead of (21), then the sequence $\{u^{(v)}\}$ converges to the unique solution $u^*$ of $F(u) = 0$.

**Proof**  The unique solution $u^*$ in $\Omega$ of Equation (1) satisfies the equation

$$A(u^*)u^* + G(u^*) - s = 0,$$

and the iterate $u^{(v+1)}$ satisfies the equation

$$A(u^{(v)})u^{(v+1)} + G(u^{(v+1)}) = F_v(u^{(v+1)}) + s,$$

where the Euclidean norm of $F_v(u^{(v+1)})$ satisfies the inequality (28).

Subtracting Equation (31) from Equation (30) and taking into account of the identity (12), we can write

$$A(u^*)u^* + G(u^*) - A(u^{(v)})u^{(v+1)} - G(u^{(v+1)}) = -F_v(u^{(v+1)}),$$

as

$$A(u^*)(u^* - u^{(v+1)}) + (A(u^*) - A(u^{(v)}))u^{(v+1)} + G(u^*) - G(u^{(v+1)}) = -F_v(u^{(v+1)}).$$

Thus, we have

$$\langle (A(u^*) - A(u^{(v)}))u^{(v+1)} , u^* - u^{(v+1)} \rangle + \langle (A(u^*) - A(u^{(v)}))u^{(v+1)} , u^* - u^{(v+1)} \rangle$$

$$+ \langle G(u^*) - G(u^{(v+1)}) , u^* - u^{(v+1)} \rangle = \langle -F_v(u^{(v+1)}) , u^* - u^{(v+1)} \rangle.$$

The estimates (15), (17) and (18) in the previous section can be used to yield

$$\langle -F_v(u^{(v+1)}) , u^* - u^{(v+1)} \rangle$$

$$\geq \frac{1}{2} \sum_{j=1}^{N} \sum_{i=1}^{N} \sigma_{ij}(u^*_{ij}) |\nabla h(u^*_{ij} - u^{(v+1)}_{ij})|^2 + \sum_{j=1}^{N} \sigma_{N+j}^*|u^*_{N+j} - u^{(v+1)}_{N+j}|^2.$$
Taking into account of assumption (ii), Equation (32) becomes

$$\langle -F_v(u^{(v+1)}, u^* - u^{(v+1)}) \rangle$$

$$\geq h^2 \sum_{j=1}^{N} \sum_{i=1}^{N} |\sigma_{ij}(u^*_j) - \sigma_{ij}(u^{(v)}_{ij})||\nabla_h(u^*_j) - \nabla_h(u^{(v+1)}_{ij})|$$

$$- h^2 \sum_{j=1}^{N} |\sigma_{N+1j}(u^*_j) - \sigma_{N+1j}(u^{(v)}_{N+1j})||u^*_N - u^{(v+1)}_{Nj}|$$

$$- \frac{\Lambda(\beta + \varepsilon_{v+1}\widetilde{\beta})}{2} \sum_{j=1}^{N} \sum_{i=1}^{N} \left( \frac{|u^*_j - u^{(v)}_{ij}|^2}{\alpha} + \alpha |\nabla_h(u^*_j - u^{(v+1)}_{ij})|^2 \right)$$

$$\geq h^2 \sigma_{\min} \sum_{j=1}^{N} \sum_{i=1}^{N} (|\nabla_x(u^*_j - u^{(v+1)}_{ij})|^2 + |\nabla_y(u^*_j - u^{(v+1)}_{ij})|^2) + \sigma_{\min} \sum_{j=1}^{N} |u^*_{Nj} - u^{(v+1)}_{Nj}|^2$$

$$+ \sigma_{\min} \sum_{i=1}^{N} |u^*_i - u^{(v+1)}_{iN}|^2 - \frac{\Lambda(\beta + \varepsilon_{v+1}\widetilde{\beta})}{\alpha} h^2 \sum_{j=1}^{N} \sum_{i=1}^{N} |u^*_j - u^{(v)}_{ij}|^2$$

