Uniqueness of discrete solutions of nonmonotone PDEs without a globally fine mesh condition

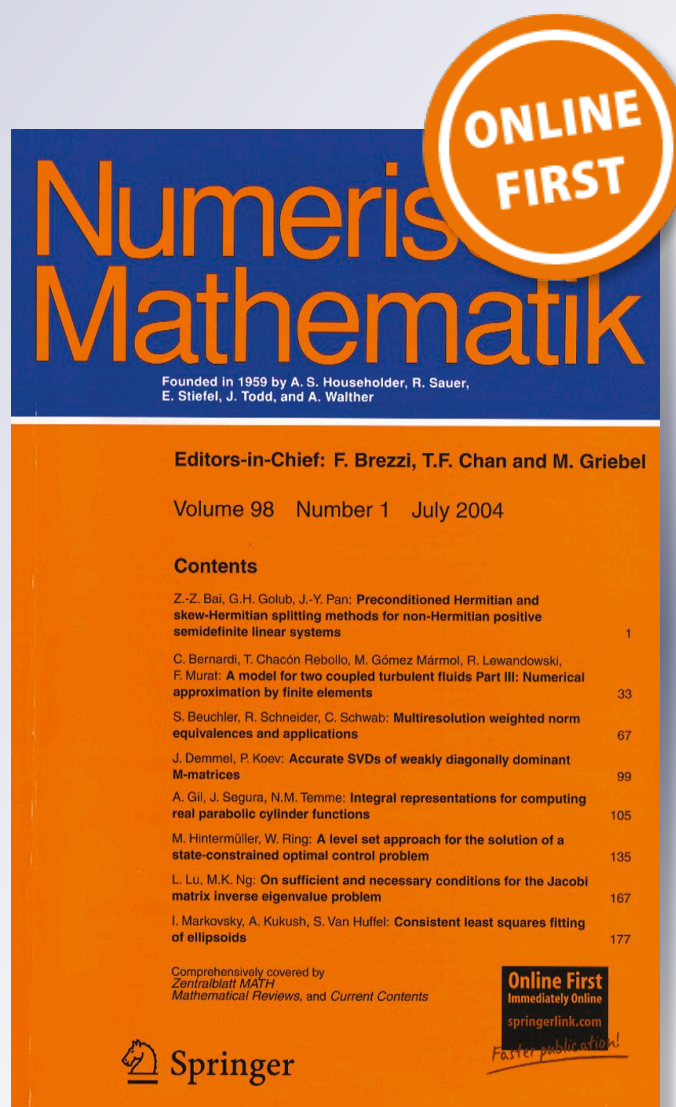## Sara Pollock & Yunrong Zhu

Contents

Comprehensively covered by
Zentralblatt MATH
Mathematical Reviews, and Current Contents

Online First
Immediately Online
springerlink.com

🌱 Springer

🌱 Springer

Springer

Uniqueness of discrete solutions of nonmonotone PDEs without a globally fine mesh condition

Sara Pollock[1] · **Yunrong Zhu**[2]

Abstract Uniqueness of the finite element solution for nonmonotone quasilinear problems of elliptic type is established in one and two dimensions. In each case, we prove a comparison theorem based on locally bounding the variation of the discrete solution over each element. The uniqueness follows from this result, and does not require a globally small meshsize.

Mathematics Subject Classification 65N30 · 35J65

1 Introduction

We consider the piecewise linear finite element (FE) approximation of the quasilinear elliptic partial differential equation (PDE)

$$-\text{div}(\kappa(x, u)\nabla u) = f \ \text{ in } \ \Omega. \tag{1.1}$$

The following assumptions on the diffusion coefficient $\kappa(x, s)$, for $x \in \Omega$ and $s \in \mathbb{R}$ are made throughout the remainder of the paper. Further assumptions on the source

✉ Sara Pollock
sara.pollock@wright.edu

Yunrong Zhu
zhuyunr@isu.edu

1 Department of Mathematics and Statistics, Wright State University, Dayton, OH 45435, USA

2 Department of Mathematics and Statistics, Idaho State University, Pocatello, ID 83209, USA

function f are stated as needed in the following sections, along with the boundary conditions corresponding to (1.1).

Assumption 1 Assume $\kappa(x, s)$ is a Carathéodory function, and assume there are constants $0 < k_\alpha < k_\beta$ with

$$k_\alpha < \kappa(x, s) < k_\beta, \tag{1.2}$$

for all $s \in \mathbb{R}$, and a.e. $x \in \Omega$.

Further assume the diffusion coefficient κ satisfies the following Lipschitz condition.

Assumption 2 There is a constant $L_0 > 0$ with

$$|\kappa(x, s) - \kappa(x, t)| \le L_0|s - t|, \quad \text{for all } s, t \in \mathbb{R}, \text{ a.e. } x \in \Omega. \tag{1.3}$$

Existence and uniqueness of smooth solutions, $u \in C^2(\bar{\Omega})$, to the continuous problem (1.1) were established for the Dirichlet problem in [7]. Those results were extended in [9] to cover existence and uniqueness of weak solutions $u \in H^1(\Omega)$ under mixed Dirichlet/Robin conditions on the boundary, and for Dirichlet problems in [3]. Uniqueness results and comparison principles for more general classes of quasilinear problems are also found in Chapter 10 of [8], and in [2], and the references therein. As demonstrated by the counterexamples shown in [4], uniqueness for the discrete problem may fail, even where it holds for the corresponding continuous problem.

Let $\Omega \subset \mathbb{R}^n$, $n \ge 1$, be an open bounded polygonal domain, (an interval, if $n = 1$) with boundary $\partial\Omega$. Let $\mathcal{V}$ denote the discrete space of continuous piecewise linear functions subordinate to a conforming simplicial partition $\mathcal{T}$ of Ω such that functions $v \in \mathcal{V}$ vanish on the prescribed Dirichlet part of the boundary. We consider the approximation of solutions to the weak form of (1.1) by functions $u \in \mathcal{V}$. We refer to Theorem 1.2 of [4] and Theorem 2.6 of [9] to establish the existence of a solution to the discrete Dirichlet and respectively mixed Dirichlet/Neumann problems. Succinctly, let $f \in H^{-1}(\Omega)$ in the Dirichlet case, or $f \in L_2(\Omega)$ for the case of mixed boundary conditions, with bounded, measurable data on the boundary. Then under Assumptions 1 and 2, there exists a solution $u \in \mathcal{V}$ to (1.1) satisfying the prescribed conditions on the boundary. The upper bound k_β in Assumption 1 is used in establishing existence, and not elsewhere in the remainder of this paper.

We now turn our attention to sufficient conditions to establish uniqueness of the discrete solution. In the recent investigation [1], regarding the effects of numerical quadrature on nonmonotone nonlinear problems over $\Omega \subset \mathbb{R}^d$, $d \le 3$, the condition of a globally sufficiently fine mesh as required in [4] is referred to as essential for establishing the uniqueness of the discrete solution. Within the context of adaptive finite element methods, convergence of a Discontinuous Galerkin method for this class of problems is shown in [5], under the assumption that the mesh is globally sufficiently fine; more precisely, under the assumptions of [6] which develops local convergence theory for the piecewise linear finite element solution. However, recent work by one of the present authors on the numerical solution of the same class of nonmonotone problems by means of adaptive regularized methods [12–14], demonstrates that a

solution to the discrete problem may be recovered under only local adaptive mesh refinement. This raises the question of whether the solution to the discrete problem can be shown unique under only local criteria. This paper answers the question in the positive.

Inspired by the analysis of [4], the discrete finite element solution is shown here to be unique so long as the variation of the discrete solution over each mesh element satisfies a given bound related to the constants in Assumptions 1 and 2. To our knowledge, this is the first result indicating uniqueness of the discrete solution for this class of problems that does not require a globally fine mesh. The results here are viewed as an important step towards establishing the convergence of an efficiently computed discrete solution by means of adaptive mesh refinement. The issue of convergence of the discrete solution to the PDE solution under adaptive refinement is not addressed here, as the analysis of appropriate a posteriori error indicators is nontrivial and merits separate investigation.

Our approach to establishing uniqueness under local conditions is based upon first establishing a comparison theorem from which the uniqueness of the solution immediately follows. Among the advantages of this approach over the iterative arguments of [1,6] are the weaker requirements both on the problem data and the fineness of the discretization. Here we require the coefficient $\kappa(x, s)$ to be Lipschitz with respect to its second argument, which agrees with the conditions on κ for the continuous problem to have a unique solution [9]. In contrast, κ is assumed twice continuously differentiable in [6] and [1, Theorem 5], where uniqueness is established following the local convergence of Newton iterations. Moreover, [1, Theorem 4] addresses uniqueness of solutions without the iterative argument, but based on a priori estimates requiring a smallness assumption on the H_2 norm of the PDE solution u.

