

Multilevel Preconditioning for Variational Problems

Angela Kunoth

Abstract For the numerical solution of linear variational problems involving elliptic partial differential operators in $n \geq 2$ space dimensions, iterative solution schemes are indispensable on account of their problem size. Our guiding principle is to devise iterative solvers which are optimal in the number of arithmetic operations, i.e., which are of linear complexity in the total number of unknowns. For these algorithms, asymptotically optimal preconditioners are required. The class of preconditioners for which this can be shown are of multilevel type, requiring nested approximation spaces to approximate the solution of the system on a fine user-specified grid. For smooth solutions of standard second and fourth order elliptic PDEs (partial differential equations) in variational form, approximations based on tensor products of higher-order B-splines yield high accuracy.

For such problem classes, this survey collects the main ingredients for multilevel preconditioners in terms of higher order B-splines. There are three types of multilevel preconditioners for which asymptotic optimality can be shown. Two of them, the so-called additive preconditioners, are specified for isogeometric analysis involving linear elliptic partial differential operators in terms of variants of the BPX (Bramble-Pasciak-Xu) preconditioner and wavelet preconditioners. The third class are the so-called multiplicative preconditioners, specifically, multigrid methods.

An essential ingredient for all these multilevel preconditioners are intergrid operators which transform vectors or matrices between grids of different grid spacing. For higher order B-splines, these intergrid operators can be derived from their refinement relations. In addition to a presentation of the theoretical ingredients, the performance of the different preconditioners will be demonstrated by some numerical examples.

A. Kunoth (✉)

Mathematisches Institut, Universität zu Köln, Weyertal 86–90, 50931 Köln, Germany
e-mail: kunoth@math.uni-koeln.de

1 Introduction

For variational systems involving linear elliptic partial differential equations (PDEs) in $n \geq 2$ space dimensions, a standard finite element or finite difference discretization on a uniform grid with grid spacing $0 < h < 1$ leads to the problem to numerically solve a large ill-conditioned system of linear equations. This is due to the fact that PDE operators have positive order $2r$, i.e., $r = 1$ for second order or $r = 2$ for fourth order operators. Ill-conditioned means that the system matrix \mathbf{A}_h exhibits a spectral condition number $\kappa_2(\mathbf{A}_h)$ which is proportional to h^{-2r} , i.e., $\kappa_2(\mathbf{A}_h) \sim h^{-2r}$. Here and in the following, the relation $a \sim b$ stands for $a \lesssim b$ and $b \lesssim a$ where the latter inequality means that b can be bounded by some constant times a uniformly in all parameters on which a and b may depend. Since the convergence speed of any iterative solution scheme depends on the spectral condition number, the scheme will therefore become prohibitively slow. This effect becomes even worse when h is chosen smaller, in order to obtain more accurate approximations; the number of unknowns N increases like $N \sim h^{-n}$ and, thus, the system size also increases accordingly.

On the other hand, solutions of elliptic PDEs typically exhibit a multiscale behaviour. Enhancing iterative methods by multilevel ingredients, therefore, enables one to achieve much more efficient solution schemes. Ultimately, one strives for an ‘optimally efficient scheme’. This means that one can solve the problem with fine grid accuracy with an amount of arithmetic operations that is proportional to the number of unknowns on this grid. The first such methods which were proven to provide an asymptotically optimal iterative scheme were geometric multigrid algorithms [9]. The basic idea of these schemes is to successively solve smaller approximations of the linear system on the finest grid. These can often be interpreted as discretizations with respect to coarser grids. Their iterative solutions can be seen as approximating the inverses of the system matrices applied to the right hand side on the different grids, thereby reducing the spectral condition number of the original system matrix. This idea has, therefore, suggested the term ‘preconditioner’. We call a linear operator \mathbf{C}_h an (*asymptotically*) *optimal preconditioner* if its set-up, storage and application is of optimal linear complexity in the number of unknowns N and if $\kappa_2(\mathbf{C}_h \mathbf{A}_h) \sim 1$ independent of h .

The search for such optimal preconditioners was a major topic for numerical elliptic PDE solvers in the 1980s. The goal was to better understand the ingredients which made a preconditioner optimal. Specifically, one aimed at finding directly applicable versions which could be interpreted as a change of basis. With the arrival of the hierarchical basis preconditioner [54], extending an idea of Babuška from the univariate case, a simple preconditioner became available. Although it is not optimal—one still has logarithmic growth in the grid spacing, $\kappa_2(\mathbf{C}_h \mathbf{A}_h) \sim \log h^{-2r}$, in the bivariate case, and exponential growth for $n = 3$ —its simplicity still makes it popular up to now [42]. Another multilevel preconditioner was presented in [3, 4]. Within the last years, (tensor product) hierarchical B-spline discretizations were increasingly employed in the area of isogeometric analysis, mainly in the context

of deriving discretizations which can be locally adapted to singularities, see, e.g., [36, 51], the survey [5] and the references therein.

At the end of the 1980s, another methodology to derive preconditioners via space decompositions and subspace corrections was developed, see the review [52]. The BPX (Bramble-Pasciak-Xu) preconditioner proposed first in [10] was numerically observed to be optimal; it is based on a hierarchical generator system with grid-dependent weights. With techniques from approximation theory, its optimality was theoretically established in [21, 45]. Since then, its range of application has been widened extensively. For second and fourth order elliptic problems on the sphere a BPX-type preconditioner was developed and its optimality proved for triangular finite elements in [43]. The survey article [53] records extensions of the BPX and of multigrid preconditioners to $H(\text{grad})$, $H(\text{curl})$, and $H(\text{div})$ systems on adaptive and unstructured grids; to name just two extensions.

Multigrid preconditioners for isogeometric analysis were presented in [27], whereas domain decomposition type preconditioners were proposed in [6]. Within the class of domain decomposition methods, also tearing and interconnecting methods (FETI-BETI) are important [34]. Note that not all of these preconditioners are or have been proven to be asymptotically optimal.

In this survey, we present the main ideas of the BPX preconditioner from [11] in the context of isogeometric analysis, employing tensor products of higher order B-splines in Sect. 2. We will see that the main theoretical tool to prove optimality of the BPX preconditioner are multilevel characterizations of the underlying energy space, so-called norm equivalences between Sobolev space norms and weighted sequence norms, describing their subspace contributions. We will also see that the main computational ingredients are linear intergrid operators which map vectors and matrices between grids of different grid spacings.

At about the same time, wavelets as a special example of a multiscale orthogonal basis of $L_2(\mathbb{R})$ with compact support were constructed [25]. While initially mainly developed and used for signal analysis and image compression, wavelets were soon discovered to provide optimal preconditioners in the above sense for second order elliptic boundary value problems [21, 33]. However, the fact that one cannot exploit L_2 -orthogonality for elliptic boundary value problems together with the difficulty that the L_2 -orthogonal Daubechies wavelets are only given implicitly led to the search for variants which are more practical for the numerical solution of PDEs. Soon, it was realized that biorthogonal spline-wavelets as developed in [17] are better suited since they allow one to work with piecewise polynomials for the discretization.

The principal and crucial property to prove optimality of a wavelet preconditioner are norm equivalences between Sobolev norms and sequence norms of weighted wavelet expansion coefficients. I mentioned before that multilevel characterizations for energy spaces are the crucial tool for proving optimality of the BPX preconditioner. Once a wavelet basis is available, one can represent the subspace contributions appearing there explicitly in terms of weighted wavelet coefficients. On this basis, optimal conditioning of the resulting linear system of equations can be achieved by applying the FWT (Fast Wavelet Transform) to a single-scale

discretization on a uniform grid. While again multilevel characterizations of the underlying energy space play the crucial theoretical role for proving optimality, the intergrid operators which perform bases changes between different levels of resolution are the main practical ingredient for the efficiency of the FWT.

Nowadays, the terminology ‘wavelets’ is used in a more general sense that originally in [25]; we consider classes of multiscale bases with three main features: (R) they provide Riesz bases for the underlying function spaces, (L) the basis functions are local, and (CP) they exhibit cancellation properties. These will be detailed in Sect. 3.

After the initial results concerning optimal preconditioning with wavelets of local support in [21], research on employing wavelets for numerically solving elliptic PDEs went into different directions. One problem was that the original constructions in [17, 25] and many others were based on using the Fourier transform so that these constructions provide bases only for function spaces on all of \mathbb{R} , on the torus or, by tensorization, on \mathbb{R}^n . In contrast, PDEs naturally live on a bounded domain $\Omega \subset \mathbb{R}^n$. In order for wavelets to be employed for numerical PDEs, there arose the need for constructions of wavelets on bounded intervals and domains without, of course, loosing the crucial properties (R), (L) and (CP). The first such systematic construction of biorthogonal spline-wavelets on $[0, 1]$ of arbitrary order and, by tensor products, on $[0, 1]^n$, was provided in [22]. Different domain decomposition approaches yield constructions of biorthogonal wavelets on domains which can be represented as unions of parametric mappings of $[0, 1]^n$ [13, 23, 24, 40], see also [50] for details. Once such bases are available, the absolute value of the condition numbers of (systems of) elliptic PDEs can be ameliorated significantly by further inexpensive linear transformations taking into account a setup of the system matrices on the coarsest grid, a so-called operator-based preconditioning [12, 46]. A more recent survey on the results of wavelet-based preconditioning with extensions to PDE-constrained control problems can be found in [39].

Aside from optimal preconditioning, the built-in potential of local adaptivity for wavelets plays a prominent role when solving elliptic or parabolic PDEs with non-smooth solutions, on account of the fact that wavelets provide a locally supported Riesz basis for a whole function space. This issue is extensively addressed in the more recent survey paper [48]. In addition to the material in [26], there are at least four extensive surveys on wavelet and multiscale methods for more general PDEs addressing, among other things, the connection between adaptivity and nonlinear approximation and the evaluation of nonlinearities [15, 18–20].

In this article, I want to remain focused on multilevel preconditioning with higher-order discretizations for smooth solutions for which uniform grids provide a user-specified accuracy.

Isogeometric analysis is an the increasingly popular field in which higher order B-Splines are employed to reach this accuracy. Another such area is mathematical finance; specifically, option pricing problems. The fair pricing of an American option can in a standard model be formulated as a parabolic boundary value problem involving Black–Scholes’ equation with a free boundary. The aim is to compute the free boundary, the optimal exercise price of the option, together with the solution of

the PDE, the value of the option. For American put options, one generally does not have closed form solutions so that one has to resort to numerical schemes. In the simplest Black–Scholes’ model, the volatility is assumed to be constant.

Numerical schemes for American option pricing are typically based on finite difference approaches, see, e.g., [1] and the references therein. In [14] and [44], corresponding multigrid methods are developed. However, here one is not only interested in the solution of the PDE, the option prize. For developing hedging strategies with options, it is of even more importance to accurately compute first and second derivatives of the solution, the so-called Greek letters. The availability of a smooth approximation enables to compute pointwise derivatives and, therefore, avoids additional numerical approximations of derivatives. Therefore, the idea in [30, 31] was to employ higher order B-splines for the solution of the variational inequality derived from Black-Scholes equation for American options. This approach enabled us to pointwise differentiate the option prize and, therefore, achieve a high accuracy approximation for the Greek letters. For the numerical solution, we extended the ideas about monotone multigrid methods with linear finite elements to solve variational inequalities from [35] to higher order B-splines. Here the intergrid operators mentioned before which stem from the refinement relations for B-splines had to be adapted in order not to violate the variational inequality; a task which is for higher order B-splines much more difficult than for linear finite elements since they are not interpolatory any more. Specific care has to be taken when deriving these intergrid operators in order not to violate the inequality constraints. The key feature was to replace function values by B-spline expansion coefficients which remain of the same sign because of the positivity of B-splines. Naturally, these ideas can be extended to obstacle problems which lead to variational inequalities and for which higher order approximations are sought.

