Preprint
UCRL-JC-148376

On Mesh-Independent Convergence of an Inexact Newton-Multigrid Algorithm

Peter N. Brown, Panayot Vassilevski, and Carol S. Woodward.

This article was submitted to
SIAM Journal on Scientific Computing

U.S. Department of Energy

Lawrence Livermore National Laboratory

May 2002

ON MESH–INDEPENDENT CONVERGENCE OF AN INEXACT NEWTON–MULTIGRID ALGORITHM

PETER N. BROWN, PANAYOT S. VASSILEVSKI, AND CAROL S. WOODWARD

Abstract. In this paper we revisit and prove optimal order and mesh–independent convergence of an inexact Newton method where the linear Jacobian systems are solved with multigrid techniques. This convergence is shown using Banach spaces and the norm, $\max\{\|\cdot\|_1,\ \|\cdot\|_{0,\infty}\}$, a stronger norm than is used in previous work. These results are valid for a class of second order, semi–linear, finite element, elliptic problems posed on quasi–uniform grids. Numerical results are given which validate the theory.

1. Introduction

In this paper we revisit the problem of mesh independence and optimal order convergence of inexact Newton MG (multigrid) methods for solving finite element, second order, nonlinear, elliptic equations. Previous work has shown mesh-independent convergence of exact Newton methods for such equations discretized by finite differences [1, 2]. In this work, we relax the need for an exact Newton method and show mesh-independent convergence for an inexact Newton method in which multigrid techniques are used to solve the linear Jacobian systems. This work is valid for these equations discretized by finite element methods on quasi–uniform grids. In order to prove such a convergence result, one needs to uniformly control the maximum (L_∞) norm of the iterates with respect to the mesh parameter $h \mapsto 0$. That is, classical convergence results, such as in [5], do not apply directly. Alternatively, one could exploit sophisticated L_p–estimates as in [15], see also [3]. To simplify the presentation we present our results for a model semi–linear second order elliptic problem which allows us to avoid the L_p estimates. In what follows $\|\cdot\|_s$ stands for the Sobolev space H^s–norms, and $H_0^1 = H_0^1(\Omega)$ is the subspace of $H^1 = H^1(\Omega)$ for a given polygonal domain Ω with vanishing traces on $\partial\Omega$.

The model problem of interest has the form: *Find $u \in H_0^1(\Omega) \cap L_\infty(\Omega)$ which satisfies*

$$(\mathcal{L}(u), \varphi) \equiv \int_\Omega [a(x)\nabla u \cdot \nabla\varphi + f(x,\ u)\varphi]\ dx = 0, \quad \textit{for all } \varphi \in H_0^1(\Omega), \tag{1.1}$$

Date: July 19, 1999, January 9, 2002 – beginning; Today is July 23, 2002.

1991 *Mathematics Subject Classification.* 65F10, 65N20, 65N30.

Key words and phrases. inexact Newton multigrid, semi–linear elliptic problems, finite elements, optimal order and mesh independent convergence.

This work was performed under the auspices of the U.S. Department of Energy by University of California Lawrence Livermore National Laboratory under contract No. W-7405-Eng-48.

where the nonlinear function f is sufficiently smooth, and the coefficient function $a = a(x)$ is bounded with $a(x) \geq a_0 = \text{const} > 0$ in Ω. The coefficient a can also be a symmetric positive definite matrix $(a_{ij}(x))$. Later for simplicity we will assume that $f(x, u) = -f - b(u)$ where $f \in L_2(\Omega)$ and $b_u(v) \leq 0$ for $v \in L_\infty(\Omega)$.

The existing convergence proofs (mostly their assumptions) for inexact Newton methods (cf. [4], [10], [13], [16]) do not apply directly, since we need convergence in a stronger norm than is typically considered. For example, in [4] a Hilbert space setting is used, and it is assumed that the nonlinear operator has a Jacobian that is Lipschitz continuous in a strong norm (as a mapping from $\| \cdot \|_1 \mapsto \| \cdot \|_1$ in the particular applications). We instead use a Banach space setting and incorporate the L_∞ norm of the functions in the norm of the nonlinear operator range space. We note that for the above model semi–linear elliptic problem, one can actually prove convergence of inexact Newton MG methods in H^1 (without using maximum norms) by modifying the argument found in [11]. However, we are able to prove more, namely, optimal convergence in the stronger norm, $\max\{\| \cdot \|_1, \ \| \cdot \|_{L_\infty}\}$, exploiting only the Banach (and not Hilbert) space setting.

The purpose of this paper is to make the statement of assumptions and their verification in proper norms that guarantee convergence of a modified inexact Newton MG algorithm independent of mesh parameter and with optimal work per iteration. The results we prove show that if one applies a W–cycle with sufficiently many smoothing iterations for computing inexact Newton directions the resulting method converges linearly with mesh-independent rate of convergence and optimal cost per iteration. The same results hold for a cascadic multigrid iteration (a coarse–to–fine cycle with conjugate gradient smoothing and where the number of smoothing iterations grows geometrically from the fine to coarse levels). Actually, our numerical experiments for some model test problems show that even a standard V–cycle MG provides optimal and mesh–independent convergence of the resulting inexact Newton MG method.

The remainder of the paper is as follows. In Section 2 we formulate a fairly general modified inexact Newton algorithm, state abstract assumptions, and prove its local convergence. In Section 3 we study the discretized nonlinear problem and in the next section verify the assumptions for our model second order semi–linear elliptic problem. In Section 5 we prove that one sufficiently accurate, cascadic MG–cycle or W–cycle, MG step can be applied to compute the inexact Newton iterate. Thus the cost per iterate is optimal (of order of the number of degrees of freedom). The following section contains some numerical illustrations of the convergence behavior of the method. The last section makes some concluding remarks and extensions. Finally, in an Appendix we summarize some results regarding MG convergence of the residuals (not the iterates). The presented analysis assumes full (H^2) regularity of the linearized second order elliptic problem.

2. Problem formulation

Consider the nonlinear system

$$F(u) = 0, \tag{2.1}$$

where $F: \mathcal{X} \mapsto \mathcal{Y}$, $\mathcal{X}$ and both $\mathcal{Y}$ are Banach spaces. We will denote their norms with the same symbol $\|\cdot\|$ since it will be clear from the context which one is actually meant. The induced operator norms for mappings between $\mathcal{X}$ and $\mathcal{Y}$ and vice versa will also use the same symbol $\|\cdot\|$ without causing any ambiguity. In our application the spaces $\mathcal{X} = \mathcal{X}_h$ and $\mathcal{Y} = \mathcal{Y}_h$ will depend on a mesh parameter, but their norms will be induced from an infinite dimensional space (hence will be mesh independent). Our approach then will be to verify the assumptions stated below with mesh–independent constants. This will imply mesh–independent convergence at the end. The key part in the presented theory is that one needs to construct inexact Newton iterates (see Algorithm 2.1 below) where the "inexactness" is controlled by a mesh–independent tolerance η achievable with an optimal cost. We later show this tolerance is achieved with optimal cost for certain MG cycles (see Section 5). Another key point in the theory is the assumption on the initial iterate; namely, that it is feasible to find an initial iterate which is close to the discrete solution in a stong (residual) norm and provide a constructive (practical) algorithm to compute it. This is verified in Theorem 5.1.

We now state the main assumptions:

Assumption 2.1.

(A1) *there is $u^\star \in \mathcal{X}$ such that $F(u^\star) = 0$;*

(A2) *for any u in a neighborhood of $u^\star$ there is a linear mapping $F'(u): \mathcal{X} \mapsto \mathcal{Y}$ such that for any small $\epsilon > 0$ there is a $\delta > 0$ for which*

$$\|F(u) - F(u^\star) - F'(u^\star)(u - u^\star)\| \le \epsilon \|u - u^\star\|$$

whenever $\|u - u^\star\| < \delta$.

(A3) *the derivative $F'(u)$ is invertible and $(F'(u))^{-1}$ is a bounded linear operator $\mathcal{Y} \mapsto \mathcal{X}$, for any u in a neighborhood of $u^\star$, that is,*

$$\|(F'(u))^{-1}\| \le \mu, \tag{2.2}$$

for some constant μ. In addition, we assume that the mapping $(F'(u))^{-1}$ is continuous in u (in a neighborhood of $u^\star$). That is, for any $\epsilon > 0$ there is a $\delta > 0$ such that

$$\|I - F'(u^\star)\,(F'(u))^{-1}\| < \epsilon,$$

and

$$\|I - (F'(u))^{-1}\,F'(u^\star)\| < \epsilon,$$

whenever $\|u - u^\star\| < \delta$. We note that implicit in this assumption is the fact that $F'(u)$ is one-to-one and onto as a mapping from $\mathcal{X} \mapsto \mathcal{Y}$ whenever $\|u - u^\star\| < \delta$.

We consider the modified inexact Newton algorithm given below.

Algorithm 2.1 (Modified Inexact Newton method)**.**

Consider the sequence of iterates $\{u^k\}$ *generated by*

$$\begin{array}{ll} (i) & \text{choose an initial guess } u^0; \\ (ii) & \text{For } k = 0, 1, \dots \text{ until convergence} \\ & \quad \text{- find an } s^k \text{ such that} \\ & \qquad F^{'}(u^0)s^k = -F(u^k) + r^k, \ \text{where } \|r^k\| \leq \eta\|F(u^k)\|; \\ & \quad \text{- set } u^{k+1} = u^k + s^k. \end{array} \tag{2.3}$$

One may also assume:

(A4) the residuals r^k satisfy for an $\eta < 1$, (in addition to $\|r^k\| \leq \eta\|F(u^k)\|$), the estimate

$$\left\|(F^{'}(u^\star))^{-1}r^k\right\| \leq \eta \ \left\|(F^{'}(u^\star))^{-1}F(u^k)\right\|. \tag{2.4}$$

The question is *when will such an iteration converge?* We show the following result.

