Chapter 13
Space-Time Finite Element Methods for Parabolic Evolution Problems with Variable Coefficients

Ulrich Langer, Martin Neumüller, and Andreas Schafelner

Abstract We introduce a completely unstructured, conforming space-time finite element method for the numerical solution of parabolic initial-boundary value problems with variable in space and time, possibly discontinuous diffusion coefficients. Discontinuous diffusion coefficients allow the treatment of moving interfaces. We show stability of the method and an a priori error estimate, including the case of local stabilizations which are important for adaptivity. To study the method in practice, we consider several typical model problems in one, two, and three spatial dimensions. The implementation of our space-time finite element method is fully parallelized with MPI. Extensive numerical tests were performed to study the convergence behavior of the stabilized space-time finite element discretization method and the scaling properties of the parallel AMG-preconditioned GMRES solver that we use to solve the huge system of space-time finite element equations.

13.1 Introduction

When we deal with the simulation of physical problems like transient diffusion problems, heat-conduction problems, or electromagnetic eddy current problems, the governing partial differential equations (PDEs) are often of parabolic type. Thus, the development of efficient numerical schemes for solving parabolic equations is of great importance. The standard approach to the numerical solution of parabolic PDEs uses some time-stepping method applied to the large-scale system of ordinary differential equations arising from a semi-discretization in the spatial variables, e.g., by means of the Finite Element Method (FEM); see, e.g., [42]. Another approach

U. Langer (✉) · M. Neumüller
Institute of Computational Mathematics, Johannes Kepler University Linz, Linz, Austria
e-mail: ulanger@numa.uni-linz.ac.at; neumueller@numa.uni-linz.ac.at

A. Schafelner
Doctoral Program "Computational Mathematics", Johannes Kepler University Linz, Linz, Austria
e-mail: andreas.schafelner@dk-compmath.jku.at

first discretizes the parabolic problem with respect to time by some time-stepping method, and then perform a discretization of the resulting elliptic problems in the spatial variables. This approach is sometimes also called Rothe's method, see, e.g., [25]. There are many papers on the more recent continuous or discontinuous Galerkin or Galerkin-Petrov (cG, cGP, dG, dGP) methods based on time-slices; see, e.g., [1, 4, 6, 14, 27, 31–34, 37] and the references therein. These methods are closely related to classical time-integration methods that can be solved in a time-stepping procedure. This is a sequential procedure that is not suited for parallel computing. In order overcome this drawback on massively parallel computers, time-parallel solvers, see, e.g. [12, 15, 18, 45], or time-parallel time-integration methods like PARAREAL [28] have been developed. An excellent historical overview of 50 years of time-parallel integration methods is given in [16]. An alternative approach consists in a full space-time discretization at once by treating time just as another space variable, i.e., we solve a problem with one dimension more. In fact, this approach to the numerical solution of transient problems in space and time simultaneously is not new, but becomes now a really hot topic in connection with the availability of massively parallel computers with many thousands of cores. Besides the overview paper [16] that focuses on time-parallel integration methods, we refer to [41] that provides an overview of the latest developments in this field with focus on completely unstructured space-time methods and simultaneous space-time adaptivity that treats time just as another variable.

In this paper, we will focus on space-time finite element methods that use really unstructured simplicial space-time meshes. The motivation behind this is that, for elliptic problems, there exist plenty of efficient and, most important, parallel solving methods. If we would be able to derive a stable discrete bilinear form, for which we can prove coercivity (ellipticity) with respect to some mesh-dependent norm in the space-time FE-space, then we can expect that we can efficiently solve the space-time problem fully in parallel as in the elliptic case. In this way, we can overcome the curse of sequentiality of the time-stepping methods. Another reason for the space-time approach is the fact that we are not restricted to a special structure of the mesh. This means that we can apply adaptive mesh refinement in space and time simultaneously. Last but not least, we can easily deal with moving interfaces and computational domains, where the coefficients of the PDE and/or the spatial domain $\Omega(t)$ depend on the time as well. Moreover, optimization problems constrained by a parabolic initial-boundary value problem lead to optimality conditions that can very efficiently be solved by space-time methods.

As already mentioned above, these advantages of space-time methods together with the common availability of massively parallel computers have led to a revival of space-time methods. This especially concerns space-time methods that are based on completely unstructured space-time meshes produced, e.g., by simultaneous space-time adaptivity; see [41] for a review of recent publications on this topic. For instance, Steinbach introduced a inf-sup-stable Petrov-Galerkin method [39], whereas Toulopoulos used bubble functions to stabilize a space-time finite element method [43]. In the context of using Isogeometric Analysis (IgA) as space-time discretization method, Langer et al. proposed an upwind-stabilized space-time

method for parabolic evolution equations [26]; see also PhD thesis [29] by Moore. Similar stabilized space-time finite element schemes have recently been developed in [5, 10, 30]. In [5], beside the upwind-stabilization scheme, Bank, Vassilevski and Zikatanov proposed and analyzed new EAFE (edge average finite element) schemes for parabolic convection-diffusion-reaction problems, whereas Devaud and Schwab [10] introduced upwind-stabilized schemes with mesh grading in time dealing with time singularities and hp schemes leading to exponential convergence.

The main aim of this paper is to generalize the results for the stabilized space-time scheme proposed in [26], where the authors use IgA for the discretization, and the corresponding stabilized space-time FE scheme considered by Moore [30] to the case of moving interfaces, i.e., t-dependent, discontinuous diffusion coefficients, and the possibility to choose local (element-wise) upwind test functions of the form $v_h + \theta_K h_K \partial_t v_h$ depending on the mesh-size h_K of an element K from the finite element mesh. This localization of the upwind-stabilization is very important for adaptivity that produces a family of shape regular meshes. We will use an unstructured conforming FEM to discretize the parabolic initial-boundary value problem, which we specify in the following. Let $Q := \Omega \times (0, T)$ be the space-time cylinder, with $\Omega \subset \mathbf{R}^d$, $d \in \{1, 2, 3\}$, being a sufficiently smooth and bounded spatial domain, and $T > 0$ being the final time. Furthermore, let $\Sigma := \partial\Omega \times (0, T)$, $\overline{\Sigma}_0 := \overline{\Omega} \times \{0\}$ and $\overline{\Sigma}_T := \overline{\Omega} \times \{T\}$ such that $\partial Q = \Sigma \cup \overline{\Sigma}_0 \cup \overline{\Sigma}_T$. Then we consider the following model problem that can formally be written as follows: Given f, g, ν and u_0, find u such that (s.t.)

$$\frac{\partial u}{\partial t}(x, t) - \mathrm{div}_x(\nu(x, t)\nabla_x u(x, t)) = f(x, t), \quad (x, t) \in Q, \tag{13.1}$$

$$u(x, t) = g(x, t) = 0, \quad (x, t) \in \Sigma, \tag{13.2}$$

$$u(x, 0) = u_0(x), \quad x \in \overline{\Omega}, \tag{13.3}$$

where the diffusion coefficient (reluctivity in electromagnetics) ν is a given uniformly positive and bounded coefficient function. The dependence of ν not only on space but also on time enables us to model moving interfaces. Note that we do not require ν to be smooth. In fact, we will admit discontinuities for ν. For simplicity, we assume homogeneous Dirichlet boundary conditions.

The paper is structured in the following way: In Sect. 13.2, we provide a space-time variational formulation of the parabolic initial boundary value problem (13.1)–(13.3), and we recall some existence, uniqueness and regularity results for weak solutions in appropriate space-time Sobolev spaces. Section 13.3 is devoted to the derivation and analysis of a new locally stabilized space-time finite element scheme. Moreover, we derive a priori discretization error estimates. In Sect. 13.4, we present four typical test cases for which we have performed extensive numerical studies, and we discuss the numerical results. Section 13.5 draws some conclusions, and provides an outlook on the future work.

13.2   The Space-Time Variational Formulation

In order to derive well-posed space-time variational formulations in space-time Sobolev spaces, we follow the classical approach developed in the monograph [24] by Ladyžhenskaya, Solonnikov and Uraltseva, and in the lecture notes [23] by Ladyžhenskaya. Let us first define the proper function spaces.

Definition 13.1 Let $L_2(Q)$ be the space of square integrable functions in the space-time cylinder Q. Then we define the following Sobolev (Hilbert) spaces

$$H_0^1(Q) = W_{2,0}^1(Q) := \{u \in L_2(Q) : \nabla u \in [L_2(Q)]^{d+1} \text{ and } u = 0 \text{ on } \Sigma\},$$

$$H^{1,0}(Q) = W_2^{1,0}(Q) := \{u \in L_2(Q) : \nabla_x u \in [L_2(Q)]^d\},$$

$$H_0^{1,0}(Q) = W_{2,0}^{1,0}(Q) := \{u \in H^{1,0}(Q) : u = 0 \text{ on } \Sigma\},$$

equipped with the usual scalar products and norms, as well as the Banach space

$$V_2(Q) := \{u \in H^{1,0}(Q) : |u|_Q < \infty\},$$

with the subspaces

$$V_{2,0}(Q) := \{u \in H_0^{1,0}(Q) : |u|_Q < \infty\},$$

$$V_2^{1,0}(Q) := \{u \in V_2(Q) : \lim_{\Delta t \to 0} \|u(\cdot, t + \Delta t) - u(\cdot, t)\|_{L_2(\Omega)} = 0, \text{ uniformly on } [0, T]\},$$

$$V_{2,0}^{1,0}(Q) := V_2^{1,0}(Q) \cap H_0^{1,0}(Q),$$

where the norm $|\cdot|_Q$ is defined by

$$|u|_{Q_t} := \max_{0 \le \tau \le t} \|u(\cdot, \tau)\|_{L_2(\Omega)} + \|\nabla_x u\|_{Q_t},$$

and $Q_t = \Omega \times (0, t)$ denotes a truncated space-time cylinder. Here, the appearing differential operators are defined as follows:

$$\nabla = (\nabla_x, \nabla_t)^T, \quad \nabla_x = (\partial_{x_1}, \ldots, \partial_{x_d})^T \quad \text{and} \quad \nabla_t = (\partial_t).$$

Multiplying the PDE (13.1) by a test function $v \in \hat{H}_0^1(Q) := \{v \in H_0^1(Q) : v = 0 \text{ on } \Sigma_T\}$, integrating over the complete space-time domain (cylinder) $Q = \Omega \times (0, T)$, integrating by parts with respect to time and space once, and incorporating the initial and boundary conditions, we immediately arrive at the following space-time variational formulation of the initial-boundary value problem (13.1)–(13.3): find a function $u \in H_0^{1,0}(Q)$ such that

$$a(u, v) = l(v), \quad \forall v \in \hat{H}_0^1(Q), \tag{13.4}$$

where the bilinear form $a(\cdot, \cdot)$ and the linear form $l(\cdot)$ are defined by the identities

$$a(u, v) = \int_Q (-u\partial_t v + v(x, t)\nabla_x u \nabla_x v) \, dx dt$$

and

$$l(v) = \int_\Omega u_0(x) \, v(x, 0) \, dx + \int_Q f v \, dx dt,$$

respectively. A solution u of the space-time variational (13.4) is called generalized (weak) solution of the parabolic initial-boundary value problem (13.1)–(13.3) in the space $u \in H_0^{1,0}(Q)$.