While these results were very satisfying from a numerical point of view, the underlying Black-Scholes model has a severe deficiency in assuming that the volatility is constant. Particularly, an effect called volatility smile was observed in [32]. There are several approaches to estimate the volatility from observed stock data. Therefore, we adopted Heston’s approach [28] to model the volatility to satisfy a stochastic differential equation. For Heston’s model and one asset, the standard Itô approach yields a variational inequality in two space variables with a ‘mildly’ nonsymmetric parabolic differential operator. The ideas presented in [41] was as follows. A variational inequality for the American option pricing problem with Heston’s model was discretized in terms of linear finite elements with respect to space. The resulting linear inequality system was solved in each time step with optimal linear computational complexity using a projective Gauss-Seidel scheme together with a monotone multigrid method. For this, again appropriate intergrid operators were constructed. Unfortunately, due to page limitations, I am not able to present these approaches here and have to refer to [30, 31, 41].

Once optimal preconditioners are available, for any of the problems described above, one can construct for a standard elliptic PDE optimal iterative schemes to achieve discretization error accuracy on the finest level with optimal linear complexity as follows. The idea, the so-called *nested iteration*, has been employed

together with multigrid methods for a long time: starting on a coarse grid, one iteratively solves the preconditioned linear system of equations up to discretization error accuracy on this level. The resulting solution is transformed to the next finer grid, employing again the linear intergrid operator from the coarse to the finer grid. In this way, only a fixed number of iterations are necessary on each grid, independent of that grid spacing, see, e.g., [39]. One should note that, in contrast, iterating only on the finest grid introduces an additional $\log N$ factor, where N denotes the amount of unknowns on the finest grid.

The structure of this survey paper is as follows. Section 2 is devoted to the description of the BPX preconditioner for isogeometric analysis together with the proof of its optimality and corresponding numerical results. Section 3 is concerned with wavelet approximations of solutions of PDEs, and the derivation of Fast Wavelet Transforms for optimal preconditioning. We conclude in Sect. 4 with a short summary and some outlook.

2 BPX Preconditioning for Isogeometric Analysis

In this section, we consider linear elliptic PDEs in the framework of isogeometric analysis, combining modern techniques from computer aided design with higher order approximations of the solution. We treat the physical domain by means of a regular B-spline mapping from the parametric domain $\hat{\Omega} = (0, 1)^n$, $n \geq 2$, to the physical domain Ω . The numerical solution of the PDE is computed by means of tensor product B-splines mapped onto the physical domain. We will construct additive BPX-type multilevel preconditioners and show that they are asymptotically optimal. This means that the spectral condition number of the resulting preconditioned stiffness matrix is independent of the grid spacing h . Together with a nested iteration scheme, this enables an iterative solution scheme of optimal linear complexity. The theoretical results are substantiated by numerical examples in two and three space dimensions. The results of this section are essentially contained in [11].

We consider linear elliptic partial differential operators of order $2r = 2, 4$ on the domain Ω in variational form: for given $f \in H^{-r}(\Omega)$, find $u \in H_0^r(\Omega)$ such that

$$a(u, v) = \langle f, v \rangle \quad \text{for all } v \in H_0^r(\Omega) \quad (1)$$

holds. Here the energy space is $H_0^r(\Omega)$, a subset of the Sobolev space $H^r(\Omega)$, the space of square integrable functions with square integrable derivatives up to order r , containing homogeneous Dirichlet boundary conditions for $r = 1$ and homogeneous Dirichlet and Neumann derivatives for $r = 2$. The bilinear form $a(\cdot, \cdot)$ is derived from the linear elliptic PDE operator in a standard fashion, see, e.g., [8]. For example, the Laplacian is represented as $a(v, w) = \int_{\Omega} \nabla v \cdot \nabla w \, dx$. In order for the problem to be well-posed, we require the bilinear form $a(\cdot, \cdot) : H_0^r(\Omega) \times H_0^r(\Omega) \rightarrow \mathbb{R}$ to be symmetric, continuous and coercive on $H_0^r(\Omega)$. With $\langle \cdot, \cdot \rangle$, we denote on the right hand side of (1) the dual form between $H^{-r}(\Omega)$ and $H_0^r(\Omega)$. Our

model problem (1) covers the second order Laplacian with homogeneous boundary conditions

$$-\Delta u = f \quad \text{on } \Omega, \quad u|_{\partial\Omega} = 0, \tag{2}$$

as well as fourth order problems with corresponding homogeneous Dirichlet boundary conditions,

$$\Delta^2 u = f \quad \text{on } \Omega, \quad u|_{\partial\Omega} = \mathbf{n} \cdot \nabla u|_{\partial\Omega} = 0 \tag{3}$$

where $\partial\Omega$ denotes the boundary of Ω and \mathbf{n} the outward normal derivative at $\partial\Omega$. These PDEs serve as prototypes for more involved PDEs like Maxwell’s equation or PDEs for linear and nonlinear elasticity. The reason we formulate these model problems of order $2r$ involving the parameter r is that this exhibits more clearly the order of the operator and the scaling in the subsequently used characterization of Sobolev spaces $H^r(\Omega)$. Thus, for the remainder of this section, the parameter $2r$ denoting the order of the PDE operator is fixed.

The assumptions on the bilinear form $a(\cdot, \cdot)$ entail that there exist constants $0 < c_A \leq C_A < \infty$ such that the induced self-adjoint operator $\langle Av, w \rangle := a(v, w)$ satisfies the isomorphism relation

$$c_A \|v\|_{H^r(\Omega)} \leq \|Av\|_{H^{-r}(\Omega)} \leq C_A \|v\|_{H^r(\Omega)}, \quad v \in H_0^r(\Omega). \tag{4}$$

If the precise format of the constants in (4) does not matter, we abbreviate this relation as $\|v\|_{H^r(\Omega)} \lesssim \|Av\|_{H^{-r}(\Omega)} \lesssim \|v\|_{H^r(\Omega)}$, or shortly as

$$\|Av\|_{H^{-r}(\Omega)} \sim \|v\|_{H^r(\Omega)}. \tag{5}$$

Under these conditions, Lax-Milgram’s theorem guarantees that, for any given $f \in H^{-r}(\Omega)$, the operator equation derived from (1)

$$Au = f \quad \text{in } H^{-r}(\Omega) \tag{6}$$

has a unique solution $u \in H_0^r(\Omega)$, see, e.g., [8].

In order to approximate the solution of (1) or (6), we choose a finite-dimensional subspace $V_h \subset H_0^r(\Omega)$. We will construct these approximation spaces by using tensor products of B-splines as specified next.

2.1 B-Spline Discretizations

Our construction of optimal multilevel preconditioners will rely on tensor products so that principally any space dimension $n \in \mathbb{N}$ is permissible as long as storage permits; the examples cover the cases $n = 2, 3$. As discretization space, we choose

in each spatial direction B-splines of the same degree p on uniform grids and with maximal smoothness. We begin with the univariate case and define B-splines on the interval $[0, 1]$ recursively with respect to their degree p . Given this positive integer p and some $m \in \mathbb{N}$, we call $\mathcal{E} := \{\xi_1, \dots, \xi_{m+p+1}\}$ a p -open knot vector if the knots are chosen such that

$$0 = \xi_1 = \dots = \xi_{p+1} < \xi_{p+2} < \dots < \xi_m < \xi_{m+1} = \dots = \xi_{m+p+1} = 1, \quad (7)$$

i.e., the boundary knots 0 and 1 have multiplicity $p + 1$ and the interior knots are single. For \mathcal{E} , B-spline functions of degree p are defined following the well-known Cox-de Boor recursive formula, see [7]. Starting point are the piecewise constants for $p = 0$ (or characteristic functions)

$$N_{i,0}(\zeta) = \begin{cases} 1, & \text{if } 0 \leq \xi_i \leq \zeta < \xi_{i+1} < 1, \\ 0, & \text{otherwise,} \end{cases} \quad (8)$$

with the modification that the last B-spline $N_{m,0}$ is defined also for $\zeta = 1$. For $p \geq 1$ the B-splines are defined as

$$N_{i,p}(\zeta) = \frac{\zeta - \xi_i}{\xi_{i+p} - \xi_i} N_{i,p-1}(\zeta) + \frac{\xi_{i+p+1} - \zeta}{\xi_{i+p+1} - \xi_{i+1}} N_{i+1,p-1}(\zeta), \quad \zeta \in [0, 1], \quad (9)$$

with the same modification for $N_{m,p}$. Alternatively, one can define the B-splines explicitly by applying divided differences to truncated powers [7]. This gives a set of m B-splines that form a basis for the space of *splines*, that is, piecewise polynomials of degree p with $p - 1$ continuous derivatives at the internal knots ξ_ℓ for $\ell = p + 2, \dots, m$. (Of course, one can also define B-splines on a knot sequence with multiple internal knots which entails that the spline space is not of maximal smoothness.) For $p = 1$, the B-splines are at least $C^0([0, 1])$ which suffices for the discretization of elliptic PDEs of order 2, and for $p = 2$ they are $C^1([0, 1])$ which suffices for $r = 2$. By construction, the B-spline $N_{i,p}$ is supported in the interval $[\xi_i, \xi_{i+p+1}]$.

These definitions are valid for an arbitrary spacing of knots in \mathcal{E} (7). Recall that smooth solutions of elliptic PDEs can be approximated best with discretizations on a uniform grid. Therefore, we assume from now on that the grid is *uniform*, i.e., $\xi_{\ell+1} - \xi_\ell = h$ for all $\ell = p + 1, \dots, m$.

For n space dimensions, we employ tensor products of the one-dimensional B-splines. We take in each space dimension a p -open knot vector \mathcal{E} and define on the closure of the parametric domain $\hat{\Omega} = (0, 1)^n$ (which we also denote by $\hat{\Omega}$ for simplicity of presentation) the spline space

$$\begin{aligned} S_h(\hat{\Omega}) &:= \text{span} \left\{ B_i(\mathbf{x}) := \prod_{\ell=1}^n N_{i_\ell,p}(x_\ell), i = 1, \dots, N := mn, \mathbf{x} \in \hat{\Omega} \right\} \\ &=: \text{span} \left\{ B_i(\mathbf{x}), i \in \mathbb{I}, \mathbf{x} \in \hat{\Omega} \right\}. \end{aligned} \quad (10)$$

In the spirit of isogeometric analysis, we suppose that the computational domain Ω can also be described in terms of B-splines. We assume that the computational domain Ω is the image of a mapping $\mathbf{F} : \hat{\Omega} \rightarrow \Omega$ with $\mathbf{F} := (F_1, \dots, F_n)^T$ where each component F_i of \mathbf{F} belongs to $S_{\bar{h}}(\hat{\Omega})$ for some given \bar{h} . In many applications, the geometry can be described in terms of a very coarse mesh, namely, $\bar{h} \gg h$. Moreover, we suppose that \mathbf{F} is invertible and satisfies

$$\|D^\alpha \mathbf{F}\|_{L_\infty(\hat{\Omega})} \sim 1 \text{ for } |\alpha| \leq r. \tag{11}$$

This assumption on the geometry can be weakened in that the mapping \mathbf{F} can be a piecewise C^∞ function on the mesh with respect to \bar{h} , independent of h , or the domain Ω may have a multi-patch representation. This means that one can allow Ω also to be the union of domains Ω_k where each one is parametrized by a spline mapping of the parametric domain $\hat{\Omega}$.

We now define the approximation space for (6) as

$$V_h^r := \{v_h \in H_0^r(\Omega) : v_h \circ \mathbf{F} \in S_h(\hat{\Omega})\}. \tag{12}$$

We will formulate three important properties of this approximation space which will play a crucial role later for the construction of the BPX-type preconditioners. The first one is that we suppose from now on that the B-spline basis is *normalized* with respect to L_2 , i.e.,

$$\|B_i\|_{L_2(\hat{\Omega})} \sim 1, \text{ and, thus, also } \|B_i \circ \mathbf{F}^{-1}\|_{L_2(\Omega)} \sim 1 \text{ for all } i \in \mathbb{I}. \tag{13}$$

Then one can derive the following facts [11].

