
Learning without the Phase: Regularized PhaseMax Achieves Optimal Sample Complexity

Fariborz Salehi

Department of Electrical Engineering
Caltech
fsalehi@caltech.edu

Ehsan Abbasi

Department of Electrical Engineering
Caltech
eabbasi@caltech.edu

Babak Hassibi

Department of Electrical Engineering
Caltech
hassibi@caltech.edu*

Abstract

The problem of estimating an unknown signal, $\mathbf{x}_0 \in \mathbb{R}^n$, from a vector $\mathbf{y} \in \mathbb{R}^m$ consisting of m magnitude-only measurements of the form $y_i = |\mathbf{a}_i \mathbf{x}_0|$, where \mathbf{a}_i 's are the rows of a known measurement matrix \mathbf{A} , is a classical problem known as phase retrieval. This problem arises when measuring the phase is costly or altogether infeasible. In many applications in machine learning, signal processing, statistics, etc., the underlying signal has certain structure (sparse, low-rank, finite alphabet, etc.), opening up the possibility of recovering \mathbf{x}_0 from a number of measurements smaller than the ambient dimension, i.e., $m < n$. Ideally, one would like to recover the signal from a number of phaseless measurements that is on the order of the "degrees of freedom" of the structured \mathbf{x}_0 . To this end, inspired by the PhaseMax algorithm, we formulate a convex optimization problem, where the objective function relies on an initial estimate of the true signal and also includes an additive regularization term to encourage structure. The new formulation is referred to as **regularized PhaseMax**. We analyze the performance of regularized PhaseMax to find the minimum number of phaseless measurements required for perfect signal recovery. The results are asymptotic and are in terms of the geometrical properties (such as the Gaussian width) of certain convex cones. When the measurement matrix has i.i.d. Gaussian entries, we show that our proposed method is indeed order-wise optimal, allowing perfect recovery from a number of phaseless measurements that is only a constant factor away from the optimal number of measurements required when phase information is available. We explicitly compute this constant factor, in terms of the quality of the initial estimate, by deriving the exact phase transition. The theory well matches empirical results from numerical simulations.

1 Introduction

Recovering an unknown signal or model given a limited number of linear measurements is an important problem that appears in many applications. Researchers have developed various methods

*This work was supported in part by the National Science Foundation under grants CNS-0932428, CCF-1018927, CCF-1423663 and CCF-1409204, by a grant from Qualcomm Inc., by NASA's Jet Propulsion Laboratory through the President and Director's Fund, and by King Abdullah University of Science and Technology.

with rigorous theoretical guarantees for perfect signal reconstruction, e.g. [5, 16, 38, 43]. However, there are many practical scenarios in which the signal should be reconstructed from nonlinear measurements. In particular, in many physical devices, measuring the phase is expensive or even infeasible. For instance, detection devices such as CCD cameras and photosensitive films cannot measure the phase of a light wave and instead measure the photon flux [22].

The fundamental problem of recovering a signal from magnitude-only measurements is known as *phase retrieval*. It has a rich history and occurs in many areas in engineering and applied sciences such as medical imaging [15], X-ray crystallography [27], astronomical imaging [17], and optics [45]. Due to the loss of phase information, signal reconstruction from magnitude-only measurements can be quite challenging. Therefore, despite a variety of proposed methods and analysis frameworks, phase retrieval still faces fundamental theoretical and algorithmic challenges.

Recently, convex methods have gained significant attention to solve the phase retrieval problem. The first convex-relaxation-based methods were based on semidefinite programs [7, 10] and resorted to the idea of *lifting* [2, 8, 23, 36] the signal from a vector to a matrix to linearize the quadratic constraints. While the convex nature of this formulation allows theoretical guarantees, the resulting algorithms are computationally inefficient since the number of unknowns is effectively squared. This makes these approaches intractable when the system dimension is large.

Introduced in two independent papers [3, 19], PhaseMax is a novel convex relaxation for phase retrieval which works in the original n -dimensional parameter space. Since it does not require lifting and does not square the number of unknowns, it is appealing in practice. It does, however, require an initial estimate of the signal. Preliminary theoretical analysis [3, 13, 19, 21] indicates the method achieves perfect recovery for an order optimal number of random measurements. The exact phase transition for PhaseMax has been recently computed in a sequence of papers, first for the case of real measurements [14] and then for the case of complex ones [35].

Non-convex methods for phase retrieval have a long history [18]. Recent non-convex methods start with a careful initialization [25, 26, 28] and update the solution iteratively using a gradient-descent-like scheme. Examples of such methods include Wirtinger flow algorithms [9, 12, 37], truncated amplitude flow [46], and alternating minimization [29, 49]. Despite having lower computational cost, precise theoretical analysis of such algorithms seems very technically challenging.

All the aforementioned algorithms essentially demonstrate that a signal of dimension n can be perfectly recovered through $m > Cn$ amplitude-only measurements, where $C > 1$ is a constant that depends on the algorithm as well as the measurement vectors. However, many interesting signals in practice contain fewer degrees of freedom than the ambient dimension (sparse signals, low-rank matrices, finite alphabet signals, etc.). Such low-dimensional structures open up the possibility of perfect signal recovery with a number of measurements significantly smaller than n .

1.1 Summary of contributions

In this paper we propose a new approach for recovering *structured* signals. Inspired by the PhaseMax algorithm, we introduce a new convex formulation and investigate necessary and sufficient conditions, in terms of the number of measurements, for perfect recovery. We refer to this new framework as *regularized PhaseMax*. The constrained set in this optimization is obtained by relaxing the non-convex equality constraints in the original phase retrieval problem to convex inequality constraints. The objective function consists of two terms. One is a linear functional that relies on an initial estimate of the true signal which must be externally provided. The second term is an additive regularization term that is formed based on a priori structural information about the signal.

We utilize the recently developed Convex Gaussian Min-Max Theorem (CGMT) [39] to precisely compute the necessary and sufficient number of measurements for perfect signal recovery when the entries of the measurement matrix are i.i.d. Gaussian. To the extent of our knowledge, this is the first convex optimization formulation for the problem of structured signal recovery given phaseless linear Gaussian measurements that provably requires an order optimal number of measurements. In this paper we focus on real signals and real measurements. The complex case is more involved, requires a different analysis, and will be considered in a separate work. Through our analysis, we make the following main contributions:

- We first provide a sufficient recovery condition, in Section 3.1, in terms of the number of measurements, for perfect signal recovery. We use this to infer that our proposed method is order-wise optimal.
- We characterize the exact phase transition behavior for the class of absolutely scalable regularization functions.
- We apply our findings to two special examples: unstructured signal recovery and sparse recovery. We observe that the theory well matches the result of numerical simulations for these two examples.

1.2 Prior work

Phase retrieval for structured signals has gained significant attention in recent years. A review of all of the results is beyond the scope of this paper, and we instead briefly mention some of the most relevant literature for the Gaussian measurement model. Oymak et. al. [30] analyzed the performance of the regularized PhaseLift algorithm and observed that the required sample complexity is of a suboptimal order compared to the optimal number of measurements required when phase information is available. For the special case of sparse phase retrieval similar results have been reported in [24] which indicates $\mathcal{O}(k^2 \log(n))$ measurements are required for recovering of a k -sparse signal, using regularized PhaseLift. Recently, there has been a stream of work on solving phase retrieval using non-convex methods [6, 47]. In particular, Soltanolkotabi [37] has shown that amplitude-based Wirtinger flow can break the $\mathcal{O}(k^2 \log(n))$ barrier. We also note that the paper [20] analyzed the PhaseMax algorithm with ℓ_1 regularizer and observed that it achieves perfect recovery with $\mathcal{O}(k \log(n/k))$ samples, provided a well-correlated initialization point.

2 Preliminaries

2.1 Problem setup

Let $\mathbf{x}_0 \in \mathbb{R}^n$ denote the the underlying *structured* signal. We consider the *real* phase retrieval problem with the goal of recovering \mathbf{x}_0 from m magnitude-only measurements of the form,

$$y_i = |\mathbf{a}_i^\top \mathbf{x}_0|, \quad i = 1, 2, \dots, m, \quad (1)$$

where $\{\mathbf{a}_i \in \mathbb{R}^n\}_{i=1}^m$ is the set of (known) measurement vectors. In practice, this set is identified based on the experimental settings; however, throughout this paper (for our analysis purposes) we assume that the \mathbf{a}_i 's are drawn independently from a Gaussian distribution with mean zero and covariance matrix \mathbf{I}_n . In order to exploit the structure of the signal we assume $f(\cdot)$ is a *convex* function that measures the "complexity" of the structured solution. The regularized PhaseMax algorithm also relies on an initial estimate of the true signal. Here, \mathbf{x}_{init} is used to represent this initial guess. Our analysis is based on the critical assumption that both \mathbf{x}_{init} and \mathbf{x}_0 are **independent** of all the measurement vectors. The constraint set in generalized PhaseMax is derived by simply relaxing the equality constraints in (1) into *convex* inequality constraints. We introduce the following convex optimization problem to recover the signal:

$$\begin{aligned} \hat{\mathbf{x}} = \underset{\mathbf{x} \in \mathbb{R}^n}{\operatorname{argmin}} \quad & L_\lambda(\mathbf{x}) = -\mathbf{x}_{\text{init}}^\top \mathbf{x} + \lambda f(\mathbf{x}) \\ \text{subject to:} \quad & |\mathbf{a}_i^\top \mathbf{x}| \leq y_i, \quad \text{for } 1 \leq i \leq m. \end{aligned} \quad (2)$$

The function f is assumed to be sign invariant, i.e., $f(\mathbf{x}) = f(-\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$ ($-\mathbf{x}$ has the same "complexity" as \mathbf{x} .) Note that because of the global phase ambiguity of measurements in (1), we can only estimate \mathbf{x}_0 up to a sign. Up to this sign ambiguity, we can use the normalized mean squared error (NMSE), defined as $\frac{\|\hat{\mathbf{x}} - \mathbf{x}_0\|^2}{\|\mathbf{x}_0\|^2}$, to measure the performance of the solution. In this paper we investigate the conditions under which the optimization program (2) uniquely identifies the true signal, i.e., $\hat{\mathbf{x}} = \mathbf{x}_0$ (up to the sign). Our results are asymptotic which is valid when $m, n \rightarrow \infty$.

2.2 Background on convex analysis

Our results give the required number of measurements as a function of certain geometrical properties of the descent cone of the objective function. Here, we recall these definitions from convex analysis.

Definition 1. (Descent cone) For a function $R : \mathbb{R}^n \rightarrow \mathbb{R}$ the descent(tangent) cone at point \mathbf{x} is defined as,

$$T_R(\mathbf{x}) = \text{cone}(\{\mathbf{z} \in \mathbb{R}^n : R(\mathbf{x} + \mathbf{z}) \leq R(\mathbf{x})\}) , \quad (3)$$

where $\text{cone}(\mathcal{S})$ denotes the closed conical hull of the set \mathcal{S} .

Definition 2. Let \mathcal{S} be a closed convex set in \mathbb{R}^n . For $\mathbf{x} \in \mathbb{R}^n$ the projection of \mathbf{x} on \mathcal{S} , denoted by $\Pi_{\mathcal{S}}(\mathbf{x})$, is defined as follows,

$$\Pi_{\mathcal{S}}(\mathbf{x}) := \underset{\mathbf{y} \in \mathcal{S}}{\text{argmin}} \|\mathbf{x} - \mathbf{y}\| , \quad (4)$$

where $\|\cdot\|$ is the Euclidean norm. The distance function is defined as: $\text{dist}_{\mathcal{S}}(\mathbf{x}) = \|\mathbf{x} - \Pi_{\mathcal{S}}(\mathbf{x})\|$.

Definition 3. (Statistical dimension) [1] The statistical dimension of a closed convex cone \mathcal{C} in \mathbb{R}^n is defined as,

$$d(\mathcal{C}) = \mathbb{E}_{\mathbf{g}} [\|\Pi_{\mathcal{C}}(\mathbf{g})\|^2] , \quad (5)$$

where $\mathbf{g} \in \mathbb{R}^n$ is a standard normal vector.

The statistical dimension canonically extends the dimension of linear spaces to convex cones. This quantity has been extensively studied in linear inverse problems. It is well-known that as $n \rightarrow \infty$, $m > d(T_{L_{\lambda}}(\mathbf{x}_0))$ is the necessary and sufficient condition for perfect signal recovery under noiseless linear Gaussian measurements [11, 38]. Our analysis indicates that given *phaseless* linear measurements, the regularized PhaseMax algorithm requires $\mathcal{O}(d(T_{L_{\lambda}}(\mathbf{x}_0)))$ measurements for perfect signal reconstruction. Therefore, it is order-wise optimal in that sense.

