Numerical Solution of Parabolic Equations in High Dimensions

Tobias von Petersdorff
Department of Mathematics
University of Maryland
College Park, MD 20742
USA

Christoph Schwab[1]
Seminar for Applied Mathematics
ETH Zürich
CH-8092 Zürich
Switzerland

Abstract

We consider the numerical solution of diffusion problems in $(0,T)\times\Omega$ for $\Omega\subset\mathbb{R}^d$ and for $T>0$ in dimension $d\geq 1$. We use a wavelet based sparse grid space discretization with meshwidth h and order $p\geq 1$, and hp discontinuous Galerkin time-discretization of order $r=O(|\log h|)$ on a geometric sequence of $O(|\log h|)$ many time steps. The linear systems in each time step are solved iteratively by $O(|\log h|)$ GMRES iterations with a wavelet preconditioner. We prove that this algorithm gives an $L^2(\Omega)$-error of $O(N^{-p})$ for $u(x,T)$ where N is the total number of operations, provided that the initial data satisfies $u_0\in H^\varepsilon(\Omega)$ with $\varepsilon>0$ and that $u(x,t)$ is smooth in x for $t>0$. Numerical experiments in dimension d up to 20 confirm the theory.

[1]supported in part under the programme "Computational challenges in partial differential equations" at the Isaac Newton Institute, Cambridge, and under the IHP network "Breaking Complexity" of the EC (contract number HPRN-CT-2002-00286) with support by the Swiss Federal Office for Science and Education under grant No. BBW 02.0418.

1 Introduction

The numerical solution of parabolic evolution problems by Finite Elements in a domain $\Omega \subset \mathbb{R}^d$ and by implicit time-stepping in the interval $(0, T)$ is used in numerous applications. There exists a sizeable and well-developed literature on the numerical analysis of discretization schemes, see [17] and the references there. For the solution of the linear system at each time step efficient solvers are available, e.g., based on suitable multilevel schemes. Most of these developments have been focussed on problems in dimension $d \leq 3$.

In some applications, however, the efficient numerical solution of parabolic problems in dimensions $d > 3$ is necessary. We mention here only the pricing of contracts on baskets of d assets, e.g., for an index where d can be as large as 50, and the Kolmogoroff equations for diffusions in high dimensions.

Here, the straightforward application of standard numerical schemes fails due to the so-called 'curse of dimension': the number of degrees of freedom on a tensor product Finite Element mesh of width h in dimension d grows like $O(h^{-d})$ as $h \to 0$. This observation has led to the belief that parabolic problems in dimension d larger than 3 can in effect not be solved by conventional, deterministic methods. Therefore Monte Carlo methods are used where the error decreases like $O(N^{-1/2})$ if one uses a work of N operations. This holds for any $d \geq 1$, but only in a probabilistic sense.

In this paper we describe a Finite Element algorithm for parabolic equations in high dimensions with an error of $O(N^{-p})$ for a work of N operations. Here p is the degree of the finite elements which can be any integer ≥ 1. The method is based on two observations:

(i) To reduce the number of degrees of freedom in high dimensions, so-called *sparse tensor product* Finite Element spaces are used (see, e.g. [2, 5, 7] and the references there). Their number of degrees of freedom grows like $O(h^{-1} |\log h|^{d-1})$ as $h \to 0$, instead of $O(h^{-d})$ for the full tensor product spaces. At the same time, the approximation rate in $H^1(\Omega)$ for elements of degree $p \geq 1$ and smooth functions is $O(h^p)$, the same as for full tensor product spaces. As we show in Proposition 3.2, this result requires more regularity than $H^{p+1}(\Omega)$ for the approximated function, and the amount of extra regularity increases with d. In the contract pricing problem mentioned above, the initial data u_0 of the problem (the pay-off function) is usually not smooth (typically $u_0 \in H^{3/2-\varepsilon}(\Omega)$ for $\varepsilon > 0$). However, the solution operator $E(t)$ of the parabolic problem is an analytic semigroup and increases the smoothness of the solution $u(\cdot, t)$ for $t > 0$. We prove that this parabolic smoothing effect suffices for optimal convergence of sparse space discretizations at $T > 0$ for any d, even for initial data that are just in $L^2(\Omega)$.

(ii) Even with a sparse space discretizations the number $\widehat{N}_L$ of spatial degrees of freedom is substantial if d is large. Reducing the number of time steps (and thus, the number of spatial problems to be solved) to pass from $t = 0$ to the final time T is therefore essential. Time analyticity of $E(t)$ implies analytic time regularity of the solution $u(t)$ for $t > 0$, but not uniformly in $(0, T)$. As was shown in [13], this allows to construct hp discontinuous Galerkin (DG) time-stepping schemes with *exponential convergence* in the number of spatial problems.

We analyze the fully discrete method with sparse tensor product Finite Elements of degree $p \geq 1$ and meshwidth h in space, and hp DG discretization in time. Because of the exponential convergence of the DG method in time it is sufficient to use $O(|\log h|)$ time intervals, and polynomial degree $O(|\log h|)$ in time. We then obtain at the final time T for $u(x, T)$ an L^2 error of $O(h^{\theta_0 p + \delta})$ where $\theta_0 \in (0, 1]$ is related to the regularity of the elliptic problem in Ω, and $\delta = p/((p + 1)d - 1)$. The case that $u(x, t)$ is smooth in x for all $t > 0$ corresponds to $\theta_0 = 1$.

For each DG time step we have to solve a linear system of size $(r + 1)\widehat{N}_L$. We can decouple this and obtain $r + 1$ linear systems of size $\widehat{N}_L$. Each of these $r + 1$ linear systems is of the same form as for the backward Euler method, but contains complex numbers. We solve these linear systems iteratively with GMRES and a wavelet preconditioner, and show that $O(|\log h|)$ iterations are sufficient.

The resulting algorithm requires $N = O(h^{-1} |\log h|^{2d+6})$ operations. In the case where $u(x, t)$ is smooth in x for all $t > 0$ (corresponding to $\theta_0 = 1$) we obtain that the L^2 error of $u(x, T)$ is bounded by $Ch^{p+\delta} \leq C' N^{-p}$.

Rather than covering the most general parabolic problems, we consider here the following model problem: In the d-dimensional unit cube $\Omega = (0,1)^d$, we consider

$$\frac{\partial u}{\partial t} + Au = g \quad \text{in } \ (0,T) \times \Omega\,, \tag{1.1}$$

$$u = 0 \quad \text{on } \ (0,T) \times \partial\Omega\,, \tag{1.2}$$

with the initial condition

$$u(0) = u_0 \ \text{ in } \ \Omega\,. \tag{1.3}$$

Here A is a second order elliptic differential operator in divergence form

$$Au = -\nabla \cdot \mathbf{D}(x)\,\nabla u + c(x)\,u \tag{1.4}$$

with coefficients $\mathbf{D} \in C^\infty(\overline{\Omega})^{d\times d}_{\text{sym}}$, $c \in C^\infty(\overline{\Omega})$ which are analytic in $\overline{\Omega}$ and satisfy, for all $x \in \overline{\Omega}$,

$$\forall \xi \in \mathbb{R}^d : \ \xi^\top \mathbf{D}(x)\xi \geq \gamma\,|\xi|^2\,, \tag{1.5}$$

$$c(x) \geq -\,\kappa > -\infty, \tag{1.6}$$

with constants $\gamma > 0$ and κ independent of x. We emphasize that A in (1.4) is self-adjoint only to reduce technicalities in the numerical analysis. Our algorithm works also for non self-adjoint operators A with first order terms and for time-dependent coefficients of the form $c(x,t) = c_1(x)c_2(t)$, $d_{ij}(x,t) = d_{ij,1}(x)d_{ij,2}(t)$; first order advection terms are admissible in all our results except in one nonsmooth data error estimate, whereas time-dependent coefficients would require minor modifications in the convergence proofs. The convergence rates and the complexity of our algorithm remain essentially unchanged. Finally, it is sufficient to assume in (1.4) that $c(x) \geq 0$, since the substitution $w = \exp(-\kappa t)u$ implies

$$w' + (A+\kappa)w = \exp(-\kappa t)g \text{ in } J \tag{1.7}$$

and $A + \kappa$ is of the form (1.4) with $\kappa = 0$ in (1.6).

The outline of the paper is as follows: in Section 2, we present an abstract parabolic framework, Section 3 is devoted to the space discretization by means of sparse tensor products of finite element spaces. Section 4 presents an hp-time stepping scheme for parabolic problems and exponential convergence results. Section 5 addresses the fully discrete approximation with hp-time stepping and sparse grids in space and the practical realization of the time-stepping scheme, in particular preconditioning and incomplete GRMES iterations for the linear systems of equations. Section 6 presents numerical results.

2 Abstract Parabolic Equations

For a variational formulation of (1.1)–(1.6) in $\Omega = (0,1)^d$, we require Sobolev spaces. By $H = L^2(\Omega)$ we denote the square integrable functions in Ω and by $H^s(\Omega)$, $s \geq 0$, the usual Sobolev-spaces; we need also $V = \{v \in H^1(\Omega)\colon\ v|_{\partial\Omega} = 0\}$. We identify H with its dual, $H = H^*$. Then

$$V \overset{d}{\hookrightarrow} H = L^2(\Omega) = H^* \overset{d}{\hookrightarrow} V^*\,. \tag{2.1}$$

with dense injection and the operator A in (1.4) is in $\mathcal{L}(V, V^*)$. We denote by $(\cdot,\cdot)_{V\times V^*}$ the extension of the H inner product $(\cdot,\cdot) : H \times H \to \mathbb{R}$ to $V \times V^*$, and denote by $\|\cdot\|$, $\|\cdot\|_V$, $\|\cdot\|_{V^*}$ the norms in H, V, V^*, respectively. With $A \in \mathcal{L}(V, V^*)$ we associate the bilinear form $a(\cdot,\cdot)\colon\ V \times V \to \mathbb{C}$ via

$$a(u,v) := (Au, v)_{V^*\times V},\ u, v \in V,\ Au \in V^*\,. \tag{2.2}$$

The form $a(\cdot,\cdot)$ is continuous,

$$\forall u, v \in V : \quad |a(u,v)| \leq \alpha\,\|u\|_V\,\|v\|_V \tag{2.3}$$

and coercive because of (1.5,1.6): there is $\beta > 0$ such that

$$\forall u \in V : \quad a(u,u) \geq \beta\,\|u\|_V^2 \tag{2.4}$$

for some $0 < \beta \le \alpha < \infty$.

Then $A \in \mathcal{L}(V, V^*)$ is an isomorphism and

$$\|A\|_{\mathcal{L}(V,V^*)} \le \alpha, \quad \|A^{-1}\|_{\mathcal{L}(V^*,V)} \le 1/\beta\,.$$

Then (1.1)–(1.6) is equivalent to the abstract ordinary differential equation: for $t \in J = (0, T)$,

$$u'(t) + Au = g \quad \text{in} \;\; V^*\,, \tag{2.5}$$

$$u(0) = u_0 \quad \text{in} \;\; H\,, \tag{2.6}$$

where $H = L^2(\Omega)$, $V = H_0^1(\Omega)$, $V^* = H^{-1}(\Omega) = (H_0^1(\Omega))^*$ and u' is understood in the weak sense, i.e. for $u \in L^2(J,V) \cap H^1(J,V^*)$ we have $u' \in L^2(J,V^*)$. The variational form of (2.5), (2.6) reads: given

$$u_0 \in H, \quad g \in L^2(J, V^*)\,, \tag{2.7}$$

find $u \in L^2(J,V) \cap H^1(J,V^*)$ such that $u(0) = u_0$ and

$$-\int_J (u(t), v)\, \varphi'(t) dt + \int_J a(u,v)\, \varphi(t) dt = \int_J (g(t), v)_{V^* \times V}\, \varphi\, dt \tag{2.8}$$

for every $v \in V$, $\varphi \in C_0^\infty(J)$. The initial condition is well-defined in H since (e.g. [6])

$$L^2(J;V) \cap H^1(J;V^*) \subset C^0(\overline{J};H) \tag{2.9}$$

Also from [6] we have that problem (2.7), (2.8) has a unique solution $u(t)$ and

$$\|u\|_{C(\overline{J};H)} + \|u\|_{L^2(J;V)} + \|u'\|_{L^2(J;V^*)} \le C(\|g\|_{L^2(J;V^*)} + \|u_0\|_H)\,. \tag{2.10}$$

We now introduce a scale H_s of Sobolev spaces adapted to the operator A in (1.4). By the spectral theorem, A admits a countable family of eigenpairs (λ_k, φ_k), with real eigenvalues $0 < \lambda_1 \le \lambda_2 \le \cdots \le \lambda_k \le \ldots$ accumulating only at infinity and eigenfunctions $\varphi_k \in V$. We can assume that the φ_k are orthonormal in H, i.e. $(\varphi_k, \varphi_l) = \delta_{kl}$. We then have Parseval's equation

$$\forall u \in H: \; \|u\|_H^2 = \sum_{k=1}^\infty |(u, \varphi_k)|^2 \tag{2.11}$$

and

$$u = \lim_{N\to\infty} \sum_{k=1}^N (u, \varphi_k)\, \varphi_k \;\; \text{in} \;\; H\,. \tag{2.12}$$

We now define, for $s \ge 0$, the scale of spaces H_s by

$$H_s = \Big\{ u \in L^2(\Omega):\; \sum_{k=1}^\infty (\lambda_k)^s |(u,\varphi_k)|^2 < \infty \Big\} \tag{2.13}$$

equipped with the norm

$$\|u\|_{H_s} = \Big(\sum_{k=1}^\infty (\lambda_k)^s\, |(u, \varphi_k)|^2 \Big)^{\frac{1}{2}}\,. \tag{2.14}$$

Furthermore we define for $s < 0$ the spaces $H_s := H_{-s}^*$ by duality. For $s > \frac{1}{2}$, let us denote by H_D^s the functions in $H^s(\Omega)$ with homogeneous Dirichlet condition:

$$H_D^s(\Omega) := \{\, v \in H^s(\Omega) \mid v|_\Gamma = 0 \,\}.$$

We then have the following characterization of H_s:

Proposition 2.1.

$$H_s = \begin{cases} H^s(\Omega) & \textit{for } 0 \le s < \frac{1}{2} \\ H_{00}^{1/2}(\Omega) & \textit{for } s = \frac{1}{2} \\ H_D^s(\Omega) & \textit{for } \frac{1}{2} < s \le 2 \end{cases}.$$

Proof. The case $s = 0$ follows from Parseval's identity. The case $s = 1$ follows from $\|u\|_{H_1}^2 = (Au, u)$ and (1.3), (1.4). Then we obtain the result for $0 \le s \le 1$ by interpolation. For $s = 2$ we use the fact that Ω is convex, hence $Au \in L^2(\Omega)$ implies $u \in H^2(\Omega)$. Interpolation then gives the range $1 \le s \le 2$. ☐

Remark 2.2. There is some $s_0 > 2$ such that for $s > s_0$ the spaces H_s and $H_D^s(\Omega)$ will be different: The proof of Lemma 3.1 in [17] gives that e.g.

$$H_3 = A^{-1}H_1 = A^{-1}H_D^1 = \{\, v \mid Av \in H^1(\Omega), Av|_\Gamma = 0, v|_\Gamma = 0 \,\}$$

We see that the space H_3 has an additional boundary condition $Av|_\Gamma = 0$. On the other hand there may be functions $v \notin H^3(\Omega)$ which satisfy $Av \in H^1(\Omega)$ because of singularities in the solution at the points where the boundary is not smooth.

The regularity (2.7) of the data u_0, g is sufficient for existence. In order to prove convergence rates for discretizations we will require slightly higher regularity, namely $u_0 \in H_{\varepsilon_1}$ and $g \in L^\infty(J; H_{-1+\varepsilon_2})$ with $\varepsilon_1, \varepsilon_2 > 0$.

The problem (1.1)–(1.3) with $g = 0$ admits the solution $u(t) = E(t)u_0$ where the evolution operator

$$E(t)u_0 := \sum_{k=1}^{\infty} e^{-\lambda_k t}(u_0, \varphi_k)\varphi_k$$

satisfies the following estimates which can be verified directly [14]

Proposition 2.3. *There are C, $\widehat{d} > 0$ such that for $t > 0$, $l \in \mathbb{N}_0$, $\tau \ge \theta \ge -1$ holds*

$$\|E^{(l)}(t)\|_{\mathcal{L}(H_\theta, H_\tau)}^2 \le C\widehat{d}^{2l+\tau-\theta}\Gamma(2l+1+\tau-\theta)t^{-(2l+\tau-\theta)}. \tag{2.15}$$

For $s \ge \sigma \ge 0$ there holds

$$\|E(t)u_0\|_{H_s} \le Ct^{-(s-\sigma)/2}\|u_0\|_{H_\sigma}$$

3 Space Discretization

3.1 Wavelets in $\mathbb{R}$

In the interval $I = (0, 1)$, we define the mesh $\mathcal{T}^\ell$ given by the nodes $j2^{-\ell}$, $j = 0, \ldots, 2^\ell$, with the meshwidth $h_\ell = 2^{-\ell}$. We define $\mathcal{V}^\ell$ as the space of piecewise polynomials of degree $p \ge 1$ on the mesh $\mathcal{T}^\ell$ which are in $C^{p'-1}([0, 1])$ with $1 \le p' \le p$ and vanish at the endpoints 0, 1. We write $N^\ell = \dim \mathcal{V}^\ell$, $M^\ell := N^\ell - N^{\ell-1}$, $N^{-1} := 0$; then $N^\ell = O(2^\ell)$, $\ell = 0, 1, 2, \ldots$. We employ a wavelet basis ψ_j^ℓ, $j = 1, \ldots, M^\ell$, $\ell = 0, 1, 2, \ldots$ of $\mathcal{V}^\ell$ with the properties:

$$\mathcal{V}^\ell = \mathrm{span}\{\psi_j^\ell \,|\, 0 \le \ell \le L;\ 1 \le j \le M^\ell\}, \quad \mathrm{diam}\,(\mathrm{supp}\ \psi_j^\ell) \le C\,2^{-\ell}\,. \tag{3.1}$$

Any function $v \in \mathcal{V}^L$ has the representation

$$v = \sum_{\ell=0}^{L} \sum_{j=1}^{M^\ell} v_j^\ell\, \psi_j^\ell \tag{3.2}$$

with $v_j^\ell = (v, \widetilde{\psi}_j^\ell)$ where $\widetilde{\psi}_j^\ell$ are the so-called dual wavelets. For $v \in V$ one obtains the series

$$v = \sum_{\ell=0}^{\infty} \sum_{j=1}^{M^\ell} v_j^\ell\, \psi_j^\ell \tag{3.3}$$

which converges in $L^2(I)$ and in $H_0^1(I)$. Moreover, there holds the **norm equivalence**

$$c_1\|v\|_{H_\theta}^2 \le \sum_{\ell=0}^{\infty} \sum_{j=1}^{M^\ell} |v_j^\ell|^2\, 2^{2\ell\theta} \le c_2\|v\|_{H_\theta}^2,\ 0 \le \theta \le 1\,. \tag{3.4}$$

For $v \in L^2(I)$ we can define a projection $P_L\colon L^2(I) \to \mathcal{V}^L$ by truncating (3.3):

$$P_L v := \sum_{\ell=0}^{L} \sum_{j=1}^{M^\ell} v_j^\ell \, \psi_j^\ell \,, \; P_{-1} := 0 \,. \tag{3.5}$$

This projection satisfies the **approximation property**

$$\|u - P_L u\|_{H_\theta} \le c\, 2^{-(t-\theta)L} \|u\|_{H^t(I)}, \; 0 \le \theta \le 1, \; \theta \le t \le p+1 \,. \tag{3.6}$$

The increment or detail spaces $\mathcal{W}^\ell$ are defined by

$$\begin{cases} \mathcal{W}^\ell := & \operatorname{span}\{\psi_j^\ell : 1 \le j \le M^\ell\}, \; \ell = 1, 2, 3, \ldots \\ \mathcal{W}^0 := & \mathcal{V}^0 \,. \end{cases} \tag{3.7}$$

Then

$$\mathcal{V}^\ell = \mathcal{V}^{\ell-1} \oplus \mathcal{W}^\ell \text{ for } \ell \ge 1, \quad \text{and } \; \mathcal{V}^\ell = \mathcal{W}^0 \oplus \cdots \oplus \mathcal{W}^\ell, \ell \ge 0 \,. \tag{3.8}$$

and $Q_\ell := P_\ell - P_{\ell-1}$ is a projection from $L^2(I)$ to $\mathcal{W}^\ell$.

