# Iteratively Reweighted Least Squares Minimization for Sparse Recovery

INGRID DAUBECHIES
*Princeton University*

RONALD DEVORE
*University of South Carolina*

MASSIMO FORNASIER
*Austrian Academy of Sciences*

AND

C. SİNAN GÜNTÜRK
*Courant Institute*

**Abstract**

Under certain conditions (known as the *restricted isometry property*, or RIP) on the $m \times N$ matrix $\Phi$ (where $m < N$), vectors $x \in \mathbb{R}^N$ that are sparse (i.e., have most of their entries equal to 0) can be recovered exactly from $y := \Phi x$ even though $\Phi^{-1}(y)$ is typically an $(N-m)$–dimensional hyperplane; in addition, $x$ is then equal to the element in $\Phi^{-1}(y)$ of minimal $\ell_1$-norm. This minimal element can be identified via linear programming algorithms. We study an alternative method of determining $x$, as the limit of an *iteratively reweighted least squares* (IRLS) algorithm. The main step of this IRLS finds, for a given weight vector $w$, the element in $\Phi^{-1}(y)$ with smallest $\ell_2(w)$-norm. If $x^{(n)}$ is the solution at iteration step $n$, then the new weight $w^{(n)}$ is defined by $w_i^{(n)} := [|x_i^{(n)}|^2 + \varepsilon_n^2]^{-1/2}$, $i = 1, \dots, N$, for a decreasing sequence of adaptively defined $\varepsilon_n$; this updated weight is then used to obtain $x^{(n+1)}$ and the process is repeated. We prove that when $\Phi$ satisfies the RIP conditions, the sequence $x^{(n)}$ converges for all $y$, regardless of whether $\Phi^{-1}(y)$ contains a sparse vector. If there is a sparse vector in $\Phi^{-1}(y)$, then the limit is this sparse vector, and when $x^{(n)}$ is sufficiently close to the limit, the remaining steps of the algorithm converge exponentially fast (*linear convergence* in the terminology of numerical optimization). The same algorithm with the "heavier" weight $w_i^{(n)} = [|x_i^{(n)}|^2 + \varepsilon_n^2]^{-1+\tau/2}$, $i = 1, \dots, N$, where $0 < \tau < 1$, can recover sparse solutions as well; more importantly, we show its local convergence is superlinear and approaches a *quadratic* rate for $\tau$ approaching 0. © 2009 Wiley Periodicals, Inc.

## 1 Introduction

Let $\Phi$ be an $m \times N$ matrix with $m < N$ and let $y \in \mathbb{R}^m$. (In the compressed sensing application that motivated this study, $\Phi$ typically has full rank,

Communications on Pure and Applied Mathematics, Vol. LXIII, 0001–0038 (2010)

i.e., $\mathrm{Ran}(\Phi) = \mathbb{R}^m$. We shall implicitly assume, throughout the paper, that this is the case. Our results still hold for the case where $\mathrm{Ran}(\Phi) \subsetneq \mathbb{R}^m$, with the proviso that $y$ must then lie in $\mathrm{Ran}(\Phi)$.)

The linear system of equations

$$\Phi x = y \tag{1.1}$$

is underdetermined and has infinitely many solutions. If $\mathcal{N} := \mathcal{N}(\Phi)$ is the null space of $\Phi$ and $x_0$ is any solution to (1.1), then the set $\mathcal{F}(y) := \Phi^{-1}(y)$ of all solutions to (1.1) is given by $\mathcal{F}(y) = x_0 + \mathcal{N}$.

In the absence of any other information, no solution to (1.1) is to be preferred over any other. However, many scientific applications work under the assumption that the desired solution $x \in \mathcal{F}(y)$ is either sparse or well approximated by (a) sparse vector(s). Here and later, we say a vector *has sparsity* $k$ (or is $k$*-sparse*) if it has *at most* $k$ nonzero coordinates. Suppose then that we know that the desired solution of (1.1) is $k$-sparse, where $k < m$ is known. How could we find such an $x$? One possibility is to consider any set $T$ of $k$ column indices and find the least squares solution $x^T := \mathrm{argmin}_{z \in \mathcal{F}(y)} \|\Phi_T z - y\|_{\ell_2^m}$, where $\Phi_T$ is obtained from $\Phi$ by setting to 0 all entries that are not in columns from $T$. Finding $x^T$ is numerically simple . After finding each $x^T$, we choose the particular set $T^*$ that minimizes the residual $\|\Phi_T z - y\|_{\ell_2^m}$. This would find a $k$-sparse solution (if it exists), $x^* = x^{T^*}$. However, this naive method is numerically prohibitive when $N$ and $k$ are large, since it requires solving $\binom{N}{k}$ least squares problems.

An attractive alternative to the naive minimization is its convex relaxation that consists in selecting the element in $\mathcal{F}(y)$ that has minimal $\ell_1$-norm:

$$x := \underset{z \in \mathcal{F}(y)}{\mathrm{argmin}} \|z\|_{\ell_1^N}. \tag{1.2}$$

Here and later we use the $\ell_p$-norms

$$\|x\|_{\ell_p} := \|x\|_{\ell_p^N} := \begin{cases} (\sum_{j=1}^N |x_j|^p)^{1/p}, & 0 < p < \infty, \\ \max_{j=1,\dots,N} |x_j|, & p = \infty. \end{cases} \tag{1.3}$$

Under certain assumptions on $\Phi$ and $y$ that we shall describe in Section 2, it is known that (1.2) has a unique solution (which we shall denote by $x^*$), and that, when there is a $k$-sparse solution to (1.1), (1.2) will find this solution [7, 8, 21, 22]. Because problem (1.2) can be formulated as a linear program, it is numerically tractable.

Solving underdetermined systems by $\ell_1$-minimization has a long history. It is at the heart of many numerical algorithms for approximation, compression, and statistical estimation. The use of the $\ell_1$-norm as a sparsity-promoting functional can be found first in reflection seismology and in deconvolution of seismic traces [17, 39, 40]. Rigorous results for $\ell_1$-minimization began to appear in the late 1980s, with Donoho and Stark [24] and Donoho and Logan [23]. Applications for $\ell_1$-minimization in statistical estimation began in the mid-1990s with the introduction of the LASSO and related formulations [41] (iterative soft-thresholding), also known as "basis pursuit" [16], proposed in compression applications for extracting the sparsest signal representation from highly overcomplete frames. Around the same time other signal processing groups started using $\ell_1$-minimization for the analysis of sparse signals; see, e.g., [34]. The applications and understanding of $\ell_1$-minimization saw a dramatic increase in the last five years [6, 7, 8, 9, 21, 22, 25, 26], with the development of fairly general mathematical frameworks in which $\ell_1$-minimization, known heuristically to be sparsity promoting, can be *proved* to recover sparse solutions *exactly*. We shall not trace all the relevant results and applications; a detailed history is beyond the scope of this introduction. We refer the reader to the survey papers [1, 4]. The reader can also find a comprehensive collection of the ongoing recent developments at the website `http://www.dsp.ece.rice.edu/cs/`. In fact, $\ell_1$-minimization has been so surprisingly effective in several applications that Candès, Wakin, and Boyd call it the "modern least squares" in [10]. We thus clearly need efficient algorithms for the minimization problem (1.2).

Several alternatives to (1.2) (see, e.g., [27, 33]) have been proposed as possibly more efficient numerically, or simpler to implement by nonexperts, than standard algorithms for linear programming (such as interior point or barrier methods). In this paper we clarify fine convergence properties of one such alternative method, called *iteratively reweighted least squares minimization* (IRLS). It begins with the following observation (see Section 2 for details). If (1.2) has a solution $x^*$ that has no vanishing coordinates, then the (unique!) solution $x^w$ of the weighted least squares problem

$$x^w := \operatorname*{argmin}_{z \in \mathcal{F}(y)} \|z\|_{\ell_2^N(w)}, \quad w := (w_1, \dots, w_N), \quad \text{where } w_j := |x_j^*|^{-1}, \tag{1.4}$$

coincides with $x^*$.[1]

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 (1.4) 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

[1] The following argument provides a short proof by contradiction of this statement. Assume that $x^*$ is not the $\ell_2^N(w)$-minimizer. Then there exists $\eta \in \mathcal{N}$ such that

$$\|x^* + \eta\|^2_{\ell_2^N(w)} < \|x^*\|^2_{\ell_2^N(w)}$$

or equivalently

$$\frac{1}{2}\|\eta\|^2_{\ell_2^N(w)} < -\sum_{j=1}^N w_j \eta_j x_j^* = -\sum_{j=1}^N \eta_j \operatorname{sign}(x_j^*).$$

However, because $x^*$ is an $\ell_1$-minimizer, we have $\|x^*\|_{\ell_1} \le \|x^* + h\eta\|_{\ell_1}$ for all $h \neq 0$; taking $h$ sufficiently small, this implies $\sum_{j=1}^N \eta_j \operatorname{sign}(x_j^*) = 0$, a contradiction.

time in the approximation practice in the doctoral thesis of Lawson in 1961 [32], in the form of an algorithm for solving uniform approximation problems, in particular by Chebyshev polynomials, by means of limits of weighted $\ell_p$-norm solutions. This iterative algorithm is now well-known in classical approximation theory as Lawson's algorithm. In [18] it is proved that this algorithm has in principle a linear convergence rate (see also [3] for a set of conditions on the weights for the convergence of the algorithm). In the 1970s extensions of Lawson's algorithm for $\ell_p$-minimization, and in particular $\ell_1$-minimization, were proposed. In signal analysis, IRLS was proposed as a technique to build algorithms for sparse signal reconstruction in [30]. Perhaps the most comprehensive mathematical analysis of the performance of IRLS for $\ell_p$-minimization was given in the work of Osborne [35].

Osborne proves that a suitable IRLS method is convergent for $1 < p < 3$. For $p = 1$, if $w^n$ denotes the weight at the $n^{\text{th}}$ iteration and $x^n$ the minimal weighted least squares solution for this weight, then the algorithm considered by Osborne defines the new weight $w^{n+1}$ coordinatewise as $w_j^{n+1} := |x_j^n|^{-1}$. His main conclusion in this case is that if the $\ell_1$-minimization problem (1.2) has a unique, nondegenerate solution, then the IRLS procedure converges with first-order rate, i.e., linearly (exponentially), with a constant "contraction factor."

The proof given in [35] is local, assuming that the iterations begin with a vector sufficiently close to the solution. The analysis does not take into consideration what happens if one of the coordinates vanishes at some iteration $n$, i.e., $x_j^n = 0$. Taking this to impose that the corresponding weight component $w_j^{n+1}$ must "equal" $\infty$ leads to $x_j^{n+1} = 0$ at the next iteration as well; this then persists in all later iterations. If $x_j^* = 0$, all is well, but if there is an index $j$ for which $x_j^* \neq 0$, yet $x_j^n = 0$ at some iteration step $n$, then this "infinite weight" prescription leads to problems. In practice, this is avoided by changing the definition of the weight at coordinates $j$ where $x_j^n = 0$. (See [33] and [11, 29] where a variant for total variation minimization is studied; such modified algorithms need no longer converge to $x^*$, however. See also the probabilistic approach and analysis given in [28] on the singularity problem.)

The uniqueness assumption for the minimum $\ell_1$-norm solution $x^*$ implies that $x^*$ is $k$-sparse for $k \leq m$ (see Remark 2.2). On the other hand, the nondegeneracy assumption of [35] amounts to saying that $x^*$ has exactly $m$ nonzero coordinates and seems restrictive.

The purpose of the present paper is to put forward an IRLS algorithm that gives a reweighting without infinite components in the weight, and to provide an analysis of this algorithm, with various results about its convergence and rate of convergence. It turns out that care must be taken in just how the new weight $w^{n+1}$ is derived from the solution $x^n$ of the current weighted least squares problem. To manage this difficulty, we shall consider a very specific recipe for generating the weights. Other recipes are certainly possible.

Given a real number $\varepsilon > 0$ and a weight vector $w \in \mathbb{R}^N$, with $w_j > 0$, $j = 1, \dots, N$, we define

$$\mathcal{J}(z, w, \varepsilon) := \frac{1}{2}\left[\sum_{j=1}^{N} z_j^2 w_j + \sum_{j=1}^{N}(\varepsilon^2 w_j + w_j^{-1})\right], \quad z \in \mathbb{R}^N. \tag{1.5}$$

Given $w$ and $\varepsilon$, the element $z \in \mathbb{R}^N$ that minimizes $\mathcal{J}$ is unique because $\mathcal{J}$ is strictly convex.

Our algorithm will use an alternating method for choosing weights and minimizers based on the functional $\mathcal{J}$. To describe this, 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 $i^{\text{th}}$-largest element of the set $\{|z_j|,\ j = 1, \dots, N\}$, and a vector $v$ is $k$-sparse if and only if $r(v)_{k+1} = 0$.

*Algorithm* 1. We initialize by taking $w^0 := (1, \dots, 1)$. We also set $\varepsilon_0 := 1$. We then recursively define for $n = 0, 1, \dots,$

$$x^{n+1} := \underset{z \in \mathcal{F}(y)}{\operatorname{argmin}}\ \mathcal{J}(z, w^n, \varepsilon_n) = \underset{z \in \mathcal{F}(y)}{\operatorname{argmin}} \|z\|_{\ell_2(w^n)} \tag{1.6}$$

and

$$\varepsilon_{n+1} := \min\left(\varepsilon_n, \frac{r(x^{n+1})_{K+1}}{N}\right), \tag{1.7}$$

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

$$w^{n+1} := \underset{w > 0}{\operatorname{argmin}}\ \mathcal{J}(x^{n+1}, w, \varepsilon_{n+1}). \tag{1.8}$$

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

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

$$x^{n+1} = D_n \Phi^{\mathsf{T}} (\Phi D_n \Phi^{\mathsf{T}})^{-1} y, \tag{1.9}$$

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

$$w_j^{n+1} = [(x_j^{n+1})^2 + \varepsilon_{n+1}^2]^{-1/2}, \quad j = 1, \dots, N. \tag{1.10}$$

We shall prove several results about the convergence and rate of convergence of this algorithm. This will be done under the following assumption on $\Phi$.

*Property* (Restricted Isometry Property (RIP)). We say that the matrix $\Phi$ satisfies the restricted isometry property of order $L$ with constant $\delta \in (0, 1)$ if for each vector $z$ with sparsity $L$ we have

$$(1 - \delta)\|z\|_{\ell_2^N} \le \|\Phi z\|_{\ell_2^m} \le (1 + \delta)\|z\|_{\ell_2^N}. \tag{1.11}$$

The RIP was introduced by Candès and Tao [8, 9] in their study of compressed sensing and $\ell_1$-minimization. It has several analytical and geometrical interpretations that will be discussed in Section 3. To mention just one of these results (see [19]), it is known that if $\Phi$ has the RIP of order $L := J + J'$, with

$$\delta < \frac{\sqrt{J'} - \sqrt{J}}{\sqrt{J'} + \sqrt{J}}$$

(here $J' > J$) and if (1.1) has a $J$-sparse solution $z \in \mathcal{F}(y)$, then this solution is the unique $\ell_1$-minimizer in $\mathcal{F}(y)$. (This can still be sharpened: in [5], Candès showed that if $\mathcal{F}(y)$ contains a $J$-sparse vector, and if $\Phi$ has RIP of order $2J$ with $\delta < \sqrt{2} - 1$, then that $J$-sparse vector is unique and is the unique $\ell_1$-minimizer in $\mathcal{F}(y)$.)

