# ADAPTIVE WAVELET METHODS FOR LINEAR–QUADRATIC ELLIPTIC CONTROL PROBLEMS: CONVERGENCE RATES*

WOLFGANG DAHMEN† AND ANGELA KUNOTH‡

**Abstract.** We propose an adaptive algorithm based on wavelets for the fast numerical solution of control problems governed by elliptic boundary value problems with distributed or Neumann boundary control. A quadratic cost functional that may involve fractional Sobolev norms of the state and the control is to be minimized subject to linear constraints in weak form. Placing the problem into the framework of (biorthogonal) wavelets allows to formulate the functional and the constraints equivalently in terms of $\ell_2$–norms of wavelet expansion coefficients and constraints in form of an $\ell_2$ automorphism. The resulting first order necessary conditions are then derived as a (still infinite) system in $\ell_2$. Applying the machinery developed in [CDD1, CDD2], we propose an adaptive method which can be interpreted as an inexact gradient method, where in each iteration step the primal and the adjoint system needs to be solved up to a prescribed accuracy. In particular, we show that the adaptive algorithm is asymptotically optimal, that is, the convergence rate achieved for computing the solution up to a desired target tolerance is asymptotically the same as the wavelet–best $N$–term approximation of the solution, and the total computational work is proportional to the number of computational unknowns.

**Key words.** Optimal control, elliptic boundary value problem, wavelets, infinite $\ell_2$–system, preconditioning, adaptive refinements, inexact iterations, convergence, convergence rates, optimal complexity.

**AMS subject classifications.** 65K10, 65N99, 93B40.

**1. Introduction.** Recently a new type of adaptive wavelet methods for the numerical solution of a wide class of variational problems have been developed and analyzed in a series of papers [CDD1, CDD2, CDD3]. These methods have been shown to exhibit asymptotically computational complexity in the following sense. If the solution can be approximated (using ideal complete information) by $N$ terms from the underlying wavelet basis with accuracy $\mathcal{O}(N^{-s})$ (in the energy norm) then the scheme recovers for a certain range of decay rates $s$, depending on the wavelet basis, the solution with any desired target accuracy $\varepsilon$ at a computational expense that stays proportional to $\varepsilon^{-1/s}$, uniformly in $\varepsilon$ and matches in this sense the optimal work/accuracy rate of best $N$–*term approximation.*

Moreover, the underlying analysis has lead to a new algorithmic paradigm that can be summarized as follows.

(i) Establish *well posedness* of the underlying variational problem, which is to identify a Hilbert space (energy space) for which the operator induced by the variational problem is boundedly invertible as a mapping from this Hilbert space onto its dual.

(ii) Transform the original problem into an equivalent one that is now well posed in the Euclidean metric $\ell_2$. This is usually done by finding a *wavelet basis* that is a Riesz basis for the energy space.

(iii) Exploit (ii) so as to devise an iterative scheme for the (still infinite dimensional) transformed problem on $\ell_2$ that has a fixed error reduction per step.

*This work has been supported in part by the Deutsche Forschungsgemeinschaft SFB 611, Universität Bonn, and the TMR network "Wavelets in Numerical Simulation".

†Institut für Geometrie und Praktische Mathematik, RWTH Aachen, 52056 Aachen, Germany, dahmen@igpm.rwth-aachen.de, www.igpm.rwth-aachen.de/dahmen

‡Institut für Angewandte Mathematik, Universität Bonn, Wegelerstr. 6, 53115 Bonn, Germany, kunoth@iam.uni-bonn.de, www.iam.uni-bonn.de/∼kunoth.

(iv) Perform the ideal iteration from (iii) approximately by *adaptively* applying the involved operators in wavelet coordinates within suitable dynamically updated accuracy tolerances.

The objective of this paper is to explore the use of such concepts in the context of *optimal control problems* with PDE constraints. We are primarily motivated by the following two aspects. By their very nature such control problems tend to have a rather demanding computational complexity so that the use of schemes that minimize computational complexity is very tempting. The second reason is the fact that since the above paradigm tries to stay with the infinite dimensional well–posed problem as long as possible, it turns out to inherit the stability of the infinite dimensional problem in the following sense. Compatibility conditions on finite dimensional trial spaces that may arise in coupled problems, such as in the form of the LBB condition for saddle point problems, do not arise in the adaptive context, see [CDD2, DDU]. Moreover, the fact that suitable scalings of one and the same wavelet basis form Riesz bases for a whole range of Sobolev spaces allows one to treat in a convenient way a variety of such norms in the objective functional which pose difficulties in conventional settings.

In order to bring out the basic mechanisms, we deliberately confine the discussion to rather simple types of control problems with linear constraints, including Dirichlet and Neumann problems with distributed or Neumann boundary controls. The setting will be described in Section 2 along with some examples that will guide the subsequent developments. While the above paradigm has been developed mainly for PDEs or singular integral equations, the first issue will be to formulate the optimal control problem in a way that allows us to branch into the above road map. This will be done in Section 3 that provides the background for (ii). In Section 4 we briefly collect some relevant facts from [CDD2, CDD3] that will later be used for (iv). One major task in the present context is the formulation of a convergent (ideal) iteration (iii) and a way that makes (iv) feasible. This is the objective of Section 5. Having started out from a rather general setting we will have by then narrowed down step by step requirements on the computational ingredients that will imply optimal complexity at the end and guide the construction of the scheme. The complexity analysis in Section 6 will finally allow us to identify specific evaluation schemes that will be seen to render our adaptive solver for the optimal control problems under consideration to have optimal work/accuracy rates in the above sense. It is perhaps worth noting that the analysis brings out some distinctions between the inherent computational complexity of problems with distributed versus Neumann boundary control.

Throughout the paper, we will employ the following notational conventions, unless specific constants have to be identified: The relation $a \sim b$ stands for $a \lesssim b$ and $a \gtrsim b$, where the latter relation means that $b$ can be estimated from above by a constant multiple of $a$ independent of all parameters on which $a$ or $b$ may depend.

**2. Problems in Optimal Control.** We shall be concerned with the following abstract class of problems in optimal control that will serve as a first simple model for studying adaptive solution concepts in such a context. Several specifications will guide the subsequent analysis.

**2.1. Abstract Linear–Quadratic Control Problems.** Let $Y$ and $Q$ denote the *state* and the *control space*, respectively, which are assumed to be (closed subspaces of) Hilbert spaces, with topological duals $Y', Q'$ and associated dual forms $\langle\cdot,\cdot\rangle_{Y'\times Y}$, $\langle\cdot,\cdot\rangle_{Q'\times Q}$. When there is no risk of confusion we write briefly $\langle\cdot,\cdot\rangle$. In many applications the *states* $y$ are measured in a weaker norm corresponding here to

a Hilbert space $Z$ hosting the observed data $y_*$. In contrast, the regularity imposed on the *control* $u$, represented here by a Hilbert space $U$, is often higher than that required in a natural variational formulation. Thus, we shall assume the validity of the continuous embeddings

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

All norms will be indexed by the respective space notation.

Denoting by $T : Y \to Z$ a continuous linear operator

$$\|Ty\|_Z \lesssim \|y\|_Y, \tag{2.2}$$

mapping states into the observation space $Z$, our objective is to minimize quadratic functionals of the form

$$J(y,u) = \frac{1}{2}\|Ty - y_*\|_Z^2 + \frac{\omega}{2}\|u\|_U^2, \tag{2.3}$$

subject to linear constraints, that will be described next. We shall assume that $a(\cdot,\cdot) : Y \times Y \to \mathbb{R}$ is a bilinear $Y$*–elliptic* form, i.e.,

$$a(v,v) \sim \|v\|_Y^2, \qquad v \in Y. \tag{2.4}$$

It will sometimes be convenient to refer to the linear operator $A : Y \to Y'$ defined by $\langle Ay, v\rangle = a(y,v)$, $v \in Y$.

The last ingredient is a linear continuous operator $E : Q \to Y'$, describing an action on the control.

The abstract *linear–quadratic control problem* can now be formulated as follows.

**(ACP):** *For given observations $y_* \in Z$, a right hand side $f \in Y'$ and a weight parameter $\omega > 0$, minimize the quadratic functional (2.3) over $(y,u) \in Y \times Q$ subject to the linear constraints*

$$a(y,v) = \langle f + Eu, v\rangle, \quad v \in Y. \tag{2.5}$$

REMARK 2.1. *Of course, when the observed data are compatible in the sense that $y_* \equiv TA^{-1}f$, (ACP) has the trivial solution $u \equiv 0$ yielding $J(y,u) \equiv 0$, which can be used to test solution schemes.*

**2.2. Some Examples.** In all the following $\Omega \subset \mathbb{R}^d$ denotes a bounded Lipschitz domain. The choice $Z = U = L_2(\Omega)$ in the functional (2.3) is classical (see [Li]), perhaps partly due to the difficulty of evaluating the norms that could be termed *natural* (such as fractional trace norms) with regard to the underlying variational formulation, namely, the norms $\|\cdot\|_Y, \|\cdot\|_Q$. Here we explicitly allow for employing also natural norms for observing the state $y$, unless, for statistical reasons, measurements are only meaningful in weaker norms such as $L_2$. It will be seen below that Sobolev or even Besov norms on $\Omega$ or (part of) its boundary $\Gamma = \partial\Omega$ for a certain range of regularity scales can be dealt with by our approach.

Although the problems with *distributed control* are perhaps rather of academic nature, they serve as good illustrations for the essential mechanisms.

**2.2.1. Dirichlet Problem with Distributed Control.** In our first example we consider such a distributed control problem with the following identification of the above ingredients:

$$a(v,w) := \int_\Omega \nabla v \cdot \nabla w \, dx, \quad Y = H_0^1(\Omega), \; Q = H^{-1}(\Omega) = Y'. \tag{2.6}$$

This gives rise to constraints whose strong form is given by the standard second order Dirichlet problem with distributed control,

$$\begin{aligned} -\Delta y &= f + u && \text{in } \Omega, \\ y &= 0 && \text{on } \partial\Omega. \end{aligned} \tag{2.7}$$

Admissible choices for $Z, U$, satisfying (2.1), are then

$$Z := H_{00}^s(\Omega), \quad 0 \le s \le 1, \quad U = H^t(\Omega) := (H_{00}^{-t}(\Omega))', \quad -1 \le t \le 0, \tag{2.8}$$

where $H_{00}^s(\Omega)$ consists of those elements in $H^s(\Omega)$ whose trivial extension by zero belongs to $H^s(\mathbb{R}^d)$. Thus, for $s < 1$ the states are measured in a weaker norm while for $t > -1$ additional smoothness is imposed on the control when compared with the natural norms. In particular, the classical case $U = Z = L_2(\Omega)$ is covered. In all these cases the operators $T, E$ are the canonical injections $T = I$, $E = I$, which, for the regularity scales in (2.8), are indeed bounded.

**2.2.2. Neumann Problem with Distributed Control.** Choosing

$$a(v,w) := \int_\Omega (\nabla v \cdot \nabla w + vw) dx, \quad Y := H^1(\Omega), \quad Q = (H^1(\Omega))' = Y', \tag{2.9}$$

(2.4) holds. Denoting by $\gamma$ the trace operator to $\partial\Omega$, mapping functions in $Y = H^1(\Omega)$ to $H^{1/2}(\partial\Omega)$, we consider next the constraint

$$a(y,v) = \langle \tilde{f}, v\rangle + \int_{\partial\Omega} g(\gamma v)\, ds + \langle u, v\rangle \qquad \text{for all } v \in Y \tag{2.10}$$

and for given $\tilde{f} \in Y'$, $g \in H^{-1/2}(\partial\Omega)$. Its strong form is the second order non–homogeneous Neumann problem with distributed control

$$\begin{aligned} -\Delta y + y &= \tilde{f} + u && \text{in } \Omega, \\ \frac{\partial y}{\partial n} &= g && \text{on } \partial\Omega, \end{aligned} \tag{2.11}$$

where $\frac{\partial}{\partial n}$ is the normal derivative in the direction of the outward normal. The constraints (2.10) can be formulated as an operator equation

$$Ay = f + u, \tag{2.12}$$

where the data $f$ is defined by $\langle f, v\rangle := \langle \tilde{f}, v\rangle + \int_{\partial\Omega} g(\gamma v)\, ds$ and $A$ is boundedly invertible from $Y$ to $Y'$.

In analogy to (2.8) we can take here

$$Z = H^s(\Omega), \quad 0 \le s \le 1, \quad U = (H^t(\Omega))', \quad 0 \le t \le 1. \tag{2.13}$$

Again $T = I$ and $E = I$ are then the canonical injections.

One can also prescribe as observations boundary conditions of Dirichlet type $y_*$ on $\partial\Omega$ in which case the natural observation space is $Z = H^{1/2}(\partial\Omega)$. Then $T : H^1(\Omega) \to H^{1/2}(\partial\Omega)$ coincides with the trace operator. In this case, the optimal control problem is to steer the states towards Dirichlet boundary conditions while the constraints (2.11) involve Neumann boundary conditions.

**2.2.3. Neumann Problem with Boundary Control.** Let now the boundary $\partial\Omega$ be decomposed into two parts $\overline{\partial\Omega} = \overline{\Gamma_N} \cup \overline{\Gamma_c}$, where $\Gamma_c$ has nonvanishing $d-1$ dimensional measure. For $a(\cdot,\cdot)$ from (2.9), consider the constraint

$$a(y,v) = \langle \tilde{f}, v\rangle + \int_{\Gamma_c} g(\gamma v)\, ds + \int_{\Gamma_c} u(\gamma v)\, ds \qquad \text{for all } v \in Y := H^1(\Omega) \tag{2.14}$$

and given $\tilde{f} \in Y'$, $g \in (H^{1/2}(\Gamma_c))'$, whose strong form is the second order Neumann problem

$$\begin{aligned} -\Delta y + y &= \tilde{f} \qquad \text{in } \Omega, \\ \frac{\partial y}{\partial n} &= \begin{cases} 0 & \text{on } \Gamma_N, \\ g+u & \text{on } \Gamma_c. \end{cases} \end{aligned} \tag{2.15}$$

In order to identify the remaining ingredients, note first that, for the right hand side of (2.14) to be well defined, the control must belong to $Q = (H^{1/2}(\Gamma_c))'$. Thus, the operator $E$ is the adjoint of the trace operator $\gamma$ to the *control boundary* $\Gamma_c$, defined as

$$\langle Eq, w\rangle_{(H^1(\Omega))'\times H^1(\Omega)} := \int_{\Gamma_c} q(\gamma w)\, ds. \tag{2.16}$$

That is, $E : (H^{1/2}(\Gamma_c))' \to (H^1(\Omega))'$ is an extension operator to $\Omega$. Thus, the formulation of the constraint as an operator equation reads in this case

$$Ay = f + Eu. \tag{2.17}$$

As in the previous cases, one could choose $Z$ to be a space defined on $\Omega$. A more frequent practical situation is to approximate prescribed conditions for the state on some part of the boundary.

To this end, denote by $\Gamma_\circ \subseteq \partial\Omega$ an *observation boundary* (again with strictly positive measure) and by $T : H^1(\Omega) \to H^{1/2}(\Gamma_\circ)$ the trace operator to this part of the boundary. Then the natural choice for $Z$ is $H^{1/2}(\Gamma_\circ)$. For the control, we have $Q = (H^{1/2}(\Gamma_c))'$ so that $U = L_2(\Gamma_c)$ would require the optimal control to be somewhat smoother. For these choices, the functional (2.3) is of the form

$$J(y,u) = \frac{1}{2}\|Ty - y_*\|^2_{H^{1/2}(\Gamma_\circ)} + \frac{\omega}{2}\|u\|^2_{L_2(\Gamma_c)}. \tag{2.18}$$

Again we could take $Z = H^s(\Gamma_\circ)$ for $0 \le s \le 1/2$ instead. For the choice $Z = L_2(\Gamma_\circ)$ and $U = L_2(\Gamma_c)$, the functional (2.3) with constraints (2.14) has been treated in [BKR] by employing an adaptive finite element solver. The case $\Gamma_\circ = \Gamma_c = \partial\Omega$ and $Z = U = L_2(\partial\Omega)$ is classical [Li].

REMARK 2.2. *For linear–quadratic elliptic problems with Dirichlet boundary controls the constraints are usually formulated as saddle point problems, see, e.g., [K2], which do no longer satisfy the ellipticity condition (2.4). The techniques developed below can also be extended to this situation, see [CDD2]. However, in order to make the basic mechanisms as transparent as possible we confine the present discussion to the case of elliptic constraints.*

**3. Reformulation of (ACP).** The standard approach to control problems like (ACP) would be to derive the necessary conditions for optimality in terms of an adjoint equation in the functional analytic setting, see e.g. [Li], and to discretize the resulting conditions by choosing suitable *finite dimensional* trial spaces. Here we will deviate from such a procedure in several ways. The first step is to transform the original problem (ACP) into an *equivalent* (hence still infinite dimensional) one which is, however, formulated entirely in $\ell_2$. This formulation will be seen to offer the following advantages:

- all the previous special cases take a *unified format*. All norms (including those with negative order or fractional trace norms) are represented by $\ell_2$–norms.
- there is no need for inverting ill–conditioned linear systems;
- it provides the foundation for adaptive solution strategies;
- aside from complexity issues such adaptive strategies have stabilizing effects in cases where usually discretizations have to obey compatibility constraints such as the LBB–condition.

