

Numerical methods for sparse recovery

Massimo Fornasier

Abstract. These lecture notes are an introduction to methods recently developed for performing numerical optimizations with linear model constraints and additional sparsity conditions to solutions, i.e. we expect solutions which can be represented as sparse vectors with respect to a prescribed basis. Such a type of problems has been recently greatly popularized by the development of the field of nonadaptive compressed acquisition of data, the so-called *compressed sensing*, and its relationship with ℓ_1-minimization. We start our presentation by recalling the mathematical setting of compressed sensing as a reference framework for developing further generalizations. In particular we focus on the analysis of algorithms for such problems and their performances. We introduce and analyse the homotopy method, the iteratively reweighted least square method, and the iterative hard thresholding algorithm. We will see that the properties of convergence of these algorithms to solutions depends very much on special spectral properties (Restricted Isometry Property or Null Space Property) of the matrices which define the linear models. This provides a link to the courses of Holger Rauhut and Jared Tanner who will address in detail the analysis of such properties from different points of view. The concept of sparsity does not necessarily affect the entries of a vector only, but it can also be applied, for instance, to their variation. We will show that some of the algorithms proposed for compressed sensing are in fact useful for optimization problems with total variation constraints. Usually these optimizations on continuous domains are related to the calculus of variations on bounded variation (BV) functions and to geometric measure theory, which will be the objects of the course by Antonin Chambolle. In the second part of the lecture notes we address sparse optimizations in Hilbert spaces, and especially for situations when no Restricted Isometry Property or Null Space Property are assumed. We will be able to formulate efficient algorithms based on iterative soft-thresholding also for such situations, althought their analysis will require different tools, typically from nonsmooth convex analysis. The course by Ronny Ramlau, Gerd Teschke, and Mariya Zhariy addresses further developments of these algorithms towards regularization in nonlinear inverse problems as well as adaptive strategies. A common feature of the illustrated algorithms will be their variational nature, in the sense that they are derived as minimization strategies of given energy functionals. Not only the variational framework allows us to derive very precise statements about the convergence properties of these algorithms, but it also provides the algorithms with an intrinsic robustness. We will finally address large scale computations, showing how we can define domain decomposition strategies for these nonsmooth optimizations, for problems coming from compressed sensing and ℓ_1-minimization as well as for total variation minimization problems.
The first part of the lecture notes is elementary and it does not require more than the basic knowledge of notions of linear algebra and standard inequalities. The second part of the course is slightly more advanced and addresses problems in Hilbert spaces, and we will make use of more advanced concepts of nonsmooth convex analysis.

Key words. Numerical methods for sparse optimization, calculus of variations, algorithms for nonsmooth optimization.

AMS classification. 15A29, 65K10, 90C25, 52A41, 49M30,.

Acknowledgement

I'm very grateful to the colleagues, Antonin Chambolle, Ronny Ramlau, Holger Rauhut, Jared Tanner, Gerd Teschke, and Mariya Zhariy for their very generous contribution in accepting our invitation to prepare and present a course at RICAM for this Summer School. It is a honor to us to have all of them at our institute for such an occasion. Furthermore I would like to thank especially Wolfgang Forsthuber, Madgalena Fuchs, Andreas Langer, and Annette Weihs for the enormous help they provided to us for the organization of the Summer School. We acknowledge the financial support of RICAM, of the Doctoral Program in Computational Mathematics of the Johannes Kepler University of Linz, and the support by the project Y 432-N15 START-Preis "Sparse Approximation and Optimization in High Dimensions".

1 An Introduction to Sparse Recovery

1.1 Notations

In the following we collect general notations. More specific notations will be introduced and recalled in the following sections.
We will consider $\mathbb{R}^N$ as a Banach space endowed with different norms. In particular, later we use the ℓ_p-norms

$$\|x\|_p := \|x\|_{\ell_p} := \|x\|_{\ell_p^N} := \begin{cases} \left(\sum_{i=1}^N |x_j|^p\right)^{1/p}, & 0 < p < \infty, \\ \max_{j=1,\dots,N} |x_j|, & p = \infty. \end{cases} \tag{1.1}$$

Associated to these norms we denote their unit balls by $B_{\ell_p} := B_{\ell_p^N} := \{x \in \mathbb{R} : \|x\|_p \leq 1\}$ and the balls of radius R by $B_{\ell_p}(R) := B_{\ell_p^N}(R) := R \cdot B_{\ell_p^N}$. Associated to a closed convex body $0 \in \Omega \subset \mathbb{R}^N$, we define its polar set by $\Omega^\circ = \{y \in \mathbb{R}^N : \sup_{x \in \Omega} \langle x, y\rangle \leq 1\}$. This allow us to define an associated norm $\|x\|_\Omega = \sup_{y \in \Omega^\circ} \langle x, y \rangle$.
The index set $\mathcal{I}$ is supposed countable and we will consider the $\ell_p(\mathcal{I})$ spaces of p-summable sequences as well. Their norm are defined as usual and similarly to the case of $\mathbb{R}^N$. We use the same notations B_{ℓ_p} for the $\ell_p(\mathcal{I})$-balls as for the ones in $\mathbb{R}^N$. With A we will denote usually a $m \times N$ real matrix, $m, N \in \mathbb{N}$ or an operator $A : \ell_2(\mathcal{I}) \to Y$. We denote with A^* the adjoint matrix or with K^* the adjoint of an operator K. We will always work on real vector spaces, hence, in finite dimensions, A^* usually coincides with the transposed matrix of A. The norm of an operator $K : X \to Y$ acting between two Banach spaces is denoted by $\|K\|_{X \to Y}$; for matrices the norm $\|A\|$ denotes the spectral norm. The support of a vector $x \in \mathbb{R}^N$, i.e., the set of coordinates which are not zero, is denoted by $\operatorname{supp}(x)$.
We will consider index sets $\Lambda \subset \mathcal{I}$ and their complements $\Lambda^c = \mathcal{I} \setminus \Lambda$. The symbols $|\Lambda|$ and $\#\Lambda$ are used indifferently for indicating the cardinality of Λ. We use the notation A_Λ to indicate a submatrix extracted from A by retaining only the columns indexed in Λ as well as the restrictions u_Λ of vectors u to the index set Λ.
Positive constants used in estimates are denoted as usual by $c, C, \tilde{c}, \tilde{C}, c_0, C_0, c_1, C_1, c_2, C_2, \dots$.

1.2 A Toy Mathematical Model for Sparse Recovery

1.2.1 Adaptive and compressed acquisition

Let $k \in \mathbb{N}$, $k \leq N$ and

$$\Sigma_k := \{x \in \mathbb{R}^N : \#\operatorname{supp}(x) \leq k\},$$

is the set of vectors with at most k nonzero entries, which we will call *k-sparse vectors*. The *k-best approximation error* that we can achieve in this set to a vector $x \in \mathbb{R}^N$ with respect to a suitable space quasi-norm $\|\cdot\|_X$ is defined by given by

$$\sigma_k(x)_X = \inf_{z\in\Sigma_k} \|x-z\|_X.$$

Example 1.1 Let $r(x)$ be the nonincreasing rearrangement of x, i.e., $r(x) = (|x_{i_1}|, \dots, |x_{i_N}|)^T$ and $|x_{i_j}| \geq |x_{i_{j+1}}|$ for $j = 1, \dots, N-1$. Then it is straightforward to check that

$$\sigma_k(x)_{\ell_p^N} := \left(\sum_{j=k+1}^{N} r_j(x)^p \right)^{1/p}, \quad 1 \leq p < \infty.$$

In particular, the vector $x_{[k]}$ derived from x by setting to zero all the $N-k$ smallest entries in absolute value is called *the best k-term approximation* to x and it coincides with

$$x_{[k]} = \arg\min_{z\in\Sigma_k} \|x-z\|_{\ell_p^N}. \tag{1.2}$$

for any $1 \leq p < \infty$.

Lemma 1.2 *Let $r = \frac{1}{q} - \frac{1}{p}$ and $x \in \mathbb{R}^N$. Then*

$$\sigma_k(x)_{\ell_p} \leq \|x\|_{\ell_q} k^{-r}, \quad k = 1, 2, \dots, N.$$

Proof. Let Λ be the set of indexes of the k-largest entries of x in absolute value. If $\varepsilon = r_k(x)$, then

$$\varepsilon \leq \|x\|_{\ell_q} k^{-\frac{1}{q}}.$$

Therefore

$$\begin{aligned}\sigma_k(x)_{\ell_p}^p = \sum_{j\notin\Lambda} |x_j|^p &\leq \sum_{j\notin\Lambda} \varepsilon^{p-q} |x_j|^q \\ &\leq \|x\|_{\ell_q}^{p-q} k^{-\frac{p-q}{q}} \|x\|_{\ell_q}^q,\end{aligned}$$

which implies

$$\sigma_k(x)_{\ell_p} \leq \|x\|_{\ell_q} k^{-r}.$$

□

The computation the best k-term approximation of $x \in \mathbb{R}^N$, in general requires the search of the largest entries of x in absolute value, and therefore the testing of all the entries of the vector x. This procedure is *adaptive*, since it depends on the particular vector.

1.2.2 Nonadaptive and compressed acquisition: compressed sensing

One would like to describe a *linear encoder* which allows to compute approximatively k measurements $(y_1, \dots, y_k)^T$ and a nearly optimal approximation of x in the following sense:

> Provided a set $K \subset \mathbb{R}^N$, there exists a linear map $A : \mathbb{R}^N \to \mathbb{R}^m$, with $m \approx k$ and a possibly nonlinear map $\Delta : \mathbb{R}^m \to \mathbb{R}^N$ such that
>
> $$\|x - \Delta(Ax)\|_X \le C\sigma_k(x)_X$$
>
> for all $x \in K$.

Note that the way we encode $y = Ax$ is via a prescribed map A which is independent of x as well as the decoding procedure Δ might depend on A, but not on x. This is why we call this strategy a *nonadaptive (or universal) and compressed acquisition* of x. Note further that we would like to recover an approximation to x from nearly k-linear measurements which is of the order of the k-best approximation error. In this sense we say that the performances of the encoder/decoder system (A, Δ) is nearly optimal.

1.2.3 Optimal performances of encoder/decoder pairs

Let us define $\mathcal{A}_{m,N}$ the set of all encoder/decoder pairs (A, Δ) with A a $m \times N$ matrix and Δ any function. We wonder whether there exists such a nearly optimal pair as claimed above. Let us fix $m \le N$ two natural integers, and $K \subset \mathbb{R}^N$. For $1 \le k \le N$ we denote

$$\sigma_k(K)_X := \sup_{x \in K} \sigma_k(x)_X, \text{ and } E_m(K)_X := \inf_{(A,\Delta) \in \mathcal{A}_{m,N}} \sup_{x \in K} \|x - \Delta(A(x))\|_X.$$

We would like to find the largest k such that

$$E_m(K)_X \le C_0 \sigma_k(K)_X.$$

We respond to this question in the particular case where $K = B_{\ell_1}$ and $X = \ell_2^N$. This setting will turn out to be particularly useful later on and it is already sufficient for showing that, unfortunately, it is impossible to reconstruct $x \in B_{\ell_1}$ with an accuracy asymptotically (for m, N larger and larger) of the order of the k-best approximation error in ℓ_2^M if $k = m$, but it is necessary to have a slightly larger number of measurements, i.e., $k = m - \varepsilon(m, N)$.

The proper estimation of $E_m(K)_X$ turns out to be linked to the geometrical concept of *Gelfand width*.

Definition 1.3 Let K be a compact set in X. Then the *Gelfand width* of K of order m is

$$d^m(K)_X := \inf_{\substack{Y \subset X \\ \operatorname{codim}(Y) \leq m}} \sup\{\|x\|_X : x \in K \cap Y\}.$$

We have the following fundamental equivalence.

Proposition 1.4 *Let $K \subset \mathbb{R}^N$ any closed compact set for which $K = -K$ and such that there exists a constant $C_0 >$ for which $K + K \subset C_0 K$. If $X \subset \mathbb{R}^N$ is a normed space, then*

$$d^m(K)_X \leq E_m(K)_X \leq C_0 d^m(K)_X.$$

Proof. Let $\mathcal{N} = \ker A$. Note that $Y = \mathcal{N}$ has codimension less or equal to m. Conversely, given any space $Y \subset \mathbb{R}^N$ of codimension less or equal to m, we can associate a matrix A whose rows are a basis for $Y^\perp$, With this identification we see that

$$d^m(K)_X = \inf_{A \in \mathbb{R}^{m \times N}} \sup\{\|\eta\|_X : \eta \in \mathcal{N} \cap K\}.$$

If (A, Δ) is and encoder/decoder pair in $\mathcal{A}_{m,N}$ and $z = \Delta(0)$, then for any $\eta \in \mathcal{N}$ we have also $-\eta \in \mathcal{N}$. It follows that either $\|\eta - z\|_X \geq \|\eta\|_X$ or $\|-\eta - z\|_X \geq \|\eta\|_X$. Indeed, if we assumed that both are false then

$$\|2\eta\|_X = \|\eta - z + z + \eta\|_X \leq \|\eta - z\|_X + \|-\eta - z\|_X < 2\|\eta\|_X,$$

which is not possible. Since $K = -K$ we conclude that

$$\begin{aligned} d^m(K)_X &= \inf_{A \in \mathbb{R}^{m \times N}} \sup\{\|\eta\|_X : \eta \in \mathcal{N} \cap K\} \\ &\leq \sup_{\eta \in \mathcal{N} \cap K} \|\eta - z\|_X \\ &= \sup_{\eta \in \mathcal{N} \cap K} \|\eta - \Delta(A\eta)\|_X \\ &\leq \sup_{x \in K} \|\eta - \Delta(Ax)\|_X \end{aligned}$$

By taking the infimum over all $(A, \Delta) \in \mathcal{A}_{m,N}$ we obtain

$$d^m(K)_X \leq E_m(K)_X.$$

To prove the other inequality, choose an optimal Y such that

$$d^m(K)_X = \sup\{\|x\|_X : x \in Y \in \mathcal{N} \cap K\}.$$

Let us denote the solution affine space $\mathcal{F}(y) := \{x : Ax = y\}$. Let us define a decoder as follows: If $\mathcal{F}(y) \cap C \neq \emptyset$ then we take $\bar{x}(y) \in \mathcal{F}(y)$ and $\Delta(y) = \bar{x}(y)$. If $\mathcal{F}(y) \cap = \emptyset$ then $\Delta(y) \in \mathcal{F}(y)$. Hence, we can estimate

$$\begin{aligned} E_m(K)_X &= \inf_{(A,\Delta)\in\mathcal{A}_{m,N}} \sup_{x\in K} \|x - \Delta(Ax)\|_X \\ &\leq \sup_{x,x'\in\mathcal{F}(y)\cap K} \|x - x'\|_X \\ &\leq \sup_{\eta\in C_0(\mathcal{N}\cap K)} \|\eta\|_X \leq C_0 d^m(K)_X. \end{aligned}$$

□

The following result was proven in the relevant work of Kashin, Garnaev, and Gluskin [45, 46, 50] already in the '70s and '80s. See [20, 33] for a description of the relationship between this result and the more modern point of view related to compressed sensing.

Theorem 1.5 *The Gelfand width of ℓ_q^N-balls in ℓ_p^N for $1 \leq q < p \leq 2$ are estimated by*

$$C_1 \Psi(m, N, p, q) \leq d^m(B_{\ell_q})_{\ell_p} \leq C_2 \Psi(m, N, p, q),$$

where

$$\begin{aligned} \Psi(m, N, p, q)) &= \min\left\{1, N^{1-\frac{1}{q}} m^{-\frac{1}{2}}\right\}^{\frac{1/q-1/p}{1/q-1/2}}, \quad 1 < q < p \leq 2 \\ \Psi(m, N, 2, 1)) &= \min\left\{1, \sqrt{\tfrac{\log(N/m)+1}{m}}\right\}, \qquad q = 1 \textit{ and } p = 2. \end{aligned}$$

From Proposition 1.4 and Theorem 1.5 we obtain

$$\tilde{C}_1 \Psi(m, N, p, q) \leq E_m(B_{\ell_q})_{\ell_p} \leq \tilde{C}_2 \Psi(m, N, p, q).$$

In particular, for $q = 1$ and $p = 2$, we obtain, for m, N large enough, the estimate

$$\tilde{C}_1 \sqrt{\frac{\log(N/m) + 1}{m}} \leq E_m(B_{\ell_1})_{\ell_2}.$$

If we wanted to enforce

$$E_m(B_{\ell_1})_{\ell_2} \leq C \sigma_k(B_{\ell_1})_{\ell_2},$$

then Lemma 1.2 would imply

$$\sqrt{\frac{\log(N/m) + 1}{m}} \leq C_0 k^{-\frac{1}{2}}, \text{ or } k \leq C_0 \frac{m}{\log \frac{N}{m} + 1}.$$

Hence, we proved the following

Corollary 1.6 *For m, N fixed, there exists an optimal encoder/decoder pair $(A, \Delta) \in \mathcal{A}_{m,N}$, in the sense that*

$$E_m(B_{\ell_1})_{\ell_2} \le C\sigma_k(B_{\ell_1})_{\ell_2},$$

only if

$$k \le C_0 \frac{m}{\log \frac{N}{m} + 1}, \tag{1.3}$$

for some constant $C_0 > 0$ independent of m, N.

The next section is devoted to the the construction of optimal encoder/decoder pairs $(A, \Delta) \in \mathcal{A}_{m,N}$ as stated in the latter corollary.

1.3 Survey on Mathematical Analysis of Compressed Sensing

In following section we want to show that under a certain property, called the *Restricted Isometry Property* (RIP) for a matrix A,

> The decoder, which we call ℓ_1-*minimization*,
>
> $$\Delta(y) = \arg \min_{Az=y=Ax} \|x\|_{\ell_1^N} \tag{1.4}$$
>
> performs
>
> $$\|x - \Delta(y)\|_{\ell_1^N} \le C_1 \sigma_k(x)_{\ell_1^N}, \tag{1.5}$$
>
> as well as
>
> $$\|x - \Delta(y)\|_{\ell_2^N} \le C_2 \frac{\sigma_k(x)_{\ell_1^N}}{k^{1/2}}, \tag{1.6}$$
>
> for all $x \in \mathbb{R}^N$.

Note that by (1.6) we immediately obtain

$$E_m(B_{\ell_1^N})_{\ell_2^N} \le C_0 k^{-1/2},$$

implying once again (1.3). Hence, the following question we will address is the existence of matrices A with RIP for which k is optimal, i.e.,

$$k \asymp \frac{m}{\log N/m + 1}.$$

1.3.1 An intuition why ℓ_1-minimization works well

In this section we would like to provide an intuitive explanation of the stable results (1.5) and (1.6) provided by ℓ_1-minimization (1.4) in recovering vectors from partial linear measurements. Equations (1.5) and (1.6) ensure in particular that if the vector x is k-*sparse*, then ℓ_1-minimization (1.4) will be able to recover it *exactly* from m linear measurements y obtained via the matrix A. This result is quite surprising because the problem of recovering a sparse vector, or the solution of the following optimization

$$\min \|x\|_{\ell_0^N} \text{ subject to } Ax = y, \tag{1.7}$$

is know [56, 58] to be *NP-complete*[1] whereas ℓ_1-minimization is a convex problem which can be solved at any prescribed accuracy in polynomial time. For instance interior-point methods are guaranteed to solve the problem to a fixed precision in time $\mathcal{O}(m^2 N^{1.5})$ [60]. The first intuitive approach to this surprising result is by interpreting ℓ_1-minimization as the *convexification* of the problem (1.7).

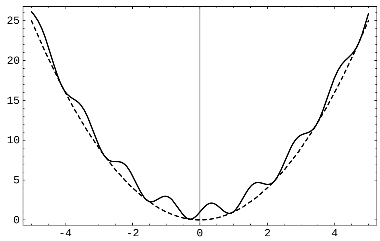


Figure 1.1 A non convex function f and a convex approximation $g \leq f$ from below.

If we were interested to solve an optimization problem

$$\min f(x) \text{ subject to } x \in \mathcal{C},$$

where f is a nonconvex function and $\mathcal{C}$ is a closed convex set, it might be convenient to recast the problem by considering its convexification, i.e.,

$$\min \bar{f}(x) \text{ subject to } x \in \mathcal{C},$$

where $\bar{f}$ is called the *convex relaxation* or the *convex envelop* of f and it is given by

$$\bar{f}(x) := \sup\{g(x) \leq f(x) : g \text{ is a convex function}\}.$$

The motivation of this choice is simply geometrical. While f can have many minimizers on $\mathcal{C}$, its convex envelop $\bar{f}$ has global minimizers (but not strictly local ones), and such global minima are likely to be in a neighborhood of a global minimum of f, see Figure 1.1. One rewrites

$$\|x\|_{\ell_0^N} := \sum_{j=1}^N |x_j|_0, \quad |t|_0 := \begin{cases} 0, & t = 0 \\ 1, & 0 < t \leq 1 \end{cases}.$$

[1] In general its resolution has a complexity which is growing with a rate faster than any polynomial, for instance exponentially, in the dimension m, N of the problem.

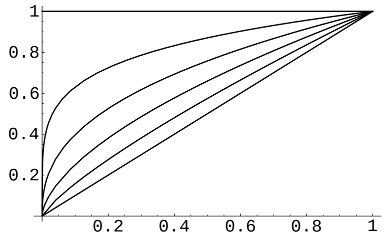


Figure 1.2 The absolute value function $|\cdot|$ is the convex relaxation of the function $|\cdot|_0$ on $[0,1]$.

Its convex envelope in $B_{\ell_\infty^N}(R) \cap \{z : Az = y\}$ is bounded below by $\frac{1}{R}\|x\|_{\ell_1^N} := \frac{1}{R}\sum_{j=1}^N |x_j|$, see Figure 1.2. This observation gives already a first impression of the motivation why ℓ_1-minimization can help in approximating sparse solutions of $Ax = y$. However, it is not yet clear when a global minimizer of

$$\min \|x\|_{\ell_1^N} \text{ subject to } Ax = y, \tag{1.8}$$

does really coincide with a solution to (1.7). Again a simple geometrical reasoning can help us to get a feeling about more general principles which will be addressed more formally in the following sections.

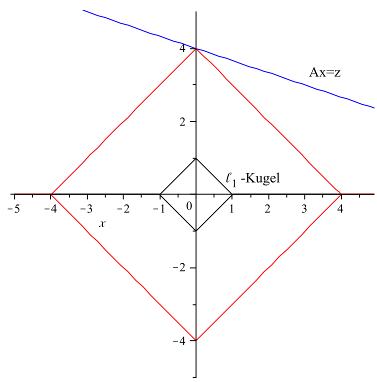


Figure 1.3 The ℓ_1-minimizer within the affine space of solutions of the linear system $Ax = y$ coincides with the sparsest solution.

Assume for a moment that $N = 2$ and $m = 1$. Hence we are dealing with an affine space of solutions $\mathcal{F}(y) = \{z : Az = y\}$ which is just a line in $\mathbb{R}^2$. When we search for the ℓ_1- norm minimizers among the elements $\mathcal{F}(y)$ (see Figure 1.3), we immediately realize that, except for pathological situations where $\mathcal{N} = \ker A$ is parallel to one of the faces of the polytope $B_{\ell_1^2}$, there is a unique solution which coincides also with a solution with minimal number of nonzero entries. Therefore, if we exclude situations in which there exists $\eta \in \mathcal{N}$ such that $|\eta_1| = |\eta_2|$ or, equivalently, we assume that

$$|\eta_i| < |\eta_{\{1,2\}\setminus\{i\}}| \tag{1.9}$$

for all $\eta \in \mathcal{N}$ and for one $i = 1, 2$, then the solution to (1.8) is a solution to (1.7)! Note also that, if we give a uniform probability distribution to the angle in $[0, 2\pi]$ formed by $\mathcal{N}$ and any of the coordinate axes, then we realize that the pathological situation of violating (1.9) has null probability. Of course, in higher dimension such simple reasoning becomes more involved, since the number of faces and edges of an ℓ_1^N-ball $B_{\ell_1^N}$ becomes larger larger and one should cumulate the probabilities of different angles with respect to possible affine spaces of codimension $N-m$. However, condition (1.9) is the right prototype of a property (we call it the *Null Space Property* (NSP) and we describe it in detail in the next section) which guarantees, also in higher dimension, that the solution to (1.8) is a solution to (1.7).

1.3.2 Restricted Isometry Property and Null Space Property

Definition 1.7 One says that $A \in \mathbb{R}^{m \times N}$ has the *Null Space Property* (NSP) of order k for $0 < \gamma < 1$ if

$$\|\eta_\Lambda\|_{\ell_1^N} \leq \gamma \|\eta_{\Lambda^c}\|_{\ell_1^N},$$

for all sets $\Lambda \subset \{1, \dots, N\}$, $\#\Lambda \leq k$ and for all $\eta \in \mathcal{N} = \ker A$.

Note that this definition greatly generalizes condition (1.9) which we introduced by our simple and rough geometrical reasoning in $\mathbb{R}^2$. However, let us stress that in order to address stability properties such as (1.5) and (1.6) (and not only the exact recovery of sparse vectors), it will not be sufficient to require $\|\eta_\Lambda\|_{\ell_1^N} < \|\eta_{\Lambda^c}\|_{\ell_1^N}$, but also a gap $\|\eta_\Lambda\|_{\ell_1^N} \leq \gamma\|\eta_{\Lambda^c}\|_{\ell_1^N}$ provided by the introduction of a constant $\gamma < 1$ is fundamental. We need further to introduce a related property for matrices.

Definition 1.8 One says that $A \in \mathbb{R}^{m \times N}$ has the RIP of order K if there exists $0 < \delta_K < 1$ such that

$$(1 - \delta_K)\|z\|_{\ell_2^N} \leq \|Az\|_{\ell_2^N} \leq (1 + \delta_K)\|z\|_{\ell_2^M},$$

for all $z \in \Sigma_K$.

The RIP turns out to be very useful in the analysis of stability of certain algorithms as we will show in Section 2.1.4. The RIP is also introduced because it implies the *Null Space Property*, and when dealing with random matrices (see Section 1.3.4) it is more easily addressed. Indeed we have:

Lemma 1.9 *Assume that $A \in \mathbb{R}^{m \times N}$ has the RIP of order $K = k + h$ with $0 < \delta_K < 1$. Then A has the NSP of order k and constant $\gamma = \sqrt{\frac{k}{h}\frac{1+\delta_K}{1-\delta_K}}$.*

Proof. Let $\Lambda \subset \{1,\dots,N\}$, $\#\Lambda \le k$. Define $\Lambda_0 = \Lambda$ and $\Lambda_1, \Lambda_2, \dots, \Lambda_s$ disjoint sets of indexes of size at most h, associated to a decreasing rearrangement of the entries of $\eta \in \mathcal{N}$. Then, by using Cauchy-Schwarz inequality, the RIP twice, the fact that $A\eta = 0$, and eventually the triangle inequality, we have the following sequence of inequalities:

$$\begin{aligned}
\|\eta_\Lambda\|_{\ell_1^N} &\le \sqrt{k}\|\eta_\Lambda\|_{\ell_2^N} \le \sqrt{k}\|\eta_{\Lambda_0\cup\Lambda_1}\|_{\ell_2^N} \\
&\le (1-\delta_K)^{-1}\sqrt{k}\|A\eta_{\Lambda_0\cup\Lambda_1}\|_{\ell_2^N} = (1-\delta_K)^{-1}\sqrt{k}\|A\eta_{\Lambda_2\cup\Lambda_3\cup\cdots\cup\Lambda_s}\|_{\ell_2^N} \\
&\le (1-\delta_K)^{-1}\sqrt{k}\sum_{j=2}^{s}\|A\eta_{\Lambda_j}\|_{\ell_2^N} \le \frac{1+\delta_K}{1-\delta_K}\sqrt{k}\sum_{j=2}^{s}\|\eta_{\Lambda_j}\|_{\ell_2^N}. \qquad (1.10)
\end{aligned}$$

Note now that $i \in \Lambda_{j+1}$ and $\ell \in \Lambda_j$ imply by construction of $\Lambda_j's$ by decreasing rearrangement of the entries of η

$$|\eta_i| \le |\eta_\ell|.$$

By taking the sum over ℓ first and than the ℓ_2^N-norm over i we get

$$|\eta_i| \le h^{-1}\|\eta_{\Lambda_j}\|_{\ell_1^N}, \text{ and } \|\eta_{\Lambda_{j+1}}\|_{\ell_2^N} \le h^{-1/2}\|\eta_{\Lambda_j}\|_{\ell_1^N}.$$

By using the latter estimates in (1.10) we obtain

$$\|\eta_\Lambda\|_{\ell_1^N} \le \frac{1+\delta_K}{1-\delta_K}\sqrt{\frac{k}{h}}\sum_{j=1}^{s-1}\|\eta_{\Lambda_j}\|_{\ell_1^N} \le \left(\frac{1+\delta_K}{1-\delta_K}\sqrt{\frac{k}{h}}\right)\|\eta_{\Lambda^c}\|_{\ell_1^N}.$$

□

The RIP property does imply the NSP, but the converse is not true. Actually the RIP is significantly more restrictive. A very detailed discussion on the limitations provided by the RIP will be the object of the course by Jared Tanner.

1.3.3 Performances of ℓ_1-minimization as an optimal decoder

In this section we address the proofs of the approximation properties (1.5) and (1.6).

Theorem 1.10 *Let $A \in \mathbb{R}^{m\times N}$ which satisfies the RIP of order $2k$ with $\delta_{2k} \le \delta < \frac{\sqrt{2}-1}{\sqrt{2}+1}$ (or simply A satisfies the NSP of order k with constant $\gamma = \frac{1+\delta}{1-\delta}\sqrt{\frac{1}{2}}$) , then the decoder Δ as in* (1.4) *satisfies* (1.5).

Proof. By Lemma 1.9 we have

$$\|\eta_\Lambda\|_{\ell_1} \le \frac{1+\delta}{1-\delta}\sqrt{\frac{1}{2}}\|\eta_{\Lambda^c}\|_{\ell_1^N},$$

for all $\Lambda \subset \{1, \ldots, N\}$, $\#\Lambda \leq k$ and $\eta \in \mathcal{N} = \ker A$. Let $x^* = \Delta(Ax)$, so that $\eta = x^* - x \in \mathcal{N}$, and

$$\|x^*\|_{\ell_1^N} \leq \|x\|_{\ell_1^N}.$$

One denotes now with Λ the set of the k-largest entries of x in absolute value. One has

$$\|x^*_\Lambda\|_{\ell_1^N} + \|x^*_{\Lambda^c}\|_{\ell_1^N} \leq \|x_\Lambda\|_{\ell_1^N} + \|x_{\Lambda^c}\|_{\ell_1^N}.$$

It follows immediately by triangle inequality

$$\|x_\Lambda\|_{\ell_1^N} - \|\eta_\Lambda\|_{\ell_1^N} + \|\eta_{\Lambda^c}\|_{\ell_1^N} - \|x_{\Lambda^c}\|_{\ell_1^N} \leq \|x_\Lambda\|_{\ell_1^N} + \|x_{\Lambda^c}\|_{\ell_1^N}.$$

Hence

$$\|\eta_{\Lambda^c}\|_{\ell_1^N} \leq \|\eta_\Lambda\|_{\ell_1^N} + 2\|x_{\Lambda^c}\|_{\ell_1^N} \leq \frac{1+\delta}{1-\delta}\sqrt{\frac{1}{2}}\|\eta_{\Lambda^c}\|_{\ell_1^N} + 2\sigma_k(x)_{\ell_1},$$

or, equivalently,

$$\|\eta_{\Lambda^c}\|_{\ell_1^N} \leq \frac{2}{1 - \frac{1+\delta}{1-\delta}\sqrt{\frac{1}{2}}}\sigma_k(x)_{\ell_1}. \tag{1.11}$$

In particular, note that by $\delta < \frac{\sqrt{2}-1}{\sqrt{2}+1}$ we have $\frac{1+\delta}{1-\delta}\sqrt{\frac{1}{2}} < 1$. Eventually we conclude with the estimates

$$\begin{aligned} \|x - x^*\|_{\ell_1^N} &= \|\eta_\Lambda\|_{\ell_1^N} + \|\eta_{\Lambda^c}\|_{\ell_1^N} \\ &\leq \left(\frac{1+\delta}{1-\delta}\sqrt{\frac{1}{2}} + 1\right)\|\eta_{\Lambda^c}\|_{\ell_1^N} \\ &\leq C_1\sigma_k(x)_{\ell_1}, \end{aligned}$$

where $C_1 := \left[\frac{2\left(\frac{1+\delta}{1-\delta}\sqrt{\frac{1}{2}}+1\right)}{1-\frac{1+\delta}{1-\delta}\sqrt{\frac{1}{2}}}\right]$. □

Similarly we address the second estimate (1.6).

Theorem 1.11 *Let $A \in \mathbb{R}^{m\times N}$ which satisfies the RIP of order $3k$ with $\delta_{3k} \leq \delta < \frac{\sqrt{2}-1}{\sqrt{2}+1}$, then the decoder Δ as in* (1.4) *satisfies* (1.6).

Proof. Let $x^* = \Delta(Ax)$. As we proceeded in Lemma 1.9, we denote $\eta = x^* - x \in \mathcal{N}$, $\Lambda_0 = \Lambda$ the set of the $2k$-largest entries of η in absolute value, and Λ_j of size at most k composed of nonincreasing rearrangement entries. Then

$$\|\eta_\Lambda\|_{\ell_2^N} \leq \frac{1+\delta}{1-\delta}k^{-\frac{1}{2}}\|\eta_{\Lambda^c}\|_{\ell_1^N},$$

Note now that by Lemma 1.2 and by Lemma 1.9

$$\begin{aligned}\|\eta_{\Lambda^c}\|_{\ell_2^N} &\leq (2k)^{-\frac{1}{2}}\|\eta\|_{\ell_1^N} = (2k)^{-1/2}\left(\|\eta_\Lambda\|_{\ell_1^N} + \|\eta_{\Lambda^c}\|_{\ell_1^N}\right)\\ &\leq (2k)^{-1/2}\left(C\|\eta_{\Lambda^c}\|_{\ell_1^N} + \|\eta_{\Lambda^c}\|_{\ell_1^N}\right)\\ &\leq \frac{C+1}{\sqrt{2}}k^{-1/2}\|\eta_{\Lambda^c}\|_{\ell_1^N},\end{aligned}$$

for a suitable constant $C > 0$. Note that, being Λ set of the $2k$-largest entries of η in absolute value, one has also

$$\|\eta_{\Lambda^c}\|_1 \leq \|\eta_{(\operatorname{supp} x_{[2k]})^c}\|_1 \leq \|\eta_{(\operatorname{supp} x_{[k]})^c}\|_1, \tag{1.12}$$

where $x_{[k]}$ is the best k-term approximation to x. The use of this latter estimate, combined with the inequality (1.11) finally gives

$$\begin{aligned}\|x - x^*\|_{\ell_2^N} &= \|\eta_\Lambda\|_{\ell_2^N} + \|\eta_{\Lambda^c}\|_{\ell_2^N}\\ &\leq Ck^{-1/2}\|\eta_{\Lambda^c}\|_{\ell_1^N}\\ &\leq \tilde{C}_2 k^{-1/2}\sigma_k(x)_{\ell_1}.\end{aligned}$$

□

We would like to conclude this section by mentioning a further stability property of ℓ_1-minimization as established in [12].

Theorem 1.12 *Let $A \in \mathbb{R}^{m\times N}$ which satisfies the RIP of order $4k$ with δ_{4k} sufficiently small. Assume further that $Ax + e = y$ where e is a measurement error. Then the decoder Δ as the further enhanced stability property:*

$$\|x - \Delta(y)\|_{\ell_2^N} \leq C_3\left(\sigma_k(x)_{\ell_2^N} + \frac{\sigma_k(x)_{\ell_1^N}}{k^{1/2}} + \|e\|_{\ell_2^N}\right). \tag{1.13}$$

1.3.4 Random matrices and optimal RIP

In this section we would like to mention how for different classes of random matrices it is possible to show that the RIP property can hold with optimal constants, i.e.,

$$k \asymp \frac{m}{\log N/m + 1}.$$

at least with high probability. This implies in particular, that such matrices exist, they are frequent, but they are given to us only with an uncertainty.

Gaussian and Bernoulli random matrices

Let (Ω, ρ) be a probability measure space and X a random variable on (Ω, ρ). One can define a random matrix $A(\omega)$, $\omega \in \Omega^{mN}$, as the matrix whose entries are independent realizations of X. We assume further that $\|A(\omega)x\|_{\ell_2^N}^2$ has expected value $\|x\|_{\ell_2^N}^2$ and

$$\mathbb{P}\left(\left|\|A(\omega)x\|_{\ell_2^N}^2 - \|x\|_{\ell_2^N}^2\right| \geq \varepsilon\|x\|_{\ell_2^N}^2\right) \leq 2e^{-mc_0(\varepsilon)}, \quad 0 < \varepsilon < 1. \tag{1.14}$$

Example 1.13 Here we collect two of the most relevant examples for which the concentration property (1.14) holds:

1. One can choose, for instance, the entries of A as i.i.d. Gaussian random variables, $A_{ij} \sim \mathcal{N}(0, \frac{1}{m})$, and $c_0(\varepsilon) = \varepsilon^2/4 - \varepsilon^3/6$. This can be shown by using Chernoff inequalities and a comparison of the moments of a Bernoulli random variable to those of a Gaussian random variable;

2. One can also use matrices where the entries are independent realizations of ± 1 Bernoulli random variables

$$A_{ij} = \begin{cases} +1/\sqrt{m}, & \text{with probability } \frac{1}{2} \\ -1/\sqrt{m}, & \text{with probability } \frac{1}{2} \end{cases}.$$

Then we have the following result, shown, for instance in [3].

Theorem 1.14 *Suppose that* m, N *and* $0 < \delta < 1$ *are fixed. If* $A(\omega), \omega \in \Omega^{mN}$ *is a random matrix of size* $m \times N$ *with the concentration property* (1.14)*, then there exist constants* $c_1, c_2 > 0$ *depending on* δ *such that the RIP holds for* $A(\omega)$ *with constant* δ *and* $k \leq c_1 \frac{m}{\log(N/m)+1}$ *with probability exceeding* $1 - 2e^{-c_2 m}$.

An extensive discussion on RIP properties of different matrices, for instance partial Fourier matrices or structured matrices, will be provided in the course by Holger Rauhut.

2 Numerical Methods for Compressed Sensing

The previous sections showed that ℓ_1-minimization performs very well in recovering sparse or approximately sparse vectors from undersampled measurements. In applications it is important to have fast methods for actually solving ℓ_1-minimization or to have similar guarantees of stability. Two such methods – the homotopy (LARS) method introduced by [35, 64], the iteratively reweighted least square method (IRLS) [30], and the iterative hard thresholding algorithm [6, 7] – will be explained in more detail below.

As a first remark, the ℓ_1-minimization problem

$$\min \|x\|_1 \quad \text{subject to } Ax = y \tag{2.15}$$

is in the real case equivalent to the linear program

$$\min \sum_{j=1}^{N} v_j \quad \text{subject to} \quad v \geq 0,\ (A| - A)v = y. \tag{2.16}$$

The solution x^* to (2.15) is obtained from the solution v^* of (2.16) via $x^* = (I| - I)v^*$. Any linear programming method may therefore be used for solving (2.15). The simplex method as well as interior point methods apply in particular [60], and standard software may be used. (In the complex case, (2.15) is equivalent to a second order cone program (SOCP) and can be solved with interior point methods as well.) However, such methods and software are of general purpose and one may expect that methods specialized to (2.15) outperform such existing standard methods. Moreover, standard software often has the drawback that one has to provide the full matrix rather than fast routines for matrix-vector multiplication which are available for instance in the case of partial Fourier matrices. In order to obtain the full performance of such methods one would therefore need to re-implement them, which is a daunting task because interior point methods usually require much fine tuning. On the contrary the two specialized methods described below are rather simple to implement and very efficient. Many more methods are available nowadays, including greedy methods, such as Orthogonal Matching Pursuit [74] and CoSaMP [73]. However, only the three methods below are explained in detail because they highlight the fundamental concepts which are useful to comprehend also other algorithms.

2.1 Direct and Iterative Methods

2.1.1 The Homotopy Method

The homotopy method – or modified LARS – [34, 35, 62, 64] solves (2.15) in the real-valued case. One considers the ℓ_1-regularized least squares functionals

$$F_\lambda(x) = \frac{1}{2}\|Ax - y\|_2^2 + \lambda\|x\|_1, \quad x \in \mathbb{R}^N,\ \lambda > 0, \tag{2.17}$$

and its minimizer x_λ. When $\lambda = \hat{\lambda}$ is large enough then $x_{\hat{\lambda}} = 0$, and furthermore, $\lim_{\lambda \to 0} x_\lambda = x^*$, where x^* is the solution to (2.15). The idea of the homotopy method is to trace the solution x_λ from $x_{\hat{\lambda}} = 0$ to x^*. The crucial observation is that the solution path $\lambda \mapsto x_\lambda$ is piecewise linear, and it is enough to trace the endpoints of the linear pieces.

The minimizer of (2.17) can be characterized using the subdifferential, which is defined for a general convex function $F : \mathbb{R}^N \to \mathbb{R}$ at a point $x \in \mathbb{R}^N$ by

$$\partial F(x) = \{v \in \mathbb{R}^N, F(y) - F(x) \geq \langle v, y - x\rangle \ \text{ for all } y \in \mathbb{R}^N\}.$$

Clearly, x is a minimizer of F if and only if $0 \in \partial F(x)$. The subdifferential of F_λ is given by

$$\partial F_\lambda(x) = A^*(y - Ax) + \lambda \partial \|x\|_1$$

where the subdifferential of the ℓ_1-norm is given by

$$\partial \|x\|_1 = \{v \in \mathbb{R}^N : v_\ell \in \partial |x_\ell|, \ell = 1, \dots, N\}$$

with the subdifferential of the absolute value being

$$\partial |z| = \begin{cases} \{\operatorname{sgn}(z)\}, & \text{if } z \neq 0, \\ [-1, 1] & \text{if } z = 0. \end{cases}$$

The inclusion $0 \in \partial F_\lambda(x)$ is equivalent to

$$(A^*(y - Ax))_\ell = \lambda \operatorname{sgn}(x_\ell) \qquad \text{if } x_\ell \neq 0, \tag{2.18}$$

$$|(A^*(y - Ax)_\ell| \leq \lambda \qquad \text{if } x_\ell = 0, \tag{2.19}$$

for all $\ell = 1, \dots, N$.

As already mentioned above the homotopy method starts with $x^{(0)} = x_\lambda = 0$. By conditions (2.18) and (2.19) the corresponding λ can be chosen as $\lambda = \lambda^{(0)} = \|A^* y\|_\infty$. In the further steps $j = 1, 2, \dots$ the algorithm computes minimizers $x^{(1)}, x^{(2)}, \dots$ and maintains an active (support) set T_j. Denote by

$$c^{(j)} = A^*(y - Ax^{(j-1)})$$

the current residual vector. The columns of the matrix A are denoted by $a_\ell, \ell = 1, \dots, N$ and for a subset $T \subset \{1, \dots, N\}$ we let A_T be the submatrix of A corresponding to the columns indexed by T.

Step 1: Let

$$\ell^{(1)} := \arg \max_{\ell = 1, \dots, N} |(A^* y)_\ell| = \arg \max_{\ell = 1, \dots, N} |c_\ell^{(1)}|.$$

One assumes here and also in the further steps that the maximum is attained at only one index ℓ. The case that the maximum is attained simultaneously at two or more indeces ℓ (which almost never happens) requires more complications that we would like to avoid here. One may refer to [35] for such details.

Now set $T_1 = \{\ell^+\}$. The vector $d \in \mathbb{R}^N$ describing the direction of the solution (homotopy) path has components

$$d_{\ell^{(1)}}^{(1)} = \|a_\ell\|_2^{-2} \operatorname{sgn}((Ay)_\ell), \quad d_\ell^{(1)} = 0, \ell \neq \ell^+.$$

The first linear piece of the solution path then takes the form

$$x = x(\gamma) = x^{(0)} + \gamma d^{(1)} = \gamma d^{(1)}, \quad \gamma \in [0, \gamma^{(1)}].$$

One verifies with the definition of $d^{(1)}$ that (2.18) is always satisfied for $x = x(\gamma)$ and $\lambda = \lambda(\gamma) = \lambda^{(0)} - \gamma$, $\gamma \in [0, \lambda^{(0)}]$. The next breakpoint is found by determining the maximal $\gamma = \gamma^{(1)} > 0$ for which (2.19) is satisfied, which is

$$\gamma^{(1)} = \min_{\ell \neq \ell^{(1)}} \left\{ \frac{\lambda^{(0)} - c_\ell^{(1)}}{1 - (A^* A\, d^{(1)})_\ell}, \frac{\lambda^{(0)} + c_\ell^{(1)}}{1 + (A^* A\, d^{(1)})} \right\}, \tag{2.20}$$

where the minimum is taken only over positive arguments. Then $x^{(1)} = x(\gamma^{(1)}) = \gamma^{(1)} d^{(1)}$ is the next minimizer of F_λ for $\lambda = \lambda^{(1)} := \lambda^{(0)} - \gamma^{(1)}$. This $\lambda^{(1)}$ satisfies $\lambda^{(1)} = \|c^{(1)}\|_\infty$. Let $\ell^{(2)}$ be the index where the minimum in (2.20) is attained (where we again assume that the minimum is attained only at one index) and put $T_2 = \{\ell^{(1)}, \ell^{(2)}\}$.

Step j: Determine the new direction $d^{(j)}$ of the homotopy path by solving

$$A_{T_j}^* A_{T_j} d_{T_j}^{(j)} = \operatorname{sgn}(c_{T_j}^{(j)}), \tag{2.21}$$

which is a linear system of equations of size at most $|T_j| \times |T_j|$. Outside the components in T_j one sets $d_\ell^{(j)} = 0$, $\ell \notin T_j$. The next piece of the path is then given by

$$x(\gamma) = x^{(j-1)} + \gamma d^{(j)}, \quad \gamma \in [0, \gamma^{(j)}].$$

The maximal γ such that $x(\gamma)$ satisfies (2.19) is

$$\gamma_+^{(j)} = \min_{\ell \notin T_j} \left\{ \frac{\lambda^{(j-1)} - c_\ell^{(j)}}{1 - (A^* A d^{(j)})_\ell}, \frac{\lambda^{(j-1)} + c_\ell^{(j)}}{1 + (A^* A d^{(j)})_\ell} \right\}. \tag{2.22}$$

The maximal γ such that $x(\gamma)$ satisfies (2.18) is determined as

$$\gamma_-^{(j)} = \min_{\ell \in T_j} \{ -x_\ell^{(j-1)} / d_\ell^{(j)} \}. \tag{2.23}$$

Both in (2.22) and (2.23) the minimum is taken only over positive arguments. The next breakpoint is given by $x^{(j+1)} = x(\gamma^{(j)})$ with $\gamma^{(j)} = \min\{\gamma_+^{(j)}, \gamma_-^{(j)}\}$. If $\gamma_+^{(j)}$ determines the minimum then the index $\ell_+^{(j)} \notin T_j$ providing the minimum in (2.22) is added to the active set, $T_{j+1} = T_j \cup \{\ell_+^{(j)}\}$. If $\gamma^{(j)} = \gamma_-^{(j)}$ then the index $\ell_-^{(j)} \in T_j$ is removed from the active set, $T_{j+1} = T_j \setminus \{\ell_-^{(j)}\}$. Further, one updates $\lambda^{(j)} = \lambda^{(j-1)} - \gamma^{(j)}$. By construction $\lambda^{(j)} = \|c^{(j)}\|_\infty$.

The algorithm stops when $\lambda^{(j)} = \|c^{(j)}\|_\infty = 0$, i.e., when the residual vanishes, and output $x^* = x^{(j)}$. Indeed, this happens after a finite number of steps. [35] proved the following result.

Theorem 2.1 *If in each step the minimum in* (2.22) *and* (2.23) *is attained in only one index ℓ, then the homotopy algorithm as described yields the minimizer of the ℓ_1-minimization problem* (2.15).

