

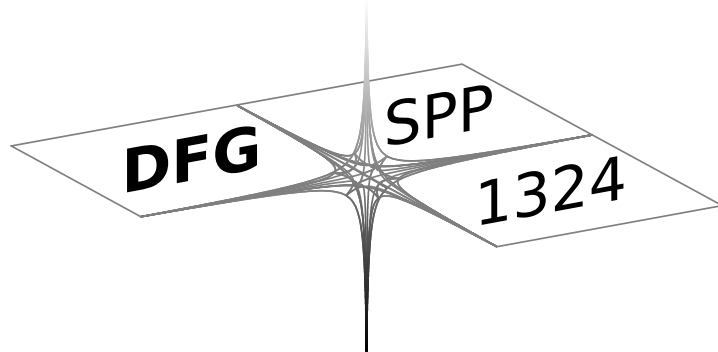
# DFG-Schwerpunktprogramm 1324

„Extraktion quantifizierbarer Information aus komplexen Systemen“

## Adaptive Wavelet Methods and Sparsity Reconstruction for Inverse Heat Conduction Problems

T. Bonesky, S. Dahlke, P. Maass, T. Raasch

Preprint 5



Edited by

AG Numerik/Optimierung  
Fachbereich 12 - Mathematik und Informatik  
Philipps-Universität Marburg  
Hans-Meerwein-Str.  
35032 Marburg

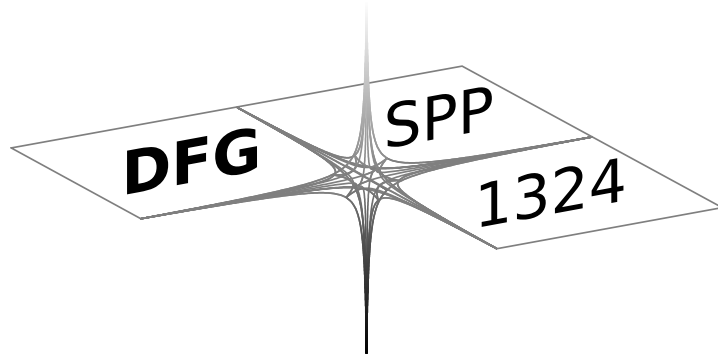
# DFG-Schwerpunktprogramm 1324

„Extraktion quantifizierbarer Information aus komplexen Systemen“

## Adaptive Wavelet Methods and Sparsity Reconstruction for Inverse Heat Conduction Problems

T. Bonesky, S. Dahlke, P. Maass, T. Raasch

Preprint 5



The consecutive numbering of the publications is determined by their chronological order.

The aim of this preprint series is to make new research rapidly available for scientific discussion. Therefore, the responsibility for the contents is solely due to the authors. The publications will be distributed by the authors.

# Adaptive Wavelet Methods and Sparsity Reconstruction for Inverse Heat Conduction Problems

Thomas Bonesky<sup>†</sup>, Stephan Dahlke<sup>§</sup>, Peter Maass<sup>†</sup>, Thorsten Raasch<sup>§</sup>

January 9, 2009

## Abstract

This paper is concerned with the numerical treatment of inverse heat conduction problems. In particular, we combine recent results on the regularization of ill-posed problems by iterated soft shrinkage with adaptive wavelet algorithms for the forward problem. The analysis is applied to an inverse parabolic problem that stems from the industrial process of melting iron ore in a steel furnace. Some numerical experiments that confirm the applicability of our approach are presented.

**AMS Subject classification:** 46N10, 47A52, 49M99, 65F20, 65F50, 65M32, 65M60, 65N12, 65T60

**Keywords:** Regularization of ill-posed problems, sparsity, adaptive numerical schemes, parabolic partial differential equations, iterated soft shrinkage.

## 1 Introduction

Typical applications in medical imaging (computer tomography), geophysics (analysis of seismic waves) or industrial production processes require to determine some critical parameters which cannot be measured directly, so that we are faced with an *inverse problem*. In its most general sense, an inverse problem amounts to solving an operator equation

$$Ax = y \tag{1}$$

with measured noisy data  $y^\delta$ ,  $\|y - y^\delta\| \leq \delta$  where  $A$  denotes a linear or non-linear operator between Hilbert spaces  $X$  and  $Y$ . Typical operator equations, such as parameter identification problems for partial differential equations or integral equations with compact operators are ill-posed in the sense that the operator  $A$  is not boundedly invertible. For an introduction to regularization techniques for such ill-posed problems see [18, 25].

We will investigate an inverse heat conduction problem, see Section 2, which arises from monitoring a steel furnace. The critical structures are localized hot spots on the inner wall of the furnace. Hence, the evaluation of the operator  $A$ , which is called the *forward problem*,

---

<sup>†</sup>Center for Industrial Mathematics / Fachbereich 3, University of Bremen, Postfach 33 04 40, 28334 Bremen, Germany, {tbonesky,pmaass}@math.uni-bremen.de;

Thomas Bonesky was supported by Deutsche Forschungsgemeinschaft, grant number MA 1657/15-1; Peter Maass was supported by Deutsche Forschungsgemeinschaft, SPP 1180, grant number MA 1657/17-1

<sup>§</sup>Philipps-Universität Marburg, Fachbereich 12, Hans-Meerwein-Str., 35032 Marburg, Germany, {dahlke,raasch}@mathematik.uni-marburg.de;

Stephan Dahlke was supported by Deutsche Forschungsgemeinschaft, grant number DA 360/12-1

Thorsten Raasch was supported by Deutsche Forschungsgemeinschaft, grant number DA 360/7-1

requires to solve a parabolic partial differential equation. In this paper we combine an adaptive forward solver for the underlying PDE with an iterated soft shrinkage approach for obtaining a sparse approximation of the searched for quantity.

Sparsity concepts are successful if the generalized solution of (1) has a sparse expansion with respect to a given basis, i.e. if the solution can be approximated well by a small selection of basis elements. In our present application, due to the particular shape of the hot spots, the generalized solution is a smooth function with isolated singularities in space-time. Therefore, the sparsity assumption is satisfied with respect to wavelet bases. The mathematical formulation leads to minimizing Tikhonov functionals with  $\ell_p$ -penalty term,  $1 \leq p < 2$ ,

$$J(x) = \|Ax - y^\delta\|^2 + \alpha \|x\|_{\mathbf{w},p}^p,$$

see Section 3 for details. The regularization properties of such an approach for linear operator equations as well as numerical schemes for minimizing such Tikhonov functionals by iterated soft shrinkage methods have been introduced in the pioneering paper [16].

This concept has been generalized in various directions, e.g. to non-linear operators or general Banach spaces. Moreover, a regularization theory for sparsity constraints in connection with inexact, adaptive operator evaluations has been recently developed, see [1, 30]. These papers give theoretical results on how to choose the accuracy of the adaptive forward solver in order to obtain optimal convergence rates. The numerical results of these papers are restricted to some rather straight forward academic problems and the adaptive solver is treated as a black box. In this paper, we exemplify the use of an adaptive wavelet solver for an inverse heat conduction problem. Hence, we will assume, that the operator  $A$  will be evaluated only approximately. The adaptive forward solver  $[Ax]_\varepsilon$  will obey a pointwise error estimate

$$\|[Ax]_\varepsilon - Ax\| \leq \varepsilon.$$

Let us emphasize, that we do not assume an error bound for the operator norm  $\|[A]_\varepsilon - A\|$  or similar.

Adaptive schemes are designed for the constructive approximation of objects that are implicitly given by a linear or non-linear operator equation. Essentially, an adaptive scheme is an updating strategy. Based on an a posteriori error estimator, an underlying grid or an underlying function space is only refined in regions where the local error is quite large. By now, the most powerful results concerning adaptive schemes have been derived for boundedly invertible operator equations, see, e.g., [8, 9]. Unfortunately, as already outlined above, this assumption is usually violated for inverse problems. Nevertheless, the realization of a regularization scheme requires the solution of the forward problem and its adjoint, and at least for these issues, adaptive strategies may be employed. Adaptive wavelet schemes for the numerical treatment of the forward problem are advantageous since the sparsity constraint for the inverse problem is defined in a wavelet basis as well.

This paper is organized as follows. In Section 2, we present a detailed description of our model problem. Then, in Section 3, we discuss the iterated soft shrinkage regularization scheme including the case of adaptive operator evaluations and we briefly recall the setting of adaptive wavelet methods as far as it is needed for our purposes. In particular, we discuss adaptive schemes for elliptic and parabolic partial differential equations. Section 4 is devoted to the numerical treatment of our model problem. Several reconstructions from noisy data are presented for a representative subset of operating parameters. We finish with some concluding remarks in Section 5.

## 2 An Inverse Heat Conduction Problem

The inverse heat conduction problem (IHCP) investigated in this paper is motivated by an application from monitoring steel production. The life span of a steel furnace is determined by some critical thickness of its outer wall, which decreases continuously due to the physical and chemical processes of the melting iron ore.

Therefore the inner shape of the wall is of specific interest. However, it is impossible to make measurements inside the furnace. Hence, this information has to be determined via indirect measurements on the outside of the furnace wall.

Mathematically we treat this problem as an IHCP on a ring-shaped domain, which we assume to lay entirely inside the existing wall of the furnace. Hence, we can assume a homogeneous conductivity on this domain. Hot spots on the inner boundary indicate some damage during the life span of the furnace. The temperature  $g$  on the inner boundary is the searched for quantity of this inverse problem.

The problem is modelled in the following way. We assume a ring-shaped domain  $\Omega = \{x \in \mathbb{R}^2 \mid 0 < r_0 < \|x\| < r_1\}$ , with inner boundary  $\Gamma_0 = \{x \in \mathbb{R}^2 \mid \|x\| = r_0\}$  and outer boundary  $\Gamma_1 = \{x \in \mathbb{R}^2 \mid \|x\| = r_1\}$ , see Figure 1. Further we formulate the forward problem as:

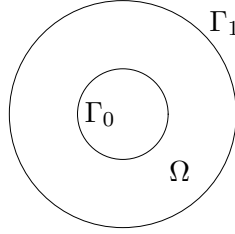


Figure 1: The ring-shaped domain  $\Omega$ .

$$\begin{aligned}
 \frac{\partial u}{\partial t} - \kappa \Delta u &= 0 && \text{in } (0, T) \times \Omega \\
 u &= g && \text{on } (0, T) \times \Gamma_0 \\
 -\kappa \frac{\partial u}{\partial \nu} &= h && \text{on } (0, T) \times \Gamma_1 \\
 u(0, \cdot) &= u_0 && \text{in } \Omega.
 \end{aligned} \tag{2}$$

We assume that the heat conductivity  $\kappa$  of the furnace wall is known, as well as the initial temperature  $u_0$ . We expect that  $\kappa$  is constant and for simplicity we set  $\kappa = 1$ . In addition to this, we assume that we can measure the heat flux  $h$  through  $\Gamma_1$  and the temperature  $\tilde{h}$  on  $\Gamma_1$  during the whole time interval  $(0, T)$ . The latter will constitute the additional input information for the formulation of the inverse problem. Problem (2) can now be reformulated as an operator equation

$$L(g; h, u_0) = u, \tag{3}$$

with a non-linear operator  $L$  depending on some fixed  $h$  and  $u_0$  assigning a function  $g$  on the inner boundary  $\Gamma_0$  to a solution  $u$  of (2). This operator is affine-linear in  $g$ . We need to subtract the inhomogeneous part  $u_c := L(0; h, u_0)$  in order to obtain a linear operator equation as follows.

