Space-Time Adaptive Wavelet Methods for Parabolic Evolution Problems

C. Schwab and R. Stevenson[2]

---

[2]Korteweg-de Vries Institute for Mathematics, Plantage Muidergracht 24, 1018 TV Amsterdam, The Netherlands

Space-Time Adaptive Wavelet Methods for Parabolic Evolution Problems

C. Schwab and R. Stevenson[2]
Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

Abstract

With respect to space-time tensor-product wavelet bases, parabolic initial boundary value problems are equivalently formulated as bi-infinite matrix problems. Adaptive wavelet methods are shown to yield sequences of approximate solutions which converge at the optimal rate. In case the spatial domain is of product type, the use of spatial tensor product wavelet bases is proved to overcome the so-called curse of dimensionality, i.e., the reduction of the convergence rate with increasing spatial dimension.

Keywords: Parabolic differential equations, wavelets, adaptivity, optimal computational complexity, best N-term approximation, matrix compression

Subject Classification: 35K10, 41A25, 46B28, 65N99, 65T60

---

[2]Korteweg-de Vries Institute for Mathematics, Plantage Muidergracht 24, 1018 TV Amsterdam, The Netherlands

1. INTRODUCTION

Let V, H be separable Hilbert spaces such that $V \hookrightarrow H$ with dense embedding. Identifying H with its (anti-) dual, we obtain the Gelfand triple $V \hookrightarrow H \hookrightarrow V'$. We use the notation $\langle \cdot, \cdot \rangle_H$ both to denote the scalar product on $H \times H$ and its unique extension by continuity to the duality pairing on $V' \times V$.

Let $0 < T < \infty$ and denote, for a.e. $t \in I := [0, T]$, by $a(t; \cdot, \cdot)$ a sesqui-linear form on $V \times V$ such that for any $\eta, \zeta \in V$, $t \mapsto a(t; \eta, \zeta)$ is measurable on I, and such that for some constants $M_a, \alpha > 0$, $\lambda \in \mathbb{R}$ and for a.e. $t \in I$,

$$(1.1) \qquad |a(t; \eta, \zeta)| \leq M_a \|\eta\|_V \|\zeta\|_V \quad (\eta, \zeta \in V) \quad (boundedness),$$

$$(1.2) \qquad \Re a(t; \eta, \eta) + \lambda \|\eta\|_H^2 \geq \alpha \|\eta\|_V^2 \quad (\eta \in V) \quad (coercivity).$$

For a.e. $t \in I$, let $A(t) \in \mathcal{L}(V, V')$ be defined by

$$\langle A(t)\eta, \zeta \rangle_H = a(t; \eta, \zeta).$$

Given $g \in L_2(I; V')$ and $h \in H$, we are interested in solving the parabolic problem

$$(1.3) \qquad \frac{du}{dt}(t) + A(t)u(t) = g(t) \quad \text{in } V', \quad u(0) = h \text{ in } H.$$

We think of $A(t)$ as a linear, scalar differential or integrodifferential operator of order $2m > 0$ on a bounded domain $\Omega \subset \mathbb{R}^n$ in variational form (systems of equations will not impose any difficulties apart from a more involved notation). Then, $H = L_2(\Omega)$ and $V = H^m(\Omega)$ or a closed subspace incorporating homogeneous Dirichlet boundary conditions.

A classical approach to the numerical solution of (1.3) is the *Method of Lines* which reduces (1.3) by spatial semidiscretization in Ω with N degrees of freedom to a system of N coupled ordinary differential equations to be solved numerically in $(0, T)$ (see, e.g., [Tho06]). Conversely, in *Rothe's Method* (see e.g. [Lan01]) (1.3) is reduced by time semidiscretization to a sequence of coupled spatial, elliptic problems to be solved. Both these approaches, and the more recently proposed *discontinuous Galerkin method* (see, e.g., [EJT85]), are essentially time marching methods. Many results are known about efficient and reliable a posteriori error estimators (e.g. [EJ95, Ver96]), which are the basis of any adaptive solution method. Examples of such methods can be found in [EJ91, Pic98, CF04, Raa07].

The ultimate aim of adaptive methods is to achieve an approximate solution with error below a prescribed tolerance at the expense of, up to an absolute multiple, minimal amount of computer time and storage. Due to the character of time stepping this seems hard to realize and, unlike for elliptic problems, so far no optimality results are known to us.

In this work we follow an alternative approach. We give a space-time variational formulation of the parabolic problem (1.3) and prove that it defines a boundedly invertible linear operator between $\mathcal{X}$ and $\mathcal{Y}'$, with $\mathcal{X}$ and $\mathcal{Y}$ being (intersections of) tensor products of certain temporal and spatial Hilbert spaces. By equipping these spaces with wavelet bases and taking tensor products of these, the parabolic problem is reformulated as a well-posed bi-infinite matrix vector problem on the sequence space ℓ_2. We solve this problem using an adaptive wavelet method introduced by Cohen, Dahmen and DeVore in [CDD01, CDD02]. In a natural norm associated to the problem, we show that the approximations produced by this method converge with the same rate as the sequence of best approximations from the span of the best N products of temporal and spatial wavelets, in linear complexity.

While keeping discrete solutions on all time levels is prohibitive for time marching methods, thanks to the use of tensorized multi-level bases our method produces approximations simultaneously in space and time without penalty in complexity because of the additional time dimension.

The idea of using tensorized multi-level bases in space and time for solving parabolic problems was exploited earlier in [GO07], mainly in a non-adaptive setting (sparse-grids). An alternative space-time adaptive wavelet solver based on more heuristic arguments was presented in [AKV06]. Numerical results presented in these references indicate the potential of this approach.

Our approach is modular in the sense that the space discretization of $A(t)$ in (1.3) can be based on any 'spatial' wavelet system satisfying a set of conditions which we specify. We distinguish two particular cases: (A) the case of isotropically supported, piecewise polynomial wavelets in Ω, and (B) the case when $\Omega = (0,1)^n$ with possibly large n, where we use tensorized univariate spline wavelets as in [vPS04, GO07] which do not suffer from the curse of dimension for large n, thereby generalizing [DSS08] to the parabolic case.

The structure of this paper is as follows: in Section 2, we reformulate abstract well-posed linear operator equations involving separable Hilbert spaces $\mathcal{X}$ and $\mathcal{Y}$ as a well-posed bi-infinite matrix equation using Riesz bases of $\mathcal{X}$ and $\mathcal{Y}$ and in Section 3, we briefly elaborate on approximation classes related to best N-term approximations in such bases. Section 4 collects known results on adaptive wavelet methods, from [CDD01, CDD02, GHS07]; particular attention is paid to quantitative versions of s^*-admissibility and s^*-compressibility of matrix representations of the operator of interest. Section 5 addresses the reformulation of the parabolic problem (1.3) and establishes the well-posedness of the parabolic operator equation in the corresponding space-time Hilbert spaces, and Section 6 presents its reformulation as bi-infinite matrix equation in ℓ_2 with particular attention to the Riesz-constants. Section 7 gives the best possible rates of space-time tensor wavelet approximations for the solution and, in particular, establishes in case (B) above the absence of the curse of dimensionality. Section 8 establishes sufficient conditions on spatial and temporal Riesz bases for the s^*-computability of the space-time operator and Section 9 addresses in case (B) the dependence of the constants in the convergence and complexity bounds on the space dimension n. The Appendix contains an proof of well-posedness of our space-time variational formulation of (1.3).

In this paper, by $C \lesssim D$ we will mean that C can be bounded by a multiple of D, independently of parameters which C and D may depend on, in particular the space dimension n, apart from dependencies that are mentioned explicitly. Obviously, $C \gtrsim D$ is defined as $D \lesssim C$, and $C \eqsim D$ as $C \lesssim D$ and $C \gtrsim D$.

2. Well-posed operator equations and their reformulation as bi-infinite matrix vector problems

Let $\mathcal{X}, \mathcal{Y}$ be separable Hilbert spaces over $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Let us assume that we have available a Riesz basis $\Psi^{\mathcal{X}} = \{\psi_\lambda^{\mathcal{X}} : \lambda \in \nabla_{\mathcal{X}}\}$ for $\mathcal{X}$, meaning that the synthesis operator

$$s_{\Psi^{\mathcal{X}}} : \ell_2(\nabla_{\mathcal{X}}) \to \mathcal{X} : \mathbf{c} \mapsto \mathbf{c}^\top \Psi^{\mathcal{X}} := \sum_{\lambda \in \nabla_{\mathcal{X}}} c_\lambda \psi_\lambda^{\mathcal{X}}$$

is boundedly invertible. By identifying $\ell_2(\nabla_{\mathcal{X}})$ with its (anti-)dual, its adjoint, known as the analysis operator, reads as

$$s'_{\Psi^{\mathcal{X}}} : \mathcal{X}' \to \ell_2(\nabla_{\mathcal{X}}) : g \mapsto [g(\psi_\lambda^{\mathcal{X}})]_{\lambda \in \nabla_{\mathcal{X}}}.$$

Similarly, let $\Psi^{\mathcal{Y}} = \{\psi_\lambda^{\mathcal{Y}} : \lambda \in \nabla_{\mathcal{Y}}\}$ be a Riesz basis for $\mathcal{Y}$, with synthesis operator $s_{\Psi^{\mathcal{Y}}}$ and adjoint $s'_{\Psi^{\mathcal{Y}}}$. For both $\Psi^{\mathcal{X}}$ and $\Psi^{\mathcal{Y}}$ we have in mind suitable wavelet bases.

Now let $B \in \mathcal{L}(\mathcal{X}, \mathcal{Y}')$ be boundedly invertible. Given an $f \in \mathcal{Y}'$, we are interested in solving the operator equation of finding $u \in \mathcal{X}$ such that

$$Bu = f.$$

Writing $u = s_{\Psi^{\mathcal{X}}} \mathbf{u}$, this problem is equivalent to the bi-infinite matrix vector problem

$$(2.1) \qquad\qquad\qquad \mathbf{B}\mathbf{u} = \mathbf{f},$$

where $\mathbf{f} = s'_{\Psi^{\mathcal{Y}}} f = [f(\psi_\lambda^{\mathcal{Y}})]_{\lambda \in \nabla_{\mathcal{Y}}} \in \ell_2(\nabla_{\mathcal{Y}})$, and the *"stiffness"* or *system matrix*

$$\mathbf{B} = s'_{\Psi^{\mathcal{Y}}} B s_{\Psi^{\mathcal{X}}} = [(B\psi_\mu^{\mathcal{X}})(\psi_\lambda^{\mathcal{Y}})]_{\lambda \in \nabla_{\mathcal{Y}}, \mu \in \nabla_{\mathcal{X}}} \in \mathcal{L}(\ell_2(\nabla_{\mathcal{X}}), \ell_2(\nabla_{\mathcal{Y}}))$$

is boundedly invertible. Introducing the sesquilinear form

$$b : \mathcal{X} \times \mathcal{Y} \to \mathbb{K} : (w, v) \to (Bw)(v),$$

we will also use the notations

$$\mathbf{B} = b(\Psi^{\mathcal{X}}, \Psi^{\mathcal{Y}}) \quad \text{and} \quad \mathbf{f} = f(\Psi^{\mathcal{Y}}).$$

With the Riesz constants

$$\Lambda_{\Psi^{\mathcal{X}}}^{\mathcal{X}} := \|s_{\Psi^{\mathcal{X}}}\|_{\ell_2(\nabla_{\mathcal{X}}) \to \mathcal{X}} = \sup_{0 \neq \mathbf{c} \in \ell_2(\nabla_{\mathcal{X}})} \frac{\|\mathbf{c}^\top \Psi^{\mathcal{X}}\|_{\mathcal{X}}}{\|\mathbf{c}\|_{\ell_2(\nabla_{\mathcal{X}})}},$$

$$\lambda_{\Psi^{\mathcal{X}}}^{\mathcal{X}} := \|s_{\Psi^{\mathcal{X}}}^{-1}\|_{\mathcal{X} \to \ell_2(\nabla_{\mathcal{X}})}^{-1} = \inf_{0 \neq \mathbf{c} \in \ell_2(\nabla_{\mathcal{X}})} \frac{\|\mathbf{c}^\top \Psi^{\mathcal{X}}\|_{\mathcal{X}}}{\|\mathbf{c}\|_{\ell_2(\nabla_{\mathcal{X}})}},$$

and $\Lambda_{\Psi^{\mathcal{Y}}}^{\mathcal{Y}}$ and $\lambda_{\Psi^{\mathcal{Y}}}^{\mathcal{Y}}$ defined analogously, obviously it holds that

$$(2.2) \qquad\qquad \|\mathbf{B}\|_{\ell_2(\nabla_{\mathcal{X}}) \to \ell_2(\nabla_{\mathcal{Y}})} \leq \|B\|_{\mathcal{X} \to \mathcal{Y}'} \Lambda_{\Psi^{\mathcal{X}}}^{\mathcal{X}} \Lambda_{\Psi^{\mathcal{Y}}}^{\mathcal{Y}},$$

$$(2.3) \qquad\qquad \|\mathbf{B}^{-1}\|_{\ell_2(\nabla_{\mathcal{Y}}) \to \ell_2(\nabla_{\mathcal{X}})} \leq \frac{\|B^{-1}\|_{\mathcal{Y}' \to \mathcal{X}}}{\lambda_{\Psi^{\mathcal{X}}}^{\mathcal{X}} \lambda_{\Psi^{\mathcal{Y}}}^{\mathcal{Y}}}.$$

Some examples of operators B and spaces $\mathcal{X}$ and $\mathcal{Y}$ that one may have in mind are

- $(Bw)(v) = \int_\Omega \nabla w \cdot \nabla v$, $\mathcal{X} = \mathcal{Y} = H_0^1(\Omega)$ (Poisson problem),
- $(B(\vec{w}, p))(\vec{v}, q) = \int_\Omega \nabla \vec{w} : \nabla \vec{v} - \int_\Omega p \, \text{div} \, \vec{v} - \int_\Omega q \, \text{div} \vec{w}$, $\mathcal{X} = \mathcal{Y} = H_0^1(\Omega)^n \times L_{2,0}(\Omega)$ for a domain $\Omega \subset \mathbb{R}^n$ (Stokes problem),
- $(Bw)(v) = \frac{1}{4\pi} \int_{\partial\Omega} \int_{\partial\Omega} \frac{(w(y) - w(x))(v(y) - v(x))}{|x - y|^3} dx dy$, $\mathcal{X} = \mathcal{Y} = H^{\frac{1}{2}}(\partial\Omega)/\mathbb{R}$, $\Omega \subset \mathbb{R}^3$ (hypersingular boundary integral equation).

In this paper, we will see that also parabolic problems fit into this framework. In that case the spaces $\mathcal{X}$ and $\mathcal{Y}$ will not be equal.

3. Best N-term approximation and approximation classes

The most economical approximations for $\mathbf{u}$ are best N-term approximations $\mathbf{u}_N$, i.e., vectors that on their supports of length $N \in \mathbb{N}_0$ coincide with the N largest coefficients in modulus of $\mathbf{u}$. For $s > 0$, the approximation class $\mathcal{A}_\infty^s(\ell_2(\nabla_{\mathcal{X}})) := \{\mathbf{v} \in \ell_2(\nabla_{\mathcal{X}}) : \|\mathbf{v}\|_{\mathcal{A}_\infty^s(\ell_2(\nabla_{\mathcal{X}}))} < \infty\}$, where

$$\|\mathbf{v}\|_{\mathcal{A}_\infty^s(\ell_2(\nabla_{\mathcal{X}}))} := \sup_{\varepsilon > 0} \varepsilon \times [\min\{N \in \mathbb{N}_0 : \|\mathbf{v} - \mathbf{v}_N\|_{\ell_2(\nabla_{\mathcal{X}})} \le \varepsilon\}]^s$$

gathers under one roof all $\mathbf{v}$ whose best N-term approximations converge to $\mathbf{v}$ with rate s.

Generally, best N-term approximations cannot be realized in practice, in particular not in the situation that the vector $\mathbf{u}$ to be approximated is only implicitly given as the solution of the bi-infinite matrix vector problem (2.1). Our aim is to construct a practical method that produces approximations to $\mathbf{u}$ which, whenever $\mathbf{u} \in \mathcal{A}_\infty^s(\ell_2(\nabla_{\mathcal{X}}))$ for some $s > 0$, converge with this rate s in linear computational complexity.

