Space-time hp-approximation of parabolic equations

D. Devaud and Ch. Schwab

Research Report No. 2017-50
October 2017

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

Noname manuscript No.
(will be inserted by the editor)

Space-time hp-approximation of parabolic equations

Denis Devaud · Christoph Schwab

the date of receipt and acceptance should be inserted later

Abstract A new space-time finite elements methods (FEM) approximation for the solution of parabolic partial differential equations (PDE) is introduced. Considering a mesh-dependent norm, it is first shown that the discrete bilinear form is coercive and continuous, yielding existence and uniqueness of the associated solution. In a second step, error estimates in this norm are derived. In particular, we show that combining low-order elements for the space variable together with an hp-approximation of the problem with respect to the temporal variable allows us to decrease the optimal convergence rates for the approximation of elliptic problems only by a logarithmic factor. For simultaneous space-time hp-discretization in both, the spatial as well as the temporal variable, overall exponential convergence in mesh-degree dependent norms on the space-time cylinder is proved, under analytic regularity assumptions on the solution with respect to the spatial variable. Numerical results for linear model problems confirming exponential convergence are presented.

Mathematics Subject Classification (2000) 65M12 · 65M15 · 65M60

1 Introduction

Let $D \subset \mathbb{R}^d$ $(d = 1, 2, 3)$ be a bounded convex polyhedral domain, $0 < T < \infty$ and $f \in L^2(Q)$, where $Q := D \times I := D \times (0, T)$. Furthermore, let $\mathcal{A} \in \left(W^{2,\infty}(D)\right)^{2\times 2}$

Denis Devaud
Institut de Statistique, Université de Neuchâtel, Avenue de Bellevaux 51, 2000 Neuchâtel
E-mail: denis.devaud@unine.ch

Christoph Schwab
Seminar for Applied Mathematics, ETH Zürich, Rämistrasse 101, 8092 Zürich
E-mail: christoph.schwab@sam.math.ethz.ch

be symmetric and such that there exist bounds $0 < \lambda_- \leq \lambda_+ < \infty$ satisfying

$$\sum_{i,j=1}^{d} \mathcal{A}_{i,j}(x)\xi_i\xi_j \geq \lambda_- |\xi|^2, \qquad \forall \xi \in \mathbb{R}^d,\ x \in D, \tag{1a}$$

$$\sup_{x\in D} \left|\mathcal{A}_{i,j}(x)\right| \leq \lambda_+, \qquad \forall i,j \in \{1,\ldots,d\}, \tag{1b}$$

i.e. $\mathcal{A}$ is continuous and uniformly positive definite.

We are then interested in solving the following problem: find $u : Q \to \mathbb{R}$ such that

$$\partial_t u(x,t) - \operatorname{div}(\mathcal{A}(x)\nabla_x u(x,t)) = f(x,t), \qquad (x,t) \in Q, \tag{2a}$$

$$u(x,t) = 0, \qquad (x,t) \in \Gamma \cup \Gamma_0, \tag{2b}$$

where $\Gamma := \partial D \times I$ and $\Gamma_0 := D \times \{0\}$. The system (2a)-(2b) is used to model for instance heat diffusion in an material. We point out that the matrix $\mathcal{A}$ allows for heterogeneous material properties.

Standard approximation techniques to solve these equations are based on time stepping or time marching schemes, such as the method of lines where the evolution equation is first discretized in space, yielding a system of ordinary differential equations (ODE). The resulting system is then approximated using an ODE solver such as Runge-Kutta methods or backward differentiation formulas [6,21]. Time marching schemes are "oblivious" in the sense that they only track the current state and, possibly (in multistep methods), the state of a few previous time steps. In many applications, however, the solution of the evolution problem over the entire time horizon is of interest. We mention only optimal control problems, data assimilation and uncertainty quantification. Space-time methods, on the contrary, interpret the evolution problem as an operator equation on suitable Bochner spaces. We refer to [18,19,10] and the references there for such formulations, for linear and certain nonlinear parabolic evolution and control problems. The numerical solution of such operator equation formulations results in iterative schemes which approximate the solution of the evolution problem on the entire space-time cylinder. At first sight, this produces massively larger volumes of data. However, parabolic regularity and multiresolution space-times data compression can result in competitive schemes. In recent years, several algorithms based on these ideas have emerged. We mention the parareal method, which aims at improved computational efficiency by time-parallel integration. Here, no data compression of the space-time solution is foreseen. Space-time adaptive, compressive discretization based on suitable multiresolution representation of approximate solutions achieves compression at runtime, and produces optimally compressed, approximate solutions, with optimality with respect to best n-term approximation rates. We refer the reader to [7] for a survey on such algorithms. More related to our method, *hp*-FEM time-discretizations based on discontinuous Galerkin (dG) approximation were proposed in [15,16,22]. There, it was proved that parabolic (analytic) regularity allows *exponential convergence* of *hp*-time (semi) discretizations. Several algorithms based on a continuous approximation (cG) in time were also analyzed, based on saddle-point formulations of the evolution equations with respect to the time variable. We mention [1,2,14,18,13]. In particular, in [18] the authors prove that it is possible to adaptively compute approximate wavelet solutions of parabolic equations with a rate depending only on the spatial discretization. Our goal in the present paper is to derive a similar result based on non-adaptive FEM solutions. To this end, we develop a coercive space-time variational formulation. Due to the saddle point nature of the continuous problem which is due to the non self-adjointness of the parabolic operator, in [12], Langer, Moore and Neumüller present a method based on a space-time isogeometric analysis in mesh-dependent norms for the approximation of (2a)-(2b). In the present paper, we present an hp-error analysis of coercive space-time formulations which are similar to the one used in this reference, and in [20]. In particular, it is based on a coercive space-time variational formulation in a mesh-degree dependent family of norms.

Assuming that the right-hand side f of (2a) is smooth in time, we obtain that the associated solution exhibits the same regularity except for a potential algebraic singularity at the initial time due an incompatibility between f and the initial condition. As discussed in [5], this lack of regularity can be overcome by considering a continuous hp-approximation for the time variable. Considering such technique in time together with a standard low order h-FEM for the spatial component, we show that solutions of the heat equation can be approximated with the same rate as elliptic equations up to a logarithmic factor with respect to the total number of degrees of freedom. This is a consequence of the exponential convergence of the time hp-approximation. Furthermore, we also discuss higher order approximation for the spatial variable in a two dimensional polygonal domain. In particular, considering an hp-approximation for both the time and space discretization, we show that it is possible to obtain exponential convergence with respect to the total number of degrees of freedom.

The paper is organized as follows. In the following section, we present a weak formulation associated to the heat equation (2a)-(2b). Well-posedness of the variational problem and a regularity result are discussed. In Section 3, our space-time discretization is introduced and we provide convergence rates for the discrete solution. Finally, a numerical example is presented in Section 4.

Throughout this manuscript, we use the notation $A \approx B$ if there exist two constants $c_1, c_2 > 0$ independent of A and B such that $c_1 A \leq B \leq c_2 A$. In integrals over space-time domains, we omit the differential $dxdt$ to lighten notation. Furthermore, for two vector spaces A and B, $A \otimes B$ refers to either the algebraic tensor product space in the case of finite dimensional spaces or to the Hilbertian tensor product space if A and B are separable Hilbert spaces. Given normed vector spaces A and B, $\mathcal{L}(A, B)$ denotes the space of linear continuous functionals from A to B. In the case $A = B$, we simply write $\mathcal{L}(A)$ and $A^* := \mathcal{L}(A, \mathbb{R})$.

2 Space-time weak formulation

2.1 Problem formulation

For $l, k \in \mathbb{N}_0$, let

$$
\begin{aligned}
H^{l,k}(Q) &:= \left\{ v \in L^2(Q) \;\middle|\; \partial_x^\alpha v \in L^2(Q),\ 0 \leq |\alpha| \leq l,\ \partial_t^i v \in L^2(Q),\ 0 \leq i \leq k \right\}, \\
H_{\otimes}^{l,k}(Q) &:= \left\{ v \in L^2(Q) \;\middle|\; \partial_x^\alpha \partial_t^i v \in L^2(Q),\ 0 \leq |\alpha| \leq l,\ 0 \leq i \leq k \right\}.
\end{aligned}
$$

Here $L^2(Q)$ is the usual space of square integrable functions over Q and for a multi-index $\alpha = (\alpha_1, \dots, \alpha_d)$, $|\alpha| := \alpha_1 + \dots + \alpha_d$ and $\partial_x^\alpha := \frac{\partial^{|\alpha|}}{\partial_{x_1}^{\alpha_1} \dots \partial_{x_d}^{\alpha_d}}$. We equip $H^{l,k}(Q)$ and $H_\otimes^{l,k}(Q)$ with the norms

$$\|v\|^2_{H^{l,k}(Q)} := \sum_{0\le j\le l} |v|^2_{j,0,Q} + \sum_{0\le i\le k} |v|^2_{0,i,Q}\,, \qquad \|v\|^2_{H_\otimes^{l,k}(Q)} := \sum_{\substack{0\le j\le l\\ 0\le i\le k}} |v|^2_{j,i,Q}\,,$$

where $|v|^2_{j,i,Q} := \sum_{|\alpha|=j} \left\|\partial_x^\alpha \partial_t^i v\right\|^2_{L^2(Q)}$. Furthermore, for $\Gamma_T := D \times \{T\}$, let

$$\begin{aligned}
H_0^{1,0}(Q) &:= \left\{ v \in H^{1,0}(Q) \middle| \, v|_\Gamma = 0 \right\},\\
H_{0,\overline{0}}^{1,1}(Q) &:= \left\{ v \in H_0^{1,0}(Q) \cap H^{0,1}(Q) \middle| \, v|_{\Gamma\cup\Gamma_T} = 0 \right\},\\
H_{0,\underline{0}}^{1,1}(Q) &:= \left\{ v \in H_0^{1,0}(Q) \cap H^{0,1}(Q) \middle| \, v|_{\Gamma\cup\Gamma_0} = 0 \right\}.
\end{aligned}$$

We point out that $H^{l,k}(Q) = H^l(D)\otimes L^2(I) \cap L^2(D)\otimes H^k(I)$ and $H_\otimes^{l,k}(Q) = H^l(D)\otimes H^k(I)$. Moreover, the embeddings $H_\otimes^{l,l}(Q) \subset H^l(Q) \subset H^{l,l}(Q)$ are continuous for all $l \in \mathbb{N}_0$, where $H^l(Q)$ denotes the classical Sobolev spaces over Q. The first inclusion is strict if $l \ge 1$ while the second one is strict in the case $l \ge 2$. The spaces $H^{l,k}(Q)$ and $H_\otimes^{l,k}(Q)$ are usually called mixed Sobolev spaces in the literature. The weak formulation of (2a)-(2b) is then: find $u \in H_0^{1,0}(Q) \cap L^2(D; \mathcal{C}(\overline{I}))$ such that

$$a(u,v) = l(v), \qquad \forall v \in H_{0,\overline{0}}^{1,1}(Q), \tag{3}$$

where

$$a(u,v) := -\int_Q u\partial_t v + \int_Q \mathcal{A}\nabla_x u \cdot \nabla_x v, \qquad l(v) := \int_Q fv.$$

It was shown in [11, Theorem 2.3 p.115] that (3) is well posed, i.e. it admits a unique solution which depends continuously on the data f. Furthermore under the assumption that $u \in H_{0,\underline{0}}^{1,1}(Q) \cap H^{2,0}(Q)$, we obtain using integration by parts in (3) that for all $v \in \mathcal{C}_0^\infty(Q) \subset H_{0,\overline{0}}^{1,1}(Q)$

$$\int_Q fv = l(v) = \mathfrak{a}(u,v) = -\int_Q u\partial_t v + \int_Q \mathcal{A}\nabla_x u \cdot \nabla_x v = \int_Q \left(\partial_t u - \operatorname{div}\left(\mathcal{A}\nabla_x u\right)\right) v.$$

Since $\mathcal{C}_0^\infty(Q)$ is dense in $L^2(Q)$, we obtain that (2a) holds in $L^2(Q)$ under theses assumptions.