The transformation hinges on the availability of appropriate *wavelet bases* which will be described next.

**3.1. Wavelet Coordinates.** In the following we shall assume that for each Hilbert space $H \in \{Y, Q\}$ we have a collection of functions

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

– *a wavelet basis* – with the following properties at our disposal. $I\!I_H$ is an infinite index set whose elements $\lambda$ encode different features such as *scale* $|\lambda|$ and spatial location $k = k(\lambda)$. In the simplest case of wavelets on the real line one has $\psi_{H,\lambda} = 2^{j/2}\psi(2^j \cdot -k)$, $j, k \in \mathbb{Z}$, normalized in $L_2$. Thus $\lambda$ represents $(j, k)$ and $|\lambda| = j$. We dispense at this point with further technical details about the actual construction of such wavelet bases but collect only those properties that are relevant in the present context.

**Locality (L):** The functions $\psi_{H,\lambda}$ are local, and the widths of their support are decreasing with growing discretization level $|\lambda|$,

$$\operatorname{diam}(\operatorname{supp}\psi_{H,\lambda}) \sim 2^{-|\lambda|}. \tag{3.2}$$

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

$$\langle v, \psi_{H,\lambda}\rangle \lesssim 2^{-|\lambda|(d/2+\tilde{m})}|v|_{W_\infty(\operatorname{supp}\,\psi_{H,\lambda})}, \tag{3.3}$$

where $d$ is the dimension of the underlying domain or manifold. Thus, integrating against a wavelet has the effect of taking an $\tilde{m}$th order difference which annihilates the smooth part of $v$. In fact, this is typically realized (for wavelets defined on Euclidean domains) by constructing $\Psi_H$ 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 $\langle \psi_{H,\lambda}, \tilde{\psi}_{H,\nu}\rangle = \delta_{\lambda,\nu}$, $\lambda, \nu \in I\!I_H$. Here and in the sequel the tilde is to express that the collection is a dual basis to a primal one for the space identified by the subscript. The role of dual bases will be addressed again below.

This cancellation property entails quasi–sparse representations of a wide class of operators.

**Riesz basis property (R):** This is perhaps the most crucial requirement. Every $v \in H$ has a unique expansion in terms of $\Psi_H$,

$$v = \sum_{\lambda \in \mathcal{I}_H} v_\lambda \, \psi_{H,\lambda} =: \mathbf{v}^T \, \Psi_H, \quad \mathbf{v} := (v_\lambda)_{\lambda \in \mathcal{I}_H}, \tag{3.4}$$

and its expansion coefficients satisfy the following *norm equivalence*: There exist finite positive constants $c_H, C_H$ such that

$$c_H \|\mathbf{v}\|_{\ell_2(\mathcal{I}_H)} \le \|\mathbf{v}^T \Psi_H\|_H \le C_H \|\mathbf{v}\|_{\ell_2(\mathcal{I}_H)}, \quad \mathbf{v} \in \mathcal{I}_H. \tag{3.5}$$

Thus, wavelet expansions induce isomorphisms between certain function and sequence spaces.

By duality arguments one can show that (3.5) is equivalent to the existence of a biorthogonal collection

$$\tilde{\Psi}_H := \{\tilde{\psi}_{H,\lambda} : \lambda \in \mathcal{I}_H\} \subset H', \quad \langle \psi_{H,\lambda}, \tilde{\psi}_{H,\mu} \rangle = \delta_{\lambda,\mu}, \qquad \lambda, \mu \in \mathcal{I}_H, \tag{3.6}$$

which is a Riesz basis in $H'$, i.e.,

$$C_H^{-1} \|\tilde{\mathbf{v}}\|_{\ell_2(\mathcal{I})} \le \|\tilde{\mathbf{v}}^T \tilde{\Psi}_H\|_{H'} \le c_H^{-1} \|\tilde{\mathbf{v}}\|_{\ell_2(\mathcal{I})} \tag{3.7}$$

holds for any $\tilde{v} = \tilde{\mathbf{v}}^T \, \tilde{\Psi}_H \in H'$, see e.g. [D1, D3, D4, K1].

We shall need a little more information about the way how bases with the above properties are constructed. In all our examples the Hilbert space $H \in \{Y, Q, Z, U\}$ is actually (a closed subspace of) a Sobolev space $H^s = H^s(G)$ or its dual (possibly determined by homogeneous boundary conditions) where $G$ is either the domain $\Omega$ or (part of) its boundary. The basis $\Psi_H$ for $H$ is then typically obtained from an *anchor* basis $\Psi = \{\psi_\lambda : \lambda \in \mathcal{I} = \mathcal{I}_H\}$ which is a Riesz basis for $L_2(G)$, i.e., $\|\psi_\lambda\|_{L_2(G)} \sim 1$, whose dual basis $\tilde{\Psi}$ is therefore also a Riesz basis for $L_2(G)$. In fact, $\Psi$ 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 Sobolev (sub-)spaces $H^s$, for $0 < s < \gamma, \tilde{\gamma}$, respectively. From this fact one derives then that for each $s \in (-\tilde{\gamma}, \gamma)$ the collection

$$\Psi_s := \{2^{-s|\lambda|} \psi_\lambda : \lambda \in \mathcal{I}\} =: \mathbf{D}^{-s} \Psi \tag{3.8}$$

is a Riesz basis for $H^s$ (with the above interpretation of $H^s$ as a dual when $s$ is negative) [D1], i.e., there exist positive constants $c_s, C_s$ such that

$$c_s \, \|\mathbf{v}\|_{\ell_2(\mathcal{I})} \le \|\mathbf{v}^T \Psi_s\|_{H^s} \le C_s \, \|\mathbf{v}\|_{\ell_2(\mathcal{I})}, \quad \mathbf{v} \in \ell_2(\mathcal{I}), \tag{3.9}$$

holds for each $s \in (-\tilde{\gamma}, \gamma)$. Analogous relations hold for $\tilde{\Psi}$ with reversed roles of $\gamma$ and $\tilde{\gamma}$. We shall make use of the following consequence of this fact. For $t \in (-\tilde{\gamma}, \gamma)$ the mapping

$$D^t : v = \mathbf{v}^T \Psi \to (\mathbf{D}^t \mathbf{v})^T \Psi = \mathbf{v}^T \mathbf{D}^t \Psi = \sum_{\lambda \in \mathcal{I}} v_\lambda \, 2^{t|\lambda|} \psi_\lambda \tag{3.10}$$

acts as a shift operator between Sobolev scales, i.e.

$$\|D^t v\|_{H^s} \sim \|v\|_{H^{s+t}} \sim \|\mathbf{D}^{s+t} \mathbf{v}\|_{\ell_2(\mathcal{I})}, \quad \text{provided that} \quad s, \; s+t \in (-\tilde{\gamma}, \gamma). \tag{3.11}$$

Concrete constructions of wavelet bases with the above properties for parameters $\gamma, \tilde{\gamma}$ ranging in most cases up to $3/2$ on bounded Euclidean domains and also on closed

piecewise parametrically defined manifolds can be found in [CTU, CM, DKU, DS1, DS2, DSt]. Note that in the above examples the relevant Sobolev regularity indices range between $-1$ and 1 so that these bases allow us to exploit relations like (3.11) when the metrics in the spaces $Z$ and $U$ differ from the natural norms in the way indicated above. So we shall henceforth assume the validity of the above properties (L), (CP), (R) in appropriate ranges as detailed in the next section.

In the sequel, it will be convenient to make systematic use of the following shorthand notation that has been already employed to some extent above. We will view $\Psi$ both as in (3.1) as a *collection* of functions as well as a (possibly infinite) (column) *vector* containing all functions always assembled in some fixed order. For a countable collection of functions $\Theta$ and some single function $\sigma$, the quantities $\langle\Theta,\sigma\rangle$ and $\langle\sigma,\Theta\rangle$ are to be understood as the column, respectively row, vector with entries $\langle\theta,\sigma\rangle$, respectively $\langle\sigma,\theta\rangle$, $\theta\in\Theta$. For two collections $\Theta,\Sigma$, the term $\langle\Theta,\Sigma\rangle$ 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 expansion coefficients in (3.4) and (3.7) can explicitly be expressed as $\mathbf{v}^T=\langle v,\tilde\Psi\rangle$ and $\tilde{\mathbf{v}}=\langle\Psi,\tilde v\rangle$. Furthermore, the *biorthogonality* or *duality conditions* (3.6) can be reexpressed as $\langle\Psi,\tilde\Psi\rangle=\mathbf{I}$ with the infinite identity matrix.

The last important ingredient concerns *wavelet representations* of operators. Suppose that $c(\cdot,\cdot)$ is a bilinear form on the product of Hilbert spaces $H\times M$ with bases $\Psi_H,\Psi_M$, respectively. Let $L: H\to M'$, $L': M\to H'$ be defined by $\langle Lv,w\rangle=c(v,w)=\langle v,L'w\rangle$. We can then represent $Lv\in M'$ in terms of the basis $\tilde\Psi_M$ for $M'$ which is dual to $\Psi_M$, i.e.,

$$\begin{aligned}Lv=L(\mathbf{v}^T\Psi_H)&=\langle L(\mathbf{v}^T\Psi_H),\Psi_M\rangle\tilde\Psi_M=c(\mathbf{v}^T\Psi_H,\Psi_M)\tilde\Psi_M=\mathbf{v}^Tc(\Psi_H,\Psi_M)\tilde\Psi_{M'}\\&=\big((c(\Psi_H,\Psi_M))^T\mathbf{v}\big)^T\tilde\Psi_M=(\langle\Psi_M,L\Psi_H\rangle\mathbf{v})^T\tilde\Psi_M.\end{aligned}\tag{3.12}$$

Thus, the expansion coefficients of $Lv$ (in the basis that spans the range space of $L$) are obtained by applying the *infinite* matrix $\mathbf{L}:=\langle\Psi_M,L\Psi_H\rangle=(c(\Psi_H,\Psi_M))^T$ to the coefficient vector of $v$, a fact that will be used frequently below. This matrix is referred to as the (standard) representation of $L$ with respect to the wavelet bases $\Psi_H,\Psi_M$ of the underlying spaces.

For general surveys on the application of wavelets to operator equations, we refer to [Co, D2, D3].

**3.2. Equivalent Control Problems in $\ell_2$.** Now we are in the position to transform the abstract control problem (ACP) into wavelet coordinates. We begin with the constraints (2.5). Following the above recipe (3.12), i.e., expanding $y$ in $\Psi_Y$ and $u$ in $\Psi_Q$, and testing with the elements of $\Psi_Y$, (2.5) takes the form

$$\mathbf{A}\boldsymbol{y}=\boldsymbol{f}+\mathbf{E}\check{\mathbf{u}},\tag{3.13}$$

where

$$\mathbf{A}:=a(\Psi_Y,\Psi_Y),\qquad\mathbf{E}:=\langle\Psi_Y,E\Psi_Q\rangle,\qquad\boldsymbol{f}:=\langle\Psi_Y,f\rangle.\tag{3.14}$$

(Since it will be more convenient later to work with a scaled version $\mathbf{u}$ of $\check{\mathbf{u}}$, we reserve the symbol $\mathbf{u}$ for that purpose.) To simplify the notation we shall suppress in the following the subscripts $\ell_2(I\!I)$ and write briefly $\|\cdot\|:=\|\cdot\|_{\ell_2(I\!I)}$ because we shall only

be dealing with Euclidean norms and the ranges of indices $\mathcal{I}$ will always be clear from the context.

It is well known that the ellipticity (2.4) and the Riesz basis property (R) (3.5) (for $H = Y$) imply the following fact, see e.g. [D2].

REMARK 3.1. *The matrix* $\mathbf{A}$ *is a boundedly invertible mapping of* $\ell_2(\mathcal{I}_Y)$ *onto itself, i.e., there exist finite positive constants* $c_{\mathbf{A}}, C_{\mathbf{A}}$ *such that*

$$c_{\mathbf{A}}\|\mathbf{v}\| \leq \|\mathbf{A}\mathbf{v}\| \leq C_{\mathbf{A}}\|\mathbf{v}\|, \qquad \mathbf{v} \in \ell_2(\mathcal{I}_Y). \tag{3.15}$$

Similarly, since the observation data $y_*$ and $Ty$ are supposed to belong to the space $Z$, it is natural to expand these quantities in terms of the basis $\Psi_Z$ (whose precise form will be derived shortly with the aid of (3.8)). In fact, for $y = \mathbf{y}^T\Psi_Y$ one has by (3.12)

$$Ty = \mathbf{y}^T\langle T\Psi_Y, \tilde{\Psi}_Z\rangle\Psi_Z, \qquad y_* = \langle y_*, \tilde{\Psi}_Z\rangle\Psi_Z,$$

which means that the representation of $T$ and the coordinates of $y_*$ are given by

$$\mathbf{T} := \langle\tilde{\Psi}_Z, T\Psi_Y\rangle, \qquad \mathbf{y}_Z := \langle\tilde{\Psi}_Z, y_*\rangle. \tag{3.16}$$

Now we still have to specify $\Psi_Z$ depending on the choice of $Z$. According to the above examples, two essentially different cases arise.

**Case (I):** The space $Z$ belongs to the Sobolev scale over $\Omega$, see (2.8), (2.13). As mentioned in the previous section, the basis $\Psi_Z$ can then be taken as a scaled version of $\Psi_Y$, i.e. there exists a diagonal matrix $\mathbf{D}_Z$ such that

$$\Psi_Z = \mathbf{D}_Z\Psi_Y, \qquad \tilde{\Psi}_Z = \mathbf{D}_Z^{-1}\tilde{\Psi}_Y, \tag{3.17}$$

which clearly forms a dual pair for $Z$. In this case $T$ was just the canonical injection so that

$$\mathbf{T} = \langle\tilde{\Psi}_Z, \Psi_Y\rangle = \mathbf{D}_Z^{-1}\langle\tilde{\Psi}_Y, \Psi_Y\rangle = \mathbf{D}_Z^{-1}, \quad \mathbf{y}_Z = \mathbf{D}_Z^{-1}\langle\tilde{\Psi}_Y, y_*\rangle =: \mathbf{D}_Z^{-1}\mathbf{y}_*. \tag{3.18}$$

If $Z$ does not coincide with $Y$, $Z$ induces a weaker topology than $Y$ so that the entries of the diagonal matrix $\mathbf{D}_Z$ increase in scale. For instance, for $Y = H^t$, $Z = H^s$, $s \leq t$, one has $(\mathbf{D}_Z)_{\lambda,\lambda} \sim 2^{(t-s)|\lambda|}$. Recall that in this case the mapping $E$ is also just the identity ((2.7), (2.12)), i.e., $Q = Y'$ and $\Psi_Q$ should span the range of $A$. So we might as well take

$$\Psi_Q := \tilde{\Psi}_Y, \quad \mathbf{E} = \mathbf{I}, \tag{3.19}$$

in this case. Thus, $\mathbf{T}$ and $\mathbf{E}$ are obviously bounded on $\ell_2(\mathcal{I}_Y)$, $\ell_2(\mathcal{I}_Q)$, respectively.

When $U$ induces a strictly stronger topology than $Q$, we can make use of the shift relations (3.11) which say that there exists a diagonal matrix $\mathbf{D}_U$ such that

$$\|u\|_U \sim \|\mathbf{D}_U^{-1}\check{\mathbf{u}}\|. \tag{3.20}$$

Again, since $\tilde{\Psi}_Y$ is a Riesz basis for $Q = Y'$ and $U$ induces a stronger topology, the diagonal entries of $\mathbf{D}_U$ are non–increasing in scale.

We have collected now all the ingredients to formulate a counterpart to (2.3) in the present case.

