

Phase Retrieval Meets Statistical Learning Theory: A Flexible Convex Relaxation

Sohail Bahmani

Justin Romberg

School of ECE
Georgia Institute of Technology
{sohail.bahmani, jrom}@ece.gatech.edu

Abstract

We propose a flexible convex relaxation for the phase retrieval problem that operates in the natural domain of the signal. Therefore, we avoid the prohibitive computational cost associated with “lifting” and semidefinite programming (SDP) in methods such as *PhaseLift* and compete with recently developed non-convex techniques for phase retrieval. We relax the quadratic equations for phaseless measurements to inequality constraints each of which representing a symmetric “slab”. Through a simple convex program, our proposed estimator finds an extreme point of the intersection of these slabs that is best aligned with a given *anchor vector*. We characterize geometric conditions that certify success of the proposed estimator. Furthermore, using classic results in statistical learning theory, we show that for random measurements the geometric certificates hold with high probability at an optimal sample complexity. Phase transition of our estimator is evaluated through simulations. Our numerical experiments also suggest that the proposed method can solve phase retrieval problems with coded diffraction measurements as well.

1 Introduction

Let $\mathbf{x}_\star \in \mathbb{C}^N$ be a signal that we would like to recover from noisy phaseless measurements

$$b_i = |\mathbf{a}_i^* \mathbf{x}_\star|^2 + \xi_i \quad i = 1, 2, \dots, M, \quad (1)$$

Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS) 2017, Fort Lauderdale, Florida, USA. JMLR: W&CP volume 54. Copyright 2017 by the author(s).

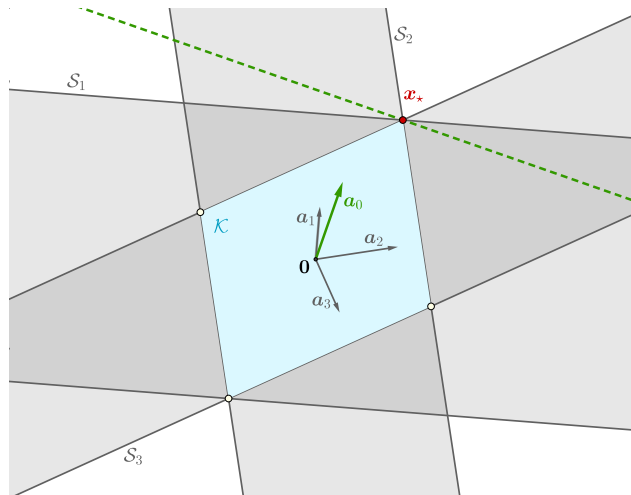


Figure 1: A two-dimensional illustration of slabs intersecting at \mathbf{x}_\star

where the measurement vectors $\mathbf{a}_i \in \mathbb{C}^N$ are given. To solve this *phase retrieval* problem with provable accuracy, different methods that rely on semidefinite relaxation have been proposed previously (e.g., Candès et al., 2013; Candès and Li, 2014; Waldspurger et al., 2015). While these methods are guaranteed to produce an accurate solution in polynomial time, they are not scalable due to the use of semidefinite programming (SDP). This drawback of SDP-based methods has motivated development of alternative non-convex methods that operate in the natural domain of the signal and exhibit better scalability (e.g., Netrapalli et al., 2013; Candès et al., 2015b). In this paper we follow a completely different approach and propose a convex relaxation of the phase retrieval problem that not only produces accurate solutions but also is scalable. Compared to the non-convex phase retrieval methods our approach inherits the flexibility of convex optimization both in analysis and application.

The geometric idea at the core of our proposed method is the following. Relaxing each measurement equation

in (1) to an inequality $|\mathbf{a}_i^* \mathbf{x}_*|^2 \leq b_i$ creates a symmetric slab \mathcal{S}_i of feasible solutions as illustrated in Figure 1. Collectively, these slabs describe a “complex polytope” \mathcal{K} of feasible solutions. In the noiseless regime (i.e., $\xi_i = 0$ for all i), the target signal \mathbf{x}_* would be one of the extreme points of \mathcal{K} . To distinguish \mathbf{x}_* among all of the extreme points, our idea is to find a hyperplane tangent to \mathcal{K} at \mathbf{x}_* . The crucial ingredient in this approach is an “anchor” vector $\mathbf{a}_0 \in \mathbb{C}^N \setminus \{\mathbf{0}\}$ that acts as the normal for the desired tangent hyperplane and it is required to have a non-vanishing correlation with \mathbf{x}_* in the sense that

$$|\mathbf{a}_0^* \mathbf{x}_*| \geq \delta \|\mathbf{a}_0\|_2 \|\mathbf{x}_*\|_2, \quad (2)$$

for some absolute constant $\delta \in (0, 1)$. The above geometric intuition is explained in more detail in Section 3.1. While our main result simply assumes that the anchor vector is given by an oracle, which may use the existing measurements, we discuss in Section 1.1 some realistic scenarios where a valid anchor vector exists or can be computed.

We assume that the noise is non-negative (i.e., $\xi_i \geq 0$) and we have $\|\boldsymbol{\xi}\|_\infty \leq \eta^{-1} \|\mathbf{x}_*\|_2^2$ for some constant $\eta > 0$. Note that the non-negativity of the noise can be dropped at the cost of a slight reduction in the effective signal-to-noise ratio. In particular, one can add the noise upperbound (i.e., $\eta^{-1} \|\mathbf{x}_*\|_2^2$) to each measurement to ensure the non-negativity. Throughout we treat \mathbb{C}^N as an inner-product space over \mathbb{R} equipped with the symmetric inner-product

$$\langle \cdot, \cdot \rangle : (\mathbf{x}_1, \mathbf{x}_2) \mapsto \operatorname{Re}(\mathbf{x}_1^* \mathbf{x}_2).$$

Clearly, in this setting \mathbb{C}^N will be a $2N$ -dimensional vector space.

With these assumptions in place, we propose the solution to the convex program¹

$$\begin{aligned} & \max_{\mathbf{x}} \langle \mathbf{a}_0, \mathbf{x} \rangle \\ & \text{subject to } |\mathbf{a}_i^* \mathbf{x}|^2 \leq b_i \quad 1 \leq i \leq M, \end{aligned} \quad (3)$$

as a computationally efficient estimator for \mathbf{x}_* . Of course, the points equal to \mathbf{x} up to a global phase, namely,

$$\mathbb{T}\mathbf{x} \stackrel{\text{def}}{=} \{\omega \mathbf{x} : |\omega| = 1\},$$

yield the same phaseless measurements. Therefore, the goal is merely to estimate a point in $\mathbb{T}\mathbf{x}_*$ accurately from the phaseless measurements (1).

In Lemma 2, below in Section 3, we establish a geometric condition that is sufficient to guarantee accurate estimation of \mathbf{x}_* via the convex program (3).

The sufficient condition given by Lemma 2 can be interpreted in terms of (non-)existence of a particularly constrained halfspace that includes all of the points $\mathbf{a}_i \mathbf{a}_i^* \mathbf{x}_*$. For random measurement vectors \mathbf{a}_i , this interpretation resembles the model and theory of *linear classifiers* studied in *statistical learning theory*, albeit in an unusual regime. Borrowing classic results from this area (summarized in Appendix A), we show that with high probability (3) produces an accurate estimate of \mathbf{x}_* .

Specifically, in our main result, Theorem 1 in Section 3, we show that drawing

$$M \gtrsim N + \log \frac{1}{\varepsilon},$$

i.i.d. random measurements, with the hidden constant factor on the right-hand side depending on δ , would suffice for the conditions of Lemma 2 to hold with probability $\geq 1 - \varepsilon$. Consequently, solution $\hat{\mathbf{x}}$ of (3) would obey

$$\|\hat{\mathbf{x}} - \mathbf{x}_*\|_2 \lesssim \eta^{-1} \|\mathbf{x}_*\|_2.$$

1.1 Choosing the anchor vector

Our approach critically depends on the choice of the anchor vector \mathbf{a}_0 that obeys an inequality of the form (2). Below we discuss two interesting scenarios where such a vector would be accessible.

