QUASI-LINEAR COMPRESSED SENSING

MARTIN EHLER[*], MASSIMO FORNASIER[†], AND JULIANE SIGL[‡]

Abstract. Inspired by significant real-life applications, in particular, sparse phase retrieval and sparse pulsation frequency detection in Asteroseismology, we investigate a general framework for *compressed sensing*, where the measurements are *quasi-linear*. We formulate natural generalizations of the well-known *Restricted Isometry Property* (RIP) towards nonlinear measurements, which allow us to prove both unique identifiability of sparse signals as well as the convergence of recovery algorithms to compute them efficiently. We show that for certain randomized quasi-linear measurements, including Lipschitz perturbations of classical RIP matrices and phase retrieval from random projections, the proposed restricted isometry properties hold with high probability. We analyze a generalized *Orthogonal Least Squares* (OLS) under the assumption that magnitudes of signal entries to be recovered decay fast. Greed is good again, as we show that this algorithm performs efficiently in phase retrieval and Asteroseismology. For situations where the decay assumption on the signal does not necessarily hold, we propose two alternative algorithms, which are natural generalizations of the well-known *iterative hard and soft-thresholding*. While these algorithms are rarely successful for the mentioned applications, we show their strong recovery guarantees for quasi-linear measurements which are Lipschitz perturbations of RIP matrices.

Key words. compressed sensing, restricted isometry property, greedy algorithm, quasi-linear, iterative thresholding

AMS subject classifications. 94A20, 47J25, 15B52

1. Introduction. Compressed sensing addresses the problem of recovering nearly-sparse signals from vastly incomplete measurements [11, 12, 14, 15, 21]. By using the prior assumptions on the signal, the number of measurements can be well below the Shannon sampling rate and effective reconstruction algorithms are available. The standard compressed sensing approach deals with *linear* measurements. The success of signal recovery algorithms often relies on the so-called *Restricted Isometry Property* (RIP) [12, 15, 27, 35, 38, 39], which is a near-identity spectral property of small submatrices of the measurement Gramian. The RIP condition is satisfied with high probability and nearly optimal number of measurements for a large class of random measurements [3, 4, 14, 35, 38], which explains the popularity of all sorts of random sensing approaches. The most effective recovery algorithms are based either on a greedy approach or on variational models, such as ℓ_1-norm minimization, leading to suitable iterative thresholded gradient descent methods. In the literature of mathematical signal processing, greedy algorithms for sparse recovery originates from the so-called Matching Pursuit [33], although several predecessors were well-known in other communities. Among astronomers and asteroseismologists, for instance, *Orthogonal Least Squares* (OLS) [31] was already in use in the '60s for the detection of significant frequencies of star light-spectra (the so-called pre-whitening) [5]. We refer to [34, 43] for more recent developments on greedy approaches. Iterative thresholding algorithms have instead a variational nature and they are designed to minimize the discrepancy with respect to the measurements and simultaneously to promote sparsity by iterated thresholding operations. We refer to [10, 18, 26] and references therein for more details on such iterative schemes for sparse recovery from linear measurements.

Often models of physical measurements in the applied sciences and engineering, however, are not linear and it is of utmost interest to investigate to which extent the theory of compressed sensing can be generalized to nonlinear measurements. Two relevant real-life applications in physics can be mentioned, asteroseismic light measurements [1] to determine the shape of pulsating stars and magnitude measurements in phase retrieval problems important to diffraction imaging and X-ray crystallography [22, 24, 29]. There are already several recent attempts towards nonlinear compressed sensing, for instance the work by Blumensath et al. [8, 9], quadratic measurements are considered in [32], and further nonlinear inverse problems are analyzed in [36, 37]. Phase retrieval is an active field of research nowadays and has been addressed by related approaches [2, 6, 13, 23, 41].

---

[*]Helmholtz Zentrum München, Institute of Computational Biology, Ingolstädter Landstrasse 1, 85764 Neuherberg, Germany, (`martin.ehler@helmholtz-muenchen.de`).

[†]Technische Universität München, Fakultät für Mathematik, Boltzmannstr. 3, 85748 Garching bei München, (`massimo.fornasier@ma.tum.de`)

[‡]Technische Universität München, Fakultät für Mathematik, Boltzmannstr. 3, 85748 Garching bei München, (`sigl.juliane@gmail.com`)

In the present paper we provide a more unified view, by restricting the possible nonlinearity of the measurements to quasi-linear maps, which are sufficiently smooth, at least Lipschitz, and they fulfill generalized versions of the classical RIP. In contrast to the situation of linear measurements, the nonlinearity of the measurements actually plays in a differing manner within different recovery algorithms. Therefore it is necessary to design corresponding forms of RIP depending on the recovery strategies used, see conditions (3.1), (3.8) for a greedy algorithm and (4.2), (4.11) for iterative thresholding algorithms. In particular, we show that for certain randomized quasi-linear measurements, including Lipschitz perturbations of classical RIP matrices and phase retrieval from random projections, the proposed restricted isometry properties hold with high probability. While for the phase retrieval problem the stability results in [23] are restricted to the real setting, we additionally extend them to the complex case.

Algorithmically we first focus on a generalized *Orthogonal Least Squares* (OLS). Such a greedy approach was already proposed in [9], although there no analysis of convergence was yet provided. We show within the framework of quasi-linear compressed sensing problems the recovery guarantees of this algorithm, by taking inspiration from [19], where a similar analysis is performed for linear measurements. The greedy algorithm we propose works for both types of applied problems mentioned above, i.e., Asteroseismology and phase retrieval. Let us stress that for the latter and for signals which have rapidly decaying nonincreasing rearrangement, few iterations of this greedy algorithm are sufficient to obtain a good recovery accuracy. Hence, our approach seems very competitive compared to the semi-definite program used in [13] for phase retrieval, by recasting the problem into a demanding optimization on matrices.

The greedy strategy as derived here, however, also inherits two drawbacks: (1) the original signal is required to satisfy the mentioned decay conditions, and (2) the approach needs careful implementations of multivariate global optimization to derive high accuracy for signal recovery.

To possibly circumvent those drawbacks, we then explore alternative strategies, generalizing iterative hard- and soft-thresholding methods, which allow us to recover nearly-sparse signals not satisfying the decay assumptions. The results we present for hard-thresholding are mainly based on Blumensath's findings in [8]. For iterative soft-thresholding, we prove in an original fashion the convergence of the algorithm towards a limit point and we bound its distance to the original signal. While iterative thresholding algorithms are rarely successful for phase retrieval problems, we show their strong recovery guarantees for quasi-linear measurements which are Lipschitz perturbations of RIP matrices. We further emphasize in our numerical experiments that different iterative algorithms based on contractive principles do provide rather diverse success recovery results for the same problem, especially when nonlinearities are involved: this is due to the fact that the basins of attraction towards fixed points of the corresponding iterations can be significantly different. In our view, this is certainly sufficient motivation to explore several algorithmic approaches and not restricting ourselves just to a favorite one.

As we clarified above, each algorithmic approach requires a different treatment of the nonlinearity of the measurements with the consequent need of defining corresponding generalizations of the RIP. Hence, we develop the presentation of our results according to the different algorithms, starting first with the generalized Orthogonal Least Squares, and later continuing with the iterative thresholding algorithms. Along the way, we present examples of applications and we show how to fulfill the required deterministic conditions of convergence by randomized quasi-linear measurements. The outline of the paper is as follows: In Section 2, we introduce the nonlinear compressed sensing problem, and in Section 3 we derive a greedy scheme for nearly sparse signal reconstruction. We show applications of this algorithm in Section 3.2.2 to analyze simulated asteroseismic data towards the detection of frequency pulsation of stars. In Section 3.3.1, we discuss refinements and changes needed for the phase retrieval problem and also provide numerical experiments. For signals not satisfying the decay assumptions needed for the greedy algorithm to converge, iterative thresholding algorithms are discussed in Section 4.

2. Quasi-linear compressed sensing.

2.1. The nonlinear model. In the by now classical compressed sensing framework, an unknown nearly sparse signal $\hat{x} \in \mathbb{R}^d$ is to be reconstructed from n linear measurements, with $n \ll d$, and one models this setting as the solution of a linear system

$$A\hat{x} + e = b,$$

where e is some noise term and the i-th row of $A \in \mathbb{R}^{n \times d}$ corresponds to the i-th linear measurement on the unknown signal $\hat{x}$ with outcome b_i. We say that A satisfies the Restricted Isometry Property (RIP) of order k with $0 < \delta_k < 1$ if

$$(1 - \delta_k)\|x\| \leq \|Ax\| \leq (1 + \delta_k)\|x\|, \tag{2.1}$$

for all $x \in \mathbb{R}^d$ with at most k nonzero entries. We call such vectors k-sparse. If A satisfies the RIP of order $2k$ and $\delta_{2k} < \frac{2}{3+\sqrt{7/4}} \approx 0.4627$, then signal recovery is possible up to noise level and k-term approximation error. It should be mentioned that large classes of random matrices $A \in \mathbb{R}^{n \times d}$ satisfy the RIP with high probability for the (nearly-)optimal dimensionality scaling $k = O\left(\frac{n}{1+\log(d/n)^\alpha}\right)$. We refer to [4, 12, 14, 15, 21, 35, 38] for the early results and [28] for a recent extended treatise.

Many real-life applications in physics and biomedical sciences, however, carry some strongly nonlinear structure, so that the linear model is not suited anymore, even as an approximation. Towards the definition of a nonlinear framework for compressed sensing, we shall consider for $n \ll d$ a map $A : \mathbb{R}^d \to \mathbb{R}^n$, which is not anymore necessarily linear, and aim at reconstructing $\hat{x} \in \mathbb{R}^d$ from the measurements $b \in \mathbb{R}^n$ given by

$$A(\hat{x}) + e = b. \tag{2.2}$$

Similarly to linear problems, also the unique and stable solution of the equation (2.2) is in general an impossible task, unless we require certain a priori assumptions on $\hat{x}$, and some stability properties similar to (3.1) for the nonlinear map A. As the variety of possible nonlinearities is extremely vast, it is perhaps too ambitious to expect that generalized RIP properties can be verified for any type of nonlinearity. As a matter of fact, and as we shall show in details below, most of the nonlinear models with stability properties which allow for nearly sparse signal recovery, have a smooth quasi-linear nature. With this we mean that there exists a Lipschitz map $F : \mathbb{R}^d \to \mathbb{R}^{n \times d}$ such that $A(x) = F(x)x$, for all $x \in \mathbb{R}^d$. However, in the following we will use and explicitly highlight this quasi-linear structure only when necessary. Our first approach

---

ℓ_p-**greedy algorithm**:
Input: $A : \mathbb{R}^d \to \mathbb{R}^n$, $b \in \mathbb{R}^n$
Initialize $x^{(0)} = 0 \in \mathbb{R}^d$, $\Lambda^{(0)} = \emptyset$
for $j = 1, 2, \ldots$ *until some stopping criterion is met* **do**
$\quad$ **for** $l \notin \Lambda^{(j-1)}$ **do**
$\quad\quad \Lambda^{(j-1,l)} := \Lambda^{(j-1)} \cup \{l\}$

$$x^{(j,l)} := \arg \min_{\{x : \text{supp}(x) \subset \Lambda^{(j-1,l)}\}} \left\| A(x) - b \right\|_{\ell_p}$$

$\quad$ **end**
$\quad$ Find index that minimizes the error:

$$l_j := \arg \min_l \left\| A(x^{(j,l)}) - b \right\|_{\ell_p}$$

$\quad$ Update: $x^{(j)} := x^{(j,l_j)}$, $\Lambda^{(j)} := \Lambda^{(j-1,l_j)}$
end
Output: $x^{(1)}$, $x^{(2)}, \ldots$

Algorithm 1: The ℓ_p-greedy algorithm terminates after finitely many steps, but we need to solve $d - j$ many j-dimensional optimization problem in the j-th step. If we know that $b = A(\hat{x}) + e$ holds and $\hat{x}$ is k-sparse, then the stopping criterion $j \leq k$ appears natural, but can also be replaced with other conditions.

---

towards the solution of (2.2) will be based on a greedy principle, since it is also perhaps the most intuitive one: we search first for the best 1-sparse signal which is minimizing the discrepancy with respect to the measurements and then we seek for a next best matching 2-sparse signal having as one of the active entries the one previously detected, and so on. This method is formally summarized in the ℓ_p-norm matching greedy

Algorithm 1. For the sake of clarity, we mention that $\|x\|$ denotes the standard Euclidean norm of any vector $x \in \mathbb{R}^d$, while $\|x\|_{\ell_p} = \left(\sum_{i=1}^{d} |x_i|^p\right)^{1/p}$ is its ℓ_p-norm for $1 \leq p < \infty$. Moreover, when dealing with matrices, we denote with $\|A\| = \|A\|_2$ the spectral norm of the matrix A and with $\|A\|_{HS}$ its Hilbert-Schmidt norm.

3. Greed is good - again.

3.1. Deterministic conditions I. Greedy algorithms have already proven useful and efficient for many sparse signal reconstruction problems in a linear setting, cf. [44], and we refer to [7] for a more recent treatise. Before we can state our reconstruction result here, we still need some preparation. The nonincreasing rearrangement of $x \in \mathbb{R}^d$ is defined as

$$r(x) = (|x_{j_1}|, \ldots, |x_{j_d}|)^\top, \quad \text{where} \quad |x_{j_i}| \geq |x_{j_{i+1}}|, \text{ for } i = 1, \ldots, d-1.$$

For $0 < \kappa < 1$, we define the class of κ-rapidly decaying vectors in $\mathbb{R}^d$ by

$$\mathcal{D}_\kappa = \{x \in \mathbb{R}^d : r_{j+1}(x) \leq \kappa r_j(x), \text{ for } j = 1, \ldots, d-1\}.$$

Given $x \in \mathbb{R}^d$, the vector $x_{\{j\}} \in \mathbb{R}^d$ is the best j-sparse approximation of x, i.e., it consists of the j largest entries of x in absolute value and zeros elsewhere. Signal recovery is possible under decay and stability conditions using the ℓ_p-greedy Algorithm 1, which is a generalized Orthogonal Least Squares [31]:

THEOREM 3.1. *Let $b = A(\hat{x}) + e$, where $\hat{x} \in \mathbb{R}^d$ is the signal to be recovered and $e \in \mathbb{R}^n$ is a noise term. Suppose further that $1 \leq k \leq d$, $r_k(\hat{x}) \neq 0$, and $1 \leq p < \infty$. If the following conditions hold,*

(i)   *there are $\alpha_k, \beta_k > 0$ such that, for all k-sparse $y \in \mathbb{R}^d$,*

$$\alpha_k \|\hat{x}_{\{k\}} - y\| \leq \|A(\hat{x}_{\{k\}}) - A(y)\|_{\ell_p} \leq \beta_k \|\hat{x}_{\{k\}} - y\|, \tag{3.1}$$

(ii)   *$\hat{x} \in \mathcal{D}_\kappa$ such that $\kappa < \frac{\tilde{\alpha}_k}{\sqrt{\tilde{\alpha}_k^2 + (\beta_k + 2L_k)^2}}$, where $0 < \tilde{\alpha}_k \leq \alpha_k - 2\|e\|_{\ell_p}/r_k(\hat{x})$ and $L_k \geq 0$ with*
$$\|A(\hat{x}) - A(\hat{x}_{\{k\}})\|_{\ell_p} \leq L_k \|\hat{x} - \hat{x}_{\{k\}}\|,$$
then the ℓ_p-greedy Algorithm 1 yields a sequence $(x^{(j)})_{j=1}^{k}$ satisfying $\operatorname{supp}(x^{(j)}) = \operatorname{supp}(\hat{x}_{\{j\}})$ and

$$\|x^{(j)} - \hat{x}\| \leq \|e\|_{\ell_p}/\alpha_k + \kappa^j r_1(\hat{x})\sqrt{2}\left(1 + \frac{\beta_k + 2L_k}{\alpha_k}\right).$$

If $\hat{x}$ is k-sparse, then $\|x^{(k)} - \hat{x}\| \leq \|e\|_{\ell_p}/\alpha_k$.

