

# The importance of phase in complex compressive sensing

Laurent Jacques\* and Thomas Feuillen\*

January 17, 2020

## Abstract

We consider the question of estimating a real low-complexity signal (such as a sparse vector or a low-rank matrix) from the phase of complex random measurements. We show that in this *phase-only compressive sensing* (PO-CS) scenario, we can perfectly recover such a signal with high probability and up to global unknown amplitude if the sensing matrix is a complex Gaussian random matrix and if the number of measurements is large compared to the complexity level of the signal space. Our approach proceeds by recasting the (non-linear) PO-CS scheme as a linear compressive sensing model built from a signal normalization constraint, and a phase-consistency constraint imposing any signal estimate to match the observed phases in the measurement domain. Practically, stable and robust signal direction estimation is achieved from any *instance optimal* algorithm of the compressive sensing literature (such as basis pursuit denoising). This is ensured by proving that the matrix associated with this equivalent linear model satisfies with high probability the restricted isometry property under the above condition on the number of measurements. We finally observe experimentally that robust signal direction recovery is reached at about twice the number of measurements needed for signal recovery in compressive sensing.

## 1 Introduction

About 40 years ago, Oppenheim and his collaborators [1, 2] determined that, under certain conditions and up to a global unknown amplitude, a band-limited signal can be exactly recovered from the phase of its Fourier transform; one can reconstruct the *direction* of this signal. The authors also provided a practical iterative algorithm for this purpose. Since the observed signal lies at the intersection of the set of band-limited signals and a specific phase-consistency set — the set of signals matching, in the frequency space, the phase of the observed signal — this algorithm is designed to find this intersection, when it is unique, by alternate projections onto these two convex sets (the POCS method).

---

\*LJ and TF are with the ICTEAM institute in UCLouvain, Belgium. E-mail: {laurent.jacques, thomas.feuillen}@uclouvain.be. LJ is funded by the Belgian F.R.S.-FNRS.

More recently, Boufounos [3] considered a similar question in the context of complex compressive sensing where measurements of a sparse signal are obtained from its multiplication with a complex, fat sensing matrix. Since keeping the phase of such complex measurements naturally extends the formalism of one-bit compressive sensing (which keeps only the sign of real random projections) [4, 5, 6], the author demonstrated that, in the case of complex Gaussian random sensing matrices, one can estimate the direction of such a signal from the phase of the compressive measurements (or *phase-only* compressive sensing – PO-CS). The estimation error then provably decays when the number of measurements increases. Boufounos actually proved that, with high probability (w.h.p.) and up to a controlled distortion, the discrepancy between the phase of complex Gaussian random projections of two sparse vectors encodes their Euclidean distance. Recently, pursuing the correspondence with one-bit CS and extending its central  $(\ell_1, \ell_2)$ -restricted isometry property ( $(\ell_1, \ell_2)$ -RIP) [7] to complex sensing matrices, Feuillen and co-authors established that a single-step procedure called projected back-projection (PBP) — an estimation of the signal direction by hard-thresholding the multiplication of the PO-CS measurements by the adjoint sensing matrix [8] — also yields small and decaying reconstruction error. Of interest for our work, Boufounos also designed in [4] a greedy algorithm delivering an estimate that is both sparse and phase-consistent with the unknown signal; this signal and the estimate share identical phases in the random projection domain. However, none of the theoretical approaches above can explain why this specific algorithm succeeds numerically in perfectly estimating the observed signal direction.

With this work, we bring two main contributions to this context. First, we demonstrate that, in a noiseless scenario, perfect estimation of a signal direction from the phase of its random, complex Gaussian measurements is possible, not only for sparse signals but for any signals belonging to a symmetric, low-complexity conic set (or *cone*) of reduced dimensionality, including (union of) subspaces, the set of sparse signals, the set of model-based or group sparse vectors [9, 10] or the set of low-rank matrices [11]. This result strikingly differs from known reconstruction guarantees in the context of (real) one-bit CS [4, 5, 6]; in this case, the direction of a low-complexity signal can only be estimated up to a lower-bounded error [5].

Second, we show that one can reconstruct the direction of a low-complexity signal from any *instance optimal* algorithms of the CS literature whose error guarantees are controlled by the restricted isometry property of the sensing matrix [12, 13, 14]. Using such algorithms, we can bound the reconstruction error of the direction of a low-complexity signal observed in a PO-CS model. This error bound is *(i) non-uniform*, in the sense that the estimation is possible, w.h.p., given the observed low-complexity signal, and *(ii) stable and robust* as the instance optimality of the selected algorithm allows for both a bounded noise on the observed measurement phases and an error on the modeling of the signal by an element of a low-complexity set. Experimentally, we observe that the number of measurements required for robust estimation of signal direction from noiseless phase-only observation is about twice the one needed for signal estimation in the case of (linear) CS.

Our approach is thus complementary to [8]; in this work, provided that the sensing matrix respects the  $(\ell_1, \ell_2)$ -RIP, the PBP algorithm yields a bounded reconstruction error for the estimation of the direction of any unknown signal in (noisy) PO-CS. The bound thus holds *uniformly*. However, the observed signal must be sparse and the reconstruction error does not vanish in absence of measurement noise.

Regarding the mathematical tools, all our developments rely on recasting of the PO-CS recovery problem as a linear compressive sensing problem. This one is built from a signal normalization constraint — as allowed from the independence of the PO-CS model to the signal amplitude — and a phase-consistency constraint imposing to match the observed measurement phases. We show that in noiseless PO-CS, provided that the number of measurements is large compared to the signal set complexity, the sensing matrix associated with this equivalent formulation respects the restricted isometry property over the signal set; the distance between two low-complexity vectors is encoded in the one of their projections by this matrix (see Sec. 2). Moreover, when a bounded noise corrupts the PO-CS model, this property is preserved if the noise variations are small. All our proofs are self-contained and mostly rely on extending to the complex field previously introduced random embeddings, such as the (local) sign-product embedding [6, 15, 7].

The rest of this paper is structured as follows. We first end this introduction by defining some useful notations and conventions. In Sec. 2, we recall a few fundamental results of the compressive sensing theory that are subsequently used in this work. Sec. 3 defines the PO-CS problem, its reformulation as a signal recovery problem observed by a linear model and provides the first guarantees on the exact estimation of a low-complexity signal in the case of noiseless PO-CS. Sec. 4 extends these guarantees to the case where the PO-CS model is corrupted by a bounded noise and the unknown signal is well represented in a low-complexity domain with a bounded modeling error. For clarity, the proofs of our main results are inserted in Sec. 5. In Sec. 6, we provide several experiments leveraging the instance optimality of the basis pursuit denoising algorithm. We first establish from which number of measurements one can perfectly estimate the direction of a sparse signal in PO-CS with complex Gaussian random sensing matrices. We then measure the accuracy of such an estimate under noise corruption for various noise levels. We conclude this paper in Sec. 7 with a few open questions and perspectives.

**Notations and conventions:** Hereafter, the symbols  $C, c > 0$  represent absolute constants whose exact value may change from one instance to the other. We denote matrices and vectors with bold symbols, *e.g.*,  $\Phi \in \mathbb{C}^{m \times n}$ ,  $\mathbf{x} \in \mathbb{C}^n$ , and scalar values with light symbols. We will often use the following quantities:  $[d] := \{1, \dots, d\}$  for  $d \in \mathbb{N}$ ;  $\mathbf{i} = \sqrt{-1}$ ;  $\Re\{\lambda\}$  (or  $\lambda^{\Re}$ ) and  $\Im\{\lambda\}$  (or  $\lambda^{\Im}$ ) are the real and imaginary part of  $\lambda \in \mathbb{C}$ , respectively, and  $\lambda^*$  is its complex conjugate;  $\mathbf{1}_d := (1, \dots, 1)^\top \in \mathbb{R}^d$  and  $\mathbf{I}_d$  is the  $d \times d$  identity matrix;  $\mathbf{A}^*$  is the adjoint (conjugate transpose) of  $\mathbf{A}$ ;  $\text{supp } \mathbf{x} = \{i : x_i \neq 0\}$ ;  $|\mathcal{S}|$  is the cardinality of a finite set  $\mathcal{S}$ ;  $\mathbf{A}_\Omega$  is matrix formed by restricting the  $d$  columns of  $\mathbf{A}$  to those indexed in  $\Omega \subset [d]$ ;  $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^d x_i^* y_i$  is the scalar product of  $\mathbf{x}, \mathbf{y} \in \mathbb{C}^d$ ;  $\|\mathbf{x}\|_p = (\sum_{i=1}^d |x_i|^p)^{1/p}$  is the  $\ell_p$ -norm of  $\mathbf{x}$  ( $p \geq 1$ ), with  $\|\mathbf{x}\|_\infty = \max_i |x_i|$  and  $\|\mathbf{x}\|_0 := |\text{supp}(\mathbf{x})|$ ;  $\mathcal{K} - \mathcal{K}' := \{\mathbf{u} - \mathbf{v} : \mathbf{u} \in \mathcal{K}, \mathbf{v} \in \mathcal{K}'\}$  and  $\mathbb{R}\mathbf{x}$  denote the Minkowski difference of  $\mathcal{K}, \mathcal{K}' \subset \mathbb{R}^d$  and the span of  $\mathbf{x} \in \mathbb{R}^d$ , respectively. The symbol “ $\sim$ ” is used to describe either the distribution law of a random quantity (*e.g.*,  $X \sim \mathcal{N}(0, 1)$ ), or the equivalence of distribution between two random quantities (*e.g.*,  $Y \sim X$ ). Given a (scalar) random distribution  $\mathcal{D}$ , we denote by  $\mathcal{D}^{m \times n}$  the  $m \times n$  random matrix distribution generating matrices with entries independently and identically distributed (i.i.d.) as  $\mathcal{D}$ , and we omit the superscript  $n$  for vectors (for which  $n = 1$ ); for instance, the  $m \times n$  real and complex Gaussian random matrices distributed as  $\mathcal{N}^{m \times n}(\mu', \sigma^2)$  and  $\mathcal{N}_{\mathbb{C}}^{m \times n}(\mu, 2\sigma^2) \sim \mathcal{N}^{m \times n}(\mu^{\Re}, \sigma^2) + \mathbf{i}\mathcal{N}^{m \times n}(\mu^{\Im}, \sigma^2)$ , respec-

tively, for some means  $\mu, \mu'$  and variance  $\sigma^2$ . An  $s$ -sparse  $\mathbf{x}$  vector belongs to the set  $\Sigma_s^n := \{\mathbf{u} \in \mathbb{R}^n, \|\mathbf{u}\|_0 \leq s\}$ .

## 2 Preliminaries

Let us first provide the key principles of CS theory. In the complex field, the compressive sensing of a signal  $\mathbf{x} \in \mathbb{R}^n$  with a complex sensing matrix  $\mathbf{A} = \mathbf{A}^{\Re} + i\mathbf{A}^{\Im} \in \mathbb{C}^{m \times n}$  corresponds to acquiring  $m$  noisy complex measurements through the model [13]

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \boldsymbol{\epsilon}, \quad (1)$$

for some complex measurement noise  $\boldsymbol{\epsilon} \in \mathbb{C}^m$ . In this model, we assume  $\mathbf{x}$  real as we can always recast the sensing of  $\mathbf{x} \in \mathbb{C}^n$  as the one of  $(\Re(\mathbf{x})^\top, \Im(\mathbf{x})^\top)^\top \in \mathbb{R}^{2n}$  with the sensing matrix  $(\mathbf{A}^{\Re}, -\mathbf{A}^{\Im}) + i(\mathbf{A}^{\Im}, \mathbf{A}^{\Re}) \in \mathbb{C}^{m \times 2n}$ .

The signal  $\mathbf{x}$  can be recovered from the measurement vector  $\mathbf{y}$ , if  $\mathbf{A}$  is “well conditioned” with respect the sparse signal set  $\Sigma_s^n$  [12, 13]. This happens for instance if, for some  $0 < \delta < 1$ ,  $\mathbf{A}$  satisfies the restricted isometry property of order  $2s$ , or  $\text{RIP}(\Sigma_{2s}^n, \delta)$ , defined by

$$(1 - \delta)\|\mathbf{u}\|^2 \leq \|\mathbf{A}\mathbf{u}\|^2 \leq (1 + \delta)\|\mathbf{u}\|^2, \quad \forall \mathbf{u} \in \Sigma_{2s}^n. \quad (2)$$

For instance, this property holds with high probability if  $\mathbf{A}$  is a real (or complex [13])  $m \times n$  random Gaussian matrix, and if  $m \geq C\delta^{-2}s \log(n/\delta s)$  [16].

