

LEAST SQUARES SOLVERS FOR ILL-POSED PDES THAT ARE CONDITIONALLY STABLE

WOLFGANG DAHMEN¹, HARALD MONSUUR² AND ROB STEVENSON^{2,*} 

Abstract. This paper is concerned with the design and analysis of least squares solvers for ill-posed PDEs that are conditionally stable. The norms and the regularization term used in the least squares functional are determined by the ingredients of the conditional stability assumption. We are then able to establish a general error bound that, in view of the conditional stability assumption, is qualitatively the best possible, without assuming consistent data. The price for these advantages is to handle dual norms which reduces to verifying suitable inf-sup stability. This, in turn, is done by constructing appropriate Fortin projectors for all sample scenarios. The theoretical findings are illustrated by numerical experiments.

Mathematics Subject Classification. 35B30, 35B35, 35B45, 35R25, 65F08, 65J20, 65M12, 65N12.

Received July 15, 2022. Accepted May 30, 2023.

1. INTRODUCTION

In this paper a general approach is developed for the numerical solution of ill-posed boundary value problems $Au = f$, where $A \in \mathcal{L}(X, V)$ for Hilbert spaces X and V , that are *conditionally stable*. The latter means that for all $u \in X$ that satisfy an *a priori* bound of the form $\|Lu\|_H \leq C$, for some $L \in \mathcal{L}(X, H)$ and a Hilbert space H , there is a continuous dependency of the solution on the data in the sense that, for some $\eta = \eta_C: \mathbb{R}^+ \rightarrow \mathbb{R}^+$ with $\lim_{t \downarrow 0} \eta(t) = 0$, it holds that $j(u) \leq \eta(\|Au\|_V)$ for some $j: X \rightarrow \mathbb{R}_+$. In applications $j(\cdot)$ is a (semi-) norm that is weaker than the norm on X , and, *e.g.*, $\eta(t) = t^s$ or even only $\eta(t) = (-1/\log t)^s$ for some $s \in (0, 1]$. Conditional stability has been established for various ill-posed PDEs including data-assimilation and Cauchy boundary data problems for Poisson's, heat and wave equations.

For finite dimensional subspaces $X^\delta \subset X$ (“ δ ” refers to “discrete”), we approximate u by the minimizer u_ε^δ over X^δ of the regularized least-squares functional $z \mapsto \sqrt{\|Az - f\|_V^2 + \varepsilon^2 \|Lz\|_H^2}$. For a suitable selection of ε , it will be shown that both $\|L(u - u_\varepsilon^\delta)\|_H$ is uniformly bounded, so that $j(u - u_\varepsilon^\delta) \leq \eta(\|f - Au_\varepsilon^\delta\|_V)$, and $\|f - Au_\varepsilon^\delta\|_V$ is bounded by an absolute multiple of $\|f - Au\|_V + \min_{z \in X^\delta} \|u - z\|_X$, being the sum of the consistency error

Keywords and phrases. Tikhonov regularization, least squares methods, conditional stability, dual norms, inf-sup stability, Fortin projectors, mixed formulations, *a posteriori* bounds for residuals.

¹ Mathematics Department, University of South Carolina, Columbia, SC 29208, USA.

² Korteweg-de Vries (KdV) Institute for Mathematics, University of Amsterdam, P.O. Box 94248, 1090 GE Amsterdam, The Netherlands.

*Corresponding author: rob.p.stevenson@gmail.com

and the error of best approximation. Consequently, we will achieve qualitatively the best possible bound on the error quantity $j(u - u_\varepsilon^\delta)$ that can be expected for $u_\varepsilon^\delta \in X^\delta$ in view of the conditional stability estimate.

In applications often V is of product form $\prod_i V_i$, and so $A = (A_i)$ and $f = (f_i)$, with one or more V_i being a Sobolev space with negative smoothness index, which is a natural space for a forcing term of a PDE, or a fractional Sobolev space on (a part of) the boundary of the computational domain, which is a natural space for a boundary datum. The norms on such spaces cannot be evaluated exactly.

An option to deal with a Sobolev norm of negative smoothness index is to replace it by an L_2 -norm. This, however, requires more smoothness of the data and more regularity of X^δ , *e.g.*, a C^1 - instead of a C^0 -finite element space, whereas it is not ensured that the error benefits from smallness of the residual in a stronger norm. Similar disadvantages are connected to the replacement of fractional Sobolev norms by (weighted) L_2 -norms.

Our approach to deal with V_i being a Sobolev space with negative smoothness index, *i.e.*, a V_i being of the form Y_i' , is to replace $\|\cdot\|_{Y_i'}$ in the least-squares functional by a discrete dual norm $\|\cdot\|_{Y_i^{\delta'}}$, where $Y_i^\delta = Y_i^\delta(X^\delta) \subset Y_i$ is such that $\|A_i \cdot\|_{Y_i'}$ is equivalent to $\|A_i \cdot\|_{Y_i^{\delta'}}$ on X^δ , and $\dim Y_i^\delta$ is proportional to $\dim X^\delta$. The first property is known to be equivalent to existence of a (uniformly bounded) Fortin projector $Y_i \rightarrow Y_i^\delta$.

By introducing the Riesz lift of the corresponding residual $f_i - A_i u_\varepsilon^\delta \in Y_i^{\delta'}$ as an independent variable, the resulting least squares problem has an equivalent formulation as a mixed system, which does not involve the dual norm $\|\cdot\|_{Y_i^{\delta'}}$, and which is Ladyshenskaja–Babuška–Brezzi (LBB) stable by virtue of the existence of the Fortin projector. In many cases, one can construct a $G_{Y_i}^\delta: Y_i^{\delta'} \rightarrow Y_i^\delta$ (known as a preconditioner) with $(G_{Y_i}^\delta)^{-1}v(v)$ equivalent to $\|v\|_{Y_i}^2$, and whose application can be performed in linear complexity, in which case one can efficiently eliminate the additional variable and so retrieves a symmetric positive definite system.

We handle fractional Sobolev norms in the same manner. Viewing a fractional Sobolev space V_i , with either a positive or negative smoothness index, as the dual of $Y_i := V_i'$, first we construct a (uniformly bounded) Fortin projector $Y_i \rightarrow Y_i^\delta$, and second, to avoid having to compute fractional norms with opposite index of arguments from Y_i^δ , we use a preconditioner $G_{Y_i}^\delta: Y_i^{\delta'} \rightarrow Y_i^\delta$ with $(G_{Y_i}^\delta)^{-1}v(v)$ equivalent to $\|v\|_{Y_i}^2$.

The steps to handle dual or fractional norms, mentioned above, make our approach practically feasible without compromising its attractive theoretical properties. Indeed, still one obtains a bound on the error quantity $j(u - u_\varepsilon^\delta)$ that is qualitatively the best possible.

We exemplify our approach by constructing Fortin interpolators and preconditioners for the examples of the Cauchy problem for Poisson's equation, and data-assimilation problems for wave- and heat-equations. Furthermore, for those examples we illustrate our theoretical findings with numerical results.

Our approach to minimize a regularized least squares functional is of course not new. Not making use of conditional stability, in [5, 8, 9] this method was analyzed for a regularizing term $\varepsilon^2 \|z\|_X^2$, instead of our choice $\varepsilon^2 \|Lz\|_H^2$ suggested by the conditional stability condition. By replacing the test function from X by minus this test function, a non-symmetric mixed system on $X \times V$ is obtained that is coercive, with a coercivity constant that is, however, proportional to ε^2 . With this formulation the notion of stability is fully due to the stabilization term. By our approach to guarantee LBB-stability, the operator A contributes to the stability of the least-squares problem. It results in a proof of convergence rates in the error quantity associated to the conditional stability estimate that seems new.

The use of conditional stability estimates for the numerical solution of various ill-posed PDEs has been advocated in series of papers [11–19]. In those works a control functional is minimized under the constraint that the state satisfies the PDE. Instead of adding Tikhonov stabilization at the continuous level, mesh-dependent stabilization terms tailored to the application at hand are added to the finite element discretization.

1.1. Layout

In Section 2 we recall the concept of conditional stability for ill-posed problems $Au = f$. Under the provision that approximations u_ε^δ to u from finite dimensional spaces X^δ are available that satisfy a certain quasi-optimal error bound in an ε -dependent energy norm, it will be demonstrated that, for a judiciously chosen regularization

parameter ε , a qualitatively best possible upper bound holds for the error quantity $j(u - u_\varepsilon^\delta)$. In Section 3 we present several classical examples to which the theory applies. In Section 4 the aforementioned quasi-optimal error bound will be demonstrated for $u_\varepsilon^\delta \in X^\delta$ being the minimizer of a regularized least squares functional, in which, under inf-sup conditions, dual norms are replaced by discrete dual norms. The resulting least-squares problem has an equivalent formulation as a mixed system. Section 5 is devoted to a reformulation of the mixed problem as a symmetric positive definite variational problem, based on uniform preconditioners that serve as approximate Riesz lifters in those Hilbert space components that require the use of dual norms. In Section 6 we discuss *a posteriori* residual estimators. In Section 7 we verify the validity of the critical inf-sup conditions for all sample problems, the results in preceding sections hinge upon. Here the central work horse are suitable Fortin operators. Finally, in Section 8 we present numerical results for our three sample problems.

1.2. Notation

In this work, by $C \lesssim D$ we will mean that C can be bounded by a multiple of D , *independently* of parameters which C and D may depend on, as the discretisation index δ , the tolerance τ for the consistency error, and the regularization parameters ε and ζ . Obviously, $C \gtrsim D$ is defined as $D \lesssim C$, and $C \approx D$ as $C \lesssim D$ and $C \gtrsim D$.

For normed linear spaces E and F , by $\mathcal{L}(E, F)$ we will denote the normed linear space of bounded linear mappings $E \rightarrow F$, and by $\text{Lis}(E, F)$ its subset of boundedly invertible linear mappings $E \rightarrow F$. We write $E \hookrightarrow F$ to denote that E is continuously embedded into F . For convenience only, we exclusively consider linear spaces over the scalar field $\mathbb{R}$.

The set $[0, \infty)$ will be denoted by $\mathbb{R}^+$.

2. PROBLEM SETTING AND MAIN RESULT

For Hilbert spaces X and V , we consider operators $A \in \mathcal{L}(X, V)$ which are neither assumed to be injective nor to have a dense range in V . We study the problem of the (approximate) reconstruction of $u \in X$ from its image Au assuming we are only given a perturbation f of Au for which

$$\|f - Au\|_V \leq \tau \quad (2.1)$$

holds for some tolerance $\tau \geq 0$ which we assume to be known. Since in particular we do not assume bounded invertibility of A our problem is ill-posed, even for $\tau = 0$. Although for convenience we refer in the following to u as *the* solution of our recovery problem, one should bear in mind that for $\tau > 0$ there may be multiple $u \in X$ that satisfy (2.1). Our results will be valid uniformly in those u .

Since Tikhonov [40] with *stability* for ill-posed problems, usually called *conditional stability*, one understands *some* continuous dependency of the solution upon the data, typically with respect to a weaker metric than that induced by $\|\cdot\|_X$, under the assumption that a bound on the solution itself is available. More specifically, similar to [11], we assume existence of an $L \in \mathcal{L}(X, H)$, where H is some additional Hilbert space, such that the following assumption is valid:

Assumption 2.1 (Conditional stability). *The pair*

$$(A, L) \in \mathcal{L}(X, V \times H) \text{ is injective,}$$

and there exists a $j: X \rightarrow \mathbb{R}^+$, and for any $\mathcal{C} > 0$, a non-decreasing $\eta = \eta_{\mathcal{C}}: \mathbb{R}^+ \rightarrow \mathbb{R}^+$ with $\lim_{t \downarrow 0} \eta(t) = 0$, such that for $z \in X$ with $\|Lz\|_H \leq \mathcal{C}$, it holds that

$$j(z) \leq \eta_{\mathcal{C}}(\|Az\|_V). \quad (2.2)$$

Typically, j is a norm or a semi-norm on a Hilbert space $\tilde{H} \hookrightarrow X$. In the first case, A is injective. Several examples will be given in Section 3.

For $\varepsilon > 0$, we set

$$\|\cdot\|_\varepsilon := \sqrt{\|A \cdot\|_V^2 + \varepsilon^2 \|L \cdot\|_H^2}. \quad (2.3)$$

To use precisely the ingredients of the conditional stability condition in the definition of $\|\cdot\|_\varepsilon$ will be seen to be essential in what follows. Moreover, notice that $\|\cdot\|_\varepsilon$ is a norm on X by our assumption of (A, L) being injective.

Thinking of ε being small, tacitly we will always assume that $\varepsilon \|L\|_{\mathcal{L}(X, H)} \lesssim 1$ so that, since A is bounded, $\|\cdot\|_\varepsilon \lesssim \|\cdot\|_X$.

Given a *finite dimensional* subspace X^δ of X , in Sections 4 and 5 we show how to compute for each $\varepsilon > 0$ a $u_\varepsilon^\delta \in X^\delta$ satisfying

$$\|u - u_\varepsilon^\delta\|_\varepsilon \lesssim \tau + \min_{z \in X^\delta} \|u - z\|_\varepsilon + \varepsilon \|Lu\|_H. \quad (2.4)$$

If (2.2) is valid for $L = 0$, and thus $\eta_{\mathcal{C}} \equiv \eta$ is independent of $\mathcal{C}$, then one speaks about *unconditional stability* of (2.1)¹. In this case $\|\cdot\|_\varepsilon = \|A \cdot\|_V$ is ε -independent, and so will be u_ε^δ .

Theorem 2.2. Assume (2.4) and recall Assumption 2.1. For $L \neq 0$, let $\varepsilon = \varepsilon(\tau, \delta) > 0$ be such that for some $\mathcal{C} \geq \|Lu\|_H$,

$$\tau + \min_{z \in X^\delta} \|u - z\|_\varepsilon \lesssim \varepsilon \mathcal{C} \quad \text{and} \quad \varepsilon \|Lu\|_H \lesssim \tau + \min_{z \in X^\delta} \|u - z\|_X. \quad (2.5)$$

Then

$$j(u - u_\varepsilon^\delta) \leq \eta_{\mathcal{C}}(\|A(u - u_\varepsilon^\delta)\|_V), \quad (2.6)$$

and

$$\|A(u - u_\varepsilon^\delta)\|_V \lesssim \tau + \min_{z \in X^\delta} \|u - z\|_X. \quad (2.7)$$

Proof. When $L = 0$, (2.6) holds unconditionally, and $\|\cdot\|_\varepsilon = \|A \cdot\|_V \lesssim \|\cdot\|_X$ so that (2.7) follows directly from (2.4). So let $L \neq 0$. From (2.4) and the lower bound from (2.5) on ε , one derives $\|L(u - u_\varepsilon^\delta)\|_H \leq \varepsilon^{-1} \|u - u_\varepsilon^\delta\|_\varepsilon \lesssim \mathcal{C}$. Conditional stability then ensures

$$j(u - u_\varepsilon^\delta) \leq \eta_{\mathcal{C}}(\|A(u - u_\varepsilon^\delta)\|_V).^2$$

Again (2.4), and the upper bound on ε from (2.5) show that

$$\|A(u - u_\varepsilon^\delta)\|_V \leq \|u - u_\varepsilon^\delta\|_\varepsilon \lesssim \tau + \min_{z \in X^\delta} \|u - z\|_X,$$

which completes the proof. $\square$

Notice that when $\min_{z \in X^\delta} \|u - z\|_X$ in (2.5) is replaced by an upper bound, then one arrives at (2.7) with $\min_{z \in X^\delta} \|u - z\|_X$ replaced by that upper bound.

In the examples given in Section 3, $\eta_{\mathcal{C}}(t)$ will be of the form $\mathcal{O}((t + \mathcal{C})^{1-\sigma} t^\sigma)$ or $\mathcal{O}((t + \mathcal{C})(\log(1 + \frac{\mathcal{C}}{t}))^{-\sigma})$ for some $\sigma \in (0, 1)$, or $\mathcal{O}(t)$.

Remark 2.3 (Optimality of estimates). The bound (2.7) on $\|A(u - u_\varepsilon^\delta)\|_V$, (valid because of the upper bound on ε), by a multiple of the sum of the (maximal) *consistency error* τ and the *approximation error* $\min_{z \in X^\delta} \|u - z\|_X$, is qualitatively the best that can be expected for a numerical approximation from X^δ . Since, thanks to the lower bound on ε , at the same time $\|L(u - u_\varepsilon^\delta)\|_H \lesssim \mathcal{C}$, we obtain the generally qualitatively best possible upper bound for $j(u - u_\varepsilon^\delta)$ that is permitted by the conditional stability estimate.

¹Not to be confused with *well-posedness*, with which we mean $A \in \mathcal{L}(X, V)$. An unconditionally stable problem where $\text{ran } A$ is closed is also benign in the sense that then, by an application of the open mapping theorem, equation (2.1) is *well-posed in least-squares sense*, i.e., $A^*A \in \mathcal{L}(X, X)$, and thus is also not of our primary interest.

²Although in applications, the decay of $\eta_{\mathcal{C}}(t)$ for $t \downarrow 0$ is faster when simultaneously $\mathcal{C} = \mathcal{C}(t) \downarrow 0$, by the very character of a conditional stability estimate $\|Lz\|_H \downarrow 0$ cannot be expected to follow from $\|Az\|_V \downarrow 0$.

That being said, inserting the upper bound in (2.7) into (2.6) can nevertheless provide a pessimistic bound. The reason is that perturbations in the data enter the approximation by solving a discretized problem, whose conditioning is for coarser and coarser meshes usually increasingly better than that of the infinite dimensional problem whose behaviour is captured by the conditional stability estimate.

Remark 2.4 (Selection of ε when $L \neq 0$). Let (X^δ) be a family of finite dimensional subspaces of X such that for some $s > 0$ for general, sufficiently smooth $u \in X$ it holds $\min_{z \in X^\delta} \|u - z\|_X \approx (\dim X^\delta)^{-s}$. Then, in view of (2.5), an obvious choice is to take $\varepsilon \approx \tau + (\dim X^\delta)^{-s}$. Because of a lacking smoothness of u , it might be, however, that this ε decays too fast for $\dim X^\delta \rightarrow \infty$ which then would manifest itself by an increase of $\|Lu_\varepsilon^\delta\|_H$. Indeed, recall that the sole reason for imposing the lower bound on ε in (2.5) is to prevent an unbounded growth of $\|Lu_\varepsilon^\delta\|_H$, and thus of $\|L(u - u_\varepsilon^\delta)\|_H$, which would jeopardize a meaningful application of the conditional stability estimate. The value of $\|Lu_\varepsilon^\delta\|_H$, however, can be monitored and so the choice of ε can be adapted when such a growth of $\|Lu_\varepsilon^\delta\|_H$ is observed.

In various numerical experiments we observed that regularization is actually not needed at all, whereas for other data ε equal to τ was close to the experimentally found best regularization parameter. In our tests, where u was smooth, we did not encounter an example where it was helpful to take ε equal to τ plus a “mesh-dependent” term that approximates $\min_{z \in X^\delta} \|u - z\|_X$. A probable explanation is the better conditioning of the discretized problems on coarser meshes.