3 Main Results

In this section we present the main results of the paper which provide us with the required number of measurements for perfect signal recovery in the regularized PhaseMax optimization (2). This gives the value $m_0 = m_0(n, \mathbf{x}_0, \mathbf{x}_{\text{init}}, \lambda)$, such that the regularized PhaseMax algorithm uniquely identifies the underlying signal \mathbf{x}_0 with high probability whenever $m > m_0$.

In Section 3.1, we find sufficient conditions for recovery of the underlying signal. Theorem 1 provides an upper bound on the number of measurements that is equal to a constant factor times the statistical dimension of the descent cone, $d(T_{L_{\lambda}}(\mathbf{x}_0))$. Therefore, even though our analysis is not exact in this section, it leads us to the important observation that our proposed method is order-wise optimal in terms of the required sample complexity for perfect signal reconstruction.

In Section 3.2, we provide an exact analysis for the phase transition behavior of regularized PhaseMax when the regularizer is an absolutely scalable function. We apply this result to the case of unstructured phaseless recovery as well as sparse phaseless recovery to compute the exact phase transitions. We then compare the result of theory with the empirical results from numerical simulations.

3.1 Sufficient recovery condition

Let $\mathbf{P} := \frac{1}{\|\mathbf{x}_0\|^2} \mathbf{x}_0 \mathbf{x}_0^T$ and $\mathbf{P}^{\perp} := \mathbf{I} - \mathbf{P}$ denote the projectors onto the span of \mathbf{x}_0 and its orthogonal complement, respectively, where $\|\cdot\|$ denotes the ℓ_2 -norm of the vectors. We also define $d^{(n)} := d(T_{L_{\lambda}}(\mathbf{x}_0))$ as the statistical dimension of the descent cone of the objective function at point \mathbf{x}_0 . Our analysis rigorously characterizes the phase transition behavior of the regularized PhaseMax in the large system limit, i.e., when $n \rightarrow \infty$, while m and $d^{(n)}$ grow at a proportional ratio $\delta = \frac{m}{d^{(n)}}$. δ is often called the oversampling ratio. Here, the superscript (n) is used to denote the elements of a sequence. To streamline the notations, we often drop this when understood from the context.

Theorem 1 provides sufficient conditions for the successful recovery of \mathbf{x}_0 . The recovery threshold depends on λ and the initialization vector, \mathbf{x}_{init} . We define $\rho_{\text{init}} := \mathbf{x}_{\text{init}}^T \mathbf{x}_0$ to quantify the caliber of the initial estimate. Due to the sign invariance property of the solution, we can assume without loss of generality that $\rho_{\text{init}} \geq 0$. Before stating the theorem, we shall introduce the function $R(\cdot) : (2, +\infty) \rightarrow \mathbb{R}_+$.

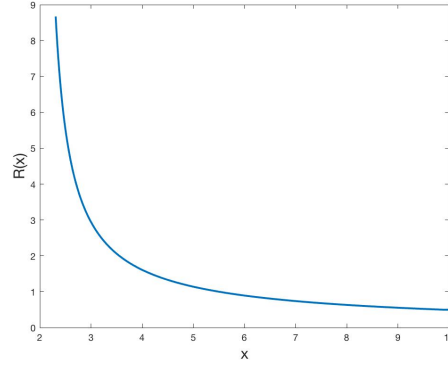


Figure 1: $R(x)$ for different values of x . R is a monotonically decreasing function.

Definition 4. For $x > 2$, $R(x)$ is the unique nonzero solution of the following equation:

$$t^2 = \frac{x}{\pi}((1 + t^2)\text{atan}(t) - t) . \quad (6)$$

Figure 1 depicts the evaluation of the function $R(x)$ for different input values x . As observed, $R(x)$ is a decreasing function with respect to x , and it approaches zero as x grows to infinity. It can be shown that for large values of the input x , $R(x)$ decays with the rate $\frac{1}{x}$.

Theorem 1 (Sufficient recovery condition). *For a fixed oversampling ratio $\delta > 2$, the regularized PhaseMax optimization (2) perfectly recovers the target signal (in the sense that $\lim_{n \rightarrow \infty} \mathbb{P}\{\|\hat{\mathbf{x}} - \mathbf{x}_0\|^2 > \epsilon \|\mathbf{x}_0\|^2\} = 0$, for any fixed $\epsilon > 0$) if,*

$$R(\delta) < \sup_{\mathbf{v} \in \partial L_\lambda(\mathbf{x}_0)} \frac{\|\mathbf{P}\mathbf{v}\|}{\|\mathbf{P}^\perp \mathbf{v}\|} , \quad (7)$$

where $\partial L_\lambda(\mathbf{x}_0)$ denotes the sub-differential set of the objective function $L_\lambda(\cdot)$ at point \mathbf{x}_0 .

It is worth noting that $\partial L_\lambda(\mathbf{x}_0)$ is a convex and compact set, and it can be expressed in terms of the sub-differential of the regularization function $\partial f(\mathbf{x}_0)$ as following,

$$\partial L_\lambda(\mathbf{x}_0) = \{\lambda \mathbf{u} - \mathbf{x}_{\text{init}} : \mathbf{u} \in \partial f(\mathbf{x}_0)\} . \quad (8)$$

Observe that since $R(\cdot)$ is a monotonically decreasing function, the inequality (7) gives a lower bound for the oversampling ratio δ . Indeed, we can restate the result in terms of this lower bound as the following corollary:

Corollary 1. *If there exist a fixed constant $\tau > 0$ such that,*

$$\sup_{\mathbf{v} \in \partial L_\lambda(\mathbf{x}_0)} \frac{\|\mathbf{P}\mathbf{v}\|}{\|\mathbf{P}^\perp \mathbf{v}\|} > \tau, \quad (9)$$

then the regularized PhaseMax optimization (2) has perfect recovery for $\delta > C$, where C is a constant that only depends on τ .

Proof. It is an immediate consequence of Theorem 1 by choosing $C = R^{-1}(\tau)$ and noting that $R(\cdot)$ is monotonically decreasing. \square

This result indicates that if \mathbf{x}_{init} and λ are chosen in such a way that the inequality (9) is satisfied for some positive constant τ , then one needs $m > Cd^{(n)}$ measurement samples for perfect recovery, where C is a constant and $d^{(n)} (= d)$ is the statistical dimension of the descent cone of the objective function at point \mathbf{x}_0 . As motivating examples, we use Theorem 1 to find upper bounds on the phase transition when \mathbf{x}_0 has no structure or it is a sparse signal.

Example 1: Assume the target signal \mathbf{x}_0 has no a priori structure. The objective function in this case would be $L(\mathbf{x}) = -\mathbf{x}_{\text{init}}^\top \mathbf{x}$, and $\partial L(\mathbf{x}_0) = \{-\mathbf{x}_{\text{init}}\}$. It can be shown that the statistical dimension is

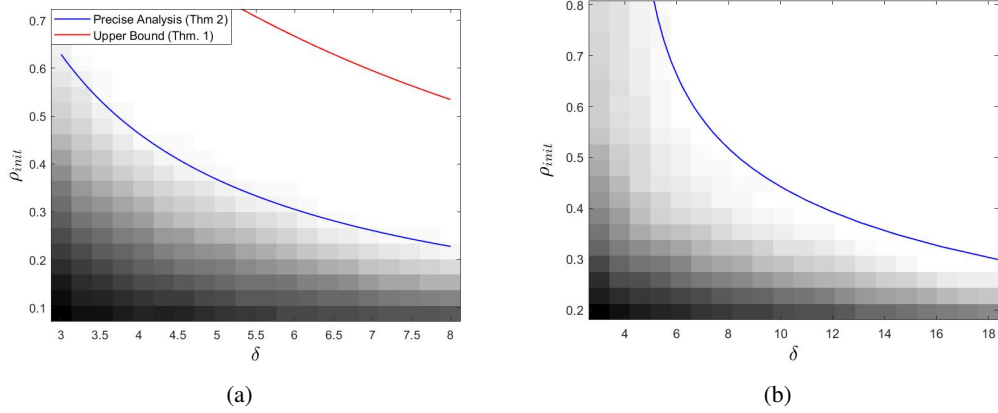


Figure 2: Phase transition regimes for the regularized PhaseMax problem in terms of the oversampling ratio δ and $\rho_{\text{init}} = \mathbf{x}_{\text{init}}^T \mathbf{x}_0$, for the cases of \mathbf{x}_0 with (a) no structure and (b) sparse signal recovery. The blue lines indicate the theoretical estimate for the phase transition derived from Theorem 2. The red line in (a) correspond to the upper bound calculated by Theorem 1. In the simulations we used signals of size $n = 128$. The result is averaged over 10 independent realization of the measurements.

$d^{(n)} = n - 1/2$. Due to the absence of the regularization term in this case, without loss of generality we can assume $\|\mathbf{x}_0\| = \|\mathbf{x}_{\text{init}}\| = 1$. Theorem 1 provides the following sufficient condition for perfect recovery:

$$\frac{\|\mathbf{P}\mathbf{x}_{\text{init}}\|}{\|\mathbf{P}^\perp \mathbf{x}_{\text{init}}\|} = \frac{\rho_{\text{init}}}{\sqrt{1 - \rho_{\text{init}}^2}} > R(\delta). \quad (10)$$

This indicates $\mathcal{O}(n)$ measurements is sufficient for perfect recovery as long as $\rho_{\text{init}} \geq \rho_0$, where $\rho_0 > 0$ is a constant that does not approach zero as $n \rightarrow \infty$. The exact phase transition for the unstructured case (PhaseMax) has been derived in [14] which is compatible with this result. Figure 2(a) shows the result of numerical simulation for different values of δ and ρ_{init} , when $n = 128$. As depicted in the figure, the sufficient recovery condition from Theorem 1 is approximately a factor of 2 away from the actual phase transition.

Example 2: Let \mathbf{x}_0 be a k -sparse signal. In this case we use $\|\cdot\|_1$ as the regularization function. We show in Section 5.5 that if $\lambda > \frac{c}{\sqrt{k}}$, then $d^{(n)} \leq Ck \log(n/k)$, for some constants $c, C > 0$. This matches the well-known order for the statistical dimension derived in the compressive sensing literature [38].

Moreover, in order to satisfy the condition in Corollary 1 we need to have $\frac{\rho_{\text{init}}}{\|\mathbf{x}_0\|_1} > (1 + \epsilon)\lambda$, for some $\epsilon > 0$. Therefore, \mathbf{x}_0 can be perfectly recovered having $\mathcal{O}(k \log(n/k))$ samples when the hyper-parameter λ is tuned properly, i.e., $\frac{c}{\sqrt{k}} < \lambda < \frac{\rho_{\text{init}}}{\|\mathbf{x}_0\|_1}$. Figure 3(a) compares this upper bounds with the precise analysis that we will show in Section 3.2. As depicted in this figure, the sufficient recovery condition is a valid upper bound on the phase transition, but it is not sharp.

3.2 Precise phase transition

So far, we have provided a sufficient condition for perfect signal recovery in the regularized PhaseMax. In this section we give the exact phase transition, i.e., the minimum number of measurements m_0 required for perfect recovery of the unknown vector \mathbf{x}_0 . For our analysis, we assume that the function $f(\mathbf{x})$ is absolutely homogeneous (scalable), i.e., $f(\tau \cdot \mathbf{x}) = |\tau| \cdot f(\mathbf{x})$, for any scalar τ . This covers a large range of regularization functions such as norms and semi-norms. Let $\partial L_\lambda^\perp(\mathbf{x}_0) \subset \mathbb{R}^n$ denote the projection of the sub-differential set into the orthogonal complement of \mathbf{x}_0 , i.e.,

$$\partial L_\lambda^\perp(\mathbf{x}_0) = \{\mathbf{P}^\perp \mathbf{u} : \mathbf{u} \in \partial L_\lambda(\mathbf{x}_0)\}, \quad (11)$$

which is a convex and compact set. To state the result in a general framework, we require a further assumption on functions $L_\lambda^{(n)}(\cdot)$.

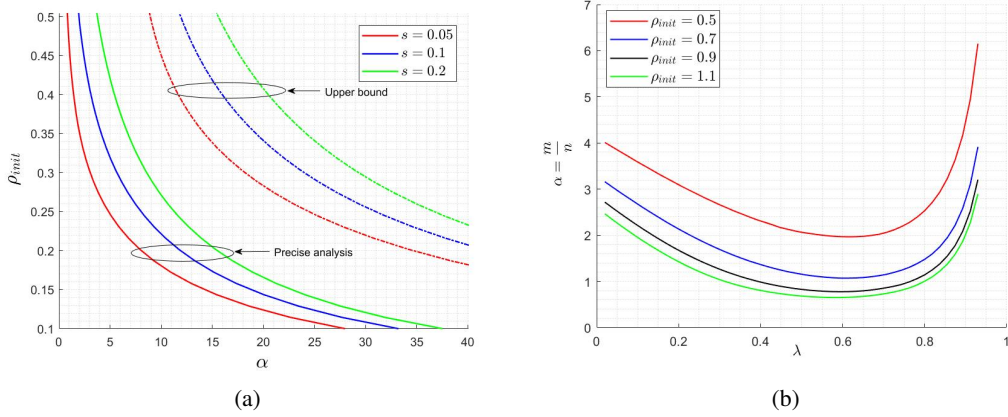


Figure 3: (a) Comparing the upper bounds on the phase transition, derived by Theorem 1 (dashed lines) and the precise phase transition by Theorem 2 (solid lines), for three values of the sparsity factor $s = k/n$. (b) The phase transition behavior as a function of the regularization parameter λ , derived from the result of Theorem 2.

