Fast Iterative Solution of Elliptic Control Problems in Wavelet Discretization*

Carsten Burstedde Angela Kunoth

September 13, 2005

Abstract

We investigate wavelet methods for the efficient numerical solution of a class of control problems constrained by a linear elliptic boundary value problem where the cost functional may contain fractional Sobolev norms of the control and the state. Starting point is the formulation of the infinite-dimensional control problem in terms of (boundary-adapted biorthogonal spline-)wavelets, involving only ℓ_2 norms of wavelet expansion coefficients (where different norms are realized by a diagonal scaling together with a Riesz map) and constraints in form of an ℓ_2 isomorphism. The coupled system of equations resulting from optimization is solved by an inexact conjugate gradient (CG) method for the control, which involves the approximate inversion of the primal and the adjoint operator using again CG iterations. Starting from a coarse discretization level, we use nested iteration to solve the coupled system on successively finer uniform discretizations up to discretization error accuracy on each level. The resulting inexact CG scheme is a 'fast solver': it is of asymptotic optimal complexity in the sense that the overall computational effort to compute the solution up to discretization error on the finest grid is proportional to the number of unknowns on that grid, a consequence of grid–independent condition numbers of the linear operators in wavelet coordinates.

In the numerical examples we study the choice of different norms and the regularization parameter in the cost functional and their effect on the solution. Moreover, for different situations the performance of the fully iterative inexact CG scheme is investigated, confirming the theoretical results.

Keywords: Control problem, fractional Sobolev norms, elliptic PDE constraints, distributed control, biorthogonal spline–wavelets, optimal preconditioning, fast iterative method, inexact conjugate gradient (CG) method, nested iteration, parameter studies.

MSC (2000) classification: 65K10, 65N99, 93B40.

1 Introduction

The efficient numerical solution of a single elliptic linear partial differential equation (PDE) has been the subject of numerous studies over the past decades. Specifically, preconditioners based on multigrid [Br, Brm, Ha2], multilevel [BPX, DK1, O] or wavelet approaches

*This work was supported by the Deutsche Forschungsgemeinschaft (SFB 611) at the Universität Bonn.

[DK1] were developed for iteratively solving the linear system resulting from discretization on uniform grids. These preconditioners are asymptotically optimal in the sense that the linear system can be solved in an amount of arithmetic operations that is *linear* in the number of unknowns, see e.g. [Br, BCD, D2, DK1, Ha2, O], motivating the terminology 'fast (iterative) solvers'.

For control problems constrained by linear elliptic PDEs, the optimality conditions lead to a coupled *system* of PDEs involving additional unknowns for which it is even more tempting to exploit multiscale preconditioners. Multigrid methods on uniform grids were proposed early in [Ha1] and then more recently in [Bo, DMSch], taking into account specific requirements stemming from the control framework. Also domain decomposition methods have been proposed recently for preconditioning such control problems [HN].

Employing wavelets for the solution of PDE-constrained control problems offers several interesting perspectives. Firstly, wavelets form Riesz bases for a whole range of Sobolev norms. In the objective functional balancing data fidelity and regularization, such a variety of norms may be used as an additional *modeling tool*. In particular, fractional Sobolev norms (like the $H^{1/2}$ norm on part of the boundary which appears as the natural norm for problems with Dirichlet boundary control [K1]) may be considered whose realization in conventional settings poses severe difficulties. Secondly, posed in an appropriate framework, the resulting coupled system of linear equations involving state, costate and control as variables is *well–conditioned* in ℓ_2, addressing the above mentioned point of optimal preconditioning and entailing that the convergence speed of iterative solvers like the conjugate gradient method does not depend on the discretization. A third motivation for using wavelets is the built–in potential to locally adapt discretizations to singularities in the data, the coefficients or the domain. Techniques from nonlinear approximation theory have been used for proving convergence of adaptive schemes including optimal convergence rates for linear and nonlinear PDEs in [CDD1, CDD2] and for linear–quadratic elliptic control problems in [DK2, K2]. Further aspects on employing wavelets for the solution of PDEs (including stability issues for stationary saddle point problems and evolution problems such as conservation laws) and related mathematical concepts may be found in the surveys [BCD, Co, CDD3, D2, D3]. For PDE-constrained control problems, adaptive schemes based on finite elements were proposed in [BKR], where, however, so far neither convergence proofs nor rates are available.

As wavelets provide a very powerful tool from an analysis point of view, it is not surprising that their actual construction on bounded and complicated domains requires more sophisticated techniques and that implementations lag behind, in particular, for systems of PDEs such as PDE–constrained control problems. Past realizations have shown that there is ample room for improvements. This concerns, in particular, the early adaptations of the construction of biorthogonal spline–wavelets on the real line based on Fourier techniques [CDF] to bounded intervals and their effect on the L_2 stability constants of the wavelet bases [DKU1, DKU2] (see the discussion at the end of Section 3.1.1 on our particular choice of wavelets). This in turn directly affects the eigenvalues of the resulting stiffness matrices. The absolute numbers of their (in terms of discretization levels) already uniformly bounded condition numbers can be further reduced by taking into account a more sophisticated diagonal scaling as induced by a norm equivalence of the form (3.3)

below, namely, involving the structure of the energy inner product defined by the elliptic boundary value problem [Ba]. An additional substantial improvement concerns inexpensive local basis transformations in terms of singular value decompositions of selected parts of the representation of the elliptic operator [Bu]. The discretizations used here are based on tensor products of one–dimensional wavelets which enables us to solve problems in arbitrary spatial dimensions as long as storage permits.

The focus of this paper is the full iterative solution of PDE–constrained elliptic control problems in wavelet discretization on uniform grids using *conjugate gradient (CG)* schemes. We restrict ourselves to the simplest class of linear–quadratic elliptic control problems with *distributed* control since then the constraints can still be formulated weakly in terms of a single elliptic PDE. We derive a *fully iterative scheme* which consists of an *inexact CG scheme* for the control as an outer iteration in which two systems for the state and the costate are solved approximately in interior iterations by CG schemes. In addition, we use a nested iteration strategy, starting on some coarse refinement level and iterating up to matching discretization error accuracy on each level before prolongating all quantities to the next higher level. The overall scheme is shown to be of optimal linear complexity in the total number of unknowns on the finest grid which justifies to call it a 'fast' algorithm.

Fully iterative schemes were proposed previously in [K1] for optimally preconditioned systems of saddle point problems arising from boundary control problems on uniform grids where, however, the outer scheme consisted of a gradient method which required adaptation of step size parameters for convergence. The present paper proposes a fully iterative method employing inner and outer CG schemes with optimal preconditioning which does not require the selection of such parameters to guarantee convergence.

We distinguish in this paper two slightly different scenarios of PDE-constrained control problems. In the first class, a specific cost functional results from a physical background of the problem and one mandatorily has to compute the solution to this particular problem. Typically in this case the norms in the cost functional are L_2 norms or first order Sobolev norms. We call this class of problems the *mandatory class*. In the second class, the cost functional is formulated by finding a good compromise between tracking given data and some regularization, thereby resulting in a well-posed optimal control problem. Here the specific selection of the norms and the regularization parameter becomes part of the modeling process itself, see, e.g. [BBDM]. Imposing a Sobolev norm on the control enforces additional smoothness but typically these norms do not have a physically motivated representer, in particular, when the smoothness order is not an integer. In the latter case, there is no unique definition of the Sobolev norms, so that there is some ambiguity in representing these norms, anyway. We call this class of problems the *ambiguous case*.

The topics we study in the numerical examples fall into two categories. The first category concerns the effect of *modeling* the optimal control problem by varying the norms for the state and the control in the objective functional (2.2), the role of equivalent norms, and the choice of the regularization parameter balancing the norms for data fidelity of the state and cost of the control. In particular, smoothness of Sobolev norms provides additional modeling parameters which allow to affect contributions on different length scales and which is, therefore, of more flexibility than a single weight parameter. Such enrichment for the model was beneficial in scattered data fitting [CK] and image processing

problems [CDLL]. The second category of numerical results deals with the performance and convergence history of the fully iterative scheme consisting of inexact CG methods in outer and inner iterations for systems in wavelet coordinates. We view these numerical examples also as prototype problems for wavelet-based adaptive schemes whose numerical performance will be reported elsewhere.

The remainder of this paper is structured as follows. In Section 2, we introduce the class of control problems discussed here and their appropriate weak formulation. In Section 3.1 we recall the basic facts for biorthogonal wavelets and discuss some norm equivalences that are used in the sequel. Section 3.2 is devoted to the formulation of a control problem in wavelet coordinates, followed in Section 3.3 by a derivation of the resulting necessary conditions. Section 4 contains a nested iteration algorithm combined with outer and inner conjugate gradient iterations which is shown to be asymptotically optimal in the number of unknowns on the finest grid. Numerical results discussing quality of solutions, iteration numbers and convergence histories are provided in Section 5.

Throughout the paper, we use the relation $a \sim b$ to express $a \lesssim b$ and $a \gtrsim b$, which means that a can be estimated from above and below by a constant multiple of b independent of all parameters on which a or b may depend.

2 Linear–Quadratic Elliptic Control Problems

2.1 Abstract Formulation

Let y denote the *state* belonging to the *state space* Y which is supposed to be a Hilbert space with topological dual Y' and dual form $\langle\cdot,\cdot\rangle_{Y'\times Y}$, or shortly $\langle\cdot,\cdot\rangle$. In particular, Y will be a (closed subspace of a) Sobolev space. By u we will always denote a *control* which in the case of *distributed* control problems discussed below will have to be in Y'. Often in the objective functional state and control are measured in norms different from the ones of their natural spaces, Y and Y' here. The *observation space* Z for the state is possibly a less regular Sobolev space while the control is measured in a space U with possibly higher regularity than Y'. We assume that there are *continuous embeddings* $Y \hookrightarrow Z$ and $U \hookrightarrow Y'$ such that

$$\|w\|_Z \lesssim \|w\|_Y\,, \quad w \in Y\,, \qquad \|v\|_{Y'} \lesssim \|v\|_U\,, \quad v \in U\,. \tag{2.1}$$

We define a *cost functional* of tracking type in terms of the state and control variables y and u as

$$\mathcal{J}(y,u) := \frac{1}{2}\,\|y - y_*\|_Z^2 + \frac{\omega}{2}\,\|u\|_U^2\,, \tag{2.2}$$

where $y_* \in Z$ are given *observation data* to be matched, and the regularization parameter $\omega > 0$ determines the relative weight of the control term. Let $a(v,v) : Y \times Y \to \mathbb{R}$ be a continuous and Y–elliptic bilinear form, i. e., $a(v,v) \sim \|v\|_Y^2$, $v \in Y$. It defines a linear operator $A : Y \to Y'$ by $\langle Av, w\rangle := a(v,w)$ which is *boundedly invertible* as an operator from Y to Y', i.e.,

$$\|Av\|_{Y'} \sim \|v\|_Y\,, \quad v \in Y\,. \tag{2.3}$$

We consider the following abstract *linear–quadratic control problem* with distributed control as a *reference model* for the subsequent derivation:

(ACP) *For given target data $y_* \in Z$, right hand side $f \in Y'$ and weight parameter $\omega > 0$, minimize the quadratic functional* (2.2) *over $(y,u) \in Y \times U$ subject to the linear constraints $a(y,v) = \langle f+u, v\rangle$ for all $v \in Y$ or, equivalently, subject to the linear operator equation*

$$Ay = f + u. \tag{2.4}$$

2.2 Example Problems

2.2.1 Dirichlet Problem

We always denote by $\Omega \subset \mathbb{R}^n$ a bounded Lipschitz domain with boundary $\partial\Omega$. The prototype for a Dirichlet problem with distributed control can be formulated for $a(v,w) := \int_\Omega \nabla v \cdot \nabla w \, dx$, $Y := H_0^1(\Omega)$ implying $Y' = H^{-1}(\Omega) = (H_0^1(\Omega))'$, which corresponds to the standard weak form of the elliptic boundary value problem

$$\begin{aligned} -\Delta y &= f + u && \text{in } \Omega\,, \\ y &= 0 && \text{on } \partial\Omega\,. \end{aligned}$$

Choices for Z, U which are admissible according to (2.1) are

$$Z := H_{00}^s(\Omega), \quad 0 \le s \le 1, \qquad U = H^{-t}(\Omega) := (H_{00}^t(\Omega))', \quad 0 \le t \le 1, \tag{2.5}$$

where $H_{00}^s(\Omega)$ is the intersection of $H^s(\Omega)$ with those functions whose trivial extension by zero is in $H^s(\mathbb{R}^n)$. The case $s = 1$ or $t = 1$ corresponds to choosing a natural norm for this problem while $s = t = 0$ leads to the classical case of $L_2(\Omega)$ norms [Li].

