Iterated hard shrinkage for minimization problems with sparsity constraints

Kristian Bredies, Dirk A. Lorenz *

June 26, 2006

Abstract

A new iterative algorithm for the solution of minimization problems which involve sparsity constraints in form of ℓ^p-penalties is proposed. In contrast to the well-known algorithm considered by Daubechies, Defrise and De Mol, it uses hard instead of soft thresholding. It is shown that the hard thresholding algorithm is a special case of the generalized conditional gradient method. Convergence properties of the generalized conditional gradient method with quadratic discrepancy term are analyzed. This leads to strong convergence of the iterates with convergence rates $\mathcal{O}(n^{-1/2})$ and $\mathcal{O}(\lambda^n)$ for $p = 1$ and $1 < p \leq 2$ respectively. Numerical experiments on image deblurring and backwards heat condution illustrate the performance of the algorithm.

Keywords: sparsity constraints, iterated hard shrinkage, generalized conditional gradient method, convergence analysis

AMS Subject Classification: 46N10, 49M05, 65K10

1 Introduction

This article deals with the solution of minimization problem which involve so-called sparsity constraints. Sparsity has been found as a powerful tool in several problems in recent years. It has been recognized, that sparsity is an important structure in many applications ranging from image processing to problems from engineering sciences. Throughout the article the following example will be used for illustration: Minimize the functional

$$\Psi(u) = \frac{\|Ku - f\|^2}{2} + \sum_k w_k |\langle u,\, \psi_k\rangle|^p \tag{1}$$

where $K : \mathcal{H}_1 \to \mathcal{H}_2$ is an operator between two Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, $\{\psi_k\}$ is an orthonormal basis of $\mathcal{H}_1$, $w_k > w_0 > 0$ is a weighting sequence, and for the exponent it holds $1 \leq p \leq 2$. In the following we will use the abbreviation $\langle u,\, \psi_k\rangle = u_k$ for the coefficients of u with respect to the basis $\{\psi_k\}$.

Problem of this type arise in different contexts:

*kbredies@math.uni-bremen.de, dlorenz@math.uni-bremen.de, Center for Industrial Mathematics / Fachbereich 3, University of Bremen, Postfach 33 04 40, 28334 Bremen

Sparse inverse problems [6]. Here K is a compact operator and one aims at solving the equation $Ku = f$. Further more one assumes that the right hand side is not know precisely but up to a certain precision $\|f - \bar{f}\| \leq \delta$. Since K is compact the solution of the problem $Ku = f$ is ill posed. A way out is to regularize the inversion of K by prior knowledge. As proved in [6] the minimization of the above functional Ψ provides a regularization which promotes sparsity of the solution in the basis $\{\psi_k\}$

Image deblurring [1]. Consider an image f which is degraded by blurring and noise, i.e. $f = K\bar{f} + \delta$. A standard Tikhonov regularization with a quadratic penalty $\Phi(u) = |u|^2_{H^s}$ would lead to a smooth minimizer with still blurred edges. A regularization better adapted to the situation of images is the penalization with the total variation $\Phi(u) = |u|_{TV}$ or (better suited for computation) the Besov semi-norm $\Phi(u) = |u|_{B^1_{1,1}}$, while the latter can be expressed precisely as in (1) with a wavelet base $\{\psi_k\}$.

Sparse representations in dictionaries [7]. The minimization problem (1) appears as the so called basis pursuit in the problem of finding sparse representations in dictionaries. Assume we have a noisy signal $f \in \mathbf{R}^n$ and seek for an approximation which is composed by a small number of "atoms" $\{\psi_k\}_{k=1,\dots,N} \in \mathbf{R}^n$. This can be stated as a constrained minimization problem

$$\min_{a \in \mathbf{R}^N} \sum |a_k| \text{ subject to } \|Da - f\|^2 \leq \delta$$

where $D = [\psi_1 \cdots \psi_N] \in \mathbf{R}^{n \times N}$. The unconstrained problem with Lagrange multiplier λ (depending on y and δ)

$$\min_{a \in \mathbf{R}^N} \sum |a_k| + \lambda \|Da - f\|^2$$

has precisely the same form as (1). See also [13, 17].

Operator equations with sparse frame expansions [16]. One can drop the assumption that the solution has a sparse representation in a given basis and consider the solution to be sparse in a given frame $\{\eta_k\}$ (see [5] for an introduction to frames). If one wants to solve the equation $Ku = f$ under the assumption that u has a sparse representation in the given frame, i.e. $u = \sum_k a_k \eta_k$ with only a few $a_k \neq 0$, one solves the minimization problem

$$\min_{a \in \ell^2} \frac{\|KFa - f\|^2}{2} + w_0 \sum_k |a_k|$$

($a = (a_k)$ and $Fa = \sum_k a_k \eta_k$).

One well understood algorithm for the minimization of Ψ is the iterated soft shrinkage algorithm introduced independently in [12] and [15]. The algorithm is analyzed in [1, 6] while in [6] the authors also show convergence of the algorithm in the infinite dimensional setting. While the question of convergence is answered it is still open how fast the iteration converges. Up to our knowledge no convergence rates have been proven for the iterated soft shrinkage. The main contribution of this article is a new minimization algorithm for which convergence rates are proved. Namely we prove that our algorithm converges linearly for $p > 1$ and like $n^{-1/2}$ for $p = 1$.

The article is organized as follows: In Section 2 we introduce our new algorithm. Section 3 is devoted to the analysis of convergence rates of the generalized conditional gradient method for functionals $F + \Phi$ with a smooth part F and a non-smooth part Φ. We consider a special case, adapted to the above minimization problem. The application of the results to the special problem (1) is given in Section 4 where explicit rates of converges for the iterated hard shrinkage algorithm are given. Section 5 presents numerical experiments on the convergence of our algorithm and compares our algorithm and the iterated soft shrinkage.

2 An iterative hard shrinkage algorithm

We state the problem of minimizing (1), i.e.

$$\min_{u \in \mathcal{H}_1} \ \Psi(u) \ ,$$

with the help of the basis expansion as

$$\min_{u \in \ell^2} \ \sum_k \frac{(Ku - f)_k^2}{2} + w_k |u_k|^p. \tag{2}$$

Remark 1. Note that we reformulated the problem with the operator $\{u_k\} \mapsto K \sum_k u_k \psi_k$ mapping from ℓ^2 to $\mathcal{H}_2$ also denoted with K.

To state our algorithm we introduce the following functions and constants: We denote $S_0 = \left(\frac{\|f\|^2}{2w_0}\right)^{1/p}$ and the functions

$$\varphi_p(x) = \begin{cases} |x|^p & \text{for } |x| \le S_0 \\ \frac{p}{2S_0^{2-p}} \left(x^2 + (\frac{2}{p} - 1)S_0^2\right) & \text{for } |x| > S_0 \end{cases} \tag{3}$$

and

$$H_p(x) = \begin{cases} \left(\frac{|x|}{p}\right)^{1/(p-1)} \operatorname{sgn}(x) & \text{for } |x| \le pS_0^{p-1} \\ \frac{S_0^{2-p} x}{p} & \text{for } |x| > pS_0^{p-1} \end{cases} \tag{4}$$

where we formally set

$$|x|^{\frac{1}{0}} = \begin{cases} 0 & \text{if } |x| < 1 \\ 1 & \text{if } |x| = 1 \\ \infty & \text{if } |x| > 1. \end{cases}$$

Note that φ_p is the usual power for small values and becomes quadratic outside of $[-S_0, S_0]$ in a $\mathcal{C}^1$-way. The function H_p is a kind of shrinkage for small values (remember that $1 \le p \le 2$) and linear outside of $[-S_0, S_0]$. See Figure 1 and Figure 2 for illustrations of φ_p and H_p, respectively.

Our algorithm is based on the following simple observation: It is clear that $\Psi(u^*) \le \Psi(0) = \frac{1}{2}\|f\|^2$ for the minimizer u^* of (2). Since Ψ fulfills the coercivity condition $\|u\|^p \le \frac{1}{w_0}\Psi(u)$ (which is proven in Proposition 1), we know $\|u^*\| \le$

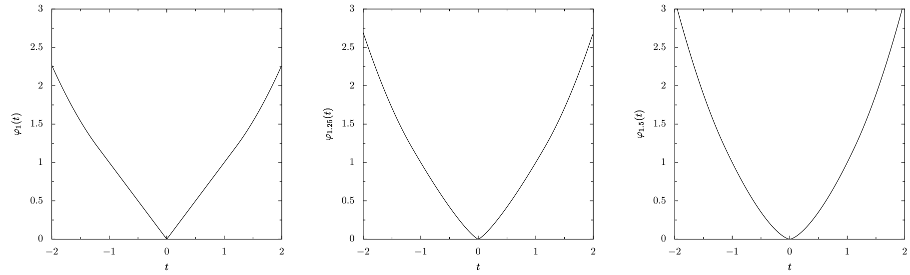

Figure 1: The function φ_p for $p = 1, 1.25, 1.5$.

