

6-24-2016

One-Bit Compressive Sensing of Dictionary-Sparse Signals

Richard Baraniuk
Rice University

Simon Foucart
University of Georgia

Deanna Needell
Claremont McKenna College

Yaniv Plan
University of British Columbia

Mary Wootters
University of Michigan - Ann Arbor

Recommended Citation

R. Baraniuk, S. Foucart, D. Needell, Y. Plan, and M. Wootters. "One-Bit Compressive Sensing of Dictionary-Sparse Signals." 2016.

This Article - preprint is brought to you for free and open access by the CMC Faculty Scholarship at Scholarship @ Claremont. It has been accepted for inclusion in CMC Faculty Publications and Research by an authorized administrator of Scholarship @ Claremont. For more information, please contact scholarship@cuc.claremont.edu.

One-Bit Compressive Sensing of Dictionary-Sparse Signals

R. Baraniuk, S. Foucart, D. Needell, Y. Plan, and M. Wootters

Abstract

One-bit compressive sensing has extended the scope of sparse recovery by showing that sparse signals can be accurately reconstructed even when their linear measurements are subject to the extreme quantization scenario of binary samples—only the sign of each linear measurement is maintained. Existing results in one-bit compressive sensing rely on the assumption that the signals of interest are sparse in some fixed orthonormal basis. However, in most practical applications, signals are sparse with respect to an overcomplete dictionary, rather than a basis. There has already been a surge of activity to obtain recovery guarantees under such a generalized sparsity model in the classical compressive sensing setting. Here, we extend the one-bit framework to this important model, providing a unified theory of one-bit compressive sensing under dictionary sparsity. Specifically, we analyze several different algorithms—based on convex programming and on hard thresholding—and show that, under natural assumptions on the sensing matrix (satisfied by Gaussian matrices), these algorithms can efficiently recover analysis-dictionary-sparse signals in the one-bit model.

Key words and phrases: compressive sensing, quantization, one-bit compressive sensing, tight frames, convex optimization, thresholding.

1 Introduction

The basic insight of compressive sensing is that a small number of linear measurements can be used to reconstruct sparse signals. In traditional compressive sensing, we wish to reconstruct an s -sparse¹ signal $\mathbf{x} \in \mathbb{R}^N$ from linear measurements of the form

$$(1) \quad \mathbf{y} = \mathbf{A}\mathbf{x} \in \mathbb{R}^m \quad (\text{or its corrupted version } \mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e}),$$

where \mathbf{A} is a $m \times N$ measurement matrix. A significant body of work over the past decade has demonstrated that the s -sparse (or nearly s -sparse) signal \mathbf{x} can be accurately and efficiently recovered from its measurement vector $\mathbf{y} = \mathbf{A}\mathbf{x}$ when \mathbf{A} has independent Gaussian entries, say, and when $m \asymp s \log(N/s)$ [DSP, EK12, FR13].

¹A signal $\mathbf{x} \in \mathbb{R}^N$ is called s -sparse if $\|\mathbf{x}\|_0 := |\text{supp}(\mathbf{x})| \leq s \ll N$.

This basic model has been extended in several directions. Two important ones—which we focus on in this work—are (a) extending the set of signals to include the larger and important class of *dictionary sparse* signals and (b) considering highly quantized measurements as in *one-bit compressive sensing*.

Both of these settings have important practical applications and have received much attention in the past few years. However, to the best of our knowledge, they have not been considered together before. In this work, we extend the theory of one-bit compressive sensing to dictionary sparse signals. Below, we briefly review the background on these notions, set up notation, and outline our contributions.

1.1 One-bit measurements

In practice, each entry $y_i = \langle \mathbf{a}_i, \mathbf{x} \rangle$ (where \mathbf{a}_i denotes the i th row of \mathbf{A}) of the measurement vector in (1) needs to be quantized. That is, rather than observing $\mathbf{y} = \mathbf{A}\mathbf{x}$, one observes $\mathbf{y} = Q(\mathbf{A}\mathbf{x})$ instead, where $Q : \mathbb{R}^m \rightarrow \mathcal{A}$ denotes the quantizer that maps each entry of its input to a corresponding quantized value in an alphabet \mathcal{A} . The so-called *one-bit compressive sensing* [BB08] problem refers to the case when $|\mathcal{A}| = 2$ and one wishes to recover \mathbf{x} from its heavily quantized (one bit) measurements $\mathbf{y} = Q(\mathbf{A}\mathbf{x})$. The simplest quantizer in the one-bit case uses the alphabet $\mathcal{A} = \{-1, 1\}$ and acts by taking the sign of each component as

$$(2) \quad y_i = Q(\langle \mathbf{a}_i, \mathbf{x} \rangle) = \text{sgn}(\langle \mathbf{a}_i, \mathbf{x} \rangle),$$

which we denote in shorthand by $\mathbf{y} = \text{sgn}(\mathbf{A}\mathbf{x})$. Since the publication of [BB08] in 2008, several efficient methods, both iterative and optimization-based, have been developed to recover the signal \mathbf{x} (up to normalization) from its one-bit measurements (see e.g. [PV13a, PV13b, GNJN13, JLBB13, YYO12, JDDV13]). In particular, it is shown [JLBB13] that the direction of any s -sparse signal \mathbf{x} can be estimated by some $\hat{\mathbf{x}}$ produced from \mathbf{y} with accuracy

$$\left\| \frac{\mathbf{x}}{\|\mathbf{x}\|_2} - \frac{\hat{\mathbf{x}}}{\|\hat{\mathbf{x}}\|_2} \right\|_2 \leq \varepsilon$$

when the number of measurements is at least

$$m = \Omega\left(\frac{s \ln(N/s)}{\varepsilon}\right).$$

Notice that with measurements of this form, we can only hope to recover the direction of the signal, not the magnitude. However, we can recover the entire signal if we allow for *thresholded measurements* of the form

$$(3) \quad y_i = \text{sgn}(\langle \mathbf{a}_i, \mathbf{x} \rangle - \tau_i).$$

In practice, it is often feasible to obtain quantized measurements of this form, and they have been studied before. Existing works using measurements of the form (3) have also allowed for *adaptive* thresholds; that is, the τ_i can be chosen adaptively based on y_j for $j < i$. The goal of those works was to improve the convergence rate, i.e., the dependence on ε in the number of measurements m . It is known that a dependence of $\Omega(1/\varepsilon)$ is necessary with nonadaptive measurements, but recent work on Sigma-Delta quantization [SWY15] and other schemes [BFN⁺14, KSW14] have shown how to break this barrier using measurements of the form (3) with adaptive thresholds.

In this article, we do not focus on the decay rate (the dependence on ε), nor do we consider adaptive measurements. However, we do consider nonadaptive measurements both of the form (2) and (3). This allows us to provide results on reconstruction of the magnitude of signals, as well as the direction.

1.2 Dictionary Sparsity

Although the classical setting assumes that the signal \mathbf{x} itself is sparse, most signals of interest are not immediately sparse. In the straightforward case, a signal may be instead sparse after some transform; for example, images are known to be sparse in the wavelet domain, sinusoidal signals in the Fourier domain, and so on [Dau92]. Fortunately, the classical framework extends directly to this model, since the product of a Gaussian matrix and an orthonormal basis is still Gaussian.

However, in many practical applications the situation is not so straightforward, and the signals of interest are sparse not in an orthonormal basis but rather in a redundant (highly overcomplete) dictionary; this is known as *dictionary sparsity*. Signals in radar and sonar systems, for example, are sparsely represented in Gabor frames, which are highly overcomplete and far from orthonormal [FS98]. Images may be sparsely represented in curvelet frames [CDDY00, CD04], undecimated wavelet frames [SED04], and other frames which by design are highly redundant. Such redundancy allows for sparser representations and a wider class of signal representations. Even in the Fourier domain, utilizing an oversampled DFT allows for much more realistic and practical signals to be represented. For these reasons, recent research has extended the compressive sensing framework to the setting where the signals of interest are sparsified by overcomplete tight frames (see e.g. [RSV08, CENR10, GNE⁺14, Fou16]).

Throughout this article, we consider a dictionary $\mathbf{D} \in \mathbb{R}^{n \times N}$ which is assumed to be a tight frame, in the sense that

$$\mathbf{D}\mathbf{D}^* = \mathbf{I}_n.$$

To distinguish between the signal and its sparse representation, we write $\mathbf{f} \in \mathbb{R}^n$ for the signal of interest and $\mathbf{f} = \mathbf{D}\mathbf{x}$, where $\mathbf{x} \in \mathbb{R}^N$ is a sparse coefficient vector. We then acquire the samples of the form $\mathbf{y} = \mathbf{A}\mathbf{f} = \mathbf{A}\mathbf{D}\mathbf{x}$ and attempt to recover the signal \mathbf{f} . Note that, due to the redundancy

of \mathbf{D} , we do not hope to be able to recover a unique coefficient vector \mathbf{x} . In other words, even when the measurement matrix \mathbf{A} is well suited for sparse recovery, the product \mathbf{AD} may have highly correlated columns, making recovery of \mathbf{x} impossible. With the introduction of a noninvertible sparsifying transform \mathbf{D} , it becomes important to distinguish between two related but distinct notions of sparsity. Precisely, we say that

- \mathbf{f} is s -synthesis-sparse if $\mathbf{f} = \mathbf{D}\mathbf{x}$ for some s -sparse $\mathbf{x} \in \mathbb{R}^N$;
- \mathbf{f} is s -analysis-sparse if $\mathbf{D}^*\mathbf{f} \in \mathbb{R}^N$ is s -sparse.

We note that analysis sparsity is a stronger assumption because, assuming analysis sparsity, one can always take $\mathbf{x} = \mathbf{D}^*\mathbf{f}$ in the synthesis sparsity model. See [EMR07] for an introduction to the analysis sparse model in compressive sensing (also called the analysis cosparsity model).

Instead of exact sparsity, it is often more realistic to study *effective sparsity*. We call a coefficient vector $\mathbf{x} \in \mathbb{R}^N$ effectively s -sparse if

$$\|\mathbf{x}\|_1 \leq \sqrt{s}\|\mathbf{x}\|_2,$$

and we say that

- \mathbf{f} is effectively s -synthesis-sparse if $\mathbf{f} = \mathbf{D}\mathbf{x}$ for some effectively s -sparse $\mathbf{x} \in \mathbb{R}^N$;
- \mathbf{f} is effectively s -analysis-sparse if $\mathbf{D}^*\mathbf{f} \in \mathbb{R}^N$ is effectively s -sparse.

We use the notation

$$\begin{aligned} \Sigma_s^N &\text{ for the set of } s\text{-sparse coefficient vectors in } \mathbb{R}^N, \\ \Sigma_s^{N,\text{eff}} &\text{ for the set of effectively } s\text{-sparse coefficient vectors in } \mathbb{R}^N. \end{aligned}$$

We also use the notation B_2^n for the set of signals with ℓ_2 -norm at most 1 (i.e., the unit ball in ℓ_2^n) and S^{n-1} for the set of signals with ℓ_2 -norm equal to 1 (i.e., the unit sphere in ℓ_2^n).

It is now well known that, if \mathbf{D} is a tight frame and \mathbf{A} satisfies analogous conditions to those in the classical setting (e.g., has independent Gaussian entries), then a signal \mathbf{f} which is (effectively) analysis- or synthesis-sparse can be accurately recovered from traditional compressive sensing measurements $\mathbf{y} = \mathbf{A}\mathbf{f} = \mathbf{AD}\mathbf{x}$ (see e.g. [RSV08, Blu11, CENR10, DNW12, GNE⁺14, NDEG13, PE13, Fou16]).

1.3 One-bit measurements with dictionaries: our setup

In this article, we study one-bit compressive sensing for dictionary-sparse signals. Precisely, our aim is to recover signals $\mathbf{f} \in \mathbb{R}^n$ from the binary measurements

$$y_i = \text{sgn}\langle \mathbf{a}_i, \mathbf{f} \rangle, \quad i = 1, \dots, m,$$

or

$$y_i = \text{sgn}(\langle \mathbf{a}_i, \mathbf{f} \rangle - \tau_i), \quad i = 1, \dots, m,$$

when these signals are sparse with respect to a dictionary \mathbf{D} .

As in Section 1.2, there are several ways to model signals which are sparse with respect to \mathbf{D} . In this work, two different signal classes are considered. For the first one, which is more general, our results are based on convex programming. For the second one, which is a more restrictive, we can obtain results using a computationally simpler algorithm based on hard thresholding.