If the algorithm is stopped earlier at some iteration j then obviously it yields the minimizer of $F_\lambda = F_{\lambda^{(j)}}$. In particular, obvious stopping rules may also be used to solve the problems

$$\min \|x\|_1 \quad \text{subject to } \|Ax - y\|_2 \leq \epsilon \tag{2.24}$$

$$\text{and} \quad \min \|Ax - y\|_2 \quad \text{subject to } \|x\|_1 \leq \delta. \tag{2.25}$$

The second of these is called the LASSO [72].

The LARS (least angle regression) algorithm is a simple modification of the homotopy method, which only adds elements to the active set in each step. So $\gamma_-^{(j)}$ in (2.23) is not considered. (Sometimes the homotopy method is therefore also called modified LARS.) Clearly, LARS is not guaranteed any more to yield the solution of (2.15). However, it is observed empirically – and can be proven rigorously in certain cases [34] – that often in sparse recovery problems, the homotopy method does never remove elements from the active set, so that in this case LARS and homotopy perform the same steps. It is a crucial point that if the solution of (2.15) is k-sparse and the homotopy method never removes elements then the solution is obtained after precisely k-steps. Furthermore, the most demanding computational part at step j is then the solution of the $j \times j$ linear system of equations (2.21). In conclusion, the homotopy and LARS methods are very efficient for sparse recovery problems.

2.1.2 Iteratively Reweighted Least Squares

In this section we want to present an iterative algorithm which, under the condition that A satisfies the NSP, is guaranteed to reconstruct vectors with the same approximation guarantees (1.5) as ℓ_1-minimization. Moreover, we will also show that such algorithm has a guaranteed linear rate of convergence which, with a minimal modification, can be improved to a superlinear rate. We need to make first a brief introduction which hopefully will shed light on the basic principles of this algorithm and their interplay with sparse recovery and ℓ_1-minimization.

Denote $\mathcal{F}(y) = \{x : Ax = y\}$ and $\mathcal{N} = \ker A$. Let us start with a few non-rigorous observations; next we will be more precise. For $t \neq 0$ we simply have

$$|t| = \frac{t^2}{|t|}.$$

Hence, an ℓ_1-minimization can be recasted into a weighted ℓ_2-minimization, and we may expect

$$\arg \min_{x \in \mathcal{F}(y)} \sum_{j=1}^N |x_j| \approx \arg \min_{x \in \mathcal{F}(y)} \sum_{j=1}^N x_j^2 |x_j^*|^{-1},$$

as soon as x^* is the wanted ℓ_1-norm minimizer. Clearly the advantage is that minimizing a smooth quadratic function $|t|^2$ is better than addressing the minimization of the nonsmooth function $|t|$. However, the obvious drawbacks are that neither we dispose of x^* a priori (this is the vector we are interested to compute!) nor we can expect that $x_j^* \neq 0$ for all $i = 1, \dots, N$, since we hope for k-sparse solutions. Hence, we start assuming that we dispose of a good approximation w_j^n of $|(x_j^*)^2 + \epsilon_n^2|^{-1/2} \approx |x_j^*|^{-1}$ and we compute

$$x^{n+1} = \arg \min_{x \in \mathcal{F}(y)} \sum_{j=1}^N x_j^2 w_j^n, \tag{2.26}$$

then we up-date $\epsilon_{n+1} \leq \epsilon_n$, we define

$$w_j^{n+1} = |(x_j^n)^2 + \epsilon_{n+1}^2|^{-1/2}, \tag{2.27}$$

and we iterate the process. The hope is that a proper choice of $\epsilon_n \to 0$ will allow for the computation of an ℓ_1-minimizer, althought such limit property is far from being obvious. The next sections will help us to describe the right mathematical setting where such limit is justified.

The relationship between ℓ_1-minimization and reweighted ℓ_2-minimization

Lemma 2.2 *An element $x^* \in \mathcal{F}(y)$ has minimal ℓ_1-norm among all elements $z \in \mathcal{F}(y)$ if and only if*

$$|\sum_{x_i^* \neq 0} \operatorname{sgn}(x_i^*)\eta_i| \leqslant \sum_{x_i^* = 0} |\eta_i|, \quad \eta \in \mathcal{N}. \tag{2.28}$$

Moreover, x^* is unique if and only if we have strict inequality for all $\eta \in \mathcal{N}$ which are not identically zero.

Proof. If $x \in \mathcal{F}(y)$ has minimum ℓ_1-norm, then we have, for any $\eta \in \mathcal{N}$ and any $t \in \mathbb{R}$,

$$\sum_{i=1}^N |x_i + t\eta_i| \geqslant \sum_{i=1}^N |x_i|. \tag{2.29}$$

Fix $\eta \in \mathcal{N}$. If t is sufficiently small then $x_i + t\eta_i$ and x_i will have the same sign $s_i := \operatorname{sgn}(x_i)$ whenever $x_i \neq 0$. Hence, (2.29) can be written as

$$t \sum_{x_i \neq 0} s_i \eta_i + \sum_{x_i = 0} |t\eta_i| \geqslant 0.$$

Choosing t of an appropriate sign, we see that (2.28) is a necessary condition.

For the opposite direction, we note that if (2.28) holds then for each $\eta \in \mathcal{N}$, we have

$$\sum_{i=1}^N |x_i| \quad = \quad \sum_{x_i \neq 0} s_i x_i = \sum_{x_i \neq 0} s_i(x_i + \eta_i) - \sum_{x_i \neq 0} s_i \eta_i$$

$$\leqslant \sum_{x_i \neq 0} s_i(x_i + \eta_i) + \sum_{x_i=0} |\eta_i| \leqslant \sum_{i=1}^{N} |x_i + \eta_i|, \tag{2.30}$$

where the first inequality uses (2.28).

If x is unique then we have strict inequality in (2.29) and hence subsequently in (2.28). If we have strict inequality in (2.28) then the subsequent strict inequality in (2.30) implies uniqueness. □

Next, consider the minimization in a weighted $\ell_2(w)$-norm. Suppose that the weight w is *strictly positive* which we define to mean that $w_j > 0$ for all $j \in \{1, \dots, N\}$. In this case, $\ell_2(w)$ is a Hilbert space with the inner product

$$\langle u, v \rangle_w := \sum_{j=1}^{N} w_j u_j v_j. \tag{2.31}$$

Define

$$x^w := \arg \min_{z \in \mathcal{F}(y)} \|z\|_{\ell_2^N(w)}. \tag{2.32}$$

Because the $\|\cdot\|_{\ell_2^N(w)}$-norm is strictly convex, the minimizer x^w is necessarily unique; it is completely characterized by the orthogonality conditions

$$\langle x^w, \eta \rangle_w = 0, \quad \forall \eta \in \mathcal{N}. \tag{2.33}$$

Namely, x^w necessarily satisfies (2.33); on the other hand, any element $z \in \mathcal{F}(y)$ that satisfies $\langle z, \eta \rangle_w = 0$ for all $\eta \in \mathcal{N}$ is automatically equal to x^w.

A fundamental relationship between ℓ_1-minimization and weighted ℓ_2-minimization is easily shown, which might seem totally unrelated at first sight, due to the different characterization of respective minimizers.

Lemma 2.3 *Assume that x^* is an ℓ_1-minimizer and that x^* has no vanishing coordinates. Then the (unique) solution x^w of the weighted least squares problem*

$$x^w := \arg \min_{z \in \mathcal{F}(y)} \|z\|_{\ell_2^N(w)}, \quad w := (w_1, \dots, w_N), \quad \textit{where } w_j := |x_j^*|^{-1},$$

coincides with x^*.

Proof. Assume that x^* is not the $\ell_2^N(w)$-minimizer. Then there exists $\eta \in \mathcal{N}$ such that $0 < \langle x^*, \eta \rangle_w = \sum_{j=1}^N w_j \eta_j x_j^* = \sum_{j=1}^N \eta_j \operatorname{sgn}(x_j^*)$. However, by Lemma 2.2 and because x^* is an ℓ_1-minimizer, we have $\sum_{j=1}^N \eta_j \operatorname{sgn}(x_j^*) = 0$, a contradiction. □

An iteratively re-weighted least square algorithm

Since we do not know x^*, this observation cannot be used directly. However, it leads to the following paradigm for finding x^*. We choose a starting weight w^0 and solve the weighted ℓ_2 minimization for this weight. We then use this solution to define a new weight w^1 and repeat this process. An IRLS algorithm of this type appears for the first time in the approximation practice in the Ph.D. thesis of Lawson in 1961 [52], in the form of an algorithm for solving uniform approximation problems, in particular by Chebyshev polynomials, by means of limits of weighted ℓ_p–norm solutions. This iterative algorithm is now well-known in classical approximation theory as Lawson's algorithm. In [19] it is proved that this algorithm has in principle a linear convergence rate. In the 1970s extensions of Lawson's algorithm for ℓ_p-minimization, and in particular ℓ_1-minimization, were proposed. In signal analysis, IRLS was proposed as a technique to build algorithms for sparse signal reconstruction in [47]. Perhaps the most comprehensive mathematical analysis of the performance of IRLS for ℓ_p-minimization was given in the work of Osborne [63]. However, the interplay of NSP, ℓ_1-minimization, and a reweighted least square algorithm has been clarified only recently in the work [30]. In the following we describe the essential lines of the analysis of this algorithm, by taking advantage of results and terminology already introduced in previous sections. Our analysis of the algorithm (2.26) and (2.27) starts from the observation that

$$|t| = \min_{w>0} \frac{1}{2}\left(wt^2 + w^{-1}\right),$$

the minimum being reached for $w = \frac{1}{|t|}$. Inspired by this simple relationship, given a real number $\epsilon > 0$ and a weight vector $w \in \mathbb{R}^N$, with $w_j > 0$, $j = 1, \dots, N$, we define

$$\mathcal{J}(z, w, \epsilon) := \frac{1}{2}\left[\sum_{j=1}^{N} z_j^2 w_j + \sum_{j=1}^{N}(\epsilon^2 w_j + w_j^{-1})\right], \quad z \in \mathbb{R}^N. \tag{2.34}$$

The algorithm roughly described in (2.26) and (2.27) can be recasted as an alternating method for choosing minimizers and weights based on the functional $\mathcal{J}$.

To describe this more rigorously, we define for $z \in \mathbb{R}^N$ the nonincreasing rearrangement $r(z)$ of the absolute values of the entries of z. Thus $r(z)_i$ is the i-th largest element of the set $\{|z_j|,\ j = 1, \dots, N\}$, and a vector v is k-sparse iff $r(v)_{k+1} = 0$.

Algorithm 1. We initialize by taking $w^0 := (1, \ldots, 1)$. We also set $\epsilon_0 := 1$. We then recursively define for $n = 0, 1, \ldots,$

$$x^{n+1} := \arg \min_{z \in \mathcal{F}(y)} \mathcal{J}(z, w^n, \epsilon_n) = \arg \min_{z \in \mathcal{F}(y)} \|z\|_{\ell_2(w^n)} \tag{2.35}$$

and

$$\epsilon_{n+1} := \min(\epsilon_n, \frac{r(x^{n+1})_{K+1}}{N}), \tag{2.36}$$

where K is a fixed integer that will be described more fully later. We also define

$$w^{n+1} := \arg \min_{w>0} \mathcal{J}(x^{n+1}, w, \epsilon_{n+1}). \tag{2.37}$$

We stop the algorithm if $\epsilon_n = 0$; in this case we define $x^j := x^n$ for $j > n$. However, in general, the algorithm will generate an infinite sequence $(x^n)_{n \in N}$ of distinct vectors.

Each step of the algorithm requires the solution of a weighted least squares problem. In matrix form

$$x^{n+1} = D_n^{-1} A^* (A D_n^{-1} A^*)^{-1} y, \tag{2.38}$$

where D_n is the $N \times N$ diagonal matrix whose j-th diagonal entry is w_j^n and A^* denotes the transpose of the matrix A. Once x^{n+1} is found, the weight w^{n+1} is given by

$$w_j^{n+1} = [(x_j^{n+1})^2 + \epsilon_{n+1}^2]^{-1/2}, \quad j = 1, \ldots, N. \tag{2.39}$$

Preliminary results

We first make some comments about the decreasing rearrangement $r(z)$ and the j-term approximation errors for vectors in $\mathbb{R}^N$. We have the following lemma:

Lemma 2.4 *The map $z \mapsto r(z)$ is Lipschitz continuous on $(\mathbb{R}^N, \|\cdot\|_{\ell_\infty})$: for any $z, z' \in \mathbb{R}^N$, we have*

$$\|r(z) - r(z')\|_{\ell_\infty} \leqslant \|z - z'\|_{\ell_\infty}. \tag{2.40}$$

Moreover, for any j, we have

$$|\sigma_j(z)_{\ell_1} - \sigma_j(z')_{\ell_1}| \leqslant \|z - z'\|_{\ell_1}, \tag{2.41}$$

and for any $J > j$, we have

$$(J - j) r(z)_J \leqslant \|z - z'\|_{\ell_1} + \sigma_j(z')_{\ell_1}. \tag{2.42}$$

Proof. For any pair of points z and z', and any $j \in \{1, \ldots, N\}$, let Λ be a set of $j-1$ indices corresponding to the $j-1$ largest entries in z'. Then

$$r(z)_j \leqslant \max_{i \in \Lambda^c} |z_i| \leqslant \max_{i \in \Lambda^c} |z'_i| + \|z - z'\|_{\ell_\infty} = r(z')_j + \|z - z'\|_{\ell_\infty}. \tag{2.43}$$

We can also reverse the roles of z and z'. Therefore, we obtain (2.40). To prove (2.41), we approximate z by a j-term best approximation $z_{[j]} \in \Sigma_j$ of z' in ℓ_1. Then

$$\sigma_j(z)_{\ell_1} \leqslant \|z - z_{[j]}\|_{\ell_1} \leqslant \|z - z'\|_{\ell_1} + \sigma_j(z')_{\ell_1},$$

and the result follows from symmetry.

To prove (2.42), it suffices to note that $(J-j)\, r(z)_J \leqslant \sigma_j(z)_{\ell_1}$. □

Our next result is an approximate reverse triangle inequality for points in $\mathcal{F}(y)$. Its importance to us lies in its implication that whenever two points $z, z' \in \mathcal{F}(y)$ have close ℓ_1-norms and one of them is close to a k-sparse vector, then they necessarily are close to each other. (Note that it also implies that the other vector must then also be close to that k-sparse vector.) This is a geometric property of the null space.

Lemma 2.5 *Assume that the NSP holds for some L and $\gamma < 1$. Then, for any $z, z' \in \mathcal{F}(y)$, we have*

$$\|z' - z\|_{\ell_1} \leqslant \frac{1+\gamma}{1-\gamma} \left(\|z'\|_{\ell_1} - \|z\|_{\ell_1} + 2\sigma_L(z)_{\ell_1} \right). \tag{2.44}$$

Proof. Let T be a set of indices of the L largest entries in z. Then

$$\begin{aligned} \|(z'-z)_{T^c}\|_{\ell_1} &\leqslant \|z'_{T^c}\|_{\ell_1} + \|z_{T^c}\|_{\ell_1} \\ &= \|z'\|_{\ell_1} - \|z'_T\|_{\ell_1} + \sigma_L(z)_{\ell_1} \\ &= \|z\|_{\ell_1} + \|z'\|_{\ell_1} - \|z\|_{\ell_1} - \|z'_T\|_{\ell_1} + \sigma_L(z)_{\ell_1} \\ &= \|z_T\|_{\ell_1} - \|z'_T\|_{\ell_1} + \|z'\|_{\ell_1} - \|z\|_{\ell_1} + 2\sigma_L(z)_{\ell_1} \\ &\leqslant \|(z'-z)_T\|_{\ell_1} + \|z'\|_{\ell_1} - \|z\|_{\ell_1} + 2\sigma_L(z)_{\ell_1}. \end{aligned} \tag{2.45}$$

Using the NSP, this gives

$$\|(z'-z)_T\|_{\ell_1} \leqslant \gamma \|(z'-z)_{T^c}\|_{\ell_1} \leqslant \gamma(\|(z'-z)_T\|_{\ell_1} + \|z'\|_{\ell_1} - \|z\|_{\ell_1} + 2\sigma_L(z)_{\ell_1}). \tag{2.46}$$

In other words,

$$\|(z'-z)_T\|_{\ell_1} \leqslant \frac{\gamma}{1-\gamma} (\|z'\|_{\ell_1} - \|z\|_{\ell_1} + 2\sigma_L(z)_{\ell_1}). \tag{2.47}$$

Using this, together with (2.45), we obtain

$$\|z'-z\|_{\ell_1} = \|(z'-z)_{T^c}\|_{\ell_1} + \|(z'-z)_T\|_{\ell_1} \leqslant \frac{1+\gamma}{1-\gamma} (\|z'\|_{\ell_1} - \|z\|_{\ell_1} + 2\sigma_L(z)_{\ell_1}), \tag{2.48}$$

as desired. □

By using the previous lemma we obtain the following estimate.

Lemma 2.6 *Assume that the NSP holds for some L and $\gamma < 1$. Suppose that $\mathcal{F}(y)$ contains an L-sparse vector. Then this vector is the unique ℓ_1-minimizer in $\mathcal{F}(y)$; denoting it by x^*, we have moreover, for all $v \in \mathcal{F}(y)$,*

$$\|v - x^*\|_{\ell_1} \leqslant 2\,\frac{1+\gamma}{1-\gamma}\,\sigma_L(v)_{\ell_1}\,. \tag{2.49}$$

Proof. For the time being, we denote the L-sparse vector in $\mathcal{F}(y)$ by x_s.
Applying (2.44) with $z' = v$ and $z = x_s$, we find

$$\|v - x_s\|_{\ell_1} \leqslant \frac{1+\gamma}{1-\gamma}[\|v\|_{\ell_1} - \|x_s\|_{\ell_1}]\,;$$

since $v \in \mathcal{F}(y)$ is arbitrary, this implies that $\|v\|_{\ell_1} - \|x_s\|_{\ell_1} \geqslant 0$ for all $v \in \mathcal{F}(y)$, so that x_s is an ℓ_1-norm minimizer in $\mathcal{F}(y)$.

If x' were another ℓ_1-minimizer in $\mathcal{F}(y)$, then it would follow that $\|x'\|_{\ell_1} = \|x_s\|_{\ell_1}$, and the inequality we just derived would imply $\|x' - x_s\|_{\ell_1} = 0$, or $x' = x_s$. It follows that x_s is the unique ℓ_1-minimizer in $\mathcal{F}(y)$, which we denote by x^*, as proposed earlier.

Finally, we apply (2.44) with $z' = x^*$ and $z = v$, and we obtain

$$\|v - x^*\| \leqslant \frac{1+\gamma}{1-\gamma}(\|x^*\|_{\ell_1} - \|v\|_{\ell_1} + 2\sigma_L(v)_{\ell_1}) \leqslant 2\frac{1+\gamma}{1-\gamma}\sigma_L(v)_{\ell_1}\,,$$

where we have used the ℓ_1-minimization property of x^*. □

Our next set of remarks centers around the functional $\mathcal{J}$ defined by (2.34). Note that for each $n = 1, 2, \dots,$ we have

$$\mathcal{J}(x^{n+1}, w^{n+1}, \epsilon_{n+1}) = \sum_{j=1}^{N}[(x_j^{n+1})^2 + \epsilon_{n+1}^2]^{1/2}. \tag{2.50}$$

We also have the following monotonicity property which holds for all $n \geqslant 0$:

$$\mathcal{J}(x^{n+1}, w^{n+1}, \epsilon_{n+1}) \leqslant \mathcal{J}(x^{n+1}, w^{n}, \epsilon_{n+1}) \leqslant \mathcal{J}(x^{n+1}, w^{n}, \epsilon_{n}) \leqslant \mathcal{J}(x^{n}, w^{n}, \epsilon_{n}). \tag{2.51}$$

Here the first inequality follows from the minimization property that defines w^{n+1}, the second inequality from $\epsilon_{n+1} \leqslant \epsilon_n$, and the last inequality from the minimization property that defines x^{n+1}. For each n, x^{n+1} is completely determined by w^n; for $n = 0$, in particular, x^1 is determined solely by w^0, and independent of the choice of $x^0 \in \mathcal{F}(y)$. (With the initial weight vector defined by $w^0 = (1, \dots, 1)$, x^1 is the classical minimum ℓ_2-norm element of $\mathcal{F}(y)$.) The inequality (2.51) for $n = 0$ thus holds for arbitrary $x^0 \in \mathcal{F}(y)$.

Lemma 2.7 *For each $n \geqslant 1$ we have*

$$\|x^n\|_{\ell_1} \leqslant \mathcal{J}(x^1, w^0, \epsilon_0) =: \mathcal{A} \tag{2.52}$$

and

$$w_j^n \geqslant \mathcal{A}^{-1}, \quad j = 1, \dots, N. \tag{2.53}$$

Proof. The bound (2.52) follows from (2.51) and

$$\|x^n\|_{\ell_1} \leqslant \sum_{j=1}^{N} [(x_j^n)^2 + \epsilon_n^2]^{1/2} = \mathcal{J}(x^n, w^n, \epsilon_n).$$

The bound (2.53) follows from $(w_j^n)^{-1} = [(x_j^n)^2 + \epsilon_n^2]^{1/2} \leqslant \mathcal{J}(x^n, w^n, \epsilon_n) \leqslant \mathcal{A}$, where the last inequality uses (2.51). □

Convergence of the algorithm

In this section, we prove that the algorithm converges. Our starting point is the following lemma that establishes $(x^n - x^{n+1}) \to 0$ for $n \to \infty$.

Lemma 2.8 *Given any $y \in \mathbb{R}^m$, the x^n satisfy*

$$\sum_{n=1}^{\infty} \|x^{n+1} - x^n\|_{\ell_2}^2 \leqslant 2\mathcal{A}^2. \tag{2.54}$$

where $\mathcal{A}$ is the constant of Lemma 2.7. In particular, we have

$$\lim_{n \to \infty} (x^n - x^{n+1}) = 0. \tag{2.55}$$

Proof. For each $n = 1, 2, \dots,$ we have

$$\begin{aligned} 2[\mathcal{J}(x^n, w^n, \epsilon_n) - \mathcal{J}(x^{n+1}, w^{n+1}, \epsilon_{n+1})] &\geqslant 2[\mathcal{J}(x^n, w^n, \epsilon_n) - \mathcal{J}(x^{n+1}, w^n, \epsilon_n)] \\ &= \langle x^n, x^n \rangle_{w^n} - \langle x^{n+1}, x^{n+1} \rangle_{w^n} \\ &= \langle x^n + x^{n+1}, x^n - x^{n+1} \rangle_{w^n} \\ &= \langle x^n - x^{n+1}, x^n - x^{n+1} \rangle_{w^n} \\ &= \sum_{j=1}^{N} w_j^n (x_j^n - x_j^{n+1})^2 \\ &\geqslant \mathcal{A}^{-1} \|x^n - x^{n+1}\|_{\ell_2}^2, \end{aligned} \tag{2.56}$$

where the third equality uses the fact that $\langle x^{n+1}, x^n - x^{n+1} \rangle_{w^n} = 0$ (observe that $x^{n+1} - x^n \in \mathcal{N}$ and invoke (2.33)), and the inequality uses the bound (2.53) on the weights. If we now sum these inequalities over $n \geqslant 1$, we arrive at (2.54). □

From the monotonicity of ϵ_n, we know that $\epsilon := \lim_{n\to\infty} \epsilon_n$ exists and is non-negative. The following functional will play an important role in our proof of convergence:

$$f_\epsilon(z) := \sum_{j=1}^{N} (z_j^2 + \epsilon^2)^{1/2}. \tag{2.57}$$

Notice that if we knew that x^n converged to x then, in view of (2.50), $f_\epsilon(x)$ would be the limit of $\mathcal{J}(x^n, w^n, \epsilon_n)$. When $\epsilon > 0$ the functional f_ϵ is strictly convex and therefore has a unique minimizer

$$x^\varepsilon := \arg \min_{z\in\mathcal{F}(y)} f_\epsilon(z). \tag{2.58}$$

This minimizer is characterized by the following lemma:

Lemma 2.9 *Let* $\varepsilon > 0$ *and* $z \in \mathcal{F}(y)$. *Then* $z = x^\varepsilon$ *if and only if* $\langle z, \eta\rangle_{\widetilde{w}(z,\varepsilon)} = 0$ *for all* $\eta \in \mathcal{N}$, *where* $\widetilde{w}(z,\varepsilon)_i = [z_i^2 + \varepsilon^2]^{-1/2}$.

Proof. For the "only if" part, let $z = x^\varepsilon$ and $\eta \in \mathcal{N}$ be arbitrary. Consider the analytic function

$$G_\varepsilon(t) := f_\varepsilon(z + t\eta) - f_\varepsilon(z).$$

We have $G_\varepsilon(0) = 0$, and by the minimization property $G_\varepsilon(t) \geqslant 0$ for all $t \in \mathbb{R}$. Hence, $G'_\varepsilon(0) = 0$. A simple calculation reveals that

$$G'_\varepsilon(0) = \sum_{j=1}^{N} \frac{\eta_i z_i}{[z_i^2 + \varepsilon^2]^{1/2}} = \langle z, \eta\rangle_{\widetilde{w}(z,\varepsilon)},$$

which gives the desired result.

For the "if" part, assume that $z \in \mathcal{F}(y)$ and $\langle z, \eta\rangle_{\widetilde{w}(z,\varepsilon)} = 0$ for all $\eta \in \mathcal{N}$, where $\widetilde{w}(z,\varepsilon)$ is defined as above. We shall show that z is a minimizer of f_ϵ on $\mathcal{F}(y)$. Indeed, consider the convex univariate function $[u^2 + \epsilon^2]^{1/2}$. For any point u_0 we have from convexity that

$$[u^2 + \epsilon^2]^{1/2} \geqslant [u_0^2 + \epsilon^2]^{1/2} + [u_0^2 + \epsilon^2]^{-1/2} u_0 (u - u_0), \tag{2.59}$$

because the right side is the linear function which is tangent to this function at u_0. It follows that for any point $v \in \mathcal{F}(y)$ we have

$$f_\epsilon(v) \geqslant f_\epsilon(z) + \sum_{j=1}^{N} [z_j^2 + \epsilon^2]^{-1/2} z_j (v_j - z_j) = f_\epsilon(z) + \langle z, v - z\rangle_{\tilde{w}(z,\varepsilon)} = f_\epsilon(z), \tag{2.60}$$

where we have used the orthogonality condition (2.66) and the fact that $v - z$ is in $\mathcal{N}$. Since v is arbitrary, it follows that $z = x^\varepsilon$, as claimed. □

We now give the convergence of the algorithm.

Theorem 2.10 *Let K (the same index as used in the update rule* (2.36)*) be chosen so that A satisfies the Null Space Property of order K, with $\gamma < 1$. Then, for each $y \in \mathbb{R}^m$, the output of Algorithm 1 converges to a vector $\bar{x}$, with $r(\bar{x})_{K+1} = N \lim_{n\to\infty} \varepsilon_n$ and the following hold:*
(i) *If $\varepsilon = \lim_{n\to\infty} \varepsilon_n = 0$, then $\bar{x}$ is K-sparse; in this case there is therefore a unique ℓ_1-minimizer x^*, and $\bar{x} = x^*$; moreover, we have, for $k \leqslant K$, and any $z \in \mathcal{F}(y)$,*

$$\|z - \bar{x}\|_{\ell_1} \leqslant c\sigma_k(z)_{\ell_1}, \quad \text{with } c := \frac{2(1+\gamma)}{1-\gamma} \tag{2.61}$$

(ii) *If $\varepsilon = \lim_{n\to\infty} \varepsilon_n > 0$, then $\bar{x} = x^\varepsilon$;*
(iii) *In this last case, if γ satisfies the stricter bound $\gamma < 1 - \frac{2}{K+2}$ (or, equivalently, if $\frac{2\gamma}{1-\gamma} < K$), then we have, for all $z \in \mathcal{F}(y)$ and any $k < K - \frac{2\gamma}{1-\gamma}$, that*

$$\|z - \bar{x}\|_{\ell_1} \leqslant \tilde{c}\sigma_k(z)_{\ell_1}, \quad \text{with } \tilde{c} := \frac{2(1+\gamma)}{1-\gamma}\left[\frac{K - k + \frac{3}{2}}{K - k - \frac{2\gamma}{1-\gamma}}\right] \tag{2.62}$$

As a consequence, this case is excluded if $\mathcal{F}(y)$ contains a vector of sparsity $k < K - \frac{2\gamma}{1-\gamma}$.

Note that the approximation properties (2.61) and (2.62) are exactly of the same order as the one (1.5) provided by ℓ_1-minimization. However, in general, $\bar{x}$ is not necessarily an ℓ_1-minimizer, unless it coincides with a sparse solution.
The constant $\tilde{c}$ can be quite reasonable; for instance, if $\gamma \leqslant 1/2$ and $k \leqslant K - 3$, then we have $\tilde{c} \leqslant 9\,\frac{1+\gamma}{1-\gamma} \leqslant 27$.

Proof. Note that since $\varepsilon_{n+1} \leq \varepsilon_n$, the ε_n always converge. We start by considering the case $\epsilon := \lim_{n\to\infty} \epsilon_n = 0$.

Case $\epsilon = 0$: In this case, we want to prove that x^n converges , and that its limit is an ℓ_1-minimizer. Suppose that $\epsilon_{n_0} = 0$ for some n_0. Then by the definition of the algorithm, we know that the iteration is stopped at $n = n_0$, and $x^n = x^{n_0}$, $n \geqslant n_0$. Therefore $\bar{x} = x^{n_0}$. From the definition of ε_n, it then also follows that $r(x^{n_0})_{K+1} = 0$ and so $\bar{x} = x^{n_0}$ is K-sparse. As noted in Lemma 2.6, if a K-sparse solution exists when A satisfies the NSP of order K with $\gamma < 1$, then it is the unique ℓ_1-minimizer. Therefore, $\bar{x}$ equals x^*, this unique minimizer.

Suppose now that $\epsilon_n > 0$ for all n. Since $\varepsilon_n \to 0$, there is an increasing sequence of indices (n_i) such that $\varepsilon_{n_i} < \epsilon_{n_i - 1}$ for all i. By the definition (2.36) of $(\varepsilon_n)_{n\in\mathcal{N}}$, we must have $r(x^{n_i})_{K+1} < N\epsilon_{n_i-1}$ for all i. Noting that $(x^n)_{n\in\mathcal{N}}$ is a bounded sequence, there exists a subsequence $(p_j)_{j\in\mathcal{N}}$ of $(n_i)_{i\in\mathcal{N}}$ such that $(x^{p_j})_{j\in\mathcal{N}}$ converges to a point

$\widetilde{x} \in \mathcal{F}(y)$. By Lemma 2.4, we know that $r(x^{p_j})_{K+1}$ converges to $r(\widetilde{x})_{K+1}$. Hence we get

$$r(\widetilde{x})_{K+1} = \lim_{j\to\infty} r(x^{p_j})_{K+1} \leqslant \lim_{j\to\infty} N\epsilon_{p_j-1} = 0, \tag{2.63}$$

which means that the support-width of $\widetilde{x}$ is at most K, i.e. $\widetilde{x}$ is K-sparse. By the same token used above, we again have that $\widetilde{x} = x^*$, the unique ℓ_1-minimizer. We must still show that $x^n \to x^*$. Since $x^{p_j} \to x^*$ and $\epsilon_{p_j} \to 0$, (2.50) implies $\mathcal{J}(x^{p_j}, w^{p_j}, \epsilon_{p_j}) \to \|x^*\|_{\ell_1}$. By the monotonicity property stated in (2.51), we get $\mathcal{J}(x^n, w^n, \varepsilon_n) \to \|x^*\|_{\ell_1}$. Since (2.50) implies

$$\mathcal{J}(x^n, w^n, \varepsilon_n) - N\varepsilon_n \leqslant \|x^n\|_{\ell_1} \leqslant \mathcal{J}(x^n, w^n, \varepsilon_n), \tag{2.64}$$

we obtain $\|x^n\|_{\ell_1} \to \|x^*\|_{\ell_1}$. Finally, we invoke Lemma 2.5 with $z' = x^n$, $z = x^*$, and $k = K$ to get

$$\limsup_{n\to\infty} \|x^n - x^*\|_{\ell_1} \leqslant \frac{1+\gamma}{1-\gamma}\left(\lim_{n\to\infty} \|x^n\|_{\ell_1} - \|x^*\|_{\ell_1}\right) = 0, \tag{2.65}$$

which completes the proof that $x^n \to x^*$ in this case.

Finally, (2.61) follows from (2.49) of Lemma 2.6 (with $L = K$), and the observation that $\sigma_n(z) \geqslant \sigma_{n'}(z)$ if $n \leqslant n'$.

Case $\epsilon > 0$**:** We shall first show that $x^n \to x^\varepsilon$, $n \to \infty$, with x^ε as defined by (2.58). By Lemma 2.7, we know that $(x^n)_{n=1}^\infty$ is a bounded sequence in $\mathbb{R}^N$ and hence this sequence has accumulation points. Let (x^{n_i}) be any convergent subsequence of (x^n) and let $\widetilde{x} \in \mathcal{F}(y)$ be its limit. We want to show that $\widetilde{x} = x^\varepsilon$.

Since $w_j^n = [(x_j^n)^2 + \epsilon_n^2]^{-1/2} \leqslant \epsilon^{-1}$, it follows that $\lim_{i\to\infty} w_j^{n_i} = [(\widetilde{x}_j)^2 + \epsilon^2]^{-1/2} = \widetilde{w}(\widetilde{x}, \varepsilon)_j =: \widetilde{w}_j$, $j = 1, \ldots, N$. On the other hand, by invoking Lemma 2.8, we now find that $x^{n_i+1} \to \widetilde{x}$, $i \to \infty$. It then follows from the orthogonality relations (2.33) that for every $\eta \in \mathcal{N}$, we have

$$\langle \widetilde{x}, \eta \rangle_{\widetilde{w}} = \lim_{i\to\infty} \langle x^{n_i+1}, \eta \rangle_{w^{n_i}} = 0. \tag{2.66}$$

Now the "if" part of Lemma 2.9 implies that $\widetilde{x} = x^\varepsilon$. Hence x^ε is the unique accumulation point of $(x^n)_{n\in\mathcal{N}}$ and therefore its limit. This establishes (ii).

To prove the error estimate (2.62) stated in (iii), we first note that for any $z \in \mathcal{F}(y)$, we have

$$\|x^\varepsilon\|_{\ell_1} \leqslant f_\epsilon(x^\varepsilon) \leqslant f_\epsilon(z) \leqslant \|z\|_{\ell_1} + N\epsilon, \tag{2.67}$$

where the second inequality uses the minimizing property of x^ε. Hence it follows that $\|x^\varepsilon\|_{\ell_1} - \|z\|_{\ell_1} \leqslant N\epsilon$. We now invoke Lemma 2.5 to obtain

$$\|x^\varepsilon - z\|_{\ell_1} \leqslant \frac{1+\gamma}{1-\gamma}[N\epsilon + 2\sigma_k(z)_{\ell_1}]. \tag{2.68}$$

From Lemma 2.4 and (2.36), we obtain

$$N\epsilon = \lim_{n\to\infty} N\epsilon_n \leqslant \lim_{n\to\infty} r(x^n)_{K+1} = r(x^\varepsilon)_{K+1}. \tag{2.69}$$

It follows from (2.42) that

$$\begin{aligned}(K+1-k)N\epsilon &\leqslant (K+1-k)r(x^\varepsilon)_{K+1} \\ &\leqslant \|x^\varepsilon - z\|_{\ell_1} + \sigma_k(z)_{\ell_1} \\ &\leqslant \frac{1+\gamma}{1-\gamma}[N\epsilon + 2\sigma_k(z)_{\ell_1}] + \sigma_k(z)_{\ell_1},\end{aligned} \tag{2.70}$$

where the last inequality uses (2.68). Since by assumption on K, we have $K-k > \frac{2\gamma}{1-\gamma}$, i.e. $K+1-k > \frac{1+\gamma}{1-\gamma}$, we obtain

$$N\epsilon + 2\sigma_k(z)_{\ell_1} \leqslant \frac{2(K-k)+3}{(K-k)-\frac{2\gamma}{1-\gamma}}\,\sigma_k(z)_{\ell_1}.$$

Using this back in (2.68), we arrive at (2.62).

Finally, notice that if $\mathcal{F}(y)$ contains a k-sparse vector (with $k < K - \frac{2\gamma}{1-\gamma}$), then we know already that this must be the unique ℓ_1-minimizer x^*; it then follows from our arguments above that we must have $\varepsilon = 0$. Indeed, if we had $\varepsilon > 0$, then (2.70) would hold for $z = x^*$; since x^* is k-sparse, $\sigma_k(x^*)_{\ell_1} = 0$, implying $\varepsilon = 0$, a contradiction with the assumption $\varepsilon > 0$. This finishes the proof.

□

Local linear rate of convergence

It is instructive to show a further very interesting result concerning the local rate of convergence of this algorithm, which makes heavily use of the NSP as well as the optimality properties we introduced above. One assumes here that $\mathcal{F}(y)$ contains the k-sparse vector x^*. The algorithm produces the sequence x^n, which converges to x^*, as established above. One denotes the (unknown) support of the k-sparse vector x^* by T.

We introduce an auxiliary sequence of error vectors $\eta^n \in \mathcal{N}$ via $\eta^n := x^n - x^*$ and

$$E_n := \|\eta^n\|_{\ell_1} = \|x^* - x^n\|_{\ell_1^N}.$$

We know that $E_n \to 0$.

The following theorem gives a bound on the rate of convergence of E_n to zero.

Theorem 2.11 *Assume A satisfies NSP of order K with constant γ such that $0 < \gamma < 1 - \frac{2}{K+2}$. Suppose that $k < K - \frac{2\gamma}{1-\gamma}$, $0 < \rho < 1$, and $0 < \gamma < 1 - \frac{2}{K+2}$ are such that*

$$\mu := \frac{\gamma(1+\gamma)}{1-\rho}\left(1 + \frac{1}{K+1-k}\right) < 1.$$

Assume that $\mathcal{F}(y)$ contains a k-sparse vector x^* and let $T = \operatorname{supp}(x^*)$. Let n_0 be such that

$$E_{n_0} \leqslant R^* := \rho \min_{i \in T} |x_i^*|. \tag{2.71}$$

Then for all $n \geqslant n_0$, we have

$$E_{n+1} \leqslant \mu\, E_n. \tag{2.72}$$

Consequently x^n converges to x^* exponentially.

Proof. We start with the relation (2.33) with $w = w^n$, $x^w = x^{n+1} = x^* + \eta^{n+1}$, and $\eta = x^{n+1} - x^* = \eta^{n+1}$, which gives

$$\sum_{i=1}^{N} (x_i^* + \eta_i^{n+1})\eta_i^{n+1} w_i^n = 0.$$

Rearranging the terms and using the fact that x^* is supported on T, we get

$$\sum_{i=1}^{N} |\eta_i^{n+1}|^2 w_i^n = -\sum_{i \in T} x_i^* \eta_i^{n+1} w_i^n = -\sum_{i \in T} \frac{x_i^*}{[(x_i^n)^2 + \epsilon_n^2]^{1/2}} \eta_i^{n+1}. \tag{2.73}$$

Prove of the theorem is by induction. One assumes that we have shown $E_n \leqslant R^*$ already. We then have, for all $i \in T$,

$$|\eta_i^n| \leqslant \|\eta^n\|_{\ell_1^N} = E_n \leqslant \rho |x_i^*| \ ,$$

so that

$$\frac{|x_i^*|}{[(x_i^n)^2 + \epsilon_n^2]^{1/2}} \leqslant \frac{|x_i^*|}{|x_i^n|} = \frac{|x_i^*|}{|x_i^* + \eta_i^n|} \leqslant \frac{1}{1-\rho}, \tag{2.74}$$

and hence (2.73) combined with (2.74) and NSP gives

$$\sum_{i=1}^{N} |\eta_i^{n+1}|^2 w_i^n \leqslant \frac{1}{1-\rho} \|\eta_T^{n+1}\|_{\ell_1} \leqslant \frac{\gamma}{1-\rho} \|\eta_{T^c}^{n+1}\|_{\ell_1}$$

At the same time, the Cauchy-Schwarz inequality combined with the above estimate yields

$$\begin{aligned} \|\eta_{T^c}^{n+1}\|_{\ell_1}^2 &\leqslant \left(\sum_{i \in T^c} |\eta_i^{n+1}|^2 w_i^n \right) \left(\sum_{i \in T^c} [(x_i^n)^2 + \epsilon_n^2]^{1/2} \right) \\ &\leqslant \left(\sum_{i=1}^{N} |\eta_i^{n+1}|^2 w_i^n \right) \left(\sum_{i \in T^c} [(\eta_i^n)^2 + \epsilon_n^2]^{1/2} \right) \\ &\leqslant \frac{\gamma}{1-\rho} \|\eta_{T^c}^{n+1}\|_{\ell_1} \left(\|\eta^n\|_{\ell_1} + N\epsilon_n \right). \end{aligned} \tag{2.75}$$

If $\eta_{T^c}^{n+1} = 0$, then $x_{T^c}^{n+1} = 0$. In this case x^{n+1} is k-sparse and the algorithm has stopped by definition; since $x^{n+1} - x^*$ is in the null space $\mathcal{N}$, which contains no k-sparse elements other than 0, we have already obtained the solution $x^{n+1} = x^*$. If $\eta_{T^c}^{n+1} \neq 0$, then after canceling the factor $\|\eta_{T^c}^{n+1}\|_{\ell_1}$ in (2.75), we get

$$\|\eta_{T^c}^{n+1}\|_{\ell_1} \leqslant \frac{\gamma}{1-\rho} \left(\|\eta^n\|_{\ell_1} + N\epsilon_n \right) ,$$

and thus

$$\|\eta^{n+1}\|_{\ell_1} = \|\eta_T^{n+1}\|_{\ell_1} + \|\eta_{T^c}^{n+1}\|_{\ell_1} \leqslant (1+\gamma)\|\eta_{T^c}^{n+1}\|_{\ell_1} \leqslant \frac{\gamma(1+\gamma)}{1-\rho}\left(\|\eta^n\|_{\ell_1} + N\epsilon_n\right). \tag{2.76}$$

Now, we also have by (2.36) and (2.42)

$$N\epsilon_n \leqslant r(x^n)_{K+1} \leqslant \frac{1}{K+1-k}(\|x^n - x^*\|_{\ell_1} + \sigma_k(x^*)_{\ell_1}) = \frac{\|\eta^n\|_{\ell_1}}{K+1-k}, \tag{2.77}$$

since by assumption $\sigma_k(x^*) = 0$. This, together with (2.76), yields the desired bound,

$$E_{n+1} = \|\eta^{n+1}\|_{\ell_1} \leqslant \frac{\gamma(1+\gamma)}{1-\rho}\left(1 + \frac{1}{K+1-k}\right)\|\eta^n\|_{\ell_1} = \mu E_n.$$

In particular, since $\mu < 1$, we have $E_{n+1} \leqslant R^*$, which completes the induction step. It follows that $E_{n+1} \leqslant \mu E_n$ for all $n \geqslant n_0$. □

A surprising superlinear convergence promoting ℓ_τ-minimization for $\tau < 1$

The linear rate (2.72) can be improved significantly, by a very simple modification of the rule of updating the weight:

$$w_j^{n+1} = \left((x_j^{n+1})^2 + \epsilon_{n+1}^2\right)^{-\frac{2-\tau}{2}}, \quad j = 1, \ldots, N, \text{ for any } 0 < \tau < 1.$$

This corresponds to the substitution of the function $\mathcal{J}$ with

$$\mathcal{J}_\tau(z,w,\epsilon) := \frac{\tau}{2}\left[\sum_{j=1}^N z_j^2 w_j + \sum_{j=1}^N \left(\epsilon^2 w_j + \frac{2-\tau}{\tau}\frac{1}{w_j^{\frac{\tau}{2-\tau}}}\right)\right], \quad z \in \mathbb{R}^N, w \in \mathbb{R}_+^N, \epsilon \in \mathbb{R}_+.$$

Surprisingly the rate of local convergence of this modified algorithm is superlinear; the rate is larger for smaller τ, increasing to approach a quadratic regime as $\tau \to 0$. More precisely the local error $E_n := \|x^n - x^*\|_{\ell_\tau^N}^\tau$ satisfies

$$E_{n+1} \leqslant \mu(\gamma,\tau) E_n^{2-\tau}, \tag{2.78}$$

where $\mu(\gamma,\tau) < 1$ for $\gamma > 0$ sufficiently small. The validity of (2.78) is restricted to x^n in a (small) ball centered at x^*. In particular, if x^0 is close enough to x^* then (2.78) ensures the convergence of the algorithm to the k-sparse solution x^*.

Numerical results

In this section we present numerical experiments that illustrate that the bounds derived in the theoretical analysis do manifest themselves in practice.

We start with numerical results that confirm the linear rate of convergence of our iteratively re-weighted least square algorithm for ℓ_1-minimization, and its robust recovery of sparse vectors. In the experiments we used a matrix A of dimensions $m \times N$ and Gaussian $\mathcal{N}(0, 1/m)$ i.i.d. entries. We already mentioned that such matrices are known to possess (with high probability) the RIP property with optimal bounds. In Figure 2.1 we depict the approximation error to the unique sparsest solution shown in Figure 2.2, and the instantaneous rate of convergence. The numerical results both confirm the expected linear rate of convergence and the robust reconstruction of the sparse vector.

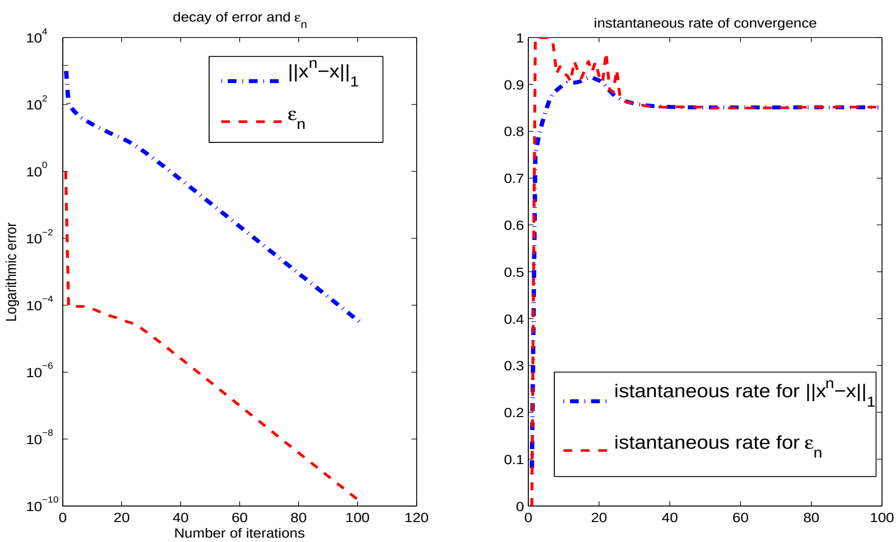


Figure 2.1 An experiment, with a matrix A of size 250×1500 with Gaussian $\mathcal{N}(0, \frac{1}{250})$ i.i.d. entries, in which recovery is sought of the 45-sparse vector x^* represented in Figure 2.2 from its image $y = Ax$. Left: plot of $\log_{10}(\|x^n - x^*\|_{\ell_1})$ as a function of n, where the x^n are generated by Algorithm 1, with ϵ_n defined adaptively, as in (2.36). Note that the scale in the ordinate axis does not report the logarithm $0, -1, -2, \ldots,$ but the corresponding accuracies $10^0, 10^{-1}, 10^{-2}, \ldots$ for $\|x^n - x^*\|_{\ell_1}$. The graph also plots ϵ_n as a function of n. Right: plot of the ratios $\|x^n - x^{n+1}\|_{\ell_1} / \|x^n - x^{n-1}\|_{\ell_1}$, and $(\varepsilon_n - \varepsilon_{n+1})/(\epsilon_{n-1} - \epsilon_n)$ for the same examples.

Next, we compare the linear convergence achieved with ℓ_1-minimization with the superlinear convergence obtained by the iteratively re-weighted least square algorithm promoting ℓ_τ-minimization.