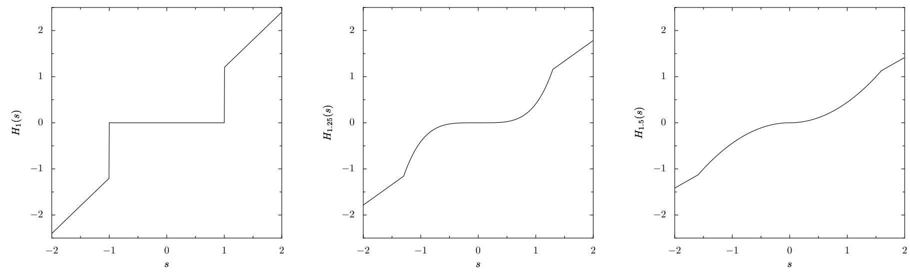

Figure 2: The function H_p for $p = 1, 1.25, 1.5$.

$(\frac{\|f\|^2}{2w_0})^{1/p}$. Hence the minimizer of Ψ does not change if we change the functional in the following way:

$$\min_u \ \tilde{\Psi}(u), \qquad \text{where} \qquad \tilde{\Psi}(u) := \sum_k \frac{(Ku-f)_k^2}{2} + w_k \varphi_p(u_k). \tag{5}$$

The minimization algorithm, which turns out to be a special case of the generalized conditional gradient algorithm now reads as follows:

Algorithm 1.

1. **Initialization.** Set $u^0 = 0$ and $n = 0$.

2. **Direction determination.** For $u^n \in \ell^2$ calculate

   $$v^n = \mathbf{H}_{p,w}\big(-K^*(Ku^n - f)\big)$$

   where

   $$\mathbf{H}_{p,w}\big(-K^*(Ku^n - f)\big)_k = H_p\Big(\frac{-\big(K^*(Ku^n - f)\big)_k}{w_k}\Big) \tag{6}$$

   with H_p according to (4).

3. **Step size determination.** Calculate s_n according to

   $$s_n = \min\Big\{1, \frac{\sum_{k=1}^\infty w_k\big(\varphi_p(u_k^n) - \varphi_p(v_k^n)\big) + \big(K^*(Ku^n - f)\big)_k (u_k^n - v_k^n)}{\|K(v^n - u^n)\|^2}\Big\} \tag{7}$$

4. **Iteration.** Set $u^{n+1} = u^n + s_n(v^n - u^n)$, $n := n+1$ and continue with Step 2.

Remark 2. This algorithm is very simple and hence easy to implement. We just need the functions φ_p and H_p available (which can be implemented explicitly) and of course the application of the operators K and K^*. In comparison to, for example the Landweber algorithm, this algorithm additionally requires the pointwise evaluation of H_p and φ_p which can be done rather fast. Moreover, since the iteration procedure in Step 4 is just a convex combination, we can reuse $K(v^n - u^n)$ for the computation of Ku^{n+1}, so we have to compute only one evaluation of K and K^* in each iteration, respectively.

Remark 3. The term

$$-D_n = \sum_{k=1}^{\infty} w_k\big(\varphi_p(u_k^n) - \varphi_p(v_k^n)\big) + \big(K^*(Ku^n - f)\big)_k(u_k^n - v_k^n)$$

in Step 3 of the algorithm can be used as an a-posteriori error bound on the distance to the minimizer, i.e. it holds $-D_n \geq \tilde{\Psi}(u^n) - \min_{u\in\ell^2} \tilde{\Psi}(u)$. So one can use the stopping criterion $-D_n < \varepsilon$ to assures that the minimal value is reached up to a certain tolerance $\varepsilon > 0$ in case of convergence, see Appendix A for details.

Remark 4. Note that if $p = 1$ the penalty functional

$$\Phi(u) = \sum_{k=1}^{\infty} w_k |u_k|$$

is non-differentiable. A common workaround for the lack of differentiability was to regularize the modulus function by the differentiable function

$$\varphi_{1,\varepsilon}(t) = \sqrt{t^2 + \varepsilon^2}$$

with a small $\varepsilon > 0$ (see e.g. [18] where this way was introduced for regularizing the TV norm). This always introduced some deviation to the real solution and posed numerical difficulties for very small ε. Especially the desired property of sparseness of the solution is lost.

In the algorithm presented above, we do in some sense the opposite: We modify the modulus function for large values in order to make the generalized conditional gradient method applicable (see later section for details). In this case, the modification is outside of the domain relevant for the minimization problem (2) and the solutions obtained are the exact sparse solutions.

The main result now are the convergence of the sequences generated by Algorithm 1 and an estimate on the distance to the true minimizer.

Theorem 1. *If $1 < p \leq 2$, then $u^n \to u^*$ to the unique minimizer of* (5) *in ℓ^2 with linear convergence speed, i.e.*

$$\|u^n - u^*\| \leq C\lambda^n$$

with a $0 < \lambda < 1$.

If $p = 1$ and K is injective, then $u^n \to u^*$ in ℓ^2 with convergence speed

$$\|u^n - u^*\| \leq Cn^{-1/2} \ .$$

Moreover, u^* is also a solution of the minimization problem (2).

The proof can be divided into two parts: First we examine the convergence of a general minimization algorithm, namely the generalized conditional gradient algorithm (cf. [2]) with discrepancy term $F(u) = \frac{1}{2}\|Ku - f\|^2$ and derive convergence rates for this procedure under certain conditions. We then apply these results to the special functional of type (5) and verify that the convergence criteria are satisfied.

3 Convergence analysis of generalized conditional gradient methods

The aim of this section is to provide convergence results for a general descent algorithm which turns out to be Algorithm 1 in the special case of the minimization problem (5). Its purpose is to solve the minimization problem

$$\min_{u \in \mathcal{H}_1} \Psi(u) \quad , \quad \Psi(u) = \frac{\|Ku - f\|^2}{2} + \Phi(u) \tag{8}$$

in a Hilbert space $\mathcal{H}_1$, with a linear and continuous operator $K : \mathcal{H}_1 \to \mathcal{H}_2$ and some suitable, convex Φ.

The algorithm is inspired by the generalized conditional gradient method [2] which addresses the minimization of general functionals of the form

$$\Psi(u) = F(u) + \Phi(u)$$

where F is smooth, but non-convex and Φ is convex but possibly non-smooth, resulting in a non-convex non-smooth Ψ. Here, we consider the special case where $F(u) = \frac{1}{2}\|Ku - f\|^2$, so problem (8) is convex, but still possibly non-smooth.

The generalized conditional gradient method applied to (8) and an explicit step-size rule gives the following algorithm:

Algorithm 2.

1. **Initialization.** Set $n = 0$ and choose u^0 such that $\Phi(u^0) < \infty$.

2. **Direction search.** For $n \geq 0$, calculate a minimizer of the approximate problem
$$\min_{v \in \mathcal{H}_1} \langle K^*(Ku^n - f), v \rangle + \Phi(v) \tag{9}$$
and denote a solution by v^n.

3. **Step size rule.** Choose the step size s_n according to
$$s_n = \min \left\{ 1, \frac{\Phi(u^n) - \Phi(v^n) + \langle Ku^n - f, K(u^n - v^n) \rangle}{\|K(v^n - u^n)\|^2} \right\} . \tag{10}$$

4. **Iteration.** Set $u^{n+1} = u^n + s_n(v^n - u^n)$, $n := n + 1$ and continue with Step 2.

In order to apply the algorithm, we have to ensure that the approximate problem (9) in Step 2 always has a solution. This is the case if the following conditions are satisfied:

Condition 1. *Let* $\Phi : H \to \mathbf{R} \cup \{\infty\}$ *fulfill*

- Φ *is proper, convex, and lower semi-continuous,*
- $\partial\Phi$ *is onto with* $(\partial\Phi)^{-1}$ *bounded, i.e.* $\Phi(u)/\|u\| \to \infty$ *whenever* $\|u\| \to \infty$.

In the following, we assume that Condition 1 on Φ is always satisfied.

From the minimization property (9) immediately follows the inequality:

$$D_n = \Phi(v^n) - \Phi(u^n) + \langle K^*(Ku^n - f),\, v^n - u^n \rangle \leq 0 \; . \tag{11}$$

This term plays a central role in the convergence analysis of this algorithm. Since we regard it as a generalization, the ideas utilized in the following are inspired by [10] where the analysis is carried out for the well-known conditional gradient method.

To prove convergence for Algorithm 2 we first derive that D_n serves as an estimate of the distance to the minimal value of Ψ in (8):

Lemma 1. *Denote by* $r_n = \Psi(u^n) - \min_{u \in \mathcal{H}_1} \Psi(u)$. *Then we have* $D_n \leq -r_n$. *In particular,* $D_n = 0$ *if and only if* u^n *is a solution of minimization problem* (8).

