ERROR ANALYSIS OF FINITE ELEMENT APPROXIMATIONS OF DIFFUSION COEFFICIENT IDENTIFICATION FOR ELLIPTIC AND PARABOLIC PROBLEMS[*]

BANGTI JIN[†] AND ZHI ZHOU[‡]

Abstract. In this work, we present a novel error analysis for recovering a spatially dependent diffusion coefficient in an elliptic or parabolic problem. It is based on the standard regularized output least-squares formulation with an $H^1(\Omega)$ seminorm penalty and then discretized using the Galerkin finite element method with conforming piecewise linear finite elements for both state and coefficient and backward Euler in time in the parabolic case. We derive a priori weighted $L^2(\Omega)$ estimates where the constants depend only on the given problem data for both elliptic and parabolic cases. Further, these estimates also allow deriving standard $L^2(\Omega)$ error estimates under a positivity condition that can be verified for certain problem data. Numerical experiments are provided to complement the error analysis.

Key words. parameter identification, finite element approximation, error estimate, Tikhonov regularization

AMS subject classifications. 65M30, 65M15, 65M60

DOI. 10.1137/20M134383X

1. Introduction. This work is concerned with error analysis of Galerkin approximations of regularized formulations for recovering a spatially dependent diffusion coefficient q for elliptic and parabolic problems. Let $\Omega \subset \mathbb{R}^d$ $(d = 1, 2, 3)$ be a convex polyhedral domain with a boundary $\partial\Omega$. Consider the following elliptic boundary value problem:

$$(1.1) \qquad \begin{cases} -\nabla \cdot (q\nabla u) = f & \text{in } \Omega, \\ \qquad\qquad u = 0 & \text{on } \partial\Omega, \end{cases}$$

where the function f denotes a given source term. The solution to problem (1.1) is denoted by $u(q)$ to indicate its dependence on the coefficient q. The inverse problem is to recover the exact diffusion coefficient $q^\dagger(x)$ from the pointwise observation z^δ, with a noise level δ, i.e.,

$$(1.2) \qquad \|z^\delta - u(q^\dagger)\|_{L^2(\Omega)} \leq \delta.$$

Throughout, the diffusion coefficient q is sought within the admissible set $\mathcal{A}$, defined by

$$(1.3) \qquad \mathcal{A} = \{q \in H^1(\Omega) : \ c_0 \leq q \leq c_1 \text{ a.e. in } \Omega\}$$

for some positive constants $c_0, c_1 > 0$.

[†]Department of Computer Science, University College London, Gower Street, London WC1E 6BT, UK (b.jin@ucl.ac.uk, bangti.jin@gmail.com).

[‡]Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong (zhizhou@polyu.edu.hk).

119

Problem (1.1) is the steady state of the following parabolic initial-boundary value problem:

$$(1.4) \quad \begin{cases} \partial_t u - \nabla \cdot (q \nabla u) = f & \text{in } \Omega \times (0, T], \\ u(0) = u_0 & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega \times (0, T], \end{cases}$$

where $T > 0$ is the final time. The functions f and u_0 are the given source term and initial condition, respectively. The corresponding inverse problem is to recover the spatially dependent diffusion coefficient $q^\dagger(x)$ from the distributed observation z^δ over $\Omega \times (T - \sigma, T)$ (for some measurement window $0 \le \sigma < T$), with a noise level δ, i.e.,

$$(1.5) \quad \|z^\delta - u(q^\dagger)\|_{L^2(T-\sigma,T;L^2(\Omega))} \le \delta.$$

The elliptic problem (1.1) and parabolic problem (1.4) describe many important physical processes, and the related inverse problems are exemplary for parameter identifications for PDEs (see the monographs [6, 22] for overviews). For example, (1.1) is often used to model the behavior of a confined inhomogeneous aquifer, where u represents the piezometric head, f is the recharge, and q is hydraulic conductivity (or transmissivity in the two-dimensional case); see [18, 36] for parameter identifications in hydrology. See also [5] for related coupled-physics inverse problems arising in medical imaging.

Due to the ill-posed nature of inverse problems, regularization, especially Tikhonov regularization, is customarily employed for constructing numerical approximations (see, e.g., [13, 23]). Commonly used stabilizing terms include $H^1(\Omega)$ and total variation seminorms, which are suitable for recovering smooth and nonsmooth diffusion coefficients, respectively. The well-posedness and convergence (with respect to the noise level) was studied [1, 9, 20, 29], and, further, convergence rates (with respect to δ) were derived under various "source" conditions, e.g., variational inequalities or conditional stability estimates [27]. In practice, the regularized formulations are further discretized, often with the Galerkin finite element method (FEM), due to its flexibility with domain geometry and low-regularity problem data. The discretization step necessarily introduces additional errors which impact the reconstruction quality. Several studies [20, 27, 37] have analyzed the convergence with respect to the discretization parameter(s), e.g., mesh size h and time step size τ, but without error bounds.

So far, only very few results were available on error bounds of approximate solutions. This is attributed to strong nonlinearity of the forward (parameter-to-state) map, low regularity of noisy data z^δ, and delicate interplay between different parameters, (noise level, regularization parameter, and discretization parameters). Falk [16] analyzed a Galerkin discretization of the standard output least-squares formulation for the elliptic inverse problem (with a Neumann boundary condition) and derived a rate $O(h^r + h^{-2}\delta)$ in the $L^2(\Omega)$ norm, where r is the polynomial degree of the finite element space and h is the mesh size. This result is derived by assuming sufficiently high regularity of the coefficient $q^\dagger$ and a certain structural condition on the gradient field; see details in Remark 3.2. In the elliptic case, there are also several results for other discrete formulations: [32] for upwind finite difference approximation of a transport equation (without noise), [3, 26] for the equation error approach (EEA) (the fidelity in the negative $H^1(\Omega)$ norm and $H^1(\Omega)$ penalty), and [28] for the EEA in a mixed formulation. However, no regularization was taken into account in the works [16, 28, 32], and thus the corresponding discrete formulations can suffer from

numerical instability. The EEA works only with the case $z^\delta \in H^1(\Omega)$, and so is the error analysis. For the regularized problem, Wang and Zou [35] derived first convergence rates (in weighted norms) for both elliptic and parabolic cases (equipped with a zero Neumann boundary condition) with either pointwise or gradient observations. In the elliptic case, the analysis employs the test function $\varphi = \frac{q^\dagger - q_h^*}{q^\dagger} e^{-2\alpha_0 c_0^{-1} u(q^\dagger)}$ (with q_h^* being a discrete minimizer, α_0 a parameter in the structural condition; cf. Remark 3.2 and c_0 lower bound on $q^\dagger$) and assumes regularity on both state u and coefficient $q^\dagger$; in the parabolic case, it requires a more involved test function. However, no estimate in the usual $L^2(\Omega)$ was given, and, further, the analysis in the parabolic case requires the measurement in the entire time interval $(0, T)$. Deckelnick and Hinze [10] studied the elliptic inverse problem of recovering matrix valued coefficients using the $L^2(\Omega)$ penalty in the H-convergence framework and in the two-dimensional case proved an $L^2(\Omega)$ estimate $O(\delta^{\frac{1}{2}})$, where the coefficient q is discretized using variational discretization. The estimate was derived under a projected source condition.

In this work, we present a novel approach to derive convergence rates for the standard regularized output least-squares formulation discretized by Galerkin FEM. The approach employs the test function $\varphi = \frac{q^\dagger - q_h^*}{q^\dagger} u(q^\dagger)$ for both elliptic and parabolic cases, inspired by the recent work [7] (on the Hölder stability of the elliptic inverse problems). It enables us to derive convergence rates in a new weighted $L^2(\Omega)$ norm for both elliptic and parabolic cases, extending the prior result for the time-fractional diffusion equation [25]. Further, we derive estimates in the usual $L^2(\Omega)$ norm, under suitable positivity conditions, which hold for a class of problem data. In the parabolic case, we relax the restriction in [35] (and also [25]) on the time horizon for the measurement from $[0, T]$ to a subinterval $[T - \sigma, T]$ for any $0 \le \sigma < T$ and the regularity assumption on the true coefficient $q^\dagger$ from $W^{2,\infty}(\Omega)$ to $W^{1,\infty}(\Omega) \cap H^2(\Omega)$. This former is achieved by a new weighting in the time direction and the latter by discrete maximal L^p regularity for parabolic problems. In the course of error analysis, no regularity assumption is made on the state u and no additional temporal regularity on the observation z^δ than $L^2(T - \sigma, T; L^2(\Omega))$, and, furthermore, no source type condition is imposed, as is usually done for parameter identifications [14, 27]. To the best of our knowledge, they are first error estimates of the kind for the concerned inverse conductivity problems.

The rest of the paper is organized as follows. In section 2, we describe useful facts about the Galerkin FEM. Then in sections 3 and 4, we describe and analyze the finite element approximations for the elliptic and parabolic inverse problems, respectively. Finally, in section 5, we present numerical results to complement the analysis. We conclude with useful notation. For any $k \ge 0$ and $p \ge 1$, the space $W^{k,p}(\Omega)$ denotes the standard Sobolev spaces of the kth order, and we write $H^k(\Omega)$, when $p = 2$ [2]. The notation $(\cdot, \cdot)$ denotes the $L^2(\Omega)$ inner product. For the analysis of parabolic problems, we use the Bochner spaces $W^{k,p}(0, T; B)$ etc., with B being a Banach space. Throughout, the notation c, with or without a subscript, denotes a generic constant which may change at each occurrence, but it is always independent of the following parameters: regularization parameter γ, mesh size h, time step size τ, and noise level δ.

2. Finite element approximations. Now we recall briefly the Galerkin FEM approximation. Let $\mathcal{T}_h$ be a shape regular quasi-uniform triangulation of the domain Ω into d-simplexes, denoted by T, with a mesh size h. Over $\mathcal{T}_h$, we define a continuous piecewise linear finite element space X_h by

$$X_h = \left\{ v_h \in H_0^1(\Omega) : \ v_h|_T \text{ is a linear function } \forall T \in \mathcal{T}_h \right\}$$

and similarly the space V_h by

$$V_h = \left\{ v_h \in H^1(\Omega) : \ v_h|_T \text{ is a linear function } \forall T \in \mathcal{T}_h \right\}.$$

The spaces X_h and V_h will be employed to approximate the state u and the diffusion coefficient q, respectively. First, we introduce useful operators on X_h and V_h. We define the $L^2(\Omega)$ projection $P_h : L^2(\Omega) \to X_h$ by

$$(P_h \varphi, \chi) = (\varphi, \chi) \quad \forall \chi \in X_h.$$

Note that the operator P_h satisfies the following error estimates [34, p. 32]: For any $s \in [1, 2]$,

$$(2.1) \quad \|P_h \varphi - \varphi\|_{L^2(\Omega)} + h\|\nabla(P_h \varphi - \varphi)\|_{L^2(\Omega)} \le h^s \|\varphi\|_{H^s(\Omega)} \quad \forall \varphi \in H^s(\Omega) \cap H_0^1(\Omega).$$

Let $\mathcal{I}_h$ be the Lagrange interpolation operator associated with the finite element space V_h. Then it satisfies the following error estimates for $s = 1, 2$ and $1 \le p \le \infty$ (with $sp > d$) [15, Theorem 1.103]:

$$(2.2) \quad \|v - \mathcal{I}_h v\|_{L^p(\Omega)} + h\|v - \mathcal{I}_h v\|_{W^{1,p}(\Omega)} \le ch^s \|v\|_{W^{s,p}(\Omega)} \quad \forall v \in W^{s,p}(\Omega).$$

Further, for any q, we define a discrete operator $A_h(q) : X_h \to X_h$ by

$$(2.3) \quad (A_h(q)v_h, \chi) = (q\nabla v_h, \nabla \chi) \quad \forall v_h, \chi \in X_h.$$

3. Elliptic case. In this section, we derive error estimates for the elliptic inverse problem.

3.1. Finite element approximation. First, we describe the regularized formulation and its finite element approximation. To recover the diffusion coefficient q in the elliptic system (1.1), we employ the standard output least-squares formulation with an $H^1(\Omega)$ seminorm penalty,

$$(3.1) \quad \min_{q \in \mathcal{A}} J_\gamma(q) = \frac{1}{2}\|u(q) - z^\delta\|_{L^2(\Omega)}^2 + \frac{\gamma}{2}\|\nabla q\|_{L^2(\Omega)}^2,$$

where the admissible set $\mathcal{A}$ is defined by (1.3) and $u(q) \in H_0^1(\Omega)$ satisfies the variational problem

$$(3.2) \quad (q\nabla u(q), \nabla v) = (f, v) \quad \forall v \in H_0^1(\Omega).$$

The $H^1(\Omega)$ seminorm penalty is suitable for recovering a smooth diffusion coefficient. The scalar $\gamma > 0$ is the regularization parameter, controlling the strength of the penalty [23]. Using standard argument in calculus of variation, it can be verified that for every $\gamma > 0$, problem (3.1)–(3.2) has at least one global minimizer q^*, and, further, the sequence of minimizers converges subsequentially in $H^1(\Omega)$ to a minimum seminorm solution as the noise level δ tends to zero provided that γ is chosen appropriately in accordance with δ, i.e., $\lim_{\delta \to 0^+} \gamma(\delta)^{-1}\delta^2 = \lim_{\delta \to 0^+} \gamma(\delta) = 0$; see, e.g., [14, 23]. In this work, we focus on the a priori choice $\gamma \sim \delta^2$ (cf. Remark 3.3 below), which is generally sufficient to ensure the noise level condition (1.2). In practice, one may also employ a posteriori rules. One popular choice is the discrepancy principle [23, 31]: Given some $\tau > 1$, it determines the largest $\gamma > 0$ such that

$$\|u(q^*) - z^\delta\|_{L^2(\Omega)} \le \tau\delta,$$

in line with the a priori knowledge (1.2).