In Figure 2.3 we are interested in the comparison of the rate of convergence when our algorithm is used for different choices of $0 < \tau \leqslant 1$. For $\tau = 1, .8, .6$ and $.56$, the figure shows the error, as a function of the iteration step n, for the iterative algorithm, with different fixed values of τ. For $\tau = 1$, the rate is linear, as in Figure

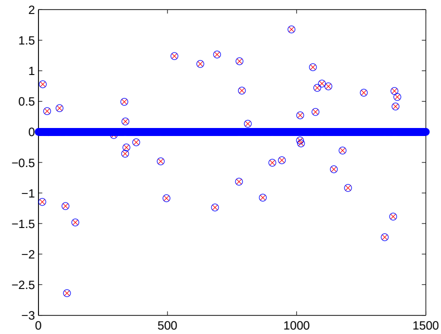


Figure 2.2 The sparse vector used in the example illustrated in Figure 2.1. This vector has dimension 1500, but only 45 non-zero entries.

2.1. For the smaller values $\tau = .8, .6$ and $.56$ the iterations initially follow the same linear rate; once they are sufficiently close to the sparse solution, the convergence rate speeds up dramatically, suggesting we have entered the region of validity of (2.78). For smaller values of τ numerical experiments do not always lead to convergence: in some cases the algorithm never get to the neighborhood of the solution where convergence is ensured. However, in this case a combination of initial iterations with the ℓ_1-inspired IRLS (for which we always have convergence) and later iterations with ℓ_τ-inspired IRLS for smaller τ allow again for a very fast convergence to the sparsest solution; this is illustrated in Figure 2.3 for the case $\tau = .5$.

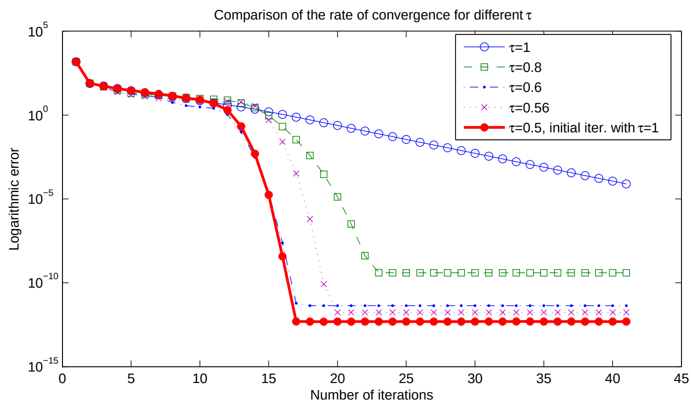


Figure 2.3 We show the decay of logarithmic error, as a function of the number of iterations of the algorithm for different values of τ (1, 0.8, 0.6, 0.56). We show also the results of an experiment in which the initial 10 iterations are performed with $\tau = 1$ and the remaining iterations with $\tau = 0.5$.

Enhanced recovery in compressed sensing and relationship with other work

In this section, we shortly report a phenomenon of enhancing of rate of recovery as depicted in Figure 2.4. As shown there, the IRLS with weights that gradually moved from an ℓ_1- to an ℓ_τ-minimization goal produce a higher experimentally determined probability of successful recovery as a function of k.

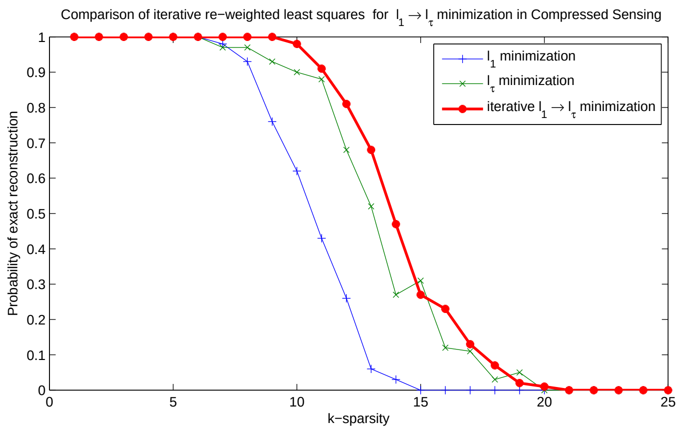


Figure 2.4 The (experimentally determined) probability of successful recovery of a sparse 250-dimensional vector x, with sparsity k, from its image $y = Ax$, as a function of k. In these experiments the matrix A is 50×250 dimensional, with i.i.d. Gaussian $\mathcal{N}(0, \frac{1}{50})$ entries. The matrix is generated once; then, for each sparsity value k shown in the plot, 500 attempts were made, for randomly generated k-sparse vectors x. Two different IRLS algorithms were compared, one with weights inspired by ℓ_1-minimization and the other with weights that gradually moved from an ℓ_1- to an ℓ_τ-minimization goal, with final $\tau = 0.5$.

Some open problems

1. In practice this algorithm appears very robust and its convergence is either linear or even superlinear when properly tuned as previously indicated. However, such guarantees of rate of convergence are valid only in a neighborhood of a solution which is presently very difficult to estimate. This lack of knowledge does not allow us yet to estimate properly the complexity of this method.

2. For $\tau < 1$ the algorithm seems converging properly when τ is not too small, but when, say, $\tau < 0.5$, then the algorithm tends to fail to reach the region of guaranteed convergence. It is an open problem to characterize very sharply such phase transition, and heuristic methods of avoidance of local minima are also of great interest.

3. While error guarantees of the type (1.5) are given, it is open whether (1.6) and (1.13) can hold for this algorithm.

2.1.3 Extensions to the Minimization Functionals with Total Variation Terms

In concrete applications, e.g., for image processing, one might be interested to recover at best a digital image provided only partial linear or nonlinear measurements, possibly corrupted by noise. Given the observation that natural and man-made images can be characterized by a relatively small number of edges and extensive relatively uniform parts, one may want to help the reconstruction by imposing that the interesting solution is the one which matches the given data and has also a few discontinuities localized on sets of lower dimension.

In the context of *compressed sensing* as described in the previous sections, we have already clarified that the minimization of ℓ_1-norms occupies a fundamental role for the promotion of sparse solutions. This understanding furnishes an important interpretation of *total variation minimization*, i.e., the minimization of the L^1-norm of derivatives [68], as a regularization technique for image restoration. The problem can be modelled as follows; let $\Omega \subset \mathbb{R}^d$, for $d = 1,2$ be a bounded open set with Lipschitz boundary, and $\mathcal{H} = L^2(\Omega)$. For $u \in L^1_{loc}(\Omega)$

$$V(u,\Omega) := \sup \left\{ \int_\Omega u \operatorname{div} \varphi \, dx : \varphi \in \left[C^1_c(\Omega)\right]^d, \|\varphi\|_\infty \le 1 \right\}$$

is the variation of u. Further, $u \in BV(\Omega)$, the space of bounded variation functions [1, 38], if and only if $V(u,\Omega) < \infty$. In this case, we denote $|D(u)|(\Omega) = V(u,\Omega)$. If $u \in W^{1,1}(\Omega)$ (the Sobolev space of L^1-functions with L^1-distributional derivatives), then $|D(u)|(\Omega) = \int_\Omega |\nabla u| \, dx$. We consider as in [16, 77] the minimization in $BV(\Omega)$ of the functional

$$\mathcal{J}(u) := \|Ku - g\|_2^2 + 2\alpha \, |D(u)| \, (\Omega), \tag{2.79}$$

where $K : L^2(\Omega) \to L^2(\Omega)$ is a bounded linear operator, $g \in L^2(\Omega)$ is a datum, and $\alpha > 0$ is a fixed *regularization parameter*. Several numerical strategies to perform efficiently total variation minimization have been proposed in the literature. However, we will discuss in the following only how to adapt an iteratively reweighted least square algorithm to this particular situation. For simplicity, we would like to work on a discrete setting and we refer to the course presented by Antonin Chambolle for more details related to the continuous setting [16, 43].

Let us fix the main notations. Since we are interested in a discrete setting we define the *discrete d-orthotope* $\Omega = \{x_1^1 < \ldots < x_{N_1}^1\} \times \ldots \times \{x_1^d < \ldots < x_{N_d}^d\} \subset \mathbb{R}^d$, $d \in \mathbb{N}$ and the considered function spaces are $\mathcal{H} = \mathbb{R}^{N_1 \times N_2 \times \ldots \times N_d}$, where $N_i \in \mathbb{N}$ for $i = 1, \ldots, d$. For $u \in \mathcal{H}$ we write $u = u(x_{\mathbf{i}})_{\mathbf{i} \in \mathcal{I}}$ with

$$\mathcal{I} := \prod_{k=1}^d \{1, \ldots, N_k\}$$

and

$$u(x_{\mathbf{i}}) = u(x^1_{i_1}, \ldots, x^d_{i_d})$$

where $i_k \in \{1, \dots, N_k\}$ and $(x_{\mathbf{i}})_{\mathbf{i}\in\mathcal{I}} \in \Omega$. Then we endow $\mathcal{H}$ with the norm

$$\|u\|_{\mathcal{H}} = \|u\|_2 = \left(\sum_{\mathbf{i}\in\mathcal{I}} |u(x_{\mathbf{i}})|^2\right)^{1/2} = \left(\sum_{x\in\Omega} |u(x)|^2\right)^{1/2}.$$

We define the scalar product of $u, v \in \mathcal{H}$ as

$$\langle u, v\rangle_{\mathcal{H}} = \sum_{\mathbf{i}\in\mathcal{I}} u(x_{\mathbf{i}})v(x_{\mathbf{i}})$$

and the scalar product of $p, q \in \mathcal{H}^d$ as

$$\langle p, q\rangle_{\mathcal{H}^d} = \sum_{\mathbf{i}\in\mathcal{I}} \langle p(x_{\mathbf{i}}), q(x_{\mathbf{i}})\rangle_{\mathbb{R}^d}$$

with $\langle y, z\rangle_{\mathbb{R}^d} = \sum_{j=1}^d y_j z_j$ for every $y = (y_1, \dots, y_d) \in \mathbb{R}^d$ and $z = (z_1, \dots, z_d) \in \mathbb{R}^d$. We will consider also other norms, in particular

$$\|u\|_p = \left(\sum_{\mathbf{i}\in\mathcal{I}} |u(x_{\mathbf{i}})|^p\right)^{1/p}, \quad 1 \le p < \infty,$$

and

$$\|u\|_\infty = \sup_{\mathbf{i}\in\mathcal{I}} |u(x_{\mathbf{i}})|.$$

We denote the discrete gradient ∇u by

$$(\nabla u)(x_{\mathbf{i}}) = ((\nabla u)^1(x_{\mathbf{i}}), \dots, (\nabla u)^d(x_{\mathbf{i}}))$$

with

$$(\nabla u)^j(x_{\mathbf{i}}) = \begin{cases} u(x^1_{i_1}, \dots, x^j_{i_j+1}, \dots, x^d_{i_d}) - u(x^1_{i_1}, \dots, x^j_{i_j}, \dots, x^d_{i_d}) & \text{if } i_j < N_j \\ 0 & \text{if } i_j = N_j \end{cases}$$

for all $j = 1, \dots, d$ and for all $\mathbf{i} = (i_1, \dots, i_d) \in \mathcal{I}$.
Let $\varphi : \mathbb{R} \to \mathbb{R}$, we define for $\omega \in \mathcal{H}^d$

$$\varphi(|\omega|)(\Omega) = \sum_{\mathbf{i}\in\mathcal{I}} \varphi(|\omega(x_{\mathbf{i}})|) = \sum_{x\in\Omega} \varphi(|\omega(x)|),$$

where $|y| = \sqrt{y_1^2 + \dots + y_d^2}$. In particular we define the *total variation* of u by setting $\varphi(s) = s$ and $\omega = \nabla u$, i.e.,

$$|\nabla u|(\Omega) := \sum_{\mathbf{i}\in\mathcal{I}} |\nabla u(x_{\mathbf{i}})| = \sum_{x\in\Omega} |\nabla u(x)|.$$

For an operator K we denote K^* its adjoint. Further we introduce the *discrete divergence* $\operatorname{div} : \mathcal{H}^d \to \mathcal{H}$ defined, in analogy with the continuous setting, by $\operatorname{div} = -\nabla^*$ (∇^* is the adjoint of the gradient ∇). The discrete divergence operator is explicitly given by

$$
\begin{aligned}
(\operatorname{div} p)(x_{\mathbf{i}}) \quad = & \begin{cases} p^1(x^1_{i_1},\dots,x^d_{i_d}) - p^1(x^1_{i_1-1},\dots,x^d_{i_d}) & \text{if } 1 < i_1 < N_1 \\ p^1(x^1_{i_1},\dots,x^d_{i_d}) & \text{if } i_1 = 1 \\ -p^1(x^1_{i_1-1},\dots,x^d_{i_d}) & \text{if } i_1 = N_1 \end{cases} \\
& + \ldots + \begin{cases} p^d(x^1_{i_1},\dots,x^d_{i_d}) - p^d(x^1_{i_1},\dots,x^d_{i_d-1}) & \text{if } 1 < i_d < N_d \\ p^d(x^1_{i_1},\dots,x^d_{i_d}) & \text{if } i_d = 1 \\ -p^d(x^1_{i_1},\dots,x^d_{i_d-1}) & \text{if } i_d = N_d, \end{cases}
\end{aligned}
$$

for every $p = (p^1,\dots,p^d) \in \mathcal{H}^d$ and for all $\mathbf{i} = (i_1,\dots,i_d) \in \mathcal{I}$. (Note that if we considered discrete domains Ω which are not discrete d-orthotopes, then the definitions of gradient and divergence operators should be adjusted accordingly.) We will use the symbol 1 to indicate the constant vector with entry values 1 and 1_D to indicate the characteristic function of the domain $D \subset \Omega$. We are interested in the minimization of the functional

$$\mathcal{J}(u) := \|Ku - g\|_2^2 + 2\alpha\, |\nabla(u)|\,(\Omega), \tag{2.80}$$

where $K \in \mathcal{L}(\mathcal{H})$ is a linear operator, $g \in \mathcal{H}$ is a datum, and $\alpha > 0$ is a fixed constant. In order to guarantee the existence of minimizers for (2.80) we assume that:

(C) $\mathcal{J}$ is coercive in $\mathcal{H}$, i.e., there exists a constant $C > 0$ such that $\{\mathcal{J} \le C\} := \{u \in \mathcal{H} : \mathcal{J}(u) \le C\}$ is bounded in $\mathcal{H}$.

It is well known that if $1 \notin \ker(K)$ then condition (C) is satisfied, see [77, Proposition 3.1], and we will assume this condition in the following.

Similarly to (2.34) for the minimization of the ℓ_1-norm, we consider the augmented functional

$$\mathcal{J}(u,w) := \|Ku - g\|_2^2 + \alpha \left(\sum_{x \in \Omega} w(x) |\nabla u(x)|^2 + \frac{1}{w(x)} \right). \tag{2.81}$$

We used again the notation $\mathcal{J}$ with the clear understanding that when applied to one variable only refers to (2.80), otherwise to (2.81). Then, as the IRLS method for compressed sensing, we consider the following

Algorithm 2. We initialize by taking $w^0 := 1$. We also set $1 > \epsilon > 0$. We then recursively define for $n = 0, 1, \ldots,$

$$u^{n+1} := \arg\min_{u \in \mathcal{H}} \mathcal{J}(u, w^n) \tag{2.82}$$

and

$$w^{n+1} := \arg\min_{\max(\varepsilon, \min(w, 1/\varepsilon))} \mathcal{J}(u^{n+1}, w). \tag{2.83}$$

Note that, by considering the Euler-Lagrange equations, (2.82) is equivalent to the solution of the following linear second order partial difference equation

$$\mathrm{div}\,(w^n \nabla u) - \frac{2}{\alpha} K^*(Ku - g) = 0, \tag{2.84}$$

which can be solved, e.g., by a preconditioned conjugate gradient method. Note that $\varepsilon \leq w^n \leq 1/\varepsilon$ and therefore the equation can be recasted into a symmetric positive definite linear system. Moreover, as perhaps already expected, the solution to (2.83) is explicitly computed by

$$w^{n+1} = \max\left(\varepsilon, \min\left(\frac{1}{|\nabla u^{n+1}|}, 1/\varepsilon\right)\right).$$

For the sake of the analysis of the convergence of this algorithm, let us introduce the following function:

$$\varphi_\varepsilon(z) = \begin{cases} \frac{1}{2\varepsilon} z^2 + \frac{\varepsilon}{2} & 0 \leq z \leq \varepsilon \\ z & \varepsilon \leq z \leq 1/\varepsilon \\ \frac{\varepsilon}{2} z^2 + \frac{1}{2\varepsilon} & z \geq 1/\varepsilon. \end{cases}$$

Note that

$$\varphi_\varepsilon(z) \geq |z|,$$

and

$$|z| = \lim_{\varepsilon \to 0} \varphi_\varepsilon(z).$$

We consider the following functional:

$$\mathcal{J}_\varepsilon(u) := \|Ku - g\|_2^2 + 2\alpha \varphi_\varepsilon(|\nabla(u)|)(\Omega), \tag{2.85}$$

Which is clearly approximating $\mathcal{J}$ from above, i.e.,

$$\mathcal{J}_\varepsilon(u) \geq \mathcal{J}(u), \text{ and } \lim_{\varepsilon \to 0} \mathcal{J}_\varepsilon(u) = \mathcal{J}(u). \tag{2.86}$$

Moreover, since $\mathcal{J}_\varepsilon$ is convex and smooth, by taking the Euler-Lagrange equations, we have that u_ε is a minimizer for $\mathcal{J}_\varepsilon$ if and only if

$$\operatorname{div}\left(\frac{\varphi_\varepsilon'(|\nabla u|)}{|\nabla u|}\nabla u\right) - \frac{2}{\alpha}K^*(Ku-g) = 0, \qquad (2.87)$$

We have the following result of convergence of the algorithm.

Theorem 2.12 *The sequence $(u^n)_{n\in\mathbb{N}}$ has subsequences that converge to a minimizer $u_\varepsilon := u^\infty$ of $\mathcal{J}_\varepsilon$. If the minimizer was unique, then the full sequence would converge to it.*

Proof. Observe that

$$\begin{aligned}\mathcal{J}(u^n,w^n) - \mathcal{J}(u^{n+1},w^{n+1}) &= \underbrace{\left(\mathcal{J}(u^n,w^n) - \mathcal{J}(u^{n+1},w^n)\right)}_{A_n} \\ &+ \underbrace{\left(\mathcal{J}(u^{n+1},w^n) - \mathcal{J}(u^{n+1},w^{n+1})\right)}_{B_n} \geq 0.\end{aligned}$$

Therefore $\mathcal{J}(u^n,w^n)$ is a nonincreasing sequence and moreover it is bounded from below, since

$$\inf_{\varepsilon_h \leq w \leq 1/\varepsilon_h}\left(\sum_{x\in\Omega} w(x)|\nabla u(x)|^2 + \frac{1}{w(x)}\right) \geq 0.$$

This implies that $\mathcal{J}(u^n,w^n)$ converges. Moreover, we can write

$$B_n = \sum_{x\in\Omega} c(w^n(x), |\nabla u^{n+1}(x)|) - c(w^{n+1}(x), |\nabla u^{n+1}(x)|),$$

where $c(t,z) := tz^2 + \frac{1}{t}$. By Taylor's formula, we have

$$c(w^n,z) = c(w^{n+1},z) + \frac{\partial c}{\partial t}(w^{n+1},z)(w^n - w^{n+1}) + \frac{1}{2}\frac{\partial^2 c}{\partial t^2}(\xi,z)|w^n - w^{n+1}|^2,$$

for $\xi \in \operatorname{conv}(w^n, w^{n+1})$. By definition of w^{n+1}, and taking into account that $\varepsilon \leq w^{n+1} \leq \frac{1}{\varepsilon}$, we have

$$\frac{\partial c}{\partial t}(w^{n+1}, |\nabla u^{n+1}(x)|)(w^n - w^{n+1}) \geq 0,$$

and $\frac{\partial^2 c}{\partial t^2}(t,z) = \frac{2}{t^3} \geq 2\varepsilon^3$, for any $t \leq 1/\varepsilon$. This implies that

$$\mathcal{J}(u^n,w^n) - \mathcal{J}(u^{n+1},w^{n+1}) \geq B_n \geq \varepsilon^3 \sum_{x\in\Omega}|w^n(x) - w^{n+1}(x)|^2,$$

and since $\mathcal{J}(u^n, w^n)$ is convergent, we have

$$\|w^n - w^{n+1}\|_2 \to 0, \tag{2.88}$$

for $n \to \infty$. Since u^{n+1} is a minimizer of $\mathcal{J}(u, w^n)$ it solves the following system of variational equations

$$\sum_{x\in\Omega} \left(w^n \nabla u^{n+1}(x) \cdot \nabla \varphi(x) + \frac{2}{\alpha}(Ku^{n+1} - g)(x) K\varphi(x) \right) \tag{2.89}$$

for all $\varphi \in \mathcal{H}$. Therefore we can write

$$\begin{aligned} &\sum_{x\in\Omega} \left(w^{n+1} \nabla u^{n+1}(x) \cdot \nabla \varphi(x) + \frac{2}{\alpha}(Ku^{n+1} - g)(x) K\varphi(x) \right) \\ = \; &\sum_{x\in\Omega} (w^{n+1} - w^n) \nabla u^{n+1}(x) \cdot \nabla \varphi(x), \end{aligned}$$

and

$$\begin{aligned} &\left| \sum_{x\in\Omega} \left(w^{n+1} \nabla u^{n+1}(x) \cdot \nabla \varphi(x) + \frac{2}{\alpha}(Ku^{n+1} - g)(x) K\varphi(x) \right) \right| \\ \le \; & C \|w^{n+1} - w^n\|_2 \|\nabla u^{n+1}\|_2 \|\nabla \varphi\|_2. \end{aligned}$$

(Remind that every norm is equivalent in finite dimensions!) By monotonicity of $(\mathcal{J}(u^{n+1}, w^{n+1}))_n$, and since $w^{n+1} = \frac{\varphi'_\varepsilon(|\nabla u^{n+1}|)}{|\nabla u^{n+1}|}$, we have

$$\mathcal{J}(u^0, w^0) \ge \mathcal{J}(u^{n+1}, w^{n+1}) = \mathcal{J}_\varepsilon(u^{n+1}) \ge \mathcal{J}(u^{n+1}) \ge c_1 |\nabla u|(\Omega) \ge c_2 \|\nabla u^{n+1}\|_2.$$

Moreover, since $\mathcal{J}_\varepsilon(u^{n+1}) \ge \mathcal{J}(u^{n+1})$ and $\mathcal{J}$ is coercive, by condition (C), we have that $\|u^{n+1}\|_2$ and $\|\nabla u^{n+1}\|_2$ are bounded uniformly with respect to n. Therefore, using (2.88), we can conclude that

$$\begin{aligned} &\left| \sum_{x\in\Omega} \left(w^{n+1} \nabla u^{n+1}(x) \cdot \nabla \varphi(x) + \frac{2}{\alpha}(Ku^{n+1} - g)(x) K\varphi(x) \right) \right| \\ \le \; & \|w^{n+1} - w^n\|_2 \|\nabla u^{n+1}\|_2 \|\nabla \varphi\|_2 \to 0, \end{aligned}$$

for $n \to \infty$, and there exists a subsequence $(u^{(n_k)})_k$ that converges in $\mathcal{H}$ to a function u^∞. Since $w^{n+1} = \frac{\varphi'_\varepsilon(|\nabla u^{n+1}|)}{|\nabla u^{n+1}|}$, and by taking the limit for $n \to \infty$, we obtain that in fact

$$\operatorname{div}\left(\frac{\varphi'_\varepsilon(|\nabla u^\infty|)}{|\nabla u^\infty|} \nabla u^\infty \right) - \frac{2}{\alpha} K^*(Ku^\infty - g) = 0, \tag{2.90}$$

The latter are the Euler-Lagrange equations associated to the functional $\mathcal{J}_\varepsilon$ and therefore u^∞ is a minimizer of $\mathcal{J}_\varepsilon$. □

It is left as a – not simple – exercise the proof of the following result. One has to make use of the monotonicity of the approximation (2.86), of the coerciveness of $\mathcal{J}$ (property (C)), and of the continuity of $\mathcal{J}_\varepsilon$. See also [25] for more general tools from so-called Γ-*convergence* for achieving such variational limits.

Proposition 2.13 *Let us assume that $(\varepsilon_h)_h$ is a sequence of positive numbers monotonically converging to zero. The accumulation points of the sequence $(u_{\varepsilon_h})_h$ of minimizers of $\mathcal{J}_{\varepsilon_h}$ are minimizers of $\mathcal{J}$.*

Let us note a few differences between Algorithm 1 and Algorithm 2. In Algorithm 1 we have been able to establish a rule of up-dating the parameter ϵ according to the iterations. This was not done for Algorithm 2, where we consider the limit for $\varepsilon \to 0$ only at the end. It is an interesting open question how can we simultaneously address a choice of a decreasing sequence $(\varepsilon_n)_n$ during the iterations and show directly the convergence of a minimizer of $\mathcal{J}$.

Figure 2.5 Fragments of the frescoes.

A relevant application

In this section we would like to report the surprising applicative results from the work [43], where the IRLS for total variation minimization has been used for vector-valued functions.

On March 11, 1944, the famous Eremitani Church in Padua (Italy) was destroyed in an Allied bombing along with the inestimable frescoes by Andrea Mantegna *et al.* contained in the Ovetari Chapel. In the last 60 years, several attempts have been made to restore the fresco fragments (Figure 2.5) by traditional methods, but without much success. An efficient pattern recognition algorithm was used to map the original position and orientation of the fragments, based on comparisons with an old gray level image of the fresco prior to the damage. This innovative technique allowed for the partial reconstruction of the frescoes. In Figure 2.6 we show a sample of the results due to this computer-assisted restoration.

Figure 2.6 On the left the scene "St. James Led to Martyrdom" with a few fragments localized by the computer assisted recollocation. On the right, we point out a particular of the scene.

Unfortunately, the surface covered by the colored fragments is only 77 m^2, while the original area was of several hundreds. This means that we could reconstruct so far only a fraction (less than 8%) of this inestimable artwork. In particular the original color of the blanks is not known. This begs the question of whether it is possible to estimate *mathematically* the original colors of the frescoes by making use of the potential information given by the available fragments and the gray level of the pictures taken before the damage.

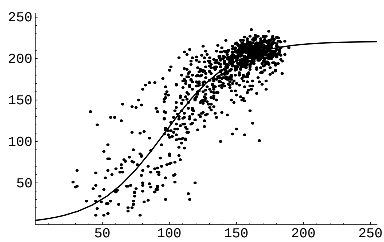


Figure 2.7 Estimate of the nonlinear curve L from a distribution of points with coordinates given by the linear combination $\xi_1 r + \xi_2 g + \xi_3 b$ of the (r, g, b) color fragments (abscissa) and by the corresponding underlying gray level of the original photographs dated to 1920 (ordinate).

We get our inspiration from physics: it is common experience that in an inhomogeneous material heat diffuses anisotropically from heat sources; the mathematical (partial differential) equations that govern this phenomenon are well-known. In turn similar equations (see (2.84)) can be used to diffuse the color (instead of the heat) from the ‘color-sources’, which are the placed fragments, keeping into account the inhomogeneity due to the gradients provided by the known gray levels. We describe formally the model as follows. A color image can be modeled as a function $u : \Omega \subset \mathbb{R}^2 \to \mathbb{R}^3_+$, so that, to each “point” $\mathbf{x} \in \Omega$ of the image, one associates the vector $u(\mathbf{x}) = (r(\mathbf{x}), g(\mathbf{x}), b(\mathbf{x})) \in \mathbb{R}^3_+$ of the color represented by the different channels, for instance, red, green, and blue. The gray level of an image can be described as non-linear projection of the colors $\mathcal{L}(r, g, b) := L(\xi_1 r + \xi_2 g + \xi_3 b)$, $(r, g, b) \in \mathbb{R}^3_+$, where $\xi_1, \xi_2, \xi_3 > 0$, $\xi_1 + \xi_2 + \xi_3 = 1$, and $L : \mathbb{R} \to \mathbb{R}$ is a suitable non-negative increasing function. For example Figure 2.7. describes the typical shape of an L function, which is estimated by fitting a distribution of data from the real color fragments, see Figure 2.6. However, it is always possible to re-equalize the grey level in such a way $L(\xi) = \xi$. In this case the function $\mathcal{L}$ is simply a linear projection. The

Figure 2.8 The first column illustrates two different data for the recolorization problem. The second column illustrates the corresponding recolorized solution.

recolorization is modeled as the minimum (color image) solution of the functional

$$\mathcal{J}(u) = \mu \int_{\Omega\setminus D} |u(x) - \bar{u}(x)|^2 dx + \int_D |\mathcal{L}(u(x)) - \bar{v}(x)|^2 dx + \int_\Omega \sum_{\ell=1}^{3} |\nabla u^\ell(x)| dx, \tag{2.91}$$

where we want to reconstruct the vector valued function $u := (u^1, u^2, u^3) : \Omega \subset \mathbb{R}^2 \to \mathbb{R}^3$ (for RGB images) from a given observed couple of color/gray level functions $(\bar{u}, \bar{v})$. The observed function $\bar{u}$ is assumed to represent correct information, *e.g.*, the given colors, on $\Omega\setminus D$, and $\bar{v}$ the result of the *nonlinear projection* $\mathcal{L} : \mathbb{R}^3 \to \mathbb{R}$, *e.g.*, the gray level, on D. Note also that we consider the total variation of each of the three color components. In case $\mathcal{L}$ is linear (e.g., after re-equalization of the gray level), the functional $\mathcal{J}$, suitably discretized, can be recasted into the form (2.80). Hence the method previously described can be applied. See Figure 2.8 for a sample of the mathematical recolorization in the real-life problem.

2.1.4 Iterative Hard Thresholding

In this section we address the following

> **Algorithm 3.** We initialize by taking $x^0 = 0$. We iterate
>
> $$x^{n+1} = \mathbb{H}_k(x^n + A^*(y - Ax^n)), \tag{2.92}$$
>
> where
>
> $$\mathbb{H}_k(x) = x_{[k]}, \tag{2.93}$$
>
> is the operator which returns the best k-term approximation to x, see (1.2).

Note that if x^* is k-sparse and $Ax^* = y$, then x^* is a fixed point of

$$x^* = \mathbb{H}_k(x^* + A^*(y - Ax^*)).$$

This algorithm can be seen as a minimizing method for the functional

$$\mathcal{J}(x) = \|y - Ax\|^2_{\ell_2^N} + 2\alpha \|x\|_{\ell_0^N}, \tag{2.94}$$

for a suitable $\alpha = \alpha(k) > 0$ or equivalently for the solution of the optimization problem

$$\min_x \|y - Ax\|^2_{\ell_2^N} \text{ subject to } \|x\|_{\ell_0^N} \leq k.$$

Actually, it was shown [6] that if $\|A\| < 1$ then this algorithm converges to a local minimizer of (2.94). We would like to analyze this algorithm following [7] in the case A satisfies the RIP. We start with a few technical lemmas which shed light on fundamental properties of RIP matrices and sparse approximations, as established in [73].

Lemma 2.14 *For all index sets $\Lambda \subset \{1,\dots,N\}$ and all A for which the RIP holds with order $k = |\Lambda|$, we have*

$$\|A_\Lambda^* y\|_{\ell_2^N} \leq (1+\delta_k)\|y\|_{\ell_2^N}, \tag{2.95}$$

$$(1-\delta_k)^2\|x_\Lambda\|_{\ell_2^N} \leq \|A_\Lambda^* A_\Lambda x_\Lambda\|_{\ell_2^N} \leq (1+\delta_k)^2\|x_\Lambda\|_{\ell_2^N} \tag{2.96}$$

and

$$\|(I - A_\Lambda^* A_\Lambda)x_\Lambda\|_{\ell_2^N} \leq \delta_k^2 \|x_\Lambda\|_{\ell_2^N}. \tag{2.97}$$

Furthermore, for two disjoint sets Λ_1 and Λ_2 and all A for which the RIP holds with order $k = |\Lambda_1 \cup \Lambda_2|$,

$$\|A_{\Lambda_1}^* A_{\Lambda_2} x_{\Lambda_2}\|_{\ell_2^N} \leq \delta_k^2 \|x_\Lambda\|_{\ell_2^N}. \tag{2.98}$$

Proof. The proof of (2.95)-(2.97) is straightforward and it is left to the reader. For (2.98), just note that $A_{\Lambda_1}^* A_{\Lambda_2}$ is a submatrix of $A_{\Lambda_1\cup\Lambda_2}^* A_{\Lambda_1\cup\Lambda_2} - I$, and therefore $\|A_{\Lambda_1}^* A_{\Lambda_2}\| \leq \|I - A_{\Lambda_1\cup\Lambda_2}^* A_{\Lambda_1\cup\Lambda_2}\|$. One concludes by (2.97). □

Lemma 2.15 *Suppose the matrix A satisfies the RIP of order k with constant $\delta_k > 0$. Then for all vectors x, the following bound holds*

$$\|Ax\|_{\ell_2^N} \leq (1+\delta_k)\left(\|x\|_{\ell_2^N} + \frac{\|x\|_{\ell_1^N}}{k^{1/2}}\right). \tag{2.99}$$

Proof. In this proof we consider $\mathbb{R}^N$ as a Banach space endowed with several different norms. In particular, the statement of the lemma can be regarded as a result about the operator norm of A as a map between two Banach spaces. For a set $I \subset \{1,2,\dots,N\}$, we consider $B_{\ell_2^I}$ the ℓ_2-norm unit ball of vectors supported in I and we define the convex set

$$S = \operatorname{conv}\left\{\bigcup_{|I|\leq k} B_{\ell_2^I}\right\} \subset \mathbb{R}^N.$$

The set S can be consider the unit ball of a norm $\|\cdot\|_S$ on $\mathbb{R}^N$, and the upper bound of the RIP property a statement about the norm of A between $S := (\mathbb{R}^N, \|\cdot\|_S)$ and $\ell_2^N := (\ell_2^N, \|\cdot\|_{\ell_2^N})$, i.e., (with a slight abuse of notation)

$$\|A\|_{S\to\ell_2^N} = \max_{x\in S}\|Ax\|_{\ell_2^N} \leq (1+\delta_k).$$

Let us define a second convex body,

$$K = \left\{x : \|x\|_{\ell_2^N} + \frac{\|x\|_{\ell_1^N}}{k^{1/2}} \leq 1\right\} \subset \mathbb{R}^N,$$

and we consider, analogously (and with the same abuse of notation), the operator norm

$$\|A\|_{K\to\ell_2^N} = \max_{x\in K}\|Ax\|_{\ell_2^N}.$$

The content of the lemma is the claim that

$$\|A\|_{K\to\ell_2^N} \le \|A\|_{S\to\ell_2^N}.$$

To establish this point it is sufficient to check that $K \subset S$. To do that we do prove the reverse inclusion of the polar sets, i.e.,

$$S^\circ \subset K^\circ.$$

Remind that the polar set of $\Omega \subset \mathbb{R}^N$ is

$$\Omega^\circ := \{y : \sup_{x\in\Omega}\langle x, y\rangle \le 1\}.$$

If Ω is convex than $\Omega^{\circ\circ} = \Omega$. Moreover, the norm associated to a convex body Ω can also be expressed by

$$\|x\|_\Omega = \sup_{y\in\Omega^\circ}\langle x, y\rangle.$$

In particular, the norm with unit ball S° is easily calculated as

$$\|x\|_{S^\circ} = \max_{|I|\le k}\|x_I\|_2.$$

Now, consider a vector x in the unit ball S° and let I be the support of the k-best approximation of x. We must have

$$\|x_{I^c}\|_\infty \le \frac{1}{\sqrt{k}},$$

otherwise $|x_i| > \frac{1}{\sqrt{k}}$ for all $i \in I$, but then $\|x\|_{S^\circ} \ge \|x_I\|_2 > 1$, a contradiction. Therefore, we can write

$$x = x_I + x_{I^c} \in B_{\ell_2^N} + \frac{1}{\sqrt{k}} B_{\ell_\infty^N}.$$

But the set on the right-hand side is precisely K° since

$$\begin{aligned}\sup_{y\in K^\circ}\langle x, y\rangle = \|x\|_K &= \|x\|_{\ell_2^N} + \frac{\|x\|_{\ell_1^N}}{k^{1/2}} \\ &= \sup_{y\in B_{\ell_2^N}}\langle x, y\rangle + \sup_{z\in\frac{1}{k^{1/2}}B_{\ell_\infty^N}}\langle x, z\rangle = \sup_{y\in B_{\ell_2^N}+\frac{1}{k^{1/2}}B_{\ell_\infty^N}}\langle x, y\rangle.\end{aligned}$$

In summary $S^\circ \subset K^\circ$. □

Lemma 2.16 *For any x we denote $x^{[k]} = x - x_{[k]}$. Let*

$$y = Ax + e = Ax_{[k]} + Ax^{[k]} + e = Ax_{[k]} + \tilde{e}.$$

If A has the RIP of order k, then the norm of the error $\tilde{e}$ can be bounded by

$$\|\tilde{e}\|_{\ell_2^N} \le (1+\delta_k)\left(\sigma_k(x)_{\ell_2^N} + \frac{\sigma_k(x)_{\ell_1^N}}{\sqrt{k}}\right) + \|e\|_{\ell_2^N}. \tag{2.100}$$

Proof. Decompose $x = x_{[k]} + x^{[k]}$ and $\tilde{e} = Ax^{[k]} + e$. To compute the norm of the error term, we simply apply the triangle inequality and Lemma 2.15. □

After this collection of technical results, we are able to establish a first convergence result.

Theorem 2.17 *Given a noisy observation $y = Ax + e$, where x is k-sparse. If A has the RIP or order $3k$ and constant $\delta_{3k}^2 < \frac{1}{\sqrt{32}}$, then, at iteration n, Algorithm 2 will recover an approximation x^n satisfying*

$$\|x - x^n\|_{\ell_2^N} \le 2^{-n}\|x\|_{\ell_2^N} + 5\|e\|_{\ell_2^N}. \tag{2.101}$$

Furthermore, after at most

$$n^* = \left\lceil \log_2\left(\frac{\|x\|_{\ell_2^N}}{\|e\|_{\ell_2^N}}\right)\right\rceil \tag{2.102}$$

iterations, the algorithm estimates x with accuracy

$$\|x - x^{n^*}\|_{\ell_2^N} \le 6\|e\|_{\ell_2^N}. \tag{2.103}$$

Proof. Let us denote $z^n := x^n + A^*(y - Ax^n)$, $r^n = x - x^n$, and $B^n := \operatorname{supp}(r^n))$. By triangle inequality we can write

$$\|x - x^{n+1}\|_{\ell_2^N} \le \|x_{B^{n+1}} - z^n_{B^{n+1}}\|_{\ell_2^N} + \|x^{n+1}_{B^{n+1}} - z^n_{B^{n+1}}\|_{\ell_2^N}.$$

By application of $\mathbb{H}_k$, $x^{n+1} = \mathbb{H}_k(z^n)$. This implies $\|x^{n+1}_{B^{n+1}} - z^n_{B^{n+1}}\|_{\ell_2^N} \le \|x_{B^{n+1}} - z^n_{B^{n+1}}\|_{\ell_2^N}$, and

$$\|x - x^{n+1}\|_{\ell_2^N} \le 2\|x_{B^{n+1}} - z^n_{B^{n+1}}\|_{\ell_2^N}.$$

We can also write

$$z^n_{B^{n+1}} = x^n_{B^{n+1}} + A^*_{B^{n+1}}Ar^n + A^*_{B^{n+1}}e.$$

We then have

$$\begin{aligned}
\|x - x^{n+1}\|_{\ell_2^N} &\le 2\|x_{B^{n+1}} - x^n_{B^{n+1}} - A^*_{B^{n+1}}Ar^n - A^*_{B^{n+1}}e\|_{\ell_2^N} \\
&\le 2\|(I - A^*_{B^{n+1}}A_{B^{n+1}})r^n\|_{\ell_2^N} + 2\|A^*_{B^{n+1}}A_{B^n \setminus B^{n+1}}r^n\|_{\ell_2^N} + 2\|A^*_{B^{n+1}}e\|_{\ell_2^N}
\end{aligned}$$

Note that $|B^n| \leq 2k$ and that $|B^{n+1} \cup B^n| \leq 3k$. By an application of the bounds in Lemma 2.14, and by using the fact that $\delta_{2k} \leq \delta_{3k}$ (note that a $2k$-sparse vector is also $3k$-sparse)

$$\|r^{n+1}\|_{\ell_2^N} \leq 2\delta_{2k}^2 \|r^n_{B^{n+1}}\|_{\ell_2^N} + 2\delta_{3k}^2 \|r^n_{B^n \setminus B^{n+1}}\|_{\ell_2^N} + 2(1+\delta_{2k})\|e\|_{\ell_2^N}$$

Moreover $\|r^n_{B^{n+1}}\|_{\ell_2^N} + \|r^n_{B^n \setminus B^{n+1}}\|_{\ell_2^N} \leq \sqrt{2}\|r^n\|_{\ell_2^N}$. Therefore we have the bound

$$\|r^{n+1}\|_{\ell_2^N} \leq 2\sqrt{2}\delta_{3k}^2 \|r^n\|_{\ell_2^N} + 2(1+\delta_{3k})\|e\|_{\ell_2^N}.$$

By assumption $\delta_{3k}^2 < \frac{1}{\sqrt{32}}$ and $2\sqrt{2}\delta_{3k}^2 < \frac{1}{2}$. (Note that here we could simply choose any value $\delta_{3k}^2 < \frac{1}{\sqrt{8}}$ and obtain a slightly different estimate!) The we get the recursion

$$\|r^{n+1}\|_{\ell_2^N} \leq 2^{-1}\|r^n\|_{\ell_2^N} + 2.17\|e\|_{\ell_2^N},$$

which iterated (note that $x^0 = 0$ and $2.17\sum_{n=0}^{\infty} 2^{-n} \leq 4.34$) gives

$$\|r^{n+1}\|_{\ell_2^N} \leq 2^{-n}\|x\|_{\ell_2^N} + 4.34\|e\|_{\ell_2^N}.$$

This is precisely the bound we were looking for. The rest of the statements of the theorem are left as an exercise. □

We have also the following result.

Corollary 2.18 *Given a noisy observation $y = Ax + e$, where x is* an arbitrary vector. *If A has the RIP or order $3k$ and constant $\delta_{3k}^2 < \frac{1}{32}$, then, at iteration n, Algorithm 2 will recover an approximation x^n satisfying*

$$\|x - x^n\|_{\ell_2^N} \leq 2^{-n}\|x\|_{\ell_2^N} + 6\left(\sigma_k(x)_{\ell_2^N} + \frac{\sigma_k(x)_{\ell_1^N}}{\sqrt{k}} + \|e\|_{\ell_2^N}\right). \tag{2.104}$$

Furthermore, after at most

$$n^* = \left\lceil \log_2\left(\frac{\|x\|_{\ell_2^N}}{\|e\|_{\ell_2^N}}\right)\right\rceil \tag{2.105}$$

iterations, the algorithm estimates x with accuracy

$$\|x - x^{n^*}\|_{\ell_2^N} \leq 7\left(\sigma_k(x)_{\ell_2^N} + \frac{\sigma_k(x)_{\ell_1^N}}{\sqrt{k}} + \|e\|_{\ell_2^N}\right). \tag{2.106}$$

Proof. We first note

$$\|x - x^n\|_{\ell_2^N} \leq \sigma_k(x)_{\ell_2^N} + \|x_{[k]} - x^n\|_{\ell_2^N}.$$

The proof now follows by bounding $\|x_{[k]} - x^n\|_{\ell_2^N}$. For this we simply apply Theorem 2.17 to $x_{[k]}$ with $\tilde{e}$ instead of e, and use Lemma 2.16 to bound $\|\tilde{e}\|_{\ell_2^N}$. The rest is left as an exercise. □

A brief discussion

This algorithm has error reduction guarantee from the very beginning of the iteration, and it is robust to noise, i.e., an estimate of the type (1.13) holds. Moreover, each iteration costs mainly as much as an application of A^*A. At first glance this algorithm is greatly superior with respect to IRLS; however, we have to stress that IRLS can converge superlinearly and a fine analysis of its complexity is widely open.

3 Numerical Methods for Sparse Recovery

In the previous chapters we put most of the emphasis on finite dimensional linear problems (also of relatively small size) where the model matrix A has the RIP or the NSP. This setting is suitable for applications in coding/decoding or compressed acquisition problems, hence from human-made problems coming from technology, while it does not fit many possible applications where we are interested to recover quantities from partial real-life measurements. In this case we may need to work with large dimensional problems (even infinite dimensional) where the model linear (or nonlinear) operator which defines the measurements has not such nice properties as the RIP and NSP. A typical example of such situation is the one reported in the previous chapter related to the color recovery from a real-life restoration problem.

Here and later we are concerned with the more general setting and the efficient minimization of functionals of the type:

$$\mathcal{J}(u) := \|Ku - y\|_Y^2 + 2\|(\langle u, \tilde{\psi}_\lambda \rangle)_{\lambda \in \mathcal{I}}\|_{\ell_{1,\alpha}(\mathcal{I})}, \tag{3.107}$$

where $K : X \to Y$ is a bounded linear operator acting between two separable Hilbert spaces X and Y, $y \in Y$ is a given measurement, and $\Psi := \{\psi_\lambda\}_{\lambda \in \mathcal{I}}$ is a prescribed countable basis for X with associated dual $\tilde{\Psi} := \{\tilde{\psi}_\lambda\}_{\lambda \in \mathcal{I}}$. For $1 \leq p < \infty$, the sequence norm $\|\mathbf{u}\|_{\ell_{p,\alpha}(\mathcal{I})} := \left(\sum_{\lambda \in \mathcal{I}} |u_\lambda|^p \alpha_\lambda\right)^{1/p}$ is the usual norm for weighted p-summable sequences, with weight $\alpha = (\alpha_\lambda)_{\lambda \in \mathcal{I}} \in \mathbb{R}_+^{\mathcal{I}}$, such that $\alpha_\lambda \geq \bar{\alpha} > 0$. Associated to the basis, we are given the synthesis map $F : \ell_2(\mathcal{I}) \to X$ defined by

$$F\mathbf{u} := \sum_{\lambda \in \mathcal{I}} u_\lambda \psi_\lambda, \quad \mathbf{u} \in \ell_2(\mathcal{I}). \tag{3.108}$$

We can re-formulate equivalently the functional in terms of sequences in $\ell_2(\mathcal{I})$ as follows:

$$\mathcal{J}(\mathbf{u}) := \mathcal{J}_\alpha(\mathbf{u}) = \|(K \circ F)\mathbf{u} - y\|_Y^2 + 2\|\mathbf{u}\|_{\ell_{1,\alpha}(\mathcal{I})}. \tag{3.109}$$

For ease of notation let us write $A := K \circ F$. Such a functional turns out to be very useful in many practical problems, where one cannot observe directly the quantities of most interest; instead their values have to be inferred from their effect on observable quantities. When this relationship between the observable y and the interesting quantity u is (approximately) linear the situation can be modeled mathematically by the equation

$$y = Ku\,, \tag{3.110}$$

If K is a "nice" (e.g., well-conditioned), easily invertible operator, and if the data y are free of noise, then this is a well-known task which can be addressed with standard numerical analysis methods. Often, however, the mapping K is not invertible or ill-conditioned. Moreover, typically (3.110) is only an idealized version in which noise has been neglected; a more accurate model is

$$y = Ku + e\,, \tag{3.111}$$

in which the data are corrupted by an (unknown) noise e. In order to deal with this type of reconstruction problem a *regularization* mechanism is required [37]. Regularization techniques try, as much as possible, to take advantage of (often vague) prior knowledge one may have about the nature of u, which is embedded into the model. The approach modelled by the functional $\mathcal{J}$ in (3.107) is indeed tailored to the case when u can be represented by a *sparse* expansion, i.e., when u can be represented by a series expansion (3.108) with respect to an orthonormal basis (or a frame [27]) that has only a small number of large coefficients. The previous chapters should convince the reader that imposing an additional ℓ_1-norm term as in (3.108) has indeed the effect of sparsifying possible solutions. Hence, we model the sparsity constraint by a regularizing ℓ_1-term in the functional to be minimized; of course, we could consider also a minimization of the type (2.94), but that has the disadvantage of being nonconvex and not being necessarily robust to noise, when no RIP conditions are imposed on the model operator A.

In the following we will not use anymore the bold form $\mathbf{u}$ for a sequence in $\ell_2(\mathcal{I})$, since here and later we will exclusively work with the space $\ell_2(\mathcal{I})$.