2.2 A regularity result

In this section, we derive a regularity result with respect to the time variable. To treat the time regularity, we consider the following classes of functions. We point out that the usual regularity theory for solutions of parabolic equations deals with analytic right-hand side functions. This is a special case of the following result for $\delta = 1$. The motivation for developing a theory for more general forcing terms comes from other PDEs. Indeed, in [4], it has been shown that solutions of Euler equations for incompressible fluids fulfill such type of regularity for some $\delta > 1$ even in the case of analytic right-hand side. Furthermore, for $\delta > 1$ it is possible to construct functions of Gevrey type with compact support, allowing then to localize any global forcing function while keeping smoothness properties.

Definition 1 (Gevrey class) Let $I = (a, b)$ be a bounded interval. Then for $k \in \mathbb{N}_0$, $\delta \geq 1$ and $\theta \geq 0$, we say that f is of *Gevrey type* (k, δ, θ) and we write $f \in H^k\left(D; \mathscr{G}^{\delta,\theta}(I)\right)$ if $f \in H^k\left(D; \mathcal{C}(\overline{I})\right)$ and if there exist $C_f, d_f > 0$ such that

$$\left(\sum_{0\leq|\alpha|\leq k} \left\|\partial_x^\alpha \partial_t^l f(t)\right\|^2_{L^2(D)}\right)^{1/2} \leq C_f d_f^l \Gamma(l+1)^\delta (t-a)^{\theta-l}, \qquad l \in \mathbb{N},\ t \in I, \quad (4)$$

where $\Gamma(s)$ denotes the Gamma function for $s > 0$. We also define $H^1_0\left(D; \mathscr{G}^{\delta,\theta}(I)\right)$ as the space of functions $f \in H^1\left(D; \mathscr{G}^{\delta,\theta}(I)\right)$ such that $f|_{\partial D} = 0$ in $\mathcal{C}(\overline{I})$. Moreover we say that f is of *Gevrey type* (k, δ) and write $f \in H^k\left(D; \mathscr{G}^{\delta}(I)\right)$ if there exist $C_f, d_f > 0$ such that

$$\left(\sum_{0\leq|\alpha|\leq k} \left\|\partial_x^\alpha \partial_t^l f(t)\right\|^2_{L^2(D)}\right)^{1/2} \leq C_f d_f^l \Gamma(l+1)^\delta, \qquad l \in \mathbb{N}_0,\ t \in \overline{I}. \qquad (5)$$

Proposition 1 *Let* $f \in L^2\left(D; \mathscr{G}^{\delta}(I)\right)$ *for some* $\delta \geq 1$ *and* $\mathcal{A} \in \left(W^{2,\infty}(D)\right)^{2\times 2}$ *satisfy* (1a)-(1b). *Then*

$$u \in L^2\left(D; \mathscr{G}^{\delta,1}(I)\right) \cap H^1_0\left(D; \mathscr{G}^{\delta,1/2}(I)\right) \cap H^{2,1}_{\otimes}(Q),$$

where u *denotes the solution of* (2a)-(2b) *associated to* f.

Proof The proof of $u \in L^2\left(D; \mathscr{G}^{\delta,1}(I)\right) \cap H^1_0\left(D; \mathscr{G}^{\delta,1/2}(I)\right)$ can be found in [5, Proposition 3.6.7]. To obtain that $u \in H^{2,1}_{\otimes}(Q)$, we have used that $u \in H^1(I; D(A))$, where $D(A)$ denotes the domain of the operator $Av := -\operatorname{div}(\mathcal{A}\nabla_x v)$ for all $v \in H^1_0(D)$. Indeed this follows the same lines as the proof of Proposition 3.6.7 in [5]. Then, from [9, Theorem 3.2.1.2], we have that $D(A) = H^2(D) \cap H^1_0(D)$ and the result follows since $H^1(I; H^2(D))$ is isomorphic to $H^{2,1}_{\otimes}(Q)$.

3 Space-time discretization

3.1 The discrete problem

In the following, we present a full space-time discretization of (3). It is based on the idea that the continuous space has a specific tensor structure. We show that the bilinear form is coercive (Lemma 1) and continuous (Lemma 2) with respect to a mesh-dependent norm. Using these results, we obtain the best approximation property (10) in this norm. The ideas used in the following are based on the development presented in [12], where a space-time isogeometric analysis approximation is considered.

For $M_t \in \mathbb{N}$, let $\mathcal{T}^t_{M_t} := \{I_n\}_{n=1}^{M_t}$ and $Q_n := D \times I_n$ for $n \in \{1, \dots, M_t\}$, where

$$0 =: t_0 < t_1 < \cdots < t_{M_t} := T, \qquad I_n := (t_{n-1}, t_n), \qquad n = 1, \dots, M_t.$$

We assume that $h_t^n := t_n - t_{n-1} \leq 2$ and that there exists $C_t \geq 1$ such that

$$h_t^n \leq h_t^{n+1} \leq C_t h_t^n, \qquad n = 1, \dots, M_t - 1. \tag{6}$$

This assumption is fulfilled for uniform as well as for geometric meshes. Given $p_t, p_x \in \mathbb{N}$ and a mesh $\mathcal{T}^x$ of D, let

$$\begin{aligned}
\mathcal{S}_0^{p_x}(D, \mathcal{T}^x) &:= \left\{ v \in H_0^1(D) \middle| \; v|_K \in \mathcal{P}^{p_x}(K), \; \forall K \in \mathcal{T}^x \right\}, \\
\mathcal{S}^{p_t}(I, \mathcal{T}^t_{M_t}) &:= \left\{ v \in H^1(I) \middle| \; v|_{I_n} \in \mathcal{P}^{p_t}(I_n), \; n = 1, \dots, M_t \right\},
\end{aligned}$$

where $\mathcal{P}^{p_x}(K)$ denotes the set of polynomials of degree at most p_x over $K \in \mathcal{T}^x$ and $\mathcal{P}^{p_t}(I_n)$ the set of polynomials of degree at most p_t over I_n, $n = 1, \dots, M_t$. Moreover, $\mathcal{S}_{\underline{0}}^{p_t}(I, \mathcal{T}^t_{M_t}) := \left\{ v \in \mathcal{S}^{p_t}(I, \mathcal{T}^t_{M_t}) \middle| \; v(0) = 0 \right\}$. We point out that the polynomial degrees p_t and p_x are fixed over the whole partition. In the following, we will consider $p_x = 1$. The case $p_x > 1$ as well as non-constant polynomial degrees over the mesh $\mathcal{T}^x$ are discussed in Section 3.3. We then define the mesh function $\mathfrak{h}_t \in \mathcal{S}^1(I, \mathcal{T}^t_{M_t})$ through

$$\mathfrak{h}_t(t_n) = h_t^n + h_t^{n+1}, \qquad n = 0, \dots, M_t,$$

where $h_t^0 = h_t^{M_t+1} = 0$. For $t \in I_n$, we then have

$$\begin{aligned}
\mathfrak{h}_t(t) &= \frac{1}{h_t^n} \left(\mathfrak{h}_t(t_{n-1})(t_n - t) + \mathfrak{h}_t(t_n)(t - t_{n-1}) \right) \\
&= \frac{1}{h_t^n} \left((h_t^n)^2 + h_t^{n-1} t_n - h_t^{n+1} t_{n-1} + t(h_t^{n+1} - h_t^{n-1}) \right),
\end{aligned} \tag{7}$$

so that

$$h_t^n \leq \mathfrak{h}_t(t) \leq 2 C_t h_t^n, \qquad \forall t \in I_n, \; n = 1, \dots, M_t. \tag{8}$$

On the space $\mathcal{V}_{hp} := \mathcal{S}_0^{p_x}(D, \mathcal{T}^x) \otimes \mathcal{S}_{\underline{0}}^{p_t}(I, \mathcal{T}^t_{M_t})$, we define for $0 < \theta < \frac{2\lambda_-}{C_t \lambda_+}$ and $v_{hp} \in \mathcal{V}_{hp}$ the norm

$$\begin{aligned}
\left\| v_{hp} \right\|_{hp}^2 := &\left(\lambda_- - \frac{C_t \theta \lambda_+}{2 p_t^2} \right) \left\| \nabla_x v_{hp} \right\|_{L^2(Q)}^2 + \frac{\theta}{p_t^2} \left\| (\mathfrak{h}_t)^{1/2} \partial_t v_{hp} \right\|_{L^2(Q)}^2 \\
&+ \frac{1}{2} \int_D \left| v_{hp}(x, T) \right|^2 dx.
\end{aligned}$$

If $\|v\|_{hp} = 0$ for some $v \in H^{1,1}_{0,\underline{0}}(Q)$, it means that v is constant over Q. Due to the initial and boundary conditions, it holds $v = 0$ in Q. Since the other properties can be obtained easily, we have that $\|\cdot\|_{hp}$ defines a norm on $H^{1,1}_{0,\underline{0}}(Q)$ and in particular it is a norm on $\mathcal{V}_{hp}$. Furthermore, it can be shown that it is equivalent to the $\|\cdot\|_{H^{1,1}(Q)}$ norm on $H^{1,1}_{0,\underline{0}}(Q)$. We then introduce the bilinear and linear forms $\mathfrak{a}_{hp} : \mathcal{V}_{hp} \times \mathcal{V}_{hp} \to \mathbb{R}$ and $l_{hp} : \mathcal{V}_{hp} \to \mathbb{R}$ defined as

$$\begin{aligned} \mathfrak{a}_{hp}(v_{hp}, w_{hp}) &:= \int_Q \left[\partial_t v_{hp} w_{hp} + \left(\mathcal{A}\nabla_x v_{hp}\right) \cdot \nabla_x w_{hp}\right] \\ &\quad + \frac{\theta}{p_t^2} \int_Q \mathfrak{h}_t \left[\partial_t v_{hp} \partial_t w_{hp} + \left(\mathcal{A}\nabla_x v_{hp}\right) \cdot \nabla_x \partial_t w_{hp}\right], && \forall v_{hp}, w_{hp} \in \mathcal{V}_{hp}, \\ l_{hp}(w_{hp}) &:= \int_Q f \left[w_{hp} + \frac{\theta}{p_t^2} \mathfrak{h}_t \partial_t w_{hp}\right], && \forall w_{hp} \in \mathcal{V}_{hp}. \end{aligned}$$

On the space $\mathcal{V}_{hp,*} := H^{1,1}_{0,\underline{0}}(Q) + \mathcal{V}_{hp}$, we define the norm

$$\|v\|^2_{hp,*} := \|v\|^2_{hp} + \frac{p_t^2}{\theta} \left\| (\mathfrak{h}_t)^{-1/2} v \right\|^2_{L^2(Q)}, \qquad \forall v \in \mathcal{V}_{hp,*}.$$

Using the same argument as above, we obtain that $\|\cdot\|_{hp,*}$ indeed defines a norm on $\mathcal{V}_{hp,*}$. We then have the following important lemmas.

