Technische Universität Graz

Adaptive least-squares space-time finite element methods

C. Köthe, R. Löscher, O. Steinbach

Berichte aus dem
Institut für Angewandte Mathematik

Bericht 2023/6

Technische Universität Graz

Adaptive least-squares space-time finite element methods

C. Köthe, R. Löscher, O. Steinbach

Berichte aus dem
Institut für Angewandte Mathematik

Bericht 2023/6

Technische Universität Graz
Institut für Angewandte Mathematik
Steyrergasse 30
A 8010 Graz

WWW: http://www.applied.math.tugraz.at

Adaptive least-squares space-time finite element methods

Christian Köthe, Richard Löscher, Olaf Steinbach

Institut für Angewandte Mathematik, TU Graz,

Steyrergasse 30, 8010 Graz, Austria

c.koethe@tugraz.at, loescher@math.tugraz.at, o.steinbach@tugraz.at

Abstract

We consider the numerical solution of an abstract operator equation $Bu = f$ by using a least-squares approach. We assume that $B : X \to Y^*$ is an isomorphism, and that $A : Y \to Y^*$ implies a norm in Y, where X and Y are Hilbert spaces. The minimizer of the least-squares functional $\frac{1}{2}\,\|Bu - f\|^2_{A^{-1}}$, i.e., the solution of the operator equation, is then characterized by the gradient equation $Su = B^*A^{-1}f$ with an elliptic and self-adjoint operator $S := B^*A^{-1}B : X \to X^*$. When introducing the adjoint $p = A^{-1}(f - Bu)$ we end up with a saddle point formulation to be solved numerically by using a mixed finite element method. Based on a discrete inf-sup stability condition we derive related a priori error estimates. While the adjoint p is zero by construction, its approximation p_h serves as a posteriori error indicator to drive an adaptive scheme when discretized appropriately. While this approach can be applied to rather general equations, here we consider second order linear partial differential equations, including the Poisson equation, the heat equation, and the wave equation, in order to demonstrate its potential, which allows to use almost arbitrary space-time finite element methods for the adaptive solution of time-dependent partial differential equations.

Keywords: Least-squares methods, space-time finite element methods, a posteriori error indicator, adaptivity, Poisson equation, heat equation, wave equation

1 Introduction

The use of Galerkin finite element methods for the numerical solution of partial differential equations is well established, see, e.g., the text books [11, 39], and many others. However, for time dependent problems it is a common procedure to first discretize the spatial part using a finite element method, and then applying either a time stepping method, e.g., an explicit or implicit Euler scheme, see, e.g., the monograph [45], or discontinuous Galerkin methods in time, see, e.g., [33], and the references given therein. Recently, the interest of discretizing space and time at once has been rising, resulting in so called space-time discretization methods, see, e.g., [26]. Although, the discretization in space and time leads to larger systems of algebraic equations to be solved, these methods bring the advantage of having full control of the discretization in space and time simultaneously, allowing for space-time adaptivity. Moreover, space-time methods offer more flexibility in the construction of efficient solvers than time stepping methods, since preconditioning and parallelization in the space-time domain is applicable, and mandatory, see, e.g., [22, 29] for space-time solvers in the case of parabolic equations. The derivation of space-time formulations usually results in Petrov-Galerkin schemes, see, e.g., [1, 38, 40, 46] in the case of the heat equation, where for the numerical treatment it is crucial to establish a related discrete inf-sup stability condition. This becomes even more involved in the case of the wave equation where a CFL condition is required, e.g., [42]. Possible approaches to overcome such a restriction is the use of discontinuous Galerkin methods, e.g., [17, 33, 34], or introducing a suitable transformation operator such as the (modified) Hilbert transformation [30, 35, 42]. Another approach is to replace the direct variational formulation by a least-squares/minimal residual equation. This has been studied in the context of first order least-squares systems in, e.g., [19, 20, 21, 36], and in the context of minimal residual Petrov–Galerkin discretizations in [1, 44], just to mention a few. For the latter approach it is well known that the Galerkin discretization results in a mixed system, where the second variable is the Riesz lift of the residual of the primal variable. This is also the point of view we will take.

From the well-established theory of least-squares methods, with a tremendous overview by Bochev and Gunzberger [6], see also [4] and [3, 5], least-squares formulations of second order partial differential equations come with the advantage of offering an error estimator for free, but also double the degrees of freedom. To apply the theory, it is of main interest to consider so called practical methods, that allow to measure the residual in a localizable norm, i.e., in a non-negative Sobolev norm allowing at least L^2 regularity. Trimmed to this interest, Führer and Karkulik introduced a first order system least-squares method (FOSLS) for parabolic problems [20], showing the applicability and power of least-squares methods in the space-time setting. However, the reformulation as a first order system comes with the fact that one has to assume a higher regularity on the source term in order to show convergence of the method. This is unnatural to a certain extent, since source terms of partial differential equations are usually considered as functionals acting on the test space and thus belong to Sobolev spaces of negative order. At this point we want to mention that recently in [19, 21], FOSLS with a source term in a negative order Sobolev space were analyzed by replacing the load by its finite element approximation.

In this paper, we will formulate and analyze a least-squares method for the solution of abstract operator equations, that overcomes this problem, i.e., we will be able to phrase the method, even when the source is of minimal regularity. For X and Y being Sobolev spaces of non-negative order, we will consider the problem to find $u \in X$ such that $Bu = f \in Y^*$, where $B : X \to Y^*$ is an isomorphism. Then the problem is equivalent to minimize the residual $\frac{1}{2}\|Bv - f\|^2_{Y^*}$ over all $v \in X$. As Y^* is the dual space of Y, the residual is measured in a Sobolev norm with negative index. The main idea is now based on lifting the negative order Sobolev norm using a bijective operator $A : Y \to Y^*$ in order to keep the method practical. Then the problem is equivalent to minimize the functional $\frac{1}{2}\|Bv - f\|^2_{A^{-1}}$, for which the minimizer $u \in X$ is characterized as the unique solution of the gradient equation $B^*A^{-1}(Bu - f) = 0$. Introducing the auxiliary variable $p = A^{-1}(f - Bu) \in Y$ we end up with solving a saddle point formulation. This saddle point point problem is of similar shape as those obtained from optimal control with energy regularization, see [27, 28, 31]. The discretization follows by standard means, using conforming finite element spaces of lowest order. To ensure uniqueness, we will need a discrete inf-sup stability condition. Though, $p \equiv 0$ on the continuous level, for the discrete Lagrange multiplier in general it holds that $p_h \neq 0$. Thus, it can be used to define an a posteriori error estimator which is shown to be efficient and reliable using a so called saturation assumption [13, 16]. The idea of lifting the norm was already considerd in [7, 8] for elliptic partial differential equations, or more recently in [32] for a larger class of problems, and in [14] adding a stabilization term when considering ill-posed problems. In [1], this concept was already applied in the context of space-time methods to parabolic equations, with a discretization on a tensor product mesh. The novelty of our method is the application to space-time problems on completely unstructured meshes, and the practicability even when using negative Sobolev norms as well as the proof of an efficient and reliable error estimator in this setting. Moreover, the use of completely unstructured space-time meshes allows for full space-time adaptivity, and later on, for a parallel solution of the resulting linear systems of algebraic equations.

This paper is organized as follows. In Section 2 we consider the least-squares approach for the solution of an operator equation $Bu = f$ in an abstract sense. We derive the saddle point formulation and show ellipticity of the related operator $S = B^*A^{-1}B$. Based on a discrete inf-sup condition we show unique solvability of the discrete system and derive related a priori error estimates. Using a so called saturation assumption we are able to prove that the use of the dual p_h as error estimator is efficient and reliable. In Section 3 we apply the approach to the Poisson equation, showing that it is also related to the well known $h - \frac{h}{2}$ estimator [12, 18]. Section 4 deals with the application of the approach to the heat equation, while the wave equation is considered in Section 5. In all cases we provide numerical examples to illustrate our theoretical findings. Finally, in Section 6 we give a conclusion, where we mention possible extensions and further work which needs to be done.

2 Abstract setting

Let $X \subset H \subset X^*$ and $Y \subset H \subset Y^*$ be Gelfand triples of Hilbert spaces, where X^*, Y^* are the duals of X, Y with respect to H and with the duality pairing $\langle f, q\rangle_H$ for $f \in Y^*$ and $q \in Y$. Let $A : Y \to Y^*$ be a bounded linear operator which is assumed to be self-adjoint and elliptic in Y. Therefore, A implies a norm in Y, i.e., $\|q\|_Y := \sqrt{\langle Aq, q\rangle_H}$ for $q \in Y$. For $f \in Y^*$, the norm is given by duality,

$$\|f\|_{Y^*} := \sup_{0 \neq q \in Y} \frac{\langle f, q\rangle_H}{\|q\|_Y}\,. \tag{2.1}$$

Lemma 2.1 *The dual norm* (2.1) *allows the representations*

$$\|f\|_{Y^*}^2 = \|p_f\|_Y^2 = \langle Ap_f, p_f\rangle_H = \langle f, p_f\rangle_H = \langle A^{-1}f, f\rangle_H, \tag{2.2}$$

where $p_f \in Y$ is the unique solution of the variational formulation

$$\langle Ap_f, q\rangle_H = \langle f, q\rangle_H \quad \text{for all } q \in Y. \tag{2.3}$$

Proof. Since $A : Y \to Y^*$ is bounded and elliptic, $p_f = A^{-1}f$ is well defined as unique solution of the variational formulation (2.3). From the definition (2.1), we first have, note that $\|f\|_{Y^*} > 0$ implies $\|p_f\|_Y > 0$,

$$\|f\|_{Y^*} = \sup_{0\neq q\in Y} \frac{\langle f, q\rangle_H}{\|q\|_Y} \geq \frac{\langle f, p_f\rangle_H}{\|p_f\|_Y}, \quad \text{i.e.,} \quad \langle f, p_f\rangle_H \leq \|f\|_{Y^*}\|p_f\|_Y \,.$$

Hence we obtain

$$\|p_f\|_Y^2 = \langle Ap_f, p_f\rangle_H = \langle f, p_f\rangle_H \leq \|f\|_{Y^*}\|p_f\|_Y, \quad \text{i.e.,} \quad \|p_f\|_Y \leq \|f\|_{Y^*} \,.$$

On the other hand, (2.1) implies, when using (2.3),

$$\|f\|_{Y^*} = \sup_{0\neq q\in Y} \frac{\langle f, q\rangle_H}{\|q\|_Y} = \sup_{0\neq q\in Y} \frac{\langle Ap_f, q\rangle_H}{\|q\|_Y} \leq \|p_f\|_Y \,,$$

and hence, $\|f\|_{Y^*} = \|p_f\|_Y$ follows. With this we finally obtain

$$\|f\|_{Y^*}^2 = \|p_f\|_Y^2 = \langle Ap_f, p_f\rangle_H = \langle f, p_f\rangle_H = \langle f, A^{-1}f\rangle_H.$$

∎

Let $B : X \to Y^*$ be a bounded linear operator which satisfies an inf-sup condition, i.e., there exist positive constants c_1^B and c_2^B such that

$$\|Bv\|_{Y^*} \leq c_2^B \, \|v\|_X, \quad \sup_{0\neq q\in Y} \frac{\langle Bv, q\rangle_H}{\|q\|_Y} \geq c_1^B \, \|v\|_X \quad \text{for all } v \in X. \tag{2.4}$$

In addition, we assume that B is surjective. Then, $B : X \to Y^*$ is an isomorphism. Due to the assumptions made we conclude unique solvability of the operator equation to find $u \in X$ such that $Bu = f$ in Y^* is satisfied. The solution of the operator equation $Bu = f$ in Y^* is equivalent to the minimization of a quadratic functional for $v \in X$,

$$\begin{aligned} \mathcal{J}(v) &= \frac{1}{2}\,\|Bv - f\|_{Y^*}^2 = \frac{1}{2}\,\langle A^{-1}(Bv - f), Bv - f\rangle_H \\ &= \frac{1}{2}\,\langle B^*A^{-1}Bv, v\rangle_H - \langle B^*A^{-1}f, v\rangle_H + \frac{1}{2}\,\langle A^{-1}f, f\rangle_H, \end{aligned}$$

whose minimizer $u \in X$ is given as solution of the gradient equation

$$B^*A^{-1}(Bu - f) = 0 \,,$$

i.e., we have to solve the operator equation

$$Su := B^*A^{-1}Bu = B^*A^{-1}f\,. \tag{2.5}$$

Note that $B^* : Y \to X^*$ is the adjoint of $B : X \to Y^*$, i.e.,

$$\langle B^*q, v\rangle_H := \langle q, Bv\rangle_H \quad \text{for all } v \in X,\ q \in Y,$$

satisfying

$$\|B^*q\|_{X^*} = \sup_{0\neq v\in X} \frac{\langle B^*q, v\rangle_H}{\|v\|_X} = \sup_{0\neq v\in X} \frac{\langle q, Bv\rangle_H}{\|v\|_X} \leq \sup_{0\neq v\in X} \frac{\|q\|_Y\|Bv\|_{Y^*}}{\|v\|_X} \leq c_2^B\,\|q\|_Y$$

for all $q \in Y$.

Lemma 2.2 *The operator $S := B^*A^{-1}B : X \to X^*$ is bounded and elliptic, i.e.,*

$$\|Su\|_{X^*} \leq c_2^S\,\|u\|_X, \quad \langle Su, u\rangle_H \geq c_1^S\,\|u\|_X^2 \quad \textit{for all}\, u \in X,$$

where $c_2^S = [c_2^B]^2$, $c_1^S = [c_1^B]^2$.