Inequalities (19) and (20) and Property (II) and, hence, inequality (29), imply (see formula (22))

$$\varepsilon_{v+1}\|u^* - u^{(v+1)}\|_h$$

$$\geq \langle -F_v(u^{(v+1)}, u^* - u^{(v+1)}) \rangle$$

$$\geq h^2 \sigma_{\min} \sum_{j=1}^{N} \sum_{i=1}^{N} (|\nabla_x(u^*_j - u^{(v+1)}_{ij})|^2 + |\nabla_y(u^*_j - u^{(v+1)}_{ij})|^2) + \sigma_{\min} \sum_{j=1}^{N} |u^*_{Nj} - u^{(v+1)}_{Nj}|^2$$

$$+ \sigma_{\min} \sum_{i=1}^{N} |u^*_i - u^{(v+1)}_{iN}|^2 - \frac{\Lambda(\beta + \varepsilon_{v+1}\widetilde{\beta})}{\alpha} h^2 \sum_{j=1}^{N} \sum_{i=1}^{N} |u^*_j - u^{(v)}_{ij}|^2$$

where $\alpha$ is as yet an undetermined positive number.

Taking into account of assumption (ii), Equation (32) becomes

$$\varepsilon_{v+1}\|u^* - u^{(v+1)}\|_h$$

$$\geq \langle -F_v(u^{(v+1)}, u^* - u^{(v+1)}) \rangle$$

$$\geq h^2 \sigma_{\min} \sum_{j=1}^{N} \sum_{i=1}^{N} (|\nabla_x(u^*_j - u^{(v+1)}_{ij})|^2 + |\nabla_y(u^*_j - u^{(v+1)}_{ij})|^2) + \sigma_{\min} \sum_{j=1}^{N} |u^*_{Nj} - u^{(v+1)}_{Nj}|^2$$

$$+ \sigma_{\min} \sum_{i=1}^{N} |u^*_i - u^{(v+1)}_{iN}|^2 - \frac{\Lambda(\beta + \varepsilon_{v+1}\widetilde{\beta})}{\alpha} h^2 \sum_{j=1}^{N} \sum_{i=1}^{N} |u^*_j - u^{(v)}_{ij}|^2$$
\[
- \frac{\Lambda(\beta + \varepsilon_{v+1}\tilde{\beta})\alpha}{2} h^2 \sum_{j=1}^{N} \sum_{i=1}^{N} (|\nabla_x(u_{ij}^* - u_{ij}^{(v+1)})|^2
+ |\nabla_y(u_{ij}^* - u_{ij}^{(v+1)})|^2) + c\|u^* - u^{(v+1)}\|^2_h.
\]

Hence
\[
\varepsilon_{v+1}\|u^* - u^{(v+1)}\|_h
\geq c\|u^* - u^{(v+1)}\|_h^2 - \frac{\Lambda(\beta + \varepsilon_{v+1}\tilde{\beta})\alpha}{2} h^2 \sum_{j=1}^{N} \sum_{i=1}^{N} |u_{ij}^* - u_{ij}^{(v)}|^2
+ \left(\sigma_{\min} - \frac{\Lambda(\beta + \varepsilon_{v+1}\tilde{\beta})\alpha}{2}\right) h^2 \sum_{j=1}^{N} \sum_{i=1}^{N} (|\nabla_x(u_{ij}^* - u_{ij}^{(v+1)})|^2 + |\nabla_y(u_{ij}^* - u_{ij}^{(v+1)})|^2).
\]

Choosing
\[
\alpha = \frac{2\sigma_{\min}}{\Lambda(\beta + \varepsilon_{v+1}\tilde{\beta})},
\]
we obtain
\[
\|u^* - u^{(v+1)}\|_h^2 - \frac{\Lambda^2(\beta + \varepsilon_{v+1}\tilde{\beta})^2}{2\sigma_{\min} c}\|u^* - u^{(v)}\|_h^2 \leq \frac{\varepsilon_{v+1}}{c}\|u^* - u^{(v+1)}\|_h.
\]

Suppose that \(\{\varepsilon_v\}\) is a sequence of positive numbers such that \(\varepsilon_v \to 0\) for \(v \to \infty\) and that condition (26) holds.