According to $0 < \tilde{\alpha}_k \leq \alpha_k - 2\|e\|_{\ell_p}/r_k(\hat{x})$, the noise term e must be small and we implicitly suppose that $r_k(\hat{x}) \neq 0$. Otherwise, we can simply choose a smaller k. If the signal $\hat{x}$ is k-sparse and the noise term e equals zero, then the ℓ_p-greedy Algorithm 1 in Theorem 3.1 yields $x^{(k)} = \hat{x}$.

REMARK 3.2. *A similar greedy algorithm was proposed in [9] for nonlinear problems, and our main contribution here is the careful analysis of its signal recovery capabilities. Conditions of the type (3.1) have also been used in [8], but with additional restrictions, so that the constants α_k and β_k must be close to each other. In our Theorem 3.1, we do not need any constraints on such constants, because the decay conditions on the signal can compensate for this. A similar relaxation using decaying signals was proposed in [19] for linear operators A, but even there the authors still assume $\beta_k/\alpha_k < 2$. We do not require here any of such conditions.*

The proof of Theorem 3.1 extends the preliminary results obtained in [42], which we split and generalize in the following two lemmas:

LEMMA 3.3. *If $\hat{x} \in \mathbb{R}^d$ is contained in $\mathcal{D}_\kappa$, then*

$$\|\hat{x} - \hat{x}_{\{j\}}\| < r_{j+1}(\hat{x})\frac{1}{\sqrt{1-\kappa^2}} \leq r_j(\hat{x})\frac{\kappa}{\sqrt{1-\kappa^2}}. \tag{3.2}$$

The proof of Lemma 3.3 is a straightforward calculation, in which the geometric series is used, so we omit the details.

LEMMA 3.4. *Fix $1 \leq k \leq d$ and suppose that $r_k(\hat{x}) \neq 0$. If $\hat{x} \in \mathbb{R}^d$ is contained in $\mathcal{D}_\kappa$ with $\kappa < \frac{\tilde{\alpha}}{\sqrt{\tilde{\alpha}^2 + (\beta + 2L)^2}}$, where $0 < \tilde{\alpha} \leq \alpha - 2\|e\|_{\ell_p}/r_k(\hat{x})$, for some $\alpha, \beta, L > 0$, then, for $j = 1, \ldots, k$,*

$$\alpha r_j(\hat{x}) > 2\|e\|_{\ell_p} + 2L\|\hat{x} - \hat{x}_{\{k\}}\| + \beta\|\hat{x}_{\{j\}} - \hat{x}_{\{k\}}\|. \tag{3.3}$$

Proof. It is sufficient to consider $\hat{x}$, which are not j-sparse. A short calculation reveals that the condition on κ implies

$$\alpha - 2\|e\|_{\ell_p}/r_k(\hat{x}) - \frac{\kappa}{\sqrt{1-\kappa^2}}(\beta + 2L) > 0.$$

We multiply the above inequality by $r_j(\hat{x})$, so that $r_j(\hat{x}) \geq r_k(\hat{x})$ yields

$$\alpha r_j(\hat{x}) - 2\|e\|_{\ell_p} - 2L\frac{\kappa}{\sqrt{1-\kappa^2}}r_k(\hat{x}) - \beta\frac{\kappa}{\sqrt{1-\kappa^2}}r_j(\hat{x}) > 0.$$

Lemma 3.3 now implies (3.3). □

We can now turn to the proof of Theorem 3.1:

Proof. [Proof of Theorem 3.1] We must check that the index set selected by the ℓ_p-greedy algorithm matches the location of the nonzero entries of $\hat{x}_{\{k\}}$. We use induction and observe that nothing needs to be checked for $j = 0$. In the induction step, we suppose that $\Lambda^{(j-1)} \subset \text{supp}(\hat{x}_{\{j-1\}})$. Let us choose $l \notin \text{supp}(\hat{x}_{\{j\}})$. The lower inequality in (3.1) yields

$$\begin{aligned}
\|b - A(x^{(j,l)})\|_{\ell_p} &\geq -\|e\|_{\ell_p} + \|A(\hat{x}) - A(x^{(j,l)})\|_{\ell_p} \\
&\geq -\|e\|_{\ell_p} - \|A(\hat{x}) - A(\hat{x}_{\{k\}})\|_{\ell_p} + \|A(\hat{x}_{\{k\}}) - A(x^{(j,l)})\|_{\ell_p} \\
&\geq -\|e\|_{\ell_p} - L_k\|\hat{x} - \hat{x}_{\{k\}}\| + \alpha_k\|\hat{x}_{\{k\}} - x^{(j,l)}\| \\
&\geq -\|e\|_{\ell_p} - L_k\|\hat{x} - \hat{x}_{\{k\}}\| + \alpha_k r_j(\hat{x}),
\end{aligned}$$

where the last inequality is a crude estimate based on $l \notin \text{supp}(\hat{x}_{\{j\}})$. Thus, Lemma 3.4 implies

$$\|b - A(x^{(j,l)})\|_{\ell_p} > \|e\|_{\ell_p} + (\beta_k + L_k)\|\hat{x} - \hat{x}_{\{k\}}\| + \beta_k\|\hat{x} - \hat{x}_{\{j\}}\|. \tag{3.4}$$

On the other hand, the minimizing property of $x^{(j)}$ and the upper inequality in (3.1) yield

$$\begin{aligned}
\|b - A(x^{(j)})\|_{\ell_p} &\leq \|b - A(\hat{x}_{\{j\}})\|_{\ell_p} \\
&\leq \|e\|_{\ell_p} + \|A(\hat{x}) - A(\hat{x}_{\{j\}})\|_{\ell_p} \\
&\leq \|e\|_{\ell_p} + \|A(\hat{x}) - A(\hat{x}_{\{k\}})\|_{\ell_p} + \|A(\hat{x}_{\{k\}}) - A(\hat{x}_{\{j\}})\|_{\ell_p} \\
&\leq \|e\|_{\ell_p} + L_k\|\hat{x} - \hat{x}_{\{k\}}\| + \beta_k\|\hat{x}_{\{k\}} - \hat{x}_{\{j\}}\|.
\end{aligned}$$

The last line and (3.4) are contradictory if $x^{(j)} = x^{(j,l)}$, so that we must have $x^{(j)} = x^{(j,l)}$, for some $l \in \text{supp}(\hat{x}_{\{j\}})$, which concludes the part about the support.

Next, we shall derive the error bound. Standard computations yield

$$\begin{aligned}
\|x^{(j)} - \hat{x}\| &\leq \|x^{(j)} - \hat{x}_{\{k\}}\| + \|\hat{x}_{\{k\}} - \hat{x}\| \\
&\leq 1/\alpha_k\|A(x^{(j)}) - A(\hat{x}_{\{k\}})\|_{\ell_p} + \|\hat{x}_{\{k\}} - \hat{x}\| \\
&\leq 1/\alpha_k\|A(x^{(j)}) - A(\hat{x})\|_{\ell_p} + 1/\alpha_k\|A(\hat{x}) - A(\hat{x}_{\{k\}})\|_{\ell_p} + \|\hat{x}_{\{k\}} - \hat{x}\| \\
&\leq 1/\alpha_k\|A(\hat{x}_{\{j\}}) - A(\hat{x})\|_{\ell_p} + \|e\|_{\ell_p}/\alpha_k + L_k/\alpha_k\|\hat{x} - \hat{x}_{\{k\}}\| + \|\hat{x}_{\{k\}} - \hat{x}\| \\
&\leq 1/\alpha_k\|A(\hat{x}_{\{j\}}) - A(\hat{x}_{\{k\}})\|_{\ell_p} + \|e\|_{\ell_p}/\alpha_k + 1/\alpha_k\|A(\hat{x}_{\{k\}}) - A(\hat{x})\|_{\ell_p} \\
&\quad + L_k/\alpha_k\|\hat{x} - \hat{x}_{\{k\}}\| + \|\hat{x}_{\{k\}} - \hat{x}\| \\
&\leq \beta_k/\alpha_k\|\hat{x}_{\{j\}} - \hat{x}_{\{k\}}\| + \|e\|_{\ell_p}/\alpha_k + 2L_k/\alpha_k\|\hat{x} - \hat{x}_{\{k\}}\| + \|\hat{x}_{\{k\}} - \hat{x}\| \\
&\leq \left(\beta_k/\alpha_k + 2L_k/\alpha_k + 1\right)\|\hat{x}_{\{k\}} - \hat{x}\| + \|e\|_{\ell_p}/\alpha_k \\
&\leq \left(\beta_k/\alpha_k + 2L_k/\alpha_k + 1\right)r_{j+1}(\hat{x})\frac{1}{\sqrt{1-\kappa^2}} + \|e\|_{\ell_p}/\alpha_k \\
&\leq \left(\beta_k/\alpha_k + 2L_k/\alpha_k + 1\right)\kappa^j r_1(\hat{x})\frac{1}{\sqrt{1-\kappa^2}} + \|e\|_{\ell_p}/\alpha_k.
\end{aligned}$$

Few rather rough estimates yield

$$\frac{\beta_k/\alpha_k + 2L_k/\alpha_k + 1}{\sqrt{1-\kappa^2}} \leq \sqrt{2}(1 + \frac{\beta_k + 2L_k}{\alpha_k})$$

which concludes the proof. □

3.2. Examples of inspiring applications. In this subsection we present examples where Algorithm 1 can be successfully used. We start with an abstract example of a nonlinear Lipschitz perturbation of a linear model and then we consider a relevant real-life example from Asteroseismology.

3.2.1. Lipschitz perturbation of a RIP matrix. As an explicit example of A matching the requirements of Theorem 3.1 with high probability, we propose Lipschitz perturbations of RIP matrices:

PROPOSITION 3.5. *If A is chosen as*

$$A(x) := A_1 x + \epsilon f(\|x - x_0\|)A_2 x, \tag{3.5}$$

where $A_1 \in \mathbb{R}^{n \times d}$ satisfies the RIP (2.1) of order k and constant $0 < \delta_k < 1$, $x_0 \in \mathbb{R}^d$ is some reference vector, $f : [0, \infty) \to \mathbb{R}$ is a bounded Lipschitz continuous function, ϵ is a sufficiently small scaling factor, and $A_2 \in \mathbb{R}^{n \times d}$ arbitrarily fixed, then there are constants $\alpha, \beta > 0$, such that the assumptions in Theorem 3.1 hold for $p = 2$.

Proof. We first check on α. If L denotes the Lipschitz constant of f, then we obtain

$$
\begin{aligned}
\|A(\hat{x}) - A(y)\| &= \|A_1\hat{x} - A_1 y + \epsilon f(\|\hat{x} - x_0\|)A_2\hat{x} - \epsilon f(\|y - x_0\|)A_2 y\| \\
&= \|A_1\hat{x} - A_1 y + \epsilon\big[f(\|\hat{x} - x_0\|)A_2\hat{x} - f(\|y - x_0\|)A_2\hat{x} \\
&\quad + f(\|y - x_0\|)A_2\hat{x} - f(\|y - x_0\|)A_2 y\big]\| \\
&\leq (1 + \delta_k)\|\hat{x} - y\| + \epsilon L \big|\|\hat{x} - x_0\| - \|y - x_0\|\big|\|A_2\|_2\|\hat{x}\| + \epsilon B\|A_2\|_2\|\hat{x} - y\| \\
&\leq (1 + \delta_k + \epsilon L\|A_2\|_2\|\hat{x}\| + \epsilon B\|A_2\|_2)\|\hat{x} - y\|,
\end{aligned}
$$

where we have used the reverse triangular inequality and $B = \sup\{|f(\|x - x_0\|)| : x \in \mathbb{R}^d \text{ is } k\text{-sparse}\}$. Thus, we can choose $\beta := 1 + \delta_k + \epsilon L\|A_2\|_2 r + B\|A_2\|_2$, where $r \geq \|\hat{x}\|$.

Next, we derive a suitable α. For k-sparse $y \in \mathbb{R}^d$, we derive similarly to the above calculations

$$
\begin{aligned}
\|A(\hat{x}) - A(y)\| &\geq \|A_1\hat{x} - A_1 y\| - \epsilon\|f(\|\hat{x} - x_0\|)A_2\hat{x} - f(\|y - x_0\|)A_2 y\| \\
&\geq (1 - \delta_k)\|\hat{x} - y\| - \epsilon L\|A_2\|_2\|\hat{x}\|\|\hat{x} - y\| - \epsilon B\|A_2\|_2\|\hat{x} - y\| \\
&= (1 - \delta_k - \epsilon\|A_2\|_2(L\|\hat{x}\| + B))\|\hat{x} - y\|.
\end{aligned}
$$

If ϵ is sufficiently small, we can choose $\alpha := 1 - \delta_k - \epsilon\|A_2\|_2(Lr + B)$. □ Any matrix satisfying the RIP of order k with high probability, for instance being within certain classes of random matrices [28], induces maps A via Proposition 3.5 that satisfy the assumptions of Theorem 3.1. Notice that the form of nonlinearity considered in (3.5) is actually quasi-linear, i.e., $A(x) = F(x)x$, where $F(x) = A_1 + \epsilon f(\|x - x_0\|)A_2$.

3.2.2. Quasi-linear compressed sensing in Asteroseismology. Asteroseismology studies the oscillation occurring inside variable pulsating stars as seismic waves [1]. Some regions of the stellar surface contract and heat up while others expand and cool down in a regular pattern causing observable changes in the light intensity. This also means that areas of different temperature correspond to locations of different expansion of the star and characterize its shape. Through the analysis of the frequency spectra it is possible to determine the internal stellar structure. Often complex pulsation patterns with multiperiodic oscillations are observed and their identification is needed.

