Eidgenössische Technische Hochschule Zürich

Ecole polytechnique fédérale de Zurich
Politecnico federale di Zurigo
Swiss Federal Institute of Technology Zurich

Space-time wavelet finite element method for parabolic equations

R. Andreev*

Revised: May 2011

Research Report No. 2010-20
July 2010

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

*Support in part by SNF Grant No. PDFMP2-127034/1

Sparse space-time finite element discretization of parabolic equations

Roman Andreev*
Seminar for Applied Mathematics,
Rämistrasse 101, 8092 Zürich, Switzerland

April 28, 2011

Abstract

For a class of linear parabolic equations we propose a nonadaptive sparse space-time Galerkin least squares discretization. We formulate criteria on the trial and test spaces for the well-posedness of the corresponding Galerkin least squares solution. In order to obtain discrete stability uniformly in the discretization parameters, we allow test spaces which are suitably larger than the trial space. The problem is then reduced to a finite, overdetermined linear system of equations by a choice of bases. We present several strategies that render the resulting normal equations well-conditioned uniformly in the discretization parameters. The numerical solution is then shown to converge quasi-optimally to the exact solution in the natural space for the original equation. Numerical examples for the heat equation confirm the theory.

1 Introduction

Linear parabolic equations constitute an important class of evolutionary partial differential equations and are used to model various phenomena in physics, financial engineering, chemistry, biology, etc. Consequently, a sizable body of methods for their numerical solution has been developed (Lang, 2001; Thomée, 2006). However, most of these methods are variations on the "method of lines", and therefore fail to produce a priori error bounds under minimal regularity assumptions and are intrinsically hard to parallelize. To date, several authors have proposed simultaneous space-time discretization schemes for parabolic initial boundary value problems, we refer to (Horton & Vandewalle, 1995; Babuška & Janik, 1990; Schötzau, 1999; Griebel & Oeltz, 2007; Schwab & Stevenson, 2009; Stevenson & Chegini, 2010) and further references therein. Using the variational formulation of (Schwab & Stevenson, 2009) we present a non-adaptive sparse space-time discretization procedure which builds upon fairly classical numerical tools of the finite element method, allows to prove quasi-optimality of the numerical solution and has the potential to significantly reduce the computational cost for sufficiently smooth solutions. The approach is fundamentally different

*andreevr@math.ethz.ch, support by SNF Grant No. PDFMP2-127034/1

from (Stevenson & Chegini, 2010) where a wavelet adaptive scheme involving a tensor construction of suitable wavelet bases is applied to the resulting *biinfinite* normal equations; or from (Babuška & Janik, 1990) where quasi-optimality is obtained for an *hp* continuous Galerkin time stepping with a constant *not* uniform in the discretization parameter, see (Babuška & Janik, 1990, Theorem 3.6.1), exploiting the setting of time-*independent* coefficients in an essential manner. In Example 4.3 we indicate, but do not elaborate on further here, how the construction of wavelets can be effectively bypassed by means of the well-known BPX operator. Moreover, we work with space- and time-dependent inputs throughout. We mention a few further connections to existing works: a general framework for least-squares finite element methods is advocated in (Bochev & Gunzburger, 2009), but our exposition is self-contained and confined to Theorem 3.1 and Section 4; in the language of (Demkowicz & Gopalakrishnan, 2011), we identify some test spaces in which "optimal test functions" for the parabolic problem can be sufficiently well approximated; the preconditioner proposed here for the discrete system an instance of the methodology of "operator" (Hiptmair, 2006) or "canonical" (Mardal & Winther, 2010) preconditioning.

The paper is organized as follows. In Section 2 we introduce the problem and present a well-posed variational formulation in suitable Bochner spaces. In Section 3 we motivate a notion of Galerkin least squares solution w.r.t. a pair of trial and test spaces. We present abstract criteria for its unique existence, and the quasi-optimality property. In Section 4 we discuss how, once a suitable pair of trial and test spaces is available, the Galerkin least squares solution can be obtained via discrete linear least squares equations, for which we construct optimal preconditioners. Section 5 introduces and investigates the properties of a particular construction of trial and test spaces based on space-time sparse tensor product spaces, to which the abstract theory of Section 3 applies. In Section 6 we specialize on the one-dimensional heat equation and give specific examples of suitable tensor product bases. We exemplify the theory developed in these sections in Section 7 by providing a series of numerical examples illuminating various aspects of the method. Section 8 summarizes the results of the paper.

We now introduce some notation used throughout. By Id we denote the identity or the injection map between spaces which will be clear from the context. For $i, j \in \mathbb{Z}$ we set $\delta_{ij} := 1$ if $i = j$ and $\delta_{ij} := 0$ if $i \neq j$. By $\mathbb{N}$ ($\mathbb{R}_+$) we denote the set of positive integers (reals) and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ ($\mathbb{R}_{\geq 0} := \mathbb{R}_+ \cup \{0\}$). Boldface lower and upper case letters denote vectors and matrices, possibly (bi-)infinite, e.g. $\mathbf{u} \in \mathbb{R}^M$, $\mathbf{B} \in \mathbb{R}^{N \times M}$, where $M, N \in \mathbb{N} \cup \{\infty\}$; if $M = \infty$ then $\mathbb{R}^M$ is the set of sequences over $\mathbb{R}$, etc. For $N \in \mathbb{N}_0$ we define $[N] := \{n \in \mathbb{N} : n \leq N\}$ and $[N] := \mathbb{N}$ if $N = \infty$. We write $\|\cdot\|_{\mathbf{M}}$ for the vector norm induced by a symmetric positive (semi-)definite matrix $\mathbf{M} \in \mathbb{R}^{M \times M}$, $M \in \mathbb{N} \cup \{\infty\}$, i.e., $\|\mathbf{u}\|_{\mathbf{M}}^2 := \mathbf{u}^\top \mathbf{M} \mathbf{u}$ for $\mathbf{u} \in \mathbb{R}^M$; we set $\ell^2_{\mathbf{M}}([M]) := \{\mathbf{u} \in \mathbb{R}^M : \|\mathbf{u}\|_{\mathbf{M}} < \infty\}$ and $\ell^2([N]) := \{\mathbf{v} \in \mathbb{R}^N : \|\mathbf{v}\|^2_{\ell^2([N])} := \sum_{n \in [N]} |\mathbf{v}_n|^2 < \infty\}$ with the obvious norms. The condition number of a matrix $\mathbf{B} \in \mathbb{R}^{N \times M}$ with respect to the norms of $\ell^2([M])$ and $\ell^2([N])$, if $\mathbf{B}$ is invertible, is denoted by $\varkappa_2(\mathbf{B}) \in \mathbb{R} \cup \{\infty\}$. By $\otimes$ we denote the tensor product of Hilbert spaces, or their elements. For a Banach space $\mathcal{Z}$ (here, all Banach spaces are over the field of reals) we denote by $\|\cdot\|_{\mathcal{Z}}$ its norm and by $\mathcal{Z}'$ its continuous dual, i.e., the space of linear continuous real valued functionals on $\mathcal{Z}$; by $\langle \cdot, \cdot \rangle_{\mathcal{Z}' \times \mathcal{Z}}$ we denote the duality pairing on $\mathcal{Z}' \times \mathcal{Z}$. By $\langle \cdot, \cdot \rangle_{\mathcal{Z}}$ we denote the scalar product on a Hilbert space $\mathcal{Z}$. If a Banach

space V embeds continuously into another, H, we write $V \hookrightarrow H$. The symbol $\cong$ denotes identification via an isomorphism. For two Banach spaces $\mathcal{Z}$, $\widetilde{\mathcal{Z}}$ we denote by $\mathcal{L}(\mathcal{Z}, \widetilde{\mathcal{Z}})$ the space of linear continuous maps $\mathcal{Z} \to \widetilde{\mathcal{Z}}$ equipped with the operator norm $\|\cdot\|_{\mathcal{L}(\mathcal{Z},\widetilde{\mathcal{Z}})}$.

2 Problem setting and variational formulation

Let V and H be separable Hilbert spaces over reals s.t. $V \hookrightarrow H \cong H' \hookrightarrow V'$ is a Gelfand triple, i.e., the embeddings are continuous and dense, and the scalar product $\langle\cdot,\cdot\rangle_H$ is compatible with the duality pairing on $V' \times V$. Let $0 < T_{\text{end}} < \infty$ and set $J := (0, T_{\text{end}})$. We will work in the following setting.

Assumption 2.1. *For (a.e.) $t \in J$ we are given a bilinear form $a_t(\cdot,\cdot)$ on $V \times V$, such that for all $\zeta, \eta \in V$*

- $J \ni t \mapsto a_t(\zeta,\eta) \in \mathbb{R}$ *is measurable,*
- $|a_t(\zeta,\eta)| \le a_{\max}\|\eta\|_V\|\zeta\|_V$,
- $a_t(\zeta,\zeta) \ge a_{\min}\|\zeta\|_V^2 - c_0\|\zeta\|_H^2$,

where $0 < a_{\min} \le a_{\max} < \infty$ and $c_0 \ge 0$ are fixed. For simplicity, in the following we assume $c_0 = 0$, see discussion in (Schwab & Stevenson, 2009, Proof of Theorem 5.1). We further restrict our attention to symmetric operators, i.e., $a_t(\zeta,\eta) = a_t(\eta,\zeta)$ for all $\zeta, \eta \in V$.

Concerning the last two assumptions we remark that in the following, merely Theorem 3.9 is bounded by this restriction, which, however, can also be removed. Consider now the following abstract parabolic equation

$$\partial_t u(t) + a_t(u(t),\cdot) = g(t) \in V', \quad u(0) = h \in H, \tag{2.1}$$

for a.e. $t \in J$, where $g : J \to V'$ and $h \in H$ are given, and $u : J \to V'$ is the unknown. Here, and in the following, $\partial_t\cdot$ denotes the (weak) partial derivative w.r.t. the temporal variable.

Example 2.2. *Let $D \subset \mathbb{R}^d$, $d \in \mathbb{N}$ be an open domain with a Lipschitz boundary. Consider*

$$\partial_t u(t,x) - \operatorname{div}(q(t,x)\operatorname{grad} u(t,x)) = g(t,x), \quad (t,x) \in J \times D, \tag{2.2}$$

$$u(x,0) = h(x), \quad x \in D, \tag{2.3}$$

$$u(t,x) = 0, \quad (t,x) \in \partial J \times D, \tag{2.4}$$

where $q \in L^\infty(J \times D)$ is a space- and time-dependent coefficient q which may describe e.g. heat conductivity or material permeability, and the differential operators div *and* grad *are w.r.t. the spatial variable $x \in D$. The natural choice here is $V = H_0^1(D)$ and $H = L^2(D)$. For (a.e.) $t \in J$ define $a_t : V \times V \to \mathbb{R}$ by*

$$a_t(\zeta,\eta) = \int_D q(t,x)\operatorname{grad}\zeta(x)\cdot\operatorname{grad}\eta(x)dx \quad \textit{for all} \quad \zeta,\eta \in V. \tag{2.5}$$

Assumption 2.5 is satisfied if we assume

$$0 < a_{\min} := \operatorname*{ess\,inf}_{J\times D} q \le \operatorname*{ess\,sup}_{J\times D} q =: a_{\max} < \infty. \tag{2.6}$$

In order to obtain a well-posed variational formulation of (2.1) we follow (Schwab & Stevenson, 2009) and introduce the spaces

$$\mathcal{X} := L^2(J,V) \cap H^1(J,V') \cong (L^2(J) \otimes V) \cap (H^1(J) \otimes V')$$
$$\mathcal{Y} := \mathcal{Y}_1 \times \mathcal{Y}_2 := L^2(J,V) \times H \cong (L^2(J) \otimes V) \times H$$

with norms $\|\cdot\|_{\mathcal{X}}$ and $\|\cdot\|_{\mathcal{Y}}$ given by

$$\|u\|_{\mathcal{X}}^2 := \|u\|_{L^2(J,V)}^2 + \|\partial_t u\|_{L^2(J,V')}^2, \quad \|v\|_{\mathcal{Y}}^2 := \|v_1\|_{L^2(J,V)}^2 + \|v_2\|_H^2 \tag{2.7}$$

for all $u \in \mathcal{X}$ and $v = (v_1, v_2) \in \mathcal{Y}$. We note that

$$\sup_{0 \le t \le T_{\text{end}}} \sup_{u \in \mathcal{X} \setminus \{0\}} \frac{\|u(t)\|_H}{\|u\|_{\mathcal{X}}} < \infty, \tag{2.8}$$

i.e., $\mathcal{X} \hookrightarrow C^0(\overline{J}, H)$, in particular the trace map $(\cdot)|_{t=0} : \mathcal{X} \to H$, $u \mapsto u|_{t=0} = u(0) \in H$ is well-defined and continuous (Lions & Magenes, 1972, Chapter 1). Define the linear operator $B : \mathcal{X} \to \mathcal{Y}'$ by

$$(Bu)(v) = \int_J \{\langle \partial_t u(t), v_1(t) \rangle_{V' \times V} + a_t(u(t), v_1(t))\}\, dt + \langle u(0), v_2 \rangle_H \tag{2.9}$$

for all $u \in \mathcal{X}$ and $v = (v_1, v_2) \in \mathcal{Y}$, as well as the functional $f : \mathcal{Y} \to \mathbb{R}$ by

$$f(v) = \int_J \langle g(t), v_1(t) \rangle_{V' \times V} dt + \langle h, v_2 \rangle_H \tag{2.10}$$

for all $v = (v_1, v_2) \in \mathcal{Y}$. It is easy to check that $B \in \mathcal{L}(\mathcal{X}, \mathcal{Y}')$ and $f \in \mathcal{Y}'$. The variational formulation of (2.1) now reads:

$$\text{find} \quad u \in \mathcal{X} \quad \text{s.t.} \quad (Bu)(v) = f(v) \quad \forall v \in \mathcal{Y}. \tag{2.11}$$

We recall (Schwab & Stevenson, 2009, Theorem 5.1) for future reference.

Theorem 2.3. *The linear operator $B : \mathcal{X} \to \mathcal{Y}'$ and its inverse are continuous. In particular,* (2.11) *is well-posed (in Hadamard sense).*

3 Abstract stability results

The starting point for a stable discretization of (2.11) w.r.t. a pair of closed subspaces $\mathcal{U} \subseteq \mathcal{X}$ and $\mathcal{V} \subseteq \mathcal{Y}$ will be the *inf-sup condition* for the bilinear form $\langle B\cdot, \cdot \rangle_{\mathcal{Y}' \times \mathcal{Y}}$ on $\mathcal{U} \times \mathcal{V}$, i.e.,

$$\inf_{u \in \mathcal{U} \setminus \{0\}} \sup_{v \in \mathcal{V} \setminus \{0\}} \frac{\langle Bu, v \rangle_{\mathcal{Y}' \times \mathcal{Y}}}{\|u\|_{\mathcal{X}} \|v\|_{\mathcal{Y}}} =: \gamma_{\mathcal{U},\mathcal{V}} > 0. \tag{3.1}$$

Assuming existence of such subspaces, we motivate in Theorem 3.1 a notion of an approximate "Galerkin least squares solution" to (2.11) in the trial space $\mathcal{U} \subseteq \mathcal{X}$ w.r.t. a sufficiently large test space $\mathcal{V} \subseteq \mathcal{Y}$ and an auxiliary norm on $\mathcal{Y}$. This leads to Definition 3.2. After some preparations, we give an abstract criterion for (3.1) in Theorem 3.9.