REMARK 3.2. *Under the assumption (I) the quadratic functional*

$$\widetilde{\mathbf{J}}(\mathbf{y},\check{\mathbf{u}}) := \frac{1}{2}\|\mathbf{D}_Z^{-1}(\mathbf{y}-\mathbf{y}_*)\|^2 + \frac{\omega}{2}\|\mathbf{D}_U^{-1}\check{\mathbf{u}}\|^2, \tag{3.21}$$

*is equivalent to the functional $J(y,u)$ from (2.3) in the following sense: There exist finite positive constants $c_J, C_J$ such that for any $y = \mathbf{y}^T\Psi_Y \in Y$, $y_* = (\mathbf{D}_Z^{-1}\mathbf{y}_*)^T\Psi_Z \in Z$ and any $u = \check{\mathbf{u}}^T\Psi_Q \in U$, one has*

$$c_J\,\widetilde{\mathbf{J}}(\mathbf{y},\check{\mathbf{u}}) \le J(y,u) \le C_J\,\widetilde{\mathbf{J}}(\mathbf{y},\check{\mathbf{u}}). \tag{3.22}$$

*Proof.* The fact that $\|\mathbf{D}_Z^{-1}(\mathbf{y}-\mathbf{y}_*)\| \sim \|Ty - y_*\|_Z$ follows immediately from (3.18) and the fact that $\Psi_Z$ is a Riesz basis for $Z$. The assertion is then an immediate consequence of (3.20). □

**Case (II):** The spaces $Z$ and $U$ are defined over different domains than $\Omega$, such as traces as explained at the end of Section 2.2.2 and in Section 2.2.3. Here we cannot simply interrelate the bases for $Y$ and $Z$. In fact, they will generally be quite independent of each other. However, we may still identify the *range* $R$ of the mapping $T$ when acting on $Y$ as a *natural* space on the observation side. Thus, we can begin again with a Riesz basis $\Psi_R$ for this natural space and can, as in all the above examples, generate a Riesz basis for $Z$ through a proper scaling

$$\Psi_Z = \mathbf{D}_Z\Psi_R, \quad \tilde{\Psi}_Z = \mathbf{D}_Z^{-1}\tilde{\Psi}_R, \tag{3.23}$$

where, by the same reasoning as above, the diagonal entries of $\mathbf{D}_Z$ increase in scale. Replacing (3.17) by (3.23), we obtain as before

$$\mathbf{T} = \mathbf{D}_Z^{-1}\langle\tilde{\Psi}_R, T\Psi_Y\rangle, \quad y_* = (\mathbf{D}_Z^{-1}\mathbf{y}_*)^T\Psi_Z, \tag{3.24}$$

to conclude that

$$\|\mathbf{D}_Z^{-1}(\mathbf{T}\mathbf{y}-\mathbf{y}_*)\| \sim \|Ty - y_*\|_Z \quad \text{for all} \quad y = \mathbf{y}^T\Psi_Y. \tag{3.25}$$

Likewise, the Riesz basis $\Psi_U$ for $U$ can be taken as a scaled version of $\Psi_Q$,

$$\Psi_U = \mathbf{D}_U\Psi_Q, \tag{3.26}$$

for some diagonal matrix $\mathbf{D}_U$ whose entries do not increase in scale since $U$ induces an at most stronger topology than $Q$. As before we infer that (3.20) holds again.

REMARK 3.3. *Under the assumption (II) the functional*

$$\widetilde{\mathbf{J}}(\mathbf{y},\check{\mathbf{u}}) := \frac{1}{2}\|\mathbf{D}_Z^{-1}(\mathbf{T}\mathbf{y}-\mathbf{y}_*)\|^2 + \frac{\omega}{2}\|\mathbf{D}_U^{-1}\check{\mathbf{u}}\|^2, \tag{3.27}$$

*with $\mathbf{T}, \mathbf{y}_*$ and $\mathbf{D}_Z, \mathbf{D}_U$ given by (3.24), (3.23) and (3.26), respectively, is equivalent to $J(y,u)$ in the above sense.*

At this point we cannot further specify the representation of $\mathbf{E}$ from (3.14) but remark that the continuity of the operators $E$ and $T$ (see, e.g., [K2] for references to trace theorems) combined with the property (R) of the wavelet bases ensures that also in the case (II) there exist finite positive constants $C_{\mathbf{T}}, C_{\mathbf{E}}$ such that

$$\|\mathbf{T}\mathbf{v}\| \le C_{\mathbf{T}}\|\mathbf{v}\|, \qquad \|\mathbf{E}\mathbf{v}\| \le C_{\mathbf{E}}\|\mathbf{v}\| \tag{3.28}$$

for $\mathbf{v} \in \ell_2(I_Y)$ and $\mathbf{v} \in \ell_2(I_Q)$, respectively.

In summary, we arrive at the following abstract control problem as a *discretized control problem* over $\ell_2$.

**(DCP):** *Minimize the quadratic functional (3.27) (respectively (3.21) in the case (I)) subject to the constraint (3.13) with* $\mathbf{A}, \mathbf{T}, \mathbf{E}$ *defined as above.*

In view of Remarks 3.2, 3.3, the minimizer of (3.21) or (3.27) is related to the minimizer of the original functional (2.7) as follows. For some $c^* \in [c_J, C_J]$ one has $c^*\widetilde{\mathbf{J}}(\hat{\mathbf{y}}, \hat{\mathbf{u}}) = J(y, u)$ for the minimizing pair $(y, u)$ and its wavelet coordinates $(\hat{\mathbf{y}}, \hat{\mathbf{u}})$, so that the minimizer $(\mathbf{y}, \check{\mathbf{u}})$ of (3.27) satisfies $\widetilde{\mathbf{J}}(\mathbf{y}, \check{\mathbf{u}}) \leq c_*^{-1} J(y, u)$. Moreover, in the case of compatible data $y_* = TA^{-1}f$ the respective minimizers coincide. In that sense the minimization of either one of the new functionals (3.21) or (3.27) is understood to capture the essential features of the original extremal problem. Since estimates for the constants in the norm equivalences (3.5) are known for concrete examples of wavelets, the value for the functional (2.7) can in case (I) be directly computed. In case (II), one would also need estimates for the constants in the trace theorems to determine the constants in (3.28).

We shall derive next several equivalent formulations of (DCP), which the subsequent numerical treatment will be based upon. By standard arguments (see e.g. [Li]), the unique minimum for (DCP) is obtained by solving the first order necessary conditions for $\widetilde{\mathbf{J}}$ which can e.g. be derived by first eliminating $\mathbf{y}$ in (3.27). In view of (3.15), we can invert (3.13) to obtain $\mathbf{y} = \mathbf{A}^{-1}\boldsymbol{f} + \mathbf{A}^{-1}\mathbf{E}\check{\mathbf{u}}$. Substituting this into (3.27) yields a functional which only depends on $\check{\mathbf{u}}$,

$$\check{\mathbf{J}}(\check{\mathbf{u}}) \;:=\; \frac{1}{2}\|\mathbf{D}_Z^{-1}\left(\mathbf{T}\mathbf{A}^{-1}\mathbf{E}\check{\mathbf{u}} - (\mathbf{y}_* - \mathbf{T}\mathbf{A}^{-1}\boldsymbol{f})\right)\|^2 + \frac{\omega}{2}\|\mathbf{D}_U^{-1}\check{\mathbf{u}}\|^2. \tag{3.29}$$

Abbreviating

$$\check{\mathbf{Z}} := \mathbf{D}_Z^{-1}\mathbf{T}\mathbf{A}^{-1}\mathbf{E}, \qquad \check{\mathbf{G}} := \mathbf{D}_Z^{-1}(\mathbf{y}_* - \mathbf{T}\mathbf{A}^{-1}\boldsymbol{f}), \tag{3.30}$$

$\check{\mathbf{J}}$ is of the form

$$\check{\mathbf{J}}(\check{\mathbf{u}}) \;=\; \frac{1}{2}\|\check{\mathbf{Z}}\check{\mathbf{u}} - \check{\mathbf{G}}\|^2 + \frac{\omega}{2}\|\mathbf{D}_U^{-1}\check{\mathbf{u}}\|^2. \tag{3.31}$$

This is a standard least squares functional whose minimizer is characterized by the normal equations which were in the present format derived in [K2].

PROPOSITION 3.4. *The functional* $\check{\mathbf{J}}$ *is twice differentiable on* $\ell_2(I_Q)$ *with first and second variation given by*

$$\delta\check{\mathbf{J}}(\check{\mathbf{u}}) = (\check{\mathbf{Z}}^T\check{\mathbf{Z}} + \omega\mathbf{D}_U^{-2})\check{\mathbf{u}} - \check{\mathbf{Z}}^T\check{\mathbf{G}}, \qquad \delta^2\check{\mathbf{J}}(\check{\mathbf{u}}) = \check{\mathbf{Z}}^T\check{\mathbf{Z}} + \omega\mathbf{D}_U^{-2}. \tag{3.32}$$

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

In fact,

$$\check{\mathbf{Q}} := \check{\mathbf{Z}}^T\check{\mathbf{Z}} + \omega\mathbf{D}_U^{-2} \tag{3.33}$$

is positive definite since $\check{\mathbf{Z}}^T\check{\mathbf{Z}}$ is at least positive semi–definite and $\mathbf{D}_U^{-2}$ is a diagonal matrix with strictly positive diagonal entries. Thus, the solution of (DCP) is uniquely determined by solving $\delta\check{\mathbf{J}}(\check{\mathbf{u}}) = 0$, i.e., the system

$$\check{\mathbf{Q}}\check{\mathbf{u}} = \check{\mathbf{g}} \qquad \text{where} \qquad \check{\mathbf{g}} := \check{\mathbf{Z}}^T\check{\mathbf{G}}. \tag{3.34}$$

The system (3.34) would offer a natural link to the setting in [CDD2] because it is symmetric positive definite and, thus, invites the application of gradient iterations which could then be carried out approximately by adaptive applications of $\check{\mathbf{Q}}$. However, there are still two obstructions, namely, (a) the spectral condition of $\check{\mathbf{Q}}$ and (b) the difficulty in applying $\check{\mathbf{Q}}$ and evaluating the right hand side due to the inverses involved in the definition of $\check{\mathbf{Z}}$ and $\check{\mathbf{G}}$.

As for (a), when using the natural norms in (2.3), the mapping $\check{\mathbf{Q}}$ is indeed boundedly invertible on $\ell_2(I\!I_Q)$ and ready for applying an iterative scheme. In the other cases, note that $\check{\mathbf{Z}}$ has a bounded $\ell_2$–norm since $\mathbf{D}_Z^{-1}$ has nonincreasing diagonal entries when the scale grows and $\mathbf{T}, \mathbf{E}, \mathbf{A}^{-1}$ are bounded on $\ell_2$, see (3.15), (3.28). However, $\mathbf{D}_U^{-2}$ will increase whenever $U$ has a strictly stronger topology than $Q$. In order to cover this more general situation we can precondition the system through the following scaling. Defining

$$\mathbf{u} := \mathbf{D}_U^{-1}\check{\mathbf{u}}, \qquad \mathbf{Z} := \check{\mathbf{Z}}\mathbf{D}_U, \qquad \mathbf{Q} := \mathbf{D}_U\check{\mathbf{Q}}\mathbf{D}_U = \mathbf{Z}^T\mathbf{Z} + \omega\mathbf{I}, \tag{3.35}$$

straightforward calculations show that (3.34) is equivalent to

$$\mathbf{Q}\mathbf{u} = \mathbf{g} := \mathbf{Z}^T\mathbf{G}. \tag{3.36}$$

Since now $\mathbf{Z}$ is $\ell_2$–bounded, the third relation in (3.35) shows that $\mathbf{Q}$ has uniformly bounded condition numbers. Consequently, there exist finite positive constants $c_{\mathbf{Q}}, C_{\mathbf{Q}}$ such that

$$c_{\mathbf{Q}}\|\mathbf{v}\| \le \|\mathbf{Q}\mathbf{v}\| \le C_{\mathbf{Q}}\|\mathbf{v}\|, \qquad \mathbf{v} \in \ell_2, \tag{3.37}$$

where we can actually take $c_{\mathbf{Q}} = \omega$. Of course, (3.36) yields the unique minimizer of the functional

$$\mathbf{J}(\mathbf{u}) \;:=\; \frac{1}{2}\|\mathbf{Z}\mathbf{u} - \mathbf{G}\|^2 + \frac{\omega}{2}\|\mathbf{u}\|^2, \tag{3.38}$$

which corresponds to normalizing the controls from the beginning in the basis $\Psi_U$.

In view of (3.37), there exists a fixed positive parameter $\alpha$ such that the *gradient iteration*

$$\mathbf{u}^{k+1} = \mathbf{u}^k + \alpha(\mathbf{g} - \mathbf{Q}\mathbf{u}^k) \tag{3.39}$$

reduces the error in each step by at least a factor $\rho < 1$, i.e.

$$\|\mathbf{u} - \mathbf{u}^{k+1}\| \le \rho\|\mathbf{u} - \mathbf{u}^k\|, \quad k = 0, 1, 2, \ldots, \tag{3.40}$$

where $\mathbf{u}$ is the exact solution of (3.36). Our ultimate goal is to carry out this iteration approximately with dynamically updated accuracy tolerances.

**4. The Basic Concepts.** In this section we collect the main conceptual tools from [CDD2, CDD3] that will be needed to treat (3.36) for the solution of (DCP) and thereby tackle obstruction (b), namely, the application of $\mathbf{Q}$ and the evaluation of the right hand side $\mathbf{g}$.

**4.1. Perturbed Iterations.** The basic strategy applies, in principle, to any system of the form

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

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

$$c_{\mathbf{M}}\|\mathbf{v}\| \le \|\mathbf{M}\mathbf{v}\| \le C_{\mathbf{M}}\|\mathbf{v}\|, \quad \mathbf{v} \in \ell_2, \tag{4.2}$$

for some finite positive constants $c_{\mathbf{M}}, C_{\mathbf{M}}$, as well as

$$\rho := \|\mathbf{I} - \alpha\mathbf{M}\| < 1 \tag{4.3}$$

for some positive number $\alpha$. Clearly, due to the positive definiteness of $\mathbf{Q}$ and by (3.37), $\mathbf{M} = \mathbf{Q}$ falls into this category.

Given (4.3), the gradient iteration

$$\mathbf{q}^{k+1} = \mathbf{q}^k + \alpha(\mathbf{z} - \mathbf{M}\mathbf{q}^k), \quad k = 0, 1, 2, \ldots, \tag{4.4}$$

will converge with a fixed error reduction rate $\rho < 1$ per step. Of course, this iteration cannot be carried out exactly because $\mathbf{M}$ is an infinite matrix and the data $\mathbf{z}$ could be an infinite array. However, one can hope that perturbed iterations with dynamical accuracy tolerances that are suitably updated in the course of the iteration will still converge. Thus, we need a routine with the following property.

RES $[\eta, \mathbf{M}, \mathbf{z}, \mathbf{v}] \to \mathbf{r}_\eta$ DETERMINES FOR A GIVEN TOLERANCE $\eta > 0$ A FINITELY SUPPORTED SEQUENCE $\mathbf{r}_\eta$ SATISFYING

$$\|\mathbf{z} - \mathbf{M}\mathbf{v} - \mathbf{r}_\eta\| \le \eta. \tag{4.5}$$

There is a further ingredient whose role is at this stage not apparent yet but which will eventually play a crucial role in controlling the complexity of the scheme.

COARSE $[\eta, \mathbf{w}] \to \mathbf{w}_\eta$ DETERMINES FOR ANY FINITELY SUPPORTED INPUT VECTOR $\mathbf{w}$ A VECTOR $\mathbf{w}_\eta$ WITH SMALLEST POSSIBLE SUPPORT SUCH THAT

$$\|\mathbf{w} - \mathbf{w}_\eta\| \le \eta. \tag{4.6}$$

The precise description of COARSE can be found in [CDD1]. The idea is to sort the entries of $\mathbf{w}$ by size and to subtract squares of their moduli starting from the smallest one until the sum reaches $\eta^2$. The sorting usually introduces a logarithmic term of the size of $\mathbf{w}$. A quasi–sorting based on binary binning can be shown to avoid the logarithmic term at the expense of the resulting support size being at most a fixed constant of the minimal size, see [B]. This will suffice for the subsequent analysis so that it is justified to suppress logarithmic terms in the sequel.

Let us suppose for the moment that the routine RES is already at our disposal. We shall first devise the precise form of a perturbed iteration that converges in the following sense. For every target accuracy $\varepsilon$ it produces after finitely many steps a finitely supported approximate solution with accuracy $\varepsilon$.

Following [CDD2], to arrive at the right interplay between the routines RES and COARSE, we need the following control parameter. Given (an estimate of) the reduction rate $\rho$ and the step size parameter $\alpha$ from (4.3), let

$$K := \min\{\ell \in I\!N : \rho^{\ell-1}(\alpha\ell + \rho) \le \tfrac{1}{10}\}. \tag{4.7}$$

(Here the upper bound $(10)^{-1}$ stems from the analysis in [CDD2] and will be used again below.) Denoting in the following always by $\mathbf{q}$ the exact solution of (4.1), a perturbed version of (4.4) can now be formulated as follows.

SOLVE $[\varepsilon, \mathbf{M}, \mathbf{z}, \overline{\mathbf{q}}^0, \varepsilon_0] \to \overline{\mathbf{q}}_\varepsilon$

(I) FIX A TARGET ACCURACY $\varepsilon > 0$. GIVEN AN INITIAL GUESS $\overline{\mathbf{q}}^0$ ALONG WITH AN ERROR BOUND $\|\mathbf{q} - \overline{\mathbf{q}}^0\| \le \varepsilon_0$, SET $j = 0$.

(II) IF $\varepsilon_j \le \varepsilon$, STOP AND SET $\overline{\mathbf{q}}_\varepsilon := \overline{\mathbf{q}}^j$. OTHERWISE SET $\mathbf{v}^0 := \overline{\mathbf{q}}^j$.

(II.1) FOR $k = 0, \dots, K-1$ COMPUTE RES $[\rho^k \varepsilon_j, \mathbf{M}, \mathbf{z}, \mathbf{v}^k] \to \mathbf{r}^k$ AND

$$\mathbf{v}^{k+1} := \mathbf{v}^k + \alpha \mathbf{r}^k. \tag{4.8}$$

(II.2) APPLY COARSE $[\frac{2}{5}\varepsilon_j, \mathbf{v}^K] \to \overline{\mathbf{q}}^{j+1}$; SET $\varepsilon_{j+1} := \frac{1}{2}\varepsilon_j$, $j+1 \to j$ AND GO TO (II).

