

# AN IMPROVED METHOD FOR SOLVING QUASILINEAR CONVECTION DIFFUSION PROBLEMS ON A COARSE MESH

SARA POLLOCK

**ABSTRACT.** A method is developed for solving quasilinear convection diffusion problems starting on a coarse mesh where the data and solution-dependent coefficients are unresolved, the problem is unstable and approximation properties do not hold. The Newton-like iterations of the solver are based on the framework of regularized pseudo-transient continuation where the proposed time integrator is a variation on the Newmark strategy, designed to introduce controllable numerical dissipation and to reduce the fluctuation between the iterates in the coarse mesh regime where the data is rough and the linearized problems are badly conditioned and possibly indefinite. An algorithm and updated marking strategy is presented to produce a stable sequence of iterates as boundary and internal layers in the data are captured by adaptive mesh partitioning. The method is suitable for use in an adaptive framework making use of local error indicators to determine mesh refinement and targeted regularization. Derivation and  $q$ -linear local convergence of the method is established, and numerical examples demonstrate the theory including the predicted rate of convergence of the iterations.

## 1. INTRODUCTION

This paper builds on the framework of [21] and develops a nonlinear solver suitable for use in adaptive methods for quasilinear elliptic problems. The method is developed to stabilize the linearizations of nonlinear diffusion and convection diffusion problems, especially when there may be steep internal or boundary layers in the problem data. The sequence of linear problems encountered by a Newton-like method under these circumstances takes the form of convection diffusion or reaction convection diffusion and the sequence of approximate solutions is subject to spikes, overshoots and spurious oscillations in the convection-dominated regime.

The present approach builds on a regularized version of the pseudo-transient continuation method as in for instance [1, 19, 6, 12, 7, 21] and the references therein, and on each mesh refinement seeks a steady-state solution of the nonlinear evolution problem  $\partial/\partial t(Ru) + g(u(t)) = 0$ , for positive semidefinite regularization term  $R$  in the interest of solving  $g(u) = 0$ . In this analysis, further stabilization is introduced into the time discretization to address the problem of nonphysical oscillations. Time discretizations featuring controllable high-frequency numerical dissipation are well known in the finite element analysis of structures as in for instance [20, 16, 15, 17, 4], and a variation related to these methods referred to here as the  $\sigma$ -split Newmark update is presently introduced. This method makes use of the controllable dissipation of the Newmark update and further controls fluctuations between the iterates by freezing a small fraction the linearization about a point with favorable properties. The  $\sigma$ -split Newmark method is derived,  $q$ -linear local convergence with a predictable rate is established, and the method is demonstrated on three variations of a model problem with steep internal layers.

---

*Date:* July 6, 2018.

*Key words and phrases.* Nonlinear diffusion, convection diffusion, quasilinear equations, pseudo-transient continuation, Newton-like methods, adaptive methods.

The goal of the solver in the adaptive setting is to start on a coarse mesh where the problem data is generally not resolved and produce a sequence of transitional states which may not be accurate solutions to the coarse mesh problems, but do allow *a posteriori* error indicators to detect the layers present in the problem data and refine the mesh to be globally fine enough for stability and locally fine enough to achieve accuracy and efficiency. The sequence of approximate problems can be classified into three phases, the initial phase where the mesh is globally too coarse and quadrature error is high and many features of the problem data remain undetected by the discretization, the pre-asymptotic phase where the mesh is fine enough for stability but the problem data is only partially resolved, and the asymptotic regime where the standard existence, uniqueness and approximation properties hold. In the coarse mesh regime the solver uses as much stabilization as necessary to produce smooth transitional solutions; in the pre-asymptotic regime the solver adaptively reduces the added stabilization increasing both accuracy and the convergence rate as the data is resolved and the approximate problem becomes less rough; in the asymptotic phase the solver limits to a standard Newton method where the initial guess interpolated from the previous refinement is a good approximation to the solution and the iterations converge quadratically.

The requirements of the adaptive method are that a local error indicator is computed on each refinement and a *reasonable* marking strategy in the sense of [22] is employed to determine the next mesh refinement. A modification of the standard adaptive marking strategy is proposed in which the marked set is determined by two parts: one that refines the elements with the largest indicators, and the other that refines a subset of the coarsest elements of the mesh when the residual from the final Newton-like iteration is significant. The method reported here improves on [21] both in terms of efficiency and in terms of the strengths of near-singularities it is able to resolve.

The remainder of the paper is organized as follows. Section 3 shows the derivation of the pseudo-transient Newmark and  $\sigma$ -split Newmark methods, (2.8) and (2.9). Section 4 demonstrates local  $q$ -linear convergence of the residual for these methods, both with rate  $q = 1 - 1/\gamma$ , with  $\gamma$  the parameter from the Newmark update. Section 5 describes a basic algorithm to implement the solver in an adaptive method, and Section 6 contains the results of numerical experiments using the described adaptive algorithm and (2.9).

The following notation is used in the remainder of the paper. The function  $g(u)$  refers to a specific problem or problem class and  $g(x)$  is used in the formal discussion of Newton-like methods. The  $n$ th iteration subordinate to the  $k$ th partition  $\mathcal{T}_k$  is denoted  $x_k^n$ , while  $x^n$  is the  $n$ th iteration on a fixed partition and  $x_k$  is the final iteration on the  $k$ th mesh, taken as the approximate solution on  $\mathcal{T}_k$ . In defining the weak and bilinear forms in the next section  $(u(x), v(x)) = \int_{\Omega} u(x)v(x) dx$ , and similarly for vector-valued functions.

## 2. TARGET PROBLEM CLASS

The nonlinear solver is developed to approximate solutions to the nonlinear problem  $g(u) = 0$ , for polygonal domain  $\Omega$  and  $g : X \rightarrow Y^*$  with  $g'(u) \in \mathcal{L}(X, Y^*)$  for real Banach spaces  $X$  and  $Y$ , particularly where  $g(u)$  takes the form of a quasilinear convection diffusion or diffusion problem in divergence form

$$g(u) := -\operatorname{div}(\kappa(u)\nabla u) + b(u) \cdot \nabla u - f(x) = 0 \text{ in } \Omega \subset \mathbb{R}^2, \quad u = 0 \text{ on } \partial\Omega. \quad (2.1)$$

or

$$g(u) := -\operatorname{div}(\kappa(u)\nabla u) - f(x) = 0 \text{ in } \Omega \subset \mathbb{R}^2, \quad u = 0 \text{ on } \partial\Omega. \quad (2.2)$$

Multiplication by a test function and integration by parts over the divergence term yields the weak form of (2.1)

$$B(u, v) = (\kappa(u)\nabla u, \nabla v) + (b(u) \cdot \nabla u, v) \text{ for all } V \in Y^*, \quad (2.3)$$

and linearizing about  $u$  at  $w \in X$  yields the bilinear form induced by  $g'(u)$

$$B'(u; w, v) = (\kappa(u)\nabla w, \nabla v) + (\kappa'(u)w\nabla u, \nabla v) + (b(u) \cdot \nabla w, v) + (b'(u)w\nabla u, v). \quad (2.4)$$

For  $f \in L_2(\Omega) \cap L_\infty(\Omega)$  and  $\kappa(u)$  bounded away from zero with  $\kappa(s)$ ,  $\kappa'(s)$  and  $\kappa''(s)$  bounded as in [3], then (2.2) has a unique solution  $u \in W^{1,p}(\Omega)$ , with  $2 < p < \infty$ . Both (2.1) and (2.2) fit into the context of [24] with the assumption that  $\kappa(u)$  is bounded and  $g'(u) : H_0^1(\Omega) \rightarrow H^{-1}(\Omega)$  is an isomorphism, in which case the solution  $u$  is an isolated solution.

The discretized equation is, find  $u_h \in X_h$  such that  $B(u_h, v) = 0$  for all  $v \in Y_h$  where  $X_h \subset X$  and  $Y_h \subset Y$  are discrete finite element spaces with respect to triangulation  $\mathcal{T}_h$ , where the family of triangulations  $\{\mathcal{T}_h\}_{0 < h < 1}$  is regular and quasi-uniform in the sense of [5]. Existence, uniqueness and approximation properties of the discrete problems induced by (2.1) and (2.2) are found in [3], [24] and the references therein, assuming the mesh is fine enough. The problem is then to start on a coarse mesh; one that is not fine enough in terms of data resolution, stability or approximation properties, and build one that is. The goal of the solver is to navigate from a coarse to a sufficiently fine mesh where the approximation properties do hold, and to do so by building an both an efficient adaptive mesh and a reasonable initial guess to start the Newton-like iterations on each refinement along the way.

The linearization of (2.2) has the form of a convection-diffusion equation, and (2.4) the linearized form of (2.1) has the form of a convection reaction diffusion equation  $\alpha(w, v) + \beta(w, v) + \gamma(w, v)$  with

$$\alpha(w, v) := (\kappa(u)\nabla w, \nabla v) \quad (2.5)$$

$$\beta(w, v) := (\kappa'(u)w\nabla u, \nabla v) + (b(u) \cdot \nabla w, v) \quad (2.6)$$

$$\gamma(w, v) = (b'(u)w\nabla u, v). \quad (2.7)$$

Using a standard Newton method to solve the nonlinear (2.3) with (2.4) generally does not work when there are steep layers present in the problem data, and coarse mesh approximations of the problem are observed to be indefinite and ill-conditioned, consistent with the observations in [14]. Many of the problems encountered including formation of spurious spikes, overshoots and instability are symptomatic of the corresponding linear convection-dominated problems [9]. Techniques from the finite element analysis of structures [20, 16, 15, 17, 4] use numerical integrators featuring high-frequency dissipation to capture lower frequency modes of the solution. The approach here investigates the use of the Newmark update and a stabilized variation of it as the time-integrators of a pseudo-transient continuation-like method as in [8, 1, 6, 19, 21].