If the RIP is verified over  $\Sigma_{2s}^n$  with  $\delta < 1/\sqrt{2}$  [17] (see also [7, Thm 6]), then, for some  $C, D > 0$  and  $\varepsilon \geq \|\boldsymbol{\epsilon}\|$ , the *basis pursuit denoising* [18] estimate  $\hat{\mathbf{x}} = \Delta_{\Sigma_s^n}(\mathbf{y}, \mathbf{A}; \varepsilon)$  with

$$\Delta_{\Sigma_s^n}(\mathbf{y}, \mathbf{A}; \varepsilon) := \arg \min_{\mathbf{u}} \|\mathbf{u}\|_1 \quad \text{s.t.} \quad \|\mathbf{A}\mathbf{u} - \mathbf{y}\| \leq \varepsilon, \quad (\text{BPDN})$$

satisfies

$$\|\mathbf{x} - \hat{\mathbf{x}}\| \leq Cs^{-1}\|\mathbf{x} - \mathbf{x}_s\|_1 + D\varepsilon, \quad (3)$$

with  $\mathbf{x}_s$  the closest  $s$ -sparse signal to  $\mathbf{x}$  (that is, its best  $s$ -sparse approximation). Consequently, if  $\mathbf{x} \in \Sigma_s^n$ ,  $\boldsymbol{\epsilon} = \mathbf{0}$ , and  $\varepsilon = 0$ , we get perfect signal recovery ( $\hat{\mathbf{x}} = \mathbf{x}$ ). Similar robustness and stability results can be achieved for a series of other algorithms — such as orthogonal matching pursuit (OMP), compressive sampling matching pursuit (CoSaMP), or iterative hard thresholding (IHT) — as soon as  $\mathbf{A}$  respects the RIP for sparsity levels equal to a few multiples of  $s$  and for small enough  $\delta$  [19, 20, 7, 21]

In this paper, we are interested in more general low-complexity signal spaces than the set of sparse vectors. We open our study, for instance, to (union of) subspaces of  $\mathbb{R}^n$ , such as band-limited or sparse signals, or signals displaying more complex sparsity patterns such as model-based or group sparsity [10, 9], and to the set  $\Lambda_r^n$  of  $\sqrt{n} \times \sqrt{n}$  rank- $r$  matrices (for  $n$  a squared integer). We thus focus on *symmetric cones*  $\mathcal{K}$  (for which  $-\mathcal{K} = \mathcal{K}$  and  $\lambda\mathcal{K} \subset \mathcal{K}$  for all  $\lambda > 0$ ) that occupy a “small volume” of  $\mathbb{R}^n$ . This can be measured by the (localized) Gaussian mean width  $w(\mathcal{K} \cap \mathbb{S}^{n-1})$  of  $\mathcal{K}$ , with

$$w(\mathcal{K}') := \mathbb{E} \sup_{\mathbf{u} \in \mathcal{K}'} \langle \mathbf{g}, \mathbf{u} \rangle, \quad \text{for } \mathcal{K}' \subset \mathbb{R}^n \text{ and } \mathbf{g} \sim \mathcal{N}^n(0, 1). \quad (4)$$

This quantity encodes the intrinsic dimension of  $\mathcal{K}$ ; for instance,  $w^2(\mathcal{L} \cap \mathbb{S}^{n-1}) \leq CL$  for a subspace  $\mathcal{L} \subset \mathbb{R}^n$  of dimension  $L$  (such as the space of band-limited signals of  $\mathbb{R}^n$  whose

discrete cosine transform (DCT) is zero over the last  $n - L$  frequencies),  $w^2(\Sigma_s^n \cap \mathbb{S}^{n-1}) \leq Cs \log(n/s)$  — that is the quantity driving the number of measurements for sparse signal recovery in CS —, and  $w(\Lambda_r^n \cap \mathbb{S}^{n-1}) \leq Cr\sqrt{n}$  [22, 6, 23] (see also [24, Table 1] for other examples of low-complexity sets).

It has now been established that, for a variety of low-complexity, symmetric cones  $\mathcal{K}$ , a signal  $\mathbf{x} \in \mathcal{K}$  can be recovered from the noisy measurement vector  $\mathbf{y}$  in (1). For each of these sets, there exists an algorithm  $\Delta_{\mathcal{K}} : \mathbb{C}^m \times \mathbb{C}^{m \times n} \times \mathbb{R}_+ \rightarrow \mathbb{R}^n$  that provides an *instance optimal* estimate  $\hat{\mathbf{x}} = \Delta_{\mathcal{K}}(\mathbf{y}, \mathbf{A}; \varepsilon)$  of  $\mathbf{x}$  verifying, for some  $C, D > 0$ ,

$$\|\mathbf{x} - \hat{\mathbf{x}}\| \leq Ce_0(\mathbf{x}, \mathcal{K}) + D\varepsilon, \quad (5)$$

where  $e_0$ , which vanishes if  $\mathbf{x} \in \mathcal{K}$ , stands for a *modeling error* in the approximation of  $\mathbf{x}$  by an element of  $\mathcal{K}$ . For instance, we have  $e_0(\mathbf{x}, \Sigma_s^n) = \|\mathbf{x} - \mathbf{x}_s\|_1/\sqrt{s}$  or  $e_0(\mathbf{x}, \Sigma_s^n) = \|\mathbf{x} - \mathbf{x}_s\|_2$  [13, 14], and  $e_0(\mathbf{X}, \Lambda_r^n) = \|\mathbf{X} - \mathbf{X}_r\|_*/\sqrt{r}$  with  $\|\cdot\|_*$  the nuclear norm and  $\mathbf{X}_r$  is the best rank- $r$  approximation of  $\mathbf{X} \in \mathbb{R}^{\sqrt{n} \times \sqrt{n}}$  [25].

This instance optimality holds if, for some  $0 < \delta_0 < 1$  associated with  $\Delta_{\mathcal{K}}$  and  $\mathcal{K}$  (for instance,  $\delta_0 = 1/\sqrt{2}$  for BPDN and  $\mathcal{K} = \Sigma_s^n$ ) and  $0 < \delta < \delta_0$ ,  $\mathbf{A}$  respects the general RIP( $\mathcal{K} - \mathcal{K}, \delta$ ) [14] defined by

$$(1 - \delta)\|\mathbf{u}\|^2 \leq \|\mathbf{A}\mathbf{u}\|^2 \leq (1 + \delta)\|\mathbf{u}\|^2, \quad \forall \mathbf{u} \in \mathcal{K} - \mathcal{K}, \quad (6)$$

where  $\mathcal{K} - \mathcal{K}$  is the Minkowski difference of  $\mathcal{K}$  with itself (*e.g.*,  $\mathcal{K} - \mathcal{K} = \Sigma_{2s}^n$  for  $\mathcal{K} = \Sigma_s^n$ ).

We introduce the concept of *recoverable* set to capture these last considerations.

**Definition 2.1.** *A symmetric cone  $\mathcal{K} \subset \mathbb{R}^n$  is called  $(\Delta_{\mathcal{K}}, \delta_0)$ -recoverable if there exists an algorithm  $\Delta_{\mathcal{K}} : \mathbb{C}^m \times \mathbb{C}^{m \times n} \times \mathbb{R}_+ \rightarrow \mathbb{R}^n$  such that, for some  $0 < \delta_0 < 1$ , any matrix  $\mathbf{A} \in \mathbb{C}^{m \times n}$  satisfying the RIP( $\mathcal{K} - \mathcal{K}, \delta$ ) with  $0 < \delta < \delta_0$ , and any noise  $\boldsymbol{\epsilon} \in \mathbb{C}^m$  with  $\|\boldsymbol{\epsilon}\| \leq \varepsilon$ , the vector  $\hat{\mathbf{x}} = \Delta_{\mathcal{K}}(\mathbf{A}\mathbf{x} + \boldsymbol{\epsilon}, \mathbf{A}; \varepsilon) \in \mathbb{R}^n$  estimates any signal  $\mathbf{x} \in \mathbb{R}^n$  with bounded modeling error  $e_0(\mathbf{x}, \mathcal{K})$  according to the instance optimality condition (5).*

To give a few examples of sets  $\mathcal{K}$  and algorithms, we can mention that a mixed norm penalization of BPDN is adapted for signals displaying structured sparsity [10], the recovery of low-rank matrices is ensured either by replacing the  $\ell_1$ -norm of the BPDN program by the nuclear norm [11, 25], or by resorting to specific iterative algorithms such as Singular Value Projection (SVP) [26] or conjugate gradient iterative hard thresholding (CGIHT) [27].

Concerning the existence of matrices satisfying (6), Mendelson and Pajor [28] proved the following crucial result that we will use in our proofs.

**Theorem 2.2** (Adapted from<sup>1</sup> [28, Thm 2.1]). *Given a cone  $\mathcal{K}' \subset \mathbb{R}^n$  and  $\delta > 0$ , if*

$$m \geq C\delta^{-2}w^2(\mathcal{K}' \cap \mathbb{S}^{n-1}), \quad (7)$$

*then the matrix  $\frac{1}{\sqrt{m}}\Phi$  with  $\Phi \sim \mathcal{N}^{m \times n}(0, 1)$  respects the RIP( $\mathcal{K}', \delta$ ) with probability exceeding  $1 - C \exp(-c\delta^2 m)$ .*

<sup>1</sup>Thm 2.1 in [28] is actually valid for any sub-Gaussian random matrices. It is also restricted to subsets of  $\mathbb{S}^{n-1}$ , which can be easily extended to cones as in Thm 2.2.

This result, which extends easily to complex Gaussian random matrices<sup>2</sup>, shows that, w.h.p., we can stably recover all signals of a low-complexity cone  $\mathcal{K}$  from  $m$  Gaussian random measurements provided  $m \geq C\delta^{-2}w^2((\mathcal{K} - \mathcal{K}) \cap \mathbb{S}^{n-1})$ . For similar guarantees in the context of other (structured) sensing matrices (*e.g.*, partial Fourier measurements) and specific low-complexity spaces, the interested reader may consult, for instance, [13] for a comprehensive overview of the literature.

### 3 Phase-only Compressive Sensing

We now depart from the linear CS model and consider the possibility of (partly) recovering a signal  $\mathbf{x} \in \mathbb{R}^n$  from the *phase* of random projections achieved with a complex sensing matrix  $\mathbf{A} \in \mathbb{C}^{m \times n}$ . For clarity, we first consider the case where  $\mathbf{x}$  belongs to a low-complexity set  $\mathcal{K} \subset \mathbb{R}^n$  and the phase measurements are exact. We thus study the non-linear, noiseless phase-only compressive sensing (PO-CS) model

$$\mathbf{z} = \text{sign}_{\mathbb{C}}(\mathbf{A}\mathbf{x}), \quad (8)$$

with  $\text{sign}_{\mathbb{C}} \lambda := \lambda/|\lambda|$  if  $\lambda \in \mathbb{C} \setminus \{0\}$ , and  $\text{sign}_{\mathbb{C}} 0 := 0$ . We develop in the next section a noisy variant of (8), and cover the case where  $\mathbf{x}$  is only known to be close to  $\mathcal{K}$  (which induces a small modeling error  $e_0$  in (5)).

This model encompasses the one considered by Oppenheim in [1] when  $\mathcal{K}$  is the subspace of band-limited signals and  $\mathbf{A}$  is the Fourier sensing matrix. Moreover, since  $\text{sign}_{\mathbb{C}} \lambda = \text{sign} \lambda$  for  $\lambda \in \mathbb{R}$ , PO-CS naturally generalizes one-bit CS to the complex field [4, 3].

We remove the ambiguity raised by the unobserved signal amplitude in (8) by arbitrary specifying

$$\|\mathbf{A}\mathbf{x}\|_1 = \kappa\sqrt{m}, \quad (9)$$

for some  $\kappa > 0$ . This  $\ell_1$ -norm normalization of  $\Phi\mathbf{x}$  is reminiscent of the one imposed in one-bit CS where the signal amplitude is unknown [6]. It can also be understood by the fact that, for  $\|\mathbf{x}\| = 1$ , and  $\mathbf{A} = \Phi/\sqrt{m}$  with  $\Phi \sim \mathcal{N}_{\mathbb{C}}^{m \times n}(0, 2)$ ,  $\mathbb{E}\|\mathbf{A}\mathbf{x}\|_1 = \kappa\sqrt{m}$  with  $\kappa := \sqrt{\pi/2}$  (see the proof of Lemma 5.2). Imposing (9) is thus close to enforcing  $\|\mathbf{x}\| = 1$  for complex Gaussian random matrices.

Following [3], this normalization allows us to recast the PO-CS model as a linear CS model. Indeed, introducing  $\boldsymbol{\alpha}_{\mathbf{z}} := \mathbf{A}^*\mathbf{z}/(\kappa\sqrt{m}) \in \mathbb{C}^m$ , we see that the *phase-consistency* and normalization constraints defined by, respectively,

$$\text{diag}(\mathbf{z})^* \mathbf{A}\mathbf{u} \in \mathbb{R}_+^m, \quad \langle \boldsymbol{\alpha}_{\mathbf{z}}, \mathbf{u} \rangle = 1, \quad (10)$$

are respected for  $\mathbf{u} = \mathbf{x}$ . Since  $\mathbf{x}$  is real, this is equivalent to

$$\begin{cases} \mathbf{H}_{\mathbf{z}}\mathbf{u} = \mathbf{0}, & \langle \boldsymbol{\alpha}_{\mathbf{z}}^{\Re}, \mathbf{u} \rangle = 1, & \langle \boldsymbol{\alpha}_{\mathbf{z}}^{\Im}, \mathbf{u} \rangle = 0 \\ (\mathbf{D}_{\mathbf{z}}^{\Re} \mathbf{A}^{\Re} + \mathbf{D}_{\mathbf{z}}^{\Im} \mathbf{A}^{\Im})\mathbf{u} > \mathbf{0}, \end{cases} \quad (11)$$

with  $\mathbf{u} \in \mathbb{R}^n$ , and  $\mathbf{D}_{\mathbf{v}} := \text{diag}(\mathbf{v})$ ,  $\mathbf{H}_{\mathbf{v}} := \Im(\mathbf{D}_{\mathbf{v}}^* \mathbf{A}) = \mathbf{D}_{\mathbf{v}}^{\Re} \mathbf{A}^{\Im} - \mathbf{D}_{\mathbf{v}}^{\Im} \mathbf{A}^{\Re}$  for  $\mathbf{v} \in \mathbb{C}^m$ .