Lemma 1 *It holds*

$$\mathfrak{a}_{hp}(v_{hp}, v_{hp}) \geq \left\|v_{hp}\right\|^2_{hp}, \qquad \forall v_{hp} \in \mathcal{V}_{hp}.$$

Proof For $v_{hp} \in \mathcal{V}_{hp}$, it holds using integration by parts in time and the boundary conditions

$$\begin{aligned} \mathfrak{a}_{hp}(v_{hp}, v_{hp}) &= \frac{1}{2} \int_Q \partial_t \left|v_{hp}\right|^2 + \int_Q \left(\mathcal{A}\nabla_x v_{hp}\right) \cdot \nabla_x v_{hp} \\ &\quad + \frac{\theta}{p_t^2} \left\| (\mathfrak{h}_t)^{1/2} \partial_t v_{hp} \right\|^2_{L^2(Q)} + \frac{\theta}{2p_t^2} \int_Q \mathfrak{h}_t \partial_t \left(\mathcal{A}\nabla_x v_{hp}\right) \cdot \nabla_x v_{hp} \\ &= \frac{1}{2} \int_D \left|v_{hp}(x,T)\right|^2 dx + \int_Q \left(\mathcal{A}\nabla_x v_{hp}\right) \cdot \nabla_x v_{hp} \\ &\quad + \frac{\theta}{2p_t^2} \int_D \mathfrak{h}_t(T) \left(\mathcal{A}(x)\nabla_x v_{hp}(x,T)\right) \cdot \nabla_x v_{hp}(x,T) dx \\ &\quad - \frac{\theta}{2p_t^2} \int_Q (\mathfrak{h}_t)' \left(\mathcal{A}\nabla_x v_{hp}\right) \cdot \nabla_x v_{hp} + \frac{\theta}{p_t^2} \left\| (\mathfrak{h}_t)^{1/2} \partial_t v_{hp} \right\|^2_{L^2(Q)} \\ &\geq \frac{1}{2} \int_D \left|v_{hp}(x,T)\right|^2 dx + \lambda_- \left\|\nabla_x v_{hp}\right\|^2_{L^2(Q)} \\ &\quad + \frac{\theta}{p_t^2} \left\| (\mathfrak{h}_t)^{1/2} \partial_t v_{hp} \right\|^2_{L^2(Q)} - \frac{\theta}{2p_t^2} \int_Q (\mathfrak{h}_t)' \left(\mathcal{A}\nabla_x v_{hp}\right) \cdot \nabla_x v_{hp}. \end{aligned}$$

Using (1b), (6) and (7), we have

$$\begin{aligned} \int_Q (\mathfrak{h}_t)' \left(\mathcal{A}\nabla_x v_{hp}\right) \cdot \nabla_x v_{hp} &= \sum_{n=1}^{M_t} \left(\frac{h_t^{n+1} - h_t^{n-1}}{h_t^n}\right) \int_{Q_n} \left(\mathcal{A}\nabla_x v_{hp}\right) \cdot \nabla_x v_{hp} \\ &\leq C_t \lambda_+ \sum_{n=1}^{M_t} \int_{Q_n} \left|\nabla_x v_{hp}\right|^2 = C_t \lambda_+ \left\|\nabla_x v_{hp}\right\|^2_{L^2(Q)}, \end{aligned}$$

and the result follows from

$$-\frac{\theta}{2p_t^2} \int_Q (\mathfrak{h}_t)' \left(\mathcal{A}\nabla_x v_{hp}\right) \cdot \nabla_x v_{hp} \geq -\frac{C_t \theta \lambda_+}{2p_t^2} \left\|\nabla_x v_{hp}\right\|^2_{L^2(Q)}.$$

Lemma 2 *There exists $\gamma > 0$ such that*

$$\mathfrak{a}_{hp}(v, w_{hp}) \le \gamma \|v\|_{hp,*} \left\|w_{hp}\right\|_{hp}, \qquad \forall v \in \mathcal{V}_{hp,*},\ w_{hp} \in \mathcal{V}_{hp}.$$

Proof For $v \in \mathcal{V}_{hp,*}$ and $w_{hp} \in \mathcal{V}_{hp}$, it holds

$$\begin{aligned}
\mathfrak{a}_{hp}(v, w_{hp}) &= -\int_Q v\partial_t w_{hp} + \int_D v(x,T) w_{hp}(x,T) dx + \int_Q (\mathcal{A}\nabla_x v)\cdot\nabla_x w_{hp} \\
&\quad + \frac{\theta}{p_t^2}\int_Q \mathfrak{h}_t \partial_t v \partial_t w_{hp} + \frac{\theta}{p_t^2}\int_Q \mathfrak{h}_t \left(\mathcal{A}\nabla_x v\right)\cdot \nabla_x \partial_t w_{hp} \\
&\le \left(\frac{p_t^2}{\theta}\left\|(\mathfrak{h}_t)^{-1/2} v\right\|_{L^2(Q)}\right)^{1/2}\left(\frac{\theta}{p_t^2}\left\|(\mathfrak{h}_t)^{1/2}\partial_t w_{hp}\right\|^2_{L^2(Q)}\right)^{1/2} \\
&\quad + 2\left(\frac{1}{2}\int_D |v(x,T)|^2\, dx\right)^{1/2}\left(\frac{1}{2}\int_D \left|w_{hp}(x,T)\right|^2 dx\right)^{1/2} \\
&\quad + \lambda_+ \|\nabla_x v\|_{L^2(Q)}\left\|\nabla_x w_{hp}\right\|_{L^2(Q)} \\
&\quad + \left(\frac{\theta}{p_t^2}\left\|(\mathfrak{h}_t)^{1/2}\partial_t v\right\|^2_{L^2(Q)}\right)^{1/2}\left(\frac{\theta}{p_t^2}\left\|(\mathfrak{h}_t)^{1/2}\partial_t w_{hp}\right\|^2_{L^2(Q)}\right)^{1/2} \\
&\quad + \lambda_+ \|\nabla_x v\|_{L^2(Q)}\left(\frac{\theta}{p_t^2}\left\|\mathfrak{h}_t\partial_t\nabla_x w_{hp}\right\|_{L^2(Q)}\right).
\end{aligned}$$

We then need to estimate the last term. From the inverse estimate [17, Theorem 3.91 p.148], it holds

$$\left\|v_p'\right\|^2_{L^2(I_n)} \le \frac{12 p_t^4}{(h_t^n)^2}\|v_p\|^2_{L^2(I_n)}, \qquad \forall v_p \in \mathcal{S}^{p_t}(I, \mathcal{T}^t_{M_t}),\ n = 1,\ldots,M_t.$$

Since this result remains valid for functions taking values in a Hilbert space, we obtain

$$\begin{aligned}
\frac{1}{p_t^4}\left\|\mathfrak{h}_t\partial_t\nabla_x w_{hp}\right\|^2_{L^2(Q)} &= \sum_{n=1}^{M_t}\frac{1}{p_t^4}\int_{Q_n}(\mathfrak{h}_t)^2\left|\partial_t\nabla_x w_{hp}\right|^2 \\
&\le \sum_{n=1}^{M_t}\frac{4C_t^2\,(h_t^n)^2}{p_t^4}\int_{Q_n}\left|\partial_t\nabla_x w_{hp}\right|^2 \le 48 C_t^2\left\|\nabla_x w_{hp}\right\|^2_{L^2(Q)}.
\end{aligned}$$

Hence

$$\begin{aligned}
\mathfrak{a}_{hp}(v, w_{hp}) &\le \left(\frac{p_t^2}{\theta}\left\|(\mathfrak{h}_t)^{-1/2} v\right\|_{L^2(Q)}\right)^{1/2}\left(\frac{\theta}{p_t^2}\left\|(\mathfrak{h}_t)^{1/2}\partial_t w_{hp}\right\|^2_{L^2(Q)}\right)^{1/2} \\
&\quad + 2\left(\frac{1}{2}\int_D |v(x,T)|^2\, dx\right)^{1/2}\left(\frac{1}{2}\int_D \left|w_{hp}(x,T)\right|^2 dx\right)^{1/2} \\
&\quad + \lambda_+ \|\nabla_x v\|_{L^2(Q)}\left\|\nabla_x w_{hp}\right\|_{L^2(Q)} \\
&\quad + \left(\frac{\theta}{p_t^2}\left\|(\mathfrak{h}_t)^{1/2}\partial_t v\right\|^2_{L^2(Q)}\right)^{1/2}\left(\frac{\theta}{p_t^2}\left\|(\mathfrak{h}_t)^{1/2}\partial_t w_{hp}\right\|^2_{L^2(Q)}\right)^{1/2} \\
&\quad + \lambda_+ \|\nabla_x v\|_{L^2(Q)}\left(4\sqrt{3}C_t\theta\left\|\nabla_x w_{hp}\right\|_{L^2(Q)}\right) \\
&\le \frac{2\max\left(1, \lambda_+\left(1 + 4\sqrt{3}C_t\theta\right)\right)}{\min\left(1, \lambda_- - \frac{C_t\lambda_+\theta}{2p_t^2}\right)}\|v\|_{hp,*}\left\|w_{hp}\right\|_{hp} =: \gamma\|v\|_{hp,*}\left\|w_{hp}\right\|_{hp}.
\end{aligned}$$