The main result of this paper (Theorem 5.3) is that whenever $\Phi$ satisfies the RIP of order $K + K'$ (for some $K' > K$) and $\delta$ sufficiently close to 0, then Algorithm 1 converges to a solution $\bar{x}$ of (1.1) for each $y \in \mathbb{R}^m$. Moreover, if there is a solution $z$ to (1.1) that has sparsity $k \leq K - \kappa$, then $\bar{x} = z$. Here $\kappa > 1$ depends on the RIP constant $\delta$ and can be made arbitrarily close to 1 when $\delta$ is made small. The result cited in our previous paragraph implies that in this case $\bar{x} = x^*$, where $x^*$ is the $\ell_1$-minimal solution to (1.1).

A second part of our analysis concerns rates of convergence. We shall show that if (1.1) has a $k$-sparse solution with, e.g., $k \leq K - 4$ and if $\Phi$ satisfies the RIP of order $3K$ with $\delta$ sufficiently close to 0, then Algorithm 1 converges exponentially fast to $\bar{x} = x^*$. Namely, once $x^{n_0}$ is sufficiently close to its limit $\bar{x}$, we have

$$\|\bar{x} - x^{n+1}\|_{\ell_1^N} \leq \mu \|\bar{x} - x^n\|_{\ell_1^N}, \quad n \geq n_0, \tag{1.12}$$

where $\mu < 1$ is a fixed constant (depending on $\delta$). From this result it follows that we have exponential convergence to $\bar{x}$ whenever $\bar{x}$ is $k$-sparse; however, we have no real information on how long it will take before the iterates enter the region where we can control $\mu$. (Note that this is similar to convergence results for the interior point algorithms that can be used for direct $\ell_1$-minimization.)

The potential of IRLS algorithms, tailored to mimic $\ell_1$-minimization and so recover sparse solutions, has recently been investigated numerically by Chartrand and several co-authors [12, 13, 15]. Our work provides proofs of several findings listed in these works.

One of the virtues of our approach is that, with minor technical modifications, it allows a similar detailed analysis of IRLS algorithms with weights that promote the *nonconvex* optimization of $\ell_\tau$-norms for $0 < \tau < 1$. We can show not only that these algorithms can again recover sparse solutions, but also that their local rate of convergence is superlinear and tends to quadratic when $\tau$ tends to 0. Thus we also justify theoretically the recent numerical results by Chartrand et al. concerning such nonconvex $\ell_\tau$-norm optimization [12, 13, 14, 38].

In the next section we make some remarks about $\ell_1$- and weighted $\ell_2$-minimization, which we shall use in our proof. In the following section, we recall the restricted isometry property and the null space property, including some of its consequences that are important to our analysis. In Section 4 we gather some preliminary results we shall need to prove our main convergence result, Theorem 5.3. We then turn to the issue on rate of convergence in Section 6. In Section 7 we generalize the convergence results obtained for $\ell_1$-minimization to the case of $\ell_\tau$-spaces for $0 < \tau < 1$; in particular, we show, with Theorem 7.9, the local superlinear convergence of the IRLS algorithm in this setting. We conclude the paper with a short section dedicated to a few numerical examples that dovetail nicely with the theoretical results.

## 2 Characterization of $\ell_1$- and Weighted $\ell_2$-Minimizers

We fix $y \in \mathbb{R}^m$ and consider the underdetermined system $\Phi x = y$. Given a norm $\|\cdot\|$, the problem of minimizing $\|z\|$ over $z \in \mathcal{F}(y)$ can be viewed as a problem of approximation. Namely, for any $x_0 \in \mathcal{F}(y)$, we can characterize the minimizers in $\mathcal{F}(y)$ as exactly those elements $z \in \mathcal{F}(y)$ that can be written as $z = x_0 + \eta$, with $\eta$ a best approximation to $-x_0$ from $\mathcal{N}$. In this way one can characterize minimizers $z$ from classical results on best approximation in normed spaces. We consider two examples of this in the present section, corresponding to the $\ell_1$-norm and the weighted $\ell_2(w)$-norm.

Throughout this paper, we shall denote by $x$ any element from $\mathcal{F}(y)$ that has smallest $\ell_1$-norm, as in (1.2). When $x$ is unique, we shall emphasize this by denoting it by $x^*$. We begin with the following well-known lemma (see, for example, Pinkus [36]), which characterizes the minimal $\ell_1$-norm elements from $\mathcal{F}(y)$.

LEMMA 2.1 *An element $x \in \mathcal{F}(y)$ has minimal $\ell_1$-norm among all elements $z \in \mathcal{F}(y)$ if and only if*

$$\left|\sum_{x_i \neq 0} \operatorname{sign}(x_i)\eta_i\right| \leq \sum_{x_i = 0} |\eta_i| \quad \forall \eta \in \mathcal{N}. \tag{2.1}$$

*Moreover, $x$ is unique if and only if we have strict inequality in* (2.1) *for all $\eta \in \mathcal{N}$ that are not identically* 0.

PROOF: We give the simple proof for completeness of this paper. If $x \in \mathcal{F}(y)$ has minimum $\ell_1$-norm, then we have, for any $\eta \in \mathcal{N}$ and any $t \in \mathbb{R}$,

$$\sum_{i=1}^{N} |x_i + t\eta_i| \geq \sum_{i=1}^{N} |x_i|. \tag{2.2}$$

Fix $\eta \in \mathcal{N}$. If $t$ is sufficiently small, then $x_i + t\eta_i$ and $x_i$ will have the same sign $s_i := \operatorname{sign}(x_i)$ whenever $x_i \neq 0$. Hence (2.2) can be written as

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

Choosing $t$ of an appropriate sign, we see that (2.1) is a necessary condition. If $x$ is unique, then for all $\eta \in \mathcal{N} \setminus \{0\}$ we have strict inequality in (2.2), and subsequently in (2.1).

For the opposite direction, we note that if (2.1) holds, then for each $\eta \in \mathcal{N}$, we have

$$\begin{aligned}\sum_{i=1}^{N} |x_i| &= \sum_{x_i \neq 0} s_i x_i = \sum_{x_i \neq 0} s_i (x_i + \eta_i) - \sum_{x_i \neq 0} s_i \eta_i \\ &\leq \sum_{x_i \neq 0} s_i (x_i + \eta_i) + \sum_{x_i = 0} |\eta_i| \leq \sum_{i=1}^{N} |x_i + \eta_i|,\end{aligned} \tag{2.3}$$

where (2.3) uses (2.1). Hence $x$ has minimal $\ell_1$-norm in $\mathcal{F}(y)$. If we have strict inequality in (2.1) for $\eta \in \mathcal{N} \setminus \{0\}$, then the subsequent strict inequality in (2.3) implies uniqueness of $x$. □

*Remark* 2.2. Lemma 2.1 has the following implication (see also [36]): If $x = x^*$ is the unique $\ell_1$-minimizer in $\mathcal{F}(y)$, then $|\{i : x_i^* = 0\}| \geq \dim \mathcal{N}$. To see this, note that for any set $Z$ of indices such that $|Z| < \dim \mathcal{N}$, there exists a vector $\eta \in \mathcal{N}$, $\eta \neq 0$, with $\eta_i = 0$ for all $i \in Z$. If $Z := \{i : x_i^* = 0\}$ has fewer than $\dim \mathcal{N}$ elements, then such an $\eta$ would give

$$\sum_{x_i^* = 0} |\eta_i| = 0,$$

contradicting the strict inequality in (2.1) implied by uniqueness.

In other words, assuming $\Phi$ is full rank so that $\dim \mathcal{N} = N - m$, we find that unique $\ell_1$-minimizers are $k$-sparse for some $k \leq m$. □

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

$$\langle u, v \rangle_w := \sum_{j=1}^{N} w_j u_j v_j. \tag{2.4}$$

We define

$$x^w := \underset{z \in \mathcal{F}(y)}{\operatorname{argmin}} \|z\|_{\ell_2^N(w)}. \tag{2.5}$$

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

$$\langle x^w, \eta \rangle_w = 0 \quad \forall \eta \in \mathcal{N}. \tag{2.6}$$

Namely, $x^w$ necessarily satisfies (2.6); 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 $x^w$.

| | |
|---|---|
| $z$ | an (arbitrary) element of $\mathcal{F}(y)$ |
| $x$ | any solution of $\min_{z\in\mathcal{F}(y)} \|z\|_{\ell_1}$ |
| $x^*$ | unique solution of $\min_{z\in\mathcal{F}(y)} \|z\|_{\ell_1}$ (notation used only when the minimizer is unique) |
| $x^w$ | unique solution of $\min_{z\in\mathcal{F}(y)} \|z\|_{\ell_2(w)}$, $w_j > 0$ for all $j$ |
| $\bar{x}$ | limit of Algorithm 1 |
| $x^\varepsilon$ | unique solution of $\min_{z\in\mathcal{F}(y)} f_\varepsilon(z)$; see (5.5) |

TABLE 2.1. Notation for solutions and minimizers.

At this point, we present a table of the notation we use in this paper, Table 2.1, to denote various kinds of minimizers and other solutions alike (such as limits of algorithms).

## 3 Restricted Isometry and the Null Space Properties

To analyze the convergence of Algorithm 1, we shall impose the restricted isometry property (RIP) already mentioned in the introduction, or a slightly weaker version, the *null space property*, which will be defined below. Recall that $\Phi$ satisfies the RIP of order $L$ for $\delta \in (0,1)$ (see (1.11)) if and only if

$$(1-\delta)\|z\|_{\ell_2^N} \le \|\Phi z\|_{\ell_2^m} \le (1+\delta)\|z\|_{\ell_2^N} \quad \text{for all } L\text{-sparse } z. \tag{3.1}$$

It is known that many families of matrices satisfy the RIP. While there are deterministic families that are known to satisfy the RIP, the largest range of $L$ (asymptotically, as $N \to \infty$, with, e.g., $m/N$ kept constant) is obtained (to date) by using random families. For example, random families in which the entries of the matrix $\Phi$ are independent realizations of a (fixed) Gaussian or Bernoulli random variable are known to have the RIP with high probability for each $L \le c_0(\delta)m/\log N$ (see [2, 8, 9, 37] for a discussion of these results).

We shall say that $\Phi$ has the *null space property* (NSP) of order $L$ for $\gamma > 0$ if

$$\|\eta_T\|_{\ell_1} \le \gamma \|\eta_{T^c}\|_{\ell_1} \tag{3.2}$$

for all sets $T$ of cardinality not exceeding $L$ and all $\eta \in \mathcal{N}$.[2] Here and later, we denote by $\eta_S$ the vector obtained from $\eta$ by setting to 0 all coordinates $\eta_i$ for $i \notin S \subset \{1,2,\dots,N\}$; $T^c$ denotes the complement of the set $T$. It is shown in lemma 4.1 of [19] that if $\Phi$ has the RIP of order $L := J + J'$ for a given $\delta \in (0,1)$, where $J, J' \ge 1$ are integers, then $\Phi$ has the NSP of order $J$ for

$$\gamma := \frac{1+\delta}{1-\delta}\sqrt{\frac{J}{J'}}.$$

Note that if $J'$ is sufficiently large, then $\gamma < 1$.

[2] This definition of the null space property is a slight variant of that given in [19] but is more convenient for the results in the present paper.

Another result in [19] (see also Lemma 4.3 below) states that in order to guarantee that a $k$-sparse vector $x^*$ is the unique $\ell_1$-minimizer in $\mathcal{F}(y)$, it is sufficient that $\Phi$ have the NSP of order $L \geq k$ and $\gamma < 1$. (In fact, the argument in [9], proving that for $\Phi$ with the RIP, $\ell_1$-minimization identifies sparse vectors in $\mathcal{F}(y)$, can be split into two steps: one that implicitly derives the NSP from the RIP, and the remainder of the proof, which uses only the NSP.)

Note that if the NSP holds for some order $L_0$ and constant $\gamma_0$ (not necessarily $< 1$), then, by choosing $a > 0$ sufficiently small, one can ensure that $\Phi$ has the NSP of order $L = aL_0$ with constant $\gamma < 1$ (see [19] for details). So the effect of requiring that $\gamma < 1$ is tantamount to reducing the range of $L$ appropriately.

When proving results on the convergence of our algorithm later in this paper, we shall state them under the assumptions that $\Phi$ has the NSP for some $\gamma < 1$ and an appropriate value of $L$. Using the observations above, they can easily be rephrased in terms of RIP bounds for $\Phi$.

## 4 Preliminary Results

We first make some comments about the decreasing rearrangement $r(z)$ and the $j$-term approximation errors for vectors in $\mathbb{R}^N$. Let us denote by $\Sigma_k$ the set of all $x \in \mathbb{R}^N$ such that $\#(\operatorname{supp}(x)) \leq k$. For any $z \in \mathbb{R}^N$ and any $j = 1, \dots, N$, we denote by

$$\sigma_j(z)_{\ell_1} := \inf_{w \in \Sigma_j} \|z - w\|_{\ell_1^N} \tag{4.1}$$

the $\ell_1$-error in approximating a general vector $z \in \mathbb{R}^N$ by a $j$-sparse vector. Note that these approximation errors can be written as a sum of entries of $r(z)$: $\sigma_j(z)_{\ell_1} = \sum_{\nu > j} r(z)_\nu$. We have the following lemma:

LEMMA 4.1 *The map $z \mapsto r(z)$ is Lipschitz-continuous on $(\mathbb{R}^N, \|\cdot\|_{\ell_\infty})$ with Lipschitz constant* 1; *i.e., for any $z, z' \in \mathbb{R}^N$, we have*

$$\|r(z) - r(z')\|_{\ell_\infty} \leq \|z - z'\|_{\ell_\infty}. \tag{4.2}$$

*Moreover, for any $j$, we have*

$$|\sigma_j(z)_{\ell_1} - \sigma_j(z')_{\ell_1}| \leq \|z - z'\|_{\ell_1}, \tag{4.3}$$

*and for any $J > j$, we have*

$$(J - j) r(z)_J \leq \|z - z'\|_{\ell_1} + \sigma_j(z')_{\ell_1}. \tag{4.4}$$

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

$$r(z)_j \leq \max_{i \in \Lambda^c} |z_i| \leq \max_{i \in \Lambda^c} |z'_i| + \|z - z'\|_{\ell_\infty} = r(z')_j + \|z - z'\|_{\ell_\infty}. \tag{4.5}$$

We can also reverse the roles of $z$ and $z'$. Therefore, we obtain (4.2). To prove (4.3), we approximate $z$ by a $j$-term best approximation $u \in \Sigma_j$ of $z'$ in $\ell_1$. Then

$$\sigma_j(z)_{\ell_1} \leq \|z - u\|_{\ell_1} \leq \|z - z'\|_{\ell_1} + \sigma_j(z')_{\ell_1}, \tag{4.6}$$

and the result follows from symmetry.