<sup>2</sup>Simply observe that for  $\Phi \sim \mathcal{N}_{\mathbb{C}}(0, 2)$  and real  $\mathbf{x}$ ,  $\|\Phi\mathbf{x}\|^2 = \|\bar{\Phi}\mathbf{x}\|^2$  with  $\bar{\Phi} := [(\Phi^{\Re})^{\top}, (\Phi^{\Im})^{\top}]^{\top} \sim \mathcal{N}^{2m \times n}(0, 1)$ .

A meaningful estimate  $\hat{\mathbf{x}} \in \mathbb{R}^n$  of  $\mathbf{x}$  should thus respect the constraints (11). Interestingly, if we discard the second line of (11), the remaining constraints amount to imposing

$$\mathbf{A}_z \mathbf{u} = \mathbf{e}_1 = (1, 0, \dots, 0)^\top = \mathbf{A}_z \mathbf{x}, \quad (12)$$

with

$$\mathbf{A}_v := \begin{pmatrix} (\boldsymbol{\alpha}_v^{\Re})^\top \\ (\boldsymbol{\alpha}_v^{\Im})^\top \\ \mathbf{H}_v \end{pmatrix} = \begin{pmatrix} \frac{1}{\kappa\sqrt{m}} ((\mathbf{v}^{\Re})^\top \mathbf{A}^{\Re} + (\mathbf{v}^{\Im})^\top \mathbf{A}^{\Im}) \\ \frac{1}{\kappa\sqrt{m}} ((\mathbf{v}^{\Re})^\top \mathbf{A}^{\Im} - (\mathbf{v}^{\Im})^\top \mathbf{A}^{\Re}) \\ \mathbf{D}_v^{\Re} \mathbf{A}^{\Im} - \mathbf{D}_v^{\Im} \mathbf{A}^{\Re} \end{pmatrix} \in \mathbb{R}^{(m+2) \times n}, \quad \mathbf{v} \in \mathbb{C}^m, \quad (13)$$

and where the equality  $\mathbf{A}_z \mathbf{x} = \mathbf{e}_1$  reformulates the hypothesis  $\|\mathbf{A}\mathbf{x}\|_1 = \kappa\sqrt{m}$  since  $\mathbf{H}_z \mathbf{x} = \mathbf{0}$ .

In other words, a subset of the constraints in (11) leads to the equivalent CS recovery problem (12) where we aim at estimating  $\mathbf{x}$  from  $\mathbf{A}_z \mathbf{x} = \mathbf{e}_1$ . Therefore, if one can show that, for some  $0 < \delta < 1$ ,  $\mathbf{A}_z$  respects w.h.p. the RIP( $\mathcal{K} - \mathcal{K}, \delta$ ) defined in (6), then, any estimate  $\hat{\mathbf{x}} \in \mathcal{K}$  satisfying (11) is sure to recover the direction of  $\mathbf{x}$  since

$$\|\hat{\mathbf{x}} - \mathbf{x}\|^2 \leq \frac{1}{1-\delta} \|\mathbf{A}_z \hat{\mathbf{x}} - \mathbf{A}_z \mathbf{x}\| = 0.$$

Moreover, as imposing both (12) and  $\hat{\mathbf{x}} \in \mathcal{K}$  often leads to impractical algorithms (in fact, an NP-hard, constrained  $\ell_0$  minimization for the recovery of sparse signals), a more practical result derives easily from the model (12) and from the definition of recoverable set introduced in Sec. 2.

**Theorem 3.1** (Perfect signal direction estimation in noiseless PO-CS). *Let  $\mathcal{K} \subset \mathbb{R}^n$  be a symmetric cone that is  $(\Delta_{\mathcal{K}}, \delta_0)$ -recoverable with the instance optimal algorithm  $\Delta_{\mathcal{K}}$ . Given  $\mathbf{A} \in \mathbb{C}^{m \times n}$ , we can recover the direction of  $\mathbf{x} \in \mathcal{K}$  from its  $m$  phase-only measurements  $\mathbf{z} = \text{sign}_{\mathbb{C}}(\mathbf{A}\mathbf{x})$  if the matrix  $\mathbf{A}_z \in \mathbb{R}^{(m+2) \times n}$  built from  $\mathbf{A}$  and  $\mathbf{z}$  in (13) respects the RIP( $\mathcal{K} - \mathcal{K}, \delta$ ) with  $0 < \delta < \delta_0$ . More precisely, under the hypothesis  $\mathbf{A}_z \mathbf{x} = \mathbf{e}_1$  (which corresponds to  $\|\mathbf{A}\mathbf{x}\|_1 = \kappa\sqrt{m}$ ), we have  $\hat{\mathbf{x}} = \mathbf{x}$  if  $\hat{\mathbf{x}} = \Delta_{\mathcal{K}}(\mathbf{e}_1, \mathbf{A}_z; 0)$ .*

*Proof.* The proof is a direct translation of Def. 2.1; in (3),  $e_0(\mathbf{x}, \mathcal{K}) = 0$  since  $\mathbf{x} \in \mathcal{K}$ , and  $\varepsilon = 0$  according to the equivalent noiseless sensing model (12).  $\square$

**Remark 3.2.** *We can question the specific choice of the normalization (9) as a way to complement the constraint  $\mathbf{H}_z \mathbf{u} = \mathbf{0}$  (and thus  $\mathbf{H}_z$ ) to reach a RIP matrix  $\mathbf{A}_z$ . For instance, we could replace  $\boldsymbol{\alpha}_z$  in (13) by an independent real Gaussian random vector  $\mathbf{g}$  and remove the signal normalization ambiguity by arbitrarily assuming  $\langle \mathbf{g}, \mathbf{x} \rangle = 1$  (which should happen with probability 1). In this case,  $\mathbf{A}'_z := (\mathbf{g}^\top, \mathbf{H}_z^\top)^\top$  would offer another equivalent linear sensing model  $\mathbf{A}'_z \mathbf{u} = \mathbf{e}_1$ , hoping then to prove that  $\mathbf{A}'_z$  respects the RIP. However, the use of (9) is critical. As developed in the proof of the next theorem (see Sec. 5.2), for any low-complexity vector  $\mathbf{u}$  and in the case where  $\mathbf{A} = \Phi/\sqrt{m}$  with  $\Phi \sim \mathcal{N}_{\mathbb{C}}^{m \times n}(0, 1)$ ,  $\langle \boldsymbol{\alpha}_z^{\Re}, \mathbf{u} \rangle$  in the first component of  $\mathbf{A}_z \mathbf{u}$  estimates the projection  $\langle \frac{\mathbf{x}}{\|\mathbf{x}\|}, \mathbf{u} \rangle$ . Moreover, since  $\mathbf{H}_z \mathbf{x} = \mathbf{0}$ ,  $\|\mathbf{H}_z \mathbf{u}\|^2$  approximates  $\|\mathbf{u} - \langle \frac{\mathbf{x}}{\|\mathbf{x}\|}, \mathbf{u} \rangle \frac{\mathbf{x}}{\|\mathbf{x}\|}\|^2$ . By Pythagoras, we can then prove that, w.h.p.,  $\|\mathbf{A}_z \mathbf{u}\|^2 = (\langle \boldsymbol{\alpha}_z^{\Re}, \mathbf{u} \rangle)^2 + \|\mathbf{H}_z \mathbf{u}\|^2$  encodes  $\|\mathbf{u}\|^2$  up to a controlled multiplicative distortion.*

The key result of this work (proved in Sec. 5) consists in showing that the matrix  $\mathbf{A}_z$  in (13) built from the association of a complex Gaussian random matrix  $\mathbf{A}$  and a signal  $\mathbf{x} \in \mathbb{R}^n$ , respects, w.h.p., the RIP.

**Theorem 3.3.** *Given a symmetric cone  $\mathcal{K} \subset \mathbb{R}^n$ ,  $\delta > 0$ ,  $\mathbf{A} = \Phi/\sqrt{m}$  with  $\Phi \sim \mathcal{N}_{\mathbb{C}}^{m \times n}(0, 2)$ ,  $\mathbf{x} \in \mathbb{R}^n$ , and  $\mathbf{A}_z$  defined in (13) from  $\mathbf{A}$  and  $\mathbf{z} = \text{sign}_{\mathbb{C}}(\mathbf{A}\mathbf{x})$  with  $\kappa = \sqrt{\pi/2}$ , if*

$$m \geq C(1 + \delta^{-2})w^2((\mathcal{K} - \mathbb{R}\mathbf{x}) \cap \mathbb{S}^{n-1}), \quad (14)$$

*then, with probability exceeding  $1 - C \exp(-c\delta^2 m)$ ,  $\mathbf{A}_z$  satisfies the RIP( $\mathcal{K}, \delta$ ).*

Therefore, if  $\mathbf{x} \in \mathcal{K}$  and  $\mathcal{K}$  is symmetric (as for the sets of sparse signals and low-rank matrices, or for any union of subspaces), we have  $\mathcal{K} - \mathcal{K} - \mathbb{R}\mathbf{x} \subset \mathcal{K}^{(3)} := \mathcal{K} + \mathcal{K} + \mathcal{K}$ , and we deduce that, from Thm 3.1 and Thm 3.3, we can exactly recover w.h.p. the direction of a signal  $\mathbf{x} \in \mathcal{K}$  from the phase of its complex Gaussian random measurements provided  $m$  is large compared to the intrinsic complexity of  $\mathcal{K}^{(3)}$  — as measured by  $w^2(\mathcal{K}^{(3)} \cap \mathbb{S}^{n-1})$ .

In Sec. 6, we illustrate this result in the case of phase-only complex Gaussian random measurements of sparse signals, with a recovery of their direction ensured by BPDN. In particular, we show there that PO-CS requires about twice the number of measurements required for perfect signal recovery in linear CS. This is naively expected as the model (8) provides  $m$  constraints (associated with  $m$  phases) compared to (1) that delivers  $2m$  independent observations.

## 4 Robust and Stable Signal Direction Estimation

Recovering a low-complexity signal from its phase-only observations can be made both robust and stable: we can allow for some noise in the PO-CS model (8), as well as some signal model mismatch — assuming  $\mathbf{x} \notin \mathcal{K}$  with small  $e_0(\mathbf{x}, \mathcal{K})$  in (5) — while still accurately estimating the direction of  $\mathbf{x}$ .

Given an unknown signal  $\mathbf{x} \in \mathbb{R}^n$ , we consider the noisy PO-CS model

$$\mathbf{z} = \text{sign}_{\mathbb{C}}(\mathbf{A}\mathbf{x}) + \boldsymbol{\epsilon} = \mathbf{z}_0 + \boldsymbol{\epsilon}, \quad (15)$$

where  $\boldsymbol{\epsilon} \in \mathbb{C}^m$  stands for some bounded noise with  $\|\boldsymbol{\epsilon}\|_{\infty} \leq \tau$  for some level  $\tau > 0$ . We note that this model encompasses a direct disturbance of the phase of (8); if  $\mathbf{z} = e^{i\boldsymbol{\xi}} \odot \text{sign}_{\mathbb{C}}(\mathbf{A}\mathbf{x})$ , where the exponential applies componentwise,  $\odot$  is the Hadamard product, and  $\|\boldsymbol{\xi}\|_{\infty} \leq \tau$ , then  $\mathbf{z} = \text{sign}_{\mathbb{C}}(\mathbf{A}\mathbf{x}) + \boldsymbol{\epsilon}$ , with  $\boldsymbol{\epsilon} := (e^{i\boldsymbol{\xi}} - 1) \odot \text{sign}_{\mathbb{C}}(\mathbf{A}\mathbf{x})$  such that  $\|\boldsymbol{\epsilon}\|_{\infty} \leq \tau$  (since  $|e^{i\mu} - 1| \leq |\mu|$  for  $\mu \in \mathbb{R}$ ).

As the model (15) does not depend on the signal amplitude, we can still work under the hypothesis that  $\|\mathbf{A}\mathbf{x}\|_1 = \kappa\sqrt{m}$  for some  $\kappa > 0$ . From Lemma 5.2 and (19), this shows that, for  $\mathbf{A} = \Phi/\sqrt{m}$  with  $\Phi \sim \mathcal{N}_{\mathbb{C}}(0, 2)$ ,  $\delta > 0$  and  $\kappa = \sqrt{\pi/2}$ ,  $(1 + \delta)^{-1} \leq \|\mathbf{x}\| \leq (1 - \delta)^{-1}$  with probability exceeding  $1 - C \exp(-c\delta^2 m)$ .