We refer to [42] for a detailed mathematical formulation of the model connecting the instantaneous star shape and its actual light intensity at different frequencies. Here we limit ourselves to a schematic description where we assume the star being a two dimensional object with a pulsating shape contour. Let the function $u(\varphi)$ describe the star shape contour, for a parameter $-1 \leq \varphi \leq 1$, which also simultaneously represents the temperature (or emitted wavelength) on the stellar surface at some fixed point in time. Its oscillatory behavior yields
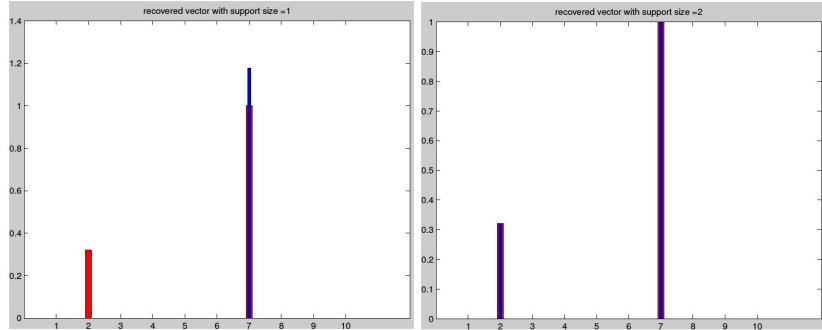
$$u(\varphi) = \sum_{i=1}^{d} x_i \sin((2\pi\varphi + \theta)i),$$

for some coefficient vector $x = (x_1, \ldots, x_d)$ and some inclination angle θ. This vector x needs to be reconstructed from the instantaneous light measurements $b = (b_1, \ldots, b_n)$, modeled in [42] by the formula

$$b_l = \frac{\sqrt{\pi}}{2d + 1} \sum_{j=-d}^{d} \omega_l \big(f_j \sum_{k=1}^{d} x_k \sin((2\pi\tfrac{j}{d} + \theta)k)\big) f_j \sum_{i=1}^{d} x_i \sin((2\pi\tfrac{j}{d} + \theta)i), \quad l = 1, \ldots, n, \tag{3.6}$$
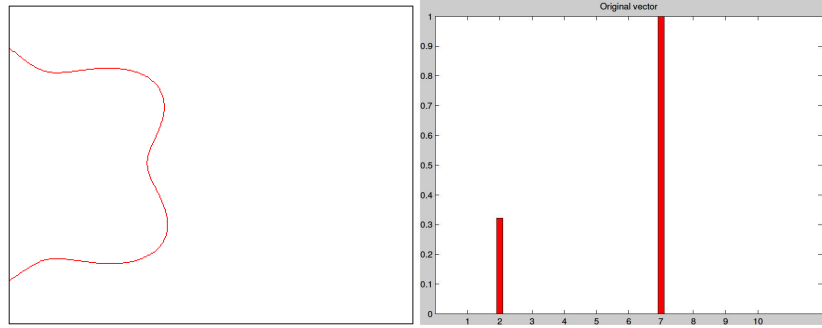
(a) Original shape by means of u and the corresponding 2-sparse Fourier coefficients x.

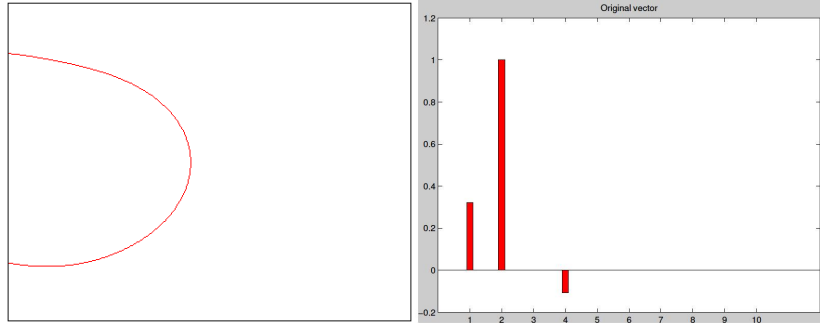(b) 1- and 2-sparse reconstruction of the Fourier coefficients.

FIG. 1. *Original 2-sparse signal and reconstructed pulsation patterns. Red corresponds to the original signal, the reconstruction is given in blue.*

so that we suppose that u is sampled at j/d. Here, f is a correction factor modeling *limb darkening*, i.e., the fading intensity of the light of the star towards its limb, and $\omega_l(\cdot)$ is some partition of unity modeling the wavelength range of each telescope sensor, see [42] for details. Notice that the light intensity data (3.6) at different frequency bands (corresponding to different ω_l) are obtained through the quasi-linear measurements $b = A(x) = F(x)x$ with

$$ F(x)_{l,i} := \frac{\sqrt{\pi}}{2d+1} \sum_{j=-d}^{d} \omega_l \big(f_j \sum_{k=1}^{d} x_k \sin((2\pi \frac{j}{d} + \theta)k) \big) f_j \sin((2\pi \frac{j}{d} + \theta)i), $$

and one wants to reconstruct a vector x matching the data with few nonzero entries. In fact it is rather accepted in the Asteroseismology community that only few low frequencies of the star shape (when its contour shape is expanded in spherical harmonics) are relevant [1].

We do not claim that the model (3.6) matches all of the assumptions in Theorem 3.1, but we shall observe that Algorithm 1 for $p = 2$ can be used to identify the instantaneous pulsation patterns of simulated light intensity data, when low frequencies are activated. This is a consequence of the fact that different low frequency activations result in sufficiently uncorrelated measurements in the data to be distinguished by the greedy algorithm towards recovery. For the numerical experiments, the ambient dimension is $d = 800$ and we make $n = 13$ measurements, see [42] for details on the choice of ω_l, f, and the used multivariate optimization routines. We generate 2- and 3-sparse signals and apply Algorithm 1 in Figs. 1 and 2 to reconstruct the signal. The greedy strategy identifies one additional location of the solution's support at each iteration step and finds the correct signal after 2 and 3 steps, respectively. In Fig. 3, we generated a signal whose entries decay rapidly, so that higher frequencies have lower magnitudes. We show the reconstruction of the shape after 3 iterations of the greedy algorithm. As expected, higher frequencies are suppressed and we obtain a low-pass filter approximation of the original shape.
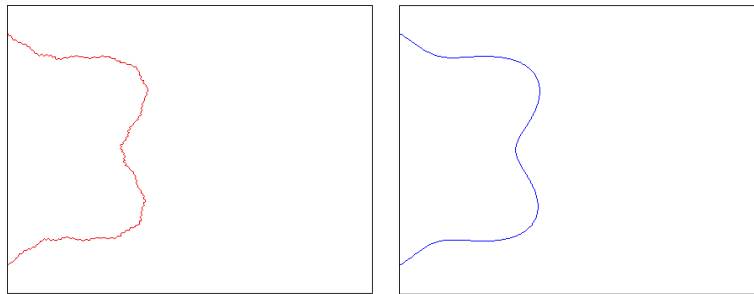
(a) Original shape by means of u and the corresponding 3-sparse Fourier coefficients x.
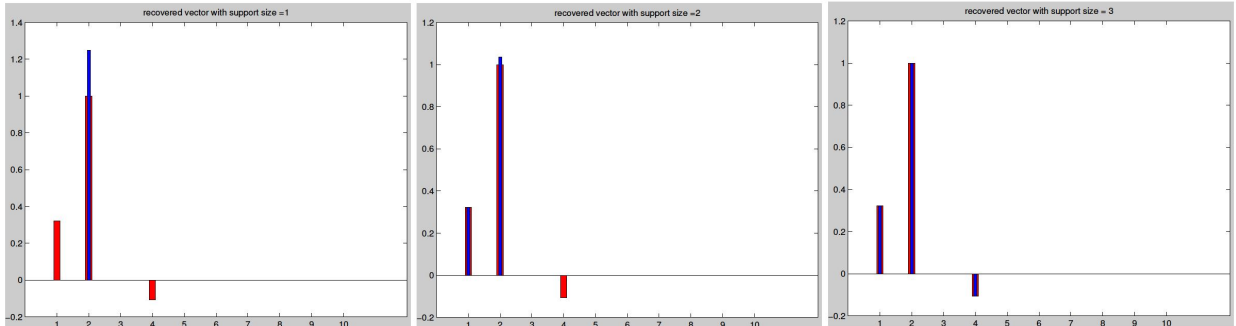
(b) 1-, 2-, and 3-sparse reconstruction.

FIG. 2. *Original 3-sparse signal and reconstructed pulsation patterns. Red corresponds to the original signal, the reconstruction is given in blue.*

(a) original shape   (b) reconstructed shape

FIG. 3. *The original signal is rapidly decaying with higher frequencies and we reconstruct 3 of its largest entries through the greedy algorithm. As desired, the reconstructed shape is a smoothened version of the original one.*

3.3. Deterministic conditions II. Experiments in X-ray crystallography and diffraction imaging require signal reconstruction from magnitude measurements, usually in terms of light intensities. We do not present the explicit physical models, which would go beyond the scope of the present paper, but refer to the literature instead. It seems impossible to provide a comprehensive list of references, so we only mention [22, 24, 29] for some classical algorithms.

Let $\hat{x} \in \mathbb{R}^d$ be some signal that we need to reconstruct from measurements $b = \left(|\langle a_i, \hat{x}\rangle|^2\right)_{i=1}^n$, where we selected a set of measurement vectors $\{a_i : i = 1, \ldots n\} \subset \mathbb{R}^d$. In other words, we have phaseless measurements and need to reconstruct $\pm\hat{x}$. It turns out surprisingly that the above framework of reconstruction from nonlinear sensing can be modified to fit the phase retrieval problem. Let us stress that so far the most efficient and stable recovery procedures are based on semi-definite programming, as used in [13], by

recasting the problem into a perhaps demanding optimization on matrices. In this section we show that there is no need to linearize the problem by lifting the dimensionality towards low-rank matrix recovery, but is is sufficient to address a plain sparse vector recovery in the fully nonlinear setting.

In models relevant to optical measurements like diffraction imaging and X-ray crystallography, we must deal with the complex setting, in which $x \in \mathbb{C}^d$ is at most determined up to multiplication by a complex unit. We shall state our findings for the real case first and afterwards discuss the extensions to complex vector spaces.

3.3.1. Quasi-linear compressed sensing in phase retrieval. Let $\{A_i : i = 1, \ldots, n\} \subset \mathbb{R}^{d \times d}$ be a collection of measurement matrices. We consider the map

$$A : \mathbb{R}^d \to \mathbb{R}^n, \quad x \mapsto A(x) = F(x)x, \text{ where } F(x) = \begin{pmatrix} x^* A_1 \\ \vdots \\ x^* A_n \end{pmatrix}, \tag{3.7}$$

and we aim at reconstructing a signal vector $\hat{x} \in \mathbb{R}^d$ from measurements $b = A(\hat{x})$. Since $A(x) = A(-x)$, the vector $\hat{x}$ can at best be determined up to its sign, and the lower bound in (3.1) cannot hold, not allowing us to use directly Theorem 3.1. However, we notice that for special classes of A, for instance, when $\{A_i : i = 1, \ldots, n\}$ are independent Gaussian matrices, the lower bound in (3.1) holds with high probability as long as y stays away from $-\hat{x}$, see the heuristic probability transitions of validity of (3.1) shown in Fig. 4. Hence, there is the hope that the greedy algorithm can nevertheless be successful, because it proceeds by selecting first the largest components of the expected solution, hence orienting the reconstruction precisely towards the direction within the space where actually (3.1) holds with high probability, in a certain sense realizing a self-fulfilling prophecy: the algorithm goes only where it is supposed to work. In order to make this geometric intuition more explicit we shall modify the deterministic conditions of Theorem 3.1 accordingly, so that we cover the above setting as well.

Under slightly different deterministic conditions, we derive a recovery result very similar to Theorem 3.1:

THEOREM 3.6. *Let A be given by (3.7) and $b = A(\hat{x}) + e$, where $\hat{x} \in \mathbb{R}^d$ is the signal to be recovered and $e \in \mathbb{R}^n$ is a noise term. Suppose further that $1 \le k \le d$, $r_k(\hat{x}) \neq 0$, and $1 \le p < \infty$. If the following conditions are satisfied,*

(i)   there are constants $\alpha_k, \beta_k > 0$, such that, for all k-sparse $y \in \mathbb{R}^d$,

$$\alpha_k \|\hat{x}_{\{k\}} \hat{x}_{\{k\}}^* - yy^*\|_{HS} \le \|A(\hat{x}_{\{k\}}) - A(y)\|_{\ell_p} \le \beta_k \|\hat{x}_{\{k\}} \hat{x}_{\{k\}}^* - yy^*\|_{HS}, \tag{3.8}$$

(ii) $\hat{x} \in \mathcal{D}_\kappa$ with $\kappa < \frac{\tilde{\alpha}_k}{\sqrt{\tilde{\alpha}_k^2 + 2(\beta_k + 2L_k)^2}}$, where $0 < \tilde{\alpha}_k \le \alpha_k - 2\|e\|_{\ell_p}/r_k(\hat{x})$ and $L_k \ge 0$ with $\|A(\hat{x}) - A(\hat{x}_{\{k\}})\|_{\ell_p} \le L_k \|\hat{x}\hat{x}^* - \hat{x}_{\{k\}}\hat{x}_{\{k\}}^*\|_{HS}$,

then the ℓ_p-greedy Algorithm 1 yields a sequence $(x^{(j)})_{j=1}^k$ satisfying $\mathrm{supp}(x^{(j)}) = \mathrm{supp}(\hat{x}_{\{j\}})$ and

$$\|x^{(j)}x^{(j)^*} - \hat{x}\hat{x}^*\|_{HS} \le \|e\|_{\ell_p}/\alpha_k + \kappa^j r_1(\hat{x})\sqrt{3}(1 + \frac{\beta_k + 2L_k}{\alpha_k}).$$

If $\hat{x}$ is k-sparse, then $\|x^{(k)} - \hat{x}\| \le \|e\|_{\ell_p}/\alpha_k$.

REMARK 3.7. *Note that (3.8) resembles the restricted isometry property for rank minimization problems, in which α_k and β_k are required to be close to each other, see [30, 40] and references therein. In Theorem 3.6, we can allow for any pair of constants and compensate deviation between α_k and β_k by adding the decay condition on the signal. In other words, we shift conditions on the measurements towards conditions on the signal.*

The structure of the proof of Theorem 3.6 is almost the same as the one for Theorem 3.1, so that we first derive results similar to the Lemmas 3.3 and 3.4:

LEMMA 3.8. *If $\hat{x} \in \mathbb{R}^d$ is contained in $\mathcal{D}_\kappa$, then*

$$\|\hat{x}\hat{x}^* - \hat{x}_{\{j\}}\hat{x}_{\{j\}}^*\|_{HS} < \sqrt{2}\|\hat{x}\| r_{j+1}(\hat{x}) \frac{1}{\sqrt{1 - \kappa^2}} \le \sqrt{2}\|\hat{x}\| r_j(\hat{x}) \frac{\kappa}{\sqrt{1 - \kappa^2}}. \tag{3.9}$$