Theorem 2.1. *Assume assumptions (A1)–(A3) hold and let* η*,* t *satisfying* $0 \leq \eta < t < 1$ *be given. Then there is an* $\epsilon > 0$*, such that, if* $\|u^0 - u^\star\| < \epsilon$*, then the sequence of iterates* $\{u^k\}$ *generated by* (2.3) *converges to* $u^\star$*. Moreover, the convergence is linear in the sense that*

$$\|u^{k+1} - u^\star\|_\star \leq t\|u^k - u^\star\|_\star, \tag{2.5}$$

where $\|v\|_\star = \|F^{'}(u^\star)v\|$ *provided that the initial iterate* u^0 *satisfies the estimate*

$$\mu\|F^{'}(u^\star)(u^0 - u^\star)\| < \epsilon, \tag{2.6}$$

where μ *is from (2.2). If assumption (A4) also is satisfied, then the following convergence estimate in the original norm holds:*

$$\|u^{k+1} - u^\star\| \leq t\|u^k - u^\star\|. \tag{2.7}$$

Proof. The proof follows the lines of the proof given in [10] (for $\mathcal{X} = \mathcal{Y}$) and [16]. From (A3) and (2.2), we have

$$\|v\| \leq \mu \ \|v\|_\star, \quad \text{for all } v \in \mathcal{X}. \tag{2.8}$$

Since $0 < \eta < t$, there is a $\gamma > 0$ such that

$$\gamma + \mu\gamma(\gamma + 1) + (\gamma + 1)\eta(1 + \mu\gamma) < t. \tag{2.9}$$

Based on the properties of $F^{'}$ and $(F^{'})^{-1}$, $(A2)$ and $(A3)$, now choose an $\epsilon > 0$ sufficiently small such that

$$\|I - (F^{'}(v))^{-1}F^{'}(u^\star)\| < \gamma, \tag{2.10}$$

$$\|I - (F^{'}(u^\star))(F^{'}(v))^{-1}\| < \gamma, \text{ and} \tag{2.11}$$

$$\|F(v) - F(u^\star) - F^{'}(u^\star)(v - u^\star)\| \leq \gamma\|v - u^\star\|, \tag{2.12}$$

all three estimates holding if $\|v - u^\star\| < \epsilon$.

Assume $\|u^0 - u^\star\| < \epsilon$. The proof proceeds by induction. Since $F^{'}(u^0)$ is one-to-one and onto by assumption (A3), the system $F^{'}(u^0)s = -F(u^0)$ has a solution and it is possible to find an s^0 such that $F^{'}(u^0)s^0 = -F(u^0) + r^0$ with $\|r^0\| \le \eta\|F(u^0)\|$. We then define u^1 by

$$u^1 = u^0 - P^{-1}F(u^0) + P^{-1}r^0,$$

where $P = F^{'}(u^0)$. Since $\|u^0 - u^\star\| < \epsilon$, P^{-1} exists and (2.10) (2.11) as well as (2.12) hold for $v = u^0$.

Next,

$$u^1 - u^\star = u^0 - u^\star - P^{-1}F(u^0) + (P^{-1} - (F^{'}(u^\star))^{-1})r^0 + (F^{'}(u^\star))^{-1}r^0.$$

Replace $F(u^0)$ by $F^{'}(u^\star)(u^0 - u^\star) + \left[F(u^0) - F(u^\star) - F^{'}(u^\star)(u^0 - u^\star)\right]$. Then

$$\begin{aligned} u^1 - u^\star &= u^0 - u^\star - P^{-1}F^{'}(u^\star)(u^0 - u^\star) - P^{-1}\left[F(u^0) - F(u^\star) - F^{'}(u^\star)(u^0 - u^\star)\right] \\ &\quad + \left(P^{-1} - (F^{'}(u^\star))^{-1}\right) r^0 + (F^{'}(u^\star))^{-1}r^0. \end{aligned} \tag{2.13}$$

So

$$\begin{aligned} F^{'}(u^\star)(u^1 - u^\star) &= \left[I - F^{'}(u^\star)P^{-1}\right] F^{'}(u^\star)(u^0 - u^\star) \\ &\quad - F^{'}(u^\star)P^{-1}\left[F(u^0) - F(u^\star) - F^{'}(u^\star)(u^0 - u^\star)\right] \\ &\quad + \left(F^{'}(u^\star)P^{-1} - I\right) r^0 + r^0. \end{aligned}$$

Therefore

$$\begin{aligned} \|u^1 - u^\star\|_\star &\le \|I - F^{'}(u^\star)P^{-1}\|\|u^0 - u^\star\|_\star + \|F^{'}(u^\star)P^{-1}\|\ \gamma\mu\ \|u^0 - u^\star\|_\star \\ &\quad + \|F^{'}(u^\star)P^{-1}\|\ \|r^0\| \\ &\le \gamma\ \|u^0 - u^\star\|_\star + \mu\gamma\ (\gamma + 1)\ \|u^0 - u^\star\|_\star + (\gamma + 1)\|r^0\|. \end{aligned} \tag{2.14}$$

Since

$$\begin{aligned} \|r^0\| &\le \eta\|F(u^0)\| \\ &\le \eta\left(\|F(u^0) - F(u^\star) - F^{'}(u^\star)(u^0 - u^\star)\| + \|F^{'}(u^\star)(u^0 - u^\star)\|\right) \\ &\le \eta\ (\gamma\|u^0 - u^\star\| + \|u^0 - u^\star\|_\star) \\ &\le \eta\ (1 + \gamma\mu)\|u^0 - u^\star\|_\star, \end{aligned}$$

we have from (2.14) that

$$\begin{aligned} \|u^1 - u^\star\|_\star &\le \left[\gamma + \mu\gamma(\gamma + 1) + (\gamma + 1)\eta(1 + \mu\gamma)\right]\|u^0 - u^\star\|_\star \\ &\le t\|u^0 - u^\star\|_\star. \end{aligned}$$

Next, from (2.13) and if we assume (2.4), one has

$$\begin{aligned}\|u^1 - u^\star\| &\leq \|I - P^{-1}F^{'}(u^\star)\|\|u^0 - u^\star\| \\ &\quad + \|P^{-1}F^{'}(u^\star)\| \left[\|(F^{'}(u^\star))^{-1}\|\|F(u^0) - F(u^\star) - F^{'}(u^\star)(u^0 - u^\star)\| \right.\\ &\qquad\qquad\qquad \left. + \|(F^{'}(u^\star))^{-1}r^0\|\right] \\ &\leq [\gamma + \mu(1+\gamma)\gamma]\,\|u^0 - u^\star\| + (1+\gamma)\eta\,\|(F^{'}(u^\star))^{-1}F(u^0)\| \\ &\leq [\gamma + \mu(1+\gamma)\gamma]\,\|u^0 - u^\star\| \\ &\quad + (1+\gamma)\left(\eta\,\mu\|F(u^0) - F(u^\star) - F^{'}(u^\star)(u^0 - u^\star)\| + \eta\,\|u^0 - u^\star\|\right) \\ &\leq [\gamma + \mu(1+\gamma)\gamma + \eta(\mu\gamma + 1)(1+\gamma)]\,\|u^0 - u^\star\| \\ &\leq t\|u^0 - u^\star\| < \epsilon,\end{aligned}$$

using the fact that

$$(F^{'}(u^\star))^{-1}F(u^0) = (F^{'}(u^\star))^{-1}[F(u^0) - F(u^\star) - F^{'}(u^\star)(u^0 - u^\star)] + (u^0 - u^\star).$$

However, in some cases we cannot ensure the estimate (2.4) with an optimal cost. Thus, only $\|r^k\| \leq \eta\,\|F(u^k)\|$ is guaranteed. In such a case we have assumed that the initial iterate satisfies estimate (2.6) (in addition to $\|u^0 - u^\star\| < \epsilon$). The latter estimate guarantees that all iterates u^k, $k \geq 1$, are in the ball $\|u^k - u^\star\| < \epsilon$ and the induction argument works. Indeed,

$$\|u^k - u^\star\| \leq \mu\,\|u^k - u^\star\|_\star \leq \mu t^k\|u^0 - u^\star\|_\star \leq \mu\|u^0 - u^\star\|_\star < \epsilon.$$

Hence, in either case, (2.10)–(2.12) hold for $v = u^k$, $k > 0$, and the proof can be completed by induction. □

Corollary 2.1. *One can rewrite step (ii) of Algorithm 2.1 in the following more traditional form:*

$$F^{'}(u^k)s^k = -F(u^k) + \widehat{r}^k, \tag{2.15}$$

where $\widehat{r}^k = r^k + \left(F^{'}(u^k) - F^{'}(u^0)\right)s^k$*. Based on Theorem 2.1 one has the estimate*

$$\begin{aligned}\|\widehat{r}^k\| &\leq \|r^k\| + \|\left(F^{'}(u^\star) - F^{'}(u^0)\right)s^k\| + \|\left(F^{'}(u^k) - F^{'}(u^\star)\right)s^k\| \\ &\leq \|r^k\| + 2\gamma\|F^{'}(u^\star)s^k\| \\ &\leq \|r^k\| + 2\gamma\|F^{'}(u^\star)\left(F^{'}(u^0)\right)^{-1}\|\,\|F^{'}(u^0)s^k\| \\ &\leq \|r^k\| + 2\gamma(1+\gamma)\,\|F^{'}(u^0)s^k\| \\ &= \|r^k\| + 2\gamma(1+\gamma)\,\|r^k - F(u^k)\| \\ &\leq (\eta + 2\gamma(1+\gamma)(\eta+1))\,\|F(u^k)\|.\end{aligned}$$

That is, for sufficiently small γ *one can guarantee an estimate of the form*

$$\|\widehat{r}^k\| \leq \widehat{\eta}\,\|F(u^k)\|, \tag{2.16}$$

with an $\widehat{\eta} < 1$*. Note that the argument goes both ways. If one computes inexact Newton directions* s^k *based on* (2.15) *such that* (2.16) *holds, then it is equivalent to think that* s_k *has been computed as in step (ii) of Algorithm 2.1 with an* $\eta = \widehat{\eta} + 2\gamma(1+\gamma)(\widehat{\eta}+1)$ *assuming that* u^k *and* u^0 *are sufficiently close to* $u^\star$ *such that (A3) holds. Therefore, Theorem* (2.1) *holds with the more traditional version of the inexact Newton method (that is, inexact Newton direction computed as in* (2.15) *satisfying* (2.16)*).*

3. Application to second order semi–linear elliptic problems

We will be interested in the following model second order semi–linear elliptic problem: Find u such that

$$\mathcal{L}u \equiv \sum_{i,j} -\frac{\partial}{\partial x_i}\left(a_{ij}(x)\frac{\partial u}{\partial x_j}\right) - b(x,u) = f(x), \text{ in } \Omega \subset \mathcal{R}^d, \tag{3.1}$$

subject to homogeneous Dirichlet boundary conditions $u = 0$ on $\partial\Omega$. We denote the solution to (3.1) in what follows by $u^\star$.