The definition of l_{hp} yields for $f \in L^2(Q)$ and $w_{hp} \in \mathcal{V}_{hp}$

$$l_{hp}(w_{hp}) \leq 2 \max\left(1, \max_{t \in \overline{I}} \sqrt{\frac{\theta \mathfrak{h}_t(t)}{p_t^2}}\right) \|f\|_{L^2(Q)} \left(\left\|w_{hp}\right\|_{hp}^2 + \left\|w_{hp}\right\|_{L^2(Q)}^2\right)^{1/2}.$$

Since $\mathcal{V}_{hp}$ is finite dimensional, all the norms are equivalent on this space and so $l_{hp} \in {\mathcal{V}_{hp}}^*$. Hence it follows from Lax-Milgram's Theorem [3, 2.5 p.38] that, assuming $f \in L^2(Q)$, there exists a unique $u_{hp} \in \mathcal{V}_{hp}$ satisfying

$$\mathfrak{a}_{hp}(u_{hp}, v_{hp}) = l_{hp}(v_{hp}), \qquad \forall v_{hp} \in \mathcal{V}_{hp}. \tag{9}$$

Lemma 3 *Let $u \in H_0^{1,0}(Q)$ be the unique solution of (3) and assume that $u \in H_{0,\underline{0}}^{1,1}(Q) \cap H^{2,0}(Q)$. Then*

$$\mathfrak{a}_{hp}(u, v_{hp}) = l_{hp}(v_{hp}), \qquad \forall v_{hp} \in \mathcal{V}_{hp}.$$

Proof We know that if $u \in H_{0,\underline{0}}^{1,1}(Q) \cap H^{2,0}(Q)$, then (2a) holds in $L^2(Q)$. Hence multiplying this equation by $v_{hp} + \frac{\theta}{p_t^2} \mathfrak{h}_t \partial_t v_{hp}$ for some $v_{hp} \in \mathcal{V}_{hp}$, integrating over Q and performing integration by parts with respect to x yields the desired result.

It follows from the previous lemma that the following Galerkin orthogonality holds if $u \in H_{0,\underline{0}}^{1,1}(Q) \cap H^{2,0}(Q)$

$$\mathfrak{a}_{hp}(u - u_{hp}, v_{hp}) = 0, \qquad \forall v_{hp} \in \mathcal{V}_{hp},$$

so that for any $w_{hp} \in \mathcal{V}_{hp}$ it holds

$$\begin{aligned}\left\|w_{hp} - u_{hp}\right\|_{hp}^2 &\leq \mathfrak{a}_{hp}(w_{hp} - u_{hp}, w_{hp} - u_{hp}) = \mathfrak{a}_{hp}(w_{hp} - u, w_{hp} - u_{hp}) \\ &\leq \gamma \left\|w_{hp} - u\right\|_{hp,*} \left\|w_{hp} - u_{hp}\right\|_{hp}.\end{aligned}$$

Hence we obtain the following best approximation property for all $w_{hp} \in \mathcal{V}_{hp}$

$$\left\|u - u_{hp}\right\|_{hp} \leq \left\|u - w_{hp}\right\|_{hp} + \left\|u_{hp} - w_{hp}\right\|_{hp} \leq (1 + \gamma) \left\|u - w_{hp}\right\|_{hp,*}. \tag{10}$$

3.2 Error estimates in the $\|\cdot\|_{hp,*}$ norm

As a consequence of (10), we need to derive error estimates in the $\|\cdot\|_{hp,*}$ norm in order to obtain convergence rates for the discrete solution in the $\|\cdot\|_{hp}$ norm. Using that for a.e. $x \in D$

$$v(x,T) = \int_0^T \partial_t v(x,s) ds \leq T^{1/2} \left(\int_0^T \left|\partial_t v(x,s)\right|^2 ds\right)^{1/2}, \qquad \forall v \in H_{\cdot,\underline{0}}^{0,1}(Q),$$

we have

$$\int_D \left|v(x,T)\right|^2 dx \leq T \left\|\partial_t v\right\|_{L^2(Q)}^2,$$

where $H^{0,1}_{\cdot,\underline{0}}(Q) := \{v \in H^{0,1}(Q) |\ v(x,0) = 0,\ \text{for a.e. } x \in D\}$. Hence for all $v \in H^{1,1}_{0,\underline{0}}(Q)$ it holds

$$\|v\|^2_{hp,*} \leq \lambda_- \|\nabla_x v\|^2_{L^2(Q)} + T\|\partial_t v\|^2_{L^2(Q)} + \frac{\theta}{p_t^2}\left\|(\mathfrak{h}_t)^{1/2}\partial_t v\right\|^2_{L^2(Q)} + \frac{p_t^2}{\theta}\left\|(\mathfrak{h}_t)^{-1/2} v\right\|^2_{L^2(Q)}.$$

Let us consider the weighted mixed norm

$$\|v\|^2_{H^{l,1}_{\otimes,\mathfrak{h}_t,p_t}(Q)} := \sum_{0\leq|\alpha|\leq l}\left(\frac{1}{p_t^2}\left\|(\mathfrak{h}_t)^{1/2}\partial_x^\alpha\partial_t v\right\|^2_{L^2(Q)} + p_t^2\left\|(\mathfrak{h}_t)^{-1/2}\partial_x^\alpha v\right\|^2_{L^2(Q)}\right).$$

From (6) and (8), $\|\cdot\|_{H^{l,1}_{\otimes,\mathfrak{h}_t,p_t}(Q)}$ is a norm on $H^{l,1}_\otimes$ equivalent to $\|\cdot\|_{H^{l,1}_\otimes(Q)}$. Then there exists $C > 0$ such that $\|v\|_{hp,*} \leq C\|v\|_{hp,**}$, where

$$\|v\|^2_{hp,**} := \|\nabla_x v\|^2_{L^2(Q)} + \|\partial_t v\|^2_{L^2(Q)} + \|v\|^2_{H^{0,1}_{\otimes,\mathfrak{h}_t,p_t}(Q)}, \qquad \forall v \in H^{1,1}_{0,\underline{0}}(Q).$$

We then only need to derive error estimates in the $\|\cdot\|_{hp,**}$ norm. Let $\pi^1_{\mathcal{T}^x} : H^2(D)\cap H^1_0(D) \to \mathcal{S}^1_0(D,\mathcal{T}^x)$ and $\pi^{p_t}_{\mathcal{T}^t_{M_t}} : H^1_{\underline{0}}(I) \to \mathcal{S}^{p_t}_{\underline{0}}(I,\mathcal{T}^t_{M_t})$ be the interpolation operators defined in [6, Corollary 1.109 p.61] and [17, Theorem 3.17 p.76], respectively, where $H^1_{\underline{0}}(I) := \{v \in H^1(I) |\ v(0) = 0\}$. Then for $\Pi^{1,p_t}_{\mathcal{T}^x,\mathcal{T}^t_{M_t}} := \pi^1_{\mathcal{T}^x} \otimes \pi^{p_t}_{\mathcal{T}^t_{M_t}} : H^{1,1}_{0,\underline{0}}(Q) \cap H^{2,1}_\otimes(Q) \to \mathcal{V}_{hp}$, we write

$$\mathrm{Id} - \Pi^{1,p_t}_{\mathcal{T}^x,\mathcal{T}^t_{M_t}} = \left(\mathrm{Id} - \pi^1_{\mathcal{T}^x}\right)\otimes \mathrm{Id} + \pi^1_{\mathcal{T}^x}\otimes\left(\mathrm{Id} - \pi^{p_t}_{\mathcal{T}^t_{M_t}}\right).$$

Using the tensor structure of the continuous space, it is possible to estimate the errors for the space and the time approximations separately. Indeed, since $\pi^1_{\mathcal{T}^x} \in \mathcal{L}(L^2(D)) \cap \mathcal{L}(H^1_0(D))$, we obtain that there exists $C > 0$ such that for all $v \in H^{1,1}_{0,\underline{0}}(Q) \cap H^{2,1}_\otimes(Q)$

$$\left\|v - \Pi^{1,p_t}_{\mathcal{T}^x,\mathcal{T}^t_{M_t}} v\right\|_{hp,**} \leq \left\|v - \left(\pi^1_{\mathcal{T}^x}\otimes \mathrm{Id}\right)v\right\|_{hp,**} + C\left\|v - \left(\mathrm{Id}\otimes\pi^{p_t}_{\mathcal{T}^t_{M_t}}\right)v\right\|_{hp,**}. \quad (11)$$

Lemma 4 *Let $D \subset \mathbb{R}^d$ $(d = 1,2,3)$ be a bounded convex polyhedral domain and $\mathcal{T}^x$ be a quasi-uniform mesh of D with meshsize h_x. Then there exists $C > 0$ such that for all $v \in H^{1,1}_{0,\underline{0}}(Q) \cap H^{2,1}_\otimes(Q)$ it holds*

$$\left\|v - \left(\pi^1_{\mathcal{T}^x}\otimes \mathrm{Id}\right)v\right\|_{hp,**} \leq Ch_x\left(\|v\|_{H^{2,0}(Q)} + h_x\left(\|v\|_{H^{2,1}_\otimes(Q)} + \|v\|_{H^{2,1}_{\otimes,\mathfrak{h}_t,p_t}(Q)}\right)\right).$$

Proof In [6, Example 1.111 (i)], it has been shown that

$$\left\|\tilde{v} - \pi^1_{\mathcal{T}^x}\tilde{v}\right\|_{L^2(D)} + h_x\left|\tilde{v} - \pi^1_{\mathcal{T}^x}\tilde{v}\right|_{H^1(D)} \leq Ch_x^2|\tilde{v}|_{H^2(D)}, \qquad \forall \tilde{v} \in H^2(D).$$

Following the lines of the proof of this result, we see that the above norms can be replaced by Bochner-Sobolev norms with values in a Hilbert space. Hence for every $v \in H^{1,1}_{0,\underline{0}}(Q) \cap H^{2,1}_{\otimes}(Q)$ it holds

$$\begin{aligned} \left| v - \left(\pi^1_{\mathcal{T}^x} \otimes \mathrm{Id}\right) v \right|_{1,0,Q} &\le C h_x \left| v \right|_{2,0,Q} , \\ \left\| v - \left(\pi^1_{\mathcal{T}^x} \otimes \mathrm{Id}\right) v \right\|_{0,1,Q} &\le C h_x^2 \left| v \right|_{2,1,Q} . \end{aligned}$$

Denoting

$$\|v\|^2_{H^1_{\mathfrak{h}_t,p_t}(I)} := \frac{1}{p_t^2} \left\| (\mathfrak{h}_t)^{1/2} \partial_t v \right\|^2_{L^2(I)} + p_t^2 \left\| (\mathfrak{h}_t)^{-1/2} v \right\|^2_{L^2(I)} ,$$

we have from (6) and (8) that $\|\cdot\|_{H^1_{\mathfrak{h}_t,p_t}(I)}$ is a norm on $H^1(I)$ equivalent to $\|\cdot\|_{H^1(I)}$. The result then follows from

$$\left\| v - \left(\pi^1_{\mathcal{T}^x} \otimes \mathrm{Id}\right) v \right\|_{H^{0,1}_{\otimes,\mathfrak{h}_t,p_t}(Q)} \le C h_x^2 \left\| v \right\|_{H^{2,1}_{\otimes,\mathfrak{h}_t,p_t}(Q)} .$$