The results above additionally require a globally small enough meshsize $h < h_0$ and the quasi-uniformity of the mesh, meaning the ratio $h_{\mathcal{T}}/h_T$ of the maximum element diameter compared with the diameter of each element is uniformly bounded by some constant C. This constant appears in bounding the radius of converence of the Newton iterations away from zero, and in bounding the norm of discrete functions in [1, Theorem 4]. In the current manuscript, we avoid the use of standard scaling arguments which introduce a global meshsize, instead relying on direct elementwise integration. We emphasize however that similarly to [4], we do not establish the computability of the finite element solution here, only the uniqueness of the discrete solution. The computability of solutions on nonuniform adaptive meshes which do not uphold the quasi-uniformity condition is demonstrated by the first author in [13,14], relying on a regularized adaptive method to enter the domain of convergence of a Newton-like iteration.

We remark that discrete maximum principles have been established for this problem and some of its generalizations without requirements on the global meshsize, as in [11,15] and the references therein. These are important tools for studying nonlinear problems because they provide a priori bounds on the discrete solution. However, maximum principles are not sufficient to establish uniqueness for the nonlinear PDE, even in the continuous setting.

The remainder of the paper is organized as follows. In Sect. 2 we establish a comparison theorem followed by a uniqueness result for the one dimensional Dirichlet

problem. The analysis is then extended to the case of mixed boundary conditions. In Sect. 2.1 we demonstrate the local criterion for uniqueness on a counterexample where two discrete solutions are known to exist. Next, in Sect. 3 we first review some useful facts about the finite element basis functions, then in 3.2 we prove a comparison theorem followed by a uniqueness result for the mixed Dirichlet/Neumann problem. Finally, we demonstrate the local uniqueness condition on a known two dimensional counterexample with two discrete solutions.

2 Uniqueness in the one dimensional case

The one dimensional quasilinear equation is given by the ordinary differential equation

$$-\frac{d}{dx}\left(\kappa(x,u)\frac{d}{dx}u\right) = f, \text{ in } \Omega \subset \mathbb{R}, \tag{2.1}$$

subject to appropriate boundary conditions as described in the main theorems below.

As in [4] let $\Omega = (a,b)$, with a subdivision

$$a = a_0 < a_1 < \cdots < a_{n-1} < a_n = b, \tag{2.2}$$

where the mesh spacing is not assumed to be uniform. Define the discrete space $\mathcal{V}_0 \subset H_0^1(\Omega)$ as the piecewise linear Lagrange finite element space subordinate to subdivision (2.2), and subject to homogeneous Dirichlet boundary conditions: $v(a) = v(b) = 0$ for all $v \in \mathcal{V}_0$. Let (u,v) denote the integral $\int_\Omega uv$, and let $v' = dv/dx$.

Theorem 3 (Comparison Theorem in 1D) *Assume f_1, $f_2 \in H^{-1}(\Omega)$ satisfy*

$$f_1 \leq f_2, \tag{2.3}$$

in the $H^{-1}(\Omega)$ sense, that is $\int_\Omega f_1 v \leq \int_\Omega f_2 v$ for all non-negative functions $v \in H_0^1(\Omega)$. Let $u_i \in \mathcal{V}_0$, $i = 1,2$, denote the solution to the respective weak forms of (2.1), with homogeneous Dirichlet boundary conditions. In particular

$$(\kappa(x,u_i)u_i', v') = (f_i, v), \text{ for all } v \in \mathcal{V}_0, \quad i = 1,2. \tag{2.4}$$

Define the intervals $\mathcal{I}_k = (a_{k-1}, a_k)$, $k = 1, \ldots, n$, and let $h_k = a_k - a_{k-1}$, the length of each respective interval $\mathcal{I}_k$, $k = 1, \ldots, n$. Then under the condition

$$\max_{0 \leq k \leq n-1} |u_2(a_k) - u_2(a_{k+1})| < \frac{2k_\alpha}{L_0}, \tag{2.5}$$

it holds that

$$u_1 \leq u_2. \tag{2.6}$$

The proof follows similarly to Theorem 2.1 in [4] with an updated choice of test-function and the avoidance of global norm estimates.

Proof For any test function $v \in \mathcal{V}_0$, subtracting the cases $i = 1$ and $i = 2$ in (2.4) yields

$$(\kappa(x, u_1)(u_1 - u_2)', v') = \big((\kappa(x, u_2) - \kappa(x, u_1))u_2', v'\big) + (f_1 - f_2, v). \quad (2.7)$$

We prove the theorem by contradiction. Suppose that $w(x) = u_1(x) - u_2(x) > 0$ anywhere on $a \leq x \leq b$, then we can define the test function $v \in \mathcal{V}_0$ as follows. Let i denote the smallest index such that $w(a_i) > 0$; and, let j be the smallest index with $j \geq i$, and such that $w(a_{j+1}) \leq 0$. Then define the piecewise linear function v by its nodal values

$$\begin{cases} v(a_k) = 1, \ k = i, \ldots, j, \\ v(a_k) = 0, \ \text{otherwise.} \end{cases} \quad (2.8)$$

As u_1 and u_2 satisfy the same boundary conditions, we have $v(a_0) = 0$ and $v(a_n) = 0$ so that v is a valid test-function, that is $v \in \mathcal{V}_0$. As $v' = 0$ on $\Omega \setminus \{\mathcal{I}_i \cup \mathcal{I}_{j+1}\}$, the left hand side of (2.7) satisfies

$$\begin{aligned}
(\kappa(x, u_1)(u_1 - u_2)', v') &= \int_{\mathcal{I}_i} \kappa(x, u_1)w'v' + \int_{\mathcal{I}_{j+1}} \kappa(x, u_2)w'v' \\
&= \int_{\mathcal{I}_i} \kappa(x, u_1)\frac{w(a_i) - w(a_{i-1})}{h_i^2} + \int_{\mathcal{I}_{j+1}} \kappa(x, u_1)\frac{w(a_j) - w(a_{j+1})}{h_{j+1}^2} \\
&\geq k_\alpha \left\{ \frac{w(a_i) - w(a_{i-1})}{h_i} + \frac{w(a_j) - w(a_{j+1})}{h_{j+1}} \right\}, \quad (2.9)
\end{aligned}$$

where the ellipticity condition (1.2) was applied, as was the fact $w(a_i) > w(a_{i-1})$, and $w(a_j) > w(a_{j+1})$, from the definition of v, alongside $v' = 1/h_i$ on $\mathcal{I}_i$, and $v' = -1/h_{j+1}$ on $\mathcal{I}_{j+1}$.

Similarly, the first term on the right hand side of (2.7) satisfies

$$\begin{aligned}
&\big((\kappa(x, u_2) - \kappa(x, u_1))u_2', v'\big) \\
&= \int_{\mathcal{I}_i} (\kappa(x, u_2) - \kappa(x, u_1))u_2'v' + \int_{\mathcal{I}_{j+1}} (\kappa(x, u_2) - \kappa(x, u_1))u_2'v' \\
&= \int_{\mathcal{I}_i} (\kappa(x, u_2) - \kappa(x, u_1))\frac{u_2(a_i) - u_2(a_{i-1})}{h_i^2} \\
&\quad - \int_{\mathcal{I}_{j+1}} (\kappa(x, u_2) - \kappa(x, u_1))\frac{u_2(a_{j+1}) - u_2(a_j)}{h_{j+1}^2} \\
&\leq \frac{L_0}{h_i^2}|u_2(a_i) - u_2(a_{i-1})| \left(\int_{\mathcal{I}_i} |w| \right) + \frac{L_0}{h_{j+1}^2}|u_2(a_{j+1}) - u_2(a_j)| \left(\int_{\mathcal{I}_{j+1}} |w| \right), \\
&\hspace{11cm} (2.10)
\end{aligned}$$

where both the Lipschitz property (1.3) of κ, and the value of v on $\mathcal{I}_i$ and $\mathcal{I}_{j+1}$, as in (2.9), were applied. Noticing that $w(a_{i-1}) \leq 0$ and $w(a_i) > 0$, and similarly of $w(a_{j+1})$ and $w(a_j)$, it follows from the piecewise linearity of w that

$$\int_{\mathcal{I}_i} |w| \leq \frac{h_i}{2}(w(a_i) - w(a_{i-1})), \text{ and } \int_{\mathcal{I}_{j+1}} |w| \leq \frac{h_{j+1}}{2}(w(a_j) - w(a_{j+1})).$$
(2.11)