Non-negative signals: Perhaps the simplest scenario is when the target signal \mathbf{x}_* is known to be real and non-negative. In usual imaging modalities these model assumptions are realistic as natural images are typically represented by pixel intensities. For these types of signals we can choose $\mathbf{a}_0 = \frac{1}{\sqrt{N}} \mathbf{1}$ for which we obtain $|\mathbf{a}_0^* \mathbf{x}_*| = \|\mathbf{x}_*\|_1 / \sqrt{N}$. Then, for (2) to hold it suffices that $\|\mathbf{x}_*\|_1 \geq \delta \sqrt{N} \|\mathbf{x}_*\|_2$ for some absolute constant $\delta \in (0, 1)$. In particular, we need \mathbf{x}_* to have at least $\delta^2 N$ non-zero entries.

Random measurements: A more interesting scenario is when we can construct the vector \mathbf{a}_0 from the (random) measurements. An effective strategy is to set \mathbf{a}_0 to be the principal eigenvector of the matrix $\boldsymbol{\Sigma} = \frac{1}{M} \sum_{i=1}^M b_i \mathbf{a}_i \mathbf{a}_i^*$. The principal eigenvector of $\boldsymbol{\Sigma}$ and its “truncated” variants have been used previously for initialization of the *Wirtinger Flow* algorithm (Candès et al., 2015b) and its refined versions (Chen and Candès, 2015; Zhang et al., 2016). For example, the following result is shown in Candès et al. (2015b, Section VII.H).

Lemma 1 (Candès et al. (2015b)). *For $1 \leq i \leq M$ let b_i be the phaseless measurements obtained from*

¹In the real case, (3) reduces to a linear program.

i.i.d. vectors $\mathbf{a}_i \sim \text{Normal}(\mathbf{0}, \frac{1}{2}\mathbf{I}) + \nu\text{Normal}(\mathbf{0}, \frac{1}{2}\mathbf{I})$ and no noise. If $M \gtrsim N \log N$ and \mathbf{a}_0 is the principal eigenvector of

$$\Sigma = \frac{1}{M} \sum_{i=1}^M b_i \mathbf{a}_i \mathbf{a}_i^*,$$

then (2) holds with probability $\geq 1 - O(N^{-2})$.

While Lemma 1 can be refined or extended in various ways, we do not pursue these paths in this paper.

1.2 Related work

There is a large body of research on phase retrieval addressing various aspects of the problem (see (Jaganathan et al., 2015) and references therein). However, we focus only on the relevant results mostly developed in recent years. Perhaps, among the most important developments are *PhaseLift* and similar methods that cast the phase retrieval problem as a particular semidefinite program (Candès et al., 2013; Candès and Li, 2014; Waldspurger et al., 2015). The main idea used by Candès et al. (2013) and Candès and Li (2014) is that by *lifting* the unknown signal using the transformation $\mathbf{x}\mathbf{x}^* \mapsto \mathbf{X}$, the (noisy) phaseless measurements (1) that are quadratic in \mathbf{x}_* can be converted to linear measurements of the rank-one positive semidefinite matrix $\mathbf{X}_* = \mathbf{x}_* \mathbf{x}_*^*$. With this observation, these SDP-based methods aim to solve the corresponding linear equations using the trace-norm to induce the rank-one structure in the solution. Inspired by the well-known convex relaxation of Max-Cut problem, *PhaseCut* method (Waldspurger et al., 2015) considers the measurement phases as the unknown variables and applies a similar lifting transform to formulate a different semidefinite relaxation for phase retrieval. While these SDP-based methods are shown to produce accurate estimates of \mathbf{X}_* at optimal sample complexity for certain random measurement models, they become computationally prohibitive in medium- to large-scale problems where SDP is practically inefficient.

More recently, there has been a growing interest in non-convex iterative methods for phase retrieval (see e.g., Netrapalli et al., 2013; Candès et al., 2015b; Schniter and Rangan, 2015; Chen and Candès, 2015; Zhang et al., 2016; Wang and Giannakis, 2016; Sun et al., 2016). These methods typically operate in the natural space of the signal and thus do not suffer the drawbacks of the SDP-based methods. With a specific initialization Netrapalli et al. (2013) establish some accuracy guarantees for a variant of the classic methods by Gerchberg and Saxton (1972); Fienup (1982) that iteratively update the estimate assuming the measurements' phase match that of the previous iterate.

The established sample complexity is (nearly) optimal in the dimension of the target signal, but it does not vary gracefully with the prescribed precision. Phase retrieval via the *Wirtinger Flow* (WF), a non-convex gradient descent method at core, is proposed by Candès et al. (2015b). It is shown that for random measurements that have Normal distribution or certain *coded diffraction patterns*, with an appropriate initialization the WF iterates exhibit the linear rate of convergence to the target signal. More recent work on the WF methods introduce better initialization by excluding the outlier measurements and achieve the optimal sample complexity (Chen and Candès, 2015; Zhang et al., 2016). The WF class of algorithms and our proposed method both achieve optimal sample complexity (up to the constant factor) and have low computational cost. However, the WF methods need careful tuning of a step size parameter and their convergence analysis often relies on Gaussian measurements. This is partly because establishing robustness of non-convex methods generally requires stronger conditions. Our method provably works for a broader set of measurement distributions, has no tuning parameters, and can be implemented in various convex optimization software.

Shortly after a draft of this manuscript was first posted online, a few independent papers proposed and analyzed the same method and its variants. Goldstein and Studer (2016), who dubbed (3) *PhaseMax*, obtained sharper constants in the sample complexity by assuming a stronger condition that the anchor is independent of the measurements. Alternative proofs and variations that rely on matrix concentration inequalities appeared later in (Hand and Voroninski, 2016c,b,a). Another distinctive feature of our analysis compared to the mentioned results is that it is less sensitive to measurement distribution as it relies on VC-type bounds.

1.3 Variations and Extensions

In this section we discuss several different ways to extend the proposed method that we leave for future research. While the core geometric idea still applies, some modifications of our theoretical arguments would be necessary to analyze these extensions.

The gross noise model considered in this paper can be pessimistic in scenarios where we have random noise or deterministic noise with a different type of bound. In these scenarios, augmenting the estimator by a noise regularization term could result in accuracy bounds that gracefully vary with the considered noise.

Another interesting extension to the proposed method, is to adapt the current theory to the case of block-wise independent measurements as in coded diffraction imaging. Our numerical experiments in Section (2)

suggest that the proposed method still performs well with these structured measurements. Nevertheless, to extend the analysis we may need to revise the current simple arguments based on Vapnik-Chervonenkis theory using more sophisticated tools from the theory of empirical processes.

Finally, we believe that our proposed method is flexible in the sense that it allows to incorporate a structural properties of the signal relatively easily. In particular, it would be interesting to analyze a variant of the proposed estimator that induces sparsity through ℓ_1 -norm regularization.

2 Numerical experiments

We evaluated the performance of our proposed method on synthetic data with the target signal $\mathbf{x}_* \sim \text{Normal}(\mathbf{0}, \frac{1}{2}\mathbf{I}) + \nu \text{Normal}(\mathbf{0}, \frac{1}{2}\mathbf{I})$ and measurements $\mathbf{a}_i \stackrel{\text{i.i.d.}}{\sim} \text{Normal}(\mathbf{0}, \frac{1}{2}\mathbf{I}) + \nu \text{Normal}(\mathbf{0}, \frac{1}{2}\mathbf{I})$ all having $N = 500$ coordinates. The noisy measurements follow (1) with the uniform noise $\xi_i \stackrel{\text{i.i.d.}}{\sim} \text{Uniform}([0, \eta^{-1}])$ in one experiment and the Gaussian noise $\xi_i \stackrel{\text{i.i.d.}}{\sim} \text{Normal}(0, \sigma^2)$ in the other. For the latter noise model we replaced any negative b_i by $b_i = 0$ to avoid negative measurements and defined the input signal-to-noise ratio as $\text{SNR} \stackrel{\text{def}}{=} 10 \log_{10} \frac{\|\mathbf{x}_*\|_2^4}{\sigma^2}$. The vector \mathbf{a}_0 is constructed as in initialization of the Wirtinger Flow mentioned in Lemma 1 through 50 iterations of the power method. We implemented the convex program (3) by TFOCS (Becker et al., 2011) with smoothing parameter $\mu = 2 \times 10^{-3}$ and at most 500 iterations. Figure 2 illustrates the 0.9-quantile and median of the relative error $\min_{\phi \in [0, 2\pi]} \|\hat{\mathbf{x}} - e^{i\phi} \mathbf{x}_*\|_2 / \|\mathbf{x}_*\|_2$ observed over 100 trials of our algorithm for different sampling ratios $\frac{M}{N}$ between 2 and 17. The plots also show the effect of different levels of noise on the relative error for both of the considered noise models.