Remark 2.5 (The case that (A, L) is only closed). This section started by assuming a pair $(A, L) \in \mathcal{L}(X, V \times H)$. If for some Hilbert space $\tilde{X}$, $(A, L): \tilde{X} \supset \text{dom}(A, L) \rightarrow V \times H$ is only linear and closed, then by defining $X := \{z \in \tilde{X}: (Az, Lz) \in V \times H\}$ equipped with the graph norm, we are back in the situation required for Assumption 2.1.

Remark 2.6 (About closedness of $\text{ran}(A, L)$). By injectivity of (A, L) from Assumption 2.1, the open mapping theorem shows that for any $\varepsilon > 0$, closedness of $\text{ran}(A, \varepsilon L) \subset V \times H$ is equivalent to $(A, \varepsilon L) \in \mathcal{L}\text{is}(X, \text{ran}(A, \varepsilon L))$, being equivalent to $\|\cdot\|_\varepsilon \approx \|\cdot\|_X$ (obviously generally *dependent* on ε). The latter shows that closedness of $\text{ran}(A, \varepsilon L)$ implies that $(X, \|\cdot\|_\varepsilon)$ is a Hilbert space, as well as that closedness of $\text{ran}(A, \varepsilon L)$ is equivalent to closedness of $\text{ran}(A, L)$.

Conversely, if $\text{ran}(A, L)$ is *not* closed, so that $\|\cdot\|_\varepsilon$ is *not* equivalent $\|\cdot\|_X$, then from $\|\cdot\|_\varepsilon \lesssim \|\cdot\|_X$ and the open mapping theorem it follows that $(X, \|\cdot\|_\varepsilon)$ is *not* a Hilbert space.

An advantage offered by a pair (A, L) with closed $\text{ran}(A, L)$ is that one can bound the condition number of the linear system that determines this approximation (see Rem. 5.3). In that context note that if Assumption 2.1 holds for some $(A, L) \in \mathcal{L}(X, V \times H)$, then it holds also for $(A, \text{Id}) \in \mathcal{L}(X, V \times X)$ with $j^{\text{new}} := j$ and $\eta_C^{\text{new}} := \eta_C \|L\|_{\mathcal{L}(V, H)}$, where now $\text{ran}(A, \text{Id})$ is closed. Despite this advantage we will not insist on closedness of $\text{ran}(A, L)$ in what follows. One argument for that is the following. If unconditional stability holds, then for computing u_ε^δ using $(A, 0)$ no regularization will be needed.

3. APPLICATIONS

Example 3.1 (Cauchy problem for Poisson’s equation). Let $\Omega \subset \mathbb{R}^d$ be a Lipschitz domain, and let Σ and Σ^c be open, measurable subsets of $\partial\Omega$ with $\Sigma \cap \Sigma^c = \emptyset$, $\overline{\Sigma} \cup \overline{\Sigma^c} = \partial\Omega$, and $|\Sigma| > 0$. Informally, the Cauchy problem asks for finding a solution to

$$-\Delta u = f_I \text{ on } \Omega, \quad u = f_D \text{ on } \Sigma, \quad \frac{\partial u}{\partial n} = f_N \text{ on } \Sigma, \quad (3.1)$$

i.e., Dirichlet and Neumann conditions are imposed on the same boundary portion of non-vanishing measure. To avoid unnecessarily restrictive assumptions on the data under which the formulation (3.1) would be meaningful, and to identify an appropriate least squares functional we employ the following rigorous weak formulation. Given $f = (f_I, f_D, f_N) \in H_{0, \Sigma^c}^1(\Omega)' \times H^{\frac{1}{2}}(\Sigma) \times H^{-\frac{1}{2}}(\Sigma)$, with $g_{f_I, f_N} := v \mapsto f_I(v) + \int_\Sigma f_N v \, ds \in H_{0, \Sigma^c}^1(\Omega)'$, we search a solution of

$$Au = (B_1 u, B_2 u) = (g_{f_I, f_N}, f_D), \quad (3.2)$$

where $(B_1, B_2) \in \mathcal{L}(\underbrace{H^1(\Omega)}_{X:=}, \underbrace{H_{0,\Sigma^c}^1(\Omega)' \times H^{\frac{1}{2}}(\Sigma)}_{V:=})$ is defined by $B_2 := \gamma_\Sigma$, being the trace operator on Σ , and B_1 represents the negative Laplacian as a mapping from X to $H_{0,\Sigma^c}^1(\Omega)'$

$$(B_1 z)(v) := \int_{\Omega} \nabla z \cdot \nabla v \, dx \quad (z \in X, v \in H_{0,\Sigma^c}^1(\Omega)).$$

Here $H_{0,\Sigma^c}^1(\Omega)$ is the closure in $H^1(\Omega)$ of the smooth functions on $\Omega \cup \Sigma$ with compact support, and $H^{-\frac{1}{2}}(\Sigma)$ is the dual of $\tilde{H}^{\frac{1}{2}}(\Sigma)$, the latter in the literature also denoted by $H_{00}^{\frac{1}{2}}(\Sigma)$. The dual of $H^{\frac{1}{2}}(\Sigma)$ is denoted by $\tilde{H}^{-\frac{1}{2}}(\Sigma)$.

For this problem it is known that $\ker A = 0$ and, when $|\Sigma^c| > 0$, $\text{ran } A \subsetneq \overline{\text{ran } A} = V$ (e.g., [9], Prop. 3.2). In Theorems 1.7, 1.9 and Remark 1.8 from [1], the following conditional *interior* and conditional *global* stability results have been established:

- (i) For $\omega \subset \Omega$ with $\text{dist}(\omega, \Sigma^c) > 0$, there exists a $\sigma \in (0, 1)$ such that

$$\|z\|_{L_2(\omega)} \lesssim (\|Az\|_V + \|z\|_{L_2(\Omega)})^{1-\sigma} \|Az\|_V^\sigma \quad (z \in X).$$

- (ii) There exists a $\sigma \in (0, 1)$ such that for $z \neq 0$

$$\|z\|_{L_2(\Omega)} \lesssim (\|Az\|_V + \|z\|_{H^1(\Omega)}) \left(\log \left(1 + \frac{\|z\|_{H^1(\Omega)}}{\|Az\|_V} \right) \right)^{-\sigma} \quad (z \in X).$$

Notice that (i) and (ii) are of the form as in Assumption 2.1 where $\|Lz\|_H$ and $\eta_{\mathcal{C}}(t)$ read as $\|z\|_{L_2(\Omega)}$ and $\mathcal{O}((t + \mathcal{C})^{1-\sigma} t^\sigma)$, or $\|z\|_{H^1(\Omega)}$ and $\mathcal{O}((t + \mathcal{C})(\log(1 + \mathcal{C}/t))^{-\sigma})$, respectively.

Example 3.2 (Data-assimilation for the heat equation). Let $\Omega \subset \mathbb{R}^d$ be a domain, $0 < T_1 < T_2 < T$, and $\emptyset \neq \omega \subset \Omega$ open. With $I := (0, T)$, given $(f, g) \in L_2(I; H^{-1}(\Omega)) \times L_2(I \times \omega)$, the data-assimilation problem reads as finding u with $\partial_t u - \Delta_x u = f$ on $I \times \Omega$, and $u|_{I \times \omega} = g$, or, more precisely, $Au = (f, g)$, where $A = (B, \Gamma_{I \times \omega}) \in \mathcal{L}(\underbrace{L_2(I; H^1(\Omega)) \cap H^1(I; H^{-1}(\Omega))}_{X:=}, \underbrace{L_2(I; H^{-1}(\Omega)) \times L_2(I \times \omega)}_{V:=})$ is defined by $\Gamma_{I \times \omega} z = z|_{I \times \omega}$,

and $(Bz)(v) := \int_I \int_{\Omega} \partial_t z v + \nabla_x z \cdot \nabla_x v \, dx \, dt$.

The following conditional stability estimates can be found in [13]:

- (a) For a bounded $\tilde{\omega} \Subset \Omega$, there exists a $\sigma \in (0, 1)$ such that

$$\|z\|_{L_2((T_1, T_2); H^1(\tilde{\omega}))} \lesssim (\|Az\|_V + \|z\|_{L_2(I \times \Omega)})^{1-\sigma} \|Az\|_V^\sigma \quad (z \in X).$$

- (b) When one has the additional information that $u = 0$ on $I \times \partial\Omega$, then X should be redefined as $X := L_2(I; H_0^1(\Omega)) \cap H^1(I; H^{-1}(\Omega))$. For Ω being a bounded convex polytope, now it holds that

$$\|z\|_{L_2((T_1, T); H^1(\Omega)) \cap H^1((T_1, T); H^{-1}(\Omega))} \lesssim \|Az\|_V \quad (z \in X).$$

In Case (a), Assumption 2.1 is valid with $\|Lz\|_H$ and $\eta_{\mathcal{C}}(t)$ reading as $\|z\|_{L_2(I \times \Omega)}$ and $\mathcal{O}((t + \mathcal{C})^{1-\sigma} t^\sigma)$, whilst Case (b) concerns *unconditional* (Lipschitz) stability, i.e., $L = 0$ and $\eta(t) = \mathcal{O}(t)$. To show the latter, it remains to verify that A is injective in Case (b). Suppose it is not, and let $0 \neq z \in X$ with $\|Az\|_V = 0$. From $X \hookrightarrow C([0, T]; L_2(\Omega))$ ([31], Ch. 1, Thm. 3.1), there exists an open interval $J \subset (0, T)$ such that for any $t \in J$, $\|z(t, \cdot)\|_{L_2(\Omega)} \neq 0$. On the other hand, the above estimate shows that $\|z\|_{L_2((T_1, T); H_0^1(\Omega)) \cap H^1((T_1, T); H^{-1}(\Omega))} = 0$ for any $T_1 \in (0, T)$. From $L_2((T_1, T); H_0^1(\Omega)) \cap H^1((T_1, T); H^{-1}(\Omega)) \hookrightarrow C([T_1, T]; L_2(\Omega))$ we arrive at a contradiction.

Example 3.3 (Data-assimilation for the wave equation ([18], Rem. A.5)). Let $\Omega \subset \mathbb{R}^d$ be a domain with a smooth boundary, and, for some $T > 0$, let $I := (0, T)$. Let $\omega \subset \bar{\Omega}$, and assume that $I \times \omega$ satisfies the *Geometric Control Condition* [3, 4]. Roughly speaking, it means that all geometric optic rays in $I \times \Omega$, taking into account their reflections at the lateral boundary, intersect the set $I \times \omega$.

Given $(f, g, h) \in V := H^{-1}(I \times \Omega) \times L_2(I \times \partial\Omega) \times L_2(I \times \omega)$, the data assimilation problem reads as finding u that satisfies

$$\begin{aligned} (\square u)(v) &:= \int_I \int_{\Omega} -\partial_t u \partial_t v + \nabla_x u \cdot \nabla_x v \, dx \, dt = f(v) \quad (v \in H_0^1(I \times \Omega)), \\ \gamma_{I \times \partial\Omega} u &:= u|_{I \times \partial\Omega} = g, \quad \text{and} \quad \Gamma_{I \times \omega} u := u|_{I \times \omega} = h. \end{aligned}$$

With $A := (\square, \gamma_{I \times \partial\Omega}, \Gamma_{I \times \omega})$ and $X = \{z \in L_2(I \times \Omega) : Az \in V\}$ equipped with the graph norm, or its completion when $A: L_2(I \times \Omega) \supset D_A \rightarrow V$ is not closed, the following *unconditional (Lipschitz) stability* is valid:

$$\|z\|_{L_{\infty}(I; L_2(\Omega))} + \|\partial_t z\|_{L_2(I; H^{-1}(\Omega))} \lesssim \|Az\|_V \quad (z \in X). \quad (3.3)$$

Remark 3.4 (The condition of $\partial\Omega$ being smooth). For finite element computations, for $d > 1$ a problem with the setting of Example 3.3 is the condition of $\partial\Omega$ being smooth. It is used in the derivation of both the *Distributed Observability Estimate* $\|z(0, \cdot)\|_{L_2(\Omega)} + \|\partial_t z(0, \cdot)\|_{L_2(\Omega)} \lesssim \|\Gamma_{I \times \omega} z\|_{L_2(I \times \omega)}$ for functions z that satisfy $\square z = 0$ and $\gamma_{I \times \partial\Omega} z = 0$, see *e.g.*, [30], and the *Energy Estimate*

$$\|z\|_{L_{\infty}(I; L_2(\Omega))} + \|\partial_t z\|_{L_2(I; H^{-1}(\Omega))} \lesssim \|z(0, \cdot)\|_{L_2(\Omega)} + \|\partial_t z(0, \cdot)\|_{H^{-1}(\Omega)} + \|\gamma_{I \times \partial\Omega} z\|_{L_2(I \times \partial\Omega)} + \|\square z\|_{H^{-1}(I \times \Omega)},$$

see Proposition A.1 from [18], which builds on Theorem 2.1 and 2.3 of [29]. From both these estimates, one easily derives (3.3) (cf. proof of Thm. 2.2 of [16] (see however [18], Rem. 2.6)). In [16] it is claimed that under stronger conditions on ω , the Distributed Observability Estimate can also be valid for polytopal Ω . See also [20] for the related boundary controllability of the wave equation on domains with corners. Assuming that on such domains also an energy estimate is valid, possibly with weaker norms on the left-hand side, in the case that these norms do control $\|z\|_{L_2(I \times \Omega)}$, proof of Theorem 2.2 of [16] will give an unconditional stability estimate (3.3) with those norms on the left-hand side.

Notice that only in Example 3.1(ii), $\text{ran}(A, L)$ is closed.

Other examples of conditional stability include, *e.g.*, *data-assimilation for the Poisson equation* [14], *the backward heat equation* [26], and *the heat and wave equations with lateral Cauchy data* ([28] and Thm. 3.4.1 of [27]).

4. LEAST SQUARES APPROXIMATION

We adhere to the assumptions in Section 2. For a given suitable finite dimensional space X^{δ} , we propose, as an approximate “solution” to (2.1), the unique minimizer $u_{\varepsilon}^{\delta} \in X^{\delta}$ of the *regularized* least-squares functional

$$z \mapsto \sqrt{\|Az - f\|_V^2 + \varepsilon^2 \|Lz\|_H^2}.$$

As we have seen in applications, however, the space V , or one or more components of V when it is a product space, are equipped with a norm that cannot be evaluated. For example it can be either a dual norm, *i.e.*, a norm of type $\sup_{0 \neq v \in Y} \frac{|\cdot(v)|}{\|v\|_Y}$, or a fractional Sobolev norm, the latter which typically arises with the enforcement of boundary conditions.

For the dual norm case we will see that under an inf-sup or LBB condition, for minimizing the least-squares functional over X^{δ} the supremum over Y can be replaced by a supremum over a suitable finite dimensional space, which makes it computable assuming $\|\cdot\|_Y$ can be evaluated.

Dealing with reflexive spaces, any norm can be viewed as a dual norm. Applying this to a fractional Sobolev norm, the space Y is a fractional Sobolev space with smoothness index of opposite sign. At a first glance this

might not seem to be helpful, but we saw that under an LBB condition the norm on the latter space needs to be evaluated for arguments from a finite dimensional subspace only. Moreover, as we will see, it suffices to compute a norm on this finite dimensional subspace that is only (uniformly) *equivalent* to the norm on Y .

In view of above comments, we consider our least squares problem in the following setting that covers all envisioned scenarios. Specifically, V may be the product of a dual space and a space whose norm is easy to evaluate. So, for some Hilbert spaces Y and W , let $V = Y' \times W$, $A = (B, C)$, so that (2.3) takes the form

$$\|\cdot\|_\varepsilon := \sqrt{\|B \cdot\|_{Y'}^2 + \|C \cdot\|_W^2 + \varepsilon^2 \|L \cdot\|_H^2}, \quad (4.1)$$

and let $f = (g, h)$, so that the aforementioned least-squares functional reads as $z \mapsto \sqrt{\|Bz - g\|_{Y'}^2 + \|Cz - h\|_W^2 + \varepsilon^2 \|Lz\|_H^2}$. We assume that the $\|\cdot\|_W$ -norm can be evaluated, ignoring possible quadrature issues. The analysis of cases where the triple (V, A, f) has none or multiple components of type either (Y', B, g) or (W, C, h) causes no additional problems.

To avoid the exact evaluation of the $\|\cdot\|_{Y'}$ -norm, given a family $(X^\delta)_{\delta \in \Delta}$ of finite dimensional subspaces of X , let $(Y^\delta)_{\delta \in \Delta}$ be a family of finite dimensional subspaces of Y such that

$$\varrho := \inf_{\delta \in \Delta} \inf_{\{z \in X^\delta : Bz \neq 0\}} \sup_{0 \neq v \in Y^\delta} \frac{|(Bz)(v)|}{\|Bz\|_{Y'} \|v\|_Y} > 0, \quad (4.2)$$

and let $\langle \cdot, \cdot \rangle_{Y^\delta}$ be an inner product on Y^δ with associated norm $\|\cdot\|_{Y^\delta}$ that satisfies

$$\|\cdot\|_{Y^\delta} \approx \|\cdot\|_Y \quad \text{on } Y^\delta \quad (4.3)$$

(uniformly in $\delta \in \Delta$). A choice $\|\cdot\|_{Y^\delta} \neq \|\cdot\|_Y$ is useful when also the $\|\cdot\|_Y$ -norm cannot be (easily) evaluated. As we will see in Section 5, another reason for taking a suitable $\|\cdot\|_{Y^\delta} \neq \|\cdot\|_Y$ is when the stiffness matrix corresponding to $\langle \cdot, \cdot \rangle_{Y^\delta}$ can be efficiently inverted in which case the approximation u_ε^δ defined in the next theorem can be found as the solution of a symmetric positive definite system instead of a saddle-point system.

Theorem 4.1. *Let (4.2) and (4.3) be valid. For u_ε^δ being the unique solution of*

$$u_\varepsilon^\delta := \operatorname{argmin}_{z \in X^\delta} \left\{ \sup_{0 \neq v \in Y^\delta} \frac{|(Bz - g)(v)|^2}{\|v\|_{Y^\delta}^2} + \|Cz - h\|_W^2 + \varepsilon^2 \|Lz\|_H^2 \right\}, \quad (4.4)$$

it holds that

$$\|u - u_\varepsilon^\delta\|_\varepsilon \lesssim \|Bu - g\|_{Y'} + \|Cu - h\|_W + \min_{z \in X^\delta} \|u - z\|_\varepsilon + \varepsilon \|Lu\|_H, \quad (4.5)$$

i.e., equation (2.4) is valid.

Proof. Initially we assume that $\|\cdot\|_{Y^\delta} = \|\cdot\|_Y$ on Y^δ , and postpone the discussion about the case where we have only a uniform equivalence (4.3) until the end of the proof.