We omit the straight-forward proof and state the second lemma that is needed:
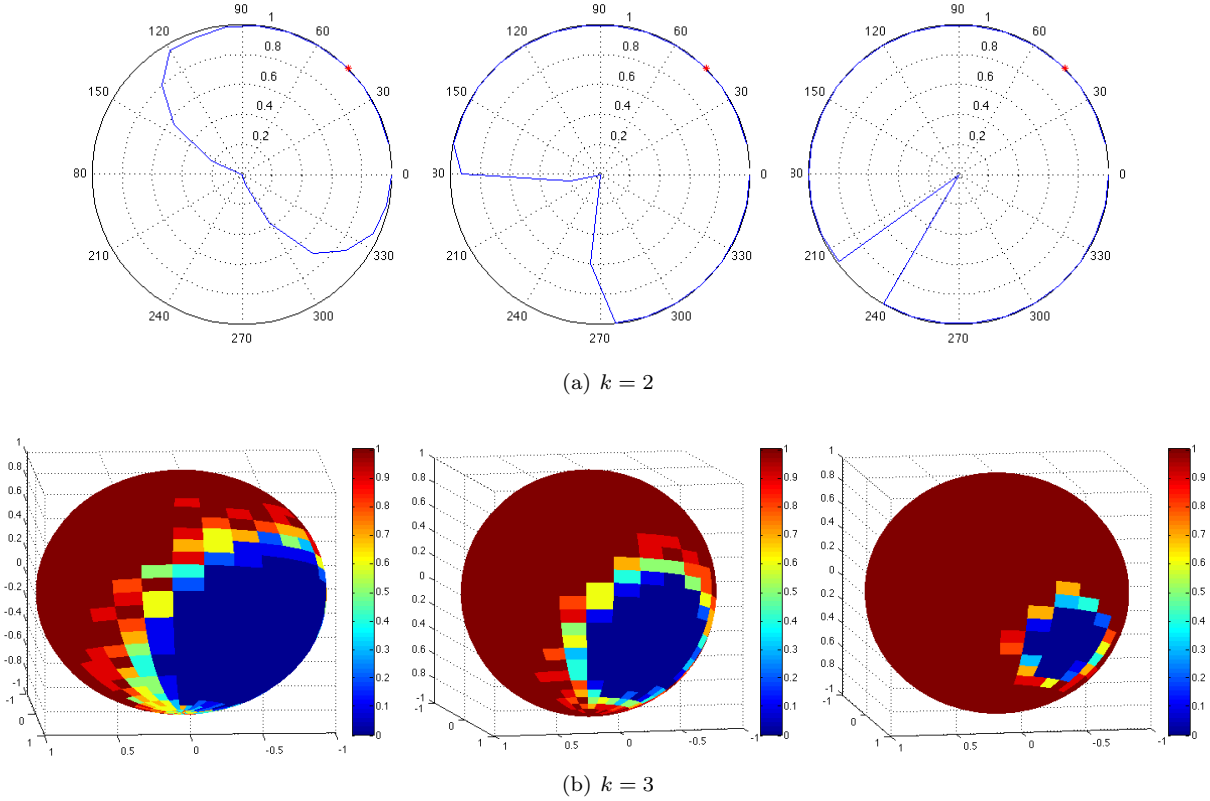
(a) $k = 2$

(b) $k = 3$

FIG. 4. *The map A is chosen as in (3.7) with independent Gaussian matrices $\{A_i\}_{i=1}^n$. We plotted the success rates of the lower bound in (3.1) for k-sparse $\hat{x} \in S^{d-1}$ marked in red and y running through S^{d-1} with the same sparsity pattern, so that both vectors can be visualized on a $k-1$-dimensional sphere (here for $k = 1, 2$). Parameters are $d = 80$, $n = 30$, and α decreases from left to right.*

LEMMA 3.9. *Fix $1 \le k \le d$ and suppose that $r_k(\hat{x}) \ne 0$. If $\hat{x} \in \mathbb{R}^d$ is contained in $\mathcal{D}_\kappa$ with $\kappa < \frac{\tilde{\alpha}}{\sqrt{\tilde{\alpha}^2 + 2(\beta + 2L)^2}}$, where $0 < \tilde{\alpha} \le \alpha - 2\|e\|_{\ell_p}/r_k(\hat{x})$, for some $\alpha, \beta, L > 0$, then, for $j = 1, \ldots, k$,*

$$\alpha \|\hat{x}\| r_j(\hat{x}) > 2\|e\|_{\ell_p} + \beta \|\hat{x}_{\{j\}} \hat{x}_{\{j\}}^* - \hat{x}_{\{k\}} \hat{x}_{\{k\}}^*\|_{HS} + 2L\|\hat{x}_{\{k\}} \hat{x}_{\{k\}}^* - \hat{x}\hat{x}^*\|_{HS} \tag{3.10}$$

Proof. [Proof of Lemma 3.9] As in the proof of Lemma 3.4, the conditions on κ imply

$$\alpha - \frac{2\|e\|_{\ell_p}}{\|\hat{x}\| r_k(\hat{x})} - \frac{\kappa}{\sqrt{1 - \kappa^2}} \sqrt{2}(\beta + 2L) > 0.$$

Multiplying by $\|\hat{x}\| r_j(\hat{x})$ and applying Lemma 3.8 yield

$$\alpha \|\hat{x}\| r_j(\hat{x}) - 2\|e\|_{\ell_p} - (\beta + 2L)\|\hat{x}_{\{j\}} \hat{x}_{\{j\}}^* - \hat{x}\hat{x}^*\|_{HS} > 0.$$

We can further estimate

$$\alpha \|\hat{x}\| r_j(\hat{x}) - 2\|e\|_{\ell_p} - \beta \|\hat{x}_{\{j\}} \hat{x}_{\{j\}}^* - \hat{x}_{\{k\}} \hat{x}_{\{k\}}^*\|_{HS} - 2L\|\hat{x}_{\{k\}} \hat{x}_{\{k\}}^* - \hat{x}\hat{x}^*\|_{HS} > 0,$$

which concludes the proof. □

Proof. [Proof of Theorem 3.6] As in the proof of Theorem 3.1, we must check that the index set selected by the ℓ_p-greedy Algorithm 1 matches the location of the nonzero entries of $\hat{x}_{\{k\}}$. Again, we use induction and the initialization $j = 0$ is trivial. Now, we suppose that $\Lambda^{(j-1)} \subset \text{supp}(\hat{x}_{\{j-1\}})$ and choose $l \notin \text{supp}(\hat{x}_{\{j\}})$.

The lower bound in (3.8) yields

$$\|A(x^{(j,l)}) - A(\hat{x})\|_{\ell_p} \geq \|A(x^{(j,l)}) - A(\hat{x}_{\{k\}})\|_{\ell_p} - \|A(\hat{x}_{\{k\}}) - A(\hat{x})\|_{\ell_p}$$

$$\geq \alpha_k \|x^{(j,l)}x^{(j,l)^*} - \hat{x}_{\{k\}}\hat{x}_{\{k\}}^*\|_{HS} - L_k \|\hat{x}_{\{k\}}\hat{x}_{\{k\}}^* - \hat{x}\hat{x}^*\|_{HS}$$

$$\geq \alpha \|\hat{x}_{\{k\}}\|r_j(\hat{x}) - L_k \|\hat{x}_{\{k\}}\hat{x}_{\{k\}}^* - \hat{x}\hat{x}^*\|_{HS},$$

which is due to $l \notin \mathrm{supp}(\hat{x}_{\{j\}})$, so that one row and one column of $x^{(j,l)}x^{(j,l)^*}$ corresponding to one of the j-largest entries of $\hat{x}$ are zero. Lemma 3.9 implies

$$\|A(x^{(j,l)}) - A(\hat{x})\|_{\ell_p} > 2\|e\|_{\ell_p} + \beta_k \|\hat{x}_{\{j\}}\hat{x}_{\{j\}}^* - \hat{x}_{\{k\}}\hat{x}_{\{k\}}^*\|_{HS} + L_k \|\hat{x}_{\{k\}}\hat{x}_{\{k\}}^* - \hat{x}\hat{x}^*\|_{HS}. \tag{3.11}$$

On the other hand, the minimizing property of $x^{(j)}$ and the Condition (3.8) imply

$$\|A(x^{(j)}) - A(\hat{x})\|_{\ell_p} \leq \|e\|_{\ell_p} + \|A(x^{(j)}) - b\|_{\ell_p}$$

$$\leq \|e\|_{\ell_p} + \|A(\hat{x}_{\{j\}}) - b\|_{\ell_p}$$

$$\leq 2\|e\|_{\ell_p} + \|A(\hat{x}_{\{j\}}) - A(\hat{x}_{\{k\}})\|_{\ell_p} + \|A(\hat{x}_{\{k\}}) - A(\hat{x})\|_{\ell_p}$$

$$\leq 2\|e\|_{\ell_p} + \beta_k \|\hat{x}_{\{j\}}\hat{x}_{\{j\}}^* - \hat{x}_{\{k\}}\hat{x}_{\{k\}}^*\|_{HS} + L_k \|\hat{x}_{\{k\}}\hat{x}_{\{k\}}^* - \hat{x}\hat{x}^*\|_{HS}.$$

The latter inequality implies with (3.11) that $x^{(j)} = x^{(j,l)}$, for all $l \in \mathrm{supp}(\hat{x}_{\{j\}})$, which concludes the part about the support.

Next, we shall verify the error bound. We obtain

$$\|x^{(j)}x^{(j)^*} - \hat{x}\hat{x}^*\|_{HS} \leq \|x^{(j)}x^{(j)^*} - \hat{x}_{\{k\}}\hat{x}_{\{k\}}^*\|_{HS} + \|\hat{x}_{\{k\}}\hat{x}_{\{k\}}^* - \hat{x}\hat{x}^*\|_{HS}$$

$$\leq 1/\alpha_k \|A(x^{(j)}) - A(\hat{x}_{\{k\}})\|_{\ell_p} + \|\hat{x}_{\{k\}}\hat{x}_{\{k\}}^* - \hat{x}\hat{x}^*\|_{HS}$$

$$\leq 1/\alpha_k \|A(x^{(j)}) - A(\hat{x})\|_{\ell_p} + 1/\alpha_k \|A(\hat{x}) - A(\hat{x}_{\{k\}})\|_{\ell_p}$$
$$+ \|\hat{x}_{\{k\}}\hat{x}_{\{k\}}^* - \hat{x}\hat{x}^*\|_{HS}$$

$$\leq 1/\alpha_k \|A(\hat{x}_{\{j\}}) - A(\hat{x})\|_{\ell_p} + \|e\|_{\ell_p}/\alpha_k + (L_k/\alpha_k + 1)\|\hat{x}\hat{x}^* - \hat{x}_{\{k\}}\hat{x}_{\{k\}}^*\|_{HS}$$

$$\leq 1/\alpha_k \|A(\hat{x}_{\{j\}}) - A(\hat{x}_{\{k\}})\|_{\ell_p} + 1/\alpha_k \|A(\hat{x}_{\{k\}}) - A(\hat{x})\|_{\ell_p} + \|e\|_{\ell_p}/\alpha_k$$
$$+ (L_k/\alpha_k + 1)\|\hat{x}\hat{x}^* - \hat{x}_{\{k\}}\hat{x}_{\{k\}}^*\|_{HS}$$

$$\leq \beta_k/\alpha_k \|\hat{x}_{\{j\}}\hat{x}_{\{j\}}^* - \hat{x}_{\{k\}}\hat{x}_{\{k\}}^*\|_{HS} + \|e\|_{\ell_p}/\alpha_k$$
$$+ (2L_k/\alpha_k + 1)\|\hat{x}\hat{x}^* - \hat{x}_{\{k\}}\hat{x}_{\{k\}}^*\|_{HS}$$

$$\leq (\beta_k/\alpha_k + 2L_k/\alpha_k + 1)\|\hat{x}\hat{x}^* - \hat{x}_{\{j\}}\hat{x}_{\{j\}}^*\|_{HS} + \|e\|_{\ell_p}/\alpha_k$$

$$\leq (\beta_k/\alpha_k + 2L_k/\alpha_k + 1)\sqrt{2}r_{j+1}(\hat{x})\frac{1}{\sqrt{1-\kappa^2}} + \|e\|_{\ell_p}/\alpha_k$$

$$\leq (\beta_k/\alpha_k + 2L_k/\alpha_k + 1)\kappa^j\sqrt{2}r_1(\hat{x})\frac{1}{\sqrt{1-\kappa^2}} + \|e\|_{\ell_p}/\alpha_k.$$

Some rough estimates yield

$$\frac{\beta_k/\alpha_k + 2L_k/\alpha_k + 1}{\sqrt{1-\kappa^2}}\sqrt{2} \leq \sqrt{3}(1 + \frac{\beta_k + 2L_k}{\alpha_k}),$$

which concludes the proof. □

REMARK 3.10. *The greedy Algorithm 1 can also be performed in the complex setting. The complex version of Theorem 3.6 also holds when recovery of $\pm\hat{x}$ is replaced by a complex unit vector times $\hat{x}$.*

3.4. Signal recovery from random measurements. We aim at choosing $\{A_i : i = 1, \ldots, n\}$ in a suitable random way, so that the conditions in Theorem 3.6 are satisfied with high probability. Indeed, the upper bound in (3.8) is always satisfied for some $C > 0$, because we are in a finite-dimensional regime. In this section we are interested in one rank matrices $A_i = a_i a_i^*$, for some vector $a_i \in \mathbb{R}^d$, $i = 1, \ldots, n$, because then

$$A(\hat{x}) = (\hat{x}^* A_1 \hat{x}, \ldots, \hat{x}^* A_n \hat{x})^* = (|\langle a_1, \hat{x}\rangle|^2, \ldots, |\langle a_n, \hat{x}\rangle|^2)^\top \tag{3.12}$$

models the phase retrieval problem.

3.4.1. Real random measurement vectors. To check on the assumptions in Theorem 3.6, we shall draw at random the measurement vectors $\{a_i : i = 1, \ldots, n\}$ from probability distributions to be characterized next. We say that a random vector $a \in \mathbb{R}^d$ satisfies the *small-ball assumption* if there is a constant $c > 0$, such that, for all $z \in \mathbb{R}^d$ and $\varepsilon > 0$,

$$\mathbb{P}\big(|\langle a, z \rangle| \leq \varepsilon \|z\|\big) \leq c\varepsilon.$$

Moreover, we say that a is *isotropic* if $\mathbb{E}|\langle a, z \rangle|^2 = \|z\|^2$, for all $z \in \mathbb{R}^d$. The vector a is said to be *L-subgaussian* if, for all $z \in \mathbb{R}^d$ and $t \geq 1$,

$$\mathbb{P}\big(|\langle a, z \rangle| \geq tL\|z\|\big) \leq 2e^{-t^2/2}.$$

Eldar and Mendelson derived the following result:

THEOREM 3.11 ([23]). *Let $\{a_i : i = 1, \ldots, n\}$ be a set of independent copies of a random vector $a \in \mathbb{R}^d$ that is isotropic, L-subgaussian, and satisfies the small-ball assumption. Then there are positive constants $c_1, \ldots, c_4$ such that, for all $t \geq c_1$ with $n \geq kc_2 t^3 \log(ed/k)$, and for all k-sparse $x, y \in \mathbb{R}^d$,*

$$\sum_{i=1}^{n} \big||\langle a_i, x \rangle|^2 - |\langle a_i, y \rangle|^2\big| \geq c_4 \|x - y\|\|x + y\| \tag{3.13}$$

with probability of failure at most $2e^{-kc_3 t^2 \log(ed/k)}$.

The uniform distribution on the sphere and the Gaussian distribution on $\mathbb{R}^d$ induce random vectors satisfying the assumptions of Theorem 3.11. However, at first glance, the above theorem does not help us directly, because we are seeking for an estimate involving the Hilbert-Schmidt norm. It is remarkable though that

$$\|xx^* - yy^*\|_{HS} \leq \|x - y\|\|x + y\| \leq \sqrt{2}\|xx^* - yy^*\|_{HS}. \tag{3.14}$$

The relation (3.14) then yields that also the lower bound in (3.8) is satisfied for any constant $\alpha \leq c_4$, so that Theorem 3.6 can be used for $p = 1$. Thus, if the signal $\hat{x}$ is sparse and satisfies decay conditions matching the constants in (3.8), then Algorithm 1 for $p = 1$ recovers $x^{(k)} = \pm\hat{x}$.

3.4.2. Complex random measurement vectors. In the following and at least for the uniform distribution on the sphere, we shall generalize Theorem 3.11 to the complex setting, so that the assumptions of Theorem 3.6 hold with high probability.

THEOREM 3.12. *If $\{a_i : i = 1, \ldots, n\}$ are independent uniformly distributed vectors on the unit sphere, then there is a constant $\alpha > 0$ such that, for all k-sparse $x, y \in \mathbb{C}^d$ and $n \geq c_1 k \log(ed/k)$,*

$$\sum_{i=1}^{n} \big||\langle a_i, x \rangle|^2 - |\langle a_i, y \rangle|^2\big| \geq \alpha n \|xx^* - yy^*\|_{HS}$$

with probability of failure at most e^{-nc_2}.