3.2 Examples of wavelets

We give an example for $p = p' = 1$, i.e., for piecewise linear continuous functions on $[0,1]$ vanishing at the endpoints $0, 1$. Since there is no nonzero function on the whole interval $[0,1]$ we now define the mesh $\mathcal{T}^\ell$ for $\ell \ge 0$ by the nodes $x_j^\ell := j2^{-\ell-1}$ with $j = 0, \ldots, 2^{\ell+1}$. We have $N_\ell = 2^{\ell+1} - 1$ and $M_\ell = 2^\ell$.

We define the wavelets ψ_j^ℓ for level $\ell = 0, 1, 2, \ldots$, $j = 1, \ldots, M_\ell$: for $\ell = 0$ we have $N_0 = M_0 = 1$ and ψ_1^0 is the function with value c_0 at $x_1^0 = \frac{1}{2}$.

For $\ell \ge 1$ we have $N_\ell = 2^{\ell+1} - 1$. and we let $c_\ell := 2^{-\ell/2}, \ell = 0, 1, \ldots.$ Then the wavelet ψ_1^ℓ has values $\psi_1^\ell(x_1^\ell) = 2c_\ell$, $\psi_1^\ell(x_2^\ell) = -c_\ell$ and zero at all other nodes. The wavelet $\psi_{M_\ell}^\ell$ has values $\psi_{M_\ell}^\ell(x_{N_\ell}^\ell) = 2c_\ell$, $\psi_{M_\ell}^\ell(x_{N_\ell-1}^\ell) = -c_\ell$ and zero at all other nodes. The wavelet ψ_j^ℓ with $1 < j < M_\ell$ has values $\psi_j^\ell(x_{2j-2}^\ell) = -c_\ell$, $\psi_j^\ell(x_{2j-1}^\ell) = 2c_\ell$, $\psi_j^\ell(x_{2j}^\ell) = -c_\ell$ and zero at all other nodes.

3.3 Sparse tensor product spaces and approximation rates

In $\Omega = I^d = (0,1)^d$, $d > 1$ we define the subspace V^L as the tensor product of the one-dimensional spaces:

$$V^L := \mathcal{V}^L \otimes \cdots \otimes \mathcal{V}^L \tag{3.9}$$

which can be written using (3.8) as

$$V^L = \sum_{0 \le \ell_i \le L} \mathcal{W}^{\ell_1} \otimes \cdots \otimes \mathcal{W}^{\ell_d}. \tag{3.10}$$

The space V^L has $O(2^{\ell d})$ degrees of freedom and is too costly if d is large. We shall use the sparse tensor product space

$$\begin{aligned} \widehat{V}^L \quad &:= \operatorname{span}\Big\{ \psi_{j_1}^{\ell_1}(x_1) \ldots \psi_{j_d}^{\ell_d}(x_d) \mid 1 \le j_i \le M^{\ell_i}, \ell_1 + \cdots + \ell_d \le L \Big\} \\ &= \sum_{0 \le \ell_1 + \cdots + \ell_d \le L} \mathcal{W}^{\ell_1} \otimes \cdots \otimes \mathcal{W}^{\ell_d}. \end{aligned} \tag{3.11}$$

As $L \to \infty$, we have $N_L := \dim(V^L) = O(2^{dL})$, and $\widehat{N}_L := \dim(\widehat{V}^L) = O(L^{d-1}\, 2^L)$, i.e. the spaces $\widehat{V}^L$ have considerably smaller dimension than V^L. On the other hand, they do have similar approximation properties as V^L, provided the function to be approximated is sufficiently smooth: To characterize the smoothness we introduce the spaces $\mathcal{H}^k$ with square integrable mixed k-th derivatives: Let $\mathcal{H}^0 := L^2(\Omega)$, and define for integer $k \ge 1$

$$\mathcal{H}^k := \{\, u \in H_0^1(\Omega) \mid \mathcal{D}^\alpha u \in L^2(\Omega), \; 0 \le \alpha_i \le k \,\} \tag{3.12}$$

equipped with the norm

$$\|u\|_{\mathcal{H}^k} := \Big(\sum_{\substack{0\le\alpha_i\le k\\1\le i\le d}} \|\mathcal{D}^\alpha u\|^2_{L^2(\Omega)} \Big)^{\frac{1}{2}}. \tag{3.13}$$

We then define $\mathcal{H}^s$ for arbitrary $s \geq 0$ by interpolation.

For a function $v \in L^2(\Omega)$ we have as a consequence of (3.3), (3.9)

$$v(x) = \sum_{\ell_1,\dots,\ell_d\ge 0} \; \sum_{1\le j_k\le n_{\ell_k}} v^{\ell_1\dots\ell_d}_{j_1\dots j_d}\, \psi^{\ell_1}_{j_1}(x_1)\dots\psi^{\ell_d}_{jd}(x_d). \tag{3.14}$$

We then define the sparse projection operator $\widehat{P}^L \colon L^2(\Omega) \to \widehat{V}^L$ by truncating the wavelet expansion:

$$(\widehat{P}_L v)(x) := \sum_{\substack{0\le \ell_1+\cdots+\ell_d\le L\\ 1\le j_k\le n_{\ell_k},k=1,\dots,d}} v^{\ell_1\dots\ell_d}_{j_1\dots j_d}\, \psi^{\ell_1}_{j_1}(x_1)\dots\psi^{\ell_d}_{jd}(x_d), \tag{3.15}$$

$$\widehat{P}_L = \sum_{0\le\ell_1+\cdots+\ell_d\le L} Q_{\ell_1}\otimes\cdots\otimes Q_{\ell_d}. \tag{3.16}$$

We next establish some properties of the sparse grid projection $\widehat{P}_L \colon V \to \widehat{V}^L$.

Proposition 3.1. *(Stability of $\widehat{P}_L$) For $0 \le \theta \le 1$ and $v \in H_\theta$ we have*

$$\|\widehat{P}_L v\|_{H_\theta} \le C\,\|v\|_{H_\theta}\,. \tag{3.17}$$

Proof. For $\theta = 0$, we have with

$$|||v|||_0^2 := \sum_{\substack{\ell_k=0\\k=1,\dots,d}}^{\infty} \; \sum_{1\le j_k\le n_{\ell_k}} \left| v^{\ell_1\dots\ell_d}_{j_1\dots j_d}\right|^2 \tag{3.18}$$

that

$$\|\widehat{P}_L v\|_{H_0} \le C_1 \left|\left|\left| \widehat{P}_L v \right|\right|\right|_0 \le C_1\,|||v|||_0 \le C_2\,\|v\|_{L^2(\Omega)}\,.$$

We also have from the norm equivalence (3.4) that for every $v \in H^1_0(\Omega)$:

$$\|v\|^2_{H_1} \le C_3 \sum_{\substack{\ell_k=0\\k=1,\dots,d}}^{\infty} \; \sum_{1\le j_k\le n_{\ell_k}} \left| v^{\ell_1\dots\ell_d}_{j_1\dots j_d}\right|^2 (1+2^{\ell_1}+\cdots+2^{\ell_d}) =: |||v|||_1^2\,. \tag{3.19}$$

It follows

$$\|\widehat{P}_L v\|_{H^1} \le C_3 \left|\left|\left| \widehat{P}_L v \right|\right|\right|_1 \le C_3\,|||v|||_1 \le C_4 \|v\|_{H^1}\,.$$

Interpolation gives (3.17). $\square$

Proposition 3.2. *(Approximation property of $\widehat{P}_L$) Assume that the component spaces $\mathcal{V}^\ell$ of $\widehat{V}^L$ have the approximation property* (3.6)*. Then for $0 \le s < p' + \frac{1}{2}$ and $s < t \le p+1$*

$$\|u - \widehat{P}_L u\|_{H^s(\Omega)} \le \begin{cases} Ch^{p+1}\,|\log h|^{(d-1)/2}\,\|u\|_{\mathcal{H}^{p+1}} & \text{if } s=0 \text{ and } t=p+1 \\ Ch^{t-s}\|u\|_{\mathcal{H}^t} & \text{otherwise} \end{cases} \tag{3.20}$$

Proof. We follow [5, Prop. 6] and [7]. We first consider the case $0 \le t < p+1$. In the one-dimensional case, for $0 \le t < p+1$, we have

$$\sum_{\ell=0\dots\infty} 2^{2t\ell}\,\|Q_\ell u\|^2 \le C\,\|u\|^2_{H^t(I)}\,. \tag{3.21}$$

Using tensor product arguments (cf. e.g. [7]) we obtain that, for $u \in \mathcal{H}^t$ and $0 \le t < p+1$,

$$\sum_{\ell_1,\dots,\ell_d=0}^{\infty} 2^{2t(\ell_1+\cdots+\ell_d)}\,\|Q_{\ell_1}\otimes\cdots\otimes Q_{\ell_d} u\|^2 \le C\,\|u\|^2_{\mathcal{H}^t}\,.$$

As in [5] this implies, for $t < p+1$ and $\ell := (\ell_1, \dots, \ell_d)$,

$$\|u - \widehat{P}_L u\|^2_{H^s(\Omega)} \le C \|u\|^2_{\mathcal{H}^t} \max_{|\ell|>L} 2^{2s|\ell|_\infty - 2t|\ell|} \le C' \|u\|^2_{\mathcal{H}^t} C' 2^{2(s-t)L}$$

as the maximum is attained at e.g. $\ell = (L+1, 0, \dots, 0)$.

In the case of $t = p+1$ we have in the one-dimensional case instead of (3.21) only $2^{t\ell} \|Q_\ell u\| \le C \|u\|_{H^t(I)}$, and one obtains with tensor product arguments as in [5]

$$\|Q_{\ell_1} \otimes \cdots \otimes Q_{\ell_d} u\| \le C 2^{-|\ell| t} \|u\|_{\mathcal{H}^t}$$

and from that

$$\|u - \widehat{P}_L u\|^2_{H^s(\Omega)} \le C \|u\|^2_{\mathcal{H}^t} \sum_{|\ell|>L} 2^{2s|\ell|_\infty - 2t|\ell|} = C \|u\|^2_{\mathcal{H}^t} 2^{2(s-t)L} \sum_{m=L+1}^{\infty} 2^{2(s-t)(m-L)} A_m$$

with $A_m := \sum_{|\ell|=m} 2^{2s(|\ell|_\infty - m)}$. For $s = 0$ we have $A_m \le C m^{d-1}$ whereas, for $s > 0$, $A_m \le C$ holds. □

Remark 3.3. This result and the proof also apply for discontinuous wavelets with $p \ge 0$, $p' = 0$, e.g. the Haar wavelets with $p = p' = 0$. The case of the Haar wavelets also illustrates that the logarithmic term is necessary in the case of $s = 0$ and $t = p+1$: Consider the function $f(x) = x_1 \cdots x_d$. The wavelet coefficients of f satisfy $f^{l_1 \dots l_d}_{j_1 \dots j_d} = c 2^{-3/2(\ell_1 + \dots + \ell_d)}$ and therefore

$$\left\|f - \widehat{P}_L f\right\|^2 = \sum_{\ell_1 + \cdots + \ell_d > L} \sum_{j_1, \dots, j_d} \left|f^{l_1 \dots l_d}_{j_1 \dots j_d}\right|^2 = \sum_{\ell_1 + \cdots + \ell_d > L} c^2 2^{-2(\ell_1 + \cdots + \ell_d)}$$

Now we see that already $\sum_{\ell_1 + \cdots + \ell_d = L+1} 2^{-2(\ell_1 + \cdots + \ell_d)} \ge C 2^{-2L} L^{d-1}$ which shows that the convergence rate in (3.20) is sharp. For the case $d = 2$ see also [7].

Remark 3.4. Sparse grid spaces based on interpolation (see e.g. [2]) do not exhibit semiorthogonality like our wavelet based sparse spaces. For such interpolation-based sparse grid spaces there is an additional logarithmic term in the H^1 approximation rate.

Remark 3.5. We can express the convergence rates (3.20) in terms of the number of degrees of freedom using $\widehat{N}_L = O(h^{-1} |\log h|^{d-1})$, yielding a bound $O(\widehat{N}_L^{t-s} (\log \widehat{N}_L)^\beta)$ with β depending on d. It was shown in [4] for $p = 1$ that one can avoid this logarithmic term by using smaller spaces than $\widehat{V}^L$ and higher regularity of u than $\mathcal{H}^{p+1}$.

3.4 Regularity of parabolic problems and sparse approximation rates

To characterize the approximation properties of $\widehat{V}^L$ we define the scale of interpolation spaces

$$X_{\theta,p} := (H^1_0(\Omega), \mathcal{H}^{p+1})_{\theta,2} \quad 0 \le \theta \le 1 \,. \tag{3.22}$$

We have $\mathcal{H}^{1+\theta p} \subset X_{\theta,p} \subset H^{1+\theta p}_D$ for $0 \le \theta \le 1$ where the inclusions are strict for $0 < \theta < 1$. From $H^1_D = X_{0,p}$ and $H^{d(p+1)}_D \subset \mathcal{H}^{p+1} = X_{1,p}$ we obtain with interpolation that

$$H^{1-\theta+\theta d(p+1)}_D \subset X_{\theta,p} \tag{3.23}$$

for $0 \le \theta \le 1$.

For $u \in H^1_0(\Omega)$ we have from Proposition 3.1 that $\left\|u - \widehat{P}_L u\right\|_{H^1(\Omega)} \le C \|u\|_{H^1(\Omega)}$. Interpolating the bound with (3.20) for $s = 1$, $t = p+1$ gives for $u \in X_{\theta,p}$

$$\left\|u - \widehat{P}_L u\right\|_{H^1(\Omega)} \le C h^{\theta p} \|u\|_{X_{\theta,p}} \tag{3.24}$$

Now let $u(t)$ be the solution of the parabolic problem (1.1)–(1.3) with $u_0 \in L^2(\Omega)$ and $g = 0$. To estimate the approximation rate $\left\| u(t_0) - \widehat{P}_L u(t_0) \right\|_{H^1(\Omega)}$ for $t_0 > 0$ we use that $u(t_0) \in H_s$ for any $s > 0$. For a smooth domain Ω we would have

$$H_{(p+1)d} \subset H^{(p+1)d} \subset \mathcal{H}^{p+1} = X_{\theta,p} \text{ with } \theta = 1.$$

But this is in general not true for the domain $\Omega = (0,1)^d$ as the boundary is not smooth: since $H_s = A^{-s/2}L^2(\Omega)$ this space can contain functions which are not in $H^s(\Omega)$ for $s > 2$: there may exist singular functions $v \notin H^s(\Omega)$ in H_s such that $Av \in H^{s-2}(\Omega)$. Since Ω is convex we always have $H_2 = A^{-1}L^2(\Omega) = H_D^2(\Omega) \subset X_{\theta,p}$ for $\theta = \big(d(p+1)-1\big)^{-1}$ by (3.23).

Hence there always exists some $\theta_0 \in (0,1]$ such that

$$H_{(p+1)d} \subset X_{\theta_0,p}. \tag{3.25}$$

which depends on the singularity functions for the operator A at the singular points of $\partial\Omega$. If $\theta_0 < 1$, we expect the reduced convergence rate $h^{\theta_0 p}$ instead of h^p.

Remark 3.6. In certain cases, we obtain in (3.24) the full approximation rate h^p. Let us consider $A = -\Delta u + cu$ in $\Omega = (0,1)^d$ where c is smooth on $\overline{\Omega}$. We denote by $T = \mathbb{R}/(2\mathbb{Z})$ the interval $[-1,1]$ with the boundary points identified. A function v on I^d can be extended to an even function v^e on T^d by $v^e(x) := v((|x_1|,\dots,|x_d|))$. Similarly we can define the antisymmetric extension v^o by $v^o(x) := \operatorname{sign}(x_1)\cdots\operatorname{sign}(x_d)v((|x_1|,\dots,|x_d|))$. We now assume that c^e is smooth on T^d. Then we may assume without loss of generality that $c \geq c_0 > 0$ because of (2.13). We define the operator $\widehat{A}^{-1}$ as the solution operator of the problem $-\Delta u + c^e u = f$ on T^d. For $f \in L^2(I^d)$ and $u = A^{-1}f$, we then have $u^o = \widehat{A}^{-1}f^o$. Therefore,

$$H_s = \left\{ \left(\widehat{A}^{-s/2}f^o\right)\Big|_{I^d} \;\Big|\; f \in L^2(I^d) \right\} \subset \left\{ \left(\widehat{A}^{-s/2}f\right)\Big|_{I^d} \;\Big|\; f \in L^2(T^d) \right\} = H^s(I^d).$$

Hence $\theta_0 = 1$ in (3.25), under the assumption that c^e is smooth. We can relax this assumption. Since the solution of an elliptic boundary value problem is smooth at points where the boundary is smooth we only have to assume that c^e is smooth in a small neighborhood of the singular part of the boundary.

Remark 3.7. If the principal part of A is different from $-\Delta$ (even if A has constant coefficients) we can no longer expect that $\theta_0 = 1$, and we will get a lower approximation rate $O(h^{\theta_0 p})$. But this also happens for the full grid space V^L: if the function u has only regularity $u \in H^{1+s_0-\varepsilon}(\Omega)$ with $s_0 < p$ then we obtain only the lower approximation rate $\|u - P_L u\|_{H^1(\Omega)} \leq C h^{s_0-\varepsilon}$ instead of $O(h^p)$.

3.5 Approximation of the elliptic problem

In $\Omega = I^d$ consider for the operator A in (1.4) the problem

$$Au = f \text{ in } \Omega,\; u|_{\partial\Omega} = 0 \tag{3.26}$$

associated with (1.1). In weak form:

$$u \in H_0^1(\Omega):\; a(u,v) = (f,v) \quad \forall v \in H_0^1(\Omega). \tag{3.27}$$

The corresponding solution operator T is continuous, i.e. $u = Tf : H_{-1} \to H_1$ boundedly.

