ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ - ΤΜΗΜΑ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ
Archimedes Center for Modeling, Analysis & Computation
UNIVERSITY OF CRETE - DEPARTMENT OF APPLIED MATHEMATICS
Archimedes Center for Modeling, Analysis & Computation

Consistent Discretizations for Vanishing Regularization Solutions to Image Processing Problems

Theodoros Katsaounis and Stephen Keeling and Michael Plexousakis

DRAFT

Consistent Discretizations for Vanishing Regularization Solutions to Image Processing Problems

Theodoros D. Katsaounis[1], Stephen L. Keeling[2] and Michaelos Plexousakis[3]

Abstract. A model problem is used to represent a typical image processing problem of reconstructing an unknown in the face of incomplete data. A consistent discretization for a vanishing regularization solution is defined so that, in the absence of noise, limits first with respect to regularization and then with respect to grid refinement agree with a continuum counterpart defined in terms of a saddle point formulation. It is proved and demonstrated computationally for an artificial example and for a realistic example with magnetic resonance images that a mixed finite element discretization is consistent in the sense defined here. On the other hand, it is demonstrated computationally that a standard finite element discretization is not consistent, and the reason for the inconsistency is suggested in terms of theoretical and computational evidence.

Keywords: consistent discretization, vanishing regularization, saddle point problem, mixed finite elements

1 Introduction

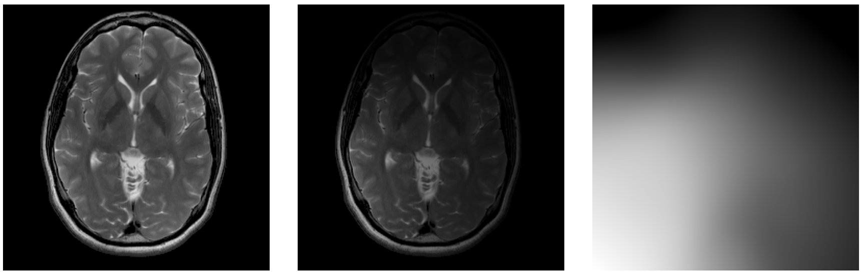
Many image processing problems involve reconstructing an unknown in the face of incomplete data [13]. A reconstruction may be formulated in terms of an ill-posed problem which is stabilized through a regularized minimization of the problem residual. To counter the effect of noise in measured data, one naturally applies nontrivial regularization. For a fixed noise level and for fixed regularization, standard discretizations of the associated optimality system have well known convergence characteristics with respect to grid refinement [3]. Also, it is well known in the infinite dimensional setting that a regularized solution converges to a vanishing regularization solution provided one coordinates the rates at which regularization and noise are simultaneously reduced [14]. As discretization of data can be seen as a type of noise, it is a significant finding in the present work that a vanishing regularization solution to certain discretized formulations may in fact converge properly with grid refinement, but this convergence does not hold in general. Numerical methods which do not exhibit this sensitivity to vanishing regularization are regarded in this work as *consistent discretizations of vanishing regularization solutions*. The purpose of this paper is to demonstrate a consistent and an inconsistent discretization of a vanishing regularization solution to a model problem. Also, convergence of the consistent approach with respect to grid refinement is proved, while the reason for the failure of the inconsistent approach is suggested in terms of theoretical and computational evidence.

Based upon the regularized imaging reconstructions and vanishing regularization solutions presented for example in [13] and [11], the concept of consistent discretization of a vanishing regularization solution is demonstrated here for the model problem which is illustrated in Fig. 1. Here it is required to estimate a smooth modulation field u from a given unmodulated image κ and a given modulated image $\tilde{u}$ satisfying $\kappa u \approx \tilde{u}$. It is assumed that the data $\tilde{u}$ and κ have

[1]Department of Applied Mathematics, University of Crete, Heraklion, 71409 Crete, Greece; email: thodoros@tem.uoc.gr, tel: +30-2810-393723, fax: +30-2810-393701.

[2]Institut für Mathematik und Wissenschaftliches Rechnen, Karl-Franzens-Universität Graz, Heinrichstraße 36, 8010 Graz, Austria; email: stephen.keeling@uni-graz.at, tel: +43-316-3805165, fax: +43-316-3809827; supported by the Austrian Science Fund *Fond zur Förderung der Wissenschaftlichen Forschung* (FWF) under grant SFB F032 ("Mathematical Optimization and Applications in Biomedical Sciences" http://math.uni-graz.at/mobis).

[3]Department of Applied Mathematics, University of Crete, 71409 Heraklion, Greece, email: plex@tem.uoc.gr, tel: +30-2810-393709, fax: +30-2810-393701.



(a) unmodulated κ **(b)** modulated $\tilde{u}$ **(c)** modulation u

Figure 1: (a) An unmodulated image κ and (b) a modulated image $\tilde{u}$ are used to estimate (c) the smooth modulation u.

the same support

$$S_{\rm d} \subsetneq \Omega, \quad |\Omega| > |S_{\rm d}| = |S_{\rm d}^\circ| > 0 \tag{1}$$

with characteristic function,

$$\chi_{\rm d} = \begin{cases} 1, & S_{\rm d} \\ 0, & \text{else} \end{cases} \tag{2}$$

so the ill-posed computation of u may be stabilized through regularization by minimizing a functional of the form

$$J^\epsilon(u) = \frac{1}{2}\int_\Omega \left[|\kappa u - \tilde{u}|^2 + \epsilon|\nabla^2 u|^2\right] \tag{3}$$

As the focus is placed here on vanishing regularization solutions in the absence of noise, it is assumed that there exist $\tilde{u}_{\rm d}, u_{\rm d} \in H^2(\Omega)$ and $\kappa_{\rm d} \in W^{2,\infty}(\Omega)$ with $\kappa_{\rm d}^{-1} \in W^{2,\infty}(\Omega)$ which satisfy

$$\chi_{\rm d}\kappa_{\rm d} = \kappa, \quad \chi_{\rm d}\tilde{u}_{\rm d} = \tilde{u}, \quad \kappa_{\rm d} u_{\rm d} = \tilde{u}_{\rm d}. \tag{4}$$

Otherwise, it is assumed that the data satisfy

$$0 < \kappa_0 \le \kappa \le \kappa_1, \quad 0 < \tilde{u}_0 \le \tilde{u} \le \tilde{u}_1 \quad \text{in} \quad S_{\rm d}. \tag{5}$$

The measured data κ_h and $\tilde{u}_h$ are assumed to be given as pixelwise constant approximations of κ and $\tilde{u}$, respectively.

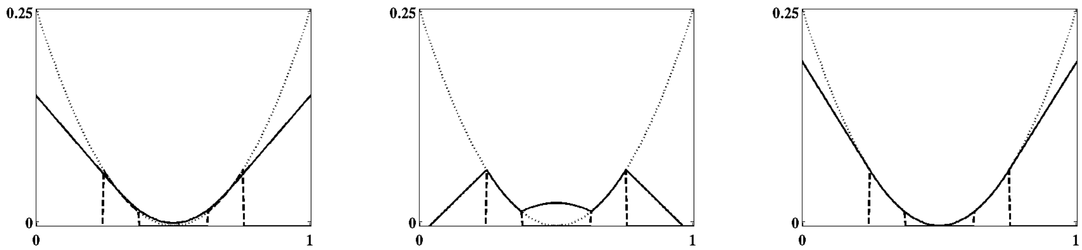
In the next sections, it is seen that the unique minimizer $u^\epsilon \in H^4(\Omega)$ of the functional J^ϵ in (3) is characterized by the strong necessary optimality condition,

$$\epsilon\Delta^2 u + \kappa^2 u = \kappa\tilde{u} \quad \text{in} \quad \Omega \tag{6}$$

where the domain of Δ^2 is understood to impose natural boundary conditions. Also, u^ϵ converges weakly in $H^2(\Omega)$ to a unique limit $u^\star$ as $\epsilon \to 0$. The unique limit $u^\star$ is termed the vanishing regularization solution and is shown below to be characterized directly as a weak solution to the saddle point problem:

$$\left\{ \begin{array}{rcl} (1-\chi_{\rm d})\lambda + \kappa^2 u &=& \kappa\tilde{u} \\ \kappa^2\lambda + \Delta^2 u &=& 0 \end{array} \right. \quad \text{in} \quad \Omega \tag{7}$$

In the numerical investigations below, (6) is discretized first with a standard finite element method in which smooth quadratic splines are applied, and second with a mixed finite element method in which smooth quadratic splines are used for the primal variable and piecewise constants are used for the dual variable. Also, the data are approximated in the usual way as pixelwise constant. Representative results are shown in Fig. 2. Here, $S_{\rm d} = [\frac{2}{8}, \frac{3}{8}] \cup [\frac{5}{8}, \frac{6}{8}]$, $\kappa = \chi_{\rm d}$ and the dashed curve is $\tilde{u} = \chi_{\rm d} u_{\rm q}$, where the quadratic curve $u_{\rm q}(x) = (x - \frac{1}{2})^2$ is shown dotted.



(a) $\epsilon = \mathcal{O}(1)$, standard **(b)** $\epsilon \ll 1$, standard **(c)** $\epsilon \ll 1$, mixed

Figure 2: Results obtained with a standard finite element method are shown (a) for $\epsilon = \mathcal{O}(1)$ and (b) for $\epsilon \ll 1$. Also, results for a mixed finite element method are shown (c) for $\epsilon \ll 1$. Here, $\kappa = \chi_\text{d}$ and the dashed curve is $\tilde{u} = \chi_\text{d} u_\text{q}$, where the quadratic curve u_q is shown dotted. The solid curves are numerical solutions to (6). For $\epsilon \ll 1$, only the mixed finite element solution accurately approximates the solution to (6).

With χ_0 being the characteristic function for $[\frac{1}{4}, \frac{3}{4}]$, the vanishing regularization solution satisfying (7) is given by $u^\star(x) = (x - \frac{1}{2})^2 \chi_0 + (\frac{1}{2}|x - \frac{1}{2}| - \frac{1}{16})(1 - \chi_0)$. Note that $u^\star$ agrees with u_q in the convex hull of S_d and is linear in $\Omega \backslash S_\text{d}$. The solid curves are numerical solutions to (6). For $\epsilon = \mathcal{O}(1)$, all numerical solutions agree qualitatively with the standard finite element result shown in Fig. 2a, which is noticably smoothed in relation to the vanishing regularization solution. For $\epsilon \ll 1$, the mixed finite element result in Fig. 2c accurately approximates $u^\star$, while the standard finite element result in Fig. 2b departs tremendously from the correct vanishing regularization solution. In the next sections, the accuracy of the mixed finite element approach described here computationally is proved theoretically, and the inaccuracy of the standard finite element is suggested in terms of the convergence framework employed.

Note that all the numerical methods considered provide results such as seen in Fig. 2c when the data κ and $\tilde{u}$ are integrated exactly without quadrature error. However, measured images are universally stored using an array of values representing pixelwise averaged intensities, as assumed for κ_h and $\tilde{u}_h$. Therefore, a significant result of this work is that there exist numerical methods which accommodate the usual data format of image processing while providing consistent discretizations of vanishing regularization solutions. See also [6], [9] and [10] for error estimates for elliptic problems in the presence of quadrature errors.