Under this hypothesis, given the definition of  $\mathbf{A}_v$  for any  $\mathbf{v} \in \mathbb{C}^m$  in (12), we find

$$\mathbf{A}_z \mathbf{x} = \mathbf{A}_{z_0} \mathbf{x} + \mathbf{A}_{\boldsymbol{\epsilon}} \mathbf{x} = \mathbf{e}_1 + \boldsymbol{\epsilon}_x,$$

with  $\boldsymbol{\epsilon}_x := \mathbf{A}_{\boldsymbol{\epsilon}} \mathbf{x}$ . In other words, if  $\|\boldsymbol{\epsilon}_x\| \leq \varepsilon$  for some noise level  $\varepsilon > 0$ , the signal  $\mathbf{x}$  respects the following  $\ell_2$ -fidelity constraint:

$$\|\mathbf{A}_z \mathbf{u} - \mathbf{e}_1\| \leq \varepsilon. \quad (16)$$



Moreover, in the case where  $\mathbf{A}$  is a normalized complex Gaussian matrix, it is easy to estimate  $\varepsilon$ .

**Lemma 4.1.** *Given a signal  $\mathbf{x} \in \mathbb{R}^n$ , and its noisy PO-CS measurements  $\mathbf{z} = \text{sign}_{\mathbb{C}}(\mathbf{A}\mathbf{x}) + \boldsymbol{\epsilon}$  with  $\mathbf{A} = \boldsymbol{\Phi}/\sqrt{m}$ ,  $\boldsymbol{\Phi} \sim \mathcal{N}_{\mathbb{C}}^{m \times n}(0, 2)$ , and  $\|\boldsymbol{\epsilon}\|_{\infty} \leq \tau$ , if  $\|\mathbf{A}\mathbf{x}\|_1 = \kappa\sqrt{m}$  with  $\kappa = \sqrt{\pi}/2$ , then, given  $\delta > 0$ ,*

$$\|\mathbf{A}_{\boldsymbol{\epsilon}}\mathbf{x}\| \leq \sqrt{2}\tau \frac{1+\delta}{1-\delta}$$

with probability exceeding  $1 - C \exp(-c\delta^2 m)$ .

*Proof.* If  $\boldsymbol{\Phi} \sim \mathcal{N}_{\mathbb{C}}^{m \times n}(0, 2)$ , Lemma 5.2 and the  $\ell_2$ -norm concentration of Gaussian random projections [29, 16] show that we have jointly

$$\left| \frac{1}{\kappa m} \|\boldsymbol{\Phi}\mathbf{x}\|_1 - \|\mathbf{x}\| \right| \leq \delta \|\mathbf{x}\|, \quad \left| \frac{1}{m} \|\boldsymbol{\Phi}\mathbf{x}\|^2 - \|\mathbf{x}\|^2 \right| \leq \delta \|\mathbf{x}\|^2,$$

with probability exceeding  $1 - C \exp(-c\delta^2 m)$ . Under this event, for  $\mathbf{A} = \boldsymbol{\Phi}/\sqrt{m}$ , and under the hypothesis  $\sqrt{m}\|\mathbf{A}\mathbf{x}\|_1 = \|\boldsymbol{\Phi}\mathbf{x}\|_1 = \kappa m$ , we find from (12)

$$\begin{aligned} \|\mathbf{A}_{\boldsymbol{\epsilon}}\mathbf{x}\|^2 &= |\langle \boldsymbol{\alpha}_{\boldsymbol{\epsilon}}, \mathbf{x} \rangle|^2 + \|\mathbf{H}_{\boldsymbol{\epsilon}}\mathbf{x}\|^2 = \frac{1}{\kappa^2 m} |\langle \boldsymbol{\epsilon}, \mathbf{A}\mathbf{x} \rangle|^2 + \|\Im(\text{diag}(\boldsymbol{\epsilon})^* \mathbf{A}\mathbf{x})\|^2 \\ &\leq \frac{\tau^2}{\kappa^2 m} \|\mathbf{A}\mathbf{x}\|_1^2 + \tau^2 \|\mathbf{A}\mathbf{x}\|^2 \leq \tau^2 (1 + (1 + \delta) \|\mathbf{x}\|^2) \leq \tau^2 (1 + \frac{1+\delta}{(1-\delta)^2 \kappa^2 m^2} \|\boldsymbol{\Phi}\mathbf{x}\|_1^2) \\ &\leq \frac{2(1+\delta)^2}{(1-\delta)^2} \tau^2. \end{aligned}$$

where we used Holder's inequality for the first inequality. This concludes the proof.  $\square$

From these observations, the definition of recoverable set in Sec. 2 provides the following theorem.

**Theorem 4.2** (Signal direction estimation in noisy PO-CS). *Let  $\mathcal{K} \subset \mathbb{R}^n$  be a symmetric cone that is  $(\Delta_{\mathcal{K}}, \delta_0)$ -recoverable with the instance optimal algorithm  $\Delta_{\mathcal{K}}$ . Given  $\mathbf{A} \in \mathbb{C}^{m \times n}$ , the signal  $\mathbf{x} \in \mathbb{R}^n$  observed from its  $m$  noisy PO-CS measurements  $\mathbf{z} = \text{sign}_{\mathbb{C}}(\mathbf{A}\mathbf{x}) + \boldsymbol{\epsilon}$ , and the hypothesis that  $\|\mathbf{A}\mathbf{x}\|_1 = \kappa\sqrt{m}$  for some  $\kappa > 0$ , if the matrix  $\mathbf{A}_{\mathbf{z}} \in \mathbb{R}^{(m+2) \times n}$  built from  $\mathbf{A}$  and  $\mathbf{z}$  respects the RIP( $\mathcal{K} - \mathcal{K}, \delta$ ) with  $0 < \delta < \delta_0$ , and if  $\|\mathbf{A}_{\boldsymbol{\epsilon}}\mathbf{x}\| \leq \varepsilon$  for some  $\varepsilon > 0$ , then the vector  $\hat{\mathbf{x}} = \Delta_{\mathcal{K}}(\mathbf{e}_1, \mathbf{A}_{\mathbf{z}}; \varepsilon)$  estimates  $\mathbf{x}$  according to the bound (5).*

Interestingly, the RIP condition imposed on  $\mathbf{A}_{\mathbf{z}}$  by this theorem can be met for complex Gaussian random matrices. Indeed, in the case where  $\mathbf{A} = \boldsymbol{\Phi}/\sqrt{m}$  with  $\boldsymbol{\Phi} \sim \mathcal{N}_{\mathbb{C}}^{m \times n}(0, 2)$ , if  $m$  is sufficiently large and the noise level  $\tau$  is small enough, the following theorem shows that w.h.p. the matrix  $\mathbf{A}_{\mathbf{z}} = \mathbf{A}_{\mathbf{z}_0} + \mathbf{A}_{\boldsymbol{\epsilon}}$  respects the RIP over  $\mathcal{K}$  with a constant reduced by  $\tau$ .

**Theorem 4.3.** *Given  $\delta, \tau > 0$  with  $0 < \delta + 9\tau < 1$ ,  $\mathbf{x} \in \mathbb{R}^n$ ,  $\mathbf{A} = \boldsymbol{\Phi}/\sqrt{m}$  with  $\boldsymbol{\Phi} \sim \mathcal{N}_{\mathbb{C}}(0, 2)$ ,  $\boldsymbol{\epsilon} \in \mathbb{C}^m$  with  $\|\boldsymbol{\epsilon}\|_{\infty} \leq \tau$ , and  $\mathbf{z} = \text{sign}_{\mathbb{C}}(\mathbf{A}\mathbf{x}) + \boldsymbol{\epsilon}$ , if*

$$m \geq C(1 + \delta^{-2})w((\mathcal{K} - \mathbb{R}\mathbf{x}) \cap \mathbb{S}^{n-1})^2, \quad (17)$$

then, the matrix  $\mathbf{A}_{\mathbf{z}}$  defined in (13) from  $\mathbf{z}$  and  $\mathbf{A}$  with  $\kappa = \sqrt{\pi}/2$  satisfies the RIP( $\mathcal{K}, \delta + 9\tau$ ) with probability exceeding  $1 - C \exp(-c\delta^2 m)$ .

Therefore, in the case of PO-CS associated with a complex Gaussian random matrix, and if the symmetric cone  $\mathcal{K}$  is  $(\Delta_{\mathcal{K}}, \delta_0)$ -recoverable, as soon as  $m$  is large compared to  $(1 + \delta^{-2})w((\mathcal{K} - \mathbb{R}\mathbf{x}) \cap \mathbb{S}^{n-1})^2$  with  $\delta + 9\tau \leq \delta_0$ , then, Thm 4.2 and Thm 4.3 tell us that we can stably and robustly recover the direction of  $\mathbf{x} \in \mathbb{R}^n$  — in the sense of (5) — from its PO-CS measurements.

**Remark 4.4.** *While we did not optimize it in the corresponding proof, the condition imposed on  $\tau$  in Thm. 4.3 makes sense. Indeed, combining Thm 4.2 and Thm 4.3 we conclude that the estimation of the signal direction is possible if  $\tau < \delta_0/9 = O(1)$  provided  $m$  is large enough (as imposed by setting  $\delta = \delta_0 - 9\tau$  in (17)). Moreover, for the noisy PO-CS model  $\mathbf{z} = e^{i\xi} \odot \text{sign}_{\mathbb{C}}(\mathbf{A}\mathbf{x})$ , if  $\xi_k \sim_{\text{i.i.d.}} \mathcal{U}([-\pi, \pi])$  for  $k \in [m]$ , then the phase of each component of  $\mathbf{z}$  is uniformly distributed over  $[-\pi, \pi]$  whatever the value of  $\mathbf{x}$ ; estimating  $\mathbf{x}/\|\mathbf{x}\|$  is then impossible. In this case, we have  $\tau = \|(e^{i\xi} - \mathbf{1}) \odot \text{sign}_{\mathbb{C}}(\mathbf{A}\mathbf{x})\|_{\infty} = \|e^{i\xi} - \mathbf{1}\|_{\infty} > \|\xi\|_{\infty}$ , with  $\|\xi\|_{\infty}$  arbitrary close to  $\pi$  (w.h.p., for large value of  $m$ ). Although the question of the existence of a robust algorithm for  $\tau \in [\delta_0/9, \pi]$  remains open, we thus see that imposing  $\tau = O(1)$  to recover the signal direction is realistic.*

## 5 Proofs

This section is devoted to proving the RIP of a matrix built in (13) from a complex Gaussian random matrix  $\mathbf{A}$  and a signal  $\mathbf{x} \in \mathbb{R}^n$ , in the context of noiseless and noisy PO-CS. We first introduce a few useful lemmata.

### 5.1 Auxiliary lemmata

We first need this classical result from Ledoux and Talagrand [30, Eq. 1.6].

**Lemma 5.1.** *If the function  $F : \mathbb{R}^n \rightarrow \mathbb{R}$  is Lipschitz with constant  $\lambda$ , i.e.,  $|F(\mathbf{u}) - F(\mathbf{u}')| \leq \lambda\|\mathbf{u} - \mathbf{u}'\|$  for all  $\mathbf{u}, \mathbf{u}' \in \mathbb{R}^n$ , then, for  $r > 0$  and  $\gamma \sim \mathcal{N}^n(0, 1)$ ,*

$$\mathbb{P}(|F(\gamma) - \mathbb{E}(F(\gamma))| > r) \leq 2 \exp(-\frac{1}{2}r^2\lambda^{-2}). \quad (18)$$

The following lemma characterizes the concentration of the random variable  $\|\Phi\mathbf{z}\|_1$  for a complex Gaussian random matrix  $\Phi$  given  $\mathbf{z} \in \mathbb{R}^n$ .

**Lemma 5.2.** *Given  $\delta > 0$ ,  $\mathbf{z} \in \mathbb{R}^n$ , and  $\Phi \sim \mathcal{N}_{\mathbb{C}}^{m \times n}(0, 2)$ , we have*

$$\mathbb{P}\left[\left|\frac{1}{\kappa m}\|\Phi\mathbf{z}\|_1 - \|\mathbf{z}\|\right| > \delta\|\mathbf{z}\|\right] \leq C \exp(-c\delta^2 m). \quad (19)$$

*Proof.* By homogeneity of (19), we can assume that  $\|\mathbf{z}\| = 1$ . In this case,  $\Phi\mathbf{z} \sim \mathbf{g} = \mathbf{g}^{\Re} + i\mathbf{g}^{\Im}$ , with  $\mathbf{g}^{\Re}, \mathbf{g}^{\Im} \sim_{\text{i.i.d.}} \mathcal{N}_{\mathbb{C}}^m(0, 1)$ . In this context,  $\frac{1}{\kappa m}\|\Phi\mathbf{z}\|_1 \sim \frac{1}{\kappa m}\|\mathbf{g}^{\Re} + i\mathbf{g}^{\Im}\|_1 = \sum_k X_k$ , where each independent random variable  $X_k := ((g_k^{\Re})^2 + (g_k^{\Im})^2)^{1/2}$  follows a Rayleigh distribution with unit scale parameter [31], which gives  $\mathbb{E}X_k = \sqrt{\pi/2} = \kappa$ . Therefore, since each  $X_k$  is a sub-Gaussian random variable [29], the concentration property of the sum of  $m$  such random variables provides [29, 8], for  $t > 0$ ,

$$\mathbb{P}\left[\left|\|\mathbf{g}^{\Re} + i\mathbf{g}^{\Im}\|_1 - \kappa m\right| = \left|\sum_k X_k - \mathbb{E}X_k\right| > t\kappa\right] \leq C \exp(-c\frac{t^2}{m}).$$

The result follows from a simple change of variable. □

We now show that the Gaussian mean width of a bounded set projected on a subspace is bounded by twice the width of the original set.