Let $\widehat{u}^L = \widehat{R}_L u \in \widehat{V}^L$ be the sparse Galerkin approximation of u, defined by

$$\widehat{u}^L \in \widehat{V}^L:\; a(\widehat{u}^L, \widehat{v}^L) = (f, \widehat{v}^L) \quad \forall \widehat{v}^L \in \widehat{V}^L \tag{3.28}$$

and denote by $\widehat{T}_L f = \widehat{R}_L u$ the approximate solution operator. We have for $0 \leq \theta \leq 1$, using (3.24),

$$\|u - \widehat{u}^L\|_{H^1(\Omega)} \leq C\, h^{\theta p} \|u\|_{X_{\theta,p}}. \tag{3.29}$$

Now a standard duality argument gives

$$\begin{aligned}\|u-\widehat{u}^L\| &\le \sup_{v\in L^2(\Omega)} \frac{(u-\widehat{u}^L,v)}{\|v\|} = \sup_{v\in L^2(\Omega)} \frac{a(u-\widehat{u}^L, w_v-\widehat{P}_L w_v)}{\|v\|}\\ &\le C\,\|u-\widehat{u}^L\|_{H^1(\Omega)} \sup_{v\in L^2(\Omega)} \frac{\|w_v-\widehat{P}_L w_v\|_{H^1(\Omega)}}{\|v\|},\end{aligned} \tag{3.30}$$

where w_v denotes the solution of

$$A^*\,w_v = v \text{ in } \Omega,\ \ w_v|_{\partial\Omega} = 0\,. \tag{3.31}$$

Since Ω is convex and $c\in C^\infty(\overline{\Omega})$, we have $w_v\in H^2(\Omega)$ and

$$\|w_v\|_{H^2(\Omega)} \le C\,\|v\|_{L^2(\Omega)}. \tag{3.32}$$

Using (3.24) and (3.23) we obtain

$$\|w_v-\widehat{P}_L w_v\|_{H^1(\Omega)} \le C\,h^{\theta' p}\,\|w_v\|_{H^{1-\theta'+\theta'(p+1)d}}\,, \quad 0\le\theta'\le 1\,. \tag{3.33}$$

With $\theta' := 1/[(p+1)d-1]$ we get from (3.33) and (3.32) that $\|w_v-\widehat{w}_v^L\|_{H^1(\Omega)} \le Ch^\delta$ where

$$\delta := \frac{p}{(p+1)d-1}. \tag{3.34}$$

Now (3.30) and (3.29) give for $\widehat{u}^L = \widehat{R}_L u$ the convergence rate

$$\left\|u-\widehat{R}_L u\right\| \le C\,h^{\theta p+\delta}\,\|u\|_{X_{\theta,p}}\,. \tag{3.35}$$

Remark 3.8. For the Galerkin approximation of the elliptic problem (3.27) with a function u^L in the full space V^L we obtain the approximation rates $\|u-u^L\|_{H^1(\Omega)} \le Ch^p\|u\|_{H^{p+1}(\Omega)}$ and $\|u-u^L\|_{L^2(\Omega)} \le Ch^{p+1}\|u\|_{H^{p+1}(\Omega)}$. The Galerkin approximation $\widehat{u}^L$ on the sparse grid gives the same convergence rate in the H^1-norm as the full grid approximation. For the L^2-error, however, we obtain for the sparse grid solution only the lower rate $O(\widehat{h}^{p+\delta})$ compared with the full grid approximation. The reason for this is that the H^2-regularity (3.32) of the adjoint problem only yields a rate h^δ on the sparse grid but h^1 on the full grid.

3.6 Spatial semidiscretization

We semidiscretize (2.8) in space: We choose an approximation for the initial value

$$\widehat{u}_0^L = \widehat{P}_L\,u_0\,. \tag{3.36}$$

Then the **solution $\widehat{u}^L$ of the spatially semidiscrete problem** is defined using a Galerkin approximation in space: Find $\widehat{u}^L(t) : J\to\widehat{V}^L$ such that $\widehat{u}^L(0)=\widehat{u}_0^L$ and such that

$$\left(\frac{d}{dt}\,\widehat{u}^L, v^L\right) + a(\widehat{u}^L, v^L) = (g(t), v^L) \quad \forall v^L\in\widehat{V}^L\,. \tag{3.37}$$

We first consider the homogeneous equation with $g(t)=0$. In the case of smooth initial data we have the following result.

Theorem 3.9. *Assume that* (3.25) *holds with* $0<\theta_0\le 1$. *Consider* (2.8), (3.37) *with* $g=0$, $p\ge 1$ *and assume that* $u_0\in H_{(p+1)d}$. *Then*

$$\|u(t)-\widehat{u}^L(t)\| \le C\,h^{\theta_0 p+\delta}\,\|u_0\|_{H_{(p+1)d}}\,. \tag{3.38}$$

Proof. The proof follows [17], Theorem 3.1. We use that by (3.35)

$$\|(T-\widehat{T}_L)f\| \le C\,h^{\theta_0 p+\delta}\|Tf\|_{X_{\theta_0,p}}\,, \tag{3.39}$$

and first consider instead of (3.36) the initial value $\widehat{u}_0^L = \widehat{P}_L^0 u_0$, with $\widehat{P}_L^0$ the $L^2(\Omega)$-projection onto $\widehat{V}^L$. Since $\|\widehat{u}^L(t)\| \le \|\widehat{u}^L(0)\|$ this causes an error contribution to (3.38) which can be estimated by

$$\left\|\widehat{P}_L^0 u_0 - \widehat{P}_L u_0\right\| \le 2\left\|\widehat{P}_L u_0 - u_0\right\| \le Ch^{p+1}\left|\log h\right|^{(d-1)/2}\|u_0\|_{\mathcal{H}^{p+1}} \tag{3.40}$$

and, using (3.20) with $s = 0$, $t = 1/d$ and $H^1(\Omega) \subset \mathcal{H}^{1/d}$ we get

$$\left\|\widehat{P}_L^0 u_0 - \widehat{P}_L u_0\right\| \le 2\left\|\widehat{P}_L u_0 - u_0\right\| \le Ch^{1/d}\|u_0\|_{\mathcal{H}^{1/d}} \le Ch^{1/d}\|u_0\|_{H^1(\Omega)}. \tag{3.41}$$

Interpolating between (3.40) and (3.41), we obtain

$$\left\|\widehat{P}_L^0 u_0 - \widehat{P}_L u_0\right\| \le Ch^{\theta_0 p + \theta_0(1-1/d)+1/d}\left|\log h\right|^{\theta_0(d-1)/2}\|u_0\|_{X_{\theta_0,p}}.$$

This implies

$$\left\|\widehat{P}_L^0 u_0 - \widehat{P}_L u_0\right\| \le Ch^{\theta_0 p+\delta}\|u_0\|_{X_{\theta_0,p}} \le Ch^{\theta_0 p+\delta}\|u_0\|_{H_{(p+1)d}}$$

since $\delta \le 1/d$ for $d \ge 1$.

The error $e(t) = u(t) - \widehat{u}^L(t)$ satisfies with $\rho = -(\widehat{T}_L - T)Au = (T - \widehat{T}_L)u_t$

$$\widehat{T}_L\, e_t + e = \rho, \quad \widehat{T}_L\, e(0) = 0. \tag{3.42}$$

Lemma 3.4 in [17] states that

$$\|e(t)\| \le C \sup_{s\le t}(s\|\rho_t(s)\| + \|\rho(s)\|). \tag{3.43}$$

The assertion then follows from

$$\|\rho(s)\| \le Ch^{\theta_0 p+\delta}\|u(s)\|_{X_{\theta_0,p}} \le Ch^{\theta_0 p+\delta}\|u(s)\|_{H_{(p+1)d}} \le C\,h^{\theta_0 p+\delta}\,\|u(0)\|_{H_{(p+1)d}}$$

and

$$\begin{aligned} s\|\rho_t(s)\| &\le s\,C\,h^{\theta_0 p+\delta}\|u_t(s)\|_{X_{\theta_0,p}} \le s\,C\,h^{\theta_0 p+\delta}\,\|u_t(s)\|_{H_{(p+1)d}} \\ &= s\,C\,h^{\theta_0 p+\delta}\|Au(s)\|_{H_{(p+1)d}} \le s\,C\,h^{\theta_0 p+\delta}\,s^{-1}\|u(0)\|_{H_{(p+1)d}}. \end{aligned}$$

□

Estimate (3.38) assumed that $u_0 \in H_{(p+1)d}$. Note that this not only requires smoothness in the interior, but also that u_0 satisfies the compatibility $u_0 = 0, Au_0 = 0, \ldots, A^k u_0 = 0$ on $\partial\Omega$ for some integer k (see Remark 2.2).

Next we prove an error bound valid for $u_0 \in L^2(\Omega)$.

Theorem 3.10. *Assume that* (3.25) *holds with* $0 < \theta_0 \le 1$. *Consider* (2.8), (3.37) *with* $p \ge 1$, $g \equiv 0$ *and* $u_0^L = \widehat{P}_L\, u_0$. *Then there is* $C > 0$ *such that, for any* $t > 0$,

$$\|u(t) - \widehat{u}^L(t)\| \le C\,h^{\theta_0 p+\delta}\,t^{-(p+1)d/2}\,\|u_0\|. \tag{3.44}$$

Proof. We follow the proof of [17], Theorem 3.2. Using (3.35) with $\theta = 0$ and Proposition 2.3 we obtain

$$\|\widehat{R}_L u(t) - u(t)\| \le C\,h^{\delta}\,\|u(t)\|_{H^1(\Omega)} \le C\,h^{\delta}\,t^{\frac{1}{2}}\,\|u_0\|.$$

yielding the following bound for $e(t) = u(t) - \widehat{u}^L(t)$:

$$\|e(t)\| \le C\,h^{\delta}\,t^{-\frac{1}{2}}\|u_0\|. \tag{3.45}$$

We define the error operators $\widehat{F}_L(t)$ by

$$e(t) = \widehat{F}_L(t)\,u_0^L = \widehat{E}_L(t)\,\widehat{P}_L\,u_0 - E(t)\,\widehat{P}_L\,u_0 = \widehat{u}_L(t) - u(t)\,, \tag{3.46}$$

where $\widehat{E}_L(t)$ denotes the solution operator of the semidiscrete problem with space $\widehat{V}^L$. With $\widehat{F}_L(t)$, the claim (3.44) may be rewritten as

$$\|\widehat{F}_L(t)\, u_0^L\| \le C\, h^{\theta_0 p+\delta}\, t^{-(p+1)d/2}\, \|u_0\|\,. \tag{3.47}$$

We have [17], p. 42:

$$\widehat{F}_L(t) = \widehat{F}_L(\tfrac{t}{2})\, E(\tfrac{t}{2}) + E(\tfrac{t}{2})\, \widehat{F}_L(\tfrac{t}{2}) + \widehat{F}_L(\tfrac{t}{2})^2\,.$$

Furthermore, using Theorem 3.9,

$$\big\|\widehat{F}_L(\tfrac{t}{2})\, E(\tfrac{t}{2})\, u_0^L\big\| \le C\, h^{\theta_0 p+\delta}\, \big\|E(\tfrac{t}{2})\, u_0^L\big\|_{H_{(p+1)d}} \le C\, h^{\theta_0 p+\delta}\, t^{-(p+1)d/2}\, \|u_0^L\|\,.$$

Since $A = A^*$ we have $(E(\frac{t}{2})\, \widehat{F}_L(\frac{t}{2}))^* = \widehat{F}_L(\frac{t}{2})\, E(\frac{t}{2})$ and it follows that

$$\big\|E(\tfrac{t}{2})\, \widehat{F}_L(\tfrac{t}{2})\, u_0^L\big\| \le\ C\, h^{\theta_0 p+\delta}\, t^{-(p+1)d/2}\, \|u_0^L\|\,,$$

and altogether

$$\|\widehat{F}_L(t)\, u_0^L\| \le\ C\, h^{\theta_0 p+\delta}\, t^{-(p+1)d/2}\, \|u_0^L\| + C\, h^{\delta}\, t^{-\frac{1}{2}}\, \big\|\widehat{F}_L(\tfrac{t}{2})\, u_0\big\|\,.$$

Iteration gives, for any integer $s \ge 1$,

$$\|\widehat{F}_L(t)\, u_0^L\| \le C\, h^{\theta_0 p+\delta}\, t^{-(p+1)d/2}\, \|u_0\| + C\, (h^{\delta}\, t^{-\frac{1}{2}})^s\, \big\|\widehat{F}_L(\tfrac{t}{2^s})\, u_0\big\|\,.$$

We choose s such that $\delta s \ge \theta_0 p + \delta$, and we find, using $\|\widehat{F}_L(\frac{t}{2^s})v\| \le 2\|v\|$, that $\|\widehat{F}_L(t)u_0^L\| \le C\, h^{\theta_0 p+\delta}\, t^{-(p+1)d/2}\|u_0\|$ which completes the proof. ☐

3.7 Inhomogeneous Problems

We now consider the inhomogeneous problem (2.5) with a nonzero function $g(x,t)$.

Following [17] we obtain a result which gives the same convergence rate at time t as in Theorem 3.10 if we assume that g is sufficiently smooth in $[t-\varepsilon, t]$:

Theorem 3.11. *Assume that* (3.25) *holds with* $0 < \theta_0 \le 1$. *Consider* (2.8), (3.37) *with* $p \ge 1$, *and* $u_0^L = \widehat{P}_L\, u_0$. *Then, for* $t > \varepsilon$, *we have*

$$\big\|\widehat{u}^L(t) - u(t)\big\| \le C_\varepsilon h^{\theta_0 p+\delta}\Big[\|u_0\| + \int_0^t \|f\|\, ds + \int_{t-\varepsilon}^t \big(\|u\|_{X_{\theta_0,p}} + \|u_t\|_{X_{\theta_0,p}}\big) ds\Big] \tag{3.48}$$

Proof. As in the proof Theorem 2.3 in [17], we have, with $e = u - \widehat{u}^L$, $\rho = \widehat{R}^L u - u$,

$$\|e(t)\| \le \|e(0)\| + C\Big(\|\rho(0)\| + \int_0^t \|\rho_t\|\, ds\Big), \tag{3.49}$$

$$\|\rho(0)\| = \Big\|(\widehat{R}^L - I)u_0\Big\| \le Ch^{\theta_0 p+\delta}\, \|u_0\|_{X_{\theta_0,p}}\,, \tag{3.50}$$

$$\|\rho_t\| = \Big\|(\widehat{R}^L - I)u_t\Big\| \le Ch^{\theta_0 p+\delta}\, \|u_t\|_{X_{\theta_0,p}}\,, \tag{3.51}$$

yielding

$$\big\|u - \widehat{u}^L\big\| \le \big\|\widehat{u}_0^L - u\big\| + ch^{\theta_0 p+\delta}\Big(\|u_0\|_{X_{\theta_0,p}} + \int_0^t \|u_t\|_{X_{\theta_0,p}}\Big). \tag{3.52}$$

Now we proceed as in Theorem 3.6 in [17] and write $u = u_1 + u_2 + u_3$ using cutoff functions so that u_2 satisfies a homogenous problem, and u_1, u_3 satisfy inhomogenous problems with zero initial data with $f_1 := u_{1,t} - Au_1 = 0$ for $t \le t_0 - \varepsilon$, $f_3 := u_{3,t} - Au_3 = 0$ for $t \ge t_0 - \frac{3}{4}\varepsilon$. Now we use (3.52) for u_1 and Theorem 3.10 for u_2. For the error $e_3(t)$ corresponding to f_3 the argument in [17] gives, with Theorem 3.10,

$$\|e_3(t)\| \le Ch^{\theta_0 p+\delta} \int_0^{t_0-3\varepsilon/4} \|f_3(s)\|\, ds \le Ch^{\theta_0 p+\delta}\Big(\|u_0\| + \int_0^{t_0} \|f\|\, ds\Big)\,.$$

☐

4 Time Discretization

In this section we analyze the time discretization of the parabolic problem. We wish to apply our error analysis to two situations: (i) the continuous problem (2.5), (2.6) where the DG-discretization leads to a semidiscrete problem (continuous in space, discrete in time), and (ii) to the spatially discrete problem (3.37) where the time discretization leads to a fully discrete problem (see Section 5). In order to accommodate both cases we introduce the abstract Gelfand triple

$$\mathcal{V} \overset{d}{\hookrightarrow} \mathcal{H} \cong \mathcal{H}^* \overset{d}{\hookrightarrow} \mathcal{V}^* \,, \tag{4.1}$$

where in case (i) $\mathcal{V} = V$, $\mathcal{H} = H$ and in case (ii) $\mathcal{V} = \widehat{V}^L$ equipped with $\|\cdot\|_V$ and $\mathcal{H} = \widehat{V}^L$ equipped with $\|\cdot\|_H$. We define the scale of spaces $\mathcal{H}_s$ such that $\mathcal{H}_0 = \mathcal{H}$, $\mathcal{H}_1 = \mathcal{V}$, $\mathcal{H}_{-1} = \mathcal{V}^*$, defining the intermediate values of $s \in [-1, 1]$ by interpolation.

We assume that $A \in \mathcal{L}(\mathcal{V}, \mathcal{V}^*)$ and $\langle Au, u\rangle \geq \alpha \|u\|_{\mathcal{V}}^2$ for all $u \in \mathcal{V}$ and consider the abstract parabolic problem

$$\begin{aligned} u'(t) + Au(t) =&\; g(t) \qquad 0 < t < T < \infty \,, &(4.2)\\ u(0) =&\; u_0 &(4.3)\end{aligned}$$

with u_0, g as in (2.7).

Solutions $u(t)$ of this problem are analytic functions of $t \in (0, T)$ if $g(t)$ is analytic. We build therefore a high order time semi-discretization of (4.2) and prove its exponential convergence. We shall apply this time discretization to the spatially semidiscrete problem in Section 5 below.

