The Arithmetic Mean Solver in Lagged Diffusivity Method for Nonlinear Diffusion Equations

Emanuele Galligani

Department of Mathematics "G. Vitali", University of Modena and Reggio Emilia, Via Campi 213/b, I-41125, Modena, Italy

Abstract This paper deals with the solution of nonlinear system arising from finite difference discretization of nonlinear diffusion convection equations by the lagged diffusivity functional iteration method combined with different inner iterative solvers. The analysis of the whole procedure with the splitting methods of the Arithmetic Mean (AM) and of the Alternating Group Explicit (AGE) has been developed. A comparison in terms of number of iterations has been done with the BiCG-STAB algorithm. Some numerical experiments have been carried out and they seem to show the effectiveness of the lagged diffusivity procedure with the Arithmetic Mean method as inner solver.

Keywords Lagged Diffusivity, Arithmetic Mean Method, Nonlinear Finite Difference Systems

1. Introduction

We consider a nonlinear diffusion convection equation where the diffusion coefficient, denoted by σ, depends on the solution.

When we use a finite difference discretization, this elliptic equation supplemented by a suitable boundary condition, can be transcribed into a nonlinear system of algebraic equations.

We wish to compute a solution of this system of nonlinear equations with a common iterative procedure in which the nonlinear term, corresponding to the discretization of the diffusivity σ, may be evaluated at the previous iteration (see[22]). In literature, this approach of *nonlinearity lagging* in the diffusivity term is denoted as Lagged Diffusivity Fixed Point Iteration or Lagged Diffusivity Functional Iteration.

In Section 2, a *model problem* described by a nonlinear diffusion convection equation subject to homogeneous Dirichlet boundary conditions is presented and a finite difference discretization is described. Then, the lagged diffusivity procedure for the solution of this nonlinear difference system is stated.

Since, a purpose here is to re-examine the lagged diffusivity procedure for solving the system of nonlinear difference equations of elliptic type in the context of Parallel Computing, the linear difference system that arises at each iteration of the lagged diffusivity procedure is solved with the iterative splitting methods of the Arithmetic

Mean introduced in[20-21] and of the Alternating Group Explicit (AGE) introduced by Evans (see[1-3]).

Thus, the *outer* iterates of the lagged diffusivity procedure are the approximate solutions of the linear systems computed with an *inner* iterative solver; a criterion for acceptability of these approximate solutions is given. A stopping rule for the lagged diffusivity procedure is also given.

Section 3 is devoted to remind the Arithmetic Mean method and AGE method for the solution of linear systems with block tridiagonal coefficient matrix.

In Section 4, the convergence of the lagged diffusivity iteration method is analysed under mild and reasonable assumptions imposed on the diffusivity σ using well known standard techniques (see[16]).

In Section 5, numerical experiments show the behaviour of the inner-outer iterations of the procedure. In this section a comparison of the lagged diffusivity iteration method with different iterative solvers is also presented. The Arithmetic Mean and the AGE methods are compared, in terms of number of iterations, with BiCG-STAB method (see[23]).

2. The Lagged Diffusivity Functional Iteration Method

Consider the model problem described by the nonlinear diffusion convection equation

$$-div(\sigma\nabla u) + \boldsymbol{p}\cdot\nabla u + qu = f, \qquad (1)$$

where $u = u(x,y)$ is the density function at the point (x,y) of a diffusion medium R, $\sigma = \sigma(u) > 0$ is the diffusion coefficient or diffusivity and is dependent on the solution u, $q = q(x,y) \geq 0$ is the absorption term, the velocity vector $\boldsymbol{p} = (p_1, p_2)^T$ is assumed to be constant and $f(x,y)$ is a real valued sufficiently smooth function.

---

* Corresponding author:
emanuele.galligani@unimore.it (Emanuele Galligani)

In the boundary ∂R of R, equation (1) can be supplemented by a homogeneous Dirichlet boundary condition of the form
$$u(x, y) = 0. \qquad (2)$$

In the following, we suppose R to be a rectangular domain with boundary ∂R and we assume that the functions σ, q and f satisfy the "smoothness" conditions:

(i) the function $\sigma = \sigma(u)$ is continuous in u; the functions $q(x,y)$ and $f(x,y)$ are continuous in x,y respectively;

(ii) there exist two positive constants σmin and σmax such that
$$0 < \sigma_{min} \leq \sigma(u) \leq \sigma_{max},$$
uniformly in u; in addition, $q(x,y) \geq q$ min ≥ 0;

(iii) for fixed $(x,y) \in R$, the function $\sigma(u)$ satisfies Lipschitz condition in u with constant Γ (uniformly in x, y), $\Gamma > 0$.

The nonlinearity introduced by the u-dependence of the coefficient $\sigma(u)$ requires that, in general, the solution of equation (1) is approximated by numerical methods.

We superimpose on $R \cup \partial R$ a grid of points $R_h \cup \partial R_h$; the set of the internal points R_h of the grid are the mesh points (x_i, y_j), for $i = 1, ..., N$ and $j = 1, ..., M$, with uniform mesh size h along x and y directions respectively, i.e. $x_{i+1} = x_i + h$ and $y_{j+1} = y_j + h$ for $i = 0, ..., N$, $j = 0, ..., M$.

Thus, at the mesh points of $R \cup \partial R$, (x_i, y_j), for $i = 0, ..., N+1$, $j = 0, ..., M+1$, the solution $u(x_i, y_j)$ is approximated by a grid function u_{ij} defined on $R_h \cup \partial R_h$ and satisfying the boundary condition (2) on ∂R_h.

In order to approximate partial derivatives in (1) we shall make use of difference quotients of grid functions. The forward, backward and centered difference quotients with respect to x and to y of the grid function u_{ij} at the mesh point (x_i, y_j), are, respectively:
$$\Delta_x u_{ij} = \frac{u_{i+1j} - u_{ij}}{h}, \quad \Delta_y u_{ij} = \frac{u_{ij+1} - u_{ij}}{h},$$
$$\nabla_x u_{ij} = \frac{u_{ij} - u_{i-1j}}{h}, \quad \nabla_y u_{ij} = \frac{u_{ij} - u_{ij-1}}{h},$$
$$\delta_x u_{ij} = \frac{1}{2}\left(\Delta_x u_{ij} + \nabla_x u_{ij}\right),$$
$$\delta_y u_{ij} = \frac{1}{2}\left(\Delta_y u_{ij} + \nabla_y u_{ij}\right),$$

while the centered second difference quotient with respect to x and to y can be written
$$\delta_x^2 u_{ij} = \nabla_x \Delta_x u_{ij} = \Delta_x \nabla_x u_{ij},$$
$$\delta_y^2 u_{ij} = \nabla_y \Delta_y u_{ij} = \Delta_y \nabla_y u_{ij}.$$
This notation was introduced by Courant et al. in[4].