Proof. Choose a u^* which satisfies

$$\Psi(u^*) = \min_{u \in \mathcal{H}_1} \Psi(u) \; .$$

Since the minimization problem is well-posed (see [11], for example), such an u^* can always be found.

First observe that

$$\langle K^*(Ku^n - f),\, u^* - u^n \rangle \leq \frac{\|Ku^* - f\|^2}{2} - \frac{\|Ku^n - f\|^2}{2} \; .$$

Now

$$\begin{aligned}
\Phi(v^n) - \Phi(u^n) + &\langle K^*(Ku^n - f),\, v^n - u^n \rangle \\
&= \Phi(v^n) - \Phi(u^*) + \langle K^*(Ku^n - f),\, v^n - u^* \rangle \\
&\qquad + \Phi(u^*) - \Phi(u^n) + \langle K^*(Ku^n - f),\, u^* - u^n \rangle \\
&\leq \Phi(u^*) - \Phi(u^n) + \langle K^*(Ku^n - f),\, u^* - u^n \rangle \\
&\leq (F + \Phi)(u^*) - (F + \Phi)(u^n) = -r_n
\end{aligned}$$

by the minimizing property of v^n and the above. The characterization is a consequence of the first order necessary condition

$$u^n \text{ optimal} \quad \Rightarrow \quad \Phi(v^n) - \Phi(u^n) + \langle K^*(Ku^n - f),\, v^n - u^n \rangle \geq 0$$

(cf. [2]) and of the fact that $r_n \geq 0$. $\square$

Remark 5. One immediate consequence is that the step size rule (10) always produces $s_n \in [0, 1]$ and $s_n = 0$ if and only if u^n is a solution of the problem.

Remark 6. The above algorithm can be interpreted as a modification of the steepest descent/Landweber algorithm for the minimization of $\frac{1}{2}\|Ku - f\|^2$. Denote by T the (set-valued) solution operator of the minimization problem (9).

The steepest descent algorithm produces iterates $u^{n+1} = u^n + s_n(v^n - u^n)$ according to

$$v^n = -K^*(Ku^n - f) \qquad\qquad s_n = \frac{\langle Ku^n - f,\, K(u^n - v^n)\rangle}{\|K(v^n - u^n)\|^2} \ .$$

In comparison, Algorithm 2 also produces in the same manner, with similar directions and step sizes:

$$v^n \in T\big(-K^*(Ku^n - f)\big)$$
$$s_n = \min\Big\{1, \frac{\Phi(u^n) - \Phi(v^n) + \langle Ku^n - f,\, K(u^n - v^n)\rangle}{\|K(v^n - u^n)\|^2}\Big\}.$$

Note that in the generalized conditional gradient algorithm, the descent direction of the steepest descent of the quadratic part F is applied to a generally non-linear operator. Likewise, the step size is essentially the one used in the steepest descent algorithm, except for the presence of Φ. Finally, in the iteration step we can only allow convex combinations, therefore it differs with respect to this restriction.

Now for the convergence analysis we note that this generalized conditional gradient algorithm has very convenient descent properties.

Lemma 2. *Denote by* $r_n = \Psi(u^n) - \min_{u \in \mathcal{H}_1} \Psi(u)$. *Then*

$$r_{n+1} - r_n \leq \frac{-r_n^2}{2\|K(v^n - u^n)\|^2} \ .$$

Proof. First note that

$$\begin{aligned}
&F\big(u^n + s_n(v^n - u^n)\big) - F(u^n) \\
&\quad = \frac{\|K\big(u^n + s_n(v^n - u^n)\big) - f\|^2}{2} - \frac{\|Ku^n - f\|^2}{2} \\
&\quad = s_n\langle Ku^n - f,\, K(v^n - u^n)\rangle + \frac{s_n^2\|K(v^n - u^n)\|^2}{2}
\end{aligned}$$

and since Φ is convex

$$\Phi\big(u^n + s_n(v^n - u^n)\big) - \Phi(u^n) \leq s_n\big(\Phi(v^n) - \Phi(u^n)\big) \ .$$

Putting both together we get

$$\begin{aligned}
\Psi(u^{n+1}) - \Psi(u^n) \leq s_n\big(\Phi(v^n) - \Phi(u^n) + \langle K^*(Ku^n - f),\, v^n - u^n\rangle\big)& \\
+ \frac{s_n^2\|K(v^n - u^n)\|^2}{2}& \ .
\end{aligned}$$

We will now make use of D_n as defined in (11). First assume that $-D_n \geq \|K(v^n - u^n)\|^2$. Then the step size rule (10) yields $s_n = 1$ and it follows

$$r_{n+1} - r_n \leq D_n - \frac{D_n}{2} \leq \frac{-D_n^2}{2\|K(v^n - u^n)\|^2} \ .$$

In the case where $-D_n \le \|K(v^n - u^n)\|^2$ we have $s_n = -D_n/\|K(v^n - u^n)\|^2$, thus

$$r_{n+1} - r_n \le \frac{-D_n^2}{\|K(v^n - u^n)\|^2} + \frac{D_n^2}{2\|K(v^n - u^n)\|^2} = \frac{-D_n^2}{2\|K(v^n - u^n)\|^2} \ .$$

Finally, due to Lemma 1 it follows $-D_n^2 \le -r_n^2$ which implies the desired inequality. □

Such an estimate immediately implies that the distances to the minimum behave like $\mathcal{O}(n^{-1})$.

Lemma 3. *The distances to the minimum r_n satisfy*

$$r_n \le Cn^{-1}$$

for some $C > 0$ which is independent of n.

Proof. Due to Lemma 2 it is immediate that $r_{n+1} \le r_n \le r_0 = \Psi(u^0)$. Since Φ is coercive, there has to be a $C_1 > 0$ such that $\|u^n\| \le C_1$ for all n. From convex analysis we know that the solution operator of the minimization problem in Step 2 is bounded, whenever the property $\Phi(u)/\|u\| \to \infty$ if $\|u\| \to \infty$ is satisfied (see [14], for example). Thus, it follows that $\|v^n\| \le C_2$ for some constant $C_2 > 0$. This gives the estimate

$$\|K(v^n - u^n)\|^2 \le \|K\|^2 (C_1 + C_2)^2 \ .$$

The following is a widely know trick for the estimation of the distance to the minimum. You can find a similar proof e.g. in [9]. Using Lemma 2 again gives

$$\begin{aligned} \frac{1}{r_{n+1}} - \frac{1}{r_n} = \frac{r_n - r_{n+1}}{r_{n+1} r_n} \ge &\frac{r_n^2}{2\|K(v^n - u^n)\|^2 r_{n+1} r_n} \\ &\ge 2\|K\|^{-2}(C_1 + C_2)^{-2} = C_3 > 0 \end{aligned}$$

and summing up yields

$$\frac{1}{r_n} - \frac{1}{r_0} = \sum_{i=0}^{n-1} \frac{1}{r_{i+1}} - \frac{1}{r_i} \ge C_3(n-1) \ .$$

Finally, since $C_3 > 0$, we conclude

$$r_n \le \big(C_3(n-1) + \tfrac{1}{r_0}\big)^{-1} \le Cn^{-1} \ .$$

□

Theorem 2. *The sequence $\{u^n\}$ generated by Algorithm 2 possesses a weakly convergent subsequence whose limit is a solution of the minimization problem* $\min_{u \in \mathcal{H}_1} \Psi(u)$. *On the other hand, each weak accumulation point of $\{u^n\}$ is a solution.*

Additionally, if K is injective or Φ is strictly convex, then the solution u^* of the minimization problem $\min_{u \in \mathcal{H}_1} \Psi(u)$ *is unique and the sequence $\{u^n\}$ generated by the Algorithm 2 converges weakly to u^*.*

Proof. From Lemma 3 we know that $\Psi(u^n) - \min_{u\in\mathcal{H}_1}\Psi(u) \leq Cn^{-1}$. Since Φ is coercive, we again get a bound $\|u^n\| \leq C_1$ for all n. Thus, we can extract a weakly convergent subsequence, not relabeled, with limit u^*. Since each subsequence of $\{u^n\}$ is a minimizing sequence and since Ψ is convex and lower semi-continuous, the functional is also weakly lower semi-continuous which implies

$$\Psi(u^*) \leq \lim_{n\to\infty}\Psi(u^n) = \min_{u\in\mathcal{H}_1}\ \Psi(u)$$

whenever a subsequence of $\{u^n\}$ converges weakly to a $u^* \in \mathcal{H}_1$.