We define $\mathcal{L}: \mathcal{V} \equiv H^1_0(\Omega) \mapsto \mathcal{V}' = H^{-1}(\Omega)$. Let $V_h \subset \mathcal{V}$ be a finite element space of continuous piecewise polynomials such that $\overline{\cup\{V_h,\ h \le h_0\}} = \mathcal{V}$. We also assume that the corresponding mesh is quasi-uniform, which implies certain inverse inequalities for the finite element functions. In particular, for a uniform constant C, we will have $\|v\|_1 \le Ch^{-1}\|v\|_0$ for all $v \in V_h$.

The Galerkin operators induced by $\mathcal{L}$ on V_h, denoted by $Av - b(v)$, are defined as follows

$$(Av - b(v),\ \varphi) = \sum_{i,j}(a_{i,j}(x)\frac{\partial v}{\partial x_i},\ \frac{\partial \varphi}{\partial x_j}) - (b(x,v),\ \varphi), \quad \text{for all } v,\ \varphi \in V_h.$$

Similarly, the Jacobian of $Av - b(v)$ at v, a linear operator denoted by $A - b_u(v)$ ($b_u(v) \equiv \frac{\partial b(v)}{\partial u}$), is defined by

$$(A\xi - b_u(v)\xi,\ \varphi) = \sum_{i,j}(a_{i,j}(x)\frac{\partial \xi}{\partial x_i},\ \frac{\partial \varphi}{\partial x_j}) - (\frac{\partial b(x,v)}{\partial u}\xi,\ \varphi), \quad \text{for all } \xi,\ \varphi \in V_h.$$

The ellipticity assumption means that the principal part, A, is H^1–bounded and coercive; that is, for two uniform constant $0 < a_0 \le a_1$ the following estimates hold,

$$\begin{aligned} a_0\|\psi\|_1^2 &\le (A\psi,\ \psi), && \text{for all } \psi \in H^1_0(\Omega),\\ (A\psi,\ \varphi) &\le a_1\|\psi\|_1\|\varphi\|_1, && \text{for all } \psi,\ \varphi \in H^1_0(\Omega). \end{aligned} \tag{3.2}$$

This holds if the coefficient matrix $(a_{ij}(x))$ is symmetric, positive definite, and bounded from above uniformly in Ω.

The discrete counterpart of (3.1) reads: find $u_h \in V_h$ such that

$$\begin{aligned} (L_h u_h,\ \varphi) &\equiv (Au_h - b(u_h),\ \varphi)\\ &\equiv \sum_{i,j}(a_{i,j}(x)\frac{\partial u_h}{\partial x_i},\ \frac{\partial \varphi}{\partial x_j}) - (b(x,u_h),\ \varphi) = (f,\ \varphi), \quad \text{for all } \varphi \in V_h. \end{aligned} \tag{3.3}$$

In the notation of the previous section, the discrete nonlinear problem (3.3) can be rewritten as

$$F_h(u_h) \equiv L_h(u_h) - Q_h f = 0, \tag{3.4}$$

where $Q_h:\ L_2(\Omega) \mapsto V_h$ is the L_2–projection. In what follows we denote the solution of (3.4) by $u_h^\star$. The derivative $F_h'(v)$, $v \in V_h$, is defined variationally as

$$(F_h'(v)\psi,\ \varphi) \equiv (A\psi - b_u(v)\ \psi,\ \varphi), \quad \text{for all } \psi,\ \varphi \in V_h.$$

We next define the Banach spaces $\mathcal{X}$ and $\mathcal{Y}$. Note that these depend on the mesh parameter $h \mapsto 0$, and that the respective nonlinear operator and its derivative depend on h. To indicate this we use an h subscript, i.e., $F = F_h$ and $F^{'} = F_h^{'}$. We note that $F_h^{'} = Q_h F^{'}$ when restricted to V_h.

Definition 3.1 (Discrete Banach spaces)**.**

- *$\mathcal{X} = V_h$ with a norm $\|\cdot\|$, such that $\|\psi\| \leq C\|F^{'}(u_h^\star)\psi\|_0$ for any $\psi \in V_h$. We will later show that under certain regularity assumptions (see below* (3.1)*), as is well–known, one has with a mesh–independent constant C,*

$$max\{\|\psi\|_1,\ \|\psi\|_{L_\infty}\} \leq C\|F_h^{'}(u_h^\star)\psi\|_0,$$

  so the latter norm is one possible candidate.
- *$\mathcal{Y} = V_h$ equipped with $\|\cdot\| = \|\cdot\|_0$.*

Above, $\|\cdot\|_s$ stands for the Sobolev space H^s–norms, and L_∞ on V_h is actually the maximum norm (since the functions in V_h are continuous).

If one chooses $\|\psi\| = \|F^{'}(u_h^\star)\psi\|_0$, for a norm in $\mathcal{X}$, it is obvious that

$$\mu \equiv \max\{\|F^{'}(u_h^\star)\|,\ \|(F^{'}(u_h^\star))^{-1}\|\} = 1.$$

Note that this is a mesh–dependent norm. In what follows, however, we will use the (mesh–independent) norm in $\mathcal{X}$

$$\|\psi\| \equiv \max\{\|\psi\|_1,\ \|\psi\|_\infty\}.$$

Assumption 3.1. *We assume that the nonlinear boundary value problem has a solution $u \in H_0^1(\Omega) \cap H^2(\Omega)$. The latter implies that $u \in L_\infty(\Omega)$ as well.*

In addition, we assume:

(i) Ω is a bounded convex polygon and the principal linear elliptic part A of $\mathcal{L}$ is H^2–regular; that is, for any $g \in L^2(\Omega)$ the solution of the linear boundary value problem

$$(Aw,\ v) = (g,\ v) \quad \text{for all } v \in H_0^1(\Omega),$$

satisfies the *a priori* estimate for a constant $C_R > 0$ (independent of the r.h.s. g),

$$\|w\|_2 \leq C_R\|g\|_0.$$

(ii) The function $b(x,\ u)$ is continuously differentiable; that is, $\frac{\partial b(x,\ v)}{\partial u}$ exists near the exact solution u of (3.1) and is uniformly Lipschitz in $x \in \Omega$ as a function of $v \in \mathcal{R}$ (in a neighborhood of u). The Lipschitz constant is denoted by L in what follows.

(iii) The function $|\frac{\partial b(x,\ u)}{\partial u}|$ is bounded in Ω.

(iv) Finally, we assume that $\frac{\partial b(x,\ v)}{\partial u} \leq 0$, again, in a neighborhood of u.

Lemma 3.1. *Under the assumptions (i)–(iv), the discrete problem* (3.4) *has a unique solution $u_h^\star$. Moreover the following error estimates hold:*

$$\|u^\star - u_h^\star\|_1 \leq Ch\|u\|_2,\ \|u^\star - u_h^\star\|_\infty \leq Ch^\alpha\|u\|_2,$$

for some positive α (< 1).

Proof. The proof is based on a standard argument (see, Appendix I). □

Next, we prove some auxiliary results.

Lemma 3.2. *Consider the linear boundary value problem, for a given* $y \in L_\infty(\Omega)$ *such that* $\|y - u^\star\|_\infty \leq \delta$ *for* $\delta > 0$,

$$(Aw - b_u(y)w,\ v) = (g,\ v), \quad \textit{for all } v \in H_0^1(\Omega).$$

Then, $w \in H^2(\Omega) \cap H_0^1(\Omega)$ *and the following a priori estimate holds*

$$\|w\|_2 \leq C_\delta \|g\|_0, \quad C_\delta = C_R(1 + L\delta + \|b_u(u^\star)\|_\infty).$$

Proof. It is clear that with $c_0 :\ c_0\|w\|_1^2 \leq (Aw,\ w)$ and c_F the Poincaré constant $c_F\|w\|_0^2 \leq \|w\|_1^2$, the following obvious estimate holds (recall that $b_u(y) \leq 0$),

$$c_0 c_F \|w\|_0^2 \leq (Aw,\ w) \leq (g,\ w) \leq \|g\|_0 \|w\|_0,$$

that is,

$$\|w\|_0 \leq C\|g\|_0,$$

with a uniform constant C. Then, since w solves the linear problem

$$Aw = q \equiv g + b_u(y)w \in L_2(\Omega),$$

using its regularity, we obtain the estimate,

$$\begin{aligned} \|w\|_2 &\leq C\|g + b_u(y)w\|_0 \\ &\leq C\left[\|g\|_0 + \|b_u(y)\|_\infty \|w\|_0\right] \\ &\leq C\left[1 + \|b_u(y)\|_\infty\right] \|g\|_0. \end{aligned}$$

Next, we have assumed that $\|y - u^\star\|_\infty < \delta$, for a given δ, so that

$$\begin{aligned} \|b_u(y)\|_\infty &\leq L\|y - u^\star\|_\infty + \|b_u(u^\star)\|_\infty \\ &\leq C_\delta. \end{aligned}$$

The constant C_δ depends on L, the Lipschitz constant of $b_u(\cdot)$, and on the L_∞ bound of $b_u(u^\star)$, all fixed in our application. □

Lemma 3.3. *Let* $v \in V_h$ *be such that* $\|v - u_h^\star\| \leq \delta$ *for* $u_h^\star$ *the exact solution of the discrete nonlinear problem (3.4). Then the solution of the discrete linear problem: given* $r \in L_2(\Omega)$ *find* $\psi \in V_h$ *such that*

$$(F_h^{'}(v)\psi,\ \varphi) \equiv (A\psi - b_u(v)\psi,\ \varphi) = (r,\ \varphi), \quad \textit{for all } \varphi \in V_h.$$

satisfies the a priori estimate,

$$\max\{\|\psi\|_1,\ \|\psi\|_\infty\} \leq C\|r\|_0,$$

with a uniform constant C.