3.1 Iterative Soft-Thresholding in Hilbert Spaces

Several authors have proposed independently an iterative soft-thresholding algorithm to approximate minimizers $u^* := u^*_\alpha$ of the functional in (3.108), see [35, 41, 70, 71]. More precisely, u^* is the limit of sequences $u^{(n)}$ defined recursively by

$$u^{(n+1)} = \mathbb{S}_\alpha \left[u^{(n)} + A^* y - A^* A u^{(n)} \right] \;, \tag{3.112}$$

starting from an arbitrary $u^{(0)}$, where $\mathbb{S}_\alpha$ is the soft-thresholding operation defined by $\mathbb{S}_\alpha(u)_\lambda = S_{\alpha_\lambda}(u_\lambda)$ with

$$S_\tau(x) = \begin{cases} x - \tau & x > \tau \\ 0 & |x| \le \tau \\ x + \tau & x < -\tau \end{cases}. \tag{3.113}$$

This is our starting point and the reference iteration on which we want to work out several innovations. Strong convergence of this algorithm was proved in [28], under the assumption that $\|A\| < 1$ (actually, convergence can be shown also for $\|A\| < \sqrt{2}$ [21]; nevertheless, the condition $\|A\| < 1$ is by no means a restriction, since it can always be met by a suitable rescaling of the functional J, in particular of K, y, and α). Soft-thresholding plays a role in this problem because it leads to the unique minimizer of a functional combining ℓ_2 and ℓ_1-norms, i.e., (see Lemma 3.1)

$$\mathbb{S}_\alpha(a) = \arg\min_{u \in \ell_2(\mathcal{I})} \left(\|u - a\|^2 + 2\|u\|_{1,\alpha} \right). \tag{3.114}$$

We will call the iteration (3.112) the *iterative soft-thresholding algorithm* or the *thresholded Landweber iteration* (ISTA).

In this section we would like to provide the analysis of the convergence of this algorithm. Due to the lack of assumptions such as the RIP or the NSP, the methods we use comes exclusively from convex analysis and we cannot take advantage of relatively simple estimates as we did for the convergence analysis of Algorithms 1,3.

3.1.1 The Surrogate Functional

The first relevant observation is that the algorithm can be recasted into an iterated minimization of a properly augmented functional, which we call the *surrogate functional* of $\mathcal{J}$, and it is defined by

$$\mathcal{J}^S(u,a) := \|Au - y\|_Y^2 + 2\|u\|_{\ell_{1,\alpha}(\mathcal{I})} + \|u - a\|_{\ell_2(\mathcal{I})}^2 - \|Au - Aa\|_Y^2. \tag{3.115}$$

Assume here and later that $\|A\| < 1$. Observe that

$$\|u - a\|_{\ell_2(\mathcal{I})}^2 - \|Au - Aa\|_Y^2 \ge C\|u - a\|_{\ell_2(\mathcal{I})}^2, \tag{3.116}$$

for $C = (1 - \|A\|^2) > 0$. Hence

$$\mathcal{J}(u) = \mathcal{J}^S(u,u) \le \mathcal{J}^S(u,a), \tag{3.117}$$

and

$$\mathcal{J}^S(u,a) - \mathcal{J}^S(u,u) \ge C\|u - a\|_{\ell_2(\mathcal{I})}^2. \tag{3.118}$$

In particular, $\mathcal{J}^S$ is strictly convex with respect to u and it has a unique minimizer with respect to u once a is fixed. We have the following technical lemmas.

Lemma 3.1 *The soft-thresholding operator is the solution of the following optimization problem:*

$$\mathbb{S}_\alpha(a) = \arg\min_{u \in \ell_2(\mathcal{I})} \left(\|u - a\|^2 + 2\|u\|_{1,\alpha} \right).$$

Proof. By componentwise optimization, we can reduce the problem to a scalar problem, i.e., we need to show that

$$S_{\alpha_\lambda}(a_\lambda) = \operatorname*{argmin}_x (x - a_\lambda)^2 + 2\alpha_\lambda |x|,$$

which is shown by a simple direct computation. Let x^* be the minimizer. It is clear that $\operatorname{sgn}(x^*)\operatorname{sgn}(a_\lambda) \geq 0$ otherwise the function is increased.. Hence we need to optimize $(x-a_\lambda)^2 + 2\alpha_\lambda \operatorname{sgn}(a_\lambda)x$ which has minimum at $\bar{x} = (a_\lambda - \operatorname{sgn}(a_\lambda)\alpha_\lambda)$. If $|a_\lambda| > \alpha_\lambda$ than $x^* = \bar{x}$. Otherwise $\operatorname{sgn}(\bar{x})\operatorname{sgn}(a_\lambda) < 0$ and $\bar{x}$ cannot be the minimizer, and we have to choose $x^* = 0$. □

Lemma 3.2 *We can express the optimization of $\mathcal{J}^S(u,a)$ with respect to u explicitly by*

$$\mathbb{S}_\alpha(a + A^*(y - Aa)) = \arg\min_{u \in \ell_2(\mathcal{I})} \mathcal{J}^S(u,a).$$

Proof. By developing the norm squares in (3.115) it is a straightforward computation to show

$$\mathcal{J}^S(u,a) = \|u - (a + A^*(y - Aa))\|^2_{\ell_2(\mathcal{I})} + 2\|u\|_{1,\alpha} + \Phi(a, A, y),$$

where $\Phi(a, A, y)$ is a function which does not depend on u. The statement follows now from an application of Lemma 3.1 and by the observation that the addition of constants to a functional does not modify its minimizer. □

3.1.2 The Algorithm and Preliminary Convergence Properties

By Lemma 3.2 we achieve

Algorithm 4. We initialize by taking any $u^{(0)} \in \ell_2(\mathcal{I})$. We iterate

$$\begin{aligned} u^{(n+1)} &= \mathbb{S}_\alpha\left[u^{(n)} + A^*y - A^*Au^{(n)}\right] \\ &= \arg\min_{u \in \ell_2(\mathcal{I})} \mathcal{J}^S(u, u^{(n)}). \end{aligned}$$

Lemma 3.3 *The sequence* $\mathcal{J}(u^{(n)})$ *is nondecreasing. Moreover* $(u^{(n)})_n$ *is bounded in* $\ell_2(\mathcal{I})$ *and*

$$\lim_{n\to\infty} \|u^{(n+1)} - u^{(n)}\|^2_{\ell_2(\mathcal{I})} = 0. \tag{3.119}$$

Proof. Let us consider the estimates

$$\begin{aligned}\mathcal{J}(u^{(n)}) &= \mathcal{J}^S(u^{(n)}, u^{(n)})\\ &\geq \mathcal{J}^S(u^{(n+1)}, u^{(n)})\\ &\geq \mathcal{J}^S(u^{(n+1)}, u^{(n+1)}) = \mathcal{J}(u^{(n+1)}),\end{aligned}$$

Hence, the sequence $\mathcal{J}(u^{(n)})$ is nondecreasing, and

$$\mathcal{J}(u^{(0)}) \geq \mathcal{J}(u^{(n)}) \geq 2\bar{\alpha}\|u^{(n)}\|_{\ell_1(\mathcal{I})} \geq 2\bar{\alpha}\|u^{(n)}\|_{\ell_2(\mathcal{I})}.$$

Therefore, $(u^{(n)})_n$ is bounded in $\ell_2(\mathcal{I})$. By (3.192), we have

$$\mathcal{J}(u^{(n)}) - \mathcal{J}(u^{(n)}) \geq C\|u^{(n)} - u^{(n+1)}\|^2_{\ell_2(\mathcal{I})}.$$

Since $\mathcal{J}(u^{(n)}) \geq 0$ is a decreasing sequence and is bounded below, it also converges, and

$$\lim_{n\to\infty} \|u^{(n+1)} - u^{(n)}\|^2_{\ell_2(\mathcal{I})} = 0.$$

□

This lemma already gives strong hints that the algorithm converges. In particular, two successive iterations become closer and closer (3.119), and by the uniform boundedness of $(u^{(n)})_n$, we know already that there are weakly converging subsequences. However, in order to conclude the convergence of the full sequence to a minimizer of $\mathcal{J}$ we need more technical work.

3.1.3 Weak Convergence of the Algorithm

As as simple exercise we state the following

Lemma 3.4 *The operator* $\mathbb{S}_\alpha$ *is nonexpansive, i.e.,*

$$\|\mathbb{S}_\alpha(u) - \mathbb{S}_\alpha(a)\|_{\ell_2(\mathcal{I})} \leq \|u - a\|_{\ell_2(\mathcal{I})}, \tag{3.120}$$

for all $u, a \in \ell_2(\mathcal{I})$.

Proof. Sketch: reason again componentwise and distinguish cases whether u_λ and/or a_λ are smaller or larger than the threshold $\pm\alpha_\lambda$. □

Moreover, we can characterize minimizers of $\mathcal{J}$ in the following way.

Proposition 3.5 *Define*

$$\Gamma(u) = \mathbb{S}_\alpha \left[u + A^* y - A^* A u \right].$$

Then the set of minimizers of $\mathcal{J}$ *coincides with the set* $\mathrm{Fix}(\Gamma)$ *of fixed points of* Γ. *In particular, since* $\mathcal{J}$ *is a coercive functional, it has minimizers, and therefore* Γ *has fixed points.*

Proof. Assume that u is the minimizer of $\mathcal{J}^S(\cdot, a)$. Let us now observe, first of all, that

$$\begin{aligned} \mathcal{J}^S(u+h,a) &= \mathcal{J}^S(u,a) + 2\langle h, u - a - A^*(y - Aa)\rangle \\ &+ \sum_{\lambda\in\mathcal{I}} 2\alpha_\lambda(|u_\lambda + h_\lambda| - |u_\lambda|) + \|h\|^2_{\ell_2(\mathcal{I})}. \end{aligned}$$

We define now $\mathcal{I}_0 = \{u_\lambda = 0\}$ and $\mathcal{I}_1 = \mathcal{I} \setminus \mathcal{I}_0$. Since, by Lemma 3.2 we have $u = \mathbb{S}_\alpha(a + A^*(y - Aa))$, and substituting it for u, we then have

$$\begin{aligned} \mathcal{J}^S(u+h,a) - \mathcal{J}^S(u,a) &= \|h\|^2_{\ell_2(\mathcal{I})} + \sum_{\lambda\in\mathcal{I}_0} \left[2\alpha_\lambda|h_\lambda| - 2h_\lambda(a_\lambda - A^*(y - Aa_\lambda))\right] \\ &+ \sum_{\lambda\in\mathcal{I}_1} \left[2\alpha_\lambda|u_\lambda + h_\lambda| - 2\alpha_\lambda|u_\lambda| + h_\lambda(-2\alpha_\lambda \operatorname{sgn}(u_\lambda))\right]. \end{aligned}$$

If $\lambda \in \mathcal{I}_0$ then $|a_\lambda - A^*(y - Aa_\lambda)| \le \alpha_\lambda$, so that $2\alpha_\lambda|h_\lambda| - 2h_\lambda(a_\lambda - A^*(y - Aa_\lambda) \ge 0$. If $\lambda \in \mathcal{I}_1$, we distinguish two cases: if $u_\lambda > 0$, then

$$2\alpha_\lambda|u_\lambda + h_\lambda| - 2\alpha_\lambda|u_\lambda| + h_\lambda(-2\alpha_\lambda \operatorname{sgn}(u_\lambda)) = 2\alpha_\lambda[|u_\lambda + h_\lambda| - (u_\lambda + h_\lambda)] \ge 0.$$

If $u_\lambda < 0$, then

$$2\alpha_\lambda|u_\lambda + h_\lambda| - 2\alpha_\lambda|u_\lambda| + h_\lambda(-2\alpha_\lambda \operatorname{sgn}(u_\lambda)) = 2\alpha_\lambda[|u_\lambda + h_\lambda| + (u_\lambda + h_\lambda)] \ge 0.$$

It follows

$$\mathcal{J}^S(u+h,a) - \mathcal{J}^S(u,a) \ge \|h\|^2_{\ell_2(\mathcal{I})}. \tag{3.121}$$

Let us assume now that

$$u = S_\alpha \left[u + A^* y - A^* A u \right].$$

Then u is the minimizer of $\mathcal{J}^S(\cdot, u)$, and therefore

$$\mathcal{J}^S(u+h,u) \ge \mathcal{J}^S(u,u) + \|h\|^2_{\ell_2(\mathcal{I})}.$$

Observing now that $\mathcal{J}(u) = J^S(u,u)$ and that $\mathcal{J}^S(u+h,u) = \mathcal{J}(u+h) + \|h\|^2_{\ell_2(\mathcal{I})} - \|Ah\|^2_Y$, we conclude that $\mathcal{J}(u+h) \ge \mathcal{J}(u) + \|Ah\|^2_Y$ for every f. Hence u is a minimizer of $\mathcal{J}$. Vice versa, if u is a minimizer of $\mathcal{J}$, then it is a minimizer of $\mathcal{J}^S(\cdot, u)$, and hence a fixed point of Γ. ∎

We need now to recall an important and well-known result related to iterations of nonexpansive maps [61]. We report it without proof; a simplified version of it can be also found in the Appendix B of [28].

Theorem 3.6 (Opial's Theorem) *Let the mapping* Γ *from* $\ell_2(\mathcal{I})$ *to itself satisfy the following conditions:*

(i) Γ *is nonexpansive, i.e.,* $\|\Gamma(u) - \Gamma(a)\|_{\ell_2(\mathcal{I})} \leq \|u - a\|_{\ell_2(\mathcal{I})}$*, for all* $u, a \in \ell_2(\mathcal{I})$*;*

(ii) Γ *is asymptotically regular, i.e.,* $\|\Gamma^{n+1}(u) - \Gamma^n(u)\|_{\ell_2(\mathcal{I})} \to 0$ *for* $n \to \infty$*;*

(iii) the set $\mathrm{Fix}(\Gamma)$ *of its fixed points is not empty.*

Then, for all u*, the sequence* $(\Gamma^n(u))_n$ *converges weakly to a fixed point in* $\mathrm{Fix}(\Gamma)$.

Eventually we have the weak convergence of the algorithm.

Theorem 3.7 *For any initial choice* $u^{(0)} \in \ell_2(\mathcal{I})$*, Algorithm 4 produces a sequence* $(u^{(n)})_n$ *which converges weakly to a minimizer of* $\mathcal{J}$*.*

Proof. It is sufficient to observe that, due to our previous results, Lemma 3.4, Lemma 3.3, and Proposition 3.5, and the assumption $\|A\| < 1$, the map $\Gamma(u) = \mathbb{S}_\alpha\left[u + A^*y - A^*Au\right]$ fulfills the requirements of Opial's Theorem. □

3.1.4 Strong Convergence of the Algorithm

In this section we shall prove the convergence of the successive iterates $u^{(n)}$ not only in the weak topology, but also in norm. Let us start by introducing some useful notations:

$$u^* = w - \lim_n u^{(n)}, \quad \xi^{(n)} = u^{(n)} - u^*, \quad h = u^* + A^*(y - Au^*).$$

We split again the proof into several intermediate lemmas.

Lemma 3.8 *We have*

$$\|A\xi^{(n)}\|_Y^2 \to 0,$$

for $n \to \infty$.

Proof. Since

$$\xi^{(n+1)} - \xi^{(n)} = \mathbb{S}_\alpha(h + (I - A^*A)\xi^{(n)}) - \mathbb{S}_\alpha(h) - \xi^{(n)},$$

and $\|\xi^{(n+1)} - \xi^{(n)}\|_{\ell_2(\mathcal{I})} = \|u^{(n+1)} - u^{(n)}\|_{\ell_2(\mathcal{I})} \to 0$ for $n \to \infty$, we have

$$\|\mathbb{S}_\alpha(h + (I - A^*A)\xi^{(n)}) - \mathbb{S}_\alpha(h) - \xi^{(n)}\|_{\ell_2(\mathcal{I})} \to 0, \tag{3.122}$$

for $n \to \infty$, and hence, also

$$\max(0, \|\xi^n\|_{\ell_2(\mathcal{I})} - \|\mathbb{S}_\alpha(h + (I - A^*A)\xi^{(n)}) - \mathbb{S}_\alpha(h)\|_{\ell_2(\mathcal{I})}) \to 0, \tag{3.123}$$

for $n \to \infty$. Since $\mathbb{S}_\alpha$ is nonexpansive we have

$$\|\mathbb{S}_\alpha(h + (I - A^*A)\xi^{(n)}) - \mathbb{S}_\alpha(h)\|_{\ell_2(\mathcal{I})} \leq \|(I - A^*A)\xi^{(n)}\|_{\ell_2(\mathcal{I})} \leq \|\xi^{(n)})\|_{\ell_2(\mathcal{I})};$$

therefore the "$\max$" in (3.123) can be dropped, and it follows that

$$\|\xi^n\|_{\ell_2(\mathcal{I})} - \|(I - A^*A)\xi^{(n)}\|_{\ell_2(\mathcal{I})} \to 0, \tag{3.124}$$

for $n \to \infty$. Because

$$\|\xi^n\|_{\ell_2(\mathcal{I})} + \|(I - A^*A)\xi^{(n)}\|_{\ell_2(\mathcal{I})} \leq 2\|\xi^n\|_{\ell_2(\mathcal{I})} = 2\|u^{(n)} - u^*\|_{\ell_2(\mathcal{I})} \leq C,$$

(Remind that $u^{(n)}$ is uniformly bounded by Lemma 3.3.), we obtain

$$\|\xi^n\|^2_{\ell_2(\mathcal{I})} - \|(I - A^*A)\xi^{(n)}\|^2_{\ell_2(\mathcal{I})} \to 0,$$

for $n \to \infty$ by (3.124). The inequality

$$\|\xi^n\|^2_{\ell_2(\mathcal{I})} - \|(I - A^*A)\xi^{(n)}\|^2_{\ell_2(\mathcal{I})} = 2\|A\xi^{(n)}\|^2_Y - \|A^*A\xi^{(n)}\|^2_{\ell_2(\mathcal{I})} \geq \|A\xi^{(n)}\|^2_Y,$$

then implies the statement. □

The previous lemma allows us to derive the following fundamental property.

Lemma 3.9 *For h given as above, $\|\mathbb{S}_\alpha(h + \xi^{(n)}) - \mathbb{S}_\alpha(h) - \xi^{(n)}\|_{\ell_2(\mathcal{I})} \to 0$, for $n \to \infty$.*

Proof. We have

$$\begin{aligned}
& \|\mathbb{S}_\alpha(h + \xi^{(n)}) - \mathbb{S}_\alpha(h) - \xi^{(n)}\|_{\ell_2(\mathcal{I})} \\
\leq\ & \|\mathbb{S}_\alpha(h + (I - A^*A)\xi^{(n)}) - \mathbb{S}_\alpha(h) - \xi^{(n)}\|_{\ell_2(\mathcal{I})} \\
+\ & \|\mathbb{S}_\alpha(h + \xi^{(n)}) - \mathbb{S}_\alpha(h + (I - A^*A)\xi^{(n)}\|_{\ell_2(\mathcal{I})} \\
\leq\ & \|\mathbb{S}_\alpha(h + (I - A^*A)\xi^{(n)}) - \mathbb{S}_\alpha(h) - \xi^{(n)}\|_{\ell_2(\mathcal{I})} + \|A^*A\xi^{(n)}\|_{\ell_2(\mathcal{I})}.
\end{aligned}$$

Both terms tend to 0, the first because of (3.122) and the second because of Lemma 3.8. □

Lemma 3.10 *If for some $a \in \ell_2(\mathcal{I})$ and some sequence $(v^n)_n$, $w - \lim_n v^n = 0$, and $\lim_n \|\mathbb{S}_\alpha(a + v^{(n)}) - \mathbb{S}_\alpha(a) - v^{(n)}\|_{\ell_2(\mathcal{I})} = 0$, then $\|v^n\|_{\ell_2(\mathcal{I})} \to 0$, for $n \to \infty$.*

Proof. Let us define a finite set $\mathcal{I}_0 \subset \mathcal{I}$ such that $\sum_{\lambda \in \mathcal{I} \setminus \mathcal{I}_0} |a_\lambda|^2 \leq \left(\frac{\bar{\alpha}}{4}\right)^2$, where $\bar{\alpha} = \inf_\lambda \alpha_\lambda$. Because this is a finite set $\sum_{\lambda \in \mathcal{I}_0} |v^n_\lambda|^2 \to 0$ for $n \to \infty$, and hence we can concentrate on $\sum_{\lambda \in \mathcal{I} \setminus \mathcal{I}_0} |v^n_\lambda|^2$ only. For each n, we split $\mathcal{I}_1 = \mathcal{I} \setminus \mathcal{I}_0$ into two subsets: $\mathcal{I}_{1,n} = \{\lambda \in \mathcal{I}_1 : ||v^n_\lambda + a_\lambda| < \alpha_\lambda\}$ and $\tilde{\mathcal{I}}_{1,n} = \mathcal{I}_1 \setminus \mathcal{I}_{1,n}$. If $\lambda \in \mathcal{I}_{1,n}$ then

$S_{\alpha_\lambda}(a_\lambda + v_\lambda^{(n)}) = S_{\alpha_\lambda}(a_\lambda) = 0$ (since $|a_\lambda| \leq \frac{\bar{\alpha}}{4} \leq \alpha_\lambda$), so that $|S_{\alpha_\lambda}(a_\lambda + v_\lambda^n) - S_{\alpha_\lambda}(a_\lambda) - v_\lambda^n| = |v_\lambda^n|$. It follows

$$\sum_{\lambda \in \mathcal{I}_{1,n}} |v_\lambda^n|^2 \leq \sum_{\lambda \in \mathcal{I}} |S_{\alpha_\lambda}(a_\lambda + v_\lambda^n) - S_{\alpha_\lambda}(a_\lambda) - v_\lambda^n|^2 \to 0,$$

for $n \to \infty$. It remains to prove that $\sum_{\lambda \in \tilde{\mathcal{I}}_{1,n}} |v_\lambda^n|^2 \to 0$ as $n \to \infty$. If $\lambda \in \mathcal{I}_1$ and $|a_\lambda + v_\lambda^n| \geq \alpha_\lambda$, then $|v_\lambda^n| \geq |a_\lambda + v_\lambda^n| - |a_\lambda| \geq \alpha_\lambda - \frac{\bar{\alpha}}{4} > \frac{\bar{\alpha}}{4} \geq |a_\lambda|$, so that $a_\lambda + v_\lambda^n$ and v_λ^n have the same sign. In particular, $\alpha_\lambda - \frac{\bar{\alpha}}{4} > |a_\lambda|$ implies $\alpha_\lambda - |a_\lambda| \geq \frac{\bar{\alpha}}{4}$. It follows that

$$\begin{aligned} |v_\lambda^n - S_{\alpha_\lambda}(a_\lambda + v_\lambda^n) + S_{\alpha_\lambda}(a_\lambda)| &= |v_\lambda^n - S_{\alpha_\lambda}(a_\lambda + v_\lambda^n)| \\ &= |v_\lambda^n - (a_\lambda + v_\lambda^n) + \alpha_\lambda \operatorname{sgn}(v_\lambda^n)| \\ &\geq \alpha_\lambda - |a_\lambda| \geq \frac{\bar{\alpha}}{4}. \end{aligned}$$

This implies that

$$\sum_{\lambda \in \tilde{\mathcal{I}}_{1,n}} |S_{\alpha_\lambda}(a_\lambda + v_\lambda^n) - S_{\alpha_\lambda}(a_\lambda) - v_\lambda^n|^2 \geq \left(\frac{\bar{\alpha}}{4}\right)^2 |\tilde{\mathcal{I}}_{1,n}|.$$

But, $\sum_{\lambda \in \tilde{\mathcal{I}}_{1,n}} |S_{\alpha_\lambda}(a_\lambda + v_\lambda^n) - S_{\alpha_\lambda}(a_\lambda) - v_\lambda^n|^2 \geq \left(\frac{\bar{\alpha}}{4}\right)^2 \to 0$ for $n \to \infty$ and therefore $\tilde{\mathcal{I}}_{1,n}$ must be empty for n large enough. □

The combination of Lemma 3.8 and Lemma 3.9, together with the weak convergence Theorem 3.7 allows us to have norm convergence.

Theorem 3.11 *For any initial choice $u^{(0)} \in \ell_2(\mathcal{I})$, Algorithm 4 produces a sequence $(u^{(n)})_n$ which converges strongly to a minimizer u^* of $\mathcal{J}$.*

3.2 Principles of Acceleration

Recently, also the qualitative convergence properties of iterative soft-thresholding have been investigated. Note first that the aforementioned condition or $\|A\| < 1$ (or even $\|A\| < \sqrt{2}$) does not guarantee contractivity of the iteration operator $I - A^*A$, since A^*A may not be boundedly invertible. The insertion of $\mathbb{S}_\alpha$ does not improve the situation since $\mathbb{S}_\alpha$ is nonexpansive, but also noncontractive. Hence, for any minimizer u^* (which is also a fixed point of (3.112)), the estimate

$$\|u^* - u^{(n+1)}\|_{\ell_2(\mathcal{I})} \leq \left\|(I - A^*A)(u^* - u^{(n)})\right\|_{\ell_2(\mathcal{I})} \leq \|I - A^*A\| \|u^* - u^{(n)}\|_{\ell_2(\mathcal{I})} \quad (3.125)$$

does not give rise to a linear error reduction. However, under additional assumptions on the operator A or on minimizers u^*, linear convergence of (3.112) can be easily ensured. In particular, if A fulfills the so-called *finite basis injectivity* (FBI) condition (see [10] where this terminology is introduced), i.e., for any finite set $\Lambda \subset \mathcal{I}$, the restriction A_Λ is injective,then (3.112) converges linearly to a minimizer u^* of $\mathcal{J}$. The following simple argument shows indeed that the FBI condition implies linear error reduction as soon as $\|A\| < 1$. In that case, we have strong convergence (Theorem 3.11) of the $u^{(n)}$ to a finitely supported limit sequence u^*. We can therefore find a finite index set $\Lambda \subset \mathcal{I}$ such that all iterates $u^{(n)}$ and u^* are supported in Λ. By the FBI condition, A_Λ is injective and hence $A^*A|_{\Lambda\times\Lambda}$ is boundedly invertible, so that $I - A^*_\Lambda A_\Lambda$ is a contraction on $\ell_2(\Lambda)$. Using

$$u_\Lambda^{(n+1)} = \mathbb{S}_\alpha\big(u_\Lambda^{(n)} + A^*_\Lambda(v - A_\Lambda u_\Lambda^{(n)})\big)$$

and an analogous argument as in (3.125), it follows that $\|u^* - u^{(n+1)}\|_{\ell_2(\mathcal{I})} \leq \gamma \|u^* - u^{(n)}\|_{\ell_2(\mathcal{I})}$, where $\gamma = \max\{|1 - \|(A^*A|_{\Lambda\times\Lambda})^{-1}\|^{-1}|, |\|A^*A|_{\Lambda\times\Lambda}\| - 1|\} \in (0,1)$. Typical examples where $A = K \circ F$ fulfills the FBI condition arise when K is injective and Ψ is either a Riesz basis for X or a so-called FBI frame, i.e., each finite subsystem of Ψ is linearly independent. However, depending on Λ, the matrix $A^*A|_{\Lambda\times\Lambda}$ can be arbitrarily badly conditioned, resulting in a constant error reduction γ, arbitrarily close to 1.

However, it is possible to show that for several FBI operators K and for certain choices of Ψ, the matrix A^*A can be preconditioned by a matrix $D^{-1/2}$, resulting in the matrix $D^{-1/2}A^*AD^{-1/2}$, in such a way that any restriction $(D^{-1/2}A^*AD^{-1/2})_{\Lambda\times\Lambda}$ turns out to be well-conditioned as soon as $\Lambda \subset \mathcal{I}$ is a small set, but independently of its "location" within $\mathcal{I}$. Let us remark that, in particular, we do not claim to be able to have full well-conditioned matrices (as it happens in well-posed problems [23, 24] by simple diagonal preconditioning), but that only small arbitrary finite dimensional submatrices are indeed well-conditioned. Let us say that one can promote a "local" well-conditioning of the matrices.

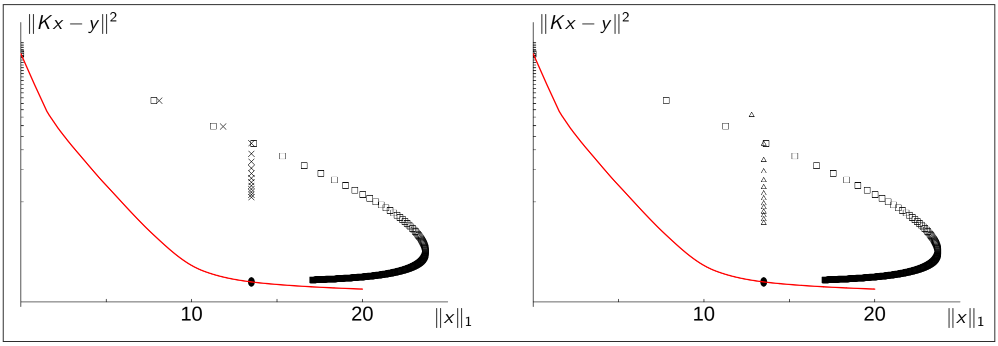
Typically one considers injective (non local) compact operators K with Schwartz kernel having certain polynomial decay properties of the derivatives, i.e.,

$$Ku(x) = \int_\Omega \Phi(x,\xi)u(\xi)d\xi, \quad x \in \tilde{\Omega},$$

for $\tilde{\Omega}, \Omega \subset \mathbb{R}^d$, $u \in X := H^t(\Omega)$, and

$$|\partial_x^\alpha \partial_\xi^\beta \Phi(x,\xi)| \leq c_{\alpha,\beta}|x-\xi|^{-(d+2t+|\alpha|+|\beta|)}, \quad t \in \mathbb{R}, \text{ and multi-indexes } \alpha, \beta \in \mathbb{N}^d.$$

Moreover, for the proper choice of the discrete matrix $A^*A := F^*K^*KF$, one uses multiscale bases Ψ, such as wavelets, which do make a good job in this situation. We refer the reader to [22] for more details.



(a) (b)

Figure 3.1 The path, in the $\|x\|_1$ vs. $\|Kx - y\|^2$ plane, followed by the iterates $u^{(n)}$ of three different iterative algorithms. The operator K and the data y are taken from a seismic tomography problem [53]. The boxes (in both (a) and (b)) correspond to the thresholded Landweber algorithm. In this example, iterative thresholded Landweber (3.112) first overshoots the ℓ_1 norm of the limit (represented by the fat dot), and then requires a large number of iterations to reduce $\|u^{(n)}\|_1$ again (500 are shown in this figure). In (a) the crosses correspond to the path followed by the iterates of the projected Landweber iteration (which is given as in (3.126) for $\beta^{(n)} = 1$); in (b) the triangles correspond to the projected steepest descent iteration (3.126); in both cases, only 15 iterates are shown. The discrepancy decreases more quickly for projected steepest descent than for the projected Landweber algorithm. The solid line corresponds to the limit *trade-off curve*, generated by $u^*(\bar{\alpha})$ for decreasing values of $\bar{\alpha} > 0$. The vertical axes uses a logarithmic scale for clarity.

3.2.1 From the Projected Gradient Method to the Iterative Soft-Thresholding with Decreasing Thresholding Parameter

With such a “local” well-conditioning, it should also be clear that iterating on small sets Λ will also improve the convergence rate. Unfortunately, the iterative soft-thresholding does not act initially on small sets (see also Figure 3.4 and Figure 3.5), but it rather starts iterating on relatively large sets, slowly shrinking to the size of the support of the limit u^*.

Let us take a closer look at the characteristic dynamics of Algorithm 4 in Figure 3.1. Let us assume for simplicity here that $\alpha_\lambda = \bar{\alpha} > 0$ for all $\lambda \in \mathcal{I}$. As this plot of the discrepancy $\mathcal{D}(u^{(n)}) = \|Ku^{(n)} - y\|_Y^2 = \|Au^{(n)} - y\|_Y^2$ versus $\|u^{(n)}\|_1$ shows, the algorithm converges initially relatively fast, then it overshoots the value $\|u^*\|_1$ and it takes very long to re-correct back. In other words, starting from $u^{(0)} = 0$, the algorithm generates a path $\{u^{(n)}; n \in \mathbb{N}\}$ that is initially fully contained in the ℓ_1-ball $B_R := B_{\ell_1(\mathcal{I})}(R) := \{u \in \ell_2(\Lambda); \|u\|_1 \leq R\}$, with $R := \|u^*\|_1$. Then it gets out of the ball to slowly inch back to it in the limit.

The way to avoid this long “external” detour was proposed in [29] by forcing the successive iterates to remain within the ball B_R. One method to achieve this is to substitute for the thresholding operations the projection $\mathbb{P}_{B_R}$, where, for any closed convex set C, and any u, we define $\mathbb{P}_C(u)$ to be the unique point in C for which the ℓ_2-distance to u is minimal. With a slight abuse of notation, we shall denote $\mathbb{P}_{B_R}$ by $\mathbb{P}_R$; this will not cause confusion, because it will be clear from the context whether the subscript of $\mathbb{P}$ is a set or a positive number.

Furthermore, modifying the iterations by introducing an adaptive “descent parameter” $\beta^{(n)} > 0$ in each iteration, defining $u^{(n+1)}$ by

$$u^{(n+1)} = \mathbb{P}_R \left[u^{(n)} + \beta^{(n)} A^*(y - Au^{(n)}) \right], \tag{3.126}$$

does lead, in numerical simulations, to much faster convergence. The typical dynamics of this modified algorithm are illustrated in Figure 3.1(b), which clearly shows the larger steps and faster convergence (when compared with the projected Landweber iteration in Fig. 3.1(a) which is for $\beta^{(n)} = 1$). We shall refer to this modified algorithm as the *projected gradient iteration* or the *projected steepest descent* (PSD). The motivation of the faster convergence behavior is the fact that we never leave the target ℓ_1-ball, and we tend not to iterate on large index sets. On the basis of this intuition we find even more promising results for an ‘interior’ algorithm in which we still project on ℓ_1-balls, but now with a slowly increasing radius, i.e.

$$u^{(n+1)} = \mathbb{P}_{R^{(n)}} \left(u^{(n)} + \beta^{(n)} A^*(y - Au^{(n)}) \right) \qquad \text{and} \qquad R^{(n+1)} = (n+1)R/N \tag{3.127}$$

where N is the prescribed maximum number of iterations (the origin is chosen as the starting point of this iteration). The better performance of this algorithm can be explained by the fact that the projection $\mathbb{P}_R(u)$ onto an ℓ_1-ball of radius R do coincide

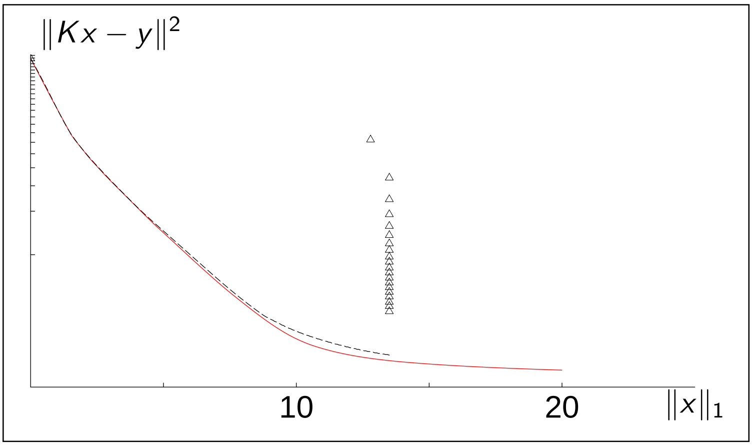


Figure 3.2 Trade-off curve and its approximation with algorithm (3.127) in 200 steps.

with a thresholding $\mathbb{P}_R(u) = \mathbb{S}_{\alpha(u;R)}(u)$ for a suitable thresholding parameter $\alpha = \alpha(u;R)$ depending on u and R, which is larger for smaller R.

Lemma 3.12 *For any fixed $a \in \ell_2(\mathcal{I})$ and for $\tau > 0$, $\|\mathbb{S}_\tau(a)\|_1$ is a piecewise linear, continuous, decreasing function of τ; moreover, if $a \in \ell_1(\Lambda)$ then $\|\mathbb{S}_0(a)\|_{\ell_1(\mathcal{I})} = \|a\|_{\ell_1(\mathcal{I})}$ and $\|\mathbb{S}_\tau(a)\|_{\ell_1(\mathcal{I})} = 0$ for $\tau \geq \max_\lambda |a_\lambda|$.*

Proof. $\|\mathbb{S}_\tau(a)\|_{\ell_1(\mathcal{I})} = \sum_\lambda |S_\tau(a_\lambda)| = \sum_\lambda S_\tau(|a_\lambda|) = \sum_{|a_\lambda|>\tau}(|a_\lambda| - \tau)$; the sum in the right hand side is finite for $\tau > 0$. □

A schematic illustration is given in Figure 3.3.

Lemma 3.13 *If $\|a\|_{\ell_1(\mathcal{I})} > R$, then the $\ell_2(\mathcal{I})$ projection of a on the ℓ_1-ball with radius R is given by $\mathbb{P}_R(a) = \mathbb{S}_\mu(a)$ where μ (depending on a and R) is chosen such that $\|\mathbb{S}_\mu(a)\|_{\ell_1(\mathcal{I})} = R$. If $\|a\|_{\ell_1(\mathcal{I})} \leq R$ then $\mathbb{P}_R(a) = \mathbb{S}_0(a) = a$.*

Proof. Suppose $\|a\|_{\ell_1(\mathcal{I})} > R$. Because, by Lemma 3.12, $\|\mathbb{S}_\mu(a)\|_{\ell_1(\mathcal{I})}$ is continuous in μ and $\|\mathbb{S}_\mu(a)\|_{\ell_1(\mathcal{I})} = 0$ for sufficiently large μ, we can choose μ such that $\|\mathbb{S}_\mu(a)\|_{\ell_1(\mathcal{I})} = R$. (See Figure 3.3.) On the other hand, $u^* = \mathbb{S}_\mu(a)$ is the unique minimizer of $\|u - a\|^2_{\ell_2(\mathcal{I})} + 2\mu\|u\|_{\ell_1(\mathcal{I})}$ (see Lemma 3.1), i.e.,

$$\|u^* - a\|^2_{\ell_2(\mathcal{I})} + 2\mu\|u^*\|_{\ell_1(\mathcal{I})} < \|u - a\|^2_{\ell_2(\mathcal{I})} + 2\mu\|u\|_{\ell_1(\mathcal{I})}$$

for all $u \neq i^*$. Since $\|u^*\|_{\ell_1(\mathcal{I})} = R$, it follows that

$$\forall u \in B_R,\ u \neq u^* \ : \qquad \|u^* - a\|^2 < \|u - a\|^2$$

Hence u^* is closer to a than any other u in B_R. In other words, $\mathbb{P}_R(a) = u^* = \mathbb{S}_\mu(a)$. □

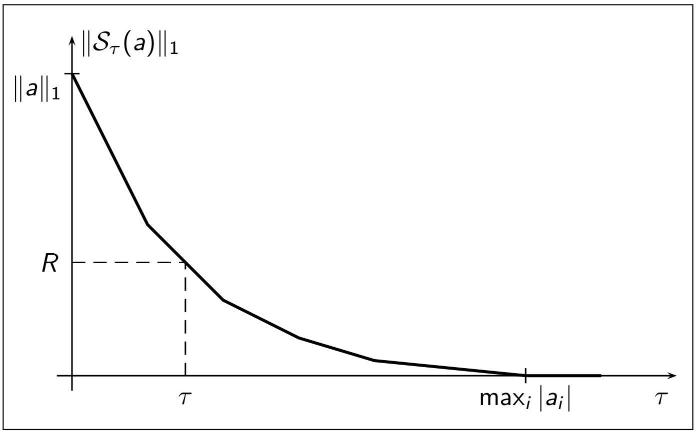


Figure 3.3 For a given vector $a \in \ell_2$, $\|\mathbb{S}_\tau(a)\|_1$ is a piecewise linear continuous and decreasing function of τ (strictly decreasing for $\tau < \max_i |a_i|$) . The knots are located at $\{|a_i|, i : 1 \dots m\}$ and 0. Finding τ such that $\|\mathbb{S}_\tau(a)\|_{\ell_1(\mathcal{I})} = R$ ultimately comes down to a linear interpolation. The figure is made for the finite dimensional case.

This in particular implies that the algorithm (3.127) iterates initially on very small sets which inflate by growing during the process and approach the size of the support of the target minimizer u^*. Unlike the thresholded Landweber iteration and the projected steepest descent [28, 29], unfortunately there is no proof yet of convergence of this 'interior' algorithm, being a very interesting open problem.

However, we can provide an algorithm which mimics the behavior of (3.127), i.e., it starts with large thresholding parameters $\alpha^{(n)}$ and geometrically reduces them during the iterations to a target limit $\alpha > 0$, for which the convergence is guaranteed:

$$u^{(n+1)} = \mathbb{S}_{\alpha^{(n)}} \left[u^{(n)} + A^* y - A^* A u^{(n)} \right]. \tag{3.128}$$

For matrices A for which the restrictions $A^*A|_{\Lambda \times \Lambda}$ are uniformly well-conditioned with respect to Λ of small size, our analysis provides also a prescribed linear rate of convergence of the iteration (3.128).

3.2.2 Sample of Analysis of Acceleration Methods

Technical lemmas

We are particularly interested in computing approximations with the smallest possible number of nonzero entries. As a benchmark, we recall that the most economical approximations of a given vector $v \in \ell_2(\mathcal{I})$ are provided again by the *best k-term approximations* $v_{[k]}$, defined by discarding in v all but the $k \in \mathbb{N}_0$ largest coefficients in absolute value. The error of best N-term approximation is defined as

$$\sigma_k(v)_{\ell_2} := \|v - v_{[k]}\|_{\ell_2(\mathcal{I})}. \tag{3.129}$$

The subspace of all ℓ_2 vectors with best k-term approximation rate $s > 0$, i.e., $\sigma_k(v)_{\ell_2} \lesssim k^{-s}$ for some decay rate $s > 0$, is commonly referred to as the *weak* ℓ_τ *space* $\ell_\tau^w(\mathcal{I})$, for $\tau = (s + \frac{1}{2})^{-1}$, which, endowed with

$$|v|_{\ell_\tau^w(\mathcal{I})} := \sup_{k \in \mathbb{N}_0} (k+1)^s \sigma_k(v)_{\ell_2}, \tag{3.130}$$

becomes the quasi-Banach space $(\ell_\tau^w(\mathcal{I}), |\cdot|_{\ell_\tau^w(\mathcal{I})})$. Moreover, for any $0 < \epsilon \leq 2 - \tau$, we have the continuous embedding $\ell_\tau(\mathcal{I}) \hookrightarrow \ell_\tau^w(\mathcal{I}) \hookrightarrow \ell_{\tau+\epsilon}(\mathcal{I})$, justifying why $\ell_\tau^w(\mathcal{I})$ is called weak $\ell_\tau(\mathcal{I})$.

When it comes to the concrete computations of good approximations with a small number of active coefficients, one frequently utilizes certain thresholding procedures. Here small entries of a given vector are simply discarded, whereas the large entries may be slightly modified. In this paper, we shall make use of *soft-thresholding* that we already introduced in (3.113). It is well-known, see [28], that $\mathbb{S}_\alpha$ is non-expansive for any $\alpha \in \mathbb{R}_+^{\mathcal{I}}$,

$$\|\mathbb{S}_\alpha(v) - \mathbb{S}_\alpha(w)\|_{\ell_2(\mathcal{I})} \leq \|v - w\|_{\ell_2(\mathcal{I})}, \quad \text{for all } v, w \in \ell_2(\mathcal{I}). \tag{3.131}$$

Moreover, for any fixed $x \in \mathbb{R}$, the mapping $\tau \mapsto S_\tau(x)$ is Lipschitz continuous with

$$|S_\tau(x) - S_{\tau'}(x)| \leq |\tau - \tau'|, \quad \text{for all } \tau, \tau' \geq 0. \tag{3.132}$$

We readily infer the following technical estimate.

Lemma 3.14 *Assume* $v \in \ell_2(\mathcal{I})$, $\alpha, \beta \in \mathbb{R}_+^{\mathcal{I}}$ *such that* $\bar{\alpha} = \inf_\lambda \alpha_\lambda = \inf_\lambda \beta_\lambda = \bar{\beta} > 0$, *and define* $\Lambda_{\bar{\alpha}}(v) := \left\{\lambda \in \mathcal{I} : |v_\lambda| > \bar{\alpha}\right\}$. *Then*

$$\|\mathbb{S}_\alpha(v) - \mathbb{S}_\beta(v)\|_{\ell_2(\mathcal{I})} \leq \left(\#\Lambda_{\bar{\alpha}}(v)\right)^{1/2} \max_{\lambda \in \Lambda_{\bar{\alpha}}(v)} |\alpha_\lambda - \beta_\lambda|. \tag{3.133}$$

Proof. By (3.132) we have the estimate

$$\begin{aligned}
\|\mathbb{S}_\alpha(v) - \mathbb{S}_\beta(v)\|_{\ell_2(\mathcal{I})} &= \left(\sum_{\lambda \in \mathcal{I}} |S_{\alpha_\lambda}(v_\lambda) - S_{\beta_\lambda}(v_\lambda)|^2\right)^{1/2} \\
&= \left(\sum_{\lambda \in \{\mu \in \mathcal{I} : |v_\mu| > \min\{\alpha_\mu, \beta_\mu\}\}} |S_{\alpha_\lambda}(v_\lambda) - S_{\beta_\lambda}(v_\lambda)|^2\right)^{1/2} \\
&= \left(\sum_{\lambda \in \Lambda_{\bar{\alpha}}(v)} |S_{\alpha_\lambda}(v_\lambda) - S_{\beta_\lambda}(v_\lambda)|^2\right)^{1/2} \\
&\leq \left(\#\Lambda_{\bar{\alpha}}(v)\right)^{1/2} \max_{\lambda \in \Lambda_{\bar{\alpha}}(v)} |\alpha_\lambda - \beta_\lambda|
\end{aligned}$$

□

Let $v \in \ell_\tau^w(\mathcal{I})$, it is well-known [31, §7]

$$\#\Lambda_{\bar\alpha}(v) \le C|v|^\tau_{\ell_\tau^w(\mathcal{I})}\bar\alpha^{-\tau}, \tag{3.134}$$

and, for $\alpha_\lambda = \bar\alpha$ for all $\lambda \in \mathcal{I}$, we have

$$\|v - \mathbb{S}_\alpha(v)\|_{\ell_2(\mathcal{I})} \le C|v|^{\tau/2}_{\ell_\tau^w(\mathcal{I})}\bar\alpha^{1-\tau/2}, \tag{3.135}$$

where the constants are given by $C = C(\tau) > 0$. Let $v \in \ell_0(\mathcal{I}) := \cap_{\tau>0}\ell_\tau^w(\mathcal{I})$, and $|v|_{\ell_0} := \#\operatorname{supp}(v) < \infty$. Then we have the straightforward estimate

$$\#\Lambda_{\bar\alpha}(v) \le |v|_{\ell_0} \tag{3.136}$$

and, for $\alpha_\lambda = \bar\alpha$ for all $\lambda \in \mathcal{I}$, we have

$$\|v - \mathbb{S}_\alpha(v)\|_{\ell_2(\mathcal{I})} \le |v|^{1/2}_{\ell_0}\bar\alpha, \tag{3.137}$$

which is easily shown by a direct computation. In the sequel, we shall also use the following support size estimate.

