SIAM J. NUMER. ANAL.
Vol. 33, No. 5, pp. 1759–1777, October 1996

TWO-GRID DISCRETIZATION TECHNIQUES FOR LINEAR AND NONLINEAR PDEs*

JINCHAO XU†

Abstract. A number of finite element discretization techniques based on two (or more) subspaces for nonlinear elliptic partial differential equations (PDEs) is presented. Convergence estimates are derived to justify the efficiency of these algorithms. With the new proposed techniques, solving a large class of nonlinear elliptic boundary value problems will not be much more difficult than the solution of one linearized equation. Similar techniques are also used to solve nonsymmetric and/or indefinite linear systems by solving symmetric positive definite (SPD) systems. For the analysis of these two-grid or multigrid methods, optimal $\mathcal{L}^p$ error estimates are also obtained for the classic finite element discretizations.

Key words. error estimates, finite elements, iterative methods, multigrid, nonlinear, nonsymmetric, two-grid

AMS subject classifications. 65F10, 65N22, 65N30, 65M50, 65N55

1. Introduction. The main purpose of this paper is to present some discretization techniques based on two (or more) finite element subspaces for solving partial differential equations (PDEs). Examples under our study here for this technique are linear as well as nonlinear second-order elliptic boundary value problems. Inspired by Xu [1–2] for a method to solve nonsymmetric and indefinite linear algebraic systems, we employ two finite element subspaces, $\mathcal{V}_H$ and $\mathcal{V}_h$ (with mesh size $h \ll H$), in our discretization schemes. On the coarser space $\mathcal{V}_H$, we use the standard finite element discretization to obtain a rough approximation $u_H \in \mathcal{V}_H$ and then solve a linearized equation based on u_H to produce a corrected solution $u^h \in \mathcal{V}_h$. A remarkable fact about this simple technique is that the space $\mathcal{V}_H$ can be extremely coarse (in contrast to $\mathcal{V}_h$) and still maintain the optimal accuracy. For example, if the piecewise linear finite element is used for a semilinear equation, u^h is asymptotically as accurate as the standard (nonlinear) finite element discretization in the finer space $\mathcal{V}_h$ if $H = O(h^{\frac{1}{4}})$. Moreover, if two linearized systems are solved on $\mathcal{V}_h$, it suffices to take $H = O(h^{\frac{1}{8}})$. This means that solving a nonlinear equation is not much more difficult than solving one linear equation, since $\dim \mathcal{V}_H \ll \dim \mathcal{V}_h$ and the work for solving u_H is relatively negligible.

The two-grid algorithm is also extended to multiple subspaces $\mathcal{V}_{h_j}$ and optimal estimates are obtained with, for example, $h_1 = O(H^4|\log H|)$ and $h_j = h_{j-1}^2|\log h_{j-1}|$ $(2 \le j \le J)$. This type of algorithm is related to the so-called *projective Newton method* studied by Vainikko [3] and Witsch [4]. A convergence analysis of this algorithm was given in Rannacher [5] and recently in [6]. Similar methods have also been studied by Bank [7] for the multigrid iterative solution of the nonlinear algebraic systems resulting from the standard finite element discretization. For other multigrid methods for nonlinear problems, we refer, for example, to Brandt [8], Hackbusch [9], and Reusken [10] and the references cited therein.

Our method is also, in a way, related to the so-called *mesh independence principle* (MIP) that has been studied, for example, in [6], [11–15] (and other references cited therein) for solving nonlinear differential equations by the Newton iterations. MIP refers to the fact that the number of Newton iterations in solving the discretized (by finite difference or finite element) nonlinear differential equation is asymptotically independent of the discretization parameters such as the mesh size. With the method in this paper, a stronger MIP holds: only one Newton iteration is sufficient. (In this statement, of course, we did not consider the number of Newton

*Received by the editors June 22, 1992; accepted for publication (in revised form) November 7, 1994.
†Department of Mathematics, Penn State University, University Park, PA 16802 (xu@math.psu.edu). This work was partially supported by the National Science Foundation and the Schlumberger Foundation.

iterations needed to solve the nonlinear system from the coarse grid, which requires very little work as compared with the one Newton iteration in the fine grid.)

The error analysis of our two-grid methods is based on some $\mathcal{L}^p$ and $\mathcal{W}_p^1$ estimates ($2 \le p \le \infty$) for the standard finite element discretization. The cases $p = 2$ and $p = \infty$ have been studied by many authors (cf. Schultz [16], Douglas and Dupont [17], Nitsche [18], Johnson and Thomée [19], Rannacher [20], Frehse [21], Mittelmann [22], Frehse and Rannacher [23–24], Rannacher [5], Nitsche [25], Dobrowolski and Rannacher [26]), but the case $p \neq 2$ or ∞ cannot be found in the literature. Because of their independent theoretical interests (in addition to the application in this paper), we shall a give a detailed derivation of these estimates (in §3). One main idea in our analysis is to linearize the nonlinear PDEs at the exact solution and consider its finite element discretizations. Such an idea has been used in most of the aforementioned papers, but our analysis appears to be much simpler. One observation that plays a significant role in our analysis is that the finite element approximation of the nonlinear equation is "superclose" to the finite element approximation to the aforementioned linearized equation. As a result, the analysis of nonlinear problems is, in a straightforward fashion, reduced to the analysis of linear problems. For this reason, we shall give a brief presentation on the finite element theory for linear problems (in §2) and, in particular, on some estimates for some discrete Green functions.

It is also observed that the finite element solution of a nonsymmetric and/or indefinite linear equation is "superclose" to the finite element solution of some symmetric positive definite (SPD) equation. These facts are useful for both the theoretical analysis and the design of efficient solvers for the resultant algebraic systems. In fact, we shall present several algorithms based on such considerations for nonsymmetric and/or indefinite linear systems (in §4).

For simplicity of exposition, only the scalar equations will be considered in this paper, but the techniques and the corresponding results are extended to certain systems of equations in a very straightforward fashion. For clarity of presentation, we will only consider two-dimensional problems, as many results for three-dimensional linear problems needed in our nonlinear analysis are not readily available in the literature. It is well known that, in the finite element theory, special care needs to be taken near the curved part of the boundary in order to achieve the best approximation for higher-order elements. For the sake of simplicity, we will not get into the technical details in this direction and our presentation will be made solely for polygonal domains. For technical reasons, for some of the results in the paper we will further assume the polygonal domain is also convex.

Therefore, we assume that Ω is a convex polygonal domain in the plane. For $p \ge 1$ and integer $m \ge 0$, let $\mathcal{W}_p^m(\Omega)$ be the standard Sobolev space with a norm $\|\cdot\|_{m,p}$ given by

$$\|v\|_{m,p}^p = \sum_{|\alpha| \le m} \|D^\alpha v\|_{\mathcal{L}^p(\Omega)}^p.$$

For $p = 2$, we denote $\mathcal{H}^m(\Omega) = \mathcal{W}_2^m(\Omega)$ and $\mathcal{H}_0^1(\Omega) = \{v \in \mathcal{H}^1(\Omega) : v|_{\partial\Omega} = 0\}$, where $v|_{\partial\Omega} = 0$ is in the sense of *trace*, $\|\cdot\|_m = \|\cdot\|_{m,2}$ and $\|\cdot\| = \|\cdot\|_{0,2}$. The space $\mathcal{H}^{-1}(\Omega)$, the dual of $\mathcal{H}_0^1(\Omega)$, will also be used.

Throughout this paper, we shall use the letter C or c (with or without subscripts) to denote a generic positive constant which may stand for different values at its different occurrences. When it is not important to keep track of these constants, we shall conceal the letter C or c into the notation $\lesssim$ or $\gtrsim$. Here

$$x \lesssim y \quad \text{means} \quad x \le Cy \quad \text{and} \quad x \gtrsim y \quad \text{means} \quad x \ge cy.$$

The rest of the paper is organized as follows. In §2, linear PDEs and finite element discretizations will be discussed. Section 3 is devoted to finite element discretization for nonlinear elliptic boundary value problems and the error estimate in $\mathcal{L}^p$-norm. Some two-grid methods for nonsymmetric and/or indefinite linear PDEs will be discussed in §4. The main two-grid and some multigrid algorithms will be presented in §5.

2. Linear problems and finite element discretizations. In this section, we shall discuss finite element discretizations for nonsymmetric and/or indefinite linear PDEs. These results, mostly well known, lay down the groundwork for the further analysis of nonlinear problems.