The basic idea behind the proof is to use a suitable intermediary $u_\varepsilon \in X$ to estimate then $\|u - u_\varepsilon\|_\varepsilon$ and $\|u_\varepsilon - u_\varepsilon^\delta\|_\varepsilon$. Ideally, u_ε should be the minimizer over $z \in X$ of $\|Bz - g\|_{Y'}^2 + \|Cz - h\|_W^2 + \varepsilon^2 \|Lz\|_H^2$. Since we are not insisting of $\operatorname{ran}(B, C, L)$ being closed (cf. Rem. 2.6), this minimizer would not necessarily exist. Therefore, more work is required to identify such uniform intermediaries *via* a perturbation that ensures closedness. To that end, for $\zeta \geq 0$ we equip X with the additional norm

$$\|\cdot\|_{\varepsilon, \zeta} := \sqrt{\|B \cdot\|_{Y'}^2 + \|C \cdot\|_W^2 + \varepsilon^2 \|L \cdot\|_H^2 + \zeta^2 \|\cdot\|_X^2}.$$

Since $\operatorname{ran}(B, C, L, \operatorname{Id})$ is closed, for $\zeta > 0$ the minimizer $u_{\varepsilon, \zeta}$ over $z \in X$ of

$$G_{\varepsilon, \zeta}(z) := \|Bz - g\|_{Y'}^2 + \|Cz - h\|_W^2 + \varepsilon^2 \|Lz\|_H^2 + \zeta^2 \|z\|_X^2$$

does exist uniquely as the solution of the Euler–Lagrange equations

$$\langle B\tilde{z}, Bu_{\varepsilon,\zeta} - g \rangle_{Y'} + \langle C\tilde{z}, Cu_{\varepsilon,\zeta} - h \rangle_W + \varepsilon^2 \langle L\tilde{z}, Lu_{\varepsilon,\zeta} \rangle_H + \zeta^2 \langle \tilde{z}, u_{\varepsilon,\zeta} \rangle_X = 0 \quad (\tilde{z} \in X).$$

To deal with the inner product in Y' we employ the Riesz *lifter* (= inverse Riesz *map*) $R_Y: Y' \rightarrow Y$, defined for $f \in Y'$ by

$$\langle \tilde{z}, R_Y f \rangle_Y = f(\tilde{z}), \quad \forall \tilde{z} \in Y,$$

and introduce the lifted residual

$$v_{\varepsilon,\zeta} := R_Y(g - Bu_{\varepsilon,\zeta}). \quad (4.6)$$

We can then write

$$\langle B\tilde{z}, Bu_{\varepsilon,\zeta} - g \rangle_{Y'} = -\langle v_{\varepsilon,\zeta}, R_Y B\tilde{z} \rangle_Y = -(B\tilde{z})(v_{\varepsilon,\zeta}),$$

from which one infers that $(u_{\varepsilon,\zeta}, v_{\varepsilon,\zeta}) \in X \times Y$ is the (unique) solution of

$$\underbrace{(B\tilde{z})(v_{\varepsilon,\zeta}) + (Bu_{\varepsilon,\zeta})(\tilde{v}) + \langle \tilde{v}, v_{\varepsilon,\zeta} \rangle_Y - \langle C\tilde{z}, Cu_{\varepsilon,\zeta} \rangle_W - \varepsilon^2 \langle L\tilde{z}, Lu_{\varepsilon,\zeta} \rangle_H - \zeta^2 \langle \tilde{z}, u_{\varepsilon,\zeta} \rangle_X}_{d_{\varepsilon,\zeta}((u_{\varepsilon,\zeta}, v_{\varepsilon,\zeta}), (\tilde{z}, \tilde{v}))} := \ell(\tilde{z}, \tilde{v}) := g(\tilde{v}) - \langle C\tilde{z}, h \rangle_W \quad ((\tilde{z}, \tilde{v}) \in X \times Y).$$

The symmetric bilinear form $d_{\varepsilon,\zeta}$ on $((X, \|\cdot\|_{\varepsilon,\zeta}) \times Y) \times ((X, \|\cdot\|_{\varepsilon,\zeta}) \times Y)$ is bounded (uniformly in $\varepsilon > 0$ and $\zeta \geq 0$).

Similarly as in the continuous case, for $\delta \in \Delta$ the minimizer $u_{\varepsilon,\zeta}^\delta$ over X^δ of

$$G_{\varepsilon,\zeta}^\delta(z) := \sup_{0 \neq v \in Y^\delta} \frac{|(Bz - g)(v)|^2}{\|v\|_Y^2} + \|Cz - h\|_W^2 + \varepsilon^2 \|Lz\|_H^2 + \zeta^2 \|z\|_X^2$$

exists uniquely. Thanks to $\dim X^\delta < \infty$, this holds even true for $\zeta = 0$ (where $u_{\varepsilon,0}^\delta = u_\varepsilon^\delta$ is the solution of (4.4))

Defining in analogy to (4.6),

$$\langle \tilde{z}, v_{\varepsilon,\zeta}^\delta \rangle_Y = (g - Bu_{\varepsilon,\zeta}^\delta)(\tilde{z}), \quad \tilde{z} \in Y^\delta,$$

the pair $(u_{\varepsilon,\zeta}^\delta, v_{\varepsilon,\zeta}^\delta) \in X^\delta \times Y^\delta$ solves the Galerkin system

$$d_{\varepsilon,\zeta}((u_{\varepsilon,\zeta}^\delta, v_{\varepsilon,\zeta}^\delta), (\tilde{z}, \tilde{v})) = \ell(\tilde{z}, \tilde{v}) \quad ((\tilde{z}, \tilde{v}) \in X^\delta \times Y^\delta). \quad (4.7)$$

We will demonstrate later below that

$$\inf_{\delta \in \Delta, \varepsilon > 0, \zeta \geq 0} \inf_{0 \neq (z,v) \in X^\delta \times Y^\delta} \sup_{0 \neq (\tilde{z}, \tilde{v}) \in X^\delta \times Y^\delta} \frac{d_{\varepsilon,\zeta}((z,v), (\tilde{z}, \tilde{v}))}{\sqrt{\|z\|_{\varepsilon,\zeta}^2 + \|v\|_Y^2} \sqrt{\|\tilde{z}\|_{\varepsilon,\zeta}^2 + \|\tilde{v}\|_Y^2}} > 0. \quad (4.8)$$

Assuming the validity of (4.8) for the moment, the symmetry and boundedness of $d_{\varepsilon,\zeta}$, then shows that for $\zeta > 0$,

$$\|v_{\varepsilon,\zeta} - v_{\varepsilon,\zeta}^\delta\|_Y + \|u_{\varepsilon,\zeta} - u_{\varepsilon,\zeta}^\delta\|_{\varepsilon,\zeta} \lesssim \min_{(\tilde{z}, \tilde{v}) \in X^\delta \times Y^\delta} \left\{ \|v_{\varepsilon,\zeta} - \tilde{v}\|_Y + \|u_{\varepsilon,\zeta} - \tilde{z}\|_{\varepsilon,\zeta} \right\}. \quad (4.9)$$

uniformly in ε (see *e.g.*, [39], Thm. 3.1).

Now, for any $u \in X$ we simply estimate $\|u - u_{\varepsilon,\zeta}^\delta\|_{\varepsilon,\zeta} \leq \|u - u_{\varepsilon,\zeta}\|_{\varepsilon,\zeta} + \|u_{\varepsilon,\zeta} - u_{\varepsilon,\zeta}^\delta\|_{\varepsilon,\zeta}$ and deduce first from (4.9) that we have for the second term

$$\begin{aligned} \|u_{\varepsilon,\zeta} - u_{\varepsilon,\zeta}^\delta\|_{\varepsilon,\zeta} &\lesssim \|v_{\varepsilon,\zeta}\|_Y + \min_{\tilde{z} \in X^\delta} \|u_{\varepsilon,\zeta} - \tilde{z}\|_{\varepsilon,\zeta} \\ &\leq \|g - Bu_{\varepsilon,\zeta}\|_{Y'} + \|u - u_{\varepsilon,\zeta}\|_{\varepsilon,\zeta} + \min_{\tilde{z} \in X^\delta} \|u - \tilde{z}\|_{\varepsilon,\zeta} \end{aligned}$$

$$\leq \|g - Bu\|_{Y'} + 2\|u - u_{\varepsilon,\zeta}\|_{\varepsilon,\zeta} + \min_{\tilde{z} \in X^\delta} \|u - \tilde{z}\|_{\varepsilon,\zeta}. \quad (4.10)$$

Regarding the first term, an application of the triangle inequality and optimality of $u_{\varepsilon,\zeta}$ give

$$\begin{aligned} \|u - u_{\varepsilon,\zeta}\|_{\varepsilon,\zeta} &\leq \sqrt{G_{\varepsilon,\zeta}(u)} + \sqrt{G_{\varepsilon,\zeta}(u_{\varepsilon,\zeta})} \\ &\leq 2\sqrt{G_{\varepsilon,\zeta}(u)} = 2\sqrt{\|Bu - g\|_{Y'}^2 + \|Cu - h\|_W^2 + \varepsilon^2\|Lu\|_H^2 + \zeta^2\|u\|_X^2}, \end{aligned}$$

which together with (4.10) yields

$$\|u - u_{\varepsilon,\zeta}^\delta\|_{\varepsilon,\zeta} \lesssim \|Bu - g\|_{Y'} + \|Cu - h\|_W + \min_{z \in X^\delta} \|u - z\|_{\varepsilon,\zeta} + \varepsilon\|Lu\|_H + \zeta\|u\|_X \quad (4.11)$$

(uniformly in $\delta \in \Delta$, $\varepsilon, \zeta > 0$).

Now fixing $\delta \in \Delta$ and $\varepsilon > 0$, the norms $\|\cdot\|_{\varepsilon,\zeta}$ on X^δ are equivalent uniformly in $\zeta \in [0, 1]$, and $\lim_{\zeta \downarrow 0} \min_{z \in X^\delta} \|u - z\|_{\varepsilon,\zeta} = \min_{z \in X^\delta} \|u - z\|_\varepsilon$. Introducing the operator $D_{\varepsilon,\zeta}^\delta \in \mathcal{L}(X^\delta \times Y^\delta, (X^\delta \times Y^\delta)')$, defined by

$$(D_{\varepsilon,\zeta}^\delta(u_{\varepsilon,\zeta}^\delta, v_{\varepsilon,\zeta}^\delta))(\tilde{z}, \tilde{v}) := d_{\varepsilon,\zeta}((u_{\varepsilon,\zeta}^\delta, v_{\varepsilon,\zeta}^\delta), (\tilde{z}, \tilde{v})),$$

and writing $(D_{\varepsilon,\zeta}^\delta)^{-1} = (D_{\varepsilon,0}^\delta)^{-1} + (D_{\varepsilon,\zeta}^\delta)^{-1}(D_{\varepsilon,0}^\delta - D_{\varepsilon,\zeta}^\delta)(D_{\varepsilon,0}^\delta)^{-1}$, the uniform stability (4.8) and $\lim_{\zeta \downarrow 0} D_{\varepsilon,\zeta}^\delta = D_{\varepsilon,0}^\delta$ show that $\lim_{\zeta \downarrow 0} u_{\varepsilon,\zeta}^\delta = u_\varepsilon^\delta$. We conclude that by taking the limit for $\zeta \downarrow 0$ in (4.11) the proof of (4.5) for the case that $\|\cdot\|_{Y^\delta} = \|\cdot\|_Y$ is completed as soon as we have confirmed the validity of (4.8).

The latter statement (4.8) left to be shown is equivalent to uniform stability of the variational problem (4.7), i.e.,

$$\|u_{\varepsilon,\zeta}^\delta\|_{\varepsilon,\zeta} + \|v_{\varepsilon,\zeta}^\delta\|_Y \lesssim \|\ell\|_{((X^\delta, \|\cdot\|_{\varepsilon,\zeta}) \times Y^\delta)'} \quad (4.12)$$

(uniformly in $\delta, \varepsilon > 0$, and $\zeta \geq 0$) which requires utilizing (4.2). To that end, we introduce as additional variables $\theta = -Cu_{\varepsilon,\zeta}^\delta$, $\mu = -\varepsilon Lu_{\varepsilon,\zeta}^\delta$, and $\chi = -\zeta u_{\varepsilon,\zeta}^\delta$ and the bilinear form

$$e\left(z, \left(\tilde{v}, \tilde{\theta}, \tilde{\mu}, \tilde{\chi}\right)\right) := (Bz)(\tilde{v}) + \left\langle \tilde{\theta}, Cz \right\rangle_W + \varepsilon \langle \tilde{\mu}, Lz \rangle_H + \zeta \langle \tilde{\chi}, z \rangle_X \quad (4.13)$$

over $X^\delta \times (Y^\delta \times W \times H \times X)$. Then, (4.7) is equivalent to finding $(u_{\varepsilon,\zeta}^\delta, v_{\varepsilon,\zeta}^\delta, \theta, \mu, \chi) \in X^\delta \times Y^\delta \times W \times H \times X$ such that

$$\left\langle \left(\tilde{v}, \tilde{\theta}, \tilde{\mu}, \tilde{\chi}\right), (v_{\varepsilon,\zeta}^\delta, \theta, \mu, \chi) \right\rangle_{Y \times W \times H \times X} + e\left(u_{\varepsilon,\zeta}^\delta, \left(\tilde{v}, \tilde{\theta}, \tilde{\mu}, \tilde{\chi}\right)\right) + e\left(\tilde{z}, (v_{\varepsilon,\zeta}^\delta, \theta, \mu, \chi)\right) = \ell(\tilde{z}, \tilde{v}) \quad (4.14)$$

for all $(\tilde{z}, \tilde{v}, \tilde{\theta}, \tilde{\mu}, \tilde{\chi}) \in X^\delta \times Y^\delta \times W \times H \times X$.

Recall from (4.2) that for any $\sigma \in (0, 1)$, given $z \in X^\delta$ we can find $\tilde{v} \in Y^\delta$ with $\|\tilde{v}\|_Y = \|Bz\|_{Y'}$ and $(Bz)(\tilde{v}) \geq \sigma\rho\|Bz\|_{Y'}^2$. Then take $(\tilde{\theta}, \tilde{\mu}, \tilde{\chi}) = (Cz, \varepsilon Lz, \zeta z)$, to conclude that

$$\begin{aligned} e\left(z, \left(\tilde{v}, \tilde{\theta}, \tilde{\mu}, \tilde{\chi}\right)\right) &\geq \sigma\rho\|Bz\|_{Y'}^2 + \|Cz\|_W^2 + \varepsilon^2\|Lz\|_H^2 + \zeta^2\|z\|_X^2 \\ &\geq \sigma\rho\|z\|_{\varepsilon,\zeta} \sqrt{\|\tilde{v}\|_Y^2 + \|\tilde{\theta}\|_W^2 + \|\tilde{\mu}\|_H^2 + \|\tilde{\chi}\|_X^2}. \end{aligned}$$

From this LBB stability we conclude that (4.14) is indeed uniformly stable, i.e., that $\|u_{\varepsilon,\zeta}^\delta\|_{\varepsilon,\zeta} + \|v_{\varepsilon,\zeta}^\delta\|_Y + \|\theta\|_W + \|\mu\|_H + \|\chi\|_X \lesssim \|\ell\|_{(X^\delta, \|\cdot\|_{\varepsilon,\zeta}) \times Y^\delta}'$, which implies (4.12).

Finally, we discuss the case that $\|\cdot\|_{Y^\delta} \neq \|\cdot\|_Y$. We write $Y = Y^\delta \oplus (Y^\delta)^\perp \simeq Y^\delta \times (Y^\delta)^\perp$, and replace for the Y^δ -component the inner product $\langle \cdot, \cdot \rangle_Y$ by $\langle \cdot, \cdot \rangle_{Y^\delta}$. Then we are back in the case that $\|\cdot\|_{Y^\delta} = \|\cdot\|_Y$ on Y^δ , and the proof so far shows that u_ε^δ defined in (4.4) satisfies (4.5) with the Y' -norm at its right-hand side and in the definition of $\|\cdot\|_\varepsilon$ at its left- and right-hand side being defined in terms of the so modified Y -norm. Since by our assumption (4.3) this modified Y -norm, and so its resulting dual norm, is equivalent to the original norm, uniformly in $\delta \in \Delta$, the proof is completed. $\square$

In view of (4.7) and the last paragraph in the above proof, note that the least squares approximation u_ε^δ defined in (4.4) can be computed as the first component of the pair $(u_\varepsilon^\delta, v_\varepsilon^\delta) \in X^\delta \times Y^\delta$ that solves the *mixed system*

$$\begin{cases} (B\tilde{z})(v_\varepsilon^\delta) + (Bu_\varepsilon^\delta)(\tilde{v}) + \langle \tilde{v}, v_\varepsilon^\delta \rangle_{Y^\delta} \\ - \langle C\tilde{z}, Cu_\varepsilon^\delta \rangle_W - \varepsilon^2 \langle L\tilde{z}, Lu_\varepsilon^\delta \rangle_H = g(\tilde{v}) - \langle C\tilde{z}, h \rangle_W \end{cases} \quad ((\tilde{z}, \tilde{v}) \in X^\delta \times Y^\delta). \quad (4.15)$$

For a suitable choice of $\langle \cdot, \cdot \rangle_{Y^\delta}$, in Section 5 we will reduce this system to a symmetric positive definite system for the primal variable u_ε^δ .

Remark 4.2 (Relation to existing work). Suppose that the norm $\|\cdot\|_X$ is computable (which is not the case in Example 3.2) and assume for simplicity that $\|\cdot\|_{Y^\delta} = \|\cdot\|_Y$. Then by replacing $\varepsilon^2 \|Lz\|_H^2$ by the stronger regularizing term $\varepsilon^2 \|z\|_X^2$, instead of minimizing (4.4) as in Theorem 4.1 one may minimize the least-squares functional

$$\hat{u}_\varepsilon^\delta := \operatorname{argmin}_{z \in X^\delta} \left\{ \sup_{0 \neq v \in Y^\delta} \frac{|(Bz - g)(v)|^2}{\|v\|_Y^2} + \|Cz - h\|_W^2 + \varepsilon^2 \|z\|_X^2 \right\}.$$

Upon replacing $\varepsilon^2 \langle L \cdot, L \cdot \rangle_H$ by $\varepsilon^2 \langle \cdot, \cdot \rangle_X$ in (4.7), and the test function $\tilde{z}$ by $-\tilde{z}$, this amounts to solving for $(\hat{u}_\varepsilon^\delta, \hat{v}_\varepsilon^\delta) \in X^\delta \times Y^\delta$ from

$$-(B\tilde{z})(\hat{v}_\varepsilon^\delta) + (B\hat{u}_\varepsilon^\delta)(\tilde{v}) + \langle \tilde{v}, \hat{v}_\varepsilon^\delta \rangle_Y + \langle C\tilde{z}, C\hat{u}_\varepsilon^\delta \rangle_W + \varepsilon^2 \langle \tilde{z}, \hat{u}_\varepsilon^\delta \rangle_X = g(\tilde{v}) + \langle C\tilde{z}, h \rangle_W \quad ((\tilde{z}, \tilde{v}) \in X^\delta \times Y^\delta). \quad (4.16)$$

The bilinear form on $X \times Y$ on the left-hand side of (4.16) is bounded, uniformly in $\delta \in \Delta$. Moreover, already *without* assuming (4.2), it is also *coercive*, although with the unfavorable coercivity constant $\asymp \varepsilon^2$. With $(\hat{u}_\varepsilon, \hat{v}_\varepsilon) \in X \times Y$ denoting the solution of (4.16) under testing with all $(\tilde{z}, \tilde{v}) \in X \times Y$, one arrives at the estimate

$$\|\hat{v}_\varepsilon - \hat{v}_\varepsilon^\delta\|_Y + \|\hat{u}_\varepsilon - \hat{u}_\varepsilon^\delta\|_X \lesssim \varepsilon^{-2} \inf_{(\tilde{z}, \tilde{v}) \in X^\delta \times Y^\delta} \{\|\hat{v}_\varepsilon - \tilde{v}\|_Y + \|\hat{u}_\varepsilon - \tilde{z}\|_X\}. \quad (4.17)$$

This approach to solve (4.16) without ensuring the LBB stability (4.2) has been suggested in [5, 9]. In Theorem 2.7 from [9] it was shown that $\lim_{\varepsilon \downarrow 0} (\hat{u}_\varepsilon, \hat{v}_\varepsilon) = (u, 0)$ in $X \times Y$, where u is the minimal norm solution of $(Bu, Cu) = (g, h)$, under the assumption that a solution exists. To conclude from (4.17) convergence of numerical approximations u_ε^δ towards u requires making assumptions on the behavior of higher order norms of $(\hat{u}_\varepsilon, \hat{v}_\varepsilon)$ as functions of $\varepsilon \downarrow 0$ (cf. [5], Thm. 4.1). On the other hand, this approach does not assume any conditional stability, and it concerns convergence in norm instead of convergence in the functional j associated to a conditional stability assumption.