Then (see [12]), we may assume that there exists an integer \(v_0\) such that
\[
\frac{\Lambda^2(\beta + \varepsilon_{v+1}\tilde{\beta})^2}{2\sigma_{\min} c} \leq \gamma < 1 \quad \text{for all} \quad v > v_0.
\]

Moreover, since the grid function \(\{u_{ij}^{(v+1)}\}\) belongs to \(\Omega\), we may assume that
\[
\frac{1}{c}\|u^* - u^{(v+1)}\|_h \leq \frac{1}{c}2\rho \equiv a.
\]

Thus, we obtain the inequality
\[
\|u^* - u^{(v+1)}\|_h^2 \leq \gamma\|u^* - u^{(v)}\|_h^2 + a\varepsilon_{v+1}, \quad v > v_0,
\]
and iteratively
\[
\|u^* - u^{(v_0+\mu)}\|_h^2 \leq \gamma^\mu\|u^* - u^{(v_0)}\|_h^2 + a\sum_{j=1}^{\mu} \gamma^{\mu-j}\varepsilon_{v_0+j}, \quad \mu = 1, 2, \ldots
\]

Since \(\varepsilon_v \to 0\) it follows from the general Toeplitz lemma [16, p. 399] that
\[
\lim_{v \to \infty} \|u^* - u^{(v)}\|_h^2 = 0.
\]

Hence, the sequence \(\{u^{(v)}\}\) of approximate solutions converges to the solution \(u^*\) of the system (1).
Let us now turn to the numerical solution of the weakly nonlinear system (27) with a Newton iterative method.

For the solution of systems of nonlinear equations of the form (1) with a Newton iterative method a common set of Standard Assumptions on $F(u)$ is the following.

- Equation (1) has a solution $u^*$ in an open domain $K$ of $\mathbb{R}^n$. For any vector $u$ and $u^{(v)}$ belonging to $K$:
  - $A(u^{(v)})$ is a block tridiagonal matrix of order $n$ for any iterate $u^{(v)}$. The diagonal blocks are square (although not necessarily all of the same order) tridiagonal submatrices, and the off-diagonal blocks are diagonal submatrices.
  - The matrix $A(u^{(v)})$ is irreducibly diagonally dominant [21, pp. 18, 23] and has positive diagonal entries and nonpositive off-diagonal entries for all the mesh spacings sufficiently small and for all the iterates $u^{(v)} \in K$.
  - $G(u)$ is a continuously differentiable diagonal mapping on $K$, i.e. the $k$th component $G_k$ of $G$ is a function of only the $k$th variable $u_k$, $k = 1, \ldots, n$; furthermore $G'(u) \geq 0$ for all $u \in K$.

Thus, $A(u^{(v)})$ is an irreducible nonsingular $M$-matrix [15, p. 108] and $F'_v(u) = A(u^{(v)}) + G'(u)$ is also an irreducible $M$-matrix with $F'_v(u)^{-1} \leq A(u^{(v)})^{-1}$ for all $u$ and for all the iterates $u^{(v)}$ (see, e.g. [15, p. 109]).

A less restrictive assumption on $G(u)$ may be considered (see, e.g. [6]).

However, with the assumption of $-G(u)$ being diagonal, it is possible to express the rate of change due to a reaction of many reaction diffusion processes in realistic applications. (See e.g. [1,2,10,11,13,14,17,18].)

As a specific example of a solver of Equation (27), we consider the Newton-arithmetic mean (AM) method and its simplified version developed in [6–9].

The Newton-AM method belongs to the general class of Newton-Iterative methods, as, for example, the well-known Newton-SOR method (see [16, Sections 7.4, 10.3]).

The Newton-AM method incorporates at each stage of Newton’s method, the AM method as inner iterative solver for the linear Newton equation [8]. When we apply this method to the weakly nonlinear system (27), we can reduce it to a two-stage iterative method as that described in [6,9].

The above standard assumptions on $F_v(u)$ permit us to analyse the convergence of the Newton-AM method and of the simplified Newton-AM method [6,8,9].

In [6–9], we have also discussed the questions concerning with the global convergence of the Newton-AM method, using the line search strategy, and the monotone convergence of the simplified Newton-AM method.

These methods are well suited for implementation on parallel computers and are significantly efficient when the nonlinear differential operator in Equation (2) is general, i.e. it is composed of a diffusion term and an advection term (see [8]). The advection term may be present in a reaction diffusion equation if the reactions are taking place in a flowing fluid.