Now we can formulate the finite element discretization of problem (3.1)–(3.2),

$$(3.3) \qquad \min_{q_h \in \mathcal{A}_h} J_{\gamma,h}(q_h) = \frac{1}{2}\|u_h(q_h) - z^\delta\|_{L^2(\Omega)}^2 + \frac{\gamma}{2}\|\nabla q_h\|_{L^2(\Omega)}^2,$$

subject to $q_h \in \mathcal{A}_h$ and $u_h(q_h)$ satisfying

$$(3.4) \qquad (q_h \nabla u_h(q_h), \nabla v_h) = (f, v_h) \quad \forall v_h \in X_h.$$

The discrete admissible set $\mathcal{A}_h$ is taken to be

$$(3.5) \qquad \mathcal{A}_h := \mathcal{A} \cap V_h = \{q_h \in V_h : \ c_0 \le q_h(x) \le c_1 \text{ in } \Omega\}.$$

For the discrete problem (3.3)–(3.4), the following existence and convergence results hold. For any fixed $h > 0$, there exists at least one minimizer q_h^* to problem (3.3)–(3.4). Further, the sequence of minimizers $\{q_h^*\}_{h>0}$ contains a subsequence that converges in $H^1(\Omega)$ to a minimizer to problem (3.1)–(3.2). The proof follows by a standard argument from calculus of variation and the density of the space V_h in $H^1(\Omega)$ and thus is omitted; see [21, 37] for related analysis.

3.2. Error estimates. Now we establish an error estimate of the numerical approximation (3.3)–(3.4) with respect to the exact conductivity $q^\dagger$. We shall make the following assumption on the problem data.

Assumption 3.1. The exact diffusion coefficient $q^\dagger$ and source term f satisfy $q^\dagger \in H^2(\Omega) \cap W^{1,\infty}(\Omega) \cap \mathcal{A}$ and $f \in L^\infty(\Omega)$.

Under Assumption 3.1, there holds (see [30, Lemma 2.1] and [19, Theorems 3.3 and 3.4])

$$(3.6) \qquad u \in H_0^1(\Omega) \cap H^2(\Omega) \cap W^{1,\infty}(\Omega).$$

Note that this regularity result requires only $q^\dagger \in W^{1,\infty}(\Omega) \cap \mathcal{A}$.

The following a priori estimate holds. The proof is identical with that for [35, Lemma 5.2] but with the Dirichlet boundary condition in place of the Neumann one (see also Lemma A.1 for related argument). The proof requires the estimate $\|q - \mathcal{I}_h q\|_{L^2(\Omega)} \le ch^2$ due to the use of $u_h(\mathcal{I}_h q)$ as an intermediate solution and thus the condition $q^\dagger \in H^2(\Omega)$ in Assumption 3.1. See the proof in Lemma A.1 and [35, Lemma 5.2] for details.

LEMMA 3.1. *Let $q^\dagger \in \mathcal{A}$ be the exact diffusion coefficient, $u(q^\dagger)$ the solution to problem* (3.2), *and Assumption* 3.1 *be fulfilled. Let $(q_h^*, u_h(q_h^*)) \in \mathcal{A}_h \times X_h$ be a solution of problem* (3.3)–(3.4). *Then there holds*

$$\|u_h(q_h^*) - u(q^\dagger)\|_{L^2(\Omega)} + \gamma^{\frac{1}{2}}\|\nabla q_h^*\|_{L^2(\Omega)} \le c(h^2 + \delta + \gamma^{\frac{1}{2}}).$$

Now we state the main result of this section, i.e., a weighted error estimate for the Galerkin approximation q_h^*. The positivity of the weight $q^\dagger|\nabla u(q^\dagger)|^2 + fu(q^\dagger)$ will be discussed below.

THEOREM 3.2. *Let Assumption* 3.1 *be fulfilled. Let $q^\dagger$ be the exact diffusion coefficient, $u(q^\dagger)$ the solution to problem* (3.2), *and $q_h^* \in \mathcal{A}_h$ a minimizer of problem* (3.3)–(3.4). *Then with $\eta = h^2 + \delta + \gamma^{\frac{1}{2}}$, there holds*

$$\int_\Omega (q^\dagger - q_h^*)^2 \big(q^\dagger|\nabla u(q^\dagger)|^2 + fu(q^\dagger)\big)\, dx \le c(h\gamma^{-\frac{1}{2}}\eta + \min(h + h^{-1}\eta, 1))\gamma^{-\frac{1}{2}}\eta.$$

Proof. For any test function $\varphi \in H_0^1(\Omega)$, we have the following splitting (with $u = u(q^\dagger)$):

$$((q^\dagger - q_h^*)\nabla u, \nabla \varphi) = ((q^\dagger - q_h^*)\nabla u, \nabla(\varphi - P_h\varphi)) + (q^\dagger \nabla u - q_h^* \nabla u, \nabla P_h \varphi).$$

Applying integration by parts and the variational formulations of u and $u_h(q_h^*)$ to the first and second terms, respectively, leads to

$$((q^\dagger - q_h^*)\nabla u, \nabla \varphi) = -(\nabla \cdot ((q^\dagger - q_h^*)\nabla u), \varphi - P_h\varphi) + (q_h^* \nabla(u_h(q_h^*) - u), \nabla P_h \varphi)$$
$$(3.7) \qquad =: \mathrm{I}_1 + \mathrm{I}_2.$$

Next we bound the two terms. Direct computation with the triangle inequality gives

$$\|\nabla \cdot ((q^\dagger - q_h^*)\nabla u)\|_{L^2(\Omega)} \leq \|\nabla q^\dagger\|_{L^\infty(\Omega)}\|\nabla u\|_{L^2(\Omega)} + \|q^\dagger - q_h^*\|_{L^\infty(\Omega)}\|\Delta u\|_{L^2(\Omega)}$$
$$+ \|\nabla q_h^*\|_{L^2(\Omega)}\|\nabla u\|_{L^\infty(\Omega)}.$$

In view of the regularity estimate (3.6) and the box constraint of the admissible set $\mathcal{A}$, we derive

$$\|\nabla \cdot ((q^\dagger - q_h^*)\nabla u)\|_{L^2(\Omega)} \leq c + \|\nabla q_h^*\|_{L^2(\Omega)}\|\nabla u\|_{L^\infty(\Omega)} \leq c(1 + \|\nabla q_h^*\|_{L^2(\Omega)}).$$

This and the Cauchy–Schwarz inequality imply that the term I_1 is bounded by

$$|\mathrm{I}_1| \leq c(1 + \|\nabla q_h^*\|_{L^2(\Omega)})\|\varphi - P_h\varphi\|_{L^2(\Omega)}.$$

Now we choose the test function φ to be $\varphi \equiv \frac{q^\dagger - q_h^*}{q^\dagger} u$, and then direct computation gives

$$\nabla \varphi = \left(q^{\dagger -1}\nabla(q^\dagger - q_h^*) - q^{\dagger -2}(q^\dagger - q_h^*)\nabla q^\dagger\right)u + q^{\dagger -1}(q^\dagger - q_h^*)\nabla u,$$

which implies $\varphi \in H_0^1(\Omega)$. By the box constraint of the admissible set $\mathcal{A}$ and the regularity estimate (3.6), we have

$$\|\nabla \varphi\|_{L^2(\Omega)} \leq c\big[(1 + \|\nabla q_h^*\|_{L^2(\Omega)})\|u\|_{L^\infty(\Omega)} + \|\nabla u\|_{L^2(\Omega)}\big]$$
$$\leq c(1 + \|\nabla q_h^*\|_{L^2(\Omega)}).$$

Now the approximation property of the projection operator P_h in (2.1) implies

$$\|\varphi - P_h\varphi\|_{L^2(\Omega)} \leq ch\|\nabla \varphi\|_{L^2(\Omega)} \leq ch(1 + \|\nabla q_h^*\|_{L^2(\Omega)}).$$

Thus, in view of Lemma 3.1, the term I_1 in (3.7) can be bounded by

$$(3.8) \qquad |\mathrm{I}_1| \leq ch(1 + \|\nabla q_h^*\|_{L^2(\Omega)})^2 \leq ch(1 + \gamma^{-1}\eta^2) \leq ch\gamma^{-1}\eta^2.$$

For the term I_2, by the triangle inequality, inverse inequality on the space X_h, the $L^2(\Omega)$ stability of P_h, and Lemma 3.1, we have

$$\|\nabla(u - u_h(q_h^*))\|_{L^2(\Omega)} \leq \|\nabla(u - P_h u)\|_{L^2(\Omega)} + h^{-1}\|P_h u - u_h(q_h^*)\|_{L^2(\Omega)}$$
$$\leq c(h + h^{-1}\|u - u_h(q_h^*)\|_{L^2(\Omega)}) \leq c(h + h^{-1}\eta).$$

Meanwhile, clearly, there holds $\|\nabla(u - u_h(q_h^*))\|_{L^2(\Omega)} \leq c$, and hence the Cauchy–Schwarz inequality and Lemma 3.1 imply

$$\mathrm{I}_2 \leq \|\nabla(u - u_h(q_h^*))\|_{L^2(\Omega)}\|\nabla \varphi\|_{L^2(\Omega)}$$
$$(3.9) \qquad \leq c\min(h + h^{-1}\eta, 1)(1 + \|\nabla q_h^*\|_{L^2(\Omega)})$$
$$\leq c\min(h^{-1}\eta, 1)\gamma^{-\frac{1}{2}}\eta.$$

The estimates (3.8) and (3.9) together imply

$$((q^\dagger - q_h^*)\nabla u, \nabla\varphi) \le c(h\gamma^{-\frac{1}{2}}\eta + \min(h + h^{-1}\eta, 1))\gamma^{-\frac{1}{2}}\eta.$$

Now we claim the identity

$$((q^\dagger - q_h^*)\nabla u, \nabla\varphi) = \frac{1}{2}\int_\Omega \left(\frac{q^\dagger - q_h^*}{q^\dagger}\right)^2 \left(q^\dagger|\nabla u|^2 + fu\right) \mathrm{d}x,$$

which together with the preceding estimate leads directly to the desired assertion. It remains to show the claim. Actually, integration by parts yields

$$((q^\dagger - q_h^*)\nabla u, \nabla\varphi) = -\left(\nabla\left(\frac{q^\dagger - q_h^*}{q^\dagger}\right), q^\dagger\varphi\nabla u\right) - \left(\frac{q^\dagger - q_h^*}{q^\dagger}\varphi, \nabla\cdot(q^\dagger\nabla u)\right).$$

By the governing equation for u, $f = -\nabla\cdot(q^\dagger\nabla u)$, we deduce

$$((q^\dagger - q_h^*)\nabla u, \nabla\varphi) = \frac{1}{2}((q^\dagger - q_h^*)\nabla u, \nabla\varphi) - \frac{1}{2}\left(\nabla\left(\frac{q^\dagger - q_h^*}{q^\dagger}\right), q^\dagger\varphi\nabla u\right) + \frac{1}{2}\left(\frac{q^\dagger - q_h^*}{q^\dagger}\varphi, f\right).$$

Then plugging in $\varphi = \frac{q^\dagger - q_h^*}{q^\dagger}u$ and collecting the terms gives the claim and completes the proof. □

The next result gives an $L^2(\Omega)$ error estimate. The notation $\mathrm{dist}(x, \partial\Omega)$ denotes the distance of $x \in \Omega$ to the boundary $\partial\Omega$.