Lemma 3.15 *Let $v \in \ell_\tau^w(\mathcal{I})$ and $w \in \ell_2(\mathcal{I})$ with $\|v - w\|_{\ell_2(\mathcal{I})} \le \epsilon$. Assume $\alpha = (\alpha_\lambda)_{\lambda\in\mathcal{I}} \in \mathbb{R}_+^{\mathcal{I}}$ and $\inf_\lambda \alpha_\lambda = \bar\alpha > 0$. Then it holds*

$$\#\operatorname{supp}\mathbb{S}_\alpha(w) \le \#\Lambda_{\bar\alpha}(w) \le \frac{4\epsilon^2}{\bar\alpha^2} + 4C|v|^\tau_{\ell_\tau^w(\mathcal{I})}\bar\alpha^{-\tau}, \tag{3.138}$$

where $C = C(\tau) > 0$. In particular if $v \in \ell_0(\mathcal{I})$ then the estimate is refined

$$\#\operatorname{supp}\mathbb{S}_\alpha(w) \le \#\Lambda_{\bar\alpha}(w) \le \frac{4\epsilon^2}{\bar\alpha^2} + |v|_{\ell_0(\mathcal{I})}. \tag{3.139}$$

Proof. We consider two sets $\mathcal{I}_1 = \{\lambda \in \mathcal{I} : |w_\lambda| \ge \bar\alpha, \text{ and } |v_\lambda| > \bar\alpha/2\}$, and $\mathcal{I}_2 = \{\lambda \in \mathcal{I} : |w_\lambda| \ge \bar\alpha, \text{ and } |v_\lambda| \le \bar\alpha/2\}$. Then from (3.134)

$$\#\mathcal{I}_1 \le \#\{\lambda \in \mathcal{I} : |v_\lambda| > \bar\alpha/2\} \le 2^\tau C|v|^\tau_{\ell_\tau^w(\mathcal{I})}\bar\alpha^{-\tau} \le 4C|v|^\tau_{\ell_\tau^w(\mathcal{I})}\bar\alpha^{-\tau},$$

and

$$(\bar\alpha/2)^2(\#\mathcal{I}_2) \le \sum_{\lambda\in\mathcal{I}_2}|v_\lambda - w_\lambda|^2 \le \varepsilon^2.$$

These estimates imply (3.138), and similarly one gets (3.139). □

Decreasing iterative soft-thresholding

For threshold parameters $\alpha, \alpha^{(n)} \in \mathbb{R}_+^{\mathcal{I}}$, where $\alpha^{(n)} \geq \alpha$, i.e., $\alpha_\lambda^{(n)} \geq \alpha_\lambda$ for all $\lambda \in \Lambda$, and $\bar{\alpha} = \inf_{\lambda \in \mathcal{I}} \alpha_\lambda > 0$, we consider the iteration

Algorithm 5.

$$u^{(0)} = 0, \quad u^{(n+1)} = \mathbb{S}_{\alpha^{(n)}}\big(u^{(n)} + A^*(y - Au^{(n)})\big), \quad n = 0, 1, \dots \tag{3.140}$$

which we call the *decreasing iterative soft-thresholding algorithm* (D-ISTA).

Theorem 3.16 *Let $\|A\| < \sqrt{2}$ and let $\bar{u} := (I - A^*A)u^* + A^*y \in \ell_\tau^w(\mathcal{I})$ for some $0 < \tau < 2$. Moreover, let $L = L(\alpha) := \frac{4\|u^*\|^2_{\ell_2(\mathcal{I})}}{\bar{\alpha}^2} + 4C\|\bar{u}\|^\tau_{\ell_\tau^w(\mathcal{I})}\bar{\alpha}^{-\tau}$, and assume that for $S^* := \operatorname{supp} u^*$ and all finite subsets $\Lambda \subset \mathcal{I}$ with at most $\#\Lambda \leq 2L$ elements, the operator $(I - A^*A)|_{(S^* \cup \Lambda) \times (S^* \cup \Lambda)}$ is contractive on $\ell_2(S^* \cup \Lambda)$, i.e., $\|(I - A^*A)|_{S^* \cup \Lambda \times S^* \cup \Lambda} w\|_{\ell_2(S^* \cup \Lambda)} \leq \gamma_0 \|w\|_{\ell_2(S^* \cup \Lambda)}$, for all $w \in \ell_2(S^* \cup \Lambda)$, or*

$$\|(I - A^*A)|_{S^* \cup \Lambda \times S^* \cup \Lambda}\| \leq \gamma_0, \tag{3.141}$$

where $0 < \gamma_0 < 1$. Then, for any $\gamma_0 < \gamma < 1$, the iterates $u^{(n)}$ from (3.140) fulfill $\# \operatorname{supp} u^{(n)} \leq L$ and they converge to u^* at a linear rate

$$\|u^* - u^{(n)}\|_{\ell_2(\mathcal{I})} \leq \gamma^n \|u^*\|_{\ell_2(\mathcal{I})} =: \epsilon_n \tag{3.142}$$

whenever the $\alpha^{(n)}$ are chosen according to

$$\alpha_\lambda \leq \alpha_\lambda^{(n)} \leq \alpha_\lambda + (\gamma - \gamma_0)L^{-1/2}\epsilon_n, \text{ for all } \lambda \in \Lambda. \tag{3.143}$$

Proof. We develop the proof by induction. For the initial iterate, we have $u^{(0)} = 0$, so that $\# \operatorname{supp} u^{(0)} \leq L$ and (3.142) is trivially true. Assume as an induction hypothesis that $S^{(n)} := \operatorname{supp}(u^{(n)})$ is such that $\#S^{(n)} \leq L$, and $\|u^* - u^{(n)}\|_{\ell_2(\mathcal{I})} \leq \epsilon_n$. Abbreviating $w^{(n)} := u^{(n)} + A^*(y - Au^{(n)})$, by $\|A^*A\| \leq 2$ and the induction hypothesis, it follows that

$$\|\bar{u} - w^{(n)}\|_{\ell_2(\mathcal{I})} = \big\|(I - A^*A)(u^* - u^{(n)})\big\|_{\ell_2(\mathcal{I})} \leq \|u^* - u^{(n)}\|_{\ell_2(\mathcal{I})} \leq \epsilon_n. \tag{3.144}$$

Hence, using (3.138), we obtain the estimate

$$\#S^{(n+1)} = \# \operatorname{supp} \mathbb{S}_{\alpha^{(n)}}(w^{(n)}) \leq \Lambda_{\bar{\alpha}}(w^{(n)}) \leq \frac{4\epsilon_n^2}{\bar{\alpha}^2} + 4C|\bar{u}|^\tau_{\ell_\tau^w(\mathcal{I})}\bar{\alpha}^{-\tau} \leq L. \tag{3.145}$$

Since also $\#S^{(n)} \leq L$ by induction hypothesis, the set $\Lambda^{(n)} := S^{(n)} \cup S^{(n+1)}$ has at most $2L$ elements, so that, by assumption, $(I - A^*A)|_{S\cup\Lambda^{(n)}\times S\cup\Lambda^{(n)}}$ is contractive with contraction constant γ_0. Using the identities

$$\begin{aligned} u^*_{S\cup\Lambda^{(n)}} &= \mathbb{S}_\alpha(\bar{u}_{S\cup\Lambda^{(n)}}) \\ &= \mathbb{S}_\alpha(u^*_{S\cup\Lambda^{(n)}} + A^*_{S\cup\Lambda^{(n)}}(y - A_{S\cup\Lambda^{(n)}}u^*_{S\cup\Lambda^{(n)}})), \end{aligned}$$

and

$$\begin{aligned} u^{(n+1)}_{S\cup\Lambda^{(n)}} &= \mathbb{S}_{\alpha^{(n)}}(w^{(n)}_{S\cup\Lambda^{(n)}}) \\ &= \mathbb{S}_{\alpha^{(n)}}(u^{(n)}_{S\cup\Lambda^{(n)}} + A^*_{S\cup\Lambda^{(n)}}(y - A_{S\cup\Lambda^{(n)}}u^{(n)}_{S\cup\Lambda^{(n)}})), \end{aligned}$$

it follows from (3.131), (3.133), (3.125), and $\alpha^{(n)} \geq \alpha$ that

$$\begin{aligned} &\|u^* - u^{(n+1)}\|_{\ell_2(\mathcal{I})} \\ &= \|(u^* - u^{(n+1)})_{S\cup\Lambda^{(n)}}\|_{\ell_2(S\cup\Lambda^{(n)})} \\ &= \left\|\mathbb{S}_\alpha(\bar{u}_{S\cup\Lambda^{(n)}}) - \mathbb{S}_{\alpha^{(n)}}(w^{(n)}_{S\cup\Lambda^{(n)}})\right\|_{\ell_2(S\cup\Lambda^{(n)})} \\ &\leq \left\|\mathbb{S}_\alpha(\bar{u}_{S\cup\Lambda^{(n)}}) - \mathbb{S}_\alpha(w^{(n)}_{S\cup\Lambda^{(n)}})\right\|_{\ell_2(S\cup\Lambda^{(n)})} + \left\|\mathbb{S}_\alpha(w^{(n)}_{S\cup\Lambda^{(n)}}) - \mathbb{S}_{\alpha^{(n)}}(w^{(n)}_{S\cup\Lambda^{(n)}})\right\|_{\ell_2(S\cup\Lambda^{(n)})} \\ &\leq \left\|(I - A^*A|_{S\cup\Lambda^{(n)}\times S\cup\Lambda^{(n)}})(u^* - u^{(n)})_{S\cup\Lambda^{(n)}}\right\|_{\ell_2(S\cup\Lambda^{(n)})} \\ &\qquad\qquad + \left(\#\Lambda_{\bar{\alpha}}(w^{(n)})\right)^{1/2}\left(\max_{\lambda\in\Lambda_{\bar{\alpha}}(w^{(n)})}|\alpha_\lambda - \alpha^{(n)}_\lambda|\right) \\ &\leq \gamma_0\epsilon_n + \left(\#\Lambda_{\bar{\alpha}}(w^{(n)})\right)^{1/2}\left(\max_{\lambda\in\Lambda_{\bar{\alpha}}(w^{(n)})}|\alpha_\lambda - \alpha^{(n)}_\lambda|\right). \end{aligned}$$

Using (3.145) we obtain $\|u - u^{(n+1)}\|_{\ell_2(\mathcal{I})} \leq \gamma_0\epsilon_n + \sqrt{L}\left(\max_{\lambda\in\Lambda_{\bar{\alpha}}(w^{(n)})}|\alpha^{(n)}_\lambda - \alpha_\lambda|\right)$, and, since the $\alpha^{(n)}$ are chosen according to (3.143), the claim follows. □

Note that assumption (3.141) in finite dimension essentially coincides with the request that the matrix A satisfies the RIP (see Lemma 2.14). With these results at hand and those related to RIP matrices in finite dimension, we are in the situation of estimating the relevant parameters in order to apply Theorem 3.16 when we are dealing with a compressed sensing problem. We proceed to a numerical comparison of the algorithm D-ISTA in (3.140) and the iterative soft-thresholding ISTA. In Figure 3.4 we show the behavior of the algorithms in the computation of a sparse minimizer u^* for A being a 500×2500 matrix with i.i.d. Gaussian entries, $\alpha = 10^{-3}$, $\gamma_0 = 0.1$ and $\gamma = 0.95$. In Figure 3.5 we show the behavior of the algorithms in the same situation but for parameters $\alpha = 10^{-4}$, $\gamma_0 = 0.01$ and $\gamma = 0.998$. In both the cases, related to small values of α (we reiterate that a small range of α is the most crucial situation for the efficiency of iterative algorithms, see Section 3.2.2), ISTA tends to iterate initially on vectors with a large number of nonzero entries, while D-ISTA inflates slowly the support size of the iterations to eventually converge to the right support of u^*. The iteration on an inflating support allows D-ISTA to take advantage of the local well-conditioning of the matrix A from the very beginning of the iterations. This effect results in a *controlled* linear rate of convergence which is much steeper than the one of ISTA. In particular in Figure 3.5 after 1500 D-ISTA has correctly detected the support of the minimizer u^* and reached already an accuracy of $10^{-0.5}$, whereas it is clear that the convergence of ISTA is simply dramatically slow.

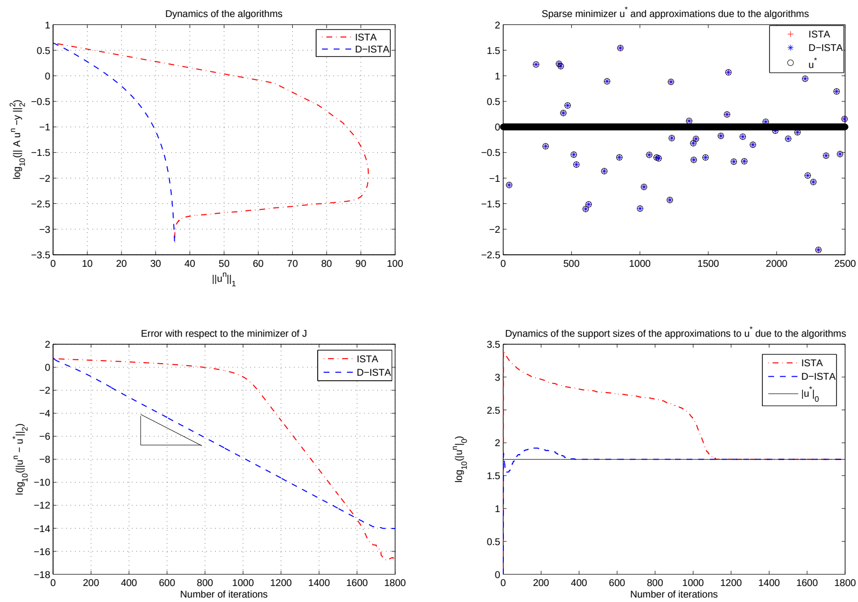


Figure 3.4 We show the behavior of the algorithms ISTA and D-ISTA in the computation of a sparse minimizer u^* for A being a 500×2500 matrix with i.i.d. Gaussian entries, $\alpha = 10^{-3}$, $\gamma_0 = 0.1$ and $\gamma = 0.95$. In the top left figure we presents the dynamics of the algorithms in the plane $\|u\|_{\ell_1} - \log(\|Au - y\|_2^2)$. On the bottom left, we show the absolute error to the precomputed minimizer u^* with respect to the number of iterations. On the bottom right we show how the size of the supports of the iterations grow with the number of iterations. The figure on the top right shows the vector u^*, and the approximations due to the algorithms. In this case both the algorithms approximate with very high accuracy the minimizer u^* after 1800 iterations.

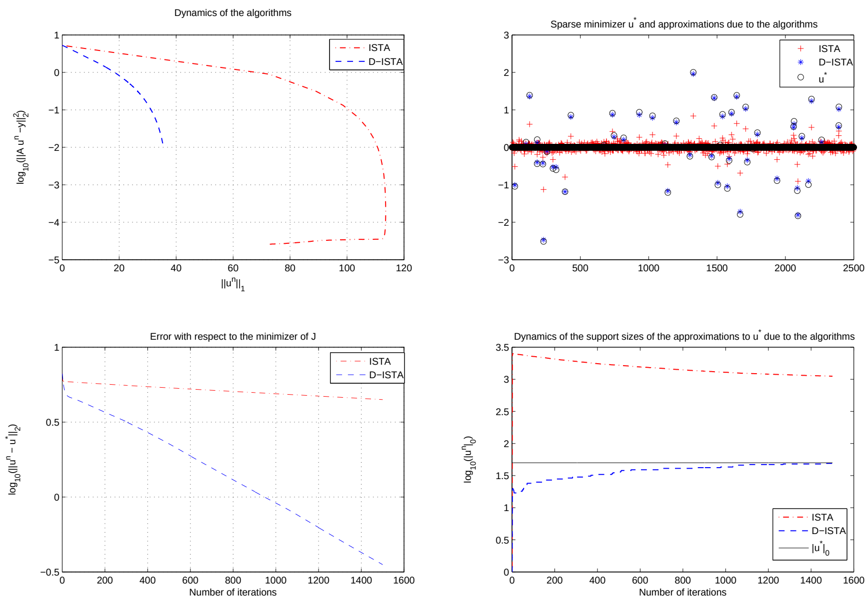


Figure 3.5 We show the behavior of the algorithms ISTA and D-ISTA in the computation of a sparse minimizer u^* for A being a 500×2500 matrix with i.i.d. Gaussian entries, $\alpha = 10^{-4}$, $\gamma_0 = 0.01$ and $\gamma = 0.998$. In the top left figure we presents the dynamics of the algorithms in the plane $\|u\|_{\ell_1} - \log(\|Au - y\|_2^2)$. On the bottom left, we show the absolute error to the precomputed minimizer u^* with respect to the number of iterations. On the bottom right we show how the size of the supports of the iterations grow with the number of iterations. The figure on the top right shows the vector u^*, and the approximations due to the algorithms. In this case D-ISTA detects the right support of u^* after 1500 iterations, whereas ISTA keeps dramatically far behind.

Related work

There exist by now several iterative methods that can be used for the minimization problem (3.109) in *finite dimensions*. We shall account a few of the most recently analyzed and discussed:

(a) the *GPSR-algorithm* (gradient projection for sparse reconstruction), another iterative projection method, in the auxiliary variables $x, y \geq 0$ with $u = x - y$ [39].

(b) the $\ell_1 - \ell_s$ *algorithm*, an interior point method using preconditioned conjugate gra- dient substeps (this method solves a linear system in each outer iteration step) [51].

(c) *FISTA* (fast iterative soft-thresholding algorithm) is a variation of the iterative soft-thresholding [5]. Define the operator $\Gamma(u) = \mathbb{S}_\alpha(u + A^*(y - Au))$. The

FISTA is defined as the iteration, starting for $u^{(0)} = 0$,

$$u^{(n+1)} = \Gamma\left(u^{(n)} + \frac{t^{(n)} - 1}{t^{(n+1)}}\left(u^{(n)} - u^{(n-1)}\right)\right),$$

where $t^{(n+1)} = \frac{1+\sqrt{1+4(t^{(n)})^2}}{2}$ and $t^{(0)} = 1$.

As is addressed in the recent paper [54] which accounts a very detailed comparison of these different algorithms, they do perform quite well when the regularization parameter α is sufficiently large, with a small advantage for GPSR. When α gets quite small all the algorithms, except for FISTA, deteriorate significantly their performances. Moreover, local conditioning properties of the linear operator A seem particularly affecting the performances of iterative algorithms.

While these methods are particularly suited for finite dimensional problems, it would be interesting to produce an effective strategy, for any range of the parameter α, for a large class of infinite dimensional problems. In the recent paper [22] the following ingredients are combined for this scope:

- *multiscale preconditioning* allows for a local well-conditioning of the matrix A and therefore reproduces at infinite dimension the conditions of best performances for iterative algorithms;
- *adaptivity* combined with a *decreasing thresholding strategy* allow for a *controlled* inflation of the support size of the iterations, promoting the minimal computational cost in terms of number of algebraic equations, as well as the exploitation from the very beginning of the iteration of the local well-conditioning of the matrix A.

In [66] the authors propose as well an adaptive method similar to [22] where, instead of the soft-thresholding, a *coarsening function*, i.e., a compressed hard-thresholding procedure, is implemented. The emphasis in the latter contribution is on the regularization properties of such an adaptive method which does not dispose of a reference energy functional (3.107), and it will be the object of the lectures presented by R. Ramlau, G. Teschke, and M. Zhariy.

3.3 Domain Decomposition Methods for ℓ_1-Minimization

Besides the elegant mathematics needed for the convergence proof, one of the major features of Algorithm 4 is its simplicity, also in terms of implementation. Indeed thresholding methods combined with wavelets have been often presented, e.g., in image processing, as a possible good alternative to total variation minimization which requires instead, as we already discussed in the previous sections, the solution of a degenerate partial differential equation. As pointed out in the previous sections, in general iterative soft-thresholding can converge very slowly.

In particular, it is practically not possible to use such an algorithm when the dimension of the problem is really large, unless we provide all the modifications we accounted above. And still for certain very large scale problems, this might not be enough. For that we need to consider further dimensionality reduction techniques. In this section we introduce a sequential domain decomposition method for the linear inverse problem with sparsity constraints modelled by (3.109). The goal is to join the simplicity of Algorithm 4 with a dimension reduction technique provided by a decomposition which will improve the convergence and the complexity of the algorithm without increasing the sophistication of the algorithm.

For simplicity, we start by decomposing the "domain" of the sequences $\mathcal{I}$ into two disjoint sets $\mathcal{I}_1, \mathcal{I}_2$ so that $\mathcal{I} = \mathcal{I}_1 \cup \mathcal{I}_2$. The extension to decompositions into multiple subsets $\mathcal{I} = \mathcal{I}_1 \cup \cdots \cup \mathcal{I}_\mathcal{N}$ follows from an analysis similar to the basic case $\mathcal{N} = 2$ Associated to a decomposition $\mathcal{C} = \{\mathcal{I}_1, \mathcal{I}_2\}$ we define the *extension operators* E_i : $\ell_2(\mathcal{I}_i) \to \ell_2(\mathcal{I})$, $(E_i v)_\lambda = v_\lambda$, if $\lambda \in \mathcal{I}_i$, $(E_i v)_\lambda = 0$, otherwise, $i = 1, 2$. The adjoint operator, which we call the *restriction operator*, is denoted by $R_i := E_i^*$. With these operators we may define the functional $\mathcal{J}(u_1, u_2)$, $\mathcal{J} : \ell_2(\mathcal{I}_1) \times \ell_2(\mathcal{I}_2) \to \mathbb{R}$, given by

$$\mathcal{J}(u_1, u_2) := \mathcal{J}(E_1 u_1 + E_2 u_2).$$

For the sequence u_i we use the notation $u_{\lambda,i}$ in order to denote its components. We want to formulate and to analyze the following algorithm: Pick an initial $E_1 u_1^{(0)} + E_2 u_2^{(0)} := u^{(0)} \in \ell_1(\mathcal{I})$, for example $u^{(0)} = 0$, and iterate

$$\begin{cases} u_1^{(n+1)} = \arg\min_{v_1 \in \ell_2(\mathcal{I}_1)} \mathcal{J}(v_1, u_2^{(n)}) \\ u_2^{(n+1)} = \arg\min_{v_2 \in \ell_2(\mathcal{I}_2)} \mathcal{J}(u_1^{(n+1)}, v_2) \\ u^{(n+1)} := E_1 u_1^{(n+1)} + E_2 u_2^{(n+1)}. \end{cases} \tag{3.146}$$

Let us observe that $\|E_1 u_1 + E_2 u_2\|_{\ell_1(\mathcal{I})} := \|u_1\|_{\ell_1(\mathcal{I}_1)} + \|u_2\|_{\ell_1(\mathcal{I}_2)}$, hence

$$\arg \min_{v_1 \in \ell_2(\mathcal{I}_1)} \mathcal{J}(v_1, u_2^{(n)}) = \arg \min_{v_1 \in \ell_2(\mathcal{I}_1)} \|(g - AE_2 u_2^{(n)}) - AE_1 v_1\|_Y^2 + \tau \|v_1\|_1.$$

A similar formulation holds for $\arg\min_{v_2 \in \ell_2(\mathcal{I}_2)} \mathcal{J}(u_1^{(n+1)}, v_2)$. This means that the solution of the local problems on $\mathcal{I}_i$ is of the *same* kind as the original problem $\arg\min_{u \in \ell_2(\mathcal{I})} \mathcal{J}(u)$, but the dimension for each has been reduced. Unfortunately the functionals $\mathcal{J}(u, u_2^{(n)})$ and $\mathcal{J}(u_1^{(n+1)}, v)$ do not need to have a unique minimizer. Therefore the formulation as in (3.162) is not in principle well defined. In the following we will consider a particular choice of the minimizers and in particular we will implement Algorithm 4 in order to solve each local problem. This choice leads to the following algorithm.

Algorithm 6. Pick an initial $E_1 u_1^{(0)} + E_2 u_2^{(0)} := u^{(0)} \in \ell_1(\mathcal{I})$, for example $u^{(0)} = 0$, and iterate

$$\begin{cases} \begin{cases} u_1^{(n+1,0)} = u_1^{(n,L)} \\ u_1^{(n+1,\ell+1)} = \mathbb{S}_\alpha \left(u_1^{(n+1,\ell)} + R_1 A^*((y - AE_2 u_2^{(n,M)}) - AE_1 u_1^{(n+1,\ell)}) \right) \\ \ell = 0, \dots, L-1 \end{cases} \\ \begin{cases} u_2^{(n+1,0)} = u_2^{(n,M)} \\ u_2^{(n+1,\ell+1)} = \mathbb{S}_\alpha \left(u_2^{(n+1,\ell)} + R_2 A^*((y - AE_1 u_1^{(n+1,L)}) - AE_2 u_2^{(n+1,\ell)}) \right) \\ \ell = 0, \dots, M-1 \end{cases} \\ u^{(n+1)} := E_1 u_1^{(n+1,L)} + E_2 u_2^{(n+1,M)}. \end{cases} \tag{3.147}$$

Of course, for $L = M = \infty$ the previous algorithm realizes a particular instance of (3.162). However, in practice we will never execute an infinite number of inner iterations and therefore it is important to analyze the convergence of the algorithm when $L, M \in \mathbb{N}$ are finite.

At this point the question is whether algorithm (3.147) really converges to a minimizer of the original functional $\mathcal{J}$. This is the scope of the following sections. Only for ease of notation, we assume now that the thresholding parameter $\alpha > 0$ is a scalar, hence $\mathbb{S}_\alpha(u)$ acts on (u) with the same thresholding $S_\alpha(u_\lambda)$ for each vector component u_λ.

3.3.1 Weak Convergence of the Sequential Algorithm

A main tool in the analysis of non-smooth functionals and their minima is the concept of subdifferential. We introduce it already in the presentation of the homotopy method in Section 2.1.1. Recall that for a convex functional F on some Banach space V its subdifferential $\partial F(x)$ at a point $x \in V$ with $F(x) < \infty$ is defined as the set

$$\partial F(x) \;=\; \{x^* \in V^*, x^*(z-x) + F(x) \;\leq\; F(z) \text{ for all } z \in V\},$$

where V^* denotes the dual space of V. It is obvious from this definition that $0 \in \partial F(x)$ if and only if x is a minimizer of F.

Example 3.17 Let $V = \ell_1(\mathcal{I})$ and $F(x) := \|x\|_1$ is the ℓ_1 norm. We have

$$\partial \| \cdot \|_1(x) \;=\; \{\xi \in \ell_\infty(\mathcal{I}) : \; \xi_\lambda \in \partial | \cdot |(x_\lambda), \lambda \in \mathcal{I}\} \tag{3.148}$$

where $\partial | \cdot |(z) = \{\text{sgn}(z)\}$ if $z \neq 0$ and $\partial | \cdot |(0) = [-1, 1]$.

By observing that $\partial(\|A\cdot -y\|_Y^2)(u) = \{2A^*(Au-y)\}$ and by an application of [36, Proposition 5.2] combined with the example above, we obtain the following characterizations of the subdifferentials of $\mathcal{J}$ and $\mathcal{J}^S$.

Lemma 3.18 *i) The subdifferential of $\mathcal{J}$ at u is given by*

$$\begin{aligned}\partial\mathcal{J}(u) &= 2A^*(Au-y) + 2\alpha\partial\|\cdot\|_1(u)\\ &= \{\xi \in \ell_\infty(\mathcal{I}) : \xi_\lambda \in [2A^*(Au-y)]_\lambda + 2\alpha\partial|\cdot|(u_\lambda)\}.\end{aligned}$$

ii) The subdifferential of $\mathcal{J}^S$ with respect to the sole component u is given by

$$\begin{aligned}\partial_u\mathcal{J}^S(u,a) &= -2(a + A^*(y-Aa)) + 2u + 2\alpha\partial\|\cdot\|_1(u)\\ &= \{\xi \in \ell_\infty(\mathcal{I}) : \xi_\lambda \in [-2(a + A^*(y-Aa))]_\lambda + 2u_\lambda + 2\alpha\partial|\cdot|(u_\lambda)\}.\end{aligned}$$

In light of Lemma 3.2 we can reformulate Algorithm 5 by

$$\begin{cases}\begin{cases}u_1^{(n+1,0)} = u_1^{(n,L)}\\ u_1^{(n+1,\ell+1)} = \arg\min_{u_1\in\ell_2(\mathcal{I}_1)} \mathcal{J}^S(E_1u_1 + E_2u_2^{(n,M)}, E_1u_1^{(n+1,\ell)} + E_2u_2^{(n,M)})\\ \ell = 0,\dots,L-1\end{cases}\\ \begin{cases}u_2^{(n+1,0)} = u_2^{(n,M)}\\ u_2^{(n+1,\ell+1)} = \arg\min_{u_2\in\ell_2(\mathcal{I}_2)} \mathcal{J}^S(E_1u_1^{(n+1,L)} + E_2u_2, E_1u_1^{(n+1,L)} + E_2u_2^{(n+1,\ell)})\\ \ell = 0,\dots,M-1\end{cases}\\ u^{(n+1)} := E_1u_1^{(n+1,L)} + E_2u_2^{(n+1,M)}.\end{cases} \tag{3.149}$$

Before we actually start proving the weak convergence of the algorithm in (3.210) we recall the following definition [67].

Definition 3.19 Let V be a topological space and $\mathcal{A} = (A_n)_{n\in\mathbb{N}}$ a sequence of subsets of V. The subset $A \subseteq V$ is called the *limit of the sequence* $\mathcal{A}$, and we write $A = \lim_n A_n$, if

$$A = \{a \in V : \exists a_n \in A_n, a = \lim_n a_n\}.$$

The following observation will be useful for us, see, e.g., [67, Proposition 8.7].

Lemma 3.20 *Assume that Γ is a convex function on $\mathbb{R}^M$ and $(x_n)_{n\in\mathbb{N}} \subset \mathbb{R}^M$ a convergent sequence with limit x such that $\Gamma(x_n), \Gamma(x) < \infty$. Then the subdifferentials satisfy*

$$\lim_{n\to\infty} \partial\Gamma(x_n) \subseteq \partial\Gamma(x).$$

In other words, the subdifferential $\partial\Gamma$ of a convex function is an outer semicontinuous set-valued function.

Theorem 3.21 (Weak convergence) *The algorithm in* (3.210) *produces a sequence* $(u^{(n)})_{n\in\mathbb{N}}$ *in* $\ell_2(\mathcal{I})$ *whose weak accumulation points are minimizers of the functional* $\mathcal{J}$*. In particular, the set of the weak accumulation points is non-empty and if* $u^{(\infty)}$ *is a weak accumulation point then*

$$u^{(\infty)} = \mathbb{S}_\alpha(u^{(\infty)} + A^*(g - Au^{(\infty)})).$$

Proof. Let us first observe that by (3.191)

$$\begin{aligned}\mathcal{J}(u^{(n)}) = \mathcal{J}^S(u^{(n)}, u^{(n)}) &= \mathcal{J}^S(E_1 u_1^{(n,L)} + E_2 u_2^{(n,M)}, E_1 u_1^{(n,L)} + E_2 u_2^{(n,M)}) \\ &= \mathcal{J}^S(E_1 u_1^{(n,L)} + E_2 u_2^{(n,M)}, E_1 u_1^{(n+1,0)} + E_2 u_2^{(n,M)}).\end{aligned}$$

By definition of $u_1^{(n+1,1)}$ and its minimal properties in (3.210) we have

$$\begin{aligned}&\mathcal{J}^S(E_1 u_1^{(n,L)} + E_2 u_2^{(n,M)}, E_1 u_1^{(n+1,0)} + E_2 u_2^{(n,M)}) \\ \geq\ &\mathcal{J}^S(E_1 u_1^{(n+1,1)} + E_2 u_2^{(n,M)}, E_1 u_1^{(n+1,0)} + E_2 u_2^{(n,M)}).\end{aligned}$$

Again, an application of (3.191) gives

$$\begin{aligned}&\mathcal{J}^S(E_1 u_1^{(n+1,1)} + E_2 u_2^{(n,M)}, E_1 u_1^{(n+1,0)} + E_2 u_2^{(n,M)}) \\ \geq\ &\mathcal{J}^S(E_1 u_1^{(n+1,1)} + E_2 u_2^{(n,M)}, E_1 u_1^{(n+1,1)} + E_2 u_2^{(n,M)})\end{aligned}$$

Putting in line these inequalities we obtain

$$\mathcal{J}(u^{(n)}) \geq \mathcal{J}^S(E_1 u_1^{(n+1,1)} + E_2 u_2^{(n,M)}, E_1 u_1^{(n+1,1)} + E_2 u_2^{(n,M)}).$$

In particular, from (3.192) we have

$$\mathcal{J}(u^{(n)}) - \mathcal{J}^S(E_1 u_1^{(n+1,1)} + E_2 u_2^{(n,M)}, E_1 u_1^{(n+1,1)} + E_2 u_2^{(n,M)}) \geq C\|u_1^{(n+1,1)} - u_1^{(n+1,0)}\|^2_{\ell_2(\mathcal{I}_1)}.$$

By induction we obtain

$$\begin{aligned}\mathcal{J}(u^{(n)}) &\geq \mathcal{J}^S(E_1 u_1^{(n+1,1)} + E_2 u_2^{(n,M)}, E_1 u_1^{(n+1,1)} + E_2 u_2^{(n,M)}) \geq \dots \\ &\geq \mathcal{J}^S(E_1 u_1^{(n+1,L)} + E_2 u_2^{(n,M)}, E_1 u_1^{(n+1,L)} + E_2 u_2^{(n,M)}) \\ &= \mathcal{J}(E_1 u_1^{(n+1,L)} + E_2 u_2^{(n,M)}),\end{aligned}$$

and

$$\mathcal{J}(u^{(n)}) - \mathcal{J}(E_1 u_1^{(n+1,L)} + E_2 u_2^{(n,M)}) \geq C \sum_{\ell=0}^{L-1} \|u_1^{(n+1,\ell+1)} - u_1^{(n+1,\ell)}\|^2_{\ell_2(\mathcal{I}_1)}.$$

By definition of $u_2^{(n+1,1)}$ and its minimal properties we have

$$\begin{aligned} &\mathcal{J}^S(E_1u_1^{(n+1,L)} + E_2u_2^{(n,M)}, E_1u_1^{(n+1,L)} + E_2u_2^{(n,M)}) \\ \geq\ &\mathcal{J}^S(E_1u_1^{(n+1,L)} + E_2u_2^{(n+1,1)}, E_1u_1^{(n+1,L)} + E_2u_2^{(n+1,0)}). \end{aligned}$$

By similar arguments as above we find

$$\mathcal{J}(u^{(n)}) \geq \mathcal{J}^S(E_1u_1^{(n+1,L)} + E_2u_2^{(n+1,M)}, E_1u_1^{(n+1,L)} + E_2u_2^{(n+1,M)}) = \mathcal{J}(u^{(n+1)}), \tag{3.150}$$

and

$$\begin{aligned} &\mathcal{J}(u^{(n)}) - \mathcal{J}(u^{(n+1)}) \\ \geq\ &C\left(\sum_{\ell=0}^{L-1}\|u_1^{(n+1,\ell+1)} - u_1^{(n+1,\ell)}\|^2_{\ell_2(\mathcal{I}_1)} + \sum_{m=0}^{M-1}\|u_2^{(n+1,m+1)} - u_2^{(n+1,m)}\|^2_{\ell_2(\mathcal{I}_2)}\right). \end{aligned} \tag{3.151}$$

From (3.193) we have $\mathcal{J}(u^{(0)}) \geq \mathcal{J}(u^{(n)}) \geq 2\alpha\|u^{(n)}\|_{\ell_1(\mathcal{I})} \geq 2\alpha\|u^{(n)}\|_{\ell_2(\mathcal{I})}$. This means that $(u^{(n)})_{n\in\mathbb{N}}$ is uniformly bounded in $\ell_2(\mathcal{I})$, hence there exists a weakly convergent subsequence $(u^{(n_j)})_{j\in\mathbb{N}}$. Let us denote $u^{(\infty)}$ the weak limit of the subsequence. For simplicity, we rename such subsequence by $(u^{(n)})_{n\in\mathbb{N}}$. Moreover, since the sequence $(\mathcal{J}(u^{(n)}))_{n\in\mathbb{N}}$ is monotonically decreasing and bounded from below by 0, it is also convergent. From (3.194) and the latter convergence we deduce

$$\left(\sum_{\ell=0}^{L-1}\|u_1^{(n+1,\ell+1)} - u_1^{(n+1,\ell)}\|^2_{\ell_2(\mathcal{I}_1)} + \sum_{m=0}^{M-1}\|u_2^{(n+1,m+1)} - u_2^{(n+1,m)}\|^2_{\ell_2(\mathcal{I}_2)}\right) \to 0, \quad n\to\infty. \tag{3.152}$$

In particular, by the standard inequality $(a^2+b^2) \geq \frac{1}{2}(a+b)^2$ for $a, b > 0$ and the triangle inequality, we have also

$$\|u^{(n)} - u^{(n+1)}\|_{\ell_2(\mathcal{I})} \to 0, \quad n \to \infty. \tag{3.153}$$

We would like now to show that

$$0 \in \lim_{n\to\infty} \partial\mathcal{J}(u^{(n)}) \subset \partial\mathcal{J}(u^{(\infty)}).$$

To this end, and in light of Lemma 3.18, we reason componentwise. By definition of $u_1^{(n+1,L)}$ we have

$$\begin{aligned} 0 \in [-2(u_1^{(n+1,L-1)} + R_1A^*((y - AE_2u_2^{(n,M)}) - AE_1u_1^{(n+1,L-1)}))]_\lambda \\ + 2u_{\lambda,1}^{(n+1,L)} + 2\alpha\partial|\cdot|(u_{\lambda,1}^{(n+1,L)}), \end{aligned}$$

for $\lambda \in \mathcal{I}_1$, and by definition of $u_2^{(n+1,M)}$ we have

$$0 \in [-2(u_2^{(n+1,M-1)} + R_2 A^*((y - AE_1 u_1^{(n+1,L)}) - AE_2 u_2^{(n+1,M-1)}))]_\lambda + 2u_{\lambda,2}^{(n+1,M)} + 2\alpha\partial|\cdot|(u_{\lambda,2}^{(n+1,M)}),$$

for $\lambda \in \mathcal{I}_2$. Let us compute $\partial\mathcal{J}(u^{(n+1)})_\lambda$,

$$\partial\mathcal{J}(u^{(n+1)})_\lambda = [-2A^*(y - AE_1 u_1^{(n+1,L)} - AE_2 u_2^{(n+1,M)})]_\lambda + 2\alpha\partial|\cdot|(u_{\lambda,i}^{(n+1,K)}), \tag{3.154}$$

where $\lambda \in \mathcal{I}_i$ and $K = L, M$ for $i = 1, 2$ respectively. We would like to find a $\xi_\lambda^{(n+1)} \in \partial\mathcal{J}(u^{(n+1)})_\lambda$ such that $\xi_\lambda^{(n+1)} \to 0$ for $n \to \infty$. By (3.154) we have that for $\lambda \in \mathcal{I}_1$

$$0 = [-2(u_1^{(n+1,L-1)} + R_1 A^*((y - AE_2 u_2^{(n,M)}) - AE_1 u_1^{(n+1,L-1)}))]_\lambda + 2u_{\lambda,1}^{(n+1,L)} + 2\alpha\xi_{\lambda,1}^{(n+1)},$$

for a $\xi_{\lambda,1}^{(n+1)} \in \partial|\cdot|(u_{\lambda,1}^{(n+1,L)})$, and, by (3.154), for $\lambda \in \mathcal{I}_2$

$$0 = [-2(u_2^{(n+1,M-1)} + R_2 A^*((y - AE_1 u_1^{(n+1,L)}) - AE_2 u_2^{(n+1,M-1)}))]_\lambda + 2u_{\lambda,2}^{(n+1,M)} + 2\alpha\xi_{\lambda,2}^{(n+1)},$$

for a $\xi_{\lambda,2}^{(n+1)} \in \partial|\cdot|(u_{\lambda,2}^{(n+1,M)})$. Thus by adding zero from (3.154) as represented by the previous two formulas, we can choose

$$\begin{aligned}\xi_\lambda^{(n+1)} &= 2(u_{\lambda,1}^{(n+1,L)} - u_{\lambda,1}^{(n+1,L-1)}) + [R_1 A^* AE_1(u_1^{(n+1,L)} - u_1^{(n+1,L-1)})]_\lambda \\ &+ [R_1 A^* AE_2(u_2^{(n+1,M)} - u_1^{(n,M)})]_\lambda,\end{aligned}$$

if $\lambda \in \mathcal{I}_1$ and

$$\xi_\lambda^{(n+1)} = 2(u_{\lambda,2}^{(n+1,M)} - u_{\lambda,2}^{(n+1,M-1)}) + [R_2 A^* AE_1(u_2^{(n+1,M)} - u_1^{(n+1,M-1)})]_\lambda,$$

if $\lambda \in \mathcal{I}_2$. For both these choices, from (3.195) and (3.196), and by continuity of A, we have $\xi_\lambda^{(n+1)} \to 0$ for $n \to \infty$. By continuity of A, weak convergence of $u^{(n)}$ (which implies componentwise convergence), and Lemma 3.20 we obtain

$$0 \in \lim_{n\to\infty} \partial\mathcal{J}(u^{(n)})_\lambda \subset \partial\mathcal{J}(u^{(\infty)})_\lambda, \quad \forall\lambda \in \mathcal{I}.$$

It follows from Lemma 3.18 that $0 \in \partial\mathcal{J}(u^{(\infty})$. By the properties of the subdifferential we have that $u^{(\infty)}$ is a minimizer of $\mathcal{J}$. Of course, the reasoning above holds for any weakly convergent subsequence and therefore all weak accumulation points of the original sequence $(u^{(n)})_n$ are minimizers of $\mathcal{J}$.

Similarly, by taking now the limit for $n \to \infty$ in (3.154) and (3.154), and by using (3.195) we obtain

$$0 \in [-2(R_1 u^{(\infty)} + R_1 A^*((y - AE_2 R_2 u^{(\infty)}) - AE_1 R_1 u^{(\infty)}))]_\lambda + 2u_\lambda^{(\infty)} + 2\alpha\partial|\cdot|(u_\lambda^{(\infty)}),$$

for $\lambda \in \mathcal{I}_1$ and

$$0 \in [-2(R_2 u^{(\infty)} + R_2 A^*((y - AE_1 R_1 u^{(\infty)}) - AE_2 R_2 u^{(\infty)})]_\lambda + 2u_\lambda^{(\infty)} + 2\alpha\partial|\cdot|(u_\lambda^{(\infty)}).$$

for $\lambda \in \mathcal{I}_2$. In other words, we have

$$0 \in \partial_u \mathcal{J}^S(u^{(\infty)}, u^{(\infty)}).$$

An application of Lemma 3.18 and Proposition 3.5 imply

$$u^{(\infty)} = \mathbb{S}_\alpha(u^{(\infty)} + A^*(y - Au^{(\infty)})).$$

□

Remark 3.22 1. Because $u^{(\infty)} = \mathbb{S}_\alpha(u^{(\infty)} + A^*(y - Au^{(\infty)}))$, we could infer the minimality of $u^{(\infty)}$ by invoking Proposition 3.5. In the previous proof we wanted to present an alternative argument based on differential inclusions.

2. Since $(u^{(n)})_{n\in\mathbb{N}}$ is bounded and (3.195) holds, also $(u_i^{n,\ell})_{n,\ell}$ are bounded for $i = 1, 2$.

3.3.2 Strong Convergence of the Sequential Algorithm

In this section we want to show that the convergence of a subsequence $(u^{n_j})_j$ to any accumulation point $u^{(\infty)}$ holds not only in the weak topology, but also in the Hilbert space $\ell_2(\mathcal{I})$ norm. Let us define

$$\begin{aligned} \eta^{(n+1)} &:= u_1^{(n+1,L)} - u_1^{(\infty)}, \qquad \eta^{(n+1/2)} := u_1^{(n+1,L-1)} - u_1^{(\infty)}, \\ \mu^{(n+1)} &:= u_2^{(n+1,M)} - u_2^{(\infty)}, \qquad \mu^{(n+1/2)} := u_2^{(n+1,M-1)} - u_2^{(\infty)}, \end{aligned}$$

where $u_i^{(\infty)} := R_i u^{(\infty)}$. From Theorem 3.35 we also have

$$u_i^{(\infty)} = \mathbb{S}_\alpha(\underbrace{u_i^{(\infty)} + R_i T^*(g - TE_1 u_1^{(\infty)} - TE_2 u_2^{(\infty)})}_{:=h_i}), \quad i = 1, 2.$$

Let us also denote $h := E_1 h_1 + E_2 h_2$ and $\xi^{(n)} := E_1 \eta^{(n+1/2)} + E_2 \mu^{(n+1/2)}$.

For the proof of strong convergence we need the following technical lemmas. Their strategy of their proofs is similar to that of Lemma 3.8 and Lemma 3.9.

Lemma 3.23 $\|A\xi^{(n)}\|_Y^2 \to 0$ *for* $n \to \infty$.