Assumption 1 (Asymptotic functionals) *We say Assumption 1 holds if the following uniform convergences exist, as $n \rightarrow \infty$,*

$$\begin{aligned} \beta - \mathbb{E} \left[\frac{1}{\sqrt{n}} \mathbf{h}^\top \Pi_{\partial L_\lambda(\mathbf{x}_0)} \left(\frac{\beta}{\sqrt{n}} \mathbf{h} \right) \right] &\xrightarrow{\text{Unif.}} F_\lambda(\beta), \text{ and,} \\ \mathbb{E} \left[\text{dist}_{\partial L_\lambda(\mathbf{x}_0)} \left(\frac{\beta}{\sqrt{n}} \mathbf{h} \right) \right] &\xrightarrow{\text{Unif.}} G_\lambda(\beta), \end{aligned} \quad (12)$$

where $\mathbf{h} \in \mathbb{R}^n$ has i.i.d. standard normal entries and $F_\lambda, G_\lambda : \mathbb{R}_+ \rightarrow \mathbb{R}$ denote the functions that the sequences uniformly converge to.

One can show that, under some mild conditions on the regularization function $f(\cdot)$, Assumption 1 holds and also $F_\lambda(\beta) = G_\lambda(\beta)G'_\lambda(\beta)$, where $G'_\lambda(\cdot)$ denotes the derivative of the function $G_\lambda(\cdot)$. This assumption especially holds for the class of separable regularizers, where $f(\mathbf{v}) = \sum_i \tilde{f}(v_i)$ (e.g. ℓ_1 norm for the case of sparse phase-retrieval). Later in this section, we will see validity of this assumption for two examples discussed earlier in Section 3. Our precise phase transition results indicate the required number of measurements as the solution of a set of two nonlinear equations with two unknowns. We define a new parameter $\alpha := \frac{m}{n}$, where $\alpha_{\text{opt}} = \frac{m_0}{n}$ indicates the exact phase transition of the regularized PhaseMax optimization. The following theorem gives an implicit formula to derive α_{opt} .

Theorem 2 (Precise phase transition). *Let $\hat{\mathbf{x}}$ be the solution to the regularized PhaseMax optimization (2) with the objective function $L_\lambda(\mathbf{x}) = -\mathbf{x}_{init}^\top \mathbf{x} + \lambda f(\mathbf{x})$, where the convex function $f(\cdot)$ is absolutely homogeneous and Assumption 1 holds. The regularized PhaseMax optimization would perfectly recover the target signal \mathbf{x}_0 if and only if:*

1. $\alpha > \alpha_{\text{opt}}$, where α_{opt} is the solution of the following system of non-linear equations with two unknowns, α and β ,

$$\begin{cases} -G_\lambda(\beta) L_\lambda(\mathbf{x}_0) = \tan\left(\frac{\pi}{\alpha\beta} F_\lambda(\beta)\right) (G_\lambda^2(\beta) - \beta F_\lambda(\beta)), \\ \tan\left(\frac{\pi}{\alpha\beta} F_\lambda(\beta)\right) (G_\lambda(\beta) + \frac{\pi}{\alpha\beta} F_\lambda(\beta) L_\lambda(\mathbf{x}_0)) = \frac{\pi}{\alpha\beta} F_\lambda(\beta) G_\lambda(\beta), \end{cases} \quad (13)$$

2. and, $L_\lambda(\mathbf{x}_0) < L_\lambda(0) = 0$.

where the functions $F_\lambda(\cdot)$ and $G_\lambda(\cdot)$ are defined in (12).

A few remarks are in place for this theorem:

[Solving equations (13)] The system of nonlinear equations (13) only involves two scalars β and α , and the functions $F_\lambda(\beta)$ and $G_\lambda(\beta)$ are determined by the objective function $L_\lambda(\mathbf{x})$. For our numerical simulations in the examples of Section 3.2.1 and Section 3.2.2, we used a fixed-point

iteration method that can quickly find the solution given a proper initialization.

[Tuning λ] Theorem 2 requires the objective function to satisfy $L_\lambda(\mathbf{x}_0) = \lambda f(\mathbf{x}_0) - \rho_{\text{init}} < 0$. Therefore, it is necessary to choose λ in such a way that $\lambda < \rho_{\text{init}}/f(\mathbf{x}_0)$. Some additional assumptions on the unknown vector \mathbf{x}_0 enables us to calculate the proper range for λ . For instance, if we consider a random ensemble for \mathbf{x}_0 where the non-zero entries of \mathbf{x}_0 are Gaussian (or other) random variables, $\mathbb{E}[f(\mathbf{x}_0)]$ gives a reasonable estimation on $f(\mathbf{x}_0)$ that can help us choosing λ appropriately. We will see an example of such case in section 3.2.2. Figure 3(b) shows an example of how the phase transition of the regularized PhaseMax, or equivalently the required sample complexity, behaves as a function of the hyper-parameter λ .

In the next sections, we use the result of Theorem 2 to compute the exact phase transition for the case of unstructured signal as well as the sparse signal recovery. Since the regularizer $f(\mathbf{x})$ is absolutely scalable, for both examples, we assume that $\|\mathbf{x}_0\| = 1$.

3.2.1 Unstructured signal recovery

When there is no a priori information about the structure of the target signal, we use the following optimization (PhaseMax) for signal recovery:

$$\begin{aligned} \hat{\mathbf{x}} = \underset{\mathbf{x} \in \mathbb{R}^n}{\operatorname{argmin}} \quad & L(\mathbf{x}) = -\mathbf{x}_{\text{init}}^\top \mathbf{x} \\ \text{subject to:} \quad & |\mathbf{a}_i^\top \mathbf{x}| \leq y_i, \text{ for } 1 \leq i \leq m. \end{aligned} \quad (14)$$

Due to the absence of the regularization term, without loss of generality we can assume $\|\mathbf{x}_{\text{init}}\| = 1$. Moreover, $L(\mathbf{x}_0) = -\rho_{\text{init}}$ which indicates that the second condition in Theorem 2. To apply the result of our theorem, we first compute explicit formulas for the functions $F_\lambda(\beta)$, and $G_\lambda(\beta)$, as follows,

$$F_\lambda(\beta) = \beta, \quad G_\lambda(\beta) = \sqrt{\beta^2 + 1 - \rho_{\text{init}}^2}. \quad (15)$$

We can now form the system of nonlinear equations (13) as follows,

$$\begin{cases} \sqrt{\beta^2 + 1 - \rho_{\text{init}}^2} \frac{\rho_{\text{init}}}{1 - \rho_{\text{init}}^2} = \tan(\frac{\pi}{\alpha}), \\ \tan(\frac{\pi}{\alpha}) (\sqrt{\beta^2 + 1 - \rho_{\text{init}}^2} - \frac{\pi \rho_{\text{init}}}{\alpha}) = \frac{\pi}{\alpha} \sqrt{\beta^2 + 1 - \rho_{\text{init}}^2}. \end{cases} \quad (16)$$

Finally, solving equations (16) yields the following necessary and sufficient condition for perfect recovery,

$$\frac{\pi}{\alpha \tan(\pi/\alpha)} > 1 - \rho_{\text{init}}^2, \quad (17)$$

which also verifies the result of [14].

Figure 2(a) shows the result of numerical simulations of running the PhaseMax algorithm for different values of ρ_{init} and δ . The intensity level of the color of each square in Figure 2, represents the error of PhaseMax in recovering \mathbf{x}_0 . As seen in the figure, although our theoretical results has been established for the asymptotic setting (when the problem dimensions approach infinity), the blue line, which is derived from (17), reasonably predicts the phase transition for $n = 128$. The sufficient conditions that is derived from Theorem 1 is also depicted by the red line in the same figure.

3.2.2 Sparse recovery

We consider the case where the target signal \mathbf{x}_0 is sparse with k non-zero entries. The convex function $f(\mathbf{x}) = \frac{1}{\sqrt{n}} \|\mathbf{x}\|_1$, which is known to be a proper regularizer that enforces sparsity [41], is used in the regularized PhaseMax optimization to recover \mathbf{x}_0 ,

$$\begin{aligned} \hat{\mathbf{x}} = \underset{\mathbf{x} \in \mathbb{R}^n}{\operatorname{argmin}} \quad & L_\lambda(\mathbf{x}) = -\mathbf{x}_{\text{init}}^\top \mathbf{x} + \frac{\lambda}{\sqrt{n}} \|\mathbf{x}\|_1 \\ \text{subject to:} \quad & |\mathbf{a}_i^\top \mathbf{x}| \leq y_i, \text{ for } 1 \leq i \leq m. \end{aligned} \quad (18)$$

To streamline notations, we assume the non-zero entries of \mathbf{x}_0 are the first k entries and decompose vector $\mathbf{v} \in \mathbb{R}^n$ as $\mathbf{v} = \begin{bmatrix} \mathbf{v}^\Delta \\ \mathbf{v}^{\Delta^c} \end{bmatrix}$, where $\mathbf{v}^\Delta \in \mathbb{R}^k$ denotes the first k entries of \mathbf{v} , and $\mathbf{v}^{\Delta^c} \in \mathbb{R}^{n-k}$ is the remaining $n - k$ entries. As $m, n \rightarrow \infty$, we would like to apply the result of Theorem 2

to compute the exact phase transition. Due to the rotational invariance property of the Gaussian distribution, it can be shown that multiplying the last $(n - k)$ entries of \mathbf{x}_{init} , by a unitary matrix $\mathbf{U} \in \mathbb{R}^{(n-k) \times (n-k)}$ does not change the phase transition behavior in (2). Hence, we can assume the entries of $\mathbf{x}_{\text{init}}^{\Delta^c}$ have Gaussian distribution, i.e.,

$$\mathbf{x}_{\text{init}} = \begin{bmatrix} \mathbf{x}_{\text{init}}^{\Delta} \\ \mathbf{x}_{\text{init}}^{\Delta^c} \end{bmatrix}, \quad \text{and} \quad \mathbf{x}_{\text{init}}^{\Delta^c} = \frac{1}{\sqrt{n-k}} \|\mathbf{x}_{\text{init}}^{\Delta^c}\| \mathbf{g}, \quad (19)$$

where $\mathbf{g} \in \mathbb{R}^{n-k}$ has standard normal entries. This observation enables us to establish the following lemma:

Lemma 1. *Consider the optimization problem (18) to recover the k -sparse signal \mathbf{x}_0 . We assume the entries of \mathbf{x}_{init} are distributed as in (19) and define $\tilde{\rho} := \frac{1}{\sqrt{k}} \text{sign}(\mathbf{x}_0^{\Delta})^{\top} \mathbf{x}_{\text{init}}^{\Delta}$, where $\text{sign}(\cdot)$ denotes the component-wise sign function. Then, Assumption 1 holds with:*

$$\begin{aligned} F_{\lambda}(\beta) &= \beta(s + 2(1-s) \cdot Q(\frac{\lambda}{\sqrt{\beta^2 + \frac{\|\mathbf{x}_{\text{init}}^{\Delta^c}\|^2}{1-s}})}) , \\ G_{\lambda}^2(\beta) &= s \cdot (\beta^2 + \lambda^2) + \|\mathbf{x}_{\text{init}}^{\Delta}\|^2 - 2\lambda\sqrt{s}\tilde{\rho} - L^2(\mathbf{x}_0) \\ &\quad + (1-s)(\beta^2 + \frac{\|\mathbf{x}_{\text{init}}^{\Delta^c}\|^2}{1-s}) \cdot \mathbb{E}_H[\text{shrink}^2(H, \frac{\lambda}{\sqrt{\beta^2 + \frac{\|\mathbf{x}_{\text{init}}^{\Delta^c}\|^2}{1-s}})}] \end{aligned} \quad (20)$$

where $Q(\cdot)$ is the tail distribution of the standard normal distribution, H has standard normal distribution and $s := k/n$ is the sparsity factor. The shrinkage function $\text{shrink}(\cdot, \cdot) : \mathbb{R} \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is defined as:

$$\text{shrink}(x, \tau) = (|x| - \tau) \mathbf{1}\{|x| \geq \tau\}. \quad (21)$$

It is worth noting that the function $\text{shrink}(\cdot, \cdot)$ also appeared in computing the statistical dimension for ℓ_1 regularization (see Section 5.5) which indicates some implicit relation to α_{opt} .