COROLLARY 3.3. *Let Assumption* 3.1 *be fulfilled, and assume that there exists some $\beta \ge 0$ such that*

(3.10) $$(q^\dagger|\nabla u(q^\dagger)|^2 + fu(q^\dagger))(x) \ge c\,\mathrm{dist}(x, \partial\Omega)^\beta \quad a.e. \ in \ \Omega.$$

Then the approximation q_h^* satisfies

$$\|q^\dagger - q_h^*\|_{L^2(\Omega)} \le c((h\gamma^{-\frac{1}{2}}\eta + \min(h^{-1}\eta, 1))\gamma^{-\frac{1}{2}}\eta)^{\frac{1}{2(1+\beta)}}.$$

In particular, for any $\delta > 0$, the choices $\gamma \sim \delta^2$ and $h \sim \sqrt{\delta}$ imply

$$\|q^\dagger - q_h^*\|_{L^2(\Omega)} \le c\delta^{\frac{1}{4(1+\beta)}}.$$

Proof. Let $u = u(q^\dagger)$. Then it follows from Theorem 3.2 that

$$\int_\Omega (q^\dagger - q_h^*)^2\left(q^\dagger|\nabla u|^2 + fu\right)\mathrm{d}x \le c(h\gamma^{-\frac{1}{2}}\eta + \min(h + h^{-1}\eta, 1))\gamma^{-\frac{1}{2}}\eta.$$

Then we decompose the domain Ω into two disjoint sets $\Omega = \Omega_\rho \cup \Omega_\rho^c$,

$$\Omega_\rho = \{x \in \Omega: \ \mathrm{dist}(x, \partial\Omega) \ge \rho\} \quad \text{and} \quad \Omega_\rho^c = \Omega\backslash\Omega_\rho,$$

where the constant $\rho > 0$ is to be chosen. On the subdomain Ω_ρ, we have

$$\int_{\Omega_\rho}(q^\dagger - q_h^*)^2\,\mathrm{d}x \le \rho^{-\beta}\int_{\Omega_\rho}(q^\dagger - q_h^*)^2\rho^\beta\,\mathrm{d}x$$

$$\le \rho^{-\beta}\int_{\Omega_\rho}(q^\dagger - q_h^*)^2\mathrm{dist}(x, \partial\Omega)^\beta\,\mathrm{d}x$$

$$\le c\rho^{-\beta}\int_{\Omega_\rho}(q^\dagger - q_h^*)^2\left(q^\dagger|\nabla u|^2 + fu\right)\mathrm{d}x$$

$$\le c\rho^{-\beta}(h\gamma^{-\frac{1}{2}}\eta + \min(h + h^{-1}\eta, 1))\gamma^{-\frac{1}{2}}\eta.$$

Meanwhile, by the box constraint of $\mathcal{A}$ and $\mathcal{A}_h$, we have

$$\int_{\Omega_\rho^c} (q^\dagger - q_h^*)^2 \, \mathrm{d}x \leq c|\Omega_\rho^c| \leq c\rho.$$

Then the desired result follows by balancing $\rho^{-\beta}(h\gamma^{-\frac{1}{2}}\eta + \min(h + h^{-1}\eta, 1))\gamma^{-\frac{1}{2}}\eta$ with ρ. $\qquad\square$

Remark 3.1. The positivity condition (3.10) has been established in [7, Lemma 3.7] for $\beta = 2$ provided that Ω is a Lipschitz domain, $q^\dagger \in \mathcal{A}$, and $f \in L^2(\Omega)$ with $f \geq c_f > 0$. Moreover, the condition with $\beta = 0$ holds provided that the domain Ω is $C^{2,\alpha}$, $q^\dagger \in C^{1,\alpha}(\overline{\Omega})$, and $f \in C^{0,\alpha}(\overline{\Omega})$ with $\alpha > 0$ and $f \geq c_f > 0$ [7, Lemma 3.3]. In the latter case, by choosing $\gamma \sim \delta^2$ and $h \sim \sqrt{\delta}$, we obtain

$$\|q^\dagger - q_h^*\|_{L^2(\Omega)} \leq c\delta^{\frac{1}{4}}.$$

Qualitatively, this estimate agrees with the conditional stability estimates in [7]. Indeed, for $u(q_1), u(q_2) \in H^2(\Omega) \cap H_0^1(\Omega)$ with $\|u(q_1) - u(q_2)\|_{L^2(\Omega)} \leq \delta$, [7, Theorem 3.2] implies

$$\|q_1 - q_2\|_{L^2(\Omega)} \leq c\|u(q_1) - u(q_2)\|_{H^1(\Omega)}^{\frac{1}{2}}.$$

This, the Gagliardo–Nirenberg interpolation inequality [8]

$$\|u\|_{H^1(\Omega)} \leq \|u\|_{L^2(\Omega)}^{\frac{1}{2}} \|u\|_{H^2(\Omega)}^{\frac{1}{2}}$$

and the regularity estimates $u(q_1), u(q_2) \in H^2(\Omega)$ directly give

$$\|q_1 - q_2\|_{L^2(\Omega)} \leq c(\|u(q_1)\|_{H^2(\Omega)} + \|u(q_2)\|_{H^2(\Omega)})^{\frac{1}{4}} \|u(q_1) - u(q_2)\|_{L^2(\Omega)}^{\frac{1}{4}}$$
$$\leq c\|u(q_1) - u(q_2)\|_{L^2(\Omega)}^{\frac{1}{4}} \leq c\delta^{\frac{1}{4}}.$$

There is a growing interest in using conditional stability estimates to derive convergence rates for continuous regularized formulations for inverse problems; see, e.g., [12] and references therein. However, analogous results for discretization errors based on conditional stability still seem to be missing.

Remark 3.2. Several alternative structural conditions have been proposed for deriving convergence rates, and it is instructive to compare these conditions for the elliptic inverse problem. One condition (with a Neumann boundary condition) is given by [16]

(3.11) $$\nabla u(q^\dagger)(x) \cdot \nu \geq c > 0 \quad \text{a.e. in } \Omega$$

for some constant c and constant vector ν or the less restrictive condition [28] $\max(|\nabla u|, \Delta u) \geq c > 0$ a.e. in Ω. Either condition implies the positivity condition (3.10) with $\beta = 0$ provided that u and f have the same sign (e.g., by the weak maximum principle for elliptic PDEs). Wang and Zou [35] derived an error estimate under a weaker assumption $\alpha_0|\nabla u|^2 \geq f$ a.e. in Ω. However, this condition is not positively homogeneous (with respect to problem data f).

Remark 3.3. Falk [16] proposed a numerical scheme for the elliptic inverse problem with a Neumann boundary condition, based on the output least-squares formulation, by looking for $u_h(q_h^*) \in V_h^r$ (continuous piecewise polynomials of degree r) and $q_h^* \in V_h^{r+1}$. If assumption (3.11) holds, $u \in W^{r+3,\infty}(\Omega)$, and $q^\dagger \in H^{r+1}(\Omega) \cap \mathcal{A}$, then

$$\|q^\dagger - q_h^*\|_{L^2(\Omega)} \leq c(h^r + \delta h^{-2}).$$

If $r = 1$ and $h \sim \delta^{\frac{1}{3}}$, it implies an error $O(\delta^{\frac{1}{3}})$. This better rate is obtained by the stronger regularity assumption $u(q^\dagger) \in W^{4,\infty}(\Omega)$ than that in Theorem 3.2.

Remark 3.4. Theorem 3.2 provides guidance for choosing the algorithmic parameters: Given the noise level δ, we may choose $\gamma \sim \delta^2$ and $h \sim \delta^{\frac{1}{2}}$. The choice $\gamma \sim \delta^2$ differs from the usual condition for Tikhonov regularization, i.e., $\lim_{\delta \to 0^+} \frac{\delta^2}{\gamma(\delta)} = 0$, but it agrees with that with conditional stability (see, e.g., [12, Theorems 1.1 and 1.2]).

4. Parabolic case. Now we analyze the parabolic inverse problem. For a function $v(x,t) : \Omega \times (0,T) \to \mathbb{R}$, we shall write $v(t) = v(\cdot, t)$ as a vector valued function on $(0,T)$ below.

4.1. Finite element approximation. To recover the diffusion coefficient $q^\dagger$ in (1.4), we employ the standard output least-squares formulation with an $H^1(\Omega)$ seminorm penalty,

$$(4.1) \qquad \min_{q \in \mathcal{A}} J_\gamma(q) = \frac{1}{2}\|u(q) - z^\delta\|^2_{L^2(T-\sigma,T;L^2(\Omega))} + \frac{\gamma}{2}\|\nabla q\|^2_{L^2(\Omega)},$$

where the admissible set $\mathcal{A}$ is given by (1.3) and $u(q)$ satisfies the variational problem
(4.2)
$$(\partial_t u(q), v) + (q\nabla u(q), \nabla v) = (f, v) \quad \forall v \in H_0^1(\Omega), t \in (0,T], \quad \text{with} \quad u(0) = u_0.$$

Now we describe a discretization of problem (4.1)–(4.2) based on the Galerkin FEM in space and the backward Euler method in time. Specifically, we partition the time interval $[0, T]$ uniformly, with grid points $t_n = n\tau$, $n = 0, \dots, N$ and a time step size $\tau = T/N$. The fully discrete scheme for problem (1.4) reads as follows: Given $U_h^0 = P_h u_0 \in X_h$, find $U_h^n \in X_h$ such that

$$(4.3) \qquad (\bar{\partial}_\tau U_h^n, \chi) + (q\nabla U_h^n, \nabla\chi) = (f(t_n)\chi) \quad \forall \chi \in X_h, \quad n = 1, 2, \dots, N,$$

where $\bar{\partial}_\tau \varphi^n = \frac{\varphi^n - \varphi^{n-1}}{\tau}$ denotes the backward Euler approximation to $\partial_t \varphi(t_n)$ (with the shorthand $\varphi^n = \varphi(t_n)$). Using operator $A_h(q)$ in (2.3), we rewrite (4.3) as

$$\bar{\partial}_\tau U_h^n + A_h(q)U_h^n = P_h f(t_n), \quad n = 1, 2, \dots, N.$$

Then the finite element discretization of problem (4.1)–(4.2) reads

$$(4.4) \qquad \min_{q_h \in \mathcal{A}_h} J_{\gamma,h,\tau}(q_h) = \tau \sum_{n=N_\sigma}^{N} \|U_h^n(q_h) - z_n^\delta\|^2_{L^2(\Omega)} + \frac{\gamma}{2}\|\nabla q_h\|^2_{L^2(\Omega)},$$

with

$$(4.5) \qquad\qquad z_n^\delta = \tau^{-1}\int_{t_{n-1}}^{t_n} z^\delta(t)\mathrm{d}t,$$

where the discrete admissible set $\mathcal{A}_h$ is given by (3.5) and $U_h^n(q_h) \in X_h$ satisfies $U_h^0 = P_h u_0$ and

$$(4.6) \qquad \bar{\partial}_\tau U_h^n(q_h) + A_h(q_h)U_h^n(q_h) = P_h f(t_n), \quad n = 1, 2, \dots, N+1.$$

Throughout, we assume that $N_\sigma = (T - \sigma)/\tau + 1$ is an integer. Analogous to the elliptic case, the following existence and convergence results hold. If $u_0 \in H_0^1(\Omega)$ and $f \in C([0,T]; L^2(\Omega))$, for every $h, \tau > 0$, there exists at least one minimizer $q_h^* \in \mathcal{A}_h$ (which depends also on τ) to problem (4.4)–(4.6), and, furthermore, the sequence of minimizers $\{q_h^*\}_{h,\tau>0}$ contains a subsequence that converges in $H^1(\Omega)$ to a minimizer of problem (4.1)–(4.2) as $h, \tau \to 0^+$; see [20, 27] for a proof.

4.2. Error estimates. Now we derive error estimates of approximations q_h^* under the following regularity condition on the problem data.

Assumption 4.1. The diffusion coefficient $q^\dagger$, initial data u_0, and source term f satisfy $q^\dagger \in H^2(\Omega) \cap W^{1,\infty}(\Omega) \cap \mathcal{A}$, $u_0 \in H^2(\Omega) \cap H_0^1(\Omega) \cap W^{1,\infty}(\Omega)$, $f \in L^\infty((0,T) \times \Omega) \cap C^1([0,T]; L^2(\Omega)) \cap W^{2,1}(0,T; L^2(\Omega))$.