Proof. Since

$$\begin{aligned}
\eta^{(n+1)} - \eta^{(n+1/2)} = \mathbb{S}_\alpha(h_1 &+ (I - R_1 A^* A E_1)\eta^{(n+1/2)} \\
&- R_1 A^* A E_2 \mu^{(n)}) - \mathbb{S}_\alpha(h_1) - \eta^{(n+1/2)}, \\
\mu^{(n+1)} - \mu^{(n+1/2)} = \mathbb{S}_\alpha(h_2 &+ (I - R_2 A^* A E_2)\mu^{(n+1/2)} \\
&- R_2 A^* A E_1 \eta^{(n+1)}) - \mathbb{S}_\alpha(h_2) - \mu^{(n+1/2)},
\end{aligned}$$

and $\|\eta^{(n+1)} - \eta^{(n+1/2)}\|_{\ell_2(\mathcal{I}_1)} = \|u_1^{(n+1,L)} - u_1^{(n+1,L-1)}\|_{\ell_2(\mathcal{I}_1)} \to 0$, $\|\mu^{(n+1)} - \mu^{(n+1/2)}\|_{\ell_2(\mathcal{I}_1)} = \|u_2^{(n+1,M)} - u_2^{(n+1,M-1)}\|_{\ell_2(\mathcal{I}_2)} \to 0$ by (3.195), we have

$$\begin{aligned}
&\|\mathbb{S}_\alpha(h_1 + (I - R_1 A^* A E_1)\eta^{(n+1/2)} - R_1 A^* A E_2 \mu^{(n)}) - \mathbb{S}_\alpha(h_1) - \eta^{(n+1/2)}\|_{\ell_2(\mathcal{I}_1)} \\
\geq\ & \Big| \|\mathbb{S}_\alpha(h_1 + (I - R_1 A^* A E_1)\eta^{(n+1/2)} - R_1 A^* A E_2 \mu^{(n)}) \\
&\qquad - \mathbb{S}_\alpha(h_1)\|_{\ell_2(\mathcal{I}_1)} - \|\eta^{(n+1/2)}\|_{\ell_2(\mathcal{I}_1)} \Big| \to 0, \qquad (3.155)
\end{aligned}$$

and

$$\begin{aligned}
&\|\mathbb{S}_\alpha(h_2 + (I - R_2 A^* A E_2)\mu^{(n+1/2)} - R_2 A^* A E_1 \eta^{(n+1)}) - \mathbb{S}_\alpha(h_2) - \mu^{(n+1/2)}\|_{\ell_2(\mathcal{I}_2)} \\
\geq\ & \Big| \|\mathbb{S}_\alpha(h_2 + (I - R_2 A^* A E_2)\mu^{(n+1/2)} - R_2 A^* A E_1 \eta^{(n+1)}) \\
&\qquad - \mathbb{S}_\alpha(h_2)\|_{\ell_2(\mathcal{I}_2)} - \|\mu^{(n+1/2)}\|_{\ell_2(\mathcal{I}_2)} \Big| \to 0. \qquad (3.156)
\end{aligned}$$

By nonexpansiveness of $\mathbb{S}_\alpha$ we have the estimates

$$\begin{aligned}
&\|\mathbb{S}_\alpha(h_2 + (I - R_2 A^* A E_2)\mu^{(n+1/2)} - R_2 A^* A E_1 \eta^{(n+1)}) - \mathbb{S}_\alpha(h_2)\|_{\ell_2(\mathcal{I}_2)} \\
\leq\ & \|(I - R_2 A^* A E_2)\mu^{(n+1/2)} - R_2 A^* A E_1 \eta^{(n+1)}\|_{\ell_2(\mathcal{I}_2)} \\
\leq\ & \|(I - R_2 A^* A E_2)\mu^{(n+1/2)} - R_2 A^* A E_1 \eta^{(n+1/2)}\|_{\ell_2(\mathcal{I}_2)} \\
+\ & \underbrace{\|R_2 A^* A E_1(\eta^{(n+1/2)}) - \eta^{(n+1)})\|_{\ell_2(\mathcal{I}_2)}}_{:=\varepsilon^{(n)}}.
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
&\|\mathbb{S}_\alpha(h_1 + (I - R_1 A^* A E_1)\eta^{(n+1/2)} - R_1 A^* A E_2 \mu^{(n)}) - \mathbb{S}_\alpha(h_1)\|_{\ell_2(\mathcal{I}_1)} \\
\leq\ & \|(I - R_1 A^* A E_1)\eta^{(n+1/2)} - R_1 A^* A E_2 \mu^{(n)}\|_{\ell_2(\mathcal{I}_1)} \\
\leq\ & \|(I - R_1 A^* A E_1)\eta^{(n+1/2)} - R_1 A^* A E_2 \mu^{(n+1/2)}\|_{\ell_2(\mathcal{I}_1)} \\
+\ & \underbrace{\|R_1 A^* A E_2(\mu^{(n+1/2)} - \mu^{(n)})\|_{\ell_2(\mathcal{I}_1)}}_{\delta^{(n)}}.
\end{aligned}$$

Combining the previous inequalities, we obtain the estimates

$$\begin{aligned}
& \|\mathbb{S}_\alpha(h_1 + (I - R_1 A^* A E_1)\eta^{(n+1/2)} - R_1 A^* A E_2 \mu^{(n)} - \mathbb{S}_\alpha(h_1)\|^2_{\ell_2(\mathcal{I}_1)} \\
+ & \|\mathbb{S}_\alpha(h_2 + (I - R_2 A^* A E_2)\mu^{(n+1/2)} - R_2 A^* A E_1 \eta^{(n+1)} - \mathbb{S}_\alpha(h_2)\|^2_{\ell_2(\mathcal{I}_2)} \\
\leq & \|(I - R_1 A^* A E_1)\eta^{(n+1/2)} - R_1 A^* A E_2 \mu^{(n)}\|^2_{\ell_2(\mathcal{I}_1)} \\
+ & \|(I - R_2 A^* A E_2)\mu^{(n+1/2)} - R_2 A^* A E_1 \eta^{(n+1)}\|^2_{\ell_2(\mathcal{I}_2)} \\
= & \Big(\|(I - R_1 A^* A E_1)\eta^{(n+1/2)} - R_1 A^* A E_2 \mu^{(n+1/2)}\|_{\ell_2(\mathcal{I}_1)} \\
+ & \|R_1 A^* A E_2(\mu^{(n+1/2)} - \mu^{(n)})\|_{\ell_2(\mathcal{I}_1)}\Big)^2 \\
+ & \Big(\|(I - R_2 A^* A E_2)\mu^{(n+1/2)} - R_2 A^* A E_1 \eta^{(n+1/2)}\|_{\ell_2(\mathcal{I}_2)} \\
+ & \|R_2 A^* A E_1(\eta^{(n+1/2)}) - \eta^{(n+1)})\|_{\ell_2(\mathcal{I}_2)}\Big)^2 \\
\leq & \|(I - A^* A)\xi^{(n)}\|^2_{\ell_2(\mathcal{I})} + ((\varepsilon^{(n)})^2 + (\delta^{(n)})^2 + C'(\varepsilon^{(n)} + \delta^{(n)})) \\
\leq & \|\xi^{(n)}\|^2_{\ell_2(\mathcal{I})} + ((\varepsilon^{(n)})^2 + (\delta^{(n)})^2 + C'(\varepsilon^{(n)} + \delta^{(n)}))
\end{aligned}$$

The constant $C' > 0$ is due to the boundedness of $u^{(n,\ell)}$. Certainly, by (3.195), for every $\varepsilon > 0$ there exists n_0 such that for $n > n_0$ we have $(\varepsilon^{(n)})^2 + (\delta^{(n)})^2 + C'(\varepsilon^{(n)} + \delta^{(n)})) \leq \varepsilon$. Therefore, if

$$\begin{aligned}
& \|\mathbb{S}_\alpha(h_1 + (I - R_1 A^* A E_1)\eta^{(n+1/2)} - R_1 A^* A E_2 \mu^{(n)} - \mathbb{S}_\alpha(h_1)\|^2_{\ell_2(\mathcal{I}_1)} \\
+ & \|\mathbb{S}_\alpha(h_2 + (I - R_2 A^* A E_2)\mu^{(n+1/2)} - R_2 A^* A E_1 \eta^{(n+1)} - \mathbb{S}_\alpha(h_2)\|^2_{\ell_2(\mathcal{I}_2)} \geq \|\xi^{(n)}\|^2_{\ell_2(\mathcal{I})},
\end{aligned}$$

then

$$\begin{aligned}
0 \leq & \|(I - R_1 A^* A E_1)\mu^{(n+1/2)} - R_1 A^* A E_2 \mu^{(n)}\|^2_{\ell_2(\mathcal{I}_1)} \\
+ & \|(I - R_2 A^* A E_2)\mu^{(n+1/2)} - R_2 A^* A E_1 \eta^{(n+1)}\|^2_{\ell_2(\mathcal{I})} - \|\xi^{(n)}\|^2_{\ell_2(\mathcal{I})} \\
\leq & (\varepsilon^{(n)})^2 + (\delta^{(n)})^2 + C'(\varepsilon^{(n)} + \delta^{(n)})) \leq \varepsilon
\end{aligned}$$

If, instead, we have

$$\begin{aligned}
& \|\mathbb{S}_\alpha(h_1 + (I - R_1 A^* A E_1)\eta^{(n+1/2)} - R_1 A^* A E_2 \mu^{(n)}) - \mathbb{S}_\alpha(h_1)\|^2_{\ell_2(\mathcal{I}_1)} \\
+ & \|\mathbb{S}_\alpha(h_2 + (I - R_2 A^* A E_2)\mu^{(n+1/2)} - R_2 A^* A E_1 \eta^{(n+1)}) - \mathbb{S}_\alpha(h_2)\|^2_{\ell_2(\mathcal{I}_2)} < \|\xi^{(n)}\|^2_{\ell_2(\mathcal{I})},
\end{aligned}$$

then by (3.155) and (3.156)

$$\begin{aligned}
\|\xi^{(n)}\|^2_{\ell_2(\mathcal{I})} &- \Big(\|(I - R_1A^*AE_1)\mu^{(n+1/2)} - R_1A^*AE_2\mu^{(n)})\|^2_{\ell_2(\mathcal{I}_1)} \\
&+ \|(I - R_2A^*AE_2)\mu^{(n+1/2)} - R_2A^*AE_1\eta^{(n+1)})\|^2_{\ell_2(\mathcal{I}_2)}\Big) \\
&\leq \|\xi^{(n)}\|^2_{\ell_2(\mathcal{I})} \\
&- \|\mathbb{S}_\alpha(h_1 + (I - R_1A^*AE_1)\eta^{(n+1/2)} - R_1A^*AE_2\mu^{(n)}) - \mathbb{S}_\alpha(h_1)\|^2_{\ell_2(\mathcal{I}_1)} \\
&- \|\mathbb{S}_\alpha(h_2 + (I - R_2A^*AE_2)\mu^{(n+1/2)} - R_2A^*AE_1\eta^{(n+1)}) - \mathbb{S}_\alpha(h_2)\|^2_{\ell_2(\mathcal{I}_2)} \\
&= \Big|\|\xi^{(n)}\|^2_{\ell_2(\mathcal{I})} \\
&- \|\mathbb{S}_\alpha(h_1 + (I - R_1A^*AE_1)\eta^{(n+1/2)} - R_1A^*AE_2\mu^{(n)}) - \mathbb{S}_\alpha(h_1)\|^2_{\ell_2(\mathcal{I}_1)} \\
&- \|\mathbb{S}_\alpha(h_2 + (I - R_2A^*AE_2)\mu^{(n+1/2)} - R_2A^*AE_1\eta^{(n+1)}) - \mathbb{S}_\alpha(h_2)\|^2_{\ell_2(\mathcal{I}_2)}\Big| \\
&\leq \Big|\|\eta^{(n+1/2)}\|^2_{\ell_2(\mathcal{I}_1)} \\
&- \|\mathbb{S}_\alpha(h_1 + (I - R_1A^*AE_1)\eta^{(n+1/2)} - R_1A^*AE_2\mu^{(n)}) - \mathbb{S}_\alpha(h_1)\|^2_{\ell_2(\mathcal{I}_1)}\Big| \\
&+ \Big|\|\mu^{(n+1/2)}\|^2_{\ell_2(\mathcal{I}_2)} \\
&- \|\mathbb{S}_\alpha(h_2 + (I - R_2A^*AE_2)\mu^{(n+1/2)} - R_2A^*AE_1\eta^{(n+1)}) - \mathbb{S}_\alpha(h_2)\|^2_{\ell_2(\mathcal{I}_2)}\Big| \leq \varepsilon
\end{aligned}$$

for n large enough. This implies

$$\begin{aligned}
\lim_{n\to\infty}\Big[\|\xi^{(n)}\|^2_{\ell_2(\mathcal{I})} &- \Big(\|(I - R_1A^*AE_1)\mu^{(n+1/2)} - R_1A^*AE_2\mu^{(n)}\|^2_{\ell_2(\mathcal{I}_1)} \\
&+ \|(I - R_2A^*AE_2)\mu^{(n+1/2)} - R_2A^*AE_1\eta^{(n+1)}\|^2_{\ell_2(\mathcal{I}_2)}\Big)\Big] = 0
\end{aligned}$$

Observe now that

$$\begin{aligned}
&\|(I - R_1A^*AE_1)\mu^{(n+1/2)} - R_1A^*AE_2\mu^{(n)}\|^2_{\ell_2(\mathcal{I}_1)} \\
+ &\|(I - R_2A^*AE_2)\mu^{(n+1/2)} - R_2A^*AE_1\eta^{(n+1)}\|^2_{\ell_2(\mathcal{I}_2)} \\
\leq &(\|(I - R_1A^*AE_1)\mu^{(n+1/2)} - R_1A^*AE_2\mu^{(n+1/2)}\|_{\ell_2(\mathcal{I}_1)} + \delta^{(n)})^2 \\
+ &(\|(I - R_2A^*AE_2)\mu^{(n+1/2)} - R_2A^*AE_1\eta^{(n+1/2)}\|_{\ell_2(\mathcal{I}_2)} + \varepsilon^{(n)})^2 \\
\leq &\|(I - A^*A)\xi^{(n)}\|^2_{\ell_2(\mathcal{I})} + \Big((\varepsilon^{(n)})^2 + (\delta^{(n)})^2 + 2C'(\varepsilon^{(n)} + \delta^{(n)})\Big),
\end{aligned}$$

for a suitable constant $C' > 0$ as above. Therefore we have

$$\begin{aligned}
\|\xi^{(n)}\|^2_{\ell_2(\mathcal{I})} &- \Big(\|(I - R_1 A^* A E_1)\mu^{(n+1/2)} - R_1 A^* A E_2 \mu^{(n)})\|^2_{\ell_2(\mathcal{I}_1)} \\
&+ \|(I - R_2 A^* A E_2)\mu^{(n+1/2)} - R_2 A^* A E_1 \eta^{(n+1)})\|^2_{\ell_2(\mathcal{I}_2)}\Big) \\
&\geq \|\xi^{(n)}\|^2_{\ell_2(\mathcal{I})} - \|(I - A^*A)\xi^{(n)}\|^2_{\ell_2(\mathcal{I})} - \Big((\varepsilon^{(n)})^2 + (\delta^{(n)})^2 + 2C'(\varepsilon^{(n)} + \delta^{(n)})\Big) \\
&= 2\|A\xi^{(n)}\|^2_Y - \|A^*A\xi^{(n)}\|^2_{\ell_2(\mathcal{I})} - \Big((\varepsilon^{(n)})^2 + (\delta^{(n)})^2 + 2C'(\varepsilon^{(n)} + \delta^{(n)})\Big) \\
&\geq \|A\xi^{(n)}\|^2_Y - \Big((\varepsilon^{(n)})^2 + (\delta^{(n)})^2 + 2C'(\varepsilon^{(n)} + \delta^{(n)})\Big).
\end{aligned}$$

This implies $\|A\xi^{(n)}\|^2_Y \to 0$ for $n \to \infty$.

□

Lemma 3.24 *For* $h = E_1 h_1 + E_2 h_2$, $\|\mathbb{S}_\alpha(h + \xi^{(n)}) - \mathbb{S}_\alpha(h) - \xi^{(n)}\|_{\ell_2(\mathcal{I})} \to 0$, *for* $n \to \infty$.

Proof. We have

$$\begin{aligned}
&\mathbb{S}_\alpha(h + \xi^{(n)} - A^*A\xi^{(n)}) \\
= \;& E_1\Big(\mathbb{S}_\alpha(h_1 + (I - R_1 A^* A E_1)\eta^{(n+1/2)} - R_1 A^* A E_2 \mu^{(n+1/2)})\Big) \\
+ \;& E_2\Big(\mathbb{S}_\alpha(h_2 + (I - R_2 A^* A E_2)\mu^{(n+1/2)} - R_2 A^* A E_1 \eta^{(n+1/2)})\Big)
\end{aligned}$$

Therefore, we can write

$$\begin{aligned}
&\mathbb{S}_\alpha(h + \xi^{(n)} - A^*A\xi^{(n)}) \\
= \;& E_1\Big[\mathbb{S}_\alpha(h_1 + (I - R_1 A^* A E_1)\eta^{(n+1/2)} - R_1 A^* A E_2 \mu^{(n)}) \\
+ \;& \mathbb{S}_\alpha(h_1 + (I - R_1 A^* A E_1)\eta^{(n+1/2)} - R_1 A^* A E_2 \mu^{(n+1/2)}) \\
- \;& \mathbb{S}_\alpha(h_1 + (I - R_1 A^* A E_1)\eta^{(n+1/2)} - R_1 A^* A E_2 \mu^{(n)})\Big] \\
+ \;& E_2\Big[\mathbb{S}_\alpha(h_2 + (I - R_2 A^* A E_2)\mu^{(n+1/2)} - R_2 A^* A E_1 \eta^{(n+1)}) \\
+ \;& \mathbb{S}_\alpha(h_2 + (I - R_2 A^* A E_2)\mu^{(n+1/2)} - R_2 A^* A E_1 \eta^{(n+1/2)}) \\
- \;& \mathbb{S}_\alpha(h_2 + (I - R_2 A^* A E_2)\mu^{(n+1/2)} - R_2 A^* A E_1 \eta^{(n+1)})\Big].
\end{aligned}$$

By using the nonexpansiveness of $\mathbb{S}_\alpha$, the boundedness of the operators E_i, R_i, A^*A, and the triangle inequality we obtain,

$$\begin{aligned}
& \|\mathbb{S}_\alpha(h+\xi^{(n)}) - \mathbb{S}_\alpha(h) - \xi^{(n)}\|_{\ell_2(\mathcal{I})} \\
\leq\ & \|\mathbb{S}_\alpha(h+\xi^{(n)} - A^*A\xi^{(n)}) - \mathbb{S}_\alpha(h) - \xi^{(n)}\|_{\ell_2(\mathcal{I})} \\
+\ & \|\mathbb{S}_\alpha(h+\xi^{(n)}) - \mathbb{S}_\alpha(h+\xi^{(n)} - A^*A\xi^{(n)})\|_{\ell_2(\mathcal{I})} \\
\leq\ & \Bigg(\underbrace{\|\mathbb{S}_\alpha(h_1 + (I - R_1A^*AE_1)\eta^{(n+1/2)} - R_1A^*AE_2\mu^{(n)}) - \mathbb{S}_\alpha(h_1) - \eta^{(n+1/2)}\|_{\ell_2(\mathcal{I}_1)}}_{:=A^{(n)}} \\
+\ & \underbrace{\|\mathbb{S}_\alpha(h_2 + (I - R_2A^*AE_2)\mu^{(n+1/2)} - R_2A^*AE_1\eta^{(n+1)}) - \mathbb{S}_\alpha(h_2) - \mu^{(n+1/2)}\|_{\ell_2(\mathcal{I}_2)}}_{:=B^{(n)}} \\
+\ & \underbrace{\|\mu^{(n+1/2)} - \mu^{(n)}\|_{\ell_2(\mathcal{I}_2)} + \|\eta^{(n+1)} - \eta^{(n+1/2)}\|_{\ell_2(\mathcal{I}_1)}}_{:=C^{(n)}} \\
+\ & \underbrace{\|A^*A\xi^{(n)}\|_{\ell_2(\mathcal{I})}}_{:=D^{(n)}} \Bigg).
\end{aligned}$$

The quantities $A^{(n)}, B^{(n)}$ vanish for $n \to \infty$ because of (3.155) and (3.156). The quantity $C^{(n)}$ vanishes for $n \to \infty$ because of (3.195), and $D^{(n)}$ vanishes $n \to \infty$ thanks to Lemma 3.23. □

By combining the previous technical achievements, we can now state the strong convergence.

Theorem 3.25 (Strong convergence) *Algorithm 5 produces a sequence $(u^{(n)})_{n\in\mathbb{N}}$ in $\ell_2(\mathcal{I})$ whose strong accumulation points are minimizers of the functional $\mathcal{J}$. In particular, the set of strong accumulation points is non-empty.*

Proof. Let $u^{(\infty)}$ be a weak accumulation point and let $(u^{(n_j)})_{j\in\mathbb{N}}$ be a subsequence weakly convergent to $u^{(\infty)}$. Let us denote the latter sequence $(u^{(n)})_{n\in\mathbb{N}}$ again. With the notation used in this section, by Theorem 3.35 and (3.195) we have that $\xi^{(n)} = E_1\eta^{(n+1/2)} + E_2\mu^{(n+1/2)}$ weakly converges to zero. By Lemma 3.24 we have $\lim_{n\to\infty} \|\mathbb{S}_\alpha(h+\xi^{(n)}) - \mathbb{S}_\alpha(h) - \xi^{(n)}\|_{\ell_2(\mathcal{I})} = 0$. From Lemma 3.10 we conclude that $\xi^{(n)} = E_1\eta^{(n+1/2)} + E_2\mu^{(n+1/2)}$ converges to zero strongly. Again by (3.195) we have that $(u^{(n)})_{n\in\mathbb{N}}$ converges to $u^{(\infty)}$ strongly. □

3.3.3 A Parallel Domain Decomposition Method

The most natural modification to (3.147) in order to obtain a parallelizable algorithm is to substitute the term $u^{(n+1,L)}$ with $R_1 u^{(n)}$ in the second inner iterations. This makes the inner iterations on $\mathcal{I}_1$ and $\mathcal{I}_2$ mutually independent, hence executable by two processors at the same time. We obtain the following algorithm: Pick an initial $u^{(0)} \in \ell_1(\mathcal{I})$, for example $u^{(0)} = 0$, and iterate

$$\begin{cases} \begin{cases} u_1^{(n+1,0)} = R_1 u^{(n)} \\ u_1^{(n+1,\ell+1)} = \mathbb{S}_\alpha\left(u_1^{(n+1,\ell)} + R_1 A^*((g - AE_2 R_2 u^{(n)}) - AE_1 u_1^{(n+1,\ell)})\right) \\ \ell = 0, \dots, L-1 \end{cases} \\ \begin{cases} u_2^{(n+1,0)} = R_2 u^{(n)} \\ u_2^{(n+1,\ell+1)} = \mathbb{S}_\alpha\left(u_2^{(n+1,\ell)} + R_2 A^*((g - AE_1 R_1 u^{(n)}) - AE_2 u_2^{(n+1,\ell)})\right) \\ \ell = 0, \dots, M-1 \end{cases} \\ u^{(n+1)} := E_1 u_1^{(n+1,L)} + E_2 u_2^{(n+1,M)}. \end{cases} \tag{3.157}$$

The behavior of this algorithm is somehow bizzare. Indeed, the algorithm usually alternates between the two subsequences given by $u^{(2n)}$ and its consecutive iteration $u^{(2n+1)}$. These two sequences are complementary, in the sense that they encode independent patterns of the solution. In particular, for $u^{(\infty)} = u' + u''$, $u^{(2n)} \approx u'$ and $u^{(2n+1)} \approx u''$ for n not too large. During the iterations and for n large the two subsequences slowly approach to each other, merging the complementary features and shaping the final limit which usually coincides with the wanted minimal solution, see Figure 3.6. Unfortunately, this “oscillatory behavior” makes impossible to prove monotonicity of the sequence $(\mathcal{J}(u^{(n)}))_{n\in\mathbb{N}}$ and we have no proof of convergence. However, since the subsequences are early indicating different features of

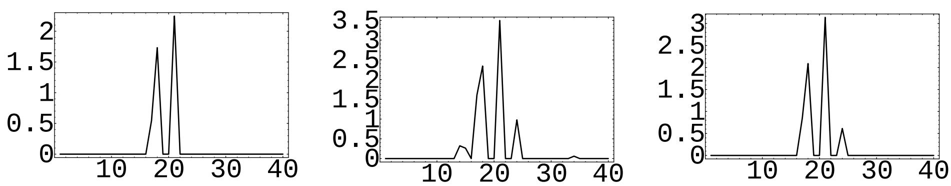


Figure 3.6 On the left we show $u^{(2n)}$, in the center $u^{(2n+1)}$, and on the right $u^{(\infty)}$. The two consecutive iterations contain different features which will appear in the solution.

the final limit, we may modify the algorithm by substituting $u^{(n+1)} := E_1 u_1^{(n+1,L)} + E_2 u_2^{(n+1,M)}$ with $u^{(n+1)} := \frac{(E_1 u_1^{(n+1,L)} + E_2 u_2^{(n+1,M)}) + u^{(n)}}{2}$ that is the average of the current iteration and the previous one. This enforces an early merging of complementary features and leads to the following algorithm:

Algorithm 7. Pick an initial $u^{(0)} \in \ell_1(\mathcal{I})$, for example $u^{(0)} = 0$, and iterate

$$\begin{cases} \begin{cases} u_1^{(n+1,0)} = R_1 u^{(n)} \\ u_1^{(n+1,\ell+1)} = \mathbb{S}_\alpha \left(u_1^{(n+1,\ell)} + R_1 A^*((g - AE_2 R_2 u^{(n)}) - AE_1 u_1^{(n+1,\ell)}) \right) \\ \ell = 0, \dots, L-1 \end{cases} \\ \begin{cases} u_2^{(n+1,0)} = R_2 u^{(n)} \\ u_2^{(n+1,\ell+1)} = \mathbb{S}_\alpha \left(u_2^{(n+1,\ell)} + R_2 A^*((g - AE_1 R_1 u^{(n)}) - AE_2 u_2^{(n+1,\ell)}) \right) \\ \ell = 0, \dots, M-1 \end{cases} \\ u^{(n+1)} := \frac{(E_1 u_1^{(n+1,L)} + E_2 u_2^{(n+1,M)}) + u^{(n)}}{2}. \end{cases} \tag{3.158}$$

The proof of strong convergence of this algorithm is very similar to the one of Algorithm 6. For the details, we refer the reader to [42].

3.4 Domain Decomposition Methods for Total Variation Minimization

We would like to continue our parallel discussion of ℓ_1-minimization as well as total variation minimization as we did in Section 2.1.3. In particular, we would like to show that also for total variation minimization it is possible to formulate domain decomposition methods. Several numerical strategies to perform efficiently total variation minimization have been proposed in the literature as well. Without claiming of being exhaustive, we list a few of the relevant methods, ordered by their chronological appearance:

(i) the linearization approach of Vogel et al. [32] and of Chambolle and Lions [16] by iteratively re-weighted least squares, see also Section 2.1.3;

(ii) the primal-dual approach of Chan et al. [17];

(iii) variational approximation via locally quadratic functionals as in the work of Vese et al. [2, 77];

(iv) iterative thresholding algorithms based on projections onto convex sets as in the work of Chambolle [14] as well as in the work of Combettes and Wajs [21];

(v) iterative minimization of the Bregman distance as in the work of Osher et al. [65] (also notice the very recent Bregman split approach [48]);

(vi) graph cuts [15, 26] for the minimization of (2.79) with $T = I$ (the identity operator) and an anisotropic total variation;

(vii) the approach proposed by Nesterov [59] and its modifications by Weiss et al. [78].

These approaches differ significantly, and they provide a convincing view of the interest this problem has been able to generate and of its applicative impact. However, because of their iterative-sequential formulation, none of the mentioned methods is able to address in real-time, or at least in an acceptable computational time, extremely large problems, such as 4D imaging (spatial plus temporal dimensions) from functional magnetic-resonance in nuclear medical imaging, astronomical imaging or global terrestrial seismic tomography. For such large scale simulations we need to address methods which allow us to reduce the problem to a finite sequence of subproblems of a more manageable size, perhaps computable by one of the methods listed above. With this aim we introduced subspace correction and domain decomposition methods both for ℓ_1-norm and total variation minimizations [42, 44, 69]. We address the interested reader to the broad literature included in [44] for an introduction to domain decompositions methods both for PDEs and convex minimization.

Difficulty of the problem

Due to the nonsmoothness and nonadditivity of the total variation with respect to a nonoverlapping domain decomposition (note that the total variation of a function on the whole domain equals the sum of the total variations on the subdomains plus the size of the jumps at the interfaces [44, formula (3.4)]; this is one of the main differences to the situation we already encountered with ℓ_1-minimization), one encounters additional difficulties in showing convergence of such decomposition strategies to global minimizers. In particular, we stress very clearly that well-known approaches as in [13, 18, 75, 76] are not directly applicable to this problem, because either they do address additive problems (as the one of ℓ_1-minimization) or smooth convex minimizations, which is *not* the case of total variation minimization. Moreover the interesting solutions may be discontinuous, e.g., along curves in 2D. These discontinuities may cross the interfaces of the domain decomposition patches. Hence, the crucial difficulty is the correct numerical treatment of interfaces, with the preservation of crossing discontinuities and the correct matching where the solution is continuous instead, see [44, Section 7.1.1].

The work [44] was particularly addressed to *nonoverlapping* domain decompositions $\Omega_1 \cup \Omega_2 \subset \Omega \subset \bar{\Omega}_1 \cup \bar{\Omega}_2$ and $\Omega_1 \cap \Omega_2 = \emptyset$. Associated to the decomposition define $V_i = \{u \in L^2(\Omega) : \operatorname{supp}(u) \subset \Omega_i\}$, for $i = 1, 2$; note that $L^2(\Omega) = V_1 \oplus V_2$. With this splitting we wanted to minimize $\mathcal{J}$ by suitable instances of the following alternating algorithm: Pick an initial $V_1 \oplus V_2 \ni u_1^{(0)} + u_2^{(0)} := u^{(0)}$, for example $u^{(0)} = 0$, and iterate

$$\begin{cases} u_1^{(n+1)} \approx \arg\min_{v_1 \in V_1} \mathcal{J}(v_1 + u_2^{(n)}) \\ u_2^{(n+1)} \approx \arg\min_{v_2 \in V_2} \mathcal{J}(u_1^{(n+1)} + v_2) \\ u^{(n+1)} := u_1^{(n+1)} + u_2^{(n+1)}. \end{cases}$$

In [44] an implementation of this algorithm is proposed, it is guaranteed to converge, and to decrease the objective energy $\mathcal{J}$ monotonically. One could prove its convergence to minimizers of $\mathcal{J}$ only under technical conditions on the interfaces of the subdomains. However, in numerical experiments, the algorithm seems always converging robustly to the expected minimizer. This discrepancy between theoretical analysis and numerical evidences motivated the investigation on *overlapping* domain decompositions. The hope is that the redundancy given by overlapping patches and the avoidance of boundary interfaces could allow for a technically easier theoretical analysis.

The approach, results, and technical issues

In this section we show how to adapt Algorithm 5 and Algorithm 6 to the case of an *overlapping* domain decompositions for total variation minimization. The setting of an overlapping domain decomposition eventually provides us with a framework in which one can successfully prove its convergence to minimizers of $\mathcal{J}$, both in its sequential and parallel forms. Let us stress that to our knowledge this is the first method which addresses a domain decomposition strategy for total variation minimization with a formal theoretical justification of convergence. It is important to mention that there are other very recent attempts of addressing domain decomposition methods for total variation minimization with successful numerical results [57].

The analysis is performed for a discrete approximation of the continuous functional (2.79), for ease again denoted $\mathcal{J}$ in (2.80). Essentially we approximate functions u by their sampling on a regular grid and their gradient Du by finite differences ∇u. It is well-known that such a discrete approximation Γ-converges to the continuous functional (see [9]). In particular, discrete minimizers of (2.80), interpolated by piecewise linear functions, converge in weak-$*$-topology of BV to minimizers of the functional (2.79) in the continuous setting. Of course, when dealing with numerical solutions, only the discrete approach matters together with its approximation properties to the continuous problem. However, the need of working in the discrete setting is not only practical, it is also topological. In fact bounded sets in BV are (only) weakly-$*$-compact, and this property is fundamental for showing that certain sequences have converging subsequences. Unfortunately, the weak-$*$-topology of BV is "too weak" for our purpose of proving convergence of the domain decomposition algorithm; for instance, the trace on boundary sets is *not* a continuous operator with respect to this topology. This difficulty can be avoided, for instance, by Γ-approximating the functional (2.79) by means of quadratic functionals (as in [2, 16, 77]) and working with the topology of $W^{1,2}(\Omega)$, the Sobolev space of L^2-functions with L^2-distributional first derivatives. However, this strategy changes the singular nature of the problem which makes it both interesting and difficult. Hence, the discrete approach has the virtues of being practical for numerical implementations, of correctly approximating the continuous setting, and of retaining the major features which makes the problem interesting. Note further that in the discrete setting where topological issues are not a concern anymore, also the dimension d can be arbitrary, contrary to the continuous setting where the dimension d has to be linked to boundedness properties of the operator T, see [77, property H2, pag. 134]. For ease of presentation, and in order to avoid unnecessary technicalities, we limit our analysis to splitting the problem into two subdomains Ω_1 and Ω_2. This is by no means a restriction. The generalization to multiple domains comes quite natural in our specific setting, see also [44, Remark 5.3]. When dealing with discrete subdomains Ω_i, for technical reasons, we will require a certain splitting property for the total variation, i.e.,

$$|\nabla u|(\Omega) = |\nabla u|_{\Omega_1}|(\Omega_1)+c_1(u|_{(\Omega_2\setminus\Omega_1)\cup\Gamma_1}), \quad |\nabla u|(\Omega) = |\nabla u|_{\Omega_2}|(\Omega_2)+c_2(u|_{(\Omega_1\setminus\Omega_2)\cup\Gamma_2}), \tag{3.159}$$

where c_1 and c_2 are suitable functions which depend only on the restrictions $u|_{(\Omega_2\setminus\Omega_1)\cup\Gamma_1}$ and $u|_{(\Omega_1\setminus\Omega_2)\cup\Gamma_2}$ respectively, see (3.166) (symbols and notations are clarified once for all in the following section). Note that this formula is the discrete analogous of [44, formula (3.4)] in the continuous setting. The simplest examples of discrete domains with such a property are discrete d-dimensional rectangles (d-*orthotopes*). Hence, for ease of presentation, we will assume to work with d-orthotope domains, also noting that such decompositions are already sufficient for any practical use in image processing, and stressing that the results can be generalized also to subdomains with different shapes as long as (3.159) is satisfied.

Additional notations

Additionally to the notations already introduced in Section 2.1.3 for the total variation minimization setting, we consider also the closed convex set

$$K := \left\{ \operatorname{div} p : p \in \mathcal{H}^d, |p(x)|_\infty \leq 1 \text{ for all } x \in \Omega \right\},$$

where $|p(x)|_\infty = \max\left\{|p^1(x)|,\dots,|p^d(x)|\right\}$, and denote $P_K(u) = \operatorname{argmin}_{v\in K} \|u - v\|_2$ the *orthogonal projection onto* K.

3.4.1 The Overlapping Domain Decomposition Algorithm

As before we are interested in the minimization of the functional

$$\mathcal{J}(u) := \|Ku - g\|_2^2 + 2\alpha \left|\nabla(u)\right|(\Omega), \tag{3.160}$$

where $K \in \mathcal{L}(\mathcal{H})$ is a linear operator, $g \in \mathcal{H}$ is a datum, and $\alpha > 0$ is a fixed constant. and we assume that $1 \notin \ker(K)$.

Now, instead of minimizing (3.160) on the whole domain we decompose Ω into two overlapping subdomains Ω_1 and Ω_2 such that $\Omega = \Omega_1 \cup \Omega_2$, $\Omega_1 \cap \Omega_2 \neq \emptyset$, and (3.159) is fulfilled. For consistency of the definitions of gradient and divergence, we assume that also the subdomains Ω_i are discrete d-orthotopes as well as Ω, stressing that this is by no means a restriction, but only for ease of presentation. Due to this domain decomposition $\mathcal{H}$ is split into two closed subspaces $V_j = \{u \in \mathcal{H} : \operatorname{supp}(u) \subset \Omega_j\}$, for $j = 1, 2$. Note that $\mathcal{H} = V_1 + V_2$ is not a direct sum of V_1 and V_2, but just a linear sum of subspaces. Thus any $u \in \mathcal{H}$ has a nonunique representation

$$u(x) = \begin{cases} u_1(x) & x \in \Omega_1 \setminus \Omega_2 \\ u_1(x) + u_2(x) & x \in \Omega_1 \cap \Omega_2 \\ u_2(x) & x \in \Omega_2 \setminus \Omega_1 \end{cases}, \quad u_i \in V_i, \quad i = 1, 2. \tag{3.161}$$

We denote by Γ_1 the interface between Ω_1 and $\Omega_2 \setminus \Omega_1$ and by Γ_2 the interface between Ω_2 and $\Omega_1 \setminus \Omega_2$ (the interfaces are naturally defined in the discrete setting).

We introduce the trace operator of the restriction to a boundary Γ_i

$$\mathrm{Tr}\,|_{\Gamma_i} \colon V_i \to \mathbb{R}^{\Gamma_i}, \;\; i = 1, 2$$

with $\mathrm{Tr}\,|_{\Gamma_i} v_i = v_i\,|_{\Gamma_i}$, the restriction of v_i on Γ_i. Note that $\mathbb{R}^{\Gamma_i}$ is as usual the set of maps from Γ_i to $\mathbb{R}$. The trace operator is clearly a linear and continuous operator. We additionally fix a *bounded uniform partition of unity* (BUPU) $\{\chi_1, \chi_2\} \subset \mathcal{H}$ such that

(a) $\mathrm{Tr}\,|_{\Gamma_i} \chi_i = 0$ for $i = 1, 2$,

(b) $\chi_1 + \chi_2 = 1$,

(c) $\operatorname{supp} \chi_i \subset \Omega_i$ for $i = 1, 2$,

(d) $\max\{\|\chi_1\|_\infty, \|\chi_2\|_\infty\} = \kappa < \infty$.

We would like to solve

$$\mathrm{argmin}_{u \in \mathcal{H}}\, \mathcal{J}(u)$$

by picking an initial $V_1 + V_2 \ni \tilde{u}_1^{(0)} + \tilde{u}_2^{(0)} := u^{(0)} \in \mathcal{H}$, e.g., $\tilde{u}_i^{(0)} = 0, i = 1, 2$, and iterate

$$\begin{cases} u_1^{(n+1)} \approx \mathrm{argmin}_{\substack{v_1 \in V_1 \\ \mathrm{Tr}|_{\Gamma_1} v_1 = 0}} \mathcal{J}(v_1 + \tilde{u}_2^{(n)}) \\ u_2^{(n+1)} \approx \mathrm{argmin}_{\substack{v_2 \in V_2 \\ \mathrm{Tr}|_{\Gamma_2} v_2 = 0}} \mathcal{J}(u_1^{(n+1)} + v_2) \\ u^{(n+1)} := u_1^{(n+1)} + u_2^{(n+1)} \\ \tilde{u}_1^{(n+1)} := \chi_1 \cdot u^{(n+1)} \\ \tilde{u}_2^{(n+1)} := \chi_2 \cdot u^{(n+1)}. \end{cases} \tag{3.162}$$

Note that we are minimizing over functions $v_i \in V_i$ for $i = 1, 2$ which vanish on the interior boundaries, i.e., $\mathrm{Tr}\,|_{\Gamma_i} v_i = 0$. Moreover $u^{(n)}$ is the sum of the local minimizers $u_1^{(n)}$ and $u_2^{(n)}$, which are not uniquely determined on the overlapping part. Therefore we introduced a suitable correction by χ_1 and χ_2 in order to force the subminimizing sequences $(u_1^{(n)})_{n \in \mathbb{N}}$ and $(u_2^{(n)})_{n \in \mathbb{N}}$ to keep uniformly bounded. This issue will be explained in detail below, see Lemma 3.36. From the definition of χ_i, $i = 1, 2$, it is clear that

$$u_1^{(n+1)} + u_2^{(n+1)} = u^{(n+1)} = (\chi_1 + \chi_2) u^{(n+1)} = \tilde{u}_1^{(n+1)} + \tilde{u}_2^{(n+1)}.$$

Note that in general $u_1^{(n)} \neq \tilde{u}_1^{(n)}$ and $u_2^{(n)} \neq \tilde{u}_2^{(n)}$. In (3.162) we use "≈" (the approximation symbol) because in practice we never perform the exact minimization. In the following section we discuss how to realize the approximation to the individual subspace minimizations.

3.4.2 Local Minimization by Lagrange Multipliers

Let us consider, for example, the subspace minimization on Ω_1

$$\operatorname{argmin}_{\substack{v_1\in V_1\\ \mathrm{Tr}|_{\Gamma_1} v_1=0}} \mathcal{J}(v_1+u_2) = \operatorname{argmin}_{\substack{v_1\in V_1\\ \mathrm{Tr}|_{\Gamma_1} v_1=0}} \|Kv_1-(g-Ku_2)\|_2^2+2\alpha\left|\nabla(v_1+u_2\mid_{\Omega_1})\right|(\Omega). \tag{3.163}$$

First of all, observe that $\{u\in\mathcal{H} : \mathrm{Tr}\,|_{\Gamma_1} u = \mathrm{Tr}\,|_{\Gamma_1} u_2,\ \mathcal{J}(u)\le C\} \subset \{\mathcal{J}\le C\}$, hence the former set is also bounded by assumption (C) and the minimization problem (3.163) has solutions.

It is useful to us to consider again a surrogate functional $\mathcal{J}_1^s$ of $\mathcal{J}$: Assume $a, u_1 \in V_1$, $u_2\in V_2$, and define

$$\mathcal{J}_1^s(u_1+u_2,a) := \mathcal{J}(u_1+u_2)+\|u_1-a\|_2^2-\|K(u_1-a)\|_2^2. \tag{3.164}$$

A straightforward computation shows that

$$\mathcal{J}_1^s(u_1+u_2,a) = \|u_1-(a+(K^*(g-Ku_2-Ka))\mid_{\Omega_1})\|_2^2+2\alpha\left|\nabla(u_1+u_2)\right|(\Omega)+\Phi(a,g,u_2),$$

where Φ is a function of a, g, u_2 only. Note that now the variable u_1 is not anymore effected by the action of K. Consequently, we want to realize an approximate solution to (3.163) by using the following algorithm: For $u_1^{(0)} = \tilde{u}_1^{(0)} \in V_1$,

$$u_1^{(\ell+1)} = \operatorname{argmin}_{\substack{u_1\in V_1\\ \mathrm{Tr}|_{\Gamma_1} u_1=0}} \mathcal{J}_1^s(u_1+u_2,u_1^{(\ell)}), \quad \ell\ge 0. \tag{3.165}$$

Additionally in (3.165) we can restrict the total variation on Ω_1 only, since we have

$$\left|\nabla(u_1+u_2)\right|(\Omega) \quad = \quad \left|\nabla(u_1+u_2)\mid_{\Omega_1}\right|(\Omega_1)+c_1(u_2|_{(\Omega_2\setminus\Omega_1)\cup\Gamma_1}). \tag{3.166}$$

where we used (3.159) and the assumption that u_1 vanishes on the interior boundary Γ_1. Hence (3.165) is equivalent to

$$\operatorname{argmin}_{\substack{u_1\in V_1\\ \mathrm{Tr}|_{\Gamma_1} u_1=0}} \mathcal{J}_1^s(u_1+u_2,u_1^{(\ell)}) = \operatorname{argmin}_{\substack{u_1\in V_1\\ \mathrm{Tr}|_{\Gamma_1} u_1=0}} \|u_1-z_1\|_2^2+2\alpha\left|\nabla(u_1+u_2)\mid_{\Omega_1}\right|(\Omega_1),$$

where $z_1 = u_1^{(\ell)}+(K^*(g-Ku_2-Ku_1^{(\ell)}))\mid_{\Omega_1}$. Similarly the same arguments work for the second subproblem.

Before proving the convergence of this algorithm, we need to clarify first how to practically compute $u_1^{(\ell+1)}$ for $\tilde{u}_1^{(\ell)}$ given. To this end we need to introduce further notions and to recall some useful results.

Generalized Lagrange multipliers for nonsmooth objective functions

We consider the following problem

$$\operatorname{argmin}_{x \in V} \{F(x) : Gx = b\}, \tag{3.167}$$

where $G : V \to V$ is a linear operator on V. We have the following useful result.

Theorem 3.26 *[49, Theorem 2.1.4, p. 305] Let* $N = \{G^* \lambda : \ \lambda \in V\} = \operatorname{Range}(G^*)$. *Then,* $x_0 \in \{x \in V : \ G(x) = b\}$ *solves the constrained minimization problem* (3.167) *if and only if*

$$0 \in \partial F(x_0) + N.$$

We want to exploit Theorem 3.26 in order to produce an algorithmic solution to each iteration step (3.165), which practically stems from the solution of a problem of this type

$$\operatorname{argmin}_{\substack{u_1 \in V_1 \\ \operatorname{Tr}|_{\Gamma_1} u_1 = 0}} \|u_1 - z_1\|_2^2 + 2\alpha \left|\nabla (u_1 + u_2 \mid_{\Omega_1})\right| (\Omega_1).$$

It is well-known how to solve this problem if $u_2 \equiv 0$ in $\bar{\Omega}_1$ and the trace condition is not imposed. For the general case we propose the following solution strategy. In what follows all the involved quantities are restricted to Ω_1, e.g., $u_1 = u_1 \mid_{\Omega_1}, u_2 = u_2 \mid_{\Omega_1}$.

Theorem 3.27 (Oblique thresholding) *For* $u_2 \in V_2$ *and for* $z_1 \in V_1$ *the following statements are equivalent:*

(i) $u_1^* = \operatorname{argmin}_{\substack{u_1 \in V_1 \\ \operatorname{Tr}|_{\Gamma_1} u_1 = 0}} \|u_1 - z_1\|_2^2 + 2\alpha \left|\nabla (u_1 + u_2)\right| (\Omega_1)$*;*

(ii) there exists $\eta \in \operatorname{Range}(\operatorname{Tr}|_{\Gamma_1})^* = \{\eta \in V_1$ *with* $\operatorname{supp}(\eta) = \Gamma_1\}$ *such that* $0 \in u_1^* - (z_1 - \eta) + \alpha \partial_{V_1} \left|\nabla (\cdot + u_2)\right| (\Omega_1)(u_1^*)$*;*

(iii) there exists $\eta \in V_1$ *with* $\operatorname{supp}(\eta) = \Gamma_1$ *such that* $u_1^* = (I - P_{\alpha K})(z_1 + u_2 - \eta) - u_2 \in V_1$ *and* $\operatorname{Tr}|_{\Gamma_1} u_1^* = 0$*;*

(iv) there exists $\eta \in V_1$ *with* $\operatorname{supp}(\eta) = \Gamma_1$ *such that* $\operatorname{Tr}|_{\Gamma_1} \eta = \operatorname{Tr}|_{\Gamma_1} z_1 + \operatorname{Tr}|_{\Gamma_1} P_{\alpha K}(\eta - (z_1 + u_2))$ *or equivalently* $\eta = (\operatorname{Tr}|_{\Gamma_1})^* \operatorname{Tr}|_{\Gamma_1} (z_1 + P_{\alpha K}(\eta - (z_1 + u_2)))$.

We call the solution operation provided by this theorem an *oblique thresholding*, in analogy to the terminology for ℓ_1-minimization (see Lemma 3.1), because it performs a thresholding of the derivatives, i.e., it sets to zero most of the derivatives of $u = u_1 + u_2 \approx z_1$ on Ω_1, provided u_2 which is a fixed vector in V_2.

Proof. Let us show the equivalence between (i) and (ii). The problem in (i) can be reformulated as

$$u_1^* = \operatorname{argmin}_{u_1 \in V_1} \{F(u_1) := \|u_1 - z_1\|_2^2 + 2\alpha \left|\nabla (u_1 + u_2)\right| (\Omega_1), \operatorname{Tr}|_{\Gamma_1} u_1 = 0\}. \tag{3.168}$$

Recall that $\mathrm{Tr}\,|_{\Gamma_1}\colon V_1 \to \mathbb{R}^{\Gamma_1}$ is a surjective map with closed range. This means that $(\mathrm{Tr}\,|_{\Gamma_1})^*$ is injective and that $\mathrm{Range}(\mathrm{Tr}\,|_{\Gamma_1})^* = \{\eta \in V_1 \text{ with } \mathrm{supp}(\eta) = \Gamma_1\}$ is closed. Using Theorem 3.26 the optimality of u_1^* is equivalent to the existence of $\eta \in \mathrm{Range}(\mathrm{Tr}\,|_{\Gamma_1})^*$ such that

$$0 \in \partial_{V_1} F(u_1^*) + 2\eta. \tag{3.169}$$

Due to the continuity of $\|u_1 - z_1\|_2^2$ in V_1, we have, by [36, Proposition 5.6], that

$$\partial_{V_1} F(u_1^*) = 2(u_1^* - z_1) + 2\alpha\partial_{V_1} |\nabla(\cdot + u_2)|\,(\Omega_1)(u_1^*). \tag{3.170}$$

Thus, the optimality of u_1^* is equivalent to

$$0 \in u_1^* - z_1 + \eta + \alpha\partial_{V_1} |\nabla(\cdot + u_2)|\,(\Omega_1)(u_1^*). \tag{3.171}$$

This concludes the equivalence of (i) and (ii). Let us show now that (iii) is equivalent to (ii). The condition in (iii) can be rewritten as

$$\xi^* = (I - P_{\alpha K})(z_1 + u_2 - \eta), \qquad \xi^* = u_1^* + u_2.$$

Since $|\nabla(\cdot)| \geq 0$ is 1-homogeneous and lower-semicontinuous, by [44, Example 4.2.2], the latter is equivalent to

$$0 \in \xi^* - (z_1 + u_2 - \eta) + \alpha\partial_{V_1} |\nabla(\cdot)|\,(\Omega_1)(\xi^*),$$

and equivalent to (ii). Note that in particular we have $\partial_{V_1} |\nabla(\cdot)|\,(\Omega_1)(\xi^*) = \partial_{V_1} |\nabla(\cdot + u_2)|\,(\Omega_1)(u_1^*)$, which is easily shown by a direct computation from the definition of subdifferential.

We prove now the equivalence between (iii) and (iv). We have

$$\begin{aligned} u_1^* &= (I - P_{\alpha K})(z_1 + u_2 - \eta) - u_2 \in V_1, \quad \eta \in V_1 \text{ with } \mathrm{supp}(\eta) = \Gamma_1, \mathrm{Tr}\,|_{\Gamma_1} u_1^* = 0 \\ &= z_1 - \eta - P_{\alpha K}(z_1 + u_2 - \eta). \end{aligned}$$

By applying $\mathrm{Tr}\,|_{\Gamma_1}$ to both sides of the latter equality we get

$$0 = \mathrm{Tr}\,|_{\Gamma_1} z_1 - \mathrm{Tr}\,|_{\Gamma_1} \eta - \mathrm{Tr}\,|_{\Gamma_1} P_{\alpha K}(z_1 + u_2 - \eta).$$

By observing that $-\mathrm{Tr}\,|_{\Gamma_1} P_{\alpha K}(\xi) = \mathrm{Tr}\,|_{\Gamma_1} P_{\alpha K}(-\xi)$, we obtain the fixed point equation

$$\mathrm{Tr}\,|_{\Gamma_1} \eta = \mathrm{Tr}\,|_{\Gamma_1} z_1 + \mathrm{Tr}\,|_{\Gamma_1} P_{\alpha K}(\eta - (z_1 + u_2)). \tag{3.172}$$

Conversely, since all the considered quantities in

$$(I - P_{\alpha K})(z_1 + u_2 - \eta) - u_2$$

are in V_1, the whole expression is an element in V_1 and hence u_1^* as defined in (iii) is an element in V_1 and $\mathrm{Tr}\,|_{\Gamma_1} u_1^* = 0$. This shows the equivalence between (iii) and (iv) and therewith finishes the proof. □

We wonder now whether any of the conditions in Theorem 3.27 is indeed practically satisfied. In particular, we want to show that $\eta \in V_1$ as in (iii) or (iv) of the previous theorem is provided as the limit of the following iterative algorithm:

$$\eta^{(0)} \in V_1, \operatorname{supp} \eta^{(0)} = \Gamma_1 \quad \eta^{(m+1)} = (\mathrm{Tr}\,|_{\Gamma_1})^* \,\mathrm{Tr}\,|_{\Gamma_1} \left(z_1 + P_{\alpha K}(\eta^{(m)} - (z_1 + u_2))\right), \quad m \geq 0 \tag{3.173}$$

Proposition 3.28 *The following statements are equivalent:*

(i) there exists $\eta \in V_1$ such that $\eta = (\mathrm{Tr}\,|_{\Gamma_1})^* \,\mathrm{Tr}\,|_{\Gamma_1} (z_1 + P_{\alpha K}(\eta - (z_1 + u_2)))$ (which is in turn the condition (iv) of Theorem 3.27)

(ii) the iteration (3.173) converges to any $\eta \in V_1$ that satisfies (3.172).