We split the operator in (3) into a linear part generating a solution  $u_l$  and a constant part represented by a solution  $u_c$

$$\begin{aligned}
 L(g; h, u_0) &= L(g; 0, 0) + L(0; h, u_0) \\
 &= u_l + u_c.
 \end{aligned}$$

Here  $u_l$  indicates the solution of

$$\begin{aligned} \frac{\partial u}{\partial t} - \Delta u &= 0 & \text{in } (0, T) \times \Omega \\ u &= g & \text{on } (0, T) \times \Gamma_0 \\ -\frac{\partial u}{\partial \nu} &= 0 & \text{on } (0, T) \times \Gamma_1 \\ u(0, \cdot) &= 0 & \text{in } \Omega. \end{aligned} \tag{4}$$

Defining the IHCP we have to restrict the solution  $u = u_l + u_c$  of (2) to the outer boundary  $\Gamma_1$ . Finally, we obtain

$$L(g; h, u_0)|_{(0, T) \times \Gamma_1} = u_l|_{(0, T) \times \Gamma_1} + u_c|_{(0, T) \times \Gamma_1} = \tilde{h},$$

which allows us to define

$$Kg = \tilde{h} - u_c|_{(0, T) \times \Gamma_1} =: \hat{h},$$

where  $K$  denotes the linear solution operator of (4) followed by the restriction to  $(0, T) \times \Gamma_1$ .

Note, that our additional temperature measurements  $\tilde{h}$  on the outer wall correspond to  $Kg$  after subtracting the restriction of  $L(0; h, u_0)$  on the boundary. The operator equation  $Kg = \hat{h}$  can now be handled with linear regularization theory.

## 2.1 Mapping Properties of the Forward Operator

As a basis for the following arguments, we shall clarify first the mapping properties of the linear forward operator  $K : g \mapsto L(g; 0, 0)|_{(0, T) \times \Gamma_1}$ . In particular, we have to address the question between which function spaces the boundedness of  $K$  may be expected to hold. Since the Dirichlet data  $g$  under consideration are potentially corrupted by measurement noise, the most attractive domain of definition for  $K$  would be  $L_2(0, T; L_2(\Gamma_0))$ . However, it turns out that the boundedness of  $K$  on that space can not be verified by a standard approach as soon as the spatial dimension is greater or equal to 2 and  $\Omega$  is nontrivially shaped. This is related with the fact that a variational formulation of the problem (4) will involve extensions of Dirichlet boundary data on  $\Gamma_0$  into the interior of  $\Omega$  which are only bounded on  $H^s(\Gamma_0)$  for  $s > 0$ .

**Remark 2.1** *If the spatial domain  $\Omega$  was not the ring-shaped domain but one-dimensional, i.e., in the case of the sideways heat equation*

$$\begin{aligned} u_t - u_{xx} &= 0 & \text{in } (0, T) \times (0, 1) \\ u &= g & \text{on } (0, T) \times \{0\} \\ u_x &= 0 & \text{on } (0, T) \times \{1\} \\ u(0, \cdot) &= 0 & \text{in } (0, 1), \end{aligned}$$

*it is well-known that the forward operator  $K : g \mapsto u(\cdot, \cdot; g)|_{x=1}$  would be bounded from  $L_2(0, T)$  to  $L_2(0, T)$  as a convolution operator with smooth kernel [17, 32].*

It is therefore desirable to verify the boundedness of  $K$  between spaces that are only slightly smaller than  $L_2(0, T; L_2(\Gamma_i))$ . In the sequel, we will confine the discussion to the proof that  $K$  is bounded as an operator between spaces of weakly differentiable functions. Following [24, Ch. 4], for an open set or a closed manifold  $S \subset \mathbb{R}^n$  and  $s \geq 0$ , let  $L_2(0, T; H^s(S))$  be the space of all measurable mappings  $w : (0, T) \rightarrow H^s(S)$  with finite norm

$$\|w\|_{L_2(0, T; H^s(S))} := \left( \int_0^T \|w(t, \cdot)\|_{H^s(S)}^2 dt \right)^{1/2}.$$



Moreover, for  $r \in \mathbb{N}_0$ , let  $H^r(0, T; L_2(S))$  be the space of all mappings  $w \in L_2(0, T; L_2(S))$  whose  $r$ -th distributional derivative  $\frac{\partial^r w}{\partial t^r}$  is also contained in  $L_2(0, T; L_2(S))$ , with norm

$$\|w\|_{H^r(0, T; L_2(S))} := \|w\|_{L_2(0, T; L_2(S))} + \left\| \frac{\partial^r w}{\partial t^r} \right\|_{L_2(0, T; L_2(S))}.$$

For noninteger  $r \geq 0$ , the space  $H^r(0, T; L_2(S))$  shall be defined by interpolation. Given  $r, s \geq 0$ , let us finally set

$$H^{r,s}((0, T) \times S) := H^r(0, T; L_2(S)) \cap L_2(0, T; H^s(S))$$

with associated norm  $\|w\|_{H^{r,s}((0, T) \times S)} := \|w\|_{H^r(0, T; L_2(S))} + \|w\|_{L_2(0, T; H^s(S))}$ .

The following trace theorem for functions from  $H^{r,s}((0, T) \times \Omega)$  can be found in [24, Chapter 4, Section 2]:

**Theorem 2.2** *Assume that  $\Omega \subset \mathbb{R}^n$  is an open, bounded set with smooth boundary  $\partial\Omega$  and let  $Q := (0, T) \times \Omega$ . For a given boundary part  $\Gamma \subset \partial\Omega$ , let  $\Sigma := (0, T) \times \Gamma$ .*

(i) *For  $r \geq 0$ ,  $s > \frac{1}{2}$  and any  $k \in \mathbb{N}_0$  with  $k < s - \frac{1}{2}$ , there exists a bounded surjective trace operator*

$$u \mapsto \frac{\partial^k u}{\partial \nu^k} \Big|_{\Sigma} : H^{r,s}(Q) \rightarrow H^{r-(k+\frac{1}{2})\frac{r}{s}, s-(k+\frac{1}{2})}(\Sigma). \quad (5)$$

(ii) *For  $r > \frac{1}{2}$ ,  $s \geq 0$  and any  $j \in \mathbb{N}_0$  with  $j < r - \frac{1}{2}$ , there exists a bounded surjective trace operator*

$$u \mapsto \frac{\partial^j u}{\partial t^j} \Big|_{t=0} : H^{r,s}(Q) \rightarrow H^{s-(j+\frac{1}{2})\frac{s}{r}}(\Omega). \quad (6)$$

(iii) *If  $r, s > 0$  are such that  $\frac{1}{r} + \frac{1}{s} < 2$ , then the traces of  $u \in H^{r,s}(Q)$  fulfill the compatibility conditions*

$$\frac{\partial^j}{\partial t^j} \left( \frac{\partial^k u}{\partial \nu^k} \Big|_{\Sigma} \right) \Big|_{t=0} = \gamma_k \left( \frac{\partial^j u}{\partial t^j} \Big|_{t=0} \right), \quad \text{for all } j, k \in \mathbb{N}_0 \text{ with } \frac{j}{r} + \frac{k}{s} < 1 - \frac{1}{2} \left( \frac{1}{r} + \frac{1}{s} \right), \quad (7)$$

where  $\gamma_k : H^t(\Omega) \rightarrow H^{t-k-1/2}(\Sigma)$  is the  $k$ -th normal derivative operator,  $k < t - 1/2$ .

As an immediate consequence of Theorem 2.2(i), for all  $r \geq 0$  and  $s > \frac{1}{2}$ , we can derive the existence of a bounded Dirichlet trace lifting

$$E_{\Gamma_0} : H^{r(1-1/(2s)), s-1/2}((0, T) \times \Gamma_0) \rightarrow H^{r,s}(Q),$$

with  $(E_{\Gamma_0} g)|_{(0, T) \times \Gamma_0} = g$  for all  $g \in H^{r(1-1/(2s)), s-1/2}((0, T) \times \Gamma_0)$ . What is more, in the case of a ring-shaped domain  $\Omega$  and  $s > \frac{3}{2}$ , it can be assumed that the Neumann trace  $\frac{\partial}{\partial \nu} (E_{\Gamma_0} g)|_{(0, T) \times \Gamma_1}$  completely vanishes. This property can be assured, e.g., by smoothly damping an extension of  $g$  in the vicinity of  $\Gamma_1$ , since  $\Gamma_0$  and  $\Gamma_1$  are separated sets.

For the analysis of the forward operator  $K$ , the special case  $r = 1$  and  $s = 2$  will be of particular importance. The following theorem clarifies the mapping properties of  $K$  in that situation.

**Theorem 2.3** *Assume that  $\Omega$  is the ring-shaped domain and let  $Q := (0, T) \times \Omega$ . Moreover, let  $V_i := H^{3/4, 3/2}((0, T) \times \Gamma_i)$ ,  $i \in \{0, 1\}$ , and let  $\tilde{V}_0 := \{g \in V_0 : g(0, \cdot) = 0\}$ . Then for each  $g \in \tilde{V}_0$ , (4) has a unique solution  $u = u(\cdot, \cdot; g) \in H^{1,2}(Q)$  which linearly and continuously depends on  $g$ ,*

$$\|u\|_{H^{1,2}(Q)} \leq C \|g\|_{V_0}, \quad (8)$$

with a universal constant  $C > 0$ . Moreover, the operator  $K$ , mapping Dirichlet data on the inner boundary  $\Gamma_0$  to Dirichlet data on the outer boundary  $\Gamma_1$ , defines a bounded operator

$$K : \tilde{V}_0 \rightarrow V_1.$$

In order to prove Theorem 2.3, we shall apply a well-known technique to transform the inhomogeneous Dirichlet data into a nontrivial right-hand side of the differential equation. For parabolic initial-boundary value problems with homogeneous boundary data, existence, uniqueness and stability of solutions are guaranteed by the following theorem, see [21, Ch. IV, §9] for details.

**Theorem 2.4** *Assume that  $\Omega \subset \mathbb{R}^n$  is an open, bounded set with smooth boundary  $\partial\Omega = \bar{\Gamma}_0 \cup \bar{\Gamma}_1$  and let  $Q := (0, T) \times \Omega$ . Moreover, let  $f \in L_2(0, T; L_2(\Omega))$  and  $u_0 \in H^1(\Omega)$  with  $u_0|_{\Gamma_0} = 0$ . Then the boundary value problem*

$$\begin{aligned} \frac{\partial \tilde{u}}{\partial t} - \Delta \tilde{u} &= f && \text{in } (0, T) \times \Omega \\ \tilde{u} &= 0 && \text{on } (0, T) \times \Gamma_0 \\ -\frac{\partial \tilde{u}}{\partial \nu} &= 0 && \text{on } (0, T) \times \Gamma_1 \\ \tilde{u}(0, \cdot) &= u_0 && \text{in } \Omega \end{aligned} \tag{9}$$

has a unique solution  $\tilde{u} \in H^{1,2}(Q)$  which depends affine linearly and continuously on the driving term and on the initial data,

$$\|\tilde{u}\|_{H^{1,2}(Q)} \leq C(\|f\|_{L_2(0,T;L_2(\Omega))} + \|u_0\|_{H^1(\Omega)}). \tag{10}$$

With Theorem 2.4 at hand, we are in the position to prove Theorem 2.3.

