Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Seminar for Applied Mathematics

Space-time hp-approximation of parabolic equations in H 1/2

D. Devaud

Research Report No. 2018-01
January 2018

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

SPACE-TIME HP-APPROXIMATION OF PARABOLIC EQUATIONS IN $H^{1/2}$

DENIS DEVAUD*

Key words. Parabolic partial differential equations, space-time approximation, *hp*-refinements, exponential convergence, a priori error estimates.

AMS subject classifications. 35K10,46B28,65M12,65M15,65M60

Abstract. We analyze a class of variational space-time discretizations for a broad class of initial boundary value problems for linear, parabolic evolution equations.

The space-time variational formulation is based on fractional Sobolev spaces of order 1/2 and the Riemann-Liouville derivative of order 1/2 with respect to the temporal variable. It accomodates general, conforming space discretizations and naturally accomodates discretization of infinite horizon evolution problems. We prove an inf-sup condition for *hp*-time semidiscretizations with an explicit expression of stable testfunctions given in terms of Hilbert transforms of the corresponding trial functions; inf-sup constants are independent of temporal order and the time-step sequences, allowing quasioptimal, high-order discretizations on graded time step sequences, and also *hp*-time discretizations. For solutions exhibiting Gevrey regularity in time and taking values in certain weighted Bochner spaces, we establish novel exponential convergence estimates in terms of N_t, the number of (elliptic) spatial problems to be solved. The space-time variational setting allows general space discretizations, and, in particular, for spatial *hp*-FEM discretizations.

We report numerical tests of the method for model problems in one space dimension with typical singular solutions in the spatial and temporal variable. *hp*-discretizations in both, spatial and temporal variable, are used without any loss of stability, resulting in overall exponential convergence of the space-time discretization.

1. Introduction.

We consider linear, parabolic evolution problems which are set in a pair of separable Hilbert spaces $(V, (\cdot,\cdot)_V)$ and $(H, (\cdot,\cdot)_H)$ such that the embedding $V \subset H$ is continuous, dense and compact and assume without loss of generality that $\|\cdot\|_H \leq \|\cdot\|_V$. Identifying H with its dual H^*, we have the Gelfand triple $V \subseteq H \subseteq V^*$. Given a linear self-adjoint operator $\mathcal{A} \in \mathcal{L}(V, V^*)$ and a time horizon $0 < T \leq \infty$, our goal is to solve the initial-value parabolic problem

$$\mathfrak{B}u := \partial_t u + \mathcal{A}u = f, \tag{1.1a}$$

$$u(0) = 0, \tag{1.1b}$$

in $I = (0, T)$. In (1.1a), $\mathcal{A}$ is meant as a linear, strongly elliptic operator of order $2m$ and V as a closed subspace of the Sobolev space $H^m(D)$ taking into account essential boundary conditions. Here $D \subseteq \mathbb{R}^d$ is a bounded domain for $d = 1, 2, 3$. These equations can be used for instance to model heat conduction in a possibly heterogeneous material or flow of an incompressible Newtonian fluid through the Stokes equations.

In order to approximate the solution of (1.1a)-(1.1b), the usual approach is to consider time stepping schemes [25]. The idea is to first discretize the equations in space so as to obtain a system of ordinary differential equations (ODEs). In a second step, a solver for ODEs is used. On the other hand, first discretizing in space and then in time yields the so-called Rothe's method. The algorithm that we introduce is based on a different approach and belongs to the class of space-time methods. In that setting, the problem is considered over the space-time cylinder $Q = I \times D$ and the solution is sought in a suitable Bochner space. This approach has already been well-studied in different contexts. We mention among others the parareal method which aims at time-parallel integration [10]. Closer to what we present, several methods based on finite element (FEM) approximation have been derived. For instance, in [20, 21, 28], the authors use discontinuous elements for the discretization of the temporal component. In particular, they show that exponential convergence can be obtained in that setting. Another algorithm which was introduced by Schwab and Stevenson in [23] considers adaptive wavelet methods. Based on this, it is possible to recover the optimal convergence rate associated to the underlying elliptic problem for the fully space-time discrete problem. We consider a discretization based on continuous FEM approximation for the time variable. Several algorithms using this idea already exist [1, 2, 14, 23].

The first ingredient to derive a space-time approximation is to introduce a space-time weak formulation associated to (1.1a)-(1.1b). In [23], the authors present a formulation whose solution belongs to $L^2(I; V) \cap H^1(I; V^*)$ and prove that the problem is well-posed. A different approach was presented by Langer, Moore

*Institut de Statistique, Université de Neuchâtel, Avenue de Bellevaux 51, 2000 Neuchâtel.

and Neumüller in [14]. There, it is shown that the bilinear form associated to the problem is coercive and continuous with respect to a mesh-dependent norm, yielding that it admits a unique solution. Our approach is based on the weak formulation used in [15], which was first introduced by Fontes in [9]. Considering (partial) integration by parts for the time derivative, we obtain a formulation containing derivatives of fractional order (see Section 2). The main advantage of such approach is that it is possible to prove inf-sup stability of the bilinear form using a continuous linear operator from the trial to the test space (Proposition 2.5). This important feature is then used to build inf-sup stable pairs of discrete spaces. A time regularity result for the solution of (1.1a)-(1.1b) is also discussed. Considering a smooth forcing term, the associated solution exhibits the same regularity in time except for a potential algebraic singularity at the initial time due to an incompatibility between f and the initial condition. This lack of regularity does not allow us to obtain a high-order method based on for instance p-FEM approximation. Instead, an hp-discretization has to be considered and exponential convergence for the approximation of the temporal component can be proven (Theorem 4.4). Combining this result with both low-order elements (Section 4.2.1) and hp-approximation in the case of the two-dimensional heat equation (Section 4.2.2), it is possible to derive convergence rates for the fully discrete scheme. In particular, we show that space-time hp-approximation allow for exponential convergence (Theorem 4.10).

In the next section, well-posedness of the continuous problem is discussed. Then, continuous piecewise polynomial approximation in the fractional order Sobolev space $H^{1/2}_{00}(I)$ is discussed in Section 3. Space-time discretization of parabolic equations is described in the penultimate section and we conclude with numerical results in Section 5.

Throughout the paper, the set of all positive integers $1, 2, \dots$ is denoted $\mathbb{N}$ and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. For $n \in \mathbb{N}$ and two vectors $a, b \in \mathbb{R}^n$, we write $a \geq b$ $(a > b)$ if $a_i \geq b_i$ $(a_i > b_i)$ for all $i = 1, \dots, n$. Furthermore, for $k \in \mathbb{N}$, $a + k = (a_1 + k, \dots, a_n + k)$. Given two Banach spaces $(A, \|\cdot\|_A)$ and $(B, \|\cdot\|_B)$, $\mathcal{L}(A, B)$ denotes the space of continuous linear functionals from A to B and $\mathcal{B}(A, B)$ the space of bounded bilinear forms from $A \times B$ to $\mathbb{R}$. Finally, $[A, B]_\theta$ refers to the complex interpolation space of order $\theta \in [0, 1]$ between A and B [5, Chapter 4].

2. Linear parabolic evolution problems. The framework associated to the parabolic equations (1.1a)-(1.1b) is presented in this section. This formulation was first introduced by Fontes in [9] and is based on fractional order Bochner-Sobolev spaces [12, Section 2.5]. Definitions of these spaces and several different characterizations are presented in Section 2.2. The main advantage of that setting is that it is possible to show an inf-sup condition using a bounded linear operator between the trial and the test space (Proposition 2.5). This allows us then to build explicitly inf-sup stable pairs of discrete spaces as explained in Section 5. The bilinear form associated to the weak formulation of (1.1a)-(1.1b) is obtained using fractional integration by parts on the time variable. This is discussed in the following section.

2.1. Fractional calculus. The theory presented here is based on [9, 15, 19]. To obtain well-posedness of our problem on $\mathbb{R}_+$, we relate it to the problem stated on $\mathbb{R}$. However, the theory for parabolic equations stated over the real axis is not discussed in this paper and we refer to [7, 8] for more details. For a Hilbert space $(W, (\cdot,\cdot)_W)$, let us define

$$\mathcal{F}(\mathbb{R}; W) := \left\{ g \in \mathcal{C}^\infty(\mathbb{R}; W) \;\middle|\; \|g\|_{H^s(\mathbb{R};W)} < \infty, \ \forall s \in \mathbb{R} \right\}, \tag{2.1a}$$

$$\mathcal{F}(\mathbb{R}_+; W) := \left\{ g \in \mathcal{C}^\infty(\mathbb{R}_+; W) \;\middle|\; \exists \tilde{g} \in \mathcal{F}(\mathbb{R}; W) : g = \tilde{g}|_{\mathbb{R}_+} \right\}, \tag{2.1b}$$

where $H^s(\mathbb{R}; W)$ denote the Bessel potential spaces for $s \in \mathbb{R}$ [12, Definition 5.6.2]. In order to treat initial conditions, we also consider

$$\mathcal{F}_0(\mathbb{R}_+; W) := \{ g \in \mathcal{C}^\infty(\mathbb{R}_+; W) \mid E_0 g \in \mathcal{F}(\mathbb{R}; W) \},$$

where E_0 denotes the extension by zero operator of functions defined on $\mathbb{R}_+$ to $\mathbb{R}$. We also introduce the restriction to $\mathbb{R}_+$ operator $R_>$. Furthermore, the dual spaces of $\mathcal{F}(\mathbb{R}_+; W)$ and $\mathcal{F}_0(\mathbb{R}_+; W)$ are denoted $\mathcal{F}_0'(\mathbb{R}_+; W) := \mathcal{F}(\mathbb{R}_+; W)^*$ and $\mathcal{F}'(\mathbb{R}_+; W) := \mathcal{F}_0(\mathbb{R}_+; W)^*$, respectively, where W has been identified with its dual using Riesz representation theorem. The Riemann-Liouville fractional derivatives of order $0 < \alpha < 1$

are then defined as

$$(D_+^\alpha v)(t) := \frac{1}{\Gamma(1-\alpha)} \frac{d}{dt} \int_0^t (t-s)^{-\alpha} v(s) ds, \qquad \forall v \in \mathcal{F}_0(\mathbb{R}_+; W), \tag{2.2a}$$

$$(D_-^\alpha v)(t) := \frac{1}{\Gamma(1-\alpha)} \frac{d}{dt} \int_t^\infty (s-t)^{-\alpha} v(s) ds, \qquad \forall v \in \mathcal{F}(\mathbb{R}_+; W). \tag{2.2b}$$

Due to their convenient treatment of boundary conditions, Caputo fractional derivatives are more widely used in the literature than the Riemann-Liouville ones. However, we point out that both derivatives coincide in our context [13, (2.4.6) and (2.4.7)]. The following property holds for all $0 < \alpha < 1$, $f \in \mathcal{F}_0(\mathbb{R}_+; W)$ and $g \in \mathcal{F}(\mathbb{R}_+; W)$

$$\int_{\mathbb{R}_+} \left(D_+^\alpha f, g\right)_W dt = \int_{\mathbb{R}_+} \left(f, D_-^\alpha g\right)_W dt.$$

Using this we can define the $\mathcal{F}_0'(\mathbb{R}_+; W)$ and $\mathcal{F}'(\mathbb{R}_+; W)$ distribution derivatives D_+^α and D_-^α as

$$\begin{aligned} \left\langle D_+^\alpha \Phi, f\right\rangle_{\mathcal{F}_0'(\mathbb{R}_+;W)} &:= \left\langle \Phi, D_-^\alpha f\right\rangle_{\mathcal{F}_0'(\mathbb{R}_+;W)}, &\qquad \forall \Phi \in \mathcal{F}_0'(\mathbb{R}_+; W),\ f \in \mathcal{F}(\mathbb{R}_+; W), \\ \left\langle D_-^\alpha \Psi, g\right\rangle_{\mathcal{F}'(\mathbb{R}_+;W)} &:= \left\langle \Psi, D_+^\alpha g\right\rangle_{\mathcal{F}'(\mathbb{R}_+;W)}, &\qquad \forall \Psi \in \mathcal{F}'(\mathbb{R}_+; W),\ g \in \mathcal{F}_0(\mathbb{R}_+; W). \end{aligned}$$

Hence we have that the following operators are continuous [9, (2.24)-(2.28)]

$$\begin{aligned} &D_+^\alpha : \mathcal{F}_0(\mathbb{R}_+; W) \to \mathcal{F}_0(\mathbb{R}_+; W), &\qquad &D_+^\alpha : \mathcal{F}_0'(\mathbb{R}_+; W) \to \mathcal{F}_0'(\mathbb{R}_+; W), \\ &D_-^\alpha : \mathcal{F}(\mathbb{R}_+; W) \to \mathcal{F}(\mathbb{R}_+; W), &\qquad &D_-^\alpha : \mathcal{F}'(\mathbb{R}_+; W) \to \mathcal{F}'(\mathbb{R}_+; W). \end{aligned}$$

Furthermore, we will need the Hilbert transform defined for functions $v \in \mathcal{F}(\mathbb{R}; W)$ as

$$(\mathfrak{H} v)(t) := \frac{1}{\pi} \,\text{p.\,v.} \int_{-\infty}^\infty \frac{v(s)}{t-s} ds := \frac{1}{\pi} \lim_{\varepsilon \to 0} \left(\int_{-\infty}^{t-\varepsilon} \frac{v(s)}{t-s} ds + \int_{t+\varepsilon}^\infty \frac{v(s)}{t-s} ds \right), \tag{2.3}$$

where p. v. denotes the Cauchy principal value. It holds that $\mathfrak{H} \in \mathcal{L}\left(\mathcal{F}(\mathbb{R}; W)\right)$ [7, Section 3.2.2].