We also evaluated our method using noiseless measurements with coded diffraction patterns as described by Candès et al. (2015a). Specifically, with indices $i = (k, \ell)$ for $1 \leq k \leq N$ and $1 \leq \ell \leq L$, we used measurements of the form $\mathbf{a}_i = \mathbf{f}_k \circ \phi_\ell$ which is the pointwise product of the k -th discrete Fourier basis vector (i.e., \mathbf{f}_k) and a random modulation pattern with i.i.d. symmetric Bernoulli entries (i.e., ϕ_ℓ). The target signal is an $N = 960 \times 1280 \approx 1.2 \times 10^6$ pixel image of a Persian Leopard.² We used $L = 20$ independent coded diffraction patterns $\{\phi_\ell\}_{1 \leq \ell \leq L}$. Therefore, the total number of (scalar) measurements is $M = LN \approx 2.5 \times 10^7$. Similar to the first simulation, the vector \mathbf{a}_0 is constructed as the (approximate) prin-

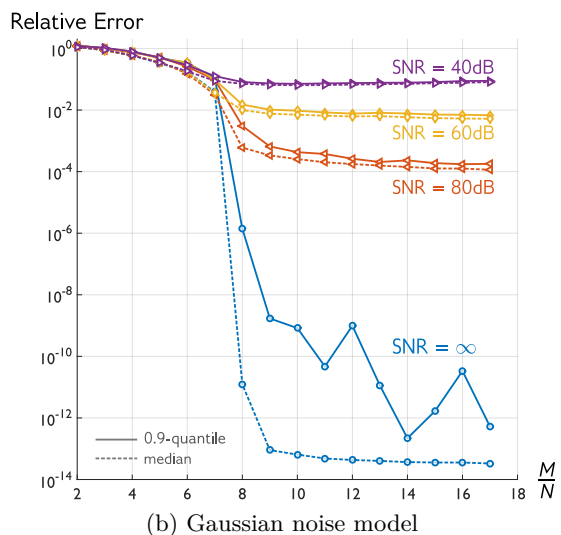
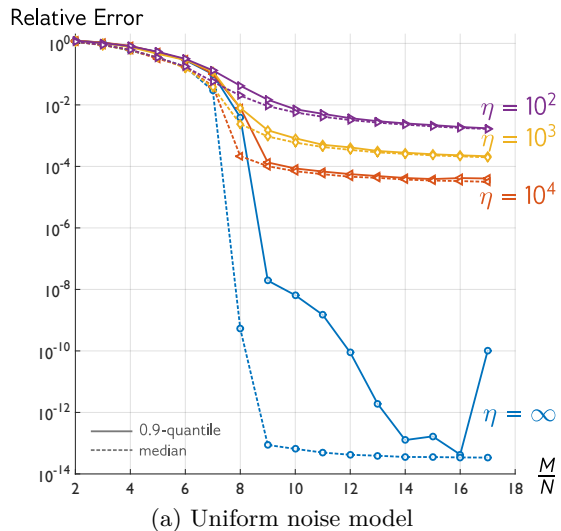


Figure 2: Empirical relative error of the proposed method at different sampling ratios (i.e., $\frac{M}{N}$) and noise levels with $N = 500$

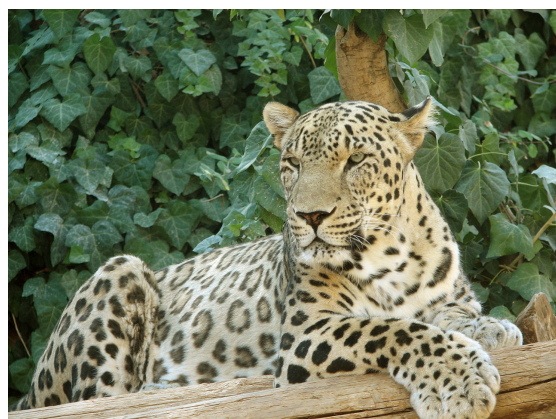


Figure 3: Persian Leopard at 960×1280 resolution. Relative error is 8.2×10^{-8} .

²Available online at:

https://upload.wikimedia.org/wikipedia/commons/thumb/7/7d/Persian_Leopard_sitting.jpg/1280px-Persian_Leopard_sitting.jpg

cipal eigenvector of $\frac{1}{M} \sum_i b_i \mathbf{a}_i \mathbf{a}_i^*$ through 50 iterations of the power method. The convex program is also solved using TFOCS, but this time with smoothing parameter $\mu = 10^{-6}$ and restricting the total number of forward and adjoint coded diffraction operator to 500. The recovered image is depicted in Figure 3 which has a relative error of about 8.2×10^{-8} .

3 Theoretical Analysis

In this section we provide the precise statement of the our results and their proofs. For the sake of simplicity in notation and derivation, but without loss of generality, we make the following assumptions. We assume that $\alpha_0^* \mathbf{x}_*$ is a positive real number since any point in $\mathbb{T} \mathbf{x}_*$ is a valid target. Furthermore, we assume that \mathbf{x}_* is unit-norm (i.e., $\|\mathbf{x}_*\|_2 = 1$) and thus the bound on the noise reduces to $\|\boldsymbol{\xi}\|_\infty \leq \eta^{-1}$. We first establish, in Lemma 2, a geometric condition for success of phase retrieval through (3). Then we use this lemma to prove our main result for random measurements in Theorem 1. We also rely on tools from statistical learning theory that are outlined in Appendix A.

3.1 Geometry of intersecting slabs

To understand the geometry of (3) it is worthwhile to first consider the noiseless scenario. The feasible set is the intersection of the sets

$$\mathcal{S}_i = \left\{ \mathbf{x} \in \mathbb{C}^N : |\mathbf{a}_i^* \mathbf{x}|^2 \leq b_i \right\}$$

corresponding to the pairs (\mathbf{a}_i, b_i) for $i = 1, 2, \dots, M$. The sets \mathcal{S}_i are effectively symmetric ‘‘complex slabs’’. Denote their intersection by

$$\mathcal{K} = \bigcap_{i=1}^M \mathcal{S}_i.$$

In (3) the objective function is linear, thus its maximizer is an extreme point of the convex constraint set \mathcal{K} . Clearly, \mathbf{x}_* as well as any other point in $\mathbb{T} \mathbf{x}_*$ are extreme points of \mathcal{K} . However, \mathcal{K} typically has other extreme points that are not equivalent to \mathbf{x}_* . Intuitively, using the non-vanishing correlation of \mathbf{a}_0 with \mathbf{x}_* , the convex program (3) is effectively eliminating the superfluous extreme points of \mathcal{K} . The geometric interpretation is that the hyperplane normal to \mathbf{a}_0 that passes through \mathbf{x}_* is also tangent to \mathcal{K} , as Figure 1 suggests. It is not difficult to show that an analogous interpretation from the dual point of view is that \mathbf{a}_0 is in the interior of the conical hull cone $\{\mathbf{a}_i \mathbf{a}_i^* \mathbf{x}_*\}_{1 \leq i \leq N}$.