We have numerically computed the solution of the nonlinear system (20). Figure 2(a), and Figure 2(b) shows the error of regularized PhaseMax over a range of ρ_{init} and δ . The comparison between our upper bound derived from Theorem 1 and precise analysis of Theorem 2 is depicted in Figure 3(a) for three values of the sparsity factor $s = 0.05, 0.1, 0.2$. Observe that the upper bound is only a constant factor away from the precise phase transition, while its derivation involves simpler formulas. Finally, Figure 3(b), illustrates impact of the regularization parameter λ on the phase transition of the regularized PhaseMax optimization for four values of ρ_{init} . The values of λ in this figure are normalized by $\frac{\rho_{\text{init}}\sqrt{n}}{\|\mathbf{x}_0\|}$, which is the maximum acceptable value of λ in the regularized PhaseMax.

4 Conclusion and Future Directions

In this paper, we introduced a new convex optimization framework, *regularized PhaseMax*, to solve the structured phase retrieval problem. We have shown that, given a proper initialization, the regularized PhaseMax optimization perfectly recovers the underlying signal from a number of phaseless measurements that is only a constant factor away from the number of measurements required when the phase information is available. We explicitly computed this constant factor.

An important (yet still open) research problem is to investigate the required sample complexity to construct a proper initialization vector, \mathbf{x}_{init} . As an example, for the case of sparse phase retrieval, even though our analysis indicates that $\mathcal{O}(k \log \frac{n}{k})$ is the required sample complexity of the regularized PhaseMax optimization, the best known initialization technique [6] needs $\mathcal{O}(k^2 \log n)$ samples to generate a meaningful initialization, which is suboptimal. An important future direction is to study initialization techniques that break this sample complexity barrier, or to use information theoretic arguments (as in [28]) to show that the sample complexity for the initialization cannot be improved.

To form the objective function in the regularized PhaseMax, we exploited some a-priori knowledge about the structure of the underlying signal. In many practical settings such prior information is not available. There has been some interesting recent publications (e.g. [4, 48]) which introduce efficient algorithms to learn the structure of the underlying signal. An interesting research direction is to investigate new optimization framework that does not rely on the prior information about the structure of the underlying signal.

References

- [1] Dennis Amelunxen, Martin Lotz, Michael B McCoy, and Joel A Tropp. Living on the edge: Phase transitions in convex programs with random data. *Information and Inference: A Journal of the IMA*, 3(3):224–294, 2014.
- [2] Sohail Bahmani and Justin Romberg. Efficient compressive phase retrieval with constrained sensing vectors. In *Advances in Neural Information Processing Systems*, pages 523–531, 2015.
- [3] Sohail Bahmani and Justin Romberg. Phase retrieval meets statistical learning theory: A flexible convex relaxation. In *Artificial Intelligence and Statistics*, pages 252–260, 2017.
- [4] Milad Bakhshizadeh, Arian Maleki, and Shirin Jalali. Compressive phase retrieval of structured signals. In *2018 IEEE International Symposium on Information Theory (ISIT)*, pages 2291–2295. IEEE, 2018.
- [5] Mohsen Bayati and Andrea Montanari. The lasso risk for gaussian matrices. *IEEE Transactions on Information Theory*, 58(4):1997–2017, 2012.
- [6] T Tony Cai, Xiaodong Li, Zongming Ma, et al. Optimal rates of convergence for noisy sparse phase retrieval via thresholded wirtinger flow. *The Annals of Statistics*, 44(5):2221–2251, 2016.
- [7] Emmanuel J Candes, Yonina C Eldar, Thomas Strohmer, and Vladislav Voroninski. Phase retrieval via matrix completion. *SIAM review*, 57(2):225–251, 2015.
- [8] Emmanuel J Candes, Xiaodong Li, and Mahdi Soltanolkotabi. Phase retrieval from coded diffraction patterns. *Applied and Computational Harmonic Analysis*, 39(2):277–299, 2015.
- [9] Emmanuel J Candes, Xiaodong Li, and Mahdi Soltanolkotabi. Phase retrieval via wirtinger flow: Theory and algorithms. *IEEE Transactions on Information Theory*, 61(4):1985–2007, 2015.
- [10] Emmanuel J Candes, Thomas Strohmer, and Vladislav Voroninski. Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming. *Communications on Pure and Applied Mathematics*, 66(8):1241–1274, 2013.
- [11] Venkat Chandrasekaran, Benjamin Recht, Pablo A Parrilo, and Alan S Willsky. The convex geometry of linear inverse problems. *Foundations of Computational mathematics*, 12(6):805–849, 2012.
- [12] Yuxin Chen and Emmanuel Candes. Solving random quadratic systems of equations is nearly as easy as solving linear systems. In *Advances in Neural Information Processing Systems*, pages 739–747, 2015.
- [13] Oussama Dhifallah and Yue M Lu. Fundamental limits of phasemax for phase retrieval: A replica analysis. In *Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), 2017 IEEE 7th International Workshop on*, pages 1–5. IEEE, 2017.
- [14] Oussama Dhifallah, Christos Thrampoulidis, and Yue M Lu. Phase retrieval via polytope optimization: Geometry, phase transitions, and new algorithms. *arXiv preprint arXiv:1805.09555*, 2018.
- [15] Martin Dierolf, Andreas Menzel, Pierre Thibault, Philipp Schneider, Cameron M Kewish, Roger Wepf, Oliver Bunk, and Franz Pfeiffer. Ptychographic x-ray computed tomography at the nanoscale. *Nature*, 467(7314):436–439, 2010.
- [16] David L Donoho, Arian Maleki, and Andrea Montanari. Message-passing algorithms for compressed sensing. *Proceedings of the National Academy of Sciences*, 106(45):18914–18919, 2009.
- [17] C Fienup and J Dainty. Phase retrieval and image reconstruction for astronomy. *Image Recovery: Theory and Application*, pages 231–275, 1987.
- [18] James R Fienup. Phase retrieval algorithms: a comparison. *Applied optics*, 21(15):2758–2769, 1982.

- [19] Tom Goldstein and Christoph Studer. Phasemax: Convex phase retrieval via basis pursuit. *IEEE Transactions on Information Theory*, 2018.
- [20] Paul Hand and Vladislav Voroninski. Compressed sensing from phaseless gaussian measurements via linear programming in the natural parameter space. *arXiv preprint arXiv:1611.05985*, 2016.
- [21] Paul Hand and Vladislav Voroninski. An elementary proof of convex phase retrieval in the natural parameter space via the linear program phasemax. *arXiv preprint arXiv:1611.03935*, 2016.
- [22] Kishore Jaganathan, Yonina C Eldar, and Babak Hassibi. Phase retrieval: An overview of recent developments. *arXiv preprint arXiv:1510.07713*, 2015.
- [23] Kishore Jaganathan, Samet Oymak, and Babak Hassibi. Sparse phase retrieval: Uniqueness guarantees and recovery algorithms. *IEEE Transactions on Signal Processing*, 65(9):2402–2410, 2017.
- [24] Xiaodong Li and Vladislav Voroninski. Sparse signal recovery from quadratic measurements via convex programming. *SIAM Journal on Mathematical Analysis*, 45(5):3019–3033, 2013.
- [25] Yue M Lu and Gen Li. Phase transitions of spectral initialization for high-dimensional nonconvex estimation. *arXiv preprint arXiv:1702.06435*, 2017.
- [26] Junjie Ma, Ji Xu, and Arian Maleki. Optimization-based amp for phase retrieval: The impact of initialization and ℓ_2 -regularization. *arXiv preprint arXiv:1801.01170*, 2018.
- [27] Rick P Millane. Phase retrieval in crystallography and optics. *JOSA A*, 7(3):394–411, 1990.
- [28] Marco Mondelli and Andrea Montanari. Fundamental limits of weak recovery with applications to phase retrieval. *arXiv preprint arXiv:1708.05932*, 2017.
- [29] Praneeth Netrapalli, Prateek Jain, and Sujay Sanghavi. Phase retrieval using alternating minimization. In *Advances in Neural Information Processing Systems*, pages 2796–2804, 2013.
- [30] Samet Oymak, Amin Jalali, Maryam Fazel, Yonina C Eldar, and Babak Hassibi. Simultaneously structured models with application to sparse and low-rank matrices. *IEEE Transactions on Information Theory*, 61(5):2886–2908, 2015.
- [31] Samet Oymak, Christos Thrampoulidis, and Babak Hassibi. The squared-error of generalized lasso: A precise analysis. *arXiv preprint arXiv:1311.0830*, 2013.
- [32] Ralph Tyrell Rockafellar. *Convex analysis*. Princeton university press, 2015.
- [33] Mark Rudelson and Roman Vershynin. Sparse reconstruction by convex relaxation: Fourier and gaussian measurements. In *Information Sciences and Systems, 2006 40th Annual Conference on*, pages 207–212. IEEE, 2006.
- [34] Walter Rudin et al. *Principles of mathematical analysis*, volume 3. McGraw-hill New York, 1976.
- [35] Fariborz Salehi, Ehsan Abbasi, and Babak Hassibi. A precise analysis of phasemax in phase retrieval. In *Information Theory (ISIT), 2018 IEEE International Symposium on*, pages 976–980. IEEE, 2018.
- [36] Fariborz Salehi, Kishore Jaganathan, and Babak Hassibi. Multiple illumination phaseless super-resolution (mips) with applications to phaseless doa estimation and diffraction imaging. In *Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on*, pages 3949–3953. IEEE, 2017.
- [37] Mahdi Soltanolkotabi. Structured signal recovery from quadratic measurements: Breaking sample complexity barriers via nonconvex optimization. *arXiv preprint arXiv:1702.06175*, 2017.
- [38] Mihailo Stojnic. Various thresholds for ℓ_1 -optimization in compressed sensing. 2009.

- [39] Christos Thrampoulidis, Ehsan Abbasi, and Babak Hassibi. Precise error analysis of regularized m-estimators in high-dimensions. *IEEE Transactions on Information Theory*, 2018.
- [40] Christos Thrampoulidis, Samet Oymak, and Babak Hassibi. Regularized linear regression: A precise analysis of the estimation error. In *Conference on Learning Theory*, pages 1683–1709, 2015.
- [41] Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288, 1996.
- [42] Joel A Tropp. Convex recovery of a structured signal from independent random linear measurements. In *Sampling Theory, a Renaissance*, pages 67–101. Springer, 2015.
- [43] Joel A Tropp and Anna C Gilbert. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Transactions on information theory*, 53(12):4655–4666, 2007.
- [44] Roman Vershynin. High-dimensional probability. *An Introduction with Applications*, 2016.
- [45] Adriaan Walther. The question of phase retrieval in optics. *Journal of Modern Optics*, 10(1):41–49, 1963.
- [46] Gang Wang, Georgios B Giannakis, and Yonina C Eldar. Solving systems of random quadratic equations via truncated amplitude flow. *IEEE Transactions on Information Theory*, 64(2):773–794, 2018.
- [47] Gang Wang, Liang Zhang, Georgios B Giannakis, Mehmet Akcakaya, and Jie Chen. Sparse phase retrieval via truncated amplitude flow. *IEEE Transactions on Signal Processing*, 66(2):479–491, 2016.
- [48] Shanshan Wu, Alexandros G Dimakis, Sujay Sanghavi, Felix X Yu, Daniel Holtmann-Rice, Dmitry Storchus, Afshin Rostamizadeh, and Sanjiv Kumar. The sparse recovery autoencoder. *arXiv preprint arXiv:1806.10175*, 2018.
- [49] Teng Zhang. Phase retrieval using alternating minimization in a batch setting. *arXiv preprint arXiv:1706.08167*, 2017.

5 Appendix

5.1 Convex Gaussian min-max theorem

Our analysis is based on the recently developed Convex Gaussian Min-max Theorem (CGMT) [39, 40]. The CGMT associates with a Primary Optimization (PO) problem an Auxiliary Optimization (AO) problem from which we can investigate various properties of the primary optimization, such as phase transitions. In particular, the (PO) and the (AO) problems are defined respectively as follows:

$$\Phi(\mathbf{G}) := \min_{\mathbf{w} \in \mathbf{S}_{\mathbf{w}}} \max_{\mathbf{u} \in \mathbf{S}_{\mathbf{u}}} \mathbf{u}^T \mathbf{G} \mathbf{w} + \psi(u, w), \quad (22a)$$

$$\phi(\mathbf{g}, \mathbf{h}) := \min_{\mathbf{w} \in \mathbf{S}_{\mathbf{w}}} \max_{\mathbf{u} \in \mathbf{S}_{\mathbf{u}}} \|\mathbf{w}\| \mathbf{g}^T \mathbf{u} - \|\mathbf{u}\| \mathbf{h}^T \mathbf{w} + \psi(u, w), \quad (22b)$$

where $\mathbf{G} \in \mathbb{R}^{m \times n}$, $\mathbf{g} \in \mathbb{R}^m$, $\mathbf{h} \in \mathbb{R}^n$, $\mathbf{S}_{\mathbf{w}} \subset \mathbb{R}^n$, $\mathbf{S}_{\mathbf{u}} \subset \mathbb{R}^m$ and $\psi : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$. Denote $\mathbf{w}_{\Phi} := \mathbf{w}_{\Phi}(\mathbf{G})$ and $\mathbf{w}_{\phi} := \mathbf{w}_{\phi}(\mathbf{g}, \mathbf{h})$ any optimal minimizers in (22a) and (22b), respectively. The following lemma is a result of CGMT [39].