The next section develops the two following methods to improve the convergence of the coarse-mesh iterates in the sequence of linearized problems. A positive semidefinite  $R$  is used to target specific degrees of freedom for regularization, and the role of  $R$  may be seen either from the homotopy perspective as a modification of the path between an initial  $u^0$  and  $u^*$  that solves  $g(u^*) = 0$ , or from the regularization viewpoint as a penalty against certain characteristics of the iterates. In what follows,  $R$  is chosen based on the Laplacian to penalize against high curvature, with degrees of freedom selected for regularization based on an *a posteriori* error indicator. The selective approach to

regularization distinguishes the method presented here and in [21] from the method of pseudo-transient continuation found elsewhere in the literature. For  $\gamma \geq 1$ , a Newmark update generalizing a backward Euler discretization yields the iteration

$$(\alpha_n R + \gamma g'(x^n)) w^n = -g(x^n), \quad x^{n+1} = x^n + w^n. \quad (2.8)$$

The proposed  $\sigma$ -split Newmark update yields the iteration,

$$(\alpha_n R + \gamma \{(1 - \sigma)g'(\bar{x}) + \sigma g'(x^n)\}) w^n = -g(x^n), \quad x^{n+1} = x^n + w^n, \quad (2.9)$$

and analysis and demonstration of this last method is the focus of the remainder of the paper. Local convergence of (2.8) is discussed along the way and local convergence of (2.9) is established by a perturbation of that result. The method based on (2.9) was also tested and found effective on a shifted  $p$ -Laplacian

$$g(u) = -\operatorname{div} \{ (\varepsilon + |\nabla u - a|)^{p-2} (\nabla u - a) \} - f, \quad (2.10)$$

and the related

$$g(u) = -\operatorname{div} \left\{ \left( b + \frac{1}{\varepsilon + |\nabla u - a|^2} \right) (\nabla u - a) \right\} - f, \quad (2.11)$$

problems similar to those investigated in [2] and [13], which reside outside the current target class but still benefit from the regularization techniques described here when starting the adaptive method from a coarse mesh, especially if the coefficient  $\varepsilon \ll 1$ .

### 3. CONTINUATION METHODS

The homotopy or pseudo-transient continuation method of stabilizing the Newton-like iterations for finding the solution  $x^*$  of  $g(x) = 0$  is developed by discretizing the ODE

$$\frac{\partial(Rx)}{\partial t} + g(x(t)) = 0, \quad x(0) = x^0, \quad (3.1)$$

with a positive semidefinite linear functional  $R$ . In much of the literature,  $R$  is taken to be the identity or a scaled version thereof [8, 19], and the references therein; however, the theory is developed for positive-definite functionals other than the identity [1] where it is referred to as the ‘‘s’’ method, and positive semi-definite functionals in [6, 21]. This idea can be generalized to discretizing the ODE based on the normal equations formulation of (3.1)

$$\frac{\partial}{\partial t} R^* R x + g'(x(t))^* g(x(t)) = 0, \quad x(0) = x^0, \quad (3.2)$$

with adjoint  $R^*$  the formal adjoint of  $R$  and  $g'(x)^*$  the formal adjoint of  $g'(x)$ . As shown in [21], the discretization of (3.2) corresponds to the method of Tikhonov regularization.

As discussed in for instance [1, 8, 19, 6], letting  $x^n$  a finite dimensional approximation to  $x(t_n)$  with  $\Delta t_n = t_{n+1} - t_n$ , the standard method is to discretize (3.1) by a backward-Euler approximation to  $\partial/\partial t(Rx)$  and a linearization of  $g'(x^{n+1})$  about  $x^n$ . In the case of discretizing (3.2),  $g'(x^{n+1})^*$  is approximated by  $g'(x^n)^*$ . To increase stability of the linearized system, other discretizations of (3.2) are presently considered. The backward-Euler time discretization is replaced by the more general Newmark update, and the linearization of  $g(x^{n+1})$  is split about two distinct points, one the latest iterate and the other yielding a Jacobian with favorable properties.

By the original method of backward-Euler discretization and first order Taylor expansion about the previous iterate, the resulting Newton-like iteration is given by

$$\left( \frac{1}{\Delta t_n} R + g'(x^n) \right) w^n = -g(x^n). \quad x^{n+1} = x^n + w^n. \quad (3.3)$$

This method increases the stability of the linear system for positive definite and possibly semi-definite  $g'(u)$ , but runs the risk of shifting the spectrum of the approximate Jacobian towards zero, making the condition significantly worse in the case that  $g'(u)$  is indefinite, resulting in large fluctuations between the iterates. To cope with this situation which occurs in the coarse mesh approximation of quasilinear problems, the iteration based on the normal equations enforcing the shift of the spectrum away from zero is given by

$$\left( \frac{1}{\Delta t_n} R^* R + g'(x^n)^* g'(x^n) \right) w^n = -g'(x^n)g(x^n), \quad x^{n+1} = x^n + w^n. \quad (3.4)$$

This method is introduced in [21], where it is shown that (3.4) is also found by minimizing the Tikhonov functional  $G_\alpha(w)$  for

$$G_\alpha(w) = \|g'(x^n)w + g(x^n)\|_0^2 + \alpha_n \|Rw\|_0^2, \quad (3.5)$$

with  $\alpha_n = 1/\Delta t_n$  and  $\|\cdot\|_0$  the  $L_2$  norm. As in [10], the necessary and sufficient condition for the minimizer  $w$  is  $G'_\alpha(w) = 0$ , yielding (3.4). While successfully increasing the stability of the system, this method remains unsatisfactory due to the increased complexity of solving the system based on the normal equations. A new method is now introduced which adds stability while preserving the sparsity of the system. The cost, as shown in Section 4 is trading the asymptotically quadratic convergence of (3.3) for  $q$ -linear convergence. The proposed algorithm of Section 5 updates the parameter  $\gamma$  of the Newmark method on each mesh refinement as the sequence of linear systems stabilize until the method reduces to the original (3.3) for which quadratic convergence of the error is observed.

**3.1. Time discretization by the Newmark method.** The Newmark method [20] discretizes the time derivative  $\dot{u} = \partial u / \partial t$  by

$$x^{n+1} - x^n = \Delta t_n \{ (1 - \gamma)\dot{x}^n + \gamma\dot{x}^{n+1} \}. \quad (3.6)$$

For  $\gamma = 1$ , (3.6) reduces to the backward Euler discretization described above. As discussed in [20] this time integrator is second order accurate for  $\gamma = 1/2$  and introduces nonphysical damping of the high frequency modes for  $\gamma > 1/2$  proportional to  $\gamma - 1/2$ . More sophisticated  $\alpha$  and  $\theta$  collocation methods as in [16, 15, 4] incorporating the Newmark update are designed to further control numerical dissipation across targeted frequencies. In the current context the improved resolution and risk of overshoot of these methods designed with two time derivatives in mind is potentially of interest but their implementation in a pseudo-transient continuation setting is beyond the scope of this article. For stabilizing the transitional states of the sequence of coarse-mesh problems, damping of the high-frequency oscillations is the desirable property and more important than a higher order of accuracy.

Solving (3.6) for  $\dot{x}^{n+1}$

$$\dot{x}^{n+1} = \frac{x^{n+1} - x^n}{\gamma \Delta t_n} - \frac{(1 - \gamma)}{\gamma} \dot{x}^n. \quad (3.7)$$

Applying  $R\dot{x}^n = -g(x^n)$  and  $R\dot{x}^{n+1} = -g(x^{n+1})$  to (3.7)

$$\frac{1}{\gamma \Delta t} R(x^{n+1} - x^n) = \frac{(1 - \gamma)}{\gamma} g(x^n) - g(x^{n+1}) \quad (3.8)$$

Linearizing  $g(x^{n+1})$  about  $x^n$ , obtain the Newton-like iteration

$$\left( \frac{1}{\Delta t_n} R + \gamma g'(x^n) \right) w^n = -g(x^n), \quad x^{n+1} = x^n + w^n. \quad (3.9)$$

For  $\gamma > 1$  this method comparatively scales down the influence of the right-hand side data; the danger is still allowing and arguably increasing the likeliness of shifting an indefinite left-hand side operator spectrum towards zero. By the same approach, the iteration based on the normal equations is given by

$$\left( \frac{1}{\Delta t_n} R^* R + \gamma g'(x^n)^* g'(x^n) \right) w^n = -g'(x^n)g(x^n), \quad x^{n+1} = x^n + w^n. \quad (3.10)$$

Due to the overdamping effect of setting the parameter  $\gamma > 1/2$ , the iterations (3.9) and (3.10) stabilize the solution in some situations where the iteration (3.4) is infeasible. Similarly to (3.4)- (3.5), the solution  $w^n$  of (3.10) is seen to be the minimizer of the generalized Tikhonov-type functional

$$G_\alpha(w) = \|g'(x^n)w + \frac{1}{\gamma}g(x^n)\|_0^2 + \frac{\alpha_n}{\gamma}\|Rw\|_0^2, \quad (3.11)$$

with  $\alpha_n = 1/\Delta t_n$ .

**3.2. The Newmark update with  $\sigma$ -splitting.** The Newark discretization of the previous section successfully introduces high-frequency dissipation increasing the stability of the linearized system but suffers the same drawback as the backward-Euler discretization requiring the formulation based on the normal equations in the case of a possibly indefinite Jacobian to control highly unstable sequences of iterates. The  $\sigma$ -split Newmark update runs without the use of the normal equations: effectively freezing a small fraction of the Jacobian at a point with favorable properties dramatically reduces the fluctuations between iterates in the coarse mesh regime where boundary and internal layers are only partially resolved.

As seen in the derivation, the method can be thought of as splitting the linearization of  $g'(x^{n+1})$  about two points, or more precisely as approximating  $\Delta x^n$  by a combination of  $\dot{x}^{n+1}$ ,  $\dot{x}^n$  and  $\Delta \dot{x}^n$  where the first is used in the backward-Euler update, the first two are used in the Newmark update with the second adding control of numerical dissipation; and the third is introduced here to reduce the fluctuation of the Jacobian outside the domain of convergence of the Newton-like iterations. The method is derived as follows, and is demonstrated in Section 6 to work without the use of the normal equations in situations where other methods including those involving the normal equations are seen to fail.