The first class consists of signals $\mathbf{f} \in (\mathbf{D}^*)^{-1} \Sigma_s^{N, \text{eff}}$ that are effectively s -analysis-sparse, i.e., they satisfy

$$(4) \quad \|\mathbf{D}^* \mathbf{f}\|_1 \leq \sqrt{s} \|\mathbf{D}^* \mathbf{f}\|_2.$$

This occurs, of course, when $\mathbf{D}^* \mathbf{f}$ is genuinely sparse (analysis sparsity) and this is realistic if we are working e.g. with piecewise-constant images, since they are sparse after application of the total variation operator. We consider effectively sparse signals since genuine analysis sparsity is unrealistic when \mathbf{D} has columns in general position, as it would imply that \mathbf{f} is orthogonal to too many columns of \mathbf{D} .

The second class consists of signals $\mathbf{f} \in \mathbf{D}(\Sigma_s^N) \cap (\mathbf{D}^*)^{-1} \Sigma_{\kappa s}^{N, \text{eff}}$ that are both s -synthesis-sparse and κs -analysis-sparse for some $\kappa \geq 1$. This will occur as soon as the signals are s -synthesis-sparse, provided we utilize suitable dictionaries $\mathbf{D} \in \mathbb{R}^{n \times N}$. One could take, for instance, the matrix of an equiangular tight frame when $N = n + k$, $k = \text{constant}$. Other examples of suitable dictionaries found in [KNW15] include harmonic frames again with $N = n + k$, $k = \text{constant}$, as well as Fourier and Haar frames with constant redundancy factor N/n .

Figure 1 summarizes the relationship between the various domains we deal with.

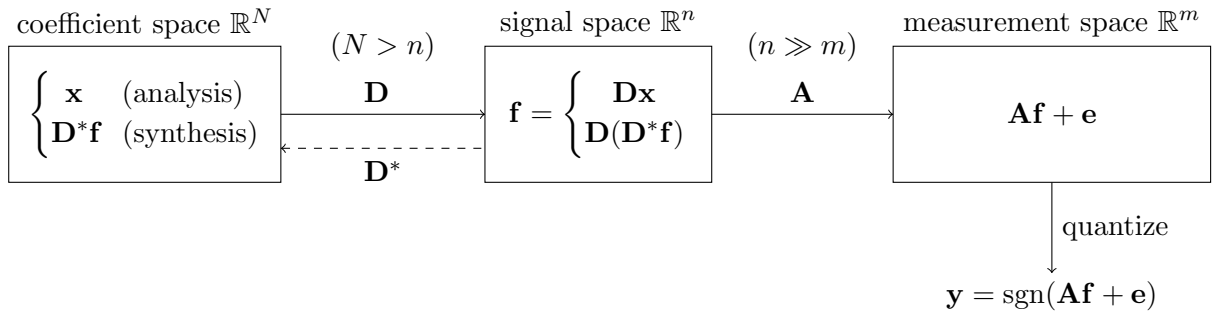


Figure 1: The coefficient, signal, and measurement domains.

1.4 Contributions

Our main results demonstrate that one-bit compressive sensing is viable even when the sparsifying transform is an overcomplete dictionary. As outlined in Section 1.1, we consider both the challenge of recovering the direction $\mathbf{f}/\|\mathbf{f}\|_2$ of a signal \mathbf{f} , as well as the challenge of recovering the entire signal (direction and magnitude). Using measurements of the form $y_i = \text{sgn}\langle \mathbf{a}_i, \mathbf{f} \rangle$, we can recover the direction but not the magnitude; using measurements of the form $y_i = \text{sgn}(\langle \mathbf{a}_i, \mathbf{f} \rangle - \tau_i)$, we may recover both.

In (one-bit) compressive sensing, two standard families of algorithms are (a) algorithms based on convex programming, and (b) algorithms based on thresholding. In this article, we analyze algorithms from both classes. One reason to study multiple algorithms is to give a more complete landscape of this problem. Another reason is that the different algorithms come with different trade-offs (between computational complexity and the strength of assumptions required), and it is valuable to explore this space of trade-offs.

Recovering the direction. First, we show that the direction of a dictionary sparse signal can be estimated from one-bit measurements of the type $\text{sgn}(\mathbf{A}\mathbf{f})$. We consider two algorithms; our first approach is based on linear programming, and our second is based on hard thresholding. The linear programming approach is more computationally demanding, but applies to a broader class of signals.

In Section 3, we prove that both of these approaches are effective, provided the sensing matrix \mathbf{A} satisfies certain properties. In Section 2, we state that these properties are in fact satisfied by a matrix \mathbf{A} populated with independent Gaussian entries. We combine all of these results to prove the statement below. As noted above, the different algorithms require different definitions of “dictionary sparsity”. In what follows, γ, C, c refer to absolute numerical constants.

Theorem 1 (Informal statement of direction recovery). Let $\varepsilon > 0$, let $m \geq C\varepsilon^{-7}s \ln(eN/s)$, and let $\mathbf{A} \in \mathbb{R}^{m \times n}$ be populated by independent standard normal random variables. Then, with failure probability at most $\gamma \exp(-c\varepsilon^2 m)$, any dictionary sparse² signal $\mathbf{f} \in \mathbb{R}^n$ observed via $\mathbf{y} = \text{sgn}(\mathbf{A}\mathbf{f})$ can be approximated by the output $\widehat{\mathbf{f}}$ of an efficient algorithm with error

$$\left\| \frac{\mathbf{f}}{\|\mathbf{f}\|_2} - \frac{\widehat{\mathbf{f}}}{\|\widehat{\mathbf{f}}\|_2} \right\|_2 \leq \varepsilon.$$

²Here, “dictionary sparsity” means effective s -analysis-sparsity if $\widehat{\mathbf{f}}$ is produced by convex programming and genuine s -synthesis sparsity together with effective κs -analysis-sparsity if $\widehat{\mathbf{f}}$ is produced by hard thresholding.

Recovering the whole signal. By using one-bit measurements of the form $\text{sgn}(\mathbf{A}\mathbf{f} - \boldsymbol{\tau})$, where τ_1, \dots, τ_m are properly normalized Gaussian random thresholds, we are able to recover not just the direction but also the magnitude of a dictionary-sparse signal \mathbf{f} .

We consider three algorithms; our first approach is based on linear programming, our second approach on second-order cone programming, and our third approach on hard thresholding. Again, there are different trade-offs to the different algorithms. As above, the approach based on hard thresholding is more efficient, while the approaches based on convex programming apply to a broader signal class. There is also a trade-off between linear programming and second-order cone programming: the second-order cone program requires knowledge of $\|\mathbf{f}\|_2$ while the linear program does not (although it does require a loose bound), but the second-order cone programming approach applies to a slightly larger class of signals.

We show in Section 4 that all three of these algorithms are effective when the sensing matrix \mathbf{A} is populated with independent Gaussian entries and when the thresholds τ_i are also independent Gaussian random variables. We combine the results of Section 4 in the following theorem.

Theorem 2 (Informal statement of signal estimation). Let $\varepsilon, r, \sigma > 0$, let $m \geq C\varepsilon^{-9}s \ln(eN/s)$, and let $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\boldsymbol{\tau} \in \mathbb{R}^m$ be populated by independent mean-zero normal random variables with variance 1 and σ^2 , respectively. Then, with failure probability at most $\gamma \exp(-c\varepsilon^2 m)$, any dictionary sparse³ signal $\mathbf{f} \in \mathbb{R}^n$ with $\|\mathbf{f}\|_2 \leq r$ observed via $\mathbf{y} = \text{sgn}(\mathbf{A}\mathbf{f} - \boldsymbol{\tau})$ is approximated by the output $\hat{\mathbf{f}}$ of an efficient algorithm with error

$$\|\mathbf{f} - \hat{\mathbf{f}}\|_2 \leq \varepsilon r.$$

We have not spelled out the dependence of the number of measurements and the failure probability on the parameters r and σ : as long as they are roughly the same order of magnitude, the dependence is absorbed in the constants C and c (see Section 4 for precise statements). As outlined earlier, an estimate of r is required to implement the second-order cone program, but the other two algorithms do not require such an estimate.

1.5 Discussion and future directions

The purpose of this work is to demonstrate that techniques from one-bit compressive sensing can be effective for the recovery of dictionary-sparse signals and we propose several algorithms to accomplish this for various notions of dictionary sparsity. Still, some interesting future directions remain.

³See footnote 2.

Firstly, we do not believe that the dependence on ε above is optimal. We do believe instead that a logarithmic dependence on ε for the number of measurements (or equivalently an exponential decay in the oversampling factor $\lambda = m/(s \ln(eN/s))$ for the recovery error) is possible by choosing the thresholds τ_1, \dots, τ_m adaptively. This would be achieved by adjusting the method of [BFN⁺14], but with the strong proviso of exact sparsity.

Secondly, it is worth asking to what extent the trade-offs between the different algorithms reflect reality. In particular, is it only an artifact of the proof that the simpler algorithm based on hard thresholding applies to a narrower class of signals?

1.6 Organization

The remainder of the paper is organized as follows. In Section 2, we outline some technical tools upon which our results rely, namely some properties of Gaussian random matrices . In Section 3, we consider recovery of the direction $\mathbf{f}/\|\mathbf{f}\|$ only and we propose two algorithms to achieve it. In Section 4, we present three algorithms for the recovery of the entire signal \mathbf{f} . Finally, in Section 5, we provide proofs for the results outlined in Section 2.

2 Technical ingredients

In this section, we highlight the theoretical properties upon which our results rely. Their proofs are deferred to Section 5 so that the reader does not lose track of our objectives. The first property we put forward is an adaptation to the dictionary case of the so-called sign product embedding property (the term was coined in [JDDV13] but the result originally appeared in [PV13b]).

Theorem 3 (D-SPEP). Let $\delta > 0$, let $m \geq C\delta^{-7}s \ln(eN/s)$, and let $\mathbf{A} \in \mathbb{R}^{m \times n}$ be populated by independent standard normal random variables. Then, with failure probability at most $\gamma \exp(-c\delta^2 m)$, the renormalized matrix $\mathbf{A}' := (\sqrt{2/\pi}/m)\mathbf{A}$ satisfies the s th order sign product embedding property adapted to $\mathbf{D} \in \mathbb{R}^{n \times N}$ with constant δ — **D-SPEP**(s, δ) for short — i.e.,

$$(5) \quad |\langle \mathbf{A}'\mathbf{f}, \text{sgn}(\mathbf{A}'\mathbf{g}) \rangle - \langle \mathbf{f}, \mathbf{g} \rangle| \leq \delta$$

holds for all $\mathbf{f}, \mathbf{g} \in \mathbf{D}(\Sigma_s^N) \cap S^{n-1}$.

Remark 1. The power δ^{-7} is unlikely to be optimal. At least in the nondictionary case, i.e., when $\mathbf{D} = \mathbf{I}_n$, it can be reduced to δ^{-2} , see [BL15].

As an immediate consequence of **D-SPEP**, setting $\mathbf{g} = \mathbf{f}$ in (5) allows one to deduce a variation of the classical restricted isometry property adapted to \mathbf{D} , where the inner norm becomes the ℓ_1 -norm (we mention in passing that this variation could also be deduced by other means).

Corollary 4 (D-RIP₁). Let $\delta > 0$, let $m \geq C\delta^{-7}s \ln(eN/s)$, and let $\mathbf{A} \in \mathbb{R}^{m \times n}$ be populated by independent standard normal random variables. Then, with failure probability at most $\gamma \exp(-c\delta^2 m)$, the renormalized matrix $\mathbf{A}' := (\sqrt{2/\pi}/m)\mathbf{A}$ satisfies the s th-order ℓ_1 -restricted isometry property adapted to $\mathbf{D} \in \mathbb{R}^{n \times N}$ with constant δ — **D-RIP₁**(s, δ) for short — i.e.,

$$(6) \quad (1 - \delta)\|\mathbf{f}\|_2 \leq \|\mathbf{A}'\mathbf{f}\|_1 \leq (1 + \delta)\|\mathbf{f}\|_2$$

holds for all $\mathbf{f} \in \mathbf{D}(\Sigma_s^N)$.

The next property we put forward is an adaptation of the tessellation of the “effectively sparse sphere” (see [PV14]) to the dictionary case. In what follows, given a (noninvertible) matrix \mathbf{M} and a set K , we denote by $\mathbf{M}^{-1}(K)$ the preimage of K with respect to \mathbf{M} .