Applying (2.11) to (2.10), yields

$$\begin{aligned}((\kappa(x, u_2) - \kappa(x, u_1))u_2', v') &\leq \frac{L_0}{2h_i}(w(a_i) - w(a_{i-1}))|u_2(a_i) - u_2(a_{i-1})| \\ &+ \frac{L_0}{2h_{j+1}}(w(a_j) - w(a_{j+1}))|u_2(a_j) - u_2(a_{j+1})|.\end{aligned}$$
(2.12)

Putting together (2.9) and (2.12) with (2.7) yields

$$\begin{aligned}k_\alpha &\left\{\frac{w(a_i) - w(a_{i-1})}{h_i} + \frac{w(a_j) - w(a_{j+1})}{h_{j+1}}\right\} \\ &\leq \frac{L_0}{2h_i}(w(a_i) - w(a_{i-1}))|u_2(a_i) - u_2(a_{i-1})| \\ &+ \frac{L_0}{2h_{j+1}}(w(a_j) - w(a_{j+1}))|u_2(a_j) - u_2(a_{j+1})| + (f_1 - f_2, v).\end{aligned}$$
(2.13)

Rearranging terms yields

$$\begin{aligned}\frac{1}{h_i} &\left\{k_\alpha - \frac{L_0}{2}|u_2(a_i) - u_2(a_{i-1})|\right\}(w(a_i) - w(a_{i-1})) \\ &+ \frac{1}{h_{j+1}}\left\{k_\alpha - \frac{L_0}{2}|u_2(a_{j+1}) - u_2(a_j)|\right\}(w(a_j) - w(a_{j+1})) \\ &\leq (f_1 - f_2, v).\end{aligned}$$
(2.14)

Noting again the terms $(w(a_i) - w(a_{i-1}))$ and $(w(a_j) - w(a_{j+1}))$, are both strictly positive, and $(f_1 - f_2, v) \leq 0$, this yields a contradiction under the condition (2.5), from which it is clear the test-function v cannot be so defined, and the conclusion follows. □

The main purpose of Theorem 3 is to establish the following immediate corollary, namely the uniqueness of the one dimensional linear FE solution without a global meshsize condition. The sufficient condition used here bounds the change in the FE solution over a single mesh element. This of course can be controlled by *locally* reducing the meshsize. As such, the result is suitable for an adaptive mesh refinement algorithm.

Corollary 1 (Uniqueness of the FE Solution in 1D) *Let $f \in H^{-1}(\Omega)$, and let $u \in \mathcal{V}_0$ denote a solution to the weak form of* (2.1)*, with homogeneous Dirichlet boundary conditions. In particular*

$$(\kappa(x, u)u', v') = (f, v), \text{ for all } v \in \mathcal{V}_0. \tag{2.15}$$

Then under the condition

$$\max_{0 \le k \le n-1} |u(a_k) - u(a_{k+1})| < \frac{2k_\alpha}{L_0}, \tag{2.16}$$

(cf., (2.5)*), the linear FE solution u is the unique solution to* (2.15)*.*

Proof Let $f_1 = f_2 = f$ in Theorem 3. Then, let u_1 and u_2 be two respective solutions to the problem (2.15). Applying the results of Theorem 3 yields both $u_1 \le u_2$ and $u_2 \le u_1$. □

Notably, the argument does not depend on the global meshsize, rather the variation in the discrete piecewise linear solution over each element. As per the following remark, this can be rephrased as a condition involving the local meshsize.

Remark 1 The condition (2.16) is local, and can be checked computationally. These two qualities make it amenable for use in an adaptive algorithm to assure uniqueness of the discrete solution. The condition can equivalently be written as

$$|u'_2|_{\mathcal{I}_k} < \frac{2k_\alpha}{h_k L_0}, \text{ for all } k = 1, \ldots n.$$

In other words, the mesh should be *locally* fine, where the slope of the solution is steep.

The next corollary generalizes Theorem 3 to the case of mixed Dirichlet-Neumann boundary conditions.

Corollary 2 (Comparison Theorem for Mixed Boundary Conditions in 1D) *Assume $f_1, f_2 \in L_2(\Omega)$ satisfy $f_1 \le f_2$ in Ω. Define $\mathcal{V}_{0,b}$ as the piecewise linear Lagrange finite element space subordinate to subdivision* (2.2)*, with $v(b) = 0$ for all $v \in \mathcal{V}_{0,b}$. Let $u_i \in \mathcal{V}_{0,b}$, $i = 1, 2$ denote the solution to the respective weak forms of* (2.1) *with the mixed boundary conditions*

$$u(b) = 0, \text{ and } \kappa(a, u(a))u'(a) = \psi_a, \tag{2.17}$$

for a given $\psi_a \in \mathbb{R}$. In particular

$$(\kappa(x, u_i)u'_i, v') = (f_i, v) + \psi_a v(a), \text{ for all } v \in \mathcal{V}_{0,b}, \quad i = 1, 2. \tag{2.18}$$

Define the intervals $\mathcal{I}_k$, and their respective lengths h_k as in Theorem 3, for $k = 1, \ldots, n$. Then under the condition (2.5)*, it holds that $u_1 \le u_2$.*

The proof follows again by construction of an appropriate test function v.

Proof For any test function $v \in \mathcal{V}_{0,b}$, subtracting the cases $i = 1$ and $i = 2$ in (2.18) yields

$$(\kappa(x, u_1)(u_1 - u_2)', v') = \big((\kappa(x, u_2) - \kappa(x, u_1))u_2', v'\big) + (f_1 - f_2, v). \quad (2.19)$$

As before, we define $w = u_1 - u_2$ and prove the result by contradiction. Supposing that $w(x) > 0$ anywhere on $a \le x \le b$, two cases are now possible. In the first case, there is a least index $i < n$ with $w(a_{i-1}) \le 0$ and $w(a_i) > 0$. Then, because $w(b) = 0$, there is certain to be a least index $j \ge i$ for which $w(a_j) > 0$ and $w(a_{j+1}) \le 0$. The proof then follows that of Theorem 3, with v chosen by (2.8).

The second case is characterized by $w(a_k) > 0$ for all $k \le j$, for some $j \le n - 1$. In particular, there is some smallest index $j < n$ such that $w(a_j) > 0$, and $w(a_k) \le 0$ for all $k = j + 1, \ldots, n$. It remains then to establish the result assuming this is the case. Define the test function $v \in \mathcal{V}_{0,b}$ by

$$\begin{cases} v(a_k) = 1, \ k = 0, \ldots, j, \\ v(a_k) = 0, \ k = j + 1, \ldots, n. \end{cases} \quad (2.20)$$

Noting that $v' = -1/h_{j+1}$ on $\mathcal{I}_{j+1}$, and $v' = 0$, otherwise, the argument follows as before with the following estimates. Inequality (2.9) reduces to

$$\begin{aligned}
(\kappa(x, u_1)w', v') &= \int_{\mathcal{I}_{j+1}} \kappa(x, u_2) \frac{w(a_j) - w(a_{j+1})}{h_{j+1}^2} \\
&\ge k_\alpha \left\{ \frac{w(a_j) - w(a_{j+1})}{h_{j+1}} \right\}.
\end{aligned} \quad (2.21)$$

Inequality (2.10) reduces to

$$((\kappa(x, u_2) - \kappa(x, u_1))u_2', v') \le \frac{L_0}{2h_{j+1}}(w(a_j) - w(a_{j+1}))|u_2(a_j) - u_2(a_{j+1})|. \quad (2.22)$$

Putting together (2.21) and (2.22) with (2.20), and rearranging terms, yields

$$\frac{1}{h_{j+1}} \left\{ k_\alpha - \frac{L_0}{2}|u_2(a_{j+1}) - u_2(a_j)| \right\} (w(a_j) - w(a_{j+1})) \le (f_1 - f_2, v). \quad (2.23)$$

Again noting the strict positivity of $(w(a_j) - w(a_{j+1}))$, a contradiction is encountered under the satisfaction of the condition (2.5), from which the conclusion follows. $\quad \square$

As in the case of homogeneous Dirichlet boundary conditions, the main purpose of Corollary 2 is to establish the uniqueness of the piecewise linear FE solution under conditions suitable for adaptive mesh refinement. The next corollary for the case of mixed boundary conditions establishes the analogous result to Corollary 1.