Proof. For $u \in X$ we first have, using (2.2) for $f = Bu$,

$$\begin{aligned}\|Su\|_{X^*} &= \sup_{0\neq v\in X} \frac{\langle Su, v\rangle_H}{\|v\|_X} = \sup_{0\neq v\in X} \frac{\langle A^{-1}Bu, Bv\rangle_H}{\|v\|_X} \\ &\leq \sup_{0\neq v\in X} \frac{\|A^{-1}Bu\|_Y\|Bv\|_{Y^*}}{\|v\|_X} \leq \sup_{0\neq v\in X} \frac{\|Bu\|_{Y^*}\|Bv\|_{Y^*}}{\|v\|_X} \leq [c_2^B]^2\,\|u\|_X\,.\end{aligned}$$

In addition, we define $p_u = A^{-1}Bu \in Y$ to obtain

$$\|p_u\|_Y^2 = \langle Ap_u, p_u\rangle_H = \langle B^*A^{-1}Bu, u\rangle_H = \langle Su, u\rangle_H.$$

From the inf-sup stability condition in (2.4) we then conclude

$$c_1^B\,\|u\|_X \leq \sup_{0\neq v\in Y} \frac{\langle Bu, v\rangle_H}{\|v\|_Y} = \sup_{0\neq v\in Y} \frac{\langle Ap_u, v\rangle_H}{\|v\|_Y} \leq \|p_u\|_Y,$$

i.e.,

$$[c_1^B]^2\,\|u\|_X^2 \leq \|p_u\|_Y^2 \;=\; \langle Su, u\rangle_H.$$

■

In fact, $S := B^*A^{-1}B : X \to X^*$ defines an equivalent norm in X,

$$\|v\|_S := \sqrt{\langle Sv, v\rangle_H} = \sqrt{\langle A^{-1}Bv, Bv\rangle_H} = \|Bv\|_{Y^*}\,,$$

satisfying

$$c_1^B\,\|v\|_X \leq \|v\|_S \leq c_2^B\,\|v\|_X \quad \text{for all } v \in X. \tag{2.6}$$

The variational formulation of the operator equation (2.5) is to find $u \in X$ such that

$$\langle Su, v\rangle_H = \langle B^* A^{-1} f, v\rangle_H \quad \text{for all } v \in X, \tag{2.7}$$

which is uniquely solvable for all $f \in Y^*$.

Let $X_H = \text{span}\{\varphi_k\}_{k=1}^{M_X} \subset X$ be some finite dimensional ansatz space which is defined with respect to some admissible decomposition of the computational domain into shape-regular simplicial finite elements of mesh size H, see, e.g., [11, 39]. Then the Galerkin variational formulation of (2.7) is to find $u_H \in X_H$ such that

$$\langle Su_H, v_H\rangle_H = \langle B^* A^{-1} f, v_H\rangle_H \quad \text{for all } v_H \in X_H. \tag{2.8}$$

Using standard arguments, we conclude Cea's lemma,

$$\|u - u_H\|_S \leq \inf_{v_H \in X_H} \|u - v_H\|_S\,, \tag{2.9}$$

and convergence $u_H \to u$ in X follows from an approximation property of X_H.

Since the operator $S = B^* A^{-1} B$ does not allow a direct evaluation in general, we have to define a suitable approximation. For $u \in X$ we write $Su = B^* A^{-1} Bu = B^* p_u$, where $p_u = A^{-1} Bu$. In fact, $p_u \in Y$ is the unique solution of the variational formulation

$$\langle Ap_u, q\rangle_H = \langle Bu, q\rangle_H \quad \text{for all } q \in Y.$$

Let $Y_h = \text{span}\{\psi_i\}_{i=1}^{M_Y} \subset Y$ be a finite dimensional ansatz space which is defined with respect to some mutually different decomposition of the computational domain into finite elements of mesh size h. Then we define $p_{uh} \in Y_h$ as unique solution of the Galerkin variational formulation

$$\langle Ap_{uh}, q_h\rangle_H = \langle Bu, q_h\rangle_H \quad \text{for all } q_h \in Y_h,$$

and we define $\widetilde{S}u := B^* p_{uh}$ as approximation of $Su = B^* p_u$. We immediately have the bounds

$$\|p_{uh}\|_Y^2 = \langle Ap_{uh}, p_{uh}\rangle_H = \langle Bu, p_{uh}\rangle_H \leq \|Bu\|_{Y^*} \|p_{uh}\|_Y \leq c_2^B\, \|u\|_X \|p_{uh}\|_Y,$$

i.e.,

$$\|p_{uh}\|_Y \leq c_2^B\, \|u\|_X,$$

and

$$\|\widetilde{S}u\|_{X^*} = \|B^* p_{uh}\|_{X^*} \leq c_2^B\, \|p_{uh}\|_Y \leq [c_2^B]^2\, \|u\|_X, \tag{2.10}$$

as well as the error estimate

$$\|(S - \widetilde{S})u\|_{X^*} = \|B^*(p_u - p_{uh})\|_{X^*} \leq c_2^B\, \|p_u - p_{uh}\|_Y \leq c_2^B \inf_{q_h \in Y_h} \|p_u - q_h\|_Y\,. \tag{2.11}$$

In the same way we define $p_{fh} \in Y_h$ as unique solution of the variational formulation

$$\langle Ap_{fh}, q_h\rangle_H = \langle f, q_h\rangle_H \quad \text{for all } q_h \in Y_h,$$

in order to define $B^* p_{fh}$ as approximation of $B^* p_f = B^* A^{-1} f$, i.e., $p_f = A^{-1} f$. Hence, instead of (2.8) we now consider the perturbed variational formulation to find $\widetilde{u}_H \in X_H$ such that

$$\langle \widetilde{S} \widetilde{u}_H, v_H \rangle_H = \langle B^* p_{fh}, v_H \rangle_H \quad \text{for all } v_H \in X_H. \tag{2.12}$$

To ensure unique solvability of (2.12), we assume the discrete inf-sup stability condition

$$c_S \, \|v_H\|_X \leq \sup_{0 \neq q_h \in Y_h} \frac{\langle B v_H, q_h \rangle_H}{\|q_h\|_Y} \quad \text{for all } v_H \in X_H. \tag{2.13}$$

Lemma 2.3 *Assume the discrete inf-sup stability condition* (2.13) *to be satisfied. Then the approximate operator $\widetilde{S}$ is discrete elliptic in X_H, i.e.,*

$$\langle \widetilde{S} v_H, v_H \rangle_H \geq [c_S]^2 \, \|v_H\|_X^2 \quad \textit{for all } v_H \in X_H. \tag{2.14}$$

Proof. We first note that $\widetilde{S} v_H = B^* p_{v_H h}$ where $p_{v_H h} \in Y_h$ solves

$$\langle A p_{v_H h}, q_h \rangle_H = \langle B v_H, q_h \rangle_H \quad \text{for all } q_h \in Y_h.$$

Hence we obtain

$$\langle \widetilde{S} v_H, v_H \rangle_H = \langle B^* p_{v_H h}, v_H \rangle_H = \langle B v_H, p_{v_H h} \rangle_H = \langle A p_{v_H h}, p_{v_H h} \rangle_H = \|p_{v_H h}\|_Y^2 \, .$$

On the other hand, the discrete inf-sup stability condition (2.13) implies

$$c_S \, \|v_H\|_X \leq \sup_{0 \neq q_h \in Y_h} \frac{\langle B v_H, q_h \rangle_H}{\|q_h\|_Y} = \sup_{0 \neq q_h \in Y_h} \frac{\langle A p_{v_H h}, q_h \rangle_H}{\|q_h\|_Y} \leq \|p_{v_H h}\|_Y,$$

and hence,

$$c_S^2 \, \|v_H\|_X^2 \; \leq \; \|p_{v_H h}\|_Y^2 \; = \; \langle \widetilde{S} v_H, v_H \rangle_H$$

follows. ∎

The discrete ellipticity estimate (2.14) ensures not only unique solvability of the perturbed variational formulation (2.12), using the Strang lemma, see, e.g., [11, 39], we can prove the following error estimate.

Lemma 2.4 *For the unique solution $\widetilde{u}_H \in X_H$ of the perturbed variational formulation* (2.12) *there holds the error estimate*

$$\begin{aligned} \|u - \widetilde{u}_H\|_X \;\; &\leq \;\; \left(1 + 2 \, \frac{[c_2^B]^2}{[c_S]^2} \right) \frac{c_2^B}{c_1^B} \inf_{v_H \in X_H} \|u - v_H\|_X \\ &\qquad + \frac{c_2^B}{[c_S]^2} \left[\inf_{q_h \in Y_h} \|p_u - q_h\|_Y + \inf_{q_h \in Y_h} \|p_f - q_h\|_Y \right]. \end{aligned} \tag{2.15}$$

Proof. When considering the difference of the variational formulations (2.8) and (2.12), this gives

$$\langle Su_H - \widetilde{S}\widetilde{u}_H, v_H\rangle_H = \langle B^*(p_f - p_{fh}), v_H\rangle_H \quad \text{for all } v_H \in X_H.$$

From the discrete ellipticity (2.14) we then conclude

$$\begin{aligned}
[c_S]^2 \, \|u_H - \widetilde{u}_H\|_X^2 &\leq \langle \widetilde{S}(u_H - \widetilde{u}_H), u_H - \widetilde{u}_H\rangle_H \\
&= \langle (\widetilde{S} - S)u_H, u_H - \widetilde{u}_H\rangle_H + \langle B^*(p_f - p_{fh}), u_H - \widetilde{u}_H\rangle_H \\
&\leq \|(\widetilde{S} - S)u_H\|_{X^*}\|u_H - \widetilde{u}_H\|_X + c_2^B \, \|p_f - p_{fh}\|_Y \|u_H - \widetilde{u}_H\|_X,
\end{aligned}$$

i.e.,

$$[c_S]^2 \, \|u_H - \widetilde{u}_H\|_X \leq \|(\widetilde{S} - S)u_H\|_{X^*} + c_2^B \, \|p_f - p_{fh}\|_Y.$$

We further have, using Lemma 2.2 and (2.10),

$$\begin{aligned}
\|(\widetilde{S} - S)u_H\|_{X^*} &\leq \|(\widetilde{S} - S)u\|_{X^*} + \|(\widetilde{S} - S)(u - u_H)\|_{X^*} \\
&\leq \|(\widetilde{S} - S)u\|_{X^*} + 2\,[c_2^B]^2 \, \|u - u_H\|_X,
\end{aligned}$$

and by using the triangle inequality,

$$\|u - \widetilde{u}_H\|_X \leq \left(1 + 2\,\frac{[c_2^B]^2}{[c_S]^2}\right) \|u - u_H\|_X + \frac{1}{[c_S]^2}\Big[\|(\widetilde{S} - S)u\|_{X^*} + c_2^B \, \|p_f - p_{fh}\|_Y\Big].$$

The assertion now follows from the norm equivalence inequalites (2.6), Cea's lemma (2.9), and the error estimate (2.11). ■

The perturbed variational formulation (2.12) is, using $p_h := p_{fh} - p_{\widetilde{u}_H h} \in Y_h$, equivalent to the coupled variational formulation to find $(\widetilde{u}_H, p_h) \in X_H \times Y_h$ such that

$$\langle Ap_h, q_h\rangle_H + \langle B\widetilde{u}_H, q_h\rangle_H = \langle f, q_h\rangle_H, \quad \langle p_h, Bv_H\rangle_H = 0 \tag{2.16}$$

is satisfied for all $(v_H, q_h) \in X_H \times Y_h$. This is equivalent to the linear system of algebraic equations,

$$\begin{pmatrix} A_h & B_h \\ B_h^\top & \end{pmatrix} \begin{pmatrix} \underline{p} \\ \underline{u} \end{pmatrix} = \begin{pmatrix} \underline{f} \\ \underline{0} \end{pmatrix}, \tag{2.17}$$

where, for $i, j = 1, \ldots, M_Y$ and $k = 1, \ldots, M_X$,

$$A_h[j, i] = \langle A\psi_i, \psi_j\rangle_H, \quad B_h[j, k] = \langle B\varphi_k, \psi_j\rangle_H, \quad f_j = \langle f, \psi_j\rangle_H.$$

Since A_h is symmetric and positive definite, and hence invertible, we conclude the Schur complement system

$$B_h^\top A_h^{-1} B_h \underline{u} = B_h^\top A_h^{-1} \underline{f} \tag{2.18}$$

which is the matrix representation of the perturbed variational formulation (2.12). Note that the Schur complement matrix $S_h := B_h^\top A_h^{-1} B_h$ is symmetric and positive definite.

Moreover, we observe that (2.16) is the Galerkin formulation of the coupled variational formulation to find $(u, p) \in X \times Y$ such that

$$\langle Ap, q\rangle_H + \langle Bu, q\rangle_H = \langle f, q\rangle_H, \quad \langle p, Bv\rangle_H = 0 \tag{2.19}$$

is satisfied for all $(v, q) \in X \times Y$, i.e., of the coupled operator equation

$$Ap + Bu = f, \quad B^*p = 0\,. \tag{2.20}$$

Due to $p := A^{-1}(f - Bu)$, this is (2.5), where by construction we have $p \equiv 0$. If the Galerkin matrix B_h is invertible, so is $B_h^\top$, and hence $p_h \equiv 0$ follows in this particular case. But in general, the discrete inf-sup stability condition (2.13) involves spaces such that $\dim(X_H) \neq \dim(Y_h)$. Thus, B_h is a not a square matrix and hence not invertible, so that $p_h \in Y_h$ is not zero, and we can use p_h to define an a posteriori error indicator for $\|u - \widetilde{u}_H\|_X$.