*Proof of Theorem 2.3.* We start by showing uniqueness and existence of a solution  $u \in H^{1,2}(Q)$  to (4) for given Dirichlet data  $g \in \tilde{V}_0$ .

Concerning uniqueness, assume that  $u, v \in H^{1,2}(Q)$  solve (4). Then  $u - v$  solves (9) with  $f = 0$  and  $u_0 = 0$ , so that Theorem 2.4 implies  $u - v = 0$ .

For the existence of solutions to (4), consider a trace lifting  $G := E_{\Gamma_0}g \in H^{1,2}(Q)$  of  $g$ , with continuous dependence  $\|G\|_{H^{1,2}(Q)} \leq C\|g\|_{V_0}$ . Therefore, it holds that  $f := \frac{\partial}{\partial t}G - \Delta G \in L_2(0, T; L_2(\Omega))$  with  $\|f\|_{L_2(0,T;L_2(\Omega))} \leq C'\|g\|_{V_0}$ , and Theorem 2.2(ii) yields that  $u_0 := -G(0, \cdot) \in H^1(\Omega)$  with  $\|u_0\|_{H^1(\Omega)} \leq C''\|g\|_{V_0}$ . Moreover, since  $r = 1$  and  $s = 2$  imply  $\frac{1}{r} + \frac{1}{s} = \frac{3}{2} < 2$ , an application of Theorem 2.2(iii) for  $j = k = 0$  yields  $u_0|_{\Gamma_0} = 0$ . Therefore, by Theorem 2.4, (9) has a unique solution  $\tilde{u} \in H^{1,2}(Q)$  with linear and continuous dependence on  $g$ ,  $\|\tilde{u}\|_{H^{1,2}(Q)} \leq C'''\|g\|_{V_0}$ , so that  $u := \tilde{u} + G \in H^{1,2}(Q)$  solves (4) with  $\|u\|_{H^{1,2}(Q)} \leq C''''\|g\|_{V_0}$ .

Since  $Kg = u|_{(0,T) \times \Gamma_1}$ , the boundedness of  $K$  now follows from the boundedness of the trace operator  $H^{1,2}(Q) \rightarrow V_1$ , see Theorem 2.2(i).  $\blacksquare$

**Remark 2.5** *By the density of  $D(K) := \tilde{V}_0$  in  $L_2(0, T; L_2(\Gamma_0))$ , the forward operator  $K$  can also be regarded as a linear, densely defined operator  $K : D(K) \subset L_2(0, T; L_2(\Gamma_0)) \rightarrow L_2(0, T; L_2(\Gamma_1))$ . However, it is not clear whether  $K$  has a bounded extension to the full space  $L_2(0, T; L_2(\Gamma_0))$ .*

## 2.2 Mapping Properties of the Adjoint Operator

To use a classical Tikhonov regularization or, as we will do in the following, a sparsity reconstructing shrinkage method, the adjoint operator  $K^*$  is needed. In general, for Hilbert spaces  $X$  and  $Y$ , the adjoint operator of a densely defined operator  $K : X \supset D(K) \rightarrow Y$  is given via the relation

$$\langle x, K^*y \rangle_X = \langle Kx, y \rangle_Y, \quad \text{for all } x \in D(K), y \in D(K^*), \tag{11}$$

where  $D(K^*) := \{y \in Y : x \mapsto \langle Kx, y \rangle_Y \text{ is continuous on } D(K)\}$ .

In order to concretely identify the mapping properties of  $K^*$ , consider the following *adjoint* boundary value problem

$$\begin{aligned} \frac{\partial v}{\partial t} + \Delta v &= 0 & \text{in } (0, T) \times \Omega \\ v &= 0 & \text{on } (0, T) \times \Gamma_0 \\ -\frac{\partial v}{\partial \nu} &= h & \text{on } (0, T) \times \Gamma_1 \\ v(T, \cdot) &= 0 & \text{in } \Omega \end{aligned} \tag{12}$$

and its solution operator  $M : h \mapsto \frac{\partial v}{\partial \nu}|_{(0, T) \times \Gamma_0}$ . In the sequel, we shall clarify the mapping properties of  $M$  and its relation to  $K$ .

Note first that by an application of Theorem 2.2(i) for  $k = 1$ , for all  $r \geq 0$  and  $s > \frac{3}{2}$ , we can deduce the existence of a bounded Neumann trace lifting

$$F_{\Gamma_1} : H^{r(1-3/(2s)), s-3/2}((0, T) \times \Gamma_1) \rightarrow H^{r, s}(Q),$$

with  $(\frac{\partial}{\partial \nu} F_{\Gamma_1} h)|_{(0, T) \times \Gamma_1} = h$  for all  $h \in H^{r(1-3/(2s)), s-3/2}((0, T) \times \Gamma_1)$ . Moreover, similar to the properties of  $E_{\Gamma_0}$ , we can assume that the Dirichlet data  $(F_{\Gamma_1} h)|_{(0, T) \times \Gamma_0}$  completely vanish.

Again we shall be interested in the special case  $r = 1$  and  $s = 2$ , and the following mapping properties of  $M$  can be shown in analogy to Theorem 2.3:

**Theorem 2.6** *Assume that  $\Omega$  is the ring-shaped domain and let  $Q := (0, T) \times \Omega$ . Moreover, let  $W_i := H^{1/4, 1/2}((0, T) \times \Gamma_i)$ ,  $i \in \{0, 1\}$ . Then for each  $h \in W_1$ , (12) has a unique solution  $v = v(\cdot, \cdot; h) \in H^{1, 2}(Q)$  which linearly and continuously depends on  $h$ ,*

$$\|v\|_{H^{1, 2}(Q)} \leq C \|h\|_{W_1}, \tag{13}$$

with a universal constant  $C > 0$ . Moreover, the operator  $M$ , mapping Neumann data on the outer boundary  $\Gamma_1$  to Neumann data on the inner boundary  $\Gamma_0$ , defines a bounded operator

$$M : W_1 \rightarrow W_0.$$

*Proof.* The uniqueness of solutions  $v \in H^{1, 2}(Q)$  to (12) for given Neumann data  $h \in W_1$  can be shown as in Theorem 2.3. For the existence of a solution  $v$ , let  $H := F_{\Gamma_1} h(T - \cdot) \in H^{1, 2}(Q)$  be a Neumann trace lifting of  $h(T - \cdot)$ , with continuous dependence  $\|H\|_{H^{1, 2}(Q)} \leq C \|h\|_{W_1}$ . Therefore, it holds that  $f := \frac{\partial}{\partial t} H - \Delta H \in L_2(0, T; L_2(\Omega))$  with  $\|f\|_{L_2(0, T; L_2(\Omega))} \leq C' \|h\|_{W_1}$ , and an application of Theorem 2.2(ii) with  $j = 0$  yields that  $v_0 := -H(0, \cdot) \in H^1(\Omega)$  and  $\|v_0\|_{H^1(\Omega)} \leq C'' \|h\|_{W_1}$ . Moreover, by the mapping properties of  $F_{\Gamma_1}$ , we have  $v_0|_{\Gamma_0} = 0$ . Theorem 2.4 then guarantees that the auxiliary problem

$$\begin{aligned} \frac{\partial \tilde{v}}{\partial t} - \Delta \tilde{v} &= f & \text{in } (0, T) \times \Omega \\ \tilde{v} &= 0 & \text{on } (0, T) \times \Gamma_0 \\ -\frac{\partial \tilde{v}}{\partial \nu} &= 0 & \text{on } (0, T) \times \Gamma_1 \\ \tilde{v}(0, \cdot) &= v_0 & \text{in } \Omega \end{aligned}$$

has a unique solution  $\tilde{v} \in H^{1, 2}(Q)$ , with linear and continuous dependence on  $h$ ,  $\|\tilde{v}\|_{H^{1, 2}(Q)} \leq C''' \|h\|_{W_1}$ , so that  $v := (\tilde{v} + H)(T - \cdot) \in H^{1, 2}(Q)$  solves (12) with  $\|v\|_{H^{1, 2}(Q)} \leq C'''' \|h\|_{W_1}$ .

Since  $Mh = \frac{\partial v}{\partial \nu}|_{(0, T) \times \Gamma_0}$ , the boundedness of  $M$  follows from the boundedness of the Neumann trace operator  $H^{1, 2}(\Omega) \rightarrow W_0$ , see Theorem 2.2(ii).  $\blacksquare$

Therefore, see Remark 2.5, also  $M$  may be regarded as a densely defined, potentially unbounded operator from  $D(M) := W_1 \hookrightarrow L_2(0, T; L_2(\Gamma_1))$  to  $L_2(0, T; L_2(\Gamma_0))$ . We shall now verify that  $K$  in fact coincides with  $M^*$  on  $\tilde{V}_0 = D(K) \subset D(M^*)$ , readily justifying why (12) may be called the adjoint problem.

**Theorem 2.7** *It holds that  $D(K) \subset D(M^*)$  and  $M^*|_{D(K)} = K$ .*

*Proof.* Abbreviate  $X := L_2(0, T; L_2(\Gamma_0))$  and  $Y := L_2(0, T; L_2(\Gamma_1))$ . We will prove that

$$\langle Kg, h \rangle_Y = \langle g, Mh \rangle_X, \quad \text{for all } g \in D(K) \subset X, h \in D(M) \subset Y, \quad (14)$$

from which it immediately follows that  $h \mapsto \langle g, Mh \rangle_X = \langle Kg, h \rangle_Y$  is continuous on  $D(M)$  in the topology of  $Y$ , i.e.,  $D(K) \subset D(M^*)$ . Moreover, as claimed, (14) implies that  $M^*$  coincides with  $K$  on  $D(K)$ .

For given Dirichlet and Neumann data  $g \in \tilde{V}_0$  and  $h \in W_1$ , Theorems 2.3 and 2.6 imply that (4) and (12) have solutions  $u, v \in H^{1,2}(Q)$ , respectively. By using  $u(0, \cdot) = v(T, \cdot) = 0$  and the weak temporal differentiability of  $u$  and  $v$ , we compute that

$$\begin{aligned} 0 &= \int_{\Omega} (u(T, x)v(T, x) - u(0, x)v(0, x)) \, dx \\ &= \int_{\Omega} \int_0^T \frac{\partial}{\partial t} (u(t, x)v(t, x)) \, dt \, dx \\ &= \int_0^T \int_{\Omega} \left( \frac{\partial u}{\partial t}(t, x)v(t, x) + \frac{\partial v}{\partial t}(t, x)u(t, x) \right) \, dx \, dt, \end{aligned}$$

where in the last step, we are allowed to change the order of integration since the integrand is contained in  $L_1(Q)$  due to  $u, v \in H^{1,2}(Q)$ . By (4), (12) and an application of Green's formula in space, see [24, Ch. 2], it follows that

$$\begin{aligned} 0 &= \int_0^T \int_{\Omega} (v(t, x)\Delta u(t, x) - u(t, x)\Delta v(t, x)) \, dx \, dt \\ &= \int_0^T \int_{\partial\Omega} (v(t, x)\frac{\partial u}{\partial \nu}(t, x) - u(t, x)\frac{\partial v}{\partial \nu}(t, x)) \, d\Gamma \, dt, \end{aligned}$$

so that an insertion of the boundary data finally yields (14),

$$\begin{aligned} 0 &= - \int_0^T \int_{\Gamma_0} u(t, x)\frac{\partial v}{\partial \nu}(t, x) \, dx \, dt - \int_0^T \int_{\Gamma_1} u(t, x)\frac{\partial v}{\partial \nu}(t, x) \, dx \, dt \\ &= -\langle g, Mh \rangle_X + \langle Kg, h \rangle_Y. \quad \blacksquare \end{aligned}$$

**Remark 2.8** *The Sobolev spaces introduced above are not preferable for treating the inverse problem numerically. First of all, we need to tackle measurement noise. By using the classical deterministic noise model we replace the image space  $V_1$  by  $L_2(0, T; L_2(\Gamma_1))$ . Moreover, we quote [2, p. R120], which addresses a parameter identification problem for an elliptic differential equation: ‘Working with the  $L_2(\partial\Omega)$ , norm instead of  $H^{1/2}$ , is motivated by the high cost of calculating fractional Sobolev space norms, as well as the possible lack of differentiability of the operator norm.’. Hence, we will use  $L_2$ -norms for our numerical applications.*

These investigations show, that it will be necessary to have stable and efficient strategies for solving the parabolic problems (4) and (12). For reasons that will be explained below, our method of choice will be an adaptive wavelet scheme for solving these PDEs.