In the case that no particular initial guess is known, step (I) is replaced by the default

(I)' FIX A TARGET ACCURACY $\varepsilon > 0$. SET $j = 0$ AND

$$\overline{\mathbf{q}}^0 = \mathbf{0}, \quad \varepsilon_0 := c_{\mathbf{M}}^{-1} \|\mathbf{z}\|. \tag{4.9}$$

In this case we use the short notation SOLVE $[\varepsilon, \mathbf{M}, \mathbf{z}] \to \overline{\mathbf{q}}_\varepsilon$.

The choice of the interior tolerance $\rho^k \varepsilon_j$ in step (II.1) yields the following estimate from [CDD2] regarding the final iterate $\mathbf{v}^K$ resulting from step (II.1). Inserting the exact iterate of (4.4) with initial value $\overline{\mathbf{w}}^j$ denoted by $\overline{\mathbf{v}}^K(\overline{\mathbf{w}}^j)$, we get

$$\begin{aligned} \|\mathbf{v}^K - \mathbf{q}\| &\le \|\mathbf{v}^K - \overline{\mathbf{v}}^K(\overline{\mathbf{w}}^j)\| + \|\overline{\mathbf{v}}^K(\overline{\mathbf{w}}^j) - \mathbf{q}\| \\ &\le \alpha K \rho^{K-1} \varepsilon_j + \rho^K \|\overline{\mathbf{w}}^j - \mathbf{q}\| \le (\alpha K + \rho) \rho^{K-1} \varepsilon_j. \end{aligned} \tag{4.10}$$

Employing the choice of $K$ in (4.7), this yields

$$\|\mathbf{v}^K - \mathbf{q}\| \le \frac{\varepsilon_j}{10} \tag{4.11}$$

The particular form of the constants for the interior estimates that can be seen in (4.10) will be employed later in Section 6.

Straightforward perturbation arguments reveal the following result, see [CDD2, CDD3].

PROPOSITION 4.1. *The iterates* $\overline{\mathbf{q}}^j$ *generated by* SOLVE $[\varepsilon, \mathbf{M}, \mathbf{z}]$ *satisfy*

$$\|\mathbf{q} - \overline{\mathbf{q}}^j\| \le \varepsilon_j, \qquad j \in I\!N_0, \tag{4.12}$$

*where* $\varepsilon_j = 2^{-j}\varepsilon_0$.

Of course, the estimates for $\alpha$ rely on the constants in the norm equivalences (3.5) and in the relation (4.2). So there may be only a poor estimate for $\rho$ which, in turn, gives rise to an overly pessimistic choice for the number $K$ defined in (4.7) of perturbed iterations in each block (II) of SOLVE prior to a coarsening step. Therefore, we recall from [CDD3] that step (II) can be terminated based on monitoring the approximate residuals as follows. By (4.2), we have

$$\|\mathbf{q} - \mathbf{v}\| \le c_{\mathbf{M}}^{-1} \|\mathbf{z} - \mathbf{M}\mathbf{v}\|. \tag{4.13}$$

Choose any $\bar{\rho} < 1$ and define $\bar{K}$ by (4.7) with respect to $\bar{\rho}$. Replacing $\rho$ by $\bar{\rho}$ in the definition of the tolerances in step (II) of SOLVE would take $M := \max\{K, \bar{K}\}$ steps to ensure that in the $(j+1)$st call of (II) $\|\mathbf{q} - \mathbf{v}^M\| \le \varepsilon_j/10$. Now suppose that the $\rho$ is expected to be a too pessimistic estimate of the true reduction rate. Choosing e.g. $\bar{\rho} := 1/2$ and setting $\eta_k := 2^{-k}\varepsilon_j = \bar{\rho}^{-k}\varepsilon_j$ as tolerances in the $(j+1)$st call of (II), we infer from (4.13) and (4.5) that

$$\|\mathbf{q} - \mathbf{v}^k\| \le c_{\mathbf{M}}^{-1} \|\mathbf{z} - \mathbf{M}\mathbf{v}^k\| \le c_{\mathbf{M}}^{-1} \left(\eta_k + \|\mathbf{r}^k\|\right) =: \delta_k,$$

where $\mathbf{r}^k$ is the approximate residual produced in step (II.1) of SOLVE. By the previous remarks, we can terminate the iteration in step (II) of SOLVE when either $k = K-1$ or the *computable a–posteriori bound* $\delta_k$ satisfies $\delta_k \leq \varepsilon_j/10$, which might happen much earlier than predicted by (4.7). Of course, the constant $c_{\mathbf{M}}$ is usually also only estimated. However, a poor estimate enters the above a–posteriori termination criterion in a less severe way than a poor estimate for $\rho$. Nevertheless, in order to keep the exposition as simple as possible, we confine the subsequent discussion to the above version of SOLVE, bearing in mind that variants of the above sort are automatically covered by the complexity analysis.

**4.2. Complexity Analysis.** Of course, the main issues are the actual realization of the routine RES, and to understand its complexity. The realization will depend on the concrete application, which here will be the control problem (DCP). Here we outline first a suitable framework for the complexity analysis. Striving for schemes that are in some sense *optimal*, the meaning of optimality has to be clarified first.

We say that the scheme SOLVE has an *optimal work/accuracy rate s* if the following is true: Whenever the error of best $N$–term approximation

$$\sigma_N(\mathbf{q}) := \|\mathbf{q} - \mathbf{q}_N\| := \min_{\#\operatorname{supp}\mathbf{v} \leq N} \|\mathbf{q} - \mathbf{v}\| \tag{4.14}$$

decays like $\mathcal{O}(N^{-s})$, then the solution $\overline{\mathbf{q}}_\varepsilon$ is produced by SOLVE at an expense that also stays proportional to $\varepsilon^{-1/s}$ and in that sense matches the best $N$–term rate. Clearly this implies that $\#\operatorname{supp}\overline{\mathbf{q}}_\varepsilon$ also stays proportional to $\varepsilon^{-1/s}$. Thus, our benchmark is 'best $N$–term approximation' which is obviously the best one can hope for.

Clearly this best $N$–term approximation $\mathbf{q}_N$ of $\mathbf{q}$ is given by taking the $N$ largest terms in modulus from $\mathbf{q}$. When $\mathbf{q}$ is the (unknown) solution of (4.1) this knowledge is certainly not available. Nevertheless, the formulation of appropriate complexity criteria will be based on a characterization of those sequences $\mathbf{v}$ for which the best $N$–term approximation error decays at a particular rate (*Lorentz spaces*). Following [CDD1], consider sequences that are *sparse* in the sense that for any given threshold $0 < \eta \leq 1$, say, the number of terms exceeding that threshold is controlled by some function of this threshold. Specifically, set for some $0 < \tau < 2$

$$\ell_\tau^w := \{\mathbf{v} \in \ell_2 : \ \#\{\lambda \in \mathcal{I} : |v_\lambda| > \eta\} \leq C_{\mathbf{v}}\, \eta^{-\tau}, \ \text{for all } 0 < \eta \leq 1\}, \tag{4.15}$$

i.e. the set $\ell_\tau^w$ consists of all those sequences $\mathbf{v} \in \ell_2$ for which there exists a constant $C_{\mathbf{v}}$ such that for all $0 < \eta \leq 1$ the number of terms $v_\lambda$ whose moduli exceed the threshold $\eta$ is bounded by $C_{\mathbf{v}}\eta^{-\tau}$. Note that this determines a strict subspace of $\ell_2$ only when $\tau < 2$, and the sequence is the sparser the smaller $\tau$ is. Denote for a given $\mathbf{v} \in \ell_\tau^w$ by $C_{\mathbf{v}}$ the smallest constant for which (4.15) holds. Then one has

$$|\mathbf{v}|_{\ell_\tau^w} := \sup_{n \in I\!N} n^{1/\tau} v_n^* = C_{\mathbf{v}}^{1/\tau}, \tag{4.16}$$

where $\mathbf{v}^* = (v_n^*)_{n \in I\!N}$ is a non–decreasing rearrangement of $\mathbf{v}$. The quantity

$$\|\mathbf{v}\|_{\ell_\tau^w} := \|\mathbf{v}\| + |\mathbf{v}|_{\ell_\tau^w} \tag{4.17}$$

can be shown to be a quasi–norm for $\ell_\tau^w$ [CDD1]. Furthermore, because of the following continuous embeddings

$$\ell_\tau \subset \ell_\tau^w \subset \ell_{\tau+\varepsilon} \subset \ell_2 \qquad \text{for } \tau < \tau + \varepsilon < 2, \tag{4.18}$$

$\ell_\tau^w$ is very close to $\ell_\tau$ which justifies to call it *weak* $\ell_\tau$. Now we can recall the following result from [CDD1] which relates the sequences in $\ell_\tau^w$ to best $N$–term approximation.

PROPOSITION 4.2. *Let positive real numbers $s$ and $\tau$ be related by*

$$\frac{1}{\tau} = s + \frac{1}{2}. \tag{4.19}$$

*Then a sequence $\mathbf{v}$ belongs to $\ell_\tau^w$ if and only if*

$$\|\mathbf{v} - \mathbf{v}_N\| \lesssim N^{-s} \qquad \text{and} \qquad \sigma_N(\mathbf{v}) \lesssim N^{-s}\|\mathbf{v}\|_{\ell_\tau^w}, \tag{4.20}$$

*where as before $\mathbf{v}_N$ denotes a best $N$–term approximation of $\mathbf{v}$.*

Depending on the space $H$ which is characterized by the wavelet basis $\Psi_H$, the fact that an array of wavelet coefficients $\mathbf{v}$ belongs to $\ell_\tau$ is typically equivalent to the fact that the expansion $\mathbf{v}^T\Psi_H$ belongs to a certain Besov space which describes a much weaker regularity measure than a Sobolev space of corresponding order. In view of (4.18), Proposition 4.2 therefore expresses how much loss of regularity can be compensated by best $N$-term approximation, i.e., by judiciously placing the degrees of freedom in a nonlinear way so as to retain a certain optimal order of error decay. We shall return to this issue later.

As will be seen in Theorem 4.3 below, a key criterion for a scheme SOLVE to exhibit an optimal work/accuracy rate can now be formulated through the following property of the respective residual approximation.

**$\tau^*$–Sparsity:** *The routine* RES *is called $\tau^*$–sparse for some $0 < \tau^* < 2$ if the following holds: Whenever the solution $\mathbf{q}$ of (4.1) belongs to $\ell_\tau^w$ for some $\tau^* < \tau < 2$, then for any $\mathbf{v}$ with finite support the output $\mathbf{r}_\eta$ of* RES $[\eta, \mathbf{M}, \mathbf{z}, \mathbf{v}]$ *satisfies*

*(i)*
$$\begin{aligned} \|\mathbf{r}_\eta\|_{\ell_\tau^w} &\lesssim \max\{\|\mathbf{v}\|_{\ell_\tau^w}, \|\mathbf{q}\|_{\ell_\tau^w}\}, \\ \#\operatorname{supp}\mathbf{r}_\eta &\lesssim \eta^{-1/s}\max\{\|\mathbf{v}\|_{\ell_\tau^w}^{1/s}, \|\mathbf{q}\|_{\ell_\tau^w}^{1/s}\}, \end{aligned} \tag{4.21}$$

*where $s$ and $\tau$ are related as before by (4.19);*

*(ii) the number of floating point operations needed to compute $\mathbf{r}_\eta$ stays proportional to $\#\operatorname{supp}\mathbf{r}_\eta$.*

*Furthermore, the constants in (i), (ii) depend only on $\tau$ as $\tau \to \tau^*$.*

In this context we shall always make the following tacit assumption. Given data are always be considered completely accessible. In practical terms this may mean that, depending on some final target accuracy (in view of (3.37)) sufficiently many of the corresponding coefficients of explicitly given data are determined in a preprocessing step and then ordered by size, so that COARSE can be applied to a finitely supported array. For notational simplicity we shall not distinguish between the ideal exact data and such an approximation.

The following result can then be extracted from the analysis in [CDD2] (see also [CDD3] for nonlinear problems) and has been employed already in [DUV].

THEOREM 4.3. *If* RES *is $\tau^*$–sparse and if the exact solution $\mathbf{q}$ of (4.1) belongs to $\ell_\tau^w$ for some $\tau > \tau^*$, then for every $\varepsilon > 0$ algorithm* SOLVE $[\varepsilon, \mathbf{M}, \mathbf{z}]$ *produces after finitely many steps an output $\overline{\mathbf{q}}_\varepsilon$ (which, according to Proposition 4.1, always satisfies $\|\mathbf{q} - \overline{\mathbf{q}}_\varepsilon\| < \varepsilon$) with the following properties: For $s$ and $\tau$ related by (4.19), one has*

$$\#\operatorname{supp}\overline{\mathbf{q}}_\varepsilon \lesssim \varepsilon^{-1/s}\|\mathbf{q}\|_{\ell_\tau^w}^{1/s}, \qquad \|\overline{\mathbf{q}}_\varepsilon\|_{\ell_\tau^w} \lesssim \|\mathbf{q}\|_{\ell_\tau^w}, \tag{4.22}$$

*and the number of floating point operations needed to compute $\overline{\mathbf{q}}_\varepsilon$ remains proportional to $\#\operatorname{supp}\overline{\mathbf{q}}_\varepsilon$.*

Thus, $\tau^*$-sparsity of the routine RES implies asymptotically optimal work/accuracy rates of the scheme SOLVE for a certain range of decay rates given by $\tau^*$. We stress that the algorithm itself does *not* require any a–priori knowledge about the solution such as its actual best $N$–term approximation rate. Theorem 4.3 also shows that controlling the $\ell_\tau^w$–norm of the quantities generated in the course of the computation is crucial. This finally explains the role of COARSE in step (II.2) of SOLVE through the following result from [CDD1].

LEMMA 4.4 (Coarsening Lemma). *Let $\mathbf{v} \in \ell_\tau^w$ and let $\mathbf{w}$ be any finitely supported approximation such that $\|\mathbf{v}-\mathbf{w}\| \le \frac{1}{5}\eta$. Then the output $\mathbf{w}_\eta$ of* COARSE$[\frac{4}{5}\eta, \mathbf{w}]$ *satisfies*

$$\#\operatorname{supp}\mathbf{w}_\eta \lesssim \|\mathbf{v}\|_{\ell_\tau^w}^{1/\tau}\,\eta^{-1/s}, \quad \|\mathbf{v}-\mathbf{w}_\eta\| \lesssim \eta, \quad \text{and} \quad \|\mathbf{w}_\eta\|_{\ell_\tau^w} \lesssim \|\mathbf{v}\|_{\ell_\tau^w}. \tag{4.23}$$

Thus, knowing an error bound for a given finitely supported approximation $\mathbf{w}$, a certain coarsening using only knowledge about $\mathbf{w}$, produces a new approximation to (the possibly unknown) $\mathbf{v}$ which gives rise to with a slightly larger error but realizes up to a uniform constant the optimal relation between support and accuracy. In the scheme SOLVE this means that by the coarsening step the $\ell_\tau^w$–norms of the iterates $\mathbf{v}^K$ are controlled. Recall from (4.11) that the choice of the constant $K$ in (4.7), which controls the number of iterations in step (II.1), guarantees that in the $(j+1)$st outer iteration of SOLVE the iterate $\mathbf{v}^K$ satisfies $\|\mathbf{q}-\mathbf{v}^K\| \le \frac{1}{10}\varepsilon_j$. The threshold $\frac{2}{5}\varepsilon_j$ in step (II.2) assures, on account of (4.23), that the error after coarsening is still bounded by $\frac{1}{2}\varepsilon_j$. At the same time, if $\mathbf{q} \in \ell_\tau^w$, then $\|\overline{\mathbf{q}}^j\|_{\ell_\tau^w}$ remains bounded and $\#\operatorname{supp}\overline{\mathbf{q}}^j$ increases at most like $\varepsilon_j^{-1/\tau}$ which is the best possible $N$–term rate for sequences in $\ell_\tau^w$. Thus to ensure an overall optimal work/accuracy rate one has to show that the $\ell_\tau^w$–norms of the intermediate iterates $\mathbf{v}^k$ in step (II.1) of SOLVE cannot grow too much which is indeed guaranteed by $\tau^*$–sparsity.

The remainder of this paper is now devoted to the construction and analysis of a concrete realization of SOLVE – termed $\text{SOLVE}_{\text{DCP}}$ – for the problem (DCP) such that the corresponding routine $\text{RES}_{\text{DCP}}$ is $\tau^*$-sparse.

**5. The Scheme $\text{SOLVE}_{\text{DCP}}$.** Since $\mathbf{Q}$ from (3.35) involves several inverses of matrices it is not so clear how to realize a residual approximation in an economical way – recall obstruction (b) in Section 3.2. Therefore we shall first consider several equivalent formulations of (DCP).