2.1. Linear elliptic partial differential operators. Let α, β, γ (with the ranges in $\mathbb{R}^{2\times 2}$, $\mathbb{R}^2$, and $\mathbb{R}^1$, respectively) be smooth functions on $\overline{\Omega}$ satisfying, for some positive constant α_0, that

$$\xi^T\alpha(x)\xi \geq \alpha_0|\xi|^2 \quad \forall\, \xi \in \mathbb{R}^2.$$

We shall study the two linear operators[1]

$$\mathcal{L}v = -\mathrm{div}(\alpha(x)\nabla v) \quad \text{and} \quad \hat{\mathcal{L}}v = \mathcal{L}v + \beta(x)\cdot\nabla v + \gamma(x)v. \tag{2.1}$$

Obviously $\mathcal{L} : \mathcal{H}_0^1(\Omega) \to \mathcal{H}^{-1}(\Omega)$ is an isomorphism. Our basic assumption is that $\hat{\mathcal{L}} : \mathcal{H}_0^1(\Omega) \to \mathcal{H}^{-1}(\Omega)$ is also an isomorphism. (A simple sufficient condition for this assumption to be satisfied is that $\gamma(x) \geq 0$.) An application of the open-mapping theorem yields

$$\|v\|_1 \lesssim \|\hat{\mathcal{L}}v\|_{-1} \quad \forall\, v \in \mathcal{H}_0^1(\Omega).$$

It is easy to see that if $\hat{\mathcal{L}}$ satisfies the above assumption and the above estimate, so does its formal adjoint

$$\hat{\mathcal{L}}^*u = -\mathrm{div}(\alpha(x)\nabla u + \beta(x)u) + \gamma(x)u.$$

Namely, $\hat{\mathcal{L}}^* : \mathcal{H}_0^1(\Omega) \to \mathcal{H}^{-1}(\Omega)$ is also an isomorphism.

Corresponding to $\mathcal{L}$ and $\hat{\mathcal{L}}$, we define two bilinear forms, for $u, v \in \mathcal{H}_0^1(\Omega)$, as

$$A(u,v) = \int_\Omega \alpha(x)\nabla u\cdot\nabla v\,dx, \quad \hat{A}(u,v) = A(u,v) + \int_\Omega ((\beta\cdot\nabla u)v + \gamma(x)uv)\,dx. \tag{2.2}$$

We shall often use the following well-known regularity result (cf. Grisvard [27]).

LEMMA 2.1. *If $u \in \mathcal{H}_0^1(\Omega)$ and $\hat{\mathcal{L}}u \in \mathcal{L}^q(\Omega)$ for $1 < q \leq 2$, then $u \in \mathcal{W}_q^2(\Omega)$ and*

$$\|u\|_{2,q} \leq C\|\hat{\mathcal{L}}u\|_{0,q}$$

for some positive constant C depending on q, the coefficients of $\hat{\mathcal{L}}$, and the domain Ω.

2.2. Finite element discretizations and Galerkin projections. We assume that Ω is partitioned by a quasi-uniform triangulation $T_h = \{\tau_i\}$. By this we mean that τ_i's are triangles of size h with $h \in (0,1)$ and $\bar{\Omega} = \cup_i\bar{\tau}_i$ and there exist constants C_0 and C_1 not depending on h such that each element τ_i is contained in (contains) a disc of radius C_1h (respectively, C_0h).

For a given triangulation T_h, a finite element space $\mathcal{V}_h \subset \mathcal{V} \equiv \mathcal{H}_0^1(\Omega)$ is defined by

$$\mathcal{V}_h = \{v \in C(\bar{\Omega}) : v|_\tau \in \mathcal{V}_\tau^r \quad \forall\, \tau \in T_h,\, v|_{\partial\Omega} = 0\},$$

[1]Nonsymmetric and/or indefinite operators or bilinear forms will in general be denoted by symbols with "hats."

where $\mathcal{V}_\tau^r$ is the space of polynomial of degree not greater than a positive integer r. For a given $v \in C(\bar{\Omega})$, $v_I \in \mathcal{V}_h$ will denote the standard nodal value interpolation of v.

It is well known that (cf. [28]) $\mathcal{V}_h$ satisfies the approximation property

$$\inf_{\chi \in \mathcal{V}_h} \{\|v - \chi\|_{0,q} + h\|v - \chi\|_{1,q}\} \lesssim h^{k-2/p+2/q} |v|_{k,p} \tag{2.3}$$

for all $v \in \mathcal{W}_p^k(\Omega) \cap \mathcal{H}_0^1(\Omega)$, $2 \le k \le r+1$, and $1 \le p \le q \le \infty$.

Let $P_h : \mathcal{V} \longrightarrow \mathcal{V}_h$ be the standard Galerkin projection defined by

$$A(P_h v, \chi) = A(v, \chi) \quad \forall\, \chi \in \mathcal{V}_h. \tag{2.4}$$

Using Lemma 2.1 and a standard duality argument, we have

$$\|v - P_h v\| \lesssim h\|v\|_1 \quad \forall\, v \in \mathcal{V}. \tag{2.5}$$

It is proved in Rannacher and Scott [29] that

$$\|P_h v\|_{1,p} \lesssim \|v\|_{1,p} \quad \forall\, v \in \mathcal{V} \cap \mathcal{W}_p^1,\ 1 < p \le \infty. \tag{2.6}$$

For the nonsymmetric and/or indefinite problems, the following result (based on Schatz [30]), is of fundamental importance.

LEMMA 2.2. *If h is sufficiently small, then*

$$\|v_h\|_1 \lesssim \sup_{\varphi \in \mathcal{V}_h} \frac{\hat{A}(v_h, \varphi)}{\|\varphi\|_1} \quad \text{and} \quad \|v_h\|_1 \lesssim \sup_{\varphi \in \mathcal{V}_h} \frac{\hat{A}(\varphi, v_h)}{\|\varphi\|_1} \quad \forall\, v_h \in \mathcal{V}_h. \tag{2.7}$$

The same results are also valid for ε sufficiently small if $\hat{A}$ in (2.7) is replaced by $\hat{A}_\varepsilon$ defined by

$$\hat{A}_\epsilon(u, v) = \int_\Omega (\alpha_\epsilon(x)\nabla u \cdot \nabla v + (\beta_\epsilon \cdot \nabla u)v + \gamma_\epsilon(x)uv)\, dx$$

with the functions $\alpha_\epsilon, \beta_\epsilon, \gamma_\epsilon \in L_\infty(\Omega)$ satisfying

$$\|\alpha - \alpha_\varepsilon\|_{0,\infty} + \|\beta - \beta_\varepsilon\|_{0,\infty} + \|\gamma - \gamma_\varepsilon\|_{0,\infty} = \delta_\varepsilon,$$

where $\delta_\epsilon = o(1)$ as $\varepsilon \to 0$.

Proof. Since $\hat{\mathcal{L}} : \mathcal{H}_0^1(\Omega) \to \mathcal{H}^{-1}(\Omega)$ is an isomorphism, we have

$$\|v_h\|_1 \lesssim \sup_{w \in \mathcal{V}} \frac{\hat{A}(v_h, w)}{\|w\|_1}.$$

Note that by definition and (2.5)

$$\begin{aligned}
\hat{A}(v_h, P_h w) &= \hat{A}(v_h, w) - \hat{A}(v_h, w - P_h w) \\
&= \hat{A}(v_h, w) + (A - \hat{A})(v_h, w - P_h w) \\
&\ge \hat{A}(v_h, w) - c\|v_h\|_1\|w - P_h w\| \\
&\ge \hat{A}(v_h, w) - c_1 h\|v_h\|_1\|w\|_1.
\end{aligned}$$

The proof of the first estimate in (2.7) then follows by using the fact that $\|P_h w\|_1 \lesssim \|w\|_1$. The proof of the second estimate is similar.

For the form $\hat{A}_\epsilon(\cdot,\cdot)$, it follows from the assumption that

$$\hat{A}_\epsilon(v_h, \phi) \geq \hat{A}(v_h, \phi) - c\delta_\epsilon \|v_h\|_1 \|\phi\|_1.$$

The desired result then easily follows if ϵ is sufficiently small. □

Now, define $\hat{P}_h : \mathcal{V} \longrightarrow \mathcal{V}_h$ by

$$\hat{A}(\hat{P}_h v, \chi) = \hat{A}(v, \chi) \quad \forall\, \chi \in \mathcal{V}_h. \tag{2.8}$$

Lemma 2.3 follows from (2.7) and Lemma 2.1.

LEMMA 2.3. *If h is sufficiently small, then $\hat{P}_h$ is well defined and*

$$\|u - \hat{P}_h u\| + h\|u - \hat{P}_h u\|_1 \lesssim h \inf_{\chi\in\mathcal{V}_h} \|u - \chi\|_1 \quad \forall\, u \in \mathcal{V}.$$

More generally, we have (cf. Brenner and Scott [31]) the following.

LEMMA 2.4. *The projection $\hat{P}_h$ admits the estimate*

$$\|u - \hat{P}_h u\|_{0,p} \lesssim h^{r+1}\|u\|_{r+1,p}, \; 2 \leq p < \infty,$$

$$\|u - \hat{P}_h u\|_{0,\infty} \lesssim h^{r+1}|\log h|\|u\|_{r+1,\infty},$$

$$\|u - \hat{P}_h u\|_{1,p} \lesssim \inf_{\chi\in\mathcal{V}_h} \|u - \chi\|_{1,p} \lesssim h^{r+1}\|u\|_{r+1,p}, \; 2 \leq p \leq \infty.$$

We now introduce some discrete Green functions. Given $z \in \Omega$, the Green functions g_h^z, $\hat{g}_h^z \in \mathcal{V}_h$, are defined by

$$A(g_h^z, \chi) = (\partial\chi)(z), \quad \hat{A}(\hat{g}_h^z, \chi) = (\partial\chi)(z) \quad \forall\, \chi \in \mathcal{V}_h \tag{2.9}$$

where ∂ represents either $\frac{\partial}{\partial x_1}$ or $\frac{\partial}{\partial x_2}$.