For the uniqueness statement, note that if K is injective, then $F(u)$ is strictly convex. Thus, according to the assumptions, either F or Φ is strictly convex and in particular the sum $\Psi = F + \Phi$. Therefore, Ψ has to admit a unique minimizer u^*. The above and the usual subsequence argument then yields that u^n converges to weakly u^*. □

In many cases, strong convergence can also be established. For this purpose, we consider the functional

$$R(v) = \langle K^*(Ku^* - f),\, v - u^*\rangle + \Phi(v) - \Phi(u^*) \tag{12}$$

at a minimizer u^*. Note that if u^* is a solution, then $-K^*(Ku^* - f) \in \partial\Phi(u^*)$, so R can be interpreted as some kind of Bregman distance at u^* with respect to Φ. In particular, $\Psi(v) - \min_{u\in H}\Psi(u) \geq R(v) \geq 0$, so $R(v) = 0$ if v is a minimizer of Ψ.

Theorem 3. *Let $\{u^n\}$ be a sequence generated by Algorithm 2.*

If, for a $u^* \in \mathcal{H}_1$ and a closed subspace $M \subset \mathcal{H}_1$ we have for each $L > 0$

$$\|v - u^*\| \leq L \quad \Rightarrow \quad R(v) \geq c(L)\|P_M(v - u^*)\|^2$$

with some $c(L) > 0$, then $P_M(u^n) \to P_M(u^*)$ in $\mathcal{H}_1$.

If, moreover, $M^\perp$ is finite-dimensional, then there still exists a subsequence of $\{u^n\}$ which converges strongly to a solution. In particular, if K is injective, then $u^n \to u^*$ to the unique solution with convergence speed

$$\|u^n - u^*\| \leq Cn^{-1/2}\ .$$

In the case $M = \mathcal{H}_1$, the minimizer is unique regardless of K and we can improve the convergence speed to

$$\|u^n - u^*\| \leq C\lambda^{n/2} \quad , \quad r_n \leq r_0\lambda^n$$

with some $0 < \lambda < 1$.

Proof. From Lemma 3 we know that $r_n = \Psi(u^n) - \min_{u\in\mathcal{H}_1}\ \Psi(u)$ converges to zero with estimate $r_n \leq C_1 n^{-1}$. It is also clear that $\|u^n\| \leq C_2$, so we can find a $L > 0$ such that $\|u^n - u^*\| \leq L$. By assumption and convexity of F,

$$r_n \geq \Phi(u^n) - \Phi(u^*) + \langle K^*(Ku^* - f),\, u^n - u^*\rangle \geq c(L)\|P_M(u^n - u^*)\|^2$$

which implies the convergence $P_M(u^n) \to P_M(u^*)$ with rate

$$\|P_M(u^n) - P_M(u^*)\| \leq \sqrt{\frac{C_1}{c(L)}n^{-1/2}}\ .$$

From Theorem 2 we know there is a weakly convergent subsequence of u^n which converges to a solution. Denote this subsequence also by u^n and its weak limit by u^{**}. If $M^\perp$ is finite-dimensional, it follows $P_{M^\perp}(u^n) \to P_{M^\perp}(u^{**})$. By above $P_M(u^n) \to P_M(u^*)$ and in particular $P_M(u^{**}) = P_M(u^*)$. So it follows

$$u^n = P_M(u^n) + P_{M^\perp}(u^n) \to P_M(u^{**}) + P_M(u^{**}) = u^{**} \ .$$

The convergence statement in case of uniqueness then follows from the usual subsequence argument.

Now assume that K is injective. We renorm $\operatorname{range}(K) \subset \mathcal{H}_2$ according to $|||v||| = \|v\| + \|P_M K^{-1} v\|$ and verify that $\operatorname{range}(K)$ is complete under this norm: For a Cauchy sequence $\{v^n\}$ the sequence $\tilde{v}^n = P_M K^{-1} v^n$ is also a Cauchy sequence in $\mathcal{H}_1$, hence $\tilde{v}^n \to \tilde{v}$ which implies $K\tilde{v}^n \to K\tilde{v}$. Now $v^n - K\tilde{v}^n$ is Cauchy in the finite-dimensional space $\operatorname{range}(KP_{M^\perp})$, thus moreover $v^n - K\tilde{v}^n \to Kw$ for a $w \in M^\perp$. This gives the convergence $v^n \to K(u + w)$.

It follows from $|||Ku||| \leq (\|K\| + 1)\|u\|$ that K is continuous between the Banach spaces $\mathcal{H}_1 \to \operatorname{range}(K)_{|||\cdot|||}$ and also bijective, so by the open mapping theorem (cf. [8]), K is continuously invertible which means that

$$\|u\| \leq C_3\big(\|Ku\| + \|P_M u\|\big)$$

for each $u \in \mathcal{H}_1$.

Since u^* is optimal, $-K^*(Ku^* - f) \in \partial\Phi(u^*)$ meaning that

$$\begin{aligned} r_n = \frac{\|Ku^n - f\|^2}{2} + \Phi(u^n) - \frac{\|Ku^* - f\|^2}{2} - \Phi(u^*) \\ \geq \frac{\|Ku^n - f\|^2 - \|Ku^* - f\|^2 - 2\langle Ku^* - f,\, K(u^n - u^*)\rangle}{2} \\ = \frac{\|K(u^n - u^*)\|^2}{2} \ . \end{aligned}$$

Together with the above,

$$\begin{aligned} \|u^n - u^*\|^2 \leq 2C_3^2\big(\|K(u^n - u^*)\|^2 + \|P_M(u^n - u^*)\|^2\big) \\ \leq 2C_1C_3^2\big(c(L)^{-1} + 2\big)n^{-1} \end{aligned}$$

which proves the asserted convergence rate $n^{-1/2}$.

For the remaining statement, note that if $M = \mathcal{H}_1$, then the assumptions imply that $R(v) > 0$ for $v \neq u^*$. But since $R(v) = 0$ for each solution v, the solution u^* has to be unique. Now to prove the linear convergence speed, we first show that the solution operator T of the minimization problem (9) is locally Lipschitz continuous at u^*. Choose a $u \in \mathcal{H}_1$ with $\|u - u^*\| \leq L$ and denote by $v \in T(u)$. Since v is a solution of the minimization problem, it holds

$$\Phi(v) - \Phi(u^*) + \langle K^*(Ku - f),\, v - u^*\rangle \leq 0$$

thus

$$\begin{aligned} \|K\|^2\|u - u^*\|\|v - u^*\| \geq \langle K^*K(u^* - u),\, v - u^*\rangle \\ \geq \Phi(v) - \Phi(u^*) + \langle K^*(Ku^* - f),\, v - u^*\rangle = R(v) \geq c(L)\|v - u^*\|^2 \ . \end{aligned}$$

It follows that

$$\|v - u^*\| \le \frac{\|K\|^2}{c(L)} \|u - u^*\| \ .$$

Now we can estimate

$$\begin{aligned} 2\|K(v^n - u^n)\|^2 &\le 2\|K\|^2 (\|u^n - u^*\| + \|v^n - u^*\|)^2 \\ &\le 2\|K\|^2 \Big(1 + \frac{\|K\|^2}{c(L)}\Big)^2 \|u^n - u^*\|^2 = C_4 \|u^n - u^*\|^2 \ . \end{aligned}$$

Plugging this and the above estimate with $M = \mathcal{H}_1$ into Lemma 2 gives

$$r_{n+1} \le r_n \Big(1 - \frac{c(L)\|u^n - u^*\|^2}{2\|K(v^n - u^n)\|^2}\Big) \le r_n \Big(1 - \frac{c(L)}{C_4}\Big) \le \lambda r_n$$

with a $0 < \lambda < 1$. This proves

$$r_n \le \lambda^n r_0 \ .$$

Finally, the above estimate on the norm yields

$$\|u^n - u^*\| \le \sqrt{\frac{r_n}{c(L)}} \le C \lambda^{n/2}$$

with $C = \sqrt{r_0/c(L)}$. $\square$

4 Convergence rates for iterated hard shrinkage

In this section, we will show that the algorithm in the previous section yields Algorithm 1 when applied to functionals of the type (5), i.e.

$$F(u) = \frac{1}{2}\|Ku - f\|^2 \quad , \quad \Phi(u) = \sum_{k=1}^{\infty} w_k \varphi_p(u_k)$$

in $\mathcal{H}_1 = \ell^2$ and with φ_p according to (3).

But before we turn to the convergence proofs, let us justify the modifications on the problem by showing that (5) yields the same minimizers as (2).