4.1 Time regularity

The solution operator of the parabolic problem (4.2), (4.3) generates an analytic semigroup $E(t)$, i.e. the solution $u(t)$ becomes analytic in t for $t > 0$, provided the data $g(t) \in L^\infty(J, \mathcal{V}^*)$ is an analytic function of $t \in [0, T]$ taking values in $\mathcal{H}_\theta$ for some $\theta > -1$. We quantify the time analyticity of g by assuming from now on that there are constants C_g, $\widehat{d}_g$ such that

$$\|g^{(l)}(t)\|_{\mathcal{H}_\theta} \leq C_g\, (\widehat{d}_g)^l\, l! \|g(t)\|_{L^\infty(0,T,\mathcal{H}_\theta)} \quad \text{for all } t \in [0, T] \text{ and all } l \in \mathbb{N}_0, \quad -1 \leq \theta \leq 1. \tag{4.4}$$

The solution $u(t)$ of (4.2), (4.3) is a mild solution (see [9]) and can be represented as

$$u(t) = E(t)u_0 + \int_0^t E(t-s)\, g(s)\, ds,\; 0 \leq t \leq T. \tag{4.5}$$

To address the time-analyticity of $u(t)$, we write $u(t) = u_1(t) + u_2(t)$ with

$$\begin{aligned} u_1'(t) + Au_1(t) &= 0, & u_1(0) &= u_0 \,, &(4.6)\\ u_2'(t) + Au_2(t) &= g(t), & u_2(0) &= 0 \,. &(4.7)\end{aligned}$$

By (2.15) with $\tau = 1$ we have

Proposition 4.1. *For $u_0 \in \mathcal{H}_\theta$, $0 \leq \theta \leq 1$,*

$$u_1(t) = E(t)u_0$$

and there are $C, \widehat{d} > 0$ such that for all $l \in \mathbb{N}_0$, $t > 0$ holds

$$\|u_1^{(l)}(t)\|_{\mathcal{V}}^2 \leq C\widehat{d}^{\,2l+1-\theta}\,\Gamma(2l + 2 - \theta)\, t^{-(2l+1)+\theta} \|u_0\|_{\mathcal{H}_\theta}^2 \,. \tag{4.8}$$

Next, we have for $u_2(t)$ in (4.7):

Proposition 4.2. *Assume that g satisfies* (4.4). *Then,*

$$u_2(t) = \int_0^t E(t-s)\, g(s) ds, \quad 0 \le t \le T\,, \tag{4.9}$$

and there are $C, \widehat{d} > 0$ such that for all $l \in \mathbb{N}_0$, $0 < t \le \min\{1, T\}$ and $0 \le \theta \le 1$

$$\|u_2^{(l)}(t)\|_{\mathcal{V}}^2 \le \tilde{C}_g \widehat{d}_g^{2l}\, \Gamma(2l+2-\theta) t^{-2l+\theta} \|g\|^2_{L^\infty(0,T;\mathcal{H}_{-1+\theta})}. \tag{4.10}$$

Proof. We set $\mathcal{V}_\theta = \mathcal{H}_{-1+2\theta}$ for $0 \le \theta \le 1$ and $\mathcal{V} = \mathcal{V}_1 = \mathcal{H}_1$ in this proof to simplify notation. From (4.5) we have for $l \ge 1$

$$u_2^{(l)}(t) = \sum_{i=0}^{l-1} E^{(i)}(t)\, g^{(l-1-i)}(0) + \int_0^t E(s)\, g^{(l)}(t-s) ds, \quad l \ge 1$$

and we estimate

$$\begin{aligned} \|u_2^{(l)}(t)\|_{\mathcal{V}} \le & \sum_{i=0}^{l-1} \|E^{(i)}\|_{\mathcal{L}(\mathcal{V}_\theta, \mathcal{V})} \|g^{(l-1-i)}(0)\|_{\mathcal{V}_\theta} + \int_0^t \|E(s)\|_{\mathcal{L}(\mathcal{V}_\theta, \mathcal{V})} \|g^{(l)}(t-s)\|_{\mathcal{V}_\theta}\, ds \\ =: & \; S + I\,. \end{aligned}$$

We estimate S. By (2.15) (with θ replaced by $-1+2\theta$ and with $\tau = 1$)

$$\|E^{(i)}\|_{\mathcal{L}(\mathcal{V}_\theta, \mathcal{V})} \le C \widehat{d}^{i+1-\theta} \Gamma(2i+3-2\theta)^{1/2} t^{-(i+1-\theta)}.$$

Using $\Gamma(2z) = \pi^{-\frac{1}{2}}\, 2^{2z-1}\, \Gamma(z)\, \Gamma(z+\frac{1}{2})$ with $z = i + 3/2 - \theta$ gives

$$\Gamma(2i+3-2\theta)^{1/2} \le C 2^{i+1-\theta}\, \Gamma(i+3/2-\theta)^{1/2} \Gamma(i+2-\theta)^{1/2} \le \; C 2^{i+1-\theta}\, \Gamma(i+5/2-\theta)\,.$$

With (4.4), we estimate

$$S \le C \sum_{i=0}^{l-1} \widehat{d}^{i+1-\theta}\, \Gamma(2i+3-2\theta)^{\frac{1}{2}}\, t^{-(i+1)+\theta}\, C_g \|g(t)\|_{\mathcal{V}_\theta}\, \widehat{d}_g^{l-1-i} \cdot \Gamma(l-i)\,.$$

and, using the log-convexity of the Gamma function,

$$\begin{aligned} S \le & \; C\, C_g \|g(t)\|_{\mathcal{V}_\theta} \sum_{i=0}^{l-1} (2\widehat{d})^{\,i+1-\theta}\, \Gamma(i+5/2-\theta)\, t^{-(i+1)+\theta}\, \widehat{d}_g^{l-(i+1)}\, \Gamma(l-i) \\ \le & \; C\, C_g \|g(t)\|_{\mathcal{V}_\theta}\, \max\{2\widehat{d}, d_g\}^{l+1}\, \Gamma(l+2-\theta) \sum_{i=0}^{l-1} \frac{\Gamma(i+2-\theta)\Gamma(l-i)}{\Gamma(l+2-\theta)} \cdot t^{-(i+1)+\theta} \\ \le & \; C\, C_g\, \widehat{d}_1^{l+1}\, \|g(t)\|_{\mathcal{V}_\theta}\, \Gamma(l+1) \sum_{i=0}^{l-1} t^{-(i+1)+\theta}\,. \end{aligned}$$

For $0 < t < \min\{1, T\}$, we find that there are $C, \widehat{d}_1 > 0$ such that, for all $l \ge 0$,

$$S \le C\, C_g\, (\widehat{d}_1)^{\,l+1}\, \|g(t)\|_{\mathcal{V}_\theta}\, \Gamma(l+1)\, t^{-l+\theta}\,,$$

and $\Gamma(l+1)^2 \le \Gamma(l+1)\, \Gamma\big(l+\frac{3}{2}\big) \le C\Gamma(2l+2)\, 2^{-2(l+1)}$ gives

$$S^2 \le C_3\, (\widehat{d}_2)^{\,2l+1-2\theta}\, \|g(t)\|^2_{\mathcal{V}_\theta}\, t^{-2l+2\theta}\, \Gamma(2l+2).$$

Analogously, we get from (2.15) with $l = 0$, $0 \le \theta \le 1$, the bound

$$\begin{aligned} I \le\ & C_2 \int_0^t s^{-1+\theta} \, \|g(t-s)\|_{\mathcal{V}_\theta} \, ds \, C_g \, (\widehat{d}_g)^{\,l} \, l! \\ \le\ & C_2 \, C_g \, (\widehat{d}_1)^{\,l+\frac{1}{2}-\theta} \, \Gamma(l+1) \, t^\theta \, \|g\|_{L^\infty((0,T);\mathcal{V}_\theta)} \, . \end{aligned} \tag{4.11}$$

Squaring and adding the bound, the assertion follows if we replace θ by $\theta/2$ and adjust $\widehat{d}$. $\square$

Combining (4.8) and (4.10), we get the following result.

Corollary 4.3. *Assume that $u_0 \in H_\theta$ and that g satisfies (4.4). Then there exist $\tilde{C}_g, \widehat{d}_g$ such that the following hold.*

i) *For $0 < t < \min(1, T)$, $0 \le \theta \le 1$ and every integer $l \ge 0$,*

$$\|u^{(l)}(t)\|^2_{\mathcal{V}} \le \tilde{C}_g \, \widehat{d}_g^{\,2l} \, \Gamma(2l+2-\theta) \Big\{ t^{-2l-1+\theta} \|u_0\|^2_{\mathcal{H}_\theta} + t^{-2l+\theta} \|g\|^2_{L^\infty(0,T;\mathcal{H}_{-1+\theta})} \Big\} \tag{4.12}$$

ii) *Further, if $0 < a < b \le \min(1, T)$, then for any integer $l \ge 1$ and $0 < \theta \le 1$,*

$$\begin{aligned} \int_a^b \|u^{(l)}(t)\|^2_{\mathcal{V}} \, dt \le \tilde{C}_g \, \widehat{d}_g^{\,2l} \, \Gamma(2l+2-\theta) \, \cdot \\ \Big\{ a^{-2l+\theta} \|u_0\|^2_{\mathcal{H}_\theta} + a^{-2l+\theta} (b-a) \|g\|^2_{L^\infty(0,T;\mathcal{H}_{-1+\theta})} \Big\} \, . \end{aligned}$$

iii) *If $s \ge 1$ is arbitrary, and $0 < \theta \le 1$, $0 < a < b \le \min(1, T)$,*

$$\|u\|^2_{H^s((a,b);\mathcal{V})} \le \tilde{C}_g \, \widehat{d}_g^{\,2s} \, \Gamma(2s+2-\theta) a^{-2s+\theta} \Big\{ \|u_0\|^2_{\mathcal{H}_\theta} + \|g\|^2_{L^\infty(0,T;\mathcal{H}_{-1+\theta})} \Big\} \, .$$

4.2 Discontinuous Galerkin time discretization

We discretize problem (4.2) in time. To this end let $\mathcal{M}$ be a partition of $(0, T)$ into $M(\mathcal{M})$ timesteps $\{I_m\}_{m=1}^M$, $I_m = (t_{m-1}, t_m)$, $1 \le m \le M$, of size $k_m = t_m - t_{m-1}$. Define the one-sided limits of $u \in \mathcal{H}$ (or $\mathcal{V}$) as

$$\begin{aligned} u_m^+ &:= \lim_{s \to 0^+} \; u(t_m + s), \quad 0 \le m \le M-1 \, , \\ u_m^- &:= \lim_{s \to 0^+} \; u(t_m - s), \quad 1 \le m \le M \, , \end{aligned} \tag{4.13}$$

and $[u]_m := u_m^+ - u_m^-$, $1 \le m \le M-1$.

Proposition 4.4. *[13] The weak solution $u \in L^2(J;\mathcal{V}) \cap H^1(J;\mathcal{V}^*)$ of (4.2) satisfies*

$$B_{\mathrm{DG}}(u, v) = \; (u_0, v_0^+) + \sum_{m=1}^M \int_{I_m} (g, v)_{\mathcal{V}^* \times \mathcal{V}} \, dt \tag{4.14}$$

for all

$$v \in C_b(\mathcal{M}; \mathcal{V}) := \big\{ u : J \to \mathcal{V} : u\big|_{I_m} \in C^0(\overline{I}_m; \mathcal{V}), \, I_m \in \mathcal{M} \big\},$$

where

$$B_{\mathrm{DG}}(u, v) := \sum_{m=1}^M \int_{I_m} \big\{ (u', v)_{\mathcal{V}^* \times \mathcal{V}} + a(u, v) \big\} \, dt + (u_0^+, v_0^+)_{\mathcal{H}} + \sum_{m=2}^M \big([u]_{m-1}, v_{m-1}^+ \big)_{\mathcal{H}} \, .$$

The one-sided limits in (4.13), (4.14) are well-defined due to (2.9).

With time step k_m we associate an order $r_m \geq 0$, and define the semidiscrete space

$$\mathcal{S}^{\mathbf{r}}(\mathcal{M};\mathcal{V}) = \left\{u : J \to \mathcal{V} : u\big|_{I_m} \in \mathcal{P}^{r_m}(I_m;\mathcal{V}),\ 1 \leq m \leq M\right\}, \tag{4.15}$$

with the order vector $\mathbf{r} = (r_1, \ldots, r_M)$. If the orders are uniform, i.e. $r_m = r$ for all m, we write $\mathcal{S}^r(\mathcal{M};\mathcal{V})$. The number of unknown coefficient functions in $\mathcal{V}$ of $u \in \mathcal{S}^{\mathbf{r}}(\mathcal{M};\mathcal{V}) = \mathcal{S}^{\mathbf{r}}(\mathcal{M}) \otimes \mathcal{V}$ is

$$\dim\left(\mathcal{S}^{\mathbf{r}}(\mathcal{M})\right) = \sum_{m=1}^{M} (r_m + 1)\,. \tag{4.16}$$

The **DG time-stepping scheme** is given by: find $U \in \mathcal{S}^{\mathbf{r}}(\mathcal{M};\mathcal{V})$ such that

$$B_{\mathrm{DG}}(U, W) = (u_0, W_0^+) + \sum_{m=1}^{M} \int_{I_m} (g, W)_{\mathcal{V}^* \times \mathcal{V}}\, dt \tag{4.17}$$

for all $W \in \mathcal{S}^{\mathbf{r}}(\mathcal{M};\mathcal{V})$.

Problem (4.17) has a unique solution U which can be obtained by successively solving M spatial problems for the $r_m + 1$, $m = 1, \ldots, M$, unknown coefficient functions in $\mathcal{V}$ (see Section 4). From (4.14) and (4.17) we have the Galerkin orthogonality

$$B_{\mathrm{DG}}(u - U, W) = 0 \quad \forall W \in \mathcal{S}^{\mathbf{r}}(\mathcal{M};\mathcal{V})\,. \tag{4.18}$$

On $I = (-1, 1)$, define for $r \geq 0$ and $u \in C^0(\overline{I};\mathcal{V})$ the projector $\Pi^r u \in \mathcal{P}^r(I;\mathcal{V})$ by

$$\int_I (u - \Pi^r u, q)_{\mathcal{H}}\, dt = 0 \quad \forall q \in \mathcal{P}^{r-1}(I;\mathcal{V}), \quad (\Pi^r u)(1) = u(1) \in \mathcal{V}\,. \tag{4.19}$$

If $r = 0$, the first condition is void. On time interval (a, b) of length $k = b - a > 0$ we define $\Pi^r_{(a,b)}$ by

$$\Pi^r_{(a,b)} u = \left(\Pi^r(u \circ Q)\right) \circ Q^{-1} \tag{4.20}$$

where $Q : (-1, 1) \to (a, b)$ is given by $\xi \longmapsto x = \frac{1}{2}\,(a + b + \xi k)$. The global DG-interpolant of $u \in C_b(\mathcal{M};\mathcal{V})$ is then given by

$$\mathcal{I}u \in \mathcal{S}^{\mathbf{r}}(\mathcal{M};\mathcal{V}) : \mathcal{I}u\big|_{I_m} = \Pi^{r_m}_{I_m}\left(u\big|_{I_m}\right). \tag{4.21}$$

Theorem 4.5. *The DG-solution U of (3.38) satisfies*

$$\frac{1}{2}\left\|(u - U)^-_M\right\|^2_{\mathcal{H}} + \left\|u - U\right\|^2_{L^2(J;\mathcal{V})} \leq \left(1 + \frac{\alpha}{\beta}\right)^2 \left\|u - \mathcal{I}u\right\|^2_{L^2(J;\mathcal{V})}\,. \tag{4.22}$$

Proof. We have for all $U, W \in \mathcal{S}^{\mathbf{r}}(\mathcal{M};\mathcal{V})$:

$$\begin{aligned}
|||U - W|||^2 &:= \int_J \left\|U - W\right\|^2_{\mathcal{V}} dt + \frac{1}{2}\left\|(U - W)^-_M\right\|^2_{\mathcal{H}} \\
&\leq \frac{1}{2}\left\|(U - W)_0\right\|^2_{\mathcal{H}} + \frac{1}{2}\sum_{m=1}^{M-1} \left\|[U - W]_m\right\|^2_{\mathcal{H}} + \frac{1}{2}\left\|(U - W)^-_M\right\|^2_{\mathcal{H}} + \int_J \left\|U - W\right\|^2_{\mathcal{V}} dt \\
&\leq \frac{1}{\beta}\, B_{\mathrm{DG}}\left(U - W, U - W\right).
\end{aligned}$$

Hence we get

$$|||U - \mathcal{I}u|||^2 \leq \beta^{-1}\, B_{\mathrm{DG}}\left(U - \mathcal{I}u, U - \mathcal{I}u\right) = \beta^{-1}\, B_{\mathrm{DG}}(u - \mathcal{I}u, U - \mathcal{I}u).$$

With $\Theta := U - \mathcal{I}u \in \mathcal{S}^{\mathbf{r}}(\mathcal{M};\mathcal{V})$ it follows that

$$\begin{aligned}
|||U - \mathcal{I}u|||^2 &\le \beta^{-1} \int_J |\{-(u-\mathcal{I}u;\Theta')_{\mathcal{H}}\} + a(u-\mathcal{I}u,\Theta)\,|\,dt \\
&+ \beta^{-1} \sum_{m=1}^{M-1} |(u-\mathcal{I}u)_m^-,\ [\Theta]_m)_{\mathcal{H}}| + \beta^{-1}\left|\left((u-\mathcal{I}u)_M^-,\ \Theta_M^-\right)_{\mathcal{H}}\right| \\
&\le \beta^{-1} \int_J |a(u-\mathcal{I}u,\Theta)|\,dt \le \frac{\alpha}{\beta} \int_J \|u-\mathcal{I}u\|_{\mathcal{V}}\, \|\Theta\|_{\mathcal{V}}\,dt\,, \\
&\le \frac{\alpha}{\beta} \Big(\int_J \|\Theta\|_{\mathcal{V}}^2\,dt\Big)^{\frac12} \Big(\int_J \|u-\mathcal{I}u\|_{\mathcal{V}}^2\,dt\Big)^{\frac12} \le \frac{\alpha}{\beta}\, |||\Theta|||\ \ \|u-\mathcal{I}u\|_{L^2(J;\mathcal{V})}\,,
\end{aligned}$$

whence we find

$$|||U - \mathcal{I}u||| \le \frac{\alpha}{\beta}\ \|u-\mathcal{I}u\|_{L^2(J;\mathcal{V})}.$$

Also

$$|||u - U||| \le |||u-\mathcal{I}u||| + |||U-\mathcal{I}u||| \le \Big(1+\frac{\alpha}{\beta}\Big)\ \|u-\mathcal{I}u\|_{L^2(J;\mathcal{V})}$$

which completes the proof. □

Remark 4.6. If we have instead of (3.5) only the (weaker) Garding inequality

$$\forall u \in \mathcal{V}:\ a(u,u) + \kappa\|u\|_{\mathcal{H}}^2 \ge \beta\,\|u\|_{\mathcal{V}}^2$$

for some $\kappa,\beta > 0$, we still obtain (4.22) with $(1+\frac{\alpha}{\beta})$ replaced by $\exp(\kappa T)(1+\frac{\alpha}{\beta})$, this follows from the substitution $w = e^{-\kappa t}u$.

We see from (4.22) that the error at $t_M = T$ of the DG solution U as well as its $L^2(J;\mathcal{V})$ error is controlled by the quality of the interpolant $\mathcal{I}u$. We show now that $\mathcal{M},\mathbf{r}$ can be chosen such that $\|u-\mathcal{I}u\|_{L^2(J;\mathcal{V})}^2$ decreases exponentially in $N = \dim(\mathcal{S}^r(\mathcal{M}))$.

Definition 4.7. *A mesh $\{I_m\}_{m=1}^M$ in $J=(0,T)$ is geometric with M time steps $I_m = (t_{m-1},t_m)$, $m=1,\dots,M$, and grading factor $\sigma \in (0,1)$, if*

$$t_0 = 0,\ \ t_m = T\sigma^{M-m}, \quad 1 \le m \le M\,. \tag{4.23}$$

Then

$$k_m = \lambda t_{m-1},\ \ \lambda = (1-\sigma)/\sigma, \quad 2 \le m \le M\,. \tag{4.24}$$

We write $\mathcal{M}_{M,\sigma}$ for such meshes.