Lemma 5 *Let $v \in H^{1,1}_{0,\underline{0}}(Q) \cap H^{2,1}_{\otimes}(Q)$ such that for all $n \in \{2, \dots, M_t\}$, there exists $s_n \in \mathbb{N}_0$ satisfying $v \in H^{1,s_n+1}_{\otimes}(Q_n)$ and $v \in H^{0,s_1+1}(Q_1)$ for some $s_1 \in \mathbb{N}_0$. Then*

$$\begin{aligned} \left\| v - \left(\mathrm{Id} \otimes \pi^{p_t}_{\mathcal{T}^t_{M_t}}\right) v \right\|^2_{hp,**} \le{}& \left| v - \left(\mathrm{Id} \otimes \pi^{p_t}_{\mathcal{T}^t_{M_t}}\right) v \right|^2_{1,0,Q_1} \\ &+ \sum_{n=2}^{M_t} \left(\frac{h_t^n}{2}\right)^{2(s_n+1)} \frac{\Gamma(p_t - s_n + 1)}{p_t(p_t+1)\Gamma(p_t + s_n + 1)} \left| v \right|^2_{1,s_n+1,Q_n} \\ &+ (4C_t + 1) \sum_{n=1}^{M_t} \left(\frac{h_t^n}{2}\right)^{2s_n} \frac{\Gamma(p_t - s_n + 1)}{\Gamma(p_t + s_n + 1)} \left| v \right|^2_{0,s_n+1,Q_n} , \end{aligned}$$

where C_t is defined in (6).

Proof We consider a similar argument to the one used in the proof of the previous lemma. Following the lines of the proof of Theorem 3.17 in [17], we see that for $j = 0, 1$ and $n = 1, \dots, M_t$

$$\begin{aligned} \left| v - \left(\mathrm{Id} \otimes \pi^{p_t}_{\mathcal{T}^t_{M_t}}\right) v \right|^2_{j,0,Q_n} &\le \left(\frac{h_t^n}{2}\right)^{2(s_n+1)} \frac{\Gamma(p_t - s_n + 1)}{p_t(p_t+1)\Gamma(p_t + s_n + 1)} \left| v \right|^2_{j,s_n+1,Q_n} , \\ \left| v - \left(\mathrm{Id} \otimes \pi^{p_t}_{\mathcal{T}^t_{M_t}}\right) v \right|^2_{j,1,Q_n} &\le \left(\frac{h_t^n}{2}\right)^{2s_n} \frac{\Gamma(p_t - s_n + 1)}{\Gamma(p_t + s_n + 1)} \left| v \right|^2_{j,s_n+1,Q_n} . \end{aligned}$$

Hence

$$\begin{aligned} \left| v - \left(\mathrm{Id} \otimes \pi^{p_t}_{\mathcal{T}^t_{M_t}}\right) v \right|^2_{1,0,Q} \le{}& \left| v - \left(\mathrm{Id} \otimes \pi^{p_t}_{\mathcal{T}^t_{M_t}}\right) v \right|^2_{1,0,Q_1} \\ &+ \sum_{n=2}^{M_t} \left(\frac{h_t^n}{2}\right)^{2(s_n+1)} \frac{\Gamma(p_t - s_n + 1)}{p_t(p_t+1)\Gamma(p_t + s_n + 1)} \left| v \right|^2_{1,s_n+1,Q_n} , \\ \left| v - \left(\mathrm{Id} \otimes \pi^{p_t}_{\mathcal{T}^t_{M_t}}\right) v \right|^2_{0,1,Q} \le{}& \sum_{n=1}^{M_t} \left(\frac{h_t^n}{2}\right)^{2s_n} \frac{\Gamma(p_t - s_n + 1)}{\Gamma(p_t + s_n + 1)} \left| v \right|^2_{0,s_n+1,Q_n} . \end{aligned}$$

Turning to the $\|\cdot\|_{H^{0,1}_{\otimes,\mathfrak{h}_t,p_t}(Q)}$ part of the norm, we have from (8)

$$\|v\|^2_{H^{0,1}_{\otimes,\mathfrak{h}_t,p_t}(Q)} \leq 2C_t \sum_{n=1}^{M_t} \left(\frac{h_t^n}{p_t^2} |v|^2_{0,1,Q_n} + \frac{p_t^2}{h_t^n} |v|^2_{0,0,Q_n} \right), \qquad \forall v \in H^{0,1}(Q),$$

so that

$$\left\| v - \left(\mathrm{Id} \otimes \pi^{p_t}_{\mathcal{T}^t_{M_t}} \right) v \right\|^2_{H^{0,1}_{\otimes,\mathfrak{h}_t,p_t}(Q)} \leq 4C_t \sum_{n=1}^{M_t} \left(\frac{h_t^n}{2} \right)^{2s_n+1} \frac{\Gamma(p_t - s_n + 1)}{\Gamma(p_t + s_n + 1)} |v|^2_{0,s_n+1,Q_n}.$$

3.2.1 hp-approximation in time

Using the previous results, we consider an approximation of the space-time solution over geometric meshes for the time variable. More precisely, let $\sigma_t \in (0,1)$ be a grading factor and for $M_t \in \mathbb{N}$, define

$$t_0^{\sigma_t} = 0, \qquad t_n^{\sigma_t} = T\sigma_t^{M_t-n}, \qquad n = 1, \ldots, M_t, \tag{12}$$

and $\mathcal{T}^t_{\sigma_t,M_t} := \{I_n^{\sigma_t}\}_{n=1}^{M_t}$, where $I_n^{\sigma_t} := (t_{n-1}^{\sigma_t}, t_n^{\sigma_t})$ for $n = 1, \ldots, M_t$. Furthermore, we write $Q_n^{\sigma_t} := D \times I_n^{\sigma_t}$ for $n = 1, \ldots, M_t$. Note that for $\kappa := \left(\sigma_t^{-1} - 1\right)$, it holds

$$h^n_{t,\sigma_t} = \kappa T \sigma_t^{M_t-(n-1)}, \qquad n = 1, \ldots, M_t. \tag{13}$$

Combining Proposition 1 and Lemma 4.3.3 in [5], we obtain the following result.

Lemma 6 *Let $f \in L^2\left(D; \mathscr{G}^\delta(I)\right)$ for some $\delta \geq 1$. Then there exists $C = C(\sigma_t) > 0$ such that for all $l \in \mathbb{N}$ and $n \in \{2, \ldots, M_t\}$, it holds*

$$\begin{aligned}
|u|^2_{0,1,Q_1^{\sigma_t}} &\leq CTd^2\sigma_t^{M_t}, \\
|u|^2_{0,l,Q_n^{\sigma_t}} &\leq CT^3 \left(\frac{d}{T}\right)^{2l} \Gamma(l+1)^{2\delta} \sigma_t^{(M_t-n+1)(3-2l)}, \\
|u|^2_{1,l,Q_n^{\sigma_t}} &\leq CT^2 \left(\frac{d}{T}\right)^{2l} \Gamma(l+1)^{2\delta} \sigma_t^{2(M_t-n+1)(1-l)},
\end{aligned}$$

where u is the solution of (3), $d := \max(1, d_f T)$ and d_f is the constant in the bound (4) associated to f.

Lemma 7 *Let $f \in L^2\left(D; \mathscr{G}^\delta(I)\right)$ for some $\delta \geq 1$. Then there exists $C > 0$ such that*

$$\left| u - \left(\mathrm{Id} \otimes \pi^{p_t}_{\mathcal{T}^t_{M_t}} \right) u \right|^2_{1,0,Q_1} \leq Ct_1,$$

where u is the solution of (3).

Proof By construction of $\pi^{p_t}_{\mathcal{T}^t_{M_t}}$, it holds [17, Theorem 3.14]

$$\left| u - \left(\mathrm{Id} \otimes \pi^{p_t}_{\mathcal{T}^t_{M_t}}\right) u \right|_{1,0,Q_1} \leq \left| u - \left(\mathrm{Id} \otimes \pi^{1}_{\mathcal{T}^t_{M_t}}\right) u \right|_{1,0,Q_1}.$$

Moreover $u \in H^1(D; \mathscr{G}^{\delta,1/2}(I))$ by Lemma 4.3.3 in [5]. Since the operator $\pi^1_{\mathcal{T}^t_{M_t}}$ maps to $\mathcal{S}^1(I, \mathcal{T}^t_{M_t})$ and using the initial condition, we obtain that for a.e. $x \in D$

$$\left(\mathrm{Id} \otimes \pi^{1}_{\mathcal{T}^t_{M_t}}\right) u(x,t) = \frac{t}{t_1} u(x, t_1), \qquad \forall t \in I_1.$$

Hence

$$\begin{aligned}
\left| u - \left(\mathrm{Id} \otimes \pi^{p_t}_{\mathcal{T}^t_{M_t}}\right) u \right|^2_{1,0,Q_1} &\leq 2 \left(|u|^2_{1,0,Q_1} + \left| \left(\mathrm{Id} \otimes \pi^{1}_{\mathcal{T}^t_{M_t}}\right) u \right|^2_{1,0,Q_1} \right) \\
&\leq C \left(\int_0^{t_1} dt + \frac{1}{t_1^2} \int_0^{t_1} t^2 dt \right) \leq C t_1,
\end{aligned}$$

where we have used that $u \in \mathcal{C}(\overline{Q})$ since $u \in H^1(D; \mathscr{G}^{\delta,1/2}(I)) \subset H^1(D; \mathcal{C}(\overline{I}))$.

