Optimized wavelet preconditioning

Angela Kunoth

Abstract The numerical solution of linear stationary variational problems involving elliptic partial differential operators usually requires iterative solvers on account of their problem size. Our guiding principle is to devise theoretically and practically efficient iterative solution schemes which are optimal in the number of arithmetic operations, i.e., of linear complexity in the total number of unknowns. For these algorithms, asymptotically optimal preconditioners are indispensable. This article collects the main ingredients for multilevel preconditioners based on wavelets for certain systems of elliptic PDEs with smooth solutions. Specifically, we consider problems from optimal control with distributed or Dirichlet boundary control constrained by elliptic PDEs. Moreover, the wavelet characterization of function space norms will also be used in modelling the control functional, thereby extending the range of applicability over conventional methods. The wavelet preconditioners are optimized for these PDE systems to exhibit small absolute condition numbers and consequently entail absolute low iteration numbers, as numerical experiments show.

1 Introduction

For variational systems involving linear elliptic partial differential equations (PDEs) with smooth solutions, standard finite element or finite difference discretizations on uniform grids lead to the problem to solve a large ill-conditioned system of linear equations, due to the fact that PDE operators have positive order. Any iterative solution scheme will therefore become prohibitively slow since its speed depends on the spectral condition number, and the effect becomes even worse when the grid becomes finer and the number of unknowns increases. But since solutions typically exhibit a multiscale behaviour, enhancing iterative methods by multilevel ingredi-

Angela Kunoth
Institut für Mathematik, Universität Paderborn, Warburger Str. 100, 33098 Paderborn, Germany,
e-mail: kunoth@math.uni-paderborn.de,
URL: http://www2.math.uni-paderborn.de/ags/kunoth/group/angelakunoth.html

R.A. DeVore, A. Kunoth (eds.), *Multiscale, Nonlinear and Adaptive Approximation*,
DOI 10.1007/978-3-642-03413-8_10,

ents have proved to achieve much more efficient solution schemes. Naturally, one strives for an 'optimally efficient scheme', meaning 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 [BH]. The basic idea of these schemes is to successively solve smaller versions of the linear system which can often be interpreted as discretizations with respect to coarser grids, thereby reducing the spectral condition number of the original system matrix and, hence, suggesting the term 'preconditioner'.

The search for optimal preconditioners was a major topic for numerical elliptic PDE solvers in the '80's. The goal was to better understand the ingredients which made a preconditioner optimal and, specifically, to find directly applicable versions which could be interpreted as a change of basis. With the arrival of the hierarchical basis preconditioner [Y], extending an idea of Babuška from the univariate case, a simple preconditioner became available. Although it is not optimal — the system matrix still exhibits a logarithmically growing spectral condition number in the bivariate case and exponential growth in three spatial dimensions — its simplicity still makes it popular up to now [MB]. During this time, a new methodology to derive preconditioners via space decomposition and subspace corrections was developed by Jinchao Xu [X1, X2]. The BPX preconditioner proposed first in [BPX] was numerically observed to be optimal; it is based on a weighted hierarchical generator system. With techniques from Approximation Theory, its optimality was theoretically established in [DK1, O]. Since then, its range of application has been widened extensively. For example, for second and fourth order elliptic problems on the sphere a BPX-type preconditioner has been developed and its optimality proved recently in [MKB]. The survey article by Jinchao Xu and coauthors in this volume records extensions of the BPX and of multigrid preconditioners to $H(\mathrm{grad})$, $H(\mathrm{curl})$, and $H(\mathrm{div})$ systems on adaptive and unstructured grids.

At about the same time, wavelets as a special example of a multiscale basis of $L_2(\mathbb{R})$ with compact support were constructed [Dau]. While initially mainly developed and used for signal analysis and image compression, wavelets were soon discovered to also provide optimal preconditioners in the above sense for second order elliptic boundary value problems [DK1, J]. However, the fact that one cannot really 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 numerical PDEs. It was soon realized that biorthogonal spline-wavelets as developed in [CDF] are better suited since they allow one to work with piecewise polynomials for the actual 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. On this basis, optimal conditioning of the resulting linear system of equations can be achieved by applying the Fast Wavelet Transform to a single-scale discretization on a uniform grid, together with an application of an appropriate diagonal matrix.

Nowadays, the terminology 'wavelets' is used in a more general sense that originally in [Dau]: we rather consider classes of multiscale bases with three main features:

(R) Riesz basis property for the underlying function spaces,
(L) locality of the basis functions, and
(CP) cancellation properties.

These will be detailed in Section 3.

After the initial results concerning optimal preconditioning with functions of local support in [DK1], research on using wavelets for numerically solving elliptic PDEs went into different directions. One problem was that the original constructions in [Dau, CDF] and many others were based on employing 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 $[0,1]$ and, by tensor products, on $[0,1]^n$, was provided in [DKU]. Different domain decomposition approaches yield constructions of biorthogonal wavelets on domains which can be represented as unions of parametric mappings of $[0,1]^n$ [CTU, DS2, DS3, KS], see the article by Helmut Harbrecht and Reinhold Schneider in this volume and also [U] 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 called operator–based preconditioning [Bu1, Pa].

Aside from optimal preconditioning, the built-in potential of local adaptivity for wavelets is playing a prominent role when solving stationary 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 article by Rob Stevenson in this volume.

In addition to the material in this volume, 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 [Co, D2, D3, D4].

In my article, I want to remain focussed on discretizations for smooth solutions (for which uniform grids give desired accuracy) since ideally an adaptive scheme should also perform numerically well for this case. Thus, in order to asses numerical tests, results for reference schemes on uniform grids should be available.

Another extremely useful application of the Riesz basis property (R) of wavelets concerns PDE-constrained control problems guided by elliptic boundary value problems. Here a quadratic optimization functional involving Sobolev norms of the state and the control of a system is to be minimized, subject to an elliptic PDE in variational form which couples state and control variables. In wavelet bases, the numerical evaluation of Sobolev norms even with fractional smoothness indices amounts to a multiplication with a diagonal basis and can be realized fast [Bu2].

This allows one to efficiently evaluate natural function space norms as they arise for PDE–constrained control, or different norms in the control functional more adequate for modelling purposes [Bu1, BK]. Conventional discretizations based on finite elements have concentrated here on evaluating function space norms with integer smoothness. Also for linear–quadratic elliptic control problems with non–smooth solutions, adaptive wavelets provide most efficient solution schemes. Convergence and optimal complexity estimates of respective adaptive wavelet methods were established in [DK3]. Among such optimization problems, boundary control problems where the control is exerted through essential boundary conditions, appear practically most often. Formulating the elliptic PDE as a saddle point problem by introducing Lagrange multipliers for the boundary conditions allows one to handle changing boundary controls in a flexible manner. Wavelet approaches for treating such more involved systems of elliptic PDEs in saddle point form have been investigated in [K1, K4] and numerically optimized in [Pa].

This paper is of a more introductory nature: its purpose is to collect the basic ingredients for wavelet preconditioners, apply them to (systems of) linear elliptic PDEs in variational form and provide some numerical results on their performance. Specifically, some effort will be spent on the description of nested iterative solution schemes for systems from PDE-constrained control problems.

The structure of this paper is as follows. First, in Section 2, some well–posed variational problems are compiled. The simplest example is a linear second order elliptic boundary value problem for which we derive two forms of an operator equation, once as a single equation and once as a saddle point system where nonhomogeneous boundary conditions are handled by means of Lagrange multipliers. Both formulations are then used as basic systems for PDE-constrained control problems, one with control through the right hand side and one involving a Dirichlet boundary control. In Section 3 necessary ingredients and basic properties of wavelets are assembled. In particular, Section 3.4 collects the essential construction principles for wavelets on bounded domains which do not rely on Fourier techniques, namely, multiresolution analyses of function spaces and the concept of stable completions. In Section 4 we formulate the problem classes introduced in Section 2 in wavelet coordinates and derive in particular for the control problems the resulting systems of linear equations arising from the optimality conditions. Section 5 is devoted to the iterative solution of these systems. We investigate fully iterative schemes on uniform grids and show that the resulting systems can be solved in the wavelet framework together with a nested iteration strategy with an amount of arithmetic operations which is proportional to the total number of unknowns on the finest grid. Numerical experiments on the performance of the solvers as well as on the modelling issue round off this contribution.

The following notations are used frequently. The relation $a \sim b$ always 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. Norms and inner products are indexed by the corresponding function space. $L_p(\Omega)$ are for $1 \le p \le \infty$ the usual Lebesgue spaces on a domain $\Omega \subset \mathbb{R}^n$, and $W_p^k(\Omega) \subset L_p(\Omega)$

denote for $k \in \mathbb{N}$ the Sobolev spaces of functions whose weak derivatives up to order k are bounded in $L_p(\Omega)$. For $p = 2$, we abbreviate $H^k(\Omega) = W_2^k(\Omega)$.

2 Systems of elliptic partial differential equations (PDEs)

We first formulate the classes of variational problems which will be investigated here in an abstract form.

2.1 Abstract operator systems

Let $\mathscr{H}$ be a Hilbert space with norm $\|\cdot\|_{\mathscr{H}}$ with normed dual $\mathscr{H}'$ endowed with the norm

$$\|w\|_{\mathscr{H}'} := \sup_{v \in \mathscr{H}} \frac{\langle v, w \rangle}{\|v\|_{\mathscr{H}}} \tag{1}$$

where $\langle \cdot, \cdot \rangle$ denotes the dual pairing between $\mathscr{H}$ and $\mathscr{H}'$.

Given $F \in \mathscr{H}'$, the goal is to find a solution to the operator equation

$$\mathscr{L} U = F \tag{2}$$

where $\mathscr{L} : \mathscr{H} \to \mathscr{H}'$ is a linear operator which is assumed to be a bounded bijection,

$$\|\mathscr{L} V\|_{\mathscr{H}'} \sim \|V\|_{\mathscr{H}}, \qquad V \in \mathscr{H}. \tag{3}$$

The operator equation (2) is *well-posed* since (3) implies for any given data $F \in \mathscr{H}'$ the existence and uniqueness of the solution $U \in \mathscr{H}$ which depends continuously on the data. Property (3) is also called *mapping property* of $\mathscr{L}$.

The examples that we consider will be such that $\mathscr{H}$ is a product space

$$\mathscr{H} := H_{1,0} \times \cdots \times H_{m,0}, \tag{4}$$

where each of the $H_{i,0} \subseteq H_i$ is a Hilbert space or a closed subspace of a Hilbert space H_i determined, for instance, by homogeneous boundary conditions. The spaces H_i will be Sobolev spaces living on a bounded domain $\Omega \subset \mathbb{R}^n$ or on (part of) its boundary. In view of the definition of $\mathscr{H}$, the elements $V \in \mathscr{H}$ will consist of m components $V = (v_1, \ldots, v_m)^T$ for which we define $\|V\|_{\mathscr{H}}^2 := \sum_{i=1}^m \|v_i\|_{H_i}^2$. The dual space $\mathscr{H}'$ is then endowed with the norm

$$\|W\|_{\mathscr{H}'} := \sup_{V \in \mathscr{H}} \frac{\langle V, W \rangle}{\|V\|_{\mathscr{H}}} \tag{5}$$

where $\langle V, W \rangle := \sum_{i=1}^m \langle v_i, w_i \rangle_i$ in terms of the dual pairing $\langle \cdot, \cdot \rangle_i$ between H_i and H_i'.

We will formulate four classes of problems which fit into this format. A recurring theme in the derivation of the system of operator equations (2) is the minimization of a quadratic functional.

2.2 A scalar elliptic boundary value problem

Denote by $\partial\Omega := \Gamma \cup \Gamma_N$ the boundary of Ω which is assumed to be piecewise smooth. We consider the scalar second order boundary value problem

$$\begin{aligned} -\nabla\cdot(\mathbf{a}\nabla y) + cy &= f \quad \text{in } \Omega, \\ y &= g \quad \text{on } \Gamma, \\ (\mathbf{a}\nabla y)\cdot\mathbf{n} &= 0 \quad \text{on } \Gamma_N, \end{aligned} \tag{6}$$

where $\mathbf{n} = \mathbf{n}(\mathbf{x})$ is the outward normal at $\mathbf{x} \in \Gamma$, $\mathbf{a} = \mathbf{a}(\mathbf{x}) \in \mathbb{R}^{n\times n}$ is symmetric, uniformly positive definite and bounded on Ω, and $c \in L_\infty(\Omega)$. Furthermore, f and g are some given right hand side and boundary data. With the usual definition of the bilinear form

$$a(v,w) := \int_\Omega (\mathbf{a}\nabla v\cdot\nabla w + cvw)\, d\mathbf{x}, \tag{7}$$

the weak formulation of (6) requires in the case $g \equiv 0$ to find $y \in \mathscr{H}$ where

$$\mathscr{H} := H^1_{0,\Gamma}(\Omega) := \{v \in H^1(\Omega) :\ v|_\Gamma = 0\}, \tag{8}$$

or

$$\mathscr{H} := \{v \in H^1(\Omega) : \int_\Omega v(\mathbf{x})\, d\mathbf{x} = 0\} \quad \text{when } \Gamma = \emptyset, \tag{9}$$

such that

$$a(y,v) = \langle v, f\rangle, \quad v \in \mathscr{H}. \tag{10}$$

Neumann–type boundary conditions on Γ_N are implicitly contained in the weak formulation (10), therefore called *natural boundary conditions*. In contrast, Dirichlet boundary conditions on Γ have to be posed explicitly, therefore called *essential boundary conditions*. The easiest way to achieve this for homogeneous Dirichlet boundary conditions when $g \equiv 0$ is to include them into the solution space as above in (8). In the nonhomogeneous case $g \not\equiv 0$ on Γ in (6) and $\Gamma \neq \emptyset$, one can reduce this to a problem with homogeneous boundary conditions by *homogenization* as follows. Let $w \in H^1(\Omega)$ be such that $w \equiv g$ on Γ. Then $\tilde{y} := y - w$ satisfies $a(\tilde{y},v) = a(y,v) - a(w,v) = \langle v, f\rangle - a(w,v) =: \langle v, \tilde{f}\rangle$ for all $v \in \mathscr{H}$ defined in (8), and on Γ one has $\tilde{y} = g - w \equiv 0$ yielding $\tilde{y} \in \mathscr{H}$. Therefore, it suffices to consider the weak form (10) with a perhaps modified right hand side. Another possibility which allows to treat the case $g \not\equiv 0$ explicitly is discussed in the next section.

The crucial properties are that the bilinear form defined in (7) is symmetric, continuous and elliptic on $\mathscr{H}$,

$$a(v,v) \sim \|v\|^2_{\mathscr{H}} \qquad \text{for any } v \in \mathscr{H}, \tag{11}$$

see, e.g., [B]. By Riesz' representation theorem, the bilinear form defines a linear operator $A : \mathscr{H} \to \mathscr{H}'$ by

$$\langle w, Av\rangle := a(v,w), \qquad v,w \in \mathscr{H}, \tag{12}$$

which is under the above assumptions an isomorphism,

$$c_A\|v\|_{\mathscr{H}} \;\le\; \|Av\|_{\mathscr{H}'} \;\le\; C_A\|v\|_{\mathscr{H}} \qquad \text{for any } v \in \mathscr{H}. \tag{13}$$

Relation (13) entails that given any $f \in \mathscr{H}'$, there exists a unique $y \in \mathscr{H}$ which solves the linear operator equation

$$Ay = f \qquad \text{in } \mathscr{H}' \tag{14}$$

derived from (10). This linear system where the operator defines a bounded bijection in the sense of (13) is the simplest case of a well-posed variational problem (2). In the notation from Section 2.1, we have here $m = 1$ and $\mathscr{L} = A$.

2.3 Saddle point problems involving essential boundary conditions

A particular saddle point problem derived from (6) shall be considered next. Since it is particularly appropriate to handle essential non–homogeneous Dirichlet boundary conditions, it will also be employed later in the context of boundary control problems.

Recall, e.g., from [B] that the solution $y \in \mathscr{H}$ of (10) is also the unique minimizer of the minimization problem

$$\inf_{v\in\mathscr{H}} \mathscr{J}(v), \qquad \mathscr{J}(v) := \frac{1}{2}a(v,v) - \langle v, f\rangle. \tag{15}$$

This means that y is a critical point for the first order variational derivative of $\mathscr{J}$, i.e., $\delta\mathscr{J}(y;v) = 0$. Here $\delta^s\mathscr{J}(v;w_1,\ldots,w_s)$ denotes the s–th variation of $\mathscr{J}$ at v in directions $w_1,\ldots,w_s$. In particular, for $s = 1$

$$\delta\mathscr{J}(v;w) := \lim_{t\to 0} \frac{\mathscr{J}(v+tw) - \mathscr{J}(v)}{t} \tag{16}$$

is the (Gateaux) derivative of $\mathscr{J}$ at v in direction w.

Generalizing (15) to the case of nonhomogeneous Dirichlet boundary conditions g, we want to minimize $\mathscr{J}$ over $v \in H^1(\Omega)$ subject to constraints in form of the essential boundary conditions $v = g$ on Γ. A standard technique from nonlinear optimization is to employ a *Lagrange multiplier* p to append the constraints to the optimization functional $\mathscr{J}$ defined in (15). Satisfying the constraint is guaranteed

by taking the supremum over all such Lagrange multipliers before taking the infimum. Thus, minimization subject to a constraint leads to the problem of finding a *saddle point* (y,p) of the *saddle point problem*

$$\inf_{v\in H^1(\Omega)} \sup_{q\in(H^{1/2}(\Gamma))'} \mathscr{J}(v) + \langle v-g, q\rangle_\Gamma. \tag{17}$$

The choice of the Lagrange multiplier space and the dual form $\langle\cdot,\cdot\rangle_\Gamma$ in (17) can be explained as follows. The boundary expression $v = g$ actually means taking the *trace* of $v \in H^1(\Omega)$ to $\Gamma \subseteq \partial\Omega$ which we explicitly write as $\gamma v := v|_\Gamma$. Classical trace theorems from, e.g., [Gr] state that for any $v \in H^1(\Omega)$ one looses '$\frac{1}{2}$ order of smoothness' when taking traces, therefore yielding $\gamma v \in H^{1/2}(\Gamma)$. Thus, when the data g is such that $g \in H^{1/2}(\Gamma)$, the expression in (17) involving the dual form $\langle\cdot,\cdot\rangle_\Gamma := \langle\cdot,\cdot\rangle_{H^{1/2}(\Gamma)\times(H^{1/2}(\Gamma))'}$ is well–defined, and so is the selection of the multiplier space $(H^{1/2}(\Gamma))'$. In case of Dirichlet boundary conditions on the whole boundary of Ω, i.e., the case $\Gamma \equiv \partial\Omega$, one can identify $(H^{1/2}(\Gamma))' = H^{-1/2}(\Gamma)$.

The above formulation (17) was first investigated in [Ba]. Another standard technique from optimization to handle minimization problems under constraints is to append the constraints to $\mathscr{J}(v)$ by means of a *penalty parameter* ε. For this approach, however, the system matrix depends on ε. So far, no optimal preconditioners have been established for this case so that we do not discuss this method here any further.

The method of appending essential boundary conditions by Lagrange multipliers is particularly appealing in connection with *fictitious domain methods* which may be used for problems with changing boundaries such as shape optimization problems. There one embeds the domain Ω into a larger, simple domain $\square$, and formulates (17) with respect to $H^1(\square)$ and dual form on the boundary Γ [K2]. One should note, however, that in the case that Γ is a proper subset of $\partial\Omega$, there may occur some ambiguity in the relation between the fictitious domain formulation and the corresponding strong form (6). In fact, without further assumptions, one cannot establish that the infimum of (17) with respect to $H^1(\square)$, when restricted to Ω, is the same as taking the infimum of (17) with respect to $H^1(\Omega)$. This is indeed guaranteed by using another set of Lagrangian multipliers. We currently investigate this for a problem from electrical impedance tomography in [KK].

In order to bring out the role of the trace operator, we define in addition to (7) a second bilinear form on $H^1(\Omega)\times(H^{1/2}(\Gamma))'$ by

$$b(v,q) := \int_\Gamma (\gamma v)(s)\, q(s)\, ds \tag{18}$$

so that the saddle point problem (17) may be rewritten as

$$\inf_{v\in H^1(\Omega)} \sup_{q\in(H^{1/2}(\Gamma))'} \mathscr{J}(v,q), \qquad \mathscr{J}(v,q) := \mathscr{J}(v) + b(v,q) - \langle g, q\rangle_\Gamma. \tag{19}$$

Determining the critical points of first order variations of $\mathscr{J}$, now with respect to both v and q, yields the system of equations that a saddle point (y,p) has to satisfy

$$\begin{aligned} a(y,v)+b(v,p) &= \langle v,f\rangle, && v\in H^1(\Omega),\\ b(y,q) &= \langle g,q\rangle_\Gamma, && q\in (H^{1/2}(\Gamma))'. \end{aligned} \tag{20}$$

Defining the linear operator $B: H^1(\Omega)\to H^{1/2}(\Gamma)$ and its adjoint $B': (H^{1/2}(\Gamma))'\to (H^1(\Omega))'$ by $\langle Bv,q\rangle_\Gamma = \langle v,B'q\rangle_\Gamma := b(v,q)$, this can be rewritten as a linear operator equation from $\mathscr{H} := H^1(\Omega)\times (H^{1/2}(\Gamma))'$ to $\mathscr{H}'$ as follows:
Given $(f,g)\in\mathscr{H}'$, find $(y,p)\in\mathscr{H}$ that solves

$$\begin{pmatrix} A & B'\\ B & 0\end{pmatrix}\begin{pmatrix} y\\ p\end{pmatrix} = \begin{pmatrix} f\\ g\end{pmatrix}. \tag{21}$$

It can be shown in the present context that the Lagrange multiplier can be determined by $p=-\mathbf{n}\cdot\mathbf{a}\nabla y$ which can be interpreted as *stress force* on the boundary [Ba].