More generally, with noisy measurements, \mathcal{K} is still a symmetric complex polytope that is convex and includes $\mathbb{T} \mathbf{x}_*$ due to non-negativity of the noise. We

would like to find conditions that guarantee that the solution to (3) is close to \mathbf{x}_* . More specifically, we would like to show that if $\hat{\mathbf{x}} = \mathbf{x}_* + \mathbf{h}$ is any solution to (3) and $t > 0$ is some constant, then with $\|\mathbf{h}\|_2 > (t\eta)^{-1}$ the inequalities

$$\begin{aligned} \langle \mathbf{a}_0, \mathbf{h} \rangle &\geq 0 \\ |\mathbf{a}_i^* (\mathbf{x}_* + \mathbf{h})|^2 &\leq |\mathbf{a}_i^* \mathbf{x}_*|^2 + \xi_i \quad 1 \leq i \leq M, \end{aligned}$$

cannot hold simultaneously. The following lemma provides the desired sufficient condition.

Lemma 2. *Let*

$$\mathcal{R}_\delta = \left\{ \mathbf{h} \in \mathbb{C}^N : \|\mathbf{h} - (\mathbf{x}_*^* \mathbf{h}) \mathbf{x}_*\|_2 \geq \delta |\operatorname{Im}(\mathbf{x}_*^* \mathbf{h})| \right\}, \quad (4)$$

and $\epsilon \geq 0$ be some constant. If every vector $\mathbf{h} \in \mathcal{R}_\delta$ with $\|\mathbf{h}\|_2 > \epsilon$ violates at least one of the inequalities

$$\begin{aligned} \langle \mathbf{a}_0, \mathbf{h} \rangle &\geq 0 \\ \langle \mathbf{a}_i \mathbf{a}_i^* \mathbf{x}_*, \mathbf{h} \rangle &\leq \frac{1}{2} \eta^{-1} \quad 1 \leq i \leq M, \end{aligned} \quad (5)$$

then any solution $\hat{\mathbf{x}}$ to (3) obeys

$$\|\hat{\mathbf{x}} - \mathbf{x}_*\|_2 \leq \epsilon.$$

Proof. It suffices to show that $\mathbf{h} = \hat{\mathbf{x}} - \mathbf{x}_*$ obeys (5) and it belongs to \mathcal{R}_δ . Given that

$$\begin{aligned} \xi_i &\geq |\mathbf{a}_i^* (\mathbf{x}_* + \mathbf{h})|^2 - |\mathbf{a}_i^* \mathbf{x}_*|^2 = |\mathbf{a}_i^* \mathbf{h}|^2 + 2\langle \mathbf{a}_i \mathbf{a}_i^* \mathbf{x}_*, \mathbf{h} \rangle \\ &\geq 2\langle \mathbf{a}_i \mathbf{a}_i^* \mathbf{x}_*, \mathbf{h} \rangle \end{aligned}$$

and $\xi_i \leq \eta^{-1}$, we have $\langle \mathbf{a}_i \mathbf{a}_i^* \mathbf{x}_*, \mathbf{h} \rangle \leq \frac{1}{2} \eta^{-1}$. Feasibility of \mathbf{x}_* also guarantees that $\langle \mathbf{a}_0, \mathbf{h} \rangle \geq 0$. Therefore, we have shown that \mathbf{h} satisfies (5).

The constraints of (3) are invariant under a global change of phase (i.e., the action of \mathbb{T}). It easily follows that the solution $\hat{\mathbf{x}}$ to (3) should obey $\operatorname{Im}(\mathbf{a}_0^* \hat{\mathbf{x}}) = 0$. Therefore, we have $\operatorname{Im}(\mathbf{a}_0^* \mathbf{h}) = 0$ as we assumed $\alpha = \alpha_0^* \mathbf{x}_* \in \mathbb{R}$. The same assumption also implies that $\mathbf{a}_0 = \alpha \mathbf{x}_* + \mathbf{a}_{0\perp}$ for $\mathbf{a}_{0\perp} = (\mathbf{I} - \mathbf{x}_* \mathbf{x}_*^*) \mathbf{a}_0$ which clearly obeys $\mathbf{x}_*^* \mathbf{a}_{0\perp} = 0$. Thus, using triangle inequality and the bound (2) we obtain

$$\begin{aligned} 0 = |\operatorname{Im}(\mathbf{a}_0^* \mathbf{h})| &= |\alpha \operatorname{Im}(\mathbf{x}_*^* \mathbf{h}) + \operatorname{Im}(\mathbf{a}_{0\perp}^* \mathbf{h})| \\ &\geq \alpha |\operatorname{Im}(\mathbf{x}_*^* \mathbf{h})| - |\operatorname{Im}(\mathbf{a}_{0\perp}^* \mathbf{h})| \\ &\geq \delta \|\mathbf{a}_0\|_2 |\operatorname{Im}(\mathbf{x}_*^* \mathbf{h})| - \|\mathbf{h}_\perp\|_2 \|\mathbf{a}_0\|_2, \end{aligned}$$

where $\mathbf{h}_\perp = (\mathbf{I} - \mathbf{x}_* \mathbf{x}_*^*) \mathbf{h} = \mathbf{h} - (\mathbf{x}_*^* \mathbf{h}) \mathbf{x}_*$. The above inequality completes the proof as it is equivalent to $\mathbf{h} \in \mathcal{R}_\delta$. \square

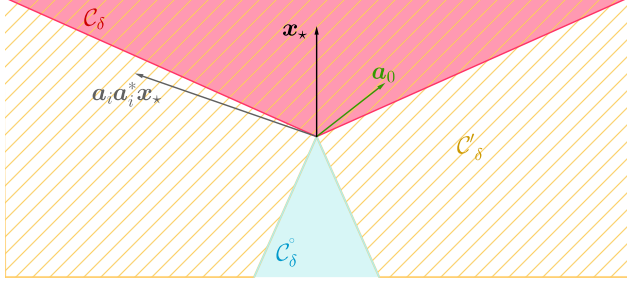


Figure 4: Relative positioning of \mathbf{a}_0 and $\mathbf{a}_i \mathbf{a}_i^* \mathbf{x}_*$ with respect to the cones \mathcal{C}_δ and \mathcal{C}'_δ

3.2 Guarantees for random measurements

In this section we will show that if the vectors \mathbf{a}_i for $1 \leq i \leq M$ are drawn from a random distribution and (2) holds for a sufficiently large constant δ , then with high probability (3) produces an accurate estimate of \mathbf{x}_* . Our strategy is to show that for a sufficiently large M the sufficient condition provided in Lemma 2 holds with high probability.

For $\delta \in (0, 1)$ let \mathcal{C}_δ be the convex cone given by

$$\mathcal{C}_\delta = \{ \mathbf{y} \in \mathbb{C}^N : \mathbf{x}_*^* \mathbf{y} \geq \delta \|\mathbf{y}\|_2 \},$$

where $\mathbf{x}_*^* \mathbf{y}$ is implicitly assumed to be a real number. The *polar cone* of a set \mathcal{C} is defined as

$$\mathcal{C}^\circ \stackrel{\text{def}}{=} \{ \mathbf{z} : \langle \mathbf{z}, \mathbf{y} \rangle \leq 0 \text{ for all } \mathbf{y} \in \mathcal{C} \}.$$

It is easy to verify that the polar cone of \mathcal{C}_δ is

$$\mathcal{C}_\delta^\circ = \left\{ \mathbf{z} \in \mathbb{C}^N : \delta \langle \mathbf{x}_*, \mathbf{z} \rangle \leq -\sqrt{1-\delta^2} \sqrt{\|\mathbf{z}\|_2^2 - |\mathbf{x}_*^* \mathbf{z}|^2} \right\}.$$

Since $\mathbf{a}_0 \in \mathcal{C}_\delta$ by assumption, it follows that for every $\mathbf{h} \in \mathcal{C}_\delta^\circ$ we have $\langle \mathbf{a}_0, \mathbf{h} \rangle \leq 0$. Therefore, the inequality $\langle \mathbf{a}_0, \mathbf{h} \rangle \geq 0$ can hold only for vectors \mathbf{z} in the closure of the complement of \mathcal{C}_δ° which we denote by

$$\mathcal{C}'_\delta = \left\{ \mathbf{z} \in \mathbb{C}^N : \delta \langle \mathbf{x}_*, \mathbf{z} \rangle \geq -\sqrt{1-\delta^2} \sqrt{\|\mathbf{z}\|_2^2 - |\mathbf{x}_*^* \mathbf{z}|^2} \right\}. \quad (6)$$

A typical positioning of \mathbf{a}_0 and $\mathbf{a}_i \mathbf{a}_i^* \mathbf{x}_*$ needed to guarantee unique recovery is illustrated in Figure 4.