5. Numerical experiments

In this section, we consider numerical experiments of the lagged diffusivity fixed point iteration for the solution on a square domain $[0, 1] \times [0, 1]$ of the problem (2)–(3).

Different functions for the nonlinearity factors $\sigma(x, u)$ and $g(x, u)$ have been considered.

We are concerned with some simple steady-state problems (2)–(3) in which reaction kinetics and diffusion are coupled.
This coupling gives rise to reaction diffusion equations which, in a one-dimensional scalar case, can look like
\[
\frac{\partial u}{\partial t} - \sigma(u) \frac{\partial^2 u}{\partial x^2} + g(u) = 0,
\]
where \( u \equiv u(x,t) \) is the concentration, \(-g(u)\) represents the kinetics and \( \sigma(u) \) is the diffusion coefficient, here taken to be concentration dependent. (When \( g(u) = -ru(1 - u/k) \), this equation is known as the Fisher–Kolmogoroff equation [14, Vol. I, p. 400]).

More specifically, we will consider reaction diffusion problems concerning the heat transfer caused by a chemical reaction (see, e.g. [17]) and the growth of spatially distributed communities (see, e.g. [2]).

The aim of this numerical study is to make a computational verification of the effectiveness of the lagged diffusivity procedure for solving nonlinear difference systems of the type (1) in an extended domain of convergence.

The source function \( s(x) \) is chosen in order to satisfy a prespecified exact solution \( u^* = \{u^*(x_i, y_j)\} \) of the nonlinear system (1), \( i, j = 1, \ldots, N \). The solution \( u^*(x,y) \) is chosen equal to \( \sin(\pi x) \sin(\pi y) \).

In the following, we list the functions \( \sigma \) and \( g \) and how they are referred.

\[
\begin{align*}
\sigma_1 &: \sigma(u) = 0.5 + 0.5u, \\
\sigma_2 &: \sigma(u) = 0.02 + 0.5u^2, \\
\sigma_3 &: \sigma(u) = \frac{1}{(0.02 + 0.5u)}, \\
g_1 &: g(u) = 100e^{0.5u}, \\
g_2 &: g(u) = \frac{10^3u}{(1 + 10u)}, \\
g_3 &: g(u) = 5u \log(1 + u).
\end{align*}
\]

We observe that

for \( g_1: g > 0, g' > 0 \) and \( g'' > 0 \), for any value of \( u \);

for \( g_2: g \geq 0 \) and \( g' > 0 \) when \( u \geq 0 \);

for \( g_3: g \geq 0, g' \geq 0 \) and \( g'' > 0 \) when \( u \geq 0 \);

and then, the functions \( g_1, g_2 \) and \( g_3 \) satisfy the standard Assumptions on \( g \) for \( u \geq 0 \).

The lagged diffusivity fixed point iteration has been implemented in a Fortran code with machine precision \( 2.2 \times 10^{-16} \).

This method stops when
\[
\varepsilon_{v+1} \leq \varepsilon,
\]
with \( \varepsilon = 10^{-4} \). An experiment has been carried out with different values of the threshold \( \varepsilon \).

At each iteration of the lagged diffusivity fixed point method, the approximate solution computed by the simplified Newton-AM method satisfies the stopping rule
\[
\|F_v(u^{(v+1)})\| \leq \varepsilon_{v+1},
\]
with \( \varepsilon_1 = 0.1\|F(u^{(0)})\| \) and
\[
\varepsilon_{v+1} = 0.5\varepsilon_v, \quad v = 1, 2, \ldots 
\]

The starting vector of the lagged diffusivity fixed point iteration \( u^{(0)} \) is the null vector \( (u^{(0)} = 0) \) or the vector whose all components are equal to 1 \( (u^{(0)} = e) \).
In the tables, \( \nu^* \) indicates the iteration of the lagged diffusivity fixed point method for which condition (33) is satisfied. The number \( k_{\text{tot}} \) is the sum of the simplified Newton method’s iterations and it is included in brackets; \( j_k \) indicates the number of iterations of the AM method for the solution of thelinear system that occurs at each iteration of the Newton method. This number of iterations of the AM method is chosen to be fixed at each simplified Newton iteration. For the details of the AM method used here, see [19].