Proposition 1. *Let problem* (2) *be given for a fixed $f \in \mathcal{H}_2$ and $1 \le p \le 2$. Then all minimizers u^* satisfy*

$$\|u^*\| \le S_0 \quad , \quad S_0 = \Big(\frac{\|f\|^2}{2w_0}\Big)^{1/p} \ .$$

Consequently, the minimizers of

$$\min_{u \in \mathcal{H}_1} \ \tilde{\Psi}(u) \quad , \quad \tilde{\Psi}(u) = \frac{\|Ku - f\|^2}{2} + \sum_{k=1}^{\infty} w_k \varphi_p(u_k)$$

coincide with the minimizers of (2) *whenever $\varphi_p(t) \ge |t|^p$ for $|t| \ge S_0$.*

Proof. Observe that

$$1 = \sum_{k=1}^{\infty} \Big(\frac{|u_k|}{\|u\|} \Big)^2 \leq \sum_{k=1}^{\infty} \Big(\frac{|u_k|}{\|u\|} \Big)^p \quad \Rightarrow \quad \|u\|^p \leq \sum_{k=1}^{\infty} |u_k|^p$$

hence the estimate follows from

$$\|u^*\|^p \leq \frac{1}{w_0} \sum_{k=1}^{\infty} w_k |u_k^*|^p \leq \frac{\Psi(0)}{w_0} = \frac{\|f\|^2}{2w_0} = S_0^p \ .$$

Further note that $\Psi(u) \leq \tilde{\Psi}(u)$ with equality if $\|u\| \leq S_0$. If u^* is a minimizer of Ψ, then we have

$$\tilde{\Psi}(u^*) = \Psi(u^*) \leq \Psi(u) \leq \tilde{\Psi}(u)$$

for all $u \in \mathcal{H}_1$. Thus, u^* is also in minimizer for $\tilde{\Psi}$. On the other hand, if u is not a minimizer for Ψ, then there exists a $u^* \in \mathcal{H}_1$ with $\|u^*\| \leq S_0$ such that

$$\tilde{\Psi}(u^*) = \Psi(u^*) < \Psi(u) \leq \tilde{\Psi}(u)$$

meaning that u is also not a minimizer for $\tilde{\Psi}$. □

Remark 7. Let us remark that φ_p as defined in (3) indeed fulfill $\varphi_p(t) \geq |t|^p$ for all $t \in \mathbf{R}$. This follows from $\varphi_p(\pm S_0) = |t|^p$ and a comparison of the derivatives, i.e.

$$|pt^{p-1}| \leq pS_0^{p-2}|t|$$

for $|t| \geq S_0$.

In order to apply the convergence results of the previous section, we have to verify that Algorithm 1 corresponds to Algorithm 2 in the case of (5). This will we done in the following. First, we check that the algorithm is indeed applicable, i.e. we show that the functional Φ meets Condition 1.

Lemma 4. *Let $\varphi : \mathbf{R} \to \mathbf{R} \cup \{\infty\}$ with $\varphi(0) = 0$ convex, lower semi-continuous, and such that $\varphi(t)/|t| \to \infty$ if $|t| \to \infty$ as well as $\varphi(t) \geq |t|^p$ for some $1 \leq p \leq 2$. Then*

$$\Phi(u) = \sum_{k=1}^{\infty} w_k \varphi(u_k)$$

is proper, convex, lower semi-continuous and fulfills $\Phi(u)/\|u\| \to \infty$ when $\|u\| \to \infty$.

Proof. To see that Φ is proper, convex and lower semi-continuous, we refer to the standard literature on convex analysis [11].

To establish the desired coercivity, suppose $\|u^n\| \to \infty$. Suppose there exists a sequence k_n such that $|u_k^n| \to \infty$. Then $\varphi(u_k^n)/|u_k^n| \to \infty$ and it follows

$$\frac{\Phi(u^n)}{\|u^n\|} \geq w_0 \varphi(u_k^n) \to \infty \ .$$

If there is no such sequence, $|u_k^n| \leq C$ for all $k, n \geq 1$ since $|u_k^n| \leq \|u^n\|$. Note that by construction $\varphi(t) \geq |t|^p$, hence

$$\|u^n\|^2 = \sum_{k=1}^{\infty} |u_k^n|^2 \leq \frac{C^{2-p}}{w_0} \sum_{k=1}^{\infty} w_k |u_k^n|^p \leq \frac{C^{2-p}}{w_0} \sum_{k=1}^{\infty} w_k \varphi(u_k^n) = \frac{C^{2-p}}{w_0} \Phi(u^n) \ .$$

Dividing by $\|u^n\|$ yields $\Phi(u^n)/\|u^n\| \to \infty$. □

We now analyze the steps of Algorithm 2 with respect to Algorithm 1. Choosing $u_0 = 0$ as initialization is feasible since always $\Phi(0) = 0$. The direction search in Step 2 amounts to solving the minimization problem

$$\min_{v \in \mathcal{H}_1} \sum_{k=1}^{\infty} \big(K^*(Ku - f)\big)_k (v_k - u_k) + w_k \varphi_p(u_k)$$

which can be done pointwise. This involves the solution of

$$\min_{t \in \mathbf{R}} \ st + \bar{w}\varphi_p(t)$$

for given $s, t \in \mathbf{R}$ and $\bar{w} > 0$. Noting the equivalence to $-\frac{s}{\bar{w}} \in \partial\varphi_p(t)$ and applying some subgradient calculus shows that a solution is indeed given by $H_p\big(-\frac{s}{\bar{w}}\big)$ according to (4). Pointwise application then gives that $v = \mathbf{H}_{p,w}\big(-K^*(Ku-f)\big)$ is a solution. Step 3 and 4 in Algorithm 1 is exactly corresponding to Step 3 and 4 in Algorithm 2 with the particular choice of Φ.

Remark 8. As Proposition 1 and Lemma 4 show, it is not necessary to modify the functional Φ in the case $1 < p \leq 2$. But if we apply Algorithm 2 to the unmodified functional, we have to evaluate $|s|^{1/(p-1)}$ for possibly great $|s|$. This might lead to numerical problems since $p \approx 1$ leads to high powers and the available range of numbers may be left.

Moreover, it is also not necessary to take a quadratic extension outside of $[-S_0, S_0]$ as done in (3). In fact, an arbitrary function φ satisfying the conditions of Proposition 1 and Lemma 4 is possible. The choice of φ however is reflected in the algorithm when $(\partial\varphi)^{-1}$ is computed. The quadratic extension in (3) leads to the linear sections in (4) which are easy to compute.

We want to apply the convergence results of Theorem 3. For establishing the estimates of the type $R(v) \geq c(L)\|v - u^*\|^2$ we need the following elementary result which is proven in Appendix B.

Lemma 5. *Let $1 < p \leq 2$. For each $C_1 > 0$ and $L > 0$ there exists a $c_1(L) > 0$ such that*

$$|t|^p - |s|^p - p\,\mathrm{sgn}(s)|s|^{p-1}(t - s) \geq c_1(L)|t - s|^2$$

for all $|s| \leq C_1$ and $|t - s| \leq L$.

Lemma 6. *Denote by u^* a solution of the minimization problem* (5) *with a modified functional φ_p meeting the requirements of Proposition 1 and consider the associated functional R according to* (12).

If $1 < p \leq 2$, then for each $L > 0$ there exists a $c(L) > 0$ such that

$$\|v - u^*\| \leq L \quad \Rightarrow \quad R(v) \geq c(L)\|v - u^*\|^2 \ .$$

If $p = 1$ then there exists a subspace $M \subset H$ with $M^\perp$ finite-dimensional such that for each $L > 0$ there exists a $c(L) > 0$ with

$$\|v - u^*\| \leq L \quad \Rightarrow \quad R(v) \geq c(L)\|P_M(v - u^*)\|^2 \ .$$

Proof. First consider the case $1 < p \leq 2$. If we have a minimizer u^* then $-K^*(Ku^* - f) \in \partial\Phi(u^*)$. The functional Φ is defined as a pointwise sum, thus, by standard arguments from convex analysis,

$$\big(-K^*(Ku^* - f)\big)_k = w_k p\,\mathrm{sgn}(u_k^*)|u_k^*|^{p-1}$$

for each $k \geq 1$.

From Proposition 1 we know that $|u_k^*| \leq S_0$ and applying Lemma 5 with $C_1 = S_0$ and an arbitrary $L > 0$ gives

$$\begin{aligned} w_k\big(\varphi_p(t) - \varphi_p(s) - p\,\mathrm{sgn}(s)|s|^{p-1}\big) \\ \geq w_0\big(|t|^p - |s|^p - p\,\mathrm{sgn}(s)|s|^{p-1}(t-s)\big) \geq w_0 c_1(L)|s-t|^2 \end{aligned}$$

for each $|s| \leq C_1$, $|s-t| \leq L$, remembering that $\varphi_p(s) = |s|^p$ for $|s| \leq C_1$ and $\varphi_p(t) \geq |t|^p$. Hence, if $\|v - u^*\| \leq L$,

$$\begin{aligned} R(v) = \sum_{k=1}^{\infty} w_k\big(\varphi_p(v_k) - \varphi_p(u_k^*) - p\,\mathrm{sgn}(u_k^*)|u_k^*|^{p-1}\big) \\ \geq w_0 c_1(L) \sum_{k=1}^{\infty} |v_k - u_k^*|^2 = c(L)\|v - u^*\|^2 \end{aligned}$$

This proves the desired statement for $1 < p \leq 2$.