4. Adaptive wavelet methods

Let $s > 0$ be such that $\mathbf{u} \in \mathcal{A}_\infty^s(\ell_2(\nabla_{\mathcal{X}}))$. In [CDD01] and [CDD02], adaptive wavelet methods for solving (2.1) were introduced. Both these methods are iterative methods. To be able to bound their complexity, one needs a suitable bound on the complexity of an approximate matrix-vector product in terms of the prescribed tolerance. We formalize this in the notion of s^*-admissibility.

Definition 4.1. $\mathbf{B} \in \mathcal{L}(\ell_2(\nabla_{\mathcal{X}}), \ell_2(\nabla_{\mathcal{Y}}))$ *is* s^*-admissible *if there exists a routine*

$$\mathbf{APPLY}_{\mathbf{B}}[\mathbf{w}, \varepsilon] \to \mathbf{z}$$

which yields, for any $\varepsilon > 0$ and any finitely supported $\mathbf{w} \in \ell_2(\nabla_{\mathcal{X}})$, a finitely supported $\mathbf{z} \in \ell_2(\nabla_{\mathcal{Y}})$ with $\|\mathbf{B}\mathbf{w} - \mathbf{z}\|_{\ell_2(\nabla_{\mathcal{Y}})} \le \varepsilon$ and for which, for any $\bar{s} \in (0, s^*)$, there exists an admissibility constant $a_{\mathbf{B},\bar{s}}$ such that $\#\mathrm{supp}\, \mathbf{z} \le a_{\mathbf{B},\bar{s}} \varepsilon^{-1/\bar{s}} \|\mathbf{w}\|_{\mathcal{A}_\infty^{\bar{s}}(\ell_2(\nabla_{\mathcal{X}}))}^{1/\bar{s}}$, and the number of arithmetic operations and storage locations used by the call $\mathbf{APPLY}_{\mathbf{B}}[\mathbf{w}, \varepsilon]$ is bounded by some absolute multiple of

$$a_{\mathbf{B},\bar{s}} \varepsilon^{-1/\bar{s}} \|\mathbf{w}\|_{\mathcal{A}_\infty^{\bar{s}}(\ell_2(\nabla_{\mathcal{X}}))}^{1/\bar{s}} + \#\mathrm{supp}\, \mathbf{w} + 1.$$

The realization of $\mathbf{APPLY}_{\mathbf{B}}[\mathbf{w}, \varepsilon]$ for system matrices $\mathbf{B}$ arising from parabolic problems will be a major topic in this work.

Remark 4.2. In Sect. 9 we investigate adaptive wavelet methods for solving parabolic problems (1.3) in high space dimensions, i.e. in $\Omega \subset \mathbb{R}^n$ with possibly large n. Introducing in Definition 4.1 the admissibility constants $a_{\mathbf{B},\bar{s}}$ will allow us to quantify the dimension-dependence in convergence and complexity estimates for the adaptive wavelet methods.

In order to approximate $\mathbf{u}$ one should be able to approximate $\mathbf{f}$. *Throughout this paper we assume availability of the following routine.*

$\mathbf{RHS}_{\mathbf{f}}[\varepsilon] \to \mathbf{f}_\varepsilon$: *For given $\varepsilon > 0$, it yields a finitely supported $\mathbf{f}_\varepsilon \in \ell_2(\nabla_{\mathcal{Y}})$ with*

$$\|\mathbf{f} - \mathbf{f}_\varepsilon\|_{\ell_2(\nabla_{\mathcal{Y}})} \le \varepsilon \quad and \quad \#\mathrm{supp}\, \mathbf{f}_\varepsilon \lesssim \min\{N : \|\mathbf{f} - \mathbf{f}_N\| \le \varepsilon\},$$

with the number of arithmetic operations and storage locations used by the call $\mathbf{RHS}_{\mathbf{f}}[\varepsilon]$ bounded by some absolute multiple of $\#\mathrm{supp}\, \mathbf{f}_\varepsilon + 1$.

In the following, we record some consequences of having $\mathbf{APPLY_B}$ and $\mathbf{RHS_f}$ routines. A proof of Proposition 4.3 can be given along the lines in [CDD01, CDD02] or [DFR+07, Prop. 3.3].

Proposition 4.3. *Let $\mathbf{B}$ in (2.1) be s^*-admissible. Then for any $\bar{s} \in (0, s^*)$, we have $\|\mathbf{B}\|_{\mathcal{A}^{\bar{s}}_\infty(\ell_2(\nabla_\mathcal{X})) \to \mathcal{A}^{\bar{s}}_\infty(\ell_2(\nabla_\mathcal{Y}))} \leq a^{\bar{s}}_{\mathbf{B},\bar{s}}$. For $\mathbf{z}_\varepsilon := \mathbf{APPLY_B}[\mathbf{w}, \varepsilon]$, we have $\|\mathbf{z}_\varepsilon\|_{\mathcal{A}^{\bar{s}}_\infty(\ell_2(\nabla_\mathcal{Y}))} \leq a^{\bar{s}}_{\mathbf{B},\bar{s}}\|\mathbf{w}\|_{\mathcal{A}^{\bar{s}}_\infty(\ell_2(\nabla_\mathcal{X}))}$.*

Using the definition of $\mathcal{A}^s_\infty(\ell_2(\nabla_\mathcal{Y}))$ and the properties of $\mathbf{RHS_f}$, we have the following corollary.

Corollary 4.4. *If, in (2.1), $\mathbf{B}$ is s^*-admissible and $\mathbf{u} \in \mathcal{A}^s_\infty(\ell_2(\nabla_\mathcal{X}))$ for $s < s^*$ then for $\mathbf{f}_\varepsilon = \mathbf{RHS_f}[\varepsilon]$, $\#\mathrm{supp}\,\mathbf{f}_\varepsilon \lesssim a_{\mathbf{B},s}\varepsilon^{-1/s}\|\mathbf{u}\|^{1/s}_{\mathcal{A}^s_\infty(\ell_2(\nabla_\mathcal{X}))}$ with the number of arithmetic operations and storage locations used by the call $\mathbf{RHS_f}[\varepsilon]$ being bounded by some absolute multiple of*

$$a_{\mathbf{B},s}\varepsilon^{-1/s}\|\mathbf{u}\|^{1/s}_{\mathcal{A}^s_\infty(\ell_2(\nabla_\mathcal{X}))} + 1.$$

Remark 4.5. Besides $\|\mathbf{f} - \mathbf{f}_\varepsilon\|_{\ell_2(\nabla_\mathcal{Y})} \leq \varepsilon$, the complexity bounds in Corollary 4.4, with $a_{\mathbf{B},s}$ reading as some constant, are essential for the use of $\mathbf{RHS_f}$ in the adaptive wavelet methods.

The following corollary of Proposition 4.3 can be used for example for the construction of valid $\mathbf{APPLY}$ and $\mathbf{RHS}$ routines in case the adaptive wavelet algorithms are applied to a preconditioned system.

Corollary 4.6. *If $\mathbf{B} \in \mathcal{L}(\ell_2(\nabla_\mathcal{X}), \ell_2(\nabla_\mathcal{Y}))$, $\mathbf{C} \in \mathcal{L}(\ell_2(\nabla_\mathcal{Y}), \ell_2(\nabla_\mathcal{Z}))$ are both s^*-admissible, then so is $\mathbf{CB} \in \mathcal{L}(\ell_2(\nabla_\mathcal{X}), \ell_2(\nabla_\mathcal{Z}))$. A valid routine $\mathbf{APPLY_{CB}}$ is*

$$(4.1) \qquad [\mathbf{w}, \varepsilon] \mapsto \mathbf{APPLY_C}\big[\mathbf{APPLY_B}[\mathbf{w}, \varepsilon/(2\|\mathbf{C}\|)], \varepsilon/2\big],$$

with admissibility constant $a_{\mathbf{CB},\bar{s}} \lesssim a_{\mathbf{B},\bar{s}}(\|\mathbf{C}\|^{1/\bar{s}} + a_{\mathbf{C},\bar{s}})$ for $\bar{s} \in (0, s^*)$.

For some $s^* > s$, let $\mathbf{C} \in \mathcal{L}(\ell_2(\nabla_\mathcal{Y}), \ell_2(\nabla_\mathcal{Z}))$ be s^*-admissible. Then for

$$(4.2) \qquad \mathbf{RHS_{Cf}}[\varepsilon] := \mathbf{APPLY_C}[\mathbf{RHS_f}[\varepsilon/(2\|\mathbf{C}\|)], \varepsilon/2],$$

there holds that $\#\mathrm{supp}\,\mathbf{RHS_{Cf}}[\varepsilon] \lesssim a_{\mathbf{B},s}(\|\mathbf{C}\|^{1/s} + a_{\mathbf{C},s})\varepsilon^{-1/s}\|\mathbf{u}\|^{1/s}_{\mathcal{A}^s_\infty(\ell_2(\nabla_\mathcal{X}))}$ and $\|\mathbf{Cf} - \mathbf{RHS_{Cf}}[\varepsilon]\|_{\ell_2(\nabla_\mathcal{Z})} \leq \varepsilon$, with the number of arithmetic operations and storage locations used by the call $\mathbf{RHS_{Cf}}[\varepsilon]$ bounded by some absolute multiple of

$$a_{\mathbf{B},s}(\|\mathbf{C}\|^{1/s} + a_{\mathbf{C},s})\varepsilon^{-1/s}\|\mathbf{u}\|^{1/s}_{\mathcal{A}^s_\infty(\ell_2(\nabla_\mathcal{X}))} + 1.$$

Remark 4.7. The properties of $\mathbf{RHS_{Cf}}$ given in the above corollary show that it is a valid routine for approximating $\mathbf{Cf}$ in the sense of Remark 4.5.

Consider first the case that $\mathbf{B}$ is *symmetric positive definite* $\mathbf{B}$, i.e., $\nabla_\mathcal{X} = \nabla_\mathcal{Y}$ and $\mathbf{B} = \mathbf{B}^* > 0$. In this situation, both adaptive wavelet methods from [CDD01, CDD02] were shown to be *quasi-optimal* in the following sense:

Theorem 4.8. *If in (2.1) $\mathbf{B}$ is s^* admissible, then for any $\varepsilon > 0$, both adaptive wavelet methods from [CDD01, CDD02] produce an approximation $\mathbf{u}_\varepsilon$ to $\mathbf{u}$ with $\|\mathbf{u} - \mathbf{u}_\varepsilon\|_{\ell_2(\nabla_\mathcal{X})} \leq \varepsilon$. If in (2.1) for some $s > 0$ it holds $\mathbf{u} \in \mathcal{A}^s_\infty(\ell_2(\nabla_\mathcal{X}))$, then $\#\mathrm{supp}\,\mathbf{u}_\varepsilon \lesssim \varepsilon^{-1/s}\|\mathbf{u}\|^{1/s}_{\mathcal{A}^s_\infty(\ell_2(\nabla_\mathcal{X}))}$ and if, moreover, $s < s^*$ then the number of arithmetic operations and storage locations required by a call of either of these adaptive wavelet solvers with tolerance ε is bounded by some multiple of*

$$\varepsilon^{-1/s}(1 + a_{\mathbf{B},s})\|\mathbf{u}\|^{1/s}_{\mathcal{A}^s_\infty(\ell_2(\nabla_\mathcal{X}))} + 1.$$

The multiples depend only on s when it tends to 0 or ∞, and on $\|\mathbf{B}\|$ and $\|\mathbf{B}^{-1}\|$ when they tend to infinity.

The method from [CDD02] consists of the application of a damped Richardson iteration to $\mathbf{Bu} = \mathbf{f}$, where the required residual computations are approximated using calls of $\mathbf{APPLY_B}$ and $\mathbf{RHS_f}$ within tolerances that decrease linearly with the iteration counter.

With the method from [CDD01], a sequence $\Xi_0 \subset \Xi_1 \subset \cdots \subset \nabla_{\mathcal{X}}$ is produced, together with corresponding (approximate) Galerkin solutions $\mathbf{u}_i \in \ell_2(\Xi_i)$. The coefficients of approximate residuals $\mathbf{f} - \mathbf{Bu}_i$ are used as indicators how to expand Ξ_i to Ξ_{i+1} such that it gives rise to an improved Galerkin approximation.

Both methods rely on a recurrent coarsening of the approximation vectors, where small coefficients are removed in order to keep an optimal balance between accuracy and support length. In [GHS07], it was shown that with the method from [CDD01] coarsening can be avoided, which gives a quantitative advantage.

The key why s^*-admissibility of $\mathbf{B}$ can be expected is the observation that for a wide class of operators the stiffness matrix with respect to suitable wavelet bases is close to a computable sparse matrix.

Definition 4.9. $\mathbf{B} \in \mathcal{L}(\ell_2(\nabla_{\mathcal{X}}), \ell_2(\nabla_{\mathcal{Y}}))$ is s^*-*computable* if, for each $N \in \mathbb{N}$, there exists a $\mathbf{B}_N \in \mathcal{L}(\ell_2(\nabla_{\mathcal{X}}), \ell_2(\nabla_{\mathcal{Y}}))$ having in each column at most N non-zero entries whose joint computation takes an absolute multiple of N operations, such that the *computability constants*

$$c_{\mathbf{B},\bar{s}} := \sup_{N \in \mathbb{N}} N \|\mathbf{B} - \mathbf{B}_N\|_{\ell_2(\nabla_{\mathcal{X}}) \to \ell_2(\nabla_{\mathcal{Y}})}^{1/\bar{s}}$$

are finite for any $\bar{s} \in (0, s^*)$.

Theorem 4.10. *An s^*-computable $\mathbf{B}$ is s^*-admissible. Moreover, for $\bar{s} < s^*$, $a_{\mathbf{B},\bar{s}} \lesssim c_{\mathbf{B},\bar{s}}$ where the constant in this estimate depends only on $\bar{s} \downarrow 0$, $\bar{s} \uparrow s^*$, and on $\|\mathbf{B}\| \to \infty$.*

This theorem is proven by the construction of a suitable $\mathbf{APPLY_B}$ routine as was done in [CDD01, §6.4], where a log factor in the complexity estimate due to sorting was removed in later studies by the application of an approximate sorting, see [Bar05, Met02, Ste03]. In [DSS08] some modifications to this approximate matrix vector routine were proposed that give rise to quantitative improvements.

Remark 4.11. Theorem 4.10 has been shown under the tacit assumption that a virtually unbounded amount of memory is available so that direct addressing can be applied. We refer to [DSS08, §5] for a further discussion of this issue.

Remark 4.12. Theorem 4.8 requires that $\mathbf{B}$ is s^*-admissible for an $s^* > s$ when $\mathbf{u} \in \mathcal{A}_\infty^s(\ell_2(\nabla_{\mathcal{X}}))$. Generally this value of s is unknown, and so the condition on s^* should be interpreted in the sense that s^* has to be larger than any s for which membership of the solution $\mathbf{u}$ in $\mathcal{A}_\infty^s(\ell_2(\nabla_{\mathcal{X}}))$ can be expected. For example, for a scalar elliptic equation of order $2m$ in n space dimensions with (isotropic) wavelets of order d, such s do not exceed $\frac{d-m}{n}$. For wavelets that have sufficiently many vanishing moments and are sufficiently regular, for a wide class of operators that include the examples mentioned in Section 2, s^*-computability and with that s^*-admissibility has been demonstrated for an s^* that indeed is sufficiently large in the aforementioned sense (see [GS06a, GS06b]).

So far we confined the discussion to stiffness matrices that are symmetric positive definite. In [Gan06], it was shown that the results concerning the method from [CDD01]/[GHS07] extend to the situation that nonsymmetric perturbations of lower order are added. The approach from [CDD02] basically applies whenever one has a linearly convergent stationary iterative scheme for $\mathbf{B}\mathbf{u} = \mathbf{f}$ available. There is, however, no recipe that yields such a scheme for general boundedly invertible $\mathbf{B}$. In particular, for the stiffness matrices $\mathbf{B}$ that will result from parabolic problems, we are not aware of such schemes.