Under Assumption 4.1, the parabolic problem (1.4) has a unique solution

$$(4.7) \qquad u \in W^{1,p}(0,T; L^q(\Omega)) \cap L^p(0,T; W^{2,q}(\Omega)) \quad \forall p, q \in (1, \infty).$$

The result follows directly from maximal L^p regularity of the parabolic equation; see, e.g., [30, Lemma 2.1]. Then by real interpolation and the Sobolev embedding theorem, we deduce

$$(4.8) \qquad u \in L^\infty(0,T; W^{1,\infty}(\Omega)).$$

Further, there holds [34, Lemma 3.2]

$$(4.9) \qquad \|\partial_t u(t)\|_{L^2(\Omega)} + t\|\partial_{tt} u(t)\|_{L^2(\Omega)} \leq c \quad \forall t \in [0, T].$$

The latter estimate immediately implies a useful uniform bound $\|u(t)\|_{H^2(\Omega)} \leq c$ since

$$\|A(q^\dagger)u(t)\|_{L^2(\Omega)} \leq \|\partial_t u(t)\|_{L^2(\Omega)} + \|f(t)\|_{L^2(\Omega)} \leq c.$$

With the choice of z_n^δ in (4.4), we have the following estimate.

LEMMA 4.1. *Let Assumption 4.1 be fulfilled. Then for z_n^δ defined in (4.5), there holds*

$$\tau \sum_{n=N_\sigma}^N \int_\Omega |u(t_n) - z_n^\delta|^2 \, dx \leq c(\tau^2 + \delta^2).$$

Proof. Let $u_n = \tau^{-1} \int_{t_{n-1}}^{t_n} u(t) \, dt$. Then we have

$$u(t_n) - u_n = \tau^{-1} \int_{t_{n-1}}^{t_n} u(t_n) - u(t) dt = \tau^{-1} \int_{t_{n-1}}^{t_n} \int_t^{t_n} \partial_s u(s) ds dt,$$

and thus by the regularity estimate (4.9),

$$\|u(t_n) - u_n\|_{L^2(\Omega)} \leq \tau^{-1} \int_{t_{n-1}}^{t_n} \int_t^{t_n} \|\partial_s u(s)\|_{L^2(\Omega)} ds dt$$
$$\leq c\tau \|\partial_t u(t)\|_{C([t_{n-1}, t_n]; L^2(\Omega))} \leq c\tau.$$

Meanwhile, by the Cauchy–Schwarz inequality, $\tau |u_n|^2 \leq \int_{t_{n-1}}^{t_n} u(t)^2 dt$. The last two estimates, the definition of the noise level δ in (1.5), and the stability estimate

$$\tau \sum_{n=N_\sigma}^N \int_\Omega |u_n - z_n^\delta|^2 \, dx \leq \int_{T-\sigma}^T \int_\Omega |u(t) - z^\delta(t)|^2 \, dx dt \leq \delta^2$$

imply the desired result immediately.   □

The next lemma gives error estimates of a fully discrete scheme for the direct problem (1.4): Find $U_h^n(q^\dagger)$ satisfying $U_h^0 = P_h u_0$ and

$$(4.10) \qquad \bar{\partial}_\tau(U_h^n(q^\dagger) - U_h^0) + A_h(q^\dagger)U_h^n(q^\dagger) = P_h f(t_n), \qquad n = 1, 2, \ldots, N.$$

It plays an important role in the error analysis below. The proof is standard but lengthy and hence is deferred to Appendix B.

LEMMA 4.2. *Let $q^\dagger$ be the exact diffusion coefficient and $u \equiv u(q^\dagger)$ the solution to problem* (4.2) *and* $\{U_h^n(q^\dagger)\}$ *the solution to problem* (4.10). *Then under Assumption* 4.1,

$$\|u(t_n) - U_h^n(q^\dagger)\|_{L^2(\Omega)} \leq c(\tau + h^2).$$

The next lemma provides an error estimate of the scheme (4.10) corresponding to the coefficient $\mathcal{I}_h q^\dagger$. It slightly relaxes the regularity assumption in [35, Lemma 6.1] from $q^\dagger \in W^{2,\infty}(\Omega)$ to $q^\dagger \in W^{1,\infty}(\Omega) \cap H^2(\Omega)$. The latter is identical with Assumption 3.1.

LEMMA 4.3. *Let $q^\dagger$ be the exact diffusion coefficient, $u \equiv u(q^\dagger)$ the solution to problem* (4.2), *and* $\{U_h^n(\mathcal{I}_h q^\dagger)\}$ *the solutions to the scheme* (4.10) *with* $\mathcal{I}_h q^\dagger$. *Then under Assumption* 4.1,

$$\tau \sum_{n=1}^{N} \|u(t_n) - U_h^n(\mathcal{I}_h q^\dagger)\|_{L^2(\Omega)}^2 \leq c(\tau^2 + h^4).$$

Proof. Note that $U_h^n(q^\dagger)$ and $U_h^n(\mathcal{I}_h q^\dagger)$, respectively, satisfy

$$A_h(q^\dagger)^{-1}\bar{\partial}_\tau U_h^n(q^\dagger) + U_h^n(q^\dagger) = A_h(q^\dagger)^{-1} P_h f(t_n), \quad n = 1, 2, \ldots, N,$$
$$A_h(\mathcal{I}_h q^\dagger)^{-1}\bar{\partial}_\tau U_h^n(\mathcal{I}_h q^\dagger) + U_h^n(\mathcal{I}_h q^\dagger) = A_h(\mathcal{I}_h q^\dagger)^{-1} P_h f(t_n), \quad n = 1, 2, \ldots, N,$$

with $U_h^0(q^\dagger) = U_h^0(\mathcal{I}_h q^\dagger) = P_h u_0$. Hence, $\rho_h^n := U_h^n(q^\dagger) - U_h^n(\mathcal{I}_h q^\dagger)$ satisfies
(4.11)
$$A_h(q^\dagger)^{-1}\bar{\partial}_\tau \rho_h^n + \rho_h^n = (A_h(q^\dagger)^{-1} - A_h(\mathcal{I}_h q^\dagger)^{-1})(P_h f(t_n) - \bar{\partial}_\tau U_h^n(\mathcal{I}_h q^\dagger)), \quad n = 1, \ldots, N,$$

with $\rho_h^0 = 0$. It follows from direct computation that

$$\begin{aligned}
(\bar{\partial}_\tau A_h(q^\dagger)^{-1}\rho_h^n, \rho_h^n) &= (\bar{\partial}_\tau A_h(q^\dagger)^{-\frac{1}{2}}\rho_h^n, A_h(q^\dagger)^{-\frac{1}{2}}\rho_h^n) \\
&= \tfrac{1}{2}\bar{\partial}_\tau\|A_h(q^\dagger)^{-\frac{1}{2}}\rho_h^n\|_{L^2(\Omega)}^2 + \tfrac{1}{2\tau}\|A_h(q^\dagger)^{-\frac{1}{2}}(\rho_h^n - \rho_h^{n-1})\|_{L^2(\Omega)}^2 \\
&\geq \tfrac{1}{2}\bar{\partial}_\tau\|A_h(q^\dagger)^{-\frac{1}{2}}\rho_h^n\|_{L^2(\Omega)}^2,
\end{aligned}$$

Then taking inner product (4.11) with ρ_h^n and by the Cauchy–Schwarz inequality, we obtain

$$\begin{aligned}
&\tfrac{1}{2}\bar{\partial}_\tau\|A_h(q^\dagger)^{-\frac{1}{2}}\rho_h^n\|_{L^2(\Omega)}^2 + \|\rho_h^n\|_{L^2(\Omega)}^2 \\
&\quad \leq \|(A_h(q^\dagger)^{-1} - A_h(\mathcal{I}_h q^\dagger)^{-1})(P_h f(t_n) - \bar{\partial}_\tau U_h^n(\mathcal{I}_h q^\dagger))\|_{L^2(\Omega)}\|\rho_h^n\|_{L^2(\Omega)} \\
&\quad \leq \tfrac{1}{2}\|(A_h(q^\dagger)^{-1} - A_h(\mathcal{I}_h q^\dagger)^{-1})(P_h f(t_n) - \bar{\partial}_\tau U_h^n(\mathcal{I}_h q^\dagger))\|_{L^2(\Omega)}^2 + \tfrac{1}{2}\|\rho_h^n\|_{L^2(\Omega)}^2.
\end{aligned}$$

Further, by Lemma A.1, we have, for any $\epsilon > 0$ and $p \geq \max(d + \epsilon, 2)$,

$$\|A_h(\mathcal{I}_h q^\dagger)^{-1} - A_h(q^\dagger)^{-1}\|_{L^p(\Omega) \to L^2(\Omega)} \leq ch^2.$$

Hence, for $p \geq \max(d + \epsilon, 2)$,

$$\bar{\partial}_\tau \|A_h(q^\dagger)^{-\frac{1}{2}} \rho_h^n\|_{L^2(\Omega)}^2 + \|\rho_h^n\|_{L^2(\Omega)}^2 \leq ch^4 \|P_h f(t_n) - \bar{\partial}_\tau U_h^n(\mathcal{I}_h q^\dagger)\|_{L^p(\Omega)}^2$$
$$\leq ch^4 \|f(t_n) - \bar{\partial}_\tau U_h^n(\mathcal{I}_h q^\dagger)\|_{L^p(\Omega)}^2,$$

where in the second line we have used the $L^p(\Omega)$ stability of P_h [11]. Then summing over n gives

$$\|A_h(q^\dagger)^{-\frac{1}{2}} \rho_h^N\|_{L^2(\Omega)}^2 + \tau \sum_{n=1}^N \|\rho_h^n\|_{L^2(\Omega)}^2$$
$$\leq ch^4 \Big(\tau \sum_{n=1}^N \|f(t_n)\|_{L^p(\Omega)}^2 + \tau \sum_{n=1}^N \|\bar{\partial}_\tau U_h^n(\mathcal{I}_h q^\dagger)\|_{L^p(\Omega)}^2 \Big).$$

Then the maximal ℓ^p regularity for the backward Euler scheme [4] implies

$$\tau \sum_{n=1}^N \|\rho_h^n\|_{L^p(\Omega)}^2 \leq ch^4 \Big(\tau \sum_{n=1}^N \|f(t_n)\|_{L^p(\Omega)}^2 + \|\nabla u_0\|_{L^p(\Omega)}^2 \Big).$$

Finally, the desired estimate follows from Lemma 4.2 and the triangle inequality. $\quad\square$

The next result gives a priori bounds on q_h^* and error estimates on the corresponding approximations $U_h^n(q_h^*)$. This result will play a crucial role in the proof of Theorem 4.5 below.

LEMMA 4.4. *Let $q^\dagger$ be the exact coefficient and $u \equiv u(q^\dagger)$ the solution to problem* (4.2). *Let $q_h^* \in \mathcal{A}_h$ be the solution to problem* (4.4)–(4.6) *and $\{U_h^n(q_h^*)\}_{n=1}^N$ the fully discrete solution to problem* (4.6). *Then under Assumption* 4.1, *there holds*

$$\tau \sum_{n=N_\sigma}^N \|U_h^n(q_h^*) - u(t_n)\|_{L^2(\Omega)}^2 + \gamma \|\nabla q_h^*\|_{L^2(\Omega)}^2 \leq c(\tau^2 + h^4 + \delta^2 + \gamma).$$

Proof. By the minimizing property of $q_h^* \in \mathcal{A}_h$, since $\mathcal{I}_h q^\dagger \in \mathcal{A}_h$, we deduce

$$J_{\gamma,h,\tau}(q_h^*) \leq J_{\gamma,h,\tau}(\mathcal{I}_h q^\dagger).$$

By the triangle inequality, we derive

$$\tau \sum_{n=N_\sigma}^N \|U_h^n(q_h^*) - u(t_n)\|_{L^2(\Omega)}^2$$
$$\leq c\tau \sum_{n=N_\sigma}^N \|U_h^n(q_h^*) - z_n^\delta\|_{L^2(\Omega)}^2 + c\tau \sum_{n=N_\sigma}^N \|z_n^\delta - u(t_n)\|_{L^2(\Omega)}^2.$$

These two inequalities together imply

$$\tau \sum_{n=N_\sigma}^{N} \|U_h^n(q_h^*) - u(t_n)\|_{L^2(\Omega)}^2 + \gamma\|\nabla q_h^*\|_{L^2(\Omega)}^2$$

$$\leq c\tau \sum_{n=N_\sigma}^{N} \|U_h^n(\mathcal{I}_h q^\dagger) - z_n^\delta\|_{L^2(\Omega)}^2 + c\gamma\|\nabla\mathcal{I}_h q^\dagger\|_{L^2(\Omega)}^2 + c\tau \sum_{n=N_\sigma}^{N} \|z_n^\delta - u(t_n)\|_{L^2(\Omega)}^2$$

$$\leq c\tau \sum_{n=N_\sigma}^{N} \|U_h^n(\mathcal{I}_h q^\dagger) - u(t_n)\|_{L^2(\Omega)}^2 + c\gamma\|\nabla\mathcal{I}_h q^\dagger\|_{L^2(\Omega)}^2 + c\tau \sum_{n=N_\sigma}^{N} \|z_n^\delta - u(t_n)\|_{L^2(\Omega)}^2$$

$$\leq c\tau \sum_{n=N_\sigma}^{N} \|U_h^n(\mathcal{I}_h q^\dagger) - u(t_n)\|_{L^2(\Omega)}^2 + c\gamma\|\nabla\mathcal{I}_h q^\dagger\|_{L^2(\Omega)}^2 + c(\delta^2 + \tau^2),$$

where the last line follows from Lemma 4.1. Since $q^\dagger \in W^{1,\infty}(\Omega)$ by Assumption 4.1, $\|\nabla\mathcal{I}_h q^\dagger\|_{L^2(\Omega)} \leq c$; cf. (2.2). Combining the preceding estimates with Lemma 4.3 completes the proof. □

Now we give the main result of this section, i.e., error estimate of the numerical approximation $q_h^* \in \mathcal{A}_h$, with the weight involving $q^\dagger|\nabla u(t_n)|^2 + (f(t_n) - \partial_t u(t_n))u(t_n)$, whose positivity will be analyzed below in section 4.3.