2.2.2 Neumann Problem

Define $a(v,w) := \int_\Omega (\nabla v \cdot \nabla w + vw) dx$ on $Y := H^1(\Omega)$ and consider as constraint for given data $\tilde{f} \in (H^1(\Omega))'$, $g \in H^{-1/2}(\partial\Omega)$

$$a(y,v) = \langle \tilde{f}, v\rangle + \int_{\partial\Omega} gv \, ds + \langle u, v\rangle \qquad \text{for all } v \in Y\,. \tag{2.6}$$

This equation can be derived from the strong form of the non–homogeneous Neumann problem

$$\begin{aligned} -\Delta y + y &= \tilde{f} + u && \text{in } \Omega\,, \\ \tfrac{\partial y}{\partial n} &= g && \text{on } \partial\Omega\,, \end{aligned}$$

where $\frac{\partial}{\partial n}$ denotes the normal derivative in the direction of the outward normal. Abbreviating $\langle f, v\rangle := \langle \tilde{f}, v\rangle + \int_{\partial\Omega} gv \, ds$, the constraints (2.6) can be rephrased as an operator equation

$$Ay = f + u, \tag{2.7}$$

where A is an isomorphism from Y to Y'. Analogous to (2.5) we can take here

$$Z := H^s(\Omega), \quad 0 \le s \le 1, \qquad U = (H^t(\Omega))', \quad 0 \le t \le 1, \tag{2.8}$$

where again $s = t = 1$ corresponds to choosing the natural norms for y and u.

Control problems involving Neumann boundary control can also be formulated in the present framework, see e.g. [Li, DK2]. Dirichlet boundary controls treated by saddle point formulations in wavelet discretizations where the control is measured in $H^{1/2}(\Gamma)$ have been investigated in [K1]. This paper was motivated by the results in [GL] where an H^1 norm for the control was treated approximately in a finite element context.

3 Control Problems in Wavelet Coordinates

In view of the examples in the previous section, for the class of ambiguous problems we want to consider control problems where the *choice* of the norms in the objective functional is part of the modeling process itself. The abstract control problem (ACP) will be formulated as a (still infinite dimensional) control problem in wavelet coordinates which is posed entirely in ℓ_2, where all norms appearing in the cost functional will be weighted ℓ_2 norms and where the representation of the constraints (2.4) will be well–conditioned in ℓ_2. We briefly recall some basic facts concerning wavelets and norm equivalences between Sobolev spaces and discrete norms in terms of weighted wavelet coefficients.

3.1 Wavelet Bases

3.1.1 Basic Properties

Suppose that there is for each Hilbert space $H \in \{Y, Z, U\}$ a *wavelet basis*

$$\Psi_H := \{\psi_{H,\lambda} : \lambda \in \mathcal{I}_H\} \subset H\,, \tag{3.1}$$

indexed by elements from an infinite set $\mathcal{I}_H$ whose elements λ comprise different information such as the *refinement scale* or *level of resolution* $j =: |\lambda|$ and a spatial location $\mathbf{k} = \mathbf{k}(\lambda) \in \mathbb{Z}^n$. The wavelet basis Ψ_H is to have the following properties. First, it is a *Riesz basis* for H: every $v \in H$ has a unique expansion in terms of Ψ_H,

$$v = \sum_{\lambda \in \mathcal{I}_H} v_\lambda \, \psi_{H,\lambda} =: \mathbf{v}^T \, \Psi_H, \quad \mathbf{v} := (v_\lambda)_{\lambda \in \mathcal{I}_H}, \tag{3.2}$$

and its expansion coefficients satisfy a *norm equivalence*

$$\|\mathbf{v}\| \;\sim\; \|\mathbf{v}^T \Psi_H\|_H, \quad \mathbf{v} \in \ell_2(\mathcal{I}_H), \tag{3.3}$$

where here and in the sequel we will write ℓ_2 norms as $\|\cdot\| := \|\cdot\|_{\ell_2(\mathcal{I}_H)}$. Second, Ψ_H is *local*, that is, $\operatorname{diam}(\operatorname{supp}\psi_{H,\lambda}) \sim 2^{-|\lambda|}$. We will view Ψ_H both as a *collection* of functions as in (3.1) as well as a (possibly infinite) column *vector*. For a countable collection of functions

Θ and some single function σ, the term $\langle \Theta, \sigma \rangle$ denotes the column vector with entries $\langle \theta, \sigma \rangle$, $\theta \in \Theta$. For two collections Θ, Σ, the quantity $\langle \Theta, \Sigma \rangle$ is then a matrix with entries $(\langle \theta, \sigma \rangle)_{\theta \in \Theta,\ \sigma \in \Sigma}$, and for a (possibly infinite) matrix $\mathbf{C}$ one derives $\langle \mathbf{C}\Theta, \Sigma \rangle = \mathbf{C}\langle \Theta, \Sigma \rangle$ and $\langle \Theta, \mathbf{C}\Sigma \rangle = \langle \Theta, \Sigma \rangle \mathbf{C}^T$. By duality, (3.3) is equivalent to the existence of a collection of functions which is *dual* or *biorthogonal* to Ψ_H,

$$\tilde{\Psi}_H := \{\tilde{\psi}_{H,\lambda} : \lambda \in I\!I_H\} \subset H', \quad \langle \Psi, \tilde{\Psi} \rangle = \mathbf{I} \tag{3.4}$$

(with the infinite identity matrix $\mathbf{I}$) which is a Riesz basis for H', that is, for any $\tilde{v} = \tilde{\mathbf{v}}^T \tilde{\Psi}_H \in H'$ one has

$$\|\tilde{\mathbf{v}}\| \ \sim \ \|\tilde{\mathbf{v}}^T \tilde{\Psi}_H\|_{H'}, \tag{3.5}$$

see [D1, D3].

For the problem at hand, 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, and H^s denotes the dual of H^{-s} for $s < 0$). The wavelet basis Ψ_H for H is then obtained from an *anchor basis* $\Psi = \{\psi_\lambda : \lambda \in I\!I = I\!I_H\}$. This basis Ψ is a Riesz basis for $L_2(\Omega)$, that is, Ψ is scaled such that $\|\psi_\lambda\|_{L_2(\Omega)} \sim 1$. Its dual basis $\tilde{\Psi}$ is also a Riesz basis for $L_2(\Omega)$, and 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 $\Psi_s := \{2^{-s|\lambda|}\psi_\lambda : \lambda \in I\!I\} =: \mathbf{D}^{-s}\Psi$ is a Riesz basis for H^s,

$$\|\mathbf{v}\| \sim \ \|\mathbf{v}^T \Psi_s\|_{H^s}, \quad \mathbf{v} \in \ell_2(I\!I), \tag{3.6}$$

holds for each $s \in (-\tilde{\gamma}, \gamma)$ [D1]. Analogously, $\tilde{\Psi}_s := \{2^{s|\lambda|}\,\tilde{\psi}_\lambda : \lambda \in I\!I\} = \mathbf{D}^s\,\tilde{\Psi}$ forms a Riesz basis of H^s for $s \in (-\gamma, \tilde{\gamma})$.

Constructions of wavelet bases with the above properties for parameters $\gamma, \tilde{\gamma} \leq 3/2$ on a bounded Lipschitz domain Ω can be found in [DKU2, DSt] which suffice for the above mentioned examples where the Sobolev indices range between -1 and 1. These bases are further modified in [Bu] to improve the absolute values of the Riesz constants in the norm equivalences (3.3). Out of the variety of different wavelet bases that exist, our preference are biorthogonal spline-wavelets: we can work computationally with local linear combinations of continuous piecewise polynomials while still being able to exploit the full functional analytic background provided by the theory. Multiwavelets provide a different class of wavelets, see, e.g., [DDLY, DGH] which also possesses several nice features for the solution of operator equations, see, e.g. [ABGV].

3.1.2 Norm Equivalences and Riesz Maps

An important feature to establish norm equivalences (3.6) for a whole range $s \in (-\tilde{\gamma}, \gamma)$ of Sobolev spaces is the scaling provided by $\mathbf{D}^{-s}$. However, there are several other norms which are *equivalent* to $\|\cdot\|_{H^s}$ and which may be employed in the cost functional (2.2). Although the quadratic nature of the functional is still maintained, their selection precisely determines the solution of (ACP). On the other hand, for noninteger smoothness order,

there are different ways to define Sobolev spaces, such as using interpolation spaces between integer order cases, or intrinsic norms defined by double integrals, or by extensions combined with Fourier transforms [Ad].

In order to address this mathematical modeling issue, 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 $(\cdot,\cdot)_{L_2}$ (mass matrix) defined as

$$\mathbf{R} := (\Psi, \Psi)_{L_2} = \langle \Psi, \Psi \rangle. \tag{3.7}$$

Expanding Ψ in terms of $\tilde{\Psi}$ and recalling the duality (3.4), this also means

$$\mathbf{I} = \langle \Psi, \tilde{\Psi} \rangle = \langle \langle \Psi, \Psi \rangle \tilde{\Psi}, \tilde{\Psi} \rangle = \mathbf{R} \langle \tilde{\Psi}, \tilde{\Psi} \rangle \qquad \text{or} \qquad \mathbf{R}^{-1} = \langle \tilde{\Psi}, \tilde{\Psi} \rangle. \tag{3.8}$$

Another interpretation is that $\mathbf{R}$ is the matrix performing the change of basis from $\tilde{\Psi}$ to Ψ, that is, $\Psi = \mathbf{R}\tilde{\Psi}$. For any $w = \mathbf{w}^T \Psi \in L_2$, one obtains the identities

$$\|w\|_{L_2}^2 = (\mathbf{w}^T \Psi, \mathbf{w}^T \Psi)_{L_2} = \mathbf{w}^T \langle \Psi, \Psi \rangle \, \mathbf{w} = \mathbf{w}^T \mathbf{R} \mathbf{w} = \|\mathbf{R}^{1/2} \mathbf{w}\|^2 =: \|\hat{\mathbf{w}}\|^2. \tag{3.9}$$

Thus, 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 get from (3.6) with $s = 0$

$$c_0^2 \, \|\mathbf{w}\|^2 \;\le\; \|w\|_{L_2}^2 \;\le\; C_0^2 \, \|\mathbf{w}\|^2. \tag{3.10}$$

Thus, the *condition number* $\kappa(\Psi)$ of the wavelet basis can be derived in terms of the extreme eigenvalues of $\mathbf{R}$,

$$\kappa(\Psi) := \left(\frac{C_0}{c_0} \right)^2 = \frac{\lambda_{\max}(\mathbf{R})}{\lambda_{\min}(\mathbf{R})} = \kappa(\mathbf{R}) \sim 1, \tag{3.11}$$

where $\kappa(\mathbf{R})$ also denotes the spectral condition number of $\mathbf{R}$ and the last relation is assured by the asymptotic estimate (3.10). However, the absolute constants will have an impact on numerical results in specific cases.

