

# Operator Equations, Mixed Constraints, Frame-Based Iterative Concepts and Applications

*Habilitationsschrift*

Dr. rer. nat. Gerd Teschke

Zentrum für Technomathematik  
Universität Bremen  
PF 33 04 40  
28334 Bremen  
Germany

20. Juni 2005



*for Caroline*





## Acknowledgments

I gratefully acknowledge the contributions of all the people who supported me during the work on this thesis. First of all, I want to thank my parents for being with me all the time. Special thanks are devoted to all my collaborators who directed my attention to the topics that are partially presented here in the thesis and who taught me to see the world not only as a mathematician.

I am forced to thank especially Peter Maaß for supporting and encouraging me in the last years. A very special thank to Ingrid Daubechies for drawing my attention to very interesting topics in harmonic analysis and inverse problems. She gave me the opportunity to work in Princeton; very kindly she has created a nice and stimulating atmosphere in which I felt comfortable - thanks Ingrid. Moreover, I want to thank Stephan Dahlke who directed me on the path of frames, function spaces, and operator equations. He accompanied me from the beginning on and he taught me a lot. A special thank is also devoted to Ronny Ramlau, my local collaborator. Together with him we developed the iterative concepts for nonlinear problems. Furthermore I want to thank Christine DeMol, Gabriele Steidl, Luminita Vese, Michel Defrise, Matthias Holschneider, Bruno Torrèsani, Massimo Fornasier, Matthias Räscher, Dirk Lorenz, and many more for inspiring discussions related to the topic my this thesis.

Finally, I want to thank Caroline for her love and emotional support.



## **Abstract**

The primary goal of this thesis is to develop new iterative concepts for solving linear and nonlinear operator equations in its variational form. The basic novel ingredients are multi frames and mixed sparsity and smoothness constraints. The secondary goal consists of elaborating special properties of the schemes and applying them in the context of image and signal processing, inverse problems, harmonic analysis and machine learning.

## **Zusammenfassung**

Der Hauptzweck dieser Arbeit besteht in der Entwicklung neuer iterativer Konzepte zur Lösung linearer und nichtlinearer Operatorgleichungen. Die Neuheit zeichnet sich dadurch aus, dass in der Variationsformulierung der Aufgabe gemischte Randbedingungen betrachtet und Multi Frames eingesetzt werden. Darüber hinaus werden Eigenschaften der entwickelten Verfahren analysiert und die Methoden dann in Bereichen wie Bild und Signalverarbeitung, Inverse Probleme, Harmonischer Analysis und Lern- bzw. Klassifikationsmaschinen angewendet.



# Contents

<b>1</b>	<b>Introduction</b>	<b>11</b>
<b>2</b>	<b>Linear Operator Equations and Iterative Concepts</b>	<b>17</b>
2.1	Preliminaries . . . . .	17
2.1.1	Frames . . . . .	18
2.1.2	Linear Problems . . . . .	19
2.1.3	Smoothness, Sparsity, and other Constraints . . . . .	20
2.2	Multi-Frames and Mixed Constraints . . . . .	22
2.2.1	Formulation of the Variational Problem . . . . .	22
2.2.2	Minimization of Surrogate Functionals . . . . .	25
2.2.3	Convergence Analysis . . . . .	28
2.2.4	Regularization Properties . . . . .	34
<b>3</b>	<b>Applications I: Linear Problems</b>	<b>37</b>
3.1	Compression for Audio and Image Signals . . . . .	37
3.1.1	Application to Audio Coding . . . . .	37
3.1.2	Application to Image Restoration and Compression . . . . .	38
3.2	Image Decomposition and Restoration Problems . . . . .	43
3.2.1	Variational Problem with Smoothness and Sparsity Constraints . . . . .	43
3.2.2	Approximation of the Solution . . . . .	45
3.2.3	Numerics and Improvements by Redundancy . . . . .	48
3.3	Acceleration of Support Vector Machines . . . . .	57
3.3.1	On Support Vector Machines and its Reduction . . . . .	57
3.3.2	Wavelet Frame Approximated Support Vector Machine . . . . .	60
3.3.3	Numerical Verifications . . . . .	67
<b>4</b>	<b>Nonlinear Operator Equations and Iterative Concepts</b>	<b>71</b>
4.1	Nonlinear Problems and Quadratic Constraints . . . . .	71
4.1.1	Formulation of the Variational Problem . . . . .	73
4.1.2	Proper Surrogate Functionals . . . . .	74
4.1.3	Minimization of Surrogate Functionals . . . . .	79
4.1.4	Convergence Analysis . . . . .	84
4.1.5	Regularization Properties . . . . .	89
4.2	Multi-Frames and Mixed One-Homogeneous Constraints . . . . .	92
4.2.1	Scope of the Problem . . . . .	92

4.2.2	Proper Surrogate Functionals . . . . .	94
4.2.3	Minimization yields Projected Fixed Point Iterations . . . . .	99
4.2.4	Convergence Analysis . . . . .	106
4.2.5	A Regularization result . . . . .	108
<b>5</b>	<b>Applications II: Nonlinear Problems</b>	<b>111</b>
5.1	Damped Landweber Fixed Point Iteration for SPECT . . . . .	111
5.2	Computation of Optimally Localized Wavelets . . . . .	114
5.2.1	Motivation and Basic Formulas . . . . .	114
5.2.2	Optimally Localized States and a General Existence Theorem . . . . .	116
5.2.3	Numerical Approximation of Localized States . . . . .	122
5.3	Thresholding Landweber Fixed Point Iteration – An Illustration . . . . .	126
5.4	Acceleration of Support Vector Machines . . . . .	130
<b>6</b>	<b>Perspectives on Adaptive Frame Strategies</b>	<b>133</b>
6.1	Brief Review on an Adaptive Frame Method . . . . .	134
6.2	Adaptivity for Solving Variational Problems . . . . .	136
	<b>Literature</b>	<b>139</b>

# Chapter 1

## Introduction

New and promizing developments especially in applied mathematics are made by merging theories and applications of surrounding areas. Quite recently, first progress was made while bringing together aspects from operator equations, inverse problems, harmonic analysis, frame theory, convex and non-convex analysis etc. The merging of these building blocks from several areas is often caused by and has impact in many important applications like medicine, physics, astrophysics, machine learning, and in the wide range of image and signal processing. Naturally, one is faced with a variety of novel and very interesting mathematical questions, e.g. how to solve an operator equation by means of frames, or how to solve the problem when being restricted to certain ‘non-classical’ constraints. A related but more applied question in this context might be: how to build sparse and fast variants of support vector machines.

In this thesis we shall mainly consider the following situation: given a linear or nonlinear operator equation being potentially ill-posed. Basically, our first step is to find some feasible way of rewriting the problem in its variational form. Furthermore, involving certain properties on the solution to be approximated we answer the question what is a proper analytical and a numerically thrifty way for solving the ‘restricted’ variational formulation. The *novelty of this thesis* is that we chose the concept of surrogate functionals for recasting the problem and for discretizing we develop the concept of multi frames, i.e. we abstain from preselecting a particular basis. Instead we allow highly redundant dictionaries of frames. This assures sort of ‘optimal’ approximation of the solution. Moreover, we involve constraints, e.g. a mixture of smoothness and sparsity, that are beyond classical theories. In particular, we consider multi constraints of non-quadratic functionals. These two new ingredients in combination with the nonlinearity of the operator equation under consideration extend the classical known theory substantially. The capabilities of the resulting strategies are demonstrated in diverse applications. Moreover, from the point of view of inverse problems, we ask of course also for additional very important features such as convergence and regularization properties of the constructed schemes.

*In a nutshell:* the primary goal of this thesis is to develop numerical schemes for solving linear and nonlinear operator equations whereas the secondary goal consists of elaborating special properties of the schemes and applying all in the context of image and signal processing, inverse problems, harmonic analysis and machine learning.

We are especially interested in mathematical problems where the features or signals of interest cannot be observed directly, but have to be interfered from other observable quantities. At least, one may find some nonlinear relationship between the feature modeled by a function  $v$ , and the derived quantities modeled by another function  $z$  such that we may formulate the problem by an operator equation

$$F(v) = z .$$

Such a problem makes only sense when everything is placed in an adequate setting. Often, when dealing with real data, the observation  $f$  are not exactly equal to  $z$ , but a distortion of  $z$ . In this situation, we classify the problem ill-posed if the solution of the problem does not depend continuously on the observations. In such an ill-posed case the solution might differ substantially from the searched quantity, even if there was only a little distortion in the data  $z$ . In order to stabilize or to circumvent these effects, one has to use so-called regularization concepts. In linear and nonlinear lore, there still exists a number of methods, e.g. Tikhonov regularization and iterative strategies like Landweber methods, Levenberg-Marquardt methods, Gauss-Newton, conjugate gradients etc, that are regularization schemes for this kind of problems, see e.g. [Lan51, Sch98, EHN96a, Han95, Ram]. The computation of a minimizer of the Tikhonov functional with some quadratic constraint

$$\Phi(v) = \|f - F(v)\|^2 + \alpha\|v\|^2$$

is difficult due to the nonlinearity of the operator. Contrary to the linear case where the functional is convex and the minimizer can be computed via

$$v = (F^*F + \alpha I)^{-1}F^*f ,$$

the Tikhonov functional for the nonlinear case is non-convex and might thus have several local minimizers, and the results of the minimization routines might strongly depend on the initial guess. As still mentioned, alternatively to Tikhonov strategies one may also use iterative concepts which produce an approximation of the solution within each iteration step. In the ill-posed situation, the iteration has to be terminated adequately. The stopping index plays then the role of the regularization parameter. However, to show for iterative schemes convergence rates and regularization properties is more difficult than for Tikhonov based concepts. Nevertheless, since iterative techniques are usually not too difficult to perform, they are mostly used for many applications in the range of inverse problems as well as in image and signal processing.

Quite recently, a first bridging between Tikhonov functionals and Landweber-like iterations was established in [DDD04]. For some linear operator  $F$ , functionals of the form

$$\Phi(v) = \|f - F(v)\|^2 + \alpha\|v\|_{\ell_{p,w}}^p$$

are considered. Applying methods from surrogate variational calculus, the minimization amounts to a Landweber iteration with some ‘general’ shrinkage operation applied in each iteration step,

$$v_{k+1} = \mathbf{S}_{\alpha, \mathbf{w}}(v_k + F^*(f - Fv_k)) .$$



In the quadratic case ( $p = 2$ ) this is nothing than a classical damped Landweber iteration, i.e. starting with some Tikhonov functional, a version of Landwebers iteration is constructed. This may allow for  $p = 2$  in some sense a dual consideration of regularization results for Tikhonov and Landweber methods.

Grabbing now the vision and the spine of [DDD04], one may ask the following natural questions which we aim to answer in this thesis at hand:

- *Why not represent the solution of the linear problem by multi frames instead of using some preselected basis ?*

This makes sense when searching for an ‘optimal’ representations of the solution, e.g. searching for the sparsest representation. Motivations in this direction are given by recent developments in approximation theory using highly redundant dictionaries of atoms.

- *When dealing with multi frames, what about a mixing of constraints on the individual frame coefficients ?*

Developments in convex and non-convex analysis provide certain tools that seem to be promizing for treating penalties that are beyond the quadratic ones.

- *Main question: Is there a natural way to extend the whole machinery to nonlinear operator equations ?*

Typically, most of the practical problems are modeled by some nonlinear relationship. This requires totally new concepts in order to involve the operator under consideration in its full nonlinearity. Our approach provides new methods to solve such problems. Moreover, we may allow the solution to have a ‘mixed’ (sparse and/or smooth) representation by means of atoms coming from a multi frame dictionary.

- *What about numerical improvements by involving adaptive strategies ?*

Recent progress, see e.g. [Ste03], for solving operator equations by means of frames in an adaptive framework suggests a possible strategy on how to proceed for variational problems considered here. In this thesis we do not focus on precise elaborations on that topic but we want to sketch some ideas.

The organization of the thesis follows these questions and is essentially based on seven papers:

*Summary Chapter 2.* In this chapter we are concerned with linear inverse problems where the solution is assumed to have sparse and/or smooth expansion with respect to several bases or frames. We develop a regularization scheme which is sort of Landweber iteration with specific frame-wise  $\ell_p$ -thresholding in each step. The work was mainly inspired and driven by discussions on audio coding with B. Torr  sani, see also [MT05, JT05], and uses technical concepts for linear inverse problems with sparsity constraints, see [DDD04], acquired during my sabbatical at Princeton University where I worked with I. Daubechies

on related problems. The results shown here are published in [Tes05b].

*Summary Chapter 3.* The theory for linear problems elaborated in the previous Chapter 2 will here be applied in several fields, such as audio and image coding, published in [Tes05b], image decomposition and restoration problems (texture analysis), published in [DT04, DT05], and finally, for speeding-up Reduced Support Vector Machines, published in [RRTV05].

*Summary Chapter 4.* This chapter is devoted to the main question: development of schemes for nonlinear problems, and is split in two sections. In the first Section 4.1, we are interested in algorithms for the computation of a minimizer of the associated non-convex Tikhonov functional. Basically, as in the Chapter 2, we aim to use the techniques of surrogate functionals but now in order to introduce convex replacement functionals that are better suited as the non-convex one and where the sequence of the minimizers converge to a minimizer of the original problem. Moreover, assuming certain smoothness conditions on the problem we may state regularization properties. This theory was developed in collaboration with R. Ramlau and is published in [RT04].

The second Section 4.2 is concerned with nonlinear inverse problems where the solution, as in Chapter 2, is assumed to have a sparse expansion with respect to several preassigned bases or frames. We develop a new scheme which allows to minimize a non-convex Tikhonov functional where the usual quadratic regularization term is replaced by a one-homogeneous (typically weighted  $\ell_p$ ) penalties on the coefficients (or isometrically transformed coefficients) of such multi-frame expansions. The computation of the solution amounts in this setting to a system of Landweber-fixed-point iterations with projections ( $\ell_p$ -thresholding) applied in each fixed-point iteration step. The here presented theory is published in [Tes05c].

*Summary Chapter 5.* Within this chapter we apply the theory for nonlinear problems established in Sections 4.1 and 4.2 to the analysis of Single Photon Emission Computerized Tomography (SPECT), the construction of optimally localized coherent states, image decomposition tasks, and to support vector machines. In the first two sections, we want to apply the machinery developed in the sections 4.1. The aim is to demonstrate the capabilities and the performance of our algorithm in solving a full nonlinear ill-posed SPECT problem. Moreover, we shall see that also the computation of optimally localized states is an example where the elaborated scheme of Section 4.1 can be usefully applied. We are concerned with localization properties of coherent states. Instead of looking in the context of classical uncertainty relations we consider more ‘generalized’ localization quantities. This is done by introducing measures on the reproducing kernel. Beside proving the existence of such optimally localized states, we shall see that the numerical computation (approximation) fits into the class of variational problems with quadratic constraints. The results written up here are published in [HT05]. In section three and four of this chapter we apply the concepts developed in Section 4.2. At first, we apply the techniques to nonlinear image deformation problems (which is somehow artificial but for illustrating purposes well suited). Secondly, we consider as in Section 3.3 the problem of accelerating support vector machines but we discuss here the full nonlinear problem, i.e. simultaneously reducing the number of set vectors and sparsely approximating them.

*Summary Chapter 6.* The last chapter is devoted to adaptivity. We briefly sketch on how adaptivity can be incorporated in the presented iteration schemes. Firstly, based on [Ste03], we review recently developed frame techniques for operator equations. The iteration methods mentioned there turn out to be Landweber iterations which are under consideration in this thesis. We discuss in brief the relations and possible extensions.



## Chapter 2

# Linear Operator Equations and Iterative Concepts

### 2.1 Preliminaries

Recent studies in the field of signal processing and inverse problems have shown the importance of sparse representations for various tasks, such as signal compression, denoising etc. Typically, such sparse representations are achieved by using a suitable orthonormal basis in the underlying function space. However, recent developments also indicate that redundant systems, such as frames, or dictionaries of ‘waveform’ systems may yield a gain in this context.

When dealing with dictionaries of ‘waveform’ systems, there exist several methods, e.g. best orthogonal basis, matching pursuit, basis pursuit etc., see, e.g., [CDS95], that allow a decomposition of a signal into an ‘optimal’ superposition of dictionary elements, where optimal means having the smallest  $\ell_1$  norm of coefficients among all such decompositions. At least basis pursuit in highly over-complete dictionaries leads to very large scale optimization problems (but can be attacked by linear programming, i.e. by interior-point methods).

In this chapter we develop a new iterative method for finding the  $\ell_p$ -optimal decomposition ( $1 \leq p \leq 2$ ) of a given signal into dictionary building blocks. The skeletal idea of this scheme was originally discovered for solving linear inverse problems with one sparsity constraint, see [DDD04]. But instead of preselecting one orthonormal basis or frame only, we typically assume that the signal might be a superposition  $n$  different components and thus, we pick a dictionary consisting of a family of  $n$  frames. Moreover, we combine this with an inverse problem, namely assuming that we have not observed the signal directly, but only other quantities that are linearly related to the signal.

The advantage of the proposed method is that for achieving convergence of the iteration process, we do not need to require any further assumptions on the preselected family of frames (e.g. such as incoherence). Moreover, since each individual frame is separately penalized we may mix the constraints, i.e. the penalties may vary from  $\ell_1$ -sparsity to  $\ell_2$ -quadratic smoothness constraints. A similar attempt where mixed (sparsity and smoothness) constraints were used was made, e.g., in [DT04, DT05, DD04a], but these approaches involve one single basis/frame only.

The remaining chapter is organized follows: at first, we review a few facts on frames, linear inverse problems and constraints. Then, in Section 2.2, we present the main result, which is the introduction of what we call the multi-frame concept, the resulting variational problem, its minimization, the convergence and the stability analysis.

### 2.1.1 Frames

A frame  $\{\phi_\lambda\}_{\lambda \in \Lambda}$  in a Hilbert space  $\mathcal{H}$  is a set of vectors for which there exists constants  $A, B > 0$  such that, for all  $v \in \mathcal{H}$ ,

$$A\|v\|_{\mathcal{H}}^2 \leq \sum_{\lambda \in \Lambda} |\langle v, \phi_\lambda \rangle_{\mathcal{H}}|^2 \leq B\|v\|_{\mathcal{H}}^2 ,$$

see for the roots of frames [DS52]. Frames are typically ‘over-complete’, i.e. for a given vector  $v \in \mathcal{H}$ , one can find many different sequences  $g \in \ell_2$  of coefficients so that

$$v = \sum_{\lambda \in \Lambda} g_\lambda \phi_\lambda . \quad (2.1.1)$$

A few of them have special properties for which they are preferred, e.g. a sequence with minimal  $\ell_2$  norm. The problem of finding sequences  $g$  can be considered as an inverse problem. To this end, let us consider the operator  $F$  (often called the frame operator) that maps a function  $v \in \mathcal{H}$  to the element  $Fv$  of  $\ell_2$  by  $Fv = \{\langle v, \phi_\lambda \rangle_{\mathcal{H}}\}_{\lambda \in \Lambda}$ . The adjoint  $F^*$  maps a sequence  $g \in \ell_2$  to the element  $F^*g$  of  $\mathcal{H}$  via  $F^*g = \sum_{\lambda \in \Lambda} g_\lambda \phi_\lambda$ , i.e. solving (2.1.1) amounts to solving  $F^*g = v$ . In order to show how to solve the last equation and to highlight the relation to standard frame lore, we observe that, for  $v \in \mathcal{H}$ , one has

$$F^*Fv = \sum_{\lambda \in \Lambda} \langle v, \phi_\lambda \rangle_{\mathcal{H}} \phi_\lambda ;$$

for  $g \in \ell_2$ , the sequence  $FF^*g$  is given by

$$(FF^*g)_\eta = \sum_{\lambda \in \Lambda} g_\lambda \langle \phi_\lambda, \phi_\eta \rangle_{\mathcal{H}} .$$

In this context, the sequence  $g$  of minimum  $\ell_2$ -norm satisfying (2.1.1) is given by  $g^\dagger = (F^*)^\dagger v$ . Standard frame concepts suggest  $g^\dagger = F(F^*F)^{-1}v$ , so that  $(F^*)^\dagger = F(F^*F)^{-1}$  in this case. The latter equation holds true since this inverse problem is well-posed: even though  $\mathcal{N}(F^*) \neq \{0\}$ , the operator  $F^*F$  has its spectrum completely within the interval  $[A, B]$ . The spectrum of the operator  $FF^*$  has a gap between the eigenvalue 0 and the remainder of the spectrum, which is contained in  $[A, B]$  (the operator  $FF^*$  becomes only invertible if the frame satisfies special properties, e.g. if  $\{\phi_\lambda\}_{\lambda \in \Lambda}$  forms a Riesz basis for  $\mathcal{H}$ ). The relation between  $F^*F$  and  $FF^*$  is now as follows: since  $F^*F$  is invertible, every  $v \in \mathcal{H}$  has expansions

$$v = \sum_{\lambda \in \Lambda} \langle v, (F^*F)^{-1} \phi_\lambda \rangle_{\mathcal{H}} \phi_\lambda = \sum_{\lambda \in \Lambda} \langle v, \phi_\lambda \rangle_{\mathcal{H}} (F^*F)^{-1} \phi_\lambda .$$

These expansions are only useful if it is possible to calculate  $(F^*F)^{-1}\phi_\lambda$  (the so-called ‘canonical’ dual frame). Often it is convenient (and also more efficient) to employ an iterative reconstruction method, which is usually called the frame algorithm: given a relaxation parameter  $0 < \gamma < 2/B$ , set  $\delta = \max\{|1 - \gamma A|, |1 - \gamma B|\} < 1$ . Let  $v_0 = 0$  and define the iteration

$$v_{m+1} = v_m + \gamma F^* F(v - v_m) ,$$

for which  $\|v - v_m\|_{\mathcal{H}} \leq \delta^m \|v\|_{\mathcal{H}}$ . Let us now rewrite the frame algorithm (based on  $F^*F$ ) by means of the Gram matrix  $(FF^*)_{\lambda,\eta} = \langle \phi_\lambda, \phi_\eta \rangle_{\mathcal{H}}$ . Suppose that  $v_m = \sum_{\lambda \in \Lambda} (g_m)_\lambda \phi_\lambda$  with coefficient sequence  $g_m$ . Then

$$g_{m+1} = g_m + \gamma F(v - F^* g_m) \quad (2.1.2)$$

since the coefficients of  $v_1 = \gamma F^* F v$  are  $\gamma \langle v, \phi_\lambda \rangle_{\mathcal{H}}$  and since

$$F^* F v_m = \sum_{\eta \in \Lambda} \left\langle \sum_{\lambda \in \Lambda} (g_m)_\lambda \phi_\lambda, \phi_\eta \right\rangle \phi_\eta = \sum_{\eta \in \Lambda} (F F^* g_m)_\eta \phi_\eta .$$

Iteration (2.1.2) is nothing than a Landweber iteration, which is a linear regularization scheme (assumed  $0 < \gamma < 2/\|F^*\|^2$ ) and minimizes the discrepancy  $\|v - F^* g\|_{\mathcal{H}}^2$ . Consequently, the iterates of (2.1.2) approximate the minimum  $\ell_2$ -norm sequence  $\{\langle v, (F^*F)^{-1}\phi_\lambda \rangle_{\mathcal{H}}\}_{\lambda \in \Lambda}$ .

It is now often of interest to find sequences that are sparser than the minimum  $\ell_2$ -norm solution. For instance, one may know a priori that  $v$  is a noisy version of a linear combination of  $\phi_\lambda$  with a coefficient sequence with small  $\ell_p$ -norm ( $p = 1$  or, more general,  $1 \leq p \leq 2$ ). In this situation, it makes sense to compute some  $g$  that minimizes

$$\|v - F^* g\|_{\mathcal{H}}^2 + \alpha \|g\|_{\ell_p}^p .$$

For  $p = 2$ , a possible way to approach the minimizer for the last problem is given by damped Landweber iterations (for  $\|F\|^2 < B'$ )

$$g_{m+1} = \frac{1}{B' + \alpha} (B' g_m + F(v - F^* g_m)) ;$$

and for  $p = 1$ , by Landweber iterations with shrinkage operation in each step

$$g_{m+1} = S_{\frac{\alpha}{2B'}} (g_m + (B')^{-1} F(v - F^* g_m)) .$$

## 2.1.2 Linear Problems

We abstain from introducing all the basic facts of linear inverse problems in its full completeness. We restrict ourselves to the facts that are really under consideration here and would rather refer the reader to the abundant literature concerned with this topic, e.g., [EHN96b, Kre89, Lou89] and many more.

As we have seen in the last section, the approximation of the dual frame  $(F^*F)^{-1}\phi_\lambda$  or of the sequence  $\{\langle v, (F^*F)^{-1}\phi_\lambda \rangle_{\mathcal{H}}\}_{\lambda \in \Lambda}$  can be directly related to solving a linear inverse problem in its variational form. However, in many applications, the features or signals

of interest cannot be observed directly, but have to be inferred from other, observable quantities. Very often, there is a linear relationship between the feature modeled by a function  $v$ , and the derived quantities modeled by another function  $z$ , i.e. we can write the problem of inferring  $v$  from  $z$  as

$$Av = z .$$

This equation and the task of solving it makes only sense when everything is placed in an adequate setting. The observations (data), which we shall model by yet another function,  $f$ , are typically not exactly equal to  $z = Av$ , but a distortion of  $z$ . Often the distortion is modeled by an additive noise error term  $e$  (from which one typically assumes that it can be measured by its  $L_2$ -norm),

$$f = z + e = Av + e .$$

Therefore it is customary to take as the image space  $L_2$ ; even if the true images  $z$  lie in a much smaller space. Thus, we shall always assume that  $A$  is a bounded operator from  $\mathcal{H}$  to  $\mathcal{H}'$  (think of  $\mathcal{H}' = L_2$ ). To find an estimate for  $v$  from observed  $f$ , one can minimize the discrepancy

$$\|f - Av\|_{\mathcal{H}'}^2 .$$

The minimizer of the discrepancy is called the pseudo-solution of the inverse problem. If  $A$  has a trivial null-space, the unique minimizer is given by  $(A^*A)^{-1}A^*f$ ; if the null-space is non-trivial, one picks the unique element  $z^\dagger$  of minimum norm. This function is called the generalized solution. Even when  $A^*A$  is not invertible,  $z^\dagger = A^\dagger f$  is well-defined for all  $f$  with  $A^*f \in \mathcal{R}(A^*A)$ . But the generalized inverse may be unbounded, then the problem is ill-posed. In such cases, it has to be replaced by bounded approximants, so that numerically stable solutions can be used as meaningful approximant. This is the goal of regularization. A regularized version in its variational form, the so-called Tikhonov functional, is given by

$$\|f - Av\|_{\mathcal{H}'}^2 + \alpha \|v\|_{\mathcal{H}}^2 .$$

Let us now put the frame concept and regularization theory together, i.e. given an observation  $f$ , the task is to search for a sequence  $g$  of dual frame coefficients for our feature  $v$ . The variational problem of this inverse problems then takes the form

$$\Phi(g) = \|f - AF^*g\|_{\mathcal{H}'}^2 + \alpha \|g\|_{\ell_p^p}^p , \quad (2.1.3)$$

where we allow  $1 \leq p \leq 2$ ; the cases  $p < 2$  promote sparse representations of  $v$  (typically one would pick  $p = 1$ ). This kind of variational problem can be solved by applying the results and methods presented in [DDD04].

### 2.1.3 Smoothness, Sparsity, and other Constraints

In this section, we briefly recall some facts on wavelets and their capabilities for the characterization of smoothness spaces, i.e. for so-called Besov spaces (note that certain scales of norms or semi-norms of Besov spaces can also be seen as sparsity measures). Moreover, we shortly explain how frames may characterizes smoothness spaces.



### Wavelets and Besov Scales

Especially important for our approaches are the smoothness characterization properties of wavelets: one can determine the membership of a function in many different smoothness functional spaces by examining the decay properties of its wavelets coefficients. For a comprehensive introduction and overview on this topic we would refer the reader to the abundant literature, see e.g. [Dau92, Dau93, CDF92, Dah96, DJP92, DJP88, FJ90, Tri78].

Suppose  $H$  is a Hilbert space. Let  $\{V_j\}$  be a sequence of closed nested subspaces of  $H$  whose union is dense in  $H$  while their intersection is zero. In addition,  $V_0$  is shift-invariant and  $f \in V_j \leftrightarrow f(2^j \cdot) \in V_0$ , so that the sequence  $\{V_j\}$  forms a multi-resolution analysis. In many cases of practical relevance the spaces  $V_j$  are spanned by single scale bases  $\Phi_j = \{\phi_{j,k} : k \in I_j\}$  which are uniformly stable. Successively updating a current approximation in  $V_j$  to a better one in  $V_{j+1}$  can be facilitated if stable bases  $\Psi_j = \{\psi_{j,k} : k \in J_j\}$  for some complement  $W_j$  of  $V_j$  in  $V_{j+1}$  are available. Hence, any  $f_n \in V_n$  has an alternative multi-scale representation  $f_n = \sum_{k \in I_0} f_{0,k} \phi_{0,k} + \sum_{j=0}^n \sum_{k \in J_j} f_{j,k} \psi_{j,k}$ . The essential constraint on the choice of  $W_j$  is that  $\Psi = \bigcup_j \Psi_j$  forms a Riesz-basis of  $H$ , i.e. every  $f \in H$  has a unique expansion

$$f = \sum_j \sum_{k \in J_j} \langle f, \tilde{\psi}_{j,k} \rangle \psi_{j,k} \quad \text{such that} \quad \|f\|_H \sim \left( \sum_j \sum_{k \in J_j} |\langle f, \tilde{\psi}_{j,k} \rangle|^2 \right)^{\frac{1}{2}}, \quad (2.1.4)$$

where  $\tilde{\Psi}$  forms a bi-orthogonal system and is in fact also a Riesz-basis for  $H$ , see, e.g., [Dau92].

For our approach we assume that any function (image)  $f \in L_2(I)$  can be extended periodically to all of  $\mathbb{R}^2$ . Here  $I$  is assumed to be the unit square  $(0, 1]^2 = \Omega$ . Throughout this paper we only consider compactly supported tensor product wavelet systems (based on Daubechies' orthogonal wavelets, see [Dau93], or symmetric bi-orthogonal wavelets by Cohen, Daubechies, and Feauveau, see [CDF92]).

We are finally interested in characterizations of Besov spaces, see, e.g., [Tri78]. For  $\beta > 0$  and  $0 < p, q \leq \infty$  the Besov space  $B_q^\beta(L_p(\Omega))$  of order  $\beta$  is the set of functions

$$B_q^\beta(L_p(\Omega)) = \{f \in L_p(\Omega) : |f|_{B_q^\beta(L_p(\Omega))} < \infty\},$$

where  $|f|_{B_q^\beta(L_p(\Omega))} = \left( \int_0^\infty (t^{-\beta} \omega_l(f; t)_p)^q dt / t \right)^{1/q}$  and  $\omega_l$  denotes the  $l$ -th modulus of smoothness,  $l > \beta$ . These spaces are endowed with the norm  $\|f\|_{B_q^\beta(L_p(\Omega))} = \|f\|_{L_p(\Omega)} + |f|_{B_q^\beta(L_p(\Omega))}$ . (For  $p < 1$ , this is not a norm, strictly speaking, and the Besov spaces are complete topological vector spaces but no longer Banach spaces, see [DeV98] for details, including the characterization of these spaces by wavelets.) What is important to us is that one can determine whether a function is in  $B_q^\beta(L_p(\Omega))$  simply by examining its wavelet coefficients. The case  $p = q$ , on which we shall focus, is the easiest. Suppose that  $\phi$  has  $R$  continuous derivatives and  $\psi$  has vanishing moments of order  $M$ . Then, as long as  $\beta < \min(R, M)$ , one has in, two dimensions, for all  $f \in B_p^\beta(L_p(\Omega))$ , the following norm equivalence (denoted by  $\sim$ )

$$|f|_{B_p^\beta(L_p(\Omega))} \sim \left( \sum_{\lambda} 2^{|\lambda|s p} |f_{\lambda}|^p \right)^{1/p} \quad \text{with} \quad f_{\lambda} := \langle f, \tilde{\psi}_{\lambda} \rangle, \quad s = \beta + 1 - 2/p \quad \text{and} \quad |\lambda| = j. \quad (2.1.5)$$

In what follows, we shall always use the equivalent weighted  $\ell_p$ -norm of the  $\{f_\lambda\}$  instead of the standard Besov norm; with a slight abuse of notation we shall continue to denote it by the same symbol, however. When  $p = q = 2$ , the space  $B_2^\beta(L_2(\Omega))$  is the Bessel potential space  $H^\beta(\Omega)$ . In analogy with the special case of Bessel potential spaces  $H^\beta(\Omega)$ , the Besov space  $B_p^\beta(L_p(\Omega))$  with  $\beta < 0$  can be viewed as the dual space of  $B_{p'}^{\beta'}(L_{p'}(\Omega))$ , where  $\beta' = -\beta$  and  $1/p + 1/p' = 1$ .

### Characterizations by Frames

Not only wavelet bases provide reasonable characterizations of function spaces but also frame based characterizations are possible. First basics on frames were introduced in Section 2.1.1. Here we just want to give an idea on how frames might characterize function spaces.

Fundamental developments on modern frame theory can be found in a series of papers [H.G86, HK88, HK89a, HK89b, HK92]. This very aesthetic and subtle theory is essentially based on group theory and is a tool to construct so-called *coorbit spaces* which are defined by collecting all functions for which the associated wavelet transform is contained in some (weighted)  $L_p$ -space (which can be seen as Besov- and Modulation spaces etc). The basic idea of characterizations is that for a judicious discretization of the group representation one may obtain desired frames for these coorbit spaces. Once we have frames for these coorbit spaces at hand, we might pick the associated coefficient sequence space norm in order to add adequate constraints to our variational formulation of the inverse problem. At this point we wish to remark, that frames offer much more freedom in sense that one is no longer restricted to the whole Euclidean plane. In recently published papers, see [DST04a, DST04b], we have extended the group based frame concept to bounded domains and manifolds (e.g. the sphere) which offers especially for certain inverse problems a much better suited representation of the solution. Because of the complexity of the construction process of frames, the related coorbit spaces and sequence space characterizations, we abstain from a detailed review and refer the reader to [DST04a, DST04b]. The essential message is that involving constraints that are typically given by weighted frame coefficient sequence space norms is allowed and naturally suggests the usage of adequate (families of) frames. Adequate frame means here a judiciously discretized family of analyzing atoms.

## 2.2 Multi-Frames and Mixed Constraints

Instead of using one single frame only, we aim now to represent the function we are searching for by means of several frames. This makes sense since for certain classes of signals it often seems that one single frame is not always best suited (in the sense of locally best sparse approximation).

### 2.2.1 Formulation of the Variational Problem

Let  $\mathcal{H}$  and  $\mathcal{H}'$  as before. The suggested multi-frame setting requires the presence of a finite family of frames  $\{\phi_\lambda^i\}_{\lambda \in \Lambda, i \in \mathcal{I}}$  where each individual collection  $\{\phi_\lambda^i\}_{\lambda \in \Lambda}$  ( $i = 1, 2, \dots, n$ ) is

a frame for  $\mathcal{H}$ . For each frame we have the frame operator

$$F_i : \mathcal{H} \rightarrow \ell_2, \text{ via } v \mapsto v^i := \{\langle v, \phi_\lambda^i \rangle\}_{\lambda \in \Lambda_i} . \quad (2.2.1)$$

All the frame operators may now be related by considering the following composition operator

$$K : (\ell_2)^n = \ell_2 \times \dots \times \ell_2 \rightarrow \mathcal{H}, \text{ via } (v^1, \dots, v^n) \mapsto \sum_{i=1}^n F_i^* v^i \quad (2.2.2)$$

or more general, if we additionally involve a linear inverse problem by some bounded linear operator  $A : \mathcal{H} \rightarrow \mathcal{H}'$

$$K_A : (\ell_2)^n = \ell_2 \times \dots \times \ell_2 \rightarrow \mathcal{H}', \text{ via } (v^1, \dots, v^n) \mapsto \sum_{i=1}^n A F_i^* v^i . \quad (2.2.3)$$

For later use have to compute the adjoint and a bound for the operator norm of  $K_A$ .

**Lemma 2.2.1** *The adjoint operator is given by*

$$K_A^* : \mathcal{H}' \mapsto (\ell_2)^n, \text{ via } g \mapsto K_A^* g = (F_1 A^* g, \dots, F_n A^* g) . \quad (2.2.4)$$

Moreover, if we assume that  $\|A\| < \tilde{C}$  and that  $B_i$  denotes the upper frame bound for  $F_i$ , then

$$\|K_A\| < \tilde{C} \sqrt{B_1 + \dots + B_n} . \quad (2.2.5)$$

*Proof.* First, we note that the Hilbert space  $(\ell_2)^n$  is endowed with the scalar

$$\langle g, h \rangle_{(\ell_2)^n} = \langle g^1, h^1 \rangle_{\ell_2} + \dots + \langle g^n, h^n \rangle_{\ell_2}$$

and thus the associated norm is given by

$$\|g\|_{(\ell_2)^n}^2 = \|g^1\|_{\ell_2}^2 + \dots + \|g^n\|_{\ell_2}^2 .$$

For  $f = (f^1, \dots, f^n) \in (\ell_2)^n$  and  $h \in \mathcal{H}'$  the adjoint operator (2.2.4) can now be easily derived:

$$\begin{aligned} \langle K_A f, h \rangle_{\mathcal{H}'} &= \sum_{i=1}^n \langle A F_i^* f^i, h \rangle_{\mathcal{H}'} = \sum_{i=1}^n \langle f^i, F_i A^* h \rangle_{\ell_2} \\ &= \langle f, (F_1 A^* h, \dots, F_n A^* h) \rangle_{(\ell_2)^n} = \langle f, K_A^* h \rangle_{(\ell_2)^n} \end{aligned}$$

and bound (2.2.5) follows then directly by

$$\begin{aligned} \|K_A^* f\|_{(\ell_2)^n}^2 &= \|F_1 A^* f\|_{\ell_2}^2 + \dots + \|F_n A^* f\|_{\ell_2}^2 \\ &\leq \|F_1\|^2 \|A\|^2 \|f\|_{\mathcal{H}}^2 + \dots + \|F_n\|^2 \|A\|^2 \|f\|_{\mathcal{H}}^2 , \end{aligned}$$

i.e.

$$\|K_A\| < \tilde{C} \sqrt{B_1 + \dots + B_n} .$$

□

With this specific operator  $K_A$  we may now formulate the following variational problem

$$\Phi(g) := \|f - K_A g\|_{\mathcal{H}'}^2 + \alpha \cdot |||g|||, \quad (2.2.6)$$

where  $g = (g^1, \dots, g^n)$ ,  $|||g||| := (|g^1|_{p_1, w_1}, \dots, |g^n|_{p_n, w_n})$  with  $|\cdot|_{p_i, w_i}$  denoting a weighted  $\ell_{p_i}$  (semi)-norm, and  $\alpha = (\alpha_1, \dots, \alpha_n)$  represent  $n$  positive regularization parameters. In principle, we restrict ourselves also to  $1 \leq p_i \leq 2$  with not necessarily requiring  $p_i = p_j$ . Thus  $\Phi$  is in principal no longer homogeneous and this complicates the choice of  $\alpha$ .

A strategy for solving this kind of problem for  $n = 1$ ,  $1 \leq p_1 \leq 2$  and  $\{\phi_\lambda\}_{\lambda \in \Lambda}$  being a basis (also concepts for frames) is shown in [DDD04], for  $n = 2$ , concepts are suggested in [DT04, DT05, DD04a]. For our purposes, we will follow the techniques introduced there but the specialty here is that each component  $g^i$  of  $g$  is represented by another frame, i.e. we are searching for an approximation of  $v$  which is a composition of different frames (or different bases).

In what follows, we propose a strategy how compute or to approximate the vector of sequences,  $g$ . First, we show that  $\Phi$  is a convex functional. For general frame systems one cannot expect uniqueness (or strict convexity) since in principal we have  $\ker(F_i^*) \neq \{0\}$ , i.e. even when  $\mathcal{N}(A) = \{0\}$  we have no chance.

**Lemma 2.2.2** *The functional  $\Phi$  is convex.*

*Proof.* Consider  $g = (g^1, \dots, g^n) \in (\ell_2)^n$  and  $h = (h^1, \dots, h^n) \in (\ell_2)^n$  and some  $\tau \in (0, 1)$ . Then,

$$\begin{aligned} \Delta &= \Phi(\tau g + (1 - \tau)h) - \tau\Phi(g) - (1 - \tau)\Phi(h) \\ &= -\tau(1 - \tau)\|K_A g - K_A h\|_{\mathcal{H}'}^2 + \alpha \cdot (|||\tau g + (1 - \tau)h||| - \tau|||g||| - (1 - \tau)|||h|||) . \end{aligned}$$

The second term is non-positive since a Banach (semi) norm is convex, and the first term is also non-positive. Consequently,  $\Delta \leq 0$  and  $\Phi$  is convex.  $\square$

The next step is to construct a surrogate or so-called replacement functional for  $\Phi$  from which we expect a simplification of the minimization process. The overall goal is to avoid the appearance of  $\|K_A g\|_{\mathcal{H}'}^2$  which typically causes a non-linear coupling of all the frame coefficients we aim to compute. Defining a constant  $C := \tilde{C}\sqrt{B_1 + \dots + B_n}$  the standard Gaussian surrogate for the data discrepancy takes the following form

$$\Gamma^{sur}(g; a) = \|f - K_A g\|_{\mathcal{H}'}^2 + C^2 \|g - a\|_{(\ell_2)^n}^2 - \|K_A g - K_A a\|_{\mathcal{H}'}^2$$

for some auxiliary element  $a \in (\ell_2)^n$ .

**Lemma 2.2.3** *The functional  $\Gamma^{sur}(g; a)$  is a proper surrogate for  $\|f - K_A g\|_{\mathcal{H}'}^2$ .*

*Proof.* ‘Proper’ in this context means that the problem remains convex and

$$\Gamma^{sur}(g; a) - \|f - K_A g\|_{\mathcal{H}'}^2 \geq 0 .$$

To this end, consider

$$C^2 \|g - a\|_{(\ell_2)^n}^2 - \|K_A g - K_A a\|_{\mathcal{H}'}^2 = \langle (C^2 - K_A^* K_A)(g - a), g - a \rangle_{(\ell_2)^n} . \quad (2.2.7)$$

Defining  $L := \sqrt{C^2 - K_A^* K_A}$ , we observe for  $\tau \in (0, 1)$  and  $g, h \in (\ell_2)^n$

$$\begin{aligned} \|L(\tau g + (1 - \tau)h - a)\|_{(\ell_2)^n}^2 - \tau \|L(g - a)\|_{(\ell_2)^n}^2 - (1 - \tau) \|L(h - a)\|_{(\ell_2)^n}^2 \\ = -\tau(1 - \tau) \|L(g - h)\|_{(\ell_2)^n}^2 . \end{aligned}$$

Since  $C^2 - K_A^* K_A$  and therewith  $L$  are strictly positive operators, (2.2.7) is strictly convex and positive for  $g \neq a$ .  $\square$

Now we are able to define the global surrogate for  $\Phi$ :

$$\Phi^{sur}(g; a) := \Gamma^{sur}(g, a) + \alpha \cdot \|g\| , \quad (2.2.8)$$

satisfying

$$\Phi^{sur}(g; g) = \Phi(g) , \quad \Phi^{sur}(g; a) \geq \Phi(g) \quad \text{for all } a \in (\ell_2)^n . \quad (2.2.9)$$

The definition of our surrogate functional (2.2.8) suggests the following iteration in order to approach the minimizer  $g^*$  of the initial problem (2.2.6): starting from an arbitrarily chosen  $g_0$ , we determine the minimizer  $g_1$  of (2.2.8) for  $a = g_0$ ; each successive iterate  $g_m$  is then the minimizer for  $g$  of (2.2.8) anchored at the previous iterate  $a = g_{m-1}$ :

$$g_0 \text{ arbitrary ; } g_{m+1} = \arg \min_g \Phi^{sur}(g; g_m) \quad m = 0, 1, \dots \quad (2.2.10)$$

## 2.2.2 Minimization of Surrogate Functionals

The general principles of minimizing (2.2.8) explored in [DDD04] essentially apply here. The difference is that we deal instead with one single frame with  $n$  frames. For sake of illustrating the ideas, we limit ourselves to the case  $p_j = 1$  for  $j = 1, \dots, n$ . The other cases  $1 < p_i \leq 2$  (not necessarily requiring  $p_i = p_j$ ) cause no additional problems and can be treated analogously in the same manner and is therefore left to the reader.

First we discuss the minimization of (2.2.8) for some generic  $a \in (\ell_2)^n$ . The surrogate

functional has the following form

$$\begin{aligned}
\Phi^{sur}(g; a) &= \|f\|_{\mathcal{H}'}^2 - 2\langle g, K_A^* f \rangle_{(\ell_2)^n} + \alpha \cdot |||g||| \\
&\quad C^2 \|g - a\|_{(\ell_2)^n}^2 + 2\langle g, K_A^* K_A a \rangle_{(\ell_2)^n} - \|K_A a\|_{\mathcal{H}'}^2 \\
&= C^2 \|g\|_{(\ell_2)^n}^2 - 2\langle g, K_A^* f + C^2 a - K_A^* K_A a \rangle_{(\ell_2)^n} + \alpha \cdot |||g||| \\
&\quad + \|f\|_{\mathcal{H}'}^2 + C^2 \|a\|_{(\ell_2)^n}^2 - \|K_A a\|_{\mathcal{H}'}^2 \\
&= \sum_{i=1}^n \sum_{\lambda \in \Lambda_i} (C^2 (g_\lambda^i)^2 - 2g_\lambda^i [F_i A^* f + C^2 a^i - F_i A^* K_A a]_\lambda + \alpha_i |g_\lambda^i|) \\
&\quad + \|f\|_{\mathcal{H}'}^2 + C^2 \|a\|_{(\ell_2)^n}^2 - \|K_A a\|_{\mathcal{H}'}^2,
\end{aligned} \tag{2.2.11}$$

where we have used the shorthand  $g_\lambda^i$  for  $\langle g, \phi_\lambda^i \rangle$  (and implicitly assumed that we are dealing with real functions; otherwise one needs to parametrize by modulus and phase). The latter variational equation for the  $g_\lambda^i$  decouple. The summand is differentiable in  $g_\lambda^i$  except at  $g_\lambda^i = 0$ . To overcome this drawback we introduce set-valued derivatives, i.e. we allow  $\text{sign}(0) \in [-1, 1]$ . Then the minimization reduces to solving

$$g_\lambda^i + \frac{\alpha_i}{2C^2} \text{sign}(g_\lambda^i) = C^{-2} [F_i A^* f + C^2 a^i - F_i A^* K_A a]_\lambda. \tag{2.2.12}$$

Denoting the soft-shrinkage operator by  $S_t$  with shrinkage parameter  $t$ , we obtain an explicit expression for the coefficients

$$g_\lambda^i = S_{\frac{\alpha_i}{2C^2}} (C^{-2} [F_i A^* f + C^2 a^i - F_i A^* K_A a]_\lambda). \tag{2.2.13}$$

Let us now introduce with a slight abuse of notation the soft-shrinkage operation for some  $f \in \varepsilon$  acting component-wise

$$S_t(f) = \{S_t(f_\lambda)\}_{\lambda \in \Lambda}.$$

With this shorthand we may introduce the combined shrinkage operator for some vector of sequences  $(f^1, \dots, f^n) \in (\ell_2)^n$  and a multi parameter  $t = (t_1, \dots, t_n)$

$$\mathbf{S}_t(f) = (S_{t_1}(f^1), \dots, S_{t_n}(f^n)).$$

In this setting the minimizer  $g$  for (2.2.8) can be written in the much simpler form

$$g = \mathbf{S}_{\frac{\alpha}{2C^2}} (C^{-2} [K_A^* f + C^2 a - K_A^* K_A a]). \tag{2.2.14}$$

We summarize our findings for the particular case  $p_i = 1$ ,  $w_i = 1$  for  $i = 1, \dots, n$ :

**Proposition 2.2.1** *Suppose the operator  $A$  maps a Hilbert space  $\mathcal{H}$  to another Hilbert space  $\mathcal{H}'$ , with  $\|A\| < \tilde{C}$ , and suppose we are given  $n$  frames where the respective frame operators  $F_i$  map  $\mathcal{H}$  to  $\varepsilon$  with upper frame bounds  $B_i$ , and suppose  $f$  is an element of  $\mathcal{H}'$ .*

Pick  $p_i = 1$ ,  $w_i = 1$  for  $i = 1, \dots, n$ , and  $a \in (\ell_2)^n$ . If  $\Phi^{sur}(g; a)$  is defined as in (2.2.8) on  $(\ell_2)^n$ , then  $\Phi^{sur}(g; a)$  has a unique minimizer in  $(\ell_2)^n$ . This minimizer is given by

$$g = \mathbf{S}_{\frac{\alpha}{2C^2}} \left( C^{-2} [K_A^* f + C^2 a - K_A^* K_A a] \right) . \quad (2.2.15)$$

For all  $h \in (\ell_2)^n$ , one has

$$\Phi^{sur}(g + h; a) \geq \Phi^{sur}(g; a) + C^2 \|h\|_{(\ell_2)^n}^2 .$$

*Proof.* We observe that

$$\Phi^{sur}(g+h; a) - \Phi^{sur}(g; a) = C^2 \|h\|_{(\ell_2)^n}^2 + 2 \langle h, C^2 g - C^2 a - K_A^* (f - K_A) \rangle_{(\ell_2)^n} + \alpha (|||g + h||| - |||g|||) .$$

Defining sets  $\Lambda_i^0 := \{\lambda \in \Lambda_i \mid g_\lambda^i = 0\}$ , and  $\Lambda_i^1 := \Lambda_i \setminus \Lambda_i^0$  and substituting (2.2.12) for  $g_\lambda^i$ , we recast the latter equation

$$\begin{aligned} \Phi^{sur}(g + h; a) - \Phi^{sur}(g; a) &= C^2 \|h\|_{(\ell_2)^n}^2 + \sum_{i=1}^n \sum_{\lambda \in \Lambda_i^0} \{ \alpha_i |h_\lambda^i| - 2h_\lambda^i [F_i A^* f + C^2 a^i - F_i A^* K_A a]_\lambda \} \\ &\quad + \sum_{i=1}^n \sum_{\lambda \in \Lambda_i^1} \alpha_i \{ |g_\lambda^i + h_\lambda^i| - |g_\lambda^i| - h_\lambda^i \text{sign}(g_\lambda^i) \} . \end{aligned}$$

For  $\lambda \in \Lambda_i^0$  we have  $[F_i A^* f + C^2 a^i - F_i A^* K_A a]_\lambda \leq \alpha_i/2$ , so that

$$\alpha_i |h_\lambda^i| - 2h_\lambda^i [F_i A^* f + C^2 a^i - F_i A^* K_A a]_\lambda \geq 0 .$$

For  $\lambda \in \Lambda_i^1$ , we consider two cases: if  $g_\lambda^i > 0$ , then

$$|g_\lambda^i + h_\lambda^i| - |g_\lambda^i| - h_\lambda^i \text{sign}(g_\lambda^i) = |g_\lambda^i + h_\lambda^i| - (g_\lambda^i + h_\lambda^i) \geq 0;$$

if  $g_\lambda^i < 0$ , then

$$|g_\lambda^i + h_\lambda^i| - |g_\lambda^i| - h_\lambda^i \text{sign}(g_\lambda^i) = |g_\lambda^i + h_\lambda^i| + (g_\lambda^i + h_\lambda^i) \geq 0,$$

which proves the assertion. □

Proposition 2.2.1 directly carries over to iteration (2.2.10):

**Corollary 2.2.1** *Make the same assumptions as in Proposition 2.2.1. Pick  $g_0 \in (\ell_2)^n$ , and define the functions  $g_m$  by the algorithm (2.2.10). Then*

$$g_{m+1} = \mathbf{S}_{\frac{\alpha}{2C^2}} \left( C^{-2} [K_A^* f + C^2 g_m - K_A^* K_A g_m] \right) . \quad (2.2.16)$$

### 2.2.3 Convergence Analysis

In this section we consider the convergence of the proposed iteration (2.2.10):

**Theorem 2.2.1** *Suppose the operator  $A$  maps a Hilbert space  $\mathcal{H}$  to another Hilbert space  $\mathcal{H}'$ , with  $\|A\| < \tilde{C}$ , and suppose we are given  $n$  frames where the respective frame operators  $F_i$  map  $\mathcal{H}$  to  $\varepsilon$  with upper frame bounds  $B_i$ , and suppose  $f$  is an element of  $\mathcal{H}'$ . Then the sequence of iterates*

$$g_{m+1} = \mathbf{S}_{\frac{\alpha}{2C^2}} \left( C^{-2} [K_A^* f + C^2 g_m - K_A^* K_A g_m] \right) \quad , \quad m = 1, 2, \dots \quad ,$$

with  $g_0$  arbitrarily chosen in  $(\ell_2)^n$ , converges in norm to a minimizer of the functional

$$\Phi(g) = \|f - K_A g\|_{\mathcal{H}'}^2 + \alpha \cdot \|g\| \quad .$$

First, we prove weak convergence, and we show that the weak limit is a minimizer for  $\Phi$ ; and next, we show that the convergence holds also in norm.

With the following shorthand

$$\mathbf{T}g = \mathbf{S}_{\frac{\alpha}{2C^2}} \left( C^{-2} [K_A^* f + C^2 g - K_A^* K_A g] \right) \quad ,$$

i.e.  $g_m = \mathbf{T}^m g_0$ , we may formulate the weak convergence result as follows:

**Proposition 2.2.2** *The sequence  $\mathbf{T}^m g_0$ ,  $n = 1, 2, \dots$  converges weakly, and its limit is a fixed point for  $\mathbf{T}$ .*

This result can be achieved by applying Opial's Theorem, see [Opi67]:

**Theorem 2.2.2 (Opial)** *Let the mapping  $\mathbf{A}$  from  $\mathcal{H}$  to  $\mathcal{H}$  satisfy the following conditions:*

- i)  $\mathbf{A}$  is non-expansive, i.e. for all  $v, w \in \mathcal{H}$ ,  $\|\mathbf{A}v - \mathbf{A}w\| \leq \|v - w\|$ ,
- ii)  $\mathbf{A}$  is asymptotically regular: for all  $v \in \mathcal{H}$ ,  $\|\mathbf{A}^{n+1}v - \mathbf{A}^n v\| \xrightarrow{n \rightarrow \infty} 0$ ,
- iii) the set  $\mathcal{F}$  of fixed points of  $\mathbf{A}$  in  $\mathcal{H}$  is not empty.

*Then, for all  $v \in \mathcal{H}$ , the sequence  $\{\mathbf{A}^n v\}_{n \in \mathbb{N}}$  converges weakly to a fixed point in  $\mathcal{F}$ .*

In order to prove Proposition 2.2.2, we apply Theorem 2.2.2 to  $\mathbf{T}$ . To this end, we have to verify conditions i), ii) and iii). We do this by the following series of lemmas.