Theorem 3.1. *Let $\mathcal{U} \subseteq \mathcal{X}$ and $\mathcal{V} \subseteq \mathcal{Y}$ be closed subspaces. Assume that the* inf-sup *condition* (3.1) *holds. Further, let $\langle \cdot, \cdot \rangle'_{\mathcal{Y}}$ be a scalar product on $\mathcal{Y}$, such that the induced norm $\|\cdot\|'_{\mathcal{Y}}$ on $\mathcal{Y}$ is equivalent to $\|\cdot\|_{\mathcal{Y}}$, that is there exist constants $0 < c \leq C < \infty$ s.t.*

$$\forall v \in \mathcal{Y}: \quad c\|v\|'_{\mathcal{Y}} \leq \|v\|_{\mathcal{Y}} \leq C\|v\|'_{\mathcal{Y}}. \tag{3.2}$$

Then there exists a unique $u' \in \mathcal{U}$ which satisfies

$$u' = \operatorname*{arg\,min}_{w \in \mathcal{U}} \sup_{v \in \mathcal{V}\setminus\{0\}} \frac{|\langle Bw - f, v\rangle_{\mathcal{Y}' \times \mathcal{Y}}|}{\|v\|'_{\mathcal{Y}}}. \tag{3.3}$$

Moreover, the a priori quasi-optimality estimate

$$\|u - u'\|_{\mathcal{X}} \leq \left(1 + \left(1 + \frac{C}{c}\right) \frac{\|B\|_{\mathcal{L}(\mathcal{X},\mathcal{Y}')}}{\gamma_{\mathcal{U},\mathcal{V}}}\right) \inf_{w \in \mathcal{U}} \|u - w\|_{\mathcal{X}} \tag{3.4}$$

holds, where $u = B^{-1}f \in \mathcal{X}$ solves (2.11).

Proof. For each $u \in \mathcal{X}$ let $v_u^\star \in \mathcal{V}$ denote the unique element which satisfies $\langle v_u^\star, v\rangle'_{\mathcal{Y}} = \langle Bu, v\rangle_{\mathcal{Y}' \times \mathcal{Y}}$ for all $v \in \mathcal{V}$, which is well-defined by the Riesz representation theorem. By (3.2) the bound $\|v_u^\star\|'_{\mathcal{Y}} \leq C\|B\|_{\mathcal{L}(\mathcal{X},\mathcal{Y}')}\|u\|_{\mathcal{X}}$ holds. Moreover, we have

$$\|v_w^\star - v_u^\star\|'_{\mathcal{Y}} = \sup_{v \in \mathcal{V}\setminus\{0\}} \frac{|\langle Bw - Bu, v\rangle_{\mathcal{Y}' \times \mathcal{Y}}|}{\|v\|'_{\mathcal{Y}}} \quad \text{for all} \quad w, u \in \mathcal{X}, \tag{3.5}$$

which, by means of (3.1), implies that the map

$$\mathcal{U} \to \mathcal{V}^\star := \{v_u^\star : u \in \mathcal{U}\} \subseteq \mathcal{V}, \quad w \mapsto v_w^\star \tag{3.6}$$

is injective and thus bijective. Let now $u := B^{-1}f$. Since $\mathcal{V}^\star \subseteq \mathcal{Y}$ is a closed subspace, there exists a unique $v^\star \in \mathcal{V}^\star$ which minimizes $v^\star \mapsto \|v^\star - v_u^\star\|'_{\mathcal{Y}}$. By bijectivity of (3.6) there exists a unique $u' \in \mathcal{U}$ such that $v_{u'}^\star = v^\star$, which is thus the unique solution to (3.3) by the characterization (3.5) with $u = B^{-1}f$. To show (3.4), we observe that for all $w \in \mathcal{U}$ we have

$$\begin{aligned}\gamma_{\mathcal{U},\mathcal{V}}\|w - u'\|_{\mathcal{X}} &\leq \sup_{v \in \mathcal{V}\setminus\{0\}} \frac{\langle B(w - u'), v\rangle_{\mathcal{Y}' \times \mathcal{Y}}}{\|v\|_{\mathcal{Y}}} \leq \|B(w - u)\|_{\mathcal{Y}'} + \frac{1}{c}\|v_u^\star - v_{u'}^\star\|'_{\mathcal{Y}} \\ &\leq \left(1 + \frac{C}{c}\right)\|B\|_{\mathcal{L}(\mathcal{X},\mathcal{Y}')}\|u - w\|_{\mathcal{X}},\end{aligned}$$

using (3.1), (3.5), $\|v_u^\star - v_{u'}^\star\|'_{\mathcal{Y}} \leq \|v_u^\star - v_w^\star\|'_{\mathcal{Y}}$ and (3.2). This implies (3.4). □

We remark that we are interested in using norms $\|\cdot\|'_{\mathcal{Y}}$ on $\mathcal{Y}$ which are numerically easily accessible, for instance via sequence norms generated by a choice of a Riesz basis on $\mathcal{Y}$, see Example 4.2. Theorem 3.1 motivates the following definition.

Definition 3.2. *Let $\mathcal{U} \subseteq \mathcal{X}$ and $\mathcal{V} \subseteq \mathcal{Y}$ be closed subspaces. Let $\|\cdot\|'_{\mathcal{Y}} \sim \|\cdot\|_{\mathcal{Y}}$ be equivalent norms. We call the solution of* (3.3) *the Galerkin least squares solution, provided it exists and is unique.*

For our purposes it is useful to introduce a notation for the *inf-sup condition* for bilinear forms, such as $\langle B\cdot,\cdot\rangle_{\mathcal{Y}'\times\mathcal{Y}}$ in (3.1). This generalizes the notion of the "$\mathcal{K}$-condition" of Babuška & Janik (1990), introduced in the context of parabolic equations to describe stability of the duality paring $\langle\cdot,\cdot\rangle_{V'\times V}$ w.r.t. a finite element space $W\subseteq V$. This is the subject of the following definition.

Definition 3.3. *Let $\widetilde{\mathcal{Z}}$ and $\mathcal{Z}$ be Banach spaces where a pairing $\langle\cdot,\cdot\rangle_{\widetilde{\mathcal{Z}}\times\mathcal{Z}}:\widetilde{\mathcal{Z}}\times\mathcal{Z}\to\mathbb{R}$ is defined. Let $\widetilde{Z}\subseteq\widetilde{\mathcal{Z}}$, $Z\subseteq\mathcal{Z}$ be subspaces. We define $\mathcal{K}_{\widetilde{\mathcal{Z}}\times\mathcal{Z}}(\widetilde{Z},Z)\in\mathbb{R}_{\geq 0}$ w.r.t. the pairing $\langle\cdot,\cdot\rangle_{\widetilde{\mathcal{Z}}\times\mathcal{Z}}$ by*

$$\mathcal{K}_{\widetilde{\mathcal{Z}}\times\mathcal{Z}}(\widetilde{Z},Z):=\inf_{\widetilde{z}\in\widetilde{Z}\setminus\{0\}}\sup_{z\in Z\setminus\{0\}}\frac{\langle\widetilde{z},z\rangle_{\widetilde{\mathcal{Z}}\times\mathcal{Z}}}{\|\widetilde{z}\|_{\widetilde{\mathcal{Z}}}\|z\|_{\mathcal{Z}}}. \tag{3.7}$$

Example 3.4. *We give some examples for the above definition, in particular illustrating typical pairings $\langle\cdot,\cdot\rangle_{\widetilde{\mathcal{Z}}\times\mathcal{Z}}$ to occur in the following.*

1. *For a Hilbert space $\mathcal{Z}$, subspaces $W,Z\subseteq\mathcal{Z}$ we have for any $0\leq\kappa\leq 1$*

$$\mathcal{K}_{\mathcal{Z}\times\mathcal{Z}}(W,Z)\geq\kappa\quad\Leftrightarrow\quad\inf_{z\in Z}\|w-z\|_{\mathcal{Z}}^2\leq(1-\kappa)\|w\|_{\mathcal{Z}}^2\ \forall w\in W \tag{3.8}$$

   with the scalar product $\langle\cdot,\cdot\rangle_{\mathcal{Z}}$ as pairing.

2. *If $\mathcal{Z}\hookrightarrow\widetilde{\mathcal{Z}}$ are Banach spaces then for all subspaces $W\subseteq Z\subseteq\mathcal{Z}$ we have $\mathcal{K}_{\widetilde{\mathcal{Z}}\times\mathcal{Z}}(W,Z)\geq\mathcal{K}_{\widetilde{\mathcal{Z}}\times\mathcal{Z}}(Z,Z)\geq\mathcal{K}_{\widetilde{\mathcal{Z}}\times\mathcal{Z}}(Z,W)$.*

3. *For a Banach space $\mathcal{Z}$, $\mathcal{K}_{\mathcal{Z}'\times\mathcal{Z}}(U,W)\in[0,1]$ for all $U\subseteq\mathcal{Z}'$, $W\subseteq\mathcal{Z}$, where $\langle\cdot,\cdot\rangle_{\mathcal{Z}'\times\mathcal{Z}}$ is the duality pairing.*

4. *The* inf-sup condition (3.1) *can be restated as $\gamma_{\mathcal{U},\mathcal{V}}=\mathcal{K}_{\mathcal{X}\times\mathcal{Y}}(\mathcal{U},\mathcal{V})>0$ w.r.t. the pairing $\langle\cdot,\cdot\rangle_{\mathcal{X}\times\mathcal{Y}}:=\langle B\cdot,\cdot\rangle_{\mathcal{Y}'\times\mathcal{Y}}$.*

The following proposition is a generalization of the fact that $H^1(D)$ stability of the $L^2(D)$-orthogonal projection onto a subspace $W\subseteq H^1(D)$ implies $\mathcal{K}_{(H^1(D))'\times H^1(D)}(W,W)>0$, see also (McLean & Steinbach, 1999, Lemma 3.3) for a comparable statement in the context of boundary element methods. The stability of the $L^2(D)$-orthogonal projector can in turn be verified by checking local mesh criteria (Bramble *et al.*, 2002).

Proposition 3.5. *Let $\widetilde{\mathcal{Z}}$ and $\mathcal{Z}$ be Hilbert spaces where a pairing $\langle\cdot,\cdot\rangle_{\widetilde{\mathcal{Z}}\times\mathcal{Z}}$ is defined. Let $U\subseteq\widetilde{\mathcal{Z}}$, $W\subseteq\mathcal{Z}$ be subspaces. Assume that $T:\widetilde{\mathcal{Z}}\to W$ satisfies*

$$\langle\widetilde{z},w\rangle_{\widetilde{\mathcal{Z}}\times\mathcal{Z}}=\langle T\widetilde{z},w\rangle_{\mathcal{Z}}\quad\text{for all}\quad\widetilde{z}\in\widetilde{\mathcal{Z}},w\in W. \tag{3.9}$$

For any $C>0$, the following are equivalent:

i) $\|u\|_{\widetilde{\mathcal{Z}}}\leq C\|Tu\|_{\mathcal{Z}}$ *for all* $u\in U$.

ii) $\mathcal{K}_{\widetilde{\mathcal{Z}}\times\mathcal{Z}}(U,W)\geq C^{-1}$.

ii') $\mathcal{K}_{\widetilde{\mathcal{Z}}\times\mathcal{Z}}(U,Z)\geq C^{-1}\mathcal{K}_{\mathcal{Z}\times\mathcal{Z}}(W,Z)$ *for all subspaces* $Z\subseteq\mathcal{Z}$.

In particular, for all subspaces $Z \subseteq \mathcal{Z}$,

$$\mathcal{K}_{\widetilde{\mathcal{Z}}\times\mathcal{Z}}(U,Z) \geq \mathcal{K}_{\widetilde{\mathcal{Z}}\times\mathcal{Z}}(U,W)\mathcal{K}_{\mathcal{Z}\times\mathcal{Z}}(W,Z). \tag{3.10}$$

Now assume additionally that $\mathcal{Z} \hookrightarrow \mathcal{H} \cong \mathcal{H}' \hookrightarrow \mathcal{Z}' = \widetilde{\mathcal{Z}}$ forms a Gelfand triple of separable Hilbert spaces. Further, let $Q : \mathcal{Z} \to W$ satisfy

$$\langle u, Qz\rangle_{\mathcal{Z}'\times\mathcal{Z}} = \langle u, z\rangle_{\mathcal{Z}'\times\mathcal{Z}} \quad \text{for all} \quad u \in U, z \in \mathcal{Z}. \tag{3.11}$$

If $U \cong W \subseteq \mathcal{Z}$ is closed, then Q is the $\mathcal{H}$-orthogonal projection onto W. Moreover, i) *is equivalent to stability of Q in $\mathcal{Z}$, i.e.,*

iii) $\|Qz\|_{\mathcal{Z}} \leq C\|z\|_{\mathcal{Z}}$ for all $z \in \mathcal{Z}$.