Under the assumptions that

$$u_0 \in L_2(\Omega) \quad \text{and} \quad f \in L_{2,1}(Q) := \{v : Q \to \mathbf{R} : \int_0^T \|v(\cdot, t)\|_{L_2(\Omega)} \, dt < \infty\}, \tag{13.5}$$

and that

$$0 < \underline{v} \le v(x, t) \le \overline{v}, \quad \text{for almost all } (x, t) \in Q, \tag{13.6}$$

with positive constants $\underline{v}$ and $\overline{v}$, the following theorem was proven by means of Galerkin's method and appropriate a priori estimates in [23]:

Theorem 13.1 ([23, Chapter III, Thm. 3.1]) *Under the conditions* (13.5) *and* (13.6)*, the space-time variational problem* (13.4) *has at least one generalized (weak) solution in* $H_0^{1,0}(Q)$.

Definition 13.2 ([23, Chapter III]) A generalized solution $u \in H_0^{1,0}(Q)$ of the space-time variational problem (13.4) is a called a generalized solution in $V_{2,0}^{1,0}(Q)$, if $u \in V_{2,0}^{1,0}(Q)$ and if it fulfills the energy-balance equation

$$\frac{1}{2}\|u(\cdot, t)\|_{L_2(\Omega)}^2 + \int_{Q_t} v(x, \tau)|\nabla_x u|^2 \, dx d\tau = \frac{1}{2}\|u(\cdot, 0)\|_{L_2(\Omega)}^2 + \int_{Q_t} f u \, dx d\tau.$$

and the identity

$$\int_\Omega u(x, t) \ v(x, t) \, dx - \int_\Omega u_0 v(x, 0) \, dx$$

$$+ \int_{Q_t} -u\partial_t v + v\nabla_x u \nabla_x v \, dx d\tau = \int_{Q_t} f \, v \, dx d\tau,$$

for all $v \in H_0^1(Q)$ and any $t \in (0, T)$.

Theorem 13.2 ([23, Chapter III, Thm. 3.2]) *If the assumptions* (13.5) *and* (13.6) *are fulfilled, then any generalized solution of the space-time variational problem* (13.4) *in* $H_0^{1,0}(Q)$ *is the generalized solution in* $V_{2,0}^{1,0}(Q)$ *and it is unique in* $H_0^{1,0}(Q)$.

Corollary 13.1 *If the assumptions* (13.5) *and* (13.6) *hold, then there exists a unique generalized solution* $u \in V_{2,0}^{1,0}(Q)$ *to the the space-time variational problem* (13.4).

Remark 13.1 For the case $\nu = 1$, $f \in L_2(Q)$ and $u_0 \in H_0^1(\Omega)$, Ladyžhenskaya proved in [23, Chapter III, Thm. 2.1] that the generalized solution u of (13.4) belongs to space $H_0^{\Delta,1}(Q) = W_{2,0}^{\Delta,1}(Q) = \{v \in H_0^1(Q) : \Delta_x v \in L_2(Q)\}$, and u continuously depends on t in the norm of the space $H_0^1(\Omega)$. If $\partial\Omega \in C^2$, then $u \in W_{2,0}^{2,1}(Q)$.

   More regularity results can already be found in the classical monograph [24] and in the more recent references [46] and [22]. The last reference provides an overview on maximal parabolic regularity results; see also [27]. The space-time finite element scheme that we are going to derive is consistent for solutions u of (13.4) that have at least piecewise partial time derivative $\partial_t u$ in L_2 and fluxes $\nu\nabla_x u$ in $H(\mathrm{div}_x) = \{v = (v_1, \ldots, v_d) \in [L_2(Q)]^d : \mathrm{div}_x v \in L_2(Q)\}$. This is ensured in the case of maximal parabolic regularity where $\partial_t u \in L_2(Q)$ and $\mathrm{div}_x(\nu\nabla_x u) \in L_2(Q)$, i.e. $\partial_t u - \mathrm{div}_x(\nu\nabla_x u) = f$ in $L_2(Q)$. We emphasize that we need this property only element-wise for deriving a consistent scheme.

13.3   The Space-Time Finite Element Scheme

From the previous section, we know that there exists a unique generalized solution of the initial-boundary value problem (13.1) in $H_0^{1,0}(Q) \cap V_{2,0}^{1,0}(Q)$ that may have more regularity due to more regularity of the data, see Remark 13.1 and the references mentioned above. The goal of this section is to derive a consistent and stable space-time finite element scheme with a discrete (mesh-dependent) bilinear form $a_h(\cdot, \cdot)$ that is coercive (elliptic) on the space-time finite element spaces and bounded on extended spaces with respect to appropriately chosen, mesh-dependent norms. These properties ensure existence and uniqueness of a finite element solution, and, together with appropriate interpolation respectively approximation error estimates, a priori discretization error estimates for sufficiently smooth solutions.

   Similar to Langer et al. in [26], we use special time-upwind test functions, but in contrast to [26] the time-upwind test functions are now locally scaled by the element mesh-size in order to handle adaptivity. First, we need a regular or, at least, a shape regular triangulation $\mathscr{T}_h$ of the space-time cylinder Q; see, e.g., [7, 9] for details. We now formally define this triangulation as $\mathscr{T}_h := \{K : K \subset Q, K \text{ open}\}$ such that $\overline{Q} = \bigcup_{K \in \mathscr{T}_h} \overline{K}$, with $K \cap K' = \emptyset$ for $K \neq K' \in \mathscr{T}_h$, and the usual conditions

imposed on a regular or a shape regular triangulation are fulfilled [7, 9]. On each of these elements K, we now define individual *time-upwind test functions*

$$v_{h,K}(x,t) := v_h(x,t) + \theta_K h_K \, \partial_t v_h(x,t), \text{ for all } (x,t) \in K,$$

where θ_K is a local positive parameter that will be defined later, and $h_K :=$ $diam(K)$. Here, v_h is some test function from a standard conforming space-time finite element space $V_{0h} = \{v \in C(\overline{Q}) : v(x_K(\cdot)) \in \mathbf{P}_p(\hat{K}), \forall K \in \mathscr{T}_h, v = 0 \text{ on } \overline{\Sigma} \cup \overline{\Sigma}_0\}$, where $x_K(\cdot)$ is the map from the reference element $\hat{K}$ to the finite element $K \in \mathscr{T}_h$, and $\mathbf{P}_p(\hat{K})$ is the space of polynomials of the degree p on the reference element $\hat{K}$. For simplicity, throughout this paper and, in particular, in our numerical experiments in Sect. 13.4, we use affine-linear mappings $x_K(\cdot)$ and simplicial elements. From now on, unless specified otherwise, all functions depend on both space and time variables. So, we can omit the arguments.

In this section, we will also use the following spaces:

$$H^{1,1}_{0,\underline{0}}(Q) := \{u \in L_2(Q) : \nabla_x u \in [L_2(Q)]^d, \ \partial_t u \in L_2(Q) \text{ and } u|_{\Sigma \cup \Sigma_0} = 0\},$$

$$H^{2,1}_{0,\underline{0}}(\mathscr{T}_h) := \{v \in H^{1,1}_{0,\underline{0}}(Q) : v|_K \in H^{2,1}(K), \ \forall K \in \mathscr{T}_h\},$$

$$W^1_\infty(\mathscr{T}_h) := \{v \in L_\infty(Q) : v|_K \in W^1_\infty(K), \ \forall K \in \mathscr{T}_h\},$$

where $H^{2,1}(K) := \{v \in L_2(K) : \partial_t v, \partial_{x_i} v, \partial_{x_i} \partial_{x_j} v \in L_2(K) \text{ and } \partial_t v \in L_2(K)\}$. For the sake of convenience, we now consider homogeneous initial conditions, i.e., $u_0 = 0$ on Ω. Furthermore, we assume that $v \in W^1_\infty(\mathscr{T}_h)$, and that the PDE has a sufficiently smooth solution u, e.g., $u \in H^{2,1}_{0,\underline{0}}(\mathscr{T}_h)$; cf. also our discussion in Sect. 13.2. Now we first multiply the PDE (13.1) by the space-time test function $v_{h,K}$, and then integrate over a single element K. Summing up over all elements and applying integration by parts in the principle term, we obtain

$$\sum_{K \in \mathscr{T}_h} \int_K (\partial_t u v_h + \theta_K h_K \partial_t u \partial_t v_h + \nu \nabla_x u \cdot \nabla_x v_h + \theta_K h_K \nu \nabla_x u \cdot \nabla_x (\partial_t v_h)) \, \mathrm{d}(x,t)$$

$$- \sum_{K \in \mathscr{T}_h} \int_{\partial K} (\nu \nabla_x u \cdot n_x v_h + \theta_K h_K \nu \nabla_x u \cdot n_x \partial_t v_h) \, \mathrm{d}s_{(x,t)} = l_h(v_h) \quad (13.7)$$

with the linear form

$$l_h(v_h) := \sum_{K \in \mathscr{T}_h} \int_K f \, (v_h + \theta_K h_K \partial_t v_h) \, \mathrm{d}(x,t). \quad (13.8)$$

For the exact solution u of (13.1), we know that the fluxes have to be continuous across the boundaries of the elements $K \in \mathscr{T}_h$. This observation means that

$$\sum_{K \in \mathscr{T}_h} \int_{\partial K} \nu \nabla_x u \cdot n_x v_h \, \mathrm{d}s_{(x,t)} = 0$$

for all test functions $v_h \in V_{0h}$. We mention that v_h is zero on Σ, and that n_x vanishes on $\overline{\Sigma} \cup \overline{\Sigma}_0$. Therefore, the first boundary term completely disappears from (13.7), but, in general, not the second term, since $\theta_K h_K$ varies from element to element. We now arrived at the consistency identity

$$a_h(u, v_h) = l_h(v_h), \quad \forall v_h \in V_{0h}, \tag{13.9}$$

that holds for a sufficiently smooth solution u, e.g., $u \in H_{0,\underline{0}}^{2,1}(\mathcal{T}_h)$, where the discrete (mesh-dependent) bilinear form $a_h(\cdot, \cdot)$ is defined by the identity

$$
\begin{aligned}
a_h(u, v_h) := & \sum_{K \in \mathcal{T}_h} \int_K (\partial_t u \, v_h + \theta_K h_K \, \partial_t u \, \partial_t v_h) \, \mathrm{d}(x, t) \\
& + \sum_{K \in \mathcal{T}_h} \int_K (\nu \nabla_x u \cdot \nabla_x v_h + \theta_K h_K \, \nu \nabla_x u \cdot \nabla_x(\partial_t v_h)) \, \mathrm{d}(x, t) \\
& - \sum_{K \in \mathcal{T}_h} \int_{\partial K} \theta_K h_K \, \nu \nabla_x u \cdot n_x \, \partial_t v_h \, \mathrm{d}s_{(x,t)},
\end{aligned} \tag{13.10}
$$

and the linear form $l_h(\cdot)$ is defined by (13.8), with given $\nu \in W_\infty^1(\mathcal{T}_h)$ and $f \in L_2(Q)$.