Providing a discretization error $O(h^2)$, the finite difference approximation of (1) in (x_i, y_j), $i = 1, ..., N, j = 1, ..., M$, is given by
$$-\Delta_x\left(\sigma(u_{ij})\nabla_x u_{ij}\right) - \Delta_y\left(\sigma(u_{ij})\nabla_y u_{ij}\right) +$$
$$+ p_1 \delta_x u_{ij} + p_2 \delta_y u_{ij} + q(x_i, y_j)u_{ij} = f(x_i, y_j), \quad (3)$$
By ordering in a row lexicographic order the mesh points $P_l = (x_i, y_j)$, (i.e., $l=(j-1)\cdot N+i$ with $j = 1, ..., M$, and $i = 1, ..., N$), we can write the vector $\boldsymbol{u}$ of components u_{ij} and the difference equations (3) as the nonlinear system
$$\boldsymbol{F}(\boldsymbol{u}) \equiv A(\boldsymbol{u})\boldsymbol{u} - \boldsymbol{f} = \boldsymbol{0}, \qquad (4)$$
where the matrix $A(\boldsymbol{u})$ is of order $n = M \times N$ and has the block

tridiagonal form; the M diagonal blocks are tridiagonal matrices of order N and the M-1 sub- and super- diagonal blocks are diagonal matrices of order N.

The five nonzero elements of $A(\boldsymbol{u})$ corresponding to u_{ij-1}, u_{i-1j}, u_{ij}, u_{i+1j} and u_{ij+1} respectively, are $-(B_{ij} + \tilde{B}_{ij})$, $-(L_{ij} + \tilde{L}_{ij})$, $(D_{ij} + \hat{D}_{ij})$, $-(R_{ij} + \tilde{R}_{ij})$, and $-(T_{ij} + \tilde{T}_{ij})$, where
$$L_{ij} \equiv L_{ij}(\boldsymbol{u}) = \frac{1}{h^2}\sigma(u_{ij}), \quad R_{ij} \equiv R_{ij}(\boldsymbol{u}) = \frac{1}{h^2}\sigma(u_{i+1j}), (5)$$
$$B_{ij} \equiv B_{ij}(\boldsymbol{u}) = \frac{1}{h^2}\sigma(u_{ij}), \quad T_{ij} \equiv T_{ij}(\boldsymbol{u}) = \frac{1}{h^2}\sigma(u_{ij+1}),$$
and
$$\tilde{L}_{ij} = \frac{p_1}{2h}, \quad \tilde{R}_{ij} = -\frac{p_1}{2h}, \quad \tilde{B}_{ij} = \frac{p_2}{2h}, \quad \tilde{T}_{ij} = -\frac{p_2}{2h},$$
$$D_{ij} \equiv D_{ij}(\boldsymbol{u}) = B_{ij} + L_{ij} + R_{ij} + T_{ij}, \qquad (6)$$
$$\hat{D}_{ij} = q(x_i, y_j).$$
The matrix $A(\boldsymbol{u})$ is an irreducible matrix ([24, p. 18]).

Providing that the mesh spacing h is sufficiently small, i.e.,
$$h < min\left\{\frac{2\sigma_{min}}{|p_1|}, \frac{2\sigma_{min}}{|p_2|}\right\}, \qquad (7)$$
the matrix $A(\boldsymbol{u})$ is strictly (when $q(x_i, y_j) > 0$) or irreducibly (when $q(x_i, y_j) = 0$) diagonally dominant ([24, p. 23]) and has positive diagonal elements, $a_{rr}(\boldsymbol{u}) > 0$, $r = 1, ..., n$, and nonpositive off diagonal elements $a_{rs}(\boldsymbol{u}) \leq 0$, $r \neq s$, with $r,s = 1, ..., n$; therefore $A(\boldsymbol{u})$ is an $\mathcal{M}$-matrix ([24, p. 91] or [18, p. 110]).

In the case of diffusion equation ($\boldsymbol{p} = \boldsymbol{0}$), the matrix $A(\boldsymbol{u})$ is also symmetric; then it is a symmetric and positive definite matrix ([24, p. 91]).

The vector $\boldsymbol{f}$ in (4) has components $f_l = f(x_i, y_j)$ for $i = 1, ..., N, j = 1, ..., M$ and $l = (j-1)\cdot N+i$.

We remark that while the grid function u_{ij} is defined on the whole mesh region $R_h \cup \partial R_h$, the vector $\boldsymbol{u} \in \mathbb{R}^n$ represents the grid function $\{u_{ij}\}$ defined only on the interior mesh points R_h.

Here, we suppose that a solution $\boldsymbol{u}^*$ of the system (4) exists in $\mathcal{B}$ (see Section 4).

For solving the nonlinear system (4) the easiest and maybe the most common method is *to lag* the nonlinear term in (4) generating an iterative procedure denoted as Lagged Diffusivity Functional Iteration.

With this iterative procedure the nonlinear system (4) can be solved via a sequence of systems of linear equations.

Specifically, given a sequence of positive numbers $\{\varepsilon_\nu\}$ such that $\varepsilon_\nu \to 0$ as $\nu \to \infty$ and an initial estimate $\boldsymbol{u}^{(0)}$ of the solution $\boldsymbol{u}^*$ of the system (4), we generate a sequence of iterates $\{\boldsymbol{u}^{(\nu)}\}$, $\nu = 0, 1, ...$, with the following rule for the transition from a current iteration $\boldsymbol{u}^{(\nu)}$ to the new iterate $\boldsymbol{u}^{(\nu+1)}$:

● Find an approximate solution $\boldsymbol{u}^{(\nu+1)}$ of the linear system
$$\boldsymbol{F}_\nu(\boldsymbol{u}) \equiv A(\boldsymbol{u}^{(\nu)})\boldsymbol{u} - \boldsymbol{f} = \boldsymbol{0}, \qquad (8)$$
with the criterion for acceptability of the solution on the norm
$$\|\boldsymbol{F}_\nu(\boldsymbol{u}^{(\nu+1)})\| \leq \varepsilon_{\nu+1}. \qquad (9)$$

Then, the lagged diffusivity procedure is composed by an *outer* iteration that generates the sequence $\{u^{(v)}\}$ and by an *inner* iterative solver of the linear system (8). This solver must be particularly well suited for implementation on parallel computers.

The termination criterion for the outer iteration is provided by the following stopping rule

$$\varepsilon_{v+1} \leq \underline{\varepsilon}, \quad (10)$$

where ε_{v+1} decreases as $\varepsilon_{v+1} = 0.5\ \varepsilon_v$, $v = 1,2,...$, with $\varepsilon_1 = 0.1\|F\left(u^{(0)}\right)\|$ and $\underline{\varepsilon}$ is a prespecified threshold.