Theorem 1 *Let $\{B_i\}_{i \in \mathbb{I}}$ be the B-spline basis defined in (10) and normalized as in (13), $N = \#\mathbb{I}$ and V_h^r as in (12). Then we have*

(S) Uniform stability with respect to $L_2(\Omega)$

For any $\mathbf{c} \in \ell_2(\mathbb{I})$,

$$\left\| \sum_{i=1}^N c_i B_i \circ \mathbf{F}^{-1} \right\|_{L_2(\Omega)}^2 \sim \sum_{i=1}^N |c_i|^2 =: \|\mathbf{c}\|_{\ell_2}^2, \quad \mathbf{c} := (c_i)_{i=1, \dots, N}; \tag{14}$$

(J) Direct or Jackson estimates

$$\inf_{v_h \in V_h^r} \|v - v_h\|_{L_2(\Omega)} \lesssim h^s |v|_{H^s(\Omega)} \text{ for any } v \in H^s(\Omega), \quad 0 \leq s \leq r, \tag{15}$$

where $|\cdot|_{H^s(\Omega)}$ denotes the Sobolev seminorm of highest weak derivatives s ;

(B) Inverse or Bernstein estimates

$$\|v_h\|_{H^s(\Omega)} \lesssim h^{-s} \|v_h\|_{L_2(\Omega)} \text{ for any } v_h \in V_h^r \text{ and } 0 \leq s \leq r. \tag{16}$$

In all these estimates, the constants are independent of h but may depend on \mathbf{F} , i.e., Ω , the polynomial degree p and the spatial dimension n .

In the next section, we construct BPX-type preconditioners for (6) in terms of approximations with (12) and show their optimality.

2.2 Additive Multilevel Preconditioners

The construction of optimal preconditioners are based on a *multiresolution analysis* of the underlying energy function space $H_0^r(\Omega)$. As before, $2r \in \{2, 4\}$ stands for the order of the PDEs we are solving, and is always kept fixed.

We first describe the necessary ingredients within an abstract basis-free framework, see, e.g., [18]. Afterwards, we specify the realization for the parametrized tensor product spaces in (12).

Let \mathcal{V} be a sequence of strictly nested spaces V_j , starting with some fixed coarsest index $j_0 > 0$, determined by the polynomial degree p which determines the support of the basis functions (which also depends on Ω), and terminating with a highest resolution level J ,

$$V_{j_0} \subset V_{j_0+1} \subset \cdots \subset V_j \subset \cdots \subset V_J \subset H_0^r(\Omega). \quad (17)$$

The index j denotes the level of resolution defining approximations on a grid with dyadic grid spacing $h = 2^{-j}$, i.e., we use from now on the notation V_j instead of V_h to indicate different grid spacings. Then, V_J will be the space relative to the finest grid 2^{-J} . We associate with \mathcal{V} a sequence of linear projectors $\mathcal{P} := \{P_j\}_{j \geq j_0}$ with the following properties.

Properties 2 We assume that

(P1) P_j maps $H_0^r(\Omega)$ onto V_j ,

(P2) $P_j P_\ell = P_j$ for $j \leq \ell$,

(P3) \mathcal{P} is uniformly bounded on $L_2(\Omega)$, i.e., $\|P_j\|_{L_2(\Omega)} \lesssim 1$ for any $j \geq j_0$ with a constant independent of j .

These conditions are satisfied, for example, for $L_2(\Omega)$ -orthogonal projectors, or, in the case of splines, for the quasi-interpolant proposed and analyzed in [47, Chapter 4]. The second condition (P2) ensures that the differences $P_j - P_{j-1}$ are also projectors for any $j > j_0$. We define next a sequence $\mathcal{W} := \{W_j\}_{j \geq j_0}$ of complement spaces

$$W_j := (P_{j+1} - P_j)V_{j+1} \quad (18)$$

which then yields the direct (but not necessarily orthogonal) decomposition

$$V_{j+1} = V_j \oplus W_j. \quad (19)$$

Thus, for the finest level J , we can express V_J in its *multilevel decomposition*

$$V_J = \bigoplus_{j=j_0-1}^{J-1} W_j \tag{20}$$

upon setting $W_{j_0-1} := V_{j_0}$. Setting also $P_{j_0-1} := 0$, the corresponding multilevel representation of any $v \in V_J$ is then

$$v = \sum_{j=j_0}^J (P_j - P_{j-1})v. \tag{21}$$

We now have the following result which will be used later for the proof of the optimality of the multilevel preconditioners.

Theorem 3 *Let \mathcal{P}, \mathcal{V} be as above where, in addition, we require that for each V_j , $j_0 \leq j \leq J$, a Jackson and Bernstein estimate as in Theorem 1 (J) and (B) hold with $h = 2^{-j}$. Then one has the function space characterization*

$$\|v\|_{H^r(\Omega)} \sim \left(\sum_{j=j_0}^J 2^{2rj} \|(P_j - P_{j-1})v\|_{L_2(\Omega)}^2 \right)^{1/2} \quad \text{for any } v \in V_J. \tag{22}$$

Such a result holds for much larger classes of function spaces, Sobolev or even Besov spaces which are subsets of $L_q(\Omega)$ for general q , possibly different from 2 and for any function $v \in H^r(\Omega)$, then with an infinite sum on the right hand side, see, e.g., [18]. The proof of Theorem 3 for such cases heavily relies on tools from approximation theory and can be found in [21, 45].

Next we demonstrate how to exploit the norm equivalence (22) in the construction of an optimal multilevel preconditioner. Define for any $v, w \in V_J$ the linear self-adjoint positive-definite operator $C_J : V_J \rightarrow V_J$ given by

$$(C_J^{-1}v, w)_{L_2(\Omega)} := \sum_{j=j_0}^J 2^{2rj} ((P_j - P_{j-1})v, (P_j - P_{j-1})w)_{L_2(\Omega)}, \tag{23}$$

which we call a multilevel *BPX-type preconditioner*. Let $A_J : V_J \rightarrow V_J$ be the finite-dimensional operator defined by $(A_Jv, w)_{L_2(\Omega)} := a(v, w)$ for all $v, w \in V_J$, the approximation of A in (6) with respect to V_J .

Theorem 4 *With the same prerequisites as in Theorem 3, C_J is an asymptotically optimal symmetric preconditioner for A_J , i.e., $\kappa_2(C_J^{1/2}A_JC_J^{1/2}) \sim 1$ with constants independent of J .*

Proof For the parametric domain $\hat{\Omega}$, the result was proved independently in [21, 45] and is based on the combination of (22) together with the well-posedness of the

continuous problem (6). The result on the physical domain follows then together with (11). \square

Concrete realizations of the preconditioner defined in (23) based on B-splines lead to representations of the complement spaces W_j whose bases are called *wavelets*. For these, efficient implementations of optimal linear complexity involving the Fast Wavelet Transform can be derived explicitly, see Sect. 3.

However, since the order of the PDE operator r is positive, one can use here the argumentation from [10] which will allow to work with the same basis functions as for the spaces V_j . The first part of the argument relies on the assumption that the P_j are L_2 -orthogonal projectors. For a clear distinction, we shall use the notation O_j for L_2 -orthogonal projectors and reserve the notation P_j for the linear projectors with Properties 2. Then, the BPX-type preconditioner (23) (using the same symbol C_J for simplicity) reads as

$$C_J^{-1} := \sum_{j=j_0}^J 2^{2jr} (O_j - O_{j-1}), \quad (24)$$

which is by Theorem 4 a BPX-type preconditioner for the self-adjoint positive definite operator A_J . By the orthogonality of the projectors O_j , we can immediately derive from (24) that

$$C_J = \sum_{j=j_0}^J 2^{-2jr} (O_j - O_{j-1}). \quad (25)$$

Since $r > 0$, by rearranging the sum, the exponentially decaying scaling factors allow one to replace C_J by the spectrally equivalent operator

$$C_J = \sum_{j=j_0}^J 2^{-2jr} O_j. \quad (26)$$

Recall that two linear operators $\mathcal{A} : V_J \rightarrow V_J$ and $\mathcal{B} : V_J \rightarrow V_J$ are *spectrally equivalent* if they satisfy

$$(\mathcal{A}v, v)_{L_2(\Omega)} \sim (\mathcal{B}v, v)_{L_2(\Omega)}, \quad v \in V_J, \quad (27)$$

with constants independent of J . Thus, the realization of the preconditioner is reduced to a computation in terms of the bases of the spaces V_j instead of W_j . The orthogonal projector O_j can, in turn, be replaced by a simpler local operator which is spectrally equivalent to O_j , see [37] and the derivation below.

Up to this point, the introduction to multilevel preconditioners has been basis-free. We now show how this framework can be used to construct a BPX-preconditioner for the linear system (6). Based on the definition (12), we construct

a sequence of spaces satisfying (17) such that $V_J = V_h^r$. In fact, we suppose that for each space dimension we have a sequence of p -open knot vectors $\mathcal{E}_{j_0,\ell}, \dots, \mathcal{E}_{J,\ell}$, $\ell = 1, \dots, n$, which provide a uniform partition of the interval $[0, 1]$ such that $\mathcal{E}_{j,\ell} \subset \mathcal{E}_{j+1,\ell}$ for $j = j_0, j_0 + 1, \dots, J$. In particular, we assume that $\mathcal{E}_{j+1,\ell}$ is obtained from $\mathcal{E}_{j,\ell}$ by dyadic refinement, i.e., the grid spacing for $\mathcal{E}_{j,\ell}$ is proportional to 2^{-j} for each $\ell = 1, \dots, n$. In view of the assumptions on the parametric mapping \mathbf{F} , we assume that $\bar{h} = 2^{-j_0}$, i.e., \mathbf{F} can be represented in terms of B-splines on the coarsest level j_0 . By construction, we have now achieved that

$$S_{j_0}(\hat{\Omega}) \subset S_{j_0+1}(\hat{\Omega}) \subset \dots \subset S_J(\hat{\Omega}).$$

Setting $V_j^r := \{v \in H_0^1(\Omega) : v \circ \mathbf{F} \in S_j(\hat{\Omega})\}$, we arrive at a sequence of nested spaces

$$V_{j_0}^r \subset V_{j_0+1}^r \subset \dots \subset V_J^r.$$

Setting $\mathbb{I}_j := \{1, \dots, \dim S_j(\hat{\Omega})\}$, we denote by B_i^j , $i \in \mathbb{I}_j$, the set of L_2 -normalized B-spline basis functions for the space $S_j(\hat{\Omega})$. Define now the positive definite operator $P_j : L_2(\Omega) \rightarrow V_j^r$ as

$$P_j := \sum_{i \in \mathbb{I}_j} (\cdot, B_i^j \circ \mathbf{F}^{-1})_{L_2(\Omega)} B_i^j \circ \mathbf{F}^{-1}. \tag{28}$$

Corollary 5 *For the basis $\{B_i^j \circ \mathbf{F}^{-1}, i \in \mathbb{I}_j\}$, the operators P_j and the L_2 -projectors O_j are spectrally equivalent for any j .*

Proof The assertion follows by combining (11), (14), with Remark 3.7.1 from [37], see [10] for the main ingredients. □

Finally, we obtain an explicit representation of the preconditioner C_J in terms of the mapped spline bases for V_j^r , $j = j_0, \dots, J$,

$$C_J = \sum_{j=j_0}^J 2^{-2jr} \sum_{i \in \mathbb{I}_j} (\cdot, B_i^j \circ \mathbf{F}^{-1})_{L_2(\Omega)} B_i^j \circ \mathbf{F}^{-1}. \tag{29}$$

Note that this preconditioner involves *all* B-splines from all levels j with an appropriate scaling, i.e., in fact a properly scaled *generating system* for V_J^r .

Remark 6 The *hierarchical basis (HB) preconditioner* introduced for $n = 2$ in [54] for piecewise linear B-splines fits into this framework by choosing *Lagrangian interpolants* in place of the projectors P_j in (23). However, since these operators do not satisfy (P3) in Properties 2, they do not yield an asymptotically optimal preconditioner for $n \geq 2$. Specifically, for $n = 3$, this preconditioner does not have an effect at all.

So far we have not explicitly addressed the dependence of the preconditioned system on p . Since all estimates in Theorem 1 which enter the proof of optimality depend on p , it is to be expected that the absolute values of the condition numbers, i.e., the values of the constants, depend on and increase with p . Indeed, in the next section, we show some numerical results which also aim at studying this dependence.

2.3 Realization of the BPX Preconditioner and Numerical Results

Now we are in the position to describe the concrete implementation of the BPX preconditioner. Its main ingredient are linear *intergrid operators* which map vectors and matrices between different grids. Specifically, we need to define prolongation and restriction operators.