**Lemma 5.3.** *For a bounded subset  $\mathcal{K}' \subset \mathbb{R}^n$ , and a  $L$ -dimensional subspace  $\mathcal{L} \subset \mathbb{R}^n$  related to the projecting matrix  $\mathbf{P} \in \mathbb{R}^{L \times n}$ , we have*

$$w(\mathbf{P}(\mathcal{K}')) \leq 2w(\mathcal{K}'). \quad (20)$$

*Proof.* Let us define  $\mathcal{L}^\perp$  and  $\mathbf{P}^\perp$  the subspace orthogonal to  $\mathcal{L}$  and its  $(n - L) \times n$  projection matrix, respectively, so that  $(\mathbf{P}^\perp)^\top \mathbf{P}^\perp + \mathbf{P}^\top \mathbf{P} = \mathbf{I}_n$ . Then, for  $\mathbf{g} \sim \mathcal{N}^L(0, 1)$ , which can always be written as  $\mathbf{g} = \mathbf{P}\mathbf{g}'$  for some  $\mathbf{g}' \sim \mathcal{N}^n(0, 1)$  (by marginalization of the Gaussian distribution), we have

$$\begin{aligned} w(\mathbf{P}\mathcal{K}') &= \mathbb{E} \sup_{\mathbf{u} \in \mathcal{K}'} \langle \mathbf{g}, \mathbf{P}\mathbf{u} \rangle = \mathbb{E} \sup_{\mathbf{u} \in \mathcal{K}'} \langle \mathbf{P}\mathbf{g}', \mathbf{P}\mathbf{u} \rangle = \mathbb{E} \sup_{\mathbf{u} \in \mathcal{K}'} \langle \mathbf{P}^\top \mathbf{P}\mathbf{g}', \mathbf{u} \rangle \\ &\leq \mathbb{E} \sup_{\mathbf{u} \in \mathcal{K}'} \max_{\sigma \in \{\pm 1\}} \sigma \langle (\mathbf{P}^\perp)^\top \mathbf{P}^\perp \mathbf{g}', \mathbf{u} \rangle + \langle \mathbf{P}^\top \mathbf{P}\mathbf{g}', \mathbf{u} \rangle \\ &\leq \mathbb{E} \sup_{\mathbf{u} \in \mathcal{K}'} \sum_{\sigma \in \{\pm 1\}} \langle \sigma (\mathbf{P}^\perp)^\top \mathbf{P}^\perp \mathbf{g}' + \mathbf{P}^\top \mathbf{P}\mathbf{g}', \mathbf{u} \rangle \\ &\leq \sum_{\sigma \in \{\pm 1\}} \mathbb{E} \sup_{\mathbf{u} \in \mathcal{K}'} \langle \sigma (\mathbf{P}^\perp)^\top \mathbf{P}^\perp \mathbf{g}' + \mathbf{P}^\top \mathbf{P}\mathbf{g}', \mathbf{u} \rangle = 2w(\mathcal{K}'), \end{aligned}$$

where the last line used the fact that  $\pm(\mathbf{P}^\perp)^\top \mathbf{P}^\perp \mathbf{g}' + \mathbf{P}^\top \mathbf{P}\mathbf{g}' \sim \mathcal{N}^n(0, 1)$ .  $\square$

Finally, this fourth lemma introduces a complex, local sign-product embedding [6, 15]; it shows that projecting the random measurements of any low-complexity vector onto the phase-only measurements of another fixed vector is close, up to a rescaling, to the scalar product between both vectors.

**Lemma 5.4.** *Given a symmetric cone  $\mathcal{K} \subset \mathbb{R}^n$ ,  $\delta > 0$ , and  $\mathbf{z} \in \mathbb{S}^{n-1}$ , if*

$$m \geq C\delta^{-2}w^2(\mathcal{K} \cap \mathbb{S}^{n-1}), \quad (21)$$

*then, with probability exceeding  $1 - C \exp(-c\delta^2 m)$ , the random matrix  $\Phi \sim \mathcal{N}_{\mathbb{C}}^{m \times n}(0, 2)$  satisfies*

$$\left| \frac{1}{\kappa m} \langle \text{sign}_{\mathbb{C}}(\Phi \mathbf{z}), \Phi \mathbf{u} \rangle - \langle \mathbf{z}, \mathbf{u} \rangle \right| \leq \delta \|\mathbf{u}\|, \quad \forall \mathbf{u} \in \mathcal{K}. \quad (22)$$

*Proof.* Since (22) is homogeneous in  $\mathbf{u}$  and  $\mathcal{K}$  is conic, we can assume  $\|\mathbf{u}\| = 1$ . Given  $\mathcal{K}^* = \mathcal{K} \cap \mathbb{S}^{n-1}$ , we must prove that  $\mathbb{P}(E > \delta) \leq C \exp(-c\delta^2 m)$  with

$$E := \sup_{\mathbf{u} \in \mathcal{K}^*} \left| \frac{1}{\kappa m} \langle \text{sign}_{\mathbb{C}}(\Phi \mathbf{z}), \Phi \mathbf{u} \rangle - \langle \mathbf{z}, \mathbf{u} \rangle \right|.$$

Using the decomposition  $\mathbf{u} = \mathbf{u}^\parallel + \mathbf{u}^\perp$  with  $\mathbf{u}^\parallel := \langle \mathbf{z}, \mathbf{u} \rangle \mathbf{z}$  and  $\mathbf{u}^\perp := (\mathbf{u} - \langle \mathbf{z}, \mathbf{u} \rangle \mathbf{z})$  (with  $\langle \mathbf{u}^\parallel, \mathbf{u}^\perp \rangle = 0$ ), and the triangular inequality, we find

$$\begin{aligned} E &\leq \sup_{\mathbf{u} \in \mathcal{K}^*} \left| \langle \mathbf{z}, \mathbf{u} \rangle \right| \left| \frac{1}{\kappa m} \|\Phi \mathbf{z}\|_1 - 1 \right| + \frac{1}{\kappa m} \sup_{\mathbf{u} \in \mathcal{K}^*} \left| \langle \text{sign}_{\mathbb{C}}(\Phi \mathbf{z}), \Phi \mathbf{u}^\perp \rangle \right| \\ &\leq \underbrace{\left| \frac{1}{\kappa m} \|\Phi \mathbf{z}\|_1 - 1 \right|}_{A(\Phi)} + \underbrace{\frac{1}{\kappa m} \sup_{\mathbf{u} \in \mathcal{K}^*} \left| \langle \text{sign}_{\mathbb{C}}(\Phi \mathbf{z}), \Phi \mathbf{u}^\perp \rangle \right|}_{B(\Phi)}. \end{aligned}$$

Moreover, since  $\mathbb{P}(E > \delta) \leq \mathbb{P}(2 \max(A, B) > \delta) \leq \mathbb{P}(A > \delta/2) + \mathbb{P}(B > \delta/2)$ , and  $\mathbb{P}(A > \delta/2) \leq C \exp(-c\delta^2 m)$  from Lemma 5.2, we only have to prove that  $\mathbb{P}(B > \delta/2) \leq C \exp(-c\delta^2 m)$  if (21) holds.

From the rotational invariance of the Gaussian distribution,  $\Phi$  and  $\Phi\mathbf{R}$  have the same distribution for any rotation matrix  $\mathbf{R} \in \mathbb{R}^{n \times n}$ , and  $\mathbb{P}(B(\Phi) > \delta/2) = \mathbb{P}(B(\Phi\mathbf{R}) > \delta/2)$ . Since  $\mathbf{z} \in \mathbb{S}^{n-1}$  is fixed, we decide to set  $\mathbf{R}$  such that  $\mathbf{R}\mathbf{z} = \mathbf{e}_1$  and  $(\mathbf{R}\mathbf{v})_1 = 0$  for all  $\mathbf{v} \in (\mathbb{R}\mathbf{z})^\perp := \{\mathbf{u} \in \mathbb{R}^n : \langle \mathbf{u}, \mathbf{z} \rangle = 0\}$ . Then, defining the restriction matrix  $\mathbf{S} \in \mathbb{R}^{(n-1) \times n}$  such that  $\mathbf{S}\mathbf{w} = (w_2, \dots, w_n)^\top \in \mathbb{R}^{n-1}$  for any  $\mathbf{w} \in \mathbb{R}^n$ , we observe that, for all  $\mathbf{v} \in (\mathbb{R}\mathbf{z})^\perp$ ,  $\Phi\mathbf{R}\mathbf{v} = \mathbf{G}\mathbf{S}\mathbf{R}\mathbf{v}$  with  $\mathbf{G} = \Phi\mathbf{S}^\top \sim \mathcal{N}_{\mathbb{C}}^{m \times (n-1)}(0, 2)$  independent of  $\mathbf{g} := \Phi\mathbf{R}\mathbf{z} = \Phi\mathbf{e}_1 \sim \mathcal{N}_{\mathbb{C}}^m(0, 2)$ .

Therefore, since  $\mathbf{u}^\perp \in (\mathbb{R}\mathbf{z})^\perp$  and  $\mathbf{S}\mathbf{R}\mathbf{u}^\perp = \mathbf{S}\mathbf{R}\mathbf{u}$  by design of  $\mathbf{S}$  and  $\mathbf{R}$ , we can write

$$|\langle \text{sign}_{\mathbb{C}}(\Phi\mathbf{R}\mathbf{z}), \Phi\mathbf{R}\mathbf{u}^\perp \rangle| = |\langle \text{sign}_{\mathbb{C}}(\mathbf{g}), \mathbf{G}\mathbf{S}\mathbf{R}\mathbf{u}^\perp \rangle| = |\langle \text{sign}_{\mathbb{C}}(\mathbf{g}), \mathbf{G}\mathbf{S}\mathbf{R}\mathbf{u} \rangle|, \quad (23)$$

From the independence of  $\mathbf{g}$  and  $\mathbf{G}$ , the equivalence (23) allows us to condition the random variable  $B(\Phi\mathbf{R})$  to the value of  $\mathbf{g}$  while preserving the distribution of  $\mathbf{G}$ . We thus focus on bounding  $\mathbb{P}(B(\Phi\mathbf{R}) > \delta/2 \mid \mathbf{g})$  and eventually use  $\mathbb{P}(B(\Phi\mathbf{R}) > \delta/2) = \mathbb{E}\mathbb{P}(B(\Phi\mathbf{R}) > \delta/2 \mid \mathbf{g})$  by expectation over  $\mathbf{g}$ .

For  $\mathbf{g}$  fixed,  $|\langle \text{sign}_{\mathbb{C}}(\mathbf{g}), \mathbf{G}\mathbf{S}\mathbf{R}\mathbf{u} \rangle| = |\langle \mathbf{G}^* \text{sign}_{\mathbb{C}}(\mathbf{g}), \mathbf{S}\mathbf{R}\mathbf{u} \rangle|$  is distributed as  $\sqrt{m} |\langle \gamma, \mathbf{S}\mathbf{R}\mathbf{u} \rangle|$  with  $\gamma \sim \mathcal{N}_{\mathbb{C}}^{n-1}(0, 2)$ , since  $\|\text{sign}_{\mathbb{C}}(\mathbf{g})\| = \sqrt{m}$ . Therefore, using  $|\langle \mathbf{a}, \mathbf{b} \rangle| \leq 2 \max(|\langle \mathbf{a}^{\Re}, \mathbf{b} \rangle|, |\langle \mathbf{a}^{\Im}, \mathbf{b} \rangle|)$  for any  $\mathbf{a} \in \mathbb{C}^d$  and  $\mathbf{b} \in \mathbb{R}^d$ , we find

$$\begin{aligned} \mathbb{P}(B(\Phi\mathbf{R}) > \frac{\delta}{2} \mid \mathbf{g}) &= \mathbb{P}\left(\frac{1}{\kappa\sqrt{m}} \sup_{\mathbf{u} \in \mathcal{K}^*} |\langle \gamma, \mathbf{S}\mathbf{R}\mathbf{u} \rangle| \geq \frac{\delta}{2} \mid \mathbf{g}\right) \\ &\leq \mathbb{P}\left(\frac{1}{\kappa\sqrt{m}} \sup_{\mathbf{u} \in \mathcal{K}^*} \max(|\langle \gamma^{\Re}, \mathbf{S}\mathbf{R}\mathbf{u} \rangle|, |\langle \gamma^{\Im}, \mathbf{S}\mathbf{R}\mathbf{u} \rangle|) \geq \frac{\delta}{4} \mid \mathbf{g}\right) \\ &\leq 2\mathbb{P}(F(\mathbf{g}') \geq \frac{\delta}{4} \mid \mathbf{g}), \end{aligned}$$

where we defined  $F : \mathbf{v} \in \mathbb{R}^{n-1} \rightarrow F(\mathbf{v}) := \sup_{\mathbf{u} \in \mathcal{K}^*} |\langle \mathbf{v}, \mathbf{S}\mathbf{R}\mathbf{u} \rangle| / (\kappa\sqrt{m})$ , and we used the union bound and the fact that  $\gamma^{\Re}, \gamma^{\Im} \sim_{\text{i.i.d.}} \mathbf{g}' \sim \mathcal{N}^{n-1}(0, 1)$ .