We define $\gamma = \max\{1,\lambda\}$. We will also use variable orders $\{r_m\}_{m=1}^M$ on $\mathcal{M}_{M,\sigma}$:

Definition 4.8. *The order vector $\{r_m\}_{m=1}^M$ is linear with slope $\mu > 0$ on $\mathcal{M}_{M,\sigma}$, if $r_1 = 0$ and if $r_m = \lfloor \mu m \rfloor$, $m=2,\dots,M$.*

We estimate the approximation error

$$\|u-\mathcal{I}u\|_{L^2(J;\mathcal{V})}^2 = \|u-\Pi_{I_1}^{r_1}u\|_{L^2(I_1;\mathcal{V})}^2 + \sum_{m=2}^M \|u-\Pi_{I_m}^{r_m}u\|_{L^2(I_m;\mathcal{V})}^2 \tag{4.25}$$

in each time step. We start with the first time step I_1 and recall that $r_1 = 0$.

Lemma 4.9. *For $u \in \mathcal{H}_\theta$ and $g(t)$ satisfying (4.4) for some $0<\theta\le 1$, we have for $0<k\le 1$*

$$\|u-\Pi_{I_1}^{r_1}u\|_{L^2(I_1,\mathcal{V})}^2 \le C\left(k^\theta\,\|u_0\|_{\mathcal{H}_\theta}^2 + k^{\theta+1}\,\|g\|_{L^\infty(J;\mathcal{H}_{-1+\theta})}^2\right).$$

Proof. Recall that $r_1 = 0$ and that by (4.19) $\Pi^0_{I_1} u = u(k)$. As in (4.6), (4.7), $u = u_1 + u_2$ and

$$\int_0^k \|u(t) - u(k)\|^2_{\mathcal{V}}\, dt \le C \sum_{\ell=1}^4 T_\ell\,,$$

$$T_1 = \int_0^k \|u_1(t)\|^2_{\mathcal{V}}\, dt,\ T_2 = \int_0^k \|u_1(k)\|^2_{\mathcal{V}}\, dt,\ T_3 = \int_0^k \|u_2(t)\|^2_{\mathcal{V}}\, dt,\ T_4 = \int_0^k \|u_2(k)\|^2_{\mathcal{V}}\, dt\,.$$

T_1, T_2 are estimated as in [13]; consider T_3: using (4.12) with $l = 0$ gives for $0 < \theta \le 1$

$$T_3 = \int_0^k \|u_2(t)\|^2_{\mathcal{V}}\, dt \le C \int_0^k t^\theta\, dt\, \|g\|^2_{L^\infty(0,k;\mathcal{H}_{-1+\theta})}.$$

Analogously, from (4.10) with $l = 0$ it follows that

$$T_4 = k\, \|u_2(k)\|^2_{\mathcal{V}} \le C k^{1+\theta}\, \|g\|^2_{L^\infty(0,k;\mathcal{H}_{-1+\theta})}.$$

□

For the intervals I_m with $m \ge 2$ we have the following result.

Lemma 4.10. *Assume* (4.4). *Then for every $\alpha \in (0, 1)$ there exist constants $C, \widehat{d} > 0$ such that for $I_m \in \mathcal{M}_{M,\sigma}$, $r_m \ge 1$, $m \ge 2$, and for $0 \le \theta \le 1$,*

$$\left\|u - \Pi^{r_m}_{I_m} u\right\|^2_{L^2(I_m;\mathcal{V})} \le C\sigma^{(M-m+1)\theta} (f_{\gamma\widehat{d}}(\alpha))^{r_m} \left\{\left\|u_0\right\|^2_{\mathcal{H}_\theta} + \left\|g\right\|^2_{L^\infty(J;\mathcal{H}_{-1+\theta})}\right\} \tag{4.26}$$

where $\gamma = \max\{1, (1 - \sigma)/\sigma\}$ and

$$f_\eta(\alpha) = \eta^{2\alpha}\, \frac{(1 - \alpha)^{1-\alpha}}{(1 + \alpha)^{1+\alpha}}\,.$$

Proof. We write I in place of I_m and t in place of t_{m-1}, and α, r, s in the following calculations. Then, as in [13],

$$\begin{aligned}\left\|u - \Pi^r_I u\right\|^2_{L^2(I;\mathcal{V})} &\le C\, \frac{\Gamma(r + 1 - s)}{r^2\, \Gamma(r + 1 + s)} \left(\frac{k}{2}\right)^{2(s+1)} \left\|u\right\|^2_{H^{s+1}(I;\mathcal{V})} \\ &\le C\, \frac{\Gamma(r + 1 - s)}{r^2\, \Gamma(r + 1 + s)} \left(\frac{\lambda}{2}\right)^{2(s+1)} t^{2s+2} \left\|u\right\|^2_{H^{s+1}(I;\mathcal{V})}\,.\end{aligned}$$

By Corollary 4.3 iii) (with $s + 1$ in place of s) we find, for any $s > 0$ and $r \ge 1$, $0 \le \theta \le 1$:

$$\left\|u - \Pi^r_I u\right\|^2_{L^2(I;\mathcal{V})} \le C\, \frac{\Gamma(r + 1 - s)}{r^2\, \Gamma(r + 1 + s)} \left(\frac{\gamma \widehat{d}_1}{2}\right)^{2s} \Gamma(2s + 1)\, t^\theta \Big\{\left\|u_0\right\|^2_{\mathcal{H}_\theta} + \left\|g\right\|^2_{L^\infty(J;\mathcal{H}_{-1+\theta})}\Big\}\,.$$

Choosing $s = \alpha r$ with $0 < \alpha < 1$ and using Stirling's formula gives

$$\frac{\Gamma(r + 1 - s)}{\Gamma(r + 1 + s)}\, \Gamma(2s + 1) \le \sqrt{r}\, 2^{2s} \left[\frac{(1 - \alpha)^{1-\alpha}}{(1 + \alpha)^{1+\alpha}}\right]^r,$$

and the claim follows. □

Theorem 4.11. *Assume that the initial value problem* (4.2) *is discretized using* (4.17) *on a geometric partition $\mathcal{M}_{M,\sigma}$ with $0 < \sigma < 1$ and the order vector $\mathbf{r}$ is linear with slope*

$$\mu > \max\Big\{1,\ \frac{\theta |\ln(\sigma)|}{|\ln(f_{\gamma\widehat{d}}(\alpha^*))|}\Big\},\ \alpha^* = (1 + \gamma^2 \widehat{d}^2)^{-\frac{1}{2}} \tag{4.27}$$

where $\widehat{d}$ is, as in Propositions 4.1, 4.2, depending only on u_0 and g. Then there exist $b, C > 0$ independent of M, α, β such that for $0 < \theta \leq 1$,

$$\left\|(u-U)_M^-\right\|_{\mathcal{H}}^2 + \left\|u-U\right\|_{L^2(J;\mathcal{V})}^2 \leq C\Big(1+\frac{\alpha}{\beta}\Big)^2 e^{-b\sqrt{N}}\Big\{\|u_0\|_{\mathcal{H}_\theta}^2 + \left\|g\right\|_{L^\infty(0,T;\mathcal{H}_{-1+\theta})}^2\Big\} \tag{4.28}$$

where $b = c\theta|\ln(\sigma)|/\sqrt{\mu}$ and $N = \dim(\mathcal{S}^{\mathbf{r}}(\mathcal{M}_{\mathcal{M},\sigma}))$ is the number of spatial problems to be solved in the DG time discretization.

Proof. We consider only $T = 1$, the general case is obtained by scaling. Further, by (4.22) it suffices to bound (4.25). We apply for time step I_1 Lemma 4.9 and for time interval I_m, $m \geq 2$, Lemma 4.10. This gives, with $k_1 = t_1 = \sigma^n$ (cf. (4.23)),

$$\left\|u-\mathcal{I}u\right\|_{L^2(J;\mathcal{V})}^2 \leq C\Big\{\sigma^{(M-1)\theta} + \sum_{m=2}^{M} \sigma^{(M-m+1)\theta} f_{\gamma d}(\alpha_m)^{r_m}\Big\}\Big\{\|u_0\|_{\mathcal{H}_\theta}^2 + \left\|g\right\|_{L^\infty(0,T;\mathcal{H}_{-1+\theta})}^2\Big\}.$$

Now we select $\alpha_m = \alpha^* = (1+\gamma^2 d^2)^{-\frac{1}{2}}$. Then $f_{\gamma d}(\alpha^*) = f_{\min} < 1$, and, by (4.27),

$$\sigma^{(2-m)\theta}\, f_{\min}^{r_m} \leq C\,\sigma^{2\theta}\,\Big(\frac{f_{\min}^{\mu}}{\sigma^\theta}\Big)^m \leq C\,\sigma^{2\theta}\,q^m,\ q<1\,;$$

hence

$$\left\|u-\mathcal{I}u\right\|_{L^2(J;\mathcal{V})}^2 \leq C\sigma^{(M-1)\theta}\Big\{1+\sigma^{2\theta}\sum_{m=2}^{M} q^m\Big\}\Big\{\|u_0\|_{\mathcal{H}_\theta}^2 + \left\|g\right\|_{L^\infty(0,T;\mathcal{H}_{-1+\theta})}^2\Big\}$$

from where the assertion follows on noticing that $N \leq C\mu M^2$ as $M \to \infty$. □

In the following we discuss the convergence of three modifications of DG time-stepping.

Remark 4.12. Instead of the linear order vector with slope μ in Definition 4.8, we can also choose the same polynomial degree $r_m = r = \lfloor \mu M \rfloor$ on all intervals $I_m \in \mathcal{M}_{M,\sigma}$, $m = 1, ..., M$.

In this case (4.28) holds with $b > 0$ for any $\mu > 0$; i.e., (4.27) is not necessary, see [13].

Remark 4.13. We can also achieve convergence of DG time-stepping by mesh refinement while keeping the polynomial degree r fixed. It was shown in [13] that for the regularity (4.8), (4.10) of the exact solution for some $0 < \theta \leq 1$, $\mathcal{M}$ must be algebraically graded as follows:

$$t_m = h(\frac{mT}{M}), m = 0, ..., M \quad \text{with} \quad h(t) = t^{(2r+3)/\theta}. \tag{4.29}$$

We then have the algebraic convergence

$$\left\|(u-U)_M^-\right\|_{\mathcal{H}} + \left\|u-U\right\|_{L^2(J;\mathcal{V})} \leq C\Big(1+\frac{\alpha}{\beta}\Big)M^{-(r+1)} \tag{4.30}$$

where C depends on $u_0 \in \mathcal{H}_\theta$, g and on r. The case $r = 0$ corresponds to the backward Euler scheme.

Remark 4.14. We can use a single time step, i.e., $M = 1$, and increase the polynomial degree r. If u is analytic in $[0, T]$ with values in $\mathcal{V}$, this results in exponential convergence $O(\exp(-bN))$ where N is the number of spatial problems to be solved [13].

5 Discretization in space and time

5.1 Fully discrete problem

We discretize the parabolic problem (2.5), (2.6) in time with a hp-discontinuous Galerkin method using a geometric mesh $\mathcal{M}_{M,\sigma}$. For simplicity, we choose the DG time-stepping with uniform degree vector r

as in Remark 4.12. We further choose $\mu = 1$ and a geometric time-step sequence $\mathcal{M}_{M,\sigma}$ in $(0, T)$ with grading factor $\sigma \in (0, 1)$ and $M = r$ time steps.

The space discretization will be performed in $\Omega = (0, 1)^d$ with the sparse grid subspace $\widehat{V}^L$ of V of meshwidth $h = 2^{-L}$, $L > 0$.

We now define the **solution $\widehat{U}^L$ of the fully discrete problem** as follows:
Find $\widehat{U}^L \in \mathcal{S}^r(\mathcal{M}_{r,\sigma}; \widehat{V}^L)$ such that

$$B_{\mathrm{DG}}(\widehat{U}^L, \widehat{W}) = (u_0, \widehat{W}_0^+) + \sum_{m=1}^{M} \int_{I_m} (g, \widehat{W})_{V^* \times V}\, dt \quad \forall \widehat{W} \in \mathcal{S}^r(\mathcal{M}_{r,\sigma}; \widehat{V}^L). \tag{5.1}$$

5.2 Error analysis

Let us estimate the error

$$\|u(T) - \widehat{U}^L(T)\| \qquad T > 0. \tag{5.2}$$

If we denote by $\widehat{u}^L$ the semidiscrete solution in (3.36), (3.37), we have

$$\|u(T) - \widehat{U}^L(T)\| \leq \|u(T) - \widehat{u}^L(T)\| + \|\widehat{u}^L(T) - \widehat{U}^L\|\,. \tag{5.3}$$

The first error term was estimated in Theorem 3.10. For the second error term we observe that the spatially semidiscrete problem (3.36), (3.37) fits into the abstract framework of Section 4: we keep $\mathcal{H} = L^2(\Omega)$ as pivot space and select $\mathcal{V} = \widehat{V}^L \subset L^2(\Omega)$, equipped with the $H_0^1(\Omega)$-norm. Then the bilinear form $a(\cdot, \cdot) : \widehat{V}^L \times \widehat{V}^L \to \mathbb{R}$ induces an operator $\widehat{A}^L : \widehat{V}^L \to (\widehat{V}^L)' = \mathcal{V}^*$ and the semidiscrete problem (3.36), (3.37) reads

$$(\widehat{u}^L)'(t) + \widehat{A}^L\, \widehat{u}^L(t) = \widehat{P}_L^* g \;\text{ in }\; J = (0, T)\,, \tag{5.4}$$

with initial condition

$$\widehat{u}^L(0) = \widehat{P}_L\, u_0 := \widehat{u}_0^L. \tag{5.5}$$

We then obtain the following error estimate for the fully discrete solution $\widehat{U}^L(T)$.

Theorem 5.1. *Assume that the initial data u_0 of* (1.1), (1.2) *belongs to $H_\theta = (L^2(\Omega), H_0^1(\Omega))_{\theta,2}$ for some $0 < \theta \leq 1$. Then, the error* (5.2) *of the fully discrete Galerkin scheme* (5.1) *with uniform order r, geometric time step sequence $\mathcal{M}_{r,\sigma}$ in $J = (0, T)$ and sparse grids in space with meshwidth $h = 2^{-L}$ satisfies the error estimate*

$$\|u(T) - \widehat{U}^L(T)\| \leq C_1(u_0, g) h^{\theta_0 p + \delta} + C_2(u_0, g) e^{-br}\,. \tag{5.6}$$

Note that $C_1(u_0, g) = \|u_0\| + \|g\|_{L^2([0,T];V^*)} + C_g$ where C_g measures the additional spatial regularity of g in $[T - \delta, T]$. We have $C_2(u_0, g) = \|u_0\|_{H_\theta} + \|g\|_{L^\infty([0,T];H_{-1+\theta})} + C_g'$ where C_g' depends on $\widehat{d}_g$.

Corollary 5.2. *If we choose $r = O(|\log h|)$ then*

$$\|u(T) - \widehat{U}^L(T)\| \leq C(u_0, g) h^{\theta_0 p + \delta} \tag{5.7}$$

where $C(u_0, g) = C_1(u_0, g) + C_2(u_0, g)$.

5.3 Derivation of the linear system

In each of the M time steps (5.1) amounts to the solution of a linear system of size

$$(r + 1)\widehat{N}_L = (r + 1) O(h^{-1} |\log h|^{d-1}) \tag{5.8}$$

which depends on the time step k and h and which we now derive.

Let $\mathcal{Q} := \mathcal{P}_r(I_m, \widehat{V}^L)$, equipped with the norm of $L^2([t_{m-1}, t_m]) \otimes V$. In time step m of the algorithm we have to determine $\widehat{U}^L_m := \widehat{U}^L\Big|_{I_m} \in \mathcal{Q}$ which by(5.1) satisfies, for all $W \in \mathcal{Q}$,

$$\int_{t_{m-1}}^{t_m} \Big[\Big((\widehat{U}^L_m)', W\Big) + a(\widehat{U}^L_m, W)\Big] dt + \Big(\widehat{U}^L_m(t_{m-1}), W(t_{m-1})\Big)$$
$$= \int_{t_{m-1}}^{t_m} (g, W)_{V^* \times V} dt + \Big(\widehat{U}^L_{m-1}(t_{m-1}), W(t_{m-1})\Big) \quad (5.9)$$

where the expression $\widehat{U}^L_0(t_0)$ is defined to mean the initial value $\widehat{u}^L_0$.

Let $\{\varphi\}_{j=0}^{r_m}$ be a basis of the polynomial space $\mathcal{P}_{r_m}(-1, 1)$. Then the time shape functions on time interval I_m are given by $\varphi_j \circ F_m^{-1}$ where the mapping $F_m \colon (-1, 1) \to I_m$ is given by

$$t = F_m(\tau) = \frac{1}{2}(t_{m-1} + t_m) + \frac{1}{2} k_m \tau, \qquad k_m = t_m - t_{m-1}, \qquad \tau \in (-1, 1).$$

If we write $\widehat{U}^L_m(x, t)$ and $\widehat{W}$ in (5.9) as

$$\widehat{U}^L_m(x, t) = \sum_{j=0}^{r_m} \widehat{U}^L_{m,j}(x)(\varphi_j \circ F_m^{-1})(t), \quad \widehat{W}(x, t) = \sum_{j=0}^{r_m} \widehat{W}^L_{m,j}(x)(\varphi_j \circ F_m^{-1})(t), \quad (5.10)$$

the variational problem (5.9) has the following form:
Find $(\widehat{U}^L_{m,j})_{j=0}^r \in (\widehat{V}^L)^{r+1}$ such that for every $(\widehat{W}_i)_{i=0}^r \in (\widehat{V}^L)^{r+1}$

$$\sum_{i,j=0}^r \left(C_{ij} \cdot (\widehat{U}^L_{m,j}, \widehat{W}_i)_H + \frac{k_m}{2} G_{ij} \cdot a(\widehat{U}^L_{m,j}, \widehat{W}_i) \right) = \sum_{i=0}^r \left(\frac{k_m}{2} f^1_{m,i}(\widehat{W}_i) + f^2_{m,i}(\widehat{W}_i) \right), \quad (5.11)$$

where (see [18])

$$f^1_{m,i}(v) := \Big(\int_{-1}^1 (g \circ F_m)\varphi_i d\tau, v\Big)_H, \quad f^2_{m,i}(v) := \varphi_i(-1)\big(\widehat{U}^L_{m-1}(t_{m-1}), v\big)_H$$
$$C_{ij} := \int_{-1}^1 \varphi'_j \varphi_i d\tau + \varphi_j(-1)\varphi_i(-1), \quad G_{ij} = \int_{-1}^1 \varphi_j \varphi_i d\tau. \quad (5.12)$$

Equation (5.11) is a linear system of size $(r + 1)\widehat{N}_L$ to be solved in each time step $m = 1, \dots, M$. We will drop the subscript m for sake of readability. Denoting by $\mathbf{M}$ and $\mathbf{A}$ the mass and stiffness matrix of $\widehat{V}^L$ with respect to $(\cdot, \cdot)_H$ and $a(\cdot, \cdot)$, respectively, (5.11) takes the matrix form

$$\mathbf{R}\mathbf{u} = \mathbf{f}, \quad (5.13)$$
$$\mathbf{R} = \mathbf{C} \otimes \mathbf{M} + \frac{k}{2} \mathbf{G} \otimes \mathbf{A}, \qquad \mathbf{f} = \frac{k}{2}\mathbf{f}^1 + \mathbf{f}^2, \quad (5.14)$$

where $\mathbf{u}$ denotes the coefficient vector of $\widehat{U}^L_m \in \mathcal{Q}$.