2.2. The spaces $H^{1/2}(\mathbb{R}_+; W)$ and $H_{00}^{1/2}(\mathbb{R}_+; W)$. The weak formulation associated to (1.1a)-(1.1b) is based on the Sobolev spaces of fractional order $H^{1/2}$. The key property to link these spaces to the theory developed in the previous section is that they can be defined equivalently using the Riemann-Liouville fractional derivatives of order $\alpha = 1/2$ (Proposition 2.4).

Definition 2.1 (Sobolev-Slobodeckij spaces). *Let $-\infty \le a < b \le \infty$ and $I = (a, b)$.*

1. *$H^{1/2}(I; W)$ is the set of functions $v \in L^2(I; W)$ for which the norm*

$$\|v\|_{H^{1/2}(I;W)}^2 := \|v\|_{L^2(I;W)}^2 + \int_I \int_I \frac{\|v(s) - v(t)\|_W^2}{|s-t|^2} ds dt \tag{2.4}$$

   is finite.

2. *$H_{00}^{1/2}(I; W)$ is the set of functions $v \in L^2(I; W)$ for which the following norm is finite*

$$\|v\|_{H_{00}^{1/2}(I;W)}^2 := \|v\|_{H^{1/2}(I;W)}^2 + \int_I \frac{\|v(t)\|_W^2}{t-a} dt + \int_I \frac{\|v(t)\|_W^2}{b-t} dt, \tag{2.5}$$

   with the convention that the boundary terms do not appear in the case $a = \infty$ and/or $b = \infty$.

Proposition 2.2. *Let $-\infty \le a < b \le \infty$ and $I = (a, b)$.*

1. *A function v belongs to the interpolation space $\widetilde{H}^{1/2}(I; W) := [L^2(I; W), H^1(I; W)]_{1/2}$ if and only if $v \in H^{1/2}(I; W)$ and the norms $\|\cdot\|_{\widetilde{H}^{1/2}(I;W)}$ and $\|\cdot\|_{H^{1/2}(I;W)}$ are equivalent.*
2. *A function v belongs to the interpolation space $\widetilde{H}_{00}^{1/2}(I; W) := [L^2(I; W), H_0^1(I; W)]_{1/2}$ if and only if $v \in H_{00}^{1/2}(I; W)$ and the norms $\|\cdot\|_{\widetilde{H}_{00}^{1/2}(I;W)}$ and $\|\cdot\|_{H_{00}^{1/2}(I;W)}$ are equivalent.*

Proof. If we can prove 1, then 2 follows from [16, Théorème 11.7 p.72]. The case $a = -\infty$ and $b = \infty$ is treated in [12, Theorem 5.6.9 p.452]. In the other cases it is possible to find an operator $E \in \mathcal{L}(\widetilde{H}^{1/2}(I;W), \widetilde{H}^{1/2}(\mathbb{R};W))$ [16, Théorème 8.1 p.42]. The result then follows from [17, Theorem 3.18 p.81]. ∎

An important remark is that the space $H_{00}^{1/2}(I;W)$ is strictly included in the space $H_0^{1/2}(I;W)$, i.e. the closure of $\mathcal{C}_0^\infty(\mathbb{R}_+;W)$ in the $H^{1/2}(I;W)$ norm. Moreover the embeddings $\mathcal{F}_0(\mathbb{R}_+;W) \subset H_{00}^{1/2}(I;W)$ and $\mathcal{F}(\mathbb{R}_+;W) \subset H^{1/2}(I;W)$ are dense [9, Lemma 3.7]. The following important integration by parts property can be found in [15, Lemma 2.7].

LEMMA 2.3. *For $v \in H_{00}^{1/2}(\mathbb{R}_+;W)$, it holds*

$$\langle Dv, w\rangle_{\mathcal{F}_0'(\mathbb{R}_+;W)} = \int_{\mathbb{R}_+} \left(D_+^{1/2} v, D_-^{1/2} w\right)_W dt, \qquad \forall w \in \mathcal{F}(\mathbb{R}_+;W), \tag{2.6}$$

where the $D_+^{1/2}$ derivative has to be understood in the $\mathcal{F}_0'(\mathbb{R}_+;W)$ sense.

We also have the following characterizations of the spaces $H_{00}^{1/2}(\mathbb{R}_+;W)$ and $H^{1/2}(\mathbb{R}_+;W)$ [9, Lemmas 3.8 and 3.9]. Together with the previous lemma, it allows us to derive the weak formulation of our problem and obtain that the associated bilinear form is continuous.

PROPOSITION 2.4. *The following properties hold:*

1. *A function $v \in L^2(\mathbb{R}_+;W)$ belongs to $H^{1/2}(\mathbb{R}_+;W)$ if and only if its $\mathcal{F}'(\mathbb{R}_+;W)$-derivative $D_-^{1/2}v$ belongs to $L^2(\mathbb{R}_+;W)$. Moreover, the norm*

$$\|v\|_{H_-^{1/2}(\mathbb{R}_+;W)}^2 := \|v\|_{L^2(\mathbb{R}_+;W)}^2 + \left\|D_-^{1/2}v\right\|_{L^2(\mathbb{R}_+;W)}^2, \qquad \forall v \in H^{1/2}(\mathbb{R}_+;W),$$

is equivalent to $\|\cdot\|_{H^{1/2}(\mathbb{R}_+;W)}$ defined in (2.4).

2. *A function $v \in L^2(\mathbb{R}_+;W)$ belongs to $H_{00}^{1/2}(\mathbb{R}_+;W)$ if and only if its $\mathcal{F}_0'(\mathbb{R}_+;W)$-derivative $D_+^{1/2}v$ belongs to $L^2(\mathbb{R}_+;W)$. Moreover, the norm*

$$\|v\|_{H_+^{1/2}(\mathbb{R}_+;W)}^2 := \|v\|_{L^2(\mathbb{R}_+;W)}^2 + \left\|D_+^{1/2}v\right\|_{L^2(\mathbb{R}_+;W)}^2, \qquad \forall v \in H_{00}^{1/2}(\mathbb{R}_+;W),$$

is equivalent to $\|\cdot\|_{H_{00}^{1/2}(\mathbb{R}_+;W)}$ defined in (2.5).

2.3. Initial-value problems. Associated to $\mathcal{A}$, we define the bilinear form $\mathfrak{a} \in \mathcal{B}(V,V)$ as $\mathfrak{a}(v,w) := \langle \mathcal{A}v, w\rangle_{V^*}$ for all $v, w \in V$ and assume that there exist $0 < \lambda_- \le \lambda_+ < \infty$ such that

$$\mathfrak{a}(v,w) \le \lambda_+ \|v\|_V \|w\|_V, \qquad \forall v, w \in V, \tag{2.7a}$$

$$\mathfrak{a}(v,v) \ge \lambda_- \|v\|_V^2, \qquad \forall v \in V, \tag{2.7b}$$

i.e. $\mathfrak{a}$ is continuous and coercive on V, respectively. We see that (2.7a) follows from the assumption that $\mathcal{A} \in \mathcal{L}(V, V^*)$. Note that here λ_- is chosen as the largest possible constant such that (2.7b) holds. We introduce the normed vector spaces

$$\mathcal{X} := H_{00}^{1/2}(\mathbb{R}_+;H) \cap L^2(\mathbb{R}_+;V), \qquad \|v\|_{\mathcal{X}}^2 := \|v\|_{H_{00}^{1/2}(\mathbb{R}_+;H)}^2 + \|v\|_{L^2(\mathbb{R}_+;V)}^2, \qquad \forall v \in \mathcal{X},$$

$$\mathcal{Y} := H^{1/2}(\mathbb{R}_+;H) \cap L^2(\mathbb{R}_+;V), \qquad \|v\|_{\mathcal{Y}}^2 := \|v\|_{H^{1/2}(\mathbb{R}_+;H)}^2 + \|v\|_{L^2(\mathbb{R}_+;V)}^2, \qquad \forall v \in \mathcal{Y}.$$

Note that $\mathcal{X}$ and $\mathcal{Y}$ are isomorphic to $(H_{00}^{1/2}(\mathbb{R}_+) \otimes H) \cap (L^2(\mathbb{R}_+) \otimes V)$ and $(H^{1/2}(\mathbb{R}_+) \otimes H) \cap (L^2(\mathbb{R}_+) \otimes V)$, respectively [15, (3.1)]. This will be used later in the approximation theory to split the error estimation between the spatial and the temporal components. The weak formulation of (1.1a)-(1.1b) is then: given $f \in \mathcal{Y}^*$, find $u \in \mathcal{X}$ such that

$$\mathfrak{b}(u,v) := \int_{\mathbb{R}_+} \left[\left(D_+^{1/2} u, D_-^{1/2} v\right)_H + \mathfrak{a}(u,v)\right] dt = \langle f, v\rangle_{\mathcal{Y}^*}, \qquad \forall v \in \mathcal{Y}. \tag{2.8}$$

This is justified by Lemma 2.3. Due to Proposition 2.4, it is clear that $\mathfrak{b}$ is a continuous bilinear form. Let us define for $\eta \in \mathbb{R}$ the operator

$$\mathcal{H}^{\eta} := R_{>}\left(\cos(\pi\eta)\mathrm{Id} + \sin(\pi\eta)\mathfrak{H}\right) E_0,$$

where $\mathfrak{H}$ is defined in (2.3). Then $\mathcal{H}^{\eta} \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ [7, Section 3.5] and we obtain the following important result. The well-posedness of (2.8) has been shown in [9, Section 4.1], while the inf-sup property can be found in [15, Section 2.9]. In [7] both results have been obtain using a different approach based on splitting the space $H^{1/2}(\mathbb{R}; H)$ into symmetric and antisymmetric functions. In particular, in [15] the inf-sup condition (2.9) has been obtained only for small values of η. The general result can however be found in [7, Theorem 3.5.2].

Proposition 2.5. *Assume that (2.7a)-(2.7b) hold. Then for all $f \in \mathcal{Y}^*$, (2.8) admits a unique solution $u \in \mathcal{X}$. Furthermore for all $\eta \in \left(0, \frac{1}{2}\right)$, there exists $\beta = \beta(\eta) >$ such that*

$$\mathfrak{b}(v, \mathcal{H}^{-\eta} v) \geq \beta(\eta) \left\|v\right\|_{\mathcal{X}} \left\|\mathcal{H}^{-\eta} v\right\|_{\mathcal{Y}}, \qquad \forall v \in \mathcal{X}. \tag{2.9}$$

We point out that (2.9) is a stronger condition than the usual inf-sup inequality. It provides us with an explicit way of defining inf-sup stable pairs of discrete spaces. Indeed, considering a discrete trial space, an associated test space can be chosen as the image of the trial space by $\mathcal{H}^{-\eta}$ for some $\eta \in \left(0, \frac{1}{2}\right)$. This is explained in Section 4. It is also used in [7, Theorem 3.5.2] to prove that (2.8) is well-posed.

2.4. Regularity results. We present here some new regularity results for solutions of linear parabolic equations. Since it is not the main focus of this paper, we refer to [7, Section 3.6] for the associated proofs. The usual regularity theory for parabolic equations is based on the assumption that the right-hand side is analytic with respect to its time variable [20]. In that case, it can be shown that the solution inherits this regularity except for a potential algebraic singularity at the initial time due to an incompatibility between f and the initial condition. We extend this theory to Gevrey classes. Such regularity allows for instance to localize right-hand side functions since they can have compact supports, which is not the case of analytic functions. The technique used to prove this result is the same as in [20], namely it is based on the spectral decomposition of the operator $\mathcal{A}$.