A remedy proposed in [CDD02] is to apply the adaptive schemes to the *normal equations*

$$(4.3) \qquad \mathbf{B}^*\mathbf{B}\mathbf{u} = \mathbf{B}^*\mathbf{f}.$$

instead. Clearly the operator $\mathbf{B}^*\mathbf{B} \in \mathcal{L}(\ell_2(\nabla_\mathcal{X}), \ell_2(\nabla_\mathcal{X}))$ is boundedly invertible, symmetric positive definite, with $\|\mathbf{B}^*\mathbf{B}\|_{\ell_2(\nabla_\mathcal{X}) \to \ell_2(\nabla_\mathcal{X})} \leq \|\mathbf{B}\|^2_{\ell_2(\nabla_\mathcal{X}) \to \ell_2(\nabla_\mathcal{Y})}$ and $\|(\mathbf{B}^*\mathbf{B})^{-1}\|_{\ell_2(\nabla_\mathcal{X}) \to \ell_2(\nabla_\mathcal{X})} \leq \|\mathbf{B}^{-1}\|^2_{\ell_2(\nabla_\mathcal{Y}) \to \ell_2(\nabla_\mathcal{X})}$. Now let $\mathbf{u} \in \mathcal{A}^s_\infty(\ell_2(\nabla_\mathcal{X}))$, and for some $s^* > s$, let $\mathbf{B}$ and $\mathbf{B}^*$ be s^*-admissible. By Corollary 4.6, with $\mathbf{B}^*$ in place of $\mathbf{C}$, a valid $\mathbf{RHS}_{\mathbf{B}^*\mathbf{f}}$ routine is given by (4.2), and $\mathbf{B}^*\mathbf{B}$ is s^*-admissible with a valid $\mathbf{APPLY}_{\mathbf{B}^*\mathbf{B}}$ routine given by (4.1). A combination of Theorem 4.8 and Corollary 4.6 yields the following result.

Theorem 4.13. *For any $\varepsilon > 0$, the adaptive wavelet methods from [CDD02] or [CDD01]/[GHS07] applied to the normal equations (4.3) using above $\mathbf{APPLY}_{\mathbf{B}^*\mathbf{B}}$ and $\mathbf{RHS}_{\mathbf{B}^*f}$ routines produce an approximation $\mathbf{u}_\varepsilon$ to $\mathbf{u}$ with $\|\mathbf{u} - \mathbf{u}_\varepsilon\|_{\ell_2(\nabla_\mathcal{X})} \leq \varepsilon$. If for some $s > 0$, $\mathbf{u} \in \mathcal{A}^s_\infty(\ell_2(\nabla_\mathcal{X}))$, then $\#\mathrm{supp}\,\mathbf{u}_\varepsilon \lesssim \varepsilon^{-1/s}\|\mathbf{u}\|^{1/s}_{\mathcal{A}^s_\infty(\ell_2(\nabla_\mathcal{X}))}$, with constant only dependent on s when it tends to 0 or ∞, and on $\|\mathbf{B}\|$ and $\|\mathbf{B}^{-1}\|$ when they tend to infinity.*

If $s < s^*$, then the number of arithmetic operations and storage locations required by a call of either of these adaptive wavelet methods with tolerance ε is bounded by some multiple of

$$1 + \varepsilon^{-1/s}(1 + a_{\mathbf{B},s}(1 + a_{\mathbf{B}^*,s}))\|\mathbf{u}\|^{1/s}_{\mathcal{A}^s_\infty(\ell_2(\nabla_\mathcal{X}))}$$

where this multiple only depends on s when it tends to 0 or ∞, and on $\|\mathbf{B}\|$ and $\|\mathbf{B}^{-1}\|$ when they tend to infinity.

5. Variational formulation of the parabolic problem

The variational form of (1.3) will be based on the space

$$(5.1) \qquad \begin{aligned} \mathcal{X} &= L_2(I;V) \cap H^1(I;V') \\ &= \{v : v \in L_2(I;V), v, \tfrac{dv}{dt} \in L_2(I;V')\} \end{aligned}$$

equipped with the graph norm $\| \circ \|_\mathcal{X}$ given by

$$(5.2) \qquad \|v\|_\mathcal{X} := \left(\|v\|^2_{L_2(I;V)} + \|\tfrac{dv}{dt}\|^2_{L_2(I;V')}\right)^{\frac{1}{2}}.$$

It is known that $\mathcal{X} \hookrightarrow C([0,T];H)$ (e.g. [DL92, Ch.XVIII, §1, Th.1]). Inspection of the proof reveals that

$$(5.3) \qquad M_e := \sup_{0 \neq w \in \mathcal{X}} \frac{\|w(0)\|_H}{\|w\|_\mathcal{X}}$$

is bounded uniformly in the choice of $V \hookrightarrow H$, and is only dependent on T when it tends to zero.

By integrating (1.3) over $t \in I$ we arrive at the variational formulation of the initial-boundary value problem (1.3): find

$$(5.4) \qquad u \in \mathcal{X}: \qquad b(u, v) = f(v) \qquad (v = (v_1, v_2) \in \mathcal{Y})$$

where the "test space" $\mathcal{Y}$ is

$$(5.5) \qquad \qquad \mathcal{Y} = L_2(I; V) \times H$$

equipped with norm $\| \circ \|_{\mathcal{Y}}$ given by $\|v\|_{\mathcal{Y}}^2 = \|v_1\|_{L_2(I;V)}^2 + \|v_2\|_H^2$, and the bilinear form $b(\cdot, \cdot) : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ is defined by

$$(5.6) \qquad b(w, (v_1, v_2)) := \int_I \langle \tfrac{dw}{dt}(t), v_1(t) \rangle_H + a(t; w(t), v_1(t)) dt + \langle w(0), v_2 \rangle_H,$$

and the "load functional" $f(\cdot) : \mathcal{Y} \to \mathbb{R}$ is given by

$$(5.7) \qquad f(v) := \int_I \langle g(t), v_1(t) \rangle_H dt + \langle h, v_2 \rangle_H \quad \text{for} \quad v = (v_1, v_2) \in \mathcal{Y}.$$

Theorem 5.1. *The operator $B \in \mathcal{L}(\mathcal{X}, \mathcal{Y}')$ defined by $(Bw)(v) = b(w, v)$ with $b(\cdot, \cdot)$, $\mathcal{X}$ and $\mathcal{Y}$ from (5.6), (5.1) and (5.5) is boundedly invertible. With $\varrho := \sup_{0 \neq v \in V} \frac{\|v\|_H}{\|v\|_V}$, the norm of B or B^{-1} can be bounded by an increasing function of M_a, M_e and λT, or of α^{-1}, M_a, M_e and λT and $\lambda \varrho^2$, respectively.*

Although formulated slightly differently, a proof of this theorem can be found in [DL92, Ch.XVIII, §3] and [Wlo82, Ch.IV, §26] (without the statement about the bounds on the norms of B and B^{-1} though). In Appendix A, we give an alternative, shorter proof of this result based on a well-known characterization of bounded invertibility of linear operators between Hilbert spaces in terms of three conditions on the associated sesqui-linear form.

6. PARABOLIC PROBLEMS AS BI-INFINITE MATRIX VECTOR EQUATIONS

In order to construct Riesz bases for $\mathcal{X}$ and $\mathcal{Y}$, we use that

$$\mathcal{X} = (L_2(I) \otimes V) \cap (H^1(I) \otimes V') \text{ and } \mathcal{Y} = (L_2(I) \otimes V) \times H.$$

Let

$$\Sigma = \{\sigma_\mu : \mu \in \nabla_x\} \subset V$$

be a collection of functions that is a *normalized Riesz basis for H, that renormalized in V or V', is a Riesz basis for both these spaces.* Let

$$\Theta = \{\theta_\lambda : \lambda \in \nabla_t\} \subset H^1(I)$$

be a collection of functions that is a *normalized Riesz basis for $L_2(I)$, that renormalized in $H^1(I)$, is a Riesz basis for that space.* From [GO95, Prop. 1 and 2], it follows that then the collection $\Theta \otimes \Sigma$ normalized in $\mathcal{X}$, i.e., the collection

$$(t, x) \mapsto \left\{ \frac{\theta_\lambda(t) \sigma_\mu(x)}{\sqrt{\|\sigma_\mu\|_V^2 + \|\theta_\lambda\|_{H^1(I)}^2 \|\sigma_\mu\|_{V'}^2}} : (\lambda, \mu) \in \nabla_{\mathcal{X}} := \nabla_t \times \nabla_x \right\}$$

is a Riesz basis for $\mathcal{X}$, and that $(\Theta \otimes \Sigma) \times \Sigma$ normalized in $\mathcal{Y}$, i.e., the collection

$$(t, x) \mapsto \left\{ \left(\frac{\theta_\lambda(t) \sigma_\mu(x)}{\|\sigma_\mu\|_V}, 0 \right) : (\lambda, \mu) \in \nabla_t \times \nabla_x \right\} \cup \left\{ (0, \sigma_\mu(x)) : \mu \in \nabla_x \right\}$$

is a Riesz basis for $\mathcal{Y}$. Moreover, denoting the Riesz basis for V' consisting of the collection Σ normalized in V' as $[\Sigma]_{V'}$, and similarly for the other collections and spaces, with the notations introduced in Sect. 2 it holds that

$$(6.1) \qquad \Lambda^{\mathcal{X}}_{[\Theta \otimes \Sigma]_{\mathcal{X}}} \leq \max(\Lambda^{L_2(I)}_{\Theta} \Lambda^{V}_{[\Sigma]_V}, \Lambda^{H^1(I)}_{[\Theta]_{H^1(I)}} \Lambda^{V'}_{[\Sigma]_{V'}}),$$

$$(6.2) \qquad \lambda^{\mathcal{X}}_{[\Theta \otimes \Sigma]_{\mathcal{X}}} \geq \min(\lambda^{L_2(I)}_{\Theta} \lambda^{V}_{[\Sigma]_V}, \lambda^{H^1(I)}_{[\Theta]_{H^1(I)}} \lambda^{V'}_{[\Sigma]_{V'}}),$$

$$(6.3) \qquad \Lambda^{\mathcal{Y}}_{[(\Theta \otimes \Sigma) \times \Sigma]_{\mathcal{Y}}} \leq \max(\Lambda^{L_2(I)}_{\Theta} \Lambda^{V}_{[\Sigma]_V}, \Lambda^{H}_{\Sigma}),$$

$$(6.4) \qquad \lambda^{\mathcal{Y}}_{[(\Theta \otimes \Sigma) \times \Sigma]_{\mathcal{Y}}} \geq \min(\lambda^{L_2(I)}_{\Theta} \lambda^{V}_{[\Sigma]_V}, \lambda^{H}_{\Sigma}).$$

Denoting with $\|\Sigma\|_V$ the $\nabla_x \times \nabla_x$ diagonal matrix with diagonal entries $\|\sigma_\mu\|_V$ ($\mu \in \nabla_x$), and similarly for the other collections and spaces, the stiffness or system matrix $\mathbf{B}$ corresponding to the variational form (5.6) and the Riesz bases $[\Theta \otimes \Sigma]_{\mathcal{X}}$, $[(\Theta \otimes \Sigma) \times \Sigma]_{\mathcal{Y}}$ for $\mathcal{X}$ and $\mathcal{Y}$ is given by

$$\mathbf{B} = b([\Theta \otimes \Sigma]_{\mathcal{X}}, [(\Theta \otimes \Sigma) \times \Sigma]_{\mathcal{Y}}) = \begin{bmatrix} \mathrm{Id}_t \otimes \|\Sigma\|_V^{-1} & 0 \\ 0 & \mathrm{Id}_x \end{bmatrix} \circ$$

$$\begin{bmatrix} \langle \Theta', \Theta \rangle_{L_2(I)} \otimes \langle \Sigma, \Sigma \rangle_H + \int_I a(t, \Theta \otimes \Sigma, \Theta \otimes \Sigma) dt \\ \langle \Theta(0) \otimes \Sigma, \Sigma \rangle_H \end{bmatrix} \circ \|\Theta \otimes \Sigma\|_{\mathcal{X}}^{-1}.$$

Writing the solution u of (5.4) as $u = \mathbf{u}^\top [\Theta \otimes \Sigma]_{\mathcal{X}}$, we conclude that $\mathbf{u}$ is the solution of $\mathbf{Bu} = \mathbf{f}$ with

$$(6.5) \qquad \mathbf{f} = \begin{bmatrix} \int_I \langle g(t), \Theta \otimes [\Sigma]_V \rangle_H dt \\ \langle h, \Sigma \rangle_H \end{bmatrix}.$$

The matrix $\langle \Theta(0) \otimes \Sigma, \Sigma \rangle_{L_2(\Omega)} \|\Theta \otimes \Sigma\|_{\mathcal{X}}^{-1}$ can be written as $\langle \Sigma, \Sigma \rangle_H \mathbf{R}(\Theta, \Sigma)$, where $\mathbf{R} = \mathbf{R}(\Theta, \Sigma) \in \mathcal{L}(\ell_2(\nabla_t \times \nabla_x), \ell_2(\nabla_x))$ is given by

$$\mathbf{R}_{\mu, (\lambda, \nu)} = \begin{cases} \dfrac{\theta_\lambda(0)}{\sqrt{\|\sigma_\mu\|_V^2 + \|\theta_\lambda\|_{H^1(I)}^2 \|\sigma_\mu\|_{V'}^2}} & \text{when } \mu = \nu, \\ 0 & \text{otherwise.} \end{cases}$$

Introducing $\mathbf{D}_1 := (\|\Theta\|_{H^1(I)} \otimes \|\Sigma\|_{V'}) \|\Theta \otimes \Sigma\|_{\mathcal{X}}^{-1}$ and $\mathbf{D}_2 := (\mathrm{Id}_t \otimes \|\Sigma\|_V) \|\Theta \otimes \Sigma\|_{\mathcal{X}}^{-1}$, both being diagonal matrices with entries in modulus less than 1, $\mathbf{B}$ can be written as

$$(6.6)$$
$$\begin{bmatrix} \langle [\Theta]'_{H^1(I)}, \Theta \rangle_{L_2(I)} \otimes \langle [\Sigma]_{V'}, [\Sigma]_V \rangle_{L_2(\Omega)} \mathbf{D}_1 + \int_I a(t, \Theta \otimes [\Sigma]_V, \Theta \otimes [\Sigma]_V) dt \mathbf{D}_2 \\ \langle \Sigma, \Sigma \rangle_H \mathbf{R}(\Theta, \Sigma) \end{bmatrix}.$$

7. BEST POSSIBLE RATES IN $\mathcal{X}$

To solve the parabolic problem (5.4), we propose to apply the adaptive wavelet algorithms from [CDD02] or [CDD01]/[GHS07] to the normal equations $\mathbf{B}^* \mathbf{Bu} = \mathbf{B}^* \mathbf{f}$, where $\mathbf{B}$ is from (6.6) and $\mathbf{f}$ from (6.5). As we learned in Sect. 4, to conclude that these methods perform quasi-optimally, we have to show s^*-admissibility of both $\mathbf{B}$ and $\mathbf{B}^*$ for an s^* *larger than any s for which membership of* $\mathbf{u} \in \mathcal{A}^s_\infty(\ell_2(\nabla_{\mathcal{X}}))$ *can be expected*. In this section, we study which rates of best N-term approximation can at best be expected, under the provision that the solution u is sufficiently smooth.

For any $\Lambda \subset \nabla_t$, let $Q_{\Lambda,t} : L_2(I) \to \ell_2(\Lambda) : u \mapsto \sum_{\lambda \in \Lambda} u_\lambda \theta_\lambda$, where $\sum_{\lambda \in \nabla_t} u_\lambda \theta_\lambda$ is the unique expansion of u with respect to Θ. Since $u_\lambda = \langle u, \tilde{\theta}_\lambda \rangle_{L_2(I)}$ where $\tilde{\Theta} = \{\tilde{\theta}_\lambda : \lambda \in \nabla_t\}$ is the Riesz basis of $L_2(I)$ that is biorthogonal to Θ, we call $Q_{\Lambda,t}$ the $L_2(I)$-*biorthogonal projector* associated to Θ and index set $\Lambda \subset \nabla_t$.