2 Optimality and Saddle Point Conditions

Since the functional J^ϵ of (3) is quadratic in u, its variational properties are given in a rather staightforward way in terms of forms which will be used throughout this work. The Gâteaux derivative of J^ϵ is obtained as $\delta J^\epsilon / \delta u(u; v) = \epsilon a(u, v) + b(u, v) - d(v)$, where a, b and d are given by

$$a(u, v) = \int_\Omega \nabla^2 u : \nabla^2 v \tag{8}$$

$$b(u, v) = \int_\Omega \kappa^2 u v \tag{9}$$

$$d(v) = \int_\Omega \kappa \tilde{u} v \tag{10}$$

The necessary optimality condition for the minimization of J^ϵ is given by

$$\epsilon a(u, v) + b(u, v) = d(v), \quad \forall v \in H^2(\Omega) \tag{11}$$

whose solvability is guaranteed by Theorem 1 below. For this, the kernel of the form a is identified as the set of linear functions,

$$\mathcal{L} = \{\alpha_0 + \boldsymbol{\alpha} \cdot \boldsymbol{x} : \alpha_0 \in \mathbb{R}, \boldsymbol{\alpha} \in \mathbb{R}^n\}.$$

Lemma 1 *There exist constants c_1 and c_2 such that*

$$c_1\|v\|^2_{H^2(\Omega)} \le a(v,v) + b(v,v) \le c_2\|v\|^2_{H^2(\Omega)}, \quad \forall v \in H^2(\Omega). \tag{12}$$

Proof: Boundedness of the forms a and b follows readily with (5). Coercivity is established using a Poincaré argument since (1) and (5) imply that b coerces the kernel $\mathcal{L}$ of a. ∎

Theorem 1 *There exists a unique $u^\epsilon \in H^2(\Omega)$ satisfying (11) which is the unique minimizer for J^ϵ in (3).*

Proof: Boundedness of the form d follows with (5). By Lemma 1, the left side of (11) is bounded in terms of and coerces the $H^2(\Omega)$-norm. By the Lax-Milgram Lemma [3], (11) is uniquely solvable. For an arbitrary $v \in H^2(\Omega)$, it follows with (11) and (12) that $2[J^\epsilon(u^\epsilon + v) - J^\epsilon(u^\epsilon)] = \epsilon a(v,v) + b(v,v) \ge c_1\|v\|^2_{H^2(\Omega)}$. Thus, $J^\epsilon(u^\epsilon + v) > J^\epsilon(u^\epsilon)$ for every nontrivial $v \in H^2(\Omega)$. ∎

The following vanishing regularization result can be established by verifying conditions of results in [14] in the more general setting, but the result can be established more directly as follows.

Theorem 2 *The solutions $\{u^\epsilon\}$ to (11) converge weakly in $H^2(\Omega)$ to a unique limit $u^\star$ as $\epsilon \to 0$.*

Proof: It will be first be shown that $\{u^\epsilon\}$ is bounded in $H^2(\Omega)$ independently of ϵ. By (4), u_{d} satisfies $b(v, u_{\mathrm{d}}) = d(v)$, $\forall v \in H^2(\Omega)$, which can be subtracted from (11) with $v = w^\epsilon = u^\epsilon - u_{\mathrm{d}}$ to obtain $\epsilon a(w^\epsilon, w^\epsilon) + b(w^\epsilon, w^\epsilon) = -\epsilon a(u_{\mathrm{d}}, w^\epsilon)$. It follows that $a(w^\epsilon, w^\epsilon) \le |a(u_{\mathrm{d}}, w^\epsilon)|$ or $a(w^\epsilon, w^\epsilon) \le a(u_{\mathrm{d}}, u_{\mathrm{d}})$ and that $b(w^\epsilon, w^\epsilon) \le \epsilon|a(u_{\mathrm{d}}, w^\epsilon)|$ or $b(w^\epsilon, w^\epsilon) \le \frac{1}{2}a(u_{\mathrm{d}}, u_{\mathrm{d}}) + \frac{1}{2}a(w^\epsilon, w^\epsilon)$ for $\epsilon \le 1$. Hence, $\frac{1}{2}a(w^\epsilon, w^\epsilon) + b(w^\epsilon, w^\epsilon) \le \frac{3}{2}a(u_{\mathrm{d}}, u_{\mathrm{d}})$, and with (12), w^ϵ is bounded in $H^2(\Omega)$. Thus, u^ϵ is bounded in $H^2(\Omega)$ in terms of w^ϵ and u_{d}. Therefore, as $\epsilon \to 0$, there is a subsequence $\{u^{\epsilon_l}\}$ which converges weakly in $H^2(\Omega)$ to a limit $u^\star \in H^2(\Omega)$.

If there were another weak limit, $\hat{u}^{\hat{\epsilon}_l} \rightharpoonup \hat{u}^\star$, then according to (11), $w^\star = u^\star - \hat{u}^\star$ would satisfy $b(w^\star, w^\star) \overset{l\to\infty}{\longleftarrow} b(w^\star, u^{\epsilon_l} - u^{\hat{\epsilon}_l}) = \hat{\epsilon}_l a(u^{\hat{\epsilon}_l}, w^\star) - \epsilon_l a(u^{\epsilon_l}, w^\star) \overset{l\to\infty}{\longrightarrow} 0$. By (1) and (5), $w^\star = 0$ holds in S_{d}. Hence, $d(w^\star) = 0$ and $b(w^\star, \phi) = 0$, $\forall \phi \in L^2(\Omega)$. It follows that $a(w^\star, w^\star) \overset{l\to\infty}{\longleftarrow} a(u^{\epsilon_l} - u^{\hat{\epsilon}_l}, w^\star) = \{\epsilon_l^{-1}[d(w^\star) - b(w^\star, u^{\epsilon_l})]_{=0} - \hat{\epsilon}_l^{-1}[d(w^\star) - b(u^{\hat{\epsilon}_l}, w^\star)]_{=0}\} = 0$. By (12), $\|w^\star\|_{H^2(\Omega)} = 0$, and therefore the weak limit $u^\star = \hat{u}^\star$ is unique. ∎

A direct characterization of the limit $u^\star$ to which the primal variables $\{u^\epsilon\}$ converge is facilitated by writing the optimality condition for the minimization of J_ϵ as (15) below based upon Fenchel Duality [4]. This approach also allows a characterization of the limit $\lambda^\star$ to which the dual variables $\{\lambda^\epsilon\}$ converge. The following spaces emerge naturally:

$$L^2_{\mathrm{d}}(\Omega) = \{v \in L^2(\Omega) : (1 - \chi_{\mathrm{d}})v = 0\} \tag{13}$$

$$H_0(\Delta^2) = \{w \in H^2(\Omega) : \Delta^2 w \in L^2(\Omega), a(w,\phi) = (\Delta^2 w, \phi)_{L^2(\Omega)}, \forall \phi \in H^2(\Omega)\} \tag{14}$$

Note that the conditions specified for $H_0(\Delta^2)$ guarantee that the boundary conditions of (6) are satisfied in the sense seen below in (16).

Theorem 3 *There exists a unique solution $(\lambda^\epsilon, u^\epsilon) \in L^2_{\mathrm{d}}(\Omega) \times H_0(\Delta^2)$ to*

$$\left\{ \begin{array}{rcll} -\epsilon b(\mu, \lambda^\epsilon) + b(\mu, u^\epsilon) & = & d(\mu), & \forall \mu \in L^2_{\mathrm{d}}(\Omega) \\ b(\lambda^\epsilon, v) + a(u,v) & = & 0, & \forall v \in H^2(\Omega) \end{array} \right. \tag{15}$$

where u^ϵ is the unique solution to (11).

Proof: For any $\mu \in L^2(\Omega)$, decompose $\mu = \mu_1 + \mu_2$ with $\mu_1 \in L^2_{\rm d}(\Omega)$ and $\mu_2 \in L^2_{\rm d}(\Omega)^\perp$. Then $d(\mu_2) = 0$ and $b(\mu_2, v) = 0$, $\forall v \in H^2(\Omega)$, imply that the first equation in (15) must hold $\forall \mu \in L^2(\Omega)$ and in particular $\forall \mu \in H^2(\Omega)$. Setting $\mu = v$ and eliminating λ^ϵ from (15) shows that u^ϵ must be the unique solution to (11) guaranteed by Theorem 1. Recalling (4) and setting $\mu = \epsilon\lambda^\epsilon + \chi_{\rm d}(u_{\rm d} - u^\epsilon)$ in the first equation of (15) shows that $\lambda^\epsilon \in L^2_{\rm d}(\Omega)$ is uniquely determined by $\lambda^\epsilon = \chi_{\rm d}(u^\epsilon - u_{\rm d})/\epsilon$. Again, writing the first equation for arbitrary $\mu \in H^2(\Omega) \subset L^2(\Omega)$ and combining this with (11) shows that $(\lambda^\epsilon, u^\epsilon)$ also satisfies the second equation in (15). According to elliptic regularity results [8], the solution to (11) satisfies $u^\epsilon \in H^4(\Omega)$. Also, the traces in the following integration of (11) by parts are adequately defined:

$$\epsilon\left\{(\Delta^2 u^\epsilon, \phi)_{L^2(\Omega)} + \int_{\partial\Omega} [\nabla\phi \cdot \partial_n \nabla u^\epsilon - \phi \partial_n \Delta u^\epsilon]\right\} = d(\phi) - b(u^\epsilon, \phi), \quad \forall \phi \in H^2(\Omega) \tag{16}$$

Choosing ϕ in (16) to be concentrated on $\partial\Omega$ shows that the boundary integral in (16) must vanish. Subtracting the result from (11) shows that $a(u^\epsilon, \phi) = (\Delta^2 u^\epsilon, \phi)$, $\forall \phi \in H^2(\Omega)$, and hence $u^\epsilon \in H_0(\Delta^2)$. ■

It will next be shown that in the limit of vanishing regularization, solutions to (15) suffer a loss of regularity as they converge to a unique limit characterized by (18) below. Specifically, while $\{u^\epsilon\} \subset H_0(\Delta^2)$ and $\{\lambda^\epsilon\} \subset L_{\rm d}(\Delta^2)$ hold, the limits satisfy only $u^\star \in H^2(\Omega)$ and $\lambda^\star \in H^{-2}_{\rm d}(\Omega)$, where the Hilbert space $H^{-2}_{\rm d}(\Omega)$ is a subspace of $H^{-2}(\Omega)$ defined here as the completion of $L^2_{\rm d}(\Omega)$ with respect to the norm

$$\|\mu\|_{H^{-2}_{\rm d}(\Omega)} = \sup_{v \in H^2(\Omega)} \frac{\int_\Omega \chi_{\rm d} \mu v}{\|v\|_{H^2(\Omega)}} \tag{17}$$

Theorem 4 *The unique solution $(\lambda^\epsilon, u^\epsilon)$ to the system (15) converges weakly in $H^{-2}_{\rm d}(\Omega) \times H^2(\Omega)$ as $\epsilon \to 0$ to a unique limit $(\lambda^\star, u^\star)$ satisfying*

$$\left\{\begin{array}{rcll} b(\mu, u) &=& d(\mu) & \forall \mu \in H^{-2}_{\rm d}(\Omega) \\ b(\lambda, v) + a(u, v) &=& 0 & \forall v \in H^2(\Omega) \end{array}\right. \tag{18}$$

Proof: By Theorem 3, u^ϵ is the unique solution to (11), so it follows with Theorem 2 that the sequence $\{u^\epsilon\}$ converges weakly in $H^2(\Omega)$ to the limit $u^\star$. With this weak convergence of $\{u^\epsilon\}$,

$$\int_\Omega \chi_{\rm d} \lambda^\epsilon \phi = b(\lambda^\epsilon, \phi/\kappa^2_{\rm d}) = -a(u^\epsilon, \phi/\kappa^2_{\rm d}) \stackrel{\epsilon\to 0}{\longrightarrow} -a(u^\star, \phi/\kappa^2_{\rm d}), \quad \forall \phi \in H^2(\Omega)$$

where (4) has been used. It follows that $\{\lambda^\epsilon\}$ converges weakly in $H^{-2}_{\rm d}(\Omega)$, so let $\lambda^\star \in H^{-2}_{\rm d}(\Omega)$ denote the weak limit, where $\int_\Omega \chi_{\rm d} \lambda^\star \phi$ agrees with the rightmost term in the last equation. Thus, $\lambda^\star$ is given uniquely in terms of $u^\star$. Furthermore,

$$b(\lambda^\star, \psi) = \int_\Omega \kappa^2 \lambda^\star \psi = \int_\Omega \chi_{\rm d} \lambda^\star (\psi \kappa^2_{\rm d}) = -\int_\Omega \nabla^2 u^\star : \nabla^2 \psi = -a(u^\star, \psi) \quad \forall \psi \in H^2(\Omega)$$

shows that $(\lambda^\star, u^\star)$ satisfies the second equation in (18). Using the weak convergence of $\{u^\epsilon\}$, then the first equation of (15) and finally the weak convergence of $\{\lambda^\epsilon\}$, it follows that

$$\int_\Omega \kappa^2 (u^\star - u_{\rm d}) v = \lim_{\epsilon\to 0} \int_\Omega \kappa^2 (u^\epsilon - u_{\rm d}) v = \lim_{\epsilon\to 0} \epsilon \int_\Omega \chi_{\rm d} \lambda^\epsilon (\kappa^2_{\rm d} v) = \int_\Omega \chi_{\rm d} \lambda^\star (\kappa^2_{\rm d} v) \lim_{\epsilon\to 0} \epsilon = 0 \quad \forall v \in H^2(\Omega).$$

