A mesh independent superlinear algorithm for some nonlinear nonsymmetric elliptic systems

by I. Antal¹, J. Karátson¹

Keywords: outer-inner iteration, Newton’s method, conjugate gradient method, nonlinear nonsymmetric elliptic systems, superlinear convergence, mesh independence

1 Introduction

Nonlinear elliptic transport systems arise in various problems in applied mathematics, most often leading to large-scale problems owing to the huge number of equations, see e.g. [14, 17, 18]. For large-scale elliptic problems, iterative processes are the most widespread solution methods, which often rely on Hilbert space theory when mesh independence is desired. (See e.g. [7, 11, 12] and the authors’ papers [3, 10].)

We consider elliptic transport systems with coupling in the nonlinear reaction terms, for which polynomial growth is allowed, and suitable coercivity is prescribed which can be naturally satisfied when the problem arises from time discretization of parabolic problems. We propose an outer-inner (damped inexact Newton plus PCG) iteration for the finite element discretization of the problem, and prove mesh independent superlinear convergence for both the outer and inner iterations. Numerical experiments strengthen our theoretical results.

2 The problem

We consider nonlinear elliptic transport systems of the form

\[
\begin{align*}
-\text{div} \left( K_i \nabla u_i \right) + b_i \cdot \nabla u_i + f_i(x, u_1, \ldots, u_l) &= g_i \quad (i = 1, \ldots, l) \\
\quad u_i|_{\partial \Omega} &= 0
\end{align*}
\]

on a bounded domain \( \Omega \subset \mathbb{R}^d \) \((d = 2 \text{ or } 3)\) under the following assumptions:

ASSUMPTIONS BVP.

(i) (Smoothness:) \( K_i \in L^\infty(\Omega), \quad b_i \in C^1(\overline{\Omega})^d \) and \( g_i \in L^2(\Omega) \) \((i = 1, \ldots, l)\), further, the function \( f = (f_1, \ldots, f_l) : \Omega \times \mathbb{R}^l \to \mathbb{R}^l \) is measurable and bounded w.r. to the variable \( x \in \Omega \) and \( C^1 \) in the variable \( \xi \in \mathbb{R}^l \).

(ii) (Coercivity:) there is \( m > 0 \) such that \( K_i \geq m \) holds for all \( i = 1, \ldots, l \), further, using the notation \( f'_\xi(x, \xi) := \frac{\partial f(x, \xi)}{\partial \xi} \),

\[
f'_\xi(x, \xi) \eta \cdot \eta - \frac{1}{2} \left( \max_i \text{ div } b_i(x) \right) |\eta|^2 \geq 0
\]
for any \( (x, \xi) \in \Omega \times \mathbb{R}^l \) and \( \eta \in \mathbb{R}^l \).

(iii) (Local Lipschitz continuity:) let \( 3 \leq p \) (if \( d = 2 \)) or \( 3 \leq p \leq 6 \) (if \( d = 3 \)), then there exist constants \( c_1, c_2 \geq 0 \) such that for any \( (x, \xi_1) \) and \( (x, \xi_2) \in \Omega \times \mathbb{R}^l \),

\[
\left\| f'_\xi(x, \xi_1) - f'_\xi(x, \xi_2) \right\| \leq (c_1 + c_2 \max(|\xi_1|, |\xi_2|)^{p-3})|\xi_1 - \xi_2|.
\]

We note that assumption (iii) implies the estimates

\[
\left\| f'_\xi(x, \xi) \right\| \leq c_3 + c_4|\xi|^{p-2}, \quad |f(x, \xi)| \leq c_5 + c_6|\xi|^{p-1}
\]

for any \( (x, \xi) \in \Omega \times \mathbb{R}^l \).

Systems of the form (1) arise e.g. from the time discretization of nonlinear reaction-convection-diffusion (transport) systems

\[
\begin{aligned}
\frac{\partial c_i}{\partial t} - \operatorname{div} (K_i \nabla c_i) + b_i \cdot \nabla c_i + R_i(x, c_1, \ldots, c_l) &= 0 \\
c_i |_{\partial \Omega} &= 0
\end{aligned}
\]

\( (i = 1, \ldots, l). \) (4)

In many real-life problems, e.g. where \( c_i \) are concentrations of chemical species, such systems may consist of a huge number of equations [18]. Using a time discretization with sufficiently small steplength \( \tau \), the obtained nonlinear elliptic systems satisfy the coercivity assumptions above.

For brevity, we write (1) as

\[
-\operatorname{div} (K \nabla u) + b \cdot \nabla u + f(x, u) = g
\]

\( u_{|\partial \Omega} = 0 \)

using obvious notations.

### 3 Weak formulation and properties

The real Hilbert space \( H^1_0(\Omega)^l \) is endowed with the inner product

\[
\langle u, v \rangle_{H^1_0} := \sum_{i=1}^l \int_{\Omega} \nabla u_i \cdot \nabla v_i.
\]

For any \( u \in H^1_0(\Omega)^l \) let

\[
\langle F(u), v \rangle_{H^1_0} = \int_{\Omega} \sum_{i=1}^l \left( K_i \nabla u_i \cdot \nabla v_i + (b_i \cdot \nabla u_i) v_i + f_i(x, u) v_i \right)
\]

\( \equiv \int_{\Omega} \left( K \nabla u \cdot \nabla v + (b \cdot \nabla u) \cdot v + f(x, u) \cdot v \right) \quad (v \in H^1_0(\Omega)^l). \) (7)
This relation defines an operator \( F : H_0^1(\Omega)^l \to H_0^1(\Omega)^l \) via the Riesz representation theorem, since for any fixed \( u \in H_0^1(\Omega)^l \) the r.h.s. integral defines a bounded linear functional on \( H_0^1(\Omega)^l \). The latter is seen in a standard way [8], using the growth condition on \( f \) in (3). Here we rely on the Sobolev embedding theorems [1]: if \( p^* := +\infty \) (if \( d = 2 \)) or \( p^* := 6 \) (if \( d = 3 \)), then for all \( p \leq p^* \) we have the embedding and corresponding estimate

\[
H_0^1(\Omega) \subset L^p(\Omega), \quad \|v\|_{L^p(\Omega)} \leq C_p \cdot \|v\|_{H_0^1(\Omega)} \quad (v \in H_0^1(\Omega)). \tag{8}
\]

with some constant \( C_p > 0 \).