Definition 2.6 (Gevrey class). *Let $I = (a, b) \subset \mathbb{R}$ be an interval for some $-\infty < a < b \leq \infty$. For $\delta \geq 1$ and $\theta \geq 0$, we say that v is of* Gevrey type (δ, θ) *and we write $v \in \mathscr{G}^{\delta,\theta}(I; H)$ if $v \in \mathcal{C}(\overline{I}; H)$ and there exist $C_v, d_v > 0$ such that*

$$\left\|v^{(l)}(t)\right\|_H \leq C_v d_v^l \Gamma(l+1)^{\delta} (t-a)^{\theta - l}, \qquad \forall l \in \mathbb{N},\ t \in I, \tag{2.10}$$

where $\Gamma(s)$ denotes the Gamma function for $s > 0$. Moreover we say that v is of Gevrey type δ *and write $u \in \mathscr{G}^{\delta}(I; H)$ if $v \in \mathcal{C}(\overline{I}; H)$ and there exist $C_v, d_v > 0$ such that*

$$\left\|v^{(l)}(t)\right\|_H \leq C_v d_v^l \Gamma(l+1)^{\delta}, \qquad \forall l \in \mathbb{N}_0,\ t \in \overline{I}. \tag{2.11}$$

If I is bounded, then $\mathscr{G}^{\delta,\bar{\theta}}(I; H) \subset \mathscr{G}^{\delta,\theta}(I; H)$ for all $0 \leq \theta \leq \bar{\theta} < \infty$. We point out that the bound (2.10) is assumed to hold only for $l \in \mathbb{N}$. Hence we control the derivatives of the function but do not enforce that $v(a) = 0$. This point is crucial since it allows us in principle to treat solutions of parabolic equations with non-homogeneous initial value. In order to allow $v(a) \neq 0$, condition (2.10) could be equivalently replaced by

$$\left\|v^{(l)}(t)\right\|_H \leq C_v d_v^l \Gamma(l+1)^{\delta} (t-a)^{\min(0, \theta - l)}, \qquad \forall l \in \mathbb{N}_0,\ t \in I.$$

Proposition 2.7. *Let $f \in L^2(\mathbb{R}_+; H)$ satisfy $\operatorname{ess\,supp} f \subset \overline{I} = [0, T]$ and $f \in \mathscr{G}^{\delta}(I; H)$ for some $0 < T < \infty$ and $\delta \geq 1$. Then for any $T < T_\infty \leq \infty$ we have for $I_\infty = (T, T_\infty)$*

$$u|_I \in \mathscr{G}^{\delta,1}(I; H) \cap \mathscr{G}^{\delta,1/2}(I; V), \qquad u|_{I_\infty} \in \mathscr{G}^{1,1/2}(I_\infty; H) \cap \mathscr{G}^{1,0}(I_\infty; V),$$

where $u \in \mathcal{X}$ is the solution of (2.8) associated to f.

Remark 2.8. *The previous proposition implies that for every T_∞, there exist $C_u, d_u > 0$ such that for all $l \in \mathbb{N}$ it holds*

$$\left\|u^{(l)}(t)\right\|_H \leq C_u d_u^l \Gamma(l+1)(t-T)^{1/2-l}, \qquad \left\|u^{(l)}(t)\right\|_V \leq C_u d_u^l \Gamma(l+1)(t-T)^{-l}, \qquad \forall t \in I_\infty.$$

It is important to point out that in this case neither C_u nor d_u depend on T_∞.

3. hp-approximation in $H^{1/2}(I;H)$. As we will see in Section 4, equation (2.9) allows us to build explicitly inf-sup stable pairs of discrete spaces. Using this approach yields quasi-optimality of the discrete solution of (2.8). Hence to obtain convergence rates for the discretization, it is sufficient to investigate approximation properties of the space $\mathcal{X}$. In particular, we define in the following an interpolation operator and derive associated error estimates in the $H_{00}^{1/2}(I;H)$-norm. In this section, $(H, (\cdot,\cdot)_H)$ denotes a generic Hilbert space.

3.1. Piecewise polynomial approximation of functions in $H^{1/2}(I;H)$. For a bounded interval $I = (a,b)$ and $M \in \mathbb{N}$, let $\mathcal{T} = \{I_n\}_{n=1}^M$ be a partition of I, where $I_n = (t_{n-1}, t_n)$ and $a = t_0 < t_1 < \cdots < t_M = b$. Furthermore we denote $h_n = t_n - t_{n-1}$ for $n = 1, \ldots, M$. Associated to $\mathcal{T}$, we consider a vector $\mathbf{p} = \{p_n\}_{n=1}^M \in \mathbb{N}^M$ of polynomial degrees and define for $l \in \mathbb{N}_0$

$$S^{\mathbf{p},l}(I,\mathcal{T};H) := \left\{ v \in H^l(I;H) \;\middle|\; v|_{I_n} \in \mathcal{P}^{p_n}(I_n;H),\ n = 1,\ldots,M \right\},$$

where $\mathcal{P}^{p_n}(I_n;H)$ denotes the space of polynomials of degree at most p_n over I_n taking values in H. In order to analyze piecewise polynomial approximation of functions in $H^{1/2}(I;H)$, we define for $k_n \in \mathbb{N}_0$ and an interval I_n

$$|v|^2_{H_J^{k_n}(I_n;H)} := \int_{I_n} \left\|v^{(k_n)}(t)\right\|_H^2 (t-t_{n-1})^{k_n-1/2}(t_n-t)^{k_n-1/2} dt.$$

This quantity is a seminorm for a so-called Jacobi-weighted Sobolev space, whence the subscript J. When taken over the reference interval $\hat{I} := (-1,1)$, the weight in the integral is the one associated with the orthogonal Jacobi polynomials for a specific choice of parameters. This idea is strongly used in the error analysis and was originally introduced by Guo and Babuška [3, 4] and Guo and Heuer [11]. We then introduce an interpolation operator over $\hat{I}$ and using an affine mapping between every I_n and $\hat{I}$, it yields an operator $\pi_{I_n}^{p_n}$ such that $\left(\pi_{I_n}^{p_n} v\right)(t_{n-1}) = v(t_{n-1})$ and $\left(\pi_{I_n}^{p_n} v\right)(t_n) = v(t_n)$ for every $v \in \mathcal{C}(\overline{I_n};H)$ and $n = 1,\ldots,M$. A global interpolation operator is then defined for $v \in \mathcal{C}(\overline{I};H)$ as $\pi_{\mathcal{T}}^{\mathbf{p}} v|_{I_n} = \pi_{I_n}^{p_n} v$. The proofs of the following results can be found in [7, Chapter 4].

Proposition 3.1. *Let I, $\mathcal{T}$ and $\mathbf{p} \in \mathbb{N}^M$ be as above. Let further $v \in \mathcal{C}\left(\overline{I};H\right)$ and assume that for every $n \in \{1,\ldots,M\}$ there exists $k_n \in \mathbb{N}$ such that $p_n + 1 \geq k_n$ and $|v|_{H_J^{k_n}(I_n;H)} < \infty$. Then it holds*

$$\|v - \pi_{\mathcal{T}}^{\mathbf{p}} v\|^2_{H_{00}^{1/2}(I;H)} \leq C \sum_{n=1}^M \frac{\Gamma(p_n-k_n+2)}{\Gamma(p_n+k_n+1)} |v|^2_{H_J^{k_n}(I_n;H)}, \tag{3.1a}$$

$$\|v - \pi_{\mathcal{T}}^{\mathbf{p}} v\|^2_{L^2(I;H)} \leq C \sum_{n=1}^M \left(\frac{h_n}{2}\right) \frac{\Gamma(p_n-k_n+2)\Gamma(p_n+1)}{\Gamma(p_n+k_n+1)\Gamma(p_n+2)} |v|^2_{H_J^{k_n}(I_n;H)}, \tag{3.1b}$$

for a constant $C > 0$. Moreover $v(t_n) = \left(\pi_{\mathcal{T}}^{\mathbf{p}} v\right)(t_n)$ for all $n = 0,\ldots,M$.

Corollary 3.2. *Let I, $\mathcal{T}$ and $\mathbf{p} \in \mathbb{N}^M$ be as above. Furthermore let $v \in \mathcal{C}\left(\overline{I};H\right)$ and assume that for every $n \in \{1,\ldots,M\}$ there exists $k_n \in \mathbb{N}$ such that $p_n + 1 \geq k_n$ and $v|_{I_n} \in H^{k_n}(I_n;H)$. Then it holds*

$$\|v - \pi_{\mathcal{T}}^{\mathbf{p}} v\|^2_{H_{00}^{1/2}(I;H)} \leq C \sum_{n=1}^M \left(\frac{h_n}{2}\right)^{2k_n-1} \frac{\Gamma(p_n-k_n+2)}{\Gamma(p_n+k_n+1)} |v|^2_{H^{k_n}(I_n;H)}, \tag{3.2a}$$

$$\|v - \pi_{\mathcal{T}}^{\mathbf{p}} v\|^2_{L^2(I;H)} \leq C \sum_{n=1}^M \left(\frac{h_n}{2}\right)^{2k_n} \frac{\Gamma(p_n-k_n+2)\Gamma(p_n+1)}{\Gamma(p_n+k_n+1)\Gamma(p_n+2)} |v|^2_{H^{k_n}(I_n;H)}, \tag{3.2b}$$

for a constant $C > 0$. Moreover $v(t_n) = (\pi_{\mathcal{T}}^{\mathbf{p}} v)(t_n)$ for all $n = 0, \dots, M$.

Proof. It follows from the previous proposition and $\max_{t\in I_n}(t-t_{n-1})^{k_n-1/2}(t_n-t)^{k_n-1/2} = \left(\frac{h_n}{2}\right)^{2k_n-1}$. ∎

REMARK 3.3. *Using a similar argument we obtain*

$$\|u-\pi_{\mathcal{T}}^{\mathbf{p}}u\|^2_{H^{1/2}_{00}(I;H)} \le C\left(|u|^2_{H^1_J(I_1;H)} + \sum_{n=2}^{M}\left(\frac{h_n}{2}\right)^{2k_n+1}\frac{\Gamma(p_n-k_n+1)}{\Gamma(p_n+k_n+2)}|u|^2_{H^{k_n+1}(I_n;H)}\right),$$

$$\|u-\pi_{\mathcal{T}}^{\mathbf{p}}u\|^2_{L^2(I;H)} \le C\left(\frac{h_1}{2}|u|^2_{H^1_J(I_1;H)} + \sum_{n=2}^{M}\left(\frac{h_n}{2}\right)^{2k_n+2}\frac{\Gamma(p_n-k_n+1)\Gamma(p_n+1)}{\Gamma(p_n+k_n+2)\Gamma(p_n+2)}|u|^2_{H^{k_n+1}(I_n;H)}\right),$$

with $p_n \ge \max(k_n, 1) \in \mathbb{N}_0$ for $n = 1, \dots, M$. This estimate is particularly useful if u contains a singularity at the endpoint a so that $u|_{I_1} \notin H^1(I_1; H)$.

REMARK 3.4. *Using the bounds [7, (A.1.4)] on the Gamma function, if $v \in H^k(I;H)$ for some $k \in \mathbb{N}$ and choosing $h_n = h$ and $p_n = p \ge k-1$ for all $n \in \{1,\dots,M\}$, we obtain that there exists $C > 0$ such that*

$$\|v-\pi_{\mathcal{T}}^{\mathbf{p}}v\|_{H^{1/2}_{00}(I;H)} \le C\left(\frac{he}{2p}\right)^{k-1/2}|v|_{H^k(I;H)},$$

$$\|v-\pi_{\mathcal{T}}^{\mathbf{p}}v\|_{L^2(I;H)} \le C\left(\frac{he}{2p}\right)^{k}|v|_{H^k(I;H)}.$$

This allows us to derive convergence rates for h- and p-refinements in time; see Remark 4.5.

Using Remark 3.3, it is possible to derive high-order convergence for the approximation of the solution of (2.8). As discussed in Section 2.4, assuming that the right-hand side is smooth in time, then the associated solution will also be smooth except for a potential algebraic singularity at the initial time due to an incompatibility between f and the initial condition. The weighted seminorm allows us then to treat this singularity.