Theorem 1. *Suppose that noisy phaseless measurements of a unit vector \mathbf{x}_* as in (1) are given under bounded non-negative noise $\xi_1, \xi_2, \dots, \xi_M \in [0, \eta^{-1}]$. Let \mathbf{a}_0 be an anchor vector obeying (2) for some constant $\delta \in (0, 1)$ and define \mathcal{R}_δ and \mathcal{C}'_δ respectively as in (4) and (6). Furthermore, given a constant $t > 0$, suppose that for $1 \leq i \leq M$ the measurement vectors \mathbf{a}_i are i.i.d. copies of a random variable $\mathbf{a} \in \mathbb{C}^N$ that obeys*

$$\inf_{\substack{\mathbf{h} \in \mathcal{C}'_\delta \cap \mathcal{R}_\delta \\ \|\mathbf{h}\|_2 > (t\eta)^{-1}}} \mathbb{P} \left(\langle \mathbf{a} \mathbf{a}^* \mathbf{x}_*, \mathbf{h} \rangle > \frac{1}{2} \eta^{-1} \right) \geq p_{\min}(\delta, t),$$

for a constant $p_{\min}(\delta, t) \in (0, 1)$ depending only on δ and t .³ For any $\varepsilon > 0$, if we have

$$M \stackrel{p_{\min}(\delta, t)}{\gtrsim} N + \log \frac{1}{\varepsilon},$$

with the hidden constant factor inversely related to $p_{\min}(\delta, t)$, then with probability $\geq 1 - \varepsilon$ the estimate $\hat{\mathbf{x}}$ obtained through (3) obeys

$$\|\hat{\mathbf{x}} - \mathbf{x}_*\|_2 \leq (t\eta)^{-1}.$$

Proof. Let $\mathbf{h} = \hat{\mathbf{x}} - \mathbf{x}_*$. It suffices to show that for any $\mathbf{h} \in \mathcal{C}'_\delta \cap \mathcal{R}_\delta$ with $\|\mathbf{h}\|_2 > (t\eta)^{-1}$ there exists at least one $1 \leq i \leq M$ such that $\langle \mathbf{a}_i \mathbf{a}_i^* \mathbf{x}_*, \mathbf{h} \rangle > \frac{1}{2} \eta^{-1}$. Specifically, we would like to show that with high probability $\sum_{i=1}^M \mathbb{1}(\langle \mathbf{a}_i \mathbf{a}_i^* \mathbf{x}_*, \mathbf{h} \rangle > \frac{1}{2} \eta^{-1}) > 0$. Denote the empirical probability of $\langle \mathbf{a} \mathbf{a}^* \mathbf{x}_*, \mathbf{h} \rangle > \frac{1}{2} \eta^{-1}$ by

$$\hat{p}_M(\mathbf{h}) = \frac{1}{M} \sum_{i=1}^M \mathbb{1} \left(\langle \mathbf{a}_i \mathbf{a}_i^* \mathbf{x}_*, \mathbf{h} \rangle > \frac{1}{2} \eta^{-1} \right),$$

which is an approximation of the true probability of the event denoted by

$$\begin{aligned} p(\mathbf{h}) &= \mathbb{E} \mathbb{1} \left(\langle \mathbf{a} \mathbf{a}^* \mathbf{x}_*, \mathbf{h} \rangle > \frac{1}{2} \eta^{-1} \right) \\ &= \mathbb{P} \left(\langle \mathbf{a} \mathbf{a}^* \mathbf{x}_*, \mathbf{h} \rangle > \frac{1}{2} \eta^{-1} \right). \end{aligned}$$

Considering the set of binary functions

$$\mathcal{F} = \left\{ \mathbf{z} \mapsto \mathbb{1} \left(\langle \mathbf{z}, \mathbf{h} \rangle > \frac{1}{2} \eta^{-1} \right) : \mathbf{h} \in \mathcal{C}'_\delta \cap \mathcal{R}_\delta \text{ and } \|\mathbf{h}\|_2 > (t\eta)^{-1} \right\}$$

whose shatter coefficient is denoted by $s(\mathcal{F}, M)$, a direct application of Theorem 2 in Appendix A shows that

$$\sup_{\substack{\mathbf{h} \in \mathcal{C}'_\delta \cap \mathcal{R}_\delta \\ \|\mathbf{h}\|_2 > (t\eta)^{-1}}} |\hat{p}_M(\mathbf{h}) - p(\mathbf{h})| \leq \sqrt{\frac{8 \log \frac{s(\mathcal{F}, M)}{\varepsilon}}{M}}$$

with probability $\geq 1 - \varepsilon$. Since \mathcal{F} is a subset of \mathcal{H} the set of indicators of all half-spaces (with a common offset), it has a smaller VC-dimension than \mathcal{H} . Moreover, it is well-known—as a direct implication of Radon’s theorem (see e.g., Matoušek, 2002)—that the VC-dimension of half-spaces indicators is no more than the ambient dimension. In particular, we have $\dim_{\text{VC}}(\mathcal{F}) \leq \dim_{\text{VC}}(\mathcal{H}) \leq 2N$ as our domain is effectively a $2N$ -dimensional real vector space. Therefore, invoking Lemma 4 below we obtain

$$\sup_{\substack{\mathbf{h} \in \mathcal{C}'_\delta \cap \mathcal{R}_\delta \\ \|\mathbf{h}\|_2 > (t\eta)^{-1}}} |\hat{p}_M(\mathbf{h}) - p(\mathbf{h})| \leq \sqrt{\frac{16N \log \frac{\varepsilon M}{2N} + 8 \log \frac{8}{\varepsilon}}{M}}.$$

³Clearly, the best $p_{\min}(\delta, t)$ decreases as t increases.

Now, because $p(\mathbf{h}) \geq p_{\min}(\delta, t)$ for all $\mathbf{h} \in \mathcal{C}'_\delta \cap \mathcal{R}_\delta$ with $\|\mathbf{h}\|_2 > (t\eta)^{-1}$, the above inequality implies that

$$\inf_{\substack{\mathbf{h} \in \mathcal{C}'_\delta \cap \mathcal{R}_\delta \\ \|\mathbf{h}\|_2 > (t\eta)^{-1}}} \widehat{p}_M(\mathbf{h}) \geq p_{\min}(\delta, t) - \sqrt{\frac{16N \log \frac{eM}{2N} + 8 \log \frac{8}{\varepsilon}}{M}}.$$