**Lemma 2.2.4** *The operator  $\mathbf{S}$  is non-expansive, i.e. for all  $v, w \in (\ell_2)^n$ ,*

$$\|\mathbf{S}_t(v) - \mathbf{S}_t(w)\|_{(\ell_2)^n} \leq \|v - w\|_{(\ell_2)^n} \quad .$$

*Proof.* The results is obtained by applying the fact that each single shrinkage operator is non-expansive, see, e.g., [DDD04],

$$\begin{aligned} \|\mathbf{S}_t(v) - \mathbf{S}_t(w)\|_{(\ell_2)^n}^2 &= \sum_{i=1}^n \|S_t(v^i) - S_t(w^i)\|_{\varepsilon}^2 = \sum_{i=1}^n \sum_{\lambda \in \Lambda_i} |S_t(v_{\lambda}^i) - S_t(w_{\lambda}^i)|^2 \\ &\leq \sum_{i=1}^n \sum_{\lambda \in \Lambda_i} |v_{\lambda}^i - w_{\lambda}^i|^2 = \|v - w\|_{(\ell_2)^n}^2 \quad . \end{aligned}$$



□

**Lemma 2.2.5** *The mapping  $\mathbf{T}$  is non-expansive, i.e. for all  $v, w \in (\ell_2)^n$*

$$\|\mathbf{T}v - \mathbf{T}w\|_{(\ell_2)^n} \leq \|v - w\|_{(\ell_2)^n}.$$

*Proof.* By Lemma 2.2.4 we have

$$\begin{aligned} \|\mathbf{T}v - \mathbf{T}w\|_{(\ell_2)^n} &\leq \|(I - C^{-2}K_A^*K_A)(v - w)\|_{(\ell_2)^n} \\ &\leq \|v - w\|_{(\ell_2)^n} \end{aligned}$$

since we have chosen  $C$  such that  $\|K_A\| < C$ . □

Hence,  $\mathbf{T}$  satisfies condition *i*) in Theorem 2.2.2. Next, we verify condition *ii*):

**Lemma 2.2.6** *The sequences  $\{\Phi(g_m)\}_{m \in \mathbb{N}}$  and  $\{\Phi^{sur}(g_{m+1}; g_m)\}_{m \in \mathbb{N}}$  are non-increasing.*

*Proof.* By the definition of  $L$  we have

$$\Phi(g_{m+1}) + \|L(g_{m+1} - g_m)\|_{(\ell_2)^n}^2 = \Phi^{sur}(g_{m+1}; g_m) \leq \Phi^{sur}(g_m; g_m) = \Phi(g_m)$$

and

$$\Phi^{sur}(g_{m+2}; g_{m+1}) \leq \Phi(g_{m+1}) \leq \Phi(g_{m+1}) + \|L(g_{m+1} - g_m)\|_{(\ell_2)^n}^2 = \Phi^{sur}(g_{m+1}; g_m).$$

□

**Lemma 2.2.7** *The series  $\sum_{m=0}^{\infty} \|g_{m+1} - g_m\|_{(\ell_2)^n}^2$  is convergent.*

*Proof.* Since  $L$  is a strictly positive operator, we have

$$\sum_{m=0}^N \|g_{m+1} - g_m\|_{(\ell_2)^n}^2 \leq \frac{1}{M} \sum_{m=0}^N \|L(g_{m+1} - g_m)\|_{(\ell_2)^n}^2$$

where  $M$  is a strictly lower bound for  $L^*L$ . By Lemma 2.2.6,

$$\sum_{m=0}^N \|L(g_{m+1} - g_m)\|_{(\ell_2)^n}^2 \leq \sum_{m=0}^N (\Phi(g_m) - \Phi(g_{m+1})) \leq \Phi(g_0),$$

regardless of the choice of  $N \in \mathbb{N}$  and the infinite series converges. □

Consequently, we have that

**Lemma 2.2.8** *The mapping  $\mathbf{T}$  is asymptotically regular, i.e.*

$$\|\mathbf{T}^{m+1}g_0 - \mathbf{T}^m g_0\|_{(\ell_2)^n} \rightarrow 0 \quad \text{for } n \rightarrow \infty.$$

We finalize the proof of Theorem 2.2.2 with verifying condition *iii*):

**Lemma 2.2.9** *The  $\|g_m\|_{(\ell_2)^n}$  are bounded uniformly in  $n$ .*

*Proof.* By Lemma 2.2.6 and since  $\alpha_i > 0$  for  $i = 1, \dots, n$ , we have

$$\alpha \cdot |||g_m||| \leq \Phi(g_m) \leq \Phi(g_0), \text{ i.e. } \left| |||g_m||| \right|_{\ell_1} \leq \frac{1}{\min_i \alpha_i} \Phi(g_0).$$

Hence the  $g_m$  are uniformly bounded. Moreover, since

$$\begin{aligned} \|g_m\|_{(\ell_2)^n}^2 &= \sum_{i=1}^n \|g_m^i\|_{\varepsilon}^2 = \sum_{i=1}^n \sum_{\lambda \in \Lambda_i} |(g_m^i)_{\lambda}|^2 = \sum_{i=1}^n \max_{\lambda \in \Lambda_i} |(g_m^i)_{\lambda}| \cdot |g_m^i|_{1,1} \\ &= \left( \max_{\lambda \in \Lambda_1} |(g_m^1)_{\lambda}|, \dots, \max_{\lambda \in \Lambda_n} |(g_m^n)_{\lambda}| \right) \cdot |||g_m||| \\ &\leq |||g_m||| \cdot |||g_m||| \leq \left| |||g_m||| \right|_{\ell_1}^2, \end{aligned}$$

we also have a uniform bound on the  $\|g_m\|_{(\ell_2)^n}^2$ .  $\square$

**Lemma 2.2.10** *Suppose the mapping  $\mathbf{A}$  from  $\mathcal{H}$  to  $\mathcal{H}$  satisfies the conditions *i*) and *ii*) in Theorem 2.2.2. Then, if a subsequence of  $\{\mathbf{A}^n v\}_{n \in \mathbb{N}}$  converges weakly in  $\mathcal{H}$ , then its limit is a fixed point of  $\mathbf{A}$ .*

**Lemma 2.2.11** *The set of fixed points of  $\mathbf{T}$  is not empty.*

*Proof.* By Lemma 2.2.9, the  $\mathbf{T}^m g_0$  are uniformly bounded in  $m$ . By the Banach-Alaoglu Theorem, the sequence has a weak accumulation point. By Lemma 2.2.10, this weak accumulation point is a fixed point for  $\mathbf{T}$  and consequently, the set of fixed points of  $\mathbf{T}$  is not empty.  $\square$

Finally, by Lemmas 2.2.5, 2.2.8, and 2.2.11, we have shown Theorem 2.2.2. Moreover, we can show that this fixed point is also a minimizer for  $\Phi$ :

**Proposition 2.2.3** *A fixed point for  $\mathbf{T}$  is a minimizer for the functional  $\Phi$ .*

*Proof.* If  $g_{\star} = \mathbf{T}g_{\star}$ , then by Proposition 2.2.1, we have that  $g_{\star}$  is a minimizer for  $\Phi^{sur}(g; g_{\star})$ . Moreover, for all  $h \in (\ell_2)^n$ ,

$$\Phi^{sur}(g_{\star} + h; g_{\star}) \geq \Phi^{sur}(g_{\star}; g_{\star}) + C^2 \|h\|_{(\ell_2)^n}^2.$$

With  $\Phi^{sur}(g_{\star}; g_{\star}) = \Phi(g_{\star})$  and  $\Phi^{sur}(g_{\star} + h; g_{\star}) = \Phi(g_{\star} + h) + C^2 \|h\|_{(\ell_2)^n}^2 - \|K_A h\|_{\mathcal{H}'}^2$ , we deduce that, for all  $h \in (\ell_2)^n$ ,

$$\Phi(g_{\star} + h) \geq \Phi(g_{\star}) + \|K_A h\|_{\mathcal{H}'}^2,$$

which proves that  $g_\star$  is also a minimizer for  $\Phi$ .  $\square$

Next, we shall prove that the convergence of  $\{g_m\}_{m \in \mathbb{N}}$  holds also in the Hilbert space norm  $\|\cdot\|_{(\ell_2)^n}$ . Let us introduce the following shorthands

$$g_\star = w - \lim_{m \rightarrow \infty} g_m, \quad u_m = g_m - g_\star, \quad h = g_\star + C^{-2}K_A^*(f - K_A g_\star). \quad (2.2.17)$$

**Lemma 2.2.12**  $\|K_A u_m\|_{(\ell_2)^n} \rightarrow 0$  for  $m \rightarrow \infty$ .

*Proof.* First, observe that with

$$\begin{aligned} u_{m+1} &= g_{m+1} - g_\star = \mathbf{S}(g_m + C^{-2}K_A^*(f - K_A g_m)) - \mathbf{S}(h) \\ &= \mathbf{S}(h + (I - C^{-2})u_m) - \mathbf{S}(h) \end{aligned}$$

one has

$$u_{m+1} - u_m = \mathbf{S}(h + (I - C^{-2}K_A^*K_A)u_m) - \mathbf{S}(h) - u_m$$

and since  $\|u_{m+1} - u_m\|_{(\ell_2)^n} = \|g_{m+1} - g_m\|_{(\ell_2)^n} \rightarrow 0$  for  $n \rightarrow \infty$  (by Lemma 2.2.8), we have

$$\|\mathbf{S}(h + (I - C^{-2}K_A^*K_A)u_m) - \mathbf{S}(h) - u_m\|_{(\ell_2)^n} \rightarrow 0 \quad \text{for } n \rightarrow \infty. \quad (2.2.18)$$

By triangle inequality,

$$\left| \|u_m\|_{(\ell_2)^n} - \|\mathbf{S}(h + (I - C^{-2}K_A^*K_A)u_m) - \mathbf{S}(h)\|_{(\ell_2)^n} \right| \rightarrow 0 \quad \text{for } n \rightarrow \infty. \quad (2.2.19)$$

By Lemma 2.2.4, we have

$$\|\mathbf{S}(h + (I - C^{-2}K_A^*K_A)u_m) - \mathbf{S}(h)\|_{(\ell_2)^n} \leq \|(I - C^{-2}K_A^*K_A)u_m\|_{(\ell_2)^n} \leq \|u_m\|_{(\ell_2)^n}$$

and thus the modulus in (2.2.19) can be dropped, which implies

$$\|u_m\|_{(\ell_2)^n} - \|(I - C^{-2}K_A^*K_A)u_m\|_{(\ell_2)^n} \rightarrow 0 \quad \text{for } n \rightarrow \infty. \quad (2.2.20)$$

Since  $\|u_m\|_{(\ell_2)^n} + \|(I - C^{-2}K_A^*K_A)u_m\|_{(\ell_2)^n} \leq 2\|g_m - g_\star\|_{(\ell_2)^n} \leq 2(\|g_\star\|_{(\ell_2)^n} + \sup_m \|g_m\|_{(\ell_2)^n}) = \tau$  where  $\tau$  is finite by Lemma 2.2.9, we obtain by (2.2.20)

$$\begin{aligned} 0 &\leq \|u_m\|_{(\ell_2)^n}^2 - \|(I - C^{-2}K_A^*K_A)u_m\|_{(\ell_2)^n}^2 \\ &\leq \tau (\|u_m\|_{(\ell_2)^n} - \|(I - C^{-2}K_A^*K_A)u_m\|_{(\ell_2)^n}) \rightarrow 0 \quad \text{for } n \rightarrow \infty. \end{aligned}$$

The inequality

$$\begin{aligned} \|u_m\|_{(\ell_2)^n}^2 - \|(I - C^{-2}K_A^*K_A)u_m\|_{(\ell_2)^n}^2 &= 2C^{-2}\|K_A u_m\|_{(\ell_2)^n}^2 - \|C^{-2}K_A^*K_A u_m\|_{(\ell_2)^n}^2 \\ &\geq C^{-2}\|K_A u_m\|_{(\ell_2)^n}^2 \end{aligned}$$

finally proves the assertion.  $\square$

**Lemma 2.2.13**  $\|\mathbf{S}(h + u_m) - \mathbf{S}(h) - u_m\|_{(\ell_2)^n} \rightarrow 0$  for  $n \rightarrow \infty$ .

*Proof.*

$$\begin{aligned} \|\mathbf{S}(h + u_m) - \mathbf{S}(h) - u_m\|_{(\ell_2)^n} &\leq \|\mathbf{S}(h + (I - C^{-2}K_A^*K_A)u_m) - \mathbf{S}(h) - u_m\|_{(\ell_2)^n} \\ &\quad + \|\mathbf{S}(h + u_m) - \mathbf{S}(h + (I - C^{-2}K_A^*K_A)u_m)\|_{(\ell_2)^n} \\ &\leq \|\mathbf{S}(h + (I - C^{-2}K_A^*K_A)u_m) - \mathbf{S}(h) - u_m\|_{(\ell_2)^n} \\ &\quad + \|C^{-2}K_A^*K_A u_m\|_{(\ell_2)^n} . \end{aligned}$$

The assertion follows because of Lemma 2.2.12 and (2.2.18).  $\square$

The next lemma establishes norm convergence.

**Lemma 2.2.14** *If for some  $h \in (\ell_2)^n$ , and some sequence  $\{w_m\}_{m \in \mathbb{N}}$  with  $w = \lim_{m \rightarrow \infty} w_m = 0$  and  $\lim_{m \rightarrow \infty} \|\mathbf{S}(h + w_m) - \mathbf{S}(h) - w_m\|_{(\ell_2)^n} = 0$  then  $\|w_m\|_{(\ell_2)^n} \rightarrow 0$  for  $m \rightarrow \infty$ .*

*Proof.* First, note again that

$$\|w_m\|_{(\ell_2)^n}^2 = \sum_{i=1}^n \sum_{\lambda \in \Lambda_i} |(w_m^i)_\lambda|^2 .$$

For each index  $i$  we define finite sets  $\Lambda_i^0 \subset \Lambda_i$ , so that  $\sum_{\lambda \in \Lambda_i \setminus \Lambda_i^0} |h_\lambda^i|^2 \leq \left(\frac{\alpha_i}{4C^2}\right)^2$ . Since each  $\Lambda_i^0$  is finite, we have by the weak convergence of the  $w_m$  that

$$\sum_{\lambda \in \Lambda_i^0} |(w_m^i)_\lambda|^2 \rightarrow 0 \text{ for } m \rightarrow \infty$$

holds for every  $i$ , thus

$$\sum_{i=1}^n \sum_{\lambda \in \Lambda_i^0} |(w_m^i)_\lambda|^2 \rightarrow 0 \text{ for } m \rightarrow \infty .$$

Let us now focus on the remaining sums  $\sum_{\lambda \in \Lambda_i \setminus \Lambda_i^0} |(w_m^i)_\lambda|^2$ : for each  $i$  and each  $m$  we split  $\Lambda_i^1 = \Lambda_i \setminus \Lambda_i^0$  into two subsets:

$$\Lambda_i^{1,m} = \{\lambda \in \Lambda_i^1 : |(w_m^i)_\lambda + h_\lambda^i| < \alpha_i/2C^2\} \text{ and } \tilde{\Lambda}_i^{1,m} = \Lambda_i^1 \setminus \Lambda_i^{1,m} .$$

If  $\lambda \in \Lambda_i^{1,m}$ , then  $S_{\frac{\alpha_i}{2C^2}}((w_m^i)_\lambda + h_\lambda^i) = S_{\frac{\alpha_i}{2C^2}}(h_\lambda^i) = 0$  such that

$$|(w_m^i)_\lambda - S_{\frac{\alpha_i}{2C^2}}((w_m^i)_\lambda + h_\lambda^i) + S_{\frac{\alpha_i}{2C^2}}(h_\lambda^i)| = |(w_m^i)_\lambda|$$

and thus

$$\begin{aligned} \sum_{i=1}^n \sum_{\lambda \in \Lambda_i^{1,m}} |(w_m^i)_\lambda|^2 &= \sum_{i=1}^n \sum_{\lambda \in \Lambda_i^{1,m}} |(w_m^i)_\lambda - S_{\frac{\alpha_i}{2C^2}}((w_m^i)_\lambda + h_\lambda^i) + S_{\frac{\alpha_i}{2C^2}}(h_\lambda^i)|^2 \\ &\leq \sum_{i=1}^n \sum_{\lambda \in \Lambda_i} |(w_m^i)_\lambda - S_{\frac{\alpha_i}{2C^2}}((w_m^i)_\lambda + h_\lambda^i) + S_{\frac{\alpha_i}{2C^2}}(h_\lambda^i)|^2 \\ &= \|\mathbf{S}(h + u_m) - \mathbf{S}(h) - u_m\|_{(\ell_2)^n}^2 \rightarrow 0 \text{ for } m \rightarrow \infty . \end{aligned} \quad (2.2.21)$$

It remains to consider the case  $\lambda \in \tilde{\Lambda}_i^{1,m}$ . In this situation,  $|(w_m^i)_\lambda + h_\lambda^i| \geq \alpha_i/2C^2$ , and hence we have for all  $i$  and  $m$

$$|(w_m^i)_\lambda| \geq |(w_m^i)_\lambda + h_\lambda^i| - |h_\lambda^i| \geq \alpha_i/2C^2 - \alpha_i/4C^2 = \alpha_i/4C^2 \geq |h_\lambda^i| ,$$

implying that  $(w_m^i)_\lambda + h_\lambda^i$  and  $h_\lambda^i$  have the same sign and it follows that for  $\lambda \in \tilde{\Lambda}_i^{1,m}$

$$\begin{aligned} |(w_m^i)_\lambda - S_{\frac{\alpha_i}{2C^2}}((w_m^i)_\lambda + h_\lambda^i) + S_{\frac{\alpha_i}{2C^2}}(h_\lambda^i)| &= |(w_m^i)_\lambda - S_{\frac{\alpha_i}{2C^2}}((w_m^i)_\lambda + h_\lambda^i)| \\ &= |(w_m^i)_\lambda - ((w_m^i)_\lambda + h_\lambda^i) + \alpha_i/2C^2 \text{sign}((w_m^i)_\lambda)| \\ &\geq \alpha_i/2C^2 - |h_\lambda^i| \geq \alpha_i/4C^2 . \end{aligned}$$

Consequently, for each individual  $i$  and  $m$  we deduce,

$$\sum_{\lambda \in \tilde{\Lambda}_i^{1,m}} |(w_m^i)_\lambda - S_{\frac{\alpha_i}{2C^2}}((w_m^i)_\lambda + h_\lambda^i) + S_{\frac{\alpha_i}{2C^2}}(h_\lambda^i)|^2 \geq (\alpha_i/4C^2)^2 |\tilde{\Lambda}_i^{1,m}| ,$$

but since  $\|\mathbf{S}(h + u_m) - \mathbf{S}(h) - u_m\|_{(\ell_2)^n}^2 \rightarrow 0$  for  $m \rightarrow \infty$ , necessarily implying

$$\|S_{\frac{\alpha_i}{2C^2}}(h^i + u_m^i) - S_{\frac{\alpha_i}{2C^2}}(h^i) - u_m^i\|_\varepsilon^2 \rightarrow 0 \quad \text{for } m \rightarrow \infty ,$$

there exists an index  $m^*$  uniform in  $i$  so that for  $m > m^*$ ,

$$\sum_{\lambda \in \tilde{\Lambda}_i^{1,m}} |(w_m^i)_\lambda - S_{\frac{\alpha_i}{2C^2}}((w_m^i)_\lambda + h_\lambda^i) + S_{\frac{\alpha_i}{2C^2}}(h_\lambda^i)|^2 < (\alpha_i/4C^2)^2 ,$$

which implies that for all  $i$  the sets  $\tilde{\Lambda}_i^{1,m}$  are empty when  $m > m^*$ . Consequently,

$$\sum_{i=1}^n \sum_{\lambda \in \tilde{\Lambda}_i^{1,m}} |(w_m^i)_\lambda|^2 = 0 \quad \text{for } m > m^* ,$$

which completes the proof since we have

$$\|w_m\|_{(\ell_2)^n}^2 = \sum_{i=1}^n \left( \sum_{\lambda \in \Lambda_i^0} |(w_m^i)_\lambda|^2 + \sum_{\lambda \in \Lambda_i^1} |(w_m^i)_\lambda|^2 + \sum_{\lambda \in \tilde{\Lambda}_i^{1,m}} |(w_m^i)_\lambda|^2 \right) \rightarrow 0 \quad \text{for } m \rightarrow \infty .$$

□

Setting  $w_m = u_m$  and  $h$  and  $u_m$  as in (2.2.17), we have shown that

$$\|g_m - g_\star\|_{(\ell_2)^n} \rightarrow 0 \quad \text{for } m \rightarrow \infty .$$

Moreover, setting  $v_m = \sum_{i=1}^n (g_m^i)_\lambda \phi_\lambda^i$  and  $v_\star = \sum_{i=1}^n (g_\star^i)_\lambda \phi_\lambda^i$ , the estimate

$$\|v_m - v_\star\|_{\mathcal{H}} \leq \left( n \max_i \{B_i\} \right)^{1/2} \|g_m - g_\star\|_{(\ell_2)^n}$$

ensures that the convergence holds also in the  $\mathcal{H}$ -norm topology.

### 2.2.4 Regularization Properties

The choice of the damping/threshold vector  $\alpha$  may be of course chosen by the user. But in a more general context, this parameter can be considered as compression or regularization parameter. In terms of compression, this vector controls the sparsity to be attained and therewith the approximation quality. But in a typical situation of an inverse problem, considering some nontrivial operator  $A$  and having only noisy data at hand, the vector  $\alpha$  plays the most important role in computing stabilized solutions. In this case, if the ‘error’  $e = f - K_A g_o$  tends to zero, we wish our estimate for the solution of the inverse problem tend to  $g_o$ , since the minimizer of  $\Phi(g)$  differs from  $g_o$  for  $\alpha \neq 0$ . In inverse problems lore, this means to identify a functional relation between  $\alpha$  and the noise floor  $\delta$ , i.e.  $\alpha = \alpha(\delta)$  with  $\alpha(\delta) \rightarrow 0$  and  $\|g^{*,\alpha(\delta)} - g_o\| \rightarrow 0$  as  $\alpha \rightarrow 0$ . If we find a parameter rule achieving this, then the suggested iteration scheme will regularize the ill-posed problem. In [DDD04], a regularization theorem is provided for the univariate case, i.e. for  $n = 1$  and, for reasons of simplicity, for the unique situation, i.e. for  $1 < p \leq 2$  or  $\mathcal{N}(A) = \{0\}$ . However, in our context we always have to face the fact that  $\mathcal{N}(K_A)$  is nontrivial as long as we deal with frames, i.e. even if  $\mathcal{N}(A) = \{0\}$ . Thus it is only reasonable to show that we approach one solution  $g^\dagger$  when passing to the limit  $\delta \rightarrow 0$ . Moreover, we limit the analysis to the homogeneous case  $p_i = 1$  ( $i = 1, \dots, n$ ). For a non-homogeneous mixing of penalties the analysis requires a slightly different analysis for proving that one really approaches solutions with minimal penalty value.

**Theorem 2.2.1** *Let  $p_i = 1$  ( $i = 1, \dots, n$ ), and  $f \in \mathcal{H}'$  with  $\|f - z\|_{\mathcal{H}'} \leq \delta$ ,  $\alpha_{\min}(\delta) = \min_j \{\alpha_j(\delta)\}$ ,  $\alpha_{\max}(\delta) = \max_j \{\alpha_j(\delta)\}$ , and assume  $\alpha(\delta) = (\alpha_1(\delta), \dots, \alpha_n(\delta))$  is chosen such that*

$$\alpha(\delta) \xrightarrow{\delta \rightarrow 0} 0, \quad \delta^2 / \alpha_{\min}(\delta) \xrightarrow{\delta \rightarrow 0} 0, \quad \alpha_{\max}(\delta) / \alpha_{\min}(\delta) \xrightarrow{\delta \rightarrow 0} 1.$$

*Then every sequence  $\{g^{*,\alpha(\delta)}\}$  of minimizers of the functional  $\Phi(g)$  where  $\delta \rightarrow 0$  and  $\alpha = \alpha(\delta)$  has a convergent subsequence. The limit of every convergent subsequence is a solution of  $K_A g = z$  with minimal values of  $\left| \left| \left| \cdot \right| \right| \right|_{\ell_1}$ .*

*Proof.* As  $g^{*,\alpha(\delta)} = ((g^{*,\alpha(\delta)})^1, \dots, (g^{*,\alpha(\delta)})^n)$  is a minimizer of  $\Phi$ , we have

$$\|f - K_A g^{*,\alpha(\delta)}\|_{\mathcal{H}'}^2 + \alpha \cdot \left| \left| \left| g^{*,\alpha(\delta)} \right| \right| \right| \leq \delta^2 + \alpha \cdot \left| \left| \left| g^\dagger \right| \right| \right|. \quad (2.2.22)$$

Thus, by the made assumptions on  $\alpha(\delta)$ , we achieve

$$\lim_{\delta \rightarrow 0} K_A g^{*,\alpha(\delta)} = z.$$

Again by (2.2.22),

$$\left| \left| \left| g^{*,\alpha(\delta)} \right| \right| \right|_{\ell_1} \leq \frac{\delta^2}{\alpha_{\min}(\delta)} + \frac{\alpha_{\max}(\delta)}{\alpha_{\min}(\delta)} \left| \left| \left| g^\dagger \right| \right| \right|_{\ell_1}$$

implying,

$$\limsup_{\delta \rightarrow 0} \|g\|_{(\ell_2)^n} \leq \limsup_{\delta \rightarrow 0} \left| \left| \left| g^{*,\alpha(\delta)} \right| \right| \right|_{\ell_1} \leq \left| \left| \left| g^\dagger \right| \right| \right|_{\ell_1},$$

i.e.  $\|g^{*,\alpha(\delta)}\|_{(\ell_2)^n}$  are uniformly bounded. Consequently, the sequence has a weakly convergent subsequence (again denoted by  $\{g^{*,\alpha(\delta)}\}$ ) with weak limit  $g^\circ$ ,

$$g^\circ = w - \lim_{\delta \rightarrow 0} g^{*,\alpha(\delta)} .$$

Assume now  $g^\dagger$  is a solution of the inverse problem with minimal values of  $\left| \left| \left| \cdot \right| \right| \right|_{\ell_1}$ . Because  $g^{*,\alpha(\delta)}$  converge weakly to  $g^\circ$ , for all  $\lambda \in \Lambda_i$  ( $i = 1, \dots, n$ ),

$$(g^{*,\alpha(\delta)})_\lambda^i \rightarrow (g^\circ)_\lambda^i \text{ as } \delta \rightarrow 0 ,$$

we may use Fatou's lemma to obtain

$$\left| \left| \left| g^\circ \right| \right| \right|_{\ell_1} \leq \limsup_{\delta \rightarrow 0} \left| \left| \left| g^{*,\alpha(\delta)} \right| \right| \right|_{\ell_1}$$

and thus

$$\limsup_{\delta \rightarrow 0} \left| \left| \left| g^{*,\alpha(\delta)} \right| \right| \right|_{\ell_1} \leq \left| \left| \left| g^\dagger \right| \right| \right|_{\ell_1} \leq \left| \left| \left| g^\circ \right| \right| \right|_{\ell_1} .$$

Hence  $g^\circ$  is also a solution with minimal values of  $\left| \left| \left| \cdot \right| \right| \right|_{\ell_1}$ . □





# Chapter 3

## Applications I: Linear Problems

### 3.1 Compression for Audio and Image Signals

This section is devoted to show the usefulness of the proposed multi-frame scheme. We present a little number of numerical experiments from different perspectives: convergence, sparsity, approximation quality, and applicability to synthetic data, real audio data, and images.

The overall configuration of our algorithm is as follows: for sake of simplicity, we pick as our underlying frames a wavelet basis (Haar system) and a (non-local) Fourier basis only. Hence,  $B_1 = B_2 = 1$ . In the examples, we restrict ourselves to  $A = I$ . Consequently, the constant  $C$  in our Gaussian surrogate is not allowed to be equal or smaller than  $\sqrt{2}$ . We aim to achieve sparsity in both representations, i.e. we set  $p_1 = p_2 = 1$ ; moreover, we do not involve additional penalty weight sequences, i.e.  $w_1 = w_2 = 1$ . The variational problem is thus simply given by

$$\Phi(g^1, g^2) = \|f - (F_1^* g^1 + F_2^* g^2)\|^2 + \alpha_1 \|g^1\|_{\ell_1} + \alpha_2 \|g^2\|_{\ell_1} ,$$

and the minimization by Gaussian surrogates yields the following iteration

$$\begin{pmatrix} (g^1)_{m+1} \\ (g^2)_{m+1} \end{pmatrix} = \begin{pmatrix} S_{\alpha_1/2C^2} (C^{-2}\{F_1 f + C^2(g^1)_m - F_1 F_1^*(g^1)_m - F_1 F_2^*(g^2)_m\}) \\ S_{\alpha_2/2C^2} (C^{-2}\{F_2 f + C^2(g^1)_m - F_2 F_1^*(g^1)_m - F_2 F_2^*(g^2)_m\}) \end{pmatrix} .$$

Since we deal with bases only, the application of  $F_1 F_1^*$ ,  $F_1 F_2^*$ ,  $F_2 F_1^*$ , and  $F_2 F_2^*$  simplifies to the discrete decomposition and reconstruction schemes. If one really goes beyond bases, i.e. using frames, one indeed has to compute (approximate) all the (mixed) gram matrices. This might be of course costly but can be optimized by picking localized and reasonably incoherent frames. In case the frames are not reasonably incoherent, the scheme amounts to averaging over all the components and then, all the sequences  $g^i$  contain very similar informations.

#### 3.1.1 Application to Audio Coding

Let us now denote by  $\{\phi_\lambda^1\}$  the Fourier system and with  $\{\phi_\lambda^2\}$  the wavelet system. Then we consider two frame operators  $F_1 : v \mapsto \{\langle v, \phi_\lambda^1 \rangle\} = g^1$  and  $F_2 : v \mapsto \{\langle v, \phi_\lambda^2 \rangle\} = g^2$ .

*A synthetic Example:* In this example we have simulated a signal  $f$  that is a composition of two different components: a harmonic wave and noisy perturbation within the interval  $[350, 400]$ . As a sampled discrete vector it has a total number of 641 coefficients in the time-domain representation. This discrete vector is used as input for our algorithm. The results for  $\alpha_1 = \alpha_2 = 0.2$  are visualized in Figure 3.1. We find that involving the Haar wavelet basis and the Fourier basis splits the signal in very sparse and well separated components. The sparsity evolution graph shows the rapid decay of the number of Fourier coefficients which is by the optimal matching of Fourier basis atoms.

*Example “Glockenspiel”:* This data set represents a real audio signal consisting of tonal components and a sequence of (bell) attacks. We again try to apply Haar wavelet and Fourier splitting. For  $\alpha_1 = 0.02$  and  $\alpha_2 = 0.01$  the results are shown in Figure 3.2. As expected, the Haar system captures all the bell attacks very well, and, moreover, the Fourier system the tonal components. We admit that sparsity could be obviously improved just by taking local Fourier systems (as done in [MT05]). The sparsity evolution graph shows the rapid decay of the number of wavelet coefficients which can be explained by a fast “bell attacks” localization process through the iteration.

We summarize, whenever the dictionary consists of complementary frames, the proposed algorithm produces a sparse representation in which the individual components overlap inconsiderably.

### 3.1.2 Application to Image Restoration and Compression

In what follows we provide evidence that the machinery can naturally be applied for image restoration and compression tasks. We consider as test data “Part of woman image” and verify that non-optimally chosen families of frames may achieve the same reconstruction results but that the sparsity gets essentially worse (what is typically expected).

*Example Wavelet-Wavelet:* In this case, see Figure 3.3 ( $\alpha_1 = 10$ ,  $\alpha_2 = 20$ ), we have picked a Daubechies-6-wavelet and the Haar wavelet system. Both bases capture local structures but of different smoothness. In order to provide a comparison of sparsity with the next example, the parameters  $\alpha_1$  and  $\alpha_2$  are chosen such that similar SNR’s and relative approximation errors are achieved.

*Example Wavelet-Fourier:* This example is the same as the latter one except that we have exchanged the Daubechies-6-wavelet system with the Fourier system (here for reasons just explained we have chosen  $\alpha_1 = 15$  and  $\alpha_2 = 13$ ). As we may observe in Figure 3.4 (and since the Haar and the Fourier system are complementary), we achieve much better sparsity as before. The complementary selectivity of wavelets and harmonics manifests here very visible when splitting into local jumps and oscillatory components.

We finally conclude that even for image restoration/compression tasks (here only denoising is illustrated, but deblurring – or more generally: inverting operator equations – is by construction possible) the proposed method has demonstrated its capabilities.

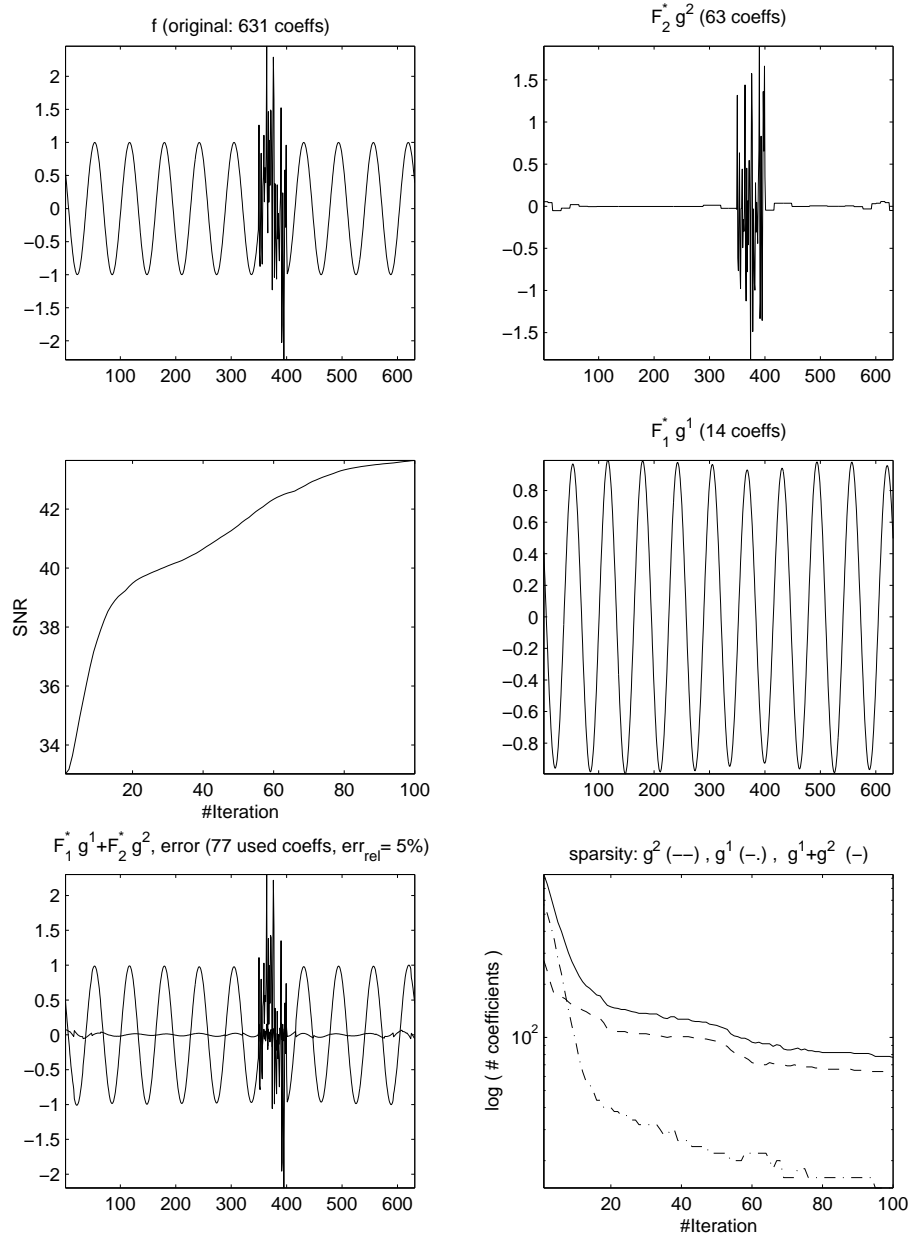


Figure 3.1: From top left to up right: synthetic data, Haar wavelet component ( $g^2$ ) (in time domain) after 100 iterations, SNR evolution through the iteration process, Fourier component ( $g^1$ ) (in time domain) after 100 iterations, reconstruction and error after 100 iterations, and sparsity evolution through the iteration process.

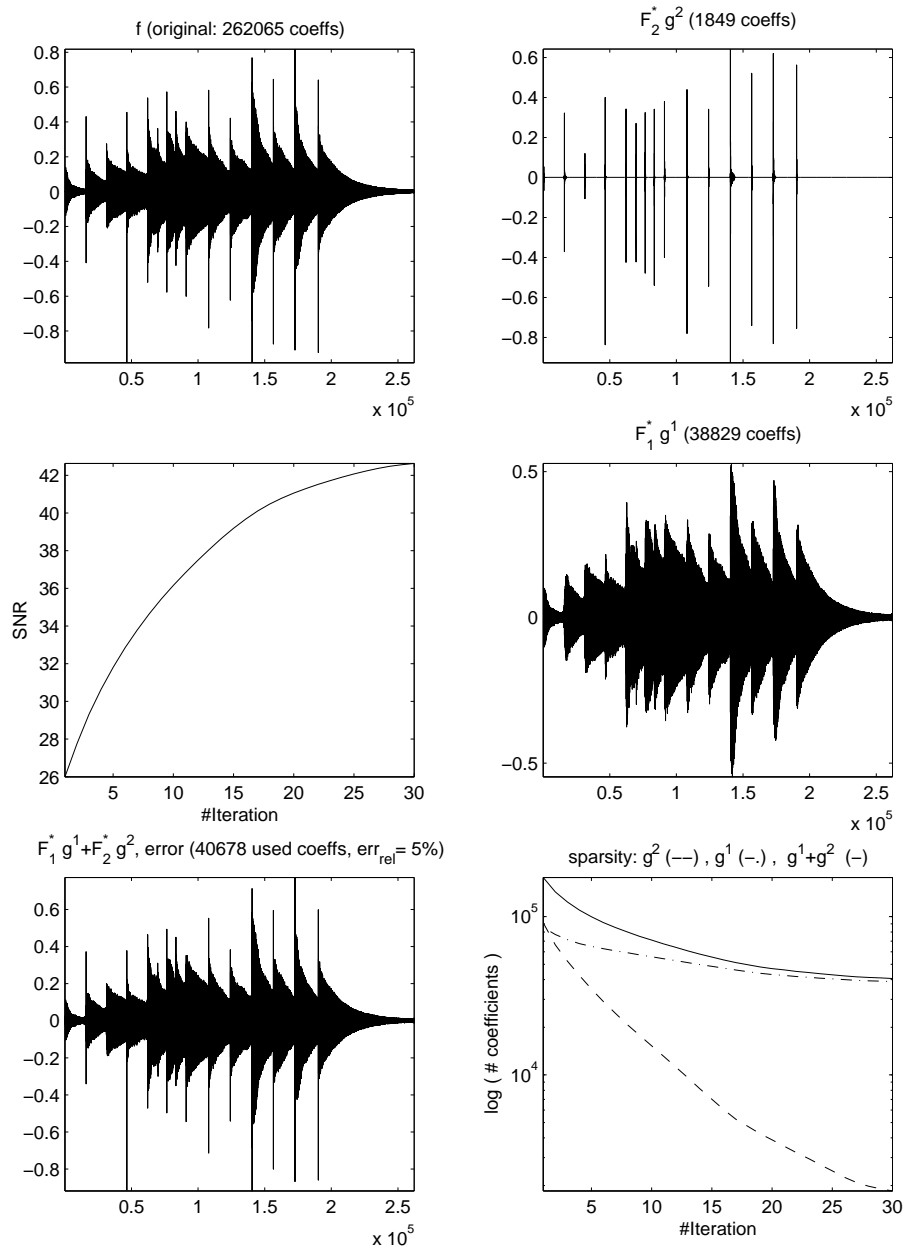


Figure 3.2: From top left to up right: “Glockenspiel” data, Haar wavelet component ( $g^2$ ) (in time domain) after 30 iterations, SNR evolution through the iteration process, Fourier component ( $g^1$ ) (in time domain) after 30 iterations, reconstruction and error after 30 iterations, and sparsity evolution through the iteration process.

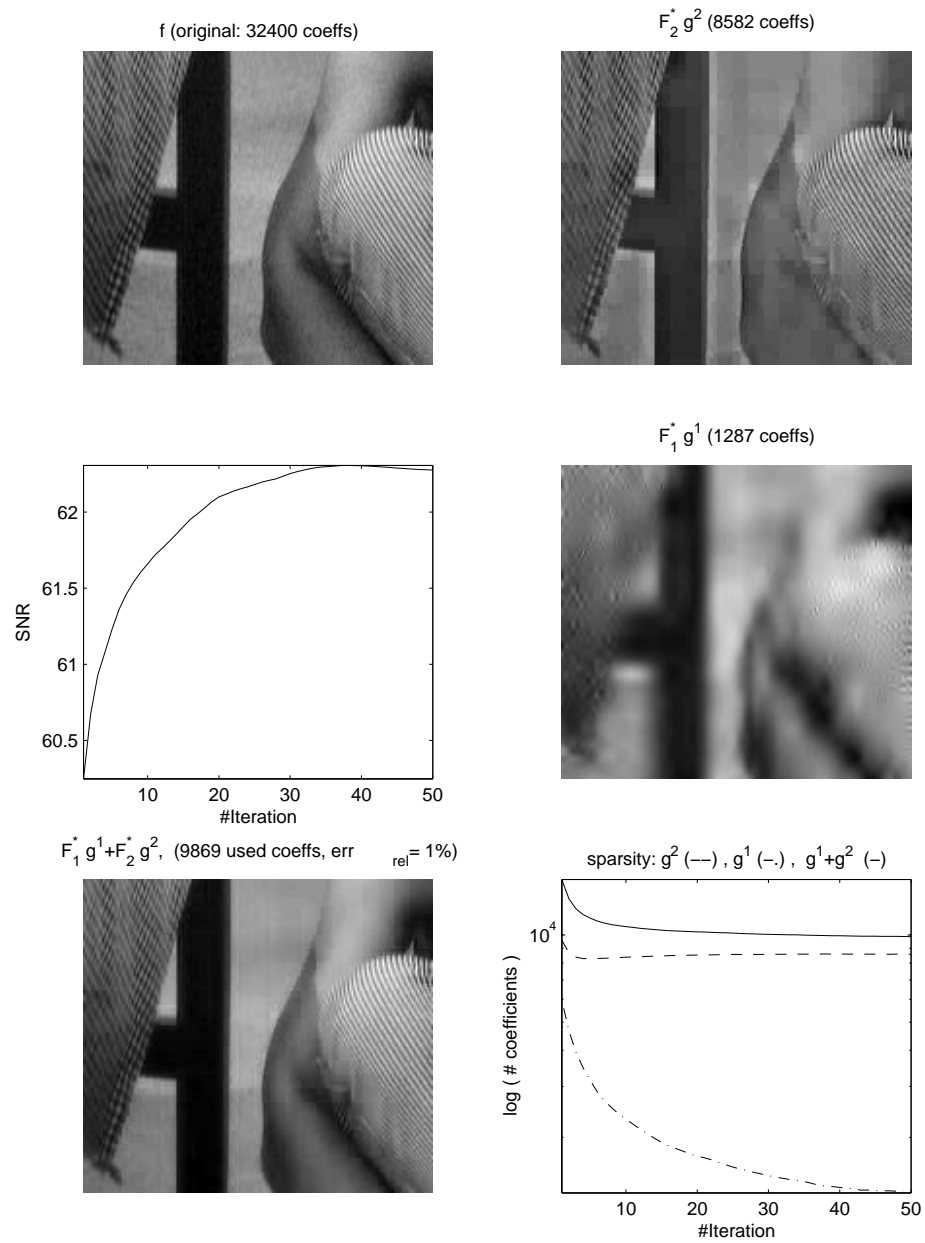


Figure 3.3: “Part of woman image”, Haar wavelet component ( $g^2$ ) (in time domain) after 50 iterations, SNR evolution through the iteration process, Daubechies-6-wavelet component ( $g^1$ ) (in time domain) after 50 iterations, reconstruction and error after 50 iterations, and sparsity evolution through the iteration process.

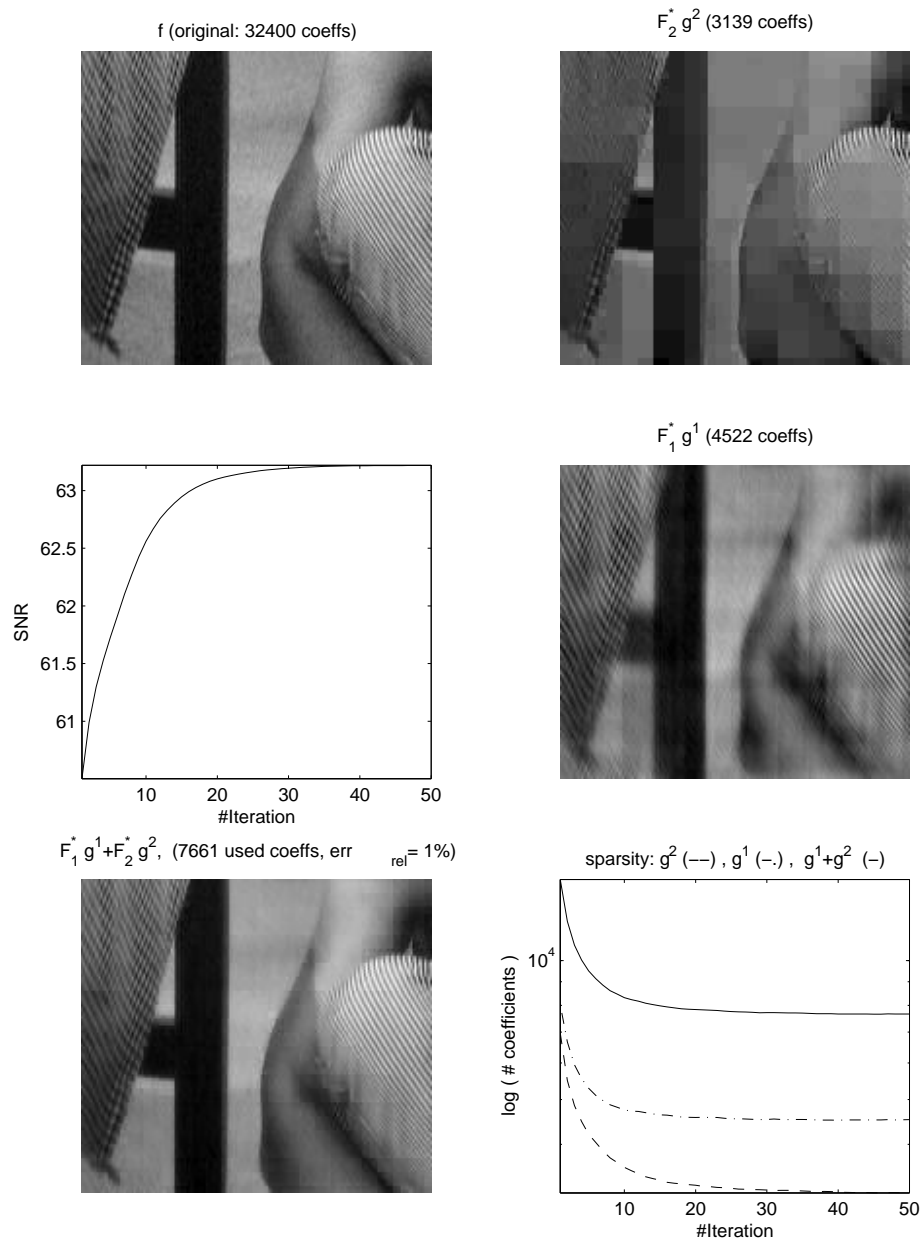


Figure 3.4: “Part of woman image”, Haar wavelet component ( $g^2$ ) (in time domain) after 50 iterations, SNR evolution through the iteration process, Fourier component ( $g^1$ ) (in time domain) after 50 iterations, reconstruction and error after 50 iterations, and sparsity evolution through the iteration process.

## 3.2 Image Decomposition and Restoration Problems

Inspired by recent papers of Vese–Osher [OV02] and Osher–Solé–Vese [OSV02] we present a wavelet–based treatment of variational problems arising in the field of image processing. In particular, we follow their approach and discuss a special class of variational functionals that induce a decomposition of images into oscillating and cartoon components and possibly an appropriate ‘noise’ component. In the setting of [OV02] and [OSV02], the cartoon component of an image is modeled by a  $BV$  function; the corresponding incorporation of  $BV$  penalty terms in the variational functional leads to PDE schemes that are numerically intensive. By replacing the  $BV$  penalty term by a  $B_1^1(L_1)$  term (which amounts to a slightly stronger constraint on the minimizer), and writing the problem in a wavelet framework, we obtain elegant and numerically efficient schemes with results very similar to those obtained in [OV02] and [OSV02]. This approach allows us, moreover, to incorporate general bounded linear blur operators into the problem so that the minimization leads to a simultaneous decomposition, deblurring and denoising.

### 3.2.1 Variational Problem with Smoothness and Sparsity Constraints

In general, an important problem in image processing is the restoration of the ‘true’ image from an observation. In almost all applications the observed image is a noisy and blurred version of the true image. In principle, the restoration task can be understood as an inverse problem, i.e. one can attack it by solving a related variational problem.

Here we focus on a special class of variational problems which induce a decomposition of images in oscillating and cartoon components; the cartoon part is ideally piecewise smooth with possible abrupt edges and contours; the oscillation part on the other hand ‘fills’ in the smooth regions in the cartoon with texture -like features. Several authors, e.g. [OV02, OSV02], propose to model the cartoon component by the space  $BV$  which induces a penalty term that allows edges and contours in the reconstructed cartoon images. However, the minimization of variational problems of this type usually results in PDE based schemes that are numerically intensive.

The main goal is to provide a computationally thriftier algorithm by using a wavelet–based scheme that solves not the same but a very similar variational problem, in which the  $BV$ –constraint, which cannot easily be expressed in the wavelet domain, is replaced by a  $B_1^1(L_1)$ –term, i.e. a slightly stricter constraint (since  $B_1^1(L_1) \subset BV$  in two dimensions). Moreover, we can allow the involvement of general linear bounded blur operators, which extends the range of application. By applying recent results, see [DDD04], we show convergence of the proposed scheme.

In order to give a brief description of the underlying variational problems, we recall the methods proposed in [OV02, OSV02]. They follow the idea of Y. Meyer [Mey02], proposed as an improvement on the total variation framework of L. Rudin, S. Osher and E. Fatemi [ROF92]. In principle, the models can be understood as a decomposition of an image  $f$  into  $f = u + v$ , where  $u$  represents the cartoon part and  $v$  the texture part. In

the Vese–Osher model, see [OV02], the decomposition is induced by solving

$$\inf_{u, g_1, g_2} G_p(u, g_1, g_2) , \quad \text{where} \quad (3.2.1)$$

$$G_p(u, g_1, g_2) = \int_{\Omega} |\nabla u| + \lambda \|f - (u + \operatorname{div} g)\|_{L_2(\Omega)}^2 + \mu \|g\|_{L_p(\Omega)} ,$$

with  $f \in L_2(\Omega)$ ,  $\Omega \subset \mathbb{R}^2$ , and  $v = \operatorname{div} g = \operatorname{div}(g_1, g_2)$ . The first term is the total variation of  $u$ . If  $u \in L_1$  and  $|\nabla u|$  is a finite measure on  $\Omega$ , then  $u \in BV(\Omega)$ . This space allows discontinuities, therefore edges and contours generally appear in  $u$ . The second term represents the restoration discrepancy; to penalize  $v$ , the third term approximates (by taking  $p$  finite) the norm of the space of oscillating functions introduced by Y. Meyer (with  $p = \infty$ ) which is in some sense dual to  $BV(\Omega)$ . (For details we refer the reader to [Mey02].) Setting  $p = 2$  and  $g = \nabla P + Q$ , where  $P$  is a single-valued function and  $Q$  is a divergence-free vector field, it is shown in [OSV02] that the  $v$ -penalty term can be expressed by

$$\|g\|_{L_2(\Omega)} = \left( \int_{\Omega} |\nabla(\Delta)^{-1} v|^2 \right)^{1/2} = \|v\|_{H^{-1}(\Omega)} .$$

(The  $H^{-1}$  calculus is allowed as long as we deal with oscillatory texture/noise components that have zero mean.) With these assumptions, the variational problem (3.2.1) simplifies to solving

$$\inf_{u, g_1, g_2} G_2(u, v) , \quad \text{where} \quad (3.2.2)$$

$$G_2(u, v) = \int_{\Omega} |\nabla u| + \lambda \|f - (u + v)\|_{L_2(\Omega)}^2 + \mu \|v\|_{H^{-1}(\Omega)} .$$

In general, one drawback is that the minimization of (3.2.1) or (3.2.2) leads to numerically intensive schemes.

Instead of solving problem (3.2.2) by means of finite difference schemes, we propose a wavelet-based treatment. We are encouraged by the fact that elementary methods based on wavelet shrinkage solve similar extremal problems where  $BV(\Omega)$  is replaced by the Besov space  $B_1^1(L_1(\Omega))$ . Since  $BV(\Omega)$  can not be simply described in terms of wavelet coefficients, it is not clear that  $BV(\Omega)$  minimizers can be obtained in this way. Yet, it is shown in [CDPX99], exploiting  $B_1^1(L_1(\Omega)) \subset BV(\Omega) \subset B_1^1(L_1(\Omega))$  – *weak*, that methods using Haar systems provide near  $BV(\Omega)$  minimizers. So far there exists no similar result for general (in particular smoother) wavelet systems. We shall nevertheless use wavelets that have more smoothness/vanishing moments than Haar wavelets, because we expect them to be better suited to the modeling of the smooth parts in the cartoon image. Though we may not obtain provable ‘near-best- $BV$ -minimizers’, we hope to nevertheless not be ‘too far off’. Limiting ourselves to the case  $p = 2$ , replacing  $BV(\Omega)$  by  $B_1^1(L_1(\Omega))$ , and, moreover, extending the range of applicability by incorporating a bounded linear operator  $K$ , we end up with the following variational problem:

$$\inf_{u, v} \mathcal{F}_f(v, u) , \quad \text{where}$$

$$\mathcal{F}_f(v, u) = \|f - K(u + v)\|_{L_2(\Omega)}^2 + \gamma \|v\|_{H^{-1}(\Omega)}^2 + 2\alpha |u|_{B_1^1(L_1(\Omega))} .$$



### 3.2.2 Approximation of the Solution

As stated in Section 3.2.1, we aim to solve

$$\inf_{u,v} \mathcal{F}_f(v, u) , \quad \text{where} \quad (3.2.3)$$

$$\mathcal{F}_f(v, u) = \|f - K(u + v)\|_{L_2(\Omega)}^2 + \gamma \|v\|_{H^{-1}(\Omega)}^2 + 2\alpha |u|_{B_1^1(L_1(\Omega))} .$$

At first, we may observe the following

**Lemma 3.2.1** *If the null-space  $\mathcal{N}(K)$  of the operator  $K$  is trivial, then the variational problem (3.2.3) has a unique minimizer.*

This can be seen as follows:

$$\begin{aligned} \mathcal{F}_f(\mu(v, u) + (1 - \mu)(v', u')) - \mu \mathcal{F}_f(v, u) - (1 - \mu) \mathcal{F}_f(v', u') = \\ -\mu(1 - \mu) \left( \|K(u - u' + v - v')\|_{L_2(\Omega)}^2 + \gamma \|v - v'\|_{H^{-1}(\Omega)}^2 \right) \\ + 2\alpha \left( |\mu u + (1 - \mu)u'|_{B_1^1(L_1(\Omega))} - \mu |u|_{B_1^1(L_1(\Omega))} - (1 - \mu) |u'|_{B_1^1(L_1(\Omega))} \right) \end{aligned} \quad (3.2.4)$$

with  $0 < \mu < 1$ . Since the Banach norm is convex the right hand side of (3.2.4) is non-positive, i.e.  $\mathcal{F}_f$  is convex. Since  $\mathcal{N}(K) = \{0\}$ , the term  $\|K(u - u' + v - v')\|$  can be zero only if  $u - u' + v - v' = 0$ , moreover,  $\|v - v'\|$  is zero only if  $v - v' = 0$ . Hence, (3.2.4) is strictly convex.  $\square$

In order to solve this problem by means of wavelets we have to switch to the sequence space formulation. When  $K$  is the identity operator the problem simplifies to

$$\inf_{u,v} \left\{ \sum_{\lambda \in J} (|f_\lambda - (u_\lambda + v_\lambda)|^2 + \gamma 2^{-2|\lambda|} |v_\lambda|^2 + 2\alpha |u_\lambda|) \right\} , \quad (3.2.5)$$

where  $J = \{\lambda = (i, j, k) : k \in J_j, j \in \mathbb{Z}, i = 1, 2, 3\}$  is the index set used in our separable setting. The minimization of (3.2.5) is straightforward, since it decouples into easy one-dimensional minimizations. This results in an explicit shrinkage scheme, presented also in [DT04]:

**Proposition 3.2.1** *Let  $f$  be a given function. The functional (3.2.5) is minimized by the parametrized class of functions  $\tilde{v}_{\gamma, \alpha}$  and  $\tilde{u}_{\gamma, \alpha}$  given by the following non-linear filtering of the wavelet series of  $f$ :*

$$\tilde{v}_{\gamma, \alpha} = \sum_{\lambda \in J_{j_0}} (1 + \gamma 2^{-2|\lambda|})^{-1} [f_\lambda - S_{\alpha(2^{2|\lambda|} + \gamma)/\gamma}(f_\lambda)] \psi_\lambda$$

and

$$\tilde{u}_{\gamma, \alpha} = \sum_{k \in I_{j_0}} \langle f, \tilde{\phi}_{j_0, k} \rangle \phi_{j_0, k} + \sum_{\lambda \in J_{j_0}} S_{\alpha(2^{2|\lambda|} + \gamma)/\gamma}(f_\lambda) \psi_\lambda ,$$

where  $S_t$  denotes the soft-shrinkage operator,  $J_{j_0}$  all indices  $\lambda$  for scales larger than  $j_0$  and  $I_{j_0}$  the indices  $\lambda$  for the fixed scale  $j_0$ .

In the case where  $K$  is not the identity operator the minimization process results in a coupled system of nonlinear equations for the wavelet coefficients  $u_\lambda$  and  $v_\lambda$ , which is not as straightforward to solve. To overcome this problem, we adapt an iterative approach. As in [DDD04] we derive the iterative algorithm from a sequence of so-called surrogate functionals that are each easy to minimize, and for which one hopes that the successive minimizers have the minimizing element of (3.2.3) as limit. However, contrary to [DDD04] our variational problem has mixed quadratic and non-quadratic penalties. This requires a slightly different use of surrogate functionals. In [DD04b, DD04a] a similar  $u + v$  problem is solved by an approach that combines  $u$  and  $v$  into one vector-valued function  $(u, v)$ . This leads to alternating iterations with respect to  $u$  and  $v$  simultaneously. It can be shown that the minimizers of the resulting alternating algorithm strongly converge to the desired unique solution, [DD04b].