Lemma 2.5 *Let* $(\widetilde{u}_H, p_h) \in X_H \times Y_h$ *be the unique solution of* (2.16). *Then,*

$$\|p_h\|_Y \leq \|u - \widetilde{u}_H\|_S \leq c_2^B\, \|u - \widetilde{u}_H\|_X\,. \tag{2.21}$$

Proof. When subtracting the Galerkin formulation (2.16) from (2.19) for $q = q_h \in Y_h \subset Y$ and $v = v_H \in X_H \subset X$, this gives the Galerkin orthogonalities

$$\langle A(p - p_h), q_h\rangle_H + \langle B(u - \widetilde{u}_H), q_h\rangle_H = 0, \quad \langle p - p_h, Bv_H\rangle_H = 0 \tag{2.22}$$

for all $(v_H, q_h) \in X_H \times Y_h$. In particular for $p \equiv 0$ this gives

$$\langle Ap_h, q_h\rangle_H = \langle B(u - \widetilde{u}_H), q_h\rangle_H, \quad \langle Bv_H, p_h\rangle_H = 0 \quad \text{for all } (v_H, q_h) \in X_H \times Y_h.$$

Hence, when choosing $q_h = p_h \in Y_h$, we further conclude

$$\|p_h\|_Y^2 = \langle Ap_h, p_h\rangle_H = \langle B(u - \widetilde{u}_H), p_h\rangle_H \leq \|B(u - \widetilde{u}_H)\|_{Y^*}\|p_h\|_Y,$$

i.e.,

$$\|p_h\|_Y \leq \|B(u - \widetilde{u}_H)\|_{Y^*} = \|u - \widetilde{u}_H\|_S,$$

and the assertion finally follows from (2.6). ■

While the upper estimate (2.21) shows the efficiency of the error estimator $\|p_h\|_Y$, reliability is more involved. For this we introduce an ansatz space $X_{\overline{H}} \subset X$ such that $X_H \subset X_{\overline{H}}$ is satisfied. As in (2.13) we assume the discrete inf-sup stability condition

$$\overline{c}_S\, \|v_{\overline{H}}\|_X \leq \sup_{0 \neq q_h \in Y_h} \frac{\langle Bv_{\overline{H}}, q_h\rangle_H}{\|q_h\|_Y} \quad \text{for all } v_{\overline{H}} \in X_{\overline{H}}. \tag{2.23}$$

Due to $X_H \subset X_{\overline{H}}$ we have that (2.13) is a direct consequence of (2.23). Using (2.23) we can determine $(\widetilde{u}_{\overline{H}}, \overline{p}_h) \in X_{\overline{H}} \times Y_h$ as unique solution of the variational formulation such that

$$\langle A\overline{p}_h, q_h\rangle_H + \langle B\widetilde{u}_{\overline{H}}, q_h\rangle_H = \langle f, q_h\rangle_H, \quad \langle \overline{p}_h, Bv_{\overline{H}}\rangle_H = 0 \tag{2.24}$$

is satisfied for all $(v_{\overline{H}}, q_h) \in X_{\overline{H}} \times Y_h$.

Lemma 2.6 *Let* $(\widetilde{u}_H, p_h) \in X_H \times Y_h$ *and* $(\widetilde{u}_{\overline{H}}, \overline{p}_h) \in X_{\overline{H}} \times Y_h$ *be the unique solutions of the Galerkin variational formulations* (2.16) *and* (2.24), *respectively. Assume the saturation assumption*

$$\|u - \widetilde{u}_{\overline{H}}\|_X \leq \eta \, \|u - \widetilde{u}_H\|_X \quad \text{for some } \eta \in (0,1). \tag{2.25}$$

Then the error estimator $\|p_h\|_Y$ *is reliable, satisfying*

$$\|u - \widetilde{u}_H\|_X \leq \frac{1}{1-\eta} \frac{c_2^B}{\overline{c}_S^2} \, \|p_h\|_Y \, . \tag{2.26}$$

Proof. Subtracting the Galerkin variational formulation (2.24) from (2.16), this gives the Galerkin orthogonality

$$\langle B(\widetilde{u}_{\overline{H}} - \widetilde{u}_H), q_h\rangle_H = \langle A(p_h - \overline{p}_h), q_h\rangle_H \quad \text{for all } q_h \in Y_h.$$

From the discrete inf-sup stability condition (2.23) we then conclude, recall $\widetilde{u}_H - \widetilde{u}_{\overline{H}} \in X_{\overline{H}}$,

$$\overline{c}_S \, \|\widetilde{u}_H - \widetilde{u}_{\overline{H}}\|_X \leq \sup_{0 \neq q_h \in Y_h} \frac{\langle B(\widetilde{u}_{\overline{H}} - \widetilde{u}_H), q_h\rangle_H}{\|q_h\|_Y} = \sup_{0 \neq q_h \in Y_h} \frac{\langle A(p_h - \overline{p}_h), q_h\rangle_H}{\|q_h\|_Y} \leq \|p_h - \overline{p}_h\|_Y \, .$$

On the other hand, using the above Galerkin orthogonality and the second equation in (2.24) for $v_{\overline{H}} = \widetilde{u}_{\overline{H}} - \widetilde{u}_H$, this gives

$$\begin{aligned} \|p_h - \overline{p}_h\|_Y^2 &= \langle A(p_h - \overline{p}_h), p_h - \overline{p}_h\rangle_H = \langle B(\widetilde{u}_{\overline{H}} - \widetilde{u}_H), p_h - \overline{p}_h\rangle_H \\ &= \langle B(\widetilde{u}_{\overline{H}} - \widetilde{u}_H), p_h\rangle_H \leq c_2^B \, \|\widetilde{u}_{\overline{H}} - \widetilde{u}_H\|_X \|p_h\|_Y. \end{aligned}$$

Hence we obtain

$$\|\widetilde{u}_{\overline{H}} - \widetilde{u}_H\|_X^2 \leq \frac{1}{\overline{c}_S^2} \, \|p_h - \overline{p}_h\|_Y^2 \leq \frac{c_2^B}{\overline{c}_S^2} \, \|\widetilde{u}_{\overline{H}} - \widetilde{u}_H\|_X \|p_h\|_Y,$$

i.e.,

$$\|\widetilde{u}_{\overline{H}} - \widetilde{u}_H\|_X \leq \frac{c_2^B}{\overline{c}_S^2} \, \|p_h\|_Y \, .$$

Using the triangle inequality and the saturation assumption (2.25) we finally have

$$\|u - \widetilde{u}_H\|_X \leq \|u - \widetilde{u}_{\overline{H}}\|_X + \|\widetilde{u}_{\overline{H}} - \widetilde{u}_H\|_X \leq \eta \, \|u - \widetilde{u}_H\|_X + \frac{c_2^B}{\overline{c}_S^2} \, \|p_h\|_Y,$$

from which the assertion follows. ■

It remains to define, for a given ansatz space $X_{\overline{H}}$, the test space Y_h such that the discrete inf-sup condition (2.23) and therefore (2.13) are satisfied. This can be achieved by assuming a sufficiently rich test space Y_h, as stated in the following abstract approach. But, as we will see later, this is not always required, since one may establish (2.23) in a different way.

Theorem 2.7 *For a given finite element space $X_{\overline{H}} \subset X$ let $Y_h \subset Y$ such that*

$$\sup_{v_{\overline{H}} \in X_{\overline{H}}} \inf_{q_h \in Y_h} \|p_{v_{\overline{H}}} - q_h\|_Y \leq \delta \, \|p_{v_{\overline{H}}}\|_Y = \delta \, \|v_{\overline{H}}\|_S \tag{2.27}$$

is satisfied for some $\delta \in (0,1)$. Then there holds the discrete inf-sup stability condition (2.23), *i.e.,*

$$c_1^B (1-\delta) \, \|v_{\overline{H}}\|_X \leq \sup_{q_h \in Y_h} \frac{\langle B v_{\overline{H}}, q_h \rangle_H}{\|q_h\|_Y} \quad \textit{for all } v_{\overline{H}} \in X_{\overline{H}}. \tag{2.28}$$

Proof. For an arbitrary but fixed $v_{\overline{H}} \in X_{\overline{H}}$ we define $p_{v_{\overline{H}}} = A^{-1} B v_{\overline{H}} \in Y$ and as in the proof of Lemma 2.2 we conclude

$$\|p_{v_{\overline{H}}}\|_Y^2 = \|v_{\overline{H}}\|_S^2 = \langle B v_{\overline{H}}, p_{v_{\overline{H}}} \rangle_H \, .$$

In addition we define $p_{v_{\overline{H}}h} \in Y_h$ as unique solution of the Galerkin variational formulation

$$\langle A p_{v_{\overline{H}}h}, q_h \rangle_H = \langle B v_{\overline{H}}, q_h \rangle_H = \langle A p_{v_{\overline{H}}}, q_h \rangle_H \quad \text{for all } q_h \in Y_h,$$

satisfying the bound

$$\|p_{v_{\overline{H}}h}\|_Y \leq \|p_{v_{\overline{H}}}\|_Y,$$

and Cea's lemma,

$$\|p_{v_{\overline{H}}} - p_{v_{\overline{H}}h}\|_Y \leq \inf_{q_h \in Y_h} \|p_{v_{\overline{H}}} - q_h\|_Y \, .$$

From (2.27) we then obtain

$$\|p_{v_{\overline{H}}} - p_{v_{\overline{H}}h}\|_Y \leq \delta \, \|p_{v_{\overline{H}}}\|_Y \, .$$

Hence we can write

$$\begin{aligned}
\langle B v_{\overline{H}}, p_{v_{\overline{H}}h} \rangle_H &= \langle A p_{v_{\overline{H}}}, p_{v_{\overline{H}}h} \rangle_H = \langle A p_{v_{\overline{H}}}, p_{v_{\overline{H}}} \rangle_H - \langle A p_{v_{\overline{H}}}, p_{v_{\overline{H}}} - p_{v_{\overline{H}}h} \rangle_H \\
&\geq \|p_{v_{\overline{H}}}\|_Y^2 - \|p_{v_{\overline{H}}}\|_Y \|p_{v_{\overline{H}}} - p_{v_{\overline{H}}h}\|_Y \geq (1-\delta) \, \|p_{v_{\overline{H}}}\|_Y^2 \\
&\geq (1-\delta) \, \|p_{v_{\overline{H}}}\|_Y \|p_{v_{\overline{H}}h}\|_Y,
\end{aligned}$$

i.e.,

$$(1-\delta) \, \|v_{\overline{H}}\|_S = (1-\delta) \, \|p_{v_{\overline{H}}}\|_Y \leq \frac{\langle B v_{\overline{H}}, p_{v_{\overline{H}}h} \rangle_H}{\|p_{v_{\overline{H}}h}\|_Y},$$

implying the inf-sup stability condition (2.28). ■

In some applications, e.g., when considering space-time finite element methods for the heat equation as in [40], the discrete inf-sup condition can be established when using a (discretization dependent) norm on the ansatz space X. In particular, let $\|\cdot\|_{X,h} : X \to \mathbb{R}$ define a norm on X_H, satisfying $\|v\|_{X,h} \leq \|v\|_X$ for all $v \in X$ and assume that

$$\widetilde{c}_S \, \|v_H\|_{X,h} \leq \sup_{0 \neq q_h \in Y_h} \frac{\langle B v_H, q_h \rangle_H}{\|q_h\|_Y} \quad \text{for all } v_H \in X_H. \tag{2.29}$$

Then the following stability and error estimates hold true.

Lemma 2.8 *Assume the discrete inf-sup stability condition* (2.29)*. Then the approximate operator* $\widetilde{S}$ *is discrete elliptic, satisfying*

$$\langle \widetilde{S} v_H, v_H \rangle_H \geq [\widetilde{c}_S]^2 \, \|v_H\|^2_{X,h} \quad \textit{for all } v_H \in X_H,$$

and the perturbed variational formulation (2.12) *admits a unique solution. Furthermore, there holds the error estimate*

$$\|u - \widetilde{u}_H\|_{X,h} \leq \left(1 + \frac{2c_2^B}{\widetilde{c}_S}\right) \inf_{v_H \in X_H} \|u - v_H\|_X. \tag{2.30}$$

Proof. The proof of the discrete ellipticity estimate follows the lines of the proof of Lemma 2.3, replacing the discrete inf-sup stability condition (2.13) by (2.29). This already ensures unique solvability of (2.12), as $\|\cdot\|_{X,h}$ defines a norm on $X_H \subset X$. To derive the error estimate, let $v_H \in X_H$ be arbitrary but fixed. First note that we have

$$\|u - \widetilde{u}_H\|_{X,h} \leq \|u - v_H\|_{X,h} + \|v_H - \widetilde{u}_H\|_{X,h} \leq \|u - v_H\|_X + \|v_H - \widetilde{u}_H\|_{X,h} \, .$$

When using the discrete inf-sup stability condition (2.29), and (2.16), for the second term we further have

$$\begin{aligned} \widetilde{c}_S \, \|v_H - \widetilde{u}_H\|_{X,h} &\leq \sup_{0 \neq q_h \in Y_h} \frac{\langle B(v_H - \widetilde{u}_H), q_h \rangle_H}{\|q_h\|_Y} = \sup_{0 \neq q_h \in Y_h} \frac{\langle B v_H - (f - A p_h), q_h \rangle_H}{\|q_h\|_Y} \\ &= \sup_{0 \neq q_h \in Y_h} \frac{\langle B(v_H - u) + A p_h, q_h \rangle_H}{\|q_h\|_Y} \leq c_2^B \, \|u - v_H\|_X + \|p_h\|_Y . \end{aligned}$$

Using $\langle p_h, B v_H \rangle_H = 0$ for all $v_H \in X_H$, see (2.16), we can further estimate

$$\|p_h\|^2_Y = \langle A p_h, p_h \rangle_H = \langle B(u - \widetilde{u}_H), p_h \rangle_H = \langle B(u - v_H), p_h \rangle_H \leq c_2^B \, \|u - v_H\|_X \|p_h\|_Y ,$$

i.e., $\|p_h\|_Y \leq c_2^B \, \|u - v_H\|_X$. Since $v_H \in X_H$ was arbitrary, this concludes the proof. ∎

In what follows we will discuss several applications of this abstract setting. Although our main interest is in the discretization of time dependent partial differential equations such as the heat and the wave equation, we will first consider an elliptic problem in order to present the main ideas of this approach for examples which are well known in literature. Crucial is the choice of the finite element spaces $X_{\overline{H}}$ and Y_h such that the discrete inf-sup condition (2.23) is satisfied. As we will see, and depending on the particular partial differential equation to be solved, we may even consider $X_{\overline{H}} = Y_h$, or we may consider stable pairs of finite element functions which are defined with respect to the same finite element mesh. In the most general case we may choose first a finite element space X_H, then a space $X_{\overline{H}}$ to ensure the saturation condition (2.25), and finally Y_h to satisfy the discrete inf-sup stability condition (2.23), e.g., we may define the ansatz space Y_h with respect to a sufficient small mesh size $h < c_0 \overline{H}$ in order to meet (2.27).