Lemma 2. *Consider the two optimizations (22a) and (22b). Let $\mathbf{S}_{\mathbf{w}}, \mathbf{S}_{\mathbf{u}}$ be convex sets where at least one of them is compact, ψ be continuous and convex-concave on $\mathbf{S}_{\mathbf{w}} \times \mathbf{S}_{\mathbf{u}}$, and, \mathbf{G}, \mathbf{g} and \mathbf{h} all have entries iid standard normal. Suppose there exist a unique α such that in the limit of $n \rightarrow \infty$ it holds in probability that $\|\mathbf{w}_{\phi}(\mathbf{g}, \mathbf{h})\| \rightarrow \alpha$. Then, the same holds for $\mathbf{w}_{\Phi}(\mathbf{G})$ and we have $\|\mathbf{w}_{\Phi}(\mathbf{G})\| \rightarrow \alpha$.*

CGMT essentially replaces the optimization in (22a) with the one in (22b) (which we refer to as the Auxiliary Optimization) which is simpler to analyze. This is the main step in our proofs.

5.1.1 Convergence analysis of the Auxiliary Optimization

After applying CGMT in the proof of Theorem 2, we aim to analyze the auxiliary optimization. In this path, we replace several functions with their limits in probability. This can be done through the same tricks used in section A.4 of [39] and Lemma B.1 in the same paper. Here, we state the following lemma without proof.

Lemma 3 (Min-convergence – Open Sets). *Consider a sequence of proper, convex stochastic functions $M_n : (0, \infty) \rightarrow \mathcal{R}$, and, a deterministic function $M : (0, \infty) \rightarrow \mathcal{R}$, such that:*

1. $M_n(x) \xrightarrow{P} M(x)$, for all $x > 0$,
2. there exists $z > 0$ such that $M(x) > \inf_{x>0} M(x)$ for all $x \geq z$.

Then, $\inf_{x>0} M_n(x) \xrightarrow{P} \inf_{x>0} M(x)$.

The objective function in our optimization problems satisfy the assumptions of this lemma at the points that we replace them with their limits.

5.2 Proof of Theorem 1

In order to establish the result we use the following lemma which provides an equivalent optimization that has the same error performance as PhaseMax, and is the key ingredient in deriving the main results of the paper.

Lemma 4 (Equivalent Optimization). *Consider the regularized PhaseMax problem introduced in Section 2.1. As $n \rightarrow \infty$, the error performance converges in probability as follows:*

$$\left(\frac{\|\hat{\mathbf{x}} - \mathbf{x}_0\|^2}{\|\mathbf{x}_0\|^2} \right) - (\|\mathbf{w}^*\|^2 + (1 - s^*)^2) \xrightarrow{n \rightarrow \infty} 0. \quad (23)$$

Here $s^* \in \mathbb{R}$, and $\mathbf{w}^* \in \mathbb{R}^n$ are the unique optimizers of the following optimization program,

$$\begin{aligned} \min_{s \in \mathbb{R}} \min_{\mathbf{w} \in \mathbb{R}^n, \mathbf{w} \perp \mathbf{x}_0} & -\mathbf{x}_{init}^T (s\mathbf{x}_0 + \mathbf{w}) + \lambda f(s\mathbf{x}_0 + \mathbf{w}) \\ \text{subject to:} & \mathbf{h}^T \mathbf{w} \geq \sqrt{m \mathbf{c}_d(s, \|\mathbf{w}\|)} \end{aligned} \quad (24)$$

where $\mathbf{h} \in \mathbb{R}^n$ has i.i.d. standard normal entries and the function $\mathbf{c}_d : \mathbb{R} \times \mathbb{R}_+ \rightarrow \mathbb{R}$ is defined as,

$$\mathbf{c}_d(s, r) = \frac{1}{\pi} [((1+s)^2 + r^2) \operatorname{atan}\left(\frac{r}{1+s}\right) + ((1-s)^2 + r^2) \operatorname{atan}\left(\frac{r}{1-s}\right) - 2r]. \quad (25)$$

The full technical details of obtaining this result is explained in Section 5.3. In short, to show the equivalence, we start from (2) and define new variables $s := \mathbf{x}_0^\top \mathbf{x}$ and $\mathbf{w} = \mathbf{P}^\perp \mathbf{x}$. Then reformulate it as an unconstrained optimization using Lagrange multipliers. The result is a consequence of applying CGMT (Lemma 2, see Appendix 5.1) with some simplifications.

Consider the following optimization:

$$\begin{aligned} \min_{s \in \mathbb{R}} \min_{\mathbf{w} \in \mathbb{R}^n, \mathbf{w} \perp \mathbf{x}_0} \quad & -\mathbf{x}_{\text{init}}^\top (s\mathbf{x}_0 + \mathbf{w}) + \lambda f(s\mathbf{x}_0 + \mathbf{w}) \\ \text{subject to:} \quad & \mathbf{h}^\top \mathbf{w} \geq \sqrt{m \mathbf{c}_d(s, \|\mathbf{w}\|)}. \end{aligned} \quad (26)$$

The result of Lemma 4 established that as $n \rightarrow \infty$, the error performance of the regularized PhaseMax converges to the error performance in (26). The following corollary indicates the necessary and sufficient condition for perfect recovery:

Corollary 2. *As $n \rightarrow \infty$, \mathbf{x}_0 is the unique solution of the regularized PhaseMax optimization, if and only if $(s^*, \mathbf{w}^*) = (1, \mathbf{0})$ be the unique optimizer of the equivalent optimization (26).*

Proof. This is an immediate result of Lemma 4, noticing that $\|\hat{\mathbf{x}} - \mathbf{x}_0\| = 0$ is the condition for perfect recovery. \square

Let us proceed onwards with analyzing (26). For simplicity, we assume $\|\mathbf{x}_0\| = 1$. Define a new function $\hat{f}(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}$ as follows,

$$\hat{f}(\mathbf{x}) = f(\mathbf{x}_0) + \max_{\mathbf{v} \in \partial f(\mathbf{x}_0)} \mathbf{v}^\top (\mathbf{x} - \mathbf{x}_0). \quad (27)$$

$\partial f(\mathbf{x}_0)$ is the sub-differential set of function f at point \mathbf{x}_0 which is a convex and compact set. $\hat{f}(\cdot)$ is basically the first-order approximation of the regularization function $f(\cdot)$ at point \mathbf{x}_0 . This replacement cannot be done in general, but since we are only investigating the phase transition regime where the norm of the error, $\|\hat{\mathbf{x}} - \mathbf{x}_0\|$, approaches to zero, we may perform this exchange. To investigate the phase transition behavior in (26), we bound $1 - s$ and $\|\mathbf{w}\|$ to a small neighborhood of 0. Therefore, it is valid to replace f with \hat{f} in that small neighborhood around \mathbf{x}_0 . Reformulating the optimization using this replacement would give us the following,

$$\begin{aligned} \min_{s \in \mathbb{R}, \mathbf{w} \perp \mathbf{x}_0} \max_{\mathbf{v} \in \partial f(\mathbf{x}_0)} \quad & -\mathbf{x}_{\text{init}}^\top (s\mathbf{x}_0 + \mathbf{w}) + \lambda \hat{f}(s\mathbf{x}_0 + \mathbf{w}) + \lambda \mathbf{v}^\top ((s-1)\mathbf{x}_0 + \mathbf{w}) \\ \text{subject to:} \quad & \mathbf{h}^\top \mathbf{w} \geq \sqrt{m \mathbf{c}_d(s, \|\mathbf{w}\|)}. \end{aligned} \quad (28)$$

We add the constant term, $\mathbf{x}_{\text{init}}^\top \mathbf{x}_0 - \lambda f(\mathbf{x}_0)$, to the objective function and reformulate the maximization in terms of $\partial L(\mathbf{x}_0)$ as follows,

$$\begin{aligned} \min_{s \in \mathbb{R}, \mathbf{w} \perp \mathbf{x}_0} \max_{\mathbf{v} \in \partial L(\mathbf{x}_0)} \quad & (s-1)\mathbf{x}_0^\top \mathbf{v} + \mathbf{w}^\top \mathbf{v} \\ \text{subject to:} \quad & \mathbf{h}^\top \mathbf{w} \geq \sqrt{m \mathbf{c}_d(s, \|\mathbf{w}\|)}. \end{aligned} \quad (29)$$

If $|s| > 1$ in (29), we have the following inequalities:

$$\|\mathbf{w}\|^2 d(T_L(\mathbf{x}_0)) \geq (\mathbf{h}^\top \mathbf{w})^2 \geq m \mathbf{c}_d(s, \|\mathbf{w}\|) > \frac{m}{2} \|\mathbf{w}\|^2. \quad (30)$$

The first inequality is due to the fact that $\mathbf{x} - \mathbf{x}_0 = (s-1)\mathbf{x}_0 + \mathbf{w}$ is in $T_L(\mathbf{x}_0)$ (the descent cone of the objective at point \mathbf{x}_0). The second inequality appeared as a constraint in the optimization

problem (29). The last inequality is true since $\mathbf{c}_d(s, r) > r^2/2$, when $|s| \geq 1$. Therefore, using the assumption $\delta = \frac{m}{d} > 2$, it can be shown that the feasible set of (29) is nonempty if and only if $|s| \leq 1$.

Since the regularized PhaseMax optimization is convex, in order to show that $s^* = 1$ and $\mathbf{w}^* = \mathbf{0}$ are the unique optimizers of (29), it is sufficient to check the optimality condition in a small neighborhood of $(s^* = 1, \mathbf{w}^* = \mathbf{0})$. We also use the following approximation of the function $\mathbf{c}_d(s, r)$ which is valid in a small neighborhood around the point $(s, r) = (1, 0)$:

$$\mathbf{c}_d(s, r) = \frac{1}{\pi} [((1-s)^2 + r^2) \text{atan}\left(\frac{r}{1-s}\right) - r(1-s)] . \quad (31)$$

Next, for fixed $|s| < 1$, we will find an upper bound for $r := \|\mathbf{w}\|$ such that s and \mathbf{w} satisfy the constraint in (29). We use the following inequalities:

$$r^2 d(T_L(\mathbf{x}_0)) \geq (\mathbf{h}^\top \mathbf{w})^2 \geq m \mathbf{c}_d(s, r) \Rightarrow r^2 \geq \delta \mathbf{c}_d(s, r) . \quad (32)$$

Replacing the approximation (31) for $\mathbf{c}_d(s, r)$, when $s \uparrow 1$ we have,

$$r \leq R(\delta)(1-s) , \quad (33)$$

where $R(\delta)$ is the unique nonzero solution of the following nonlinear equation:

$$t^2 = \frac{\delta}{\pi} ((1+t^2) \text{atan}(t) - t) . \quad (34)$$

We are now at the stage to establish the result of Theorem 1. Assume $\tilde{\mathbf{v}} \in \partial L(\mathbf{x}_0)$ achieves the supremum in (7) (Note that $\tilde{\mathbf{v}}$ always exist because the set $\partial L(\mathbf{x}_0)$ is compact). $\tilde{\mathbf{v}}$ then satisfies the following conditions:

1. $\mathbf{x}_0^\top \tilde{\mathbf{v}} < 0$,
2. $\|\mathbf{P}\tilde{\mathbf{v}}\| > R(\delta) \|\mathbf{P}^\perp \tilde{\mathbf{v}}\|$.