We will follow a different approach here, in which we first solve the quadratic problem for  $v$ , and then construct an iteration scheme for  $u$ . To this end, we introduce the differential operator  $T := (-\Delta)^{1/2}$ . Setting  $v = Th$  the variational problem (3.2.3) reads as

$$\inf_{(u,h)} \mathcal{F}_f(h, u) , \quad \text{with} \quad (3.2.6)$$

$$\mathcal{F}_f(h, u) = \|f - K(u + Th)\|_{L_2(\Omega)}^2 + \gamma \|h\|_{L_2(\Omega)}^2 + 2\alpha |u|_{B_1^1(L_1(\Omega))} .$$

Minimizing (3.2.6) with respect to  $w$  results in

$$\tilde{h}_\gamma(f, u) = (T^* K^* K T + \gamma)^{-1} T^* K^* (f - Ku)$$

or equivalently

$$\tilde{v}_\gamma(f, u) = T(T^* K^* K T + \gamma)^{-1} T^* K^* (f - Ku) .$$

Inserting this explicit expression for  $\tilde{h}_\gamma(f, u)$  in (3.2.6) and defining

$$f_\gamma := T_\gamma f, \quad T_\gamma^2 := I - K T (T^* K^* K T + \gamma)^{-1} T^* K^* , \quad (3.2.7)$$

we obtain

$$\mathcal{F}_f(\tilde{h}_\gamma(f, u), u) = \|f_\gamma - T_\gamma K u\|_{L_2(\Omega)}^2 + 2\alpha |u|_{B_1^1(L_1(\Omega))} . \quad (3.2.8)$$

Thus, the remaining task is to solve

$$\inf_u \mathcal{F}_f(\tilde{h}_\gamma(f, u), u) , \quad \text{where} \quad (3.2.9)$$

$$\mathcal{F}_f(\tilde{h}_\gamma(f, u), u) = \|f_\gamma - T_\gamma K u\|_{L_2(\Omega)}^2 + 2\alpha |u|_{B_1^1(L_1(\Omega))} .$$

The corresponding variational equations in the sequence space representation are

$$\forall \lambda : (K^* T_\gamma^2 K u)_\lambda - (K^* f_\gamma)_\lambda + \alpha \text{sign}(u_\lambda) = 0 .$$

This gives a coupled system of nonlinear equations for  $u_\lambda$ . For this reason we construct surrogate functionals that remove the influence of  $K^* T_\gamma^2 K u$ . First, we choose a constant  $C$  such that  $\|K^* T_\gamma^2 K\| < C$ . Since  $\|T_\gamma\| \leq 1$ , it suffices to require that  $\|K^* K\| < C$ . Then we define the functional

$$\Phi(u; a) := C \|u - a\|_{L_2(\Omega)}^2 - \|T_\gamma K(u - a)\|_{L_2(\Omega)}^2$$

which depends on an auxiliary element  $a \in L_2(\Omega)$ . We observe that  $\Phi(u, a)$  is strictly convex in  $u$  for any  $a$ . Since  $K$  can be rescaled, we limit our analysis without loss of generality to the case  $C = 1$ . We finally add  $\Phi(u; a)$  to  $\mathcal{F}_f(\tilde{h}_\gamma(f, u), u)$  and obtain the following surrogate functional

$$\begin{aligned} \mathcal{F}_f^{sur}(\tilde{h}_\gamma(f, a), u; a) &= \mathcal{F}_f(\tilde{h}_\gamma(f, u), u) + \Phi(u; a) \\ &= \sum_{\lambda} \{u_{\lambda}^2 - 2u_{\lambda}(a + K^*T_{\gamma}^2(f - Ka))_{\lambda} + 2\alpha|u_{\lambda}|\} \\ &\quad + \|f_{\gamma}\|_{L_2(\Omega)}^2 + \|a\|_{L_2(\Omega)}^2 - \|T_{\gamma}Ka\|_{L_2(\Omega)}^2 . \end{aligned} \quad (3.2.10)$$

The advantage of minimizing (3.2.10) is that the variational equations for  $u_{\lambda}$  decouple. The summands of (3.2.10) are differentiable in  $u_{\lambda}$  except at the point of non-differentiability. The variational equations for each  $\lambda$  are now given by

$$u_{\lambda} + \alpha \text{sign}(u_{\lambda}) = (a + K^*T_{\gamma}^2(f - Ka))_{\lambda} .$$

This results in an explicit soft-shrinkage operation for  $u_{\lambda}$

$$u_{\lambda} = S_{\alpha}((a + K^*T_{\gamma}^2(f - Ka))_{\lambda}) .$$

The next proposition summarizes our findings; it is the specialization to our particular case of a more general theorem in [DDD04].

**Proposition 3.2.2** *Suppose  $K$  is a linear bounded operator modeling the blur, with  $K$  maps  $L_2(\Omega)$  to  $L_2(\Omega)$  and  $\|K^*K\| < 1$ . Moreover, assume  $T_{\gamma}$  is defined as in (3.2.7) and the functional  $\mathcal{F}_f^{sur}(\tilde{h}, u; a)$  is given by*

$$\mathcal{F}_f^{sur}(\tilde{h}_\gamma(f, u), u; a) = \mathcal{F}_f(\tilde{h}_\gamma(f, u), u) + \Phi(u; a) .$$

*Then, for arbitrarily chosen  $a \in L_2(\Omega)$ , the functional  $\mathcal{F}_f^{sur}(\tilde{h}_\gamma(f, u), u; a)$  has a unique minimizer in  $L_2(\Omega)$ . The minimizing element is given by*

$$\tilde{u}_{\gamma, \alpha} = \mathbf{S}_{\alpha}(a + K^*T_{\gamma}^2(f - Ka)) ,$$

*where the operator  $\mathbf{S}_{\alpha}$  is defined component-wise by*

$$\mathbf{S}_{\alpha}(x) = \sum_{\lambda} S_{\alpha}(x_{\lambda})\psi_{\lambda} .$$

The proof follows from [DDD04]. One can now define an iterative algorithm by repeated minimization of  $\mathcal{F}_f^{sur}$ :

$$u^0 \text{ arbitrary ; } u^n = \arg \min_u \left( \mathcal{F}_f^{sur}(\tilde{h}_\gamma(f, u), u; u^{n-1}) \right) \quad n = 1, 2, \dots \quad (3.2.11)$$

The convergence result of [DDD04] can again be applied directly:

**Theorem 3.2.1** *Suppose  $K$  is a linear bounded operator, with  $\|K^*K\| < 1$ , and that  $T_\gamma$  is defined as in (3.2.7). Then the sequence of iterates*

$$u_{\gamma,\alpha}^n = \mathbf{S}_\alpha(u_{\gamma,\alpha}^{n-1} + K^*T_\gamma^2(f - Ku_{\gamma,\alpha}^{n-1})) , \quad n = 1, 2, \dots ,$$

*with arbitrarily chosen  $u^0 \in L_2(\Omega)$ , converges in norm to a minimizer  $\tilde{u}_{\gamma,\alpha}$  of the functional*

$$\mathcal{F}_f(\tilde{h}_\gamma(f, u), u) = \|T_\gamma(f - Ku)\|_{L_2(\Omega)}^2 + 2\alpha|u|_{B_1^1(L_1(\Omega))} .$$

*If  $\mathcal{N}(T_\gamma K) = \{0\}$ , then the minimizer  $\tilde{u}_{\gamma,\alpha}$  is unique, and every sequence of iterates converges to  $\tilde{u}_{\gamma,\alpha}$  in norm.*

Combining the result of Theorem 3.2.1 and the representation for  $\tilde{v}$  we summarize how the image can finally be decomposed in cartoon and oscillating components.

**Corollary 3.2.1** *Assume that  $K$  is a linear bounded operator modeling the blur, with  $\|K^*K\| < 1$ . Moreover, if  $T_\gamma$  is defined as in (3.2.7) and if  $\tilde{u}_{\gamma,\alpha}$  is the minimizing element of (3.2.9), obtained as a limit of  $u_{\gamma,\alpha}^n$  (see Theorem 3.2.1), then the variational problem*

$$\inf_{(u,h)} \mathcal{F}_f(h, u), \text{ with } \mathcal{F}_f(h, u) = \|f - K(u + Th)\|_{L_2(\Omega)}^2 + \gamma\|h\|_{L_2(\Omega)}^2 + 2\alpha|u|_{B_1^1(L_1(\Omega))}$$

*is minimized by the class*

$$(\tilde{u}_{\gamma,\alpha}, (T^*K^*KT + \gamma)^{-1}T^*K^*(f - K\tilde{u}_{\gamma,\alpha})) .$$

*where  $\tilde{u}_{\gamma,\alpha}$  is the unique limit of the sequence*

$$u_{\gamma,\alpha}^n = \mathbf{S}_\alpha(u_{\gamma,\alpha}^{n-1} + K^*T_\gamma^2(f - Ku_{\gamma,\alpha}^{n-1})) , \quad n = 1, 2, \dots .$$

### 3.2.3 Numerics and Improvements by Redundancy

The non-linear filtering rule of Proposition 3.2.1 gives explicit descriptions of  $\tilde{v}$  and  $\tilde{u}$  that are computed by fast discrete wavelet schemes. However, non-redundant filtering very often creates artifacts in terms of undesirable oscillations, which manifest themselves as ringing and edge blurring. Poor directional selectivity of traditional tensor product wavelet bases likewise cause artifacts. In this section we discuss various refinements on the basic algorithm that address this problem. In particular, we shall use redundant translation invariant schemes, complex wavelets, and additional edge dependent penalty weights.

#### Translation invariance by cycle-spinning

Assume that we are given an image with  $2^M$  rows of  $2^M$  pixels, where the gray value of each pixel gives an average of  $f$  on a square  $2^{-M} \times 2^{-M}$ , which we denote by  $f_k^M$ , with  $k$  a double index running through all the elements of  $\{0, 1, \dots, 2^M - 1\} \times \{0, 1, \dots, 2^M - 1\}$ . A traditional wavelet transform then computes  $f_l^j, d_l^{j,i}$  with  $j_0 \leq j \leq M, i = 1, 2, 3$  and  $l \in \{0, 1, \dots, 2^j - 1\} \times \{0, 1, \dots, 2^j - 1\}$  for each  $j$ , where the  $f_l^j$  stand for an average of  $f$  on mostly localized on (and indexed by) the squares  $[l_1 2^{-j}, (l_1 + 1) 2^{-j}] \times [l_2 2^{-j}, (l_2 + 1) 2^{-j}]$ ,

and the  $d_l^{j,i}$  stand for the different species of wavelets (in two dimensions, there are three) in the tensor product multi-resolution analysis. Because the corresponding wavelet basis is not translation invariant, Coifman and Donoho proposed in [CD95] to recover translation invariance by averaging over the  $2^{2(M+1-j_0)}$  translates of the wavelet basis; since many wavelets occur in more than one of these translated bases (in fact, each  $\psi_{j,i,k}(x - 2^M n)$  in exactly  $2^{2(j+1-j_0)}$  different bases), the average over all these bases uses only  $(M+1-j_0)2^{2M}$  different basis functions (and not  $2^{4(M+1-j_0)}$  = number of bases  $\times$  number of elements in each basis). This approach is called *cycle-spinning*. Writing, with a slight abuse of notation,  $\psi_{j,i,k+2^{j-M}n}$  for the translate  $\psi_{j,i,k}(x - 2^M n)$ , this average can then be written as

$$f^M = 2^{-2(M+1-j_0)} \sum_{l_1, l_2=0}^{2^M-1} \left\{ f_{l_2-M+j_0}^{j_0} \phi_{j_0, l_2-M+j_0} + \sum_{j=j_0}^{M-1} 2^{2(j-j_0)} \sum_{i=1}^3 d_{l_2-M+j}^{j,i} \psi_{j,i, l_2-M+j} \right\}.$$

Carrying out our nonlinear filtering in each of the bases and averaging the result then corresponds to applying the corresponding nonlinear filtering on the (much smaller number of) coefficients in the last expression. This is the standard way to implement thresholding on cycle-spinned representations.

The resulting sequence space representation of the variational functional (3.2.5) has to be adapted to the redundant representation of  $f$ . To this end, we note that the Besov penalty term takes the form

$$|f|_{B_p^\beta(L_p)} \sim \left( \sum_{j \geq j_0, i, k} 2^{(js+2(j-M))} |\langle f, \tilde{\psi}_{j,i,k} 2^{j-M} \rangle|^p \right)^{1/p}.$$

The norms  $\|\cdot\|_{L_2}^2$  and  $\|\cdot\|_{H^{-1}}^2$  change similarly. Consequently, we obtain the same minimization rule but with respect to a richer class of wavelet coefficients.

### Directional sensitivity by frequency projections

It has been shown by several authors [Kin99, Sel01, FvSCB00] that if one treats positive and negative frequencies separately in the one-dimensional wavelet transform (resulting in complex wavelets), the directional selectivity of the corresponding two-dimensional multi-resolution analysis is improved. This can be done by applying the following orthogonal projections:

$$\begin{aligned} \mathcal{P}^+ &: L_2 \rightarrow L_{2,+} = \{f \in L_2 : \text{supp } \hat{f} \subseteq [0, \infty)\} \\ \mathcal{P}^- &: L_2 \rightarrow L_{2,-} = \{f \in L_2 : \text{supp } \hat{f} \subseteq (-\infty, 0]\}. \end{aligned}$$

The projectors  $\mathcal{P}^+$  and  $\mathcal{P}^-$  may be either applied to  $f$  or to  $\{\phi, \tilde{\phi}\}$  and  $\{\psi, \tilde{\psi}\}$ . In a discrete framework these projections have to be approximated. This has been done in different ways in the literature. In [Kin99, Sel01] Hilbert transform pairs of wavelets are used. In [FvSCB00]  $f$  is projected (approximately) by multiplying with shifted generator symbols in the frequency domain. We follow the second approach, i.e.

$$(P^+ f)^\wedge(\omega) := \hat{f}(\omega) H(\omega - \pi/2) \quad \text{and} \quad (P^- f)^\wedge(\omega) := \hat{f}(\omega) H(\omega + \pi/2),$$

where  $f$  denotes the function to be analyzed and  $H$  is the low-pass filter for a conjugate quadrature mirror filter pair. One then has

$$\hat{f}(\omega) = (B^+ P^+ f)^\wedge(\omega) + (B^- P^- f)^\wedge(\omega) , \quad (3.2.12)$$

where the back-projections are given by

$$(B^+ f)^\wedge = \overline{\hat{f} H(\cdot - \pi/2)} \quad \text{and} \quad (B^- f)^\wedge = \overline{\hat{f} H(\cdot + \pi/2)}$$

respectively. This technique provides us with a simple multiplication scheme in Fourier, or equivalently, a convolution scheme in time domain. In a separable two-dimensional framework the projections need to be carried out in each of the two frequency variables, resulting in four approximate projection operators  $P^{++}, P^{+-}, P^{-+}, P^{--}$ . Because  $f$  is real, we have

$$(P^{++} f)^\wedge(-\omega) = \overline{(P^{--} f)^\wedge(\omega)} \quad \text{and} \quad (P^{+-} f)^\wedge(-\omega) = \overline{(P^{-+} f)^\wedge(\omega)} ,$$

so that the computation of  $P^{-+} f$  and  $P^{--} f$  can be omitted. Consequently, the modified variational functional takes the form

$$\begin{aligned} \mathcal{F}_f(u, v) &= 2 (\|P^{++}(f - (u + v))\|_{L_2}^2 + \|P^{+-}(f - (u + v))\|_{L_2}^2) + \\ &\quad 2\lambda (\|P^{++}v\|_{H^{-1}}^2 + \|P^{+-}v\|_{H^{-1}}^2) + 2\alpha |u|_{B_{1(L_1)}^1} \\ &\leq 2 (\|P^{++}(f - (u + v))\|_{L_2}^2 + \|P^{+-}(f - (u + v))\|_{L_2}^2) + \\ &\quad 2\lambda (\|P^{++}v\|_{H^{-1}}^2 + \|P^{+-}v\|_{H^{-1}}^2) + \\ &\quad 4\alpha (|P^{++}u|_{B_{1(L_1)}^1} + |P^{+-}u|_{B_{1(L_1)}^1}) , \end{aligned}$$

which can be minimized with respect to  $\{P^{++}v, P^{++}u\}$  and  $\{P^{+-}v, P^{+-}u\}$  separately. The projections are be complex-valued, so that the thresholding operator needs to be adapted. Parameterizing the wavelet coefficients by modulus and angle and minimizing yields the following filtering rules for the projections of  $\tilde{v}_{\gamma,\alpha}$  and  $\tilde{u}_{\gamma,\alpha}$  (where  $\cdot\cdot$  stands for any combination of  $+$ ,  $-$ )

$$P^{\cdot\cdot} \tilde{v}_{\gamma,\alpha} = \sum_{\lambda \in J_{j_0}} (1 + \gamma 2^{-2|\lambda|})^{-1} [P^{\cdot\cdot} f_\lambda - S_{\alpha(2^{2|\lambda|} + \gamma)/\gamma}(|P^{\cdot\cdot} f_\lambda|) e^{i\omega(P^{\cdot\cdot} f)}] \psi_\lambda$$

and

$$P^{\cdot\cdot} \tilde{u}_{\gamma,\alpha} = \sum_{k \in I_{j_0}} \langle P^{\cdot\cdot} f, \tilde{\phi}_{j_0,k} \rangle \phi_{j_0,k} + \sum_{\lambda \in J_{j_0}} (1 + \gamma 2^{-2|\lambda|})^{-1} S_{\alpha(2^{2|\lambda|} + \gamma)/\gamma}(|P^{\cdot\cdot} f_\lambda|) e^{i\omega(P^{\cdot\cdot} f)} \psi_\lambda .$$

Finally, we have to apply the back-projections to obtain the minimizing functions

$$\tilde{v}_{\gamma,\alpha}^{BP} = B^{++} P^{++} \tilde{v}_{\gamma,\alpha} + B^{--} \overline{P^{++} \tilde{v}_{\gamma,\alpha}} + B^{+-} P^{+-} \tilde{v}_{\gamma,\alpha} + B^{-+} \overline{P^{+-} \tilde{v}_{\gamma,\alpha}}$$

and

$$\tilde{u}_{\gamma,\alpha}^{BP} = B^{++} P^{++} \tilde{u}_{\gamma,\alpha} + B^{--} \overline{P^{++} \tilde{u}_{\gamma,\alpha}} + B^{+-} P^{+-} \tilde{u}_{\gamma,\alpha} + B^{-+} \overline{P^{+-} \tilde{u}_{\gamma,\alpha}} .$$

### Weighted penalty functions

In order to improve the capability of preserving edges we additionally introduce a positive weight sequence  $w_\lambda$  in the  $H^{-1}$  penalty term. Consequently, we aim at minimizing a slightly modified sequence space functional

$$\sum_{\lambda \in J} (|f_\lambda - (u_\lambda + v_\lambda)|^2 + \gamma 2^{-2|\lambda|} w_\lambda |v_\lambda|^2 + 2\alpha |u_\lambda| \cdot 1_{\{\lambda \in J_{j_0}\}}) . \quad (3.2.13)$$

The resulting texture and cartoon components take the form

$$\tilde{v}_{\gamma, \alpha}^w = \sum_{\lambda \in J_{j_0}} (1 + \gamma w_\lambda 2^{-2|\lambda|})^{-1} [f_\lambda - S_{\alpha(2^{2|\lambda|} + \gamma w_\lambda)/\gamma w_\lambda}(f_\lambda)] \psi_\lambda$$

and

$$\tilde{u}_{\gamma, \alpha}^w = \sum_{k \in I_{j_0}} \langle f, \tilde{\phi}_{j_0, k} \rangle \phi_{j_0, k} + \sum_{\lambda \in J_{j_0}} S_{\alpha(2^{2|\lambda|} + \gamma w_\lambda)/\gamma w_\lambda}(f_\lambda) \psi_\lambda .$$

The main goal is to introduce a control parameter that depends on the local structure of  $f$ . The local penalty weight  $w_\lambda$  should be large in the presence of an edge and small otherwise; the result of this weighting is to enhance the sensitivity of  $u$  near edges. In order to do this, we must first localize the edges, which we do by a procedure similar to an edge detection algorithm in [MZ92]. This scheme rests on the analysis of the cycle-spinned wavelet coefficients  $f_\lambda$  at or near the same location but at different scales. We expect that the  $f_\lambda$  belonging to fine decomposition scales contain informations of edges (well localized) as well as oscillating components. Oscillating texture components typically show up in fine scales only; edges on the other hand leave a signature of larger wavelet coefficients through a wider range of scales. We thus apply the following not very sophisticated edge detector. Suppose that  $f \in V_M$  and  $j_e$  denotes some ‘critical’ scale, then for a certain range of scales  $|\lambda| = |(i, j, k)| = j \in \{j_0, \dots, j_1 - j_e - 2, j_1 - j_e - 1\}$  we mark all positions  $k$  where  $|f_\lambda|$  is larger than a level dependent threshold parameter  $t_j$ . Here the value  $t_j$  is chosen proportional to the mean value of all wavelet coefficients of level  $j$ . We say that  $|f_\lambda|$  represents an edge if  $k$  was marked for all  $j \in \{j_0, \dots, j_1 - j_e - 2, j_1 - j_e - 1\}$ . Finally, we adaptively choose the penalty sequence by setting

$$w_\lambda = \begin{cases} \Theta_\lambda & \text{if } j \in \{M - 1, \dots, j_1 - j_e\} \text{ and } k \text{ was marked as an edge ,} \\ \vartheta_\lambda & \text{otherwise ,} \end{cases}$$

where  $\vartheta_\lambda$  is close to one and  $\Theta_\lambda$  is much larger in order to penalize the corresponding  $v_\lambda$ 's.

### Numerical Results

Now, we present the numerical experiments obtained with our wavelet-based scheme. We start with the case where  $K$  is the identity operator. In order to show how the nonlinear (redundant) wavelet scheme acts on piecewise constant functions we decompose a geometric image (representing cartoon components only) with sharp contours, see Figure 3.5. We

observe that  $\tilde{u}$  represents the cartoon part very well. The texture component  $\tilde{v}$  (plus a constant for illustration purposes) contains only some very weak contour structures.

Next, we demonstrate the performance of the Haar shrinkage algorithm successively incorporating redundancy and local penalty weights. The redundancy is implemented by cycle spinning as describe in Section 3.2.3. The local penalty weights are computed the following way: firstly, we apply the shrinkage operator  $\mathbf{S}$  to  $f$  with a level dependent threshold (the threshold per scale is equal to two times the mean value of all the wavelet coefficients of the scale under consideration). Secondly, the non zero values of  $S_{\text{threshold}}(f_\lambda)$  per scale indicate where  $w_\lambda$  is set to  $\Theta_\lambda = 1 + C'$  (here  $C' = 10$ , moreover, we set  $w_\lambda$  equal to  $\vartheta_\lambda = 1$  elsewhere). The coefficients  $S_{\text{threshold}}(f_\lambda)$  for the first two scales of a segment of a woman image are visualized in Figure 3.6. In Figure 3.7, we present our numerical results. The upper row shows the original and the noisy image. The next row visualizes the results for non-redundant Haar shrinkage (Method A). The third row shows the same but incorporating cycle spinning (Method B), and the last row shows the incorporation of cycle spinning and local penalty weights. Each extension of the shrinkage method improves the results. This is also be confirmed by comparing the signal-to-noise-ratios (which is here defined as follows:  $SNR(f, g) = 10 \log_{10}(\|f\|^2 / \|f - g\|^2)$ ), see Table 3.1.

The next experiment is done on a fabric image, see Figure 3.8. But in contrast to the examples before, we present here the use of frequency projection as introduced in Section 3.2.3. The numerical result shows convincingly that the texture component can be also well separated from the cartoon part.

In order to compare the performance with the Vese–Osher TV model and with the Vese–Solé–Osher  $H^{-1}$  model we apply our scheme to a woman image (the same that was used in [OV02, OSV02]), see Figure 3.9. We obtain very similar results as obtained with the TV model proposed in [OV02]. Compared with the results obtained with the  $H^{-1}$  model proposed in [OSV02] we observe that our reconstruction of the texture component contains much less cartoon information. In terms of computational cost we have observed that even in the case of applying cycle spinning and edge enhancement our proposed wavelet shrinkage scheme is less time consuming than the Vese–Solé–Osher  $H^{-1}$  restoration scheme, see table 3.2, even when the wavelet method is implemented in Matlab, which is slower than the compiled version for the Vese–Solé–Osher scheme.

We end this section with presenting an experiment where  $K$  is not the identity operator. In our particular case  $K$  is a convolution operator with Gaussian kernel. The implementation is simply done in Fourier space. The upper row in Figure 3.10 shows the original  $f$  and the blurred image  $Kf$ . The lower row visualizes the results: the cartoon component  $\tilde{u}$ , the texture component  $\tilde{v}$ , and the sum of both  $\tilde{u} + \tilde{v}$ . One may clearly see that the deblurred image  $\tilde{u} + \tilde{v}$  contains (after a small number of iterations) more small scale details than  $Kf$ . This definitely shows the capabilities of the proposed iterative deblurring scheme (3.2.11).





Figure 3.5: From left to right: initial geometric image  $f$ ,  $\tilde{u}$ ,  $\tilde{v} + 150$ , computed with Db3 in the translation invariant setting,  $\alpha = 0.5$ ,  $\gamma = 0.01$ .



Figure 3.6: Left: noisy segment of a woman image, middle and right: first two scales of  $\mathbf{S}(f)$  inducing the weight function  $w$ .

Haar Shrinkage	$\text{SNR}(f, f_\varepsilon)$	$\text{SNR}(f, u + v)$	$\text{SNR}(f, u)$
Method A	20,7203	18,3319	16,0680
Method B	20,7203	21,6672	16,5886
Method C	20,7203	23,8334	17,5070

Table 3.1: Signal-to-noise ratios of the several decomposition methods (Haar shrinkage, translation invariant Haar shrinkage, translation invariant Haar shrinkage with edge enhancement).

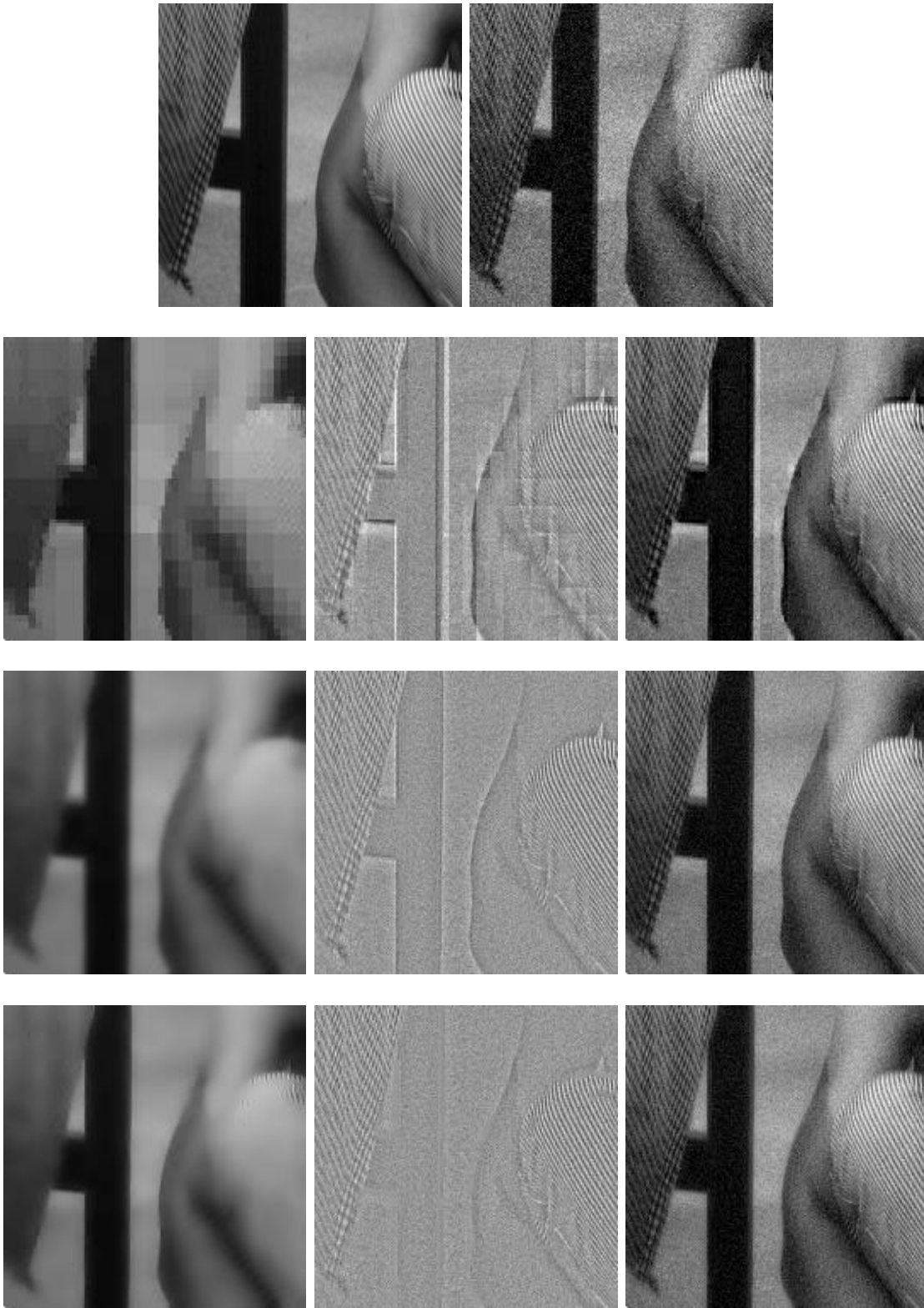


Figure 3.7: Top: initial and noisy image, 2nd row: non-redundant Haar shrinkage (Method A), 3rd row: translation invariant Haar shrinkage (Method B), bottom: translation invariant Haar shrinkage with edge enhancement (Method C); 2nd-4th row from left to right:  $\tilde{u}$ ,  $\tilde{v} + 150$  and  $\tilde{u} + \tilde{v}$ ,  $\alpha = 0.5$ ,  $\gamma = 0.0001$ , computed with Haar wavelets and critical scale  $j_e = -3$ .

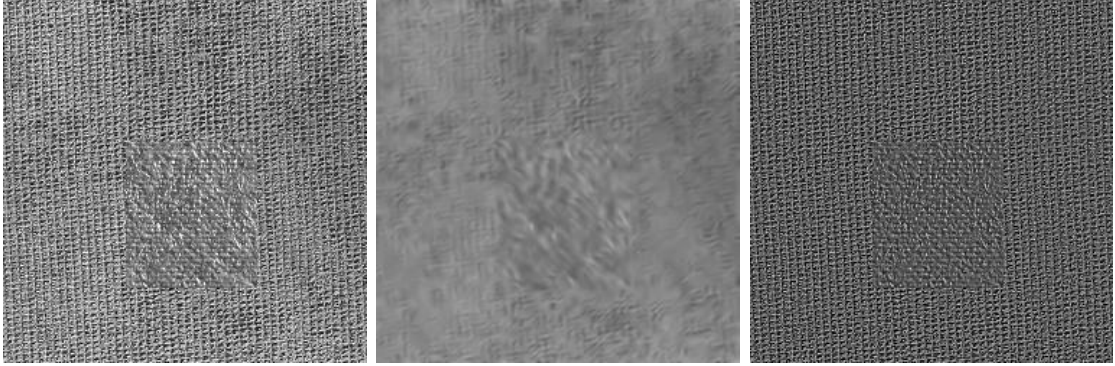


Figure 3.8: From left to right: initial fabric image  $f$ ,  $\tilde{u}$ ,  $\tilde{v} + 150$ , computed with Db4 incorporating frequency projections,  $\alpha = 0.8$ ,  $\gamma = 0.002$ .



Figure 3.9: Top from left to right: initial woman image  $f$ ,  $\tilde{u}$  and  $\tilde{v} + 150$ , computed with Db10 (Method C),  $\alpha = 0.5$ ,  $\gamma = 0.002$ ; bottom from left to right:  $u$  and  $v$  obtained by the Vese–Osher TV model and the  $v$  component obtained by the Vese–Solé–Osher  $H^{-1}$  model.

Data basis	"Barbara" image (512x512 pixel)
Hardware Architecture	PC
Operating System	linux
OS Distribution	redhat7.3
Model	PC, AMD Athlon-XP
Memory Size (MB)	1024
Processor Speed (MHz)	1333
Number of CPUs	1
Computational cost	(average over 10 runs)
PDE scheme in Fortran (compiler f77)	56,67 sec
wavelet shrinkage Method A (Matlab)	4,20 sec
wavelet shrinkage Method B (Matlab)	24,78 sec
wavelet shrinkage Method C (Matlab)	26,56 sec

Table 3.2: Comparison of computational cost of the PDE- and the wavelet-based methods.

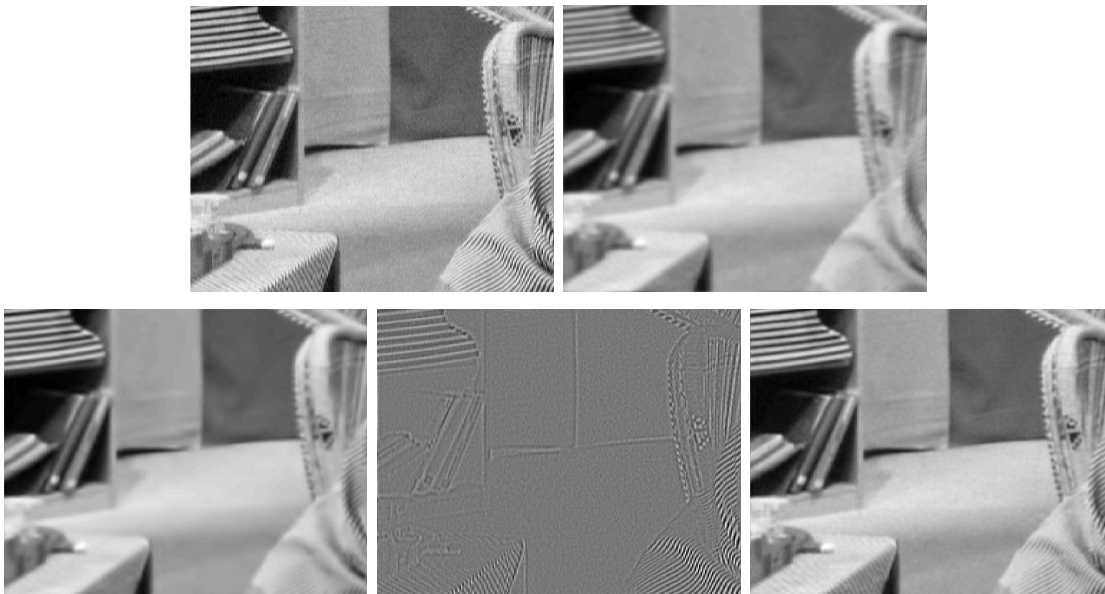


Figure 3.10: Top from left to right: initial image  $f$ , blurred image  $Kf$ ; bottom from left to right: deblurred  $\tilde{u}$ , deblurred  $\tilde{v} + 150$ , deblurred  $\tilde{u} + \tilde{v}$ , computed with Db3 using the iterative approach,  $\alpha = 0.2$ ,  $\gamma = 0.001$ .

### 3.3 Acceleration of Support Vector Machines

In this section, we apply the iterative strategy developed in Chapter 2 for reducing the runtime computational complexity of a Support Vector Machine classifier by an automatic and very efficient training. We propose an wavelet frame transformation of the Reduced Support Vector Machine. To achieve high run-time efficiency, the complexity of the classifier is made dependent on the input image patch. The fast classification uses a hierarchical evaluation over the number and as novelty over different levels of approximation accuracy of the Reduced Set Vectors. For non-symmetric data we achieve an early rejection of easy to discriminate vectors. In contrast to former methods the trade-off between accuracy and speed is very continuous. We compute a Haar-like structure of the Reduced Set Vectors that enables a very fast Support Vector Machine kernel evaluation by use of Integral Images. We apply this algorithm to the problem of face detection in images, but it can also be used for other image based classifications. It is shown in the experiments that this novel algorithm provides, for a comparable accuracy, a 15 fold speed-up over the Reduced Support Vector Machine and a 530 fold speed-up over the Support Vector Machine. The proposed face detector application gains real-time performance by a high accuracy.

#### 3.3.1 On Support Vector Machines and its Reduction

Image based classification tasks are time sensitive, e.g. detecting a specific object in an image, like a face is a computationally expensive task, as all the pixels of the image are potential object centers. Hence all the pixels have to be classified. Therefore, a method to increase the classification speed is based on a cascaded evaluation of hierarchical filters: pixels easy to discriminate are classified by simple and fast filters and pixels that resemble the object of interest are classified by more involved and slower filters. In the area of face detection, this method was independently introduced by Keren *et al.* [KOG01], by Romdhani *et al.* [RTSB01] and by Viola and Jones [VJ02].

The detector from Keren *et al.* [KOG01] assumes that the negative examples (i.e. the non-faces) are modeled by a Boltzmann distribution and that they are smooth. This assumption could increase the number of false positive in presence of a cluttered background. Romdhani *et al.* [RTSB01] use a Cascaded Reduced Set Vectors expansion of a Support Vector Machine (SVM) [Vap98]. The speed bottleneck of [RTSB01] is that at least one convolution of a  $20 \times 20$  filter has to be carried out on the full image, resulting in a computationally expensive evaluation of the kernel with an image patch. Viola & Jones [VJ02] use Haar-like oriented edge filters having a block like structure enabling a very fast evaluation by use of an Integral Image. These filters are weak, in the sense that their discrimination power is low. They are selected, among a finite set, by the Ada-boost algorithm that yields the ones with the best discrimination. A drawback of their approach is that it is not clear that the cascade achieves optimal performances. Practically, the training proceeds by trial and error, and often, the number of filters per stage must be manually selected so that the false positive rate decreases smoothly. Another drawback of the method is that the set of available filters is limited and manually selected. Additionally, the training of the classifier is very slow, as every filter (and there are about  $10^5$  of them) is evaluated on the whole set of training examples, and this is done every time

a filter is added to a stage of the cascade.

Here we present a novel efficient classification algorithm based on following features:

1. Use of a SVM classifier that is known to have optimal generalization capabilities.
2. To achieve high run-time efficiency we use a reduced set of Support Vectors [RTSB01].
3. The high run-time efficiency is also gained by a coarse-to-fine cascaded complexity of the classifier. For non-symmetric data (i.e. only few positives to many negatives) we achieve an early rejection of easy to discriminate vectors. This is realized with an only as fine as necessary approximated classifier by:
  - (a) a hierarchical evaluation over the number of RSV's (only as many RSV's as necessary) similar to [RTSB01] and as novelty
  - (b) over the levels of the approximation accuracy of the RSV's (only as fine as necessary approximation accuracy of the RSV's).
4. We constrain the RSV's to have a Haar-like block structure. Similarly to [VJ02], we use the Integral Image method introduced in [Cro84] to achieve high speed-ups, because this block structure enables a very fast kernel evaluation.
5. We use wavelet frame theory for gaining a near-optimal approximation of RSV's. The proposed learning stage is straightforward, automatic and does not require the manual selection of ad-hoc parameters, as opposed to the Viola and Jones method [VJ02].

The novelty to [RRV05] is 3. (b) and 5. by replacing the former ASA optimization using morphological filters. The difficulties was to result in the global optimum approximation in general and not in a local minimum. The other problem was to adjust the optimal approximation accuracy, because only one approximation level was used.

Let us now briefly introduce the terms of Support Vector Machines (SVM) and let us outline the usage of an approximation of SVMs called Reduced Set Vector Machines (RVM), see [SMB<sup>+</sup>99]. To this end, suppose that we have a labeled training set consisting of a series of  $20 \times 20$  image patches  $\mathbf{x}_i \in \mathcal{X}$  (arranged in a 400 dimensional vector) along with their class label  $y_i \in \{\pm 1\}$ . Support Vector classifiers implicitly map the data  $\mathbf{x}_i$  into a dot product space  $F$  via a (usually nonlinear) map  $\Phi : \mathcal{X} \rightarrow F$ ,  $\mathbf{x} \mapsto \Phi(\mathbf{x})$ . Often,  $F$  is referred to as the *feature space*. Although  $F$  can be high-dimensional, it is usually not necessary to explicitly work in that space [BGV92]. There exists a class of kernels  $k(\mathbf{x}, \mathbf{x}')$  which can be shown to compute the dot products in associated feature spaces, i.e.  $k(\mathbf{x}, \mathbf{x}') = \langle \Phi(\mathbf{x}), \Phi(\mathbf{x}') \rangle$ . It is shown in [Vap98] that the training of a SVM classifier provides a classifier with the *largest* margin, i.e. with the *best* generalization performances for the given training data and the given kernel. Thus, the classification of an image patch  $\mathbf{x}$  by an SVM classification function, with  $N_x$  support vectors  $\mathbf{x}_i$  with non-null coefficients  $\alpha_i$  and with a threshold  $b$ , is expressed as follows:

$$y = \text{sign} \left( \sum_i^{N_x} \alpha_i k(\mathbf{x}_i, \mathbf{x}) + b \right) \quad (3.3.1)$$

A kernel often used, and used here, is the Gaussian Radial Basis Function Kernel:

$$k(\mathbf{x}_i, \mathbf{x}) = \exp\left(\frac{-\|\mathbf{x}_i - \mathbf{x}\|^2}{2\sigma^2}\right) \quad (3.3.2)$$

The Support Vectors (SV) form a subset of the training vectors. The classification of one patch by an SVM is slow because there are many support vectors. The SVM can be approximated by a Reduced Set Vector (RVM) expansion [SMB<sup>+</sup>99]. We denote by  $\Psi_1 \in F$ , the vector normal to the separating hyper-plane of the SVM, and by  $\Psi'_{N_z} \in F$ , the vector normal to the RVM with  $N_z$  vectors:

$$\Psi_1 = \sum_{i=1}^{N_x} \alpha_i \Phi(\mathbf{x}_i), \quad \Psi'_{N_z} = \sum_{i=1}^{N_z} \beta_i \Phi(\mathbf{z}_i), \quad \text{with } N_z \ll N_x \quad (3.3.3)$$

The  $\mathbf{z}_i$  are the *Reduced Set Vectors* and are found by minimizing

$$\|\Psi_1 - \Psi'_{N_z}\|^2$$

with respect to  $\mathbf{z}_i$  and to  $\beta_i$ . They have the particularity that they can take any values, they are not limited to be one of the training vectors, as for the support vectors. Hence, much less Reduced Set Vectors are needed to approximate the SVM. For instance, an SVM with more than 8000 Support Vectors can be accurately approximated by an RVM with 100 Reduced Set Vectors. The second advantage of RVM is that they provide a hierarchy of classifiers. It was shown in [RTSB01] that the first Reduced Set Vector is the one that discriminates the data the most; and the second Reduced Set Vector is the one that discriminates most of the data that were mis-classified by the first Reduced Set Vector, etc. This hierarchy of classifiers is obtained by first finding  $\beta_1$  and  $\mathbf{z}_1$  that minimizes  $\|\Psi_1 - \beta_1 \Phi(\mathbf{z}_1)\|^2$ . Then the Reduced Set Vector  $k$  is obtained by minimizing  $\|\Psi_k - \beta_k \Phi(\mathbf{z}_k)\|^2$ , where  $\Psi_k = \Psi_1 - \sum_{i=1}^{k-1} \beta_i \Phi(\mathbf{z}_i)$ .

Then, Romdhani *et al.* used in [RTSB01] a *Cascaded Evaluation* based on an early rejection principle, to that the number of Reduced Set Vectors necessary to classify a patch is, on average, much less than the number of Reduced Set Vectors,  $N_z$ . So, the classification of a patch  $\mathbf{x}$  by an RVM with  $j$  Reduced Set Vector is:

$$y_j(\mathbf{x}) = \text{sign}\left(\sum_{i=1}^j \beta_{j,i} k(\mathbf{x}, \mathbf{z}_i) + b_j\right) \quad (3.3.4)$$

This approach provides a significant speedup over the SVM (by a factor of 30), but is still not fast enough, as the image has to be convolved, at least by a  $20 \times 20$  filter. The algorithm presented in this paper improves this method because it does not require to perform this convolution explicitly. Indeed, it approximates the Reduced Set Vectors by Haar-like filters and compute the evaluation of a patch using an Integral Image of the input image. An Integral Image [VJ02] is used to compute the sum of the pixels in a rectangular area of the input image in constant time, by just four additions. They can be used to compute very efficiently the dot product of an image patch with an image that has a block-like structure, i.e. rectangles of constant values.

### 3.3.2 Wavelet Frame Approximated Support Vector Machine

RVM provide a hierarchy of classifier of increasing complexity. Their use for fast face detection is demonstrated in [RTSB01]. But is still not fast enough, as the image has to be convolved, at least by a  $n \times n$  filter ( $n$  size of the filter, in our case  $n = 20$ ).

To achieve high run-time efficiency we use a coarse-to-fine cascaded complexity of the classifier as explained in 3.3.2. For this hierarchical approach we compute once at the learning stage different levels of approximations of the RSV's using over-complete wavelets (3.3.2). The obtained Wavelet Approximated Reduced Set Vectors (W-RSV) have a Haar-like structure. This rectangle structure enables a fast SVM kernel evaluation by use of Integral Images as motivated in 3.3.2.

#### Reduced Support Vector Evaluation by Integral Images

As it is explained in [RRV05], the speed bottleneck of the Cascaded Reduced Set Vector classifier is the computation of the kernel of a patch,  $\mathbf{x}$  with a Reduced Set Vector,  $\mathbf{z}_k$ . In the case of the Gaussian kernel

$$k(\mathbf{x}, \mathbf{z}_k) = \exp \left( \frac{-\|\mathbf{x} - \mathbf{z}_k\|^2}{2 \sigma^2} \right),$$

the computational load is spent in evaluating the norm of the difference between a patch and a Reduced Set Vector (see Equation (3.3.2)). This norm can be expanded as follows:

$$\|\mathbf{x} - \mathbf{z}_k\|^2 = \mathbf{x}'\mathbf{x} - 2\mathbf{x}'\mathbf{z}_k + \mathbf{z}_k'\mathbf{z}_k \quad (3.3.5)$$

As  $\mathbf{z}_k$  is independent of the input image, it can be pre-computed,  $\mathbf{x}'\mathbf{x}$  is efficiently computed using the Integral Image, finally, the computational load is determined by the term  $2\mathbf{x}'\mathbf{z}_k$ .

Let us now briefly describe the method of *Integral Images* which effectively reduces the computational load, see [Cro84], [VJ02]: The value of the integral image,  $ii$ , at point  $(x, y)$ , see Figure (3.11), is the sum of all the pixels, in the input image  $i$ , above and to the left of  $(x, y)$ :

$$ii(x, y) = \sum_{a \leq x, b \leq y} i(a, b). \quad (3.3.6)$$

Consequently, we have

$$\sum_{\substack{x_1 < a \leq x_4, \\ y_1 < b \leq y_4}} i(a, b) = ii(x_4, y_4) - ii(x_2, y_2) - ii(x_3, y_3) + ii(x_1, y_1) \quad (3.3.7)$$

Moreover, an integral image can be recursively computed by

$$\begin{aligned} s(x, y) &= s(x, y-1) + i(x, y) \\ ii(x, y) &= ii(x-1, y) + s(x, y), \end{aligned} \quad (3.3.8)$$

where  $s(x, y)$  is the cumulative column sum.



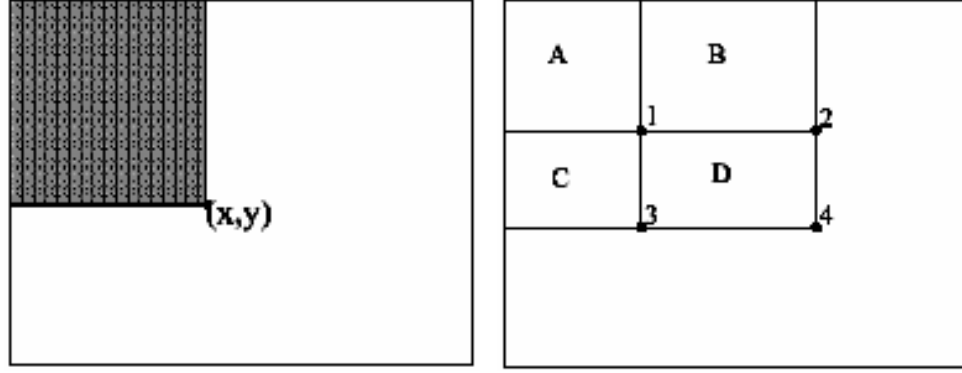


Figure 3.11: Definition of Integral Images (*left*) and computation of the sum of a rectangle area from the input image (*right*) by  $D = ii(4) - ii(2) - ii(3) + ii(1)$

### Wavelet-Shrinkage and Haar-Frame Reduced Set Vectors

In contrast to other approaches we do not use a wavelet transform of the input space as a pre-processing during the working process [Kar05, ZS02, CGT99]. The novelty is, that we apply the over-complete wavelet transform (OCWT) at the learning stage. Our approach proposes a wavelet transform of the Reduced Support Vector Machine itself.

The reason to proceed this way is because non-redundant representations often creates artifacts caused by the restricted grid of the wavelet basis. For our purpose, it is essential to pick a representation that optimally meets the local image structure (see Figure 3.12). The OCWT has its origin in translation invariance, i.e. representing the image by all possible shifted versions of the underlying wavelet basis.

In order to make full usage of the concept of integral images it would be desirable to approximate the computed RSV's,  $\mathbf{z}$  by block-wise structured images that are not too far off while keeping the number of rectangular regions with constant gray value much smaller than in  $\mathbf{z}$ . Roughly speaking, we are searching for an approximation of a given image  $\mathbf{z}$  by a piecewise block structured image  $\mathbf{u}$  which is as sparse as possible. As we have seen in Section 3.2, this optimization problem can be casted in the following variational form

$$\min_{\mathbf{u}} \|\mathbf{z} - \mathbf{u}\|_{L_2}^2 + 2\alpha |\mathbf{u}|_{B_1^1(L_1)} , \quad (3.3.9)$$

where  $B_1^1(L_1)$  denotes a particular Besov semi-norm; for an overview we again refer the reader to [Tri78, ST87] and for a detailed discussion of the problem to [CDPX99]. The Besov (semi) norm of a given function can be expressed by means of its wavelet coefficients and, moreover, in two dimensions the Besov penalty is nothing else than a  $\ell_1$  constraint on the wavelet coefficients (promoting sparsity as required).

The minimization of (3.3.9) is easily obtained, see again Section 3.2: at first, we may completely represent (3.3.9) by means of the associated wavelet coefficients,

$$\min_{\hat{u}_\lambda} \sum_{\lambda \in \Lambda} \{ (z_\lambda - u_\lambda)^2 + 2\alpha |u_\lambda| \} . \quad (3.3.10)$$

Since the wavelet basis is linearly independent, we can minimize summand-wise and obtain the following explicit expression for the optimum  $u_\lambda$ , see, e.g. [DDD04, DT04,

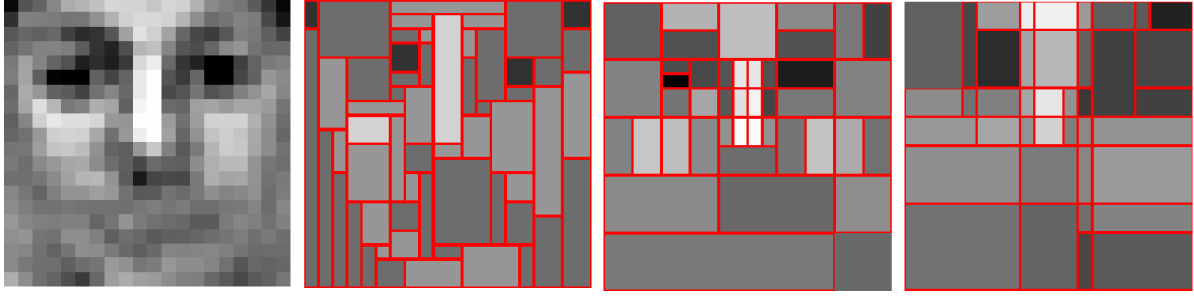


Figure 3.12: Example for a Reduced Set Vector (*from left to right*) and different Haar-like approximations: morphological filter [RRV05], standard wavelet transform and more optimal shifted wavelet transform

DT05],

$$u_\lambda = S_\alpha(z_\lambda) = \text{sign}(z_\lambda) \max\{|z_\lambda| - \alpha, 0\} , \quad (3.3.11)$$

where  $S_\alpha$  is the shorthand for the soft-shrinkage operation with threshold  $\alpha$ . Consequently, the optimum  $\mathbf{u}$  is simply obtained by soft-shrinking the wavelet coefficients of  $\mathbf{z}$ , i.e.

$$\mathbf{u} = S_\alpha(\mathbf{z}) = \sum_{\lambda \in \Lambda} S_\alpha(z_\lambda) \psi_\lambda.$$

As we have seen, so far we have involved just an individual wavelet basis. This, of course, allows a very fast wavelet representation of an image which is computed by fast discrete wavelet schemes. However, non-redundant representations and filtering very often creates artifacts in terms of undesirable oscillations or non-optimally represented details, which manifest themselves as ringing and edge blurring. For our purpose, it is essential to pick a representation that optimally meets the local image structure. The most promising method for adequately solving this kind of problem has its origin in translation invariance (the method of cycle spinning, see, e.g. [CD95]), i.e. representing the image by all possible shifted versions of the underlying (Haar) wavelet basis. But contrary to the idea of introducing redundancy by averaging over all possible representations of  $\mathbf{z}$  (i.e. really dealing with frames), we just aim to pick only that one which is optimally suited for our given image.

### Hyper-plane Approximation by Wavelet Shrinkage

Once we have approximated the Support Vectors of the SVM by the W-RSV's, the question arises whether the hyper-plane approximation  $\Psi''_{N_z} = \sum_{i=1}^{N_z} \beta_i \Phi(\mathbf{u}_i)$  is close to  $\Psi_{N_x}$ , i.e. we have consider the quantity

$$\|\Psi''_{N_z} - \Psi_{N_x}\| .$$

We firstly have computed the Reduced Set Vectors,  $\mathbf{z}_i$  by minimizing  $\|\Psi'_{N_z} - \Psi_{N_x}\|^2$  with respect to  $\mathbf{z}_i$  and to  $\beta_i$  (demonstrated in [RRV05] and [RTSB01]). Consequently, by triangle inequality it remains to show that we may reasonably control

$$\|\Psi''_{N_z} - \Psi'_{N_z}\| .$$

With the help of Cauchy-Schwarz and using a kernel function with  $k(\mathbf{z}_i, \mathbf{z}_i) = 1$  (such as the Gaussian Kernel 3.3.2 chosen in this paper) we obtain

$$\begin{aligned}
\|\Psi''_{N_z} - \Psi'_{N_z}\|^2 &\leq \sum_{i,j=1}^{N_z} |\beta_i| |\beta_j| \|\Phi(\mathbf{z}_i) - \Phi(\mathbf{u}_i)\| \|\Phi(\mathbf{z}_j) - \Phi(\mathbf{u}_j)\| \\
&= 2 \sum_{i,j=1}^{N_z} |\beta_i| |\beta_j| \sqrt{(1 - k(\mathbf{z}_i, \mathbf{u}_i))(1 - k(\mathbf{z}_j, \mathbf{u}_j))}. \\
&= 2 \left( \sum_{i=1}^{N_z} |\beta_i| \sqrt{(1 - k(\mathbf{z}_i, \mathbf{u}_i))} \right)^2. \tag{3.3.12}
\end{aligned}$$

Thus, we may control the approximation error for all  $i = 1, \dots, N_z$  by

$$1 - k(\mathbf{z}_i, \mathbf{u}_i) = 1 - \exp\left(\frac{-\|\mathbf{z}_i - \mathbf{u}_i\|^2}{2\sigma^2}\right) = \frac{\|\mathbf{z}_i - \mathbf{u}_i\|^2}{2\sigma^2} + \mathcal{O}(\|\cdot\|^4), \tag{3.3.13}$$

i.e. that the data misfit discrepancy is directly controlled by the  $L_2$  distance (which is minimized by (3.3.11) in the  $\ell_1$  sense) of the sparse approximation  $\mathbf{u}_i$  of  $\mathbf{z}_i$ . In other words, up to terms of higher order we have achieved the best approximation of the RVM under sparsity constraints, i.e.

$$\|\Psi''_{N_z} - \Psi'_{N_z}\| \leq \sigma^{-2} \sum_{i=1}^{N_z} |\beta_i| \|\mathbf{z}_i - \mathbf{u}_i\|. \tag{3.3.14}$$

Let us now consider more in detail the relation between the approximation error and the threshold parameter  $\alpha$ . This is important in order to control the trade-off between the sparsity and the approximation. To this end, observe that the argument in the Gaussian can be expressed by means of the corresponding wavelet coefficients, i.e.

$$\|\mathbf{z}_i - \mathbf{u}_i\|^2 = \sum_{\lambda \in \Lambda} (z_{i,\lambda} - S_\alpha(z_{i,\lambda}))^2,$$

and assuming  $\mathbf{z}$  has  $2^M \times 2^M$  pixel, we have the following bound

$$1 - k(\mathbf{z}_i, \mathbf{u}_i) \leq 1 - \prod_{\lambda \in \Lambda} \exp\left(\frac{-\alpha^2}{2\sigma^2}\right) = 1 - \exp\left(\frac{-2^{2M}\alpha^2}{2\sigma^2}\right).$$

Consequently, the worst case error is given by

$$\|\Psi''_{N_z} - \Psi'_{N_z}\|^2 \leq 2 \left[ 1 - \exp\left(\frac{-2^{2M}\alpha^2}{2\sigma^2}\right) \right] \left( \sum_{i=1}^{N_z} |\beta_i| \right)^2.$$

Let us denote this error bound by  $E$ . Then,  $E \rightarrow 0$  as  $\alpha \rightarrow 0$ , and for each  $\alpha > 0$ , we see that  $2(\sum_{i=1}^{N_z} |\beta_i|)^2 \geq E > 0$ . Hence, the price for sparsity (reduction of computational

complexity) is the approximation quality which can be easily controlled: assume a certain approximation error  $E$  is allowed, then the sparsity scaling constraint  $\alpha$  must fulfill

$$0 < \alpha^2 \leq -\ln \left( 1 - \frac{E}{2 \left[ \sum_{i=1}^{N_z} |\beta_i| \right]^2} \right)^{2\sigma^2 2^{-2M}}.$$

Neglecting the error terms of higher order, the worst case error reduces to

$$\|\Psi''_{N_z} - \Psi'_{N_z}\| \lesssim 2^M \alpha \sigma^{-1} \sum_{i=1}^{N_z} |\beta_i|$$

and thus,

$$0 < \alpha \leq \frac{\sigma 2^{-M} E}{\sum_{i=1}^{N_z} |\beta_i|}.$$

### Algorithmic Considerations

We first compute the Reduced Set Vectors,  $\mathbf{z}_i$  by minimizing  $\|\Psi'_{N_z} - \Psi_{N_x}\|^2$  like demonstrated in [RTSB01, RRV05]. Then we achieve the Wavelet Approximated Reduced Set Vectors (W-RSV's),  $\mathbf{u}_i^l$  using local best shift approximations of the RSV's,  $\mathbf{z}_i$  ( $i = 1, \dots, N_z$ ) at the coarse-to-fine approximation levels  $l = 1, \dots, N_l$ . This hierarchy of classifiers is obtained by first finding the first W-RSV,  $\mathbf{u}_1^l$  and  $\beta_1^l$  that minimizes  $\|\Psi_{N_z}^0 - \beta_1^l \Phi(\mathbf{u}_1^l)\|^2$  where  $\Psi_{N_z}^0 = \sum_{i=1}^{N_x} \alpha_i \Phi(\mathbf{x}_i)$  at the first approximation level  $l = 1$ . Then the Wavelet Approximated Reduced Set Vector  $\mathbf{u}_i^l$  is obtained by minimizing the distance  $\delta_i^l$  to the SVM hyper-plane

$$\delta_i^l = \|\Psi_{i-1}^l - \beta_i^l \Phi(\mathbf{u}_i^l)\|^2, \text{ where } \Psi_{i-1}^l = \Psi_{N_z}^{l-1} - \sum_{k=1}^{i-1} \beta_k^l \Phi(\mathbf{u}_k^l). \quad (3.3.15)$$

This evaluation continues until  $i = N_z$  and we start with the next finer approximation level  $l = l + 1$ , until  $l = N_l$  is reached.