### 3 Regularization and Adaptive Forward Solver

The considered IHCP described in Section 2 is an inverse and ill-posed parameter identification problem. In terms of inverse problems, this can be formulated as a linear operator equation

$$K\mathbf{x} = y,$$

where evaluating the operator  $K : \ell_2 \rightarrow Y$  requires to solve a parabolic partial differential equation followed by a restriction to the outer boundary. Our investigation is motivated by the need to monitor steel furnaces. High temperatures in certain regions of the furnace wall indicate critical operating conditions. In particular we search for hot spots at a circle inside the furnace wall. Those hot spots will appear locally at this inner circle. The shape of the hot spot will change depending on whether it is created from a crack or a larger damaged part on the inner surface of the furnace. Hence, it is feasible to assume that the searched for temperature distribution has a sparse representation in a basis, which is well suited for capturing the overall, smooth temperature distribution as well as some very localized structures with very few coefficients. This leads to the choice of a wavelet basis.

We will apply Tikhonov regularization with a wavelet sparsity constraint. I.e. we expand the searched for quantity  $g = \sum x_\lambda \psi_\lambda$  in a wavelet basis and determine the coefficient sequence  $\mathbf{x} \in \ell_2$  by minimizing the functional

$$\Gamma_{\alpha w, p, \delta}(\mathbf{x}) = \|K\mathbf{x} - y^\delta\|^2 + \alpha \|\mathbf{x}\|_{\mathbf{w}, p}^p, \quad (15)$$

with

$$\|\mathbf{x}\|_{\mathbf{w}, p}^p = \sum_{\lambda} w_{\lambda} |x_{\lambda}|^p. \quad (16)$$

The operator  $K$  is given by an inverse wavelet transform followed by  $A$ .

We will now summarize the basic construction of adaptive wavelet solvers for evaluating the forward operator  $K$  and the convergence properties of iterated soft shrinkage methods with adaptive operator evaluation.

#### 3.1 Adaptive Iterated Soft Shrinkage

A minimizer of (15) can be obtained by iterated soft shrinkage

$$\begin{aligned} \mathbf{x}^0 & \text{arbitrary} \\ \mathbf{x}^{n+1} &= \mathbf{S}_{\alpha w, p}(\mathbf{x}^n - K^*(K\mathbf{x}^n - y^\delta)) \\ &= \{S_{\alpha w_{\lambda}, p}((\mathbf{x}^n - K^*(K\mathbf{x}^n - y^\delta))_{\lambda})\}_{\lambda}, \end{aligned} \quad (17)$$

with the shrinkage functions

$$S_{\alpha w_{\lambda}, p}(x) = \begin{cases} \operatorname{sgn}(x) [|x| - \frac{\alpha w_{\lambda}}{2}]_+ & p = 1 \\ G_{\alpha w_{\lambda}, p}^{-1}(x) & 1 < p \leq 2 \end{cases} \quad (18)$$

and

$$G_{\alpha w_{\lambda}, p}(x) = x + \frac{\alpha w_{\lambda} p}{2} \operatorname{sgn}(x) |x|^{p-1}.$$

The convergence of iterated soft shrinkage methods for  $1 \leq p \leq 2$  and linear operators  $K$  to a minimizer of (15) was proved in [16]. The convergence can be shown under the following assumptions:

- $\Gamma_{\alpha\mathbf{w},p,\delta}$  as defined in (15) has a unique minimizer,
- the exponent satisfies  $1 \leq p \leq 2$ ,
- the weights satisfy  $0 < w \leq w_\lambda$  for all  $\lambda$ .

In addition regularization properties of the minimizer were proved [16, Theorem 4.1]. Subsequently this basic theory was extended to non-linear operators, general Banach spaces etc., see e.g. [3–5, 11].

For complex applications, such as parameter identification problems for partial differential equations, we need to design efficient algorithms for the evaluation of the operator  $K$ . Hence, we will utilize adaptive operator evaluations  $[K\mathbf{x}]_\varepsilon$ , where the theoretical results are restricted to the choice  $1 < p \leq 2$ . We only present the basic facts and refer the interested reader to [1] for details.

First of all, we want to emphasize again that an adaptive scheme does not use a fixed operator approximation. The pointwise approximation of the solution operator may vary for every evaluation. An adaptive scheme approximates the exact evaluation  $K\mathbf{x}$  within a tolerance of  $\varepsilon$ :

$$\|[K\mathbf{x}]_\varepsilon - K\mathbf{x}\| \leq \varepsilon.$$

The same assumption we make for the adjoint operator  $K^*$ . We assume

$$\|[K^*z]_\varepsilon - K^*z\| \leq \varepsilon.$$

Iteration (17) reformulated for adaptive operator evaluations then looks like

$$\begin{aligned} \mathbf{x}^0 & \quad \text{arbitrary} \\ \mathbf{x}^{n+1} & = \mathbf{S}_{\alpha\mathbf{w},p}(\mathbf{x}^n - [K^*([K\mathbf{x}^n]_\varepsilon - y^\delta)]_\varepsilon). \end{aligned} \quad (19)$$

If we denote the minimizer of the exact iteration (17) by  $\mathbf{x}_{\alpha,p}^\delta$ , then we can define the distance between this minimizer and the iterates generated by iteration (19) as

$$d^n = \|\mathbf{x}^n - \mathbf{x}_{\alpha,p}^\delta\|.$$

To show a decrease of these distances, we use the contraction property of the operator  $\mathbf{S}_{\alpha\mathbf{w},p}$ . The basic estimate in [1] states that the iterates of the adaptive soft shrinkage method stay within a ball of radius

$$R = \|\mathbf{x}^0\| + 2\left[\frac{1}{w}\left(\frac{\delta^2}{\alpha} + \|\mathbf{x}^\dagger\|_{\mathbf{w},p}^p\right)\right]^{1/p} + \|y^\delta\| + 3\delta,$$

which implies

$$d^n \leq \frac{3\varepsilon}{1+C} + \frac{1}{1+C}d^{n-1},$$

with

$$C = C(\alpha, \mathbf{w}, p, R) = \frac{\alpha w p (p-1)}{2} R^{p-2}.$$

**Remark 3.1** *The constant  $R$  contains the norm of the unknown solution  $\mathbf{x}^\dagger$ . To be able to calculate concrete values for  $C$ , it is necessary to estimate the constant  $R$  using only quantities, which are actually available. There are at least two possibilities. Firstly, there is a coarser estimate for  $R$ , which contains the norm of  $y^\delta$  instead of the norm of  $\mathbf{x}^\dagger$ , see [1]. The second possibility is to use some a priori knowledge on  $\mathbf{x}^\dagger$ . For instance assume  $\|\mathbf{x}^\dagger\|_{\mathbf{w},p} < \rho$ , with some finite  $\rho$ .*

The decrease of the distances  $d^n$  ensures that the iterates are located in a ball with radius  $r(\varepsilon) = \frac{6}{C}\varepsilon$  around  $x_\alpha^\delta$  after a finite number of iterations. As proved in [1, Proposition 3.12], this is the case after at most

$$N = \left\lceil \frac{\log\left(\frac{r(\varepsilon)}{d^0}\right)}{\log\left(\frac{C+2}{2(C+1)}\right)} \right\rceil$$

iterations.

So far we only considered the case of fixed parameters  $\{\alpha, \delta\}$ . To obtain a regularization result for the adaptive case, we have to investigate the behavior of the iterates, in case that  $\delta$  tends to zero. For the proof of a regularization result it is crucial that the values of  $R$  are bounded for  $\delta$  tending to zero. Obviously, this is the case if  $\lim_{\delta \rightarrow 0} \frac{\delta^2}{\alpha} = 0$ . Coupling the parameters  $\delta$ ,  $\alpha$  and  $\varepsilon$  in a proper way finally gives the regularization result also in the adaptive case, see [1, Theorem 4.2].

**Theorem 3.2** *Assume that  $K : \ell_2 \rightarrow Y$  is a bounded linear operator mapping into a Hilbert space  $Y$  with  $\|K\| < 1$ . Further let  $1 < p \leq 2$  and  $0 < w \leq w_\lambda$ . Let  $K\bar{\mathbf{x}} = y$ , with the ideal solution  $\bar{\mathbf{x}}$  and denote by  $\mathbf{x}^\dagger$  the unique minimum- $\|\cdot\|_{\mathbf{w},p}$ -solution of  $K\mathbf{x} = y$ . Let further the following assumptions be valid:*

- $\lim_{\delta \rightarrow 0} \alpha(\delta) = 0$
- $\lim_{\delta \rightarrow 0} \frac{\delta^2}{\alpha(\delta)} = 0$

and choose  $\varepsilon = \mathcal{O}(\alpha^\tau)$  with  $\tau > 1$ . Then we have

$$\lim_{\delta \rightarrow 0} \left( \sup_{\|K\bar{\mathbf{x}} - y^\delta\| \leq \delta} \|\mathbf{x}^{N(\delta), \delta} - \mathbf{x}^\dagger\| \right) = 0,$$

where  $\mathbf{x}^{N(\delta), \delta} = \mathbf{x}^N$  denotes the first iterate of iteration (19) such that  $d^N = \|\mathbf{x}^N - \mathbf{x}_\alpha^\delta\| < r(\varepsilon)$ .

## 3.2 Wavelets

In this section, we briefly recall the wavelet setting as far as it is needed for our purposes. In its most general sense, a *wavelet* basis  $\Psi = \{\psi_\lambda : \lambda \in \mathcal{J}\}$  is a Riesz basis for  $L_2(\Omega)$ , where  $\Omega$  denotes a domain in  $\mathbb{R}^d$  or a closed manifold. The indices  $\lambda \in \mathcal{J}$  typically encode several types of information, namely the *scale* (often denoted  $|\lambda|$ ), the spatial location and also the type of the wavelet. For instance, on the real line  $\lambda$  can be identified with  $(j, k)$ , where  $j = |\lambda|$  denotes the dyadic refinement level and  $2^{-j}k$  signifies the location of the wavelet.