We briefly discuss the properties of B representing the trace operator. Classical trace theorems from, e.g., [Gr], state that for any $f\in H^s(\Omega)$, $1/2<s<3/2$, one has

$$\|f|_\Gamma\|_{H^{s-1/2}(\Gamma)} \lesssim \|f\|_{H^s(\Omega)}. \tag{22}$$

Conversely, for every $g\in H^{s-1/2}(\Gamma)$, there exists some $f\in H^s(\Omega)$ such that $f|_\Gamma = g$ and

$$\|f\|_{H^s(\Omega)} \lesssim \|g\|_{H^{s-1/2}(\Gamma)}. \tag{23}$$

Note that the range of s extends accordingly if Γ is more regular. Estimate (22) immediately entails for $s=1$ that $B: H^1(\Omega)\to H^{1/2}(\Gamma)$ is continuous. Moreover, the second property (23) means B is surjective, i.e., $\operatorname{range} B = H^{1/2}(\Gamma)$ and $\ker B' = \{0\}$. This yields that the *inf–sup condition*

$$\inf_{q\in(H^{1/2}(\Gamma))'}\sup_{v\in H^1(\Omega)}\frac{\langle Bv,q\rangle_\Gamma}{\|v\|_{H^1(\Omega)}\,\|q\|_{(H^{1/2}(\Gamma))'}} \gtrsim 1 \tag{24}$$

is satisfied.

In the notation from Section 2.1, the system (21) is a saddle point problem on $\mathscr{H} = Y\times Q$. Thus, we identify $Y = H^1(\Omega)$ and $Q = (H^{1/2}(\Gamma))'$ and linear operators $A: Y\to Y'$ and $B: Y\to Q'$.

The abstract theory of saddle point problems states that existence and uniqueness of a solution pair $(y,p)\in\mathscr{H}$ of (21) holds if and only if A and B are continuous, A is invertible on $\ker B\subseteq Y$ and the range of B is closed in Q', see, e.g., [B, BF] and the article in this volume by Ricardo Nochetto and coauthors. The properties for B and the continuity for A have been assured above. In addition, we will always deal here with operators A which are invertible on $\ker B$ which cover the standard cases of the Laplacian ($\mathbf{a}=I$ and $c\equiv 0$) and the Helmholtz operator ($\mathbf{a}=I$ and $c=1$).

Consequently, the operator

$$\mathscr{L} := \begin{pmatrix} A & B'\\ B & 0\end{pmatrix}: \mathscr{H}\to\mathscr{H}' \tag{25}$$

is linear bijection, and one has the mapping property

$$\left\| \mathscr{L} \begin{pmatrix} v \\ q \end{pmatrix} \right\|_{\mathscr{H}'} \sim \left\| \begin{pmatrix} v \\ q \end{pmatrix} \right\|_{\mathscr{H}} \tag{26}$$

for any $(v,q) \in \mathscr{H}$ with constants depending on upper and lower bounds for A, B. Finally, the operator equation (21) is established as a well-posed variational problem in the sense of Section 2.1: for given $(f,g) \in \mathscr{H}'$, there exists a unique solution $(y,p) \in \mathscr{H} = Y \times Q$ which depends continuously on the data.

2.4 PDE-constrained control problems: Distributed control

An important class of problems where the numerical solution of systems (14) or (21) is required repeatedly are control problems with PDE-constraints. Using the notation from Section 2.2, consider as a guiding model the objective to minimize a quadratic functional of the form

$$\mathscr{J}(y,u) = \frac{1}{2}\|y - y_*\|_{\mathscr{Z}}^2 + \frac{\omega}{2}\|u\|_{\mathscr{U}}^2, \tag{27}$$

subject to the linear constraints

$$Ay = f + u \qquad \text{in } H' \tag{28}$$

where $A : H \to H'$ is defined as above in (12) satisfying (13), and $f \in H$ is given. The space H is in this subsection defined as in (8) or in (9), and we reserve the symbol $\mathscr{H}$ for a resulting product space later.

In order for a solution y of (28), the *state* of the system, to be well–defined, the problem formulation has to ensure that the unknown *control* u appearing in the right hand side of (28) is at least in H'. This can be achieved by choosing the *control space* $\mathscr{U}$ whose norm appears in (27) such that it is as least as smooth as H'. The second ingredient in the optimization functional (27) is a data fidelity term which enforces a match of the system state y to some prescribed target state y_*, measured in some norm which is typically weaker than $\|\cdot\|_H$. Thus, we require that the *observation space* $\mathscr{Z}$ and the control space $\mathscr{U}$ are such that the continuous embeddings

$$\|v\|_{H'} \lesssim \|v\|_{\mathscr{U}}, \quad v \in \mathscr{U}, \qquad \|v\|_{\mathscr{Z}} \lesssim \|v\|_H, \quad v \in H, \tag{29}$$

hold. Mostly one has investigated the simplest cases of norms which occur for $\mathscr{U} = \mathscr{Z} = L_2(\Omega)$ and which are covered by these assumptions [Li]. The parameter $\omega > 0$ balances the norms in (27), the data fidelity term and the control.

Since the control appears in all of the right hand side of (28), such control problems are called *distributed* control problems. Although their practical value is of a rather limited nature, distributed control problems help to bring out the basic mechanisms. Note that when the observed data is *compatible* in the sense that $y_* \equiv A^{-1}f$, the control problem yields the trivial control $u \equiv 0$ which implies $\mathscr{J}(y,u) \equiv 0$.

Solution schemes for the control problem (27) subject to the constraints (28) can be based on the system of operator equations derived next by the same variational principles as employed in the previous section, using a Lagrange multiplier p to enforce the constraints. Defining the Lagrangian functional

$$\mathrm{Lagr}(y,p,u) := \mathscr{J}(y,u) + \langle p, Ay - f - u\rangle \tag{30}$$

on $H \times H \times H'$, the first order necessary conditions or *Karush-Kuhn-Tucker (KKT) conditions* $\delta\,\mathrm{Lagr}(x) = 0$ for $x = p, y, u$ can be derived as

$$\begin{aligned} Ay &= f + u \\ A'p &= -S(y - y_*) \\ \omega R u &= p. \end{aligned} \tag{31}$$

Here the linear operators S and R can be interpreted as Riesz operators defined by the inner products $(\cdot,\cdot)_{\mathscr{Z}}$ and $(\cdot,\cdot)_{\mathscr{U}}$. The system (31) may be written in saddle point form as

$$\mathscr{L}V := \begin{pmatrix} \mathscr{A} & \mathscr{B}' \\ \mathscr{B} & 0 \end{pmatrix} V := \begin{pmatrix} S & 0 & A' \\ 0 & \omega R & -I \\ A & -I & 0 \end{pmatrix} \begin{pmatrix} y \\ u \\ p \end{pmatrix} = \begin{pmatrix} Sy_* \\ 0 \\ f \end{pmatrix} =: F \tag{32}$$

on $\mathscr{H} := H \times H \times H'$. Here we can also allow that $\mathscr{Z}$ in (27) is a *trace space* on part of the boundary $\partial\Omega$ as long as the corresponding condition (29) is satisfied [K3]. The class of control problems where the control is exerted through Neumann boundary conditions can also be written in this form since in this case the control still appears on the right hand side of a single operator equation of a form like (28), see [DK3]. Well-posedness of the system (32) can now be established by applying the conditions for saddle point problems stated in Section 2.3.

A few statements on the *model* of the control problem should be made. While the PDE constraints (28) that govern the system are fixed, there is in many applications some ambiguity with respect to the choice of the spaces $\mathscr{Z}$ and $\mathscr{U}$. L_2 norms are easily realized in finite element discretizations, although in some applications smoother norms for the observation $\|\cdot\|_{\mathscr{Z}}$ or for the control $\|\cdot\|_{\mathscr{U}}$ are desirable. This is the case, for instance, in temperature cooling processes where also the gradient of the temperature of a material is to be controlled. Once $\mathscr{Z}$ and $\mathscr{U}$ are fixed, there is only a single parameter ω to balance the two norms in (27). *Modelling* the objective functional is therefore an issue where more flexibility may be advantageous. Specifically in a multiscale setting, one may want to weight contributions on different scales by multiple parameters.

The wavelet setting which we describe in Section 3 allows for this flexibility. It is based on formulating the norms in the objective functional in terms of weighted wavelet coefficient sequences which are equivalent to the norms for $\mathscr{Z}$, $\mathscr{U}$ and which, in addition, support an efficient numerical implementation. Once wavelet discretizations are introduced, we formulate below control problems with such objective functionals.

2.5 PDE-constrained control problems: Dirichlet boundary control

Practically the most relevant control problems are problems with Dirichlet boundary control. They can be posed using the saddle point formulation from Section 2.3.

We consider as an illustrative guiding model the problem to minimize for some given data y_* the quadratic functional

$$\mathscr{J}(y,u) = \frac{1}{2}\|y-y_*\|_{\mathscr{Z}}^2 + \frac{\omega}{2}\|u\|_{\mathscr{U}}^2, \tag{33}$$

where, adhering to the notation in Section 2.2 the state y and the control u are coupled through the linear second order elliptic boundary value problem

$$\begin{aligned} -\nabla\cdot(\mathbf{a}\nabla y) + ky &= f \qquad \text{in } \Omega,\\ y &= u \qquad \text{on } \Gamma,\\ (\mathbf{a}\nabla y)\cdot\mathbf{n} &= 0 \qquad \text{on } \Gamma_N. \end{aligned} \tag{34}$$

The appearance of the control u as a Dirichlet boundary condition in (34) is referred to as a *Dirichlet boundary control*. In view of the treatment of essential Dirichlet boundary conditions in the context of saddle point problems derived in Section 2.3, we write the PDE constraints (34) in the operator form (21) on $Y\times Q$ where $Y = H^1(\Omega)$ and $Q = (H^{1/2}(\Gamma))'$.

The model control problem with Dirichlet boundary control then reads as follows: Minimize for given data $y_*\in\mathscr{Z}$ and $f\in Y'$ the quadratic functional

$$\mathscr{J}(y,u) = \frac{1}{2}\|y-y_*\|_{\mathscr{Z}}^2 + \frac{\omega}{2}\|u\|_{\mathscr{U}}^2 \tag{35}$$

subject to

$$\begin{pmatrix} A & B'\\ B & 0\end{pmatrix}\begin{pmatrix} y\\ p\end{pmatrix} = \begin{pmatrix} f\\ u\end{pmatrix}. \tag{36}$$

In view of the problem formulation in Section 2.4 and the discussion of the choice of the observation space $\mathscr{Z}$ and the control space, here we require analogously that $\mathscr{Z}$ and $\mathscr{U}$ are such that the continuous embeddings

$$\|v\|_{Q'} \lesssim \|v\|_{\mathscr{U}}, \quad v\in\mathscr{U}, \qquad \|v\|_{\mathscr{Z}} \lesssim \|v\|_Y, \quad v\in Y, \tag{37}$$

hold. Also the case of observations on part of the boundary $\partial\Omega$ can be taken into account [K4]. It should be mentioned that the simple choice $\mathscr{U} = L_2(\Gamma)$ which is used in many applications of Dirichlet control problems is *not* covered here. Indeed, there may arise the problem of well-posedness as follows. The constraints (34) or, in weak form (21), guarantee a unique weak solution $y\in Y = H^1(\Omega)$ *provided* that the boundary term u satisfies $u\in Q' = H^{1/2}(\Gamma)$. In the framework of control problems, this smoothness of u therefore has to be required either by the choice of $\mathscr{U}$ or by the choice of $\mathscr{Z}$ (such as $\mathscr{Z} = H^1(\Omega)$) which would assure $By\in Q'$. In the latter case, we could relax condition (37) on $\mathscr{U}$.

By variational principles, we can derive as before the first order necessary conditions for a coupled *system* of saddle point problems. Well-posedness of this system can again be established by applying the conditions for saddle point problems from Section 2.3 where the inf-sup condition for the saddle point problem (21) yields an inf-sup condition for the exterior saddle point problem of interior saddle point problems [K1].

3 Wavelets

The numerical solution of the afore-mentioned classes of problems hinges on the availability of appropriate wavelet bases for the function spaces under consideration which are all specific Hilbert spaces on the domain or on (part of) its boundary.

3.1 Basic properties

For the above classes of problems, we need to have a wavelet basis at our disposal for each occurring function space. A *wavelet basis* for a Hilbert space H is here understood as a collection of functions

$$\Psi_H := \{\psi_{H,\lambda} : \lambda \in \mathcal{I}_H\} \subset H \tag{38}$$

which are indexed by elements λ from an infinite index set $\mathcal{I}_H$. 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 dimensions, 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 $\|\cdot\|_{L_2}$. On bounded domains, the structure of the functions is essentially the same up to modifications near the boundary.

The three crucial properties that we will assume the wavelet basis to have for the sequel are the following.

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

$$v = \sum_{\lambda \in \mathcal{I}_H} v_\lambda \, \psi_{H,\lambda} =: \mathbf{v}^T \Psi_H, \quad \mathbf{v} := (v_\lambda)_{\lambda \in \mathcal{I}_H}, \tag{39}$$

and its expansion coefficients satisfy a *norm equivalence*: for any $\mathbf{v} = \{v_\lambda : \lambda \in \mathcal{I}_H\}$ one has

$$c_H \, \|\mathbf{v}\|_{\ell_2(\mathcal{I}_H)} \;\le\; \|\mathbf{v}^T \Psi_H\|_H \;\le\; C_H \, \|\mathbf{v}\|_{\ell_2(\mathcal{I}_H)}, \quad \mathbf{v} \in \ell_2(\mathcal{I}_H), \tag{40}$$

where $0 < c_H \le C_H < \infty$. This means that wavelet expansions induce *isomorphisms* between certain function spaces and sequence spaces. We write ℓ_2 norms without subscripts as $\|\cdot\| := \|\cdot\|_{\ell_2(\mathcal{I}_H)}$ when the index set is clear from the context. If the precise constants do not matter, we write the norm equivalence (40) shortly as

$$\|\mathbf{v}\| \ \sim \ \|\mathbf{v}^T \Psi_H\|_H, \quad \mathbf{v} \in \ell_2(\mathcal{I}_H). \tag{41}$$

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

$$\operatorname{diam}(\operatorname{supp} \psi_{H,\lambda}) \ \sim \ 2^{-|\lambda|}. \tag{42}$$

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

$$\langle v, \psi_{H,\lambda} \rangle \lesssim 2^{-|\lambda|(n/2-n/p+\tilde{m})} |v|_{W_p^{\tilde{m}}(\operatorname{supp} \psi_{H,\lambda})}. \tag{43}$$

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 $\tilde{S}_j := \operatorname{span}\{\tilde{\psi}_{H,\lambda} : |\lambda| < j\}$ contain all polynomials of order $\tilde{m}$. Here *dual basis* means that $\langle \psi_{H,\lambda}, \tilde{\psi}_{H,\nu} \rangle = \delta_{\lambda,\nu}$, $\lambda, \nu \in \mathcal{I}_H$.

A few remarks on these properties should be made. In (R), the norm equivalence (41) is crucial since it means complete 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 (40) may be neglected in a controllable way. Moreover, (CP) can be used to derive quasi–sparse representations of a wide class of operators, see the article by Rob Stevenson in this volume.

By duality arguments one can show that (40) is equivalent to the existence of a biorthogonal collection which is *dual* or *biorthogonal* to Ψ_H,

$$\tilde{\Psi}_H := \{\tilde{\psi}_{H,\lambda} : \lambda \in \mathcal{I}_H\} \subset H', \quad \langle \psi_{H,\lambda}, \tilde{\psi}_{H,\mu} \rangle = \delta_{\lambda,\mu}, \qquad \lambda, \mu \in \mathcal{I}_H, \tag{44}$$

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

$$C_H^{-1} \|\tilde{\mathbf{v}}\| \ \le \ \|\tilde{\mathbf{v}}^T \tilde{\Psi}_H\|_{H'} \ \le \ c_H^{-1} \|\tilde{\mathbf{v}}\|, \tag{45}$$

see [D1, D3]. Here and in the sequel 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 $\tilde{\Psi}_H = \Psi_{H'}$.

Above in (40), we have already introduced the following shorthand notation which simplifies the presentation of many terms. We will view Ψ_H both as in (38) 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 to be understood as 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* (44) can be expressed shortly as

$$\langle \Psi, \tilde{\Psi} \rangle = \mathbf{I} \tag{46}$$

with the infinite identity matrix $\mathbf{I}$.

Wavelets with the above properties can actually be obtained in the following way. In particular, this concerns a scaling depending on the regularity of the space under consideration. In our case, H will always be a Sobolev space $H^s = H^s(\Omega)$ or a closed subspace of $H^s(\Omega)$ determined by homogeneous boundary conditions, or its dual. For $s < 0$, H^s is interpreted as above as the dual of H^{-s}.

We typically obtain the wavelet basis Ψ_H for H from an *anchor basis* $\Psi = \{\psi_\lambda : \lambda \in I\!I = I\!I_H\}$ 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, for $0 < s < \gamma, \tilde{\gamma}$, respectively. Consequently, one can derive that for each $s \in (-\tilde{\gamma}, \gamma)$ the collection

$$\Psi_s := \{2^{-s|\lambda|}\psi_\lambda : \lambda \in I\!I\} =: \mathbf{D}^{-s}\Psi \tag{47}$$

is a Riesz basis for H^s [D1]. This means that there exist positive finite constants c_s, C_s such that

$$c_s \|\mathbf{v}\| \ \le\ \|\mathbf{v}^T \Psi_s\|_{H^s} \ \le\ C_s \|\mathbf{v}\|, \quad \mathbf{v} \in \ell_2(I\!I), \tag{48}$$

holds for each $s \in (-\tilde{\gamma}, \gamma)$. Such a scaling represented by a diagonal matrix $\mathbf{D}^s$ introduced in (47) will play an important role later on. The analogous expression in terms of the dual basis reads

$$\tilde{\Psi}_s := \{2^{s|\lambda|}\,\tilde{\psi}_\lambda : \lambda \in I\!I\} = \mathbf{D}^s \tilde{\Psi}, \tag{49}$$

where $\tilde{\Psi}_s$ forms a Riesz basis of H^s for $s \in (-\gamma, \tilde{\gamma})$. This entails the following fact. For $t \in (-\tilde{\gamma}, \gamma)$ the mapping

$$D^t : v = \mathbf{v}^T \Psi \mapsto (\mathbf{D}^t \mathbf{v})^T \Psi = \mathbf{v}^T \mathbf{D}^t \Psi = \sum_{\lambda \in I\!I} v_\lambda\, 2^{t|\lambda|}\, \psi_\lambda \tag{50}$$

acts as a shift operator between Sobolev scales which means that

$$\|D^t v\|_{H^s} \sim \|v\|_{H^{s+t}} \sim \|\mathbf{D}^{s+t}\mathbf{v}\|, \quad \text{if } s,\ s+t \in (-\tilde{\gamma},\gamma). \tag{51}$$

Concrete constructions of wavelet bases with the above properties for parameters $\gamma,\tilde{\gamma} \le 3/2$ on a bounded Lipschitz domain Ω can be found in [DKU, DSt]. This suffices for the above mentioned examples where the relevant Sobolev regularity indices range between -1 and 1.

3.2 Norm equivalences and Riesz maps

As we have seen, the scaling provided by $\mathbf{D}^{-s}$ is an important feature to establish norm equivalences (48) for the range $s \in (-\tilde{\gamma},\gamma)$ of Sobolev spaces H^s. However, there are several other norms which are *equivalent* to $\|\cdot\|_{H^s}$ which may later be used in the objective functional (27) in the context of control problems. This issue addresses the *mathematical model* which we briefly discuss now.