Starting with the Newmark update (3.6), each of the time derivative terms on the right is split into two parts

$$\begin{aligned} x^{n+1} - x^n &= \Delta t_n \{ (1 - \gamma)\sigma \dot{x}^n + (1 - \gamma)(1 - \sigma)\dot{x}^n + \gamma\sigma \dot{x}^{n+1} + \gamma(1 - \sigma)\dot{x}^{n+1} \} \\ &= \Delta t_n \{ (1 - \gamma\sigma)\dot{x}^n + \gamma(1 - \sigma)(\dot{x}^{n+1} - \dot{x}^n) + \gamma\sigma \dot{x}^{n+1} \}. \end{aligned} \quad (3.12)$$

Solving for  $\dot{x}^{n+1}$  and applying the relation  $R\dot{x}^{n+1} = -g(x^{n+1})$

$$\frac{1}{\gamma\sigma\Delta t_n} R\Delta x^n - \frac{(1 - \sigma)}{\sigma} R(\dot{x}^{n+1} - \dot{x}^n) = \frac{(1 - \gamma\sigma)}{\gamma\sigma} \dot{x}^n - g(x^{n+1}). \quad (3.13)$$

Applying the relation  $R\dot{x}^n = -g(x^n)$  on the left and linearizing  $g(x^{n+1})$  about  $x^n$  on the right

$$\left( \frac{1}{\gamma\sigma\Delta t_n} R + g'(x^n) \right) \Delta x^n + \frac{(1 - \sigma)}{\sigma} (g(x^{n+1}) - g(x^n)) = \frac{-1}{\gamma\sigma} g(x^n). \quad (3.14)$$

Linearizing both  $g(x^{n+1})$  and  $g(x^n)$  about fixed  $\bar{x}$ ,  $g(x^{n+1}) - g(x^n) = g'(\bar{x})\Delta x^n$ . Applying this and multiplying through by  $\gamma\sigma$

$$\left( \frac{1}{\Delta t_n} R + \gamma \{ (1 - \sigma)g'(\bar{x}) + \sigma g'(x^n) \} \right) \Delta x^n = -g(x^n). \quad (3.15)$$

Equation (3.15) is the basis for the  $\sigma$ -split Newmark update. It contains four parameters:  $\alpha_n = 1/\Delta t_n$ ,  $\gamma$ ,  $\sigma$  and  $\bar{x}$ ; while some guidance is provided, a detailed analysis of these parameters will be addressed in subsequent work by the author. In the current results  $\bar{x} = 0$  is used, and  $\bar{x} = x^0$  the solution from the previous refinement interpolated onto the current mesh has also been observed to work;  $\alpha_n$  is chosen as in [21] for the backward-Euler discretization,  $\gamma \geq 1$  may be set adaptively, increased to add stability then decreased to speed convergence; and  $\sigma \geq \sigma_0 \in (0, 1)$  is adaptively sent towards one based on the norm of the latest residual on each Newton-like iteration. It is observed in numerical experiments that  $\sigma$  should be close to one, and as seen in Section 4 this is a necessary for the asymptotic  $q$ -linear convergence at the rate  $q = 1 - 1/\gamma$ , in agreement with the rate found using the Newmark discretization without the  $\sigma$ -splitting.

#### 4. LOCAL CONVERGENCE

Local convergence of the residual is established for both algorithms (2.8) based on the Newmark update, and (2.9) based on the Nemark update with  $\sigma$ -splitting of the Jacobian. The second result follows as a perturbation of the first for  $\sigma$  close to unity, relying on a weaker set of assumptions. Both results require the same Lipschitz condition on the Jacobian. Denote the open ball about  $x$  by  $B(x, \varepsilon) = \{y \mid \|x - y\| < \varepsilon\}$ .

**Assumption 4.1.** *There exist  $\omega_L, \varepsilon > 0$  so that for all  $x, y \in B(x^*, \varepsilon)$*

$$\|g'(x) - g'(y)\| \leq \omega_L \|x - y\| \text{ for all } x, y \in B(x^*, \varepsilon).$$

**Assumption 4.2.** *(c.f. Assumptions 2.2-2.3 of [6], and Assumptions 4.1 and 4.4 of [21]). There exists  $\beta > 0$  so that for positive semidefinite  $R$ ,  $\gamma > 1$ , and for all  $0 < \alpha_n < \alpha_M$ , then for all  $x \in B(x^*, \varepsilon)$ :*

- 1)  $\alpha_n R + \gamma g'(x)$  is invertible.
- 2)  $\|(\alpha_n R + \gamma g'(x))^{-1}\| \leq M_\gamma$ .
- 3)  $\|(\alpha_n R)(\alpha_n R + \gamma g'(x))^{-1}\| \leq \frac{1}{1 + \beta\gamma/\alpha_n}$ .

First,  $q$ -linear convergence of the residual is established with rate  $q = 1 - 1/\gamma$ , for  $\gamma > 1$ . Linear and asymptotic quadratic convergence for the case  $\gamma = 1$  is shown in [21], following from convergence of the error. In the present discussion, convergence of the error is neither used nor shown: it is observed in numerical experiments that the residual decreases at the predicted rate over iterations where  $\|x^{n+1} - x^n\|$  may not yet be decreasing; in contrast, for the case  $\gamma = 1$  the same quantity displays asymptotically quadratic convergence to zero together with the residual. The following proof is a variation of Theorem 2.12 of [7], where Assumptions 4.1 and 4.2 replace the affine contravariant Lipschitz condition used in that version of the Newton-Mysovskikh Theorem for the standard Newton method.

**Theorem 4.3.** *Let  $\alpha_n \leq \|g(x^n)\| \leq \alpha_M$  and let Assumptions 4.1 and 4.2 hold. Define the open set  $\mathcal{S}$  by*

$$\mathcal{S} = \left\{ x \in B(x^*, \varepsilon) \mid \|g(x)\| < \frac{2\beta\gamma}{2 + \beta\gamma^2\omega_L M_\gamma^2} \right\}, \quad (4.1)$$

and suppose  $x^n \in \mathcal{S}$ ; then by iteration (2.8),  $x^{n+1} \in \mathcal{S}$ , and the sequence of residuals converges  $q$ -linearly to zero with asymptotic rate  $q = 1 - 1/\gamma$ .

*Proof.* Let  $\Delta x^n = x^{n+1} - x^n$ . Then iteration (2.8) is given by

$$(\alpha_n R + \gamma g'(x^n)) \Delta x^n = -g(x^n), \quad x^{n+1} = x^n + \Delta x^n. \quad (4.2)$$

By the integral mean-value theorem

$$\begin{aligned} g(x^n + \lambda \Delta x^n) &= g(x^n) + \int_0^\lambda g'(x^n + t \Delta x^n) \Delta x^n dt \\ &= g(x^n) + \lambda g'(x^n) \Delta x^n + \int_0^\lambda (g'(x^n + t \Delta x^n) - g'(x^n)) \Delta x^n dt \\ &= \left(1 - \frac{\lambda}{\gamma}\right) g(x^n) + \left(\frac{\lambda}{\gamma}\right) \alpha_n R (\alpha_n R + \gamma g'(x^n))^{-1} g(x^n) \\ &\quad + \int_0^\lambda (g'(x^n + t \Delta x^n) - g'(x^n)) \Delta x^n dt \end{aligned} \quad (4.3)$$

Applying Assumption 4.2 (3) to the second term and the Lipschitz condition 4.1 to the third term of (4.3)

$$\|g(x^n + \lambda \Delta x^n)\| \leq \left(1 - \frac{\lambda}{\gamma}\right) \|g(x^n)\| + \frac{\lambda}{\gamma(1 + \beta\gamma/\alpha_n)} \|g(x^n)\| + \frac{\lambda^2 \omega_L}{2} \|\Delta x^n\|^2. \quad (4.4)$$

By the iteration (4.2) and Assumption 4.2 (2),  $\|\Delta x\|^2 \leq M_\gamma^2 \|g(x^n)\|^2$ , yielding

$$\|g(x^n + \lambda \Delta x^n)\| \leq \|g(x^n)\| \left( \left(1 - \frac{\lambda}{\gamma}\right) + \frac{\lambda \alpha_n}{\gamma(\alpha_n + \beta\gamma)} + \frac{\omega_L \lambda^2 M_\gamma^2}{2} \|g(x^n)\| \right). \quad (4.5)$$

By the assumption  $\alpha_n \leq \|g(x^n)\|$ , and for  $\lambda \in [0, 1]$

$$\|g(x^n + \lambda \Delta x^n)\| \leq \|g(x^n)\| \left( \left(1 - \frac{\lambda}{\gamma}\right) + \left(\frac{\lambda}{\gamma}\right) \left(\frac{1}{\beta\gamma} + \frac{\gamma \omega_L M_\gamma^2}{2}\right) \|g(x^n)\| \right), \quad (4.6)$$

for each  $\lambda \in [0, 1]$  such that  $x^n + t \Delta x^n \in \mathcal{S}$  for all  $t \in [0, \lambda]$ . By the logic of [7] Theorem 2.12, proceed by contradiction and assume  $x^{n+1} \notin \mathcal{S}$ ; then there is a smallest  $\bar{\lambda} \in (0, 1]$  with  $g(x^n + \bar{\lambda} \Delta x^n) \in \partial \mathcal{S}$ . For that  $\bar{\lambda}$

$$\begin{aligned} \|g(x^n + \bar{\lambda} \Delta x^n)\| &\leq \|g(x^n)\| \left( \left(1 - \frac{\bar{\lambda}}{\gamma}\right) + \left(\frac{\bar{\lambda}}{\gamma}\right) \left(\frac{1}{\beta\gamma} + \frac{\gamma \omega_L M_\gamma^2}{2}\right) \|g(x^n)\| \right) \\ &< \|g(x^n)\|, \end{aligned} \quad (4.7)$$

a contradiction. This shows  $x^{n+1} \in \mathcal{S}$ . To establish the rate of convergence, set  $\lambda = 1$  in 4.6.

$$\|g(x^{n+1})\| \leq \|g(x^n)\| \left( \left(1 - \frac{1}{\gamma}\right) + \left(\frac{1}{\gamma}\right) \left(\frac{1}{\beta\gamma} + \frac{\gamma \omega_L M_\gamma^2}{2}\right) \|g(x^n)\| \right), \quad (4.8)$$

which shows both that  $\|g(x^{n+1})\| < \|g(x^n)\|$  and the asymptotic  $q$ -linear rate of  $q = 1 - 1/\gamma$  as  $\|g(x^n)\| \rightarrow 0$ .  $\square$

**Remark 4.4.** *It is observed that the method converges with the predicted rate when the iterates  $x^n \notin \mathcal{S}$  as given by (4.1). The lapse in the theory appears to be the bound on  $\|x^n\|$  which apparently converges within a smaller set as compared to the residual.*



The asymptotic convergence of the  $\sigma$ -split algorithm is addressed in the next theorem as a perturbation of Theorem 4.3. While it is not necessary in the proof for  $g'(\bar{x})$  to be positive definite,  $\bar{x}$  is ideally chosen so that  $g'(\bar{x})$  improves the condition of the approximate Jacobian, allowing Assumptions 4.5 to hold with better constants than 4.2.