We will not discuss at this point any technical description of the basis  $\Psi$ . Instead we assume that the domain  $\Omega$  under consideration enables us to construct a wavelet basis  $\Psi$  with the following properties:

- the wavelets are *local* in the sense that

$$\text{diam}(\text{supp}\psi_\lambda) \sim 2^{-|\lambda|}, \quad \lambda \in \mathcal{J};$$

- the wavelets satisfy the *cancellation property*

$$|\langle v, \psi_\lambda \rangle| \lesssim 2^{-|\lambda|\tilde{m}} \|v\|_{H^{\tilde{m}}(\text{supp}\psi_\lambda)},$$

where  $\tilde{m}$  denotes some suitable parameter, and

- the wavelet basis induces characterizations of Besov spaces of the form

$$\|f\|_{B_q^s(L_p(\Omega))} \sim \left( \sum_{|\lambda|=j_0}^{\infty} 2^{j(s+d(\frac{1}{2}-\frac{1}{p}))q} \left( \sum_{\lambda \in \mathcal{J}, |\lambda|=j} |\langle f, \tilde{\psi}_\lambda \rangle|^p \right)^{q/p} \right)^{1/q}, \quad (20)$$

where  $s > d\left(\frac{1}{p} - 1\right)_+$  and  $\tilde{\Psi} = \{\tilde{\psi}_\lambda : \lambda \in \mathcal{J}\}$  denotes the *dual basis*

$$\langle \psi_\lambda, \tilde{\psi}_\nu \rangle = \delta_{\lambda,\nu}, \quad \lambda, \nu \in \mathcal{J}.$$

**Remark 3.3** *i) By exploiting the norm equivalence (20) and using the fact that  $B_2^s(L_2(\Omega)) = H^s(\Omega)$ , a simple rescaling immediately yields a Riesz basis for  $H^s$ . We shall also assume that Dirichlet boundary conditions can be included, so that a characterization of the form (20) also holds for  $H_0^s(\Omega)$ .*

*ii) Suitable constructions of wavelets on domains can be found, e.g., in [6, 13–15]. We also refer to [7] for a detailed discussion.*

### 3.3 Adaptive Wavelet Schemes for Elliptic Problems

In this section, we briefly recall how wavelets can be used to treat elliptic operator equations of the form

$$\mathcal{A}u = f, \quad (21)$$

where we will assume  $\mathcal{A}$  to be a boundedly invertible operator from some Hilbert space  $X$  into its normed dual  $X'$ , i.e.,

$$\|\mathcal{A}u\|_{X'} \sim \|u\|_X, \quad u \in X.$$

We shall only discuss the basic ideas. For further information, the reader is referred to [8, 9, 12]. In our application,  $X$  will typically be a Sobolev space  $H^s(\Omega)$ . We shall mainly focus on the special case where

$$a(v, w) := \langle \mathcal{A}v, w \rangle$$

defines a *symmetric* bilinear form on  $X$  which is *elliptic* in the sense that

$$a(v, v) \sim \|v\|_X^2. \quad (22)$$

Usually, operator equations of the form (21) are solved by a Galerkin scheme, i.e., one defines an increasing sequence of finite dimensional approximation spaces  $S_{\Lambda_l} := \text{span}\{\eta_\mu : \mu \in \Lambda_l\}$ , where  $S_{\Lambda_l} \subset S_{\Lambda_{l+1}}$ , and projects the problem onto these spaces, i.e.,

$$\langle \mathcal{A}u_{\Lambda_l}, v \rangle = \langle f, v \rangle \quad \text{for all } v \in S_{\Lambda_l}.$$

To compute the actual Galerkin approximation, one has to solve a linear system

$$\mathbf{G}_{\Lambda_l} \mathbf{c}_{\Lambda_l} = \mathbf{f}_{\Lambda_l}, \quad \mathbf{G}_{\Lambda_l} = (\langle \mathcal{A}\eta_{\mu'}, \eta_\mu \rangle)_{\mu, \mu' \in \Lambda_l}, \quad (\mathbf{f}_{\Lambda_l})_\mu = \langle f, \eta_\mu \rangle, \quad \mu \in \Lambda_l.$$

Then the question arises how to choose the approximation spaces in a suitable way, for doing that in a somewhat clumsy fashion would yield a very inefficient scheme. One natural idea



would be to use an *adaptive* scheme, i.e., an updating strategy which essentially consists of the following three steps:

$$\begin{array}{ccc}
\text{solve} & - & \text{estimate} & - & \text{refine} \\
\mathbf{G}_{\Lambda_l} \mathbf{c}_{\Lambda_l} = \mathbf{f}_{\Lambda_l} & & \|u - u_{\Lambda_l}\| =? & & \text{add functions} \\
& & \text{a posteriori} & & \text{if necessary.} \\
& & \text{error estimator} & & 
\end{array}$$

Already the second step is highly nontrivial since the exact solution  $u$  is unknown, so that clever a posteriori error estimators are needed. Then another challenging task is to show that the refinement strategy leads to a convergent scheme and to estimate its order of convergence, if possible. In recent years, it has been shown that both tasks can be solved if wavelets are used as basis functions for the Galerkin scheme as we shall now explain.

The first step is to transform (21) into a discrete problem. By using the norm equivalences (20), it is easy to see that (21) is equivalent to

$$\mathbf{A} \mathbf{u} = \mathbf{f},$$

where

$$\mathbf{A} := \mathbf{D}^{-1} \langle \mathcal{A}\Psi, \Psi \rangle^T \mathbf{D}^{-1}, \quad \mathbf{u} := \mathbf{D} \mathbf{c}, \quad \mathbf{f} := \mathbf{D}^{-1} \langle f, \Psi \rangle^T, \quad \mathbf{D} = (2^{-s|\lambda|} \delta_{\lambda, \lambda'})_{\lambda, \lambda' \in \mathcal{J}},$$

and the computation of the Galerkin approximation amounts to solving the system

$$\mathbf{A}_\Lambda \mathbf{u}_\Lambda = \mathbf{f}_\Lambda := \mathbf{f}|_\Lambda, \quad \mathbf{A}_\Lambda := (2^{-s(|\lambda|+|\nu|)} \langle \psi_\lambda, \mathcal{A}\psi_\nu \rangle)_{\lambda, \nu \in \Lambda}.$$

Now, ellipticity (22) and the norm equivalences (20) yield

$$\|\mathbf{u} - \mathbf{u}_\Lambda\|_{\ell_2} \sim \|\mathbf{A}(\mathbf{u} - \mathbf{u}_\Lambda)\|_{\ell_2} \sim \|\mathbf{f} - \mathbf{A}(\mathbf{u}_\Lambda)\|_{\ell_2} \sim \|\mathbf{r}_\Lambda\|_{\ell_2},$$

so that the  $\ell_2$ -norm of the *residual*  $\mathbf{r}_\Lambda$  serves as an a posteriori error estimator. Each individual coefficient  $(\mathbf{r}_\Lambda)_\lambda$  can be viewed as a local error indicator. Therefore a natural adaptive strategy would consist in catching the bulk of the residual, i.e., to choose the new index set  $\hat{\Lambda}$  such that

$$\|\mathbf{r}_\Lambda|_{\hat{\Lambda}}\|_{\ell_2} \geq \zeta \|\mathbf{r}_\Lambda\|_{\ell_2}, \quad \text{for some } \zeta \in (0, 1).$$

However, such a scheme would not be implementable since the residual involves infinitely many coefficients. To transform this idea into an implementable scheme, the following three subroutines are needed:

- **RHS** $[\varepsilon, \mathbf{g}] \rightarrow \mathbf{g}_\varepsilon$ : determines for  $\mathbf{g} \in \ell_2(\mathcal{J})$  a finitely supported  $\mathbf{g}_\varepsilon \in \ell_2(\mathcal{J})$  such that

$$\|\mathbf{g} - \mathbf{g}_\varepsilon\|_{\ell_2(\mathcal{J})} \leq \varepsilon;$$

- **APPLY** $[\varepsilon, \mathbf{N}, \mathbf{v}] \rightarrow \mathbf{w}_\varepsilon$ : determines for  $\mathbf{N} \in L(\ell_2(\mathcal{J}))$  and for a finitely supported  $\mathbf{v} \in \ell_2(\mathcal{J})$  a finitely supported  $\mathbf{w}_\varepsilon$  such that

$$\|\mathbf{N}\mathbf{v} - \mathbf{w}_\varepsilon\|_{\ell_2(\mathcal{J})} \leq \varepsilon;$$

- **COARSE** $[\varepsilon, \mathbf{v}] \rightarrow \mathbf{v}_\varepsilon$ : determines for a finitely supported  $\mathbf{v} \in \ell_2(\mathcal{J})$  a finitely supported  $\mathbf{v}_\varepsilon \in \ell_2(\mathcal{J})$  with at most  $N$  significant coefficients, such that

$$\|\mathbf{v} - \mathbf{v}_\varepsilon\|_{\ell_2(\mathcal{J})} \leq \varepsilon. \quad (23)$$

Moreover,  $N \lesssim N_{\min}$  holds,  $N_{\min}$  being the minimal number of entries for which (23) is valid.

Then, employing the key idea outlined above, the resulting fundamental algorithm reads as follows:

**ALGORITHM 1.** **SOLVE** $[\varepsilon, \mathbf{A}, \mathbf{f}] \rightarrow \mathbf{u}_\varepsilon$ :

Fix target accuracy  $\varepsilon$ ,  $\Lambda_0 := \emptyset$ ,  $\mathbf{r}_{\Lambda_0} := \mathbf{f}$ ,  $\varepsilon_0 := \|\mathbf{f}\|_{\ell_2}$ ,  $j := 0$

While  $\varepsilon_j > \varepsilon$  do

$j := j + 1$

$\varepsilon_j := 2^{-(j+1)}\|\mathbf{f}\|_{\ell_2}$ ,  $\Lambda_{j,0} := \Lambda_j$ ,  $\mathbf{u}_{j,0} := \mathbf{u}_j$ ;

  For  $k = 1, \dots, M$  do

    Compute Galerkin approximation  $\mathbf{u}_{\Lambda_{j,k-1}}$  for  $\Lambda_{j,k-1}$ ;

    Compute  $\tilde{\mathbf{r}}_{\Lambda_{j,k-1}} := \mathbf{RHS}[c_1\varepsilon_{j+1}, \mathbf{f}] - \mathbf{APPLY}[c_1\varepsilon_{j+1}, \mathbf{A}, \mathbf{u}_{\Lambda_{j,k-1}}]$ ;

    Compute smallest set  $\Lambda_{j,k}$  s.t.  $\|\tilde{\mathbf{r}}_{\Lambda_{j,k-1}}|_{\Lambda_{j,k}}\|_{\ell_2} \geq \frac{1}{2}\|\tilde{\mathbf{r}}_{\Lambda_{j,k-1}}\|_{\ell_2}$ ;

  od

**COARSE** $[c_2\varepsilon_{j+1}, \mathbf{u}_{\Lambda_{j,k}}] \rightarrow (\Lambda_{j+1}, \mathbf{u}_{j+1})$

od

**Remark 3.4** *i) We shall not discuss in detail the concrete numerical realization of the three fundamental subroutines. The subroutine **COARSE** consists of a thresholding step, whereas **RHS** essentially requires the computation of a best  $n$ -term approximation. The most complicated building block is **APPLY**. Let us just mention that for elliptic operators with Schwartz kernels, the cancellation property of wavelets can be used to establish its existence. For further details, the reader is referred to [8, 9, 31].*

*ii) In **ALGORITHM 1**,  $c_1$  and  $c_2$  denote some suitably chosen constants whose concrete values depend on the problem at hand. Also the parameter  $M$  has to be chosen in a suitable way. We refer again to [8] for details.*

It can be shown that **ALGORITHM 1** has the following basic properties:

- **ALGORITHM 1** is guaranteed to converge for a huge class of problems, i.e.

$$\|\mathbf{u} - \mathbf{u}_\varepsilon\| \lesssim \varepsilon;$$

- The order of convergence of **ALGORITHM 1** is optimal in the sense that it asymptotically realizes the convergence order of best  $n$ -term wavelet approximation, i.e., if the best  $n$ -term approximation satisfies  $\mathcal{O}(n^{-s})$ , then

$$\|\mathbf{u} - \mathbf{u}_\varepsilon\| = \mathcal{O}((\#\text{supp}\mathbf{u}_\varepsilon)^{-s});$$

- The number of arithmetic operations stays proportional to the number of unknowns, that is, the number of flops needed to compute  $\mathbf{u}_\varepsilon$  satisfies  $\mathcal{O}(\#\text{supp}\mathbf{u}_\varepsilon)$ .