3 Elliptic Dirichlet boundary value problem

As first example we consider the Dirichlet boundary value problem for the Poisson equation,

$$-\Delta u(x) = f(x) \quad \text{for } x \in \Omega, \quad u(x) = 0 \quad \text{for } x \in \Gamma, \tag{3.1}$$

where $\Omega \subset \mathbb{R}^n$, $n = 2, 3$, is a bounded Lipschitz domain with boundary $\Gamma = \partial\Omega$. With respect to the abstract setting we have

$$X = Y = H_0^1(\Omega), \quad A = B = -\Delta : H_0^1(\Omega) \to H^{-1}(\Omega), \quad \|v\|_X = \|\nabla v\|_{L^2(\Omega)}.$$

We obviously have (2.4) with $c_1^B = c_2^B = 1$. In this case, the abstract variational formulation (2.19) reads to find $(u, p) \in H_0^1(\Omega) \times H_0^1(\Omega)$ such that

$$\int_\Omega \nabla p(x) \cdot \nabla q(x)\, dx + \int_\Omega \nabla u(x) \cdot \nabla q(x)\, dx = \langle f, q\rangle_\Omega, \quad \int_\Omega \nabla v(x) \cdot \nabla p(x)\, dx = 0 \tag{3.2}$$

is satisfied for all $(v, q) \in H_0^1(\Omega) \times H_0^1(\Omega)$. Let $X_h := S_h^1(\Omega) \cap H_0^1(\Omega) = \text{span}\{\varphi_k\}_{k=1}^{M_X}$, where $S_h^1(\Omega) = \text{span}\{\varphi_k\}_{k=1}^{\widetilde{M}_X}$ denotes the standard finite element ansatz space of piecewise linear and continuous basis functions, which are defined with respect to some admissible locally quasi-uniform decomposition of Ω into shape regular simplicial finite elements τ_ℓ^h of local mesh size h_ℓ. By definition we have, for $0 \neq u_h \in X_h$,

$$\|\nabla u_h\|_{L^2(\Omega)} = \frac{\langle \nabla u_h, \nabla u_h\rangle_{L^2(\Omega)}}{\|\nabla u_h\|_{L^2(\Omega)}} \leq \sup_{0\neq q_h \in X_h} \frac{\langle \nabla u_h, \nabla q_h\rangle_{L^2(\Omega)}}{\|\nabla q_h\|_{L^2(\Omega)}},$$

i.e., (2.23) for $Y_h = X_{\overline{H}} = X_h$. In the case $X_H = X_{\overline{H}} = X_h$ we finally obtain, due to $p_h = 0$, the standard finite element formulation for the Poisson equation, i.e.,

$$\int_\Omega \nabla u_h(x) \cdot \nabla q_h(x)\, dx = \langle f, q_h\rangle_\Omega \quad \text{for all } q_h \in X_h. \tag{3.3}$$

In order to use $\|\nabla p_h\|_{L^2(\Omega)}$ as a reliable and efficient error estimator we need to introduce a finite element space X_H which is defined with respect to a coarser mesh. In fact, one may first define the ansatz space X_H, and afterwards Y_h by applying some additional refinements. Hence we consider the mixed variational formulation to find $(\widetilde{u}_H, p_h) \in X_H \times Y_h$ such that

$$\int_\Omega \nabla p_h(x) \cdot \nabla q_h(x)\, dx + \int_\Omega \nabla \widetilde{u}_H(x) \cdot \nabla q_h(x)\, dx = \langle f, q_h\rangle_\Omega, \quad \int_\Omega \nabla v_H(x) \cdot \nabla p_h(x)\, dx = 0 \tag{3.4}$$

is satisfied for all $(v_H, q_h) \in X_H \times Y_h$. According to Lemma 2.4 we conclude the error estimate

$$\|\nabla(u - \widetilde{u}_H)\|_{L^2(\Omega)} \leq c_1 \inf_{v_H \in X_H} \|\nabla(u - v_H)\|_{L^2(\Omega)} + c_2 \inf_{q_h \in Y_h} \|\nabla(u - q_h)\|_{L^2(\Omega)}.$$

Hence, assuming $u \in H_0^1(\Omega) \cap H^s(\Omega)$ for some $s \in [1,2]$ we finally obtain the error estimate

$$\|\nabla(u - \widetilde{u}_H)\|_{L^2(\Omega)} \leq c\, H^{s-1}\, |u|_{H^s(\Omega)}. \tag{3.5}$$

Note that the saturation assumption (2.25) now reads, for $u_h \in X_h = Y_h$,

$$\|\nabla(u - u_h)\|_{L^2(\Omega)} \leq \eta\, \|\nabla(u - \widetilde{u}_H)\|_{L^2(\Omega)} \quad \text{for some } \eta \in (0,1).$$

For the particular choice $H = 2h$, i.e., one additional refinement to define Y_h from X_H, this is obviously related to the $h - \frac{h}{2}$ error estimator, e.g., [12, 18]. Although this error estimator is well established in literature, we just present one example solving the mixed Galerkin finite element scheme (3.4). Note that in contrast to the standard $h - \frac{h}{2}$ error estimator we compute $(p_h, \widetilde{u}_H)$ at once solving (3.4).

In what follows we will use the error estimator

$$\eta_h^2 := \|\nabla p_h\|^2_{L^2(\Omega)} = \int_\Omega |\nabla p_h(x)|^2\, dx = \sum_{\ell=1}^{N_h} \int_{\tau_\ell^h} |\nabla p_h(x)|^2\, dx = \sum_{\ell=1}^{N_h} \eta^2_{h,\ell}$$

with the local error indicators

$$\eta^2_{h,\ell} = \int_{\tau_\ell^h} |\nabla p_h(x)|^2\, dx\,.$$

In order to define local indicators for the error $\|\nabla(u - \widetilde{u}_H)\|_{L^2(\Omega)}$ we sum up all local contributions,

$$\eta^2_{H,j} = \sum_{\tau_\ell^h \subset \tau_j^H} \eta^2_{h,\ell}\,.$$

For the adaptive refinement we use the Dörfler marking [15] with the parameter $\theta = 0.5$. The marked elements are then refined using newest vertex bisection. All computations were done using the finite element software Netgen/NGSolve [37]. The resulting linear systems were solved using the sparse direct solver package Umfpack.

As example we consider the L shaped domain

$$\Omega = (-1,1)^2 \setminus ([0,1] \times [-1,0]) \subset \mathbb{R}^2,$$

and we consider the Dirichlet boundary value problem (3.1) with the exact solution, given in polar coordinates,

$$u(r,\varphi) = r^{2/3} \sin \frac{2}{3}\varphi, \quad u \in H^s(\Omega),\ s < \frac{5}{3}. \tag{3.6}$$

Due to the reduced regularity we expect a reduced order of convergence of $\mathcal{O}(H^{2/3})$ for a uniform refinement when measuring the energy error $\|\nabla(u - \widetilde{u}_H)\|_{L^2(\Omega)}$, and of $\mathcal{O}(H^{4/3})$ when considering the L^2 error $\|u - \widetilde{u}_H\|_{L^2(\Omega)}$. When using the adaptive refinement strategy as described, we recover the optimal convergence of $\mathcal{O}(H)$ in the energy norm and of $\mathcal{O}(H^2)$ in $L^2(\Omega)$, respectively. The related numerical results are given in Fig. 1. In Fig. 2 we present a comparison of the error $\|\nabla(u - \widetilde{u}_H)\|_{L^2(\Omega)}$ and the error estimator $\eta = \|\nabla p_h\|_{L^2(\Omega)}$ which shows that the error estimator is effective.

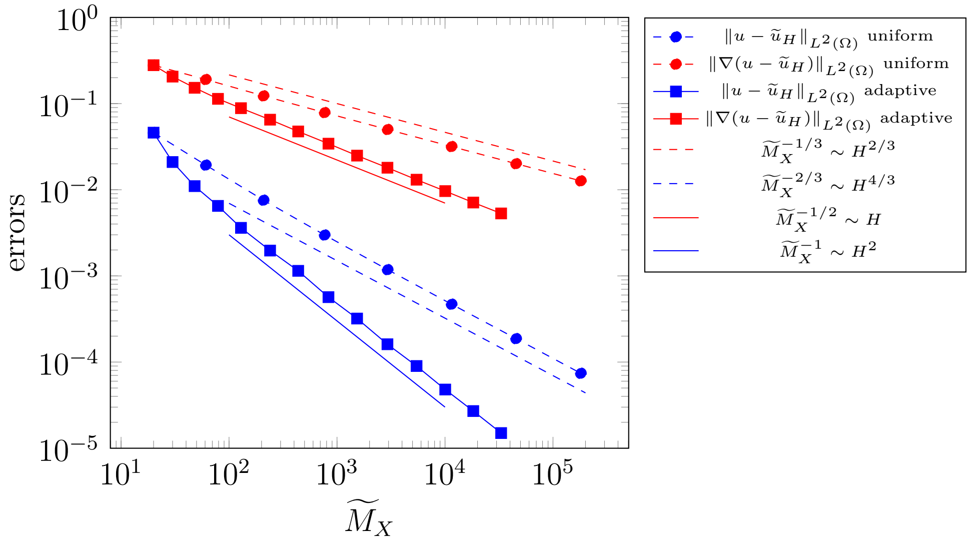


Figure 1: Convergence behavior for the approximation of the singular function (3.6).

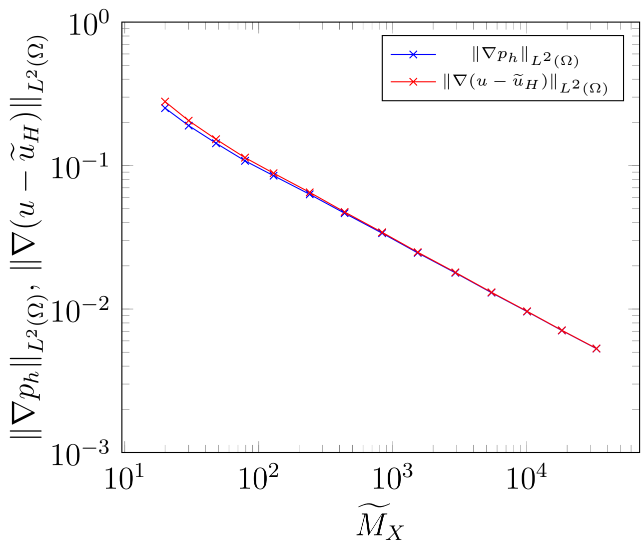


Figure 2: Comparison of the error $\|\nabla(u-\widetilde{u}_H)\|_{L^2(\Omega)}$ and the error estimator $\|\nabla p_h\|_{L^2(\Omega)}$.