Remark 13.2 We can derive a scheme that is equivalent to (13.10). In particular, instead of applying integration by parts on both principal terms, we only apply it to the first principal term and keep the second. Hence, we obtain another consistency identity

$$\tilde{a}_h(u, v_h) = l_h(v_h), \quad \forall v_h \in V_{0h},$$

that holds for a solution u of (13.4) that only belongs to $H_{0,\underline{0}}^{L,1}(\mathcal{T}_h) := \{u \in L_2(Q) : (\partial_t u)|_K \in L_2(K), (\nu \nabla_x u)|_K \in H(\mathrm{div}_x, K), u = 0 \text{ on } \overline{K} \cap (\Sigma \cup \Sigma_0) \, \forall K \in \mathcal{T}_h\}$, where

$$
\begin{aligned}
\tilde{a}_h(u, v_h) := & \sum_{K \in \mathcal{T}_h} \int_K (\partial_t u \, v_h + \theta_K h_K \, \partial_t u \, \partial_t v_h) \, \mathrm{d}(x, t) \\
& + \sum_{K \in \mathcal{T}_h} \int_K (\nu \nabla_x u \cdot \nabla_x v_h - \theta_K h_K \, \mathrm{div}_x(\nu \nabla_x u) \partial_t v_h) \, \mathrm{d}(x, t)
\end{aligned}
$$

$$\tag{13.11}$$

with given $\nu \in W_\infty^1(\mathcal{T}_h)$ and $f \in L_2(Q)$, and l_h as in (13.8). We mention that $u \in H_{0,\underline{0}}^{L,1}(\mathcal{T}_h)$ is ensured in the case of maximal parabolic regularity where u belongs $H^{L,1}(Q) := \{v \in H^1(Q) : Lu := \mathrm{div}_x(\nu \nabla_x u) \in L_2(Q)\}$.

Remark 13.3 If the test functions $v_h \in V_{0h}$ are continuous and piecewise linear ($p = 1$), then the term in (13.10) containing $\nabla_x(\partial_t v_h)$ vanishes in all elements $K \in \mathcal{T}_h$, since it only contains mixed second order derivatives of the test functions.

Now we look for a Galerkin approximation $u_h \in V_{0h}$ to the generalized solution u of the initial boundary value problem (13.1)–(13.3) using the variational identity (13.9), i.e., find $u_h \in V_{0h}$ such that

$$a_h(u_h, v_h) = l_h(v_h), \quad \forall v_h \in V_{0h}, \tag{13.12}$$

with $a_h(\cdot, \cdot)$ and $l_h(\cdot)$ as defined above by (13.10) and (13.8), respectively. In Sect. 13.2, we already showed existence and uniqueness of a weak solution to the initial-boundary value problem (13.1)–(13.3). However, our finite element scheme (13.12) is based on a mesh-dependent bilinear form $a_h(\cdot, \cdot)$. Thus, we have to investigate the stability of the space-time finite element scheme. More precisely, we will show ellipticity of the bilinear form $a_h(\cdot, \cdot) : V_{0h} \times V_{0h} \to \mathbf{R}$ w.r.t. the mesh-dependent norm

$$\|v_h\|_h^2 := \sum_{K \in \mathcal{T}_h} \left[\|\nu^{1/2} \nabla_x v_h\|_{L_2(K)}^2 + \theta_K h_K \|\partial_t v_h\|_{L_2(K)}^2 \right] + \frac{1}{2} \|v_h\|_{L_2(\Sigma_T)}^2. \tag{13.13}$$

This implies existence and uniqueness of the finite element solution $u_h \in V_{0h}$ of (13.12). For the following derivations, we assume that our triangulation $\mathcal{T}_h$ of Q is *shape regular* such that local approximation error estimates are available [7, 9]. A shape regular triangulation $\mathcal{T}_h$ of Q is called quasi-uniform, if there exists a constant c_u such that

$$h_K \le h \le c_u h_K, \quad \text{for all } K \in \mathcal{T}_h,$$

where $h = \max_{K \in \mathcal{T}_h} h_K$. Moreover, we introduce localized bounds for our coefficient function ν, i.e.,

$$\underline{\nu}_K \le \nu(x, t) \le \overline{\nu}_K, \quad \text{for almost all } (x, t) \in K \text{ and for all } K \in \mathcal{T}_h, \tag{13.14}$$

where $\underline{\nu}_K \ge \underline{\nu}$ and $\overline{\nu}_K \le \overline{\nu}$ are positive constants on every $K \in \mathcal{T}_h$. In the following, we need some inverse inequalities for functions from finite element spaces.

Lemma 13.1 *There exist generic positive constants $c_{I,1}$ and $c_{I,2}$ such that*

$$\|v_h\|_{L_2(\partial K)} \le c_{I,1} h_K^{-1/2} \|v_h\|_{L_2(K)}, \tag{13.15}$$

$$\|\nabla v_h\|_{L_2(K)} \le c_{I,2} h_K^{-1} \|v_h\|_{L_2(K)} \tag{13.16}$$

for all $v_h \in V_{0h}$ and for all $K \in \mathcal{T}_h$.

Proof For (13.15); see e.g. [11, 35], and for (13.16) see e.g. [7, 9, 11]. ☐

From $\nabla = (\nabla_x, \partial_t)^T$ and (13.16), we can immediately deduce

$$\|\partial_t v_h\|_{L_2(K)} \le c_{I,2} h_K^{-1} \|v_h\|_{L_2(K)}. \tag{13.17}$$

The above inequalities hold for the standard L_2-norms. However, we will also need such a result in some scaled norm.

Lemma 13.2 *Let $v \in W_\infty^1(\mathcal{T}_h)$ be a given, uniformly positive function. Then*

$$\|v\|_{L_2^\nu(K)}^2 = \int_K \nu(x, t) \, |v(x, t)|^2 \, \mathrm{d}(x, t)$$

is a norm, and the inverse estimate

$$\|\partial_t v_h\|_{L_2^\nu(K)} \le \|\nabla v_h\|_{L_2^\nu(K)} \le c_{I,\nu} h_K^{-1} \|v_h\|_{L_2^\nu(K)}$$

holds for all $v_h \in V_{0h}$ and for all $K \in \mathcal{T}_h$, with $c_{I,\nu} := (\overline{\nu}_K/\underline{\nu}_K)^{1/2} c_{I,2}$.

Proof See [36]. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ □

We note that, in practical applications, it is clear that $1 \le \overline{\nu}_K/\underline{\nu}_K$ is close to 1. Below, we will also need the estimate

$$\|\partial_t \partial_{x_i} v_h\|_{L_2^\nu(K)} \le c_{I,\nu} h_K^{-1} \|\partial_{x_i} v_h\|_{L_2^\nu(K)}, \tag{13.18}$$

which obviously holds for all $v_h \in V_{0h}$ and for all $K \in \mathcal{T}_h$. Moreover, we need the following inverse inequality.

Lemma 13.3 *Let $v \in W_\infty^1(\mathcal{T}_h)$ be a given uniformly positive function. Let $W_h|_K := \{w_h : w_h = \nabla_x v_h, v_h \in V_{0h}|_K\}$. Then the inverse estimate*

$$\|\mathrm{div}_x(\nu w_h)\|_{L_2(K)} \le c_{I,3} h_K^{-1} \|\nu w_h\|_{L_2(K)}, \forall w_h \in W_h|_K \tag{13.19}$$

holds, where $c_{I,3}$ is a positive constant that is independent of h_K.

Proof First, we know that $V_{0h}|_K$ is a finite-dimensional space spanned by the local shape functions $\{p^{(i)}\}_{i \in \overline{\omega}_K}$, where $\overline{\omega}_K$ is the index set of local degrees of freedom. Hence, the space $W_h|_K$ is also finite-dimensional and spanned by the generating system $\{\nabla_x p^{(i)}\}_{i \in \overline{\omega}_K}$. Moreover, for a fixed v, each product $z_h := v \, w_h$ can be represented by means of a non-necessary unique linear combination $\{v \, \nabla_x p^{(i)}\}_{i \in \overline{\omega}_K}$ on K. We denote this space by $Z_h(K) := \mathrm{span}_{i \in \overline{\omega}_K} \{v \, \nabla_x p^{(i)}\}$. Using Cauchy's inequality, we obtain

$$\|\mathrm{div}_x z_h\|_{L_2(K)}^2 = \int_K |\mathrm{div}_x z_h|^2 \, \mathrm{d}(x, t) = \int_K |\sum_{i=1}^d \partial_{x_i} z_{h,i}|^2 \, \mathrm{d}(x, t)$$

$$\le d \int_K \sum_{i=1}^d |\partial_{x_i} z_{h,i}|^2 \, \mathrm{d}(x, t) = d \sum_{i=1}^d \|\partial_{x_i} z_{h,i}\|_{L_2(K)}^2,$$

for all $z_h \in Z_h(K)$. Now, by a simple scaling argument, we can estimate each element in the sum, and obtain

$$d \sum_{i=1}^{d} \|\partial_{x_i} z_{h,i}\|^2_{L_2(K)} \leq d \sum_{i=1}^{d} C^2 h_K^{-2} \|z_{h,i}\|^2_{L_2(K)}$$

$$= dC^2 h_K^{-2} \|z_h\|^2_{L_2(K)},$$

where C is a positive constant that is independent of h_K. Taking the square root and setting $c_{I,3} := C\sqrt{d}$ closes the proof. $\qquad\square$

Lemma 13.3 gives information how the two norms in (13.19) scale w.r.t. the mesh-size h_K. However, the estimate (13.19) is not sharp w.r.t. the constant.

Lemma 13.4 *Let the assumptions of Lemma 13.3 hold. Then*

$$\|\text{div}_x(\nu w_h)\|_{L_2(K)} \leq c_{opt} \|\nu w_h\|_{L_2(K)}, \forall w_h \in W_h|_K,$$

with $c^2_{opt} = \sup_{0 \neq z_h \in Z_h(K)} \frac{\|\text{div}_x(z_h)\|^2_{L_2(K)}}{\|z_h\|^2_{L_2(K)}} \leq C^2 d h_K^{-1}$.

Proof See [36]. $\qquad\square$

Remark 13.4 We note that the constant c_{opt} in Lemma 13.4 is not only optimal, but also computable. If $z_h \in Z_h(K)$, then, by definition, we have the representation

$$z_h(x, t) = \sum_{j \in \tilde{\omega}_K} \tilde{z}_j \tilde{q}^{(j)}, \tag{13.20}$$

where $\{\tilde{q}^{(j)}\}_{j \in \tilde{\omega}_K}$ forms some basis of $Z_h(K)$. Once some basis is chosen, we can rewrite

$$\|z_h\|^2_{L_2(K)} = (z_h, z_h)_{L_2(K)} \quad \text{and} \quad \|\text{div}_x z_h\|^2_{L_2(K)} = \underbrace{(\text{div}_x z_h, \text{div}_x z_h)_{L_2(K)}}_{=:b(z_h, z_h)}$$

in the form

$$(y_h, z_h)_{L_2(K)} = (M_h \underline{y}, \underline{z}) \quad \text{and} \quad b(y_h, z_h) = (B_h \underline{y}, \underline{z}),$$

with the element mass matrix $M_h = (M_{ij} = (\tilde{q}^{(j)}, \tilde{q}^{(i)})_{L_2(K)})_{i,j \in \tilde{\omega}_K}$ and the element div_x-stiffness matrix $B_h = (B_{ij} = b(\tilde{q}^{(j)}, \tilde{q}^{(i)})_{L_2(K)})_{i,j \in \tilde{\omega}_K}$, respectively. Here, the vectors $\underline{y}$ and $\underline{z}$ are the vector of coefficients in the representation (13.20) w.r.t. the chosen basis $\{\tilde{q}^{(j)}\}_{j \in \tilde{\omega}_K}$. Using this matrix representation, we immediately get

$$c^2_{opt} = \sup_{0 \neq z_h \in Z_h(K)} \frac{\|\text{div}_x(z_h)\|^2_{L_2(K)}}{\|z_h\|^2_{L_2 QT(K)}} = \sup_{\underline{z} \in \mathbf{R}^{N_K} = |\tilde{\omega}_K|} \frac{(B_h \underline{z}, \underline{z})_{\ell_2}}{(M_h \underline{z}, \underline{z})_{\ell_2}}.$$

Hence, c_{opt}^2 is the *largest eigenvalue* of the generalized eigenvalue problem

$$B_h \underline{z} = \lambda M_h \underline{z},$$

that can easily be computed.