1. Set  $\Psi_{N_z}^0 = \sum_{i=1}^{N_x} \alpha_i \Phi(\mathbf{x}_i)$  and  $\forall_{i=1, \dots, N_z} : \mathbf{r}_i^1 = \mathbf{z}_i$ , where  $\mathbf{z}_i$  are the Reduced Set Vectors.
2. Start at the first approximation level  $l = 1$ .
3. Start with (the residual of) the first RSV  $\mathbf{r}_i^l, i = 1$ .
4. Evaluate  $\forall_s : \tilde{\mathbf{u}}^s = (W^s)^{-1} S_\alpha (W^s \mathbf{r}_i^l)$  where  $W^s$  is the wavelet decomposition and  $(W^s)^{-1}$  the reconstruction with a shifted wavelet basis by the two dimensional shift  $s \in \{1, 2, \dots, 2^J\} \times \{1, 2, \dots, 2^J\}$ . For a  $20 \times 20$  patch size a shift  $J = 3$  is sufficient.  $S_\alpha$  is the Shrinkage function with the sparsity parameter  $\alpha$  (see ?? and ??).
5. Evaluate  $\forall_s : \Delta_\delta^s = \delta_{i-1}^l - \delta_i^l$  where  $\delta_0^l = \delta_{N_z}^{l-1}$  and the number of operations  $\Delta_\omega^s = 4 * \# [\tilde{\mathbf{u}}^s] + v(\tilde{\mathbf{u}}^s)$  where  $\# [\tilde{\mathbf{u}}^s]$  is the number of piecewise constant rectangles and  $v(\tilde{\mathbf{u}}^s)$  the number of gray values of  $\tilde{\mathbf{u}}^s$  (see 3.3.2).

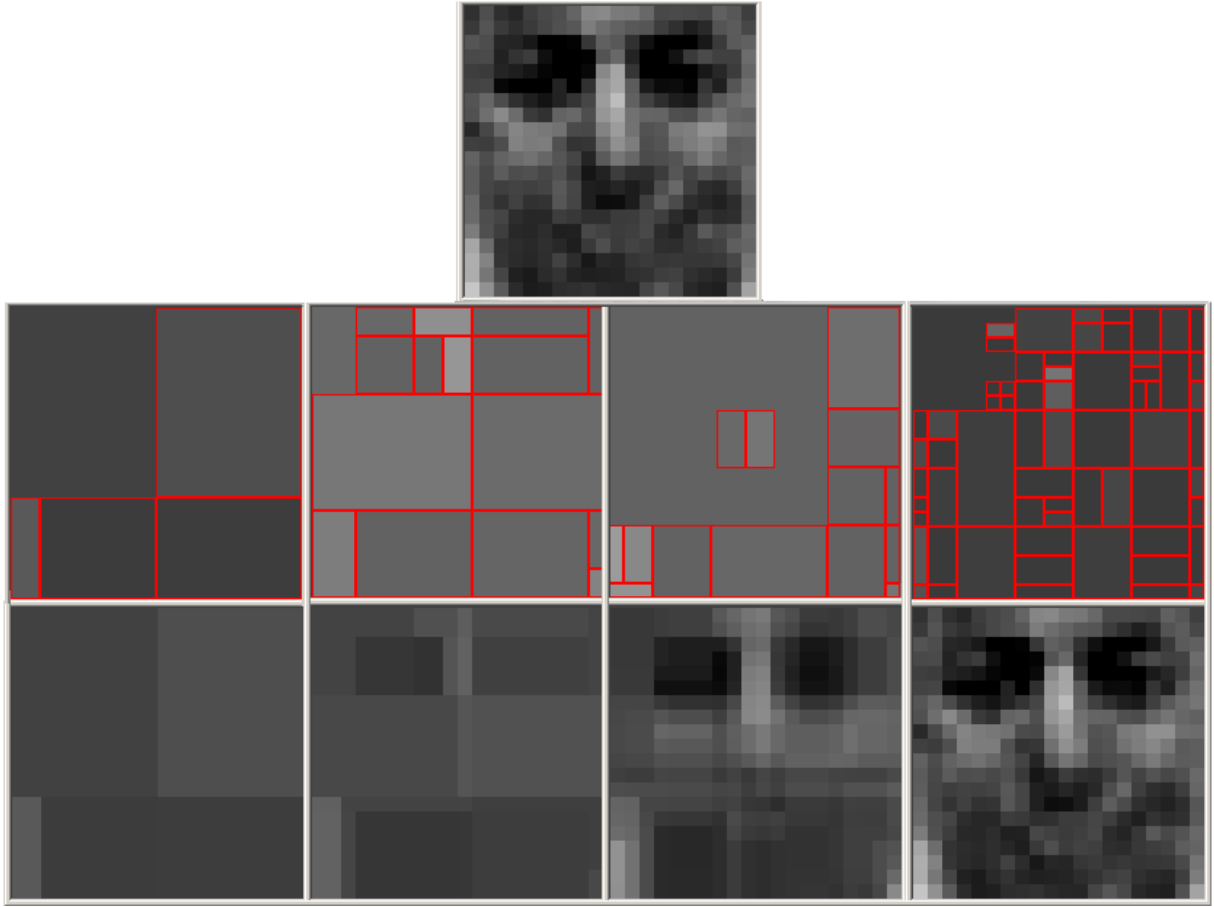


Figure 3.13: Example for a RSV  $\mathbf{z}_i$  (top) and its W-RSV  $\mathbf{u}_i^l$  at different approximation levels (middle row, left to right): e.g.  $\mathbf{u}_i^l$ ,  $l = 1, 2, 10, 19$ ). The bottom row, (left to right) shows the sum of the W-RSV's over the approximation levels (e.g.  $\sum_{l=1}^n \mathbf{u}_i^l$  with  $n = 1, 2, 10, 19$ )

6. Select the best shift  $s^*$ , where the decrease of the hyper-plane distance  $\Delta_\delta^s$  per needed number of operations  $\Delta_\omega^s$  is maximal.
7. Set  $\mathbf{u}_i^l = \tilde{\mathbf{u}}^{s^*}$  and save the rectangle structure for each approximation level of  $\mathbf{u}_i^l$  separately and set the new residual  $\mathbf{r}_i^{l+1} = \mathbf{r}_i^l - \mathbf{u}_i^l$ .
8. Set  $i = i + 1$  and goto 4. until  $i > N_z$ , otherwise set  $l = l + 1$  and goto 3. until  $l > N_l$ .

Using this algorithm we obtain from each RSV,  $\mathbf{z}_i$   $N_l$  levels of Wavelet Approximated RSV's,  $\mathbf{u}_i^l$  (see Figure 3.13 middle row). It is noticed that the approximation level  $l + 1$  of the W-RSV is not computed by a finer approximation of the original RSV,  $\mathbf{z}_i$  (e.g. by increasing sparsity parameter  $\alpha$ ). Instead the algorithm achieves the approximation  $\mathbf{u}_i^{l+1}$  from the residual  $\mathbf{r}_i^{l+1} = \mathbf{z}_i - \sum_{h=1}^l \mathbf{u}_i^h$ . Thus  $\sum_{l=1}^{N_l} \mathbf{u}_i^l$  converge to  $\mathbf{z}_i$  if  $N_l \rightarrow \infty$  (see

3.13 right column). We call it a local best shift method because the shift  $s^*$  is in general different for each approximation level. It is also noticed, that the rectangle structure of  $\mathbf{u}_i^l$  is evaluated and stored during the training and applied at the classification process for each  $l$  separately, because  $\# \left[ \sum_{h=1}^l \mathbf{u}_i^h \right] < \sum_{h=1}^l \# [\mathbf{u}_i^h]$ . As seen in 3.13 (*bottom row*) we obtain more rectangles, because the rectangles overlay by adding the approximations levels.

### Detection Process

To achieve high run-time efficiency we use a coarse-to-fine cascaded complexity of the classifier. The aim is an early rejection of easy to discriminate vectors (e.g. simple as non-faces to classify parts of the image). It is achieved by using an only as fine as necessary approximated classification hyper-plane by:

1. a hierarchical evaluation over the number  $i$  of W-RSV's  $\mathbf{u}_i^l$  (only as many approximated RSV's as necessary, similarly to [RTSB01]) and
2. a hierarchical evaluation over the levels of the approximation accuracy  $l$  of the W-RSV's  $\mathbf{u}_i^l$  (only a as fine as necessary approximation of the RSV's).

We realized this cascaded detection process using two loops. One inner loop over the number and an outer loop over the levels of the approximation accuracy of the RSV's. Thus, the classification of an image patch  $\mathbf{x}$  is expressed by the W-RVM classification function, with  $N_z^l$  ( $N_z^l \leq N_z$ ) W-RSV's,  $\mathbf{u}_i^l$  and their coefficients  $\beta^{l,i}$  for each approximation level,  $l = 1, \dots, N_l$  and the thresholds  $b_i^l$ , as follows:

$$y_i^l(\mathbf{x}) = \text{sign} \left( \sum_{h=1}^{l-1} \sum_{j=1}^{N_z^l} \beta_{h,j}^{l,i} k(\mathbf{x}, \mathbf{u}_j^h) + \sum_{j=1}^i \beta_{l,j}^{l,i} k(\mathbf{x}, \mathbf{u}_j^l) + b_i^l \right) \quad (3.3.16)$$

We gain the following detection algorithm:

1. Start at the first approximation level  $l = 1$
2. Start with the fist RSV,  $\mathbf{u}_1^l$  at the level  $l$
3. Evaluate  $y_i^l(\mathbf{x})$  for the input patch  $\mathbf{x}$  using (3.3.16)
4. If  $y_i^l < 0$  then the patch is classified as a non-face and the evaluation stops.
5. Set  $i = i + 1$  and goto 3. until  $i > N_z^l$ , otherwise set  $l = l + 1$  and goto 2. until  $l > N_l$ .

This hierarchical evaluation is a more efficient way to reject more locations by only few operations. The second reason is, that the adjustment of an optimal approximation accuracy was very sensitive in [RRV05], if only one approximation level is used. Now the evaluation applies automatically the most efficient approximation accuracy, dependent on the input image patch during the working process. In contrast to former methods the trade-off between accuracy and speed is very continuous.

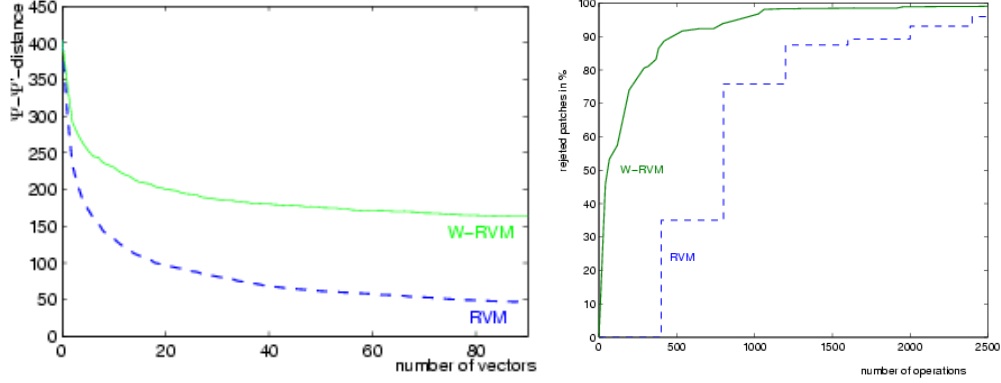


Figure 3.14: *Left:*  $\Psi_1 - \Psi'_{N_z}$  distance as function of the number of vectors  $N_z$  for the RVM (dashed line), and the W-RVM (solid line). *Right:* Percentage of rejected non-face patches as a function of the number of operations required.

However, the question arise whether to turn to the next approximation level ( $l++$ ) or to keep on incorporating more vectors at the current level ( $i++$ ). The optimum  $N_z^l$  for each level  $l$  in 3.3.16 can be evaluated once on a training set, by  $N_z^l = i$  if

$$\frac{\Delta_\omega(\mathbf{u}_{i+1}^l)}{r(\mathbf{u}_{i+1}^l)} > \frac{\Delta_\omega(\mathbf{u}_1^{l+1})}{r(\mathbf{u}_1^{l+1})},$$

where  $r$  is the number of rejections and  $\Delta_\omega$  the number of operations of  $\mathbf{u}_i^l$ . The number of approximation levels is adjusted by the shrinkage parameter  $\alpha$ . A low reduction of wavelets coefficients (fine approximation) causes more approximations levels. But the number of levels is not so decisive because more levels only mean that the classification process will change earlier to the next approximation level (smaller  $N_z^l$ , see above). Hence the method is not very sensitive concerning  $\alpha$ . Using this classification method we gain a minimal number of operations per rejection.

### 3.3.3 Numerical Verifications

For the face detector we used a training set that contains several thousand images downloaded from the World Wide Web. The training set includes 3500,  $20 \times 20$ , face patches and 20000 non-face patches and, the validation set, 1000 face patches, and 100,000 non-face patches. The SVM computed on the training set yielded about 8000 support vectors that we approximated by 90 Wavelet Approximated Reduced Set Vectors (W-RSV's) at five approximation levels by the method detailed in the previous section.

The first plot in Figure 3.14 shows the evolution of the approximation of the SVM by the RVM and by the W-RVM (in terms of the distance  $\Psi - \Psi'$ ) as a function of the number of vectors used. It can be seen that for a given accuracy more Wavelet Approximated Set Vectors are needed to approximate the SVM than for the RVM. However, as is seen of the second plot, for a given computational load, the W-RVM rejects much more non-face patches than the RVM. This explains the improved run-time performances of the W-RVM.

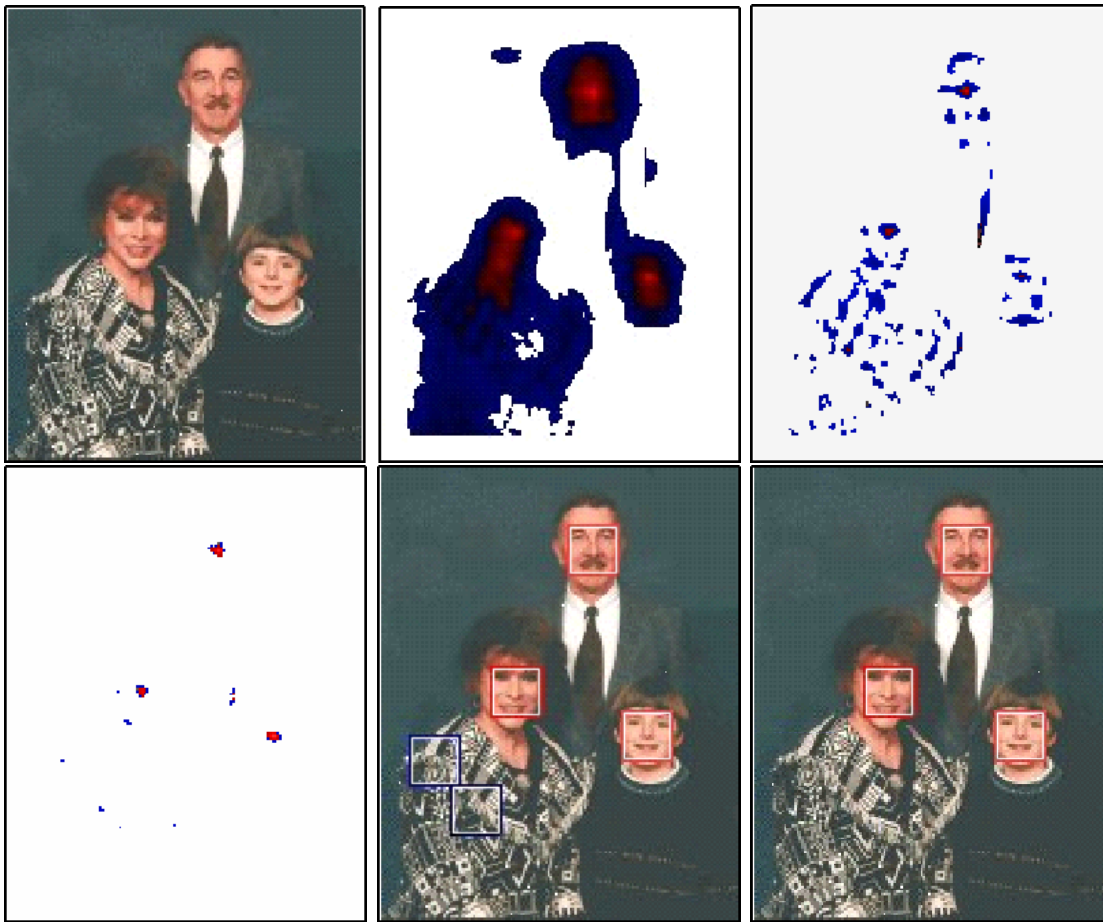


Figure 3.15: From top left to up right: input image followed by images showing the amount of rejected pixels at the 1<sup>st</sup>, 3<sup>rd</sup> and 50<sup>th</sup> stages of the cascade. The white pixels are rejected and the darkness of a pixel is proportional to the output of the W-RVM evaluation. The penultimate image shows a box around the pixels alive at the end of the the W-RSV's and the last image, after the full SVM is applied

Additionally, it can be seen that the curve is more smooth for the W-RVM, hence a better trade-off between accuracy and speed can be obtained by the W-RVM. Figure 3.15 shows an example of face detection in an image using the W-RVM. As the stages in the cascade increase fewer and fewer patches are evaluated. At the last W-RSV, only 5 pixels have to be classified using the full SVM.

Figure 3.16 shows the ROCs, computed on the validation set, of the SVM, the RVM and the W-RVM. It can be seen that the accuracies of the three classifiers are similar without (left plot) and almost equal with (right plot) the final SVM classification for the remaining patches.

Table 3.3 compares the accuracy and the average time required to evaluate the patches of the validation set. As can be seen, the novel W-RVM approach provides a significant speed-up (530-fold over the SVM and more than 15-fold over the RVM), for no substantial loss of accuracy.



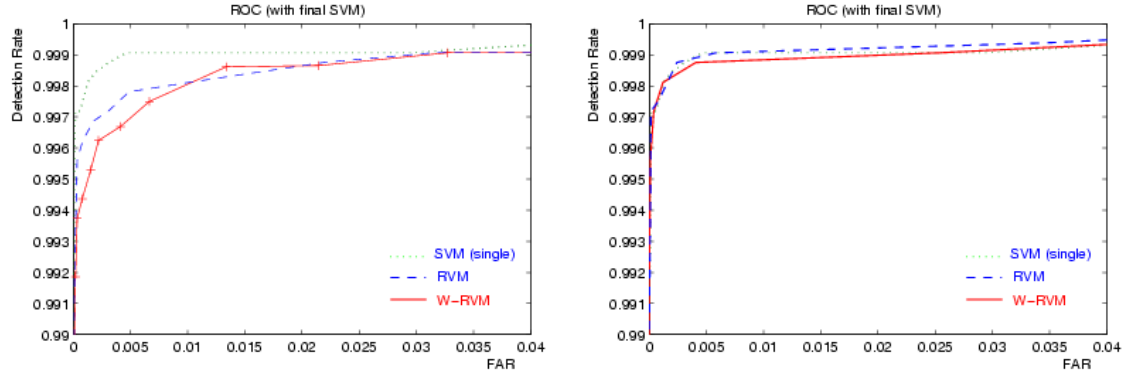


Figure 3.16: ROCs for a set of the SVM, the RVM and the W-RVM (*top*) without and (*bottom*) with the final SVM classification for the remaining patches. The FAR is related to non-face patches

method	FRR	FAR	time per patch
SVM	1.4%	0.002%	$787.34\mu s$
RVM	1.5%	0.001%	$22.51\mu s$
W-RVM	1.4%	0.002%	$1.48\mu s$

Table 3.3: Comparison of accuracy and speed improvement of the W-RVM to the RVM and SVM



## Chapter 4

# Nonlinear Operator Equations and Iterative Concepts

This chapter is devoted to the construction of Tikhonov-based iteration schemes for solving nonlinear operator equations.

### 4.1 Nonlinear Problems and Quadratic Constraints

We consider the computation of an approximation to a solution of a nonlinear operator equation

$$T(x) = y , \quad (4.1.1)$$

where  $T : X \rightarrow Y$  is an ill-posed operator between Hilbert spaces  $X, Y$ . If only noisy data  $y^\delta$  with

$$\|y^\delta - y\| \leq \delta \quad (4.1.2)$$

are available, problem (4.1.1) has to be stabilized by regularization methods. In recent years, many of the well known methods for linear ill-posed problems have been generalized to nonlinear operator equations. Unfortunately, it turns out that convergence and convergence rates can be shown only under severe restrictions to the operator for most methods. For example, convergence for Landweber method can be shown only if the operator fulfills

$$\|T(x) - T(\tilde{x}) - T'(\tilde{x})(x - \tilde{x})\| \leq \eta \|x - \tilde{x}\| \quad \text{with } \eta < 1/2 , \quad (4.1.3)$$

whereas convergence rates are only available if, for a solution  $x^\dagger$  of (4.1.1), there exists a family of bounded operators  $R_x$  with

$$T'(x) = R_x T'(x^\dagger) \quad \text{and} \quad \|I - R_x\| \leq K \|x - x^\dagger\| .$$

For other prominent iterative methods like Gauss-Newton [Bak92, BNS97], Levenberg-Marquardt [Han97a], conjugate gradient [Han97b] and Newton-like methods [Kal97, DES98], convergence can be shown under similar restrictions as (4.1.3). To obtain convergence rates, much stronger restrictions have to be assumed.

An alternative to the above mentioned iterative methods is Tikhonov regularization, where an approximation to the solution of (4.1.1) is obtained by minimizing the Tikhonov functional  $J_\alpha(x)$ ,

$$J_\alpha(x) = \|y^\delta - T(x)\|^2 + \alpha\|x - \bar{x}\|^2, \quad (4.1.4)$$

$$x_\alpha^\delta = \arg \min_x J_\alpha(x). \quad (4.1.5)$$

The advantage of Tikhonov regularization is that convergence of the method, i.e.  $x_\alpha^\delta \rightarrow x^\dagger$  for  $\delta \rightarrow 0$  and an appropriate parameter choice  $\alpha = \alpha(\delta)$  holds under weak assumptions to the operator, see, e.g., [EHN96a], and convergence rates are obtained for Fréchet differentiable operators with Lipschitz continuous derivative. However, the difficulties for Tikhonov regularization are a proper choice of the regularization parameter [Sch93, Ram02a] and the computation of the minimizer of the Tikhonov functional. As the functional is no longer convex for nonlinear operators  $T$ ,  $J_\alpha$  can even have local minimizers, and classical optimization routines might fail. Recently, we have introduced iterative methods for the minimization of the Tikhonov functional that reconstruct a global minimizer of the Tikhonov functional provided a smoothness assumption  $x^\dagger - \bar{x} = T'(x^\dagger)^*\omega$  with small  $\|\omega\|$  holds. We wish to remark that it might be difficult to show such smoothness conditions for practical problems, and for exponentially ill-posed problems Hölder-type smoothness conditions will not hold, see [Hoh97]. Thus it would be advantageous to construct iterative methods that reconstruct a minimizer of the Tikhonov functional under different assumptions. But this seems to remain a pipe dream: even here in this paper we had to incorporate some smoothness conditions to prove global minimizing properties of the reconstructed solution. However, all the here made assumptions on  $T$  are within the frame of nonlinear technologies and they are not that strong than for most of the above quoted iterative schemes.

Here we focus on the development of a method that *always* finds a *critical point* of the Tikhonov functional. Under additional assumptions on the operator and a smoothness condition on the solution we can then assure that this critical point is a *global minimizer* of  $J_\alpha$ .

The basic idea for our new iteration scheme goes as follows: consider the Tikhonov variational formulation of the inverse problem. Due to the nonlinearity, a direct reconstruction of the global minimizer is not possible. That's why we aim to solve instead of the pure Tikhonov functional a sequence of so-called surrogate or replacement functionals. This idea is borrowed from linear regularization methods with general and mixed smoothness constraints, see e.g. [DDD04, DT04, DT05]. The intention in [DDD04, DT04, DT05] is to decouple the variational equations with respect to the basis coefficients of the solution caused by the linear operator. The cost of dealing with a decoupled system of equations is an iteration process from which strong convergence properties can be shown. The situation in the nonlinear case is completely different and due to the impact of the Fréchet derivative one cannot expect to end up with similar schemes than in [DDD04, DT04, DT05]. However, the basic advantage of using replacement functionals is that each of the functionals is under certain conditions on the construction process globally convex. The minimization results then in an easy fixed point iteration. Defining now

an iteration process by iteratively minimizing a sequence of replacement functionals, we can show that the sequence of minimizing elements of each individual fixed point iteration converges in norm towards a critical point of the Tikhonov functional of the nonlinear inverse problem. Imposing additional assumptions (on the quadratic remainder of the Taylor series expression of our operator under consideration, and a smoothness condition) we obtain a uniqueness result, i.e. we are able to show that the reconstructed critical point is a global minimizer. Finally, applying a proper parameter choice rule, we are able to adopt classical convergence/order optimality results for Tikhonov regularization methods.

The remaining chapter is organized as follows: In Section 4.2.1, we state the scope of the problem. In Section 4.2.2, we explain how the replacement functionals are constructed and we minimize them in Section 4.2.3. The main result of the paper is presented in Section 4.2.4: strong convergence of the iterates towards a global minimizer. We end this paper with Section 5.3 in which we demonstrate the capabilities of the proposed scheme by solving the nonlinear SPECT problem.

### 4.1.1 Formulation of the Variational Problem

We consider the problem of deriving a minimizer of the Tikhonov functional

$$J_\alpha(x) = \|y^\delta - T(x)\|^2 + \alpha\|x - \bar{x}\|^2. \quad (4.1.6)$$

Due to the nonlinearity of the operator  $T$ , the minimizer of the functional might not be unique, or there might exist even local minimizers, such that a standard minimizing algorithm can fail in reconstructing a global minimizer. In order to obtain an easier problem which hopefully has a unique solution, we replace the functional  $J_\alpha$  by

$$J_\alpha^s(x, a) := \|y^\delta - T(x)\|^2 + \alpha\|x - \bar{x}\|^2 + C\|x - a\|^2 - \|T(x) - T(a)\|^2 \quad (4.1.7)$$

and proceed as follows:

1. Pick  $x_0$  and some proper constant  $C > 0$
2. Derive a sequence  $\{x_k\}_{k=0,1,\dots}$  by the iteration:

$$x_{k+1} = \arg \min_x J_\alpha^s(x, x_k) \quad k = 0, 1, 2, \dots$$

The overall goal of this paper is to show that the sequence  $\{x_k\}_{k=0,1,\dots}$  converges in norm topology towards a global minimizer of the Tikhonov functional (4.2.2).

In order to achieve this result we proceed in two steps: First, we aim to show norm convergence of the iterates  $x_k$  towards a critical point of the Tikhonov functional. In a second step, we verify that the reconstructed critical point is equal to a global minimizer of the Tikhonov functional. To make this program running, we have to restrict ourselves as follows:

- For the first step we limit the analysis to nonlinear operators  $T$  for which

$$x_k \xrightarrow{w} x \implies T(x_k) \rightarrow T(x) \text{ and } T'(x_k)^* z \rightarrow T'(x)^* z \text{ for all } z, \quad (4.1.8)$$

$$\|T'(x) - T'(\tilde{x})\| \leq L\|x - \tilde{x}\|. \quad (4.1.9)$$

It may happen that  $T$  already meets these conditions as an operator from  $X \rightarrow Y$ . If not, this can be achieved by assuming more regularity of the solution, i.e. we have to change the domain of  $T$  a little. To this end, let us assume that there exists a function space  $X^s$ , and a compact embedding operator  $i^s : X^s \rightarrow X$ . Now we can consider

$$\tilde{F} = F \circ i^s : X^s \longrightarrow Y.$$

We obtain

$$\|\tilde{T}'(x) - \tilde{T}'(\tilde{x})\| \leq L\|x - \tilde{x}\|_X \leq L\|x - \tilde{x}\|_{X^s}. \quad (4.1.10)$$

If now  $x_k \xrightarrow{w} x$  in  $X^s$ , then  $x_k \rightarrow x$  in  $X$  and, moreover, (4.1.10) yields  $\tilde{T}'(x_k) \rightarrow \tilde{T}'(x)$  and  $\tilde{T}'(x_k)^* \rightarrow \tilde{T}'(x)^*$  in the operator norm. This argument applies to arbitrary nonlinear continuous and Fréchet differentiable operators  $F : X \rightarrow Y$  with continuous Lipschitz derivative as long as a function space  $X^s$  with compact embedding  $i^s$  to  $X$  is available.

- To process the second step, we additionally impose that  $x^\dagger$  fulfills a smoothness condition,  $T$  is twice differentiable, and that

$$\|T(x) - T(\tilde{x}) - T'(\tilde{x})(x - \tilde{x})\| \leq \|T(x) - T(\tilde{x})\|, \quad (4.1.11)$$

which is a condition on the quadratic remainder of the Taylor series expansion of  $T$ .

### 4.1.2 Proper Surrogate Functionals

By the definition of the replacement or so-called surrogate functional  $J_\alpha^s$  in (4.2.4) it is not clear whether it is positive definite or even bounded from below. This will be clarified in this section, i.e. we will show that this is the case provided the constant  $C$  is chosen properly.

For given  $\alpha > 0$  and  $x_0$  we define a ball  $K_r(\bar{x})$  with radius  $r$  around  $\bar{x}$ , where the radius is given by

$$r^2 := \begin{cases} \frac{\|y^\delta - T(x_0)\|^2 + \alpha\|x_0 - \bar{x}\|^2}{\alpha} & \text{for } \alpha < 1 \\ \|y^\delta - T(x_0)\|^2 + \alpha\|x_0 - \bar{x}\|^2 & \text{for } \alpha \geq 1 \end{cases}. \quad (4.1.12)$$

This obviously ensures,  $x_0 \in K_r(\bar{x})$ . Furthermore, we define the constant  $C$  by

$$C := \max \left\{ 4, 2 \left( \sup_{x \in K_r(\bar{x})} \|T'(x)\| \right)^2, 2L \sqrt{\|y^\delta - T(x_0)\|^2 + \alpha\|x_0 - \bar{x}\|^2} \right\}, \quad (4.1.13)$$

where  $L$  is the Lipschitz constant of the Fréchet derivative of  $T$ . We assume that  $x_0$  was chosen such that  $r < \infty$  and  $C < \infty$ .

**Lemma 1** *Let  $r$  and  $C$  be chosen by (4.2.6), (4.2.7). Then*

$$C\|x - x_0\|^2 - \|T(x) - T(x_0)\|^2 \geq 0 \quad (4.1.14)$$

for all  $x \in K_r(\bar{x})$ , and, thus,  $J_\alpha(x) \leq J_\alpha^s(x, x_0)$ .

*Proof.* By Taylors expansion we have

$$T(x + h) = T(x) + \int_0^1 T'(x + \tau h)h \, d\tau$$

and thus

$$\|T(x) - T(x + h)\| \leq \int_0^1 \|T'(x + \tau h)\| \|h\| \, d\tau \leq \sup_{x \in K_r(\bar{x})} \|T'(x)\| \|h\| .$$

Consequently, we get for all  $x \in K_r(\bar{x})$

$$\begin{aligned} C\|x - x_0\|^2 - \|T(x) - T(x_0)\|^2 &\geq C\|x - x_0\|^2 - \left( \sup_{x \in K_r(\bar{x})} \|T'(x)\| \right)^2 \|x - x_0\|^2 \\ &= \frac{C}{2} \|x - x_0\|^2 \geq 0 , \end{aligned}$$

and the functional  $J_\alpha^s(x, x_0)$  is positive for all  $x \in K_r(\bar{x})$ . □

Next, we show that this carries over to all of the iterates:

**Proposition 4.1.1** *Let  $x_0, \alpha$  be given and  $r, C$  be defined by (4.2.6), (4.2.7). Then the functionals  $J_\alpha^s(x, x_k)$  are bounded from below for all  $k \in \mathbb{N}$  and have thus minimizers. For the minimizer  $x_{k+1}$  of  $J_\alpha^s(x, x_k)$  holds  $x_{k+1} \in K_r(\bar{x})$ .*

*Proof.* The proof will be done by induction. For  $k = 1$ , we show in a first step that  $J_\alpha^s(x, x_0)$  is bounded from below. We have

$$\|y^\delta - T(x)\|^2 = \|y^\delta - T(x_0)\|^2 + \|T(x_0) - T(x)\|^2 + 2\langle y^\delta - T(x_0), T(x_0) - T(x) \rangle . \quad (4.1.15)$$

Thus,

$$\begin{aligned} J_\alpha^s(x, x_0) - \alpha\|x - \bar{x}\|^2 &= \|y^\delta - T(x_0)\|^2 + 2\langle y^\delta - T(x_0), T(x_0) - T(x) \rangle + C\|x - x_0\|^2 \\ &\geq \|y^\delta - T(x_0)\|^2 - 2\|y^\delta - T(x_0)\| \|T(x_0) - T(x)\| + C\|x - x_0\|^2 . \end{aligned} \quad (4.1.16)$$

Again by Taylor expansion, we get

$$\|T(x_0) - T(x)\| \leq \|T'(x_0)\| \|x_0 - x\| + \frac{L}{2} \|x_0 - x\|^2 . \quad (4.1.17)$$

Now let us assume that  $J_\alpha^s(x, x_0)$  is not bounded from below. As  $T$  is continuous, there exists a sequence  $\{x_l\}_{l \in \mathbb{N}}$  with  $\|x_l\| \rightarrow \infty$  and  $J_\alpha^s(x_l, x_0) \rightarrow -\infty$ . In particular, for  $l$  large enough, follows from (4.2.12)

$$\|T(x_0) - T(x_l)\| \leq L\|x_0 - x_l\|^2 ,$$

and combining this estimate with (4.2.11) yields

$$J_\alpha^s(x_l, x_0) - \alpha \|x_l - \bar{x}\|^2 \geq \|y^\delta - T(x_0)\|^2 - 2L\|y^\delta - T(x_0)\| \|x_l - x_0\|^2 + C\|x_l - x_0\|^2.$$

From the definition of  $C$  in (4.2.7) follows  $2L\|y^\delta - T(x_0)\| \leq C$  and thus

$$J_\alpha^s(x_l, x_0) - \alpha \|x_l - \bar{x}\|^2 \geq \|y^\delta - T(x_0)\|^2 \geq 0,$$

in contradiction to our assumption  $J_\alpha^s(x_l, x_0) \rightarrow -\infty$ , and thus  $J_\alpha^s(x, x_0)$  is bounded from below. By the same argument, we find  $J_\alpha^s(x_l, x_0) \geq \alpha \|x_l - \bar{x}\|^2 \rightarrow \infty$  for any sequence  $x_l$  with  $\|x_l\| \rightarrow \infty$  and thus the functional is coercive and has a minimizer  $x_1$ .

As in (4.2.11), we get by using (4.2.12)

$$\begin{aligned} J_\alpha^s(x_1, x_0) - \alpha \|x_1 - \bar{x}\|^2 &\geq \|y^\delta - T(x_0)\|^2 + 2\langle y^\delta - T(x_0), T(x_0) - T(x_1) \rangle + C\|x_1 - x_0\|^2 \\ &\geq \|y^\delta - T(x_0)\|^2 - 2\|y^\delta - T(x_0)\| \|T'(x_0)\| \|x_1 - x_0\| \\ &\quad - L\|y^\delta - T(x_0)\| \|x_1 - x_0\|^2 + C\|x_1 - x_0\|^2 \end{aligned}$$

By (4.2.7), we have  $C/2 \geq L\|y^\delta - T(x_0)\|$ , and thus

$$J_\alpha^s(x_1, x_0) - \alpha \|x_1 - \bar{x}\|^2 \geq \|y^\delta - T(x_0)\|^2 - 2\|y^\delta - T(x_0)\| \|T'(x_0)\| \|x_1 - x_0\| + \frac{C}{2} \|x_1 - x_0\|^2. \quad (4.1.18)$$

As  $x_0 \in K_r(\bar{x})$ , it follows from (4.2.7) that  $\|T'(x_0)\| \leq \sqrt{C/2}$  holds, and we get finally

$$\begin{aligned} J_\alpha^s(x_1, x_0) - \alpha \|x_1 - \bar{x}\|^2 &\geq \|y^\delta - T(x_0)\|^2 - 2\frac{\sqrt{C}}{\sqrt{2}} \|y^\delta - T(x_0)\| \|x_1 - x_0\| + \frac{C}{2} \|x_1 - x_0\|^2 \\ &= \left( \|y^\delta - T(x_0)\| - \frac{\sqrt{C}}{\sqrt{2}} \|x_1 - x_0\| \right)^2 \geq 0. \end{aligned} \quad (4.1.19)$$

In particular, it follows for  $\alpha < 1$

$$\begin{aligned} \alpha \|x_1 - \bar{x}\|^2 &\stackrel{(4.2.13)}{\leq} J_\alpha^s(x_1, x_0) = \min_x J_\alpha^s(x, x_0) \leq J_\alpha^s(x_0, x_0) \\ &= \|y^\delta - T(x_0)\|^2 + \alpha \|x_0 - \bar{x}\|^2, \end{aligned}$$

i.e.

$$\|x_1 - \bar{x}\|^2 \leq \frac{\|y^\delta - T(x_0)\|^2 + \alpha \|x_0 - \bar{x}\|^2}{\alpha} = r^2,$$

and for  $\alpha \geq 1$

$$\begin{aligned} \|x_1 - \bar{x}\|^2 &\leq \alpha \|x_1 - \bar{x}\|^2 \stackrel{(4.2.13)}{\leq} J_\alpha^s(x_1, x_0) \leq J_\alpha^s(x_0, x_0) \\ &= \|y^\delta - T(x_0)\|^2 + \alpha \|x_0 - \bar{x}\|^2 = r^2, \end{aligned}$$

and thus  $x_1 \in K_r(\bar{x})$ .

By Lemma 11 it follows that  $C\|x_1 - x_0\|^2 - \|T(x_1) - T(x_0)\|^2 \geq 0$  and  $J_\alpha(x) \leq J_\alpha^s(x, x_0)$ , and we get

$$\|y^\delta - T(x_1)\|^2 \leq J_\alpha(x_1) \leq J_\alpha^s(x_1, x_0) \leq J_\alpha^s(x_0, x_0) \leq \|y^\delta - T(x_0)\|^2 + \alpha \|x_0 - \bar{x}\|^2,$$



and combining this estimate with the definition of  $C$  in (4.2.7) yields

$$2L\|y^\delta - T(x_1)\| \leq 2L\sqrt{\|y^\delta - T(x_0)\|^2 + \alpha\|x_0 - \bar{x}\|^2} \leq C. \quad (4.1.20)$$

Now let us assume that the following properties hold for all  $i = 1, \dots, k-1$ :

$$x_i \in K_r(\bar{x}) \quad (4.1.21)$$

$$C\|x_i - x_{i-1}\| - \|T(x_i) - T(x_{i-1})\| \geq 0 \quad (4.1.22)$$

$$2L\|y^\delta - T(x_i)\| \leq C, \quad (4.1.23)$$

where  $x_i$  denotes a minimizer of the functional  $J_\alpha^s(x, x_{i-1})$ . For  $i = 1$ , these properties have already been shown. As for the case  $i = 1$ , we have to show that the functional  $J_\alpha^s(x, x_{k-1})$  has a minimizer. First, we show that it is bounded from below: As in (4.2.11) we get

$$J_\alpha^s(x, x_{k-1}) - \alpha\|x - \bar{x}\|^2 \geq \|y^\delta - T(x_{k-1})\|^2 - 2\|y^\delta - T(x_{k-1})\|\|T(x_{k-1}) - T(x)\| + C\|x - x_{k-1}\|^2 \quad (4.1.24)$$

By Taylor expansion, we get

$$\|T(x_{k-1}) - T(x)\| \leq \|T'(x_{k-1})\|\|x_{k-1} - x\| + \frac{L}{2}\|x_{k-1} - x\|^2. \quad (4.1.25)$$

Now let us assume that  $J_\alpha^s(x, x_{k-1})$  is not bounded from below. As  $T$  is continuous, there exists a sequence  $\{x_l\}_{l \in \mathbb{N}}$  with  $\|x_l\| \rightarrow \infty$  and  $J_\alpha^s(x_l, x_{k-1}) \rightarrow -\infty$ . In particular, for  $l$  large enough, follows from (4.2.17)

$$\|T(x_{k-1}) - T(x_l)\| \leq L\|x_{k-1} - x_l\|^2,$$

and combining this estimate with (4.2.17) yields

$$J_\alpha^s(x_l, x_{k-1}) - \alpha\|x_l - \bar{x}\|^2 \geq \|y^\delta - T(x_{k-1})\|^2 - 2L\|y^\delta - T(x_{k-1})\|\|x_l - x_{k-1}\|^2 + C\|x_l - x_{k-1}\|^2.$$

From (4.2.16) follows  $2L\|y^\delta - T(x_{k-1})\| \leq C$  and thus

$$J_\alpha^s(x_l, x_{k-1}) - \alpha\|x_l - \bar{x}\|^2 \geq \|y^\delta - T(x_{k-1})\|^2 \geq 0,$$

in contradiction to our assumption  $J_\alpha^s(x_l, x_{k-1}) \rightarrow -\infty$ , and thus  $J_\alpha^s(x, x_{k-1})$  is bounded from below. By the same argument, we find  $J_\alpha^s(x_l, x_{k-1}) \geq \alpha\|x_l - \bar{x}\|^2 \rightarrow \infty$ , for any sequence  $x_l$  with  $\|x_l\| \rightarrow \infty$  and thus the functional is coercive and has a minimizer  $x_k$ . As in (4.2.17), we get by using (4.2.17)

$$\begin{aligned} J_\alpha^s(x_k, x_{k-1}) - \alpha\|x_k - \bar{x}\|^2 &\geq \|y^\delta - T(x_{k-1})\|^2 + 2\langle y^\delta - T(x_{k-1}), T(x_{k-1}) - T(x_k) \rangle \\ &\quad + C\|x_k - x_{k-1}\|^2 \\ &\geq \|y^\delta - T(x_{k-1})\|^2 - 2\|y^\delta - T(x_{k-1})\|\|T'(x_{k-1})\|\|x_k - x_{k-1}\| \\ &\quad - L\|y^\delta - T(x_{k-1})\|\|x_k - x_{k-1}\|^2 + C\|x_k - x_{k-1}\|^2. \end{aligned}$$

By (4.2.7) and assumption (4.2.16) we have  $C/2 \geq L\|y^\delta - T(x_{k-1})\|$ , and thus

$$\begin{aligned} J_\alpha^s(x_k, x_{k-1}) - \alpha\|x_k - \bar{x}\|^2 &\geq \|y^\delta - T(x_{k-1})\|^2 - 2\|y^\delta - T(x_{k-1})\|\|T'(x_{k-1})\|\|x_k - x_{k-1}\| \\ &\quad + \frac{C}{2}\|x_k - x_{k-1}\|^2. \end{aligned}$$

As  $x_{k-1} \in K_r(\bar{x})$ , it follows from (4.2.7) that  $\|T'(x_{k-1})\| \geq \sqrt{C/2}$  holds, and we get finally

$$\begin{aligned} J_\alpha^s(x_k, x_{k-1}) - \alpha \|x_k - \bar{x}\|^2 &\geq \|y^\delta - T(x_{k-1})\|^2 - 2 \frac{\sqrt{C}}{\sqrt{2}} \|y^\delta - T(x_{k-1})\| \|x_k - x_{k-1}\| \\ &\quad + \frac{C}{2} \|x_k - x_{k-1}\|^2 \\ &= \left( \|y^\delta - T(x_{k-1})\| - \frac{\sqrt{C}}{\sqrt{2}} \|x_k - x_{k-1}\| \right)^2 \geq 0. \end{aligned} \quad (4.1.26)$$

In particular, it follows for  $\alpha < 1$  by assumption (4.2.15)

$$\begin{aligned} \alpha \|x_k - \bar{x}\|^2 &\stackrel{(4.2.18)}{\leq} J_\alpha^s(x_k, x_{k-1}) = \min_x J_\alpha^s(x, x_{k-1}) \leq J_\alpha^s(x_{k-1}, x_{k-1}) \\ &= \|y^\delta - T(x_{k-1})\|^2 + \alpha \|x_{k-1} - \bar{x}\|^2 \\ &\leq \|y^\delta - T(x_{k-1})\|^2 + \alpha \|x_{k-1} - \bar{x}\|^2 \\ &\quad + C \|x_{k-1} - x_{k-2}\|^2 - \|T(x_{k-1}) - T(x_{k-2})\|^2 \\ &= J_\alpha^s(x_{k-1}, x_{k-2}) \leq J_\alpha^s(x_{k-2}, x_{k-2}) \leq \cdots \leq J_\alpha^s(x_0, x_0) \\ &= \|y^\delta - T(x_0)\|^2 + \alpha \|x_0 - \bar{x}\|^2 \end{aligned}$$

i.e.

$$\|x_k - \bar{x}\|^2 \leq \frac{\|y^\delta - T(x_0)\|^2 + \alpha \|x_0 - \bar{x}\|^2}{\alpha} \leq r^2,$$

and in the same way follows for  $\alpha \geq 1$

$$\begin{aligned} \|x_k - \bar{x}\|^2 \leq \alpha \|x_k - \bar{x}\|^2 &\stackrel{(4.2.18)}{\leq} J_\alpha^s(x_k, x_{k-1}) \leq J_\alpha^s(x_{k-1}, x_{k-1}) \leq \cdots \leq J_\alpha^s(x_0, x_0) \\ &= \|y^\delta - T(x_0)\|^2 + \alpha \|x_0 - \bar{x}\|^2 \leq r^2, \end{aligned}$$

and thus  $x_k \in K_r(\bar{x})$ .

As in Lemma 11, it follows  $C\|x_k - x_{k-1}\|^2 - \|T(x_k) - T(x_{k-1})\|^2 \geq 0$  and  $J_\alpha(x) \leq J_\alpha^s(x, x_{k-1})$ , and we get

$$\begin{aligned} \|y^\delta - T(x_k)\|^2 &\leq J_\alpha(x_k) \leq J_\alpha^s(x_k, x_{k-1}) \leq J_\alpha^s(x_{k-1}, x_{k-1}) \leq \cdots \leq J_\alpha^s(x_0, x_0) \\ &= \|y^\delta - T(x_0)\|^2 + \alpha \|x_0 - \bar{x}\|^2, \end{aligned} \quad (4.1.27)$$

and combining this estimate with the definition of  $C$  (4.2.7) yields

$$2L\|y^\delta - T(x_k)\| \leq 2L\sqrt{\|y^\delta - T(x_0)\|^2 + \alpha\|x_0 - \bar{x}\|^2} \leq C, \quad (4.1.28)$$

i.e. we have shown that the assumptions (4.2.14)-(4.2.16) hold also for  $i = k$ .  $\square$

**Corollary 2** *The sequences of functionals  $\{J_\alpha(x_k)\}_{k=0,1,2,\dots}$  and  $\{J_\alpha^s(x_{k+1}, x_k)\}_{k=0,1,2,\dots}$  are non-increasing.*

*Proof.* This follows now by  $J_\alpha(x_{k+1}) \leq J_\alpha^s(x_{k+1}, x_k) \leq J_\alpha^s(x_k, x_k) = J_\alpha(x_k) \leq J_\alpha^s(x_k, x_{k-1})$ .  $\square$

### 4.1.3 Minimization of Surrogate Functionals

In this section, we elaborate necessary conditions for a minimizer of the functional  $J_\alpha^s(x, x_{k-1})$ . Moreover, we prove that  $J_\alpha^s(x, x_k)$  is globally convex for each  $k = 0, 1, 2, \dots$ .

**Lemma 3** *The derivative  $DJ_\alpha^s(x, a)h$  of  $J_\alpha^s(x, a)$  is given by*

$$DJ_\alpha^s(x, a)h = -2\langle T'(x)^*(y^\delta - T(a)) + (Ca + \alpha\bar{x}) - (C + \alpha)x, h \rangle. \quad (4.1.29)$$

*Proof.* It is

$$J_\alpha^s(x + h, a) = \|y^\delta - T(x + h)\|^2 + \alpha\|x - \bar{x} + h\|^2 + C\|x - a + h\|^2 - \|T(x + h) - T(a)\|^2.$$

By Taylor's expansion,  $T(x + h) = T(x) + T'(x)h + O(\|h\|^2)$ , we get

$$\begin{aligned} J_\alpha^s(x + h, a) &= \|y^\delta - T(x) - T'(x)h + O(\|h\|^2)\|^2 + \alpha\|x - \bar{x} + h\|^2 + C\|x - a + h\|^2 \\ &\quad - \|T(x) - T(a) + T'(x)h + O(\|h\|^2)\|^2 \\ &= \|y^\delta - T(x)\|^2 + \|T'(x)h\|^2 - 2\langle y^\delta - T(x), T'(x)h \rangle \\ &\quad + \alpha(\|x - \bar{x}\|^2 + \|h\|^2 + 2\langle x - \bar{x}, h \rangle) + C(\|x - a\|^2 + \|h\|^2 + 2\langle x - a, h \rangle) \\ &\quad - (\|T(x) - T(a)\|^2 + \|T'(x)h\|^2 + 2\langle T(x) - T(a), T'(x)h \rangle) + O(\|h\|^2). \end{aligned}$$

It follows

$$\begin{aligned} \frac{J_\alpha^s(x + h, a) - J_\alpha^s(x, a)}{2} &= -\langle T'(x)^*(y^\delta - T(x)), h \rangle + \alpha\langle x - \bar{x}, h \rangle + C\langle x - a, h \rangle \\ &\quad - \langle T'(x)^*(T(x) - T(a)), h \rangle + O(\|h\|^2) \\ &= -\langle T'(x)^*(y^\delta - T(a)) - \alpha(x - \bar{x}) - C(x - a), h \rangle + O(\|h\|^2) \\ &= -\langle T'(x)^*(y^\delta - T(a)) - (C + \alpha)x + \alpha\bar{x} + Ca, h \rangle + O(\|h\|^2). \end{aligned}$$

and thus the derivative is given by (4.1.29).  $\square$

The necessary condition for a minimum of (4.2.4) thus reads as

$$x = \underbrace{\frac{1}{C + \alpha} (T'(x)^*(y^\delta - T(a)) + \alpha\bar{x} + Ca)}_{=: \Phi_\alpha(x, a)}. \quad (4.1.30)$$

To minimize (4.2.4), we will use a fixed point iteration for  $\Phi_\alpha(x, a)$ . As  $J_\alpha^s(x, a)$  has by Proposition 4.2.1 a minimizer, (4.1.30) has at least one fixed point. It remains to show that  $\Phi_\alpha(x, a)$  is a contraction operator:

**Lemma 4** *The operator  $\Phi_\alpha(x, a)$  is a contraction, i.e.  $\|\Phi_\alpha(x, a) - \Phi_\alpha(\tilde{x}, a)\| \leq q\|x - \tilde{x}\|$ , if*

$$q := \frac{L}{C + \alpha} \sqrt{J_\alpha(a)} < 1.$$

*Proof.* We have  $\Phi_\alpha(x, a) - \Phi_\alpha(\tilde{x}, a) = \frac{1}{C+\alpha}(T'(x) - T'(\tilde{x}))^*(y^\delta - T(a))$ , and by using the Lipschitz-continuity of  $T'$  we get

$$\begin{aligned} \|\Phi_\alpha(x, a) - \Phi_\alpha(\tilde{x}, a)\| &= \frac{1}{C+\alpha} \|T'(x) - T'(\tilde{x})\| \|y^\delta - T(a)\| \\ &\leq \frac{L}{C+\alpha} \|y^\delta - T(a)\| \|x - \tilde{x}\| \leq \frac{L}{C+\alpha} \sqrt{J_\alpha(a)} \|x - \tilde{x}\|. \end{aligned}$$

□

**Proposition 4.1.2** *In our algorithm, the operator  $\Phi_\alpha(x, x_k)$  is for all  $k = 0, 1, 2, \dots$  and all  $\alpha \geq 0$  a contraction.*

*Proof.* By the definition of  $C$  in (4.2.7), Lemma 15 (setting  $a = x_0$ ), we deduce that  $\Phi_\alpha(x, x_0)$  is a contraction with

$$q = \frac{L}{C+\alpha} \sqrt{J_\alpha(x_0)} = \frac{C}{2(C+\alpha)} \leq \frac{1}{2} < 1.$$

With the help of Corollary 12, we complete the proof

$$\begin{aligned} \|\Phi_\alpha(x, x_k) - \Phi_\alpha(\tilde{x}, x_k)\| &\leq \frac{L}{C+\alpha} \sqrt{J_\alpha(x_k)} \|x - \tilde{x}\| \leq \frac{L}{C+\alpha} \sqrt{J_\alpha(x_{k-1})} \|x - \tilde{x}\| \\ &\leq \dots \frac{L}{C+\alpha} \sqrt{J_\alpha(x_0)} \|x - \tilde{x}\|. \end{aligned} \quad (4.1.31)$$

□

Up to here, we do know that our fixed point iteration for (4.1.30) converges towards a critical point of  $J_\alpha^s(x, x_k)$ .