If $M = \frac{8}{p_{\min}^2(\delta, t)} (c \cdot 2N + 2 \log \frac{8}{\varepsilon})$, then we have

$$\begin{aligned} \log \frac{eM}{2N} + \frac{\log \frac{8}{\varepsilon}}{2N} &= \log \frac{8e}{p_{\min}^2(\delta, t)} + \log \left(c + \frac{\log \frac{8}{\varepsilon}}{N} \right) \\ &\quad + \frac{\log \frac{8}{\varepsilon}}{2N} \\ &\leq \log \frac{8e}{p_{\min}^2(\delta, t)} + \frac{c}{2} + \frac{\log \frac{8}{\varepsilon}}{2N} - 1 \\ &\quad + \log 2 + \frac{\log \frac{8}{\varepsilon}}{2N} \\ &< \log \frac{8e}{p_{\min}^2(\delta, t)} + \frac{c}{2} + \frac{\log \frac{8}{\varepsilon}}{N}, \end{aligned}$$

where we used the inequality $\log u - \log 2 = \log \frac{u}{2} \leq \frac{u}{2} - 1$ in the second line. Setting $c = 2 \log \frac{8e}{p_{\min}^2(\delta, t)}$, it follows that

$$\begin{aligned} \frac{16N \log \frac{eM}{2N} + 8 \log \frac{8}{\varepsilon}}{M} &= \frac{16N}{M} \left(\log \frac{eM}{2N} + \frac{\log \frac{8}{\varepsilon}}{2N} \right) \\ &< \frac{16N}{M} \left(2 \log \frac{8e}{p_{\min}^2(\delta, t)} + \frac{\log \frac{8}{\varepsilon}}{N} \right) = p_{\min}^2(\delta, t) \end{aligned}$$

and thus we can guarantee that

$$\inf_{\substack{\mathbf{h} \in \mathcal{C}'_\delta \cap \mathcal{R}_\delta \\ \|\mathbf{h}\|_2 > (t\eta)^{-1}}} \widehat{p}_M(\mathbf{h}) > p_{\min}(\delta, t) - p_{\min}(\delta, t) = 0.$$

This immediately implies that for $M \gtrsim N + \log \frac{1}{\varepsilon}$ we have

$$\inf_{\substack{\mathbf{h} \in \mathcal{C}'_\delta \cap \mathcal{R}_\delta \\ \|\mathbf{h}\|_2 > (t\eta)^{-1}}} \sum_{i=1}^M \mathbb{1}(\langle \mathbf{a}_i \mathbf{a}_i^* \mathbf{x}_*, \mathbf{h} \rangle > \frac{1}{2} \eta^{-1}) = M \widehat{p}_M(\mathbf{h}) > 0,$$

as desired. \square

We can consider the case of measurements with normal distribution as a concrete example. To apply the Theorem 1, it suffices to quantify the constant $p_{\min}(\delta, t)$ which can be achieved through Lemma 3 below.

Lemma 3. *If $\mathbf{a} \sim \text{Normal}(\mathbf{0}, \frac{1}{2}\mathbf{I}) + i\text{Normal}(\mathbf{0}, \frac{1}{2}\mathbf{I})$ and \mathbf{x}_* is a unit vector, then for every $\mathbf{h} \in \mathcal{C}'_\delta \cap \mathcal{R}_\delta$ with $\|\mathbf{h}\|_2 > (t\eta)^{-1}$ we have*

$$\mathbb{P} \left(\langle \mathbf{a} \mathbf{a}^* \mathbf{x}_*, \mathbf{h} \rangle > \frac{1}{2} \eta^{-1} \right) \geq \left(\frac{1}{2} - \frac{\sqrt{1-\delta^2}}{2} \right) e^{-2\sqrt{2}\delta^{-2}t}.$$

Proof. We can decompose \mathbf{h} as $\mathbf{h} = (\mathbf{x}_*^* \mathbf{h}) \mathbf{x}_* + \mathbf{h}_\perp$, where $\mathbf{x}_*^* \mathbf{h}_\perp = \mathbf{0}$. Therefore, we have $\langle \mathbf{a} \mathbf{a}^* \mathbf{x}_*, \mathbf{h} \rangle = \langle \mathbf{x}_*, \mathbf{h} \rangle |\mathbf{a}^* \mathbf{x}_*|^2 + \text{Re}(\overline{\mathbf{a}^* \mathbf{x}_*} \mathbf{a}^* \mathbf{h}_\perp)$. Using the facts that $\mathbf{x}_*^* \mathbf{h}_\perp = \mathbf{0}$ and $\mathbf{a} \sim \text{Normal}(\mathbf{0}, \frac{1}{2}\mathbf{I}) + i\text{Normal}(\mathbf{0}, \frac{1}{2}\mathbf{I})$ it is straightforward to show that $\langle \mathbf{x}_*, \mathbf{h} \rangle |\mathbf{a}^* \mathbf{x}_*|^2 + \text{Re}(\overline{\mathbf{a}^* \mathbf{x}_*} \mathbf{a}^* \mathbf{h}_\perp)$ has the same distribution as $\frac{1}{2} \langle \mathbf{x}_*, \mathbf{h} \rangle \|\mathbf{g}_2\|^2 + \frac{1}{2} \|\mathbf{h}_\perp\|_2 \mathbf{g}_1^\top \mathbf{g}_2$ where $\mathbf{g}_1, \mathbf{g}_2 \in \mathbb{R}^2$ are independent standard Normal random variables:

$$\begin{aligned} &\mathbb{P} \left(\langle \mathbf{a} \mathbf{a}^* \mathbf{x}_*, \mathbf{h} \rangle > \frac{1}{2} \eta^{-1} \right) \\ &= \mathbb{P} \left(\langle \mathbf{x}_*, \mathbf{h} \rangle \|\mathbf{g}_2\|_2^2 + \|\mathbf{h}_\perp\|_2 \mathbf{g}_1^\top \mathbf{g}_2 > \eta^{-1} \right). \end{aligned}$$

Since \mathbf{g}_2 has a standard normal distribution, its norm and (normalized) direction are independent. Thus, we can treat $-\mathbf{g}_1^\top \frac{\mathbf{g}_2}{\|\mathbf{g}_2\|_2} = g$ as a standard Normal scalar which is independent of $\|\mathbf{g}_2\|_2 = v \sim \text{Rayleigh}(1)$. Therefore, we have

$$\begin{aligned} &\mathbb{P} \left(\langle \mathbf{a} \mathbf{a}^* \mathbf{x}_*, \mathbf{h} \rangle > \frac{1}{2} \eta^{-1} \right) \\ &= \mathbb{P} \left(\langle \mathbf{x}_*, \mathbf{h} \rangle \|\mathbf{g}_2\|_2^2 + \|\mathbf{h}_\perp\|_2 \|\mathbf{g}_2\|_2 \mathbf{g}_1^\top \frac{\mathbf{g}_2}{\|\mathbf{g}_2\|_2} > \eta^{-1} \right) \\ &= \mathbb{P} \left(\langle \mathbf{x}_*, \mathbf{h} \rangle v - \eta^{-1} v^{-1} > \|\mathbf{h}_\perp\|_2 g \right). \end{aligned}$$

Since $\mathbf{h} \in \mathcal{R}_\delta$ and $\|\mathbf{h}\|_2 > (t\eta)^{-1}$ we have

$$\begin{aligned} (t\eta)^{-2} &< \|\mathbf{h}\|_2^2 = \|\mathbf{h}_\perp\|_2^2 + (\text{Im}(\mathbf{x}_*^* \mathbf{h}))^2 + \langle \mathbf{x}_*, \mathbf{h} \rangle^2 \\ &\leq (1 + \delta^{-2}) \|\mathbf{h}_\perp\|_2^2 + \langle \mathbf{x}_*, \mathbf{h} \rangle^2. \end{aligned} \quad (7)$$

We consider two cases depending on $\|\mathbf{h}_\perp\|_2 = 0$ or not. If $\|\mathbf{h}_\perp\|_2 = 0$, then $|\langle \mathbf{x}_*, \mathbf{h} \rangle| > (t\eta)^{-1}$. The fact that $\mathbf{h} \in \mathcal{C}'_\delta$ as well, implies that $\langle \mathbf{x}_*, \mathbf{h} \rangle$ is non-negative and thereby $\langle \mathbf{x}_*, \mathbf{h} \rangle > (t\eta)^{-1}$. Consequently, we have

$$\begin{aligned} &\mathbb{P} \left(\langle \mathbf{a} \mathbf{a}^* \mathbf{x}_*, \mathbf{h} \rangle > \frac{1}{2} \eta^{-1} \right) \\ &= \mathbb{P} \left(\langle \mathbf{x}_*, \mathbf{h} \rangle v - \eta^{-1} v^{-1} > \|\mathbf{h}_\perp\|_2 g \right) \\ &\geq \mathbb{P} \left(t^{-1} v - v^{-1} > 0 \right) = \mathbb{P}(v > \sqrt{t}) = e^{-\frac{t}{2}}. \end{aligned}$$