It has been shown that

$$\sup_{z\in\overline{\Omega}} \|g_h^z\|_{1,1} \lesssim |\log h|. \tag{2.10}$$

A proof of (2.10) can be obtained by using the weighted-norm technique; we refer, for example, to Rannacher and Scott [29] and Thomée, Xu, and Zhang [32] for details.

Similarly, we have

$$\sup_{z\in\overline{\Omega}} \|\hat{g}_h^z\|_{1,1} \lesssim |\log h|. \tag{2.11}$$

Although estimate (2.11) can be established directly as for g_h^z, it can be easily deduced from estimate (2.10). In fact, by definition, it is easy to see that $\|g_h^z - \hat{g}_h^z\|_1 \lesssim \|\hat{g}_h^z\|$. A standard duality argument (using Lemma 2.1) shows that $\|\hat{g}_h^z\| \lesssim h\, \|\hat{g}_h^z\|_1 + 1$. Estimate (2.11) then follows from (2.10).

The following results show that P_h and $\hat{P}_h$ are "superclose" in $\mathcal{H}^1(\Omega)$- and $\mathcal{W}_\infty^1(\Omega)$-norms.

LEMMA 2.5. *Assume that P_h and $\hat{P}_h$ are defined by* (2.4) *and* (2.8), *respectively*; *then*

$$\|P_h u - \hat{P}_h u\|_1 \lesssim \|u - \hat{P}_h u\|,$$

$$\|P_h u - \hat{P}_h u\|_{1,\infty} \lesssim |\log h|\|u - \hat{P}_h u\|_{0,\infty}.$$

Proof. By definition

$$A(P_h u - \hat{P}_h u, \chi) = (A - \hat{A})(u - \hat{P}_h u, \chi) \quad \forall\, \chi \in \mathcal{V}_h.$$

The desired estimates then follow by taking $\chi = P_h u - \hat{P}_h u$ and $\chi = g_h^z$, respectively. □

3. Classic finite element discretizations and $\mathcal{L}^p$ estimates for nonlinear problems. This section is devoted to the standard finite element discretization for nonlinear elliptic boundary value problems. Existence and uniqueness of the finite element solution will be discussed and error estimates in $\mathcal{L}^p$-norms will be derived.

3.1. A model problem and its finite element discretization. We consider the second-order quasi-linear elliptic problem

$$\begin{cases} -\mathrm{div}(F(x,u,\nabla u)) + g(x,u,\nabla u) = 0 & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases} \tag{3.1}$$

We assume that $F(x,y,z) : \overline{\Omega} \times \mathbb{R}^1 \times \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ and $g(x,y,z) : \overline{\Omega} \times \mathbb{R}^1 \times \mathbb{R}^2 \longrightarrow \mathbb{R}^1$ are smooth functions and that (3.1) has a solution $u \in \mathcal{H}_0^1(\Omega) \cap \mathcal{W}_{2+\epsilon}^2(\Omega)$ (for some $\varepsilon > 0$).

For any $w \in \mathcal{W}_\infty^1(\Omega)$, we denote

$$a(w) = D_z F(x,w,\nabla w) \in \mathbb{R}^{2\times 2}, \quad b(w) = D_y F(x,w,\nabla w) \in \mathbb{R}^2,$$
$$c(w) = D_z g(x,w,\nabla w) \in \mathbb{R}^2, \quad d(w) = D_y g(x,w,\nabla w) \in \mathbb{R}^1.$$

The linearized operator $\mathcal{L}$ at w (namely, the Fréchet derivative of $\mathcal{L}$ at w) is then given by

$$\mathcal{L}'(w)v = -\mathrm{div}(a(w)\nabla v + b(w)v) + c(w)\nabla v + d(w)v.$$

Our basic assumptions are, first of all, for the solution u of (3.1)

$$\xi^T a(u)\xi \geq \alpha_0 |\xi|^2 \quad \forall\, \xi \in \mathbb{R}^2, \quad x \in \overline{\Omega}$$

for some constant $\alpha_0 > 0$ and, secondly, $\mathcal{L}'(u) : \mathcal{H}_0^1(\Omega) \to \mathcal{H}^{-1}(\Omega)$ is an isomorphism. As a result of these assumptions, u must be an isolated solution of (3.1).

For convenience of exposition, we introduce two parameters δ_1 and δ_2 as follows:

$$\delta_2 = \begin{cases} 0 & \text{if } D_z^2 F(x,y,z) \equiv 0,\ D_z^2 g(x,y,z) \equiv 0, \\ 1 & \text{otherwise} \end{cases}$$

and

$$\delta_1 = \begin{cases} 0 & \text{if } \quad \delta_2 = 0,\ D_y D_z F(x,y,z) \equiv 0,\ D_y D_z g(x,y,z) \equiv 0, \\ 1 & \text{otherwise.} \end{cases}$$

If $\delta_2 = 0$ and $\delta_1 = 1$, then (3.1) is mildly nonlinear for which

$$\mathcal{L}(u) = -\mathrm{div}(\alpha(x,u)\nabla u + \beta(x,u)) + \gamma(x,u) \cdot \nabla u + g(x,u).$$

If $\delta_1 = \delta_2 = 0$, (3.1) is semilinear for which

$$\mathcal{L}(u) = -\mathrm{div}(\alpha(x)\nabla u + \beta(x,u)) + g(x,u). \tag{3.2}$$

Setting

$$A(v,\varphi) = (F(\cdot,v,\nabla v), \nabla\varphi) + (g(\cdot,v,\nabla v),\varphi),$$

then the solution u of (3.1) satisfies

$$A(u,\chi) = 0 \quad \forall\, \chi \in \mathcal{V}.$$

The classic finite element approximation of (3.1) is to find $u_h \in \mathcal{V}_h$ such that

$$A(u_h, \chi) = 0 \quad \forall\, \chi \in \mathcal{V}_h. \tag{3.3}$$

Introducing the bilinear form (induced by $\mathcal{L}'(w)$)

$$A'(w; v, \varphi) = (a(w)\nabla v + b(w)v, \nabla\varphi) + (c(w)\cdot\nabla v + d(w)v, \varphi),$$

we have the following.

LEMMA 3.1. *For any $v, v_h, \chi \in \mathcal{V}$,*

$$A(v_h, \chi) = A(v, \chi) + A'(v; v_h - v, \chi) + R(v, v_h, \chi). \tag{3.4}$$

Thus $u_h \in \mathcal{V}_h$ solves (3.3) *if and only if*

$$A'(u; u - u_h, \chi) = R(u, u_h, \chi) \quad \forall\, \chi \in \mathcal{V}_h. \tag{3.5}$$

Here the remainder R satisfies, with $e_h = v - v_h$, $1/p + 1/q = 1$, $p, q \geq 1$,

$$|R(v, v_h, \chi)| \leq C(K)(\|e_h\|_{0,2p}^2 + \delta_1\|e_h\nabla e_h\|_{0,p} + \delta_2\|\nabla e_h\|_{0,2p}^2)\ \|\nabla\chi\|_{0,q} \tag{3.6}$$

for any given $K > 0$ and the functions v and v_h satisfying $\|v\|_{1,\infty} + \|v_h\|_{1,\infty} \leq K$.

Proof. Set $\eta(t) = A(v + t(v_h - v), \chi)$. Equation (3.4) follows from the elementary identity

$$\eta(1) = \eta(0) + \eta'(0) + \int_0^1 \eta''(t)(1-t)\,dt$$

with

$$R(v, v_h, \chi) = \int_0^1 \eta''(t)(1-t)\,dt.$$

A straightforward calculation shows that

$$\begin{aligned}\eta''(t) = {} & ((F_{zz}\nabla e_h)\nabla e_h + 2F_{yz}\nabla e_h e_h + F_{yy}e_h^2, \nabla\chi)\\ & + (g_{zz}\nabla e_h\cdot\nabla e_h + 2g_{yz}\nabla e_h e_h + g_{yy}e_h^2, \chi).\end{aligned}$$

The estimate for R follows with an appropriate constant C satisfying

$$C \geq \max_{x\in\bar{\Omega}, |y|\leq K, |z|\leq K} (|F_{zz}| + 2|F_{yz}| + |F_{yy}| + |g_{zz}| + 2|g_{yz}| + |g_{yy}|). \qquad \square$$

LEMMA 3.2. *Let $\hat{P}_h$ be the projection with respect to the bilinear form $A'(u; \cdot, \cdot)$. Then, if h is sufficiently small, the finite element equation* (3.3) *has a solution u_h satisfying*

$$\|u_h - \hat{P}_h u\|_{1,\infty} \leq h^\sigma \quad \textit{and} \quad \|u - u_h\|_{1,\infty} \lesssim h^\sigma \tag{3.7}$$

for some $\sigma > 0$. Furthermore there exists a constant $\eta > 0$ such that u_h is the only solution satisfying

$$\|u_h - u\|_{1,\infty} \leq \eta.$$

Similar existence and uniqueness results can be found in the literature; cf. Rannacher [5]. For completeness, we shall now include a simple proof.