Now, we are in the position to proof the following coercivity lemma that is crucial for our approach.

Lemma 13.5 *There exits a positive constant μ_c such that*

$$a_h(v_h, v_h) \geq \mu_c \|v_h\|_h^2, \quad \forall v_h \in V_{0h},$$

with $\mu_c = \min_{K \in \mathscr{T}_h} \{1 - c_{I,3} \sqrt{\frac{\bar{v}_K \theta_K}{4h_K}}\} \geq \frac{1}{2}$ provided that $\theta_K \leq \frac{h_K}{c_{I,3}^2 \bar{v}_K}$. For instance, $\theta_K = \frac{h_K}{c_{I,3}^2 \bar{v}_K}$ yields $\mu_c = \frac{1}{2}$.

Proof Integration by parts in the last term of (13.10) yields

$$a_h(v_h, v_h) = \sum_{K \in \mathscr{T}_h} \left[\int_K \frac{1}{2} \partial_t(v_h^2) \, d(x, t) + \theta_K h_K \|\partial_t v_h\|_{L_2(K)}^2 + \int_K v |\nabla_x v_h|^2 \, d(x, t) \right.$$

$$\left. - \int_K \theta_K h_K \, \text{div}_x(v \nabla_x v_h) \partial_t v_h \, d(x, t) \right].$$

Now using Gauss' theorem and the facts that v_h is continuous across the element boundary and that $n_t = 0$ on Σ, we obtain

$$a_h(v_h, v_h) = \frac{1}{2} \left(\|v_h\|_{L_2(\Sigma_T)}^2 - \|v_h\|_{L_2(\Sigma_0)}^2 \right) + \sum_{K \in \mathscr{T}_h} \left[\theta_K h_K \|\partial_t v_h\|_{L_2(K)}^2 \right.$$

$$\left. + \int_K v |\nabla_x v_h|^2 \, d(x, t) - \int_K \theta_K h_K \, \text{div}_x(v \nabla_x v_h) \partial_t v_h \, d(x, t) \right].$$

The first, second and third term already appear in the definition of our mesh-dependent norm (13.13). It remains to estimate the last term. Using the Cauchy-Schwarz inequality, Lemma 13.3, and a scaled Young's inequality, we arrive at the estimate

$$|\theta_K h_K \int_K \text{div}_x(v \nabla_x v_h) \, \partial_t v_h \, d(x, t)|$$

$$\leq c_{I,3} \left(\frac{\varepsilon \bar{v}_K \theta_K}{2h_K} \|\nabla_x v_h\|_{L_2(K)}^2 + \frac{1}{2\varepsilon} \theta_K h_K \|\partial_t v_h\|_{L_2(K)}^2 \right).$$

This estimate and the fact that $v_h = 0$ on Σ_0 immediately yield the estimate

$$a_h(v_h, v_h) \geq \frac{1}{2}\|v_h\|^2_{L_2(\Sigma_T)} + \sum_{K \in \mathscr{T}_h} \Big[\big(1 - \frac{c_{I,3}}{2\varepsilon}\big)\theta_K h_K \|\partial_t v_h\|^2_{L_2(K)}$$

$$+ \big(1 - \varepsilon\frac{c_{I,3}\overline{\nu}_K\theta_K}{2h_K}\big)\|\nabla_x v_h\|^2_{L_2^\nu(K)}\Big].$$

We now choose $\varepsilon = \sqrt{h_K/(\theta_K \overline{\nu}_K)}$ and obtain

$$a_h(v_h, v_h) \geq \min_{K \in \mathscr{T}_h} \left(1 - c_{I,3}\sqrt{\frac{\theta_K \overline{\nu}_K}{4h_K}}\right)$$

$$\times \left(\sum_{K \in \mathscr{T}_h} \big[\|\nabla_x v_h\|^2_{L_2^\nu(K)} + \theta_K h_K \|\partial_t v_h\|^2_{L_2(K)}\big] + \frac{1}{2}\|v_h\|^2_{L_2(\Sigma_T)}\right)$$

$$\geq \mu_c \|v_h\|^2_h,$$

which concludes the proof. □

Remark 13.5 The above proof holds for any polynomial degree $p \geq 1$ and any fixed, uniformly positive $\nu \in L_\infty(Q)$. However, for the special case $p = 1$ and $\nu|_K = const$, the proof is trivial since $\partial_t(\nabla_x v_h) \equiv 0$ and $\nu|_K \Delta_x v_h \equiv 0$. Hence, the identity

$$a_h(v_h, v_h) = \sum_{K \in \mathscr{T}_h} \frac{1}{2}\int_{\partial K} v_h^2 n_t \, ds_{(x,t)} + \theta_K h_K \|\partial_t v_h\|^2_{L_2(K)} + \|\nu^{1/2}\nabla_x v_h\|^2_{L_2(K)} = \|v_h\|^2_h$$

holds, i.e., $\mu_c = 1$. Thus, for this special case, the choice of θ_K has no influence on the coercivity (ellipticity) of the space-time finite element method.

Lemma 13.5 already ensures uniqueness of the finite element solution $u_h \in V_{0h}$. Furthermore, the space V_{0h} is finite-dimensional. Hence, uniqueness implies existence of finite element solution $u_h \in V_{0h}$ of (13.9). For the special case of uniform meshes and uniform θ, i.e., $h_K = h$ and $\theta_K = \theta$ for all $K \in \mathscr{T}_h$, and $\nu \equiv 1$, a proof for ellipticity with a *mesh-independent* constant was done by Langer, Moore and Neumüller in [26] and by Moore in [29]. For a second special case, where θ_K vanishes, i.e., $\theta_K = \theta = 0$ for all $K \in \mathscr{T}_h$, Steinbach has shown existence and uniqueness of solutions to both the continuous and discrete version of (13.9) on the basis of Banach-Nečas-Babuška's theorem in [39]. In addition, both papers also include a priori discretization error estimates, where Steinbach's estimate is based on a discrete inf-sup condition. To derive an a priori error estimate w.r.t. the mesh dependent norm (13.13), we need to show that our bilinear form $a_h(\cdot, \cdot)$ is uniformly

bounded on $V_{0h,*} \times V_{0h}$, where $V_{0h,*} = H_0^{1,0}(Q) \cap H^2(\mathcal{T}_h) + V_{0h}$ is equipped with the norm

$$\|v\|_{h,*}^2 = \frac{1}{2}\|v\|_{L_2(\Sigma_T)}^2 + \sum_{K \in \mathcal{T}_h} \left[\theta_K h_K \|\partial_t v\|_{L_2(K)}^2 + \|\nabla_x v\|_{L_2^\nu(K)}^2 \right.$$
$$\left. + (\theta_K h_K)^{-1}\|v\|_{L_2(K)}^2 + \theta_K h_K |v|_{H^2(K)}^2 \right] \quad (13.21)$$

Moreover, we will make use of the following scaled trace inequality.

Lemma 13.6 *There exists a positive constants $c_{Tr} > 0$ such that*

$$\|v\|_{L_2(\partial K)}^2 \le 2c_{Tr}^2 h_K^{-1} \left(\|v\|_{L_2(K)}^2 + h_K^2 \|\nabla v\|_{L_2(K)}^2 \right) \quad (13.22)$$

for all $v \in H^1(K)$ and for all $K \in \mathcal{T}_h$.

Proof See, e.g., [35]. $\square$

Lemma 13.7 *The bilinear form $a_h(\cdot, \cdot)$ is uniformly bounded on $V_{0h,*} \times V_{0h}$, i.e.,*

$$|a_h(u, v_h)| \le \mu_b \|u\|_{h,*} \|v_h\|_h, \quad \forall u \in V_{0h,*}, v_h \in V_{0h},$$

where $\mu_b = \max_{K \in \mathcal{T}_h} \left\{ 2(1+\theta_K h_K^{-1} c_{Tr}^2 \overline{\nu}_K^2 \underline{\nu}_K^{-1}), 2c_{Tr}^2 \overline{\nu}_K^2, 2+c_{I,1}^2, 1+(c_{I,\nu}\theta_K)^2 \right\}^{1/2}$ that is uniformly bounded provided that $\theta_K = \mathcal{O}(h_K)$.

Proof We will estimate the bilinear form (13.10) term by term. Since $V_{0h} \subset H_{0,0}^{1,1}(Q)$, we can apply integration by parts and the Cauchy-Schwarz inequality to the first term, and obtain

$$\left| \sum_{K \in \mathcal{T}_h} \int_K \partial_t u v_h \, d(x, t) \right| \le \sum_{K \in \mathcal{T}_h} \left[\left((\theta_K h_K)^{-1}\|u\|_{L_2(K)}^2\right)^{1/2} \left(\theta_K h_K \|\partial_t v_h\|_{L_2(K)}^2\right)^{1/2} \right]$$
$$+ \left(\|u\|_{L_2(\Sigma_T)}^2\right)^{1/2} \left(\|v_h\|_{L_2(\Sigma_T)}^2\right)^{1/2}.$$

For the second and third term, applying the Cauchy-Schwarz inequality to each term of the sum, we immediately get the estimates

$$\left| \theta_K h_K \int_K \partial_t u \partial_t v_h \, d(x, t) \right| \le \left(\theta_K h_K \|\partial_t u\|_{L_2(K)}^2\right)^{1/2} \left(\theta_K h_K \|\partial_t v_h\|_{L_2(K)}^2\right)^{1/2},$$

$$\left| \int_K \nu \nabla_x u \nabla_x v_h \, d(x, t) \right| \le \left(\|\nabla_x u\|_{L_2^\nu(K)}^2\right)^{1/2} \left(\|\nabla_x v_h\|_{L_2^\nu(K)}^2\right)^{1/2},$$

respectively. For the fourth term, we use again Cauchy-Schwarz' inequality, the inverse estimate (13.18), and obtain

$$\left| \theta_K h_K \int_K \nu \nabla_x u \nabla_x (\partial_t v_h) \, d(x,t) \right|$$

$$\leq \left(\|\nabla_x u\|^2_{L_2^\nu(K)} \right)^{1/2} \left((\theta_K h_K)^2 \sum_{i=1}^d c^2_{I,\nu} h_K^{-2} \|\partial_{x_i} v_h\|^2_{L_2^\nu(K)} \right)^{1/2}$$

$$= \left(\|\nabla_x u\|^2_{L_2^\nu(K)} \right)^{1/2} \left((c_{I,\nu} \theta_K)^2 \|\nabla_x v_h\|^2_{L_2^\nu(K)} \right)^{1/2}.$$

For the last term, we apply Cauchy-Schwarz' inequality, the trace inequalities (13.15) and (13.22), and get

$$\left| \theta_K h_K \int_{\partial K} \nu \nabla_x u \cdot n_x \partial_t v_h \, ds_{(x,t)} \right|$$

$$\leq \left(2\theta_K \overline{\nu}_K^2 c^2_{Tr} h_K^{-1} \left[\|\nabla_x u\|^2_{L_2(K)} + h_K^2 \sum_{i=1}^d \|\nabla \partial_{x_i} u\|^2_{L_2(K)} \right] \right)^{1/2} \left(\theta_K h_K c^2_{I,1} \|\partial_t v_h\|^2_{L_2(K)} \right)^{1/2}$$

$$\leq \left(2\theta_K c^2_{Tr} \frac{\overline{\nu}_K^2}{\underline{\nu}_K} h_K^{-1} \|\nabla_x u\|^2_{L_2^\nu(K)} + 2c^2_{Tr} \overline{\nu}_K^2 \theta_K h_K |u|^2_{H^2(K)} \right)^{1/2} \left(c^2_{I,1} \theta_K h_K \|\partial_t v_h\|^2_{L_2(K)} \right)^{1/2}.$$