Here \( \text{err} \) denotes the computed relative error in \( l_2(R_h) \) norm,
\[
\text{err} = \| u^{(\nu^*)} - u^* \|_h,
\]
and \( \text{diff} \) indicates the last difference of iterations in \( l_2(R_h) \) norm,
\[
\text{diff} = \| u^{(\nu^*)} - u^{(\nu^*-1)} \|_h.
\]
While, with \( \text{res} \) and \( \text{res}_0 \) we indicate the residual and the initial residual in the Euclidean norm:
\[
\text{res} = \| F(u^{(\nu^*)}) \|, \quad \text{res}_0 = \| F(u^{(0)}) \|.
\]

The term \( 4.81(-9) \) indicates \( 4.81 \times 10^{-9} \).

From the numerical experiments, we can draw the conclusions as below.

- We observe that, since \( \varepsilon_{\nu+1} \) decreases, for \( \nu \) increasing, as Equation (34) and the lagged diffusivity iteration method stops at the iteration \( \nu^* \) when the criterium for \( \varepsilon_{\nu+1} \) in Equation (33) is satisfied, we have
  \[
  \varepsilon_{\nu^*+1} = \frac{1}{2} \varepsilon_{\nu^*} = \frac{1}{2^2} \varepsilon_{\nu^*-1} = \cdots = \frac{1}{2^{\nu^*}} \varepsilon_1 \leq \varepsilon,
  \]
  where we set \( \varepsilon_1 = 0.1 \| F(u^{(0)}) \|. \) Then
  \[
  \nu^* > \log_2 \left( \frac{\varepsilon_1}{\varepsilon} \right).
  \]
In the experiments, we obtain
  \[
  \nu^* = \left\lfloor \log_2 \left( \frac{\varepsilon_1}{\varepsilon} \right) \right\rfloor.
  \]

### Table 1. Results for different functions \( g(u) \) and \( \sigma(u) \) with \( u^{(0)} = 0 \).

| \( g(u) \) | \( \nu^* \) (ktot) | err   | res   | res0   | diff   |
|------------|----------------|-------|-------|--------|--------|
| \( \sigma(u) = \sigma 1 \) |                |       |       |        |        |
| \( g_1 \)  | 24 (528)       | 4.81  | 1.07  | 9545.60| 4.93   |
| \( g_2 \)  | 25 (1283)      | 1.13  | 1.86  | 20,077.85| 1.14  |
| \( g_3 \)  | 22 (2135)      | 2.05  | 9.18  | 2234.89| 1.66   |
| \( \sigma(u) = \sigma 2 \) |                |       |       |        |        |
| \( g_1 \)  | 23 (208)       | 1.72  | 1.72  | 8102.34| 7.28   |
| \( g_2 \)  | 25 (661)       | 8.36  | 8.64  | 18,815.92| 1.35  |
| \( g_3 \)  | 20 (6006)      | 2.17  | 2.17  | 933.26 | 3.36   |
| \( \sigma(u) = \sigma 3 \) |                |       |       |        |        |
| \( g_1 \)  | 28 (1793)      | 2.17  | 2.17  | 233,897.67| 8.27  |
| \( g_2 \)  | 28 (3077)      | 2.28  | 2.28  | 237,686.05| 2.27  |
| \( g_3 \)  | 28 (3987)      | 3.41  | 3.41  | 233,112.89| 3.41  |

\( N = 256, \varepsilon = 10^{-4}, j_k = 20 \) and \( u^{(0)} = 0 \).
• We also observe that, generally, in the experiments the outer residual $\|F(u^{(0)})\|$ has the same order of $\varepsilon$ and the error in the discrete $I_2(R_h)$ norm has order of $h\varepsilon$.  

- When in Equation (2) $\sigma(u) = \sigma 2$ and $g(u) = g 3$, the diffusion term and the reaction term are small (Degenerate Equation). If the initial iterate $u^{(0)}$ is zero (Table 1), we have a nonmonotone decreasing of $\|F(u^{(0)})\|$ that produces a large number of iterations of the simplified Newton method; e.g. we have $\|F(u^{(0)})\| = 933.26$ and $\|F(u^{(1)})\| = 8514.10$. The suggestion is to change the initial vector; the same case with $u^{(0)} = e$ produces better results (see Table 2) and a monotone decreasing of $\|F(u^{(0)})\|$. 