Let us characterize the random variable  $F(\mathbf{g}')$ . We first observe that

$$\mathbb{E}F(\mathbf{g}') = \frac{1}{\kappa\sqrt{m}} \mathbb{E} \sup_{\mathbf{u} \in \mathcal{K}^*} |\langle \mathbf{g}', \mathbf{S}\mathbf{R}\mathbf{u} \rangle| = \frac{1}{\kappa\sqrt{m}} \mathbb{E} \sup_{\mathbf{u} \in \mathcal{K}^*} \langle \mathbf{g}', \mathbf{S}\mathbf{R}\mathbf{u} \rangle = \frac{1}{\kappa\sqrt{m}} w(\mathcal{K}'),$$

where  $\mathcal{K}' := (\mathbf{S}\mathbf{R}\mathcal{K}^*)$ , the width  $w$  is defined in (4), and we used the symmetry of  $\mathcal{K}$ .

Since  $\mathbf{S}$  is a projector with  $\mathbf{S}^\top \mathbf{S} + \mathbf{e}_1 \mathbf{e}_1^\top = \mathbf{I}_n$ , Lemma 5.3 tells us that

$$\kappa\sqrt{m} \mathbb{E}F(\mathbf{g}') \leq w(\mathcal{K}') \leq 2w(\mathbf{R}\mathcal{K}^*) = 2w(\mathcal{K}^*), \quad (24)$$

since the Gaussian mean width is invariant under rotation.

Moreover,  $F$  is Lipschitz with constant  $\lambda = 1/(\kappa\sqrt{m})$  since  $\|\mathbf{S}\mathbf{R}\mathbf{u}\| \leq 1$  for all  $\mathbf{u} \in \mathcal{K}^*$ ,

$$|F(\mathbf{v}) - F(\mathbf{v}')| \leq \frac{1}{\kappa\sqrt{m}} \sup_{\mathbf{u} \in \mathcal{K}^*} |\langle \mathbf{v} - \mathbf{v}', \mathbf{S}\mathbf{R}\mathbf{u} \rangle| \leq \frac{1}{\kappa\sqrt{m}} \|\mathbf{v} - \mathbf{v}'\|.$$

We can thus invoke Lemma 5.1 on  $F$  and  $\mathbf{g}'$  to conclude that, from (24) and for  $r > 0$ ,

$$\begin{aligned} \mathbb{P}\left(\frac{1}{\kappa\sqrt{m}} \mathbb{E} \sup_{\mathbf{u} \in \mathcal{K}^*} |\langle \mathbf{g}', \mathbf{S}\mathbf{R}\mathbf{u} \rangle| \geq \frac{2}{\kappa\sqrt{m}} w(\mathcal{K}^*) + r \mid \mathbf{g}\right) \\ \leq \mathbb{P}\left(\frac{1}{\kappa\sqrt{m}} \mathbb{E} \sup_{\mathbf{u} \in \mathcal{K}^*} |\langle \mathbf{g}', \mathbf{S}\mathbf{R}\mathbf{u} \rangle| \geq \mathbb{E}F(\mathbf{g}') + r \mid \mathbf{g}\right) = \mathbb{P}(F(\mathbf{g}') \geq \mathbb{E}F(\mathbf{g}') + r \mid \mathbf{g}) \\ \leq \mathbb{P}(|F(\mathbf{g}') - \mathbb{E}F(\mathbf{g}')| > r \mid \mathbf{g}) \leq 2 \exp(-\frac{1}{2} r^2 \kappa^2 m). \end{aligned}$$

Taking, *e.g.*,  $r = \delta/8$  and  $m \geq 64 \delta^{-2} \kappa^{-2} w(\mathcal{K}^*)$ , gives  $r + \frac{2}{\kappa\sqrt{m}} w(\mathcal{K}^*) \leq \delta/4$  so that, by expectation over  $\mathbf{g}$ , the proof is concluded from

$$\mathbb{P}(B > \frac{\delta}{2}) = \mathbb{E} \mathbb{P}(B > \frac{\delta}{2} \mid \mathbf{g}) \leq 2\mathbb{E} \mathbb{P}\left(\frac{1}{\kappa\sqrt{m}} \mathbb{E} \sup_{\mathbf{u} \in \mathcal{K}^*} |\langle \mathbf{g}', \mathbf{S}\mathbf{R}\mathbf{u} \rangle| \geq \frac{\delta}{4} \mid \mathbf{g}\right) \leq C \exp(-c\delta^2 m). \quad \square$$

## 5.2 Proof of Thm 3.3

Given a symmetric cone  $\mathcal{K} \in \mathbb{R}^n$ ,  $\mathbf{x} \in \mathbb{R}^n$  (with  $\bar{\mathbf{x}} := \mathbf{x}/\|\mathbf{x}\|$ ),  $\mathbf{A} = \Phi/\sqrt{m}$  with  $\Phi \sim \mathcal{N}_{\mathbb{C}}^{m \times n}(0, 2)$ , and  $\mathbf{z} = \text{sign}_{\mathbb{C}}(\mathbf{A}\mathbf{x})$ , we can now determine under which conditions and with which probability the matrix  $\mathbf{A}_{\mathbf{z}}$  defined in (12) satisfies the RIP over  $\mathcal{K}$  for some distortion  $\delta > 0$ . This amounts to showing that, under the condition (14) and with probability exceeding  $1 - C \exp(-c\delta^2 m)$ ,

$$\left| \|\mathbf{A}_{\mathbf{z}}\mathbf{u}\|^2 - 1 \right| \leq \delta, \quad \forall \mathbf{u} \in \mathcal{K}^* := \mathcal{K} \cap \mathbb{S}^{n-1},$$

where we can assume  $\|\mathbf{u}\| = 1$  since  $\mathcal{K}$  is a cone.

Given the definition of  $\alpha_{\mathbf{z}}$  and  $\mathbf{H}_{\mathbf{z}}$  in (12), for  $\mathbf{u} \in \mathcal{K}^*$ , we see that

$$\begin{aligned} \|\mathbf{A}_{\mathbf{z}}\mathbf{u}\|^2 &= \left\| \begin{bmatrix} \alpha_{\mathbf{z}}^{\Re} \\ \alpha_{\mathbf{z}}^{\Im} \end{bmatrix}^{\top} \mathbf{u} \right\|^2 + \|\mathbf{H}_{\mathbf{z}}\mathbf{u}\|^2 = |\langle \alpha_{\mathbf{z}}, \mathbf{u} \rangle|^2 + \|\mathbf{H}_{\mathbf{z}}\mathbf{u}\|^2 \\ &= \frac{1}{\kappa^2 m^2} |\langle \text{sign}_{\mathbb{C}}(\Phi\mathbf{x}), \Phi\mathbf{u} \rangle|^2 + \|\mathbf{H}_{\mathbf{z}}\mathbf{u}\|^2. \end{aligned} \quad (25)$$

We first bound the first term of (25). Under the condition  $m \geq C(1+\delta^{-2})w^2(\mathcal{K} \cap \mathbb{S}^{n-1})$ , which is involved by (14), Lemma 5.4 and Thm 2.2 inform us that the joint event

$$\mathcal{E} : \quad \left| \frac{1}{\kappa m} \langle \text{sign}_{\mathbb{C}}(\Phi\mathbf{x}), \Phi\mathbf{v} \rangle - \langle \bar{\mathbf{x}}, \mathbf{v} \rangle \right| \leq \frac{3}{8}\delta \quad \text{and} \quad \left| \frac{1}{m} \|\Phi\mathbf{u}\|^2 - 1 \right| \leq \kappa^2 - 1, \quad \forall \mathbf{v} \in \mathcal{K}^*,$$

holds (by union bound) with probability exceeding  $1 - C \exp(-c\delta^2 m)$ . Since for any  $A, B \in \mathbb{C}$ ,  $\left| |A|^2 - |B|^2 \right| \leq |A+B||A-B|$ ,  $\mathcal{E}$  involves

$$\left| \frac{1}{\kappa^2 m^2} |\langle \text{sign}_{\mathbb{C}}(\Phi\mathbf{x}), \Phi\mathbf{u} \rangle|^2 - |\langle \bar{\mathbf{x}}, \mathbf{u} \rangle|^2 \right| \leq \frac{3\delta}{8} \left( 1 + \left| \frac{1}{\kappa m} \langle \text{sign}_{\mathbb{C}}(\Phi\mathbf{x}), \Phi\mathbf{u} \rangle \right| \right) \leq \frac{3\delta}{8} \left( 1 + \frac{1}{\kappa\sqrt{m}} \|\Phi\mathbf{u}\| \right) \leq \frac{3}{4}\delta. \quad (26)$$

We now focus on bounding the second term of (25). We first note that for any  $\mathbf{u} \in \mathcal{K}^*$  decomposed as  $\mathbf{u} = \mathbf{u}^{\parallel} + \mathbf{u}^{\perp}$  with  $\mathbf{u}^{\parallel} = \langle \mathbf{u}, \bar{\mathbf{x}} \rangle \bar{\mathbf{x}}$  and  $\mathbf{u}^{\perp} = (\mathbf{u} - \langle \mathbf{u}, \bar{\mathbf{x}} \rangle \bar{\mathbf{x}})$ , we have  $\mathbf{H}_{\mathbf{z}}\mathbf{u} = \mathbf{H}_{\mathbf{z}}\mathbf{u}^{\perp}$  since  $\mathbf{H}_{\mathbf{z}}\bar{\mathbf{x}} = \mathbf{0}$ . We are going to show that

$$p_{\mathbf{z}}(\Phi) := \mathbb{P} \left[ \exists \mathbf{u} \in \mathcal{K}^*, \left| \|\mathbf{H}_{\mathbf{z}}\mathbf{u}^{\perp}\|^2 - \|\mathbf{u}^{\perp}\|^2 \right| > \frac{1}{4}\delta \|\mathbf{u}^{\perp}\|^2 \right] \leq C \exp(-c\delta^2 m), \quad (27)$$

provided (14) is satisfied.

Using the rotational invariance of the Gaussian distribution, we note that  $p_{\mathbf{z}}(\Phi) = p_{\mathbf{z}}(\Phi\mathbf{R})$  for any rotation matrix  $\mathbf{R} \in \mathbb{R}^{n \times n}$ . We proceed similarly to the proof of Lemma 5.4 and take  $\mathbf{R}$  such that  $\mathbf{R}\mathbf{x} = \|\mathbf{x}\|\mathbf{e}_1$  and  $(\mathbf{R}\mathbf{v})_1 = 0$  for all  $\mathbf{v} \in (\mathbb{R}\mathbf{x})^{\perp}$ ; we thus find that for all  $\mathbf{v} \in (\mathbb{R}\mathbf{x})^{\perp}$ ,  $\Phi\mathbf{R}\mathbf{v} = \mathbf{G}\mathbf{S}\mathbf{R}\mathbf{v}$  with  $\mathbf{G} = \Phi\mathbf{S}^{\top} \sim \mathcal{N}_{\mathbb{C}}^{m \times (n-1)}(0, 2)$  independent of  $\mathbf{g} := \|\mathbf{x}\|^{-1}\Phi\mathbf{R}\mathbf{x} = \Phi\mathbf{e}_1 \sim \mathcal{N}_{\mathbb{C}}^m(0, 2)$ .