3. Iterative Parallel Solution of the Linear Systems

In this section, we remind the block form of the Alternating Group Explicit (AGE) and of the Arithmetic Mean methods for the solution of the linear system

$$Ax = b, \quad (11)$$

when the $n \times n$ matrix A has the block tridiagonal form

$$A = \begin{pmatrix} B_1 & C_1 & & & & \\ A_2 & B_2 & C_2 & & & \\ & A_3 & B_3 & C_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & A_{M-1} & B_{M-1} & C_{M-1} \\ & & & & A_M & B_M \end{pmatrix} \quad (12)$$

Here, each square block B_i, $i = 1, ..., M$, is a nonsingular $N \times N$ matrix and the blocks A_i and C_i ($i = 1, ..., M$; $A_1 = C_M = 0$) are square matrices of order N ($n = N \times M$).

The AGE method consists in considering the following splitting of the matrix A

$$A = G_1 + G_2,$$

where G_1 and G_2 are the following matrices

$$G_1 = \begin{pmatrix} B_1 & C_1 & & & & & \\ A_2 & B_2' & & & & & \\ & & B_3' & C_3 & & & \\ & & A_4 & B_4' & & & \\ & & & & \ddots & & \\ & & & & & B_{M-1}' & C_{M-1}' \\ & & & & & A_M & B_M' \end{pmatrix},$$

$$G_2 = \begin{pmatrix} B_1' & & & & & \\ & B_2' & C_2 & & & \\ & A_3 & B_3' & & & \\ & & & \ddots & & \\ & & & & B_{M-2}' & C_{M-2} \\ & & & & A_{M-1} & B_{M-1}' \\ & & & & & & B_M' \end{pmatrix},$$

with $B_i' = \frac{1}{2}B_i$, $i = 1, ..., M$.

Thus, starting from a vector $x^{(0)}$, the AGE method generates a sequence of iterates $\{x^{(k)}\}$ as follows; for $r \geq 0$ and $k = 0, 1, ...$, until convergence:

$$(G_1 + rI_n)x^{(k+1/2)} = (rI_n - G_2)x^{(k)} + b,$$
$$(G_2 + rI_n)x^{(k+1)} = (rI_n - G_1)x^{(k+1/2)} + b. \quad (13)$$

Here, I_n is the identity matrix of order n. The AGE method is convergent when the matrices G_1 and G_2 are symmetric and positive definite with $r \geq 0$. In this case it is proved that the optimal choice for r is $r = \sqrt{ab}$ where

$$0 < a \leq \lambda(G_1) \leq b, \qquad 0 < a \leq \lambda(G_2) \leq b,$$

and $\lambda(G_s)$, $s = 1, 2$, are the eigenvalues of the matrix G_s (see, e.g.[1]).

Furthermore, if the matrix A is irreducibly (or strictly) diagonally dominant with positive diagonal elements, $a_{ii} > 0$, and nonpositive off diagonal elements, $a_{ij} \leq 0$, $i \neq j$, and

$$r > \max_{1 \leq i \leq n} \frac{a_{ii}}{2}, \quad \left(or \quad r \geq \max_{1 \leq i \leq n} \frac{a_{ii}}{2} \right), \quad (14)$$

then the AGE method is convergent.

Indeed, from the hypotheses we have that the matrix A is an $\mathcal{M}$-matrix. Set $P_s = G_s + rI_n$ and $Q_s = rI_n - G_s$, $s = 1, 2$. The choice of r in (14) yields that the matrices P_s are strictly (or irreducibly) diagonally dominant and have positive diagonal elements and nonpositive off diagonal elements, then P_s are $\mathcal{M}$-matrices with $P_s^{-1} \geq 0$. From the hypotheses on the matrix A, the matrices Q_s are nonnegative $(Q_s \geq 0, \ s = 1, 2)$.

Thus, the iteration matrix T of the AGE method is

$$T = P_2^{-1}Q_1P_1^{-1}Q_2 \geq 0.$$

Set

$$P^{-1} = 2r\ P_2^{-1}P_1^{-1}, \quad (15)$$

we have $P^{-1} \geq 0$ (P is an $\mathcal{M}$-matrix) and $T = I_n - P^{-1}A$ holds.

Indeed, set $A = P_1 - Q_2 = P_2 - Q_1$, a multiplicative splitting method can be written

$$P_1 x^{(k+1/2)} = (P_1 - A)x^{(k)} + b,$$
$$P_2 x^{(k+1)} = (P_2 - A)x^{(k+1/2)} + b.$$

Then,

$$x^{(k+1)} = (I_n - P_2^{-1}A)(I_n - P_1^{-1}A)x^{(k)}$$
$$+((I_n - P_2^{-1}A)P_1^{-1} + P_2^{-1})\ b,$$

can be seen as a splitting method (P, Q), i.e., $A = P - Q$, by setting

$$P^{-1} = \left((I_n - P_2^{-1}A)P_1^{-1} + P_2^{-1}\right) = P_2^{-1}(P_1 + P_2 - A)P_1^{-1}.$$

Since $P_s = G_s + rI_n$, $s = 1, 2$, and $A = G_1 + G_2$, then we have the expression (15) for P^{-1}.

Now, the proof runs as that of the Regular Splitting Theorem in[18, p. 119].

The Arithmetic Mean method uses the following two splittings of the matrix A

$$A = H_1 + K_1, \quad A = H_2 + K_2,$$

where H_1 and H_2 are the following matrices

$$H_1 = \begin{pmatrix} B_1 & C_1 & & & & \\ A_2 & B_2 & & & & \\ & & B_3 & C_3 & & \\ & & A_4 & B_4 & & \\ & & & & \ddots & \\ & & & & B_{M-1} & C_{M-1} \\ & & & & A_M & B_M \end{pmatrix},$$

$$H_2 = \begin{pmatrix} B_1 & & & & & \\ & B_2 & C_2 & & & \\ & A_3 & B_3 & & & \\ & & & \ddots & & \\ & & & & B_{M-2} & C_{M-2} \\ & & & & A_{M-1} & B_{M-1} \\ & & & & & B_M \end{pmatrix},$$

and $K_1 = A - H_1$, $K_2 = A - H_2$.

We suppose M even. If M is odd, we can proceed in a similar way.

Thus, starting from a vector $\boldsymbol{x}^{(0)}$, the method of the Arithmetic Mean generates a sequence of iterates $\{\boldsymbol{x}^{(k)}\}$ as follows; for $\rho \geq 0$ and $k = 0, 1, ...,$ until convergence:

$$(H_1 + \rho I_n)\boldsymbol{z}^{[1]} = (\rho I_n - K_1)\boldsymbol{x}^{(k)} + \boldsymbol{b},$$
$$(H_2 + \rho I_n)\boldsymbol{z}^{[2]} = (\rho I_n - K_2)\boldsymbol{x}^{(k)} + \boldsymbol{b}, \qquad (16)$$
$$\boldsymbol{x}^{(k+1)} = \tfrac{1}{2}\,(\boldsymbol{z}^{[1]} + \boldsymbol{z}^{[2]}).$$

In the paper[21], it is proved that the block form of the Arithmetic Mean method above described, is convergent when the matrix A is:

● irreducibly (strictly) diagonally dominant with positive diagonal entries and nonpositive off diagonal elements with $\rho > 0$ ($\rho \geq 0$);

● positive definite but not symmetric with $\rho > \rho^*$ where

$$\rho^* = \max\{\frac{|\lambda_{min}\,(F_1)|}{|\lambda_{min}\,(\tilde{A})|}, \frac{|\lambda_{min}\,(F_2)|}{|\lambda_{min}\,(\tilde{A})|}\}$$

Here $\lambda(V)$ denotes an eigenvalue of a matrix V and

$$F_s = H_s H_s^T - K_s K_s^T, \qquad s=1,2,$$

and $\tilde{A}$ is the symmetric positive definite matrix $\tilde{A} = A + A^T$;

● symmetric positive definite with $\rho \geq 0$.