To prove (4.4), it suffices to note that $(J - j)r(z)_J \leq \sigma_j(z)_{\ell_1}$. □

Our next result is an approximate reverse triangle inequality for points in $\mathcal{F}(y)$ when $\Phi$ has the NSP. Its importance to us lies in its implication that whenever two points $z, z' \in \mathcal{F}(y)$ have close $\ell_1$-norms and one of them is compressible (i.e., close to a sparse vector), then they necessarily are close to each other. (Note that it also implies that the other vector must be close to that sparse vector.) This is a geometric property of the null space $\mathcal{N}$.

LEMMA 4.2 *Assume that* (3.2) *holds for some $L$ and $\gamma < 1$. Then for any $z, z' \in \mathcal{F}(y)$, we have*

$$\|z' - z\|_{\ell_1} \leq \frac{1+\gamma}{1-\gamma}(\|z'\|_{\ell_1} - \|z\|_{\ell_1} + 2\sigma_L(z)_{\ell_1}). \tag{4.7}$$

PROOF: Let $T$ be a set of indices of the $L$ largest entries in $z$. Then

$$\begin{aligned}
\|(z'-z)_{T^c}\|_{\ell_1} &\leq \|z'_{T^c}\|_{\ell_1} + \|z_{T^c}\|_{\ell_1} \\
&= \|z'\|_{\ell_1} - \|z'_T\|_{\ell_1} + \sigma_L(z)_{\ell_1} \\
&= \|z\|_{\ell_1} + \|z'\|_{\ell_1} - \|z\|_{\ell_1} - \|z'_T\|_{\ell_1} + \sigma_L(z)_{\ell_1} \\
&= \|z_T\|_{\ell_1} - \|z'_T\|_{\ell_1} + \|z'\|_{\ell_1} - \|z\|_{\ell_1} + 2\sigma_L(z)_{\ell_1} \\
&\leq \|(z'-z)_T\|_{\ell_1} + \|z'\|_{\ell_1} - \|z\|_{\ell_1} + 2\sigma_L(z)_{\ell_1}. \qquad (4.8)
\end{aligned}$$

Using (3.2), this gives

$$\begin{aligned}
\|(z'-z)_T\|_{\ell_1} &\leq \gamma\|(z'-z)_{T^c}\|_{\ell_1} \\
&\leq \gamma(\|(z'-z)_T\|_{\ell_1} + \|z'\|_{\ell_1} - \|z\|_{\ell_1} + 2\sigma_L(z)_{\ell_1}). \qquad (4.9)
\end{aligned}$$

In other words,

$$\|(z'-z)_T\|_{\ell_1} \leq \frac{\gamma}{1-\gamma}(\|z'\|_{\ell_1} - \|z\|_{\ell_1} + 2\sigma_L(z)_{\ell_1}). \tag{4.10}$$

Using this, together with (4.8), we obtain

$$\begin{aligned}
\|z'-z\|_{\ell_1} &= \|(z'-z)_{T^c}\|_{\ell_1} + \|(z'-z)_T\|_{\ell_1} \\
&\leq \frac{1+\gamma}{1-\gamma}(\|z'\|_{\ell_1} - \|z\|_{\ell_1} + 2\sigma_L(z)_{\ell_1}), \qquad (4.11)
\end{aligned}$$

as desired. □

This result then allows the following simple proof of some of the results of [19]:

LEMMA 4.3 *Assume that* (3.2) *holds for some $L$ and $\gamma < 1$. Suppose that $\mathcal{F}(y)$ contains an $L$-sparse vector. Then this vector is the unique $\ell_1$-minimizer in $\mathcal{F}(y)$; denoting it by $x^*$, we have, moreover, for all $v \in \mathcal{F}(y)$,*

$$\|v - x^*\|_{\ell_1} \leq 2\,\frac{1+\gamma}{1-\gamma}\,\sigma_L(v)_{\ell_1}. \tag{4.12}$$

PROOF: For the time being, we denote the $L$-sparse vector in $\mathcal{F}(y)$ by $x_s$. Applying (4.7) with $z' = v$ and $z = x_s$, we find

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

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

If $x'$ were another $\ell_1$-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 $\ell_1$-minimizer in $\mathcal{F}(y)$, which we denote by $x^*$, as proposed earlier.

Finally, we apply (4.7) with $z' = x^*$ and $z = v$, and we obtain

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

where we have used the $\ell_1$-minimization property of $x^*$. □

Our next set of remarks centers around the functional $\mathcal{J}$ defined by (1.5). First, note that for each $n = 1, 2, \dots,$ we have

$$\mathcal{J}(x^{n+1}, w^{n+1}, \varepsilon_{n+1}) = \sum_{j=1}^{N} [(x_j^{n+1})^2 + \varepsilon_{n+1}^2]^{1/2}. \tag{4.13}$$

Second, we have the following monotonicity property: for all $n \ge 0$,

$$\begin{aligned} \mathcal{J}(x^{n+1}, w^{n+1}, \varepsilon_{n+1}) &\le \mathcal{J}(x^{n+1}, w^{n}, \varepsilon_{n+1}) \\ &\le \mathcal{J}(x^{n+1}, w^{n}, \varepsilon_{n}) \\ &\le \mathcal{J}(x^{n}, w^{n}, \varepsilon_{n}). \end{aligned} \tag{4.14}$$

Here the first inequality follows from the minimization property that defines $w^{n+1}$, the second inequality from $\varepsilon_{n+1} \le \varepsilon_n$, 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 $\ell_2$-norm element of $\mathcal{F}(y)$.) Inequality (4.14) for $n = 0$ thus holds for arbitrary $x^0 \in \mathcal{F}(y)$.

LEMMA 4.4 *For each $n \ge 1$ we have*

$$\|x^n\|_{\ell_1} \le \mathcal{J}(x^1, w^0, \varepsilon_0) =: A \tag{4.15}$$

*and*

$$w_j^n \ge A^{-1}, \quad j = 1, \dots, N. \tag{4.16}$$

PROOF: The bound (4.15) follows from (4.14) and

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

The bound (4.16) follows from

$$(w_j^n)^{-1} = [(x_j^n)^2 + \varepsilon_n^2]^{1/2} \le \mathcal{J}(x^n, w^n, \varepsilon_n) \le A,$$

where the last inequality uses (4.14). □

## 5 Convergence of the Algorithm

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

LEMMA 5.1 *Given any $y \in \mathbb{R}^m$, the $x^n$ satisfy*

$$\sum_{n=1}^{\infty} \|x^{n+1} - x^n\|_{\ell_2}^2 \le 2A^2, \tag{5.1}$$

*where $A$ is the constant of Lemma* 4.4. *In particular, we have*

$$\lim_{n\to\infty} (x^n - x^{n+1}) = 0. \tag{5.2}$$

PROOF: For each $n = 1, 2, \dots,$ we have

$$\begin{aligned}
&2[\mathcal{J}(x^n, w^n, \varepsilon_n) - \mathcal{J}(x^{n+1}, w^{n+1}, \varepsilon_{n+1})]\\
&\quad \ge 2[\mathcal{J}(x^n, w^n, \varepsilon_n) - \mathcal{J}(x^{n+1}, w^n, \varepsilon_n)]\\
&\quad = \langle x^n, x^n \rangle_{w^n} - \langle x^{n+1}, x^{n+1} \rangle_{w^n}\\
&\quad = \langle x^n + x^{n+1}, x^n - x^{n+1} \rangle_{w^n}\\
&\quad = \langle x^n - x^{n+1}, x^n - x^{n+1} \rangle_{w^n}\\
&\quad = \sum_{j=1}^{N} w_j^n (x_j^n - x_j^{n+1})^2\\
&\quad \ge A^{-1} \|x^n - x^{n+1}\|_{\ell_2}^2, \qquad (5.3)
\end{aligned}$$

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 (2.6)), and the last inequality uses the bound (4.16) on the weights. If we now sum these inequalities over $n \ge 1$ and note that $\mathcal{J}(x^1, w^1, \varepsilon_1) \le A$, we arrive at (5.1). □

From the monotonicity of $\varepsilon_n$ implied by (1.7), we know that

$$\varepsilon := \lim_{n\to\infty} \varepsilon_n \tag{5.4}$$

exists and is nonnegative. The following functional will play an important role in our proof of convergence:

$$f_\varepsilon(z) := \sum_{j=1}^{N} (z_j^2 + \varepsilon^2)^{1/2}. \tag{5.5}$$

Notice that if $x^n$ converges to a vector $x$, then in view of (4.13), $f_\varepsilon(x)$ is equal to the limit of $\mathcal{J}(x^n, w^n, \varepsilon_n)$.

When $\varepsilon > 0$, the functional $f_\varepsilon$ is strictly convex, and therefore has a unique minimizer on $\mathcal{F}(y)$. Let

$$x^\varepsilon := \operatorname*{argmin}_{z\in\mathcal{F}(y)} f_\varepsilon(z). \tag{5.6}$$

This minimizer is characterized by the following lemma:

LEMMA 5.2 *Let $\varepsilon > 0$ and $z \in \mathcal{F}(y)$. Then $z = x^\varepsilon$ if and only if*

$$\langle z, \eta\rangle_{\widetilde{w}(z,\varepsilon)} = 0 \quad \forall \eta \in \mathcal{N}, \tag{5.7}$$

*where $\widetilde{w}(z,\varepsilon)_i = [z_i^2 + \varepsilon^2]^{-1/2}$, $i = 1, \dots, N$.*

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

$$G_\varepsilon(t) := f_\varepsilon(z + t\eta) - f_\varepsilon(z). \tag{5.8}$$

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

$$G'_\varepsilon(0) = \sum_{i=1}^{N} \frac{\eta_i z_i}{[z_i^2 + \varepsilon^2]^{1/2}} = \langle z, \eta\rangle_{\widetilde{w}(z,\varepsilon)}, \tag{5.9}$$

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 $f_\varepsilon$ on $\mathcal{F}(y)$. Indeed, consider the convex univariate function $[u^2 + \varepsilon^2]^{1/2}$. For any $u_0$, we have from convexity that

$$[u^2 + \varepsilon^2]^{1/2} \ge [u_0^2 + \varepsilon^2]^{1/2} + [u_0^2 + \varepsilon^2]^{-1/2} u_0 (u - u_0), \tag{5.10}$$

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

$$\begin{aligned} f_\varepsilon(v) &\geq f_\varepsilon(z) + \sum_{i=1}^{N} [z_i^2 + \varepsilon^2]^{-1/2} z_i (v_i - z_i) \\ &= f_\varepsilon(z) + \langle z, v - z \rangle_{\widetilde{w}(z,\varepsilon)} \\ &= f_\varepsilon(z), \end{aligned} \tag{5.11}$$

where we have used the orthogonality condition (5.7) for $\eta = v - z \in \mathcal{N}$. Since $v$ is arbitrary, it follows that $z = x^\varepsilon$, as claimed. $\square$

We now show the convergence of the algorithm.

THEOREM 5.3 *Let $K$ (the same index as used in the update rule* (1.7)*) be chosen so that $\Phi$ satisfies the null space property* (3.2) *of order $K$, with $\gamma < 1$. Then, for each $y \in \mathbb{R}^m$, the output of Algorithm* 1 *converges to a vector $\bar{x} \in \mathcal{F}(y)$ 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 $\ell_1$-minimizer $x^*$, and $\bar{x} = x^*$. Moreover, we have, for $k \leq K$, and any $z \in \mathcal{F}(y)$,*

$$\|z - \bar{x}\|_{\ell_1} \leq c\sigma_k(z)_{\ell_1} \quad \textit{with } c := \frac{2(1+\gamma)}{1-\gamma}. \tag{5.12}$$

(ii) *If $\varepsilon := \lim_{n\to\infty} \varepsilon_n > 0$, then $\bar{x} = x^\varepsilon$.*

(iii) *In this last case, if $\gamma$ 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*

$$\|z - \bar{x}\|_{\ell_1} \leq \tilde{c}\sigma_k(z)_{\ell_1} \quad \textit{with } \tilde{c} := \frac{2(1+\gamma)}{1-\gamma}\left[\frac{K - k + 3/2}{K - k - 2\gamma/(1-\gamma)}\right]. \tag{5.13}$$

(iv) *If $\mathcal{F}(y)$ contains a vector $z$ of sparsity $k < K - \frac{2\gamma}{1-\gamma}$, then $\varepsilon = 0$ and $\bar{x} = x^* = z$.*

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

*Case* $\varepsilon = 0$. In this case, we want to prove that $x^n$ converges and that its limit is an $\ell_1$-minimizer. Suppose that $\varepsilon_{n_0} = 0$ for some $n_0$. Then by the definition of the algorithm, we know that the iteration is stopped at $n = n_0$, and $x^n = x^{n_0}$, $n \geq n_0$. Therefore $\bar{x} = x^{n_0}$. From the definition (1.7) of $\varepsilon_n$, 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 Section 3 and Lemma 4.3, if a $K$-sparse solution exists when $\Phi$ satisfies the NSP of order $K$ with $\gamma < 1$, then it is the unique $\ell_1$-minimizer. Therefore, $\bar{x}$ equals $x^*$, this unique minimizer.

Suppose now that $\varepsilon_n > 0$ for all $n$. Since $\varepsilon_n \to 0$, there is an increasing sequence of indices $(n_i)$ such that $\varepsilon_{n_i} < \varepsilon_{n_i - 1}$ for all $i$. Definition (1.7) of $(\varepsilon_n)_{n\in\mathbb{N}}$ implies that $r(x^{n_i})_{K+1} < N\varepsilon_{n_i-1}$ for all $i$. Noting that $(x^n)_{n\in\mathbb{N}}$ is a bounded sequence (see Lemma 4.4), there exists a subsequence $(p_j)_{j\in\mathbb{N}}$ of $(n_i)_{i\in\mathbb{N}}$ such that $(x^{p_j})_{j\in\mathbb{N}}$ converges to a point $\widetilde{x} \in \mathcal{F}(y)$. By Lemma 4.1, we know that $r(x^{p_j})_{K+1}$ converges to $r(\widetilde{x})_{K+1}$. Hence we get

$$r(\widetilde{x})_{K+1} = \lim_{j\to\infty} r(x^{p_j})_{K+1} \leq \lim_{j\to\infty} N\varepsilon_{p_j - 1} = 0, \tag{5.14}$$

which means that $\widetilde{x}$ is $K$-sparse. By the same token used above, we again have that $\widetilde{x} = x^*$, the unique $\ell_1$-minimizer. We must still show that $x^n \to x^*$. Since $x^{p_j} \to x^*$ and $\varepsilon_{p_j} \to 0$, (4.13) implies $\mathcal{J}(x^{p_j}, w^{p_j}, \varepsilon_{p_j}) \to \|x^*\|_{\ell_1}$. By the monotonicity property stated in (4.14), we get $\mathcal{J}(x^n, w^n, \varepsilon_n) \to \|x^*\|_{\ell_1}$. Since (4.13) implies

$$\mathcal{J}(x^n, w^n, \varepsilon_n) - N\varepsilon_n \leq \|x^n\|_{\ell_1} \leq \mathcal{J}(x^n, w^n, \varepsilon_n), \tag{5.15}$$

we obtain $\|x^n\|_{\ell_1} \to \|x^*\|_{\ell_1}$. Finally, we invoke Lemma 4.2 with $z' = x^n$, $z = x^*$, and $L = K$ to get

$$\limsup_{n\to\infty} \|x^n - x^*\|_{\ell_1} \leq \frac{1+\gamma}{1-\gamma}\left(\lim_{n\to\infty} \|x^n\|_{\ell_1} - \|x^*\|_{\ell_1}\right) = 0, \tag{5.16}$$

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

Finally, (5.12) follows from (4.12) of Lemma 4.3 (with $L = K$), and the observation that $\sigma_n(z) \geq \sigma_{n'}(z)$ if $n \leq n'$.