Proof. Note that based on Lemma (3.1) $u_h^\star$ exists and approximates the continuous solution $u^\star$ (for sufficiently small h). Then by assumption $\|v - u_h^\star\| < \delta$. So, $\|v - u^\star\| < \delta_1 \equiv \delta + Ch^\alpha \|u\|_2$. That is, v is in a small neighborhood of $u^\star$. For δ_1 small enough (that is, for δ and h small enough) assumptions 3.1 (i)–(iv) hold.

The first estimate $\|\psi\|_1 \leq \frac{C_F}{c_0}\|r\|_0$ is standard, using the ellipticity assumption (3.2), $b_u(v) \leq 0$ (assumption (iv)) and the Poincaré inequality; that is,

$$\begin{aligned} a_0\|\psi\|_1^2 &\leq (A\psi,\ \psi) \\ &\leq (F_h^{'}(v)\psi,\ \psi) \\ &= (r,\ \psi) \\ &\leq \|r\|_0\|\psi\|_0 \\ &\leq C_F\|r\|_0\|\psi\|_1. \end{aligned}$$

Consider now the second order linear elliptic problem,

$$(Aw - b_u(v)w,\ v) = (r,\ v), \qquad \text{for all } v \in H_0^1(\Omega).$$

We proved that $\|w\|_2 \leq C\|r\|_0$ (see Lemma 3.2) with a uniform constant C. Then the following error estimate is standard

$$\|\psi - I_h w\|_1 \leq Ch\|w\|_2 \leq Ch\|r\|_0.$$

Here, I_h is the nodal interpolation operator. Using the well-known inverse inequality for finite element functions, (valid since we have assumed a quasi-uniform mesh), $(\Omega \subset \mathcal{R}^d,\ d = 2, 3)$,

$$\|\psi - I_h w\|_\infty \leq C|\log h|^{1-\frac{1}{d}} h^{1-\frac{d}{2}} \|\psi - I_h w\|_1,$$

one gets

$$\begin{aligned} \|\psi\|_\infty &\leq \|w\|_\infty + \|\psi - I_h w\|_\infty \\ &\leq C\|w\|_2 + C(|\log h|^{1-\frac{1}{d}} h^{1-\frac{d}{2}})\|\psi - I_h w\|_1 \\ &\leq C\|w\|_2 + C(|\log h|^{1-\frac{1}{d}} h^{1-\frac{d}{2}})h\|w\|_2 \\ &\leq C\|w\|_2 \\ &\leq C\|r\|_0. \end{aligned}$$

□

4. Verifying assumptions (A1)–(A3)

Next we verify the main assumptions (A1), (A2), and (A3) of 2.1. We first note that assumption (A1) was verified in Lemma 3.1.

To verify (A2) it is equivalent to prove the following lemma.

Lemma 4.1. *For any $\varphi \in V_h$, the following estimate holds*

$$(F_h(v) - F_h(u_h^\star) - F_h^{'}(u_h^\star)(v - u_h^\star),\ \varphi) \leq \gamma\|u_h^\star - v\|_\infty\|\varphi\|_0.$$

if $\|v - u_h^\star\| < \delta$ and δ is sufficiently small.

Proof. One has,

$$\begin{aligned}(F_h(v) - F_h(u_h^\star) - F_h^{'}(u_h^\star)(v-u_h^\star),\ \varphi) &= (-b(x,\ v) + b(x,\ u_h^\star) - b_u(u_h^\star)(u_h^\star - v),\ \varphi)\\ &\leq L\|u_h^\star - v\|_\infty \|u_h^\star - v\|_0 \|\varphi\|_0\\ &\leq L\|u_h^\star - v\|_\infty\, C_F\, \|u_h^\star - v\|_1 \|\varphi\|_0\\ &\leq \gamma \|u_h^\star - v\|_1 \|\varphi\|_0.\end{aligned}$$

which can be ensured if given δ sufficiently small. □

Lemma 4.2. *The term*

$$\mu = \sup_{r\in L_2(\Omega)} \frac{\max\{\|(F_h^{'}(v))^{-1}Q_h r\|_1,\ \|(F_h^{'}(v))^{-1}Q_h r\|_\infty\}}{\|r\|_0}$$

is uniformly bounded in terms of $h \mapsto 0$ for $v \in V_h$ such that $\|v - u_h^\star\| \leq \delta$.

Proof. This result follows from Lemma 3.3. □

This verifies estimate (2.2) in (A3). The remaining part of (A3) requires continuity of the derivatives and of their inverses. We show the equivalent estimates:

Lemma 4.3. *For any $w \in V_h$,*

$$\|(F_h^{'}(u_h^\star) - F_h^{'}(v))w\|_0 \leq \gamma \|F_h^{'}(v)w\|_0, \tag{4.1}$$

and

$$\|((F_h^{'}(u_h^\star))^{-1} - (F_h^{'}(v))^{-1})w\| \leq \gamma \|(F_h^{'}(u_h^\star))^{-1}w\|, \tag{4.2}$$

if $v \in V_h$ is such that $\|u_h^\star - v\| < \delta$ for δ sufficiently small.

Proof. Given $w \in L_2(\Omega)$, let $p_h \in V_h$ and $q_h \in V_h$ solve the linear problems,

$$(F_h^{'}(u_h^\star)q_h,\ \varphi) = (w,\ \varphi), \quad \text{for all } \varphi \in V_h,$$

and

$$(F_h^{'}(v)p_h,\ \varphi) = (w,\ \varphi), \quad \text{for all } \varphi \in V_h.$$

Note that p_h solves the problem,

$$(F_h^{'}(u_h^\star)p_h,\ \varphi) = (w,\ \varphi) - ((b_u(u_h^\star) - b_u(v))p_h,\ \varphi), \quad \text{for all } \varphi \in V_h.$$

Therefore, the difference $p_h - q_h$ solves the problem

$$(F_h^{'}(u_h^\star)(p_h - q_h),\ \varphi) = -((b_u(u_h^\star) - b_u(v))p_h,\ \varphi), \quad \text{for all } \varphi \in V_h.$$

Therefore, based on Lemma 3.3 one has the uniform bound

$$\begin{aligned}\max\{\|p_h - q_h\|_1,\ \|p_h - q_h\|_\infty\} &\leq C\|(b_u(u_h^\star) - b_u(v))p_h\|_0\\ &\leq CL\|u_h^\star - v\|_\infty \|p_h\|_0\\ &\leq CL\|u_h^\star - v\|_\infty \|p_h - q_h\| + CL\|u_h^\star - v\|_\infty \|q_h\|_0\\ &\leq \delta\|p_h - q_h\| + CL\|u_h^\star - v\|_\infty \|q_h\|.\end{aligned}$$

This verifies (4.2):

$$\begin{aligned} \| \left((F_h'(u_h^\star))^{-1} - (F_h'(v))^{-1} \right) w\| &\leq \tfrac{C}{1-\delta} \|u_h^\star - v\|_\infty \|(F_h'(u_h^\star))^{-1} w\| \\ &\leq \gamma \|(F_h'(u_h^\star))^{-1} w\|, \end{aligned}$$

if $CL\|u_h^\star - v\|_\infty \leq CL\delta < 1$, which is the case if $\|u_h^\star - v\|_\infty < \delta$ is sufficiently small.

For (4.1) one proceeds as follows,

$$\begin{aligned} (F_h'(u_h^\star)w - F_h'(v)w,\ \psi) &= -((b_u(u_h^\star) - b_u(v))w,\ \psi) \\ &\leq L\|u_h^\star - v\|_\infty \|w\|_0 \|\psi\|_0 \\ &\leq L\|u_h^\star - v\|_\infty\ \mu\ \|F_h'(v)w\|_0\ \|\psi\|_0. \end{aligned}$$

In the last line we used (the verified) estimate (2.2). That is,

$$\|F_h'(u_h^\star)w - F_h'(v)w\|_0 \leq L\|u_h^\star - v\|_\infty \mu \|F_h'(u_h^\star)w\|_0 \leq \gamma \|F_h'(v)w\|_0,$$

if $\|u_h^\star - v\|_\infty$ is sufficiently small and this verifies (4.1). □

Note that the estimates in Lemma 4.3 are equivalent to the ones listed in (A3) (simply let $w := \left(F_h'(v)\right)^{-1} w$ in (4.1) and $w := F_h'(u_h^\star)w$ in (4.2)).

5. Application to Inexact Newton–Multigrid algorithms

In this section we consider MG solution algorithms for computing the inexact Newton direction. A modified inexact Newton–MG algorithm starting with v^0 close enough to $u_h^\star$ requires finding $s^k = \xi \in V_h$ such that

$$F_h'(v^0)\xi = -F_h(v^k) + r^k, \qquad \text{where } \|r^k\|_0 \leq \eta \|F_h(v^k)\|_0,$$

and then setting $v^{k+1} = v^k + \xi$.