Remark 3.1 *Instead of $X = Y = H^1_0(\Omega)$, $S = B^*A^{-1}B = -\Delta$, we may also consider the abstract setting with $X = H_\Delta(\Omega) := \{v \in H^1_0(\Omega) : \Delta v \in L^2(\Omega)\}$, $Y = L^2(\Omega)$, $B = -\Delta : H_\Delta(\Omega) \to Y$, and $A = I : L^2(\Omega) \to L^2(\Omega)$. Then we have that $S = B^*A^{-1}B = \Delta^2$ is the Bi-Laplacian. For a conforming finite element discretization in this case one may use tensor product meshes and quadratic B splines to define X_h, and piecewise constant basis functions to define Y_h. For a related approach in the case of a distributed optimal control problem, see, e.g., [9, 10]. Alternatively, and applying integration by parts twice, one may exchange the roles of $X = L^2(\Omega)$ and $Y = H_\Delta(\Omega)$, in order to consider, e.g., less regular Dirichlet boundary conditions $u = g \in L^2(\Gamma)$, see, e.g., [2]. In addition, one can also apply the abstract theory as presented in this paper for the analysis of related boundary element methods [41].*

4 Parabolic Dirichlet boundary value problem

As an example for a parabolic partial differential equation we consider the Dirichlet problem for the heat equation,

$$\begin{aligned} \partial_t u(x,t) - \Delta_x u(x,t) &= f(x,t) && \text{for } (x,t) \in Q := \Omega \times (0,T), && (4.1a)\\ u(x,t) &= 0 && \text{for } (x,t) \in \Sigma := \Gamma \times (0,T), && (4.1b)\\ u(x,0) &= 0 && \text{for } x \in \Omega, && (4.1c) \end{aligned}$$

where $\Omega \subset \mathbb{R}^n$, $n = 1, 2, 3$ is a bounded Lipschitz domain with boundary $\Gamma = \partial\Omega$, and $T > 0$ is a given time horizon. In view of the abstract setting we have the Bochner spaces

$$X := L^2(0,T; H^1_0(\Omega)) \cap H^1_{0,}(0,T; H^{-1}(\Omega)),\ Y := L^2(0,T; H^1_0(\Omega)),\ H := L^2(Q),$$

where $H^1_{0,}(0,T; H^{-1}(\Omega)) := \{u \in L^2(Q) : \partial_t u \in L^2(0,T; H^{-1}(\Omega)), u(x,0) = 0, x \in \Omega\}$. The operators $A : Y \to Y^*$ and $B : X \to Y^*$ are defined in a variational sense satisfying

$$\langle Ap, q\rangle_Q := \langle \nabla_x p, \nabla_x q\rangle_{L^2(Q)}, \quad \langle Bu, q\rangle_Q := \langle \partial_t u, q\rangle_Q + \langle \nabla_x u, \nabla_x q\rangle_{L^2(Q)}$$

for all $p, q \in Y$ and $u \in X$. The corresponding norms read

$$\|p\|_Y := \|\nabla_x p\|_{L^2(Q)}, \quad \|u\|_X := \sqrt{\|\partial_t u\|^2_{Y^*} + \|\nabla_x u\|^2_Y}, \quad \|\partial_t u\|_{Y^*} = \|\nabla_x w_u\|_{L^2(Q)},$$

where $w_u \in Y$ is the unique solution of the variational problem

$$\langle \nabla_x w_u, \nabla_x q\rangle_{L^2(Q)} = \langle \partial_t u, q\rangle_Q \quad \text{for all } q \in Y. \tag{4.2}$$

The operator $A : Y \to Y^*$ is self-adjoint, bounded and elliptic. Moreover, $B : X \to Y^*$ is bounded with $c^B_2 = \sqrt{2}$, i.e.,

$$\|Bv\|_{Y^*} \le \sqrt{2}\, \|v\|_X \quad \text{for all } v \in X.$$

In order to prove an inf-sup stability condition, for $u \in X$ we first define $w_u \in Y$ as in (4.2). For $q := u + w_u \in Y$ we then have

$$\begin{aligned}
\langle Bu, q\rangle_Q &= \langle Bu, u + w_u\rangle_Q \\
&= \langle \partial_t u, u\rangle_Q + \langle \nabla_x u, \nabla_x u\rangle_{L^2(Q)} + \langle \partial_t u, w_u\rangle_Q + \langle \nabla_x u, \nabla_x w_u\rangle_{L^2(Q)} \\
&= \langle \nabla w_u, \nabla_x u\rangle_{L^2(\Omega)} + \langle \nabla_x u, \nabla_x u\rangle_{L^2(Q)} + \langle \nabla_x w_u, \nabla_x w_u\rangle_{L^2(Q)} + \langle \nabla_x u, \nabla_x w_u\rangle_{L^2(Q)} \\
&= \langle \nabla_x (u + w_u), \nabla_x (u + w_u)\rangle_{L^2(\Omega)} \\
&= \|q\|_Y^2 .
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\|q\|_Y^2 &= \|\nabla(u + w_u)\|^2_{L^2(Q)} \\
&= \langle \nabla_x u, \nabla_x u\rangle_{L^2(\Omega)} + \langle \nabla_x w_u, \nabla_x w_u\rangle_{L^2(Q)} + 2\,\langle \nabla_x w_u, \nabla_x u\rangle_{L^2(Q)} \\
&= \langle \nabla_x u, \nabla_x u\rangle_{L^2(\Omega)} + \langle \nabla_x w_u, \nabla_x w_u\rangle_{L^2(Q)} + 2\,\langle \partial_t u, u\rangle_Q \\
&\geq \|\nabla_x u\|_Y^2 + \|\nabla_x w_u\|_Y^2 \\
&= \|u\|_X^2,
\end{aligned}$$

implies

$$\langle Bu, u + w_u\rangle_Q \geq \|u\|_X \|u + w_u\|_Y ,$$

and therefore the inf-sup stability condition

$$\|u\|_X \leq \sup_{0 \neq q \in Y} \frac{\langle Bu, q\rangle_Q}{\|q\|_Y} \quad \text{for all } u \in X$$

follows with $c_1^B = 1$. Note that this estimate is improved than originally derived in [40]. Moreover, $B : X \to Y^*$ is surjective, see, e.g., [27]. Thus, we can apply the abstract theory from Section 2. The variational formulation (2.19) now reads to find $(u, p) \in X \times Y$ such that

$$\begin{aligned}
\int_Q \nabla_x p(x,t) \cdot \nabla_x q(x,t)\, dx\, dt \qquad & \\
+ \int_Q \Big[\partial_t u(x,t)\, q(x,t) + \nabla_x u(x,t) \cdot \nabla_x q(x,t)\Big] dx\, dt &= \langle f, q\rangle_Q, \qquad (4.3) \\
\int_Q \Big[\partial_t v(x,t)\, p(x,t) + \nabla_x v(x,t) \cdot \nabla_x p(x,t)\Big] dx\, dt &= 0
\end{aligned}$$

is satisfied for all $(v, q) \in X \times Y$.

For the discretization of (4.3) we first consider an ansatz space $X_H = \text{span}\{\varphi_k\}_{k=1}^{M_X} \subset X$ of piecewise linear and continuous basis functions φ_k which are defined with respect to some admissible and locally quasi-uniform decomposition of the space-time domain Q into shape regular simplicial finite elements q_ℓ^H of mesh size H_ℓ. In addition we define the ansatz space $Y_h = \text{span}\{\psi_i\}_{i=1}^{M_Y} \subset Y$ of piecewise linear and continuous basis functions ψ_i which are defined with respect to some refined decomposition of the space-time domain Q into finite elements q_i^h of local mesh size h_i. From a practical point of view, we may use one additional refinement of the mesh which was used to define X_H, i.e., for $q_i^h \subset q_\ell^H$ we have $h_i = \frac{1}{2} H_\ell$. As an alternative, we may also use second order form functions on q_ℓ^H to define Y_h. In both cases we have the inclusion $X_H \subset Y_h$ which enables us to prove a discrete inf-sup stability condition. For any $u \in X$ we define $w_{uh} \in Y_h$ as the unique solution of the Galerkin variational formulation

$$\langle \nabla_x w_{uh}, \nabla_x q_h \rangle_{L^2(Q)} = \langle \partial_t u, q_h \rangle_Q \quad \text{for all } q_h \in Y_h.$$

With this we define the discrete norm

$$\|u\|_{X,h} := \sqrt{\|u\|_Y^2 + \|w_{uh}\|_Y^2} \leq \|u\|_X,$$

and as in the continuous case, instead of $w_u \in Y$ we now use $w_{u_H h} \in Y_h$, we can prove the discrete inf-sup stability condition

$$\|u_H\|_{X,h} \leq \sup_{0 \neq q_h \in Y_h} \frac{\langle B u_H, q_h \rangle_Q}{\|q_h\|_Y} \quad \text{for all } u_H \in X_H.$$

Hence we conclude unique solvability of the mixed space-time finite element variational formulation to find $(\widetilde{u}_H, p_h) \in X_H \times Y_h$ such that

$$\begin{aligned} \int_Q \nabla_x p_h(x,t) \cdot \nabla_x q_h(x,t) \, dx \, dt \qquad & & \\ + \int_Q \Big[\partial_t \widetilde{u}_H(x,t) \, q_h(x,t) + \nabla_x \widetilde{u}_H(x,t) \cdot \nabla_x q_h(x,t) \Big] dx \, dt &=& \langle f, q_h \rangle_Q, \\ \int_Q \Big[\partial_t v_H(x,t) \, p_h(x,t) + \nabla_x v_H(x,t) \cdot \nabla_x p_h(x,t) \Big] dx \, dt &=& 0 \end{aligned} \tag{4.4}$$

is satisfied for all $(v_H, q_h) \in X_H \times Y_h$. The related error estimate now follows from the abstract estimate (2.30), i.e., with $c_2^B = \sqrt{2}$ and $\widetilde{c}_S = 1$ this gives

$$\|u - \widetilde{u}_H\|_{X,h} \leq \big(1 + 2\sqrt{2}\big) \inf_{v_H \in X_H} \|u - v_H\|_X \, .$$

Hence, assuming $u \in H^s(Q)$ for some $s \in [1,2]$ we conclude the error estimate, e.g., [40],

$$\|\nabla_x (u - \widetilde{u}_H)\|_{L^2(Q)} \leq \|u - \widetilde{u}_H\|_{X,h} \leq c H^{s-1} |u|_{H^s(Q)}.$$

We also define the ansatz space $X_{\overline{H}} := Y_h \cap X$ and compute $(\widetilde{u}_{\overline{H}}, \overline{p}_h) \in X_{\overline{H}} \times Y_h$ such that

$$\begin{aligned} \int_Q \nabla_x \overline{p}_h(x,t) \cdot \nabla_x q_h(x,t) \, dx \, dt \qquad & & \\ + \int_Q \Big[\partial_t \widetilde{u}_{\overline{H}}(x,t) \, q_h(x,t) + \nabla_x \widetilde{u}_{\overline{H}}(x,t) \cdot \nabla_x q_h(x,t) \Big] dx \, dt &=& \langle f, q_h \rangle_Q, \\ \int_Q \Big[\partial_t v_{\overline{H}}(x,t) \, \overline{p}_h(x,t) + \nabla_x v_{\overline{H}}(x,t) \cdot \nabla_x \overline{p}_h(x,t) \Big] dx \, dt &=& 0 \end{aligned} \tag{4.5}$$

is satisfied for all $(v_{\overline{H}}, q_h) \in X_{\overline{H}} \times Y_h$. As before we can prove the discrete inf-sup stability condition

$$\|u_{\overline{H}}\|_{X,h} \leq \sup_{0 \neq q_h \in Y_h} \frac{\langle B u_{\overline{H}}, q_h \rangle_Q}{\|q_h\|_Y} \quad \text{for all } u_{\overline{H}} \in X_{\overline{H}},$$

which ensures unique solvability of (4.5), and we write the abstract saturation condition (2.25) as

$$\|u - \widetilde{u}_{\overline{H}}\|_{X,h} \leq \eta \, \|u - \widetilde{u}_H\|_{X,h} \quad \text{for some } \eta \in (0,1).$$

As in the case of the Poisson equation we now use global error estimator

$$\eta_h^2 = \|\nabla_x p_h\|^2_{L^2(Q)} = \int_Q |\nabla_x p_h(x,t)|^2 \, dx \, dt = \sum_{\ell=1}^{N_h} \int_{q_\ell^h} |\nabla_x p_h(x,t)|^2 \, dx \, dt = \sum_{\ell=1}^{N_h} \eta_{h,\ell}^2$$

with the local error indicators

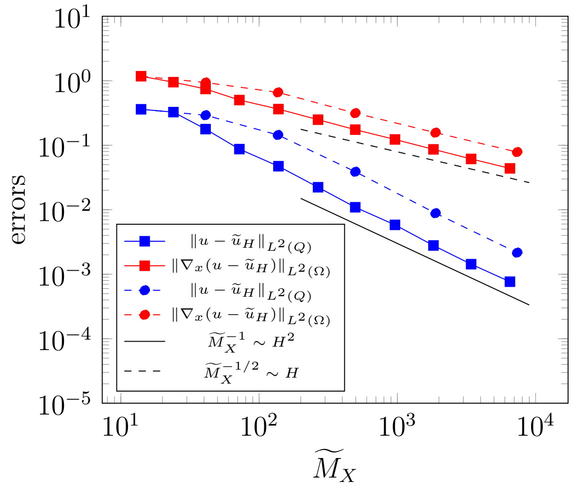
$$\eta_{h,\ell}^2 = \int_{q_\ell^h} |\nabla_x p_h(x,t)|^2 \, dx \, dt, \quad \eta_{H,j}^2 = \sum_{\tau_\ell^h \subset \tau_j^H} \eta_{h,\ell}^2$$

to drive an adaptive refinement algorithm, using Dörfler marking with $\theta = 0.5$ and newest vertex bisection, [15], and we compare these results with those obtained by a uniform refinement strategy. We consider the Galerkin variational formulation (4.4) for the ansatz space $X_H = S_H^1(Q) \cap X$ of piecewise linear and continuous basis functions, and the test space $Y_H = S_H^2(Q) \cap Y$ of piecewise second order and continuous basis functions. The latter corresponds to the test space $Y_h = S_h^1(Q) \cap Y$ of piecewise linear continuous basis functions, which are defined with respect to a refined mesh with local mesh size $h = \frac{1}{2}H$. As before, all computations were done in Netgen/NGSolve [37] where we used the sparse direct solver package Umfpack to solve the resulting linear system.