Notice, however, that if conditional stability is available and the weaker notion of convergence in the associated functional is relevant, then the difference between (4.17) and the quasi-optimal bound (4.9) shows that is indeed very rewarding to select Y^δ dependent on X^δ such that (4.2) is valid.

Uniform discrete inf-sup stability (4.2) plays a pivotal role in our approach. We recall that it is equivalent to the existence of uniformly bounded “Fortin operators”. Precisely, the following statement holds.

Theorem 4.3 ([38], Prop. 5.1). *For general $B \in \mathcal{L}(X, Y')$, and closed subspaces X^δ and Y^δ of Hilbert spaces X and Y with $BX^\delta \neq \{0\}$ and $Y^\delta \neq \{0\}$, let*

$$Q^\delta \in \mathcal{L}(Y, Y^\delta) \quad \text{with} \quad (BX^\delta)((\operatorname{Id} - Q^\delta)Y) = 0. \quad (4.18)$$

Then $\xi^\delta := \inf_{\{z \in X^\delta : Bz \neq 0\}} \sup_{0 \neq v \in Y^\delta} \frac{|(Bz)(v)|}{\|Bz\|_{Y'} \|v\|_Y} \geq \|Q^\delta\|_{\mathcal{L}(Y, Y)}^{-1}$.

Conversely, when $\xi^\delta > 0$, and $\operatorname{ran} B$ is closed or $\dim X^\delta < \infty$, then there exists a Q^δ as in (4.18), being even a projector, with $\|Q^\delta\|_{\mathcal{L}(Y, Y)} = 1/\xi^\delta$.

Using this theorem, in Section 7 we will verify condition (4.2) for the examples from Section 3.

Finally, we note that the condition (4.2) may be slightly relaxed:

Remark 4.4 (Slight relaxation of inf-sup condition (4.2)). By an application of Young's inequality, it is not difficult to show that the reasoning that provided the LBB stability of the bilinear form e from (4.13), and consequently the proof of Theorem 4.1, still applies when, for some constant $\mu < 2\sqrt{\varrho}$, (4.2) is relaxed to $\sup_{0 \neq v \in Y^\delta} \frac{|(Bz)(v)|}{\|v\|_Y} \geq \varrho \|Bz\|_{Y'} - \mu \sqrt{\|Cz\|_W^2 + \varepsilon^2 \|Lz\|_H^2}$.

5. CHOICE OF $\langle \cdot, \cdot \rangle_{Y^\delta}$, AND THE REFORMULATION OF (4.15) AS A SYMMETRIC POSITIVE DEFINITE SYSTEM

Let $G_Y^\delta = G_Y^{\delta'} \in \mathcal{L}is(Y^{\delta'}, Y^\delta)$ be such that

$$\|G_Y^\delta f\|_Y^2 \approx f(G_Y^\delta f) \quad (f \in Y^{\delta'}, \delta \in \Delta). \quad (5.1)$$

Assuming its application can be computed “efficiently”, such an operator G_Y^δ is often called a (uniform) *preconditioner*. See Remark 5.2 below for scenarios where such preconditioners are available. This G_Y^δ can be employed to *define*

$$\langle v, \tilde{v} \rangle_{Y^\delta} := \left((G_Y^\delta)^{-1} \tilde{v} \right)(v) \quad (v, \tilde{v} \in Y^\delta), \quad (5.2)$$

which, in view of (5.1), yields $\|\cdot\|_{Y^\delta} \approx \|\cdot\|_Y$ on Y^δ , i.e., (4.3). The definition of $\langle \cdot, \cdot \rangle_{Y^\delta}$ shows that G_Y^δ is the Riesz lifter associated to the Hilbert space $(Y^\delta, \langle \cdot, \cdot \rangle_{Y^\delta})$. Since a Riesz lifter is an isometry, it follows that

$$\sup_{0 \neq v \in Y^\delta} \frac{|f(v)|^2}{\|v\|_{Y^\delta}^2} = \|G_Y^\delta f\|_{Y^\delta}^2 = f(G_Y^\delta f) \quad (f \in Y^{\delta'}).$$

We conclude that for $\langle \cdot, \cdot \rangle_{Y^\delta}$ as in (5.2), u_ε^δ defined in (4.4) satisfies

$$u_\varepsilon^\delta = \operatorname{argmin}_{z \in X^\delta} \left\{ (Bz - g)(G_Y^\delta(Bz - g)) + \|Cz - h\|_W^2 + \varepsilon^2 \|Lz\|_H^2 \right\},$$

meaning that $u_\varepsilon^\delta \in X^\delta$ solves the Euler–Lagrange equations

$$\left(Bu_\varepsilon^\delta \right) (G_Y^\delta B\tilde{z}) + \langle C\tilde{z}, Cu_\varepsilon^\delta \rangle_W + \varepsilon^2 \langle L\tilde{z}, Lu_\varepsilon^\delta \rangle_H = g(G_Y^\delta B\tilde{z}) + \langle C\tilde{z}, h \rangle_W \quad (\tilde{z} \in X^\delta). \quad (5.3)$$

When G_Y^δ can be applied in linear complexity, the iterative solution of this symmetric positive definite system can be expected to be more efficient than the iterative solution of the mixed system (4.15) with $\langle \cdot, \cdot \rangle_{Y^\delta} = \langle \cdot, \cdot \rangle_Y$. So even in the case that $\langle \cdot, \cdot \rangle_Y$ can be efficiently evaluated, it can be helpful to replace $\|\cdot\|_Y$ by the above $\|\cdot\|_{Y^\delta}$ in the definition (4.4) of the regularized least squares approximation u_ε^δ .

Remark 5.1. We arrived at (5.3) without introducing the Riesz lift of $g - Bu_\varepsilon^\delta \in Y^{\delta'}$ as a separate variable. Alternatively, eliminating this variable v_ε^δ from the mixed system (4.15) using the definition (5.2) of $\langle \cdot, \cdot \rangle_{Y^\delta}$ also results in (5.3).

Remark 5.2 (Some scenarios where uniform preconditioners are available). Continuing the discussion preceding Theorem 4.1, uniform preconditioners G_Y^δ of linear complexity are for example available when Y is a Sobolev space of either positive or negative possibly non-integer smoothness index, and the Y^δ are common finite element spaces. Wavelet preconditioners can be applied, but also multi-level preconditioners that make solely use of nodal bases as the BPX preconditioner [10] for positive smoothness indices, and the preconditioner from [22] for negative smoothness indices. Another option for negative smoothness indices is given by the preconditioner from [37] based on the “operator preconditioning” framework where the “opposite order” operator of linear complexity uses a multi-level hierarchy.

Remark 5.3 (Conditioning of (5.3)). In the case that $\text{ran}(B, C, L)$ is closed, then one has $\varepsilon^2 \lesssim \frac{(B \cdot)(G_Y^\delta B \cdot) + (C \cdot, C \cdot)_W + \varepsilon^2 (L \cdot, L \cdot)_H}{\|\cdot\|_X^2} \lesssim 1$ on X^δ (cf. Rem. 2.6). So for $G_X^\delta = G_X^{\delta'} \in \mathcal{L}(X^{\delta'}, X^\delta)$ with $\|G_X^\delta \cdot\|_X^2 \approx \cdot(G_X^\delta \cdot)$ on $X^{\delta'}$, the condition number of the system matrix corresponding to (5.3) preconditioned by G_X^δ is $\lesssim \varepsilon^{-2}$, directly tying solver efficiency to regularization.

6. A POSTERIORI RESIDUAL ESTIMATION

In view of Theorem 2.2, and the inequality $\|A(u - u_\varepsilon^\delta)\|_V \leq \tau + \|f - Au_\varepsilon^\delta\|_V$, it is desirable to have an *a posteriori* estimate for $\|f - Au_\varepsilon^\delta\|_V$. Indeed, for suitable ε it will give rise to a computable upper bound for $j(u - u_\varepsilon^\delta)$.

Considering $V = Y' \times W$, $A = (B, C)$, and $f = (g, h)$, the issue is to approximate $\|g - Bu_\varepsilon^\delta\|_{Y'}$. This is where Fortin operators again come into play. Recall from Theorem 4.3 that (4.2) guarantees the existence of a family of uniformly bounded Fortin operators $(Q^\delta)_{\delta \in \Delta}$. We can then write $g - Bu_\varepsilon^\delta = Q^{\delta'}(g - Bu_\varepsilon^\delta) + (\text{Id} - Q^\delta)'g$, and obtain

$$\begin{aligned} \|g - Bu_\varepsilon^\delta\|_{Y'} &\leq \|Q^\delta\|_{\mathcal{L}(Y, Y)} \sup_{0 \neq v \in Y^\delta} \frac{|(g - Bu_\varepsilon^\delta)(v)|}{\|v\|_Y} + \underbrace{\|(\text{Id} - Q^\delta)'g\|_{Y'}}_{\text{osc}^\delta(g) :=} \\ &\stackrel{(4.3)}{\approx} \sup_{0 \neq v \in Y^\delta} \frac{|(g - Bu_\varepsilon^\delta)(v)|}{\|v\|_{Y^\delta}} + \text{osc}^\delta(g). \end{aligned}$$

Neglecting the term $\text{osc}^\delta(g)$, known as *data oscillation*, as well as the factor $\|Q^\delta\|_{\mathcal{L}(Y, Y)}$, and using that $\sup_{0 \neq v \in Y^\delta} \frac{|(g - Bu_\varepsilon^\delta)(v)|}{\|v\|_Y} \approx \sup_{0 \neq v \in Y^\delta} \frac{|(g - Bu_\varepsilon^\delta)(v)|}{\|v\|_{Y^\delta}}$, we will use

$$\sqrt{\sup_{0 \neq v \in Y^\delta} \frac{|(g - Bu_\varepsilon^\delta)(v)|^2}{\|v\|_{Y^\delta}^2} + \|h - Cu_\varepsilon^\delta\|_W^2}$$

to estimate $\|f - Au_\varepsilon^\delta\|_V$.

Having solved u_ε^δ from either (4.15) or (5.3), this estimator can be computed at the expense of computing only a few inner products. Indeed with $v_\varepsilon^\delta \in Y^\delta$ defined by $\langle v_\varepsilon^\delta, v \rangle_{Y^\delta} = (g - Bu_\varepsilon^\delta)(v)$ ($v \in Y^\delta$), the supremum under the square root equals $\|v_\varepsilon^\delta\|_{Y^\delta}^2$. When u_ε^δ is determined by solving the mixed system (4.15), this v_ε^δ is the second component of the solution. When u_ε^δ is determined by solving the symmetric positive definite system (5.3), it holds that $v_\varepsilon^\delta := G_Y^\delta(g - Bu_\varepsilon^\delta)$ and $\|v_\varepsilon^\delta\|_{Y^\delta}^2 = (g - Bu_\varepsilon^\delta)(G_Y^\delta(g - Bu_\varepsilon^\delta))$.

Since $(\text{Id} - Q^\delta)'BX^\delta = 0$, it holds that

$$\text{osc}^\delta(g) \leq \|Q^\delta\|_{\mathcal{L}(Y, Y)} \min_{z \in X^\delta} \|g - Bz\|_{Y'},$$

so that the term that we have ignored is in any case $\mathcal{O}(\min_{z \in X^\delta} \|g - Bz\|_{Y'})$. Of course this is not completely satisfactory, because it does not ensure that our estimator of $\|f - Au_\varepsilon^\delta\|_V$ is reliable. In many cases for a suitable choice of $(Y^\delta)_{\delta \in \Delta}$, the Fortin interpolators can be constructed such that, in any case for smooth u and g , $\text{osc}^\delta(g)$ is of higher order than $\min_{z \in X^\delta} \|u - z\|_X$, in which case it can be expected that $\text{osc}^\delta(g)$ is asymptotically negligible. We do not discuss this issue further.

7. VERIFICATION OF THE INF-SUP CONDITION

In this section we establish the validity of the inf-sup condition (4.2) for the examples discussed earlier.

7.1. Verification of (4.2) for the Cauchy problem for Poisson's equation (Example 3.1)

Let $\Omega \subset \mathbb{R}^d$ be a polytope, and recall that $A = (B_1, B_2) \in \mathcal{L}(H^1(\Omega), (H_{0,\Sigma^c}^1(\Omega) \times \tilde{H}^{-\frac{1}{2}}(\Sigma)))'$ where $(B_1 z)(v) = \int_{\Omega} \nabla z \cdot \nabla v \, dx$, and $B_2 = \gamma_{\Sigma}$ (see notations introduced around Eq. (4.1)). By viewing B_2 here as an operator in $\mathcal{L}(H^1(\Omega), (\tilde{H}^{-\frac{1}{2}}(\Sigma)))'$, i.e., by using that $(\tilde{H}^{-\frac{1}{2}}(\Sigma))' = H^{\frac{1}{2}}(\Sigma)$, we will avoid having to evaluate $H^{\frac{1}{2}}(\Sigma)$ -norms of residuals.

Let $(\mathcal{T}^{\delta})_{\delta \in \Delta}$ being a family of conforming, uniformly shape regular partitions of $\bar{\Omega}$ into (closed) d -simplices, with $\mathcal{F}(\mathcal{T}^{\delta})$ denoting the set of the (closed) faces of $\mathcal{T}^{\delta}$, and $\partial \mathcal{T}^{\delta} := \cup \mathcal{F}(\mathcal{T}^{\delta})$ its skeleton. Assuming $\bar{\Sigma}$ to be the union of some $e \in \mathcal{F}(\mathcal{T}^{\delta})$, we take

$$X^{\delta} = \mathcal{S}_{\mathcal{T}^{\delta}}^{0,1} := \{z \in C(\bar{\Omega}) : z|_T \in \mathcal{P}_1(T) \ (T \in \mathcal{T}^{\delta})\},$$

but expect that similar results can be shown for higher order finite element spaces.

We will approximate the solution of the Cauchy problem by the minimizer over X^{δ} of the regularized least-squares functional that, in an abstract setting, was introduced and analyzed in Section 4. To show the inf-sup condition (4.2) for the triple B , $(X^{\delta})_{\delta \in \Delta}$, and a suitable family of test spaces $(Y_1^{\delta} \times Y_2^{\delta})_{\delta \in \Delta} \subset Y_1 \times Y_2 := H_{0,\Sigma^c}^1(\Omega) \times \tilde{H}^{-\frac{1}{2}}(\Sigma)$, it suffices to show such inf-sup conditions for $(B_i, (X^{\delta})_{\delta \in \Delta}, (Y_i^{\delta})_{\delta \in \Delta})$ and $i \in \{1, 2\}$ separately, which will be done in Propositions 7.1 and 7.2. Indeed, uniformly bounded $Q_i^{\delta} \in \mathcal{L}(Y_i, Y_i^{\delta})$ with $(B_i X^{\delta})((\text{Id} - Q_i^{\delta})Y_i) = 0$ give uniformly bounded $Q^{\delta} := Q_2^{\delta} \times Q_1^{\delta} \in \mathcal{L}(Y_1 \times Y_2, Y_1^{\delta} \times Y_2^{\delta})$ with $(BX^{\delta})((\text{Id} - Q^{\delta})(Y_1 \times Y_2)) = 0$.

Proposition 7.1. *For each $\delta \in \Delta$, let $\mathcal{T}_s^{\delta}$ be a refinement of $\mathcal{T}^{\delta}$ such that³ each $T \in \mathcal{T}^{\delta}$ is subdivided into a uniformly bounded number of uniformly shape regular d -simplices, and, when $d > 1$, $\mathcal{T}_s^{\delta}$ has a vertex interior to each $e \in \mathcal{F}(\mathcal{T}^{\delta})$ with $e \not\subset \bar{\Sigma}$. Then for*

$$Y_1^{\delta} := \mathcal{S}_{\mathcal{T}_s^{\delta}}^{0,1} \cap H_{0,\Sigma^c}^1(\Omega),$$

it holds that

$$\inf_{\delta \in \Delta} \inf_{\{z \in X^{\delta} : B_1 z \neq 0\}} \sup_{\{0 \neq v \in Y_1^{\delta}\}} \frac{|(B_1 z)(v)|}{\|B_1 z\|_{H_{0,\Sigma^c}^1(\Omega)'} \|v\|_{H^1(\Omega)}} > 0.$$

Proof. For $d > 1$, let J^{δ} denote a Scott–Zhang quasi-interpolator (cf. [34]) mapping into $X^{\delta} \cap H_{0,\Sigma^c}^1(\Omega) \subset Y_1^{\delta}$, i.e., one that preserves homogeneous boundary conditions on Σ^c . Using the technique applied in [6], for $v \in H_{0,\Sigma^c}^1(\Omega)$ we define

$$v^{\delta} = Q_1^{\delta} v := J^{\delta} v + \sum_{\{e \in \mathcal{F}(\mathcal{T}^{\delta}) : e \not\subset \bar{\Sigma}\}} \frac{\int_e (\text{Id} - J^{\delta}) v \, ds}{\int_e \phi_e \, ds} \phi_e \in Y_1^{\delta},$$

where $\phi_e \in Y_1^{\delta}$ is such that ϕ_e has values in $[0, 1]$, ϕ_e is 1 at a vertex of $\mathcal{T}_s^{\delta}$ interior to e , and ϕ_e vanishes outside $\cup \omega_e$, where $\omega_e := \{T \in \mathcal{T}^{\delta} : e \subset \partial T\}$. Then for each $e \in \mathcal{F}(\mathcal{T}^{\delta})$, it holds that $\int_e v - v^{\delta} \, ds = 0$, and so for $z \in X^{\delta}$,

$$(B_1 z)(v - v^{\delta}) = \sum_{T \in \mathcal{T}^{\delta}} \left\{ - \int_T \Delta z (v - v^{\delta}) \, dx + \int_{\partial T} \frac{\partial z}{\partial n} (v - v^{\delta}) \, ds \right\} = 0.$$

So Q_1^{δ} is a Fortin interpolator into Y_1^{δ} . In view of Theorem 4.3 we need to show $\sup_{\delta \in \Delta} \|Q_1^{\delta}\|_{\mathcal{L}(H_{0,\Sigma^c}^1(\Omega), H_{0,\Sigma^c}^1(\Omega))} < \infty$.