We have the following inequalities:

$$\begin{aligned} \min_{|s| \leq 1, \mathbf{w} \perp \mathbf{x}_0} \max_{\mathbf{v} \in \partial L(\mathbf{x}_0)} (s-1)\mathbf{x}_0^\top \mathbf{v} + \mathbf{w}^\top \mathbf{v} &\geq \min_{|s| \leq 1, \mathbf{w} \perp \mathbf{x}_0} (s-1)\mathbf{x}_0^\top \tilde{\mathbf{v}} + \mathbf{w}^\top \tilde{\mathbf{v}} \\ &\geq \min_{|s| \leq 1, \mathbf{w} \perp \mathbf{x}_0} (1-s) \|\mathbf{P}\tilde{\mathbf{v}}\| - \|\mathbf{w}\| \|\mathbf{P}^\perp \tilde{\mathbf{v}}\| , \end{aligned} \quad (35)$$

where for the first inequality, we used the fact that maximization over \mathbf{v} gives a larger value compare to choosing the specific vector $\tilde{\mathbf{v}}$. The second inequality we used Cauchy-Schwarz to bound $\mathbf{w}^\top \mathbf{v}$ from below. When $s \uparrow 1$ we use the approximation (32) which bounds $\|\mathbf{w}\|$ from above. Therefore, we have:

$$(1-s) \|\mathbf{P}\tilde{\mathbf{v}}\| - \|\mathbf{w}\| \|\mathbf{P}^\perp \tilde{\mathbf{v}}\| > (1-s)(\|\mathbf{P}\tilde{\mathbf{v}}\| - R(\delta) \|\mathbf{P}^\perp \tilde{\mathbf{v}}\|) > 0. \quad (36)$$

And this would give the final result that $s^* = 1, \mathbf{w}^* = \mathbf{0}$ is the unique solution of (29). The perfect recovery in the generalized PhaseMax follows from the result of Corollary 2.

5.3 Proof of Lemma 4

Define matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ with i^{th} row equal to the measurement vector \mathbf{a}_i , for $i = 1, 2, \dots, m$. Let $\mathbf{y} := |\mathbf{A}\mathbf{x}_0| \in \mathbb{R}^m$ denote the measurement values. To streamline our analysis, we assume $\|\mathbf{x}_0\| = 1$. One can rewrite the constraint set of the optimization problem (2) as followings,

$$|\mathbf{A}\mathbf{x}| \leq \mathbf{y} \Leftrightarrow -\mathbf{A}\mathbf{x} + \mathbf{y} \geq \mathbf{0} , \text{ and } \mathbf{A}\mathbf{x} + \mathbf{y} \geq \mathbf{0}, \quad (37)$$

where all the inequalities are component-wise. Exploiting the Lagrange multipliers, we can reformulate the generalized PhaseMax optimization as,

$$\min_{\mathbf{x} \in \mathbb{R}^n} \max_{\boldsymbol{\mu}, \boldsymbol{\eta} \in \mathbb{R}_+^m} -\mathbf{x}_{\text{init}}^\top \mathbf{x} + \lambda f(\mathbf{x}) + (\boldsymbol{\mu} - \boldsymbol{\eta})^\top \mathbf{A} \mathbf{x} - (\boldsymbol{\mu} + \boldsymbol{\eta})^\top \mathbf{y}, \quad (38)$$

where μ_i and η_i are Lagrange multipliers for the inequalities $\mathbf{a}_i^\top \mathbf{x} \leq y_i$ and $\mathbf{a}_i^\top \mathbf{x} \geq -y_i$, respectively. Assume $y_i > 0$ (which happens almost surely), these two inequalities cannot be active at the same time. Therefore, at least one of μ_i and η_i must be equal to 0, for every $i = 1, 2, \dots, m$. Hence, we have $\boldsymbol{\mu} + \boldsymbol{\eta} = |\boldsymbol{\mu} - \boldsymbol{\eta}|$. Here $|\cdot|$ denotes the component-wise absolute value function. Define $\mathbf{v} := \boldsymbol{\mu} - \boldsymbol{\eta} \in \mathbb{R}^m$ and rewrite the optimization in terms of \mathbf{v} gives the following,

$$\min_{\mathbf{x} \in \mathbb{R}^n} \max_{\mathbf{v} \in \mathbb{R}^m} -\mathbf{x}_{\text{init}}^\top \mathbf{x} + \lambda f(\mathbf{x}) + \mathbf{v}^\top \mathbf{A} \mathbf{x} - |\mathbf{v}|^\top |\mathbf{A} \mathbf{x}_0|. \quad (39)$$

Since the term $|\mathbf{v}|^\top |\mathbf{A} \mathbf{x}_0|$ depends on the matrix \mathbf{A} , it is not possible to apply the CGMT to the bilinear form $\mathbf{v}^\top \mathbf{A} \mathbf{x}$. In order to apply CGMT, we use the following key decomposition for \mathbf{x} :

$$\mathbf{x} = s \mathbf{x}_0 + \mathbf{w}, \quad (40)$$

where $s = \mathbf{x}_0^\top \mathbf{x} \in \mathbb{R}$ is a scalar and the vector $\mathbf{w} = \mathbf{P}^\perp \mathbf{x} \in \mathbb{R}^n$ is orthogonal to \mathbf{x}_0 . We can rewrite the optimization problem (39) in terms of s and \mathbf{w} ,

$$\min_{s \in \mathbb{R}, \mathbf{w} \perp \mathbf{x}_0} \max_{\mathbf{v} \in \mathbb{R}^m} -s \rho_{\text{init}} - \mathbf{x}_{\text{init}}^\top \mathbf{w} + \lambda f(s \mathbf{x}_0 + \mathbf{w}) + \mathbf{v}^\top \mathbf{A} \mathbf{w} + s \mathbf{v}^\top \mathbf{A} \mathbf{x}_0 - |\mathbf{v}|^\top |\mathbf{A} \mathbf{x}_0|, \quad (41)$$

where $\rho_{\text{init}} = \mathbf{x}_{\text{init}}^\top \mathbf{x}_0$. Next, we use the following property of Gaussian matrices.

Lemma 5. Let $\mathbf{G} \in \mathbb{R}^{m \times n}$ be a random matrix with i.i.d. standard normal entries, and $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ are such that $\mathbf{u} \perp \mathbf{v}$. The random vectors $\mathbf{G} \mathbf{u}$ and $\mathbf{G} \mathbf{v}$ are independent.

Proof. Let $\mathbf{G} = [g_{i,j}]_{m \times n}$ and define $\mathbf{a} = \mathbf{G} \mathbf{u}$, and $\mathbf{b} = \mathbf{G} \mathbf{v}$. Since \mathbf{G} has Gaussian entries \mathbf{a}, \mathbf{b} are Gaussian vectors in \mathbb{R}^m . Therefore, to show their independence it is sufficient to show that $\mathbb{E}[\mathbf{a} \mathbf{b}^\top] = \mathbf{0}_{m \times m}$.

$$\mathbb{E}[a_i b_j] = \sum_{k=1}^n \sum_{l=1}^n u_k v_l \mathbb{E}[g_{i,k} g_{j,l}] = \begin{cases} \sum_{k=1}^n u_k v_k = 0, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}, \quad (42)$$

where we used the fact that $\mathbf{u}^\top \mathbf{v} = \sum_{k=1}^n u_k v_k = 0$.

□

Using the result of Lemma 5, the random vectors $\mathbf{A} \mathbf{x}_0$ and $\mathbf{A} \mathbf{w}$ are independent. So, we are allowed to change the matrix \mathbf{A} in the bilinear form $\mathbf{v}^\top \mathbf{A} \mathbf{w}$ with its independent copy $\mathbf{H} \in \mathbb{R}^{m \times n}$ which also has i.i.d. standard normal entries. We also define $\mathbf{q} = \mathbf{A} \mathbf{x}_0 \in \mathbb{R}^m$, which is independent of \mathbf{H} . Note that since \mathbf{A} has i.i.d. normal entries and $\|\mathbf{x}_0\| = 1$, the entries of \mathbf{q} also has i.i.d. standard normal distribution. We can rewrite the optimization as follows:

$$\min_{s \in \mathbb{R}, \mathbf{w} \perp \mathbf{x}_0} \max_{\mathbf{v} \in \mathbb{R}^m} -s \rho_{\text{init}} - \mathbf{x}_{\text{init}}^\top \mathbf{w} + \lambda f(s \mathbf{x}_0 + \mathbf{w}) + \mathbf{v}^\top \mathbf{H} \mathbf{w} + s \mathbf{v}^\top \mathbf{q} - |\mathbf{v}|^\top |\mathbf{q}|. \quad (43)$$

Now, we would like to apply the CGMT framework in Lemma 2 to equation (43), in order to replace the bilinear form $\mathbf{v}^\top \mathbf{H} \mathbf{w}$ with two linear forms $\|\mathbf{v}\| \mathbf{h}^\top \mathbf{w} + \mathbf{v}^\top \mathbf{g} \|\mathbf{w}\|$. But this lemma requires the set that we optimize \mathbf{w} over to be compact. In order to be able to apply CGMT, we enforce an "artificial" bound on the norm of \mathbf{w} . Note that our goal is to prove that eventually, $\hat{\mathbf{w}}$ converges to a finite number α^* . We define $K_\alpha = \alpha^* + \Delta$ for some $\Delta > 0$ and also the compact set $\mathcal{S}_\mathbf{w} = \{\mathbf{w} | \mathbf{w} \perp \mathbf{x}_0, \|\mathbf{w}\| \leq K_\alpha\}$. Let $\hat{\mathbf{w}}^{\text{temp}}$ to be the optimizer to the version of (43) where we

optimize \mathbf{w} over $\mathcal{S}_{\mathbf{w}}$. It is simple to verify that if $\|\hat{\mathbf{w}}^{\text{temp}}\| \xrightarrow{\mathbb{P}} \alpha^*$, then $\|\hat{\mathbf{w}}\| \xrightarrow{\mathbb{P}} \alpha^*$. This means that if in the final equation, we get a unique finite solution for the asymptotic behaviour of $\|\hat{\mathbf{w}}\|$ (which is what we do), the proof goes through and we can apply CGMT.

Now that this concern is taken care of, the following corollary will be the result of applying CGMT to the equation (43).

Corollary 3. *Let $\hat{\mathbf{x}}$ be the unique optimizer of the generalized PhaseMax algorithm (2). As $n \rightarrow \infty$ the error performance converges in probability as follows:*

$$\left(\frac{\|\hat{\mathbf{x}} - \mathbf{x}_0\|^2}{\|\mathbf{x}_0\|^2} \right) - (\|\mathbf{w}^*\|^2 + (1 - s^*)^2) \xrightarrow{n \rightarrow \infty} 0, \quad (44)$$

where s^*, \mathbf{w}^* are the unique optimizers of the following (auxiliary) optimization:

$$\min_{s \in \mathbb{R}, \mathbf{w} \perp \mathbf{x}_0} \max_{\mathbf{v} \in \mathbb{R}^m} -s\rho_{\text{init}} - \mathbf{x}_{\text{init}}^T \mathbf{w} + \lambda f(s\mathbf{x}_0 + \mathbf{w}) - \|\mathbf{v}\| \|\mathbf{h}^T \mathbf{w} + \mathbf{v}^T \mathbf{g}\| \|\mathbf{w}\| + s\mathbf{v}^T \mathbf{q} - |\mathbf{v}^T \mathbf{q}|. \quad (45)$$

$\mathbf{h} \in \mathbb{R}^n$ and $\mathbf{g} \in \mathbb{R}^m$ are random vectors with i.i.d. standard normal entries.

We proceed onwards with analyzing (45). Observe that if we fix $|\mathbf{v}|$, then the optimal \mathbf{v} satisfies $\text{sign}(\mathbf{v}) = \text{sign}(\|\mathbf{w}\|\mathbf{g} + s\mathbf{q})$ ($\text{sign}(\cdot)$ are component-wise functions) which simplifies the optimization to the following,

$$\min_{s \in \mathbb{R}, \mathbf{w} \perp \mathbf{x}_0} \max_{\mathbf{v} \in \mathbb{R}^m} -s\rho_{\text{init}} - \mathbf{x}_{\text{init}}^T \mathbf{w} + \lambda f(s\mathbf{x}_0 + \mathbf{w}) - \|\mathbf{v}\| \|\mathbf{h}^T \mathbf{w} + |\mathbf{v}|^T (\|\mathbf{q}\| + \|\mathbf{w}\|\mathbf{g} - |\mathbf{q}|\|), \quad (46)$$

By fixing the norm of \mathbf{v} and optimizing over its direction, the optimization problem (46) can be reduced to the following:

$$\begin{aligned} \min_{s \in \mathbb{R}, \mathbf{w} \perp \mathbf{x}_0} & -s\rho_{\text{init}} - \mathbf{x}_{\text{init}}^T \mathbf{w} + \lambda f(s\mathbf{x}_0 + \mathbf{w}) \\ \text{subject to: } & \mathbf{h}^T \mathbf{w} \geq \|\{s\mathbf{q} + \|\mathbf{w}\|\mathbf{g} - |\mathbf{q}|\}_+\|, \end{aligned} \quad (47)$$

where, for a vector \mathbf{c} , we let $\{\mathbf{c}\}_+$ denote the component-wise positive part function, with i^{th} entry equal to $\max(0, c_i)$. Next, note that \mathbf{q} and \mathbf{g} are independent vectors in \mathbb{R}^m with i.i.d. standard normal entries. We introduce the function $\mathbf{c}_{\mathbf{d}}(s, r)$ as follows:

Definition 5. *The function $\mathbf{c}_{\mathbf{d}} : \mathbb{R} \times \mathbb{R}_+ \rightarrow \mathbb{R}$ is defined as,*

$$\mathbf{c}_{\mathbf{d}}(s, r) = \mathbb{E}_{\mathbf{X}_1, \mathbf{X}_2} [\{|s\mathbf{X}_1 + r\mathbf{X}_2| - |\mathbf{X}_1|\}_+^2], \quad (48)$$

where $\mathbf{X}_1, \mathbf{X}_2 \stackrel{i.i.d.}{\sim} \mathcal{N}(0, 1)$.