Now let $p = 1$ and u^* be a minimizer. Then we know, analogly to the above, that

$$\big(-K^*(Ku^* - f)\big)_k \in w_k \partial\varphi_1(u_k^*)$$

for each $k \geq 1$. Since $\xi = -K^*(Ku^* - f) \in \ell^2$ we have $\xi_k \to 0$ for $k \to \infty$. Hence, we can choose a k_0 such that $|\xi_k| \leq \frac{w_0}{2}$ for $k \geq k_0$. Observe that $\partial\varphi_1$ is monotone and coincides with $\mathrm{sgn}(\cdot)$ in a neighborhood of 0 with $\partial\varphi_1(0) = [-1, 1]$, so $u_k^* = 0$ for $k \geq k_0$ since the opposite leads to a contradiction. Thus,

$$R(v) = \sum_{k=1}^{\infty} w_k\big(\varphi_1(v_k) - \varphi_1(u_k^*)\big) - \xi_k(v_k - u_k^*) \geq \sum_{k=k_0}^{\infty} w_k \varphi(v_k) - \xi_k v_k \ .$$

Due to the construction of φ_1 we can estimate $|t| \leq \varphi_1(t)$ which further leads to

$$R(v) \geq \sum_{k=k_0}^{\infty} w_k|v_k| - v_k\xi_k \geq w_0 \sum_{k=k_0}^{\infty} \frac{|v_k|}{2} = \frac{w_0}{2} \sum_{k=k_0}^{\infty} |v_k - u_k^*| \ .$$

Define

$$M = \{u \in \ell^2 \ : \ u_k = 0 \text{ for } k < k_0\}$$

which is clearly a closed subspace of ℓ^2 with finite-dimensional complement. Choose a v with $\|v - u^*\| \leq L$, then $|v_k - u_k|^2 \leq L^{-1}|v_k - u_k|$ so with $c(L) = w/(2L)$ we finally have

$$R(v) \geq c(L)\|P_M(v - u^*)\|^2 \ .$$

$\square$

Collecting the results of this section, we are able to apply Theorem 3 which finally gives our main convergence result for the iterated hard thresholding procedure in Algorithm 1:

Theorem 4. *If $1 < p \leq 2$, then the descent algorithm produces a sequence $\{u^n\}$ which converges linearly to the unique minimizer u^*, i.e.*

$$\|u^n - u^*\| \leq C\lambda^n$$

for a $0 \leq \lambda < 1$.

If $p = 1$ *and* K *is injective, then the descent algorithm produces a sequence* $\{u^n\}$ *which converges to the unique minimizer* u^* *in norm with speed*

$$\|u^* - u^n\| \leq C n^{-1/2} \ .$$

5 Numerical experiments

To illustrate the convergence behaviour of the iterated hard thresholding algorithm as stated in Algorithm 1, we performed numerical tests on two linear model problems and compared the results to the iterated soft thresholding algorithm introduced in [12] and [15]. Our primary aim is to demonstrate the applicability of the new algorithm, we thus perform the experiments for problems well-known in image processing and inverse problems.

The first model problem we tested is an image deblurring (deconvolution) problem with known kernel and penalization with respect to the coefficients in a Haar wavelet basis. The problem was discretized to a rectangular grid of points, i.e. we consider $u = \{u_{ij}\}$ with $1 \leq i \leq N$ and $1 \leq j \leq M$ and pose the minimization problem

$$\min_u \sum_{i=1}^{N} \sum_{j=1}^{M} \frac{\big((u * g)_{ij} - f_{ij}\big)^2}{2} + \sum_{k=1}^{NM} w_k |\langle u, \psi_k \rangle|^p$$

where ψ_k is the discrete two-dimensional wavelet basis spanning the space $\mathbf{R}^{NM}$ and $*$ denotes the usual discrete convolution with a kernel $g = \{g_{ij}\}$.

For our computations, we used an out-of-focus kernel g with radius $r = 6$ pixels which has been normalized to $\sum |g| = 0.99$ such that the associated operator's norm is strictly below 1. The original image of size 256×256 pixels with values in $[0, 1]$ was blurred with that kernel and in one case disturbed with Gaussian noise of variance η (see Figure 3). Then, Algorithm 1 as well as the iterated soft shrinkage procedure was applied for a suitable number of iterations with a wavelet decomposition up to level 5 and the following parameters:

1. $p = 1$, $w_k = 0.00002$, $\eta = 0$
2. $p = 1.75$, $w_k = 0.005$, $\eta = 0.025$

For the comparison of these two methods, we plotted the functional values at each iteration step. The results for the two deblurring tests are depicted in Figure 4.

Note that if $1 < p < 2$, we observed that the hard thresholding iteration usually performs faster than the soft thresholding algorithm although we did not optimize for computational speed. The reason for this is that the iterated soft thresholding algorithm demands the solution of

$$x + \bar{w} p \operatorname{sgn} x |x|^{p-1} = y$$

for all basis coefficients in each iteration step which has to be done by numerical approximation. Experiments show that it is necessary to compute the solution sufficiently precise since otherwise an increase of the functional value is possible. This is a relatively time-consuming task even if the quadratically

u_{orig} $\quad$ f_1 $\quad$ f_2

Figure 3: The images used as measured data f for the deblurring problem. From left to right, the original image, the blurred image without noise and the noisy blurred image used for test 1 and 2 respectively, are depicted.

convergent Newton method is used. In the iterated hard thresholding method (Algorithm 1), only the evaluation of $|x|^{1/(p-1)}$ is necessary which can usually be done significantly faster.

The second model problem we considered was solving the backwards heat equation in one dimension with sparsity constraints in a point basis. In this experiment we investigated the role of p and its influence on the performance of the algorithm.

The continuous model reads as follows: Consider an initial condition $u^0 \in \mathrm{L}^2([0,1])$ and the one dimensional heat equation with Dirichlet boundary conditions:

$$\begin{aligned} u_t &= u_{xx} \quad \text{for } (t,x) \in [0,T] \times [0,1] \\ u(0,x) &= u^0(x) \\ u(t,0) &= u(t,1) = 0. \end{aligned}$$

With K we denote the operator which maps the initial condition onto the solution of the above equation at time T. The problem of finding the initial condition u^0 from the measurement of the heat distribution f at time T is thus formulated as solving

$$Ku^0 = f.$$

For the discretized model, we choose $u^0 = \{u_k\} \in \mathbf{R}^N$, data $f = \{f_j\} \in \mathbf{R}^N$ where u_k stands for the value of u^0 at point $x_k = (k-1)/(N-1)$. In the case of the heat equation, the solution matrix for the forward problem reads as

$$K_{kj} = \frac{2}{N} \sum_{l=1}^{\infty} e^{-\pi^2 l^2 T} \sin(\pi l x_k) \sin(\pi l x_j).$$

The minimization problem then reads as

$$\min_u \sum_{j=1}^{M} \frac{\big((Ku)_j - f_j\big)^2}{2} + \sum_{k=1}^{N} w_k |u_k|^p \ .$$

To test the algorithms, we created an initial distribution u^0 with one spike. The data $f = Ku + \delta$ is degraded with a relative error of 15% (see Figure 5).

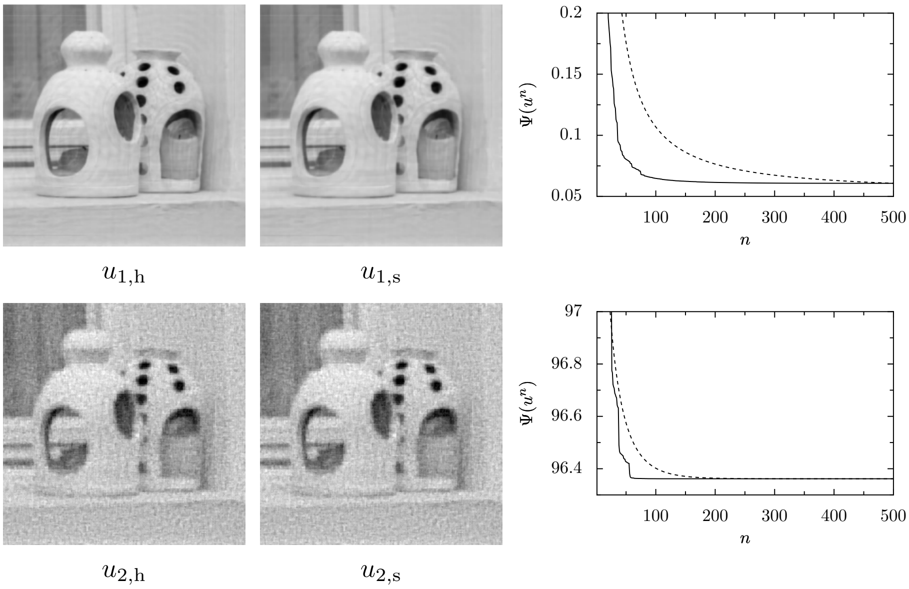

Figure 4: The results of the deblurring tests. In the left column you can see the reconstructed u after 500 iterations with the iterated hard thresholding method while in the middle column, the reconstruction u after the same amount of iterations with the iterated soft thresholding procedure is depicted. On the right hand side, a comparison of the descent of the functional values $\Psi(u^n)$ for both methods is shown. The solid and dashed lines represent the values of Ψ for the iterated hard and soft thresholding procedure, respectively.