For a given $\mu \in H^{-2}_{\rm d}(\Omega)$, choose $\{\mu_k\} \subset L^2_{\rm d}(\Omega)$ to converge to μ in $H^{-2}_{\rm d}(\Omega)$ as $k \to \infty$, and then choose $\{v_{kl}\} \subset H^2(\Omega)$ to converge to μ_k in $L^2(\Omega)$ as $l \to \infty$. Setting $v = v_{kl}$ and $\phi = \kappa^2_{\rm d}(u^\star - u_{\rm d}) \in H^2(\Omega)$ in the last equation and then recalling (4), (9) and (10) gives

$$0 = \int_\Omega \kappa^2 (u^\star - u_{\rm d}) v_{kl} = \int_\Omega \chi_{\rm d} \phi v_{kl} \stackrel{l\to\infty}{\longrightarrow} \int_\Omega \chi_{\rm d} \phi \mu_k \stackrel{k\to\infty}{\longrightarrow} \int_\Omega \chi_{\rm d} \phi \mu = b(\mu, u^\star) - d(\mu)$$

which shows that $(\lambda^\star, u^\star)$ satisfies the first equation in (18). ■

Note that (18) may be seen as a condition for a saddle point of the Lagrangian functional,

$$L(v,\mu) = \tfrac{1}{2}a(v,v) + b(\mu,v) - d(\mu), \qquad v \in H^2(\Omega), \quad \mu \in H^2_{\rm d}(\Omega)$$

which solves

$$\min_{v\in H^2(\Omega)} a(v,v) \quad \text{subject to} \quad b(\mu,v) = d(\mu), \quad \forall \mu \in H^{-2}_{\rm d}(\Omega) \tag{19}$$

Before proceeding to the numerical analysis of vanishing regularization solutions, it is first indicated that the theory of saddle points in [7] is only partially applicable in this work. In particular, the loss of regularity in the limit of vanishing regularization is not obtained with the alternative theory. Furthermore, in the approximation of saddle points by regularization on p. 65 of [7], the term $-\epsilon b(\mu,\lambda^\epsilon)$ above in (15) is replaced by a term $-\epsilon c(\mu,\lambda^\epsilon)$ where c is a coercive bilinear form on the space containing $\lambda^\star$. The simplest choice for this bilinear form c in the present context is the inner product on $H^{-2}_{\rm d}(\Omega)$, which can be given explicitly as follows. For a given $\lambda \in H^{-2}_{\rm d}(\Omega)$ define the operator C so that $C\lambda \in H^2(\Omega)$ is the solution to

$$(C\lambda, v)_{H^2(\Omega)} = (\lambda, v)_{H^{-2}_{\rm d}(\Omega), H^2(\Omega)}, \quad \forall v \in H^2(\Omega). \tag{20}$$

Then

$$(\mu,\lambda)_{H^{-2}_{\rm d}(\Omega)} = (C\lambda, C\mu)_{H^2(\Omega)} = c(\mu,\lambda), \quad \lambda,\mu \in H^{-2}_{\rm d}(\Omega) \tag{21}$$

However, if (15) is so modified, then the new equation is the optimality condition for a functional $\tilde{J}^\epsilon(v) = \epsilon a(v,v) + b(C^{-1}(v-u_{\rm d}), (v-u_{\rm d}))$, which does not agree with J^ϵ. Thus, the approach of this section has been used. Nevertheless, the theory of [7] is applied later for the convergence of numerical approximations.

3 Analysis of a Mixed Finite Element Method

For $\Omega = (0,1)$ and with cell interfaces $x_i = ih$, $i = 0,\dots,N$, $h = 1/N$, the following spline bases are defined,

$$S^{(k)}_h(\Omega) = \left\{ s \in \mathcal{P}^k([x_{i-1},x_i]), i = 1,\dots,N, s \in C^{k-1}(\Omega) \right\}, \quad k = 0,1,\dots \tag{22}$$

where $\mathcal{P}^k(D)$ is the set of polynomials of degree k on D and cellwise constant functions are meant for the case $k=0$. For $\Omega = (0,1)^n$, $n > 1$, $S^{(k)}_h(\Omega)$ is understood as tensor products of such spline bases. Also, $S^{(k)}_{h,\rm d}(\Omega)$ is the subspace of splines in $S^{(k)}_h(\Omega)$ which are supported purely on $S_{\rm d}$. As seen above in the analysis of the saddle point formulation of optimality, the spaces $H^2(\Omega)$ and $H^{-2}_{\rm d}(\Omega)$ are the suitable settings for the primal variable and the dual variable, respectively. For the mixed finite element method analyzed here, $H^2(\Omega)$ is approximated by $S^{(2)}_h(\Omega)$ and $H^{-2}_{\rm d}(\Omega)$ is approximated by $S^{(0)}_{h,\rm d}(\Omega)$. It is assumed that the data are approximated with pixelwise averages

$$\chi_{\rm d}, \kappa_h, \tilde{u}_h \in S^{(0)}_{h,\rm d}(\Omega) \tag{23}$$

which converge in $L^\infty(\Omega)$,

$$\|\kappa - \kappa_h\|_{L^\infty(\Omega)} \overset{k\to\infty}{\longrightarrow} 0, \quad \|\tilde{u} - \tilde{u}_h\|_{L^\infty(\Omega)} \overset{k\to\infty}{\longrightarrow} 0. \tag{24}$$

and also satisfy the bounds of (5). With these data the integrals (9) and (10) are approximated according to

$$b_h(u,v) = \int_\Omega \kappa_h^2 uv \tag{25}$$

and

$$d_h(v) = \int_\Omega \kappa_h \tilde{u}_h v. \tag{26}$$

Then a mixed finite element discretization of the primal dual problem (15) is to compute $(\lambda_h, u_h) \in S^{(0)}_{h,\mathrm{d}}(\Omega) \times S^{(2)}_h(\Omega)$ so that

$$\left\{ \begin{array}{rcll} -\epsilon b_h(\mu_h, \lambda^\epsilon_h) + b_h(\mu_h, u^\epsilon_h) & = & d_h(\mu_h), & \forall \mu_h \in S^{(0)}_{h,\mathrm{d}}(\Omega) \\ b_h(\lambda^\epsilon_h, v_h) + a(u^\epsilon_h, v_h) & = & 0, & \forall v_h \in S^{(2)}_h(\Omega) \end{array} \right. \tag{27}$$

The following technical lemmas are used below.

Lemma 2 *b_h is symmetric and positive definite on $S^{(0)}_{h,\mathrm{d}}(\Omega)$.*

Proof: b_h is clearly symmetric on $S^{(0)}_h(\Omega)$. Since κ_h and $\tilde{u}_h$ satisfy the bounds of (5), $b_h(\mu_h, \mu_h) \geq \kappa_0 \|\mu_h\|^2_{L^2(\Omega)}$ holds $\forall \mu_h \in S^{(0)}_h(\Omega)$. ■

With Lemma 2, define the solution operator $B_h : L^2(\Omega) \to S^{(0)}_{h,\mathrm{d}}(\Omega)$ by

$$v \in L^2(\Omega), \quad b_h(\mu_h, B_h v) = b_h(\mu_h, v), \quad \forall \mu_h \in S^{(0)}_{h,\mathrm{d}}(\Omega) \tag{28}$$

Since for an arbitrary $\tilde{\mu}_h \in S^{(0)}_{h,\mathrm{d}}(\Omega)$ a corresponding $\mu_h = \tilde{\mu}_h/\kappa^2_h \in S^{(0)}_{h,\mathrm{d}}(\Omega)$ may be chosen for (28), it follows that $B_h v$ is the $L^2(\Omega)$ projection of $v \in L^2(\Omega)$ into $S^{(0)}_{h,\mathrm{d}}(\Omega)$. Thus, let $\sigma(h)$ satisfy $\sigma(h) \overset{h\to 0}{\longrightarrow} 0$ and

$$\|\chi_\mathrm{d}(B_h v - v)\|_{L^2(\Omega)} \leq \sigma(h) \|v\|_{L^2(\Omega)} \tag{29}$$

The next lemma shows that the restriction of B_h to $S^{(2)}_h(\Omega)$ maps onto $S^{(0)}_{h,\mathrm{d}}(\Omega)$.

Lemma 3 *For every $\nu_h \in S^{(0)}_{h,\mathrm{d}}(\Omega)$ there is a $u_h \in S^{(2)}_h(\Omega)$ satisfying $b_h(B_h u_h, \mu_h) = b_h(\nu_h, \mu_h)$, $\forall \mu_h \in S^{(0)}_{h,\mathrm{d}}(\Omega)$.*

Proof: Let χ_i be the characteristic function for the ith cell so that $\{\chi_i\}$ is a basis for $S^{(0)}_h(\Omega)$. Given a (tensor product of a) canonical spline supported on three adjacent cells, let translations $\{s_j\}$ form a basis for $S^{(2)}_h(\Omega)$, which is higher dimensional that $S^{(0)}_h(\Omega)$. Define the matrices $B = \{\int_\Omega \kappa^2_h \chi_i s_j\}$, $P = \{\int_\Omega \chi_i s_j\}$ and $C = \{\int_\Omega \kappa^2_h \chi_i \chi_j\}$ which satisfy $B = CP$. For $\nu_h = \sum_i \nu_i \chi_i$ and $u_h = \sum_j u_j s_j$ with coefficient vectors $\boldsymbol{\nu} = \{\nu_i\}$ and $\boldsymbol{u} = \{u_j\}$, the equation to solve for $\boldsymbol{u}$ is $CP\boldsymbol{u} = C\boldsymbol{\nu}$. For this, define a square submatrix of P by $\tilde{P} = \{\int_\Omega \chi_i s_j : j \in J\}$ where splines $\{s_j : j \in J\}$ have the bulk of their mass in some ith cell, i.e., not outside Ω. A direct calculation shows that for $n = 1$ the submatrix is given with the lexicographical ordering by $\tilde{P} = \mathrm{tridiag}\{\frac{1}{6}, \frac{2}{3}, \frac{1}{6}\}$, which is diagonally dominant and hence invertible. For $n > 1$ the submatrix is given by a product of diagonally dominant matrices, e.g., for $n = 2$, $\tilde{P} = \tilde{P}_y \tilde{P}_x$ $\tilde{P}_x = \mathrm{tridiag}\{\frac{1}{6}, \frac{2}{3}, \frac{1}{6}\}$, $\tilde{P}_y = \mathrm{tridiag}\{\frac{1}{6}, 0, \ldots, 0, \frac{2}{3}, 0, \ldots, 0, \frac{1}{6}\}$. Define $\tilde{\boldsymbol{u}} = \tilde{P}^{-1}\boldsymbol{\nu}$ and pad $\tilde{\boldsymbol{u}}$ with zeros for the indices $j \notin J$ to construct a $\boldsymbol{u}$ satisfying $P\boldsymbol{u} = \boldsymbol{\nu}$ and hence $CP\boldsymbol{u} = C\boldsymbol{\nu}$. ■

It is further assumed that $\kappa_h \tilde{u}_h$ approximates $\kappa \tilde{u} = \kappa^2 u_\mathrm{d}$ sufficiently, so let $\tau(h)$ satisfy $\tau(h) \overset{h\to 0}{\longrightarrow} 0$ and:

$$\sup_{\mu_h \in S^{(0)}_{h,\mathrm{d}}(\Omega)} \frac{|d(\mu_h) - d_h(\mu_h)|}{\|\mu_h\|_{H^{-2}_\mathrm{d}(\Omega)}} = \sup_{\mu_h \in S^{(0)}_{h,\mathrm{d}}(\Omega)} \|\mu_h\|^{-1}_{H^{-2}_\mathrm{d}(\Omega)} \int_\Omega \mu_h(\kappa^2 u_\mathrm{d} - \kappa_h \tilde{u}_h) \leq \tau(h) \|u_\mathrm{d}\|_{H^2(\Omega)}. \tag{30}$$

It will be proven in this section that there exists a $(\lambda^\star_h, u^\star_h) \in S^{(0)}_{h,\mathrm{d}}(\Omega) \times S^{(2)}_h(\Omega)$ such that the solution $(\lambda^\epsilon_h, u^\epsilon_h)$ to (27) satisfies $(\lambda^\epsilon_h, u^\epsilon_h) \overset{\epsilon\to 0}{\longrightarrow} (\lambda^\star_h, u^\star_h)$, and that $(\lambda^\star_h, u^\star_h) \overset{h\to 0}{\longrightarrow} (\lambda^\star, u^\star)$, where $(\lambda^\star, u^\star)$ is given by Theorem 4.