**Proposition 4.1.3** *The necessary equation (4.1.30) for a minimum of the functional  $J_\alpha^s(x, x_k)$  has a unique fixed point, and the fixed point iteration converges towards the minimizer.*

*Proof.* To prove this Proposition, we have to investigate the Taylor expansion of  $J_\alpha^s$  more closely. By Taylor's expansion for  $T$  and the Lipschitz-continuity of  $T'$  we get

$$T(x+h) = T(x) + T'(x)h + R(x, h) \quad (4.1.32)$$

with

$$\|R(x, h)\| \leq \frac{L}{2} \|h\|^2. \quad (4.1.33)$$

As in the proof of Lemma 13 we get

$$\begin{aligned} J_\alpha^s(x+h, x_k) &= J_\alpha^s(x, x_k) + DJ_\alpha^s(x, x_k)h \\ &\quad - 2\langle y^\delta - T(x), R(x, h) \rangle - 2\langle T(x) - T(x_k), R(x, h) \rangle + (\alpha + C)\|h\|^2 \\ &= J_\alpha^s(x, x_k) + DJ_\alpha^s(x, x_k)h - 2\langle y^\delta - T(x_k), R(x, h) \rangle + (\alpha + C)\|h\|^2, \end{aligned} \quad (4.1.34)$$

and by using  $C \geq 2L\|y^\delta - T(x_k)\|$  follows

$$\begin{aligned} -2\langle y^\delta - T(x_k), R(x, h) \rangle + (\alpha + C)\|h\|^2 &\geq -2\|y^\delta - T(x_k)\|\|R(x, h)\| + (\alpha + C)\|h\|^2 \\ &\geq (-L\|y^\delta - T(x_k)\| + \alpha + C)\|h\|^2 \\ &\geq (C/2 + \alpha)\|h\|^2. \end{aligned} \quad (4.1.35)$$

Now assume  $\tilde{x}$  is a critical point of  $J_\alpha^s$ , i.e.  $DJ_\alpha^s(\tilde{x}, x_k)h = 0$  for all  $h$ . Consequently, by (4.1.34), (4.1.35) we have

$$J_\alpha^s(\tilde{x} + h, x_k) \geq J_\alpha^s(\tilde{x}, x_k) + (C/2 + \alpha)\|h\|^2,$$

and in particular

$$J_\alpha^s(\tilde{x} + h, x_k) > J_\alpha^s(\tilde{x}, x_k) \quad \text{for all } h \neq 0. \quad (4.1.36)$$

Thus, every critical point is a global minimizer of  $J_\alpha^s(x, x_k)$ , and, again by (4.1.36), there exists only one global minimizer.  $\square$

By assuming more regularity on  $T$  it is possible to sharpen the above given statement:

**Proposition 4.1.4** *Let  $T$  be a twice continuously differentiable operator. Then the functional  $J_\alpha^s(x, x_k)$  is strictly convex.*

*Proof.* With a slight abuse of notation we set  $J_\alpha^s(x) := J_\alpha^s(x, x_k)$ . By (4.1.34) we have

$$J_\alpha^s(x + h) = J_\alpha^s(x) + DJ_\alpha^s(x)h + g_\alpha(x, h), \quad (4.1.37)$$

where  $g_\alpha(x, h)$  is defined by

$$g_\alpha(x, h) := -2\langle y^\delta - T(x_k), R(x, h) \rangle + (\alpha + C)\|h\|^2. \quad (4.1.38)$$

For strict convexity, we have to show that

$$J_\alpha^s((1 - \lambda)x_1 + \lambda x_2) < (1 - \lambda)J_\alpha^s(x_1) + \lambda J_\alpha^s(x_2)$$

holds for  $\lambda \in (0, 1)$  and arbitrary  $x_1, x_2$ . We have

$$\begin{aligned} J_\alpha^s((1 - \lambda)x_1 + \lambda x_2) &= J_\alpha^s(x_1 + \lambda(x_2 - x_1)) = J_\alpha^s(x_2 + (1 - \lambda)(x_1 - x_2)) \\ &= (1 - \lambda)J_\alpha^s(x_1 + \lambda(x_2 - x_1)) + \lambda J_\alpha^s(x_2 + (1 - \lambda)(x_1 - x_2)) \end{aligned} \quad (4.1.39)$$

and with

$$\begin{aligned} J_\alpha^s(x_1 + \lambda(x_2 - x_1)) &= J_\alpha^s(x_1) + \lambda DJ_\alpha^s(x_1)(x_2 - x_1) \\ &\quad + g_\alpha(x_1, \lambda(x_2 - x_1)) \\ J_\alpha^s(x_2 + (1 - \lambda)(x_1 - x_2)) &= J_\alpha^s(x_2) + (1 - \lambda)DJ_\alpha^s(x_2)(x_1 - x_2) \\ &\quad + g_\alpha(x_2, (1 - \lambda)(x_1 - x_2)) \end{aligned}$$

we obtain

$$\begin{aligned} J_\alpha^s((1-\lambda)x_1 + \lambda x_2) &= (1-\lambda)J_\alpha^s(x_1) + \lambda J_\alpha^s(x_2) \\ &\quad + \lambda(1-\lambda)[DJ_\alpha^s(x_1) - DJ_\alpha^s(x_2)](x_2 - x_1) \\ &\quad + (1-\lambda)g_\alpha(x_1, \lambda(x_2 - x_1)) + \lambda g_\alpha(x_2, (1-\lambda)(x_1 - x_2)). \end{aligned}$$

Thus  $J_\alpha^s$  is strict convex if for all  $\lambda \in (0, 1)$

$$\begin{aligned} D(x_1, x_2, \lambda) &:= \lambda(1-\lambda)[DJ_\alpha^s(x_1) - DJ_\alpha^s(x_2)](x_2 - x_1) \\ &\quad + (1-\lambda)g_\alpha(x_1, \lambda(x_2 - x_1)) + \lambda g_\alpha(x_2, (1-\lambda)(x_1 - x_2)) < 0. \end{aligned}$$

We have

$$\begin{aligned} \frac{DJ_\alpha^s(x_1) - DJ_\alpha^s(x_2)}{2}(x_2 - x_1) &= -\langle T'(x_1)^*(y^\delta - T(x_k)) + Cx_k + \alpha\bar{x} - (C + \alpha)x_1, x_2 - x_1 \rangle \\ &\quad + \langle T'(x_2)^*(y^\delta - T(x_k)) + Cx_k + \alpha\bar{x} - (C + \alpha)x_2, x_2 - x_1 \rangle \\ &= -(C + \alpha)\|x_2 - x_1\|^2 \\ &\quad - \langle (T'(x_1) - T'(x_2))^*(y^\delta - T(x_k)), x_2 - x_1 \rangle \\ &= -(C + \alpha)\|x_2 - x_1\|^2 \\ &\quad - \langle y^\delta - T(x_k), T'(x_1) - T'(x_2)(x_2 - x_1) \rangle. \end{aligned}$$

As  $T$  is twice continuously Fréchet differentiable, it is

$$T'(x_1) = T'(x_2) + \int_0^1 T''(x_2 + \tau(x_1 - x_2))(x_1 - x_2, \cdot) d\tau$$

and thus,

$$\begin{aligned} [DJ_\alpha^s(x_1) - DJ_\alpha^s(x_2)](x_2 - x_1) &= -2(C + \alpha)\|x_2 - x_1\|^2 + \\ &\quad 2\langle y^\delta - T(x_k), \int_0^1 T''(x_2 + \tau(x_1 - x_2))(x_1 - x_2)^2 d\tau \rangle, \end{aligned} \tag{4.1.40}$$

where we have used the shorthand  $T''(\cdot)(h, h) = T''(\cdot)(h)^2$ . Again, as  $T$  is twice continuously Fréchet-differentiable, the function  $R(x, h)$  in (4.2.27) is given by

$$R(x, h) = \int_0^1 (1 - \tau)T''(x + \tau h)h^2 d\tau,$$

and thus we obtain

$$\begin{aligned} R(x_1, \lambda(x_2 - x_1)) &= \lambda^2 \int_0^1 (1 - \tau)T''(x_1 + \tau\lambda(x_2 - x_1))(x_2 - x_1)^2 d\tau \\ &= \int_{1-\lambda}^1 (\tau - (1 - \lambda))T''(x_2 + \tau(x_1 - x_2))(x_1 - x_2)^2 d\tau \end{aligned} \tag{4.1.41}$$

and in the same way

$$R(x_2, (1 - \lambda)(x_1 - x_2)) = \int_0^{1-\lambda} (1 - \lambda - \tau) T''(x_2 + \tau(x_1 - x_2))(x_1 - x_2)^2 d\tau . \quad (4.1.42)$$

Combining definition (4.2.27) and equations (4.2.29), (4.2.30) and (4.2.31) yields

$$D(x_1, x_2, \lambda) = -\lambda(1 - \lambda)(C + \alpha)\|x_1 - x_2\|^2 + 2\langle y^\delta - T(x_k), T(x_1, x_2, \lambda) \rangle , \quad (4.1.43)$$

where

$$\begin{aligned} T(x_1, x_2, \lambda) &:= \lambda(1 - \lambda) \int_0^1 T''(x_2 + \tau(x_1 - x_2))(x_1 - x_2)^2 d\tau \\ &\quad - (1 - \lambda) \int_{1-\lambda}^1 (\tau - (1 - \lambda)) T''(x_2 + \tau(x_1 - x_2))(x_1 - x_2)^2 d\tau \\ &\quad - \lambda \int_0^{1-\lambda} (1 - \lambda - \tau) T''(x_2 + \tau(x_1 - x_2))(x_1 - x_2)^2 d\tau . \end{aligned}$$

The functional  $T(x_1, x_2, \lambda)$  can now be recasted as follows

$$\begin{aligned} T(x_1, x_2, \lambda) &= \lambda(1 - \lambda) \int_0^{1-\lambda} T''(x_2 + \tau(x_1 - x_2))(x_1 - x_2)^2 d\tau \\ &\quad + \lambda(1 - \lambda) \int_{1-\lambda}^1 T''(x_2 + \tau(x_1 - x_2))(x_1 - x_2)^2 d\tau \\ &\quad - (1 - \lambda) \int_{1-\lambda}^1 (\tau - (1 - \lambda)) T''(x_2 + \tau(x_1 - x_2))(x_1 - x_2)^2 d\tau \\ &\quad - \lambda \int_0^{1-\lambda} (1 - \lambda - \tau) T''(x_2 + \tau(x_1 - x_2))(x_1 - x_2)^2 d\tau \\ &= \lambda \int_0^{1-\lambda} \tau T''(x_2 + \tau(x_1 - x_2))(x_1 - x_2)^2 d\tau \\ &\quad + (1 - \lambda) \int_{1-\lambda}^1 (1 - \tau) T''(x_2 + \tau(x_1 - x_2))(x_1 - x_2)^2 d\tau . \end{aligned}$$

In order to estimate  $\|T(x_1, x_2, \lambda)\|$  it is necessary to estimate the integrals separately. Due to the Lipschitz-continuity of the first derivative, the second derivative can be globally

estimated by  $\|T''(x)\| \leq L$ , and it follows

$$\begin{aligned} \lambda \left\| \int_0^{1-\lambda} \tau T''(x_2 + \tau(x_1 - x_2))(x_1 - x_2)^2 d\tau \right\| &\leq \lambda \frac{(1-\lambda)^2}{2} L \|x_1 - x_2\|^2, \\ (1-\lambda) \left\| \int_{1-\lambda}^1 (1-\tau) T''(x_2 + \tau(x_1 - x_2))(x_1 - x_2)^2 d\tau \right\| &\leq (1-\lambda) \frac{\lambda^2}{2} L \|x_1 - x_2\|^2 \end{aligned}$$

and thus

$$\|T(x_1, x_2, \lambda)\| \leq \frac{\lambda(1-\lambda)}{2} L \|x_1 - x_2\|^2. \quad (4.1.44)$$

Combining (4.2.32) and (4.2.33) yields for  $\lambda \in (0, 1)$

$$\begin{aligned} D(x_1, x_2, \lambda) &\leq -\lambda(1-\lambda)(C + \alpha) \|x_1 - x_2\|^2 + 2\|y^\delta - T(x_k)\| \|T(x_1, x_2, \lambda)\| \\ &\leq -\lambda(1-\lambda)(C + \alpha) \|x_1 - x_2\|^2 + \frac{\lambda(1-\lambda)}{2} 2L \|y^\delta - T(x_k)\| \|x_1 - x_2\|^2 \\ &\stackrel{(4.2.19)}{\leq} -\lambda(1-\lambda) \left( \frac{C}{2} + \alpha \right) \|x_1 - x_2\|^2 \leq -\lambda(1-\lambda) \frac{C}{2} \|x_1 - x_2\|^2 < 0, \end{aligned}$$

and thus the functional is strictly convex.  $\square$

#### 4.1.4 Convergence Analysis

Within this section we aim to show that the sequence of iterates  $x_k$  converges strongly towards a minimizer of the Tikhonov functional. To achieve norm convergence, we prove some preliminary Lemmas.

**Lemma 5** *The sequence of iterates  $\{x_k\}_{k=0,1,2,\dots}$  has a weakly convergent subsequence.*

*Proof.* This is an immediate consequence of Proposition 4.2.1, in which it is shown that for  $k = 0, 1, 2, \dots$  the iterates  $x_k$  are contained in  $K_r(\bar{x})$ , i.e.  $\|x_{k+1} - \bar{x}\|_X \leq r$  or equivalently  $\|x_{k+1}\|_X \leq r + \|\bar{x}\|_X < \infty$ . Since the iterates are uniformly bounded, we deduce that there exists at least one accumulation point  $x_\alpha^*$  with  $x_{k,l} \xrightarrow{w} x_\alpha^*$ , where  $x_{k,l}$  denotes a subsequence of  $x_k$ .  $\square$

**Lemma 6** *The sequence  $\{\|x_{k+1} - x_k\|\}_{k=0,1,2,\dots}$  converges to zero.*

*Proof.* With the help of Corollary 12, we observe that

$$\begin{aligned} 0 &\leq \sum_k^N \{C\|x_{k+1} - x_k\|^2 - \|T(x_{k+1}) - T(x_k)\|^2\} \\ &\leq \sum_k^N \{J_\alpha^s(x_{k+1}, x_k) - J_\alpha(x_{k+1})\} \leq \sum_k^N \{J_\alpha(x_k) - J_\alpha(x_{k+1})\} \\ &= J_\alpha(x_0) - J_\alpha(x_{N+1}) \leq J_\alpha(x_0), \end{aligned}$$



i.e. the finite sums are uniformly bounded (independent on  $N$ ). By the Taylor expansion of  $T$ , we have

$$\|T(x_{k+1}) - T(x_k)\| \leq \int_0^1 \|T'(x_k + \tau(x_{k+1} - x_k))\| \|x_{k+1} - x_k\| d\tau \leq \frac{C}{2} \|x_{k+1} - x_k\| ,$$

and thus

$$0 \leq \frac{C}{2} \|x_{k+1} - x_k\|^2 \leq C \|x_{k+1} - x_k\|^2 - \|T(x_{k+1}) - T(x_k)\|^2 \longrightarrow 0$$

as  $k \rightarrow \infty$  and the assertion follows.  $\square$

**Lemma 7** *Every subsequence of  $x_k$  has a convergent subsequence  $x_{k,l}$  that converges strongly towards a function  $x_\alpha^*$ , and  $x_\alpha^*$  satisfies the necessary condition for a minimizer of the Tikhonov functional:*

$$\alpha(x_\alpha^* - \bar{x}) = T'(x_\alpha^*)^*(y^\delta - T(x_\alpha^*)) . \quad (4.1.45)$$

*Proof.* According to (4.1.30), the minimizer  $x_{k+1}$  of  $J_\alpha^s(x, x_k)$  fulfills

$$x_{k+1} = \frac{1}{C + \alpha} (Cx_k + T'(x_{k+1})^*(y^\delta - T(x_k)) + \alpha\bar{x}) .$$

Thus,

$$\begin{aligned} x_{k+1} - x_k &= -\frac{\alpha}{\alpha + C} x_k \\ &\quad + \frac{1}{C + \alpha} (T'(x_k)^*(y^\delta - T(x_k)) + \alpha\bar{x} + (T'(x_{k+1}) - T'(x_k))^*(y^\delta - T(x_k))) \end{aligned} \quad (4.1.46)$$

and, moreover, by Lemma 20,  $\|x_{k+1} - x_k\| \rightarrow 0$ , and thus

$$\|(T'(x_{k+1}) - T'(x_k))^*(y^\delta - T(x_k))\| \leq L\|x_k - x_{k+1}\| \|y^\delta - T(x_0)\| \rightarrow 0 .$$

It follows by taking the limit  $k \rightarrow \infty$  in (4.2.35),

$$0 = \lim_{k \rightarrow \infty} (\alpha(\bar{x} - x_k) + T'(x_k)^*(y^\delta - T(x_k))) . \quad (4.1.47)$$

As the sequence  $x_k$  is bounded, every subsequence has a weakly convergent subsequence. Let  $x_{k,l}$  be an arbitrary weakly convergent subsequence with weak limit  $x_\alpha^*$  (for simplicity, we will denote this sequence by  $x_k$ , too). Since

$$T'(x_k)^*(y^\delta - T(x_k)) = T'(x_k)^*(y^\delta - T(x_\alpha^*)) + T'(x_k)^*(T(x_\alpha^*) - T(x_k)) ,$$

and because of  $\|T'(x_k)^*(T(x_\alpha^*) - T(x_k))\| \leq 2C\|T(x_\alpha^*) - T(x_k)\| \rightarrow 0$  and by assumption (4.2.5), i.e.  $T'(x_k)^*(y^\delta - T(x_\alpha^*)) \rightarrow T'(x_\alpha^*)^*(y^\delta - T(x_\alpha^*))$ , we consequently obtain

$$\lim_{k \rightarrow \infty} T'(x_k)^*(y^\delta - T(x_k)) = T'(x_\alpha^*)^*(y^\delta - T(x_\alpha^*)) . \quad (4.1.48)$$

Combining (4.2.37) with (4.2.35) proves that  $x_{k,l}$  converges, and as  $x_\alpha^*$  is the weak limit of the sequence,  $x_{k,l} \rightarrow x_\alpha^*$ . Equation (4.2.34) follows by taking the limit in (4.2.36).  $\square$

In principle, the limits of different convergent subsequences of  $x_k$  can be different. Let  $x_{k,l} \times_\alpha^*$  be a subsequence of  $x_k$ , and denote by  $\tilde{x}_{k,l}$  the predecessor of  $x_{k,l}$  in  $x_k$ , i.e.  $x_{k,l} = x_i$  and  $\tilde{x}_{k,l} = x_{i-1}$ . Then we observe

$$J_\alpha^s(x_{k,l}, \tilde{x}_{k,l}) \rightarrow J_\alpha(x_\alpha^*)$$

Moreover, as we have  $J_\alpha^s(x_{k+1}, x_k) \leq J_\alpha(x_k, x_{k-1})$  for all  $k$ , it turn out that the value of the Tikhonov functional for every limit  $x_\alpha^*$  of a convergent subsequence stays the same:

$$J_\alpha(x_\alpha^*) = \text{const} . \quad (4.1.49)$$

We will now give a simple criterion that ensures convergence of the whole sequence  $x_k$ .

**Theorem 8** *Assume that there exists at least one isolated limit  $x_\alpha^*$  of a subsequence  $x_{k,l}$  of  $x_k$ . Then  $x_k \rightarrow x_\alpha^*$  holds.*

*Proof.* By  $x_\alpha^*$  we will denote the isolated limit of the sequence  $x_{k,l}$ . Let  $M$  denote the set of all limits of subsequences of the sequence  $\{x_k\}$ , and  $M_1 := M \setminus \{x_\alpha^*\}$ . Setting  $r = \text{dist}(x_\alpha^*, M_1)/2$ , we define

$$\begin{aligned} B_r &:= \{x_k : \|x_k - x_\alpha^*\| \leq r\} \\ \bar{B}_r &:= \{x_k : x_k \notin K_r\} . \end{aligned}$$

Now let us assume  $M_1 \neq \emptyset$ . Then both  $B_r$ ,  $\bar{B}_r$  contain infinitely many elements. In particular, there exist infinitely many pairs of iterates  $x_k, x_{k+1}$  with  $x_k \in K_r$  and  $x_{k+1} \in \bar{B}_r$ , and we can define a subsequence  $\tilde{x}_k$  by picking all pairs  $x_k \in B_r$  and  $x_{k+1} \in \bar{B}_r$  out of the sequence  $\{x_k\}_{k \in \mathbb{N}}$ , i.e.

$$\tilde{x}_{2l} = x_k \in B_r \quad (4.1.50)$$

$$\tilde{x}_{2l+1} = x_{k+1} \in \bar{B}_r$$

Because of Lemma 20 we observe  $\|x_{2l} - x_{2l+1}\| \rightarrow 0$ , and with (4.1.50) follows that the elements of  $\tilde{x}_l$  come arbitrary close to  $\partial B_r = \{x : \|x - x_\alpha^*\| = r\}$ , i.e.

$$\lim_{l \rightarrow \infty} \|\tilde{x}_l - x_\alpha^*\| = r . \quad (4.1.51)$$

According to Lemma 21, every subsequence of  $x_k$  has a convergent subsequence. Let  $\tilde{x}_{l,k}$  be a convergent subsequence of  $\tilde{x}_l$  with limit  $\tilde{x}_\alpha^*$ . Because of (4.1.51) holds  $\tilde{x}_\alpha^* \in \partial B_r$ . On the other hand, as  $x_\alpha^* \neq \tilde{x}_\alpha^*$ , we have  $\tilde{x}_\alpha^* \in M_1$ , which is a contradiction to  $\text{dist}(x_\alpha^*, M_1) = 2r$ .

We conclude  $M_1 = \emptyset$ , i.e.  $x_\alpha^*$  is the only limit of convergent subsequences of  $x_k$ . As by Lemma 21 every subsequence of  $x_k$  has a subsequence that converges towards  $x_\alpha^*$ , the

whole sequence converges towards  $x_\alpha^*$  by the convergence principles.  $\square$

On the other hand, we conclude the sequence  $x_k$  can only not converge if the Tikhonov functional has a dense set of critical points, and the belonging functional values are constant.

By the following Proposition, the fixed point  $x_\alpha^*$  is also a minimizer for the functional  $J_\alpha^s(x, x_\alpha^*)$ .

**Proposition 4.1.5** *The accumulation point  $x_\alpha^*$  is a minimizer for the functional  $J_\alpha^s(x, x_\alpha^*)$ .*

*Proof.* We aim to show that for all  $h \in X$ ,

$$J_\alpha^s(x_\alpha^* + h, x_\alpha^*) \geq J_\alpha^s(x_\alpha^*, x_\alpha^*) + \frac{C}{2} \|h\|^2 .$$

This is obtained by making use of

$$J_\alpha^s(x_\alpha^* + h, x_\alpha^*) = J_\alpha^s(x_\alpha^*, x_\alpha^*) + 2\langle y^\delta - T(x_\alpha^*), T(x_\alpha^*) - T(x_\alpha^* + h) \rangle + 2\langle \alpha(x_\alpha^* - \bar{x}), h \rangle + (\alpha + C) \|h\|^2$$

and

$$\alpha(x_\alpha^* - \bar{x}) = T'(x_\alpha^*)^*(y^\delta - T(x_\alpha^*)) .$$

With the Lipschitz-continuity of  $T'$  this results in

$$\begin{aligned} J_\alpha^s(x_\alpha^* + h, x_\alpha^*) &\geq J_\alpha^s(x_\alpha^*, x_\alpha^*) - 2\|y^\delta - T(x_\alpha^*)\| \|T(x_\alpha^*) - T(x_\alpha^* + h) + T'(x_\alpha^*)h\| \\ &\quad + (\alpha + C) \|h\|^2 \\ &\geq J_\alpha^s(x_\alpha^*, x_\alpha^*) - 2\frac{C}{2L} \frac{L}{2} \|h\|^2 + (\alpha + C) \|h\|^2 \\ &= J_\alpha^s(x_\alpha^*, x_\alpha^*) + \frac{C}{2} \|h\|^2 + \alpha \|h\|^2 + J_\alpha^s(x_\alpha^*, x_\alpha^*) + \frac{C}{2} \|h\|^2 . \end{aligned}$$

$\square$

Equation (4.2.34) states that our algorithm reconstructs at least a critical point of the Tikhonov functional. In general, a critical point will not always be a minimizer of the Tikhonov functional. However, we will give a condition that ensures this property. Namely, if we impose the condition (4.1.11) and do assume that the solution  $x^\dagger$  fulfills a smoothness condition, then we can show that every critical point of the Tikhonov functional is a global minimizer. We wish to remark that (4.1.11) is a rather strong condition. However, conditions of this type have been used earlier, e.g. for Landweber iteration [HNS95, Ram99] and for Levenberg-Marquardt iteration [Han97a].

**Theorem 9** *Let  $T$  be a twice Fréchet differentiable operator with (4.1.11). If a smoothness condition*

$$x^\dagger - \bar{x} = T'(x^\dagger)^* \omega , \quad L\|\omega\| < 1/3 \quad (4.1.52)$$

*holds, and the regularization parameter is chosen with*

$$\alpha = \delta/\eta \text{ and } \eta \leq \|\omega\| \quad (4.1.53)$$

*then (4.2.34) has a unique solution. Thus the minimizer of the Tikhonov-functional is unique, too.*

*Proof.* Let  $x_\alpha^\delta$  denote a global minimizer of the Tikhonov functional, and  $x_\alpha^*$  be a critical point. With  $T(x_\alpha^*) = T(x_\alpha^\delta) + T'(x_\alpha^\delta)(x_\alpha^* - x_\alpha^\delta) + R(x_\alpha^\delta, x_\alpha^* - x_\alpha^\delta)$  we obtain

$$\begin{aligned} \|y^\delta - T(x_\alpha^*)\|^2 - \|y^\delta - T(x_\alpha^\delta)\|^2 &= \|T(x_\alpha^\delta) - T(x_\alpha^*)\|^2 + 2\langle y^\delta - T(x_\alpha^\delta), T(x_\alpha^\delta) - T(x_\alpha^*) \rangle \\ &= \|T(x_\alpha^\delta) - T(x_\alpha^*)\|^2 + 2\langle T'(x_\alpha^\delta)^*(y^\delta - T(x_\alpha^\delta)), x_\alpha^\delta - x_\alpha^* \rangle - 2\langle y^\delta - T(x_\alpha^\delta), R(x_\alpha^\delta, x_\alpha^* - x_\alpha^\delta) \rangle \\ &\stackrel{(4.2.34)}{=} \|T(x_\alpha^\delta) - T(x_\alpha^*)\|^2 + 2\alpha\langle x_\alpha^\delta - \bar{x}, x_\alpha^\delta - x_\alpha^* \rangle - 2\langle y^\delta - T(x_\alpha^\delta), R(x_\alpha^\delta, x_\alpha^* - x_\alpha^\delta) \rangle. \end{aligned}$$

Because of

$$\alpha\|x_\alpha^* - \bar{x}\|^2 - \alpha\|x_\alpha^\delta - \bar{x}\|^2 = \alpha\|x_\alpha^* - x_\alpha^\delta\|^2 + 2\alpha\langle x_\alpha^* - x_\alpha^\delta, x_\alpha^\delta - \bar{x} \rangle$$

it follows that

$$\begin{aligned} J_\alpha(x_\alpha^*) - J_\alpha(x_\alpha^\delta) &= \|y^\delta - T(x_\alpha^*)\|^2 - \|y^\delta - T(x_\alpha^\delta)\|^2 + \alpha\|x_\alpha^* - \bar{x}\|^2 - \alpha\|x_\alpha^\delta - \bar{x}\|^2 \\ &= \|T(x_\alpha^\delta) - T(x_\alpha^*)\|^2 + \alpha\|x_\alpha^* - x_\alpha^\delta\|^2 - 2\langle y^\delta - T(x_\alpha^\delta), R(x_\alpha^\delta, x_\alpha^* - x_\alpha^\delta) \rangle. \end{aligned} \quad (4.1.54)$$

By the same argument, we get

$$J_\alpha(x_\alpha^\delta) - J_\alpha(x_\alpha^*) = \|T(x_\alpha^\delta) - T(x_\alpha^*)\|^2 + \alpha\|x_\alpha^* - x_\alpha^\delta\|^2 - 2\langle y^\delta - T(x_\alpha^*), R(x_\alpha^*, x_\alpha^\delta - x_\alpha^*) \rangle. \quad (4.1.55)$$

Now, adding (4.1.54) and (4.1.55) yields

$$\begin{aligned} 0 &= 2\|T(x_\alpha^\delta) - T(x_\alpha^*)\|^2 + 2\alpha\|x_\alpha^* - x_\alpha^\delta\|^2 \\ &\quad - 2\langle y^\delta - T(x_\alpha^\delta), R(x_\alpha^\delta, x_\alpha^* - x_\alpha^\delta) \rangle - 2\langle y^\delta - T(x_\alpha^*), R(x_\alpha^*, x_\alpha^\delta - x_\alpha^*) \rangle. \end{aligned} \quad (4.1.56)$$

For twice continuous differentiable operators, the quadratic remainder of the Taylor series is given by

$$\begin{aligned} R(x_\alpha^\delta, x_\alpha^* - x_\alpha^\delta) &= \int_0^1 (1-\tau)T''(x_\alpha^\delta + \tau(x_\alpha^* - x_\alpha^\delta))(x_\alpha^* - x_\alpha^\delta, x_\alpha^* - x_\alpha^\delta) d\tau \\ R(x_\alpha^*, x_\alpha^\delta - x_\alpha^*) &= \int_0^1 (1-\tau)T''(x_\alpha^* + \tau(x_\alpha^\delta - x_\alpha^*))(x_\alpha^\delta - x_\alpha^*, x_\alpha^\delta - x_\alpha^*) d\tau. \end{aligned}$$

Setting  $\tau = 1 - \tau'$  and  $h = x_\alpha^\delta - x_\alpha^*$  and we obtain

$$R(x_\alpha^\delta, x_\alpha^* - x_\alpha^\delta) = \int_0^1 \tau' T''(x_\alpha^* + \tau' h)(h, h) d\tau'$$

and thus

$$\begin{aligned}
\langle y^\delta - T(x_\alpha^\delta), R(x_\alpha^\delta, x_\alpha^* - x_\alpha^\delta) \rangle &= \langle y^\delta - T(x_\alpha^\delta), \int_0^1 \tau T''(x_\alpha^* + \tau h)(h, h) d\tau \rangle \\
&= \langle y^\delta - T(x_\alpha^\delta), \int_0^1 (\tau - 1) T''(x_\alpha^* + \tau h)(h, h) d\tau \rangle + \langle y^\delta - T(x_\alpha^\delta), \int_0^1 T''(x_\alpha^* + \tau h)(h, h) d\tau \rangle \\
&= -\langle y^\delta - T(x_\alpha^\delta), R(x_\alpha^*, h) \rangle + \langle y^\delta - T(x_\alpha^\delta), \int_0^1 T''(x_\alpha^* + \tau h)(h, h) d\tau \rangle \tag{4.1.57}
\end{aligned}$$

Inserting (4.1.57) in (4.1.56) yields

$$\begin{aligned}
0 &= 2\|T(x_\alpha^\delta) - T(x_\alpha^*)\|^2 + 2\alpha\|x_\alpha^* - x_\alpha^\delta\|^2 - 2\langle T(x_\alpha^*) - T(x_\alpha^\delta), R(x_\alpha^*, x_\alpha^\delta - x_\alpha^*) \rangle \\
&\quad - 2\langle y^\delta - T(x_\alpha^\delta), \int_0^1 T''(x_\alpha^* + \tau h)(h, h) d\tau \rangle \\
&\geq 2\|T(x_\alpha^\delta) - T(x_\alpha^*)\|^2 + 2\alpha\|x_\alpha^* - x_\alpha^\delta\|^2 - 2\|T(x_\alpha^*) - T(x_\alpha^\delta)\| \|R(x_\alpha^*, x_\alpha^\delta - x_\alpha^*)\| \\
&\quad - 2\|y^\delta - T(x_\alpha^\delta)\| \left\| \int_0^1 T''(x_\alpha^* + \tau h)(h, h) d\tau \right\|
\end{aligned}$$

By (4.1.11) we conclude  $\|R(x_\alpha^*, x_\alpha^\delta - x_\alpha^*)\| \leq \|T(x_\alpha^\delta) - T(x_\alpha^*)\|$ , and from the smoothness condition (4.1.52), see [EHN96a] p.246, it follows

$$\|y^\delta - T(x_\alpha^\delta)\| \leq \delta + 2\alpha\|\omega\| \stackrel{(4.1.53)}{\leq} 3\alpha\|\omega\|.$$

Altogether we get

$$\begin{aligned}
0 &\geq 2\|T(x_\alpha^\delta) - T(x_\alpha^*)\|^2 + 2\alpha\|x_\alpha^* - x_\alpha^\delta\|^2 - 2\|T(x_\alpha^*) - T(x_\alpha^\delta)\| \|R(x_\alpha^*, x_\alpha^\delta - x_\alpha^*)\| \\
&\quad - 2\|y^\delta - T(x_\alpha^\delta)\| \left\| \int_0^1 T''(x_\alpha^* + \tau h)(h, h) d\tau \right\| \\
&\geq (2 - 6L\|\omega\|)\alpha\|x_\alpha^* - x_\alpha^\delta\|^2 \geq 0,
\end{aligned}$$

and thus we have shown  $(2 - 6L\|\omega\|)\alpha\|x_\alpha^* - x_\alpha^\delta\|^2 = 0$ , and because of (4.1.52) holds  $x_\alpha^* = x_\alpha^\delta$ .  $\square$

### 4.1.5 Regularization Properties

Conditions (4.1.11), (4.1.52) ensure the convergence of our algorithm towards the unique minimizer of the Tikhonov functional. Using a proper parameter choice rule for the regularization parameter gives convergence/convergence rates for Tikhonov regularization. We might recall a few well known parameter rules.

- (I) Let  $T$  be a weakly sequentially closed operator, and the regularization parameter  $\alpha$  chosen such that  $\alpha(\delta) \rightarrow 0$  and  $\delta^2/\alpha \rightarrow 0$  as  $\delta \rightarrow 0$ . Then every sequence  $x_{\alpha_k}^{\delta_k}$  with  $\delta_k$  has a convergent subsequence that converges toward an  $\bar{x}$ - minimum norm solution  $x^\dagger$ . In particular, if a smoothness condition (4.1.52) holds, and the regularization parameter is chosen by  $\alpha = \delta/\eta$ ,  $\eta \leq \|\omega\|$ , then we obtain a convergence rate of  $\mathcal{O}(\sqrt{\delta})$  [EHN96a].
- (II) Let  $T$  be a Fréchet differentiable operator with (4.1.9). Moreover, assume that  $x^\dagger$  fulfills a smoothness condition  $x^\dagger - \bar{x} = (T'(x^\dagger)^* T'(x^\dagger))^\nu \omega$  for  $\nu \in [1/2, 1]$  with  $L\|\omega\| < 1/3$ . If the parameter is chosen by  $\alpha \sim \delta^{2/(2\nu+1)}$ , then we obtain a convergence rate of  $\mathcal{O}(\delta^{2\nu/(2\nu+1)})$  [EHN96a].
- (III) (Morozov's discrepancy principle) Let  $T$  be a twice continuous differentiable operator with (4.1.9), and assume  $x^\dagger - \bar{x} = T'(x^\dagger)^* \omega$  with  $L\|\omega\| \leq 0.241$ . Then there exists a regularization parameter  $\alpha \leq \delta/\eta$ ,  $\eta \leq \|\omega\|$  with

$$\delta \leq \|y^\delta - T(x_\alpha^\delta)\| \leq c\delta, \quad (4.1.58)$$

and for a belonging minimizer holds  $\|x_\alpha^\delta - x^\dagger\| = \mathcal{O}(\sqrt{\delta})$ . A regularization parameter fulfilling (4.1.58) can be found by testing  $\|y^\delta - T(x_{\alpha_k}^\delta)\|$  for a sequence  $\alpha_k = \alpha_0 q^k$  with appropriate chosen  $\alpha_0$  and  $q < 1$ , see [Ram03].

Please note that if a solution fulfills smoothness condition from (II), then, for properly scaled  $T'(x^\dagger)$ , also a smoothness condition (4.1.52) holds. Thus, if (4.1.11) holds, all rules are conform with the requirements of our minimization algorithm. Combining all ingredients and picking a proper parameter rule we may provide the following algorithm which uses our iteration routine **TIREFU** (**T**ikhonov **R**Eplacement **F**unctional) for solving the nonlinear problem  $T(x) = y$  with  $\|y - y^\delta\| \leq \delta$ . The exact way for computing a solution  $x_\alpha^*$  goes as follows (applying III):

- Define a sequence  $\{\alpha_n\}$  with  $\alpha_n \xrightarrow{n \rightarrow \infty} 0$ , pick some  $r$  and set  $x_0 = \bar{x}$  (initial value  $x_0$  for the outer iteration)
- **while**  $\|T(x_\alpha^*) - y^\delta\| > r \cdot \delta$ 
  - $\alpha = \alpha_n$
  - pick an admissible  $C$
  - $[x_\alpha^*] = \mathbf{TIREFU}(T, y^\delta, C, \alpha, x_0)$ :
 
$$x_{k+1} = \arg \min_x J_\alpha^s(x, x_k) \text{ (solved by a **Fixed Point Iteration**)}$$

$$x_\alpha^* = \lim_{k \rightarrow \infty} x_k$$
  - $x_0 = x_\alpha^*$
- end**

For this algorithm we may now formulate the following optimality result:

**Theorem 10** *Assume that (4.1.11) holds. Then Tikhonov regularization with one of the parameter rules I-III, where the minimizers are computed by **TIREFU**, is an optimal regularization method.*

Since in any numerical realization we cannot treat infinite series (computing limits), we additionally have to incorporate a stopping rule. If  $\Phi_\alpha(x, a)$  denotes the operator defined in (4.1.30), then the algorithm reads as follows:

- Define a sequence  $\{\alpha_n\}$  with  $\alpha_n \xrightarrow{n \rightarrow \infty} 0$ , pick some  $r$ , tolerances  $\tau_1$  and  $\tau_2$ , set  $x_\alpha^\star = \bar{x}$
- **while**  $\|T(x_\alpha^\star) - y^\delta\| > r \cdot \delta$ 
  - $\alpha = \alpha_n$
  - pick an admissible  $C$
  - $[x_\alpha^\star] = \mathbf{TIREFU}(T, y^\delta, C, \alpha, x_0, \tau_1, \tau_2)$ 
    - $k = 0$
    - while**  $\|x_{k+1} - x_k\| > \tau_1$ 
      - $l = 0, x_{k,0} = x_k$
      - Repeat**
      - $l = l + 1$
      - $x_{k,l} = \Phi_\alpha(x_{k,l-1}, x_k)$
      - Until**  $\|x_{k,l} - x_{k,l+1}\| \leq \tau_2$
      - $x_{k+1} = x_{k,l}$
      - $k = k + 1$
      - end**
    - $x_\alpha^\star = x_k$
  - $x_0 = x_\alpha^\star$
- end**

As we have pointed out, the strongest limitation of **TIREFU** is condition (4.1.11). However, this condition was only used once at the very end of our analysis, and we expect that it will be possible to weaken the condition. As Landweber iteration and Levenberg-Marquardt iteration work under a similar condition, we might compare **TIREFU** with these methods. Landweber iteration is known to be a slow method, and as we use fixed point methods, we do expect that **TIREFU** will be faster. Moreover, using our optimization routine with rule II, we obtain an optimal method for  $\nu \in [1/2, 1]$ . In contrast, to obtain convergence rates, Landweber requires an additional conditions  $T'(x) = R_x T'(x^\dagger)$ , where  $R_x$  is a family of bounded operators with  $\|I - R_x\| \leq K\|x - x^\dagger\|$ . This condition is even more restrictive than (4.1.11). In addition, convergence rates are only available for  $0 < \nu \leq 1/2$ . As for Levenberg-Marquardt, it is only known that the iteration is a regularization method under a condition slightly more restrictive as (4.1.11), and so far, nothing is known on convergence rates. Thus we might conclude that **TIREFU** works under less restrictive conditions.

## 4.2 Multi-Frames and Mixed One-Homogeneous Constraints

### 4.2.1 Scope of the Problem

As in the previous section, we consider again the computation of an approximation to a solution of a nonlinear operator equation

$$T(x) = y, \quad (4.2.1)$$

where  $T : X \rightarrow Y$  is an (ill-posed) operator between Hilbert spaces  $X, Y$ . If only noisy data  $y^\delta$  with  $\|y^\delta - y\| \leq \delta$  are available, problem (4.2.1) has to be stabilized by regularization methods. Many of the well known methods for linear ill-posed problems have been generalized to nonlinear operator equations. But so far all the proposed schemes for nonlinear problems incorporate at most quadratic regularization. In many applications the solution is assumed to have sparse expansion which immediately leads to the involvement of nonquadratic penalties, e.g.  $\ell_p$  norms with  $p < 2$ . In linear lore, this problem is still solved, see [DDD04]. In nonlinear inverse problems there is a very recent first attempt, see [RT05a], which solves nonlinear operator equations with sparsity constraints. However, recent developments indicate that (highly) redundant systems, such as frames or systems of frames may yield a gain in this context (optimal representation/decomposition of the solution to be reconstructed). When dealing with dictionaries of frame systems, there exist certain methods, e.g. such as basis pursuit [CDS95], that allow a decomposition of signals/functions into an optimal superposition of dictionary elements, where optimal means having smallest  $\ell_1$  norm of coefficients among all such decompositions. In [Tes05b], we have presented a method which combines an iterated thresholding scheme for solving linear inverse problems while requiring that the solution is assumed to have a sparse expansion in a multi-frame dictionary. Here we now also assume that the solution has a sparse expansion in a multi-frame dictionary but we aim now to extend the theory to nonlinear inverse problems with mixed multi-sparsity constraints. The main result, coming out by combing previously elaborated results ([DT04, DT05], [RT05b, RT05a], and [Tes05b]) is the development of a new method which is sort of projected thresholding Landweber iteration for solving a system of fixed point equations.

As in [Tes05b], let us assume we are given a finite family of preassigned frames  $\{\phi_\lambda^i\}_{\lambda \in \Lambda_i, i \in \mathcal{I}} \subset X$ ,  $n = \text{card}(\mathcal{I})$ , for which we have associated frame operators

$$F_i : X \rightarrow \ell_2 \text{ via } F_i x = \{\langle x, \phi_\lambda^i \rangle\}_{\lambda \in \Lambda_i} \text{ with } A_i \cdot I \leq F_i^* F_i \leq B_i \cdot I.$$

The variational formulation of the nonlinear inverse problem in a multi-frame setting with so-called multi-sparsity, or more general, multi-one-homogeneous constraints can be now casted as follows: find sequences of coefficients  $\mathbf{g} = (\mathbf{g}_1, \dots, \mathbf{g}_n) \in (\ell_2)^n$  such that

$$J_\alpha(\mathbf{g}) = \|y^\delta - T(K\mathbf{g})\|_Y^2 + 2\alpha \cdot \Psi_L(\mathbf{g}) \quad (4.2.2)$$



is minimized, where  $\alpha = (\alpha_1, \dots, \alpha_n)$  and  $\Psi_{\mathbf{L}}(\mathbf{g}) = (\Psi_1(\mathbf{L}_1 \mathbf{g}_1), \dots, \Psi_n(\mathbf{L}_n \mathbf{g}_n))$ . In our case,  $K\mathbf{g} = K(\mathbf{g}_1, \dots, \mathbf{g}_n) = \sum_{i \in \mathcal{I}} F_i^* \mathbf{g}_i$ , but one could also involve, as in [Tes05b], additional linear and bounded operators  $E_i$ , i.e.  $K_E(\mathbf{g}_1, \dots, \mathbf{g}_n) = \sum_{i \in \mathcal{I}} E_i F_i^* \mathbf{g}_i$ . Moreover, the  $\Psi_i$  stand for positive, one-homogeneous, lower semi-continuous and convex penalties (which are usually some weighted  $\ell_p$  norms of the frame coefficients), and the infinite matrices  $\mathbf{L}_i$  are restricted to be isometric mappings. In particular, we also need to require,

$$\|\mathbf{g}\|_{(\ell_2)^n} \leq \|\Psi_{\mathbf{L}}(\mathbf{g})\|_{\ell_1} . \quad (4.2.3)$$

The strategies for nonlinear cases suggested in [RT05b, RT05a], seem to be also adequate when dealing with multi-sparsity, or more general, with multi-one-homogeneous constraints. Before sketching the idea, we need to clarify the  $(\ell_2)^n$ -framework. First, for sake of simplicity, we restrict ourselves to  $E_i = I$ , for all  $i$ . Note that the suggested theory applies without any changes also to  $E_i \neq I$ . For the preassigned frame operators  $F_i : X \rightarrow \ell_2$ ,

$$K : \ell_2 \times \dots \times \ell_2 \rightarrow X \text{ via } (\ell_2)^n \ni \mathbf{g} = (\mathbf{g}_1, \dots, \mathbf{g}_n) \mapsto \sum_{i \in \mathcal{I}} F_i^* \mathbf{g}_i ,$$

where the Hilbert space  $(\ell_2)^n$  is endowed with the scalar  $\langle \mathbf{g}, \mathbf{h} \rangle_{(\ell_2)^n} = \langle \mathbf{g}_1, \mathbf{h}_1 \rangle_{\ell_2} + \dots + \langle \mathbf{g}_n, \mathbf{h}_n \rangle_{\ell_2}$  and thus the associated norm is given by  $\|\mathbf{g}\|_{(\ell_2)^n}^2 = \|\mathbf{g}_1\|_{\ell_2}^2 + \dots + \|\mathbf{g}_n\|_{\ell_2}^2$ . Moreover,

$$\langle K\mathbf{f}, h \rangle_X = \langle \mathbf{f}, (F_1 h, \dots, F_n h) \rangle_{(\ell_2)^n} = \langle \mathbf{f}, K^* h \rangle_{(\ell_2)^n} ,$$

and thus,

$$\|K\| \leq \sqrt{B_1 + \dots + B_n} =: B .$$

The general idea for solving the nonlinear inverse problem in a multi-frame setting goes now as follows: we replace (4.2.2) by a sequence of functionals from which we hope that they are easier to treat and that the sequence of minimizers converge in some sense to, at least, a critical point of (4.2.2). To be more concrete, for  $\mathbf{g} \in (\ell_2)^n$  and some auxiliary  $\mathbf{a} \in (\ell_2)^n$ , we introduce

$$J_{\alpha}^s(\mathbf{g}, \mathbf{a}) := J_{\alpha}(\mathbf{g}) + C\|\mathbf{g} - \mathbf{a}\|_{(\ell_2)^n}^2 - \|T(K\mathbf{g}) - T(K\mathbf{a})\|_Y^2 \quad (4.2.4)$$

and create an iteration process by:

1. Pick  $\mathbf{g}_0 \in (\ell_2)^n$  and some proper constant  $C > 0$
2. Derive a sequence  $\{\mathbf{g}_k\}_{k=0,1,\dots}$  by the iteration:

$$\mathbf{g}_{k+1} = \arg \min_{\mathbf{g}_k \in (\ell_2)^n} J_{\alpha}^s(\mathbf{g}, \mathbf{g}_k) \quad k = 0, 1, 2, \dots$$

In order to avoid ambiguity, we will always denote  $(\mathbf{g}_i)_k \in \ell_2$  as the  $k$ -th iterate of the  $i$ -th component of  $\mathbf{g}_k \in (\ell_2)^n$ , i.e.  $(\ell_2)^n \ni \mathbf{g}_k = ((\mathbf{g}_1)_k, \dots, (\mathbf{g}_n)_k)$ , and a particular coefficient of the  $k$ -th iterate with respect to some index  $\lambda \in \Lambda_i$  is then denoted by  $(\mathbf{g}_{\lambda,i})_k$ ; in its full glory we may thus write the  $k$ -th iterate  $\mathbf{g}_k = (\{(\mathbf{g}_{\lambda,1})_k\}_{\lambda \in \Lambda_1}, \dots, \{(\mathbf{g}_{\lambda,n})_k\}_{\lambda \in \Lambda_n}) \in (\ell_2)^n$ . As we shall see later on, in order to prove norm convergence of the iterates  $\mathbf{g}_k$  towards

a critical point of  $J_\alpha$ , we have to restrict ourselves to a class of nonlinear problems for which all of the following three requirements hold true,

$$\begin{aligned} \mathbf{g}_k &\xrightarrow{w} \mathbf{g} \implies T(K\mathbf{g}_k) \rightarrow T(K\mathbf{g}) \quad , \\ F_j T'(K\mathbf{g}_k)^* z &\rightarrow F_j T'(K\mathbf{g})^* z \quad , \text{ for all } z \text{ and } j \quad , \\ \|T'(K\mathbf{g}) - T'(K\mathbf{g}')\| &\leq LB \|\mathbf{g} - \mathbf{g}'\|_{(\ell_2)^n} \quad . \end{aligned} \quad (4.2.5)$$

It may happen that  $T$  already meets these conditions as an operator from  $X \rightarrow Y$ . If not, we may proceed as stated in the last section.

The remaining section is organized as follows: In Section 4.2.2, we explain how the replacement functionals are constructed and discuss the well-posedness of the resulting problem. In Section 4.2.3, we derive conditions on the minimizing elements. The main result of the paper is presented in Section 4.2.4: strong convergence of the iterates towards a critical point. We end with giving a regularization theorem.

### 4.2.2 Proper Surrogate Functionals

By the definition of  $J_\alpha^s$  in (4.2.4) it is not clear whether the functional is positive definite or even bounded from below. This will be clarified in this section, i.e. we will show that this is the case provided the constant  $C$  is chosen properly.

For given multi-parameter  $\alpha \in \mathbb{R}_+^n$  and  $\mathbf{g}_0 \in (\ell_2)^n$  we may define a ball

$$K_r := \{\mathbf{g} \in (\ell_2)^n : \|\Psi_L(\mathbf{g})\|_{\ell_1} \leq r\} \quad ,$$

where the radius  $r$  is given by

$$r := J_\alpha(\mathbf{g}_0) / (2 \min\{\alpha_i\}). \quad (4.2.6)$$

This obviously ensures,  $\mathbf{g}_0 \in K_r$ . Furthermore, we define the constant  $C$  by

$$C := 2B^2 \max \left\{ \left( \sup_{\mathbf{g} \in K_r} \|T'(K\mathbf{g})\| \right)^2, L \sqrt{\|y^\delta - T(K\mathbf{g}_0)\|^2 + 2\alpha \cdot \Psi_L(\mathbf{g}_0)} \right\}, \quad (4.2.7)$$

where  $L$  is the Lipschitz constant of the Fréchet derivative of  $T$ . We assume that  $\mathbf{g}_0$  was chosen such that  $r < \infty$  and  $C < \infty$ .

**Lemma 11** *Let  $r$  and  $C$  be chosen by (4.2.6), (4.2.7). Then, for all  $\mathbf{g} \in K_r$ ,*

$$C \|\mathbf{g} - \mathbf{g}_0\|_{(\ell_2)^n}^2 - \|T(K\mathbf{g}) - T(K\mathbf{g}_0)\|_Y^2 \geq 0 \quad (4.2.8)$$

*and thus,  $J_\alpha(\mathbf{g}) \leq J_\alpha^s(\mathbf{g}, \mathbf{g}_0)$ .*

*Proof.* By Taylors expansion we have

$$T(K\mathbf{g} + K\mathbf{h}) = T(K\mathbf{g}) + \int_0^1 T'(K\mathbf{g} + \tau K\mathbf{h}) K\mathbf{h} d\tau$$

and thus we get with  $\mathbf{h} = \mathbf{g}_0 - \mathbf{g}$

$$\begin{aligned} \|T(K\mathbf{g}) - T(K\mathbf{g}_0)\|_Y &\leq \int_0^1 \|T'(K\mathbf{g} + \tau K(\mathbf{g}_0 - \mathbf{g}))\| \|K(\mathbf{g}_0 - \mathbf{g})\|_X d\tau \\ &\leq \sup_{\mathbf{g} \in K_r} \|T'(K\mathbf{g})\| \|K(\mathbf{g}_0 - \mathbf{g})\|_X \\ &\leq \sup_{\mathbf{g} \in K_r} \|T'(K\mathbf{g})\| B \|\mathbf{g}_0 - \mathbf{g}\|_{(\ell_2)^n} \end{aligned}$$

Consequently, we get for all  $\mathbf{g} \in K_r$

$$\begin{aligned} C\|\mathbf{g} - \mathbf{g}_0\|_{(\ell_2)^n}^2 - \|T(K\mathbf{g}) - T(K\mathbf{g}_0)\|_Y^2 &\geq \\ C\|\mathbf{g} - \mathbf{g}_0\|_{(\ell_2)^n}^2 - B^2 \left( \sup_{\mathbf{g} \in K_r} \|T'(K\mathbf{g})\| \|\mathbf{g} - \mathbf{g}_0\|_{(\ell_2)^n} \right)^2 & \\ = \frac{C}{2} \|\mathbf{g} - \mathbf{g}_0\|_{(\ell_2)^n}^2 \geq 0, & \end{aligned}$$

and the functional  $J_\alpha^s(\mathbf{g}, \mathbf{g}_0)$  is non-negative for all  $\mathbf{g} \in K_r$ .  $\square$

Next, we show that this carries over to all of the iterates:

**Proposition 4.2.1** *Let  $\mathbf{g}_0, \alpha$  be given and  $r, C$  be defined by (4.2.6), (4.2.7). Then the functionals  $J_\alpha^s(\mathbf{g}, \mathbf{g}_k)$  are bounded from below for all  $k \in \mathbb{N}$  and have thus minimizers. For the minimizer  $\mathbf{g}_{k+1}$  of  $J_\alpha^s(\mathbf{g}, \mathbf{g}_k)$  holds  $\mathbf{g}_{k+1} \in K_r$ .*

*Proof.* The proof will be done by induction. For  $k = 1$ , we show in a first step that  $J_\alpha^s(\mathbf{g}, \mathbf{g}_0)$  is bounded from below. We have

$$\begin{aligned} \|y^\delta - T(K\mathbf{g})\|_Y^2 &= \|y^\delta - T(K\mathbf{g}_0)\|_Y^2 + \|T(K\mathbf{g}_0) - T(K\mathbf{g})\|_Y^2 \\ &\quad + 2\langle y^\delta - T(K\mathbf{g}_0), T(K\mathbf{g}_0) - T(K\mathbf{g}) \rangle_Y. \end{aligned} \quad (4.2.9)$$

Thus,

$$\begin{aligned} J_\alpha^s(\mathbf{g}, \mathbf{g}_0) - 2\alpha \cdot \Psi_L(\mathbf{g}) &= \|y^\delta - T(K\mathbf{g}_0)\|_Y^2 + 2\langle y^\delta - T(K\mathbf{g}_0), T(K\mathbf{g}_0) - T(K\mathbf{g}) \rangle_Y \\ &\quad + C\|\mathbf{g} - \mathbf{g}_0\|_{(\ell_2)^n}^2 \end{aligned} \quad (4.2.10)$$

$$\begin{aligned} &\geq \|y^\delta - T(K\mathbf{g}_0)\|_Y^2 - 2\|y^\delta - T(K\mathbf{g}_0)\|_Y \|T(K\mathbf{g}_0) - T(K\mathbf{g})\|_Y \\ &\quad + C\|\mathbf{g} - \mathbf{g}_0\|_{(\ell_2)^n}^2. \end{aligned} \quad (4.2.11)$$

Again by Taylor expansion,

$$\|T(K\mathbf{g}_0) - T(K\mathbf{g})\|_Y \leq B\|T'(K\mathbf{g}_0)\| \|\mathbf{g}_0 - \mathbf{g}\|_{(\ell_2)^n} + \frac{B^2 L}{2} \|\mathbf{g}_0 - \mathbf{g}\|_{(\ell_2)^n}^2. \quad (4.2.12)$$

Now let us assume that  $J_\alpha^s(\mathbf{g}, \mathbf{g}_0)$  is not bounded from below, e.g. there exists a sequence  $\mathbf{g}_l$  such that  $J_\alpha^s(\mathbf{g}_l, \mathbf{g}_0) \rightarrow -\infty$ . This can only hold if  $\|T(K\mathbf{g}_0) - T(K\mathbf{g}_l)\| \rightarrow \infty$ , and

because of (4.2.12) follows  $\|\mathbf{g}_l\|_{(\ell_2)^n} \rightarrow \infty$  as well. In particular, for  $l$  large enough, we derive from (4.2.12)

$$\|T(K\mathbf{g}_0) - T(K\mathbf{g}_l)\|_Y \leq B^2 L \|\mathbf{g}_0 - \mathbf{g}_l\|_{(\ell_2)^n}^2,$$

and combining this estimate with (4.2.11) yields

$$J_\alpha^s(\mathbf{g}_l, \mathbf{g}_0) - 2\alpha \cdot \Psi_L(\mathbf{g}_l) \geq \|y^\delta - T(K\mathbf{g}_0)\|_Y^2 - 2B^2 L \|y^\delta - T(K\mathbf{g}_0)\|_Y \|\mathbf{g}_l - \mathbf{g}_0\|_{(\ell_2)^n}^2 + C \|\mathbf{g}_l - \mathbf{g}_0\|_{(\ell_2)^n}^2.$$

From the definition of  $C$  in (4.2.7) follows  $2B^2 L \|y^\delta - T(K\mathbf{g}_0)\|_Y \leq C$  and thus

$$J_\alpha^s(\mathbf{g}_l, \mathbf{g}_0) - 2\alpha \cdot \Psi_L(\mathbf{g}_l) \geq \|y^\delta - T(K\mathbf{g}_0)\|_Y^2 \geq 0,$$

in contradiction to our assumption  $J_\alpha^s(\mathbf{g}_l, \mathbf{g}_0) \rightarrow -\infty$ , and thus  $J_\alpha^s(\mathbf{g}, \mathbf{g}_0)$  is bounded from below. By the same argument, we find  $J_\alpha^s(\mathbf{g}_l, \mathbf{g}_0) \geq 2\alpha \cdot \Psi_L(\mathbf{g}_l)$  for any sequence  $\mathbf{g}_l$  with  $\|\mathbf{g}_l\|_{(\ell_2)^n} \rightarrow \infty$ , and by (4.2.3) we conclude  $J_\alpha^s(\mathbf{g}_l, \mathbf{g}_0) \rightarrow \infty$ , i.e. the functional is coercive and has a minimizer  $\mathbf{g}_1$ .