For a properly chosen $\xi \in V_h$, we need an estimate of the form,

$$|(A\xi - b_u(v)\xi + A\widehat{v} - b(x,\ \widehat{v}) - f,\ \varphi)| \leq \eta\ \|A\widehat{v} - b(x,\ \widehat{v}) - f\|_0 \|\varphi\|_0, \quad \text{for all } \varphi \in V_h, \tag{5.1}$$

uniformly in $\widehat{v}$ with $v \in V_h$ close enough to $u_h^\star$. In our application $v = v^0$ and $\widehat{v} = v^k$.

Let $\widehat{\xi} \in V_h$ be the solution of the following discrete linear problem,

$$(A\widehat{\xi} - b_u(v)\widehat{\xi},\ \varphi) = -(A\widehat{v} - b(x,\ \widehat{v}) - f,\ \varphi), \quad \text{for all } \varphi \in V_h. \tag{5.2}$$

Next we consider three MG algorithms for computing the inexact Newton directions $s^k = \xi$.

V–cycle MG. We first use a MG procedure that produces computationally inexpensive ξ close to $\widehat{\xi}$ based on few ($m \geq 1$) V–cycle steps applied to the linear problem (5.2). The following MG V–cycle convergence estimate is well–known (note that we have assumed $b_u(v) \leq 0$ which implies positive definiteness of $A - b_u(v)$)

$$((A\widehat{\xi} - b_u(v))(\widehat{\xi} - \xi),\ \widehat{\xi} - \xi) \leq q^m\ ((A - b_u(v))\widehat{\xi},\ \widehat{\xi}).$$

Here, $q \in (0, 1)$ stands for the convergence factor of the V–cycle MG, and $\xi = \xi^m$ is the mth iterate ($\xi^0 = 0$). In the present setting q is independent of $h \mapsto 0$.

Since, $\widehat{\xi}$ solves the above linear problem, the following *a priori* estimate holds,

$$((A - b_u(v))\widehat{\xi},\ \widehat{\xi}) \le C\|A\widehat{v} - b(x,\ \widehat{v}) - f\|_0^2, \quad C = \frac{C_F}{a_0}.$$

Combining the last two estimates one ends up with

$$((A - b_u(v))(\widehat{\xi} - \xi),\ \widehat{\xi} - \xi) \le q^m C\|A\widehat{v} - b(x,\ \widehat{v}) - f\|_0^2.$$

The desired estimate (5.1) can be rewritten as

$$|(A(\xi - \widehat{\xi}) - b_u(v)(\xi - \widehat{\xi}),\ \varphi)| \le \eta\ \|A\widehat{v} - b(x,\ \widehat{v}) - f\|_0\|\varphi\|_0.$$

We have, with $B = A - b_u(v)$ and $w = \widehat{\xi} - \xi$, using the inverse estimate $\|\varphi\|_1 \le Ch^{-1}\|\varphi\|_0$,

$$\begin{aligned}(Bw,\ \varphi) &\le (Bw,\ w)^{\frac{1}{2}}(B\varphi,\ \varphi)^{\frac{1}{2}}\\ &\le C(\sqrt{q})^m\|A\widehat{v} - b(x,\ \widehat{v}) - f\|_0(B\varphi,\ \varphi)^{\frac{1}{2}}\\ &\le C(\sqrt{q})^m\|A\widehat{v} - b(x,\ \widehat{v}) - f\|_0\|\varphi\|_1\\ &\le C(\sqrt{q})^m\|A\widehat{v} - b(x,\ \widehat{v}) - f\|_0\ h^{-1}\|\varphi\|_0\end{aligned}$$

which is the desired estimate (5.1) with $\eta = Cq^{\frac{m}{2}}h^{-1}$.

A mesh independent bound for η can generally be achieved if the number of MG V–cycles is $m = \mathcal{O}(\log \frac{1}{h})$. This leads to a nearly optimal method.

Cascadic MG cycle. Alternatively, one can use a full MG, which provides approximation ξ to $\widehat{\xi}$ of order h (that is, of order of the discretization error) then one can get $\sqrt{q}^m h^{-1}$ uniformly bounded in terms of h and hence ensure a $\eta < 1$ in optimal complexity, cf., e.g., Shaidurov [14] and Bornemann and Deuflhard [6] for a cascadic version of the full MG. The following estimates are valid, for the cascadic multigrid, which involves smoothing using ν_k CG iterations at grid k, where $\nu_k \simeq 2^{\beta(\ell-k)}\nu$, $k = 0$ is coarsest level, $\ell > 0$ is the finest mesh level, (for properly chosen $\beta = \beta_k$ to ensure optimal complexity)

$$\|\xi - \widehat{\xi}\|_1 \le Ch\frac{1}{2\nu + 1}\|B\widehat{\xi}\|_0.$$

Then, as above, $((Bw,\ w) \le C\|w\|_1^2)$, with $w = \xi - \widehat{\xi}$

$$\begin{aligned}(Bw,\ \varphi) &\le (Bw,\ w)^{\frac{1}{2}}(B\varphi,\ \varphi)^{\frac{1}{2}}\\ &\le Ch\tfrac{1}{2\nu+1}\|B\widehat{\xi}\|_0\|\varphi\|_1\\ &\le Ch\tfrac{1}{2\nu+1}\|A\widehat{v} - b(x,\ \widehat{v}) - f\|_0\|\varphi\|_1\\ &\le Ch\tfrac{1}{2\nu+1}\|A\widehat{v} - b(x,\ \widehat{v}) - f\|_0\|\varphi\|_1\\ &\le Ch\tfrac{1}{2\nu+1}\|A\widehat{v} - b(x,\ \widehat{v}) - f\|_0\ h^{-1}\|\varphi\|_0\\ &\le C\tfrac{1}{2\nu+1}\|A\widehat{v} - b(x,\ \widehat{v}) - f\|_0\ \|\varphi\|_0.\end{aligned}$$

We formulate then the first result.

Theorem 5.1. *Let Assumption 3.1 (i)–(iv) hold. For η and t such that $0 < \eta < t < 1$ there is an $\epsilon > 0$ and a $h_0 > 0$ such that for any $h < h_0$ the following is true:*

- *The family of discrete operators* $\{F_h^{'}(v)\}$ *for* $v \in H_0^1(\Omega)$ *and* $\|v - u^\star\| < \epsilon$ *has uniformly bounded inverses (as operators from* $V_h,\ \|\cdot\| = \max\{\|\cdot\|_1,\ \|\cdot\|_{L_\infty}\}$ *to* $V_h,\ \|.\|_0$*); that is, for a constant* $\mu = \mu(h_0,\ \epsilon)$ *one has* $\|\left(F_h^{'}(v)\right)^{-1}\| \leq \mu$*. This implies that Algorithm 2.1 is feasible.*
- *For any initial iterate* $u_h^0 \in V_h$ *such that* $\mu\|F_h^{'}(u_h^\star)(u_h^0 - u_h^\star)\|_0 < \epsilon$ *and* $\|u_h^0 - u_h^\star\| < \epsilon$*, the sequence of iterates* u_h^k *generated by the modified inexact Newton method from Algorithm 2.1 (*$F = F_h$*), where the inexact Newton directions* $s^k = \xi$ *are computed via the cascadic MG, converges in* $\|\cdot\| = \max\{\|\cdot\|_1,\ \|\cdot\|_{L_\infty}\}$ *to* $u_h^\star$*.*
- *The convergence is linear in the sense that*

$$\|u_h^k - u_h^\star\| \leq \mu\|F_h^{'}(u_h^\star)\left(u_h^k - u_h^\star\right)\|_0 \leq t^k\ \mu\ \|F_h^{'}(u_h^\star)\left(u_h^0 - u_h^\star\right)\|_0 \leq t^k\ \epsilon.$$

In addition the cost per iteration is optimal; that is, proportional to the number of unknowns (degrees of freedom in V_h*). Finally, the assumption that one can choose the initial iterate* u_h^0*, as indicated above, is feasible and we provide a constructive (practical) algorithm to compute it.*

Proof. The uniform boundedness of the inverses $F_h^{'}(v)$ is Lemma 4.2. It is clear then (based on Lemmas 3.1, 4.1 and 4.3) that one can choose h_0 and ϵ sufficiently small such that for any $h < h_0$ if $\|v - u_h^\star\| < \epsilon$ the Assumption 2.1 can be guaranteed and Theorem 2.1 will show the desired convergence if we can also guarantee that the estimate

$$\mu\ \|F_h^{'}(u_h^\star)\left(u_h^0 - u_h^\star\right)\| \leq \epsilon,$$

is feasible uniformly in $h \mapsto 0$.