From the independence of  $\mathbf{g}$  and  $\mathbf{G}$ , we can condition  $p_{\mathbf{z}}$  to the value of  $\mathbf{g}$  without altering the distribution of  $\mathbf{G}$ , and eventually computing this probability by expectation over  $\mathbf{g}$  from  $p_{\mathbf{z}}(\Phi\mathbf{R}) = \mathbb{E} p_{\mathbf{z}}(\Phi\mathbf{R}|\mathbf{g})$ . In this context, defining  $\mathbf{z}' := \text{sign}_{\mathbb{C}}(\mathbf{A}\mathbf{R}\mathbf{x})$  and using the properties of  $\mathbf{R}$ , we have  $\|\mathbf{u}^{\perp}\|^2 = \|\mathbf{R}\mathbf{u}^{\perp}\|^2 = \|\mathbf{S}\mathbf{R}\mathbf{u}\|^2$ . Moreover  $\mathbf{H}_{\mathbf{z}'}\mathbf{R}\mathbf{u}^{\perp} = \frac{1}{\sqrt{m}}\mathbf{G}'\mathbf{S}\mathbf{R}\mathbf{u}^{\perp} = \frac{1}{\sqrt{m}}\mathbf{G}'\mathbf{S}\mathbf{R}\mathbf{u}$  with  $\mathbf{G}' := (\mathbf{D}_{\mathbf{z}'}^{\Re}\mathbf{G}^{\Im} - \mathbf{D}_{\mathbf{z}'}^{\Im}\mathbf{G}^{\Re}) \sim \mathcal{N}^{m \times (n-1)}(0, 1)$  since  $(\mathbf{D}_{\mathbf{z}'}^{\Re})^2 + (\mathbf{D}_{\mathbf{z}'}^{\Im})^2 = \mathbf{I}_m$ . Therefore

$$\begin{aligned} p_{\mathbf{z}}(\Phi\mathbf{R}|\mathbf{g}) &= \mathbb{P} \left[ \exists \mathbf{u} \in \mathcal{K}^*, \left| \|\mathbf{H}_{\mathbf{z}'}\mathbf{R}\mathbf{u}^{\perp}\|^2 - \|\mathbf{u}^{\perp}\|^2 \right| > \frac{1}{4}\delta \|\mathbf{u}^{\perp}\|^2 \mid \mathbf{g} \right] \\ &= \mathbb{P} \left[ \exists \mathbf{u} \in \mathcal{K}^*, \left| \frac{1}{m} \|\mathbf{G}'\mathbf{S}\mathbf{R}\mathbf{u}\|^2 - \|\mathbf{S}\mathbf{R}\mathbf{u}\|^2 \right| > \frac{1}{4}\delta \|\mathbf{S}\mathbf{R}\mathbf{u}\|^2 \mid \mathbf{g} \right]. \end{aligned}$$

By homogeneity, since  $\mathcal{K}$  is a cone, we thus get

$$\begin{aligned} p_z(\Phi\mathbf{R}) &= \mathbb{P}[\exists \mathbf{v} \in \mathbf{SR}(\mathcal{K}^*), |\frac{1}{m}\|\mathbf{G}'\mathbf{v}\|^2 - \|\mathbf{v}\|^2| > \frac{1}{4}\delta\|\mathbf{v}\|^2|\mathbf{g}] \\ &= \mathbb{P}[\exists \mathbf{v} \in \mathbf{SR}(\mathcal{K}), |\frac{1}{m}\|\mathbf{G}'\mathbf{v}\|^2 - \|\mathbf{v}\|^2| > \frac{1}{4}\delta\|\mathbf{v}\|^2|\mathbf{g}], \end{aligned}$$

hence showing that  $p_z(\Phi\mathbf{R})$  is the failing probability of  $\frac{1}{\sqrt{m}}\mathbf{G}'$  satisfying the RIP over  $\mathcal{K}' = \mathbf{SR}(\mathcal{K})$  with constant  $\delta/4$ . We can thus invoke Thm. 2.2 and see that (27) holds, *i.e.*,  $p_z(\Phi\mathbf{R}) = \mathbb{E}p_z(\Phi\mathbf{R}|\mathbf{g}) \leq C \exp(-c\delta^2m)$ , provided that

$$m \geq C\delta^{-2}w^2(\mathbf{SR}(\mathcal{K}) \cap \mathbb{S}^{n-2}).$$

For  $\mathcal{K}^\perp := \{\mathbf{u} - \langle \bar{\mathbf{x}}, \mathbf{u} \rangle \bar{\mathbf{x}} : \mathbf{u} \in \mathcal{K}\} \subset \mathcal{K} - \mathbb{R}\bar{\mathbf{x}}$ , the isotropy and monotonicity of the Gaussian mean width [22, Sec. 3.2] involve

$$w(\mathbf{SR}(\mathcal{K}) \cap \mathbb{S}^{n-2}) = w((0 \oplus \mathbf{SR}(\mathcal{K})) \cap \mathbb{S}^{n-1}) = w(\mathbf{R}(\mathcal{K}^\perp) \cap \mathbb{S}^{n-1}) \leq w((\mathcal{K} - \mathbb{R}\bar{\mathbf{x}}) \cap \mathbb{S}^{n-1}),$$

where we define  $0 \oplus \mathcal{S} := \{(0, v) : v \in \mathcal{S}\} \subset \mathbb{R}^n$  for any set  $\mathcal{S} \subset \mathbb{R}^{n-1}$ .

We conclude that, provided  $m \geq C\delta^{-2}w((\mathcal{K} - \mathbb{R}\bar{\mathbf{x}}) \cap \mathbb{S}^{n-1})^2$ , which is verified from (14), the event

$$\mathcal{E}_\perp : \quad \forall \mathbf{u} \in \mathcal{K} \cap \mathbb{S}^{n-1}, \quad \left| \frac{1}{m}\|\mathbf{H}_z\mathbf{u}^\perp\|^2 - \|\mathbf{u}^\perp\|^2 \right| \leq \frac{1}{4}\delta\|\mathbf{u}^\perp\|^2 \quad (28)$$

holds with probability exceeding  $1 - C \exp(-c\delta^2m)$ .

Finally, from (26) and (28), if the joint event  $\mathcal{E} \cap \mathcal{E}_\perp$  holds, which occurs with probability exceeding  $1 - C \exp(-c\delta^2m)$ , then, for all  $\mathbf{u} \in \mathcal{K} \cap \mathbb{S}^{n-1}$ ,

$$\begin{aligned} \|\mathbf{A}_z\mathbf{u}\|^2 &= \frac{1}{\kappa^2m^2}|\langle \text{sign}_{\mathbb{C}}(\Phi\mathbf{x}), \Phi\mathbf{u} \rangle|^2 + \frac{1}{m}\|\mathbf{H}_z\mathbf{u}\|^2 \\ &\leq |\langle \bar{\mathbf{x}}, \mathbf{u} \rangle|^2 + \frac{3}{4}\delta + \|\mathbf{u}^\perp\|^2 + \frac{1}{4}\delta \\ &= \|\mathbf{u}\|^2 + \delta = 1 + \delta. \end{aligned}$$

We find similarly that  $\|\mathbf{A}_z\mathbf{u}\|^2 \geq 1 - \delta$ , which provides the final result.

### 5.3 Proof of Thm 4.3

From Thm 3.3, we already know that, under the condition (17),  $\mathbf{A}_{z_0}$  with  $z_0 = \text{sign}_{\mathbb{C}}(\mathbf{A}\mathbf{x})$  respects the RIP( $\mathcal{K}, \delta$ ) with probability exceeding  $1 - C \exp(-c\delta^2m)$ , which means that the event

$$\mathcal{E}_0 : \quad 1 - \delta \leq \|\mathbf{A}_{z_0}\mathbf{u}\|^2 \leq 1 + \delta, \quad \forall \mathbf{u} \in \mathcal{K},$$

holds with that probability.

Moreover, from Thm 2.2, provided  $m \geq C\delta^2w(\mathcal{K} \cap \mathbb{S}^{n-1})$ , which holds if (17) is verified, and with a failing probability smaller than  $C \exp(-c\delta^2m)$ , this other event is respected:

$$\mathcal{E}_1 : \quad 1 - \delta \leq \frac{1}{m}\|\Phi\mathbf{u}\|^2 \leq 1 + \delta, \quad \forall \mathbf{u} \in \mathcal{K},$$

Let us assume that both events hold, which under the condition (17) happens with probability greater than  $1 - C \exp(-c\delta^2m)$ . Then, from the definition of  $\mathbf{A}_\epsilon$ ,  $\boldsymbol{\alpha}_\epsilon$ , and  $\mathbf{H}_\epsilon$ , we compute that, for any  $\mathbf{u} \in \mathcal{K} \cap \mathbb{S}^{n-1}$ ,

$$\begin{aligned} \|\mathbf{A}_\epsilon\mathbf{u}\| &\leq \frac{1}{\kappa m}|\langle \boldsymbol{\epsilon}, \Phi\mathbf{u} \rangle| + \|\mathbf{H}_\epsilon\mathbf{u}\| = \frac{1}{\kappa m}|\langle \boldsymbol{\epsilon}, \Phi\mathbf{u} \rangle| + \frac{1}{\sqrt{m}}\|\Im(\mathbf{D}_\epsilon^*\Phi\mathbf{u})\| \\ &\leq \frac{1}{\kappa}\left(\frac{1}{\sqrt{m}}\|\boldsymbol{\epsilon}\|\right)\left(\frac{1}{\sqrt{m}}\|\Phi\mathbf{u}\|\right) + \frac{1}{\sqrt{m}}\|\mathbf{D}_\epsilon^*\Phi\mathbf{u}\| \\ &\leq \frac{\tau}{\kappa\sqrt{m}}\|\Phi\mathbf{u}\| + \frac{\tau}{\sqrt{m}}\|\Phi\mathbf{u}\| \leq \tau\sqrt{2}\left(\frac{1}{\kappa} + 1\right) = 2\sqrt{2}\tau. \end{aligned}$$

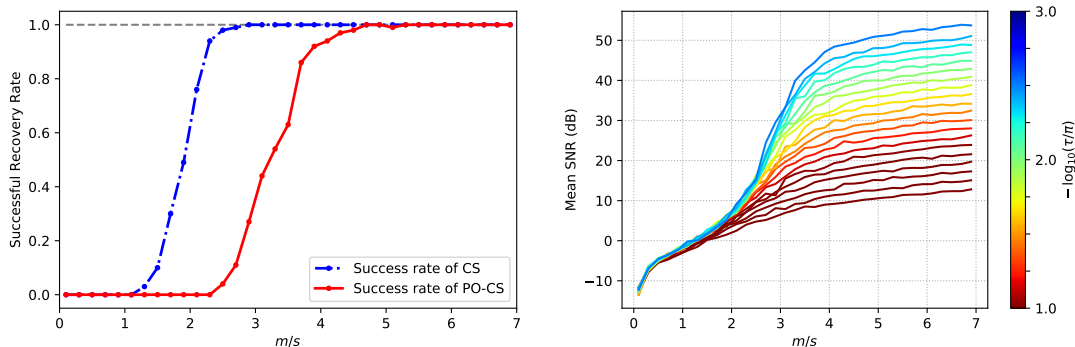


Figure 1: (left) Rate of successful recovery of  $s$ -sparse signals in function of  $m/s$  for PO-CS (in red) under the hypothesis (9), and in linear CS (dashed blue). (right, best viewed in color) Average SNR in function of  $m/s$  for the estimation of the direction of  $s$ -sparse signals in PO-CS. The color coding of the SNR curves corresponds to the noise level  $\tau$ , as measured by  $-\log_{10} \tau/\pi \in [1, 3]$ . All figures were generated by Matplotlib [32].

Therefore, since  $0 < \delta < 1$  and  $\tau < 1/9$ ,

$$\begin{aligned} \|\mathbf{A}_z \mathbf{u}\|^2 &\leq (\|\mathbf{A}_{z_0} \mathbf{u}\| + \|\mathbf{A}_\epsilon \mathbf{u}\|)^2 \leq (\sqrt{1+\delta} + 2\sqrt{2}\tau)^2 \\ &= 1 + \delta + 8\tau^2 + 4\sqrt{2}\tau\sqrt{1+\delta} \leq 1 + \delta + 8\tau(1 + \tau) < 1 + \delta + 9\tau, \end{aligned}$$

and, since  $\sqrt{1-\delta} - 2\sqrt{2}\tau > 0$  if  $0 < \delta + 9\tau < 1$ ,

$$\begin{aligned} \|\mathbf{A}_z \mathbf{u}\|^2 &\geq (\|\mathbf{A}_{z_0} \mathbf{u}\| - \|\mathbf{A}_\epsilon \mathbf{u}\|)^2 \geq (\sqrt{1-\delta} - 2\sqrt{2}\tau)^2 \\ &= 1 - \delta - 4\sqrt{2}\tau\sqrt{1-\delta} \geq 1 - \delta - 4\sqrt{2}\tau > 1 - (\delta + 9\tau), \end{aligned}$$

which finally establishes that  $\mathbf{A}_z$  is  $\text{RIP}(\mathcal{K}, \delta + 9\tau)$ .

## 6 Numerical Experiments

In this section, we test the ability to recover the direction of a sparse vector from the phase of its complex compressive observations.

As a first experiment, we compare in a noiseless scenario the sample complexities of both PO-CS and (linear) CS for sparse signal recovery. We thus study from which number of measurements a sparse vector (or its direction for PO-CS) can be perfectly recovered. The tested experimental conditions are as follows. We have randomly generated ( $s = 10$ )-sparse vectors in  $\mathbb{R}^{(n=100)}$  for a range of measurements  $m \in [1, 70]$ . Each observed sparse vector  $\mathbf{x}_0$  was simply generated by randomly picking a 10-sparse support in the  $\binom{100}{10}$  available supports, and by setting the non-zero vector components i.i.d. as  $\mathcal{N}(0, 1)$ . For each value of  $m$ , we generated the complex Gaussian random matrix as  $\Phi \sim \mathcal{N}_{\mathbb{C}}^{m \times n}(0, 2)$ . As described in Sec. 2, the  $\mathbf{x}_0$  was normalized so that  $\|\Phi \mathbf{x}_0\|_1 = \kappa m$  with  $\kappa = \sqrt{\pi/2}$ . Moreover, for each  $m$ , the signal  $\mathbf{x}_0$  and the matrix  $\Phi$  were randomly regenerated 100 times for reaching valid recovery statistics.

For the reconstruction procedure, we used the basis pursuit denoising (BPDN) program to solve  $\Delta_{\mathcal{K}}(\mathbf{A}_z, \mathbf{e}_1; 0)$ , in comparison with the estimate produced by  $\Delta_{\mathcal{K}}(\mathbf{A}, \mathbf{y}; 0)$

for the CS model  $\mathbf{y} = \Phi \mathbf{x}_0$ . Numerically, BPDN was solved by using the Python implementation of the SPGL1 toolbox<sup>3</sup> [33], together with the Numpy module [34]. In PO-CS and in CS, a reconstruction was considered as successful when the estimate  $\hat{\mathbf{x}}$  reached a relative error  $\|\mathbf{x}_0 - \hat{\mathbf{x}}\|/\|\mathbf{x}_0\| \leq 10^{-3}$  (*i.e.*, a 60 dB SNR).