Now combining the above estimates, applying Cauchy's inequality and gathering all similar items, we finally arrive at the estimate

$$|a_h(u, v_h)|$$

$$\leq \left(\|u\|^2_{L_2(\Sigma_T)} + \sum_{K \in \mathcal{T}_h} \left[\theta_K h_K \|\partial_t u\|^2_{L_2(K)} + 2(1 + \theta_K c^2_{Tr} \frac{\overline{\nu}_K^2}{\underline{\nu}_K} h_K^{-1}) \|\nabla_x u\|^2_{L_2^\nu(K)} \right. \right.$$

$$\left. \left. + (\theta_K h_K)^{-1} \|u\|^2_{L_2(K)} + 2c^2_{Tr} \overline{\nu}_K^2 \theta_K h_K |u|^2_{H^2(K)} \right] \right)^{1/2}$$

$$\times \left(\|v_h\|^2_{L_2(\Sigma_T)} + \sum_{K \in \mathcal{T}_h} \left[(2 + c^2_{I,1}) \theta_K h_K \|\partial_t v_h\|^2_{L_2(K)} \right. \right.$$

$$\left. \left. + (1 + (c_{I,1}\theta_K)^2) \|\nabla_x v_h\|^2_{L_2^\nu(K)} \right] \right)^{1/2}$$

$$\leq \quad \mu_b \|u\|_{h,*} \|v_h\|_h,$$

with $\mu_b := \max_{K \in \mathcal{T}_h} \{ 2(1 + \theta_K h_K^{-1} c^2_{Tr} \frac{\overline{\nu}_K^2}{\underline{\nu}_K}), 2c^2_{Tr} \overline{\nu}_K^2, 2 + c^2_{I,1}, 1 + (c_{I,\nu}\theta_K)^2 \}^{1/2}$. Choosing now $\theta_K = \mathcal{O}(h_K)$ ensures the uniform boundedness of the constant μ_b.

□

Remark 13.6 Choosing θ_K as in Lemma 13.5, i.e., $\theta_K = h_K/(c_{I,3}^2\overline{\nu}_K)$, we obtain $\mu_c = 1/2$ and $\mu_b = \max_{K\in\mathscr{T}_h}\left\{2(1 + \frac{\overline{\nu}_K c_{Tr}^2}{\underline{\nu}_K c_{I,3}^2}), 2c_{Tr}^2\overline{\nu}_K^2, 2 + c_{I,1}^2, 1 + (\frac{c_{I,\nu}h_K}{c_{I,3}^2\overline{\nu}_K})^2\right\}^{1/2}$.

Remark 13.7 If we consider the bilinear form from Remark 13.2, we can derive an equivalent statement, but in a different norm $\|\cdot\|_{h,*}$ defined as

$$\|v\|_{h,*}^2 = \frac{1}{2}\|v\|_{L_2(\Sigma_T)}^2 + \sum_{K\in\mathscr{T}_h}\left[\theta_K h_K\|\partial_t v\|_{L_2(K)}^2 + \|\nabla_x v\|_{L_2^\nu(K)}^2\right.$$
$$\left. + (\theta_K h_K)^{-1}\|v\|_{L_2(K)}^2 + \theta_K h_K\|\text{div}_x(\nu\nabla_x v)\|_{L_2(K)}^2\right].$$

By the same arguments as in the proof above, we estimate the first three terms in (13.11) by

$$\left|\sum_{K\in\mathscr{T}_h}\int_K \partial_t u v_h \,\mathrm{d}(x,t)\right| \leq \sum_{K\in\mathscr{T}_h}\left[\left((\theta_K h_K)^{-1}\|u\|_{L_2(K)}^2\right)^{1/2}\left(\theta_K h_K\|\partial_t v_h\|_{L_2(K)}^2\right)^{1/2}\right]$$
$$+\left(\|u\|_{L_2(\Sigma_T)}^2\right)^{1/2}\left(\|v_h\|_{L_2(\Sigma_T)}^2\right)^{1/2},$$

$$\left|\theta_K h_K\int_K \partial_t u \partial_t v_h \,\mathrm{d}(x,t)\right| \leq \left(\theta_K h_K\|\partial_t u\|_{L_2(K)}^2\right)^{1/2}\left(\theta_K h_K\|\partial_t v_h\|_{L_2(K)}^2\right)^{1/2},$$

$$\left|\int_K \nu\nabla_x u \nabla_x v_h \,\mathrm{d}(x,t)\right| \leq \left(\|\nabla_x u\|_{L_2^\nu(K)}^2\right)^{1/2}\left(\|\nabla_x v_h\|_{L_2^\nu(K)}^2\right)^{1/2}.$$

For the fourth term, we just apply the Cauchy-Schwarz inequality to each term of the sum to obtain

$$\left|\theta_K h_K\int_K \text{div}_x(\nu\nabla_x u)\,\partial_t v_h \,\mathrm{d}(x,t)\right| \leq \left(\theta_K h_K\|\text{div}_x(\nu\nabla_x u)\|_{L_2(K)}^2\right)^{1/2}\left(\theta_K h_K\|\partial_t v_h\|_{L_2(K)}^2\right)^{1/2}.$$

Now, combining the above estimates, applying Cauchy's inequality and reordering the terms, we finally obtain the estimate

$$|a_h(u,\,v_h)|$$
$$\leq \left(\|u\|_{L_2(\Sigma_T)}^2 + \sum_{K\in\mathscr{T}_h}\left[\theta_K h_K\|\partial_t u\|_{L_2(K)}^2 + \|\nabla_x u\|_{L_2^\nu(K)}^2\right.\right.$$
$$\left.\left. + (\theta_K h_K)^{-1}\|u\|_{L_2(K)}^2 + \theta_K h_K\|\text{div}_x(\nu\nabla_x u)\|_{L_2(K)}^2\right]\right)^{1/2}$$
$$\times\left(\|v_h\|_{L_2(\Sigma_T)}^2 + \sum_{K\in\mathscr{T}_h}\left[3\theta_K h_K\|\partial_t v_h\|_{L_2(K)}^2 + \|\nabla_x v_h\|_{L_2^\nu(K)}^2\right]\right)^{1/2}$$
$$\leq 3\|u\|_{h,*}\|v_h\|_h.$$

Thus, the bilinear form (13.11) is bounded for all choices of θ_K.

Remark 13.8 As in Remark 13.5, we can provide a simplified estimate for the special case $p = 1$ and $v|_K = v_K = const$. The first three terms can be estimated as in the above proof. The fourth term completely vanishes, since $\nabla_x(\partial_t v_h) = 0$. For the fifth term, we use the fact that $\partial_t v_h = const$ on $K \in \mathscr{T}_h$, Gauss' theorem and the Cauchy-Schwarz inequality, obtaining

$$|\theta_K h_K \int_{\partial K} v_K \nabla_x u \cdot n_x \partial_t v_h \, ds_{(x,t)}| \leq \left(\theta_K h_K v_K^2 \|\Delta_x u\|_{L_2(K)}^2\right)^{1/2} \left(\theta_K h_K \|\partial_t v_h\|_{L_2(K)}^2\right)^{1/2}.$$

Gathering the terms from the proof and the above estimate, we get

$$
\begin{aligned}
|a_h(u, v_h)| \leq &\left(\|u\|_{L_2(\Sigma_T)}^2 + \sum_{K \in \mathscr{T}_h} \left[\theta_K h_K \|\partial_t u\|_{L_2(K)}^2 + \|\nabla_x u\|_{L_2^\nu(K)}^2\right. \right. \\
&\left. \left. + (\theta_K h_K)^{-1} \|u\|_{L_2(K)}^2 + v_K^2 \theta_K h_K \|\Delta_x u\|_{L_2(K)}^2\right]\right)^{1/2} \\
&\times \left(\|v_h\|_{L_2(\Sigma_T)}^2 \right. \\
&\left. + \sum_{K \in \mathscr{T}_h} \left[3\theta_K h_K \|\partial_t v_h\|_{L_2(K)}^2 + \|\nabla_x v_h\|_{L_2^\nu(K)}^2\right]\right)^{1/2} \\
\leq &\max_{K \in \mathscr{T}_h}\{3, v_K^2\}^{1/2} \|u\|_{h,*} \|v_h\|_h.
\end{aligned}
$$

We immediately deduce that this new constant $\tilde{\mu}_b = \max_{K \in \mathscr{T}_h}\{3, v_K^2\}^{1/2}$ is also independent of h_K for all choices of positive θ_K, $K \in \mathscr{T}_h$.

Coercivity, boundedness, and consistency of the bilinear form $a_h(\cdot, \cdot)$ immediately yield a Céa-like estimate of the discretization error in the norm $\|\cdot\|_h$ by the best approximation error in the norm $\|\cdot\|_{h,*}$.

Lemma 13.8 *Let the assumptions of the coercivity Lemma 13.5 and the boundedness Lemma 13.7 hold, and let the solution u of the space-time variational problem (13.4) belong to $H_{0,\underline{0}}^{2,1}(\mathscr{T}_h)$. Then the discretization error estimate*

$$\|u - u_h\|_h \leq \left(1 + \frac{\mu_b}{\mu_c}\right) \inf_{v_h \in V_{0h}} \|u - v_h\|_{h,*} \tag{13.23}$$

hold, where $u_h \in V_{0h}$ denotes the solution of the space-time finite element scheme (13.12), and the norms $\|\cdot\|_h$ and $\|\cdot\|_{h,*}$ are defined by (13.13) and (13.21), respectively.

Proof First, from the consistency identity (13.9) and the space-time finite element scheme (13.12), we immediately deduce Galerkin orthogonality, i.e.,

$$a_h(u - u_h, v_h) = 0, \quad \forall v_h \in V_{0h}. \tag{13.24}$$

We start with the triangle inequality for the discretization error, i.e.,

$$\|u - u_h\|_h \le \|u - v_h\|_h + \|v_h - u_h\|_h.$$

Applying ellipticity proved in Lemma 13.5, the Galerkin orthogonality (13.24) and the generalized boundedness from Lemma 13.7 to the second term, we get

$$\mu_c \|v_h - u_h\|_h^2 \le a_h(v_h - u_h, v_h - u_h) = a_h(v_h - u, v_h u - u_h)$$
$$\le \mu_b \|v_h - u\|_{h,*} \|v_h - u_h\|_h.$$

Inserting this estimate in the triangle inequality above, we obtain

$$\|u - u_h\|_h \le \|u - v_h\|_h + \frac{\mu_b}{\mu_c} \|v_h - u\|_{h,*}. \tag{13.25}$$

Since $\|u - v_h\|_h \le \|v_h - u\|_{h,*}$, we immediately get the Céa-like estimate (13.23).
□

Remark 13.9 Remark 13.7 immediately implies that the Céa-like estimate (13.23) is also valid for solutions u from $H_{0,\underline{0}}^{L,1}(\mathcal{T}_h)$ provided that the norm $\|\cdot\|_{h,*}$ is now defined as in Remark 13.7.