Theorem 5 (Tessellation). Let $\varepsilon > 0$, let $m \geq C\varepsilon^{-6}s \ln(eN/s)$, and let $\mathbf{A} \in \mathbb{R}^{m \times n}$ be populated by independent standard normal random variables. Then, with failure probability at most $\gamma \exp(-c\varepsilon^2 m)$, the rows $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$ of \mathbf{A} ε -tessellate the effectively s -analysis-sparse sphere — we write that \mathbf{A} satisfies **D-TEs**(s, ε) for short — i.e.,

$$(7) \quad [\mathbf{f}, \mathbf{g} \in (\mathbf{D}^*)^{-1}(\Sigma_s^{N, \text{eff}}) \cap S^{n-1} : \text{sgn}\langle \mathbf{a}_i, \mathbf{f} \rangle = \text{sgn}\langle \mathbf{a}_i, \mathbf{g} \rangle \text{ for all } i = 1, \dots, m] \implies [\|\mathbf{f} - \mathbf{g}\|_2 \leq \varepsilon].$$

3 Signal estimation: direction only

In this whole section, given a measurement matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ with rows $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$, the signals $\mathbf{f} \in \mathbb{R}^n$ are acquired via $\mathbf{y} = \text{sgn}(\mathbf{A}\mathbf{f}) \in \{-1, +1\}^m$, i.e.,

$$y_i = \text{sgn}\langle \mathbf{a}_i, \mathbf{f} \rangle, \quad i = 1, \dots, m.$$

Under this model, all $c\mathbf{f}$ with $c > 0$ produce the same one-bit measurements, so one can only hope to recover the direction of \mathbf{f} . We present two methods to do so, one based on linear programming and the other one based on hard thresholding.

3.1 Linear programming

Given a signal $\mathbf{f} \in \mathbb{R}^n$ observed via $\mathbf{y} = \text{sgn}(\mathbf{A}\mathbf{f})$, the optimization scheme we consider here consists in outputting the signal \mathbf{f}_{lp} solution of

$$(8) \quad \underset{\mathbf{h} \in \mathbb{R}^n}{\text{minimize}} \|\mathbf{D}^*\mathbf{h}\|_1 \quad \text{subject to} \quad \text{sgn}(\mathbf{A}\mathbf{h}) = \mathbf{y}, \quad \|\mathbf{A}\mathbf{h}\|_1 = 1.$$

This is in fact a linear program (and thus may be solved efficiently), since the condition $\text{sgn}(\mathbf{A}\mathbf{h}) = \mathbf{y}$ reads

$$y_i(\mathbf{A}\mathbf{h})_i \geq 0 \quad \text{for all } i = 1, \dots, m,$$

and, under this constraint, the condition $\|\mathbf{A}\mathbf{h}\|_1 = 1$ reads

$$\sum_{i=1}^m y_i (\mathbf{A}\mathbf{h})_i = 1.$$

Theorem 6. If $\mathbf{A} \in \mathbb{R}^{m \times n}$ satisfies both \mathbf{D} -TES($36s, \varepsilon$) and \mathbf{D} -RIP $_1(25s, 1/5)$, then any effectively s -analysis-sparse signal $\mathbf{f} \in (\mathbf{D}^*)^{-1}\Sigma_s^{N, \text{eff}}$ observed via $\mathbf{y} = \text{sgn}(\mathbf{A}\mathbf{f})$ is directionally approximated by the output \mathbf{f}_{lp} of the linear program (8) with error

$$\left\| \frac{\mathbf{f}}{\|\mathbf{f}\|_2} - \frac{\mathbf{f}_{\text{lp}}}{\|\mathbf{f}_{\text{lp}}\|_2} \right\|_2 \leq \varepsilon.$$

Proof. The main step is to show that \mathbf{f}_{lp} is effectively $36s$ -analysis-sparse when \mathbf{D} -RIP $_1(t, \delta)$ holds with $t = 25s$ and $\delta = 1/5$. Then, since both $\mathbf{f}/\|\mathbf{f}\|_2$ and $\mathbf{f}_{\text{lp}}/\|\mathbf{f}_{\text{lp}}\|_2$ belong to $(\mathbf{D}^*)^{-1}\Sigma_{36s}^{N, \text{eff}} \cap \mathcal{S}^{n-1}$ and have the same sign observations, \mathbf{D} -TES($36s, \varepsilon$) implies the desired conclusion. To prove the effective analysis-sparsity of \mathbf{f}_{lp} , we first estimate $\|\mathbf{A}\mathbf{f}\|_1$ from below. For this purpose, let T_0 denote an index set of t largest absolute entries of $\mathbf{D}^*\mathbf{f}$, T_1 an index set of next t largest absolute entries of $\mathbf{D}^*\mathbf{f}$, T_2 an index set of next t largest absolute entries of $\mathbf{D}^*\mathbf{f}$, etc.. We have

$$\begin{aligned} \|\mathbf{A}\mathbf{f}\|_1 &= \|\mathbf{A}\mathbf{D}\mathbf{D}^*\mathbf{f}\|_1 = \left\| \mathbf{A}\mathbf{D} \left(\sum_{k \geq 0} (\mathbf{D}^*\mathbf{f})_{T_k} \right) \right\|_1 \geq \|\mathbf{A}\mathbf{D}((\mathbf{D}^*\mathbf{f})_{T_0})\|_1 - \sum_{k \geq 1} \|\mathbf{A}\mathbf{D}((\mathbf{D}^*\mathbf{f})_{T_k})\|_1 \\ &\geq (1 - \delta) \|\mathbf{D}((\mathbf{D}^*\mathbf{f})_{T_0})\|_2 - \sum_{k \geq 1} (1 + \delta) \|\mathbf{D}((\mathbf{D}^*\mathbf{f})_{T_k})\|_2, \end{aligned}$$

where the last step used \mathbf{D} -RIP $_1(t, \delta)$. We notice that, for $k \geq 1$,

$$\|\mathbf{D}((\mathbf{D}^*\mathbf{f})_{T_k})\|_2 \leq \|(\mathbf{D}^*\mathbf{f})_{T_k}\|_2 \leq \frac{1}{\sqrt{t}} \|(\mathbf{D}^*\mathbf{f})_{T_{k-1}}\|_1,$$

from where it follows that

$$(9) \quad \|\mathbf{A}\mathbf{f}\|_1 \geq (1 - \delta) \|\mathbf{D}((\mathbf{D}^*\mathbf{f})_{T_0})\|_2 - \frac{1 + \delta}{\sqrt{t}} \|\mathbf{D}^*\mathbf{f}\|_1.$$

In addition, we observe that

$$\begin{aligned} \|\mathbf{D}^*\mathbf{f}\|_2 &= \|\mathbf{f}\|_2 = \|\mathbf{D}\mathbf{D}^*\mathbf{f}\|_2 = \left\| \mathbf{D} \left(\sum_{k \geq 0} (\mathbf{D}^*\mathbf{f})_{T_k} \right) \right\|_2 \leq \|\mathbf{D}((\mathbf{D}^*\mathbf{f})_{T_0})\|_2 + \sum_{k \geq 1} \|\mathbf{D}((\mathbf{D}^*\mathbf{f})_{T_k})\|_2 \\ &\leq \|\mathbf{D}((\mathbf{D}^*\mathbf{f})_{T_0})\|_2 + \frac{1}{\sqrt{t}} \|\mathbf{D}^*\mathbf{f}\|_1. \end{aligned}$$

In view of the effective sparsity of $\mathbf{D}^*\mathbf{f}$, we obtain

$$\|\mathbf{D}^*\mathbf{f}\|_1 \leq \sqrt{s} \|\mathbf{D}^*\mathbf{f}\|_2 \leq \sqrt{s} \|\mathbf{D}((\mathbf{D}^*\mathbf{f})_{T_0})\|_2 + \sqrt{s/t} \|\mathbf{D}^*\mathbf{f}\|_1,$$

hence

$$(10) \quad \|\mathbf{D}((\mathbf{D}^*\mathbf{f})_{T_0})\|_2 \geq \frac{1 - \sqrt{s/t}}{\sqrt{s}} \|\mathbf{D}^*\mathbf{f}\|_1.$$

Substituting (10) in (9) yields

$$(11) \quad \|\mathbf{A}\mathbf{f}\|_1 \geq \left((1 - \delta)(1 - \sqrt{s/t}) - (1 + \delta)(\sqrt{s/t}) \right) \frac{1}{\sqrt{s}} \|\mathbf{D}^*\mathbf{f}\|_1 = \frac{2/5}{\sqrt{s}} \|\mathbf{D}^*\mathbf{f}\|_1,$$

where we have used the values $t = 25s$ and $\delta = 1/5$. This lower estimate for $\|\mathbf{A}\mathbf{f}\|_1$, combined with the minimality property of \mathbf{f}_{lp} , allows us to derive that

$$(12) \quad \|\mathbf{D}^*\mathbf{f}_{\text{lp}}\|_1 \leq \|\mathbf{D}^*(\mathbf{f}/\|\mathbf{A}\mathbf{f}\|_1)\|_1 = \frac{\|\mathbf{D}^*\mathbf{f}\|_1}{\|\mathbf{A}\mathbf{f}\|_1} \leq (5/2)\sqrt{s}.$$

Next, with \widehat{T}_0 denoting an index set of t largest absolute entries of $\mathbf{D}^*\mathbf{f}_{\text{lp}}$, \widehat{T}_1 an index set of next t largest absolute entries of $\mathbf{D}^*\mathbf{f}_{\text{lp}}$, \widehat{T}_2 an index set of next t largest absolute entries of $\mathbf{D}^*\mathbf{f}_{\text{lp}}$, etc., we can write

$$\begin{aligned} 1 &= \|\mathbf{A}\mathbf{f}_{\text{lp}}\|_1 = \|\mathbf{A}\mathbf{D}\mathbf{D}^*\mathbf{f}_{\text{lp}}\|_1 = \left\| \mathbf{A}\mathbf{D} \left(\sum_{k \geq 0} (\mathbf{D}^*\mathbf{f}_{\text{lp}})_{\widehat{T}_k} \right) \right\|_1 \leq \sum_{k \geq 0} \left\| \mathbf{A}\mathbf{D} \left((\mathbf{D}^*\mathbf{f}_{\text{lp}})_{\widehat{T}_k} \right) \right\|_1 \\ &\leq \sum_{k \geq 0} (1 + \delta) \left\| \mathbf{D} \left((\mathbf{D}^*\mathbf{f}_{\text{lp}})_{\widehat{T}_k} \right) \right\|_2 = (1 + \delta) \left[\left\| (\mathbf{D}^*\mathbf{f}_{\text{lp}})_{\widehat{T}_0} \right\|_2 + \sum_{k \geq 1} \left\| (\mathbf{D}^*\mathbf{f}_{\text{lp}})_{\widehat{T}_k} \right\|_2 \right] \\ &\leq (1 + \delta) \left[\|\mathbf{D}^*\mathbf{f}_{\text{lp}}\|_2 + \frac{1}{\sqrt{t}} \|\mathbf{D}^*\mathbf{f}_{\text{lp}}\|_1 \right] \leq (1 + \delta) \left[\|\mathbf{D}^*\mathbf{f}_{\text{lp}}\|_2 + (5/2)\sqrt{s/t} \right]. \end{aligned}$$

This chain of inequalities shows that

$$(13) \quad \|\mathbf{D}^*\mathbf{f}_{\text{lp}}\|_2 \geq \frac{1 - (5/2)\sqrt{s/t}}{1 + \delta} = \frac{5}{12}.$$

Combining (12) and (13), we obtain

$$\|\mathbf{D}^*\mathbf{f}_{\text{lp}}\|_1 \leq 6\sqrt{s}\|\mathbf{D}^*\mathbf{f}_{\text{lp}}\|_2.$$

In other words, $\mathbf{D}^*\mathbf{f}_{\text{lp}}$ is effectively $36s$ -sparse, which is what was needed to conclude the proof. \square

Remark 2. We point out that if \mathbf{f} was genuinely, instead of effectively, s -analysis-sparse, then a lower bound of the type (11) would be immediate from the \mathbf{D} -RIP₁. We also point out that our method of proving that the linear program outputs an effectively analysis-sparse signal is new even in the case $\mathbf{D} = \mathbf{I}_n$. In fact, it makes it possible to remove a logarithmic factor from the number of measurements in this “nondictionary” case, too (compare with [PV13a]). Furthermore, it allows for an analysis of the linear program (8) only based on deterministic conditions that the matrix \mathbf{A} may satisfy.