Lemma 4 *For h sufficiently small there exist constants $\tilde{c}_1$ and $\tilde{c}_2$ independent of h such that*

$$\tilde{c}_1\|v\|^2_{H^2(\Omega)} \le a(v,v) + b_h(B_hv, B_hv) \le \tilde{c}_2\|v\|^2_{H^2(\Omega)}, \quad \forall v \in H^2(\Omega). \tag{31}$$

Proof: By (28), $|b_h(B_hv, B_hv)| = |b_h(v, B_hv)| \le [b_h(v,v)]^{\frac{1}{2}}[b_h(B_hv, B_hv)]^{\frac{1}{2}}$ so

$$|b_h(B_hv, B_hv)| \le b_h(v,v), \quad \forall v \in H^2(\Omega) \tag{32}$$

With the bounds (5) for κ_h and $\tilde{u}_h$, (12) and (32) give $\tilde{c}_2 = c_2$ in (31). Then using (24), (28), (29) and (32),

$$\begin{aligned} |b_h(B_hv, B_hv) - b(v,v)| &\le |b_h(v, B_hv) - b(v, B_hv)| + |b(v, B_h) - b(v,v)| \\ &\le \|\kappa_h^2 - \kappa^2\|_{L^\infty(\Omega)}\|v\|_{L^2(\Omega)}\|B_hv\|_{L^2(\Omega)} \\ &\qquad + \|\kappa^2\|_{L^\infty(\Omega)}\|v\|_{L^2(\Omega)}\|\chi_{\mathrm{d}}(B_hv - v)\|_{L^2(\Omega)} \\ &\le \|\kappa_h^2 - \kappa^2\|_{L^\infty(\Omega)}\|v\|_{L^2(\Omega)}[\|v\|_{L^2(\Omega)} + \|\chi_{\mathrm{d}}(B_hv - v)\|_{L^2(\Omega)}] \\ &\qquad + \|\kappa^2\|_{L^\infty(\Omega)}\|v\|_{L^2(\Omega)}\|\chi_{\mathrm{d}}(B_hv - v)\|_{L^2(\Omega)} \\ &\overset{h\to 0}{\longrightarrow} 0 \end{aligned}$$

and coercivity follows with (12) for h sufficiently small and $\tilde{c}_1 \in (0, c_1)$. ∎

Theorem 5 *There exists a unique $(\lambda_h^\epsilon, u_h^\epsilon) \in S^{(0)}_{h,\mathrm{d}}(\Omega) \times S^{(2)}_h(\Omega)$ satisfying (27).*

Proof: Applying (28) to the first equation of (27) gives $b_h(\mu_h, u_h^\epsilon) = b_h(\mu_h, B_hu_h^\epsilon)$ or

$$-\epsilon b_h(\mu_h, \lambda_h^\epsilon) + b_h(\mu_h, B_hu_h^\epsilon) = d_h(\mu_h), \quad \forall \mu_h \in S^{(0)}_{h,\mathrm{d}}(\Omega) \tag{33}$$

Applying (28) to the second equation of (27) gives $b_h(\lambda_h^\epsilon, v_h) = b_h(\lambda_h^\epsilon, B_hv_h)$, and it follows with $\mu_h = B_hv_h$ in (33) that

$$\begin{aligned} b_h(B_hv_h, B_hu_h^\epsilon) - d_h(B_hv_h) + \epsilon a(u_h^\epsilon, v_h) &= \epsilon b_h(B_hv_h, \lambda_h^\epsilon) + \epsilon a(u_h^\epsilon, v_h) \\ &= \epsilon b_h(v_h, \lambda_h^\epsilon) + \epsilon a(u_h^\epsilon, v_h) = 0, \quad \forall v_h \in S^{(2)}_h(\Omega) \end{aligned}$$

where the second equation of (27) gives the last equality. Hence u_h^ϵ must satisfy

$$b_h(B_hu_h^\epsilon, B_hv_h) + \epsilon a(u_h^\epsilon, v_h) = d_h(B_hv_h), \quad \forall v_h \in S^{(2)}_h(\Omega) \tag{34}$$

By (4), (5), (26) and (29),

$$\begin{aligned} |d_h(B_hv_h)| &\le \kappa_1\tilde{u}_1\|\chi_{\mathrm{d}}B_hv_h\|_{L^2(\Omega)} \le \kappa_1\tilde{u}_1[\|\chi_{\mathrm{d}}v_h\|_{L^2(\Omega)} + \|\chi_{\mathrm{d}}(B_hv_h - v_h)\|_{L^2(\Omega)}] \\ &\le \kappa_1\tilde{u}_1[1 + \sigma(h)]\|v_h\|_{H^2(\Omega)}, \quad \forall v \in S^{(2)}_h(\Omega). \end{aligned}$$

Thus, with Lemma 4, it follows with the Lax-Milgram Lemma [3] that u_h^ϵ is uniquely defined by (34). So with Lemma 2, λ_h^ϵ is uniquely determined by (33). ∎

The vanishing regularization solution $u_h^\star$, whose existence is established next in Theorem 6, is characterized precisely below in (36).

Theorem 6 *The solutions $\{u_h^\epsilon\}$ to (27) converge to a unique limit $u_h^\star \in S^{(2)}_h(\Omega)$ as $\epsilon \to 0$.*

Proof: It will be first be shown that $\{u_h^\epsilon\}$ is bounded in $H^2(\Omega)$ independently of ϵ. By Lemma 2, let $\nu_h \in S^{(0)}_{h,\mathrm{d}}(\Omega)$ be chosen so that $b_h(\nu_h, \mu_h) = d_h(\mu_h)$, $\forall \mu_h \in S^{(0)}_{h,\mathrm{d}}(\Omega)$, and then by Lemma 3, let $u_{h,\mathrm{d}}$ be chosen to satisfy $b_h(B_hu_{h,\mathrm{d}}, \mu_h) = b_h(\nu_h, \mu_h) = d_h(\mu_h)$, $\forall \mu_h \in S^{(0)}_{h,\mathrm{d}}(\Omega)$. Setting $\mu_h = B_hv_h$, $v_h \in S^{(2)}_h(\Omega)$, in this equation and subtracting from (34) with $v_h = w_h^\epsilon = u_h^\epsilon - u_{h,\mathrm{d}}$

gives $\epsilon a(w_h^\epsilon, w_h^\epsilon) + b_h(B_h w_h^\epsilon, B_h w_h^\epsilon) = -\epsilon a(u_{h,\mathrm{d}}, w_h^\epsilon)$. It follows that $a(w_h^\epsilon, w_h^\epsilon) \le |a(u_{h,\mathrm{d}}, w_h^\epsilon)|$ or $a(w_h^\epsilon, w_h^\epsilon) \le a(u_{h,\mathrm{d}}, u_{h,\mathrm{d}})$ and that $b_h(B_h w_h^\epsilon, B_h w_h^\epsilon) \le \epsilon |a(u_{h,\mathrm{d}}, w_h^\epsilon)|$ or $b_h(B_h w_h^\epsilon, B_h w_h^\epsilon) \le \frac{1}{2}a(u_{h,\mathrm{d}}, u_{h,\mathrm{d}}) + \frac{1}{2}a(w_h^\epsilon, w_h^\epsilon)$ for $\epsilon \le 1$. Hence, $\frac{1}{2}a(w_h^\epsilon, w_h^\epsilon) + b_h(B_h w_h^\epsilon, B_h w_h^\epsilon) \le \frac{3}{2}a(u_{h,\mathrm{d}}, u_{h,\mathrm{d}})$, and with (31), w_h^ϵ is bounded in $H^2(\Omega)$. Thus, u_h^ϵ is bounded in $H^2(\Omega)$ in terms of w_h^ϵ and $u_{h,\mathrm{d}}$. Since $S_h^{(2)}(\Omega)$ is finite dimensional, there is a subsequence $\{u_h^{\epsilon_l}\}$ which converges in $H^2(\Omega)$ to a limit $u^\star \in S_h^{(2)}(\Omega)$ as $\epsilon \to 0$.

If there were another limit, $\hat{u}_h^{\hat{\epsilon}_l} \to \hat{u}_h^\star$, then according to (34), $w_h^\star = u_h^\star - \hat{u}_h^\star$ would satisfy $b_h(B_h w_h^\star, B_h w_h^\star) \overset{l\to\infty}{\longleftarrow} b_h(B_h w_h^\star, B_h u_h^{\epsilon_l} - u_h^{\hat{\epsilon}_l}) = \hat{\epsilon}_l a(u_h^{\hat{\epsilon}_l}, w_h^\star) - \epsilon_l a(u_h^{\epsilon_l}, w_h^\star) \overset{l\to\infty}{\longrightarrow} 0$. By Lemma 2, $B_h w_h^\star = 0$. It follows that $a(w_h^\star, w_h^\star) \overset{l\to\infty}{\longleftarrow} a(u_h^{\epsilon_l} - u_h^{\hat{\epsilon}_l}, w_h^\star) = \{\epsilon_l^{-1}[d_h(B_h w_h^\star) - b_h(B_h w_h^\star, B_h u_h^{\epsilon_l})]_{=0} - \hat{\epsilon}_l^{-1}[d_h(w_h^\star) - b_h(B_h u_h^{\hat{\epsilon}_l}, B_h w_h^\star)]_{=0}\} = 0$. With (31), $\|w_h^\star\|_{H^2(\Omega)} = 0$, and therefore the weak limit $u_h^\star = \hat{u}_h^\star$ is unique. ∎

For Theorem 7 below, the following constructions analogous to (20) are used. For a given $\mu_h \in S_{h,\mathrm{d}}^{(0)}(\Omega)$ define the operator C_h so that $C_h \mu_h \in S_h^{(2)}(\Omega)$ is the solution to

$$(C_h \mu_h, v_h)_{H^2(\Omega)} = (\mu_h, v_h)_{L^2(\Omega)}, \quad \forall v_h \in S_h^{(2)}(\Omega) \tag{35}$$

Then

$$\sup_{v_h \in S_h^{(2)}(\Omega)} \frac{(\mu_h, v_h)_{L^2(\Omega)}}{\|v_h\|_{H^2(\Omega)}} = \sup_{v_h \in S_h^{(2)}(\Omega)} \frac{(v_h, C_h\mu_h)_{H^2(\Omega)}}{\|v_h\|_{H^2(\Omega)}} \le \|C_h \mu_h\|_{H^2(\Omega)}$$

where equality holds for $v_h = C_h \mu_h$. As seen in the proof of Lemma 3, the matrix P has maximal rank, and hence $C_h \mu_h = 0$ if and only if $\mu_h = 0$. Thus, the following is a norm on $S_{h,\mathrm{d}}^{(0)}(\Omega)$,

$$\|\mu_h\|_{S_{h,\mathrm{d}}^{(0)}(\Omega)} = \sup_{v_h \in S_h^{(2)}(\Omega)} \frac{(\mu_h, v_h)_{L^2(\Omega)}}{\|v_h\|_{H^2(\Omega)}} = \|C_h \mu_h\|_{H^2(\Omega)}$$

with associated scalar product,

$$(\mu_h, \lambda_h)_{S_{h,\mathrm{d}}^{(0)}(\Omega)} = (C_h \lambda_h, C_h \mu_h)_{H^2(\Omega)} = (\mu_h, C_h \lambda_h)_{L^2(\Omega)}, \quad \lambda_h, \mu_h \in S_{h,\mathrm{d}}^{(0)}(\Omega).$$

Next, the vanishing regularization solution is characterized.