Since $V_j^r \subset V_{j+1}^r$, each B-spline B_i^j on level j can be represented by a linear combination of B-splines B_k^{j+1} on level $j+1$. Arranging the B-splines in the set $\{B_i^j, i \in \mathbb{I}_j\}$ into a vector \mathbf{B}^j in a fixed order, this relation denoted as *refinement relation* can be written as

$$\mathbf{B}^j = \mathbf{I}_j^{j+1} \mathbf{B}^{j+1} \quad (30)$$

with *prolongation operator* \mathbf{I}_j^{j+1} from the trial space V_j^r to the trial space V_{j+1}^r . The restriction \mathbf{I}_{j+1}^j is then simply defined as the transposed operator, i.e., $\mathbf{I}_{j+1}^j = (\mathbf{I}_j^{j+1})^T$. In case of piecewise linear B-splines, this definition coincides with the well known prolongation and restriction operators from finite element textbooks obtained by interpolation, see, e.g., [8].

We will exemplify the construction in case of quadratic and cubic B-splines on the interval, see, e.g., [7], as follows. We equidistantly subdivide the interval $[0, 1]$ into 2^j subintervals and obtain 2^j and $2^j + 1$, respectively, B-splines for $p = 2, 3$ and the corresponding quadratic and cubic spline space V_j^r which is given on this partition, respectively, see Fig. 1 for an illustration. Note that the two boundary functions which do not vanish at the boundary were removed in order to guarantee that $V_j^r \subset H_0^p(\Omega)$. Moreover, recall that the B-splines are L_2 normalized according to (13) which means that B_i^j is of the form $B_i^j(\zeta) = 2^{j/2} B(2^j \zeta - i)$ if B_i^j is an interior function, and correspondingly for the boundary functions.

In order to derive the explicit form of the discretized BPX-preconditioner, for given functions $u_J, v_J \in V_J$ with expansion coefficients $u_{J,k}$ and $v_{J,\ell}$, respectively, we conclude from (29) that

$$\begin{aligned} (C_J u_J, v_J)_{L_2(\Omega)} &= \sum_{k,\ell \in \mathbb{I}_J} u_{J,k} v_{J,\ell} (C_J (B_k^J \circ \mathbf{F}^{-1}), B_\ell^J \circ \mathbf{F}^{-1})_{L_2(\Omega)} \\ &= \sum_{k,\ell \in \mathbb{I}_J} u_{J,k} v_{J,\ell} \sum_{j=j_0}^J 2^{-2jr} \sum_{i \in \mathbb{I}_j} (B_k^J \circ \mathbf{F}^{-1}, B_i^j \circ \mathbf{F}^{-1})_{L_2(\Omega)} \\ &\quad \times (B_i^j \circ \mathbf{F}^{-1}, B_\ell^J \circ \mathbf{F}^{-1})_{L_2(\Omega)}. \end{aligned}$$

Next, one can introduce the mass matrix $\mathbf{M}_J = [(B_k^J \circ \mathbf{F}^{-1}, B_\ell^J \circ \mathbf{F}^{-1})_{L_2(\Omega)}]_{k,\ell}$ and obtains by the use of restrictions and prolongations

$$(C_J u_J, v_J)_{L_2(\Omega)} = \sum_{j=j_0}^J 2^{-2jr} \mathbf{u}_j^T \mathbf{M}_j \mathbf{I}_j^J \mathbf{I}_j^j \mathbf{M}_j \mathbf{v}_j.$$

The mass matrices which appear in this expression can be further suppressed since \mathbf{M}_j is spectrally equivalent to the identity matrix. Finally, the *discretized BPX-preconditioner* to be implemented is of the simple form

$$\mathbf{C}_J = \sum_{j=j_0}^J 2^{-2jr} \mathbf{I}_j^J \mathbf{I}_j^j, \quad (31)$$

involving only restrictions and prolongations. A further simple improvement can be obtained by replacing the scaling factor 2^{-2jr} by $\text{diag}(\mathbf{A}_j)^{-1}$, where $\text{diag}(\mathbf{A}_j)$ denotes the diagonal matrix built from the diagonal entries of the stiffness matrix \mathbf{A}_j . This diagonal scaling has the same effect as the levelwise scaling by 2^{-2jr} but improves the condition numbers considerably, particularly if parametric mappings are involved. Thus, the discretized BPX-preconditioner takes on the form

$$\mathbf{C}_J = \sum_{j=j_0}^J \mathbf{I}_j^J \text{diag}(\mathbf{A}_j)^{-1} \mathbf{I}_j^j \quad (32)$$

which we will use in the subsequent computations presented in Tables 1 and 2. If the condition number $\kappa(\mathbf{A}_{j_0})$ is already high in absolute numbers on the coarsest level j_0 , it is worth to use its exact inverse on the coarse grid, i.e., to apply instead of (32) the operator

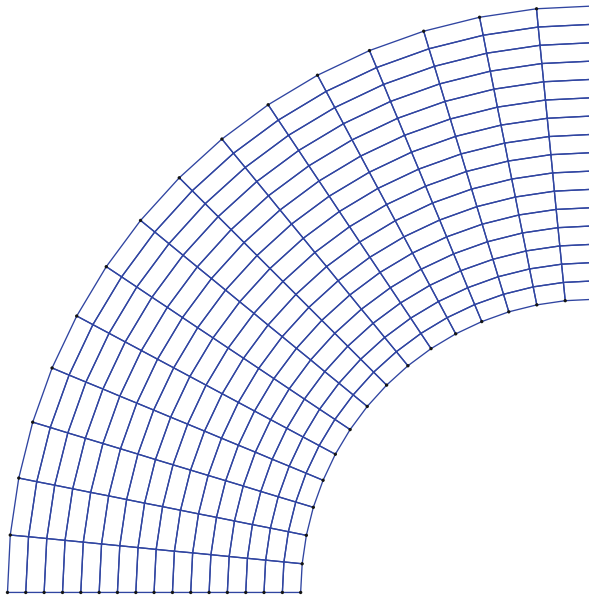
$$\mathbf{C}_J = \mathbf{I}_{j_0}^J \mathbf{A}_{j_0}^{-1} \mathbf{I}_{j_0}^{j_0} + \sum_{j=j_0+1}^J \mathbf{I}_j^J \text{diag}(\mathbf{A}_j)^{-1} \mathbf{I}_j^j,$$

Table 1 Condition numbers of the BPX-preconditioned Laplacian on $\hat{\Omega} = (0, 1)^n$ for $n = 1, 2, 3$

Level	Interval ($n = 1$)				Square ($n = 2$)				Cube ($n = 3$)			
	$p = 1$	$p = 2$	$p = 3$	$p = 4$	$p = 1$	$p = 2$	$p = 3$	$p = 4$	$p = 1$	$p = 2$	$p = 3$	$p = 4$
	3	7.43	3.81	7.03	5.93	5.93	7.31	22.8	133	3.49	39.5	356
4	8.87	4.40	9.47	7.81	5.00	9.03	40.2	225	4.85	50.8	624	9478
5	10.2	4.67	11.0	9.36	5.70	9.72	51.8	293	5.75	56.6	795	11,887
6	11.3	4.87	12.1	10.7	6.27	10.1	58.7	340	6.40	59.7	895	13,185
7	12.2	5.00	12.7	11.5	6.74	10.4	63.1	371	6.91	61.3	961	13,211
8	13.0	5.10	13.0	11.9	7.14	10.5	66.0	391	7.34	62.2	990	13,234
9	13.7	5.17	13.2	12.1	7.48	10.6	68.0	403	7.70	62.6	1016	13,255
10	14.2	5.22	13.4	12.2	7.77	10.6	69.3	411	7.99	62.9	1040	–

Table 2 Condition numbers of the BPX-preconditioned Laplacian on the analytic arc seen on the right hand side. The bracketed numbers are the related condition numbers without preconditioning

Level	$p = 1$	$p = 2$	$p = 3$	$p = 4$
3	5.04 (21.8)	12.4 (8.64)	31.8 (31.8)	184 (184)
4	11.1 (90.2)	16.3 (34.3)	54.7 (32.9)	291 (173)
5	25.3 (368)	19.0 (139)	70.1 (98.9)	376 (171)
6	31.9 (1492)	21.4 (560)	79.2 (401)	436 (322)
7	37.4 (6015)	23.1 (2255)	84.4 (1620)	471 (1297)
8	42.1 (241,721)	24.3 (9062)	87.3 (6506)	490 (5217)
9	45.7 (969,301)	25.2 (36,353)	89.0 (26,121)	500 (20,945)
10	48.8 (388,690)	25.9 (145,774)	90.1 (104,745)	505 (83,975)



see [12, 46]. Another substantial improvement of the BPX-preconditioner can be achieved by replacing the diagonal scaling on each level by, e.g., a SSOR preconditioning as follows. We decompose the system matrix as $\mathbf{A}_j = \mathbf{L}_j + \mathbf{D}_j + \mathbf{L}_j^T$ with the diagonal matrix \mathbf{D}_j , the lower triangular part \mathbf{L}_j , and the upper triangular part \mathbf{L}_j^T . Then we replace the diagonal scaling on each level of the BPX-preconditioner (32) by the SSOR preconditioner, i.e., instead of (32) we apply the preconditioner

$$\mathbf{C}_J = \sum_{j=j_0}^J \mathbf{I}_j^J (\mathbf{D}_j + \mathbf{L}_j)^{-T} \mathbf{D}_j (\mathbf{D}_j + \mathbf{L}_j)^{-1} \mathbf{I}_j. \tag{33}$$

In doing so, the condition numbers can be improved impressively. In Table 3, we list the ℓ_2 -condition numbers for the BPX-preconditioned Laplacian in case of cubic B-splines in two spatial dimensions. By comparing the numbers with those found in Tables 1 and 2 one can infer that the related condition numbers are all reduced by a factor about five. Note that the setup, storage and application of the operator defined in (33) is still of optimal linear complexity.

Finally, we provide numerical results in order to demonstrate the preconditioning and to specify the dependence on the spatial dimension n and the spline degree p . We consider an approximation of the homogeneous Dirichlet problem for the Poisson equation on the n -dimensional unit cube $\hat{\Omega} = (0, 1)^n$ for $n = 1, 2, 3$. The mesh on level j is obtained by subdividing the cube j -times dyadically into 2^n subcubes of mesh size $h_j = 2^{-j}$. On this subdivision, we consider the B-splines of degree $p = 1, 2, 3, 4$ as defined in Sect. 2.1. The ℓ_2 -condition numbers of the related stiffness matrices, preconditioned by the BPX-preconditioner (32), are shown in Table 1. The condition numbers seem to be independent of the level j , but they depend on the spline degree p and the space dimension n for $n > 1$. For fourth order problems on the sphere, corresponding results for the bi-Laplacian with and without BPX preconditioning were presented in [43].

We study next the dependence of the condition numbers on the parametric mapping \mathbf{F} . We consider the case $n = 2$ in case of a smooth mapping (see the plot on the right hand side of Table 2 for an illustration of the mapping). As one can see from Table 2, the condition numbers are at most about a factor of five higher than the related values in Table 1. Nearly the same observation holds if we replace the parametric mapping by a C^0 -parametrization which maps the unit square onto an L-shaped domain, see [11].

If we consider a singular map \mathbf{F} , that is, a mapping that does not satisfy (11), the condition numbers grow considerably as expected, see [11]. But even in this case, the BPX-preconditioner with SSOR acceleration (33) is able to drastically reduce the condition numbers of the system matrix in all examples, see Table 3.

Table 3 Condition numbers of the BPX-preconditioned Laplacian for cubic B-splines on different geometries in case of using a BPX-SSOR preconditioning on each level

Level	Square		C^0 -map	Singular C^1 -map
			of the L-shape	of the L-shape
3	3.61	3.65	3.67	3.80
4	6.58	6.97	7.01	7.05
5	8.47	10.2	10.2	14.8
6	9.73	13.1	13.2	32.2
7	10.5	14.9	15.2	77.7
8	11.0	15.9	16.3	180
9	11.2	16.5	17.0	411
10	11.4	16.9	17.7	933

3 Wavelets and the FWT (Fast Wavelet Transform)

Returning to the abstract setting at the beginning of Sect. 2.2, it will now be specified how to construct and realize the FWT (fast wavelet transform) for preconditioning. I would like to emphasize at this point that wavelets are, more importantly than for preconditioning, an appropriate theoretical tool to derive adaptive discretizations in case of singularities in the domain or data. In fact, they enable one to not only prove convergence of a corresponding adaptive scheme but also optimal complexity of these schemes, see [16] and several subsequent papers or [48] for a more recent survey.