Corollary 3 (Uniqueness of the FE Solution under Mixed BC in 1D) *Let $f \in L_2(\Omega)$, and let $u \in \mathcal{V}_{0,b}$ denote a solution to the weak form of* (2.1), *with the mixed boundary conditions* (2.17). *In particular*

$$(\kappa(x, u)u', v') = (f, v) + \psi_a v(a), \text{ for all } v \in \mathcal{V}_{0,b}. \qquad (2.24)$$

Then, under the condition (2.16), *the linear FE solution u is the unique solution to* (2.24).

Proof Let $f_1 = f_2 = f$ in Corollary 2. Then let u_1 and u_2 be two respective solutions to (2.24). Applying the results of Corollary 2 yields the result. □

Remark 2 Corollaries 2 and 3 trivially generalize to the mixed boundary value problem with a homogeneous Dirichlet condition at $x = a$ and a Neumann condition at $x = b$.

2.1 Addressing the counterexample of André and Chipot

The analysis of [4] not only established the uniqueness of the finite element solution in one and two dimensions under the sufficient condition of a globally small meshsize, but also provided a counterexample. In this example, a diffusion coefficient $\kappa(x, u)$ for which $-(\kappa(x, u)u')' = f$ has two discrete solutions, u and $2u$, is constructed on a set partition. Here, we show the counterexample in question violates the condition (2.16). We refer the reader to Section 2 of [4] for the complete details of the example, and explain here how the local condition (2.16) detects the possibility of nonuniqueness.

The diffusion coefficient $\kappa(x, z)$ is constructed to satisfy $1/3 \leq \kappa(x, z) \leq 1$, for all $z \in \mathbb{R}$ and a.e. $x \in [0, 1]$. That is, (1.2) is satisfied with $k_\alpha = 1/3$. The coefficient $\kappa(x, z)$ is defined in terms of a function $\phi(t) \in C^\infty(0, 1)$ which satisfies $\phi'(0) = \phi'(1) = 0$ and

$$\phi(0) = \phi(1) = 1, \quad 1/3 \leq \phi(t) \leq 1, \text{ and } \int_0^1 \phi(t) = 1/2. \qquad (2.25)$$

A first solution u defined on a uniform partition corresponding to the notation of (2.2), is assumed to satisfy $u(0) = u(1) = 0$, and $u(a_i) = u_i > 0$, $i = 1, \ldots, n - 1$. As demonstrated in equation (2.38) in [4], on the first interval, the derivative of $\kappa(x, z)$ with respect to its second argument satisfies

$$\left| \frac{\partial}{\partial z} \kappa(x, z) \right| = \left| \frac{\phi(t) - \phi(0)}{t u_1} \right|, \quad a_0 \leq x \leq a_1, \quad u(x) \leq z(x) \leq 2u(x). \qquad (2.26)$$

However, the properties of ϕ given by (2.25) require that the magnitude of the slope of the secant line, $|\phi(0) - \phi(t)|/t > 2/3$ for some $t \in (a_0, a_1]$, for $z(x)$ in the given range. This is seen by comparison against the slope of the linear function $\psi_0(t)$ given by $\psi_0(0) = 1$ and $\psi_0(1) = 1/3$. This implies a lower bound on the Lipschitz constant, namely $L_0 > (2/3)/u_1$. The local condition for uniqueness in the present discussion given by (2.16), requires on the first element of the partition, $(u_1 - 0) < 2k_\alpha/L$.

Applying $k_\alpha = 1/3$ and $L_0 > (2/3)/u_1$, the condition for uniqueness on the first element requires

$$u_1 < \frac{2k_\alpha}{L_0} < \frac{(2/3)u_1}{(2/3)} = u_1. \tag{2.27}$$

As such, the local condition (2.16) detects the possibility of a nonunique solution.

Remark 3 It is further noted that if (2.25) is modified so that $k \le \phi(t) \le 1$, for any $0 < k < 1/3$, then $k_\alpha = k$, and $L_0 > (1 - k)/u_1$, by the same reasoning as above. Then in place of (2.27) we have $u_1 < 2ku_1/(1 - k) < u_1$, violating the condition for uniqueness. More generally, for $0 < k < 1/2$, the magnitude of the slope of the secant line from $(0, \phi(0))$ to $(t, \phi(t))$ must be at least $|\phi(0) - \phi(t)|/t > (1 - k)^2/(1 - 2k)$ for some $t \in (0, 1)$ in order to simultaneously satisfy the first and last conditions of (2.25). This is found by considering the area below the curve $\psi(t) = \max\{k, 1 - st\}$ from $t = 0$ to $t = 1$ with s defined so the area $\int_0^1 \psi(t) = 1/2$, namely, $s = -(1 - k)^2/(1 - 2k)$. Then $L_0 > |s|/u_1$ and $2k/L_0 < u_1 \cdot 2k(1 - 2k)/(1 - k^2) < u_1$.

3 Uniqueness in the two-dimensional case

Let $\Omega \subset \mathbb{R}^2$ be an open bounded polyhedral domain, with boundary $\partial\Omega = \Gamma_D \cup \Gamma_N$, where Γ_D has positive measure, and $\Gamma_N = \partial\Omega \setminus \Gamma_D$. Homogeneous Dirichlet boundary conditions will be imposed on Γ_D, and Neumann conditions will be imposed on Γ_N. We next establish the uniqueness of the piecewise linear finite element solution to the discrete approximation corresponding to (1.1), in two dimensions, under the assumptions 1 and 2. For simplicity of defining the finite element solution space, the discussion assumes a mixed boundary condition, with a homogeneous Dirichlet part. However, the method of the proof trivially generalizes to allow nonhomogeneous Dirichlet data, or pure Dirichlet conditions, where $\Gamma_D = \partial\Omega$.

Let $\mathcal{T}$ be a conforming triangulation of domain Ω by triangles that exactly captures the boundary $\partial\Omega$, and each of Γ_D and Γ_N. The mesh is assumed to satisfy the following acuteness condition, meaning the smallest angle is bounded away from zero.

Assumption 4 (*Mesh regularity*) There are numbers $0 < t_{min} < t_{max}$, for which the interior angles θ_i, $i = 1, 2, 3$, of each $T \in \mathcal{T}$ satisfy

$$t_{min} \le \theta_i \le t_{max} < \pi/2, \; i = 1, 2, 3. \tag{3.1}$$

Similar assumptions on the angles of the mesh are made in [4] (Section 4); see also [10] concerning a related class of quasilinear problems featuring a gradient-dependent nonlinearity and mixed boundary conditions. The acuteness of each triangle disallows the orthogonality of the gradients of the standard finite element basis functions. The smallest-angle assumption preserves the condition of the mesh, and necessarily for the discussion that follows, the minimum ratio of edges in a given triangle. To that end, for each $T \in \mathcal{T}$, denote the minimum ratio of sines of interior angles by

$$\gamma_T := \min_{i,j\in\{1,2,3\}} \frac{\sin(\theta_i)}{\sin(\theta_j)}, \quad \text{for } \theta_i, \ i = 1, 2, 3, \quad \text{the angles of } T, \tag{3.2}$$

$$c_T := \min_{i=1,2,3} \cos(\theta_i). \tag{3.3}$$

Under Assumption 4, $\sin(t_{min}) < \gamma_T \leq 1$, and $\cos(t_{max}) < c_T < \cos(t_{min})$, for each $T \in \mathcal{T}$.

Let $\mathcal{D}$ be the collection of vertices or nodes of partition $\mathcal{T}$, where each $d \in \mathcal{D}$ has coordinates $d = (x, y) \in \overline{\Omega}$. Define the discrete space $\mathcal{V}_{0,D}$ as space of the piecewise linear functions subordinate to partition $\mathcal{T}$, that vanish on Γ_D. Without confusion, the following Sect. 3.1 represents $a = (x, y) \in \Omega$, as a coordinate representation, whereas the remainder of the text follows the convention $x \in \Omega$ as a point in the domain with two associated physical coordinates.