One can (theoretically) choose u_h^0 close to $u_h^\star$ such that

$$\mu\|F_h^{'}(u_h^\star)(u_h^0 - u_h^\star)\|_0 \leq \delta\ \|F_h^{'}(u_h^\star)u_h^\star\|_0$$

for any given small $\delta > 0$ (see estimate (5.5)). Theoretically, this means that if one knew the right hand side g of the problem $F_h^{'}(u_h^\star)u_h^\star = g$ and the actions of the linear operator $F_h^{'}(u_h^\star)$ were computationally available, one could have chosen u_h^0 as an approximation to $u_h^\star$ obtained by our MG cycle. The question then is if $\|F_h^{'}(u_h^\star)u_h^\star\|_0$ will stay bounded when $u_h^\star \mapsto u^\star$ (that is, when $h \mapsto 0$). We have, (noting that on V_h, $F_h^{'}(u_h^\star) = Q_h F^{'}(u_h^\star)$)

$$\begin{aligned}\|F_h^{'}(u_h^\star)u_h^\star\|_0 &\leq \|Q_h F^{'}(u_h^\star)(u_h^\star - u^\star)\|_0 + \|Q_h F^{'}(u_h^\star)u^\star\|_0\\ &\leq \|b_u(u_h^\star)\|_\infty\|u_h^\star - u^\star\|_0 + \|Q_h A(u^\star - u_h^\star)\|_0 + \|Au^\star\|_0 + \|b_u(u_h^\star)\|_\infty\|u^\star\|_0\\ &\leq C\|u^\star\|_2 + \|Q_h A(u^\star - u_h^\star)\|_0.\end{aligned}$$

The only thing that remains to be seen is that $\|Q_h A(u^\star - u_h^\star)\|_0$ stays bounded. It is easy to see that $\|Q_h A(u^\star - u_h^\star)\|_0 \leq C\|u^\star\|_2$. Indeed, for any $\psi \in L_2(\Omega)$, using the H^1–boundedness of A and an inverse inequality for $Q_h\psi \in V_h$, one has

$$\begin{aligned}(Q_h A(u^\star - u_h^\star),\ \psi) &= (A(u^\star - u_h^\star),\ Q_h\psi)\\ &\leq a_1\|u^\star - u_h^\star\|_1\ \|Q_h\psi\|_1\\ &\leq Ch\ \|u^\star\|_2\ Ch^{-1}\|\psi\|_0\\ &\leq C\|u^\star\|_2\|\psi\|_0.\end{aligned}$$

In practice, the actions of $F_h'(u_h^\star)$ and the r.h.s. g are not available, but one can instead use their coarse approximations. More specifically let define H such that $H < h_0$ (but independent of $h \mapsto 0$). Solve the discrete nonlinear problem $F_H(u_H^\star) = 0$ in the corresponding coarse finite element space V_H. Consider then the following (fine–grid) linear problem (note that $u_H^\star \in V_H \subset V_h$)

$$((A - b_u(u_H^\star))\overline{u}_h^\star,\ \varphi) = (\overline{g},\ \varphi) \equiv (f + b(u_H^\star) - b_u(u_H^\star)u_H^\star,\ \varphi), \quad \text{for all } \varphi \in V_h. \tag{5.3}$$

This problem approximates the problem $F_h(u_h^\star) = 0$ which rewritten reads,

$$(F_h'(u_H^\star)u_h^\star,\ \varphi) = (f + b(u_h^\star) - b_u(u_H^\star)u_h^\star,\ \varphi), \text{ for all } \varphi \in V_h.$$

Since the difference $\overline{u}_h^\star - u_h^\star$ solves

$$(F_h'(u_H^\star)(\overline{u}_h^\star - u_h^\star),\ \varphi) = (b(u_H^\star) - b(u_h^\star) - b_u(u_H^\star)(u_H^\star - u_h^\star),\ \varphi), \text{ for all } \varphi \in V_h,$$

one gets the estimate

$$(F_h'(u_H^\star)(\overline{u}_h^\star - u_h^\star),\ \varphi) \leq L\|u_H^\star - u_h^\star\|_\infty \|u_H^\star - u_h^\star\|_0 \|\varphi\|_0.$$

That is,

$$\|(F_h'(u_H^\star)(\overline{u}_h^\star - u_h^\star)\|_0 \leq L\|u_H^\star - u_h^\star\|_\infty \|u_H^\star - u_h^\star\|_0.$$

Now, since $h < H < h_0$ the error estimates from Lemma 3.1 hold. Therefore,

$$\|(F_h'(u_H^\star)(\overline{u}_h^\star - u_h^\star)\|_0 \leq CH^{1+\alpha}\|u^\star\|_2^2. \tag{5.4}$$

Let finally u_h^0 be an approximation to $\overline{u}_h^\star$ obtained by our MG algorithm applied to (5.3). Note that u_h^0 is computationally available and computed in optimal cost. We have the convergence estimate of the cascadic MG cycle,

$$\|F_h'(u_H^\star)(\overline{u}_h^\star - u_h^0)\|_0 \leq C\frac{1}{2\nu + 1}\ \|\overline{g}\|_0 = C\frac{1}{2\nu + 1}\ \|f + b(u_H^\star) - b_u(u_H^\star)u_H^\star\|_0.$$

The final estimate we need reads (note that $\|u_H^\star - u_h^\star\| \leq CH^{1+\alpha}\|u^\star\|_2 < \epsilon$ hence (2.10) holds with a $\gamma < 1$):

$$\begin{aligned}\|F_h'(u_h^\star)(u_h^\star - u_h^0)\|_0 &\leq (1+\gamma)\|F_h'(u_H^\star)(u_h^\star - u_h^0)\|_0 \\ &\leq 2\left[\|F_h'(u_H^\star)(u_h^\star - \overline{u}_h^\star)\|_0 + \|F_h'(u_H^\star)(\overline{u}_h^\star - u_h^0)\|_0\right] \\ &\leq 2\left[CH^{1+\alpha}\|u^\star\|_2^2 + C\tfrac{1}{2\nu+1}\ \|f + b(u_H^\star) - b_u(u_H^\star)u_H^\star\|_0\right] \\ &\leq \tfrac{\epsilon}{\mu},\end{aligned}$$

which can be ensured if both H and $\frac{1}{2\nu+1}$ are sufficiently small. We again note that H and hence $u_H^\star$ depend only on η and t, hence the terms that involve $u_H^\star$ can be considered fixed (for fixed η and t). □

In [11] the cascadic MG has been extended for semi–linear elliptic PDEs of the form we considered. It provides approximation of order h starting from a coarse grid $H = \mathcal{O}(\sqrt{h})$ and at every finer level one solves the corresponding linearized problem with geometrically decreasing number of CG iterations using the interpolated final iterate from the previous coarse mesh as an initial iterate for the CG method. Early results on multilevel inexact Newton methods in the form of nested iteration or cascadic iteration are found in [5].

Remark 5.1. *Note that our coarse mesh size H used in the proof of Theorem 5.1 is fixed (independent of $h \mapsto 0$). If one makes it mesh dependent (e.g., as in* [15], [3] *and* [11]*) then strictly speaking one does not really have control on the cost.*

W–cycle MG. Finally, if one uses W–cycle MG with sufficiently many smoothing iterations $\nu \geq 1$ the following L_2 reduction of the MG residual can be proved based on well–known arguments (see Appendix II):

$$\|A(\xi - \widehat{\xi}) - b_u(v)(\xi - \widehat{\xi})\|_0 \leq C \frac{1}{1+\nu} \|A\widehat{v} - b(x,\ \widehat{v}) - f\|_0. \tag{5.5}$$

This is exactly our desired estimate (5.1) with a mesh-independent $\eta = C\frac{1}{1+\nu} < 1$.

We summarize:

Theorem 5.2. *Let Assumption 3.1 (i)–(iv) hold. For η and t such that $0 < \eta < t < 1$ there is an $\epsilon > 0$ and a $h_0 > 0$ such that for any $h < h_0$ the following is true:*

- *The family of discrete operators $\{F_h'(v)\}$ for $v \in H_0^1(\Omega)$ and $\|v - u^\star\| < \epsilon$ has uniformly bounded inverses (as operators from $V_h,\ \|\cdot\| = \max\{\|\cdot\|_1,\ \|\cdot\|_{L_\infty}\}$ to $V_h,\ \|.\|_0$); that is, for a constant $\mu = \mu(h_0,\ \epsilon)$ one has $\|\left(F_h'(v)\right)^{-1}\| \leq \mu$. This implies that Algorithm 2.1 is feasible.*
- *For any initial iterate $u_h^0 \in V_h$ such that $\mu\|F_h'(u_h^\star)(u_h^0 - u_h^\star)\|_0 < \epsilon$ and $\|u_h^0 - u_h^\star\| < \epsilon$, the sequence of iterates u_h^k generated by the modified inexact Newton method from Algorithm 2.1 ($F = F_h$), where the inexact Newton directions $s^k = \xi$ are computed via the W–cycle MG with sufficiently many smoothing steps (depending only on ϵ), converges in $\|\cdot\| = \max\{\|\cdot\|_1,\ \|\cdot\|_{L_\infty}\}$ to $u_h^\star$.*
- *The convergence is linear in the sense that*

$$\|u_h^k - u_h^\star\| \leq \mu\|F_h'(u_h^\star)\left(u_h^k - u_h^\star\right)\|_0 \leq t^k\ \mu\ \|F_h'(u_h^\star)\left(u_h^0 - u_h^\star\right)\|_0 \leq t^k\ \epsilon.$$

In addition the cost per iteration is optimal; that is, proportional to the number of unknowns (degrees of freedom in V_h). Finally, the assumption that one can choose the initial iterate u_h^0 as indicated above is feasible and we provide a constructive (practical) algorithm to compute it.

Proof. We omit the proof of this second result as it is essentially the same as for Theorem 5.1. ☐

6. Numerical illustration

In this section, we illustrate the results of the previous section on two test problems. The first problem is covered by our theoretical results, while the second is not. However, the numerical results suggest that the theory is true more generally, and we make some concluding remarks regarding extensions in the following section. We have tested the following nonlinear second order elliptic PDEs:

Problem 1:

$$-\Delta u - b(u) = f(x,y),\ (x,y) \in \Omega = (0,\ 2) \times (0,\ 1),$$

subject to homogeneous Dirichlet boundary conditions ($u = 0$ on $\partial\Omega$). The nonlinear function was

$$-b(u) = \begin{cases} e^{-u}u^3, & u \geq 0, \\ 0, & u < 0. \end{cases}$$

Problem 2:

$$-\nabla(a(u)\nabla u) - b(u) = f(x,y),\ (x,y) \in \Omega = (0,\ 2) \times (0,\ 1),$$

subject to homogeneous Dirichlet boundary conditions ($u = 0$ on $\partial\Omega$). The nonlinear functions were

$$a(u) = \frac{1}{\sqrt{\epsilon + u^2}},\ \epsilon = 0.001,\ \ b(u) = \begin{cases} e^{-u}u^3, & u \geq 0, \\ 0, & u < 0. \end{cases}$$

In all cases, the right–hand side f was chosen to match the exact solution

$$u^\star = 10\psi(x)\, y(1-y),\ \text{where}\ \psi(x) = \begin{cases} x^2(1-x)^2, & x \in [0,\ 1], \\ (x-1)^2(2-x)^2, & x \in [1,\ 2]. \end{cases}$$

The discrete problems are obtained by using bilinear basis functions on rectangular elements of mesh size $h = \frac{1}{32}, \frac{1}{64}, \frac{1}{128}, \frac{1}{256}$. The coarsening is based on element agglomeration which leads to an algebraic multigrid (AMGe) which for the model uniform rectangular mesh and smooth PDE coefficients performs pretty much like a standard geometric multigrid. We have chosen the r.h.s. of the discrete problems such that the exact discrete solution $u_h^\star$ matches $u^\star$ pointwise at the mesh points. Thus we know exactly the discrete solution and can measure the algebraic error $u_h^k - u_h^\star$ in various norms.