Remark 5.3. If the coefficients are independent of t, we obtain with the temporal shape functions $\varphi_i(\tau) = (i+1/2)^{1/2} L_i(\tau)$ and with L_i denoting the i-th Legendre polynomial on $(-1, 1)$ (normalized such that $L_i(1) = 1$), that $\mathbf{G} = \mathbf{I}$ in (5.12) and

$$C_{ij} = \sigma_{ij}(i + \tfrac{1}{2})^{1/2}(j + \tfrac{1}{2})^{1/2}, \quad \sigma_{ij} = \begin{cases} (-1)^{i+j} & \text{if } j > i \\ 1 & \text{otherwise} \end{cases}, \quad i, j = 0, \dots, r. \quad (5.15)$$

If the coefficients of A depend on t, $\mathbf{C}$ has to be computed by numerical quadrature.

From now on we will use temporal shape functions $\varphi_i(\tau) = (i + 1/2)^{1/2} L_i(\tau)$.

5.4 Decoupling

As observed in [13] the system (5.13) of size $(r+1)\widehat{N}_L$ can be reduced to solving $r+1$ linear systems of size $\widehat{N}_L$: we use the Schur decomposition $\mathbf{C} = \mathbf{Q}\mathbf{T}\mathbf{Q}^H$ with a unitary matrix $\mathbf{Q}$ and an upper triangular matrix $\mathbf{T}$ which has the eigenvalues $\lambda_1, \ldots, \lambda_{r+1}$ of $\mathbf{C}$ on the diagonal. Multiplying (5.13) by $\mathbf{Q}^H \otimes \mathbf{I}$ from the left gives

$$\left(\mathbf{T} \otimes \mathbf{M} + \frac{k}{2}\mathbf{I} \otimes \mathbf{A}\right)\mathbf{w} = \mathbf{g}$$

with

$$\mathbf{w} := (\mathbf{Q}^H \otimes \mathbf{I})\mathbf{x}, \qquad \mathbf{g} := (\mathbf{Q}^H \otimes \mathbf{I})\mathbf{f}$$

This system is block-upper-triangular: with $\mathbf{w} = (\mathbf{w}_0, \ldots, \mathbf{w}_r)$ we obtain the solution by solving

$$\left(\lambda_{j+1}\mathbf{M} + \frac{k}{2}\mathbf{A}\right)\mathbf{w}_j = \mathbf{s}_j \quad \text{for } j = r, r-1, \ldots, 0 \tag{5.16}$$

where

$$\mathbf{s}_j := \mathbf{g}_j - \sum_{l=j+1}^{r} \mathbf{T}_{j+1,l+1}\mathbf{M}\mathbf{w}_l.$$

For each DG time step, we have to solve the $r+1$ linear systems in (5.16). Each of these linear systems is of the same type as in the backward Euler method where the matrix is $\mathbf{M} + k\mathbf{A}$. Therefore an implementation of the DG method (5.1) is very similar to an implementation of the backward Euler method.

If the coefficients of the operator A are independent of t, the Schur decomposition of the $(r+1)\times(r+1)$ matrix $\mathbf{C}$ needs to be computed once at the beginning of the time stepping algorithm.

Remark 5.4. Note that the $r+1$ linear systems (5.16) have to be solved sequentially. As described in [13], there is an alternative scheme which uses the matrix $\mathbf{Y}$ of eigenvectors instead of $\mathbf{Q}$ (assuming that $\mathbf{C}$ is diagonalizable). This leads to a system which is block-diagonal (instead of block-upper-triangular). It can be solved by solving $r+1$ linear systems of size $\widehat{N}_L \times \widehat{N}_L$ *in parallel*. However, the condition number of $\mathbf{Y}$ increases rapidly with r. If one solves the linear systems exactly this only causes a magnification of the round-off error and works well in practice for values $r < 10$. If one solves the linear system with incomplete iterations the error will be multiplied by the condition number of $\mathbf{Y}$, and one has to increase the number of iterations very rapidly with increasing r to compensate. With the choice $r = O(|\log h|)$ suggested by Corollary 5.2 we would not be able to obtain an overall complexity of $O(h^{-1}\,|\log h|^c)$.

5.5 Iterative Solution of Linear Equations

By (5.16), a time step of order r amounts to solving $r+1$ linear systems with coefficient matrix

$$\mathbf{B} := \lambda\mathbf{M} + \frac{k}{2}\mathbf{A} \tag{5.17}$$

where λ is an eigenvalue of $\mathbf{C}$ in (5.15). We solve the equations (5.16) approximately with incomplete GMRES iteration, causing an additional error in the overall scheme which we analyze here together with the overall complexity. Throughout, we denote by $\|\cdot\|$ the 2-norm of a vector or a matrix and we use the notation $\|\mathbf{w}\|_{\mathbf{B}} := (\mathbf{w}^H\mathbf{B}\mathbf{w})^{1/2}$.

5.5.1 Eigenvalues of C

For the convergence analysis of the GMRES method we will need the following properties of the eigenvalues of the matrix $\mathbf{C}$ from (5.15):

Lemma 5.5. *The eigenvalues $\lambda_j^{(r)}$ of the matrix $\mathbf{C}$ from (5.15) satisfy for $r = 0, 1, 2, \ldots$*

$$\operatorname{Re}\lambda_j^{(r)} \geq C_1\big|\lambda_j^{(r)}\big|r^{-\alpha}, \qquad j = 1, \ldots, r+1 \tag{5.18}$$

$$\big|\lambda_j^{(r)}\big| \geq C_2 r^{\tilde{\alpha}} \tag{5.19}$$

with $\alpha = 2$, $\tilde{\alpha} = 0$ and constants $C_1, C_2 > 0$ independent of r. Furthermore, the matrix $\mathbf{T}$ of the Schur decomposition $\mathbf{C} = \mathbf{QTQ}^H$ satisfies

$$\|\mathbf{T}\|_2 \leq Cr^2. \tag{5.20}$$

Proof. Let $\mathcal{P}_r$ denote the space of complex-valued polynomials of degree $\leq r$, let $\|g\|^2 = \int_{-1}^{1} |g(t)|^2 \, dt$. Because of (5.12), an eigenvalue λ of the matrix $\mathbf{C}$ corresponds to the existence of a nonzero polynomial $p \in \mathcal{P}_r$ such that

$$\int_{-1}^{1} p'\overline{q}\, dt + p(-1)\overline{q(-1)} = \lambda \int_{-1}^{1} p\overline{q}\, dt \qquad \text{for all } q \in \mathcal{P}_r. \tag{5.21}$$

Using this with $q = p$ and taking the real and imaginary parts gives

$$\tfrac{1}{2} |p(-1)|^2 + \tfrac{1}{2} |p(1)|^2 = \mathrm{Re}(\lambda) \|p\|^2 \tag{5.22}$$

$$\mathrm{Im} \int_{-1}^{1} p'\overline{p}\, dt = \mathrm{Im}(\lambda) \|p\|^2 \tag{5.23}$$

where we used $\mathrm{Re}(p'\overline{p}) = \frac{1}{2}(p\overline{p})'$ for the real part. Using (5.21) with $q(t) = (1+t)p'(t)$ gives

$$\int_{-1}^{1} (1+t)\, |p'|^2 \, dt = \lambda S \tag{5.24}$$

where $S := \displaystyle\int_{-1}^{1} (1+t)\, p\overline{p}'\, dt$. As $\lambda \neq 0$ ((4.17) always has a unique solution) we have that

$$\mathrm{Re}\, S = \frac{\mathrm{Re}(\lambda)}{|\lambda|^2} \int_{-1}^{1} (1+t)\, |p'|^2 \, dt \geq 0 \tag{5.25}$$

since (5.22) shows $\mathrm{Re}\,\lambda \geq 0$. Integrating by parts we obtain

$$0 \leq \mathrm{Re}\, S = \Big[(1+t)\, \tfrac{1}{2} p\overline{p}\Big]_{-1}^{1} - \int_{-1}^{1} \tfrac{1}{2} p\overline{p}\, dt = |p(1)|^2 - \tfrac{1}{2} \|p\|^2 . \tag{5.26}$$

From (5.23) we get with the inverse inequality (e.g. [15])

$$|\mathrm{Im}(\lambda)|\, \|p\|^2 \leq \left| \int_{-1}^{1} p'\overline{p}\, dt \right| \leq \|p\|\, \|p'\| \leq \sqrt{3} r^2 \|p\|^2 . \tag{5.27}$$

Now (5.26) and (5.22) give

$$|\mathrm{Im}(\lambda)|\, \|p\|^2 \leq \sqrt{3} r^2 2\, |p(1)|^2 \leq 4\sqrt{3} r^2\, \mathrm{Re}(\lambda)\, \|p\|^2 . \tag{5.28}$$

Inequality (5.19) is proved in [13]. We have

$$\left| \int_{-1}^{1} p'\overline{q}\, dt + p(-1)\overline{q(-1)} \right| \leq Cr^2 \|p\|\, \|q\|$$

using the inverse inequality and $|p(-1)| \leq C' r \|p\|$ (see [13]). This shows $\|\mathbf{C}\|_2 \leq r^2$ and (5.20) follows since $\mathbf{Q}$ is unitary . ☐

Remark 5.6. The estimates (5.18), (5.19) are not sharp. Results of computations for $r = 1, \ldots, 50$ performed with 50 digits of accuracy are shown in Figure 1. They suggest that (5.18) seems to hold with $\alpha = \frac{2}{3}$, and (5.19) with $\tilde{\alpha} = 1$.

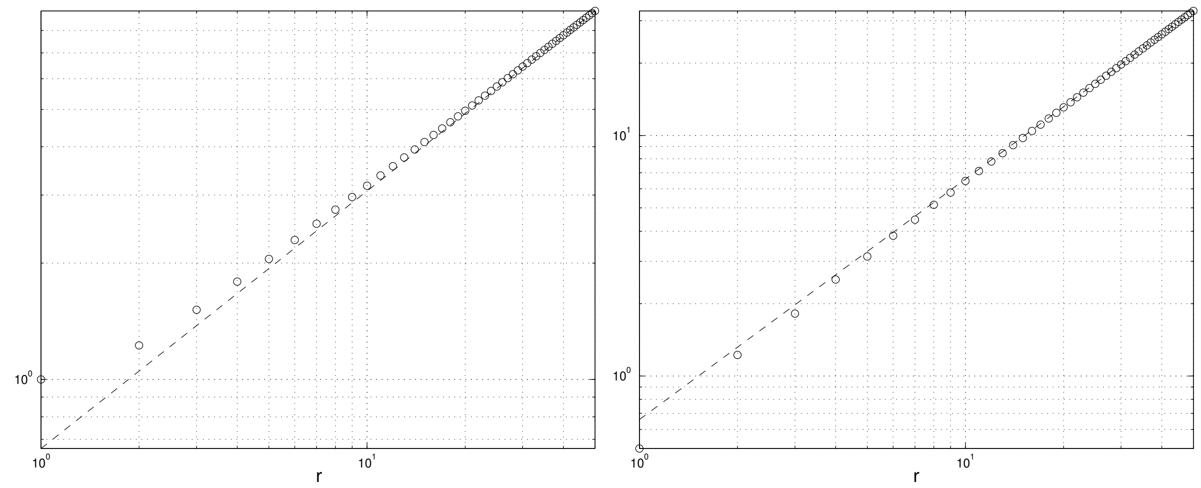

Figure 1: *Left:* $\max_j \dfrac{|\lambda_j^{(r)}|}{\operatorname{Re}\lambda_j^{(r)}}$ for $r = 1, \ldots, 50$, dashed line is $Cr^{2/3}$. *Right:* $\min_j |\lambda_j^{(r)}|$ for $r = 1, \ldots, 50$, dashed line is Cr.

5.5.2 Preconditioning

The norm equivalence (3.4) with $\theta = 0$ implies for every $v \in \widehat{V}^L$ of the form (3.3) with coefficient vector $\mathbf{v} = (v_j^\ell)$

$$C_1 \|v\|^2 \le \mathbf{v}^H \mathbf{M} \mathbf{v} \le C_2 \|v\|^2 \tag{5.29}$$

with constants C_1, C_2 independent of L. Let $\mathbf{D}_A$ denote the diagonal matrix with entries $2^{2l_1} + \cdots + 2^{2l_d}$ for an index corresponding to level $(l_1, \ldots, l_d)$. Then (2.3), (2.4) and (3.4) with $\theta = 1$ imply that

$$C_1 \mathbf{v}^H \mathbf{D}_A \mathbf{v} \le \mathbf{v}^H \mathbf{A} \mathbf{v} \le C_2 \mathbf{v}^H \mathbf{D}_A \mathbf{v} \tag{5.30}$$

with constants C_1, C_2 independent of L.

Let $\|\mathbf{w}\|_{\mathbf{D}_A} := (\mathbf{w}^H \mathbf{D}_A \mathbf{w})^{1/2}$, $\|\mathbf{w}\|_{\mathbf{D}_A^{-1}} := (\mathbf{w}^H \mathbf{D}_A^{-1} \mathbf{w})^{1/2}$. For $v \in \widehat{V}^L$ with coefficient vector $\mathbf{v}$ and $f \in (\widehat{V}^L)^*$ with coefficient vector $\mathbf{f}$ we then have

$$\|v\|_V \sim \|\mathbf{v}\|_{\mathbf{D}_A}, \qquad \|f\|_{V^*} \sim \|\mathbf{f}\|_{\mathbf{D}_A^{-1}} \tag{5.31}$$

where the norm equivalences hold with constants independent of L.

We now define for preconditioning the diagonal matrix $\mathbf{S}$ and the scaled matrix $\widehat{\mathbf{B}}$ as

$$\mathbf{S} := \Big(\operatorname{Re}(\lambda)\mathbf{I} + \frac{k}{2}\mathbf{D}_A\Big)^{1/2}, \qquad \widehat{\mathbf{B}} := \mathbf{S}^{-1}\mathbf{B}\mathbf{S}^{-1}. \tag{5.32}$$

Lemma 5.7. *For the linear system $\widehat{\mathbf{B}}\widehat{\mathbf{x}} = \widehat{\mathbf{b}}$ let $\widehat{\mathbf{x}}_j$ denote the iterates obtained by the restarted GMRES(m_0) method with initial guess $\widehat{\mathbf{x}}_0$. Then*

$$\left\|\widehat{\mathbf{b}} - \widehat{\mathbf{B}}\mathbf{x}_j\right\| \le C(1 - cr^{-2\alpha})^j \left\|\widehat{\mathbf{b}} - \widehat{\mathbf{B}}\mathbf{x}_0\right\| \tag{5.33}$$

Let $\mathbf{x}_j = \mathbf{S}^{-1}\widehat{\mathbf{x}}_j$, $\mathbf{b} = \mathbf{S}\widehat{\mathbf{b}}$. Then

$$\|\mathbf{b} - \mathbf{B}\mathbf{x}_j\|_{\mathbf{D}_A^{-1}} \le Ch^{-1}(1 + C_1 k r^{\alpha - \tilde{\alpha}})^{1/2}(1 - cr^{-2\alpha})^j \|\mathbf{b} - \mathbf{B}\mathbf{x}_0\|_{\mathbf{D}_A^{-1}} \tag{5.34}$$

with $\alpha, \tilde{\alpha}$ from (5.18), and C, c independent of L, k, r.

Proof. Since $\operatorname{Re}(\mathbf{x}^H i \operatorname{Im}(\lambda)\mathbf{M}\mathbf{x}) = 0$ we obtain from (5.29), (5.30) that

$$\operatorname{Re}(\mathbf{x}^H \mathbf{B}\mathbf{x}) \ge C\mathbf{x}^H \mathbf{S}^2 \mathbf{x} \quad \forall \mathbf{x} \in \mathbb{C}^{\widehat{N}_L}$$

implying with $\mathbf{y} = \mathbf{S}\mathbf{x}$

$$\mathrm{Re}(\mathbf{y}^H \widehat{\mathbf{B}} \mathbf{y}) \geq C_3 \|\mathbf{y}\|^2 \quad \forall \mathbf{y} \in \mathbb{C}^{\widehat{N}_L}. \tag{5.35}$$

We have

$$\left|\mathbf{x}^H \mathbf{B} \mathbf{y}\right| = \left|\lambda \mathbf{x}^H \mathbf{M} \mathbf{y}\right| + \frac{k}{2}\left|\mathbf{x}^H \mathbf{A}^L \mathbf{y}\right| \leq C |\lambda| \|\mathbf{x}\| \|\mathbf{y}\| + C\frac{k}{2}\left\|\mathbf{D}_A^{1/2}\mathbf{x}\right\| \left\|\mathbf{D}_A^{1/2}\mathbf{y}\right\|$$

With $\mathbf{D} := |\lambda| \mathbf{I} + \frac{k}{2}\mathbf{D}_A$ we then get

$$\left|\mathbf{x}^H \mathbf{B} \mathbf{y}\right| \leq C(\mathbf{x}^H \mathbf{D} \mathbf{x})^{1/2} (\mathbf{y}^H \mathbf{D} \mathbf{y})^{1/2}$$

Now we use (5.18) and obtain

$$\left|\mathbf{x}^H \mathbf{B} \mathbf{y}\right| \leq C r^{\alpha} (\mathbf{x}^H \mathbf{S}^2 \mathbf{x})^{1/2} (\mathbf{y}^H \mathbf{S}^2 \mathbf{y})^{1/2}$$

or

$$\left|\mathbf{x}^H \widehat{\mathbf{B}} \mathbf{y}\right| \leq C_4 r^{\alpha} \|\mathbf{x}\| \|\mathbf{y}\|. \tag{5.36}$$

Inequalities (5.35) and (5.36) can be stated as

$$\lambda_{\min}\Big((\widehat{\mathbf{B}} + \widehat{\mathbf{B}}^H)/2\Big) \geq C_3, \qquad \left\|\widehat{\mathbf{B}}\right\| \leq C_4 r^{\alpha}$$

According to [3] the non-restarted GMRES method for the matrix $\widehat{\mathbf{B}}$ yields iterates $\mathbf{x}_m$ and residuals $\mathbf{r}_m$ satisfying

$$\|\mathbf{r}_m\| \leq \Big(1 - \frac{C_3^2}{r^{2\alpha} C_4^2}\Big)^{m/2} \|\mathbf{r}_0\|$$

which shows (5.33) (the proof in [3] is given for real matrices, but all arguments carry over to the complex case).