**Proposition 3.1** The operator \( F : H_0^1(\Omega)^l \to H_0^1(\Omega)^l \) is Gateaux differentiable and satisfies

\[
\langle F'(u)h, h \rangle_{H_0^1} \geq m \|h\|^2_{H_0^1} \quad (u, h \in H_0^1(\Omega)^l), \tag{9}
\]

further, \( F' \) is locally Lipschitz continuous, namely,

\[
\|F'(u) - F'(v)\| \leq L(r) \|u - v\|_{H_0^1}
\]

for all \( u, v \in H_0^1(\Omega)^l \) with \( \|u\|_{H_0^1} \leq r, \|u\|_{H_0^1} \leq r \), where

\[
L(r) := c_1 C_3^2 + c_2 C_p^p r^{p-3} \quad (r > 0). \tag{10}
\]

**Proof.** (1) The Gateaux differentiability needs to be proved only for the nonlinear part, but then it follows e.g. from [8]. Using the divergence theorem and assumption (ii), we obtain

\[
\langle F'(u)h, h \rangle_{H_0^1} = \int_\Omega \left( K \|\nabla h\|^2 + f'_\xi(x, u) h \cdot h - \frac{1}{2} \sum_{i=1}^l (\text{div } b_i) h_i^2 \right) \geq \int_\Omega |\nabla h|^2.
\]

(2) Assumption (iii) implies for any \((x, \xi_1)\) and \((x, \xi_2)\) \(\in \Omega \times \mathbb{R}^l\) and \(\eta, \zeta \in \mathbb{R}^l\),

\[
\left| \left( f'_\xi(x, \xi_1) - f'_\xi(x, \xi_2) \right) \eta \cdot \zeta \right| \leq \left( c_1 + c_2 \max \{ |\xi_1|, |\xi_2| \}^{p-3} \right) \xi_1 - \xi_2 |\eta| |\zeta|,
\]

hence for all \( u, v, h, z \in H_0^1(\Omega)^l \)

\[
\left| \langle (F'(u) - F'(v))h, z \rangle_{H_0^1} \right| = \left| \int_\Omega (f'_\xi(x, u) - f'_\xi(x, v)) h \cdot z \right| \\
\leq \int_\Omega \left( c_1 + c_2 \max \{ |u|, |v| \}^{p-3} \right) |u - v| |h| |z|
\]

\[
\leq c_1 \|u - v\|_{L^p} \|h\|_{L^3} \|z\|_{L^3} + c_2 \left( \max \|u\|_{L^p}, \|v\|_{L^p} \right)^{p-3} \|u - v\|_{L^p} \|h\|_{L^p} \|z\|_{L^p}
\]

where for any \( u \in H_0^1(\Omega)^l \), \( \|u\|_{L^p} \equiv \|u\|_{L^p(\Omega)^l} := (\int_\Omega |u|^p)^{1/p} \), and in the last estimate Hölder’s inequality has been used for the cases \( \frac{1}{p} + \frac{1}{q} + \frac{1}{3} = 1 \) and \( \frac{p-3}{p} + \frac{1}{p} + \frac{1}{p} + \frac{1}{p} = 1 \). Then (8) yields

\[
\left| \langle (F'(u) - F'(v))h, z \rangle_{H_0^1} \right| \\
\leq c_1 C_3^2 \|u - v\|_{H_0^1} \|h\|_{H_0^1} \|z\|_{H_0^1} + c_2 C_p^p \left( \max \|u\|_{H_0^1}, \|v\|_{H_0^1} \right)^{p-3} \|u - v\|_{H_0^1} \|h\|_{H_0^1} \|z\|_{H_0^1},
\]

3
hence
\[ \|F'(u) - F'(v)\| = \sup_{h, z \in H^1_0(\Omega)} \left| \langle (F'(u) - F'(v))h, z \rangle_{H^1_0} \right| \]
\[ \leq \left( c_1 C_3^q + c_2 C_p^p \left( \max \|u\|_{H^1_0}, \|v\|_{H^1_0} \right)^{p-3} \right) \|u - v\|_{H^1_0}. \]

**Proposition 3.2** System (1) has a unique weak solution, i.e., \( u \in H^1_0(\Omega) \) satisfying
\[ \langle F(u), v \rangle_{H^1_0} = \int_\Omega g \cdot v \quad (v \in H^1_0(\Omega)). \]

**Proof.** The coercivity (9) implies that for all \( u \in H^1_0(\Omega) \) the operator \( F'(u) \) is regular, i.e. maps onto \( H^1_0(\Omega) \), further,
\[ \|F'(u)h\|_{H^1_0} \geq m\|h\|_{H^1_0} \quad (h \in H^1_0(\Omega)). \] (11)

Then a well-posedness theorem, which follows from [13], provides a unique solution for equation \( F(u) = f \) or
\[ \langle F(u), v \rangle_{H^1_0} = \langle f, v \rangle_{H^1_0} \quad (v \in H^1_0(\Omega)) \]
where the vector \( f \in H^1_0(\Omega) \) satisfies \( \langle f, v \rangle_{H^1_0} = \int_\Omega g \cdot v \quad (v \in H^1_0(\Omega)) \), and the existence of \( f \) follows from the Riesz representation theorem. \( \Box \)