**5.1. Auxiliary Formulations of (DCP).** It will be helpful to derive equivalent formulations that better support numerical realizations. Substituting as before $\mathbf{u} := \mathbf{D}_U^{-1}\breve{\mathbf{u}}$, we define the Lagrangian and introduce as an additional variable the Lagrange parameter $\mathbf{p} \in \ell_2(\mathcal{I}_Y)$,

$$\text{Lagr}(\mathbf{y},\mathbf{p},\mathbf{u}) := \widetilde{\mathbf{J}}(\mathbf{y},\mathbf{u}) + \langle \mathbf{p}, \mathbf{A}\mathbf{y} - \boldsymbol{f} - \mathbf{E}\mathbf{D}_U\mathbf{u}\rangle, \tag{5.1}$$

where $\widetilde{\mathbf{J}}(\mathbf{y},\mathbf{u}) = \frac{1}{2}\|\mathbf{D}_Z^{-1}(\mathbf{T}\mathbf{y}-\mathbf{y}_*)\|^2 + \frac{\omega}{2}\|\mathbf{u}\|^2$ has been defined in (3.27). Straightforward calculations yield the first–order Euler–Lagrange equations whose solution also yields the minimizer of (3.38), see e.g. [Z] or [K2].

REMARK 5.1. *The solution* $\mathbf{u}$ *of the system (3.36) is a component of the solution* $(\mathbf{y}, \mathbf{p}, \mathbf{u})$ *of the weakly coupled system of Euler equations in wavelet coordinates*

$$\textbf{(EE)} \qquad \begin{aligned} \mathbf{A}\mathbf{y} &= \boldsymbol{f} + \mathbf{E}\mathbf{D}_U\mathbf{u} \\ \mathbf{A}^T\mathbf{p} &= -\mathbf{T}^T\mathbf{D}_Z^{-2}\,(\mathbf{T}\mathbf{y} - \mathbf{y}_*) \qquad (5.2) \\ \omega\mathbf{u} &= \mathbf{D}_U\mathbf{E}^T\mathbf{p}. \qquad (5.3) \end{aligned}$$

The first equation of (EE) is often denoted as the *state* or *primal equation* while the second equation is called the *costate* or *adjoint equation.*

Note that either a scaled version of the Lagrange multiplier $\mathbf{p}$ or of its image under $\mathbf{E}^T$ (recall the cases (I), (II) from Section 3.2) agrees with the optimal control.

The system (EE) can, of course, be reformulated as a saddle point problem

$$\begin{pmatrix} \omega\mathbf{I} & \mathbf{0} & -\mathbf{D}_U\mathbf{E}^T \\ \mathbf{0} & \mathbf{T}^T\mathbf{D}_Z^{-2}\mathbf{T} & \mathbf{A}^T \\ -\mathbf{E}\mathbf{D}_U & \mathbf{A} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \mathbf{y} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{T}^T\mathbf{D}_Z^{-2}\mathbf{y}_* \\ \boldsymbol{f} \end{pmatrix}. \qquad (5.4)$$

In particular, in the case (I) when using natural norms in (2.3), i.e. $\mathbf{D}_U = \mathbf{D}_Z = \mathbf{I}$, we have

$$\begin{pmatrix} \omega\mathbf{I} & \mathbf{0} & -\mathbf{I} \\ \mathbf{0} & \mathbf{I} & \mathbf{A}^T \\ -\mathbf{I} & \mathbf{A} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \mathbf{y} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{y}_* \\ \boldsymbol{f} \end{pmatrix}. \qquad (5.5)$$

Due to the trivial form of the upper left two by two block, this system satisfies the *inf–sup* condition and, therefore, defines a boundedly invertible mapping on $\ell_2(I\!I_Y)^3$. Thus, one can immediately apply the results from [DDU] on adaptive Uzawa iterations for well posed saddle point problems. Corresponding optimal complexity estimates apply whenever the matrix $\mathbf{A}$ is compressible, which is the case in all the above examples.

In general, however, when $\mathbf{T}$ is a trace operator, the block $\mathbf{T}^T\mathbf{D}_Z^{-2}\mathbf{T}$ has a non-trivial kernel. In order to apply the Uzawa strategy, one first has to stabilize the system so as to have a well-defined Schur complement. This can be done along the lines described in [DDU]. Here we prefer the formally somewhat different approach based on the system (EE). This approach also applies in principle to constraints in form of a saddle point system as pointed out in Remark 2.2.

Our strategy for approximating in each step the *residual* $\mathbf{g} - \mathbf{Q}\mathbf{u}^k$ will be based upon the following observation, namely, that the residual of (3.36) is just the residual of the third equation in (EE).

LEMMA 5.2. *For any* $\mathbf{v} \in \ell_2(I\!I_Q)$*, one has the representation*

$$\mathbf{Q}\mathbf{v} - \mathbf{g} = \omega\mathbf{v} - \mathbf{E}_U^T\mathbf{p}, \quad \textit{where } \mathbf{E}_U := \mathbf{E}\mathbf{D}_U, \qquad (5.6)$$

*where* $\mathbf{p}$ *is the solution of the first two equations in (EE), i.e., for any given* $\mathbf{v}$*, the sequence* $\mathbf{p}$ *is determined by solving*

$$\mathbf{A}^T\mathbf{p} = -\mathbf{T}_Z^T\,(\mathbf{T}_Z\mathbf{y} - \mathbf{y}_Z)\,, \quad \textit{where } \mathbf{A}\mathbf{y} = \boldsymbol{f} + \mathbf{E}_U\mathbf{v}, \quad \mathbf{T}_Z := \mathbf{D}_Z^{-1}\mathbf{T}, \qquad (5.7)$$

*and* $\mathbf{y}_Z = \mathbf{D}_Z^{-1}\mathbf{y}_*$ *is defined in (3.18).*

*Proof.* By definition we infer from (3.35) and (3.30) that

$$\mathbf{Q}\mathbf{v} - \mathbf{g} = \omega\mathbf{v} + \mathbf{Z}^T(\mathbf{Z}\mathbf{v} - \mathbf{D}_Z^{-1}(\mathbf{y}_* - \mathbf{T}\mathbf{A}^{-1}\boldsymbol{f})).$$

Recalling the definition of $\mathbf{Z}$ from (3.35), the second term on the right hand side of this equality can be written as

$$\mathbf{Z}^T\mathbf{D}_Z^{-1}\mathbf{T}\mathbf{A}^{-1}\mathbf{E}\mathbf{D}_U\mathbf{v} - \mathbf{Z}^T\mathbf{D}_Z^{-1}(\mathbf{y}_* - \mathbf{T}\mathbf{A}^{-1}\boldsymbol{f}) \\ = \mathbf{Z}^T\mathbf{D}_Z^{-1}\mathbf{T}\mathbf{A}^{-1}(\mathbf{E}\mathbf{D}_U\mathbf{v} + \boldsymbol{f}) - \mathbf{Z}^T\mathbf{D}_Z^{-1}\mathbf{y}_*.$$

Thus, taking $\mathbf{y}$ as the solution of the second equation in (5.7), this reduces to

$$-\mathbf{Z}^T\mathbf{D}_Z^{-1}(\mathbf{y}_* - \mathbf{T}\mathbf{y}) = -\mathbf{D}_U\mathbf{E}^T\mathbf{A}^{-T}\mathbf{T}^T\mathbf{D}_Z^{-2}(\mathbf{y}_* - \mathbf{T}\mathbf{y}) = -\mathbf{D}_U\mathbf{E}^T\mathbf{p},$$

where we have used the first equation in (5.7) and the definition of $\mathbf{T}_Z$. This finishes the proof. □

Since the entries of $\mathbf{D}_Z^{-1}, \mathbf{D}_U$ are nonincreasing in scale (and assuming without loss of generality that they are all bounded by one), we infer from (3.28) that one still has

$$\|\mathbf{T}_Z\mathbf{v}\| \le C_{\mathbf{T}}\|\mathbf{v}\|, \quad \|\mathbf{E}_U\mathbf{v}\| \le C_{\mathbf{E}}\|\mathbf{v}\|. \tag{5.8}$$

It will be convenient in the sequel to be able to refer to the equations (EE) with the notation from Lemma 5.2 as the system

$$\textbf{(EEn)} \qquad \begin{aligned} \mathbf{A}\mathbf{y} &= \boldsymbol{f} + \mathbf{E}_U\mathbf{u} \\ \mathbf{A}^T\mathbf{p} &= -\mathbf{T}_Z^T(\mathbf{T}_Z\mathbf{y} - \mathbf{y}_Z) \\ \omega\mathbf{u} &= \mathbf{E}_U^T\mathbf{p}. \end{aligned}$$

**5.2. Realization of the Routine $\text{RES}_{\text{DCP}}$.** The realization of the routine $\text{RES}_{\text{DCP}}$ for the problem (3.36) will be based on the residual representation (5.6) in Lemma 5.2. However, this requires solving the two auxiliary systems in (EEn). Since the residual has to be only approximated, these systems will have to be solved only approximately. These approximate solutions, in turn, will be provided again by calls of the scheme SOLVE but this time with respect to suitable residual schemes tailored to the systems in (EEn). In all our examples the matrix $\mathbf{A}$ appearing in (EEn) is symmetric positive definite and the choice of wavelet bases ensures the validity of (3.15). Hence, (4.2) and (4.3) hold for $\mathbf{M} = \mathbf{A}$ and $\mathbf{M} = \mathbf{A}^T$ so that the scheme SOLVE can indeed be invoked. We hasten to mention, however, that the symmetry and positive definiteness of $\mathbf{A}$ is not essential. As long as (3.15) holds, which means that the operator equation induced by the constraints is well–posed, (which is still the case, e.g., for many saddle point problems) we can multiply the systems in (EEn) by $\mathbf{A}^T$, respectively $\mathbf{A}$, to arrive at a least squares formulation with $\mathbf{M} = \mathbf{A}^T\mathbf{A}$ or $\mathbf{M} = \mathbf{A}\mathbf{A}^T$, still satisfying (4.2) but now yielding symmetric positive definite systems to ensure (4.3). However, in order to keep the exposition as simple as possible, we confine the following discussion to the case that $\mathbf{A}$ already satisfies (in addition to (3.15)) (4.3).

Note also that, although we conceptually use the fact that a gradient iteration for the reduced problem (3.36) reduces the error for $\mathbf{u}$ in each step by a fixed amount, the use of (EEn) for the evaluation of the residuals will generate as byproducts approximate solutions to the full Euler–Lagrange system, i.e., we shall obtain approximations to the exact solution triple $(\mathbf{y}, \mathbf{p}, \mathbf{u})$ of (EEn).

Under this hypothesis, we have to formulate next the ingredients for suitable versions $\text{SOLVE}_{\text{PRM}}$ and $\text{SOLVE}_{\text{ADJ}}$ of SOLVE for the systems in (EEn). Specifically, this requires identifying residual schemes $\text{RES}_{\text{PRM}}$ and $\text{RES}_{\text{ADJ}}$ for the systems $\text{SOLVE}_{\text{PRM}}$

and SOLVE$_{\text{ADJ}}$. The main task in both cases is to apply operators $\mathbf{A}, \mathbf{A}^T, \mathbf{T}_Z, \mathbf{T}_Z^T$, $\mathbf{E}_U$ and $\mathbf{E}_U^T$. Again we assume for the moment that routines for the application of these operators are available, i.e., that for any $\mathbf{L} \in \{\mathbf{A}, \mathbf{A}^T, \mathbf{T}_Z, \mathbf{T}_Z^T, \mathbf{E}_U, \mathbf{E}_U^T\}$ we have a scheme with the following property at our disposal. We shall later discuss their concrete realization.

APPLY$[\eta, \mathbf{L}, \mathbf{v}] \to \mathbf{w}_\eta$ DETERMINES FOR ANY FINITELY SUPPORTED INPUT VECTOR $\mathbf{v}$ AND ANY TOLERANCE $\eta > 0$ A FINITELY SUPPORTED OUTPUT $\mathbf{w}_\eta$ WHICH SATISFIES

$$\|\mathbf{L}\mathbf{v} - \mathbf{w}_\eta\| \le \eta. \tag{5.9}$$

The scheme SOLVE$_{\text{PRM}}$ for the first system in (EEn) is now defined by

$$\text{SOLVE}_{\text{PRM}}[\eta, \mathbf{A}, \mathbf{E}_U, \boldsymbol{f}, \mathbf{v}, \overline{\mathbf{y}}^0, \varepsilon_0] := \text{SOLVE}[\eta, \mathbf{A}, \boldsymbol{f} + \mathbf{E}_U\mathbf{v}, \overline{\mathbf{y}}^0, \varepsilon_0], \tag{5.10}$$

where $\overline{\mathbf{y}}^0$ is an initial guess for the solution $\mathbf{y}$ of $\mathbf{A}\mathbf{y} = \boldsymbol{f} + \mathbf{E}_U\mathbf{v}$ with accuracy $\varepsilon_0$ and where the scheme RES for step (II) in SOLVE is in this case realized by a new routine RES$_{\text{PRM}}$ which is defined as follows.

RES$_{\text{PRM}}[\eta, \mathbf{A}, \mathbf{E}_U, \boldsymbol{f}, \mathbf{v}, \overline{\mathbf{y}}] \to \mathbf{r}_\eta$ DETERMINES FOR ANY POSITIVE TOLERANCE $\eta$, A GIVEN FINITELY SUPPORTED $\mathbf{v}$ AND ANY FINITELY SUPPORTED INPUT $\overline{\mathbf{y}}$ A FINITELY SUPPORTED APPROXIMATE RESIDUAL $\mathbf{r}_\eta$ SATISFYING (4.5), THAT IS,

$$\|\boldsymbol{f} + \mathbf{E}_U\mathbf{v} - \mathbf{A}\overline{\mathbf{y}} - \mathbf{r}_\eta\| \le \eta, \tag{5.11}$$

AS FOLLOWS:

(I) APPLY$[\frac{1}{3}\eta, \mathbf{A}, \overline{\mathbf{y}}] \to \mathbf{w}_\eta$;

(II) COARSE$[\frac{1}{3}\eta, \boldsymbol{f}] \to \boldsymbol{f}_\eta$;
APPLY$[\frac{1}{3}\eta, \mathbf{E}_U, \mathbf{v}] \to \mathbf{z}_\eta$;

(III) SET $\mathbf{r}_\eta := \boldsymbol{f}_\eta + \mathbf{z}_\eta - \mathbf{w}_\eta$.

In fact, noting that $\boldsymbol{f} + \mathbf{E}_U\mathbf{v} - \mathbf{A}\overline{\mathbf{y}} - \mathbf{r}_\eta = (\boldsymbol{f} - \boldsymbol{f}_\eta) + (\mathbf{E}_U\mathbf{v} - \mathbf{z}_\eta) + (\mathbf{w}_\eta - \mathbf{A}\overline{\mathbf{y}})$, by triangle inequality (5.11) is an immediate consequence of the choice of the tolerances in steps (I) – (III) of RES$_{\text{PRM}}$.

Similarly, we need a version of SOLVE for the approximate solution of the second system in (EEn), $\mathbf{A}^T\mathbf{p} = -\mathbf{T}_Z^T(\mathbf{T}_Z\mathbf{y} - \mathbf{y}_Z)$, which therefore depends on $\mathbf{y}_Z = \mathbf{D}_Z^{-1}\mathbf{y}_*$, an approximate solution $\overline{\mathbf{y}}$ of the primal system and possibly on some initial guess $\overline{\mathbf{p}}^0$ with accuracy $\varepsilon_0$. Specifically, we set here

$$\text{SOLVE}_{\text{ADJ}}[\eta, \mathbf{A}, \mathbf{T}_Z, \mathbf{y}_Z, \overline{\mathbf{y}}, \overline{\mathbf{p}}^0, \varepsilon_0] := \text{SOLVE}[\eta, \mathbf{A}^T, -\mathbf{T}_Z^T(\mathbf{T}_Z\overline{\mathbf{y}} - \mathbf{y}_Z), \overline{\mathbf{p}}^0, \varepsilon_0]. \tag{5.12}$$

As usual we will assume that the data $\boldsymbol{f}, \mathbf{y}_Z$ are approximated in a preprocessing step with sufficient accuracy (depending on the final target accuracy for solving (3.36)) by finite arrays whose entries are ordered by size and hence can be treated by COARSE.

Again we have to identify a suitable residual approximation scheme RES$_{\text{ADJ}}$ for step (II) of this version of SOLVE where the main issue is the approximate evaluation of the right hand side. The routine RES$_{\text{ADJ}}$ is defined as follows.

RES$_{\text{ADJ}}[\eta, \mathbf{A}, \mathbf{T}_Z, \mathbf{y}_Z, \overline{\mathbf{y}}, \overline{\mathbf{p}}] \to \mathbf{r}_\eta$ DETERMINES FOR ANY POSITIVE TOLERANCE $\eta$, GIVEN FINITELY SUPPORTED DATA $\overline{\mathbf{y}}, \mathbf{y}_Z$ AND ANY FINITELY SUPPORTED INPUT $\overline{\mathbf{p}}$ AN APPROXIMATE RESIDUAL $\mathbf{r}_\eta$ SATISFYING (4.5), I.E.,

$$\|-\mathbf{T}_Z^T(\mathbf{T}_Z\overline{\mathbf{y}} - \mathbf{y}_Z) - \mathbf{A}^T\overline{\mathbf{p}} - \mathbf{r}_\eta\| \le \eta, \tag{5.13}$$

AS FOLLOWS:

(I) APPLY$[\frac{1}{3}\eta, \mathbf{A}^T, \overline{\mathbf{p}}] \to \mathbf{w}_\eta$;

(II) APPLY$[\frac{1}{6C_\mathbf{T}}\eta, \mathbf{T}_Z, \overline{\mathbf{y}}] \to \mathbf{z}_\eta$ WITH $C_\mathbf{T}$ FROM (5.8);
COARSE$[\frac{1}{6C_\mathbf{T}}\eta, \mathbf{y}_Z] \to (\mathbf{y}_Z)_\eta$; SET $\mathbf{d}_\eta := (\mathbf{y}_Z)_\eta - \mathbf{z}_\eta$;
APPLY$[\frac{1}{3}\eta, \mathbf{T}_Z^T, \mathbf{d}_\eta] \to \mathbf{v}_\eta$;

(III) SET $\mathbf{r}_\eta := \mathbf{v}_\eta - \mathbf{w}_\eta$.