Proof. For fixed $x, y \in \mathbb{R}^d$, the results in [13] imply that there are constants $c_1, c > 0$ such that, for all $t > 0$,

$$\sum_{i=1}^{n} \big||\langle a_i, x \rangle|^2 - |\langle a_i, y \rangle|^2\big| \geq 1/\sqrt{(2)}(c_1 - t)n\|xx^* - yy^*\|_{HS} \tag{3.15}$$

with probability of failure at most $2e^{-nct^2}$. If both x and y are k-sparse, then the union of their supports induces an at most $2k$-dimensional coordinate subspace, so that also $xx^* - yy^*$ can be reduced to a $2k \times 2k$ matrix, by eliminating rows and columns that do not belong to the indices of the subspace. Results in [13] can be used to derive that the estimate (3.15) holds uniformly for elements in this subspace when $n \geq c_3 t^{-2} \log(t^{-1})2k$, for some constant $c_3 > 0$. There are at most $\binom{d}{2k} \leq (ed/(2k))^{2k}$ many of such coordinate subspaces, see also [4]. Therefore, by a union bound, the probability of failure is at most

$$(ed/(2k))^{2k}2e^{-nct^2} = 2e^{-n\left(ct^2 - \frac{2k \log(ed/(2k))}{n}\right)}.$$

Thus, if also $n \geq \frac{1}{ct^2 - c_2} \log(ed/2k)2k$ with $ct^2 - c_2 > 0$, then we have the desired result. $\square$

REMARK 3.13. *We want to point out that the use of the term $\|x - y\|\|x + y\|$ limits Theorem 3.11 to the real setting. Our observation (3.14) was the key to derive the analog result in the complex setting.*

3.4.3. Rank-m projectors as measurements. A slightly more general phase retrieval problem was discussed in [2], where the measurement A in (3.7) is given by $A_i = \frac{d}{m}P_{V_i}$, $i = 1, \ldots, n$, and each P_{V_i} is an orthogonal projector onto an m-dimensional linear subspace V_i of $\mathbb{R}^d$. The set of all m-dimensional linear subspaces $\mathcal{G}_{m,d}$ is a manifold endowed with the standard normalized Haar measure σ_m:

THEOREM 3.14 ([2]). *There is a constant $u_m > 0$, only depending on m, such that the following holds: for $0 < r < 1$ fixed, there exist constants $c(r), C(r) > 0$, such that, for all $n \geq c(r)d$ and $\{V_j : j = 1, \ldots n\} \subset \mathcal{G}_{m,d}$ independently chosen random subspaces with identical distribution σ_m, the inequality*

$$\sum_{i=1}^{n} \left| x^*A_i x - y^*A_i y \right| \geq u_m(1-r)n\|xx^* - yy^*\|_2, \tag{3.16}$$

for all $x, y \in \mathbb{R}^d$, holds with probability of failure at most $e^{-C(r)n}$. Since the rank of $xx^* - yy^*$ is at most 2, its Hilbert-Schmidt norm is bounded by $\sqrt{2}$ times the operator norm. Thus, we have the lower ℓ_1-bound in (3.8) for $k = d$ when $n \geq c(r)d$, and the random choice of subspaces (and hence orthogonal projectors) enables us to apply Theorem 3.6. The result for k-sparse signals is a consequence of Theorem 3.14:

COROLLARY 3.15. *There is a constant $u_m > 0$, only depending on m, such that the following holds: for $0 < r < 1$ fixed, there exist constants $c(r), C(r) > 0$, such that, for all $n \geq c(r)k\log(ed/k)$ and $\{V_j : j = 1, \ldots n\} \subset \mathcal{G}_{m,d}$ independently chosen random subspaces with identical distribution σ_m, the inequality*

$$\sum_{i=1}^{n} \left| x^*A_i x - y^*A_i y \right| \geq u_m(1-r)n\|xx^* - yy^*\|_{HS}, \tag{3.17}$$

for all k-sparse $x, y \in \mathbb{R}^d$, holds with probability of failure at most $e^{-C(r)n}$.

Proof. The lower bound on n in Theorem 3.14 is not needed when the vectors x and y are fixed in (3.16), cf. [2]. If both x, y are supported in one fixed coordinate subspace of dimension $2k$, then the proof of Theorem 3.14 in [2], see also [13], yields that (3.17) holds uniformly for this subspace provided $n \geq c(r)2k$.

Similar to the proof of Theorem 3.12, we shall use Theorem 3.14 with a $2k$ coordinate subspace and then apply a union bound by counting the number of such subspaces. Indeed, since $xx^* - yy^*$ can be treated as a $2k \times 2k$ matrix by removing zero rows and columns, Theorem 3.14 implies (3.17) for all $x, y \in \mathbb{R}^d$ supported in a fixed coordinate subspace of dimension $2k$ with probability of failure at most $e^{-C(r)n}$ when $n \geq c(r)2k$. Again, we have used that $xx^* - yy^*$ has rank at most two, so that the Hilbert-Schmidt norm is bounded by $\sqrt{2}$ times the operator norm. The remaining part can be copied from the end of the proof of Theorem 3.12. □

REMARK 3.16. *It is mentioned in [2] already that Theorem 3.14 also holds in the complex setting. Therefore, Corollary 3.15 has a complex version too, and our present results hold for complex rank-m projectors.*

3.4.4. Nearly isometric random maps. To conclude the discussion on random measurements for phase retrieval, we shall generalize some results from [40] to sparse vectors. Also, we want to present a framework, in which $\{A_i : i = 1, \ldots, n\}$ in (3.12) can be chosen as a set of independent random matrices with independent Gaussian entries. Let $\mathcal{A}$ be a random map that takes values in linear maps from $\mathbb{R}^{d \times d}$ to $\mathbb{R}^n$. Then $\mathcal{A}$ is called nearly isometrically distributed if, for all $X \in \mathbb{R}^{d \times d}$,

$$\mathbb{E}\|\mathcal{A}(X)\|^2 = \|X\|_{HS}^2, \tag{3.18}$$

and, for all $0 < \epsilon < 1$, we have

$$(1 - \epsilon)\|X\|_{HS}^2 \leq \|\mathcal{A}(X)\|^2 \leq (1 + \epsilon)\|X\|_{HS}^2 \tag{3.19}$$

with probability of failure at most $2e^{-nf(\epsilon)}$, where $f : (0,1) \to \mathbb{R}_{>0}$ is an increasing function.

Note that the definition of nearly isometries in [40] is more restrictive, but we can find several examples there. For instance, if $\{A_i : i = 1, \ldots, n\}$ in (3.12) are independent matrices with independent standard Gaussian entries, then the map

$$\mathcal{A} : \mathbb{R}^{d \times d} \to \mathbb{R}^n, \qquad \mathcal{A}(X) := \frac{1}{\sqrt{n}} \begin{pmatrix} \text{trace}(A_1^*X) \\ \vdots \\ \text{trace}(A_n^*X) \end{pmatrix}$$

is nearly isometrically distributed, see [17, 40].

The following theorem fits into our setting, and it should be mentioned that we will only use (3.18) and (3.19) for symmetric matrices X of rank at most 2. So, we could even further weaken the notion of nearly isometric distributions accordingly.

THEOREM 3.17. *Fix* $1 \leq k \leq d$. *If* $\mathcal{A}$ *is a nearly isometric random map from* $\mathbb{R}^{d \times d}$ *to* $\mathbb{R}^n$ *and* $A(x) := \mathcal{A}(xx^*)$, *then there are constants* $c_1, c_2 > 0$, *such that, uniformly for all* $0 < \delta < 1$ *and all* k-*sparse* $x, y \in \mathbb{R}^d$,

$$(1-\delta)\|xx^* - yy^*\|_{HS} \leq \|A(x) - A(y)\| \leq (1+\delta)\|xx^* - yy^*\|_{HS} \qquad (3.20)$$

with probability of failure at most $2e^{-n\left(f(\delta/2) - \frac{4k+2}{n}\log(65/\delta) - \frac{2k}{n}\log(ed/2k)\right)}$.

As a consequence of Theorem 3.17, we can fix δ and derive two constants $c_1, c_2 > 0$ depending only on δ and $f(\delta/2)$, such that (3.20) holds for all k-sparse $x, y \in \mathbb{R}^d$ in a uniform fashion with probability of failure at most $e^{-c_1 n}$ whenever $n \geq c_2 k \log(ed/k)$. The analogous result for not necessarily sparse vectors is derived in [40].

Proof. We fix an index set $\mathcal{I}$ of $2k$ coordinates in $\mathbb{R}^d$, denote the underlying coordinate subspace by V, and define

$$\mathcal{X} = \{X \in \mathbb{R}^{d \times d} : X = X^*, \ \text{rank}(X) \leq 2, \ \|X\|_{HS} = 1, \ X_{i,j} = 0, \ \text{if } i \notin \mathcal{I} \text{ or } j \notin \mathcal{I}\}.$$

Any element $X \in \mathcal{X}$ can be written as $X = axx^* + byy^*$ such that $a^2 + b^2 = 1$ and $x, y \in V$ are orthogonal unit norm vectors. In order to build a covering of $\mathcal{X}$, we start with a covering of the $2k$-dimensional unit sphere S_V in V. Indeed, there is a finite set $\mathcal{N}_1 \subset S_V$ of cardinality at most $(1 + \frac{64}{\delta})^{2k} \leq (\frac{65}{\delta})^{2k}$, such that, for every $x \in S_V$, there is $y \in \mathcal{N}_1$ with $\|x - y\| \leq \delta/32$, see, for instance, [45, Lemma 5.2]. We can also uniformly cover $[-1, 1]$ with a finite set of cardinality $\frac{32}{\delta}$, so that the error is bounded by $\delta/16$. Thus, we can cover $\mathcal{X}$ with a set $\mathcal{N}$ of cardinality at most $(\frac{32}{\delta})^2 (\frac{65}{\delta})^{4k} \leq (\frac{65}{\delta})^{4k+2}$, such that, for every $X = axx^* + byy^* \in \mathcal{X}$, there is $Y = a_0 x_0 x_0 + b_0 y_0 y_0^* \in \mathcal{N}$ with

$$|a - a_0| \leq \delta/16, \quad |b - b_0| \leq \delta/16, \quad \|x - x_0\| \leq \delta/32, \quad \|y - y_0\| \leq \delta/32. \qquad (3.21)$$

We can choose $\varepsilon = \delta/2$, so that (3.19) holds uniformly on $\mathcal{N}$ with probability of failure at most $2(\frac{65}{\delta})^{4k+2} e^{-nf(\delta/2)}$. Taking the square root yields with at most the same probability of failure that

$$1 - \delta/2 \leq \|\mathcal{A}(Y)\| \leq 1 + \delta/2$$

holds uniformly for all $Y \in \mathcal{N}$.

We now define the random variable

$$M = \max\{K \geq 0 : \|\mathcal{A}(X)\| \leq K\|X\|_{HS}, \ \text{for all } X \in \mathcal{X}\} \qquad (3.22)$$

and consider an arbitrary $X = axx + byy^* \in \mathcal{X}$. Then there is $Y = a_0 x_0 x_0 + b_0 y_0 y_0^* \in \mathcal{N}$ such that (3.21) is satisfied, so that we can further estimate

$$\|\mathcal{A}(X)\| \leq \|\mathcal{A}(Y)\| + \|\mathcal{A}(axx + byy^* - a_0 x_0 x_0 - b_0 y_0 y_0^*)\|$$
$$\leq 1 + \delta/2 + \|\mathcal{A}(axx - a_0 x_0 x_0^*)\| + \|\mathcal{A}(byy - b_0 y_0 y_0^*)\|.$$

Although $axx - a_0 x_0 x_0^*$ and $byy - b_0 y_0 y_0^*$ may not be elements in $\mathcal{X}$, a simple normalization allows us to apply the bound (3.22), so that we obtain

$$\|\mathcal{A}(X)\| \leq 1 + \delta/2 + M\|axx - a_0 x_0 x_0^*\|_{HS} + M\|byy - b_0 y_0 y_0^*\|_{HS}$$
$$\leq 1 + \delta/2 + M|a|\|xx^* - x_0 x_0^*\|_{HS} + M|a - a_0|\|x_0 x_0^*\|_{HS} +$$
$$M|b|\|yy^* - y_0 y_0^*\|_{HS} + M|b - b_0|\|y_0 y_0^*\|_{HS}$$
$$\leq 1 + \delta/2 + M\delta/4.$$

For the last inequality, we have used (3.14). By choosing $X \in \mathcal{X}$ with $\|\mathcal{A}(X)\| = M$, we derive $M \leq 1 + \delta/2 + M\delta/4$, which implies $M \leq 1 + \delta$. Thus, we obtain $\|\mathcal{A}(X)\| \leq 1 + \delta$.
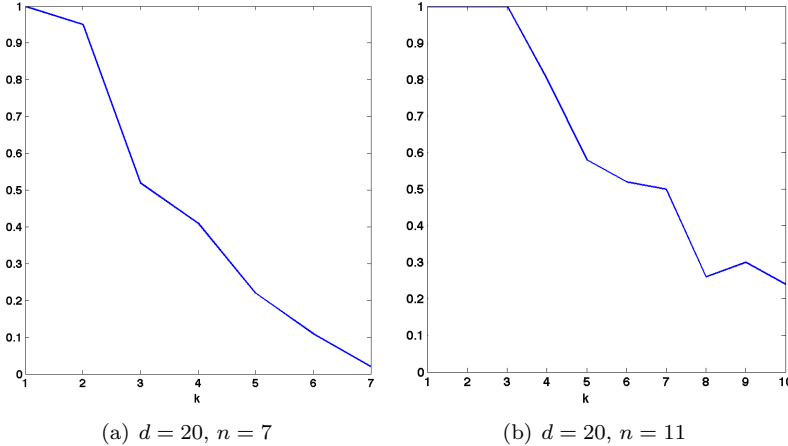
(a) $d = 20$, $n = 7$ (b) $d = 20$, $n = 11$

FIG. 5. *Signal reconstruction rates vs. sparsity k. Reconstruction is repeated* 50 *times for each signal and k to derive stable recovery rates. For small k, we almost certainly recover the signal, and, as expected, increasing k leads to decreased success rates.*

The lower bound is similarly derived by

$$\|\mathcal{A}(X)\| \geq \|\mathcal{A}(Y)\| - \|\mathcal{A}(Y - X)\| \geq 1 - \delta/2 - (1 + \delta)\delta/4 \geq 1 - \delta.$$

So far, we have the estimate (3.20) in a uniform fashion for all x, y in some fixed coordinate subspace of dimension $2k$ with probability of failure at most $2(\frac{65}{\delta})^{4k+2} e^{-nf(\delta/2)}$. Again, we derive the union bound by counting subspaces as in the proof of Theorem 3.12. There are at most $(ed/(2k))^{2k}$ many subspaces, so that we can conclude the proof. $\square$

3.5. Numerical experiments for greedy phase retrieval. We shall follow the model in (3.12) and study signal reconstruction rates for random choices of measurement vectors $\{a_i : i = 1, \ldots n\}$ chosen as independent standard Gaussian vectors. For a fixed number of measurements n, we shall study the signal recovery rate depending on the sparsity k. We expect to have high success rates for small k and decreased rates when k grows. Figure 5 shows results of numerical experiments consistent with our theoretical findings. We must point out though that each step of the greedy algorithm requires solving a nonconvex global optimization problem. Here, we used standard optimization routines that may yield results that are not optimal. Better outcomes can be expected when applying more sophisticated global optimization methods, for instance, based on adaptive grids and more elaborate analysis of functions of few parameters in high dimensions, cf. [16].