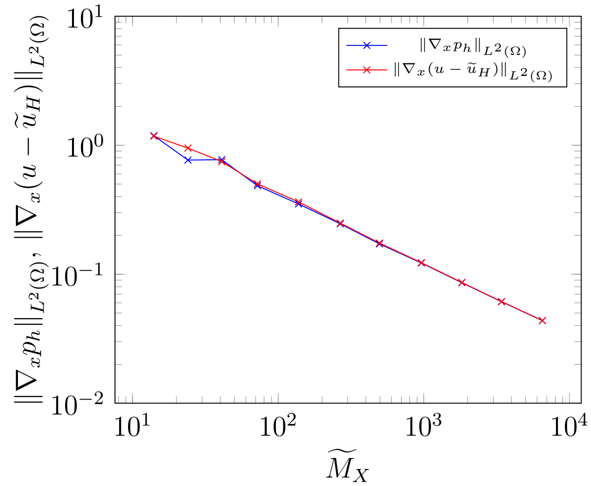
In the first example we use the one-dimensional spatial domain $\Omega = (0,3)$ and the time horizon $T = 6$, i.e., we have the space-time domain $Q := (0,3) \times (0,6) \subset \mathbb{R}^2$. As exact solution we consider the smooth function

$$u(x,t) := \begin{cases} \frac{1}{2}(t - x - 2)^3 (x - t)^3 \sin \frac{\pi}{3} x & \text{for } x \leq t \text{ and } t - x \leq 2, \\ 0 & \text{else,} \end{cases} \tag{4.6}$$

and we compute $f = \partial_t u - \Delta_x u$ accordingly. Since u is smooth we expect optimal orders of convergence, i.e., $\mathcal{O}(H^2)$ when measuring the error in $L^2(Q)$, and $\mathcal{O}(H)$ in the energy norm $\|\nabla_x \cdot \|_{L^2(Q)}$. These rates are confirmed by the numerical results for both a uniform and an adaptive refinement strategy, as shown in Fig. 3(a). In Fig. 3(b) we present a comparison of the error $\|\nabla_x (u - \widetilde{u}_H)\|_{L^2(Q)}$ with the error estimator $\|\nabla_x p_h\|_{L^2(Q)}$, where the curves are almost parallel. Moreover, in Table 1 we provide a comparison of the errors $\|\nabla_x (u - \widetilde{u}_H)\|_{L^2(Q)}$ for both the uniform and adaptive refinement strategy. We observe that in the adaptive case less degrees of freedom are needed to reach the same level of accuracy than in the case of a uniform refinement. Finally, in Fig. 4 we present the related finite



(a) Errors $\|\nabla_x(u - \widetilde{u}_H)\|_{L^2(Q)}$ and $\|u - \widetilde{u}_H\|_{L^2(Q)}$ for uniform and adaptive refinement strategies.

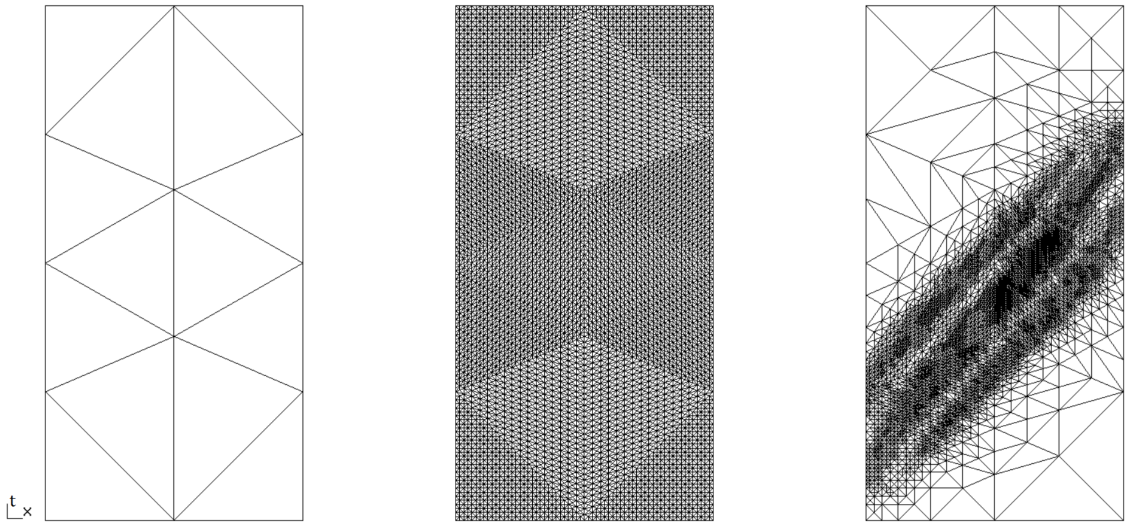


(b) History of the norms $\|\nabla_x p_h\|_{L^2(Q)}$, $\|\nabla_x(u - \widetilde{u}_H)\|_{L^2(Q)}$ for an adaptive refinement strategy.

Figure 3: Convergence results in the case of a smooth solution for the heat equation.

uniform			adaptive		
L	$\widetilde{M}_X$	$\|\nabla_x(u - \widetilde{u}_H)\|_{L^2(Q)}$	L	$\widetilde{M}_X$	$\|\nabla_x(u - \widetilde{u}_H)\|_{L^2(Q)}$
0	14	1.175e+00	0	14	1.175e+00
1	41	9.391e-01	2	41	7.463e-01
2	137	6.597e-01	4	138	3.637e-01
3	497	3.144e-01	6	496	1.745e-01
4	1889	1.568e-01	8	1825	8.636e-02
5	7361	7.840e-02	10	6524	4.377e-02

Table 1: Comparison of the errors between uniform and adaptive refinement.



(a) Initial mesh (b) Uniform refinement $L = 5$ (c) Adaptive refinement $L = 10$

Figure 4: Space-time finite element meshes.

element meshes. Note that the solution is smooth but behaves similar as a wave. This motivates to use a mesh which is adaptive in space and time.

As a second example we consider the heat equation (4.1) in the space-time domain $Q = (0,1)^2 \subset \mathbb{R}^2$, i.e., $\Omega = (0,1)$ and $T = 1$ with homogeneous Dirichlet and initial conditions, but a discontinuous right hand side

$$f(x,t) = \begin{cases} 1 & (x,t) \in \left\{ (x,t) \in (0,1) \times \left(\frac{1}{10}, \frac{1}{2}\right) : \ x - \frac{1}{10} \leq t \leq x - \frac{1}{20} \right\}, \\ 0 & \text{else}, \end{cases} \tag{4.7}$$

as depicted in Fig. 6(a). Note that a similar example was also considered in [20]. Since the exact solution is unknown, in Fig. 5 we provide the results for the error estimator $\|\nabla_x p_h\|_{L^2(Q)}$ which is equivalent to the error $\|\nabla_x (u - \widetilde{u}_H)\|_{L^2(Q)}$. As already observed in [20, Section 5.2.3], in the case of uniform refinement we obtain a reduced rate of $\widetilde{M}_X^{-\frac{1}{4}}$ which corresponds to $\mathcal{O}(H^{\frac{1}{2}})$, while in the case of an adaptive refinement we have the optimal rate of $\widetilde{M}_X^{-\frac{1}{2}}$, i.e., $\mathcal{O}(H)$. In Fig. 6(b) we provide the adaptive mesh at level $L = 9$.

As a last example for the heat equation we consider again the space-time domain $Q = (0,1)^2$, with $f(x,t) = 2$ for $(x,t) \in Q$, and $u_0(x) = 1$ for $x \in (0,1)$. Obviously, we have $u_0 \in L^2(0,1)$, but $u_0 \notin H_0^1(\Omega)$, i.e., there is no compatibility with the homogeneous Dirichlet boundary conditions at $t = 0$. This results in a reduced order of convergence. In this example, we choose the parameter $\theta = 0.9$ in the Dörfler marking strategy. Since the exact solution is not known, in Fig. 7 we provide the results for the error indicator $\|\nabla_x p_h\|_{L^2(Q)}$ which is equivalent to the error $\|\nabla_x (u - \widetilde{u}_H)\|_{L^2(Q)}$. We observe a rate of $\mathcal{O}(H^{0.08})$ in the case of a uniform refinement, and of $\mathcal{O}(H^{0.16})$ for the adaptive refinement strategy. These rates coincide with those observed in [20, Section 5.2.4].

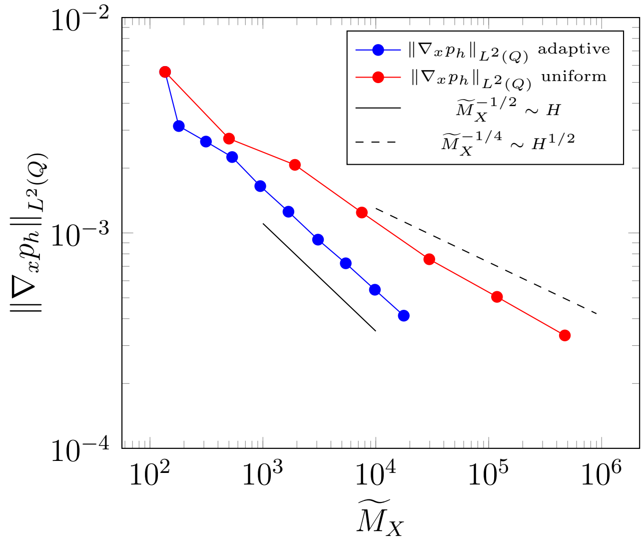
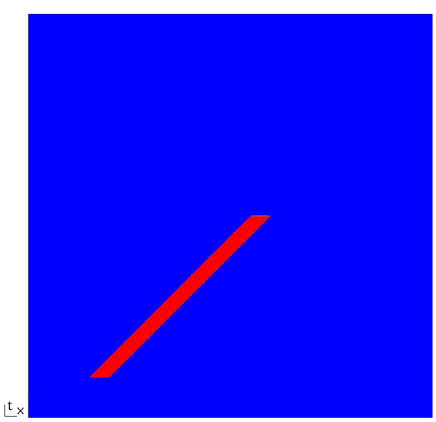
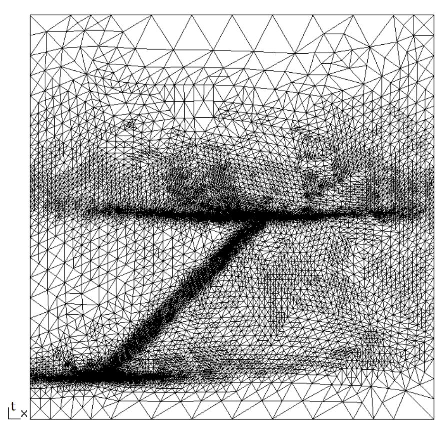


Figure 5: Error estimator $\|\nabla_x p_h\|_{L^2(Q)}$ in the case of a discontinuous right hand side.



(a) Discontinuous right hand side (4.7)



(b) Adaptive mesh with $\widetilde{M}_X = 17689$

Figure 6: Singular solution of the heat equation for discontinuous right hand side.

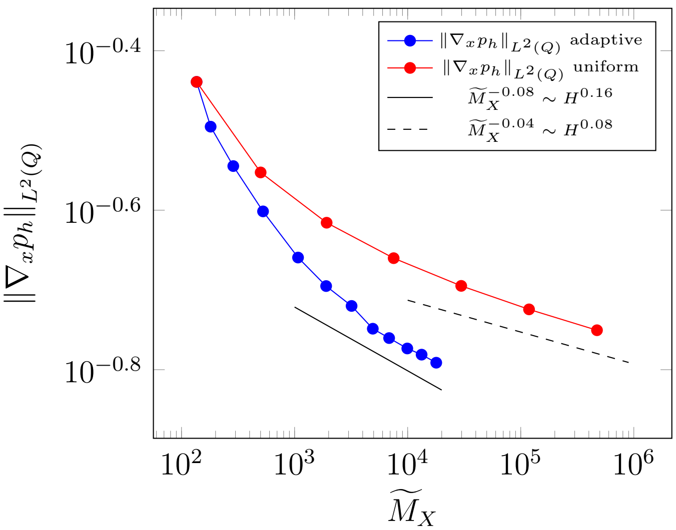


Figure 7: Error estimator $\|\nabla_x p_h\|_{L^2(Q)}$ in the case of non compatible initial conditions $u_0 \notin H^1_0(\Omega)$.

Remark 4.1 *Similar as in the case for the Poisson equation we may also define*

$$X := \Big\{ v \in L^2(0,T; L^2(\Omega)) \cap H^1_{0,}(0,T; H^{-1}(\Omega)) : \partial_t v - \Delta_x v \in L^2(Q) \Big\}, \quad Y = L^2(Q).$$

In this case we have $B = \partial_t - \Delta_x : X \to L^2(Q)$, and $A := I : L^2(Q) \to L^2(Q)$. For a conforming space-time finite element discretization one can use tensor-product meshes with quadratic B splines in the spatial directions, and piecewise linear continuous basis functions in the temporal direction.