Proposition 2 *Let us assume that $f \in L^2(D; \mathscr{G}^\delta(I))$ for some $\delta \geq 1$. For $\sigma_t \in (0,1)$ and $M_t \in \mathbb{N}$, let further $\mathcal{T}^t_{\sigma_t, M_t}$ be defined through (12). Then there exists $\mu_{t,0} \geq 1$ such that choosing the vector of temporal polynomial orders to depends polynomially on the time step, i.e., $p_t = \left\lfloor \mu_t M_t^\delta \right\rfloor$ for any sufficiently large slope $\mu_t > \mu_{t,0}$, there exists $C > 0$ satisfying*

$$\left\| u - \left(\mathrm{Id} \otimes \pi^{p_t}_{\mathcal{T}^t_{\sigma_t, M_t}}\right) u \right\|_{hp,**} \leq C \sigma_t^{\frac{M_t}{2}}.$$

Proof Let us denote $I := \|u - \left(\mathrm{Id} \otimes \pi^{p_t}_{\mathcal{T}^t_{\sigma_t, M_t}}\right) u\|^2_{hp,**}$. Then, from Lemmas 5, 6 and 7, taking $s_1 = 0$, we have

$$\begin{aligned}
I \leq &\left| u - \left(\mathrm{Id} \otimes \pi^{p_t}_{\mathcal{T}^t_{\sigma_t, M_t}}\right) u \right|^2_{1,0,Q_1} + (4C_t + 1) |u|^2_{0,1,Q_1^{\sigma_t}} \\
&+ \sum_{n=2}^{M_t} \left(\frac{h^n_{t,\sigma_t}}{2} \right)^{2(s_n+1)} \frac{\Gamma(p_t - s_n + 1)}{p_t(p_t+1)\Gamma(p_t + s_n + 1)} |u|^2_{1,s_n+1,Q_n} \\
&+ (4C_t + 1) \sum_{n=2}^{M_t} \left(\frac{h^n_{t,\sigma_t}}{2} \right)^{2s_n} \frac{\Gamma(p_t - s_n + 1)}{\Gamma(p_t + s_n + 1)} |u|^2_{0,s_n+1,Q_n} \\
&\leq C t_1^{\sigma_t} + C \sigma_t^{M_t} \\
&+ C \sum_{n=2}^{M_t} \left(\frac{d h^n_{t,\sigma_t}}{2T} \right)^{2(s_n+1)} \frac{\Gamma(p_t - s_n + 1)\Gamma(s_n+2)^{2\delta}}{p_t(p_t+1)\Gamma(p_t + s_n + 1)} \sigma_t^{-2(M_t - n + 1)s_n} \\
&+ C \sum_{n=2}^{M_t} \left(\frac{d h^n_{t,\sigma_t}}{2T} \right)^{2s_n} \left(\frac{d}{T} \right)^2 \frac{\Gamma(p_t - s_n + 1)\Gamma(s_n+2)^{2\delta}}{\Gamma(p_t + s_n + 1)} \sigma_t^{(M_t - n + 1)(1 - 2s_n)}.
\end{aligned}$$

From (13), we have

$$\left(\frac{dh_{t,\sigma_t}^n}{2T}\right)^{2(s_n+1)}\sigma_t^{-2(M_t-n+1)s_n}=\left(\frac{\kappa d}{2}\right)^{2(s_n+1)}\sigma_t^{2(M_t-n+1)},$$
$$\left(\frac{dh_{t,\sigma_t}^n}{2T}\right)^{2s_n}\left(\frac{d}{T}\right)^2\sigma_t^{(M_t-n+1)(1-2s_n)}=\left(\frac{2}{\kappa T}\right)^2\left(\frac{\kappa d}{2}\right)^{2(s_n+1)}\sigma_t^{(M_t-n+1)}.$$

Choosing $s_n=n+1$ for all $n\in\{1,\dots,M_t\}$, there exists $C>0$ satisfying

$$\begin{aligned}\left\|u-\left(\mathrm{Id}\otimes\pi^{p_t}_{\mathcal{T}^t_{\sigma_t,M_t}}\right)u\right\|^2_{hp,**}&\leq C\sigma_t^{M_t}\left(1+\sum_{n=2}^{M_t}\left(\frac{\kappa d}{2\sqrt{\sigma_t}}\right)^{2n}\frac{\Gamma(p_t-n)\Gamma(n+3)^{2\delta}}{\Gamma(p_t+n+2)}\right)\\&=C\sigma_t^{M_t}\left(1+\sum_{n=1}^{M_t}\alpha^{2n}\frac{\Gamma(p_t-n)}{\Gamma(p_t+n+2)}\Gamma(n+3)^{2\delta}\right),\end{aligned}$$

where $\alpha:=\frac{\kappa d}{2\sqrt{\sigma_t}}$. From [5, (A.1.3),(A.1.4)], it holds for $p_t=\left\lfloor\mu_tM_t^\delta\right\rfloor$

$$\begin{aligned}\Gamma(n+3)&\leq Cn^{5/2}\left(\frac{n}{e}\right)^n\leq Cn^{5/2}\left(\frac{M_t}{e}\right)^n,\\\frac{\Gamma(p_t-n)}{\Gamma(p_t+n+2)}&=\frac{\Gamma(\left\lfloor\mu_tM_t^\delta\right\rfloor-n)}{\Gamma(\left\lfloor\mu_tM_t^\delta\right\rfloor+n+2)}\\&\leq e\left(\frac{\left\lfloor\mu_tM_t^\delta\right\rfloor}{e}\right)^{-2(1+n)}\leq e\left(\frac{\mu_tM_t^\delta}{2e}\right)^{-2(1+n)},\end{aligned}$$

where we have used $\left\lfloor\mu_tM_t^\delta\right\rfloor\geq\frac{\mu_tM_t^\delta}{2}$. Hence for every

$$\mu_t>\mu_{t,0}:=\frac{2\alpha}{e^{\delta-1}},$$

it holds

$$\begin{aligned}\sum_{n=1}^{M_t}\alpha^{2n}\frac{\Gamma(\left\lfloor\mu_tM_t^\delta\right\rfloor-n+2)}{\Gamma(\left\lfloor\mu_tM_t^\delta\right\rfloor+n)}\Gamma(n+1)^{2\delta}&\leq C\left(\mu_tM_t^\delta\right)^{-2}\sum_{n=1}^{M_t}\alpha^{2n}\left(\frac{2e}{\mu_tM_t^\delta}\frac{M_t^\delta}{e^\delta}\right)^{2n}n^{5\delta}\\&\leq C\left(\mu_tM_t^\delta\right)^{-2}\sum_{n=1}^{\infty}n^{5\delta}\left(\frac{2\alpha}{\mu_te^{\delta-1}}\right)^{2n},\end{aligned}$$

and it can be shown using for instance a ratio test that the series converges. Wrapping up everything, we obtain the desired result.

Theorem 1 *Let us assume that $f\in L^2(D;\mathscr{G}^\delta(I))$ for some $\delta\geq 1$. For $\sigma_t\in(0,1)$ and $M_t\in\mathbb{N}$, let further $\mathcal{T}^t_{\sigma_t,M_t}$ be defined through* (12) *and choose $h_x\approx\sigma_t^{\frac{M_t}{2}}$. Then there exists $\mu_{t,0}\geq 1$ such that choosing $p_t=\left\lfloor\mu_tM_t^\delta\right\rfloor$ for any $\mu_t>\mu_{t,0}$, there exists $C>0$ satisfying*

$$\left\|u-u_{hp}\right\|_{hp}\leq C\log(N)^{\frac{\delta(d+1)+1}{d}}N^{-\frac{1}{d}},\tag{14}$$

where $N=\dim\left(\mathcal{V}_{hp}\right)$.

Proof Let $\mu_{t,0} > 0$ be defined in Proposition 2. From $h_x \approx \sigma_t^{\frac{M_t}{2}}$, (10), (11), Lemma 4 and Proposition 2, there exists $C > 0$ such that

$$\left\|u - u_{hp}\right\|_{hp} \leq C \left(h_x \left(\|u\|_{H^{2,0}(Q)} + h_x \left(\|u\|_{H^{2,1}_{\otimes}(Q)} + \|u\|_{H^{2,1}_{\otimes,\mathfrak{h}_t,p_t}(Q)} \right) \right) + \sigma_t^{\frac{M_t}{2}} \right)$$
$$\leq C h_x \left(\max \left(1, \|u\|_{H^{2,1}_{\otimes}(Q)} \right) + h_x \|u\|_{H^{2,1}_{\otimes,\mathfrak{h}_t,p_t}(Q)} \right).$$

Furthermore

$$\begin{aligned} h_x^2 \|u\|^2_{H^{2,1}_{\otimes,\mathfrak{h}_t,p_t}(Q)} &= h_x^2 \sum_{0\leq|\alpha|\leq 2} \left(\frac{1}{p_t^2} \left\| \mathfrak{h}_t^{1/2} \partial_x^\alpha \partial_t u \right\|^2_{L^2(Q)} + p_t^2 \left\| \mathfrak{h}_t^{-1/2} \partial_x^\alpha u \right\|^2_{L^2(Q)} \right) \\ &\leq h_x^2 \left(\frac{h^{M_t}_{t,\sigma_t}}{p_t^2} + \frac{p_t^2}{h^1_{t,\sigma_t}} \right) \sum_{0\leq|\alpha|\leq 2} \left(\left\| \partial_x^\alpha \partial_t u \right\|^2_{L^2(Q)} + \left\| \partial_x^\alpha u \right\|^2_{L^2(Q)} \right) \\ &\leq C \sigma_t^{M_t} \left(\frac{1}{M_t^{2\delta}} + \frac{M_t^{2\delta}}{\sigma_t^{M_t}} \right) \|u\|^2_{H^{2,1}_{\otimes}(Q)} \leq C M_t^{2\delta} \|u\|^2_{H^{2,1}_{\otimes}(Q)}, \end{aligned}$$

so that
$$\left\|u - u_{hp}\right\|_{hp} \leq C M_t^{\delta} h_x \max \left(1, \|u\|_{H^{2,1}(Q)} \right).$$

From $N_x := \dim \left(\mathcal{S}_0^1(D, \mathcal{T}^x) \right) \approx (h_x)^{-d}$, $N_t \approx M_t^{\delta+1}$ and $h_x \approx \sigma_t^{\frac{M_t}{2}}$, we have

$$N := N_x N_t \approx (h_x)^{-d} N_t \approx \sigma_t^{-\frac{dM_t}{2}} M_t^{\delta+1}$$

so that $N^{-\frac{1}{d}} \approx \sigma_t^{\frac{M_t}{2}} M_t^{-\frac{\delta+1}{d}}$ and $\log(N) \approx M_t$, whence

$$M_t^{\delta} h_x \leq C \sigma_t^{\frac{M_t}{2}} M_t^{\delta} = C M_t^{\frac{\delta+1}{d}+\delta} \sigma_t^{\frac{M_t}{2}} M_t^{-\frac{\delta+1}{d}} \approx \left(\frac{\log(N)^{\delta+1}}{N} \right)^{\frac{1}{d}} \log(N)^{\delta}.$$

Remark 1 Following the lines of the proof, we see that a sufficient condition for (14) to hold is that $u \in L^2 \left(D; \mathscr{G}^{\delta,1}(I) \right) \cap H_0^1 \left(D; \mathscr{G}^{\delta,1/2}(I) \right) \cap H^{2,1}_{\otimes}(Q)$.