**Assumption 4.5.** (c.f. Assumption 4.2). *There exist  $\beta > 0$  so that for  $0 < \sigma_0 \leq \sigma \leq 1$ , fixed  $\bar{x}$ , positive semidefinite  $R$ ,  $\gamma > 1$ , and for all  $0 < \alpha_n < \alpha_M$ , then for all  $x \in B(x^*, \varepsilon)$ :*

- 1)  $\alpha_n R + \gamma\{(1 - \sigma)g'(\bar{x}) - \sigma g'(x)\}$  is invertible.
- 2)  $\|(\alpha_n R + \gamma\{(1 - \sigma)g'(\bar{x}) - \sigma g'(x)\})^{-1}\| \leq M_{\gamma\sigma}$ .
- 3)  $\|(\alpha_n R)(\alpha_n R + \gamma\{1 - \sigma\}g'(\bar{x}) - \sigma g'(x))^{-1}\| \leq \frac{1}{1 + \beta\gamma/\alpha_n}$ .

As with Assumption 4.2 (3), the third clause agrees with the similar stability bound of [6] Assumption 2.3, with  $\alpha_n$  replaced by  $\alpha_n/\gamma$ . Local convergence of the  $\sigma$ -split Newmark algorithm is established with the same asymptotic rate as in the previous result.

**Theorem 4.6.** *Let  $\alpha_n \leq \|g(x^n)\| \leq \alpha_M$  and let Assumptions 4.1 and 4.5 hold. Define  $\sigma$  by*

$$\sigma = \max \left\{ \sigma_0, 1 - \frac{\|g(x^n)\|}{K_0} \right\}, \quad (4.9)$$

for a given  $K_0$ . Then there exists  $\delta_1 > 0$  such that for  $x^n$  in the open set  $\mathcal{S}$  given by

$$\mathcal{S} = \left\{ x \in B(x^*, \varepsilon) \mid \|g(x)\| < \delta_1 \right\}, \quad (4.10)$$

and  $x^{n+1}$  defined by iteration (2.9), it hold that  $x^{n+1} \in \mathcal{S}$ , and the sequence of residuals converges  $q$ -linearly to zero with asymptotic rate  $q = 1 - 1/\gamma$ .

*Proof.* Let  $\Delta x^n = x^{n+1} - x^n$ . Then iteration (2.9) is given by

$$(\alpha_n R + \gamma\{(1 - \sigma)g'(\bar{x}) + \sigma g'(x^n)\}) \Delta x^n = -g(x^n), \quad x^{n+1} = x^n + \Delta x^n. \quad (4.11)$$

Much of the proof parallels Theorem 4.3, and is summarized here. Starting with the integral mean-value theorem

$$\begin{aligned} g(x^n + \lambda \Delta x^n) &= g(x^n) + \int_0^\lambda g'(x^n + t \Delta x^n) \Delta x^n dt \\ &= g(x^n) + \lambda g'(x^n) \Delta x^n + \int_0^\lambda (g'(x^n + t \Delta x^n) - g'(x^n)) \Delta x^n dt. \end{aligned} \quad (4.12)$$

By iteration (4.11)

$$g'(x^n) \Delta x^n = -\frac{1}{\gamma\sigma} g(x^n) - \frac{1}{\gamma\sigma} \alpha_n R \Delta x^n - \frac{1 - \sigma}{\sigma} g'(\bar{x}) \Delta x^n. \quad (4.13)$$

Bounding the second term on the right of (4.13) by Assumptions 4.5 (3)

$$\begin{aligned} \|\alpha_n R \Delta x^n\| &= \|\alpha_n R (\alpha_n R + \gamma\{(1 - \sigma)g'(\bar{x}) + \sigma g'(x^n)\})^{-1} g(x^n)\| \\ &\leq \frac{1}{1 + \beta\gamma/\alpha_n} \|g(x^n)\|, \end{aligned} \quad (4.14)$$

and by Assumptions 4.1 and 4.5 (2), then applying  $\alpha_n \leq \|g(x^n)\|$  and  $\lambda \in [0, 1]$

$$\begin{aligned} \|g(x^n + \lambda\Delta x^n)\| &\leq \|g(x^n)\| \left\{ \left(1 - \frac{\lambda}{\gamma\sigma}\right) + \frac{(1-\sigma)}{\sigma} \lambda M_{\gamma\sigma} \|g'(\bar{x})\| + \left(\frac{\lambda}{\gamma\sigma}\right) \frac{1}{(1 + \beta\gamma/\alpha_n)} \right. \\ &\quad \left. + \frac{\omega_L \lambda^2 M_{\gamma\sigma}^2}{2} \|g(x^n)\| \right\} \\ &\leq \|g(x^n)\| \left\{ \left(1 - \frac{\lambda}{\gamma\sigma}\right) + \frac{(1-\sigma)}{\sigma} \lambda M_{\gamma\sigma} \|g'(\bar{x})\| \right. \\ &\quad \left. + \left(\frac{\lambda}{\beta\gamma^2\sigma}\right) \|g(x^n)\| + \frac{\lambda\omega_L M_{\gamma\sigma}^2}{2} \|g(x^n)\| \right\}. \end{aligned} \quad (4.15)$$

Applying (4.9) to the quantity  $(1-\sigma)/\sigma$  under the assumption  $1 - \|g(x^n)\|/K_0 \geq \sigma_0$  and expanding in orders of  $\|g(x^n)\|/K_0 < 1$ ,

$$\begin{aligned} \frac{1-\sigma}{\sigma} &= \frac{\|g(x^n)\|}{K_0} \frac{1}{(1 - \|g(x^n)\|/K_0)} = \frac{\|g(x^n)\|}{K_0} \left(1 + \mathcal{O}\left(\frac{\|g(x^n)\|}{K_0}\right)\right) \\ &= \frac{\|g(x^n)\|}{K_0} + \mathcal{O}\left(\frac{\|g(x^n)\|}{K_0}\right)^2. \end{aligned} \quad (4.16)$$

Similarly

$$\frac{1}{\sigma} = 1 + \frac{\|g(x^n)\|}{K_0} + \mathcal{O}\left(\frac{\|g(x^n)\|}{K_0}\right)^2. \quad (4.17)$$

Applying (4.16) and (4.17) to (4.15), then for any fixed  $P > 3$

$$\begin{aligned} \|g(x^n + \lambda\Delta x^n)\| &\leq \|g(x^n)\| \left\{ \left(1 - \frac{\lambda}{\gamma}\right) + \frac{\lambda}{P\gamma} \|g(x^n)\| \right. \\ &\quad \left. \times P \left( \frac{K_0(\beta\gamma^2\omega_L M_{\gamma\sigma}^2/2 + 1) + \beta\gamma(M_{\gamma\sigma}\|g(\bar{x})\| - 1)}{K_0\beta\gamma} \right) + \mathcal{O}(\|g(x^n)\|^2) \right\}. \end{aligned} \quad (4.18)$$

Supposing

$$\|g(x^n)\| \leq \delta_0 := \frac{K_0\beta\gamma/P}{K_0(\beta\gamma^2\omega_L M_{\gamma\sigma}^2/2 + 1) + \beta\gamma(M_{\gamma\sigma}\|g(\bar{x})\| - 1)}, \quad (4.19)$$

it follows that

$$\|g(x^n + \lambda\Delta x^n)\| \leq \|g(x^n)\| \left(1 - \frac{(P-1)\lambda}{P\gamma} + \mathcal{O}(\|g(x^n)\|^2)\right), \quad (4.20)$$

so there exists  $\delta_1 \in (0, \delta_0]$  for which  $\|g(x^n)\| < \delta_1$  implies

$$\|g(x^n + \lambda\Delta x^n)\| < \|g(x^n)\| \left(1 - \frac{(P-2)\lambda}{P\gamma}\right). \quad (4.21)$$

The result follows by the same logic as in the proof of Theorem 4.3.  $\square$

**Remark 4.7.** *It is observed in some problems that the residual  $\|g(x^n)\|$  can decrease at a rate slightly better than the one predicted when  $\sigma \neq 1$ , whereas for  $\sigma = 1$  the rate is as predicted.*

## 5. ALGORITHM

The solver using the  $\sigma$ -Newmark iteration (2.9) may be implemented in an adaptive method according to the following basic algorithm. To exploit both the stability of method (2.9) and the quadratic convergence of (3.3), the parameters  $\gamma$  and  $R$  are updated on each refinement, and the parameters  $\alpha$  and  $\sigma$  are updated on each iteration of the solver. The main steps of the adaptive method are as follows with the exit criteria and parameter updates specified below. An example of a targeted regularization term  $R$  based on *a posteriori* error indicators is given in Section 6.