Fig. 1(left) displays the success rate of both approaches, as measured from the fraction of successful reconstructions over the 100 trials in function of  $m$ . We observe that, as for the linear case, there exists a minimal measurement number from which the signal direction  $\mathbf{x}_0/\|\mathbf{x}_0\|$  is perfectly estimated, roughly from  $m/s \geq 5$ , with a success-failure transition around  $m/s \simeq 3.4$ . This is larger than the perfect reconstruction rate reached by CS from  $m/s \geq 2.5$  with a transition point at  $m/s \simeq 1.8$ . However, this increase is meaningful. The CS reconstruction is achieved from  $m$  complex measurements, equivalent to  $2m$  real observations of  $\mathbf{x}_0$ , while the constraints brought by the model (12) correspond to a measurement model containing  $m + 2$  real observations (by considering the first two rows of  $\mathbf{A}_z$  associated with the normalization constraints). Therefore,  $m$  observations have been lost from the phase-only sensing model. We can thus expect a ratio of (at least) 2 between the sample complexities of PO-CS and CS. This ratio could possibly be reduced by exploiting the unused the last constraint of (11).

In a second experiment, we have measured the robustness of PO-CS to the addition of a noise with bounded complex amplitude, as developed in Sec. 4. We have considered the same experimental setup as above, corrupting the noisy PO-CS model (15) with a uniform additive noise  $\boldsymbol{\epsilon} \in \mathbb{C}^m$  with  $\epsilon_k \sim_{\text{i.i.d.}} \mathcal{U}(\tau\mathcal{B})$  for  $k \in [m]$ ,  $\mathcal{B} = \{\lambda \in \mathbb{C} : |\lambda| \leq 1\}$ . Since Rem. 4.4 shows the existence of an additive noise with level greater than  $\pi$  for which reconstructing the signal direction is impossible, we have set the noise level of this experiment to  $\tau = \pi/10^\alpha$  for  $\alpha$  evenly sampled in  $[1, 3]$ . The reconstruction procedure was performed using the basis pursuit denoising program  $\Delta_{\mathcal{K}}(\mathbf{A}_z, \mathbf{e}_1; \varepsilon)$  for  $\varepsilon$  set<sup>4</sup> to the oracle value  $\varepsilon = \|\mathbf{A}_z \mathbf{x}\|$ , and we solved BPDN with the SPGL1 python toolbox. As above, the results have been averaged over 100 trials for each association of the parameters  $m$  and  $\tau$ , the quantities  $\Phi$ ,  $\mathbf{x}$ , and  $\boldsymbol{\epsilon}$  being randomly regenerated at each trial.

In Fig. 1(right), we plot the evolution of the average signal-to-noise ratio (SNR)  $20 \log_{10} \|\mathbf{x}\|/\|\mathbf{x} - \hat{\mathbf{x}}\|$  (under the normalization hypothesis  $\|\mathbf{A}\mathbf{x}\|_1 = \kappa\sqrt{m}$ ) as a function of  $m/s$ . For each sampled value of  $\tau$ , we display one average SNR curve colored according to the colorbar on the right. As expected, from  $m/s \simeq 3$ , the quality of the BPDN estimate improves, with an SNR level growing linearly with the increase of  $-\log_{10} \tau/\pi$  (*i.e.*, for  $\tau$  decaying logarithmically). This is involved by taking the logarithm of both sides of the instance optimality relation (3) respected by BPDN.

## 7 Conclusion and Perspectives

With this work, we have shown that one can perfectly reconstruct the direction of a low-complexity signal belonging to a symmetric cone from the phase of its complex Gaussian random projections, *i.e.*, in the non-linear PO-CS model (8). Inspired by [3], this is achieved by formulating an equivalent linear sensing model defined from a signal normalization constraint, and a phase-consistency constraint with the observed signal phases in

<sup>3</sup><https://github.com/drrelyea/spgl1>

<sup>4</sup>We postpone to a future work the definition of an accurate estimator of this quantity.



the random projection domain. This reformulation allows us to perform this reconstruction with any instance optimal algorithm provided that the associated sensing matrix respects the RIP property over the considered low-complexity signal set. In addition, we inherit the stability and the robustness from such algorithms; the PO-CS model can be corrupted by a bounded noise with small noise level, and the observed signal is only required to be well approximated by an element of a low-complexity set (with small modeling error). We proved that a complex Gaussian random sensing matrix fulfills w.h.p. the conditions allowing such a reconstruction; it leads to an equivalent sensing matrix (in the reformulated model) that respects w.h.p. the restricted isometry property provided that the number of measurements is large compared to the complexity of the signal set, with an additional dependence to the noise level for noisy PO-CS.

We can think of (at least) two open questions that are worth being investigated in future work. First, all the reconstruction guarantees involving the considered random matrices are non-uniform; they are valid w.h.p. given the observed signal. However, up to an increase of the exponent of  $1/\delta$  in the sample-complexity condition (14), and in the case where the signal belongs to a low-complexity set  $\mathcal{K}$  (zero modeling error), it should be possible to use the sign-product embedding [7] and its extension to the complex field [8] in order to ensure that, w.h.p.,  $\mathbf{A}_z$  (with  $z = \text{sign}_{\mathbb{C}}(\mathbf{A}\mathbf{x})$ ) is RIP for all  $\mathbf{x} \in \mathcal{K}$ .

Second, in several imaging applications (such as radar, magnetic resonance imaging, computed tomography), the complex sensing matrix representing the observational model is structured and amounts to (randomly) sub-sampling the rows of a Fourier matrix. In these contexts, phase-only compressive sensing provides an appealing sensing alternative, for instance, making the measurements insensitive to large amplitude variations and easing the measurement quantization. In correspondence with the inspirational work of Oppenheim [2, 1], a critical open question is thus to extend the RIP stated in Thm 3.3 and Thm 4.3 to partial random Fourier matrices or to other structured random sensing constructions. A first breakthrough in this direction would be to verify this extension for partial Gaussian circulant matrices, which satisfy w.h.p. the  $(\ell_1, \ell_2)$ -RIP and are applicable to one-bit CS [37].

## References

- [1] A. V. Oppenheim, M. H. Hayes, and J. S. Lim, "Iterative procedures for signal reconstruction from fourier transform phase," *Optical Engineering*, vol. 21, no. 1, p. 211122, feb 1982.
- [2] A. Oppenheim and J. Lim, "The importance of phase in signals," *Proceedings of the IEEE*, vol. 69, no. 5, pp. 529–541, 1981.
- [3] P. T. Boufounos, "Sparse signal reconstruction from phase-only measurements." Citeseer, 2013.
- [4] P. T. Boufounos and R. G. Baraniuk, "1-bit compressive sensing," in *2008 42nd Annual Conference on Information Sciences and Systems*. IEEE, 2008, pp. 16–21.
- [5] 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 Transactions on Information Theory*, vol. 59, no. 4, pp. 2082–2102, 2013.
- [6] Y. Plan and R. Vershynin, "Robust 1-bit compressed sensing and sparse logistic regression: A convex programming approach," *IEEE Transactions on Information Theory*, vol. 59, no. 1, pp. 482–494, 2012.

- [7] S. Foucart, “Flavors of compressive sensing,” in *International Conference Approximation Theory*. Springer, 2016, pp. 61–104.
- [8] T. Feuillen, M. E. Davies, L. Vandendorpe, and L. Jacques, “ $(\ell_1, \ell_2)$ -RIP and Projected Back-Projection Reconstruction for Phase-Only Measurements,” *preprint*, 2019.
- [9] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, “Model-based compressive sensing,” *IEEE Transactions on information theory*, vol. 56, no. 4, pp. 1982–2001, 2010.
- [10] U. Ayaz, S. Dirksen, and H. Rauhut, “Uniform recovery of fusion frame structured sparse signals,” *Applied and Computational Harmonic Analysis*, vol. 41, no. 2, pp. 341–361, sep 2016.
- [11] M. Fazel, “Matrix rank minimization with applications,” Ph.D. dissertation, Stanford University, 2002.
- [12] E. J. Candès and T. Tao, “Decoding by linear programming,” vol. 51, no. 12, p. 4203, 2005.
- [13] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*. Springer New York, 2013.
- [14] Y. Traonmilin and R. Gribonval, “Stable recovery of low-dimensional cones in hilbert spaces: One RIP to rule them all,” *Applied and Computational Harmonic Analysis*, vol. 45, no. 1, pp. 170–205, jul 2018.
- [15] L. Jacques, K. Degraux, and C. De Vleeschouwer, “Quantized iterative hard thresholding: Bridging 1-bit and high-resolution quantized compressed sensing,” *arXiv preprint arXiv:1305.1786*, 2013.
- [16] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, “A simple proof of the restricted isometry property for random matrices,” *Constructive Approximation*, vol. 28, no. 3, pp. 253–263, 2008.
- [17] T. T. Cai and A. Zhang, “Sparse representation of a polytope and recovery of sparse signals and low-rank matrices,” *IEEE Transactions on Information Theory*, vol. 60, no. 1, pp. 122–132, 2013.
- [18] S. S. Chen, D. L. Donoho, and M. A. Saunders, “Atomic decomposition by basis pursuit,” vol. 43, no. 1, pp. 129–159, 2001.
- [19] J. A. Tropp and A. C. Gilbert, “Signal recovery from random measurements via orthogonal matching pursuit,” *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, dec 2007.
- [20] T. Blumensath and M. E. Davies, “Iterative hard thresholding for compressed sensing,” *Applied and computational harmonic analysis*, vol. 27, no. 3, pp. 265–274, 2009.
- [21] D. Needell and J. A. Tropp, “Cosamp: Iterative signal recovery from incomplete and inaccurate samples,” *Applied and computational harmonic analysis*, vol. 26, no. 3, pp. 301–321, 2009.
- [22] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky, “The convex geometry of linear inverse problems,” *Foundations of Computational mathematics*, vol. 12, no. 6, pp. 805–849, 2012.
- [23] E. J. Candès and Y. Plan, “Tight oracle inequalities for low-rank matrix recovery from a minimal number of noisy random measurements,” *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 2342–2359, 2011.
- [24] L. Jacques and V. Cambareri, “Time for dithering: fast and quantized random embeddings via the restricted isometry property,” *Information and Inference: A Journal of the IMA*, vol. 6, no. 4, pp. 441–476, apr 2017.
- [25] M. Fazel, E. Candès, B. Recht, and P. Parrilo, “Compressed sensing and robust recovery of low rank matrices,” in *2008 42nd Asilomar Conference on Signals, Systems and Computers*. IEEE, 2008, pp. 1043–1047.
- [26] P. Jain, R. Meka, and I. S. Dhillon, “Guaranteed rank minimization via singular value projection,” in *Advances in Neural Information Processing Systems*, 2010, pp. 937–945.
- [27] J. D. Blanchard, J. Tanner, and K. Wei, “CGIHT: conjugate gradient iterative hard thresholding for compressed sensing and matrix completion,” *Information and Inference*, p. iav011, nov 2015.

- [28] S. Mendelson, A. Pajor, and N. Tomczak-Jaegermann, “Uniform uncertainty principle for bernoulli and subgaussian ensembles,” *Constructive Approximation*, vol. 28, no. 3, pp. 277–289, feb 2008.
- [29] R. Vershynin, “Introduction to the non-asymptotic analysis of random matrices,” in *Compressed Sensing*, Y. C. Eldar and G. Kutyniok, Eds. Cambridge University Press, 2012, pp. 210–268.
- [30] M. Ledoux and M. Talagrand, *Probability in Banach Spaces: isoperimetry and processes*. Springer Science & Business Media, 2013.
- [31] A. Papoulis and S. U. Pillai, *Probability, random variables, and stochastic processes*. Tata McGraw-Hill Education, 2002.
- [32] J. D. Hunter, “Matplotlib: A 2d graphics environment,” *Computing in Science & Engineering*, vol. 9, no. 3, pp. 90–95, 2007.
- [33] E. van den Berg and M. P. Friedlander, “Probing the pareto frontier for basis pursuit solutions,” *SIAM Journal on Scientific Computing*, vol. 31, no. 2, pp. 890–912, jan 2009.
- [34] S. Van Der Walt, S. C. Colbert, and G. Varoquaux, “The numpy array: a structure for efficient numerical computation,” *Computing in Science & Engineering*, vol. 13, no. 2, p. 22, 2011.
- [35] F. Pfeiffer, C. Kottler, O. Bunk, and C. David, “Hard x-ray phase tomography with low-brilliance sources,” *Physical Review Letters*, vol. 98, no. 10, mar 2007.
- [36] A. González, L. Jacques, C. D. Vleeschouwer, and P. Antoine, “Compressive optical deflectometric tomography: A constrained total-variation minimization approach,” *Inverse Problems and Imaging*, vol. 8, no. 2, pp. 421–457, may 2014.
- [37] S. Dirksen, H. C. Jung, and H. Rauhut, “One-bit compressed sensing with partial gaussian circulant matrices,” *Information and Inference: A Journal of the IMA*, oct 2019.