To confirm the validity of (5.13), note that by steps (I) – (III) of $\text{RES}_{\text{ADJ}}$

$$\begin{aligned}-\mathbf{T}_Z^T(\mathbf{T}_Z\overline{\mathbf{y}} - \mathbf{y}_Z) - \mathbf{A}^T\overline{\mathbf{p}} - \mathbf{r}_\eta \\ = -\mathbf{T}_Z^T\left((\mathbf{T}_Z\overline{\mathbf{y}} - \mathbf{y}_Z) - \mathbf{d}_\eta\right) + (\mathbf{T}_Z^T\mathbf{d}_\eta - \mathbf{v}_\eta) + (\mathbf{w}_\eta - \mathbf{A}^T\overline{\mathbf{p}}),\end{aligned}$$

so that (5.13) follows, in view of (5.8) and the tolerances above, by triangle inequality.

Recall that the exact solution $\mathbf{u}$ of (3.36) is the third component of the solution triple $(\mathbf{y}, \mathbf{p}, \mathbf{u})$ of the Euler–Lagrange system (EEn). We shall consistently use this notation for the exact solutions of the respective systems. We are now in a position to define the residual scheme for the version of SOLVE applied to (3.36). We shall refer to this specification as $\text{SOLVE}_{\text{DCP}}$. Likewise the corresponding residual scheme is denoted by $\text{RES}_{\text{DCP}}$. We shall use the constants from (3.15) and (5.8). Since the scheme is based on Lemma 5.2, it will therefore involve several parameters stemming from the auxiliary systems (EEn).

$\text{RES}_{\text{DCP}}[\eta, \mathbf{Q}, \mathbf{g}, \tilde{\mathbf{y}}, \delta_y, \tilde{\mathbf{p}}, \delta_p, \mathbf{v}, \delta_v] \to (\mathbf{r}_\eta, \tilde{\mathbf{y}}, \delta_y, \tilde{\mathbf{p}}, \delta_p)$ DETERMINES FOR ANY APPROXIMATE SOLUTION TRIPLE $(\tilde{\mathbf{y}}, \tilde{\mathbf{p}}, \mathbf{v})$ OF THE SYSTEM (EEN) SATISFYING

$$\|\mathbf{y} - \tilde{\mathbf{y}}\| \le \delta_y, \quad \|\mathbf{p} - \tilde{\mathbf{p}}\| \le \delta_p, \quad \|\mathbf{w} - \mathbf{v}\| \le \delta_v, \tag{5.14}$$

AN APPROXIMATE RESIDUAL $\mathbf{r}_\eta$ SUCH THAT

$$\|\mathbf{g} - \mathbf{Q}\mathbf{v} - \mathbf{r}_\eta\| \le \eta. \tag{5.15}$$

MOREOVER, THE INITIAL APPROXIMATIONS $\tilde{\mathbf{y}}, \tilde{\mathbf{p}}$ ARE OVERWRITTEN BY NEW APPROXIMATIONS $\tilde{\mathbf{y}}, \tilde{\mathbf{p}}$ SATISFYING (5.14) WITH NEW BOUNDS $\delta_y$ AND $\delta_p$ DEFINED IN (5.17) BELOW, AS FOLLOWS:

(I) $\text{SOLVE}_{\text{PRM}}[\frac{c_\mathbf{A}\eta}{3C_\mathbf{E}C_\mathbf{T}^2}, \mathbf{A}, \boldsymbol{f}, \mathbf{v}, \tilde{\mathbf{y}}, \delta_y] \to \mathbf{y}_\eta$;

(II) $\text{SOLVE}_{\text{ADJ}}[\frac{\eta}{3C_\mathbf{E}}, \mathbf{A}, \mathbf{T}_Z, \mathbf{y}_Z, \mathbf{y}_\eta, \tilde{\mathbf{p}}, \delta_p] \to \mathbf{p}_\eta$;

(III) APPLY$[\frac{\eta}{3}, \mathbf{E}_U^T, \mathbf{p}_\eta] \to \mathbf{q}_\eta$;

(IV) SET $\mathbf{r}_\eta := \mathbf{q}_\eta - \omega\mathbf{v}$;

(V) SET

$$\xi_y := \frac{C_\mathbf{E}}{c_\mathbf{A}}\,\delta_v + \frac{c_\mathbf{A}}{3C_\mathbf{E}C_\mathbf{T}^2}\,\eta, \quad \xi_p := \frac{C_\mathbf{T}^2 C_\mathbf{E}}{c_\mathbf{A}^2}\,\delta_v + \frac{2}{3C_\mathbf{E}}\,\eta, \tag{5.16}$$

AND REPLACE $\tilde{\mathbf{y}}, \delta_y$ AS WELL AS $\tilde{\mathbf{p}}, \delta_p$ BY THE NEW VALUES

$$\begin{aligned}\tilde{\mathbf{y}} &:= \text{COARSE}[4\xi_y, \mathbf{y}_\eta], & \delta_y &:= 5\,\xi_y, \\ \tilde{\mathbf{p}} &:= \text{COARSE}[4\xi_p, \mathbf{p}_\eta], & \delta_p &:= 5\,\xi_p.\end{aligned} \tag{5.17}$$

((5.16) already indicates the conditions on the tolerance $\eta$ and the accuracy bound $\delta_v$ under which the new error bounds in (5.17) are actually tighter. The precise relation

between $\eta$ and $\delta_v$ in the context of SOLVE$_{\text{DCP}}$ will emerge from the complexity analysis in Section 6, see (6.2) below.) Let us confirm the claimed estimates (5.15) and (5.17). To this end, let for any given input $\mathbf{v}$ the exact solution to the first system in (EEn) be denoted by $\mathbf{y_v}$. Moreover, let $\mathbf{p_v}$ be the exact solution of the second system in (EEn) with $\mathbf{y}$ replaced by $\mathbf{y_v}$. Finally, let $\hat{\mathbf{p}}$ be the exact solution of the second system but with $\mathbf{y}$ replaced by the approximate solution $\mathbf{y}_\eta$ of the first equation in (EEn). By step (IV) in RES$_{\text{DCP}}$ and (5.6), we have

$$\mathbf{g} - \mathbf{Q}\mathbf{v} - \mathbf{r}_\eta = \mathbf{E}_U^T\mathbf{p_v} - \mathbf{q}_\eta = \mathbf{E}_U^T(\mathbf{p_v} - \mathbf{p}_\eta) + \mathbf{E}_U^T\mathbf{p}_\eta - \mathbf{q}_\eta.$$

Hence it follows that

$$\|\mathbf{g} - \mathbf{Q}\mathbf{v} - \mathbf{r}_\eta\| \le \frac{\eta}{3} + C_{\mathbf{E}}\|\mathbf{p}_\eta - \mathbf{p_v}\|. \tag{5.18}$$

In order to estimate the second term, note that

$$\mathbf{p_v} - \hat{\mathbf{p}} = \mathbf{A}^{-T}\mathbf{T}_Z^T\mathbf{T}_Z(\mathbf{y_v} - \mathbf{y}_\eta),$$

and therefore, by (3.15), (5.8) and step (I),

$$\|\mathbf{p_v} - \hat{\mathbf{p}}\| \le c_{\mathbf{A}}^{-1}C_{\mathbf{T}}^2\|\mathbf{y_v} - \mathbf{y}_\eta\| \le \frac{\eta}{3C_{\mathbf{E}}}. \tag{5.19}$$

Thus, by step (II) and (5.19), $\|\mathbf{p}_\eta - \mathbf{p_v}\| \le \frac{2\eta}{3C_{\mathbf{E}}}$, which together with (5.18) confirms (5.15).

Adhering to the above notational conventions, the first system in (EEn) yields $\mathbf{y} - \mathbf{y_v} = \mathbf{A}^{-1}\mathbf{E}_U(\mathbf{w} - \mathbf{v})$ so that by (5.14), (3.15) and (5.8)

$$\|\mathbf{y} - \mathbf{y}_\eta\| \le \|\mathbf{y} - \mathbf{y_v}\| + \|\mathbf{y_v} - \mathbf{y}_\eta\| \le \frac{C_{\mathbf{E}}}{c_{\mathbf{A}}}\,\delta_v + \frac{c_{\mathbf{A}}}{3C_{\mathbf{E}}C_{\mathbf{T}}^2}\,\eta, \tag{5.20}$$

which is the value of $\xi_y$ in step (V). Likewise we infer from the second system in (EEn) that

$$\mathbf{p} - \mathbf{p}_\eta = \mathbf{p} - \hat{\mathbf{p}} + \hat{\mathbf{p}} - \mathbf{p}_\eta = \mathbf{A}^{-T}\mathbf{T}_Z^T\mathbf{T}_Z(\mathbf{y} - \mathbf{y}_\eta) + \hat{\mathbf{p}} - \mathbf{p}_\eta.$$

Hence, by (3.15), (3.28) and step (II), we obtain

$$\|\mathbf{p} - \mathbf{p}_\eta\| \le \frac{C_{\mathbf{T}}^2}{c_{\mathbf{A}}}\,\xi_y + \frac{\eta}{3C_{\mathbf{E}}} = \frac{C_{\mathbf{T}}^2C_{\mathbf{E}}}{c_{\mathbf{A}}^2}\,\delta_v + \frac{2}{3C_{\mathbf{E}}}\,\eta, \tag{5.21}$$

which is the value of $\xi_p$ in step (V). The estimates (5.14) with the new bounds defined in (5.17) are now an immediate consequence of the coarsening step in (V) and the triangle inequality. This concludes the confirmation of all estimates stated in RES$_{\text{DCP}}$.

It remains to initialize the scheme SOLVE$_{\text{DCP}}$. Again we assume that $\boldsymbol{f}$ and $\mathbf{y}_Z$ are given and fully accessible. Choosing $\overline{\mathbf{u}}^0 \equiv \mathbf{0}$ we infer from (3.37), (3.30) and (3.36) that

$$\begin{aligned}
\|\overline{\mathbf{u}}^0 - \mathbf{u}\| &\le c_{\mathbf{Q}}^{-1}\|\mathbf{Q}\overline{\mathbf{w}}^0 - \mathbf{g}\| = c_{\mathbf{Q}}^{-1}\|\mathbf{g}\| \\
&= c_{\mathbf{Q}}^{-1}\|\mathbf{E}_U^T\mathbf{A}^{-T}\mathbf{T}_Z^T(\mathbf{y}_Z - \mathbf{T}_Z\mathbf{A}^{-1}\boldsymbol{f})\| \\
&\le \frac{C_{\mathbf{E}}C_{\mathbf{T}}}{c_{\mathbf{A}}}\left(\|\mathbf{y}_Z\| + \frac{C_{\mathbf{T}}}{c_{\mathbf{A}}}\|\boldsymbol{f}\|\right) \\
&=: \varepsilon_0.
\end{aligned} \tag{5.22}$$

Moreover, for $\tilde{\mathbf{y}}^0 := \mathbf{0}$ one has

$$\|\mathbf{y} - \tilde{\mathbf{y}}^0\| = \|\mathbf{A}^{-1}(\boldsymbol{f} + \mathbf{E}_U\mathbf{w})\| \le c_{\mathbf{A}}^{-1}\left(\|\boldsymbol{f}\| + C_{\mathbf{E}}c_{\mathbf{Q}}^{-1}\|\mathbf{g}\|\right)$$
$$\le c_{\mathbf{A}}^{-1}\left(\|\boldsymbol{f}\| + C_{\mathbf{E}}\varepsilon_0\right) =: \delta_{y,0}. \tag{5.23}$$

Similarly, for $\tilde{\mathbf{p}}^0 := \mathbf{0}$ we obtain

$$\|\tilde{\mathbf{p}}^0 - \mathbf{p}\| = \|\mathbf{A}^{-T}\mathbf{T}_Z^T(\mathbf{T}_Z\mathbf{y} - \mathbf{y}_Z)\| \le c_{\mathbf{A}}^{-1}(C_{\mathbf{T}}^2\delta_{y,0} + C_{\mathbf{T}}\|\mathbf{y}_Z\|) =: \delta_{p,0}. \tag{5.24}$$

The scheme SOLVE$_{\text{DCP}}$ takes now the following form with the error reduction factor $\rho = \rho(\mathbf{Q})$ from (3.40) and $K$ given by (4.7) with $\alpha$ from (3.39).

SOLVE$_{\text{DCP}}[\varepsilon, \mathbf{Q}, \mathbf{g}] \to \overline{\mathbf{u}}_\varepsilon$

(I) LET $\overline{\mathbf{q}}^0 := \mathbf{0}$ AND LET $\varepsilon_0$ BE GIVEN BY (5.22). MOREOVER, LET $\tilde{\mathbf{y}} := \mathbf{0}$, $\tilde{\mathbf{p}} := \mathbf{0}$ AND SET $j = 0$. FINALLY, LET $\delta_y := \delta_{y,0}, \delta_p := \delta_{p,0}$ BE DEFINED BY (5.23), (5.24), RESPECTIVELY.

(II) IF $\varepsilon_j \le \varepsilon$, STOP AND SET $\overline{\mathbf{u}}_\varepsilon := \overline{\mathbf{u}}^j$, $\overline{\mathbf{y}}_\varepsilon = \tilde{\mathbf{y}}$, $\overline{\mathbf{p}}_\varepsilon = \tilde{\mathbf{p}}$.
OTHERWISE SET $\mathbf{v}^0 := \overline{\mathbf{u}}^j$.

(II.1) FOR $k = 0, \dots, K-1$, COMPUTE
RES$_{\text{DCP}}[\rho^k\varepsilon_j, \mathbf{Q}, \mathbf{g}, \tilde{\mathbf{y}}, \delta_y, \tilde{\mathbf{p}}, \delta_p, \mathbf{v}^k, \delta_k] \to (\mathbf{r}^k, \tilde{\mathbf{y}}, \delta_y, \tilde{\mathbf{p}}, \delta_p)$,
WHERE $\delta_0 := \varepsilon_j$ AND $\delta_k := \rho^{k-1}(\alpha k + \rho)\varepsilon_j$;
SET

$$\mathbf{v}^{k+1} := \mathbf{v}^k + \alpha\mathbf{r}^k. \tag{5.25}$$

(II.2) APPLY COARSE$[\frac{2}{5}\varepsilon_j, \mathbf{v}^K] \to \overline{\mathbf{u}}^{j+1}$; SET $\varepsilon_{j+1} := \frac{1}{2}\varepsilon_j$, $j+1 \to j$ AND GO TO (II).

(The particular choice of the interior tolerance $\delta_k$ in step (II.1) is based the estimate (4.10).) Since when overwriting $\tilde{\mathbf{y}}$, $\tilde{\mathbf{p}}$ at the last stage prior to the termination of SOLVE$_{\text{DCP}}$ one has $\delta_v \le \varepsilon$, $\eta \le \varepsilon$, the following fact is an immediate consequence of (5.17).

REMARK 5.3. *The outputs $\overline{\mathbf{y}}_\varepsilon$ and $\overline{\mathbf{p}}_\varepsilon$ produced by* SOLVE$_{\text{DCP}}$ *in addition to $\mathbf{u}_\varepsilon$ are approximations to the exact solutions $\mathbf{y}, \mathbf{p}$ of (EEn) satisfying*

$$\|\mathbf{y} - \overline{\mathbf{y}}_\varepsilon\| \le 5\varepsilon\left(\frac{C_{\mathbf{E}}}{c_{\mathbf{A}}} + \frac{c_{\mathbf{A}}}{3C_{\mathbf{E}}C_{\mathbf{T}}^2}\right), \tag{5.26}$$

$$\|\mathbf{p} - \overline{\mathbf{p}}_\varepsilon\| \le 5\varepsilon\left(\frac{C_{\mathbf{T}}^2C_{\mathbf{E}}}{c_{\mathbf{A}}^2} + \frac{2}{3C_{\mathbf{E}}}\right). \tag{5.27}$$

**6. The Complexity of SOLVE$_{\text{DCP}}$.** In view of the definition of RES$_{\text{PRM}}$ and RES$_{\text{ADJ}}$ entering RES$_{\text{DCP}}$, the scheme SOLVE$_{\text{DCP}}$ ultimately hinges on the availability of suitable schemes APPLY for the operators $\mathbf{L} \in \{\mathbf{A}, \mathbf{A}^T, \mathbf{T}_Z, \mathbf{T}_Z^T, \mathbf{E}_U, \mathbf{E}_U^T\}$. We shall adhere to our strategy of narrowing down step by step the requirements on our algorithmic ingredients and wish to identify first conditions on the APPLY schemes that ensure $\tau^*$–sparsity of RES$_{\text{DCP}}$ as formulated in Section 4.2. It will not surprise that the key requirement is that the approximate application of each of these operators has a work/accuracy rate that is within some range comparable to best $N$–term approximation. Precisely, we say that APPLY$[\cdot, \mathbf{L}, \cdot]$ is $\tau^*$–*efficient* for some $0 < \tau^* < 2$ if for

any finitely supported $\mathbf{v} \in \ell_\tau^w$, for $0 < \tau^* < \tau < 2$, the output $\mathbf{w}_\eta$ of $\text{APPLY}[\eta, \mathbf{L}, \mathbf{v}]$ satisfies

$$\|\mathbf{w}_\eta\|_{\ell_\tau^w} \lesssim \|\mathbf{v}\|_{\ell_\tau^w}, \quad \#\text{supp}\,\mathbf{w}_\eta \lesssim \eta^{-1/s}\|\mathbf{v}\|_{\ell_\tau^w}^{1/s} \quad \eta \to 0, \tag{6.1}$$

where the constants depend only on $\tau$ as $\tau \to \tau^*$ and where $s$ is related to $\tau$ by (4.19). Moreover, the number of floating point operations needed to compute $\mathbf{w}_\eta$ is to remain proportional to $\#\text{supp}\,\mathbf{w}_\eta$.