3.2. Exponential convergence of Gevrey functions. As Proposition 2.7 suggests, solutions of parabolic equations typically contain an algebraic singularity at the initial time. To circumvent the lack of regularity of functions in $\mathscr{G}^{\delta,\theta}(I;H)$, a graded mesh towards the singularity is considered. More specifically, for $\sigma \in (0,1)$ and $M \in \mathbb{N}$, we define

$$t_{\sigma,0} = a, \qquad t_{\sigma,m} = a + (b-a)\sigma^{M-m}, \qquad m = 1,\dots,M, \tag{3.3}$$

and $\mathcal{T}_\sigma := \{I_{\sigma,m}\}_{m=1}^M$, where $I_{\sigma,m} := (t_{\sigma,m-1}, t_{\sigma,m})$ for $m = 1,\dots,M$. It yields that for $\kappa = \sigma^{-1} - 1$

$$h_{\sigma,m} = t_{\sigma,m} - t_{\sigma,m-1} = \kappa(b-a)\sigma^{M-m+1}. \tag{3.4}$$

Furthermore, given $\mu > 0$ and $\delta \ge 1$, we consider the vector of polynomial degrees $\mathbf{p}_{\mu,\delta} := \{p_{\mu,\delta,m}\}_{m=1}^M$ defined through

$$p_{\mu,\delta,1} = 1, \qquad p_{\mu,\delta,m} = \max\left(1, \lfloor \mu m^\delta \rfloor\right), \qquad m = 2,\dots,M. \tag{3.5}$$

We then denote $N = \dim(S^{\mathbf{p}_{\mu,\delta}}(I,\mathcal{T}_\sigma))$. Since $\mathscr{G}^{\delta,\theta}(I;H) \subset \mathcal{C}\left(\overline{I_{\sigma,1}};H\right)$, the interpolation operator $\pi_{\mathcal{T}_\sigma}^{\mathbf{p}_{\mu,\delta}}$ is well-defined for functions in this space. In order to simplify the notations, we write $\pi_{\mu,\sigma,\delta} = \pi_{\mathcal{T}_\sigma}^{\mathbf{p}_{\mu,\delta}}$. We will need the following lemma in order to prove exponential convergence of the hp-approximation of Gevrey functions.

LEMMA 3.5. *Let $\alpha > 0$. Then there exists $\mu_0 \ge 1$ such that for any $M \in \mathbb{N}$ and defining $p_n = \lfloor \mu n^\delta \rfloor$ for some $\mu > \mu_0$ and $n = 1,\dots,M$, it holds*

$$\sum_{n=1}^{M}\alpha^{2n}\frac{\Gamma(p_n-n+1)}{\Gamma(p_n+n+2)}\Gamma(n+2)^{2\delta} \le C,$$

for a constant $C = C(\alpha,\mu,\delta) > 0$ independent of M.

Proof. From [7, (A.1.3),(A.1.4)], it holds for $\mu \geq 1$ and $n \in \mathbb{N}$

$$\Gamma(n+2) < \sqrt{2\pi}e3^{3/2}n^{n+3/2}e^{-n},$$
$$\frac{\Gamma(p_n - n + 1)}{\Gamma(p_n + n + 2)} \leq \frac{\Gamma(\mu n^\delta - n + 1)}{\Gamma(\mu n^\delta + n + 1)} \leq e\left(\frac{\mu n^\delta}{e}\right)^{-2n}.$$

It follows that there exists $C > 0$ such that for all $n = 1, \ldots, M$

$$\alpha^{2n}\frac{\Gamma(p_n - n + 1)}{\Gamma(p_n + n + 2)}\Gamma(n+2)^{2\delta} \leq Cn^{3\delta}\left(\frac{\alpha}{\mu e^{\delta-1}}\right)^{2n}.$$

Defining

$$\mu_0 = \max\left(1, \frac{\alpha}{e^{\delta-1}}\right), \tag{3.6}$$

it follows that for any $\mu > \mu_0$, we have

$$\frac{\alpha}{\mu e^{\delta-1}} =: \zeta \in (0,1).$$

Hence we get

$$\sum_{n=1}^{M}\alpha^{2n}\frac{\Gamma(p_n - n + 1)}{\Gamma(p_n + n + 2)}\Gamma(n+2)^{2\delta} \leq C\sum_{n=1}^{M}n^{3\delta}\zeta^n \leq C\sum_{n=1}^{\infty}n^{3\delta}\zeta^n,$$

and it can be shown using for instance a ratio test that the last series converges. ∎

PROPOSITION 3.6. *Let $u \in \mathscr{G}^{\delta,\theta}(I;H)$ for some $\theta > 1/4$, $\delta \geq 1$ and assume that the constants C_u and d_u in (2.10) are independent of $|I|$. For $\sigma \in (0,1)$ and $M \in \mathbb{N}$, define $\mathcal{T}_\sigma$ through (3.3). Then there exists $\mu_0 \geq 1$ such that defining $\mathbf{p}_{\mu,\delta}$ through (3.5) for any $\mu > \mu_0$, it holds*

$$\|u - \pi_{\mu,\sigma,\delta}u\|_{H^{1/2}_{00}(I;H)} \leq C\,|I|^\theta \exp\left(-\beta N^{\frac{1}{\delta+1}}\right), \tag{3.7}$$

for some constants $C, \beta > 0$ independent of $|I|$.

Proof. On $I_{\sigma,1}$, we have the following bound [7, Lemma 4.3.1]

$$|u|^2_{H^1_J(I_{\sigma,1};H)} \leq C(b-a)^{2\theta}\sigma^{2\theta(M-1)},$$

for a constant $C > 0$. Furthermore, from [7, Lemma 4.3.3], we have that there exist $C, d > 0$ such that for all $l \in \mathbb{N}$ and $m \in \{2, \ldots, M\}$ it holds

$$|u|^2_{H^l(I_{\sigma,m};H)} \leq C(b-a)^{2\theta+1}\left(\frac{d}{b-a}\right)^{2l}\Gamma(l+1)^{2\delta}\sigma^{(M-m+1)(2(\theta-l)+1)}.$$

From (3.4), we then have

$$\left(\frac{h_{\sigma,n}}{2}\right)^{2k_n+1}|u|^2_{H^{k_n+1}(I_{\sigma,n};H)} \leq C(\theta,\sigma)d(b-a)^{2\theta}\left(\frac{\kappa d}{2}\right)^{2k_n+1}\Gamma(k_n+2)^{2\delta}\sigma^{2\theta(M-n+1)},$$

so that Remark 3.3 yields

$$\|u - \pi_{\mu,\sigma,\delta}u\|^2_{H^{1/2}_{00}(I;H)} \leq C(b-a)^{2\theta}\sigma^{2\theta M}\left(1 + \sum_{n=2}^{M}\left(\frac{\kappa d}{2}\right)^{2k_n}\frac{\Gamma(p_n - k_n + 1)}{\Gamma(p_n + k_n + 2)}\Gamma(k_n+2)^{2\delta}\sigma^{-2\theta n}\right).$$

We then choose $k_n = n$ and apply Lemma 3.5 with $\alpha = \frac{\kappa d}{2\sigma^\theta}$. Hence there exists a constant $C(\theta,\sigma,\delta,\mu) > 0$ such that

$$\|u - \pi_{\mu,\sigma,\delta} u\|^2_{H^{1/2}_{00}(I;H)} \leq C(b-a)^{2\theta}\sigma^{2\theta M} = C(b-a)^{2\theta}\exp\left(-2\theta\left|\log(\sigma)\right| M\right).$$

The result then follows with

$$\beta = \frac{\theta\left|\log(\sigma)\right|}{\mu^{\frac{1}{\delta+1}}}, \tag{3.8}$$

since

$$N = 1 + \sum_{n=2}^{M}\left\lfloor \mu n^\delta \right\rfloor \leq \mu\left(\sum_{n=1}^{M} n^\delta\right) \leq \mu M^{\delta+1}.$$

□

Proposition 3.7. *Let $u \in \mathscr{G}^{\delta,\theta}(I;H)$ for some $\theta \geq 0$, $\delta \geq 1$ and assume that the constants C_u and d_u in (2.10) are independent of $|I|$. For $\sigma \in (0,1)$ and $M \in \mathbb{N}$, define $\mathcal{T}_\sigma$ through (3.3). Then there exists $\mu_0 \geq 1$ such that defining $\mathbf{p}_{\mu,\delta}$ through (3.5) for any $\mu > \mu_0$, it holds*

$$\|u - \pi_{\mu,\sigma,\delta} u\|_{L^2(I;H)} \leq C\left|I\right|^{1/2}\max\left(\left|I\right|^\theta, 1\right)\exp\left(-\beta N^{\frac{1}{\delta+1}}\right), \tag{3.9}$$

for some constants $C, \beta > 0$.

Proof. The argument is the same as for the proof of Proposition 3.6 except that the estimate of $|u|_{H^1_J(I_{\sigma,1};H)}$ is replaced by [7, Lemma 4.3.2]

$$\|u - \pi_{\mu,\sigma,\delta} u\|^2_{L^2(I_{\sigma,1};H)} \leq C(b-a)\sigma^{M-1}.$$

□

Propositions 3.6 and 3.7 are very similar and since $\|v\|_{L^2(I;H)} \leq \|v\|_{H^{1/2}_{00}(I;H)}$ for all $v \in H^{1/2}_{00}(I;H)$, the second is a consequence of the first one. However, we point out that Proposition 3.6 does not allow for values $\theta \in [0,1/4]$ while Proposition 3.7 does. Furthermore the reason for making the constant explicit in $|I|$ in (3.7) and (3.9) is that we will use these results for the approximation of the solution of (2.8) which is defined on the unbounded domain $\mathbb{R}_+$.

4. Space-time Petrov-Galerkin approximation. In this section, we introduce a space-time discretization for the solution of (2.8) and derive associated convergence rates. Let $\left(\Theta^l_{\mathcal{X}}\right)_{l\in\mathbb{N}} \subset H^{1/2}_{00}(\mathbb{R}_+)$ and $\left(\Sigma^m_V\right)_{m\in\mathbb{N}} \subset V$ be two sequences of finite dimensional dense spaces. Based on this, we define for $l \in \mathbb{N}$ and $\eta \in \left(0, \frac{1}{2}\right)$

$$\mathcal{X}^l := \Theta^l_{\mathcal{X}} \otimes \Sigma^l_V, \qquad \mathcal{Y}^l := \Theta^l_{\mathcal{Y}} \otimes \Sigma^l_V,$$

where $\Theta^l_{\mathcal{Y}} := \mathcal{H}^{-\eta}\left(\Theta^l_{\mathcal{X}}\right) = \left\{\mathcal{H}^{-\eta}\left(v^l\right) \;\middle|\; v^l \in \Theta^l_{\mathcal{X}}\right\}$. Then the sequence $\left(\mathcal{X}^l\right)_{l\in\mathbb{N}}$ is dense in $\mathcal{X}$ [27, Theorem 3.12]. Furthermore, we have the following result which is a direct consequence of Proposition 2.5. We point out that the choice of η is arbitrary and only influences the inf-sup constant associated to the bilinear form $\mathfrak{b}$ [7, Remark 3.4.2].

Proposition 4.1. *Assume that (2.7a)-(2.7b) hold and let $\mathcal{X}^l$ be as above, $\eta \in \left(0, \frac{1}{2}\right)$ and $\Theta^l_{\mathcal{Y}} := \mathcal{H}^{-\eta}\left(\Theta^l_{\mathcal{X}}\right)$. Then for any $f \in \mathcal{Y}^*$ and $l \in \mathbb{N}$, there exists a unique $u^l \in \mathcal{X}^l$ such that*

$$\mathfrak{b}(u^l, v^l) = \langle f, v^l\rangle_{\mathcal{Y}^*}, \qquad \forall v^l \in \mathcal{Y}^l. \tag{4.1}$$

Moreover, there exists $C = C(\eta) > 0$ such that

$$\left\|u - u^l\right\|_{\mathcal{X}} \leq C \inf_{w^l\in\mathcal{X}^l}\left\|u - w^l\right\|_{\mathcal{X}}, \tag{4.2}$$

where $u \in \mathcal{X}$ is the unique solution of (2.8) associated to f.