We first consider norm equivalences for the L_2 norm. Let as before Ψ be the anchor wavelet basis for L_2 for which the *Riesz operator* $\mathbf{R} = \mathbf{R}_{L_2}$ is the (infinite) Gramian matrix with respect to the inner product $(\cdot,\cdot)_{L_2}$ defined as

$$\mathbf{R} := (\Psi,\Psi)_{L_2} = \langle \Psi,\Psi\rangle. \tag{52}$$

Expanding Ψ in terms of $\tilde{\Psi}$ and recalling the duality (46), this yields

$$\mathbf{I} = \langle \Psi,\tilde{\Psi}\rangle = \big\langle \langle \Psi,\Psi\rangle\tilde{\Psi},\tilde{\Psi}\big\rangle = \mathbf{R}\langle\tilde{\Psi},\tilde{\Psi}\rangle \qquad \text{or} \qquad \mathbf{R}^{-1} = \langle\tilde{\Psi},\tilde{\Psi}\rangle. \tag{53}$$

$\mathbf{R}$ may be interpreted as the transformation matrix for the change of basis from $\tilde{\Psi}$ to Ψ, that is, $\Psi = \mathbf{R}\tilde{\Psi}$. For any $w = \mathbf{w}^T\Psi \in L_2$, we now obtain the identities

$$\|w\|_{L_2}^2 = (\mathbf{w}^T\Psi,\mathbf{w}^T\Psi)_{L_2} = \mathbf{w}^T\langle\Psi,\Psi\rangle\,\mathbf{w} = \mathbf{w}^T\mathbf{R}\mathbf{w} = \|\mathbf{R}^{1/2}\mathbf{w}\|^2 =: \|\hat{\mathbf{w}}\|^2. \tag{54}$$

Expanding w with respect to the basis $\hat{\Psi} := \mathbf{R}^{-1/2}\Psi = \mathbf{R}^{1/2}\tilde{\Psi}$, that is, $w = \hat{\mathbf{w}}^T\hat{\Psi}$, yields $\|w\|_{L_2} = \|\hat{\mathbf{w}}\|$. On the other hand, we obtain from (48) with $s = 0$

$$c_0^2\,\|\mathbf{w}\|^2 \ \le\ \|w\|_{L_2}^2 \ \le\ C_0^2\,\|\mathbf{w}\|^2. \tag{55}$$

From this we can derive the *condition number* $\kappa(\Psi)$ of the wavelet basis in terms of the extreme eigenvalues of $\mathbf{R}$ by defining

$$\kappa(\Psi) := \left(\frac{C_0}{c_0}\right)^2 = \frac{\lambda_{\max}(\mathbf{R})}{\lambda_{\min}(\mathbf{R})} = \kappa(\mathbf{R}) \sim 1, \tag{56}$$

where $\kappa(\mathbf{R})$ also denotes the spectral condition number of $\mathbf{R}$ and where the last relation is assured by the asymptotic estimate (55). However, the absolute constants will have an impact on numerical results in each individual case.

For a Hilbert space H denote by Ψ_H a wavelet basis for H satisfying (R), (L), (CP) with a corresponding dual basis $\tilde{\Psi}_H$. The (infinite) Gramian matrix with respect to the inner product $(\cdot,\cdot)_H$ inducing $\|\cdot\|_H$ which is defined by

$$\mathbf{R}_H := (\Psi_H, \Psi_H)_H \tag{57}$$

will be also called *Riesz operator*. The space L_2 is covered trivially by $\mathbf{R}_0 = \mathbf{R}$. For any function $v := \mathbf{v}^T\Psi_H \in H$ we have then the identity

$$\begin{aligned}\|v\|_H^2 = (v,v)_H &= (\mathbf{v}^T\Psi_H, \mathbf{v}^T\Psi_H)_H = \mathbf{v}^T(\Psi_H,\Psi_H)_H\,\mathbf{v}\\ &= \mathbf{v}^T\mathbf{R}_H\mathbf{v} = \|\mathbf{R}_H^{1/2}\mathbf{v}\|^2.\end{aligned} \tag{58}$$

Note that in general $\mathbf{R}_H$ may not be explicitly computable, in particular, when H is a fractional Sobolev space.

Again referring to (48), we obtain as in (56) for the more general case

$$\kappa(\Psi_s) := \left(\frac{C_s}{c_s}\right)^2 = \frac{\lambda_{\max}(\mathbf{R}_{H^s})}{\lambda_{\min}(\mathbf{R}_{H^s})} = \kappa(\mathbf{R}_{H^s}) \sim 1 \quad \text{for each } s \in (-\tilde{\gamma},\gamma). \tag{59}$$

Thus, all Riesz operators on the applicable scale of Sobolev spaces are spectrally equivalent. Moreover, comparing (59) with (56), we get

$$\frac{c_s}{C_0}\,\|\mathbf{R}^{1/2}\mathbf{v}\| \;\le\; \|\mathbf{R}_{H^s}^{1/2}\mathbf{v}\| \;\le\; \frac{C_s}{c_0}\,\|\mathbf{R}^{1/2}\mathbf{v}\|. \tag{60}$$

Of course, in practice, the constants appearing in this equation may be much sharper, as the bases for Sobolev spaces with different exponents are only obtained by a diagonal scaling which preserves much of the structure of the original basis for L_2.

We summarize these results for further reference.

Proposition 3.1. *In the above notation, we have for any $v = \mathbf{v}^T\Psi_s \in H^s$ the norm equivalences*

$$\|v\|_{H^s} = \|\mathbf{R}_{H^s}^{1/2}\mathbf{v}\| \sim \|\mathbf{R}^{1/2}\mathbf{v}\| \sim \|\mathbf{v}\| \qquad \textit{for each } s \in (-\tilde{\gamma},\gamma). \tag{61}$$

3.3 Representation of operators

A final ingredient concerns the *wavelet representation* of linear operators in terms of wavelets. Let H, V be Hilbert spaces with wavelet bases Ψ_H, Ψ_V and corresponding duals $\tilde{\Psi}_H$, $\tilde{\Psi}_V$, and suppose that $\mathscr{L} : H \to V$ is a linear operator with dual $\mathscr{L}' : V' \to H'$ defined by $\langle v, \mathscr{L}'w\rangle := \langle \mathscr{L}v, w\rangle$ for all $v \in H$, $w \in V$.

We shall make frequent use of this representation and its properties.

Remark 3.1. The wavelet representation of $\mathscr{L} : H \to V$ with respect to the bases $\Psi_H, \tilde{\Psi}_V$ of H, V', respectively, is given by

$$\mathbf{L} := \langle \tilde{\Psi}_V, \mathscr{L}\Psi_H \rangle, \quad \mathscr{L}v = (\mathbf{L}\mathbf{v})^T \Psi_V. \tag{62}$$

Thus, the expansion coefficients of $\mathscr{L}v$ in the basis that spans the range space of $\mathscr{L}$ are obtained by applying the *infinite* matrix $\mathbf{L} = \langle \tilde{\Psi}_V, \mathscr{L}\Psi_H \rangle$ to the coefficient vector of v. Moreover, boundedness of $\mathscr{L}$ implies boundedness of $\mathbf{L}$ in ℓ_2, i.e.,

$$\|\mathscr{L}v\|_V \lesssim \|v\|_H, \quad v \in H, \quad \text{implies} \quad \|\mathbf{L}\| := \sup_{\|\mathbf{v}\|_{\ell_2(I\!I_H)} \le 1} \|\mathbf{L}\mathbf{v}\|_{\ell_2(I\!I_V)} \lesssim 1. \tag{63}$$

Proof. Any image $\mathscr{L}v \in V$ can naturally be expanded with respect to Ψ_V as $\mathscr{L}v = \langle \mathscr{L}v, \tilde{\Psi}_V \rangle \Psi_V$. Expanding in addition v in the basis Ψ_H, $v = \mathbf{v}^T \Psi_H$ yields

$$\mathscr{L}v = \mathbf{v}^T \langle \mathscr{L}\Psi_H, \tilde{\Psi}_V \rangle \Psi_V = (\langle \mathscr{L}\Psi_H, \tilde{\Psi}_V \rangle^T \mathbf{v})^T \Psi_V = (\langle \tilde{\Psi}_V, \mathscr{L}\Psi_H \rangle \mathbf{v})^T \Psi_V. \tag{64}$$

As for (63), we can infer from (40) and (62) that

$$\|\mathbf{L}\mathbf{v}\|_{\ell_2(I\!I_V)} \sim \|(\mathbf{L}\mathbf{v})^T \Psi_V\|_V = \|Lv\|_V \lesssim \|v\|_H \sim \|\mathbf{v}\|_{\ell_2(I\!I_H)},$$

which confirms the claim. □

3.4 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., [D2]. Their cornerstones are *multiresolution analyses* of the function spaces under consideration 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* $\mathscr{S}$ 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,

$$S_{j_0} \subset S_{j_0+1} \subset \ldots \subset S_j \subset S_{j+1} \subset \ldots L_2, \qquad \operatorname{clos}_{L_2}\Big(\bigcup_{j=j_0}^{\infty} S_j \Big) = L_2. \tag{65}$$

The index j is the refinement level which appeared already in the elements of the index set $I\!I$ in (38), 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

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

Typically the multiresolution spaces S_j have the form

$$S_j = S(\Phi_j), \qquad \Phi_j = \{\phi_{j,k} : k \in \Delta_j\}, \tag{66}$$

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

$$\|\mathbf{c}\|_{\ell_2(\Delta_j)} \sim \|\mathbf{c}^T \Phi_j\|_{L_2}, \qquad \mathbf{c} = \{c_k\}_{k\in\Delta_j} \in \ell_2(\Delta_j), \tag{67}$$

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 (66).

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

$$\operatorname{diam}(\operatorname{supp}\phi_{j,k}) \sim 2^{-j}. \tag{68}$$

It follows from (67) that they are scaled such that

$$\|\phi_{j,k}\|_{L_2} \sim 1 \tag{69}$$

holds. It is known that nestedness (65) together with stability (67) implies the existence of matrices $\mathbf{M}_{j,0} = (m^j_{r,k})_{r\in\Delta_{j+1},k\in\Delta_j}$ such that the two-scale relation

$$\phi_{j,k} = \sum_{r\in\Delta_{j+1}} m^j_{r,k}\phi_{j+1,r}, \quad k \in \Delta_j, \tag{70}$$

is satisfied. We can essentially simplify the subsequent presentation of the material by viewing (70) as a matrix–vector equation which then attains the compact form

$$\Phi_j = \mathbf{M}^T_{j,0}\Phi_{j+1}. \tag{71}$$

Any set of functions satisfying an equation of this form, the *refinement* or *two–scale relation*, will be called *refinable*.

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 (67) that

$$\|\mathbf{M}_{j,0}\| = \mathcal{O}(1), \quad j \geq j_0, \tag{72}$$

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 $\mathscr{S}$ 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.,

$$S(\Phi_{j+1}) = S(\Phi_j) \oplus S(\Psi_j) \tag{73}$$

where

$$\Psi_j = \{\psi_{j,k} : k \in \nabla_j\}, \qquad \nabla_j := \Delta_{j+1} \setminus \Delta_j. \tag{74}$$

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 (67). Since (73) implies $S(\Psi_j) \subset S(\Phi_{j+1})$, the functions in Ψ_j must also satisfy a matrix–vector relation of the form

$$\Psi_j = \mathbf{M}_{j,1}^T \Phi_{j+1} \tag{75}$$

with a matrix $\mathbf{M}_{j,1}$ of size $(\#\Delta_{j+1}) \times (\#\nabla_j)$. Furthermore, (73) is equivalent to the fact that the linear operator composed of $\mathbf{M}_{j,0}$ and $\mathbf{M}_{j,1}$,

$$\mathbf{M}_j = (\mathbf{M}_{j,0}, \mathbf{M}_{j,1}), \tag{76}$$

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

$$\|\mathbf{M}_j\|, \|\mathbf{M}_j^{-1}\| = \mathscr{O}(1), \quad j \to \infty. \tag{77}$$

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

$$\mathbf{G}_j = \mathbf{M}_j^{-1} = \begin{pmatrix} \mathbf{G}_{j,0} \\ \mathbf{G}_{j,1} \end{pmatrix}. \tag{78}$$

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 L_2 *orthogonal wavelets* [Dau]. A systematic construction of more general $\mathbf{M}_j$, $\mathbf{G}_j$ for spline-wavelets can be found in [DKU], see also [D2] 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 (77) 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 prob-

lem 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.

The concept of stable completions has been introduced in [CDP] for which a special case is known as Sweldens' *lifting scheme*. Of course, constructions that yield compactly supported wavelets are particularly suited for computations in numerical analysis.

Combining the two–scale relations (71) and (75), one can see that $\mathbf{M}_j$ performs a change of bases in the space S_{j+1},

$$\begin{pmatrix} \Phi_j \\ \Psi_j \end{pmatrix} = \begin{pmatrix} \mathbf{M}_{j,0}^T \\ \mathbf{M}_{j,1}^T \end{pmatrix} \Phi_{j+1} = \mathbf{M}_j^T \Phi_{j+1}. \tag{79}$$

Conversely, applying the inverse of $\mathbf{M}_j$ to both sides of (79) results in the *reconstruction identity*

$$\Phi_{j+1} = \mathbf{G}_j^T \begin{pmatrix} \Phi_j \\ \Psi_j \end{pmatrix} = \mathbf{G}_{j,0}^T \Phi_j + \mathbf{G}_{j,1}^T \Psi_j. \tag{80}$$

An example of the structure of the matrices $\mathbf{M}_j$ and $\mathbf{G}_j$ is given in Figure 1.

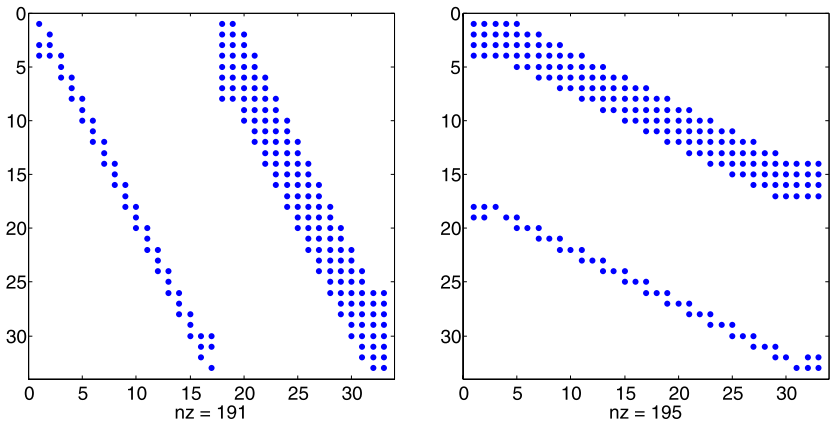

Fig. 1 Nonzero pattern of matrices $\mathbf{M}_j$ (left) and $\mathbf{G}_j$ (right) for boundary-adapted B-splines of order $d = 2$ as generators and duals of order $\tilde{d} = 4$

Fixing a *finest resolution level* J, one can repeat the decomposition (73) 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,

$$S(\Phi_J) = S(\Phi_{j_0}) \oplus \bigoplus_{j=j_0}^{J-1} S(\Psi_j). \tag{81}$$

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

$$v = (\mathbf{c}_J)^T \Phi_J = \sum_{k \in \Delta_J} c_{J,k} \phi_{J,k} \tag{82}$$

as well as in its *multi–scale form*

$$v = (\mathbf{c}_{j_0})^T \Phi_{j_0} + (\mathbf{d}_{j_0})^T \Psi_{j_0} + \cdots + (\mathbf{d}_{J-1})^T \Psi_{J-1} \tag{83}$$

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

$$\Psi^J := \Phi_{j_0} \cup \bigcup_{j=j_0}^{J-1} \Psi_j =: \bigcup_{j=j_0-1}^{J-1} \Psi_j \tag{84}$$

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*,

$$\mathbf{T}_J : \ell_2(\Delta_j) \to \ell_2(\Delta_j), \qquad \mathbf{d}^J \mapsto \mathbf{c}_J, \tag{85}$$

where

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

The previous relations (79) and (80) indicate that this will involve the matrices $\mathbf{M}_j$ and $\mathbf{G}_j$. In fact, $\mathbf{T}_J$ has the representation

$$\mathbf{T}_J = \mathbf{T}_{J,J-1} \cdots \mathbf{T}_{J,j_0}, \tag{86}$$

where each factor has the form

$$\mathbf{T}_{J,j} := \begin{pmatrix} \mathbf{M}_j & \mathbf{0} \\ \mathbf{0} & \mathbf{I}^{(\#\Delta_J - \#\Delta_{j+1})} \end{pmatrix} \in \mathbb{R}^{(\#\Delta_J) \times (\#\Delta_J)}. \tag{87}$$

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

$$\begin{array}{ccccccccc} \mathbf{c}_{j_0} & \xrightarrow{\mathbf{M}_{j_0,0}} & \mathbf{c}_{j_0+1} & \xrightarrow{\mathbf{M}_{j_0+1,0}} & \mathbf{c}_{j_0+2} & \longrightarrow \cdots & \mathbf{c}_{J-1} & \xrightarrow{\mathbf{M}_{J-1,0}} & \mathbf{c}_J \\ & \nearrow_{\mathbf{M}_{j_0,1}} & & \nearrow_{\mathbf{M}_{j_0+1,1}} & & \nearrow \cdots & & \nearrow_{\mathbf{M}_{J-1,1}} & \\ \mathbf{d}_{j_0} & & \mathbf{d}_{j_0+1} & & \mathbf{d}_{j_0+2} & & \mathbf{d}_{J-1} & & \end{array} \tag{88}$$

Accordingly, the inverse transform $\mathbf{T}_J^{-1}$ can be written also in product structure (86) in reverse order involving the matrices $\mathbf{G}_j$ as follows:

$$\mathbf{T}_J^{-1} = \mathbf{T}_{J,j_0}^{-1} \cdots \mathbf{T}_{J,J-1}^{-1}, \tag{89}$$

where each factor has the form

$$\mathbf{T}_{J,j}^{-1} := \begin{pmatrix} \mathbf{G}_j & \mathbf{0} \\ \mathbf{0} & \mathbf{I}^{(\#\Delta_J - \#\Delta_{j+1})} \end{pmatrix} \in \mathbb{R}^{(\#\Delta_J)\times(\#\Delta_J)}. \tag{90}$$

The corresponding pyramid scheme is then

$$\begin{array}{ccccccccc} & & \mathbf{G}_{J-1,0} & & \mathbf{G}_{J-2,0} & & & & \mathbf{G}_{j_0,0} \\ \mathbf{c}_J & \longrightarrow & \mathbf{c}_{J-1} & \longrightarrow & \mathbf{c}_{J-2} & \longrightarrow & \cdots & \longrightarrow & \mathbf{c}_{j_0} \\ & & \searrow & & \searrow & & \searrow \quad \cdots \quad \searrow & & \\ & & \mathbf{d}_{J-1} & & \mathbf{d}_{J-2} & & \mathbf{d}_{J-1} & & \mathbf{d}_{j_0} \end{array} \tag{91}$$

Remark 3.2. Property (77) 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 1 at the end of this section, spectral condition numbers for the Fast Wavelet Transform for different constructions of biorthogonal wavelets on the interval computed in [Pa] 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 $J \longrightarrow \infty$ in (84),

$$\begin{aligned} \Psi &:= \bigcup_{j=j_0-1}^{\infty} \Psi_j = \{\psi_{j,k} : (j,k) \in I\!I\}, \qquad \Psi_{j_0-1} := \Phi_{j_0} \\ I\!I &:= \big\{\{j_0\} \times \Delta_{j_0}\big\} \cup \bigcup_{j=j_0}^{\infty} \big\{\{j\} \times \nabla_j\big\}. \end{aligned} \tag{92}$$

The next theorem from [D1] illustrates the relation between Ψ and $\mathbf{T}_J$.

Theorem 3.1. *The multiscale transformations $\mathbf{T}_J$ are well–conditioned in the sense*

$$\|\mathbf{T}_J\|, \|\mathbf{T}_J^{-1}\| = \mathcal{O}(1), \quad J \geq j_0, \tag{93}$$

if and only if the collection Ψ defined by (92) is a Riesz basis for L_2, i.e., every $v \in L_2$ has unique expansions

$$v = \sum_{j=j_0-1}^{\infty} \langle v, \tilde{\Psi}_j \rangle \Psi_j = \sum_{j=j_0-1}^{\infty} \langle v, \Psi_j \rangle \tilde{\Psi}_j, \tag{94}$$

where $\tilde{\Psi}$ defined analogously as in (92) is also a Riesz basis for L_2 which is biorthogonal or dual to Ψ,

$$\langle \Psi, \tilde{\Psi} \rangle = \mathbf{I} \tag{95}$$

such that

$$\|v\|_{L_2} \sim \|\langle \tilde{\Psi}, v\rangle\|_{\ell_2(I)} \sim \|\langle \Psi, v\rangle\|_{\ell_2(I)}. \tag{96}$$

We briefly explain next how the functions in $\tilde{\Psi}$, denoted as *wavelets dual to* Ψ, or *dual wavelets*, can be determined. Assume that there is a second multiresolution $\tilde{\mathcal{S}}$ of L_2 satisfying (65) where

$$\tilde{S}_j = S(\tilde{\Phi}_j), \qquad \tilde{\Phi}_j = \{\tilde{\phi}_{j,k} : k \in \Delta_j\} \tag{97}$$

and $\{\tilde{\Phi}_j\}_{j=j_0}^{\infty}$ is uniformly stable in j in the sense of (67). Let the functions in $\tilde{\Phi}_j$ also have compact support satisfying (68). Furthermore, suppose that the biorthogonality conditions

$$\langle \Phi_j, \tilde{\Phi}_j\rangle = \mathbf{I} \tag{98}$$

hold. We will often refer to Φ_j as the *primal* and to $\tilde{\Phi}_j$ as the *dual generators*. The nestedness of the $\tilde{S}_j$ and the stability again implies that $\tilde{\Phi}_j$ is refinable with some matrix $\tilde{\mathbf{M}}_{j,0}$, similar to (71),

$$\tilde{\Phi}_j = \tilde{\mathbf{M}}_{j,0}^T \tilde{\Phi}_{j+1}. \tag{99}$$

The problem of determining biorthogonal wavelets now consists in finding bases $\Psi_j, \tilde{\Psi}_j$ for the complements of $S(\Phi_j)$ in $S(\Phi_{j+1})$, and of $S(\tilde{\Phi}_j)$ in $S(\tilde{\Phi}_{j+1})$, such that

$$S(\Phi_j) \perp S(\tilde{\Psi}_j), \qquad S(\tilde{\Phi}_j) \perp S(\Psi_j) \tag{100}$$

and

$$S(\Psi_j) \perp S(\tilde{\Psi}_r), \quad j \neq r, \tag{101}$$

holds. The connection between the concept of stable completions and the dual generators and wavelets is made by the following result which is a special case from [CDP].