4 FEM discretization and Newton iteration

Let us consider the FEM discretization of system (5) in some FEM subspace
\[ V_h = \text{span}\{\varphi_1, \ldots, \varphi_N\} \subset H^1_0, \]
where \( \varphi_i \) are linearly independent. We seek the FEM solution \( u_h \in V_h \):
\[ \langle F(u_h), v_h \rangle_{H^1_0} = \int_\Omega g \cdot v_h \quad (v \in V_h). \]

Defining the operator \( F_h : V_h \rightarrow V_h \) and the function \( g_h \in V_h \) by the identities \( \langle F_h(u_h), v_h \rangle_{H^1_0} = \langle F(u_h), v_h \rangle_{H^1_0} \) (\( \forall v_h \in V_h \)) and \( \langle g_h, v_h \rangle_{H^1_0} = \int_\Omega g \cdot v_h \) (\( \forall v_h \in V_h \)), respectively, we can write our problem as
\[ F_h(u_h) = g_h \] (12)
in \( V_h \), which requires the solution of an \( N \times N \) nonlinear algebraic system for the coefficient vector of \( u_h \).

We apply the damped inexact Newton method (DIN) for the iterative solution of problem (12). The construction and convergence is established as follows:
Theorem 4.1 Let Assumptions BVP hold. Let \( u_0 \in V_h \) be arbitrary, and let us define \( R_0 := 2m^{-1}\|F_h(u_0) - g_h\|_{H^1_0} + \|u_0\|_{H^1_0} \) and \( L := L(R_0) \) with the function \( L(r) \) defined in (10). The DIN iteration defines a sequence \( (u_n) \subset V_h \) constructed recursively as

\[
\begin{cases}
  u_{n+1} = u_n + \tau_n p_n \quad (n \in \mathbb{N}), & \text{where} \\
  \|F_h(u_n)p_n + (F_h(u_n) - g_h)\|_{H^1_0} \leq \delta_n \|F_h(u_n) - g_h\|_{H^1_0} \quad \text{with} \quad 0 < \delta_n \leq \delta_0 < 1 \\
  \tau_n = \min \{1, \frac{1-\delta_n}{(1+\delta_n)^2} L\|F_h(u_n) - g_h\|_{H^1_0} \}.
\end{cases}
\]

Then

\[
\|u_n - u_h\|_{H^1_0} \leq m^{-1}\|F_h(u_n) - g_h\|_{H^1_0} \rightarrow 0 \text{ monotonically.}
\]

In particular, if

\[
\delta_n \leq \text{const} \cdot \|F_h(u_n) - g_h\|_{H^1_0}^\gamma \quad \text{with some} \quad 0 < \gamma \leq 1
\]

then the convergence is locally of order \( 1+\gamma \), that is, the convergence is linear for \( n_0 \) steps until \( \|F_h(u_n) - g_h\|_{H^1_0} \leq \varepsilon \) where \( \varepsilon \leq (1 - \delta_0)^{m^2/\pi L} \), and further on (as \( \tau_n \equiv 1 \))

\[
\|u_n - u_h\|_{H^1_0} \leq d_1 q^{(1+\gamma)n-n_0}
\]

with some \( d_1 > 0, \quad 0 < q < 1. \)

Proof. We have seen in the proof of Proposition 3.2 that for all \( u \in H^1_0(\Omega) \) the operator \( F'(u) \) is regular and (11) holds. Further, by Proposition 3.1, \( F' \) is locally Lipschitz continuous. These properties are inherited with the same constants by the operator \( F_h \) in \( V_h \) by definition, and they imply the given convergence estimates of the DIN method (see e.g. [8], Theorem 5.12 and Remark 5.17). In particular, as pointed out in the cited remark, \( u_n \) satisfies the a priori estimate

\[
\|u_n\|_{H^1_0} \leq R_0
\]

with \( R_0 \) given in the theorem, hence the formulation involves the global Lipschitz constant \( L := L(R_0) \).

Remark 4.1 (Mesh independence.) Let \( r_n := \|F_h(u_n) - g_h\|_{H^1_0} \). As shown by [8, Theorem 5.12], the linear convergence factor for the first \( n_0 \) steps depends on \( L, m, r_0 \) and \( \delta_0 \), whereas \( d_1 \) and \( q \) in the superlinear estimate (13) depends on \( L, m, r_{n_0} \) and the prescribed sequence \( \delta_n \). If, for a sequence of FEM subspaces \( V_h \) such that \( h \rightarrow 0 \), we define \( u_0 \in V_h \) as the projection of a fixed function in \( H^1_0(\Omega) \), e.g. \( u_0 := 0 \), then \( r_0 \) is bounded in \( h \) and the other constants \( L, m, \delta_0 \) and \( \delta_n \) are given independently of \( h \). Further, \( r_{n_0} \) can be prescribed by the choice of the step when we start the undamped part of the iteration. Hence the convergence rate of the DIN iteration is bounded mesh independently.
5 Solution of the linearized problems: inner CG iterations

Let \( u_n \) be constructed in the DIN iteration, and let us consider the linearized problem

\[
F'_h(u_n)p_h = r_h
\]

(15)

(where \( r_h := g_h - F_h(u_n) \)), which is equivalent to the FEM solution in \( V_h \) of the linear elliptic problem

\[
-\text{div} (K_i \nabla p_i) + b_i \cdot \nabla p_i + \sum_{j=1}^l \partial_j f_i(x, u_n)p_j = r_i \quad (i = 1, \ldots, l)
\]

\[p_i|_{\partial\Omega} = 0\]  