**Algorithm 5.1** (Basic algorithm for (2.9)). *Start with initial  $x^0$ ,  $\gamma$ ,  $\sigma_0$ . On partition  $\mathcal{T}_k$ ,  $k = 0, 1, 2, \dots$*

- 1) Compute  $R$ ,  $g'(\bar{x})$ .
- 2) Set  $\alpha_0 = \|g(x^0)\|$ .
- 3) While Exit criteria 5.2 are not met on iteration  $n - 1$ :
  - (i) Set  $\sigma$  according to (4.9).
  - (ii) Solve  $(\alpha_n R + \gamma\{\sigma g'(x^n) + (1 - \sigma)g'(\bar{x})\})\Delta x^n = g(x^n)$ .
  - (iii) Update  $x^{n+1} = x^n + \Delta x^n$ .
  - (iv) Update  $\alpha_n$ .
- 4) Update  $\gamma$  for partition  $\mathcal{T}_{k+1}$  according to (5.3)

**Criteria 5.2** (Exit criteria). *Given a user set tolerance  $tol$ , an accepted rate of convergence given by  $q_{acc} = 1 - 1/(M\gamma)$  for some constant  $M$ , e.g.,  $M = 2$ , and a maximum number of iterations either chosen as a constant or based on the predicted rate of convergence, exit the solver on partition  $\mathcal{T}_k$  after calculating iterate  $x^{n+1}$  when one of the conditions holds.*

- 1)  $\|g(x^{n+1})\| \leq tol$ .
- 2) (i)  $\|g(x^{n+1})\| < \|g(x^n)\|$ , AND  
 (ii)  $\|g(x^n)\| < \min\{\|g(x^0)\|, \|g(x_{k-1})\|\}$ , AND  
 (iii)  $\|g(x^{n+1})\|/\|g(x^n)\| < q_{acc}$ , AND  
 (iv)  $\|g(x^{n+1})\|/\|g(x^n)\| > \|g(x^n)\|/\|g(x^{n-1})\|$ .
- 3) Maximum number of iterations exceeded.

In terms of the three phases of the solution process in [21], the final asymptotic regime is characterized by the iterations terminating by Criteria 5.2 (1); iterations in the pre-asymptotic regime terminate with Criteria 5.2 (2); and iterations terminate with a mix of Criteria 5.2 (2) and (3) in the initial phase.

The second exit criterion 5.2 (2) merits some explanation, as it allows the iteration to end once reduction of the residual has slowed indicating the iterate has attained a reasonably stable configuration from which a good prediction about where to refine the mesh may be determined. The first two clauses require the residual is decreasing, and has decreased below the level given by the previous iterate with at least as much decrease from the initial iteration on the current partition and the residual from the previous partition, if it is well defined. The third criterion requires error reduction at or close to the predicted rate, and the last that the error reduction between iterates is slowing down. These four criteria together assure the sequence of transitional states is not getting further in the sense of the solver's residual from a converged solution, and prevent situations such as spikes propagating indefinitely across a sequence of partitions. While spikes, overshoots or other undesirable characteristics may propagate through several refinements, such solutions will eventually not reduce the residual at the specified rate. When the adaptive

mesh is fine enough, these characteristics are observed to smooth out, and otherwise the solver eventually fails and the solution is reset and started again on a finer mesh.

One of the main improvements of (2.9) over the regularized method of [21] is the generation of more stable sequences of transitional iterates through the pre-asymptotic phase. The sequence of pre-asymptotic approximate problems only partially resolve the data and as internal layers are uncovered over several refinements the condition of these problems is generally bad. The high-frequency dissipation of the Newmark strategy combined with the stabilization of the iterates produced by the  $\sigma$ -splitting allow sequences of solutions to propagate through this regime and the mesh to be marked via error indicators leading to the accurate solution of the problem on otherwise coarser meshes than possible if starting the solver on a mesh that is uniformly fine enough to resolve the data. In order to make use of the stability and dissipation properties of larger values of  $\gamma = \gamma_k$  as well as the asymptotically quadratic convergence if the iterations are stable for  $\gamma = 1$ , the following minimal guidelines are presented, based on the termination criteria above.

**Update 5.3** (Newmark parameter  $\gamma$ ). . *Generally, if the predicted error reduction rate is achieved,  $\gamma$  should be reduced; and if the iteration fails, more stability is needed and  $\gamma$  should be increased.*

- (1) *Exit criterion 5.2 (1): If  $\gamma_k > 1$  set  $1 \leq \gamma_{k+1} < \gamma_k$ .*
- (2) *Exit criterion 5.2 (2): If  $\|g(x^{n+1})\|/\|g(x^n)\|$  is within tolerance of  $1 - 1/\gamma_k$ , set  $1 \leq \gamma_{k+1} < \gamma_k$ .*
- (3) *Exit criterion 5.2 (3): Set  $\gamma_{k+1} > \gamma_k$ .*

In practice, the residual tends to get reduced below tolerance once  $\gamma = 1$ . In the results of Section 6,  $\gamma$  is decreased by two when the target rate is achieved, decreased by one if a stable rate lower than predicted is achieved, increased by one when the solver fails, and by two if the maximum number of iteration is exceeded while the iterations are converging below the acceptable rate. An initial  $\gamma_0$  should be chosen large enough to see error reduction on the initial mesh, and not significantly larger.

The regularization parameter  $\alpha_n = 1/\Delta t_n$  is updated by the method described in [21], repeated here for convenience. In accordance with the convergence Theorems 4.3 and 4.6, this strategy assures  $\alpha_n \leq \|g(x^n)\|$  so long as the residual is decreasing.

**Update 5.4** (Tikhonov regularization parameter  $\alpha$ ). *Set  $\beta_0 = 1$ . For  $n \geq 1$ ,*

$$\alpha_n = \beta_n \|g(x^n)\|, \text{ with } \beta_n = \frac{\|g(x^n)\|}{\|g(x^{n-1})\|}.$$

*To reduce rapid fluctuation of  $\beta_n$ , correct to ensure  $\beta_{n-1}/2 \leq \beta_n \leq 1$  in the case that  $\|g(x^n)\| < \|g(x^{n-1})\|$  and  $\beta_n \leq 2\beta_{n-1}$  if  $\|g(x^n)\| > \|g(x^{n-1})\|$ .*

**5.1. Marking strategy.** An *a posteriori* error indicator  $\eta_T$ ,  $T \in \mathcal{T}_k$  is assumed available to determine adaptive mesh refinement and as one option for determining a targeted regularization term  $R$ . In the current results, standard local residual-based element indicators as in for instance [23, 11] are used for both of these purposes, further described in Section 6. Other approaches to solver- and problem-specific regularization and mesh refinement are currently under investigation by the author.

The goal of the marking strategy is to build a mesh that is globally fine enough for stability, and locally as fine as necessary to achieve the desired accuracy. To improve the efficiency of the method by increasing the stability of the transitional states in the pre-asymptotic phase, the following marking strategy is presented. Based on exit criteria 5.2 there are three possible outcomes of the nonlinear solve on refinement  $\mathcal{T}_k$ .

- 1) Exit criterium 5.2 (1): The iterate  $x_k$  has residual  $\|g(x_k)\| \leq tol$ .
- 2) Exit criterium 5.2 (2): The iterate  $x_k$  has residual  $\|g(x_k)\| > tol$ .
- 3) Exit criterium 5.2 (3): The solver failed and  $x_k$  is reset to zero.

In the case of 5.2 (3), the coarsest set of elements in the mesh are refined. Failure of the solver reflects a *global* rather than a *local* problem. Starting on a sufficiently coarse mesh, several resets are expected.

In the case of 5.2 (2) the error indicators  $\eta_T$  are computed, and the mesh is refined both according to the elements with the largest local indicators, and according to the coarsest elements with the largest local indicators. This strategy allows local refinement to take place in order to capture boundary and internal layers to attain eventual accuracy and efficiency while also building the adaptive mesh fine enough to achieve stability. In meshes that are too coarse to resolve the data and where the approximate problem may have coefficients based on highly inaccurate approximate solutions, nonphysical overshoots often develop in the iterates; while the error indicators in the vicinity of these spikes may be high, refining primarily in these regions exacerbates the problem. As in case 5.2 (3), a large residual  $\|g(u_k)\|$  predicts a global issue with the mesh. However, valuable information about the near-singularities in the problem data can be predicted from the non-converged iterates, so some local refinement can build a more efficient mesh.

In the case of 5.2 (1) the error indicators are computed, and the mesh is refined with respect to the largest local indicators. Any *reasonable* marking procedure [22] for cases 5.2 (1)-(2) may be applied; in particular for case 5.2 (2) both the element with the largest local indicator must be marked, as well as the coarsest element with the largest local indicator. In the current results, the Dörfler marking strategy is used with parameter  $\theta$ , which for case 5.2 (2) is split into  $\theta = \theta_C + \theta_F$  and the marked sets  $\mathcal{M}_F \subset \mathcal{T}_k$  and  $\mathcal{M}_C \subset \mathcal{T}_k$  are chosen by sets of least cardinality with

$$\sum_{T \in \mathcal{M}_F} \eta_T^2 \geq \theta_F \sum_{T \in \mathcal{T}_k} \eta_T^2, \quad \sum_{T \in \mathcal{M}_C} \eta_T^2 \geq \theta_C \sum_{T \in \mathcal{T}_k} \eta_T^2, \quad (5.1)$$

and the marked set  $\mathcal{M} = \mathcal{M}_F \cup \mathcal{M}_C$ . In case 5.2 (1),  $\theta_F = \theta$  and  $\theta_C = 0$ . An heuristic choice of  $\Phi(\theta) = \theta_C$  is given in Section 6.