Now let the Hilbert space $V \in \{Z, U\}$ determine the metric employed in the cost functional (2.2). Denote by Ψ_V a wavelet basis for V satisfying (R), (L) with a corresponding dual basis $\tilde{\Psi}_V$. The (infinite) Gramian matrix with respect to the inner product $(\cdot,\cdot)_V$ inducing $\|\cdot\|_V$ defined by

$$\check{\mathbf{R}}_V := (\Psi_V, \Psi_V)_V \tag{3.12}$$

will also be called *Riesz operator*. The space L_2 is covered trivially since $\check{\mathbf{R}}_{L_2} = \mathbf{R}$. For any function $v := \mathbf{v}^T \Psi_V \in V$ we then have the identity

$$\|v\|_V^2 = (v, v)_V = (\mathbf{v}^T \Psi_V, \mathbf{v}^T \Psi_V)_V = \mathbf{v}^T (\Psi_V, \Psi_V)_V \, \mathbf{v} = \mathbf{v}^T \check{\mathbf{R}}_V \mathbf{v} = \|\check{\mathbf{R}}_V^{1/2} \mathbf{v}\|^2. \tag{3.13}$$

In general, $\check{\mathbf{R}}_V$ may not be explicitly computable, particularly, when V is a fractional order Sobolev space. Therefore, the question arises how to mimic the effect of measuring in $\|\cdot\|_V$ as close as possible (aside from the issue, however, whether $\|\cdot\|_V$ is indeed the 'right' norm for whatever effect one wants to control in (2.2)).

Again referring to (3.6), we obtain as in (3.11) for Sobolev spaces

$$\kappa(\Psi_s) := \left(\frac{C_s}{c_s}\right)^2 = \frac{\lambda_{\max}(\check{\mathbf{R}}_{H^s})}{\lambda_{\min}(\check{\mathbf{R}}_{H^s})} = \kappa(\check{\mathbf{R}}_{H^s}) \sim 1 \qquad \text{for each } s \in (-\tilde{\gamma}, \gamma). \tag{3.14}$$

Thus, all Riesz operators on the applicable scale of Sobolev spaces are spectrally equivalent. Moreover, comparing (3.14) with (3.11), we get

$$\frac{c_s}{C_0}\,\|\mathbf{R}^{1/2}\mathbf{v}\| \;\le\; \|\check{\mathbf{R}}_{H^s}^{1/2}\mathbf{v}\| \;\le\; \frac{C_s}{c_0}\,\|\mathbf{R}^{1/2}\mathbf{v}\| \tag{3.15}$$

and the following result.

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

$$\|v\|_{H^s} = \|\check{\mathbf{R}}_{H^s}^{1/2}\mathbf{v}\| \sim \|\mathbf{R}^{1/2}\mathbf{v}\| \qquad \textit{for each } s \in (-\tilde{\gamma}, \gamma). \tag{3.16}$$

It should be pointed out that $\mathbf{R}$ and its inverse are directly accessible for computations, while $\check{\mathbf{R}}_{H^s}$ for noninteger s is usually not.

3.1.3 Representation of Linear Operators

Finally, the *wavelet representation* of a linear operator in terms of wavelets can be derived as follows [DK2]. 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 $B : H \to V'$ is a linear operator with dual $B' : V \to H'$ defined by $\langle v, B'w\rangle := \langle Bv, w\rangle$ for all $v \in H$, $w \in V$. Then $Bv = w \in V'$ can be represented as $\mathbf{Bv} = \mathbf{w}$ in terms of the wavelet coefficients $\mathbf{v}$ for v (expanded in Ψ_H) and $\mathbf{w}$ (in terms of $\tilde{\Psi}_V$), where

$$\mathbf{B} := \langle \Psi_V, B\Psi_H\rangle. \tag{3.17}$$

3.2 Control Problems in ℓ_2

We will now consider discrete formulations of the abstract control problem (ACP) in wavelet coordinates. Consider the linear elliptic constraints (2.4). Following the recipe from Section 3.1.3, i.e., expanding $y = \mathbf{y}^T\Psi_Y$ and $u = \check{\mathbf{u}}^T\tilde{\Psi}_Y$, and testing with the elements of Ψ_Y, (2.4) attains the form

$$\mathbf{Ay} = \mathbf{f} + \check{\mathbf{u}}, \tag{3.18}$$

where

$$\mathbf{A} := a(\Psi_Y, \Psi_Y), \qquad \mathbf{f} := \langle \Psi_Y, f\rangle. \tag{3.19}$$

(Later it will be more convenient to work with a scaled version $\mathbf{u}$ of $\check{\mathbf{u}}$, reserving the symbol $\mathbf{u}$ for that purpose.) The ellipticity of A (2.3) together with the Riesz basis property (R) of Ψ_Y imply the following well known fact.

Proposition 3.2 *The matrix* $\mathbf{A}$ *is a boundedly invertible mapping from* $\ell_2(I\!I_Y)$ *onto itself, that is, there exists constants* $0 < c_{\mathbf{A}} \leq C_{\mathbf{A}} < \infty$ *such that for any* $\mathbf{v} \in \ell_2(I\!I_Y)$

$$c_{\mathbf{A}}\|\mathbf{v}\| \;\leq\; \|\mathbf{A}\mathbf{v}\| \;\leq\; C_{\mathbf{A}}\|\mathbf{v}\|. \tag{3.20}$$

Note that (3.20) entails that any finite section of $\mathbf{A}$ is also uniformly well-conditioned.

As for the cost functional (2.2), the question arises whether the problem formulation requires the norms $\|\cdot\|_Z$, $\|\cdot\|_U$ to be evaluated *exactly* (as for the mandatory problem class), or whether it suffices to consider *equivalent norms* which model the effect induced by higher or lower order Sobolev smoothness (as in the ambiguous problem class). Recall that the question of equivalent norms arises anyway as soon as Z or U are different from a Sobolev space with integer smoothness parameter, as there are different ways to define Sobolev spaces for these cases [Ad]. In view of the norm equivalences (3.16) and Proposition 3.1, the following more sophisticated construction of Riesz-type operators is introduced in [Bu].

Definition 3.3 *Define for* $Z \in \{L_2, H^1\}$ $\mathbf{R}_Z := \check{\mathbf{R}}_Z$ *with* $\check{\mathbf{R}}_Z$ *from* (3.12). *For* $Z = H^s$ *and* $s \in (0,1)$ *not an integer, let in view of* (3.13) $\mathbf{R}_Z$ *be constructed such that* $\|\mathbf{R}_{H^s}^{1/2}\mathbf{v}\| = \|v\|_{H^s}$ *if* $v = \mathbf{v}^T\Psi_{H^s}$ *is a constant function. This can be achieved by taking a certain convex combination between the* $\mathbf{R}_Z$ *for the computable cases when* $Z \in \{L_2, H^1\}$. *For non-positive smoothness, we work with a generalization of the inversion in* (3.8) *which has the properties* $\mathbf{R}_U = \check{\mathbf{R}}_U$ *for* $U \in \{L_2, (H^1)'\}$ *and* $\|\mathbf{R}_U^{1/2}\mathbf{v}\| = \|v\|_{(H^t)'}$ *for* $U = (H^t)'$, $t \in (0,1)$, *and constant functions* $v = \mathbf{v}^T\tilde{\Psi}_{H^{-t}}$. *In all other cases, we have* $\|\mathbf{R}_Z\mathbf{v}\| \sim \|\check{\mathbf{R}}_Z\mathbf{v}\|$ *and* $\|\mathbf{R}_U\mathbf{v}\| \sim \|\check{\mathbf{R}}_U\mathbf{v}\|$ *for all* $\mathbf{v} \in \ell_2$.

Although for a constant function a variation of s does not have an influence on $\|v\|_{H^s}$, its absolute value depends on s. The above definition of $\mathbf{R}_Z, \mathbf{R}_U$ therefore contains a calibration such that they exactly represent constants in the sense (3.13). Note that the application of $\mathbf{R}_U$ actually means to solve a well-conditioned linear system.

Let now the state and the control be expanded as in (3.18). In terms of wavelet bases, the canonical injections implied by the embedding (2.1) correspond to a multiplication by a diagonal matrix. That is, denoting by Ψ_Z, Ψ_U wavelet bases for Z, U, let $\mathbf{D}_Z, \mathbf{D}_U$ be such that

$$\Psi_Z = \mathbf{D}_Z\Psi_Y, \qquad \tilde{\Psi}_Y = \mathbf{D}_U\Psi_U. \tag{3.21}$$

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

$$\|\mathbf{D}_Z^{-1}\|, \|\mathbf{D}_U^{-1}\| \leq C. \tag{3.22}$$

For example, for $Y = H^\alpha, Z = H^\beta$, or for $Y' = H^{-\alpha}$, $U = H^{-\beta}$, $0 \leq \beta \leq \alpha$, $\mathbf{D}_Z, \mathbf{D}_U$ have entries $(\mathbf{D}_Z)_{\lambda,\lambda} = (\mathbf{D}_U)_{\lambda,\lambda} = (\mathbf{D}^{\alpha-\beta})_{\lambda,\lambda} = 2^{(\alpha-\beta)|\lambda|}$. Writing $y, y_* \in Z$ as $y = \mathbf{y}^T\Psi_Y = (\mathbf{D}_Z^{-1}\mathbf{y})^T\Psi_Z$ and $y_* = \langle y_*, \tilde{\Psi}_Z\rangle\Psi_Z =: (\mathbf{D}_Z^{-1}\mathbf{y}_*)^T\Psi_Z = \mathbf{y}_*^T$ we obtain, in view of (3.16) and the discussion around Definition 3.3

$$\|y - y_*\|_Z \sim \|\mathbf{R}_Z^{1/2}\mathbf{D}_Z^{-1}(\mathbf{y} - \mathbf{y}_*)\| \tag{3.23}$$

where equality holds for Z a Sobolev space with integer smoothness and for constant functions for all choices of Z. As for representing $\|\cdot\|_U$, recall that u is expanded in the dual basis $\tilde{\Psi}_Y$. Consequently, for $u = \check{\mathbf{u}}^T \tilde{\Psi}_Y$, we have in view of (3.21) and Definition 3.3

$$\|u\|_U \sim \|\mathbf{R}_U^{1/2} \mathbf{D}_U \check{\mathbf{u}}\| =: \|\mathbf{R}_U^{1/2} \mathbf{u}\|, \tag{3.24}$$

that is, $\mathbf{u} = \mathbf{D}_U \check{\mathbf{u}}$ are in fact the expansion coefficients with respect to Ψ_U. Again, for the case of U being a Sobolev space with integer smoothness, or for constant functions in case U is arbitrary, the first equivalence in (3.24) is an equality sign.