(16)

where \( r_i = g_i + \text{div} (K_i \nabla u_{n,i}) - b_i \cdot \nabla u_{n,i} - f_i(x, u_n) \). Denoting by \( c \) and \( d \) the coefficient vectors of \( p_h \) and \( r_h \), respectively, and by \( L_h^{(n)} \) the stiffness matrix corresponding to the linear problem (16), we need to solve the linear algebraic system

\[
L_h^{(n)} c = d.
\]

(17)

We propose a preconditioned conjugate gradient method to solve (17). We define our preconditioners based on the following equivalent operator: letting

\[
S_i u_i := -\text{div} (K_i \nabla u_i) + h_i u_i \quad (i = 1, \ldots, l)
\]

(18)

(for \( u_i|_{\partial\Omega} = 0 \)), where \( h_i \in L^\infty(\Omega) \) and \( h_i \geq 0 \), we define the independent \( l \)-tuple of elliptic operators

\[
Su = (S_1 u_1, \ldots, S_l u_l).
\]

(19)

We consider the preconditioned form of the algebraic system (17):

\[
S_h^{-1} L_h^{(n)} c = f
\]

(20)

(with \( f = S_h^{-1} c \)), where \( S_h \) denotes the stiffness matrix of \( S \) in the same FEM subspace \( V_h \). This preconditioning leads to the FEM solutions in \( V_h \) of independent symmetric auxiliary linear elliptic problems of the form \( S z_k = f_k \). For such problems various fast solvers are available, in particular when \( K_i \) and \( h_i \) are also constant (like fast Fourier transform, cyclic reduction or multigrid, see e.g. [9, 15, 16]), which turn \( S_h \) into an efficient preconditioner.

Our goal is to apply a suitable CG iteration to (20).

5.1 Conjugate gradient algorithms for nonsymmetric linear problems

In this subsection we summarize the required results on the conjugate gradient method based on [4]. Let us consider a nonsymmetric linear algebraic system

\[
Au = b
\]

(21)
with given \( A \in \mathbb{R}^{N \times N}, \ b \in \mathbb{R}^N \). Let \( \langle \cdot, \cdot \rangle \) be a given inner product on \( \mathbb{R}^N \) and, denoting by \( A^* \) the adjoint of \( A \) w.r.t. this inner product, assume that

\[
A + A^* > 0. \tag{22}
\]

There exist several CG algorithms for such nonsymmetric systems (see e.g. \([2, 6]\)). One of the most widespread ways is to consider the normal (or symmetrized) equation and apply a symmetric CG method. This leads to the following algorithm: let \( u_0 \in \mathbb{R}^N \) be arbitrary, \( r_0 := Au_0 - b, \ s_0 := d_0 := A^*r_0 \); for given \( d_k, u_k, r_k \) and \( s_k \), we let

\[
\begin{aligned}
z_k &= Ad_k, \\
\alpha_k &= \frac{\langle r_k, z_k \rangle}{\|z_k\|^2}, \quad u_{k+1} = u_k - \alpha_k d_k, \quad r_{k+1} = r_k - \alpha_k z_k; \\
\beta_k &= \frac{\|s_{k+1}\|^2}{\|s_k\|^2}, \quad d_{k+1} = s_{k+1} + \beta_k d_k.
\end{aligned} \tag{23}
\]

Let us consider the decomposition \( A = I + C \). Using the notation \( \nu := \min_{x \in \mathbb{R}^N} \frac{\|Az\|^2}{\|z\|^2}, \) the error vector \( r_k := Au_k - b \) satisfies

\[
\left( \frac{\|r_k\|}{\|r_0\|} \right)^{1/k} \leq \frac{2}{k
^{\nu} \sum_{i=1}^k \left( |\lambda_i(C^* + C)| + \lambda_i(C^*C) \right)} \quad (k = 1, 2, \ldots, N). \tag{24}
\]

The above result has a mesh independent bound when suitably applied to elliptic systems. Let us consider the Dirichlet problem

\[
\begin{aligned}
L_iu &\equiv -\text{div} \left( K_i \nabla u_i \right) + b_i \cdot \nabla u_i + \sum_{j=1}^l V_{ij}u_j = g_i \quad (i = 1, \ldots, l) \\
u_i |\partial \Omega &= 0
\end{aligned} \tag{25}
\]

on a bounded domain \( \Omega \subset \mathbb{R}^d \), where \( K_i \) is as in Assumptions BVP, \( b_i \in C^1(\overline{\Omega})^d \), \( g_i \in L^2(\Omega), \ V_{ij} \in L^\infty(\Omega) \), and we assume that \( b_i \) and the matrix \( V = \{V_{ij}\}_{i,j=1}^l \) satisfy the coercivity property

\[
\lambda_{\min}(V + V^T) - \max_i \text{div } b_i \geq 0 \tag{26}
\]

pointwise on \( \Omega \), where \( \lambda_{\min} \) denotes the smallest eigenvalue. (Then system (25) has a unique weak solution \( u \in H^1_0(\Omega)^l \).) Let us choose a FEM subspace \( V_h = \text{span}\{\varphi_1, \ldots, \varphi_N\} \subset H^1_0(\Omega)^l \) and look for the solution of the corresponding algebraic system \( L_h u = b \). We define the preconditioning operator (19) and the corresponding inner product on \( H^1_0(\Omega)^l \)

\[
\langle u, v \rangle_S := \int_{\Omega} \sum_{i=1}^l (K_i \nabla u_i \cdot \nabla v_i + h_i u_i v_i)
\]

which is equivalent to (6). We propose the stiffness matrix \( S_h \) of \( S \) in \( V_h \) as preconditioner for system \( L_h u = b \), and solve the preconditioned system \( S_h^{-1}L_h u = S_h^{-1}b \) using the CG algorithm (23) with the \( S_h \)-inner product and with the cast \( A = S_h^{-1}L_h \) and \( A^* = S_h^{-1}L_h^T \).
Then the following mesh independent superlinear convergence result holds, given in terms of the compact operator $Q_S$ defined via

$$
(Q_S u, v)_S = \sum_{i=1}^{l} \int_{\Omega} \left( (b_i \cdot \nabla u_i) v_i + \sum_{j=1}^{l} V_{ij} u_j - h_i u_i \right) v_i \equiv \int_{\Omega} \left( (b \cdot \nabla u) \cdot v + (V - hI) u \cdot v \right)
$$