THEOREM 4.5. *Let $q^\dagger$ be the exact diffusion coefficient, $u \equiv u(q^\dagger)$ the solution to problem* (4.2), *and $q_h^* \in \mathcal{A}_h$ a solution to problem* (4.4)–(4.6). *Then under Assumption 4.1, with $\eta = \tau + h^2 + \delta + \gamma^{\frac{1}{2}}$, there holds*

$$\tau^3 \sum_{j=N_\sigma+1}^{N} \sum_{i=N_\sigma+1}^{j} \sum_{n=i}^{j} \int_\Omega \left(\frac{q^\dagger - q_h^*}{q^\dagger}\right)^2 \left(q^\dagger|\nabla u(t_n)|^2 + (f(t_n) - \partial_t u(t_n))u(t_n)\right) \mathrm{d}x$$
$$\leq c(h\gamma^{-\frac{1}{2}}\eta + \min(1, h^{-1}\eta))\gamma^{-\frac{1}{2}}\eta.$$

Proof. For any test function $\varphi \in H_0^1(\Omega)$, we have

$$((q^\dagger - q_h^*)\nabla u(t_n), \nabla\varphi) = -(\nabla \cdot ((q^\dagger - q_h^*)\nabla u(t_n)), \varphi - P_h\varphi)$$
$$+ (q_h^*\nabla(U_h^n(q_h^*) - u(t_n)), \nabla P_h\varphi)$$
$$+ (q^\dagger\nabla u(t_n) - q_h^*\nabla U_h^n(q_h^*), \nabla P_h\varphi) = \sum_{i=1}^{3} \mathrm{I}_i^n.$$

Throughout, the test function φ is taken to be $\varphi \equiv \varphi^n = \frac{q^\dagger - q_h^*}{q^\dagger}u(t_n)$. Then repeating the argument in Theorem 3.2 with the regularity estimates (4.7) and (4.8) and the approximation property of P_h in (2.1) yields
(4.12)
$$\|\nabla\varphi^n\|_{L^2(\Omega)} \leq c(1 + \|\nabla q_h^*\|_{L^2(\Omega)}) \quad \text{and} \quad \|P_h\varphi^n - \varphi^n\|_{L^2(\Omega)} \leq ch(1 + \|\nabla q_h^*\|_{L^2(\Omega)}).$$

Next we bound the three terms separately. By Assumption 4.1 (and hence the regularity estimates (4.7) and (4.8)) and the box constraint of $q^\dagger$ and q_h^*, the term I_1^n is bounded by

$$|\mathrm{I}_1^n| \leq ch(1 + \|\nabla q_h^*\|_{L^2(\Omega)})^2 \leq ch(1 + \gamma^{-1}\eta^2) \leq ch\gamma^{-1}\eta^2.$$

For the term I_2^n, by the triangle inequality, inverse inequality, and $L^2(\Omega)$ stability of the operator P_h in (2.1), we deduce

$$
\begin{aligned}
\|\nabla(u(t_n) &- U_h^n(q_h^*))\|_{L^2(\Omega)} \\
&\leq \|\nabla(u(t_n) - P_h u(t_n))\|_{L^2(\Omega)} + h^{-1}\|P_h u(t_n) - U_h^n(q_h^*)\|_{L^2(\Omega)} \\
&\leq c(h + h^{-1}\|P_h u(t_n) - U_h^n(q_h^*)\|_{L^2(\Omega)}).
\end{aligned}
$$

Meanwhile, by the energy argument in Lemma 4.3, we deduce

$$
\tau \sum_{n=1}^N \|\nabla U_h^n(q_h^*)\|_{L^2(\Omega)}^2 \leq c\Big(\sum_{n=1}^N \|f(t_n)\|_{L^2(\Omega)}^2 + \|\nabla u_0\|_{L^2(\Omega)}^2 \Big) \leq c.
$$

By this and the regularity estimate (4.9), $\tau \sum_n \|\nabla(u(t_n) - U_h^n(q_h^*))\|_{L^2(\Omega)}^2 \leq c$. Consequently, the Cauchy–Schwarz inequality, Lemma 4.4, and (4.12) imply

$$
\begin{aligned}
\tau \sum_{n=N_\sigma}^N I_2^n &\leq \tau \sum_{n=1}^N \|\nabla(u(t_n) - U_h^n(q_h^*))\|_{L^2(\Omega)} \|\nabla \varphi^n\|_{L^2(\Omega)} \\
&\leq c \min\left(1, h + h^{-1}\Big(\tau \sum_{n=N_\sigma}^N \|u(t_n) - U_h^n(q_h^*)\|_{L^2(\Omega)}^2 \Big)^{\frac{1}{2}} \right)(1 + \|\nabla q_h^*\|_{L^2(\Omega)}) \\
&\leq c \min(1, h + h^{-1}\eta)\gamma^{-\frac{1}{2}}\eta \leq c \min(1, h^{-1}\eta)\gamma^{-\frac{1}{2}}\eta.
\end{aligned}
$$

Next we bound the term I_3^n. It follows from the variational formulations (4.2) and (4.6) that

$$
\begin{aligned}
I_3^n &= (q^\dagger \nabla u(t_n) - q_h^* \nabla U_h^n(q_h^*), \nabla P_h \varphi^n) \\
&= (\bar{\partial}_\tau U_h^n(q_h^*) - \partial_t u(t_n), P_h \varphi^n) \\
&= (\bar{\partial}_\tau [U_h^n(q_h^*) - u(t_n)], P_h \varphi^n) + (\bar{\partial}_\tau u(t_n) - \partial_t u(t_n), P_h \varphi^n) =: I_{3,1}^n + I_{3,2}^n.
\end{aligned}
$$

It remains to bound the two terms $I_{3,1}^n$ and $I_{3,2}^n$ separately. Note that

$$
\bar{\partial}_\tau u(t_n) - \partial_t u(t_n) = \tau^{-1} \int_{t_{n-1}}^{t_n} \partial_s u(s) - \partial_t u(t_n) \mathrm{d}s.
$$

Thus, by the regularity estimate (4.9), we have for $n \geq 2$

$$
\|\bar{\partial}_\tau u(t_n) - \partial_t u(t_n)\|_{L^2(\Omega)} \leq \frac{1}{\tau} \int_{t_{n-1}}^{t_n} \int_s^{t_n} \|u''(\xi)\|_{L^2(\Omega)} \, \mathrm{d}\xi \, \mathrm{d}s \leq c\tau t_{n-1}^{-1} \leq c\tau t_n^{-1}
$$

and for $n = 1$

$$
\|\bar{\partial}_\tau u(\tau) - \partial_t u(\tau)\|_{L^2(\Omega)} \leq c.
$$

Consequently, there holds

$$
|I_{3,2}^n| \leq \|\bar{\partial}_\tau u(t_n) - \partial_t u(t_n)\|_{L^2(\Omega)} \|P_h \varphi^n\|_{L^2(\Omega)} \leq c\tau t_n^{-1}, \quad n = 1, 2, \ldots, N,
$$

and

$$
\left| \tau^3 \sum_{j=N_\sigma+1}^N \sum_{i=N_\sigma+1}^j \sum_{n=i}^j I_{3,2}^n \right| \leq c\tau \int_{T-\sigma}^T \int_{T-\sigma}^t \int_s^t \xi^{-1} \, \mathrm{d}\xi \mathrm{d}s \mathrm{d}t \leq c\tau.
$$

Meanwhile, since $U_h^0(q_h^*) = U_h^0$ and $u(0) = u_0$, the summation by parts formula yields

$$(4.13) \qquad \tau \sum_{n=i}^{j} I_{3,1}^n = \tau \sum_{n=i}^{j} (\bar{\partial}_\tau [U_h^n(q_h^*) - u(t_n)], P_h \varphi^n)$$

$$= (U_h^j(q_h^*) - u(t_j), P_h\varphi^j) - (U_h^{i-1}(q_h^*) - u(t_{i-1}), P_h\varphi^i)$$

$$- \tau \sum_{n=i}^{j-1} (U_h^n(q_h^*) - u(t_n), \bar{\partial}_\tau P_h \varphi^{n+1}).$$

For the first two terms, by the Cauchy–Schwarz inequality and Hölder's inequality, we have

$$\left| \tau^2 \sum_{j=N_\sigma+1}^{N} \sum_{i=N_\sigma+1}^{j} (U_h^j(q_h^*) - u(t_j), P_h\varphi^j) - (U_h^{i-1}(q_h^*) - u(t_{i-1}), P_h\varphi^i) \right|$$

$$\leq c \left(\tau \sum_{n=N_\sigma+1}^{N} \|U_h^n(q_h^*) - u(t_n)\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}} \leq c\eta$$

since by (4.9), $\|P_h\varphi^i\|_{L^2(\Omega)} \leq c$. Meanwhile, by using the regularity estimate (4.9) and the box constraint, we have

$$\|\bar{\partial}_\tau P_h \varphi^n\|_{L^2(\Omega)} = \tau^{-1} \left\| \int_{t_{n-1}}^{t_n} P_h \frac{q^\dagger - q_h^*}{q_\dagger} \partial_t u(t) dt \right\|_{L^2(\Omega)} \leq c\tau^{-1} \int_{t_{n-1}}^{t_n} \|\partial_t u(t)\|_{L^2(\Omega)} dt \leq c,$$

and hence

$$\left| \tau^3 \sum_{j=N_\sigma+1}^{N} \sum_{i=N_\sigma+1}^{j} \sum_{n=i}^{j-1} (U_h^n(q_h^*) - u(t_n), \bar{\partial}_\tau P_h \varphi^{n+1}) \right|$$

$$\leq c\tau^3 \sum_{j=N_\sigma+1}^{N} \sum_{i=N_\sigma+1}^{j} \sum_{n=i}^{j-1} \|U_h^n(q_h^*) - u(t_n)\|_{L^2(\Omega)}$$

$$\leq c \left(\tau \sum_{n=N_\sigma+1}^{N} \|U_h^n(q_h^*) - u(t_n)\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}} \leq c\eta.$$

Finally, this and the identity

$$((q^\dagger - q_h^*)\nabla u(t_n), \nabla\varphi^n) = \frac{1}{2} \int_\Omega \left(\frac{q^\dagger - q_h^*}{q^\dagger} \right)^2 \left(q^\dagger |\nabla u(t_n)|^2 + (f(t_n) - \partial_t u(t_n))u(t_n) \right) dx$$

(cf. the proof of Theorem 3.2) lead to the desired assertion, completing the proof. □

The next result gives an $L^2(\Omega)$ estimate under a suitable positivity condition similar to (3.10). The proof is identical with that for Corollary 3.3 and thus is omitted.