4. Iterative thresholding for quasi-linear problems. Although the quasi-linear structure of the measurements does not play an explicit role in the formulation of Algorithm 1, in the examples we showed in the previous sections it was nevertheless a crucial aspect to obtain generalized RIP conditions such as (3.1) and (3.8). When using iterative thresholding algorithms, as we shall show below, the quasi-linear structure of the measurements gets into the formulation of the algorithms as well. Hence, from now on we shall be a bit more explicit on the form of nonlinearity and we consider a map $F : \mathbb{R}^d \to \mathbb{R}^{n \times d}$, and aim to reconstruct $\hat{x} \in \mathbb{R}^d$ from measurements $b \in \mathbb{R}^n$ given by

$$F(\hat{x})\hat{x} = b.$$

As guiding examples, we keep as references the maps A in Proposition 3.5, which can be written as $A(x) = F(x)x$, where $F(x) = A_1 + \epsilon f(\|x - x_0\|)A_2$.

The study of iterative thresholding algorithms that we propose below is motivated by the intrinsic limitations of Algorithm 1, in particular, its restriction to recovering only signals which have a rapid decay of their nonincreasing rearrangement, and the potential complexity explosion due to the need of performing several high-dimensional global optimizations at each greedy step.

4.1. Iterative hard-thresholding. We follow ideas in [8] and aim to use the iterative scheme

$$x^{(j+1)} := \left(x^{(j)} + \frac{1}{\mu_k}F(x^{(j)})^*(b - F(x^{(j)})x^{(j)})\right)_{\{k\}}, \tag{4.1}$$

where $\mu_k > 0$ is some parameter and, as before, $y_{\{k\}}$ denotes the best k-sparse approximation of $y \in \mathbb{R}^d$. The following theorem is a reformulation of a result by Blumensath in [8]:

THEOREM 4.1. *Let $b = A(\hat{x}) + e$, where $\hat{x} \in \mathbb{R}^d$ is the signal to be recovered and $e \in \mathbb{R}^n$ is a noise term. Suppose $1 \leq k \leq d$ is fixed. If F satisfies the following assumptions,*

(i)   there is $c > 0$ such that $\|F(\hat{x})\|_2 \leq c$,

(ii) there are $\alpha_k, \beta_k > 0$ such that, for all k-sparse $x, y, z \in \mathbb{R}^d$,

$$\alpha_k\|x - y\| \leq \|F(z)(x - y)\| \leq \beta_k\|x - y\|, \tag{4.2}$$

(iii) there is $C_k > 0$ such that, for all k-sparse $y \in \mathbb{R}^d$,

$$\|F(\hat{x}_{\{k\}}) - F(y)\|_2 \leq C_k\|\hat{x}_{\{k\}} - y\|, \tag{4.3}$$

(iv) there is $L_k > 0$ such that $\|F(\hat{x})\hat{x} - F(\hat{x}_{\{k\}})\hat{x}_{\{k\}}\| \leq L_k\|\hat{x} - \hat{x}_{\{k\}}\|$,

(v)   the constants satisfy $\beta_k^2 \leq 1/\mu_k < \frac{3}{2}\alpha_k^2 - 4\|\hat{x}_{\{k\}}\|^2 C_k^2$,

then the iterative scheme (4.1) *converges towards $x^\star$ satisfying*

$$\|x^\star - \hat{x}\| \leq \gamma\|e\| + (1 + \gamma c + \gamma L_k)\|\hat{x} - \hat{x}_{\{k\}}\|,$$

where $\gamma = \frac{2}{0.75\alpha_k^2 - 1/\mu_k - 2\|\hat{x}_{\{k\}}\|^2 C_k^2}$.

Note that (4.2) is again a RIP condition for each $F(z)$. The proof of Theorem 4.1 is based on Blumensath's findings in [8], where the nonlinear operator A is replaced by its first order approximation at $x^{(j)}$ within the iterative scheme. When dealing with the quasi-linear setting, it is natural to use $F(x^{(j)})$, so we formulated the iteration in this way already.

Proof. We first verify that the assumptions in [8, Corollary 2] are satisfied. By using (4.3), we derive, for k-sparse $y \in \mathbb{R}^d$,

$$\begin{aligned}\|F(\hat{x}_{\{k\}})\hat{x}_{\{k\}} - F(y)y - F(y)(\hat{x}_{\{k\}} - y)\| &= \|F(\hat{x}_{\{k\}})\hat{x}_{\{k\}} - F(y)\hat{x}_{\{k\}}\| \\ &\leq \|\hat{x}_{\{k\}}\|C_k\|\hat{x}_{\{k\}} - y\|.\end{aligned}$$

The assumptions of [8, Corollary 2] are satisfied, which implies that the iterative scheme (4.1) converges to some k-sparse $x^\star$ satisfying

$$\|x^\star - \hat{x}\| \leq \gamma\|b - F(\hat{x}_{\{k\}})\hat{x}_{\{k\}}\| + \|\hat{x}_{\{k\}} - \hat{x}\|,$$

where $\gamma = \frac{2}{0.75\alpha_k^2 - 1/\mu_k - 2\|\hat{x}_{\{k\}}\|^2 C_k^2}$. We still need to estimate the left term on the right-hand side. A zero addition yields

$$\begin{aligned}\|b - F(\hat{x}_{\{k\}})\hat{x}_{\{k\}}\| &\leq \|e\| + \|F(\hat{x})\hat{x} - F(\hat{x}_{\{k\}})\hat{x}_{\{k\}}\| \\ &\leq \|e\| + L_k\|\hat{x} - \hat{x}_{\{k\}}\|,\end{aligned}$$

which concludes the proof. $\square$

We shall verify that the map in Proposition 3.5 satisfies the assumptions of Theorem 4.1 at least when $\hat{x}$ is k-sparse:

EXAMPLE 4.2. *Let $F(x) = A_1 + \epsilon f(\|x - x_0\|)A_2$, so that $F(x)x = A(x)$ with A as in Proposition 3.5. By a similar proof and under the same notations we derive that an upper bound in (4.2) can be chosen as $\beta = 1 + \delta + \epsilon B\|A_2\|$, where δ is the RIP constant for A_1. For the lower bound, we compute $\alpha = 1 - \delta - \epsilon B\|A_2\|$. In other words, $F(x)$ satisfies the RIP of order k with constant $\delta + \epsilon B\|A_2\|$. In (4.3), we can choose $C = \epsilon L\|A_2\|$. Thus, if $\epsilon, B, \|A_2\|, \delta$ are sufficiently small, then the assumptions of Theorem 4.1 are satisfied.*

4.2. Iterative soft-thresholding. The type of thresholding in the scheme (4.1) of the previous section is one among many potential choices. Here, we shall discuss soft-thresholding, which is widely applied when dealing with linear compressed sensing problems. Our findings in this section are based on preliminary results in [42]. We suppose that the original signal is sparse and, in fact, we aim to reconstruct the sparsest $\hat{x}$ that matches the data $b = F(\hat{x})\hat{x}$. In other words, we intend to solve for $\hat{x} \in \mathbb{R}^d$ with $\hat{x} = \arg\min \|x\|_{\ell_0}$ subject to $F(x)x = b$. Theorem 4.1 yields that we can use iterative hard-thresholding to reconstruct the sparsest solution. Here, we shall follow a slightly different strategy. As ℓ_0-minimization is a combinatorial optimization problem and computationally cumbersome in principle, even NP-hard under many circumstances, it is common practice in compressed sensing to replace the ℓ_0-pseudo norm with the ℓ_1-norm, so that we consider the problem

$$\arg\min \|x\|_{\ell_1} \quad \text{subject to} \quad F(x)x = b. \tag{4.4}$$

It is also somewhat standard to work with an additional relaxation of it and instead solve for $\hat{x}_\alpha$ given by

$$\hat{x}_\alpha := \arg\min_{x \in \mathbb{R}^d} \mathcal{J}_\alpha(x), \qquad \text{where} \qquad \mathcal{J}_\alpha(x) := \|F(x)x - b\|_{\ell_2}^2 + \alpha\|x\|_{\ell_1}, \tag{4.5}$$

where $\alpha > 0$ is sometimes called the relaxation parameter. The optimization (4.5) allows for $F(x)x \neq b$, hence, is particularly beneficial when we deal with measurement noise, so that $b = F(\hat{x})\hat{x} + e$ and α can be suitably chosen to compensate for the magnitude of e. If there is no noise term, then (4.5) approximates (4.4) when α tends to zero. The latter is a standard result but we explicitly state this and prove it for the sake of completeness:

PROPOSITION 4.3. *Let the map $x \mapsto F(x)$ be continuous and suppose that $(\alpha_n)_{n=1}^\infty$ is a sequence of nonnegative numbers that converge towards 0. If $(\hat{x}_{\alpha_n})_{n=1}^\infty$ is any sequence of minimizers of $\mathcal{J}_{\alpha_n}$, then it contains a subsequence that converges towards a minimizer of (4.4). If the minimizer of (4.4) is unique, then the entire sequence $(\hat{x}_{\alpha_n})_{n=1}^\infty$ converges towards this minimizer.*

Proof. Let $\hat{x}$ be a minimizer of (4.4). Direct computations yield

$$\|\hat{x}_{\alpha_n}\|_{\ell_1} \leq \frac{1}{\alpha_n}\mathcal{J}_{\alpha_n}(\hat{x}_{\alpha_n}) \leq \frac{1}{\alpha_n}\mathcal{J}_{\alpha_n}(\hat{x}) = \|\hat{x}\|_{\ell_1}. \tag{4.6}$$

Thus, there is a convergent subsequence $(\hat{x}_{\alpha_{n_j}})_{j=1}^\infty \to \bar{x} \in \mathbb{R}^d$, for $j \to \infty$. Since

$$\|F(\bar{x})\bar{x} - y\| = \lim_{j\to\infty}\|F(\hat{x}_{\alpha_{n_j}})\hat{x}_{\alpha_{n_j}} - y\| \leq \lim_{j\to\infty}\mathcal{J}_{\alpha_{n_j}}(\hat{x}_{\alpha_{n_j}}) \leq \lim_{j\to\infty}\mathcal{J}_{\alpha_{n_j}}(\hat{x}) = 0$$

and (4.6) hold, $\bar{x}$ must be a minimizer of (4.4).

Now, suppose that the minimizer $\hat{x}$ of (4.4) is unique. If x_0 is an accumulation point of $(\hat{x}_{\alpha_n})_{j=1}^\infty$, then there is a subsequence converging towards x_0. The same arguments as above with the uniqueness assumption yield $x_0 = \hat{x}$, so that $(\hat{x}_{\alpha_n})_{j=1}^\infty$ is a bounded sequence with only one single accumulation point. Hence, the entire sequence converges towards $\hat{x}$. $\square$ From here on, we shall focus on (4.5), which we aim to solve at least approximately using some iterative scheme. First, we define the map

$$\mathscr{S}_\alpha : \mathbb{R}^d \to \mathbb{R}^d, \qquad x \mapsto \mathscr{S}_\alpha(x) := \arg\min_{y\in\mathbb{R}^d}\|F(x)y - b\|^2 + \alpha\|y\|_{\ell_1}.$$

To develop the iterative scheme, we present some conditions, so that $\mathscr{S}_\alpha$ is contractive:

THEOREM 4.4. *Given $b \in \mathbb{R}^n$, fix $\alpha > 0$ and suppose that there are constants $c_1, c_2, c_3, \gamma > 0$ such that, for all $x, y \in \mathbb{R}^d$,*

 (i) $\|F(x)\|_2 \leq c_1$,
 (ii) there is $z_x \in \mathbb{R}^d$ such that $\|z_x\|_{\ell_1} \leq c_2\|b\|$ and $F(x)z_x = b$,
(iii) $\|F(x) - F(y)\|_2 \leq c_3\|x - y\|$,
(iv) if y is $\frac{4}{\alpha^2}(c_1 + c_2 + c_1^2 c_2)^2\|b\|^2$-sparse, then

$$(1 - \gamma)\|y\|^2 \leq \|F(x)y\|^2 \leq (1 + \gamma)\|y\|^2,$$

 (v) the constants satisfy $\gamma < 1 - (1 + 2c_1 c_2)c_3\|b\|$,

then $\mathscr{S}_\alpha$ is a bounded contraction, so that the recursive scheme $x_\alpha^{(j+1)} := \mathscr{S}_\alpha(x_\alpha^{(j)})$ converges for any initial vector towards a point x_α satisfying the fixed point relationship

$$x_\alpha = \arg\min_{y \in \mathbb{R}^d} \|F(x_\alpha)y - b\|^2 + \alpha\|y\|_{\ell_1}. \tag{4.7}$$

REMARK 4.5. *We believe that the fixed point of (4.7) in Theorem 4.4 is close to the actual minimizer of (4.5). To support this point of view, we shall later investigate on the distance $\|x_\alpha - \hat{x}_\alpha\|$ in Theorem 4.8 and also provide some numerical experiments in Section 4.3.*

A few more comments are in order before we take care of the proof: Note that the constant c_1 must hold for the operator norm in (i), and γ in the RIP of (iv) covers only sparse vectors. Therefore, $1 + \gamma$ can be much smaller than c_1. Condition (iii) is a standard Lipschitz property. If γ is indeed less than 1, then small data b can make up for larger other constants, so that (v) can hold. The requirement (ii) is more delicate though and a rough derivation goes as follows: the data b are supposed to lie in the range of $F(x)$, which is satisfied, for instance, if $F(x)$ is onto. The pseudo-inverse of $F(x)$ then yields a vector z_x with minimal ℓ_2-norm. We can then ask for boundedness of all operator norms of the pseudo-inverses. However, we still need to bound the ℓ_1-norm by using the ℓ_2-norm, which introduces an additional factor $\sqrt{d}$.

We introduce the soft-thresholding operator $\mathbb{S}_\alpha : \mathbb{R}^d \to \mathbb{R}^d$, $x \mapsto \mathbb{S}_\alpha(x)$ given by

$$(\mathbb{S}_\alpha(x))_i = \begin{cases} x_i - \alpha/2, & \alpha/2 \le x_i \\ 0, & -\alpha/2 < x_i < \alpha/2 \ , \\ x_i + \alpha/2, & x_i \le -\alpha/2 \end{cases} \tag{4.8}$$

which we shall use in the following proof:

Proof. [Proof of Theorem 4.4] For $x \in \mathbb{R}^d$, we can apply (ii), so that $F(x)z_x = b$ and $\|z_x\|_{\ell_1} \le c_2\|b\|$, implying

$$\alpha\|\mathscr{S}_\alpha(x)\| \le \|F(x)\mathscr{S}_\alpha(x) - b\|^2 + \alpha\|\mathscr{S}_\alpha(x)\|_{\ell_1} \le \alpha\|z_x\|_{\ell_1} \le \alpha c_2\|b\|.$$

Thus, we have $\|\mathscr{S}_\alpha(x)\| \le c_2\|b\|$.