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

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

$$\check{\mathbf{J}}(\mathbf{y}, \mathbf{u}) := \frac{1}{2} \|\mathbf{R}_Z^{1/2} \mathbf{D}_Z^{-1} (\mathbf{y} - \mathbf{y}_*)\|^2 + \frac{\omega}{2} \|\mathbf{R}_U^{1/2} \mathbf{u}\|^2 \tag{3.25}$$

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

$$\mathbf{A}\mathbf{y} = \mathbf{f} + \mathbf{D}_U^{-1} \mathbf{u}. \tag{3.26}$$

Remark 3.4 The relation between (DCP) and (ACP) is the following. The constraint (3.26) is the exact representation of (2.4) in wavelet coordinates. For the objective functionals (2.2) and (3.25), there exist positive finite constants $c_J \le C_J$ such that $c_J \check{\mathbf{J}}(\mathbf{y}, \mathbf{u}) \le \mathcal{J}(y, u) \le C_J \check{\mathbf{J}}(\mathbf{y}, \mathbf{u})$ holds for any $y = \mathbf{y}^T \Psi_Y \in Y$, given $y_* = (\mathbf{D}_Z^{-1}\mathbf{y}_*)^T \Psi_Z \in Z$ and any $u = \mathbf{u}^T \Psi_U \in U$ for arbitrary choices of Z, U. (DCP) and (ACP) are *identical* for the cases for which the Riesz maps $\mathbf{R}_Z$, $\mathbf{R}_U$ exactly represent $\|\cdot\|_Z$ or $\|\cdot\|_U$. This concerns the class of *mandatory problems* involving Sobolev norms with integer smoothness order as well as general Sobolev norms of constant functions. In the *ambiguous case* involving general fractional smoothness orders, as pointed out at the beginning of Section 3.1.2, in view of the ambiguity with respect to the particular definition of a Sobolev norm we can equally well start with (DCP) as the control problem to be solved. In this case, problem (ACP) may be viewed as a motivation for the derivation of (DCP).

3.3 Optimality Conditions

Following standard lines from e.g. [Li], the unique minimum for (DCP) can be obtained by solving the first order necessary conditions for $\check{\mathbf{J}}$. In view of (3.20), we invert (3.26) and obtain $\mathbf{y} = \mathbf{A}^{-1}\mathbf{f} + \mathbf{A}^{-1}\mathbf{D}_U^{-1}\mathbf{u}$. Substitution into (3.25) yields a functional depending only on $\mathbf{u}$,

$$\mathbf{J}(\mathbf{u}) := \frac{1}{2} \|\mathbf{R}_Z^{1/2} \mathbf{D}_Z^{-1} \left(\mathbf{A}^{-1}\mathbf{D}_U^{-1}\mathbf{u} - (\mathbf{y}_* - \mathbf{A}^{-1}\mathbf{f}) \right)\|^2 + \frac{\omega}{2} \|\mathbf{R}_U^{1/2} \mathbf{u}\|^2. \tag{3.27}$$

With the abbreviations

$$\mathbf{Z} := \mathbf{R}_Z^{1/2} \mathbf{D}_Z^{-1} \mathbf{A}^{-1} \mathbf{D}_U^{-1}, \qquad \mathbf{G} := -\mathbf{R}_Z^{1/2} \mathbf{D}_Z^{-1} (\mathbf{A}^{-1}\mathbf{f} - \mathbf{y}_*), \tag{3.28}$$

the functional simplifies to

$$\mathbf{J}(\mathbf{u}) = \frac{1}{2}\,\|\mathbf{Z}\mathbf{u} - \mathbf{G}\|^2 + \frac{\omega}{2}\,\|\mathbf{R}_U^{1/2}\mathbf{u}\|^2 \tag{3.29}$$

for which the following can immediately be shown, cf. [K1].

Proposition 3.5 *The functional* $\mathbf{J}$ *is twice differentiable with first and second variation given by*

$$\delta\mathbf{J}(\mathbf{u}) = (\mathbf{Z}^T\mathbf{Z} + \omega\mathbf{R}_U)\mathbf{u} - \mathbf{Z}^T\mathbf{G}, \qquad \delta^2\mathbf{J}(\mathbf{u}) = \mathbf{Z}^T\mathbf{Z} + \omega\mathbf{R}_U. \tag{3.30}$$

Setting furthermore

$$\mathbf{Q} := \mathbf{Z}^T\mathbf{Z} + \omega\mathbf{R}_U\,, \qquad \mathbf{g} := \mathbf{Z}^T\mathbf{G}\,, \tag{3.31}$$

the unique minimizer $\mathbf{u}$ of (3.29) is given by solving $\delta\mathbf{J}(\mathbf{u}) = \mathbf{0}$ or equivalently the system

$$\mathbf{Q}\mathbf{u} = \mathbf{g}\,. \tag{3.32}$$

The following result which follows from (3.20) and (3.22) has been assured in [DK2].

Proposition 3.6 *The 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*

$$c_{\mathbf{Q}}\,\|\mathbf{v}\| \le \|\mathbf{Q}\mathbf{v}\| \le C_{\mathbf{Q}}\,\|\mathbf{v}\|, \qquad \mathbf{v} \in \ell_2. \tag{3.33}$$

Thus, since $\mathbf{Q}$ is a symmetric positive definite (infinite) matrix, a conjugate gradient (CG) scheme using (approximative) finite versions of $\mathbf{Q}$ can be used to solve (3.32), and the convergence speed does not depend on the discretization as the spectral condition number of $\mathbf{Q}$ is uniformly bounded. In order to make such iterative schemes practically feasible, of course, the explicit inversion of $\mathbf{A}$ in the definition of $\mathbf{Q}$ is replaced by an iterative solver. Here it is useful to derive an equivalent formulation of (3.32) which is based on the Lagrangian multiplier mechanism, see e.g. [Z]. In fact, defining for (DCP) the *Lagrangian* introducing the *Lagrange multiplier*, *adjoint variable* or *adjoint state* $\mathbf{p}$,

$$\mathbf{L}(\mathbf{y}, \mathbf{p}, \mathbf{u}) := \check{\mathbf{J}}(\mathbf{y}, \mathbf{u}) + \langle \mathbf{p}, \mathbf{A}\mathbf{y} - \mathbf{f} - \mathbf{D}_U^{-1}\mathbf{u}\rangle\,, \tag{3.34}$$

the first–order Euler–Lagrange equations $\delta\mathbf{L}(\mathbf{y}, \mathbf{p}, \mathbf{u}) = \mathbf{0}$ are

$$\mathbf{A}\mathbf{y} = \quad \mathbf{f} + \mathbf{D}_U^{-1}\mathbf{u}, \tag{3.35a}$$

$$\mathbf{A}^T\mathbf{p} = -\mathbf{D}_Z^{-1}\mathbf{R}_Z\mathbf{D}_Z^{-1}(\mathbf{y} - \mathbf{y}_*) \tag{3.35b}$$

$$\omega\mathbf{R}_U\mathbf{u} = \quad \mathbf{D}_U^{-1}\mathbf{p}\,, \tag{3.35c}$$

see [DK2, K1]. The first system which is just (3.26) will be referred to as the *primal system* or the *state* or *forward equation*. Accordingly, we call (3.35b) the *adjoint* or *dual system*, or the *costate equation*. The third equation (3.35c) is sometimes denoted as the *design equation* and has the following interpretation [DK2].

Proposition 3.7 *Solving for a given control vector* $\mathbf{u}$ *successively* (3.26) *for* $\mathbf{y}$ *and* (3.35b) *for* $\mathbf{p}$*, the* residual *for* (3.32) *is*

$$\mathbf{Q}\mathbf{u} - \mathbf{g} = \omega\mathbf{R}_U\mathbf{u} - \mathbf{D}_U^{-1}\mathbf{p}. \tag{3.36}$$

4 A Nested Iteration–Inexact CG Algorithm

Up to this point we have been dealing with infinite–dimensional matrices and wavelet coefficients as vectors. For the finite–dimensional situation, we derive a fully iterative scheme for (3.35) which is based on a *conjugate gradient (CG) iteration* for (3.32) as an *outer iteration* where each application of $\mathbf{Q}$ is in turn realized by solving the primal and the adjoint system (3.26) and (3.35b) also by a CG method as *inner iterations*. Since the interior systems are only solved up to discretization error accuracy, this procedure may therefore be viewed as an *inexact CG method*.

4.1 Finite Systems

From now on, let the domain be $\Omega = (0,1)^n$. The standard construction of biorthogonal spline wavelets adapted to bounded tensor product domains from [DKU1] has additionally been modified. Although finite sections of the operators $\mathbf{A}$ and $\mathbf{Q}$ have been assured to have uniformly bounded condition numbers on account of (3.20), (3.33), the absolute constants can further be improved by adapting the construction to the underlying *energy inner product*. This has been observed and exploited already in [Ba]. In [Bu], this is carried even a step further in that one can not only improve the absolute constants in the condition numbers of the operators $\mathbf{A}$ and $\mathbf{Q}$, the corresponding transformation can also be implemented efficiently in terms of a linear block transformation directly on the stiffness matrix. In order to choose finite–dimensional trial spaces for the Hilbert spaces under consideration, $H \in \{Y, Z, U\}$, we work here with uniform discretizations which corresponds in the wavelet setting to pick the index set of all indices up to some *highest refinement level* J, i.e., $\mathcal{I}_{J,H} := \{\lambda \in \mathcal{I}_H : |\lambda| \leq J\} \subset \mathcal{I}_H$ satisfying $N_{J,H} := \#\mathcal{I}_{J,H} < \infty$. The representation of operators and and vectors is then derived as in Section 3.2 with respect to this truncated index set which corresponds to deleting all entries that refer to indices λ satisfying $|\lambda| > J$. There is by construction also a *coarsest level* of resolution denoted by j_0.

Remark 4.1 Computationally the representation of an operator according to (3.17) is realized by first setting up the operator in terms of the *generator basis* on the finest level J which consists simply of tensor products of B–Splines (or linear combinations of these near the boundaries). The application of an operator in the *wavelet basis* then consists of applying this sparse matrix together with the Fast Wavelet Transform (FWT) and its transpose which in this multiplicative form needs $\mathcal{O}(N_{J,H})$ arithmetic operations and is therefore asymptotically optimal, see also [Ba, D2, DKU1].

Although the matrix $\mathbf{A}$ arising from the problems in Section 2.2 is symmetric, we continue to write $\mathbf{A}, \mathbf{A}^T$ to distinguish the two operators in (3.35a) and (3.35b). For nonsymmetric $\mathbf{A}$, the CG method for the interior systems used below in `APPLY` would have to be replaced by a GMRES or another scheme for nonsymmetric matrices. In view of the uniformly bounded condition number of $\mathbf{A}$ (3.20), one could in principle even apply the CG method to the normal equations, i.e., to $\mathbf{A}^T\mathbf{A}\mathbf{y} = \mathbf{A}^T(\mathbf{f} + \mathbf{D}_U^{-1}\mathbf{u})$.

4.2 A Basic Conjugate Gradient (CG) Method

Consider a linear system of equations

$$\mathbf{M}\mathbf{q} = \mathbf{z}, \tag{4.1}$$

where $\mathbf{M} \in \mathbb{R}^{N\times N}$ is a symmetric positive definite matrix satisfying

$$c_{\mathbf{M}}\|\mathbf{v}\| \le \|\mathbf{M}\mathbf{v}\| \le C_{\mathbf{M}}\|\mathbf{v}\|, \quad \mathbf{v} \in \mathbb{R}^N, \tag{4.2}$$

for some constants $0 < c_{\mathbf{M}} \le C_{\mathbf{M}} < \infty$ and where $\mathbf{z} \in \mathbb{R}^N$ is some given right hand side. The residual using an approximation $\tilde{\mathbf{q}}$ to $\mathbf{q}$ for (4.1) will be abbreviated as

$$\textsc{Res}(\tilde{\mathbf{q}}) := \mathbf{M}\tilde{\mathbf{q}} - \mathbf{z}. \tag{4.3}$$

We will employ a basic conjugate gradient (CG) method described next that iteratively computes an approximate solution $\mathbf{q}_K$ to (4.1) with given initial vector $\mathbf{q}_0$ and given tolerance $\varepsilon > 0$ such that

$$\|\mathbf{M}\mathbf{q}_K - \mathbf{z}\| = \|\textsc{Res}(\mathbf{q}_K)\| \le \varepsilon, \tag{4.4}$$

where K denotes the number of iterations used. Later ε will be specified depending on the discretization for which (4.1) is set up. The following scheme `CG` contains a routine `APPLY`$(\eta_k, \mathbf{M}, \mathbf{d}_k)$ which for $\mathbf{M} = \mathbf{A}, \mathbf{A}^T$ is simply the matrix–vector multiplication $\mathbf{M}\mathbf{d}_k$. Otherwise, it approximately computes this product up to a tolerance $\eta_k = \eta_k(\varepsilon)$ depending on ε.