Proof. Define a nonlinear operator $\Phi : \mathcal{V}_h \to \mathcal{V}_h$ by

$$A'(u; \Phi(v_h), \chi) = A'(u; u, \chi) - R(u, v_h, \chi) \quad \forall\, \chi \in \mathcal{V}_h.$$

By (2.7) it is easy to see that Φ is continuous. As $u \in \mathcal{H}_0^1(\Omega) \cap \mathcal{W}_{2+\varepsilon}^2(\Omega)$, there exists a $\sigma > 0$ (by Lemma 2.4) such that

$$\|u - \hat{P}_h u\|_{1,\infty} \lesssim h^\sigma.$$

Defining a set

$$B = \{v \in \mathcal{V}_h : \ \ \|v - \hat{P}_h u\|_{1,\infty} \le h^\sigma\},$$

we claim that $\Phi(B) \subset B$ for sufficiently small h. In fact,

$$A'(u; \Phi(v_h) - \hat{P}_h u, \chi) = -R(u, v_h, \chi).$$

For $v_h \in B$, taking $\chi = \hat{g}_h^z$ (defined by (2.9) with $\hat{A}(\cdot,\cdot) = A'(u;\cdot,\cdot)$) gives

$$\begin{aligned}\|\Phi(v_h) - \hat{P}_h u\|_{1,\infty} &\le C_0 |\log h|\ \|u - v_h\|_{1,\infty}^2\\ &\le 2C_0|\log h|\ (\|u - \hat{P}_h u\|_{1,\infty}^2 + \|v_h - \hat{P}_h u\|_{1,\infty}^2)\\ &\le C_1 |\log h|\ h^{2\sigma} \le h^\sigma\end{aligned}$$

if h is sufficiently small. Thus $\Phi(B) \subset B$. An application of Brouwer's fixed-point theorem shows the existence of a $u_h \in B$ so that $\Phi(u_h) = u_h$. By definition such a u_h satisfies the desired properties.

To prove the uniqueness, let u_h and $\tilde{u}_h$ be two solutions of (3.3) satisfying

$$\|u - u_h\|_{1,\infty} \le \eta \ \text{ and } \ \|u - \tilde{u}_h\|_{1,\infty} \le \eta.$$

Then

$$\int_0^1 A'(u_h + t(\tilde{u}_h - u_h); \tilde{u}_h - u_h, \chi)dt = A(\tilde{u}_h, \chi) - A(u_h, \chi) = 0.$$

By assumption and Lemma 2.2, if η is sufficiently small,

$$\|u_h - \tilde{u}_h\|_1 \lesssim \sup_{\chi \in \mathcal{V}_h} \frac{\int_0^1 A'(u_h + t(\tilde{u}_h - u_h); \tilde{u}_h - u_h, \chi)dt}{\|\chi\|_1} = 0.$$

Thus $\tilde{u}_h = u_h$. □

3.2. $\mathcal{L}^p$ estimates. We shall now derive some $\mathcal{L}^p$ estimates for the finite element approximation. The main ingredient in our approach is the following superconvergence estimate between nonlinear and linearized problems.

LEMMA 3.3. *Assume that u_h and $\hat{P}_h$ are as described in Lemma* 3.2. *Then*

$$\|u_h - \hat{P}_h u\|_1 \lesssim \|u - u_h\|_{0,4}^2 + \delta_1 \|\nabla(u - u_h)^2\| + \delta_2 \|u - u_h\|_{1,4}^2, \tag{3.8}$$

$$\|u_h - \hat{P}_h u\|_{1,\infty} \lesssim |\log h| \left(\|u - u_h\|_{0,\infty}^2 + \delta_1 \|\nabla(u - u_h)^2\|_{0,\infty} + \delta_2 \|u - u_h\|_{1,\infty}^2\right). \tag{3.9}$$

Proof. The estimate (3.8) is obtained by taking $\chi = \hat{P}_h u - u_h$ in (3.5) and applying (3.6). The second estimate is obtained by taking $\chi = \hat{g}_h^z$ in (3.5) with $\hat{A}(\cdot,\cdot) = A'(u;\cdot,\cdot)$. □

THEOREM 3.4. *Assume that $u \in \mathcal{W}_{2+\epsilon}^2(\Omega)$ and $u_h \in \mathcal{V}_h$ are the solutions of* (3.1) *and* (3.3), *respectively, that satisfy* (3.7). *Then*

$$\|u - u_h\|_{1,p} \lesssim h^r \quad \text{if } u \in \mathcal{W}_p^{r+1}(\Omega), \quad 2 \le p \le \infty, \tag{3.10}$$

$$\|u - u_h\|_{0,p} \lesssim h^{r+1} \quad \text{if } u \in \mathcal{W}_p^{r+1}(\Omega), \quad 2 \le p < \infty, \tag{3.11}$$

and

$$\|u - u_h\|_{0,\infty} \lesssim h^{r+1}|\log h| \quad \text{if } u \in \mathcal{W}_\infty^{r+1}(\Omega). \tag{3.12}$$

Proof. We shall divide the proof for (3.10) into five different cases: $p = 2$, $p = \infty$, $p = 4$, $p \geq 4$, and $2 < p < 4$. It follows from (3.8) and (3.7) that

$$\|\hat{P}_h u - u_h\|_1 \lesssim \|u - u_h\|_{1,4}^2 \lesssim \|u - u_h\|_{1,\infty} \|u - u_h\|_1 \lesssim h^\sigma \|u - u_h\|_1.$$

Thus, if h is sufficiently small,

$$\|u - u_h\|_1 \lesssim \|u - \hat{P}_h u\|_1 \lesssim \inf_{\chi \in \mathcal{V}_h} \|u - \chi\|_1.$$

This implies (3.10) for $p = 2$.

By (3.9)

$$\|u - u_h\|_{1,\infty} \lesssim \|u - \hat{P}_h u\|_{1,\infty} + |\log h| h^\sigma \|u - u_h\|_{1,\infty}.$$

Thus if h is sufficiently small,

$$\|u - u_h\|_{1,\infty} \lesssim \|u - \hat{P}_h u\|_{1,\infty} \lesssim \inf_{\chi \in \mathcal{V}_h} \|v - \chi\|_{1,\infty},$$

which implies that (3.10) holds for $p = \infty$ and, by (2.3), that

$$\|u - u_h\|_{1,\infty} \lesssim h^{r-2/p} \text{ if } u \in \mathcal{W}_p^{r+1}(\Omega) \text{ and } 2 \leq p < \infty. \tag{3.13}$$

By (3.8) and (3.9), we have

$$\begin{aligned} \|u_h - \hat{P}_h u\|_{1,4} &\lesssim \|u_h - \hat{P}_h u\|_{1,2}^{1/2} \|u_h - \hat{P}_h u\|_{1,\infty}^{1/2} \\ &\lesssim \|u - u_h\|_{1,4} \|u - u_h\|_{1,\infty}. \end{aligned}$$

Thus if h is sufficiently small, we have

$$\|u - u_h\|_{1,4} \lesssim \|u - \hat{P}_h u\|_{1,4} \lesssim \inf_{\chi \in \mathcal{V}_h} \|u - \chi\|_{1,4},$$

which implies (3.10) for $p = 4$.

Now, we assume that $4 \leq p < \infty$. Again, by (3.8) and (3.9), we have, if $u \in \mathcal{W}_p^{r+1}(\Omega)$,

$$\begin{aligned} \|u_h - \hat{P}_h u\|_{1,p} &\lesssim \|u_h - \hat{P}_h u\|_{1,2}^{2/p} \|u_h - \hat{P}_h u\|_{1,\infty}^{1-2/p} \\ &\lesssim \|u - u_h\|_{1,4}^{4/p} \|u - u_h\|_{1,\infty}^{2-4/p} \\ &\lesssim h^{4r/p} h^{(r-2/p)(2-4/p)} \lesssim h^{2r-1} \lesssim h^r. \end{aligned}$$

This proves (3.10) for $p \geq 4$.

Now, we assume that $2 < p < 4$. It follows from the previous inequalities (2.3) and (3.13) that

$$\begin{aligned} \|u_h - \hat{P}_h u\|_{1,p} &\lesssim \|u - u_h\|_{1,4}^{4/p} \|u - u_h\|_{1,\infty}^{2-4/p} \\ &\lesssim \inf_{\chi \in \mathcal{V}_h} \|u - \chi\|_{1,4}^{4/p} \|u - u_h\|_{1,\infty}^{2-4/p} \\ &\lesssim h^{(r+1/2-2/p)4/p} h^{(r-2/p)(2-4/p)} \lesssim h^{2r-2/p} \lesssim h^r. \end{aligned}$$

This proves (3.10) for $2 < p < 4$. The proof for (3.10) is then complete.