We assume that Θ is a wavelet basis constructed using dyadic refinements which is of *order* $d_t > 1$. That is, with, for $k \in \mathbb{N}_0$, $\nabla_t^{(k)}$ being the set of $\lambda \in \nabla_t$ with *level* $\mathbb{N}_0 \ni |\lambda| \leq k$, it holds that $\#\nabla_t^{(k)} \approx 2^k$. Additionally setting $\nabla_t^{(-1)} := \emptyset$, for $Q_{k,t} := Q_{\nabla_t^{(k)},t}$ we have

$$\|\mathrm{Id} - Q_{k,t}\|_{H^{d_t}(I) \to L_2(I)} \lesssim 2^{-kd_t}, \quad \|\mathrm{Id} - Q_{k,t}\|_{H^{d_t}(I) \to H^1(I)} \lesssim 2^{-k(d_t-1)}.$$

For some $m > 0$, and some bounded domain $\Omega \subset \mathbb{R}^n$, let

$$V = H^m(\Omega)$$

or a closed subspace incorporating essential boundary conditions, and

$$H = L_2(\Omega).$$

For the wavelet basis Σ in space, we consider two cases: either

(A) it is a standard isotropic wavelet basis of order $d_x > m$ constructed from a dyadic multiresolution analysis in $L_2(\Omega)$, or

(B) $\Omega = \prod_{i=1}^n (a_i, b_i)$, and Σ is the tensor product of wavelet bases Σ_i as in (A) in each of the coordinate spaces.

For $n = 1$, (A) and (B) coincide, and in the following we will consider (A) only for $n > 1$.

Remark 7.1. The intermediate case that an n-dimensional domain Ω is the product of $1 < \ell < n$ domains, so at least one of them being higher dimensional, and Σ is the ℓ-fold tensor product of wavelet bases in the coordinate spaces can easily be analyzed along similar lines.

7.1. **Best rate in case** (A)**.** For any $\Lambda \subset \nabla_x$, let $Q_{\Lambda,x} : L_2(\Omega) \mapsto \ell_2(\Lambda)$ be the $L_2(\Omega)$-biorthogonal projector associated to Σ and Λ. The assumption of Σ being of order d_x means that with $\nabla_x^{(k)}$ being the set of $\lambda \in \nabla_x$ with level $|\lambda| \leq k \in \mathbb{N}_0$, it holds that $\#\nabla_x^{(k)} \approx 2^{kn}$. Additionally setting $\nabla_x^{(-1)} := \emptyset$, for $Q_{k,x} := Q_{\nabla_x^{(k)},x}$ we have

$$\|\mathrm{Id} - Q_{k,x}\|_{H^{d_x}(\Omega) \cap V \to V} \lesssim 2^{-k(d_x-m)}, \|\mathrm{Id} - Q_{k,x}\|_{H^{d_x}(\Omega) \cap V \to V'} \lesssim 2^{-k(d_x+m)}.$$

In case $d_t < \frac{d_x-m}{n}$, by taking, for some $\varepsilon > 0$, $\ell/k \in [\frac{d_t}{d_x-m} + \varepsilon, \frac{1}{n} - \varepsilon]$, we have

$$\left\|\mathrm{Id} - \sum_{p=0}^k \sum_{q=0}^\ell (Q_{p,t} - Q_{p-1,t}) \otimes (Q_{q,x} - Q_{q-1,x})\right\|_{H^{d_t}(I) \otimes (H^{d_x}(\Omega) \cap V) \to L_2(I) \otimes V} \lesssim 2^{-kd_t}.$$

The operator $\sum_{p=0}^k \sum_{q=0}^\ell (Q_{p,t} - Q_{p-1,t}) \otimes (Q_{q,x} - Q_{q-1,x})$ is the $L_2(\Omega)$-biorthogonal projector associated to the tensor product basis $\Theta \otimes \Sigma$ and "sparse" product index set $\cup_{p=0}^k \cup_{q=0}^\ell (\nabla_t^{(p)} \backslash \nabla_t^{(p-1)}) \times (\nabla_x^{(q)} \backslash \nabla_x^{(q-1)})$, which has cardinality of order 2^k (see [BG04] + references cited there).

Similarly, for $d_t > \frac{d_x-m}{n}$, by taking $k/\ell \in [\frac{d_x-m}{d_t} + \varepsilon, n - \varepsilon]$, we obtain an error estimate of order $2^{-\ell(d_x-m)}$ for the biorthogonal projector associated to a sparse product index set with cardinality of order $2^{\ell n}$.

Finally, when $d_t = \frac{d_x - m}{n}$, by taking $k = n\ell$, the error bound is of order $2^{-\ell(d_x-m)}$ for the biorthogonal projector associated to a sparse product index set with cardinality of order $\ell 2^{\ell n}$.

In summary, assuming $u \in H^{d_t}(I) \otimes (H^{d_x}(\Omega) \cap V)$, the error in $L_2(I) \otimes V$ of a biorthogonal projection associated to a sparse product index set with cardinality of order N is of order $N^{-\min(d_t, \frac{d_x-m}{n})}$, with an additional log factor when $d_t = (d_x - m)/n$.

Similarly, still assuming that $u \in H^{d_t}(I) \otimes (H^{d_x}(\Omega) \cap V)$, with a suitable choice of sparse product index sets with cardinality of order N, the error measured in $H^1(I) \otimes V'$ in the corresponding biorthogonal projection is of order $N^{-\min(d_t-1, \frac{d_x+m}{n})}$, with an additional log factor when $d_t = (d_x - m)/n$.

By taking the union of the sparse product index sets for approximation in $L_2(I) \otimes V$ and in $H^1(I) \otimes V'$, we obtain biorthogonal projectors that, for $u \in H^{d_t}(I) \otimes (H^{d_x}(\Omega) \cap V)$, with an index set of order N give rise to an error in $\mathcal{X}$ of order $N^{-\min(d_t-1, \frac{d_x-m}{n})}$. Here we use that, after normalization, $\Theta \otimes \Sigma$ is a Riesz basis for $L_2(I) \otimes V$ and $H^1(I) \otimes V'$, meaning that when the index set is extended, the error in a biorthogonal projection can increase with at most a constant factor. Moreover, we used that if $d_t = \frac{d_x-m}{n}$ or $d_t - 1 = \frac{d_x+m}{n}$, then $\min(d_t-1, d_t, \frac{d_x-m}{n}, \frac{d_x+m}{n}) < d_t$ or $\min(d_t - 1, d_t, \frac{d_x-m}{n}, \frac{d_x+m}{n}) < d_t - 1$, respectively, so that the aforementioned log factors are irrelevant. In view of the approximation orders of the bases being applied, and the definition of $\mathcal{X}$, the derived rate

$$\min(d_t - 1, \tfrac{d_x-m}{n})$$

is the best that for general smooth functions can be expected.

The division by n in the second argument is known as the *curse of dimensionality*. In the next subsection, we will see that it can be circumvented by applying tensor product wavelets when the spatial domain is a Cartesian product of lower dimensional domains.

7.2. **Best rate in case** (B). For $1 \leq i \leq n$, let V_i be either $H^m(a_i, b_i)$ or a closed subspace incorporating essential boundary conditions. Let $\Sigma_i = \{\sigma_{i,\lambda_i} : \lambda_i \in \nabla_i\}$ be a normalized Riesz basis for $L_2(a_i, b_i)$, that renormalized in V_i or V_i' is a Riesz basis for both these spaces. For any $\Lambda_i \subset \nabla_i$, let $Q_{\Lambda_i, i} : L_2(a_i, b_i) \to \ell_2(\Lambda_i)$ be the $L_2(a_i, b_i)$-biorthogonal projectors associated to Σ_i and Λ_i. The assumption of Σ_i being of order d_x means that with $\nabla_i^{(k)}$ being the set of $\lambda_i \in \nabla_i$ with level $|\lambda_i| \leq k \in \mathbb{N}_0$, it holds that $\#\nabla_i^{(k)} \eqsim 2^k$. Additionally setting $\nabla_i^{(-1)} := \emptyset$, for $Q_{k,i} := Q_{\nabla_i^{(k)}, i}$ we have

$$\|\mathrm{Id} - Q_{k,i}\|_{H^{d_x}(a_i,b_i) \cap V_i \to V_i} \lesssim 2^{-k(d_x-m)}, \|\mathrm{Id} - Q_{k,i}\|_{H^{d_x}(a_i,b_i) \cap V_i \to V_i'} \lesssim 2^{-k(d_x+m)}.$$

The collection $\Sigma := \otimes_{i=1}^n \Sigma_i = \{\sigma_\lambda := \otimes_{i=1}^n \sigma_{i,\lambda_i} : \lambda \in \nabla_x := \prod_{i=1}^n \nabla_i\}$ is a normalized Riesz basis for $L_2(\Omega)$, that renormalized in

$$V := \cap_{i=1}^n \otimes_{j=1}^n W_{ij}, \text{ where } W_{ij} := \begin{cases} L_2(a_i, b_i) & \text{when } j \neq i, \\ V_i & \text{when } j = i, \end{cases}$$

or in V', is a Riesz basis for both these spaces.

Recall that for any $\Lambda \subset \nabla_x$, $Q_{\Lambda,x}$ denotes the $L_2(\Omega)$-orthogonal projector associated to Σ and Λ. As shown in [GK00] (cf. also [DSS08] for approximation in V'), there exist "optimized" sparse product sets $\nabla_x^{(1)} \subset \nabla_x^{(1)} \subset \cdots \subset \nabla_x$ and

$\hat{\nabla}_x^{(1)} \subset \hat{\nabla}_x^{(1)} \subset \cdots \subset \nabla_x$ with $\#\nabla_x^{(1)} \eqsim 2^k \eqsim \hat{\nabla}_x^{(k)}$, such that with $Q_{k,x} := Q_{\nabla_x^{(k)},x}$ and $\hat{Q}_{k,x} := Q_{\hat{\nabla}_x^{(k)},x}$, and

$$\mathcal{H}^{d_x}(\Omega) := \cap_{i=1}^n \otimes_{j=1}^n Z_{ij}, \text{ where } Z_{ij} := \begin{cases} L_2(a_i, b_i) & \text{when } j \neq i, \\ H^{d_x}(a_i, b_i) \cap V_i & \text{when } j = i, \end{cases}$$

it holds that

$$\|\mathrm{Id} - Q_{k,x}\|_{\mathcal{H}^{d_x}(\Omega) \to V} \lesssim 2^{-k(d_x - m)}, \|\mathrm{Id} - \hat{Q}_{k,x}\|_{\mathcal{H}^{d_x}(\Omega) \to V'} \lesssim 2^{-k(d_x + \frac{m}{n})}.$$

Now similarly to case (A), by taking the union of sparse products of the index sets $(\nabla_t^{(p)})_{0 \leq p \leq k}$ with $(\nabla_x^{(q)})_{0 \leq q \leq \ell}$ or $(\hat{\nabla}_x^{(q)})_{0 \leq q \leq \ell}$ for suitable k and ℓ, we obtain $L_2(I \times \Omega)$-biorthogonal projectors associated to $\Theta \otimes \Sigma$ that, for $u \in H^{d_t}(I) \otimes \mathcal{H}^{d_x}(\Omega)$, with an index set of order N give rise to an error in $\mathcal{X}$ of order $N^{-\min(d_t - 1, d_x - m)}$. Again, this rate

$$\min(d_t - 1, d_x - m)$$

is the best that for general smooth functions can be expected.

In view of the obtained results, we conclude that it suffices to show s^*-admissibility of both $\mathbf{B}$ and $\mathbf{B}^*$ for an

$$(7.1) \qquad s^* > \begin{cases} \min(d_t - 1, \frac{d_x - m}{n}) & \text{in case (A)}, \\ \min(d_t - 1, d_x - m) & \text{in case (B)}. \end{cases}$$

8. s^*-ADMISSIBILITY OF $\mathbf{B}$ FROM (6.6) AND OF ITS ADJOINT

For handling tensor products of matrices, we use the following proposition. Part (b) will be used in Sect. 9 where we analyze the dependence of constants on the space dimension.

Proposition 8.1. *Let for some $s^* > 0$, $\mathbf{D}$, $\mathbf{E}$ be s^*-computable. Then*
  (a) *$\mathbf{D} \otimes \mathbf{E}$ is s^*-computable with computability constant satisfying, for $0 < \bar{s} < \tilde{s} < s^*$, $c_{\mathbf{D} \otimes \mathbf{E}, \bar{s}} \lesssim (c_{\mathbf{D},\tilde{s}} c_{\mathbf{E},\tilde{s}})^{\tilde{s}/\bar{s}}$ and*
  (b) *for any $\varepsilon \in (0, s^*)$, $\mathbf{D} \otimes \mathbf{E}$ is $(s^* - \varepsilon)$–computable, with computability constant $c_{\mathbf{D} \otimes \mathbf{E}, \bar{s}}$ satisfying, for $0 < \bar{s} < s^* - \varepsilon < \tilde{s} < s^*$, $c_{\mathbf{D} \otimes \mathbf{E}, \bar{s}} \lesssim \max(c_{\mathbf{D},\tilde{s}}, 1) \max(c_{\mathbf{E},\tilde{s}}, 1)$.*

The constants absorbed in the $\lesssim$ symbol in the bounds on the computability constants in (a) and (b) are only dependent on $\tilde{s} \downarrow 0$, $\tilde{s} \to \infty$ and $\tilde{s} - \bar{s} \downarrow 0$.

Proof. (a). For $N \in \mathbb{N}$, let $\mathbf{D}_N$ and $\mathbf{E}_N$ be approximations to $\mathbf{D}$ and $\mathbf{E}$ as in the definition of s^*-computability. For $\ell \in \mathbb{N}_0$, set $\mathbf{D}_{[\ell]} = \mathbf{D}_{2^\ell}$ and $\mathbf{D}_{[-1]} := 0$, and similarly $\mathbf{E}_{[\ell]}$. Then for any $0 < \bar{s} < \tilde{s} < s^*$ and $q \in \mathbb{N}_0$,

$$\|\mathbf{D} \otimes \mathbf{E} - \sum_{\ell + m \leq q} (\mathbf{D}_{[\ell]} - \mathbf{D}_{[\ell-1]}) \otimes (\mathbf{E}_{[m]} - \mathbf{E}_{[m-1]})\| =$$

$$\|\sum_{\ell + m > q} (\mathbf{D}_{[\ell]} - \mathbf{D}_{[\ell-1]}) \otimes (\mathbf{E}_{[m]} - \mathbf{E}_{[m-1]})\| \lesssim \sum_{\ell + m > q} (c_{\mathbf{D},\tilde{s}} c_{\mathbf{E},\tilde{s}})^{\tilde{s}} 2^{-(\ell + m)\tilde{s}}$$

$$\lesssim (c_{\mathbf{D},\tilde{s}} c_{\mathbf{E},\tilde{s}})^{\tilde{s}} (q + 1) 2^{-q\tilde{s}} \lesssim (c_{\mathbf{D},\tilde{s}} c_{\mathbf{E},\tilde{s}})^{\tilde{s}} ((q + 1) 2^q)^{-\bar{s}},$$

with constants only dependent on $\tilde{s} \downarrow 0$, $\tilde{s} \to \infty$ and $\tilde{s} - \bar{s} \downarrow 0$. The number of non-zero entries in each column of $\sum_{\ell + m \leq q} (\mathbf{D}_{[\ell]} - \mathbf{D}_{[\ell-1]}) \otimes (\mathbf{E}_{[m]} - \mathbf{E}_{[m-1]})$, as well as the number of operations needed for computing them, is less than some absolute multiple of $\sum_{\ell + m \leq q} 2^{\ell + m} \eqsim (q + 1) 2^q$.