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

(I) Set $\mathbf{d}_0 := \mathbf{z} - \mathbf{M}\mathbf{q}_0$ and $\mathbf{r}_0 := -\mathbf{d}_0$. Let $k = 0$.

(II) While $\|\mathbf{r}_k\| > \varepsilon$

$$\begin{array}{llllll} \mathbf{m}_k & := & \texttt{APPLY}(\eta_k(\varepsilon), \mathbf{M}, \mathbf{d}_k) & \alpha_k & := & \dfrac{(\mathbf{r}_k)^T\mathbf{r}_k}{(\mathbf{d}_k)^T\mathbf{m}_k} \\ \mathbf{q}_{k+1} & := & \mathbf{q}_k + \alpha_k\mathbf{d}_k & \mathbf{r}_{k+1} & := & \mathbf{r}_k + \alpha_k\mathbf{m}_k \\ \beta_k & := & \dfrac{(\mathbf{r}_{k+1})^T\mathbf{r}_{k+1}}{(\mathbf{r}_k)^T\mathbf{r}_k} & \mathbf{d}_{k+1} & := & -\mathbf{r}_{k+1} + \beta_k\mathbf{d}_k \quad k := k+1 \end{array}$$

(III) Set $K := k - 1$. (4.5)

The routine `CG` computes the residual up to the stopping criterion ε. The error in the solution itself is therefore multiplied by $\|\mathbf{M}^{-1}\| = c_{\mathbf{M}}^{-1}$, that is,

$$\|\mathbf{q} - \mathbf{q}_K\| = \|\mathbf{M}^{-1}(\mathbf{z} - \mathbf{M}\mathbf{q}_K)\| \le \|\mathbf{M}^{-1}\|\,\|\textsc{Res}(\mathbf{q}_K)\| \le \varepsilon\, c_{\mathbf{M}}^{-1}. \tag{4.6}$$

We will also need to employ the scheme `APPLY` for $\mathbf{M} = \mathbf{Q}$. Since $\mathbf{Q}$ is not explicitly accessible we resort to iteratively solving interior systems to evaluate $\mathbf{A}^{-1}$ and $\mathbf{A}^{-T}$. Recall the representation (3.36) for $\mathbf{Q}\mathbf{d} - \mathbf{g}$ in terms of the solution of (3.35), i.e., $\mathbf{Q}\mathbf{d} = \mathbf{g} + \textsc{Res}(\mathbf{d})$. As the right hand side $\mathbf{g}$ also contains applications of $\mathbf{A}^{-1}$ and $\mathbf{A}^{-T}$, $\mathbf{g}$ is approximated by applying interior conjugate gradient iterations up to stopping criterion ζ. The following algorithm describes the approximate computation of $\mathbf{g}$.

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

(I) CG $[\frac{c_{\mathbf{A}}}{2C}\frac{c_{\mathbf{A}}}{C^2C_0^2}\zeta, \mathbf{0}, \mathbf{A}, \mathbf{f}] \to \mathbf{g}_1$

(II) CG $[\frac{c_{\mathbf{A}}}{2C}\zeta, \mathbf{0}, \mathbf{A}^T, -\mathbf{D}_Z^{-1}\mathbf{R}_Z\mathbf{D}_Z^{-1}(\mathbf{g}_1 - \mathbf{y}_*)] \to \mathbf{g}_2$

(III) $\mathbf{g}_\zeta := \mathbf{D}_U^{-1}\mathbf{g}_2$.

Here the tolerances used within the two conjugate gradient methods depend on the constants $c_{\mathbf{A}}, C, C_0$ from (3.20), (3.22) and (3.10), respectively. Note that, 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 is solved more accurately than the adjoint system in step (II).

Proposition 4.2 *The result* $\mathbf{g}_\zeta$ *of* RHS$[\zeta, \mathbf{A}, \mathbf{f}, \mathbf{y}_*]$ *satisfies upon completion*

$$\|\mathbf{g}_\zeta - \mathbf{g}\| \le \zeta. \tag{4.7}$$

Proof: To confirm (4.7), recalling the definition (3.31) of $\mathbf{g}$, one has by step (III) and step (II)

$$\begin{aligned} \|\mathbf{g}_\zeta - \mathbf{g}\| &\le \|\mathbf{D}_U^{-1}\|\,\|\mathbf{g}_2 - \mathbf{D}_U\mathbf{g}\| \\ &\le C\|\mathbf{A}^{-T}\|\,\|\mathbf{A}^T\mathbf{g}_2 - \mathbf{D}_Z^{-1}\mathbf{R}_Z\mathbf{D}_Z^{-1}(\mathbf{A}^{-1}\mathbf{f} - \mathbf{g}_1 + \mathbf{g}_1 - \mathbf{y}_*)\| \\ &\le \frac{C}{c_{\mathbf{A}}}\left(\frac{c_{\mathbf{A}}}{2C}\zeta + \|\mathbf{D}_Z^{-1}\mathbf{R}_Z\mathbf{D}_Z^{-1}(\mathbf{A}^{-1}\mathbf{f} - \mathbf{g}_1)\|\right). \end{aligned} \tag{4.8}$$

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

$$\begin{aligned} \|\mathbf{g}_\zeta - \mathbf{g}\| &\le \frac{C}{c_{\mathbf{A}}}\left(\frac{c_{\mathbf{A}}}{2C}\zeta + C^2C_0^2\,\|\mathbf{A}^{-1}\|\,\|\mathbf{f} - \mathbf{A}\mathbf{g}_1\|\right) \\ &\le \frac{C}{c_{\mathbf{A}}}\left(\frac{c_{\mathbf{A}}}{2C}\zeta + \frac{C^2C_0^2}{c_{\mathbf{A}}}\frac{c_{\mathbf{A}}}{2C}\frac{c_{\mathbf{A}}}{C^2C_0^2}\zeta\right) \quad = \zeta. \end{aligned} \tag{4.9}$$

□

As for the computation of an approximation $\mathbf{m}_\eta$ to the matrix–vector product $\mathbf{Qd}$, we employ the following routine which needs in the last step an appropriate approximation for $\mathbf{g}$.

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

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

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

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

Note that the tolerances differ only slightly from those in the routine RHS, although the ratio between the tolerances in step (I) and (II) is the same, namely, $c_{\mathbf{A}}(CC_0)^{-2}$. This will become clear in the proof of the following result.

Proposition 4.3 *The result* $\mathbf{m}_\eta$ *of* APPLY $[\eta, \mathbf{Q}, \mathbf{d}]$ *satisfies*

$$\|\mathbf{m}_\eta - \mathbf{Q}\mathbf{d}\| \le \eta. \tag{4.10}$$

Proof: Let us confirm that the choice of the stopping criteria in steps (I) and (II) indeed yields (4.10). Denote by $\mathbf{y_d}$ the exact solution of (3.35a) with $\mathbf{d}$ in place of $\mathbf{u}$ on the right hand side, and by $\mathbf{p_d}$ the exact solution of (3.35b) with $\mathbf{y_d}$ on the right hand side. Then we have by step (III) and (3.36) combined with (3.10) and (3.22)

$$\begin{aligned}\|\mathbf{m}_\eta - \mathbf{Q}\mathbf{d}\| &= \|\mathbf{g}_{\eta/3} - \mathbf{g} + \omega\mathbf{R}_U\mathbf{d} - \mathbf{D}_U^{-1}\mathbf{p}_\eta \ - (\mathbf{Q}\mathbf{d} - \mathbf{g})\| \\ &\le \frac{1}{3}\eta + \|\omega\mathbf{R}_U\mathbf{d} - \mathbf{D}_U^{-1}\mathbf{p}_\eta - (\omega\mathbf{R}_U\mathbf{d} - \mathbf{D}_U^{-1}\mathbf{p_d})\| \\ &\le \frac{1}{3}\eta + C\|\mathbf{p_d} - \mathbf{p}_\eta\|.\end{aligned} \tag{4.11}$$

Denoting by $\hat{\mathbf{p}}$ the exact solution of (3.35b) with $\mathbf{y}_\eta$ on the right hand side, we have $\mathbf{p_d} - \hat{\mathbf{p}} = -\mathbf{A}^{-T}\mathbf{D}_Z^{-1}\mathbf{R}_Z\mathbf{D}_Z^{-1}(\mathbf{y_d} - \mathbf{y}_\eta)$. It follows by (3.20), (3.10) and (3.22) that

$$\|\mathbf{p_d} - \hat{\mathbf{p}}\| \le \frac{C^2C_0^2}{c_\mathbf{A}}\|\mathbf{y_d} - \mathbf{y}_\eta\| \le \frac{1}{3C}\eta, \tag{4.12}$$

where the last estimate follows by the choice of the threshold in step (I). Combining (4.11) and (4.12) together with (4.7) and the stopping criterion in step (II) readily confirms the assertion,

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

□

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 (4.1), e.g., in [ES, BoF, GY]. In particular, the results in [ES] yield that for the system $\mathbf{Q}\mathbf{u} = \mathbf{g}$ one can estimate the difference between the actually computed residual $\mathbf{r}_k$ in CG $[\varepsilon, \mathbf{q}_0, \mathbf{Q}, \mathbf{g}]$ and $\mathrm{RES}(\mathbf{u}_k) = \mathbf{Q}\mathbf{u}_k - \mathbf{g}$ as

$$\|\mathbf{r}_k - \mathrm{RES}(\mathbf{u}_k)\| \le C_\mathbf{Q}\sum_{i=0}^{k-1}\eta_i|\alpha_i|\,\|\mathbf{d}_i\|, \tag{4.13}$$

where $C_\mathbf{Q}$ stems from (3.33) and $\alpha_i, \mathbf{d}_i$ are defined in (4.5). In [BoF], the thresholds $\eta_i = \eta_i(\varepsilon)$ for the inner iterations in $\mathtt{APPLY}(\eta_i, \mathbf{Q}, \mathbf{d}_i)$ are selected as $\eta_i = \varepsilon(|\alpha_i|\,\|\mathbf{d}_i\|)^{-1}$ which entails that they may be relaxed as $\|\mathbf{d}_i\|$ becomes smaller. In the scheme used here, we have nevertheless chosen the η_i to be proportional to the outer accuracy ε incorporating a safety factor accounting for the values of α_i and $\|\mathbf{d}_i\|$. In agreement with the literature cited above, we get a linear dependence of the total computing time on the number of unknowns.

Finally, we are in the position to formulate the full nested iteration strategy which employs outer and inner CG routines as follows. The scheme starts at the coarsest level of resolution j_0 with the exact $\mathbf{u}^{j_0}$ and progressively solves (3.32) with respect to each level j until the norm of the current residual is below the discretization error on that level. Since in wavelet coordinates $\|\cdot\|$ corresponds to the energy norm and since 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 which is controlled by a fixed positive constant ν. 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 (3.32) is computed exactly by QR decomposition. Subsequently, the results from level j are prolongated onto the next higher level $j+1$ which in the wavelet setting is accomplished by simply padding with zeros. Since wavelet coordinates have the character of differences, this prolongation corresponds to the exact representation in higher resolution wavelet coordinates. In summary, we formulate the *Nested–Iteration–Incomplete–Conjugate–Gradient* (NIICG) Algorithm as follows.