The proof of (3.12) is easy, since, by (3.9), we have

$$\begin{aligned}\|u-u_h\|_{0,\infty} &\le \|u-\hat{P}_h u\|_{0,\infty}+\|u_h-\hat{P}_h u\|_{1,\infty}\\ &\le \|u-\hat{P}_h u\|_{0,\infty}+|\log h|\|u-u_h\|_{1,\infty}^2.\end{aligned}$$

To prove (3.11), we shall apply a duality argument. Consider the auxiliary problem: find $w\in\mathcal{H}_0^1(\Omega)$ such that

$$A'(u;v,w)=(\varphi,v)\quad\forall\, v\in\mathcal{H}_0^1(\Omega).$$

Given $2\le p<\infty$, set $q=p/(p-1)\in(1,2]$. By Lemma 2.1

$$\|w\|_{2,q}\lesssim\|\varphi\|_{0,q}\quad\text{for }\varphi\in\mathcal{L}^q.$$

It follows that

$$\|w-P_h w\|_{1,q}\lesssim h\|w\|_{2,q}\lesssim h\|\varphi\|_{0,q}.$$

If $s>2p/(p+2)$, by (2.6) and a well-known Sobolev imbedding theorem, we have

$$\|P_h w\|_{1,s/(s-1)}\lesssim\|w\|_{1,s/(s-1)}\lesssim\|w\|_{2,q}\lesssim\|\varphi\|_{0,q}.$$

Consequently, with $s>2p/(p+2)$, by Lemma 3.1

$$\begin{aligned}(u-u_h,\varphi)=A'(u;u-u_h,w)&=A'(u;u-u_h,w-P_h w)+A'(u;u-u_h,P_h w)\\ &\lesssim\|u-u_h\|_{1,p}\|w-P_h w\|_{1,q}+\|u-u_h\|_{1,2s}^2\|P_h w\|_{1,s/(s-1)}\\ &\lesssim(h\|u-u_h\|_{1,p}+\|u-u_h\|_{1,2s}^2)\|\varphi\|_{0,q}.\end{aligned}$$

Thus

$$\|u-u_h\|_{0,p}\lesssim h\|u-u_h\|_{1,p}+\|u-u_h\|_{1,2s}^2.$$

For $p=2$, if $u\in\mathcal{W}_2^{r+1}(\Omega)$,

$$\|u-u_h\|_{0,2}\lesssim h\|u-u_h\|_{1,2}+\|u-u_h\|_{1,2+\epsilon}^2\lesssim h^{r+1}.$$

For $2<p<\infty$, since $4p/(p+2)<p$, we have

$$\|u-u_h\|_{0,p}\lesssim h\|u-u_h\|_{1,p}+\|u-u_h\|_{1,p}^2.$$

This yields the estimate (3.11). □

Remark 1. Using the estimate in Lemma 3.3, we can derive many other types of error estimates. For example, the following negative norm estimate can be obtained:

$$\|u-u_h\|_{1-k}\lesssim h^{r+k}\ ,\quad 1\le k\le r$$

provided that $\partial\Omega$ is sufficiently smooth and the elements near the curved boundary are appropriately constructed (cf. Scott [33]).

To see the above estimate, consider an auxiliary problem: find $w\in\mathcal{H}_0^1(\Omega)$ such that

$$A'(u;v,w)=(\varphi,v)\quad\forall\, v\in\mathcal{H}_0^1(\Omega).$$

By the well-known regularity estimates (cf. [34]),

$$\|w\|_{k+1}\lesssim\|\varphi\|_{k-1}\qquad\text{for }\varphi\in\mathcal{H}^{k-1}(\Omega).$$

Thus

$$
\begin{aligned}
(u-u_h,\varphi) = A'(u;u-u_h,w) &= A'(u;u-u_h,w-w_I)+A'(u;u-u_h,w_I)\\
&\lesssim \|u-u_h\|_1\|w-w_I\|_1+\|u-u_h\|_{1,4}^2\|w_I\|_1\\
&\lesssim (h^k\|u-u_h\|_1+\|u-u_h\|_{1,4}^2)\|w\|_{k+1}\\
&\lesssim h^{k+r}\|\varphi\|_{k-1}.
\end{aligned}
$$

The desired estimate then follows.

4. Two-grid discretization for non-SPD linear problems. In this section, we shall present a number of algorithms for non-SPD problems based on two finite element spaces. The idea is to reduce a non-SPD problem into a SPD problem by solving a non-SPD problem on a much smaller space.

The basic mechanisms in our approach are two quasi-uniform triangulations of Ω, T_H, and T_h, with two different mesh sizes H and h ($H > h$), and the corresponding finite element spaces $\mathcal{V}_H$ and $\mathcal{V}_h$ which will be called coarse and fine space, respectively. In the applications given below, we shall always assume that

$$
H = O(h^\lambda) \quad \text{for some } 0<\lambda<1. \tag{4.1}
$$

With the bilinear form $\hat{A}$ defined in (2.2), for h sufficiently small, let $u_h \in \mathcal{V}_h$ be the unique solution of

$$
\hat{A}(u_h,\chi)=(f,\chi) \quad \forall\, \chi\in\mathcal{V}_h
$$

and denote the bilinear form of the lower-order terms of the operator $\hat{\mathcal{L}}$ (in (2.1)) by

$$
N(v,\chi)=(\hat{A}-A)(v,\chi)=(\beta\cdot\nabla v,\chi)+(\gamma v,\chi).
$$

Let us now present our first two-grid algorithm.

ALGORITHM 4.1.
1. *Find* $u_H\in\mathcal{V}_H$ *such that* $\hat{A}(u_H,\varphi)=(f,\varphi)\quad \forall\,\varphi\in\mathcal{V}_H$.
2. *Find* $u^h\in\mathcal{V}_h$ *such that* $A(u^h,\chi)+N(u_H,\chi)=(f,\chi)\quad\forall\,\chi\in\mathcal{V}_h$.

We note that the linear system in the second step of the above algorithm is SPD.

THEOREM 4.2. *Assume* $u^h\in\mathcal{V}_h$ *is the solution obtained by Algorithm* 4.1 *for* H *sufficiently small; then*

$$
\|u_h-u^h\|_1\lesssim H^{r+1}\|u\|_{r+1} \quad \textit{and} \quad \|u-u^h\|_1\lesssim (h^r+H^{r+1})\|u\|_{r+1}
$$

provided that $u\in\mathcal{H}^{r+1}(\Omega)$.

Proof. A direct calculation and an application of Lemma 2.4 shows that

$$
\begin{aligned}
A(u_h-u^h,\chi) &= -N((I-\hat{P}_H)u_h,\chi)\\
&\lesssim \|(I-\hat{P}_H)u_h\|\ \|\chi\|_1\\
&\lesssim (H\|u-u_h\|_1+\|(I-\hat{P}_H)u\|)\ \|\chi\|_1\\
&\lesssim H^{r+1}\ \|u\|_{r+1}\ \|\chi\|_1.
\end{aligned}
$$

The desired result then follows. $\square$

Remark 2. If $\beta(x)=0$ and $r\ge 2$, we have

$$
\|P_hu-u^h\|_1\lesssim\|u-u_H\|_{-1}\lesssim H^{r+2}\|u\|_{r+2}
$$

and

$$\|u - u^h\|_1 \lesssim (h^r + H^{r+2})\|u\|_{r+2}.$$

Algorithm 4.1 can be applied in a successive fashion.

ALGORITHM 4.3. *Let $u_h^0 = 0$; assume that $u_h^k \in \mathcal{V}_h$ has been obtained, $u_h^{k+1} \in \mathcal{V}_h$ is defined as follows:*

1. *Find $e_H \in \mathcal{V}_H$ such that* $\hat{A}(e_H + u_h^k, \varphi) = (f, \varphi) \quad \forall\, \varphi \in \mathcal{V}_H$.
2. *Find $u^h \in \mathcal{V}_h$ such that* $A(u_h^{k+1}, \chi) + N(u_h^k + e_H, \chi) = (f, \chi) \quad \forall\, \chi \in \mathcal{V}_h$.

It is well known that most linear iterative methods for solving algebraic systems can be obtained by an appropriate matrix (or operator) splitting. For the nonsymmetric system under consideration, the most natural splitting would lead to the iterative method

$$A(u_h^{k+1}, \chi) + N(u_h^k, \chi) = (f, \chi) \quad \forall\, \chi \in \mathcal{V}_h.$$

This iterative scheme, however, is not convergent in general. Algorithm 4.3 may be considered as a modification of this "natural" iterative scheme with recourse to an additional coarse space.

THEOREM 4.4. *Assume $u_h^k \in \mathcal{V}_h$ is the solution obtained by Algorithm* 4.3 *for $k \geq 1$; then*

$$\|u_h - u_h^k\|_1 \lesssim H^{k+r}\|u\|_{r+1}$$

and

$$\|u - u_h^k\|_1 \lesssim (h^r + H^{k+r})\|u\|_{r+1}.$$

Proof. By definition and Lemma 2.4,

$$\begin{aligned} A(u_h - u_h^k, \chi) &= N((I - \hat{P}_H)(u_h^{k-1} - u_h), \chi) \\ &\leq \|(I - \hat{P}_H)(u_h^{k-1} - u_h)\| \ \|\chi\|_1 \\ &\lesssim H \ \|u_h^{k-1} - u_h\|_1 \ \|\chi\|_1. \end{aligned}$$

This implies

$$\|u_h - u_h^k\|_1 \lesssim H\|u_h - u_h^{k-1}\|_1.$$

Applying the above estimate successively and then using Theorem 4.2 yields

$$\|u_h - u_h^k\|_1 \lesssim H^{k-1}\|u_h - u_h^1\|_1 \lesssim H^{k+r}\|u\|_{r+1}. \qquad \Box$$

Remark 3. The SPD system in the step of Algorithm 4.3 may not be solved exactly. The corresponding algorithms can be found in Xu [1–2].