(27)

(\textbf{u}, \textbf{v} \in H^1_0(\Omega)^l)$, and denoting by $s_i(Q_S) := \lambda_i((Q^*_S + Q_S)^{1/2}$ and $\lambda_i(Q^*_S + Q_S)$ ($i = 1, 2, \ldots$) the singular values resp. ordered eigenvalues of the corresponding operators:

**Theorem 5.1** [4]. *The CG algorithm (23) with $S_h$-inner product, applied for the $N \times N$ preconditioned system $S_h^{-1}L_h \mathbf{c} = S_h^{-1}b$, yields

$$
\left( \frac{\|r_k\|_{S_h}}{\|r_0\|_{S_h}} \right)^{1/k} \leq \varepsilon_k \quad (k = 1, 2, \ldots, N)
$$

(28)

where $\varepsilon_k = \frac{2}{km^2} \sum_{i=1}^{k} \left( |\lambda_i(Q^*_S + Q_S)| + s_i(Q_S)^2 \right) \rightarrow 0$ (as $k \rightarrow \infty$) (29)

and $(\varepsilon_k)_{k \in \mathbb{N}^+}$ is a sequence independent of $n$ and $V_h$.*

**5.2 Uniform superlinear convergence of the inner CG iteration**

Based on the previous subsection, we apply the CG algorithm (23) with $S_h$-inner product to system (20). We verify that the superlinear convergence rate of this algorithm is bounded uniformly w.r.t. both the mesh and the outer Newton iterate, i.e., the sequence $\varepsilon_k$ in (28) can be replaced by a sequence $\hat{\varepsilon}_k$ which is independent of both $V_h$ and $\mathbf{c}_n$.

We rely on Theorem 5.1. Here the operator $Q_S$ in (27) now contains the Jacobian $V = f'(x, u_n)$, that is, $Q_S = Q_S(n)$ defined by

$$
(Q_S^{(n)} v, z)_S = \sum_{i=1}^{l} \int_{\Omega} \left( (b_i \cdot \nabla v_i) z_i + \sum_{j=1}^{l} \partial_j f_i(x, u_n) v_j - h_i v_i \right) z_i
$$

$$
\equiv \int_{\Omega} \left( (b \cdot \nabla v) \cdot z + (f'(x, u_n) - hI) v \cdot z \right)
$$

(\textbf{v}, \textbf{z} \in H^1_0(\Omega)^l)$.

Although Theorem 5.1 itself states mesh independence for the linear problem (25), our linearized algebraic system (20) depends on an outer Newton iterate $u_n$ constructed in a given FEM subspace. Hence even the mesh independence part itself of the following theorem does not obviously follow from Theorem 5.1. We now give our estimate involving two minimax ratios, related to the $L^2$ and $L^p$ norms, respectively.

**Theorem 5.2** *The CG algorithm (23) with $S_h$-inner product, applied for the $N \times N$ preconditioned system (20), yields

$$
\left( \frac{\|r_k\|_{S_h}}{\|r_0\|_{S_h}} \right)^{1/k} \leq \hat{\varepsilon}_k \quad (k = 1, 2, \ldots, N)
$$

(30)
with \( \hat{\varepsilon}_k := \frac{2}{km^2} \sum_{i=1}^{k} \left( C_1 \min_{H_{i-1} \in H_0^1(\Omega)} \dashv_{H_{i-1}} \max_{V \in H_{i-1}} \frac{\|v\|_{L^2(\Omega)}^2}{\|v\|_{S}^2} + C_2 \min_{H_{i-1} \in H_0^1(\Omega)} \dashv_{H_{i-1}} \max_{V \in H_{i-1}} \frac{\|v\|_{L^p(\Omega)}^2}{\|v\|_{S}^2} \right) \rightarrow 0 \) as \( k \rightarrow \infty \) (where \( H_{i-1} \) stands for an arbitrary \((i-1)\)-dimensional subspace and orthogonality is understood in \( S \)-inner product), and here the constants \( C_1, C_2 > 0 \) and hence the sequence \( (\hat{\varepsilon}_k)_{k \in \mathbb{N}} \) are independent of \( V_h \) and \( u_n \).