*Case* $\varepsilon > 0$. We shall first show that $x^n \to x^\varepsilon$, with $x^\varepsilon$ as defined by (5.6). As above, 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 $w_j^n = [(x_j^n)^2 + \varepsilon_n^2]^{-1/2} \leq \varepsilon^{-1}$, it follows that

$$\lim_{i\to\infty} w_j^{n_i} = [(\widetilde{x}_j)^2 + \varepsilon^2]^{-1/2} = \widetilde{w}(\widetilde{x}, \varepsilon)_j =: \widetilde{w}_j \in (0,\infty), \quad j = 1, \dots, N,$$

where the notation $\widetilde{w}(\cdot,\cdot)$ is as in Lemma 5.2. On the other hand, by invoking Lemma 5.1, we now find that $x^{n_i+1} \to \widetilde{x}$ as $i \to \infty$. It then follows from the orthogonality relations (2.6) that for every $i$ and every $\eta \in \mathcal{N}$, we have $\langle x^{n_i+1}, \eta\rangle_{w^{n_i}} = 0$, and therefore

$$\langle \widetilde{x}, \eta\rangle_{\widetilde{w}} = \lim_{i\to\infty} \langle x^{n_i+1}, \eta\rangle_{w^{n_i}} = 0 \quad \forall \eta \in \mathcal{N}. \tag{5.17}$$

Now the "if" part of Lemma 5.2 implies that $\widetilde{x} = x^\varepsilon$, the unique minimizer of $f_\varepsilon$ on $\mathcal{F}(y)$. Hence $(x^n)_{n\in\mathbb{N}}$ has a unique accumulation point and its limit, which we then denote by $\bar{x}$, is equal to $x^\varepsilon$. This establishes (ii).

To prove the error estimate (5.13) stated in (iii), we first note that for any $z \in \mathcal{F}(y)$, we have

$$\|x^\varepsilon\|_{\ell_1} \leq f_\varepsilon(x^\varepsilon) \leq f_\varepsilon(z) \leq \|z\|_{\ell_1} + N\varepsilon, \tag{5.18}$$

where the second inequality uses the minimizing property of $x^\varepsilon$. Hence it follows that $\|x^\varepsilon\|_{\ell_1} - \|z\|_{\ell_1} \le N\varepsilon$. We now invoke Lemma 4.2 to obtain

$$\|x^\varepsilon - z\|_{\ell_1} \le \frac{1+\gamma}{1-\gamma}[N\varepsilon + 2\sigma_k(z)_{\ell_1}], \quad k \le K. \tag{5.19}$$

From Lemma 4.1 and (1.7), we obtain

$$N\varepsilon = \lim_{n\to\infty} N\varepsilon_n \le \lim_{n\to\infty} r(x^n)_{K+1} = r(x^\varepsilon)_{K+1}. \tag{5.20}$$

It follows from (4.4) that

$$\begin{aligned}(K+1-k)N\varepsilon &\le (K+1-k)r(x^\varepsilon)_{K+1} \\ &\le \|x^\varepsilon - z\|_{\ell_1} + \sigma_k(z)_{\ell_1} \\ &\le \frac{1+\gamma}{1-\gamma}[N\varepsilon + 2\sigma_k(z)_{\ell_1}] + \sigma_k(z)_{\ell_1},\end{aligned} \tag{5.21}$$

where the last inequality uses (5.19). Since by our assumption we have $K - k > \frac{2\gamma}{1-\gamma}$, i.e., $K + 1 - k > \frac{1+\gamma}{1-\gamma}$, the inequality (5.21) yields

$$N\varepsilon \le \frac{3 + 4\gamma/(1-\gamma)}{(K-k) - 2\gamma/(1-\gamma)}\,\sigma_k(z)_{\ell_1}. \tag{5.22}$$

Using this back in (5.19), we arrive at (5.13).

Finally, notice that if $\mathcal{F}(y)$ contains a $k$-sparse vector $z$ with $k < K - \frac{2\gamma}{1-\gamma}$ (which, as shown before, must be the unique $\ell_1$-minimizer $x^*$), then (5.22) contradicts the assumption that $\varepsilon > 0$, since $\sigma_k(z)_{\ell_1} = 0$. Therefore the presence of a $k$-sparse solution implies that $\varepsilon = 0$. This finishes the proof. □

*Remark* 5.4. Let us briefly compare our analysis of the IRLS algorithm with $\ell_1$-minimization. The latter recovers a $k$-sparse solution (when one exists) if $\Phi$ has the NSP of order $K$ and $k \le K$. The analysis given in our proof of Theorem 5.3 guarantees that our IRLS algorithm recovers $k$-sparse $x$ for a slightly smaller range of values $k$ than $\ell_1$-minimization, namely for $k < K - \frac{2\gamma}{1-\gamma}$. Notice that this "gap" vanishes as $\gamma$ is decreased. Although we have no examples to demonstrate, our arguments cannot exclude the case where $\mathcal{F}(y)$ contains a $k$-sparse vector $x^*$ with $K - \frac{2\gamma}{1-\gamma} \le k \le K$ (e.g., if $\gamma \ge \frac{1}{3}$ and $k = K - 1$), and our IRLS algorithm converges to $\bar{x}$, yet $\bar{x} \ne x^*$. However, note that unless $\gamma$ is close to 1, the range of $k$-values in this "gap" is fairly small; for instance, for $\gamma < \frac{1}{3}$, this nonrecovery of a $k$-sparse $x^*$ can happen only if $k = K$.

*Remark* 5.5. The constant $c$ in (5.12) is clearly smaller than the constant $\widetilde{c}$ in (5.13); it follows that when $k < K - \frac{2\gamma}{1-\gamma}$, estimate (5.13) holds for all cases, whether $\varepsilon = 0$ or not. The constant $\widetilde{c}$ can be quite reasonable; for instance, if $\gamma \le \frac{1}{2}$ and $k \le K - 3$, then we have $\widetilde{c} \le 9\,\frac{1+\gamma}{1-\gamma} \le 27$.

# 6 Rate of Convergence

Under the conditions of Theorem 5.3 the algorithm converges to a limit $\overline{x}$; if there is a $k$-sparse vector in $\mathcal{F}(y)$ with $k < K - \frac{2\gamma}{1-\gamma}$, then this limit coincides with that $k$-sparse vector, which is then also automatically the unique $\ell_1$-minimizer $x^*$. In this section our goal is to establish a bound for the rate of convergence in both the sparse and nonsparse cases. In the latter case, the goal is to establish the rate at which $x^n$ approaches a ball of radius $C_1\sigma_k(x^*)_{\ell^1}$ centered at $x^*$. We shall work under the same assumptions as in Theorem 5.3.

## 6.1 Case of $k$-Sparse Vectors

Let us begin by assuming that $\mathcal{F}(y)$ contains the $k$-sparse vector $x^*$. The algorithm produces the sequence $x^n$, which converges to $x^*$, as established above. Let us denote 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 we set

$$E_n := \|\eta^n\|_{\ell_1}. \tag{6.1}$$

We know that $E_n \to 0$. The following theorem gives a bound on its rate of convergence.

THEOREM 6.1 *Assume that $\Phi$ satisfies the NSP of order $K$ with constant $\gamma$ such that $0 < \gamma < 1 - \frac{2}{K+2}$, and that $\mathcal{F}(y)$ contains a $k$-sparse vector $x^*$ with $k < K - \frac{2\gamma}{1-\gamma}$. Let $0 < \rho < 1$ be a given constant. Assume, in addition, that $K$, $k$, $\gamma$, and $\rho$ are such that*

$$\mu := \frac{\gamma(1+\gamma)}{1-\rho}\left(1 + \frac{1}{K+1-k}\right) < 1. \tag{6.2}$$

*Let $T = \operatorname{supp}(x^*)$ and $n_0$ be such that*

$$E_{n_0} \le R^* := \rho \min_{i \in T} |x_i^*|. \tag{6.3}$$

*Then for all $n \ge n_0$, we have*

$$E_{n+1} \le \mu E_n. \tag{6.4}$$

*Consequently, $x^n$ converges to $x^*$ exponentially.*

*Remark* 6.2. Notice that if $\gamma$ is sufficiently small, e.g., $\gamma(1+\gamma) < \frac{2}{3}$, then for any $k < K$, there is a $\rho > 0$ for which $\mu < 1$, so we have exponential convergence to $x^*$ whenever $x^*$ is $k$-sparse.

PROOF: We start with relation (2.6) with $w = w^n$, $x^w = x^{n+1} = x^* + \eta^{n+1}$, and $\eta = x^{n+1} - x^* = \eta^{n+1}$, which gives

$$\sum_{i=1}^{N} (x_i^* + \eta_i^{n+1})\eta_i^{n+1} w_i^n = 0. \tag{6.5}$$

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

$$\sum_{i=1}^{N} |\eta_i^{n+1}|^2 w_i^n = -\sum_{i \in T} x_i^* \eta_i^{n+1} w_i^n = -\sum_{i \in T} \frac{x_i^*}{[(x_i^n)^2 + \varepsilon_n^2]^{1/2}} \eta_i^{n+1}. \tag{6.6}$$

We will prove the theorem by induction. Let us assume that we have shown $E_n \le R^*$ already. We then have, for all $i \in T$,

$$|\eta_i^n| \le \|\eta^n\|_{\ell_1^N} = E_n \le \rho |x_i^*|, \tag{6.7}$$

so that

$$\frac{|x_i^*|}{[(x_i^n)^2 + \varepsilon_n^2]^{1/2}} \le \frac{|x_i^*|}{|x_i^n|} = \frac{|x_i^*|}{|x_i^* + \eta_i^n|} \le \frac{1}{1-\rho}, \tag{6.8}$$

and hence (6.6) combined with (6.8) and the NSP gives

$$\sum_{i=1}^{N} |\eta_i^{n+1}|^2 w_i^n \le \frac{1}{1-\rho} \|\eta_T^{n+1}\|_{\ell_1} \le \frac{\gamma}{1-\rho} \|\eta_{T^c}^{n+1}\|_{\ell_1}. \tag{6.9}$$

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

$$\begin{aligned}
\|\eta_{T^c}^{n+1}\|_{\ell_1}^2 &\le \Bigg(\sum_{i \in T^c} |\eta_i^{n+1}|^2 w_i^n\Bigg)\Bigg(\sum_{i \in T^c} [(x_i^n)^2 + \varepsilon_n^2]^{1/2}\Bigg) \\
&\le \Bigg(\sum_{i=1}^{N} |\eta_i^{n+1}|^2 w_i^n\Bigg)\Bigg(\sum_{i \in T^c} [(\eta_i^n)^2 + \varepsilon_n^2]^{1/2}\Bigg) \\
&\le \frac{\gamma}{1-\rho} \|\eta_{T^c}^{n+1}\|_{\ell_1} (\|\eta^n\|_{\ell_1} + N\varepsilon_n). \qquad (6.10)
\end{aligned}$$

If $\eta_{T^c}^{n+1} = 0$, then $x_{T^c}^{n+1} = 0$. In this case $x^{n+1}$ is $k$-sparse (with support contained in $T$) 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 $\eta_{T^c}^{n+1} \ne 0$, then after canceling the factor $\|\eta_{T^c}^{n+1}\|_{\ell_1}$ in (6.10), we get

$$\|\eta_{T^c}^{n+1}\|_{\ell_1} \le \frac{\gamma}{1-\rho} (\|\eta^n\|_{\ell_1} + N\varepsilon_n), \tag{6.11}$$

and thus

$$\begin{aligned}\|\eta^{n+1}\|_{\ell_1} &= \|\eta^{n+1}_T\|_{\ell_1} + \|\eta^{n+1}_{T^c}\|_{\ell_1}\\ &\le (1+\gamma)\|\eta^{n+1}_{T^c}\|_{\ell_1}\\ &\le \frac{\gamma(1+\gamma)}{1-\rho}(\|\eta^n\|_{\ell_1} + N\varepsilon_n).\end{aligned} \tag{6.12}$$

Now, we also have by (1.7) and (4.4)

$$\begin{aligned}N\varepsilon_n \le r(x^n)_{K+1} &\le \frac{1}{K+1-k}(\|x^n - x^*\|_{\ell_1} + \sigma_k(x^*)_{\ell_1})\\ &= \frac{\|\eta^n\|_{\ell_1}}{K+1-k},\end{aligned} \tag{6.13}$$

since by assumption $\sigma_k(x^*) = 0$. This, together with (6.12), yields the desired bound

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

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

*Remark* 6.3. Note that the precise update rule (1.7) for $\varepsilon_n$ does not really intervene in this analysis. If $E_{n_0} \le R^*$, then the estimate

$$E_{n+1} \le \mu_0(E_n + N\varepsilon_n) \quad \text{with } \mu_0 := \frac{\gamma(1+\gamma)}{1-\rho} \tag{6.14}$$

guarantees that all further $E_n$ will be bounded by $R^*$ as well, provided $N\varepsilon_n \le (\mu_0^{-1} - 1)R^*$. It is only in guaranteeing that (6.3) must be satisfied for some $n_0$ that the update rule plays a role: indeed, by Theorem 5.3, $E_n \to 0$ for $n \to \infty$ if $\varepsilon_n$ is updated following (1.7), so that (6.3) has to be satisfied eventually.

Other update rules may work as well. If $(\varepsilon_n)_{n\in\mathbb{N}}$ is defined so that it is a monotonically decreasing sequence with limit $\varepsilon$, then the relation (6.14) immediately implies that

$$\limsup_{n\to\infty} E_n \le \frac{\mu_0 N\varepsilon}{1-\mu_0}. \tag{6.15}$$

In particular, if $\varepsilon = 0$, then $E_n \to 0$. The rate at which $E_n \to 0$ in this case will depend on $\mu_0$ as well as on the rate with which $\varepsilon_n \to 0$. We shall not quantify this relation, except to note that if $\varepsilon_n = O(\beta^n)$ for some $\beta < 1$, then $E_n = O(n\widetilde{\mu}^n)$ where $\widetilde{\mu} = \max(\mu_0, \beta)$.