Remark 3.6. *Let $\widetilde{\mathcal{Z}}$, $\mathcal{Z}$ be Hilbert spaces and $\langle\cdot,\cdot\rangle_{\widetilde{\mathcal{Z}}\times\mathcal{Z}} : \widetilde{\mathcal{Z}}\times\mathcal{Z} \to \mathbb{R}$ a continuous bilinear form. Let $W \subseteq \mathcal{Z}$ be a subspace. It is easy to see that there is at most one $T : \widetilde{\mathcal{Z}} \to W$ satisfying* (3.9)*. Assume that W is separable, and set $N := \dim W$. Let $\{w_n\}_{n\in[N]} \subset W$ be a $\mathcal{Z}$-orthonormal basis for W. Then $\sum_{n\in[N]} w_n\langle\cdot, w_n\rangle_{\widetilde{\mathcal{Z}}\times\mathcal{Z}} \in \mathcal{L}(\widetilde{\mathcal{Z}}, W)$ can be easily verified to satisfy* (3.9)*. Thus, if $\mathcal{Z}$ is separable,* (3.9) *uniquely defines T and $T \in \mathcal{L}(\widetilde{\mathcal{Z}}, W)$.*

Proof. (Of Proposition 3.5) To see *i) $\Leftrightarrow$ ii)*, let $u \in U \setminus \{0\}$ be arbitrary. Given *i)*, the fact that $\langle u, Tu\rangle_{\widetilde{\mathcal{Z}}\times\mathcal{Z}} = \|Tu\|_{\mathcal{Z}}^2 \geq C^{-1}\|Tu\|_{\mathcal{Z}}\|u\|_{\widetilde{\mathcal{Z}}}$ and $Tu \neq 0$ show *ii)*. Conversely, assume *ii)*. Then *i)* is due to

$$C^{-1}\|u\|_{\widetilde{\mathcal{Z}}} \leq \sup_{w\in W\setminus\{0\}} \frac{\langle u, w\rangle_{\widetilde{\mathcal{Z}}\times\mathcal{Z}}}{\|w\|_{\mathcal{Z}}} = \sup_{w\in W\setminus\{0\}} \frac{\langle Tu, w\rangle_{\mathcal{Z}}}{\|w\|_{\mathcal{Z}}} \leq \|Tu\|_{\mathcal{Z}}.$$

With the choice $Z := W$, the implication *ii') $\Rightarrow$ ii)* is obvious from (3.8). We now check *ii) $\Rightarrow$ ii')*. Indeed, observing the proven implication *ii) $\Rightarrow$ i)*, we have for any subspace $Z \subseteq \mathcal{Z}$

$$\mathcal{K}_{\widetilde{\mathcal{Z}}\times\mathcal{Z}}(U,Z) = \inf_{u\in U\setminus\{0\}} \frac{\|Tu\|_{\mathcal{Z}}}{\|u\|_{\widetilde{\mathcal{Z}}}} \sup_{z\in Z\setminus\{0\}} \frac{\langle Tu, z\rangle_{\mathcal{Z}}}{\|Tu\|_{\mathcal{Z}}\|z\|_{\mathcal{Z}}} \geq C^{-1}\mathcal{K}_{\mathcal{Z}\times\mathcal{Z}}(W,Z).$$

To complete the first part of the proof in the nontrivial case where the r.h.s. of (3.10) is nonzero, it suffices to use *ii)* and *ii')* with the admissible choice $C := \mathcal{K}_{\widetilde{\mathcal{Z}}\times\mathcal{Z}}(U,W)^{-1} > 0$. This shows *i) $\Leftrightarrow$ ii) $\Leftrightarrow$ ii')* and (3.10).
We now show *i) $\Rightarrow$ iii)*. Clearly, (3.11) characterizes Q as the $\mathcal{H}$-orthogonal projection by definition of the Gelfand triple. Further, T is easily checked to be surjective. Thus, for all $z \in \mathcal{Z}$ there exists $u \in U$ with $Tu = Qz$, for which $\langle Tu, Qz\rangle_{\mathcal{Z}} = \langle u, z\rangle_{\mathcal{Z}'\times\mathcal{Z}}$ with *i)* yields *iii)*. Finally, *iii) $\Rightarrow$ i)* is clear using (3.9) and (3.11). □

Corollary 3.7. *Let $\widetilde{\mathcal{Z}}$, $\mathcal{Z}$ be separable Hilbert spaces. Let $U_i \subseteq \widetilde{\mathcal{Z}}$, $i \in \mathbb{N}$, and $W \subseteq \mathcal{Z}$ be closed subspaces. Let $T \in \mathcal{L}(\widetilde{\mathcal{Z}}, W)$ be given by* (3.9)*. Set $\kappa_i := \mathcal{K}_{\widetilde{\mathcal{Z}}\times\mathcal{Z}}(U_i, W)$, $i \in \mathbb{N}$. Assume further*

$$\langle Tu_i, Tu_j\rangle_{\mathcal{Z}} = 0 = \langle u_i, u_j\rangle_{\widetilde{\mathcal{Z}}} \quad \forall u_i \in U_i, u_j \in U_j, i \neq j. \tag{3.12}$$

Then $\mathcal{K}_{\widetilde{\mathcal{Z}}\times\mathcal{Z}}(\sum_{i\in\mathbb{N}} U_i, W) \geq \inf_{i\in\mathbb{N}} \kappa_i$.

Proof. We assume $\kappa := \inf_{i\in\mathbb{N}} \kappa_i > 0$, as the statement is trivial otherwise. Since any $u \in \sum_{i\in\mathbb{N}} U_i$ has the form $u = \sum_{i\in\mathbb{N}} u_i$ with (unique) $u_i \in U_i \setminus \{0\}$, $i \in \mathbb{N}$, we have using (3.12) and Proposition 3.5 "*ii) ⇒ i)*"

$$\|Tu\|_{\mathcal{Z}}^2 = \sum_{i\in\mathbb{N}} \|Tu_i\|_{\mathcal{Z}}^2 \geq \sum_{i\in\mathbb{N}} \kappa_i^2 \|u_i\|_{\widetilde{\mathcal{Z}}}^2 \geq \kappa \sum_{i\in\mathbb{N}} \|u_i\|_{\widetilde{\mathcal{Z}}}^2 = \kappa^2 \|u\|_{\widetilde{\mathcal{Z}}}^2.$$

Using Proposition 3.5 "*i) ⇒ ii)*" shows the claim. □

Corollary 3.8. *Let $\widetilde{\mathcal{Z}}_i$, $\mathcal{Z}_i$, $i = 1, 2$, be separable Hilbert spaces. Let $U_i \subseteq \mathcal{Z}_i'$ and $W_i \subseteq \widetilde{\mathcal{Z}}_i$, $i = 1, 2$, be subspaces. Set $\kappa_i := \mathcal{K}_{\widetilde{\mathcal{Z}}_i \times \mathcal{Z}_i}(U_i, W_i)$, $i = 1, 2$. Then $\mathcal{K}_{(\widetilde{\mathcal{Z}}_1 \otimes \widetilde{\mathcal{Z}}_2)\times(\mathcal{Z}_1\otimes\mathcal{Z}_2)}(U_1 \otimes U_2, W_1 \otimes W_2) \geq \kappa_1\kappa_2$ and $\mathcal{K}_{(\widetilde{\mathcal{Z}}_1 \times \widetilde{\mathcal{Z}}_2)\times(\mathcal{Z}_1\times\mathcal{Z}_2)}(U_1 \times U_2, W_1 \times W_2) \geq \kappa_1 + \kappa_2$.*

Proof. The proof follows similarly to Corollary 3.7 from Proposition 3.5. □

We now formulate and prove a result which will allow the construction of stable finite element spaces for solving (2.11). We note that for the presented proof it is essential that $a_t(\cdot,\cdot)$ is symmetric and elliptic in the sense of Assumption 2.1. We write $\partial_t \mathcal{U} := \{\partial_t u \in L^2(J; V') : u \in \mathcal{U}\}$ for any subspace $\mathcal{U} \subseteq \mathcal{X}$.

Theorem 3.9. *Let Assumption 2.1 hold. There exists a constant $c > 0$, only dependent on $a_t(\cdot,\cdot)$, s.t. the inf-sup condition* (3.1) *holds with $\gamma_{\mathcal{U},\mathcal{V}} \geq c\kappa_1 > 0$ for all closed subspaces $\mathcal{U} \subseteq \mathcal{X}$, $\mathcal{V} = \mathcal{V}_1 \times \mathcal{V}_2 \subseteq \mathcal{Y}$ satisfying*

1. $\kappa_1 := \mathcal{K}_{\mathcal{Y}_1'\times\mathcal{Y}_1}(\partial_t\mathcal{U}, \mathcal{V}_1) > 0$,
2. $\mathcal{U} \subseteq \mathcal{V}_1$,
3. $\mathcal{U}|_{t=0} := \{u|_{t=0} : u \in \mathcal{U}\} \subseteq \mathcal{V}_2$.

Proof. On the set $\mathcal{Z} := \mathcal{Z}_1 \times \mathcal{Z}_2 := \mathcal{Y}_1 \times \mathcal{Y}_2$ we consider $\langle\cdot,\cdot\rangle_{\mathcal{Z}}$ defined by

$$\langle z, \widetilde{z}\rangle_{\mathcal{Z}} := \int_J a_t(z_1(t), \widetilde{z}_1(t))dt + \langle z_2, \widetilde{z}_2\rangle_H \quad \forall z, \widetilde{z} \in \mathcal{Z}. \tag{3.13}$$

By Assumption 2.1, $\langle\cdot,\cdot\rangle_{\mathcal{Z}}$ is a scalar product, $(\mathcal{Z}, \langle\cdot,\cdot\rangle_{\mathcal{Z}})$ is a Hilbert space, and the induced norm $\|\cdot\|_{\mathcal{Z}}$ is equivalent to $\|\cdot\|_{\mathcal{Y}}$, and similarly for the duals $\mathcal{Z}'$ and $\mathcal{Y}'$. In particular, $\widetilde{\kappa}_1 := \mathcal{K}_{\mathcal{Z}_1'\times\mathcal{Z}_1}(\partial_t\mathcal{U}, \mathcal{V}_1) \geq \widetilde{c}\kappa_1$with the pairing $\langle\partial_t x, z_1\rangle_{\mathcal{Z}_1'\times\mathcal{Z}_1} = \int_J \langle\partial_t x(t), z_1(t)\rangle_{V'\times V} dt$, $x \in \mathcal{X}$, $z_1 \in \mathcal{Z}_1$, for a constant $\widetilde{c} > 0$ only depending on $a_t(\cdot,\cdot)$. Let $T : \mathcal{X} \to \mathcal{V} \subseteq \mathcal{Z}$ be given by

$$\langle Tx, v\rangle_{\mathcal{Z}} = \langle x, v\rangle_{\mathcal{X}\times\mathcal{Z}} := (Bx)(v) \quad \forall x \in \mathcal{X}, v \in \mathcal{V}. \tag{3.14}$$

Note that $\langle\cdot,\cdot\rangle_{\mathcal{X}\times\mathcal{Z}} : \mathcal{X} \times \mathcal{Z} \to \mathbb{R}$ is bilinear and continuous, and thus T is well-defined, see Remark 3.6. Define $I : \mathcal{X} \to \mathcal{Z}$ by $Ix := (x, x(0)) \in \mathcal{Z}$, $x \in \mathcal{X}$, which is well-defined, since $x(0) \in H$ (see (2.8)) and $\mathcal{X} \subseteq \mathcal{Y}_1$.
Let $u \in \mathcal{U}$ be arbitrary. Since for a.e. $t \in J$ there holds $\langle\partial_t u(t), u(t)\rangle_{V'\times V} = \frac{1}{2}\partial_t\|u(t)\|_H^2$, from (3.14) we obtain

$$2\langle Tu, Iu\rangle_{\mathcal{Z}} = \|u\|_{\mathcal{Z}_1}^2 + \|Iu\|_{\mathcal{Z}}^2 + \|u(T_{\text{end}})\|_H^2. \tag{3.15}$$

Due to $\mathcal{U} \subseteq \mathcal{V}_1$ and $\mathcal{U}|_{t=0} \subseteq \mathcal{V}_2$ we have $I\mathcal{U} \subseteq \mathcal{V}$. Thus, (3.14) yields

$$\|Tu - Iu\|_{\mathcal{Z}} = \sup_{v\in\mathcal{V}\setminus\{0\}} \frac{\langle Tu - Iu, v\rangle_{\mathcal{Z}}}{\|v\|_{\mathcal{Z}}} = \sup_{v\in\mathcal{V}\setminus\{0\}} \frac{\langle\partial_t u, v_1\rangle_{\mathcal{Z}_1'\times\mathcal{Z}_1}}{\|v\|_{\mathcal{Z}}} \geq \widetilde{\kappa}_1\|\partial_t u\|_{\mathcal{Z}_1'},$$

from which we obtain with (3.15)

$$\|Tu\|_{\mathcal{Z}}^2 \geq \|Tu - Iu\|_{\mathcal{Z}}^2 + \|u\|_{\mathcal{Z}_1}^2 \geq \widetilde{c}^2\kappa_1^2\|\partial_t u\|_{\mathcal{Z}_1'}^2 + \|u\|_{\mathcal{Z}_1}^2 \geq c\kappa_1\|u\|_{\mathcal{X}}^2,$$

where $c > 0$ depends only on $a_t(\cdot,\cdot)$ due to norm equivalences $\|\cdot\|_{\mathcal{Z}_1'} \sim \|\cdot\|_{\mathcal{Y}_1'}$ and $\|\cdot\|_{\mathcal{Z}_1} \sim \|\cdot\|_{\mathcal{Y}_1}$ and by definition of $\widetilde{c} > 0$. As $u \in \mathcal{U}$ was arbitrary, Proposition 3.5 "*i) ⇒ ii)*" shows $\mathcal{K}_{\mathcal{X}\times\mathcal{Y}}(\mathcal{U},\mathcal{V}) \geq c\kappa_1$, which is (see Example 3.4) the claim. □

Example 3.10. *Theorem 3.9 applies to the pair $\mathcal{U} := \mathcal{X}$, $\mathcal{V} := \mathcal{Y}$, in which case we have $\kappa_1 = 1$.*

We will apply Theorem 3.9 in a setting described more closely by the following proposition. For a subspace $E \subseteq H^1(J)$ we write $E' := \{e' \in L^2(J) : e \in E\}$, where $e' \in L^2(J)$ denotes the distributional derivative of $e \in H^1(J)$.

Proposition 3.11. *Let $E_k \subseteq H^1(J)$, $F_k \subseteq L^2(J)$, $k \in \mathbb{N}_0$, and $U_\ell \subseteq V'$, $V_\ell \subseteq V$, $\ell \in \mathbb{N}_0$, be closed subspaces. Assume that $(U_\ell)_{\ell\in\mathbb{N}_0}$ and $(V_\ell)_{\ell\in\mathbb{N}_0}$ are nested, that is $U_\ell \subseteq U_{\ell+1}$ and $V_\ell \subseteq V_{\ell+1}$, $\ell \in \mathbb{N}_0$. Set*

$$\tau := \inf_{k\in\mathbb{N}_0} \mathcal{K}_{L^2(J)\times L^2(J)}(E_k', F_k) \quad \text{and} \quad \eta := \inf_{\ell\in\mathbb{N}_0} \mathcal{K}_{V'\times V}(U_\ell, V_\ell).$$

Let $L \in \mathbb{N}_0$. Define the subspaces $\mathcal{U} \subseteq \mathcal{X}$ and $\mathcal{V}_1 \subseteq \mathcal{Y}_1$ by

$$\mathcal{U} := \sum_{0\leq k+\ell\leq L} E_k \otimes U_\ell \quad \text{and} \quad \mathcal{V}_1 := \sum_{0\leq k+\ell\leq L} F_k \otimes V_\ell, \tag{3.16}$$

where k, ℓ range over $\mathbb{N}_0$. Then $\mathcal{K}_{\mathcal{Y}_1'\times\mathcal{Y}_1}(\partial_t\mathcal{U}, \mathcal{V}_1) \geq \eta\tau$.