### 3.4 Adaptive Wavelet Schemes for Parabolic Problems

In this section, we turn to the development of adaptive wavelet-based numerical schemes for linear parabolic problems of the form

$$\begin{aligned} \frac{\partial}{\partial t} u(t, x) &= \mathcal{A}(x, \partial) u(t, x) + f(t, x) && \text{in } (0, T] \times \Omega, \\ u(t, x) &= 0 && \text{on } (0, T] \times \partial\Omega, \\ u(0, x) &= u_0(x) && \text{in } \Omega. \end{aligned} \quad (24)$$

We assume that we are given a Gelfand triple  $X \hookrightarrow V \hookrightarrow X'$  of Hilbert spaces and that  $\mathcal{A}(x, \partial) : X \rightarrow X'$  fits into the setting of Section 3.3. Moreover, we assume that the operator  $\mathcal{A} : D(\mathcal{A}) \subset V \rightarrow V$  is *sectorial* in the sense of [27].

The parabolic problem (24) may be considered as an abstract initial value problem

$$u'(t) = \mathcal{A}u(t) + f(t), \quad t \in (0, T], \quad u(0) = u_0, \quad (25)$$

for a Hilbert space-valued variable  $u : [0, T] \rightarrow V$ . For its numerical treatment, we use the Rothe method which is also known as the horizontal method of lines. Doing so, the discretization is performed in two major steps. Firstly, we consider a semidiscretization in time, where we will employ an  $S$ -stage linearly implicit scheme. We shall end up with an orbit of approximations  $u^{(n)} \in L_2(\Omega)$  at intermediate times  $t_n$  that are implicitly given via the  $S$  elliptic stage equations. In a finite element context, this very approach has already been propagated in [22, 23]. For the realization of the increment  $u^{(n)} \mapsto u^{(n+1)}$  and the spatial discretization of the stage equations, we will then employ the adaptive wavelet scheme introduced in Section 3.3 as a black box solver.

Let us start with the time discretization. We consider an  $S$ -stage linearly implicit method for the semidiscretization in time of *Rosenbrock–Wanner* (*ROW*)-type. By this we mean an iteration of the form

$$u^{(n+1)} = u^{(n)} + \sum_{i=1}^S m_i u_i \quad (26)$$

with the *stage equations*

$$\left( \frac{1}{h\gamma_{i,i}} I - J \right) u_i = \mathcal{A} \left( u^{(n)} + \sum_{j=1}^{i-1} a_{i,j} u_j \right) + f(t_n + \alpha_i h) + \sum_{j=1}^{i-1} \frac{c_{i,j}}{h} u_j + h\gamma_i g, \quad i = 1, \dots, S, \quad (27)$$

where  $J \approx \mathcal{A}$  and  $g \approx f'(t_n)$  and  $(\gamma_{i,j})_{i,j=1}^S, (a_{i,j})_{i,j=1}^S, (\alpha_{i,j})_{i,j=1}^S, (c_{i,j})_{i,j=1}^S, (m_1, \dots, m_S)^\top$  are suitably chosen parameters.

It is well-known that for a strongly  $A(\theta)$ -stable Rosenbrock method the numerical approximations according to (27) indeed converge to the exact solution as  $h \rightarrow 0$ , see [26] for details. However, a constant temporal step size  $h$  might not be the most economic choice. At least for times  $t$  close to 0 and in situations where the driving term  $f$  is not smooth at  $t$ , it is advisable to choose small values of  $h$  in order to track the behavior of the exact solution correctly. In regions where  $f$  and  $u$  are temporally smooth, larger time step sizes may be used. As a consequence, we have to employ an a posteriori temporal error estimator to control the current value of  $h$ . The traditional approach resorts to estimators for the local truncation error at  $t_n$

$$\delta_h(t_n) := \Phi^{t_n, t_n+h}(u(t_n)) - u(t_n + h),$$

where  $\Phi^{t_n, t_n+h} : V \rightarrow V$  is the increment mapping of the given Rosenbrock scheme at time  $t_n$  with step size  $h$ . For the global error at  $t = t_{n+1} = t_n + h_n$ , we have the decomposition

$$e_{n+1} = u^{(n+1)} - u(t_{n+1}) = \Phi^{t_n, t_n+h_n}(u^{(n)}) - \Phi^{t_n, t_n+h_n}(u(t_n)) + \delta_{h_n}(t_n),$$

i.e.,  $e_{n+1}$  comprises the local error at time  $t_n$  and the difference between the current Rosenbrock step  $\Phi^{t_n, t_n+h_n}(u^{(n)})$  and the virtual step  $\Phi^{t_n, t_n+h_n}(u(t_n))$  with starting point  $u(t_n)$ . Estimators for the local discretization error  $\delta_{h_n}(t_n)$  can be either based on an embedded lower order scheme or on extrapolation techniques, see [19, 20]. For applications to partial differential equations, embedding strategies yield sufficient results and thus are our method of choice.

Since the iteration (26) cannot be implemented numerically, we will now finally address the numerical approximation of all the ingredients by finite-dimensional counterparts. Precisely, we have to find approximate, computable iterates  $\tilde{u}^{(n+1)}$ , such that the additional error  $\tilde{u}^{(n+1)} - u^{(n+1)}$  introduced by the spatial discretization stays below some given tolerance  $\varepsilon$  when measured in an appropriate norm. Hence this perturbation of the virtual orbit  $\{u^{(n)}\}_{n \geq 0}$  can be interpreted as a controllable additional error of the temporal discretization. The accumulation of local perturbations in the course of the iteration is then an issue for the step size controller. In order not to spoil the convergence behavior of the unperturbed iterates  $u^{(n)}$  we will demand that  $\tilde{u}^{(n+1)} - u^{(n+1)}$  stays small in the topology of  $X$ , which results in the requirement

$$\|\tilde{u}^{(n+1)} - u^{(n+1)}\|_X \leq \varepsilon$$

for the numerical scheme, where  $\varepsilon > 0$  is the desired target accuracy. To achieve this goal, we want to use the convergent adaptive wavelet schemes as outlined in Subsection 3.3. Observe that by (26), the exact increment  $u^{(n+1)}$  differs from  $u^{(n)}$  by a linear combination of the exact solutions  $u_i$  of the  $S$  stage equations (27). The operators involved in (27) take the form

$$B_\alpha := \alpha I - \mathcal{A}, \quad \alpha \geq 0,$$

where  $\alpha = (h\gamma_{i,i})^{-1}$  for the  $i$ -th stage equation. By the estimate

$$\langle B_0 v, v \rangle \leq \langle B_\alpha v, v \rangle = \alpha \langle v, v \rangle_V + \langle B_0 v, v \rangle \leq (C\alpha + 1) \langle B_0 v, v \rangle, \quad v \in X,$$

we see that the energy norms  $\|v\|_{B_\alpha} := |\langle B_\alpha v, v \rangle|^{1/2}$  differ from  $\|v\|_{B_0} \approx \|v\|_X$  only by an  $\alpha$ -dependent constant:

$$\|v\|_{B_0} \leq \|v\|_{B_\alpha} \leq (C\alpha + 1)^{1/2} \|v\|_{B_0}, \quad v \in X.$$

Consequently, if we define

$$(\mathbf{D}_\alpha)_{\lambda, \lambda} := \|\psi_\lambda\|_{B_\alpha}, \quad \lambda \in \mathcal{J},$$

then the system  $\mathbf{D}_\alpha^{-1}\Psi$  is a Riesz basis in the energy space  $(H, \|\cdot\|_{B_\alpha})$ , with Riesz constants independent of  $\alpha \geq 0$ :

$$\|\mathbf{c}\|_{\ell_2} \sim \|\mathbf{c}^\top \mathbf{D}_\alpha^{-1}\Psi\|_{B_\alpha}, \quad \mathbf{c} \in \ell_2.$$

Therefore, we can use the Riesz basis  $\mathbf{D}_\alpha^{-1}\Psi$ ,  $\alpha = (h\gamma_{i,i})^{-1}$  as test functions in a variational formulation of (27). Abbreviating the exact right-hand side of (27) by

$$r_{i,h} := \mathcal{A}\left(u^{(n)} + \sum_{j=1}^{i-1} a_{i,j} u_j\right) + f(t_n + \alpha_i h) + \sum_{j=1}^{i-1} \frac{c_{i,j}}{h} u_j + h\gamma_i f'(t_n),$$

we get the system of equations

$$\langle B_\alpha u_i, \mathbf{D}_\alpha^{-1}\Psi \rangle^\top = \langle r_{i,h}, \mathbf{D}_\alpha^{-1}\Psi \rangle^\top. \quad (28)$$

Inserting a wavelet representation of  $u_i = (\mathbf{D}_\alpha \mathbf{u}_i)^\top \mathbf{D}_\alpha^{-1} \Psi$  into the variational formulation (28), we end up with the biinfinite linear system in  $\ell_2$

$$\mathbf{D}_\alpha^{-1} \langle B_\alpha \Psi, \Psi \rangle^\top \mathbf{D}_\alpha^{-1} \mathbf{D}_\alpha \mathbf{u}_i = \mathbf{D}_\alpha^{-1} \langle r_{i,h}, \Psi \rangle^\top. \quad (29)$$

Now we observe that problem (29) exactly fits into the setting of Subsection 3.3.

A detailed analysis of the concepts outlined above can be found in the PhD-thesis [29].

## 4 Numerical Investigations

In this section we present some numerical experiments for the IHCP from monitoring steel production introduced in Section 2. First we summarize the whole solution algorithm, which has been described in detail in the previous section. The numerical results are then presented in the second part of the current section.