In the nonlinear iterations we have selected initial iterates $u_h^0 = 10(2\theta - 1)u_h^\star$, for $\theta = \theta(x)$ being random numbers in [0,1]. We tested the performance of the inexact Newton–MG with V–cycle MG for solving the linearized systems. Since the V–cycle showed optimal mesh–independent convergence, it is clear that the W-cycle and the cascadic MG will lead to the same result. In Table 1, and Table 2, we show the number of nonlinear iterations required to achieve ℓ^2–residual error reduction by a factor 10^{-12} for Problem 1 and 10^{-6} for Problem 2, total number of linear iterations (i.e., number of V–cycles) and the algebraic errors $u_h^k - u_h^\star$, (here k stands for the final nonlinear iterate) in three norms, maximum, ℓ^2 and ℓ^2–residual norm which are discrete counterparts of the continuous L_∞ and L_2 norms. It is clear from the tables that convergence is mesh independent and of optimal cost for the inexact Newton MG method. In Problem 2, the Jacobian $F_h^{'}(u_h^k)v$ for $v \in V$ has been approximated by the linear operator corresponding to the bilinear form $(a(u_h^k)\nabla v,\ \varphi) - (b_u(u_h^k)v,\ \varphi)$ for all $\varphi \in V$. Note that this gives rise to a symmetric positive definite matrix which does not pose additional difficulties to a standard multigrid. This approximation to the true (discrete) Jacobian gives rise to one more level of inexactness in the inexact Newton method we consider. In Fig. 1 one can see the almost linear convergence of the inexact Newton MG method in the ℓ^2-residual-norm and a bit faster than

TABLE 1. Convergence of inexact Newton–MG for **Problem 1**; number of nonlinear iterations, total number of linear iterations and the algebraic errors $u_h^k - u_h^\star$.

h^{-1}	nonlinear iterations	total number of linear iterations	max–error	ℓ^2–residual error	ℓ^2–error
32	5	14	1.137573e-11	9.873159e-11	7.467105e-12
64	5	15	9.253515e-12	7.853968e-11	6.083558e-12
128	4	12	1.773736e-09	1.548300e-08	1.192745e-09
256	4	12	1.163052e-09	1.119520e-08	7.184886e-10

TABLE 2. Convergence of inexact Newton–MG for **Problem 2**; number of nonlinear iterations, total number of linear iterations and the algebraic errors $u_h^k - u_h^\star$.

h^{-1}	nonlinear iterations	total number of linear iterations	max–error	ℓ^2–residual error	ℓ^2–error
32	9	27	1.722173e-05	3.345141e-03	5.257332e-06
64	8	24	1.067139e-04	1.980442e-02	3.410634e-05
128	7	21	5.370777e-04	9.792177e-02	1.826281e-04
256	7	21	6.083837e-04	1.113324e-01	2.077409e-04

linear convergence in maximum and ℓ^2-norm. Finally, we mention that we have implemented the more traditional version of the inexact Newton MG method (see Corollary 2.1).

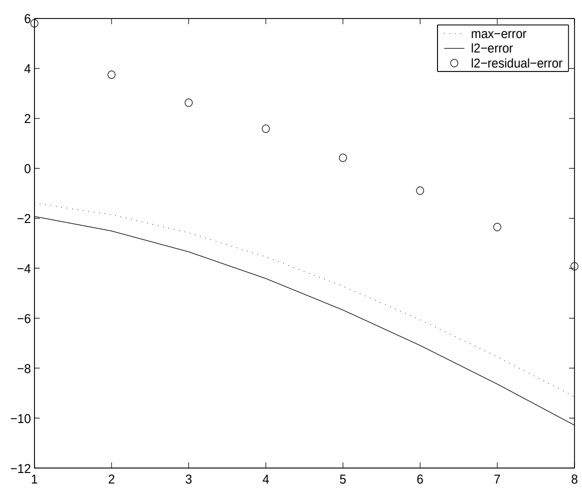

FIGURE 1. Plot of the logarithm of the errors versus the number of iterations; $h^{-1} = 64$; **Problem 2.**

7. Concluding remarks

The full regularity for the simple model case is not essential (if we give up convergence in L_∞). It may be more adequate if one considers the following more general nonlinear elliptic operator,

$$\mathcal{L}u = -\operatorname{div}(A(x)\nabla u + \underline{a}(x,\ u)) - b(x,\ u).$$

Here, in addition to the previous coefficients $A(x) = (a_{ij}(x))_{i,j=1}^d$ and $b(x,\ u)$ we have the nonlinear vector function $\underline{a}(x,\ u) = (a_i(x,\ u))_{i=1}^d$. Hence, $\operatorname{div} A(x)\nabla u = \sum\limits_i \frac{\partial}{\partial x_i}(\sum\limits_j a_{ij}(x)\frac{\partial u}{\partial x_j})$ and $\operatorname{div}\underline{a}(x,\ u) = \sum\limits_{i=1}^d \left[\frac{\partial a_i(x,\,u)}{\partial x_i} + \frac{\partial a_i}{\partial u}(x,\ u)\frac{\partial u}{\partial x_i}\right]$. We remark, that the analysis should translate in a straightforward manner to the above more general nonlinear operator under the assumptions that the nonlinear PDE has a unique solution $u \in H^2(\Omega) \cap H_0^1(\Omega)$, appropriate smoothness of the coefficients $a_i(x,\ u)$, proper sign of the derivatives of $\underline{a}$ and b which guarantee coercivity of the linearized operator, and using W–cycle MG with sufficiently many smoothing steps for computing the inexact Newton directions. Also, as is well–known, the linear finite element problems associated with the discrete Jacobians may require a sufficiently fine coarse mesh in order to have a stable discretization.

Finally, it should be possible to extend the results presented here to problems involving nonlinearities associated with the leading second order term, for example $-\operatorname{div} A\nabla u$ with $A = A(.,\ u)$. However, these problems require more sophisticated L_p–error estimates, (cf., [7] or [15], and earlier in [9]) and also require, in general, higher regularity or smoothness of the solution (see [15] and also [8]).

Appendices

Appendix I: Existence and error estimates of the discrete solution. Such results are found in many papers, see, e.g., Xu [15] (see also Chapter 7.7 of [7]). The main construction is as follows. Let $P_h:\ H^2(\Omega) \cap H_0^1(\Omega) \mapsto V_h$ be the projection with respect to the linear operator $A - b_u(u)$. That is, $v \mapsto P_h v$ is computed by solving the linear finite element problem,

$$((A - b_u(u))P_h v,\ \varphi) = ((A - b_u(u))v,\ \varphi), \quad \text{for all } \varphi \in V_h.$$

We use the fact that P_h has certain approximation properties, namely, we have $\|u - P_h u\|_\infty \le Ch^\alpha \|u\|_2$ for some $\alpha > 0$ and $\|u - P_h u\|_1 \le Ch\|u\|_2$.

Recall that in our application we use the norm $\|\cdot\| = \max\{\|\cdot\|_1,\ \|\cdot\|_\infty\}$.

Consider the following ball,

$$B = \{\psi \in V_h:\ \|\psi - P_h u\|_\infty \le h^\alpha,\ \|\psi - P_h u\|_1 \le h\}.$$

Let $\Phi:\ B \mapsto V_h$ be the nonlinear mapping defined by $\psi \in V_h \mapsto \Phi(\psi)$ as the solution of the problem,

$$((A - b_u(u))\Phi(\psi),\ \varphi) = ((A - b_u(u))u,\ \varphi) - R(u;\ \psi,\ \varphi), \quad \text{for all } \varphi \in V_h.$$

Here, $R(u;\ \psi,\ \xi)$ is the residual form coming from the Taylor series $F(\psi) = F(u) + F'(u)(\psi - u) + \dots$, in a variational form, that is,

$$R(u;\ \psi,\ \xi) = (A\psi - b(\psi),\ \xi) - (Au - b(u),\ \xi) - ((A - b_u(u))(\psi - u),\ \xi).$$

It is clear that $\Phi(\psi)$ is continuous. One has,

$$\begin{aligned} R(u;\ \psi,\ \xi) - R(u;\ \theta,\ \xi) &= -(b(\psi) - b(\theta) - b_u(u)(\psi - \theta),\ \xi) \\ &\leq C \max\{\|\psi - P_h u\|_\infty,\ \|\theta - P_h u\|_\infty,\ \|u - P_h u\|_\infty\} \|\psi - \theta\| \cdot \|\xi\|_0. \end{aligned}$$

Hence, since $\Phi(\psi) - \Phi(\theta)$ solves

$$((A - b_u(u))(\Phi(\psi) - \Phi(\theta)),\ \varphi) = -R(u;\ \psi,\ \varphi) + R(u;\ \theta,\ \varphi),$$

based on Lemma 3.3, the Lipschitz continuity of Φ (on B) follows,

$$\|\Phi(\psi) - \Phi(\theta)\| \leq Ch^\alpha \|\psi - \theta\|.$$

We used the following estimates for $R(\cdot;\ \cdot,\ \cdot)$,

$$R(u;\ \psi,\ \xi) = -(b(\psi) - b(u) - b_u(u)(\psi - u),\ \xi) \leq L\|\psi - u\|_\infty \|\psi - u\|_0 \|\xi\|_0.$$