One should note that the existence of a $\tau^*$–efficient scheme for an operator $\mathbf{L}$ has the following important implication that follows immediately from Proposition 4.2.

REMARK 6.1. *If one can find a $\tau^*$–efficient scheme for $\mathbf{L}$ then $\mathbf{L}$ is bounded on $\ell_\tau^w$ for every $\tau > \tau^*$.*

*Proof.* For convenience, the proof from [CDD1] is included here. For $\mathbf{v} \in \ell_\tau^w$ and $\eta > 0$ there exists a $\tilde{\mathbf{v}}$ with $\|\mathbf{v} - \tilde{\mathbf{v}}\| \le \eta/(2\|\mathbf{L}\|)$ and $\#\text{supp}\,\tilde{\mathbf{v}} \lesssim \eta^{-1/s}\|\mathbf{v}\|_{\ell_\tau^w}^{1/s}$. Now by definition of $\mathbf{w}_\eta = \text{APPLY}\,[\eta/2, \mathbf{L}, \tilde{\mathbf{v}}]$ and $\tau^*$–efficiency of $\mathbf{L}$ (6.1), one has for $\tau > \tau^*$, the estimate $\|\mathbf{L}\tilde{\mathbf{v}} - \mathbf{w}_\eta\| \le \eta/2$ while $\#\text{supp}\,\mathbf{w}_\eta \lesssim \eta^{-1/s}\|\tilde{\mathbf{v}}\|_{\ell_\tau^w}^{1/s} \le \eta^{-1/s}\|\mathbf{v}\|_{\ell_\tau^w}^{1/s}$. Since $\|\mathbf{L}\mathbf{v} - \mathbf{w}_\eta\| \le \eta$ we have identified a vector $\mathbf{w}_\eta$ with support $\lesssim \eta^{-1/s}$ that approximates $\mathbf{L}\mathbf{v}$ within accuracy $\eta$. Hence we can invoke Proposition 4.2 to conclude that $\|\mathbf{L}\mathbf{v}\|_{\ell_\tau^w} \lesssim \|\mathbf{v}\|_{\ell_\tau^w}$ as claimed. ∎

PROPOSITION 6.2. *If the* APPLY *schemes in* $\text{RES}_{\text{PRM}}$ *and* $\text{RES}_{\text{ADJ}}$ *are $\tau^*$–efficient for some $\tau^* < 2$, then* $\text{RES}_{\text{DCP}}$ *is $\tau^*$–sparse whenever there exists a constant $C$ such that*

$$C\eta \ge \max\{\delta_v, \delta_p\} \tag{6.2}$$

$$\max\{\|\tilde{\mathbf{p}}\|_{\ell_\tau^w}, \|\tilde{\mathbf{y}}\|_{\ell_\tau^w}, \|\mathbf{v}\|_{\ell_\tau^w}\} \le C\left(\|\mathbf{y}\|_{\ell_\tau^w} + \|\mathbf{p}\|_{\ell_\tau^w} + \|\mathbf{u}\|_{\ell_\tau^w}\right), \tag{6.3}$$

*where $\mathbf{v}$ is the current finitely supported input and where $\tilde{\mathbf{y}}, \tilde{\mathbf{p}}$ are the initial guesses for the exact solution components $(\mathbf{y}, \mathbf{p})$.*

*Proof.* Since $\text{SOLVE}_{\text{DCP}}$ actually determines an approximation to the full triple $(\mathbf{y}, \mathbf{p}, \mathbf{u})$, the notion of $\tau^*$–sparseness of $\text{RES}_{\text{DCP}}$ refers to properties of the whole triple. Thus, we have to assume that each of the solution components belongs to $\ell_\tau^w$ for some $\tau > \tau^*$. By Remark 6.1 and our hypothesis on $\tau^*$–efficiency, each $\mathbf{L} \in \{\mathbf{A}, \mathbf{A}^T, \mathbf{T}_Z, \mathbf{T}_Z^T, \mathbf{E}_U, \mathbf{E}_U^T\}$ is bounded on $\ell_\tau^w$ for $\tau > \tau^*$. Thus, for the first system in (EEn) this implies

$$\|\boldsymbol{f}\|_{\ell_\tau^w} \lesssim \|\mathbf{y}\|_{\ell_\tau^w} + \|\mathbf{u}\|_{\ell_\tau^w}. \tag{6.4}$$

Likewise we have

$$\|\mathbf{T}_Z^T\mathbf{y}_Z\|_{\ell_\tau^w} \lesssim \|\mathbf{p}\|_{\ell_\tau^w} + \|\mathbf{y}\|_{\ell_\tau^w}. \tag{6.5}$$

Now, by the assumption (6.2), the quotients $\delta_v/\eta$, $\delta_p/\eta$ are bounded. Therefore, according to step (I) in $\text{RES}_{\text{DCP}}$, the scheme $\text{SOLVE}_{\text{PRM}}$ will invoke only a uniformly bounded finite number of iteration blocks (II) with corresponding residual approximations $\text{RES}_{\text{PRM}}$. From the $\tau^*$–efficiency of $\mathbf{A}$ and $\mathbf{E}_U$ and Remark 6.1, we infer that

$$\|\mathbf{y}_\eta\|_{\ell_\tau^w} \lesssim \|\boldsymbol{f}\|_{\ell_\tau^w} + \|\mathbf{v}\|_{\ell_\tau^w} + \|\tilde{\mathbf{y}}\|_{\ell_\tau^w} \lesssim \|\mathbf{y}\|_{\ell_\tau^w} + \|\mathbf{u}\|_{\ell_\tau^w} + \|\mathbf{v}\|_{\ell_\tau^w} + \|\tilde{\mathbf{y}}\|_{\ell_\tau^w}, \tag{6.6}$$

where we have used (6.4). Likewise one concludes that the output $\mathbf{p}_\eta$ of step (II) of $\text{RES}_{\text{DCP}}$ satisfies

$$\begin{aligned}\|\mathbf{p}_\eta\|_{\ell_\tau^w} &\lesssim \|\tilde{\mathbf{p}}\|_{\ell_\tau^w} + \|\mathbf{T}_Z^T \mathbf{y}_Z\|_{\ell_\tau^w} + \|\tilde{\mathbf{y}}\|_{\ell_\tau^w} \\ &\lesssim \|\tilde{\mathbf{p}}\|_{\ell_\tau^w} + \|\tilde{\mathbf{y}}\|_{\ell_\tau^w} + \|\mathbf{p}\|_{\ell_\tau^w} + \|\mathbf{y}\|_{\ell_\tau^w}, \end{aligned} \tag{6.7}$$

where we have used (6.5) in the last step. (4.21) follows now from (6.6), (6.7) and (6.3). The second part of (4.21) and (ii) of the $\tau^*$–sparseness of $\text{RES}_{\text{DCP}}$ can be concluded from $\tau^*$–efficiency of the APPLY schemes in $\text{RES}_{\text{PRM}}$ and $\text{RES}_{\text{ADJ}}$. This confirms the claim. □

THEOREM 6.3. *Assume that the* APPLY *schemes appearing in* $\text{RES}_{\text{PRM}}$ *and* $\text{RES}_{\text{ADJ}}$ *are* $\tau^*$*-efficient for some* $\tau^* < 2$ *and that the components of the solution* $(\mathbf{y}, \mathbf{p}, \mathbf{u})$ *of (EEn) all belong to the respective space* $\ell_\tau^w$ *for some* $\tau > \tau^*$. *Then the approximate solutions* $\mathbf{y}_\varepsilon, \mathbf{p}_\varepsilon, \mathbf{u}_\varepsilon$, *produced by* $\text{SOLVE}_{\text{DCP}}$ *for any target accuracy* $\varepsilon$, *satisfy*

$$\|\mathbf{y}_\varepsilon\|_{\ell_\tau^w} + \|\mathbf{p}_\varepsilon\|_{\ell_\tau^w} + \|\mathbf{u}_\varepsilon\|_{\ell_\tau^w} \lesssim \|\mathbf{y}\|_{\ell_\tau^w} + \|\mathbf{p}\|_{\ell_\tau^w} + \|\mathbf{u}\|_{\ell_\tau^w}, \tag{6.8}$$

*and*

$$(\#\operatorname{supp}\mathbf{y}_\varepsilon) + (\#\operatorname{supp}\mathbf{p}_\varepsilon) + (\#\operatorname{supp}\mathbf{u}_\varepsilon) \lesssim \left(\|\mathbf{y}\|_{\ell_\tau^w}^{1/s} + \|\mathbf{p}\|_{\ell_\tau^w}^{1/s} + \|\mathbf{u}\|_{\ell_\tau^w}^{1/s}\right) \varepsilon^{-1/s}, \tag{6.9}$$

*where the constants depend only on* $\tau$ *when* $\tau$ *approaches* $\tau^*$. *Moreover, the number of floating point operations required during the execution of* $\text{SOLVE}_{\text{DCP}}$ *remains proportional to the right hand side of (6.9).*

*Proof.* According to Theorem 4.3, it remains to show that at each stage when $\text{RES}_{\text{DCP}}$ is called in step (II.1) of $\text{SOLVE}_{\text{DCP}}$, the hypotheses (6.2) and (6.3) in Proposition 6.2 are satisfied for some fixed constant $C$. The claim follows then from Theorem 4.3.

The validity of (6.2) is a consequence of the bounds (5.17) for the initial guesses, the values of $\eta$ and $\delta_k$ in the $k$th perturbed iteration of the $(j+1)$st call of step (II.1) of $\text{SOLVE}_{\text{DCP}}$, and the initialization bounds (5.22), (5.23) and (5.24). By the coarsening Lemma 4.4 and the coarsening in step (V) of $\text{RES}_{\text{DCP}}$, we know that

$$\|\tilde{\mathbf{y}}\|_{\ell_\tau^w} \lesssim \|\mathbf{y}\|_{\ell_\tau^w}, \quad \|\tilde{\mathbf{p}}\|_{\ell_\tau^w} \lesssim \|\mathbf{p}\|_{\ell_\tau^w}. \tag{6.10}$$

Moreover, since in the $(j+1)$st call of step (II) in $\text{SOLVE}_{\text{DCP}}$ $\mathbf{v}^K$ satisfies $\|\mathbf{u} - \mathbf{v}^K\| \leq \varepsilon_j/10$, see [CDD2] or (4.10), we conclude from step (II.2) in $\text{SOLVE}_{\text{DCP}}$ and Lemma 4.4 that

$$\|\overline{\mathbf{u}}^j\|_{\ell_\tau^w} \lesssim \|\mathbf{u}\|_{\ell_\tau^w}, \quad j \in I\!N_0. \tag{6.11}$$

Combining (6.10) and (6.11), confirms the validity of (6.3). □

Thus the practical realization of $\text{SOLVE}_{\text{DCP}}$ providing optimal work/accuracy rates for a possibly large range of decay rates of the error of best $N$–term approximation hinges on the availability of $\tau*$–efficient APPLY schemes with possibly small $\tau^*$ for the involved operators.

*Distributed Controls.* In this regard we discuss first the example in Section 2.2.1 for *natural norms*, i.e., $Z = H_0^1(\Omega)$ and $U = Y' = Q = H^{-1}(\Omega)$. In this case, one has $\mathbf{E} = \mathbf{T} = \mathbf{D}_Z = \mathbf{D}_U = \mathbf{I}$ and $\mathbf{A} = \mathbf{A}^T$. Since the identity mapping is $\tau^*$–efficient for any $\tau^* < 2$ we only have to discuss the $\tau^*$–efficiency of $\mathbf{A}$ defined by (3.14). The fact that one can indeed devise efficient schemes for the approximate application of

wavelet representations of a wide class of operators, including differential operators, is a consequence of the cancellation properties (3.3) of wavelets together with the norm equivalences (3.5) for the relevant function spaces. In fact, such representations turn out to be *quasi–sparse* in the following sense. Recall that a matrix $\mathbf{A}$ is called $s^*$*–compressible*, if for any $0 < s < s^*$ there exists a matrix $\mathbf{A}_j$ with at most $\leq \alpha_j 2^j$ nonzero entries per row and column such that

$$\|\mathbf{A} - \mathbf{A}_j\| \leq \alpha_j 2^{-sj}, \qquad j \in I\!N_0, \tag{6.12}$$

where $\{\alpha_j\}_{j \in I\!N_0}$ is any summable sequence.

Denote for a finitely supported vector $\mathbf{v}$ its best $2^j$–aproximations (given by the $2^j$ largest wavelet coefficients) by $\mathbf{v}_{[j]} := \mathbf{v}_{2^j}$. Following [CDD1], the expansion

$$\mathbf{w}_j := \mathbf{A}_j \mathbf{v}_{[0]} + \mathbf{A}_{j-1}(\mathbf{v}_{[1]} - \mathbf{v}_{[0]}) + \cdots + \mathbf{A}_0(\mathbf{v}_{[j]} - \mathbf{v}_{[j-1]}) \tag{6.13}$$

approximates $\mathbf{Av}$. In fact, combining the *a-priori* knowledge (6.12) with the *a-posteriori* information $\|\mathbf{v}_{[k]} - \mathbf{v}_{[k-1]}\|$, one can see that for any finitely supported input $\mathbf{v}$ the error $\|\mathbf{Av} - \mathbf{w}_j\|$ tends to zero when $j$ grows. Thus, given a tolerance $\eta > 0$, one chooses the smallest $j$ so that the bound for $\|\mathbf{Av} - \mathbf{w}_j\|$ is less than or equal to $\eta$. This leads to a concrete scheme with the following properties.

APPLY $[\eta, \mathbf{A}, \mathbf{v}] \to \mathbf{w}_\eta$ COMPUTES FOR A GIVEN TOLERANCE $\eta > 0$ A FINITELY SUPPORTED SEQUENCE $\mathbf{w}_\eta$ SATISFYING

$$\|\mathbf{Av} - \mathbf{w}_\eta\| \leq \eta. \tag{6.14}$$

A detailed description and analysis of this routine can be found in [CDD1]. Its implementation has been discussed in [BCDU]. The following essential complexity estimate is taken from [CDD1].

THEOREM 6.4. *If* $\mathbf{A}$ *is* $s^*$*–compressible, then* $\mathbf{A}$ *is bounded on* $\ell_\tau^w$ *for* $s < s^*$, *where* $\tau$ *and* $s$ *are related by (4.19),* $\frac{1}{\tau} = s + \frac{1}{2}$. *Moreover, for a finitely supported vector* $\mathbf{v}$ *the output* $\mathbf{w}_\eta$ *of* APPLY $[\eta, \mathbf{A}, \mathbf{v}]$ *satisfies*

$$\|\mathbf{w}_\eta\|_{\ell_\tau^w} \lesssim \|\mathbf{v}\|_{\ell_\tau^w}, \qquad \#\operatorname{supp} \mathbf{w}_\eta, \ \#\text{flops} \lesssim \eta^{-1/s} \|\mathbf{v}\|_{\ell_\tau^w}^{1/s}. \tag{6.15}$$

Thus, the above scheme APPLY is $\tau^*$-efficient for $\tau^* = (s^* + 1/2)^{-1}$ whenever $\mathbf{A}$ is $s^*$-compressible. It is known that $s^*$ is the larger the higher the regularity and the order of cancellation properties of the wavelets are for all the differential operators considered in Section 2. Bounds for $s^*$ in terms of these quantities for families of spline wavelets can be found, e.g., in [BCDU]. Hence, Theorem 6.3 ensures asymptotically optimal complexity bounds in the range $\tau > \tau^*$, i.e., the scheme $\text{SOLVE}_{\text{DCP}}$ recovers rates of the error of best $N$-term approximation of order $N^{-s}$ for $s < s^*$.