As in (4.2.11), we get by using (4.2.12),

$$\begin{aligned} J_\alpha^s(\mathbf{g}_1, \mathbf{g}_0) - 2\alpha \Psi_L(\mathbf{g}_1) &\geq \|y^\delta - T(K\mathbf{g}_0)\|_Y^2 \\ &\quad - 2B \|y^\delta - T(K\mathbf{g}_0)\|_Y \|T'(K\mathbf{g}_0)\| \|\mathbf{g}_1 - \mathbf{g}_0\|_{(\ell_2)^n} \\ &\quad - B^2 L \|y^\delta - T(K\mathbf{g}_0)\|_Y \|\mathbf{g}_1 - \mathbf{g}_0\|_{(\ell_2)^n}^2 + C \|\mathbf{g}_1 - \mathbf{g}_0\|_{(\ell_2)^n}^2. \end{aligned}$$

By (4.2.7),  $C/2 \geq B^2 L \|y^\delta - T(K\mathbf{g}_0)\|_Y$ , and thus,

$$\begin{aligned} J_\alpha^s(\mathbf{g}_1, \mathbf{g}_0) - 2\alpha \cdot \Psi_L(\mathbf{g}_1) &\geq \|y^\delta - T(K\mathbf{g}_0)\|_Y^2 \\ &\quad - 2B \|y^\delta - T(K\mathbf{g}_0)\|_Y \|T'(K\mathbf{g}_0)\| \|\mathbf{g}_1 - \mathbf{g}_0\|_{(\ell_2)^n} \\ &\quad + \frac{C}{2} \|\mathbf{g}_1 - \mathbf{g}_0\|_{(\ell_2)^n}^2. \end{aligned}$$

As  $\mathbf{g}_0 \in K_r$ , it follows from (4.2.7) that  $B \|T'(K\mathbf{g}_0)\| \leq \sqrt{C/2}$  holds, and consequently,

$$\begin{aligned} J_\alpha^s(\mathbf{g}_1, \mathbf{g}_0) - 2\alpha \cdot \Psi_L(\mathbf{g}_1) &\geq \|y^\delta - T(K\mathbf{g}_0)\|_Y^2 - 2 \frac{\sqrt{C}}{\sqrt{2}} \|y^\delta - T(K\mathbf{g}_0)\|_Y \|\mathbf{g}_1 - \mathbf{g}_0\|_{(\ell_2)^n} \\ &\quad + \frac{C}{2} \|\mathbf{g}_1 - \mathbf{g}_0\|_{(\ell_2)^n}^2 \\ &= \left( \|y^\delta - T(K\mathbf{g}_0)\|_Y - \frac{\sqrt{C}}{\sqrt{2}} \|\mathbf{g}_1 - \mathbf{g}_0\|_{(\ell_2)^n} \right)^2 \geq 0. \end{aligned}$$

In particular,

$$2 \min\{\alpha_i\} \|\Psi_L(\mathbf{g}_1)\|_{\ell_1} \leq 2\alpha \cdot \Psi_L(\mathbf{g}_1) \leq J_\alpha^s(\mathbf{g}_1, \mathbf{g}_0) = \min_{\mathbf{g}} J_\alpha^s(\mathbf{g}, \mathbf{g}_0) \leq J_\alpha^s(\mathbf{g}_0, \mathbf{g}_0) = J_\alpha(\mathbf{g}_0),$$

i.e.  $\|\Psi_L(\mathbf{g}_1)\|_{\ell_1} \leq J_\alpha(\mathbf{g}_0)/(2 \min\{\alpha_i\}) = r$ , and thus,  $\mathbf{g}_1 \in K_r$ . Next, thanks to Lemma 11,

$$C \|\mathbf{g}_1 - \mathbf{g}_0\|_{\ell_2}^2 - \|T(K\mathbf{g}_1) - T(K\mathbf{g}_0)\|_Y^2 \geq 0 \quad \text{and} \quad J_\alpha(\mathbf{g}_1) \leq J_\alpha^s(\mathbf{g}_1, \mathbf{g}_0),$$

and thus,

$$\|y^\delta - T(K\mathbf{g}_1)\|_Y^2 \leq J_\alpha(\mathbf{g}_1) \leq J_\alpha^s(\mathbf{g}_1, \mathbf{g}_0) \leq J_\alpha^s(\mathbf{g}_0, \mathbf{g}_0) \leq \|y^\delta - T(K\mathbf{g}_0)\|_Y^2 + 2\alpha \cdot \Psi_L(\mathbf{g}_0),$$

and combining this estimate with the definition of  $C$  in (4.2.7) yields

$$2B^2L\|y^\delta - T(K\mathbf{g}_1)\|_Y \leq 2B^2L\sqrt{\|y^\delta - T(K\mathbf{g}_0)\|_Y^2 + 2\alpha \cdot \Psi_L(\mathbf{g}_0)} \leq C. \quad (4.2.13)$$

For applying the induction step, assume that for all  $i = 1, \dots, k-1$ , the following properties hold true:

$$\mathbf{g}_i \in K_r \quad (4.2.14)$$

$$C\|\mathbf{g}_i - \mathbf{g}_{i-1}\|_{(\ell_2)^n}^2 - \|T(K\mathbf{g}_i) - T(K\mathbf{g}_{i-1})\|_Y^2 \geq 0 \quad (4.2.15)$$

$$2B^2L\|y^\delta - T(K\mathbf{g}_i)\|_Y \leq C, \quad (4.2.16)$$

where  $\mathbf{g}_i$  denotes a minimizer of the functional  $J_\alpha^s(\mathbf{g}, \mathbf{g}_{i-1})$ . For  $i = 1$ , these properties have already been shown. As for the case  $i = 1$ , we have to show that the functional  $J_\alpha^s(\mathbf{g}, \mathbf{g}_{k-1})$  has a minimizer. First, we show that it is bounded from below: As in (4.2.11),

$$\begin{aligned} J_\alpha^s(\mathbf{g}, \mathbf{g}_{k-1}) - 2\alpha \cdot \Psi_L(\mathbf{g}) &\geq \|y^\delta - T(K\mathbf{g}_{k-1})\|_Y^2 \\ &\quad - 2\|y^\delta - T(K\mathbf{g}_{k-1})\|_Y \|T(K\mathbf{g}_{k-1}) - T(K\mathbf{g})\|_Y \\ &\quad + C\|\mathbf{g} - \mathbf{g}_{k-1}\|_{(\ell_2)^n}^2. \end{aligned}$$

By Taylor expansion,

$$\|T(K\mathbf{g}_{k-1}) - T(K\mathbf{g})\|_Y \leq \|T'(K\mathbf{g}_{k-1})\|_Y \|\mathbf{g}_{k-1} - \mathbf{g}\|_{(\ell_2)^n} + \frac{B^2L}{2} \|\mathbf{g}_{k-1} - \mathbf{g}\|_{(\ell_2)^n}^2. \quad (4.2.17)$$

Let us now assume that  $J_\alpha^s(\mathbf{g}, \mathbf{g}_{k-1})$  is not bounded from below. As in the case  $k = 1$ , there exists a sequence  $\{\mathbf{g}_l\}_{l \in \mathbb{N}}$  with  $\|\mathbf{g}_l\|_{(\ell_2)^n} \rightarrow \infty$  and  $J_\alpha^s(\mathbf{g}_l, \mathbf{g}_{k-1}) \rightarrow -\infty$ . In particular, for  $l$  large enough, follows from (4.2.17)

$$\|T(K\mathbf{g}_{k-1}) - T(K\mathbf{g}_l)\|_Y \leq B^2L\|\mathbf{g}_{k-1} - \mathbf{g}_l\|_{(\ell_2)^n}^2,$$

and combining this estimate with (4.2.17) yields

$$\begin{aligned} J_\alpha^s(\mathbf{g}_l, \mathbf{g}_{k-1}) - 2\alpha \cdot \Psi_L(\mathbf{g}_l) &\geq \|y^\delta - T(K\mathbf{g}_{k-1})\|_Y^2 \\ &\quad - 2BL\|y^\delta - T(K\mathbf{g}_{k-1})\|_Y \|\mathbf{g}_l - \mathbf{g}_{k-1}\|_{(\ell_2)^n}^2 \\ &\quad + C\|\mathbf{g}_l - \mathbf{g}_{k-1}\|_{(\ell_2)^n}^2. \end{aligned}$$

By (4.2.16),  $2B^2L\|y^\delta - T(K\mathbf{g}_{k-1})\|_Y \leq C$  and thus

$$J_\alpha^s(\mathbf{g}_l, \mathbf{g}_{k-1}) - 2\alpha \cdot \Psi_L(\mathbf{g}_l) \geq \|y^\delta - T(K\mathbf{g}_{k-1})\|_Y^2 \geq 0,$$

in contradiction to our assumption  $J_\alpha^s(\mathbf{g}_l, \mathbf{g}_{k-1}) \rightarrow -\infty$ , and thus  $J_\alpha^s(\mathbf{g}, \mathbf{g}_{k-1})$  is bounded from below. By the same argument, we find  $J_\alpha^s(\mathbf{g}_l, \mathbf{g}_{k-1}) \geq 2\alpha \cdot \Psi_L(\mathbf{g}_l) \rightarrow \infty$  for any

sequence  $\mathbf{g}_l$  with  $\|\mathbf{g}_l\|_{(\ell_2)^n} \rightarrow \infty$  and thus the functional is coercive and has a minimizer  $\mathbf{g}_k$ . As in (4.2.17), we obtain

$$\begin{aligned} J_\alpha^s(\mathbf{g}_k, \mathbf{g}_{k-1}) - 2\alpha \cdot \Psi_L(\mathbf{g}_k) &\geq \|y^\delta - T(K\mathbf{g}_{k-1})\|_Y^2 \\ &\quad - 2B\|y^\delta - T(K\mathbf{g}_{k-1})\|_Y \|T'(K\mathbf{g}_{k-1})\| \|\mathbf{g}_k - \mathbf{g}_{k-1}\|_{(\ell_2)^n} \\ &\quad - B^2 L \|y^\delta - T(K\mathbf{g}_{k-1})\|_Y \|\mathbf{g}_k - \mathbf{g}_{k-1}\|_{(\ell_2)^n}^2 \\ &\quad + C \|\mathbf{g}_k - \mathbf{g}_{k-1}\|_{(\ell_2)^n}^2. \end{aligned}$$

By (4.2.7) and assumption (4.2.16) we have  $C/2 \geq B^2 L \|y^\delta - T(K\mathbf{g}_{k-1})\|_Y$ , and thus

$$\begin{aligned} J_\alpha^s(\mathbf{g}_k, \mathbf{g}_{k-1}) - 2\alpha \cdot \Psi_L(\mathbf{g}_k) &\geq \|y^\delta - T(K\mathbf{g}_{k-1})\|_Y^2 \\ &\quad - 2B\|y^\delta - T(K\mathbf{g}_{k-1})\|_Y \|T'(K\mathbf{g}_{k-1})\| \|\mathbf{g}_k - \mathbf{g}_{k-1}\|_{(\ell_2)^n} \\ &\quad + \frac{C}{2} \|\mathbf{g}_k - \mathbf{g}_{k-1}\|_{(\ell_2)^n}^2. \end{aligned}$$

As  $\mathbf{g}_{k-1} \in K_r$ , it follows from (4.2.7) that  $B\|T'(K\mathbf{g}_{k-1})\| \leq \sqrt{C/2}$ , and we consequently have

$$J_\alpha^s(\mathbf{g}_k, \mathbf{g}_{k-1}) - 2\alpha \cdot \Psi_L(\mathbf{g}_k) \geq \left( \|y^\delta - T(K\mathbf{g}_{k-1})\|_Y - \frac{\sqrt{C}}{\sqrt{2}} \|\mathbf{g}_k - \mathbf{g}_{k-1}\|_{(\ell_2)^n} \right)^2 \geq 0.$$

In particular, it follows by (4.2.15),

$$\begin{aligned} 2 \min\{\alpha_i\} \|\Psi_L(\mathbf{g}_k)\|_{\ell_1} &\leq 2\alpha \cdot \Psi_L(\mathbf{g}_k) \leq J_\alpha^s(\mathbf{g}_k, \mathbf{g}_{k-1}) = \min_{\mathbf{g}} J_\alpha^s(\mathbf{g}, \mathbf{g}_{k-1}) \leq J_\alpha^s(\mathbf{g}_{k-1}, \mathbf{g}_{k-1}) \\ &= J_\alpha^s(\mathbf{g}_{k-1}, \mathbf{g}_{k-2}) \leq J_\alpha^s(\mathbf{g}_{k-2}, \mathbf{g}_{k-2}) \leq \dots \leq J_\alpha^s(\mathbf{g}_0, \mathbf{g}_0) \end{aligned}$$

i.e.  $\|\Psi_L(\mathbf{g}_k)\|_{\ell_1} \leq J_\alpha(\mathbf{g}_0)/(2 \min\{\alpha_i\}) = r$ , and thus,  $\mathbf{g}_k \in K_r$ . As in Lemma 11, it follows

$$C \|\mathbf{g}_k - \mathbf{g}_{k-1}\|_{(\ell_2)^n}^2 - \|T(K\mathbf{g}_k) - T(K\mathbf{g}_{k-1})\|_Y^2 \geq 0$$

and

$$J_\alpha(\mathbf{g}) \leq J_\alpha^s(\mathbf{g}, \mathbf{g}_{k-1}),$$

and we obtain

$$\begin{aligned} \|y^\delta - T(K\mathbf{g}_k)\|_Y^2 &\leq J_\alpha(\mathbf{g}_k) \leq J_\alpha^s(\mathbf{g}_k, \mathbf{g}_{k-1}) \leq J_\alpha^s(\mathbf{g}_{k-1}, \mathbf{g}_{k-1}) \leq \dots \leq J_\alpha^s(\mathbf{g}_0, \mathbf{g}_0) \\ &= \|y^\delta - T(K\mathbf{g}_0)\|_Y^2 + 2\alpha \cdot \Psi_L(\mathbf{g}_0), \end{aligned} \quad (4.2.18)$$

and combining this estimate with the definition of  $C$  (4.2.7) yields

$$2B^2 L \|y^\delta - T(K\mathbf{g}_k)\|_Y \leq 2B^2 L \sqrt{\|y^\delta - T(K\mathbf{g}_0)\|_Y^2 + 2\alpha \cdot \Psi_L(\mathbf{g}_0)} \leq C, \quad (4.2.19)$$

i.e. we have shown that the assumptions (4.2.14)-(4.2.16) hold also for  $i = k$ .  $\square$

As an immediate consequence out of the latter proof we have

**Corollary 12** *The sequences of functionals  $\{J_\alpha(\mathbf{g}_k)\}_{k=0,1,2,\dots}$  and  $\{J_\alpha^s(\mathbf{g}_{k+1}, \mathbf{g}_k)\}_{k=0,1,2,\dots}$  are non-increasing.*

### 4.2.3 Minimization yields Projected Fixed Point Iterations

In this section, we elaborate necessary conditions for a minimizer of the functional  $J_\alpha^s(\mathbf{g}, \mathbf{a})$ .

**Lemma 13** *The necessary condition for a minimum of  $J_\alpha^s(\mathbf{g}, \mathbf{a})$  is given by*

$$0 \in -F_j T'(K\mathbf{g})^*(y^\delta - T(K\mathbf{a})) + C\mathbf{g}_j - C\mathbf{a}_j + \alpha_j \mathbf{L}_j^* \partial \Psi_j(\mathbf{L}_j \mathbf{g}_j) \quad , \text{ for all } j = 1, \dots, n . \quad (4.2.20)$$

*Proof.* In the notion of subgradients (which is allowed, see later on for a convexity result), we have for  $j = 1, \dots, n$ ,

$$\partial_j J_\alpha^s(\mathbf{g}, \mathbf{a}) = -2F_j T'(K\mathbf{g})^*(y^\delta - T(K\mathbf{a})) + 2C\mathbf{g}_j - 2C\mathbf{a}_j + 2\alpha_j \partial \Theta_j(\mathbf{g}_j) .$$

Consequently, through  $\mathbf{v} \in \partial \Theta_j(\mathbf{g}_j) \Leftrightarrow \mathbf{L}_j \mathbf{v} \in \partial \Psi_j(\mathbf{L}_j \mathbf{g}_j)$ , the necessary condition (4.2.20) follows immediately.  $\square$

Before giving an equivalent condition, we will have a closer look to the relation between the functionals  $\Psi_j$  and associated closed convex sets  $\mathcal{C}_j$ . We may consider the Fenchel or so-called dual functional of  $\Psi_j$ , which we will denote by  $\Psi_j^*$ . Since we have assumed  $\Psi_j$  to be a positive and one homogeneous functional, there exists a convex set  $\mathcal{C}_j$  such that  $\Psi_j^*$  is equal to the indicator function  $\chi_{\mathcal{C}_j}$  over  $\mathcal{C}_j$ . Moreover, in Hilbert space lore, we have total duality between convex sets and positive and one homogeneous functionals, i.e.  $\Psi_j = (\chi_{\mathcal{C}_j})^*$ .

**Lemma 14** *Let  $M_j(\mathbf{g}, \mathbf{a}) := F_j T'(K\mathbf{g})^*(y^\delta - T(K\mathbf{a}))/C + \mathbf{a}_j$ , then the necessary conditions (4.2.20) can be casted as*

$$\mathbf{g}_j = \frac{\alpha_j}{C} \mathbf{L}_j^* (I - P_{\mathcal{C}_j}) \left( \frac{C}{\alpha_j} \mathbf{L}_j M_j(\mathbf{g}, \mathbf{a}) \right) , \quad j = 1, \dots, n . \quad (4.2.21)$$

where  $P_{\mathcal{C}_j}$  is the orthogonal projection onto the convex set  $\mathcal{C}_j$ .

*Proof.* With the shorthand  $M_j(\mathbf{g}, \mathbf{a})$  we may rewrite (4.2.20) for each  $j$ ,

$$\mathbf{L}_j \frac{M_j(\mathbf{g}, \mathbf{a}) - \mathbf{g}_j}{\frac{\alpha_j}{C}} \in \partial \Psi_j(\mathbf{L}_j \mathbf{g}_j) ,$$

and thus, by standard arguments in convex analysis,

$$\frac{C}{\alpha_j} \mathbf{L}_j \mathbf{g}_j \in \frac{C}{\alpha_j} \partial \Psi_j^* \left( \mathbf{L}_j \frac{M_j(\mathbf{g}, \mathbf{a}) - \mathbf{g}_j}{\frac{\alpha_j}{C}} \right) .$$

In order to have an expression by means of projections (or generalized shrinkage operations), we expand the latter formula as follows,

$$\begin{aligned} \mathbf{L}_j \frac{M_j(\mathbf{g}, \mathbf{a})}{\frac{\alpha_j}{C}} &\in \mathbf{L}_j \frac{M_j(\mathbf{g}, \mathbf{a}) - \mathbf{g}_j}{\frac{\alpha_j}{C}} + \frac{C}{\alpha_j} \partial \Psi_j^* \left( \mathbf{L}_j \frac{M_j(\mathbf{g}, \mathbf{a}) - \mathbf{g}_j}{\frac{\alpha_j}{C}} \right) \\ &= \left( I + \frac{C}{\alpha_j} \partial \Psi_j^* \right) \left( \mathbf{L}_j \frac{M_j(\mathbf{g}, \mathbf{a}) - \mathbf{g}_j}{\frac{\alpha_j}{C}} \right) , \end{aligned}$$

which is equivalent to

$$\left(I + \frac{C}{\alpha_j} \partial \Psi_j^*\right)^{-1} \left(\mathbf{L}_j \frac{M_j(\mathbf{g}, \mathbf{a})}{\frac{\alpha_j}{C}}\right) = \mathbf{L}_j \frac{M_j(\mathbf{g}, \mathbf{a}) - \mathbf{g}_j}{\frac{\alpha_j}{C}}.$$

Again, by standard results in convex analysis, it is known that  $\left(I + \frac{C}{\alpha_j} \partial \Psi_j^*\right)^{-1}$  is nothing than the orthogonal projection onto the associated convex set  $\mathcal{C}_j$ , and hence the assertion follows,

$$\mathbf{g}_j = \frac{\alpha_j}{C} \mathbf{L}_j^* (I - P_{\mathcal{C}_j}) \left(\mathbf{L}_j \frac{M_j(\mathbf{g}, \mathbf{a})}{\frac{\alpha_j}{C}}\right).$$

□

The latter lemma states that for minimizing (4.2.4) we need to solve a system of  $n$  fixed point equations (4.2.21), which are nonlinearly coupled via the  $P_{\mathcal{C}_j}$ . To condense the notation a little, we introduce nonlinear operators (and call them generalized shrinkage operators)

$$\mathbf{S}_j := \mathbf{S}_{\alpha_j, \mathbf{L}_j, \mathcal{C}_j} = \frac{\alpha_j}{C} \mathbf{L}_j^* (I - P_{\mathcal{C}_j}) \mathbf{L}_j \frac{C}{\alpha_j}.$$

Thus, we may write

$$\mathbf{g} = (\mathbf{S}_1(M_1(\mathbf{g}, \mathbf{a})), \dots, \mathbf{S}_n(M_n(\mathbf{g}, \mathbf{a}))) .$$

Let us now consider the associated fixed point map

$$\Phi(\mathbf{g}, \mathbf{a}) = (\mathbf{S}_1(M_1(\mathbf{g}, \mathbf{a})), \dots, \mathbf{S}_n(M_n(\mathbf{g}, \mathbf{a}))) .$$

**Lemma 15** *For some generic  $\mathbf{a}$ , the operator  $\Phi(\cdot, \mathbf{a})$  is a contraction if  $B^2 L / C \sqrt{J_\alpha(\mathbf{a})} < 1$ , i.e.*

$$\|\Phi(\mathbf{g}, \mathbf{a}) - \Phi(\tilde{\mathbf{g}}, \mathbf{a})\|_{(\ell_2)^n} \leq q \|\mathbf{g} - \tilde{\mathbf{g}}\|_{(\ell_2)^n} \quad \text{if} \quad q := \frac{B^2 L}{C} \sqrt{J_\alpha(\mathbf{a})} < 1 .$$

Before proving this lemma, we need a result on projections onto convex sets.

**Lemma 16** *Let  $K$  be a closed and convex set, then the mapping  $I - P_K$  is non-expansive.*

This Lemma can be deduced by the following two standard properties of convex sets.

**Lemma 17** *Let  $K$  be a closed and convex set in some Hilbert space  $H$ , then for all  $u \in H$  and all  $k \in K$  the inequality  $\langle u - P_K u, k - P_K u \rangle \leq 0$  holds true.*

*Proof.* For all  $\lambda \in [0, 1]$  one has

$$\|u - ((1 - \lambda)P_K u + \lambda k)\|^2 \geq \|u - P_K u\|^2 .$$

Thus, for all  $\lambda \in [0, 1]$

$$-2\lambda \langle u - P_K u, k - P_K u \rangle + \lambda^2 \|k - P_K u\|^2 \geq 0,$$

and therewith we have  $\langle u - P_K u, k - P_K u \rangle \leq 0$ .

□



**Lemma 18** *Let  $K$  be a closed and convex set, then for all  $u, v \in H$  the inequality*

$$\|u - v - (P_K u - P_K v)\| \leq \|u - v\|$$

*holds true.*

*Proof.* We need to prove

$$-2\langle u - v, P_K u - P_K v \rangle + \|P_K u - P_K v\|^2 \leq 0 .$$

By Lemma 17 we have  $\langle u - P_K u, P_K v - P_K u \rangle \leq 0$ , or equivalently

$$-\langle u, P_K u \rangle + \langle u, P_K v \rangle + \|P_K u\|^2 - \langle P_K u, P_K v \rangle \leq 0 .$$

By symmetry we have

$$-\langle v, P_K v \rangle + \langle v, P_K u \rangle + \|P_K v\|^2 - \langle P_K v, P_K u \rangle \leq 0 .$$

Summing the two inequalities leads to

$$-\langle u - v, P_K u - P_K v \rangle + \|P_K u - P_K v\|^2 \leq 0 ,$$

and thus

$$-2\langle u - v, P_K u - P_K v \rangle + \|P_K u - P_K v\|^2 \leq -\|P_K u - P_K v\|^2 \leq 0 .$$

□

Thanks to Lemma 18, we still have assured Lemma 16, and with Lemma 16 at hand, we are able to prove Lemma 15.

*Proof.* We have by Lemma 16 and the Lipschitz-continuity of  $T'$ ,

$$\begin{aligned} & \|\Phi(\mathbf{g}, \mathbf{a}) - \Phi(\tilde{\mathbf{g}}, \mathbf{a})\|_{(\ell_2)^n}^2 = \\ &= \sum_{j=1}^n \|\mathbf{S}_j(\mathbf{g}, \mathbf{a}) - \mathbf{S}_j(\tilde{\mathbf{g}}, \mathbf{a})\|_{\ell_2}^2 \\ &= \sum_{j=1}^n \frac{\alpha_j}{C} \left\| (I - P_{\mathcal{C}_j}) \left( \mathbf{L}_j \frac{M_j(\mathbf{g}, \mathbf{a})}{\frac{\alpha_j j}{C}} \right) - (I - P_{\mathcal{C}_j}) \left( \mathbf{L}_j \frac{M_j(\tilde{\mathbf{g}}, \mathbf{a})}{\frac{\alpha_j}{C}} \right) \right\|_{\ell_2}^2 \\ &\leq \sum_{j=1}^n \|M_j(\mathbf{g}, \mathbf{a}) - M_j(\tilde{\mathbf{g}}, \mathbf{a})\|_{\ell_2}^2 \\ &\leq \sum_{j=1}^n \frac{B_j}{C^2} \|T'(K\mathbf{g}) - T'(K\tilde{\mathbf{g}})\|^2 \|y^\delta - T(K\mathbf{a})\|_Y^2 \\ &\leq \sum_{j=1}^n \frac{B_j L^2}{C^2} \left( \sum_{i=1}^n B_i^{1/2} \|\mathbf{g}_i - \tilde{\mathbf{g}}_i\|_{\ell_2} \right)^2 J_\alpha(\mathbf{a}) \leq \frac{B^4 L^2}{C^2} \|\mathbf{g} - \tilde{\mathbf{g}}\|_{(\ell_2)^n}^2 J_\alpha(\mathbf{a}) \end{aligned} \tag{4.2.22}$$

and the assertion follows. □

**Proposition 4.2.2** *The fixed point map  $\Phi(\mathbf{g}, \mathbf{g}_k)$  is for all  $k = 0, 1, 2, \dots$  a contraction.*

*Proof.* By the definition of  $C$  in (4.2.7) and Lemma 15 (setting  $\mathbf{a} = \mathbf{g}_0$ ), we deduce that  $\Phi(\mathbf{g}, \mathbf{g}_0)$  is a contraction with

$$q = \frac{B^2 L}{C} \sqrt{J_\alpha(\mathbf{g}_0)} \leq \frac{1}{2} < 1.$$

With the help of Corollary 12, we complete the proof

$$\begin{aligned} \|\Phi(\mathbf{g}, \mathbf{g}_k) - \Phi(\tilde{\mathbf{g}}, \mathbf{g}_k)\|_{(\ell_2)^n} &\leq \frac{B^2 L}{C} \sqrt{J_\alpha(\mathbf{g}_k)} \|\mathbf{g} - \tilde{\mathbf{g}}\|_{(\ell_2)^n} \\ &\leq \dots \leq \frac{B^2 L}{C} \sqrt{J_\alpha(\mathbf{g}_0)} \|\mathbf{g} - \tilde{\mathbf{g}}\|_{(\ell_2)^n} \leq \frac{1}{2} \|\mathbf{g} - \tilde{\mathbf{g}}\|_{\ell_2}. \end{aligned}$$

□

Up to here, we do know whether our fixed point iteration converges towards a critical point of  $J_\alpha^s(\mathbf{g}, \mathbf{g}_k)$ .

**Proposition 4.2.3** *The necessary equation (4.2.21) for a minimum of the functional  $J_\alpha^s(\mathbf{g}, \mathbf{g}_k)$  has a unique fixed point, and the fixed point iteration converges towards the minimizer.*

*Proof.* To verify this assertion, we have to investigate the Taylor expansion of  $J_\alpha^s$  more closely. By Taylor's expansion for  $T$  and the Lipschitz-continuity of  $T'$  we get

$$T(K\mathbf{g} + K\mathbf{h}) = T(K\mathbf{g}) + T'(K\mathbf{g})K\mathbf{h} + R(K\mathbf{g}, K\mathbf{h}) \quad (4.2.23)$$

with

$$\|R(K\mathbf{g}, K\mathbf{h})\|_Y \leq \frac{B^2 L}{2} \|\mathbf{h}\|_{(\ell_2)^n}^2. \quad (4.2.24)$$

Denoting with  $\nabla$  the multi-valued (sub)gradient (still having in mind that the subgradient is set-valued) and with  $\mathbf{g}_k$  the  $k$ -th iterate ( $\mathbf{g}_j$  indicates the  $j$ -th component of  $\mathbf{g}$ ),

$$\begin{aligned} J_\alpha^s(\mathbf{g} + \mathbf{h}, \mathbf{g}_k) - J_\alpha^s(\mathbf{g}, \mathbf{g}_k) &= \nabla J_\alpha^s(\mathbf{g}, \mathbf{g}_k) \cdot \mathbf{h} + C \|\mathbf{h}\|_{(\ell_2)^n}^2 - 2 \langle y^\delta - T(K\mathbf{g}_k), R(K\mathbf{g}, K\mathbf{h}) \rangle_Y \\ &\quad + 2 \sum_{j=1}^n \alpha_j \{ \Theta_j(\mathbf{g}_j + \mathbf{h}_j) - \Theta(\mathbf{g}_j) - \partial \Theta_j(\mathbf{g}_j) \mathbf{h}_j \} \\ &\geq \nabla J_\alpha^s(\mathbf{g}, \mathbf{g}_k) \cdot \mathbf{h} + C \|\mathbf{h}\|_{(\ell_2)^n}^2 - 2 \|y^\delta - T(K\mathbf{g}_k)\|_{\ell_2} \frac{B^2 L}{2} \|\mathbf{h}\|_{\ell_2}^2 \\ &\quad + 2 \sum_{j=1}^n \alpha_j \{ \Theta_j(\mathbf{g}_j + \mathbf{h}_j) - \Theta(\mathbf{g}_j) - \partial \Theta_j(\mathbf{g}_j) \mathbf{h}_j \} \\ &\geq \nabla J_\alpha^s(\mathbf{g}, \mathbf{g}_k) \cdot \mathbf{h} + \frac{C}{2} \|\mathbf{h}\|_{(\ell_2)^n}^2 \\ &\quad + 2 \sum_{j=1}^n \alpha_j \{ \Theta_j(\mathbf{g}_j + \mathbf{h}_j) - \Theta(\mathbf{g}_j) - \partial \Theta_j(\mathbf{g}_j) \mathbf{h}_j \}. \end{aligned}$$

Assuming  $\mathbf{g}$  is a critical point, i.e.  $\nabla J_\alpha^s(\mathbf{g}, \mathbf{g}_k) \cdot \mathbf{h} = 0$  for all  $\mathbf{h}$ , we have

$$J_\alpha^s(\mathbf{g} + \mathbf{h}, \mathbf{g}_k) - J_\alpha^s(\mathbf{g}, \mathbf{g}_k) \geq \frac{C}{2} \|\mathbf{h}\|_{(\ell_2)^n}^2 + 2 \sum_{j=1}^n \alpha_j \{ \Theta_j(\mathbf{g}_j + \mathbf{h}_j) - \Theta(\mathbf{g}_j) - \partial \Theta_j(\mathbf{g}_j) \mathbf{h}_j \} .$$

By the definition of subgradients (for each individual  $j$ ): an element  $\mathbf{v} \in \ell_2$  belongs to  $\partial \Theta_j(\mathbf{g}_j)$  if and only if for all  $\mathbf{x} \in \ell_2$ ,

$$\Theta_j(\mathbf{g}_j) + \langle \mathbf{v}, \mathbf{x} - \mathbf{g}_j \rangle_{\ell_2} \leq \Theta_j(\mathbf{x}) ,$$

and, in particular for  $\mathbf{x} = \mathbf{g}_j + \mathbf{h}_j$ , this yields for all  $\mathbf{v} \in \partial \Theta_j(\mathbf{g}_j)$  and all  $\mathbf{h}_j \in \ell_2$ ,

$$\Theta_j(\mathbf{g}_j) + \langle \mathbf{v}, \mathbf{h}_j \rangle_{\ell_2} \leq \Theta_j(\mathbf{g}_j + \mathbf{h}_j) \text{ or, equivalently, } 0 \leq \Theta_j(\mathbf{g}_j + \mathbf{h}_j) - \Theta_j(\mathbf{g}_j) - \partial \Theta_j(\mathbf{g}_j) \mathbf{h}_j .$$

Consequently,

$$J_\alpha^s(\mathbf{g} + \mathbf{h}, \mathbf{g}_k) - J_\alpha^s(\mathbf{g}, \mathbf{g}_k) \geq \frac{C}{2} \|\mathbf{h}\|_{(\ell_2)^n}^2 ,$$

and thus every critical point is a global minimizer of  $J_\alpha^s(\mathbf{g}, \mathbf{g}_k)$ , and, again by the latter inequality, there exists only one global minimizer.  $\square$

By assuming more regularity on  $T$ , the latter statement can be improved:

**Proposition 4.2.4** *Let  $T$  be a twice continuously differentiable operator. Then the functional  $J_\alpha^s(\mathbf{g}, \mathbf{g}_k)$  is strictly convex.*

*Proof.* Since the non-convex part of  $J_\alpha^s$  is the discrepancy  $\|y^\delta - T(K\mathbf{g})\|_Y^2$ , it remains to show that

$$J^d(\mathbf{g}) := \|y^\delta - T(K\mathbf{g})\|_Y^2 + C\|\mathbf{g} - \mathbf{g}_k\|_{\ell_2}^2 - \|T(K\mathbf{g}) - T(K\mathbf{g}_k)\|_Y^2 \quad (4.2.25)$$

is strictly convex in  $\mathbf{g}$ , i.e. we have to show that

$$J^d((1 - \lambda)\mathbf{g}_1 + \lambda\mathbf{g}_2) < (1 - \lambda)J^d(\mathbf{g}_1) + \lambda J^d(\mathbf{g}_2)$$

holds for  $\lambda \in (0, 1)$  and arbitrary  $\mathbf{g}_1, \mathbf{g}_2 \in (\ell_2)^n$ . At first, we express  $J^d$  by its Taylor expansion,

$$J^d(\mathbf{g} + \mathbf{h}) = J^d(\mathbf{g}) + \nabla J^d(\mathbf{g}) \cdot \mathbf{h} + r(\mathbf{g}, \mathbf{h}) , \quad (4.2.26)$$

where

$$r(\mathbf{g}, \mathbf{h}) := -2\langle y^\delta - T(K\mathbf{g}_k), R(K\mathbf{g}, K\mathbf{h}) \rangle_Y + C\|\mathbf{h}\|_{(\ell_2)^n}^2 . \quad (4.2.27)$$

We have

$$\begin{aligned} J^d((1 - \lambda)\mathbf{g}_1 + \lambda\mathbf{g}_2) &= J^d(\mathbf{g}_1 + \lambda(\mathbf{g}_2 - \mathbf{g}_1)) = J^d(\mathbf{g}_2 + (1 - \lambda)(\mathbf{g}_1 - \mathbf{g}_2)) \\ &= (1 - \lambda)J^d(\mathbf{g}_1 + \lambda(\mathbf{g}_2 - \mathbf{g}_1)) + \lambda J^d(\mathbf{g}_2 + (1 - \lambda)(\mathbf{g}_1 - \mathbf{g}_2)) \end{aligned} \quad (4.2.28)$$

and with

$$\begin{aligned} J^d(\mathbf{g}_1 + \lambda(\mathbf{g}_2 - \mathbf{g}_1)) &= J^d(\mathbf{g}_1) + \lambda \nabla J^d(\mathbf{g}_1) \cdot (\mathbf{g}_2 - \mathbf{g}_1) + r(\mathbf{g}_1, \lambda(\mathbf{g}_2 - \mathbf{g}_1)) \\ J^d(\mathbf{g}_2 + (1 - \lambda)(\mathbf{g}_1 - \mathbf{g}_2)) &= J^d(\mathbf{g}_2) + (1 - \lambda) \nabla J^d(\mathbf{g}_2) \cdot (\mathbf{g}_1 - \mathbf{g}_2) \\ &\quad + r(\mathbf{g}_2, (1 - \lambda)(\mathbf{g}_1 - \mathbf{g}_2)) \end{aligned}$$

we obtain

$$\begin{aligned} J^d((1-\lambda)\mathbf{g}_1 + \lambda\mathbf{g}_2) &= (1-\lambda)J^d(\mathbf{g}_1) + \lambda J^d(\mathbf{g}_2) \\ &\quad + \lambda(1-\lambda) [\nabla J^d(\mathbf{g}_1) - \nabla J^d(\mathbf{g}_2)] \cdot (\mathbf{g}_2 - \mathbf{g}_1) \\ &\quad + (1-\lambda)r(\mathbf{g}_1, \lambda(\mathbf{g}_2 - \mathbf{g}_1)) + \lambda r(\mathbf{g}_2, (1-\lambda)(\mathbf{g}_1 - \mathbf{g}_2)) . \end{aligned}$$

Thus,  $J_\alpha^s$  is strictly convex if for all  $\lambda \in (0, 1)$ ,

$$\begin{aligned} D(\mathbf{g}_1, \mathbf{g}_2, \lambda) &:= \lambda(1-\lambda) [\nabla J^d(\mathbf{g}_1) - \nabla J^d(\mathbf{g}_2)] \cdot (\mathbf{g}_2 - \mathbf{g}_1) \\ &\quad + (1-\lambda)r(\mathbf{g}_1, \lambda(\mathbf{g}_2 - \mathbf{g}_1)) + \lambda r(\mathbf{g}_2, (1-\lambda)(\mathbf{g}_1 - \mathbf{g}_2)) < 0 . \end{aligned}$$

We have

$$\begin{aligned} [\nabla J^d(\mathbf{g}_1) - \nabla J^d(\mathbf{g}_2)] \cdot (\mathbf{g}_2 - \mathbf{g}_1) &= -2C\|\mathbf{g}_2 - \mathbf{g}_1\|_{(\ell_2)^n}^2 \\ &\quad - 2\langle y^\delta - T(K\mathbf{g}_k), (T'(K\mathbf{g}_1) - T'(K\mathbf{g}_2))K(\mathbf{g}_2 - \mathbf{g}_1) \rangle_Y . \end{aligned}$$

As  $T$  is twice continuously Fréchet differentiable, it is

$$T'(K\mathbf{g}_1) = T'(K\mathbf{g}_2) + \int_0^1 T''(K\mathbf{g}_2 + \tau K(\mathbf{g}_1 - \mathbf{g}_2))(K(\mathbf{g}_1 - \mathbf{g}_2), \cdot) d\tau$$

and thus,

$$\begin{aligned} [\nabla J^d(\mathbf{g}_1) - \nabla J^d(\mathbf{g}_2)] \cdot (\mathbf{g}_2 - \mathbf{g}_1) &= \\ &= -2C\|\mathbf{g}_2 - \mathbf{g}_1\|_{(\ell_2)^n}^2 + 2\langle y^\delta - T(K\mathbf{g}_k), \int_0^1 T''(K\mathbf{g}_2 + \tau K(\mathbf{g}_1 - \mathbf{g}_2))(K(\mathbf{g}_1 - \mathbf{g}_2))^2 d\tau \rangle , \end{aligned} \tag{4.2.29}$$

where we have used the shorthand  $T''(\cdot)(\cdot, \cdot) = T''(\cdot)(\cdot)^2$ . Again, as  $T$  is twice continuously Fréchet-differentiable, the function  $R(K\mathbf{g}, K\mathbf{h})$  in (4.2.27) is given by

$$R(K\mathbf{g}, K\mathbf{h}) = \int_0^1 (1-\tau)T''(K\mathbf{g} + \tau K\mathbf{h})(K\mathbf{h})^2 d\tau ,$$

and thus we obtain

$$\begin{aligned} R(K\mathbf{g}_1, \lambda K(\mathbf{g}_2 - \mathbf{g}_1)) &= \lambda^2 \int_0^1 (1-\tau)T''(K\mathbf{g}_1 + \tau \lambda K(\mathbf{g}_2 - \mathbf{g}_1))(K(\mathbf{g}_2 - \mathbf{g}_1))^2 d\tau \\ &= \int_{1-\lambda}^1 (\tau - (1-\lambda))T''(K\mathbf{g}_2 + \tau K(\mathbf{g}_1 - \mathbf{g}_2))(K(\mathbf{g}_1 - \mathbf{g}_2))^2 d\tau \end{aligned} \tag{4.2.30}$$

and in the same way

$$R(K\mathbf{g}_2, (1-\lambda)K(\mathbf{g}_1 - \mathbf{g}_2)) = \int_0^{1-\lambda} (1-\lambda-\tau)T''(K\mathbf{g}_2 + \tau K(\mathbf{g}_1 - \mathbf{g}_2))(K(\mathbf{g}_1 - \mathbf{g}_2))^2 d\tau . \quad (4.2.31)$$

Combining definition (4.2.27) and equations (4.2.29), (4.2.30) and (4.2.31) yields

$$D(\mathbf{g}_1, \mathbf{g}_2, \lambda) = -\lambda(1-\lambda)C\|\mathbf{g}_1 - \mathbf{g}_2\|_{(\ell_2)^n}^2 + 2\langle y^\delta - T(K\mathbf{g}_k), f(\mathbf{g}_1, \mathbf{g}_2, \lambda) \rangle_Y , \quad (4.2.32)$$

where

$$\begin{aligned} f(\mathbf{g}_1, \mathbf{g}_2, \lambda) &:= \lambda(1-\lambda) \int_0^1 T''(K\mathbf{g}_2 + \tau K(\mathbf{g}_1 - \mathbf{g}_2))(K(\mathbf{g}_1 - \mathbf{g}_2))^2 d\tau \\ &\quad - (1-\lambda) \int_{1-\lambda}^1 (\tau - (1-\lambda))T''(K\mathbf{g}_2 + \tau K(\mathbf{g}_1 - \mathbf{g}_2))(K(\mathbf{g}_1 - \mathbf{g}_2))^2 d\tau \\ &\quad - \lambda \int_0^{1-\lambda} (1-\lambda-\tau)T''(K\mathbf{g}_2 + \tau K(\mathbf{g}_1 - \mathbf{g}_2))(K(\mathbf{g}_1 - \mathbf{g}_2))^2 d\tau . \end{aligned}$$

The functional  $f(\mathbf{g}_1, \mathbf{g}_2, \lambda)$  can now be recasted as follows

$$\begin{aligned} f(x_1, x_2, \lambda) &= \lambda \int_0^{1-\lambda} \tau T''(K\mathbf{g}_2 + \tau K(\mathbf{g}_1 - \mathbf{g}_2))(K(\mathbf{g}_1 - \mathbf{g}_2))^2 d\tau \\ &\quad + (1-\lambda) \int_{1-\lambda}^1 (1-\tau)T''(K\mathbf{g}_2 + \tau K(\mathbf{g}_1 - \mathbf{g}_2))(K(\mathbf{g}_1 - \mathbf{g}_2))^2 d\tau . \end{aligned}$$

In order to estimate  $\|f(\mathbf{g}_1, \mathbf{g}_2, \lambda)\|_Y$  it is necessary to estimate the integrals separately. Due to the Lipschitz-continuity of the first derivative, the second derivative can be globally estimated by  $L$ , and it follows,

$$\begin{aligned} \left\| \int_0^{1-\lambda} \tau T''(K\mathbf{g}_2 + \tau K(\mathbf{g}_1 - \mathbf{g}_2))(K(\mathbf{g}_1 - \mathbf{g}_2))^2 d\tau \right\|_Y &\leq \frac{(1-\lambda)^2}{2} B^2 L \|\mathbf{g}_1 - \mathbf{g}_2\|_{(\ell_2)^n}^2 \\ \left\| \int_{1-\lambda}^1 (1-\tau)T''(K\mathbf{g}_2 + \tau K(\mathbf{g}_1 - \mathbf{g}_2))(K(\mathbf{g}_1 - \mathbf{g}_2))^2 d\tau \right\|_Y &\leq \frac{\lambda^2}{2} B^2 L \|\mathbf{g}_1 - \mathbf{g}_2\|_{(\ell_2)^n}^2 \end{aligned}$$

and thus

$$\|f(\mathbf{g}_1, \mathbf{g}_2, \lambda)\|_Y \leq \frac{\lambda(1-\lambda)}{2} B^2 L \|\mathbf{g}_1 - \mathbf{g}_2\|_{(\ell_2)^n}^2 . \quad (4.2.33)$$

Combining (4.2.32) and (4.2.33) yields for  $\lambda \in (0, 1)$

$$\begin{aligned}
D(\mathbf{g}_1, \mathbf{g}_2, \lambda) &\leq -\lambda(1-\lambda)C\|\mathbf{g}_1 - \mathbf{g}_2\|_{(\ell_2)^n}^2 + 2\|y^\delta - T(K\mathbf{g}_k)\|_Y \|f(\mathbf{g}_1, \mathbf{g}_2, \lambda)\|_Y \\
&\leq -\lambda(1-\lambda)C\|\mathbf{g}_1 - \mathbf{g}_2\|_{(\ell_2)^n}^2 \\
&\quad + \frac{\lambda(1-\lambda)}{2} 2B^2L\|y^\delta - T(K\mathbf{g}_k)\|_Y \|\mathbf{g}_1 - \mathbf{g}_2\|_{(\ell_2)^n}^2 \\
&\stackrel{(4.2.19)}{\leq} -\lambda(1-\lambda)\frac{C}{2}\|\mathbf{g}_1 - \mathbf{g}_2\|_{(\ell_2)^n}^2 < 0,
\end{aligned}$$

and thus the functional is strictly convex.  $\square$

#### 4.2.4 Convergence Analysis

Within this section we discuss convergence properties of the proposed scheme, i.e. we aim to show that the sequence of iterates  $\{\mathbf{g}_k\}$  converges strongly towards a critical point of  $J_\alpha$ , at least.

**Lemma 19** *The sequence of iterates  $\{\mathbf{g}_k\}_{k=0,1,2,\dots}$  has a weakly convergent subsequence.*

*Proof.* This is an immediate consequence of Proposition 4.2.1, in which we have shown that for  $k = 0, 1, 2, \dots$  the iterates  $\mathbf{g}_k$  are contained in  $K_r$ , and by requirement (4.2.3),  $\|\mathbf{g}_k\|_{(\ell_2)^n} \leq r$ . Since the iterates are uniformly bounded, we deduce that there exists at least one accumulation point  $\mathbf{g}_\alpha^*$  with  $\mathbf{g}_{k,l} \xrightarrow{w} \mathbf{g}_\alpha^*$ , where  $\mathbf{g}_{k,l}$  denotes a subsequence of  $\mathbf{g}_k$ .  $\square$

**Lemma 20** *For the iterates  $\mathbf{g}_k$  holds  $\lim_{k \rightarrow \infty} \|\mathbf{g}_{k+1} - \mathbf{g}_k\|_{(\ell_2)^n} = 0$ .*

*Proof.* With the help of Corollary 12, we observe that

$$\begin{aligned}
0 &\leq \sum_{k=0}^N \{C\|\mathbf{g}_{k+1} - \mathbf{g}_k\|_{(\ell_2)^n}^2 - \|T(K\mathbf{g}_{k+1}) - T(K\mathbf{g}_k)\|_Y^2\} \\
&\leq \sum_{k=0}^N \{J_\alpha^s(\mathbf{g}_{k+1}, \mathbf{g}_k) - J_\alpha(\mathbf{g}_{k+1})\} \leq \sum_{k=0}^N \{J_\alpha(\mathbf{g}_k) - J_\alpha(\mathbf{g}_{k+1})\} \\
&= J_\alpha(\mathbf{g}_0) - J_\alpha(\mathbf{g}_{N+1}) \leq J_\alpha(\mathbf{g}_0),
\end{aligned}$$

i.e. the finite sums are uniformly bounded (independent on  $N$ ). Now, by the Taylor expansion of  $T$ , we have

$$\|T(K\mathbf{g}_{k+1}) - T(K\mathbf{g}_k)\|_Y^2 \leq \frac{C}{2} \|\mathbf{g}_{k+1} - \mathbf{g}_k\|_{(\ell_2)^n}^2,$$

and thus

$$0 \leq \frac{C}{2} \|\mathbf{g}_{k+1} - \mathbf{g}_k\|_{(\ell_2)^n}^2 \leq C\|\mathbf{g}_{k+1} - \mathbf{g}_k\|_{(\ell_2)^n}^2 - \|T(K\mathbf{g}_{k+1}) - T(K\mathbf{g}_k)\|_Y^2 \longrightarrow 0$$

as  $k \rightarrow \infty$  and the assertion follows.  $\square$

**Lemma 21** *Every subsequence of  $\mathbf{g}_k$  has a convergent subsequence  $\mathbf{g}_{k,l}$  that converges strongly towards a function  $\mathbf{g}_\alpha^*$ , and  $\mathbf{g}_\alpha^*$  satisfies the necessary condition for a minimizer of  $J_\alpha$ :*

$$F_j T'(K\mathbf{g}_\alpha^*)^*(y^\delta - T(K\mathbf{g}_\alpha^*)) \in \alpha_j \partial \Theta_j((\mathbf{g}_j)_\alpha^*) , \quad j = 1, \dots, n . \quad (4.2.34)$$

*Proof.* According to Lemma 13, the minimizer  $\mathbf{g}_{k+1}$  of  $J_\alpha^s(\mathbf{g}, \mathbf{g}_k)$  fulfills

$$0 \in F_j T'(K\mathbf{g}_{k+1})^*(y^\delta - T(K\mathbf{g}_k)) - C(\mathbf{g}_j)_{k+1} + C(\mathbf{g}_j)_k - \alpha_j \partial \Theta_j((\mathbf{g}_j)_{k+1}).$$

Thus, for all  $j = 1, \dots, n$ ,

$$\begin{aligned} (\mathbf{g}_j)_{k+1} - (\mathbf{g}_j)_k &\in \frac{1}{C} \left( F_j T'(K\mathbf{g}_{k+1})^*(y^\delta - T(K\mathbf{g}_{k+1})) - \alpha_j \partial \Theta_j((\mathbf{g}_j)_{k+1}) \right. \\ &\quad \left. + F_j T'(K\mathbf{g}_{k+1})^*(T(K\mathbf{g}_{k+1}) - T(K\mathbf{g}_k)) \right) \end{aligned} \quad (4.2.35)$$

and, moreover, by Lemma 20,

$$\|F_j T'(K\mathbf{g}_{k+1})^*(T(K\mathbf{g}_{k+1}) - T(K\mathbf{g}_k))\|_Y \leq \frac{CB_j^{1/2}}{2B} \|\mathbf{g}_{k+1} - \mathbf{g}_k\|_{(\ell_2)^n} \rightarrow 0 .$$

Passing to the limit  $k \rightarrow \infty$  in (4.2.35),

$$0 \in \lim_{k \rightarrow \infty} \left( F_j T'(K\mathbf{g}_{k+1})^*(y^\delta - T(K\mathbf{g}_{k+1})) - \alpha_j \partial \Theta_j((\mathbf{g}_j)_{k+1}) \right) . \quad (4.2.36)$$

Since  $\mathbf{g}_k$  is bounded, every subsequence has a weakly convergent subsequence. Let  $\mathbf{g}_{k,l}$  denote such a weakly convergent subsequence with weak limit  $\mathbf{g}_\alpha^*$  (for simplicity, we will denote this sequence by  $\mathbf{g}_k$ , too). Since

$$\begin{aligned} F_j T'(K\mathbf{g}_{k+1})^*(y^\delta - T(K\mathbf{g}_{k+1})) &= \\ F_j T'(K\mathbf{g}_{k+1})^*(y^\delta - T(K\mathbf{g}_\alpha^*)) &+ F_j T'(K\mathbf{g}_{k+1})^*(T(K\mathbf{g}_\alpha^*) - T(K\mathbf{g}_{k+1})) , \end{aligned}$$

and because of

$$\|F_j T'(K\mathbf{g}_{k+1})^*(T(K\mathbf{g}_\alpha^*) - T(K\mathbf{g}_{k+1}))\|_{\ell_2} \leq \frac{\sqrt{CB_j}}{\sqrt{2}B} \|T(K\mathbf{g}_\alpha^*) - T(K\mathbf{g}_{k+1})\| \rightarrow 0$$

and by assumption (4.2.5), i.e.

$$F_j T'(K\mathbf{g}_{k+1})^*(y^\delta - T(K\mathbf{g}_\alpha^*)) \rightarrow F_j T'(K\mathbf{g}_\alpha^*)^*(y^\delta - T(K\mathbf{g}_\alpha^*)),$$

we consequently obtain

$$\lim_{k \rightarrow \infty} F_j T'(K\mathbf{g}_{k+1})^*(y^\delta - T(K\mathbf{g}_{k+1})) = F_j T'(K\mathbf{g}_\alpha^*)^*(y^\delta - T(K\mathbf{g}_\alpha^*)) . \quad (4.2.37)$$

Next, we have to consider  $\lim_{k \rightarrow \infty} \partial \Theta_j((\mathbf{g}_j)_k)$ . By an elementwise consideration we have,  $\mathbf{v} \in \partial \Theta_j((\mathbf{g}_j)_k)$  if and only if for all  $\mathbf{x} \in \ell_2$  the inequality

$\Theta_j(\mathbf{x}) \geq \partial\Theta_j((\mathbf{g}_j)_k) + \langle \mathbf{v}, \mathbf{x} - (\mathbf{g}_j)_k \rangle_{\ell_2}$  holds true. The assumption that  $\Theta_j$  is lower semi-continuous and convex implies weak lower semi-continuity of all the  $\Theta_j$ , i.e.  $\Theta_j((\mathbf{g}_j)_\alpha^*) \leq \lim_{k \rightarrow \infty} \inf \Theta_j((\mathbf{g}_j)_k) \leq \lim_{k \rightarrow \infty} \Theta_j((\mathbf{g}_j)_k)$ . The same holds true for the  $\ell_2$ -inner product. Thus, we deduce that for all  $\mathbf{v} \in \lim_{k \rightarrow \infty} \Theta_j((\mathbf{g}_j)_k)$  we have  $\mathbf{v} \in \partial\Theta_j((\mathbf{g}_j)_\alpha^*)$ , i.e.  $\lim_{k \rightarrow \infty} \partial\Theta_j((\mathbf{g}_j)_k) \subseteq \partial\Theta_j((\mathbf{g}_j)_\alpha^*)$ . Combining (4.2.37) with (4.2.35) proves that  $\mathbf{g}_{k,l}$  converges, and as  $\mathbf{g}_\alpha^*$  is the weak limit of the sequence,  $\mathbf{g}_{k,l} \rightarrow \mathbf{g}_\alpha^*$ . Equations (4.2.34) follow by passing to the limit in (4.2.36).  $\square$

In principle, the limits of different convergent subsequences of  $\mathbf{g}_k$  may differ. Let  $\mathbf{g}_{k,l} \rightarrow \mathbf{g}_\alpha^*$  be a subsequence of  $\mathbf{g}_k$ , and let  $\tilde{\mathbf{g}}_{k,l}$  the predecessor of  $\mathbf{g}_{k,l}$  in  $\mathbf{g}_k$ , i.e.  $\mathbf{g}_{k,l} = \mathbf{g}_i$  and  $\tilde{\mathbf{g}}_{k,l} = \mathbf{g}_{i-1}$ . Then we observe,  $J_\alpha^s(\mathbf{g}_{k,l}, \tilde{\mathbf{g}}_{k,l}) \rightarrow J_\alpha(\mathbf{g}_\alpha^*)$ . Moreover, as we have  $J_\alpha^s(\mathbf{g}_{k+1}, \mathbf{g}_k) \leq J_\alpha^s(\mathbf{g}_k, \mathbf{g}_{k-1})$  for all  $k$ , it turns out that the value of the Tikhonov functional for every limit  $\mathbf{g}_\alpha^*$  of a convergent subsequence remains the same, i.e.  $J_\alpha(\mathbf{g}_\alpha^*) = \text{const}$ .

We may now summarize our findings and give a simple criterion that ensures strong convergence of the whole sequence  $\{\mathbf{g}_k\}$ .

**Theorem 22** *Assume that there exists at least one isolated limit  $\mathbf{g}_\alpha^*$  of a subsequence  $\mathbf{g}_{k,l}$  of  $\mathbf{g}_k$ . Then  $\mathbf{g}_k \rightarrow \mathbf{g}_\alpha^*$  as  $k \rightarrow \infty$ . The accumulation point  $\mathbf{g}_\alpha^*$  is a minimizer for the functional  $J_\alpha^s(\mathbf{g}, \mathbf{g}_\alpha^*)$ .*

*Proof.* As in the proof of Proposition 4.2.3 we obtain,  $J_\alpha^s(x_\alpha^* + h, x_\alpha^*) \geq J_\alpha^s(x_\alpha^*, x_\alpha^*) + \frac{C}{2} \|h\|^2$ . The remaining proof of norm convergence can be directly taken from [RT05b].  $\square$

## 4.2.5 A Regularization result

After stating norm convergence results for the proposed multi-frame approach for solving nonlinear operator equations, we now focus on how to optimally choose the parameter vector  $\alpha$ . As still considered in Section 2.2.4, the vector  $\alpha$  plays the most important role in computing stabilized solutions. Again we have to identify a functional relation between  $\alpha$  and the noise floor  $\delta$ , i.e.  $\alpha = \alpha(\delta)$  with  $\alpha(\delta) \rightarrow 0$  and  $\|\mathbf{g}^{*,\alpha(\delta)} - \mathbf{g}^\dagger\| \rightarrow 0$  as  $\alpha \rightarrow 0$ . If we find a parameter rule achieving this, then the suggested iteration scheme would regularize the ill-posed problem. Moreover, as long as we deal with frames, i.e. even if the inverse problem would have a unique solution, the corresponding vector of frame sequences to represent this solution will never have. Thus it is only reasonable to show that we approach one solution  $\mathbf{g}^\dagger$  when passing to the limit  $\delta \rightarrow 0$ .