### 6.2 Case of Noisy $k$-Sparse Vectors

We show here that the exponential rate of convergence to a $k$-sparse limit vector can be extended to the case where the "ideal" (i.e., $k$-sparse) target vector has been corrupted by noise and is therefore only "approximately $k$-sparse." More precisely, we no longer assume that $\mathcal{F}(y)$ contains a $k$-sparse vector; consequently, the limit $\overline{x}$ of the $x^n$ need not be an $\ell_1$-minimizer (see Theorem 5.3). If $x$ is any $\ell_1$-minimizer in $\mathcal{F}(y)$, Theorem 5.3 guarantees $\|\overline{x} - x\|_{\ell_1} \le C\sigma_k(x)_{\ell_1}$; since this is the best level of accuracy guaranteed in the limit, we are in this case interested only in how fast $x^n$ will converge to a ball centered at $x$ with radius given by some (prearranged) multiple of $\sigma_k(x)_{\ell_1}$. (Note that if $\mathcal{F}(y)$ contains several $\ell_1$-minimizers, they all lie within a distance $C'\sigma_k(x)_{\ell_1}$ of each other, so that it does not matter which $x$ we pick.) We shall express the notion that $z$ is "approximately $k$-sparse with gap ratio $C$" or a "noisy version of a $k$-sparse vector with gap ratio $C$" by the condition

$$r(z)_k \ge C\sigma_k(z)_{\ell_1}, \tag{6.16}$$

where $k$ is such that $\Phi$ has the NSP for some pair $K, \gamma$ such that $0 \le k < K - \frac{2\gamma}{1-\gamma}$ (e.g., we could have $K = k + 1$ if $\gamma < \frac{1}{2}$). If the gap ratio $C$ is much greater than the constant $\widetilde{c}$ in (5.13), then exponential convergence can be exhibited for a meaningful number of iterations. Note that this class includes perturbations of any $k$-sparse vector for which the perturbation is sufficiently small in the $\ell^1$-norm (when compared to the unperturbed $k$-sparse vector).

Our argument for the noisy case will closely resemble the case for the exact $k$-sparse vectors. However, there are some crucial differences that justify our decision to separate these two cases.

We will be interested in only the case $\varepsilon > 0$ where we recall that $\varepsilon$ is the limit of the $\varepsilon_n$ occurring in the algorithm. This assumption implies $\sigma_k(x)_{\ell_1} > 0$, and can only happen if $x$ is not $K$-sparse. (As noted earlier, the exact $k$-sparse case always corresponds to $\varepsilon = 0$ if $k < K - \frac{2\gamma}{1-\gamma}$. For $k$ in the region $K - \frac{2\gamma}{1-\gamma} \le k \le K$, both $\varepsilon = 0$ and $\varepsilon > 0$ are theoretical possibilities.)

First, we redefine

$$\eta^n := x^n - x^\varepsilon,$$

where $x^\varepsilon$ is the minimizer of $f_\varepsilon$ on $\mathcal{F}(y)$ and $\varepsilon > 0$. We know from Theorem 5.3 that $\eta^n \to 0$. We again set $E_n := \|\eta^n\|_{\ell_1}$.

THEOREM 6.4 *Given* $0 < \rho < 1$ *and integers* $k$ *and* $K$ *with* $k < K$, *assume that* $\Phi$ *satisfies the NSP of order* $K$ *with constant* $\gamma$ *such that all the conditions of Theorem* 5.3 *are satisfied and, in addition,*

$$\mu := \frac{\gamma(1+\gamma)}{1-\rho}\left(1 + \frac{1}{K+1-k}\right) < 1. \tag{6.17}$$

*Suppose $z \in \mathcal{F}(y)$ is "approximately $k$-sparse with gap ratio $C$", i.e.,*

$$r(z)_k \geq C\sigma_k(z)_{\ell_1} \tag{6.18}$$

*with $C > \widetilde{c}$, where $\widetilde{c}$ is as in Theorem 5.3. Let $T$ stand for the set of indices of the $k$ largest entries of $x^\varepsilon$, and $n_0$ be such that*

$$E_{n_0} \leq R^* := \rho \min_{i \in T} |x_i^\varepsilon| = \rho r(x^\varepsilon)_k. \tag{6.19}$$

*Then for all $n \geq n_0$, we have*

$$E_{n+1} \leq \mu E_n + B\sigma_k(z)_{\ell_1}, \tag{6.20}$$

*where $B > 0$ is a constant. Similarly, if we define $\widetilde{E}_n := \|x^n - z\|_{\ell_1}$, then*

$$\widetilde{E}_{n+1} \leq \mu \widetilde{E}_n + \widetilde{B}\sigma_k(z)_{\ell_1} \tag{6.21}$$

*for $n \geq n_0$, where $\widetilde{B} > 0$ is a constant. This implies that $x^n$ converges at an exponential (linear) rate to the ball of radius $\widetilde{B}(1-\mu)^{-1}\sigma_k(z)_{\ell_1}$ centered at $z$.*

*Remark* 6.5. Note that Theorem 5.3 trivially implies the inequalities (6.20) and (6.21) in the limit $n \to \infty$ since $E_n \to 0$, $\sigma_k(z)_{\ell_1} > 0$, and $\|\bar{x} - z\|_{\ell_1} \leq \widetilde{c}\sigma_k(z)_{\ell_1}$. However, Theorem 6.4 quantifies the event when it is guaranteed that the two measures of error, $E_n$ and $\widetilde{E}_n$, must shrink (at least) by a factor $\mu < 1$ at each iteration. As noted above, this corresponds to the range $\sigma_k(z)_{\ell_1} \lesssim E_n, \widetilde{E}_n \lesssim r(x^\varepsilon)_k$, and would be realized if, say, $z$ is the sum of a $k$-sparse vector and a fully supported "noise" vector that is sufficiently small in the $\ell_1$-norm. In this sense, the theorem shows that the rate estimate of Theorem 5.3 extends to a neighborhood of $k$-sparse vectors.

PROOF: First, note that the existence of $n_0$ is guaranteed by the fact that $E_n \to 0$ and $R^* > 0$. For the latter, note that Lemma 4.1 and Theorem 5.3 imply

$$r(x^\varepsilon)_k \geq r(z)_k - \|z - x^\varepsilon\|_{\ell_1} \geq (C - \widetilde{c})\sigma_k(z)_{\ell_1}, \tag{6.22}$$

so that $R^* \geq \rho(C - \widetilde{c})\sigma_k(z)_{\ell_1} > 0$.

We follow the proof of Theorem 6.1 and consider the orthogonality relation (6.6). Since $x^\varepsilon$ is not sparse in general, we rewrite (6.6) as

$$\sum_{i=1}^N |\eta_i^{n+1}|^2 w_i^n = -\sum_{i=1}^N x_i^\varepsilon \eta_i^{n+1} w_i^n = -\sum_{i \in T \cup T^c} \frac{x_i^\varepsilon}{[(x_i^n)^2 + \varepsilon_n^2]^{1/2}} \eta_i^{n+1}. \tag{6.23}$$

We deal with the contribution on $T$ in the same way as before:

$$\left| \sum_{i \in T} \frac{x_i^\varepsilon}{[(x_i^n)^2 + \varepsilon_n^2]^{1/2}} \eta_i^{n+1} \right| \leq \frac{1}{1-\rho} \|\eta_T^{n+1}\|_{\ell_1} \leq \frac{\gamma}{1-\rho} \|\eta_{T^c}^{n+1}\|_{\ell_1}. \tag{6.24}$$

For the contribution on $T^c$, note that

$$\beta_n := \max_{i \in T^c} \frac{|\eta_i^{n+1}|}{[(x_i^n)^2 + \varepsilon_n^2]^{1/2}} \leq \varepsilon^{-1} \|\eta^{n+1}\|_{\ell_\infty}. \tag{6.25}$$

Since $\eta^n \to 0$, we have $\beta_n \to 0$. It follows that

$$\begin{aligned}\left|\sum_{i\in T^c} \frac{x_i^\varepsilon}{[(x_i^n)^2+\varepsilon_n^2]^{1/2}}\, \eta_i^{n+1}\right| &\le \beta_n \sigma_k(x^\varepsilon)_{\ell_1}\\ &\le \beta_n(\sigma_k(z)_{\ell_1} + \|x^\varepsilon - z\|_{\ell_1})\\ &\le C_2\beta_n\sigma_k(z)_{\ell_1}, \end{aligned} \tag{6.26}$$

where the second inequality is due to Lemma 4.1, the last one to Theorem 5.3, and $C_2 = \widetilde{c} + 1$. Combining these two bounds, we get

$$\sum_{i=1}^N |\eta_i^{n+1}|^2 w_i^n \le \frac{\gamma}{1-\rho}\|\eta_{T^c}^{n+1}\|_{\ell_1} + C_2\beta_n\sigma_k(z)_{\ell_1}. \tag{6.27}$$

We combine this again with a Cauchy-Schwarz estimate to obtain

$$\begin{aligned}\|\eta_{T^c}^{n+1}\|_{\ell_1}^2 &\le \left(\sum_{i\in T^c}|\eta_i^{n+1}|^2 w_i^n\right)\left(\sum_{i\in T^c}[(x_i^n)^2+\varepsilon_n^2]^{1/2}\right)\\ &\le \left(\sum_{i=1}^N|\eta_i^{n+1}|^2 w_i^n\right)\left(\sum_{i\in T^c}[|\eta_i^n| + |x_i^\varepsilon| + \varepsilon_n]\right)\\ &\le \left(\frac{\gamma}{1-\rho}\|\eta_{T^c}^{n+1}\|_{\ell_1} + C_2\beta_n\sigma_k(z)_{\ell_1}\right)(\|\eta_{T^c}^n\|_{\ell_1} + \sigma_k(x^\varepsilon)_{\ell_1} + N\varepsilon_n)\\ &\le \left(\frac{\gamma}{1-\rho}\|\eta_{T^c}^{n+1}\|_{\ell_1} + C_2\beta_n\sigma_k(z)_{\ell_1}\right)(\|\eta_{T^c}^n\|_{\ell_1} + C_2\sigma_k(z)_{\ell_1} + N\varepsilon_n). \end{aligned} \tag{6.28}$$

It is easy to check that if $u^2 \le Au + B$, where $A$ and $B$ are positive, then $u \le A + B/A$. Applying this to $u = \|\eta_{T^c}^{n+1}\|_{\ell_1}$ in the above estimate, we get

$$\|\eta_{T^c}^{n+1}\|_{\ell_1} \le \frac{\gamma}{1-\rho}\,[\|\eta_{T^c}^n\|_{\ell_1} + C_2\sigma_k(z)_{\ell_1} + N\varepsilon_n] + C_3\beta_n\sigma_k(z)_{\ell_1}, \tag{6.29}$$

where $C_3 = C_2(1-\rho)/\gamma$. Similarly to (6.13), we also have, by combining (4.4) with (part of) the chain of inequalities (6.26),

$$\begin{aligned}N\varepsilon_n &\le r(x^n)_{K+1}\\ &\le \frac{1}{K+1-k}(\|x^n - x^\varepsilon\|_{\ell_1} + \sigma_k(x^\varepsilon)_{\ell_1})\\ &\le \frac{1}{K+1-k}(\|\eta^n\|_{\ell_1} + C_2\sigma_k(z)_{\ell_1}), \end{aligned} \tag{6.30}$$

and consequently (6.29) yields

$$
\begin{aligned}
\|\eta^{n+1}\|_{\ell_1} &\le (1+\gamma)\|\eta^{n+1}_{T^c}\|_{\ell_1} \\
&\le \frac{\gamma(1+\gamma)}{1-\rho}\left(1+\frac{1}{K+1-k}\right)\|\eta^n\|_{\ell_1} \\
&\quad + (1+\gamma)(C_3\beta_n + C_4)\,\sigma_k(z)_{\ell_1}
\end{aligned} \tag{6.31}
$$

where $C_4 = C_2\gamma(1-\rho)^{-1}(1+1/(K+1-k))$. Since the $\beta_n$ are bounded, this gives

$$
E_{n+1} \le \mu E_n + B\sigma_k(z)_{\ell_1}. \tag{6.32}
$$

It then follows that if we pick $\widetilde{\mu}$ so that $1 > \widetilde{\mu} > \mu$, and consider the range of $n > n_0$ such that $E_n \ge (\widetilde{\mu}-\mu)^{-1}B\sigma_k(z)_{\ell_1} =: r^*$, then

$$
E_{n+1} \le \widetilde{\mu} E_n. \tag{6.33}
$$

Hence we are guaranteed exponential decay of $E_n$ as long as $x^n$ is sufficiently far from its limit. The smallest possible value of $r^*$ corresponds to the case $\widetilde{\mu} \approx 1$.

To establish a rate of convergence to a comparably sized ball centered at $z$, we consider $\widetilde{E}_n = \|x^n - z\|_{\ell_1}$. It then follows that

$$
\begin{aligned}
\widetilde{E}_{n+1} &\le \|x^{n+1} - x^\varepsilon\|_{\ell_1} + \|x^\varepsilon - z\|_{\ell_1} \\
&\le \mu\|x^n - x^\varepsilon\|_{\ell_1} + B\sigma_k(z)_{\ell_1} + \widetilde{c}\sigma_k(z)_{\ell_1} \\
&\le \mu\|x^n - z\|_{\ell_1} + B\sigma_k(z)_{\ell_1} + \widetilde{c}(1+\mu)\sigma_k(z)_{\ell_1} \\
&= \mu\widetilde{E}_n + \widetilde{B}\sigma_k(z)_{\ell_1},
\end{aligned} \tag{6.34}
$$

which shows the claimed exponential decay and also that

$$
\limsup_{n\to\infty} \widetilde{E}_n \le \widetilde{B}(1-\mu)^{-1}\sigma_k(z)_{\ell_1}. \qquad \square
$$

## 7 Beyond the Convex Case: $\ell_\tau$-Minimization for $\tau < 1$

If $\Phi$ has the NSP of order $K$ with $\gamma < 1$, then (see Section 3) $\ell_1$-minimization recovers $K$-sparse solutions to $\Phi x = y$ for any $y \in \mathbb{R}^m$ that admits such a sparse solution; i.e., $\ell_1$-minimization also gives $\ell_0$-minimizers, provided their support has size at most $K$. In [31], Gribonval and Nielsen showed that in this case, $\ell_1$-minimization also gives the $\ell_\tau$-minimizers; i.e., $\ell_1$-minimization also solves nonconvex optimization problems of the type

$$
x^* = \operatorname*{argmin}_{z\in\mathcal{F}(y)} \|z\|^\tau_{\ell^N_\tau} \quad \text{for } 0 < \tau < 1. \tag{7.1}
$$

Let us first recall the results of [31] that are of most interest to us here, reformulated for our setting and notation.

LEMMA 7.1 [31, theorem 2] *Assume that* $x^*$ *is a* $K$*-sparse vector in* $\mathcal{F}(y)$ *and that* $0 < \tau \leq 1$. *If*

$$\sum_{i \in T} |\eta_i|^\tau < \sum_{i \in T^c} |\eta_i|^\tau \quad \text{or equivalently} \quad \sum_{i \in T} |\eta_i|^\tau < \frac{1}{2} \sum_{i=1}^{N} |\eta_i|^\tau$$

*for all* $\eta \in \mathcal{N}$ *and for all* $T \subset \{1, \ldots, N\}$ *with* $\#T \leq K$, *then*

$$x^* = \operatorname*{argmin}_{z \in \mathcal{F}(y)} \|z\|_{\ell_\tau^N}^\tau .$$

LEMMA 7.2 [31, theorem 5] *Let* $z \in \mathbb{R}^N$, $0 < \tau_1 \leq \tau_2 \leq 1$, *and* $K \in \mathbb{N}$. *Then*

$$\sup_{\substack{T \subset \{1,\ldots,N\} \\ \#T \leq K}} \frac{\sum_{i \in T} |z_i|^{\tau_1}}{\sum_{i=1}^N |z_i|^{\tau_1}} \leq \sup_{\substack{T \subset \{1,\ldots,N\} \\ \#T \leq K}} \frac{\sum_{i \in T} |z_i|^{\tau_2}}{\sum_{i=1}^N |z_i|^{\tau_2}} .$$

Combining these two lemmas with the observations in Section 3 leads immediately to the following result:

THEOREM 7.3 *Fix any* $0 < \tau \leq 1$. *If* $\Phi$ *satisfies the NSP of order* $K$ *with constant* $\gamma$, *then*

$$\sum_{i \in T} |\eta_i|^\tau < \gamma \sum_{i \in T^c} |\eta_i|^\tau \tag{7.2}$$

*for all* $\eta \in \mathcal{N}$ *and for all* $T \subset \{1, \ldots, N\}$ *such that* $\#T \leq K$.