3.1 Some Basic Notions

In view of the problem formulation (6), we need to have a wavelet basis for the space $H'_0(\Omega)$ and its dual. We formulate these basic properties for a general Hilbert space H , following [18] or [38]. A *wavelet basis* for H , shortly *wavelets*, is understood here as a collection of functions

$$\Psi_H := \{\psi_{H,\lambda} : \lambda \in \mathbb{I}_H\} \subset H \tag{34}$$

which are indexed by elements λ from an infinite index set \mathbb{I}_H . Each of the indices λ comprises different information $\lambda = (j, \mathbf{k}, \mathbf{e})$ such as the *refinement scale* or *level of resolution* j and a spatial location $\mathbf{k} = \mathbf{k}(\lambda) \in \mathbb{Z}^n$. In more than one space dimension, the basis functions are built from taking tensor products of certain univariate functions, and in this case the third index \mathbf{e} contains information on the *type* of wavelet. We will frequently use the symbol $|\lambda| := j$ to access the resolution level j . In the univariate case on all of \mathbb{R} , $\psi_{H,\lambda}$ is typically generated by means of shifts and dilates of a single function ψ , i.e., $\psi_\lambda = \psi_{j,k} = 2^{j/2} \psi(2^j \cdot -k)$, $j, k \in \mathbb{Z}$, normalized with respect to $\|\cdot\|_{L_2(\mathbb{R})}$. On bounded domains which can be represented

as unions of mappings of tensor product domains, the structure of the functions is essentially the same up to modifications near the boundary.

We assume that the wavelets satisfy the following crucial properties.

Riesz basis property (R): Every $v \in H$ has a unique expansion in terms of Ψ_H ,

$$v = \sum_{\lambda \in \mathbb{I}_H} v_\lambda \psi_{H,\lambda} =: \mathbf{v}^T \Psi_H, \quad \mathbf{v} := (v_\lambda)_{\lambda \in \mathbb{I}_H}, \quad (35)$$

and its expansion coefficients satisfy the *norm equivalence*

$$\|\mathbf{v}\| \sim \|\mathbf{v}^T \Psi_H\|_H, \quad \mathbf{v} \in \ell_2(\mathbb{I}_H), \quad (36)$$

where $\|\cdot\| := \|\cdot\|_{\ell_2(\mathbb{I}_H)}$.

Locality (L): The functions $\psi_{H,\lambda}$ have compact support which decreases with increasing level $j = |\lambda|$, i.e.,

$$\text{diam}(\text{supp } \psi_{H,\lambda}) \sim 2^{-|\lambda|}. \quad (37)$$

Cancellation property (CP): There exists an integer $\tilde{d} = \tilde{d}_H$ such that

$$\langle v, \psi_{H,\lambda} \rangle \lesssim 2^{-|\lambda|(n/2-n/p+\tilde{d})} |v|_{W_p^{\tilde{d}}(\text{supp } \psi_{H,\lambda})}. \quad (38)$$

This means that integrating against a wavelet has the effect of taking an \tilde{m} th order difference which annihilates the smooth part of v . This property is for wavelets defined on Euclidean domains typically realized by constructing Ψ_H in such a way that it possesses a *dual* or *biorthogonal* basis $\tilde{\Psi}_H \subset H'$ such that the multiresolution spaces $\tilde{S}_j := \text{span}\{\tilde{\psi}_{H,\lambda} : |\lambda| < j\}$ contain all polynomials of order \tilde{m} . Here *dual basis* means that $\langle \psi_{H,\lambda}, \tilde{\psi}_{H,\mu} \rangle = \delta_{\lambda,\mu}$, $\lambda, \mu \in \mathbb{I}_H$.

A few remarks on these properties should be made. In (R), the norm equivalence (36) is crucial since it means control over a function measured in $\|\cdot\|_H$ from above and below by its expansion coefficients: small changes in the coefficients only cause small changes in the function. Together with the locality (L), this also means that local changes stay local. This stability is an important feature which is used for deriving optimal preconditioners. Finally, the cancellation property (CP) entails that smooth functions have small wavelet coefficients which, on account of (36) may be neglected in a controllable way. Moreover, (CP) can be used to derive quasi-sparse representations of a wide class of operators which is important for compression purposes, see, e.g., [48].

By duality arguments one can show that (36) is equivalent to the existence of a biorthogonal collection which is *dual* or *biorthogonal* to Ψ_H ,

$$\tilde{\Psi}_H := \{\tilde{\psi}_{H,\lambda} : \lambda \in \mathbb{I}_H\} \subset H', \quad \langle \psi_{H,\lambda}, \tilde{\psi}_{H,\mu} \rangle = \delta_{\lambda,\mu}, \quad \lambda, \mu \in \mathbb{I}_H, \quad (39)$$

which is a Riesz basis for H' , that is, for any $\tilde{v} = \tilde{v}^T \tilde{\Psi}_H \in H'$ one has

$$\|\tilde{v}\| \sim \|\tilde{v}^T \tilde{\Psi}_H\|_{H'} \tag{40}$$

see [19]. Here the tilde expresses that the collection $\tilde{\Psi}_H$ is a dual basis to a primal one for the space identified by the subscript, so that $\tilde{\Psi}_H = \Psi_{H'}$. Here and in the following, we will view Ψ_H both as in (34) as a *collection* of functions as well as a (possibly infinite) column *vector* containing all functions always assembled in some fixed unspecified order. For a countable collection of functions Θ and some single function σ , the term $\langle \Theta, \sigma \rangle$ is then the column vector with entries $\langle \theta, \sigma \rangle$, $\theta \in \Theta$, and correspondingly $\langle \sigma, \Theta \rangle$ the row vector. For two collections Θ, Σ , the quantity $\langle \Theta, \Sigma \rangle$ is then a (possibly infinite) matrix with entries $(\langle \theta, \sigma \rangle)_{\theta \in \Theta, \sigma \in \Sigma}$ for which $\langle \Theta, \Sigma \rangle = \langle \Sigma, \Theta \rangle^T$. This also implies for a (possibly infinite) matrix \mathbf{C} that $\langle \mathbf{C}\Theta, \Sigma \rangle = \mathbf{C}\langle \Theta, \Sigma \rangle$ and $\langle \Theta, \mathbf{C}\Sigma \rangle = \langle \Theta, \Sigma \rangle \mathbf{C}^T$.

In this notation, the *biorthogonality* or *duality conditions* (39) can be expressed shortly as

$$\langle \Psi, \tilde{\Psi} \rangle = \mathbf{I} \tag{41}$$

with the infinite identity matrix \mathbf{I} .

Wavelets with the above properties can be obtained in the following way. We start from an *anchor basis* $\Psi = \{\psi_\lambda : \lambda \in \mathbb{I} = \mathbb{I}_H\}$ which is a Riesz basis for $L_2(\Omega)$, meaning that Ψ is scaled such that $\|\psi_\lambda\|_{L_2(\Omega)} \sim 1$. Moreover, its dual basis $\tilde{\Psi}$ is also a Riesz basis for $L_2(\Omega)$. Ψ and $\tilde{\Psi}$ are constructed in such a way that rescaled versions of *both bases* $\Psi, \tilde{\Psi}$ form Riesz bases for a whole range of (closed subspaces of) Sobolev spaces $H^s := H^s(\Omega)$, for $0 < s < \gamma, \tilde{\gamma}$, respectively, or subspaces of these. Consequently, one can derive that for each $s \in (-\tilde{\gamma}, \gamma)$ the collection

$$\Psi_s := \{2^{-s|\lambda|} \psi_\lambda : \lambda \in \mathbb{I}\} =: \mathbf{D}^{-s} \Psi \tag{42}$$

is a Riesz basis for H^s [18], i.e., it holds

$$\|\mathbf{v}\| \sim \|\mathbf{v}^T \Psi_s\|_{H^s}, \quad \mathbf{v} \in \ell_2(\mathbb{I}), \tag{43}$$

for each $s \in (-\tilde{\gamma}, \gamma)$. Such a scaling represented by a diagonal matrix \mathbf{D}^s introduced in (42) will play an important role later on. Concrete constructions of wavelet bases with the above properties for parameters $\gamma, \tilde{\gamma} \leq 5/2$ can be found in [22] which cover the situation in (6) for $r = 1, 2$.

3.2 Multiscale Decomposition of Function Spaces

In this section, the basic construction principles of the biorthogonal wavelets with properties (R), (L) and (CP) are summarized, see, e.g., [18]. Their cornerstones are *multiresolution analyses* of the function spaces under consideration as in Sect. 2.2 and the concept of *stable completions*. These concepts are free of Fourier techniques and can therefore be applied to derive constructions of wavelets on domains or manifolds which are subsets of \mathbb{R}^n .

Multiresolution of L_2 (univariate case). Practical constructions of wavelets typically start out with multiresolution analyses of function spaces. Consider a *multiresolution* \mathcal{S} of L_2 which consists of closed subspaces S_j of L_2 , called *trial spaces*, such that they are nested and their union is dense in L_2 ,

$$S_{j_0} \subset S_{j_0+1} \subset \dots \subset S_j \subset S_{j+1} \subset \dots \subset L_2, \quad \text{clos}_{L_2} \left(\bigcup_{j=j_0}^{\infty} S_j \right) = L_2. \quad (44)$$

The index j is the refinement level which appeared already in the elements of the index set \mathbb{I} in (34), starting with some coarsest level $j_0 \in \mathbb{N}_0$. We abbreviate for a finite subset $\Theta \subset L_2$ the linear span of Θ as

$$S(\Theta) = \text{span}\{\Theta\}.$$

Typically the multiresolution spaces S_j have the form

$$S_j = S(\Phi_j), \quad \Phi_j = \{\phi_{j,k} : k \in \Delta_j\}, \quad (45)$$

for some finite index set Δ_j , where the set $\{\Phi_j\}_{j=j_0}^{\infty}$ is *uniformly stable* in the sense that

$$\|\mathbf{c}\|_{\ell_2(\Delta_j)} \sim \|\mathbf{c}^T \Phi_j\|_{L_2}, \quad \mathbf{c} = \{c_k\}_{k \in \Delta_j} \in \ell_2(\Delta_j), \quad (46)$$

holds uniformly in j . Here we have used again the shorthand notation

$$\mathbf{c}^T \Phi_j = \sum_{k \in \Delta_j} c_k \phi_{j,k}$$

and Φ_j denotes both the (column) vector containing the functions $\phi_{j,k}$ as well as the set of functions (45).

The collection Φ_j is called *single scale basis* since all its elements live only on one scale j . In the present context of multiresolution analysis, Φ_j is also called *generator basis* or shortly *generators* of the multiresolution. We assume that the $\phi_{j,k}$ are compactly supported with

$$\text{diam}(\text{supp } \phi_{j,k}) \sim 2^{-j}. \quad (47)$$

It follows from (46) that they are scaled such that

$$\|\phi_{j,k}\|_{L_2} \sim 1 \quad (48)$$

holds. It is known that nestedness (44) together with stability (46) implies the existence of matrices $\mathbf{M}_{j,0}$ such that the two-scale relation

$$\Phi_j = \mathbf{M}_{j,0}^T \Phi_{j+1} \quad (49)$$

holds. Any set of functions satisfying an equation of this form, the *refinement* or *two-scale relation*, will be called *refinable*. The matrices $\mathbf{M}_{j,0}$ are up to boundary modifications exactly the intergrid operators occurring in (30).