3.1 Basic setup in two dimensions

Some standard notations for the two dimensional problem are first reviewed, for the ease of presentation to the reader. Let $a_1 = (x_1, y_1)$, $a_2 = (x_2, y_2)$, and $a_3 = (x_3, y_3)$ be a local counterclockwise numbering of the vertices of a simplex $T \in \mathcal{T}$. Let the corresponding edges $\{e_1, e_2, e_3\}$, follow a consistent local numbering, namely edge e_i is opposite vertex a_i, $i = 1, 2, 3$. Let φ_i the basis function on element $T \in \mathcal{T}$ defined by its nodal values at the vertices of T, that is

$$\varphi_i(a_j) = \begin{cases} 1, & i = j, \\ 0, & i \neq j \end{cases}, \quad i, j = 1, 2, 3.$$

The inner product between gradients of basis functions, and their respective integrals over elements $T \in \mathcal{T}$, may be computed by change of variables to a reference element $\widehat{T}$, in reference domain variables $(\widehat{x}, \widehat{y})$. Specifically, the coordinates of $\widehat{T}$ are given as $(\widehat{x}_1, \widehat{y}_1) = (0, 0)$, $(\widehat{x}_2, \widehat{y}_2) = (1, 0)$, $(\widehat{x}_3, \widehat{y}_3) = (0, 1)$. The Jacobian of the transformation between reference and physical coordinates, and the corresponding Jacobian J are given by

$$J\begin{pmatrix} \widehat{x} \\ \widehat{y} \end{pmatrix} = \begin{pmatrix} x - x_1 \\ y - y_1 \end{pmatrix}, \quad J = \begin{pmatrix} x_2 - x_1 & x_3 - x_1 \\ y_2 - y_1 & y_3 - y_1 \end{pmatrix}, \tag{3.4}$$

with $\det J = 2|T|$, where $|T|$ is the area of triangle T. The reference element $\widehat{T}$ is equipped with the nodal basis functions $\widehat{\varphi}_i$, $i = 1, 2, 3$, where

$$\widehat{\varphi}_1 = 1 - \widehat{x} - \widehat{y}, \quad \widehat{\varphi}_2 = \widehat{x}, \quad \widehat{\varphi}_3 = \widehat{y},$$

and gradients $\widehat{\nabla}$ taken with respect to the reference domain variables $\widehat{x}$ and $\widehat{y}$

$$\widehat{\nabla}\widehat{\varphi}_1 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \quad \widehat{\nabla}\widehat{\varphi}_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \widehat{\nabla}\widehat{\varphi}_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Then $\nabla\varphi_i = J^{-T}\widehat{\nabla}\widehat{\varphi}_i$ and it is useful to note that

$$\nabla\varphi_i + \nabla\varphi_j = -\nabla\varphi_k, \tag{3.5}$$

for any distinct assignment of i, j and k to the integers $\{1, 2, 3\}$. Multiplication by the inverse of the Jacobian (3.4), allows the representation of $\nabla\varphi_i^T \nabla\varphi_i$ in terms of edge-length $|e_i|$. The identity (3.5) allows the expansion

$$\nabla\varphi_k^T \nabla\varphi_k = \nabla(\varphi_i + \varphi_j)^T \nabla(\varphi_i + \varphi_j),$$

and the representation of the inner product $\nabla\varphi_i^T \nabla\varphi_j$ in terms of the three edge-lengths of associated triangle T.

$$\nabla\varphi_i^T \nabla\varphi_j = \frac{1}{|\det J|^2} \begin{cases} |e_j|^2, & i = j, \\ (|e_k|^2 - |e_i|^2 - |e_j|^2)/2, & i \neq j. \end{cases} \tag{3.6}$$

The maximum interior angle less than $\pi/2$ from Assumption 4, together with (3.6) assures the negativity of any $\nabla\varphi_i^T \nabla\varphi_j$, for $i \neq j$. Further, the integral of each inner product over element T in the physical domain then satisfies the following.

$$\int_T \nabla\varphi_i^T \nabla\varphi_i = \frac{1}{4|T|}|e_i|^2, \quad \text{and} \quad \int_T \nabla\varphi_i^T \nabla\varphi_j = -\frac{|e_i||e_j|}{4|T|}\cos\theta_k. \tag{3.7}$$

The minimum and maximum interior angle conditions in Assumption 4, together with the law of sines, allows a bound on the minimum and maximum ratio of edges in a triangle in terms of γ_T, defined in (3.2). For each $T \in \mathcal{T}$ it holds that

$$\gamma_T|e_j| \leq |e_i| \leq \gamma_T^{-1}|e_j|, \quad i, j = 1, 2, 3. \tag{3.8}$$

3.2 Uniqueness in two dimensions

We next prove two technical lemmas to support the comparison theorem in two dimensions. In contrast to the case of one dimension, both the cases of Dirichlet and mixed Dirichlet/Neumann boundary conditions are handled together. This is due in part to the additional complexities of the mesh partition in two dimensions, particularly that a given vertex may belong to many triangles $T \in \mathcal{T}$, rather than only two. Because of this, we consider a test function supported on all triangles in the domain over which the difference of solutions is positive at any vertex. In Lemma 1, we establish the necessary estimates for Theorem 5 on any triangle where the difference of solutions, denoted w, is positive on exactly one vertex; in Lemma 2, we establish the analogous estimate for the case of two positive vertices.

Lemma 1 *Let Assumptions 1 and 2 hold, and let $u_1, u_2, w, v \in \mathcal{V}_{0,D}$, subordinate to some fixed $\mathcal{T}$ and Γ_D. Suppose there is some triangle $T \in \mathcal{T}$ for which $w(x)$ is positive at one vertex, and nonpositive at the other two. Assign the set of indices $\mathcal{N}_T = \{i, j, k\}$ to take the values $\{1, 2, 3\}$ so that $w(a_i) > 0$, and $0 \geq w(a_j) \geq w(a_k)$. Define the*

test function $v(x)$ by it's nodal values by $v(a_i) = 1$, and $v(a) = 0$ at all other mesh vertices $a \in \mathcal{D}$. Then the following inequalities hold.

$$\int_T \kappa(x, u_1) \nabla w^T \nabla v \geq (w(a_i) - w(a_k)) \frac{|e_i||e_k|}{4|T|} (k_\alpha \gamma_T c_T), \qquad (3.9)$$

and

$$\int_T (\kappa(x, u_2) - \kappa(x, u_1)) \nabla u_2^T \nabla v$$
$$\leq (w(a_i) - w(a_k)) \frac{|e_i||e_k|}{4|T|} \frac{7L_0}{6} \left(1 + \gamma_T^{-1}\right) \max_{i', j'} |u_2(a_{i'}) - u_2(a_{j'})|. \qquad (3.10)$$

Proof The test function v is a basis function of $\mathcal{V}_{0,D}$, namely $v = \varphi_i$, with respect to the local numbering on T. The function w may be expanded as a linear combination of basis functions φ_i, φ_j and φ_k. An integral over T on the left hand side of (3.27) satisfies

$$\int_T \kappa(x, u_1) \nabla w^T \nabla v = \sum_{n \in \mathcal{N}_T} w(a_n) \int_T \kappa(x, u_1) \nabla \varphi_n^T \nabla \varphi_i$$
$$= w(a_i) \int_T \kappa(x, u_1) \nabla \varphi_i^T \nabla \varphi_i$$
$$+ w(a_j) \int_T \kappa(x, u_1) \nabla (\varphi_j + \varphi_k)^T \nabla \varphi_i$$
$$+ (w(a_k) - w(a_j)) \int_T \kappa(x, u_1) \nabla \varphi_k^T \nabla \varphi_i. \qquad (3.11)$$

Applying the identity (3.5), the first two terms of (3.11) may be combined. The angle condition (3.1) and the choice $w(a_k) \leq w(a_j)$ assures the last term is non-negative. Together with the lower bound from (1.2) and integration by (3.7) this yields

$$\int_T \kappa(x, u_1) \nabla w^T \nabla v = \left(w(a_i) - w(a_j)\right) \int_T \kappa(x, u_1) \nabla \varphi_i^T \nabla \varphi_i$$
$$+ \left(w(a_k) - w(a_j)\right) \int_T \kappa(x, u_1) \nabla \varphi_k^T \nabla \varphi_i$$
$$\geq \frac{k_\alpha c_T}{4|T|} \left\{ (w(a_i) - w(a_j))|e_i|^2 + (w(a_j) - w(a_k))|e_i||e_k| \right\}. \qquad (3.12)$$

Controlling the minimum ratio of edges by γ_T, as in (3.8), inequality (3.12) reduces to the first result, (3.9).