We solved the above minimization problem with the iterated hard shrinkage algorithm for $w_k = 0.03$ and different values of p, namely

$$p_1 = 1,\ p_2 = 1.01,\ p_3 = 1.5.$$

As worked out in Appendix A the values $-D_n$ as defined in (11) can be used as a stopping criterion and in this experiment we stopped the iteration if $-D_n$ becomes smaller that 10^{-8} or the maximum number of 1000 iterations is reached. Figure 6 shows the results of the minimization process together with the estimators $-D_n$. Note that the minimizers for $p = 1$ and $p = 1.01$ does not differ to much although the estimator $-D_n$ behaves very different: For $p = 1$ it oscillates heavily and is decaying slowly as the theory indicates. The slight change from $p = 1$ to $p = 1.01$ results in an estimator which is still oscillating but vanishing much faster and the algorithm stopped after 136 iterations. For $p = 1.5$ the sparsity of the reconstruction is lost but the algorithm terminated after just 29 iterations.

6 Conclusion

We proposed a new algorithm for the minimization of functionals of type

$$\sum_k \frac{(Ku - f)_k^2}{2} + w_k |u_k|^p, \quad 1 \le p \le 2$$

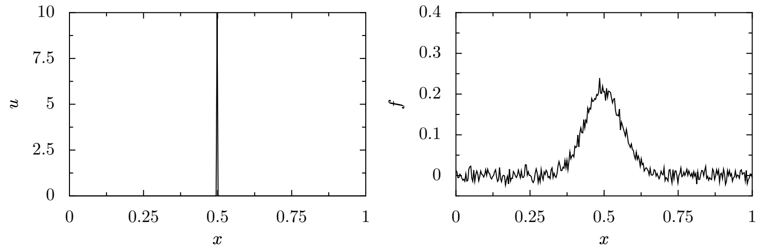

Figure 5: The data of the second model problem. From left to right: The spike initial heat distribution u and the heat distribution at time $T = .002$. (Note the different scaling.)

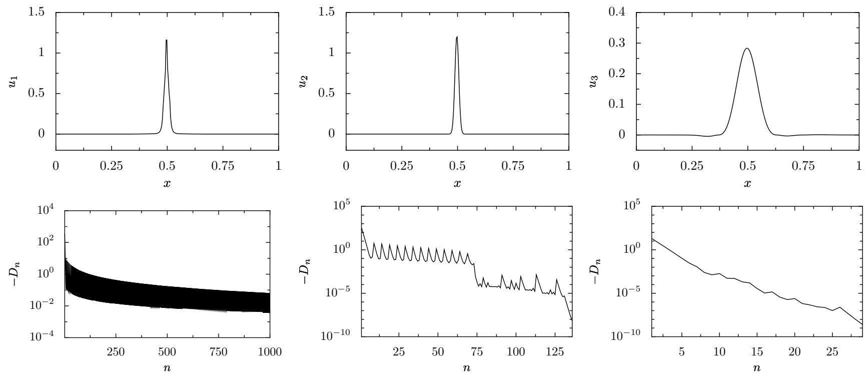

Figure 6: The results of the reconstruction of the initial condition. Top row from left to right: Reconstruction for $p = 1$, $p = 1.01$, and $p = 1.5$ respectively. Bottom row: The values of the estimator $-D_n$ in a logarithmic scale. (Note again different scaling.)

in the infinite dimensional setting. Our algorithm is based on iterated hard shrinkage. We established convergence rates for this algorithm, namely we proved convergence with rate $\mathcal{O}(n^{-1/2})$ for $p = 1$ and $\mathcal{O}(\lambda^n)$ for $1 < p \leq 2$. We remark that the iterative hard shrinkage is a discontinuous algorithm, hence convergence rates are not at all easy to establish.

For finite dimensional problems of the above type there are other algorithms with better performance (e.g. interior point methods, see [4]), but none of them has a proper foundation for the infinite dimensional setting. To our best knowledge the results stated here are the first results on convergence rates for a minimization algorithm of the above functional in infinite dimensions.

We emphasize that the convergence rate is only influenced by Φ and not by F, i.e. not by the operator K. For functionals $\Phi(u) = |u|_{H^s}$ one can expect similar convergence rates. Unfortunately, the case of total variation deblurring $\Phi(u) = |u|_{TV}$ seems not to fit into this context and further analysis in needed (while the case of the discrete TV-norm as treated in [3] goes well with this algorithm).

The change of the convergence rate from $p > 1$ to $p = 1$ is rather drastically: from linear convergence to convergence as $n^{-1/2}$ – and this is observed in the numerical experiments. To speed up the minimization for $p = 1$ it could be of interest to use the minimizer of the minimization problem for p slightly larger than 1 as initial value u_0 for the iteration with $p = 1$. Another possibility is to decrease p during the iteration and use the framework of Γ-convergence to prove convergence of the algorithm.

The new iterative hard shrinkage algorithm works well in many cases since it seems to have 'good ideas' for finding 'unconventional' descent directions. On the other hand it sometimes runs into a situation where it can not find good ways for further descent (see Figure 4: some steps reduce the functional values drastically while sometimes the functional does not decrease much for many iterations). The iterative soft shrinkage, whereas, gives well descent in every step. Hence, a combination of both may share the good features of both.

As a side result we established an estimator $-D_n$ for the distance of the nth iterate to the minimizer which can be used as a stopping criterion for iterative algorithms in case of convergence (see Remark 10 in Appendix A).

A Convergence of the a-posteriori error bound

For numerical computations, it is often useful to have an estimate on some error so one can formulate stopping criteria for iterative algorithms. As already mentioned in Remark 3, the generalized conditional gradient algorithm (Algorithm 2) involves the estimates to the distance to the minimizer $-D_n$ according to (11). The following proposition shows that they also vanish in case of convergence and therefore $-D_n < \varepsilon$ for some $\varepsilon > 0$ can be used as a stopping criterion for Algorithm 2.

Proposition 2. *Let Φ given according to Condition 1 and consider a sequence $\{u^n\}$ which is generated by Algorithm 2 for the solution of* (8).

Then

$$-D_n = \Phi(u^n) - \Phi(v^n) + \langle K^*(Ku^n - f),\, u^n - v^n\rangle \to 0$$

for $n \to \infty$.

Proof. First observe that descent property of Lemma 2 imply convergence of the functional values $\Psi(u^n)$, especially $\Psi(u^n) - \Psi(u^{n+1}) \to 0$. In Lemma 2 a slightly different version of the descent property is also proven:

$$\Psi(u^{n+1}) - \Psi(u^n) \leq \frac{-D_n^2}{2\|K(v^n - u^n)\|^2} \ .$$

Remember that Condition 1 and the descent property imply $\|v^n - u^n\| \leq C$ (cf. the proof of Lemma 3), thus

$$D_n^2 \leq 2C^2 \|K\|^2 \big(\Psi(u^n) - \Psi(u^{n+1})\big)$$

which proves the assertion. □

Remark 9. If Φ fulfills Condition 1, then it is possible to compute D_n without the knowledge of v^n with the help of the conjugate functional Φ^* (see [11] for an introduction): As already mentioned, the requirement that v^n is a solution of

$$\min_{v \in \mathcal{H}_1} \ \langle K^*(Ku^n - f),\, v\rangle + \Phi(v)$$

can equivalently be expressed by $-K^*(Ku^n - f) \in \partial\Phi(v^n)$ which is in turn, with the help of the Fenchel identity, equivalent to

$$-\Phi(v^n) - \langle K^*(Ku^n - f),\, v^n\rangle = \Phi^*\big(-K^*(Ku^n - f)\big) \ .$$

Plugging this into the definition of D_n in (11) gives

$$-D_n = \langle K^*(Ku^n - f),\, u^n\rangle + \Phi(u^n) + \Phi^*\big(-K^*(Ku^n - f)\big) \ , \qquad (13)$$

a formula where v^n no longer appears.