Then for $\psi \in B$, using the approximation property of P_h on u, and the triangle inequality $\|\psi - u\| \leq \|\psi - P_h u\| + \|u - P_h u\| \leq Ch^\alpha$, imply

$$|R(u;\ \psi,\ \xi)| \leq Ch^{2\alpha} \|\xi\|_0.$$

Note that $\Phi(\psi) - P_h u$ solves the problem,

$$((A - b_u(u))(\Phi(\psi) - P_h u),\ \varphi) = -R(u;\ \psi,\ \varphi), \text{ for all } \varphi \in V_h.$$

The latter implies, based on Lemma 3.3,

$$\|\Phi(\psi) - P_h u\| \leq C \sup_{\xi \in V_h} \frac{|R(u;\ \psi,\ \xi)|}{\|\xi\|_0} \leq Ch^{1+\alpha} < h.$$

That is, $\Phi(B) \subset B$ and Brouwer's fixed-point theorem implies existence of solution $u_h^\star = \Phi(u_h^\star)$ of the discrete nonlinear problem,

$$(Au_h^\star - b(u_h^\star),\ \varphi) - (Au - b(u),\ \varphi) = 0, \quad \text{for all } \varphi \in V_h.$$

This is true since, the above identity is equivalent to $((A - b_u(u))(u - u_h^\star),\ \varphi) = R(u;\ u_h^\star,\ \varphi)$ and the latter on the other hand is the definition of $\Phi(u_h^\star) = u_h^\star$.

The error estimate is readily obtained using the approximation property of P_h and the fact that $u_h^\star \in B$. One has,

$$\|u - u_h^\star\|_1 \leq \|u - P_h u\|_1 + \|P_h u - u_h^\star\|_1 \leq Ch\|u\|_2 + h \leq Ch.$$

Similarly, one gets $\|u - u_h^\star\|_\infty \leq Ch^\alpha$.

Appendix II: Multigrid residual convergence. This result is proved based on well–known facts; namely, an approximation property and a smoothing property, (cf. Hackbusch [12] or Brenner and Scott [7]). Typically, though the convergence of MG is studied in energy norm (or $\| \cdot \|_1$). Here, for completeness, we show a two–grid convergence result for the residual iteration matrix in L_2 (which is $(A., A.)$ or energy–square norm convergence of the iterates). Then, as is well–known, by perturbation analysis, a W–cycle MG convergence follows assuming sufficiently many smoothing steps.

Let $A_h = A - b_u(v)$ be the discrete Jacobian $F_h'(v)$. Let also $V_{2h} \subset V_h$ be a coarse finite element space. Here we show that under the assumed H^2–regularity, a standard two–grid algorithm applied to

$$(A_h \xi_h,\ \varphi) = (r,\ \varphi), \quad \text{for all } \varphi \in V_h,$$

gives iterates $\xi = \xi_{TG}$ such that

$$\|A_h(\xi_h - \xi)\|_0 \leq C \frac{1}{1+\nu} \|r\|_0,$$

where ν is the number of smoothing iterations. That is, the L_2–norm of the initial residual r (initial iterate is zero) is reduced with a factor that can be made arbitrarily small if ν is sufficiently large. Then, as is well–known, a W–cycle MG will have asymptotically the same property (for sufficiently many smoothing iterations).

Consider a standard smoothing iteration matrix $I - \omega A_h$, where $\omega \simeq \mathcal{O}(\|A_h\|^{-1}) \simeq \mathcal{O}(h^2)$ and $\omega \leq \|A_h\|^{-1}$ or slightly more general, a symmetric positive definite matrix M, such that $(A_h \phi,\ \phi) \leq (M\phi,\ \phi)$ for all $\phi \in V_h$, $\|M\|_0 \simeq h^{-2}$ and $\text{cond}(M) = \mathcal{O}(1)$. A two–grid algorithm has the following iteration matrix that relates the resulting residual $r_{TG} = A_h(\xi_h - \xi_{TG})$ and the initial residual r,

$$r_{TG} = (I - A_h M^{-1})^{\nu} (I - A_h P (A_{2h})^{-1} P^T) r.$$

Here, $P^T = Q_{2h}$ is the restriction from V_h to V_{2h} (Q_{2h} is the L_2–projection onto V_{2h}). Then $P = I$ on V_{2h} (since $V_{2h} \subset V_h$). Hence, $A_{2h} = P^T A_h P = Q_{2h} A_h$.

Using duality and based on the full regularity the following approximation property holds (cf., e.g., [12]):

$$\| \left((A_h)^{-1} - P(A_{2h})^{-1} P^T \right) r \|_0 \leq C h^2 \|r\|_0 \tag{7.1}$$

We show an estimate of the form,

$$\|r_{TG}\|_0 \leq C \frac{1}{1+\nu} \|r\|_0.$$

Based on the formula for the TG residual iteration matrix and using the approximation property (7.1) and the properties of M, the following inequalities are straightforward,

$$\begin{aligned}
\|r_{TG}\|_0 &\leq \|(I - A_h M^{-1})^\nu A_h M^{-1}\|_0 \cdot \|M\|_0 \cdot \| \left((A_h)^{-1} - P(A_{2h})^{-1} P^T\right) r\|_0 \\
&\leq \|M^{-\frac{1}{2}}(I - A_h M^{-1})^\nu A_h M^{-\frac{1}{2}}\|_0 \cdot \mathrm{cond}(M^{\frac{1}{2}}) \cdot \|M\|_0 \cdot \|((A_h)^{-1} - P(A_{2h})^{-1}P^T) r\|_0 \\
&= \|(I - M^{-\frac{1}{2}} A_h M^{-\frac{1}{2}})^\nu M^{-\frac{1}{2}} A_h M^{-\frac{1}{2}}\|_0 \cdot \mathrm{cond}(M^{\frac{1}{2}}) \cdot \\
&\qquad\qquad \|M\|_0 \cdot \| \left((A_h)^{-1} - P(A_{2h})^{-1} P^T\right) r\|_0 \\
&\leq \max_{t\in[0,\,1]} \{(1-t)^\nu t\} \cdot \mathrm{cond}(M^{\frac{1}{2}}) \cdot C h^{-2}\, C h^2 \|r\|_0 \\
&\leq C \frac{1}{1+\nu} \|r\|_0.
\end{aligned}$$

Here, we used also the fact that the eigenvalues of the symmetric operator $M^{-\frac{1}{2}} A_h M^{-\frac{1}{2}}$ are contained in $[0, 1]$.

References

[1] E. L. Allgower, S. F. McCormick, and D. V. Pryor, *A general mesh independence principle for Newton's method applied to second order boundary value problems*, Computing 23(1979), pp. 233–246.

[2] E. L. Allgower and S. F. McCormick, *Newton's method with mesh refinements for numerical solution of nonlinear two-point boundary value problems*, Numerische Mathematik 29(1978), pp. 237-260.

[3] O. Axelsson and W. Layton, *A two–level method for the discretization of nonlinear boundary value problems*, SIAM J. Numer. Anal. 33(1996), pp. 2359-2374.

[4] R. E. Bank and D. J. Rose, *Global approximate Newton methods*, Numerische Mathematik 37(1981), pp. 279-295.

[5] R. E. Bank and D. J. Rose, *Analysis of a multi-level iterative method for nonlinear finite element equations*, Mathematics of Computation 39(1982), pp. 435-465.

[6] F. A. Bornemann and P. Deuflhard, *The cascadic multigrid method for elliptic problems*, Numerische Mathematik 75(1996), pp. 135-152.

[7] S. C. Brenner and L. R. Scott, The Mathematical Theory of Finite Element Methods, Springer-Verlag, New York, 1994.

[8] A. L. Codd, T. A. Manteuffel, and S. F. McCormick, *Multilevel first-order system least squares for nonlinear partial differential equations, with applications to elliptic grid generation*, preprint (2002) (available at http://ftp.amath.colorado.edu/pubs/fosls/nonlinear.ps).

[9] J. Douglas, Jr. and T. Dupont, *A Galerkin method for a nonlinear Dirichlet problem*, Mathematics of Computation 29(1975), pp. 689-696.

[10] R. S. Dembo, S. C. Eisenstat, and T. Steihaug, *Inexact Newton methods*, SIAM J. Numer. Anal. 19(1982), pp. 400-408.

[11] L. V. Gilyova and V. V. Shaidurov, *A cascade algorithm for solving a discrete analogue of weakly nonlinear elliptic problems*, Technische Universität Dresden, 1998, preprint IOKOMO-01-98.

[12] W. Hackbusch, Multi-Grid Methods and Applications, Springer-Verlag, Berlin, 1985.

[13] L. Mansfield, *On the solution of non-linear finite element systems*, SIAM Journal Numerical Analysis 17(1980), pp. 752-765.

[14] V. V. Shaidurov, *Some estimates of the rate of convergence of the cascadic conjugate gradient method*, Computers Math. Applic. 31(1996), pp. 161-171.

[15] J. Xu, *Two–grid discretization techniques for linear and nonlinear problems*, SIAM Journal Numerical Analysis 33(1996), pp. 1759-1777.

[16] T. J. Ypma, *Local convergence of inexact Newton methods*, SIAM Journal Numerical Analysis 21(1984), pp. 583-590.

Center for Applied Scientific Computing, Lawrence Livermore National Laboratory, P.O. Box 808, L-561, Livermore, CA 94551, U.S.A.
E-mail address: pnbrown@llnl.gov, panayot@llnl.gov, cswoodward@llnl.gov

University of California
Lawrence Livermore National Laboratory
Technical Information Department
Livermore, CA 94551