COROLLARY 4.6. *Let Assumption 4.1 be fulfilled, and there exists some $\beta \geq 0$ such that*
(4.14)
$$q^\dagger(x)|\nabla u(q^\dagger)(x,t)|^2 + (f(x,t) - \partial_t u(q^\dagger)(x,t))u(q^\dagger)(x,t) \geq c \, \text{dist}(x, \partial\Omega)^\beta \quad \text{a.e. in } \Omega,$$

for any $t \in [T - \sigma, T]$. Then for any $\delta > 0$, with $\eta = \tau + h^2 + \delta + \gamma^{\frac{1}{2}}$, there holds

$$\|q^\dagger - q_h^*\|_{L^2(\Omega)} \leq c((h\gamma^{-1}\eta + \gamma^{-\frac{1}{2}} \min(1, h^{-1}\eta))\eta)^{\frac{1}{2(1+\beta)}}.$$

In particular, the choices $\gamma \sim \delta^2$, $h \sim \sqrt{\delta}$, and $\tau \sim \delta$ imply

$$\|q^\dagger - q_h^*\|_{L^2(\Omega)} \leq c\delta^{\frac{1}{4(1+\beta)}}.$$

Remark 4.1. Note that in the identity (4.13), the first two terms cannot be bounded directly since only ℓ^2 bounds are available on $(U_h^j(q_h^*) - u(t_j), P_h\varphi^j)$ and $(U_h^{i-1}(q_h^*) - u(t_{i-1}), P_h\varphi^i)$. The triple sum $\sum_{j=N_\sigma+1}^N \sum_{i=N_\sigma+1}^j \sum_{n=i}^j$ is precisely to exploit relevant ℓ^2 bounds.

Remark 4.2. The error estimate in Corollary 4.6 provides the usual $L^2(\Omega)$ error estimate. Alternatively, one obtains the $L^2(\Omega)$ estimate if the following structural condition holds: For the exact diffusion coefficient $q^\dagger$ and the corresponding state variable $u \equiv u(q^\dagger)$, there holds

$$(4.15) \quad \int_{T-\sigma}^T \int_{T-\sigma}^t \int_s^t \left(q^\dagger |\nabla u(\xi)|^2 + (f(\xi) - \partial_\xi u(\xi))u(\xi) \right) \mathrm{d}\xi \mathrm{d}s \mathrm{d}t > c_0 \qquad \text{a.e. } x \in \overline{\Omega}.$$

Similar structural conditions have been assumed in the literature, e.g., the characteristic condition [33]

$$t^{-1} \int_0^t \nabla u(q^\dagger)(x, s) \mathrm{d}s \cdot \nu \geq c > 0 \quad \text{a.e. in } \Omega \times (0, T),$$

with some constant c and vector ν, or [35, Theorem 6.4],

$$\alpha_0 |\int_0^t \nabla u(q^\dagger)(s) \mathrm{d}s|^2 + t \int_0^t (\partial_s u(q^\dagger)(s) - f(s)) \mathrm{d}s \geq 0 \quad \text{a.e. in } \Omega \times (0, T).$$

4.3. On the positivity condition (4.14). Condition (4.14) allows deriving an $L^2(\Omega)$ estimate; cf. Corollary 4.6. Now we give sufficient conditions on problem data to ensure (4.14).

PROPOSITION 4.7. *Let Ω be a bounded Lipschitz domain, $q^\dagger \in \mathcal{A} \cap W^{1,\infty}(\Omega)$, $u_0 \in H^2(\Omega) \cap H_0^1(\Omega)$, and $f \in H^1((0, T); L^2(\Omega))$. Meanwhile, assume that $f \geq c_f > 0$ and $\partial_t f \leq 0$ a.e. in $\Omega \times (0, T)$ and $u_0(x) \geq 0$, $f(x, 0) + \nabla \cdot (q^\dagger \nabla u_0(x)) \leq 0$ a.e. in Ω. Then the positivity condition (4.14) holds with $\beta = 2$, with the constant c only depending on c_0, c_1, c_f, and Ω.*

Proof. Since $u_0 \geq 0$ and $f > c_f$, the maximum principle of parabolic equations [17] implies

$$u(x, t) \geq 0 \quad \forall (x, t) \in \overline{\Omega} \times [0, T].$$

Let $w(x, t) = \partial_t u(x, t)$. Then it satisfies

$$\begin{cases} \partial_t w - \nabla \cdot (q^\dagger \nabla w) = \partial_t f & \text{in } \Omega \times (0, T], \\ w = 0 & \text{on } \partial\Omega \times (0, T], \\ w(0) = f(0) + \nabla \cdot (q^\dagger \nabla u_0) & \text{in } \Omega. \end{cases}$$

By assumption, $\partial_t f \leq 0$ in $\Omega \times (0, T]$ and $w(0) \leq 0$ in Ω. Then the parabolic maximum principle implies $\partial_t u(x, t) = w(x, t) \leq 0$ in $\overline{\Omega} \times [0, T]$. Therefore, there holds

$$(4.16)$$
$$q^\dagger(x)|\nabla u(x, t)|^2 + (f(x, t) - \partial_t u(x, t))u(x, t) \geq \min(c_0, c_f)(|\nabla u(x, t)|^2 + u(x, t)).$$

So it suffices to prove $u(x,t) \geq c\,\mathrm{dist}(x,\partial\Omega)^2$ for $(x,t) \in \Omega \times (0,T]$. For any fixed $t \in [T-\sigma, T]$, we have $f(x,t) - \partial_t u(x,t) \in L^2(\Omega)$. Now consider the elliptic problem

$$(4.17) \qquad \begin{cases} -\nabla \cdot (q^\dagger \nabla u(t)) = f(t) - \partial_t u(t) & \text{in } \Omega, \\ \qquad\qquad\qquad u(t) = 0 & \text{on } \partial\Omega. \end{cases}$$

Let $G(x,y)$ be Green's function corresponding to the elliptic operator $\nabla \cdot (q^\dagger(x)\nabla \cdot)$. Then $G(x,y)$ is nonnegative (by the maximum principle) and satisfies the following a priori estimate (see, e.g., [19, Theorem 1.1] and [7, Lemma 3.7]):

$$G(x,y) \geq c|x-y|^{2-d} \quad \text{for} \quad |x-y| \leq \rho(x) := \mathrm{dist}(x,\partial\Omega).$$

Consequently, for any $x \in \Omega$ and $t \in [T-\sigma, T]$, there holds

$$u(x,t) = \int_\Omega G(x,y)(f(y,t) - \partial_t u(y,t))\,dy \geq \int_\Omega G(x,y)f(y,t)\,dy \geq c_f \int_\Omega G(x,y)\,dy$$

$$\geq c_f \int_{|x-y|<\rho(x)/2} G(x,y)\,dy \geq c \int_{|x-y|<\rho(x)/2} |x-y|^{2-d}\,dy \geq c\rho(x)^2.$$

This completes the proof of the proposition. $\qquad\square$

The next result gives sufficient conditions for the positivity condition (4.14) with $\beta = 0$ under stronger regularity assumptions on the problem data.

PROPOSITION 4.8. *Let Ω be a bounded $C^{2,\alpha}$ domain, $f \in C^1([0,T]; C^{0,\alpha}(\overline{\Omega}))$ with $f \geq c_f > 0$, $\partial_t f \leq 0$ in $\overline{\Omega} \times [0,T]$, and $u_0 \in C^{2,\alpha}(\overline{\Omega}) \cap H^1_0(\Omega)$ with $u_0 \geq 0$ in Ω. Moreover, assume $q^\dagger \in \mathcal{A} \cap C^{1,\alpha}(\overline{\Omega})$ and $f(x,0) + \nabla \cdot (q^\dagger \nabla u_0(x)) \leq 0$ in Ω. Then the positivity condition (4.14) holds with $\beta = 0$, with the constant only depending on c_0, c_1, c_f, Ω, and $\|q^\dagger\|_{C^{1,\alpha}(\overline{\Omega})}$.*

Proof. By the argument in the proof of Proposition 4.7, we have $\partial_t u \in C([0,T]; C^{0,\alpha}(\overline{\Omega}))$ and $\partial_t u \leq 0$ for all $(x,t) \in \overline{\Omega} \times (0,T)$. Hence, the inequality (4.16) is still valid. Now it suffices to show that for any $(x,t) \in \overline{\Omega} \times [T-\sigma, T]$, there holds $|\nabla u(x,t)|^2 + u(x,t) \geq c > 0$. Note that $u(x,t)$ is the solution of the elliptic problem (4.17) with a $C^{0,\alpha}(\overline{\Omega})$ source term $f(t) - \partial_t u(t) \geq f(t) \geq c_f$. Then the desired result follows from Schauder estimates and a standard compactness argument. For the details, see the proof of [7, Lemma 3.3]. $\qquad\square$

5. Numerical results. In this section, we present several numerical experiments to complement the analysis. Throughout, the discrete optimization problem is solved by the conjugate gradient method, which converges within tens of iterations. The lower and upper bounds in the admissible set $\mathcal{A}$ are taken to be 0.5 and 5, respectively, and are enforced by a projection step. In the elliptic case, the noisy data z^δ are generated by

$$z^\delta(x) = u(q^\dagger)(x) + \varepsilon \sup_{x\in\Omega} |u(q^\dagger)|\xi(x),$$

where ξ follows the standard Gaussian distribution and $\varepsilon > 0$ denotes the (relative) noise level and similarly for the parabolic case. The noisy data z^δ is first generated on a fine mesh and then interpolated to a coarse spatial/temporal mesh for the inversion step. All the computations are carried out on a personal laptop with MATLAB 2019.

TABLE 5.1
Numerical results for Example 5.1: convergence with respect to ε, with γ and h initialized to 5.00e-8 and 2.50e-2.

ε	5.00e-2	3.00e-2	1.00e-2	5.00e-3	3.00e-3	1.00e-3	5.00e-4	
e_q	2.52e-1	2.56e-1	8.08e-2	4.84e-2	4.06e-2	1.63e-2	8.43e-3	0.76
e_u	2.10e-3	9.89e-4	2.54e-4	1.20e-4	7.45e-5	2.06e-5	8.46e-6	1.16

(a) ε=1e-3   (b) ε=1e-2   (c) ε=5e-2
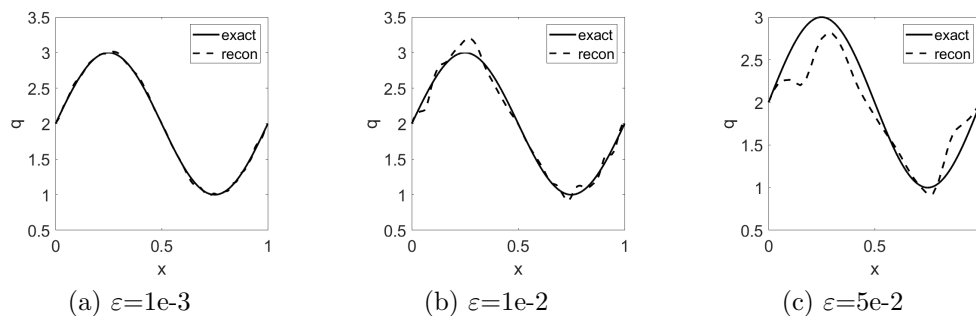
FIG. 5.1. *Numerical reconstructions for Example 5.1 at three noise levels.*

5.1. Numerical results for elliptic problems. First, we give one- and two-dimensional elliptic examples.

Example 5.1. $\Omega = (0,1)$, $q^\dagger(x) = 2 + \sin 2\pi x$ and $f \equiv 1$. The exact data are generated on a fine mesh with a mesh size $h = 1/3200$.

The numerical results for Example 5.1 are summarized in Table 5.1, where the numbers in the last column denote convergence rates with respect to the noise level δ, i.e., the exponent α in $O(\delta^\alpha)$. In the tables, e_q and e_u are defined by

$$e_q = \|q_h^* - q^\dagger\|_{L^2(\Omega)} \quad \text{and} \quad e_u = \|u_h(q_h^*) - u(q^\dagger)\|_{L^2(\Omega)},$$

respectively. For the convergence with respect to ε, the regularization parameter γ and mesh size h are taken to be $\gamma = c_\gamma \varepsilon^2$ and $h = c_h \varepsilon^{\frac{1}{2}}$, respectively, as suggested by Corollary 3.3, where the constant c_γ is determined by trial and error. Figure 5.1(a) indicates that the error e_q decays to zero as the noise level ε decreases to zero, with an empirical rate $O(\delta^{0.76})$. Meanwhile, the numerical experiment shows that the weight $|\nabla u(q^\dagger)|^2 + f u(q^\dagger)$ in the error estimate is indeed strictly positive over the domain $\overline{\Omega}$, even though both components have vanishing points. Thus, by Theorem 3.2 and Corollary 3.3, the predicted rate is $O(\delta^{\frac{1}{4}})$, which is much lower than the empirical rate $O(\delta^{0.76})$, indicating the suboptimality of the predicted rate. The error e_u converges slightly faster than first order. See also Figure 5.1 for an illustration of the reconstructions at three different noise levels.

Example 5.2. $\Omega = (0,1)^2$, $q^\dagger(x_1, x_2) = 1 + x_2(1 - x_2)\sin \pi x_1$, and $f \equiv 1$. The data are generated on a fine mesh with a mesh size $h = 1/200$.