By an application of the trace theorem and a homogeneity argument, the construction of a Scott–Zhang quasi-interpolator shows that for arbitrary $T_e \in \mathcal{T}^{\delta}$ with $T_e \subset \omega_e$, with $h_e := \text{diam}(e)$ it holds that

$$\|(\text{Id} - J^{\delta})v\|_{L_2(e)} \lesssim h_e^{-\frac{1}{2}} \|(\text{Id} - J^{\delta})v\|_{L_2(T_e)} + h_e^{\frac{1}{2}} \|(\text{Id} - J^{\delta})v\|_{H^1(T_e)} \lesssim h_e^{\frac{1}{2}} |v|_{H^1(\omega_{T_e})},$$

³Obviously, $\mathcal{T}_s^{\delta}$ can be a further refinement of a partition that satisfies the listed requirements. Indeed, such a further refinement renders a Y_1^{δ} that is only larger.

where $\omega_T := \cup\{T' \in \mathcal{T}^\delta : |T \cap T'| > 0\}$. From $\int_e \phi_e \, ds \gtrsim h_e^{d-1}$, $\|\phi_e\|_{H^1(\Omega)} \lesssim h_e^{\frac{d}{2}-1}$, and $|e|^{\frac{1}{2}} \lesssim h_e^{\frac{d-1}{2}}$, we conclude

$$\begin{aligned} \|Q_1^\delta v\|_{H^1(\Omega)} &\leq \|J^\delta v\|_{H^1(\Omega)} + \sum_{\{e \in \mathcal{F}(\mathcal{T}^\delta) : e \notin \bar{\Sigma}\}} \frac{|\int_e (\text{Id} - J^\delta) v \, ds|}{|\int_e \phi_e \, ds|} \|\phi_e\|_{H^1(\Omega)} \\ &\lesssim \|v\|_{H^1(\Omega)} + \sum_{\{e \in \mathcal{F}(\mathcal{T}^\delta) : e \notin \bar{\Sigma}\}} h_e^{1-d} h_e^{\frac{d}{2}-1} h_e^{\frac{d-1}{2}} h_e^{\frac{1}{2}} |v|_{H^1(\omega_{T_e})} \lesssim \|v\|_{H^1(\Omega)}. \end{aligned} \quad \square$$

Proposition 7.2. For $\delta \in \Delta$, let $\mathcal{E}_s^\delta$ be a red-refinement of the partition $\mathcal{E}^\delta := \{e \in \mathcal{F}(\mathcal{T}^\delta) : e \subset \bar{\Sigma}\}$ of Σ , i.e., in $\mathcal{E}_s^\delta$ each $(d-1)$ -simplex $e \in \mathcal{E}^\delta$ has been split into 2^{d-1} subsimplices that are similar to e . Then with

$$Y_2^\delta = \mathcal{S}_{\mathcal{E}_s^\delta}^{-1,0} := \{v \in L_2(\Sigma) : v|_{\tilde{e}} \in \mathcal{P}_0(\tilde{e}) \ (\tilde{e} \in \mathcal{E}_s^\delta)\},$$

it holds that

$$\inf_{\delta \in \Delta} \inf_{\{z \in X^\delta : \gamma_\Sigma z \neq 0\}} \sup_{\{0 \neq v \in Y_2^\delta\}} \frac{|\int_\Sigma z v \, ds|}{\|\gamma_\Sigma z\|_{H^{\frac{1}{2}}(\Sigma)} \|v\|_{\tilde{H}^{-\frac{1}{2}}(\Sigma)}} > 0.$$

Proof. Although formulated differently, a proof of this statement can be found in Theorem 4.1 and Lemma 5.6 of [36]. We summarize the main steps. With $\Phi^\delta = \{\phi_\nu\}$ denoting the nodal basis for $\mathcal{S}_{\mathcal{E}^\delta}^{0,1}$, there exists a collection $\Psi^\delta = \{\psi_\nu\} \subset Y_2^\delta$ for which $\langle \phi_\nu, \psi_{\nu'} \rangle_{L_2(\Sigma)} \approx \delta_{\nu\nu'} \|\phi_\nu\|_{L_2(\Sigma)} \|\psi_{\nu'}\|_{L_2(\Sigma)}$ and $\text{supp } \psi_\nu \subseteq \text{supp } \phi_\nu$. The resulting biorthogonal Fortin projector Q_2^δ with $\text{ran } Q_2^\delta = \text{span } \Psi^\delta$ and $\text{ran}(\text{Id} - Q_2^\delta) = (\mathcal{S}_{\mathcal{E}^\delta}^{0,1})^\perp_{L_2(\Sigma)}$ is uniformly bounded w.r.t. $L_2(\Sigma)$, and, thanks to the approximation properties of $\mathcal{S}_{\mathcal{E}^\delta}^{0,1}$, its adjoint is uniformly bounded w.r.t. $H^1(\Sigma)$, and therefore also w.r.t. $H^{\frac{1}{2}}(\Sigma)$. We conclude that $\sup_{\delta \in \Delta} \|Q_2^\delta\|_{\mathcal{L}(\tilde{H}^{-\frac{1}{2}}(\Sigma), \tilde{H}^{-\frac{1}{2}}(\Sigma))} < \infty$ as required. $\square$

Remark 7.3. The proof of the above proposition hinges on the construction of a (uniformly locally supported) basis dual to Φ^δ from the span of piecewise constants w.r.t. some refined partition. Such a construction is not restricted to red-refinement, and obviously it applies when a deeper than red-refinement is applied as with the application of the bisect(5) refinement rule for two-dimensional Σ (see e.g., [32], Rem. 1).

Next, in order to avoid the evaluation of the $\tilde{H}^{-\frac{1}{2}}(\Sigma)$ -norm of arguments from Y_2^δ , let $G_2^\delta = G_2^{\delta'} \in \mathcal{L}(\text{Is}(Y_2^{\delta'}, Y_2^\delta))$ be such that $\|G_2^\delta f\|_{\tilde{H}^{-\frac{1}{2}}(\Sigma)}^2 \approx f(G_2^\delta f)$ ($f \in Y_2^{\delta'}$), and whose application can be performed in linear complexity (examples in [22, 37]). We equip Y_2^δ with $((G_2^\delta)^{-1} \cdot)(\cdot)^{\frac{1}{2}} \approx \|\cdot\|_{\tilde{H}^{-\frac{1}{2}}(\Sigma)}$.

In conclusion we have that the least squares approximation u_ε^δ from (4.4) of the solution of (3.2) is given as the minimizer over $z \in X^\delta$ of

$$\sup_{0 \neq (v_1, v_2) \in Y_1^\delta \times Y_2^\delta} \frac{|\int_\Omega \nabla z \cdot \nabla v_1 \, dx + \int_\Sigma \gamma_\Sigma z v_2 \, ds - (f_I(v_1) + \int_\Sigma f_N v_1 + f_D v_2 \, ds)|^2}{\|v_1\|_{H^1(\Omega)}^2 + \left((G_2^\delta)^{-1} v_2\right)(v_2)} + \varepsilon^2 \|z\|_H^2, \quad (7.1)$$

where $H = L_2(\Omega)$ or $H = H^1(\Omega)$ for Case (i) or Case (ii), of Example 3.1 in Section 3, respectively. According to (4.15) this u_ε^δ can be computed as the first component of the solution $(u_\varepsilon^\delta, v_{\varepsilon,1}^\delta, v_{\varepsilon,2}^\delta) \in X^\delta \times Y_1^\delta \times Y_2^\delta$ of

$$\begin{aligned} \int_\Omega \nabla \tilde{z} \cdot \nabla v_{\varepsilon,1}^\delta \, dx + \int_\Sigma \gamma_\Sigma \tilde{z} v_{\varepsilon,2}^\delta \, ds + \int_\Omega \nabla u_\varepsilon^\delta \cdot \nabla \tilde{v}_1 \, dx + \int_\Sigma \gamma_\Sigma u_\varepsilon^\delta \tilde{v}_2 \, ds + \langle \tilde{v}_1, v_{\varepsilon,1}^\delta \rangle_{H^1(\Omega)} \\ + \left((G_2^\delta)^{-1} \tilde{v}_2\right)(v_{\varepsilon,2}^\delta) - \varepsilon^2 \langle \tilde{z}, u_\varepsilon^\delta \rangle_H = f_I(\tilde{v}_1) + \int_\Sigma f_N \tilde{v}_1 + f_D \tilde{v}_2 \, ds \end{aligned}$$

for all $(\tilde{z}, \tilde{v}_1, \tilde{v}_2) \in X^\delta \times Y_1^\delta \times Y_2^\delta$, which after elimination of $v_{\varepsilon,2}^\delta$ reads as finding $(u_\varepsilon^\delta, v_{\varepsilon,1}^\delta) \in X^\delta \times Y_1^\delta$ that satisfies

$$\begin{aligned} \int_{\Omega} \nabla \tilde{z} \cdot \nabla v_{\varepsilon,1}^\delta dx + \int_{\Omega} \nabla u_\varepsilon^\delta \cdot \nabla \tilde{v}_1 dx + \langle \tilde{v}_1, v_{\varepsilon,1}^\delta \rangle_{H^1(\Omega)} - \int_{\Sigma} \gamma_{\Sigma} u_\varepsilon^\delta G_2^\delta \gamma_{\Sigma} \tilde{z} ds - \varepsilon^2 \langle u_\varepsilon^\delta, \tilde{z} \rangle_H \\ = f_I(\tilde{v}_1) + \int_{\Sigma} f_N \tilde{v}_1 - f_D G_2^\delta \gamma_{\Sigma} \tilde{z} ds \quad ((\tilde{z}, \tilde{v}_1) \in X^\delta \times Y_1^\delta). \end{aligned} \quad (7.2)$$

For this u_ε^δ the bound on $\|u - u_\varepsilon^\delta\|_\varepsilon$ from Theorem 4.1 applies, and, with the specification of j and η according to the estimates in Example 3.1 (i) or (ii), so does Theorem 2.2.

Remark 7.4. For completeness, we recall that in order to enhance the efficiency of an iterative solution process, additionally in (7.1) one may replace $\|v_1\|_{H^1(\Omega)}^2$ by an equivalent expression $((G_1^\delta)^{-1}v_1)(v_1)$, where $G_1^\delta = G_1^{\delta'} \in \mathcal{L}(\text{is}(Y_1^{\delta'}, Y_1^\delta))$ is such that $\|G_1^{\delta'} f\|_{H^1(\Omega)}^2 \approx f(G_1^{\delta'} f)$ ($f \in Y_1^{\delta'}$), and whose application can be performed in linear complexity (e.g., a multigrid preconditioner). With this change, $\langle \tilde{v}_1, v_{\varepsilon,1}^\delta \rangle_{H^1(\Omega)}$ in (7.2) reads as $((G_1^\delta)^{-1}\tilde{v}_1)(v_{\varepsilon,1}^\delta)$, and the unknown $v_{\varepsilon,1}^\delta$ can be eliminated which results in the symmetric positive definite system of finding $u_\varepsilon^\delta \in X^\delta$ that satisfies

$$(B_1 u_\varepsilon^\delta)(G_1^\delta B_1 \tilde{z}) + \int_{\Sigma} \gamma_{\Sigma} u_\varepsilon^\delta G_2^\delta \gamma_{\Sigma} \tilde{z} ds + \varepsilon^2 \langle u_\varepsilon^\delta, \tilde{z} \rangle_H = f_I(G_1^\delta B_1 \tilde{z}) + \int_{\Sigma} f_N G_1^\delta B_1 \tilde{z} + f_D G_2^\delta \gamma_{\Sigma} \tilde{z} ds \quad (\tilde{z} \in X^\delta),$$

cf. equation (5.3).

Remark 7.5 (Other approaches). In [12], the Dirichlet boundary condition was enforced by the so-called Nitsche method, which avoids the treatment of fractional Sobolev norms. Under the additional regularity condition that $u \in H^2(\Omega)$, an error estimate was derived based on the conditional stability estimates (i) or (ii).

In [7], $B_2 = \gamma_{\Sigma}$ was viewed as a map in $\mathcal{L}(H^1(\Omega), L_2(\Sigma))$, i.e., the discrepancy between the Dirichlet data and the trace of the approximate solution was measured in the weaker $L_2(\Sigma)$ -norm. For data such that the Cauchy problem has a (unique) solution, the solution of the resulting regularized least squares problem with $X^\delta = H^1(\Omega)$ (so without discretization) was shown to converge to the exact solution for $\varepsilon \rightarrow 0$.

7.2. Verification of (4.2) for the data-assimilation problem for the heat equation (Example 3.2)

Let $\Omega \subset \mathbb{R}^d$ be a polytope, and recall that $A = (B, \Gamma_{I \times \omega}) \in \mathcal{L}(X, V)$, where $I = (0, T)$, $X = L_2(I; H^1(\Omega)) \cap H^1(I; H^{-1}(\Omega))$ (Case (a)) or $X = L_2(I; H_0^1(\Omega)) \cap H^1(I; H^{-1}(\Omega))$ (Case (b)), $V = Y' \times L_2(I \times \omega)$, $Y = L_2(I; H_0^1(\Omega))$, $\Gamma_{I \times \omega} u = u|_{I \times \omega}$, and

$$(Bu)(v) = \int_I \int_{\Omega} \partial_t u v + \nabla_x u \cdot \nabla_x v dx dt.$$

Proposition 7.6. Let $(\mathcal{T}^\delta)_{\delta \in \Delta}$ be a family of conforming, uniformly shape regular partitions of $\bar{\Omega}$ into d -simplices. For each $\delta \in \Delta$, let $\mathcal{T}_s^\delta$ be a refinement of $\mathcal{T}^\delta$ such that each $T \in \mathcal{T}^\delta$ is subdivided into a uniformly bounded number of uniformly shape regular d -simplices and

- (1) when $d > 1$, $\mathcal{T}_s^\delta$ has a vertex interior to each $e \in \mathcal{F}(\mathcal{T}^\delta)$ with $e \not\subset \partial\Omega$,
- (2) $\inf_{\delta \in \Delta} \inf_{T \in \mathcal{T}^\delta} \inf_{0 \neq \tilde{p} \in \mathcal{P}_1(T)} \sup_{\{0 \neq \tilde{p} \in H_0^1(T) : \tilde{p}|_{\tilde{T}} \in \mathcal{P}_1(\tilde{T}) (\tilde{T} \in \mathcal{T}_s^\delta, \tilde{T} \subset T)\}} \frac{\langle p, \tilde{p} \rangle_{L_2(T)}}{\|p\|_{L_2(T)} \|\tilde{p}\|_{L_2(T)}} > 0,$

or $\mathcal{T}_s^\delta$ is a further refinement of such a partition (cf. footnote 3). With $(I^\delta)_{\delta \in \Delta}$ being a partition of $\bar{I}$, let

$$X^\delta := X \cap \left(\mathcal{S}_{I^\delta}^{0,1} \otimes \mathcal{S}_T^{0,1} \right), \quad Y^\delta := \mathcal{S}_{I^\delta}^{-1,1} \otimes \left(\mathcal{S}_{T_s^\delta}^{0,1} \cap H_0^1(\Omega) \right).$$

Then

$$\inf_{\delta \in \Delta} \inf_{\{z \in X^\delta : Bz \neq 0\}} \sup_{\{0 \neq v \in Y^\delta\}} \frac{|(Bz)(v)|}{\|Bz\|_{Y'} \|v\|_Y} > 0,$$

i.e., equation (4.2) is valid.

Remark 7.7. When $\mathcal{T}_s^\delta$ is generated from $\mathcal{T}^\delta$ by recursive either red-refinement or newest vertex bisection, for $d \in \{1, 2, 3\}$ condition (7.6) is satisfied as soon as the refinement is sufficiently deep such that $\mathcal{T}_s^\delta$ contains a (closed) d -simplex in the interior of each $T \in \mathcal{T}^\delta$ as can be verified by a direct computation on a reference d -simplex (see [21], §5.1).

Proof. It suffices to prove the statement for X that corresponds to Case (a), so that $X^\delta := \mathcal{S}_{T_s^\delta}^{0,1} \otimes \mathcal{S}_{T_s^\delta}^{0,1}$.

Let $Q_{1,x}^\delta: H_0^1(\Omega) \rightarrow \mathcal{S}_{T_s^\delta}^{0,1} \cap H_0^1(\Omega)$ be the Fortin operator Q_1 introduced in the proof of Proposition 7.1 (for $\Sigma = \partial\Omega$). It has the properties

$$\begin{aligned} \sup_{\delta \in \Delta} \|Q_{1,x}^\delta\|_{\mathcal{L}(H_0^1(\Omega), H_0^1(\Omega))} &< \infty, \quad \int_e (\text{Id} - Q_{1,x}^\delta)v \, ds = 0 \quad (v \in H_0^1(\Omega), e \in \mathcal{F}(\mathcal{T}^\delta)), \\ \|(\text{Id} - Q_{1,x}^\delta)v\|_{L_2(T)} &\lesssim \text{diam}(T)|v|_{H^1(\omega_T)} \quad (v \in H_0^1(\Omega), T \in \mathcal{T}^\delta), \end{aligned}$$

which also implies $\|(\text{Id} - Q_{1,x}^\delta)'g\|_{H^{-1}(\Omega)} \lesssim \sum_{T \in \mathcal{T}^\delta} \text{diam}(T)^2 \|g\|_{L_2(T)}^2$ ($g \in L_2(\Omega)$).

Thanks to (7.6), Theorem 4.3 shows that there exists a projector $Q_{2,x}^\delta: L_2(\Omega) \rightarrow \{v \in \mathcal{S}_{T_s^\delta}^{0,1} : v|_{\partial\mathcal{T}^\delta} = 0\}$, where, for $T \in \mathcal{T}^\delta$, $(Q_{2,x}^\delta v)|_T$ only depends on $v|_T$, with

$$\sup_{\delta \in \Delta} \|Q_{2,x}^\delta\|_{\mathcal{L}(L_2(\Omega), L_2(\Omega))} < \infty, \quad \text{ran}(\text{Id} - Q_{2,x}^\delta) \perp_{L_2(\Omega)} \mathcal{S}_{T_s^\delta}^{-1,1},$$

so that $Q_{2,x}^\delta$ reproduces $\mathcal{S}_{T_s^\delta}^{-1,1}$, and so

$$\left\| (\text{Id} - Q_{2,x}^\delta)'g \right\|_{L_2(T)} \lesssim \text{diam}(T)|g|_{H^1(T)} \quad (g \in H^1(\Omega), T \in \mathcal{T}^\delta).$$

With $Q_x^\delta := Q_{1,x}^\delta + Q_{2,x}^\delta - Q_{2,x}^\delta Q_{1,x}^\delta: H_0^1(\Omega) \rightarrow \mathcal{S}_{T_s^\delta}^{0,1} \cap H_0^1(\Omega)$, and Q_t^δ being the $L_2(I)$ -orthogonal projector onto $\mathcal{S}_{T_s^\delta}^{-1,1}$, we take $Q^\delta := Q_t^\delta \otimes Q_x^\delta$. From

$$\|Q_{2,x}^\delta(\text{Id} - Q_{1,x}^\delta)v\|_{H^1(T)} \lesssim \text{diam}(T)^{-1} \|(\text{Id} - Q_{1,x}^\delta)v\|_{H^1(T)} \lesssim |v|_{H^1(\omega_T)} \quad (T \in \mathcal{T}^\delta),$$

one infers that $\sup_{\delta \in \Delta} \|Q^\delta\|_{\mathcal{L}(Y,Y)} < \infty$.