(b). For some fixed $\tilde{s} \in (s^* - \varepsilon, s^*)$, let $\mathbf{D}_{[\ell]} = \mathbf{D}_{\lceil 2^\ell c_{\mathbf{D},\tilde{s}} \rceil}$, $\mathbf{E}_{[\ell]} = \mathbf{E}_{\lceil 2^\ell c_{\mathbf{E},\tilde{s}} \rceil}$. Then

$$\left\| \mathbf{D} \otimes \mathbf{E} - \sum_{\ell+m \leq q} (\mathbf{D}_{[\ell]} - \mathbf{D}_{[\ell-1]}) \otimes (\mathbf{E}_{[m]} - \mathbf{E}_{[m-1]}) \right\| \lesssim ((q+1)2^q)^{-\bar{s}},$$

only dependent on $\tilde{s} \downarrow 0$, $\tilde{s} \to \infty$ and $\tilde{s} - \bar{s} \downarrow 0$. The number of non-zero entries in each column of $\sum_{\ell+m \leq q} (\mathbf{D}_{[\ell]} - \mathbf{D}_{[\ell-1]}) \otimes (\mathbf{E}_{[m]} - \mathbf{E}_{[m-1]})$, as well as the number of operations needed for computing them, is less than some absolute multiple of

$$\sum_{\ell+m \leq q} \lceil 2^\ell c_{\mathbf{D},\tilde{s}} \rceil \lceil 2^m c_{\mathbf{E},\tilde{s}} \rceil \lesssim \max(c_{\mathbf{D},\tilde{s}}, 1) \max(c_{\mathbf{E},\tilde{s}}, 1)(q+1)2^q.$$

$\square$

In view of the representation (6.6) of $\mathbf{B}$, using Corollary 4.6 and Proposition 8.1(a), for proving s^*-admissibility of $\mathbf{B}$ and $\mathbf{B}^*$ it suffices to show that $\mathbf{R}(\Theta, \Sigma)$, $\mathbf{R}(\Theta, \Sigma)^*$, $\langle \Sigma, \Sigma \rangle_{L_2(\Omega)}$ and $\int_I a(t, \Theta \otimes [\Sigma]_V, \Theta \otimes [\Sigma]_V) dt$ and its adjoint are s^*-admissible, e.g., s^*-computable, and that $\langle [\Theta]'_{H^1(I)}, \Theta \rangle_{L_2(I)}$, $\langle \Theta, [\Theta]'_{H^1(I)} \rangle_{L_2(I)}$, $\langle [\Sigma]_{V'}, [\Sigma]_V \rangle_{L_2(\Omega)}$ and $\langle [\Sigma]_V, [\Sigma]_{V'} \rangle_{L_2(\Omega)}$ are s^*-computable. This will require quite some technicalities. The conclusions, however, are summarized in the forthcoming Theorem 8.8.

We need some assumptions on the wavelet bases Θ and Σ that are collected in the following subsection. These assumptions can be met by available wavelet constructions. Concerning Θ and Σ in case (B), one can use biorthogonal ([DKU99, Pri06]), semi-orthogonal ([CQ92]) or orthogonal (piecewise polynomial) ([DGH96, DGH99, Goo00]) wavelets on the interval. Case (A) seems only relevant when Ω is not a product domain. Then to construct wavelet bases, one can resort to domain decomposition techniques. At least for differential operators in the spatial direction, for obtaining a sufficiently compressible system matrix it will turn out that in this case globally C^1 wavelets are required. The only construction known to us is that from [DS99] further investigated in [KS06].

Remark 8.2. In [HS06], for differential operators of order 2, sufficient compressibility was obtained by a more easy realizable continuous "gluing" of wavelets over the subdomain interfaces, which however have the stronger patchwise cancellation properties. Unfortunately, this approach cannot be applied here since it does not yield wavelets that are a basis for V'.

8.1. **Wavelet assumptions.** We already assumed that $\Theta = \{\theta_\lambda : \lambda \in \nabla_t\}$ is a normalized Riesz basis of $L_2(I)$, that renormalized in $H^1(I)$ is a Riesz basis of that space, and that Θ is of order $d_t > 1$. In addition to that, we will assume that the θ_λ are

(t1) *local*, i.e., $\sup_{x \in [0,1], \ell \in \mathbb{N}_0} \#\{|\lambda| = \ell : x \in \operatorname{supp} \theta_\lambda\} < \infty$ and $\operatorname{supp} \theta_\lambda \lesssim 2^{-|\lambda|}$,

(t2) *piecewise polynomial of order d_t*, where with "piecewise" is meant that the singular support consists of a uniformly bounded number of points,

(t3) *globally continuous*, specifically $\|\theta_\lambda\|_{W^k_\infty(0,1)} \lesssim 2^{|\lambda|(\frac{1}{2}+k)}$ ($k \in \{0, 1\}$),

(t4) that, for $|\lambda| > 0$, have $\tilde{d}_t \geq d_t$ *vanishing moments*.

Concerning the wavelets in space, we already assumed that $\Sigma = \{\sigma_\lambda : \lambda \in \nabla_x\}$ is a normalized Riesz basis of $L_2(\Omega)$, that renormalized in V or V' is a Riesz basis for both these spaces, and that Σ, or in case (B) each of the factors Σ_i, is of order $d_x > m$.

In addition to that, in **case (A)** for some $\mathbb{N}_0 \ni r_x \geq m - 1$ (necessarily with $r_x \leq d_x - 2$) and $\tilde{d}_x \in \mathbb{N}_0$, we will assume that the σ_λ are

(s1) *local* and *piecewise smooth*, i.e., that for any $\ell \in \mathbb{N}_0$ there exist collections $\{\Omega_{\ell,v} : v \in \mathcal{O}_\ell\}$ of disjoint, uniformly shape regular, open subdomains such that $\overline{\Omega} = \cup_{v \in \mathcal{O}_\ell} \overline{\Omega_{\ell,v}}$, $\overline{\Omega_{\ell,v}}$ is the union of some $\overline{\Omega_{\ell+1,\tilde{v}}}$, $\mathrm{diam}(\Omega_{\ell,v}) \eqsim 2^{-\ell}$, $\mathrm{supp}\,\sigma_\lambda$ is connected and is the union of a uniformly bounded number of $\overline{\Omega_{|\lambda|,v}}$, each $\overline{\Omega_{\ell,v}}$ has non-empty intersection with the supports of a uniformly bounded number of σ_λ with $|\lambda| = \ell$, and for some sufficiently large K, for $k \in \{0, K\}$,

$$\|\sigma_\lambda\|_{W_\infty^k(\Omega_{|\lambda|,v})} \lesssim 2^{|\lambda|(\frac{n}{2}+k)},$$

(s2) *globally C^{r_x}*, specifically $\|\sigma_\lambda\|_{W_\infty^k(\Omega)} \lesssim 2^{|\lambda|(\frac{n}{2}+k)}$ ($k \in \{0, r_x + 1\}$),

(s3) that, for $|\lambda| > 0$, have *cancellation properties of order $\tilde{d}_x$*, i.e.,

$$(8.1) \qquad |\int_\Omega w\sigma_\lambda| \lesssim 2^{-|\lambda|(\frac{n}{2}+k)}\|w\|_{W_\infty^k(\Omega)} \ (k \in \{0, \tilde{d}_x\},\ w \in W_\infty^k(\Omega) \cap V).$$

(s4) For handling the quadrature issue, in addition to (s1), we assume that for any ℓ and $v \in \mathcal{O}_\ell$, there exists a sufficiently smooth transformation of coordinates κ, with derivatives bounded uniformly in ℓ and v, such that for all $|\lambda| = \ell$, $(\sigma_\lambda \circ \kappa)|_{\kappa^{-1}(\Omega_{\ell,v})}$ is a polynomial of some fixed degree.

Remark 8.3. The intersection of $W_\infty^k(\Omega)$ with V in (8.1) is only meaningful when the definition of V incorporates homogeneous Dirichlet boundary conditions. Since we need that, renormalized, Σ is also a basis for V', in that case (8.1) cannot be expected to hold for any $w \in W_\infty^k(\Omega)$. Indeed think of w as being a polynomial of order $\tilde{d}_x$ not in V. By writing it in terms of the dual basis $\tilde{\Sigma}$, i.e., the collection with $\langle \tilde{\Sigma}, \Sigma \rangle_{L_2(\Omega)} = \mathrm{Id}$, and by testing the expansion with primal wavelets, we find that all coefficients for $|\lambda| > 0$ are zero, and thus that w is a (finite) linear combination of dual wavelets. We conclude that $\tilde{\Sigma} \not\subset V$, meaning that Σ cannot be a basis for V'. So in our setting that, properly scaled, Σ is a basis for both V and V', wavelets with so called "complementary boundary conditions" ([DS98]) cannot be applied. As we will see in Sect. 8.5, for differential operators in the spatial direction, this will impose a restriction on the order d_x of the wavelet basis Σ. Another restriction caused by this fact was already mentioned in Remark 8.2.

For **case** (B), we assume that each of the Σ_i satisfies the above conditions with (Ω, n) reading as $((a_i, b_i), 1)$. In this case, it is natural simply to assume that the wavelets are *piecewise polynomials of order d_x*, with those on positive levels being orthogonal to all polynomials of order $\tilde{d}_x$ that are in V. Note that the latter property together with (s2) for $s = 0$ implies (s3).

8.2. s^*-computability of $\langle [\Theta]'_{H^1(I)}, \Theta \rangle_{L_2(I)}$ and its adjoint. By (t1), (t2) and (t4), for each $\lambda \in \nabla_t$ and $\ell \in \mathbb{N}_0$, the number of $\mu \in \nabla_t$ with $|\mu| = \ell$ and $\int_I \theta'_\lambda \theta_\mu \neq 0$ or $\int_I \theta'_\mu \theta_\lambda \neq 0$ is bounded, uniformly in λ and ℓ. Indeed, $\int_I \theta'_\lambda \theta_\mu$ can only be non-zero when θ_μ does not vanish on the singular support of θ_λ, and using integration by parts, $\int_I \theta'_\mu \theta_\lambda$ can only be non-zero when θ_μ does not vanish on the singular support of θ_λ or at ∂I.

As a consequence of Θ being of order $d_t \geq 1$ we have $2^{|\lambda|} = 2^{|\lambda|}\|\theta_\lambda\|_{L_2(I)} = 2^{|\lambda|}\|(\mathrm{Id} - Q_{|\lambda|-1})\theta_\lambda\|_{L_2(I)} \lesssim \|\theta_\lambda\|_{H^1(I)}$. Using (t1) and (t3), we infer that

$$(8.2) \qquad \|\theta_\lambda\|_{H^1(I)}^{-1}|\int_I \theta'_\lambda \theta_\mu| \lesssim 2^{-|\lambda|}2^{-\max(|\lambda|,|\mu|)}2^{|\lambda|(\frac{1}{2}+1)}2^{|\mu|\frac{1}{2}} = 2^{-\left||\lambda|-|\mu|\right|\frac{1}{2}}.$$

Finally, we note that any entry of $\langle [\Theta]'_{H^1(I)}, \Theta \rangle_{L_2(I)}$ can be computed exactly in $\mathcal{O}(1)$ operations.

Standard arguments using the Schur lemma now show that $\langle [\Theta]'_{H^1(I)}, \Theta \rangle_{L_2(I)}$ and $\langle \Theta, [\Theta]'_{H^1(I)} \rangle_{L_2(I)}$ are ∞-computable.

Although not needed here, at this point we note that from $\operatorname{diam}(\operatorname{supp} \theta_\lambda) \lesssim 2^{-|\lambda|}$ ((t1)) and (t3), it follows that $\|\theta_\lambda\|_{H^1(I)} \lesssim 2^{|\lambda|}$, and thus that

$$\|\theta_\lambda\|_{H^1(I)} \eqsim 2^{|\lambda|}.$$

Remark 8.4. Suppose that instead of (t3), the θ_λ are even in $C^{r_t}(I)$ for some $r_t \in \mathbb{N}$ (necessarily with $r_t \leq d_t - 2$), i.e., that $\|\theta_\lambda\|_{W^s_\infty(0,1)} \lesssim 2^{|\lambda|(\frac{1}{2}+s)}$ for $s \in \{0, r_t+1\}$. Then by subtracting a suitable polynomial of order r_t from θ'_λ in (8.2), and using that θ_μ has $\tilde{d}_t \geq d_t \geq r_t$ vanishing moments one shows that for $|\lambda| \leq |\mu|$, $\|\theta_\lambda\|^{-1}_{H^1(I)} |\int_I \theta'_\lambda \theta_\mu| \lesssim 2^{-\left||\lambda|-|\mu|\right|(\frac{1}{2}+r_t)}$. Similarly, for $|\lambda| \geq |\mu|$, using integration by parts one obtains $\|\theta_\lambda\|^{-1}_{H^1(I)} |\int_I \theta'_\lambda \theta_\mu| \lesssim 2^{-\left||\lambda|-|\mu|\right|(\frac{3}{2}+r_t)}$ *if* $t \mapsto \theta_\lambda(t)\theta_\mu(t)$ *vanishes at* ∂I. Since the wavelets in time do not satisfy Dirichlet boundary conditions, there are, however, $\lambda, \mu \in \nabla_t$ with $|\lambda| \geq |\mu|$ for which $t \mapsto \theta_\lambda(t)\theta_\mu(t)$ does not vanish at the boundary. For those entries (8.2) cannot be improved.

8.3. s^*-computability of $\langle \Sigma, \Sigma \rangle_{L_2(\Omega)}$ and $\langle [\Sigma]_{V'}, [\Sigma]_V \rangle_{L_2(\Omega)}$ and its adjoint.

We start with **case** (A). As shown in [SS08, Lemma 3.1], from (s1) and (s2), for $|\mu| \geq |\lambda|$ we have

$$(8.3) \qquad |\langle \sigma_\lambda, \sigma_\mu \rangle_{L_2(\Omega)}| \lesssim 2^{(|\lambda|-|\mu|)(\frac{n}{2}+r_x+1)},$$

or even

$$(8.4) \qquad |\langle \sigma_\lambda, \sigma_\mu \rangle_{L_2(\Omega)}| \lesssim 2^{(|\lambda|-|\mu|)(\frac{n}{2}+\tilde{d}_x)},$$

with (8.4) being valid when σ_μ vanishes on $\operatorname{sing\,supp} \sigma_\lambda$ [in [SS08] it was assumed that the wavelets satisfy homogeneous Dirichlet boundary conditions, and that (8.1) is valid for all $w \in W^k_\infty(\Omega)$ and not only for $w \in W^k_\infty(\Omega) \cap V$. Both assumptions, however, were not used for the estimates in this case]. Now using that by the locality and the piecewise smoothness of the wavelets, for any λ and $\ell \geq |\lambda|$, the number of $|\mu| = \ell$ for which (8.4) does not hold is $\mathcal{O}(2^{(\ell-|\lambda|)(n-1)})$, it was shown that $\langle \Sigma, \Sigma \rangle_{L_2(\Omega)}$ is s^*-computable with

$$s^* = \min(\tfrac{\tilde{d}_x}{n}, \tfrac{r_x+\frac{3}{2}}{n-1}).$$

(Apply [SS08, Thms 4.1, 6.2] with "n"= 1; for the latter also (s4) is used.) We will refer to entries for which (8.4) holds as *regular entries*, and to the remaining entries, for which thus only (8.3) is available, as *singular entries*.

From (s2) and $\operatorname{diam}(\operatorname{supp} \sigma_\lambda) \lesssim 2^{-|\lambda|}$ ((s1)), we have $\|\sigma_\lambda\|_V \lesssim 2^{|\lambda|m}$. From Σ being of order $d_x \geq m$, we have $2^{|\lambda|m} \lesssim \|\sigma_\lambda\|_V$, so that

$$(8.5) \qquad \|\sigma_\lambda\|_V \eqsim 2^{|\lambda|m} \quad \text{and} \quad \|\sigma_\lambda\|_{V'} \eqsim 2^{-|\lambda|m}.$$

To see the latter, we use that Σ is a normalized Riesz basis for $L_2(\Omega)$, that renormalized in V or V', is a Riesz basis for both these spaces. The dual collection $\tilde{\Sigma}$ has therefore the same properties. From that, one easily derives that

$\|\sigma_\lambda\|_{V'} \eqsim 1/\|\tilde\sigma_\lambda\|_{V'}$ and similarly $\|\tilde\sigma_\lambda\|_{V'} \eqsim 1/\|\sigma_\lambda\|_{V'}$, so that $1 = \|\sigma_\lambda\|_{L_2(\Omega)}^2 \leq \|\sigma_\lambda\|_V \|\sigma_\lambda\|_{V'} \eqsim (\|\tilde\sigma_\lambda\|_V \|\tilde\sigma_\lambda\|_{V'})^{-1} \leq \|\tilde\sigma_\lambda\|_{L_2(\Omega)}^{-2} = 1$, i.e.,

$$(8.6) \qquad\qquad\qquad \|\sigma_\lambda\|_V \eqsim 1/\|\sigma_\lambda\|_{V'}.$$

From (8.3), (8.5) and (8.5), we conclude that

$$|[\langle[\Sigma]_{V'}, \Sigma\rangle_V)_{L_2(\Omega)}]_{\lambda,\mu}| = \frac{|\langle\sigma_\mu, \sigma_\lambda\rangle_{L_2(\Omega)}|}{\|\sigma_\mu\|_{V'}\|\sigma_\lambda\|_V} \lesssim 2^{-\left||\lambda|-|\mu|\right|(\frac{n}{2}+r_x+1)} 2^{|\lambda|m-|\mu|m}$$

$$\leq 2^{-\left||\lambda|-|\mu|\right|(\frac{n}{2}+r_x+1-m)}$$

or

$$|[\langle[\Sigma]_{V'}, \Sigma\rangle_V)_{L_2(\Omega)}]_{\lambda,\mu}| \lesssim 2^{-\left||\lambda|-|\mu|\right|(\frac{n}{2}+\tilde d_x-m)},$$

for the singular or regular entries, respectively. The analysis from [SS08] now shows that $\langle[\Sigma]_{V'}, \Sigma\rangle_V)_{L_2(\Omega)}$ is s^*-computable with

$$s^* = \min(\frac{\tilde d_x - m}{n}, \frac{r_x + \frac{3}{2} - m}{n-1}).$$

In view of the requirement (7.1) on s^*, we conclude that for case (A) we need

$$(8.7) \qquad\qquad \tilde d_x > d_x \text{ and } \frac{r_x + \frac{3}{2} - m}{n-1} > \frac{d_x - m}{n},$$

with the latter reading as $\frac{d_x - m}{n} > \frac{1}{2}$ for spline wavelets ($r_x = d_x - 2$).