3.2 Hard thresholding

Given a signal $\mathbf{f} \in \mathbb{R}^n$ observed via $\mathbf{y} = \text{sgn}(\mathbf{A}\mathbf{f})$, the hard thresholding scheme we consider here consists in constructing a signal $\mathbf{f}_{\text{ht}} \in \mathbb{R}^n$ as

$$(14) \quad \mathbf{f}_{\text{ht}} = \mathbf{D}\mathbf{z}, \quad \text{where } \mathbf{z} := H_t(\mathbf{D}^*\mathbf{A}^*\mathbf{y}).$$

Our recovery result holds for s -synthesis sparse signals that are also effectively κs -analysis-sparse for some $\kappa \geq 1$ (we discussed in the introduction some choices of dictionaries \mathbf{D} making this happen).

Theorem 7. If $\mathbf{A} \in \mathbb{R}^{m \times n}$ satisfies \mathbf{D} -SPEP($s+t, \varepsilon/8$), $t = \lceil 16\varepsilon^{-2}\kappa s \rceil$, then any s -synthesis-sparse signal $\mathbf{f} \in \mathbf{D}(\Sigma_s^N)$ with $\mathbf{D}^*\mathbf{f} \in \Sigma_{\kappa s}^{N, \text{eff}}$ observed via $\mathbf{y} = \text{sgn}(\mathbf{A}\mathbf{f})$ is directionally approximated by the output \mathbf{f}_{ht} of the hard thresholding (14) with error

$$\left\| \frac{\mathbf{f}}{\|\mathbf{f}\|_2} - \frac{\mathbf{f}_{\text{ht}}}{\|\mathbf{f}_{\text{ht}}\|_2} \right\|_2 \leq \varepsilon.$$

Proof. We assume without loss of generality that $\|\mathbf{f}\|_2 = 1$. Let $T = T_0$ denote an index set of t largest absolute entries of $\mathbf{D}^*\mathbf{f}$, T_1 an index set of next t largest absolute entries of $\mathbf{D}^*\mathbf{f}$, T_2 an index set of next t largest absolute entries of $\mathbf{D}^*\mathbf{f}$, etc.. We start by noticing that \mathbf{z} is a better t -sparse approximation to $\mathbf{D}^*\mathbf{A}^*\mathbf{y} = \mathbf{D}^*\mathbf{A}^*\text{sgn}(\mathbf{A}\mathbf{f})$ than $[\mathbf{D}^*\mathbf{f}]_T$, so we can write

$$\|\mathbf{D}^*\mathbf{A}^*\text{sgn}(\mathbf{A}\mathbf{f}) - \mathbf{z}\|_2^2 \leq \|\mathbf{D}^*\mathbf{A}^*\text{sgn}(\mathbf{A}\mathbf{f}) - [\mathbf{D}^*\mathbf{f}]_T\|_2^2,$$

i.e.,

$$\|(\mathbf{D}^*\mathbf{f} - \mathbf{z}) - (\mathbf{D}^*\mathbf{f} - \mathbf{D}^*\mathbf{A}^*\text{sgn}(\mathbf{A}\mathbf{f}))\|_2^2 \leq \|(\mathbf{D}^*\mathbf{f} - \mathbf{D}^*\mathbf{A}^*\text{sgn}(\mathbf{A}\mathbf{f})) - [\mathbf{D}^*\mathbf{f}]_T\|_2^2.$$

Expanding the squares and rearranging gives

$$(15) \quad \|\mathbf{D}^*\mathbf{f} - \mathbf{z}\|_2^2 \leq 2\langle \mathbf{D}^*\mathbf{f} - \mathbf{z}, \mathbf{D}^*\mathbf{f} - \mathbf{D}^*\mathbf{A}^*\text{sgn}(\mathbf{A}\mathbf{f}) \rangle$$

$$(16) \quad - 2\langle [\mathbf{D}^*\mathbf{f}]_T, \mathbf{D}^*\mathbf{f} - \mathbf{D}^*\mathbf{A}^*\text{sgn}(\mathbf{A}\mathbf{f}) \rangle$$

$$(17) \quad + \|[\mathbf{D}^*\mathbf{f}]_T\|_2^2.$$

To bound (17), we invoke [FR13, Theorem 2.5] and the effective analysis-sparsity of \mathbf{f} to derive

$$\|[\mathbf{D}^*\mathbf{f}]_T\|_2^2 \leq \frac{1}{4t} \|\mathbf{D}^*\mathbf{f}\|_1^2 \leq \frac{\kappa s}{4t} \|\mathbf{D}^*\mathbf{f}\|_2^2 = \frac{\kappa s}{4t} \|\mathbf{f}\|_2^2 = \frac{\kappa s}{4t}.$$

To bound (15) in absolute value, we notice that it can be written as

$$\begin{aligned} 2|\langle \mathbf{D}\mathbf{D}^*\mathbf{f} - \mathbf{D}\mathbf{z}, \mathbf{f} - \mathbf{A}^*\text{sgn}(\mathbf{A}\mathbf{f}) \rangle| &= 2|\langle \mathbf{f} - \mathbf{f}_{\text{ht}}, \mathbf{f} - \mathbf{A}^*\text{sgn}(\mathbf{A}\mathbf{f}) \rangle| \\ &= 2|\langle \mathbf{f} - \mathbf{f}_{\text{ht}}, \mathbf{f} \rangle - \langle \mathbf{A}(\mathbf{f} - \mathbf{f}_{\text{ht}}), \text{sgn}(\mathbf{A}\mathbf{f}) \rangle| \leq 2\varepsilon' \|\mathbf{f} - \mathbf{f}_{\text{ht}}\|_2, \end{aligned}$$

where the last step followed from \mathbf{D} -SPEP($s + t, \varepsilon'$), $\varepsilon' := \varepsilon/8$. Finally, (16) can be bounded in absolute value by

$$\begin{aligned} 2 \sum_{k \geq 1} |\langle [\mathbf{D}^* \mathbf{f}]_{T_k}, \mathbf{D}^*(\mathbf{f} - \mathbf{A}^* \text{sgn}(\mathbf{A}\mathbf{f})) \rangle| &= 2 \sum_{k \geq 1} |\langle \mathbf{D}([\mathbf{D}^* \mathbf{f}]_{T_k}), \mathbf{f} - \mathbf{A}^* \text{sgn}(\mathbf{A}\mathbf{f}) \rangle| \\ &= 2 \sum_{k \geq 1} |\langle \mathbf{D}([\mathbf{D}^* \mathbf{f}]_{T_k}), \mathbf{f} \rangle - \langle \mathbf{A}(\mathbf{D}([\mathbf{D}^* \mathbf{f}]_{T_k})), \text{sgn}(\mathbf{A}\mathbf{f}) \rangle| \leq 2 \sum_{k \geq 1} \varepsilon' \|\mathbf{D}([\mathbf{D}^* \mathbf{f}]_{T_k})\|_2 \\ &\leq 2\varepsilon' \sum_{k \geq 1} \|\mathbf{D}([\mathbf{D}^* \mathbf{f}]_{T_k})\|_2 \leq 2\varepsilon' \sum_{k \geq 1} \frac{\|[\mathbf{D}^* \mathbf{f}]_{T_{k-1}}\|_1}{\sqrt{t}} \leq 2\varepsilon' \frac{\|\mathbf{D}^* \mathbf{f}\|_1}{\sqrt{t}} \leq 2\varepsilon' \frac{\sqrt{\kappa S} \|\mathbf{D}^* \mathbf{f}\|_2}{\sqrt{t}} = 2\varepsilon' \sqrt{\frac{\kappa S}{t}}. \end{aligned}$$

Putting everything together, we obtain

$$\|\mathbf{D}^* \mathbf{f} - \mathbf{z}\|_2^2 \leq 2\varepsilon' \|\mathbf{f} - \mathbf{f}_{\text{ht}}\|_2 + 2\varepsilon' \sqrt{\frac{\kappa S}{t}} + \frac{\kappa S}{4t}.$$

In view of $\|\mathbf{f} - \mathbf{f}_{\text{ht}}\|_2 = \|\mathbf{D}(\mathbf{D}^* \mathbf{f} - \mathbf{z})\|_2 \leq \|\mathbf{D}^* \mathbf{f} - \mathbf{z}\|_2$, it follows that

$$\|\mathbf{f} - \mathbf{f}_{\text{ht}}\|_2^2 \leq 2\varepsilon' \|\mathbf{f} - \mathbf{f}_{\text{ht}}\|_2 + 2\varepsilon' \sqrt{\frac{\kappa S}{t}} + \frac{\kappa S}{4t}, \quad \text{i.e., } (\|\mathbf{f} - \mathbf{f}_{\text{ht}}\|_2 - \varepsilon')^2 \leq \varepsilon'^2 + 2\varepsilon' \sqrt{\frac{\kappa S}{t}} + \frac{\kappa S}{4t} \leq \left(\varepsilon' + \sqrt{\frac{\kappa S}{t}}\right)^2.$$

This implies that

$$\|\mathbf{f} - \mathbf{f}_{\text{ht}}\|_2 \leq 2\varepsilon' + \sqrt{\frac{\kappa S}{t}}.$$

Finally, since $\mathbf{f}_{\text{ht}}/\|\mathbf{f}_{\text{ht}}\|_2$ is the best ℓ_2 -normalized approximation to \mathbf{f}_{ht} , we conclude that

$$\left\| \mathbf{f} - \frac{\mathbf{f}_{\text{ht}}}{\|\mathbf{f}_{\text{ht}}\|_2} \right\|_2 \leq \|\mathbf{f} - \mathbf{f}_{\text{ht}}\|_2 + \left\| \mathbf{f}_{\text{ht}} - \frac{\mathbf{f}_{\text{ht}}}{\|\mathbf{f}_{\text{ht}}\|_2} \right\|_2 \leq 2\|\mathbf{f} - \mathbf{f}_{\text{ht}}\|_2 \leq 4\varepsilon' + 2\sqrt{\frac{\kappa S}{t}}.$$

The announced result follows from our choices of t and ε' . \square

4 Signal estimation: direction and magnitude

Since information of the type $y_i = \text{sgn}\langle \mathbf{a}_i, \mathbf{f} \rangle$ can at best allow one to estimate the direction of a signal $\mathbf{f} \in \mathbb{R}^n$, we consider in this section information of the type

$$y_i = \text{sgn}(\langle \mathbf{a}_i, \mathbf{f} \rangle - \tau_i), \quad i = 1, \dots, m,$$

for some thresholds τ_1, \dots, τ_m introduced before quantization. In the rest of this section, we give three methods for recovering \mathbf{f} in its entirety. The first one is based on linear programming, the second one on second-order cone programming, and the last one on hard thresholding.

We are going to show that using these algorithms, one can estimate both the direction and the magnitude of dictionary-sparse signal $\mathbf{f} \in \mathbb{R}^n$ given a prior magnitude bound such as $\|\mathbf{f}\|_2 \leq r$.

We simply rely on the previous results by “lifting” the situation from \mathbb{R}^n to \mathbb{R}^{n+1} , in view of the observation that $\mathbf{y} = \text{sgn}(\mathbf{A}\mathbf{f} - \boldsymbol{\tau})$ can be interpreted as

$$\mathbf{y} = \text{sgn}(\tilde{\mathbf{A}}\tilde{\mathbf{f}}), \quad \text{where} \quad \tilde{\mathbf{A}} := \begin{bmatrix} \mathbf{A} & \begin{matrix} -\tau_1/c \\ \vdots \\ -\tau_m/c \end{matrix} \end{bmatrix} \in \mathbb{R}^{m \times (n+1)}, \quad \tilde{\mathbf{f}} := \begin{bmatrix} \mathbf{f} \\ c \end{bmatrix} \in \mathbb{R}^{n+1}.$$

The following lemma will be equally useful when dealing with linear programming, second-order cone programming, or with hard thresholding schemes.