For $z \in X^\delta$ and $v \in Y$, by writing $(Bz)((\text{Id} - Q^\delta)v) = (Bz)((\text{Id} - Q_t^\delta) \otimes \text{Id} v) + (Bz)(Q_t^\delta \otimes (\text{Id} - Q_x^\delta)v)$, and by realizing that the second term equals

$$\sum_{J \in I^\delta} \int_J \sum_{T \in \mathcal{T}^\delta} \left\{ \int_T \partial_t z Q_t^\delta \otimes (\text{Id} - Q_{2,x}^\delta)(\text{Id} - Q_{1,x}^\delta)v \, dx + \int_{\partial T} \frac{\partial z}{\partial n} Q_t^\delta \otimes (\text{Id} - Q_{1,x}^\delta + Q_{2,x}^\delta Q_{1,x}^\delta)v \, ds \right\} dt$$

one infers that both terms vanish, so that Q^δ is a valid Fortin operator. $\square$

Remark 7.8. For Case (b) and assuming Ω being convex, in Theorem 5.7 of [21] the statement of Proposition 7.6 was proven without Condition (7.6), which upon assuming Condition (7.6), is, however, harmless.

Remark 7.9 (Different meshes in different time-slabs). The result of Proposition 7.6 directly extends to the situation that

$$X^\delta := \{z \in X : z|_{J \times T} \in \mathcal{P}_1(J) \otimes \mathcal{P}_1(T) \ (J \in I^\delta, T \in \mathcal{T}^\delta(J))\},$$

$$Y^\delta := \left\{ v \in Y : v|_{J \times \tilde{T}} \in \mathcal{P}_1(J) \otimes \mathcal{P}_1(\tilde{T}) \quad \left(J \in I^\delta, \tilde{T} \in \mathcal{T}_s^\delta(J) \right) \right\},$$

i.e., when the spatial meshes underlying X^δ and Y^δ possibly depend on the time interval $J \in I^\delta$. Notice that the global continuity of functions in X^δ implies that nonconformities between spatial meshes in adjacent time-slabs result in “hanging nodes”.

Remark 7.10. To end up with really general partitions of the time-space cylinder into prismatic elements, it would be necessary to allow elements $J \times T$ and $\hat{J} \times \hat{T}$ with $T \cap \hat{T}$ being a $(d-1)$ -simplex and $J \subsetneq \hat{J}$, i.e., to allow hanging nodes on interfaces perpendicular to the plane $\{0\} \times \Omega$. Uniform boundedness in $\mathcal{L}(Y, Y)$ of Fortin interpolators, needed to prove inf-sup stability, requires essentially that they map functions that are constant in x to functions that are constant in x . For such general partitions, unfortunately we do not see how this can be done.

The kind of partitions for which we are able to prove inf-sup stability do not allow local (adaptive) refinements. To circumvent this problem, in the next section we consider a reformulation of the data assimilation problem for the heat equation as a first order system.

Concluding, we have that the least squares approximation u_ε^δ from (4.4) of the solution of the data-assimilation problem for the heat equation is given as the minimizer

$$u_\varepsilon^\delta := \operatorname{argmin}_{z \in X^\delta} \left\{ \sup_{0 \neq v \in Y^\delta} \frac{|(Bz - g)(v)|^2}{\|v\|_{L_2(I; H^1(\Omega))}^2} + \|\Gamma_{I \times \omega} z - h\|_{L_2(I \times \omega)}^2 + \varepsilon^2 \|z\|_{L_2(I \times \Omega)}^2 \right\},$$

where in Case (b) the regularizing term $\varepsilon^2 \|z\|_{L_2(I \times \Omega)}^2$ can be omitted. By replacing the nominator $\|v\|_{L_2(I; H^1(\Omega))}^2$ by $((G_Y^\delta)^{-1}v)(v)$ for some uniform preconditioner $G_Y^\delta = G_Y^{\delta'} \in \mathcal{L}(\text{is}(Y^{\delta'}, Y^\delta))$, the resulting system can be reduced to a symmetric positive definite system. For Y^δ as in tensor product setting from Proposition 7.2 or in the time-slab setting from Remark 7.9, such preconditioners are easily constructed using multi-grid preconditioners in the spatial direction. For the resulting u_ε^δ , the bound on $\|u - u_\varepsilon^\delta\|_\varepsilon$ from Theorem 4.1 applies, and so do the bounds from Theorem 2.2 with the specification of j and η corresponding to Example 3.2.

7.3. Data-assimilation for the heat equation as a first order system

We reconsider the data assimilation problem $Au = (B, \Gamma_{I \times \omega})u = (g, h)$ from Example 3.2. Writing here u as u_1 , following [23, 25] for the well-posed forward heat problem we rewrite this data-assimilation problem as a first order system for $\mathbf{u} = (u_1, \mathbf{u}_2)$, where $\mathbf{u}_2 = -\nabla_x u_1$. To that end restricting (f, g) to $L_2(I \times \Omega) \times L_2(I \times \omega) \subsetneq L_2(I; H^{-1}(\Omega)) \times L_2(I \times \omega)$, we consider the system

$$\tilde{A}\mathbf{u} := (\mathbf{u}_2 + \nabla_x u_1, \operatorname{div} \mathbf{u}, \Gamma_{I \times \omega} u_1) = (0, f, g), \quad (7.3)$$

where $\operatorname{div} \mathbf{u} := \partial_t u_1 + \operatorname{div}_x \mathbf{u}_2$. With

$$\tilde{X} := \left\{ \mathbf{u} = (u_1, \mathbf{u}_2) \in L_2(I; H^1(\Omega)) \times L_2(I \times \Omega)^d : \operatorname{div} \mathbf{u} \in L_2(I \times \Omega) \right\}$$

equipped with the graph norm, where for Case (b) the space $L_2(I; H^1(\Omega))$ should be read as $L_2(I; H_0^1(\Omega))$, it holds that

$$\tilde{A} \in \mathcal{L} \left(\tilde{X}, \underbrace{L_2(I \times \Omega)^d \times L_2(I \times \Omega) \times L_2(I \times \omega)}_{\tilde{V} :=} \right).$$

We are going to derive (un)conditional stability estimates for the first order system (7.3).

For any $\mathbf{z} = (z_1, \mathbf{z}_2) \in \tilde{X}$ and $v \in L_2(I; H_0^1(\Omega))$, integration-by-parts gives

$$(Bz_1)(v) = \int_I \int_{\Omega} \partial_t z_1 v + \nabla_x z_1 \cdot \nabla_x v \, dx \, dt = \int_I \int_{\Omega} \operatorname{div} \mathbf{z} v + (\mathbf{z}_2 + \nabla_x z_1) \cdot \nabla_x v \, dx \, dt,$$

so that, thanks to $\max(\|v\|_{L_2(I \times \Omega)}, \|\nabla_x v\|_{L_2(I \times \Omega)^d}) \leq \|v\|_{L_2(I; H^1(\Omega))}$,

$$\begin{aligned} \|Bz_1\|_{L_2(I; H^{-1}(\Omega))} &= \sup_{0 \neq v \in L_2(I; H_0^1(\Omega))} \frac{\int_I \int_{\Omega} \partial_t z_1 v + \nabla_x z_1 \cdot \nabla_x v \, dx \, dt}{\|v\|_{L_2(I; H^1(\Omega))}} \\ &\leq \|\operatorname{div} \mathbf{z}\|_{L_2(I \times \Omega)} + \|\mathbf{z}_2 + \nabla_x z_1\|_{L_2(I \times \Omega)^d}, \end{aligned} \quad (7.4)$$

and so $\|Az_1\|_V = \sqrt{\|Bz_1\|_{L_2(I; H^{-1}(\Omega))}^2 + \|\Gamma_{I \times \omega} z_1\|_{L_2(I \times \omega)}^2} \lesssim \|\tilde{A}\mathbf{z}\|_{\tilde{V}}$. From the conditional or unconditional stability estimates for the second order formulation from (a) and (b), respectively, we infer that

$$\|z_1\|_{L_2((T_1, T_2); H^1(\tilde{\omega}))} \lesssim \left(\|\tilde{A}\mathbf{z}\|_{\tilde{V}} + \|z_1\|_{L_2(I \times \Omega)} \right)^{1-\sigma} \|\tilde{A}\mathbf{z}\|_{\tilde{V}}^{\sigma} \quad (7.5)$$

or

$$\|z_1\|_{L_2((T_1, T); H^1(\Omega)) \cap H^1((T_1, T); H^{-1}(\Omega))} \lesssim \|\tilde{A}\mathbf{z}\|_{\tilde{V}}, \quad (7.6)$$

in Cases (a) and (b), respectively.

To conclude (un)conditional stability for the first order system formulation in both cases, it remains to check injectivity of $(\tilde{A}, \tilde{L})$ where $\tilde{L} \in \mathcal{L}(\tilde{X}, L_2(I \times \Omega))$ is given by $\tilde{L}z := z_1$ in Case (a), and $\tilde{L} := 0$ in Case (b). In Case (a), $(\tilde{A}, \tilde{L})\mathbf{z} = 0$ implies $z_1 = 0$ and thus $\mathbf{z}_2 = -\nabla_x z_1 = 0$; and in Case (b), $\tilde{A}\mathbf{z} = 0$ implies $Az_1 = 0$, which gives $z_1 = 0$, and so $\mathbf{z}_2 = -\nabla_x z_1 = 0$.

Given a finite dimensional subspace $X^{\delta} \subset \tilde{X}$, the regularized least squares approximation $\mathbf{u}_{\varepsilon}^{\delta} \in X^{\delta}$ of the solution of (7.3) is given by

$$\operatorname{argmin}_{\mathbf{z}=(z_1, \mathbf{z}_2) \in X^{\delta}} \|\mathbf{z}_2 + \nabla_x z_1\|_{L_2(I \times \Omega)^d}^2 + \|\operatorname{div} \mathbf{z} - g\|_{L_2(I \times \Omega)}^2 + \|\Gamma_{I \times \omega} z_1 - h\|_{L_2(I \times \omega)}^2 + \varepsilon^2 \|z_1\|_{L_2(I \times \Omega)}^2,$$

where in Case (b) the regularizing term $\varepsilon^2 \|z_1\|_{L_2(I \times \Omega)}^2$ can be omitted. The bound on $\|\mathbf{u} - \mathbf{u}_{\varepsilon}^{\delta}\|_{\varepsilon}$ from Theorem 4.1 applies, and so do the bounds from Theorem 2.2 with the specification of j and η corresponding to (7.5) or (7.6) in Cases (a) and (b), respectively.

The main *advantage* of this regularized first order system least squares (FOSLS) formulation is that all components of the residual are measured in L_2 -type norms, so that there is no need to introduce one or more of these components as independent variables, and to ensure inf-sup stability by a careful selection of “trial” and “test” spaces. As a consequence *any* finite dimensional subspace $X^{\delta} \subset \tilde{X}$ can be applied. A potential *disadvantage* is that to arrive at this formulation, in (7.4) the norm $\|\operatorname{div} \mathbf{u}\|_{L_2(I; H^{-1}(\Omega))}$ was estimated on the stronger norm $\|\operatorname{div} \mathbf{u}\|_{L_2(I \times \Omega)}$, which may result in reduced convergence rates for solutions that have singularities. Experiments reported on in [23] for the well-posed forward heat equation show that the risk of getting very low rates is not imaginary.

7.4. Verification of (4.2) for the data-assimilation problem for the wave equation (Example 3.3)

Let $\Omega \subset \mathbb{R}^d$ be a polytope (cf. Rem. 3.4), and recall that $A = (\square, \gamma_{I \times \partial\Omega}, \Gamma_{I \times \omega}) \in \mathcal{L}(X, H^{-1}(I \times \Omega) \times L_2(I \times \partial\Omega) \times L_2(I \times \omega))$. We have to verify (4.2) for $B = \square \in \mathcal{L}(H^1(I \times \Omega), H^{-1}(I \times \Omega))$.

For $(\mathcal{T}^{\delta})_{\delta \in \Delta}$ being a family of conforming, uniformly shape regular partitions of $I \times \Omega$ into $(d+1)$ -simplices, we take

$$X^{\delta} := \mathcal{S}_{\mathcal{T}^{\delta}}^{0,1}.$$

Proposition 7.11. *For each $\delta \in \Delta$, let $\mathcal{T}_s^\delta$ be a refinement of $\mathcal{T}^\delta$ such that³ each $T \in \mathcal{T}^\delta$ is subdivided into a uniformly bounded number of uniformly shape regular $(d+1)$ -simplices, and $\mathcal{T}_s^\delta$ has a vertex interior to each $e \in \mathcal{F}(\mathcal{T}^\delta)$ with $e \not\subset \partial(I \times \Omega)$. Then for*

$$Y^\delta := \mathcal{S}_{\mathcal{T}_s^\delta}^{0,1} \cap H_0^1(I \times \Omega),$$

it holds that

$$\inf_{\delta \in \Delta} \inf_{\{z \in X^\delta : \square z \neq 0\}} \sup_{\{0 \neq v \in Y^\delta\}} \frac{|(\square z)(v)|}{\|\square z\|_{H^{-1}(I \times \Omega)} \|v\|_{H^1(I \times \Omega)}} > 0.$$

The proof of this proposition is similar to the proof of Proposition 7.1. The fact that the current result concerns the wave operator on $I \times \Omega$ instead of the Laplacian on Ω does not make any difference.

Concluding, the least squares approximation $u_\varepsilon^\delta \equiv u^\delta$ from (4.4) of the solution of the data-assimilation problem for the wave equation is given as the minimizer

$$u^\delta := \operatorname{argmin}_{z \in X^\delta} \left\{ \sup_{0 \neq v \in Y^\delta} \frac{|(\square z - f)(v)|^2}{\|v\|_{H^1(I \times \Omega)}^2} + \|\gamma_{I \times \partial\Omega} z - g\|_{L_2(I \times \partial\Omega)}^2 + \|\Gamma_{I \times \omega} z - h\|_{L_2(I \times \omega)}^2 \right\}.$$

By replacing the denominator $\|v\|_{H^1(I \times \Omega)}^2$ by $((G_Y^\delta)^{-1}v)(v)$ for some uniform preconditioner $G_Y^\delta = G_Y^{\delta'} \in \mathcal{L}(\mathcal{Y}^{\delta'}, Y^\delta)$, the resulting system can be reduced to a symmetric positive definite system. For the resulting u^δ , the bound on $\|u - u^\delta\|_\varepsilon$ from Theorem 4.1 applies (where $\|\cdot\|_\varepsilon$ is ε -independent because $L = 0$), and so do the bounds from Theorem 2.2 with the specification of j and η corresponding to Example 3.3.

8. NUMERICAL EXPERIMENTS

The package P1-FEM from [24] was adjusted to implement the problems in MATLAB. The finite element library NGSolve from [33] was used to implement the data assimilation problem for the heat equation in two dimensions.

We consider finite element spaces w.r.t. uniformly shape regular partitions $\mathcal{T}^\delta$ of n -dimensional bounded domains (e.g., Ω , $I \times \Omega$, or $\Gamma \subset \partial\Omega$) into n -simplices, where we restrict ourselves to partitions that are quasi-uniform. In view of the latter, we can speak of the mesh size h_δ , which number raised to the power $-n$ is proportional to $\#\mathcal{T}$ and thus to the dimension of the finite element space (of fixed order).

In this section the relative error with respect to some $j: X \rightarrow \mathbb{R}_+$ in a numerical approximation u_ε^δ to the prescribed solution u is defined as $\frac{j(u - u_\varepsilon^\delta)}{j(u)}$.

8.1. Cauchy problem for Poisson's equation

For $\Omega = (0, \pi) \times (0, 1)$, $\Sigma = (0, \pi) \times \{0\}$, and $\Sigma^c = \partial\Omega \setminus \bar{\Sigma}$, given $f = (f_I, f_D, f_N) \in (H_{0, \Sigma^c}^1(\Omega))' \times H^{\frac{1}{2}}(\Sigma) \times H^{-\frac{1}{2}}(\Sigma)$ we consider the problem of finding $u \in H^1(\Omega)$ that solves

$$-\Delta u = f_I \text{ on } \Omega, \quad u = f_D \text{ on } \Sigma, \quad \frac{\partial u}{\partial n} = f_N \text{ on } \Sigma, \quad (8.1)$$

or, more precisely, its variational formulation $(B_1 u, B_2 u) = (g_{f_I, f_N}, f_D)$ from (3.2).

We consider a sequence of uniform triangulations $(\mathcal{T}^\delta)_{\delta \in \Delta}$ of $\bar{\Omega}$, where each next triangulation is created from its predecessor by one uniform newest vertex bisection starting from an initial triangulation that consists of 12 triangles created from a subdivision of $\bar{\Omega}$ into 3 rectangles of size $\frac{\pi}{3} \times 1$ by cutting each of these rectangles along their diagonals. The three interior vertices in this initial triangulation are labelled as the “newest vertices” of all 4 triangles that contain them.

Following Section 7.1, we take $X^\delta = \mathcal{S}_{\mathcal{T}^\delta}^{0,1}(\Omega)$, and with $\mathcal{T}_s^\delta$ denoting the second successor of $\mathcal{T}^\delta$ in the sequence of triangulations, we set $Y_1^\delta = \mathcal{S}_{\mathcal{T}_s^\delta}^{0,1} \cap H_{0, \Sigma^c}^1(\Omega)$ and $Y_2^\delta = \mathcal{S}_{\mathcal{E}_s^\delta}^{-1,0}$, where $\mathcal{E}_s^\delta$ is the set of edges on $\bar{\Sigma}$ of $T \in \mathcal{T}_s^\delta$.

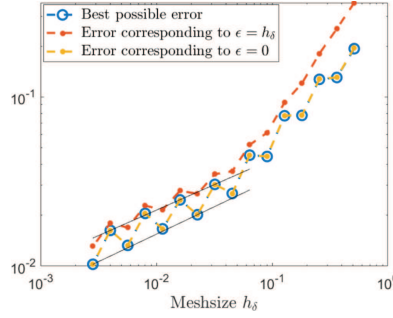


FIGURE 1. Poisson's equation. Meshsize *vs.* relative $L_2(\Omega)$ -error in case of unperturbed data. The asymptotic convergence rate in terms of $\#$ DoFs indicated by the solid straight lines is close to 0.15.

Considering the conditional stability estimate from Case (ii) in Example 3.1 (we did not test Case (i)), we recall that the idea behind our approach is to compute the minimizer u_ε^δ over X^δ of the regularized least squares functional $z \mapsto \|B_1 z - g_{f_I, f_N}\|_{H_{0, \Sigma^c}^1(\Omega)'}^2 + \|\gamma_\Sigma z - f_D\|_{\tilde{H}^{-\frac{1}{2}}(\Sigma)'}^2 + \varepsilon^2 \|z\|_{H^1(\Omega)}^2$. To make this method feasible without compromising its qualitative properties, we replace the suprema over $H_{0, \Sigma^c}^1(\Omega)$ and $\tilde{H}^{-\frac{1}{2}}(\Sigma)$ in the dual norms in the first two terms by suprema over Y_1^δ and Y_2^δ , respectively (see Props. 7.1 and 7.2). At the same time we replace the norms on $H^1(\Omega)$ and $\tilde{H}^{-\frac{1}{2}}(\Sigma)$ in the denominators by $((G_{Y_1}^\delta)^{-1} \cdot)(\cdot)^{\frac{1}{2}}$ and $((G_{Y_2}^\delta)^{-1} \cdot)(\cdot)^{\frac{1}{2}}$ for preconditioners $G_{Y_i}^\delta = G_{Y_i}^{\delta'} \in \mathcal{L}is(Y_i^{\delta'}, Y_i^\delta)$ with $\|G_{Y_1}^\delta f\|_{H^1(\Omega)}^2 \approx f(G_{Y_1}^\delta f)$ ($f \in Y_1^{\delta'}$) and $\|G_{Y_2}^\delta f\|_{\tilde{H}^{-\frac{1}{2}}(\Omega)}^2 \approx f(G_{Y_2}^\delta f)$ ($f \in Y_2^{\delta'}$). Then the resulting approximation u_ε^δ can be computed as the unique solution in X^δ of the symmetric positive definite system

$$(B_1 u_\varepsilon^\delta - g_{f_I, f_N})(G_{Y_1}^\delta B_1 \tilde{z}) + \int_{\Sigma} (\gamma_\Sigma u_\varepsilon^\delta - f_D) G_{Y_2}^\delta \gamma_\Sigma \tilde{z} \, ds + \varepsilon^2 \langle u_\varepsilon^\delta, \tilde{z} \rangle_{H^1(\Omega)} = 0 \quad (\tilde{z} \in X^\delta).$$

For $G_{Y_2}^\delta$ we take the (additive) multi-level preconditioner introduced in [22], and for $G_{Y_1}^\delta$ we use a common (multiplicative) multi-level preconditioner. Both preconditioners have linear computational complexity.