If $\|\mathbf{h}_\perp\|_2 > 0$, then we can invoke Lemma 5 in the Appendix with $\alpha = \frac{\langle \mathbf{x}_*, \mathbf{h} \rangle}{\|\mathbf{h}_\perp\|_2}$ to show that

$$\begin{aligned} &\mathbb{P} \left(\langle \mathbf{a} \mathbf{a}^* \mathbf{x}_*, \mathbf{h} \rangle > \frac{1}{2} \eta^{-1} \right) \\ &\geq \mathbb{P} \left(\frac{\langle \mathbf{x}_*, \mathbf{h} \rangle}{\|\mathbf{h}_\perp\|_2} v - \frac{\eta^{-1}}{\|\mathbf{h}_\perp\|_2} v^{-1} > g \right) \\ &= \left(\frac{1}{2} + \frac{\alpha}{2\sqrt{\alpha^2 + 1}} \right) \exp \left(-\frac{\eta^{-1}}{\|\mathbf{h}_\perp\|_2 \sqrt{\alpha^2 + 1 + \alpha}} \right). \end{aligned}$$

Then by rewriting (7) as $\frac{(t\eta)^{-2}}{\|\mathbf{h}_\perp\|_2^2} \leq 1 + \delta^{-2} + \alpha^2$ we have

$$\begin{aligned} & \mathbb{P}\left(\langle \mathbf{a}\mathbf{a}^* \mathbf{x}_*, \mathbf{h} \rangle > \frac{1}{2}\eta^{-1}\right) \\ & \geq \left(\frac{1}{2} + \frac{\alpha}{2\sqrt{\alpha^2+1}}\right) \exp\left(-\frac{\sqrt{1+\delta^{-2}+\alpha^2}}{\sqrt{\alpha^2+1}+\alpha}t\right). \end{aligned}$$

The fact that $\mathbf{h} \in \mathcal{C}'_\delta$, guarantees that $\alpha \geq -\sqrt{\delta^{-2}-1}$. Since $\frac{\alpha}{\sqrt{\alpha^2+1}}$ and $-\frac{\sqrt{1+\delta^{-2}+\alpha^2}}{\sqrt{\alpha^2+1}+\alpha}$ are both increasing in α , we obtain

$$\begin{aligned} & \mathbb{P}\left(\langle \mathbf{a}\mathbf{a}^* \mathbf{x}_*, \mathbf{h} \rangle > \frac{1}{2}\eta^{-1}\right) \\ & \geq \left(\frac{1}{2} - \frac{\sqrt{1-\delta^2}}{2}\right) \exp\left(-\frac{\sqrt{2\delta^{-2}}}{\sqrt{\delta^{-2}}-\sqrt{\delta^{-2}-1}}t\right), \\ & = \left(\frac{1}{2} - \frac{\sqrt{1-\delta^2}}{2}\right) \exp\left(-\frac{\sqrt{2}}{1-\sqrt{1-\delta^2}}t\right) \\ & \geq \left(\frac{1}{2} - \frac{\sqrt{1-\delta^2}}{2}\right) \exp\left(-2\sqrt{2}\delta^{-2}t\right). \end{aligned}$$

The above lower bound is the smaller one of the two considered cases and thus the proof is complete. \square

A Tools from statistical learning theory

For reference, here we provide some of the classic results in statistical learning theory that we employed in our analysis. We mostly follow the exposition of the subject presented by Devroye et al. (2013, chapters 13 and 14).

Definition 1 (Shatter coefficient). The n -th shatter coefficient (or growth function) of a class \mathcal{F} of binary functions $f: \mathcal{X} \rightarrow \{0, 1\}$ is defined as

$$s(\mathcal{F}, n) \stackrel{\text{def}}{=} \max_{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n \in \mathcal{X}} |\{(f(\mathbf{x}_1), f(\mathbf{x}_2), \dots, f(\mathbf{x}_n)) : f \in \mathcal{F}\}|.$$

Intuitively, the shatter coefficient $s(\mathcal{F}, n)$ is the largest number of binary patterns that the functions in \mathcal{F} can induce on n points.

Definition 2 (VC-dimension). The Vapnik-Chervonenkis (VC) dimension of a class \mathcal{F} of binary functions is the largest number n such that $s(\mathcal{F}, n) = 2^n$, namely,

$$\dim_{\text{VC}}(\mathcal{F}) \stackrel{\text{def}}{=} \max\{n : s(\mathcal{F}, n) = 2^n\}.$$

Naturally, $\dim_{\text{VC}}(\mathcal{F}) = \infty$ if $s(\mathcal{F}, n) = 2^n$ for all n .

If \mathcal{F} can induce all binary patterns on n points, \mathcal{F} is said to “shatter” n points. Therefore, the VC-dimension of \mathcal{F} is the largest number of points that \mathcal{F} can shatter.

Lemma 4 (Vapnik and Chervonenkis (1971); Sauer (1972); Shelah (1972)). For a class \mathcal{F} of binary functions with VC-dimension $d = \dim_{\text{VC}}(\mathcal{F})$ we have

$$s(\mathcal{F}, n) \leq \sum_{i=0}^d \binom{n}{i}.$$

In particular,

$$s(\mathcal{F}, n) \leq \left(\frac{en}{d}\right)^d. \quad (8)$$

The following theorem is originally due to Vapnik and Chervonenkis (1971). We restate the theorem as presented in Devroye et al. (2013).

Theorem 2 (Vapnik and Chervonenkis (1971)). Let \mathcal{F} be a class of binary functions and $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ be i.i.d. copies of an arbitrary random variable \mathbf{x} . Then for every $t > 0$ we have

$$\mathbb{P}\left(\sup_{f \in \mathcal{F}} \left| \frac{1}{n} \sum_{i=1}^n f(\mathbf{x}_i) - \mathbb{E}f(\mathbf{x}) \right| > t\right) \leq 8s(\mathcal{F}, n)e^{-nt^2/8}.$$

B Auxiliary Lemma

Lemma 5. Let $v \sim \text{Rayleigh}(1)$ and $g \sim \text{Normal}(0, 1)$ be independent random variables. Then we have

$$\begin{aligned} & \mathbb{P}(\alpha v + \beta v^{-1} > g) \\ & = \begin{cases} 1 - \frac{\sqrt{\alpha^2+1}-\alpha}{2\sqrt{\alpha^2+1}} e^{-\beta(\alpha+\sqrt{\alpha^2+1})} & \text{for } \beta \geq 0 \\ \frac{\sqrt{\alpha^2+1}+\alpha}{2\sqrt{\alpha^2+1}} e^{\beta/(\alpha+\sqrt{\alpha^2+1})} & \text{for } \beta < 0. \end{cases} \end{aligned}$$

for all $\alpha, \beta \in \mathbb{R}$.