*In addition, if* $\gamma < 1$, *and if there exists a* $K$*-sparse vector in* $\mathcal{F}(y)$, *then this* $K$*-sparse vector is the unique minimizer in* $\mathcal{F}(y)$ *of* $\|\cdot\|_{\ell_\tau}$.

At first sight, these results suggest there is nothing to be gained by carrying out $\ell_\tau$- rather than $\ell_1$-minimization; in addition, sparse recovery via the nonconvex problem (7.1) is much harder than the more easily solvable convex problem of $\ell_1$-minimization.

Yet we shall show in this section that $\ell_\tau$-minimization has unexpected benefits, and that it may be both useful *and* practically feasible via an IRLS approach. Before we start, it is expedient to introduce the following definition: we shall say that $\Phi$ has the $\tau$*-null space property* ($\tau$-NSP) of order $K$ with constant $\gamma > 0$ if, for all sets $T$ of cardinality at most $K$ and all $\eta \in \mathcal{N}$,

$$\|\eta_T\|_{\ell_\tau^N}^\tau \leq \gamma \|\eta_{T^c}\|_{\ell_\tau^N}^\tau . \tag{7.3}$$

In what follows we shall construct an IRLS algorithm for $\ell_\tau$- minimization. We shall see the following:

(a) In practice, $\ell_\tau$-minimization can be carried out by an IRLS algorithm. Hence the nonconvexity does not necessarily make the problem intractable.

(b) In particular, if $\Phi$ satisfies the $\tau$-NSP of order $K$, and if there exists a $k$-sparse vector $x^*$ in $\mathcal{F}(y)$, with $k \leq K - \kappa$ for suitable $\kappa$ given below, then the IRLS algorithm converges to $\bar{x} = x^*$, which is also the $\ell_\tau$-minimizer.

(c) Surprisingly, the rate of local convergence of the algorithm is superlinear; the rate is faster for smaller $\tau$, and increases towards quadratic as $\tau \to 0$. More precisely, we will show that the local error $E_n := \|x^n - x^*\|_{\ell_\tau^N}^\tau$ satisfies

$$E_{n+1} \le \mu(\gamma, \tau) E_n^{2-\tau}, \tag{7.4}$$

where $\mu(\gamma, \tau) < 1$ for $\gamma > 0$ sufficiently small. As before, the validity of (7.4) requires that $x^n$ be in a neighborhood of $x^*$. In particular, if $x^0$ is close enough to $x^*$, then (7.4) ensures the convergence of the algorithm to the $k$-sparse solution $x^*$.

Some of the virtues of $\ell_\tau$-minimization were recently highlighted in a series of papers [12, 13, 14, 38]. Chartrand and Staneva [14] give a fine analysis of the RIP from which they conclude not only that $\ell_\tau$-minimization recovers $k$-sparse vectors, but also that the range of $k$ for which this recovery works is larger for smaller $\tau$. Namely, for random Gaussian matrices, they prove that with high probability on the draw of the matrix, sparse recovery by $\ell_\tau$-minimization works for

$$k \le \frac{m}{c_1(\tau) + \tau c_2(\tau)\log(N/k)},$$

where $c_1(\tau)$ is bounded and $c_2(\tau)$ decreases to 0 as $\tau \to 0$. In particular, the dependence of the sparsity $k$ on the number $N$ of columns vanishes for $\tau \to 0$. These bounds give a quantitative estimate of the improvement provided by $\ell_\tau$-minimization vis a vis $\ell_1$-minimization for which the range of $k$-sparsity for having exact recovery is clearly smaller (see Figure 8.4 for a numerical illustration).

### 7.1 Some Useful Properties of $\ell_\tau$-Spaces

We start by listing in one proposition some fundamental and well-known properties of $\ell_\tau$-spaces for $0 < \tau \le 1$. For further details we refer the reader to, e.g., [20].

PROPOSITION 7.4

(i) *Assume* $0 < \tau \le 1$. *Then the map* $z \mapsto \|z\|_{\ell_\tau^N}$ *defines a quasi-norm for* $\mathbb{R}^N$*; in particular, the triangle inequality holds up to a constant, i.e.,*

$$\|u + v\|_{\ell_\tau^N} \le C(\tau)(\|u\|_{\ell_\tau^N} + \|v\|_{\ell_\tau^N}) \quad \textit{for all } u, v \in \mathbb{R}^N. \tag{7.5}$$

*If one considers the* $\tau^{th}$ *powers of the "*$\tau$*-norm," then one has the so-called* $\tau$*-triangle inequality*:

$$\|u + v\|_{\ell_\tau^N}^\tau \le \|u\|_{\ell_\tau^N}^\tau + \|v\|_{\ell_\tau^N}^\tau \quad \textit{for all } u, v \in \mathbb{R}^N. \tag{7.6}$$

(ii) *We have, for any* $0 < \tau_1 \le \tau_2 \le \infty$,

$$\|u\|_{\ell_{\tau_2}^N} \le \|u\|_{\ell_{\tau_1}^N} \quad \textit{for all } u \in \mathbb{R}^N. \tag{7.7}$$

(iii) (*Generalized Hölder inequality*) *For* $0 < \tau \leq 1$ *and* $0 < p, q < \infty$ *such that* $\frac{1}{\tau} = \frac{1}{p} + \frac{1}{q}$*, and for a positive weight vector* $w = (w_i)_{i=1}^N$*, we have*

$$\|(u_i v_i)_{i=1}^N\|_{\ell_\tau^N(w)} \leq \|u\|_{\ell_p^N(w)} \|v\|_{\ell_q^N(w)} \quad \text{for all } u, v \in \mathbb{R}^N, \tag{7.8}$$

*where* $\|v\|_{\ell_r^N(w)} := (\sum_{i=1}^N |v_i|^r w_i)^{1/r}$*, as usual, for* $0 < r < \infty$.

For technical reasons, it is often more convenient to employ the $\tau$-triangle inequality (7.6) than (7.5); in this sense, for $\ell_\tau$-minimization $\|\cdot\|_{\ell_\tau^N}^\tau$ turns out to be more natural as a measure of error than the quasi-norm $\|\cdot\|_{\ell_\tau^N}$.

In order to prove the three claims (a)–(c) listed before the start of this subsection, we also need to generalize certain results previously shown only in the $\ell_1$ setting to the $\ell_\tau$ setting. In the following we assume $0 < \tau \leq 1$. We denote by

$$\sigma_k(z)_{\ell_\tau^N} := \sum_{\nu > k} r(z)_\nu^\tau$$

the error of the best $k$-term approximation to $z$ with respect to $\|\cdot\|_{\ell_\tau^N}^\tau$. As a straightforward generalization of analogous results valid for the $\ell_1$-norm, we have the following two technical lemmas:

LEMMA 7.5 *For any* $j \in \{1, \dots, N\}$*, we have*

$$|\sigma_j(z)_{\ell_\tau^N} - \sigma_j(z')_{\ell_\tau^N}| \leq \|z - z'\|_{\ell_\tau^N}^\tau$$

*for all* $z, z' \in \mathbb{R}^N$*. Moreover, for any* $J > j$*, we have*

$$(J - j) r(z)_J^\tau \leq \sigma_j(z)_{\ell_\tau^N} \leq \|z - z'\|_{\ell_\tau^N}^\tau + \sigma_j(z')_{\ell_\tau^N}.$$

LEMMA 7.6 *Assume that* $\Phi$ *has the* $\tau$*-NSP of order* $K$ *with constant* $0 < \gamma < 1$*. Then, for any* $z, z' \in \mathcal{F}(y)$*, we have*

$$\|z' - z\|_{\ell_\tau^N}^\tau \leq \frac{1 + \gamma}{1 - \gamma} (\|z'\|_{\ell_\tau^N}^\tau - \|z\|_{\ell_\tau^N}^\tau + 2\sigma_K(z)_{\ell_\tau^N}).$$

The proofs of these lemmas are essentially identical to the ones of Lemma 4.1 and Lemma 4.2 except for substituting $\|\cdot\|_{\ell_\tau^N}^\tau$ for $\|\cdot\|_{\ell_1^N}$ and $\sigma_k(\cdot)_{\ell_\tau^N}$ for $\sigma_k(\cdot)_{\ell_1^N}$, respectively.

## 7.2 An IRLS Algorithm for $\ell_\tau$-Minimization

To define an IRLS algorithm promoting $\ell_\tau$-minimization for a generic $0 < \tau \leq 1$, we first define a $\tau$-dependent functional $\mathcal{J}_\tau$, generalizing $\mathcal{J}$:

$$\mathcal{J}_\tau(z, w, \varepsilon) := \frac{\tau}{2} \left[ \sum_{j=1}^N z_j^2 w_j + \sum_{j=1}^N \left( \varepsilon^2 w_j + \frac{2 - \tau}{\tau} \frac{1}{w_j^{\frac{\tau}{2-\tau}}} \right) \right], \tag{7.9}$$

where $z \in \mathbb{R}^N$, $w \in \mathbb{R}_+^N$, and $\varepsilon \in \mathbb{R}_+$. The desired algorithm, which we shall call Algorithm $\tau$, is then defined simply by substituting $\mathcal{J}_\tau$ for $\mathcal{J}$ in Algorithm 1, while keeping the same update rule (1.7) for $\varepsilon_n$. In particular, we have

$$w_j^{n+1} = ((x_j^{n+1})^2 + \varepsilon_{n+1}^2)^{-\frac{2-\tau}{2}}, \quad j = 1, \dots, N,$$

and

$$\mathcal{J}_\tau(x^{n+1}, w^{n+1}, \varepsilon_{n+1}) = \sum_{j=1}^{N} ((x_j^{n+1})^2 + \varepsilon_{n+1}^2)^{\frac{\tau}{2}}.$$

Fundamental properties of Algorithm $\tau$ are derived in the same way as before. In particular, the values $\mathcal{J}_\tau(x^n, w^n, \varepsilon_n)$ decrease monotonically,

$$\mathcal{J}_\tau(x^{n+1}, w^{n+1}, \varepsilon_{n+1}) \le \mathcal{J}_\tau(x^n, w^n, \varepsilon_n), \quad n \ge 0,$$

and the iterates are bounded,

$$\|x^n\|_{\ell_\tau^N}^\tau \le \mathcal{J}_\tau(x^1, w^0, \varepsilon_0) =: A_0.$$

As in Lemma 4.4, the weights are uniformly bounded from below, i.e.,

$$w_j^n \ge \widetilde{A}_0 > 0, \quad n \ge 1, \ j = 1, \dots, N.$$

Moreover, using $\mathcal{J}_\tau$ for $\mathcal{J}$ in Lemma 5.1, we can again prove the asymptotic regularity of the iterations, i.e.,

$$\lim_{n\to\infty} \|x^{n+1} - x^n\|_{\ell_2^N} = 0.$$

The first significant difference with the $\ell_1$-case arises when $\varepsilon = \lim_{n\to\infty} \varepsilon_n > 0$. In this latter situation, we need to consider the function

$$f_{\varepsilon,\tau}(z) := \sum_{j=1}^{N} (z_j^2 + \varepsilon^2)^{\frac{\tau}{2}}. \tag{7.10}$$

We denote by $\mathcal{Z}_{\varepsilon,\tau}(y)$ its set of minimizers on $\mathcal{F}(y)$ (since $f_{\varepsilon,\tau}$ is no longer convex it may have more than one minimizer). Even though every minimizer $z \in \mathcal{Z}_{\varepsilon,\tau}(y)$ still satisfies

$$\langle z, \eta \rangle_{\widetilde{w}(z,\varepsilon,\tau)} = 0 \quad \text{for all } \eta \in \mathcal{N},$$

where $\widetilde{w}(z, \varepsilon, \tau)_j = (z_j^2 + \varepsilon^2)^{(\tau-2)/2}$, $j = 1, \dots, N$, the converse need no longer be true.

The following theorem summarizes the convergence properties of Algorithm $\tau$ in the case $\tau < 1$.

THEOREM 7.7 *Let $K$ (the same parameter as in the update rule (1.7)) be chosen so that $\Phi$ satisfies the $\tau$-NSP of order $K$ with a constant $\gamma$ such that $\gamma < 1 - \frac{2}{K+2}$. For any $y \in \mathbb{R}^m$, let $\overline{\mathcal{Z}}_\tau(y)$ be the set of accumulation points of $(x^n)_{n\in\mathbb{N}}$ produced by Algorithm $\tau$, and define $\varepsilon := \lim_{n\to\infty} \varepsilon_n$. Then the following properties hold:*

(i) *If $\varepsilon = 0$, then $\overline{\mathcal{Z}}_\tau(y)$ consists of a single point $\overline{x}$; i.e., $x^n$ converges to $\overline{x}$, and $\overline{x}$ is an $\ell_\tau$-minimizer in $\mathcal{F}(y)$ that is also $K$-sparse.*

(ii) *If* $\varepsilon > 0$, *then for each* $\overline{x} \in \overline{\overline{\mathcal{Z}}}_\tau(y)$ *we have* $\langle \overline{x}, \eta \rangle_{w(\overline{x},\varepsilon,\tau)} = 0$ *for all* $\eta \in \mathcal{N}$.

(iii) *If* $\varepsilon > 0$ *and* $\overline{x} \in \overline{\overline{\mathcal{Z}}}_\tau(y) \cap \mathcal{Z}_{\varepsilon,\tau}(y)$, *then for each* $z \in \mathcal{F}(y)$ *and any* $k < K - \frac{2\gamma}{1-\gamma}$, *we have*

$$\|z - \overline{x}\|^\tau_{\ell^N_\tau} \leq C_2 \sigma_k(z)_{\ell^N_\tau}.$$

The proof of this theorem uses Lemmas 7.1 through 7.6 and follows the same arguments as for Theorem 5.3.

*Remark* 7.8. Unlike Theorem 5.3, Theorem 7.7 does not ensure that Algorithm $\tau$ converges to the sparsest or to the minimal $\ell_\tau$-norm solution. It does provide sufficient conditions that are verifiable a posteriori (e.g., $\varepsilon = 0$) for such convergence. The reason for this weaker result is the nonconvexity of $f_{\varepsilon,\tau}$. Nevertheless, as is often the case for nonconvex problems, we can establish a local convergence result that also highlights the rate we can expect for such convergence. This is the content of the following section; it will be followed by numerical results that dovetail nicely with the theoretical results.