In all our experiments, we prescribe the solution

$$u(x, y) = \sin x \sinh y + \frac{1}{9} x^2,$$

which corresponds to (exact) data

$$f = (f_I, f_D, f_N) = \left(-\frac{2}{9}, x \mapsto \frac{1}{9} x^2, x \mapsto -\sin x\right). \quad (8.2)$$

We measure the errors of numerical solutions in the relative $L_2(\Omega)$ -norm, *i.e.*, the $L_2(\Omega)$ -norm divided by the $L_2(\Omega)$ -norm of the exact solution.

8.1.1. Unperturbed data

For the case of unperturbed data, we consider two strategies for choosing the regularization parameter, *viz.*, $\varepsilon = h_\delta$, the latter being the mesh-size, and $\varepsilon = 0$, and compare the results with those obtained with the experimentally found ε that minimizes $\|u - u_\varepsilon^\delta\|_{L_2(\Omega)}$. Note that since u is smooth, the choice $\varepsilon = h_\delta$ satisfies the conditions in (2.5).

The numerical results presented in Figure 1 indicate, however, that regularization is not helpful, although it somewhat improves the conditioning of the system.

Likely the oscillations in the curves from Figure 1 are due to the different geometry of the triangulations after an even or odd number of uniform refinements. The Hölder continuous behaviour, with exponent 0.15, of

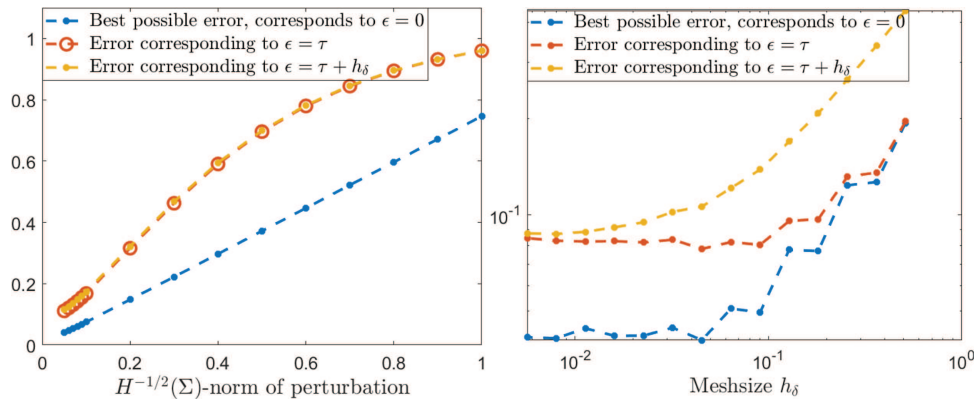


FIGURE 2. Poisson's equation. Norm of perturbation *vs.* relative $L_2(\Omega)$ -error in case of randomly perturbed Neumann data for different choices of ε . *Left:* random perturbations of varying $H^{-1/2}(\Sigma)$ -norm and fixed mesh with $\#\text{DoFs} \approx 10^5$. *Right:* random perturbations with $H^{-1/2}(\Sigma)$ -norm equal to 0.1, and varying mesh-size.

the error as function of the residual, the latter being of order h_δ , is better than the logarithmic dependence provided by the conditional stability estimate. That estimate, however, covers the case of a residual of most “nasty” type (and an infinitely fine mesh), whereas in our test, the residual is some specific function dependent on the partition and the prescribed solution.

8.1.2. Randomly perturbed data

We now perturb the Neumann datum f_N with a random piecewise constant $g \in \mathcal{S}_{\mathcal{E}_s}^{-1,0}$ with $\|g\|_{H^{-\frac{1}{2}}(\Sigma)} \approx \tau$. We achieved this by normalizing a random function in $\mathcal{S}_{\mathcal{E}_s}^{-1,0}$, taking values in $[0, 1]$, in a discrete $H^{-\frac{1}{2}}(\Sigma)$ -norm, that is uniformly equivalent to the true $H^{-\frac{1}{2}}(\Sigma)$ -norm, and then multiplying the result with τ . We used the discrete $H^{-\frac{1}{2}}(\Sigma)$ -norm constructed on p211 from [22] using results from [2].

Remark 8.1. An alternative for the latter is to construct a refinement $\mathcal{E}_{ss}^\delta$ of $\mathcal{E}_s^\delta$ such that for $g \in \mathcal{S}_{\mathcal{E}_s}^{-1,0}$, $\|g\|_{H^{-\frac{1}{2}}(\Sigma)} \approx \sup_{0 \neq v \in \mathcal{S}_{\mathcal{E}_{ss}}^{0,1} \cap H_0^1(\Sigma)} \frac{\int_\Sigma g v \, ds}{\|v\|_{H_{00}^{\frac{1}{2}}(\Sigma)}}$ (see [35], §3.1-2), after which an expression equivalent to the right-hand side can be computed using a standard multi-level preconditioner.

We compare the results obtained with the regularization strategies $\varepsilon = \tau + h_\delta$, which satisfies the conditions in (2.5), and $\varepsilon = \tau$, with those obtained with the experimentally found ε that minimizes $\|u - u_\varepsilon^\delta\|_{L_2(\Omega)}$, which also here turns out to be $\varepsilon = 0$. The results are presented in Figure 2.

We conclude that, for this problem, apparently such random perturbations are harmless, because without any regularization, for $h_\delta \downarrow 0$, which results in an increasingly ill-posed problem, their effect on the error hardly increases.

8.1.3. “Difficult” perturbations

From [1] we know that for $m \in \mathbb{N}$, the solution $u = u^{(m)}$ of the Cauchy problem (8.1) with data $f = f^{(m)} = (0, 0, f_N^{(m)})$ where $f_N^{(m)}(x) := -\sqrt{\frac{2m}{\pi}} \sin mx$, is given by $u^{(m)}(x, y) = \sqrt{\frac{2}{m\pi}} \sin mx \sinh my$. It holds that $\|f_N^{(m)}\|_{H^{-\frac{1}{2}}(\Sigma)} = 1$, $\|u^{(m)}\|_{H^1(\Omega)} \approx |u^{(m)}|_{H^1(\Omega)} = \sqrt{\frac{1}{2} \sinh 2m} \sim \frac{1}{2} e^m$ ($m \rightarrow \infty$), and $\|u^{(m)}\|_{L_2(\Omega)} \sim \frac{e^m}{2\sqrt{2m}}$ ($m \rightarrow \infty$), illustrating the strong ill-posedness of the Cauchy problem.

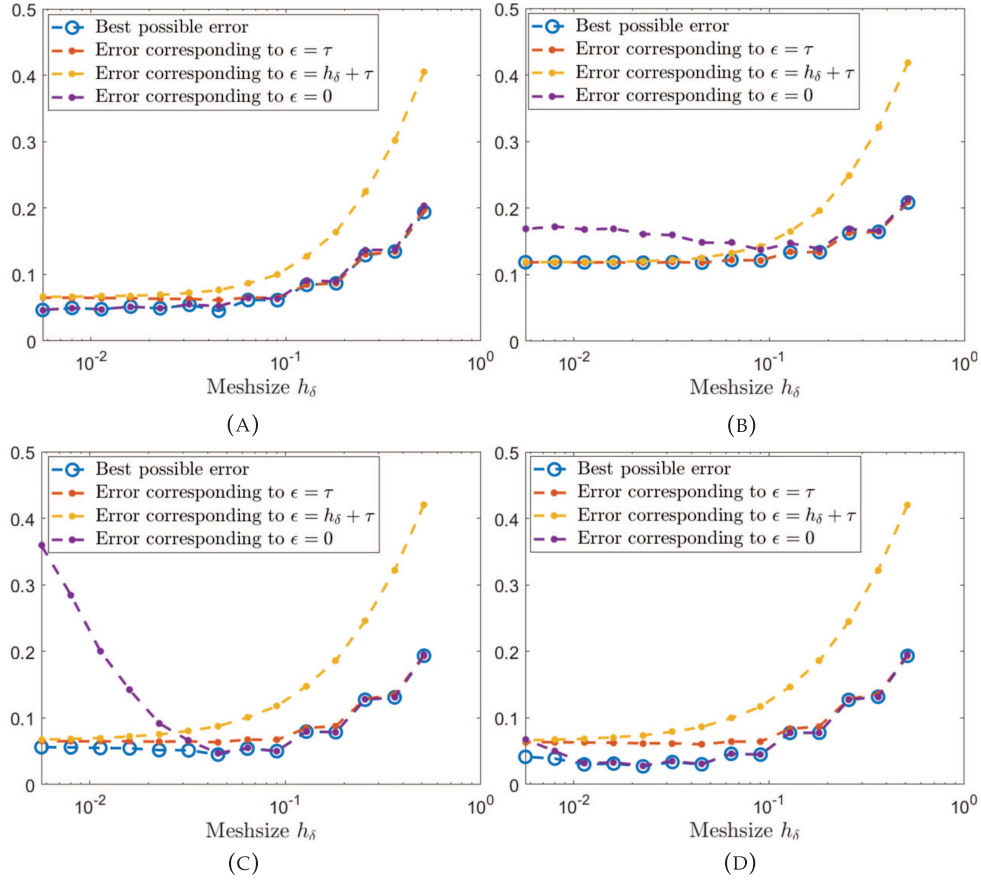


FIGURE 3. Poisson's equation. Meshsize *vs.* relative $L_2(\Omega)$ -error for different choices of ε and perturbation with $0.1\sqrt{\frac{2m}{\pi}}\sin mx$ of the Neumann datum for $m = 1, 3, 6, 16$. (A) For $m = 1$. (B) For $m = 3$. (C) For $m = 6$. (D) For $m = 16$.

We investigate our numerical solver when we perturb the exact Neumann datum from (8.2) with $0.1 * f^{(m)}$. We compare the same regularization strategies as with random perturbations. The results given in Figure 3 show that both for $m = 1, 3$ as well as for $m = 16$ regularization at most slightly improves the results. For $m = 1, 3$ this can be understood because the perturbation has an only modest effect on the solution. An explanation why for $m = 16$ regularization is hardly helpful is that on the meshes that we employed apparently the best representation of $u^{(16)}$ has a *much* smaller norm than $u^{(16)}$ itself. For the intermediate value $m = 6$, however, we clearly see that regularization *is* helpful.

8.2. Data-assimilation for the wave equation

For $\Omega = (0, 1)$, $I = (0, 1)$ and $\omega = (\frac{1}{2}, \frac{3}{4})$, given $(f, g, h) \in H^{-1}(I \times \Omega) \times L_2(I \times \partial\Omega) \times L_2(I \times \omega)$, we consider the problem of finding u that solves

$$\frac{\partial^2 u}{\partial t^2} - \Delta_x u = f \text{ on } I \times \Omega, \quad u = g \text{ on } I \times \partial\Omega, \quad u = h \text{ on } I \times \omega,$$

or, more precisely its variational formulation $Au = (f, g, h)$ with $A := (\square, \gamma_{I \times \partial\Omega}, \Gamma_{I \times \omega})$ given in Example 3.3.

We consider a sequence of uniform triangulations $(\mathcal{T}^\delta)_{\delta \in \Delta}$ of $\overline{I \times \Omega}$, where each next triangulation is created from its predecessor by one uniform newest vertex bisection starting from an initial triangulation that is created by cutting $I \times \Omega$ along both diagonals. The interior vertex in this initial triangulation is labelled as the “newest vertex” of all four triangles.

Following Section 7.4, we take $X^\delta := \mathcal{S}_{\mathcal{T}^\delta}^{0,1}(I \times \Omega)$, and with $\mathcal{T}_s^\delta$ denoting the second successor of $\mathcal{T}^\delta$ in the sequence of triangulations, we set $Y^\delta := \mathcal{S}_{\mathcal{T}_s^\delta}^{0,1} \cap H_0^1(I \times \Omega)$.

Considering the *unconditional* stability estimate (3.3) in Example 3.3, our approach is to minimize the least squares functional $\|\square z - f\|_{H^{-1}(I \times \Omega)}^2 + \|\gamma_{I \times \partial\Omega} z - g\|_{L_2(I \times \partial\Omega)}^2 + \|\Gamma_{I \times \omega} z - h\|_{L_2(I \times \omega)}^2$ over $z \in X^\delta$, so without regularization term. To make this method feasible without comprimizing its qualitative properties, first we replace the supremum from the first term by the supremum over Y^δ (see Prop. 7.11). Second, to make the computation of the resulting dual norm efficient, we introduce a preconditioner $G_Y^\delta \in \mathcal{L}is(Y^{\delta'}, Y^\delta)$ with $\|G_Y^\delta f\|_{H^1(\Omega)}^2 \approx f(G_Y^\delta f)$ ($f \in Y^{\delta'}$), and compute our approximation u^δ as the unique solution in X^δ of the symmetric positive definite system

$$(\square u^\delta - f)(G_Y^\delta \square \tilde{z}) + \langle \gamma_{I \times \partial\Omega} u^\delta - g, \gamma_{I \times \partial\Omega} \tilde{z} \rangle_{L_2(I \times \partial\Omega)} + \langle \Gamma_{I \times \omega} u^\delta - h, \Gamma_{I \times \omega} \tilde{z} \rangle_{L_2(I \times \omega)} = 0 \quad (\tilde{z} \in X^\delta).$$

For G_Y^δ we use a common (multiplicative) multi-level preconditioner.

In our experiments, we prescribe the solution

$$u(t, x) = \cos(\pi t) \sin(\pi x),$$

which corresponds to (exact) data

$$(f, g, h) = (0, 0, u|_{I \times \omega}).$$

We perform experiments with unperturbed and perturbed data. Instead of the error in the hard to evaluate norm $\|\cdot\|_{L_\infty(I; L_2(\Omega))} + \|\partial_t \cdot\|_{L_2(I; H^{-1}(\Omega))}$ from the unconditional stability estimate (3.3), we provide the *a posteriori* residual estimator from Section 6 given by

$$\sqrt{(f - \square u^\delta) \left(G_Y^\delta (f - \square u^\delta) + \|\gamma_{I \times \partial\Omega} u^\delta - g\|_{L_2(I \times \partial\Omega)}^2 + \|\Gamma_{I \times \omega} u^\delta - h\|_{L_2(I \times \omega)}^2 \right)},$$

which provides, modulo a constant factor, an upper bound for the aforementioned norm of the error up to data oscillations. We additionally provide the relative errors in the easily evaluable $L_2(I \times \Omega)$ - and $H^1(I \times \Omega)$ -norms (*i.e.*, these norms divided by the corresponding norm of the exact solution).

For the perturbed case we add to h either a constant perturbation with $L_2(I \times \omega)$ -norm equal to τ , or a random perturbation of the form $p = \tau \frac{\hat{p}}{\|\hat{p}\|_{L_2(I \times \omega)}}$, where $\hat{p}$ is a random function in X^δ with values in $[0, 1]$. We take $\tau = 0.01$. Figure 4 shows the numerical results.

In the unperturbed cases, the rates for $L_2(I \times \Omega)$ - and $H^1(I \times \Omega)$ -norms are equal to the best approximation rates in these norms.

8.3. Data-assimilation for the heat equation

For $\Omega = (0, 1)^d$, $I = (0, 1)$ and $\omega = (\frac{1}{4}, \frac{3}{4})^d$, given $(f, g) \in L_2(I; H^{-1}(\Omega)) \times L_2(I \times \omega)$, we consider the problem of finding u that solves the problem

$$\partial_t u - \Delta_x u = f \text{ on } I \times \Omega, \quad u|_{I \times \omega} = g,$$

which was discussed in Example 3.2. We consider the formulation of this problem as a first order system as analyzed in Section 7.3. Assuming $f \in L_2(I \times \Omega)$, for $\mathbf{u} = (u_1, \mathbf{u}_2) = (u, -\nabla_x u)$, it reads as $\tilde{A}\mathbf{u} := (\mathbf{u}_2 + \nabla_x u_1, \text{div } \mathbf{u}, \Gamma_{I \times \omega} u_1) = (0, f, g)$, where $\text{div } \mathbf{u} := \partial_t u_1 + \text{div}_x \mathbf{u}_2$ and $\mathbf{u} \in \tilde{X} := \{\mathbf{u} = (u_1, \mathbf{u}_2) \in L_2(I; H^1(\Omega)) \times L_2(I \times \Omega)^d : \text{div } \mathbf{u} \in L_2(I \times \Omega)\}$. Recall that we study this problem in two cases. Either we have no knowledge

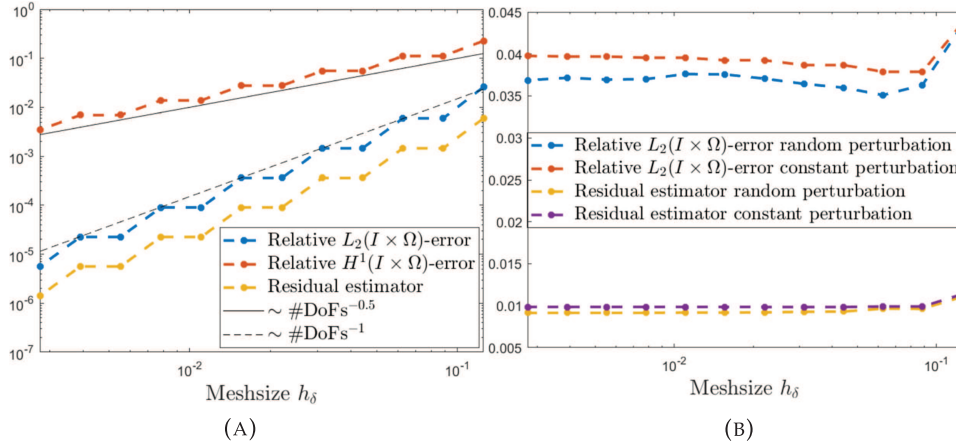


FIGURE 4. Data assimilation for the wave equation. Meshsize *vs.* relative error (or residual estimator). (A) Unperturbed data. (B) Perturbations, with $L_2(I \times \omega)$ -norm equal to 0.01, of datum h .

of u on $I \times \partial\Omega$ (Case (a)), or u is required to vanish on this lateral boundary (Case (b)). The latter problem is unconditionally stable. In this Case (b), the space $L_2(I; H^1(\Omega))$ should be read as $L_2(I; H_0^1(\Omega))$.