### 4.1 The Algorithm for Solving the IHCP

In Section 3 we described several theoretical building blocks, which are necessary for solving the IHCP stated in Section 2. Now, we put everything together and specify the whole reconstruction algorithm for a fixed regularization parameter:

1. Choose a temperature distribution  $g^0$  on  $(0, T) \times \Gamma_0$ .
2. Choose a precision  $\varepsilon$ , set  $n = 0$ .
3. Solve the direct problem (4) with  $g = g^n$ , within a precision of  $\varepsilon$ .
4. Restrict the solution  $u$  to the outer boundary  $(0, T) \times \Gamma_1$ .
5. Calculate the residual  $u|_{(0,T) \times \Gamma_1} - \hat{h}^\delta$ .
6. Solve the adjoint problem (12) with  $h = u|_{(0,T) \times \Gamma_1} - \hat{h}^\delta$ , within a precision of  $\varepsilon$ .
7. Calculate the derivative of the solution  $v$  in step 6 in normal direction  $\frac{\partial v}{\partial \nu}|_{(0,T) \times \Gamma_0}$  on the inner boundary.
8. Determine the argument of the shrinkage operator  $g^n - (\frac{\partial v}{\partial \nu}|_{(0,T) \times \Gamma_0})$ .
9. Calculate the next iterate  $g^{n+1} = \mathbf{S}_{\alpha, w, p}(g^n - (\frac{\partial v}{\partial \nu}|_{(0,T) \times \Gamma_0}))$ .
10. Go back to step 3.

Having this strategy in mind, we consider its numerical realization extensively in the next subsection.

**Remark 4.1** *A crucial point in regularization theory is choosing the regularization parameter in an optimal way. In case of classical regularization schemes, like Tikhonov regularization or Landweber iteration with exact operator evaluations, there are several a priori and a posteriori parameter choice rules, which guarantee optimal convergence rates.*

*For the adaptive theory presented above, no theory for choosing the regularization parameter is available, so far. As far as we know there are not even any a posteriori parameter choice strategies for the presented generalized Tikhonov-regularization defined by (15), (16) using exact operator evaluations. However, this topic is still a current field of research and will be considered in a forthcoming paper.*

## 4.2 Numerical Results for the IHCP

For the numerical experiments on the ring-shaped domain  $\Omega$ , we choose radii  $r_0 = 0.5$  and  $r_1 = 2$ , and the final time  $T$  is set to 1. The spaces  $L_2(0, 1; L_2(\Gamma_i)) \cong L_2((0, 1) \times \Gamma_i)$  that contain the boundary data for the forward and adjoint problem,  $i \in \{0, 1\}$ , are discretized by using biorthogonal wavelet bases  $\Psi = \{\psi_\lambda\}_{\lambda \in \mathcal{J}}$ . By the specific geometry of  $\Omega$ , we can work in both cases with an appropriately lifted tensor product  $\Psi = \Psi^{\text{time}} \otimes \Psi^{\text{angular}}$  of a boundary-adapted wavelet basis  $\Psi^{\text{time}} \subset L_2(0, 1)$  for the temporal component and a periodic wavelet basis  $\Psi^{\text{angular}} \subset L_2(0, 1)^{\text{per}}$  for the angular component, using polar coordinates. In particular, we will choose  $\Psi^{\text{time}}$  to be a linear spline wavelet basis with two vanishing moments, as constructed in [28]. The angular basis  $\Psi^{\text{angular}}$  is taken as a periodized quadratic spline wavelet basis from [10] with three vanishing moments.

The sought temperature distribution  $x \in L_2((0, 1) \times \Gamma_0)$  should be sparse in the underlying wavelet system. To this end, for simplicity, we choose  $x$  to be one of the scaling functions within  $\Psi$  on the coarsest multiresolution level  $j_0$ , see Figure 2 for a plot. Internally, we then compute

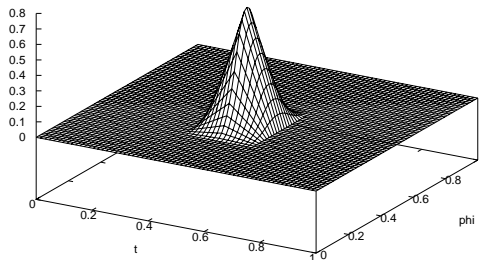


Figure 2: Exact solution  $x$  for the experiments.

the wavelet coefficient array of the observed data  $y = Ax$  as an input for the iterative shrinkage algorithm, augmented by additional noise of magnitude  $\delta \geq \|y^\delta - y\|$ .

Within the iteration, the forward and the adjoint operators are applied inexactly by an adaptive discretization of the generic parabolic problem

$$u_t = \Delta u + f, \quad u|_{(0,1) \times \Gamma_0} = 0, \quad \frac{\partial u}{\partial \nu}|_{(0,1) \times \Gamma_1} = 0, \quad u(0, \cdot) = u_0, \quad (30)$$

endowed with different driving terms  $f$  and initial data  $u_0$ , depending on whether we are solving an inhomogeneous Dirichlet or Neumann problem. For example, assume that some inhomogeneous Dirichlet data  $g$  are given for the forward operator  $K$ . Testing (4) with functions from  $V := \{v \in H^1(\Omega) \mid v|_{\Gamma_0} = 0\}$ , we observe that it suffices to look for solutions of the form  $u = w + G$ , where  $G(t) = Eg(t)$  denotes a bounded trace lifting of  $g(t) \in H^{1/2}(\Gamma_0)$  into the space  $H^1(\Omega)$ , and  $w$  solves the homogeneous problem (30) with a suitable driving term  $f$ . The concrete form of  $f$  can be derived when testing (4) with  $v \in V$ ,

$$\langle f(t), v \rangle_{V' \times V} = - \int_{\Omega} \nabla G(t) \nabla v \, dx - \left\langle \frac{\partial G(t)}{\partial t}, v \right\rangle_{V' \times V} \quad \text{for all } v \in V \text{ and a.e. } t \in (0, T).$$

In a completely analogous way, also inhomogeneous Neumann data  $h$  for the adjoint problem (12) can be transformed into an associated driving term  $\tilde{f}$  for the homogeneous problem (30). We refer to Section 2 for further details on the involved function spaces.

Using the specific geometry of  $\Omega$ , concrete extensions of the current Dirichlet and Neumann data may be obtained by multiplication with an appropriately shaped polynomial in radial direction. To be specific, Dirichlet data  $g \in L_2((0, 1) \times \Gamma_0)$  are extended into the space-time cylinder by the setting

$$G(t, \mathbf{x}) = g(t, \frac{r_0 \mathbf{x}}{\|\mathbf{x}\|})p(\|\mathbf{x}\|), \quad \text{for all } (t, \mathbf{x}) \in (0, 1) \times \Omega,$$

where  $p$  is a quadratic polynomial with  $p(r_0) = 1$  and  $p(r_1) = p'(r_1) = 0$ , ensuring that the extended boundary data vanish of second order at the exterior boundary  $\Gamma_1$ . Conversely, a global function with Neumann data  $h \in L_2((0, 1) \times \Gamma_1)$  and second-order zero boundary conditions at  $\Gamma_0$  is given by

$$H(t, \mathbf{x}) = h(t, \frac{r_1 \mathbf{x}}{\|\mathbf{x}\|})q(\|\mathbf{x}\|), \quad \text{for all } (t, \mathbf{x}) \in (0, 1) \times \Omega,$$

where  $q$  is a quadratic polynomial with  $q(r_0) = q'(r_0) = 0$  and  $q'(r_1) = 1$ .

The parabolic subproblems are treated with the Rothe method, see Subsection 3.4, using inexact linearly implicit increments [29]. For the experiments, we choose the second-order Rosenbrock scheme ROS2. The elliptic subproblems are solved by the adaptive wavelet-Galerkin algorithm from [8], up to a prescribed accuracy. Due to the fact that the temporal wavelet basis  $\Psi^{\text{time}}$  is constructed from interpolatory scaling functions, a transformation of space-time data into a spatial wavelet representation at dyadic time nodes  $t = 2^{-j}k$  is feasible via 2D wavelet transforms.

In the following, starting from  $f_0 = 0$ , we present the iterates after a certain fixed number of shrinkage iterations. The weights  $w_\lambda$  are chosen to be 1. By the norm equivalences (20) for the wavelet basis  $\Psi$ , this choice in turn corresponds to a  $B_{p,p}^{2/p-1}$  penalty term in the Tikhonov functional. The tolerance  $\varepsilon$  for the inexact forward and adjoint solves is chosen proportional to the regularization parameter  $\alpha$ , corresponding to the limit case  $\tau \rightarrow 1$  in Theorem 3.2.

We first show some results for the case  $p = 1.1$  and for certain values of the regularization parameter  $\alpha$ . For an absolute noise level  $\delta = 10^{-3}$ , the results from Figure 3 show that iterative shrinkage can locate the local features of the target function quite well. However, the amplitudes of the reconstructions are significantly smaller, than in the exact solution. There are still quite a few active wavelet coefficients in regions where the exact solution  $x$  is zero.

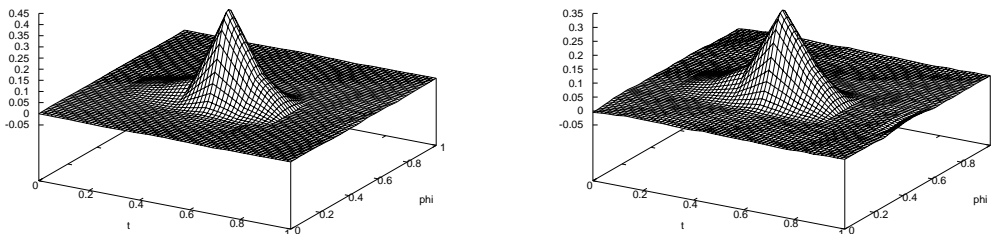


Figure 3: Reconstructions for  $p = 1.1$  after 2500 iterations ( $\delta = 0.001$ ) for  $\alpha = 9.8e - 4$  (left) and  $\alpha = 2.0e - 3$  (right)

Interestingly, the results improve considerably if iterative shrinkage is replaced by iterative *thresholding*, i.e., switching to the case  $p = 1$ . Without having theoretical results available, we choose the same dependence of the accuracies  $\varepsilon$  on  $\alpha$ . Results after 2500 thresholding iterations

can be found in Figure 4. We see that thresholding effectively reduced the number of degrees of freedom. In regions where the exact solution  $x$  is zero, also the reconstructions vanish.

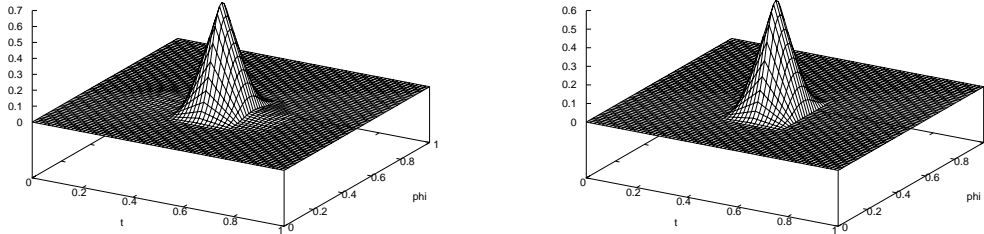


Figure 4: Reconstructions for  $p = 1$  after 2500 iterations ( $\delta = 0.001$ ) for  $\alpha = 4.9e - 4$  (*left*) and  $\alpha = 2.0e - 3$  (*right*)

In Figures 5 and 6, the absolute noise level  $\delta$  is increased to  $10^{-2}$ . We observe the same effect as for low noise. Both for  $p = 1.1$  and for  $p = 1$ , the localization of the reconstructions is quite good. But the amplitude of the exact solution is reconstructed considerably better when using thresholding iterations instead of mere shrinkage. Moreover, the reconstructions exhibit fewer active wavelet coefficients, due to the soft thresholding operations.