Proof. Consider the auxiliary space $\widetilde{\mathcal{V}}_1 := \sum_{0\leq k+\ell\leq L} E_k' \otimes V_\ell \subseteq \mathcal{Y}_1$. The claim follows using Proposition 3.5, (3.10), once $\widetilde{\eta} := \mathcal{K}_{\mathcal{Y}_1'\times\mathcal{Y}_1}(\partial_t\mathcal{U}, \widetilde{\mathcal{V}}_1) \geq \eta$ and $\widetilde{\tau} := \mathcal{K}_{\mathcal{Y}_1\times\mathcal{Y}_1}(\widetilde{\mathcal{V}}_1, \mathcal{V}_1) \geq \tau$ are established. To see $\widetilde{\eta} \geq \eta$, consider $T : \mathcal{Y}_1' \to \widetilde{\mathcal{V}}_1$, given by

$$\langle Ty_1', \widetilde{v}_1\rangle_{\mathcal{Y}_1\times\mathcal{Y}_1} = \langle y_1', \widetilde{v}_1\rangle_{\mathcal{Y}_1'\times\mathcal{Y}_1} \quad \forall y_1' \in \partial_t\mathcal{U}, \widetilde{v}_1 \in \widetilde{\mathcal{V}}_1.$$

Since $U_\ell \subseteq U_{\ell+1}$ and $V_\ell \subseteq V_{\ell+1}$, $\ell \in \mathbb{N}_0$, defining

$$G_0 := E_0' \quad \text{and} \quad G_k := E_k' \cap \left(\bigcup_{k'=0}^{k-1} E_{k'}'\right)^{\perp_{L^2(J)}} \quad \forall k \in \mathbb{N}$$

we have $\partial_t\mathcal{U} = \sum_{k=0}^L G_k \otimes U_{L-k}$ and $\widetilde{\mathcal{V}}_1 = \sum_{k=0}^L G_k \otimes V_{L-k}$. Note that $G_k \perp_{L^2(J)} G_{k'}$ for all nonnegative integers $k < k' \leq L$. In particular, for all $k = 0, 1, \ldots, L$ we have $T(G_k \otimes U_{L-k}) \subseteq G_k \otimes V_{L-k}$. Using Lemma 3.8 and Lemma 3.7 we obtain $\widetilde{\eta} \geq \eta$. Lastly, $\widetilde{\tau} \geq \tau$ follows easily from (3.8). □

4 Discrete weighted least squares formulation

In order to describe the discretization and the solution process for (2.11), we assume that we are given operators $\mathcal{N} : \mathcal{Y} \to \mathcal{Y}'$ and $\mathcal{M} : \mathcal{X} \to \mathcal{X}'$ such that $\langle\cdot,\cdot\rangle_{\mathcal{N}} := \langle\mathcal{N}\cdot,\cdot\rangle_{\mathcal{Y}'\times\mathcal{Y}}$ is a scalar product on $\mathcal{Y}$ and $\langle\cdot,\cdot\rangle_{\mathcal{M}} := \langle\mathcal{M}\cdot,\cdot\rangle_{\mathcal{X}'\times\mathcal{X}}$ is

a scalar product on $\mathcal{X}$. We assume that the induced norms $\|\cdot\|_{\mathcal{N}}$ and $\|\cdot\|_{\mathcal{M}}$ provide equivalent norms to $\|\cdot\|_{\mathcal{Y}}$ and $\|\cdot\|_{\mathcal{X}}$, i.e.,

$$\exists d_{\mathcal{N}}, D_{\mathcal{N}} > 0: \quad d_{\mathcal{N}}\|v\|_{\mathcal{Y}} \leq \|v\|_{\mathcal{N}} \leq D_{\mathcal{N}}\|v\|_{\mathcal{Y}} \quad \forall v \in \mathcal{Y} \tag{4.1}$$

and

$$\exists d_{\mathcal{M}}, D_{\mathcal{M}} > 0: \quad d_{\mathcal{M}}\|u\|_{\mathcal{X}} \leq \|u\|_{\mathcal{M}} \leq D_{\mathcal{M}}\|u\|_{\mathcal{X}} \quad \forall u \in \mathcal{X}. \tag{4.2}$$

Throughout this section we will assume that $\mathcal{U} \subseteq \mathcal{X}$ and $\mathcal{V} \subseteq \mathcal{Y}$ are closed subspaces s.t. the *inf-sup condition* (3.1) holds. Let $M := \dim \mathcal{U} \in \mathbb{N} \cup \{\infty\}$ and $N := \dim \mathcal{V} \in \mathbb{N} \cup \{\infty\}$ be the dimensions of $\mathcal{U}$ and $\mathcal{V}$. Let $\{\phi_m\}_{m\in[M]} \subset \mathcal{U}$ and $\{\psi_n\}_{n\in[N]} \subset \mathcal{V}$ be bases. We define (possibly bi- or semi-infinite) matrices $\mathbf{N} \in \mathbb{R}^{N\times N}$, $\mathbf{B} \in \mathbb{R}^{N\times M}$ and $\mathbf{M} \in \mathbb{R}^{M\times M}$ by

$$\mathbf{N}_{n',n} := (\mathcal{N}\psi_n)(\psi_{n'}),\ \mathbf{B}_{n,m} := (B\phi_m)(\psi_n),\ \mathbf{M}_{m,m'} := (\mathcal{M}\phi_{m'})(\phi_m), \tag{4.3}$$

$n, n' \in [N]$, $m, m' \in [M]$, and $\mathbf{f} \in \mathbb{R}^N$ as the (possibly infinite) column vector with components $\mathbf{f}_n := f(\psi_n)$, $n \in [N]$. We call $\mathbf{B}$ the *system matrix* and $\mathbf{f}$ the *load vector* w.r.t. the chosen bases. The norm equivalences (4.1)–(4.2) can be restated by saying that the synthesis operators $\mathcal{J} : \ell^2_{\mathbf{N}}([N]) \to \mathcal{V}$, defined by $\mathcal{J}\mathbf{v} := \sum_{n\in[N]} \mathbf{v}_n\psi_n$ for all $\mathbf{v} \in \ell^2_{\mathbf{N}}([N])$, and $\mathcal{I} : \ell^2_{\mathbf{M}}([M]) \to \mathcal{U}$, defined by $\mathcal{I}\mathbf{w} := \sum_{m\in[M]} \mathbf{w}_m\phi_m$ for all $\mathbf{w} \in \ell^2_{\mathbf{M}}([M])$ are (quasi-isometric) isomorphisms.

Example 4.1. *Define $\mathcal{M} : \mathcal{X} \to \mathcal{X}'$ as the Riesz operator, characterized by $\langle \mathcal{M}u, u'\rangle_{\mathcal{X}'\times\mathcal{X}} = \langle u, u'\rangle_{\mathcal{X}}$ for all $u, u' \in \mathcal{X}$. Then $\|\cdot\|_{\mathcal{M}} \equiv \|\cdot\|_{\mathcal{X}}$.*

Example 4.2. *Assume $\{\psi_n\}_{n\in[N]}$ can be rescaled to a Riesz basis (Christensen, 2003) for $\mathcal{Y}$, i.e., there exist constants $0 \leq \lambda \leq \Lambda$ s.t.*

$$\lambda\|\mathbf{v}\|^2_{\ell^2([N])} \leq \left\| \sum_{n\in[N]} \mathbf{v}_n \frac{\psi_n}{\|\psi_n\|_{\mathcal{Y}}} \right\|^2_{\mathcal{Y}} \leq \Lambda\|\mathbf{v}\|^2_{\ell^2([N])} \quad \forall \mathbf{v} \in \ell^2([N]). \tag{4.4}$$

Define $\mathcal{N} : \mathcal{Y} \to \mathcal{Y}'$ by $\langle \mathcal{N}\psi_n, \psi_{n'}\rangle_{\mathcal{Y}'\times\mathcal{Y}} := \delta_{nn'}c_n\|\psi_n\|^2_{\mathcal{Y}}$, $n, n' \in [N]$, where $c_n > 0$ are constants with $0 < \underline{c} := \inf_{n\in[N]} c_n \leq \sup_{n\in[N]} c_n =: \overline{c} < \infty$. Then $\mathbf{N} \in \mathbb{R}^{N\times N}$ is a nonnegative diagonal matrix. Note that (4.4) *implies*

$$\lambda/\overline{c}\|\mathbf{v}\|^2_{\mathbf{N}} \leq \|\mathcal{J}\mathbf{v}\|^2_{\mathcal{Y}} \leq \Lambda/\underline{c}\|\mathbf{v}\|^2_{\mathbf{N}} \quad \forall \mathbf{v} \in \ell^2_{\mathbf{N}}([N]),$$

and thus the norm equivalence (4.1) *holds with $d_{\mathcal{N}} = \sqrt{\underline{c}/\Lambda}$, $D_{\mathcal{N}} = \sqrt{\overline{c}/\lambda}$. This renders Riesz bases particularly well-suited for what follows.*

Example 4.3. *This example is motivated by the well-known BPX (Bramble* et al.*, 1990) operator, cf. (Xu, 1992, Section 5). Let $\{0\} = V_{-1} \subseteq V_\ell \subseteq V_{\ell+1} \subseteq V$ be closed, nested subspaces, s.t. $\bigcup_{\ell\in\mathbb{N}_0} V_\ell$ is dense in V. Let $Q_\ell : V \to V_\ell \cap (V_{\ell-1})^{\perp_H}$, $\ell \in \mathbb{N}_0$ be the H-orthogonal projector. Assume, $0 < d_V \leq D_V$ are constants s.t.*

$$\forall v \in V: \quad d_V\|v\|^2_V \leq \sum_{\ell\in\mathbb{N}_0} 2^{2\ell}\|Q_\ell v\|^2_H \leq D_V\|v\|^2_V.$$

Then, as in (Oswald, 1998, Lemma 1) it follows that also

$$\forall v \in V: \quad D_V^{-1}\|v\|_{V'}^2 \le \sum_{\ell\in\mathbb{N}_0} 2^{-2\ell}\|Q_\ell v\|_H^2 \le d_V^{-1}\|v\|_{V'}^2.$$

Further, let $E_k \subseteq H^1(J)$, $k \in \mathbb{N}_0$, be closed subspaces s.t. $\bigcup_{k\in\mathbb{N}_0} E_k$ is dense in $H^1(J)$ and $P_k : H^1(J) \to E_k$, $k \in \mathbb{N}_0$, be linear projectors satisfying $P_k P_{k'} = 0$ for all nonnegative integers $k \neq k'$. Assume that with constants $0 < d_E \le D_E$, where $E \in \{L^2(J), H^1(J)\}$, for all $e \in H^1(J)$ there holds

$$d_E\|e\|_E^2 \le \sum_{k\in\mathbb{N}_0} 2^{2kt}\|P_k e\|_{L^2(J)}^2 \le D_E\|e\|_E^2 \quad \text{for} \quad \begin{cases} t=0, E=L^2(J) \\ t=1, E=H^1(J). \end{cases}$$

Consider the operators $\mathcal{M}^+$ and $\mathcal{M}^-$, defined on $H^1(J)\otimes V$ by

$$\mathcal{M}^{\pm} = \sum_{k\in\mathbb{N}_0}\sum_{\ell\in\mathbb{N}_0} \left(2^{2\ell} + 2^{2k}2^{-2\ell}\right)^{\pm 1} P_k \otimes Q_\ell.$$

It can be easily checked that

1. *$\mathcal{M}^+$ extends continuously to an operator $\mathcal{M} \in \mathcal{L}(\mathcal{X}, \mathcal{X}')$ s.t. with $d_{\mathcal{M}} \sim \min\{d_{L^2(J)}d_V, d_{H^1(J)}D_V^{-1}\}$ and $D_{\mathcal{M}} \sim \max\{D_{L^2(J)}D_V, D_{H^1(J)}d_V^{-1}\}$ the norm equivalence (4.2) holds up to constants of the equivalence $\|\cdot\|_{\mathcal{X}}^2 \sim \|\cdot\|_{L^2(J;V)}^2 + \|\cdot\|_{H^1(J;V')}^2$.*
2. *$\mathcal{M}^-$ extends continuously to the inverse of $\mathcal{M}$.*
3. *As in (4.3), let $\mathbf{M}_\pm \in \mathbb{R}^{M\times M}$ and $\mathbf{M}_0 \in \mathbb{R}^{M\times M}$ be the matrices with components given by $(\mathcal{M}^\pm\phi_{m'})(\phi_m)$ and $\langle\phi_{m'},\phi_m\rangle_{L^2(J;H)}$, respectively. Then $\mathbf{M}_+^{-1} = \mathbf{M}_0^{-1}\mathbf{M}_-\mathbf{M}_0^{-1}$.*

Therefore, the inverse of the matrix $\mathbf{M} = \mathbf{M}_+$ of the operator $\mathcal{M}$ can be efficiently applied to a vector, as required by Algorithm 4.5. Similarly, an operator $\mathcal{N} \in \mathcal{L}(\mathcal{Y}, \mathcal{Y}')$ satisfying (4.1) can be constructed.

Proposition 4.4. *Let $\mathcal{U} \subseteq \mathcal{X}$, $\mathcal{V} \subseteq \mathcal{Y}$ be closed subspaces satisfying the* inf-sup condition (3.1). *Assume norm equivalences (4.1)–(4.2). Let $\mathbf{N}$, $\mathbf{B}$ and $\mathbf{M}$ be defined by (4.3). There holds*

$$\frac{\gamma_{\mathcal{U},\mathcal{V}}}{D_{\mathcal{N}}D_{\mathcal{M}}} =: \widetilde{\gamma} \le \frac{\|\mathbf{B}\mathbf{u}\|_{\mathbf{N}^{-1}}}{\|\mathbf{u}\|_{\mathbf{M}}} \le \widetilde{\Gamma} := \frac{\|B\|_{\mathcal{L}(\mathcal{X},\mathcal{Y}')}}{d_{\mathcal{N}}d_{\mathcal{M}}} \quad \forall \mathbf{u} \in \ell^2_{\mathbf{M}}([M]) \setminus \{0\}. \tag{4.5}$$

Moreover, with $\mathbf{M}^{\top/2}\mathbf{M}^{1/2} := \mathbf{M}$ and $\mathbf{N}^{\top/2}\mathbf{N}^{1/2} := \mathbf{N}$, we have

$$\varkappa_2 := \varkappa_2(\mathbf{M}^{-\top/2}\mathbf{B}^\top\mathbf{N}^{-1}\mathbf{B}\mathbf{M}^{-1/2}) \le \widetilde{\Gamma}/\widetilde{\gamma}. \tag{4.6}$$

Further, there is a unique minimizer $\mathbf{u}' \in \ell^2_{\mathbf{M}}([M])$ of

$$\text{find} \quad \mathbf{w} \in \ell^2_{\mathbf{M}}([M]) \quad \text{s.t.} \quad \|\mathbf{B}\mathbf{w} - \mathbf{f}\|_{\mathbf{N}^{-1}} \overset{!}{\to} \min \tag{4.7}$$

and $u' := \mathcal{I}\mathbf{u}' \in \mathcal{U}$ satisfies (3.3) with $\|\cdot\|_{\mathcal{Y}}' := \|\cdot\|_{\mathcal{N}}$.