NIICG $[\mathbf{f}, \mathbf{y}_*, J] \to \mathbf{u}^J$

(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 (3.28).
(2) Compute solution $\mathbf{u}^{j_0}$ of (3.32) by QR decomposition.

(II) While $j < J$

(1) Prolongate $\mathbf{u}^j \to \mathbf{u}_0^{j+1}$ by padding with zeros, set $j := j+1$.
(2) Compute right hand side using RHS $[\nu\, 2^{-(d-1)j}, \mathbf{A}, \mathbf{f}^j, \mathbf{y}_*^j] \to \mathbf{g}^j$.
(3) Compute solution of (3.32) using CG $[\nu\, 2^{-(d-1)j}, \mathbf{u}_0^j, \mathbf{Q}, \mathbf{g}^j] \to \mathbf{u}^j$.

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

Remark 4.4 On account of (3.20) and (3.33), 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 $\mathcal{O}(J) = \mathcal{O}(-\log \varepsilon_J)$ iterations to achieve the prescribed discretization error accuracy $\varepsilon_J := \nu\, 2^{-(d-1)J}$. As each application of $\mathbf{A}$ and $\mathbf{Q}$ requires by Remark 4.1 $\mathcal{O}(N_J)$ operations, the solution of (3.32) by CG iterating only on the finest level requires $\mathcal{O}(J\, N_J)$ arithmetic operations.

Proposition 4.5 *If the residual* (3.36) *is computed on each level j up to discretization error proportional to $2^{-(d-1)j}$ 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.*

The argumentation which was briefly sketched in [DKS] is as follows. Continuing in the line of Remark 4.4, nested iteration allows to get rid of the factor J in the total amount of operations: starting with the exact solution on the coarsest level j_0, one needs only a fixed amount of iterations to reduce the error up to ε_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 ε_j. 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 Numerical Results

Now we present several numerical results on the domain $\Omega = (0,1)^n$ which address modeling issues such as the influence of the choice of the parameters in the cost functional (2.2) or (3.25) on the quality of the solution as well as iteration and convergence histories of `NIICG` in up to three spatial dimensions.

For all experiments, the wavelets employed here are tensor products of the biorthogonal spline wavelets adapted to an interval of primal order $d = 2$ which are linear combinations of piecewise linear B–Splines and duals of order $\tilde{d} = 4$ as constructed in [DKU1]. These bases allow for norm equivalences (3.6) in the range $\gamma \in (-3/2, 3/2)$, covering, in particular $H^1(\Omega)$ and $H^{-1}(\Omega)$. The bases are further adapted to the energy inner product induced by the elliptic operator, a procedure which has been optimized in [Bu]. The coarsest resolution level $j_0 = 3$ is specified by the construction of the boundary–adapted piecewise linear wavelets on the primal side while the highest level J is only restricted by storage for higher dimensions. Throughout the experiments, instead of employing diagonal matrices $\mathbf{D}^{-s}$ with entries $(\mathbf{D}^{-s})_{\lambda,\lambda} = 2^{-s|\lambda|}$ as described in Section 3.2, we work with the entries of the diagonal of the stiffness matrix in wavelet coordinates, $(a(\psi_\lambda, \psi_\lambda))^{-s/2}$, since they provide better absolute numbers.

5.1 Modeling

We first discuss modeling issues, namely, the general behavior of the quality of $\mathbf{y}$ and $\mathbf{u}$ in view of different choices of the norms in the cost functional (3.25). We show examples in one dimension since they reveal most transparently the interplay between smoothness for observation and control.

The first example concerns the formulation of the control problem in wavelet coordinates (DCP) in a simple case where the exact solution can be determined so that we can investigate the question of equivalent norms in the objective functional between the L_2 and a weighted ℓ_2 norm in terms of Riesz maps. We consider the control problem with distributed control constrained by the Neumann problem from Section 2.2.2 with

homogeneous boundary condition $g \equiv 0$ and *constant* data $\tilde{f} = f$ and y_*. In this case, the state, adjoint state and control are constant and the solution of the correspondingly reduced system of necessary conditions

$$y = f + u, \qquad p = -(y - y_*), \qquad \omega u = p \tag{5.1}$$

is given by

$$y = \tfrac{y_* + \omega f}{1+\omega}, \qquad u = \tfrac{y_* - f}{1+\omega}, \qquad p = \omega u. \tag{5.2}$$

In view of the constant solution, the smoothness parameters s, t in (2.2), (2.8) have no influence on the solution. To confirm this, we solve the system in wavelet coordinates (3.35) for $f \equiv 1$ and $y_* \equiv -1$ and various choices of Z and U and vary the regularization parameter ω between 0 and 1. The resulting values for y, p, u are independent of Z and U and are displayed in Figure 1. We observe that we exactly reproduce these constants, confirming that the construction of Riesz maps in Definition 3.3 has the desired effect.

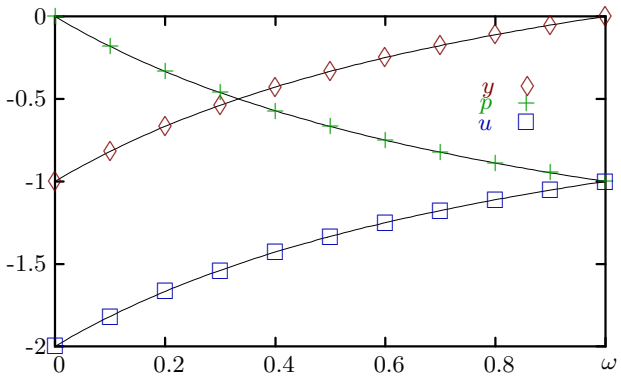

Figure 1: Solution of the control problem with Neumann constraints from Section 2.2.2 corresponding to (5.1) with constant data $f \equiv 1$, $y_* \equiv -1$ for weight parameter $\omega \in [0, 1]$ displayed on the x-axis. The exact solutions y, p, u according to (5.2) are given by the continuous lines. The corresponding discrete symbols are the results of solving (3.35) for these values of ω. For instance, for $\omega = 1$ we obtain $y \equiv 0$ and $p = u \equiv 0$, or for the degenerate case $\omega = 0$ we get $y \equiv -1$, $p \equiv 0$ and $u \equiv -2$.

The graphics for the next univariate examples are displayed for resolution level $j = 8$. The figures always show the state y on the left and the control u on the right. In the next example we solve again the control problem with the constraints from Section 2.2.2 with homogeneous Neumann boundary conditions but this time for a non–constant right hand side $f(x) := 1 + 2.3\exp(-15|x - \frac{1}{2}|)$ ('peak data'). The target function is $y_* \equiv 0$. We fix $\omega = 1$ and study the behavior of the functions y, u when varying the observation space $Z = H^s(\Omega)$ for fixed control space $U = (H^t(\Omega))'$ for $t = 0$ in Figure 2 and for $t = s$ in Figure 3. The lines in each of the figures correspond to the choice displayed within the figures in the same order. We see that in Figure 2 the variations in y and u are rather marginal for fixed U while in Figure 3 a variation of both s and t produce large amplitudes for u for s and t close to 1. The small differences in the results for y are due to a smooth y so that enforcement of increased Sobolev regularity has only very little effect.

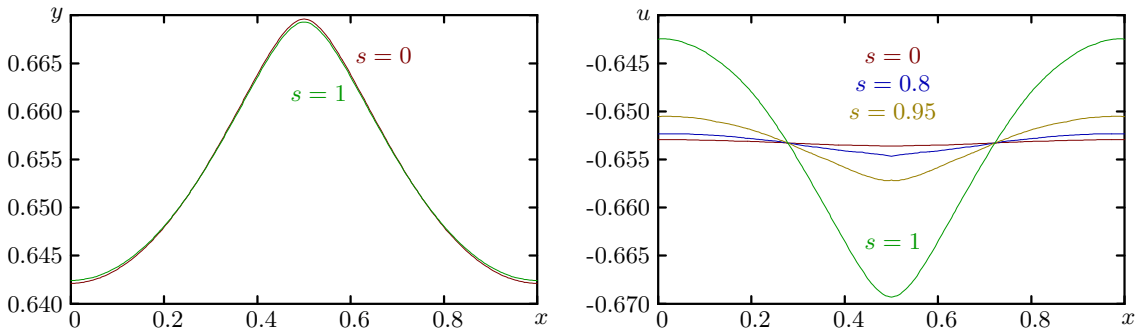

Figure 2: Solution of the control problem with Neumann constraints for peak data f and $y_* \equiv 0$ and norms in the objective functional equivalent to $U = L_2(\Omega)$, $Z = H^s(\Omega)$ and different values of s; state y (left) and control u (right).

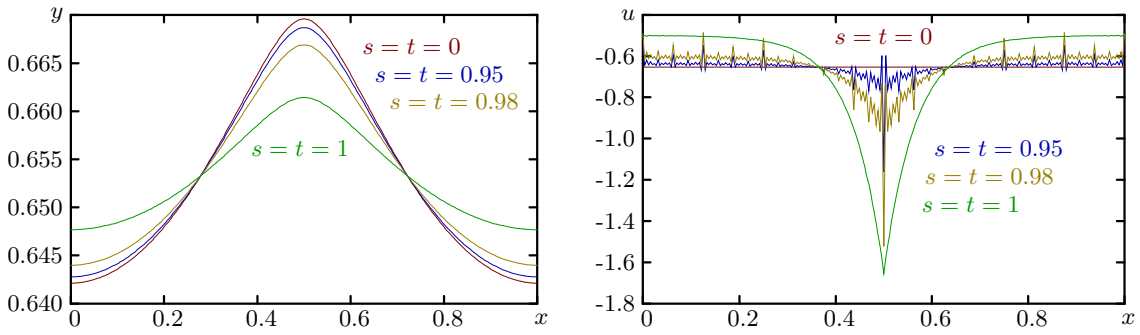

Figure 3: Solution of the control problem with Neumann constraints for peak data f, $y_* \equiv 0$, $s = t$.

For the case $s = t = 1$ we also display the results when varying the regularization parameter ω in Figure 4. This case corresponds to taking natural norms $Z = H^1(\Omega)$ and $U = (H^1(\Omega))'$ in (2.2) which are represented exactly by $\mathbf{R}_Z, \mathbf{R}_U$ according to Definition 3.3, and $\mathbf{D}_Z = \mathbf{D}_U = \mathbf{I}$. Thus, $\mathbf{Q}$ is stably invertible for any ω (including $\omega = 0$) since $\mathbf{Z}$ has a uniformly bounded condition number in this case. Although for the degenerate case $\omega = 0$ the control problem (ACP) or (DCP) is not well-posed, we can consider $\omega = 0$ for the system (3.35). Thus, for $\omega = 0$ and $y_* \equiv 0$, we expect from (3.35) that $p \equiv 0$, $y_* \equiv 0 \equiv y$, entailing $u \equiv -f$. This is exactly the case, as can be seen in the right figure in Figure 4 displaying u: the lowest line corresponds to the case $\omega = 0$ and is the same as $-f$ for the 'peak data' f. We see that in comparison with the previous results, smaller choices of ω enforce proximity to the target state $y_* \equiv 0$.