Now consider the same example but with a strictly larger space $Z \supset Y$ and a strictly smaller space $U \subset Q = Y'$. While one still has $\mathbf{E} = \mathbf{T} = \mathbf{I}$, the matrices $\mathbf{D}_Z, \mathbf{D}_U$ are nontrivial scalings of the form $\mathbf{D}_Z = \mathbf{D}^\alpha$, i.e., $(\mathbf{D}_Z)_{\lambda,\nu} = 2^{\alpha|\lambda|}\delta_{\lambda,\nu}$, and $\mathbf{D}_U = \mathbf{D}^{-\beta}$ for some positive numbers $\alpha, \beta > 0$. The system (EEn) then takes the form

$$\begin{aligned} \mathbf{Ay} &= \boldsymbol{f} + \mathbf{D}^{-\beta}\mathbf{u} \\ \mathbf{A}^T\mathbf{p} &= -\mathbf{D}^{-2\alpha}(\mathbf{y} - \mathbf{y}_*) \\ \omega\mathbf{u} &= \mathbf{D}^{-\beta}\mathbf{p}. \end{aligned} \tag{6.16}$$

First of all, the scaling smoothes the right hand sides of the first two equations. However, it also says that the components $\mathbf{p}$ and $\mathbf{u}$ belong to different sparsity classes. In fact, a diagonal scaling results in a shift between weak $\ell_\tau$–spaces. In particular, scaling with $\mathbf{D}^{-\beta}$ for $\beta > 0$ makes a sequence more compressible, i.e., results in a smaller value for $\tau$. Recall that $d$ denotes the spatial dimension of the underlying domain.

PROPOSITION 6.5. *One has that*

$$\mathbf{p} \in \ell_\tau^w \quad \textit{implies} \quad \mathbf{D}^{-\beta}\mathbf{p} \in \ell_{\tau'}^w, \quad \textit{where} \quad \frac{1}{\tau'} := \frac{1}{\tau} + \frac{\beta}{d}. \tag{6.17}$$

*Moreover, this result is sharp in the sense that for no $\tau'' < \tau'$ there holds $\mathbf{D}^{-\beta}\mathbf{p} \in \ell_{\tau''}^w$ for all $\mathbf{p} \in \ell_\tau^w$.*

*Proof.* Let $C > 1$ be a fixed constant that will later be chosen at our convenience. Let $\mathcal{P}$ be the class of those $\mathbf{p} \in \ell_\tau^w$ with $\|\mathbf{p}\|_{\ell_\tau^w} \le C^{1/\tau}$ and let $\tilde{\mathbf{p}} := \mathbf{D}^{-\beta}\mathbf{p}$. Consider the set

$$\tilde{\Lambda}_{(J)} := \{\lambda : |\tilde{p}_\lambda| \ge \eta_J\} \quad \text{where } \eta_J := 2^{-J\tau/\tau'}.$$

In view of the definition of $\ell_{\tau'}^w$ in (4.15), we have to show that the cardinality of $\tilde{\Lambda}_{(J)}$ increases at most like $2^{J\tau}$. (Standard arguments imply then that $\#\{\lambda : |\tilde{p}_\lambda| \ge \eta\} \lesssim \eta^{-\tilde{\tau}}$ for any $\eta \le 1$). To this end, we determine first which of the sets

$$\Lambda_j := \{\lambda : 2^{-j} \le |p_\lambda| < 2^{-j+1}\}$$

is always fully contained in $\tilde{\Lambda}_{(J)}$. We know from [CDD1] that $\#\Lambda_j \le C2^{j\tau}$. Without loss of generality we may assume that $\#\Lambda_j = C2^{j\tau}$ to cover the largest possible sets. Since the entries of $\tilde{\mathbf{p}}$ arise from those of $\mathbf{p}$ by scaling with the weights $2^{-\beta|\lambda|} \le 1$, the smaller the levels $|\lambda|$ in these weights, the better is the chance for $\lambda \in \Lambda_j$ to belong to $\tilde{\Lambda}_{(J)}$. Thus, to ensure that for any $\mathbf{p} \in \mathcal{P}$ the set $\Lambda_j$ is contained in $\tilde{\Lambda}_{(J)}$, we must be able to find $C2^{j\tau}$ indices $\lambda$ with possibly small $|\lambda|$ such that $2^{-\beta|\lambda|}2^{-j} \ge \eta_J$. This count clearly involves the spatial dimension $d$ since $c2^{jd}$ indices $\lambda$ of level $|\lambda| = j$ can occur. Here the constant $c$ depends on the spatial dimension of the functions whose wavelet coefficients are considered. Thus, the smallest possible maximum level $L_j$ of these indices is therefore determined by $c2^{dL_j} = C2^{j\tau}$, i.e., $L_j = j\tau/d + \left(\log_2 \frac{C}{c}\right)/d$. Assume for convenience that $C/c \ge 1$. Then, for $\lambda \in \Lambda_j$ we conclude

$$|\tilde{p}_\lambda| \ge 2^{-\beta|\lambda|}2^{-j} \ge 2^{-(\beta L_j + j)} = 2^{-j\left(\frac{d+\beta\tau}{d}\right)}(C/c)^{\beta/d} \ge 2^{-j\tau/\tau'}. \tag{6.18}$$

Thus, $\Lambda_j \subset \tilde{\Lambda}_{(J)}$ for $j \le J$. On the other hand, since for $\lambda \in \Lambda_j$ one also has $|\tilde{p}_\lambda| \le 2^{-\beta|\lambda|-j+1}$, not all of the indices in $\Lambda_j$ can be always contained in $\tilde{\Lambda}_{(J)}$ for $j \ge J$ and any choice of $\mathbf{p} \in \mathcal{P}$. To determine the maximum number of $\lambda \in \Lambda_{J+k}$ for $k \in \mathbb{N}$ that can belong to $\tilde{\Lambda}_{(J)}$ for any $\mathbf{p} \in \mathcal{P}$, we must have in view of (6.18) $|\lambda| \le \ell_k$, where $\beta\ell_k + J + k \le J\tau/\tau'$. Using that $(\tau/\tau') - 1 = \beta\tau/d$, straightforward calculations yield $\ell_k \le \frac{J\tau}{d} - \frac{k}{\beta}$. Thus, we can assign at most $2^{d\ell_k}$ scaling weights to $\Lambda_j$ to keep $|\tilde{p}_\lambda| \ge \eta_J$ for that many $\lambda \in \Lambda_{J+k}$. Moreover, the set $\Lambda_{J+k}$ is disjoint from $\tilde{\Lambda}_{(J)}$ whenever $2^{-(J+k)+1} \le \eta_J$ which is the case when $k \ge 1 + J\tau\beta/d$. Hence, we have

$$\sum_{k\in\mathbb{N}} \#(\Lambda_{J+k} \cap \tilde{\Lambda}_{(J)}) \le \sum_{k=1}^{J\tau\beta/d} 2^{d\ell_k} = \sum_{k=1}^{J\tau\beta/d} 2^{\frac{d}{\beta}\left(\frac{J\tau\beta}{d} - k\right)} \lesssim 2^{J\tau}. \tag{6.19}$$

Since $\mathcal{I} = \bigcup_{j\geq 0}\Lambda_j$ and $\sum_{j\leq J}\#\Lambda_j \lesssim 2^{J\tau}$, we conclude from (6.19) that $\#\tilde{\Lambda}_{(J)} \lesssim 2^{J\tau} = \eta^{-\tau'}$ for $\eta = 2^{-J\tau/\tau'}$. This confirms (6.17).

To verify the rest of the assertion, consider $\mathbf{p}$ whose decreasing rearrangement is given by $p_n^* = n^{-1/\tau}$ while $\tilde{p}_n^* = 2^{-\beta(j+1)}n^{-1/\tau}$ for $2^{dj} < n \leq 2^{d(j+1)}$. Then

$$\begin{aligned}\sigma_{2^{dj}}(\tilde{\mathbf{p}})^2 &= \sum_{n>2^{dj}}(\tilde{p}_n^*)^2 \gtrsim \sum_{j=J}^{\infty}\sum_{2^{jd}<n\leq 2^{d(j+1)}} n^{-2/\tau}2^{-2\beta(j+1)} \\ &\gtrsim \sum_{j=J}^{\infty}2^{-2dj\left(\frac{1}{\tau}+\frac{\beta}{d}-\frac{1}{2}\right)} \gtrsim \left(2^{-dJs'}\right)^2,\end{aligned}$$

where $s' = \frac{1}{\tau'} - \frac{1}{2}$. Thus, by Proposition 4.2, $\mathbf{p} \notin \ell_{\tau''}^w$ for any $\tau'' < \tau'$, which finishes the proof. □

Therefore, whatever the sparsity class of the adjoint variable $\mathbf{p}$ is, the third equation in (6.16) says, in view of Proposition 6.5, that the control $\mathbf{u}$ is even sparser. Thus, although the control $\mathbf{u}$ may be accurately recovered with relatively few degrees of freedom the overall solution complexity is in the above case bounded from below by the less sparse auxiliary variable $\mathbf{p}$.

As a possible remedy one might think of introducing the variable $\tilde{\mathbf{p}} = \mathbf{D}^{-\beta}\mathbf{p}$ (in order to replace $\mathbf{p}$ by a sparser variable) and rewrite the second system in (6.16) as

$$\tilde{\mathbf{A}}^T\tilde{\mathbf{p}} = \mathbf{D}^{-\beta}\mathbf{D}^{-2\alpha}(\mathbf{y}_* - \mathbf{y}), \quad \tilde{\mathbf{A}} := \mathbf{D}^{-\beta}\mathbf{A}\mathbf{D}^{\beta}.$$

To apply our complexity analysis we could assume now that $\tilde{\mathbf{p}}$ has a certain sparsity which would then naturally be the same as the sparsity of $\mathbf{u}$. But although the matrix $\tilde{\mathbf{A}}$ has still the same spectrum as $\mathbf{A}$ and hence is an $\ell_2$-automorphism for which the gradient iteration would still converge, it is presumably less compressible. In fact, the unsymmetric scaling means that we only have an estimate of the form

$$|(\tilde{\mathbf{A}})_{\lambda,\nu}| \lesssim 2^{\beta(|\nu|-|\lambda|)}2^{-\sigma||\lambda|-|\nu||},$$

where $\sigma$ results from the regularity of the wavelets and determines the original compressibility of $\mathbf{A}$.

So, in summary, in the case of a distributed control the solution complexity is not determined by the sparseness of the control but by that of the remaining variables in (EEn).

*Boundary Control.* The situation is different for the example from Section 2.2.3. Recall that in this case $Y = H^1(\Omega)$, $Q = (H^{1/2}(\Gamma_c))'$ so that $E : (H^{1/2}(\Gamma_c))' \to (H^1(\Omega))'$ is the extension operator defined by (2.16). Choosing $Z = H^s(\Omega)$ for $0 \leq s \leq 1$ as the observation space, $T$ is the canonical injection and

$$\mathbf{T} = \mathbf{I}, \qquad \mathbf{T}_Z = \mathbf{D}^{-s}. \tag{6.20}$$

Choosing bases $\Psi_Q \subset Q = (H^{1/2}(\Gamma_c))'$, $\tilde{\Psi}_Q \subset Q' = H^{1/2}(\Gamma_c)$ and $\Psi_Y \subset Y = H^1(\Omega)$, $\tilde{\Psi}_Y \subset (H^1(\Omega))'$, (2.16) and (3.14) say that $\mathbf{E}$ is given by

$$\mathbf{E} = \langle\gamma\Psi_Y, \tilde{\Psi}_Q\rangle, \quad \mathbf{E}^T = \langle\tilde{\Psi}_Q, \gamma\Psi_Y\rangle. \tag{6.21}$$

Thus, the entries of $\mathbf{E}$ are inner products of traces of wavelets on $\Omega$ with wavelets on the control boundary $\Gamma_c$. Choosing $U = L_2(\Gamma_c)$, the Euler–Lagrange system (EE)

now reads

$$\begin{aligned} \mathbf{A}\mathbf{y} &= \boldsymbol{f} + \mathbf{E}\mathbf{D}^{-1/2}\mathbf{u} \\ \mathbf{A}^T\mathbf{p} &= -\mathbf{D}^{-2s}(\mathbf{y} - \mathbf{y}_*) \\ \omega\mathbf{u} &= \mathbf{D}^{-1/2}\mathbf{E}^T\mathbf{p}. \end{aligned} \tag{6.22}$$

Recall that the sparsity of solutions of the first two systems in (6.22) is exploited by the compressibility of $\mathbf{A}$ up to the limiting index $s^*$ which depends only on the cancellation properties and the regularity of $\Psi_Y$ and not on the particular differential operator. In contrast, as shown in [M], $\mathbf{E}$ is $\tau^*$-efficient only for $\tau^* \geq 1$, i.e., $\mathbf{E}^T$ is not bounded on $\ell_\tau^w$ for $\tau < 1$. In other words, $\mathbf{E}^T$ is at most $s^*$-compressible for $s^* = 1/2$. The reason is that traces of wavelets are in general no longer wavelets so that this factor does not have any cancellation properties that help keeping the entries of $\mathbf{E}$ small. Thus, in this case, even when $\mathbf{p}$ is highly sparse in the sense that $\mathbf{p}$ belongs to $\ell_\tau^w$ for $\tau$ much smaller than 1, the application of $\mathbf{E}^T$ in the third equation of (6.22) reduces that sparsity when computing the control $\mathbf{u}$. However, the scaling $\mathbf{D}_U = \mathbf{D}^{-1/2}$ raises the order of compressibility by Proposition 6.5 to $s^* = 1$. This can be also seen directly because it enhances the decay of the entries of $\mathbf{E}^T$ along each row. Without the attenuation caused by the scaling the latter decay is weak due to the lack of vanishing moments of the traces of domain wavelets.

**7. Concluding Remarks.** We have developed a class of fully adaptive schemes for the solution of optimal control problems with elliptic boundary problems as constraints. The approach is based on a gradient iteration for the corresponding full infinite dimensional variational problem in wavelet coordinates. The numerical realization relies on the adaptive application of the involved operators within stage dependent dynamically updated tolerances. The complexity of such schemes is shown to hinge on the properties of these application routines. Concrete realizations of such schemes are exhibited in several simple cases. This sheds some light on the different inherent complexity properties of distributed versus boundary control problems also in connection with different choices of norms in the objective functional. We refer to [BK] for first numerical experiments with algorithms of the above form with uniform refinements where the influence of different norms is explored. We have not considered here so far the role of the regularization parameter $\omega$ in (2.3), (6.16) or (6.22). Its variation affects all scales simultaneously while the diagonal scalings representing different norms treat high and low frequencies differently. This issue is also addressed in the experiments in [BK]. It turns out that the scheme is robust when $\omega$ tends to zero and turns out the correct solution. This can be seen from the structure of the above scheme $\text{SOLVE}_{\text{DCP}}$ where $\omega$ only enters the update in step (IV) of $\text{RES}_{\text{DCP}}$ but leaves the convergence of the ideal iteration (3.39) unaffected. Corresponding numerical experiments for the above adaptive version will be presented and discussed in a forthcoming paper.

REFERENCES

[B] A. Barinka, Fast Evaluation Tools for Adaptive Wavelet Schemes, PhD. Dissertation, RWTH Aachen, 2003, in preparation.

[BCDU] A. Barinka, T. Barsch, Ph. Charton, A. Cohen, S. Dahlke, W. Dahmen, K. Urban, Adaptive wavelet schemes for elliptic problems — Implementation and numerical experiments, SIAM J. Sci. Comp., 23 (2001), 910–939.

[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.

[BK] C. Burstedde, A. Kunoth, Wavelet methods for linear–quadratic control problems, Manuscript.

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

[Co] A. Cohen, Numerical Analysis of Wavelet Methods, Handbook of Numerical Analysis II, vol. 8, P.G. Ciarlet, J.L. Lions (eds.), Elsevier Science Publishers, 1998.

[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 methods II — Beyond the elliptic case, Found. Computat. Math., 2 (2002), 203–245.

[CDD3] A. Cohen, W. Dahmen, R. DeVore, Adaptive wavelet scheme for nonlinear variational problems, Preprint, July 2002.

[CM] A. Cohen, R. Masson, Wavelet adaptive methods for second order elliptic problems, boundary conditions and domain decomposition, Numer. Math., 86 (2000), 193–238.

[DDU] S. Dahlke, W. Dahmen, K. Urban, Adaptive wavelet methods for saddle point problems — 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, C.I.M.E. Lecture Notes, Springer-Verlag, in print.

[DKS] W. Dahmen, A. Kunoth, R. Schneider, Wavelet least square 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, Composite wavelet bases for operator equations, Math. Comp., 68 (1999), 1533–1567.

[DS2] 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 – stability and moment conditions, SIAM J. Numer. Anal., 37 (1999), 319–325.

[DUV] W. Dahmen, K. Urban, J. Vorloeper, Adaptive wavelet methods — Basic concepts and applications to the Stokes problem, in: Wavelet Analysis, D.-X. Zhou (ed.), World Scientific, New Jersey, 2002.

[K1] A. Kunoth, Wavelet Methods — Boundary Value Problems and Control Problems, Advances in Numerical Mathematics, Teubner, 2001.

[K2] A. Kunoth, Fast iterative solution of saddle point problems in optimal control based on wavelets, Comput. Optim. Appl., 22 (2002), 225–259.

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

[M] M. Mommer, Fictitious domain – Lagrange multiplier approach: Smoothness analysis, in preparation.

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