Proof. Note that (4.5) implies that for any singular value σ of the matrix $\mathbf{N}^{-\top/2}\mathbf{B}\mathbf{M}^{-1/2}$, we have $0 < \widetilde{\gamma} \le \sigma \le \widetilde{\Gamma} < \infty$, from which (4.6) follows at once. To see (4.5), let $\mathbf{u} \in \ell^2_{\mathbf{M}}([M]) \setminus \{0\}$, and $u := \mathcal{I}\mathbf{u} \in \mathcal{U}$. Then

$$\frac{\|\mathbf{B}\mathbf{u}\|_{\mathbf{N}^{-1}}}{\|\mathbf{u}\|_{\mathbf{M}}} = \sup_{\mathbf{v}\in\ell^2_{\mathbf{N}}([N])\setminus\{0\}} \frac{\mathbf{v}^\top\mathbf{B}\mathbf{u}}{\|\mathbf{u}\|_{\mathbf{M}}\|\mathbf{v}\|_{\mathbf{N}}} = \sup_{v\in\mathcal{V}\setminus\{0\}} \frac{(Bu)(v)}{\|u\|_{\mathcal{M}}\|v\|_{\mathcal{N}}},$$

which, using (4.1)–(4.2) and (3.1) yields (4.5). Finally, existence and uniqueness of a minimizer of (4.7) can be easily checked using (4.6), and equivalence to (3.3) is seen from the identity

$$\|\mathbf{B}\mathbf{w}-\mathbf{f}\|_{\mathbf{N}^{-1}} = \sup_{\mathbf{v}\in\ell^2_{\mathbf{N}}([N])\setminus\{0\}} \frac{\mathbf{v}^\top(\mathbf{B}\mathbf{w}-\mathbf{f})}{\|\mathbf{v}\|_{\mathbf{N}}} = \sup_{v\in\mathcal{V}\setminus\{0\}} \frac{|(B\mathcal{I}\mathbf{w})(v)-f(v)|}{\|v\|_{\mathcal{N}}}$$

for all $\mathbf{w} \in \mathbb{R}^M$ with $\|\mathbf{w}\|_{\mathbf{M}} < \infty$ by definition of $\mathbf{B}$, $\mathbf{f}$ and $\mathbf{N}$. □

Note that the solution $\mathbf{u}' \in \mathbb{R}^M$ to (4.7) is the unique minimizer of

$$\|\mathbf{N}^{-\top/2}\mathbf{B}\mathbf{M}^{-1/2}\widetilde{\mathbf{u}}' - \mathbf{N}^{-\top/2}\mathbf{f}\|_2 \stackrel{!}{\to} \min \quad \text{with} \quad \widetilde{\mathbf{u}}' := \mathbf{M}^{1/2}\mathbf{u}', \tag{4.8}$$

and, equivalently, solves the normal equations

$$\mathbf{M}^{-\top/2}\mathbf{B}^\top\mathbf{N}^{-1}\mathbf{B}\mathbf{M}^{-1/2}\widetilde{\mathbf{u}}' = \mathbf{M}^{-\top/2}\mathbf{B}^\top\mathbf{N}^{-1}\mathbf{f} \quad \text{with} \quad \widetilde{\mathbf{u}}' := \mathbf{M}^{1/2}\mathbf{u}', \tag{4.9}$$

where the preconditioned matrix is symmetric positive definite with condition number $\varkappa_2$ bounded by means of (4.6). While the conjugate gradient (CG) algorithm can be applied to (4.9), we adapt the analytically equivalent least squares algorithm based on bidiagonalization of Paige & Saunders (1982). The method can be reformulated to require the computation of the action of $\mathbf{M}^{-1}$ and $\mathbf{N}^{-1}$ only (Benbow, 1999, Section 4.3), i.e., no Cholesky (or similar) decomposition of $\mathbf{M}$ or $\mathbf{N}$ needs to be computed. Since this formulation is not very well known we give the complete algorithm below.

Algorithm 4.5 (Generalized least squares)**.** *For $\mathbf{B} \in \mathbb{R}^{N\times M}$, $N, M \in \mathbb{N}_0$, $N \ge M$, of full rank, $\mathbf{M} \in \mathbb{R}^{M\times M}$ and $\mathbf{N} \in \mathbb{R}^{N\times N}$ s.p.d., $\mathbf{f} \in \mathbb{R}^N$, returns an approximate solution $\mathbf{w}_{i^\star} \in \mathbb{R}^M$ to* (4.7)

1. *(a)* $(\mathbf{v}_1, \widetilde{\mathbf{v}}_1, \beta_1) :=$ NORMALIZE$(\mathbf{f}, \mathbf{N})$
   (b) $(\mathbf{u}_1, \widetilde{\mathbf{u}}_1, \alpha_1) :=$ NORMALIZE$(\mathbf{B}^\top\mathbf{v}_1, \mathbf{M})$
   (c) $\mathbf{d}_1 := \mathbf{u}_1$, $\mathbf{w}_0 := \mathbf{0}$, $\bar{\phi}_1 = \beta_1$, $\bar{\rho}_1 = \alpha_1$
2. *For $i = 1, 2, \ldots, i^\star$ do the following steps (or until convergence)*
   (a) $(\mathbf{v}_{i+1}, \widetilde{\mathbf{v}}_{i+1}, \beta_{i+1}) :=$ NORMALIZE$(\mathbf{B}\mathbf{u}_i - \alpha_i\widetilde{\mathbf{v}}_i, \mathbf{N})$
   (b) $(\mathbf{u}_{i+1}, \widetilde{\mathbf{u}}_{i+1}, \alpha_{i+1}) :=$ NORMALIZE$(\mathbf{B}^\top\mathbf{v}_{i+1} - \beta_{i+1}\widetilde{\mathbf{u}}_i, \mathbf{M})$
   (c) $\rho_i := \sqrt{\bar{\rho}_i^2 + \beta_{i+1}^2}$, $c_i := \bar{\rho}_i/\rho_i$, $s_i := \beta_{i+1}/\rho_i$
   (d) $\theta_{i+1} := s_i\alpha_{i+1}$, $\bar{\rho}_{i+1} := -c_i\alpha_{i+1}$, $\phi_i := c_i\bar{\phi}_i$, $\bar{\phi}_{i+1} := s_i\bar{\phi}_i$
   (e) $\mathbf{w}_i := \mathbf{w}_{i-1} + (\phi_i/\rho_i)\mathbf{d}_i$, $\mathbf{d}_{i+1} := \mathbf{u}_{i+1} - (\theta_{i+1}/\rho_i)\mathbf{d}_i$

NORMALIZE : $\mathbb{R}^K \times \mathbb{R}^{K\times K} \ni (\mathbf{s}, \mathbf{S}) \mapsto (\mathbf{z}, \widetilde{\mathbf{z}}, z) \in \mathbb{R}^K \times \mathbb{R}^K \times \mathbb{R}$, *with* $\mathbf{S}$ *s.p.d.*

1. *Solve* $\mathbf{S}\mathbf{s}^\star = \mathbf{s}$ *for* $\mathbf{s}^\star$. *Set* $z := \sqrt{\mathbf{s}^\top \mathbf{s}^\star}$ *and* $(\mathbf{z}, \widetilde{\mathbf{z}}) := (z^{-1}\mathbf{s}^\star, z^{-1}\mathbf{s})$

In exact arithmetic, the CG algorithm applied to (4.9) produces after k steps an approximate solution $\mathbf{w}_k \in \mathbb{R}^M$ with $\|\mathbf{M}^{1/2}\mathbf{w}_k - \mathbf{M}^{1/2}\mathbf{u}'\|_2 \leq C\beta^k$, where $\beta = \frac{\sqrt{\varkappa_2}-1}{\sqrt{\varkappa_2}+1}$, $\mathbf{u}' \in \mathbb{R}^M$ is the exact minimizer of (4.7), and $\varkappa_2$ is the condition number given in (4.6), see (Golub & Van Loan, 1996, Theorem 10.2.6). Hence, with $w_k := \mathcal{I}\mathbf{w}_k \in \mathcal{U}$ and $u' := \mathcal{I}\mathbf{u}' \in \mathcal{U}$, we have $\|w_k - u'\|_{\mathcal{X}} \sim \|\mathbf{w}_k - \mathbf{u}'\|_{\mathbf{M}} \leq C\beta^k$. If $M < \infty$, then $\mathbf{w}_M$ is the exact solution to (4.7), which is independent of $\mathcal{M}$, and we may assume $d_{\mathcal{M}} = 1 = D_{\mathcal{M}}$ using Example 4.1. Combining with Theorem 3.1, we obtain the following result.

Proposition 4.6. *Let* $\mathcal{U} \subseteq \mathcal{X}$, $\mathcal{V} \subseteq \mathcal{Y}$ *be closed subspaces satisfying the* inf-sup condition (3.1). *Assume norm equivalences* (4.1)–(4.2). *Let* $\mathbf{w}_k \in \mathbb{R}^M$ *be the* k*-th iterate of Algorithm 4.5 applied to the tuple* $(\mathbf{B}, \mathbf{M}, \mathbf{N}, \mathbf{f})$, *and* $u := B^{-1}f \in \mathcal{X}$. *Then* $w_k := \mathcal{I}\mathbf{w}_k \in \mathcal{U}$ *satisfies*

$$\|u - w_k\|_{\mathcal{X}} \leq C(\beta^k + \inf_{w \in \mathcal{U}} \|u - w\|_{\mathcal{X}}), \quad 0 \leq \beta < 1, \quad C > 0 \tag{4.10}$$

where β *and* C *only depend on the ratios* $\gamma_{\mathcal{U},\mathcal{V}}^{-1}\|B\|_{\mathcal{L}(\mathcal{X},\mathcal{Y}')}$, $d_{\mathcal{N}}^{-1}D_{\mathcal{N}}$, *and* $d_{\mathcal{M}}^{-1}D_{\mathcal{M}}$. *If* $M < \infty$, *then* w_M *satisfies* (4.10) *with* $\beta = 0$ *and a* $C > 0$ *independent of* $\mathcal{M}$.

5 Space-time tensor product spaces

For the purpose of this section assume we are given a sequence of nested subspaces $E_k \subseteq E_{k+1} \subset H^1(J)$, $k \in \mathbb{N}_0$, and $V_\ell \subseteq V_{\ell+1} \subset V$, $\ell \in \mathbb{N}_0$, such that $\sum_{k\in\mathbb{N}_0} E_k$ is dense in $H^1(J)$ and $\sum_{\ell\in\mathbb{N}_0} V_\ell$ is dense in V. For each $L, \Delta L \in \mathbb{N}_0$ and $\rho \in \mathbb{R}_{\geq 0} \cup \{\infty\}$ (with the convention $\infty \cdot 0 = 0$) we define

FTP) full tensor product spaces $\mathcal{U}_L^{(\rho)} \subset \mathcal{X}$ and $\mathcal{V}_{L,\Delta L}^{(\rho)} \subset \mathcal{Y}$ by

$$\mathcal{U}_L^{(\rho)} = \sum_{0\leq k,\rho\ell\leq L} E_k \otimes V_\ell, \quad \mathcal{V}_{L,\Delta L}^{(\rho)} = \sum_{0\leq k,\rho\ell\leq L} (E_{k+\Delta L} \otimes V_\ell) \times V_\ell \tag{5.1}$$

STP) sparse tensor product spaces $\widehat{\mathcal{U}}_L^{(\rho)} \subset \mathcal{X}$ and $\widehat{\mathcal{V}}_{L,\Delta L}^{(\rho)} \subset \mathcal{Y}$ by

$$\widehat{\mathcal{U}}_L^{(\rho)} = \sum_{0\leq k+\rho\ell\leq L} E_k \otimes V_\ell, \quad \widehat{\mathcal{V}}_{L,\Delta L}^{(\rho)} = \sum_{0\leq k+\rho\ell\leq L} (E_{k+\Delta L} \otimes V_\ell) \times V_\ell \tag{5.2}$$

where $k, \ell \in \mathbb{N}_0$ in all cases. The choice of ρ is discussed in Section 5.1.

5.1 Convergence rates

In this section we comment on convergence rates that can be expected for the Galerkin least squares solution (3.3) due to the quasi-optimality estimate (3.4). To state these, we assume existence of one-parametric families of Hilbert spaces $G^t \subseteq G^0$, $t \geq 0$, s.t. $(G^0, G^1) = (L^2(J), H^1(J))$ and $W^s \subseteq W^{-1}$, $s \geq -1$, s.t. $(W^{-1}, W^0, W^1) = (V', H, V)$. Further we assume existence of projectors $P_k : L^2(J) \to E_k$, $k \in \mathbb{N}_0$, and $Q_\ell : V' \to V_\ell$, $\ell \in \mathbb{N}_0$, with the properties

$$\|\mathrm{Id} - P_k\|_{\mathcal{L}(G^{t^\star}, G^t)} \lesssim 2^{-k(t^\star - t)} \quad \text{and} \quad \|\mathrm{Id} - Q_\ell\|_{\mathcal{L}(W^{s^\star}, W^s)} \lesssim 2^{-\ell(s^\star - s)/m}$$

for all $0 \le t \le t^\star$ and $-1 \le s \le s^\star$ with some $t^\star \ge 1$, $s^\star \ge 1$ and "order" $2m > 0$ of the operator $a_t(\cdot,\cdot)$. For simplicity, and in order to relate to the numerical experiments below, we specialize on the case $\dim E_k \sim 2^k$, $\dim V_\ell \sim 2^{d\ell}$, $d \in \mathbb{N}$. For sufficiently smooth solutions we will see the predicted convergence rates confirmed in the numerical examples in Section 7 with $W^{-1} = H^{-1}(D)$, $W^0 = L^2(D)$ and $W^s = H^s(D) \cap H^1_0(D)$, $s \ge 1$.

5.1.1 Full tensor product spaces

For $k = L$, $\ell = \max\{\ell \in \mathbb{N}_0 : \rho\ell \le L\}$, $L \in \mathbb{N}_0$ we obtain

$$\|\mathrm{Id} - P_k \otimes Q_\ell\|_{\mathcal{L}(G^{t^\star}\otimes W^{s^\star}, G^t \otimes W^s)} \lesssim 2^{-k(t^\star - t)} + 2^{-\ell(s^\star - s)/m} \lesssim (\dim \mathcal{U}_L^{(\rho)})^{-r}$$

with an admissible rate $r \ge 0$ to be determined. The choice $\rho = \frac{s^\star - s}{m(t^\star - t)}$ equilibrates the two errors, in which case we find the admissible values

$$r \le \frac{1}{\frac{1}{t^\star - t} + \frac{dm}{s^\star - s}} = \begin{cases} 1 & \text{for} \quad G^t \otimes W^s = L^2(J) \otimes H \\ \frac{2}{3} & \text{for} \quad G^t \otimes W^s = L^2(J) \otimes V \\ \frac{3}{4} & \text{for} \quad G^t \otimes W^s = H^1(J) \otimes V', \end{cases} \tag{5.3}$$

while the choice $\rho = 1$ yields

$$r \le \frac{1}{2} \min\left\{ t^\star - t, \frac{s^\star - s}{dm} \right\} = \begin{cases} 1 & \text{for} \quad G^t \otimes W^s = L^2(J) \otimes H \\ \frac{1}{2} & \text{for} \quad G^t \otimes W^s = L^2(J) \otimes V \\ \frac{1}{2} & \text{for} \quad G^t \otimes W^s = H^1(J) \otimes V'. \end{cases} \tag{5.4}$$

In both cases we have set $t^\star = s^\star = 2$, $d = 1$ and $m = 1$.

5.1.2 Sparse tensor product spaces

For $L \in \mathbb{N}_0$, $\rho \in \mathbb{R}_{\ge 0} \cup \{\infty\}$ consider the sparse tensor projector $\mathcal{Q}_L^{(\rho)} = \sum_{0 \le k + \rho\ell \le L} (Q_k - Q_{k-1}) \otimes (P_\ell - P_{\ell-1})$. In the case $t^\star = s^\star = 2$, $d = 1$, $m = 1$ and $\rho = 1$ we may choose any $r < 1$ to obtain

$$\|\mathrm{Id} - \mathcal{Q}_L^{(\rho)}\|_{\mathcal{L}(G^{t^\star}\otimes W^{s^\star}, G^t \otimes W^s)} \lesssim (\dim \widehat{\mathcal{U}}_L^{(\rho)})^{-r} \quad \text{for} \quad \begin{cases} G^t \otimes W^s = L^2(J) \otimes V \\ G^t \otimes W^s = H^1(J) \otimes V' \end{cases}$$

by standard arguments, and $\|\mathrm{Id} - \mathcal{Q}_L\|_{\mathcal{L}(G^{t^\star}\otimes W^{s^\star}, \mathcal{X})} \lesssim (\dim \widehat{\mathcal{U}}_L^{(\rho)})^{-r}$ by combining the two cases. Compared to the full tensor case (5.4) the approximation rate attainable for smooth solutions is therefore nearly doubled, cf. (5.4). More generally, spaces of the form given in Remark 6.4 need to be used. For a detailed discussion we refer to (Schwab & Stevenson, 2009, Section 7).