**Proof.** We rely on Theorem 5.1 and prove that the sequence \( \varepsilon_k \) in (28)-(29) satisfies \( \varepsilon_k \leq \hat{\varepsilon}_k \) if \( Q_S = Q_S^{(n)} \) as above, further, that \( \hat{\varepsilon}_k \rightarrow 0 \). The divergence theorem yields for \( \mathbf{v}, \mathbf{z} \in H_0^1(\Omega) \)

\[
\int_{\Omega} (\mathbf{b}_i \cdot \nabla v_i) z_i = -\int_{\Omega} v_i (\mathbf{b}_i \cdot \nabla z_i) - \int_{\Omega} (\text{div} \mathbf{b}_i) v_i z_i, \tag{32}
\]

hence from (27) and (3)

\[
\|Q_S^{(n)} \mathbf{v}\|_S = \sup_{\mathbf{v} \in H_0^1(\Omega)} \|Q_S^{(n)} \mathbf{v}, \mathbf{z}\|_S
\]

\[
= \sup_{\mathbf{v} \in H_0^1(\Omega)} \left| \sum_{i=1}^{l} \int_{\Omega} \left( -v_i (\mathbf{b}_i \cdot \nabla z_i) + (\sum_{j=1}^{l} \partial_j f_i(x, u_n) v_j - h_i v_i - (\text{div} \mathbf{b}_i) v_i) z_i \right) \right|
\]

\[
\equiv \sup_{\mathbf{v} \in H_0^1(\Omega)} \left| \int_{\Omega} \left( -\mathbf{v} \cdot (\mathbf{b} \cdot \nabla \mathbf{z}) + (f'_\mathbf{b}(x, u_n) - (\mathbf{h} + \text{div} \mathbf{b}) I ) \mathbf{v} \cdot \mathbf{z} \right) \right|
\]

\[
\leq \sup_{\mathbf{v} \in H_0^1(\Omega)} \left( \max_i \|\mathbf{b}_i\|_{L^\infty(\Omega)} \int_{\Omega} \|v\| \|\nabla \mathbf{z}\|_{L^2(\Omega)^d} + (c_3 + \max_i \|h_i + \text{div} \mathbf{b}_i\|_{L^\infty(\Omega)}) \int_{\Omega} \|\mathbf{v}\|_{L^2(\Omega)} \|\mathbf{z}\|_{L^2(\Omega)} \right)
\]

\[
+ c_4 \int_{\Omega} \|u_n\|_{L^p(\Omega)}^{p-2} \|\mathbf{v}\|_{L^p(\Omega)} \|\mathbf{z}\|_{L^p(\Omega)} \tag{33}
\]

where in the last term Hölder’s inequality has been used for the case \( \frac{p-2}{p} + \frac{1}{p} + \frac{1}{p} = 1 \). Here we have \( \|\nabla \mathbf{z}\|_{L^2(\Omega)^d} = \|\mathbf{z}\|_{H_0^1} \leq \frac{1}{\sqrt{m}} \cdot \|\mathbf{z}\|_S = \frac{1}{\sqrt{m}} \) and, also using (8), \( \|\mathbf{z}\|_{L^p(\Omega)} \leq \frac{C_p}{\sqrt{m}} \cdot \|\mathbf{z}\|_S = \frac{C_p}{\sqrt{m}} \) for all \( p \leq p^* \). Therefore

\[
\|Q_S^{(n)} \mathbf{v}\|_S \leq \left( \frac{1}{\sqrt{m}} \max_i \|\mathbf{b}_i\|_{L^\infty(\Omega)} \|\mathbf{v}\|_{L^2(\Omega)} \right)
\]

\[
+ C_2 \sqrt{m} \left( c_3 + \max_i \|h_i + \text{div} \mathbf{b}_i\|_{L^\infty(\Omega)} \right) \|\mathbf{v}\|_{L^2(\Omega)}^2 + c_4 \frac{C_p}{\sqrt{m}} \|u_n\|_{L^p(\Omega)}^{p-2} \|\mathbf{v}\|_{L^p(\Omega)} \tag{34}
\]

Moreover, from (8) and (14)

\[
\|u_n\|_{L^p(\Omega)} \leq C_p \cdot \|u_n\|_{H^1_0} \leq C_p R_0, \tag{35}
\]
owing to the compactness of the embeddings
where (37) and (36), respectively, we obtain
where and here $K_1, K_2$ are independent of $h$ and $u_n$.