The results with respect to the choice of the observation and control space are quite different for the control problems with constraints in form of the Dirichlet problem from Section 2.2.1. We have chosen two cases where $y_* \notin Y = H_0^1(\Omega)$. In the first example displayed in Figure 5, we set the target to $y_* \equiv 1$ and choose again f to be the 'peak data'. We see here the results for state y (left) and control u (right) for different values of s for $Z = H^s(\Omega)$ and $U = L_2(\Omega)$. In Figure 6 we have switched the data to $f \equiv 1$ and y_* the 'peak data', again for $t = 0$ ($U = L_2(\Omega)$) and for different choices of s. We see that in

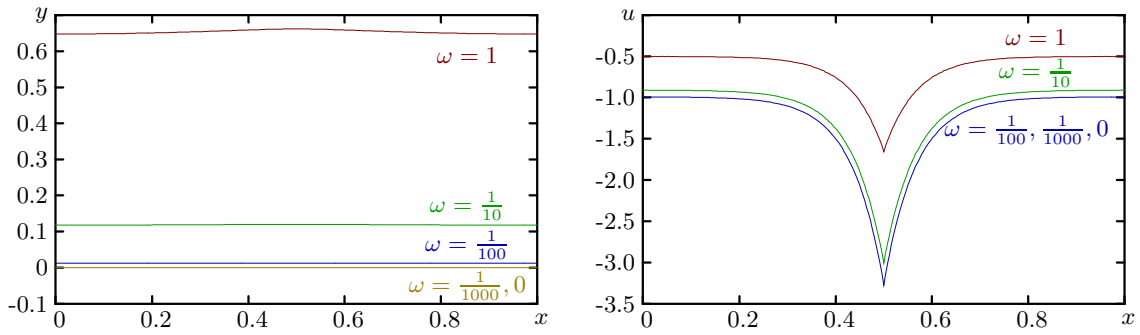

Figure 4: Solution of the control problem with Neumann constraints for peak data f, $y_* \equiv 0$, $s = t = 1$, for different values of ω. In the plots for u, the results for $\omega = 0.01, 0.001$ and 0 are visually not distinguishable. For $\omega = 0$, we obtain $u = -f$.

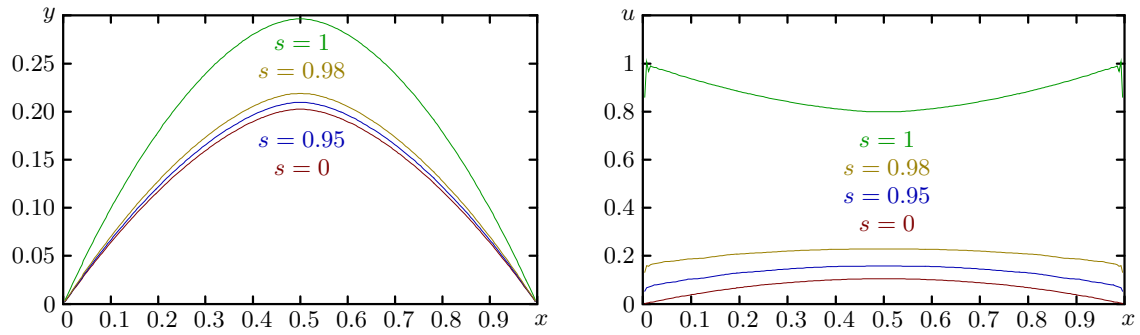

Figure 5: Solution of the control problem with Dirichlet constraints from Section 2.2.1 with peak data f and $y_* \equiv 1$, $t = 0$, for different values of s.

the first case a smoother observation space $Z = H^s(\Omega)$ pushes y towards the target but is deterred from reaching this state by the homogeneous boundary conditions enforced by $Y = H_0^1(\Omega)$. The peak in the target state in the second case has a strong effect on the control for higher s. The oscillations in u for the cases of noninteger s, t stem from the approximate computation of Riesz operators and leave room for further improvement.

Similar results can be obtained also for higher spatial dimensions. We dispense with concrete illustrations here and refer to [Bu].

5.2 Iteration Histories

Finally we give for the above example of the control problem with Neumann constraints a number of numerical results concerning the actual performance of the fully iterative scheme `NIICG` up to the three–dimensional case. Unless stated otherwise, all results are for $\omega = 1$, $y_* \equiv 0$ and the 'peak data' as right hand side.

In order to get an impression of the size of the constants $c_{\mathbf{A}}, C, C_0$ from (3.20), (3.22) and (3.10), respectively, which are appearing in `RHS` and `APPLY`, for instance in the bivariate case we obtain estimates $c_{\mathbf{A}} = 0.054$, $C = 0.153$ and $C_0 = 2.742$. Since we have employed biorthogonal wavelets based on piecewise linear continuous B–Splines, we are here in the

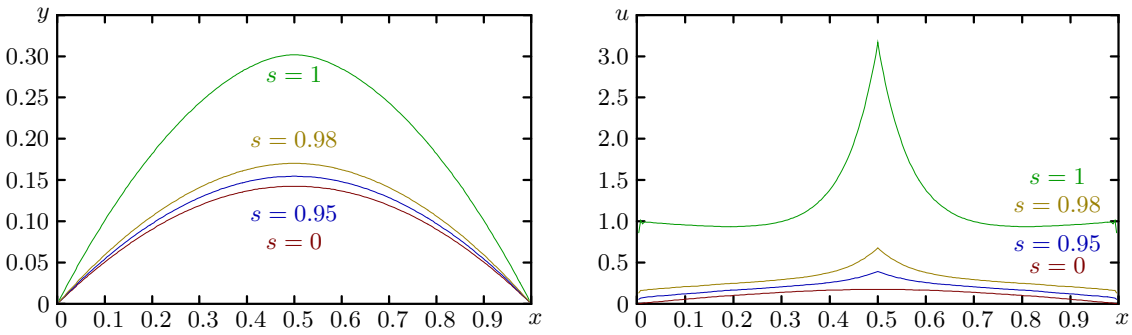

Figure 6: Solution of the control problem with Dirichlet constraints from Section 2.2.1 with data $f \equiv 1$ and peak data y_*, $t = 0$, for different values of s.

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		ex	ex	ex	ex	6.09e-03	1.67e-02	1.17e-06	7.22e-06
4	1.28e-05	1	3	0	4	2.06e-03	9.14e-03	7.16e-06	7.22e-06
5	4.40e-06	2	2	0	4	6.06e-04	4.65e-03	2.29e-06	2.30e-06
6	2.20e-06	2	2	0	4	1.68e-04	2.33e-03	1.38e-06	1.38e-06
7	1.28e-06	2	2	0	4	4.62e-05	1.16e-03	9.33e-07	9.33e-07
8	8.98e-07	2	2	1	4	1.32e-05	5.82e-04	6.07e-07	6.07e-07
9	1.15e-07	3	1	1	3	4.29e-06	2.91e-04	7.58e-08	7.58e-08
10	1.20e-07	1	3	2	4	2.05e-06	1.45e-04	7.58e-08	7.58e-08
11	7.05e-08	2	2	1	4	4.90e-07	7.26e-05	4.26e-08	4.26e-08
12	3.32e-08	2	2	1	4	2.73e-07	3.63e-05	3.22e-08	3.22e-08
13	2.34e-08	2	2	1	4	2.12e-07	1.80e-05	2.50e-08	2.50e-08
14	1.40e-08	2	2	1	4	4.00e-08	8.80e-06	2.16e-08	2.16e-08
15	4.89e-09	3	1	1	3	2.72e-08	3.93e-06	1.22e-08	1.22e-08
16	1.18e-09	5	1	1	3	7.57e-09	7.57e-09	2.56e-09	2.56e-09

Table 1: Iteration history (1D) for the control problem with Neumann constraints for $Z = H^{0.5}(\Omega)$, $U = L_2(\Omega)$.

case of functions of order $d = 2$ so that the stopping criterion for the outer iteration in `NIICG` (relative to $\|\cdot\|$ which corresponds to the energy norm) is chosen to be proportional to 2^{-j}. The parameter ν in `NIICG` is set to $\nu = 0.01$.

We display here several tables with iteration histories for different choices of norms. The setup of all the tables is as follows. Column one contains the resolution level j. The coarsest level is $j_0 = 3$ on which the primal and dual systems are solved exactly, indicated by the entry ex. For $n = 1$, the highest resolution for which the control problem is solved is $J = 16$, in the bivariate case it is $J = 10$ and in 3D $J = 7$ which corresponds to $2 \cdot 10^6$ unknowns in $\mathbf{y}$ and $\mathbf{u}$ each. The second column displays the final value of the residual of the outer `CG` scheme on this level, i.e., $\|\mathbf{r}_K^j\| = \|\text{RES}(\mathbf{u}_K^j)\|$. The next four columns show the number of outer `CG` iterations (#O) for $\mathbf{Q}$ according to the `APPLY` scheme followed by the maximum number of inner iterations for the primal system (#E), the adjoint system (#A) and the design equation (#R). Furthermore, we want to assess how well the *errors* $\mathbf{y}$ and $\mathbf{u}$ are estimated once the *residual* is forced below the target accuracy, that is, estimating the effect of (4.6). To this end, we have computed highly accurate approximations to $\mathbf{y}^J$ and $\mathbf{u}^J$ on the respective finest level of resolution and compared the restriction $R(\mathbf{y}^J)$ to

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		ex	ex	ex	ex	6.10e-03	9.90e-03	2.28e-03	2.36e-03
4	3.84e-06	8	9	8	8	2.06e-03	4.91e-03	1.10e-03	1.11e-03
5	7.15e-07	11	13	10	10	6.13e-04	2.39e-03	3.56e-04	3.57e-04
6	7.73e-07	11	14	9	9	1.67e-04	1.17e-03	1.03e-04	1.03e-04
7	6.83e-07	10	15	9	9	4.52e-05	5.83e-04	2.82e-05	2.82e-05
8	2.35e-07	10	14	9	9	1.26e-05	2.91e-04	7.65e-06	7.65e-06
9	7.23e-08	10	13	8	8	3.92e-06	1.46e-04	2.11e-06	2.11e-06
10	6.62e-08	9	12	7	7	1.64e-06	7.27e-05	6.15e-07	6.15e-07
11	3.70e-08	8	11	8	8	8.53e-07	3.64e-05	1.88e-07	1.88e-07
12	1.51e-08	7	11	7	7	1.60e-07	1.81e-05	6.40e-08	6.40e-08
13	1.07e-08	6	11	8	8	9.46e-08	9.02e-06	2.34e-08	2.34e-08
14	2.98e-09	6	11	7	7	7.38e-08	4.40e-06	7.82e-09	7.82e-09
15	1.42e-09	6	11	7	7	6.36e-08	1.97e-06	2.63e-09	2.63e-09
16	5.16e-10	6	11	7	7	5.37e-09	5.37e-09	3.68e-10	3.68e-10

Table 2: Iteration history (1D) for the control problem with Neumann constraints for $Z = H^1(\Omega)$, $U = (H^1(\Omega))'$.

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		ex	ex	ex	ex	7.01e-03	1.52e-02	1.14e-06	6.99e-06
4	5.30e-06	3	2	0	11	2.33e-03	8.05e-03	1.95e-06	2.15e-06
5	4.80e-06	2	4	0	12	6.72e-04	4.04e-03	1.77e-06	1.78e-06
6	1.43e-06	4	2	0	14	1.78e-04	2.01e-03	6.29e-07	6.30e-07
7	7.17e-07	3	3	0	13	4.67e-05	1.00e-03	3.95e-07	3.95e-07
8	5.20e-07	3	3	1	13	1.27e-05	4.88e-04	2.70e-07	2.70e-07
9	1.95e-07	5	2	1	13	3.77e-06	2.18e-04	6.07e-08	6.07e-08
10	8.66e-08	2	4	2	12	1.00e-06	1.00e-06	5.67e-08	5.67e-08

Table 3: Iteration history (2D) for the control problem with Neumann constraints for $Z = H^{0.5}(\Omega)$, $U = L_2(\Omega)$.