- When in Equation (2) $\sigma(u) = \sigma 1$ or $\sigma(u) = \sigma 3$ and $g(u) = g 3$, the diffusion term dominates the reaction term (Equation of Diffusion-type). The matrix $A(u)$ with $\sigma(u) = \sigma 3$ ($g(u) = g 3$) is more ill conditioned than $A(u)$ with $\sigma(u) = \sigma 1$. When in Equation (2) $\sigma(u) = \sigma 1$ or $\sigma(u) = \sigma 2$ and $g(u) = g 2$, the slope of $\sigma(u)$ is quasi-flat and the slope of $g(u)$ is steep. When in Equation (2) $\sigma(u) = \sigma 1$ or $\sigma(u) = \sigma 2$ and $g(u) = g 1$, the slope of $\sigma(u)$ is quasi-flat and the slope of $g(u)$ is quasi-flat with $\|G(u)\|$ large. When in Equation (2) $\sigma(u) = \sigma 3$ and $g(u) = g 1$ or $g(u) = g 2$, the diffusion term is comparable to the reaction term. The numerical results on these test problems (Tables 1 and 2) confirm the statements 1 and 2 at the beginning of the previous section. 

- In the experiment in Table 3 the function $\sigma$ is chosen dependent on $u$ and on $x$ and $y$. The range of values of this function $\sigma$ is larger than that of $\sigma 1$ and a comparison with Tables 1 and 2 confirms that we could expect a larger number of the simplified Newton iterations.

### Table 2. Results for different functions $g(u)$ and $\sigma(u)$ with $u^{(0)} = e$.

| $g(u)$ | $v^*$ (ktot) | err | res | res0 | diff |
|-------|--------------|-----|-----|-----|------|
| $\sigma(u) = \sigma 1$ | | | | | |
| $g 1$ | 31 (553) | 6.64 (−9) | 1.48 (−4) | 1.678,846.53 | 6.38 (−9) |
| $g 2$ | 31 (1291) | 1.48 (−8) | 1.38 (−4) | 1.679,411.24 | 1.48 (−8) |
| $g 3$ | 31 (1987) | 2.43 (−8) | 1.34 (−4) | 1.676,964.70 | 2.44 (−8) |

| $\sigma(u) = \sigma 2$ | | | | | |
| $g 1$ | 30 (231) | 5.52 (−9) | 1.24 (−4) | 780,956.25 | 6.09 (−9) |
| $g 2$ | 30 (684) | 1.82 (−8) | 1.13 (−4) | 781,361.58 | 1.77 (−8) |
| $g 3$ | 30 (928) | 9.14 (−8) | 1.48 (−4) | 779,441.66 | 5.32 (−8) |

| $\sigma(u) = \sigma 3$ | | | | | |
| $g 1$ | 37 (1487) | 5.17 (−9) | 1.91 (−4) | 74,825,897.36 | 5.24 (−9) |
| $g 2$ | 37 (2878) | 1.39 (−8) | 2.89 (−4) | 74,826,351.39 | 1.40 (−8) |
| $g 3$ | 37 (3600) | 2.14 (−8) | 4.18 (−4) | 74,824,482.23 | 2.14 (−8) |

$N = 256, \varepsilon = 10^{-4}, j_k = 20$ and $u^{(0)} = e$. 

### Table 3. Results for different functions $g(u)$ and $\sigma(u)$ dependent on $x, y$ and $u$.

| $g^{(0)}$ | $v^*$ (ktot) | err | res | res0 | diff |
|----------|--------------|-----|-----|-----|------|
| $u^{(0)} = 0$ | 25 (1052) | 4.63 (−9) | 1.49 (−4) | 25,787.21 | 4.59 (−9) |
| $u^{(0)} = 1$ | 25 (1788) | 1.14 (−8) | 1.80 (−4) | 32,151.45 | 1.15 (−8) |
| $g 1$ | 25 (2457) | 1.16 (−8) | 1.27 (−4) | 23,142.53 | 1.16 (−8) |
| $u^{(0)} = e$ | | | | | |
| $g 1$ | 32 (1058) | 3.89 (−9) | 1.25 (−4) | 2,777,386.07 | 3.86 (−9) |
| $g 2$ | 32 (1844) | 7.70 (−9) | 1.21 (−4) | 2,777,975.09 | 7.63 (−9) |
| $g 3$ | 32 (2428) | 1.08 (−8) | 1.19 (−4) | 2,775,500.98 | 1.08 (−8) |

$N = 256, \sigma(u,x,y) = 0.5 + 0.5u + (x + y)^2, \varepsilon = 10^{-4}$ and $j_k = 20$. 