Proposition 3.2. *Suppose that the biorthogonal collections* $\{\Phi_j\}_{j=j_0}^{\infty}$ *and* $\{\tilde{\Phi}_j\}_{j=j_0}^{\infty}$ *are both uniformly stable and refinable with refinement matrices* $\mathbf{M}_{j,0}$, $\tilde{\mathbf{M}}_{j,0}$, *i.e.,*

$$\Phi_j = \mathbf{M}_{j,0}^T \Phi_{j+1}, \qquad \tilde{\Phi}_j = \tilde{\mathbf{M}}_{j,0}^T \tilde{\Phi}_{j+1}, \tag{102}$$

and satisfy the duality condition (98). Assume that $\check{\mathbf{M}}_{j,1}$ *is any stable completion of* $\mathbf{M}_{j,0}$ *such that*

$$\check{\mathbf{M}}_j := (\mathbf{M}_{j,0}, \check{\mathbf{M}}_{j,1}) = \check{\mathbf{G}}_j^{-1} \tag{103}$$

satisfies (77).
Then

$$\mathbf{M}_{j,1} := (\mathbf{I} - \mathbf{M}_{j,0}\tilde{\mathbf{M}}_{j,0}^T)\check{\mathbf{M}}_{j,1} \tag{104}$$

is also a stable completion of $\mathbf{M}_{j,0}$, *and* $\mathbf{G}_j = \mathbf{M}_j^{-1} = (\mathbf{M}_{j,0}, \mathbf{M}_{j,1})^{-1}$ *has the form*

$$\mathbf{G}_j = \begin{pmatrix} \tilde{\mathbf{M}}_{j,0}^T \\ \check{\mathbf{G}}_{j,1} \end{pmatrix}. \tag{105}$$

Moreover, the collections of functions

$$\Psi_j := \mathbf{M}_{j,1}^T \Phi_{j+1}, \qquad \tilde{\Psi}_j := \check{\mathbf{G}}_{j,1} \tilde{\Phi}_{j+1} \tag{106}$$

form biorthogonal systems,

$$\langle \Psi_j, \tilde{\Psi}_j \rangle = \mathbf{I}, \qquad \langle \Psi_j, \tilde{\Phi}_j \rangle = \langle \Phi_j, \tilde{\Psi}_j \rangle = \mathbf{0}, \tag{107}$$

so that

$$S(\Psi_j) \perp S(\tilde{\Psi}_r), \quad j \neq r, \qquad S(\Phi_j) \perp S(\tilde{\Psi}_j), \quad S(\tilde{\Phi}_j) \perp S(\Psi_j). \tag{108}$$

In particular, the relations (98), (107) imply that the collections

$$\Psi = \bigcup_{j=j_0-1}^{\infty} \Psi_j, \qquad \tilde{\Psi} := \bigcup_{j=j_0-1}^{\infty} \tilde{\Psi}_j := \tilde{\Phi}_{j_0} \cup \bigcup_{j=j_0}^{\infty} \tilde{\Psi}_j \tag{109}$$

are biorthogonal,

$$\langle \Psi, \tilde{\Psi} \rangle = \mathbf{I}. \tag{110}$$

Remark 3.3. Note that the properties needed in addition to (110) to ensure (96) are neither properties of the complements nor of their bases $\Psi, \tilde{\Psi}$ but of the multiresolution sequences $\mathscr{S}$ and $\tilde{\mathscr{S}}$. These can be phrased as approximation and regularity properties and appear in Theorem 3.2.

We briefly recall yet another useful point of view. The operators

$$\begin{aligned} P_j v &:= \langle v, \tilde{\Phi}_j \rangle \Phi_j = \langle v, \tilde{\Psi}^j \rangle \Psi^j = \langle v, \tilde{\Phi}_{j_0} \rangle \Phi_{j_0} + \sum_{r=j_0}^{j-1} \langle v, \tilde{\Psi}_r \rangle \Psi_r \\ P'_j v &:= \langle v, \Phi_j \rangle \tilde{\Phi}_j = \langle v, \Psi^j \rangle \tilde{\Psi}^j = \langle v, \Phi_{j_0} \rangle \tilde{\Phi}_{j_0} + \sum_{r=j_0}^{j-1} \langle v, \Psi_r \rangle \tilde{\Psi}_r \end{aligned} \tag{111}$$

are projectors onto

$$S(\Phi_j) = S(\Psi^j) \qquad \text{and} \qquad S(\tilde{\Phi}_j) = S(\tilde{\Psi}^j) \tag{112}$$

respectively, which satisfy

$$P_r P_j = P_r, \quad P'_r P'_j = P'_r, \qquad r \leq j. \tag{113}$$

Remark 3.4. Let $\{\Phi_j\}_{j=j_0}^{\infty}$ be uniformly stable. The P_j defined by (111) are uniformly bounded if and only if $\{\tilde{\Phi}_j\}_{j=j_0}^{\infty}$ is also uniformly stable. Moreover, the P_j satisfy (113) if and only if the $\tilde{\Phi}_j$ are refinable as well. Note that then (98) implies

$$\mathbf{M}_{j,0}^T \tilde{\mathbf{M}}_{j,0} = \mathbf{I}. \tag{114}$$

In terms of the projectors, the uniform stability of the complement bases Ψ_j, $\tilde{\Psi}_j$ means that

$$\|(P_{j+1}-P_j)v\|_{L_2} \sim \|\langle \tilde{\Psi}_j, v\rangle\|_{\ell_2(\nabla_j)}, \quad \|(P'_{j+1}-P'_j)v\|_{L_2} \sim \|\langle \Psi_j, v\rangle\|_{\ell_2(\nabla_j)}, \quad (115)$$

so that the L_2 norm equivalence (96) is equivalent to

$$\|v\|_{L_2}^2 \sim \sum_{j=j_0}^{\infty} \|(P_j - P_{j-1})v\|_{L_2}^2 \sim \sum_{j=j_0}^{\infty} \|(P'_j - P'_{j-1})v\|_{L_2}^2 \quad (116)$$

for any $v \in L_2$, where $P_{j_0-1} = P'_{j_0-1} := 0$.

The whole concept derived so far lives from both Φ_j *and* $\tilde{\Phi}_j$. It should be pointed out that in the algorithms one actually does not need $\tilde{\Phi}_j$ explicitly for computations.

We recall next results that guarantee norm equivalences of the type (40) for Sobolev spaces.

Multiresolution of Sobolev spaces. Let now $\mathscr{S}$ be a multiresolution sequence consisting of closed subspaces of H^s with the property (65) whose union is dense in H^s. The following result from [D1] ensures under which conditions norm equivalences hold for the H^s–norm.

Theorem 3.2. *Let $\{\Phi_j\}_{j=j_0}^{\infty}$ and $\{\tilde{\Phi}_j\}_{j=j_0}^{\infty}$ be uniformly stable, refinable, biorthogonal collections and let the $P_j : H^s \to S(\Phi_j)$ be defined by (111).*
If the Jackson-type estimate

$$\inf_{v_j \in S_j} \|v - v_j\|_{L_2} \lesssim 2^{-sj}\|v\|_{H^s}, \quad v \in H^s,\ 0 < s \le \bar{d}, \quad (117)$$

and the Bernstein inequality

$$\|v_j\|_{H^s} \lesssim 2^{sj}\|v_j\|_{L_2}, \quad v_j \in S_j,\ s < \bar{t}, \quad (118)$$

hold for

$$S_j = \left\{ \begin{matrix} S(\Phi_j) \\ S(\tilde{\Phi}_j) \end{matrix} \right\} \text{ with order } \bar{d} = \left\{ \begin{matrix} d \\ \tilde{d} \end{matrix} \right\} \text{ and } \bar{t} = \left\{ \begin{matrix} t \\ \tilde{t} \end{matrix} \right\}, \quad (119)$$

then for

$$0 < \sigma := \min\{d, t\}, \qquad 0 < \tilde{\sigma} := \min\{\tilde{d}, \tilde{t}\}, \quad (120)$$

one has

$$\|v\|_{H^s}^2 \sim \sum_{j=j_0}^{\infty} 2^{2sj}\|(P_j - P_{j-1})v\|_{L_2}^2, \quad s \in (-\tilde{\sigma}, \sigma). \quad (121)$$

Recall that we always write $H^s = (H^{-s})'$ for $s < 0$.

The regularity of $\mathscr{S}$ and $\tilde{\mathscr{S}}$ is characterized by

$$t := \sup\{s : S(\Phi_j) \subset H^s,\ j \ge j_0\}, \qquad \tilde{t} := \sup\{s : S(\tilde{\Phi}_j) \subset H^s,\ j \ge j_0\} \quad (122)$$

Recalling the representation (115), we can immediately derive the following fact.

Corollary 3.1. *Suppose that the assumptions in Theorem 3.2 hold. Then we have the norm equivalence*

$$\|v\|_{H^s}^2 \sim \sum_{j=j_0-1}^{\infty} 2^{2sj} \|\langle \tilde{\Psi}_j, v\rangle\|_{\ell_2(\nabla_j)}^2, \quad s \in (-\tilde{\sigma}, \sigma). \tag{123}$$

In particular for $s = 0$ the Riesz basis property of the Ψ, $\tilde{\Psi}$ relative to L_2(96) is recovered. For many applications it suffices to have (121) or (123) only for certain $s > 0$ for which one only needs to require (117) and (118) for $\{\Phi_j\}_{j=j_0}^{\infty}$. The Jackson estimates (117) of order $\tilde{d}$ for $S(\tilde{\Phi}_j)$ imply the cancellation properties (CP) (43), see, e.g., [D4].

Remark 3.5. When the wavelets live on $\Omega \subset \mathbb{R}^n$, (117) 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 $\tilde{d}$. On account of (95), 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 (43).

The following generalizations of the discrete norms (116) are useful. Let for $s \in \mathbb{R}$

$$|||v|||_s := \left(\sum_{j=j_0}^{\infty} 2^{2sj} \|(P_j - P_{j-1})v\|_{L_2}^2 \right)^{1/2} \tag{124}$$

which by the relations (115) is also equivalent to

$$|v|_s := \left(\sum_{j=j_0-1}^{\infty} 2^{2sj} \|\langle \tilde{\Psi}_j, v\rangle\|_{\ell_2(\nabla_j)}^2 \right)^{1/2}. \tag{125}$$

In this notation, (121) and (123) read

$$\|v\|_{H^s} \sim |||v|||_s \sim |v|_s. \tag{126}$$

In terms of such discrete norms, Jackson and Bernstein estimates hold with constants equal to one.

Lemma 3.1. *[K1] Let $\{\Phi_j\}_{j=j_0}^{\infty}$ and $\{\tilde{\Phi}_j\}_{j=j_0}^{\infty}$ be uniformly stable, refinable, biorthogonal collections and let the P_j be defined by (111). Then the estimates*

$$|v - P_j v|_{s'} \le 2^{-(j+1)(s-s')} |v|_s, \qquad v \in H^s, \ s' \le s \le d, \tag{127}$$

and

$$|v_j|_s \le 2^{j(s-s')} |v_j|_{s'}, \qquad v_j \in S(\Phi_j), \ s' \le s \le d, \tag{128}$$

are valid, and correspondingly for the dual side.

The same results hold for the norm $|||\cdot|||$ defined in (124).

Reverse Cauchy–Schwarz Inequalities. The biorthogonality condition (98) implies together with direct and inverse estimates the following reverse Cauchy–Schwarz inequalities for finite–dimensional spaces [DK2]. This is one essential ingredient in proving a sufficient condition for satisfying the LBB condition in Section 4.2.

Lemma 3.2. *Let the assumptions in Theorem 3.2 be valid such that the norm equivalence (121) holds for $(-\tilde\sigma,\sigma)$ with $\sigma,\tilde\sigma$ defined in (120). Then for any $v\in S(\Phi_j)$ there exists some $\tilde v^*=\tilde v^*(v)\in S(\tilde\Phi_j)$ such that*

$$\|v\|_{H^s}\,\|\tilde v^*\|_{H^{-s}} \lesssim \langle v,\tilde v^*\rangle \tag{129}$$

for any $0\le s<\min(\sigma,\tilde\sigma)$.

The proof of this result given in [DK2] for $s=1/2$ in terms of the projectors P_j defined in (111) and corresponding duals P_j' immediately carries over to more general s. Recalling the representation (112) in terms of wavelets, the reverse Cauchy inequality (129) attains the following sharp form.

Lemma 3.3. *[K1] Let the assumptions of Lemma 3.1 hold. Then for every $v\in S(\Phi_j)$ there exists some $\tilde v^*=\tilde v^*(v)\in S(\tilde\Phi_j)$ such that*

$$|v|_s\,|\tilde v^*|_{-s} = \langle v,\tilde v^*\rangle \tag{130}$$

for any $0\le s\le\min(\sigma,\tilde\sigma)$.

Proof. Every $v\in S(\Phi_j)$ can be written as

$$v=\sum_{r=j_0-1}^{j-1}2^{sr}\sum_{k\in\nabla_r}v_{r,k}\psi_{r,k}.$$

Setting now

$$\tilde v^*:=\sum_{r=j_0-1}^{j-1}2^{-sr}\sum_{k\in\nabla_r}v_{r,k}\tilde\psi_{r,k}$$

with the same coefficients $v_{j,k}$, the definition of $|\cdot|_s$ yields by biorthogonality (110)

$$|v|_s\,|\tilde v^*|_{-s} = \sum_{r=j_0-1}^{j-1}\sum_{k\in\nabla_r}|v_{j,k}|^2.$$

Combining this with the observation

$$\langle v,\tilde v^*\rangle = \sum_{r=j_0-1}^{j-1}\sum_{k\in\nabla_r}|v_{j,k}|^2$$

confirms (130). □

Remark 3.6. The previous proof reveals that the identity (130) is also true for elements from infinite–dimensional spaces H^s and $(H^s)'$ for which Ψ and $\tilde\Psi$ are Riesz bases.

Biorthogonal wavelets on $\mathbb{R}$. The construction of biorthogonal spline-wavelets on $\mathbb{R}$ from [CDF] for $L_2=L_2(\mathbb{R})$ employs the multiresolution framework introduced

at the beginning of this section. There the $\phi_{j,k}$ are generated through the dilates and translates of a single function $\phi \in L_2$,

$$\phi_{j,k} = 2^{j/2}\phi(2^j \cdot -k). \tag{131}$$

This corresponds to the idea of a *uniform* virtual underlying grid, explaining the terminology *uniform refinements*. B–Splines on uniform grids are known to satisfy refinement relations (70) in addition to being compactly supported and having L_2–stable integer translates. For computations, they have the additional advantage that they can be expressed as piecewise polynomials. In the context of variational formulations for second order boundary value problems, a well–used example are the nodal finite elements $\phi_{j,k}$ generated by the cardinal B–Spline of order two, i.e., the piecewise linear continuous function commonly called the 'hat function'. For cardinal B–Splines as generators, a whole class of dual generators $\tilde{\phi}_{j,k}$ (of arbitrary smoothness at the expense of larger supports) can be constructed which are also generated by one single function $\tilde{\phi}$ through translates and dilates. By Fourier techniques, one can construct from $\phi, \tilde{\phi}$ then a pair of biorthogonal wavelets $\psi, \tilde{\psi}$ whose dilates and translates built as in (131) constitute Riesz bases for $L_2(\mathbb{R})$.

By taking tensor products of these functions, of course, one can generate biorthogonal wavelet bases for $L_2(\mathbb{R}^n)$.

Biorthogonal wavelets on domains. Some constructions that exist by now have as a core ingredient tensor products of one-dimensional wavelets on an *interval* derived from the biorthogonal wavelets from [CDF] on $\mathbb{R}$. On finite intervals in $\mathbb{R}$, the corresponding constructions are usually based on keeping the elements of $\Phi_j, \tilde{\Phi}_j$ supported *inside* the interval while modifying those translates overlapping the end points of the interval so as to preserve a desired degree of polynomial exactness. A general detailed construction satisfying all these requirements has been proposed in [DKU]. Here just the main ideas for constructing a biorthogonal pair $\Phi_j, \tilde{\Phi}_j$ and corresponding wavelets satisfying the above requirements are sketched, where we apply the techniques derived at the beginning of this section.

We start out with those functions from two collections of biorthogonal generators $\Phi_j^{\mathbb{R}}, \tilde{\Phi}_j^{\mathbb{R}}$ for some fixed $j \geq j_0$ living on the whole real line whose support has nonempty intersection with the interval $(0,1)$. In order to treat the boundary effects separately, we assumed that the coarsest resolution level j_0 is large enough so that, in view of (68), functions overlapping one end of the interval vanish at the other. One then leaves as many functions from the collection $\Phi_j^{\mathbb{R}}, \tilde{\Phi}_j^{\mathbb{R}}$ living in the interior of the interval untouched and modifies only those near the interval ends. Note that keeping just the restrictions to the interval of those translates overlapping the end points would destroy stability (and also the cardinality of the primal and dual basis functions living on $(0,1)$ since their supports do not have the same size). Therefore, modifications at the end points are necessary; also, just discarding them from the collections (66), (97) would produce an error near the end points. The basic idea is essentially the same for all constructions of orthogonal and biorthogonal wavelets on $\mathbb{R}$ adapted to an interval. Namely, one takes *fixed* linear combinations of all functions in $\Phi_j^{\mathbb{R}}, \tilde{\Phi}_j^{\mathbb{R}}$ living near the ends of the interval in such a way that monomials

up to the exactness order are reproduced there and such that the generator bases have the same cardinality. Because of the boundary modifications, the collections of generators are there no longer biorthogonal. However, one can show in the case of cardinal B–Splines as primal generators (which is a widely used class for numerical analysis) that biorthogonalization is indeed possible. This yields collections denoted by $\Phi_j^{(0,1)}, \tilde{\Phi}_j^{(0,1)}$ which then satisfy (98) on $(0,1)$ and all assumptions required in Proposition 3.2.

For the construction of corresponding wavelets, first an *initial* stable completion $\check{\mathbf{M}}_{j,1}$ is computed by applying Gaussian eliminations to factor $\mathbf{M}_{j,0}$ and then to find a uniformly stable inverse of $\check{\mathbf{M}}_j$. Here we exploit that for cardinal B–Splines as generators the refinement matrices $\mathbf{M}_{j,0}$ are totally positive. Thus, they can be stably decomposed by Gaussian elimination without pivoting. Application of Proposition 3.2 then gives the corresponding biorthogonal wavelets $\Psi_j^{(0,1)}, \tilde{\Psi}_j^{(0,1)}$ on $(0,1)$ which satisfy the requirements in Corollary 3.1. It turns out that these wavelets coincide in the interior of the interval again with those on all of $\mathbb{R}$ from [CDF]. An example of the primal wavelets for $d = 2$ generated by piecewise linear continuous functions is displayed in Figure 2 on the left.

After constructing these basic versions, one can then perform local transformations near the ends of the interval in order to improve the condition or L_2 stability constants, see [Bu1, Pa] for corresponding results and numerical examples.

We display spectral condition numbers for the FWT for two different constructions of biorthogonal wavelets on the interval in Table 1. 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 [DKU] for the case $d = 2, \tilde{d} = 4$ while the last column displays the condition numbers for a scaled version derived in [Bu1]. We observe that the absolute numbers stay constant and low even for high levels j. We will see later in Section 4.1 how the transformation $\mathbf{T}_J$ is used for preconditioning.

j	$\kappa_2(\mathbf{T}_{\mathrm{DKU}})$	$\kappa_2(\mathbf{T}_{\mathrm{B}})$
4	4.743e+00	4.640e+00
5	6.221e+00	6.024e+00
6	8.154e+00	6.860e+00
7	9.473e+00	7.396e+00
8	1.023e+01	7.707e+00
9	1.064e+01	7.876e+00
10	1.086e+01	7.965e+00

j	$\kappa_2(\mathbf{T}_{\mathrm{DKU}})$	$\kappa_2(\mathbf{T}_{\mathrm{B}})$
11	1.097e+01	8.011e+00
12	1.103e+01	8.034e+00
13	1.106e+01	8.046e+00
14	1.107e+01	8.051e+00
15	1.108e+01	8.054e+00
16	1.108e+01	8.056e+00

Table 1 Computed spectral condition numbers for the Fast Wavelet Transform on $[0,1]$ for different constructions of biorthogonal wavelets on the interval [Pa]

Along these lines, also biorthogonal generators and wavelets with homogeneous (Dirichlet) boundary conditions can be constructed. Since the $\Phi_j^{(0,1)}$ are locally near the boundary monomials which all vanish at $0,1$ except for one, removing the one from $\Phi_j^{(0,1)}$ which corresponds to the constant function produces a collection of generators with homogeneous boundary conditions at $0,1$. In order for the moment conditions (43) still to hold for the Ψ_j, the dual generators have to have *complementary* boundary conditions. A corresponding construction has been carried out in [DS1] and implemented in [Bu1]. Homogeneous boundary conditions of higher order can be generated accordingly.