3.3 Generalization to higher order FEM in two space dimensions

Let $D \subset \mathbb{R}^2$ be a polygonal domain and denote its vertices by A_i, $i = 1, \ldots, L$. In order to deal with polynomial approximation of order $p_x > 1$ in the spatial variable, we consider weighted Sobolev spaces. Indeed, solutions of elliptic problems will typically belong to these spaces. This is a consequence of the polygonal shape of the domain which might induce corner singularities. We then expect solutions of parabolic equations to fulfill a similar type of regularity in space. To circumvent this lack of regularity, it is possible to consider meshes which are graded towards the corners of the domain. In that particular case, it is not possible to obtain convergence rates with respect to h_x and so from now on we will only work with N_x, i.e. the number of degrees freedom related to the spatial approximation.

3.3.1 Approximation of higher order

Let us assume that the solution u of (3) satisfies

$$u = \sum_{i=0}^{L} u_i, \tag{15}$$

where $u_0 \in H^{p_x+1}(D; H^1(I))$ and for $i = 1, \ldots, L$, u_i can be written in polar coordinates as

$$u_i(x) = c_i r_i^{\gamma_i} g_i(\theta_i) \chi_i,$$

where $c_i \in H^1(I)$ and $\gamma_i > 0$ are constants, r_i denotes the distance to the vertex A_i, $\theta_i \in [0, 2\pi)$ is the angle of x with respect to A_i and a half line starting at A_i, $g_i \in W^{1,\infty}(0, 2\pi)$ is 2π-periodic and piecewise $W^{p_x+1,\infty}$ and χ_i is a $C_\infty(\overline{D})$ cut-off function. Under these assumptions, we have the following result. It can be obtained using [8, Theorem 2.1] and the same procedure as in the proof of Lemma 4. The factor $\frac{p_t}{\sqrt{h_t^{min}}}$ comes from the mesh-dependent part of the norm $\|\cdot\|_{hp,**}$.

Lemma 8 *Let u fulfill* (15)*. Then there exists a sequence of meshes $\{\mathcal{T}_n^x\}_{n\in\mathbb{N}}$ and $C > 0$ such that*

$$\left\| u - \left(\pi^{p_x}_{\mathcal{T}^x_{N_x}} \otimes \mathrm{Id} \right) u \right\|_{hp,**} \leq C \frac{p_t}{\sqrt{h_t^{min}}} N_x^{-\frac{p_x}{2}}, \tag{16}$$

where $h_t^{min} := \min_{n=1,\ldots,M_t} h_t^n$ and C depends on $|u|_{H^{p_x+1}(D)}$ and $c_i, \gamma_i, \chi_i, g_i$ for $i = 1, \ldots, L$.

Proposition 3 *Let $p_x \in \mathbb{N}$, $f \in L^2(D; \mathscr{G}^\delta(I))$ for some $\delta \geq 1$ and assume that u admits the decomposition* (15)*. For $\sigma_t \in (0,1)$ and $M_t \in \mathbb{N}$, let further $\mathcal{T}^t_{\sigma_t, M_t}$ be defined through* (12) *and choose $N_x \approx \sigma_t^{-\frac{2M_t}{p_x}}$. Then there exists $\mu_{t,0} \geq 1$ such that, choosing $p_t = \left\lfloor \mu_t M_t^\delta \right\rfloor$ for any $\mu_t > \mu_{t,0}$, there exists $C > 0$ satisfying*

$$\left\| u - u_{hp} \right\|_{hp} \leq C \log(N)^{\frac{p_x(\delta+1)}{4} + \delta} N^{-\frac{p_x}{4}}, \tag{17}$$

where $N = \dim \left(\mathcal{V}_{hp} \right)$.

Proof Let $\mu_{t,0} > 0$ be as defined in Proposition 2. From $N_x^{-\frac{p_x}{2}} \approx \sigma_t^{M_t}$, (10), (11), Proposition 2 and Lemma 8, there exists $C > 0$ such that

$$\begin{aligned} \left\| u - u_{hp} \right\|_{hp} &\leq C \left(\frac{p_t}{\sqrt{h_t^{min}}} N_x^{-\frac{p_x}{2}} + \sigma_t^{\frac{M_t}{2}} \right) \\ &\leq C \left(M_t^\delta \sigma_t^{-\frac{M_t}{2}} N_x^{-\frac{p_x}{2}} + \sigma_t^{\frac{M_t}{2}} \right) \leq C M_t^\delta \sigma_t^{\frac{M_t}{2}}. \end{aligned}$$

From $N = N_x N_t \approx \sigma_t^{-\frac{2M_t}{p_x}} M_t^{\delta+1}$, it holds $\log(N) \approx M_t$ and

$$N^{-\frac{p_x}{4}} \log(N)^{\frac{p_x(\delta+1)}{4} + \delta} \approx \sigma_t^{\frac{M_t}{2}} M_t^{-\frac{p_x(\delta+1)}{4}} M_t^{\frac{p_x(\delta+1)}{4} + \delta} = \sigma_t^{\frac{M_t}{2}} M_t^\delta,$$

and the result follows.

Remark 2 For $p_x = 1$, (17) yields

$$\left\|u - u_{hp}\right\|_{hp} \leq C \log(N)^{\frac{5\delta+1}{4}} N^{-\frac{1}{4}},$$

while inserting $d = 1$ in (14) gives

$$\left\|u - u_{hp}\right\|_{hp} \leq C \log(N)^{\frac{3\delta+1}{2}} N^{-\frac{1}{2}}.$$

Hence in the case $p_x = 1$, the result of Proposition 3 is not sharp. This is a consequence of Lemma 8 and more precisely of Theorem 2.1 from [8]. Indeed, in this paper an approximation result for continuous piecewise polynomial elements on graded meshes is obtained in the H^1-norm. However the term of $\|\cdot\|_{hp,**}$ containing the mesh-dependent norm involves only the L^2-norm in space. Hence in principle, (16) could be sharpened to

$$\left\|v - \left(\pi^{p_x}_{\mathcal{T}^x_{N_x}} \otimes \mathrm{Id}\right)v\right\|_{hp,**} \leq C\left(\frac{p_t}{\sqrt{h_t^{min}}} N_x^{-\frac{p_x+1}{2}} + N_x^{-\frac{p_x}{2}}\right),$$

allowing us to prove a better bound in Proposition 3. In fact, choosing $N_x \approx \sigma_t^{-\frac{2M_t}{p_x+1}}$ yields for $p_x \geq 1$

$$\begin{aligned}\left\|u - u_{hp}\right\|_{hp} &\leq C \max\left(\log(N)^{\frac{(p_x+1)(\delta+1)}{4}+\delta} N^{-\frac{p_x+1}{4}}, N_x^{-\frac{p_x}{2}}\right)\\ &= C \log(N)^{\frac{(p_x+1)(\delta+1)}{4}+\delta} N^{-\frac{p_x+1}{4}}.\end{aligned}$$

In that case, $p_x = 1$ gives the result from Proposition 3.

3.3.2 hp-approximation for the spatial variable

In order to deal with an *hp*-approximation in the spatial component of u, we consider the following classes of functions. The reason is that they describe the regularity of solutions of elliptic equations when the associated right-hand side is analytic and hence we expect the same kind of behavior for the spatial component of solutions of parabolic equations. However, to the best of our knowledge, this result can not be found in the literature. Let $\beta = (\beta_1, \dots, \beta_L) \in [0,1)^L$ and the weight function Φ_β defined for $x \in \overline{D}$ as

$$\Phi_\beta(x) := \prod_{j=1}^{L} (r_j(x))^{\beta_j}, \qquad r_i(x) := \min(1, |x - A_i|), \qquad i = 1, \dots, L.$$

Then for $m \geq l \geq 1$ and a Hilbert space H, $H^{m,l}_\beta(D;H)$ is the space of functions $v \in L^2(D;H)$ for which

$$\|v\|^2_{H^{m,l}_\beta(D;H)} := \|v\|^2_{H^{l-1}(D;H)} + |v|^2_{H^{m,l}_\beta(D;H)} < \infty,$$

where

$$|v|^2_{H^{m,l}_\beta(D;H)} := \sum_{k=l}^{m} \sum_{|\alpha|=k} \left\|\Phi_{\beta+k-l} \partial_x^\alpha v\right\|^2_{L^2(D;H)}.$$

Definition 2 For $l \geq 1$, $v \in \mathcal{B}^l_\beta(D;H)$ if $v \in H^{m,l}_\beta(D;H)$ for all $m \geq l$ and there exist $C > 0$ and $d \geq 1$ such that for all $k \geq l$ it holds

$$\sum_{|\alpha|=k} \left\| \Phi_{\beta+k-l} \partial_x^\alpha v \right\|_{L^2(D;H)} \leq C d^{k-l} \Gamma(k-l+1).$$

Functions in $\mathcal{B}^2_\beta(D;H)$ are globally continuous, analytic inside D and contain potential algebraic singularities at the vertices $A_1, \ldots, A_L$. We then consider an *hp*-discretization in D. Given a grading factor $\sigma_x \in (0,1)$ and $M_x \in \mathbb{N}$, we consider a mesh $\mathcal{T}^x_{\sigma_x,M_x}$ of D which is exponentially graded towards the corners and has elements of fixed size away from ∂D. The idea is similar to the construction of $\mathcal{T}^t_{\sigma_t,M_t}$. Furthermore, given $\mu_x > 0$, we consider a polynomial distribution $\mathbf{p}_x^{\mu_x}$ over $\mathcal{T}^x_{\sigma_x,M_x}$ such that $\mathbf{p}_x^{\mu_x}\big|_K = 1$ for an element $K \in \mathcal{T}^x_{\sigma_x,M_x}$ such that there exists $i \in \{1,\ldots,L\}$ with $A_i \in \overline{K}$. We then suppose that $\mathbf{p}_x^{\mu_x}$ increases linearly with slope μ_x away from the corners and is constant and proportional to M_x over the elements of fixed size of $\mathcal{T}^x_{\sigma_x,M_x}$. A precise description of how to build $\mathcal{T}^x_{\sigma_x,M_x}$ and $\mathbf{p}^x_{\sigma_x,M_x}$ and a proof of the following Lemma can be found in [17, Chapter 4]. Again, using the same procedure as in the proof of Lemma 4, we obtain the following result.