5 Hyperbolic Dirichlet boundary value problem

In this section we apply the abstract theory to the Dirichlet boundary value problem for the wave equation,

$$\Box u(x,t) := \partial_{tt} u(x,t) - \Delta_x u(x,t) = f(x,t) \quad \text{for } (x,t) \in Q := \Omega \times (0,T), \tag{5.1a}$$

$$u(x,t) = 0 \quad \text{for } (x,t) \in \Sigma := \Gamma \times (0,T), \tag{5.1b}$$

$$u(x,0) = \partial_t u(x,t)_{|t=0} = 0 \quad \text{for } x \in \Omega, \tag{5.1c}$$

where, as in the parabolic case, $\Omega \subset \mathbb{R}^n$, $n = 1,2,3$ is a bounded Lipschitz domain with boundary $\Gamma := \partial\Omega$, and $T > 0$ is a given time horizon. Following [42], the space-time variational formulation of the wave equation (5.1) is to find $u \in H^{1,1}_{0;0,}(Q)$ such that

$$-\langle \partial_t u, \partial_t q \rangle_{L^2(Q)} + \langle \nabla_x u, \nabla_x q \rangle_{L^2(Q)} = \langle f, q \rangle_Q \tag{5.2}$$

is satisfied for all $q \in H^{1,1}_{0;,0}(Q)$. Here we use the anisotropic Sobolev space

$$H^{1,1}_{0;0,}(Q) := L^2(0,T; H^1_0(\Omega)) \cap H^1_{0,}(0,T; L^2(\Omega)),$$

where $H^1_{0,}(0,T;L^2(\Omega))$ covers the zero initial condition $u(x,0) = 0$ for $x \in \Omega$, while $L^2(0,T;H^1_0(\Omega))$ includes the homogeneous Dirichlet boundary condition on Σ. Note that the second initial condition $\partial_t u(x,t)_{|t=0} = 0$ for $x \in \Omega$ enters the variational formulation (5.2) in a natural way. A norm in $H^{1,1}_{0;0,}(Q)$ is given by the graph norm

$$\|u\|_{H^{1,1}_{0;0,}(Q)} := \sqrt{\|\partial_t u\|^2_{L^2(Q)} + \|\nabla_x u\|^2_{L^2(Q)}} = |u|_{H^1(Q)}.$$

Note that $H^{1,1}_{0;,0}(Q)$ is defined accordingly, but with a zero terminal condition $q(x,T) = 0$ for $x \in \Omega$. Although the right hand side of the variational formulation is well defined for $f \in [H^{1,1}_{0;,0}(Q)]^*$, in order to ensure a unique solution $u \in H^{1,1}_{0;0,}(Q)$ we have to assume $f \in L^2(Q)$, e.g., [25, 42]. In fact, the solution operator mapping $f \in L^2(Q)$ to the solution $u \in H^{1,1}_{0;0,}(Q)$ of the variational formulation (5.2) does not define an isomorphism. Instead we have to enlarge the ansatz space in order to incorporate the second initial condition $\partial_t u(x,t)_{|t=0} = 0$ in an appropriate way. In what follows we will consider a generalized variational formulation of the wave equation, see [43]: For the enlarged space-time domain $Q_- := \Omega \times (-T,T)$, and for $u \in L^2(Q)$ we define the zero extension

$$\widetilde{u}(x,t) := \begin{cases} u(x,t) & \text{for } (x,t) \in Q, \\ 0, & \text{else.} \end{cases}$$

The application of the wave operator $\Box\widetilde{u}$ on Q_- will be formulated as a distribution, i.e., for $\varphi \in C^\infty_0(Q_-)$ we define

$$\langle \Box\widetilde{u}, \varphi\rangle_{Q_-} := \int_{Q_-} \widetilde{u}(x,t)\,\Box\varphi(x,t)\,dx\,dt = \int_Q u(x,t)\,\Box\varphi(x,t)\,dx\,dt.$$

Now we are in the position to introduce the space

$$\mathcal{H}(Q) := \Big\{ u = \widetilde{u}_{|Q} : \widetilde{u} \in L^2(Q_-),\, \widetilde{u}_{|\Omega\times(-T,0)} = 0,\, \Box\widetilde{u} \in [H^1_0(Q_-)]^* \Big\},$$

with the graph norm

$$\|u\|_{\mathcal{H}(Q)} := \sqrt{\|u\|^2_{L^2(Q)} + \|\Box\widetilde{u}\|^2_{[H^1_0(Q_-)]^*}}.$$

The normed vector space $(\mathcal{H}(Q), \|\cdot\|_{\mathcal{H}(Q)})$ is a Banach space, and it holds true that, see [43, Lemma 3.5], $H^{1,1}_{0;0,}(Q) \subset \mathcal{H}(Q)$ i.e.,

$$\|\Box\widetilde{u}\|_{[H^1_0(Q_-)]^*} \le \|u\|_{H^{1,1}_{0;0,}(Q)} \quad \text{for all } u \in H^{1,1}_{0;0,}(Q). \tag{5.3}$$

Therefore, we can consider the space

$$\mathcal{H}_{0;0,}(Q) := \overline{H^{1,1}_{0;0,}(Q)}^{\|\cdot\|_{\mathcal{H}(Q)}} \subset \mathcal{H}(Q)$$

which will serve as ansatz space. For $u \in \mathcal{H}_{0;0,}(Q)$, an equivalent norm is given as, see [43, Lemma 3.6],

$$\|u\|_{\mathcal{H}_{0;0,}(Q)} = \|\Box\widetilde{u}\|_{[H_0^1(Q_-)]^*}.$$

For given $f \in [H^{1,1}_{0;,0}(Q)]^*$ we now consider the variational formulation to find $u \in \mathcal{H}_{0;0,}(Q)$ such that

$$b(u,q) := \langle Bu, q\rangle_Q = \langle \Box\widetilde{u}, \mathcal{E}q\rangle_{Q_-} = \langle f, q\rangle_Q \quad \text{for all } q \in H^{1,1}_{0;,0}(Q), \tag{5.4}$$

where $\mathcal{E} : H^{1,1}_{0;,0}(Q) \to H^1_0(Q_-)$ is a suitable extension operator, e.g., reflection in time. Note that we have, see [43, Lemma 3.5],

$$\langle \Box\widetilde{u}, \mathcal{E}q\rangle_{Q_-} = -\langle \partial_t u, \partial_t q\rangle_{L^2(Q)} + \langle \nabla_x u, \nabla_x q\rangle_{L^2(Q)}, \quad u \in H^{1,1}_{0;0,}(Q) \subset \mathcal{H}_{0;0,}(Q), q \in H^{1,1}_{0;,0}(Q).$$

We conclude that the bilinear form within the variational formulation (5.4) is bounded for all $u \in \mathcal{H}_{0;0,}(Q)$ and $q \in H^{1,1}_{0;,0}(Q)$, and satisfies the inf-sup stability condition

$$\|u\|_{\mathcal{H}_{0;0,}(Q)} = \sup_{0\neq q\in H^{1,1}_{0;,0}(Q)} \frac{\langle \Box\widetilde{u}, \mathcal{E}q\rangle_{Q_-}}{\|q\|_{H^{1,1}_{0;,0}(Q)}} \quad \text{for all } u \in \mathcal{H}_{0;0,}(Q). \tag{5.5}$$

In view of the abstract setting we therefore have

$$X := \mathcal{H}_{0;0,}(Q), \quad Y := H^{1,1}_{0;,0}(Q), \quad H := L^2(Q), \quad B : X \to Y^*.$$

Finally, we define

$$\langle Ap, q\rangle_Q := \langle \partial_t p, \partial_t q\rangle_{L^2(Q)} + \langle \nabla_x p, \nabla_x q\rangle_{L^2(Q)} \quad \text{for all } p, q \in Y = H^{1,1}_{0;,0}(Q),$$

which corresponds to the space-time Laplacian. Thus we can apply the abstract theory as given in Section 2. The variational formulation (2.19) now reads to find $(u,p) \in X \times Y$ such that

$$\langle \partial_t p, \partial_t q\rangle_{L^2(Q)} + \langle \nabla_x p, \nabla_x q\rangle_{L^2(Q)} + \langle \Box\widetilde{u}, \mathcal{E}q\rangle_{Q_-} = \langle f, q\rangle_Q, \quad \langle \Box\widetilde{v}, \mathcal{E}p\rangle_{Q_-} = 0 \tag{5.6}$$

is satisfied for all $(v,q) \in X \times Y$.

Let $X_H \subset H^{1,1}_{0;0,}(Q) \subset \mathcal{H}_{0;0,}(Q) = X$ be the conforming finite element space of piecewise linear and continuous basis functions which are defined with respect to an admissible locally quasi-uniform decomposition of the space-time domain Q into shape-regular simplicial finite elements of mesh size H. Moreover, let $Y_h \subset H^{1,1}_{0;,0}(Q)$ be a second finite element space of piecewise linear and continuous basis functions which are defined with respect to a refined decomposition of the space-time domain into finite elements of mesh size h. The Galerkin discretization of (5.6) then reads to find $(u_H, p_h) \in X_H \times Y_h$ such that

$$\langle \partial_t p_h, \partial_t q_h\rangle_{L^2(Q)} + \langle \nabla_x p_h, \nabla_x q_h\rangle_{L^2(Q)} - \langle \partial_t u_H, \partial_t q_h\rangle_{L^2(Q)} + \langle \nabla_x u_H, \nabla_x q_h\rangle_{L^2(Q)} = \langle f, q_h\rangle_Q \tag{5.7}$$

and

$$-\langle \partial_t v_H, \partial_t p_h\rangle_{L^2(Q)} + \langle \nabla_x v_H, \nabla_x p_h\rangle_{L^2(Q)} = 0 \tag{5.8}$$

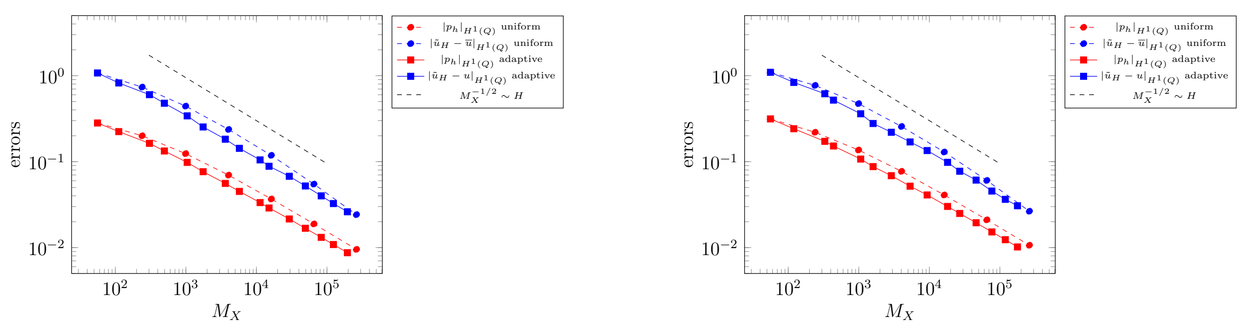
is satisfied for all $(v_H, q_h) \in X_H \times Y_h$. Using the relation $\|u\|_{\mathcal{H}_{0;0,}(Q)} \leq \|u\|_{H^{1,1}_{0;0,}(Q)} = |u|_{H^1(Q)}$ for $u \in H^{1,1}_{0;0,}(Q)$ we conclude a best approximation result.

Theorem 5.1 *Assume* (2.27) *such that the discrete inf-sup stability condition* (2.13) *is satisfied, and assume* $u \in H^{1,1}_{0;0,}(Q) \cap H^s(Q)$ *for some* $s \in [1,2]$*. For the unique solution* $(u_H, p_h) \in X_H \times Y_h$ *of* (5.7) *and* (5.8) *there holds the error estimate*

$$\|u - u_H\|_{\mathcal{H}_{0;0,}(Q)} \le \inf_{v_H \in X_H} |u - v_H|_{H^1(Q)} \le c\, H^{s-1} |u|_{H^s(Q)}.$$

If the finite element space Y_h is chosen appropriately, i.e., such that the discrete inf-sup stability condition (2.23) and the saturation condition (2.25) are satisfied, $\eta = |p_h|_{H^1(Q)}$ serves as an error indicator for $\|u - u_H\|_{\mathcal{H}_{0;0,}(Q)}$. We define the finite element spaces $X_H = S^1_H(Q) \cap X$ and $Y_h = S^1_h(Q) \cap Y$ of piecewise linear continuous basis functions with either $h = H/2$ or $h = H/4$, and we consider the Dörfler criterion [15] with parameter $\theta = 0.5$. As in the previous examples we refine all marked elements using red-green-blue refinement.

As a first example we consider the smooth solution (4.6) as for the heat equation, and we compute $f = \partial_{tt} u - \Delta_x u$. In this case we expect to see optimal orders of convergence for the error in the energy norm. In Fig. 8 we present the numerical results for both a uniform and an adaptive refinement, and for $h = H/2$ as well as for $h = H/4$. In all cases we observe a linear convergence, for both the error $|u - \widetilde{u}_H|_{H^1(Q)}$, and the estimator $|p_h|_{H^1(Q)}$. Finally, in Fig. 9 we present the initial mesh, and two adaptively refined meshes in the space-time domain.



(a) $h = H/2$ (b) $h = H/4$

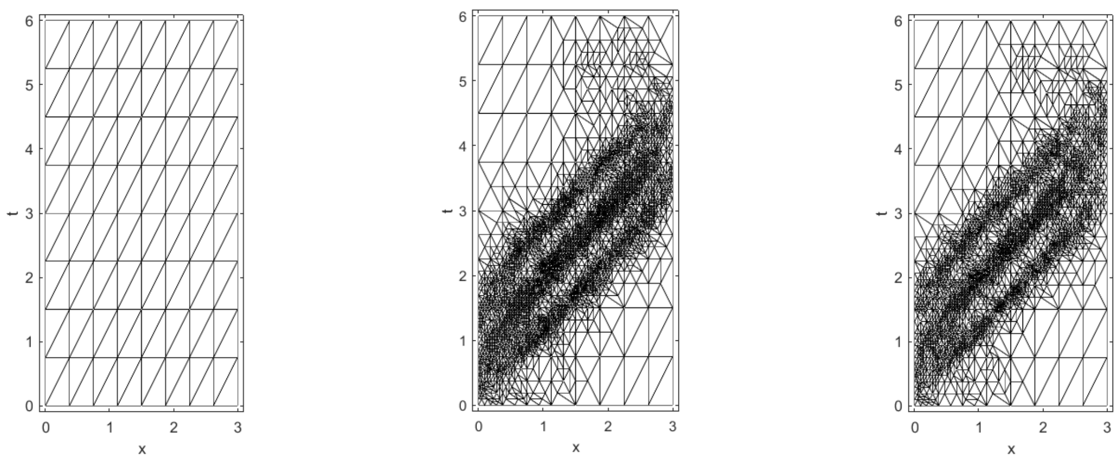
Figure 8: Errors $|u - \widetilde{u}_H|_{H^1(Q)}$ and estimators $|p_h|_{H^1(Q)}$ for adaptive and uniform refinements for different choices of h and H for the smooth solution (4.6).