By taking tensor products of the wavelets on $(0,1)$, in this manner biorthogonal wavelets for Sobolev spaces on $(0,1)^n$ with or without homogeneous boundary conditions are obtained. This construction can be further extended to any other domain or manifold which is the image of a regular parametric mapping of the unit cube. Some results on the construction of wavelets on manifolds are summarized in [D3]. There are essentially two approaches. The first idea is based on domain decomposition and consists in 'glueing' generators across interelement boundaries, see, e.g., [CTU, DS2]. These approaches all have in common that the norm equivalences (123) for $H^s = H^s(\Gamma)$ can be shown to hold only for the range $-1/2 < s < 3/2$, due to the fact that duality arguments apply only for this range because of the nature of a modified inner product to which biorthogonality refers. The other approach which overcomes the above limitations on the ranges for which the norm equivalences hold has been developed in [DS3] based on previous characterizations of function spaces as Cartesian products from [CF]. The construction in [DS3] has been optimized and implemented to construct biorthogonal wavelet bases on the sphere in [KS], see the right graphic in Figure 2. More on such constructions for boundary integral operators can be found in the article by Helmut Harbrecht and Reinhold Schneider in this volume.

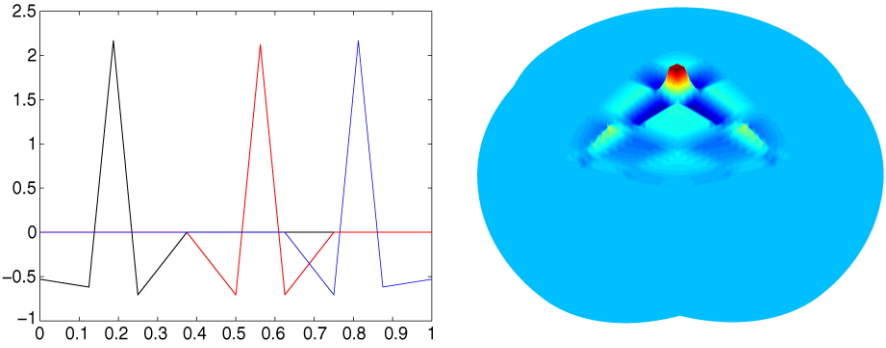

Fig. 2 Primal wavelets for $d = 2$ on $[0,1]$ (left) and on a sphere as constructed in [KS] (right)

Of course, there are also different attempts to construct wavelet bases with the above properties without using tensor products. A construction of biorthogonal spline-wavelets on triangles introduced by [Stv] has been implemented in two spa-

tial dimensions with an application to the numerical solution of a linear second order elliptic boundary value problem in [Kr].

4 Problems in wavelet coordinates

4.1 Elliptic boundary value problems

We now derive a representation of the elliptic boundary value problem from Section 2.2 in terms of (infinite) wavelet coordinates.

Let for $\mathscr{H}$ given by (8) or (9) $\Psi_{\mathscr{H}}$ be a wavelet basis with corresponding dual $\tilde{\Psi}_{\mathscr{H}}$ which satisfies the properties (R), (L) and (CP) from Section 3.1. Following the recipe from Section 3.3, expanding $y = \mathbf{y}^T \Psi_{\mathscr{H}}$, $f = \mathbf{f}^T \tilde{\Psi}_{\mathscr{H}}$ and recalling (12), the wavelet representation of the elliptic boundary value problem (14) is given by

$$\mathbf{A}\mathbf{y} = \mathbf{f} \tag{132}$$

where

$$\mathbf{A} := a(\Psi_{\mathscr{H}}, \Psi_{\mathscr{H}}), \qquad \mathbf{f} := \langle \Psi_{\mathscr{H}}, f \rangle. \tag{133}$$

Then the mapping property (13) and the Riesz basis property (R) yield the following fact.

Proposition 4.1. *The infinite matrix* $\mathbf{A}$ *is a boundedly invertible mapping from* $\ell_2 = \ell_2(\mathbb{I}_{\mathscr{H}})$ *into itself, and there exists finite positive constants* $c_{\mathbf{A}} \le C_{\mathbf{A}}$ *such that*

$$c_{\mathbf{A}} \|\mathbf{v}\| \le \|\mathbf{A}\mathbf{v}\| \le C_{\mathbf{A}} \|\mathbf{v}\|, \qquad \mathbf{v} \in \ell_2(\mathbb{I}_{\mathscr{H}}). \tag{134}$$

Proof. For any $v \in \mathscr{H}$ with coefficient vector $\mathbf{v} \in \ell_2$, we have by the lower estimates in (40), (13) and the upper inequality in (45), respectively,

$$\|\mathbf{v}\| \le c_{\mathscr{H}}^{-1} \|v\|_{\mathscr{H}} \le c_{\mathscr{H}}^{-1} c_A^{-1} \|Av\|_{\mathscr{H}'} = c_{\mathscr{H}}^{-1} c_A^{-1} \|(\mathbf{A}\mathbf{v})^T \tilde{\Psi}_{\mathscr{H}}\|_{\mathscr{H}'} \le c_{\mathscr{H}}^{-2} c_A^{-1} \|\mathbf{A}\mathbf{v}\|$$

where we have used the wavelet representation (62) for A. Likewise, the converse estimate

$$\|\mathbf{A}\mathbf{v}\| \le C_{\mathscr{H}} \|Av\|_{\mathscr{H}'} \le C_{\mathscr{H}} C_A \|v\|_{\mathscr{H}} \le C_{\mathscr{H}}^2 C_A \|\mathbf{v}\|$$

follows by the lower inequality in (45) and the upper estimates in (13) and (40). The constants appearing in (134) are therefore identified as $c_{\mathbf{A}} := c_{\mathscr{H}}^2 c_A$ and $C_{\mathbf{A}} := c_{\mathscr{H}}^2 C_A$. □

In the present situation where $\mathbf{A}$ is defined via the elliptic bilinear form $a(\cdot,\cdot)$, Proposition 4.1 entails the following result with respect to *preconditioning*. Let for $\mathbb{I} = \mathbb{I}_{\mathscr{H}}$ 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 $S_\Lambda := \operatorname{span}\Psi_\Lambda$ the respective finite-dimensional subspace of $\mathscr{H}$. For the wavelet representation of A in terms of Ψ_Λ,

$$\mathbf{A}_\Lambda := a(\Psi_\Lambda, \Psi_\Lambda), \tag{135}$$

we obtain the following result.

Proposition 4.2. *If $a(\cdot,\cdot)$ is $\mathscr{H}$-elliptic according to (11), the finite matrix $\mathbf{A}_\Lambda$ is symmetric positive definite and its spectral condition number is bounded uniformly in Λ, i.e.,*

$$\kappa_2(\mathbf{A}_\Lambda) \le \frac{C_{\mathbf{A}}}{c_{\mathbf{A}}}, \tag{136}$$

where $c_{\mathbf{A}}, C_{\mathbf{A}}$ are the constants from (134).

Proof. Clearly, since $\mathbf{A}_\Lambda$ is just a finite section of $\mathbf{A}$, we have $\|\mathbf{A}_\Lambda\| \le \|\mathbf{A}\|$. On the other hand, by assumption, $a(\cdot,\cdot)$ is $\mathscr{H}$-elliptic which entails that $a(\cdot,\cdot)$ is also elliptic on every finite subspace $S_\Lambda \subset \mathscr{H}$. Thus, we infer $\|\mathbf{A}_\Lambda^{-1}\| \le \|\mathbf{A}^{-1}\|$, and we have

$$c_{\mathbf{A}}\|\mathbf{v}_\Lambda\| \le \|\mathbf{A}_\Lambda \mathbf{v}_\Lambda\| \le C_{\mathbf{A}}\|\mathbf{v}_\Lambda\|, \qquad \mathbf{v}_\Lambda \in S_\Lambda. \tag{137}$$

Together with the definition $\kappa_2(\mathbf{A}_\Lambda) := \|\mathbf{A}_\Lambda\|\,\|\mathbf{A}_\Lambda^{-1}\|$ we obtain the claimed estimate. □

In other words, representations of A with respect to properly scaled wavelet bases for $\mathscr{H}$ entail well-conditioned system matrices $\mathbf{A}_\Lambda$ independent of Λ. This in turn means that the convergence speed of an iterative solver applied to the corresponding finite system

$$\mathbf{A}_\Lambda \mathbf{y}_\Lambda = \mathbf{f}_\Lambda \tag{138}$$

does not deteriorate as $|\Lambda| \to \infty$.

In summary, ellipticity implies stability of the Galerkin discretizations for any set $\Lambda \subset \mathbb{I}$. This is not automatically the case for any finite versions of the saddle point problems, as we will see in Section 4.2.

Fast wavelet transform. We briefly summarize how in the situation of uniform refinements, i.e., when $S(\Phi_J) = S(\Psi^J)$, the Fast Wavelet Transformation (FWT) $\mathbf{T}_J$ can be used for preconditioning linear elliptic operators, together with a a diagonal scaling induced by the norm equivalence (123) [DK1]. We recall the notation from Section 3.4 where the wavelet basis is in fact the (unscaled) anchor basis from Section 3.1. Thus, the norm equivalence (40) using the scaled wavelet basis Ψ_H is the same as (123) in the anchor basis. Recall that the norm equivalence (123) 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 $\mathbf{D}^s_{(j,k),(j',k')} = 2^{sj}\delta_{j,j'}\delta_{k,k'}$. For $\mathscr{H} \subset H^1(\Omega)$, the case $s = 1$ is relevant.

In a stable Galerkin scheme for (10) with respect to $S(\Psi^J) = S(\Psi_\Lambda)$, 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 $j_0 - 1 \le j \le J - 1$. The representation of A with respect to the (unscaled) wavelet basis Ψ^J can be expressed in terms of the Fast Wavelet Transform $\mathbf{T}_J$, that is,

$$\langle \Psi^J, A\Psi^J \rangle \;=\; \mathbf{T}_J^T \, \langle \Phi_J, A\Phi_J \rangle \, \mathbf{T}_J, \tag{139}$$

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 Fast Wavelet Transform $\mathbf{T}_J$ together with $\mathbf{D}_J$ yields that the operator

$$\mathbf{A}_J \;:=\; \mathbf{D}_J^{-1} \mathbf{T}_J^T \, \langle \Phi_J, A\Phi_J \rangle \, \mathbf{T}_J \mathbf{D}_J^{-1} \tag{140}$$

has uniformly bounded condition numbers independent of J. This can be seen by combining the properties of A according to (13) with the norm equivalences (40) and (45).

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 [Bu1]. 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 (140) [Bu1, Pa].

In Table 2 we display the condition numbers for discretizations using the weak form of the elliptic operator $-\Delta + \mathrm{id}$ 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 [Bu1].

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		

Table 2 Optimized spectral condition numbers of the operator $\mathbf{A}$ using tensor products of biorthogonal wavelets on the interval for space dimensions $n = 1,2,3$ [Bu1]

4.2 Saddle point problems

As in the previous situation, we derive a representation of the saddle point problem introduced in Section 2.3 in terms of (infinite) wavelet coordinates.

Let for $\mathscr{H} = Y \times Q$ with $Y = H^1(\Omega)$, $Q = (H^{1/2}(\Gamma))'$ two collections of wavelet bases Ψ_Y, Ψ_Q be available, each satisfying (R), (L) and (CP), with respective duals

$\tilde{\Psi}_Y$, $\tilde{\Psi}_Q$. Like before, we expand $y = \mathbf{y}^T \Psi_Y$ and $p = \mathbf{p}^T \Psi_Q$ and test with the elements from Ψ_Y, Ψ_Q. Then (21) attains the form

$$\mathbf{L}\begin{pmatrix}\mathbf{y}\\ \mathbf{p}\end{pmatrix} := \begin{pmatrix}\mathbf{A} & \mathbf{B}^T\\ \mathbf{B} & \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathbf{y}\\ \mathbf{p}\end{pmatrix} = \begin{pmatrix}\mathbf{f}\\ \mathbf{g}\end{pmatrix}, \tag{141}$$

where

$$\begin{aligned}\mathbf{A} &:= \langle \Psi_Y, A\Psi_Y\rangle \qquad & \mathbf{f} &:= \langle \Psi_Y, f\rangle,\\ \mathbf{B} &:= \langle \Psi_Q, B\Psi_Y\rangle, & \mathbf{g} &:= \langle \Psi_Q, g\rangle.\end{aligned} \tag{142}$$

In view of the above assertions, the operator $\mathbf{L}$ is an ℓ_2–automorphism, i.e., for every $(\mathbf{v},\mathbf{q}) \in \ell_2(\mathcal{I}) = \ell_2(\mathcal{I}_Y \times \mathcal{I}_Q)$ we have

$$c_{\mathbf{L}} \left\|\begin{pmatrix}\mathbf{v}\\ \mathbf{q}\end{pmatrix}\right\| \le \left\|\mathbf{L}\begin{pmatrix}\mathbf{v}\\ \mathbf{q}\end{pmatrix}\right\| \le C_{\mathbf{L}} \left\|\begin{pmatrix}\mathbf{v}\\ \mathbf{q}\end{pmatrix}\right\| \tag{143}$$

with constants $c_{\mathbf{L}}, C_{\mathbf{L}}$ only depending on $c_{\mathscr{L}}, C_{\mathscr{L}}$ from (26) and the constants in the norm equivalences (40) and (45).

For saddle point problems with an operator $\mathbf{L}$ satisfying (143), finite sections are in general not uniformly stable in the sense of (137). In fact, for discretizations on uniform grids, the validity of the corresponding mapping property relies on a suitable stability condition, see,e.g., [BF] or the article by Ricardo Nochetto and coauthors in this volume. Corresponding results derived in [DK2] are as follows.

The bilinear form $a(\cdot,\cdot)$ defined in (7) is for $c > 0$ elliptic on all of $Y = H^1(\Omega)$ and, hence, also on any finite–dimensional subspace of Y. Let there be two multiresolution analyses $\mathscr{Y}$ of $H^1(\Omega)$ and $\mathscr{Q}$ of Q where the discrete spaces are $Y_j \subset H^1(\Omega)$ and $Q_\Lambda =: Q_\ell \subset (H^{1/2}(\Gamma))'$. With the notation from Section 3.4 and in addition superscripts referring to the domain on which the functions live, these spaces are represented by

$$\begin{aligned}Y_j &= S(\Phi_j^\Omega) = S(\Psi^{j,\Omega}), \qquad & \tilde{Y}_j &= S(\tilde{\Phi}_j^\Omega) = S(\tilde{\Psi}^{j,\Omega}),\\ Q_\ell &= S(\Phi_\ell^\Gamma) = S(\Psi^{\ell,\Gamma}), & \tilde{Q}_\ell &= S(\tilde{\Phi}_\ell^\Gamma) = S(\tilde{\Psi}^{\ell,\Gamma}).\end{aligned} \tag{144}$$

Here the indices j and ℓ refer to mesh sizes on the domain and the boundary,

$$h_\Omega \sim 2^{-j} \qquad \text{and} \qquad h_\Gamma \sim 2^{-\ell}.$$

The discrete inf–sup condition, the *LBB condition*, for the pair Y_j, Q_ℓ requires that there exists a constant $\beta_1 > 0$ *independent* of j and ℓ such that

$$\inf_{q\in Q_\ell} \sup_{v\in Y_j} \frac{b(v,q)}{\|v\|_{H^1(\Omega)}\,\|q\|_{(H^{1/2}(\Gamma))'}} \ge \beta_1 > 0 \tag{145}$$

holds. We have investigated in [DK2] the general case in arbitrary spatial dimensions where the Q_ℓ are *not* trace spaces of Y_j. Employing the reverse Cauchy-Schwarz inequalities from Section 3.4, one can show that (145) is satisfied provided

that $h_\Gamma(h_\Omega)^{-1} = 2^{j-\ell} \geq c_\Omega > 1$. This is similar to the condition which was known for bivariate polygons and particular finite elements [Ba]. Although the theoretical estimates are quite pessimistic, numerical experiments for a circular boundary within a square show that the spectral condition numbers of $\mathbf{B}\mathbf{B}^T$ are still well-behaved even when this sufficient condition is violated.

It should be mentioned that the obstructions caused by the LBB condition can be avoided by means of stabilization techniques proposed, where, however, the location of the boundary of Ω relative to the mesh is somewhat constrained. A related approach which systematically avoids restrictions of the LBB type is based on least squares techniques [DKS].

It is particularly noteworthy that adaptive schemes based on wavelets can be designed in such a way that the LBB condition is *automatically* enforced. This was first observed in [DDU]. More on this subject can be found in [D4].

In order to get an impression of the value of the constants for the condition numbers for $\mathbf{A}$ in (136) and the corresponding ones for the saddle point operator on uniform grids (143), an example with $\Omega = (0,1)^2$ and Γ chosen as one face of its boundary was implemented and investigated in [Pa]. In Table 3, the spectral condition numbers of $\mathbf{A}$ and $\mathbf{L}$ with respect to two different constructions of wavelets for the case $d = 2$ and $\tilde{d} = 4$ are displayed. We see next to the first column in which the refinement level j is listed the spectral condition numbers of $\mathbf{A}$ with the wavelet construction from [DKU] denoted by $\mathbf{A}_{\mathrm{DKU}}$ and with the modification introduced in [Bu1] and a further transformation [Pa] denoted by $\mathbf{A}_{\mathrm{B}}$. The last columns contain the respective numbers for the saddle point matrix $\mathbf{L}$ where $\kappa_2(\mathbf{L}) := \sqrt{\kappa(\mathbf{L}^T\mathbf{L})}$. We observe that the spectral condition numbers stay uniformly bounded and small as j increases.

j	$\kappa_2(\mathbf{A}_{\mathrm{DKU}})$	$\kappa_2(\mathbf{A}_{\mathrm{B}})$	$\kappa_2(\mathbf{L}_{\mathrm{DKU}})$	$\kappa_2(\mathbf{L}_{\mathrm{B}})$
3	5.195e+02	1.898e+01	1.581e+02	4.147e+01
4	6.271e+02	1.066e+02	1.903e+02	1.050e+02
5	6.522e+02	1.423e+02	1.997e+02	1.399e+02
6	6.830e+02	1.820e+02	2.112e+02	1.806e+02
7	7.037e+02	2.162e+02	2.318e+02	2.145e+02
8	7.205e+02	2.457e+02	2.530e+02	2.431e+02
9	7.336e+02	2.679e+02	2.706e+02	2.652e+02

Table 3 Spectral condition numbers of the operators $\mathbf{A}$ and $\mathbf{L}$ on $\Omega = (0,1)^2$ for different constructions of biorthogonal wavelets on the interval [Pa]

4.3 Control problems: Distributed control

After these preparations, we can now discuss appropriate wavelet formulations for PDE-constrained control problems with distributed control as introduced in Section 2.4. Let for $\mathscr{V} \in \{H, \mathscr{Z}, \mathscr{U}\}$ $\Psi_{\mathscr{V}}$ denote a wavelet basis with the properties (R), (L), (CP) for $\mathscr{V}$ with dual basis $\tilde{\Psi}_{\mathscr{V}}$.

Let $\mathscr{Z}, \mathscr{U}$ satisfy the embedding (29). In terms of wavelet bases, the corresponding canonical injections correspond in view of (47) to a multiplication by a diagonal matrix. That is, let $\mathbf{D}_{\mathscr{Z}}, \mathbf{D}_H$ be such that

$$\Psi_{\mathscr{Z}} = \mathbf{D}_{\mathscr{Z}} \Psi_H, \quad \tilde{\Psi}_H = \mathbf{D}_H \Psi_{\mathscr{U}}. \tag{146}$$

Since $\mathscr{Z}$ possibly induces a weaker and $\mathscr{U}$ a stronger topology, the diagonal matrices $\mathbf{D}_{\mathscr{Z}}, \mathbf{D}_H$ are such that their entries are nondecreasing in scale, and there is a finite constant C such that

$$\|\mathbf{D}_{\mathscr{Z}}^{-1}\|, \|\mathbf{D}_H^{-1}\| \le C. \tag{147}$$

For instance, for $H = H^\alpha, \mathscr{Z} = H^\beta$, or for $H' = H^{-\alpha}$, $\mathscr{U} = H^{-\beta}$, $0 \le \beta \le \alpha$, $\mathbf{D}_{\mathscr{Z}}, \mathbf{D}_H$ have entries $(\mathbf{D}_{\mathscr{Z}})_{\lambda,\lambda} = (\mathbf{D}_H)_{\lambda,\lambda} = (\mathbf{D}^{\alpha-\beta})_{\lambda,\lambda} = 2^{(\alpha-\beta)|\lambda|}$.