Proof. We already know that the bilinear form $\mathfrak{b} : \mathcal{X}^l \times \mathcal{Y}^l \to \mathbb{R}$ is continuous. Furthermore, (2.9) and the construction of $\mathcal{Y}^l$ yield that

$$\inf_{0\neq v^l\in\mathcal{X}^l} \sup_{0\neq w^l\in\mathcal{Y}^l} \frac{\mathfrak{b}\left(v^l, w^l\right)}{\left\|v^l\right\|_{\mathcal{X}} \left\|w^l\right\|_{\mathcal{Y}}} \geq \inf_{0\neq v^l\in\mathcal{X}^l} \frac{\mathfrak{b}\left(v^l, \mathcal{H}^{-\eta}\left(v^l\right)\right)}{\left\|v^l\right\|_{\mathcal{X}} \left\|\mathcal{H}^{-\eta}\left(v^l\right)\right\|_{\mathcal{Y}}} \geq \beta(\eta).$$

This condition implies that the operator associated to $\mathfrak{b}(\cdot,\cdot)$ defined on the discrete spaces is injective. Since $\dim\left(\mathcal{X}^l\right) = \dim\left(\mathcal{Y}^l\right) < \infty$, we obtain that it is surjective which is equivalent to say that it is non-degenerate. Hence it follows from [6, Theorem 3.6] that (4.1) admits a unique solution for any $f \in \left(\mathcal{Y}^l\right)^*$ and in particular for any $f \in \mathcal{Y}^*$.

It remains to show (4.2). Note that we have the Galerkin orthogonality $\mathfrak{b}\left(u - u^l, v^l\right) = 0$ for all $v^l \in \mathcal{Y}^l$ since $\mathcal{Y}^l \subset \mathcal{Y}$. It follows from the inf-sup condition that for any $w^l \in \mathcal{X}^l$, there exists $v^l \in \mathcal{Y}^l$ such that

$$\beta(\eta)\left\|u^l - w^l\right\|_{\mathcal{X}} \left\|v^l\right\|_{\mathcal{Y}} \leq \mathfrak{b}\left(u^l - w^l, v^l\right) = \mathfrak{b}\left(u - w^l, v^l\right) \leq \|\mathfrak{b}\|_{\mathcal{B}(\mathcal{X},\mathcal{Y})} \left\|u - w^l\right\|_{\mathcal{X}} \left\|v^l\right\|_{\mathcal{Y}}.$$

Hence, for all $w^l \in \mathcal{X}^l$ it holds

$$\left\|u - u^l\right\|_{\mathcal{X}} \leq \left\|u - w^l\right\|_{\mathcal{X}} + \left\|u^l - w^l\right\|_{\mathcal{X}} \leq \left(1 + \frac{\|\mathfrak{b}\|_{\mathcal{B}(\mathcal{X},\mathcal{Y})}}{\beta(\eta)}\right)\left\|u - w^l\right\|_{\mathcal{X}}.$$

∎

Using the tensor structure of $\mathcal{X}$, if we can define two operators $\Pi_{\Theta^l_{\mathcal{X}}} \in \mathcal{L}(L^2(\mathbb{R}_+), \Theta^l_{\mathcal{X}})$ and $\Pi_{\Sigma^l_V} \in \mathcal{L}(H, \Sigma^l_V)$, then we can estimate the error bound of the full operator $\Pi^l_{\mathcal{X}} := \Pi_{\Theta^l_{\mathcal{X}}} \otimes \Pi_{\Sigma^l_V} : \mathcal{X} \to \mathcal{X}^l$ from the relation

$$(\mathrm{Id}\otimes\mathrm{Id}) - (\Pi_{\Theta^l_{\mathcal{X}}} \otimes \Pi_{\Sigma^l_V}) = \left(\left(\mathrm{Id} - \Pi_{\Theta^l_{\mathcal{X}}}\right)\otimes \Pi_{\Sigma^l_V}\right) + \left(\mathrm{Id}\otimes\left(\mathrm{Id} - \Pi_{\Sigma^l_V}\right)\right).$$

Indeed, using a triangular inequality and the continuity of $\Pi_{\Theta^l_{\mathcal{X}}}$ and $\Pi_{\Sigma^l_V}$, we obtain

$$\left\|u - \Pi^l_{\mathcal{X}} u\right\|_{\mathcal{X}} \leq C\left(\left\|u - \left(\Pi_{\Theta^l_{\mathcal{X}}}\otimes\mathrm{Id}\right)u\right\|_{\mathcal{X}} + \left\|u - \left(\mathrm{Id}\otimes\Pi_{\Sigma^l_V}\right)u\right\|_{\mathcal{X}}\right). \tag{4.3}$$

It follows that in order to obtain convergence rates for the discrete solution $u^l \in \mathcal{X}^l$, we only need to derive approximation properties for the spaces V and $H^{1/2}_{00}(\mathbb{R}_+)$.

REMARK 4.2. *We see that the construction of a basis of $\Theta^l_{\mathcal{Y}}$ necessitates to apply the operator $\mathcal{H}^{-\eta}$ to elements in $\Theta^l_{\mathcal{X}}$. In principle, this needs to be done using numerical integration, which can lead to approximation errors. However, in [7, Chapter 6], explicit formulas to build the matrices associated to $\Theta^l_{\mathcal{X}}$ and $\Theta^l_{\mathcal{Y}}$ in the case of continuous piecewise polynomial basis functions are presented.*

4.1. *hp*-approximation for the time variable. In this section, we discuss approximation for the time variable in the $H^{1/2}_{00}(\mathbb{R}_+)$-norm. This is essentially the same theory as the one developed in Section 3 except that we approximate functions on the unbounded domain $\mathbb{R}_+$. Let us assume that f satisfies the assumption of Proposition 2.7, i.e. there exist $0 < T < \infty$ and $\delta \geq 1$ such that $f \in L^2(\mathbb{R}_+; H)$, $\operatorname{ess\,supp} f \subset [0,T]$ and $f \in \mathscr{G}^\delta((0,T); H)$. We split $\mathbb{R}_+$ in three subintervals for the analysis, namely $(0,T)$, (T, T_∞) and (T_∞, ∞) for a T_∞ to be defined later.

Given two grading factors $\sigma_1, \sigma_2 \in (0,1)$ and $M \in \mathbb{N}$, we consider a mesh $\mathcal{T}_\sigma = \{I_{\sigma,n}\}_{n=1}^{2M+1}$, where $I_{\sigma,n} = (t_{\sigma,n-1}, t_{\sigma,n})$ for

$$t_{\sigma,0} = 0, \qquad t_{\sigma,m} = T\sigma_1^{M-m}, \qquad t_{\sigma,M+m} = T + M\sigma_2^{M-m}, \qquad m = 1,\ldots,M, \tag{4.4}$$

where $\sigma = (\sigma_1, \sigma_2)$. Moreover we choose $t_{\sigma,2M+1} = T + M + \sigma_2$ and $T_\infty = T + M$. The choice of $t_{\sigma,2M+1}$ is arbitrary and any constant instead of σ_2 will lead to the result presented below. Furthermore, for slope parameters $\mu_1, \mu_2 > 0$ and $\delta \geq 1$, we choose for $m = 2,\ldots,M$

$$p_{\mu,\delta,1} = p_{\mu,\delta,M+1} = p_{\mu,\delta,2M+1} = 1, \qquad p_{\mu,\delta,m} = \max\left(1, \lfloor \mu_1 m^\delta \rfloor\right), \qquad p_{\mu,\delta,M+m} = \max\left(1, \lfloor \mu_2 m \rfloor\right), \tag{4.5}$$

and denote $\mathbf{p}_{\mu,\delta} := \{p_{\mu,\delta,m}\}_{m=1}^{2M+1}$ for $\mu := (\mu_1, \mu_2)$. Let $N_t := \dim\left(S_0^{\mathbf{p}_{\mu,\delta},1}(\mathbb{R}_+, \mathcal{T}_\sigma)\right)$, where

$$S_0^{\mathbf{p}_{\mu,\delta},1}(\mathbb{R}_+, \mathcal{T}_\sigma) := \left\{ v \in H_0^1(\mathbb{R}_+) \;\middle|\; v|_{I_{\sigma,n}} \in \mathcal{P}^{p_{\mu,\delta,n}}(I_{\sigma,n}),\ n = 1, \dots, 2M+1,\ v|_{(t_{\sigma,2M+1},\infty)} = 0 \right\}.$$

We then define

$$\left(\pi_{N_t}^{\sigma,\mu,\delta} u\right)(t) := \begin{cases} \left(\pi_{I_{\sigma,n}}^{p_{\mu,\delta,n}} u\right)(t) & t \in I_{\sigma,n},\ n \in \{1, \dots, 2M\}, \\ \left(\frac{t_{\sigma,2M+1} - t}{h_{\sigma,2M+1}}\right) u(T_\infty) & t \in I_{\sigma,2M+1}, \\ 0 & t > t_{\sigma,2M+1}, \end{cases} \tag{4.6}$$

where $\pi_{I_{\sigma,n}}^{p_{\mu,\delta,n}}$ is the interpolation operator defined in Section 3.1 for $n \in \{1, \dots, 2M\}$. In other words, we use the interpolation operator built in Section 3.1 on the intervals $(0, T)$ and (T, T_∞) while on $(T_\infty, t_{\sigma,2M+1})$ we consider a linear interpolation of the function. The reason for choosing a linear approximation on the last interval and not a constant approximation is to ensure that we obtain a continuous element. On the remaining of the interval, since the solution u associated to f decays exponentially, we approximate the function by zero. In particular, we have the following lemma. The proof is omitted and we refer the reader to [7, Lemma 5.2.1].

Lemma 4.3. *Let $f \in L^2(\mathbb{R}_+; H)$ satisfy $\operatorname{ess\,supp} f \subset [0, T]$ for some $0 < T < \infty$. Then there exist $C, b > 0$ such that for any $T < T_\infty < \infty$ it holds*

$$\left\| u - \pi_{N_t}^{\sigma,\mu,\delta} u \right\|_{L^2((T_\infty,\infty);V)} \le C e^{-b(T_\infty - T)},$$
$$\left\| u - \pi_{N_t}^{\sigma,\mu,\delta} u \right\|_{H_{00}^{1/2}((T_\infty,\infty);H)} \le C e^{-b(T_\infty - T)}.$$

We point out that this result is independent of the choice of the parameters σ, μ and δ. In fact, it only depends on T_∞ and $t_{\sigma,2M+1}$. It can hence also be used in the case of h- or p-approximation to obtain convergence results. We then obtain the following important result.

Theorem 4.4. *Let $f \in L^2(\mathbb{R}_+; H)$ satisfy $\operatorname{ess\,supp} f \subset [0, T]$ and $f \in \mathscr{G}^\delta([0, T]; H)$ for some $\delta \ge 1$ and $0 < T < \infty$. Furthermore let $\mathcal{T}_\sigma$ be defined according to (4.4) for given $M \in \mathbb{N}$ and $\sigma_1, \sigma_2 \in (0, 1)$. Then there exists $\mu_0 = (\mu_{1,0}, \mu_{2,0}) > 0$ such that defining $\mathbf{p}_{\mu,\delta}$ as (4.5) for any $\mu > \mu_0$, there exist $C, b > 0$ with*

$$\left\| u - \pi_{N_t}^{\sigma,\mu,\delta} u \right\|_{\mathcal{X}} \le C \exp\left(-b N_t^{\frac{1}{\delta+1}}\right).$$

Proof. To simplify the notations, we write $I_1 = (0, T)$, $I_2 = (T, T_\infty)$ and $I_3 = (T_\infty, \infty)$. From [7, Remark A.2.5], there exists $C > 0$ such that

$$\begin{aligned} \left\| u - \pi_{N_t}^{\sigma,\mu,\delta} u \right\|_{\mathcal{X}}^2 \le C \Big(& \left\| u - \pi_{N_t}^{\sigma,\mu,\delta} u \right\|_{L^2(I_1;V)}^2 + \left\| u - \pi_{N_t}^{\sigma,\mu,\delta} u \right\|_{L^2(I_2;V)}^2 + \left\| u - \pi_{N_t}^{\sigma,\mu,\delta} u \right\|_{L^2(I_3;V)}^2 \\ & + \left\| u - \pi_{N_t}^{\sigma,\mu,\delta} u \right\|_{H_{00}^{1/2}(I_1;H)}^2 + \left\| u - \pi_{N_t}^{\sigma,\mu,\delta} u \right\|_{H_{00}^{1/2}(I_2;H)}^2 + \left\| u - \pi_{N_t}^{\sigma,\mu,\delta} u \right\|_{H_{00}^{1/2}(I_3;H)}^2 \Big). \end{aligned}$$

Note that

$$N_t = 2 + \sum_{m=2}^{M} \left(\lfloor \mu_1 m^\delta \rfloor + \lfloor \mu_2 m \rfloor \right) \le 2 \max(1, \mu_1, \mu_2) \sum_{m=1}^{M} m^\delta \le 2 \max(1, \mu_1, \mu_2) M^{\delta+1}.$$