5.2 Hierarchic tensor product Riesz bases

Let ∇_k^Θ, $k \in \mathbb{N}_0$, and ∇_ℓ^Σ, $\ell \in \mathbb{N}_0$, be index sets s.t. $\{\theta_\lambda : \lambda \in \bigcup_{k' \le k} \nabla_{k'}^\Theta\}$ is a basis for E_k, $k \in \mathbb{N}_0$ and $\{\sigma_\mu : \mu \in \bigcup_{\ell' \le \ell} \nabla_{\ell'}^\Theta\}$ is a basis for V_ℓ, $\ell \in \mathbb{N}_0$. Then $\Theta := \{\theta_\lambda\}_{\lambda \in \nabla^\Theta}$ is a basis for $H^1(J)$ and $\Sigma := \{\sigma_\mu\}_{\mu \in \nabla^\Sigma}$ is a basis for V, where

$\nabla^\Theta := \bigcup_{k\in\mathbb{N}_0} \nabla_k^\Theta$ and $\nabla^\Sigma := \bigcup_{\ell\in\mathbb{N}_0} \nabla_\ell^\Sigma$. Defining, for each $L, \Delta L \in \mathbb{N}_0$ and $\rho \in \mathbb{R}_{\geq 0} \cup \{\infty\}$, the index sets

$$\nabla_L^{(\rho)} := \bigcup_{0\leq k,\rho\ell\leq L} \nabla_k^\Theta \times \nabla_\ell^\Sigma, \quad \nabla_{L,\Delta L}^{(\rho)} := \bigcup_{0\leq k,\rho\ell\leq L} \nabla_{k+\Delta L}^\Theta \times \nabla_\ell^\Sigma$$

and

$$\widehat{\nabla}_L^{(\rho)} := \bigcup_{0\leq k+\rho\ell\leq L} \nabla_k^\Theta \times \nabla_\ell^\Sigma, \quad \widehat{\nabla}_{L,\Delta L}^{(\rho)} := \bigcup_{0\leq k+\rho\ell\leq L} \nabla_{k+\Delta L}^\Theta \times \nabla_\ell^\Sigma,$$

$\mathcal{U}_L^{(\rho)}$ is spanned by $\theta_\lambda \otimes \sigma_\mu$, $(\lambda,\mu) \in \nabla_L^{(\rho)}$, and $\mathcal{V}_{L,\Delta L}^{(\rho)}$ is spanned by $(\theta_\lambda \otimes \sigma_\mu, \sigma_{\mu'})$, $(\lambda,\mu) \in \nabla_{L,\Delta L}^{(\rho)}$, $\mu' \in \bigcup_{\rho\ell\leq L} \nabla_\ell^\Sigma$, similarly for $\widehat{\mathcal{U}}_L^{(\rho)}$ and $\widehat{\mathcal{V}}_{L,\Delta L}^{(\rho)}$.
We have indicated in Example 4.1 and Example 4.2 how suitable operators $\mathcal{N}$ and $\mathcal{M}$ satisfying (4.1)–(4.2) can be obtained. For the rest of the paper, we focus on the construction relying on Riesz bases.

Assumption 5.1. *The collection $\Sigma \subset V$ can be rescaled to a Riesz basis for H, similarly for V. The collection Θ can be rescaled to a Riesz basis for $L^2(J)$.*

With this assumption, it is easy to see that the collection

$$\left\{\left(\frac{\theta_\lambda \otimes \sigma_\mu}{\|\theta_\lambda\|_{L^2(J)}\|\sigma_\mu\|_V}, 0\right)\right\}_{(\lambda,\mu)\in\nabla^\Theta\times\nabla^\Sigma} \cup \left\{\left(0, \frac{\sigma_\mu}{\|\sigma_\mu\|_H}\right)\right\}_{\mu\in\nabla^\Sigma} \tag{5.5}$$

is a Riesz basis for $\mathcal{Y}$. From this, we construct $\mathcal{N}$ as in Example 4.2.

Assumption 5.2. *In addition to Assumption 5.1, $\Sigma \subset V$ can be rescaled to a Riesz basis for V', and $\Theta \subset H^1(J)$ can be rescaled to a Riesz basis for $H^1(J)$.*

Under Assumption 5.2, (Griebel & Oswald, 1995, Proposition 1&2) imply that the collection $\left\{c_{\lambda\mu}^{-1}\theta_\lambda \otimes \sigma_\mu : (\lambda,\mu) \in \nabla^\Theta \times \nabla^\Sigma\right\}$, where $c_{\lambda\mu}^2 = \|\theta_\mu\|_H^2\|\sigma_\mu\|_V^2 + \|\theta_\lambda\|_{H^1(J)}^2\|\sigma_\mu\|_{V'}^2$, is a Riesz basis for $\mathcal{X}$. From this, we construct $\mathcal{M}$ analogously to Example 4.2.

Remark 5.3. *Piecewise polynomial bases on triangulations satisfying Assumption 5.1 have been constructed in e.g. (Nguyen, 2005), and bases on an interval satisfying Assumption 5.2 in e.g. (Donovan* et al.*, 1996). However, the requirements of Assumption 5.2 remain, from the practical point of view, much stronger than those of Assumption 5.1. This issue has been addressed in (Stevenson & Chegini, 2010) and in Example 4.3, for more details we refer to a forthcoming work (Andreev, ip). We emphasize, however, that $\mathcal{M}$ and $\mathcal{N}$ assume different roles in the resolution of the discrete weighted least squares problem* (4.7)*. While $\mathcal{N}$ ensures well-posedness and the quality of the exact solution via the quasi-optimality estimate* (3.4)*, $\mathcal{M}$ acts mainly as a preconditioner: the exact solution of* (4.7) *corresponds to the exact solution of* (3.4) *with constants c and C determined by the choice of $\mathcal{N}$.*

5.3 Parametric representation

For each $L, \Delta L \in \mathbb{N}_0$, $\rho \in \mathbb{R}_{\geq 0} \cup \{\infty\}$ we call $\mathbf{B}_{L,\Delta L}^{(\rho)}$ and $\widehat{\mathbf{B}}_{L,\Delta L}^{(\rho)}$ the system matrices corresponding to the pairs $(\mathcal{U}_L^{(\rho)}, \mathcal{V}_{L,\Delta L}^{(\rho)})$ and $(\widehat{\mathcal{U}}_L^{(\rho)}, \widehat{\mathcal{V}}_{L,\Delta L}^{(\rho)})$, constructed

as in (4.3) w.r.t. the unscaled tensor product bases, e.g.

$$\widehat{\mathbf{B}}^{(\rho)}_{L,\Delta L} = \begin{pmatrix} \Big(\langle B(\theta_\lambda \otimes \sigma_\mu), (\theta_{\lambda'} \otimes \sigma_{\mu'}, 0)\rangle_{\mathcal{Y}'\times\mathcal{Y}}\Big)_{(\lambda',\mu')\in\widehat{\nabla}^{(\rho)}_{L,\Delta L},(\lambda,\mu)\in\widehat{\nabla}^{(\rho)}_{L}} \\ \big(\theta_\lambda(0)\langle\sigma_\mu,\sigma_{\mu'}\rangle_H\big)_{\mu'\in\bigcup_{\rho\ell\le L}\nabla^\Sigma_\ell,(\lambda,\mu)\in\widehat{\nabla}^{(\rho)}_{L}} \end{pmatrix}. \tag{5.6}$$

Similarly, the load vectors are given by

$$\widehat{\mathbf{f}}^{(\rho)}_{L,\Delta L} = \begin{pmatrix} \Big(\langle F, (\theta_{\lambda'} \otimes \sigma_{\mu'}, 0)\rangle_{\mathcal{Y}'\times\mathcal{Y}}\Big)_{(\lambda',\mu')\in\widehat{\nabla}^{(\rho)}_{L,\Delta L}} \\ \big(\langle h,\sigma_{\mu'}\rangle_H\big)_{\mu'\in\bigcup_{\rho\ell\le L}\nabla^\Sigma_\ell} \end{pmatrix}, \tag{5.7}$$

and $\mathbf{f}^{(\rho)}_{L,\Delta L}$ analogously, replacing $\widehat{\nabla}$ by ∇. For notational convenience, in an expression like $\mathbf{N}^{-\top/2}\widehat{\mathbf{B}}^{(\rho)}_{L,\Delta L}\mathbf{M}^{-1/2}$ we implicitly restrict $\mathbf{M}$ and $\mathbf{N}$ to the indices $\widehat{\nabla}^{(\rho)}_L$ and $\widehat{\nabla}^{(\rho)}_{L,\Delta L}\times\bigcup_{\rho\ell\le L}\nabla^\Sigma_\ell$, respectively. We denote by $\mathbf{u}^{(\rho)}_{L,\Delta L}$, $u^{(\rho)}_{L,\Delta L} := \mathcal{I}\mathbf{u}^{(\rho)}_{L,\Delta L}$ and $\widehat{\mathbf{u}}^{(\rho)}_{L,\Delta L}$, $\widehat{u}^{(\rho)}_{L,\Delta L} := \mathcal{I}\widehat{\mathbf{u}}^{(\rho)}_{L,\Delta L}$ the corresponding discrete Galerkin least squares solutions of the minimization problem (4.7), where $\mathbf{N}$ is defined as in (4.3) w.r.t. respective bases.

6 Application to heat conduction

In this section we discuss one specific construction of subspaces $E_k \subseteq H^1(J)$, $k \in \mathbb{N}_0$, and $V_\ell \subseteq V$, $\ell \in \mathbb{N}_0$, from which we proceed to define $\mathcal{U}^{(\rho)}_L$, $\mathcal{V}^{(\rho)}_{L,\Delta L}$ and $\widehat{\mathcal{U}}^{(\rho)}_L$, $\widehat{\mathcal{V}}^{(\rho)}_{L,\Delta L}$ as in Section 5. These are shown in Proposition 6.3 to satisfy the *inf-sup condition* (3.1) uniformly in the choice of $L, \Delta L \in \mathbb{N}_0$, $\rho \in \mathbb{R}_{\ge 0}\cup\{\infty\}$.

6.1 Temporal discretization

We define the subspaces $E_k \subset H^1(J)$, $k \in \mathbb{N}_0$, consisting of all continuous, piecewise affine functions w.r.t. a uniform partition $\mathcal{T}_k := \{t^k_n := nh_k\}_{n=0}^{2^{k+1}} \subset \overline{J}$, where $h_k := T_{\mathrm{end}}2^{-(k+1)}$, of the interval J into 2^{k+1} subintervals. In particular, $\dim E_k = 2^{k+1}+1$ and $\dim E'_k = 2^{k+1}$ for all $k \in N_0$.

Proposition 6.1. *For all $\Delta k \in \mathbb{N}_0$: $\inf_{k\in\mathbb{N}_0}\mathcal{K}_{L^2(J)}(E'_k, E_{k+\Delta k}) \ge 1 - 2^{-\Delta k}$.*

Proof. The proof is provided in Appendix A. □

6.2 Spatial discretization

For the concrete example Example 2.2 of heat conduction we assume first that $D \subset \mathbb{R}^d$ is a bounded open polytope with a polyhedral boundary. We assume that we are given a sequence of nested triangulations $\mathcal{T}_\ell$, $\ell \in \mathbb{N}_0$, of D and define $V_\ell \subset V := H^1_0(D)$, $\ell \in \mathbb{N}_0$, as the space of piecewise polynomial continuous functions on D w.r.t. $\mathcal{T}_\ell$, and in particular $V_\ell \subseteq V_{\ell+1}$, $\ell \in \mathbb{N}_0$. We make the following assumption on the sequence V_ℓ, $\ell \in \mathbb{N}_0$.

Assumption 6.2. *The $L^2(D)$-orthogonal projector $Q_\ell : L^2(D) \to V_\ell$ is stable in $H^1(D)$ uniformly in $\ell \in \mathbb{N}_0$, i.e., $\sup_{\ell\in\mathbb{N}_0}\sup_{v\in V\setminus\{0\}}\frac{\|Q_\ell v\|_V}{\|v\|_V} =: \eta^{-1} < \infty$.*

For piecewise linear continuous functions w.r.t. quasi-uniform triangulations $\{\mathcal{T}_\ell\}_{\ell\in\mathbb{N}_0}$, among others, this assumption is indeed satisfied, see e.g. (Bramble *et al.*, 2002). From the results in Section 3, we obtain the following.

Proposition 6.3. *Let Assumption 6.2 hold. Let $\Delta L \in \mathbb{N}$. For any pair of subspaces $(\mathcal{U},\mathcal{V}) \in \{(\mathcal{U}_L^{(\rho)},\mathcal{V}_{L,\Delta L}^{(\rho)}),(\widehat{\mathcal{U}}_L^{(\rho)},\widehat{\mathcal{V}}_{L,\Delta L}^{(\rho)})\}_{L\in\mathbb{N}_0}$ the* inf-sup condition (3.1) *holds with $\gamma_{\mathcal{U},\mathcal{V}} \geq \tau\eta$, where $\tau = 1-2^{-\Delta L}$ and $\eta > 0$ is as in Assumption 6.2.*

Proof. By Proposition 6.1, we have $\mathcal{K}_{L^2(J)\times L^2(J)}(E_k', E_{k+\Delta L}) \geq \tau := 1 - 2^{-\Delta L}$. Let $\eta > 0$ be as in Assumption 6.2. By Proposition 3.5, *iii)* $\Rightarrow$ *ii)*, $\inf_{\ell\in\mathbb{N}_0} \mathcal{K}_{V'\times V}(V_\ell, V_\ell) \geq \eta$. The claim follows from Proposition 3.11. □

Remark 6.4. *The result remains valid for spaces of the form $\mathcal{U} = \widehat{\mathcal{U}}_L^{(\rho_1)} + \widehat{\mathcal{U}}_L^{(\rho_2)}$ and $\mathcal{V} = \widehat{\mathcal{V}}_{L,\Delta L}^{(\rho_1)} + \widehat{\mathcal{V}}_{L,\Delta L}^{(\rho_2)}$ with $\Delta L \in \mathbb{N}$ and $\rho_1, \rho_2 \in \mathbb{R}_{\geq 0} \cup \{\infty\}$.*