Let $\widehat{\mathbf{r}}_j := \widehat{\mathbf{b}} - \widehat{\mathbf{B}}\mathbf{x}_j$, $\mathbf{r}_j := \mathbf{b} - \mathbf{B}\mathbf{x}_j = \mathbf{S}\widehat{\mathbf{r}}_j$. Hence

$$\mathbf{r}_j^H \mathbf{D}_A^{-1} \mathbf{r}_j = \widehat{\mathbf{r}}_j^H \Big(\mathrm{Re}(\lambda)\mathbf{D}_A^{-1} + \frac{k}{2}\mathbf{I}\Big)\widehat{\mathbf{r}}_j.$$

As the elements of the diagonal matrix $\mathbf{D}_A$ are between d and dh^{-2}, we have

$$\frac{\|\mathbf{r}_j\|^2_{\mathbf{D}_A^{-1}}}{\|\mathbf{r}_0\|^2_{\mathbf{D}_A^{-1}}} \leq C\frac{2d^{-1}\mathrm{Re}(\lambda) + k}{2d^{-1}\mathrm{Re}(\lambda)h^2 + k}\,\frac{\|\widehat{\mathbf{r}}_j\|^2}{\|\widehat{\mathbf{r}}_0\|^2} \leq Ch^{-2}(1 + C_1 k \mathrm{Re}(\lambda)^{-1})\frac{\|\widehat{\mathbf{r}}_j\|^2}{\|\widehat{\mathbf{r}}_0\|^2}.$$

Now we get with (5.18), (5.19) that $\mathrm{Re}(\lambda)^{-1} \leq C r^{\alpha} |\lambda|^{-1} \leq C r^{\alpha - \tilde{\alpha}}$. □

5.5.3 Fully discrete scheme with incomplete GMRES

The fully discrete scheme with incomplete GMRES iteration yields approximations $\tilde{U}_1^L, \ldots, \tilde{U}_M^L$ to $\widehat{U}_1^L, \ldots, \widehat{U}_M^L$ and proceeds at each time step $m = 1, \ldots, M$ as follows: write $\tilde{U}_m^L = \tilde{U}_{m-1}^L(t_{m-1}) + \tilde{Z}$ where $\tilde{U}^0(t_0)$ is defined as initial value $\widehat{u}_0^L$ and $\tilde{Z}$ is an approximation of the function $Z \in \mathcal{Q}$ which satisfies for all $V \in \mathcal{Q}$

$$\int_{t_{m-1}}^{t_m} \Big[\Big(Z', V\Big) + a(Z, V)\Big]dt + \Big(Z(t_{m-1}), V(t_{m-1})\Big) = \int_{t_{m-1}}^{t_m} (g, V)_{V^* \times V} dt - \int_{t_{m-1}}^{t_m} a(\tilde{U}_{m-1}^L(t_{m-1}), V) dt. \tag{5.37}$$

This corresponds to a linear system $\mathbf{R}\mathbf{z} = \mathbf{f}$. Using the Schur decomposition we obtain the $r+1$ linear systems (5.16). For each of the linear systems (5.16) with $j = r, r-1, \ldots, 0$ we use 0 as initial guess for $\mathbf{w}_0$ and apply n_G steps of GMRES(m_0), yielding an approximate solution $\tilde{\mathbf{w}}_j$. To analyze the impact of this approximation on the global accuracy, we use the stability of the hp-DG timestepping.

We now use the norm

$$\|g\|_{a,*}^2 := \sup_{v\in\widehat{V}^L} \frac{|\langle g, v\rangle|}{\|v\|_a}.$$

If we let $W = \widehat{U}_m^L$ in (5.9) we obtain with $[\widehat{U}^L]_{m-1} := \widehat{U}_m^L(t_{m-1}) - \widehat{U}_{m-1}^L(t_{m-1})$

$$\begin{aligned}\left\|[\widehat{U}^L]_{m-1}\right\|^2 - \tfrac{1}{2}\left\|\widehat{U}_{m-1}^L(t_{m-1})\right\|^2 + \tfrac{1}{2}\left\|\widehat{U}_m^L(t_m)\right\|^2 + \int_{t_{m-1}}^{t_m}\left\|\widehat{U}_m^L\right\|_V^2 dt &= \int_{t_{m-1}}^{t_m}\left(g, \widehat{U}_m^L\right)_{V^*\times V} dt \\ &\le \tfrac{1}{2}\int_{t_{m-1}}^{t_m}\|g\|_{a,*}^2\, dt + \tfrac{1}{2}\int_{t_{m-1}}^{t_m}\left\|\widehat{U}_m^L\right\|_a^2 dt\end{aligned}$$

yielding the stability estimate

$$\left\|\widehat{U}_m^L(t_m)\right\|^2 + 2\left\|[\widehat{U}^L]_{m-1}\right\|^2 + \int_{t_{m-1}}^{t_m}\left\|\widehat{U}_m^L\right\|_a^2 dt \le \left\|\widehat{U}_{m-1}^L(t_{m-1})\right\|^2 + \int_{t_{m-1}}^{t_m}\|g\|_{a,*}^2\, dt. \tag{5.38}$$

Adding these estimates for $m = 1, \dots, M$ gives

$$\left\|\widehat{U}_M^L(T)\right\|^2 + 2\sum_{m=0}^{M-1}\left\|[\widehat{U}^L]_m\right\|^2 + \int_0^T\left\|\widehat{U}^L\right\|_a^2 dt \le \left\|\widehat{u}_0^L\right\|^2 + \int_0^T\|g\|_{a,*}^2\, dt. \tag{5.39}$$

Lemma 5.8. *For* $\mathbf{u}$ *and* $\mathbf{f}$ *in* (5.13) *holds the stability bound*

$$\frac{k}{2}\|\mathbf{u}\|_{\mathbf{I}\otimes\mathbf{A}}^2 \le \|\mathbf{u}_-\|_{\mathbf{M}}^2 + \frac{k}{2}\left\|\mathbf{f}^1\right\|_{\mathbf{I}\otimes\mathbf{A}^{-1}}^2 \tag{5.40}$$

where $\mathbf{u}_-$ *denotes the coefficient vector of* $\widehat{U}_m^L(t_{m-1}) \in \widehat{V}^L$ *(here we used only the third term on the left hand side in* (5.38)*).*

Proof. Let $\mathbf{w}$ be the coefficient vector of a function $\widehat{w}^L \in \widehat{V}^L$, let $\mathbf{f}$ be the coefficient vector of a functional $\widehat{f}^L \in (\widehat{V}^L)^*$, then

$$\left\|\widehat{w}^L\right\|_V \sim \left\|\widehat{w}^L\right\|_a = \|\mathbf{w}\|_{\mathbf{A}}, \qquad \left\|\widehat{w}^L\right\|_H = \|\mathbf{w}\|_{\mathbf{M}}, \qquad \left\|\widehat{f}^L\right\|_{a,*} = \|\mathbf{f}\|_{\mathbf{A}^{-1}}$$

We can also express norms of functions of x and t in terms of the coefficient vector: We have for $\widehat{U}_m^L \in \mathcal{Q}$ and the corresponding coefficient vector $\mathbf{u}$

$$\left\|\widehat{U}_m^L\right\|_{\mathcal{Q}}^2 = \int_{t_{m-1}}^{t_m}\left\|\widehat{U}_m^L\right\|^2 dt = \frac{k}{2}\|\mathbf{u}\|_{\mathbf{I}\otimes\mathbf{A}}^2 \tag{5.41}$$

since $\left\|\varphi_i \circ F_m^{-1}\right\|_{L^2(I_m)} = \sqrt{\frac{k}{2}}$. Using this the stability estimate (5.38) implies for (5.13) the bound (5.40) □

We apply (5.34) to the error equation and obtain that

$$\left\|(\lambda_{j+1}\mathbf{M} + \tfrac{k}{2}\mathbf{A})\tilde{\mathbf{w}}_j - \tilde{\mathbf{s}}_j\right\|_{\mathbf{D}_A^{-1}} \le C_{h,k,r}q^{n_G}\|\tilde{\mathbf{s}}_j\|_{\mathbf{D}_A^{-1}} \tag{5.42}$$

where

$$\tilde{\mathbf{s}}_j := \mathbf{g}_j - \sum_{k=j+1}^{r}\mathbf{T}_{j+1,k+1}\mathbf{M}\tilde{\mathbf{w}}_k, \qquad C_{h,k,r} = Ch^{-1}(1 + k^{1/2}r), \qquad q = 1 - cr^{-2\alpha}$$

Let $\mathbf{T}_u$ denote the upper triangular part of $\mathbf{T}$, then we have $\mathbf{s} = \mathbf{g} - (\mathbf{T}_u \otimes \mathbf{M})\tilde{\mathbf{w}}$. Adding the squares of estimates (5.42) together for $r = 0, \dots, j$ gives with $\mathbf{R}_0 := (\mathbf{T}\otimes\mathbf{M} + \frac{k}{2}\mathbf{I}\otimes\mathbf{A})$

$$\|\mathbf{R}_0\tilde{\mathbf{w}} - \mathbf{g}\|_{\mathbf{I}\otimes\mathbf{D}_A^{-1}} \le C_{h,k,r}q^{n_G}\|\mathbf{g} - (\mathbf{T}_u\otimes\mathbf{M})\tilde{\mathbf{w}}\|_{\mathbf{I}\otimes\mathbf{D}_A^{-1}}. \tag{5.43}$$

Since $\|\mathbf{M}\mathbf{x}\|_{\mathbf{D}_A^{-1}} \le C\,\|\mathbf{x}\|_{\mathbf{D}_A}$ and (5.20) we have $\|(\mathbf{T}_u \otimes \mathbf{M})\tilde{\mathbf{w}}\|_{\mathbf{I}\otimes\mathbf{D}_A^{-1}} \le Cr^2\,\|\tilde{\mathbf{w}}\|_{\mathbf{I}\otimes\mathbf{D}_A}$. By (5.40) we have $\left\|\mathbf{R}^{-1}\mathbf{f}^1\right\|_{\mathbf{I}\otimes\mathbf{D}_A} \le 2k^{-1}\left\|\mathbf{f}^1\right\|_{\mathbf{I}\otimes\mathbf{D}_A^{-1}}$ and therefore $\|\tilde{\mathbf{w}}\|_{\mathbf{I}\otimes\mathbf{D}_A} \le 2k^{-1}\,\|\mathbf{R}_0\tilde{\mathbf{w}}\|_{\mathbf{I}\otimes\mathbf{D}_A^{-1}}$, yielding

$$\|\tilde{\mathbf{w}}\|_{\mathbf{I}\otimes\mathbf{D}_A} \le 2k^{-1}\,\|\mathbf{R}_0\tilde{\mathbf{w}}\|_{\mathbf{I}\otimes\mathbf{D}_A^{-1}} \le 2k^{-1}\big(\|\mathbf{g}\|_{\mathbf{I}\otimes\mathbf{D}_A^{-1}} + \|\mathbf{R}_0\tilde{\mathbf{w}}-\mathbf{g}\|_{\mathbf{I}\otimes\mathbf{D}_A^{-1}}\big). \tag{5.44}$$

Combining these estimates with (5.43) gives

$$\|\mathbf{R}_0\tilde{\mathbf{w}}-\mathbf{g}\|_{\mathbf{I}\otimes\mathbf{D}_A^{-1}} \le C_{h,k,r}q^{n_G}\Big[(1+Ck^{-1}r^2)\,\|\mathbf{g}\|_{\mathbf{I}\otimes\mathbf{D}_A^{-1}} + Ck^{-1}r^2\,\|\mathbf{R}_0\tilde{\mathbf{w}}-\mathbf{g}\|_{\mathbf{I}\otimes\mathbf{D}_A^{-1}}\Big]$$

If n_G is so large that $C_{h,k,r}q^{n_G}Ck^{-1}r^2 \le \frac{1}{2}$ we have with $C'_{h,k,r} := C_{h,k,r}k^{-1}r^2 = Ch^{-1}(1+k^{1/2}r)k^{-1}r^2$

$$\begin{aligned}\|\mathbf{R}_0\tilde{\mathbf{w}}-\mathbf{g}\|_{\mathbf{I}\otimes\mathbf{D}_A^{-1}} &\le C'_{h,k,r}q^{n_G}\,\|\mathbf{g}\|_{\mathbf{I}\otimes\mathbf{D}_A^{-1}}\\ \|\mathbf{R}\tilde{\mathbf{z}}-\mathbf{f}\|_{\mathbf{I}\otimes\mathbf{D}_A^{-1}} &\le C'_{h,k,r}q^{n_G}\,\|\mathbf{f}\|_{\mathbf{I}\otimes\mathbf{D}_A^{-1}}\end{aligned}$$

for the resulting approximation $\tilde{\mathbf{z}} := (\mathbf{Q}^H \otimes \mathbf{I})\tilde{\mathbf{w}}$ of $\mathbf{z}$ since $\mathbf{Q}$ is unitary.

To translate this estimate to DG setting, consider the residual ρ_m defined by

$$\begin{aligned}\langle\rho_m, V\rangle :=& \int_{t_{m-1}}^{t_m}\Big[\Big((\tilde{U}_m^L)', V\Big) + a(\tilde{U}_m^L, V)\Big]dt + \Big(\tilde{U}_m^L(t_{m-1}), V(t_{m-1})\Big)\\ &- \int_{t_{m-1}}^{t_m}(g,V)_{V^*\times V}\,dt - \Big(\tilde{U}_{m-1}^L(t_{m-1}), V(t_{m-1})\Big)\\ =& \int_{t_{m-1}}^{t_m}\Big[\Big(\tilde{Z}', V\Big) + a(\tilde{Z}, V)\Big]dt + \Big(\tilde{Z}(t_{m-1}), V(t_{m-1})\Big)\\ &- \int_{t_{m-1}}^{t_m}(g,V)_{V^*\times V}\,dt + \int_{t_{m-1}}^{t_m} a(\tilde{U}_{m-1}^L(t_{m-1}), V)dt.\end{aligned}$$

We have with

$$\left|\int_{t_{m-1}}^{t_m} a(\tilde{U}_{m-1}^L(t_{m-1}), V)dt\right| \le \left\|\tilde{U}_{m-1}^L(t_{m-1})\right\|_{\mathcal{Q}}\|V\|_{\mathcal{Q}} \le k^{1/2}\left\|\tilde{U}_{m-1}^L(t_{m-1})\right\|_{\mathcal{Q}}\|V\|_V$$

that

$$\|\rho_m\|_{\mathcal{Q}*} \le C'_{h,k,r}q^{n_G}\left(\|g\|_{\mathcal{Q}*} + k^{1/2}\left\|\tilde{U}_{m-1}^L(t_{m-1})\right\|_V\right) \tag{5.45}$$

5.5.4 Incomplete GMRES iteration

We now estimate the error of the approximations $\tilde{U}_1^L,\dots,\tilde{U}_M^L$ compared to $\widehat{U}_1^L,\dots,\widehat{U}_M^L$ where all linear systems (5.13) are solved exactly. The differences $\zeta_m := \tilde{U}_m^L - \widehat{U}_m^L$ satisfy

$$\begin{aligned}&\int_{t_{m-1}}^{t_m}\Big[\Big(\zeta_m', V\Big) + a(\zeta_m, V)\Big]dt + \Big(\zeta_m(t_{m-1}), V(t_{m-1})\Big)\\ &\qquad = \int_{t_{m-1}}^{t_m}(\rho_m, V)_{V^*\times V}\,dt + \Big(\zeta_{m-1}(t_{m-1}), V(t_{m-1})\Big)\end{aligned}$$

where $\zeta_0(t_0) := 0$. With (5.39) we obtain from (5.45) for $l = 1,\dots,M$

$$\begin{aligned}E_l &:= \|\zeta_m(t_m)\|_V^2 + \sum_{m=1}^{l}\|\zeta_m\|_{\mathcal{Q}}^2\,dt \le \sum_{m=1}^{l}\|\rho_m\|_{\mathcal{Q}^*}^2\,dt\\ &\le (C'_{h,k,r})^2q^{2n_G}\sum_{m=1}^{l}\Big(\|g\|_{\mathcal{Q}^*}^2 + k\big\|\widehat{U}_{m-1}^L(t_{m-1})\big\|_V^2 + k\,\|\zeta_{m-1}(t_{m-1})\|_V^2\Big)\end{aligned}$$

We denote the right hand side of (5.39) with R and obtain with the inverse inequality

$$\begin{aligned}\left\|\widehat{U}^L_{m-1}(t_{m-1})\right\|_V^2 &\le Ch^{-2}\left\|\widehat{U}^L_{m-1}(t_{m-1})\right\|_H^2 \le Ch^{-2}R\\ \|\zeta_{m-1}(t_{m-1})\|_V^2 &\le Ch^{-2}\,\|\zeta_{m-1}(t_{m-1})\|_H^2 \le Ch^{-2}E_{l-1}\end{aligned}$$

yielding

$$E_l \le (C'_{h,k,r})^2 q^{2n_G} C\Big((1+h^{-2})R + k\sum_{m=1}^{l-1} h^{-2}E_m\Big)$$

Therefore we have estimates of the form $E_l \le \mu + \nu\sum_{m=1}^{l-1} E_m$ for $l = 1, 2, \ldots$ from which we get by induction $E_l \le \mu(1+\nu)^{l-1}$. Here we have $\nu = (C'_{h,k,r})^2 q^{2n_G} h^{-2}T/M$. We choose the number n_G of GMRES steps so that $\nu = (C'_{h,k,r})^2 q^{2n_G} h^{-2} \le 1$ and get $(1+\nu)^M \le e^T$ and

$$E_M \le (C'_{h,k,r})^2 q^{2n_G} Ch^{-2}e^T R\,. \tag{5.46}$$

Finally we choose n_G large enough so that the resulting bound for $\|\zeta_M(T)\|_H = \left\|\tilde{U}^L(T) - \widehat{U}^L(T)\right\|_H$ is less than the bound in Corollary 5.2: We need n_G such that

$$C'_{h,k,r}h^{-2}q^{n_G} \le Ch^{-3}k^{-1}r^3 q^{n_G} \le Ch^{\theta_0 p+\delta} \tag{5.47}$$

Since $q = 1 - cr^{-2\alpha}$ we require $n_G \ge Cr^{2\alpha}(|\log h| + |\log k|) = C'\,|\log h|^{1+2\alpha}$ using $r = O(|\log h|)$ and $|\log k| \le Cr$. We obtain

Theorem 5.9. *Under the assumptions of Theorem 5.1 choose the number and order of time steps such that $M = r = O(|\log h|)$ and in each time step $n_G = O(|\log h|)^{1+2\alpha}$ GMRES iterations. Then the fully discrete Galerkin scheme with incomplete GMRES gives a approximate solution $\tilde{U}^L$ satisfying*

$$\left\|u(T) - \tilde{U}^L(T)\right\|_H \le C(u_0, g)h^{\theta_0 p+\delta} \tag{5.48}$$

with $C(u_0, g)$ as in Corollary 5.2 and δ as in (3.34). *For solutions which are smooth in x for $t > 0$, $\theta_0 = 1$.*

5.6 Implementation of matrix-vector products and complexity

For the iterative solution of the linear systems (5.16) with GMRES we have to compute matrix vector products with the stiffness matrix $\mathbf{A}$ and the mass matrix $\mathbf{M}$. Note that these matrices are densely populated since most basis functions have large supports. Naive implementation would therefore yield a complexity which is at least $O(\widehat{N}_L^2)$. In addition, the functions $D(x)$ and $c(x)$ used in the definition of the stiffness matrix require in general numerical integration, and the use of a standard fixed order tensor product quadrature requires $O(h^{-d})$ operations.