At each iteration k of the AGE method, we have to solve $M/2$ linear systems of order $2N$, $i = 1, 3, 5, ..., M-1$,

$$\begin{bmatrix} B_i' + rI_N & C_i \\ A_{i+1} & B_{i+1}' + rI_N \end{bmatrix}\begin{bmatrix} \boldsymbol{x}_i \\ \boldsymbol{x}_{i+1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{v}_i \\ \boldsymbol{v}_{i+1} \end{bmatrix}, \qquad (17)$$

to obtain the vector $\boldsymbol{x}^{(k+1/2)}$. The solution of (17) can be seen as a block partitioned vector

$$\boldsymbol{x}^{(k+1/2)} = (\boldsymbol{x}_1^{(k+\frac{1}{2})T}, \boldsymbol{x}_2^{(k+\frac{1}{2})T}, ..., \boldsymbol{x}_M^{(k+\frac{1}{2})T})^T,$$

where each block has N components.

Here $\boldsymbol{v} = (\boldsymbol{v}_1^T, \boldsymbol{v}_2^T, ..., \boldsymbol{v}_M^T)^T$ is the right hand side (r.h.s.) of the first equation of (13).

These $M/2$ systems can be solved simultaneously (*in parallel*).

Then, in order to obtain the new iterate $\boldsymbol{x}^{(k+1)}$ of the AGE method, we have to solve, in parallel, $M/2-1$ systems as (17) with $i = 2, 4, 6, ..., M-2$ and two linear systems of order N

$$(B_1' + rI_N)\boldsymbol{x}_1 = \boldsymbol{v}_1,$$
$$(B_M' + rI_N)\boldsymbol{x}_M = \boldsymbol{v}_M,$$

that can be solved in parallel, as well. The AGE method has an intrinsic parallelism.

In the case of the additive splitting method of the Arithmetic Mean, at each iteration k we have to solve $M-1$ linear systems of order $2N$, $i = 1, 2, 3, ..., M-1$,

$$\begin{bmatrix} B_i + \rho I_N & C_i \\ A_{i+1} & B_{i+1} + \rho I_N \end{bmatrix}\begin{bmatrix} \boldsymbol{z}_i \\ \boldsymbol{z}_{i+1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{w}_i \\ \boldsymbol{w}_{i+1} \end{bmatrix}, \qquad (18)$$

where $\boldsymbol{z} = (\boldsymbol{z}_1^T, \boldsymbol{z}_2^T, ..., \boldsymbol{z}_M^T)^T$ and $\boldsymbol{w} = (\boldsymbol{w}_1^T, \boldsymbol{w}_2^T, ..., \boldsymbol{w}_M^T)^T$ indicate the vector $\boldsymbol{z}^{[1]}$ (for $i = 1, 3, 5, ..., M-1$) and the corresponding r.h.s. of the first equation of (16) or the vector $\boldsymbol{z}^{[2]}$ (for $i = 2, 4, 6, ..., M-2$) and the corresponding r.h.s. of the second equation of (16).

Furthermore, we have to solve two linear systems of order N

$$(B_1 + \rho I_N)\boldsymbol{z}_1 = \boldsymbol{w}_1$$
$$(B_M + \rho I_N)\boldsymbol{z}_M = \boldsymbol{w}_M$$

where $\boldsymbol{z}_1$ and $\boldsymbol{z}_M$ indicate the first and the last block of the vector $\boldsymbol{z}^{[2]}$ and $\boldsymbol{w}_1$, $\boldsymbol{w}_M$ the corresponding r.h.s. of the second equation of (16).

These systems can be solved in parallel. The Arithmetic Mean method introduces also an *explicit parallelism* in order to increase the degree of multiprogramming, that is the number of processes that can be executed simultaneously ([13, p. 87]).

When the system (11) arises from the finite difference discretization of the problem (1)-(2), the diagonal blocks of A in (12) are tridiagonal while the sub and superdiagonal blocks are diagonal.

Thus, the systems (17) and (18) can be solved directly as in[9, 8] or iteratively, generating a two-stage iterative method as in[12]. A direct solution of these systems can be performed by cyclic reduction solvers ([17, p. 125],[15]) combined with an approximate Schur complement method ([17, p.123, p. 217]) or with block Gaussian elimination.

4. Analysis of the Convergence of the Lagged Diffusivity Functional Iteration Method

In this section, we prove the convergence of the sequence $\{\boldsymbol{u}^{(v)}\}$ generated by the lagged diffusivity iteration method for the solution of the system (4) under the smoothness assumptions (i)-(iii).

We define u_{ij} and v_{ij}, $i = 0, ..., N+1$ and $j = 0, ..., M+1$, the grid functions defined on $R_h \cup \partial R_h$ and satisfying the Dirichlet boundary condition on ∂R_h.

For grid functions $\{u_{ij}\}$ and $\{v_{ij}\}$ of this type, the discrete $l_2(R_h)$ inner product and norm are defined by the formulas

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

$$\| \boldsymbol{u} \|_h = \left(h^2 \sum_{i=1}^{N} \sum_{j=1}^{M} |u_{ij}|^2 \right)^{1/2} = \langle \boldsymbol{u}, \boldsymbol{u} \rangle^{1/2},$$

respectively.

We say that the grid functions $\{u_{ij}\}$ defined on $R_h \cup \partial R_h$ and vanishing on ∂R_h satisfy **Property A** if they are uniformly bounded 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 $R_h \cup \partial R_h$. The set of all grid functions which satisfy Property A is denoted by $\mathcal{B}$. Thus, $\mathcal{B}$ is the set of all grid functions $\{u_{ij}\}$ for which there exist two positive constants ϱ and β such that

$$\| \boldsymbol{u} \|_h \leq \varrho, \qquad (19)$$
$$|\nabla_x u_{ij}| \leq \beta, \qquad |\nabla_y u_{ij}| \leq \beta, \qquad (20)$$

The constant ϱ is independent of h; also β is independent of h but it depends on $\|\boldsymbol{f}\|$ (see[16]).

We assume that the system (4), $\boldsymbol{F}(\boldsymbol{u}) = \boldsymbol{0}$, where $\boldsymbol{F}(\boldsymbol{u}) \equiv A(\boldsymbol{u})\boldsymbol{u} - \boldsymbol{f}$, has at least one solution $\boldsymbol{u}^* \in \mathcal{B}$.