In **case** (B), for each of the Σ_i, we have

$$|\langle\sigma_{i,\lambda_i}, \sigma_{i,\mu_i}\rangle_{L_2(a_i,b_i)}| \lesssim 2^{-\left||\lambda_i|-|\mu_i|\right|(\frac{1}{2}+r_x+1)},$$

with the regular entries now being zero. That is, for each $\lambda_i \in \nabla_i$ and $\ell \in \mathbb{N}_0$, the number of $\mu_i \in \nabla_i$ with $|\mu_i| = \ell$ and $\langle\sigma_{i,\lambda_i}, \sigma_{i,\mu_i}\rangle_{L_2(a_i,b_i)} \neq 0$ is bounded uniformly in λ_i and ℓ. Moreover, any entry can be computed exactly at unit cost. As in Sect. 8.2, we conclude that $\langle\Sigma_i, \Sigma_i\rangle_{L_2(a_i,b_i)}$ is ∞-computable, and so is $\langle\Sigma, \Sigma\rangle_{L_2(\Omega)} = \otimes_{i=1}^n \langle\Sigma_i, \Sigma_i\rangle_{L_2(a_i,b_i)}$.

As in (8.5), we have $\|\sigma_{i,\lambda_i}\|_{V_i} \eqsim 2^{|\lambda_i|m}$, and so, cf. (8.6), $\|\sigma_\lambda\|_V^2 \eqsim \|\sigma_\lambda\|_{V'}^{-2} \eqsim \sum_{i=1}^n 4^{|\lambda_i|m}$. From $(\sum_{i=1}^n 4^{|\mu_i|m} / \sum_{i=1}^n 4^{|\lambda_i|m})^{\frac{1}{2}} \leq \prod_{i=1}^n 2^{\left||\lambda_i|-|\mu_i|\right|m}$, and the fact that from $r_x \geq m - 1$, it follows that $[2^{\left||\lambda_i|-|\mu_i|\right|m}\langle\sigma_{i,\lambda_i}, \sigma_{i,\mu_i}\rangle_{L_2(a_i,b_i)}]_{\lambda_i,\mu_i\in\nabla_i}$ is ∞-computable, we derive that $\langle[\Sigma]_{V'}, [\Sigma]_V\rangle_{L_2(\Omega)}$ and its adjoint are ∞-computable.

8.4. s^*-computability of $\mathbf{R}(\Theta, \Sigma)$ and its adjoint.

Each column of $\mathbf{R}$ contains at most 1 non-zero entry, so its application to any finitely supported vector can be performed *exactly*, taking a number of operations proportional to the length of this vector.

Concerning $\mathbf{R}(\Theta, \Sigma)^*$, from (t1) we have

$$\sup_{\ell\in\mathbb{N}_0} \#\{\lambda \in \nabla_t : |\lambda| = \ell, \theta_\lambda(0) \neq 0\} < \infty.$$

The trace theorem shows that

$$|\theta_\lambda(0)| \lesssim \sqrt{\|\theta_\lambda\|_{L_2(I)}\|\theta_\lambda\|_{H^1(I)}} = \sqrt{\|\theta_\lambda\|_{H^1(I)}}, \quad (\lambda \in \nabla_t).$$

Knowing that $\|\theta_\lambda\|_{H^1(I)} \eqsim 2^{|\lambda|}$ and $\|\sigma_\mu\|_V \eqsim \|\sigma_\mu\|_{V'}^{-1}$, we infer that

$$|\mathbf{R}_{\mu,(\lambda,\mu)}| \lesssim \frac{1}{\sqrt{2^{-|\lambda|}\|\sigma_\mu\|_V^2 + 2^{|\lambda|}\|\sigma_\mu\|_{V'}^{-2}}}.$$

By defining $\mathbf{R}_N$ by dropping all entries $\mathbf{R}_{\mu,(\lambda,\nu)}$ with $\big||\lambda| - 2\log_2(\|\sigma_\mu\|_V)\big| > N$ from *row* $\mathbf{R}_{\mu,\cdot}$, the number of non-zero entries per row of $\mathbf{R}_N$ is $\mathcal{O}(N)$, independent of the row or even of the basis Σ, where an application of the Schur lemma shows that $\|\mathbf{R} - \mathbf{R}_N\| \lesssim 2^{-N/2}$. We conclude that $\mathbf{R}(\Theta, \Sigma)^*$ is ∞-computable.

Remark 8.5. The fact that $\mathbf{R}_{\mu,(\lambda,\mu)}$ is not small if and only if $|\lambda| \approx 2\log_2(\|\sigma_\mu\|_V)$ ($\lambda \in \nabla_t$, $\mu \in \nabla_x$) shows how frequencies in the initial data $u(0)$ will transported into the interior of the space-time cylinder via the application of $\mathbf{B}^*$. Recall that $2\log_2(\|\sigma_\mu\|_V)$ is $\eqsim 2m|\mu|$ in case (A), and $\eqsim \log_2(\sum_i 2^{2m|\mu_i|})$ in case (B).

8.5. s^*-computability of $\int_I a(t, \Theta \otimes [\Sigma]_V, \Theta \otimes [\Sigma]_V)dt$ and its adjoint. Let us first consider the situation that

$$a(t, \cdot, \cdot) = a_1(t)a_2(\cdot, \cdot)$$

with $a_1 > 0$, a_1 and $1/a_1$ bounded on I, and a_2 independent of t, bounded on V ((1.1)), and coercive on V with respect to $L_2(\Omega)$ ((1.2)). We have

$$\int_I a(t, \Theta \otimes [\Sigma]_V, \Theta \otimes [\Sigma]_V)dt = \langle a_1(\cdot)\Theta, \Theta\rangle_{L_2(I)} \otimes a_2([\Sigma]_V, [\Sigma]_V),$$

and it suffices to investigate s^*-computability of both factors, and that of the adjoint of $a_2([\Sigma]_V, [\Sigma]_V)$ in case it is unsymmetric.

For a_1 being a constant, $\langle a_1\Theta, \Theta\rangle_{L_2(I)}$ is ∞-computable (cf. discussion on the computability of $\langle \Sigma, \Sigma\rangle_{L_2(\Omega)}$ in case (B)). For a non-constant, but smooth a_1, [SS08, Thms 4.1, 6.2] shows that $\langle a_1(\cdot)\Theta, \Theta\rangle_{L_2(I)}$ is s^*-computable for $s^* = \tilde{d}_t$. Since in (t4) it was assumed that $\tilde{d}_t \geq d_t$, this value of s^* suffices in view of (7.1).

Computability of $a_2([\Sigma]_V, [\Sigma]_V)$ depends on the corresponding spatial operator at hand. If, for some $m \in \mathbb{N}$, $V = H_0^m(\Omega)$, and a_2 corresponds to a *differential operator* of order $2m$ in variational form with sufficiently smooth coefficients, then from [SS08] it can be deduced that both $a_2([\Sigma]_V, [\Sigma]_V)$ and its adjoint are s^*-computable with

$$s^* = \begin{cases} \min(\frac{\tilde{d}_x}{n}, \frac{\min(r_x+1,2m)+\frac{1}{2}-m}{n-1}) & \text{in case (A)}, \\ \tilde{d}_x & \text{in case (B)}. \end{cases}$$

Here is the place where in case (A) we have to pay for having (8.1) only for w that satisfy homogeneous Dirichlet boundary conditions, which was already announced in Remark 8.3. Without this restriction, we would have obtained $s^* = \min(\frac{\tilde{d}_x}{n}, \frac{r_x+\frac{3}{2}-m}{n-1})$. To see this, it is sufficient to consider one term of the differential operator in variational form of type $\int_\Omega g\partial^\alpha v \partial^\beta w$ with g smooth and, since lower order terms do not give additional problems, $|\alpha| = |\beta| = m$. To prove the first estimate of [SS08, Lemma 3.1], the cases $(2m =)|\alpha + \beta| \geq r_x + 1$ and $|\alpha + \beta| < r_x + 1$ were distinguished. The proof in the second case is no longer valid, meaning that in the remainder of [SS08] we have to read $r_x + 1$ as $\min(r_x + 1, 2m)$. The application of [SS08, Thms 4.1, 6.2] now yields the above result.

Remark 8.6. Also the proof of the second estimate of [SS08, Lemma 3.1] dealing with regular entries is no longer valid when the wavelet involved on the highest level does not satisfy (8.1) for all $w \in W_\infty^k(\Omega)$. This affects, however, only matrix entries where such a wavelet touches the boundary. By treating those entries as singular entries, this fact has no consequences on the value of s^*.

In view of (7.1), we need

$$(8.8) \qquad \tilde{d}_x > d_x - m,$$

in case (A) already being implied by (8.7), and in case (A), additionally

$$(8.9) \qquad \frac{\min(r_x+1,2m)+\frac{1}{2}-m}{n-1} > \frac{d_x-m}{n}.$$

For $m = 1$, this additional condition is only satisfied for $n = 2$ or $n = 3$, $r_x = 1$ and $d_x = 3$.

Remark 8.7. To circumvent the problems caused by the restricted kind of cancellation properties of wavelets near the boundary, one might think of considering a boundary value problem with Neumann boundary conditions, so that $V = H^m(\Omega)$. Indeed, in that case all wavelets satisfy (8.1) for all $w \in W_\infty^k(\Omega)$. Yet, now one cannot rely on integration by parts for bounding entries of the stiffness matrix corresponding to wavelets whose supports intersect each other on $\partial\Omega$. As a consequence for this case, only $s^* = \min(\frac{\tilde{d}_x}{n}, \frac{\frac{1}{2}}{n-1})$ in case (A) can be demonstrated (still $s^* = \tilde{d}_x$ in case (B)). This value of s^* in case (A) is never larger than $\frac{d_x-m}{n}$.

If for some $m > 0$ and an n-dimensional, sufficiently smooth (for $m \le 1$ Lipschitz will suffice), closed manifold $\Gamma \subset \mathbb{R}^{n+1}$, $V = H^m(\Gamma)$, and a_2 corresponds to a *singular integral operator* of order $2m$, like the hypersingular operator, s^*-computability of $a_2([\Sigma]_V, [\Sigma]_V)$ has been shown in [GS06b] with $s^* = \min(\frac{\tilde{d}_x+m}{n}, \frac{r_x+\frac{3}{2}-t}{n-1})$ in case (A) and sufficiently smooth Γ. In view of (7.1), here we need $\tilde{d}_x > d_x - 2m$ and again $\frac{r_x+\frac{3}{2}-m}{n-1} > \frac{d_x-m}{n}$. These conditions are already implied by (8.7). As far as we know, s^*-computability in case (B) has not been investigated for integral operators.

Also for non-separable forms $a(t, \cdot, \cdot)$, s^*-computability for a sufficiently large s^* can be valid. For some $m \in \mathbb{N}$ and some bounded domain $\Omega \subset \mathbb{R}^n$, let $V = H_0^m(\Omega)$ and

$$a(t, v, w) = \int_\Omega \sum_{|\alpha|,|\beta| \le m} a_{\alpha,\beta}(t, x) \partial_x^\alpha v(t, x) \partial_x^\beta w(t, x) dx, \quad (v, w \in V)$$

for bounded $(a_{\alpha,\beta})$ with $\sum_{|\alpha|,|\beta|=m} a_{\alpha,\beta}(t, x)\xi^{\alpha+\beta} \gtrsim |\xi|^{2m}$ (a.e. $(t, x) \in I \times \Omega$, $\xi \in \mathbb{R}^n$). Then (1.1) and (1.2) are satisfied. For sufficiently smooth $(a_{\alpha,\beta})$, an application of [SS08, Thms 4.1, 6.2] shows that $\int_I a(t, \Theta \otimes [\Sigma]_V, \Theta \otimes [\Sigma]_V)dt$ and its adjoint are s^*-computable with

$$s^* = \begin{cases} \min(\tilde{d}_t, \frac{\tilde{d}_x}{n}, \frac{\min(r_x+1,2m)+\frac{1}{2}-m}{n-1}) & \text{in case (A)}, \\ \min(\tilde{d}_t, \tilde{d}_x) & \text{in case (B)}. \end{cases}$$

So we end up with the same value of s^* as in the case of having a separable form. We summarize our findings in the following theorem.

Theorem 8.8. *Consider the parabolic problem* (5.4), (5.6), (5.7) *with a form $a(\cdot, \cdot)$ corresponding to a differential or integral operator as considered in this subsection. Consider its representation $\mathbf{Bu} = \mathbf{f}$ using temporal and spatial wavelet bases Θ and Σ as in Sect. 8.1. Then for any $\varepsilon > 0$, the adaptive wavelet methods from* [CDD02] *or* [CDD01],[GHS07] *applied to the normal equations (4.3) with $\mathbf{f}$ and $\mathbf{B}$ as in (6.5) and (6.6), respectively, produce an approximation $\mathbf{u}_\varepsilon$ with*

$$\|u - \mathbf{u}_\varepsilon^\top[\Theta \otimes \Sigma]\|_{\mathcal{X}} \approx \|\mathbf{u} - \mathbf{u}_\varepsilon\| \le \varepsilon.$$

If for some $s > 0$, $\mathbf{u} \in \mathcal{A}_\infty^s(\ell_2(\nabla_\mathcal{X}))$, then $\operatorname{supp} \mathbf{u}_\varepsilon \lesssim \varepsilon^{-1/s} \|\mathbf{u}\|_{\mathcal{A}_\infty^s(\ell_2(\nabla_\mathcal{X}))}^{1/s}$.

If $s \leq \begin{cases} \min(d_t - 1, \frac{d_x - m}{n}) & \text{in case } (A), \\ \min(d_t - 1, d_x - m) & \text{in case } (B), \end{cases}$ and in case (A), $\tilde{d}_x > d_x$ and $\frac{r_x + \frac{3}{2} - m}{n-1} > \frac{d_x - m}{n}$, and for differential operators additionally $\frac{m + \frac{1}{2}}{n-1} > \frac{d_x - m}{n}$, and in case (B) and differential operators $\tilde{d}_x > d_x - m$, then the number of arithmetic operations and storage locations required by one call of the space-time adaptive solver with tolerance ε is bounded by some multiple of $\varepsilon^{-1/s} \|\mathbf{u}\|_{\mathcal{A}_\infty^s(\ell_2(\nabla_\mathcal{X}))}^{1/s} + 1$.

9. The adaptive solution of parabolic problems in high space dimensions

As we have seen, when $\Omega \subset \mathbb{R}^n$ is a product domain, say $\Omega = (0,1)^n$, and the wavelet basis is of tensor product type (i.e., case (B)), then in any case when u is sufficiently smooth, its representation $\mathbf{u} \in \mathcal{A}_\infty^{\min(d_t - 1, d_x - m)}$. So the rate of best N-term approximation does not deteriorate when n increases. What is more, the class $\mathcal{A}_\infty^{\min(d_t - 1, d_x - m)}$ also contains the representations of functions that have very limited (Sobolev) smoothness, which is the reason to study adaptive methods in the first place.