From Lemma 4.3 and since $T_\infty = T + M$, there exists $\tilde{b}_3 > 0$ such that

$$\left\| u - \pi_{N_t}^{\sigma,\mu,\delta} u \right\|_{L^2(I_3;V)} \le C_3 \exp\left(-\tilde{b}_3 (T_\infty - T)\right) = C_3 \exp\left(-\tilde{b}_3 M\right) \le C_3 \exp\left(-b_3 N_t^{\frac{1}{\delta+1}}\right),$$
$$\left\| u - \pi_{N_t}^{\sigma,\mu,\delta} u \right\|_{H_{00}^{1/2}(I_3;H)} \le C_3 \exp\left(-\tilde{b}_3 (T_\infty - T)\right) = C_3 \exp\left(-\tilde{b}_3 M\right) \le C_3 \exp\left(-b_3 N_t^{\frac{1}{\delta+1}}\right),$$

where $b_3 = \tilde{b}_3 \left(2 \max(1, \mu_1, \mu_2)\right)^{-\frac{1}{\delta+1}}$. From Proposition 2.7, it holds

$$u|_{I_1} \in \mathscr{G}^{\delta,1}(I_1; H) \cap \mathscr{G}^{\delta,1/2}(I_1; V), \qquad u|_{I_2} \in \mathscr{G}^{1,1/2}(I_2; H) \cap \mathscr{G}^{1,0}(I_2; V),$$

and Remark 2.8 yields that the constants C_u and d_u in (2.10) on the interval I_2 are independent of T_∞. From Propositions 3.6 and 3.7, there exist $\mu_{1,H}, \mu_{1,V} > 0$ such that for $\mu_1 > \max(\mu_{1,H}, \mu_{1,V}) =: \mu_{1,0}$ it holds

$$\left\| u - \pi_{N_t}^{\sigma,\mu,\delta} u \right\|_{L^2(I_1;V)} \leq C_1 \exp\left(-b_1 N_t^{\frac{1}{\delta+1}}\right),$$
$$\left\| u - \pi_{N_t}^{\sigma,\mu,\delta} u \right\|_{H_{00}^{1/2}(I_1;H)} \leq C_1 \exp\left(-b_1 N_t^{\frac{1}{\delta+1}}\right).$$

Using the same argument together with the explicit dependence on $|I_2|$ in (3.7) and (3.9), there exist $\mu_{2,H}, \mu_{2,V} > 0$ such that for $\mu_2 > \max(\mu_{2,H}, \mu_{2,V}) =: \mu_{2,0}$ it holds

$$\left\| u - \pi_{N_t}^{\sigma,\mu,\delta} u \right\|_{L^2(I_2;V)} \leq C_2 \exp\left(-b_2 \sqrt{N_t}\right) \leq C_2 \exp\left(-b_2 N_t^{\frac{1}{\delta+1}}\right),$$
$$\left\| u - \pi_{N_t}^{\sigma,\mu,\delta} u \right\|_{H_{00}^{1/2}(I_2;H)} \leq C_2 \exp\left(-b_2 \sqrt{N_t}\right) \leq C_2 \exp\left(-b_2 N_t^{\frac{1}{\delta+1}}\right).$$

Wrapping up everything, we obtain the desired result. ∎

REMARK 4.5. *Using Remark 3.4, we can also obtain convergence results for h- and p-refinements considering a quasi-uniform mesh and a fixed polynomial degree. However in the case of h-approximation, due to the fact that the solution is approximated over an unbounded domain, the convergence rate is decreased by a logarithmic factor. More precisely, assume that $f \in L^2(\mathbb{R}_+; H)$ satisfy $\operatorname{ess\,supp} f \subset= [0, T]$ for some $0 < T < \infty$ and that there exists $r \geq 0$ such that*

$$u|_I \in H^{r+1}((0,T); H) \cap H^{r+1/2}((0,T); V) \text{ and } u|_{(T,\infty)} \in H^{r+1}((T,\infty); H) \cap H^{r+1/2}((T,\infty); V). \tag{4.7}$$

Then for $p, M \in \mathbb{N}$, $h_1 = \frac{T}{M}$ and $h_2 = \frac{2\min(p,r)+1}{2M}$, we define

$$t_0 = 0, \qquad t_j = jh_1, \qquad t_{j+M} = T + jh_2, \qquad p_j = p, \qquad j = 1, \ldots, M. \tag{4.8}$$

It is then possible to show that [7, Remark 5.2.5]

$$\left\| u - \pi_{N_t}^{p} u \right\|_{\mathcal{X}} \leq C \left(\frac{\log(N_t)}{N_t}\right)^{\min(p,r)+1/2}, \tag{4.9}$$

where $N_t = Mp$ and $\pi_{N_t}^{p}$ is the interpolation operator defined in (4.6) associated to the mesh defined through (4.8) and the vector of fixed polynomial degrees equal to p.

Turning to the p-refinement case, let f be as above and assume that there exist $C, d > 0$ such that

$$|u|_{H^r((0,T);V)} \leq C d^r \Gamma(r+1), \qquad |u|_{H^r((T,\infty);V)} \leq C d^r, \qquad r \in \mathbb{N}_0.$$

For $p, M \in \mathbb{N}$, $h_1 = \frac{T}{M}$ and $h_2 = \frac{p}{M}$, the underlying fixed mesh and the vector of polynomial degrees are defined as in (4.8). It is then possible to show that [7, Remark 5.2.6]

$$\left\| u - \pi_{N_t}^{p} u \right\|_{\mathcal{X}} \leq C e^{-bp},$$

for some constants $C, b > 0$.

4.2. Space-time approximation. So far, we only derived error estimates for the approximation of the temporal variable. We now turn to the full space-time discretization. First, we derive error estimates in an abstract setting and apply this theory to space-time h-refinement. In a second step, the two-dimensional heat equation in a polygonal domain is considered. We show that using an hp-approximation in both time and space it is possible to obtain exponential convergence with respect to the total number of degrees of freedom. Since we consider in this section both space and time approximations, we use t and x as subscripts associated to the temporal and spatial component, respectively.

4.2.1. An abstract result. Let us consider a sequence of Hilbert spaces $\{X_s\}_{s\geq 0}$ such that $X_0 = H$, $X_\varepsilon = V$ for some $\varepsilon \geq 0$ and $X_{s_2} \subset X_{s_1}$ for all $0 \leq s_1 \leq s_2 < \infty$. We also assume that the norm on X_s is equivalent to the interpolation norm on $[X_{s_1}, X_{s_2}]_\theta$ for any $\varepsilon \leq s_1 < s < s_2 < \infty$ and $\theta = \frac{s-s_1}{s_2-s_1}$. The spaces X_s are then called *smoothness spaces* associated to (H, V).

Definition 4.6 (Approximation of order $\overline{s}$). *Let $\{X_s\}_{s\geq 0}$ be a sequence of* smoothness spaces *associated to (H,V) and $\varepsilon \leq \overline{s} < \infty$. Then a sequence of finite dimensional spaces $\{V_{N_x}\}_{N_x\in\mathbb{N}} \subset V$ is said to be an approximation of order $\overline{s}$ of the pair (H,V) (or equivalently (X_0, X_ε)) if for every $N_x \in \mathbb{N}$ there exists an operator $\Pi_{N_x} : V \to V_{N_x}$ such that for $\varepsilon \leq s \leq \overline{s}$ and $v \in X_s$ it holds*

$$\|v - \Pi_{N_x} v\|_V \leq C_s N_x^{-(s-\varepsilon)} \|v\|_{X_s},$$
$$\|v - \Pi_{N_x} v\|_H \leq C_s N_x^{-s} \|v\|_{X_s},$$

for a constant $C_s > 0$.

Based on this, we derive convergence rates for the h-approximation discussed in Remark 4.5. The mesh defined through (4.8) will be denoted $\mathcal{T}_{t,h}$. Furthermore, we write p_t for the vector of fixed polynomial degrees.

Proposition 4.7. *Let $\{V_{N_x}\}_{N_x\in\mathbb{N}} \subset V$ be an approximation of order $\overline{s}$ of the pair $(H,V) = (X_0, X_\varepsilon)$ and $f \in L^2(\mathbb{R}_+; H)$ satisfy* $\operatorname{ess\,supp} f \subset= [0,T]$ *for some $0 < T < \infty$. Given $p_t, N_t, N_x \in \mathbb{N}$ such that $\frac{N_t}{p_t} \in \mathbb{N}$, we denote $u_h^{(N_t,N_x)}$ the unique solution of (4.1) in the discrete space*

$$\mathcal{X}_h^{(N_t,N_x)} := S_0^{p_t,1}(\mathbb{R}_+, \mathcal{T}_{t,h}) \otimes V_{N_x}.$$

Let us assume that

$$u|_{(0,T)} \in H^{r+1}((0,T);H) \cap H^{r+1/2}((0,T);V) \text{ and } u|_{(T,\infty)} \in H^{r+1}((T,\infty);H) \cap H^{r+1/2}((T,\infty);V)$$

for some $r \geq 0$ and

$$u \in H_{00}^{1/2}(\mathbb{R}_+; X_s) \text{ and } u \in L^2(\mathbb{R}_+; X_{s+\varepsilon})$$

for some $\varepsilon \leq s \leq \overline{s} - \varepsilon$. Then there exists $C > 0$ such that

$$\left\|u - u_h^{(N_t,N_x)}\right\|_{\mathcal{X}} \leq C\left(N_x^{-s} + \left(\frac{N_t}{\log(N_t)}\right)^{-(\min(p_t,r)+1/2)}\right).$$

Proof. From (4.2) and (4.3), we have

$$\left\|u - u_h^{(N_t,N_x)}\right\|_{\mathcal{X}} \leq C\left(\left\|u - \left(\pi_{N_t}^{p_t} \otimes \mathrm{Id}\right) u\right\|_{\mathcal{X}} + \left\|u - \left(\mathrm{Id} \otimes \Pi_{N_x}\right) u\right\|_{\mathcal{X}}\right).$$

From (4.9), we have

$$\left\|u - \left(\pi_{N_t}^{p_t} \otimes \mathrm{Id}\right) u\right\|_{\mathcal{X}} \leq C\left(\frac{N_t}{\log(N_t)}\right)^{-(\min(p_t,r)+1/2)}.$$

Since $\{V_{N_x}\}_{N_x\in\mathbb{N}}$ is an approximation of order $\overline{s}$ of (H,V), we have

$$\|\mathrm{Id}_V - \Pi_{N_x}\|_{\mathcal{L}(X_{s+\varepsilon},V)} \leq C_s N_x^{-s}, \qquad \|\mathrm{Id}_H - \Pi_{N_x}\|_{\mathcal{L}(X_s,H)} \leq C_s N_x^{-s}.$$

Since $H_{00}^{1/2}(\mathbb{R}_+;H)$ and $L^2(\mathbb{R}_+;V)$ are isomorphic to $H_{00}^{1/2}(\mathbb{R}_+) \otimes H$ and $L^2(\mathbb{R}_+) \otimes V$, respectively, the result follows by interpolation [27, Theorem 8.32(b)]. ∎

Example 4.8. *Let $D \subset \mathbb{R}^d$ with $d = 1,2$ be an interval or a bounded polygonal domain in the case $d = 2$ and define $H = L^2(D)$ and $V = H_0^1(D)$. Then for every $s \geq 0$, we define $X_s = H^{ds}(D) \cap H_0^1(D)$ and choose*

$\varepsilon = 1/d$. In that case, the interpolation property follows from [26, Theorem 2 p.317]. Let us consider a quasi-uniform mesh $\mathcal{T}_{x,h}$ of D with meshwidth h_x. Given a polynomial degree $p_x \in \mathbb{N}$, we consider the space of continuous, piecewise polynomial functions

$$V_{N_x} := S_0^{p_x,1}(D, \mathcal{T}_{x,h}) := \left\{ v \in H_0^1(D) \;\middle|\; v|_T \in \mathcal{P}^{p_x}(T),\ \forall T \in \mathcal{T}_{x,h} \right\}.$$