## 6. NUMERICAL EXAMPLES

The nonlinear solver (2.9) is demonstrated on three problems with different structure in their internal layers. In Example 6.2 results are reported for a model nonlinear convection-diffusion problem of the form  $g(u) = 0$  which has a smooth sinusoidal solution. The first variation on the model problem, Example 6.3 demonstrates the algorithm on the same differential operator with the problem data chosen to generate a higher frequency solution. Example 6.4 shows the results for a related nonlinear diffusion problem with two concentric internal layers. The adaptive finite element method is implemented using the finite element library FETK [18] and a direct solver is used on each linear system. Both trial and test spaces are taken as the linear finite element space  $V_k$  consisting of Lagrange finite elements  $\mathbb{P}_1$  over partition  $\mathcal{T}_k$  that satisfy the homogeneous Dirichlet boundary conditions.

The local *a posteriori* residual-based indicator, and corresponding jump-based indicator for element  $T \in \mathcal{T}_k$  with  $h_T$  the element diameter are given by

$$\zeta_T^2(v) = h_T \|J_T(v)\|_{L_2(\partial T)}^2 \quad (6.1)$$

$$\eta_T^2(v) = \eta_{\mathcal{T}_k}^2(v, T) := h_T^2 \|g(v)\|_{L_2(T)}^2 + \zeta_T^2(v), \quad (6.2)$$

$J_T(v) := \llbracket \kappa(v) \nabla v \cdot n \rrbracket_{\partial T}$ , with jump  $\llbracket \phi \rrbracket_{\partial T} := \lim_{t \rightarrow 0} \phi(x + tn) - \phi(x - tn)$ , where  $n$  is the appropriate outward normal defined on  $\partial T$ . The error estimator on partition  $\mathcal{T}_k$  is given by the  $l_2$  sum of indicators  $\eta_{\mathcal{T}_k}^2 = \sum_{T \in \mathcal{T}_k} \eta_T^2$ , and similarly for  $\zeta_{\mathcal{T}_k}$ .

On each mesh partition  $\mathcal{T}_k$  the penalty matrix  $R = R_k$  is a localized version of the Laplacian stiffness matrix denoted  $R^{global}$ . Let  $D^{loc}$  a diagonal matrix of zeros and ones,  $v_j$  a vertex of subordinate to partition  $\mathcal{T}_k$ , and  $\zeta_T = \zeta_T(u^0)$ . Then set

$$\tilde{\psi}_k = \sqrt{\text{median}_{T \in \mathcal{T}_k}(\zeta_T)}, \quad \text{and } \psi_k = \begin{cases} \tilde{\psi}_k^{1/2}, & \text{if } \tilde{\psi}_k > 1, \\ \tilde{\psi}, & \text{otherwise,} \end{cases}$$

and

$$D_{jj}^{loc} = \begin{cases} 0, & \text{if } \zeta_T \leq \psi_k \text{ for each element that contains } v_j \text{ as a vertex,} \\ 1, & \text{otherwise.} \end{cases} \quad (6.3)$$

Then set  $R_k = D^{loc} R^{global} D^{loc}$ .

**Remark 6.1.** In [21] a similar strategy is applied using the full indicator  $\eta_T$  as opposed to the  $\zeta_T$  the jump-term alone. This modification was made to select degrees of freedom to penalize against curvature that show the greatest disparity in curvature as predicted by the jump term. The role of the selection function  $\psi$  is to regularize based on spikes in the indicators leaving the stable dof unregularized to speed convergence. As the algorithm approaches the asymptotic phase the selection process  $\psi$  becomes unimportant because  $\alpha_n \leq \|g(x^n)\| \rightarrow 0$ .

In these results the Dörfler parameters  $\theta_C + \theta_F = \theta$  with  $\theta = 0.6$ , for refinement  $k + 1$  is determined according to

$$\theta_C = \Phi(\theta) = \theta \left( \frac{1}{2} + \frac{1}{\pi} \arctan(\|g(u_k)\|/100 - \pi/2) \right). \quad (6.4)$$

This form of  $\Phi(\theta)$  is chosen to ensure a significant fraction of the marked set is from the coarse mesh until the residual drops below a given threshold. In the following examples the parameter for setting  $\sigma$  by formula (4.9) are  $\sigma \geq \sigma_0 = 0.9$  and  $K_0 = 2000$ . The tolerance for the nonlinear solver is  $tol = 10^{-7}$ .

**Example 6.2** (Nonlinear convection diffusion). Consider the nonlinear convection diffusion equation on  $\Omega = [0, 1] \times [0, 1]$ ,

$$g(u) := -\text{div}(\kappa(u) \nabla u) + b(u) \cdot \nabla u - f(x, y) = 0 \in \Omega, \quad u = 0 \text{ on } \partial\Omega. \quad (6.5)$$

The nonlinear diffusion coefficient is given by

$$\kappa(s) = k + \frac{1}{((\epsilon + (s - a)^2))}, \quad \text{with } a = 0.5, \text{ and } k = 1. \quad (6.6)$$

The nonlinear convection term is given by

$$b(s) = ((s - a), (s - a)^2)^T, \quad (6.7)$$

and load  $f(x, y)$  chosen so the exact solution  $u(x, y) = \sin \pi x \sin \pi y$ .

Existence of isolated solutions of (6.5)-(6.7) is discussed in Section 2 following [24], assuming the mesh is fine enough. This problem without the convection term is investigated in [21], with  $\varepsilon = 10^{-3}$ ; the method here has been observed to run to convergence from a mesh of 144 elements with  $\varepsilon = 8 \times 10^{-5}$ , and was tested on smaller values of  $\varepsilon$ ; however the size of the resulting meshes were unwieldy for test-problems. Data on the iterations in the pre-asymptotic regime is presented here for  $\varepsilon = 2 \times 10^{-4}$  starting on a mesh of 144 elements, with  $\gamma = 10$  and run until convergence; and for  $\varepsilon = 6 \times 10^{-4}$  run well into the asymptotic regime. In the pre-asymptotic case the iterations are terminated by exit criteria 5.2 (2), and in most cases where  $\gamma$  is reduced it is reduced by one as the iterates are converging at a rate better than predicted. On refinements 20, 21 and 28 the final iterates are converging at or close to the predicted rate  $q = 1 - 1/\gamma$ , and  $\gamma$  is reduced by two. Table 1 summarizes the computation showing the final residual  $\|g(u_k)\|$ , the final ratio  $\|g(u^{n+1})\|/\|g(u^n)\|$ , the final value of  $\sigma$  and the Newmark parameter  $\gamma$ .

Of note is the range in magnitude over which the residual shows monotonic and in fact predictable and stable decrease. On level 7, the first partition in the pre-asymptotic regime, the residual is reduced from an initial  $\|g(u^0)\| = 1564.5$ , to a final  $\|g(u^9)\| = 871.1$ , using  $\sigma = \sigma_0 = 0.9$  on each iteration. The final pre-asymptotic solve reduces the residual from  $\|g(u^0)\| = 4.6$ , to  $\|g(u^4)\| = 1.4$ , with  $\sigma > 0.996$  and approaching one. On level 7 for which  $\gamma = 20$ , the ratio of consecutive residuals stabilizes well below the predicted rate whereas in level 28, the predicted rate  $1 - 1/\gamma$ , with  $\gamma = 3$ , is achieved. Most of the pre-asymptotic solves exit after 3-4 iterations, and two of the solves take up to 9; generally the higher number of iterations corresponds to qualitative changes in the iterates such as the smoothing of spikes and changes in curvature illustrated for instance in the solution snapshots in Figure 3.

Setting  $\varepsilon = 6 \times 10^{-4}$ , a milder version of the same problem illustrates the reduction in the error from the initial through the pre-asymptotic and into the asymptotic regime which in this case runs from levels 2 to 19. The computation is again started on a mesh of 144 elements with  $\gamma = 10$ . Figure 1 shows a logarithmic plot of the  $H^1$  error against the number of elements  $n$  compared with  $n^{-1/2}$ . The increase in the error and estimator for the first several refinements demonstrates the sequences of partitions capturing the internal layer in the problem data; the first mesh in Figure 1 and corresponding solution in Figure 2 show the representative behavior as this occurs; the flatness at the top of the solution is characteristic of the penalization by the Laplacian against curvature. The second mesh of Figure 1 and corresponding solution of Figure 3 show the approximate solution as the internal layer is better resolved, but spurious nonphysical oscillations are still apparent in the iterates; this mesh further illustrates the characteristic of the layer getting captured by the mesh, but not uniformly. The final mesh of Figure 2 and solution of Figure 3 are representative of the end of the asymptotic regime where the mesh now refines to capture the layer with increasingly uniform resolution and the solution has the overall correct shape. Within several iteration the approximate solution looks like a sinusoid and the Newton-like iterations solve to tolerance. It is noted in Figure 1 that the error is stable for the first several iterations in the asymptotic regime while the error estimator reduces at a steady rate; presumably this occurs due to the presence of another nearby solution of (6.5). A similar phenomenon is observed in [13] in examples where the monotonicity is violated and the solution is known only to be locally unique.

**Example 6.3** (Higher frequency solution). *A variation of Example 6.2 is considered where the problem is given by (6.5)-(6.7), with  $\varepsilon = 6 \times 10^{-4}$ , and the load  $f(x, y)$  chosen so the exact solution is  $u = \sin(2\pi x) \sin(2\pi y)$ .*

Level	$\ g(u_k)\ $	$\frac{\ g(u^{n+1})\ }{\ g(u^n)\ }$	$\sigma_k$	$\gamma_k$
⋮				
7	871.1	0.90	0.9	20
8	786.8	0.93	⋮	20
9	647.0	0.94		19
10	535.0	0.94		18
11	445.0	0.93		17
12	368.3	0.93		16
13	305.2	0.93		15
14	275.7	0.92	⋮	14
15	231.3	0.92	0.9	13
16	165.6	0.91	0.909	12
17	137.7	0.90	0.924	11
18	122.1	0.90	0.932	10
19	98.4	0.88	0.944	9
20	75.0	0.87	0.957	8
21	59.0	0.83	0.964	6
22	44.0	0.75	0.971	4
23	23.5	0.67	0.983	3
24	14.1	0.90	0.990	3
25	7.7	0.71	0.995	3
26	4.8	0.68	0.997	3
27	2.5	0.67	0.998	3
28	1.4	0.67	0.999	3
29	1.9e-08	8.1e-03	1	1
⋮				

TABLE 1. Summary of data from the pre-asymptotic regime for Example 6.2 with  $\varepsilon = 2 \times 10^{-4}$ .