Lemma 8. For $\tilde{\mathbf{f}}, \tilde{\mathbf{g}} \in \mathbb{R}^{n+1}$ written as

$$\tilde{\mathbf{f}} := \begin{bmatrix} \mathbf{f}_{[n]} \\ f_{n+1} \end{bmatrix} \quad \text{and} \quad \tilde{\mathbf{g}} := \begin{bmatrix} \mathbf{g}_{[n]} \\ g_{n+1} \end{bmatrix}$$

with $\tilde{\mathbf{f}}_{[n]}, \tilde{\mathbf{g}}_{[n]} \in \mathbb{R}^n$ and with $f_{n+1} \neq 0, g_{n+1} \neq 0$, one has

$$\left\| \frac{\mathbf{f}_{[n]}}{f_{n+1}} - \frac{\mathbf{g}_{[n]}}{g_{n+1}} \right\|_2 \leq \frac{\|\tilde{\mathbf{f}}\|_2 \|\tilde{\mathbf{g}}\|_2}{|f_{n+1}| |g_{n+1}|} \left\| \frac{\tilde{\mathbf{f}}}{\|\tilde{\mathbf{f}}\|_2} - \frac{\tilde{\mathbf{g}}}{\|\tilde{\mathbf{g}}\|_2} \right\|_2.$$

Proof. By using the triangle inequality in \mathbb{R}^n and Cauchy–Schwarz inequality in \mathbb{R}^2 , we can write

$$\begin{aligned} \left\| \frac{\mathbf{f}_{[n]}}{f_{n+1}} - \frac{\mathbf{g}_{[n]}}{g_{n+1}} \right\|_2 &= \|\tilde{\mathbf{f}}\|_2 \left\| \frac{1/f_{n+1} \mathbf{f}_{[n]} - 1/g_{n+1} \mathbf{g}_{[n]}}{\|\tilde{\mathbf{f}}\|_2} \right\|_2 \\ &\leq \|\tilde{\mathbf{f}}\|_2 \left(\frac{1}{f_{n+1}} \left\| \frac{\mathbf{f}_{[n]}}{\|\tilde{\mathbf{f}}\|_2} - \frac{\mathbf{g}_{[n]}}{\|\tilde{\mathbf{g}}\|_2} \right\|_2 + \left| \frac{1/g_{n+1} - 1/f_{n+1}}{\|\tilde{\mathbf{f}}\|_2} \right| \|\mathbf{g}_{[n]}\|_2 \right) \\ &= \|\tilde{\mathbf{f}}\|_2 \left(\frac{1}{f_{n+1}} \left\| \frac{\mathbf{f}_{[n]}}{\|\tilde{\mathbf{f}}\|_2} - \frac{\mathbf{g}_{[n]}}{\|\tilde{\mathbf{g}}\|_2} \right\|_2 + \frac{\|\mathbf{g}_{[n]}\|_2}{|f_{n+1}| |g_{n+1}|} \left| \frac{f_{n+1}}{\|\tilde{\mathbf{f}}\|_2} - \frac{g_{n+1}}{\|\tilde{\mathbf{g}}\|_2} \right| \right) \\ &\leq \|\tilde{\mathbf{f}}\|_2 \left[\frac{1}{|f_{n+1}|^2} + \frac{\|\mathbf{g}_{[n]}\|_2^2}{|f_{n+1}|^2 |g_{n+1}|^2} \right]^{1/2} \left[\left\| \frac{\mathbf{f}_{[n]}}{\|\tilde{\mathbf{f}}\|_2} - \frac{\mathbf{g}_{[n]}}{\|\tilde{\mathbf{g}}\|_2} \right\|_2^2 + \left| \frac{f_{n+1}}{\|\tilde{\mathbf{f}}\|_2} - \frac{g_{n+1}}{\|\tilde{\mathbf{g}}\|_2} \right|^2 \right]^{1/2} \\ &= \|\tilde{\mathbf{f}}\|_2 \left[\frac{\|\tilde{\mathbf{g}}\|_2^2}{|f_{n+1}|^2 |g_{n+1}|^2} \right]^{1/2} \left\| \frac{\tilde{\mathbf{f}}}{\|\tilde{\mathbf{f}}\|_2} - \frac{\tilde{\mathbf{g}}}{\|\tilde{\mathbf{g}}\|_2} \right\|_2, \end{aligned}$$

which is the announced result. \square

4.1 Linear programming

Given a signal $\mathbf{f} \in \mathbb{R}^n$ observed via $\mathbf{y} = \text{sgn}(\mathbf{A}\mathbf{f} - \boldsymbol{\tau})$ with $\tau_1, \dots, \tau_m \sim \mathcal{N}(0, \sigma^2)$, the optimization scheme we consider here consists in outputting the signal

$$(18) \quad \mathbf{f}_{\text{LP}} = \frac{\sigma}{\hat{u}} \hat{\mathbf{h}} \in \mathbb{R}^n,$$

where $\hat{\mathbf{h}} \in \mathbb{R}^n$ and $\hat{u} \in \mathbb{R}$ are solutions of

$$(19) \quad \underset{\mathbf{h} \in \mathbb{R}^n, u \in \mathbb{R}}{\text{minimize}} \quad \|\mathbf{D}^* \mathbf{h}\|_1 + |u| \quad \text{subject to} \quad \text{sgn}(\mathbf{A} \mathbf{h} - u \boldsymbol{\tau} / \sigma) = \mathbf{y}, \quad \|\mathbf{A} \mathbf{h} - u \boldsymbol{\tau} / \sigma\|_1 = 1.$$

Theorem 9. Let $\varepsilon, r, \sigma > 0$, let $m \geq C(r/\sigma + \sigma/r)^6 \varepsilon^{-6} s \ln(eN/s)$, and let $\mathbf{A} \in \mathbb{R}^{m \times n}$ be populated by independent standard normal random variables. Furthermore, let τ_1, \dots, τ_m be independent normal random variables with mean zero and variance σ^2 that are also independent from the entries of \mathbf{A} . Then, with failure probability at most $\gamma \exp(-cm\varepsilon^2 r^2 \sigma^2 / (r^2 + \sigma^2)^2)$, any effectively s -analysis sparse $\mathbf{f} \in \mathbb{R}^n$ satisfying $\|\mathbf{f}\|_2 \leq r$ and observed via $\mathbf{y} = \text{sgn}(\mathbf{A} \mathbf{f} - \boldsymbol{\tau})$ is approximated by \mathbf{f}_{LP} given in (18) with error

$$\|\mathbf{f} - \mathbf{f}_{\text{LP}}\|_2 \leq \varepsilon r.$$

Proof. Let us introduce the “lifted” signal $\tilde{\mathbf{f}} \in \mathbb{R}^{n+1}$, the “lifted” tight frame $\tilde{\mathbf{D}} \in \mathbb{R}^{(n+1) \times (n+1)}$, and the “lifted” measurement matrix $\tilde{\mathbf{A}} \in \mathbb{R}^{m \times (n+1)}$ defined as

$$(20) \quad \tilde{\mathbf{f}} := \begin{bmatrix} \mathbf{f} \\ \sigma \end{bmatrix}, \quad \tilde{\mathbf{D}} := \left[\begin{array}{c|c} \mathbf{D} & \mathbf{0} \\ \hline \mathbf{0} & 1 \end{array} \right], \quad \tilde{\mathbf{A}} := \left[\begin{array}{c|c} & \begin{matrix} -\tau_1/\sigma \\ \vdots \\ -\tau_m/\sigma \end{matrix} \\ \hline \mathbf{A} & \end{array} \right].$$

First, we observe that $\tilde{\mathbf{f}}$ is effectively $(s+1)$ -analysis-sparse (relative to $\tilde{\mathbf{D}}$), since $\tilde{\mathbf{D}}^* \tilde{\mathbf{f}} = \begin{bmatrix} \mathbf{D}^* \mathbf{f} \\ \sigma \end{bmatrix}$, hence

$$\frac{\|\tilde{\mathbf{D}}^* \tilde{\mathbf{f}}\|_1}{\|\tilde{\mathbf{D}}^* \tilde{\mathbf{f}}\|_2} = \frac{\|\mathbf{D}^* \mathbf{f}\|_1 + \sigma}{\sqrt{\|\mathbf{D}^* \mathbf{f}\|_2^2 + \sigma^2}} \leq \frac{\sqrt{s} \|\mathbf{D}^* \mathbf{f}\|_2 + \sigma}{\sqrt{\|\mathbf{D}^* \mathbf{f}\|_2^2 + \sigma^2}} \leq \sqrt{s+1}.$$

Next, we observe that the matrix $\tilde{\mathbf{A}} \in \mathbb{R}^{m \times (n+1)}$, populated by independent standard normal random variables, satisfies $\tilde{\mathbf{D}}$ -TES $(36(s+1), \varepsilon')$, $\varepsilon' := \frac{r\sigma}{2(r^2 + \sigma^2)} \varepsilon$, and $\tilde{\mathbf{D}}$ -RIP $_1(25(s+1), 1/5)$ with failure probability at most $\gamma \exp(-cm\varepsilon'^2) + \gamma' \exp(-c'm) \leq \gamma'' \exp(-c''m\varepsilon^2 r^2 \sigma^2 / (r^2 + \sigma^2)^2)$, since $m \geq C\varepsilon'^{-6}(s+1) \ln(eN/(s+1))$ and $m \geq C(1/5)^{-7}(s+1) \ln(eN/(s+1))$ are ensured by our assumption on m . Finally, we observe that $\mathbf{y} = \text{sgn}(\tilde{\mathbf{A}} \tilde{\mathbf{f}})$ and that the optimization program (19) reads

$$\underset{\tilde{\mathbf{h}} \in \mathbb{R}^{n+1}}{\text{minimize}} \quad \|\tilde{\mathbf{D}}^* \tilde{\mathbf{h}}\|_1 \quad \text{subject to} \quad \text{sgn}(\tilde{\mathbf{A}} \tilde{\mathbf{h}}) = \mathbf{y}, \quad \|\tilde{\mathbf{A}} \tilde{\mathbf{h}}\|_1 = 1.$$

Denoting its solution as $\tilde{\mathbf{g}} = \begin{bmatrix} \mathbf{g}_{[n]} \\ g_{n+1} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{h}} \\ \hat{u} \end{bmatrix} \in \mathbb{R}^{n+1}$, Theorem 6 implies that

$$\left\| \frac{\tilde{\mathbf{f}}}{\|\tilde{\mathbf{f}}\|_2} - \frac{\tilde{\mathbf{g}}}{\|\tilde{\mathbf{g}}\|_2} \right\|_2 \leq \varepsilon'.$$

In particular, looking at the last coordinate, this inequality yields

$$\left| \frac{\sigma}{\|\tilde{\mathbf{f}}\|_2} - \frac{g_{n+1}}{\|\tilde{\mathbf{g}}\|_2} \right| \leq \varepsilon', \quad \text{hence} \quad \frac{|g_{n+1}|}{\|\tilde{\mathbf{g}}\|_2} \geq \frac{\sigma}{\|\tilde{\mathbf{f}}\|_2} - \varepsilon' \geq \frac{\sigma}{\sqrt{r^2 + \sigma^2}} - \frac{\sigma}{2\sqrt{r^2 + \sigma^2}} = \frac{\sigma}{2\sqrt{r^2 + \sigma^2}}.$$

In turn, applying Lemma 8 while taking $\mathbf{f} = \mathbf{f}_{[n]}$ and $\mathbf{f}_{\text{LP}} = (\sigma/g_{n+1})\mathbf{g}_{[n]}$ into consideration gives

$$\left\| \frac{\mathbf{f}}{\sigma} - \frac{\mathbf{f}_{\text{LP}}}{\sigma} \right\|_2 \leq \frac{\|\tilde{\mathbf{f}}\|_2}{\sigma} \frac{\|\tilde{\mathbf{g}}\|_2}{|g_{n+1}|} \varepsilon' \leq \frac{\|\tilde{\mathbf{f}}\|_2}{\sigma} \frac{2\sqrt{r^2 + \sigma^2}}{\sigma} \frac{r\sigma}{2(r^2 + \sigma^2)} \varepsilon = \frac{\|\tilde{\mathbf{f}}\|_2}{\sigma} \frac{r}{\sqrt{r^2 + \sigma^2}} \varepsilon,$$

so that

$$\|\mathbf{f} - \mathbf{f}_{\text{LP}}\|_2 \leq \|\tilde{\mathbf{f}}\|_2 \frac{r}{\sqrt{r^2 + \sigma^2}} \varepsilon \leq r\varepsilon.$$

This establishes the announced result. \square

Remark 3. The recovery scheme (19) does not require an estimation of r to be run. The recovery scheme presented next does require such an estimation. Moreover, it is a second-order cone program instead of a simpler linear program. But it has one noticeable advantage, namely that it not only applies to signals satisfying $\|\mathbf{D}^*\mathbf{f}\|_1 \leq \sqrt{s}\|\mathbf{D}^*\mathbf{f}\|_2$ and $\|\mathbf{D}^*\mathbf{f}\|_2 \leq r$ but more generally to signals satisfying $\|\mathbf{D}^*\mathbf{f}\|_1 \leq \sqrt{sr}$ and $\|\mathbf{D}^*\mathbf{f}\|_2 \leq r$. For both schemes, one needs σ to be of the same order as r for the results to become meaningful in terms of number of measurement and success probability. However, if r is only upper-estimated, then one could choose $\sigma \geq r$ and obtain a weaker recovery error $\|\mathbf{f} - \hat{\mathbf{f}}\|_2 \leq \varepsilon\sigma$ with relevant number of measurement and success probability.