- In the experiment in Table 3 the function $\sigma$ is chosen dependent on $u$ and on $x$ and $y$. The range of values of this function $\sigma$ is larger than that of $\sigma 1$ and a comparison with Tables 1 and 2 confirms that we could expect a larger number of the simplified Newton iterations.
Table 4. Results for different values of $\xi$.

| $\xi$   | $\nu^*(k_{tot})$ | err  | res  | diff |
|---------|-------------------|------|------|------|
| $g(u) = g_1$; res0 = 8102.34; |
| $10^{-2}$  | 17 (152) | 4.91 (−7) | 1.11 (−2) | 4.68 (−7) |
| $10^{-3}$  | 20 (180) | 6.12 (−8) | 1.38 (−3) | 5.83 (−8) |
| $10^{-4}$  | 23 (208) | 7.62 (−9) | 1.72 (−4) | 7.28 (−9) |
| $10^{-5}$  | 27 (245) | 4.86 (−10) | 1.10 (−5) | 4.63 (−10) |
| $g(u) = g_2$; res0 = 18,815.92; |
| $10^{-2}$  | 18 (462) | 1.75 (−6) | 1.09 (−2) | 1.87 (−6) |
| $10^{-3}$  | 21 (547) | 2.21 (−7) | 1.37 (−3) | 2.17 (−7) |
| $10^{-4}$  | 25 (661) | 1.38 (−8) | 8.64 (−5) | 1.35 (−8) |
| $10^{-5}$  | 28 (747) | 1.72 (−9) | 1.07 (−5) | 1.75 (−9) |

$g(u) = g_3$; res0 = 933.26;

$N = 256$, $\sigma(u) = \sigma_2$, $j_k = 20$ and $u^{(0)} = 0$.

Table 5. Results for different values of $N$.

| $N$   | $\nu^*(k_{tot})$ | err  | res  | res0  | diff  |
|-------|-------------------|------|------|-------|-------|
| $\sigma(u) = \sigma_1$; $g(u) = g_1$ |
| 512   | 25 (2193) | 2.41 (−9) | 1.07 (−4) | 19,054.35 | 2.45 (−9) |
| 256   | 24 (528)  | 4.81 (−9) | 1.07 (−4) | 9545.60  | 4.93 (−9) |
| 128   | 23 (128)  | 9.04 (−9) | 1.01 (−4) | 4791.20  | 1.03 (−8) |
| 64    | 22 (29)   | 1.52 (−8) | 8.60 (−5) | 2413.97  | 9.80 (−9) |
| $\sigma(u) = \sigma_2$; $g(u) = g_2$ |
| 512   | 26 (2698) | 6.96 (−9) | 8.64 (−5) | 37,558.59 | 6.99 (−9) |
| 256   | 25 (661)  | 1.38 (−8) | 8.64 (−5) | 18,815.92 | 1.35 (−8) |
| 128   | 24 (172)  | 2.57 (−8) | 8.04 (−5) | 9444.59  | 2.98 (−8) |
| 64    | 23 (53)   | 4.61 (−8) | 7.28 (−5) | 4758.82  | 5.44 (−8) |

$\sigma(u) = \sigma_3$; $g(u) = g_2$

| 512   | 29 (12,641) | 1.21 (−8) | 5.04 (−4) | 504,777.77 | 1.21 (−8) |
| 256   | 28 (3077)  | 2.28 (−8) | 4.70 (−4) | 237,686.05 | 2.27 (−8) |
| 128   | 27 (759)   | 4.14 (−8) | 4.14 (−4) | 109,421.02 | 4.08 (−8) |
| 64    | 26 (197)   | 7.28 (−8) | 3.48 (−4) | 49,538.06  | 7.33 (−8) |

$\xi = 10^{-4}$, $j_k = 20$ and $u^{(0)} = 0$.