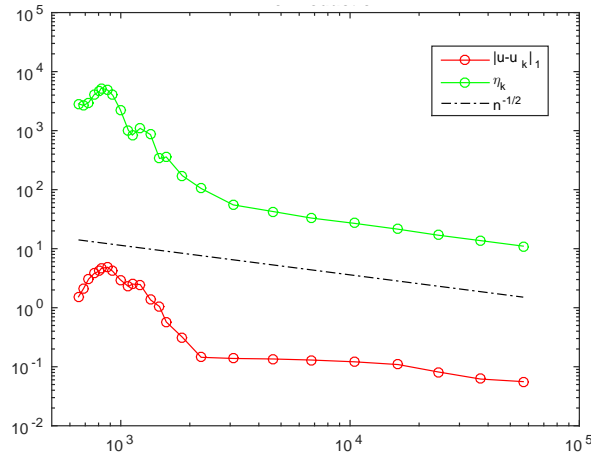


FIGURE 1.  $H^1$  error (below) and error estimator (above) against  $n^{-1/2}$  where  $n$  is the number of elements, for Example 6.2 with  $\varepsilon = 6 \times 10^{-4}$ .



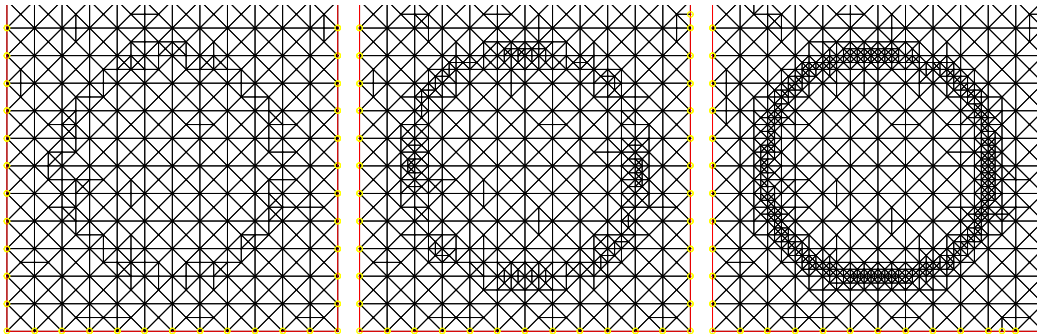


FIGURE 2. Mesh for Example 6.2 with  $\varepsilon = 6 \times 10^{-4}$  after 6, 12 and 18 adaptive refinements from an initial mesh with 144 elements.

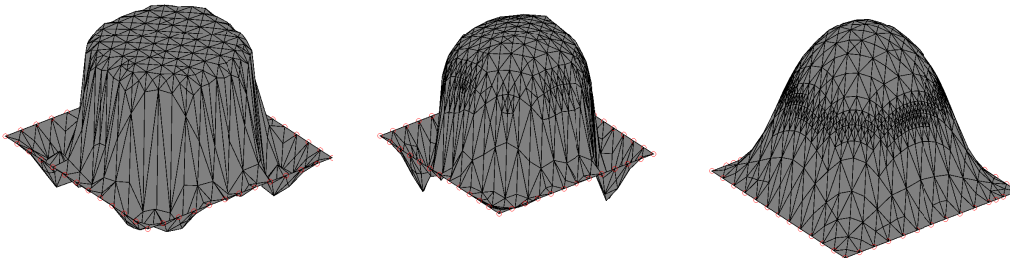


FIGURE 3. Snapshots of the finite element solution for Example 6.2 with  $\varepsilon = 6 \times 10^{-4}$  in the pre-asymptotic regime after 6, 12 and 18 adaptive refinements from an initial mesh with 144 elements.

Local uniqueness of the solution follows from [24], assuming the mesh is fine enough. This example features two disjoint internal layers in the problem data, one about each positive peak of the sinusoid. The decrease in the residual in the pre-asymptotic regime closely resembles the data presented in Table 1 for Example 6.2. For  $\varepsilon = 6 \times 10^{-4}$ , the pre-asymptotic phase for this example runs from levels 4 to 26 as opposed to 2 to 19 for the sinusoid with a single peak; and the mesh in Example 6.3 maintains a maximum meshsize of  $h = 8.6 \times 10^{-4}$ , whereas in Example 6.2, the asymptotic meshsize is  $h = 3.5 \times 10^{-3}$ , indicating two more coarse mesh refinements are taken to stabilize the iterates. This is to be expected due to the decrease in width of the corresponding layers. Mesh partitions and snapshots of the pre-asymptotic iterates are shown in Figures 4 and 5. The characteristic flatness of the peaks due to the penalization against curvature is again observed in the first two iterates, shown respectively at the 8th and 12th adaptive refinements, and resolved by the 22nd, as seen on the right. The iterates show a qualitatively different behavior than the nonphysical oscillations of Example 6.2; here, the downward peaks which should have the same magnitude as the upward peaks start as shallow and extend to their full depth as the refinements progress.

**Example 6.4** (Concentric layers). Consider the nonlinear diffusion problem on  $\Omega = [0, 1] \times [0, 1]$  with two concentric internal layers.

$$g(u) := -\operatorname{div}(\kappa(u)\nabla u) - f(x, y) = 0 \quad \in \Omega, \quad u = 0 \text{ on } \partial\Omega, \quad (6.8)$$

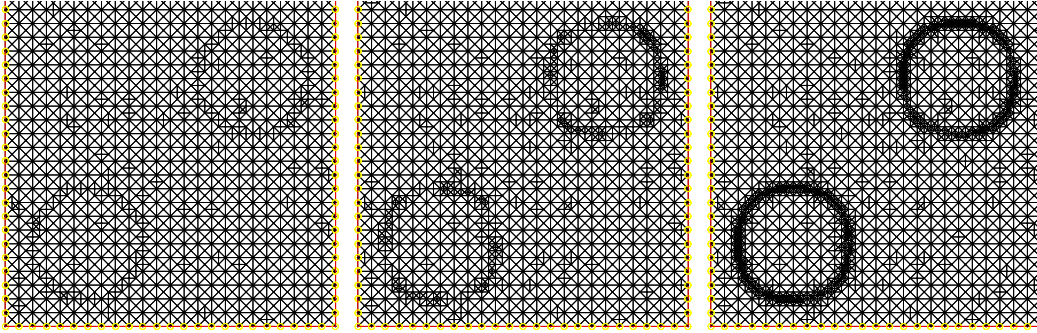


FIGURE 4. Mesh for Example 6.3 with  $\varepsilon = 6 \times 10^{-4}$  after 8, 14 and 22 adaptive refinements from an initial mesh with 144 elements.

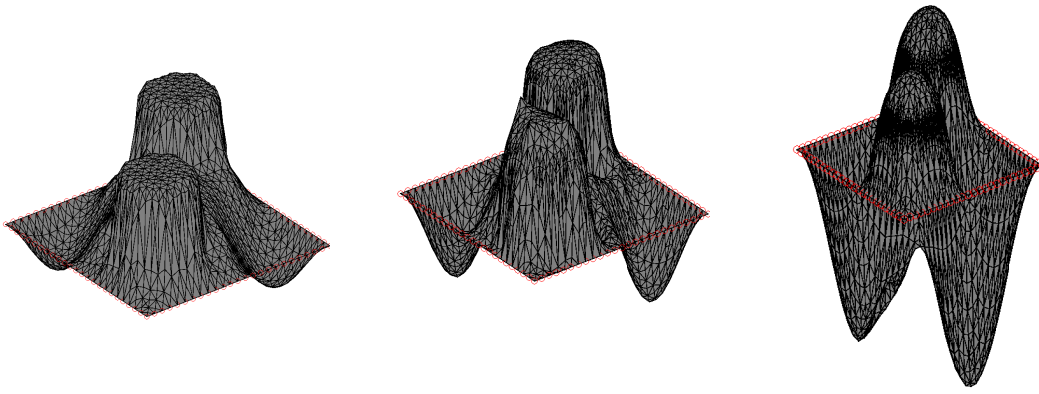


FIGURE 5. Snapshots of the finite element solution for Example 6.3 with  $\varepsilon = 6 \times 10^{-4}$  in the pre-asymptotic regime after 8, 14 and 22 adaptive refinements from an initial mesh with 144 elements.

with the nonlinear diffusion coefficient given by

$$\kappa(s) = k + \frac{1}{((\varepsilon + (s - a)^2)} + \frac{1}{((\varepsilon + (s - c)^2)}, \text{ with } a = 0.5, c = 0.8 \text{ and } k = 1. \quad (6.9)$$

The load  $f(x, y)$  is chosen so the exact solution  $u(x, y) = \sin(\pi x) \sin(\pi y)$ .