6.3 Biorthogonal spline wavelet bases

As anticipated in Section 5, in order to obtain efficient preconditioners for (4.7), we will employ Riesz bases. We give a simple construction on the interval here. For more general constructions see (Primbs, 2010) and references therein.
Set $\nabla_0^\Theta = \{(n,0)\}_{n=0}^2$, and for $k \in \mathbb{N}$ define $\nabla_k^\Theta = \{(n,k) : n \leq 2^{k+1} \text{ odd}\}$. For each $\lambda = (n,k) \in \nabla_k^\Theta$ with $k \in \mathbb{N}$, let $\theta_\lambda \in H^1(J)$ be the piecewise (w.r.t. $\mathcal{T}_k$) affine function which attains the values

$$\begin{pmatrix}\theta_\lambda(t_{n-1}^k)\\ \theta_\lambda(t_n^k)\\ \theta_\lambda(t_{n+1}^k)\end{pmatrix} = \begin{cases}\left(0,-\frac{3}{2},\frac{1}{2}\right)^\top & \text{if } t_{n-1}^k = 0,\\ \left(\frac{1}{2},-\frac{3}{2},0\right)^\top & \text{if } t_{n+1}^k = T_{\text{end}},\\ \left(\frac{1}{2},-1,\frac{1}{2}\right)^\top & \text{else,}\end{cases}$$

and zero at all other nodes. For $\lambda = (n,0) \in \nabla_0$ we define θ_λ as the standard nodal interpolant s.t. $\theta_{(n,0)}(t_{n'}^0) = \delta_{nn'}$, with linear interpolation in between. The collection $\{\theta_\lambda : \lambda \in \nabla_k, k \in \mathbb{N}_0\} \subset H^1(J)$ consists of biorthogonal piecewise linear spline wavelets and can be rescaled to a Riesz basis for the spaces for $L^2(J)$, similarly for $H^1(J)$, we refer to (Han & Shen, 2006), and thus satisfy Assumption 5.1–5.2.
In order to describe the basis for the space V, we specialize on the case that the spatial domain D is the interval $D = (-1,1) \subset \mathbb{R}$. Extension to product domains of the form $(-1,1)^d$ is straightforward; for more general constructions we refer to (Urban, 2009) and references therein. Here, we employ the same biorthogonal spline wavelet basis for the spatial discretization as in the temporal direction, adapting it slightly to conform with the homogeneous boundary conditions of $V = H_0^1(D)$. Thus, we set $\nabla_0^\Sigma := \{(1,0)\}$ and $\nabla_k^\Sigma := \{(n,k) : n \leq 2^{k+1} \text{ odd}\}$ for $k \in \mathbb{N}$ and let the collection $\{\sigma_\mu\}_{\mu\in\nabla_k^\Sigma} \subset V$ be the obvious adaptation of the functions $\{\theta_\lambda\}_{\lambda\in\nabla_k^\Theta}$ to the interval D for each $k \in \mathbb{N}_0$.
The collection Σ satisfies Assumption 5.1, but fails to satisfy Assumption 5.2. As discussed in Remark 5.3 this, however, is not critical for what follows. In fact, measuring the Riesz basis constants w.r.t. $\|\cdot\|_{V'}$ for $\{\sigma_\mu\}_{\mu\in\nabla_k^\Sigma}$ normalized in V' reveals that they deteriorate only moderately with k.

7 Numerical examples

In this section we complement the theory of previous sections by numerical examples, which demonstrate the efficiency of the proposed space-time wavelet discretization scheme and the potential of the sparse space-time Galerkin least squares method.

7.1 Preconditioning

We set $q \equiv 1$ and $T_{\mathrm{end}} = 2$ in (2.2)–(2.4), and measure the extreme singular values of the preconditioned system matrix $\mathbf{N}^{-\top/2}\mathbf{B}^{(\rho)}_{L,\Delta L}\mathbf{M}^{-1/2}$ for $\Delta L \in \{0, 1\}$. We set $\rho = \infty$, i.e., the spatial discretization is kept fixed at $\ell = 0$. Figure 7.1 illustrates that the introduction of an additional time discretization level in the test space $\mathcal{V}^{(\rho)}_{L,\Delta L}$ by choosing $\Delta L = 1$ renders the system (4.9) well-conditioned uniformly in $L \in \mathbb{N}_0$. Note that the condition number increases exponentially as $L \to \infty$ for $\Delta L = 0$ as the *temporal* resolution is increased, as was already indicated in (Babuška & Janik, 1990, Theorem 2.2.1).

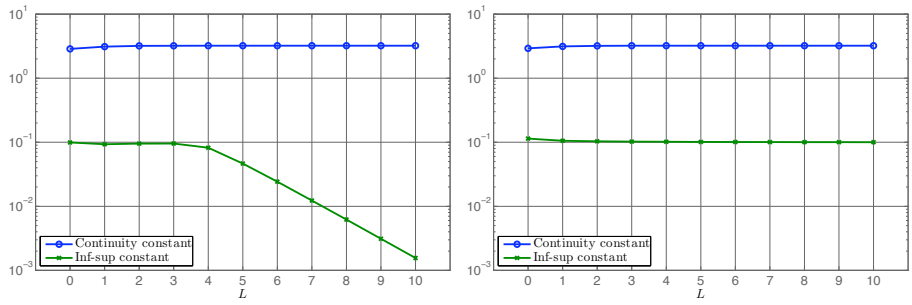

Figure 7.1: Maximal ("continuity constant") and minimal ("inf-sup constant") singular values of the preconditioned system matrix $\mathbf{B}^{(\rho)}_{L,\Delta L}$ on different discretization levels $L \in \mathbb{N}_0$ for $\rho = \infty$ with $\Delta L = 0$ (**left**) and $\Delta L = 1$ (**right**).

7.2 Smooth solution

We consider the heat equation with the solution $u(t, x) = \cos(x\pi/2)e^{-t}$, $(t, x) \in J \times D$, where $J = (0, T_{\mathrm{end}})$ with $T_{\mathrm{end}} = 2$. The initial condition $u(0, \cdot)$ is in $H^1_0(D)$, and the solution is in $H^{t^\star}(J) \otimes (H^{s^\star}(D) \cap H^1_0(D))$ for arbitrary $t^\star, s^\star \geq 1$, in particular for $t^\star = s^\star = 2$. For the definition of $\mathcal{U}^{(\rho)}_L$ and $\mathcal{V}^{(\rho)}_{L,\Delta L}$ in (5.1) we set $\Delta L = 1$. We measure the error of the Galerkin least squares solution $u^{(\rho)}_{L,\Delta L}$ for various $L \in \mathbb{N}_0$ in the space $L^2(J, H^1_0(D))$ for $\rho \in \{1, \frac{1}{2}\}$ and, in addition the error of the solution in $L^2(J, L^2(D))$ and the error in the initial condition in $L^2(D)$ for $\rho = 1$. The results are shown in Figure 7.2. For $\rho = 1$ we recover the rates given in (5.4), and for $\rho = \frac{1}{2}$ we recover the rate given in (5.3). The error in the initial condition decays with rate *one* w.r.t. the total number of degrees of freedom.

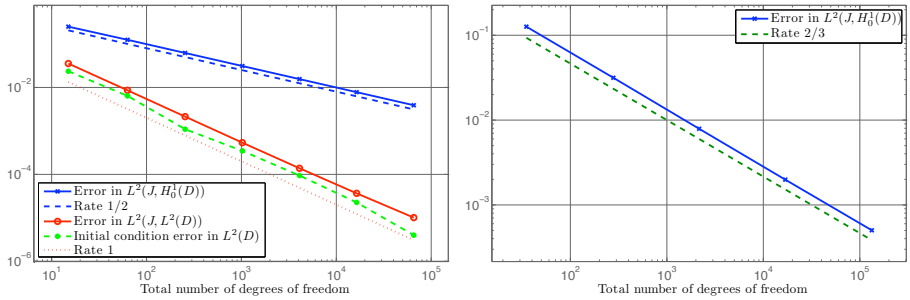

Figure 7.2: Convergence of the Galerkin least squares solution $u^{(\rho)}_{L,\Delta L}$ in different norms for a smooth problem from Section 7.2 with $\rho = 1$ (**left**) and $\rho = \frac{1}{2}$ (**right**).

7.3 Non-smooth coefficient

We consider equation (2.2)–(2.4) on $J \times D = (0, T_{\text{end}}) \times (-1, 1)$, $T_{\text{end}} = 2$ with r.h.s. $g \equiv 1$, initial condition $h \equiv 0$ and the piecewise constant $q \in L^\infty(J \times D)$ given by $q(t,x) = 1 - \frac{1}{2}\operatorname{sign}(2 + x - 2t)$, $(t,x) \in J \times D$. We set $\rho = 1$ and $\Delta L = 1$, and compute the Galerkin least squares solution $u^{(\rho)}_{L,\Delta L}$ for $L \in \{0, \dots, 7\}$ using Matlab's solver `lsqr` applied to (4.8) with relative residual tolerance 10^{-10}. The last solution $u^{(\rho)}_{7,\Delta L}$ is then used as the reference solution. The errors $u^{(\rho)}_{L,\Delta L} - u^{(\rho)}_{7,\Delta L}$ estimated in $\mathcal{X}$ using the norm equivalence (4.2), and measured in $L^2(J, H^1_0(D))$ and $L^2(J, L^2(D))$ are shown in Figure 7.3 (left) for $L \in \{0, \dots, 6\}$, together with the $L^2(D)$ error in the initial condition. All four converge; the rate, however, suffers from the low regularity of the solution. Timings of various algorithmic components for the computation of $u^{(\rho)}_{L,\Delta L}$ as function of the total number of degrees of freedom are shown in Figure 7.3 (right). As L increases, the time for the assembly of the right hand side $\mathbf{f}^{(\rho)}_{L,\Delta L}$ and for the assembly of the system matrix $\mathbf{B}^{(\rho)}_{L,\Delta L}$, which is done in parallel on 16 processes, scales linearly in the number of degrees of freedom, while the parallelization overhead in the assembly of the system matrix is visible for small L. Similarly, time per iteration of the least squares solver scales linearly in the number of degrees of freedom.

7.4 Nonsmooth initial data

In this example we choose the initial data $h \in L^2(D) \setminus H^1(D)$, namely $h(x) := \chi_{(-1,0)}(-x-1) + \chi_{(0,1)}(-x+1)$, a.e. $x \in D$, as well as $g \equiv 1$ and $q \equiv 1$. Using the Fourier series expansion of h it is easy to check that $h \in H^{1/2-\varepsilon}(D) \setminus H^{1/2}(D)$ for any $\varepsilon \in (0, \frac{1}{2})$, and the solution u does not belong to $H^1(J) \otimes (H^2(D) \cap H^1_0(D))$. With $\rho = 1$ and $\Delta L = 1$ in (5.1), the numerical solution converges slowly in $\mathcal{X}$, while for the initial data we obtain the optimal convergence rate $1/4$ in the total number of degrees of freedom, see Figure 7.4. Similar behavior arises for an *incompatible* initial data $h \in H^1(D) \setminus H^1_0(D)$.

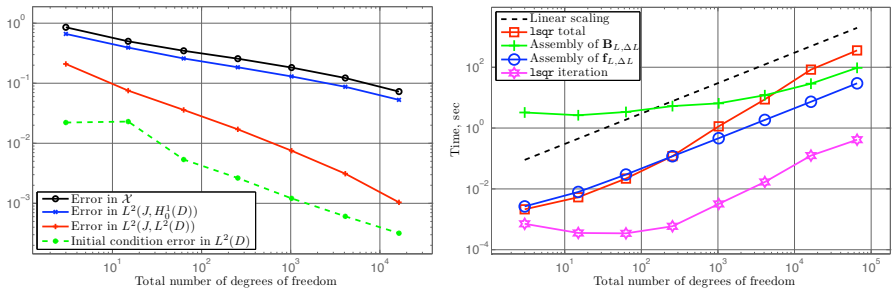

Figure 7.3: **Left**: Estimated errors of the Galerkin least squares solution for a non-smooth diffusion coefficient q, see Section 7.3. **Right**: Timings of the application of `lsqr`, assembly of the system matrix $\mathbf{B}_{L,\Delta L}^{(\rho)}$, the load vector $\mathbf{f}_{L,\Delta L}^{(\rho)}$ and of one iteration of `lsqr` as function of the total number of degrees of freedom for the example in Section 7.3.

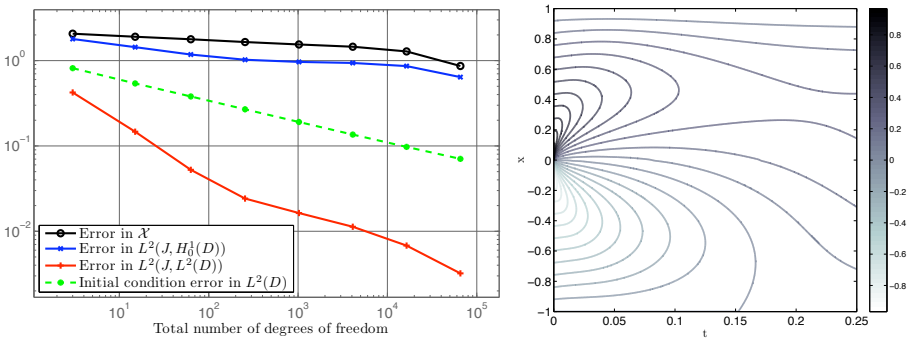

Figure 7.4: Estimated errors of the Galerkin least squares solution (**left**), and the solution (**right**) to (2.2)–(2.4) with $g \equiv 1$, $q \equiv 1$ and a nonsmooth initial data $h \notin H_0^1(D)$, see Section 7.4.

7.5 Sparse space-time tensor product

We choose $T_{\text{end}} = 2$, $q \equiv 1$ and $h \equiv 0$, and $g(t,x) = \sin(\pi t/2)^2 \cos(x+\cos(\pi t/2))$, $(t,x) \in J \times D$. The corresponding solution to (2.2)–(2.4) is displayed in Figure 7.5 (right). We set $\rho = 1$, $\Delta L = 1$ and estimate the error in $\mathcal{X}$ of the full tensor product (FTP) Galerkin least squares solution $u^{(\rho)}_{L,\Delta L}$ and the sparse tensor product (STP) Galerkin least squares solution $\widehat{u}^{(\rho)}_{L,\Delta L}$ for $L \in \{0,\dots,6\}$. To do so, we use $u^{(\rho)}_{7,\Delta L}$ as the reference solution and estimate the error in $\mathcal{X}$ using the norm equivalence (4.2). The results are shown in Figure 7.5 (left). In accordance with Section 5.1 we obtain the convergence rate 1/2 w.r.t. the total number of degrees of freedom in the FTP case and a rate which approaches *one* as L is increased in the STP case. For $L = 6$ the number of degrees of freedom used for the STP solution is over one order of magnitude smaller than for the FTP solution, without significant loss in accuracy.

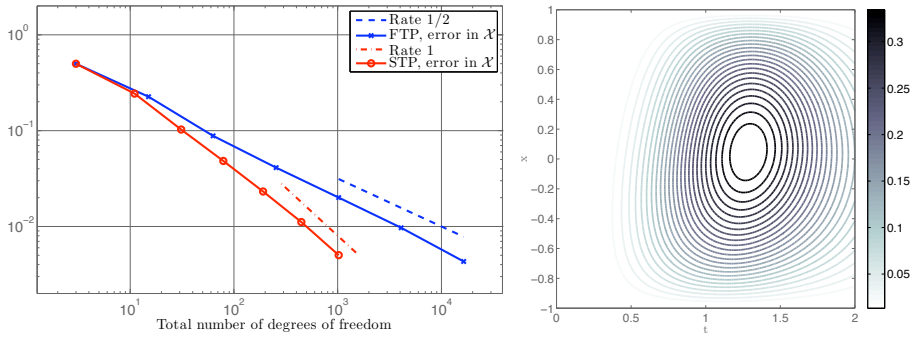

Figure 7.5: Estimated errors of the Galerkin least squares solutions $u^{(\rho)}_{L,\Delta L}$ (FTP) and $\widehat{u}^{(\rho)}_{L,\Delta L}$ (STP) (**left**), and the solution (**right**) to (2.2)–(2.4) with $q \equiv 1$, $h \equiv 0$, $\rho = 1$, $\Delta L = 1$ and a nonseparable source g as in Section 7.5.