The next theorem provides conditions on the functional relation  $\alpha(\delta)$  for which the constructed Landweber fixed point iteration with projections in each step is a regularization scheme (up to uniqueness).



**Theorem 23** *Let  $y^\delta \in Y$  with  $\|y^\delta - y\|_Y \leq \delta$ ,  $\alpha_{\min}(\delta) = \min_j \{\alpha_j(\delta)\}$ ,  $\alpha_{\max}(\delta) = \max_j \{\alpha_j(\delta)\}$ , and assume  $\alpha(\delta) = (\alpha_1(\delta), \dots, \alpha_n(\delta))$  is chosen such that*

$$\alpha(\delta) \xrightarrow{\delta \rightarrow 0} 0, \quad \delta^2/\alpha_{\min}(\delta) \xrightarrow{\delta \rightarrow 0} 0, \quad \alpha_{\max}(\delta)/\alpha_{\min}(\delta) \xrightarrow{\delta \rightarrow 0} 1.$$

*Then every sequence  $\{g^{*,\alpha(\delta)}\}$  of minimizers of the functional  $J_\alpha(\mathbf{g})$  where  $\delta \rightarrow 0$  and  $\alpha = \alpha(\delta)$  has a convergent subsequence. The limit of every convergent subsequence is a solution of  $T(K\mathbf{g}) = y$  with minimal values of  $\Psi(\mathbf{L}_j \mathbf{g}_j)$ ,  $j = 1, \dots, n$ .*

*Proof.* As  $g^{*,\alpha(\delta)} = (g_1^{*,\alpha(\delta)}, \dots, g_n^{*,\alpha(\delta)})$  is a minimizer of  $J_\alpha$ , we have

$$\|y^\delta - T(Kg^{*,\alpha(\delta)})\|_Y^2 + 2\alpha \cdot \Psi_{\mathbf{L}}(g^{*,\alpha(\delta)}) \leq \delta^2 + 2\alpha \cdot \Psi_{\mathbf{L}}(\mathbf{g}^\dagger). \quad (4.2.38)$$

Thus, by the made assumptions on  $\alpha(\delta)$ , we achieve

$$\lim_{\delta \rightarrow 0} T(Kg^{*,\alpha(\delta)}) = y.$$

Again by (4.2.38),

$$\|\Psi_{\mathbf{L}}(g^{*,\alpha(\delta)})\|_{\ell_1} \leq \frac{\delta^2}{2\alpha_{\min}(\delta)} + \frac{\alpha_{\max}(\delta)}{\alpha_{\min}(\delta)} \|\Psi_{\mathbf{L}}(\mathbf{g}^\dagger)\|_{\ell_1}$$

implying,

$$\limsup_{\delta \rightarrow 0} \|g^{*,\alpha(\delta)}\|_{(\ell_2)^n} \leq \limsup_{\delta \rightarrow 0} \|\Psi_{\mathbf{L}}(g^{*,\alpha(\delta)})\|_{\ell_1} \leq \|\Psi_{\mathbf{L}}(\mathbf{g}^\dagger)\|_{\ell_1},$$

i.e.  $\|g^{*,\alpha(\delta)}\|_{(\ell_2)^n}$  are uniformly bounded. Consequently, the sequence has a weakly convergent subsequence (again denoted by  $\{g^{*,\alpha(\delta)}\}$ ) with weak limit  $\mathbf{g}^\circ$ ,

$$\mathbf{g}^\circ = w - \lim_{\delta \rightarrow 0} g^{*,\alpha(\delta)}.$$

Since  $T$  is strongly continuous,

$$y = \lim_{\delta \rightarrow 0} T(Kg^{*,\alpha(\delta)}) = T(K\mathbf{g}^\circ),$$

i.e.  $\mathbf{g}^\circ$  is a solution of  $T(K\mathbf{g}) = y$ . Assume now  $\mathbf{g}^\dagger$  is a solution of the inverse problem with minimal values of  $\Psi_j(\mathbf{L}_j \cdot)$ . Then, since all the  $\Psi_j$  are weak semi-continuous, we deduce

$$\Psi_j(\mathbf{L}_j \mathbf{g}_j^\circ) \leq \limsup_{\delta \rightarrow 0} \Psi_j(\mathbf{L}_j g_j^{*,\alpha(\delta)}) \leq \Psi_j(\mathbf{L}_j \mathbf{g}_j^\dagger) \leq \Psi_j(\mathbf{L}_j \mathbf{g}_j^\circ) \text{ for } j = 1, \dots, n.$$

Hence  $\mathbf{g}^\circ$  is also a solution with minimal values of  $\Psi_j(\mathbf{L}_j \cdot)$ . □

### Resulting regularization method:

We may now summarize our findings and suggest the following regularization method. Assume that all the conditions we have imposed in the previous sections apply to our problem and, moreover, assume we have a parameter rule at hand that fulfills the conditions of Theorem 23. Then the algorithm goes as follows:

- Define a sequence  $\{\alpha_n\}$  satisfying the condition of Theorem 23, and pick  $r \geq 1$ ,  $\mathbf{g}_0$
- **while**  $\|y^\delta - T(K\mathbf{g}^{*,\alpha})\| > r \cdot \delta$ 
  - $\alpha = \alpha_n$
  - pick an admissible  $C$
  - $[\mathbf{g}^{*,\alpha}] = \text{Iteration}(T, y^\delta, C, \alpha, \mathbf{g}_0)$ :
 
$$\mathbf{g}_{k+1} = \arg \min_{\mathbf{g}} J_\alpha^s(\mathbf{g}, \mathbf{g}_k) \text{ (solved by a projected fixed point iteration)}$$

$$\mathbf{g}^{*,\alpha} = \lim_{k \rightarrow \infty} \mathbf{g}_k$$
  - $x_0 = x_\alpha^*$
- end**

In practice (treatment of limits), we have to incorporate stopping rules that will slightly modify this scheme:

- Define a sequence  $\{\alpha_n\}$  satisfying the condition of Theorem 23, and pick  $r \geq 1$ ,  $\mathbf{g}_0$ , and additionally two tolerances  $\tau_1, \tau_2$
- **while**  $\|y^\delta - T(K\mathbf{g}^{*,\alpha})\| > r \cdot \delta$ 
  - $\alpha = \alpha_n$
  - pick an admissible  $C$
  - $[\mathbf{g}_\alpha^*] = \text{Iteration}(T, F, y^\delta, C, \alpha, \tau_1, \tau_2)$ 
    - $k = 0$
    - while**  $\|\mathbf{g}_{k+1} - \mathbf{g}_k\|_{\ell_2} > \tau_1$ 
      - $l = 0$ ,  $\mathbf{g}_{k,0} = \mathbf{g}_k$
      - while**  $\|\mathbf{g}_{k,l} - \mathbf{g}_{k,l+1}\|_{\ell_2} > \tau_2$ 
        - $l = l + 1$
        - $\mathbf{g}_{k,l} = \Phi_{\alpha,C}(\mathbf{g}_{k,l-1}, \mathbf{g}_k)$
      - end**
      - $\mathbf{g}_{k+1} = \mathbf{g}_{k,l}$
      - $k = k + 1$
    - end**
    - $\mathbf{g}_\alpha^* = \mathbf{g}_k$
  - end**

# Chapter 5

## Applications II: Nonlinear Problems

This chapter shall illustrate the wide range of applicability where the iterative fixed point schemes developed in Sections 4.1 and 4.2 can be applied.

### 5.1 Damped Landweber Fixed Point Iteration for SPECT

In SPECT, the patient gets a radiopharmaceutical, which is distributed through the whole body by the blood flow, and is finally enriched in some areas of interest. The task is to recover the distribution of the radiopharmaceutical (or, in short of the activity function  $f$ ) from measurements of the radioactivity outside the body. In contrast to CT, where the measured intensity depends only on the intensity of the incoming X-ray and the density  $\mu$  of the tissue along the path of the X-ray, the measurement for SPECT depend on the activity function  $f$  (which describes the distribution of the radiopharmaceutical) and the density  $\mu$  of the tissue. The measured data  $y$  and the tuple  $(f, \mu)$  are connected via the attenuated Radon Transform (ATRT),

$$y = R(f, \mu)(s, \omega) = \int_{\mathbb{R}} f(s\omega^\perp + t\omega) e^{-\int_t^\infty \mu(s\omega^\perp + r\omega) dr} dt, \quad (5.1.1)$$

where  $s \in \mathbb{R}$  and  $\omega \in S^1$ . Usually both  $f$  and  $\mu$  are unknown functions, and  $R$  is a nonlinear operator. In order to invert (5.1.1), two strategies can be used. Firstly, the density distribution can be determined by the inversion of a additional CT scan (in most scanners, the CT data is gathered simultaneously). With this approach, one has to solve two linear problems, as the attenuated Radon transform is linear if  $\mu$  is assumed to be known, and currently developed inversion formulas can be used [Nat01]. However, attaching an X-ray source to a SPECT scanner makes them much more expensive. The scanning time for each patient might increase, which leads again to higher costs for each scan. Thus the second strategy, where the ATRT is treated as a nonlinear operator, seems to be promising. The drawbacks of this strategy are the non-uniqueness of the operator (which usually leads to a wrong reconstruction for the density function  $\mu$ ) and much higher computational costs for the inversion of the nonlinear operator. In

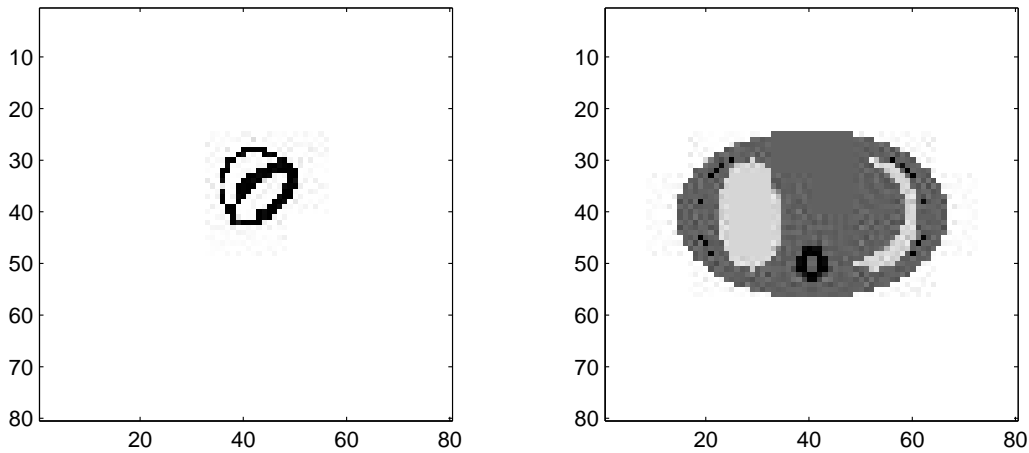


Figure 5.1: Activity function  $f_*$  (left) and attenuation function  $\mu_*$  (right)

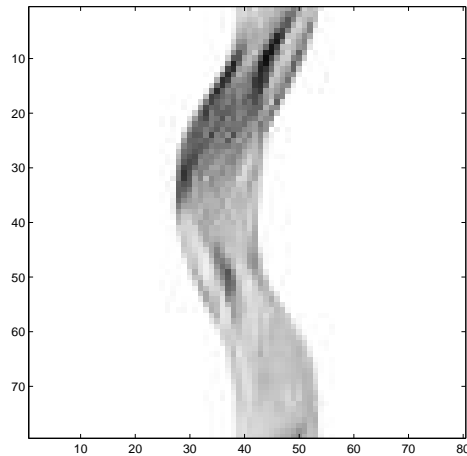


Figure 5.2: Generated data  $g(s, \omega) = R(f_*, \mu_*)(s, \omega)$ .

the last decade, several ideas for solving the nonlinear problem (5.1.1) were discussed, see, e.g., [CGLT79, MY93, WCCG96, WCNG97, RCNB00]. Dicken [Dic99] showed that Tikhonov regularization for nonlinear operators can be used for the reconstruction of the activity function. Methods for the computation of a minimizer of the Tikhonov functional were proposed in [Ram02b, Ram03, Ram04] and applied to SPECT. Here, we will only demonstrate that our method can be used for the computation of a minimizer. For the test computations, we would like to use the so called MCAT phantom [TTP<sup>+</sup>90], see Figure 5.1. The belonging sinogram data is shown in Figure 5.2.

In a first attempt, we want to compute the minimizer of the Tikhonov functional with regularization parameter  $\alpha = 3430$ . The data was contaminated with multiplicative Gaussian noise with relative error  $\delta_{rel} = 5\%$  (here  $\delta_{rel} = \|y^\delta - y\|/\|y\|$ ). The inner iteration was terminated if the relative distance of two consecutive iterates was less than  $1e - 6$ , and the outer iteration was terminated if the relative distance between two consecutive

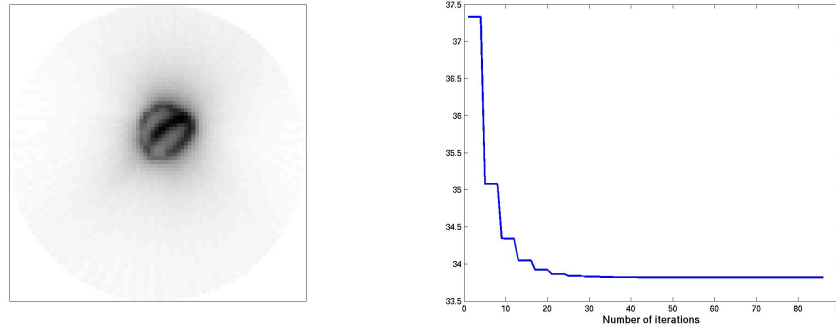


Figure 5.3: Minimizer  $f_{\alpha}^{\delta}$  of the Tikhonov functional for  $\alpha = 3430$  (l) and values of the Tikhonov functional (r)

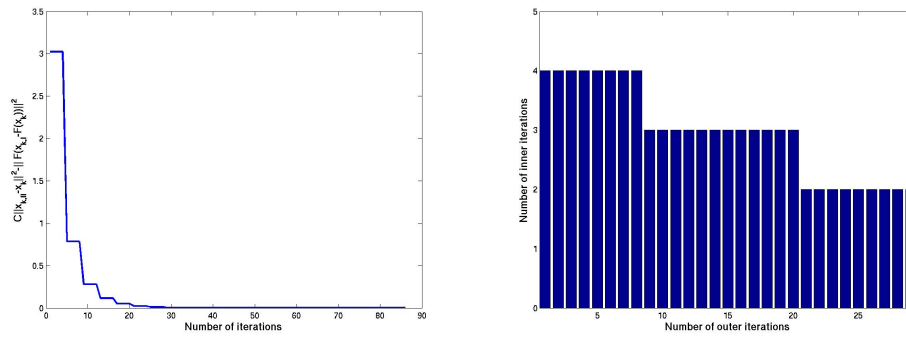


Figure 5.4: Plot of the additive term in the replacement functional (l) and number of inner iterations for each outer iteration (r)

outer iterates was less than  $1e - 5$ . After only a few iterations, the value of the Tikhonov functional remains almost constant, see Figure 5.3. The values of the additive term  $C\|x_{k,l} - x_k\| - \|F(x_{k,l}) - F(x_k)\|$ ,  $x_k = (f_k, \mu_k)$  is shown in Figure 5.4. Clearly, the additive term converges fast to zero, and thus the values of the replacement functional and the Tikhonov functional are almost the same. Moreover, it turns out that we only need a few inner iterations to achieve the required accuracy, see Figure 5.4. This actually indicates that the whole iteration itself is quite fast. In a final test computation, we used Morozov's discrepancy principle to determine an appropriate regularization parameter (see (4.1.58)). For a sequence  $\alpha_k = a_0 q^k$ ,  $k = 0, 1, \dots$ ,  $a_0 = 1000$  and  $q = 0.5$  we computed  $x_{\alpha_k}^{\delta}$  by **TIREFU**, and picked the first minimizer  $x_{\alpha_k}^{\delta}$  with (4.1.58) and  $c = 2$ . In our case, we had to compute 10 minimizing functions. The residual of the minimizer with  $\alpha = 1.95$  was smaller than  $2\delta$  for the first time, and the reconstruction was stopped. Figure 5.5 shows the results.

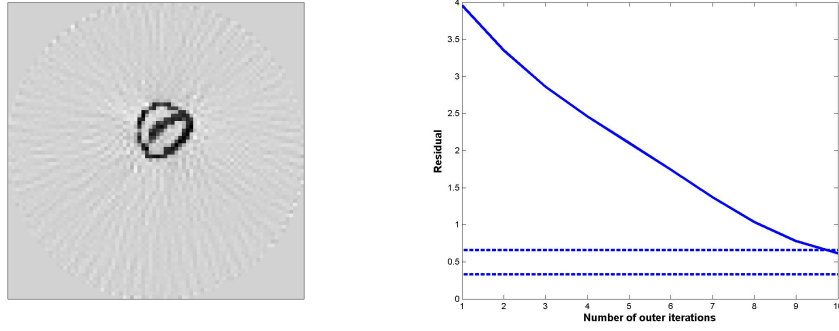


Figure 5.5: Left: Reconstruction according to Morozov's discrepancy principle. Right: Plot of the residual  $\|y^\delta - R(f_{\alpha_k}^\delta, \mu_{\alpha_k}^\delta)\|$ . The dashed lines mark the region  $[\delta, 2\delta]$ .

## 5.2 Computation of Optimally Localized Wavelets

Within this section we provide an analytical fundament which can be used to construct optimally localized coherent states. It turns out that a way to approximate these states is given by the iterative approach of Section 4.1.

### 5.2.1 Motivation and Basic Formulas

It is well known that a nonzero state cannot be arbitrarily well localized simultaneously in space and Fourier domain. This fact may be quantified by the Heisenberg uncertainty relation,

$$\Delta x \Delta k \geq 2\pi.$$

There are functions which are optimally localized in phase space in that they satisfy the inequality with the lower bound. For a detailed analysis on uncertainty relations in the context of Gabor and wavelet transforms we refer, e.g. to [GMP85, GMP86, DM95, Tes05a]. In this paper we want to discuss more general measures of uncertainty or delocalization in phase space and we shall prove the existence of optimally localized states.

In principle, general wavelet transforms

$$s \mapsto \mathcal{W}_g s, \quad \mathcal{W}_g s(x) = (U(x)g, s)_H, \quad x \in G$$

associated with the square integrable irreducible representation  $U$  of a locally compact group  $G$  provides a one to one correspondence between the state Hilbert space  $H$  and a reproducing kernel Hilbert space over the group. The reproducing kernel is up to a normalization the wavelet transform of the wavelet itself

$$\Pi = \mathcal{W}_g g.$$

This kernel can be interpreted as the Heisenberg box of the phase space. In this paper we shall be concerned with localization properties of these reproducing kernels. In fact the reproducing kernel of wavelet analysis cannot be arbitrarily well localized. For instance,

there is no wavelet such that the associated reproducing kernel is compactly supported. Obviously here is no universal way of quantifying localization. Instead we propose to quantify localization through the following and similar families of cost functionals

$$\Phi[g] = \sup_{x \in G} w(x)^{-1} |\mathcal{W}_g g(x)| \quad ,$$

where  $w$  is some positive weight, decaying “at  $\infty$ ”. We will prove that this functional and similar “localization” functionals are weakly lower semi-continuous. It therefore has a minimizer over any weak\* compact set. In other words: for any such measure of localization there is at least one optimal state.

Localization of wavelet transforms has been considered before. In [?, Dau88] the authors consider localizing wavelet coefficients with respect to some preassigned analyzing wavelet and a compact subset in the wavelet plane. In their approach the analyzing wavelet was fixed. In this paper, however, we consider the nonlinear problem of optimizing the localization of the reproducing kernel. Since the reproducing kernel depends quadratically on the underlying wavelet, this problem is by nature highly nonlinear and therefore an explicit expression of the optimal state seems to remain a pipe dream. However, in the last section we discuss a numerical procedure to approximately compute such optimally localized states.

Let us now recall the basic formulas of continuous wavelet transform associated with a square integrable group representation. We only recall the few facts that are necessary for this paper. For more details we refer to, e.g., [DM76], [Hol95]. Let  $G$  be a non-compact, locally compact group topological group and  $G \ni g \mapsto U(g)$  a unitary, strongly continuous, irreducible, square integrable representation in some Hilbert space  $H$ . The wavelet transform of  $s \in H$  with respect to  $g \in H$  is point-wise defined for  $x \in G$

$$\mathcal{W}_g s(x) = (U(x)g, s)_H.$$

The left and the right invariant Haar measures are denoted by  $d\lambda$  and  $d\rho$ . They are defined up to some positive factor. Over  $G$  we consider the two Hilbert spaces  $L^2(G, d\lambda)$  and  $L^2(G, d\rho)$ . We suppose that  $d\lambda$  and  $d\rho$  are scaled suitably so that the mapping  $s(x) \mapsto s(x^{-1})$  is an isometry between these two Hilbert spaces. A wavelet is called admissible, if  $\mathcal{W}_g g \in L^2(G, d\lambda)$ . Thanks to the formula

$$\mathcal{W}_g s(x) = \mathcal{W}_s g(\bar{x}^{-1}),$$

admissibility is also equivalent to  $\mathcal{W}_g g \in L^2(G, d\rho)$ . We denote the set of all admissible wavelets by  $A$ . For  $g, h \in A$  and  $s, u \in H$  the following equation holds

$$(\mathcal{W}_g s, \mathcal{W}_h u)_{L^2(G, d\lambda)} = C(g, h) (s, u)_H, \quad (5.2.1)$$

where  $C$  is a densely defined, closed, positive quadratic form. Its form domain is precisely  $A$ . We write  $c_{g,h} = C(g, h)$  and  $c_g = C(g, g)$ . By the first representation theorem there is a closed, linear operator  $T$  such that for all  $u \in D(T)$  and  $v \in A$  we have  $C(v, u) = (v, Tu)_H$ . The space  $A$  is in general a non-closed subspace of  $H$ . However, since  $C$  is a closed quadratic form it becomes a Hilbert space with respect to

$$(s, u)_A = (s, u)_H + C(s, s), \quad \|s\|_A^2 = \|s\|_H^2 + c_s.$$

Convergence in  $A$  will be understood with respect to this norm. From equation (5.2.1) it follows that in particular  $\mathcal{W}_g$  is for admissible, non-zero  $g$  a multiple of an isometry

$$\|\mathcal{W}_g s\|_{L^2(G, d\lambda)}^2 = c_g \|s\|_H^2.$$

The adjoint of the wavelet analysis is a wavelet synthesis

$$\mathcal{W}_g^* = \mathcal{M}_g.$$

Formally it can be written as follows

$$\mathcal{M}_h r(x) = \int_G r(x) U(x) h d\lambda(x).$$

We have for  $g, h \in A$

$$\mathcal{M}_h \mathcal{W}_g = c_{g,h} 1.$$

The combination  $\mathcal{W}_g \mathcal{M}_h$  can be written as non-commutative convolution operator. If we define on  $L^2(G, d\lambda) \times L^2(G, d\lambda)$  the (left) convolution product as

$$\Pi * r(x) = \int_G \Pi(y^{-1} \circ x) r(y) d\lambda(y),$$

we have for  $g, h \in A$

$$\mathcal{W}_g \mathcal{M}_h = \Pi *, \quad \Pi = \mathcal{W}_g h.$$

In particular we will use the following formula over and over

$$\mathcal{W}_h = \Pi * \mathcal{W}_g, \quad \Pi = c_{g,h}^{-1} \mathcal{W}_h h.$$

### 5.2.2 Optimally Localized States and a General Existence Theorem

Before establishing the general existence theorem we consider the particular case of localization measures through a weighted  $L^\infty$  norm. Consider therefore a positive function  $w : G \rightarrow R_+$ . We may suppose without loss of generality (see below) that  $w$  is symmetric

$$w(x^{-1}) = w(x).$$

Then  $w \in L^2(G, d\lambda)$  is equivalent with  $w \in L^2(G, d\rho)$ . Moreover, we consider symmetric weights, for which either of both (and hence both) of the following holds

$$w * w \in L^2(G, d\lambda), \quad w * w \in L^2(G, d\rho).$$

We call a weight satisfying these two conditions an admissible weight. A natural measure for the localization of a function  $r$  over  $G$  with respect to  $w$  is the following weighted norm

$$\sup_G w^{-1} |r|.$$



For fixed  $h \in A$  let  $\Sigma_h \subset H$  denote the affine subspace of co-dimension 1 defined through  $\Sigma_h = \{g \in H : (g, Th)_H = 1\}$ . Note that  $\Sigma_h$  may contain non-admissible vectors. The admissible wavelets in  $\Sigma_h$  are reconstruction wavelets for  $h$  :

$$\mathcal{M}_h \mathcal{W}_g = \mathcal{M}_g \mathcal{W}_h = 1.$$

We introduce the following functional on  $H$ .

$$\Phi[s] = \Phi_w[s] = \sup_G w^{-1} |\mathcal{W}_s s|,$$

whenever the right hand side is finite. In all other cases we set  $\Phi[s] = \infty$ . We now can formulate the main theorem.

Let  $w$  be a symmetric weight function such that  $w * w \in L^2(G, d\lambda)$ . Let  $h \in H$ ,  $h \neq 0$ , be such that  $\Phi[h] < \infty$ . Then there exists a wavelet  $g \in \Sigma_h$  such that for all  $u \in \Sigma_h$  we have

$$\Phi[g] \leq \Phi[u].$$

In other words, the localization functional  $\Phi$  has a minimizer in each  $\Sigma_h$ , for all  $h$ , which have some regularity as expressed through  $|\mathcal{W}_h h| \leq w$ . Note that we have to require, that  $\Phi[h] < \infty$ . This ensures, that the set of functions having a  $w$  localization is not empty. In turn, this is a requirement for  $w$  in which it should not be decaying too fast (e.g. compactly supported weights are not possible). The requirement that  $w$  is symmetric is no limitation of generality. Indeed we have

$$w(x)^{-1} |\mathcal{W}_g g(x)| = w(x)^{-1} |\mathcal{W}_g g(x^{-1})|.$$

Taking the sup over  $G$  can also be realized as sup over all  $x^{-1}$  with  $x \in G$ . Therefore

$$\sup_{x \in G} w(x)^{-1} |\mathcal{W}_g g(x)| = \sup_{x \in G} w(x^{-1})^{-1} |\mathcal{W}_g g(x)|$$

and thus

$$\Phi[g] = \sup_{x \in G} \tilde{w}(x)^{-1} |\mathcal{W}_g g(x)|, \quad \tilde{w}(x) = \min\{w(x), w(x^{-1})\}.$$

First we characterize the admissible wavelets by a localization property

**Lemma 24** *Let  $h \in A$ ,  $h \neq 0$  be given. Then  $g \in H$  is actually in  $A$  if and only if*

$$\|g\|^2 = \|\mathcal{W}_h g\|_{L^2(G, d\lambda)}^2 + \|\mathcal{W}_h g\|_{L^2(G, d\rho)}^2 < \infty.$$

*The square root of the left hand side defines a norm which on  $A$  is equivalent to the norm of  $A$  :*

$$c_h^{-1} \|\mathcal{W}_h g\|_{L^2(G, d\lambda)}^2 + \|h\|_H^{-2} \|\mathcal{W}_h g\|_{L^2(G, d\rho)}^2 = \|g\|_A^2.$$

*Proof.* Suppose  $g \in A$ . Thanks to the formula

$$\|\mathcal{W}_h g\|_{L^2(G, d\rho)}^2 = \|\mathcal{W}_g h\|_{L^2(G, d\lambda)}^2 = c_g \|h\|_H^2$$

and the isometric property of the wavelet transform we may conclude.  $\square$

The heart of the proof of the main theorem relies on the following compact embedding lemma.

**Lemma 25** *Let  $w > 0$  be a symmetric weight function with  $w \in L^2(G, d\lambda)$ , and let  $h \in A$ ,  $h \neq 0$ , be fixed. The set  $\Delta_{h,w} = \{s \in H : |\mathcal{W}_h s| \leq w\}$  is an  $A$ -compact subset of  $A$ .*

*Proof.* In view of the preceding lemma  $\Delta_{h,w} \subset A$ . Let  $s_n \in \Delta_{h,w}$  be an arbitrary sequence. We will show, that we may extract an  $A$  convergent subsequence whose limit is in  $\Delta_{h,w}$ .

Since  $G$  is locally compact we can find a sequence of compact sets  $K_m \subset G$  with  $K_m \subset K_{m+1}$  and  $\bigcup K_m = G$ . Upon choosing a subsequence we may require that

$$\int_{G \setminus K_m} w^2(d\lambda + d\rho) \leq 1/m.$$

Since the representation  $U$  is strongly continuous and since  $\Delta_{h,w}$  is bounded in  $H$  it follows, that the restrictions to each  $K_m$  of functions  $\mathcal{W}_h s_n$  are uniformly continuous. On each  $K_m$  we can therefore find a uniformly convergent subsequence. Therefore upon choosing a suitable diagonal subsequence we may suppose that for  $m' > m$

$$\int_{K_m} |\mathcal{W}_g s_m - \mathcal{W}_g s_{m'}|^2 d\lambda + \int_{K_m} |\mathcal{W}_g s_m - \mathcal{W}_g s_{m'}|^2 d\rho \leq 1/m.$$

We therefore have

$$\begin{aligned} \|s_m - s_{m'}\|_A^2 &= \int_{G \setminus K_m} |\mathcal{W}_g s_m - \mathcal{W}_g s_{m'}|^2 (d\lambda + d\rho) \\ &\quad + \int_{K_m} |\mathcal{W}_g s_m - \mathcal{W}_g s_{m'}|^2 (d\lambda + d\rho) \leq 2/m \end{aligned}$$

and thus  $s_m$  is a Cauchy sequence in  $A$  and hence in  $H$ . Since  $A$  is complete it has a limit  $s$  in  $A$ . Actually  $s \in \Delta_{h,w}$  since  $\mathcal{W}_h s_n(x) \rightarrow \mathcal{W}_h s(x)$  point-wise and therefore going to the limit in  $|\mathcal{W}_h s_n| \leq w$  we conclude that  $|\mathcal{W}_h s| \leq w$ .  $\square$

We now prove the theorem.

*Proof.* Since  $\Phi[h] < \infty$  we have  $h \in A$  and since  $h \neq 0$  we have  $c_h^{-1}h \in \Sigma_h$ . Therefore with some finite  $\gamma$

$$0 \leq \inf_{g \in \Sigma_h} \Phi[g] = \gamma < \infty.$$

We can find a sequence  $g_n \in \Sigma_h$  such that  $\lim_{n \rightarrow \infty} \Phi[g_n] = \gamma$ . Upon choosing a subsequence we may assume that

$$|\mathcal{W}_{g_n} g_n| \leq (\gamma + 1/n) w.$$

In particular it follows that  $g_n \in A$ . Since by hypothesis  $C(g_n, h) = (g_n, Th)_H = 1$ , for each  $n$  we have

$$\mathcal{W}_h g_n = \Pi * \mathcal{W}_{g_n} g_n, \quad \Pi = \mathcal{W}_h h.$$

Now, from  $|\mathcal{W}_{g_n} g_n| \leq (\gamma + 1/n) w$  and  $|\mathcal{W}_h h| \leq \Phi[h] w$  it follows that  $|\mathcal{W}_h g_n| \leq (\gamma + 1/n) \Phi[h] w * w$ . By hypothesis  $w * w \in L^2(G, d\lambda + d\rho)$  and therefore thanks to the compactness argument of lemma 25 we may find an  $A$  convergent subsequence  $g_{m(n)}$  with limit

$$\lim_{n \rightarrow \infty} g_{m(n)} = g \quad \text{convergence in } A,$$

and  $g \in \Delta_{h,w*w} \subset A$ . Now convergence in  $A$  implies convergence in  $H$  and thus

$$(g, Th)_H = \lim_{n \rightarrow \infty} (g_{m(n)}, Th)_H = 1,$$

and therefore  $g \in \Sigma_h$ . □

The results obtained so far can be generalized to a more abstract setting. Consider two Banach spaces  $B, K \subset L^2(G, d\lambda + d\rho)$  of functions over  $G$  with continuous embeddings.  $B$  should be a lattice,  $\| |s| \|_B = \|s\|_B$ . We then can define a localization with respect to  $B$  simply as

$$\Phi[g] = \|\mathcal{W}_g g\|_B.$$

We include the value  $\Phi = \infty$  in the natural way. For  $B$  and  $K$  we further suppose that the following holds:

**Invariance of  $B$**   $B$  should be  $G$  bi-invariant: for all  $y \in G$  there is an  $b > 0$  such that

$$\|s(y \circ \cdot)\|_B \leq b\|s\|_B, \quad \|s(\cdot \circ y)\|_B \leq b\|s\|_B.$$

It should be stable under inversion

$$\|s(x^{-1})\|_B \leq d\|s\|_B.$$

Then we have  $u(x) = s(y^{-1} \circ x \circ y)$  satisfies  $\|u\|_B \leq e\|s\|_B$ .

**Semi-continuity of  $B$  norm** Suppose further that the following inequality holds for the norm in  $B$ : if  $s_n \in B$  is any sequence of non-negative functions  $s_n \geq 0$  then consider  $s = \liminf_{n \rightarrow \infty} s_n$ . Then we require that

$$\|s\|_B = \|\liminf_{n \rightarrow \infty} s_n\|_B \leq \liminf_{n \rightarrow \infty} \|s_n\|_B.$$

In classical  $L^p$  spaces this is a direct consequence of Fatou's lemma

$$\int \liminf s_n d\mu \leq \liminf \int s_n d\mu.$$

**Compact embedding of  $K$**  For the space  $K$  we suppose that the following compact embedding property holds: let  $L \subset K$  be a  $K$ -bounded set. If now on each compact subset of  $G$  the set of functions  $L$  is uniformly continuous then  $L$  is precompact in  $L^2(G, d\lambda + d\rho)$ .

**Convolution mapping** The two spaces  $B$  and  $K$  are linked through the following convolution property: for fixed  $r \in B$  the convolution product with  $r$  is a linear operator

$$*_r : B \rightarrow K, \quad u \mapsto r * u,$$

and it is bounded  $\|r * u\|_K \leq d\|u\|_B$ , with  $d$  depending only on  $r$ .

**Nonempty** We suppose that there is at least one  $h \in D(T)$ ,  $h \neq 0$  such that  $\Phi[h] < \infty$ . This  $h$  is clearly admissible,  $h \in A$ , since  $B \subset L^2(G, d\lambda + d\rho)$ .

Under these conditions, the following theorem holds.

There is a  $g \in \Sigma_h \cap A$  such that for all  $u \in \Sigma_h$  we have

$$\Phi[g] \leq \Phi[u].$$

This function  $g$  then satisfies  $\mathcal{W}_h g \in K$ . Typical examples for such spaces  $B$  are as follows. For

$$\|r\|_{w,\infty} = \|w^{-1}r\|_{L^\infty(G)}$$

we obtain the results of the previous section, if we suppose

$$w * w \in L^2(G, d\lambda + d\rho).$$

A second useful family is given by

$$\|r\|_{w,2} = \|w^{-1/2}r\|_{L^2(d\lambda+d\rho)}.$$

We denote by  $B_{w,\infty}$  and  $B_{w,2}$  the associated Banach space. This means, we consider localization quantities of the form

$$\int_G w^{-1} |\mathcal{W}_g g|^2 d\lambda.$$

If  $w$  is such that

$$\eta(x) = \sup_{y \in G} \sqrt{w(x^{-1} \circ y)w(y)}$$

satisfies  $\eta \in L^2(G, d\lambda + d\rho)$ , we may estimate for  $r = w^{1/2} u$ ,  $s = w^{1/2} v$ ,  $u, v \in L^2(G, d\lambda + d\rho)$  that

$$\begin{aligned} |r * s(x)| &\leq \int w(y^{-1} \circ x) w(y) u(y^{-1} \circ x) v(y) d\lambda(y) \\ &\leq \eta(y) \|u\|_{L^2(G, d\rho)} \|v\|_{L^2(G, d\lambda)} \leq \eta(y) \|r\|_{w,2} \|s\|_{w,2} \end{aligned}$$

and thus we have  $B_{w,2} * B_{w,2} \subset B_{\eta,\infty}$ . Therefore the theorem applies and we have the existence of an optimally localized reconstruction wavelet.

We now prove the theorem. To start we analyze the mapping properties of  $\Phi$ . We denote by  $\Delta$  the domain of  $\Phi$ .

$$\Delta = \{s \in H : \Phi[s] < \infty\} \subset A.$$

**Lemma 26** *The functional  $\Phi$  is strongly  $H$  - lower semi-continuous on  $\Delta$ . More precisely for  $\Delta \ni u_n \rightarrow u \in H$  in  $H$ , we have*

$$\Phi[u] \leq \liminf_{n \rightarrow \infty} \Phi[u_n].$$

*Proof.* Since  $\mathcal{W}_{u_n} u_n \rightarrow \mathcal{W}_u u$  point-wise and thus by hypothesis of semi-continuity of the  $B$ -norm

$$\Phi[\lim_{n \rightarrow \infty} u_n] = \|\lim_{n \rightarrow \infty} \mathcal{W}_{u_n} u_n\|_B \leq \liminf_{n \rightarrow \infty} \|\mathcal{W}_{u_n} u_n\| = \liminf_{n \rightarrow \infty} \Phi[u_n]$$

□

We even have

**Lemma 27** *The functional  $\Phi$  is  $H$ -weakly lower semi-continuous on  $\Delta \setminus \{0\}$ . More precisely, for any  $H$ -weak convergent sequence  $\Delta \ni g_n \rightarrow g \in H$ ,  $g \neq 0$  with  $g_n \in \Delta$  we have  $\Phi[g] \leq \liminf \Phi[g_n]$ .*

*Proof.* Let

$$\gamma = \liminf_{n \rightarrow \infty} \Phi[g_n].$$

Clearly  $\gamma \geq 0$ . In the case that  $\gamma = \infty$  the lemma holds true and we may suppose  $0 \leq \gamma < \infty$ . We may find a subsequence which denote again by  $g_n$  with  $\Phi[g_n] < \infty$ , and  $\Phi[g_n] \rightarrow \gamma$ . By hypothesis there is an  $h \in D(T)$  with  $\Phi[h] < \infty$ . By the invariance of  $B$  it follows thanks to

$$(\mathcal{W}_{U(y)h} U(y)h)(x) = \mathcal{W}_h h(y^{-1} \circ x \circ y)$$

that the whole orbit of  $h$  has the same properties. Since the representation is irreducible and  $T$  has dense range, we may suppose that  $c_{g,h} = (g, Th)_H \neq 0$ . For  $s \in A$  we have by continuity of the convolution

$$\|\mathcal{W}_h s\|_K = \|\Pi * \mathcal{W}_s s\|_K \leq d \|\mathcal{W}_s s\|_B = d \Phi[s], \quad \Pi = c_{s,h}^{-1} \mathcal{W}_h h.$$

By weak convergence we have  $c_{g_n,h} = (g_n, Th)_H \rightarrow (g, Th)_H \neq 0$ . Therefore, since  $g_n \in A$  tanks to  $\Phi[g_n] < \infty$ , we may conclude by setting  $s = g_n$  in the formula above that  $\{\mathcal{W}_h g_n\}$  is a bounded set in  $K$  and hence it is bounded in  $H$  too. Since the representation of the wavelet transform is strongly continuous this family of functions is uniformly continuous on any compact subset of  $G$ . Because of the compact embedding property of  $K$  we can extract an  $H$  convergent subsequence  $g_{m(n)} \rightarrow g$ . Since  $\Phi$  is strongly lower continuous we have  $\Phi[g] \leq \liminf \Phi[g_n] = \gamma$ . □

Now the proof of the main theorem is easy.

*Proof.* Let  $\Phi[g_n] \rightarrow \gamma = \inf_{g \in \Sigma_h} \Phi[g]$ . Since as before  $\Phi[g_n] \leq b \|g_n\|_K$  we see that  $g_n$  is bounded in  $H$ . Thanks to the Banach-Alaoglou theorem we may extract an  $H$ -weakly convergent subsequence  $g_{m(n)} \rightarrow g$  weakly. Since  $\Phi$  is weakly lower semi-continuous we may conclude that  $\Phi[g] = \gamma$ . The set  $\Sigma_h$  is weakly closed and hence  $g \in \Sigma_h$ . □

We can even prove the following optimal localization result.

There is  $g$  with  $\|g\|_H = 1$ , such that for all  $u \in H$ , with  $\|u\|_H = 1$  we have

$$\Phi[g] \leq \Phi[u].$$

*Proof.* As before, we find a weakly convergent sequence  $g_n \rightarrow g$  weakly with  $\Phi[g] = \inf_{\|s\|=1} \Phi[s]$ . Now as in the proof of lemma 27 we see that there is a strongly convergent subsequence  $g_{m(n)} \rightarrow g$  and thus  $\|g\| = 1$ . □

### 5.2.3 Numerical Approximation of Localized States

Since there is no obvious strategy to derive the optimally localized vectors  $g$  explicitly, we aim to approximate  $\arg \inf_{\|g\|_H=1} \Phi[g]$  numerically. Here we limit ourselves to  $L^{2,w^{-1}}$  localization. Then, in accordance with Theorem 5.2.2, the optimization problem can be casted as follows

$$\Phi(g) = \|\mathcal{W}_g g\|_{L^{2,w^{-1}}}^2 + \alpha \|g\|_H^2.$$

In order to discretize the problem somehow, we may represent  $g$  by means of some frame  $\{\psi_\lambda\}_{\lambda \in \Lambda} \subset H$ , i.e.

$$g = \sum_{\lambda \in \Lambda} g_\lambda \psi_\lambda.$$

Consequently, the goal is to reconstruct a sequence  $\{g_\lambda\}_{\lambda \in \Lambda} = \mathbf{g} \in \ell_2$  for which  $\Phi(g) \leq \Phi(s)$ , for all  $s \in H$ .

Introducing for some  $x \in G$  the infinite matrix  $\mathbf{A}(x) = ((\psi_\lambda, U(x)\psi_\eta)_H)_{\lambda, \eta \in \Lambda}$ , the wavelet transform reads as  $\mathcal{W}_g g(x) = \langle \mathbf{g}, \mathbf{A}(x) \mathbf{g} \rangle_{\ell_2} =: F[\mathbf{g}](x)$ . Obviously,  $F[\mathbf{g}](e) = \|g\|_H^2$ , and thus we may write

$$\Phi(\mathbf{g}) = \|F[\mathbf{g}]\|_{L^{2,w^{-1}}}^2 + \alpha F[\mathbf{g}](e).$$

Since the optimization problem is no longer convex, we have to apply adequate strategies for nonlinear problems. We suggest to make use of a Tikhonov-based iteration method for nonlinear problems which was developed in [RT05b]. The technology to be applied here will always find a critical point of  $\Phi$ , and under additional assumptions on  $F$  and the solution one can assure that the critical point is a global minimizer.

The method borrowed from [RT05b] goes now as follows: Firstly, in order to obtain a problem which is hopefully easier to solve, we replace  $\Phi$  by

$$\Phi^s(\mathbf{g}; \mathbf{a}) := \Phi(\mathbf{g}) + C \|\mathbf{g} - \mathbf{a}\|_{\ell_2}^2 - \|F[\mathbf{g}] - F[\mathbf{a}]\|_{L^{2,w^{-1}}}^2, \quad (5.2.2)$$

where  $\mathbf{a}$  is some auxiliary element in  $\ell_2$ . So far its not clear whether  $\Phi^s$  is positive or even bounded from below. Following the lines in [RT05b], i.e. choosing for  $\alpha > 0$  a ball around the origin  $K_r$  and  $C$  adequately large (in dependence on  $F$  and  $\Phi(\mathbf{a})$ ), one can assure for all  $\mathbf{g} \in K_r$ ,  $\Phi(\mathbf{g}) \leq \Phi^s(\mathbf{g}; \mathbf{a})$ .

The iteration process is now obtained by picking some initial  $\mathbf{g}_0 = \mathbf{a}$  and therewith some proper  $C > 0$  and by deriving a sequence  $\{\mathbf{g}_k\}_{k \in \mathbf{N}}$  via

$$\mathbf{g}_{k+1} = \arg \min_{\mathbf{g}} \Phi^s(\mathbf{g}; \mathbf{g}_k).$$

From this iteration we expect convergence at least towards a critical point of  $\Phi$ . First, we have to make sure that the sequence of functionals is properly defined:

**Lemma 28** *Let  $\mathbf{a}$  be given and  $K_r$ ,  $C$  be defined as in [RT05b]. Then for all  $k \in \mathbf{N}$ ,  $\Phi^s(\mathbf{g}; \mathbf{g}_k)$  are bounded from below, and, moreover, for the minimizers  $\mathbf{g}_{k+1}$  holds  $\mathbf{g}_{k+1} \in K_r$ .*

Let now  $\mathbf{A}$  be the shorthand for  $\mathbf{A}(e)$ . A simple calculation shows:

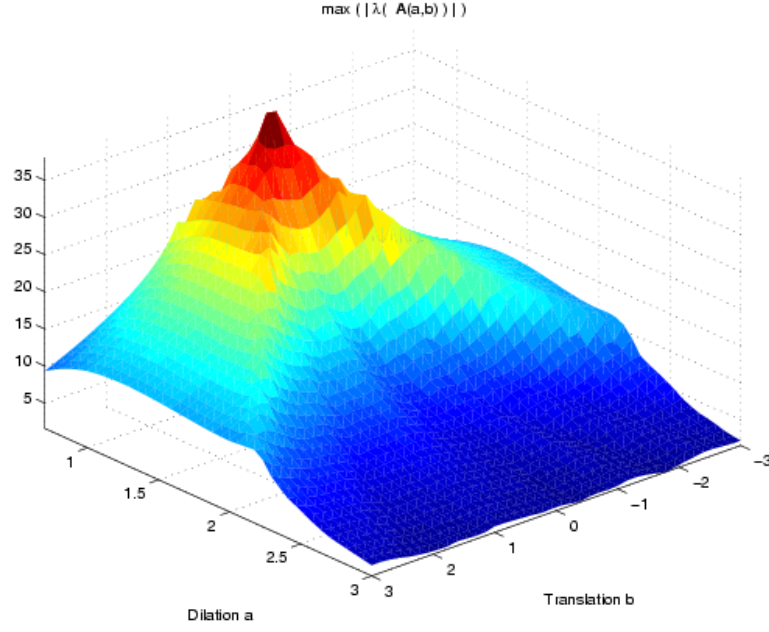


Figure 5.6: Maximal Eigenvalues of the infinite matrices  $A(a, b)$  for all the  $(a, b) \in G$  used in the frame representation.

**Lemma 29** *The necessary condition for a minimum of (5.2.2) reads as*

$$\mathbf{g} = \frac{1}{C} \{ \mathbf{g}_k - \alpha \mathbf{A} \mathbf{g} - F'[\mathbf{g}]^* (F[\mathbf{g}_k] w^{-1}) \} \quad (5.2.3)$$

The hope is that the right hand side of (5.2.3) defines a contraction. A straightforward computation shows,

$$\|\mathbf{g} - \mathbf{g}'\|_{\ell_2} \leq \frac{1}{C} \{ \alpha \|\mathbf{A}\| + 2 \|\mathbf{A}(\cdot)\|_{L^{2,w^{-1}}} \|F[\mathbf{g}_k]\|_{L^{2,w^{-1}}} \} \|\mathbf{g} - \mathbf{g}'\|_{\ell_2} .$$

To bound this quantity requires the Lipschitz-continuity of  $F'[\mathbf{g}]$ , or in other words, the finiteness of  $\|\mathbf{A}(\cdot)\|_{L^{2,w^{-1}}}$  which is difficult to prove, but can be verified numerically: we may consider the spectral radius  $\rho(\mathbf{A}(a, b))$  (for a particular frame, see below) as a function of  $(a, b) \in G$ . Figure 5.6 shows a sufficient decay of  $\rho(\mathbf{A}(a, b))$  and assures therewith that, for  $C$  large enough, the convergence of the fixed point iteration (5.2.3) towards a unique minimizer  $\mathbf{g}_{k+1}$  of  $\Phi^s(\mathbf{g}; \mathbf{g}_k)$  can be achieved. Moreover, we have with the help of [RT05b] that the sequence  $\{\mathbf{g}_k\}$  converges at least towards a critical point of  $\Phi$ . If we could impose more smoothness on  $F$  and on the solution  $\mathbf{g}$  to be reconstructed, we could also achieve uniqueness.

Next, we have to ensure that  $\|\mathbf{g}^{n+1}\|_H^2 = 1$  (the index  $n$  stands now for the fixed point iteration) holds true through the whole fixed point iteration process, i.e. we have to determine  $\alpha$  in each iteration step:

$$2\alpha \Re(\mathbf{A} \mathbf{g}^n, \mathbf{A}(\mathbf{g}_k - F'[\mathbf{g}^n]^* (F[\mathbf{g}_k] w^{-1})))_{\ell_2} = \frac{1}{C^2} \{ \alpha^2 F[\mathbf{A} \mathbf{g}^n](e) - 2\alpha \Re(\mathbf{A} \mathbf{g}^n, \mathbf{A}(\mathbf{g}_k - F'[\mathbf{g}^n]^* (F[\mathbf{g}_k] w^{-1})))_{\ell_2} + F[\mathbf{g}_k - F'[\mathbf{g}^n]^* (F[\mathbf{g}_k] w^{-1})](e) \} ,$$

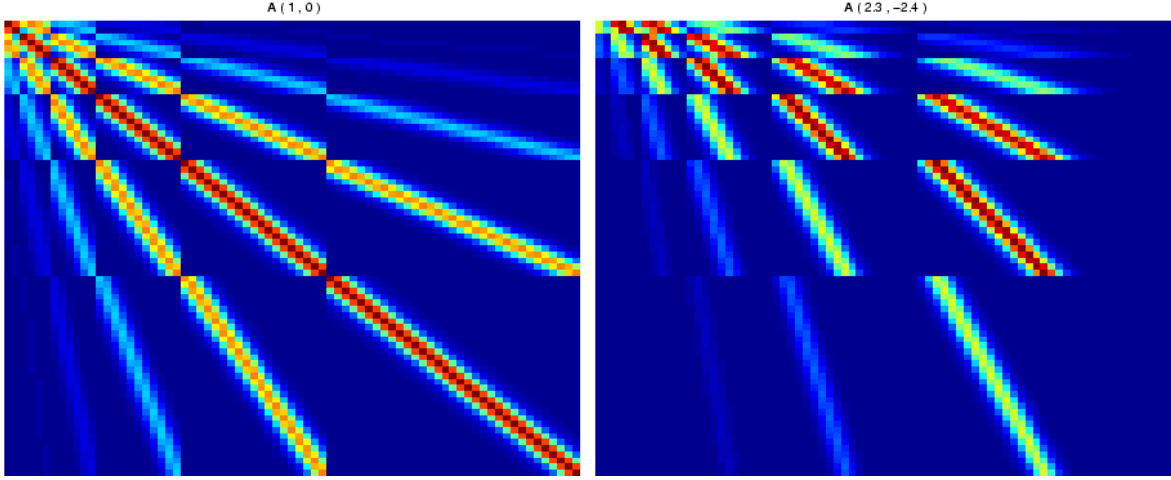


Figure 5.7: Structure of  $A(a, b)$  for two particular cases; left:  $(a, b) = (1, 0)$  and right  $(a, b) = (2.3, -2.4)$ .

i.e. finding  $\alpha = \alpha^{n+1}$  amounts to finding the roots of a real parabola. With the shorthand  $M = \Re(\mathbf{A}\mathbf{g}^n, \mathbf{A}(\mathbf{g}_k - F'[\mathbf{g}^n]^*(F[\mathbf{g}_k]w^{-1})))_{\ell_2}$ , we obtain

$$\alpha^{n+1} = \frac{M \pm (M^2 - F[\mathbf{A}\mathbf{g}^n](e)\{F[\mathbf{g}_k - F'[\mathbf{g}^n]^*(F[\mathbf{g}_k]w^{-1})](e) - C^2\})^{1/2}}{F[\mathbf{A}\mathbf{g}^n](e)} . \quad (5.2.4)$$

Now we can summarize an algorithm for computing a critical sequence  $\mathbf{g}$  for the minimization problem  $\inf_{\|\mathbf{g}\|_H=1} \Phi[\mathbf{g}]$ :

- *pick some initial  $\mathbf{g}_0$  (not too far off the expected solution) and some  $C > 0$  (large enough)*
- *compute  $\mathbf{g}_{k+1} = \arg \min_{\mathbf{g}} \Phi^s(\mathbf{g}; \mathbf{g}_k)$  via fixed point iteration (5.2.3):*
  - *compute  $\alpha^{n+1} = \max\{\alpha_1^{n+1}, \alpha_2^{n+1}\}$  via (5.2.4)*
  - *compute  $\mathbf{g}^{n+1}$  via (5.2.3)*
  - $\mathbf{g}_{k+1} = \lim_{n \rightarrow \infty} \mathbf{g}^{n+1}$

In what follows we aim to illustrate the computation of an optimally localized wavelet. For sake of simply computing the operators  $\mathbf{A}(x)$ , we have chosen a (finite dimensional) Cauchy wavelet frame  $\{\psi_\lambda\} \subset L_2(\mathbf{R})$  of order  $N$  (here  $N=3$ ). Thus,  $\mathbf{A}(x)$  can be derived for each  $x \in G$  explicitly, see Figure 5.7. The symmetric weight function in our particular case is given by

$$w^{-1}(x) = w^{-1}(a, b) = (|a| + |a|^{-1})^4 \cdot (1 + |b|(1 + |a|)^{-1})^4 ,$$

and the resulting iteration process to reconstruct at least a critical  $\mathbf{g}$  is illustrated in Figure 5.8, and the final approximation with the time representation in Figure 5.9.



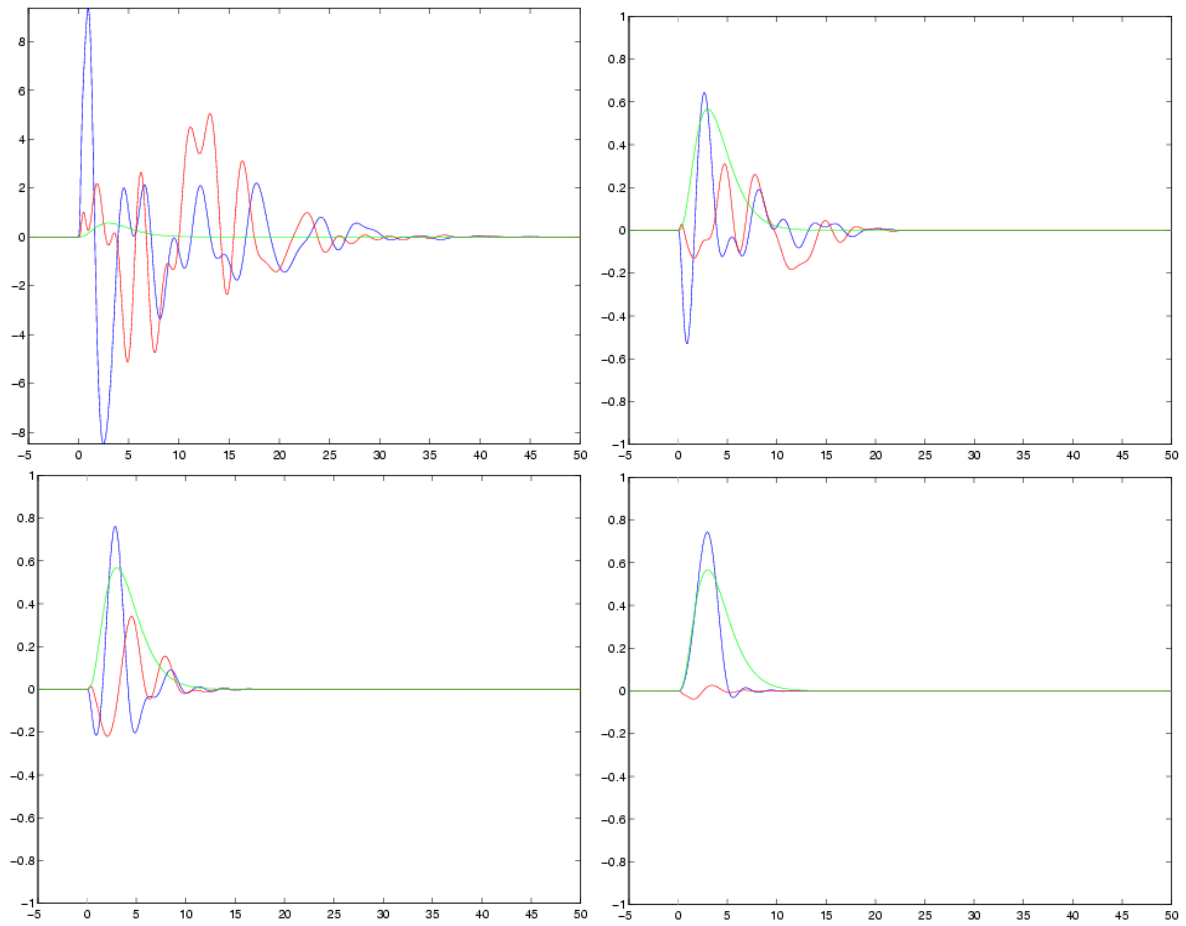


Figure 5.8: From top left to up right: Fourier representations of initial  $g_0$  (not normalized),  $g_4$ ,  $g_{10}$ , and  $g_{30}$  (blue/red - real and imaginary part; green - Cauchy wavelet).