4.2 Second-order cone programming

Given a signal $\mathbf{f} \in \mathbb{R}^n$ observed via $\mathbf{y} = \text{sgn}(\mathbf{A}\mathbf{f} - \boldsymbol{\tau})$ with $\tau_1, \dots, \tau_m \sim \mathcal{N}(0, \sigma^2)$, the optimization scheme we consider here consists in outputting the signal

$$(21) \quad \mathbf{f}_{\text{CP}} = \underset{\mathbf{h} \in \mathbb{R}^n}{\text{argmin}} \|\mathbf{D}^*\mathbf{h}\|_1 \quad \text{subject to} \quad \text{sgn}(\mathbf{A}\mathbf{h} - \boldsymbol{\tau}) = \mathbf{y}, \quad \|\mathbf{h}\|_2 \leq r.$$

Theorem 10. Let $\varepsilon, r, \sigma > 0$, let $m \geq C(r/\sigma + \sigma/r)^6(r^2/\sigma^2 + 1)\varepsilon^{-6}s \ln(eN/s)$, and let $\mathbf{A} \in \mathbb{R}^{m \times n}$ be populated by independent standard normal random variables. Furthermore, let τ_1, \dots, τ_m be independent normal random variables with mean zero and variance σ^2 that are also independent from \mathbf{A} . Then, with failure probability at most $\gamma \exp(-c'm\varepsilon^2 r^2 \sigma^2 / (r^2 + \sigma^2)^2)$, any signal $\mathbf{f} \in \mathbb{R}^n$ with $\|\mathbf{f}\|_2 \leq r$, $\|\mathbf{D}^*\mathbf{f}\|_1 \leq \sqrt{sr}$, and observed via $\mathbf{y} = \text{sgn}(\mathbf{A}\mathbf{f} - \boldsymbol{\tau})$ is approximated by \mathbf{f}_{CP} given in (21) with error

$$\|\mathbf{f} - \mathbf{f}_{\text{CP}}\|_2 \leq \varepsilon r.$$

Proof. We again use the notation (20) introducing the “lifted” objects $\tilde{\mathbf{f}}$, $\tilde{\mathbf{D}}$, and $\tilde{\mathbf{A}}$. Moreover, we set $\tilde{\mathbf{g}} := \begin{bmatrix} \mathbf{f}_{\text{CP}} \\ \sigma \end{bmatrix}$. We claim that $\tilde{\mathbf{f}}$ and $\tilde{\mathbf{g}}$ are effectively s' -analysis-sparse, $s' := (r^2/\sigma^2 + 1)(s + 1)$.

For $\tilde{\mathbf{g}}$, this indeed follows from $\|\tilde{\mathbf{D}}^*\tilde{\mathbf{g}}\|_2 = \|\tilde{\mathbf{g}}\|_2 = \sqrt{\|\mathbf{f}_{\text{CP}}\|_2^2 + \sigma^2} \geq \sigma$ and

$$\|\tilde{\mathbf{D}}^*\tilde{\mathbf{g}}\|_1 = \left\| \begin{bmatrix} \mathbf{D}^*\mathbf{f}_{\text{CP}} \\ \sigma \end{bmatrix} \right\|_1 = \|\mathbf{D}^*\mathbf{f}_{\text{CP}}\|_1 + \sigma \leq \|\mathbf{D}^*\mathbf{f}\|_1 + \sigma \leq \sqrt{sr} + \sigma \leq \sqrt{r^2 + \sigma^2} \sqrt{s + 1}.$$

We also notice that $\tilde{\mathbf{A}}$ satisfies $\tilde{\mathbf{D}}$ -TES(s', ε'), $\varepsilon' := \frac{r\sigma}{r^2 + \sigma^2}\varepsilon$, with failure probability at most $\gamma \exp(-cm\varepsilon'^2) \leq \gamma \exp(-c'm\varepsilon^2 r^2 \sigma^2 / (r^2 + \sigma^2)^2)$, since $m \geq C\varepsilon'^{-6} s' \ln(eN/s')$ is ensured by our assumption on m . Finally, we observe that both $\tilde{\mathbf{f}}/\|\tilde{\mathbf{f}}\|_2$ and $\tilde{\mathbf{g}}/\|\tilde{\mathbf{g}}\|_2$ are ℓ_2 -normalized effectively s' -analysis-sparse and have the same sign observations $\text{sgn}(\tilde{\mathbf{A}}\tilde{\mathbf{f}}) = \text{sgn}(\tilde{\mathbf{A}}\tilde{\mathbf{g}}) = \mathbf{y}$. Thus,

$$\left\| \frac{\tilde{\mathbf{f}}}{\|\tilde{\mathbf{f}}\|_2} - \frac{\tilde{\mathbf{g}}}{\|\tilde{\mathbf{g}}\|_2} \right\|_2 \leq \varepsilon'.$$

In view of Lemma 8, we derive

$$\left\| \frac{\mathbf{f}}{\sigma} - \frac{\mathbf{f}_{\text{CP}}}{\sigma} \right\|_2 \leq \frac{r^2 + \sigma^2}{\sigma^2} \varepsilon', \quad \text{hence} \quad \|\mathbf{f} - \mathbf{f}_{\text{CP}}\|_2 \leq \frac{r^2 + \sigma^2}{\sigma} \varepsilon' = r\varepsilon.$$

This establishes the announced result. \square

4.3 Hard thresholding

Given a signal $\mathbf{f} \in \mathbb{R}^N$ observed via $\mathbf{y} = \text{sgn}(\mathbf{A}\mathbf{f} - \boldsymbol{\tau})$ with $\tau_1, \dots, \tau_m \sim \mathcal{N}(0, \sigma^2)$, the hard thresholding scheme we consider here consists in outputting the signal

$$(22) \quad \mathbf{f}_{\text{HT}} = \frac{-\sigma^2}{\langle \boldsymbol{\tau}, \mathbf{y} \rangle} \mathbf{D}\mathbf{z}, \quad \mathbf{z} = H_{t-1}(\mathbf{D}^* \mathbf{A}^* \mathbf{y}).$$

Theorem 11. Let $\varepsilon, r, \sigma > 0$, let $m \geq C\kappa(r/\sigma + \sigma/r)^9 \varepsilon^{-9} s \ln(eN/s)$, and let $\mathbf{A} \in \mathbb{R}^{m \times n}$ be populated by independent standard normal random variables. Furthermore, let τ_1, \dots, τ_m be independent normal random variables with mean zero and variance σ^2 that are also independent from the entries of \mathbf{A} . Then, with failure probability at most $\gamma \exp(-cm\varepsilon^2 r^2 \sigma^2 / (r^2 + \sigma^2)^2)$, any s -synthesis sparse and effectively κs -analysis sparse signal $\mathbf{f} \in \mathbb{R}^n$ satisfying $\|\mathbf{f}\|_2 \leq r$ and observed via $\mathbf{y} = \text{sgn}(\mathbf{A}\mathbf{f} - \boldsymbol{\tau})$ is approximated by \mathbf{f}_{HT} given in (22) for $t := \lceil 16(\varepsilon'/8)^{-2} \kappa(s+1) \rceil$ with error

$$\|\mathbf{f} - \mathbf{f}_{\text{HT}}\|_2 \leq \varepsilon r.$$

Proof. We again use the notation (20) for the “lifted” objects $\tilde{\mathbf{f}}$, $\tilde{\mathbf{D}}$, and $\tilde{\mathbf{A}}$. First, we notice that $\tilde{\mathbf{f}}$ is $(s+1)$ -synthesis sparse (relative to $\tilde{\mathbf{D}}$), as well as effectively $\kappa(s+1)$ -analysis sparse, since

$$\tilde{\mathbf{D}}^* \tilde{\mathbf{f}} = \begin{bmatrix} \mathbf{D}^* \mathbf{f} \\ \sigma \end{bmatrix} \text{ satisfies}$$

$$\frac{\|\tilde{\mathbf{D}}^* \tilde{\mathbf{f}}\|_1}{\|\tilde{\mathbf{D}}^* \tilde{\mathbf{f}}\|_2} = \frac{\|\mathbf{D}^* \mathbf{f}\|_1 + \sigma}{\sqrt{\|\mathbf{D}^* \mathbf{f}\|_2^2 + \sigma^2}} \leq \frac{\sqrt{\kappa s} \|\mathbf{D}^* \mathbf{f}\|_2 + \sigma}{\sqrt{\|\mathbf{D}^* \mathbf{f}\|_2^2 + \sigma^2}} \leq \sqrt{\kappa s + 1} \leq \sqrt{\kappa(s+1)}.$$

Next, we observe that the matrix $\tilde{\mathbf{A}}$, populated by independent standard normal random variables, satisfies $\tilde{\mathbf{D}}$ -SPEP($s+1+t, \varepsilon'/8$), $\varepsilon' := \frac{r\sigma}{2(r^2 + \sigma^2)}\varepsilon$, with failure probability at most $\gamma \exp(-cm\varepsilon'^2 r^2)$,

since $m \geq C(\varepsilon'/8)^{-7}(s+1+t) \ln(e(N+1)/(s+1+t))$ is ensured by our assumption on m . Finally, since $\mathbf{y} = \text{sgn}(\tilde{\mathbf{A}}\mathbf{f})$, Theorem 7 implies that

$$\left\| \frac{\tilde{\mathbf{f}}}{\|\tilde{\mathbf{f}}\|_2} - \frac{\tilde{\mathbf{g}}}{\|\tilde{\mathbf{g}}\|_2} \right\|_2 \leq \varepsilon',$$

where $\tilde{\mathbf{g}} \in \mathbb{R}^{n+1}$ is the output of the “lifted” hard thresholding scheme. i.e.,

$$\tilde{\mathbf{g}} = \tilde{\mathbf{D}}\tilde{\mathbf{z}}, \quad \tilde{\mathbf{z}} = H_t(\tilde{\mathbf{D}}^*\tilde{\mathbf{A}}^*\mathbf{y}),$$

In particular, looking at the last coordinate, this inequality yields

$$(23) \quad \left| \frac{\sigma}{\|\tilde{\mathbf{f}}\|_2} - \frac{g_{n+1}}{\|\tilde{\mathbf{g}}\|_2} \right| \leq \varepsilon', \quad \text{hence} \quad \frac{|g_{n+1}|}{\|\tilde{\mathbf{g}}\|_2} \geq \frac{\sigma}{\|\tilde{\mathbf{f}}\|_2} - \varepsilon' \geq \frac{\sigma}{\sqrt{r^2 + \sigma^2}} - \frac{\sigma}{2\sqrt{r^2 + \sigma^2}} = \frac{\sigma}{2\sqrt{r^2 + \sigma^2}}.$$

Now let us also observe that

$$\tilde{\mathbf{z}} = H_t \left(\begin{bmatrix} \mathbf{D}^*\mathbf{A}^*\mathbf{y} \\ -\langle \boldsymbol{\tau}, \mathbf{y} \rangle / \sigma \end{bmatrix} \right) = \begin{cases} \begin{bmatrix} H_t(\mathbf{D}^*\mathbf{A}^*\mathbf{y}) \\ 0 \end{bmatrix}, \\ \text{or} \\ \begin{bmatrix} H_{t-1}(\mathbf{D}^*\mathbf{A}^*\mathbf{y}) \\ -\langle \boldsymbol{\tau}, \mathbf{y} \rangle / \sigma \end{bmatrix}, \end{cases} \quad \text{hence} \quad \tilde{\mathbf{g}} = \tilde{\mathbf{D}}\tilde{\mathbf{z}} = \begin{cases} \begin{bmatrix} \mathbf{D}(H_t(\mathbf{D}^*\mathbf{A}^*\mathbf{y})) \\ 0 \end{bmatrix}, \\ \text{or} \\ \begin{bmatrix} \mathbf{D}(H_{t-1}(\mathbf{D}^*\mathbf{A}^*\mathbf{y})) \\ -\langle \boldsymbol{\tau}, \mathbf{y} \rangle / \sigma \end{bmatrix}. \end{cases}$$

In view of (23), the latter option prevails. It is then apparent that $\mathbf{f}_{\text{HT}} = \sigma \mathbf{g}_{[n]} / g_{n+1}$. Lemma 8 gives

$$\left\| \frac{\mathbf{f}}{\sigma} - \frac{\mathbf{f}_{\text{HT}}}{\sigma} \right\|_2 \leq \frac{\|\tilde{\mathbf{f}}\|_2}{\sigma} \frac{\|\tilde{\mathbf{g}}\|_2}{|g_{n+1}|} \varepsilon' \leq \frac{\|\tilde{\mathbf{f}}\|_2}{\sigma} \frac{2\sqrt{r^2 + \sigma^2}}{\sigma} \frac{r\sigma}{2(r^2 + \sigma^2)} \varepsilon = \frac{\|\tilde{\mathbf{f}}\|_2}{\sigma} \frac{r}{\sqrt{r^2 + \sigma^2}} \varepsilon,$$

so that

$$\|\mathbf{f} - \mathbf{f}_{\text{HT}}\|_2 \leq \|\tilde{\mathbf{f}}\|_2 \frac{r}{\sqrt{r^2 + \sigma^2}} \varepsilon \leq r\varepsilon.$$

This establishes the announced result. \square

5 Postponed proofs and further remarks

This final section contains the theoretical justification of the technical properties underlying our results, followed by a few points of discussion around them.