8 Conclusions

For a class of linear parabolic equation we have proposed a finite element space-time sparse discretization, which reduces the problem to a finite, overdetermined linear system of equations in the generic case of the test space being sufficiently fine compared to the trial space. The scheme allows for space-time sparse trial and test spaces which potentially leads to a substantial reduction in the computational cost. We prove discrete stability of this discretization scheme. The corresponding normal equations can be efficiently preconditioned, e.g. by means of appropriate Riesz bases or based on the well-known BPX operator. The corresponding Galerkin least squares solution is shown to converge quasi-optimally to the continuous solution in the natural space for the original equation. The presented numerical examples illustrate and confirm the theory, but also expose some limitations, calling for space-time adaptive algorithms to capture singularities arising at the boundaries of the space-time cylinder.
In order to apply the presented Galerkin least squares discretization scheme to linear parabolic equations with more general generators satisfying the Gårding

inequality, it suffices to generalize Theorem 3.9 to include that case. Finally, we remark that we have specialized in Section 5 on finite element test *and* trial spaces that consist of functions continuous in time; this is not essential, since for a conforming method, only temporal continuity in the test space is expected. For more details on these remarks we refer to future work (Andreev, ip).
Numerous helpful comments on preliminary versions of this paper by Christoph Schwab, the anonymous referees, and several others are gratefully acknowledged.

References

Andreev, R. (i.p.) Multilevel Galerkin sparse space-time FEM for parametric parabolic problems (working title). in preparation.

Babuška, I. & Janik, T. (1990) The h-p Version of the Finite Element Method for Parabolic Equations. II. The h-p Version in Time. *Numerical Methods for Partial Differential Equations*, **6**, 343–369.

Benbow, S. J. (1999) Solving generalized least-squares problems with LSQR. *SIAM J. Matrix Anal. Appl.*, **21**, 166–177 (electronic).

Bochev, P. B. & Gunzburger, M. D. (2009) *Least-squares finite element methods*. Applied Mathematical Sciences, vol. 166. New York: Springer, pp. xxii+660.

Bramble, J. H., Pasciak, J. E. & Xu, J. (1990) Parallel multilevel preconditioners. *Math. Comp.*, **55**, 1–22.

Bramble, J. H., Pasciak, J. E. & Steinbach, O. (2002) On the stability of the L^2 projection in $H^1(\Omega)$. *Math. Comp.*, **71**, 147–156 (electronic).

Christensen, O. (2003) *An introduction to frames and Riesz bases*. Applied and Numerical Harmonic Analysis. Boston, MA: Birkhäuser Boston Inc., pp. xxii+440.

Demkowicz, L. & Gopalakrishnan, J. (2011) A class of discontinuous petrov–galerkin methods. ii. optimal test functions. *Numerical Methods for Partial Differential Equations*, **27**, 70–105.

Donovan, G. C., Geronimo, J. S. & Hardin, D. P. (1996) Intertwining multiresolution analyses and the construction of piecewise-polynomial wavelets. *SIAM J. Math. Anal.*, **27**, 1791–1815.

Golub, G. H. & Van Loan, C. F. (1996) *Matrix computations*. Johns Hopkins Studies in the Mathematical Sciences, third edn. Baltimore, MD: Johns Hopkins University Press, pp. xxx+698.

Griebel, M. & Oeltz, D. (2007) A sparse grid space-time discretization scheme for parabolic problems. *Computing*, **81**, 1–34.

Griebel, M. & Oswald, P. (1995) Tensor product type subspace splittings and multilevel iterative methods for anisotropic problems. *Adv. Comput. Math.*, **4**, 171–206.

Han, B. & Shen, Z. (2006) Wavelets with short support. *SIAM J. Math. Anal.*, **38**, 530–556 (electronic).

Hiptmair, R. (2006) Operator preconditioning. *Comput. Math. Appl.*, **52**, 699–706.

Horton, G. & Vandewalle, S. (1995) A space-time multigrid method for parabolic partial differential equations. *SIAM J. Sci. Comput.*, **16**, 848–864.

Lang, J. (2001) *Adaptive multilevel solution of nonlinear parabolic PDE systems*. Lecture Notes in Computational Science and Engineering, vol. 16. Berlin: Springer-Verlag, pp. xii+157. Theory, algorithm, and applications.

Lions, J.-L. & Magenes, E. (1972) *Non-homogeneous boundary value problems and applications. Vol. I*. New York: Springer-Verlag, pp. xvi+357. Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 181.

Mardal, K.-A. & Winther, R. (2010) Preconditioning discretizations of systems of partial differential equations. *Numer. Linear Algebra Appl.*

McLean, W. & Steinbach, O. (1999) Boundary element preconditioners for a hypersingular integral equation on an interval. *Adv. Comput. Math.*, **11**, 271–286.

Nguyen, H. (2005) Finite element wavelets for solving partial differential equations. *Ph.D. thesis*, Universiteit Utrecht.

Oswald, P. (1998) Multilevel norms for $H^{-1/2}$. *Computing*, **61**, 235–255.

Paige, C. C. & Saunders, M. A. (1982) LSQR: an algorithm for sparse linear equations and sparse least squares. *ACM Trans. Math. Software*, **8**, 43–71.

Primbs, M. (2010) New stable biorthogonal spline-wavelets on the interval. *Results Math.*, **57**, 121–162.

Schötzau, D. (1999) hp-DGFEM for parabolic evolution problems. *Ph.D. thesis*, ETH Zürich. Diss. Math.Wiss. ETH Zürich, Nr. 13041, 1999. Ref.: Christoph Schwab, Korref.: Jürg Marti, Korref.: Rolf Stenberg.

Schwab, C. & Stevenson, R. (2009) Space-time adaptive wavelet methods for parabolic evolution problems. *Math. Comp.*, **78**, 1293–1318.

Stevenson, R. & Chegini, N. G. (2010) Adaptive wavelet schemes for parabolic problems: Sparse matrices and numerical results. *Technical Report*. Korteweg-de Vries Institute for Mathematics, University of Amsterdam.

Thomée, V. (2006) *Galerkin finite element methods for parabolic problems*. Springer Series in Computational Mathematics, vol. 25, second edn. Berlin: Springer-Verlag, pp. xii+370.

Urban, K. (2009) *Wavelet methods for elliptic partial differential equations*. Numerical Mathematics and Scientific Computation. Oxford: Oxford University Press, pp. xxviii+480.

Xu, J. (1992) Iterative methods by space decomposition and subspace correction. *SIAM Rev.*, **34**, 581–613.

A Proof of Proposition 6.1

We show: for all integers $K \geq k \geq 0$ and all $e \in E_k$

$$\inf_{f \in E_K} \|e' - f\|^2_{L^2(J)} \leq 2^{-(K-k)} \left(\|e'\|^2_{L^2(J)} - \frac{1}{T} \left(\int_J e' \right)^2 \right), \tag{A.1}$$

with $J = (0,T) \subset \mathbb{R}$, and the estimate is sharp. From this, (6.1) then follows using the characterization (3.8). Let

$$\forall k \in \mathbb{N}_0 : \quad \mathcal{T}_k := \{0 = t_0^k < t_1^k < \cdots t_{N_k}^k < t_{N_k+1}^k = T\} \tag{A.2}$$

be the equidistant partition of J where $N_k := \dim E_k - 2$. Note that an $L^2(J)$-orthonormal basis for E_k' is given by the indicator functions $\chi_n^k := h_k^{-1/2} \chi_{(t_n^k, t_{n+1}^k)}$, $n = 0, \ldots, N_k$, where $h_k := \frac{T}{N_k+1}$. Fix arbitrary $k \in \mathbb{N}_0$ and $e \in E_k \subset H^1(J)$. Obviously, $e' \in L^2(J)$ is a piecewise constant function w.r.t. the partition $\mathcal{T}_K$ for any integer $K \geq k$, and thus

$$e' = \sum_{n=0}^{N_K} c_n^K \chi_n^K \quad \text{and} \quad \|e'\|^2_{L^2(J)} = \sum_{n=0}^{N_K} |c_n^K|^2 \quad \text{with} \quad \{c_n^K\}_{n=0}^{N_K} \subset \mathbb{R}.$$

We denote the jump of e' at any interior node $t_n^K \in \mathcal{T}_K \cap J$ by

$$\delta_n^K := e'(t_n^K+) - e'(t_n^K-) \quad \text{where} \quad e'(t\pm) := \lim_{h \downarrow 0} e'(t \pm h), \quad \forall t \in J$$

and let $\delta^K = (\delta_n^K)_{n=1}^{N_K} \in \mathbb{R}^{N_K}$ denote the vector collecting the jumps of e'. We can estimate the sum of squares of the jumps of e' in the interior of J by

$$\begin{aligned} \|\delta^K\|_2^2 = \|\delta^k\|_2^2 &\leq 4 \sum_{n=0}^{N_k} \|c_n^k \chi_n^k\|^2_{L^\infty(J)} - 2 \left(\|c_0^k \chi_0^k\|^2_{L^\infty(J)} + \|c_{N_k}^k \chi_{N_k}^k\|^2_{L^\infty(J)} \right) \\ &= \frac{4}{h_k} \|e'\|^2_{L^2(J)} - 2 \left(|e'(0+)|^2 + |e'(T-)|^2 \right) \end{aligned} \tag{A.3}$$

and the estimate is easily seen to be sharp for $c_n^k = c_0^k (-1)^n$. Let now $K \in \mathbb{N}_0$, $K \geq k$ be fixed. Having (A.3) in mind, we write $\delta := \delta^K$ for short. In order to estimate the l.h.s. of the assertion we construct a basis for the $L^2(J)$-orthogonal complement C_K of E_K in $E_K + E_K'$, i.e., $E_K + E_K' = E_K \oplus C_K$ with $E_K \perp_{L^2(J)} C_K$. For that purpose, we introduce the piecewise linear discontinuous $L^2(\mathbb{R})$ function

$$\psi : t \mapsto \psi(t) = \begin{cases} -1 - \frac{3}{2}t, & -1 < t < 0, \\ 1 - \frac{3}{2}t, & 0 < t < 1, \\ 0, & \text{else}, \end{cases}$$

which has jumps 1/2, 2, 1/2 at locations $t = -1, 0, 1$, respectively. We have $\|\psi\|^2_{L^2(\mathbb{R})} = \frac{1}{2}$. Moreover, ψ is orthogonal to the integer translates of the hat function $t \mapsto \max\{1 - |t-k|, 0\}$, $k \in \mathbb{Z}$. Consider the scaled translates of ψ,

$$\{\psi_n^K = \psi(\cdot/h_K - n)\}_{n=1,\ldots,N_K} \tag{A.4}$$

such that ψ_n^K is centered at t_n^K. It follows that (A.4) is a basis for C_K with the Gramian given by the tridiagonal matrix

$$M_\psi^K = \frac{1}{2} h_K \|\psi\|_{L^2(\mathbb{R})}^2 M_{\Delta\psi}^K \quad \text{where} \quad M_{\Delta\psi}^K := \begin{pmatrix} 2 & \frac{1}{2} & \\ \frac{1}{2} & \ddots & \ddots \\ & \ddots & 2 \end{pmatrix}.$$

Note that $M_{\Delta\psi}^K$ is exactly the matrix of jumps of the scaled translates (A.4) at the interior nodes $\mathcal{T}_K \cap J$. Let $d = (d_n)_{n=1}^{N_K} \in \mathbb{R}^{N_K}$ solve the equation $M_{\Delta\psi}^K d = \delta$. Note that $\|d\|_2 \leq \|\delta\|_2$, since for the spectrum of $M_{\Delta\psi}^K$ we have $\sigma(M_{\Delta\psi}^K) \subset [1, \infty)$ by the Gerschgorin disk theorem. Set $f := e' - \sum_{n=1}^{N_K} d_n \psi_n^K$. By definition of d, the function $f \in L^2(J)$ is piecewise linear w.r.t. the partition $\mathcal{T}_K$ and has no jumps in the interior of J, thus $f \in E_K$. In fact, f is the $L^2(J)$-orthogonal projection of e' onto E_K, since (A.4) is a basis for C_K. Using $M_\psi^K = \frac{1}{4} h_K M_{\Delta\psi}^K$, $\|d\|_2 \leq \|\delta\|_2$ and (A.3) we obtain

$$\|e' - f\|_{L^2(J)}^2 = d^\top M_\psi^K d = \frac{h_K}{4} d^\top \delta \leq \frac{h_K}{4} \|\delta\|_2^2 \leq \frac{h_K}{h_k} \|e'\|_{L^2(J)}^2. \tag{A.5}$$

This shows the claim if $e(T) = e(0)$. For a general $e \in E_k$ consider $w : t \mapsto w(t) = e(t) - e(0) - tb$, where $b = \frac{1}{T}(e(T) - e(0)) = \frac{1}{T} \int_J e'$. We have $\|w'\|_{L^2(J)}^2 = \|e'\|_{L^2(J)}^2 - b^2 T$, and $\inf_{f \in E_K} \|e' - f\|_{L^2(J)}^2 = \inf_{f \in E_K} \|w' + b - f\|_{L^2(J)}^2$ shows (A.1). Retracing the steps of the proof reveals that (A.1) is sharp. This completes the proof.

B List of figures

List of Figures

Research Reports

No.	Authors/Title
10-20	*R. Andreev* Space-time wavelet finite element method for parabolic equations
10-19	*V.H. Hoang and C. Schwab* Regularity and generalized polynomial chaos approximation of parametric and random 2nd order hyperbolic partial differential equations
10-18	*A. Barth, C. Schwab and N. Zollinger* Multi-Level Monte Carlo Finite Element method for elliptic PDE's with stochastic coefficients
10-17	*B. Kågström, L. Karlsson and D. Kressner* Computing codimensions and generic canonical forms for generalized matrix products
10-16	*D. Kressner and C. Tobler* Low-Rank tensor Krylov subspace methods for parametrized linear systems
10-15	*C.J. Gittelson* Representation of Gaussian fields in series with independent coefficients
10-14	*R. Hiptmair, J. Li and J. Zou* Convergence analysis of Finite Element Methods for $H(\mathrm{div};\Omega)$-elliptic interface problems
10-13	*M.H. Gutknecht and J.-P.M. Zemke* Eigenvalue computations based on IDR
10-12	*H. Brandsmeier, K. Schmidt and Ch. Schwab* A multiscale hp-FEM for 2D photonic crystal band
10-11	*V.H. Hoang and C. Schwab* Sparse tensor Galerkin discretizations for parametric and random parabolic PDEs. I: Analytic regularity and gpc-approximation
10-10	*V. Gradinaru, G.A. Hagedorn, A. Joye* Exponentially accurate semiclassical tunneling wave functions in one dimension
10-09	*B. Pentenrieder and C. Schwab* *hp*-FEM for second moments of elliptic PDEs with stochastic data. Part 2: Exponential convergence