The numerical results for Example 5.2 are presented in Table 5.2 and Figure 5.2. The empirical convergence rates for e_q and e_u with respect to δ are about $O(\delta^{0.72})$ and $O(\delta)$, respectively, which are comparable with that for Example 5.1. In either metric, the convergence is very steady. Note that for this example, the weight $q^\dagger|\nabla u(q^\dagger)|^2 + f u(q^\dagger)$ is not strictly positive over $\overline{\Omega}$ since it vanishes at two corners of the square domain Ω.

TABLE 5.2

Numerical results for Example 5.2: convergence with respect to ε, with γ and h initialized to 5e-6 and 8.33e-2.

ε	5.00e-2	3.00e-2	1.00e-2	5.00e-3	3.00e-3	1.00e-03	5.00e-4	
e_q	4.46e-2	3.17e-2	1.27e-2	6.98e-3	5.59e-3	2.64e-03	1.63e-3	0.72
e_u	7.88e-4	4.11e-4	1.20e-4	6.56e-5	3.89e-5	1.39e-05	7.72e-6	1.00

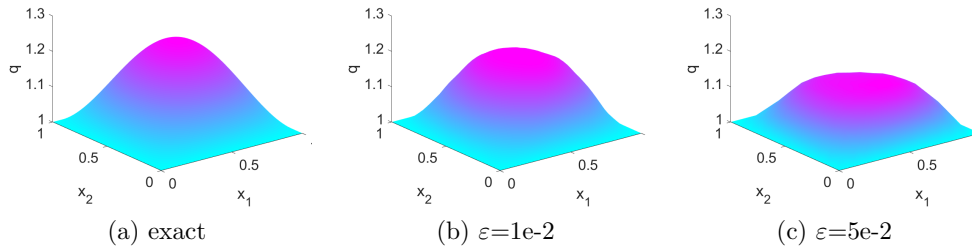
(a) exact (b) ε=1e-2 (c) ε=5e-2

FIG. 5.2. *Numerical reconstructions for Example 5.2 at two noise levels.*

TABLE 5.3

Numerical results for Example 5.3: convergence with respect ε, with γ, h, and τ are initialized with 1.00e-7, 2.50e-2, and 1/400.

ε	5.00e-2	3.00e-2	1.00e-2	5.00e-3	3.00e-3	1.00e-3	5.00e-4	
e_q	1.97e-2	1.34e-2	6.74e-3	2.58e-3	2.26e-3	8.86e-4	9.57e-4	0.71
e_u	2.31e-4	1.07e-4	8.78e-5	3.83e-5	3.68e-5	1.22e-5	1.19e-5	0.64

(a) ε=1e-3 (b) ε=1e-2 (c) ε=5e-2
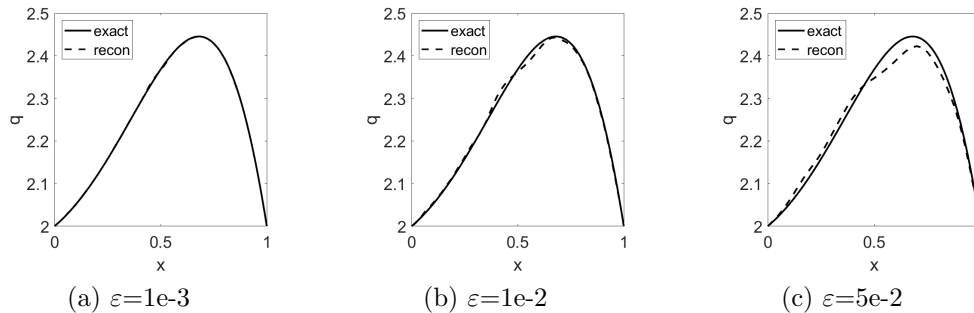
FIG. 5.3. *Numerical reconstructions for Example 5.3 at three noise levels.*

5.2. Numerical results for parabolic problems. Now we present numerical results for one- and two-dimensional parabolic problems.

Example 5.3. $\Omega = (0,1)$, $T = 0.1$, $\sigma = 0$, $q^\dagger = 2 + \sin(2\pi x)e^{-2(1-x)}$, $u_0 = \sin(\pi x)$, and $f = 4x(1-x)$. The exact data are generated on a fine mesh with $h = 1/1600$ and $\tau = 1/8000$.

The numerical results for Example 5.3 are shown in Table 5.3 and Figure 5.3, where e_q is defined as before and e_u is defined by $e_u = (\tau \sum_{n=N_\sigma}^{N} \|U_h^n(q_h^*)(t_n) - u(q^\dagger)(t_n)\|_{L^2(\Omega)}^2)^{\frac{1}{2}}$. The regularization parameter γ, the mesh size h, and the time step size τ are chosen such that they all decrease with the noise level ε, as suggested by Corollary 4.6. One can check that the positivity condition (4.14) holds, and thus Corollary 4.6 is indeed applicable. We observe a very steady convergence for both

TABLE 5.4

Numerical results for Example 5.4: convergence with respect to ε, with γ, h and τ, initialized to $1.00e$-6, $8.33e$-2, and $1/1600$.

ε	5.00e-2	3.00e-2	1.00e-2	5.00e-3	3.00e-3	1.00e-3	5.00e-4	
e_q	1.95e-2	9.54e-3	4.32e-3	2.87e-3	2.28e-3	1.37e-3	9.37e-4	0.62
e_u	3.49e-3	1.70e-3	7.52e-4	3.92e-4	2.68e-4	7.26e-5	4.17e-5	0.94

(a) exact   (b) ε=1e-2   (c) ε=5e-2
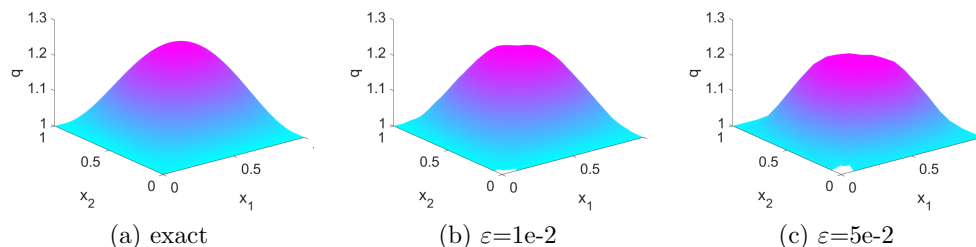
FIG. 5.4. *Numerical reconstructions for Example 5.4 at two noise levels.*

quantities e_q and e_u. The convergence rate for e_q is comparable with the elliptic cases in Examples 5.1 and 5.2; however, the rate for e_u is slightly slower, at a rate about $O(\delta^{0.64})$, when compared with the nearly $O(\delta)$ rate in Examples 5.1 and 5.2. The precise mechanism for this loss is still unclear.

Example 5.4. $\Omega = (0,1)^2$, $T = 0.1$, $q^\dagger(x_1, x_2) = 1 + (1-x_1)x_1 \sin(\pi x_2)$, $u_0(x_1, x_2) = 4x_1(1-x_1)$, and $f \equiv 1$. The exact data are generated on a finer mesh with $h = 1/200$ and $\tau = 1/12800$.

The numerical results for Example 5.4 are shown in Table 5.4 and Figure 5.4. The empirical rates with respect to ε and τ are largely comparable with the preceding examples, and the overall convergence is very steady.

In sum, the numerical experiments confirm the convergence of the Galerkin approximation in the $L^2(\Omega)$. However, the theoretical rate is still slower than the empirical one. It remains an important issue to derive sharp error estimates. In addition, it is also of interest to derive convergence rates with respect to h for the (nonlinear) optimal control problems (with fixed δ and γ), for which there seems to be no known result.

Appendix A. Basic estimates. We give an error bound on the Galerkin approximation. This estimate is used in the proof of Lemma 4.3.

LEMMA A.1. *Let* $q \in W^{1,\infty}(\Omega) \cap H^2(\Omega)$, *with* $c_0 \leq q(x) \leq c_1$ *a.e.* Ω. *Let* $u_h(q) \in X_h$ *and* $u_h(\mathcal{I}_h q) \in X_h$ *be the solutions to the variational problems*

$$(q \nabla u_h(q), \nabla v) = (f, v) \quad and \quad (\mathcal{I}_h q \nabla u_h(\mathcal{I}_h q), \nabla v) = (f, v) \quad \forall v \in X_h,$$

respectively. Then for any $\epsilon > 0$ *and* $p \geq \max(d + \epsilon, 2)$, *there holds*

$$\|u_h(q) - u_h(\mathcal{I}_h q)\|_{L^2(\Omega)} \leq ch^2 \|f\|_{L^p(\Omega)}.$$

Proof. By the definitions of $u_h(\mathcal{I}_h q)$ and $u_h(q)$, $w_h \equiv u_h(q) - u_h(\mathcal{I}_h q)$ satisfies

(A.1) $(q \nabla w_h, \nabla v) = ((\mathcal{I}_h q - q) \nabla u_h(\mathcal{I}_h q), \nabla v) \quad \forall v \in X_h.$

Since $q \geq c_0$, by the approximation property (2.2), we derive

$$c_0 \|\nabla w_h\|_{L^2(\Omega)}^2 \leq (q\nabla w_h, \nabla w_h) = ((\mathcal{I}_h q - q)\nabla u_h(\mathcal{I}_h q), \nabla w_h)$$
$$\leq c_1 \|\mathcal{I}_h q - q\|_{L^\infty(\Omega)} \|\nabla u_h(\mathcal{I}_h q)\|_{L^2(\Omega)} \|\nabla w_h\|_{L^2(\Omega)}$$
$$\leq ch\|q\|_{W^{1,\infty}(\Omega)} \|f\|_{L^2(\Omega)} \|\nabla w_h\|_{L^2(\Omega)},$$

i.e., $\|\nabla w_h\|_{L^2(\Omega)} \leq ch\|f\|_{L^2(\Omega)}$. Next, we derive the $L^2(\Omega)$ estimate by using a duality argument. Let $\psi \in H^2(\Omega) \cap H_0^1(\Omega)$ solve $(q\nabla v, \nabla \psi) = (v, w_h)$ for any $v \in H_0^1(\Omega)$. Meanwhile, we have

$$\|w_h\|_{L^2(\Omega)}^2 = (q\nabla w_h, \nabla \psi) = (q\nabla w_h, \nabla(\psi - \mathcal{I}_h \psi)) + (q\nabla w_h, \nabla \mathcal{I}_h \psi)$$
$$\leq \|q\|_{L^\infty(\Omega)} \|\nabla w_h\|_{L^2(\Omega)} \|\nabla(\psi - \mathcal{I}_h \psi)\|_{L^2(\Omega)} + (q\nabla w_h, \nabla \mathcal{I}_h \psi)$$
$$\leq ch^2 \|\psi\|_{H^2(\Omega)} \|f\|_{L^2(\Omega)} + (q\nabla w_h, \nabla \mathcal{I}_h \psi).$$

Further, using (A.1) and the a priori estimate $\|u(q)\|_{W^{1,\infty}(\Omega)} \leq c\|f\|_{L^p(\Omega)}$ for any $p \geq \max(d+\epsilon, 2)$ [30, equation (2.2)] and the estimate $\|\nabla(u(q) - u_h(q))\|_{L^2(\Omega)} \leq ch$, we obtain

$$(q\nabla w_h, \nabla \mathcal{I}_h \psi) = ((\mathcal{I}_h q - q)\nabla u_h(\mathcal{I}_h q), \nabla \mathcal{I}_h \psi)$$
$$= ((\mathcal{I}_h q - q)\nabla[u_h(\mathcal{I}_h q) - u(q)], \nabla \mathcal{I}_h \psi) + ((\mathcal{I}_h q - q)\nabla u(q), \nabla \mathcal{I}_h \psi)$$
$$\leq \|\mathcal{I}_h q - q\|_{L^\infty(\Omega)} \|\nabla(u_h(\mathcal{I}_h q) - u(q))\|_{L^2(\Omega)} \|\nabla \mathcal{I}_h \psi\|_{L^2(\Omega)}$$
$$\quad + \|\mathcal{I}_h q - q\|_{L^2(\Omega)} \|\nabla u(q)\|_{L^\infty(\Omega)} \|\nabla \mathcal{I}_h \psi\|_{L^2(\Omega)}$$
$$\leq ch^2 \|\nabla \mathcal{I}_h \psi\|_{L^2(\Omega)} \|f\|_{L^p(\Omega)} \leq ch^2 \|\psi\|_{H^2(\Omega)} \|f\|_{L^p(\Omega)}.$$

This and the regularity $\|\psi\|_{H^2(\Omega)} \leq c\|w_h\|_{L^2(\Omega)}$ lead to

$$\|u_h(q) - u_h(\mathcal{I}_h q)\|_{L^2(\Omega)} \leq ch^2 \|f\|_{L^p(\Omega)}$$

for any $p \geq \max(d+\epsilon, 2)$. This completes the proof of the lemma. $\qquad \square$

Appendix B. Proof of Lemma 4.2.