Lemma 9 *Let $\sigma_x \in (0,1)$ and assume that $u \in \mathcal{B}^2_\beta(D;H^1(I))$. Then there exists $\mu_{x,0} > 0$ such that for any $\mu_x > \mu_{x,0}$, there exist $C, b > 0$ satisfying for $M_x \in \mathbb{N}$*

$$\left\| u - \left(\pi^{\mathbf{p}_x^{\mu_x}}_{\mathcal{T}^x_{\sigma_x,M_x}} \otimes \mathrm{Id} \right) u \right\|_{hp,**} \leq C \frac{p_t}{\sqrt{h_t^{min}}} \exp\left(-b N_x^{\frac{1}{3}} \right). \tag{18}$$

Proposition 4 *Let $f \in L^2(D;\mathscr{G}^\delta(I))$ for some $\delta \geq 1$ and $u \in \mathcal{B}^2_\beta(D;H^1(I))$. For $\sigma_t, \sigma_x \in (0,1)$ and $M_t, M_x \in \mathbb{N}$, let further $\mathcal{T}^t_{\sigma_t,M_t}$ be defined through (12) and $\mathcal{T}^x_{\sigma_x,M_x}$ be as above. Moreover, for a given $\tilde{b} > \frac{|\log(\sigma_t)|}{2}$, assume that $c_1 b N_x^{\frac{1}{3}} \leq \tilde{b} M_t \leq c_2 b N_x^{\frac{1}{3}}$ for some $0 < c_1 \leq c_2 < 1$, where b is defined in (18). Then there exists $\mu_{t,0} \geq 1$ and $\mu_{x,0} > 0$ such that choosing $p_t = \left\lfloor \mu_t M_t^\delta \right\rfloor$ for any $\mu_t > \mu_{t,0}$ and $\mu_x > \mu_{x,0}$, there exists $C > 0$ satisfying*

$$\left\| u - u_{hp} \right\|_{hp} \leq C \exp\left(-c N^{\frac{1}{4+\delta}} \right), \tag{19}$$

where u_{hp} is the solution of (9) in the space $\mathcal{V}_{hp} = \mathcal{S}_0^{\mathbf{p}_x^{\mu_x}}(D, \mathcal{T}^x_{\sigma_x,M_x}) \otimes \mathcal{S}_{\underline{0}}^{p_t}(I, \mathcal{T}^t_{\sigma_t,M_t})$ and $N = \dim\left(\mathcal{V}_{hp}\right)$.

Proof Let $\mu_{t,0}$ and $\mu_{x,0}$ be defined in Proposition 2 and Lemma 9, respectively. From (10), (11) and $c_1 \frac{b}{\tilde{b}} N_x^{\frac{1}{3}} \leq M_t \leq c_2 \frac{b}{\tilde{b}} N_x^{\frac{1}{3}}$, there exists $C > 0$ such that

$$\begin{aligned}
\left\| u - u_{hp} \right\|_{hp} &\leq C \left(M_t^\delta \sigma_t^{-\frac{M_t}{2}} \exp\left(-b N_x^{\frac{1}{3}} \right) + \sigma_t^{\frac{M_t}{2}} \right) \\
&\leq C \left(\exp\left(-b N_x^{\frac{1}{3}} + \tilde{b} M_t \right) + \sigma_t^{\frac{M_t}{2}} \right) \leq C \exp\left(-c N_x^{\frac{1}{3}} \right),
\end{aligned}$$

for some $c > 0$. From $M_t \approx N_x^{\frac{1}{3}}$ and $N_t \approx M_t^{\delta+1}$, we then have $N = N_x N_t \approx N_x^{1+\frac{\delta+1}{3}}$, yielding the desired result.

4 Numerical results

In this section, we provide numerical evidence of Theorem 1. Let us consider $I = D = (0, 1)$, $\mathcal{A} = 1$ and the exact solution

$$u(x,t) := t^{3/4} x^{5/2} (1 - x). \tag{20}$$

In that setting, we have that $u \in H^2(D; \mathscr{G}^{1,3/4}(I)) \cap H_0^1(D; \mathscr{G}^{1,3/4}(I))$ so that we can apply Remark 1 with $\delta = 1$. Hence for every $\sigma_t \in (0, 1)$, there exists $\mu_{t,0} > 0$ such that defining $p_t = \lfloor \mu_t M_t \rfloor$ for any $\mu_t > \mu_{t,0}$ and considering $h_x \approx \sigma_t^{\frac{M_t}{2}}$ it holds

$$\left\| u - u_{hp} \right\|_{hp} \leq C \log(N)^{\frac{\delta(d+1)+1}{d}} N^{-\frac{1}{d}} = C \log(N)^3 N^{-1}.$$

In Figure 1, we present convergence results for different values of σ_t and μ_t. We see that in all the cases, the expected rate can be observed. Furthermore, the results indicate that the error can be decreased by decreasing σ_t. However, we point out that the value of μ_t needs to be increased whenever σ_t is decreased to obtain the desired convergence, yielding a faster increase in the total number of degrees of freedom.

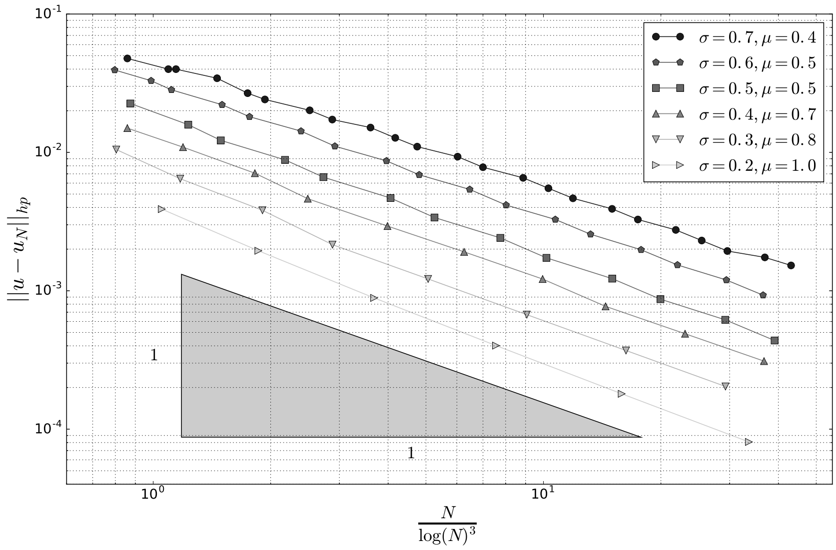

Fig. 1 Convergence of the approximation of (20) in the framework of Theorem 1 for different pairs of parameters (σ_t, μ_t). The expected convergence rate can indeed be observed for all the pairs.

5 Conclusion

In this paper, we developed a coercive space-time variational formulation for linear, parabolic evolution equations on finite time intervals. Coercivity was expressed in terms of mesh-dependent norms, which involved both, size of time-interval (time-step) and polynomial order of the time discretization. The formulation generalizes recent, "h-version" mesh-dependent formulations developed in [20, 12]. Quasi optimality of the Galerkin discretization in this setting is established to hold *for general space-discretizations*. *Exponential convergence rates* of hp time-semidiscretizations was established under Gevrey-type parabolic regularity. For full hp discretizations in both spatial and temporal variables, likewise an exponential convergence rate bound was proved, under the assumption of analytic regularity of the solution expressed in countably normed spaces as are typically encountered in elliptic regularity theory in nonsmooth domains. Numerical experiments for a model problem with a model singular solution confirmed the exponential convergence estimates.

Acknowledgements This work was performed in part while D.Devaud was PhD student at ETH Zürich. Research supported in part by the Swiss National Science Foundation under Grant No. 149819.

References

1. R. Andreev, *Stability of sparse space-time finite element discretizations of linear parabolic evolution equations*, IMA J. Numer. Anal., 33 (2013), pp. 242–260.
2. ———, *Space-time discretization of the heat equation*, Numer. Algorithms, 67 (2014), pp. 713–731.
3. D. Braess, *Finite elements*, Cambridge University Press, Cambridge, third ed., 2007. Theory, fast solvers, and applications in elasticity theory, Translated from the German by Larry L. Schumaker.
4. J.-Y. Chemin, *Perfect incompressible fluids*, vol. 14 of Oxford Lecture Series in Mathematics and its Applications, The Clarendon Press, Oxford University Press, New York, 1998. Translated from the 1995 French original by Isabelle Gallagher and Dragos Iftimie.
5. D. Devaud, *hp-approximation of linear parabolic evolution problems in $H^{1/2}$*, PhD thesis, ETH Zurich, 2017.
6. A. Ern and J.-L. Guermond, *Theory and practice of finite elements*, vol. 159 of Applied Mathematical Sciences, Springer-Verlag, New York, 2004.
7. M. J. Gander, *50 years of time parallel time integration*, in Multiple Shooting and Time Domain Decomposition Methods, Springer, 2015, pp. 69–113.
8. F. D. Gaspoz and P. Morin, *Convergence rates for adaptive finite elements*, IMA J. Numer. Anal., 29 (2009), pp. 917–936.
9. P. Grisvard, *Elliptic problems in nonsmooth domains*, vol. 24 of Monographs and Studies in Mathematics, Pitman (Advanced Publishing Program), Boston, MA, 1985.
10. A. Kunoth and C. Schwab, *Sparse adaptive tensor Galerkin approximations of stochastic PDE-constrained control problems*, SIAM/ASA J. Uncertain. Quantif., 4 (2016), pp. 1034–1059.
11. O. A. Ladyzhenskaya, *The boundary value problems of mathematical physics*, vol. 49 of Applied Mathematical Sciences, Springer-Verlag, New York, 1985. Translated from the Russian by Jack Lohwater [Arthur J. Lohwater].
12. U. Langer, S. E. Moore, and M. Neumüller, *Space-time isogeometric analysis of parabolic evolution problems*, Comput. Methods Appl. Mech. Engrg., 306 (2016), pp. 342–363.
13. S. Larsson and M. Molteni, *Numerical solution of parabolic problems based on a weak space-time formulation*, Comput. Methods Appl. Math., 17 (2017), pp. 65–84.
14. C. Mollet, *Stability of Petrov-Galerkin discretizations: application to the space-time weak formulation for parabolic evolution problems*, Comput. Methods Appl. Math., 14 (2014), pp. 231–255.
15. D. Schötzau and C. Schwab, *An hp a priori error analysis of the DG time-stepping method for initial value problems*, Calcolo, 37 (2000), pp. 207–232.
16. ———, *Time discretization of parabolic problems by the hp-version of the discontinuous Galerkin finite element method*, SIAM J. Numer. Anal., 38 (2000), pp. 837–875.

17. C. Schwab, *p- and hp-finite element methods*, Numerical Mathematics and Scientific Computation, The Clarendon Press, Oxford University Press, New York, 1998. Theory and applications in solid and fluid mechanics.
18. C. Schwab and R. Stevenson, *Space-time adaptive wavelet methods for parabolic evolution problems*, Math. Comp., 78 (2009), pp. 1293–1318.
19. ———, *Fractional space-time variational formulations of (Navier-) Stokes equations*, SIAM J. Math. Anal., 49 (2017), pp. 2442–2467.
20. O. Steinbach, *Space-time finite element methods for parabolic problems*, Comput. Methods Appl. Math., 15 (2015), pp. 551–566.
21. V. Thomée, *Galerkin finite element methods for parabolic problems*, vol. 25 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, second ed., 2006.
22. T. Werder, K. Gerdes, D. Schötzau, and C. Schwab, *hp-discontinuous Galerkin time stepping for parabolic problems*, Comput. Methods Appl. Mech. Engrg., 190 (2001), pp. 6685–6708.