The next result creates a connection between the norm of the vector $\{|s\mathbf{q} + \|\mathbf{w}\|\mathbf{g} - |\mathbf{q}|\}_+$, which appears in the constraint of (47), and the function $\mathbf{c}_{\mathbf{d}}$.

Lemma 6. $\frac{1}{m} \|\{s\mathbf{q} + \|\mathbf{w}\|\mathbf{g} - |\mathbf{q}|\}_+\|^2 \xrightarrow{\mathbb{P}} \mathbf{c}_{\mathbf{d}}(s, \|\mathbf{w}\|)$, when $m \rightarrow \infty$.

Proof. Define the vector $\mathbf{u} \in \mathbb{R}_+^m$ as,

$$\mathbf{u} := \{|s\mathbf{q} + \|\mathbf{w}\|\mathbf{g} - |\mathbf{q}|\}_+. \quad (49)$$

The entries of \mathbf{u} are i.i.d and $\mathbb{E}[u_i^2] = \mathbf{c}_{\mathbf{d}}(s, \|\mathbf{w}\|)$, for $1 \leq i \leq m$. Therefore, the weak law of large number gives the following:

$$\frac{1}{m} \|\mathbf{u}\|^2 = \frac{1}{m} \sum_{i=1}^m u_i^2 \xrightarrow{\mathbb{P}} \mathbb{E}[u_i^2] = \mathbf{c}_{\mathbf{d}}(s, \|\mathbf{w}\|). \quad (50)$$

□

To conclude the proof of Lemma 4, we exploit the result of Lemma 6 to replace $\| \{ |s\mathbf{q}| + \|\mathbf{w}\|\mathbf{g}| - |\mathbf{q}| \}_+ \|$ in (47), which gives us the following optimization:

$$\begin{aligned} \min_{s \in \mathbb{R}} \min_{\mathbf{w} \in \mathbb{R}^n, \mathbf{w} \perp \mathbf{x}_0} & -\mathbf{x}_{\text{init}}^\top (s\mathbf{x}_0 + \mathbf{w}) + \lambda f(s\mathbf{x}_0 + \mathbf{w}) \\ \text{subject to:} & \mathbf{h}^\top \mathbf{w} \geq \sqrt{m \mathbf{c}_d(s, \|\mathbf{w}\|)}. \end{aligned} \quad (51)$$

Due to lack of space, we are not going through the technical details of obtaining the convergence result in (51). The point-wise convergence of the objective functions, for fixed values of s and $\|\mathbf{w}\|$, follows from Lemma 6. To show the uniform convergence, we appeal to the convexity of the objective function. The corresponding convergence of the optimal cost follows from the uniform convergence.

The following lemma gives an explicit formula for the function \mathbf{c}_d in terms of its two input arguments.

Lemma 7.

$$\mathbf{c}_d(s, t) = \frac{1}{\pi} [((1+s)^2 + t^2) \operatorname{atan}\left(\frac{t}{1+s}\right) + ((1-s)^2 + t^2) \operatorname{atan}\left(\frac{t}{1-s}\right) - 2t]. \quad (52)$$

Proof.

$$\mathbf{c}_d(s, t) = \mathbb{E}[\{|s\mathbf{X}_1 + t\mathbf{X}_2| - |\mathbf{X}_1|\}_+^2] \quad (53)$$

$$\begin{aligned} &= \frac{1}{\pi} \int_0^\infty e^{-x_1^2/2} \int_{\frac{1-s}{t}x_1}^\infty e^{-x_2^2/2} (tx_2 - (1-s)x_1)^2 dx_2 dx_1 \\ &+ \frac{1}{\pi} \int_0^\infty e^{-x_1^2/2} \int_{-\infty}^{-\frac{1+s}{t}x_1} e^{-x_2^2/2} (tx_2 + (1+s)x_1)^2 dx_2 dx_1, \end{aligned} \quad (54)$$

where, due to symmetry, we have computed the expectation only for $\mathbf{X}_1 > 0$ and multiplied the result by two. In order to compute the integral change of variable and use the polar coordinates (r, θ) in the x_1x_2 -plane. Recall that we have $x_1 = r \cos(\theta)$, and $x_2 = r \sin(\theta)$. Applying the change of variables would result the following:

$$\mathbf{c}_d(s, t) = \frac{1}{\pi} \int_{\operatorname{atan}(\frac{1-s}{t})}^{\pi/2} \int_0^\infty r^3 e^{-r^2/2} (t \sin(\theta) - (1-s) \cos(\theta))^2 dr d\theta \quad (55)$$

$$\begin{aligned} &+ \frac{1}{\pi} \int_{\operatorname{atan}(\frac{1+s}{t})}^{\pi/2} \int_0^\infty r^3 e^{-r^2/2} (t \sin(\theta) - (1+s) \cos(\theta))^2 dr d\theta \\ &= \frac{2}{\pi} \int_{\operatorname{atan}(\frac{1-s}{t})}^{\pi/2} (t \sin(\theta) - (1-s) \cos(\theta))^2 d\theta + \frac{2}{\pi} \int_{\operatorname{atan}(\frac{1+s}{t})}^{\pi/2} (t \sin(\theta) - (1+s) \cos(\theta))^2 d\theta \end{aligned} \quad (56)$$

$$= \frac{2}{\pi} \left[\frac{t^2 + (1-s)^2}{2} \operatorname{atan}\left(\frac{t}{1-s}\right) - \frac{t(1-s)}{2} \right] + \frac{2}{\pi} \left[\frac{t^2 + (1+s)^2}{2} \operatorname{atan}\left(\frac{t}{1+s}\right) - \frac{t(1+s)}{2} \right] \quad (57)$$

$$= \frac{1}{\pi} [((1+s)^2 + t^2) \operatorname{atan}\left(\frac{t}{1+s}\right) + ((1-s)^2 + t^2) \operatorname{atan}\left(\frac{t}{1-s}\right) - 2t]. \quad (58)$$

We derived (56) using the fact that $\int_0^\infty r^3 e^{-r^2/2} dr = 2$. Computing each of the integrals with respect to θ would result in (57), and the final result in (58) is derived by simplifying the terms in (57).

□

5.4 Proof of Theorem 2

We start from the equivalent optimization derived as the result of Lemma 4, defined as,

$$\min_{s \in \mathbb{R}} \min_{\mathbf{w} \in \mathbb{R}^n, \mathbf{w} \perp \mathbf{x}_0} \lambda f(s\mathbf{x}_0 + \mathbf{w}) - s\rho_{\text{init}} - \mathbf{x}_{\text{init}}^T \mathbf{w} \quad (59)$$

$$\text{subject to:} \quad \mathbf{h}^T \mathbf{w} \geq \sqrt{m \mathbf{c}_d(s, \|\mathbf{w}\|)} . \quad (60)$$

One key idea to analyze the optimization is to replace $f(s\mathbf{x}_0 + \mathbf{w})$ with its first-order linear approximation around the point \mathbf{x}_0 . Let \hat{f} denotes the approximation function,

$$\hat{f}(\mathbf{x}) = f(\mathbf{x}_0) + \max_{\mathbf{v} \in \lambda \partial f(\mathbf{x}_0)} \mathbf{v}^T (\mathbf{x} - \mathbf{x}_0) . \quad (61)$$

Here, $\partial f(\mathbf{x}_0)$ denotes the sub-differential of $f(\cdot)$ at point \mathbf{x}_0 which is well-defined for convex functions and is a compact and convex set. Replacing $f(\cdot)$ with its approximation, enables us to precisely analyze the conditions for perfect signal recovery in the equivalent optimization (60) which determines the precise phase transition in the regularized PhaseMax optimization. This approximation is tight when the norm of the error approaches zero (which occurs in perfect recovery). We refer the interested reader to [31] for more details.

Therefore, for the rest of this section, we will analyze the following optimization,

$$\min_{\substack{s \in \mathbb{R} \\ \mathbf{w} \in \mathbb{R}^n, \mathbf{w} \perp \mathbf{x}_0}} \max_{\mathbf{v} \in \lambda \partial f(\mathbf{x}_0)} f(\mathbf{x}_0) + \lambda \mathbf{v}^T (s\mathbf{x}_0 + \mathbf{w} - \mathbf{x}_0) - s\rho_{\text{init}} - \mathbf{x}_{\text{init}}^T \mathbf{w} \\ \text{subject to:} \quad \mathbf{h}^T \mathbf{w} \geq \sqrt{m \mathbf{c}_d(s, \|\mathbf{w}\|)} , \quad (62)$$

Next, we use the dual variable β to rewrite (62) as

$$\min_{\substack{s \in \mathbb{R} \\ \mathbf{w} \in \mathbb{R}^n, \mathbf{w} \perp \mathbf{x}_0}} \max_{\substack{\mathbf{v} \in \lambda \partial f(\mathbf{x}_0) \\ \beta \geq 0}} f(\mathbf{x}_0) + \lambda \mathbf{v}^T (s\mathbf{x}_0 + \mathbf{w} - \mathbf{x}_0) - s\rho_{\text{init}} - \mathbf{x}_{\text{init}}^T \mathbf{w} - \frac{\beta}{\sqrt{n}} \mathbf{h}^T \mathbf{w} + \frac{\beta}{\sqrt{n}} \sqrt{m \mathbf{c}_d(s, \|\mathbf{w}\|)} . \quad (63)$$

In the next step, we would like to switch the minimization over \mathbf{w} with the maximization over β . But since the objective function is not convex with respect to \mathbf{w} , such an exchange wouldn't be a direct result of the Sion's min-max theorem. However, note that the initial optimization satisfies the conditions of the Sion's min-max theorem. In the asymptotic settings, using the same techniques as in [39] (see section A.2.4 in the appendix of the paper), one can show that changing the order of min and max does not change the solution of the optimization problem.

Hence, we are now able to first do the minimization over \mathbf{w} . To do this, we define $r := \|\mathbf{w}\|$ and by fixing r we are computing the minimization with respect to the direction of \mathbf{w} . The following optimization is the result of minimization over the direction of \mathbf{w} :

$$\min_{\substack{s \in \mathbb{R} \\ r \geq 0}} \max_{\substack{\mathbf{v} \in \lambda \partial f(\mathbf{x}_0) \\ \beta \geq 0}} (1-s)(\rho_{\text{init}} - \mathbf{v}^T \mathbf{x}_0) - r \cdot \|\mathbf{P}^\perp(\lambda \mathbf{v} - \mathbf{x}_{\text{init}} - \frac{\beta}{\sqrt{n}} \mathbf{h})\| + \beta \sqrt{\alpha \mathbf{c}_d(s, r)} , \quad (64)$$

where as defined in Section 3, $\alpha = \frac{m}{n}$ is the oversampling ratio, and $\mathbf{P}^\perp = \mathbf{I} - \mathbf{x}_0 \mathbf{x}_0^T$ is the projection to the orthogonal subspace of \mathbf{x}_0 .

Up to this point, the result is valid for every convex function $f(\cdot)$. But in order to continue our analysis, we need the following lemma which restricts us to a specific class of functions, i.e., the class of absolutely scalable functions.

Lemma 8. *Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a convex function such that for all $\mathbf{x} \in \mathbb{R}^n$ and $\alpha \geq 0$, $f(\alpha \mathbf{x}) = \alpha f(\mathbf{x})$. Then, for all $\mathbf{v} \in \partial f(\mathbf{x})$,*

$$\mathbf{v}^T \mathbf{x} = f(\mathbf{x}) ,$$

where $\partial f(\mathbf{x})$ is the set of sub-differentials of function f at point \mathbf{x} .