### 7.3 Local Superlinear Convergence

Throughout this section, we assume that there exists a $k$-sparse vector $x^*$ in $\mathcal{F}(y)$. We define the error vectors

$$\eta^n := x^n - x^* \in \mathcal{N}$$

as before, but now measure the error by

$$E_n := \|\eta^n\|^\tau_{\ell^N_\tau}.$$

THEOREM 7.9 *Assume that* $\Phi$ *has the* $\tau$*-NSP of order* $K$ *with constant* $\gamma \in (0,1)$ *and that* $\mathcal{F}(y)$ *contains a* $k$*-sparse vector* $x^*$ *with* $k \leq K$. *Let* $T = \operatorname{supp}(x^*)$ *and suppose that, for a given* $0 < \rho < 1$, *we have*

$$E_{n_0} \leq R^* := [\rho \min_{i \in T} |x_i^*|]^\tau \tag{7.11}$$

*and define*

$$\mu := \mu(\rho, K, \gamma, \tau, N) := 2^{1-\tau}\gamma(1+\gamma)^{2-\tau} A^\tau \left(1 + \left(\frac{N^{1-\tau}}{K+1-k}\right)^{2-\tau}\right),$$

*where*

$$A := ((1-\rho)^{2-\tau} \min_{i \in T} |x_i^*|^{1-\tau})^{-1}.$$

*If* $\mu$ *and* $R^*$ *are such that*

$$\mu {R^*}^{1-\tau} \leq 1, \tag{7.12}$$

*then for all* $n \geq n_0$ *we have*

$$E_{n+1} \leq \mu E_n^{2-\tau}. \tag{7.13}$$

PROOF: Similarly to Theorem 6.1, the proof uses induction on $n$. We assume that $E_n \le R^*$ and will derive (7.13). The first few steps of the proof are the same as before; in particular, the orthogonality relation (6.6) now reads

$$\sum_{i=1}^{N} |\eta_i^{n+1}|^2 w_i^n = -\sum_{i \in T} \frac{x_i^*}{[(x_i^n)^2 + \varepsilon_n^2]^{1-\tau/2}} \eta_i^{n+1}, \tag{7.14}$$

and we have $|\eta_i^n| \le \rho |x_i^*|$ for all $i \in T$. Next, (6.8) is replaced by the following estimate; we have, for $i \in T$,

$$\frac{|x_i^*|}{[(x_i^n)^2 + \varepsilon_n^2]^{1-\tau/2}} \le \frac{|x_i^*|}{|x_i^* + \eta_i^n|^{2-\tau}} \le \frac{|x_i^*|}{[|x_i^*|(1-\rho)]^{2-\tau}} \le A. \tag{7.15}$$

Combining (7.14) and (7.15), we obtain

$$\sum_{i=1}^{N} |\eta_i^{n+1}|^2 w_i^n \le A \Biggl( \sum_{i \in T} |\eta_i^{n+1}|^\tau \Biggr)^{1/\tau},$$

where we have also used the embedding inequality (7.7). We now apply the $\tau$-NSP to find

$$\|\eta^{n+1}\|_{\ell_2(w^n)}^{2\tau} = \Biggl( \sum_{i=1}^{N} |\eta_i^{n+1}|^2 w_i^n \Biggr)^\tau \le A^\tau \|\eta_T^{n+1}\|_{\ell_\tau^N}^\tau \le \gamma A^\tau \|\eta_{T^c}^{n+1}\|_{\ell_\tau^N}^\tau .$$

At the same time, the generalized Hölder inequality (see Proposition 7.4(iii)) for $p = 2$ and $q = \frac{2\tau}{2-\tau}$, together with the above estimates, yields

$$\begin{aligned} \|\eta_{T^c}^{n+1}\|_{\ell_\tau^N}^{2\tau} &= \|(|\eta_i^{n+1}|(w_i^n)^{-1/\tau})_{i=1}^N\|_{\ell_\tau^N(w^n;T^c)}^{2\tau} \\ &\le \|\eta^{n+1}\|_{\ell_2^N(w^n)}^{2\tau} \|((w_i^n)^{-1/\tau})_{i=1}^N\|_{\ell_{2\tau/(2-\tau)}^N(w^n;T^c)}^{2\tau} \\ &\le \gamma A^\tau \|\eta_{T^c}^{n+1}\|_{\ell_\tau^N}^\tau \|((w_i^n)^{-1/\tau})_{i=1}^N\|_{\ell_{2\tau/(2-\tau)}^N(w^n;T^c)}^{2\tau} . \end{aligned}$$

In other words,

$$\|\eta_{T^c}^{n+1}\|_{\ell_\tau^N}^\tau \le \gamma A^\tau \|((w_i^n)^{-1/\tau})_{i=1}^N\|_{\ell_{2\tau/(2-\tau)}^N(w^n;T^c)}^{2\tau} . \tag{7.16}$$

Let us now estimate the weight term. By the $\frac{\tau}{2}$-triangle inequality (7.6) and the convexity of the function $u \mapsto u^{2-\tau}$, we have

$$\begin{aligned} \|((w_i^n)^{-1/\tau})_{i=1}^N\|_{\ell_{2\tau/(2-\tau)}^N(w^n;T^c)}^{2\tau} &= \Biggl( \sum_{i=1}^{N} (|\eta_i^n|^2 + \varepsilon_n^2)^{\frac{\tau}{2}} \Biggr)^{2-\tau} \\ &\le \Biggl( \sum_{i=1}^{N} (|\eta_i^n|^\tau + \varepsilon_n^\tau) \Biggr)^{2-\tau} \\ &\le 2^{1-\tau} \Biggl( \Biggl( \sum_{i=1}^{N} |\eta_i^n|^\tau \Biggr)^{2-\tau} + N^{2-\tau} \varepsilon_n^{\tau(2-\tau)} \Biggr). \end{aligned}$$

Now, the update rule (1.7) and Lemma 7.5 along with the fact that $x^*$ is $k$-sparse result in the following estimates:

$$\begin{aligned} N^{2-\tau}\varepsilon_n^{\tau(2-\tau)} &= N^{(1-\tau)(2-\tau)}(N^\tau \varepsilon_n^\tau)^{2-\tau} \\ &\le N^{(1-\tau)(2-\tau)}(r(x^n)_{K+1}^\tau)^{2-\tau} \\ &\le \left(\frac{N^{1-\tau}}{K+1-k}(\|x^n - x^*\|_{\ell_\tau^N}^\tau + \sigma_k(x^*)_{\ell_\tau^N})\right)^{2-\tau} \\ &= \left(\frac{N^{1-\tau}}{K+1-k}\right)^{2-\tau}(\|\eta^n\|_{\ell_\tau^N}^\tau)^{2-\tau}. \end{aligned}$$

Using these estimates in (7.16) gives

$$\|\eta_{T^c}^{n+1}\|_{\ell_\tau^N}^\tau \le 2^{1-\tau}\gamma A^\tau \left(1 + \left(\frac{N^{1-\tau}}{K+1-k}\right)^{2-\tau}\right)(\|\eta^n\|_{\ell_\tau^N}^\tau)^{2-\tau},$$

and (7.13) follows by a further application of the $\tau$-NSP (as in the first inequality in (6.12)). Because of the assumption (7.12), we now have $E_{n+1} \le E_n \le R^*$, which completes the induction step. $\square$

*Remark* 7.10. In contrast to the $\ell_1$ case, we do not need $\mu < 1$ to ensure that $E_n$ decreases. In fact, all that is needed for the error reduction is $\mu E_n^{1-\tau} < 1$ for some sufficiently large $n$. In fact, $\mu$ could be quite large in cases where the smallest nonzero component of the sparse vector is very small. We have not observed this effect in our examples; we expect that our analysis, although apparently accurate in describing the rate of convergence (see Section 8), is somewhat pessimistic in estimating the coefficient $\mu$.

## 8 Numerical Results

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

### 8.1 Convergence Rates

We start with numerical results that confirm the linear rate of convergence of our IRLS algorithm for $\ell_1$-minimization and its robust recovery of sparse vectors. In our experiments the matrix $\Phi$ of dimensions $m \times N$ consisted of Gaussian $\mathcal{N}(0, 1/m)$ i.i.d. entries. Such matrices are known to satisfy (with high probability) the RIP with optimal bounds [2, 9, 37]. In Figure 8.1 we depict the error of approximation of the iterates to the unique sparsest solution shown in Figure 8.2, 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.

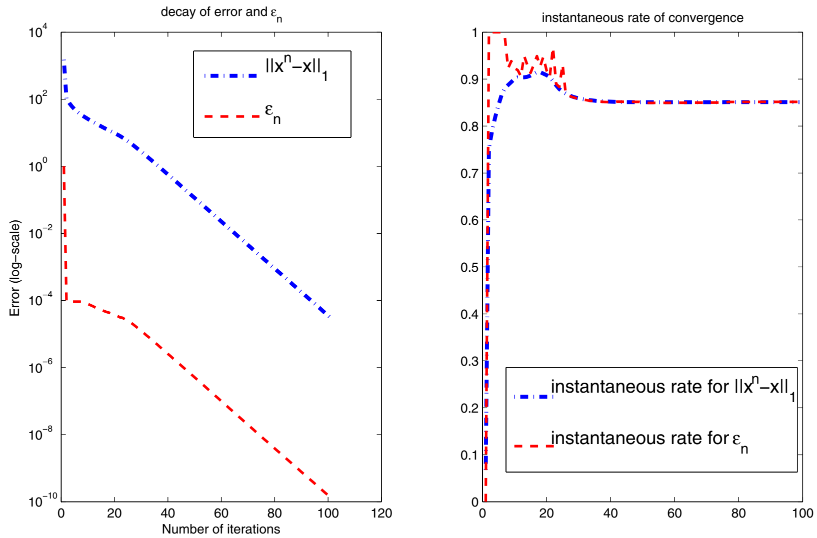


FIGURE 8.1. The results of an experiment in which the entries of a $250 \times 1500$ matrix $\Phi$ were sampled independently from the Gaussian distribution $\mathcal{N}(0, \frac{1}{250})$ and the recovery of a 45-sparse vector $x$ (shown in Figure 8.2) was sought from its image $y = \Phi x$ using the limit of $(x^n)$ produced by Algorithm 1. Left: plot of $\|x^n - x\|_{\ell_1}$ and $\varepsilon_n$ (defined adaptively as in (1.7)) as a function of $n$. Note the logaritmic scale in the ordinate axis. Right: plot of the ratios $\|x^n - x^{n+1}\|_{\ell_1}/\|x^n - x^{n-1}\|_{\ell_1}$, and $(\varepsilon_n - \varepsilon_{n+1})/(\varepsilon_{n-1} - \varepsilon_n)$ for the same experiment.

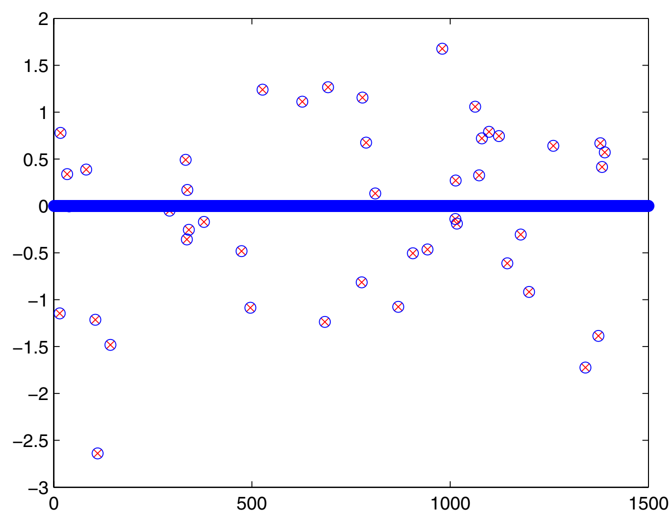

FIGURE 8.2. The sparse vector of dimension 1500 used in the example illustrated in Figure 8.1. This vector has only 45 nonzero entries.

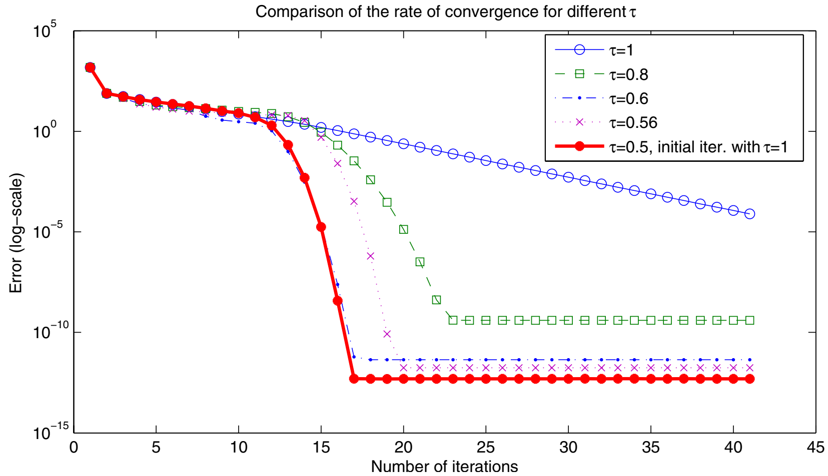


FIGURE 8.3. We show the decay of error as a function of the number of iterations of the algorithm for different values of $\tau$ (1, 0.8, 0.6, 0.56). We also show the results of an experiment in which the initial 10 iterations are performed with $\tau = 1$ and the remaining iterations with $\tau = 0.5$.

Next we compare the linear convergence achieved with the IRLS algorithm for $\ell_1$-minimization with the superlinear convergence of the IRLS algorithm for $\ell_\tau$-minimization with $\tau < 1$. In Figure 8.3 we compare the convergence behavior for different choices of $\tau \in (0, 1]$ where we again plot the approximation error (in logarithmic scale) as a function of $n$. The case $\tau = 1$ shows linear convergence as before. For the smaller values $\tau = 0.8$, 0.6, and 0.56, the error decay initially follows a linear, transient regime; however, once the iterates get sufficiently close to the sparse solution vector, the convergence is seen to speed up dramatically, suggesting that (7.13) takes effect, resulting in superlinear convergence. For smaller values of $\tau$, we have not always observed convergence. (Note that we do not have proof of global convergence.) However, a combination of initial iterations with the $\ell_1$-weighted IRLS (for which we always have convergence) and later iterations with the $\ell_\tau$-weighted IRLS for smaller $\tau$ appeared to allow again for fast convergence to the sparsest solution; this is illustrated in Figure 8.3 for the case $\tau = 0.5$.