The results in Tables 4 and 5 show, in some cases, the behaviour of the Lagged Diffusivity procedure for different values of the threshold $\xi$ and of the number $N$ of grid points per dimension. From Table 4, we see that the choice $\xi = 10^{-4}$ is appropriate.

Note

1. See, e.g. [16, p. 141].

Acknowledgements

The author is very grateful to the anonymous referees for their valuable comments and suggestions, which have improved the paper.
References

[1] R. Aris, *The Mathematical Theory of Diffusion in Permeable Catalysts*, Vols. I and II, Clarendon Press, Oxford, 1975.
[2] G. Buffoni, A. Griffa, Z. Li, and P. de Mottoni, *Spatially distributed communities: The resource–consumer system*, J. Math. Biol. 33 (1995), pp. 723–743.
[3] R. Courant, K.O. Friedrichs, and H. Lewy, *Über die partiellen Differenzengleichungen der mathematischen Physik*, Math. Ann. 100 (1928), pp. 32–74.
[4] R.S. Dembo, S.C. Eisenstat, and T. Steihaug, *Inexact Newton methods*, SIAM J. Numer. Anal. 19 (1982), pp. 400–408.
[5] S.C. Eisenstat and H.F. Walker, *Globally convergent inexact Newton methods*, SIAM J. Optim. 4 (1994), pp. 393–422.
[6] E. Galligani, *A two-stage iterative method for solving a weakly nonlinear system*, Atti del Seminario Matematico e Fisico dell’Università di Modena L, 2002, pp. 195–215.
[7] E. Galligani, *A two-stage iterative method for solving a weakly nonlinear parametrized system*, Int. J. Comput. Math. 79 (2002), pp. 1211–1224.
[8] E. Galligani, *The Newton-arithmetic mean method for the solution of systems of nonlinear equations*, Appl. Math. Comput. 134 (2003), pp. 9–34.
[9] E. Galligani, *On solving a special class of weakly nonlinear finite-difference systems*, Int. J. Comput. Math. 86 (2009), pp. 503–522.
[10] H.B. Keller and D.S. Cohen, *Some positive problems suggested by nonlinear heat generation*, J. Math. Mech. 16 (1967), pp. 1361–1376.
[11] J.P. Kernevez, *Enzyme Mathematics*, North-Holland, Amsterdam, 1980.
[12] G.H. Meyer, *The numerical solution of quasilinear elliptic equations*, in *Numerical Solution of Systems of Nonlinear Algebraic Equations*, G. Byrne and C.A. Hall, eds., Academic Press, New York, 1973, pp. 27–61.
[13] J.J. Moré, *A collection of nonlinear model problems*, in *Computational Solution of Nonlinear Systems of Equations*, E.L. Allgower and K. Georg, eds., Lectures in Applied Mathematics, Vol. 26, American Mathematical Society, Providence, RI, 1990, pp. 723–762.
[14] J.D. Murray, *Mathematical Biology*, Vols. I and II, Springer, Berlin, 2003.
[15] J.M. Ortega, *Numerical Analysis: A Second Course*, Academic Press, New York, 1972 (SIAM, Philadelphia, 1990).
[16] J.M. Ortega and W.C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970 (SIAM, Philadelphia, 2000).
[17] C.Y. Pao, *Monotone iterative methods for finite difference system of reaction diffusion equations*, Numer. Math. 46 (1985), pp. 571–586.
[18] S.T. Pohozaev, *The Dirichlet problem for the equation \( \Delta u = u^2 \)*, Sov. Math. 1 (1960), pp. 1143–1146.
[19] V. Ruggiero and E. Galligani, *A parallel algorithm for solving block tridiagonal linear systems*, Comput. Math. Appl. 24 (1992), pp. 15–21.
[20] J. Thomas, *Numerical Partial Differential Equations: Finite Difference Methods*, Springer, New York, 1995.
[21] R.S. Varga, *Matrix Iterative Analysis*, 2nd ed., Springer, Berlin, 2000.
[22] C.R. Vogel and M.E. Oman, *Iterative methods for total variation denoising*, SIAM J. Scient. Comput. 17 (1996), pp. 227–238.