A characterization of $\mathcal{A}_\infty^{\min(d_t - 1, d_x - m)}$, or of $\mathcal{A}_\infty^s$ for $s \in (0, \min(d_t - 1, d_x - m))$, in terms of Besov like spaces, and complementary regularity theorems for the parabolic problem giving sufficient conditions for the solution u to be contained in these spaces is however outside the scope of this paper.

We have seen that if $\mathbf{u} \in \mathcal{A}_\infty^s$ where s thus can be as large as $\min(d_t - 1, d_x - m)$, then the approximations to $\mathbf{u}$ produced by the adaptive wavelet algorithms converge with this rate s in linear complexity. That is, for any $\varepsilon > 0$, they produce an $\mathbf{u}_\varepsilon$ with $\|\mathbf{u} - \mathbf{u}_\varepsilon\| \leq \varepsilon$ and $\#\operatorname{supp} \mathbf{u}_\varepsilon \lesssim \varepsilon^{-1/s} \|\mathbf{u}\|_{\mathcal{A}_\infty^s}^{1/s}$, taking a number of operations $\lesssim \varepsilon^{-1/s} \|\mathbf{u}\|_{\mathcal{A}_\infty^s}^{1/s} + 1$, where moreover, $\|u - \mathbf{u}_\varepsilon^\top \Psi\|_\mathcal{X} \lesssim \varepsilon$. With a general choice of the spatial wavelets in the coordinate directions, it can be expected, however, that the hidden constants in these statements absorbed by the three "$\lesssim$"-symbols grow exponentially with n, making them of very limited practical use for larger values of n. In this section, for a family of second order parabolic problems, and for a special choice of the spatial wavelets, we will show that the aforementioned hidden constants grow at most quadratically with n, and are uniformly bounded as function of singular perturbation parameters.

Let $V = H_0^1((0,1)^n)$, and let

$$(9.1) \qquad a(t; \eta, \zeta) = a(\eta, \zeta) = \int_{(0,1)^n} c_0 \eta \zeta + \sum_{i=1}^n c_i \partial_i \eta \partial_i \zeta$$

with constants $c_0 \geq 0$ and $c_i > 0$ $(1 \leq i \leq n)$. Equipping V with *energy norm* $\|\cdot\| = a(\cdot, \cdot)^{\frac{1}{2}}$, a is bounded ((1.1)) and *elliptic* (i.e., coercive ((1.2)) with $\lambda = 0$) uniformly in n as well as in $c_0 \geq 0$ and $c_i > 0$ $(1 \leq i \leq n)$. Theorem 5.1 shows that for the operator $B \in \mathcal{L}(\mathcal{X}, \mathcal{Y}')$ defined by the corresponding parabolic problem, both B and B^{-1} are bounded *uniformly in n, $c_0 \geq 0$ and $c_i > 0$ $(1 \leq i \leq n)$* (note that this is not necessarily true when a is coercive but not elliptic, i.e., when $c_0 < 0$).

We consider the wavelet basis in space to be of the form $\Sigma = \otimes_{i=1}^n \Sigma_i$. In addition to the wavelet assumptions we made earlier on the Σ_i (being local, piecewise

polynomial of order d_x, globally C^{r_x}, having $\tilde{d}_x$ vanishing moments, and renor-
malized being a basis for $H_0^1(0,1)$ and $H^{-1}(0,1)$), here we assume that Σ_i is an
orthonormal basis for $L_2(0,1)$. Such bases of *multi*-wavelet type were constructed
in [DGH96, DGH99, Goo00]. Thanks to the orthonormality, from [GO95, Prop1
and 2] one deduces that all of $(\lambda_{[\Sigma]_V}^V)^{-1}$, $\Lambda_{[\Sigma]_V}^V$, $(\lambda_{[\Sigma]_{V'}}^{V'})^{-1}$, $\Lambda_{[\Sigma]_{V'}}^{V'}$, $(\lambda_{\Sigma}^{L_2((0,1)^n)})^{-1}$
and $\Lambda_{\Sigma}^{L_2((0,1)^n)}$ are bounded *uniformly in* n, $c_0 \geq 0$ *and* $c_i > 0$ $(1 \leq i \leq n)$,
whereas otherwise in any case some of these expressions are exponentially growing
with n. From (6.1), (6.2), (6.3) and (6.4), we conclude that $(\lambda_{[\Theta \otimes \Sigma]_{\mathcal{X}}}^{\mathcal{X}})^{-1}$, $\Lambda_{[\Theta \otimes \Sigma]_{\mathcal{X}}}^{\mathcal{X}}$,
$(\lambda_{[(\Theta \otimes \Sigma) \times \Sigma]_{\mathcal{Y}}}^{\mathcal{Y}})^{-1}$ and $\Lambda_{[(\Theta \otimes \Sigma) \times \Sigma]_{\mathcal{Y}}}^{\mathcal{Y}}$, and by (2.2) and (2.3), thus $\|\mathbf{B}\|$ and $\|\mathbf{B}^{-1}\|$ are
bounded *uniformly in* n, $c_0 \geq 0$ *and* $c_i > 0$ $(1 \leq i \leq n)$.

Recalling that we solve $\mathbf{Bu} = \mathbf{f}$ by applying an adaptive wavelet method to the
normal equations $\mathbf{B}^*\mathbf{Bu} = \mathbf{B}^*\mathbf{f}$, in view of Theorem 4.13, what is left to investigate
is, for $s \leq \min(d_t - 1, d_x - 1)$, the possible dependence on n and $c_0 \geq 0$ and $c_i > 0$
$(1 \leq i \leq n)$ of the admissibility constants $a_{\mathbf{B},s}$ and $a_{\mathbf{B}^*,s}$.

With $\mathbf{A} := a([\Sigma]_V, [\Sigma]_V)$, $\mathbf{B}$ from (6.6) now reads as

$$
\begin{bmatrix}
(\langle [\Theta]'_{H^1(I)}, \Theta \rangle_{L_2(I)} \otimes \mathrm{Id}_x) \mathbf{D}_1 + (\langle \Theta, \Theta \rangle_{L_2(I)} \otimes \mathbf{A}) \mathbf{D}_2 \\
\mathbf{R}(\Theta, \Sigma)
\end{bmatrix}.
$$

Each column of $\mathbf{D}_1$, $\mathbf{D}_2$ and $\mathbf{R}(\Theta, \Sigma)$ can be computed exactly at unit cost. The
matrices $\langle [\Theta]'_{H^1(I)}, \Theta \rangle_{L_2(I)} \otimes \mathrm{Id}_x$, $\langle \Theta, \Theta \rangle_{L_2(I)}$ and $\mathbf{R}(\Theta, \Sigma)^*$ are ∞-computable with
computability constants independent of Σ and the spatial operator. In [DSS08,
Theorem 3.5], we showed that $\mathbf{A}$ is ∞-computable with

$$(9.2) \qquad\qquad\qquad c_{\mathbf{A},s} \lesssim n$$

independent of $c_0 \geq 0$ and $c_i > 0$ $(1 \leq i \leq n)$ and only dependent on $s \to \infty$. By
Proposition 8.1(b) we infer that for any $s^* > 0$, $\langle \Theta, \Theta \rangle_{L_2(I)} \otimes \mathbf{A}$ is s^*-computable
with for $s < s^*$, $c_{\langle \Theta, \Theta \rangle_{L_2(I)} \otimes \mathbf{A}, s} \lesssim n$ only dependent on $s^* \to \infty$. Now an application
of Theorem 4.10 shows that for any $s^* > 0$, $\mathbf{B}$ and $\mathbf{B}^*$ are s^*-admissible with for
$s < s^*$, $a_{\mathbf{B},s}$, $a_{\mathbf{B}^*,s} \lesssim n$ only dependent on $s \downarrow 0$ or $s^* \to \infty$. We arrive at

Theorem 9.1. *Consider the parabolic problem (5.4), (5.6), (5.7) with $a(\cdot, \cdot)$ as in
(9.1). Consider its representation $\mathbf{Bu} = \mathbf{f}$ using temporal and spatial wavelet bases
Θ and $\Sigma = \otimes_{i=1}^n \Sigma_i$ as in Sect. 8.1, where the Σ_i are L_2-orthonormal.*

*Then for any $\varepsilon > 0$, the adaptive wavelet methods from [CDD02] or [CDD01],
[GHS07] applied to the normal equations (4.3) with $\mathbf{f}$ and $\mathbf{B}$ as in (6.5) and (6.6),
respectively, produce an approximation $\mathbf{u}_\varepsilon$ with*

$$
\|u - \mathbf{u}_\varepsilon^\top [\Theta \otimes \Sigma]\|_{\mathcal{X}} \eqsim \|\mathbf{u} - \mathbf{u}_\varepsilon\| \leq \varepsilon
$$

uniformly in n and $c_0 \geq 0$ and $c_i > 0$ $(1 \leq i \leq n)$.

*If for some $s > 0$, $\mathbf{u} \in \mathcal{A}_\infty^s(\ell_2(\nabla_{\mathcal{X}}))$, then $\operatorname{supp} \mathbf{u}_\varepsilon \lesssim \varepsilon^{-1/s} \|\mathbf{u}\|_{\mathcal{A}_\infty^s(\ell_2(\nabla_{\mathcal{X}}))}^{1/s}$, only
dependent on s when it tends to 0 or ∞, and thus uniformly in n and $c_0 \geq 0$ and
$c_i > 0$ $(1 \leq i \leq n)$.*

*If for arbitrary $s^* > 0$, $s < s^*$, then the number of operations and storage
locations required by one call of the space-time adaptive algorithm with tolerance
$\varepsilon > 0$ is bounded by some multiple of*

$$(9.3) \qquad\qquad\qquad \varepsilon^{-1/s} n^2 \|\mathbf{u}\|_{\mathcal{A}_\infty^s(\ell_2(\nabla_{\mathcal{X}}))}^{1/s} + 1,$$

where this multiple is uniformly bounded in n and $c_0 \geq 0$ and $c_i > 0$ $(1 \leq i \leq n)$ and depends only on $s \downarrow 0$ or $s^* \to \infty$.

Remark 9.2. Instead of (9.1), let us consider a form

$$\tilde{a}(t; \eta, \zeta) = \tilde{a}(\eta, \zeta) = \int_{(0,1)^n} \sum_{|\alpha|, |\beta| \leq 1} c_{\alpha, \beta} \partial^\alpha \eta \partial^\beta \zeta,$$

such that for some absolute constants $c_1, C_1 > 0$, $|\tilde{a}(\eta, \zeta)| \leq C_1 \|\eta\| \|\zeta\|$, $c_1 \|\eta\|^2 \leq \tilde{a}(\eta, \eta)$ $(\eta, \zeta \in V)$. As indicated [DSS08], if the $c_{\alpha,\beta}$ are constants, being zero when $|\alpha| = |\beta|$, then (9.2) is valid with n reading as the number of non-zero coefficients $c_{\alpha,\beta}$ (being in $[n+1, n^2+1]$). For smooth, variable coefficients $c_{\alpha,\beta}$, this holds also true when $s < s^*$ for some $s^* > d_x - 1$, at least when each of these coefficients depend on a uniformly bounded number of space variables. Finally, also first order terms can be included, at the expense of the replacement of n in (9.2) by some polynomially growing factor in n. We conclude that in all cases, Theorem 9.1 holds with n^2 in (9.3) reading as some polynomially growing factor in n.

Appendix A. Proof of Theorem 5.1

We start the proof by noting that w.l.o.g. we can take $\lambda = 0$ in (1.2). Indeed, writing $u(t) = \hat{u}(t)e^{\lambda t}$, $v_1(t) = \check{v}_1(t)e^{-\lambda t}$, and $g(t) = \hat{g}(t)e^{\lambda t}$, one easily verifies that u satisfies (5.4) if and only if $\hat{u}$ solves

$$\int_0^T \langle \tfrac{d\hat{u}}{dt}(t), \check{v}_1(t) \rangle_H + \lambda \langle \hat{u}(t), \check{v}_1(t) \rangle_H + a(t; \hat{u}(t), \check{v}_1(t)) dt + \langle \hat{u}(0), v_2 \rangle_H$$

$$= \int_0^T \langle \hat{g}(t), \check{v}_1(t) \rangle_H dt + \langle h, v_2 \rangle_H$$

for all $\check{v} = (\check{v}_1, v_2) \in \mathcal{Y}$. A straightforward calculation shows that for $u \neq 0$ and with ϱ as defined in Theorem 5.1,

$$\max(\sqrt{1 + 2\lambda^2 \varrho^4}, \sqrt{2}) \leq \|u\|_{\mathcal{X}} / \|\hat{u}\|_{\mathcal{X}} \leq e^{\lambda T} \max(\sqrt{1 + 2\lambda^2 \varrho^4}, \sqrt{2}),$$

and

$$1 \leq \|(g, h)\|_{\mathcal{Y}} / \|(\hat{g}, h)\|_{\mathcal{Y}} \leq e^{\lambda T},$$

with which the influence of $\lambda > 0$ on the norms of B and B^{-1} can be estimated.

We proceed assuming that $\lambda = 0$. It is well known (see e.g. [Bab71] or [Bra01, Th. 3.6]) that bounded invertibility of $B \in \mathcal{L}(\mathcal{X}, \mathcal{Y}')$ is equivalent to the following three conditions on $b(w, v) := (Bw)(v)$.

(A.1) $$M_b := \sup_{0 \neq w \in \mathcal{X}, \, 0 \neq w \in \mathcal{Y}} \frac{|b(w, v)|}{\|w\|_{\mathcal{X}} \|v\|_{\mathcal{Y}}} < \infty \quad (\textit{continuity}),$$

(A.2) $$\beta := \inf_{0 \neq w \in \mathcal{X}} \sup_{0 \neq v \in \mathcal{Y}} \frac{|b(w, v)|}{\|w\|_{\mathcal{X}} \|v\|_{\mathcal{Y}}} > 0 \quad (\textit{inf sup-condition}),$$

(A.3) $$\forall 0 \neq v \in \mathcal{Y}, \quad \sup_{0 \neq w \in \mathcal{X}} |b(w, v)| > 0 \quad (\textit{surjectivity}),$$

where $\|B\|_{\mathcal{X} \to \mathcal{Y}'} = M_b$ and $\|B^{-1}\|_{\mathcal{Y}' \to \mathcal{X}} = \frac{1}{\beta}$.

By assumption (1.1) and (5.3), we infer that

(A.4) $$|b(w, v)| \leq \sqrt{2 \max(1, M_a^2) + M_e^2} \, \|w\|_{\mathcal{X}} \|v\|_{\mathcal{Y}}, \quad (w \in \mathcal{X}, v \in \mathcal{Y}),$$

proving (A.1).

The adjoint $A(t)' : V \to V'$ of $A(t)$ satisfies $\langle A(t)'\eta, \zeta \rangle_H = \overline{a(t; \zeta, \eta)}$. From (1.1) and (1.2) with $\lambda = 0$, we have $\|A(t)'\|_{V \to V'} \le M_a$ and $\|(A(t)')^{-1}\|_{V' \to V} \le 1/\alpha$ for a.e. $t \in (0, T)$.