To establish the second estimate, expand u_2 as a linear combination of basis functions. Again applying $v = \varphi_i$ allows

$$\int_T (\kappa(x, u_2) - \kappa(x, u_1))\nabla u_2^T \nabla v$$

$$= \sum_{n \in \mathcal{N}_T} u_2(a_n) \int_T (\kappa(x, u_2) - \kappa(x, u_1))\nabla \varphi_n^T \nabla \varphi_i$$

$$= u_2(a_i) \int_T (\kappa(x, u_2) - \kappa(x, u_1))\nabla \varphi_i^T \nabla \varphi_i$$

$$+ u_2(a_j) \int_T (\kappa(x, u_2) - \kappa(x, u_1))\nabla (\varphi_j + \varphi_k)^T \nabla \varphi_i$$

$$+ (u_2(a_k) - u_2(a_j)) \int_T (\kappa(x, u_2) - \kappa(x, u_1))\nabla \varphi_k^T \nabla \varphi_i. \qquad (3.13)$$

Applying the identity $\nabla(\varphi_j + \varphi_k) = -\nabla \varphi_i$, to (3.13), obtain

$$\int_T (\kappa(x, u_2) - \kappa(x, u_1))\nabla u_2^T \nabla v$$

$$\leq |u_2(a_i) - u_2(a_j)| \left| \int_T (\kappa(x, u_2) - \kappa(x, u_1))\nabla \varphi_i^T \nabla \varphi_i \right|$$

$$+ |u_2(a_j) - u_2(a_k)| \left| \int_T (\kappa(x, u_2) - \kappa(x, u_1))\nabla \varphi_k^T \nabla \varphi_i \right|. \qquad (3.14)$$

Evaluating the constant inner product between basis functions by (3.6) and applying the Lipschitz condition (1.3), inequality (3.14) reduces to

$$\int_T (\kappa(x, u_2) - \kappa(x, u_1))\nabla u_2^T \nabla v$$

$$\leq \frac{L_0(|e_i|^2 + |e_i||e_k|)}{4|T|^2} \max_{i',j'} |u_2(a_{i'}) - u_2(a_{j'})| \int_T |w|$$

$$\leq \frac{L_0(|e_i||e_k|(1 + \gamma_T^{-1}))}{4|T|^2} \max_{i',j'} |u_2(a_{i'}) - u_2(a_{j'})| \int_T |w|, \qquad (3.15)$$

where the last inequality follows from (3.8).

To bound the integral of $|w|$, consider the following decomposition making use of $\varphi_i + \varphi_j + \varphi_k = 1$, $w(a_i) \geq w(a_j)$ and $w(a_j) \geq w(a_k)$.

$$|w| = |w(a_i)\varphi_i + w(a_j)\varphi_j + w(a_k)\varphi_k|$$

$$= |(w(a_i) - w(a_j))\varphi_i + (w(a_k) - w(a_j))\varphi_k + w(a_j)(\varphi_i + \varphi_j + \varphi_k)|$$

$$\leq (w(a_i) - w(a_j))\varphi_i + (w(a_j) - w(a_k))\varphi_k + (w(a_j) - w(a_k))$$

$$= (w(a_i) - w(a_j))\varphi_i + (w(a_j) - w(a_k))(1 + \varphi_k).$$

Applying $\int_T \varphi_i = \int_T \varphi_k = |T|/6$, we find

$$
\begin{aligned}
\int_T |w| &\leq (w(a_i) - w(a_j)) \frac{|T|}{6} + (w(a_j) - w(a_k)) \frac{7|T|}{6} \\
&\leq (w(a_i) - w(a_k)) \frac{7|T|}{6}.
\end{aligned} \tag{3.16}
$$

Putting together (3.15) with (3.16) yields the desired estimate (3.10), which is suitable for comparison with the first result, (3.9). □

Remark 4 The first result (3.9) of Lemma 1 (and the analogous result of the next lemma) depends both on the positivity of $\kappa(x, u_1) \nabla \varphi_i^T \nabla \varphi_i$ and the strict negativity of $\kappa(x, u_1) \nabla \varphi_k^T \nabla \varphi_i$, for $k \neq i$ to establish (3.12) from (3.11). If the scalar coefficient $\kappa(\cdot, s)$ were replaced by a 2×2 symmetric positive definite tensor coefficient, the positivity of $\nabla \varphi_i^T \kappa(x, u_1) \nabla \varphi_i$ would still hold. However, the essential negativity of the mixed product $\nabla \varphi_k^T \kappa(x, u_1) \nabla \varphi_i$ would not then be enforced. For this reason, the current results are restricted to the case of a scalar coefficient. The extension of the current results to the anisotropic problem with a tensor coefficient $\kappa(x, s)$ is considered by the authors to be an interested problem.

We next consider the analogous estimates in the case where the function w is positive at two vertices. Notably, the results in this second case closely resemble the first. There are however minor differences in calculations to obtain each estimate. These two cases together are sufficient to prove the comparison theorem that follows.

Lemma 2 *Let Assumptions 1 and 2 hold, and let $u_1, u_2, w, v \in \mathcal{V}_{0,D}$, subordinate to some fixed $\mathcal{T}$ and Γ_D. Suppose there is some triangle $T \in \mathcal{T}$ for which $w(x)$ is positive at two vertices, and nonpositive at the third. Assign the set of indices $\mathcal{N}_T = \{i, j, k\}$ to take the values $\{1, 2, 3\}$ so that $w(a_i) \geq w(a_j) > 0$, and $0 \geq w(a_k)$. Define the test function $v(x)$ by it's nodal values by $v(a_i) = v(a_j) = 1$, and $v(a) = 0$ at all other mesh vertices $a \in \mathcal{D}$. Then the following inequalities hold.*

$$
\int_T \kappa(x, u_1) \nabla w^T \nabla v \geq (w(a_i) - w(a_k)) \frac{|e_i||e_k|}{4|T|} (k_\alpha \gamma_T c_T), \tag{3.17}
$$

and

$$
\begin{aligned}
&\int_T (\kappa(x, u_2) - \kappa(x, u_1)) \nabla u_2^T \nabla v \\
&\quad \leq (w(a_i) - w(a_k)) \frac{|e_i||e_k|}{4|T|} \frac{7L_0}{6} \left(1 + \gamma_T^{-1}\right) \max_{i', j'} |u_2(a_{i'}) - u_2(a_{j'})|. \tag{3.18}
\end{aligned}
$$

Proof In this case the test function v is given by $\varphi_i + \varphi_j$. The identity (3.5) then implies $\nabla v = -\nabla \varphi_k$, yielding

$$
\begin{aligned}
\nabla w^T \nabla v &= \sum_{n \in \mathcal{N}_T} -w(a_n) \nabla \varphi_n^T \nabla \varphi_k \\
&= -(w(a_i) - w(a_j)) \nabla \varphi_i^T \nabla \varphi_k + (w(a_j) - w(a_k)) \nabla \varphi_k^T \nabla \varphi_k, \quad (3.19)
\end{aligned}
$$

where the angle condition (3.1) implies the non-negativity of the first term, and the positivity of the second is clear from the local ordering of the nodes. Applying the positivity of $\kappa(x, u_1)$ from (1.2), the expansion (3.19) and integration over T using (3.7) allows

$$
\int_T \kappa(x, u_1) \nabla w^T \nabla v \geq \frac{k_\alpha c_T}{4|T|} \left\{ (w(a_i) - w(a_j)) |e_i| |e_k| + (w(a_j) - w(a_k)) |e_k|^2 \right\}.
$$
$$(3.20)$$

As in the previous lemma, using (3.8) to bound the minimum ratio of edges in (3.20) in terms of γ_T, we have the result (3.17).