Then, there exist $c_1, c_2 > 0$ such that $c_1 N_x^{-1/d} \leq h_x \leq c_2 N_x^{-1/d}$. From [24] (for $d = 2$) and [22] (for $d = 1$), we can define for every V_{N_x} an interpolation operator Π_{N_x} such that for any $1 \leq l \leq p_x + 1$ it holds

$$\begin{aligned}
\|v - \Pi_{N_x} v\|_{L^2(D)} &\leq C h^l \|v\|_{H^l(D)} \leq C N_x^{-s} \|v\|_{X_s}, && \forall v \in H^l(D) \cap H_0^1(D),\\
\|v - \Pi_{N_x} v\|_{H^1(D)} &\leq C h^{l-1} \|v\|_{H^l(D)} \leq C N_x^{-\left(s-\frac{1}{d}\right)} \|v\|_{X_s}, && \forall v \in H^l(D) \cap H_0^1(D),
\end{aligned}$$

where $s = l/d$. It follows that the sequence of spaces $\{V_{N_x}\}_{N_x\in\mathbb{N}}$ is an approximation of order $\frac{p_x+1}{d}$ of the pair $(L^2(D), H_0^1(D))$. Hence if

$$\begin{aligned}
u|_{(0,T)} &\in H^{p_t+1}((0,T); L^2(D)) \cap H^{p_t+1/2}((0,T); H_0^1(D)),\\
u|_{(T,\infty)} &\in H^{p_t+1}((T,\infty); L^2(D)) \cap H^{p_t+1/2}((T,\infty); H_0^1(D)),
\end{aligned}$$

and

$$u \in H_{00}^{1/2}(\mathbb{R}_+; H^{p_x}(D) \cap H_0^1(D)) \text{ and } u \in L^2(\mathbb{R}_+; H^{p_x+1}(D) \cap H_0^1(D)),$$

then

$$\left\| u - u_h^{(N_t,N_x)} \right\|_{\mathcal{X}} \leq C \left(N_x^{-p_x/d} + \left(\frac{N_t}{\log(N_t)} \right)^{-(p_t+1/2)} \right). \tag{4.10}$$

Moreover if u has support in $(0,T)$, the above estimate reduces to

$$\left\| u - u_h^{(N_t,N_x)} \right\|_{\mathcal{X}} \leq C \left(N_x^{-p_x/d} + N_t^{-(p_t+1/2)} \right). \tag{4.11}$$

4.2.2. Space-time hp-approximation of the two-dimensional heat equation. Let $D \subset \mathbb{R}^2$ be a (possibly non-convex) bounded polygonal domain and $\mathcal{A}v := -\operatorname{div}_x(\tau \nabla_x v)$ for some $\tau \in \left(W^{2,\infty}(D)\right)^{2\times 2}$ such that there exists $\lambda_- > 0$ satisfying

$$\sum_{i,j=1}^{2} \tau_{i,j}(x) \xi_i \xi_j \geq \lambda_- |\xi|^2, \qquad \forall \xi \in \mathbb{R}^2,\ x \in D.$$

In that case $H = L^2(D)$, $V = H_0^1(D)$ and $\mathfrak{a}(v,w) = \int_D (\tau \nabla_x v) \cdot \nabla_x w$. The problem in the weak form then becomes: given $f \in \mathcal{Y}^*$, find $u \in \mathcal{X}$ such that

$$\int_{\mathbb{R}_+} \int_D \left[\partial_{t,+}^{1/2} u(t,x) \partial_{t,-}^{1/2} v(t,x) + (\tau(x) \nabla_x u(t,x)) \cdot \nabla_x v(t,x) \right] dx dt = \langle f, v \rangle_{\mathcal{Y}^*}, \qquad \forall v \in \mathcal{Y}, \tag{4.12}$$

where $\partial_{t,+}^{1/2}$ and $\partial_{t,-}^{1/2}$ denote the operators $D_+^{1/2}$ and $D_-^{1/2}$ with respect to the temporal variable, respectively.

In order to account for the spatial regularity, we introduce weighted Sobolev spaces which are used in the regularity theory of elliptic equations. We follow here [22, Chapter 4]. Let $A_1, \ldots, A_L$ denote the vertices of D, $\gamma = (\gamma_1, \ldots, \gamma_L) \in [0,1)^L$ and

$$\Phi_\gamma(x) := \prod_{j=1}^{L} (r_j(x))^{\gamma_j}, \qquad r_i(x) := \min(1, |x - A_i|), \qquad \forall x \in \overline{D},\ i = 1, \ldots, L. \tag{4.13}$$

Then for a Hilbert space $(H,(\cdot,\cdot)_H)$ and $m \geq l \geq 1$, we define the weighted Bochner-Sobolev space $H^{m,l}_\gamma(D;H)$ as the space of functions $v \in L^2(D;H)$ for which

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

where

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

Definition 4.9. *Let $D \subset \mathbb{R}^2$ be a bounded polygonal domain, $(H,(\cdot,\cdot)_H)$ a Hilbert space and $\gamma = (\gamma_1,\ldots,\gamma_L) \in [0,1)^L$. For $l \geq 1$, $v \in \mathcal{B}^l_\gamma(D;H)$ if $v \in H^{m,l}_\gamma(D;H)$ for all $m \geq l$ and there exist $C_v > 0$ and $d_v \geq 1$ such that for all $k \geq l$ it holds*

$$\sum_{|\alpha|=k} \|\Phi_{\gamma+k-l}\partial_x^\alpha v\|_{L^2(D;H)} \leq C_v d_v^{k-l}\Gamma(k-l+1),$$

where Φ_γ is defined in (4.13).

Functions in $\mathcal{B}^2_\gamma(D;H)$ are globally continuous, analytic inside D and contain potential algebraic singularities at the vertices $A_1,\ldots,A_L$.

For $\delta \geq 1$, $\sigma_{t,1},\sigma_{t,2} \in (0,1)$ and $\mu_{t,1},\mu_{t,2} > 0$, let $\mathcal{T}_{\sigma_t}$ and $\mathbf{p}_{\mu_t,\delta}$ be defined through (4.4) and (4.5), respectively. Here we have used the notation $\sigma_t = (\sigma_{t,1},\sigma_{t,2})$ and $\mu_t = (\mu_{t,1},\mu_{t,2})$. Furthermore, let $\mathcal{T}_{\sigma_x}$ be a proper graded mesh of D for a grading factor $\sigma_x \in (0,1)$. Given a slope parameter $\mu_x > 0$, $\mathbf{p}_{\mu_x}$ denotes a polynomial distribution associated to $\mathcal{T}_{\sigma_x}$ whose elements are increasing linearly away from the vertices $A_1,\ldots,A_L$. We refer the reader to [22, Chapter 4] for the construction of $\mathcal{T}_{\sigma_x}$ and $\mathbf{p}_{\mu_x}$. Let then

$$\mathcal{X}^N_{\sigma,\mu,\delta} := S_0^{\mathbf{p}_{\mu_t,\delta},1}(\mathbb{R}_+,\mathcal{T}_{\sigma_t}) \otimes S_0^{\mathbf{p}_{\mu_x},1}(D,\mathcal{T}_{\sigma_x}),$$

where $\sigma = (\sigma_{t,1},\sigma_{t,2},\sigma_x)$, $\mu = (\mu_{t,1},\mu_{t,2},\mu_x)$ and

$$S_0^{\mathbf{p}_{\mu_x},1}(D,\mathcal{T}_{\sigma_x}) := \left\{ v \in H^1_0(D) \;\middle|\; v|_T \in \mathcal{P}^{p_{\mu_x,T}}(T),\ \forall T \in \mathcal{T}_{\sigma_x} \right\} \subset \mathcal{C}(D).$$

In the following theorem, $N_x := \dim\left(S_0^{\mathbf{p}_{\mu_x},1}(D,\mathcal{T}_{\sigma_x})\right)$ and $N_t = \dim\left(S_0^{\mathbf{p}_{\mu_t,\delta},1}(\mathbb{R}_+,\mathcal{T}_{\sigma_t})\right)$ denote the number of degrees of freedom associated to the spatial and temporal approximations, respectively. Furthermore, $N = N_x N_t$.

Theorem 4.10. *Let $f \in L^2(\mathbb{R}_+;L^2(D))$ satisfy $\operatorname{ess\,supp} f \subset [0,T]$ and $f \in \mathscr{G}^\delta((0,T);L^2(D))$ for some $0 < T < \infty$ and $\delta \geq 1$. Furthermore assume that $u \in H^{1/2}_{00}(\mathbb{R}_+;\mathcal{B}^2_\gamma(D) \cap H^1_0(D))$. Then for grading factors $\sigma_1,\sigma_2,\sigma_x \in (0,1)$, there exist $\mu_0 = (\mu_{t,1,0},\mu_{t,2,0},\mu_{x,0}) > 0$ such that for all $\mu > \mu_0$, there exist $C,b > 0$ satisfying*

$$\|u-u_N\|_{\mathcal{X}} \leq C\left(\exp\left(-bN_t^{\frac{1}{\delta+1}}\right) + \exp\left(-bN_x^{\frac{1}{3}}\right)\right),$$

where u_N denotes the solution of (4.12) in $\mathcal{X}^N_{\sigma,\mu,\delta}$.

Proof. In [22, Chapter 4], an operator $\pi^{\sigma_x,\mu_x}_{N_x} : \mathcal{B}^2_\gamma(D) \cap H^1_0(D) \to S_0^{\mathbf{p}_{\mu_x},1}(D,\mathcal{T}_{\sigma_x})$ is introduced. It is shown that there exists $\mu_{x,0} > 0$ such that for all $\mu_x > \mu_{x,0}$, there exist $C,b > 0$ satisfying for all $v \in \mathcal{B}^2_\gamma(D) \cap H^1_0(D)$

$$\left\|v - \pi^{\sigma_x,\mu_x}_{N_x} v\right\|_{H^1(D)} \leq C\exp\left(-bN_x^{\frac{1}{3}}\right).$$

We point out that this result remains valid for functions taking values in a Hilbert space. It follows that for $u \in H^{1/2}_{00}(\mathbb{R}_+;\mathcal{B}^2_\gamma(D) \cap H^1_0(D))$, we have

$$\left\|u - \left(\mathrm{Id} \otimes \pi^{\sigma_x,\mu_x}_{N_x}\right) u\right\|_{\mathcal{X}} \leq C\exp\left(-bN_x^{\frac{1}{3}}\right).$$

The result then follows from (4.3) and Theorem 4.4. ∎

REMARK 4.11. *The argument used above can be simplified in the one-dimensional case in order to obtain similar results. Indeed, a theory can be developed based on [22, Chapter 3] for the spatial hp-refinement. In that setting, using graded meshes and linearly increasing polynomial degrees in a similar fashion as for the temporal discretization, we obtain*

$$\|u - u_N\|_{\mathcal{X}} \leq C \left(\exp\left(-bN_t^{\frac{1}{\delta+1}}\right) + \exp\left(-bN_x^{\frac{1}{2}}\right)\right),$$

assuming that f satisfies the assumption of Theorem 4.10 and that u exhibits more regularity in space.

5. Numerical results.

In order to provide numerical evidence of the results obtained in the previous section, we consider the one-dimensional heat equation with $D = (0,1)$ and $\tau = 1$. We present convergence results for h- and hp-refinements. Similar results for p-refinement can also be obtained [7, Section 6.4]. In all cases, the grading method considered is the same for the spatial and the temporal discretization. All the computations have been performed using the Python programming language.

5.1. h-refinement.

Let us consider the solutions

$$u_1(t,x) = \sin(\pi t)\sin(\pi x)\chi_{(0,1)}(t), \qquad u_2(t,x) = \cos\left(\frac{5\pi}{2}t\right)\sin(\pi x)\chi_{(0,1)}(t).$$

We use the same mesh in time and space with $M = M_x = M_t$ elements, so that $N = N_x N_t \approx N_t^2$. Furthermore the functions are piecewise analytic and have compact support in time, so that estimate (4.11) yields

$$\|u_i - u_{i,N}\|_{\mathcal{X}} \leq C N^{-\frac{1}{2}\min(p_x, p_t+1/2)}, \qquad i = 1,2.$$

We consider $p = p_t = p_x$ in the case $i = 1$ and $p = p_t = p_x - 1$ for $i = 2$ in order to see that the exponents are sharp. It gives

$$\|u_1 - u_{1,N}\|_{\mathcal{X}}^2 \leq C N^{-p}, \qquad \|u_2 - u_{2,N}\|_{\mathcal{X}}^2 \leq C N^{-(p+1/2)}.$$

The experimental convergence rates for the error in the $\mathcal{X}$-norm with respect to $\sqrt{N}$ for both cases and different values of p are presented in the following table. The estimated slopes are based on the last 5 points of the experiment. Convergence plots associated to the experimental rates obtained for u_1 and u_2 are depicted in Figure 5.1. We see that the expected convergence rates are indeed obtained.

p	u_1 - Expected	u_1 - Observed	u_2 - Expected	u_2 - Observed
1	1	0.996	1.5	1.517
2	2	1.994	2.5	2.442
3	3	2.987	3.5	3.517
4	4	3.928	4.5	4.493

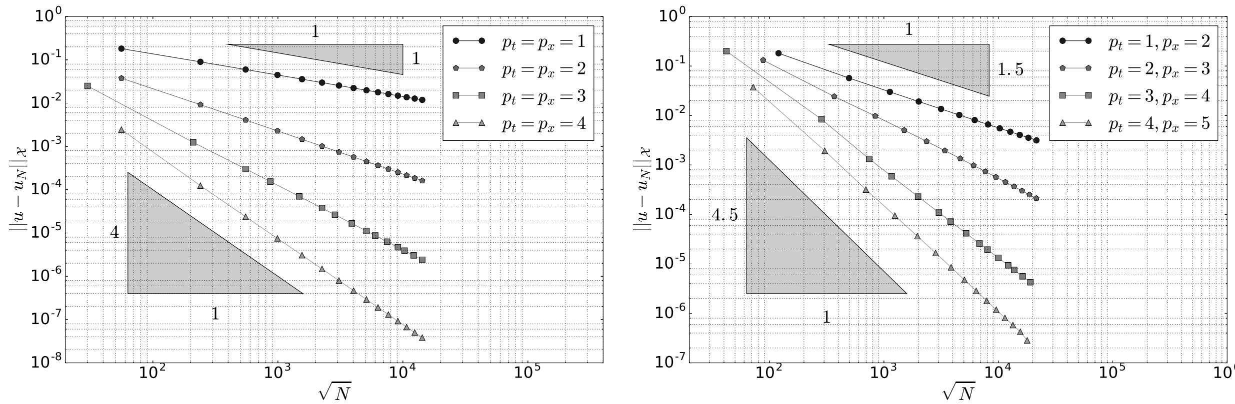
5.2. hp-refinement.