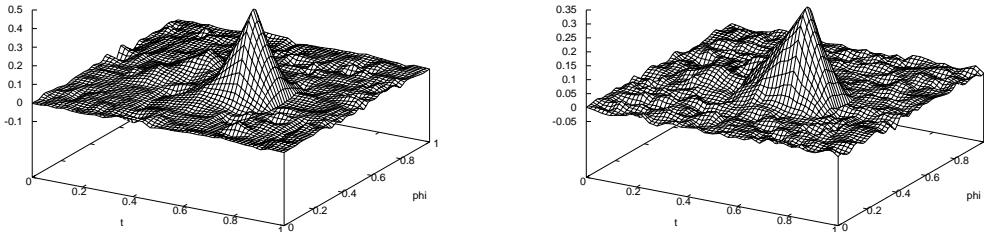


Figure 5: Reconstructions for  $p = 1.1$  after 2500 iterations ( $\delta = 0.01$ ) for  $\alpha = 9.8e - 4$  (*left*) and  $\alpha = 2.0e - 3$  (*right*)

Some figures concerning the summands of the Tikhonov functional for certain reconstructions  $x$  can be found in Table 1. As is already visible to some extent in the preceding figures, the reconstructions for  $p = 1.1$  require more active wavelet coefficients than for  $p = 1$ , in particular when the noise parameter  $\delta$  grows. As an effect, the results for  $p = 1.1$  exhibit significantly larger residual errors than those for  $p = 1$ , underlining the visual anticipation from the presented plots.

Up to a certain noise level and with an appropriate choice of the regularization parameter  $\alpha$ , we see that it is indeed possible to obtain almost perfect reconstructions of target functions that are sparse in the underlying wavelet system.



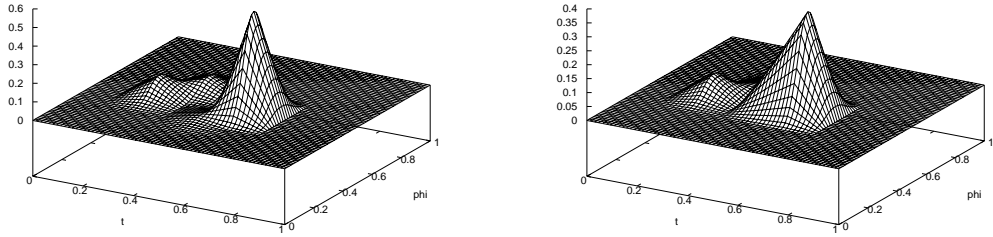


Figure 6: Reconstructions for  $p = 1$  after 2500 iterations ( $\delta = 0.01$ ) for  $\alpha = 9.8e - 4$  (left) and  $\alpha = 2.0e - 3$  (right)

$\delta$	$p$	$\alpha$	$\ Kx - y^\delta\ _{L_2}$	$\ x\ _{\mathbf{w},p}$
$10^{-3}$	1	$4.9e - 4$	$1.46e - 3$	$2.11e - 1$
	1.1	$9.8e - 4$	$3.12e - 3$	$2.07e - 1$
$10^{-2}$	1	$9.8e - 4$	$6.12e - 3$	$2.32e - 1$
	1.1	$9.8e - 4$	$5.29e - 3$	$4.40e - 1$

Table 1: Residual norms and penalty terms for certain reconstructions.

## 5 Conclusions

In this paper, we have outlined how the abstract analysis investigated in [1] can be applied to a very concrete application, i.e., to an inverse heat conduction problem which arises from monitoring steel production. The method consists of an adaptive iterated soft shrinkage procedure based on wavelets. The algorithm requires the approximate solution of the forward problem, and for this, also an adaptive wavelet scheme has been employed. The use of wavelets for both issues is advantageous for the following reasons. Firstly, wavelets are well-suited for the detection of sparse structures. Secondly, adaptive wavelet schemes for elliptic and parabolic problems are guaranteed to converge with optimal order. Thirdly, our approach avoids the complicated and inconvenient change of bases when different systems of ansatz functions are used. The numerical experiments clearly confirm the power of our approach. More numerical experiments will be presented in the near future.

## References

- [1] T. Bonesky and P. Maass, *Iterated soft shrinkage with adaptive operator evaluations*, preprint, 2008, To appear in *Inverse and Ill-Posed Problems*.
- [2] L. Borcea, *Electrical impedance tomography*, *Inverse Problems* **18** (2002), no. 6, R99–R136.
- [3] K. Bredies and D. A. Lorenz, *Iterated hard shrinkage for minimization problems with sparsity constraints*, *SIAM J. Sci. Comput.* **30** (2008), no. 2, 657–683.
- [4] ———, *Linear convergence of iterated soft-thresholding*, preprint, 2008, To appear in *Journal of Fourier Analysis and Applications*.

- [5] K. Bredies, D. A. Lorenz, and P. Maass, *A generalized conditional gradient method and its connection to an iterative shrinkage method*, To appear in *Computational Optimization and Applications*, 2008.
- [6] C. Canuto, A. Tabacco, and K. Urban, *The wavelet element method, part II: Realization and additional features in 2D and 3D*, *Appl. Comput. Harmon. Anal.* **8** (2000), 123–165.
- [7] A. Cohen, *Wavelet methods in numerical analysis*, *Handbook of Numerical Analysis* (P.G. Ciarlet and J.L. Lions, eds.), vol. VII, North-Holland, Amsterdam, 2000, pp. 417–711.
- [8] A. Cohen, W. Dahmen, and R. DeVore, *Adaptive wavelet methods for elliptic operator equations – Convergence rates*, *Math. Comput.* **70** (2001), no. 233, 27–75.
- [9] ———, *Adaptive wavelet methods II: Beyond the elliptic case*, *Found. Comput. Math.* **2** (2002), no. 3, 203–245.
- [10] A. Cohen, I. Daubechies, and J.-C. Feauveau, *Biorthogonal bases of compactly supported wavelets*, *Commun. Pure Appl. Math.* **45** (1992), 485–560.
- [11] P. L. Combettes and V. R. Wajs, *Signal recovery by proximal forward-backward splitting*, *Multiscale Model. Simul.* **4** (2005), no. 4, 1168–1200.
- [12] S. Dahlke, W. Dahmen, R. Hochmuth, and R. Schneider, *Stable multiscale bases and local error estimation for elliptic problems*, *Appl. Numer. Math.* **23** (1997), 21–48.
- [13] W. Dahmen and R. Schneider, *Wavelets with complementary boundary conditions — Function spaces on the cube*, *Result. Math.* **34** (1998), no. 3–4, 255–293.
- [14] ———, *Composite wavelet bases for operator equations*, *Math. Comput.* **68** (1999), 1533–1567.
- [15] ———, *Wavelets on manifolds I. Construction and domain decomposition*, *SIAM J. Math. Anal.* **31** (1999), 184–230.
- [16] I. Daubechies, M. Defrise, and C. De Mol, *An iterative thresholding algorithm for linear inverse problems with a sparsity constraint*, *Commun. Pure Appl. Math.* **57** (2004), no. 11, 1413–1457.
- [17] L. Eldén, *The numerical solution of a non-characteristic Cauchy problem for a parabolic equation*, *Numerical treatment of inverse problems in differential and integral equations*, *Proc. int. Workshop, Heidelberg 1982*, *Prog. Sci. Comput.*, vol. 2, 1983, pp. 246–268.
- [18] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of inverse problems*, Kluwer, 1996.
- [19] E. Hairer, S. P. Nørsett, and G. Wanner, *Solving ordinary differential equations. I: Nonstiff problems*, 2nd rev. ed., *Springer Series in Computational Mathematics*, vol. 8, Springer, Berlin, 1993.
- [20] E. Hairer and G. Wanner, *Solving ordinary differential equations. II: Stiff and differential-algebraic problems*, 2nd rev. ed., *Springer Series in Computational Mathematics*, vol. 14, Springer, Berlin, 1996.
- [21] O. A. Ladyženskaja, V. A. Solonnikov, and N. N. Ural’ceva, *Linear and quasilinear equations of parabolic type*, American Mathematical Society, Providence, Rhode Island, 1968.

- [22] J. Lang, *Adaptive multilevel solution of nonlinear parabolic PDE systems. Theory, algorithm, and applications*, Preprint SC 99–20, Konrad-Zuse-Zentrum für Informationstechnik Berlin, 1999.
- [23] ———, *Adaptive multilevel solution of nonlinear parabolic PDE systems. Theory, algorithm, and applications*, Lecture Notes in Computational Science and Engineering, vol. 16, Springer, Berlin, 2001.
- [24] J.L. Lions and E. Magenes, *Problèmes aux limites non homogènes et applications*, vol. 2, Dunod, Paris, 1968.
- [25] A. K. Louis, *Inverse und schlechtgestellte Probleme*, Teubner, Stuttgart, 1989.
- [26] C. Lubich and A. Ostermann, *Linearly implicit time discretization of nonlinear parabolic equations*, IMA J. Numer. Anal. **15** (1995), no. 4, 555–583.
- [27] A. Lunardi, *Analytic semigroups and optimal regularity in parabolic problems*, Progress in Nonlinear Differential Equations and Their Applications, vol. 16, Birkhäuser, 1995.
- [28] M. Primbs, *Stabile biorthogonale Spline-Waveletbasen auf dem Intervall*, Dissertation, Universität Duisburg-Essen, 2006.
- [29] T. Raasch, *Adaptive wavelet and frame schemes for elliptic and parabolic equations*, Dissertation, Philipps-Universität Marburg, 2007.
- [30] R. Ramlau, G. Teschke, and M. Zharyi, *A compressive Landweber iteration for solving ill-posed inverse problems*, Inv. Problems **24** (2008), no. 6, 065013.
- [31] R. Stevenson, *Adaptive solution of operator equations using wavelet frames*, SIAM J. Numer. Anal. **41** (2003), no. 3, 1074–1100.
- [32] D. V. Widder, *The heat equation*, Pure and Applied Mathematics, vol. 67, Academic Press, 1976.

# Preprint Series DFG-SPP 1324

<http://www.dfg-spp1324.de>

## Reports

- [1] R. Ramlau, G. Teschke, and M. Zhariy. A Compressive Landweber Iteration for Solving Ill-Posed Inverse Problems. Preprint 1, DFG-SPP 1324, September 2008.
- [2] G. Plonka. The Easy Path Wavelet Transform: A New Adaptive Wavelet Transform for Sparse Representation of Two-dimensional Data. Preprint 2, DFG-SPP 1324, September 2008.
- [3] E. Novak and H. Woźniakowski. Optimal Order of Convergence and (In-) Tractability of Multivariate Approximation of Smooth Functions. Preprint 3, DFG-SPP 1324, October 2008.
- [4] M. Espig, L. Grasedyck, and W. Hackbusch. Black Box Low Tensor Rank Approximation using Fibre-Crosses. Preprint 4, DFG-SPP 1324, October 2008.
- [5] T. Bonesky, S. Dahlke, P. Maass, and T. Raasch. Adaptive Wavelet Methods and Sparsity Reconstruction for Inverse Heat Conduction Problems. Preprint 5, DFG-SPP 1324, January 2009.