For the second estimate, first expand u_2 as a linear combination of basis functions, and apply $\nabla v = -\nabla \varphi_k$.

$$
\begin{aligned}
-\nabla u_2^T \nabla v &= \sum_{n \in \mathcal{N}_T} u_2(a_n) \nabla \varphi_n^T \nabla \varphi_k \\
&= (u_2(a_i) - u_2(a_j)) \nabla \varphi_i^T \nabla \varphi_k + (u_2(a_k) - u_2(a_j)) \nabla \varphi_k^T \nabla \varphi_k. \quad (3.21)
\end{aligned}
$$

Then, using the expansion of (3.21), apply the Lipschitz condition (1.3) to bound the integral over T.

$$
\begin{aligned}
\int_T (\kappa(x, u_2) &- \kappa(x, u_1)) \nabla u_2^T \nabla v \\
&= (u_2(a_i) - u_2(a_j)) \int (\kappa(x, u_1) - \kappa(x, u_2)) \nabla \varphi_i^T \nabla \varphi_k \\
&\quad + (u_2(a_k) - u_2(a_j)) \int (\kappa(x, u_1) - \kappa(x, u_2)) \nabla \varphi_k^T \nabla \varphi_k \\
&\leq |u_2(a_i) - u_2(a_j)| \frac{L_0 |e_i| |e_k|}{4|T|} \int_T |w| \\
&\quad + |u_2(a_k) - u_2(a_j)| \frac{L_0 |e_k|^2}{4|T|} \int_T |w| \\
&\leq \frac{L_0(|e_k|^2 + |e_i| |e_k|)}{4|T|^2} \max_{i', j'} |u_2(a_{i'}) - u_2(a_{j'})| \int_T |w|. \quad (3.22)
\end{aligned}
$$

The integral of $|w|$ over T is also bounded by (3.16). Putting this together with (3.22), and applying $|e_k| \leq \gamma_T^{-1} |e_i|$ from (3.8), allows the second part of the result, (3.18), which is in suitable form for comparison with the first part, (3.17). □

With the technical lemmas in hand, we are now ready to prove a comparison theorem in two dimensions. The corollary to this theorem is uniqueness of the discrete solution under mixed Dirichlet/Neumann boundary conditions.

Theorem 5 (Comparison Theorem in 2D) *Let Assumptions* 1 *and* 2 *hold, let* f_1 *and* $f_2 \in L_2(\Omega)$ *satisfy* $f_1 \leq f_2$, *a.e.* $x \in \Omega$; *and, let* $\psi \in H^{1/2}(\Gamma_N)$. *Let* $u_i \in \mathcal{V}_{0,D}$, $i = 1, 2$, *denote the solution to the respective weak forms of* (1.1), *subject to the following boundary conditions.*

$$u = 0 \text{ on } \Gamma_D, \text{ and } \kappa(x, u)\nabla u^T n = \psi \text{ on } \Gamma_N, \tag{3.23}$$

with n the outward-facing normal on Γ_N. *In particular*

$$(\kappa(x, u_i)\nabla u_i, \nabla v) = (f_i, v) + \int_{\Gamma_N} \psi v \text{ for all } v \in \mathcal{V}_0, \quad i = 1, 2. \tag{3.24}$$

Then under the condition

$$\max_{T \in \mathcal{T}} \left(c_T^{-1} \gamma_T^{-1}(1 + \gamma_T^{-1}) \max_{i', j' \in \{1, 2, 3\}} |u_2(a_{i'}) - u_2(a_{j'})| \right) < \frac{6k_\alpha}{7L_0}, \tag{3.25}$$

it holds that $u_1 \leq u_2$.

Proof By (3.24) and subtraction

$$(\kappa(u_1)\nabla w, \nabla v) = ((\kappa(u_2) - \kappa(u_1))\nabla u_2, \nabla v) + (f_1 - f_2, v), \tag{3.26}$$

where $w = u_1 - u_2$. We prove the theorem by contradiction. Assume there is at least one node $d \in \mathcal{D}$ on which $u_1(d) > u_2(d)$, i.e. $w(d) > 0$. Then the test function v may be defined by its nodal values: one at nodes where w is positive, and zero elsewhere.

$$v(d) = \begin{cases} 1, & \text{for } d \in \mathcal{D} \text{ with } w(d) > 0, \\ 0, & \text{for } d \in \mathcal{D} \text{ with } w(d) \leq 0. \end{cases}$$

Then w is zero on Γ_D so that $v \in \mathcal{V}_{0,D}$. For each triangle $T \in \mathcal{T}$, $v(a_i) = 0$, on either one, two, or all three vertices a_i, $i = 1, 2, 3$. Define the following subsets of partition $\mathcal{T}$ with respect to this property.

$$\mathcal{T}_1 = \{T \in \mathcal{T} \mid v(a_i) = 1, \text{ on exactly one vertex } a_i \text{ of } T\},$$
$$\mathcal{T}_2 = \{T \in \mathcal{T} \mid v(a_i) = 1, \text{ on exactly two vertices } a_i, a_j, i \neq j, \text{ of } T\}.$$

Due to the fact that $\Gamma_D \neq \emptyset$, it is clear that $\mathcal{T}_1 \cup \mathcal{T}_2$ is nonempty. As $\nabla w = 0$, over $\mathcal{T} \setminus \{\mathcal{T}_1 \cup \mathcal{T}_2\}$, Eq. (3.26) can then be written in terms of elementwise integration over partitions $\mathcal{T}_1$ and $\mathcal{T}_2$.

$$\sum_{T \in (\mathcal{T}_1 \cup \mathcal{T}_2)} \int_T \kappa(u_1) \nabla w^T \nabla v = \sum_{T \in (\mathcal{T}_1 \cup \mathcal{T}_2)} \int_T (\kappa(u_2) - \kappa(u_1)) \nabla u_2^T \nabla v + (f_1 - f_2, v).$$

(3.27)

Bounds on the left and right hand sides of (3.26) have been worked out by elementwise integration, first for $T \in \mathcal{T}_1$, in Lemma 1, then for $T \in \mathcal{T}_2$ in Lemma 2. Putting together the results of these two lemmas, (3.9), (3.10), (3.17) and (3.18), into (3.27), we have

$$\sum_{T \in (\mathcal{T}_1 \cup \mathcal{T}_2)} (w(a_i) - w(a_k)) \frac{|e_i||e_k|}{4|T|}$$
$$\times \left\{ k_\alpha \gamma_T c_T - \frac{7L_0}{6} \left(1 + \gamma_T^{-1}\right) \max_{i',j'} |u_2(a_{i'}) - u_2(a_{j'})| \right\}$$
$$\le (f_1 - f_2, v) \le 0.$$

(3.28)

As the terms to the left of the brackets on each line of (3.28) are strictly positive, a contradiction is attained under the condition (3.25). This establishes the comparison result. $\square$

It is emphasized that condition (3.25) is a local condition, and it differs from the one dimensional analogue (2.5), by a factor related to the mesh geometry and in particular to smallest and largest angles. As in one dimension, the two dimensional comparison theorem leads to the main result.

Corollary 4 (Uniqueness of the FE Solution in 2D) *Let Assumptions 1 and 2 hold, let $f \in L_2(\Omega)$ and $\psi \in H^{1/2}(\Gamma_N)$. Let $u \in \mathcal{V}_{0,D}$, denote the solution to the weak form of (1.1), subject to the following boundary conditions.*

$$u = 0 \text{ on } \Gamma_D, \text{ and } \kappa(x, u) \nabla u^T n = \psi \text{ on } \Gamma_N.$$

In particular

$$(\kappa(x, u) \nabla u, \nabla v) = (f, v) + \int_{\Gamma_N} \psi v \text{ for all } v \in \mathcal{V}_{0,D}.$$

(3.29)

Then under the condition

$$\max_{T \in \mathcal{T}} \left(c_T^{-1} \gamma_T^{-1} (1 + \gamma_T^{-1}) \max_{i',j' \in \{1,2,3\}} |u(a_{i'}) - u(a_{j'})| \right) < \frac{6k_\alpha}{7L_0},$$

(3.30)

(cf., (3.25)), the linear FE solution u is the unique solution to (3.29).

Proof Let $f_1 = f_2 = f$ in Theorem 5. Then, let u_1 and u_2 be two respective solutions to problem (3.24). Applying the results of Theorem 5 yields both $u_1 \le u_2$ and $u_2 \le u_1$. Notably, the argument again does not depend on either local or global meshsizes, only the variation in the discrete piecewise linear solution over each element. This can be controlled in an adaptive algorithm by monitoring the nodal values over each element and bisecting triangles where the condition for uniqueness is violated. $\square$

Remark 5 A sufficient condition for uniqueness is then given by

$$\max_{T \in \mathcal{T}} \max_{i', j' \in \{1,2,3\}} |u(a_{i'}) - u(a_{j'})| < \frac{6k_\alpha}{7L_0} \left(\frac{s_{min}^2 c_{min}}{1 + s_{min}} \right), \tag{3.31}$$

where $s_{min} = \sin(t_{min})$, and $c_{min} = \min_{T \in \mathcal{T}} c_T$, by applying the global bound $s_{min} \leq \gamma_T \leq 1$. While this condition appears easier to check if used in an adaptive algorithm, the condition (3.30) is preferred. In particular, the locally angle-dependent version better characterizes the condition for uniqueness based on both the maximum variation of solution u over an element T, together with the condition of that triangle. For instance, if the solution varies rapidly where the mesh is very regular, and the solution varies slowly where the mesh is poorly conditioned in the sense of containing small angles, then (3.31) gives an overly pessimistic requirement.