Theorem 7 *The unique solution $(\lambda_h^\epsilon, u_h^\epsilon)$ to the system (27) converges as $\epsilon \to 0$ to a unique limit $(\lambda_h^\star, u_h^\star) \in S_{h,\mathrm{d}}^{(0)}(\Omega) \times S_h^{(2)}(\Omega)$ satisfying*

$$\left\{ \begin{array}{rcll} b_h(\mu_h, u_h^\star) & = & d_h(\mu_h) & \forall \mu_h \in S_{h,\mathrm{d}}^{(0)}(\Omega) \\ b_h(\lambda_h^\star, v_h) + a(u_h^\star, v_h) & = & 0 & \forall v_h \in S_h^{(2)}(\Omega) \end{array} \right. \tag{36}$$

Proof: By Theorem 6, $\{u_h^\epsilon\}$ converges to the limit $u_h^\star$. Using this convergence of $\{u_h^\epsilon\}$ in (27),

$$\begin{aligned} (\kappa_h^2 \lambda_h^\epsilon, \mu_h)_{S_{h,\mathrm{d}}^{(0)}(\Omega)} = (\kappa_h^2 \lambda_h^\epsilon, C_h \mu_h)_{L^2(\Omega)} = b_h(\lambda_h^\epsilon, C_h \mu_h) = \\ -a(u_h^\epsilon, C_h \mu_h) \overset{\epsilon\to 0}{\longrightarrow} -a(u_h^\star, C_h\mu_h), \quad \forall \mu_h \in S_{h,\mathrm{d}}^{(0)}(\Omega) \end{aligned}$$

Thus, due to the the finite dimensionality of $S_{h,\mathrm{d}}^{(0)}(\Omega)$, $\{\kappa_h^2 \lambda_h^\epsilon\}$ converges to some $\kappa_h^2 \lambda_h^\star \in S_{h,\mathrm{d}}^{(0)}(\Omega)$. Furthermore,

$$-a(u_h^\star, v_h) = -\lim_{\epsilon\to 0} a(u_h^\epsilon, v_h) = \lim_{\epsilon \to 0} b_h(\lambda_h^\epsilon, v_h) = b_h(\lambda_h^\star, v_h), \quad \forall v_h \in S_h^{(2)}(\Omega)$$

shows that the second equation of (36) is satisfied and

$$(\kappa_h^2 \lambda_h^\star, \mu_h)_{S_{h,\mathrm{d}}^{(0)}(\Omega)} = b_h(\lambda_h^\star, C_h \mu_h) = -a(u_h^\star, C_h \mu_h), \quad \forall \mu_h \in S_{h,\mathrm{d}}^{(0)}(\Omega)$$

together with (5) imply that $\lambda_h^\star$ is uniquely determined in terms of $u_h^\star$. Using the convergence of $\{u_h^\epsilon\}$, then the first equation of (27) and finally the convergence of $\{\lambda_h^\epsilon\}$, it follows that $\forall \mu_h \in S_{h,\mathrm{d}}^{(0)}(\Omega)$,

$$b_h(\mu_h, u_h^\star) - d_h(\mu_h) = \lim_{\epsilon\to 0} b_h(\mu_h, u_h^\epsilon) - d_h(\mu_h) = \lim_{\epsilon\to 0} \epsilon b_h(\mu_h, \lambda_h^\epsilon) = b_h(\mu_h, \lambda_h^\star) \lim_{\epsilon\to 0} \epsilon = 0$$

which shows that $(\lambda^\star, u^\star)$ satisfies the first equation in (36). ■

Note that (36) may be seen as a condition for a saddle point of the Lagrangian functional,

$$L(v_h, \mu_h) = \tfrac{1}{2} a(v_h, v_h) + b_h(\mu_h, v_h) - d_h(\mu_h), \qquad v_h \in S_h^{(2)}(\Omega), \quad \mu \in S_{h,\mathrm{d}}^{(0)}(\Omega) \tag{37}$$

which solves

$$\min_{v_h \in S_h^{(2)}(\Omega)} a(v_h, v_h) \quad \text{subject to} \quad b_h(\mu_h, v_h) = d_h(\mu_h), \quad \forall \mu_h \in S_{h,\mathrm{d}}^{(0)}(\Omega) \tag{38}$$

Due to the approximation of data, the forms b_h and d_h may not be subtracted directly from their counterparts b and d for estimations of identical arguments. In particular, the conditions seen below in Lemma 6 evidently do not hold for b_h. For the analysis, it is crucial that (18) and (36) can be reformulated below as (42) and (45). For this, define

$$\hat{b}(\mu, u) = \int_\Omega \chi_\mathrm{d} \mu u \tag{39}$$

satisfying $\hat{b}(\mu, u) = (\mu, u)_{H_\mathrm{d}^{-2}(\Omega), H^2(\Omega)}$ and

$$\hat{d}(\mu) = \int_\Omega \chi_\mathrm{d} u_\mathrm{d} \mu \tag{40}$$

satisfying $\hat{d}(\mu) = \hat{b}(\mu, u_\mathrm{d})$. Then set

$$\hat{\lambda}^\star = \kappa^2 \lambda^\star \in H_\mathrm{d}^{-2}(\Omega) \tag{41}$$

and $\hat{\mu} = \kappa^2 \mu \in H_\mathrm{d}^{-2}(\Omega)$ in (18) so that $(\lambda^\star, u^\star)$ satisfies (18) if and only if $(\hat{\lambda}^\star, u^\star)$ satisfies

$$\begin{cases} \hat{b}(\hat{\mu}, u^\star) &= \hat{d}(\hat{\mu}) \quad \forall \hat{\mu} \in H_\mathrm{d}^{-2}(\Omega) \\ \hat{b}(\hat{\lambda}^\star, v) + a(u^\star, v) &= 0 \quad\quad \forall v \in H^2(\Omega) \end{cases} \tag{42}$$

Similarly, with

$$\hat{d}_h(v) = \int_\Omega \kappa_h^{-1} \tilde{u}_h v \tag{43}$$

set

$$\hat{\lambda}_h^\star = \kappa_h^2 \lambda_h^\star \in S_{h,\mathrm{d}}^{(0)}(\Omega) \tag{44}$$

and $\hat{\mu}_h = \kappa_h^2 \mu_h \in S_{h,\mathrm{d}}^{(0)}(\Omega)$ in (36) so that $(\lambda_h^\star, u_h^\star)$ satisfies (36) if and only if $(\hat{\lambda}_h^\star, u_h^\star)$ satisfies

$$\begin{cases} \hat{b}(\hat{\mu}_h, u_h^\star) &= \hat{d}_h(\hat{\mu}_h) \quad \forall \hat{\mu}_h \in S_{h,\mathrm{d}}^{(0)}(\Omega) \\ \hat{b}(\hat{\lambda}_h^\star, v_h) + a(u_h^\star, v_h) &= 0 \quad\quad \forall v_h \in S_h^{(2)}(\Omega) \end{cases} \tag{45}$$

To show that the solution $(u_h^\star, \hat{\lambda}_h^\star)$ to (45) converges with grid refinement to the solution $(u^\star, \hat{\lambda}^\star)$ to (42), it will be shown that the conditions of Theorem 1.1 on p. 114 of [7] are met, but the proof will be modified to accommodate the approximation of data. Note that $\hat{b}$ now appears in both (42) and (45), so the only new term to be estimated is $\hat{d} - \hat{d}_h$. The first condition to be proved is that the form a is coercive on

$$V = \{u \in H^2(\Omega) : \hat{b}(\mu, u) = 0, \forall \mu \in H_\mathrm{d}^{-2}(\Omega)\} \tag{46}$$

and on its approximation

$$V_h = \{u_h \in S_h^{(2)}(\Omega) : \hat{b}(\mu_h, u_h) = 0, \forall \mu_h \in S_{h,\mathrm{d}}^{(0)}(\Omega)\}. \tag{47}$$

Lemma 5 *There exists a constant α independent of h such that,*

$$a(u,u) \geq \alpha \|u\|^2_{H^2(\Omega)}, \quad \forall u \in V \tag{48}$$

and,

$$a(u_h,u_h) \geq \alpha \|u_h\|^2_{H^2(\Omega)}, \quad \forall u_h \in V_h \tag{49}$$

Proof: First, (48) is proved with a a Poincaré argument since $\hat{b}$ coerces the kernel $\mathcal{L}$ of a. For (49), note that $V_h \subset V$ does not hold since a function $v_h \in V_h$ need not to vanish in S_d but rather only to have average value zero in every cell of S_d. To prove (49), suppose $\exists u_h \in V_h$ such that $a(u_h,u_h) = 0$. Thus, $u_h \in \mathcal{L}$ is linear. Since $u_h \in V_h$, its average value on every cell in S_d is zero. These two conditions imply that $u_h = 0$, and it follows that a is positive definite on V_h.

To show that α in (49) is independent of h, suppose for the sake of contradiction that there exists a sequence $\{u_{h_k}\}$ with $u_{h_k} \in V_{h_k}$ satisfying $\|u_{h_k}\|_{H^2(\Omega)} = 1$ while $h_k = 2^{-k}h_1 \overset{k\to\infty}{\longrightarrow} 0$ and $a(u_{h_k},u_{h_k}) \overset{k\to\infty}{\longrightarrow} 0$. Then the sequence is bounded in $H^2(\Omega)$ according to $\|u_{h_k}\|_{H^2(\Omega)} = 1$, and weak (subsequential) convergence in $H^2(\Omega)$ follows. For convenience, let $\{u_{h_k}\}$ again denote the subsequence. By compactness of $H^1(\Omega)$ in $H^2(\Omega)$, the sequence converges strongly in $H^1(\Omega)$. With $a(u_{h_k},u_{h_k}) \overset{k\to\infty}{\longrightarrow} 0$, it follows that the sequence converges strongly in $H^2(\Omega)$ to some u_0. Since $a(u_0,u_0) = \lim_{k\to\infty} a(u_{h_k},u_{h_k}) = 0$, $u_0 \in \mathcal{L}$. As $S^{(0)}_{h_{k+1},\mathrm{d}}(\Omega) \subset S^{(0)}_{h_k,\mathrm{d}}(\Omega)$, it follows that $\forall l$,

$$\hat{b}(\mu_{h_l},u_0) = \lim_{k\to\infty} \hat{b}(\mu_{h_l},u_k) = 0, \quad \forall \mu_{h_l} \in S^{(0)}_{h_l,\mathrm{d}}(\Omega)$$

Since $\cup_{l=1}^{\infty} S^{(0)}_{h_l,\mathrm{d}}(\Omega)$ is dense in $L^2_\mathrm{d}(\Omega)$, the linear function u_0 vanishes in S_d and hence $u_0 = 0$. However, this contradicts $\|u_0\|_{H^{(2)}(\Omega)} = \lim_{k\to\infty} \|u_{h_k}\|_{H^2(\Omega)} = 1$. The contradiction shows that (48) holds for an α independent of h. ∎

The next condition to be demonstrated is that the form $\hat{b}$ satisfies the Ladysenskaja-Babuska-Brezzi (LBB) condition on $H^2(\Omega) \times H^{-2}_\mathrm{d}(\Omega)$ and on $S^{(2)}_h(\Omega) \times S^{(2)}_{h,\mathrm{d}}(\Omega)$.

Lemma 6 *There exists a constant β independent of h such that*

$$\sup_{v\in H^2(\Omega)} \frac{\hat{b}(\mu,v)}{\|v\|_{H^2(\Omega)}} \geq \beta \|\mu\|_{H^{-2}_\mathrm{d}(\Omega)}, \quad \forall \mu \in H^{-2}_\mathrm{d}(\Omega) \tag{50}$$

and

$$\sup_{v_h\in S^{(2)}_h(\Omega)} \frac{\hat{b}(\mu_h,v_h)}{\|v_h\|_{H^2(\Omega)}} \geq \beta \|\mu_h\|_{H^{-2}_\mathrm{d}(\Omega)}, \quad \forall \mu_h \in S^{(0)}_{h,\mathrm{d}}(\Omega) \tag{51}$$

The solution operator C defined by (20) is an isomorphism from $(H^{-2}_\mathrm{d}(\Omega), \|\cdot\|_{H^{-2}_\mathrm{d}(\Omega)})$ onto $(V^\perp, \|\cdot\|_{H^2(\Omega)})$. Similarly, the solution operator C_h defined by (35) is an isomorphism from $(S^{(0)}_{h,\mathrm{d}}(\Omega), \|\cdot\|_{H^{-2}_\mathrm{d}(\Omega)})$ onto $(V^\perp_h, \|\cdot\|_{H^2(\Omega)})$.

Proof: The condition (50) follows immediately with (17). To demonstrate (51), the approach is to use Lemma 1.1 on p. 117 of [7]. For this, it must be shown that two conditions below, (54) and (55), are satisfied for an operator Π_h from $H^2(\Omega)$ into $S^{(0)}_{h,\mathrm{d}}(\Omega)$ defined by (53) below.