Given $\alpha > -2$ and $\beta > 0$, we consider

$$f(t,x) = x^\alpha(1-x)e^{-\beta t}.$$

The solution of (4.12) associated to f contains a singularity at $t = x = 0$ and can be computed using a Fourier expansion based on the one-dimensional Laplace operator. The singularity at $x = 0$ is due to the x-dependence of f while the one at $t = 0$ follows from an incompatibility between the t-dependence of f and the initial condition. Assuming that $\beta \neq \pi k$ for all $k \in \mathbb{N}$, the solution is then given by

$$u(t,x) = 2\sum_{k=1}^{\infty}(f_{\alpha,k} - f_{\alpha+1,k})\sin(\pi k x)\left(\frac{e^{-\beta t} - e^{-\pi^2 k^2 t}}{\pi k - \beta}\right), \tag{5.1}$$



(a) (b)

Figure 5.1: Convergence of the h-approximation in time and space with respect to $\sqrt{N}$ for several polynomial degrees. The results for u_1 are presented in (a) while (b) corresponds to u_2.

where

$$f_{\alpha,k} := \int_0^1 x^\alpha \sin(\pi k x)dx = \frac{\pi k}{\alpha+2}\,{}_1F_2\left(\frac{\alpha}{2}+1; \frac{3}{2}, \frac{\alpha}{2}+2; -\frac{(\pi k)^2}{4}\right), \qquad \forall k \in \mathbb{N},$$

and ${}_1F_2$ denotes a generalized hypergeometric function [18, Section 16.2]. For our experiments, we choose $\alpha = 3/4$ and $\beta = 5$. We consider geometric meshes in time and space and use the same grading factors and slope parameters for the spatial and the temporal discretization. Let us denote $\sigma = \sigma_x = \sigma_t \in (0,1)$ and $\mu = \mu_x = \mu_t > 0$ the grading factors and slope parameters, respectively. Considering Remark 4.11, we have

$$\|u - u_N\|_{\mathcal{X}} \le C\left(\exp\left(-b\sqrt{N_t}\right) + \exp\left(-b\sqrt{N_x}\right)\right),$$

which corresponds to the case $\delta = 1$. Since $M = M_x = M_t$, we have $N = N_x N_t \approx N_x^2 \approx N_t^2$, so that

$$\|u - u_N\|_{\mathcal{X}} \le C \exp\left(-bN^{\frac{1}{4}}\right).$$

In Figure 5.2, we present convergence results for hp-discretization in time and space for different values of σ and μ. We see that that for smaller values of σ, larger values of μ need to be considered. Exponential convergence can indeed be observed.

Acknowledgement. This work was performed in part while the author was PhD student at ETH Zürich under the supervision of Prof. Ch. Schwab. Research supported in part by the Swiss National Science Foundation under Grant No. 149819. We also acknowledge helpful discussions at the Mathematisches Forschungsinstitut Oberwolfach (MFO) during the workshop Space-time Methods, 13-17 March 2017.

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, https://doi.org/10.1093/imanum/drs014, http://dx.doi.org/10.1093/imanum/drs014.

[2] R. Andreev, *Space-time discretization of the heat equation*, Numer. Algorithms, 67 (2014), pp. 713–731, https://doi.org/10.1007/s11075-013-9818-4, http://dx.doi.org/10.1007/s11075-013-9818-4.

[3] I. Babuška and B. Guo, *Direct and inverse approximation theorems for the p-version of the finite element method in the framework of weighted Besov spaces. I. Approximability of functions in the weighted Besov spaces*, SIAM J. Numer. Anal., 39 (2001/02), pp. 1512–1538, https://doi.org/10.1137/S0036142901356551, http://dx.doi.org/10.1137/S0036142901356551.

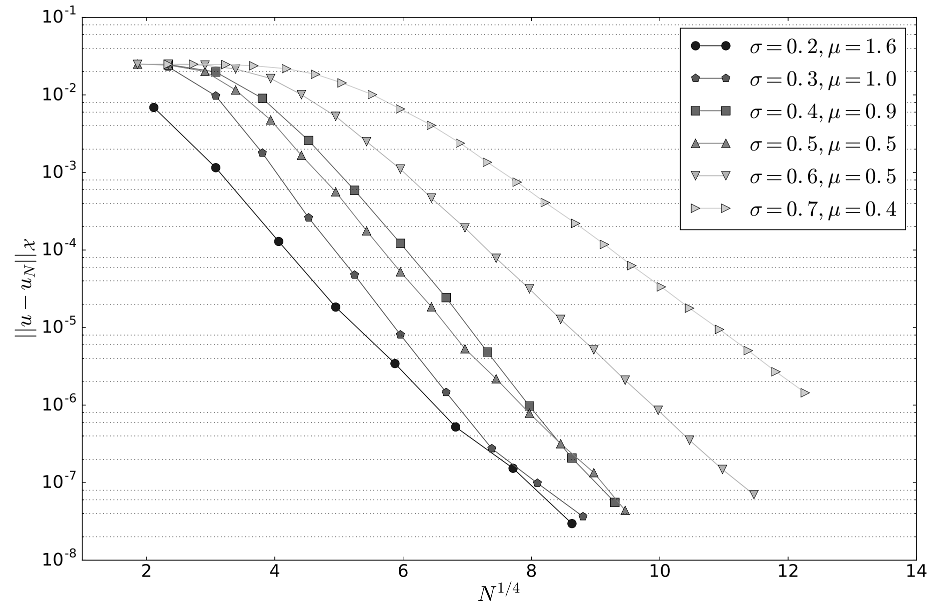


Figure 5.2: Convergence of the hp-approximation in time and space of u given by (5.1) for $\alpha = 3/4$ and $\beta = 5$. The grading factors $\sigma = \sigma_t = \sigma_x$ and slopes $\mu = \mu_t = \mu_x$ are the same for the temporal and spatial discretizations.

[4] I. Babuška and B. Guo, *Direct and inverse approximation theorems for the p-version of the finite element method in the framework of weighted Besov spaces. II. Optimal rate of convergence of the p-version finite element solutions*, Math. Models Methods Appl. Sci., 12 (2002), pp. 689–719, https://doi.org/10.1142/S0218202502001854, http://dx.doi.org/10.1142/S0218202502001854.

[5] J. Bergh and J. Löfström, *Interpolation spaces. An introduction*, Springer-Verlag, Berlin-New York, 1976. Grundlehren der Mathematischen Wissenschaften, No. 223.

[6] D. Braess, *Finite elements*, Cambridge University Press, Cambridge, third ed., 2007, https://doi.org/10.1017/CBO9780511618635, http://dx.doi.org/10.1017/CBO9780511618635. Theory, fast solvers, and applications in elasticity theory, Translated from the German by Larry L. Schumaker.

[7] D. Devaud, *hp-approximation of linear parabolic evolution problems in $H^{1/2}$*, PhD thesis, ETH Zurich, 2017, https://doi.org/10.3929/ethz-b-000173039.

[8] M. Fontes, *A monotone operator method for elliptic-parabolic equations*, Comm. Partial Differential Equations, 25 (2000), pp. 681–702, https://doi.org/10.1080/03605300008821527, http://dx.doi.org/10.1080/03605300008821527.

[9] M. Fontes, *Initial-boundary value problems for parabolic equations*, Ann. Acad. Sci. Fenn. Math., 34 (2009), pp. 583–605.

[10] M. J. Gander, *50 years of time parallel time integration*, in Multiple Shooting and Time Domain Decomposition Methods, Springer, 2015, pp. 69–113.

[11] B. Guo and N. Heuer, *The optimal rate of convergence of the p-version of the boundary element method in two dimensions*, Numer. Math., 98 (2004), pp. 499–538, https://doi.org/10.1007/s00211-004-0535-8, http://dx.doi.org/10.1007/s00211-004-0535-8.

[12] T. Hytönen, J. van Neerven, M. Veraar, and L. Weis, *Analysis in Banach spaces. Vol.I: Martingales and Littlewood-Paley theory*, vol. 63 of Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics, Springer, 2016, https://doi.org/10.1007/978-3-319-48520-1.

[13] A. A. Kilbas, H. M. Srivastava, and J. J. Trujillo, *Theory and applications of fractional differential equations*, vol. 204 of North-Holland Mathematics Studies, Elsevier Science B.V., Amsterdam, 2006.

[14] 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, https://doi.org/10.1016/j.cma.2016.03.042, http://dx.doi.org/10.1016/j.cma.2016.03.042.

[15] S. Larsson and C. Schwab, *Compressive space-time galerkin discretizations of parabolic partial differential equations*, Tech. Report 2015-04, Seminar for Applied Mathematics, ETH Zürich, Switzerland, 2015, /sam_reports/reports_final/reports2015/2015-04.pdf.

[16] J.-L. Lions and E. Magenes, *Problèmes aux limites non homogènes et applications. Vol. 1*, Travaux et Recherches Mathématiques, No. 17, Dunod, Paris, 1968.

[17] W. McLean, *Strongly elliptic systems and boundary integral equations*, Cambridge University Press, Cambridge, 2000.

[18] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, eds., *NIST handbook of mathematical functions*, U.S.

Department of Commerce, National Institute of Standards and Technology, Washington, DC; Cambridge University Press, Cambridge, 2010. With 1 CD-ROM (Windows, Macintosh and UNIX).

[19] S. G. Samko, A. A. Kilbas, and O. I. Marichev, *Fractional integrals and derivatives*, Gordon and Breach Science Publishers, Yverdon, 1993.

[20] D. Schötzau, *hp-DGFEM for parabolic evolution problems: Applications to diffusion and viscous incompressible fluid flow*, ProQuest LLC, Ann Arbor, MI, 1999, https://doi.org/10.3929/ethz-b-000173039.

[21] D. Schötzau and C. Schwab, *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, https://doi.org/10.1137/S0036142999352394, http://dx.doi.org/10.1137/S0036142999352394.

[22] 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.

[23] C. Schwab and R. Stevenson, *Space-time adaptive wavelet methods for parabolic evolution problems*, Math. Comp., 78 (2009), pp. 1293–1318, https://doi.org/10.1090/S0025-5718-08-02205-9, http://dx.doi.org/10.1090/S0025-5718-08-02205-9.

[24] L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483–493, https://doi.org/10.2307/2008497, http://dx.doi.org/10.2307/2008497.

[25] V. Thomée, *Galerkin finite element methods for parabolic problems*, vol. 25 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, second ed., 2006.

[26] H. Triebel, *Interpolation theory, function spaces, differential operators*, Johann Ambrosius Barth, Heidelberg, second ed., 1995.

[27] J. Weidmann, *Linear operators in Hilbert spaces*, vol. 68 of Graduate Texts in Mathematics, Springer-Verlag, New York-Berlin, 1980. Translated from the German by Joseph Szücs.

[28] 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, https://doi.org/10.1016/S0045-7825(01)00258-4, http://dx.doi.org/10.1016/S0045-7825(01)00258-4.