For the proof of this Proposition we need to recall some well-known notions and results.

Definition 3.29 A nonexpansive map $\mathcal{T} : \mathcal{H} \to \mathcal{H}$ is strongly nonexpansive if for $(u_n - v_n)_n$ bounded and $\|\mathcal{T}(u_n) - \mathcal{T}(v_n)\|_2 - \|u_n - v_n\|_2 \to 0$ we have

$$u_n - v_n - (\mathcal{T}(u_n) - \mathcal{T}(v_n)) \to 0, \quad n \to \infty.$$

Proposition 3.30 (Corollaries 1.3, 1.4, and 1.5 [11]) *Let $\mathcal{T} : \mathcal{H} \to \mathcal{H}$ be a strongly nonexpansive map. Then $\operatorname{fix} \mathcal{T} = \{u \in \mathcal{H} : \mathcal{T}(u) = u\} \neq \emptyset$ if and only if $(\mathcal{T}^n u)_n$ converges to a fixed point $u_0 \in \operatorname{fix} \mathcal{T}$ for any choice of $u \in \mathcal{H}$.*

Proof. (Proposition 3.28) Projections onto convex sets are strongly nonexpansive [4, Corollary 4.2.3]. Moreover, the composition of strongly nonexpansive maps is strongly nonexpansive [11, Lemma 2.1]. By an application of Proposition 3.30 we immediately have the result, since any map of the type $\mathcal{T}(\xi) = Q(\xi) + \xi_0$ is strongly nonexpansive whenever Q is (this is a simple observation from the definition of strongly nonexpansive maps). Indeed, we are looking for fixed points of $\eta = (\mathrm{Tr}\,|_{\Gamma_1})^* \,\mathrm{Tr}\,|_{\Gamma_1} (z_1 + P_{\alpha K}(\eta - (z_1 + u_2)))$ or, equivalently, of $\xi = \underbrace{(\mathrm{Tr}\,|_{\Gamma_1})^* \,\mathrm{Tr}\,|_{\Gamma_1} P_{\alpha K}}_{:=Q}(\xi) - \underbrace{((\mathrm{Tr}\,|_{\Gamma_1})^* \,\mathrm{Tr}\,|_{\Gamma_1} u_2)}_{:=\xi_0}$, where $\xi = (\mathrm{Tr}\,|_{\Gamma_1})^* \,\mathrm{Tr}\,|_{\Gamma_1} (\eta - (z_1 + u_2))$. □

Convergence of the subspace minimization

From the results of the previous section it follows that the iteration (3.165) can be explicitly computed by

$$u_1^{(\ell+1)} = S_\alpha(u_1^{(\ell)} + K^*(g - Ku_2 - Ku_1^{(\ell)}) + u_2 - \eta^{(\ell)}) - u_2, \tag{3.174}$$

where $S_\alpha := I - P_{\alpha K}$ and $\eta^{(\ell)} \in V_1$ is any solution of the fixed point equation

$$\begin{aligned}\eta &= (\mathrm{Tr}\,|_{\Gamma_1})^* \,\mathrm{Tr}\,|_{\Gamma_1} \Big((u_1^{(\ell)} K^*(g - Ku_2 - Ku_1^{(\ell)})) \\ &- P_{\alpha K}(u_1^{(\ell)} + K^*(g - Ku_2 - Ku_1^{(\ell)} + u_2 - \eta))\Big).\end{aligned}$$

The computation of $\eta^{(\ell)}$ can be implemented by the algorithm (3.173).

Proposition 3.31 *Assume $u_2 \in V_2$ and $\|K\| < 1$. Then the iteration* (3.174) *converges to a solution $u_1^* \in V_1$ of* (3.163) *for any initial choice of $u_1^{(0)} \in V_1$.*

The proof of this proposition is similar to the one of Theorem 3.7 and it is omitted.

Let us conclude this section mentioning that all the results presented here hold symmetrically for the minimization on V_2, and that the notations should be just adjusted accordingly.

3.4.3 Convergence of the Sequential Alternating Subspace Minimization

In this section we want to prove the convergence of the algorithm (3.162) to minimizers of $\mathcal{J}$. In order to do that, we need a characterization of solutions of the minimization problem (2.80) as the one provided in [77, Proposition 4.1] for the continuous setting. We specify the arguments in [77, Proposition 4.1] for our discrete setting and we highlight the significant differences with respect to the continuous one.

Characterization of solutions

We make the following assumptions:

(A_φ) $\varphi : \mathbb{R} \to \mathbb{R}$ is a convex function, nondecreasing in $\mathbb{R}^+$ such that

- (i) $\varphi(0) = 0$.
- (ii) There exist $c > 0$ and $b \geq 0$ such that $cz - b \leq \varphi(z) \leq cz + b$, for all $z \in \mathbb{R}^+$.

The particular example we have in mind is simply $\varphi(s) = s$, but we keep a more general notation for uniformity with respect to the continuous version in [77, Proposition 4.1]. In this section we are concerned with the following more general minimization problem

$$\mathrm{argmin}_{u \in \mathcal{H}} \{\mathcal{J}_\varphi(u) := \|Ku - g\|_2^2 + 2\alpha\varphi(|\nabla u|)(\Omega)\} \tag{3.175}$$

where $g \in \mathcal{H}$ is a datum, $\alpha > 0$ is a fixed constant (in particular for $\varphi(s) = s$).

To characterize the solution of the minimization problem (3.175) we use duality results from [36]. Therefore we recall the definition of the *conjugate (or Legendre transform)* of a function (for example see [36, Def. 4.1, pag. 17]):

Definition 3.32 Let V and V^* be two vector spaces placed in the duality by a bilinear pairing denoted by $\langle\cdot,\cdot\rangle$ and $\phi : V \to \mathbb{R}$ be a convex function. The *conjugate function (or Legendre transform)* $\phi^* : V^* \to \mathbb{R}$ is defined by

$$\phi^*(u^*) = \sup_{u\in V}\{\langle u, u^*\rangle - \phi(u)\}.$$

Proposition 3.33 *Let* $\zeta, u \in \mathcal{H}$. *If the assumption* (A_φ) *is fulfilled, then* $\zeta \in \partial\mathcal{J}_\varphi(u)$ *if and only if there exists* $M = (M_0, \bar{M}) \in \mathcal{H}\times\mathcal{H}^d$, $\frac{|\bar{M}(x)|}{2\alpha} \leq c_1 \in [0,+\infty)$ *for all* $x \in \Omega$ *such that*

$$\langle \bar{M}(x), (\nabla u)(x)\rangle_{\mathbb{R}^d} + 2\alpha\varphi(|(\nabla u)(x)|) + 2\alpha\varphi_1^*\left(\frac{|\bar{M}(x)|}{2\alpha}\right) = 0 \quad \textit{for all } x\in\Omega \tag{3.176}$$

$$K^*M_0 - \operatorname{div}\bar{M} + \zeta = 0 \tag{3.177}$$

$$-M_0 = 2(Ku-g), \tag{3.178}$$

where φ_1^* *is the conjugate function of* φ_1 *defined by* $\varphi_1(s) = \varphi(|s|)$, *for* $s\in\mathbb{R}$.

If additionally φ *is differentiable and* $|(\nabla u)(x)| \neq 0$ *for* $x\in\Omega$, *then we can compute* $\bar{M}$ *as*

$$\bar{M}(x) = -2\alpha\frac{\varphi'(|(\nabla u)(x)|)}{|(\nabla u)(x)|}(\nabla u)(x). \tag{3.179}$$

Remark 3.34 (i) For $\varphi(s) = s$ the function φ_1 from Proposition 3.33 turns out to be $\varphi_1(s) = |s|$. Its conjugate function φ_1^* is then given by

$$\varphi_1^*(s^*) = \sup_{s\in\mathbb{R}}\{\langle s^*, s\rangle - |s|\} = \begin{cases} 0 & \text{for } |s^*|\leq 1 \\ \infty & \text{else} \end{cases}.$$

Hence condition (3.176) specifies as follows

$$\langle \bar{M}(x), (\nabla u)(x)\rangle_{\mathbb{R}^d} + 2\alpha|(\nabla u)(x)| = 0$$

and, directly from the proof of Proposition 3.33 in the Appendix, $|\bar{M}(x)| \leq 2\alpha$ for all $x\in\Omega$.

(ii) We want to highlight a few important differences with respect to the continuous case. Due to our definition of the gradient and its relationship with the divergence operator $-\operatorname{div} = \nabla^*$ no boundary conditions are needed. Therefore condition (10) of [77, Proposition 4.1] has no discrete correspondent in our setting. The continuous total variation of a function can be decomposed into an absolute continuous part with respect to the Lebesgue measure and a singular part, whereas no singular part appears in the discrete setting. Therefore condition (6) and (7) of [77, Proposition 4.1] have not a discrete correspondent either.

(iii) An interesting consequence of Proposition 3.33 is that the map $S_\alpha = (I - P_{\alpha K})$ is bounded, i.e., $\|S_\alpha(z^k)\|_2 \to \infty$ if and only if $\|z^k\|_2 \to \infty$, for $k \to \infty$. In fact, since

$$S_\alpha(z) \quad = \quad \arg\min_{u \in \mathcal{H}} \|u - z\|_2^2 + 2\alpha|\nabla u|(\Omega),$$

from (3.177) and (3.178), we immediately obtain

$$S_\alpha(z) = z - \frac{1}{2}\operatorname{div}\bar{M},$$

and $\bar{M}$ is uniformly bounded.

Proof. (Proposition 3.33.) It is clear that $\zeta \in \partial\mathcal{J}_\varphi(u)$ if and only if $u = \operatorname{argmin}_{v\in\mathcal{H}}\{\mathcal{J}_\varphi(v) - \langle \zeta, v\rangle_{\mathcal{H}}\}$, and let us consider the following variational problem:

$$\inf_{v\in\mathcal{H}}\{\mathcal{J}_\varphi(v) - \langle\zeta, v\rangle_{\mathcal{H}}\} = \inf_{v\in\mathcal{H}}\{\|Kv - g\|_2^2 + 2\alpha\varphi(|\nabla v|)(\Omega) - \langle\zeta, v\rangle_{\mathcal{H}}\} \qquad (\mathcal{P})$$

We denote such an infimum by $\inf(\mathcal{P})$. Now we compute ($\mathcal{P}^*$) the dual of ($\mathcal{P}$). Let $\mathcal{F} : \mathcal{H} \to \mathbb{R}$, $\mathcal{G} : \mathcal{H} \times \mathcal{H}^d \to \mathbb{R}$, $\mathcal{G}_1 : \mathcal{H} \to \mathbb{R}$, $\mathcal{G}_2 : \mathcal{H}^d \to \mathbb{R}$, such that

$$\begin{aligned} \mathcal{F}(v) &= -\langle\zeta, v\rangle_{\mathcal{H}} \\ \mathcal{G}_1(w_0) &= \|w_0 - g\|_2^2 \\ \mathcal{G}_2(\bar{w}) &= 2\alpha\varphi(|\bar{w}|)(\Omega) \\ \mathcal{G}(w) &= \mathcal{G}_1(w_0) + \mathcal{G}_2(\bar{w}) \end{aligned}$$

with $w = (w_0, \bar{w}) \in \mathcal{H} \times \mathcal{H}^d$. Then the dual problem of ($\mathcal{P}$) is given by (cf. [36, p 60])

$$\sup_{p^* \in \mathcal{H}\times\mathcal{H}^d}\{-\mathcal{F}^*(\mathscr{M}^* p^*) - \mathcal{G}^*(-p^*)\} \qquad (\mathcal{P}^*)$$

where $\mathscr{M} : \mathcal{H} \to \mathcal{H} \times \mathcal{H}^d$ is defined by

$$\mathscr{M}v = (Kv, (\nabla v)^1, \ldots, (\nabla v)^d)$$

and $\mathscr{M}^*$ is its adjoint. We denote the supremum in ($\mathcal{P}^*$) by $\sup(\mathcal{P}^*)$. Using the definition of the conjugate function we compute $\mathcal{F}^*$ and $\mathcal{G}^*$. In particular

$$\mathcal{F}^*(\mathscr{M}^*p^*) = \sup_{v\in\mathcal{H}}\{\langle\mathscr{M}^*p^*, v\rangle_{\mathcal{H}} - \mathcal{F}(v)\} = \sup_{v\in\mathcal{H}}\langle\mathscr{M}^*p^* + \zeta, v\rangle_{\mathcal{H}} = \begin{cases} 0 & \mathscr{M}^*p^* + \zeta = 0 \\ \infty & \text{otherwise} \end{cases}$$

where $p^* = (p_0^*, \bar{p}^*)$ and

$$\begin{aligned} \mathcal{G}^*(p^*) &= \sup_{w\in\mathcal{H}\times\mathcal{H}^d}\{\langle p^*, w\rangle_{\mathcal{H}\times\mathcal{H}^d} - \mathcal{G}(w)\} \\ &= \sup_{w=(w_0,\bar{w})\in\mathcal{H}\times\mathcal{H}^d}\{\langle p_0^*, w_0\rangle_{\mathcal{H}} + \langle\bar{p}^*, \bar{w}\rangle_{\mathcal{H}^d} - \mathcal{G}_1(w_0) - \mathcal{G}_2(\bar{w})\} \\ &= \sup_{w_0\in\mathcal{H}}\{\langle p_0^*, w_0\rangle_{\mathcal{H}} - \mathcal{G}_1(w_0)\} + \sup_{\bar{w}\in\mathcal{H}^d}\{\langle\bar{p}^*, \bar{w}\rangle_{\mathcal{H}^d} - \mathcal{G}_2(\bar{w})\} \\ &= \mathcal{G}_1^*(p_0^*) + \mathcal{G}_2^*(\bar{p}^*) \end{aligned}$$

We have that

$$\mathcal{G}_1^*(p_0^*) = \left\langle \frac{p_0^*}{4} + g, p_0^* \right\rangle_{\mathcal{H}}$$

and (see [36])

$$\mathcal{G}_2^*(\bar{p}^*) = 2\alpha\varphi_1^* \left(\frac{|\bar{p}^*|}{2\alpha} \right)(\Omega)$$

if $\frac{|\bar{p}^*(x)|}{2\alpha} \in \operatorname{Dom}\varphi_1^*$, where φ_1^* is the conjugate function of φ_1 defined by

$$\varphi_1(s) := \varphi(|s|) \;\; s \in \mathbb{R}.$$

For ease we include in below the explicit computation of these conjugate functions. So we can write $(\mathcal{P}^*)$ in the following way

$$\sup_{p^* \in \mathcal{K}} \left\{ - \left\langle \frac{-p_0^*}{4} + g, -p_0^* \right\rangle_{\mathcal{H}} - 2\alpha\varphi_1^* \left(\frac{|\bar{p}^*|}{2\alpha} \right)(\Omega) \right\} \tag{4.180}$$

where

$$\mathcal{K} = \left\{ p^* \in \mathcal{H} \times \mathcal{H}^d : \frac{|\bar{p}^*(x)|}{2\alpha} \in \operatorname{Dom}\varphi_1^* \text{ for all } x \in \Omega, \mathscr{M}^* p^* + \zeta = 0 \right\}.$$

The function φ_1 also fulfills assumption (A_φ)(ii) (i.e., there exists $c_1 > 0, b \geq 0$ such that $c_1 z - b \leq \varphi_1(z) \leq c_1 z + b$, for all $z \in \mathbb{R}^+$). The conjugate function of φ_1 is given by $\varphi_1^*(s) = \sup_{z\in\mathbb{R}}\{\langle s, z\rangle - \varphi_1(z)\}$. Using the previous inequalities and that φ_1 is even (i.e., $\varphi_1(z) = \varphi_1(-z)$ for all $z \in \mathbb{R}$) we have

$$(\sup_{z\in\mathbb{R}}\{\langle s,z\rangle - c_1|z| + b\} \geq) \sup_{z\in\mathbb{R}}\{\langle s,z\rangle - \varphi_1(z)\} \geq \sup_{z\in\mathbb{R}}\{\langle s,z\rangle - c_1|z| - b\} = \begin{cases} -b & \text{if } |s| \leq c_1 \\ \infty & \text{else} \end{cases}. \tag{4.181}$$

In particular, one can see that $s \in \operatorname{Dom}\varphi_1^*$ if and only if $|s| \leq c_1$.

From $\mathscr{M}^* p^* + \zeta = 0$ we obtain

$$\langle \mathscr{M}^* p^*, \omega\rangle_{\mathcal{H}} + \langle \zeta, \omega\rangle_{\mathcal{H}} = \langle p^*, \mathscr{M}\omega\rangle_{\mathcal{H}^{d+1}} + \langle \zeta, \omega\rangle_{\mathcal{H}} = \langle p_0^*, K\omega\rangle_{\mathcal{H}} + \langle \bar{p}^*, \nabla\omega\rangle_{\mathcal{H}^d} + \langle \zeta, \omega\rangle_{\mathcal{H}} = 0$$

for all $\omega \in \mathcal{H}$. Then, since $\langle \bar{p}^*, \nabla\omega\rangle_{\mathcal{H}^d} = \langle -\operatorname{div}\bar{p}^*, \omega\rangle_{\mathcal{H}}$ (see Section 2.1.3), we have

$$K^* p_0^* - \operatorname{div}\bar{p}^* + \zeta = 0.$$

Hence we can write $\mathcal{K}$ in the following way

$$\mathcal{K} = \left\{ p^* = (p_0^*, \bar{p}^*) \in \mathcal{H} \times \mathcal{H}^d : \frac{|\bar{p}^*(x)|}{2\alpha} \leq c_1 \text{ for all } x \in \Omega, K^* p_0^* - \operatorname{div}\bar{p}^* + \zeta = 0 \right\}.$$

We now apply the duality results from [36, Theorem III.4.1], since the functional in $(\mathcal{P})$ is convex, continuous with respect to $\mathscr{M}v$ in $\mathcal{H} \times \mathcal{H}^d$, and $\inf(\mathcal{P})$ is finite. Then $\inf(\mathcal{P}) = \sup(\mathcal{P}^*) \in \mathbb{R}$ and $(\mathcal{P}^*)$ has a solution $M = (M_0, \bar{M}) \in \mathcal{K}$.

Let us assume that u is a solution of ($\mathcal{P}$) and M is a solution of ($\mathcal{P}^*$). From $\inf(\mathcal{P})=\sup(\mathcal{P}^*)$ we get

$$\|Ku-g\|_2^2+2\alpha\varphi(|\nabla u|)(\Omega)-\langle\zeta,u\rangle_{\mathcal{H}} = -\left\langle\frac{-M_0}{4}+g,-M_0\right\rangle_{\mathcal{H}}-2\alpha\varphi_1^*\left(\frac{|\bar{M}|}{2\alpha}\right)(\Omega) \tag{4.182}$$

where $M=(M_0,\bar{M})\in\mathcal{H}\times\mathcal{H}^d$, $\frac{|\bar{M}(x)|}{2\alpha}\leq c_1$ and $K^*M_0-\operatorname{div}\bar{M}+\zeta=0$, which verifies the direct implication of (3.177). In particular

$$-\langle\zeta,u\rangle_{\mathcal{H}}=\langle K^*M_0,u\rangle_{\mathcal{H}}-\langle\operatorname{div}\bar{M},u\rangle_{\mathcal{H}}=\langle M_0,Ku\rangle_{\mathcal{H}}+\langle\bar{M},\nabla u\rangle_{\mathcal{H}^d},$$

and

$$\|Ku-g\|_2^2+\langle M_0,Ku\rangle_{\mathcal{H}}+\langle\bar{M},\nabla u\rangle_{\mathcal{H}^d}+2\alpha\varphi(|\nabla u|)(\Omega)+\left\langle\frac{-M_0}{4}+g,-M_0\right\rangle_{\mathcal{H}}+2\alpha\varphi_1^*\left(\frac{|\bar{M}|}{2\alpha}\right)(\Omega)= \tag{4.183}$$

Let us write (4.183) again in the following form

$$\sum_{x\in\Omega}|(Ku-g)(x)|^2+\sum_{x\in\Omega}M_0(x)(Ku)(x)+\sum_{x\in\Omega}\sum_{j=1}^{d}\bar{M}^j(x)(\nabla u)^j(x)+\sum_{x\in\Omega}2\alpha\varphi(|(\nabla u)(x)|)$$
$$+\sum_{x\in\Omega}\left(\frac{-M_0(x)}{4}+g(x)\right)(-M_0(x))+\sum_{x\in\Omega}2\alpha\varphi_1^*\left(\frac{|\bar{M}(x)|}{2\alpha}\right)=0. \tag{4.184}$$

Now we have

1. $2\alpha\varphi(|(\nabla u)(x)|)+\sum_{j=1}^d\bar{M}^j(x)(\nabla u)^j(x)+2\alpha\varphi_1^*\left(\frac{|\bar{M}(x)|}{2\alpha}\right)\geq 2\alpha\varphi(|(\nabla u)(x)|)-\sum_{j=1}^d|\bar{M}^j(x)||(\nabla u)^j(x)|+2\alpha\varphi_1^*\left(\frac{|\bar{M}(x)|}{2\alpha}\right)\geq 0$ by the definition of φ_1^*, since $2\alpha\varphi_1^*\left(\frac{|\bar{M}(x)|}{2\alpha}\right)=\sup_{S\in\mathbb{R}^d}\{\langle\bar{M}^j(x),S\rangle_{\mathbb{R}^d}-2\alpha\varphi(|S|)\}=\sup_{S\in\mathbb{R}^d}\{\langle|\bar{M}^j(x)|,|S|\rangle_{\mathbb{R}^d}-2\alpha\varphi(|S|)\}$.
2. $|(Ku-g)(x)|^2+M_0(x)(Ku)(x)+\left(\frac{-M_0(x)}{4}+g(x))(-M_0(x)\right)=(((Ku)(x)-g(x)))^2+M_0(x)((Ku)(x)-g(x))+\left(\frac{M_0(x)}{2}\right)^2=\left(((Ku)(x)-g(x))+\frac{M_0(x)}{2}\right)^2\geq 0.$

Hence condition (4.183) reduces to

$$2\alpha\varphi(|(\nabla u)(x)|)+\sum_{j=1}^{d}\bar{M}^j(x)(\nabla u)^j(x)+2\alpha\varphi_1^*\left(\frac{|\bar{M}(x)|}{2\alpha}\right)=0\quad\text{for all } x\in\Omega \tag{4.185}$$

$$-M_0(x)=2((Ku)(x)-g(x))\quad\text{for all } x\in\Omega. \tag{4.186}$$

Conversely, if such an $M = (M_0, \bar{M}) \in \mathcal{H} \times \mathcal{H}^d$ with $\frac{|\bar{M}(x)|}{2\alpha} \leq c_1$ exists which fulfills conditions (3.176)-(3.178), it is clear from previous considerations that equation (4.182) holds. Let us denote the functional on the left side of (4.182) by

$$P(u) := \|Ku - g\|_2^2 + 2\alpha\varphi(|\nabla u|)(\Omega) - \langle \zeta, u \rangle_{\mathcal{H}}$$

and the functional on the right side of (4.182) by

$$P^*(M) := -\left\langle \frac{-M_0}{4} + g, -M_0 \right\rangle_{\mathcal{H}} - 2\alpha\varphi_1^* \left(\frac{|\bar{M}|}{2\alpha} \right)(\Omega).$$

We know that the functional P is the functional of $(\mathcal{P})$ and P^* is the functional of $(\mathcal{P}^*)$. Hence $\inf P = \inf(\mathcal{P})$ and $\sup P^* = \sup(\mathcal{P}^*)$. Since P is convex, continuous with respect to $\mathscr{M}u$ in $\mathcal{H} \times \mathcal{H}^d$, and $\inf(\mathcal{P})$ is finite we know from duality results [36, Theorem III.4.1] that $\inf(\mathcal{P}) = \sup(\mathcal{P}^*) \in \mathbb{R}$. We assume that M is no solution of $(\mathcal{P}^*)$, i.e., $P^*(M) < \sup(\mathcal{P}^*)$, and u is no solution of $(\mathcal{P})$, i.e, $P(u) > \inf(\mathcal{P})$. Then we have that

$$P(u) > \inf(\mathcal{P}) = \sup(\mathcal{P}^*) > P^*(M).$$

Thus (4.182) is valid if and only if M is a solution of $(\mathcal{P}^*)$ and u is a solution of $(\mathcal{P})$ which amounts to saying that $\zeta \in \partial \mathcal{J}_\varphi(u)$.

If additionally φ is differentiable and $|(\nabla u)(x)| \neq 0$ for $x \in \Omega$, we show that we can compute $\bar{M}(x)$ explicitly. From equation (3.176) (resp. (4.185)) we have

$$2\alpha\varphi_1^* \left(\frac{|-\bar{M}(x)|}{2\alpha} \right) = -\langle \bar{M}(x), (\nabla u)(x) \rangle_{\mathbb{R}^d} - 2\alpha\varphi(|(\nabla u)(x)|). \tag{4.187}$$

From the definition of conjugate function we have

$$\begin{aligned}
2\alpha\varphi_1^* \left(\frac{|-\bar{M}(x)|}{2\alpha} \right) &= 2\alpha \sup_{t \in \mathbb{R}} \left\{ \left\langle \frac{|-\bar{M}(x)|}{2\alpha}, t \right\rangle - \varphi_1(t) \right\} \\
&= 2\alpha \sup_{t \geq 0} \left\{ \left\langle \frac{|-\bar{M}(x)|}{2\alpha}, t \right\rangle - \varphi_1(t) \right\} \\
&= 2\alpha \sup_{t \geq 0} \sup_{\substack{S \in \mathbb{R}^d \\ |S| = t}} \left\{ \left\langle \frac{-\bar{M}(x)}{2\alpha}, S \right\rangle_{\mathbb{R}^d} - \varphi_1(|S|) \right\} \\
&= \sup_{S \in \mathbb{R}^d} \left\{ \left\langle -\bar{M}(x), S \right\rangle_{\mathbb{R}^d} - 2\alpha\varphi(|S|)(\Omega) \right\}.
\end{aligned} \tag{4.188}$$

Now, if $|(\nabla u)(x)| \neq 0$ for $x \in \Omega$, then it follows from (4.187) that the supremum is taken on in $S = |(\nabla u)(x)|$ and we have

$$\nabla_S(-\langle \bar{M}(x), S \rangle_{\mathbb{R}^d} - 2\alpha\varphi(|S|)(\Omega)) = 0$$

which implies

$$\bar{M}^j(x) = -2\alpha\frac{\varphi'(|(\nabla u)(x)|)}{|(\nabla u)(x)|}(\nabla u)^j(x) \quad j = 1, \ldots, d,$$

and verifies (3.179). This finishes the proof. □

Computation of conjugate functions. Let us compute the conjugate function of the convex function $\mathcal{G}_1(w_0) = \|w_0 - g\|_2^2$. From Definition 3.32 we have

$$\mathcal{G}_1^*(p_0^*) = \sup_{w_0 \in \mathcal{H}} \{\langle w_0, p_0^* \rangle_{\mathcal{H}} - \mathcal{G}_1(w_0)\} = \sup_{w_0 \in \mathcal{H}} \{\langle w_0, p_0^* \rangle_{\mathcal{H}} - \langle w_0 - g, w_0 - g \rangle_{\mathcal{H}}\}.$$

We set $H(w_0) := \langle w_0, p_0^* \rangle_{\mathcal{H}} - \langle w_0 - g, w_0 - g \rangle_{\mathcal{H}}$. To get the maximum of H we compute the Gâteaux-differential at w_0 of H,

$$H'(w_0) = p_0^* - 2(w_0 - g) = 0$$

and we set it to zero $H'(w_0) = 0$, since $H''(w_0) < 0$, and we get $w_0 = \frac{p_0}{2} + g$. Thus we have that

$$\sup_{w_0 \in \mathcal{H}} H(w_0) = \left\langle \frac{p_0^*}{4} + g, p_0^* \right\rangle_{\mathcal{H}} = \mathcal{G}_1^*(p_0^*)$$

Now we are going to compute the conjugate function of $\mathcal{G}_2(\bar{w}) = 2\alpha\varphi(|\bar{w}|)(\Omega)$. Associated to our notations we define the space $\mathcal{H}_0^+ = \mathbb{R}_0^{+N_1 \times \ldots \times N_d}$. From Definition 3.32 we have

$$\begin{aligned}
\mathcal{G}_2^*(\bar{p}^*) &= \sup_{\bar{w} \in \mathcal{H}^d} \{\langle \bar{w}, \bar{p}^* \rangle_{\mathcal{H}^d} - 2\alpha\varphi(|\bar{w}|)(\Omega)\} \\
&= \sup_{t \in \mathcal{H}_0^+} \sup_{\substack{\bar{w} \in \mathcal{H}^d \\ |\bar{w}(x)| = t(x)}} \{\langle \bar{w}, \bar{p}^* \rangle_{\mathcal{H}^d} - 2\alpha\varphi(|\bar{w}|)(\Omega)\} \\
&= \sup_{t \in \mathcal{H}_0^+} \{\langle t, |\bar{p}^*| \rangle_{\mathcal{H}} - 2\alpha\varphi(t)(\Omega)\}.
\end{aligned}$$

If φ were an even function then

$$\begin{aligned}
\sup_{t \in \mathcal{H}_0^+} \{\langle t, |\bar{p}^*| \rangle_{\mathcal{H}} - 2\alpha\varphi(t)(\Omega)\} &= \sup_{t \in \mathcal{H}} \{\langle t, |\bar{p}^*| \rangle_{\mathcal{H}} - 2\alpha\varphi(t)(\Omega)\} \\
&= 2\alpha \sup_{t \in \mathcal{H}} \left\{ \left\langle t, \frac{|\bar{p}^*|}{2\alpha} \right\rangle_{\mathcal{H}} - \varphi(t)(\Omega) \right\} \\
&= 2\alpha\varphi^* \left(\frac{|\bar{p}^*|}{2\alpha} \right)(\Omega)
\end{aligned}$$

where φ^* is the conjugate function of φ.

Unfortunately φ is not even in general. To overcome this difficulty we have to choose a function which is equal to $\varphi(s)$ for $s \geq 0$ and does not change the supremum for $s < 0$. For instance, one can choose $\varphi_1(s) = \varphi(|s|)$ for $s \in \mathbb{R}$. Then we have

$$\begin{aligned}\sup_{t \in \mathcal{H}_0^+} \{\langle t, |\bar{p}^*| \rangle_{\mathcal{H}} - 2\alpha\varphi(t)(\Omega)\} &= \sup_{t \in \mathcal{H}}\{\langle t, |\bar{p}^*| \rangle_{\mathcal{H}} - 2\alpha\varphi_1(t)(\Omega)\} \\ &= 2\alpha \sup_{t \in \mathcal{H}} \left\{ \left\langle t, \frac{|\bar{p}^*|}{2\alpha} \right\rangle_{\mathcal{H}} - \varphi_1(t)(\Omega) \right\} \\ &= 2\alpha\varphi_1^* \left(\frac{|\bar{p}^*|}{2\alpha} \right)(\Omega)\end{aligned}$$

where φ_1^* is the conjugate function of φ_1. Note that one can also choose $\varphi_1(s) = \varphi(s)$ for $s \geq 0$ and $\varphi_1(s) = \infty$ for $s < 0$.

Convergence properties

We return to the sequential algorithm (3.162). Let us explicitly express the algorithm as follows:

Algorithm 7. Pick an initial $V_1 + V_2 \ni \tilde{u}_1^{(0)} + \tilde{u}_2^{(0)} := u^{(0)} \in \mathcal{H}$, for example, $\tilde{u}_i^{(0)} = 0, i = 1, 2$, and iterate

$$\begin{cases} \begin{cases} u_1^{(n+1,0)} = \tilde{u}_1^{(n)} \\ u_1^{(n+1,\ell+1)} = \operatorname{argmin}_{\substack{u_1 \in V_1 \\ \mathrm{Tr}|_{\Gamma_1} u_1 = 0}} \mathcal{J}_1^s(u_1 + \tilde{u}_2^{(n)}, u_1^{(n+1,\ell)}) \\ \ell = 0, \ldots, L-1 \end{cases} \\ \begin{cases} u_2^{(n+1,0)} = \tilde{u}_2^{(n)} \\ u_2^{(n+1,m+1)} = \operatorname{argmin}_{\substack{u_2 \in V_2 \\ \mathrm{Tr}|_{\Gamma_2} u_2 = 0}} \mathcal{J}_2^s(u_1^{(n+1,L)} + u_2, u_2^{(n+1,m)}) \\ m = 0, \ldots, M-1 \end{cases} \\ u^{(n+1)} := u_1^{(n+1,L)} + u_2^{(n+1,M)} \\ \tilde{u}_1^{(n+1)} := \chi_1 \cdot u^{(n+1)} \\ \tilde{u}_2^{(n+1)} := \chi_2 \cdot u^{(n+1)}. \end{cases} \tag{3.189}$$

Note that we do prescribe a finite number L and M of inner iterations for each subspace respectively and that $u^{(n+1)} = \tilde{u}_1^{(n+1)} + \tilde{u}_2^{(n+1)}$, with $u_i^{(n+1)} \neq \tilde{u}_i^{(n+1)}$, $i = 1, 2$, in general. In this section we want to prove its convergence for any choice of L and M.

Observe that, for $a \in V_i$ and $\|K\| < 1$,

$$\|u_i - a\|_2^2 - \|Ku_i - Ka\|_2^2 \geq C\|u_i - a\|_2^2, \tag{3.190}$$

for $C = (1 - \|K\|^2) > 0$. Hence

$$\mathcal{J}(u) = \mathcal{J}_i^s(u, u_i) \leq \mathcal{J}_i^s(u, a), \tag{3.191}$$

and

$$\mathcal{J}_i^s(u, a) - \mathcal{J}_i^s(u, u_i) \geq C\|u_i - a\|_2^2. \tag{3.192}$$

Proposition 3.35 (Convergence properties) *Let us assume that $\|K\| < 1$. The algorithm in (3.210) produces a sequence $(u^{(n)})_{n\in\mathbb{N}}$ in $\mathcal{H}$ with the following properties:*

(i) $\mathcal{J}(u^{(n)}) > \mathcal{J}(u^{(n+1)})$ *for all $n \in \mathbb{N}$ (unless $u^{(n)} = u^{(n+1)}$);*

(ii) $\lim_{n\to\infty} \|u^{(n+1)} - u^{(n)}\|_2 = 0$*;*

(iii) the sequence $(u^{(n)})_{n\in\mathbb{N}}$ has subsequences which converge in $\mathcal{H}$.

Proof. Let us first observe that

$$\mathcal{J}(u^{(n)}) = \mathcal{J}_1^s(\tilde{u}_1^{(n)} + \tilde{u}_2^{(n)}, \tilde{u}_1^{(n)}) \quad = \quad \mathcal{J}_1^s(\tilde{u}_1^{(n)} + \tilde{u}_2^{(n)}, u_1^{(n+1,0)}).$$

By definition of $u_1^{(n+1,1)}$ and the minimal properties of $u_1^{(n+1,1)}$ in (3.210) we have

$$\mathcal{J}_1^s(\tilde{u}_1^{(n)} + \tilde{u}_2^{(n)}, u_1^{(n+1,0)}) \geq \mathcal{J}_1^s(u_1^{(n+1,1)} + \tilde{u}_2^{(n)}, u_1^{(n+1,0)}).$$

From (3.191) we have

$$\mathcal{J}_1^s(u_1^{(n+1,1)} + \tilde{u}_2^{(n)}, u_1^{(n+1,0)}) \geq \mathcal{J}_1^s(u_1^{(n+1,1)} + \tilde{u}_2^{(n)}, u_1^{(n+1,1)}) = \mathcal{J}(u_1^{(n+1,1)} + \tilde{u}_2^{(n)}).$$

Putting in line these inequalities we obtain

$$\mathcal{J}(u^{(n)}) \geq \mathcal{J}(u_1^{(n+1,1)} + \tilde{u}_2^{(n)}).$$

In particular, from (3.192) we have

$$\mathcal{J}(u^{(n)}) - \mathcal{J}(u_1^{(n+1,1)} + \tilde{u}_2^{(n)}) \geq C\|u_1^{(n+1,1)} - u_1^{(n+1,0)}\|_2^2.$$

After L steps we conclude the estimate

$$\mathcal{J}(u^{(n)}) \quad \geq \quad \mathcal{J}(u_1^{(n+1,L)} + \tilde{u}_2^{(n)}),$$

and

$$\mathcal{J}(u^{(n)}) - \mathcal{J}(u_1^{(n+1,L)} + \tilde{u}_2^{(n)}) \geq C \sum_{\ell=0}^{L-1} \|u_1^{(n+1,\ell+1)} - u_1^{(n+1,\ell)}\|_2^2.$$

By definition of $u_2^{(n+1,1)}$ and its minimal properties we have

$$\mathcal{J}(u_1^{(n+1,L)} + \tilde{u}_2^{(n)}) \geq \mathcal{J}_2^s(u_1^{(n+1,L)} + u_2^{(n+1,1)}, u_2^{(n+1,0)}).$$

By similar arguments as above we finally find the decreasing estimate

$$\mathcal{J}(u^{(n)}) \geq \mathcal{J}(u_1^{(n+1,L)} + u_2^{(n+1,M)}) = \mathcal{J}(u^{(n+1)}) = \mathcal{J}(\tilde{u}_1^{(n+1)} + \tilde{u}_2^{(n+1)}), \quad (3.193)$$

and

$$\mathcal{J}(u^{(n)}) - \mathcal{J}(u^{(n+1)})$$
$$\geq C \left(\sum_{\ell=0}^{L-1} \|u_1^{(n+1,\ell+1)} - u_1^{(n+1,\ell)}\|_2^2 + \sum_{m=0}^{M-1} \|u_2^{(n+1,m+1)} - u_2^{(n+1,m)}\|_2^2 \right), \quad (3.194)$$

which verifies (i).

From (3.193) we have $\mathcal{J}(u^{(0)}) \geq \mathcal{J}(u^{(n)})$. By the coerciveness condition (C) $(u^{(n)})_{n\in\mathbb{N}}$ is uniformly bounded in $\mathcal{H}$, hence there exists a convergent subsequence $(u^{(n_k)})_{k\in\mathbb{N}}$ and hence (iii) holds. Let us denote $u^{(\infty)}$ the limit of the subsequence. For simplicity, we rename such a subsequence by $(u^{(n)})_{n\in\mathbb{N}}$. Moreover, since the sequence $(\mathcal{J}(u^{(n)}))_{n\in\mathbb{N}}$ is monotonically decreasing and bounded from below by 0, it is also convergent. From (3.194) and the latter convergence we deduce

$$\left(\sum_{\ell=0}^{L-1} \|u_1^{(n+1,\ell+1)} - u_1^{(n+1,\ell)}\|_2^2 + \sum_{m=0}^{M-1} \|u_2^{(n+1,m+1)} - u_2^{(n+1,m)}\|_2^2 \right) \to 0, \quad n \to \infty. \quad (3.195)$$

In particular, by the standard inequality $(a^2 + b^2) \geq \frac{1}{2}(a+b)^2$ for $a, b > 0$ and the triangle inequality, we have also

$$\|u^{(n)} - u^{(n+1)}\|_2 \to 0, \quad n \to \infty. \quad (3.196)$$

This gives (ii) and completes the proof. □

The use of the partition of unity $\{\chi_1, \chi_2\}$ allows not only to guarantee the boundedness of $(u^{(n)})_{n\in\mathbb{N}}$, but also of the sequences $(\tilde{u}_1^{(n)})_{n\in\mathbb{N}}$ and $(\tilde{u}_2^{(n)})_{n\in\mathbb{N}}$.

Lemma 3.36 *The sequences $(\tilde{u}_1^{(n)})_{n\in\mathbb{N}}$ and $(\tilde{u}_2^{(n)})_{n\in\mathbb{N}}$ produced by the algorithm* (3.210) *are bounded, i.e., there exists a constant $\tilde{C} > 0$ such that $\|\tilde{u}_i^{(n)}\|_2 \leq \tilde{C}$ for $i = 1, 2$.*

Proof. From the boundedness of $(u^{(n)})_{n\in\mathbb{N}}$ we have

$$\|\tilde{u}_i^{(n)}\|_2 = \|\chi_i u^{(n)}\|_2 \leq \kappa \|u^{(n)}\|_2 \leq \tilde{C} \quad \text{for } i = 1, 2.$$

□

From Remark 3.34 (iii) we can also show the following auxiliary lemma.

Lemma 3.37 *The sequences $(\eta_1^{(n,L)})_n$ and $(\eta_2^{(n,M)})_n$ are bounded.*

Proof. From previous considerations we know that

$$u_1^{(n,L)} = S_\alpha(z_1^{(n,L-1)} + \tilde{u}_2^{(n-1)} - \eta_1^{(n,L)}) - \tilde{u}_2^{(n-1)}$$
$$u_2^{(n,M)} = S_\alpha(z_2^{(n,M-1)} + u_1^{(n,L)} - \eta_2^{(n,M)}) - u_1^{(n,L)}.$$

Assume $(\eta_1^{(n,L)})_n$ were unbounded, then by Remark 3.34 (iii), also $S_\alpha(z_1^{(n,L-1)} + \tilde{u}_2^{(n-1)} - \eta_1^{(n,L)})$ would be unbounded. Since $(\tilde{u}_2^{(n)})_n$ and $(u_1^{(n,L)})_n$ are bounded by Lemma 3.36 and formula (3.195), we have a contradiction. Thus $(\eta_1^{(n,L)})_n$ has to be bounded. With the same argument we can show that $(\eta_2^{(n,M)})_n$ is bounded. □

We can eventually show the convergence of the algorithm to minimizers of $\mathcal{J}$.

Theorem 3.38 (Convergence to minimizers) *Assume $\|K\| < 1$. Then accumulation points of the sequence $(u^{(n)})_{n\in\mathbb{N}}$ produced by algorithm (3.210) are minimizers of $\mathcal{J}$. If $\mathcal{J}$ has a unique minimizer then the sequence $(u^{(n)})_{n\in\mathbb{N}}$ converges to it.*

Proof. Let us denote $u^{(\infty)}$ the limit of a subsequence. For simplicity, we rename such a subsequence by $(u^{(n)})_{n\in\mathbb{N}}$. From Lemma 3.36 we know that $(\tilde{u}_1^{(n)})_{n\in\mathbb{N}}$, $(\tilde{u}_2^{(n)})_{n\in\mathbb{N}}$ and consequently $(u_1^{(n,L)})_{n\in\mathbb{N}}$,$(u_2^{(n,M)})_{n\in\mathbb{N}}$ are bounded. So the limit $u^{(\infty)}$ can be written as

$$u^{(\infty)} = u_1^{(\infty)} + u_2^{(\infty)} = \tilde{u}_1^{(\infty)} + \tilde{u}_2^{(\infty)} \tag{3.197}$$

where $u_1^{(\infty)}$ is the limit of $(u_1^{(n,L)})_{n\in\mathbb{N}}$, $u_2^{(\infty)}$ is the limit of $(u_2^{(n,M)})_{n\in\mathbb{N}}$, and $\tilde{u}_i^{(\infty)}$ is the limit of $(\tilde{u}_i^{(n)})_{n\in\mathbb{N}}$ for $i = 1, 2$. Now we show that $\tilde{u}_2^{(\infty)} = u_2^{(\infty)}$. By using the triangle inequality, from (3.195) it directly follows that

$$\|u_2^{(n+1,M)} - \tilde{u}_2^{(n)}\|_2 \to 0, \quad n \to \infty. \tag{3.198}$$

Moreover, since $\chi_2 \in V_2$ is a fixed vector which is independent of n, we obtain from Proposition 3.35 (ii) that

$$\|\chi_2(u^{(n)} - u^{(n+1)})\|_2 \to 0, \quad n \to \infty,$$

and hence

$$\|\tilde{u}_2^{(n)} - \tilde{u}_2^{(n+1)}\|_2 \to 0, \quad n \to \infty. \tag{3.199}$$

Putting (3.198) and (3.199) together and noting that

$$\|u_2^{(n+1,M)} - \tilde{u}_2^{(n)}\|_2 + \|\tilde{u}_2^{(n)} - \tilde{u}_2^{(n+1)}\|_2 \geq \|u_2^{(n+1,M)} - \tilde{u}_2^{(n+1)}\|_2$$

we have

$$\|u_2^{(n+1,M)} - \tilde{u}_2^{(n+1)}\|_2 \to 0, \quad n \to \infty, \tag{3.200}$$

which means that the sequences $(u_2^{(n,M)})_{n\in\mathbb{N}}$ and $(\tilde{u}_2^{(n)})_{n\in\mathbb{N}}$ have the same limit, i.e., $\tilde{u}_2^{(\infty)} = u_2^{(\infty)}$, which we denote by $u_2^{(\infty)}$. Then from (3.200) and (3.197) it directly follows that $\tilde{u}_1^{(\infty)} = u_1^{(\infty)}$.

As in the proof of the oblique thresholding theorem we set

$$F_1(u_1^{(n+1,L)}) := \|u_1^{(n+1,L)} - z_1^{(n+1,L)}\|_2^2 + 2\alpha|\nabla(u_1^{(n+1,L)} + \tilde{u}_2^{(n)}\Big|_{\Omega_1})|(\Omega_1)$$

where

$$z_1^{(n+1,L)} := u_1^{(n+1,L-1)} + (K^*(g - K\tilde{u}_2^{(n)} - Ku_1^{(n+1,L-1)}))\Big|_{\Omega_1}.$$

The optimality condition for $u_1^{(n+1,L)}$ is

$$0 \in \partial_{V_1} F_1(u_1^{(n+1,L)}) + 2\eta_1^{(n+1,L)}$$

where

$$\eta_1^{(n+1,L)} = (\mathrm{Tr}\,|_{\Gamma_1})^* \,\mathrm{Tr}\,|_{\Gamma_1} \left((z_1^{(n+1,L)}) + P_{\alpha K}(\eta_1^{(n+1,L)} - z_1^{(n+1,L)} - \tilde{u}_2^{(n)})\right).$$

In order to use the characterization of elements in the subdifferential of $|\nabla u|(\Omega)$, i.e., Proposition 3.33, we have to rewrite the minimization problem for F_1. More precisely, we define

$$\hat{F}_1(\xi_1^{(n+1,L)}) := \|\xi_1^{(n+1,L)} - \tilde{u}_2^{(n)}\Big|_{\Omega_1} - z_1^{(n+1,L)}\|_2^2 + 2\alpha|\nabla(\xi_1^{(n+1,L)})|(\Omega_1)$$

for $\xi_1^{(n+1,L)} \in V_1$ with $\mathrm{Tr}\,|_{\Gamma_1}\ \xi_1^{(n+1,L)} = \tilde{u}_2^{(n)}$. Then the optimality condition for $\xi_1^{(n+1,L)}$ is

$$0 \in \partial\hat{F}_1(\xi_1^{(n+1,L)}) + 2\eta_1^{(n+1,L)} \tag{3.201}$$

Note that indeed $\xi_1^{(n+1,L)}$ is optimal if and only if $u_1^{(n+1,L)} = \xi_1^{(n+1,L)} - \tilde{u}_2^{(n)}\Big|_{\Omega_1}$ is optimal.