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		ex	ex	ex	ex	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	2.80e-06	1.94e-04
7	3.35e-06	7	12	5	9	4.56e-05	9.73e-04	8.96e-07	1.35e-04
8	2.42e-06	7	11	5	10	1.23e-05	4.74e-04	7.06e-07	8.88e-05
9	1.20e-06	8	11	5	10	3.62e-06	2.12e-04	3.44e-07	5.14e-05
10	4.68e-07	9	10	5	9	9.60e-07	9.60e-07	1.18e-07	1.18e-07

Table 4: Iteration history (2D) for the control problem with Neumann constraints, $Z = H^1(\Omega)$, $U = (H^{0.5}(\Omega))'$.

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		ex	ex	ex	ex	6.60e-03	1.42e-02	2.50e-04	7.11e-04
4	8.49e-08	12	5	1	10	2.24e-03	7.63e-03	5.73e-05	1.06e-04
5	3.99e-08	12	4	1	10	6.42e-04	3.84e-03	1.43e-05	2.02e-05
6	1.96e-08	8	4	2	9	1.73e-04	1.92e-03	3.85e-06	5.36e-06
7	1.06e-08	7	4	2	9	4.60e-05	9.51e-04	1.14e-06	1.54e-06
8	5.67e-09	5	5	3	8	1.28e-05	4.64e-04	3.56e-07	4.49e-07
9	2.15e-09	5	5	3	8	3.47e-06	2.07e-04	1.20e-07	1.38e-07
10	1.07e-09	4	6	4	8	4.31e-07	4.31e-07	5.19e-08	5.19e-08

Table 5: Iteration history (2D) for the control problem with Neumann constraints, $Z = H^{0.5}(\Omega)$, $U = L_2(\Omega)$, $\omega = 0.01$.

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		ex	ex	ex	ex	6.97e-03	1.35e-02	1.12e-06	6.77e-06
4	6.04e-06	6	6	0	39	2.28e-03	6.96e-03	1.65e-06	1.87e-06
5	2.69e-06	7	4	0	52	6.14e-04	3.38e-03	7.04e-07	7.19e-07
6	1.77e-06	5	5	0	53	1.34e-04	1.50e-03	5.24e-07	5.26e-07
7	7.96e-07	7	4	0	52	1.32e-05	1.32e-05	2.70e-07	2.70e-07

Table 6: Iteration history (3D) for the control problem with Neumann constraints for $Z = H^{0.5}(\Omega)$, $U = L_2(\Omega)$.

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		ex	ex	ex	ex	6.82e-03	1.31e-02	1.25e-04	4.40e-04
4	1.84e-05	8	19	9	19	2.24e-03	6.78e-03	4.73e-05	3.51e-04
5	9.66e-06	10	21	10	23	6.03e-04	3.29e-03	1.15e-05	2.49e-04
6	4.08e-06	12	20	10	24	1.31e-04	1.46e-03	2.38e-06	1.50e-04
7	2.04e-06	13	21	9	25	1.31e-05	1.31e-05	5.21e-07	5.21e-07

Table 7: Iteration history (3D) for the control problem with Neumann constraints for $Z = H^1(\Omega)$, $U = (H^{0.5}(\Omega))'$.

level j with $\mathbf{y}^j$ as well as the prolongation $P(\mathbf{y}^j)$ to level J by adding zeroes with $\mathbf{y}^J$, and correspondingly for the control variable. Table 5 contains an example for regularization parameter $\omega = 0.01$; in all other cases it is set to $\omega = 1$. Comparing the results from Table 3 with the results from Table 5, we see that even for small ω the absolute numbers of the iterations stay constant. Their slight increase in the number of outer iterations may account for the fact that also the residual in the second column attains a smaller value.

In all tables, it is firstly confirmed that the number of outer iterations on each level is constant as predicted by the theory in Section 4, and similarly for the maximum number of inner iterations. Thus, the effect of the uniformly bounded condition numbers of the involved operators is clearly visible: a fixed finite number of iterations independent of the resolution suffices to reach discretization error accuracy on that level. However, as it is known, the total numbers increase depending on the spatial dimension n which is also confirmed by the numbers presented here. Moreover, one needs in average more iterations for the primal system as for the adjoint system which is enforced by the smaller absolute stopping criterion. We also see in all tables that the stopping criteria for the residuals also bound the errors times a constant factor (which naturally depends on the spatial dimension n). Thus, we can conclude that the influence of the smallest eigenvalue $c_{\mathbf{Q}}^{-1}$ on the error according to (4.6) does not harm the performance of the numerical scheme. Furthermore, by comparing the iteration numbers and errors for different choices in the functional, we see that for the case $U = L_2(\Omega)$ one usually needs the smallest number of absolute iterations.

All numerical results have been obtained on a Pentium IV PC with 2.53GHz clock speed and 1GB of memory. Execution times range from 5-20 seconds in the 1D case, less than 10 minutes for the 2D case up to 2-3 hours in three dimensions. This is a somewhat stronger increase than one would expect in view of the growing maximal number of unknowns which are proportional to 2^{nJ}. This increase in computing time depending on the spatial dimension n is caused by a more involved application of the operators as well as by an increase in the absolute values of the condition numbers. However, the latter are still uniformly bounded independent of the level of resolution.

Acknowledgments

The authors would like to thank Wolfgang Dahmen, Roland Pabel and Andreas Wedel for helpful discussions and support. We also thank the anonymous referees for their valuable suggestions.

References

[Ad] R.A. Adams, Sobolev Spaces, Academic Press, 1978.

[ABGV] B. Alpert, G. Beylkin, D. Gines, L. Vozovoi, Adaptive solution of partial differential equations in multiwavelet bases, J. Comput. Phys. 182 (2002), 149-190.

[Ba] T. Barsch, Adaptive Multiskalenverfahren für elliptische partielle Differentialgleichungen – Realisierung, Umsetzung und numerische Ergebnisse (in German), Shaker Verlag, Aachen, 2001.

[BKR] R. Becker, H. Kapp, R. Rannacher, Adaptive finite element methods for optimal control of partial differential equations: Basic concept, SIAM J. Contr. Optim., 39 (2000), 113-132.

[BBDM] T. Binder, L. Blank, W. Dahmen, W. Marquardt, Multiscale concepts for moving horizon optimization, in: *Online Optimization for Large Scale Systems*, M. Grötschel, S.O. Krumke, and J. Rambau (eds.), Springer, Berlin (2001), 341-362.

[Bo] A. Borzi, Multigrid methods for parabolic distributed optimal control problems, J. Comp. Appl. Math., 157 (2003), 365–382.

[BoF] A. Bouras, V. Frayssé, A relaxation strategy for inexact matrix–vector products for Krylov methods, Tech. Rep. TR–PA–00–15, CERFACS, Toulouse, 2000.

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

[Brm] J.H. Bramble, Multigrid Methods, Pitman, 1993.

[BCD] J.H. Bramble, A. Cohen, W. Dahmen, Multiscale Problems and Methods in Numerical Simulations, C. Canuto (ed.), Lecture Notes in Mathematics 1825, Springer, 2003.

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

[Bu] C. Burstedde, Fast Optimised Wavelet Methods for Control Problems Constrained by Elliptic PDEs, Doctorate Thesis, Universität Bonn, September 2005.

[CK] D. Castaño, A. Kunoth, Multilevel regularization of wavelet based fitting of scattered data — Some experiments, Numer. Algor. 39 (1-3), 2005, 81-96.

[CDLL] A. Chambolle, R.A. DeVore, N.–Y Lee, B.J. Lucier. Nonlinear wavelet image processing: Variational problems, compression, and noise removal through wavelet shrinkage. IEEE Trans. Image Process 7 (1998), 319-335.

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

[CDD1] A. Cohen, W. Dahmen, R. DeVore, Adaptive wavelet methods for elliptic operator equations – Convergence rates, Math. Comp., 70 (2001), 27–75.

[CDD2] A. Cohen, W. Dahmen, R. DeVore, Adaptive wavelet schemes for nonlinear variational problems, SIAM J. Numer. Anal. 41 (2003), 1785-1823.

[CDD3] A. Cohen, W. Dahmen, R. DeVore, Adaptive wavelet techniques in numerical simulation, Encyclopedia of Computational Mechanics, E. Stein, R. de Borst, T.J.H. Hughes, eds, John Wiley & Sons, 2004.

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

[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.

[DK1] W. Dahmen, A. Kunoth, Multilevel preconditioning, Numer. Math., 63 (1992), 315–344.

[DK2] W. Dahmen, A. Kunoth, Adaptive wavelet methods for linear–quadratic elliptic control problems, SIAM J. Contr. Optim. 43 (2005), 1640–1675.

[DKS] W. Dahmen, A. Kunoth, R. Schneider, Wavelet least square methods for boundary value problems, SIAM J. Numer. Anal. 39 (2002), 1985–2013.

[DKU1] W. Dahmen, A. Kunoth, K. Urban, Biorthogonal spline wavelets on the interval – Stability and moment conditions, Appl. Comput. Harm. Anal., 6 (1999), 132–196.

[DKU2] W. Dahmen, A. Kunoth, K. Urban, Wavelets in numerical analysis and their quantitative properties, in: Surface Fitting and Multiresolution Methods, A. Le Méhauté, C. Rabut and L.L. Schumaker (eds.), Vanderbilt University Press, Nashville, TN (1997), 93–130.

[DSt] W. Dahmen, R. Stevenson, Element–by–element construction of wavelets satisfying stability and moment conditions, SIAM J. Numer. Anal., 37 (1999), 319–325.

[DDLY] D.L. Donoho, N. Dyn, D. Levin, and T. P.-Y. Yu, Smooth multiwavelet duals of Alpert bases by moment-interpolating refinement, Appl. Comput. Harmon. Anal. 9 (2000), 166-203.

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

[DMSch] Th. Dreyer, B. Maar, V. Schulz, Multigrid optimization in applications, J. Comp. Appl. Maths., 120 (2000), 67–84.

[ES] J. van den Eshof, G.L.G. Sleijpen, Inexact Krylov subspace methods for linear systems, SIAM J. Matr. Anal. Appl. 26 (2004), 125-153.

[GY] G.H. Golub, Q. Ye, Inexact preconditioned conjugate gradient method with inner–outer iteration, SIAM J. Sci. Comput., 21 (2000), 1305–1320.

[GL] M. D. Gunzburger, H. C. Lee, Analysis, approximation, and computation of a coupled solid/fluid temperature control problem, Comp. Meth. Appl. Mech. Engrg., 118 (1994), 133–152.

[Ha1] W. Hackbusch, Fast solution of elliptic control problems, J. Optim. Theory Appl. 31 (1980), 565–581.

[Ha2] W. Hackbusch, Elliptic Differential Equations: Theory and Numerical Treatment, Springer, 1992.

[HN] M. Heinkenschloss, H. Nguyen, Domain decomposition preconditioners for linear-quadratic elliptic optimal control problems, CAAM TR04-20, Rice University, November 2004.

[K1] A. Kunoth, Fast iterative solution of saddle point problems in optimal control based on wavelets, Comput. Optim. Appl., 22 (2002), 225–259.

[K2] A. Kunoth, Adaptive wavelet schemes for an elliptic control problem with Dirichlet boundary control, Numer. Algor. 39 (1-3), 2005, 199-220.

[Li] J.L. Lions, Optimal Control of Systems Governed by Partial Differential Equations, Springer, Berlin, 1971.

[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.), Bulg. Acad. Sci., Sofia (1992), 203–214.

[St] R. Stevenson, Locally supported, piecewise polynomial biorthogonal wavelets on non-uniform meshes, Constr. Approx., 19 (2003), 477–508.

[Z] E. Zeidler, Nonlinear Functional Analysis and its Applications; III: Variational Methods and Optimization, Springer, 1985.

Carsten Burstedde and Angela Kunoth
Institut für Angewandte Mathematik
Wegelerstr. 6, Universität Bonn
53115 Bonn, Germany
E–Mail: {bursted,kunoth}@iam.uni-bonn.de