With $P^{(2)}_h$ defined as the projection of $H^2(\Omega)$ onto $S^{(2)}_h(\Omega)$, note that the operators C and C_h of (20) and (35) satisfy $C_h = P^{(2)}_h C$ on $S^{(0)}_{h,\mathrm{d}}(\Omega)$ according to

$$\begin{aligned}(P^{(2)}_h C\mu_h, v_h)_{H^2(\Omega)} = (C\mu_h, v_h)_{H^2(\Omega)} = \hat{b}(\mu_h,v_h) = (C_h\mu_h, v_h)_{H^2(\Omega)},& \\ \forall \mu_h \in S^{(0)}_{h,\mathrm{d}}(\Omega), \quad \forall v_h \in S^{(2)}_h(\Omega)&\end{aligned} \tag{52}$$

With the approximation property of the projection, $\|(I - P_h^{(2)})v\|_{H^2(\Omega)} \stackrel{h\to 0}{\longrightarrow} 0$, $\forall v \in H^2(\Omega)$, it follows for h sufficiently small, $\exists c > 0$ such that

$$\|C_h\mu_h\|_{H^2(\Omega)} \geq \|C\mu_h\|_{H^2(\Omega)} - \|(1-P_h^{(2)})C\mu_h\|_{H^2(\Omega)} \geq c\|C\mu_h\|_{H^2(\Omega)} > 0, \quad 0 \neq \mu_h \in S_{h,\mathrm{d}}^{(0)}(\Omega)$$

where the last inequality follows with (20) and (50). Hence, $C_h^* C_h$ is symmetric and positive definite on $S_{h,\mathrm{d}}^{(0)}(\Omega)$. So for a given $v \in H^2(\Omega)$, let $\lambda_h \in S_{h,\mathrm{d}}^{(2)}(\Omega)$ solve $(C_h\lambda_h, C_h\mu_h)_{H^2(\Omega)} = (v, C\mu_h)_{H^2(\Omega)}$, $\forall \mu_h \in S_{h,\mathrm{d}}^{(0)}(\Omega)$. Then define the operator Π_h from $H^2(\Omega)$ into $S_h^{(2)}(\Omega)$ by $\Pi_h v = C_h\lambda_h$ so that

$$(\Pi_h v, C_h\mu_h)_{H^2(\Omega)} = (v, C\mu_h)_{H^2(\Omega)}, \quad \forall \mu_h \in S_{h,\mathrm{d}}^{(0)}(\Omega), \quad \forall v \in H^2(\Omega). \tag{53}$$

With (20), (35) and (53),

$$\hat{b}(\mu_h, \Pi_h v) = (\Pi_h v, C_h\mu_h)_{H^2(\Omega)} = (C\mu_h, v)_{H^2(\Omega)} = \hat{b}(\mu_h, v), \quad \forall \mu_h \in S_{h,\mathrm{d}}^{(0)}(\Omega), \quad \forall v \in H^2(\Omega) \tag{54}$$

Thus, (54) is the first condition to be demonstrated. Setting $C_h\mu_h = \Pi_h v$ in (53) gives

$$\|\Pi_h v\|_{H^2(\Omega)}^2 = (v, \Pi_h v)_{H^2(\Omega)} \leq \|v\|_{H^2(\Omega)}\|\Pi_h v\|_{H^2(\Omega)} \quad \text{or} \quad \|\Pi_h v\|_{H^2(\Omega)} \leq \|v\|_{H^2(\Omega)} \tag{55}$$

and (55) is the second condition to be demonstrated. The condition (51) follows now from Lemma 1.1 on p. 117 of [7]. The characterizations of the operators C and C_h of (20) and (35) as isomorphisms follow from Lemma 4.1 on p. 58 of [7]. ■

The final conditions for Theorem 8 below are given as follows.

Lemma 7 *The forms $\hat{b}$, $\hat{d}$ and $\hat{d}_h$ of (39), (40) and (43) satisfy*

$$\sup_{v_h \in S_h^{(2)}(\Omega), \mu_h \in S_{h,\mathrm{d}}^{(0)}(\Omega)} \frac{|\hat{b}(\mu_h, v_h)|}{\|v_h\|_{H^2(\Omega)}\|\mu_h\|_{H_\mathrm{d}^{-2}(\Omega)}} \leq \sup_{v \in H^2(\Omega), \mu \in H_\mathrm{d}^{-2}(\Omega)} \frac{|\hat{b}(\mu, v)|}{\|v\|_{H^2(\Omega)}\|\mu\|_{H_\mathrm{d}^{-2}(\Omega)}} \leq 1 \tag{56}$$

$$\sup_{\mu \in H_\mathrm{d}^{-2}(\Omega)} \frac{|\hat{d}(\mu)|}{\|\mu\|_{H_\mathrm{d}^{-2}(\Omega)}} \leq \|u_\mathrm{d}\|_{H^2(\Omega)} \tag{57}$$

and

$$\sup_{\mu_h \in S_{h,\mathrm{d}}^{(2)}(\Omega)} \frac{|\hat{d}_h(\mu_h)|}{\|\mu_h\|_{H_\mathrm{d}^{-2}(\Omega)}} \leq [1 + \tau(h)]\|u_\mathrm{d}\|_{H^2(\Omega)}. \tag{58}$$

Proof: Since $S_h^{(2)}(\Omega) \subset H^2(\Omega)$ and $S_{h,\mathrm{d}}^{(0)}(\Omega) \subset H_\mathrm{d}^{-2}(\Omega)$, (56) follows immediately with the definition (17) of the norm on $H_\mathrm{d}^{-2}(\Omega)$ constructed in terms of $\hat{b}$ in (39). Similarly, (57) follows from (17) and (40). For (58), (30) and (57) are combined to give

$$|\hat{d}_h(\mu_h)| \leq |\hat{d}(\mu_h)| + |\hat{d}(\mu_h) + \hat{d}_h(\mu_h)| \leq (1 + \tau(h))\|u_\mathrm{d}\|_{H^2(\Omega)}\|\mu_h\|_{H_\mathrm{d}^{-2}(\Omega)} \tag{59}$$

and (58) follows. ■

Thus, all the conditions of Theorem 1.1 on p. 114 of [7] are met, and the result is summarized in Theorem 8 below, with the proof adapted for data approximation.

Theorem 8 *Under the conditions* (48), (49), (50), (51), (56) *and* (57), *the solution* $(\lambda_h^\star, u_h^\star)$ *to* (36) *converges to the solution* $(\lambda^\star, u^\star)$ *to* (18) *according to*

$$\|u^\star - u_h^\star\|_{H^2(\Omega)} + \|\lambda^\star - \lambda_h^\star\|_{H_{\rm d}^{-2}(\Omega)} \le K \left\{ \begin{array}{c} \inf_{v_h \in S_h^{(2)}(\Omega)} \|u^\star - v_h\|_{H^2(\Omega)} + \inf_{\mu_h \in S_{h,{\rm d}}^{(0)}(\Omega)} \|\lambda^\star - \mu_h\|_{H_{\rm d}^{-2}(\Omega)} \\ +\tau(h)\|u_{\rm d}\|_{H^2(\Omega)} \end{array} \right\} \tag{60}$$

for a constant K *independent of* h.

Proof: Define the inhomogeneous counterpart to V_h in (47) as

$$V_{h,{\rm d}} = \{u_h \in S_h^{(2)}(\Omega) : \hat{b}(\mu_h, u_h) = \hat{d}_h(\mu_h), \forall \mu_h \in S_{h,{\rm d}}^{(0)}(\Omega)\}.$$

For an arbitrary but fixed $v_h^{\rm d} \in V_{h,{\rm d}}$, set $v_h^0 = u_h^\star - v_h^{\rm d}$ so that by (36), $v_h^0 \in V_h$, $\hat{b}(\lambda_h^\star, v_h^0) = 0$ and

$$a(v_h^0, v_h^0) = a(u_h^\star, v_h^0) - a(v_h^{\rm d}, v_h^0) = -\hat{b}(\lambda_h^\star, v_h^0) - a(v_h^{\rm d}, v_h^0) = -a(v_h^{\rm d}, v_h^0).$$

Then take $v = v_h^0$ in the second equation of (42) and add the result to the right side of the last equation to obtain

$$a(v_h^0, v_h^0) = a(u^\star - v_h^{\rm d}, v_h^0) + \hat{b}(\lambda^\star, v_h^0).$$

Moreover, since $v_h^0 \in V_h$, $\hat{b}(\mu_h, v_h^0) = 0$ holds $\forall \mu_h \in S_{h,{\rm d}}^{(0)}(\Omega)$, and so

$$a(v_h^0, v_h^0) = a(u^\star - v_h^{\rm d}, v_h^0) + \hat{b}(\lambda^\star - \mu_h, v_h^0), \quad \forall \mu_h \in S_{h,{\rm d}}^{(0)}(\Omega).$$

The V_h-coercivity of a (48) and the continuity of a and $\hat{b}$ yield:

$$\alpha\|v_h^0\|_{H^2(\Omega)} \le \|u^\star - v_h^{\rm d}\|_{H^2(\Omega)} + \|\lambda^\star - \mu_h\|_{H_{\rm d}^{-2}(\Omega)}, \quad \forall \mu_h \in S_{h,{\rm d}}^{(0)}(\Omega).$$

Thus,

$$\begin{aligned} \|u^\star - u_h^\star\|_{H^2(\Omega)} &\le \|u^\star - v_h^{\rm d}\|_{H^2(\Omega)} + \|v_h^0\|_{H^2(\Omega)} \\ &\le (1+\alpha^{-1})\|u^\star - v_h^{\rm d}\|_{H^2(\Omega)} + \alpha^{-1}\|\lambda^\star - \mu_h\|_{H_{\rm d}^{-2}(\Omega)} \\ &\qquad \forall \mu_h \in S_{h,{\rm d}}^{(0)}(\Omega) \end{aligned}$$

and since $v_h^{\rm d} \in V_{h,{\rm d}}$ is arbitrary it follows that

$$\|u^\star - u_h^\star\|_{H^2(\Omega)} \le (1+\alpha^{-1}) \inf_{w_h^{\rm d} \in V_{h,{\rm d}}} \|u^\star - w_h^{\rm d}\|_{H^2(\Omega)} + \alpha^{-1} \inf_{\mu_h \in S_{h,{\rm d}}^{(0)}(\Omega)} \|\lambda^\star - \mu_h\|_{H_{\rm d}^{-2}(\Omega)}. \tag{61}$$

It will next be shown that the $V_{h,{\rm d}}$ term in (61) can be so estimated,

$$\inf_{w_h^{\rm d} \in V_{h,{\rm d}}} \|u^\star - w_h^{\rm d}\|_{H^2(\Omega)} \le (1+\beta^{-1}) \inf_{v_h \in S_h^{(2)}(\Omega)} \|u^\star - v_h\|_{H^2(\Omega)} + \beta^{-1}\tau(h)\|u_{\rm d}\|_{H^2(\Omega)}. \tag{62}$$

Now let $\hat{u}_h \in S_h^{(2)}(\Omega)$ be fixed but arbitrary. By Lemma 3, choose $z_h^0 + z_h = \tilde{z}_h \in S_h^{(2)}(\Omega)$, $z_h^0 \in V_h$, $z_h \in V_h^\perp$, so that

$$\hat{b}(\mu_h, z_h) = \hat{b}(\mu_h, \tilde{z}_h) = \hat{b}(\mu_h, u_h^\star - \hat{u}_h), \quad \forall \mu_h \in S_{h,{\rm d}}^{(0)}(\Omega). \tag{63}$$

By Lemma 6, C_h is an isomorphism from $(S_{h,{\rm d}}^{(0)}(\Omega), \|\cdot\|_{H_{\rm d}^{-2}(\Omega)})$ onto $(V_h^\perp, \|\cdot\|_{H^2(\Omega)})$. Hence the estimate,

$$\|C_h\mu_h\|_{H^2(\Omega)} = \sup_{v_h \in V_h^\perp} \frac{(C_h\mu_h, v_h)_{H^2(\Omega)}}{\|v_h\|_{H^2(\Omega)}} = \sup_{v_h \in V_h^\perp} \frac{\hat{b}(\mu_h, v_h)}{\|v_h\|_{H^2(\Omega)}} \ge \beta\|\mu_h\|_{H_{\rm d}^{-2}(\Omega)} \quad \forall \mu_h \in S_{h,{\rm d}}^{(0)}(\Omega)$$

is equivalent to

$$\|C_h^* v_h\|_{H_{\rm d}^{-2}(\Omega)} = \sup_{\mu_h \in S_{h,{\rm d}}^{(0)}(\Omega)} \frac{(\mu_h, C_h^* v_h)_{H_{\rm d}^{-2}(\Omega)}}{\|\mu_h\|_{H_{\rm d}^{-2}(\Omega)}} = \sup_{\mu_h \in S_{h,{\rm d}}^{(0)}(\Omega)} \frac{(C_h \mu_h, v_h)_{H^2(\Omega)}}{\|\mu_h\|_{H_{\rm d}^{-2}(\Omega)}}$$
$$= \sup_{\mu_h \in S_{h,{\rm d}}^{(0)}(\Omega)} \frac{\hat{b}(\mu_h, v_h)}{\|\mu_h\|_{H_{\rm d}^{-2}(\Omega)}} \geq \beta \|v_h\|_{H^2(\Omega)}, \quad \forall v_h \in V_h^\perp.$$