Remark 10. Equation (13) can also be used as a stopping criterion for iterative minimization algorithms in case the algorithm converges:

Assume that $u^n \to u^*$ with $\Psi(u^n) \to \Psi(u^*) = \min_{u\in\mathcal{H}_1} \Psi(u)$ and Φ satisfies Condition 1. Then $\frac{1}{2}\|Ku^n - f\|^2 \to \frac{1}{2}\|Ku^* - f\|^2$ by continuity and consequently $\Phi(u^n) \to \Phi(u^*)$ by the minimization property. Now Condition 1 on Φ implies that Φ^* is finite on the whole space $\mathcal{H}_1$: Assume the opposite, i.e. the existence of a sequence $v^n \in \mathcal{H}_1$ as well as a $v^* \in \mathcal{H}_1$ such that

$$\sup_n \ \langle v^n,\, v^*\rangle - \Phi(v^n) = \infty \ .$$

This immediately implies that $v^* \neq 0$ and $\|v^n\| \to \infty$ since Φ is bounded from below. But due to Condition 1, $\Phi(u)/\|u\| \to \infty$ which means that for each $C > \|v^*\|$ one can find an n_0 such that $-\Phi(v^n) \leq -C\|v^n\|$ for $n \geq n_0$ and, consequently,

$$\sup_{n\geq n_0} \ \langle v^n,\, v^*\rangle - \Phi(v^n) \leq (\|v^*\| - C)\|v^n\| \leq 0$$

which is a contradiction to $\Phi^*(v^*) = \infty$.

It follows that Φ^* is finite everywhere and hence continuous (see e.g. [14]). So additionally, $\Phi^*\big(-K^*(Ku^n - f)\big) \to \Phi^*\big(-K^*(Ku^* - f)\big)$ and the continuity of the scalar product leads to

$$\lim_{n\to\infty} -D_n = \langle K^*(Ku^* - f),\, u^*\rangle + \Phi(u^*) + \Phi^*\big(-K^*(Ku^* - f)\big) = 0 \ .$$

Remark 11. Let us show how the error bound $-D$ can be computed in the modified problem (5).

We first state the conjugate functionals Φ^* associated with the penalty functionals Φ. With some calculation one obtains for the conjugates of φ_p as defined in (3) that

$$\varphi_p^*(x^*) = \begin{cases} \frac{p^{-p'}}{p'}|x^*|^{p'} & \text{if } |x^*| \leq pS_0^{p-1} \\ \frac{1}{2p}S_0^{2-p}|x^*|^2 + \big((p-1)p^{p-1} - \frac{p}{2}\big)S_0^p & \text{if } |x^*| > pS_0^{p-1} \end{cases}$$

where p' denotes the dual exponent, i.e. $\frac{1}{p} + \frac{1}{p'} = 1$ and we formally define that $\frac{1}{\infty} = 0$.

Since the functional Φ is defined by pointwise summation in ℓ^2, we can also take the conjugate pointwise. This gives

$$\Phi^*(u^*) = \sum_{k=1}^{\infty} \varphi_p^*\Big(\frac{u_k^*}{w_k}\Big)$$

with the φ_p^* as above. Eventually, the error bound at some $u \in \ell^2$ can be expressed by

$$-D = \sum_{k=1}^{\infty} \big(K^*(Ku-f)\big)_k u_k + w_k \varphi_p(u_k) + \varphi_p^*\Big(\frac{-\big(K^*(Ku-f)\big)_k}{w_k}\Big) .$$

B Proof of Lemma 5

Our aim is to show that there for each $C_1 > 0$ and $L > 0$ there is a $c_1(L) > 0$ such that the estimate

$$|t|^p - |s|^p - p\,\mathrm{sgn}(s)|s|^{p-1}(t-s) \geq c_1(L)|t-s|^2$$

for all real numbers satisfying $|s| \leq C_1$ and $|t-s| \leq L$.

We will write the right hand side as

$$p\int_s^t \mathrm{sgn}(\tau)|\tau|^{p-1} - \mathrm{sgn}(s)|s|^{p-1} \,\mathrm{d}\tau$$

and estimate the integrand accordingly.

Without loss of generality, we assume $s \geq 0$ and consider $\tau \geq s$ first. Let $C_2 = C_1 + L$. Since $p \leq 2$, the function $\tau \mapsto |\tau|^{p-1}$ is concave, thus

$$(1-\theta)|s|^{p-1} + \theta C_2^{p-1} \leq |(1-\theta)s + \theta C_2|^{p-1}$$

for $\theta = (\tau - s)/(C_2 - s)$, implying that

$$|\tau|^{p-1} - |s|^{p-1} \geq (\tau - s)\frac{|s|^{p-1} + C_2^{p-1}}{C_2 - s} \geq (\tau - s)C_2^p .$$

On the other hand, for $0 \leq \tau < s$ we have

$$|\tau|^{p-1} - |s|^{p-1} \leq (p-1)|s|^{p-2}(\tau - s) \leq (p-1)C_1^{p-2}(\tau - s)$$

again by concavity and $-(p-1)|s|^{p-2}$ being the subgradient of $-|s|^{p-1}$. If $\tau < 0$, it follows

$$-|\tau|^{p-1} \leq L^{p-2}\tau \quad , \quad -|s|^{p-1} \leq -C_1^{p-2}s ,$$

hence

$$-|\tau|^{p-1} - |s|^{p-1} \leq \min\{C_1^{p-2}, L^{p-2}\}(\tau - s) .$$

Note that all inequalities involving $\tau - s$ remain true if the constant in front is smaller. Therefore, if we choose

$$c_1(L) = \frac{1}{2p}\min\{C_1^{p-2}, L^{p-2}, (p-1)C_1^{p-2}, C_2^{p-1}L^{-1}\}$$

it holds

$$|\tau|^{p-1} - |s|^{p-1} \begin{cases} \geq c_1(L)2p(\tau - s) & \text{if } \tau \geq s \\ \leq c_1(L)2p(\tau - s) & \text{if } \tau \leq s \end{cases}$$

and the integral identity finally yields

$$|t|^p - |s|^p - p\,\mathrm{sgn}(s)|s|^{p-1}(t-s) \geq c_1(L)|t-s|^2 .$$

Analog conclusions for $s < 0$ also yield this inequality.

References

[1] Julien Bect, Laure Blanc-Féraud, Gilles Aubert, and Antoine Chambolle. A l^1-unified variational framework for image reconstruction. In Tomás Pajdla and Jiri Matas, editors, *Proceeding of the European Conference on Computer Vision*, volume 3024 of *Lecture Notes in Computer Science*, pages 1–13. Springer, 2004.

[2] Kristian Bredies, Dirk A. Lorenz, and Peter Maass. A generalized conditional gradient method and its connection to an iterative shrinkage method. Accepted for publication in *Computational Optimization and Applications*, 2005.

[3] Antonin Chambolle. An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision*, 20:89–97, 2004.

[4] Scott Shaobing Chen, David L. Donoho, and Michael A. Saunders. Atomic decomposition by basis pursuit. *SIAM Journal on Scientific Computing*, 20(1):33–61, 1998.

[5] Ingrid Daubechies. *Ten Lectures on Wavelets*, volume 61 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. SIAM (Society for Industrial and Applied Mathematics), 1992.

[6] Ingrid Daubechies, Michel Defrise, and Christine De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications in Pure and Applied Mathematics*, 57(11):1413–1457, 2004.

[7] David L. Donoho, Micheal Elad, and Vladimir Temlyakov. Stable recovery of sparse overcomplete representations in the presence of noise. *IEEE Transactions on Information Theory*, 52:6–18, 2006.

[8] Nelson Dunford and Jacob T. Schwartz. *Linear Operators Part 1: General Theory*. Interscience Publishers, Inc., New York, 1957.

[9] Joseph C. Dunn. Rates of convergence for conditional gradient algorithms near singular and nonsingular extremals. *SIAM Journal on Control and Optimization*, 17(2):187–211, March 1979.

[10] Joseph C. Dunn. Convergence rates for conditional gradient sequences generated by implicit step length rules. *SIAM Journal on Control and Optimization*, 18(5):473–487, September 1980.

[11] Ivar Ekeland and Roger Temam. *Convex analysis and variational problems*. North-Holland, Amsterdam, 1976.

[12] Mário A. T. Figueiredo and Robert D. Nowak. An EM algorithm for wavelet-based image restoration. *IEEE Transactions on Image Processing*, 12(8):906–916, 2003.

[13] Rémi Gribonval and Mads Nielsen. Sparse representation in unions of bases. *IEEE Transactions of Information Theory*, 49(12):33200–3325, December 2003.

[14] Ralph Edwin Showalter. *Monotone Operators in Banach Space and Nonlinear Partial Differential Equations*, volume 49 of *Mathematical Surveys and Monographs*. American Mathematical Society, 1997.

[15] Jean-Luc Starck, Mai K. Nguyen, and Fionn Murtagh. Wavelets and curvelets for image deconvolution: a combined approach. *Signal Processing*, 83:2279–2283, 2003.

[16] Gerd Teschke. Multi-frames in thresholding iterations for nonlinear operator equations with mixed sparsity constraints. Technical report, Preprint series of the DFG priority program 1114 "Mathematical methods for time series analysis and digital image processing", July 2005.

[17] Joel A. Tropp. Just relax: Convex programming methods for identifying sparse signals in noise. *IEEE Transactions on Information Theory*, 52(3):1030–1051, March 2006.

[18] Curtis R. Vogel and Mary E. Oman. Iterative methods for total variation denoising. *SIAM Journal on Scientific Computing*, 17:227–238, 1996.