Now setting $v_i = z_i$ in (32),
\[
\int_\Omega (b_i \cdot \nabla v_i) v_i = - \int_\Omega \frac{1}{2} (\text{div } b_i) v_i^2
\]
hence
\[
|\langle Q_S^{(n)} v, v \rangle_S | = \left| \sum_{i=1}^l \int_\Omega ((b_i \cdot \nabla v_i) v_i + \sum_{j=1}^l \partial_j f_i(x, u_n) v_j - h_i v_i) v_i \right|
\]
\[
= \left| \int_\Omega ((b \cdot \nabla v) \cdot v + (f'_i(x, u_n) - hI) v \cdot v) \right|
\]
\[
\leq \int_\Omega \max_i |h_i + \frac{1}{2} \text{div } b_i| ||v||^2 + \int_\Omega (c_3 + c_4 |u_n|^{p-2}) |v|^2
\]
\[
\leq (c_3 + \max_i |h_i + \frac{1}{2} \text{div } b_i|_{L^\infty(\Omega)} ||v||^2_{L^2(\Omega)} + c_4 ||u_n||_{L^p(\Omega)}^{p-2} ||v||^2_{L^p(\Omega)}.
\]
Using (35) again, we obtain
\[
|\langle Q_S^{(n)} v, v \rangle_S | \leq K_3 ||v||^2_{L^2(\Omega)} + K_4 ||v||^2_{L^p(\Omega)}
\]
and here $K_3, K_4$ are independent of $h$ and $u_n$. Now let $H_S = H_0^1(\Omega)^l$ with the $S$-inner product. The variational characterization of the eigenvalues yields
\[
\lambda_i((Q_S^{(n)})^* + Q_S^{(n)}) = \min_{H_{i-1} \subset H_S} \max_{v \perp H_{i-1}} \frac{|\langle (Q_S^{(n)})^* + Q_S^{(n)} v, v \rangle_S |}{||v||^2_S} = 2 \min_{H_{i-1} \subset H_S} \max_{v \perp H_{i-1}} \frac{|\langle Q_S^{(n)} v, v \rangle_S |}{||v||^2_S}
\]
and
\[
s_i(Q_S^{(n)})^2 = \lambda_i((Q_S^{(n)})^* Q_S^{(n)}) = \min_{H_{i-1} \subset H_S} \max_{v \perp H_{i-1}} \frac{|\langle (Q_S^{(n)})^* Q_S^{(n)} v, v \rangle_S |}{||v||^2_S} = \min_{H_{i-1} \subset H_S} \max_{v \perp H_{i-1}} \frac{|Q_S^{(n)} v||^2}{||v||^2_S},
\]
where $H_{i-1}$ stands for an arbitrary $(i-1)$-dimensional subspace. Summing up and using (37) and (36), respectively, we obtain
\[
|\lambda_i((Q_S^{(n)})^* + Q_S^{(n)}) + s_i(Q_S^{(n)})^2 | \leq C_1 \min_{H_{i-1} \subset H_S} \max_{v \perp H_{i-1}} \frac{|\langle (Q_S^{(n)})^* Q_S^{(n)} v, v \rangle_S |}{||v||^2_S} + C_2 \min_{H_{i-1} \subset H_S} \max_{v \perp H_{i-1}} \frac{|Q_S^{(n)} v||^2}{||v||^2_S}
\]
where $C_1 = 2K_3 + K_1, C_2 = 2K_4 + K_2$. Here both terms on the r.h.s. tend to 0 as $i \to \infty$, owing to the compactness of the embeddings $H_0^1(\Omega)^l \subset L^2(\Omega)^l$ and $H_0^1(\Omega)^l \subset L^p(\Omega)^l$. (In particular, the first min-max term gives the reciprocal of the eigenvalues of $S$ in $L^2(\Omega)^l$.) That is, the sequence $(\hat{\epsilon}_k)$ is constant times the arithmetic means of a sequence that tends to zero, hence, as is well-known, $\hat{\epsilon}_k$ itself tends to zero.
Remark 5.1 (Explicit asymptotics for \( \hat{\varepsilon}_k \).) The functions \( u_n \in V_h \ (n \in \mathbb{N}^+) \) and \( u_h \in V_h \) are bounded since they are piecewise polynomials. If they are also uniformly bounded as \( h \to 0 \), which follows e.g. in the case of uniform convergence, then the term (33) can be estimated by \( c_4 (\sup \| u_n \|_{L^\infty(\Omega)}^{p-2} \| v \|_{L^2(\Omega)} \| z \|_{L^2(\Omega)} \) instead of the H"older estimate (34), i.e. this term can also be included in the \( L^2 \)-norm estimates before, and (36) is simply replaced by
\[
\| Q_n^{(n)} v \|_S^2 \leq K'_1 \| v \|_{L^2(\Omega)}^2 \tag{38}
\]
where the constant \( K'_1 \) is independent of \( h \) and \( u_n \). Under our Dirichlet boundary conditions, as pointed out in [4], the numbers \( \varrho_i \) are the reciprocals of the eigenvalues of \( S \) for which \( \varrho_i = O(i^{-2/d}) \) holds [5], hence by an elementary calculation
\[
\hat{\varepsilon}_k \leq O\left(\frac{\log k}{k}\right) \quad \text{if} \quad d = 2 \quad \text{and} \quad \hat{\varepsilon}_k \leq O\left(\frac{1}{k^{2/3}}\right) \quad \text{if} \quad d = 3.
\]

6 Numerical experiments

We have made experiments on the test system
\[
-\Delta u_i + b_i \cdot \nabla u_i + f_i(u_1, \ldots, u_l) = g_i \quad (i = 1, \ldots, l)
\]
\[ u_i|_{\partial\Omega} = 0 \]
on the domain \( \Omega = [0,1] \times [0,1] \), where \( b_i = (1,1)^T \) for all \( i \), and \( f(u) = 4A |u|^2 u \) where \( A \) is the lower triangular part of the constant 1 matrix.

The experiments were carried out in the following way:

- we used Courant elements for the FEM discretization using uniform mesh with width \( h = 1/N \), where \( N \) is the number of subintervals on the interval \([0,1] \times \{0\}\);
- the coordinates of the exact solution were chosen among the functions of form \( u(x,y) = C \cdot x(1-x)y(1-y) \) and \( u(x,y) = C \cdot \sin \pi x \sin \pi y \);
- the stopping criterion was \( \| F_h(u_n) - b_n \| \leq 10^{-6} \);
- the auxiliary problems were solved with FFT;
- we used adaptive damping parameters \( \tau_n \);
- the code was written in Matlab.
We have run the code for the system with \( l = 2, 4, 6 \) equations, respectively. The results were much similar for different \( l \) with a slight increase in number of inner iterations and large increase in computing time.

We present the results for \( l = 4 \) equations, here \( r_n := \| F_h(u_n) - g_h \|_{H^{-1}} \) and \( n_{inn} \) denotes the number of inner iterations. The superlinear phase of the outer DIN iteration starts at the 5th step. The mesh uniform behaviour of the convergence can be observed in both the outer and inner iterations.