In view of the conditional or unconditional stability estimates (7.5) or (7.6), respectively, given a finite dimensional subspace $X^\delta \subset \tilde{X}$, with $\tilde{V} := L_2(I \times \Omega)^d \times L_2(I \times \Omega) \times L_2(I \times \omega)$ our approach is to minimize the least squares functional $\|\tilde{A}\mathbf{u} - (0, f, g)\|_{\tilde{V}}^2 + \varepsilon^2 \|u_1\|_{L_2(I \times \Omega)}^2$ over $\mathbf{u} \in X^\delta$, where in Case (b) the regularization term $\varepsilon^2 \|u_1\|_{L_2(I \times \Omega)}^2$ is omitted.

In our experiments, we prescribe the solution

$$u(t, x) = (t^3 + 1) \prod_{i=1}^d \sin(\pi x_i),$$

and define the data (f, g) correspondingly. For Case (a) the errors are measured in the relative $L_2((T_1, T_2); H^1(\tilde{\omega}))$ -norm, where we take $T_1 = \frac{1}{8}$, $T_2 = \frac{7}{8}$ and $\tilde{\omega} = (\frac{1}{8}, \frac{7}{8})^d$. Instead of recording the error in the $L_2((T_1, T); H_0^1(\Omega)) \cap H^1((T_1, T); H^{-1}(\Omega))$ -norm, which is hard to evaluate, we make use of the unconditional stability estimate (7.6) for Case (b), and provide the residual $\|\tilde{A}\mathbf{u} - (0, f, g)\|_{\tilde{V}}$ which, modulo a constant factor, is an upper bound for the aforementioned norm of the error. In addition we measure relative errors in the $L_2((T_1, T); H^1(\Omega))$ -norm for $T_1 = \frac{1}{8}$ which is easy to evaluate.

8.3.1. Unperturbed data, and $\Omega = (0, 1)$

We consider a sequence of uniform triangulations $(\mathcal{T}^\delta)_{\delta \in \Delta}$ of $\overline{I \times \Omega}$, where each next triangulation is created from its predecessor by one uniform newest vertex bisection starting from an initial triangulation that is created by cutting $I \times \Omega$ along both diagonals. The interior vertex in this initial triangulation is labelled as the “newest vertex” of all four triangles. We set $X^\delta := \mathcal{S}_{\mathcal{T}^\delta}^{0,1}(I \times \Omega) \times \mathcal{S}_{\mathcal{T}^\delta}^{0,1}(I \times \Omega)^d$ in Case (a), and $X^\delta := (\mathcal{S}_{\mathcal{T}^\delta}^{0,1}(I \times \Omega) \cap L_2(I; H_0^1(\Omega))) \times \mathcal{S}_{\mathcal{T}^\delta}^{0,1}(I \times \Omega)^d$ in Case (b).

Taking unperturbed data, we consider two strategies for choosing the regularization parameter ε in Case (a), namely $\varepsilon = 0$ and $\varepsilon = h_\delta$, the latter being the mesh-size. Since u is smooth, the choice $\varepsilon = h_\delta$ satisfies the conditions in (2.5). In Figure 5, we give the relative errors for both these choices of ε , and also give the relative error and residual estimator in Case (b). In the latter unconditionally stable case no regularization is applied.

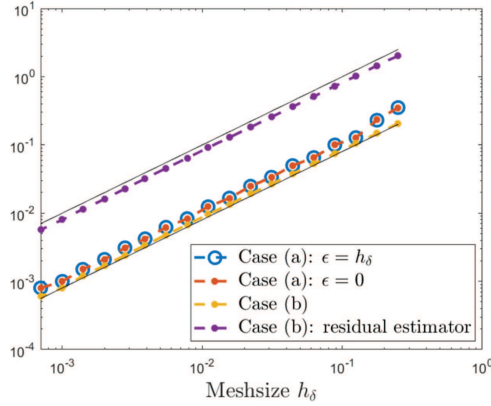


FIGURE 5. Data assimilation for the heat equation, and $\Omega = (0, 1)$. Meshsize *vs.* relative error (or residual estimator) in case of unperturbed data. Note that the norms in which the errors are measured are different for Cases (a) and (b). The asymptotic convergence rate in terms of # DoFs, indicated by the solid straight lines, is 0.5.

As in the case of the Cauchy problem for Poisson's equation, the numerical results in Figure 5 indicate that with unperturbed data regularization is not helpful. The rates for $L_2((\frac{1}{8}, \frac{7}{8}); H^1((\frac{1}{8}, \frac{7}{8})^d))$ - or $L_2((\frac{1}{8}, 1); H^1(0, 1))$ -norms are equal to the best approximation rates in these norms.

8.3.2. Randomly perturbed data, $\Omega = (0, 1)$

For $(X^\delta)_{\delta \in \Delta}$ as in Section 8.3.1, we now perturb the observational datum g with a random piecewise constant $p \in \mathcal{S}_{\mathcal{T}^\delta}^{-1,0}$ with $\|p\|_{L_2(\omega)} = \tau$. This p is constructed by normalizing a random function $\hat{p} \in \mathcal{S}_{\mathcal{T}^\delta}^{-1,0}$ and multiplying with τ . We considered the cases where $\hat{p}$ takes values in either $[0, 1]$ or $[-\frac{1}{2}, \frac{1}{2}]$.

For Case (a), we compare the results obtained with the regularization strategies $\varepsilon = \tau$, and $\varepsilon = \tau + h_\delta$, where the latter choice satisfies the conditions in (2.5), with those obtained with the experimentally found ε that minimizes $\|u - u_\varepsilon^\delta\|_{L_2(I; H^1(\Omega))}$. In addition, we present the results obtained for Case (b). The results are shown in Figure 6.

For both Cases (a) and (b), the solution is much more sensitive to random perturbations with mean $\tau/2$ than to those with mean 0. In Case (a) regularization is helpful for perturbations with mean $\tau/2$, but it is not when the mean is 0.

8.3.3. Unperturbed data, and $\Omega = (0, 1)^2$

We now consider the data-assimilation problem described in Section 8.3 for unperturbed data and the two-dimensional spatial domain $\Omega = (0, 1)^2$. We consider a sequence of conforming partitions $(\mathcal{T}^\delta)_{\delta \in \Delta}$ of $\bar{I} \times \bar{\Omega}$ into tetrahedra, where each partition consists of h_δ^{-3} cubes with sidelength h_δ that are decomposed into 6 tetrahedra using the Kuhn splitting. Since with our mesh-sizes and linear finite elements we could not clearly observe convergence in Case (a), we take quadratic elements, *i.e.*, we set $X^\delta := \mathcal{S}_{\mathcal{T}^\delta}^{0,2}(I \times \Omega) \times \mathcal{S}_{\mathcal{T}^\delta}^{0,2}(I \times \Omega)^2$ in Case (a), and $X^\delta := (\mathcal{S}_{\mathcal{T}^\delta}^{0,2}(I \times \Omega) \cap L_2(I; H_0^1(\Omega))) \times \mathcal{S}_{\mathcal{T}^\delta}^{0,2}(I \times \Omega)^2$ in Case (b).

We consider regularization parameters $\varepsilon = 0$ and $\varepsilon = h_\delta^2$ for Case (a), where the latter satisfies the conditions in (2.5), and for Case (b) apply no regularization. The results are given in Figure 7.

9. CONCLUSION

We have constructed a least squares solver for general conditionally stable ill-posed PDEs. For this solver it was demonstrated that, for a suitable regularization parameter, the error in the numerical approximation is

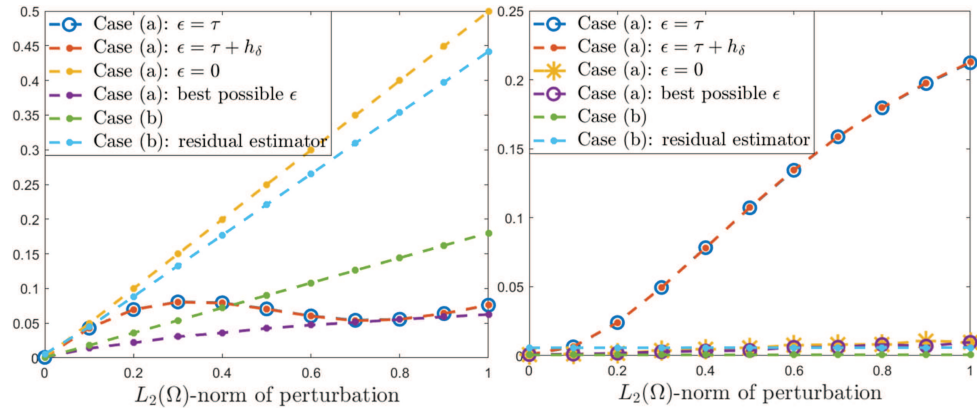


FIGURE 6. Data assimilation for the heat equation, and $\Omega = (0, 1)$. Norm of perturbation *vs.* relative error (or residual estimator). Results for random perturbation of the observational datum g , for a fixed mesh with $\#$ DoFs $\approx 10^6$ for different choices of ϵ and τ . *Left*: the case where $\hat{p}$ takes values in $[0, 1]$. *Right*: the case where $\hat{p}$ takes values in $[-1/2, 1/2]$.

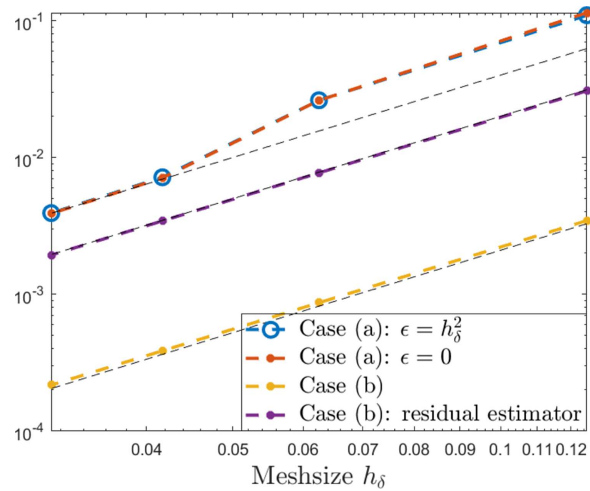


FIGURE 7. Data assimilation for the heat equation, and $\Omega = (0, 1)^2$. Meshsize *vs.* relative error (or residual estimator) in case of unperturbed data. The asymptotic convergence rate in terms of $\#$ DoFs indicated by the dashed straight lines is 1.

qualitatively the best that can be expected in view of the conditional stability estimate. In applications the least squares functional to be minimized involves negative and/or fractional Sobolev norms of residuals. It was shown that these norms can be replaced by computable quantities without compromising any of the attractive theoretical properties of the method.

The theoretical results were illustrated by numerical experiments for Poisson's equation with Cauchy data, and data-assimilation problems for both heat and wave-equation. In several examples the bounds on the error in the numerical approximation that were derived using the conditional stability estimates were pessimistic, which is not surprising since these estimates cover worst case settings. Similarly, it turns out that in many cases

better results were obtained by applying a smaller regularization parameter than predicted by the theoretical estimates.

Acknowledgements. This research has been supported in part by the NSF Grant DMS ID 2012469, by the SmartState and Williams-Hedberg Foundation, by the SFB 1481, funded by the German Research Foundation, and by the Netherlands Organization for Scientific Research (NWO) under contract. no. SH-208-11. We acknowledge the support of SURF (www.surf.nl) in using the National Supercomputer Snellius.

REFERENCES

- [1] G. Alessandrini, L. Rondi, E. Rosset and S. Vessella, The stability for the Cauchy problem for elliptic equations. *Inverse Prob.* **25** (2009) 123004.
- [2] M. Arioli and D. Loghin, Discrete interpolation norms with applications. *SIAM J. Numer. Anal.* **47** (2009) 2924–2951.
- [3] C. Bardos, Problèmes aux limites pour les équations aux dérivées partielles du premier ordre à coefficients réels; théorèmes d’approximation; application à l’équation de transport. *Ann. Sci. École Norm. Sup.* **3** (1970) 185–233.
- [4] C. Bardos, G. Lebeau and J. Rauch, Sharp sufficient conditions for the observation, control, and stabilization of waves from the boundary. *SIAM J. Control Optim.* **30** (1992) 1024–1065.
- [5] E. Bécache, L. Bourgeois, L. Franceschini and J. Dardé, Application of mixed formulations of quasi-reversibility to solve ill-posed problems for heat and wave equations: the 1D case. *Inverse Prob. Imaging* **9** (2015) 971–1002.
- [6] C. Bernardi and G. Raugel, Analysis of some finite elements for the Stokes problem. *Math. Comp.* **44** (1985) 71–79.
- [7] P.B. Bochev and M.D. Gunzburger, Least-Squares Finite Element Methods. Vol. 166 of *Applied Mathematical Sciences*. Springer, New York (2009).
- [8] L. Bourgeois and L. Chesnel, On quasi-reversibility solutions to the Cauchy problem for the Laplace equation: regularity and error estimates. *ESAIM Math. Model. Numer. Anal.* **54** (2020) 493–529.
- [9] L. Bourgeois and A. Recoquillay, A mixed formulation of the Tikhonov regularization and its application to inverse PDE problems. *ESAIM Math. Model. Numer. Anal.* **52** (2018) 123–145.
- [10] J.H. Bramble, J.E. Pasciak and J. Xu, Parallel multilevel preconditioners. *Math. Comput.* **55** (1990) 1–22.
- [11] E. Burman, Stabilised finite element methods for ill-posed problems with conditional stability, in Building Bridges: Connections and Challenges in Modern Approaches to Numerical Partial Differential Equations. Vol. 114 of *Lect. Notes Comput. Sci. Eng.* Springer, Cham (2016) 93–127.
- [12] E. Burman, The elliptic Cauchy problem revisited: control of boundary data in natural norms. *C. R. Math. Acad. Sci. Paris* **355** (2017) 479–484.
- [13] E. Burman and L. Oksanen, Data assimilation for the heat equation using stabilized finite element methods. *Numer. Math.* **139** (2018) 505–528.
- [14] E. Burman, P. Hansbo and M.G. Larson, Solving ill-posed control problems by stabilized finite element methods: an alternative to Tikhonov regularization. *Inverse Prob.* **34** (2018) 035004.
- [15] E. Burman, J. Ish-Horowicz and L. Oksanen, Fully discrete finite element data assimilation method for the heat equation. *ESAIM Math. Model. Numer. Anal.* **52** (2018) 2065–2082.
- [16] E. Burman, A. Feizmohammadi and L. Oksanen, A finite element data assimilation method for the wave equation. *Math. Comput.* **89** (2020) 1681–1709.
- [17] E. Burman, A. Feizmohammadi, A. Münch and L. Oksanen, Space time stabilized finite element methods for a unique continuation problem subject to the wave equation. *ESAIM Math. Model. Numer. Anal.* **55** (2021) S969–S991.
- [18] E. Burman, A. Feizmohammadi, A. Münch and L. Oksanen, Spacetime finite element methods for control problems subject to the wave equation. Preprint [arXiv:2109.07890](https://arxiv.org/abs/2109.07890) (2021).
- [19] E. Burman, G. Delay and A. Ern, The unique continuation problem for the heat equation discretized with a high-order space-time nonconforming method. [hal 03720960](https://arxiv.org/abs/2203.03720) (2022).
- [20] N. Burq, Contrôle de l’équation des ondes dans des ouverts comportant des coins. *Bull. Soc. Math. France* **126** (1998) 601–637.
- [21] W. Dahmen, R. Stevenson and J. Westerdiep, Accuracy controlled data assimilation for parabolic problems. *Math. Comput.* **91** (2022) 557–595.
- [22] T. Führer, Multilevel decompositions and norms for negative order Sobolev spaces. *Math. Comput.* **91** (2021) 183–218.
- [23] T. Führer and M. Karkulik, Space-time least-squares finite elements for parabolic equations. *Comput. Math. Appl.* **92** (2021) 27–36.
- [24] S. Funken, D. Praetorius and P. Wissgott, Efficient implementation of adaptive P1-FEM in Matlab. *Comput. Methods Appl. Math.* **11** (2011) 460–490.
- [25] G. Gantner and R.P. Stevenson, Further results on a space-time FOSLS formulation of parabolic PDEs. *ESAIM Math. Model. Numer. Anal.* **55** (2021) 283–299.
- [26] O. Imanuvilov and M. Yamamoto, Conditional stability in a backward parabolic system. *Appl. Anal.* **93** (2014) 2174–2198.
- [27] V. Isakov, Inverse Problems for Partial Differential Equations. Vol. 127 of *Applied Mathematical Sciences*, 2nd edition. Springer, New York (2006).

- [28] M.V. Klibanov, Estimates of initial conditions of parabolic equations and inequalities via lateral Cauchy data. *Inverse Prob.* **22** (2006) 495–514.
- [29] I. Lasiecka, J.-L. Lions and R. Triggiani, Nonhomogeneous boundary value problems for second order hyperbolic operators. *J. Math. Pures Appl.* **65** (1986) 149–192.
- [30] J. Le Rousseau, G. Lebeau, P. Terpolilli and E. Trélat, Geometric control condition for the wave equation with a time-dependent observation domain. *Anal. PDE* **10** (2017) 983–1015.
- [31] J.-L. Lions and E. Magenes, Non-homogeneous Boundary Value Problems and Applications. Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 181. Vol. I. Springer-Verlag, New York-Heidelberg (1972).
- [32] M. Page and D. Praetorius, Convergence of adaptive FEM for some elliptic obstacle problem. *Appl. Anal.* **92** (2013) 595–615.
- [33] J. Schöberl, *C++11 implementation of finite elements in ngsolve*. Technical report, Institute for Analysis and Scientific Computing, Vienna University of Technology (2014).
- [34] L.R. Scott and S. Zhang, Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comput.* **54** (1990) 483–493.
- [35] R.P. Stevenson and R. van Venetië, Uniform preconditioners for problems of negative order. *Math. Comput.* **89** (2020) 645–674.
- [36] R.P. Stevenson and R. van Venetië, Uniform preconditioners for problems of positive order. *Comput. Math. Appl.* **79** (2020) 3516–3530.
- [37] R.P. Stevenson and R. van Venetië, Uniform preconditioners of linear complexity for problems of negative order. *Comput. Methods Appl. Math.* **21** (2021) 469–478.
- [38] R.P. Stevenson and J. Westerdiep, Minimal residual space-time discretizations of parabolic equations: asymmetric spatial operators. *Comput. Math. Appl.* **101** (2021) 107–118.
- [39] R.P. Stevenson and J. Westerdiep, Stability of Galerkin discretizations of a mixed space-time variational formulation of parabolic evolution equations. *IMA J. Numer. Anal.* **41** (2021) 28–47.
- [40] A.N. Tikhonov, On the stability of inverse problems. *C. R. (Doklady) Acad. Sci. URSS (N.S.)* **39** (1943) 176–179.



Please help to maintain this journal in open access!

This journal is currently published in open access under the Subscribe to Open model (S2O). We are thankful to our subscribers and supporters for making it possible to publish this journal in open access in the current year, free of charge for authors and readers.

Check with your library that it subscribes to the journal, or consider making a personal donation to the S2O programme by contacting subscribers@edpsciences.org.

More information, including a list of supporters and financial transparency reports, is available at <https://edpsciences.org/en/subscribe-to-open-s2o>.