We expand y in Ψ_H and u in a wavelet basis $\Psi_{\mathscr{U}}$ for $\mathscr{U} \subset H'$,

$$u = \mathbf{u}^T \Psi_U = (\mathbf{D}_H^{-1}\mathbf{u})^T \Psi_{H'}. \tag{148}$$

Following the derivation in Section 4.1, the linear constraints (28) attain the form

$$\mathbf{A}\mathbf{y} = \mathbf{f} + \mathbf{D}_H^{-1}\mathbf{u} \tag{149}$$

where

$$\mathbf{A} := a(\Psi_H, \Psi_H), \qquad \mathbf{f} := \langle \Psi_H, f \rangle. \tag{150}$$

Recall that $\mathbf{A}$ has been assumed to be symmetric. The objective functional (33) is stated in terms of the norms $\|\cdot\|_{\mathscr{Z}}$ and $\|\cdot\|_{\mathscr{U}}$. For an exact representation of these norms, corresponding Riesz operators $\mathbf{R}_{\mathscr{Z}}$ and $\mathbf{R}_{\mathscr{U}}$ defined analogously to (57) would come into play which may not be explicitly computable if $\mathscr{Z}, \mathscr{U}$ are fractional Sobolev spaces. On the other hand, as mentioned before, such a cost functional in many cases serves the purpose of yielding unique solutions while there is some ambiguity in its exact formulation. Hence, in search for a formulation which best supports numerical realizations, it is often sufficient to employ norms which are *equivalent* to $\|\cdot\|_{\mathscr{Z}}$ and $\|\cdot\|_{\mathscr{U}}$. In view of the discussion in Section 3.2, we can work for the norms $\|\cdot\|_{\mathscr{Z}}$, $\|\cdot\|_{\mathscr{U}}$ only with the diagonal scaling matrices $\mathbf{D}^s$ induced by the regularity of $\mathscr{Z}, \mathscr{U}$, or we can in addition include the Riesz map $\mathbf{R}$ defined in (52). In the numerical studies in [Bu1], a somewhat better quality of the solution is observed when $\mathbf{R}$ is included. In order to keep track of the appearance of the Riesz maps in the linear systems derived below, we choose here the latter variant.

Moreover, we expand the given observation function $y_* \in \mathscr{Z}$ as

$$y_* = \langle y_*, \tilde{\Psi}_{\mathscr{Z}} \rangle \Psi_{\mathscr{Z}} =: (\mathbf{D}_{\mathscr{Z}}^{-1}\mathbf{y}_*)^T \Psi_{\mathscr{Z}} = \mathbf{y}_*^T \Psi_H. \tag{151}$$

The way the vector $\mathbf{y}_*$ is defined here for notational convenience may by itself actually have infinite norm in ℓ_2. However, its occurrence will always include premultiplication by $\mathbf{D}_{\mathscr{Z}}^{-1}$ which is therefore always well–defined. In view of (61), we obtain the relations

$$\|y - y_*\|_{\mathscr{Z}} \sim \|\mathbf{R}^{1/2}\mathbf{D}_{\mathscr{Z}}^{-1}(\mathbf{y} - \mathbf{y}_*)\| \sim \|\mathbf{D}_{\mathscr{Z}}^{-1}(\mathbf{y} - \mathbf{y}_*\|. \tag{152}$$

Note that here $\mathbf{R} = \langle \Psi, \Psi \rangle$ (and not $\mathbf{R}^{-1}$) comes into play since y, y_* have been expanded in a scaled version of the primal wavelet basis Ψ. Hence, equivalent norms for $\|\cdot\|_{\mathscr{Z}}$ may involve $\mathbf{R}$. As for describing equivalent norms for $\|\cdot\|_{\mathscr{U}}$, recall that u is expanded in the basis Ψ_U for $U \subset H'$. Consequently, $\mathbf{R}^{-1}$ is the natural matrix to take into account when considering equivalent norms, i.e., we choose here

$$\|u\|_{\mathscr{U}} \sim \|\mathbf{R}^{-1/2}\mathbf{u}\|. \tag{153}$$

Finally, we formulate the following control problem in (infinite) wavelet coordinates.

(DCP) *For given data* $\mathbf{D}_{\mathscr{Z}}^{-1}\mathbf{y}_* \in \ell_2(\mathbb{I}_{\mathscr{Z}})$, $\mathbf{f} \in \ell_2(\mathbb{I}_H)$, *and weight parameter* $\omega > 0$, *minimize the quadratic functional*

$$\check{\mathbf{J}}(\mathbf{y}, \mathbf{u}) := \tfrac{1}{2}\,\|\mathbf{R}^{1/2}\mathbf{D}_{\mathscr{Z}}^{-1}(\mathbf{y} - \mathbf{y}_*)\|^2 + \tfrac{\omega}{2}\,\|\mathbf{R}^{-1/2}\mathbf{u}\|^2 \tag{154}$$

over $(\mathbf{y}, \mathbf{u}) \in \ell_2(\mathbb{I}_H) \times \ell_2(\mathbb{I}_H)$ *subject to the linear constraints*

$$\mathbf{A}\mathbf{y} = \mathbf{f} + \mathbf{D}_H^{-1}\mathbf{u}. \tag{155}$$

Remark 4.1. Problem (DCP) can be viewed as (discretized yet still infinite–dimensional) *representation* of the linear–quadratic control problem (27) together with (28) in wavelet coordinates in the following sense. The functional $\check{\mathbf{J}}(\mathbf{y}, \mathbf{u})$ defined in (154) is equivalent to the functional $\mathscr{J}(y, u)$ from (27) in the sense that there exist constants $0 < c_J \le C_J < \infty$ such that

$$c_J\,\check{\mathbf{J}}(\mathbf{y}, \mathbf{u}) \;\le\; \mathscr{J}(y, u) \;\le\; C_J\,\check{\mathbf{J}}(\mathbf{y}, \mathbf{u}) \tag{156}$$

holds for any $y = \mathbf{y}^T\Psi_H \in H$, given $y_* = (\mathbf{D}_{\mathscr{Z}}^{-1}\mathbf{y}_*)^T\Psi_{\mathscr{Z}} \in \mathscr{Z}$ and any $u = \mathbf{u}^T\Psi_{\mathscr{U}} \in \mathscr{U}$. Moreover, in the case of compatible data $y_* = A^{-1}f$ yielding $\mathscr{J}(y, u) \equiv 0$, the respective minimizers coincide, and $\mathbf{y}_* = \mathbf{A}^{-1}\mathbf{f}$ yields $\check{\mathbf{J}}(\mathbf{y}, \mathbf{u}) \equiv \mathbf{0}$. In this sense the new functional (154) captures the essential features of the model minimization functional.

Once problem (DCP) is posed, we can apply variational principles to derive necessary and sufficient conditions for a unique solution. All control problems considered here are in fact simple in this regard, as we have to minimize a quadratic functional subject to linear constraints, for which the first order necessary conditions are also sufficient. In principle, there are two ways to derive the optimality conditions for (DCP). We have encountered in Section 2.4 already the technique via the Lagrangian.

We define for (DCP) the *Lagrangian* introducing the *Lagrange multiplier*, *adjoint variable* or *adjoint state* $\mathbf{p}$ as

$$\mathbf{Lagr}(\mathbf{y},\mathbf{p},\mathbf{u}) := \check{\mathbf{J}}(\mathbf{y},\mathbf{u}) + \langle \mathbf{p}, \mathbf{A}\mathbf{y} - \mathbf{f} - \mathbf{D}_H^{-1}\mathbf{u}\rangle. \tag{157}$$

Then the KKT conditions $\delta\,\mathbf{Lagr}(\mathbf{w}) = \mathbf{0}$ for $\mathbf{w} = \mathbf{p},\mathbf{y},\mathbf{u}$ are, respectively,

$$\mathbf{A}\mathbf{y} = \mathbf{f} + \mathbf{D}_H^{-1}\mathbf{u}, \tag{158a}$$

$$\mathbf{A}^T\mathbf{p} = -\mathbf{D}_{\mathscr{Z}}^{-1}\mathbf{R}\mathbf{D}_{\mathscr{Z}}^{-1}(\mathbf{y} - \mathbf{y}_*) \tag{158b}$$

$$\omega\mathbf{R}^{-1}\mathbf{u} = \mathbf{D}_H^{-1}\mathbf{p}. \tag{158c}$$

The first system resulting from the variation with respect to the Lagrange multiplier always recovers the original constraints (155) and will be referred to as the *primal system* or the *state equation*. Accordingly, we call (158b) the *adjoint* or *dual system*, or the *costate equation*. The third equation (158c) is sometimes denoted as the *design equation*. Although $\mathbf{A}$ is symmetric, we continue to write $\mathbf{A}^T$ for the operator of the adjoint system to distinguish it from the primal system.

The coupled system (158) is to be solved later. However, in order to derive convergent iterations and deduce complexity estimates, a different formulation will be advantageous. It is based on the fact that $\mathbf{A}$ is according to Proposition 4.1 a boundedly invertible mapping on ℓ_2. Thus, we can formally invert (149) to obtain $\mathbf{y} = \mathbf{A}^{-1}\mathbf{f} + \mathbf{A}^{-1}\mathbf{D}_H^{-1}\mathbf{u}$. Substitution into (154) yields a functional depending only on $\mathbf{u}$,

$$\mathbf{J}(\mathbf{u}) := \tfrac{1}{2}\,\|\mathbf{R}^{1/2}\mathbf{D}_{\mathscr{Z}}^{-1}\left(\mathbf{A}^{-1}\mathbf{D}_H^{-1}\mathbf{u} - (\mathbf{y}_* - \mathbf{A}^{-1}\mathbf{f})\right)\|^2 + \tfrac{\omega}{2}\,\|\mathbf{R}^{-1/2}\mathbf{u}\|^2. \tag{159}$$

Employing the abbreviations

$$\mathbf{Z} := \mathbf{R}^{1/2}\mathbf{D}_{\mathscr{Z}}^{-1}\mathbf{A}^{-1}\mathbf{D}_H^{-1}, \tag{160a}$$

$$\mathbf{G} := -\mathbf{R}^{1/2}\mathbf{D}_{\mathscr{Z}}^{-1}(\mathbf{A}^{-1}\mathbf{f} - \mathbf{y}_*), \tag{160b}$$

the functional simplifies to

$$\mathbf{J}(\mathbf{u}) = \tfrac{1}{2}\|\mathbf{Z}\mathbf{u} - \mathbf{G}\|^2 + \tfrac{\omega}{2}\,\|\mathbf{R}^{-1/2}\mathbf{u}\|^2. \tag{161}$$

Proposition 4.3. *[K3] The functional* $\mathbf{J}$ *is twice differentiable with first and second variation*

$$\delta\mathbf{J}(\mathbf{u}) = (\mathbf{Z}^T\mathbf{Z} + \omega\mathbf{R}^{-1})\mathbf{u} - \mathbf{Z}^T\mathbf{G}, \qquad \delta^2\mathbf{J}(\mathbf{u}) = \mathbf{Z}^T\mathbf{Z} + \omega\mathbf{R}^{-1}. \tag{162}$$

In particular, $\mathbf{J}$ *is convex so that a unique minimizer exists.*

Setting

$$\mathbf{Q} := \mathbf{Z}^T\mathbf{Z} + \omega\mathbf{R}^{-1}, \qquad \mathbf{g} := \mathbf{Z}^T\mathbf{G}, \tag{163}$$

the unique minimizer $\mathbf{u}$ of (161) is given by solving

$$\delta \mathbf{J}(\mathbf{u}) = \mathbf{0} \tag{164}$$

or, equivalently, the system

$$\mathbf{Q}\mathbf{u} = \mathbf{g}. \tag{165}$$

By definition (163), $\mathbf{Q}$ is a symmetric positive definite (infinite) matrix. Hence, finite versions of (165) could be solved by gradient or conjugate gradient iterative schemes. As the convergence speed of any such iteration depends on the spectral condition number of $\mathbf{Q}$, it is important to note that the following result.

Proposition 4.4. *The (infinite) matrix* $\mathbf{Q}$ *is uniformly bounded on* ℓ_2*, i.e., there exist constants* $0 < c_{\mathbf{Q}} \le C_{\mathbf{Q}} < \infty$ *such that*

$$c_{\mathbf{Q}} \|\mathbf{v}\| \le \|\mathbf{Q}\mathbf{v}\| \le C_{\mathbf{Q}} \|\mathbf{v}\|, \qquad \mathbf{v} \in \ell_2. \tag{166}$$

The proof follows from (13) and (147) [DK3]. Of course, in order to make such iterative schemes for (165) practically feasible, the explicit inversion of $\mathbf{A}$ in the definition of $\mathbf{Q}$ has to be avoided and replaced by an iterative solver in turn. This is where the system (158) will come into play. In particular, the third equation (158c) has the following interpretation which will turn out to be very useful later.

Proposition 4.5. *If we solve for a given control vector* $\mathbf{u}$ *successively (155) for* $\mathbf{y}$ *and (158b) for* $\mathbf{p}$*, then the* residual *for (165) attains the form*

$$\mathbf{Q}\mathbf{u} - \mathbf{g} = \omega \mathbf{R}^{-1}\mathbf{u} - \mathbf{D}_U^{-1}\mathbf{p}. \tag{167}$$

Proof. Solving consecutively (155) and (158b) and recalling the definitions of $\mathbf{Z}$, $\mathbf{g}$ (160a), (163) we obtain

$$\begin{aligned}
\mathbf{D}_H^{-1}\mathbf{p} &= -\mathbf{D}_H^{-1}(\mathbf{A}^{-T}\mathbf{D}_{\mathscr{Z}}^{-1}\mathbf{R}\mathbf{D}_{\mathscr{Z}}^{-1}(\mathbf{y}-\mathbf{y}_*)) \\
&= -\mathbf{Z}^T\mathbf{R}^{1/2}\mathbf{D}_{\mathscr{Z}}^{-1}(\mathbf{A}^{-1}\mathbf{f}+\mathbf{A}^{-1}\mathbf{D}_H^{-1}\mathbf{u}-\mathbf{y}_*) \\
&= \mathbf{Z}^T\mathbf{G} - \mathbf{Z}^T\mathbf{R}^{1/2}\mathbf{D}_{\mathscr{Z}}^{-1}\mathbf{A}^{-1}\mathbf{D}_H^{-1}\mathbf{u} \\
&= \mathbf{g} - \mathbf{Z}^T\mathbf{Z}\mathbf{u}.
\end{aligned}$$

Hence, the residual $\mathbf{Q}\mathbf{u} - \mathbf{g}$ attains the form

$$\mathbf{Q}\mathbf{u} - \mathbf{g} = (\mathbf{Z}^T\mathbf{Z} + \omega\mathbf{R}^{-1})\mathbf{u} - \mathbf{g} = \omega\mathbf{R}^{-1}\mathbf{u} - \mathbf{D}_H^{-1}\mathbf{p},$$

where we have used the definition of $\mathbf{Q}$ from (163). □

Having derived the optimality conditions (158), the next issue is their efficient numerical solution. In view of the fact that the system (158) still involves infinite matrices and vectors, this also raises the question how to derive computable finite versions. By now we have investigated two scenarios.

The first version with respect to *uniform discretizations* is based on choosing finite–dimensional subspaces of the function spaces under consideration. The second version which deals with *adaptive discretizations* is actually based on the infi-

nite system (158). In both scenarios, a fully iterative numerical scheme for the solution of (158) can be designed along the following lines. The basic iteration scheme is a *gradient* or *conjugate gradient iteration* for (165) as an *outer iteration* where each application of $\mathbf{Q}$ is in turn realized by solving the primal and the dual system (155) and (158b) also by a gradient or conjugate gradient method as *inner iterations*.

For *uniform* discretizations for which we wanted to test numerically the role of equivalent norms and the influence of Riesz maps in the cost functional (154), we have used in [BK] as central iterative scheme the conjugate gradient (CG) method. Since the interior systems are only solved up to discretization error accuracy, the whole procedure may therefore be viewed as an *inexact conjugate gradient (CG) method*. We stress already at this point that the iteration numbers of such a method do *not* depend on the discretization level as finite versions of all involved operators are also uniformly well–conditioned in the sense of (166). In each step of the outer iteration, the error will be reduced by a fixed factor ρ. Combined with a *nested iteration strategy*, it will be shown that this yields an asymptotically optimal method in the amount of arithmetic operations.

Starting from the infinite coupled system (158), we have investigated in [DK3] *adaptive schemes* which, given any prescribed accuracy $\varepsilon > 0$, solve (158) such that the error for $\mathbf{y}, \mathbf{u}, \mathbf{p}$ is controlled by ε. There we have used for a simpler analysis a *gradient scheme* as basic iterative scheme.

4.4 Control problems: Dirichlet boundary control

Having derived a representation in wavelet coordinates for both the saddle point problem from Section 2.3 and the PDE-constrained control problem in the previous section, an appropriate representation of the control problem with Dirichlet boundary control introduced in Section 2.5 is straightforward. In order not to be overburdened with notation, we specifically choose the control space on the boundary as $\mathscr{U} := Q(= (H^{1/2}(\Gamma))')$. For the more general situation covered by (37), a diagonal matrix with nondecreasing entries like in (146) would come into play to switch between $\mathscr{U}$ and Q. Thus, the exact wavelet representation of the constraints (36) is given by the system (141), where we exchange the given Dirichlet boundary term $\mathbf{g}$ by $\mathbf{u}$ in the present situation to express the dependence on the control in the right hand side, i.e.,

$$\mathbf{L}\begin{pmatrix}\mathbf{y}\\ \mathbf{p}\end{pmatrix} := \begin{pmatrix}\mathbf{A} & \mathbf{B}^T\\ \mathbf{B} & \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathbf{y}\\ \mathbf{p}\end{pmatrix} = \begin{pmatrix}\mathbf{f}\\ \mathbf{u}\end{pmatrix}. \tag{168}$$

The derivation of a representer of the initial objective functional (35) is under the embedding condition (37) $\|v\|_{\mathscr{Z}} \lesssim \|v\|_Y$ for $v \in Y$ now the same as in the previous section, where all reference to the space H is to be exchanged by reference to Y. We end up with the following minimization problem in wavelet coordinates for the case of Dirichlet boundary control.

(DCP) *For given data* $\mathbf{D}_{\mathscr{Z}}^{-1}\mathbf{y}_* \in \ell_2(I_{\mathscr{Z}})$, $\mathbf{f} \in \ell_2(I_Y)$, *and weight parameter* $\omega > 0$, *minimize the quadratic functional*

$$\check{\mathbf{J}}(\mathbf{y},\mathbf{u}) := \tfrac{1}{2}\,\|\mathbf{R}^{1/2}\mathbf{D}_{\mathscr{Z}}^{-1}(\mathbf{y}-\mathbf{y}_*)\|^2 + \tfrac{\omega}{2}\,\|\mathbf{R}^{-1/2}\mathbf{u}\|^2 \tag{169}$$

over $(\mathbf{y},\mathbf{u}) \in \ell_2(I_Y) \times \ell_2(I_Y)$ *subject to the linear constraints (168),*

$$\mathbf{L}\begin{pmatrix}\mathbf{y}\\ \mathbf{p}\end{pmatrix} = \begin{pmatrix}\mathbf{f}\\ \mathbf{u}\end{pmatrix}.$$

The corresponding Karush-Kuhn-Tucker conditions can be derived by the same variational principles as in the previous section by defining a Lagrangian in terms of the functional $\check{\mathbf{J}}(\mathbf{y},\mathbf{u})$ and appending the constraints (149) with the help of additional Lagrange multipliers $(\mathbf{z},\mu)^T$, see [K3]. We obtain in this case a system of coupled saddle point problems

$$\mathbf{L}\begin{pmatrix}\mathbf{y}\\ \mathbf{p}\end{pmatrix} = \begin{pmatrix}\mathbf{f}\\ \mathbf{u}\end{pmatrix} \tag{170a}$$

$$\mathbf{L}^T\begin{pmatrix}\mathbf{z}\\ \mu\end{pmatrix} = \begin{pmatrix}-\omega\mathbf{D}_{\mathscr{Z}}^{-1}\mathbf{R}\mathbf{D}_{\mathscr{Z}}^{-1}(\mathbf{y}-\mathbf{y}_*)\\ \mathbf{0}\end{pmatrix} \tag{170b}$$

$$\mathbf{u} = \mu. \tag{170c}$$

Again, the first system appearing here, the *primal system*, are just the constraints (149) while (46) will be referred to as the *dual* or *adjoint system*. The specific form of the right hand side of the dual system emerges from the particular formulation of the minimization functional (169). The (here trivial) equation (170c) stems from measuring $\mathbf{u}$ just in ℓ_2, representing measuring the control in its natural trace norm. Instead of replacing μ by $\mathbf{u}$ in (46) and trying to solve the resulting equations, (170c) will be essential to devise an inexact gradient scheme. In fact, since $\mathbf{L}$ in (149) is an invertible operator, we can rewrite $\check{\mathbf{J}}(\mathbf{y},\mathbf{u})$ by formally inverting (149) as a functional of $\mathbf{u}$, that is, $\mathbf{J}(\mathbf{u}) := \check{\mathbf{J}}(\mathbf{y}(\mathbf{u}),\mathbf{u})$ as above. The following result will be very useful for the design of the outer–inner iterative solvers

Proposition 4.6. *The first variation of* $\mathbf{J}$ *satisfies*

$$\delta\mathbf{J}(\mathbf{u}) \;=\; \mathbf{u}-\mu, \tag{171}$$

where $(\mathbf{u},\mu)$ *are part of the solution of (170). Moreover,* $\mathbf{J}$ *is convex so that a unique minimizer exists.*

Hence, equation (170c) is just $\delta\mathbf{J}(\mathbf{u}) = \mathbf{0}$. For a unified treatment below of both control problems considered in these notes, it will be useful to rewrite (170c) like in (165) as a condensed equation for the control $\mathbf{u}$ alone. We formally invert (168) and (170b) and obtain

$$\mathbf{Q}\mathbf{u} = \mathbf{g} \tag{172}$$

with the abbreviations

$$\mathbf{Q} := \mathbf{Z}^T\mathbf{Z} + \omega\mathbf{I}, \quad \mathbf{g} := \mathbf{Z}^T(\mathbf{y}_* - \mathbf{T}_\square\mathbf{L}^{-1}\mathbf{I}_\square\mathbf{f}) \tag{173}$$

and

$$\mathbf{Z} := \mathbf{T}_\square\mathbf{L}^{-1}\mathbf{I}_\square, \qquad \mathbf{I}_\square := \begin{pmatrix}\mathbf{0}\\ \mathbf{I}\end{pmatrix}, \qquad \mathbf{T}_\square := (\mathbf{T}\ \ \mathbf{0}). \tag{174}$$

Proposition 4.7. *The vector* $\mathbf{u}$ *as part of the solution vector* $(\mathbf{y},\mathbf{p},\mathbf{z},\mu,\mathbf{u})$ *of (170) coincides with the unique solution* $\mathbf{u}$ *of the condensed equations (172).*

5 Iterative solution

Each of the four problem classes discussed above lead to the problem to finally solve a system

$$\delta\mathbf{J}(\mathbf{q}) \ = \ \mathbf{0} \tag{175}$$

or, equivalently, a linear system

$$\mathbf{M}\mathbf{q} = \mathbf{b}, \tag{176}$$

where $\mathbf{M} : \ell_2 \to \ell_2$ is a (possibly infinite) symmetric positive definite matrix satisfying

$$c_\mathbf{M}\|\mathbf{v}\| \le \|\mathbf{M}\mathbf{v}\| \le C_\mathbf{M}\|\mathbf{v}\|, \quad \mathbf{v} \in \ell_2, \tag{177}$$

for some constants $0 < c_\mathbf{M} \le C_\mathbf{M} < \infty$ and where $\mathbf{b} \in \ell_2$ is some given right hand side.