Denoting by $[X, Y]$ the space of bounded linear operators from a normed linear space X into the normed linear space Y , one has that

$$\mathbf{M}_{j,0} \in [\ell_2(\Delta_j), \ell_2(\Delta_{j+1})]$$

is *uniformly sparse* which means that the number of entries in each row or column is uniformly bounded. Furthermore, one infers from (46) that

$$\|\mathbf{M}_{j,0}\| = \mathcal{O}(1), \quad j \geq j_0, \quad (50)$$

where the corresponding operator norm is defined as

$$\|\mathbf{M}_{j,0}\| := \sup_{\mathbf{c} \in \ell_2(\Delta_j), \|\mathbf{c}\|_{\ell_2(\Delta_j)}=1} \|\mathbf{M}_{j,0}\mathbf{c}\|_{\ell_2(\Delta_{j+1})}.$$

Since the union of \mathcal{S} is dense in L_2 , a basis for L_2 can be assembled from functions which span any complement between two successive spaces S_j and S_{j+1} , i.e.,

$$S(\Phi_{j+1}) = S(\Phi_j) \oplus S(\Psi_j) \quad (51)$$

where

$$\Psi_j = \{\psi_{j,k} : k \in \nabla_j\}, \quad \nabla_j := \Delta_{j+1} \setminus \Delta_j. \quad (52)$$

The functions Ψ_j are called *wavelet functions* or shortly *wavelets* if, among other conditions detailed below, the union $\{\Phi_j \cup \Psi_j\}$ is still uniformly stable in the sense of (46). Since (51) implies $S(\Psi_j) \subset S(\Phi_{j+1})$, the functions in Ψ_j must also satisfy a matrix-vector relation of the form

$$\Psi_j = \mathbf{M}_{j,1}^T \Phi_{j+1} \quad (53)$$

with a matrix $\mathbf{M}_{j,1}$ of size $(\#\Delta_{j+1}) \times (\#\nabla_j)$. Furthermore, (51) is equivalent to the fact that the linear operator composed of $\mathbf{M}_{j,0}$ and $\mathbf{M}_{j,1}$,

$$\mathbf{M}_j = (\mathbf{M}_{j,0}, \mathbf{M}_{j,1}), \quad (54)$$

is *invertible* as a mapping from $\ell_2(\Delta_j \cup \nabla_j)$ onto $\ell_2(\Delta_{j+1})$. One can also show that the set $\{\Phi_j \cup \Psi_j\}$ is uniformly stable if and only if

$$\|\mathbf{M}_j\|, \|\mathbf{M}_j^{-1}\| = \mathcal{O}(1), \quad j \rightarrow \infty. \quad (55)$$

The particular cases that will be important for practical purposes are when not only $\mathbf{M}_{j,0}$ and $\mathbf{M}_{j,1}$ are uniformly sparse but also the inverse of \mathbf{M}_j . We denote this inverse by \mathbf{G}_j and assume that it is split into

$$\mathbf{G}_j = \mathbf{M}_j^{-1} = \begin{pmatrix} \mathbf{G}_{j,0} \\ \mathbf{G}_{j,1} \end{pmatrix}. \quad (56)$$

A special situation occurs when \mathbf{M}_j is an orthogonal matrix,

$$\mathbf{G}_j = \mathbf{M}_j^{-1} = \mathbf{M}_j^T$$

which corresponds to the case of *L₂ orthogonal wavelets* [25]. A systematic construction of more general \mathbf{M}_j , \mathbf{G}_j for spline-wavelets can be found in [22], see also [18] for more examples, including the hierarchical basis.

Thus, the identification of the functions Ψ_j which span the complement of $S(\Phi_j)$ in $S(\Phi_{j+1})$ is equivalent to completing a given refinement matrix $\mathbf{M}_{j,0}$ to an invertible matrix \mathbf{M}_j in such a way that (55) is satisfied. Any such completion $\mathbf{M}_{j,1}$ is called *stable completion* of $\mathbf{M}_{j,0}$. In other words, the problem of the construction of compactly supported wavelets can equivalently be formulated as an algebraic problem of finding the (uniformly) sparse completion of a (uniformly) sparse matrix $\mathbf{M}_{j,0}$ in such a way that its inverse is also (uniformly) sparse. The fact that inverses of sparse matrices are usually dense elucidates the difficulties in the constructions. Constructions that yield compactly supported wavelets are particularly suited for computations in numerical analysis.

Combining the two-scale relations (49) and (53), one can see that \mathbf{M}_j performs a change of bases in the space S_{j+1} ,

$$\begin{pmatrix} \Phi_j \\ \Psi_j \end{pmatrix} = \begin{pmatrix} \mathbf{M}_{j,0}^T \\ \mathbf{M}_{j,1}^T \end{pmatrix} \Phi_{j+1} = \mathbf{M}_j^T \Phi_{j+1}. \quad (57)$$

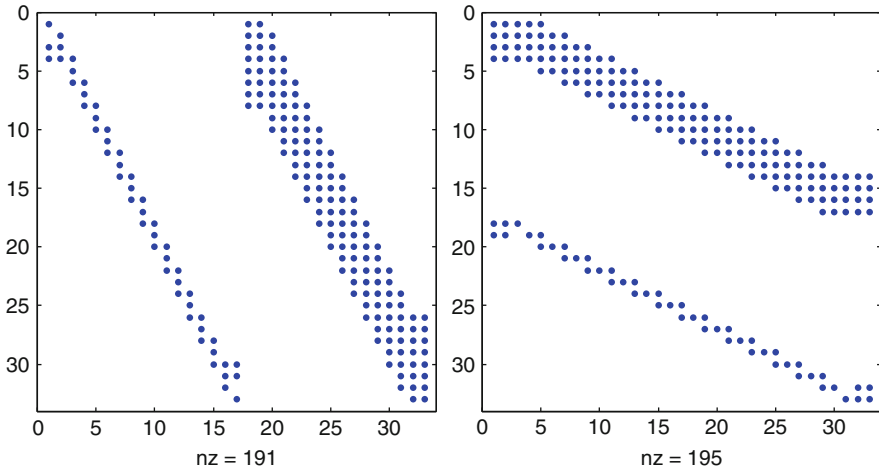


Fig. 2 Nonzero pattern of matrices \mathbf{M}_j (left) and \mathbf{G}_j (right) for boundary-adapted B-splines of order $d = 2$ (degree $p = 1$) as generators and duals of order $\tilde{d} = 4$

Conversely, applying the inverse of \mathbf{M}_j to both sides of (57) results in the *reconstruction identity*

$$\Phi_{j+1} = \mathbf{G}_j^T \begin{pmatrix} \Phi_j \\ \Psi_j \end{pmatrix} = \mathbf{G}_{j,0}^T \Phi_j + \mathbf{G}_{j,1}^T \Psi_j. \tag{58}$$

An example of the structure of the matrices \mathbf{M}_j and \mathbf{G}_j is given in Fig. 2.

Fixing a *finest resolution level* J , one can repeat the decomposition (51) so that $S_J = S(\Phi_J)$ can be written in terms of the functions from the coarsest space supplied with the complement functions from all intermediate levels,

$$S(\Phi_J) = S(\Phi_{j_0}) \oplus \bigoplus_{j=j_0}^{J-1} S(\Psi_j). \tag{59}$$

Thus, every function $v \in S(\Phi_J)$ can be written in its *single-scale representation*

$$v = (\mathbf{c}_J)^T \Phi_J = \sum_{k \in \Delta_J} c_{J,k} \phi_{J,k} \tag{60}$$

as well as in its *multi-scale form*

$$v = (\mathbf{c}_{j_0})^T \Phi_{j_0} + (\mathbf{d}_{j_0})^T \Psi_{j_0} + \dots + (\mathbf{d}_{J-1})^T \Psi_{J-1} \tag{61}$$

with respect to the *multiscale* or *wavelet basis*

$$\Psi^J := \Phi_{j_0} \cup \bigcup_{j=j_0}^{J-1} \Psi_j =: \bigcup_{j=j_0-1}^{J-1} \Psi_j \tag{62}$$

Often the single-scale representation of a function may be easier to compute and evaluate while the multi-scale representation allows one to separate features of the underlying function characterized by different length scales. Since therefore both representations are advantageous, it is useful to determine the transformation between the two representations, commonly referred to as the *Wavelet Transform*,

$$\mathbf{T}_J : \ell_2(\Delta_j) \rightarrow \ell_2(\Delta_j), \quad \mathbf{d}^J \mapsto \mathbf{c}_J, \tag{63}$$

where

$$\mathbf{d}^J := (\mathbf{c}_{j_0}, \mathbf{d}_{j_0}, \dots, \mathbf{d}_{J-1})^T.$$

The previous relations (57) and (58) indicate that this will involve the matrices \mathbf{M}_j and \mathbf{G}_j . In fact, \mathbf{T}_J has the representation

$$\mathbf{T}_J = \mathbf{T}_{J,J-1} \cdots \mathbf{T}_{J,j_0}, \tag{64}$$

where each factor has the form

$$\mathbf{T}_{J,j} := \begin{pmatrix} \mathbf{M}_j & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{(\#\Delta_j - \#\Delta_{j+1})} \end{pmatrix} \in \mathbb{R}^{(\#\Delta_j) \times (\#\Delta_j)}. \tag{65}$$

Schematically \mathbf{T}_J can be visualized as a pyramid scheme

$$\begin{array}{ccccccc} & \mathbf{M}_{j_0,0} & & \mathbf{M}_{j_0+1,0} & & & \mathbf{M}_{J-1,0} \\ \mathbf{c}_{j_0} & \longrightarrow & \mathbf{c}_{j_0+1} & \longrightarrow & \mathbf{c}_{j_0+2} & \longrightarrow & \cdots & \mathbf{c}_{J-1} & \longrightarrow & \mathbf{c}_J \\ & & & & & & & & & \\ & \mathbf{M}_{j_0,1} & & \mathbf{M}_{j_0+1,1} & & & & \mathbf{M}_{J-1,1} & & \\ & \nearrow & & \nearrow & & & \nearrow & \cdots & & \nearrow \\ \mathbf{d}_{j_0} & & \mathbf{d}_{j_0+1} & & \mathbf{d}_{j_0+2} & & & \mathbf{d}_{J-1} & & \end{array} \tag{66}$$

Accordingly, the inverse transform \mathbf{T}_J^{-1} can be written also in product structure (64) in reverse order involving the matrices \mathbf{G}_j as follows:

$$\mathbf{T}_J^{-1} = \mathbf{T}_{J,j_0}^{-1} \cdots \mathbf{T}_{J,J-1}^{-1}, \tag{67}$$

where each factor has the form

$$\mathbf{T}_{J_j}^{-1} := \begin{pmatrix} \mathbf{G}_j & \mathbf{0} \\ \mathbf{0} & \mathbf{I}^{(\#\Delta_j - \#\Delta_{j+1})} \end{pmatrix} \in \mathbb{R}^{(\#\Delta_j) \times (\#\Delta_j)}. \tag{68}$$

The corresponding pyramid scheme is then

$$\begin{array}{ccccccc} & \mathbf{G}_{J-1,0} & & \mathbf{G}_{J-2,0} & & & \mathbf{G}_{j_0,0} \\ \mathbf{c}_J & \longrightarrow & \mathbf{c}_{J-1} & \longrightarrow & \mathbf{c}_{J-2} & \longrightarrow & \cdots & \longrightarrow & \mathbf{c}_{j_0} \\ & & & & & & & & \\ & \mathbf{G}_{J-1,1} & & \mathbf{G}_{J-2,1} & & & \mathbf{G}_{j_0,1} & & \\ & \searrow & & \searrow & & & \searrow & & \searrow \\ & & \mathbf{d}_{J-1} & & \mathbf{d}_{J-2} & & \mathbf{d}_{j_0-1} & & \mathbf{d}_{j_0} \end{array} \tag{69}$$

Property (55) and the fact that \mathbf{M}_j and \mathbf{G}_j can be applied in $(\#\Delta_{j+1})$ operations uniformly in j entails that the complexity of applying \mathbf{T}_J or \mathbf{T}_J^{-1} using the pyramid scheme is of order $\mathcal{O}(\#\Delta_J) = \mathcal{O}(\dim S_J)$ uniformly in J . For this reason, \mathbf{T}_J is called the *Fast Wavelet Transform* (FWT). Note that one should *not* explicitly assemble \mathbf{T}_J or \mathbf{T}_J^{-1} . In fact, due to the particular band structure of \mathbf{M}_j and \mathbf{G}_j , this would result in matrices with $\mathcal{O}(J\#\Delta_J)$ entries. In Table 4 at the end of this section, spectral condition numbers for the FWT for different constructions of biorthogonal wavelets on the interval computed in [46] are displayed.