## 8.2 Enhanced Recovery in Compressed Sensing and Relationship with Other Work

Candès, Wakin, and Boyd [10] showed, by numerical experimentation, that iteratively reweighted $\ell_1$-minimization, with weights suggested by an $\ell_0$-minimization goal, can enhance the range of sparsity for which perfect reconstruction of sparse vectors is achievable. In experiments with iteratively reweighted $\ell_2$-minimization algorithms, Chartrand and several collaborators observed a similar significant improvement [12, 13, 14, 15, 38]; see in particular [14, sec. 4]. We also illustrate this in Figure 8.4, where we show that the range of $k$ for which exact recovery can be

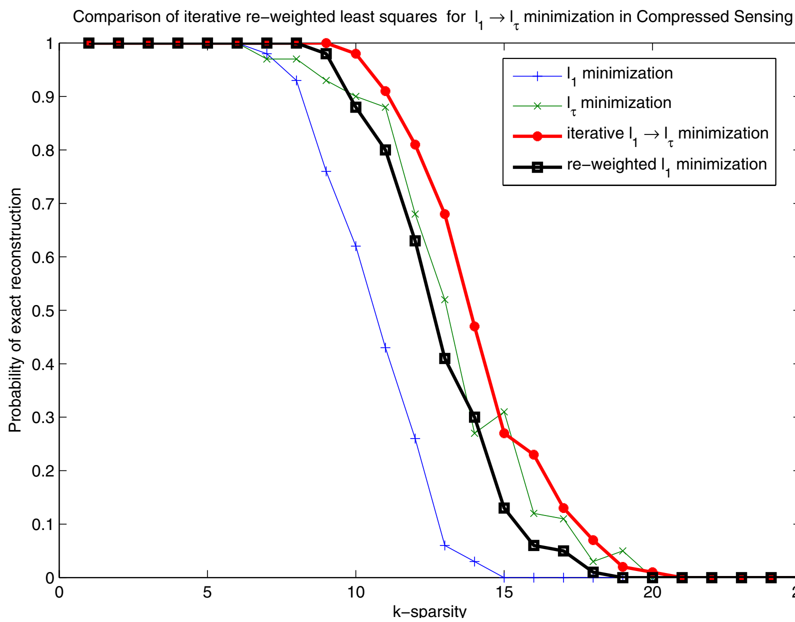


FIGURE 8.4. The (experimentally determined) probability of successful recovery of a sparse 250-dimensional vector $x$, with sparsity $k$, from its image $y = \Phi x$, as a function of $k$. In these experiments the matrix $\Phi$ is $50 \times 250$ dimensional, with i.i.d. Gaussian $\mathcal{N}(0, \frac{1}{50})$ entries. The matrix was generated once; then, for each sparsity value $k$ shown in the plot, 100 attempts were made for randomly generated $k$-sparse vectors $x$. Four different IRLS algorithms were compared: one with weights designed for $\ell_1$-minimization, one with weights designed for $\ell_\tau$-minimization with $\tau = 0.65$, the iteratively reweighted $\ell_1$-minimization algorithm proposed in [10], and a modified IRLS algorithm with weights that gradually moved from an $\ell_1$- to an $\ell_\tau$-minimization goal, with final $\tau = 0.5$. We refer to [14, sec. 4] for similar experiments (for different values of $\tau$), although realized there by fixing the sparsity first and then randomly generating matrices with an increasing number of measurements $m$.

expected is improved for our IRLS algorithm when the weights are chosen gradually moving from an $\ell_1$- to an $\ell_\tau$-minimization goal, with final $\tau = 0.5$. It is also to be noted that in practice IRLS appears to be computationally less demanding than weighted $\ell_1$-minimization; we report our findings in Figure 8.5, where a comparison of the CPU times is done by using standard packages without any specific optimization. As far as we know, there is (yet) no available analysis for reweighted $\ell_1$-minimization that is comparable to the analysis of convergence presented here of our IRLS algorithm, giving a realistic picture of the numerical computations.

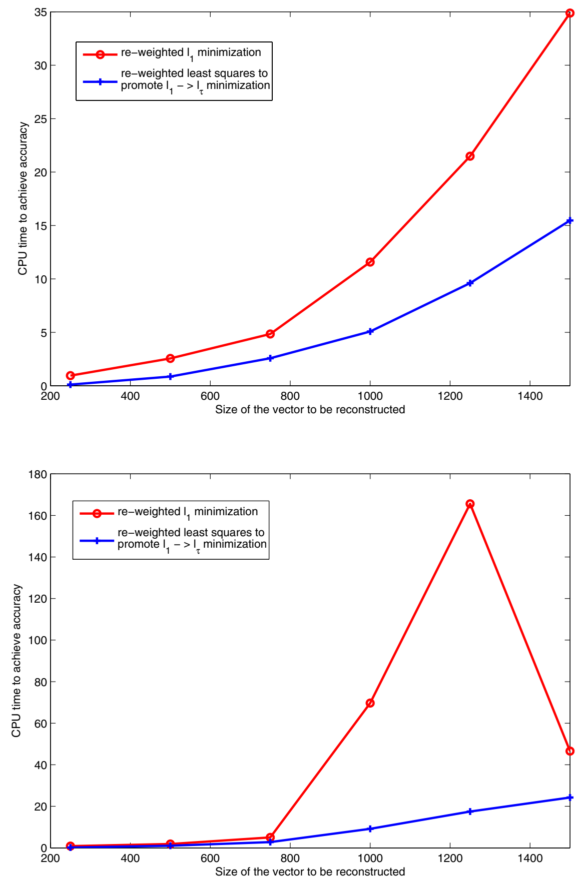


FIGURE 8.5. We compare the CPU time of the IRLS algorithm with weights that gradually move from an $\ell_1$- to an $\ell_\tau$-minimization goal, with final $\tau = 0.5$, and the CPU time of the iteratively reweighted $\ell_1$-minimization as proposed in [10] in recovering a sparse vector for different dimensions $(k, m, N)$ of the problem. In the first plot we show the computational time required by the algorithms in reconstructing with uniform accuracy $10^{-7}$ a sparse vector with $k = 5$. In these experiments the matrix $\Phi$ is $m \times N$ dimensional, with i.i.d. Gaussian $\mathcal{N}(0, \frac{1}{m})$ entries, for $N \in \{250, 500, 750, 1000, 1250, 1500\}$ and $m = N/5$. In the second plot we show the same situation but for a sparse vector with a support size $k = 12m/50$. Note that for this choice of $k$, the iteratively reweighted $\ell_1$-minimization tends to fail the reconstruction of the sparse vector, whereas the IRLS can still robustly succeed. We implemented both algorithms using Matlab without any specific optimization. In particular, for the iteratively reweighted $\ell_1$-minimization we used the CVX Matlab package by Michael Grant and Stephen Boyd, provided from `http://www.stanford.edu/~boyd/cvx/` to perform the weighted $\ell_1$-minimization at each iteration.

**Acknowledgments.** We would like to thank Yu Chen and Michael Overton for various conversations on the topic of this paper, and Rachel Ward for pointing out an improvement of Theorem 5.3.

I. Daubechies gratefully acknowledges partial support under National Science Foundation grants DMS-0504924 and DMS-0530865. R. DeVore thanks the Courant Institute for supporting an academic year visit when part of this work was done. He also gratefully acknowledges partial support by Office of Naval Research contracts ONR-N00014-03-1-0051, ONR/DEPSCoR N00014-03-1-0675, and ONR/ DEPSCoR N00014-00-1-0470; Army Research Office Contract DAAD 19-02-1-0028; and National Science Foundation grants DMS-0221642 and DMS-0200187. M. Fornasier acknowledges the financial support provided by the European Union via the Marie Curie Individual Fellowship MOIF-CT-2006-039438, and he thanks the Program in Applied and Computational Mathematics at Princeton University for its hospitality during the preparation of this work. S. Güntürk gratefully acknowledges partial support by National Science Foundation Grant CCF-0515187, the Alfred P. Sloan Research Fellowship, and the New York University Goddard Fellowship.

## Bibliography

[1] Baraniuk, R. G. Compressive sensing [Lecture notes]. *IEEE Signal Processing Magazine* **24** (2007), no. 4, 118–121.

[2] Baraniuk, R.; Davenport, M.; DeVore, R.; Wakin, M. B. A simple proof of the restricted isometry property for random matrices (aka "The Johnson-Lindenstrauss lemma meets compressed sensing"). *Constr. Approx.* **28** (2008), no. 3, 253–263.

[3] Byrd, R.; Payne, D. *Convergence of the IRLS algorithm for robust regression*. Technical Report 313, John Hopkins University, 1979.

[4] Candès, E. J. Compressive sampling. *International Congress of Mathematicians. Vol. III*, 1433–1452. European Mathematical Society, Zürich, 2006.

[5] Candès, E. J. Lecture notes of the IMA course on *Compressive Sampling and Frontiers in Signal Processing*, June 4–15, 2007. Available at: `http://www.ima.umn.edu/2006-2007/ND6.4-15.07/abstracts.html`.

[6] Candès, E. J.; Romberg, J.; Tao, T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inform. Theory* **52** (2006), no. 2, 489–509.

[7] Candès, E. J.; Romberg, J. K.; Tao, T. Stable signal recovery from incomplete and inaccurate measurements. *Comm. Pure Appl. Math.* **59** (2006), no. 8, 1207–1223.

[8] Candès, E. J.; Tao, T. Decoding by linear programming. *IEEE Trans. Inform. Theory* **51** (2005), no. 12, 4203–4215.

[9] Candès, E. J.; Tao, T. Near-optimal signal recovery from random projections: universal encoding strategies? *IEEE Trans. Inform. Theory* **52** (2006), no. 12, 5406–5425.

[10] Candès, E. J.; Wakin, M. B.; Boyd, S. P. Enhancing sparsity by reweighted $l_1$ minimization. *J. Fourier Anal. Appl.* **14** (2008), no. 5-6, 877–905.

[11] Chambolle, A.; Lions, P.-L. Image recovery via total variation minimization and related problems. *Numer. Math.* **76** (1997), no. 2, 167–188.

[12] Chartrand, R. Exact reconstructions of sparse signals via nonconvex minimization. *IEEE Signal Process. Lett.* **14** (2007), no. 10, 707–701.

[13] Chartrand, R. Nonconvex compressive sensing and reconstruction of gradient-sparse images: random vs. tomographic Fourier sampling. *IEEE International Conference on Image Processing (ICIP)* 2008. Available at: `http://math.lanl.gov/~rick/Publications/chartrand-2008-nonconvex.shtml`.

[14] Chartrand, R.; Staneva, V. Restricted isometry properties and nonconvex compressive sensing. *Inverse Problems* **24** (2008), no. 035020, 1–14.

[15] Chartrand, R.; Yin, W. Iteratively reweighted algorithms for compressive sensing. *Proceedings of the 33rd IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2008. Available at: `http://math.lanl.gov/~rick/Publications/chartrand-2008-iteratively.shtml`.

[16] Chen, S. S.; Donoho, D. L.; Saunders, M. A. Atomic decomposition by basis pursuit. *SIAM J. Sci. Comput.* **20** (1998), no. 1, 33–61 (electronic).

[17] Claerbout, J. F.; Muir, F. Robust modeling with erratic data. *Geophysics* **38** (1973), no. 5, 826–844.

[18] Cline, A. K. Rate of convergence of Lawson's algorithm. *Math. Comp.* **26** (1972), 167–176.

[19] Cohen, A.; Dahmen, W.; DeVore, R. Compressed sensing and best $k$-term approximation. *J. Amer. Math. Soc.* **22** (2009), no. 1, 211–231.

[20] DeVore, R. A. Nonlinear approximation. *Acta Numer.* **7** (1998), 51–150.

[21] Donoho, D. L. Compressed sensing. *IEEE Trans. Inform. Theory* **52** (2006), no. 4, 1289–1306.

[22] Donoho, D. L. High-dimensional centrally symmetric polytopes with neighborliness proportional to dimension. *Discrete Comput. Geom.* **35** (2006), no. 4, 617–652.

[23] Donoho, D. L.; Logan, B. F. Signal recovery and the large sieve. *SIAM J. Appl. Math.* **52** (1992), no. 2, 577–591.

[24] Donoho, D. L.; Stark, P. B. Uncertainty principles and signal recovery. *SIAM J. Appl. Math.* **49** (1989), no. 3, 906–931.

[25] Donoho, D. L.; Tanner, J. Sparse nonnegative solution of underdetermined linear equations by linear programming. *Proc. Natl. Acad. Sci. USA* **102** (2005), no. 27, 9446–9451 (electronic).

[26] Donoho, D. L.; Tanner, J. Counting faces of randomly projected polytopes when the projection radically lowers dimension. *J. Amer. Math. Soc.* **22** (2009), no. 1, 1–53.

[27] Donoho, D. L.; Tsaig, Y. Fast solution of $\ell_1$-norm minimization problems when the solution may be sparse. *IEEE Trans. Inform. Theory* **54** (2008), no. 11, 4789–4812.

[28] Figueiredo, M. A. T.; Bioucas-Dias, J. M.; Nowak, R. D. Majorization-minimization algorithms for wavelet-based image restoration. *IEEE Trans. Image Process.* **16** (2007), no. 12, 2980–2991.

[29] Fornasier, M.; March, R. Restoration of color images by vector valued BV functions and variational calculus. *SIAM J. Appl. Math.* **68** (2007), no. 2, 437–460.

[30] Gorodnitsky, I. F.; Rao, B. D. Sparse signal reconstruction from limited data using FOCUSS: a re-weighted minimum norm algorithm. *IEEE Trans. Signal Process.* **45** (1997), no. 3, 600–616.

[31] Gribonval, R.; Nielsen, M. Highly sparse representations from dictionaries are unique and independent of the sparseness measure. *Appl. Comput. Harmon. Anal.* **22** (2007), no. 3, 335–355.

[32] Lawson, C. L. *Contributions to the theory of linear least maximum approximation*. Ph.D. thesis, University of California, Los Angeles, 1961.

[33] Li, Y. A globally convergent method for $l_p$ problems. *SIAM J. Optim.* **3** (1993), no. 3, 609–629.

[34] O'Brien, M. S.; Sinclair, A. N.; Kramer, S. M. Recovery of a sparse spike time series by $\ell_1$ norm deconvolution. *IEEE Trans. Signal Process.* **42** (1994), no. 12, 3353–3365.

[35] Osborne, M. R. *Finite algorithms in optimization and data analysis*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. Wiley, Chichester, 1985.

[36] Pinkus, A. M. *On $L^1$-approximation*. Cambridge Tracts in Mathematics, 93. Cambridge University Press, Cambridge, 1989.

[37] Rudelson, M.; Vershynin, R. On sparse reconstruction from Fourier and Gaussian measurements. *Comm. Pure Appl. Math.* **61** (2008), no. 8, 1025–1045.

[38] Saab, R.; Chartrand, R.; Yilmaz, O. Stable sparse approximations via nonconvex optimization. *Proceedings of the 33rd IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 3885–3888, ICASSP, 2008.

[39] Santosa, F.; Symes, W. W. Linear inversion of band-limited reflection seismograms. *SIAM J. Sci. Statist. Comput.* **7** (1986), no. 4, 1307–1330.

[40] Taylor, H. L.; Banks, S. C.; McCoy, J. F. Deconvolution with the $\ell_1$ norm. *Geophysics* **44** (1979), no. 1, 39–52.

[41] Tibshirani, R. Regression shrinkage and selection via the lasso. *J. Roy. Statist. Soc. Ser. B* **58** (1996), no. 1, 267–288.

INGRID DAUBECHIES
Princeton University
Department of Mathematics
and Program in Applied
and Computational Mathematics
Fine Hall, Washington Road
Princeton, NJ 08544
E-mail: ingrid@math.princeton.edu

RONALD DEVORE
University of South Carolina
Department of Mathematics
Columbia, SC 29208
E-mail: devore@math.sc.edu

MASSIMO FORNASIER
Johann Radon Institute
for Computational
and Applied Mathematics
Austrian Academy of Sciences
Altenbergerstrasse 69
A-4040 Linz
AUSTRIA
E-mail: massimo.fornasier@oeaw.ac.at

C. SİNAN GÜNTÜRK
Courant Institute
251 Mercer Street
New York, NY 10012
E-mail: gunturk@courant.nyu.edu

Received June 2008.