A simple *gradient method* for solving (175) is

$$\mathbf{q}_{k+1} := \mathbf{q}_k - \alpha\,\delta\mathbf{J}(\mathbf{q}_k), \qquad k = 0,1,2,\ldots \tag{178}$$

with some initial guess $\mathbf{q}_0$. In all of the previously considered situations, it has been asserted that there exists a fixed parameter α, depending on bounds for the second variation of $\mathbf{J}$, such that (178) converges and reduces the error in each step by at least a fixed factor $\rho < 1$, i.e.,

$$\|\mathbf{q} - \mathbf{q}_{k+1}\| \le \rho\|\mathbf{q} - \mathbf{q}_k\|, \quad k = 0,1,2,\ldots, \tag{179}$$

where ρ is determined by

$$\rho := \|\mathbf{I} - \alpha\mathbf{M}\| < 1.$$

Hence, the scheme (178) is a convergent iteration for the possibly infinite system (176). Next we will need to discuss how to reduce the infinite systems to computable finite versions.

5.1 Finite systems on uniform grids

We consider finite-dimensional trial spaces with respect to uniform discretizations. For each of the Hilbert spaces H, this means in the wavelet setting to pick the index set of all indices up to some *highest refinement level* J, i.e.,

$$\mathcal{I}_{J,H} := \{\lambda \in \mathcal{I}_H : |\lambda| \le J\} \subset \mathcal{I}_H$$

satisfying $N_{J,H} := \#\mathcal{I}_{J,H} < \infty$. The representation of operators is then built as in Section 3.3 with respect to this truncated index set which corresponds to deleting all rows and columns that refer to indices λ such that $|\lambda| > J$, and correspondingly for functions. There is by construction also a *coarsest level* of resolution denoted by j_0.

Computationally the representation of operators according to (62) is in the case of uniform grids always realized as follows. First, the operator is set up in terms of the *generator basis* on the finest level J. This generator basis simply consists of tensor products of B-Splines, or linear combinations of these near the boundaries. The representation of an operator in the *wavelet basis* is then achieved by applying the Fast Wavelet Transform (FWT) which needs $\mathcal{O}(N_{J,H})$ arithmetic operations and is therefore asymptotically optimal, see, e.g., [D2, DKU, K1] and Section 3.4.

In order not to overburden the notation, let in this subsection the resulting system for $N = N_{J,H}$ unknowns again be denoted by

$$\mathbf{M}\mathbf{q} = \mathbf{b}, \tag{180}$$

where now $\mathbf{M} : \mathbb{R}^N \to \mathbb{R}^N$ is a symmetric positive definite matrix satisfying (177) on $\mathbb{R}^N$. It will be convenient to abbreviate the residual using an approximation $\tilde{\mathbf{q}}$ to $\mathbf{q}$ for (180) as

$$\text{RESD}(\tilde{\mathbf{q}}) := \mathbf{M}\tilde{\mathbf{q}} - \mathbf{b}. \tag{181}$$

We will employ a basic conjugate gradient method that iteratively computes an approximate solution $\mathbf{q}_K$ to (180) with given initial vector $\mathbf{q}_0$ and given tolerance $\varepsilon > 0$ such that

$$\|\mathbf{M}\mathbf{q}_K - \mathbf{b}\| = \|\text{RESD}(\mathbf{q}_K)\| \le \varepsilon, \tag{182}$$

where K denotes the number of iterations used. Later we specify ε depending on the discretization for which (180) is set up. The following scheme CG contains a routine $\text{APP}(\eta_k, \mathbf{M}, \mathbf{d}_k)$ which in view of the problem classes discussed above is to have the property that it approximately computes the product $\mathbf{M}\mathbf{d}_k$ up to a tolerance $\eta_k = \eta_k(\varepsilon)$ depending on ε, i.e., the output ${}_k$ of $\text{APP}(\eta_k, \mathbf{M}, \mathbf{d}_k)$ satisfies

$$\|\mathbf{m}_k - \mathbf{M}\mathbf{d}_k\| \le \eta_k. \tag{183}$$

For the cases where $\mathbf{M} = \mathbf{A}$, this is simply the matrix-vector multiplication $\mathbf{M}\mathbf{d}_k$. For the situations where $\mathbf{M}$ may involve the solution of an additional system, this multiplication will be only approximative.

CG $[\varepsilon, \mathbf{q}_0, \mathbf{M}, \mathbf{b}] \to \mathbf{q}_K$

(I) SET $\mathbf{d}_0 := \mathbf{b} - \mathbf{M}\mathbf{q}_0$ AND $\mathbf{r}_0 := -\mathbf{d}_0$. LET $k = 0$.
(II) WHILE $\|\mathbf{r}_k\| > \varepsilon$

$$\begin{aligned} \mathbf{m}_k &:= \mathrm{APP}(\eta_k(\varepsilon), \mathbf{M}, \mathbf{d}_k) \\ \alpha_k &:= \frac{(\mathbf{r}_k)^T \mathbf{r}_k}{(\mathbf{d}_k)^T \mathbf{m}_k} \qquad\qquad \mathbf{q}_{k+1} := \mathbf{q}_k + \alpha_k \mathbf{d}_k \\ \mathbf{r}_{k+1} &:= \mathbf{r}_k + \alpha_k \mathbf{m}_k \qquad\qquad \beta_k := \frac{(\mathbf{r}_{k+1})^T \mathbf{r}_{k+1}}{(\mathbf{r}_k)^T \mathbf{r}_k} \\ \mathbf{d}_{k+1} &:= -\mathbf{r}_{k+1} + \beta_k \mathbf{d}_k \\ k &:= k+1 \end{aligned} \tag{184}$$

(III) SET $K := k - 1$.

Let us briefly discuss in the case $\mathbf{M} = \mathbf{A}$ that the final iterate $\mathbf{q}_K$ indeed satisfies (182). From the newly computed iterate $\mathbf{q}_{k+1} = \mathbf{q}_k + \alpha_k \mathbf{d}_k$ it follows by applying $\mathbf{M}$ on both sides that $\mathbf{M}\mathbf{q}_{k+1} - \mathbf{b} = \mathbf{M}\mathbf{q}_k - \mathbf{b} + \alpha_k \mathbf{M}\mathbf{d}_k$ which is the same as $\mathrm{RESD}(\mathbf{q}_{k+1}) = \mathrm{RESD}(\mathbf{q}_k) + \alpha_k \mathbf{M}\mathbf{d}_k$. By the initialization for $\mathbf{r}_k$ used above, this in turn is the updating term for $\mathbf{r}_k$, hence, $\mathbf{r}_k = \mathrm{RESD}(\mathbf{q}_k)$. After the stopping criterion based on $\mathbf{r}_k$ is met, the final iterate $\mathbf{q}_K$ observes (182).

The routine CG computes the *residual* up to the stopping criterion ε. From the residual, we can in view of (177) estimate the *error* in the solution as

$$\|\mathbf{q} - \mathbf{q}_K\| = \|\mathbf{M}^{-1}(\mathbf{b} - \mathbf{M}\mathbf{q}_K)\| \le \|\mathbf{M}^{-1}\| \, \|\mathrm{RESD}(\mathbf{q}_K)\| \le \frac{\varepsilon}{c_{\mathbf{M}}}, \tag{185}$$

that is, it may deviate from the norm of the residual from a factor proportional to the smallest eigenvalue of $\mathbf{M}$.

Distributed control. Let us now apply the solution scheme to the situation from Section 4.3 where $\mathbf{Q}$ now involves the inversion of finite-dimensional systems (158a) and (158b). The material in the remainder of this subsection is essentially contained in [BK].

We begin with a specification of the approximate computation of the right hand side $\mathbf{b}$ which also contains applications of $\mathbf{A}^{-1}$.

RHS$[\zeta, \mathbf{A}, \mathbf{f}, \mathbf{y}_*] \to \mathbf{b}_\zeta$

(I) CG $[\frac{c_{\mathbf{A}}}{2C} \frac{c_{\mathbf{A}}}{C^2 C_0^2} \zeta, \mathbf{0}, \mathbf{A}, \mathbf{f}] \to \mathbf{b}_1$
(II) CG $[\frac{c_{\mathbf{A}}}{2C} \zeta, \mathbf{0}, \mathbf{A}^T, -\mathbf{D}_{\mathscr{Z}}^{-1} \mathbf{R} \mathbf{D}_{\mathscr{Z}}^{-1} (\mathbf{b}_1 - \mathbf{y}_*)] \to \mathbf{b}_2$
(III) $\mathbf{b}_\zeta := \mathbf{D}_H^{-1} \mathbf{b}_2$.

The tolerances used within the two conjugate gradient methods depend on the constants $c_{\mathbf{A}}, C, C_0$ from (13), (147) and (55), respectively. Since the additional factor $c_{\mathbf{A}}(CC_0)^{-2}$ in the stopping criterion in step (I) in comparison to step (II) is in general smaller than one, this means that the primal system needs to be solved more accurately than the adjoint system in step (II).

Proposition 5.1. *The result* $\mathbf{b}_\zeta$ *of* $\mathrm{RHS}[\zeta,\mathbf{A},\mathbf{f},\mathbf{y}_*]$ *satisfies*

$$\|\mathbf{b}_\zeta-\mathbf{b}\|\le\zeta. \tag{186}$$

Proof. Recalling the definition (163) of $\mathbf{b}$, step (III) and step (II) yield

$$\begin{aligned}\|\mathbf{b}_\zeta-\mathbf{b}\| &\le \|\mathbf{D}_H^{-1}\|\,\|\mathbf{b}_2-\mathbf{D}_H\mathbf{b}\|\\ &\le C\|\mathbf{A}^{-T}\|\,\|\mathbf{A}^T\mathbf{b}_2-\mathbf{D}_{\mathscr{Z}}^{-1}\mathbf{R}\mathbf{D}_{\mathscr{Z}}^{-1}(\mathbf{A}^{-1}\mathbf{f}-\mathbf{b}_1+\mathbf{b}_1-\mathbf{y}_*)\|\\ &\le \frac{C}{c_\mathbf{A}}\left(\frac{c_\mathbf{A}}{2C}\zeta+\|\mathbf{D}_{\mathscr{Z}}^{-1}\mathbf{R}\mathbf{D}_{\mathscr{Z}}^{-1}(\mathbf{A}^{-1}\mathbf{f}-\mathbf{b}_1)\|\right).\end{aligned} \tag{187}$$

Employing the upper bounds for $\mathbf{D}_{\mathscr{Z}}^{-1}$ and $\mathbf{R}$, we arrive at

$$\begin{aligned}\|\mathbf{b}_\zeta-\mathbf{b}\| &\le \frac{C}{c_\mathbf{A}}\left(\frac{c_\mathbf{A}}{2C}\zeta+C^2C_0^2\,\|\mathbf{A}^{-1}\|\,\|\mathbf{f}-\mathbf{A}\mathbf{b}_1\|\right)\\ &\le \frac{C}{c_\mathbf{A}}\left(\frac{c_\mathbf{A}}{2C}\zeta+\frac{C^2C_0^2}{c_\mathbf{A}}\frac{c_\mathbf{A}}{2C}\frac{c_\mathbf{A}}{C^2C_0^2}\zeta\right)\ =\zeta.\end{aligned} \tag{188}$$

□

Accordingly, an approximation $\mathbf{m}_\eta$ to the matrix-vector product $\mathbf{Q}\mathbf{d}$ is the output of the following routine APP.

APP$[\eta,\mathbf{Q},\mathbf{d}]\to\mathbf{m}_\eta$

(I) CG $[\frac{c_\mathbf{A}}{3C}\frac{c_\mathbf{A}}{C^2C_0^2}\eta,\mathbf{0},\mathbf{A},\mathbf{f}+\mathbf{D}_H^{-1}\mathbf{d}]\to\mathbf{y}_\eta$

(II) CG $[\frac{c_\mathbf{A}}{3C}\eta,\mathbf{0},\mathbf{A}^T,-\mathbf{D}_{\mathscr{Z}}^{-1}\mathbf{R}\mathbf{D}_Z^{-1}(\mathbf{y}_\eta-\mathbf{y}_*)]\to\mathbf{p}_\eta$

(III) $\mathbf{m}_\eta:=\mathbf{g}_{\eta/3}+\omega\mathbf{R}^{-1}\mathbf{d}-\mathbf{D}_H^{-1}\mathbf{p}_\eta$.

The choice of the tolerances for the interior application of CG in steps (I) and (II) will become clear from the following result.

Proposition 5.2. *The result* $\mathbf{m}_\eta$ *of* APP$[\eta,\mathbf{Q},\mathbf{d}]$ *satisfies*

$$\|\mathbf{m}_\eta-\mathbf{Q}\mathbf{d}\|\le\eta. \tag{189}$$

Proof. Denote by $\mathbf{y}_\mathbf{d}$ the exact solution of (158a) with $\mathbf{d}$ in place of $\mathbf{u}$ on the right hand side, and by $\mathbf{p}_\mathbf{d}$ the exact solution of (158b) with $\mathbf{y}_\mathbf{d}$ on the right hand side. Then we deduce from step (III) and (167) combined with (55) and (147)

$$\begin{aligned}\|\mathbf{m}_\eta-\mathbf{Q}\mathbf{d}\| &= \|\mathbf{g}_{\eta/3}-\mathbf{g}+\omega\mathbf{R}^{-1}\mathbf{d}-\mathbf{D}_U^{-1}\mathbf{p}_\eta\ -(\mathbf{Q}\mathbf{d}-\mathbf{g})\|\\ &\le \frac{1}{3}\eta+\|\omega\mathbf{R}^{-1}\mathbf{d}-\mathbf{D}_U^{-1}\mathbf{p}_\eta-(\omega\mathbf{R}^{-1}\mathbf{d}-\mathbf{D}_U^{-1}\mathbf{p}_\mathbf{d})\|\\ &\le \frac{1}{3}\eta+C\|\mathbf{p}_\mathbf{d}-\mathbf{p}_\eta\|.\end{aligned} \tag{190}$$

Denote by $\hat{\mathbf{p}}$ the exact solution of (158b) with $\mathbf{y}_\eta$ on the right hand side. Then we have $\mathbf{p}_\mathbf{d}-\hat{\mathbf{p}}=-\mathbf{A}^{-T}\mathbf{D}_Z^{-1}\mathbf{R}\mathbf{D}_Z^{-1}(\mathbf{y}_\mathbf{d}-\mathbf{y}_\eta)$. It follows by (13), (55) and (147) that

$$\|\mathbf{p_d} - \hat{\mathbf{p}}\| \leq \frac{C^2 C_0^2}{c_\mathbf{A}} \|\mathbf{y_d} - \mathbf{y}_\eta\| \leq \frac{1}{3C}\eta, \tag{191}$$

where the last estimate follows by the choice of the threshold in step (I). Finally, the combination(190) and (191) together with (186) and the stopping criterion in step (II) readily confirms that

$$\begin{aligned}\|\mathbf{m}_\eta - \mathbf{Qd}\| &\leq \frac{1}{3}\eta + C\left(\|\mathbf{p_d} - \hat{\mathbf{p}}\| + \|\hat{\mathbf{p}} - \mathbf{p}_\eta\|\right) \\ &\leq \frac{1}{3}\eta + C\left(\frac{1}{3C}\eta + \frac{1}{3C}\eta\right) = \eta.\end{aligned}$$

□

The effect of perturbed applications of $\mathbf{M}$ in CG and more general Krylov subspace schemes with respect to convergence has been investigated in a numerical linear algebra context for a given linear system (180) in several papers. Here we have chosen the η_i to be proportional to the outer accuracy ε incorporating a safety factor accounting for the values of β_i and $\|\mathbf{r}_i\|$.

Finally, we can formulate a full nested iteration strategy for finite systems (158) on uniform grids which employs outer and inner CG routines as follows. The scheme starts at the coarsest level of resolution j_0 with some initial guess $\mathbf{u}_0^{j_0}$ and successively solves (165) with respect to each level j until the norm of the current residual is below the discretization error on that level.

In wavelet coordinates, $\|\cdot\|$ corresponds to the energy norm. If we employ on the primal side for approximation linear combinations of B–splines of order d, the discretization error is for smooth solutions expected to be proportional to $2^{-(d-1)j}$. Then the refinement level is successively increased until on the finest level J a prescribed tolerance proportional to the discretization error $2^{-(d-1)J}$ is met. In the following, superscripts on vectors denote the refinement level on which this term is computed. The given data $\mathbf{y}_*^j$, $\mathbf{f}^j$ are supposed to be accessible on all levels. On the coarsest level, the solution of (165) is computed exactly up to double precision by QR decomposition. Subsequently, the results from level j are prolongated onto the next higher level $j+1$. Using wavelets, this is accomplished by simply adding zeros: wavelet coordinates have the character of differences so that this prolongation corresponds to the exact representation in higher resolution wavelet coordinates. The resulting *Nested–Iteration–Incomplete–Conjugate–Gradient* Algorithm is the following.

NIICG$[J] \to \mathbf{u}^J$

(I) INITIALIZATION FOR COARSEST LEVEL $j := j_0$

(1) COMPUTE RIGHT HAND SIDE $\mathbf{g}^{j_0} = (\mathbf{Z}^T\mathbf{G})^{j_0}$ BY QR DECOMPOSITION USING (160).

(2) COMPUTE SOLUTION $\mathbf{u}^{j_0}$ OF (165) BY QR DECOMPOSITION.

(II) WHILE $j < J$

(1) PROLONGATE $\mathbf{u}^j \to \mathbf{u}_0^{j+1}$ BY ADDING ZEROS, SET $j := j+1$.
(2) COMPUTE RIGHT HAND SIDE USING RHS $[2^{-(d-1)j}, \mathbf{A}, \mathbf{f}^j, \mathbf{y}_*^j] \to \mathbf{g}^j$.
(3) COMPUTE SOLUTION OF (165) USING CG $[2^{-(d-1)j}, \mathbf{u}_0^j, \mathbf{Q}, \mathbf{g}^j] \to \mathbf{u}^j$.

Recall that step (II.3) requires multiple calls of APP$[\eta, \mathbf{Q}, \mathbf{d}]$, which in turn invokes both CG $[\ldots, \mathbf{A}, \ldots]$ as well as CG $[\ldots, \mathbf{A}^T, \ldots]$ in each application.

On account of (13) and (166), finite versions of the system matrices $\mathbf{A}$ and $\mathbf{Q}$ have uniformly bounded condition numbers, entailing that each CG routine employed in the process reduces the error by a fixed rate $\rho < 1$ in each iteration step. Let $N_J \sim 2^{nJ}$ be the total number of unknowns (for $\mathbf{y}^J, \mathbf{u}^J$ and $\mathbf{p}^J$) on the highest level J. Employing the CG method only on the highest level, one needs $\mathcal{O}(J) = \mathcal{O}(\log \varepsilon)$ iterations to achieve the prescribed discretization error accuracy $\varepsilon_J = 2^{-(d-1)J}$. As each application of $\mathbf{A}$ and $\mathbf{Q}$ requires $\mathcal{O}(N_J)$ operations, the solution of (165) by CG only on the finest level requires $\mathcal{O}(J N_J)$ arithmetic operations.