Analogously we define

$$\hat{F}_2(\xi_2^{(n+1,M)}) := \|\xi_2^{(n+1,M)} - u_1^{(n+1,L)}\Big|_{\Omega_2} - z_2^{(n+1)}\|_2^2 + 2\alpha|\nabla(\xi_2^{(n+1,M)})|(\Omega_2)$$

for $\xi_2^{(n+1,M)} \in V_2$ with $\mathrm{Tr}\,|_{\Gamma_2}\ \xi_2^{(n+1,M)} = u_1^{(n+1,L)}$, and the optimality condition for $\xi_2^{(n+1,M)}$ is

$$0 \in \partial\hat{F}_2(\xi_2^{(n+1,M)}) + 2\eta_2^{(n+1,M)} \tag{3.202}$$

where

$$\eta_2^{(n+1,M)} = (\mathrm{Tr}\,|_{\Gamma_2})^* \,\mathrm{Tr}\,|_{\Gamma_2} \left((z_2^{(n+1,M)}) + P_{\alpha K}(\eta_2^{(n+1,M)} - z_2^{(n+1,M)} - u_1^{(n+1,L)})\right).$$

Let us recall that now we are considering functionals as in Proposition 3.33 with $\varphi(s) = s$, $K = I$, and $\Omega = \Omega_i$, $i = 1, 2$. From Proposition 3.33 and Remark 3.34 we get that $\xi_1^{(n+1,L)}$, and consequently $u_1^{(n+1,L)}$ is optimal, i.e., $-2\eta_1^{(n+1,L)} \in \partial \hat{F}_1(\xi_1^{(n+1,L)})$, if and only if there exists an $M_1^{(n+1)} = (M_{0,1}^{(n+1)}, \bar{M}_1^{(n+1)}) \in V_1 \times V_1^d$ with $|\bar{M}_1^{(n+1)}(x)| \leq 2\alpha$ for all $x \in \Omega_1$ such that

$$\langle \bar{M}_1^{(n+1)}(x), (\nabla(u_1^{(n+1,L)} + \tilde{u}_2^{(n)}))(x)\rangle_{\mathbb{R}^d} + 2\alpha\varphi(|(\nabla(u_1^{(n+1,L)} + \tilde{u}_2^{(n)}))(x)|) = 0 \tag{3.203}$$

$$-2(u_1^{(n+1,L)}(x) - z_1^{(n+1,L)}(x)) - \operatorname{div} \bar{M}_1^{(n+1)}(x) - 2\eta_1^{(n+1,L)}(x) = 0. \tag{3.204}$$

for all $x \in \Omega_1$. Analogously we get that $\xi_2^{(n+1,M)}$, and consequently $u_2^{(n+1,M)}$ is optimal, i.e., $-2\eta_2^{(n+1,M)} \in \partial \hat{F}_2(\xi_2^{(n+1,M)})$, if and only if there exists an $M_2^{(n+1)} = (M_{0,2}^{(n+1)}, \bar{M}_2^{(n+1)}) \in V_2 \times V_2^d$ with $|\bar{M}_2^{(n+1)}(x)| \leq 2\alpha$ for all $x \in \Omega_2$ such that

$$\langle \bar{M}_2^{(n+1)}(x), (\nabla(u_1^{(n+1,L)} + u_2^{(n+1,M)}))(x)\rangle_{\mathbb{R}^d} + 2\alpha\varphi(|(\nabla(u_1^{(n+1,L)} + \tilde{u}_2^{(n+1,M)}))(x)|) = 0 \tag{3.205}$$

$$-2(u_2^{(n+1,M)}(x) - z_2^{(n+1,M)}(x)) - \operatorname{div} \bar{M}_2^{(n+1)}(x) - 2\eta_2^{(n+1,M)}(x) = 0, \tag{3.206}$$

for all $x \in \Omega_2$. Since $(\bar{M}_1^{(n)}(x))_{n\in\mathcal{N}}$ is bounded for all $x \in \Omega_1$ and $(\bar{M}_2^{(n)}(x))_{n\in\mathcal{N}}$ is bounded for all $x \in \Omega_2$, there exist convergent subsequences $(\bar{M}_1^{(n_k)}(x))_{k\in\mathcal{N}}$ and $(\bar{M}_2^{(n_k)}(x))_{k\in\mathcal{N}}$. Let us denote $\bar{M}_1^{(\infty)}(x)$ and $\bar{M}_2^{(\infty)}(x)$ the respective limits of the sequences. For simplicity we rename such sequences by $(\bar{M}_1^{(n)}(x))_{n\in\mathcal{N}}$ and $(\bar{M}_2^{(n)}(x))_{n\in\mathcal{N}}$.

Note that, by Lemma 3.37 (or simply from (3.204) and (3.206)) the sequences $(\eta_1^{(n,L)})_{n\in\mathcal{N}}$ and $(\eta_2^{(n,M)})_{n\in\mathcal{N}}$ are also bounded. Hence there exist convergent subsequences which we denote, for simplicity, again by $(\eta_1^{(n,L)})_{n\in\mathcal{N}}$ and $(\eta_2^{(n,M)})_{n\in\mathcal{N}}$ with limits $\eta_i^{(\infty)}$, $i = 1, 2$. By taking in (3.203)-(3.206) the limits for $n \to \infty$ we obtain

$$\langle \bar{M}_1^{(\infty)}(x), (\nabla(u_1^{(\infty)} + u_2^{(\infty)}))(x)\rangle_{\mathbb{R}^d} + 2\alpha\varphi(|(\nabla(u_1^{(\infty)} + u_2^{(\infty)}))(x)|) = 0 \quad \text{for all } x \in \Omega_1$$
$$-2(u_1^{(\infty)}(x) - z_1^{(\infty)}(x)) - \operatorname{div} \bar{M}_1^{(\infty)}(x) - 2\eta_1^{(\infty)}(x) = 0 \quad \text{for all } x \in \Omega_1$$

$$\langle \bar{M}_2^{(\infty)}(x), (\nabla(u_1^{(\infty)} + u_2^{(\infty)}))(x)\rangle_{\mathbb{R}^d} + 2\alpha\varphi(|(\nabla(u_1^{(\infty)} + u_2^{(\infty)}))(x)|) = 0 \quad \text{for all } x \in \Omega_2$$
$$-2(u_2^{(\infty)}(x) - z_2^{(\infty)}(x)) - \operatorname{div} \bar{M}_2^{(\infty)}(x) - 2\eta_2^{(\infty)}(x) = 0 \quad \text{for all } x \in \Omega_2$$

Since $\operatorname{supp} \eta_1^{(\infty)} = \Gamma_1$ and $\operatorname{supp} \eta_2^{(\infty)} = \Gamma_2$ we have

$$\begin{aligned} \langle \bar{M}_1^{(\infty)}(x), (\nabla(u^{(\infty)})(x)\rangle_{\mathbb{R}^d} + 2\alpha\varphi(|(\nabla(u^{(\infty)})(x)|) = 0 \quad &\text{for all } x \in \Omega_1 \\ -2K^*((Ku^{(\infty)})(x) - g^{(\infty)}(x)) - \operatorname{div} \bar{M}_1^{(\infty)}(x) = 0 \quad &\text{for all } x \in \Omega_1 \setminus \Gamma_1 \end{aligned} \tag{3.207}$$

$$\begin{aligned} \langle \bar{M}_2^{(\infty)}(x), (\nabla(u^{(\infty)})(x)\rangle_{\mathbb{R}^d} + 2\alpha\varphi(|(\nabla(u^{(\infty)})(x)|) = 0 \quad &\text{for all } x \in \Omega_2 \\ -2K^*((Ku^{(\infty)})(x) - g^{(\infty)}(x)) - \operatorname{div} \bar{M}_2^{(\infty)}(x) = 0 \quad &\text{for all } x \in \Omega_2 \setminus \Gamma_2. \end{aligned} \tag{3.208}$$

Observe now that from Proposition 3.33 we also have that $0 \in \mathcal{J}(u^{(\infty)})$ if and only if there exists $M^{(\infty)} = (M_0^{(\infty)}, \bar{M}^{(\infty)})$ with $|\bar{M}_0^{(\infty)}(x)| \leq 2\alpha$ for all $x \in \Omega$ such that

$$\begin{aligned} \langle \bar{M}^{(\infty)}(x), (\nabla(u^{(\infty)})(x)\rangle_{\mathbb{R}^d} + 2\alpha\varphi(|(\nabla(u^{(\infty)})(x)|) = 0 \quad &\text{for all } x \in \Omega \\ -2K^*((Ku^{(\infty)})(x) - g^{(\infty)}(x)) - \operatorname{div} \bar{M}^{(\infty)}(x) = 0 \quad &\text{for all } x \in \Omega. \end{aligned} \tag{3.209}$$

Note that $\bar{M}_j^{(\infty)}(x), j = 1, 2$, for $x \in \Omega_1 \cap \Omega_2$ satisfies both (3.207) and (3.208). Hence let us choose

$$M^{(\infty)}(x) = \begin{cases} M_1^{(\infty)}(x) & \text{if } x \in \Omega_1 \setminus \Gamma_1 \\ M_2^{(\infty)}(x) & \text{if } x \in (\Omega_2 \setminus \Omega_1) \cup \Gamma_1 \end{cases}.$$

With this choice of $M^{(\infty)}$ equations (3.207) - (3.209) are valid and hence $u^{(\infty)}$ is optimal in Ω. □

Remark 3.39 (i) If $\nabla u^{(\infty)}(x) \neq 0$ for $x \in \Omega_j$, $j = 1, 2$, then $\bar{M}_j^{(\infty)}$ is given as in equation (3.179) by

$$\bar{M}_j^{(\infty)}(x) = -2\alpha \frac{(\nabla u^{(\infty)} \,|_{\Omega_j})(x)}{|(\nabla u^{(\infty)} \,|_{\Omega_j})(x)|}.$$

(ii) The boundedness of the sequences $(\tilde{u}_1^{(n)})_{n \in \mathbb{N}}$ and $(\tilde{u}_2^{(n)})_{n \in \mathbb{N}}$ has been technically used for showing the existence of an optimal decomposition $u^{(\infty)} = u_1^{(\infty)} + u_2^{(\infty)}$ in the proof of Theorem 3.38. Their boundedness is guaranteed as in Lemma 3.36 by the use of the partition of the unity $\{\chi_1, \chi_2\}$. Let us emphasize that there is no way of obtaining the boundedness of the local sequences $(u_1^{(n,L)})_{n \in \mathbb{N}}$ and $(u_2^{(n,M)})_{n \in \mathbb{N}}$ otherwise. In Figure 3.12 we show that the local sequences can become unbounded in case we do not modify them by means of the partition of the unity.

(iii) Note that for deriving the optimality condition (3.209) for $u^{(\infty)}$ we combined the respective conditions (3.207) and (3.208) for $u_1^{(\infty)}$ and $u_2^{(\infty)}$. In doing that, we strongly took advantage of the overlapping property of the subdomains, hence avoiding a fine analysis of $\eta_1^{(\infty)}$ and $\eta_2^{(\infty)}$ on the interfaces Γ_1 and Γ_2. This is the major advantage of this analysis with respect to the one provided in [44] for nonoverlapping domain decompositions.

3.4.4 A Parallel Algorithm

The parallel version of the previous algorithm (3.210) reads as follows:

Algorithm 9. Pick an initial $V_1 + V_2 \ni \tilde{u}_1^{(0)} + \tilde{u}_2^{(0)} := u^{(0)} \in \mathcal{H}$, for example, $\tilde{u}_i^{(0)} = 0, i = 1, 2$, and iterate

$$\begin{cases} \begin{cases} u_1^{(n+1,0)} = \tilde{u}_1^{(n)} \\ u_1^{(n+1,\ell+1)} = \operatorname{argmin}_{\substack{u_1 \in V_1 \\ \mathrm{Tr}|_{\Gamma_1} u_1 = 0}} \mathcal{J}_1^s(u_1 + \tilde{u}_2^{(n)}, u_1^{(n+1,\ell)}) \\ \ell = 0, \ldots, L-1 \end{cases} \\ \begin{cases} u_2^{(n+1,0)} = \tilde{u}_2^{(n)} \\ u_2^{(n+1,m+1)} = \operatorname{argmin}_{\substack{u_2 \in V_2 \\ \mathrm{Tr}|_{\Gamma_2} u_2 = 0}} \mathcal{J}_2^s(u_1^{(n+1,L)} + u_2, u_2^{(n+1,m)}) \\ m = 0, \ldots, M-1 \end{cases} \\ u^{(n+1)} := u_1^{(n+1,L)} + u_2^{(n+1,M)} \\ \tilde{u}_1^{(n+1)} := \chi_1 \cdot u^{(n+1)} \\ \tilde{u}_2^{(n+1)} := \chi_2 \cdot u^{(n+1)}. \end{cases} \tag{3.210}$$

As for ℓ_1-minimization also for this version the parallel algorithm is shown to converge in a similar way as its sequential counterpart.

3.4.5 Applications and Numerics

In this section we shall present the application of the sequential algorithm (3.162) for the minimization of $\mathcal{J}$ in one and two dimensions. In particular, we show how to implement the dual method of Chambolle [14] in order to compute the orthogonal projection $P_{\alpha K}(g)$ in the oblique thresholding, and we give a detailed explanation of the domain decompositions used in the numerics. Furthermore we present numerical examples for image *inpainting*, i.e., the recovery of missing parts of images by minimal total variation interpolation, and compressed sensing, in the nonadaptive compressed acquisition of images for a classical toy problem inspired by magnetic resonance imaging (MRI) [55]. The numerical examples of this section and respective Matlab codes can be found at [79].

Computation of $P_{\alpha K}(g)$

To solve the subiterations in (3.162) we compute the minimizer by means of oblique thresholding. More precisely, let us denote $u_2 = \tilde{u}_2^{(n)}$, $u_1 = u_1^{(n+1,\ell+1)}$, and $z_1 = u_1^{(n+1,\ell)} + K^*(g - Ku_2 - Ku_1^{(n+1,\ell)})$. We shall compute the minimizer u_1 of the first subminimization problem by

$$u_1 = (I - P_{\alpha K})(z_1 + u_2 - \eta) - u_2 \in V_1$$

for an $\eta \in V_1$ with $\operatorname{supp} \eta = \Gamma_1$ which fulfills

$$\mathrm{Tr}\,|_{\Gamma_1}(\eta) = \mathrm{Tr}\,|_{\Gamma_1}(z_1 + P_{\alpha K}(\eta - z_1 - u_2)).$$

Hence the element $\eta \in V_1$ is a limit of the corresponding fixed point iteration

$$\eta^{(0)} \in V_1, \operatorname{supp} \eta^{(0)} = \Gamma_1, \quad \eta^{(m+1)} = (\mathrm{Tr}\,|_{\Gamma_1})^* \,\mathrm{Tr}\,|_{\Gamma_1}\left(z_1 + P_{\alpha K}(\eta^{(m)} - z_1 - u_2)\right), \ m \geq 0. \tag{3.211}$$

Here K is defined as in Section 3.4, i.e.,

$$K = \left\{\operatorname{div} p : p \in \mathcal{H}^d, |p(x)|_\infty \leq 1 \quad \forall x \in \Omega\right\}.$$

To compute the projection onto αK in the oblique thresholding we use an algorithm proposed by Chambolle in [14]. His algorithm is based on considerations of the convex conjugate of the total variation and on exploiting the corresponding optimality condition. It amounts to compute $P_{\alpha K}(g)$ approximately by $\alpha \operatorname{div} p^{(n)}$, where $p^{(n)}$ is the nth iterate of the following semi-implicit gradient descent algorithm:

Choose $\tau > 0$, let $p^{(0)} = 0$ and, for any $n \geq 0$, iterate

$$p^{(n+1)}(x) = \frac{p^{(n)}(x) + \tau(\nabla(\operatorname{div} p^{(n)} - g/\alpha))(x)}{1 + \tau\left|(\nabla(\operatorname{div} p^{(n)} - g/\alpha))(x)\right|}.$$

For $\tau > 0$ sufficiently small, i.e., $\tau < 1/8$, the iteration $\alpha \operatorname{div} p^{(n)}$ was shown to converge to $P_{\alpha K}(g)$ as $n \to \infty$ (compare [14, Theorem 3.1]). Let us stress that we propose here this algorithm just for the ease of its presentation; its choice for the approximation of projections is of course by no means a restriction and one may want to implement other recent, and perhaps faster strategies, e.g., [15, 26, 48, 65, 78].

Domain decompositions

In one dimension the domain $\Omega = [a, b]$ is split into two overlapping intervals. Let $|\Omega_1 \cap \Omega_2| =: G$ be the size of the overlap of Ω_1 and Ω_2. Then we set $|\Omega_1| =: n_1 = \lceil \frac{N+G}{2} \rceil$, $\Omega_1 = [a, n_1]$ and $\Omega_2 = [n_1 - G + 1, b]$. The interfaces Γ_1 and Γ_2 are located in $i = n_1 + 1$ and $n_1 - G$ respectively (cf. Figure 3.8). The auxiliary functions χ_1 and χ_2 can be chosen in the following way (cf. Figure 3.7):

$$\chi_1(x_i) = \begin{cases} 1 & x_i \in \Omega_1 \setminus \Omega_2 \\ 1 - \frac{1}{G}(i - (n_1 - G + 1)) & x_i \in \Omega_1 \cap \Omega_2 \end{cases}$$

$$\chi_2(x_i) = \begin{cases} 1 & x_i \in \Omega_2 \setminus \Omega_1 \\ \frac{1}{G}(i - (n_1 - G + 1)) & x_i \in \Omega_1 \cap \Omega_2 \end{cases}.$$

Note that $\chi_1(x_i) + \chi_2(x_i) = 1$ for all $x_i \in \Omega$ (i.e for all $i = 1, \ldots, N$).

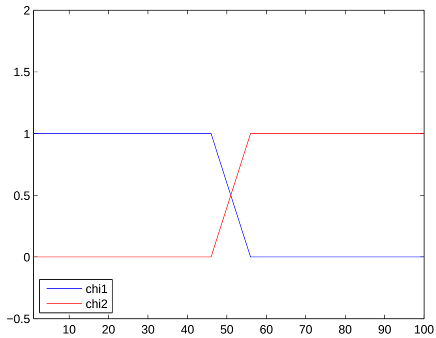


Figure 3.7 Auxiliary functions χ_1 and χ_2 for an overlapping domain decomposition with two subdomains.

In two dimensions the domain $\Omega = [a, b] \times [c, d]$ is split in an analogous way with respect to its rows. In particular we have $\Omega_1 = [a, n_1] \times [c, d]$ and $\Omega_2 = [n_1 - G + 1, b] \times [c, d]$, compare Figure 3.9. The splitting in more than two domains is done similarly:

Set $\Omega = \Omega_1 \cup \ldots \cup \Omega_{\mathcal{N}}$, the domain Ω decomposed into $\mathcal{N}$ domains Ω_i, $i = 1, \ldots, \mathcal{N}$, where Ω_i and Ω_{i+1} are overlapping for $i = 1, \ldots, \mathcal{N} - 1$. Let $|\Omega_i \cap \Omega_{i+1}| =: G$ equidistant for every $i = 1, \ldots, \mathcal{N} - 1$. Set $s = \lceil N_1/\mathcal{N} \rceil$. Then

$$\Omega_1 = [1, s + \frac{G}{2}] \times [c, d]$$

for $i = 2 : \mathcal{N} - 1$

$$\Omega_i = [(i-1)s - \frac{G}{2} + 1, is + \frac{G}{2}] \times [c, d]$$

end

$$\Omega_{\mathcal{N}} = [(\mathcal{N} - 1)s - \frac{G}{2} + 1, N_1] \times [c, d].$$

The auxiliary functions χ_i can be chosen in an analogous way as in the one dimensional case:

$$\chi_i(x_{i_1}, y_{i_2}) = \begin{cases} \frac{1}{G}(i_1 - ((i-1)s - G/2 + 1)) & (x_{i_1}, y_{i_2}) \in \Omega_{i-1} \cap \Omega_i \\ 1 & (x_{i_1}, y_{i_2}) \in \Omega_i \setminus (\Omega_{i-1} \cup \Omega_{i+1}) \\ 1 - \frac{1}{G}(i_1 - (is - G/2 + 1)) & (x_{i_1}, y_{i_2}) \in \Omega_i \cap \Omega_{i+1} \end{cases}$$

for $i = 1, \ldots, \mathcal{N}$ with $\Omega_0 = \Omega_{\mathcal{N}+1} = \emptyset$.

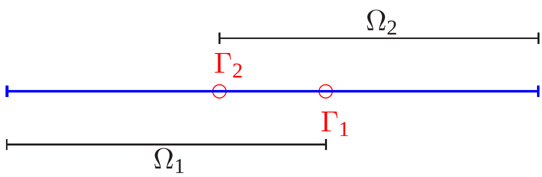


Figure 3.8 Overlapping domain decomposition in 1D.

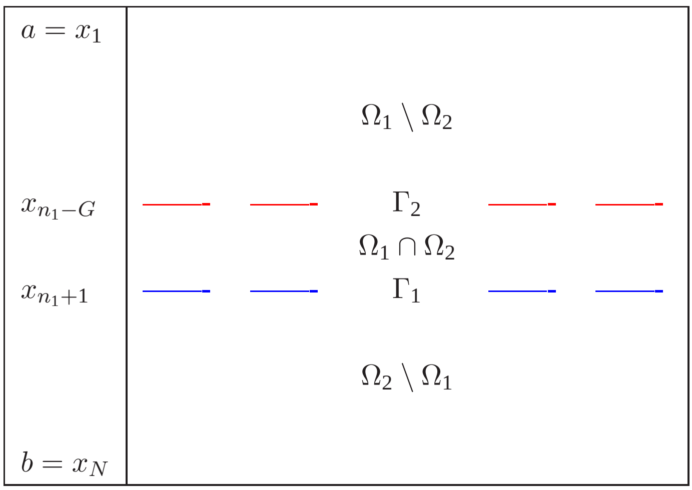


Figure 3.9 Decomposition of the image in two domains Ω_1 and Ω_2.

To compute the fixed point η of (3.172) in an efficient way we make the following considerations, which allow to restrict the computation from Ω_1 to a relatively small stripe around the interface. The fixed point η is actually supported on Γ_1 only, i.e., $\eta(x) = 0$ in $\Omega_1 \setminus \Gamma_1$. Hence, we restrict the fixed point iteration for η to a relatively small stripe $\hat{\Omega}_1 \subset \Omega_1$ Analogously, one implements the minimizations of η_2 on $\hat{\Omega}_2$.

Numerical experiments

In the following we present numerical examples for the sequential algorithm (3.210) in two particular applications: signal interpolation/image inpainting, and compressed sensing [79].

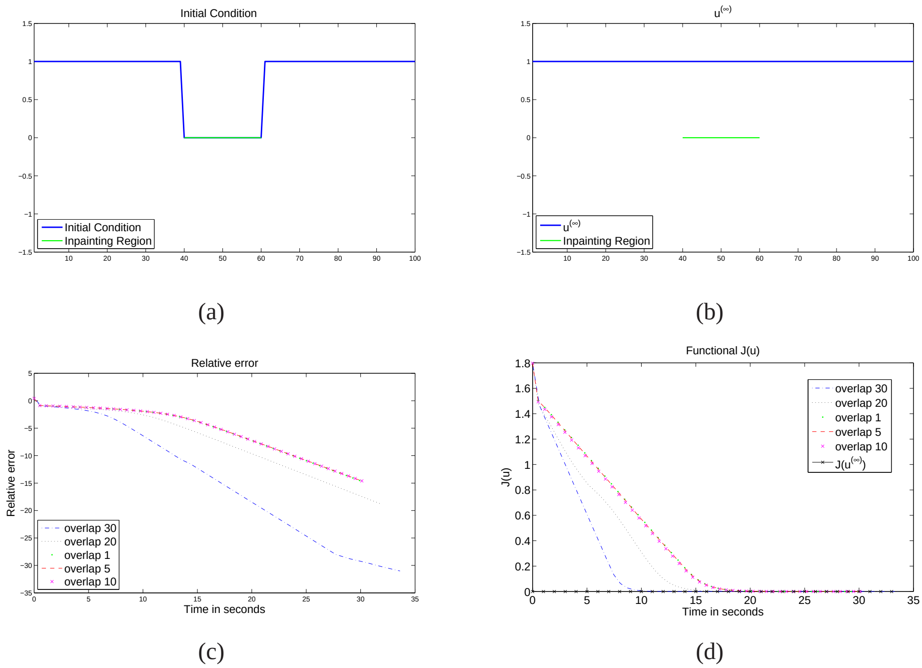


Figure 3.10 We present a numerical experiment related to the interpolation of a 1D signal by total variation minimization. The original signal is only provided outside of the green subinterval. The initial datum g is shown in (a). As expected, the minimizer $u^{(\infty)}$ is the constant vector 1, as shown in (b). In (c) and (d) we display the rates of decay of the relative error and of the value of $\mathcal{J}$ respectively, for applications of the algorithm (3.210) with different sizes G=1,5,10,20,30 of the overlapping region of two subintervals.

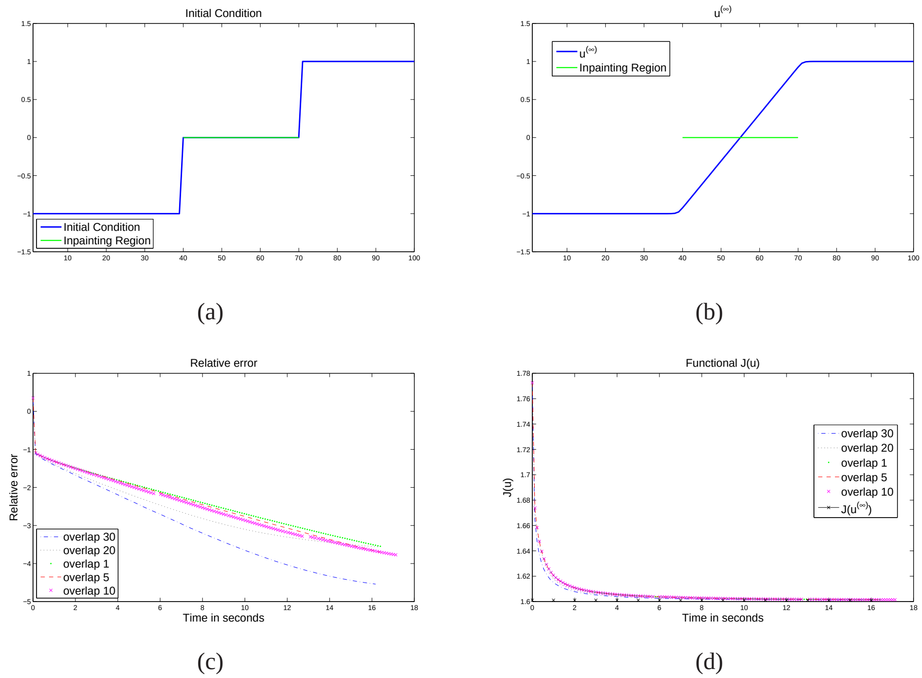
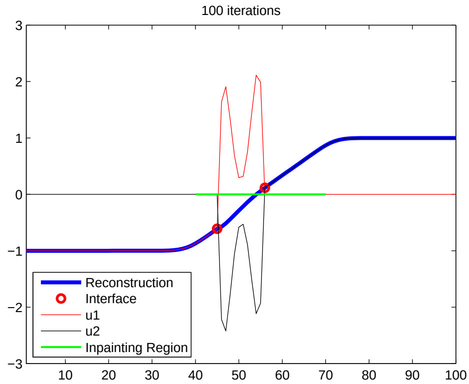


Figure 3.11 We show a second example of total variation interpolation in 1D. The initial datum g is shown in (a). As expected, a minimizer $u^{(\infty)}$ is (nearly) a piecewise linear function, as shown in (b). In (c) and (d) we display the rates of decay of the relative error and of the value of $\mathcal{J}$ respectively, for applications of the algorithm (3.210) with different sizes G=1,5,10,20,30 of the overlapping region of two subintervals.

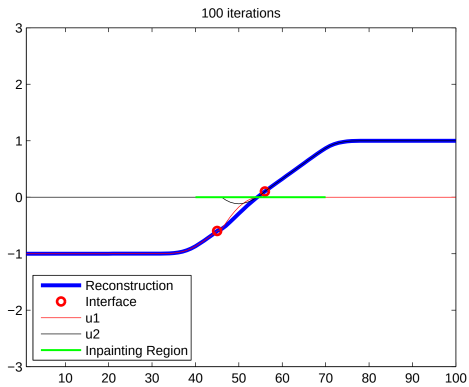
In Figure 3.10 and Figure 3.11 we show a partially corrupted 1D signal on an interval Ω of 100 sampling points, with a loss of information on an interval $D \subset \Omega$. The domain D of the missing signal points is marked with green. These signal points are reconstructed by total variation interpolation, i.e., minimizing the functional $\mathcal{J}$ in (2.80) with $\alpha = 0.4$ and $Ku = 1_{\Omega \setminus D} \cdot u$, where $1_{\Omega \setminus D}$ is the indicator function of $\Omega \setminus D$. A minimizer $u^{(\infty)}$ of $\mathcal{J}$ is precomputed with an algorithm working on the whole interval Ω without any decomposition. We show also the decay of relative error and of the value of the energy $\mathcal{J}$ for applications of algorithm (3.210) on two subdomains and with different overlap sizes $G = 1, 5, 10, 20, 30$. The fixed points η's are computed on a small interval $\hat{\Omega}_i$, $i = 1, 2$, of size 2. These results confirm the behavior of the algorithm (3.210) as predicted by the theory; the algorithm monotonically decreases $\mathcal{J}$ and computes a minimizer, independently of the size of the overlapping region. A larger overlapping region does not necessarily imply a slower convergence. In these figures we do compare the speed in terms of CPU time. In Figure 3.12 we also illustrate the effect of implementing the BUPU within the domain decomposition algorithm. In this case, with datum g as in Figure 3.11, we chose $\alpha = 1$ and an overlap of size $G = 10$. The fixed points η's are computed on a small interval $\hat{\Omega}_i$, $i = 1, 2$ respectively, of size 6. Figure 3.13 shows an example of the domain decomposition algorithm (3.210) for total variation inpainting. As for the 1D example in Figures 3.10-3.12 the operator K is a multiplier, i.e., $Ku = 1_{\Omega \setminus D} \cdot u$, where Ω denotes the rectangular image domain and $D \subset \Omega$ the missing domain in which the original image content got lost. The regularization parameter α is fixed at the value 10^{-2}. In Figure 3.13 the missing domain D is the black writing which covers parts of the image. Here, the image domain of size 449×570 pixels is split into five overlapping subdomains with an overlap size $G = 28 \times 570$. Further, the fixed points η's are computed on a small stripe $\hat{\Omega}_i$, $i = 1, \dots, 5$ respectively, of size 6×570 pixels. Finally, in Figure 3.14 we illustrate the successful application of our domain decomposition algorithm (3.210) for a compressed sensing problem. Here, we consider a medical-type image (the so-called *Logan-Shepp phantom*) and its reconstruction from only partial Fourier data. In this case the linear operator $K = S \circ \mathcal{F}$, where $\mathcal{F}$ denotes the $2D$ Fourier matrix and S is a *downsampling operator* which selects only a few frequencies as output. We minimize $\mathcal{J}$ with α set at 0.4×10^{-2}. In the application of algorithm (3.210) the image domain of size 256×256 pixels is split into four overlapping subdomains with an overlap size $G = 20 \times 256$. The fixed points η's are computed in a small stripe $\hat{\Omega}_i$, $i = 1, \dots, 4$ respectively, of size 6×256 pixels.

Bibliography

[1] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of Bounded Variation and Free Discontinuity Problems.*, Oxford Mathematical Monographs. Oxford: Clarendon Press. xviii, 2000.

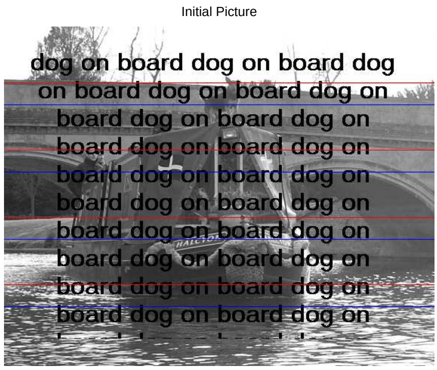


(a)



(b)

Figure 3.12 Here we present two numerical experiments related to the interpolation of a 1D signal by total variation minimization. The original signal is only provided outside of the green subinterval. On the left we show an application of algorithm (3.210) when no correction with the partition of unity is provided. In this case, the sequence of the local iterations $u_1^{(n)}, u_2^{(n)}$ is unbounded. On the right we show an application of algorithm (3.210) with the use of the partition of unity which enforces the uniform boundedness of the local iterations $u_1^{(n)}, u_2^{(n)}$.



(a)

(b)

Figure 3.13 This figure shows an application of algorithm (3.210) for image inpainting. In this simulation the problem was split into five subproblems on overlapping subdomains.

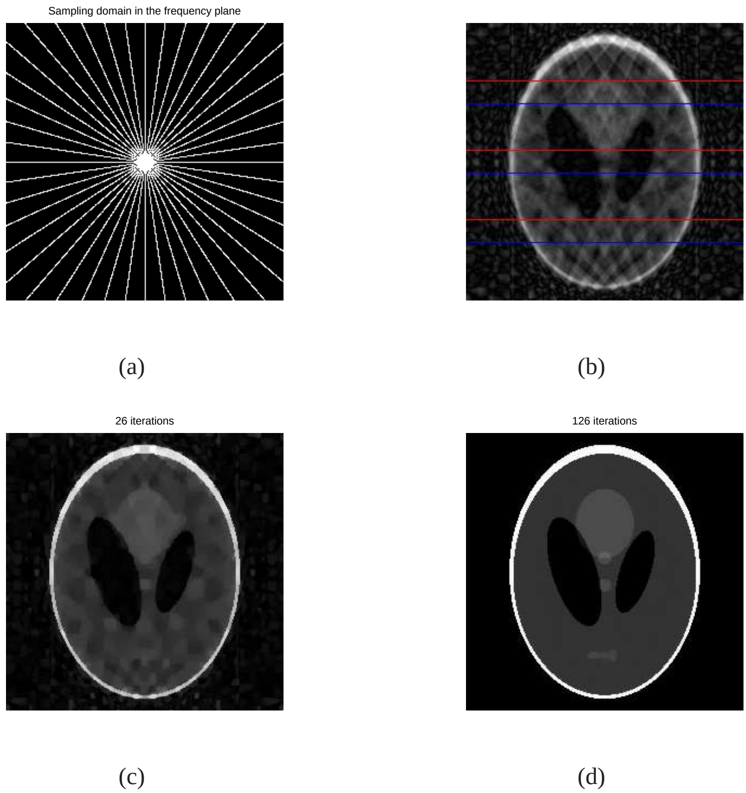


Figure 3.14 We show an application of algorithm (3.210) in a classical compressed sensing problem for recovering piecewise constant medical-type images from given partial Fourier data. In this simulation the problem was split via decomposition into four overlapping subdomains. On the top-left figure, we show the sampling data of the image in the Fourier domain. On the top-right the back-projection provided by the sampled frequency data together with the highlighted partition of the physical domain into four subdomains is shown. The bottom figures present intermediate iterations of the algorithm, i.e., $u^{(26)}$ and $u^{(125)}$.

[2] G. Aubert and P. Kornprobst, *Mathematical Problems in Image Processing. Partial Differential Equations and the Calculus of Variation*, Springer, 2002.

[3] Richard G. Baraniuk, M. Davenport, Ronald A. DeVore, and M. Wakin, *A simple proof of the restricted isometry property for random matrices*, Constr. Approx. 28 (2008), pp. 253–263.

[4] H. H. Bauschke, J. M Borwein, and A. S. Lewis, *The method of cyclic projections for closed convex sets in Hilbert space. Recent developments in optimization theory and nonlinear analysis*, (Jerusalem, 1995), 1–38, Contemp. Math., 204, Amer. Math. Soc., Providence, RI, 1997

[5] A. Beck and M. Teboulle, *A fast iterative shrinkage-thresholding algorithm for linear inverse problems*, SIAM J. Imaging Sci. **2** (2009), no. 1, 183–202.

[6] Thomas Blumensath and Mike E. Davies, *Iterative thresholding for sparse approximations*, J. Fourier Anal. Appl. 14 (2008), pp. 629–654.

[7] ______, *Iterative hard thresholding for compressed sensing*, preprint (2009).

[8] T. Bonesky, S. Dahlke, P. Maass, and T. Raasch, *Adaptive wavelet methods and sparsity reconstruction for inverse heat conduction problems*, Preprint series DFG-SPP 1324, Philipps University of Marburg (2009).

[9] A. Braides, Γ-*Convergence for Beginners*, No.22 in Oxford Lecture Series in Mathematics and Its Applications. Oxford University Press, 2002.

[10] K. Bredies and D. Lorenz, *Linear convergence of iterative soft-thresholding*, J, Fourier Anal. Appl. **14** (2008), no. 5-6, 813–837.

[11] R. E. Bruck and S. Reich, *Nonexpansive projections and resolvents of accretive operators in Banach spaces*. Houston J. Math. 3 (1977), no. 4, 459–470

[12] Emmanuel J. Candès, J. Romberg, and Terence Tao, *Stable signal recovery from incomplete and inaccurate measurements*, Comm. Pure Appl. Math. 59 (2006), pp. 1207–1223.

[13] C. Carstensen, *Domain decomposition for a non-smooth convex minimalization problems and its application to plasticity*, Numerical Linear Algebra with Applications, **4** (1998), no. 3, 177–190.

[14] A. Chambolle, *An algorithm for total variation minimization and applications*. J. Math. Imaging Vision **20** (2004), no. 1-2, 89–97.

[15] A. Chambolle, J. Darbon, *On total variation minimization and surface evolution using parametric maximum flows*, UCLA CAM report 08-19, 2008.

[16] A. Chambolle and P.-L. Lions, *Image recovery via total variation minimization and related problems.*, Numer. Math. 76 (1997), pp. 167–188.

[17] T. F. Chan, G. H. Golub, and P. Mulet, *A nonlinear primal-dual method for total variation-based image restoration*, SIAM J. Sci. Comput. **20** (1999), no. 6, 1964–1977.

[18] T. F. Chan, and T. P. Mathew, Domain decomposition algorithms, Acta Numerica **3** (1994), pp. 61–143.

[19] A. K. Cline, *Rate of convergence of Lawson's algorithm*, Math. Comp. 26 (1972), pp. 167–176.

[20] Albert Cohen, Wolfgang Dahmen, and Ronald A. DeVore, *Compressed sensing and best k-term approximation*, J. Amer. Math. Soc. 22 (2009), pp. 211–231.

[21] P. L. Combettes and V. R. Wajs, *Signal recovery by proximal forward-backward splitting*, Multiscale Model. Simul. 4 (2005), pp. 1168–1200.

[22] S. Dahlke, M. Fornasier, and T. Raasch, *Multilevel preconditioning for adaptive sparse optimization*, preprint, 2009..

[23] W. Dahmen, *Wavelet and multiscale methods for operator equations*, Acta Numerica **6** (1997), 55–228.

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

[25] G. Dal Maso, *An Introduction to* Γ*-Convergence.*, Birkhäuser, Boston, 1993.

[26] J. Darbon, and M. Sigelle, *A fast and exact algorithm for total variation minimization*, IbPRIA 2005, 3522(1), pp. 351-359, 2005.

[27] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, 1992.

[28] I. Daubechies, M. Defrise, and C. De Mol, *An iterative thresholding algorithm for linear inverse problems with a sparsity constraint*, Comm. Pure Appl. Math. 57 (2004), pp. 1413–1457.

[29] I. Daubechies, M. Fornasier, and I. Loris, *Accelerated projected gradient methods for linear inverse problems with sparsity constraints*, J. Fourier Anal. Appl. **14** (2008), no. 5-6, 764–792.

[30] I. Daubechies, R. DeVore, Massimo Fornasier, and C. Güntürk, *Iteratively re-weighted least squares minimization for sparse recovery*, Comm. Pure Appl. Math. (to appear).

[31] R. A. DeVore, *Nonlinear approximation*, Acta Numerica **7** (1998), 51–150.

[32] D. Dobson and C. R. Vogel, *Convergence of an iterative method for total variation denoising*, SIAM J. Numer. Anal. **34** (1997), no. 5, 1779–1791.

[33] David L. Donoho, *Compressed sensing*, IEEE Trans. Inform. Theory 52 (2006), pp. 1289–1306.

[34] David L. Donoho and Y. Tsaig, *Fast solution of l1-norm minimization problems when the solution may be sparse*, IEEE Trans. Inform. Theory 54 (2008), pp. 4789–4812.

[35] Bradley Efron, Trevor Hastie, Iain Johnstone, and Robert Tibshirani, *Least angle regression*, Ann. Statist. 32 (2004), pp. 407–499.

[36] I. Ekeland and R. Témam, *Convex analysis and variational problems*, SIAM, 1999.

[37] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, Springer-Verlag, 1996.

[38] L. C. Evans and R. F. Gariepy, *Measure Theory and Fine Properties of Functions*, CRC Press, 1992.

[39] M. Figueiredo, R. Nowak, and S. J. Wright, *Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems*, IEEE J. Selected Topics in Signal Process. **4** (2007), no. 1, 586–597.

[40] ______, *Sparse reconstruction by separable approximation*, IEEE Trans. Signal Proc. (2009).

[41] M. A. T. Figueiredo and R. D. Nowak, *An EM algorithm for wavelet-based image restoration.*, IEEE Trans. Image Proc. 12 (2003), pp. 906–916.

[42] M. Fornasier, *Domain decomposition methods for linear inverse problems with sparsity constraints*, Inverse Problems 23 (2007), pp. 2505–2526.

[43] M. Fornasier and R. March, *Restoration of color images by vector valued BV functions and variational calculus*, SIAM J. Appl. Math 68 (2007), pp. 437–460.

[44] M. Fornasier and C.-B. Schönlieb, *Subspace correction methods for total variation and* ℓ_1-*minimization*, to appear in SIAM J. Numer. Anal. (2009).

[45] A.Y. Garnaev and E.D. Gluskin, *On widths of the Euclidean ball*, Sov. Math., Dokl. 30 (1984), pp. 200–204.

[46] E.D. Gluskin, *Norms of random matrices and widths of finite-dimensional sets*, Math. USSR-Sb. 48 (1984), pp. 173–182.

[47] I. F. Gorodnitsky and B. D. Rao, *Sparse signal reconstruction from limited data using FOCUSS: a recursive weighted norm minimization algorithm*, IEEE Transactions on Signal Processing 45 (1997), pp. 600–616.

[48] T. Goldstein, S. Osher, *The split Bregman method for* L^1 *regularized problems*, UCLA CAM Report 08-29, 2008.

[49] J.-B. Hiriart-Urruty and C. Lemaréchal, *Convex Analysis and Minimization Algorithms I*, Vol. 305 of Grundlehren der mathematischen Wissenschaften, Springer-Verlag: Berlin, 1996

[50] B.S. Kashin, *Diameters of some finite-dimensional sets and classes of smooth functions.*, Math. USSR, Izv. 11 (1977), pp. 317–333.

[51] S.-J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, *A method for large-scale* ℓ_1-*regularized least squares problems with applications in signal processing and statistics*, IEEE Journal on Selected Topics in Signal Process. (2007).

[52] C. L. Lawson, *Contributions to the Theory of Linear Least Maximum Approximation*, 1961, Ph.D. thesis, University of California, Los Angeles.

[53] I. Loris, G. Nolet, I. Daubechies, and F. A. Dahlen, *Tomographic inversion using* ℓ_1-*norm regularization of wavelet coefficients*, Geophysical Journal International **170** (2007), no. 1, 359–370.

[54] Ignace Loris, *On the performance of algorithms for the minimization of* 1-*penalized functionals*, Inverse Problems **25** (2009), 035008.

[55] M. Lustig, D. Donoho, and J. M. Pauly, *Sparse MRI: The application of compressed sensing for rapid MR imaging*, Magnetic Resonance in Medicine **58** (2007), no. 6, 1182–1195.

[56] Stéphane G. Mallat and Zhifeng Zhang, *Matching pursuits with time-frequency dictionaries.*, IEEE Trans. Signal Process. 41 (1993), pp. 3397–3415.

[57] J. Müller, *Parallel Methods for Nonlinear Imaging Techniques*, Master thesis, University of Münster, 2008.

[58] B. K. Natarajan, *Sparse approximate solutions to linear systems.*, SIAM J. Comput. 24 (1995), pp. 227–234.

[59] Y. Nesterov, *Smooth minimization of non-smooth functions*. Mathematic Programming, Ser. A, **103** (2005), 127–152.

[60] Yurii Nesterov and Arkadii Nemirovskii, *Interior-point polynomial algorithms in convex programming*, SIAM Studies in Applied Mathematics, vol. 13, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994.

[61] Zdzisław Opial, *Weak convergence of the sequence of successive approximations for non-expansive mappings*, Bull. Amer. Math. Soc. 73 (1967), pp. 591–597.

[62] M. Osborne, Brett Presnell, and B. Turlach, *A new approach to variable selection in least squares problems*, IMA J. Numer. Anal. 20 (2000), pp. 389–403.

[63] M. R. Osborne, *Finite algorithms in optimization and data analysis*, Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics, John Wiley & Sons Ltd., Chichester, 1985.

[64] Michael Osborne, Brett Presnell, and Berwin Turlach, *On the LASSO and its dual*, J. Comput. Graph. Statist. 9 (2000), pp. 319–337.

[65] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, *An iterative regularization method for total variation-based image restoration*, Multiscale Model. Simul. **4**, no. 2 (2005) 460-489.

[66] R. Ramlau, G. Teschke, and M. Zhariy, *A compressive Landweber iteration for solving ill-posed inverse problems*, Inverse Problems **24** (2008), no. 6, 065013.

[67] R.T. Rockafellar and R.J.B. Wets, *Variational analysis*, Grundlehren der Mathematischen Wissenschaften, vol. 317, Springer-Verlag, Berlin, 1998.

[68] Leonid Rudin, Stanley Osher, and Emad Fatemi, *Nonlinear total variation based noise removal algorithms.*, Physica D 60 (1992), pp. 259–268.

[69] C.-B. Schönlieb, *Total variation minimization with an H^{-1} constraint*, CRM Series 9, Singularities in Nonlinear Evolution Phenomena and Applications proceedings, Scuola Normale Superiore Pisa 2009, 201-232.

[70] J.-L. Starck, E. J. Candès, and D. L. Donoho, *Astronomical image representation by curvelet transform*, Astronomy and Astrophysics 298 (2003), pp. 785–800.

[71] J.-L. Starck, M. K. Nguyen, and F. Murtagh, *Wavelets and curvelets for image deconvolution: a combined approach*, Signal Proc. 83 (2003), pp. 2279–2283.

[72] Robert Tibshirani, *Regression shrinkage and selection via the lasso*, J. Roy. Statist. Soc. Ser. B 58 (1996), pp. 267–288.

[73] Joel Tropp and D. Needell, *CoSaMP: Iterative signal recovery from incomplete and inaccurate samples*, Appl. Comput. Harmon. Anal. (2008), p. 30.

[74] Joel A. Tropp, *Greed is good: Algorithmic results for sparse approximation*, IEEE Trans. Inform. Theory 50 (2004), pp. 2231–2242.

[75] X.-C. Tai and P. Tseng, *Convergence rate analysis of an asynchronous space decomposition method for convex minimization*, Math. Comp. **71** (2001), no. 239, 1105–1135.

[76] X.-C. Tai and J. Xu, *Global convergence of subspace correction methods for convex optimization problems*, Math. Comp., **71** (2002), no. 237, 105–124.

[77] L. Vese, *A study in the BV space of a denoising-deblurring variational problem.*, Appl. Math. Optim. 44 (2001), pp. 131–161.

[78] P. Weiss, L. Blanc-Féraud, and G. Aubert, *Efficient schemes for total variation minimization under constraints in image processing*, SIAM J. Sci. Comput., (2009) to appear.

[79] Matlab code and numerical experiments of the methods provided in this paper can be downloaded at the web-page: `http://homepage.univie.ac.at/carola.schoenlieb/webpage$\_$tvdode/tv$\_$dode$\_$numerics.htm`

Author information

Massimo Fornasier, RICAM, Austrian Academy of Sciences, Linz, Austria.
E-mail: `massimo.fornasier@oeaw.ac.at`