Before ending this section, we present an algorithm for a symmetric and indefinite problem (namely, $\beta(x) = 0$ in (2.1)). This algorithm is based on the finite element space

$$\hat{\mathcal{V}}_h = (I - \hat{P}_H)\mathcal{V}_h.$$

ALGORITHM 4.5.

1. *Find $u_H \in \mathcal{V}_H$ such that* $\hat{A}(u_H, \varphi) = (f, \varphi) \quad \forall\, \varphi \in \mathcal{V}_H$.
2. *Find $e_h \in \hat{\mathcal{V}}_h$ such that* $A(e_h, \chi) = (f, \chi) \quad \forall\, \chi \in \hat{\mathcal{V}}_h$.
3. $u^h = u_H + e_h$.

We note that the system in the second step of Algorithm 4.5 is SPD. But since it is on the space $\hat{\mathcal{V}}_h$, this system may not be solved very easily. Nevertheless, this algorithm is of certain theoretical interest. In fact, as shown in the next theorem,

$$\|u - u^h\|_1 \lesssim (h + H^3)\|u\|_2$$

if the linear finite element is used.

THEOREM 4.6. *Assume $u^h \in \mathcal{V}_h$ is obtained by Algorithm 4.5; then*

$$\|u - u^h\|_1 \lesssim (h^r + H^{r+2})\|u\|_{r+1}.$$

Proof. As $\hat{A}$ is symmetric, so is $\hat{P}_H$. Thus

$$\hat{A}((I - \hat{P}_H)u, \chi) = (f, \chi) \quad \forall\, \chi \in \hat{\mathcal{V}}_h.$$

Therefore

$$A(u_h - (u_H + e_h), \chi) = -(\gamma(u - u_H), \chi) \lesssim H\|u - u_H\|\|\chi\|_1 \lesssim H^{r+2}\|u\|_{r+1}\|\chi\|_1,$$

where we have used the fact that $\|\chi\| \lesssim H\|\chi\|_1$ for $\chi \in \hat{\mathcal{V}}_h$. The desired result follows by taking $\chi = u_h - (u_H + e_h) \in \hat{\mathcal{V}}_h$. $\square$

5. Two-grid methods for nonlinear problems. This section is devoted to some discretization techniques based on two (or more than two) finite element subspaces. The first subsection is on some simple techniques for some mildly nonlinear equations and the rest of the section is devoted to some two-grid methods based on Newton's method for general quasi-linear equations.

5.1. Some simple two-grid methods. The techniques presented here are similar to those for algorithms for non-SPD linear problems in §4 and they will be applied to the mildly quasi-linear equation

$$\text{(5.1)} \qquad \begin{cases} -\mathrm{div}(\alpha(x,u)\nabla u + \beta(x,u)) + \gamma(x,u)\cdot\nabla u + g(x,u) = 0 & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$

This equation is a special case of (3.1) with $\delta_2 = 0$, $F(x,y,z) = \alpha(x,y)z + \beta(x,y)$, and $g(x,y,z) = \gamma(x,y)\cdot z + g(x,y)$. We assume that the early assumptions on (3.1) all hold here.

To state the algorithm we define, for $u, v, \chi \in \mathcal{W}^1_\infty(\Omega) \cap \mathcal{H}^1_0(\Omega)$,

$$\tilde{A}(u; v, \chi) = (\alpha(\cdot,u)\nabla v + \beta(\cdot,u), \nabla\chi) + (\gamma(\cdot,u)\cdot\nabla v + f(\cdot,u), \chi).$$

Our first algorithm is a nonlinear extension of Algorithm 4.1.

ALGORITHM 5.1.
1. *Find $u_H \in \mathcal{V}_H$ such that* $\quad A(u_H, \varphi) = 0 \quad \forall\, \varphi \in \mathcal{V}_H$.
2. *Find $u^h \in \mathcal{V}_h$ such that* $\quad \tilde{A}(u_H; u^h, \chi) = 0 \quad \forall\, \chi \in \mathcal{V}_h$.

THEOREM 5.2. *Assume $u^h \in \mathcal{V}_h$ is obtained by Algorithm 5.1; then*

$$\begin{aligned} \|u_h - u^h\|_1 &\lesssim H^{r+1} && \text{if } u \in \mathcal{H}^{r+1}(\Omega). \\ \|u_h - u^h\|_{1,\infty} &\lesssim H^{r+1}|\log h| && \text{if } u \in \mathcal{W}^{r+1}_\infty(\Omega). \end{aligned}$$

Proof. By definition, it follows that

$$\tilde{A}(u_H; u_h - u^h, \chi) = \tilde{A}(u_H; u_h, \chi) - \tilde{A}(u_h; u_h, \chi) \lesssim \|u_h - u_H\|_{-1}\,\|\chi\|_1.$$

Thus

$$\|u_h - u^h\|_1 \lesssim \|u_h - u_H\|_{-1} \le \|u_h - u_H\|.$$

The first estimate then follows by combining the above estimate with Lemma 2.3. Now let $\hat{g}_h^z$ be the Green function defined in (2.9) with $\hat{A}(\cdot,\cdot) = \tilde{A}(u;\cdot,\cdot)$. Then, with $e_H^h = u_h - u^h$,

$$\begin{aligned}\partial\, e_H^h(z) = \tilde{A}(u; e_H^h, \hat{g}_h^z) &= \tilde{A}(u; e_H^h, \hat{g}_h^z) - \tilde{A}(u_H; e_H^h, \hat{g}_h^z) + A(u_H, e_H^h, \hat{g}_h^z)\\ &\lesssim \|u_h - u_H\|_{0,\infty}\, \|e_H^h\|_{1,\infty}\, \|\hat{g}_h^z\|_{1,1} + \|u_h - u_H\|_{0,\infty} \|g_h^z\|_{1,1}\\ &\lesssim\; H^{r+1}|\log h|\, \|e_H^h\|_{1,\infty} + H^{r+1}|\log\ h|.\end{aligned}$$

This implies, for some constant $c > 0$,

$$\left(1 - c\, H^{r+1}|\log\ h|\right)\ \|e_H^h\|_{1,\infty} \lesssim H^{r+1}|\log\ h|.$$

The desired result then follows if H is so small that $H^{r+1}|\log\ h| \ll 1$ (see (4.1)). $\square$

Remark 4. If $r \ge 2$, we could use the negative norm estimate to conclude that

$$\|u_h - u^h\|_1 \lesssim \|u_h - u_H\|_{-1} \lesssim H^{r+2}.$$

Next we shall present an algorithm that results by combining Algorithm 5.1 with Algorithm 4.1. This algorithm reduces a nonlinear problem to a SPD linear problem and a nonlinear system of smaller size.

Define

$$A_s(u; v, \chi) = (\alpha(\cdot, u)\nabla v, \nabla\chi)$$

and

$$N(u; v, \chi) = (\beta(\cdot, u), \nabla\chi) + (\gamma(\cdot, u)\cdot\nabla v + f(\cdot, u), \chi).$$

ALGORITHM 5.3.
1. *Find* $u_H \in \mathcal{V}_H$ *such that* $A(u_H, \varphi) = 0 \quad \forall\, \varphi \in \mathcal{V}_H$.
2. *Find* $u^h \in \mathcal{V}_h$ *such that* $A_s(u_H; u^h, \chi) + N(u_H; u_H, \chi) = 0 \quad \forall\, \chi \in \mathcal{V}_h$.

THEOREM 5.4. *Assume that* $u^h \in \mathcal{V}_h$ *is obtained by Algorithm* 5.3; *then*

$$\begin{aligned}\|u_h - u^h\|_1 \quad &\lesssim \quad H^{r+1} &&\text{if } u \in \mathcal{H}^{r+1}(\Omega),\\ \|u_h - u^h\|_{1,\infty} \quad &\lesssim \quad H^{r+1}|\log\ h| &&\text{if } u \in \mathcal{W}_\infty^{r+1}(\Omega).\end{aligned}$$

Proof. By definition

$$\begin{aligned}A_s(u_H; u_h - u^h, \chi) = &\ A_s(u_H; u_h, \chi) - A_s(u_h; u_h, \chi)\\ &- N(u_h; u_h, \chi) + N(u_H; u_H, \chi)\\ \lesssim &\ \|u_H - u_h\|\ \|\chi\|_1.\end{aligned}$$

The desired result then follows easily. $\square$

5.2. Correction by one Newton iteration on the fine space. Unlike in the last subsection, the techniques here apply to the general quasi-linear equation (3.1).

Our first algorithm, roughly speaking, is to use the coarse grid approximation as an initial guess for one Newton iteration on the fine grid.

ALGORITHM 5.5.
1. *Find* $u_H \in \mathcal{V}_H$ *such that* $A(u_H, \varphi) = 0 \quad \forall \varphi \in \mathcal{V}_H$.
2. *Find* $u^h \in \mathcal{V}_h$ *such that* $A'(u_H; u^h, \chi) = A'(u_H; u_H, \chi) - A(u_H, \chi) \quad \forall \chi \in \mathcal{V}_h$.