Proposition 5.3. *If the residual (167) is computed up to discretization error proportional to $2^{-(d-1)j}$ on each level j and the corresponding solutions are taken as initial guesses for the next higher level,* NIICG *is an asymptotically optimal method in the sense that it provides the solution $\mathbf{u}^J$ up to discretization error on level J in an overall amount of $\mathcal{O}(N_J)$ arithmetic operations.*

Proof. In the above notation, nested iteration allows one to get rid of the factor J in the total amount of operations. Starting with the exact solution on the coarsest level j_0, in view of the uniformly bounded condition numbers of $\mathbf{A}$ and $\mathbf{Q}$, one needs only a fixed amount of iterations to reduce the error up to discretization error accuracy $\varepsilon_j = 2^{-(d-1)j}$ on each subsequent level j, taking the solution from the previous level as initial guess. Thus, on each level, one needs $\mathcal{O}(N_j)$ operations to realize discretization error accuracy. Since the spaces are nested and the number of unknowns on each level grows like $N_j \sim 2^{nj}$, by a geometric series argument the total number of arithmetic operations stays proportional to $\mathcal{O}(N_J)$. □

5.2 Numerical examples

5.2.1 Distributed control problem

As an illustration of the issue which norms to choose in the control functional, we consider the following example of a one-dimensional distributed control problem with the Helmholtz operator in (6) ($\mathbf{a} = I, c = 1$) and homogeneous Dirichlet boundary condition. A non–constant right hand side $f(x) := 1 + 2.3\exp(-15|x-0.5|)$ is chosen, and the target state is set to a constant $y_* \equiv 1$. We first investigate the role the different norms $\|\cdot\|_{\mathscr{Z}}$ and $\|\cdot\|_{\mathscr{U}}$ in (27), which is encoded in the diagonal matrices $\mathbf{D}_{\mathscr{Z}}, \mathbf{D}_H$ from (146), have on the solution. We see in Figure 3 for the choice $\mathscr{U} = L_2(0,1)$ and $\mathscr{Z} = H^s(0,1)$ for different values of s varying between 0 and 1

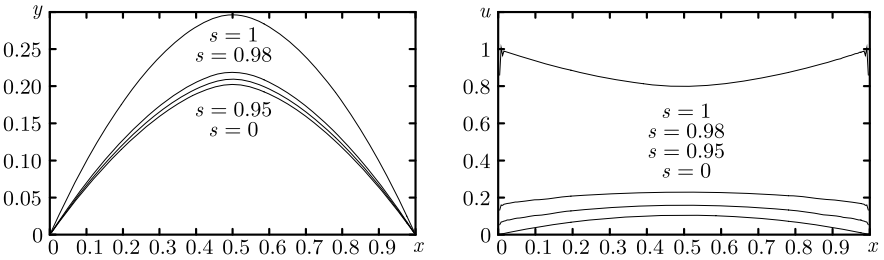

Fig. 3 Distributed control problem for elliptic PDE with Dirichlet boundary conditions, a peak as right hand side f, $y_* \equiv 1$, $\omega = 0$, $\mathscr{U} = L_2(0,1)$ and varying $\mathscr{Z} = H^s(0,1)$. Left: state y, right: control u

the solution y (left) and the corresponding control u (right) for fixed weight $\omega = 1$. As s is increased, a stronger tendency of y towards the prescribed state $y_* \equiv 1$ can be observed which is, however, deterred from reaching this state by the homogeneous boundary conditions. Extensive studies of this type can be found in [Bu1, BK].

An example displaying the performance of the proposed fully iterative scheme NIICG is shown in Table 4 for $n = 2$ and in Table 5 for $n = 3$.

j	$\|\mathbf{r}_K^j\|$	#O	#E	#A	#R	$\|R(\mathbf{y}^J) - \mathbf{y}^j\|$	$\|\mathbf{y}^J - P(\mathbf{y}^j)\|$	$\|R(\mathbf{u}^J) - \mathbf{u}^j\|$	$\|\mathbf{u}^J - P(\mathbf{u}^j)\|$
3						6.86e-03	1.48e-02	1.27e-04	4.38e-04
4	1.79e-05	5	12	5	8	2.29e-03	7.84e-03	4.77e-05	3.55e-04
5	1.98e-05	5	14	6	9	6.59e-04	3.94e-03	1.03e-05	2.68e-04
6	4.92e-06	7	13	5	9	1.74e-04	1.96e-03	2.86e-06	1.94e-04
7	3.35e-06	7	12	5	9	4.55e-05	9.73e-04	9.65e-07	1.35e-04
8	2.42e-06	7	11	5	10	1.25e-05	4.74e-04	7.59e-07	8.88e-05
9	1.20e-06	8	11	5	10	4.55e-06	2.12e-04	4.33e-07	5.14e-05
10	4.68e-07	9	10	5	9	3.02e-06	3.02e-06	2.91e-07	2.91e-07

Table 4 Iteration history for a two-dimensional distributed control problem with Neumann boundary conditions, $\omega = 1$, $\mathscr{Z} = H^1(\Omega)$, $\mathscr{U} = (H^{0.5}(\Omega))'$

j	$\|\mathbf{r}_K^j\|$	#O	#E	#A	#R	$\|R(\mathbf{y}^J) - \mathbf{y}^j\|$	$\|\mathbf{y}^J - P(\mathbf{y}^j)\|$	$\|R(\mathbf{u}^J) - \mathbf{u}^j\|$	$\|\mathbf{u}^J - P(\mathbf{u}^j)\|$
3						1.41e-04	2.92e-04	1.13e-05	2.36e-05
4	6.09e-06	10	9	1	49	1.27e-04	1.78e-04	3.46e-06	3.79e-06
5	3.25e-06	10	7	1	58	1.11e-05	6.14e-05	9.47e-07	9.53e-07
6	1.71e-06	7	6	1	57	1.00e-05	2.86e-05	5.03e-07	5.03e-07
7	8.80e-07	6	6	1	53	9.19e-06	9.19e-06	3.72e-07	3.72e-07

Table 5 Iteration history for a three-dimensional distributed control problem with Neumann boundary conditions, $\omega = 1$, $\mathscr{Z} = H^1(\Omega)$, $\mathscr{U} = (H^1(\Omega))'$

This is an example of a control problem for the Helmholtz operator with Neumann boundary conditions. The stopping criterion for the outer iteration (relative to

$\|\cdot\|$ which corresponds to the energy norm) on level j is chosen to be proportional to 2^{-j}. The second column displays the final value of the residual of the outer CG scheme on this level, i.e., $\|\mathbf{r}_K^j\| = \|\text{RESD}(\mathbf{u}_K^j)\|$. The next three columns show the number of outer CG iterations (#O) for $\mathbf{Q}$ according to the APP scheme followed by the maximum number of inner iterations for the primal system (#E), the adjoint system (#A) and the design equation (#R). We see very well the effect of the uniformly bounded condition numbers of all involved operators. The last columns display different versions of the actual error in the state $\mathbf{y}$ and the control $\mathbf{u}$ when compared to the fine grid solution (R denotes restriction of the fine grid solution to the actual grid, and P denotes prolongation). Here we can see the slight effect of the constants appearing in (185). Nevertheless the error is very well controlled by the residual.

More results for up to three spatial dimensions can be found in [Bu1, BK]. All numbers were obtained on a 3.2GHz Pentium IV computer (family 15, model 4, stepping 1, with 1MB L2 Cache).

5.2.2 Dirichlet boundary control

For the system of saddle point problems (170) arising from the control problem with Dirichlet boundary control in Section 2.5, also a fully iterative algorithm NI-ICG can be designed along the above lines with yet another level of inner iteration. Again the design equation (170c) for $\mathbf{u}$ serves as the equation for which a basic iterative scheme (178) can be posed. Of course, the CG method for $\mathbf{A}$ then has to be replaced by a convergent iterative scheme for saddle point operators $\mathbf{L}$ like Uzawa's algorithm. Also the discretization has to be chosen such that the LBB condition is satisfied, see Section 4.2. Details can be found in [K3]. Alternatively, since $\mathbf{L}$ has a uniformly bounded condition number, the CG scheme can, in principle, also be applied to $\mathbf{L}^T\mathbf{L}$. The performance of wavelet schemes on uniform grids for such systems of saddle point problems arising from optimal control has been investigated systematically in [Pa].

For illustration of the choice of different norms for the Dirichlet boundary control problem, consider the following example. We control the system through the (green) control boundary Γ in Figure 4 while a prescribed state $y_* \equiv 1$ on the (red) observation boundary Γ_y opposite the control boundary is to be achieved. The right hand side is chosen as constant $f \equiv 1$, and $\omega = 1$. Each layer in Figure 4 corresponds to the state y for different values of s when the observation term is measured in $H^s(\Gamma_y)$, that is, the objective functional (35) contains a term $\|y - y_*\|^2_{H^s(\Gamma_y)}$ for increasing s from bottom to top. We see that as the smoothness index s for the observation increases, the state moves towards the target state at the observation boundary. In comparison, in Figure 5 the weight parameter ω balancing the two terms in the functional is modified. We observe that the effect on the solution of varying s corresponds to a similar behaviour of varying the weight. However, as ω directly influences the conditioning of the system of saddle point operators, a solution scheme with fixed ω and varying s can be considered numerically more stable.

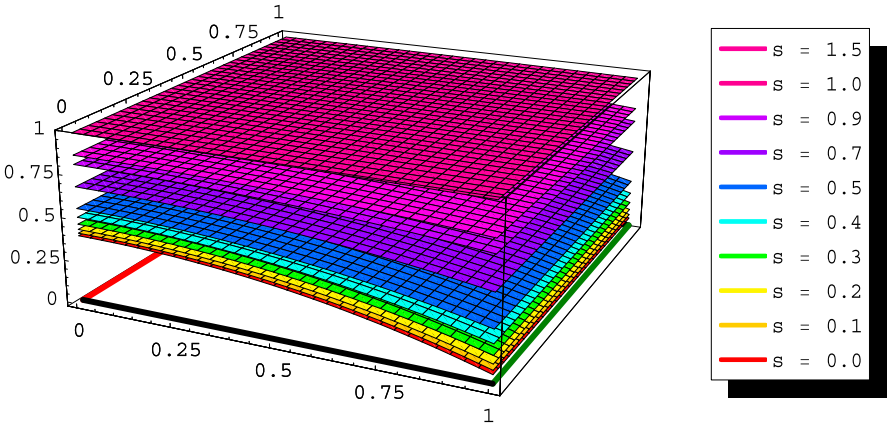

Fig. 4 State y of the Dirichlet boundary control problem using the objective functional $\mathcal{J}(y,u) = \frac{1}{2}\|y - y_*\|^2_{H^s(\Gamma_y)} + \frac{1}{2}\|u\|^2_{H^{1/2}(\Gamma)}$ for control boundary Γ (green) and observation boundary Γ_y (red) for different values of the Sobolev smoothness index s on resolution level $J = 5$ [Pa]

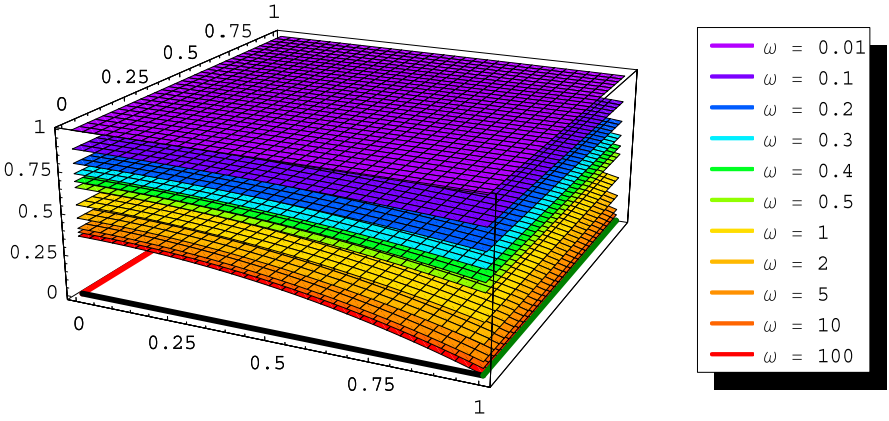

Fig. 5 State y of the Dirichlet boundary control problem using the objective functional $\mathcal{J}(y,u) = \frac{1}{2}\|y - y_*\|^2_{H^s(\Gamma_y)} + \frac{\omega}{2}\|u\|^2_{H^{1/2}(\Gamma)}$ for control boundary Γ (green) and observation boundary Γ_y (red) for different values of the weight parameter ω [Pa]

Finally, we display in Table 6 some numerical results for an elliptic control problem with Dirichlet boundary control in two spatial dimensions. Among the various iteration schemes tested, the best results with a minimal amount of iteration numbers (here: at most 2) were obtained for an inexact gradient iteration on $\mathbf{u}$ and Uzawa–type schemes with conjugate directions for each of the saddle point problems together with nested iteration.

j	$\|\mathbf{r}^j\|$	$\|\mathbf{y}-\mathbf{y}^J\|$	k_j	$\frac{\#\text{Int-It}}{k_j}$
4	1.6105e-02	7.7490e-00	0	–
5	1.6105e-02	7.7506e-00	0	–
6	6.3219e-03	1.7544e-02	2	1
7	5.8100e-03	3.3873e-02	0	–
8	1.6378e-03	3.4958e-03	2	1
9	1.8247e-03	7.4741e-03	0	–
10	4.3880e-04	9.2663e-04	2	1
11	4.6181e-04	1.8486e-03	0	–

Table 6 Dirichlet boundary control problem in two spatial dimensions with $y_{\Gamma_y} \equiv 1$, $f \equiv 1$, $\omega = 1$, $s = t = 0.5$. The table shows the number of iterations k_j needed to reduce the $\mathscr{L}$-error of $\mathbf{r}^j$ by a factor of 0.5 after prolongation of all final vectors from the previous level [Pa]

References

[Ba] I. Babuška, The finite element method with Lagrange multipliers, Numer. Math. **20** (1973), 179–192.

[B] D. Braess, Finite Elements: Theory, Fast Solvers and Applications in Solid Mechanics, 2nd ed., Cambridge University Press, Cambridge, 2001.

[BH] D. Braess, W. Hackbusch, A new convergence proof for the multigrid method including the V-cycle, SIAM J. Numer. Anal. **20** (1983), 967–975.

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

[BF] F. Brezzi, M. Fortin, Mixed and Hybrid Finite Element Methods, Springer, 1991.

[Bu1] C. Burstedde, Fast Optimised Wavelet Methods for Control Problems Constrained by Elliptic PDEs, PhD Dissertation, Mathematisch-Naturwissenschaftliche Fakultät, Universität Bonn, Germany, 2005.

[Bu2] C. Burstedde, On the numerical evaluation of fractional Sobolev norms, Comm. Pure Appl. Anal. **6**(3) (2007), 587–605.

[BK] C. Burstedde, A. Kunoth, Fast iterative solution of elliptic control problems in wavelet discretizations, J. Comp. Appl. Math. **196**(1) (2006), 299–319.

[CTU] C. Canuto, A. Tabacco, K. Urban, The wavelet element method, part I: Construction and analysis, Appl. Comput. Harm. Anal. **6** (1999), 1–52.

[CDP] J.M. Carnicer, W. Dahmen, J.M. Peña, Local decomposition of refinable spaces, Appl. Comp. Harm. Anal. **3** (1996), 127–153.

[CF] Z. Ciesielski, T. Figiel, Spline bases in classical function spaces on compact C^∞ manifolds: Part I and II, Studia Mathematica (1983), 1–58 and 95–136.

[Co] A. Cohen, Numerical Analysis of Wavelet Methods, Studies in Mathematics and its Applications 32, Elsevier, 2003.

[CDF] A. Cohen, I. Daubechies, J.-C. Feauveau, Biorthogonal bases of compactly supported wavelets, Comm. Pure Appl. Math. **45** (1992), 485–560.

[DDU] S. Dahlke, W. Dahmen, K. Urban, Adaptive wavelet methods for saddle point problems — Optimal convergence rates, SIAM J. Numer. Anal. **40** (2002), 1230–1262.

[D1] W. Dahmen, Stability of multiscale transformations, J. Four. Anal. Appl. **2** (1996), 341–361.

[D2] W. Dahmen, Wavelet and multiscale methods for operator equations, Acta Numerica (1997), 55–228.

[D3] W. Dahmen, Wavelet methods for PDEs – Some recent developments, J. Comput. Appl. Math. **128** (2001), 133–185.
[D4] W. Dahmen, Multiscale and wavelet methods for operator equations, in: Multiscale Problems and Methods in Numerical Simulation, C. Canuto (ed.), C.I.M.E. Lecture Notes in Mathematics 1825, Springer Heidelberg (2003), 31–96.
[DK1] W. Dahmen, A. Kunoth, Multilevel preconditioning, Numer. Math. **63** (1992), 315–344.
[DK2] W. Dahmen, A. Kunoth, Appending boundary conditions by Lagrange multipliers: Analysis of the LBB condition, Numer. Math. **88** (2001), 9–42.
[DK3] W. Dahmen, A. Kunoth, Adaptive wavelet methods for linear–quadratic elliptic control problems: Convergence Rates, SIAM J. Contr. Optim. **43**(5) (2005), 1640–1675.
[DKS] W. Dahmen, A. Kunoth, R. Schneider, Wavelet least squares methods for boundary value problems, SIAM J. Numer. Anal. **39** (2002), 1985–2013.
[DKU] W. Dahmen, A. Kunoth, K. Urban, Biorthogonal spline wavelets on the interval – Stability and moment conditions, Appl. Comput. Harm. Anal. **6** (1999), 132–196.
[DS1] W. Dahmen, R. Schneider, Wavelets with complementary boundary conditions — Function spaces on the cube, Results in Mathematics **34** (1998), 255–293.
[DS2] W. Dahmen, R. Schneider, Composite wavelet bases for operator equations, Math. Comp. **68** (1999), 1533–1567.
[DS3] W. Dahmen, R. Schneider, Wavelets on manifolds I: Construction and domain decomposition, SIAM J. Math. Anal. **31** (1999), 184–230.
[DSt] W. Dahmen, R. Stevenson, Element–by–element construction of wavelets satisfying stability and moment conditions, SIAM J. Numer. Anal. **37** (1999), 319–325.
[Dau] I. Daubechies, Orthonormal bases of compactly supported wavelets, Comm. Pure Appl. Math. **41** (1988), 909–996.
[Gr] P. Grisvard, Elliptic Problems in Nonsmooth Domains, Pitman, 1985.
[J] S. Jaffard, Wavelet methods for fast resolution of elliptic problems, Siam J. Numer. Anal. **29** (1992), 965–986.
[KK] P. Kantartzis, A. Kunoth, A wavelet approach for a problem in electrical impedance tomography formulated by means of a domain embedding method, Manuscript, in preparation.
[Kr] J. Krumsdorf, Finite Element Wavelets for the Numerical Solution of Elliptic Partial Differential Equations on Polygonal Domains, Diploma Thesis, Universität Bonn, 2004.
[K1] A. Kunoth, Wavelet Methods — Elliptic Boundary Value Problems and Control Problems, Advances in Numerical Mathematics, Teubner, 2001.
[K2] A. Kunoth, Wavelet techniques for the fictitious domain—Lagrange multiplier approach, Numer. Algor. **27** (2001), 291–316.
[K3] A. Kunoth, Fast iterative solution of saddle point problems in optimal control based on wavelets, Comput. Optim. Appl. **22** (2002), 225–259.
[K4] A. Kunoth, Adaptive wavelet schemes for an elliptic control problem with Dirichlet boundary control, Numer. Algor. **39**(1-3) (2005), 199–220.
[KS] A. Kunoth, J. Sahner, Wavelets on manifolds: An optimized construction, Math. Comp. **75** (2006), 1319–1349.
[Li] J.L. Lions, Optimal Control of Systems Governed by Partial Differential Equations, Springer, Berlin, 1971.
[MB] J. Maes, A. Bultheel, A hierarchical basis preconditioner for the biharmonic equation on the sphere, IMA J. Numer. Anal. **26**(3) (2006), 563–583.
[MKB] 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) (2007), 206–222.
[O] P. Oswald, On discrete norm estimates related to multilevel preconditioners in the finite element method, in: Constructive Theory of Functions, K.G. Ivanov, P. Petrushev, B. Sendov, (eds.), Proc. Int. Conf. Varna 1991, Bulg. Acad. Sci., Sofia (1992), 203–214.
[Pa] R. Pabel, Wavelet Methods for PDE Constrained Elliptic Control Problems with Dirichlet Boundary Control, Diploma Thesis, Universität Bonn, 2006. doi: 10.2370/236_232
[Stv] R. Stevenson, Locally supported, piecewise polynomial biorthogonal wavelets on nonuniform meshes, Constr. Approx. **19** (2003), 477–508.

[U] K. Urban, Wavelet Methods for Elliptic Partial Differential Equations, Oxford University Press, 2009.

[X1] J. Xu, Theory of multilevel methods, Report AM 48, Department of Mathematics, Pennsylvania State University, 1989.

[X2] J. Xu, Iterative methods by space decomposition and subspace correction, SIAM Review **34**(4) (1992), 581–613.

[Y] H. Yserentant, On the multilevel splitting of finite element spaces, Numer. Math. **49** (1986), 379–412.