<table>
<thead>
<tr>
<th>( n )</th>
<th>( N = 17 )</th>
<th>( N = 33 )</th>
<th>( N = 49 )</th>
<th>( N = 65 )</th>
<th>( N = 81 )</th>
<th>( N = 96 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r_n )</td>
<td>( n_{inn} )</td>
<td>( r_n )</td>
<td>( n_{inn} )</td>
<td>( r_n )</td>
<td>( n_{inn} )</td>
<td>( r_n )</td>
</tr>
<tr>
<td>1</td>
<td>7.3726</td>
<td>1</td>
<td>7.4081</td>
<td>1</td>
<td>7.4151</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>5.3727</td>
<td>1</td>
<td>5.3940</td>
<td>1</td>
<td>5.3982</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>3.4515</td>
<td>2</td>
<td>3.4790</td>
<td>2</td>
<td>3.4845</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>1.3288</td>
<td>1</td>
<td>1.3399</td>
<td>2</td>
<td>1.3421</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>6.6101 ( \cdot 10^{-1} )</td>
<td>2</td>
<td>3.5355 ( \cdot 10^{-1} )</td>
<td>2</td>
<td>3.5561 ( \cdot 10^{-1} )</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>2.3429 ( \cdot 10^{-1} )</td>
<td>2</td>
<td>9.2309 ( \cdot 10^{-2} )</td>
<td>5</td>
<td>9.3523 ( \cdot 10^{-2} )</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>5.7094 ( \cdot 10^{-2} )</td>
<td>5</td>
<td>1.6705 ( \cdot 10^{-2} )</td>
<td>7</td>
<td>1.6983 ( \cdot 10^{-2} )</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>3.5825 ( \cdot 10^{-3} )</td>
<td>17</td>
<td>2.2688 ( \cdot 10^{-3} )</td>
<td>17</td>
<td>2.3033 ( \cdot 10^{-3} )</td>
<td>17</td>
</tr>
<tr>
<td>9</td>
<td>3.3643 ( \cdot 10^{-4} )</td>
<td>24</td>
<td>2.8591 ( \cdot 10^{-4} )</td>
<td>24</td>
<td>2.9181 ( \cdot 10^{-4} )</td>
<td>24</td>
</tr>
<tr>
<td>10</td>
<td>3.5510 ( \cdot 10^{-5} )</td>
<td>23</td>
<td>3.7328 ( \cdot 10^{-5} )</td>
<td>37</td>
<td>3.8277 ( \cdot 10^{-5} )</td>
<td>37</td>
</tr>
<tr>
<td>11</td>
<td>4.4460 ( \cdot 10^{-6} )</td>
<td>41</td>
<td>4.9166 ( \cdot 10^{-6} )</td>
<td>49</td>
<td>5.0674 ( \cdot 10^{-6} )</td>
<td>49</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( CPU ) time(s)</th>
<th>( N = 17 )</th>
<th>( N = 33 )</th>
<th>( N = 49 )</th>
<th>( N = 65 )</th>
<th>( N = 81 )</th>
<th>( N = 96 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N = 17 )</td>
<td>( 1.1822 \cdot 10^{2} )</td>
<td>( 8.2159 \cdot 10^{2} )</td>
<td>( 4.1348 \cdot 10^{3} )</td>
<td>( N = 65 )</td>
<td>( N = 81 )</td>
<td>( N = 96 )</td>
</tr>
<tr>
<td>( n )</td>
<td>( \parallel r_n \parallel )</td>
<td>( n_{inn} )</td>
<td>( \parallel r_n \parallel )</td>
<td>( n_{inn} )</td>
<td>( \parallel r_n \parallel )</td>
<td>( n_{inn} )</td>
</tr>
<tr>
<td>1</td>
<td>7.4176</td>
<td>1</td>
<td>7.4188</td>
<td>1</td>
<td>7.4194</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>5.3997</td>
<td>1</td>
<td>5.4004</td>
<td>1</td>
<td>5.4008</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>3.4865</td>
<td>2</td>
<td>3.4874</td>
<td>2</td>
<td>3.4879</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>1.3429</td>
<td>2</td>
<td>1.3433</td>
<td>2</td>
<td>1.3435</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>3.5636 ( \cdot 10^{-1} )</td>
<td>2</td>
<td>3.5670 ( \cdot 10^{-1} )</td>
<td>2</td>
<td>3.5690 ( \cdot 10^{-1} )</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>3.399 ( \cdot 10^{-2} )</td>
<td>5</td>
<td>9.4167 ( \cdot 10^{-2} )</td>
<td>5</td>
<td>9.4280 ( \cdot 10^{-2} )</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>1.7084 ( \cdot 10^{-2} )</td>
<td>7</td>
<td>1.7132 ( \cdot 10^{-2} )</td>
<td>7</td>
<td>1.7158 ( \cdot 10^{-2} )</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>2.3158 ( \cdot 10^{-3} )</td>
<td>18</td>
<td>2.3217 ( \cdot 10^{-3} )</td>
<td>18</td>
<td>2.3249 ( \cdot 10^{-3} )</td>
<td>18</td>
</tr>
<tr>
<td>9</td>
<td>2.9276 ( \cdot 10^{-4} )</td>
<td>24</td>
<td>2.9376 ( \cdot 10^{-4} )</td>
<td>24</td>
<td>2.9430 ( \cdot 10^{-4} )</td>
<td>24</td>
</tr>
<tr>
<td>10</td>
<td>3.9288 ( \cdot 10^{-5} )</td>
<td>37</td>
<td>3.9456 ( \cdot 10^{-5} )</td>
<td>37</td>
<td>3.9548 ( \cdot 10^{-5} )</td>
<td>37</td>
</tr>
<tr>
<td>11</td>
<td>5.2105 ( \cdot 10^{-6} )</td>
<td>49</td>
<td>5.2372 ( \cdot 10^{-6} )</td>
<td>49</td>
<td>5.2519 ( \cdot 10^{-6} )</td>
<td>49</td>
</tr>
<tr>
<td>( CPU ) time(s)</td>
<td>( 1.2864 \cdot 10^{4} )</td>
<td>( 3.0766 \cdot 10^{3} )</td>
<td>( 6.2980 \cdot 10^{3} )</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
References