Suppose that the system (4) arises from the discretization of problem (1)-(2) subject to the conditions (i)-(iii) with $q \geq q_{min} > 0$ and $A(\boldsymbol{u})$ being an irreducible nonsingular $\mathcal{M}$-matrix. If

$$\frac{\Gamma^2 \beta^2}{2 q_{min} \sigma_{min}} < 1$$

for any $\boldsymbol{u} \in \mathcal{B}$, then the mapping $F(\boldsymbol{u})$ is uniformly monotone in $\mathcal{B}$. (See[16],[5] for a proof of this result and e.g.,[19, p. 141] for the definition of uniform monotonicity).

The iterate $\boldsymbol{u}^{(\nu+1)}$ of the lagged diffusivity functional iteration method satisfies the system (8) with the acceptability criterion (9), that is $\boldsymbol{u}^{(\nu+1)}$ is the solution of the linear diffusion convection equation whose diffusivity depends on the previous iterate $\boldsymbol{u}^{(\nu)}$ with inhomogeneous term $-\boldsymbol{f} - \boldsymbol{F}_\nu(\boldsymbol{u}^{(\nu+1)})$.

We assume that all the iterates $\boldsymbol{u}^{(\nu)}$ satisfy Property A. Thus in particular, by inequality (20), the backward difference quotients of each grid function are bounded and they depend on the inhomogeneous term; then, we have that there exist two constants $\beta > 0$ and $\beta_0 > 0$ such that

$$\left| \nabla_x u_{ij}^{(\nu)} \right| \leq \beta + \varepsilon_\nu \beta_0, \qquad \left| \nabla_y u_{ij}^{(\nu)} \right| \leq \beta + \varepsilon_\nu \beta_0, \qquad (21)$$

for $i = 1, ..., N+1$ and $j = 1, ..., M+1$.

Thus, we can state the following result concerning the convergence of the lagged diffusivity functional iteration method.

Theorem 1. Let $\boldsymbol{u}^*$ be a solution of the nonlinear system (4) with $F(\boldsymbol{u}) \equiv A(\boldsymbol{u})\boldsymbol{u} - \boldsymbol{f}$ arising from the finite difference discretization of problem (1)-(2) subject to the conditions (i)-(iii) with $q \geq q_{min} > 0$ and $A(\boldsymbol{u})$ being an irreducible nonsingular $\mathcal{M}$-matrix.

Assume that the mapping $F(\boldsymbol{u})$ is uniformly monotone in $\mathcal{B}$.

Suppose that $\{\varepsilon_\nu\}$ is a sequence of positive numbers such that $\varepsilon_\nu \to 0$ as $\nu \to \infty$.

Let $\boldsymbol{u}^{(0)} \in \mathcal{B}$ be arbitrary and let $\boldsymbol{u}^{(\nu+1)}$ be the solution of $\boldsymbol{F}_\nu(\boldsymbol{u}) = \boldsymbol{0}$ satisfying the condition (9) with $\boldsymbol{F}_\nu(\boldsymbol{u})$ as in (8).

If all the vectors $\boldsymbol{u}^{(\nu)}$ belong to $\mathcal{B}$ and satisfy Property A with (21) instead of (20), then the sequence $\{\boldsymbol{u}^{(\nu)}\}$ converges to $\boldsymbol{u}^*$.

Proof. The proof runs as that in[6, Theor. 1] or in[5, Theor. 1, p. 33].

A more general result on the convergence of the lagged diffusivity functional iteration method can be obtained by defining the mapping $\boldsymbol{u} = G(\boldsymbol{u})$ where

$$G(\boldsymbol{u}) = A(\boldsymbol{u})^{-1} \boldsymbol{f},$$

for all $\boldsymbol{u} \in \mathbb{R}^n$. A solution of the system (4) is a fixed point of the mapping $G(\boldsymbol{u})$.

Since the smoothness condition (iii) on $\sigma(u)$, we can have that the matrix $A(\boldsymbol{u})$ satisfies the Lipschitz-continuity condition for every bounded subset Ω of $\mathbb{R}^n$ with a Lipschitz constant Λ. Then we can write,

$$G(\boldsymbol{u}) - G(\boldsymbol{v}) = A(\boldsymbol{u})^{-1}\big(A(\boldsymbol{v}) - A(\boldsymbol{u})\big)A(\boldsymbol{v})^{-1} \boldsymbol{f}.$$

Thus, we have for $\boldsymbol{u}, \boldsymbol{v} \in \Omega$

$$\| G(\boldsymbol{u}) - G(\boldsymbol{v}) \| \leq \Lambda \| \boldsymbol{f} \| \| A(\boldsymbol{u})^{-1} \| \times$$
$$\times \| A(\boldsymbol{v})^{-1} \| \| \boldsymbol{u} - \boldsymbol{v} \| \leq \Lambda \bar{\beta}^2 \| \boldsymbol{f} \| \| \boldsymbol{u} - \boldsymbol{v} \|$$

where $\bar{\beta}$ is a bound for $\| A(\boldsymbol{u})^{-1} \|$ for any $\boldsymbol{u} \in \Omega$. Here,

$\|\cdot\|$ indicates an arbitrary vector and matrix norm.

The last inequality assures that the mapping $G(\boldsymbol{u})$ satisfies a Lipschitz condition on a bounded subset Ω of $\mathbb{R}^n$.

5. Numerical Experiments

In this section, we consider a numerical experimentation of the lagged diffusivity functional iteration method for the solution of the nonlinear system (4) generated by the finite difference discretization above described, of the elliptic problem (1)-(2). Indeed, we have to solve the system

$$A(\boldsymbol{u})\boldsymbol{u} = \boldsymbol{f}.$$

In these experiments, the vector solution $\boldsymbol{u}^*$ is prefixed and it is composed by the values of the prescribed function $u(x, y) = \sin(\pi x)\sin(\pi y)$ defined on the square $[0, 1] \times [0, 1]$.

The chosen functions for $\sigma(u)$ are

$$\sigma1 : \sigma(u) = 1 + u;$$
$$\sigma2 : \sigma(u) = \frac{a}{b + c\,u};$$
$$\sigma3 : \sigma(u) = 2(1 + 2u - 2u^2);$$

where, in the case of $\sigma2$, we have $\sigma2\_1$ for $a=3$, $b=2$, $c=1$; $\sigma2\_2$ for $a=1.5$, $b=0.1$, $c=0.9$ and $\sigma2\_3$ for $a=1$, $b=0.01$, $c=0.99$.

The vector $\boldsymbol{f}$ is computed as

$$\boldsymbol{f} \equiv \boldsymbol{f}^* = A(\boldsymbol{u}^*)\boldsymbol{u}^*,$$

where the matrix $A(\boldsymbol{u}^*)$ of order n, has elements as in (5) and (6) with $N=M$ and $n = N^2$. In all the experiments we have $N=256$ and $p_1 = p_2 \equiv p$.