To obtain a priori error estimates w.r.t. to the mesh dependent norm (13.13), we need approximation respectively interpolation error estimates for the finite element spaces V_{0h} w.r.t. the norm (13.21), which we summarize in the next Lemmas. Moreover, we need the *broken Sobolev space*

$$H^l(\mathcal{T}_h) := \{v \in L_2(Q) : v|_K \in H^l(K) \ \forall K \in \mathcal{T}_h\},$$

equipped with the *broken Sobolev (semi-)norm*

$$|v|_{H^l(\mathcal{T}_h)}^2 := \sum_{K \in \mathcal{T}_h} |v|_{H^l(K)}^2 \quad \text{and} \quad \|v\|_{H^l(\mathcal{T}_h)}^2 := \sum_{K \in \mathcal{T}_h} \|v\|_{H^l(K)}^2,$$

where l is some positive integer. For further details on such spaces, we refer to [11, 35].

Lemma 13.9 *Let l and k be positive integers with $l \ge k > (d+1)/2$, and let $v \in V_0 \cap H^k(Q) \cap H^l(\mathcal{T}_h)$, where $\{\mathcal{T}_h\}_{h>0}$ is a shape regular family of subdivisions of Q. Then there exists an interpolation operator Π_h, mapping from $V_0 \cap H^k(Q)$ to V_{0h}, such that*

$$\|v - \Pi_h v\|_{L_2(K)} \le C h_K^s |v|_{H^s(K)}, \tag{13.26}$$

$$\|\nabla(v - \Pi_h v)\|_{L_2(K)} \le C h_K^{s-1} |v|_{H^s(K)}, \tag{13.27}$$

$$|v - \Pi_h v|_{H^2(K)} \le C h_K^{s-2} |v|_{H^s(K)}, \tag{13.28}$$

where C is some generic constant independent of h_K and v, $s = \min\{l, p+1\}$, and p denotes the polynomial degree of the finite element shape functions on the reference element, and $V_0 = H^{1,1}_{0,\underline{0}}(Q)$.

Proof See e.g. [8, Theorem 4.4.4] or [9, Theorem 3.1.6].   □

Lemma 13.10 *Let the assumptions of Lemma 13.9 hold. Then the following interpolation error estimates are valid:*

$$\|v - \Pi_h v\|_{L_2(\Sigma_T)} \le c_1 \Big(\sum_{\substack{K \in \mathscr{T}_h \\ \partial K \cap \Sigma_T \ne \emptyset}} h_K^{2s-1} |v|^2_{H^s(K)} \Big)^{1/2}, \quad (13.29)$$

$$\|v - \Pi_h v\|_h \le c_2 \Big(\sum_{K \in \mathscr{T}_h} h_K^{2(s-1)} |v|_{H^s(K)} \Big)^{1/2}, \quad (13.30)$$

$$\|v - \Pi_h v\|_{h,*} \le c_3 \Big(\sum_{K \in \mathscr{T}_h} h_K^{2(s-1)} |v|_{H^s(K)} \Big)^{1/2}, \quad (13.31)$$

with positive constants c_1, c_2 and c_3 that do not depend on v or h_K provided that $\theta_K = \mathcal{O}(h_K)$ for all $K \in \mathscr{T}_h$.

Proof We start with the first estimate (13.29). We use the scaled trace inequality (13.22), and the interpolation error estimates (13.26) and (13.27), obtaining

$$\|v - \Pi_h v\|^2_{L_2(\Sigma_T)} = \sum_{\substack{K \in \mathscr{T}_h \\ \partial K \cap \Sigma_T \ne \emptyset}} \|v - \Pi_h v\|^2_{L_2(\partial K \cap \Sigma_T)} \le \sum_{\substack{K \in \mathscr{T}_h \\ \partial K \cap \Sigma_T \ne \emptyset}} \|v - \Pi_h v\|^2_{L_2(\partial K)}$$

$$\le \sum_{\substack{K \in \mathscr{T}_h \\ \partial K \cap \Sigma_T \ne \emptyset}} \Big[2 c_{Tr}^2 h_K^{-1} (\|v - \Pi_h v\|^2_{L_2(K)} + h_K^2 \|\nabla(v - \Pi_h v)\|^2_{L_2(K)}) \Big]$$

$$\le c_{Tr}^2 C^2 \sum_{\substack{K \in \mathscr{T}_h \\ \partial K \cap \Sigma_T \ne \emptyset}} [h_K^{2s-1} |v|_{H^s(K)}].$$

To prove (13.30), we use definition (13.13), assumption (13.14), the interpolation error estimate (13.27), and the above estimate (13.29), and obtain

$$\|v - \Pi_h v\|^2_h = \sum_{K \in \mathscr{T}_h} \Big[\theta_K h_K \|\partial_t(v - \Pi_h v)\|^2_{L_2(K)} + \|\nabla_x(v - \Pi_h v)\|^2_{L_2^y(K)} \Big]$$

$$+ \frac{1}{2} \|v - \Pi_h v\|^2_{L_2(\Sigma_T)}$$

$$\le \sum_{K \in \mathscr{T}_h} \Big[(C^2 \theta_K h_K + \overline{v}_K C^2 + c_1^2 h_K) h_K^{2(s-1)} |v|^2_{H^s(K)} \Big].$$

For the last estimate (13.31), we use definition (13.21), the above estimate (13.30), and the interpolation error estimate (13.28), obtaining

$$
\begin{aligned}
\|v - \Pi_h v\|_{h,*}^2 &= \|v - \Pi_h v\|_h^2 \\
&\quad + \sum_{K \in \mathscr{T}_h} \left[(\theta_K h_K)^{-1} \|v - \Pi_h v\|_{L_2(K)}^2 + \theta_K h_K |v - \Pi_h v|_{H^2(K)}^2 \right] \\
&\leq \sum_{K \in \mathscr{T}_h} \left(c_2^2 + h_K \theta_K^{-1} C^2 + \theta_K h_K^{-1} C^2 \right) h_K^{2(s-1)} |v|_{H^s(K)}^2.
\end{aligned}
$$

The special choice $\theta_K = \mathcal{O}(h_K)$ ensures that the constant c_3 is independent of h_K. □

Remark 13.10 The strong assumption $v \in H^k(Q)$ with $k > (d+1)/2$ is needed for the Lagrangian interpolation operator. However, in practical applications, where usually different materials occur, this requirement is too restrictive. In this case, the space-time cylinder $\overline{Q} = \bigcup_{i=1}^M \overline{Q_i}$ can be split into subdomains Q_i, which correspond to different materials. On each such subdomain Q_i, we can assume some regularity for the solution u, e.g., $u \in H^{\mathbf{l}}(\mathscr{T}(Q)) := \{v \in L_2(Q) : v|_{Q_i} \in H^{l_i}(Q_i), \text{ for all } i = 1, \ldots, M\}$ with some $\mathbf{l} = (l_1, \ldots, l_M) > 1$. For a similar case, Duan, Li, Tan and Zheng have shown a localized interpolation error estimate of the form

$$
\|\nabla(v - I_h v)\|_{L_2(Q)} \leq C \sum_{i=1}^M h_i^{s_i - 1} \|v\|_{H^{s_i}(Q_i)},
$$

in [13], where I_h is a special quasi-interpolation operator, and $s_i = \min\{l_i, p+1\}$.

Now we can formulate the following a priori estimate for the discretization error.

Theorem 13.3 *Let l and k be positive integers with $l \geq k > (d+1)/2$, $u \in V_0 \cap H^k(Q) \cap H^l(\mathscr{T}_h)$ be the exact solution, and $u_h \in V_{0h}$ be the solution of the finite element scheme* (13.12). *Furthermore, let the assumptions of the Lemmas 13.5 (coercivity), 13.7 (boundedness) and 13.10 (interpolation error estimates) be fulfilled. Then the a priori error estimate*

$$
\|u - u_h\|_h \leq c \left(\sum_{K \in \mathscr{T}_h} h_K^{2(s-1)} |u|_{H^s(K)}^2 \right)^{1/2} \tag{13.32}
$$

holds with $s = \min\{l, p+1\}$ and some generic positive constant c.

Proof Setting $v_h = \Pi_h u$ in (13.25), and using the interpolation error estimates (13.30) and (13.31), we obtain

$$\|u - u_h\|_h \leq \|u - \Pi_h u\|_h + \frac{\mu_b}{\mu_c}\|\Pi_h u - u\|_{h,*}$$

$$\leq \left(c_2 + c_3\frac{\mu_b}{\mu_c}\right)\left(\sum_{K\in\mathscr{T}_h} h_K^{2(s-1)}|u|^2_{H^s(K)}\right)^{1/2},$$

which proves estimate (13.32) with $c = c_2 + c_3(\mu_b/\mu_c)$. $\square$

Now we proceed with solving the discrete variational problem (13.12) that is nothing but one huge system of linear algebraic equations. Indeed, let $\{p^{(i)} : i = 1, \ldots, N_h\}$ be the finite element nodal basis of V_{0h}, i.e., $V_{0h} = \mathrm{span}\{p^{(1)}, \ldots, p^{(N_h)}\}$, where N_h is the number of all space-time unknowns (dofs). Then we can express the approximate solution u_h in terms of this basis, i.e., $u_h(x, t) = \sum_{i=1}^{N_h} u_i\, p^{(i)}(x, t)$. Furthermore, each basis function is a valid test function. Thus, we obtain N_h equations from (13.12). We can rewrite this system in terms of a system of linear algebraic equations

$$\mathbf{K}_h\mathbf{u}_h = \mathbf{f}_h, \qquad (13.33)$$

with $\mathbf{K}_h = (a_h(p^{(j)}, p^{(i)}))_{i,j=1,\ldots,N_h}$, $\mathbf{u}_h = (u_j)_{j=1,\ldots,N_h}$, $\mathbf{f}_h = (l_h(p^{(i)}))_{i=1,\ldots,N_h}$. The system matrix is non-symmetric, but positive definite due to Lemma 13.5. Indeed,

$$(\mathbf{K}_h\mathbf{v}_h, \mathbf{v}_h) = a_h(v_h, v_h) \geq \mu_c\|v_h\|^2_h > 0$$

for all $V_{0h} \ni v_h \leftrightarrow \mathbf{v}_h \in \mathbf{R}^{N_h} : \mathbf{v}_h \neq 0$. In dependence on the dimension N_h, the linear system (13.33) of algebraic equations can efficiently be solved by means of a sparse direct solver (e.g., sparse LU-factorization) or an iterative solver (e.g., preconditioned GMRES). In particular, it turns out that parallel versions of the GMRES preconditioned by algebraic multigrid can solve large-scale systems with several millions of unknowns on distributed memory computers with several hundreds of cores in a few seconds, also see Example 13.1 in Sect. 13.4.

13.4 Implementation and Numerical Results

We implemented our conforming space-time finite element scheme with the help of MFEM [21], a C++ library for finite elements. The resulting linear systems were then solved by means of the GMRES method, preconditioned by one V-algebraic multigrid (AMG) cycle of *BoomerAMG*. As a stopping criterion we used the reduction of the initial residual by a factor of 10^{-8}. These methods were provided

by the solver library hypre.[1] We note that both libraries are already fully parallelized with MPI. All numerical tests were performed on the RADON1[2] high performance computing cluster at Linz. The initial (spatial) meshes were created by NETGEN [38], and the space-time meshes were obtained by means of an algorithm provided by Karabelas and Neumüller [19]. For visualization we either use GLVis [20] or ParaView [2].