3.3 Addressing the 2D counterexample of [4]

We refer the reader to [4] for the full details of a 2D Dirichlet problem for which there are at least two discrete solutions on a given mesh. We demonstrate here the uniqueness condition (3.30) is not satisfied on the given mesh, detecting the possibility of nonuniqueness. The example and present analysis are similar to the one dimensional analogue.

In this example, Ω is the interior of the equilateral triangle with sides of length 1. The mesh $\mathcal{T}_h$ is a uniform partition, for which each element is an equilateral triangle with sides length h. The diffusion coefficient $\kappa(x, z)$ is constructed on $\mathcal{T}_h$ given a function $u \in \mathcal{V}_{0,D} \subset H_0^1(\Omega)$ that satisfies $u(x) = 0$ for $x \in \Gamma_D = \partial\Omega$, and $u(x) > 0$ for $x \in \Omega$.

The coefficient $\kappa(x, z)$ is constructed in terms of a $C^\infty(\Omega)$ function $\phi(x)$ that satisfies

$$\phi = 1 \text{ on } \Gamma, \quad |\nabla\phi| = 0 \text{ on } \Gamma, \quad \frac{1}{4} \leq \phi \leq 1, \quad \text{and} \quad \int_\Omega \phi = \frac{\sqrt{3}}{8}. \tag{3.32}$$

Moreover, ϕ is assumed invariant to the isometries of Ω and defined locally in each element T by change of variables to a reference element coinciding in shape with the global domain Ω. Local coordinates in each element T are described by $y_T : T \to \Omega$, an invertible affine maps for each element $T \in \mathcal{T}$ that assigns to each point $x \in T$ the point $y_T(x) \in \Omega$ with the same barycentric coordinates. It is outlined that

$$\frac{\partial}{\partial z}\kappa(x, z) = \frac{\phi(y_T(x)) - 1}{u(x)}, \quad u(x) \leq z(x) \leq 2u(x). \tag{3.33}$$

Now consider an element $T_1 \in \mathcal{T}$ for which two vertices lie on $\partial\Omega$, and one vertex lies in the interior of Ω. If the problem has any degrees of freedom, such an element must exist. Denote the interior vertex as a_1. Then $u_1 := u(a_1) > 0$. For clarity of presentation, assign the coordinate system (t_1, t_2) to the reference domain so that Ω

has vertices $\{(1/2, \sqrt{3}/2), (0, 0), (1, 0)\}$. Let y_T map vertex a_1 of T_1 to $(1/2, \sqrt{3}/2)$. Then $\phi(y_T(x)) = \phi((t_1, t_2))$ for each $x \in T$ and some $(t_1, t_2) \in \Omega$. Define $\check{\phi}(t_2) := \phi((1/2, t_2))$. Then because $u(x)$ is an affine function in T_1, and $u(a_2) = u(a_3) = 0$, $u(x)$ on T_1 is dependent only on reference coordinate t_2, namely, $u(y_T^{-1}(\cdot, t_2)) = (2/\sqrt{3})t_2 u_1$. Apply this now to (3.33) in triangle T_1, restricted to $x = y_T^{-1}((1/2, t_2))$, for $0 \leq t_2 \leq \sqrt{3}/2$.

$$\frac{\phi((1/2, t_2)) - 1}{u(y_T^{-1}((1/2, t_2)))} = \frac{\phi((1/2, t_2)) - \phi((1/2, 0))}{(2/\sqrt{3})t_2 u_1}$$

$$= \left(\frac{\sqrt{3}}{2u_1}\right)\frac{\check{\phi}(t_2) - \check{\phi}(0)}{t_2}. \tag{3.34}$$

The definition of ϕ in (3.32) with $k = 1/4$ makes sense for $0 < k < 1/2$, so we address k in this range. For any $0 < k < 1/2$, it suffices to consider the graph of $\check{\phi}$ on the interval $0 \leq t_2 \leq \sqrt{3}/2$. The slope of the secant line given in the right hand side term of (3.34), $(\check{\phi}(t_2) - \check{\phi}(0))/t_2$, must somewhere be greater in magnitude of the slope of the line connecting $\psi(0) = 1$ to $\psi(\sqrt{3}/2) = k$, that is $2(1 - k)/\sqrt{3}$. Then the right side of (3.34) must be greater than $(1 - k)/u_1$ for some t_2. From (3.33), this shows the Lipschitz constant $L_0 > (1 - k)/u_1$, for $k_\alpha = k$. For the equilateral mesh, $\gamma_T = 1$ and $c_T = \cos(\pi/3) = 1/2$ for all $T \in \mathcal{T}$, and condition (3.30) then requires

$$u_1 < \frac{3k}{14(1 - k)}u_1 < u_1. \tag{3.35}$$

This shows that condition (3.30) detects the possibility of nonuniqueness.

4 Conclusion

We demonstrated here uniqueness of the continuous piecewise linear finite element solution for a class of nonmonotone quasilinear elliptic problems under either Dirichlet or mixed Dirichlet/Neumann conditions. The main innovation is establishing these results without relying on the requirement of a globally fine mesh. It has been observed in practice that this condition while sufficient for the guarantee of well-posedness of the discrete problem, appears not to be necessary. We established in this analysis that it is sufficient for the variance of the solution over each element to be bounded by a multiple of the ratio of the ellipticity and Lipschitz constants for a given problem. This result is important in the analysis of adaptive methods for this class of problems, as uniqueness of the solution can now be assured based on local and computable criteria.

References

1. Abdulle, A., Vilmart, G.: A priori error estimates for finite element methods with numerical quadrature for nonmonotone nonlinear elliptic problems. Numer. Math. **121**(3), 397–431 (2012)
2. Alvino, A., Betta, M.F., Mercaldo, A.: Comparison principle for some classes of nonlinear elliptic equations. J. Differ. Equ. **249**(12), 3279–3290 (2010)
3. André, N., Chipot, M.: A remark on uniqueness for quasilinear elliptic equations. Banach Center Publ. **33**(1), 9–18 (1996)
4. André, N., Chipot, M.: Uniqueness and nonuniqueness for the approximation of quasilinear elliptic equations. SIAM J. Numer. Anal. **33**(5), 1981–1994 (1996)
5. Bi, C., Ginting, V.: A posteriori error estimates of discontinuous Galerkin method for nonmonotone quasi-linear elliptic problems. J. Sci. Comput. **55**(3), 659–687 (2013)
6. Douglas, J., Dupont, T.: A Galerkin method for a nonlinear Dirichlet problem. Math. Comput. **131**, 689–696 (1975)
7. Douglas, J., Dupont, T., Serrin, J.: Uniqueness and comparison theorems for nonlinear elliptic equations in divergence form. Arch. Ration. Mech. Anal. **42**(3), 157–168 (1971)
8. Gilbarg, D., Trudinger, N.S.: Elliptic partial differential equations of second order. Grundlehren der mathematischen Wissenschaften: 224. Springer, Berlin (1983)
9. Hlaváček, I., Křížek, M., Malý, J.: On Galerkin approximations of a quasilinear nonpotential elliptic problem of a nonmonotone type. J. Math. Anal. Appl. **184**(1), 168–189 (1994)
10. Karátson, J., Korotov, S.: Discrete maximum principles for finite element solutions of nonlinear elliptic problems with mixed boundary conditions. Numer. Math. **99**(4), 669–698 (2005)
11. Karátson, J., Korotov, S., Křížek, M.: On discrete maximum principles for nonlinear elliptic problems. Math. Comput. Simul. **76**(1–3), 99–108 (2007)
12. Pollock, S.: A regularized Newton-like method for nonlinear PDE. Numer. Funct. Anal. Optim. **36**(11), 1493–1511 (2015)
13. Pollock, S.: An improved method for solving quasilinear convection diffusion problems. SIAM J. Sci. Comput. **38**(2), A1121–A1145 (2016)
14. Pollock, S.: Stabilized and inexact adaptive methods for capturing internal layers in quasilinear PDE. J. Comput. Appl. Math. **308**, 243–262 (2016)
15. Wang, J., Zhang, R.: Maximum principles for $P1$-conforming finite element approximations of quasilinear second order elliptic equations. SIAM J. Numer. Anal. **50**(2), 626–642 (2012)