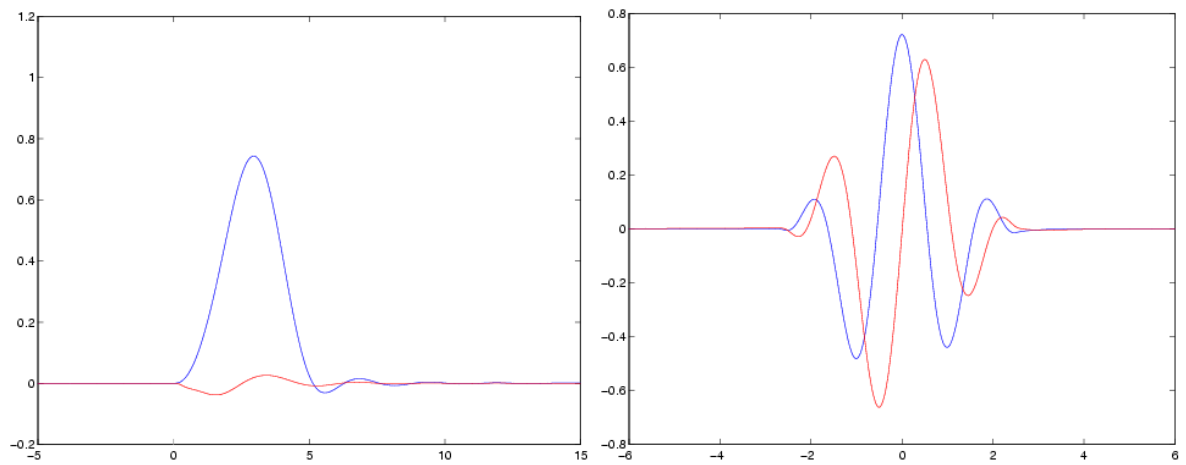


Figure 5.9: Left: Fourier representation of the approximated optimally localized coherent state; right: associated time representation.

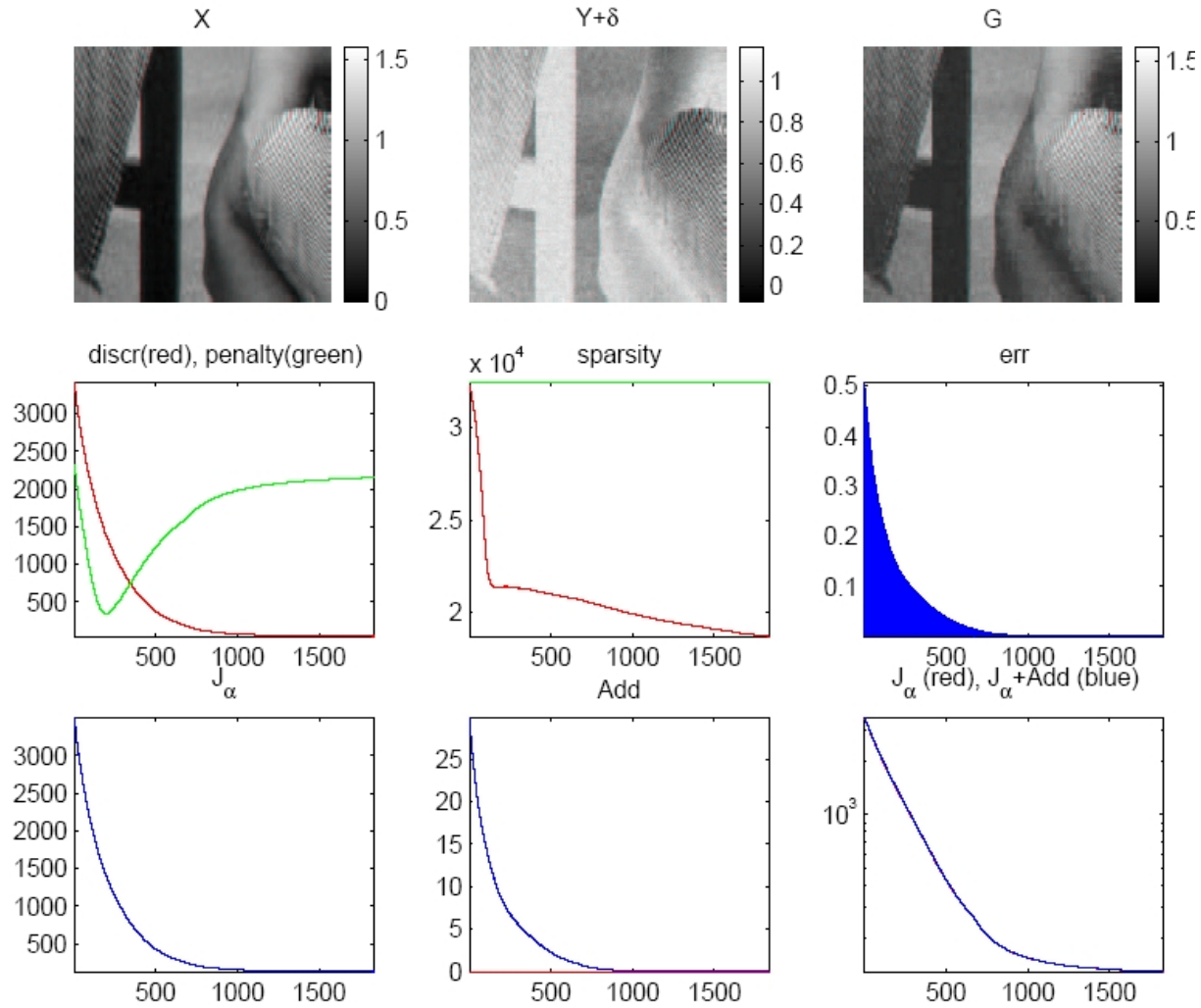


Figure 5.10: Thresholding Landweber fixed point iteration for the pixel basis and orthogonal Haar wavelet basis  $\mathbf{L} = \mathbf{W}$  and sparsity parameter  $\alpha = 0.02$ . From top left to up right: original image  $x$ ;  $T(x) + \delta = y^\delta$ ; final reconstruction of the solution; values of  $\|y^\delta - T(F^* \mathbf{g})\|_{L_2(\Omega)}^2$  (red) and  $|\mathbf{W} \mathbf{g}|_{\ell_1}$  (green) during the whole iteration process; sparsity history (red, green indicates the reference to original total number of coefficients); error plot;  $J_\alpha$ ; Gaussian surrogate term;  $J_\alpha$  (red) and  $J_\alpha^s = J_\alpha + \text{'Gaussian surrogate term'}$  (blue).

### 5.3 Thresholding Landweber Fixed Point Iteration – An Illustration

In this section, we apply the iterative machinery for solving nonlinear problems in a multi frame setting. For illustration purposes we focus on a sequence of synthetic nonlinear problems in the field of signal and image processing.

The first example is devoted to nonlinear image deformation. As the synthetic nonlinear

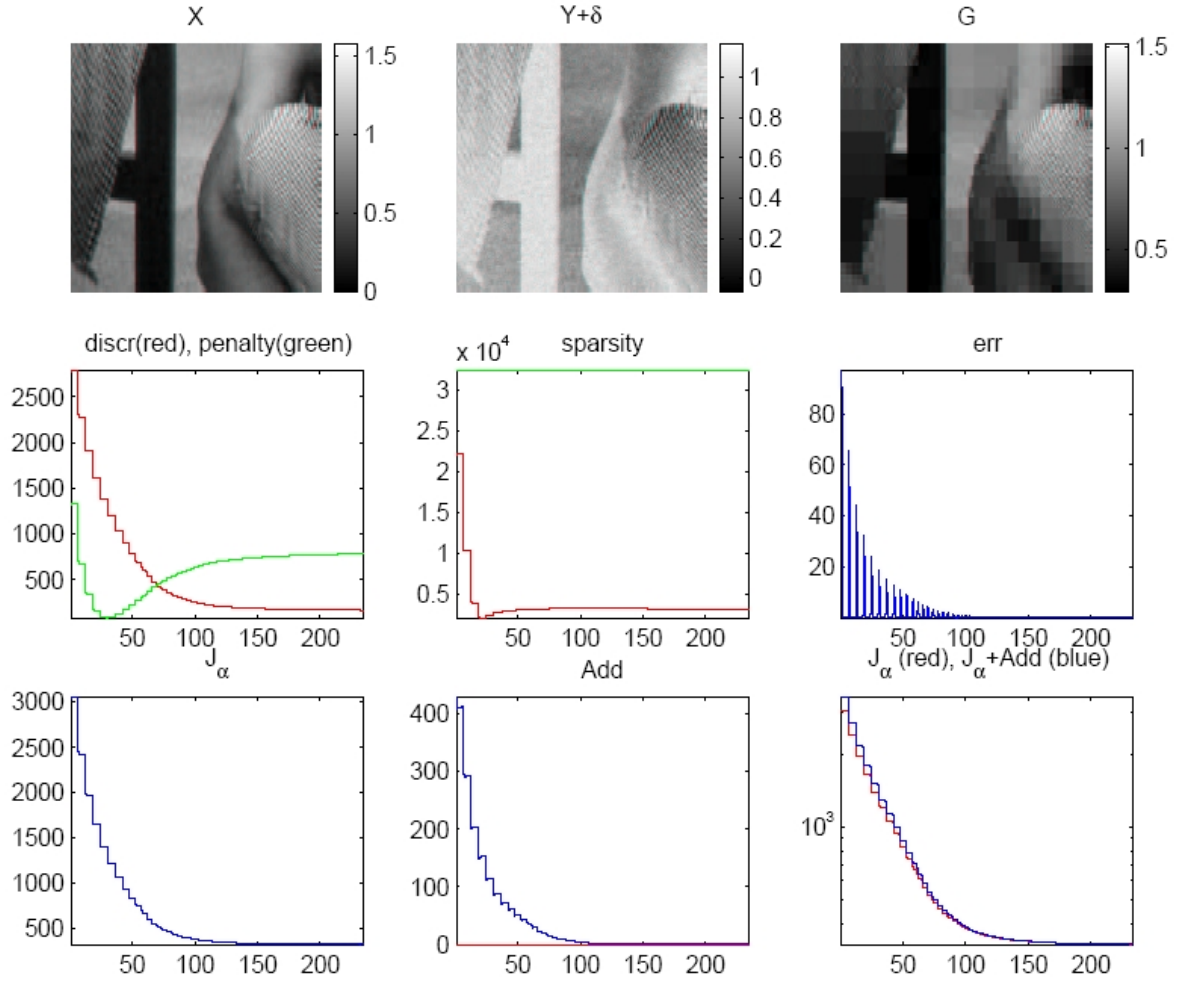


Figure 5.11: Thresholding Landweber fixed point iteration for the pixel basis and orthogonal Haar wavelet basis  $\mathbf{L} = W$  and sparsity parameter  $\alpha = 0.1$ . From top left to up right: original image  $x$ ;  $T(x) + \delta = y^\delta$ ; final reconstruction of the solution; values of  $\|y^\delta - T(F^* \mathbf{g})\|_{L_2(\Omega)}^2$  (red) and  $|W\mathbf{g}|_{\ell_1}$  (green) during the whole iteration process; sparsity history (red, green indicates the reference to original total number of coefficients); error plot;  $J_\alpha$ ; Gaussian surrogate term;  $J_\alpha$  (red) and  $J_\alpha^s = J_\alpha + \text{'Gaussian surrogate term'}$  (blue).

operator we consider

$$T(x) = \cos(x) .$$

Assuming our image is given by some  $x \in L_2(\Omega)$ , where  $\Omega = [0, 1]^2$ , then  $T$  is applied to each value  $x(k, l)$ , for all  $(k, l) \in \Omega$ ,

$$T(x(k, l)) = \cos(x(k, l)).$$

As the frame under consideration we chose the pixel basis with frame operator  $F$  and  $x = F^* \mathbf{g}$  for some  $\mathbf{g} \in \ell_2$ . Moreover, we aim to reconstruct an image while requiring

sparsity. Sparsity can be achieved when setting  $p = 1$ . However, we still know that sparsity cannot be well achieved when dealing with a pixel frame. Hence, it would be more feasible to switch to a wavelet frame (basis) when penalizing the approximation, i.e. we set  $\mathbf{L} = W$  and  $W$  denoting the orthogonal wavelet transform. Consequently, we may cast the problem as follows,

$$J_\alpha(\mathbf{g}) = \|\mathbf{y}^\delta - T(F^*\mathbf{g})\|_{L_2(\Omega)}^2 + 2\alpha|W\mathbf{g}|_{\ell_1} .$$

In the notation of Section 4.2, the Landweber iteration is then based on solving the following fixed point equation in each step,

$$\mathbf{g}_{k+1} = \frac{\alpha}{C} W^*(I - P_C) \left( \frac{C}{\alpha} WM(\mathbf{g}_{k+1}, \mathbf{g}_k) \right) = \mathbf{S}_{\alpha, W, C} (M(\mathbf{g}_{k+1}, \mathbf{g}_k)) ,$$

i.e. for each Landweber iteration we have to perform a fixed point iteration with a generalized shrinkage projection applied in each step

$$\mathbf{g}_{k+1, l+1} = \mathbf{S}_{\alpha, W, C} (M(\mathbf{g}_{k+1, l}, \mathbf{g}_k)) .$$

We finally need to derive the generalized shrinkage operator  $\mathbf{S}_{\alpha, \mathbf{L}, C}$ . Since  $p = 1$ ,

$$\Psi(W\mathbf{g}) = |W\mathbf{g}|_{\ell_1} = \sum_{\lambda \in \Lambda} |(W\mathbf{g})_\lambda| ,$$

the related convex set is then nothing else than

$$C = \{W\mathbf{g} \in \ell_2 : \sup_{\lambda \in \Lambda} |(W\mathbf{g})_\lambda| \leq 1\} .$$

This yields the componentwise acting projection  $P_C(W\mathbf{g}) = \{P_C((W\mathbf{g})_\lambda)\}_{\lambda \in \Lambda}$  with

$$P_C((W\mathbf{g})_\lambda) = \begin{cases} (W\mathbf{g})_\lambda & \text{if } |(W\mathbf{g})_\lambda| \leq 1 \\ \text{sgn}(W\mathbf{g})_\lambda & \text{if } |(W\mathbf{g})_\lambda| > 1 \end{cases} ,$$

where  $\text{sgn}(0) \in [-1, 1]$  and consequently,

$$(I - P_C)((W\mathbf{g})_\lambda) = \begin{cases} 0 & \text{if } |(W\mathbf{g})_\lambda| \leq 1 \\ \text{sgn}(W\mathbf{g})_\lambda (|(W\mathbf{g})_\lambda| - 1) & \text{if } |(W\mathbf{g})_\lambda| > 1 \end{cases} .$$

This is the well-known soft shrinkage operation with threshold 1, which we denote here by  $S_1$ . Thus,

$$\mathbf{g}_{k+1, l+1} = S_{\alpha/C} (M(\mathbf{g}_{k+1, l}, \mathbf{g}_k)) .$$

The numerical results for two different parameters  $\alpha$  are shown in Figures 5.10 and 5.11. In Figure 5.10 we have chosen  $\alpha = 0.02$ , in Figure 5.11,  $\alpha = 0.1$ . We may clearly observe that we achieve much better sparsity in the second case whereas the approximation quality is much higher in the first example, and that the number of iterations becomes less when  $\alpha$  increases.

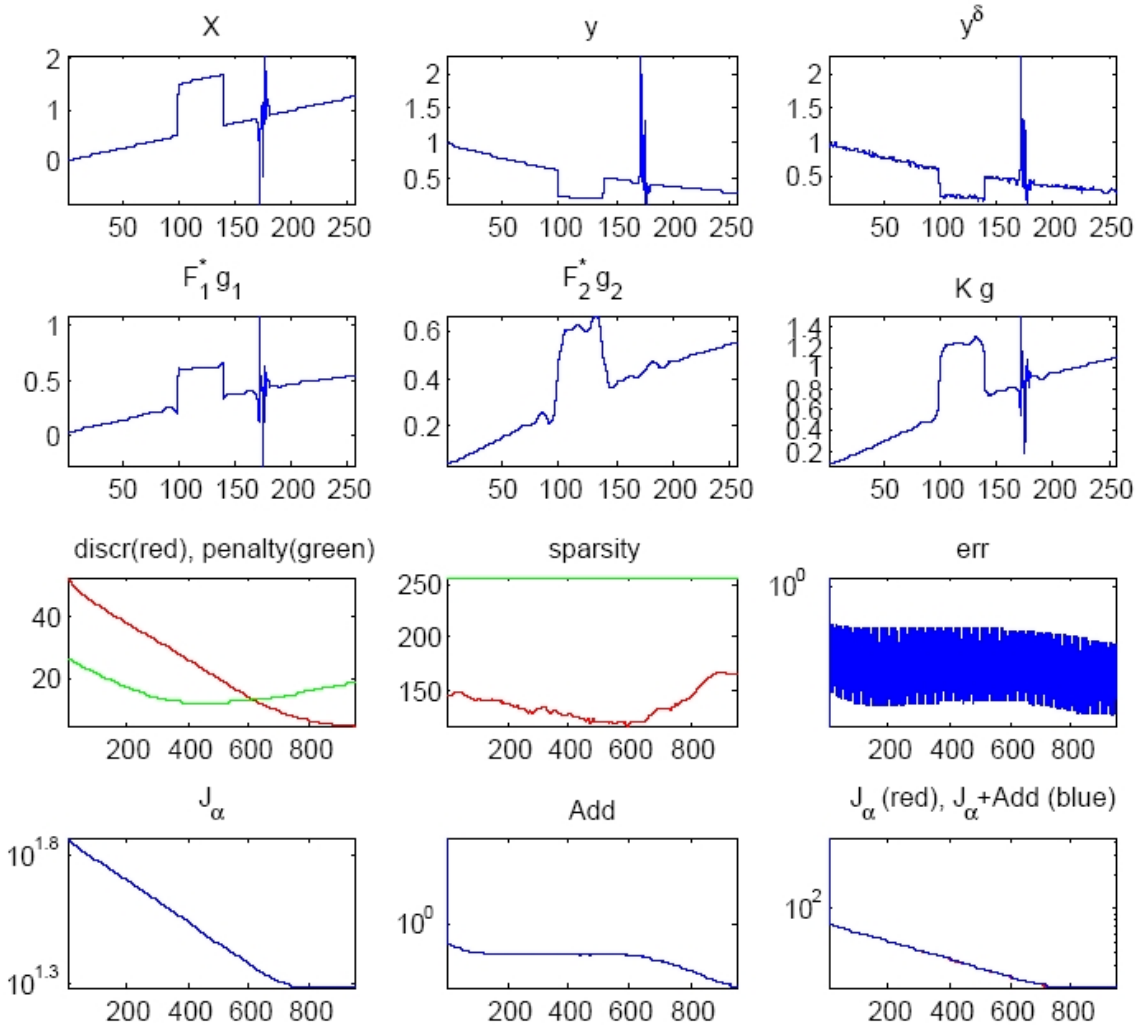


Figure 5.12: Thresholding Landweber fixed point iteration for a wavelet based dictionary ( $F_1 \sim$  Haar system,  $F_2 \sim$  Daubechies wavelet basis of order ten) and sparsity parameters  $\alpha = (0.2, 0.5)$ . From top left to up right: original data  $x$ ;  $T(x) = y$ ;  $T(x) + \delta = y^\delta$ ; final Haar reconstruction; final Db10 reconstruction; final overall reconstruction; values of  $\|y^\delta - T(F^* \mathbf{g})\|_{L_2(\Omega)}^2$  (red) and  $|\mathbf{g}_1|_{\ell_1} + |\mathbf{g}_2|_{\ell_1}$  (green) during the whole iteration process; sparsity history (red, green indicates the reference to original total number of coefficients); error plot;  $J_\alpha$ ; Gaussian surrogate term;  $J_\alpha$  (red) and  $J_\alpha^s = J_\alpha + \text{'Gaussian surrogate term'}$  (blue).

In the second illustration, we really compute a reconstruction when dealing with multi frames. For computational reasons we consider a one dimensional synthetic data set, see top left diagram in Figure 5.12. The nonlinearity comes into play by setting

$$y = T(x) = e^{-x}.$$

Our frame dictionary consists now of two different bases: Daubechies wavelet bases of

order one (Haar wavelet basis) and ten. We denote the corresponding frame operators by  $F_1$  and  $F_2$ , then

$$x = K\mathbf{g} = K(\mathbf{g}_1, \mathbf{g}_2) = F_1^* \mathbf{g}_1 + F_2^* \mathbf{g}_2 \quad .$$

Moreover, we again aim to reconstruct a sparse solution of the inverse deformation problem. Since we still deal with a wavelet based dictionary it is customary to set  $\mathbf{L}_1 = \mathbf{L}_2 = I$ . The variational problem to be minimized reads then as

$$J_\alpha(\mathbf{g}) = J_\alpha(\mathbf{g}_1, \mathbf{g}_2) = \|y^\delta - T(K(\mathbf{g}_1, \mathbf{g}_2))\|_{L_2(\Omega)}^2 + 2\alpha_1 |\mathbf{g}_1|_{\ell_1} + 2\alpha_2 |\mathbf{g}_2|_{\ell_1} \quad .$$

Hence, the resulting system of fixed point equations (to be solved in the same manner as before) is given by

$$\begin{aligned} (\mathbf{g}_1)_{k+1} &= S_{\alpha_1/C} (M_1(\mathbf{g}_{k+1}, \mathbf{g}_k)) \\ (\mathbf{g}_2)_{k+1} &= S_{\alpha_2/C} (M_2(\mathbf{g}_{k+1}, \mathbf{g}_k)) \quad . \end{aligned}$$

The results are visualized in Figure 5.12. The main observation is that we may indeed reconstruct with the proposed scheme an approximation of  $x$ . Moreover, we see that the different components of  $x$  are at most complementary covered by the two different frames: the Haar system essentially grabs the non-smooth part whereas the Db10 family describes smoother components of  $x$ . Of course, we must admit that the information is not completely split, i.e. there is still some redundant information in  $\mathbf{g}_1$  and  $\mathbf{g}_2$ . However, the reconstructed approximation of  $x$  requires even by using two bases much less coefficients (approx. 160 coefficients) than the original data set (256 coefficients).

## 5.4 Acceleration of Support Vector Machines

As in Section 3.3, we briefly consider Support Vector Machines again. But contrary to the way of applying the iterative strategy for linear problems which was essentially used to approximate the set vectors itself, we suggest now to reduce  $N_x$  and to approximate the  $\mathbf{x}_i$ 's in

$$\Psi_1(\alpha, \mathbf{x}) = \sum_{i=1}^{N_x} \alpha_i \Phi(\mathbf{x}_i)$$

simultaneously, i.e. we aim to minimize a quantity of the form

$$\|\Psi_1(\alpha, \mathbf{x}) - \Psi(\beta, \mathbf{z})\|^2 + 2\gamma_0 |\beta|_1 + 2 \sum_{i=1}^{N_x} \gamma_i |\mathbf{z}_i|_1 \rightarrow \min_{\beta, \mathbf{z}} \quad ,$$

where

$$\Psi(\beta, \mathbf{z}) = \sum_{i=1}^{N_x} \beta_i \Phi(A\mathbf{x}_i + W^{-1}\mathbf{z}_i) \quad ,$$

and  $A\mathbf{x}_i$  denotes the approximation sequences at the coarsest wavelet decomposition level, and  $W^{-1}\mathbf{z}_i$  the orthogonal complement (inverse of all the wavelet details). Following the

lines in Section 4.2, a straightforward computation results in the following system of fixed point iterations for computing the iterates

$$\begin{aligned}
& \beta_k \text{ and } \mathbf{z}_k = ((\mathbf{z}_1)_k, \dots, (\mathbf{z}_{N_x})_k) : \\
& \beta_{k+1} = \mathbf{S}_{\gamma_0/C} (\beta_k + D_\beta \Psi(\beta_{k+1}, \mathbf{z}_{k+1})^* (\Psi_1(\alpha, \mathbf{x}) - \Psi(\beta_k, \mathbf{z}_k)) / C) \\
& (\mathbf{z}_1)_{k+1} = \mathbf{S}_{\gamma_1/C} ((\mathbf{z}_1)_k + W D_{\mathbf{z}_1} \Psi(\beta_{k+1}, \mathbf{z}_{k+1})^* (\Psi_1(\alpha, \mathbf{x}) - \Psi(\beta_k, \mathbf{z}_k)) / C) \\
& \vdots \\
& (\mathbf{z}_{N_x})_{k+1} = \mathbf{S}_{\gamma_{N_x}/C} ((\mathbf{z}_{N_x})_k + W D_{\mathbf{z}_{N_x}} \Psi(\beta_{k+1}, \mathbf{z}_{k+1})^* (\Psi_1(\alpha, \mathbf{x}) - \Psi(\beta_k, \mathbf{z}_k)) / C) .
\end{aligned}$$

In accordance with the convergence results in Section 3.3 we may assure by a proper choice of  $C$  that we approach in any case a critical point, which is not the case for the reduction procedure presented in Section 3.3. However, a numerical verification is not done yet.





## Chapter 6

# Perspectives on Adaptive Frame Strategies

This chapter is devoted to the discussion of adaptive concepts for linear and nonlinear problems. This is especially motivated by the high computational complexity when solving operator equations at least with multi frame dictionaries and non-diagonal operators. In the very recent literature, e.g. [Ste03], a frame-based method for solving linear operator equations is presented. This concept is sort of extension of known wavelet based strategies, []: using a Riesz basis of wavelet type for the underlying Hilbert space, the operator equation is transformed into an equivalent matrix vector system. This system is solved iteratively, where the application of the infinite stiffness matrix is replaced by an adaptive approximation. Assuming that the stiffness matrix is sufficiently compressible, i.e. it can be sufficiently well approximated by sparse matrices, it was proved that adaptive methods have optimal computational complexity in the sense that it converges with the same rate as the best  $N$ -term approximation for the solution assuming it would be explicitly available. The condition concerning compressibility requires that the wavelets have sufficiently many vanishing moments, and that they sufficiently smooth. Except on tensor domains, wavelets that satisfy this requirement are difficult to construct. At least, when dealing with an over-complete dictionary we are beyond bases.

To this end, in many cases a frame based treatment seems to be much better suited for solving operator equations. In [Ste03], the usage of one individual frame is suggested. With this frame the operator equation is transformed into a matrix vector system, after which this system is solved iteratively by an adaptive method.

In what follows we briefly review the basic idea of Richardson iterations presented in [Ste03]. Next, we briefly sketch the relation to Landweber type iterations we have constructed in Chapter 2 and propose an adaptive iteration (for which do not give convergence results). Finally, we end this chapter with a discussion on the nonlinear case.

## 6.1 Brief Review on an Adaptive Frame Method

We review the ideas presented in [Ste03]. Given some linear operator  $L$  which is boundedly invertible from some Hilbert space  $\mathcal{H}$  into itself, we consider the problem of finding  $v$  such that

$$Lv = z .$$

The basic concept requires some additional facts on frames. Assume we are given some frame with frame operator  $F$ , such that

$$F : \mathcal{H} \rightarrow \ell_2, \quad f \mapsto \{f_\lambda = \langle f, \phi_\lambda \rangle\}_{\lambda \in \Lambda} \quad (6.1.1)$$

$$F^* : \ell_2 \rightarrow \mathcal{H}, \quad c \mapsto \sum_{\lambda \in \Lambda} c_\lambda \phi_\lambda \quad (6.1.2)$$

are bounded with norm less or equal  $\sqrt{B}$ . The composition  $F^*F$  is a positive and self-adjoint (i.e. boundedly invertible) operator. The canonical dual frame has frame bounds  $B^{-1}$ ,  $A^{-1}$ , and corresponding analysis and synthesis operators

$$\tilde{F} = F(F^*F)^{-1}, \quad \tilde{F}^* = (F^*F)^{-1}F^* . \quad (6.1.3)$$

In particular, one has the following orthogonal decomposition of  $\ell_2$

$$\ell_2 = \mathcal{R}(F) \oplus \mathcal{N}(F^*) ,$$

and

$$\mathbf{Q} := F(F^*F)^{-1}F^* : \ell_2 \rightarrow \mathcal{R}(F) ,$$

is the orthogonal projection onto  $\mathcal{R}(F)$ .

Let us now transform the operator equation in a matrix vector system. At first, write  $v = F^*\mathbf{g}$  for some  $\mathbf{g} \in \ell_2$ , where  $\mathbf{g}$  satisfies

$$\mathbf{M}\mathbf{g} = \mathbf{z} , \quad (6.1.4)$$

with

$$\mathbf{M} := FLF^* \quad \text{and} \quad \mathbf{z} := Fz .$$

From

$$\left. \begin{aligned} \tilde{F}L^{-1}\tilde{F}^*FLF^* &= \tilde{F}F^* \\ FLF^*\tilde{F}L^{-1}\tilde{F}^* &= F\tilde{F}^* \end{aligned} \right\} = \mathbf{Q} = \mathbf{I} \text{ on } \mathcal{R}(F) ,$$

we conclude that  $\mathbf{M}|_{\mathcal{R}(F)} : \mathcal{R}(F) \rightarrow \mathcal{R}(F)$  is boundedly invertible, with  $\|\mathbf{M}\| \leq B\|L\|$  and  $\|\mathbf{M}|_{\mathcal{R}(F)}^{-1}\| \leq A^{-1}\|L^{-1}\|$ , where  $\mathcal{N}(\mathbf{M}) = \mathcal{N}(F^*)$ .

Now we can write the iterative scheme to solve in infinite dimensional system  $\mathbf{M}\mathbf{g} = \mathbf{z}$ . In case  $L$  is symmetric and positive,  $\mathbf{M} = \mathbf{M}^* \geq 0$ . With  $\lambda_{\max} := \lambda_{\max}(\mathbf{M}) = \|\mathbf{M}\|$  and  $\lambda_{\min}^+ := \lambda_{\min}(\mathbf{M}|_{\mathcal{R}(F)}) = \|\mathbf{M}|_{\mathcal{R}(F)}^{-1}\|^{-1}$ , for  $0 < \eta < 2/\lambda_{\max}$ , we consider the damped Richardson iteration

$$\mathbf{g}_{k+1} = \mathbf{g}_k - \eta(\mathbf{M}\mathbf{g}_k - \mathbf{z}) . \quad (6.1.5)$$

We infer that

$$\|\mathbf{Q}(\mathbf{g} - \mathbf{g}_{k+1})\| \leq \rho \|\mathbf{Q}(\mathbf{g} - \mathbf{g}_k)\| ,$$

where  $\rho := \|(\mathbf{I} - \eta \mathbf{M})|_{\mathcal{R}(F)}\| = \max\{\eta \lambda_{\max} - 1, 1 - \eta \lambda_{\min}^+\}$ .

In case of  $L$  being non-symmetric or indefinite can be treated by considering the normal equation

$$\mathbf{M}^* \mathbf{M} \mathbf{g} = \mathbf{M}^* \mathbf{z} . \quad (6.1.6)$$

The bulk of [Ste03] is now the study of an inexact version of (6.1.5) in which the application of the infinite matrix  $\mathbf{M}$  is approximated (we neglect here the problem that errors made in  $\mathcal{R}(F^*)$  are not reduced in subsequent iterations). Obviously, since in actual computations neither we can handle an infinite vector  $\mathbf{z}$ , nor we can apply the infinite matrix  $\mathbf{M}$ , iteration (6.1.5) is not a practical algorithm. To this end, assume we have the following procedures at our disposal (for a detailed description we refer the reader to [Ste03]):

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

$$\|\mathbf{g} - \mathbf{g}_\epsilon\|_{\ell_2} \leq \epsilon \quad (6.1.7)$$

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

$$\|\mathbf{N}\mathbf{v} - \mathbf{w}_\epsilon\|_{\ell_2} \leq \epsilon \quad (6.1.8)$$

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

$$\|\mathbf{v} - \mathbf{v}_\epsilon\|_{\ell_2} \leq \epsilon . \quad (6.1.9)$$

Based on these routines, we may consider the following inexact version of the damped Richardson iteration:

**SOLVE** ( $\epsilon, \mathbf{M}, \mathbf{z}$ )  $\rightarrow \mathbf{g}_\epsilon$ :

Let  $\theta < 1/3$  and  $K \in \mathbb{N}$  be fixed such that  $3\rho^K < \theta$

$i := 0, \mathbf{g}_0 := 0, \epsilon_0 := \|\mathbf{M}|_{\mathcal{R}(F)}^{-1}\| \|\mathbf{z}\|$

while  $\epsilon_i > \epsilon$  do

$i := i + 1$

$\epsilon_i := 3\rho^K \epsilon_{i-1} / \theta$

$\mathbf{z}_i := \mathbf{APPLY}(\theta \epsilon_i / (6\eta K), \mathbf{g})$

$\mathbf{f}_{i,0} := \mathbf{g}_{i-1}$

  for  $j = 1, \dots, K$  do

$$\mathbf{f}_{i,j} := \mathbf{f}_{i,j-1} - \eta(\mathbf{APPLY}(\theta \epsilon_i / (6\eta K), \mathbf{M}, \mathbf{f}_{i,j-1}) - \mathbf{z}_i)$$

  end

$\mathbf{g}_i := \mathbf{COARSE}((1 - \theta)\epsilon, \mathbf{f}_{i,K})$

end

$$\mathbf{g}_\epsilon := \mathbf{g}_i$$

The following theorem gives the approximation result (for extensions of this theorem and a proof we refer the reader again to [Ste03]).

**Theorem 6.1.1** *Let  $\mathbf{g} \in \ell_2$  be some solution of  $\mathbf{M}\mathbf{g} = \mathbf{z}$ . Then the vectors  $\mathbf{g}_i$  and  $\mathbf{f}_{i,K}$  produced in  $\mathbf{SOLVE}(\epsilon, \mathbf{M}, \mathbf{z})$  satisfy*

$$\|\mathbf{Q}(\mathbf{g} - \mathbf{g}_i)\| \leq \epsilon_i ,$$

and thus,

$$\|\mathbf{Q}(\mathbf{g} - \mathbf{g}_\epsilon)\| \leq \epsilon .$$

## 6.2 Adaptivity for Solving Variational Problems

We focus on linear problems. In view of quadratic Tikhonov functionals we may of course consider a slightly modified operator

$$\mathbf{M} := FA^*AF^* + \alpha I$$

and may directly apply  $\mathbf{SOLVE}(\epsilon, \mathbf{M}, \mathbf{z})$  to this operator to obtain an adaptive way of deriving the solution.

However, in contrast to apply the inexact Richardson iteration to the normal equation (what quite natural in the quadratic situation), we may also investigate the relation between the Landweber iteration developed in Section 2.2. We restrict the analysis to the single frame approach (but we note that achieving optimal complexity is especially for the multi frame case of special interest). The hope is also to deduce an adaptive scheme for the linear operator setting with some non-quadratic penalty. To this end, let us consider the variational problem from Section 2.2

$$\Phi(\mathbf{g}) := \|f - K_A\mathbf{g}\|^2 + \alpha|\mathbf{g}|_p^p .$$

For  $p = 2$ , setting  $\mathbf{M} = K_A^*K_A$  and assuming  $\|A\| < 1$ , the resulting iteration reads as

$$\mathbf{g}_{i+1} = \frac{1}{B + \alpha}(K_A^*f + (B \cdot I - \mathbf{M})\mathbf{g}_i) .$$

By Theorem 2.2.1 (but is also clear since the mapping is contractive) we immediately deduce that for arbitrarily chosen  $\mathbf{g}_0$  the sequence  $\mathbf{g}_i$  converges in norm to a solution of the normal equation

$$(\mathbf{M} + \alpha I)\mathbf{g} = K_A^*f , \tag{6.2.1}$$

or explicitly

$$\mathbf{g} = \frac{1}{B + \alpha} \left( \sum_{i=0}^{\infty} \left[ \frac{BI - \mathbf{M}}{B + \alpha} \right]^i \right) K_A^*f ,$$

which can be recasted as

$$\mathbf{g} = \frac{1}{B + \alpha} \left( \sum_{i=0}^{\infty} \left[ I - \frac{\mathbf{M} + \alpha I}{B + \alpha} \right]^i \right) K_A^* f ,$$

i.e. the damped Landweber iteration is sort of damped Richardson iteration with relaxation parameter  $(B + \alpha)^{-1}$  and operator  $\mathbf{M} + \alpha$ . Thus, for  $p = 2$ , we may directly apply the adaptive concept of the latter section.

For  $1 \leq p < 2$ , we have

$$\mathbf{g}_{i+1} = S_{p, \alpha/2B} \left( \mathbf{g}_i - \frac{1}{B} (\mathbf{M} \mathbf{g}_i - K_A^* f) \right) . \quad (6.2.2)$$

Iteration (6.2.2) is nothing than a damped Richardson iteration with relaxation parameter  $B^{-1}$  and operator  $\mathbf{M}$  and where the generalized shrinkage operator with threshold  $\alpha/2B$  is applied in each step. In principle, formula (6.2.2) would suggest the following adaptive concept (where we omit detailed constants since we do not give any convergence proof here):

Let  $\mathbf{z} = K_A^* f$ , then

***shrinkSOLVE***  $(\epsilon, \mathbf{M}, \mathbf{z}) \rightarrow \mathbf{g}_\epsilon$ :

$i := 0, \mathbf{g}_0 := 0, \epsilon_0 := c_1$

**while**  $\epsilon_i > \epsilon$  **do**

$i := i + 1$

$\epsilon_i := c_2(\epsilon_{i-1})$

$\mathbf{z}_i := \mathbf{APPLY}(c_3(\epsilon_i), \mathbf{g})$

$\mathbf{h}_{i,0} := \mathbf{g}_{i-1}$

**for**  $j = 1, \dots, K$  **do**

$$\mathbf{h}_{i,j} := S_{p, \frac{\alpha}{2B}} \left( \mathbf{h}_{i,j-1} - \frac{1}{B} (\mathbf{APPLY}(c_4(\epsilon_i), \mathbf{M}, \mathbf{h}_{i,j-1}) - \mathbf{z}_i) \right)$$

**end**

$\mathbf{g}_i := \mathbf{COARSE}(c_5(\epsilon_i), \mathbf{h}_{i,K})$

**end**

$\mathbf{g}_\epsilon := \mathbf{g}_i$

**Remark 6.2.1** *In the nonlinear case, the situation is completely different since we have to treat two nested iterations in a nonlinear framework. For the single frame situation we considered the functional*

$$J_\alpha(\mathbf{g}) = \|y^\delta - T(F^* \mathbf{g})\|_{L_2(\Omega)}^2 + 2\alpha \Psi(\mathbf{L} \mathbf{g}) ,$$

where we have seen that a method to approach a solution is the Landweber iteration which is based on solving the following fixed point equation in each step,

$$\mathbf{g}_{k+1} = \mathbf{S}_{\alpha, \mathbf{L}, \mathbf{C}} \left( M(\mathbf{g}_{k+1}, \mathbf{g}_k) \right) .$$

Involving the fixed point iteration by some index  $l$ , we obtain

$$\mathbf{g}_{k+1, l+1} = \mathbf{S}_{\alpha, \mathbf{L}, \mathbf{C}} \left( \mathbf{g}_k - \frac{1}{C} \left( FT'(F\mathbf{g}_{k+1, l})^* T(F\mathbf{g}_k) - FT'(F\mathbf{g}_{k+1, l})^* y^\delta \right) \right) .$$

Since we always have to evaluate the nonlinear operators  $T'(\cdot)$  (inner iteration) and  $T(\cdot)$  (outer iteration), the technology of **APPLY** is not adequate (since its based on compressible matrices). The development of some reasonable adaptive strategy for the full nonlinear problem while using multi frames is because of the resulting computational complexity very important – but it remains for the moment a pipe dream.

# Bibliography

- [Bak92] A. W. Bakushinskii. The problem of the convergence of the iteratively regularized Gauss–Newton method. *Comput. Maths. Math. Phys.*, (32):1353–1359, 1992.
- [BGV92] B. E. Boser, I. M. Guyon, and V. N. Vapnik. A training algorithm for optimal margin classifiers. In D. Haussler, editor, *Proc. of the 5th ACM Workshop on Computational Learning Theory*, pages 144–152, Pittsburgh, PA, 1992. ACM Press.
- [BNS97] B. Blaschke, A. Neubauer, and O. Scherzer. On convergence rates for the iteratively regularized Gauss–Newton method. *IMA Journal of Numerical Analysis*, (17):421–436, 1997.
- [CD95] R.R. Coifman and D. Donoho. Translation–invariant de–noising. in *Wavelets and Statistics*, A. Antoniadis and G. Oppenheim, eds., Springer–Verlag, New York, pages 125–150, 1995.
- [CDF92] A. Cohen, I. Daubechies, and J.-C. Feauveau. Biorthogonal bases of compactly supported wavelets. *Comm. Pure Appl. Math.*, 45:485–560, 1992.
- [CDPX99] A. Cohen, R. DeVore, P. Petrushev, and H. Xu. Nonlinear Approximation and the Space  $BV(\mathbb{R}^2)$ . *American Journal of Mathematics*, (121):587–628, 1999.
- [CDS95] S.S. Chen, D.L. Donoho, and M.A. Saunders. Atomic decomposition by basis pursuit. *Report Stanford University*, 1995.
- [CGLT79] Y. Censor, D. Gustafson, A. Lent, and H. Tuy. A new approach to the emission computerized tomography problem: simultaneous calculation of attenuation and activity coefficients. *IEEE Trans. Nucl. Sci.*, (26):2275–79, 1979.
- [CGT99] G. Zikos C. Garcia and G. Tziritas. Face detection in color images using wavelet packet analysis. *IEEE Int. Conf. on Multimedia Computing and Systems*, 1999.
- [Cro84] F. Crow. Summed-area tables for texture mapping. In *Proc. of SIGGRAPH*, 18(3):207 – 212, 1984.

- [Dah96] W. Dahmen. Stability of multiscale transformations. *The Journal of Fourier Analysis and Applications*, 2:341–361, 1996.
- [Dau88] I. Daubechies. Time-frequency localisation operators: A geometric phase space approach. *IEEE Transactions on Information Theory*, (34):605–612, 1988.
- [Dau92] I. Daubechies. *Ten Lectures on Wavelets*. SIAM, Philadelphia, 1992.
- [Dau93] I. Daubechies. Wavelet transforms and orthonormal wavelet bases. *Proceedings of Symposia in Applied Mathematics*, (47), 1993.
- [DD04a] M. Defrise and C. DeMol. Inverse imaging with mixed penalties. *Conference proceedings*, 2004.
- [DD04b] M. Defrise and C. DeMol. Linear inverse problems with mixed smoothness and sparsity constraints. *Preprint*, 2004.
- [DDD04] I. Daubechies, M. Defrise, and C. DeMol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Comm. Pure Appl. Math*, 57:1413–1541, 2004.
- [DES98] P. Deuffhard, H. W. Engl, and O. Scherzer. A convergence analysis of iterative methods for the solution of nonlinear ill-posed problems under affinely invariant conditions. *Inverse Problems*, (14):1081–1106, 1998.
- [DeV98] R. DeVore. Nonlinear Approximation. *Acta Numerica*, 7:51–150, 1998.
- [Dic99] V. Dicken. A new approach towards simultaneous activity and attenuation reconstruction in emission tomography. *Inverse Problems*, 15(4):931–960, 1999.
- [DJP88] R. DeVore, B. Jawerth, and V. Popov. Interpolation of besov spaces. *Trans. Math. Soc.*, 305:397–414, 1988.
- [DJP92] R. DeVore, B. Jawerth, and V. Popov. Compression of wavelet decompositions. *Amer. J. Math.*, 114:737–785, 1992.
- [DM76] M. Duflo and C. C. Moore. On the regular representation of a nonunimodular locally compact group. *J. Funct. Anal.*, 21:209–243, 1976.
- [DM95] S. Dahlke and P. Maaß. The affine uncertainty principle in one and two dimensions. *Comp. Math. Appl.*, 30(3-6):293–305, 1995.
- [DS52] R.J. Duffin and A.C. Schäfer. A class of nonharmonic Fourier series. *Trans. Am. Math. Soc.*, 72:341–366, 1952.
- [DST04a] S. Dahlke, G. Steidl, and G. Teschke. Coorbit Spaces and Banach Frames on Homogeneous Spaces with Applications to Analyzing Functions on Spheres. *Adv. Comput. Math.*, 1-2(21):147–180, 2004.



- [DST04b] S. Dahlke, G. Steidl, and G. Teschke. Weighted Coorbit Spaces and Banach Frames on Homogeneous Spaces. *Journal of Fourier Analysis and Applications*, 5(10):507–539, 2004.
- [DT04] I. Daubechies and G. Teschke. Wavelet-based image decomposition by variational functionals. *Proc. SPIE Vol. 5266, p. 94-105, Wavelet Applications in Industrial Processing; Frederic Truchetet; Ed.*, Feb. 2004.
- [DT05] I. Daubechies and G. Teschke. Variational image restoration by means of wavelets: simultaneous decomposition, deblurring and denoising. *Applied and Computational Harmonic Analysis*, (19(1)):1–16, 2005.
- [EHN96a] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*. Kluwer, Dordrecht, 1996.
- [EHN96b] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*. Kluwer, Dordrecht, 1996.
- [FJ90] M. Frazier and B. Jawerth. A discrete transform and decompositions of distribution spaces. *J. of Functional Anal.*, 93:34–170, 1990.
- [FvSCB00] F. C. A. Fernandes, R. v. Spaendonck, M. J. Coates, and S. Burrus. Directional Complex-Wavelet Processing. *Proceedings of Society of Photo-Optical Instrumental Engineers—SPIE2000, Wavelet Applications in Signal Processing VIII*, San Diego., 2000.
- [GMP85] A. Grossmann, J. Morlet, and T. Paul. Transforms associated to squareintegrable group representations i. *Math. Phys.*, 27, 1985.
- [GMP86] A. Grossmann, J. Morlet, and T. Paul. Transforms associated to squareintegrable group representations ii. *Ann. Inst. H. Poincaré*, 45, 1986.
- [Han95] M. Hanke. *Conjugate gradient type methods for ill-posed problems*. Longman, 1995.
- [Han97a] M. Hanke. A regularizing Levenberg–Marquardt scheme, with applications to inverse groundwater filtration problems. *Inverse Problems*, (13):79–95, 1997.
- [Han97b] M. Hanke. Regularizing properties of a truncated Newton–cg algorithm for nonlinear ill-posed problems. *Num. Funct. Anal. Optim.*, (18):971–993, 1997.
- [H.G86] H.G.Feichtinger. Atomic characterization of modulation spaces through Gabor-type representations. *Proc.Conf. “Constructive Function Theory, Edmonton, Rocky Mount.J.Math.*, 149:113–126, 1986.
- [HK88] H.G.Feichtinger and K.Gröchenig. A unified approach to atomic decompositions via integrable group representations. *Proc.Conf. Function Spaces and Applications, Lund 1986, Lecture Notes in Math.*, 1302:52–73, 1988.

- [HK89a] H.G.Feichtinger and K.Gröchenig. Banach spaces related to integrable group representations and their atomic decomposition I. *J.Funct.Anal.*, 86:307–340, 1989.
- [HK89b] H.G.Feichtinger and K.Gröchenig. Banach spaces related to integrable group representations and their atomic decomposition II. *Monatsh.Math.*, 108:129–148, 1989.
- [HK92] H.G.Feichtinger and K.Gröchenig. Non-orthogonal wavelet and Gabor expansions and group representations. *in: Wavelets and Their Applications, eds. M.B.Ruskai et.al., Jones and Bartlett, Boston*, pages 353–376, 1992.
- [HNS95] M. Hanke, A. Neubauer, and O. Scherzer. A convergence analysis of the Landweber iteration for nonlinear ill-posed problems. *Numerische Mathematik*, (72):21–37, 1995.
- [Hoh97] T. Hohage. Logarithmic convergence rates of the iteratively regularized Gauss–Newton method for an inverse potential and an inverse scattering problem. *Inverse Problems*, (13):1279–1299, 1997.
- [Hol95] M. Holschneider. *Wavelets An Analysis Tool*. Clarendon Press, Oxford, 1995.
- [HT05] M. Holschneider and G. Teschke. Existence and computation of optimally localized coherent states. *DFG-SPP-1114 preprint 89, (submitted)*, 2005.
- [JT05] F. Jaillet and B. Torresani. Time–frequency jigsaw puzzle: adaptive multi-window and multilayered Gabor expansions. *preprint*, 2005.
- [Kal97] B. Kaltenbacher. Some Newton–type methods for the regularization of nonlinear ill-posed problems. *Inverse Problems*, (13):729–753, 1997.
- [Kar05] D.A. Karras. Improved defect detection in textile visual inspection using wavelet analysis and support vector machines. *ICGST International Journal on Graphics, Vision and Image Processing*, V4, 2005.
- [Kin99] N. Kinsbury. Image processing with complex wavelets. *Phil. Trans. R. Soc. Lond.*, Sept. 1999.
- [KOG01] D. Keren, M. Osadchy, and C. Gotsman. Antifaces: a novel, fast method for image detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23:747–761, July 2001.
- [Kre89] R. Kress. *Linear Integral Equation*. Springer, New York, 1989.
- [Lan51] L. Landweber. An iteration formula for fredholm integral equations of the first kind. *Amer. J. Math.*, (73):615–624, 1951.
- [Lou89] A. K. Louis. *Inverse und schlecht gestellte Probleme*. Teubner, Stuttgart, 1989.

- [Mey02] Y. Meyer. Oscillating Patterns in Image Processing and Nonlinear Evolution Equations. *University Lecture Series Volume 22*, AMS, 2002.
- [MT05] S. Molla and B. Torresani. An hybrid audio scheme using hidden markov models of waveforms. *to appear in Applied and Computational Harmonic Analysis*, 2005.
- [MY93] S. H. Manglos and T. M. Young. Constrained intraSPECT reconstructions from SPECT projections. In *Conf. Rec. IEEE Nuclear Science Symp. and Medical Imaging Conference, San Francisco, CA*, pages 1605–1609. 1993.
- [MZ92] S. Mallat and S. Zhong. Characterization of Signals from Multiscale Edges. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(7):710–732, July 1992.
- [Nat01] F. Natterer. Inversion of the attenuated radon transform. *Inverse Problems*, 17(1):113–119, 2001.
- [Opi67] Z. Opial. Weak convergence of the sequence of successive approximations for nonexpansive mappings. *Bull. Amer. Math. Soc.*, (73):591–597, 1967.
- [OSV02] S. Osher, A. Sole, and L. Vese. Image decomposition and restoration using total variation minimization and the  $H^{-1}$  norm. Technical Report 02-57, University of California Los Angeles C.A.M., 2002.
- [OV02] S. Osher and L. Vese. Modeling textures with total variation minimization and oscillating patterns in image processing. Technical Report 02-19, University of California Los Angeles C.A.M., 2002.
- [Ram] R. Ramlau. Modifizierte Landweber-Iterationen für Inverse Probleme. PhD thesis, Universität Potsdam (1997).
- [Ram99] R. Ramlau. A modified Landweber–method for inverse problems. *Numerical Functional Analysis and Optimization*, 20(1& 2), 1999.
- [Ram02a] R. Ramlau. Morozov’s discrepancy principle for Tikhonov regularization of nonlinear operators. *Numer. Funct. Anal. and Optimiz*, 23(1&2):147–172, 2002.
- [Ram02b] R. Ramlau. A steepest descent algorithm for the global minimization of the Tikhonov– functional. *Inverse Problems*, 18(2):381–405, 2002.
- [Ram03] R. Ramlau. TIGRA—an iterative algorithm for regularizing nonlinear ill–posed problems. *Inverse Problems*, 19(2):433–467, 2003.
- [Ram04] R. Ramlau. On the use of fixed point iterations for the regularization of nonlinear ill-posed problems. submitted for publication, 2004.

- [RCNB00] R. Ramlau, R. Clackdoyle, F. Noo, and G. Bal. Accurate attenuation correction in SPECT imaging using optimization of bilinear functions and assuming an unknown spatially-varying attenuation distribution. *Z. Angew. Math. Mech.*, 80(9):613–621, 2000.
- [ROF92] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variations based noise removal algorithms. *Physica D*, 60:259–268, 1992.
- [RRTV05] M. Rätsch, S. Romdhani, G. Teschke, and T. Vetter. Overcomplete Wavelet Approximation of a Support Vector Machine for Efficient Classification. *to appear in Proc. 27th Annual meeting of the German Association for Pattern Recognition*, 2005.
- [RRV05] M. Rätsch, S. Romdhani, and T. Vetter. Efficient face detection by a cascaded support vector machine using haar-like features. In *Proc. DAGM’04: 26th Pattern Recognition Symposium*, pages 62 – 70, 2005.
- [RT04] R. Ramlau and G. Teschke. Regularization of Sobolev Embedding Operators and Applications Part II: Data Driven Regularization and Applications. *Sampling Theory in Signal and Image Processing*, 3(3):225–246, 2004.
- [RT05a] R. Ramlau and G. Teschke. A Thresholding Iteration for Nonlinear Operator Equations with Sparsity Constraints. *DFG-SPP-1114 preprint (submitted for journal publication)*, 2005.
- [RT05b] R. Ramlau and G. Teschke. Tikhonov Replacement Functionals for Iteratively Solving Nonlinear Operator Equations. *DFG-SPP-1114 preprint 85 (to appear in Inverse Problems)*, 2005.
- [RTSB01] S. Romdhani, P. Torr, B. Schölkopf, and A. Blake. Computationally efficient face detection. In *Proceedings of the 8th International Conference on Computer Vision*, July 2001.
- [Sch93] O. Scherzer. The use of Morozov’s discrepancy principle for Tikhonov regularization for solving nonlinear ill-posed problems. *Computing*, (51):45–60, 1993.
- [Sch98] O. Scherzer. A modified landweber iteration for solving parameter estimation problems. *Appl. Math. Optim.*, 38:45–68, 1998.
- [Sel01] I. W. Selesnick. Hilbert transform pairs of wavelet bases. *IEEE Signal Processing Letters*, 8(6):170–173, June 2001.
- [SMB<sup>+</sup>99] B. Schölkopf, S. Mika, C. Burges, P. Knirsch, K.-R. Müller, G. Rätsch, and A. Smola. Input space vs. feature space in kernel-based methods. *IEEE Transactions on Neural Networks*, 10(5):1000 – 1017, 1999.
- [ST87] H.-J. Schmeisser and H. Triebel. *Topics in Fourier Analysis and Function Spaces*. John Wiley and Sons, New York, 1987.

- [Ste03] R. Stevenson. Adaptive solution of operator equations using wavelet frames. *SIAM J. Numer. Anal.*, (41):1074–1100, 2003.
- [Tes05a] G. Teschke. Construction of Generalized Uncertainty Principles and Wavelets in Bessel Potential Spaces. *International Journal of Wavelets, Multiresolution and Information Processing*, 3(2), 2005.
- [Tes05b] G. Teschke. Multi-Frame Representations in Linear Inverse Problems with Mixed Multi-Constraints. *DFG-SPP-1114 preprint 90 (submitted for journal publication)*, 2005.
- [Tes05c] G. Teschke. Multi-Frames in Thresholding Iterations for Nonlinear Operator Equations with Mixed Sparsity Constraints. *(submitted for journal publication)*, 2005.
- [Tri78] H. Triebel. *Interpolation Theory, Function Spaces, Differential Operators*. Verlag der Wissenschaften, Berlin, 1978.
- [TTP<sup>+</sup>90] J. A. Terry, B. M. W. Tsui, J. R. Perry, J. L. Hendricks, and G. T. Gullberg. The design of a mathematical phantom of the upper human torso for use in 3-d spect imaging research. In *Proc. 1990 Fall Meeting Biomed. Eng. Soc. (Blacksburg, VA)*, pages 1467–74. New York University Press, 1990.
- [Vap98] V. Vapnik. *Statistical Learning Theory*. Wiley, N.Y., 1998.
- [VJ02] P. Viola and M. Jones. Robust real-time object detection. *International Journal of Computer Vision*, 2002.
- [WCCG96] A. Welch, R. Clack, P. E. Christian, and G. T. Gullberg. Toward accurate attenuation correction without transmission measurements. *Journal of Nuclear Medicine*, (37):18P, 1996.
- [WCNG97] A Welch, R Clack, F Natterer, and G T Gullberg. Toward accurate attenuation correction in SPECT without transmission measurements. *IEEE Trans. Med. Imaging*, (16):532–40, 1997.
- [ZS02] G. Bebis D. DiMeo Z. Sun, R. Miller. A real-time precrash vehicle detection system. *WACV*, pages 171–176, 2002.