Example 13.1 For the first example, we consider the unit (hyper-)cube $Q = (0, 1)^{d+1}$, with $d = 2, 3$, as space-time cylinder, and choose the diffusion coefficient $v \equiv 1$. The manufactured function

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

is chosen as the exact solution, where the right-hand side is computed accordingly. This solution is highly smooth, and thus fulfills all regularity assumptions made for deriving the a prior error estimate (13.32) with optimal rates. Hence, we really can expect optimal convergence rates provided that we choose θ_K as in Remark 13.4, i.e., on each element $K \in \mathcal{T}_h$, we numerically solve a small generalized eigenvalue problem with LAPACK [3]. Indeed, Fig. 13.1b shows optimal convergence rates for all tested polynomial degrees and spatial dimensions. Moreover, we can observe from Fig. 13.1c that the preconditioned GMRES method has an optimal strong scaling behavior for systems with $N_h = 4\,601\,025$ ($p = 1, 2$) and $N_h = 5\,764\,804$ ($p = 3$) unknowns in the case $d = 3$, i.e., $Q = (0, 1)^4$. The stagnation of the scaling rate at 256 cores is due to an increased communication overhead since the problems become too small on each processor (only $\sim 15\,000$ dofs).

Example 13.2 Let us now consider an example with a moving interface in the unit hyper-cube $Q := (0, 1)^{d+1}$, with $d = 2, 3$. The moving interface is defined by the discontinuous diffusion coefficient

$$v(x, t) = \begin{cases} 1 \times 10^2, & \text{for } 2x_1 - t < \frac{1}{2}, \\ 7 \times 10^5, & \text{for } 2x_1 - t > \frac{1}{2}; \end{cases}$$

see Fig. 13.2 (left). We choose the function

$$u(x, t) = \begin{cases} \sin\left(9\pi \left(2x_1 - t - \frac{1}{2}\right)^2 (x_1 - x_1^2)\right) \sin(4\pi t) g(x), & \text{for } 2x_1 - t \leq \frac{1}{2}, \\ \sin\left(40\pi \left(2x_1 - t - \frac{1}{2}\right)^2 (x_1 - x_1^2)(t - t^2)\right) g(x), & \text{else,} \end{cases}$$
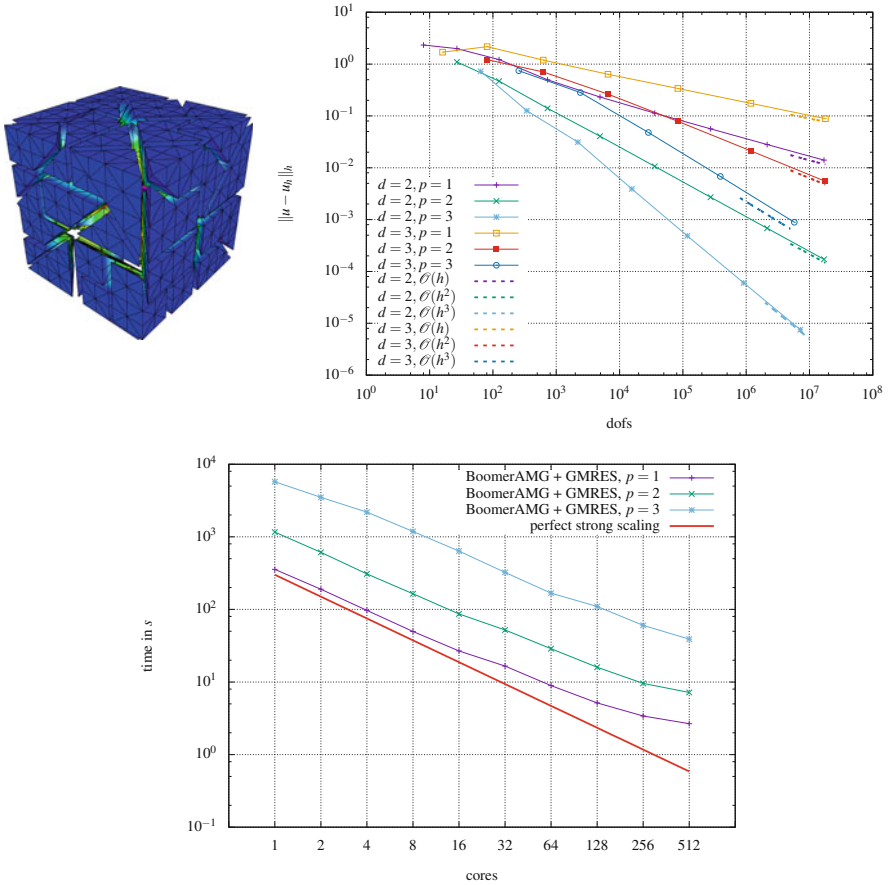

Fig. 13.1 Decomposition of the space-time cylinder into 64 subdomains for parallel computing (left); Error rates in the $\| \, . \, \|_h$-norm (right); Strong scaling of the solver for $d = 3$ and $N_h = 4\,601\,025, 4\,601\,025, 5\,764\,801$ for $p = 1, 2, 3$ (below)

with $g(x) = \prod_{i=2}^{d} \sin(\pi x_i)$, as our exact solution, and compute the corresponding right-hand side and initial data. This manufactured solution fulfills the interface conditions since this function and its first derivatives are 0 at the interface. Since this function is smooth on both sides of the moving interface, we expect optimal convergence rates; cf. Theorem 13.3. Indeed, for linear, quadratic and cubic shape functions, we observe optimal rates provided that we choose θ_K according to Remark 13.4; see Fig. 13.2 (right).

Example 13.3 For the third example, we consider the exact solution

$$u(x, t) = \left(x_1^2 - x_1 \right) \left(x_2^2 - x_2 \right) \left(t^2 - t \right) e^{-100\left((x_1 - 0.25)^2 + (x_2 - 0.25)^2 + (t - 0.25)^2 \right)}$$
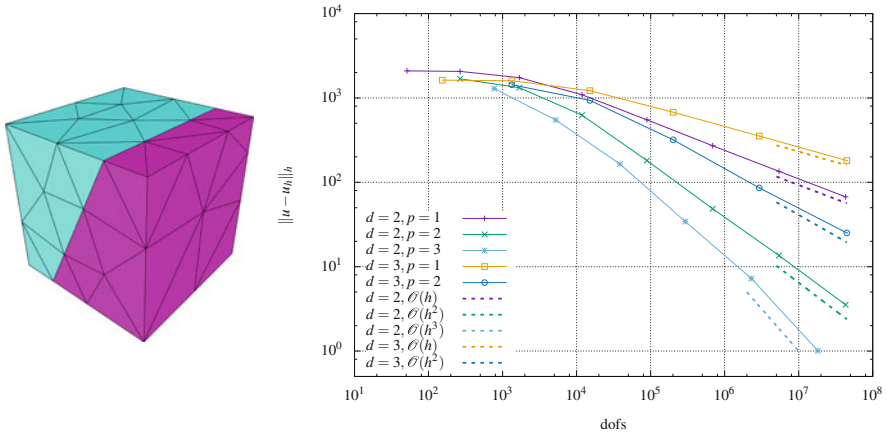
Fig. 13.2 Initial space-time mesh and diffusion coefficient $\nu(x,t)$ in color (left); Error rates in the $\|\cdot\|_h$-norm (right)
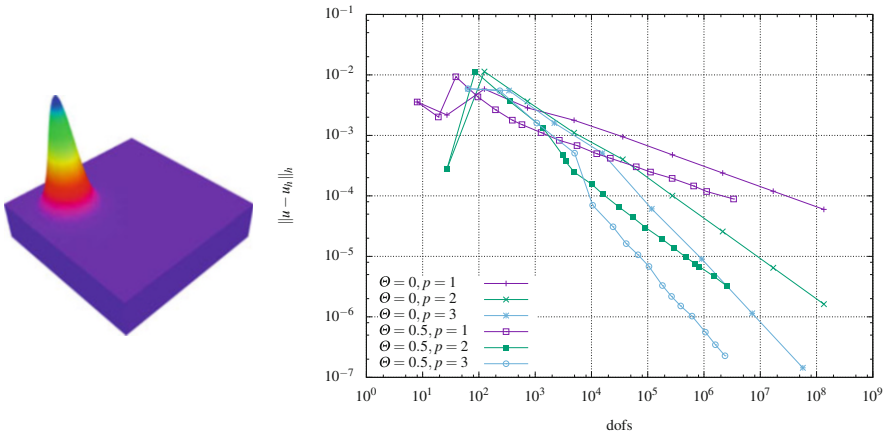

Fig. 13.3 Plot of the exact solution at $t = 0.25$ (left); Convergence rates in the $\|\cdot\|_h$-norm (right)

in the unit cube $Q = (0,1)^3$, i.e. $d = 2$, and compute the initial and boundary conditions as well as the right-hand side accordingly, where we set $\nu \equiv 1$. This function is almost zero everywhere in the space-time cylinder Q except a small area around $(0.25, 0.25, 0.25)$; see Fig. 13.3 (left). This motivates the use of an a posteriori error estimator. In particular, we use the residual-based error indicator proposed by Steinbach and Yang in [40]. For each element $K \in \mathscr{T}_h$, we compute the error indicator

$$\eta_K := \left(h_K^2 \|R_h(u_h)\|_{L_2(K)}^2 + h_K \|J_h(u_h)\|_{L_2(\partial K)}^2 \right)^{1/2},$$

where u_h is the solution of the finite element scheme (13.12), and

$$R_h(u_h) := f + \text{div}_x(\nu\nabla_x u_h) - \partial_t u_h \quad \text{in } K,$$
$$J_h(u_h) := [\nu\nabla_x u_h]_e \qquad \text{on } e \subset \partial K.$$

Here, $[\,.\,]_e$ denotes the jump across one face $e \subset \partial K$. We use a maximum marking strategy, i.e., given a parameter $\Theta \in [0, 1]$, we mark all elements whose error indicator fulfills the condition

$$\eta_K \geq \Theta \max_{K \in \mathcal{T}_h} \eta_K.$$

Unless stated otherwise, we set $\Theta = 0.5$. We note that uniform refinement is achieved by setting $\Theta = 0$. The exact solution in this example is smooth. Hence, we expect optimal convergence rates for uniform refinement after some pre-asymptotic range, which we indeed observe for all tested polynomial degrees; c.f., Fig. 13.3. For adaptive refinement, we get a better error w.r.t. to the absolute value and optimal convergence rates.

Example 13.4 For the fourth and last example, we consider the exact solution

$$u(x, t) = \sin\left(\frac{1}{\frac{1}{10\pi} + \sqrt{x_1^2 + x_2^2 + t^2}}\right),$$

in the unit cube $Q = (0, 1)^3$, i.e. $d = 2$, and compute the initial and boundary conditions as well as the right-hand side accordingly, where we set $\nu \equiv 1$. This function has a highly oscillatory behavior near the origin $(0, 0, 0)$ and is smooth everywhere else in the space-time cylinder Q; see Fig. 13.3 (left). This again motivates the use of an a posteriori error estimator. We use the same setup as in Example 13.3, i.e., the residual-based error indicator by Steinbach and Yang with a maximum marking strategy. For adaptive refinement, we recover the optimal rates for all polynomial degrees tested, whereas only reduced rates are observed for $p = 2, 3$; c.f., Fig. 13.4 (right). Moreover, we only need 47 330 dofs to obtain an energy error of the same magnitude as for 135 005 697 dofs after uniform refinement in the case $p = 1$.