Proof. We denote the standard normal cumulative distribution function and its derivative by $\Phi(\cdot)$ and $\phi(\cdot)$, respectively. Let $F(\beta) = \mathbb{P}(\alpha v + \beta v^{-1} > \gamma) = \mathbb{E}\Phi(\alpha v + \beta v^{-1})$. By Leibniz’s rule we have

$$\begin{aligned} F'(\beta) & = \mathbb{E}\left(\phi(\alpha v + \beta v^{-1}) v^{-1}\right) \\ & = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{1}{2}(\alpha v + \beta v^{-1})^2} v^{-1} \cdot v e^{-\frac{1}{2}v^2} dv \\ & = \frac{e^{-\alpha\beta}}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{1}{2}((\alpha^2+1)v^2 + \beta^2 v^{-2})} dv. \end{aligned}$$

Now let $G(\beta) = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{1}{2}((\alpha^2+1)v^2 + \beta^2 v^{-2})} dv$, so that $F'(\beta) = e^{-\alpha\beta} G(\beta)$. Using Leibniz’s rule again, we can write

$$\begin{aligned} G'(\beta) & = -\frac{\beta}{\sqrt{2\pi}} \int_0^\infty v^{-2} e^{-\frac{1}{2}((\alpha^2+1)v^2 + \beta^2 v^{-2})} dv \\ & = -\frac{\sqrt{\alpha^2+1}}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{1}{2}((\alpha^2+1)u^2 + \beta^2 u^{-2})} du \\ & = -\sqrt{\alpha^2+1} G(\beta), \end{aligned}$$

where the second line follows from the change of variable $v = \frac{\beta}{\sqrt{\alpha^2+1}}u^{-1}$. It is straightforward to show that $G(0) = \frac{1}{2\sqrt{\alpha^2+1}}$. A simple integration then yields

$$G(\beta) = G(0)e^{-\beta\sqrt{\alpha^2+1}} = \frac{1}{2\sqrt{\alpha^2+1}}e^{-\beta\sqrt{\alpha^2+1}}$$

for $\beta \geq 0$, and since $G(\beta)$ is even for all β we have

$$G(\beta) = \frac{1}{2\sqrt{\alpha^2+1}}e^{-|\beta|\sqrt{\alpha^2+1}}.$$

It then follows that

$$F'(\beta) = \frac{1}{2\sqrt{\alpha^2+1}}e^{-(\alpha\beta+|\beta|\sqrt{\alpha^2+1})}$$

Integrating again we obtain

$$F(\beta) = \begin{cases} F(0) + \frac{\sqrt{\alpha^2+1}-\alpha}{2\sqrt{\alpha^2+1}} \left(1 - e^{-\beta(\alpha+\sqrt{\alpha^2+1})}\right) & , \beta \geq 0 \\ F(0) - \frac{\sqrt{\alpha^2+1}+\alpha}{2\sqrt{\alpha^2+1}} \left(1 - e^{-\beta(\alpha-\sqrt{\alpha^2+1})}\right) & , \beta < 0. \end{cases}$$

We can calculate $F(0)$ as

$$\begin{aligned} F(0) &= \mathbb{P}(\alpha v > g) \\ &= \begin{cases} \frac{1}{2} + \frac{1}{2}\mathbb{P}(\alpha^2 v^2 > g^2) & \text{for } \alpha \geq 0 \\ \frac{1}{2}\mathbb{P}(\alpha^2 v^2 < g^2) & \text{for } \alpha < 0 \end{cases} \\ &= \begin{cases} \frac{1}{2} + \frac{1}{2}\mathbb{P}\left(\frac{\alpha^2}{\alpha^2+1} > \frac{g^2}{v^2+g^2}\right) & \text{for } \alpha \geq 0 \\ \frac{1}{2} - \frac{1}{2}\mathbb{P}\left(\frac{\alpha^2}{\alpha^2+1} > \frac{g^2}{v^2+g^2}\right) & \text{for } \alpha < 0 \end{cases} \\ &= \frac{\sqrt{\alpha^2+1} + \alpha}{2\sqrt{\alpha^2+1}}, \end{aligned}$$

where the last line follows from the fact that $\frac{g}{\sqrt{v^2+g^2}}$ has a uniform distribution over $[-1, 1]$. Replacing $F(0)$ in the expression of $F(\beta)$ and straightforward simplifications yield the desired result. \square

Acknowledgments

The authors were supported in part by ONR grant N00014-11-1-0459, NSF grants CCF-1415498 and CCF-1422540, and the Packard Foundation.

References

- S. R. Becker, E. J. Candès, and M. C. Grant. Templates for convex cone problems with applications to sparse signal recovery. *Mathematical Programming Computation*, 3(3):165–218, 2011.
- E. J. Candès and X. Li. Solving quadratic equations via PhaseLift when there are about as many equations as unknowns. *Foundations of Computational Mathematics*, 14(5):1017–1026, 2014.
- E. J. Candès, T. Strohmer, and V. Voroninski. PhaseLift: Exact and stable signal recovery from magnitude measurements via convex programming. *Communications on Pure and Applied Mathematics*, 66(8):1241–1274, 2013.
- E. J. Candès, X. Li, and M. Soltanolkotabi. Phase retrieval from coded diffraction patterns. *Applied and Computational Harmonic Analysis*, 39(2):277–299, 2015a.
- E. J. Candès, X. Li, and M. Soltanolkotabi. Phase retrieval via Wirtinger flow: Theory and algorithms. *Information Theory, IEEE Transactions on*, 61(4):1985–2007, Apr. 2015b.
- Y. Chen and E. Candès. Solving random quadratic systems of equations is nearly as easy as solving linear systems. In *Advances in Neural Information Processing Systems 28 (NIPS'15)*, pages 739–747. Curran Associates, Inc., Dec. 2015.
- L. Devroye, L. Györfi, and G. Lugosi. *A probabilistic theory of pattern recognition*, volume 31. Springer Science & Business Media, 2013.
- J. R. Fienup. Phase retrieval algorithms: a comparison. *Applied Optics*, 21(15):2758–2769, Aug 1982.
- R. W. Gerchberg and W. O. Saxton. A practical algorithm for the determination of phase from image and diffraction plane pictures. *Optik*, 35(2):237–346, 1972.
- T. Goldstein and C. Studer. PhaseMax: convex phase retrieval via basis pursuit. preprint [arXiv:1610.07531 \[cs.IT\]](https://arxiv.org/abs/1610.07531), Oct. 2016.
- P. Hand and V. Voroninski. Compressed sensing from phaseless gaussian measurements via linear programming in the natural parameter space. preprint [arXiv:1611.05985](https://arxiv.org/abs/1611.05985), 2016a.
- P. Hand and V. Voroninski. Corruption robust phase retrieval via linear programming. preprint [arXiv:1612.03547](https://arxiv.org/abs/1612.03547), 2016b.
- P. Hand and V. Voroninski. An elementary proof of convex phase retrieval in the natural parameter space via the linear program PhaseMax. preprint [arXiv:1611.03935](https://arxiv.org/abs/1611.03935), 2016c.
- K. Jaganathan, Y. C. Eldar, and B. Hassibi. Phase retrieval: An overview of recent developments. preprint [arXiv:1510.07713 \[cs.IT\]](https://arxiv.org/abs/1510.07713), 2015.
- J. Matoušek. *Lectures on discrete geometry*, volume 212 of *Graduate Texts in Mathematics*. Springer-Verlag New York, 2002.

- P. Netrapalli, P. Jain, and S. Sanghavi. Phase retrieval using alternating minimization. In *Advances in Neural Information Processing Systems 26 (NIPS 2013)*, pages 2796–2804, 2013.
- N. Sauer. On the density of families of sets. *Journal of Combinatorial Theory, Series A*, 13(1):145–147, 1972.
- P. Schniter and S. Rangan. Compressive phase retrieval via generalized approximate message passing. *IEEE Transactions on Signal Processing*, 63(4):1043–1055, Feb. 2015.
- S. Shelah. A combinatorial problem; stability and order for models and theories in infinitary languages. *Pacific J. Math.*, 41(1):247–261, 1972.
- J. Sun, Q. Qu, and J. Wright. A geometric analysis of phase retrieval. In *2016 IEEE International Symposium on Information Theory (ISIT)*, pages 2379–2383, July 2016. full-length version: [arXiv 1602.06664 \[cs.IT\]](#).
- V. N. Vapnik and A. Y. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probability & Its Applications*, 16(2):264–280, 1971.
- I. Waldspurger, A. d’Aspremont, and S. Mallat. Phase recovery, MaxCut and complex semidefinite programming. *Mathematical Programming*, 149(1-2):47–81, 2015.
- G. Wang and G. Giannakis. Solving random systems of quadratic equations via truncated generalized gradient flow. In *Advances in Neural Information Processing Systems 29 (NIPS’16)*, pages 568–576. Curran Associates, Inc., 2016.
- H. Zhang, Y. Chi, and Y. Liang. Provable non-convex phase retrieval with outliers: Median truncated wirtinger flow. In *Proceedings of The 33rd International Conference on Machine Learning (ICML’16)*, volume 48, pages 1022–1031, New York, NY, USA., 2016. JMLR W&CP.