As a second example we consider the unit square $Q = (0,1)^2$ and choose the right hand side f and the inital values $u(x,0) = u_0(x)$ and $\partial_t u(x,t)_{|t=0} = g(x)$ to be

$$f(x,t) = 2 \text{ for } (x,t) \in Q, \quad u_0(x) = 1 \text{ for } x \in (0,1), \quad g(x) = 0 \text{ for } x \in (0,1).$$

Note that the initial condition satisfies $u_0 \in L^2(0,1)$, but $u_0 \notin H^1_0(0,1)$, i.e., there is no compatibility with the homogeneous Dirichlet boundary conditions at $t = 0$. Therefore, we expect to see reduced orders of convergence. This is confirmed as shown in Fig. 10, where we obeserve a rate of $\mathcal{O}(H^{0.06})$ in the uniform refinement case, using $X_H = S^1_H(Q) \cap X$ and $Y_h = Y_H = S^2_H(Q) \cap Y$. Driving an adaptive refinement scheme with the Dörfler parameter $\theta = 0.9$ the rate could be increased to $\mathcal{O}(H^{0.1})$.

Remark 5.1 *As for the Laplace equation we can also use a boundary element least-squares formulation in the case of the wave equation [24]. In particular in the one-dimensional case one can prove the mesh condition $H > 2h$ in order to ensure stability, even for adaptive meshes.*



(a) initial, $M_X = 56$ (b) $h = H/2$, $M_X = 3624$ (c) $h = H/4$, $M_X = 2893$

Figure 9: Meshes from the adaptive refinement for the solution (4.6).

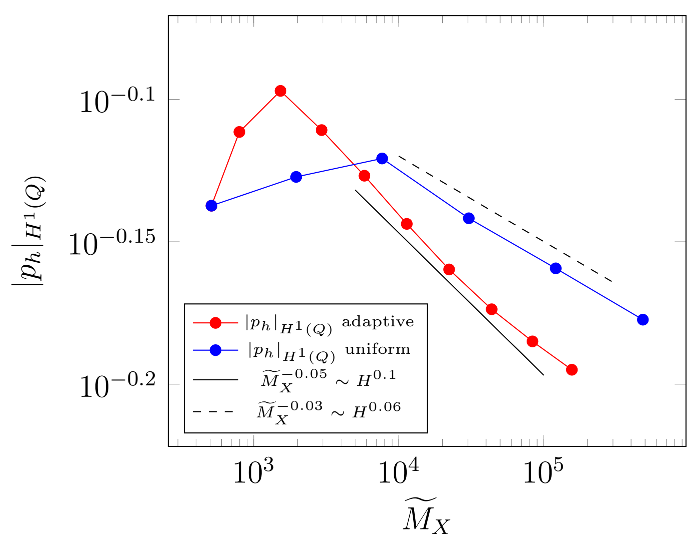


Figure 10: Error estimator for adaptive and uniform refinement in the case of a less regular solution violating the compatibility of initial and boundary conditions.

6 Conclusions

In this paper we have formulated and analyzed least-squares methods for the numerical solution of abstract operator equations $Bu = f$. Applications involve elliptic, parabolic, and hyperbolic problems, with the Poisson equation, the heat equation, and the wave equation as examples. While we assume that $B : X \to Y^*$ is an isomorphism, and B is given from the application, there is still some freedom in the choice of the underlying spaces X and Y. When considering the Laplace operator $B = -\Delta$, instead of $B : H^1_0(\Omega) \to H^{-1}(\Omega)$ we may also consider $Y = L^2(\Omega)$, implying $X = \{v \in H^1_0(\Omega) : \Delta v \in L^2(\Omega)\}$, or the other way around, i.e., an ultra weak variational formulation using $X = L^2(\Omega)$ and $Y = \{v \in H^1_0(\Omega) : \Delta v \in L^2(\Omega)\}$. It is obvious that in all of these cases we have to use appropriate inf-sup stable finite element spaces X_H and Y_h. But these approaches can be applied to the heat, and the wave equation, as well, see, e.g., [23].

In addition to those simple model problems can apply this methodology to rather general problems, including flow problems, nonlinear equations, etc. Other applications involve the least-squares formulations of boundary integral equations [24, 41].

In this paper we have considered the stability and error analysis of adaptive space-time least-squares finite element methods. It is clear that for an efficient solution of the resulting huge linear systems of algebraic equations we need to use appropriate preconditioned and parallel solution strategies. While the construction of these solvers was not within the scope of this paper, this will be done in future work. In particular, due to the structure of the linear system to be solved, we need to have preconditioners for A_h, and for the discrete Schur complement $\widetilde{S}_h = B_h^\top A_h^{-1} B_h$. Both matrices are symmetric and positive definite, independent of the particular choice of B. Possible solution strategies involve direct methods based on factorization as used in [29], or multigrid methods as considered in [22]. The use of efficient solution methods then also allows the numerical solution of partial differential equations in the four dimensional space-time domain, as we already did for optimal control problems, e.g., [28].

Acknowledgement: This work has been supported by the Austrian Science Fund (FWF) under the Grant Collaborative Research Center TRR361/F90: CREATOR Computational Electric Machine Laboratory.

References

[1] R. Andreev. Stability of sparse space-time finite element discretizations of linear parabolic evolution equations. *IMA J. Numer. Anal.*, 33(1):242–260, 2013.

[2] T. Apel, S. Nicaise, and J. Pfefferer. Adapted numerical methods for the Poisson equation with L^2 boundary data in nonconvex domains. *SIAM J. Numer. Anal.*, 55:1937–1957, 2017.

[3] P. Bochev, L. Demkowicz, J. Gopalakrishnan, and M. Gunzburger. Minimum residual and least squares finite element methods (Editorial). *Comput. Math. Appl.*, 68(11):1479, 2014.

[4] P. Bochev and M. Gunzburger. Least-squares finite element methods. In *International Congress of Mathematicians. Vol. III*, pages 1137–1162. Eur. Math. Soc., Zürich, 2006.

[5] P. Bochev and M. Gunzburger. Least-squares methods for hyperbolic problems. In *Handbook of numerical methods for hyperbolic problems*, volume 17 of *Handb. Numer. Anal.*, pages 289–317. Elsevier/North-Holland, Amsterdam, 2016.

[6] P. B. Bochev and M. D. Gunzburger. *Least-squares finite element methods*, volume 166 of *Applied Mathematical Sciences*. Springer, New York, 2009.

[7] J. Bramble, R. Lazarov, and J. Pasciak. A least-squares approach based on a discrete minus one inner product for first order systems. *Math, Comp.*, 66(219):935–955, 1997.

[8] J. H. Bramble, R. D. Lazarov, and J. E. Pasciak. Least-squares for second-order elliptic problems. *Comput. Methods Appl. Mech. Engrg.*, 152(1):195–210, 1998.

[9] S. C. Brenner. Finite element methods for elliptic distributed optimal control problems with pointwise state constraints (survey). In B. Acu, D. Danielli, M. Lewicka, A. Pati, R. V. Saraswathy, and M. Teboh-Ewungkem, editors, *Advances in Mathematical Sciences*, volume 21 of *Association for Women in Mathematics Series*, Cham, 2020. Springer.

[10] S. C. Brenner, J. Gedicke, and L.-Y. Sung. A symmetric interior penalty method for an elliptic distributed optimal control problem with pointwise state constraints. *Comput. Meth. Appl. Math.*, 23:565–589, 2023.

[11] S. C. Brenner and L. R. Scott. *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics*. Springer, New York, third edition, 2008.

[12] C. Carstensen, M. Feischl, M. Page, and D. Praetorius. Axioms of adaptivity. *Comput. Math. Appl.*, 67:1195–1253, 2014.

[13] C. Carstensen, D. Gallistl, and J. Gedicke. Justification of the saturation assumption. *Numer. Math.*, 134:1–25, 2016.

[14] W. Dahmen, H. Monsuur, and R. Stevenson. Least squares solvers for ill-posed PDEs that are conditionally stable. *ESAIM: Math. Model. Numer. Anal.*, 57:2227–2255, 2023.

[15] W. Dörfler. A convergent adaptive algorithm for Poisson's equation. *SIAM J. Numer. Anal.*, 33:1106–1124, 1996.

[16] W. Dörfler and R. H. Nochetto. Small data oscillation implies the saturation assumption. *Numer. Math.*, 91:1–12, 2002.

[17] W. Dörfler, C. Wieners, and D. Ziegler. Space-time discontinuous Galerkin methods for linear hyperbolic systems and the application to the forward problem in seismic imaging. In *Finite volumes for complex applications IX—methods, theoretical aspects, examples—FVCA 9, Bergen, Norway, June 2020*, volume 323 of *Springer Proc. Math. Stat.*, pages 477–485. Springer, Cham, 2020.

[18] S. Ferraz–Leite, C. Ortner, and D. Praetorius. Convergence of simple adaptive Galerkin schemes based on h-h/2 error estimators. *Numer. Math.*, 116:291–316, 2010.

[19] T. Führer, N. Heuer, and M. Karkulik. MINRES for second-order PDEs with singular data. *SIAM J. Numer. Anal.*, 60(3):1111–1135, 2022.

[20] T. Führer and M. Karkulik. Space-time least-squares finite elements for parabolic equations. *Comput. Math. Appl.*, 92:27–36, 2021.

[21] T. Führer. On a mixed FEM and a FOSLS with H^{-1} loads, 2023. arXiv:2210.14063.

[22] M. J. Gander and M. Neumüller. Analysis of a new space-time parallel multigrid algorithm for parabolic problems. *SIAM J. Sci. Comput.*, 38(4):A2173–A2208, 2016.

[23] J. Henning, D. Palitta, V. Simoncini, and K. Urban. An ultraweak space-time variational formulation for the wave equation: analysis and efficient numerical solution. *ESAIM Math. Model. Numer. Anal.*, 56(4):1173–1198, 2022.

[24] D. Hoonhout, R. Löscher, O. Steinbach, and C. Urzúa-Torres. Stable least-squares space-time boundary element methods for the wave equation. to be submitted, 2023.

[25] O. A. Ladyzhenskaya. *The boundary value problems of mathematical physics*, volume 49 of *Applied Mathematical Sciences*. Springer, 1985.

[26] U. Langer and O. Steinbach, editors. *Space-time methods—applications to partial differential equations*, volume 25 of *Radon Series on Computational and Applied Mathematics*. De Gruyter, Berlin, 2019.

[27] U. Langer, O. Steinbach, F. Tröltzsch, and H. Yang. Unstructured space-time finite element methods for optimal control of parabolic equations. *SIAM J. Sci. Comput.*, 43:A744–A771, 2021.

[28] U. Langer, O. Steinbach, and H. Yang. Robust space-time finite element error estimates for parabolic distributed optimal control problems with energy regularization, 2022. arXiv:2206.06455.

[29] U. Langer and M. Zank. Efficient direct space-time finite element solvers for parabolic initial-boundary value problems in anisotropic Sobolev spaces. *SIAM J. Sci. Comput.*, 43(4):A2714–A2736, 2021.

[30] R. Löscher, O. Steinbach, and M. Zank. Numerical results for an unconditionally stable space-time finite element method for the wave equation. In S. C. Brenner, E. Chung, A. Klawonn, F. Kwok, J. Xu, and J. Zou, editors, *Domain Decomposition Methods in Science and Engineering XXVI*, volume 145 of *Lecture Notes in Computational Science and Engineering*, pages 625–632, Cham, 2022. Springer.

[31] R. Löscher and O. Steinbach. Space-time finite element methods for distributed optimal control of the wave equation, 2022. arXiv:2211.02562.

[32] H. Monsuur, R. Stevenson, and J. Storn. Minimal residual methods in negative or fractional Sobolev norms, 2023. arXiv::2301.10484.

[33] M. Neumüller. *Space-Time Methods. Fast Solvers and Applications*, volume 20. Verlag der Technischen Universität Graz, 2013.

[34] M. Neumüller and O. Steinbach. Refinement of flexible space–time finite element meshes and discontinuous Galerkin methods. *Comput. Visual. Sci*, 14:189–205, 2011.

[35] I. Perugia, C. Schwab, and M. Zank. Exponential convergence of *hp*-time-stepping in space-time discretizations of parabolic PDEs. *ESAIM Math. Model. Numer. Anal.*, 57(1):29–67, 2023.

[36] A. Schafelner and P. S. Vassilevski. Numerical results for adaptive (negative norm) constrained first order system least squares formulations. *Comput. Math. Appl.*, 95:256–270, 2021.

[37] J. Schöberl. C++ Implementation of Finite Elements in NGSolve. Technical Report 30, Institute for Analysis and Scientific Computing, TU Wien, 2014.

[38] C. Schwab and R. Stevenson. Space-time adaptive wavelet methods for parabolic evolution problems. *Math. Comput.*, 78:1293–1318, 2009.

[39] O. Steinbach. *Numerical approximation methods for elliptic boundary value problems. Finite and boundary elements.* Springer, New York, 2008.

[40] O. Steinbach. Space-time finite element methods for parabolic problems. *Comput. Meth. Appl. Math.*, 15:551–566, 2015.

[41] O. Steinbach. An adaptive least squares boundary element method for elliptic boundary value problems, 2023. Berichte aus dem Institut für Angewandte Mathematik, Bericht 2023/1, TU Graz.

[42] O. Steinbach and M. Zank. Coercive space-time finite element methods for initial boundary value problems. *Electron. Trans. Numer. Anal.*, 52:154–194, 2020.

[43] O. Steinbach and M. Zank. A generalized inf-sup stable variational formulation for the wave equation. *J. Math. Anal. Appl.*, 505(1), 2022. Paper No. 125457, 24 pp.

[44] R. Stevenson and J. Westerdiep. Stability of Galerkin discretizations of a mixed space-time variational formulation of parabolic evolution equations. *IMA J. Numer. Anal.*, 41:28–47, 2021.

[45] V. Thomée. *Galerkin Finite Element Methods for Parabolic Problems*, volume 25 of *Springer Series in Computational Mathematics*. Springer, Berlin, 2006.

[46] K. Urban and A. T. Patera. An improved error bound for reduced basis approximation of linear parabolic problems. *Math. Comput.*, 83:1599–1615, 2014.