13.5 Conclusions and Future Work

In this paper, following the classical books [23] and [24], we recalled that the parabolic initial boundary value problem (13.1)–(13.3) has a unique generalized (weak) solution in $H_0^{1,0}(Q)$ that even belongs to $V_{2,0}^{1,0}(Q)$. Already Ladyžhenskaya
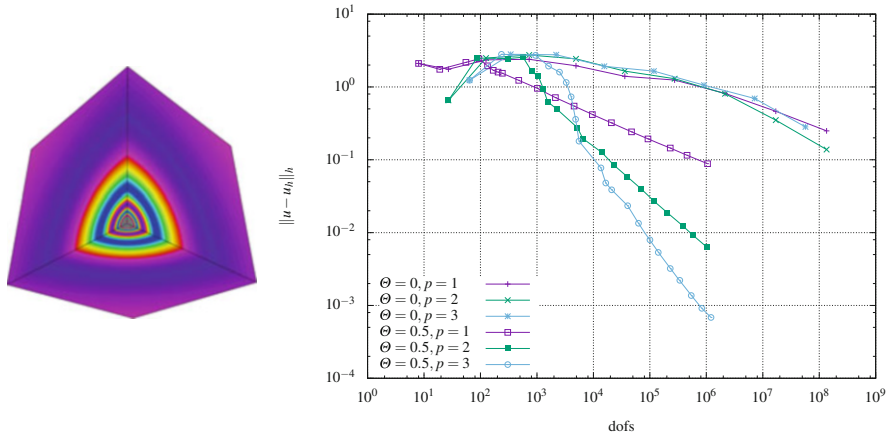
Fig. 13.4 Plot of the exact solution (left); Convergence rates in the $\| \cdot \|_h$-norm (right)

proved that, in the case $\nu = 1$, the solution u even belongs to $H^{\Delta,1}(Q)$ provided that the right-hand side $f \in L_2(Q)$ and initial conditions $u_0 \in H_0^1(\Omega)$; see [23]. This setting of the so-called maximal parabolic regularity was also considered in this paper. We again mention that we only need this property element-wise to construct a consistent and stable space-time finite element scheme. We proceeded with deriving a stable space-time finite element scheme, for which we showed coercivity (ellipticity) and boundedness on the finite element spaces respectively extended finite element spaces. These properties together with consistency and standard interpolation or quasi-interpolation error estimates led to a priori discretization error estimates in the corresponding mesh-dependent norm with optimal rates. We performed several numerical experiments with four test problems possessing different features. The first example has a smooth solution that led to optimal convergence rate as predicted by the theory. Moreover, due to the ellipticity of the bilinear form $a_h(\cdot, \cdot)$, the AMG precondition GMRES is a very efficient parallel solver. The second example has a moving interface that is given by a discontinuous diffusion coefficient $\nu(x, t)$ depending on both x and t. In the third and fourth example, we studied adaptivity based on the a posteriori residual error indicator proposed in [40]. It is clear that the interplay of adaptivity and fast parallel iterative solvers will lead to the most efficient completely unstructured adaptive space-time solvers for complicated initial-boundary value problems for linear and even non-linear parabolic partial differential equations. Adaptive Space-Time Finite Element Methods and Solvers can be useful for solving eddy current problems with moving and non-moving parts like in electrical machines. In many practical applications, one is interested in optimal control or in optimal design of electrical machines; see, e.g., [17]. Adaptive Space-Time Finite Element Methods are especially suited for solving the optimality system that is nothing but a coupled PDE system living in the space-time cylinder Q; see, e.g., [44].

References

1. Ahmed, N., Matthies, G.: Numerical study of SUPG and LPS methods combined with higher order variational time discretization schemes applied to time-dependent linear convection–diffusion–reaction equations. J. Sci. Comput. **67**(3), 988–1018 (2016)

2. Ahrens, J., Geveci, B., Law, C.: ParaView: An End-User Tool for Large Data Visualization. Elsevier, San Diego (2005)

3. Anderson, E., Bai, Z., Bischof, C., Blackford, S., Demmel, J., Dongarra, J., Du Croz, J., Greenbaum, A., Hammarling, S., McKenney, A., Sorensen, D.: LAPACK Users' Guide, 3rd edn. SIAM, Philadelphia (1999)

4. Aziz, A.K., Monk, P.: Continuous finite elements in space and time for the heat equation. Math. Comput. **52**, 255–274 (1989)

5. Bank, R.E., Vassilevski, P., Zikatanov, L.: Arbitrary dimension convection-diffusion scheme for space-time discretizations. J. Comput. Appl. Math. **310**, 19–31 (2017)

6. Bause, M., Radu, F.A., Köcher, U.: Error analysis for discretizations of parabolic problems using continuous finite elements in time and mixed finite elements in space. Numer. Math. **137**(4), 773–818 (2017)

7. Braess, D.: Finite Elemente; Theorie, schnelle Löser und Anwendungen in der Elastizitätstheorie. Springer-Verlag, Berlin (2007)

8. Brenner, S.C., Scott, L.R.: The Mathematical Theory of Finite Element Methods. Texts in Applied Mathematics, vol. 15, 3rd edn. Springer, New York (2008)

9. Ciarlet, P.G.: The Finite Element Method for Elliptic Problems. North-Holland Publishing, Amsterdam (1978)

10. Devaud, D., Schwab, C.: Space–time hp-approximation of parabolic equations. Calcolo **55**(3), 35 (2018)

11. Di Pietro, D.A., Ern, A.: Mathematical Aspects of Discontinuous Galerkin Methods. Mathématiques & Applications (Berlin) [Mathematics & Applications], vol. 69. Springer, Heidelberg (2012)

12. Dörfler, W., Findeisen, S., Wieners, C.: Space-time discontinuous Galerkin discretizations for linear first-order hyperbolic evolution. Comput. Methods Appl. Math. **16**, 409–428 (2016)

13. Duan, H., Li, S., Tan, R.C.E., Zheng, W.: A delta-regularization finite element method for a double curl problem with divergence-free constraint. SIAM J. Numer. Anal. **50**(6), 3208–3230 (2012)

14. Ern, A., Schieweck, F.: Discontinuous Galerkin method in time combined with a stabilized finite element method in space for linear first-order PDEs. Math. Comput. **85**(301), 2099–2129 (2016)

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

16. Gander, M.J.: 50 years of time parallel integration. In: Multiple Shooting and Time Domain Decomposition, pp. 69–114. Springer Verlag, Heidelberg (2015)

17. Gangl, P., Langer, U., Laurain, A., Meftahi, H., Sturm, K.: Shape optimization of an electric motor subject to nonlinear magnetostatics. SIAM J. Sci. Comput. **37**(6), B1002–B1025 (2015)

18. Hackbusch, W.: Parabolic multigrid methods. In: Glowinski, R., Lions, J.L. (eds.) Computing Methods in Applied Sciences and Engineering VI, pp. 189–197. North-Holland, Amsterdam (1984)

19. Karabelas, E., Neumüller, M.: Generating admissible space-time meshes for moving domains in d+1-dimensions. Technical Report 07, Institute of Computational Mathematics, Johannes Kepler University Linz (2015)
20. Kolev, T., Dobrev, V.: GLVis: OpenGL finite element visualization tool. https://github.com/glvis/glvis, 6 (2010)
21. Kolev, T., Dobrev, V.: MFEM: modular finite element methods library. https://github.com/mfem/mfem, 6 (2010)
22. Kunstmann, P.C., Weis, L.: Maximal L_p-regularity for Parabolic Equations, Fourier Multiplier Theorems and H^∞-functional Calculus, pp. 65–311. Springer-Verlag, Berlin (2004)
23. Ladyžhenskaya, O.A.: The Boundary Value Problems of Mathematical Physics. Applied Mathematical Sciences, vol. 49. Springer-Verlag, New York (1985)
24. Ladyzhenskaya, O.A., Solonnikov, V.A., Uraltseva, N.N.: Linear and Quasilinear Equations of Parabolic Type. AMS, Providence (1968)
25. Lang, J.: Adaptive Multilevel Solution of Nonlinear Parabolic PDE Systems: Theory, Algorithm, and Applications. LNCSE, vol. 16. Springer-Verlag, Berlin (2001)
26. Langer, U., Moore, S.E., Neumüller, M.: Space-time isogeometric analysis of parabolic evolution problems. Comput. Methods Appl. Mech. Eng. **306**, 342–363 (2016)
27. Leykekhman, D., Vexler, B.: Discrete maximal parabolic regularity for Galerkin finite element methods. Numer. Math. **135**, 923–952 (2017)
28. Lions, J.-L., Maday, Y., Turinici, G.: A "parareal" in time discretization of PDE's. C. R. Acad. Sci. Paris I **332**, 661–668 (2001)
29. Moore, S.: Nonstandard discretization strategies in isogeometric analysis for partial differential equations. PhD thesis, Johannes Kepler University Linz, Linz (2017)
30. Moore, S.: A stable space–time finite element method for parabolic evolution problems. Calcolo **55**, 18 (2018). First Online: 16 April 2018.
31. Neumüller, M.: Space-time Methods; Fast Solvers and Applications. Monographic Series TU Graz: Computation in Engineering and Science, vol. 20. TU Graz, Graz (2013)
32. Neumüller, M., Steinbach, O.: Refinement of flexible space-time finite element meshes and discontinuous Galerkin methods. Comput. Vis. Sci. **14**(5), 189–205 (2011)
33. Otoguro, Y., Takizawa, K., Tezduyar, T.E.: Space-time VMS computational flow analysis with isogeometric discretization and a general-purpose NURBS mesh generation method. Comput. Fluids **158**, 189–200 (2017)
34. Rhebergen, S., Cockburn, B., van der Vegt, J.J.W.: A space-time discontinuous Galerkin method for the incompressible Navier-Stokes equations. J. Comput. Phys. **233**, 339–358 (2013)
35. Rivière, B.: Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations. Theory and Implementation. Frontiers in Applied Mathematics, vol. 35. Society for Industrial and Applied Mathematics, Philadelphia (2008)
36. Schafelner, A.: Space-time finite element methods for parabolic inital-boundary problems. Master's thesis, Johannes Kepler University Linz (2017)
37. Schieweck, F.: A-stable discontinuous Galerkin-Petrov time discretization of higher order. J. Numer. Math. **18**(1), 25–57 (2010)
38. Schöberl, J.: NETGEN: an advancing front 2D/3D-mesh generator based on abstract rules. Comput. Vis. Sci. **1**(1), 41–52 (1997)
39. Steinbach, O.: Space-time finite element methods for parabolic problems. Comput. Methods Appl. Math. **15**(4), 551–566 (2015)
40. Steinbach, O., Yang, H.: Comparison of algebraic multigrid methods for an adaptive space-time finite-element discretization of the heat equation in 3d and 4d. Numer. Linear Algebra Appl. **25**(3) (2018). First published online as e2143, 7 February 2018
41. Steinbach, O., Yang, H.: Space–time finite element methods for parabolic evolution equations: discretization, a posteriori error estimation, adaptivity and solution. In: U. Langer, O. Steinbach (eds.) Space-Time Methods: Application to Partial Differential Equations. Radon Series on Computational and Applied Mathematics, vol. 25. pp. 226–259, de Gruyter (2019)
42. Thomée, V.: Galerkin Finite Element Methods for Parabolic Problems. Springer Series in Computational Mathematics, vol. 25, 2nd edn. Springer-Verlag, Berlin (2006)

43. Touloupoulos, I.: Stabilized space-time finite element methods of parabolic evolution problems. Technical Report 2017–19, RICAM, Linz (2017)
44. Tröltzsch, F.:   Optimal Control of Partial Differential Equations: Theory, Methods and Applications. AMS, Providence (2010)
45. Vandewalle, S.: Parallel Multigrid Waveform Relaxation for Parabolic Problems. Teubner Skripten zur Numerik. Teubner, Stuttgart (1993)
46. Wloka, J.: Partial Differential Equations. Cambridge University Press, Cambridge (1987)