Since $\cup_{j \geq j_0} S_j$ is dense in L_2 , a basis for the whole space L_2 is obtained when letting $J \rightarrow \infty$ in (62),

$$\begin{aligned} \Psi &:= \bigcup_{j=j_0-1}^{\infty} \Psi_j = \{\psi_{j,k} : (j,k) \in \mathbb{I}\}, & \Psi_{j_0-1} &:= \Phi_{j_0} \\ \mathbb{I} &:= \{j_0\} \times \Delta_{j_0} \cup \bigcup_{j=j_0}^{\infty} \{j\} \times \nabla_j. \end{aligned} \tag{70}$$

Theorem 7 ([18]) *The multiscale transformations \mathbf{T}_J are well-conditioned in the sense*

$$\|\mathbf{T}_J\|, \|\mathbf{T}_J^{-1}\| = \mathcal{O}(1), \quad J \geq j_0, \tag{71}$$

if and only if the collection Ψ defined by (70) is a Riesz basis for L_2 .

A detailed construction of the *dual wavelets* $\tilde{\Psi}$, can be found in [18, 39].

Multiresolution of Sobolev spaces. Let now \mathcal{S} be a multiresolution sequence consisting of closed subspaces of H^s with the property (44) whose union is dense in H^s . The following result from [18] ensures under which conditions norm equivalences hold for the H^s -norm.

Theorem 8 *Let $\{\Phi_j\}_{j=j_0}^\infty$ and $\{\tilde{\Phi}_j\}_{j=j_0}^\infty$ be uniformly stable, refinable, biorthogonal collections. If the Jackson-type estimate*

$$\inf_{v_j \in S_j} \|v - v_j\|_{L_2} \lesssim 2^{-sj} \|v\|_{H^s}, \quad v \in H^s, \quad 0 < s \leq \bar{d}, \tag{72}$$

and the Bernstein inequality

$$\|v_j\|_{H^s} \lesssim 2^{sj} \|v_j\|_{L_2}, \quad v_j \in S_j, \quad s < \bar{t}, \tag{73}$$

hold for

$$S_j = \left\{ \begin{matrix} S(\Phi_j) \\ S(\tilde{\Phi}_j) \end{matrix} \right\} \text{ with order } \bar{d} = \left\{ \begin{matrix} d \\ \tilde{d} \end{matrix} \right\} \text{ and } \bar{t} = \left\{ \begin{matrix} t \\ \tilde{t} \end{matrix} \right\}, \tag{74}$$

then for

$$0 < \sigma := \min\{d, t\}, \quad 0 < \tilde{\sigma} := \min\{\tilde{d}, \tilde{t}\}, \tag{75}$$

we have the norm equivalence

$$\|v\|_{H^s}^2 \sim \sum_{j=j_0-1}^\infty 2^{2sj} \| \langle \tilde{\Psi}_j, v \rangle \|_{\ell_2(\nabla_j)}^2, \quad s \in (-\tilde{\sigma}, \sigma). \tag{76}$$

For many applications it suffices to have established (76) only for certain $s > 0$ for which one only needs to require (72) and (73) for $\{\Phi_j\}_{j=j_0}^\infty$. The Jackson estimates (72) of order \tilde{d} for $S(\tilde{\Phi}_j)$ imply the cancellation properties (CP) (38), see, e.g., [20].

Remark 9 When the wavelets live on $\Omega \subset \mathbb{R}^n$, (72) means that all polynomials up to order \tilde{d} are contained in $S(\tilde{\Phi}_j)$. One also says that $S(\tilde{\Phi}_j)$ is *exact* of order \tilde{d} . Because of the biorthogonality conditions, this implies that the wavelets $\psi_{j,k}$ are orthogonal to polynomials up to order \tilde{d} or have \tilde{d} th order *vanishing moments*. By Taylor expansion, this in turn yields (38).

For a summary of different constructions of biorthogonal wavelets on \mathbb{R} and bounded domains in \mathbb{R}^n , see, e.g., [39]. We display spectral condition numbers for the FWT for two different constructions of biorthogonal wavelets on the interval in Table 4. The first column denotes the finest level on which the spectral condition numbers of the FWT are computed. The next column contains the numbers for the construction of biorthogonal spline-wavelets on the interval from [22] for the case $d = 2, \tilde{d} = 4$ while the last column displays the condition numbers for a scaled version derived in [12]. We observe that the absolute numbers stay constant and low even for high levels j . We will see later in Sect. 3.3 how the transformation \mathbf{T}_j is used for preconditioning.

Table 4 Computed spectral condition numbers for the FWT on $[0, 1]$ for different constructions of biorthogonal wavelets on the interval [46]

j	$\kappa_2(\mathbf{T}_{DKU})$	$\kappa_2(\mathbf{T}_B)$	j	$\kappa_2(\mathbf{T}_{DKU})$	$\kappa_2(\mathbf{T}_B)$
4	4.743e+00	4.640e+00	11	1.097e+01	8.011e+00
5	6.221e+00	6.024e+00	12	1.103e+01	8.034e+00
6	8.154e+00	6.860e+00	13	1.106e+01	8.046e+00
7	9.473e+00	7.396e+00	14	1.107e+01	8.051e+00
8	1.023e+01	7.707e+00	15	1.108e+01	8.054e+00
9	1.064e+01	7.876e+00	16	1.108e+01	8.056e+00
10	1.086e+01	7.965e+00			

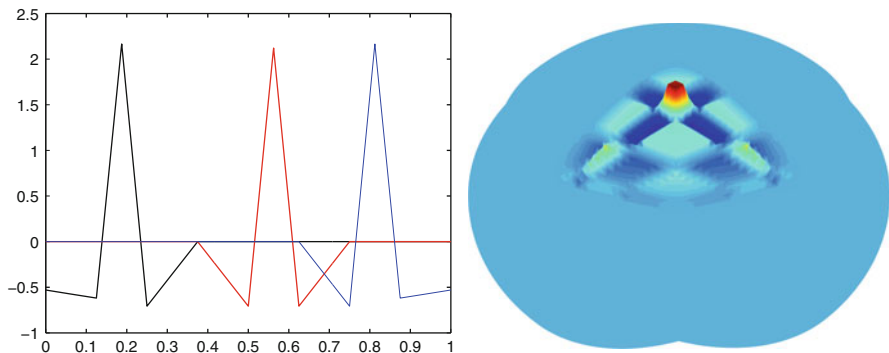


Fig. 3 Primal wavelets for $d = 2$ on $[0, 1]$ (left) and on a sphere as constructed in [40] (right)

We see in Fig. 3 some primal biorthogonal wavelets of order $d = 2$ which consist of piecewise linear B-splines. These can be employed in the construction in wavelets on manifolds [24] which were optimized and implemented to construct biorthogonal wavelet bases on the sphere in [40], see the right graphic in Fig. 3.

3.3 Elliptic Boundary Value Problems in Wavelet Coordinates

We now derive a representation of the elliptic boundary value problem (6) in terms of (initially infinite) wavelet coordinates.

Let for $H = H_0^r(\Omega)$ the collection Ψ_H be a wavelet basis with corresponding dual $\tilde{\Psi}_H$ which satisfies the properties (R), (L) and (CP) from Sect. 3.1. Expanding the solution $u = \mathbf{u}^T \Psi_H$, the right hand side $f = \mathbf{f}^T \tilde{\Psi}_H$ and recalling the definition of A in (6), the wavelet representation of the elliptic boundary value problem (6) is given by

$$\mathbf{A}\mathbf{u} = \mathbf{f} \tag{77}$$

where

$$\mathbf{A} := a(\Psi_H, \Psi_H), \quad \mathbf{f} := \langle \Psi_H, f \rangle. \tag{78}$$

Then the mapping property (5) and the Riesz basis property (R) yield the following fact.

Proposition 10 *The infinite matrix \mathbf{A} is a boundedly invertible mapping from $\ell_2 = \ell_2(\mathbb{I}_H)$ into itself, and one has*

$$\|\mathbf{v}\| \sim \|\mathbf{A}\mathbf{v}\|, \quad \mathbf{v} \in \ell_2. \tag{79}$$

The proof follows by combining the isomorphism relation (5) for the elliptic operator with the Riesz basis property (43). This entails the following consequence with respect to *preconditioning*. Let for $\mathbb{I} = \mathbb{I}_H$ the symbol Λ denote any finite subset of the index set \mathbb{I} . For the corresponding set of wavelets $\Psi_\Lambda := \{\psi_\lambda : \lambda \in \Lambda\}$, denote by $S_\Lambda := \text{span } \Psi_\Lambda$ the respective finite-dimensional subspace of H . For the wavelet representation of A in terms of Ψ_Λ , $\mathbf{A}_\Lambda := a(\Psi_\Lambda, \Psi_\Lambda)$, we then have that $\kappa_2(\mathbf{A}_\Lambda) \sim 1$ independent of Λ , on account of the ellipticity of the operator. In other words, representations of A with respect to properly scaled wavelet bases for H entail well-conditioned system matrices \mathbf{A}_Λ independent of Λ .

Fast wavelet transform. We briefly explain now in the situation of uniform refinements, i.e., when $S(\Phi_J) = S(\Psi^J)$, how the FWT \mathbf{T}_J from (64) can be used for preconditioning linear elliptic operators, together with a diagonal scaling induced by the norm equivalence (76) [21]. We recall the notation from Sect. 3.2 where the wavelet basis is in fact the (unscaled) anchor basis from Sect. 3.1. Thus, the norm equivalence (36) using the scaled wavelet basis Ψ_H is the same as (76) in the anchor basis. Recall that the norm equivalence (76) implies that every $v \in H^s$ can be expanded uniquely in terms of the Ψ and its expansion coefficients \mathbf{v} satisfy

$$\|v\|_{H^s} \sim \|\mathbf{D}^s \mathbf{v}\|_{\ell_2}$$

where \mathbf{D}^s is a diagonal matrix with entries $\mathbf{D}^s_{(j,k),(j',k')} = 2^{sj} \delta_{j,j'} \delta_{k,k'}$. Depending on the order of the elliptic operator, we have $H \subset H^r(\Omega)$ for $r = 1$ or $r = 2$. We have, therefore, already identified the diagonal (scaling) matrix \mathbf{D}_J consisting of the finite portion of the matrix $\mathbf{D} = \mathbf{D}^1$ for which $j_0 - 1 \leq j \leq J - 1$. The representation of A with respect to the (unscaled) wavelet basis Ψ^J can be expressed in terms of the FWT \mathbf{T}_J , that is,

$$\langle \Psi^J, A \Psi^J \rangle = \mathbf{T}_J^T \langle \Phi_J, A \Phi_J \rangle \mathbf{T}_J, \tag{80}$$

Table 5 Optimized spectral condition numbers of the operator \mathbf{A} using tensor products of biorthogonal wavelets on the interval for space dimensions $n = 1, 2, 3$ [12]

j	$n = 1$	$n = 2$	$n = 3$
3	22.3	9.6	18.3
4	23.9	11.8	37.1
5	25.0	14.3	39.8
6	25.7	16.0	40.9
8	26.6	18.4	
10	27.1		
12	27.3		

where Φ_J is the single-scale basis for $S(\Psi^J)$. Thus, we first set up the operator equation as in finite element settings in terms of the single-scale basis Φ_J . Applying the FWT \mathbf{T}_J together with \mathbf{D}_J yields that the operator

$$\mathbf{A}_J := \mathbf{D}_J^{-1} \mathbf{T}_J^T \langle \Phi_J, A \Phi_J \rangle \mathbf{T}_J \mathbf{D}_J^{-1} \quad (81)$$

has uniformly bounded condition numbers independent of J . This can be seen by combining the properties of A according to (5) with the norm equivalences (36) and (40).