5.1 Proof of D-SPEP

The Gaussian width turns out to be a useful tool in our proofs. For a set $K \subseteq \mathbb{R}^n$, it is defined by

$$w(K) = \mathbb{E} \left[\sup_{\mathbf{f} \in K} \langle \mathbf{f}, \mathbf{g} \rangle \right], \quad \mathbf{g} \in \mathbb{R}^n \text{ is a standard normal random vector.}$$

We isolate the following two properties.

Lemma 12. Let $K \subseteq \mathbb{R}^n$ be a linear space and $K_1, \dots, K_L \subseteq \mathbb{R}^n$ be subsets of the unit sphere S^{n-1} .

- (i) $k/\sqrt{k+1} \leq w(K \cap S^{n-1}) \leq \sqrt{k}$, $k := \dim(K)$;
- (ii) $w(K_1 \cup \dots \cup K_L) \leq \max\{w(K_1), \dots, w(K_L)\} + 3\sqrt{\ln(L)}$.

Proof. (i) By the invariance under orthogonal transformation (see [PV13b, Proposition 2.1]⁴), we can assume that $K = \mathbb{R}^k \times \{(0, \dots, 0)\}$. We then notice that $\sup_{\mathbf{f} \in K \cap S^{n-1}} \langle \mathbf{f}, \mathbf{g} \rangle = \|(g_1, \dots, g_k)\|_2$ is the ℓ_2 -norm of a standard normal random vector of dimension k . We invoke e.g. [FR13, Proposition 8.1] to derive the announced result.

(ii) Let us introduce the nonnegative random variables

$$\xi_\ell := \sup_{\mathbf{f} \in K_\ell} \langle \mathbf{f}, \mathbf{g} \rangle, \quad \ell = 1, \dots, L,$$

so that the Gaussian widths of each K_ℓ and of their union take the form

$$w(K_\ell) = \mathbb{E}(\xi_\ell), \quad \ell = 1, \dots, L, \quad \text{and} \quad w(K_1 \cup \dots \cup K_L) = \mathbb{E} \left(\max_{\ell=1, \dots, L} \xi_\ell \right).$$

By the concentration of measure inequality (see e.g. [FR13, Theorem 8.40]) applied to the function $F : \mathbf{x} \in \mathbb{R}^n \mapsto \sup_{\mathbf{f} \in K_\ell} \langle \mathbf{f}, \mathbf{x} \rangle$, which is a Lipschitz function with constant 1, each ξ_ℓ satisfies

$$\mathbb{P}(\xi_\ell \geq \mathbb{E}(\xi_\ell) + t) \leq \exp(-t^2/2).$$

Because each $\mathbb{E}(\xi_\ell)$ is no larger than $\max_\ell \mathbb{E}(\xi_\ell) = \max_\ell w(K_\ell) =: \omega$, we also have

$$\mathbb{P}(\xi_\ell \geq \omega + t) \leq \exp(-t^2/2).$$

Setting $v := \sqrt{2 \ln(L)}$, we now calculate

$$\begin{aligned} \mathbb{E} \left(\max_{\ell=1, \dots, L} \xi_\ell \right) &= \int_0^\infty \mathbb{P} \left(\max_{\ell=1, \dots, L} \xi_\ell \geq u \right) du = \left(\int_0^{\omega+v} + \int_{\omega+v}^\infty \right) \mathbb{P} \left(\max_{\ell=1, \dots, L} \xi_\ell \geq u \right) du \\ &\leq \int_0^{\omega+v} 1 du + \int_{\omega+v}^\infty \sum_{\ell=1}^L \mathbb{P}(\xi_\ell \geq u) du = \omega + v + \sum_{\ell=1}^L \int_v^\infty \mathbb{P}(\xi_\ell \geq \omega + t) dt \\ &\leq \omega + v + L \int_v^\infty \exp(-t^2/2) dt \leq \omega + v + L \frac{\exp(-v^2/2)}{v} \\ &= \omega + \sqrt{2 \ln(L)} + L \frac{1/L}{\sqrt{2 \ln(L)}} \leq \omega + c \sqrt{\ln(L)}, \end{aligned}$$

where $c = \sqrt{2} + (\sqrt{2} \ln(2))^{-1} \leq 3$. We have shown that $w(K_1 \cup \dots \cup K_L) \leq \max_\ell w(K_\ell) + 3\sqrt{\ln(L)}$, as desired. \square

⁴strictly speaking, [PV13b, Proposition 2.1] applies to the slightly different notion of mean width defined as $\mathbb{E}[\sup_{\mathbf{f} \in K-K} \langle \mathbf{f}, \mathbf{g} \rangle]$

We now turn our attention to proving the awaited theorem.

Proof of Theorem 3. According to [PV13b, Proposition 4.3], with $\mathbf{A}' := (\sqrt{2/\pi}/m)\mathbf{A}$, we have

$$|\langle \mathbf{A}'\mathbf{f}, \text{sgn}(\mathbf{A}'\mathbf{g}) \rangle - \langle \mathbf{f}, \mathbf{g} \rangle| \leq \delta,$$

for all $\mathbf{f}, \mathbf{g} \in \mathbf{D}(\Sigma_s^N) \cap S^{n-1}$ provided $m \geq C\delta^{-7}w(\mathbf{D}(\Sigma_s^N) \cap S^{n-1})^2$, so it is enough to upper-bound $w(\mathbf{D}(\Sigma_s^N) \cap S^{n-1})$ appropriately. To do so, with Σ_S^N denoting the space $\{\mathbf{x} \in \mathbb{R}^N : \text{supp}(\mathbf{x}) \subseteq S\}$ for any $S \subseteq \{1, \dots, N\}$, we use Lemma 12 to write

$$\begin{aligned} w(\mathbf{D}(\Sigma_s^N) \cap S^{n-1}) &= w\left(\bigcup_{|S|=s} \{\mathbf{D}(\Sigma_S^N) \cap S^{n-1}\}\right) \stackrel{(ii)}{\leq} \max_{|S|=s} w(\mathbf{D}(\Sigma_S^N) \cap S^{n-1}) + 3\sqrt{\ln\left(\binom{N}{s}\right)} \\ &\stackrel{(i)}{\leq} \sqrt{s} + 3\sqrt{s \ln(eN/s)} \leq 4\sqrt{s \ln(eN/s)}. \end{aligned}$$

The result is now immediate. \square

5.2 Proof of TES

We propose two approaches for proving Theorem 5. One uses again the notion of Gaussian width, the other one relies on covering numbers. The necessary results are isolated in the following lemma.

Lemma 13. The set of ℓ_2 -normalized effectively s -analysis-sparse signals satisfies

- (i) $w\left((\mathbf{D}^*)^{-1}(\Sigma_s^{N,\text{eff}}) \cap S^{n-1}\right) \leq C\sqrt{s \ln(eN/s)}$,
- (ii) $\mathcal{N}\left((\mathbf{D}^*)^{-1}(\Sigma_s^{N,\text{eff}}) \cap S^{n-1}, \rho\right) \leq \binom{N}{t} \left(1 + \frac{8}{\rho}\right)^t$, $t := \lceil 4\rho^{-2}s \rceil$.

Proof. (i) By the definition of the Gaussian width for $\mathcal{K}_s := (\mathbf{D}^*)^{-1}(\Sigma_s^{N,\text{eff}}) \cap S^{n-1}$, with $\mathbf{g} \in \mathbb{R}^n$ denoting a standard normal random vector,

$$(24) \quad w(\mathcal{K}_s) = \mathbb{E} \left[\sup_{\substack{\mathbf{D}^*\mathbf{f} \in \Sigma_s^{N,\text{eff}} \\ \|\mathbf{f}\|_2=1}} \langle \mathbf{f}, \mathbf{g} \rangle \right] = \mathbb{E} \left[\sup_{\substack{\mathbf{D}^*\mathbf{f} \in \Sigma_s^{N,\text{eff}} \\ \|\mathbf{D}^*\mathbf{f}\|_2=1}} \langle \mathbf{D}\mathbf{D}^*\mathbf{f}, \mathbf{g} \rangle \right] \leq \mathbb{E} \left[\sup_{\substack{\mathbf{x} \in \Sigma_s^{N,\text{eff}} \\ \|\mathbf{x}\|_2=1}} \langle \mathbf{D}\mathbf{x}, \mathbf{g} \rangle \right].$$

In view of $\|\mathbf{D}\|_{2 \rightarrow 2} = 1$, we have, for any $\mathbf{x}, \mathbf{x}' \in \Sigma_s^{N,\text{eff}}$ with $\|\mathbf{x}\|_2 = \|\mathbf{x}'\|_2 = 1$,

$$\begin{aligned} \mathbb{E}(\langle \mathbf{D}\mathbf{x}, \mathbf{g} \rangle - \langle \mathbf{D}\mathbf{x}', \mathbf{g}' \rangle)^2 &= \mathbb{E}[\langle \mathbf{D}\mathbf{x}, \mathbf{g} \rangle^2] + \mathbb{E}[\langle \mathbf{D}\mathbf{x}', \mathbf{g}' \rangle^2] = \|\mathbf{D}\mathbf{x}\|_2^2 + \|\mathbf{D}\mathbf{x}'\|_2^2 \leq \|\mathbf{x}\|_2^2 + \|\mathbf{x}'\|_2^2 \\ &= \mathbb{E}(\langle \mathbf{x}, \mathbf{g} \rangle - \langle \mathbf{x}', \mathbf{g}' \rangle)^2. \end{aligned}$$

Applying Slepian's lemma (see e.g. [FR13, Lemma 8.25]), we obtain

$$w(\mathcal{K}_s) \leq \mathbb{E} \left[\sup_{\substack{\mathbf{x} \in \Sigma_s^{N, \text{eff}} \\ \|\mathbf{x}\|_2=1}} \langle \mathbf{x}, \mathbf{g} \rangle \right] = w(\Sigma_s^{N, \text{eff}} \cap S^{n-1}).$$

The latter is known to be bounded by $Cs \ln(eN/s)$, see [PV13b, Lemma 2.3].