At each iteration ν, $\nu = 0,1, ...$, of the lagged diffusivity procedure, we have to solve the linear system of order n

$$A(\boldsymbol{u}^{(\nu)})\boldsymbol{u} = \boldsymbol{f},$$

with the splitting method of the Arithmetic Mean or of the Alternating Group Explicit described in a previous section (see e.g.[9],[10] for an evaluation of the block form of the Arithmetic Mean method on different parallel architectures and[7] for a description of the Fortran code implementing the method). These methods are compared with BiCG-STAB method implemented as in[14, p. 50].

We call $\boldsymbol{u}^{(\nu+1)}$ the new iteration of the lagged diffusivity procedure, computed with $k_{\nu+1}$ iterations of the inner solver such that the inner residual

$$\boldsymbol{F}_\nu(\boldsymbol{u}^{(\nu+1)}) \equiv A(\boldsymbol{u}^{(\nu)})\boldsymbol{u}^{(\nu+1)} - \boldsymbol{f}^*,$$

satisfies the condition (9)

$$\left\| \boldsymbol{F}_\nu(\boldsymbol{u}^{(\nu+1)}) \right\| \leq \varepsilon_{\nu+1},$$

with $\varepsilon_1 = 0.1 \| \boldsymbol{F}(\boldsymbol{u}^{(0)}) \|$ and

$$\varepsilon_{\nu+1} = 0.5\,\varepsilon_\nu. \qquad (22)$$

Here, $\|\cdot\|$ indicates the Euclidean norm.

The vector $\boldsymbol{F}(\boldsymbol{u}^{(0)}) = A(\boldsymbol{u}^{(0)})\boldsymbol{u}^{(0)} - \boldsymbol{f}^*$ is the initial *outer residual* and its Euclidean norm is called *res0*.

The initial vector $\boldsymbol{u}^{(0)}$ is taken as the null vector ($\boldsymbol{u}^{(0)} = \boldsymbol{0}$) or as the vector $\boldsymbol{e}$ whose all the components are equal to 1 ($\boldsymbol{u}^{(0)} = \boldsymbol{e}$).

The lagged diffusivity procedure has been implemented in a Fortran code with machine precision 2.2×10^{-16} and stops when (10) holds, i.e., for $\nu = 0,1, ...,$

$$\varepsilon_{\nu+1} \leq \underline{\varepsilon},$$

with $\underline{\varepsilon} = 10^{-4}$.

We call ν^* the iteration of the lagged diffusivity procedure for which condition (10) is satisfied. That is, the iterate $\boldsymbol{u}^{(\nu^*)}$ satisfies $\|\boldsymbol{F}_{\nu^*-1}(\boldsymbol{u}^{(\nu^*)})\| \leq \varepsilon_{\nu^*}$ $(\varepsilon_{\nu^*} > \underline{\varepsilon})$, and we have $\varepsilon_{\nu^*+1} = 0.5\,\varepsilon_{\nu^*}$ and $\varepsilon_{\nu^*+1} \leq \underline{\varepsilon}$.

In the tables, we report the number of iterations ν^* and, in brackets, the total number of iterations of the inner solver k_T, i.e.,

$$k_T = \sum_{\nu=1}^{\nu^*} k_\nu.$$

In the tables, we also report the discrete $l_2(R_h)$ norm of the error, $err = \| \boldsymbol{u}^* - \boldsymbol{u}^{\nu^*} \|_h$, the Euclidean norm of the outer residual

$$res = \left\| \boldsymbol{F}\left(\boldsymbol{u}^{(\nu^*)}\right) \right\| = \left\| A\left(\boldsymbol{u}^{(\nu^*)}\right) \boldsymbol{u}^{(\nu^*)} - \boldsymbol{f}^* \right\|,$$

and $res0$.

The symbol $*$ close to the value of res indicates that the behaviour of the norm of the outer residual $\|\boldsymbol{F}_\nu(\boldsymbol{u}^{(\nu+1)})\|$ is not monotone decreasing.

The writing *max it.* indicates that at a certain iteration ν, the maximum number of iterations of the inner solver has been reached. The maximum number of inner iterations is set equal to 20000.

The writing *n.c.* indicates that at a certain iteration ν, the condition (7) is not satisfied.

The symbol " $\checkmark$ " close to the number of inner and outer iterations denotes that at a certain iteration ν, the condition (7) is not satisfied and, in these cases, the lagged diffusivity iteration method generates the iteration by performing a prefixed number (equal to 20) of iterations of the inner iterative solver.

The writing $1.71(-8)$ indicates $1.71 \cdot 10^{-8}$.

In the tables, we indicate with *lag-AM*, *lag-AGE* and *lag-B*, the lagged diffusivity functional iteration method with the Arithmetic Mean, the Alternating Group Explicit and the BiCG-STAB method, respectively, as inner solver.

Furthermore, we observe that, since $\varepsilon_{\nu+1}$ decreases, for ν increasing, as (22) and the lagged diffusivity functional iteration method stops at the iteration ν^* when the criterion for ε_{ν^*+1} in (10) is satisfied, we have

$$\varepsilon_{\nu^*+1} = \frac{1}{2}\varepsilon_{\nu^*} = \frac{1}{2}\varepsilon_{\nu^*-1} \dots = \frac{1}{2^{\nu^*}}\varepsilon_1 \leq \underline{\varepsilon},$$

where we set $\varepsilon_1 = 0.1 \| \boldsymbol{F}(\boldsymbol{u}^{(0)}) \|$. Then,

$$\nu^* > \log_2\left(\frac{\varepsilon_1}{\underline{\varepsilon}}\right).$$

Indeed, in the experiments we obtain

$$\nu^* = \lceil \log_2\left(\frac{\varepsilon_1}{\underline{\varepsilon}}\right) \rceil.$$

Table 1. Results for different values of p for $\sigma1$

N=256	$\sigma(u) = \sigma1$	$q=0$	$\boldsymbol{u}^{(0)} = \boldsymbol{0}$	
p		$\nu^*\,(k_T)$		$res0$
	lag-AM	*lag-AGE*	*lag-B*	
500	29 (478) $\checkmark$	29 (538) $\checkmark$	29 (1399)	284368.27
400	28 (547)	28 (621)	28 (1189)	227507.49
300	28 (720)	28 (851)	28 (1118)	170650.88
200	27 (1062)	27 (1243)	27 (1175)*	113804.73
100	26 (2267)	26 (2709)	26 (1022)*	57000.26
50	25 (5131)	25 (6088)	25 (946)	28691.16
10	23 (31110)	23 (35322)	23 (1202)	6833.42
1	22 (55159)	22 (64165)	22 (1357)*	3820.91