It is known that the boundary adaptations of the generators and wavelets aggravate the absolute values of the condition numbers. Nevertheless, these constants can be substantially reduced by an operator-adapted transformation which takes into account only the coarsest discretization level and, thus, is inexpensive [12]. Numerical tests confirm that the absolute constants can further be improved by taking instead of \mathbf{D}_J^{-1} the inverse of the diagonal of $\langle \Psi^J, A \Psi^J \rangle$ for the scaling in (81) [12, 46].

In Table 5 we display the condition numbers for discretizations using the weak form of the elliptic operator $-\Delta + I$ on $(0, 1)^n$ in up to three dimensions using boundary adapted biorthogonal spline-wavelets in the case $d = 2$, $\tilde{d} = 4$ with such a scaling and additional shifts of small eigenvalues which is an inexpensive operation [12].

4 Conclusion and Outlook

The central theme of this paper was to present optimal multilevel preconditioners which enable us to reduce the spectral condition number of the system matrix to be independent of the grid spacing h . Specifically in the context of isogeometric analysis in Sect. 2, the issue arises how the spectral condition numbers depend on the polynomial degree p ; we have seen a corresponding behaviour already in Table 1. From a theoretical point of view, it is not clear yet how to estimate the dependence on p and how to remedy its influence.

Some recent results in this directions may be promising. In [29], a full geometric multigrid method was proposed for problems in isogeometric analysis with an increased number of smoothing steps to account for the dependence of the spectral condition number on the polynomial degree p . A different approach is presented with different multigrid approaches involving high-and low-order variants in [49].

I expect that the intergrid operators used in Sect. 2 will also be efficient for isogeometric collocation methods as presented in [2] although it is not clear how to prove optimality of the corresponding preconditioner.

This paper dealt with preconditioners for discretizations on uniform grids which provide best approximations for PDEs with smooth solutions. In case when the solution is not smooth, adaptivity may capture the optimal degrees of freedom when compared to a best N -term approximation. Also in an adaptive method one has to iteratively solve linear systems of equations which again requires a preconditioner. For a BPX-type or FWT preconditioner, the same principles can immediately be applied as long as the spaces generated in the adaptive process are nested as in (17). Of course, then the intergrid operators in Sect. 2.3 have to be adapted accordingly.

References

1. Y. Achdou, O. Pironneau, *Computational Methods for Option Pricing* (SIAM, Philadelphia, 2005)
2. F. Auricchio, L. Beirão da Veiga, T.J.R. Hughes, A. Reali, C. Sangalli, Isogeometric collocation methods. *Math. Models Methods Appl. Sci.* **20**(11), 2075–2107 (2010)
3. O. Axelsson, P.S. Vassilevski, Algebraic multilevel preconditioning methods I. *Numer. Math.* **56**, 157–177 (1989)
4. O. Axelsson, P.S. Vassilevski, Algebraic multilevel preconditioning methods II. *SIAM J. Numer. Anal.* **27**, 1569–1590 (1990)
5. L. Beirão da Veiga, A. Buffa, C. Sangalli, R. Vázquez, Mathematical analysis of variational isogeometric methods. *Acta Numer.* **23**, 157–287 (2014)
6. L. Beirão da Veiga, D. Cho, L.F. Pavarino, S. Scacchi, Overlapping Schwarz methods for isogeometric analysis. *SIAM J. Numer. Anal.* **50**(3), 1394–1416 (2012)
7. C. de Boor, *A Practical Guide to Splines*, rev. edn. (Springer, New York, 2001)
8. D. Braess, *Finite Elements: Theory, Fast Solvers and Applications in Solid Mechanics*, 2nd edn. (Cambridge University Press, Cambridge, 2001)
9. D. Braess, W. Hackbusch, A new convergence proof for the multigrid method including the V-cycle. *SIAM J. Numer. Anal.* **20**, 967–975 (1983)
10. J.H. Bramble, J.E. Pasciak, J. Xu, Parallel multilevel preconditioners. *Math. Comput.* **55**, 1–22 (1990)
11. A. Buffa, H. Harbrecht, A. Kunoth, G. Sangalli, BPX preconditioning for isogeometric analysis. *Comput. Methods Appl. Mech. Eng.* **265**, 63–70 (2013)
12. C. Burstedde, Fast optimised wavelet methods for control problems constrained by Elliptic PDEs, PhD Dissertation, Mathematisch-Naturwissenschaftliche Fakultät, Universität Bonn, Bonn, 2005, <http://hss.ulb.uni-bonn.de/2005/0682/0682.htm>
13. C. Canuto, A. Tabacco, K. Urban, The wavelet element method, part I: construction and analysis. *Appl. Comput. Harm. Anal.* **6**, 1–52 (1999)
14. N. Clarke, K. Parrot, Multigrid for American option pricing with stochastic volatility. *Appl. Math. Financ.* **6**, 177–195 (1999)

15. A. Cohen, *Numerical Analysis of Wavelet Methods*. Studies in Mathematics and its Applications, vol. 32 (Elsevier, Amsterdam/Boston, 2003)
16. A. Cohen, W. Dahmen, R. DeVore, Adaptive wavelet methods for elliptic operator equations—convergence rates. *Math. Comput.* **70**, 27–75 (2001)
17. A. Cohen, I. Daubechies, J.-C. Feauveau, Biorthogonal bases of compactly supported wavelets. *Commun. Pure Appl. Math.* **45**, 485–560 (1992)
18. W. Dahmen, Wavelet and multiscale methods for operator equations. *Acta Numer.* **6**, 55–228 (1997)
19. W. Dahmen, Wavelet methods for PDEs—some recent developments. *J. Comput. Appl. Math.* **128**, 133–185 (2001)
20. W. Dahmen, Multiscale and wavelet methods for operator equations, in *Multiscale Problems and Methods in Numerical Simulation*, ed. by C. Canuto. C.I.M.E. Lecture Notes in Mathematics, vol. 1825 (Springer, Heidelberg, 2003), pp. 31–96
21. W. Dahmen, A. Kunoth, Multilevel preconditioning. *Numer. Math.* **63**, 315–344 (1992)
22. W. Dahmen, A. Kunoth, K. Urban, Biorthogonal spline wavelets on the interval—stability and moment conditions. *Appl. Comput. Harm. Anal.* **6**, 132–196 (1999)
23. W. Dahmen, R. Schneider, Composite wavelet bases for operator equations. *Math. Comput.* **68**, 1533–1567 (1999)
24. W. Dahmen, R. Schneider, Wavelets on manifolds I: construction and domain decomposition. *SIAM J. Math. Anal.* **31**, 184–230 (1999)
25. I. Daubechies, Orthonormal bases of compactly supported wavelets. *Commun. Pure Appl. Math.* **41**, 909–996 (1988)
26. R.A. DeVore, A. Kunoth, *Multiscale, Nonlinear and Adaptive Approximations* (Springer, Heidelberg/New York, 2009)
27. K.P.S. Gahalaut, J.K. Kraus, S.K. Tomar, Multigrid methods for isogeometric discretization. *Comput. Methods Appl. Mech. Eng.* **253**(1), 413–425 (2013)
28. S.L. Heston, A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financ. Stud.* **6**(2), 327–343 (1993)
29. C. Hofreither, W. Zulehner, On full multigrid schemes for isogeometric analysis, NuMa-Report No. 2014–03, Johannes-Kepler-Universität Linz, May 2014, <http://www.numa.uni-linz.ac.at/Publications/>
30. M. Holtz, A. Kunoth, B-spline based monotone multigrid methods with an application to the pricing of American options, in *Multigrid, Multilevel and Multiscale Methods*, ed. by P. Wesseling, C.W. Oosterlee, P. Hemker. (Electr.) Proc. EMG, 2005
31. M. Holtz, A. Kunoth, B-spline based monotone multigrid methods. *SIAM J. Numer. Anal.* **45**(1), 1175–1199 (2007)
32. J.C. Hull, *Options, Futures and Other Derivatives* (Prentice Hall International Editions, Upper Saddle River, 2000)
33. S. Jaffard, Wavelet methods for fast resolution of elliptic problems. *SIAM J. Numer. Anal.* **29**, 965–986 (1992)
34. S.K. Kleiss, C. Pechstein, B. Jüttler, S. Tomar, IETI—iso-geometric tearing and interconnecting. *Comput. Methods Appl. Mech. Eng.* **247–248**, 201–215 (2012)
35. R. Kornhuber, *Adaptive Monotone Multigrid Methods for Nonlinear Variational Problems* (Teubner, Stuttgart, 1997)
36. R. Kraft, Adaptive and linearly independent multilevel B-splines, in *Surface Fitting and Multiresolution Methods*, ed. by A. Le Méhauté, C. Rabut, L.L. Schumaker (Vanderbilt University Press, Nashville, 1997), pp. 209–218
37. A. Kunoth, *Multilevel Preconditioning* (Shaker-Verlag, Aachen, 1994)
38. A. Kunoth, *Wavelet Methods—Elliptic Boundary Value Problems and Control Problems*. Advances in Numerical Mathematics (Teubner, Stuttgart, 2001)
39. A. Kunoth, Optimized wavelet preconditioning, in *Multiscale, Nonlinear and Adaptive Approximations*, ed. by R.A. DeVore, A. Kunoth (Springer, Heidelberg/New York, 2009), pp. 325–378
40. A. Kunoth, J. Sahner, Wavelets on manifolds: an optimized construction. *Math. Comput.* **75**, 1319–1349 (2006)

41. A. Kunoth, Chr. Schneider, K. Wiechers, Multiscale methods for the valuation of American options with stochastic volatility. *Int. J. Comput. Math.* **89**(9), 1145–1163 (2012)
42. J. Maes, A. Bultheel, A hierarchical basis preconditioner for the biharmonic equation on the sphere. *IMA J. Numer. Anal.* **26**(3), 563–583 (2006)
43. J. Maes, A. Kunoth, A. Bultheel, BPX-type preconditioners for 2nd and 4th order elliptic problems on the sphere. *SIAM J. Numer. Anal.* **45**(1), 206–222 (2007)
44. C.W. Oosterlee, On multigrid for linear complementary problems with application to American-style options. *Electron. Trans. Numer. Anal.* **15**, 165–185 (2003)
45. P. Oswald, On discrete norm estimates related to multilevel preconditioners in the finite element method, in *Constructive Theory of Functions*, ed. by K.G. Ivanov, P. Petrushev, B. Sendov. Proceedings of the International Conference Varna 1991, Bulgarian Academy of Sciences, Sofia, 1992, pp. 203–214
46. R. Pabel, Wavelet methods for PDE constrained elliptic control problems with dirichlet boundary control, Diploma Thesis, Institut für Angewandte Mathematik, Universität Bonn (2006), Shaker-Verlag Aachen (2007). doi:10.2370/236_232
47. L.L. Schumaker, *Spline Functions: Basic Theory*. Cambridge Mathematical Library, 3rd edn. (Cambridge University Press, Cambridge, 2007)
48. R. Stevenson, Adaptive wavelet methods for solving operator equations: an overview, in *Multiscale, Nonlinear and Adaptive Approximations*, ed. by R.A. DeVore, A. Kunoth (Springer, Heidelberg/New York, 2009), pp. 543–598
49. H. Sundar, G. Stadler, G. Biros, Comparison of multigrid algorithms for high-order continuous finite element discretizations, submitted, 2014. [arXiv:1402.5938]
50. K. Urban, *Wavelet Methods for Elliptic Partial Differential Equations* (Oxford University Press, Oxford, 2009)
51. A.-V. Vuong, C. Giannelli, B. Jüttler, B. Simeon, A hierarchical approach to adaptive local refinement in isogeometric analysis. *Comput. Methods Appl. Mech. Eng.* **200**(49–52), 3554–3567 (2011)
52. J. Xu, Iterative methods by space decomposition and subspace correction. *SIAM Rev.* **34**(4), 581–613 (1992)
53. J. Xu, L. Chen, R.H. Nochetto, Optimal multilevel methods for $H(\text{grad})$, $H(\text{curl})$, and $H(\text{div})$ systems on graded and unstructured grids, in *Multiscale, Nonlinear and Adaptive Approximations*, ed. by R.A. DeVore, A. Kunoth (Springer, Heidelberg/New York, 2009), pp. 599–659.
54. H. Yserentant, On the multilevel splitting of finite element spaces. *Numer. Math.* **49**, 379–412 (1986)