LEMMA 5.6. *Assume that* u^h *is the solution obtained by Algorithm* 5.5; *then*

$$\|u_h - u^h\|_1 \lesssim \|u_h - u_H\|_{0,4}^2 + \delta_1 \|(u_h - u_H)^2\|_{1,4} + \delta_2 \|u_h - u_H\|_1^2,$$

$$\|u_h - u^h\|_{1,\infty} \lesssim |\log h| (\|u_h - u_H\|_{0,\infty}^2 + \delta_1 \|(u_h - u_H)^2\|_{1,\infty} + \delta_2 \|u_h - u_H\|_{1,\infty}^2).$$

Proof. By definition and (3.4),

$$A'(u_H; u_h - u^h, \chi) = A'(u_H; u_h - u_H, \chi) + A(u_H, \chi) = -R(u_H, u_h, \chi).$$

The first estimate follows from Lemma 2.2 (with $\epsilon = H$), (2.7), and (3.6). The proof of the second estimate is similar to the proof of Theorem 5.2 by using (3.6). □

As a direct consequence of Lemma 5.6 and Theorem 3.4, we have the following.

THEOREM 5.7. *Assume* u_h *is the solution obtained by Algorithm* 5.5. *If* $u \in \mathcal{W}_4^{r+1}(\Omega)$, *then*

$$\|u_h - u^h\|_1 \lesssim (H^{2r+2} + \delta_1 H^{2r+1} + \delta_2 H^{2r}) \lesssim H^{2r}.$$

If $u \in \mathcal{W}_\infty^{r+1}(\Omega)$, *then*

$$\|u_h - u^h\|_{1,\infty} \lesssim (H^{2r+2} + \delta_1 H^{2r+1} + \delta_2 H^{2r}) |\log h| \lesssim H^{2r} |\log h|.$$

Thus if $h = O(H^{2+2/r} + \delta_1 H^{2+1/r} + \delta_2 H^{2r})$,

$$\|u - u^h\|_1 \lesssim h^r \text{ and } \|u - u^h\|_{1,\infty} \lesssim h^r |\log h|.$$

To see the efficiency of Algorithm 5.5, we single out a special case of the above theorem.

COROLLARY 5.8. *If Algorithm* 5.5 *is applied to the semilinear equation* (3.2) *with the linear finite element discretization, then*

$$\begin{array}{lll} \|u_h - u^h\|_1 & \lesssim & H^4 \quad\quad\quad\quad \text{if } u \in \mathcal{W}_4^2(\Omega), \\ \|u_h - u^h\|_{1,\infty} & \lesssim & H^4 |\log h| \quad \text{if } u \in \mathcal{W}_\infty^2(\Omega). \end{array}$$

According to Corollary 5.8, in order to obtain the optimal (or nearly optimal) approximation for the discretization u^h, it suffices to take $H = O(h^{\frac{1}{4}})$. To get an idea numerically if the fine mesh size $h = 2^{-16}$ gives $\dim \mathcal{V}_h \approx 3.3 \times 10^9$, the coarse mesh size H could be $H = h^{\frac{1}{4}} = 1/16$ which gives $\dim \mathcal{V}_H \approx 225$.

5.3. Correction by two Newton iterations on the fine space. The algorithms presented above can be greatly improved if one further Newton iteration is carried out on $\mathcal{V}_h$.

Corresponding to Algorithm 5.1, we have the following.

ALGORITHM 5.9.
1. *Find* $u_H \in \mathcal{V}_H$ *such that* $A(u_H, \varphi) = 0 \quad \forall \varphi \in \mathcal{V}_H$.
2. *Find* $u^h \in \mathcal{V}_h$ *such that* $\tilde{A}(u_H; u^h, \chi) = 0 \quad \forall \chi \in \mathcal{V}_h$.
3. *Find* $u_h^* \in \mathcal{V}_h$ *such that* $A'(u^h; u_h^*, \chi) = A'(u^h; u^h, \chi) - A(u^h, \chi) \quad \forall \chi \in \mathcal{V}_h$.

Corresponding to Algorithm 5.5, we have

ALGORITHM 5.10.
1. *Find* $u_H \in \mathcal{V}_H$ *such that* $A(u_H, \varphi) = 0 \quad \forall \varphi \in \mathcal{V}_H$.
2. *Find* $u^h \in \mathcal{V}_h$ *such that* $A'(u_H; u^h, \chi) = A'(u_H; u_H, \chi) - A(u_H, \chi) \quad \forall \chi \in \mathcal{V}_h$.
3. *Find* $u_h^* \in \mathcal{V}_h$ *such that* $A'(u^h; u_h^*, \chi) = A'(u^h; u^h, \chi) - A(u^h, \chi) \quad \forall \chi \in \mathcal{V}_h$.

With arguments similar to those in the preceding subsection, we can obtain various results as follows.

LEMMA 5.11. *For both Algorithms* 5.9 *and* 5.10,

$$\|u_h - u_h^*\|_{1,\infty} \lesssim \|u_h - u^h\|_{1,\infty}^2.$$

Thus for Algorithm 5.9

$$\|u_h - u_h^*\|_{1,\infty} \lesssim H^{2r+2}|\log h|^2,$$

and for Algorithm 5.10

$$\|u_h - u_h^*\|_{1,\infty} \lesssim (H^{4r+4} + \delta_1 H^{4r+2})|\log h|^2.$$

THEOREM 5.12. *For Algorithm* 5.9, *if* $h = O(H^{2+2/r})$, *then*

$$\|u - u_h^*\|_{1,\infty} \lesssim h^r|\log h|.$$

If $h = O(H^2)$, *then*

$$\|u - u_h^*\|_{0,\infty} \lesssim h^{r+1}|\log h|^2.$$

THEOREM 5.13. *For Algorithm* 5.10, *if* $h = O(H^{4+4/r} + \delta_1 H^{4+2/r})$, *then*

$$\|u - u_h^*\|_{1,\infty} \lesssim h^r|\log h|.$$

If $h = O(H^4 + \delta_1 H^{4-2/(r+1)})$, *then*

$$\|u - u_h^*\|_{0,\infty} \lesssim h^{r+1}|\log h|^2.$$

Again, to get an idea of the efficiency of Algorithm 5.10, we have the following.

COROLLARY 5.14. *If Algorithm* 5.10 *is applied to the semilinear equation* (3.4) *with the linear finite element discretization, then*

$$\|u_h - u^h\|_{1,\infty} \lesssim H^8|\log h|^2$$

provided that $u \in \mathcal{W}_\infty^2(\Omega)$.

5.4. Multilevel linearization. From the above discussions, it appears that the algorithm using two subspaces would suffice for most practical applications. Nevertheless, the algorithm can be made more general and perhaps more robust if multiple subspaces are used.

Assume that we are given a sequence of subspaces

$$\mathcal{V}_i = \mathcal{V}_{h_i} \subset \mathcal{V} \quad 0 \le i \le J.$$

Conceivably, we have $h_J \ll h_{J-1} \ll \cdots \ll h_0 = H \ll 1$.

ALGORITHM 5.15.
1. *Find* $u_0 \in \mathcal{V}_0$ *such that* $A(u_0, \varphi) = 0 \quad \forall\, \varphi \in \mathcal{V}_0$.
2. *For* $j = 1, 2, \ldots, J$, *find* $u_j \in \mathcal{V}_j$ *such that*

$$A'(u_{j-1}; u_j, \chi) = A'(u_{j-1}; u_{j-1}, \chi) - A(u_{j-1}, \chi) \quad \forall\, \chi \in \mathcal{V}_j.$$

The above algorithm is similar to the so-called projective Newton method [3], [4].

LEMMA 5.16. *If* H *is sufficiently small,*

$$\|u - u_j\|_{1,\infty} \lesssim h_j^r + |\log h_j|\, \|u - u_{j-1}\|_{1,\infty}^2.$$

Proof. Similar to the proof of Lemma 5.6, we have

$$A'(u_{j-1}; u - u_j, \chi) = -R(u_{j-1}, u, \chi) \quad \forall\, \chi \in \mathcal{V}_j.$$

The desired estimate can then be obtained in a way similar to proofs of Theorem 5.2 and Lemma 5.6. □

THEOREM 5.17. *If H is sufficiently small, $h_1 = (H^{2+2/r} + \delta_1 H^{2+1/r})|\log H|^{1/r}$ and*

$$c|\log h_{j-1}|^{\frac{1}{r}} h_{j-1}^2 \le h_j < h_{j-1} \quad 2 \le j \le J$$

for some appropriate positive constant c. Then

$$\|u - u_j\|_{1,\infty} \le c_1 h_j^r, \quad 1 \le j \le J. \tag{5.2}$$

Proof. By Lemma 5.11, there exists a constant $c_0 > 0$ such that

$$\|u - u_j\|_{1,\infty} \le c_0(h_j^r + |\log h_{j-1}|\|u - u_{j-1}\|_{1,\infty}^2).$$

By Theorems 3.4 and 5.7, if H is sufficiently small, the estimate (5.2) holds for $j = 0, 1$ with some constant $c_1 > 0$. Without loss of generality, we may assume $c_1 = 2c_0$. Assume now that (5.2) is valid for $j - 1$; then

$$\|u - u_j\|_{1,\infty} \le c_0(h_j^r + c_1^2|\log h_{j-1}|h_{j-1}^{2r}) \le c_0(1 + c_1^2 c^{-1})h_j^r.$$

By induction, (5.2) holds with $c = c_1^2 = 4c_0^2$. □

A weaker form of the estimate in the above theorem has been conjectured by Rannacher [5] and Bank [7] and recently proved by Rannacher [6].