Table 2. Results for different values of p for $\sigma2\_1$

N=256	$\sigma(u) = \sigma2\_1$	$q=0$	$\boldsymbol{u}^{(0)} = \boldsymbol{0}$	
p		$\nu^*\,(k_T)$		$res0$
	lag-AM	*lag-AGE*	*lag-B*	
500	29 (441)	29 (470)	29 (1189)	284348.29
400	28 (524)	28 (565)	28 (1110)	227486.24
300	28 (681)	28 (757)	28 (1108)*	170627.52
200	27 (972)	27 (1088)	27 (1050)	113777.13
100	26 (1965)	26 (2208)	26 (1034)	56960.00
50	25 (4232)	25 (4857)	25 (864)	28625.88
10	23 (24409)	23 (30782)	23 (1149)	6605.37
1	22 (60546)	22 (79276)	22 (1581)	3418.60

Table 3. Results for different values of p and $\boldsymbol{u}^{(0)}$ for $\sigma2\_2$

N=256		$\sigma(u) = \sigma2\_2$	$q=0$	
p		$\nu^*\,(k_T)$		$res0$
	lag-AM	*lag-AGE*	*lag-B*	
$\boldsymbol{u}^{(0)} = \boldsymbol{0}$				
200	*n.c.*	*n.c.*	*max it.**	121921.87
100	*n.c.*	*n.c.*	*max it.**	72982.83
50	26 (21137)	26 (70137)	26 (2767)	54613.78
10	26 (56063)	26 (343861)	26 (3907)	47663.32
1	26 (80638)	*max it.*	26 (4449)	47481.15
$\boldsymbol{u}^{(0)} = \boldsymbol{e}$				
500	35 (1272)	35 (4657) $\checkmark$	*max it.**	21347550.45
400	35 (1616)	35 (5671)	*max it.**	21581543.63
300	35 (2154)	35 (7633)	*max it*.*	21820683.14
200	35 (3120)	35 (11226)	*max it*.*	22064801.66
100	35 (6153)	35 (26619)	35 (2782)	22313735.75
50	35 (12048)	35 (57421)	35 (2724)	22439958.69
10	35 (46793)	35 (342113)	35 (4235)	22541761.97
1	35 (72524)	*max it.*	35 (4842)	22564767.64

Table 4. Results for different values of p and $\boldsymbol{u}^{(0)}$ for $\sigma2\_3$

N=256	$\sigma(u) = \sigma2\_3$		q=0	
p	$v^*(k_T)$			res0
	lag-AM	*lag-AGE*	*lag-B*	
$\boldsymbol{u}^{(0)} = \boldsymbol{0}$				
200	*n.c.*	*n.c.*	*max it.* *	746998.50
100	*n.c.*	*n.c.*	*max it.* *	744863.51
50	30 (23198)	*max it.*	*max it.* *	745422.36
10	30 (53247)	*max it.*	30 (10173)	746649.16
1	30 (102915)	*max it.*	30 (13616)	747020.37
$\boldsymbol{u}^{(0)} = \boldsymbol{e}$				
500	38 (2486)* ✓	*n.c.*	*max it.* *	147793693.53
400	38 (2516) ✓	*n.c.*	*max it.*	148074183.84
300	38 (3108) ✓	38 (68052)	*max it.* *	148355269.76
200	38 (4255)	38 (108203)	*max it.* *	148636947.90
100	38 (7153)	*max it.*	*max it.* *	148919214.91
50	38 (11826)	*max it.*	*max it.* *	149060568.19
10	38 (40302)	*max it.*	*max it.* *	149173755.93
1	38 (93196)	*max it.*	*max it.* *	149199236.03

Table 5. Results for different values of p for $\sigma3$

N=256	$\sigma(u) = \sigma3$	q=0	$\boldsymbol{u}^{(0)} = \boldsymbol{0}$	
p	$v^*(k_T)$			res0
	lag-AM	*lag-AGE*	*lag-B*	
500	29 (765)	29 (818)	29 (1161)*	284482.74
400	28 (953)	28 (1026)	28 (1081)	227640.62
300	28 (1253)	28 (1341)	28 (1039) *	170815.09
200	27 (1974)	27 (2118)	27 (1016)	114030.97
100	26 (4339)	26 (4683)	26 (924)	57411.25
50	25 (9725)	25 (10525)	25 (1054)	29460.86
10	24 (46400)	24 (53706)	24 (1565)	9468.12
1	23 (61128)	23 (71699)	*max it.* *	7559.22

Table 6. Results for different values of q

N=256	$\sigma(u) = \sigma1$	$\boldsymbol{u}^{(0)} = \boldsymbol{0}$		
q	$v^*(k_T)$			res0
	lag-AM	*lag-AGE*	*lag-B*	
$p = 20$				
0	24 (14136)	24 (18027)	24 (1069)	11991.73
10	24 (13217)	24 (16868)	24 (998)	12422.56
100	25 (8814)	25 (10794)	25 (600)	19939.78
1000	27 (2118)	27 (2383)	27 (211)	132449.08
$p = 200$				
0	27 (1062)	27 (1243)	27 (1175)	113804.73
10	27 (1061)	27 (1245)	27 (1185)*	113850.93
100	27 (1039)	27 (1232)	27 (1124)*	114914.35
1000	28 (776)	28 (928)	28 (860)	174213.87
10000	31 (267)	31 (290)	31 (128)	1293461.86
N=256	$\sigma(u) = \sigma2\_3$	$p=100$	$\boldsymbol{u}^{(0)} = \boldsymbol{e}$	
0	38 (7153)	*max it.*	*max it.* *	148919214.91
100	38 (6468)	38 (214471)	*max it.* *	148921465.88
1000	38 (3664)	38 (97128)	*max it.* *	148941809.24
10000	38 (1117)	38 (15122)	*max it.* *	149153606.71

Table 7. Some results for the error and the residuals

N=256	$\sigma(u) = \sigma1$	q=0	$\boldsymbol{u}^{(0)} = \boldsymbol{0}$
method	*err*	*res*	*res0*
p=1			
lag-AM	1.71(-8)	1.54(-4)	3820.91
lag-AGE	1.71(-8)	1.62(-4)	
lag-B	7.78(-9)	1.90(-4)	
p=50			
lag-AM	1.49(-9)	1.69(-4)	28691.16
lag-AGE	9.71(-10)	1.71(-4)	
lag-B	1.62(-10)	1.70(-4)	
p=100			
lag-AM	5.03(-10)	1.69(-4)	57000.26
lag-AGE	3.12(-10)	1.69(-4)	
lag-B	4.30(-11)	4.90(-5)	
p=300			
lag-AM	6.89(-11)	1.19(-4)	170650.88
lag-AGE	4.24(-11)	1.20(-4)	
lag-B	1.42(-11)	1.21(-4)	

6. Conclusions

From the numerical experiments the following conclusions can be drawn:

● the outer residual $\boldsymbol{res}$ has the same order of $\underline{\varepsilon}$ and the error $\boldsymbol{err}$ in the discrete $l_2(R_h)$ norm has, in worst cases, order $h\underline{\varepsilon}$;

● the AM method gives better results when the ratio between the maximum value of σ and the smallest component of $\boldsymbol{p}$ is small, that is the coefficient matrix of the linear system is strongly asymmetric (see[20]) or the deviation from asymmetry is decreasing. We define as the deviation from asymmetry of a matrix the difference between the Frobenius norms of the symmetric and nonsymmetric parts of the matrix. Furthermore, we can observe the same behaviour of the AGE method with the one of the AM method, respect to the deviation from asymmetry of the coefficient matrix that occurs at each step of the lagged diffusivity procedure;

● the lagged diffusivity functional iteration method combined with the AM or the AGE method breaks down when the coefficient matrix, at a certain iteration ν, is not an $\mathcal{M}$-matrix (n.c.). In some cases, the AGE inner solver, requires a large number of inner iterations especially for nearly symmetric matrices;

● the behaviour of the residual $\boldsymbol{res}$ when we use AM or AGE method as inner solver, is always monotone, except in one case where the lagged diffusivity procedure breaks down (the coefficient matrix is not an $\mathcal{M}$-matrix) but it has been possible to force the convergence by running a few number of iterations of the inner iterative solver. The nonmonotonicity of the residual happens at these "forced" iterations. This technique of forcing convergence is successful when condition (7) is "almost satisfied".