Local uniqueness of the solution again follows from [24], assuming the mesh is fine enough. Similarly to the error decrease in the first example, Figure 6 shows a temporary leveling off of the error while the estimator decreases at a steady rate at the beginning of the asymptotic phase indicating the presence of a nearby solution. The discontinuity in the error estimator and corresponding jump in the error is due to the final coarse mesh refinement where the solution is reset at adaptive level 22, after which the mesh maintains a maximum meshsize of  $h = 1.7 \times 10^{-3}$ , half the meshsize of Example 6.2. Once in the pre-asymptotic regime the internal layers are progressively resolved and the residual drops below tolerance converging at a quadratic rate when  $\gamma = 1$ , at adaptive level 37. Due to the aforementioned flattening of the iterates induced by the solver's stabilization, the approximate solutions gradually increase in height as the regularization parameters are reduced; as such, the inner internal layer centered at  $u = 0.8$ , is uncovered later than the outer layer centered at  $u = 0.5$ . This phenomenon is illustrated in the adaptive meshes at levels 25, 30 and 35 shown in Figure 7. The corresponding iterates shown

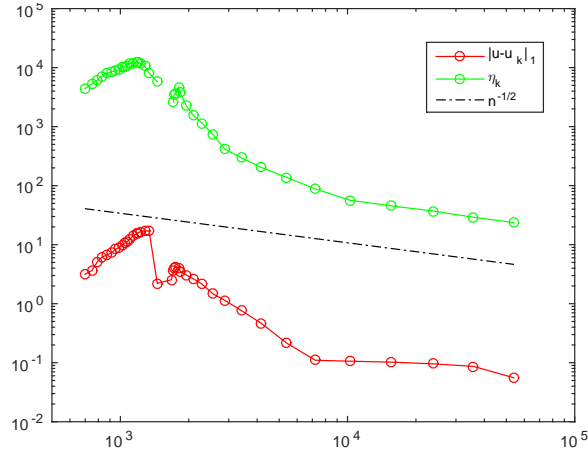


FIGURE 6.  $H^1$  error (below) and error estimator (above) against  $n^{-1/2}$  where  $n$  is the number of elements, for Example 6.4 with  $\varepsilon = 6 \times 10^{-4}$ .

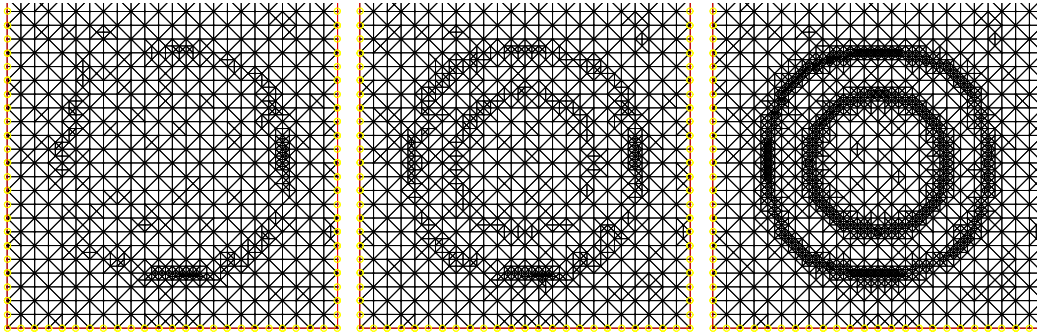


FIGURE 7. Mesh for Example 6.4 with  $\varepsilon = 6 \times 10^{-4}$  after 25, 30 and 35 adaptive refinements from an initial mesh with 144 elements.

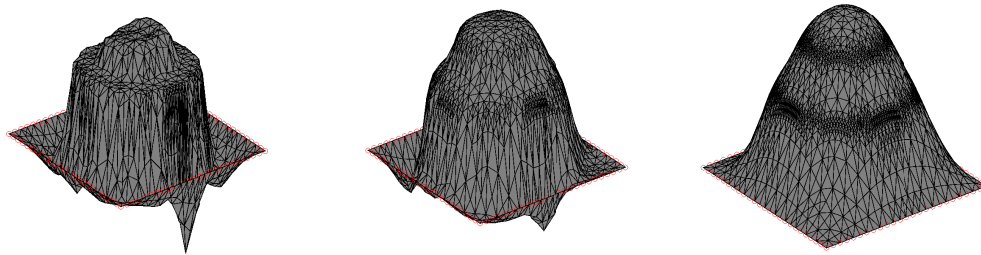


FIGURE 8. Snapshots of the finite element solution for Example 6.4 with  $\varepsilon = 6 \times 10^{-4}$  in the pre-asymptotic regime after 25, 30 and 35 adaptive refinements from an initial mesh with 144 elements.

in Figure 8 display the characteristic spikes between the outer layer and the boundary similar to those in Example 6.2, and in this case the peak of the sinusoid progressively resolves the curvature in two flatter regions.

## 7. CONCLUSION

The  $\sigma$ -split Newmark method of regularized pseudo-transient continuation is introduced to solve the nonlinear convection diffusion problem  $g(u) = 0$  starting from a coarse mesh where the problem data is not yet resolved. The method is derived and local  $q$ -linear convergence in agreement with the observed rate is established. An algorithm is presented to fit the solver into a standard adaptive method, and is demonstrated on three variations of a problem with steep internal layers. The method including the adaptive update of the solver's parameters on each refinement builds on the framework presented in [21], and is designed to effectively stabilize linearizations over rough problem data where spikes, overshoots and spurious oscillations otherwise prevent the sequence of transitional solutions from approaching an accurate approximation. The results here are an improvement first in terms of efficiency of the new solver which functions without the use of the normal equations; and second by an improved set of exit criteria for the solver which works together with an updated marking strategy to improve the stability of the method through the pre-asymptotic coarse mesh regime. Ultimately this yields a more efficient route to the asymptotic regime where the problem data is resolved and approximation properties hold. Future work will address targeted local regularization and parameter selection of the nonlinear solver, and more general types of nonlinearities.

## 8. ACKNOWLEDGMENTS

The author would like to thank William Rundell and Alan Demlow for providing input on a draft of this work.

## REFERENCES

- [1] R. BANK AND D. ROSE, *Parameter selection for Newton-like methods applicable to nonlinear partial differential equations*, SIAM J. Numer. Anal., 17 (1980), pp. 806–822.
- [2] L. BELENKI, L. DIENING, AND C. KREUZER, *Optimality of an adaptive finite element method for the  $p$ -Laplacian equation*, IMA J. Numer. Anal., 32 (2012), pp. 484–510.
- [3] G. CALOZ AND J. RAPPAZ, *Numerical analysis for nonlinear and bifurcation problems*, Handbook of Numerical Analysis, (1994).
- [4] J. CHUNG AND G. M. HULBERT, *A time integration algorithm for structural dynamics with improved numerical dissipation: The generalized- $\alpha$  method*, J. Appl. Mech., 60 (1993), pp. 371–375.
- [5] P. G. CIARLET, *Finite Element Method for Elliptic Problems*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2002.
- [6] T. S. COFFEY, C. T. KELLEY, AND D. E. KEYES, *Pseudo-transient continuation and differential-algebraic equations*, SIAM J. Sci. Comput., 25 (2003), pp. 553–569.
- [7] P. DEUFLHARD, *Newton Methods for Nonlinear Problems: Affine Invariance and Adaptive Algorithms*, Springer Publishing Company, Incorporated, 2011.
- [8] P. DEUFLHARD AND F. A. POTRA, *Asymptotic mesh independence of Newton-Galerkin methods via a refined Mysovskii theorem*, SIAM J. Numer. Anal., 29 (1992), pp. 1395–1412.
- [9] H. C. ELMAN AND A. RAMAGE, *A characterisation of oscillations in the discrete two-dimensional convection-diffusion equation*, Math. Comput., 72 (2003), pp. 263–288.
- [10] H. ENGL, M. HANKE, AND A. NEUBAUER, *Regularization of Inverse Problems*, Mathematics and Its Applications, Springer, 1996.
- [11] F. FIERRO AND A. VEESER, *On the a posteriori error analysis for equations of prescribed mean curvature*, Math. Comput., 72 (2003), pp. 1611–1634.
- [12] K. FOWLER AND C. KELLEY, *Pseudo-transient continuation for nonsmooth nonlinear equations*, SIAM J. Numer. Anal., 43 (2005), pp. 1385–1406.
- [13] E. M. GARAU, P. MORIN, AND C. ZUPPA, *Convergence of an adaptive Kačanov FEM for quasi-linear problems*, Applied Numerical Mathematics, 61 (2011), pp. 512 – 529.

- [14] G. GOLUB, D. SILVESTER, AND A. WATHEN, *Diagonal dominance and positive definiteness of upwind approximations for advection diffusion problems*, Numerical Analysis, A. R. Mitchell 75th Birthday Volume (1996), pp. 125–131. D. F. Griffiths and G. A. Watson, eds.
- [15] H. M. HILBER AND T. J. R. HUGHES, *Collocation, dissipation and overshoot for time integration schemes in structural dynamics*, Earthquake Eng. Struc., 6 (1978), pp. 99–117.
- [16] H. M. HILBER, T. J. R. HUGHES, AND R. L. TAYLOR, *Improved numerical dissipation for time integration algorithms in structural dynamics*, Earthquake Eng. Struc., 5 (1977), pp. 283–292.
- [17] C. HOFF AND P. PAHL, *Development of an implicit method with numerical dissipation from a generalized single-step algorithm for structural dynamics*, Comput. Method. Appl. M., 67 (1988), pp. 367 – 385.
- [18] M. HOLST, *Adaptive numerical treatment of elliptic systems on manifolds*, Adv. Comput Math., 15 (2001), pp. 139–191.
- [19] C. T. KELLEY AND D. KEYES, *Convergence analysis of pseudo-transient continuation*, SIAM J. Numer. Anal., 35 (1998), pp. 508–523.
- [20] N. M. NEWMARK, *A method of computation for structural dynamics*, J. Eng. Mech.-ASCE, 85 (1959), pp. 67–94.
- [21] S. POLLOCK, *A regularized Newton-like method for nonlinear PDE*, 2014. Available as arXiv:1412.6487 [math.NA].
- [22] K. G. SIEBERT, *A convergence proof for adaptive finite elements without lower bound*, IMA J. Numer. Anal., 31 (2011), pp. 947–970.
- [23] R. STEVENSON, *Optimality of a standard adaptive finite element method*, Found. Comput. Math., 7 (2007), pp. 245–269.
- [24] J. XU, *Two-grid discretization techniques for linear and nonlinear PDEs*, SIAM J. Numer. Anal., 33 (1996), pp. 1759–1777.

*E-mail address:* snpolloc@math.tamu.edu

DEPARTMENT OF MATHEMATICS, TEXAS A&M UNIVERSITY, COLLEGE STATION, TX 77843