Note that if Theorem 5.12 is applied to semilinear equations with linear finite element discretization, one may take $h_1 = H^4|\log H|$ and $h_2 = H^8|\log H|^2$.

COROLLARY 5.18. *If H is sufficiently small and, for some constants $\eta \in (0, 1)$,*

$$\eta h_{j-1} \le h_j < h_{j-1}, \qquad 1 \le j \le J,$$

then

$$\|u - u_j\|_{1,\infty} \lesssim h_j^r, \quad 0 \le j \le J.$$

A result similar to the above corollary was contained in Bank [7] on his multigrid method for solving the nonlinear Galerkin equation (3.3).

6. Concluding remarks. The algorithms studied in this paper are potentially efficient for solving a large class of linear and nonlinear problems. Although our presentation has been confined to the second-order elliptic boundary value problems, the techniques can naturally be applied to other types of problems as well. Roughly speaking, different aspects of a complex problem can be treated by spaces of different scales. In the examples studied in this paper, a very coarse grid space is sufficient for some nonsymmetric problems that are dominated (in a certain analytic sense) by their symmetric part, and is also sufficient to handle the nonlinearity for some mildly nonlinear problems. Symmetry versus nonsymmetry and linearity versus nonlinearity may not make a substantial difference on the analytical level, but their numerical approximation may differ considerably. The two-grid methods studied in this paper provide a new approach to take advantage of some "nice properties" hidden in a complex problem.

An important aspect of our two-grid algorithms is that they can be naturally applied together with multigrid and domain decomposition methods. Most domain decomposition methods, for example, are in a certain sense two-grid methods. The set of subdomains gives rise to a natural coarse grid. Hence the domain decomposition techniques fit perfectly well with our algorithms and the coarse grid plays two different important roles in such an application. Similar arguments also apply to multigrid methods. Suppose we have multiple subspaces

$\mathcal{V}_0 \subset \mathcal{V}_1 \subset \cdots \subset \mathcal{V}_j \subset \mathcal{H}_0^1$. Naturally we can choose $\mathcal{V}_H = \mathcal{V}_0$ (and $\mathcal{V}_h = \mathcal{V}_j$, of course) in our two-grid algorithms.

Applications of multigrid and domain decomposition methods with our two-grid methods for nonlinear problems are satisfying from both theoretical and practical points of view, since the systems on the fine grid are all linear and, hence, theories and numerical codes for linear problems can be adopted with few modifications.

The linear systems on the fine space in the algorithms presented in §4 are SPD and their solution methods have been well developed; we refer to Xu [35] for a summary of these methods. The linear systems from the fine space on the algorithms in §5 are mostly non-SPD and may be solved by combining the algorithms in §4. As a result, a nonlinear system on the fine space may be reduced to few SPD linear systems on the fine space together with some linear and nonlinear systems on the coarse space.

Some two-grid methods have been further improved for semilinear problems in Xu [36]. Some numerical examples on the performances of these algorithms can be found in Xu [36].

REFERENCES

[1] J. Xu, *A new class of iterative methods for nonselfadjoint or indefinite problems*, SIAM J. Numer. Anal., 29 (1992), pp. 303–319.

[2] ———, *Iterative methods by SPD and small subspace solvers for nonsymmetric or indefinite problems*, in Proc. Fifth International Symposium on Domain Decomposition Methods for Partial Differential Equations, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1992, pp. 106–118.

[3] G. Vainikko, *Galerkin's perturbation method and the general theory of approximation for nonlinear equations*, U.S.S.R. Comput. Math. and Math. Phys., 7 (1967), pp. 1–41.

[4] K. Witsch, *Projektive Newton-Verfahren und Anwendungen auf nichtlineare Randwertaufgaben*, Numer. Math., 31 (1978/79), pp. 209–230.

[5] R. Rannacher, *On finite element approximation of general boundary value problems in nonlinear elasticity*, Calcolo, 17 (1980), pp. 175–193.

[6] ———, *On the convergence of the Newton-Raphson method for strongly nonlinear finite element equations*, in Nonlinear Computational Mechanics, P. Wriggers and W. Wagner, eds., Springer-Verlag, Berlin, New York, 1991.

[7] R. Bank, *Analysis of a multilevel iterative method for nonlinear finite element equations*, Math. Comp., 39 (1982), pp. 453–465.

[8] A. Brandt, *Multilevel adaptive solutions to boundary-value problems*, Math. Comp., 31 (1977), pp. 333–390.

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

[10] A. Reusken, *Convergence of the multigrid full approximation scheme for a class of elliptic mildly nonlinear boundary value problems*, Numer. Math., 52 (1988), pp. 251–277.

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

[12] E. Allgower, K. Böhmer, F. Portra, and W. Rheinboldt, *A mesh-independence principle for operator equations and their discretizations*, SIAM J. Numer. Anal., 23 (1986), pp. 160–169.

[13] E. Allgower and K. Böhmer, *Application of the mesh independence principle to mesh refinement strategies*, SIAM J. Numer. Anal., 24 (1987), pp. 1335–1351.

[14] P. Deuflhard and F. Potra, *Asymptotic mesh independence of Newton-Galerkin methods via a refined Mysovskii theorem*, Preprint SC 90-0, Konrad-Zuse-Zentrum für Informationstechnik, Berlin, 1990.

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

[16] M. Schultz, *L^2-error bounds for the Rayleigh-Ritz-Galerkin method*, SIAM J. Numer. Anal., 8 (1971), pp. 737–748.

[17] J. Douglas and T. Dupont, *A Galerkin method for a nonlinear Dirichlet problem*, Math. Comp., 29 (1975), 689–696.

[18] J. Nitsche, *On L_∞-convergence of finite element approximations to the solution of a nonlinear boundary value problem*, Topics Numer. Anal. III, 1977.

[19] C. Johnson and V. Thomée, *Error estimates for a finite element approximation of a minimal surface*, Math. Comp., 29, (1975), pp. 343–349.

[20] R. Rannacher, *Some asymptotic estimates for finite element approximation of minimal surfaces*, RAIRO Sér. Rouge Anal. Numer., 11 (1977), pp. 181–196.

[21] J. Frehse, *Eine gleichmäßige asymptotische Gehlerabschätzung zur methode der finiten Elemente bei quasilinearen Randwertproblemen*, in Theory of Nonlinear Operators: Constructive Aspects, Tagungsband der Akademie der Wissenschaften, Berlin (DDR), 1976.

[22] H. Mittelmann, *On pointwise estimates for a finite element solution of nonlinear boundary value problems*, SIAM J. Numer. Anal., 14 (1977), pp. 773–778.

[23] J. Frehse and R. Rannacher, *Optimal uniform convergence for the finite element approximation of a quasi-linear elliptic boundary value problem*, in Proc. U.S.-Germany Symposium on Formulation and Computational Aglorithms in Finite Element Analysis, M.I.T., Cambridge, MA, 1976.

[24] ———, *Asymptotic L^∞-error estimates for linear finite element approximations of quasi-linear boundary value problems*, SIAM J. Numer. Anal., 15 (1978), pp. 418–431.

[25] J. Nitsche, *Finite element approximations for solving the elasticity problem*, in Computing Methods in Applied Sciences and Engineering, R. Glowinski and J. Lions, eds., Springer-Verlag, Berlin, New York, 1976, pp. 154–166.

[26] M. Dobrowolski and R. Rannacher, *Finite element methods for nonlinear elliptic systems of second order*, Math. Nachr., 94 (1980), pp. 155–172.

[27] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Pitman, Boston, MA, 1985.

[28] P. Ciarlet, *The Finite Element Method for Elliptic Problems*, North–Holland, New York, 1978.

[29] R. Rannacher and R. Scott, *Some optimal error estimates for piecewise linear finite element approximation*, Math. Comp., 54 (1982), pp. 437–445.

[30] A. Schatz, *An observation concerning Ritz-Galerkin methods with indefinite bilinear forms*, Math. Comp., 28 (1974), pp. 959–962.

[31] S. Brenner and R. Scott, *The Mathematical Theory of Finite Element Methods*, Springer-Verlag, Berlin, New York, 1994.

[32] V. Thomée, J. Xu, and N. Zhang, *Superconvergence of gradient in piecewise linear finite element approximation to a parabolic problem*, SIAM J. Numer. Anal., 26 (1989), pp. 389–414.

[33] R. Scott, *Interpolated boundary conditions in the finite element method*, SIAM J. Numer. Anal., 12 (1975), pp. 407–427.

[34] D. Gilbarg and N. Trudinger, *Elliptic Partial Differential Equations of Second Order*, 2nd edition, Springer-Verlag, Berlin, New York, 1983.

[35] J. Xu, *Iterative methods by space decomposition and subspace correction*, SIAM Rev., 34 (1992), pp. 581–613.

[36] ———, *A novel two-grid method for semilinear equations*, SIAM J. Sci. Comput., 15 (1994), pp. 231–237.