● the lagged diffusivity functional iteration method combined with the BiCG-STAB method, when it does not break down, requires a number of inner iterations that is not large and seems to be independent from the deviation from

asymmetry of the coefficient matrix. We observe that the failure of the lagged diffusivity procedure with the BiCG-STAB method, in most cases, happens when the decreasing of the residual res is not monotone;

● the behaviour of the lagged diffusivity procedure with AM or AGE method as iterative solver depends on the choice of the initial vector. The choice $u^{(0)} = 0$ in the cases $\sigma(u) = \sigma2\_2$ or $\sigma(u) = \sigma2\_3$ yields to negative values for $u^{(\nu)}$ for a certain ν (tipically $\nu = 1$ or $\nu = 2$) so that condition (7) is not satisfied; that is the initial iterate is too close to a region where a sufficient condition for determining the iterates of the outer procedure is not satisfied. Then, the control on condition (7) is a detection to start from another initial vector.

REFERENCES

[1]   D.J. Evans, "Group explicit iterative methods for solving large linear systems", Taylor & Francis, International Journal of Computer Mathematics, vol. 17, pp. 81-108, 1985.

[2]   D.J. Evans, W.S. Yousif, "The block alternating group explicit method (BLAGE) for the solution of elliptic difference equations", Taylor & Francis, International Journal of Computer Mathematics, vol. 22, pp. 177-185, 1987.

[3]   D.J. Evans, W.S. Yousif, "The solution of two-point boundary value problems by the alternating group explicit (AGE) method", SIAM Journal on Scientific and Statistic Computing, vol. 9, pp. 474-484, 1988.

[4]   R. Courant, K.O. Friedrichs, H. Lewy, Über die partiellen Differenzengleichungen der mathematischen Physik", Springer, Mathematische Annalen, vol. 100, pp. 32-74, 1928.

[5]   E. Galligani, "A note on the iterative solution of nonlinear steady state reaction diffusion problems", University of Modena and Reggio Emilia, Technical Report of Numerical Analysis TR NA-UniMoRE-1-2010, August 2010.

[6]   E. Galligani, "Lagged diffusivity fixed point iteration for solving steady state reaction diffusion problems", Taylor & Francis, International Journal of Computer Mathematics, in press 2012. (Available online: 2 April 2012. DOI:10.1080/00207160.2012.671478)

[7]   E. Galligani, "Analysis of the Fortran routines implementing the arithmetic mean method for the solution of block tridiagonal linear systems", University of Modena and Reggio Emilia, Technical Report of Numerical Analysis TR NA-UniMoRE-4-2012, March 2012.

[8]   E. Galligani, V. Ruggiero, "A parallel preconditioner for block tridiagonal matrices", in: Parallel Computing: Trends and Applications (G.R. Joubert, D. Trystram, F.J. Peters and D.J. Evans eds), Elsevier Science Publishers B.V., Amsterdam, pp. 113-120, 1994.

[9]   E. Galligani, V. Ruggiero, "Analysis of splitting methods for solving block tridiagonal linear systems", in: Proceedings of 2nd International Conference on Software for Multiprocessors and Supercomputers, Theory, Practice, Experience (V.P. Ivannikov, V.A. Serebriakov eds), Russian Academy of Sciences, Moscow, pp. 406-416, 1994.

[10]   E. Galligani, V. Ruggiero, "Implementation of splitting methods for solving block tridiagonal linear systems on transputers", in: Proceedings of 3rd EuroMicro Workshop on Parallel and Distributed Processing (M. Valero, A. Gonzalez eds), IEEE Compute Society Press, Los Alamitos, pp. 409-415, 1995.

[11]   E. Galligani, V. Ruggiero, "A polynomial preconditioner for block tridiagonal matrices", Taylor & Francis, Parallel Algorithms and Applications, vol. 3, pp. 227-237, 1994.

[12]   E. Galligani, V. Ruggiero, "The two-stage arithmetic mean method", Elsevier, Applied Mathematics and Computation, vol. 85, pp. 245-264, 1997.

[13]   K. Hwang, F.A. Briggs, Computer Architecture and Parallel Processing, McGraw-Hill, New York, 1984.

[14]   C.T. Kelley, Iterative Methods for Linear and Nonlinear Equations, SIAM, Philadelphia, 1995.

[15]   N.K. Madsen, G.H. Rodrigue, "Odd-even reduction for pentadiagonal matrices", in: Parallel Computers-Parallel Mathematics, Proceedings of IMACS-GI (M. Feilmeier ed.), North-Holland, Amsterdam, pp. 103-106, 1977.

[16]   G.H. Meyer, "The numerical solution of quasilinear elliptic equations", in: Numerical Solution of Systems of Nonlinear Algebraic Equations (G. Byrne, C.A. Hall eds.), Academic Press, New York, pp. 27-61, 1973.

[17]   J.M. Ortega, Introduction to Parallel and Vector Solution of Linear Systems, Plenum Press, New York, 1988.

[18]   J.M. Ortega, Numerical Analysis: A Second Course, SIAM, Philadelphia, 1990.

[19]   J.M. Ortega, W.C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, SIAM, Philadelphia, 2000.

[20]   V. Ruggiero, E. Galligani, "An iterative method for large sparse linear systems on a vector computer", Elsevier, Computers Mathematics with Applications, vol. 20, no. 1, pp. 25-28, 1990.

[21]   V. Ruggiero, E. Galligani, "A parallel algorithm for solving block tridiagonal linear systems", Elsevier, Computers Mathematics with Applications, vol. 24, no. 4, pp. 15-21, 1992.

[22]   J. Thomas, Numerical Partial Differential Equations: Finite Difference Methods, Springer, New York, 1995.

[23]   H.A. van der Vorst, "Bi-CGSTAB: a fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems", SIAM Journal on Scientific and Statistical Computing, vol. 13, pp. 631-644, 1992.

[24]   R.S. Varga, Matrix Iterative Analysis, Second Edition, Springer, Berlin, 2000.