We verify (A.2): given $0 \ne w \in \mathcal{X}$, we define $z_w(t) := (A(t)')^{-1} \frac{dw}{dt}(t)$, and select $v_w = (v_1, v_2)$ as $v_1(t) = z_w(t) + w(t)$ and $v_2 = w(0)$. From the boundedness of $(A(t)')^{-1}$ and (5.3) we then infer that

$$\text{(A.5)} \qquad \|v_w\|_{\mathcal{Y}} \le \sqrt{2 \max(\alpha^{-2}, 1) + M_e^2} \, \|w\|_{\mathcal{X}}.$$

From (1.2) and (1.1) we obtain

$$\Re\langle \tfrac{dw}{dt}(t), z_w(t) \rangle_H = \Re\langle A(t)' z_w(t), z_w(t) \rangle_H$$
$$= \Re a(t; z_w(t), z_w(t)) \ge \alpha \|z_w(t)\|_V^2 \ge \tfrac{\alpha}{M_a^2} \|\tfrac{dw}{dt}(t)\|_{V'}^2.$$

Using $a(t; w(t), z_w(t)) = \langle w(t), \frac{dw}{dt}(t) \rangle_H$, and

$$\langle \tfrac{dw}{dt}(t), w(t) \rangle_H + \langle w(t), \tfrac{dw}{dt}(t) \rangle_H = \frac{d}{dt}\langle w(t), w(t) \rangle_H,$$

we arrive at

$$\Re B(w, v_w) = \Re \int_0^T \langle \tfrac{dw}{dt}(t), z_w(t) \rangle_H + \langle \tfrac{dw}{dt}(t), w(t) \rangle_H$$
$$+ \, a(t; w(t), z_w(t)) + a(t; w(t), w(t)) dt + \|w(0)\|_H^2$$
$$\ge \int_0^T \tfrac{\alpha}{M_a^2} \|\tfrac{dw}{dt}(t)\|_{V'}^2 + \frac{d}{dt}\langle w(t), w(t) \rangle_H + \alpha \|w(t)\|_V^2 dt + \|w(0)\|_H^2$$
$$= \int_0^T \tfrac{\alpha}{M_a^2} \|\tfrac{dw}{dt}(t)\|_{V'}^2 + \alpha \|w(t)\|_V^2 dt + \|w(T)\|_H^2$$
$$\ge \min(\tfrac{\alpha}{M_a^2}, \alpha) \|w\|_{\mathcal{X}}^2 \ge \frac{\min(\tfrac{\alpha}{M_a^2}, \alpha)}{\sqrt{2 \max(\alpha^{-2}, 1) + M_e^2}} \|w\|_{\mathcal{X}} \|v_w\|_{\mathcal{Y}}.$$

Since $w \in \mathcal{X}$ was arbitrary, we obtain (A.2) with

$$\text{(A.6)} \qquad \beta \ge \frac{\min(\tfrac{\alpha}{M_a^2}, \alpha)}{\sqrt{2 \max(\alpha^{-2}, 1) + M_e^2}}.$$

To prove (A.3), let $\{\phi_i : i \in \mathbb{N}\}$ be a basis for V and let $\text{span}\{\phi_i : i = 1, ..., n\}$ be denoted as V_n. Given $0 \ne (v_1, v_2) \in \mathcal{Y} = L_2(0, T; V) \times H$ and $n \in \mathbb{N}$, we seek $z_n(t) = \sum_{i=1}^n \mathbf{z}_i^{(n)}(t)\phi_i$ such that

$$\text{(A.7)} \qquad \begin{cases} \langle \frac{dz_n}{dt}(t), \zeta_n \rangle_H + a(t; z_n(t), \zeta_n) &= a(t; v_1(t), \zeta_n), \\ z_n(0) &= \sum_{j=1}^n \mathbf{v}_{2,i}^{(n)} \phi_i, \end{cases}$$

for all $\zeta_n \in V_n$ and a.e. $t \in [0, T]$, where $\sum_{i=1}^n \mathbf{v}_{2,i}^{(n)} \phi_i \to v_2$ in H for $n \to \infty$.

Problem (A.7) can be written as

$$\begin{cases} \mathbf{M}^{(n)} \frac{d\mathbf{z}^{(n)}}{dt}(t) + \mathbf{A}^{(n)}(t)\mathbf{z}^{(n)} &= \mathbf{f}^{(n)}(t) \quad (\text{a.e. } t \in [0, T]), \\ \mathbf{z}^{(n)}(0) &= \mathbf{v}_2^{(n)}, \end{cases}$$

where

$$[\mathbf{M}^{(n)}]_{i,j} = \langle \phi_j, \phi_i \rangle_H, [\mathbf{A}^{(n)}(t)]_{i,j} = a(t; \phi_j, \phi_i), [\mathbf{f}^{(n)}(t)]_i = a(t; v_1(t), \phi_i).$$

This system of linear ODE's has a unique solution $z_n \in C([0, T]; V_n)$, with $\frac{dz_n}{dt} \in L_2(0, T; V_n)$.

Next, we show that the sequence $(z_n)_n$ is bounded in $L_2(0, T; V)$. By substituting $\zeta_n = z_n$ in (A.7), integrating over t and taking real parts, we obtain

$$\Re \int_0^T \langle \tfrac{dz_n}{dt}(t), z_n(t) \rangle_H + a(t, z_n(t), z_n(t)) dt = \Re \int_0^T a(t, v_1(t), z_n(t)) dt,$$

or

$$\|z_n(T)\|_H^2 + 2\Re \int_0^T a(t, z_n(t), z_n(t)) dt = \|z_n(0)\|_H^2 + 2\Re \int_0^T a(t, v_1(t), z_n(t)) dt.$$

We may assume that $\|z_n(0)\|_H \leq 2\|v_2\|_H$. Invoking (1.2) with $\lambda = 0$ and (1.1), we infer that for any $\varepsilon > 0$,

$$2\alpha \int_0^T \|z_n(t)\|_V^2 dt \leq 2\|v_2\|_H^2 + 2M_a \int_0^T \|v_1(t)\|_V \|z_n(t)\|_V dt$$

$$\leq 2\|v_2\|_H^2 + M_a \big[\varepsilon \int_0^T \|z_n(t)\|_V^2 dt + \varepsilon^{-1} \int_0^T \|v_1(t)\|_V^2 dt \big].$$

By choosing $\varepsilon = \alpha / M_a$, we arrive at

(A.8) $$\|z_n\|_{L_2(0,T;V)}^2 \leq \tfrac{2}{\alpha} \|v_2\|_H^2 + \tfrac{M_a^2}{\alpha^2} \|v_1\|_{L_2(0,T;V)}^2.$$

For any $1 \leq i \leq n$, $\theta \in C^1([0, T])$ with $\theta(T) = 0$, (A.7) shows that

$$\int_0^T \langle \tfrac{dz_n}{dt}(t), \phi_i \rangle_H \theta(t) dt = \int_0^T a(t; v_1(t) - z_n(t), \phi_i) \theta(t) dt,$$

and so, by integration by parts,

(A.9) $$-\int_0^T \langle z_n(t), \phi_i \rangle_H \theta'(t) dt = \langle z_n(0), \phi_i \rangle_H \theta(0) + \int_0^T a(t; v_1(t) - z_n(t), \phi_i) \theta(t) dt.$$

The boundedness of the sequence $(z_n)_n$ in the Hilbert space $L_2(0, T; V)$ implies that it has a subsequence that converges weakly to some $z \in L_2(0, T; V)$. Denoting this subsequence again as $(z_n)_n$, and taking limits in (A.9), we arrive at

(A.10) $$-\int_0^T \langle z(t), \phi_i \rangle_H \theta'(t) dt = \langle v_2, \phi_i \rangle_H \theta(0) + \int_0^T \langle A(t)(v_1(t) - z(t)), \phi_i \rangle_H \theta(t) dt$$

for any $i \in \mathbb{N}$. Since this last equation is in particular valid for any $\theta \in \mathcal{D}(]0, T[)$, we have that $\frac{dz}{dt} \in \mathcal{D}'(]0, T[; V) \hookrightarrow \mathcal{D}'(]0, T[; V')$ satisfies

$$\langle \tfrac{dz}{dt}(\theta), \phi_i \rangle_H = \langle \int_0^T A(t)(v_1(t) - z(t)) \theta(t) dt, \phi_i \rangle_H$$

or

(A.11) $$\tfrac{dz}{dt} = A(\cdot)(v_1 - z) \text{ in } \mathcal{D}'(]0, T[; V').$$

Since $v_1 - z \in L_2(0, T; V)$, and $A(\cdot) : L_2(0, T; V) \to L_2(0, T; V')$ is bounded, we conclude that $\frac{dz}{dt} \in L_2(0, T; V')$, and thus $z \in \mathcal{X}$. As a consequence, by integration by parts, in (A.10) we may replace $-\int_0^T \langle z(t), \phi_i \rangle_H \theta'(t) dt$ by $\int_0^T \langle \frac{dz}{dt}(t), \phi_i \rangle_H \theta(t) dt + \langle z(0), \phi_i \rangle_H \theta(0)$, and from (A.11) we conclude that $z(0) = v_2$. Since $\mathcal{D}(]0, T[) \otimes V$ is dense in $L_2(0, T; V)$, again (A.11) shows that

$$b(z, w) = \int_0^T a(t, v_1(t), w_1(t)) dt + \langle v_2, w_2 \rangle_H \quad \forall w = (w_1, w_2) \in \mathcal{Y}.$$

Substituting $w = (v_1, v_2)$, we conclude (A.3) which completes the proof of Theorem 5.1.

References

[AKV06]   J. Alam, N. Kevlahan and O. Vasilyev. Simultaneous space-time adaptive wavelet solution of nonlinear parabolic differential equations. *J. Comput. Phys.*, 214(2):829–857, 2006.

[Bab71]   I. Babuška. Error-bounds for finite element method. *Numer. Math.*, 16:322–333, 1970/1971.

[Bar05]   A. Barinka. *Fast Evaluation Tools for Adaptive Wavelet Schemes*. PhD thesis, RTWH Aachen, March 2005.

[BG04]   H.J. Bungartz and M. Griebel. Sparse grids. *Acta Numer.*, 13:147–269, 2004.

[Bra01]   D. Braess. *Finite Elements*. Cambridge University Press, 2001. Second edition.

[CDD01]   A. Cohen, W. Dahmen, and R. DeVore. Adaptive wavelet methods for elliptic operator equations – Convergence rates. *Math. Comp*, 70:27–75, 2001.

[CDD02]   A. Cohen, W. Dahmen, and R. DeVore. Adaptive wavelet methods II - Beyond the elliptic case. *Found. Comput. Math.*, 2(3):203–245, 2002.

[CF04]   Z. Chen and J. Feng. An adaptive finite element algorithm with reliable and efficient error control for linear parabolic problems. *Math. Comp.*, 73(247):1167–1193 (electronic), 2004.

[CQ92]   Ch. Chui and E. Quak. Wavelets on a bounded interval. In *Numerical methods in approximation theory, Vol. 9 (Oberwolfach, 1991)*, volume 105 of *Internat. Ser. Numer. Math.*, pages 53–75. Birkhäuser, Basel, 1992.

[DFR+07]   S. Dahlke, M. Fornasier, T. Raasch, R.P. Stevenson, and M. Werner. Adaptive frame methods for elliptic operator equations: The steepest descent approach. *IMA J. Numer. Math.*, 27(4):717–740, 2007.

[DGH96]   G.C. Donovan, J.S. Geronimo, and D.P. Hardin. Intertwining multiresolution analyses and the construction of piecewise-polynomial wavelets. *SIAM J. Math. Anal.*, 27(6):1791–1815, 1996.

[DGH99]   G.C. Donovan, J.S. Geronimo, and D.P. Hardin. Orthogonal polynomials and the construction of piecewise polynomial smooth wavelets. *SIAM J. Math. Anal.*, 30(5):1029–1056, 1999.

[DKU99]   W. Dahmen, A. Kunoth, and K. Urban. Biorthogonal spline-wavelets on the interval - Stability and moment conditions. *Appl. Comp. Harm. Anal.*, 6:132–196, 1999.

[DL92]   R. Dautray and J.-L. Lions. *Mathematical analysis and numerical methods for science and technology. Vol. 5.* Springer-Verlag, Berlin, 1992. Evolution problems I.

[DS98]   W. Dahmen and R. Schneider. Wavelets with complementary boundary conditions—function spaces on the cube. *Results Math.*, 34(3-4):255–293, 1998.

[DS99]   W. Dahmen and R. Schneider. Wavelets on manifolds I: Construction and domain decomposition. *SIAM J. Math. Anal.*, 31:184–230, 1999.

[DSS08]   T J. Dijkema, Ch. Schwab, and R.P. Stevenson. An adaptive wavelet method for solving high-dimensional elliptic PDEs. Technical report, January 2008. To appear.

[EJ91]   K. Eriksson and C. Johnson. Adaptive finite element methods for parabolic problems. I. A linear model problem. *SIAM J. Numer. Anal.*, 28(1):43–77, 1991.

[EJ95]   K. Eriksson and C. Johnson. Adaptive finite element methods for parabolic problems. II. Optimal error estimates in $L_\infty L_2$ and $L_\infty L_\infty$. *SIAM J. Numer. Anal.*, 32(3):706–740, 1995.

[EJT85]   K. Eriksson, C. Johnson, and V. Thomée. Time discretization of parabolic problems by the discontinuous Galerkin method. *RAIRO Modél. Math. Anal. Numér.*, 19(4):611–643, 1985.

[Gan06]   T. Gantumur. An optimal adaptive wavelet method for nonsymmetric and indefinite elliptic problems. *J. Comput. Appl. Math.*, 211(1), 90–102, 2008.

[GHS07]   T. Gantumur, H. Harbrecht, and R.P. Stevenson. An optimal adaptive wavelet method without coarsening of the iterands. *Math. Comp.*, 77:615–629, 2007.

[GS06a]   T. Gantumur and R.P. Stevenson. Computation of differential operators in wavelet coordinates. *Math. Comp.*, 75:697–709, 2006.

[GS06b]   T. Gantumur and R.P. Stevenson. Computation of singular integral operators in wavelet coordinates. *Computing*, 76:77–107, 2006.

[GK00]   M. Griebel and S. Knapek. Optimized tensor-product approximation spaces. *Constr. Approx.*, 16(4):525–540, 2000.

[GO95]   M. Griebel and P. Oswald. Tensor product type subspace splittings and multilevel iterative methods for anisotropic problems. *Adv. Comput. Math.*, 4(1–2):171–206, 1995.

[GO07]   M. Griebel and D. Oeltz. A Sparse Grid Space-Time Discretization Scheme for Parabolic Problems. *Computing*, 2007. DOI 10.1007/s00607-007-0241-3.

[Goo00]   T.N.T. Goodman. Biorthogonal refinable spline functions. In A. Cohen, C. Rabut, and L.L. Schumaker, editors, *Curve and Surface Fitting: Saint-Malo 1999*, pages 1–8, Nashville, TN, 2000. Vanderbilt University Press.

[HS06]   H. Harbrecht and R.P. Stevenson. Wavelets with patchwise cancellation properties. *Math. Comp.*, 75(256):1871–1889, 2006.

[KS06]   A. Kunoth and J. Sahner. Wavelets on manifolds: An optimized construction. *Math. Comp.*, 75:1319–1349, 2006.

[Lan01]   J. Lang. *Adaptive multilevel solution of nonlinear parabolic PDE systems*, volume 16 of *Lecture Notes in Computational Science and Engineering*. Springer-Verlag, Berlin, 2001. Theory, algorithm, and applications.

[Met02]   A. Metselaar. *Handling Wavelet Expansions in Numerical Methods*. PhD thesis, University of Twente, 2002.

[Pic98]   M. Picasso. Adaptive finite elements for a linear parabolic problem. *Comput. Methods Appl. Mech. Engrg.*, 167(3-4):223–237, 1998.

[Pri06]   M. Primbs. *Stabile biorthogonale Spline-Waveletbasen auf dem Intervall*. PhD thesis, Universität Duisburg, 2006.

[Raa07]   T. Raasch. *Adaptive Wavelet and Frame Schemes for Elliptic and Parabolic Equations*. PhD thesis, Philipps-Universität Marburg, 2007.

[SS08]   Ch. Schwab and R.P. Stevenson. Adaptive wavelet algorithms for elliptic PDEs on product domains. *Math. Comp.*, 77:71–92, 2008.

[Ste03]   R.P. Stevenson. Adaptive solution of operator equations using wavelet frames. *SIAM J. Numer. Anal.*, 41(3):1074–1100, 2003.

[Tho06]   V. Thomée. *Galerkin finite element methods for parabolic problems*, volume 25 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 2006.

[Ver96]   R. Verfürth. *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*. Wiley-Teubner, Chichester, 1996.

[vPS04]   T. von Petersdorff and Ch. Schwab. Numerical solution of parabolic equations in high dimensions. *M2AN Math. Model. Numer. Anal.*, 38(1):93–127, 2004.

[Wlo82]   J. Wloka. *Partielle Differentialgleichungen*. B. G. Teubner, Stuttgart, 1982. Sobolevräume und Randwertaufgaben.