The conditions (i) and (iii) imply

$$\|F(x)^*F(x) - F(y)^*F(y)\| \le 2c_1c_3\|x - y\|. \tag{4.9}$$

It is well-known that

$$\mathscr{S}_\alpha(x) = \mathbb{S}_\alpha\big(\xi(x)\big), \qquad \text{where} \qquad \xi(x) = (I - F(x)^*F(x))\mathscr{S}_\alpha(x) + F(x)^*b, \tag{4.10}$$

and $\mathbb{S}_\alpha$ is the soft-thresholding operator in (4.8), cf. [18]. Note that ξ can be bounded by

$$\|\xi(x)\| = \|(I - F(x)^*F(x))\mathscr{S}_\alpha(x)\| + \|F(x)^*b\|$$
$$\le c_2\|b\| + c_1^2 c_2\|b\| + c_1\|b\|$$
$$= (c_1 + c_2 + c_1^2 c_2)\|b\|.$$

It is known, cf. [20] and [25, Lemma 4.15], that the bound on ξ implies

$$\# \operatorname{supp}\big(\mathscr{S}_\alpha(x)\big) = \# \operatorname{supp}\big(\mathbb{S}_\alpha(\xi(x))\big) \le \frac{4}{\alpha^2}(c_1 + c_2 + c_1^2 c_2)^2\|b\|^2.$$

The condition (iv) can be rewritten as $\|(I - F(x)^*F(x))y\| \le \gamma\|y\|$, for suitably sparse y. Since soft-thresholding is nonexpansive, i.e.,

$$\|\mathbb{S}_\alpha(x) - \mathbb{S}_\alpha(y)\| \le \|x - y\|,$$

the identity (4.10) then yields with (4.9)

$$\|\mathscr{S}_\alpha(x) - \mathscr{S}_\alpha(y)\| \le \|\mathscr{S}_\alpha(x) - \mathscr{S}_\alpha(y) + F(x)^*b - F(y)^*b$$
$$- F(x)^*F(x)\mathscr{S}_\alpha(x) + F(y)^*F(y)\mathscr{S}_\alpha(y)\|$$
$$\le \|(I - F(x)^*F(x))(\mathscr{S}_\alpha(x) - \mathscr{S}_\alpha(y))\|$$
$$+ \|(F(y)^*F(y) - F(x)^*F(x))\mathscr{S}_\alpha(y)\| + \|(F(x)^* - F(y)^*)b\|$$
$$\le \gamma\|\mathscr{S}_\alpha(x) - \mathscr{S}_\alpha(y)\| + 2c_1c_2c_3\|b\|\|x - y\| + c_3\|x - y\|\|b\|,$$

which implies

$$\|\mathscr{S}_\alpha(x) - \mathscr{S}_\alpha(y)\| \leq \frac{2c_1c_2 + 1}{1 - \gamma} c_3 \|b\| \|x - y\|.$$

Thus, $\mathscr{S}_\alpha$ is contractive and $x_\alpha^{(j)}$ converges towards a fixed point. $\square$ If $\|b\|$ is large, then the conditions in Theorem 4.4 are extremely strong. For smaller $\|b\|$, on the other hand, we can find examples matching the requirements. Note also that the above proof reveals that x_α is at most $\lceil \frac{4}{\alpha^2}(c_1 + c_2c_4)^2 \|b\|^2 \rceil$-sparse.

EXAMPLE 4.6. *As in Example 4.2, let $F(x) = A_1 + \epsilon f(\|x - x_0\|)A_2$, so that $F(x)x = A(x)$ with A as in Proposition 3.5. We additionally suppose that $n \leq d$, that A_1 is onto, and denote the smallest eigenvalue of $A_1 A_1^*$ by $\beta > 0$. Then we can choose $c_1 = \|A_1\| + \epsilon B \|A_2\|_2$ and $c_3 = \epsilon L \|A_2\|$. If $\epsilon, B, \|A_2\|$ are sufficiently small, then the smallest eigenvalue of $F(x)F(x)^*$ is almost given by the smallest eigenvalue of $A_1 A_1^*$ and denoted by $\beta > 0$. If $F(x)^\dagger = F(x)^*(F(x)F(x)^*)^{-1}$ denotes the pseudo-inverse of $F(x)$, then $\|F(x)^\dagger\|_2 \leq c_1/\beta$. For (ii), we can define $y := F(x)^\dagger b$, so that $\|y\|_{\ell_1} \leq \sqrt{d}\|y\| \leq \sqrt{d}\frac{c_1}{\beta}\|b\|$ and $c_2 \leq \sqrt{d}\frac{c_1}{\beta}$. Still, suppose that $\epsilon, B, \|A_2\|$ are sufficiently small and also assume that A_1 satisfies the RIP with constant $0 < \delta < 1$ for sufficiently large sparsity requirements in (iv). Thus, the conditions of Theorem 4.4 are satisfied if $\|b\|$ is sufficiently small.*

Quasi-linear iterative soft-thresholding:
Input: $F : \mathbb{R}^d \to \mathbb{R}^{n \times d}$, $b \in \mathbb{R}^n$
Initialize $x^{(0)}$ as an arbitrary vector
for $j = 0, 1, 2, \ldots$ *until some stopping criterion is met* **do**

$$x_\alpha^{(j+1)} := \arg\min_{x \in \mathbb{R}^d} \mathcal{J}_\alpha^S(x, x_\alpha^{(j)}) = \mathbb{S}_\alpha\big((I - F(x_\alpha^{(j)})^* F(x_\alpha^{(j)}))x_\alpha^{(j)} + F(x_\alpha^{(j)})^* b\big)$$

end
Output: $x_\alpha^{(1)}, x_\alpha^{(2)}, \ldots$

Algorithm 2: We propose to iteratively minimize the surrogate functional, which yields a simple iterative soft-thresholding scheme.

The recursive scheme in Theorem 4.4 involving $\mathscr{S}_\alpha$ requires a minimization in each iteration step. To derive a more efficient scheme, we consider the surrogate functional

$$\mathcal{J}_\alpha^S(x, a) = \|F(a)x - b\|^2 + \alpha \|x\|_{\ell_1} + \|x - a\|^2 - \|F(a)x - F(a)a\|^2.$$

We have $\mathcal{J}_\alpha^S(x, x) = \mathcal{J}_\alpha(x)$ and propose the iterative Algorithm 2. In each iteration step, we minimize the surrogate functional in the first variable having the second one fixed with the previous iteration, which only requires a simple soft-thresholding.

Indeed, iterative soft-thresholding converges towards the fixed point x_α:

THEOREM 4.7. *Suppose that the assumptions of Theorem 4.4 are satisfied and let x_α be the k-sparse fixed point in (4.7). We define $\hat{z}_\alpha := (I - F(x_\alpha)^* F(x_\alpha))x_\alpha) + F(x_\alpha)^* b$ and $K = \frac{4\|x_\alpha\|^2}{\alpha^2} + \frac{4c}{\alpha}C$, where $C = \sup_{1 \leq l < d}(\sqrt{l+1}\|\hat{z}_\alpha - (\hat{z}_\alpha)_{\{l\}}\|_{\ell_2})$ and $c > 0$ sufficiently large. Additionally assume that*
(a) there is $0 < \tilde{\gamma} < \gamma$ such that, for all $K + k$-sparse vectors $y \in \mathbb{R}^d$,

$$(1 - \tilde{\gamma})\|y\|^2 \|F(x_\alpha)y\|^2 \leq (1 + \tilde{\gamma})\|y\|^2, \tag{4.11}$$

(b) the constants satisfy $\tilde{\gamma} + (1 + 4c_1c_2)c_3\|b\| < \gamma$.
Then by using $x_\alpha^{(0)} = 0$ as initial vector, the iterative Algorithm 2 converges towards x_α with

$$\|x_\alpha^{(j)} - x_\alpha\| \leq \gamma^j \|x_\alpha\|, \quad j = 0, 1, 2, \ldots.$$

Note that the above k is at most $\lceil \frac{4}{\alpha^2}(c_1 + c_2 + c_1^2c_2)^2 \|b\|^2 \rceil$. Also, it may be possible to choose γ a little bigger than necessary to ensure $\tilde{\gamma} < \gamma$. Condition (b) can then be satisfied when the magnitude of the data

b is sufficiently small. Moreover, if constants are suitably chosen, Example 4.6 also provides a map F that satisfies the assumptions of Theorem 4.7 when $\|b\|$ is small.

Proof. We use induction and observe that the case $j = 0$ is trivially verified. Next, we suppose that $x_\alpha^{(j)}$ satisfies $\|x_\alpha^{(j)} - x_\alpha\| \leq \gamma^j\|x_\alpha\|$ and that it has at most K nonzero entries. Our aim is now to verify that $x^{(j+1)}$ also satisfies the support condition and $\|x_\alpha^{(j+1)} - x_\alpha\| \leq \gamma^{j+1}\|x_\alpha\|$. To simplify notation let

$$f(x, y) := (I - F(x)^*F(x))y + F(x)^*b, \tag{4.12}$$

so that $\hat{z}_\alpha = f(x_\alpha, x_\alpha)$. It will be useful later to derive bounds for both terms $\|f(x_\alpha, x_\alpha) - f(x_\alpha, x_\alpha^{(j)})\|$ and $\|f(x_\alpha^{(j)}, x_\alpha^{(j)}) - f(x_\alpha^{(j)}, x_\alpha^{(j)})\|$. Therefore, we start to estimate

$$\|f(x_\alpha, x_\alpha) - f(x_\alpha, x_\alpha^{(j)})\| = \|(I - F(x_\alpha)^*F(x_\alpha))(x_\alpha - x_\alpha^{(j)})\| \leq \tilde{\gamma}\gamma^j\|x_\alpha\|, \tag{4.13}$$

where we have used (a) in the form $\|(I - F(x_\alpha)^*F(x_\alpha))y\| \leq \tilde{\gamma}\|y\|$ and the induction hypothesis.

Next, we take care of $\|f(x_\alpha, x_\alpha^{(j)}) - f(x_\alpha^{(j)}, x_\alpha^{(j)})\|$ and derive

$$
\begin{aligned}
\|f(x_\alpha, x_\alpha^{(j)}) - f(x_\alpha^{(j)}, x_\alpha^{(j)})\| &\leq \|(F(x_\alpha)^* - F(x_\alpha^{(j)})^*)b\| \\
&\quad + \|(F(x_\alpha^{(j)})^*F(x_\alpha^{(j)}) - F(x_\alpha)^*F(x_\alpha))x_\alpha^{(j)}\| \\
&\leq c_3\gamma^j\|x_\alpha\|\|b\| + \|(F(x_\alpha^{(j)})^*F(x_\alpha^{(j)}) - F(x_\alpha^{(j)})^*F(x_\alpha))x_\alpha^{(j)}\| \\
&\quad + \|(F(x_\alpha^{(j)})^*F(x_\alpha) - F(x_\alpha)^*F(x_\alpha))x_\alpha^{(j)}\| \\
&\leq c_3\gamma^j\|x_\alpha\|\|b\| + 2c_1c_3\gamma^j\|x_\alpha\|\|x_\alpha^{(j)}\| \\
&= \gamma^j\|x_\alpha\|c_3(\|b\| + 2c_1\|x_\alpha^{(j)}\|).
\end{aligned}
$$

By using (ii) and the minimizing property of x_α, we derive

$$\|x_\alpha\| \leq \|x_\alpha\|_{\ell_1} \leq \|z_{x_\alpha}\|_{\ell_1} \leq c_2\|b\|.$$

The triangular inequality then yields $\|x_\alpha^{(j)}\| \leq \gamma^j c_2\|b\| + c_2\|b\|$. Thus, we obtain the estimate

$$\|f(x_\alpha, x_\alpha^{(j)}) - f(x_\alpha^{(j)}, x_\alpha^{(j)})\| \leq \gamma^j\|x_\alpha\|c_3\|b\|(1 + 2c_1c_2(1 + \gamma^j)). \tag{4.14}$$

According to (4.13) and (4.14), the condition (b) yields

$$\|\hat{z}_\alpha - f(x_\alpha^{(j)}, x_\alpha^{(j)})\| \leq \gamma^j\|x_\alpha\|(\tilde{\gamma} + c_3\|b\|(1 + 2c_1c_2(1 + \gamma^j))) \leq \gamma^{j+1}\|x_\alpha\|. \tag{4.15}$$

Results in [25, Lemma 4.15] with $x_\alpha^{(j+1)} = \mathbb{S}_\alpha(f(x_\alpha^{(j)}, x_\alpha^{(j)}))$ imply that there is a constant $c > 0$ such that

$$\#\operatorname{supp}(x_\alpha^{(j+1)}) \leq \frac{4\gamma^{2j+2}\|x_\alpha\|^2}{\alpha^2} + \frac{4c}{\alpha}C,$$

where $C = \sup_{1 \leq l < d}(\sqrt{l+1}\|z_\alpha - (z_\alpha)_{\{l\}}\|_{\ell_2})$. Since the above right-hand side is smaller than K, we have the desired support estimate.

Next, we take care of the error bounds. Since x_α is a fixed point of (4.7), we have $x_\alpha = \mathbb{S}_\alpha(\hat{z}_\alpha)$, which we have already used in (4.10). The nonexpansiveness of $\mathbb{S}_\alpha$ yields with $\hat{z}_\alpha = f(x_\alpha, x_\alpha)$

$$
\begin{aligned}
\|x_\alpha - x_\alpha^{(j+1)}\| &\leq \|\mathbb{S}_\alpha(f(x_\alpha, x_\alpha)) - \mathbb{S}_\alpha(f(x_\alpha, x_\alpha^{(j)}))\| + \|\mathbb{S}_\alpha(f(x_\alpha, x_\alpha^{(j)})) - \mathbb{S}_\alpha(f(x_\alpha^{(j)}, x_\alpha^{(j)}))\| \\
&\leq \|f(x_\alpha, x_\alpha) - f(x_\alpha, x_\alpha^{(j)})\| + \|f(x_\alpha, x_\alpha^{(j)}) - f(x_\alpha^{(j)}, x_\alpha^{(j)})\|.
\end{aligned}
$$

The same way as for (4.15), we use the bounds in (4.13) and (4.14) with (b) to derive

$$\|x_\alpha - x_\alpha^{(j+1)}\| \leq \gamma^j\|x_\alpha\|(\tilde{\gamma} + c_3\|b\|(1 + 4c_1c_2)) \leq \gamma^{j+1}\|x_\alpha\|,$$

so that we can conclude the proof. $\square$

It remains to verify that the output x_α of the iterative soft-thresholding scheme is close to the minimizer $\hat{x}_\alpha$ of (4.5):

THEOREM 4.8. *Suppose that the assumptions of Theorem 4.7 hold, that there is a K-sparse minimizer $\hat{x}_\alpha$ of (4.5), and that $\frac{c_2 c_3}{\sqrt{1-\tilde{\gamma}}}\|b\| < 1$ holds, then we have*

$$\|x_\alpha - \hat{x}_\alpha\| \leq \frac{\sqrt{\alpha c_2 \|b\|}}{a} + \frac{c_1 + c_3 \|\hat{x}\|}{a}\|\hat{x}_\alpha - \hat{x}\|,$$

where $\hat{x}$ satisfies $F(\hat{x})\hat{x} = b$ and $a = \sqrt{1-\tilde{\gamma}} - c_2 c_3 \|b\|$. Note that Proposition 4.3 yields that the minimizer $\hat{x}_\alpha$ can approximate $\hat{x}$, so that $\|\hat{x}_\alpha - \hat{x}\|$ can become small and, hence, $\|x_\alpha - \hat{x}_\alpha\|$ must be small. It should be mentioned though that the assumptions of Theorem 4.7 depend on α because its magnitude steers the sparsity of x_α. Therefore, letting α tend to zero is quite delicate because the assumptions become stronger. Indeed, taking the limit requires that condition (iv) in Theorem 4.4 holds for all $y \in \mathbb{R}^d$, not just for sparse vectors, and the same is required for condition (a) in Theorem 4.7.