(ii) The covering number $\mathcal{N}(\mathcal{K}_s, \rho)$ is bounded above by the maximal number $\mathcal{P}(\mathcal{K}_s, \rho)$ of elements in \mathcal{K}_s that are separated by a distance ρ . We claim that $\mathcal{P}(\mathcal{K}_s, \rho) \leq \mathcal{P}(\Sigma_t^N \cap B_2^N, \rho/2)$. To justify this claim, let us consider a maximal ρ -separated set $\{\mathbf{f}^1, \dots, \mathbf{f}^L\}$ of signals in \mathcal{K}_s . For each i , let $T_i \subseteq \{1, \dots, N\}$ denote an index set of t largest absolute entries of $\mathbf{D}^* \mathbf{f}^i$. We write

$$\rho < \|\mathbf{f}^i - \mathbf{f}^j\|_2 = \|\mathbf{D}^* \mathbf{f}^i - \mathbf{D}^* \mathbf{f}^j\|_2 \leq \|(\mathbf{D}^* \mathbf{f}^i)_{T_i} - (\mathbf{D}^* \mathbf{f}^j)_{T_j}\|_2 + \|(\mathbf{D}^* \mathbf{f}^i)_{\overline{T_i}}\|_2 + \|(\mathbf{D}^* \mathbf{f}^j)_{\overline{T_j}}\|_2.$$

Invoking [FR13, Theorem 2.5], we observe that

$$\|(\mathbf{D}^* \mathbf{f}^i)_{\overline{T_i}}\|_2 \leq \frac{1}{2\sqrt{t}} \|\mathbf{D}^* \mathbf{f}^i\|_1 \leq \frac{\sqrt{s}}{2\sqrt{t}} \|\mathbf{D}^* \mathbf{f}^i\|_2 = \frac{\sqrt{s}}{2\sqrt{t}},$$

and similarly for j instead of i . Thus, we obtain

$$\rho < \|(\mathbf{D}^* \mathbf{f}^i)_{T_i} - (\mathbf{D}^* \mathbf{f}^j)_{T_j}\|_2 + \sqrt{\frac{s}{t}} \leq \|(\mathbf{D}^* \mathbf{f}^i)_{T_i} - (\mathbf{D}^* \mathbf{f}^j)_{T_j}\|_2 + \frac{\rho}{2}, \quad \text{i.e., } \|(\mathbf{D}^* \mathbf{f}^i)_{T_i} - (\mathbf{D}^* \mathbf{f}^j)_{T_j}\|_2 > \frac{\rho}{2}.$$

Since we have uncovered a set of $L = \mathcal{P}(\mathcal{K}_s, \rho)$ points in $\Sigma_t^N \cap B_2^N$ that are $(\rho/2)$ -separated, the claimed inequality is proved. We conclude by recalling that $\mathcal{P}(\Sigma_t^N \cap B_2^N, \rho/2)$ is bounded above by $\mathcal{N}(\Sigma_t^N \cap B_2^N, \rho/4)$, which is itself bounded above by $\binom{N}{t} \left(1 + \frac{2}{\rho/4}\right)^t$. \square

We can now turn our attention to proving the awaited theorem.

Proof of Theorem 5. With $\mathcal{K}_s = (\mathbf{D}^*)^{-1}(\Sigma_s^{N, \text{eff}}) \cap S^{n-1}$, the conclusion holds when $m \geq C\varepsilon^{-6} w(\mathcal{K}_s)^2$ or when $m \geq C\varepsilon^{-1} \ln(\mathcal{N}(\mathcal{K}_s, c\varepsilon))$, according to [PV14, Theorem 1.5] or to [BL15, Theorem 1.5], respectively. It now suffices to call upon Lemma 13. Note that the latter option yields better powers of ε^{-1} but less pleasant failure probability. \square

5.3 Further remarks

We conclude this theoretical section by making two noteworthy comments on the sign product embedding property and the tessellation property in the dictionary case.

Remark 4. **D-SPEP** cannot hold for arbitrary dictionary \mathbf{D} if synthesis sparsity was replaced by effective synthesis sparsity. This is because the set of effectively s -synthesis-sparse signals can be the whole space \mathbb{R}^n . Indeed, let $\mathbf{f} \in \mathbb{R}^n$ that be written as $\mathbf{f} = \mathbf{D}\mathbf{u}$ for some $\mathbf{u} \in \mathbb{R}^N$. Let also pick an $(s-1)$ -sparse vector $\mathbf{v} \in \ker \mathbf{D}$ — there are tight frame for which this is possible, e.g. the concatenation of two orthogonal matrices. For $\varepsilon > 0$ small enough, we have

$$\frac{\|\mathbf{v} + \varepsilon\mathbf{u}\|_1}{\|\mathbf{v} + \varepsilon\mathbf{u}\|_2} \leq \frac{\|\mathbf{v}\|_1 + \varepsilon\|\mathbf{u}\|_1}{\|\mathbf{v}\|_2 - \varepsilon\|\mathbf{u}\|_2} \leq \frac{\sqrt{s-1}\|\mathbf{v}\|_2 + \varepsilon\|\mathbf{u}\|_1}{\|\mathbf{v}\|_2 - \varepsilon\|\mathbf{u}\|_2} \leq \sqrt{s},$$

so that the coefficient vector $\mathbf{v} + \varepsilon\mathbf{u}$ is effectively s -sparse, hence so is $(1/\varepsilon)\mathbf{v} + \mathbf{u}$. It follows that $\mathbf{f} = \mathbf{D}((1/\varepsilon)\mathbf{v} + \mathbf{u})$ is effectively s -synthesis sparse.

Remark 5. Theorem 3 easily implies a tessellation result for $\mathbf{D}(\Sigma_s^N) \cap S^{n-1}$, the “synthesis sparse sphere”. Precisely, under the assumptions of the theorem (with a change of the constant C), **D-SPEP**($2s, \delta/2$) holds. Then, one can derive

$$[\mathbf{g}, \mathbf{h} \in \mathbf{D}(\Sigma_s) \cap S^{n-1} : \text{sgn}(\mathbf{A}\mathbf{g}) = \text{sgn}(\mathbf{A}\mathbf{h})] \implies [\|\mathbf{g} - \mathbf{h}\|_2 \leq \delta].$$

To see this, with $\varepsilon := \text{sgn}(\mathbf{A}\mathbf{g}) = \text{sgn}(\mathbf{A}\mathbf{h})$ and with $\mathbf{f} := (\mathbf{g} - \mathbf{h})/\|\mathbf{g} - \mathbf{h}\|_2 \in \mathbf{D}(\Sigma_{2s}) \cap S^{n-1}$, we have

$$\left| \frac{\sqrt{2/\pi}}{m} \langle \mathbf{A}\mathbf{f}, \varepsilon \rangle - \langle \mathbf{f}, \mathbf{g} \rangle \right| \leq \frac{\delta}{2}, \quad \left| \frac{\sqrt{2/\pi}}{m} \langle \mathbf{A}\mathbf{f}, \varepsilon \rangle - \langle \mathbf{f}, \mathbf{h} \rangle \right| \leq \frac{\delta}{2},$$

so by the triangle inequality $|\langle \mathbf{f}, \mathbf{g} - \mathbf{h} \rangle| \leq \delta$, i.e., $\|\mathbf{g} - \mathbf{h}\|_2 \leq \delta$, as announced.

Acknowledgment

The authors would like to thank the AIM SQuaRE program that funded and hosted our initial collaboration. In addition, Baraniuk is partially supported by NSF grant number CCF-1527501, ARO grant number W911NF-15-1-0316, AFOSR grant number FA9550-14-1-0088, Needell by an Alfred P. Sloan Fellowship and NSF Career grant number 1348721, Plan by NSERC grant number 22R23068, and Wootters by NSF Postdoctoral Research Fellowship grant number 1400558.

References

- [BB08] P. T. Boufounos and R. G. Baraniuk. 1-bit compressive sensing. In *Proceedings of the 42nd Annual Conference on Information Sciences and Systems (CISS)*, pages 16–21. IEEE, 2008.

- [BFN⁺14] R. G. Baraniuk, S. Foucart, D. Needell, Y. Plan, and M. Wootters. Exponential decay of reconstruction error from binary measurements of sparse signals. *arXiv preprint arXiv:1407.8246*, 2014.
- [BL15] D. Bilyk and M. T. Lacey. Random tessellations, restricted isometric embeddings, and one bit sensing. *arXiv preprint arXiv:1512.06697*, 2015.
- [Blu11] T. Blumensath. Sampling and reconstructing signals from a union of linear subspaces. *IEEE T. Inform. Theory*, 57(7):4660–4671, 2011.
- [CD04] E. J. Candès and D. L. Donoho. New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities. *Comm. Pure Appl. Math.*, 57(2):219–266, 2004.
- [CDDY00] E. J. Candès, L. Demanet, D. L. Donoho, and L. Ying. Fast discrete curvelet transforms. *Multiscale Model. Simul.*, 5:861–899, 2000.
- [CENR10] E. J. Candès, Y. C. Eldar, D. Needell, and P. Randall. Compressed sensing with coherent and redundant dictionaries. *Appl. Comp. Harm. Anal.*, 31(1):59–73, 2010.
- [Dau92] I. Daubechies. *Ten lectures on wavelets*. SIAM, Philadelphia, PA, 1992.
- [DNW12] M. Davenport, D. Needell, and M. B. Wakin. Signal space CoSaMP for sparse recovery with redundant dictionaries. *IEEE T. Inform. Theory*, 59(10):6820–6829, 2012.
- [DSP] Compressive sensing webpage. <http://dsp.rice.edu/cs>.
- [EK12] Y. C. Eldar and G. Kutyniok. *Compressed sensing: theory and applications*. Cambridge University Press, 2012.
- [EMR07] M. Elad, P. Milanfar, and R. Rubinstein. Analysis versus synthesis in signal priors. *Inverse problems*, 23(3):947, 2007.
- [Fou16] S. Foucart. Dictionary-sparse recovery via thresholding-based algorithms. *Journal of Fourier Analysis and Applications*, 22(1):6–19, 2016.
- [FR13] S. Foucart and H. Rauhut. *A mathematical introduction to compressive sensing*. Birkhäuser, 2013.
- [FS98] H. Feichtinger and T. Strohmer, editors. *Gabor Analysis and Algorithms*. Birkhäuser, 1998.
- [GNE⁺14] R. Giryes, S. Nam, M. Elad, R. Gribonval, and M. E. Davies. Greedy-like algorithms for the cospase analysis model. *Linear Algebra Appl.*, 441:22–60, 2014.
- [GNJN13] S. Gopi, P. Netrapalli, P. Jain, and A. Nori. One-bit compressed sensing: Provable support and vector recovery. In *Proceedings of the 30th International Conference on Machine Learning (ICML)*, pages 154–162, 2013.

- [JDDV13] L. Jacques, K. Degraux, and C. De Vleeschouwer. Quantized iterative hard thresholding: Bridging 1-bit and high-resolution quantized compressed sensing. In *Proceedings of the 10th International Conference on Sampling Theory and Applications (SampTA)*, pages 105–108, 2013.
- [JLBB13] L. Jacques, J. N. Laska, P. T. Boufounos, and R. G. Baraniuk. Robust 1-bit compressive sensing via binary stable embeddings of sparse vectors. *IEEE T. Inform. Theory*, 59(4):2082–2102, April 2013.
- [KNW15] F. Krahmer, D. Needell, and R. Ward. Compressive sensing with redundant dictionaries and structured measurements. *SIAM Journal on Mathematical Analysis*, 47(6):4606–4629, 2015.
- [KSW14] K. Knudson, R. Saab, and R. Ward. One-bit compressive sensing with norm estimation. *arXiv preprint arXiv:1404.6853*, 2014.
- [NDEG13] S. Nam, M. E. Davies, M. Elad, and R. Gribonval. The cospase analysis model and algorithms. *Appl. Comp. Harm. Anal.*, 34(1):30–56, 2013.
- [PE13] T. Peleg and M. Elad. Performance guarantees of the thresholding algorithm for the cospase analysis model. *IEEE T. Inform. Theory*, 59(3):1832–1845, 2013.
- [PV13a] Y. Plan and R. Vershynin. One-bit compressed sensing by linear programming. *Comm. Pure Appl. Math.*, 66(8):1275–1297, 2013.
- [PV13b] Y. Plan and R. Vershynin. Robust 1-bit compressed sensing and sparse logistic regression: A convex programming approach. *IEEE T. Inform. Theory*, 59(1):482–494, 2013.
- [PV14] Y. Plan and R. Vershynin. Dimension reduction by random hyperplane tessellations. *Discrete & Computational Geometry*, 51(2):438–461, 2014.
- [RSV08] H. Rauhut, K. Schnass, and P. Vandergheynst. Compressed sensing and redundant dictionaries. *IEEE T. Inform. Theory*, 54(5):2210–2219, 2008.
- [SED04] J.-L. Starck, M. Elad, and D. L. Donoho. Redundant multiscale transforms and their application for morphological component analysis. *Adv. Imag. Elect. Phys.*, 132, 2004.
- [SWY15] R. Saab, R. Wang, and O. Yilmaz. Quantization of compressive samples with stable and robust recovery. *arXiv preprint arXiv:1504.00087*, 2015.
- [YYO12] M. Yan, Y. Yang, and S. Osher. Robust 1-bit compressive sensing using adaptive outlier pursuit. *IEEE Transactions on Signal Processing*, 60(7):3868–3875, 2012.