Proof. If $f \equiv 0$ and $u_0 \in H^2(\Omega) \cap H_0^1(\Omega)$, the estimate can be found in [34, Theorem 3.1]. It suffices to analyze the case $u_0 = 0$ and $f \in W^{2,1}(0,T; L^2(\Omega))$. Let $A \equiv A(q^\dagger) : H_0^1(\Omega) \to H^{-1}(\Omega)$ by $(Av, \chi) = (q^\dagger \nabla v, \nabla \chi)$ for all $\chi \in H_0^1(\Omega)$. Then A generates a bounded analytic semigroup e^{-At} on $L^2(\Omega)$ and allows representing the solution $u(t)$ by

$$u(t) = \int_0^t e^{-A(t-s)} f(s)\, ds.$$

Then it follows from integration by parts that

$$\|\partial_t u(t)\|_{L^2(\Omega)} + \|Au(t)\|_{L^2(\Omega)} \leq c\|f\|_{C([0,t];L^2(\Omega))} + \int_0^t \|\partial_s f(s)\|_{L^2(\Omega)}\, ds,$$
$$\|A\partial_t u(t)\|_{L^2(\Omega)} \leq c\Big(t^{-1}\|f(0)\|_{L^2(\Omega)} + \|f'(t)\|_{L^2(\Omega)} + \int_0^t \|\partial_s^2 f(s)\|_{L^2(\Omega)}\, ds\Big).$$

The second inequality and Assumption 4.1 imply

$$\int_0^t s\|A\partial_s u(s)\|_{L^2(\Omega)}\, ds \leq ct.$$

Then by the regularity estimate (4.9) and the approximation property (2.1), we derive

$$(B.1) \qquad \|u(t) - P_h u(t)\|_{L^2(\Omega)} \le ch^2 \|u(t)\|_{H^2(\Omega)} \le ch^2 \|Au(t)\|_{L^2(\Omega)}.$$

Let u_h be the spatially semidiscrete Galerkin approximation, i.e., $\partial_t u_h + A_h u_h = P_h f$ with $u_h(0) = 0$ and $A_h \equiv A_h(q^\dagger)$, cf. (2.3). Then the difference $\zeta(t) = u_h(t) - P_h u(t)$ satisfies

$$\partial_t \zeta(t) + A_h \zeta(t) = A_h(R_h - P_h)u(t),$$

with $\zeta(0) = 0$, where $R_h : H_0^1(\Omega) \to X_h$ denotes the Ritz projection (associated with $q^\dagger$). Then (2.1) and the approximation property of R_h [34, Lemma 1.1] lead to

$$t\zeta(t) = t \int_0^t e^{-A_h(t-s)} A_h(R_h - P_h)u(s) \, \mathrm{d}s$$
$$= \int_0^t (t-s)e^{-A_h(t-s)} A_h(R_h - P_h)u(s) \, \mathrm{d}s + \int_0^t e^{-A_h(t-s)} A_h(R_h - P_h)su(s) \, \mathrm{d}s$$
$$=: \mathrm{I}_1 + \mathrm{I}_2.$$

Since $\|e^{-A_h t} A_h\|_{L^2(\Omega) \to L^2(\Omega)} \le ct^{-1}$, we deduce

$$\|\mathrm{I}_1\|_{L^2(\Omega)} \le \int_0^t \|(R_h - P_h)u(s)\|_{L^2(\Omega)} \, \mathrm{d}s \le ch^2 \int_0^t \|Au(s)\|_{L^2(\Omega)} \, \mathrm{d}s \le cth^2.$$

Similarly, integration by parts allows bounding I_2 by

$$\|\mathrm{I}_2\|_{L^2(\Omega)} \le ct\|(R_h - P_h)u(t)\|_{L^2(\Omega)} + c \int_0^t \|(R_h - P_h)\partial_s(su(s))\|_{L^2(\Omega)} \, \mathrm{d}s$$
$$\le cth^2 \|Au\|_{C([0,t];L^2(\Omega))} + ch^2 \int_0^t s\|A\partial_s(su(s))\|_{L^2} \, \mathrm{d}s \le cth^2.$$

The preceding two estimates yield $\|\zeta(t)\|_{L^2(\Omega)} \le ch^2$. This, (B.1), and the triangle inequality imply $\|u_h(t) - u(t)\|_{L^2(\Omega)} \le ch^2$. Meanwhile, repeating the argument in [24, Lemma 4.2] yields

$$\|u_h(t_n) - U_h^n(q^\dagger)\|_{L^2(\Omega)} \le c\tau \left(\|f(0)\|_{L^2(\Omega)} + \int_0^{t_n} \|\partial_s f(s)\|_{L^2(\Omega)} \, \mathrm{d}s \right) \le c\tau.$$

Then the desired assertion follows immediately by the triangle inequality. $\quad\square$

REFERENCES

[1] R. Acar, *Identification of the coefficient in elliptic equations*, SIAM J. Control Optim., 31 (1993), pp. 1221–1244, https://doi.org/10.1137/0331058.

[2] R. A. Adams and J. J. F. Fournier, *Sobolev Spaces*, 2nd ed., Elsevier/Academic Press, Amsterdam, 2003.

[3] M. F. Al-Jamal and M. S. Gockenbach, *Stability and error estimates for an equation error method for elliptic equations*, Inverse Problems, 28 (2012), 095006, https://doi.org/10.1088/0266-5611/28/9/095006.

[4] A. Ashyralyev, S. Piskarev, and L. Weis, *On well-posedness of difference schemes for abstract parabolic equations in $L^p([0,T];E)$ spaces*, Numer. Funct. Anal. Optim., 23 (2002), pp. 669–693, https://doi.org/10.1081/NFA-120016264.

[5] G. Bal and G. Uhlmann, *Reconstruction of coefficients in scalar second-order elliptic equations from knowledge of their solutions*, Comm. Pure Appl. Math., 66 (2013), pp. 1629–1652, https://doi.org/10.1002/cpa.21453.

[6] H. T. Banks and K. Kunisch, *Estimation Techniques for Distributed Parameter Systems*, Birkhäuser Boston, Boston, 1989, https://doi.org/10.1007/978-1-4612-3700-6.

[7] A. Bonito, A. Cohen, R. DeVore, G. Petrova, and G. Welper, *Diffusion coefficients estimation for elliptic partial differential equations*, SIAM J. Math. Anal., 49 (2017), pp. 1570–1592, https://doi.org/10.1137/16M1094476.

[8] H. Brezis and P. Mironescu, *Gagliardo-Nirenberg inequalities and non-inequalities: The full story*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 35 (2018), pp. 1355–1376, https://doi.org/10.1016/j.anihpc.2017.11.007.

[9] Z. Chen and J. Zou, *An augmented Lagrangian method for identifying discontinuous parameters in elliptic systems*, SIAM J. Control Optim., 37 (1999), pp. 892–910, https://doi.org/10.1137/S0363012997318602.

[10] K. Deckelnick and M. Hinze, *Convergence and error analysis of a numerical method for the identification of matrix parameters in elliptic PDEs*, Inverse Problems, 28 (2012), 115015, https://doi.org/10.1088/0266-5611/28/11/115015.

[11] J. Douglas, Jr., T. Dupont, and L. Wahlbin, *The stability in L^q of the L^2-projection into finite element function spaces*, Numer. Math., 23 (1974/75), pp. 193–197, https://doi.org/10.1007/BF01400302.

[12] H. Egger and B. Hofmann, *Tikhonov regularization in Hilbert scales under conditional stability assumptions*, Inverse Problems, 34 (2018), 115015, https://doi.org/10.1088/1361-6420/aadef4.

[13] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, Kluwer Academic, Dordrecht, the Netherlands, 1996.

[14] H. W. Engl, K. Kunisch, and A. Neubauer, *Convergence rates for Tikhonov regularisation of nonlinear ill-posed problems*, Inverse Problems, 5 (1989), pp. 523–540, http://stacks.iop.org/0266-5611/5/523.

[15] A. Ern and J.-L. Guermond, *Theory and Practice of Finite Elements*, Springer-Verlag, New York, 2004.

[16] R. S. Falk, *Error estimates for the numerical identification of a variable coefficient*, Math. Comp., 40 (1983), pp. 537–546, https://doi.org/10.2307/2007529.

[17] A. Friedman, *Remarks on the maximum principle for parabolic equations and its applications*, Pacific J. Math., 8 (1958), pp. 201–211, http://projecteuclid.org/euclid.pjm/1103040096.

[18] E. O. Frind and G. F. Pinder, *Galerkin solution of the inverse problem for aquifer transmissivity*, Water Resources Res., 9 (1973), pp. 1397–1410.

[19] M. Grüter and K.-O. Widman, *The Green function for uniformly elliptic equations*, Manuscripta Math., 37 (1982), pp. 303–342, https://doi.org/10.1007/BF01166225.

[20] S. Gutman, *Identification of discontinuous parameters in flow equations*, SIAM J. Control Optim., 28 (1990), pp. 1049–1060, https://doi.org/10.1137/0328057.

[21] M. Hinze, B. Kaltenbacher, and T. N. T. Quyen, *Identifying conductivity in electrical impedance tomography with total variation regularization*, Numer. Math., 138 (2018), pp. 723–765, https://doi.org/10.1007/s00211-017-0920-8.

[22] V. Isakov, *Inverse Problems for Partial Differential Equations*, 2nd ed., Springer, New York, 2006.

[23] K. Ito and B. Jin, *Inverse Problems: Tikhonov Theory and Algorithms*, World Scientific, Hackensack, NJ, 2015.

[24] B. Jin, B. Li, and Z. Zhou, *Numerical analysis of nonlinear subdiffusion equations*, SIAM J. Numer. Anal., 56 (2018), pp. 1–23, https://doi.org/10.1137/16M1089320.

[25] B. Jin and Z. Zhou, *Numerical Estimation of a Diffusion Coefficient in Subdiffusion*, preprint, arXiv:1909.00334, 2019.

[26] T. Kärkkäinen, *An equation error method to recover diffusion from the distributed observation*, Inverse Problems, 13 (1997), pp. 1033–1051, https://doi.org/10.1088/0266-5611/13/4/009.

[27] Y. L. Keung and J. Zou, *Numerical identifications of parameters in parabolic systems*, Inverse Problems, 14 (1998), pp. 83–100, https://doi.org/10.1088/0266-5611/14/1/009.

[28] R. V. Kohn and B. D. Lowe, *A variational method for parameter identification*, RAIRO Modél. Math. Anal. Numér., 22 (1988), pp. 119–158, https://doi.org/10.1051/m2an/1988220101191.

[29] C. Kravaris and J. H. Seinfeld, *Identification of parameters in distributed parameter systems by regularization*, SIAM J. Control Optim., 23 (1985), pp. 217–241, https://doi.org/10.1137/0323017.

[30] B. Li and W. Sun, *Maximal L^p analysis of finite element solutions for parabolic equations with nonsmooth coefficients in convex polyhedra*, Math. Comp., 86 (2017), pp. 1071–1102, https://doi.org/10.1090/mcom/3133.

[31] V. A. MOROZOV, *On the solution of functional equations by the method of regularization*, Soviet Math. Dokl., 7 (1966), pp. 414–417.

[32] G. R. RICHTER, *Numerical identification of a spatially varying diffusion coefficient*, Math. Comp., 36 (1981), pp. 375–386, https://doi.org/10.2307/2007648.

[33] X.-C. TAI AND T. KÄRKKÄINEN, *Identification of a nonlinear parameter in a parabolic equation from a linear equation*, Mat. Apl. Comput., 14 (1995), pp. 157–184.

[34] V. THOMÉE, *Galerkin Finite Element Methods for Parabolic Problems*, 2nd ed., Springer-Verlag, Berlin, 2006.

[35] L. WANG AND J. ZOU, *Error estimates of finite element methods for parameter identifications in elliptic and parabolic systems*, Discrete Contin. Dyn. Syst. Ser. B, 14 (2010), pp. 1641–1670, https://doi.org/10.3934/dcdsb.2010.14.1641.

[36] W. W. G. YEH, *Review of parameter identification procedures in groundwater hydrology: The inverse problem*, Water Resources Res., 22 (1986), pp. 95–108.

[37] J. ZOU, *Numerical methods for elliptic inverse problems*, Int. J. Comput. Math., 70 (1998), pp. 211–232, https://doi.org/10.1080/00207169808804747.