Since $z_h \in V_h^\perp$, it follows from the last estimate together with (42), (45), (30) and (63) that

$$\begin{aligned}
\beta\|z_h\|_{H^2(\Omega)} &\leq \sup_{\mu_h \in S_{h,{\rm d}}^{(0)}(\Omega)} \frac{\hat{b}(\mu_h, z_h)}{\|\mu_h\|_{H_{\rm d}^{-2}(\Omega)}} = \sup_{\mu_h \in S_{h,{\rm d}}^{(0)}(\Omega)} \frac{\hat{b}(\mu_h, u_h^\star - \hat{u}_h)}{\|\mu_h\|_{H_{\rm d}^{-2}(\Omega)}} \\
&= \sup_{\mu_h \in S_{h,{\rm d}}^{(0)}(\Omega)} \frac{\hat{b}(\mu_h, u^\star - \hat{u}_h) + \hat{b}(\mu_h, u_h^\star - u^\star)}{\|\mu_h\|_{H_{\rm d}^{-2}(\Omega)}} \\
&= \sup_{\mu_h \in S_{h,{\rm d}}^{(0)}(\Omega)} \frac{\hat{b}(\mu_h, u^\star - \hat{u}_h) + \hat{d}_h(\mu_h) - \hat{d}(\mu_h)}{\|\mu_h\|_{H_{\rm d}^{-2}(\Omega)}} \\
&\leq \|u^\star - \hat{u}_h\|_{H^2(\Omega)} + \tau(h)\|u_{\rm d}\|_{H^2(\Omega)}.
\end{aligned}$$

Thus, with $w_h^{\rm d} = z_h + \hat{u}_h$ and (45),

$$\hat{b}(\mu_h, w_h^{\rm d}) = \hat{b}(\mu_h, u_h^\star - \hat{u}_h) + \hat{b}(\mu_h, \hat{u}_h) = \hat{b}(\mu_h, u_h^\star) = \hat{d}_h(\mu_h), \quad \forall \mu_h \in S_{h,{\rm d}}^{(0)}(\Omega)$$

so $w_h^{\rm d} \in V_{h,{\rm d}}$. Furthermore,

$$\|u^\star - w_h^{\rm d}\|_{H^2(\Omega)} \leq \|u^\star - \hat{u}_h\|_{H^2(\Omega)} + \|z_h\|_{H^2(\Omega)} \leq (1+\beta^{-1})\|u^\star - \hat{u}_h\|_{H^2(\Omega)} + \beta^{-1}\tau(h)\|u_{\rm d}\|_{H^2(\Omega)}.$$

Since $\hat{u}_h \in S_h^{(2)}(\Omega)$ is arbitrary, (62) follows. It remains to estimate $\|\lambda^\star - \lambda_h^\star\|_{H_{\rm d}^{-2}(\Omega)}$. From the second equation in (42) and the second equation in (45),

$$\hat{b}(\lambda_h^\star - \mu_h, v_h) = a(u^\star - u_h^\star, v_h) + \hat{b}(\lambda^\star - \mu_h, v_h), \quad \forall v_h \in S_h^{(2)}(\Omega), \quad \forall \mu_h \in S_{h,{\rm d}}^{(0)}(\Omega)$$

Using (51) for the left side and the continuity of a and b for the right side,

$$\beta\|\lambda_h^\star - \mu_h\|_{H_{\rm d}^{-2}(\Omega)} \leq \|u^\star - u_h^\star\|_{H^2(\Omega)} + \|\lambda^\star - \mu_h\|_{H_{\rm d}^{-2}(\Omega)}$$

Hence,

$$\|\lambda^\star - \lambda_h^\star\|_{H_{\rm d}^{-2}(\Omega)} \leq \beta^{-1}\|u^\star - u_h^\star\|_{H^2(\Omega)} + (1+\beta^{-1}) \inf_{\mu_h \in S_{h,{\rm d}}^{(0)}(\Omega)} \|\lambda^\star - \mu_h\|_{H_{\rm d}^{-2}(\Omega)} \tag{64}$$

Thus, (60) follows from (61), (62) and (64). ■

Note that $S_h^{(2)}(\Omega)$ is dense in $H^2(\Omega)$, and $S_{h,{\rm d}}^{(0)}(\Omega)$ is dense in $L_{\rm d}^2(\Omega)$ and hence in $H_{\rm d}^{-2}(\Omega)$. Therefore, the desired convergence is obtained from (60).

4 Computational Results

The standard finite element method considered here for the approximation of the primal problem (11) is to compute $u_h \in S_h^{(2)}(\Omega)$ so that

$$\epsilon a(u_h, v_h) + b_h(v_h, u_h) = d_h(v_h), \quad \forall v_h \in S_h^{(2)}(\Omega) \tag{65}$$

where a, b_h and d_h are given by (8), (25), (26), respectively. With $\{s_i\}$ denoting a basis for $S_h^{(2)}(\Omega)$, define the matrices and vectors,

$$A = \{a(s_i, s_j)\}\,, \quad B = \{b_h(s_i, s_j)\}\,, \quad D = \{d_h(s_i)\}\,, \quad \boldsymbol{u}^\epsilon = \{u_i^\epsilon\}$$

where $u_h^\epsilon = \sum_i u_i^\epsilon s_i$. The linear algebraic formulation of (65) is given by

$$(\epsilon A + B)\boldsymbol{u}^\epsilon = D \tag{66}$$

which can also be written in primal-dual form as follows with the dual variable $\lambda_h^\epsilon = \sum_i \lambda_i^\epsilon s_i$, $\boldsymbol{\lambda}^\epsilon = \{\lambda_i^\epsilon\}$,

$$\begin{bmatrix} -\epsilon B & B \\ B & A \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda}^\epsilon \\ \boldsymbol{u}^\epsilon \end{bmatrix} = \begin{bmatrix} D \\ 0 \end{bmatrix} \tag{67}$$

Representative results are shown in Fig. 2 for the one-dimensional case discussed there. Solutions to (66) are shown in Fig. 2a for $\epsilon = \mathcal{O}(1)$ and in Fig. 2b for $\epsilon \ll 1$. The result for $\epsilon \ll 1$ evidently exhibits discontinuous derivatives, and it retains this appearance robustly with respect to h. These discontinuous derivatives apparently depart from a theoretical result analogous to Theorem 8 above that the limit function $u_h^\epsilon \xrightarrow{\epsilon \to 0} u_h^\star$ converges to the continuum solution $u^\star \in H^2(\Omega)$ with grid refinement. This hypothesis may be tested directly with counterparts to (36) – (38) instead of choosing ϵ vanishingly small in (66). For this, let $u_h^\star = \sum_i u_i^\star s_i$, $\boldsymbol{u}^\star = \{u_i^\star\}$ be the primal variable and $\lambda_h^\star = \sum_i \lambda_i^\star s_i$, $\boldsymbol{\lambda}^\star = \{\lambda_i^\star\}$ the dual variable in the saddle point problem

$$\begin{bmatrix} (I - BB^\dagger) & B \\ B & \omega A \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda} \\ \boldsymbol{u} \end{bmatrix} = \begin{bmatrix} D \\ 0 \end{bmatrix} \tag{68}$$

which characterizes the stationary point for the Lagrangian functional,

$$L(\boldsymbol{u}, \boldsymbol{\lambda}) = \tfrac{1}{2}\boldsymbol{u}^{\mathrm{T}} A \boldsymbol{u} + \boldsymbol{\lambda}^{\mathrm{T}}(B\boldsymbol{u} - D), \quad (I - BB^\dagger)\boldsymbol{\lambda} = 0$$

while $\boldsymbol{u}^\star$ solves

$$\min_{\boldsymbol{u}} \boldsymbol{u}^{\mathrm{T}} A \boldsymbol{u} \quad \text{subject to} \quad B\boldsymbol{u} = D.$$

Here, $B^\dagger$ denotes the pseudo-inverse of B. Since $(I - BB^\dagger)$ is the orthogonal projector onto the kernel $\mathcal{K}$ of $B = B^{\mathrm{T}}$, the complementarity condition $(I - BB^\dagger)\boldsymbol{\lambda} = 0$ ensures that $\boldsymbol{\lambda}$ has no component on $\mathcal{K}$. The numerical result is identical to that shown in Fig. 2b. On the other hand, when the data are approximated at least with pixelwise linear functions, solutions to (65) ($\epsilon \ll 1$) and (68) are identical to that shown in Fig. 2c.

The implementation of (27) will now be considered computationally. With bases for $S_{h,\mathrm{d}}^{(0)}(\Omega)$ and $S_h^{(2)}(\Omega)$ denoted by $\{\chi_i\}$ and $\{s_i\}$, respectively, define the matrices,

$$A = \{a(s_i, s_j)\}\,, \quad B = \{b_h(\chi_i, s_j)\}\,, \quad C = \{b_h(\chi_i, \chi_j)\}$$

and vectors,

$$D = \{d_h(\chi_i)\}\,, \quad \boldsymbol{u}^\epsilon = \{u_i^\epsilon\}\,, \quad \boldsymbol{\lambda}^\epsilon = \{\lambda_i^\epsilon\}$$

with $u_h^\epsilon = \sum_i u_i^\epsilon s_i$ and $\lambda_h^\epsilon = \sum_i \lambda_i^\epsilon \chi_i$ so that (27) may be written as

$$\begin{bmatrix} -\epsilon C & B \\ B^{\mathrm{T}} & A \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda}^\epsilon \\ \boldsymbol{u}^\epsilon \end{bmatrix} = \begin{bmatrix} D \\ 0 \end{bmatrix} \tag{69}$$

Since $B = CP$, where

$$P = \{(\chi_i, s_j)_{L^2(\Omega)}\}$$

is the projection of $S_h^{(2)}(\Omega)$ onto $S_h^{(0)}(\Omega)$, $\boldsymbol{\lambda}$ can be eliminated from (69) in the following form,

$$(\epsilon A + P^{\mathrm{T}} C P)\boldsymbol{u} = P^{\mathrm{T}} D \tag{70}$$

Representative results are shown in Fig. 2 for the one-dimensional case discussed there. Solutions to (69) are shown in Fig. 2a for $\epsilon = \mathcal{O}(1)$ and in Fig. 2c for $\epsilon \ll 1$. The result for $\epsilon \ll 1$ evidently agrees with the vanishing regularization solution $u^\star$, and it retains this appearance robustly with respect to h. This result is apparently consistent with the theoretical result of Theorem 8 above that the limit function $u_h^\epsilon \stackrel{\epsilon\to 0}{\longrightarrow} u_h^\star$ converges to the continuum solution $u^\star \in H^2(\Omega)$ with grid refinement. This hypothesis may be tested directly with counterparts to (36) – (38) instead of choosing ϵ vanishingly small in (69). For this, let $u_h^\star = \sum_i u_i^\star s_i$, $\boldsymbol{u}^\star = \{u_i^\star\}$ be the primal variable and $\lambda_h^\star = \sum_i \lambda_i^\star s_i$, $\boldsymbol{\lambda}^\star = \{\lambda_i^\star\}$ the dual variable in the saddle point problem

$$\begin{bmatrix} (I - BB^\dagger) & B \\ B^{\mathrm{T}} & A \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda}^\star \\ \boldsymbol{u}^\star \end{bmatrix} = \begin{bmatrix} D \\ 0 \end{bmatrix} \tag{71}$$

which characterizes the stationary point for the Lagrangian functional,

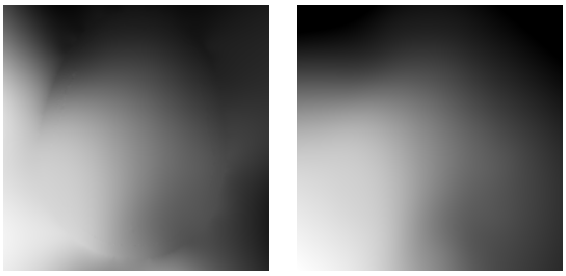
$$L(\boldsymbol{u}, \boldsymbol{\lambda}) = \tfrac{1}{2}\boldsymbol{u}^{\mathrm{T}} A \boldsymbol{u} + \boldsymbol{\lambda}^{\mathrm{T}}(B\boldsymbol{u} - D), \quad (I - BB^\dagger)\boldsymbol{\lambda} = 0$$

while $\boldsymbol{u}^\star$ solves

$$\min_U \boldsymbol{u}^{\mathrm{T}} A \boldsymbol{u} \quad \text{subject to} \quad B\boldsymbol{u} = D$$

Here, $B^\dagger$ denotes the pseudo-inverse of B. Since $(I - BB^\dagger)$ is the orthogonal projector onto the kernel $\mathcal{K}$ of B^{T}, the complementarity condition $(I - BB^\dagger)\boldsymbol{\lambda} = 0$ ensures that $\boldsymbol{\lambda}$ has no component on $\mathcal{K}$. The numerical result is identical to that shown in Fig. 2c.