For our approximate matrix-vector multiplication we assume that the functions $c(x)$ and $D_{ij}(x)$ are analytic in $x \in \overline{\Omega} = [0,1]^d$. We approximate each of the functions $c(x)$ and $D_{ij}(x)$ for $i, j = 1, \ldots, d$ by polynomials in $(x_1, \ldots, x_d)$ of degree at most q in each of the variables $x_1, \ldots, x_d$. This can be done by interpolation at Chebyshev nodes in each direction using $O(q^d)$ operations.

Let us first consider a mass matrix for a single monomial $c(x) = c_\alpha x_1^{\alpha_1}\cdots x_d^{\alpha_d}$. We exploit the tensor product structure by first computing for $k = 1, \ldots, d$ the band matrices $C^{(k)}_{i,j} = \int_0^1 \varphi_i^L(x_k)\varphi_j^L(x_k)x_k^{\alpha_k}\,dx_k$ with the scaling functions φ_j^L which form a basis for $\mathcal{V}^L$. For a function $v^L \in \widehat{V}^L$ with coefficients $v^{\ell_1\ldots\ell_d}_{j_1\ldots j_d}$ we first let $k = 1$ and iterate over all values of $(\ell_2, j_2), \ldots, (\ell_d, j_d)$. For each of those values we transform the resulting vector in (ℓ_1, j_1) from the wavelet basis to the basis of scaling functions, apply the band matrix $C^{(k)}$, and transform the result back to the wavelet basis. We then repeat this procedure with respect to the dimensions $k = 2, \ldots, d$. The total number of operations is then bounded by $Cd\widehat{N}_L$.

In this way we can implement an approximate matrix-vector product with the mass matrix $\mathbf{M}$ in $Cq^d d\widehat{N}_L$ operations, and the matrix-vector product with the stiffness matrix $\mathbf{A}$ in $Cq^d d^3\widehat{N}_L$ (since $D(x)$ is a $d \times d$ matrix).

By the analyticity assumption on $c(x), D_{ij}(x)$ the L^∞ error of the polynomial interpolations for $c(x)$ and $D(x)$ decreases exponentially with q (see Lemma 3.6 in [11]) and we can preserve the error bound in (5.48) by choosing $q = O(|\log h|)$.

Theorem 5.10. *Under the assumptions of Theorem 5.1 and 5.9 and for coefficients $c(x)$ and $D_{ij}(x)$ which are analytic in $\overline{\Omega}$ we can compute an approximation to $\tilde{U}^L(T)$ which also satisfies the bound* (5.48) *with at most $C_{p,d}h^{-1}|\log h|^{d-1+2+(1+2\alpha)+d}$ operations.*

Proof. We have $O(r)$ time steps. During each time step we solve the $r+1$ linear systems (5.16) using $n_G = O(|\log h|^{1+2\alpha})$ GMRES iterations in each case. Each GMRES iteration involves a matrix-vector product with $\mathbf{M}$ and with $\mathbf{A}$, costing $Cq^d d^3 \widehat{N}_L$ operations. Therefore the total number of operations is bounded by

$$Cr^2 n_G q^d d^3 \widehat{N}_L$$

where $r = O(|\log h|)$, $\widehat{N}_L = O(h^{-1}|\log h|^{d-1})$ and $q = O(|\log h|)$. □

Remark 5.11. We also obtain results for a fixed order $r \geq 0$ and $M \to \infty$, see Remark 4.13 (the simplest case $r = 0$ corresponds to the backward Euler method). Then we obtain with the algebraically graded time step sequence (4.29) instead of (5.6) the bound

$$\|u(T) - \widehat{U}^L(T)\| \leq C_1(u_0, g)h^{\theta_0 p+\delta} + C_2(u_0, g)M^{-r-1}.$$

To equilibrate spatial and temporal error, we use $M = O(h^{-\beta})$ time steps with $\beta = \frac{\theta_0 p+\delta}{r+1}$. We obtain from (5.47) that the number n_G of GMRES steps has to satisfy $n_G \geq Cr^{2\alpha}|\log h|$. Hence we can compute an approximation $\tilde{U}(T)$ satisfying (5.48) using a total of $Crh^{-\beta}r^{2\alpha}|\log h|$ GMRES iterations, and a total number

$$Cr^{1+2\alpha}h^{-1-\beta}|\log h|^d$$

of operations. Note that this number grows superlinearly in $\widehat{N}_L$.

Remark 5.12. We may express Theorems 5.9 and 5.10 in terms of the number $\widehat{N}_L$ of degrees of freedom: for solutions $u(x,T)$ in x for $T > 0$, we obtain for $d > 1$ an approximation $\tilde{U}^L(T)$ with

$$\left\|u(T) - \tilde{U}^L(T)\right\|_{L^2(\Omega)} \leq C(u_0, g)\widehat{N}_L^{-p-\delta+\varepsilon}$$

for any $\varepsilon > 0$ in $O(\widehat{N}_L)$ operations. This follows by absorbing powers of $\log \widehat{N}_L$ in the term $O(h^\varepsilon)$ in (5.48). In dimension $d = 1$ we have $\theta_0 = 1$, $\delta = 1$, $\widehat{N}_L = N_L$ and $h = 1/N_L$. Our scheme gives for any $T > 0$ an approximate solution $\tilde{U}^L(T)$ with $L^2(\Omega)$ convergence rate $O(N_L^{-(p+1)})$ in work $O(N_L(\log N_L)^c)$. Any time-stepping scheme of fixed order $q \geq 1$, e.g., of BDF type, will require $M = O(N_L^{(p+1)/q})$ time steps and, therefore, a total work of at least $O(N_L^{1+(p+1)/q})$ operations.

6 Numerical Results

In order to have exact solutions at our disposal, we consider the problem $u_t - \Delta u = 0$ in $\Omega = (0,1)^d$. We use $p = 1$ and the piecewise linear wavelets described in section 3.2. All computations were performed in double precision arithmetic on a PC with 2GB RAM in MATLAB 6.1.

For this problem we have, by Remark 3.6, that $\theta_0 = 1$. We compute $\tilde{U}^L(T)$, the fully discrete solution with GMRES approximation at $t = T$. Theorem 5.9 yields a convergence rate

$$\left\|u(T) - \tilde{U}^L(T)\right\|_{L^2(\Omega)} \leq C(u_0, g)h^{\theta_0 p+\delta} = Ch^{1+\frac{1}{2d-1}} \tag{6.1}$$

for $r = c|\log h|$ and $n_G = c'|\log h|^{7/3}$, using c, c' sufficiently large and the value of α suggested by Remark 5.6.

We want to illustrate the effect of $h = 2^{-L-1}$, the number M of time steps and the degree r of the DG method on the error of $u(x,T)$. Therefore we chose a large fixed value for the number n_G of GMRES iterations, so that all the errors shown in the tables below were insensitive to the iteration error.

In order to compute the error $\left\|u(T) - \tilde{U}^L(T)\right\|_{L^2(\Omega)}$ we proceed as follows: We have $\left\|u(T) - \Pi^L u(T)\right\|_{L^2(\Omega)} \leq Ch^2$ where Π^L denotes the interpolation operator for the *full grid* space V^L. Therefore it is sufficient to measure the error $E := \tilde{U}^L(T) - \Pi^L u(T)$. Since $E \in V^L$ the norm $\|E\|_{L^2(\Omega)}$ is equivalent to the norm $|\!|\!|E|\!|\!|_0$ in (3.18) which uses the wavelet coefficients. In all tables we use $|\!|\!|E|\!|\!|_0 \,/\, |\!|\!|\Pi^L u(T)|\!|\!|$ to measure the relative L^2 error.

6.1 Smooth solution

L	$r=0$	$r=1$	$r=2$	$r=3$	$r=4$	$r=5$	$r=6$
1	2.24679	0.632236	0.137173	0.161018	0.160037	0.16001	0.160017
2	2.35824	0.485767	0.0214951	0.0478435	0.0468904	0.0469531	0.0469138
3	2.38988	0.442895	0.0165634	0.0132888	0.0124359	0.0124669	0.0124489
4	2.39803	0.431726	0.0244065	0.00401679	0.00318958	0.00320808	0.00320337
5	2.40007	0.428907	0.0265104	0.00164255	0.00080374	0.000818954	0.000817462
6	2.40058	0.4282	0.0270462	0.00105846	0.000192459	0.000207351	0.000207416
7	2.40071	0.428024	0.0271778	0.000916968	$5.66855 \cdot 10^{-5}$	$5.59876 \cdot 10^{-5}$	$5.36717 \cdot 10^{-5}$
8	2.40074	0.42798	0.0272021	0.000883517	$1.50843 \cdot 10^{-5}$	$5.56368 \cdot 10^{-6}$	$5.32777 \cdot 10^{-7}$

Table 1: Smooth solution, $M = 1$ time step: relative L^2 error for $d = 5$ at $T = 0.05$

We choose the initial condition $u_0(x) = \sin(\pi x_1) \cdots \sin(\pi x_d)$. Note that $u_0 \in H_s$ for any $s > 0$, and the exact solution $u(x,t) = e^{-d\pi^2 t} \sin(\pi x_1) \cdots \sin(\pi x_d)$ is analytic with respect to $t \in [0,T]$. Due to Remark 4.14 we can use just one subinterval $[t_0, t_1] = [0,T]$ for the DG time stepping.

To show the typical convergence behavior with respect to $h = 2^{-L-1}$ and r we choose dimension $d = 5$ and $T = 0.05$. The results are shown in Table 1.

In this table the limits of the rows for $r \to \infty$ correspond to the errors of the space-semidiscretization which were analyzed in section 3. If we consider the column with $r = 6$ and compute experimental convergence rates h^α we obtain $\alpha = 1.77, 1.91, 1.96, 1.98, 1.95, 6.65$. In Theorem 3.10 we obtained $O(h^{1+1/(2d-1)}) = O(h^{10/9})$ which seems to be too pessimistic.

The limits of the columns for $L \to \infty$ correspond to the errors of the time-semidiscretization which were analyzed in section 4. The fast decay of the values in the row $L = 8$ for $r = 0, 1, 2, 3, 4$ appears to be compatible with the exponential convergence in Remark 4.14.

Theorem 5.9 considers the case where we choose r proportional to L. If we choose $r = \lceil L/2 \rceil + 1$ we obtain the experimental convergence rates h^α with $\alpha = 1.52, 1.85, 2.06, 1.99, 1.95, 1.89, 6.72$. Again, the result $O(h^{10/9})$ of the theorem seems to be too pessimistic.

6.2 Solution with singularity at $t = 0$

We now choose the initial condition $u_0(x) = 1$. Note that $u_0 \in H_\theta$ only for $\theta < \frac{1}{2}$, so we will get a solution with singular behavior for $t \to 0$. Using the sine series $1 = \sum_{k \text{ odd}}^\infty \frac{4}{k\pi} \sin(k\pi x)$ we obtain that the problem for $d = 1$ has the solution

$$v(x,t) = \sum_{k \text{ odd}}^\infty \frac{4}{k\pi} e^{-k^2\pi^2 t} \sin(k\pi x).$$

The exact solution for $d > 1$ is $u(x,t) = v(x_1,t) \cdots v(x_d,t)$. Due to the boundary incompatibility u has strong boundary layers for small t, causing high spatial approximation errors (even for a full grid approximation). Because of the singularity at $t = 0$ a single subinterval $[0,T]$ for the DG method cannot give exponential convergence, and we require the geometric time mesh analyzed in section 4. In our computations we chose the grading factor of the geometric mesh (see Definition 4.7) as $\sigma = 1/2$, and we used the same order r for all time steps, cf. Remark 4.12.

L	$r=0$	$r=1$	$r=2$	$r=3$	$r=4$
1	1.39384	0.414791	0.370531	0.370882	0.370885
2	2.07804	0.164257	0.114436	0.114707	0.114713
3	2.33455	0.0914775	0.0300100	0.0302919	0.0302921
4	2.41969	0.0823793	0.00748878	0.00761493	0.00761010
5	2.44570	0.0832414	0.00257669	0.00190387	0.00189491
6	2.45334	0.0843352	0.00216938	0.000492105	0.000474795
7	2.45554	0.0848729	0.00221850	0.000174178	0.000102090
8	2.40074	0.0850852	0.00224571	0.000114169	$3.9837 \cdot 10^{-5}$

Table 2: Singular solution, geometric time mesh with $M = r+1$ intervals: relative L^2 error for $d = 5$ at $T = 0.05$

L	$r=2$	$r=3$	$r=4$	$r=5$	$r=6$
2	1.64777	0.467259	0.450336	0.450954	0.450946
3	3.20442	0.198088	0.178191	0.178786	0.178831
4	4.83159	0.089031	0.0570031	0.0579576	0.0580453
5	6.34142	0.086463	0.0479329	0.0466059	0.0454513

Table 3: Singular solution, geometric time mesh with $M = r$ intervals: relative L^2 error for $d = 15$ at $T = 0.05$

Numerical results for $d = 5, 15, 20$ are presented in Tables 2, 3, 4. Note that the convergence behavior with respect to L and r is the similar as in Table 1: E.g., in Table 2 for $d = 5$ we obtain in the column for $r = 4$ the experimental convergence rates $\alpha = 1.69, 1.92, 1.99, 2.01, 2.00, 2.21, 1.36$ which is even better than the rate $O(h^{10/9})$ predicted in Theorem 3.10 for the space discretization.

Theorem 5.9 analyzes the algorithm with $r = O(L)$, $M = O(L)$. If we choose $r = \lceil L/2 \rceil$ we obtain for $d = 3, 5, 10$ the values shown in Table 5. Results for $d = 15, 20, 25$ are shown in Table 6. The experimental convergence rates are even better than the rate $O(h^{1+1/(2d-1)})$ of (6.1).

References

[1] H. Amann: Linear and Quasilinear Parabolic Problems 1: Abstract Linear Theory, Birkhäuser, Basel, 1995.

[2] H.-J. Bungartz, M. Griebel: A note on the complexity of solving Poisson's equation for spaces of bounded mixed derivatives, J. Complexity **15** (1999), 167–199.

[3] S. C. Eisenstat, H. C. Elman, M. H. Schultz: Variational iterative methods for nonsymmetric systems of linear equations. SIAM J. Numer. Anal. **20** (2) (1983) 345–357.

[4] M. Griebel and S. Knapek: Optimized Tensor Product Approximation Spaces, Constructive Approximation **16** (2000) 525-540.

L	$r=2$	$r=3$	$r=4$	$r=5$	$r=6$
2	2.89635	0.519846	0.518719	0.518732	0.518739
3	5.80581	0.255009	0.250773	0.250753	0.250738
4	8.88418	0.14622	0.105516	0.104774	0.106068
5	11.8563	0.193875	0.042089	0.041238	0.042886

Table 4: Singular solution, geometric time mesh with $M = r-1$ intervals: relative L^2 error for $d = 20$ at $T = 0.025$

L	$d=3$	α	$d=5$	α	$d=10$	α
1	0.235762		0.414791		1.0157	
2	0.0696534	1.76	0.164257	1.34	1.04877	
3	0.0159470	2.13	0.0300100	2.45	0.0654917	
4	0.00400648	1.99	0.00748878	2.00	0.0206558	1.66
5	0.00100063	2.00	0.00190387	1.98	0.00706261	1.55
6	0.000257924	1.96	0.000492105	1.95	0.00296293	1.25
7	$6.37306 \cdot 10^{-5}$	2.02	0.00010209	2.27		
8	$1.3902 \cdot 10^{-5}$	2.20	$3.9837 \cdot 10^{-5}$	1.36		
9	$3.91216 \cdot 10^{-6}$	1.83				

Table 5: Singular solution: relative L^2 error and experimental convergence rates h^α using $r = \lceil L/2 \rceil$, $M = r+1$, $T = 0.05$

L	$d=15$	α	$d=20$	α	$d=25$	α
1	0.738591		0.847683		0.913527	
2	0.377222		0.518732		0.642398	
3	0.154434	1.28	0.250753	1.04	0.358738	0.84054
4	0.056111	1.46	0.104774	1.25	0.17239	1.0573
5	0.019697	1.51	0.041238	1.34	0.0650509	1.4060

Table 6: Singular solution: relative L^2 error and experimental convergence rates h^α using $M = 4$, $r = 5$ and $T = 0.025$ (except for $d = 25$ and $L = 5$ where $M = 3$, $r = 4$ was used)

[5] M. Griebel, P. Oswald, T. Schiekofer: Sparse grids for boundary integral equations. Numerische Mathematik **83** (1999) 279-312

[6] J.L. Lions and E. Magenes: Non-homogeneous boundary value problems and applications I, Springer-Verlag 1972.

[7] P. Oswald: On best N-term approximation by Haar functions in H^s-norms. In: Metric Function Theory and Related Topics in Analysis (S. M. Nikolskij, B. S. Kashin, A. D. Izaak, eds.), AFC, Moscow, 1999, 137-163 (in Russian).

[8] H.C. Öttinger: Stochastic Processes in polymeric fluids, Springer Verlag 1998.

[9] A. Pazy, Semigroups of Linear Operators and Applications to Partial Differential Equations, Applied Mathematical Sciences 44, Springer-Verlag, New York 1983.

[10] T. von Petersdorff and C. Schwab: Wavelet-discretizations of parabolic integro-differential equations, SIAM J. Numer. Anal. **41** (2003) 159-180.

[11] G. Schmidlin, C. Lage and C. Schwab: Rapid solution of first kind boundary integral equations in $\mathbb{R}^3$, Engineering Analysis with Boundary Elements **27** No. 5 (2003), 469–490

[12] D. Schötzau: hp-DGFEM for Parabolic Evolution Problems, Dissertation ETH Zurich, 1999.

[13] D. Schötzau and Ch. Schwab: Time Discretization of parabolic problems by the *hp*-version of the discontinuous Galerkin finite element method SIAM J. Numer. Analysis **38** (2000) 837-875.

[14] D. Schötzau and Ch. Schwab: *hp*-Discontinuous Galerkin time-stepping for parabolic problems C.R.Acad. Sci. Paris **333** (2001) 1121-1126.

[15] Ch. Schwab: *p* and *hp* Finite Element Methods, Oxford University Press, 1998.

[16] Ch. Schwab and R.A. Todor: Sparse finite elements for stochastic elliptic problems-higher order moments (in press in Computing 2003), http://www.math.ethz.ch/research/groups/sam/reports/2003

[17] V. Thomee: Galerkin Finite Element Methods for Parabolic Problems, Springer Verlag 1997.

[18] T. Werder, D. Schötzau, K. Gerdes and C. Schwab: *hp*-Discontinuous Galerkin time-stepping for parabolic problems, Comp. Meth. Appl. Mech. Engg. **190** (2001) 6685-6708.