Proof. [Proof of Theorem 4.8] We first bound $\hat{x}_\alpha$ by

$$\alpha\|\hat{x}_\alpha\| \leq \|F(\hat{x}_\alpha)\hat{x}_\alpha - b\|^2 + \alpha\|\hat{x}_\alpha\|_{\ell_1} \leq \|F(x_\alpha)x_\alpha - b\|^2 + \alpha\|x_\alpha\| \leq \alpha\|z_{x_\alpha}\|_{\ell_1}.$$

Therefore, we have $\|\hat{x}_\alpha\| \leq c_2 \|b\|$. Since $\hat{x}_\alpha$ is K-sparse, we derive

$$\|F(x_\alpha)x_\alpha - F(\hat{x}_\alpha)\hat{x}_\alpha\| \geq \|F(x_\alpha)x_\alpha - F(x_\alpha)\hat{x}_\alpha\| - \|F(x_\alpha)\hat{x}_\alpha - F(\hat{x}_\alpha)\hat{x}_\alpha\|$$
$$\geq \sqrt{1-\tilde{\gamma}}\|x_\alpha - \hat{x}_\alpha\| - c_3\|x_\alpha - \hat{x}_\alpha\|\|\hat{x}_\alpha\|.$$

These computations and a zero addition imply

$$\|x_\alpha - \hat{x}_\alpha\| \leq \frac{1}{a}\|F(x_\alpha)x_\alpha - F(\hat{x}_\alpha)\hat{x}_\alpha\|$$
$$\leq \frac{1}{a}(\|F(x_\alpha)x_\alpha - b\| + \|F(\hat{x}_\alpha)\hat{x}_\alpha - b\|).$$

We shall now bound both terms on the right-hand side separately. The minimizing property and (ii) in Theorem 4.4 yield

$$\|F(x_\alpha)x_\alpha - b\|^2 \leq \alpha\|z_{x_\alpha}\|_{\ell_1} - \alpha\|x_\alpha\|_{\ell_1} \leq \alpha c_2 \|b\|.$$

The second term is bounded by

$$\|F(\hat{x}_\alpha)\hat{x}_\alpha - b\| \leq \|F(\hat{x}_\alpha)\hat{x}_\alpha - F(\hat{x}_\alpha)\hat{x}\| + \|F(\hat{x}_\alpha)\hat{x} - F(\hat{x})\hat{x}\|$$
$$\leq c_1\|\hat{x}_\alpha - \hat{x}\| + c_3\|\hat{x}\|\|\hat{x}_\alpha - \hat{x}\|$$
$$= (c_1 + c_3\|\hat{x}\|)\|\hat{x}_\alpha - \hat{x}\|,$$

so that we can conclude the proof. □ Alternatively, we can also bound the distance between x_α and $\hat{x}$:

PROPOSITION 4.9. *Suppose that the assumptions of Theorem 4.7 hold, that we can replace x_α in condition (a) of the latter theorem with some K-sparse $\hat{x} \in \mathbb{R}^d$ satisfying $F(\hat{x})\hat{x} = b$, and that $\frac{c_2 c_3}{\sqrt{1-\tilde{\gamma}}}\|b\| < 1$ holds, then we have*

$$\|x_\alpha - \hat{x}\| \leq \frac{\sqrt{\alpha c_2 \|b\|}}{\sqrt{1-\tilde{\gamma}} - c_2 c_3 \|b\|}.$$

Proof. We can estimate

$$\|x_\alpha - \hat{x}\| \leq \frac{1}{\sqrt{1-\tilde{\gamma}}}\|F(\hat{x})x_\alpha - F(\hat{x})\hat{x}\|$$
$$\leq \frac{1}{\sqrt{1-\tilde{\gamma}}}\big(\|F(\hat{x})x_\alpha - F(x_\alpha)x_\alpha\| + \|F(x_\alpha)x_\alpha - F(\hat{x})\hat{x}\|\big)$$
$$\leq \frac{1}{\sqrt{1-\tilde{\gamma}}}\big(c_3\|x_\alpha\|\|x_\alpha - \hat{x}\| + \sqrt{\|F(x_\alpha)x_\alpha - b\|^2 + \alpha\|x_\alpha\|_{\ell_1}}\big)$$
$$\leq \frac{1}{\sqrt{1-\tilde{\gamma}}}\big(c_3 c_2\|b\|\|x_\alpha - \hat{x}\| + \sqrt{\alpha\|z_{x_\alpha}\|_{\ell_1}}\big)$$
$$\leq \frac{1}{\sqrt{1-\tilde{\gamma}}}\big(c_3 c_2\|b\|\|x_\alpha - \hat{x}\| + \sqrt{\alpha c_2\|b\|}\big).$$

From here on, some straight-forward calculations yield the required statement. □

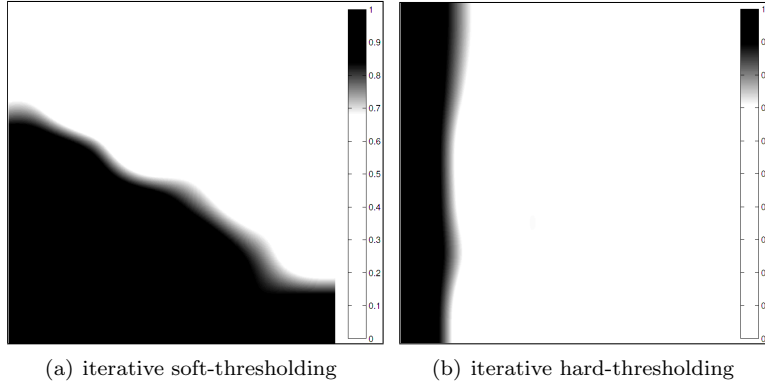(a) iterative soft-thresholding   (b) iterative hard-thresholding

Fig. 6. *Recovery rates for iterative hard- and soft-thresholding used with the measurements in the Examples 4.2 and 4.6 with $d = 80$, $n = 20$, A_1 having i.i.d. Gaussian entries, $\epsilon = 1$, and $A_2 = I$. The sparsity parameter k runs on the horizontal axis from 1 to 10, the norm of $\hat{x}$ runs on the vertical axis from 0.01 to 1. As expected, the recovery rates decrease with growing k. Consistent with the theory, we also observe decreased recovery rates for larger signal norms with soft-thresholding. Hard-thresholding appears only successful for these parameters when $k = 1$, but throughout the entire range of considered signal norms.*

4.3. Numerical experiments for iterative thresholding. Theorem 4.7 provides a simple thresh-olding algorithm to compute the fixed point x_α in (4.7) that is more efficient than the recursive scheme in Theorem 4.4. To support Theorem 4.8, we shall check numerically that x_α is indeed close to a minimizer $\hat{x}_\alpha$ of (4.5).

The quasi-linear measurements are taken from the Examples 4.2 and 4.6. The recovery rates from iterative hard- and soft-thresholding are plotted in Figure 6(a) and show a phase transition. For soft-thresholding, this transition depends on both, the sparsity level k and the measurement magnitude. Those observations are consistent with the theoretical results in Theorem 4.7 und suggest that the original signal can be recovered by iterative soft-thresholding. We also use hard-thresholding but for comparable parameter choices the signal was only recovered when $k = 1$, cf. Fig. 6(b).

The assumptions of the Theorems 4.1 and 4.7 cannot be satisfied within the phase retrieval setting, and the initial vectors $x^{(0)} = x_\alpha^{(0)} = 0$ in (4.1) and in Algorithm 2, respectively, would lead to a sequence of zero vectors. We observed numerically, that other choices of initial vectors do not yield acceptable recovery rates either, so that we did not pursue this direction.

REFERENCES

[1] C. Aerts, J. Christensen-Dalsgaard, and D. W. Kurtz, *Asteroseismology*, Springer, Berlin, 2010.
[2] C. Bachoc and M. Ehler, *Signal reconstruction from the magnitude of subspace components*, arXiv:1209.5986, (2012).
[3] B. Bah and J. Tanner, *Improved bounds on restricted isometry constants for gaussian matrices*, SIAM J. Matrix Anal., 31 (2010), pp. 2882–2898.
[4] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, *A simple proof of the restricted isometry property for random matrices*, Constr. Approx., 28 (2008), pp. 253–263.
[5] F. Barning, *The numerical analysis of the light-curve of 12 lacertae*, Bulletin of the Astronomical Institutes of the Netherlands, 17 (1963), pp. 22–28.
[6] H. H. Bauschke, P. L. Combettes, and D. R. Luke, *Phase retrieval, error reduction algorithm, and Fienup variants: A view from convex optimization*, J. Opt. Soc. Amer. A, 19 (2002), pp. 1334–1345.
[7] J. D. Blanchard, C. Cartis, J. Tanner, and A. Thompson, *Phase transitions for greedy sparse approximation algo-rithms*, Appl. Comput. Harmon. Anal., 30 (2011), pp. 188–203.

[8] T. Blumensath, *Compressed sensing with nonlinear observations and related nonlinear optimisation problems*, arXiv:1205.1650v1, (2012).

[9] T. Blumensath and M. E. Davies, *Gradient pursuit for non-linear sparse signal modelling*, Proc. European Signal Processing Conference (EUSIPCO), (2008).

[10] ———, *Iterative hard thresholding for compressed sensing*, Appl. Comput. Harmon. Anal., 27 (2009), pp. 265–274.

[11] E. J. Candes, J. Romberg, and T. Tao, *Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information*, IEEE Trans. Inform. Theory, 52 (2006), pp. 489–509.

[12] E. J. Candès, J. Romberg, and T. Tao, *Stable signal recovery from incomplete and inaccurate measurements*, Comm. Pure Appl. Math., 59 (2006), pp. 1207–1223.

[13] E. J. Candès, T. Strohmer, and V. Voroninski, *PhaseLift: Exact and stable signal recovery from magnitude measurements via convex programming*, Communications on Pure and Applied Mathematics, DOI:10.1002/cpa.21432, (2012).

[14] E. J. Candès and T. Tao, *Decoding by linear programming*, IEEE Trans. Inform. Theory, 51 (2005), pp. 4203–4215.

[15] ———, *Near-optimal signal recovery from random projections: universal encoding strategies*, IEEE Trans. Inform. Theory, 52 (2006), pp. 5406–5425.

[16] A. Cohen, R. A. DeVore, G. Petrova, and P. Wojtaszczyk, *Finding the minimum of a function*, preprint, (2013).

[17] S. Dasgupta and A. Gupta, *An elementary proof of a theorem of Johnson and Lindenstrauss*, Random Structures & Algorithms, 22 (2003), pp. 60–65.

[18] I. Daubechies, M. Defrise, and C. DeMol, *An iterative thresholding algorithm for linear inverse problems with a sparsity constraint*, Comm. Pure Appl. Math., 57 (2004), pp. 1413–1541.

[19] M. A. Davenport and M. B. Wakin, *Analysis of orthogonal matching pursuit using the restricted isometry property*, IEEE Trans. Inform. Theory, 56 (2010), pp. 4395–4401.

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

[21] D. L. Donoho, *Compressed sensing*, IEEE Trans. Inform. Theory, 52 (2006), pp. 1289–1306.

[22] J. Drenth, *Principles of Protein X-Ray Crystallography*, Springer, 2010.

[23] Y. C. Eldar and S. Mendelson, *Phase retrieval: Stability and recovery guarantees*, arXiv:1211.0872, (2012).

[24] J. R. Fienup, *Phase retrieval algorithms: a comparison*, Applied Optics, 21 (1982), pp. 2758–2769.

[25] M. Fornasier, ed., *Theoretical Foundations and Numerical Methods for Sparse Recovery*, vol. 9 of Radon Series on Computational and Applied Mathematics, De Gruyter, 2010.

[26] M. Fornasier and H. Rauhut, *Iterative thresholding algorithms*, Appl. Comput. Harmon. Anal, 25 (2008), pp. 187–208.

[27] ———, *Recovery algorithms for vector valued data with joint sparsity constraints*, SIAM J. Numer. Anal., 46 (2008), pp. 577–613.

[28] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*, Birkhauser, Boston, in press.

[29] R. W. Gerchberg and W. O. Saxton, *A practical algorithm for the determination of the phase from image and diffraction plane pictures*, Optik, 35 (1972), pp. 237–246.

[30] D. Goldfarb and S. Ma, *Convergence of fixed-point continuation algorithms for matrix rank minimization*, arXiv:0906.3499v4, (2010).

[31] C. R. Gribonval, J. Idier, and C. Soussen, *Sparse recovery conditions for orthogonal least squares*, tech. report, CRAN, INRIA Rennes-Bretagne Atlantique, IRCCyN, 2011.

[32] X. Li and V. Voroninski, *Sparse signal recovery from quadratic measurements via convex programming*, arXiv:1209.4785v1, (2012).

[33] S. Mallat and Z. Zhang, *Matching pursuits with time-frequency dictionaries*, IEEE Trans. Signal Process., 41 (1993), pp. 3397–3415.

[34] D. Needell and J. Tropp, *Cosamp: Iterative signal recovery from incomplete and inaccurate samples*, Appl. Comput. Harmon. Anal., 28 (2009), pp. 301–321.

[35] G. Pfander, H. Rauhut, and J. Tropp, *The restricted isometry property for time-frequency structured random matrices*, Prob. Theory Rel. Fields, (2012), pp. 1–31.

[36] R. Ramlau and G. Teschke, *A projection iteration for nonlinear operator equations with sparsity constraints*, Numerische Mathematik, 104 (2006), pp. 177–203.

[37] ———, *An iterative algorithm for nonlinear inverse problems with joint sparsity constraints in vector valued regimes and an application to color image inpainting*, Inverse Problems, 23 (2007), pp. 1851–1870.

[38] H. Rauhut, J. Romberg, and J. Tropp, *Restricted isometries for partial random circulant matrices*, Appl. Comput. Harmon. Anal., 32 (2012), pp. 242–254.

[39] H. Rauhut, K. Schnass, and P. Vandergheynst, *Compressed sensing and redundant dictionaries*, IEEE Trans. Inform. Theory, 54 (2008), pp. 2210–2219.

[40] B. Recht, M. Fazel, and P. Parillo, *Guaranteed minimum rank solutions to linear matrix equations via nuclear norm minimization*, SIAM Rev., 52 (2010), pp. 471–501.

[41] B. Seifert, H. Stolz, M. Donatelli, D. Langemann, and M. Tasche, *Multilevel Gauss-Newton methods for phase retrieval problems*, J. Phys. A, 39 (2006), pp. 4191–4206.

[42] J. Sigl, *Quasilinear compressed sensing.* Master's thesis, Technische Universität München, 2013.

[43] V. N. Temlyakov and P. Zheltov, *On performance of greedy algorithms*, J. Approx. Theory, 163 (2011), pp. 1134–1145.

[44] J. A. Tropp, *Greed is good: Algorithmic results for sparse approximation*, IEEE Trans. Inform. Theory, 50 (2004), pp. 2331–2242.

[45] R. Vershynin, *Introduction to the non-asymptotic analysis of random matrices*, in Compressed sensing, Theory and Applications, Y. Eldar and G. Kutyniok, eds., Cambridge University Press, 2012, ch. 5, pp. 210–268.