The standard finite element method (66) and the mixed finite element method (69) are now compared for the two-dimensional example of Figs. 1 and 3. So that an exact answer



(a) standard **(b)** mixed

Figure 3: Vanishing regularization ($\epsilon \ll 1$) estimations of u in Fig. 1c using (a) the standard finite element method (66) and (b) the mixed finite element method (69).

may be known, $\tilde{u}$ in Fig. 1b is given by the product of κ in Fig. 1a with u in Fig. 1c. For vanishingly small regularization, $\epsilon \ll 1$, the solutions to (66) and (69) are shown in Figs. 3a and 3b, respectively. Quantitative measures of the estimation error are several orders of magnitude larger for the result in Fig. 3a in relation to that in Fig. 3b, but the advantage of the mixed method in relation to the standard method is visually evident.

In an effort to understand the success of the mixed method in relation to the standard method, one notes especially the following important clue. As reported above, the results from the standard method mimic those of the mixed method if the data are represented as piecewise linear. Apparently, if the data are represented with sufficient accuracy, convergence of a regularized solution to a vanishing regularization solution can be obtained, which is consistent with the well-known result cited in Section 1 concerning the rates at which regularization and noise may be simultaneously reduced [14]. Evidently, when the data are approximated as piecewise constant, the standard method senses this approximation in the forms b_h and d_h unavoidably as inaccurate data. On the other hand, the mixed method does not sense the cellwise constant approximation of data as inaccurate, particularly in the forms d_h or $\hat{d}_h$, because the test functions are cellwise constant and satisfy

$$\hat{d}(\mu_h) = \int_\Omega \mu_h \chi_{\rm d} u_{\rm d} = \int_\Omega \mu_h \tilde{u}/\kappa = \int_\Omega \mu_h P^{(0)}_{h,{\rm d}}(\tilde{u}/\kappa) \approx \int_\Omega \mu_h \tilde{u}_h/\kappa_h = \hat{d}_h(\mu_h), \quad \forall \mu_h \in S^{(0)}_{h,{\rm d}}(\Omega) \tag{72}$$

Here $P^{(0)}_{h,{\rm d}}$ is the $L^2(\Omega)$ projection onto $S^{(0)}_{h,{\rm d}}(\Omega)$ and the approximation of the data $P^{(0)}_{h,{\rm d}}(\tilde{u}/\kappa) \approx P^{(0)}_{h,{\rm d}}\tilde{u}/P^{(0)}_{h,{\rm d}}\kappa = \tilde{u}_h/\kappa_h$ for (30) and (72) is apparently excellent. Although the projection $P^{(0)}_{h,{\rm d}}$ cannot be used similarly for the form b_h, the mixed method permits (18) and (36) with b_h and d_h to be reformulated as (42) and (45) with $\hat{b}$ and $\hat{d}_h$, where $\hat{b}$ contains no data approximation. For the standard method, the forms b_h and d_h in (65) cannot be reformulated in a way which mimics (39) and (43). Yet, for the example presented in Fig. 2, $b_h = b$ holds for all methods presented and thereby emphasizes the importance of the approximation seen in (72). For the standard method, a counterpart to (72) cannot be written in terms of a projection which is friendly to a cellwise constant approximation of data. As a result, the counterpart to (30) apparently does not hold for the standard method, which prevents the desired convergence.

In light of these observations, penalized discontinuous Galerkin methods have also been considered for this work. The general weak formulation follows that of [2], where the boundary integrals are not included here because of the natural boundary conditions. Specifically, for the one-dimensional example of Fig. 2, $\Omega = (0, 1)$ is discretized with cell interfaces $x_i = ih$, $i = 0, \dots, N$, $h = 1/N$ as described in Section 3, and the bilinear form a is modified to obtain $a_0 + a_\gamma$, defined as follows using the notation $\{u\} \equiv \frac{1}{2}(u_- + u_+)$ and $[v] \equiv v_- - v_+$,

$$\begin{aligned} a_0(u, v) &= \sum_{k=1}^{N} \int_{x_{k-1}}^{x_k} u_{xx} v_{xx} + \sum_{k=1}^{N-1} \Big\{ \{u_{xxx}\}[v] + \{v_{xxx}\}[u] - \{u_{xx}\}[v_x] - \{v_{xx}\}[u_x] \Big\}_{x_k} \\ a_\gamma(u, v) &= \gamma \sum_{k=1}^{N-1} \Big\{ h^{-3}[u][v] + h^{-1}[u_x][v_x] \Big\}_{x_k} \end{aligned}$$

While other polynomial bases have also been tested with identical results, let the Chebychev polynomials

$$P_0(t) = 1, \quad P_1(t) = t, \quad P_2(t) = 2x^2 - 1, \quad P_3(t) = 2x(2x^2 - 1) - x,$$

define the piecewise polynomial basis,

$$\{\beta_k : k = 1, \dots, 4N\} = \{P_j(2(x - x_i)/h - 1) : i = 1, \dots, N, j = 0, \dots, 3\}$$

With this basis define the matrices

$$A = A_0 + A_\gamma = \{a_0(\beta_i, \beta_j)\} + \{a_{\boldsymbol{\gamma}}(\beta_i, \beta_j)\} \in \mathbb{R}^{4N \times 4N},$$

and in terms of (25) and (26),

$$B = \{b_h(\beta_i, \beta_j)\} \in \mathbb{R}^{4N \times 4N}, \quad D = \{d_h(\beta_i)\} \in \mathbb{R}^{4N}$$

so that the approximation $u^\epsilon_h = \sum_{k=1}^{4N} u^\epsilon_k \beta_k$, $\boldsymbol{u}^\epsilon = \{u^\epsilon_k\}_{k=1}^{4N}$, to u^ϵ is determined as the solution to

$$[\epsilon A + B]\boldsymbol{u}^\epsilon = D.$$

Since the cellwise constant data are in the span of P_0 and orthogonal to $\{P_j\}_{j=1}^3$, only the P_0 modes are non-trivial in the form D. Also, only the P_0 modes are excited in the solution $\boldsymbol{u}^\epsilon$ when ϵ is especially small and γ is just large enough for $\epsilon A + B$ to be positive definite. In this case, $\boldsymbol{u}^\epsilon$ is roughly in the kernel of A_0, and A_γ becomes the discrete Laplacian with natural boundary conditions. As a result, regularization is effectively only first order, and the numerical solution u_h^ϵ is constant on $\Omega \backslash S_\mathrm{d}$, departing from all results shown in Fig. 2. On the other hand, if the data are approximated as piecewise linear, i.e., in the span of $\{P_0, P_1\}$ and orthogonal to $\{P_j\}_{j=2}^3$, then the P_0 and P_1 modes are non-trivial in the form D and excited in the solution for ϵ and γ as described above. As a result, the data approximation is sufficiently accurate and the numerical solution resembles the correct one shown in Fig. 2c. If the data are only approximated as cellwise constant, then γ can be chosen sufficiently large to obtain a correct result such as seen in Fig. 2c. However, the correct γ for such a result depends upon ϵ. Furthermore, as γ is increased ever larger, then the discontinuous Galerkin method converges to the standard finite element method, and the result resembles that shown in Fig. 2a. Instead of choosing ϵ vanishingly small to approximate the vanishing regularization solution, the primal-dual and saddle point formulations may be used as for the standard and mixed finite element methods, but the results agree with the cases described here for ϵ chosen vanishingly small.

5 Conclusions

The model problem demonstrated in Fig. 1 and formulated in (6) has been used to represent a typical image processing problem of reconstructing an unknown in the face of incomplete data. A discretization of (6) is defined to be consistent with respect to vanishing regularization if the numerical solution u_h^ϵ exhibits the convergence $u_h^\epsilon \stackrel{\epsilon \to 0}{\longrightarrow} u_h^\star$ as well as $u_h^\star \stackrel{\epsilon \to 0}{\longrightarrow} u^\star$, where the continuum solution u^ϵ to (6) exhibits the convergence $u^\epsilon \stackrel{\epsilon \to 0}{\longrightarrow} u^\star$. Since vanishing regularization is of interest only when data fidelity is sufficient, it is assumed that the data are smooth enough that the continuum limit $u^\star$ has at least the regularity of the space for the weak formulation in (11). As the primal characterization of u^ϵ in (11) permits a primal-dual formulation in (15), the vanishing regularization limit $u^\star$ may be characterized explicitly with the saddle point formulation in (18). A mixed finite element discretization of (15) is given in (27), whose vanishing regularization limit is given in (36), the counterpart to (15). The numerical solutions u_h^ϵ and $u_h^\star$ for the mixed method are shown in Theorem 8 to be consistent in the sense defined here. This consistency is also demonstrated computationally in Section 4 in terms of an artificial example and a realistic example with magnetic resonance images. On the other hand, it is demonstrated computationally that a standard finite element discretization is not consistent. The reason for the relative success of the mixed method is explained by a projection property which is friendly to cellwise constant data approximation, making the forms used for the method nearly exact. Since the counterpart forms used for the standard method suffer from significant approximation errors, the method requires that grid refinement be coordinated with the reduction of regularization. Experiments with a penalized discontinuous Galerkin method are consistent with these observations for the mixed and standard methods.

References

[1] R.A. Adams, *Sobolev Spaces*, Academic Press, New York, London, Toronto, Sydney, San Francisco, 1975.

[2] G.A. Baker, *Finite Element Methods for Elliptic Equations Using Nonconforming Elements*, Mathematics of Computation, Vol. 31, No. 137, pp. 45–59, January 1977.

[3] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.

[4] I. Ekeland and R. Témam, *Convex Analysis and Variational Problems*, Vol. 28 of Classics in Applied Mathematics, SIAM, Philadelphia, PA, English ed., 1999. (Translated from French.)

[5] L.C. Evans and R.F. Gariepy, *Measure Theory and Fine Properties of Functions*, Studies in Advanced Mathematics, CRC Press, Boca Raton, 1992.

[6] G.J. Fix, *On the effects of Quadrature in the Finite Element Method*, in *Advances in Computational Methods in Structural Mechanics and Design*, (J.T. Oden, R.W. Clough, Y. Yamamoto, Editors), pp. 44–68, The University of Alabana Press, Huntsville, 1972.

[7] V. Girault and P.-A. Raviart, *Finite Element Methods for Navier-Stokes Equations*, Springer-Verlag, Berlin, 1980.

[8] W. Hackbusch, R. Fadiman and P.D.F. Ion, *Elliptic Differential Equations: Theory and Numerical Treatment*, Springer Series in Computational Mathematics, Springer, 2010.

[9] R.J. Herbold and R.S. Varga, *The Effect of Quadrature Errors in the Numerical Solution of Two-Dimensional Boundary Values Problems by Variational Techniques*, Aequationes Math., Vol. 7, pp. 36–58, 1972.

[10] R.J. Herbold, M.H. Schultz and R.S. Varga, *The Effect of Quadrature Errors in the Numerical Solution of Boundary Value Problems by Variational Techniques*, Aequationes Math., Vol. 3, pp. 247–270, 1969.

[11] S. Fürtinger, *An Approach to Computing Binary Edge Maps for the Purpose of Registering Intensity Modulated Images*, Ph.D. Dissertation, Institute for Mathematics and Scientific Computing, Karl Franzens University of Graz, 2012.

[12] S.L. Keeling, *Generalized Rigid and Generalized Affine Image Registration and Interpolation by Geometric Multigrid*, Journal of Mathematical Imaging and Vision, Vol. 29, pp. 163–183, 2007.

[13] S.L. Keeling and G. Haase, *Geometric Multigrid for High-Order Regularizations of Early Vision Problems*, Applied Mathematics and Computation, Vol. 184, pp. 536-556, 2007.

[14] O. Scherzer, H. Grossauer, F. Lenzen, M. Grasmair and M. Haltmeier, *Variational Methods in Imaging*, Springer, 2009.

[15] C. Grossmann, H.G. Roos and M. Stynes, *Numerical Treatment of Partial Differential Equations*, Springer, 2005.