Proof. Since f is convex, for all $\mathbf{v} \in \partial f(\mathbf{x})$ and any $\epsilon < 1$, we have

$$(1-\epsilon) f(\mathbf{x}) = f((1-\epsilon)\mathbf{x}) \geq f(\mathbf{x}) - \epsilon \mathbf{v}^T \mathbf{x} . \quad (65)$$

Thus, $\epsilon f(\mathbf{x}) \leq \epsilon \mathbf{v}^T \mathbf{x}$. Choosing $\epsilon_1 = 1/2$ and $\epsilon_2 = -1/2$ yields $\mathbf{v}^T \mathbf{x} = f(\mathbf{x})$ which concludes the proof. \square

If we apply Lemma 8 to the objective function in (64), we can replace $\mathbf{v}^\top \mathbf{x}_0$ with $f(\mathbf{x}_0)$ for all $\mathbf{v} \in \partial f(\mathbf{x}_0)$, which gives the following optimization,

$$\min_{\substack{s \in \mathbb{R} \\ r \geq 0}} \max_{\beta \geq 0} -(1-s)L(\mathbf{x}_0) - r \cdot \min_{\mathbf{v} \in \lambda \mathbf{P}^\perp \partial L(\mathbf{x}_0)} \left\| \mathbf{v} - \frac{\beta}{\sqrt{n}} \mathbf{h} \right\| + \beta \sqrt{\alpha \mathbf{c}_d(s, r)}. \quad (66)$$

Recall that $(1-s)$ and $r \geq 0$ respectively represent the norm of the error in the direction of \mathbf{x}_0 and its orthogonal complement. Therefore, the perfect recovery in our optimization corresponds to the case where the optimizers are $r^* = 0$ and $s^* = 1$, and we are interested in the phase transition ratio α^* for which this happens.

We use the following approximation of the objective function near the point $(r, s) = (0, 1)$, which was introduced earlier in (31),

$$\mathbf{c}_d(s, r) = \frac{1}{\pi} [((1-s)^2 + r^2) \text{atan}\left(\frac{r}{1-s}\right) - r(1-s)]. \quad (67)$$

We define the new variable $t := \frac{r}{1-s}$ and rewrite the optimization in terms of t and s . One can show from (67) that as $r \downarrow 0$ and $s \uparrow 1$, the value of $\mathbf{c}_d(s, r)$ will only depends on the ratio t .

$$\min_{\substack{s \in \mathbb{R} \\ t \geq 0}} \max_{\beta \geq 0} \Psi(s, t, \beta) = -(1-s)L(\mathbf{x}_0) - t(1-s) \cdot \text{dist}_{\lambda \partial L^\perp(\mathbf{x}_0)}\left(\frac{\beta}{\sqrt{n}} \mathbf{h}\right) + \beta(1-s) \sqrt{\alpha((1+t^2) \text{atan}(t) - t)}, \quad (68)$$

where $\partial L^\perp(\mathbf{x}_0) = \mathbf{P}^\perp \partial L(\mathbf{x}_0)$ and $\text{dist}_S(\mathbf{x})$ is the distance function defined in Definition 1. Since we have a convex-concave objective function over three scalars, we can write the first order optimality conditions for the solutions to (68),

$$\begin{cases} \frac{\partial}{\partial \beta} \Psi(s, t, \beta) \Big|_{(s^*=1, t^*, \beta^*)} = 0 \\ \frac{\partial}{\partial s} \Psi(s, t, \beta) \Big|_{(s^*=1, t^*, \beta^*)} = 0 \\ \frac{\partial}{\partial t} \Psi(s, t, \beta) \Big|_{(s^*=1, t^*, \beta^*)} = 0 \end{cases}. \quad (69)$$

We would like to find the conditions (on α) under which the solution to (69) happens at $s^* = 1$. Therefore, we aim to solve the system of nonlinear equations (69), for three unknowns t, β and δ .

These equations can be written in the following form,

$$\begin{aligned} -t \cdot \frac{\frac{\beta}{n} \|\mathbf{h}\|^2 - \frac{h^\top}{\sqrt{n}} \Pi_{\partial L^\perp(\mathbf{x}_0)}\left(\frac{\beta}{\sqrt{n}} \mathbf{h}\right)}{\text{dist}_{\lambda \partial L^\perp(\mathbf{x}_0)}\left(\frac{\beta}{\sqrt{n}} \mathbf{h}\right)} + \sqrt{\alpha((1+t^2) \text{atan}(t) - t)} &= 0 \\ L(\mathbf{x}_0) + t \cdot \text{dist}_{\lambda \partial L^\perp(\mathbf{x}_0)}\left(\frac{\beta}{\sqrt{n}} \mathbf{h}\right) - \beta \sqrt{\alpha((1+t^2) \text{atan}(t) - t)} &= 0 \\ -\text{dist}_{\lambda \partial L^\perp(\mathbf{x}_0)}\left(\frac{\beta}{\sqrt{n}} \mathbf{h}\right) + \frac{\beta t \text{atan}(t) \sqrt{\alpha}}{\sqrt{\alpha((1+t^2) \text{atan}(t) - t)}} &= 0 \end{aligned} \quad (70)$$

Next, we are going to exploit the conditions of assumption 1 (see Section 3.2). Using theorem 5.2.2. in [44], both the functions $\text{dist}_{\lambda \partial L^\perp(\mathbf{x}_0)}\left(\frac{\beta}{\sqrt{n}} \mathbf{h}\right)$ and $\frac{h^\top}{\sqrt{n}} \Pi_{\partial L^\perp(\mathbf{x}_0)}\left(\frac{\beta}{\sqrt{n}} \mathbf{h}\right)$ converge point-wise to their expected value. Moreover, from assumption 1, we know that both $\mathbb{E}[\text{dist}_{\lambda \partial L^\perp(\mathbf{x}_0)}\left(\frac{\beta}{\sqrt{n}} \mathbf{h}\right)]$ and $\mathbb{E}[\frac{h^\top}{\sqrt{n}} \Pi_{\partial L^\perp(\mathbf{x}_0)}\left(\frac{\beta}{\sqrt{n}} \mathbf{h}\right)]$ converge uniformly to $G_\lambda(\beta)$ and $\beta - F_\lambda(\beta)$, respectively. Therefore, using the same arguments as in [39], we can replace $\text{dist}_{\lambda \partial L^\perp(\mathbf{x}_0)}\left(\frac{\beta}{\sqrt{n}} \mathbf{h}\right)$ with $F_\lambda(\beta)$ in the optimization (68), and then apply the result of theorem 7.17 in [34], we can show that $F'_\lambda(\beta) = G_\lambda(\beta) G'_\lambda(\beta)$.

Therefore, we are able to use the functions F_λ , and G_λ to rewrite the system of non-linear equations (70):

$$\begin{aligned} -t \cdot \frac{F_\lambda(\beta)}{G_\lambda(\beta)} + \sqrt{\alpha((1+t^2)\text{atan}(t) - t)} &= 0 \\ L(\mathbf{x}_0) + t \cdot G_\lambda(\beta) - \beta \sqrt{\alpha((1+t^2)\text{atan}(t) - t)} &= 0 \\ -G_\lambda(\beta) + \frac{\beta t \text{atan}(t) \sqrt{\alpha}}{\sqrt{\alpha((1+t^2)\text{atan}(t) - t)}} &= 0 \end{aligned} \quad (71)$$

By combining the first and third equations, we will get

$$t = \tan\left(\frac{\pi}{\alpha\beta} F_\lambda(\beta)\right) \quad (72)$$

Finally, using (72) in (71) reduces the number of equations to 2, and yields the following system of non-linear equations.

$$\begin{cases} G_\lambda(\beta) l = \tan\left(\frac{\pi}{\alpha\beta} F_\lambda(\beta)\right) (G_\lambda^2(\beta) - \beta F_\lambda(\beta)), \\ \tan\left(\frac{\pi}{\alpha\beta} F_\lambda(\beta)\right) (G_\lambda(\beta) - \frac{\pi l}{\alpha\beta} F_\lambda(\beta)) = \frac{\pi}{\alpha\beta} F_\lambda(\beta) G_\lambda(\beta), \end{cases} \quad (73)$$

This concludes the proof.

5.5 Computing the statistical dimension for the ℓ_1 regularization

In this section we bound the statistical dimension of the descent cone of the objective function of (2), where $f(\cdot) = \|\cdot\|_1$ is used for regularization. We assume that the underlying signal, \mathbf{x}_0 , is k -sparse and define function $L_\lambda : \mathbb{R}^n \rightarrow \mathbb{R}$ as follows,

$$L_\lambda(\mathbf{x}) := -\mathbf{x}_{\text{init}}^\top \mathbf{x} + \lambda \|\mathbf{x}\|_1, \quad (74)$$

In order to derive an upper bound for the statistical dimension $d(T_{L_\lambda}(\mathbf{x}_0))$, we first introduce another summary parameter for convex sets called the Gaussian width.

Definition 6. (Gaussian width) [44] The Gaussian width of a subset $\mathcal{T} \subset \mathbb{R}^n$ is defined as,

$$\omega(\mathcal{T}) = \mathbb{E} \sup_{\mathbf{x} \in \mathcal{T}} \langle \mathbf{x}, \mathbf{g} \rangle, \quad \text{where } \mathbf{g} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}). \quad (75)$$

We use the following proposition which indicates the relationship between the Gaussian width and statistical dimension of a convex cone.

Proposition 1. (Proposition 10.2 in [1]) Let $\mathcal{C} \subset \mathbb{R}^n$ be a convex cone. Then

$$\omega^2(\mathcal{C} \cap \mathbb{S}^{n-1}) \leq d(\mathcal{C}) \leq \omega^2(\mathcal{C} \cap \mathbb{S}^{n-1}) + 1, \quad (76)$$

where $\mathbb{S}^{n-1} \subset \mathbb{R}^n$ is the unit sphere.

The Proposition 1 shows that in order to bound the statistical dimension of a convex cone, we need to bound the squared Gaussian width of that cone. Hence, in the remaining we will bound the squared Gaussian width. We shall briefly review some known properties of Gaussian width of convex cones.

5.5.1 Some properties of Gaussian width

The Gaussian width is one of the intrinsic volumes of a body studied in combinatorial geometry. It is invariant under translation and unitary transformation and has deep connections to convex geometry. While discussing all the properties of Gaussian width is beyond the scope of this paper, we refer the interested reader to [33, 42, 44] and references therein.

Inspired by [38, 11], here we bound the Gaussian width of a cone via the distance to its polar cone. Before stating the proposition, we recall the definition of the polar cone.

Definition 7. (Polar cone) Let $\mathcal{C} \subset \mathbb{R}^n$ be a non-empty convex cone. The polar cone of \mathcal{C} , denoted by \mathcal{C}^* , is defined as follows,

$$\mathcal{C}^* = \{\mathbf{z} \in \mathbb{R}^n : \langle \mathbf{z}, \mathbf{x} \rangle \leq 0 \text{ for all } \mathbf{x} \in \mathcal{C}\}. \quad (77)$$

The following proposition establishes a connection between the Gaussian width of the cone \mathcal{C} and its polar cone \mathcal{C}^* :

Proposition 2. (Proposition 3.6 in [11]). Let \mathcal{C} be any non-empty convex cone in \mathbb{R}^n , and let $\mathbf{g} \sim \mathcal{N}(0, \mathbf{I})$ be a random Gaussian vector. Then we have the following bound:

$$\omega(\mathcal{C} \cap \mathbb{S}^{n-1}) \leq \mathbb{E}_{\mathbf{g}}[\text{dist}(\mathbf{g}, \mathcal{C}^*)], \quad (78)$$

where the $\text{dist}(\cdot, \cdot)$ function here denotes the Euclidean distance between a point and a set.

Applying Jensen's inequality will result in the following,

$$\omega^2(\mathcal{C} \cap \mathbb{S}^{n-1}) \leq \mathbb{E}_{\mathbf{g}}[\text{dist}^2(\mathbf{g}, \mathcal{C}^*)]. \quad (79)$$

This is very useful in bounding the Gaussian width of the descent cone of a convex function due to the following lemma:

Lemma 9. [32] For a convex function $f : \mathbb{R}^n \rightarrow \mathbb{R}$,

$$(T_f(\mathbf{x}))^* = \text{cone}(\partial f(\mathbf{x})), \quad (80)$$

where $\partial f(\mathbf{x})$ is the sub-differential set of function f at point \mathbf{x} .

The polar cone of $T_f(\mathbf{x})$ is also called the *normal cone*, $N_f(\mathbf{x})$, at point \mathbf{x} . Exploiting the above results, we can bound $\omega^2(T_L(\mathbf{x}_0) \cap \mathbb{S}^{n-1})$ in terms of the squared distance to the normal cone at point \mathbf{x}_0 , i.e., we have the followings,

$$d(T_{L_\lambda}(\mathbf{x}_0)) \leq \omega^2(T_{L_\lambda}(\mathbf{x}_0) \cap \mathbb{S}^{n-1}) + 1 \leq \mathbb{E}_g[\text{dist}^2(\mathbf{g}, \text{cone}(\partial L_\lambda(\mathbf{x}_0)))] + 1. \quad (81)$$

For simplicity in the remaining formulations we omit the sub-script λ and denote the objective function by L . Let Δ denote the set of coordinates where \mathbf{x}_0 is non-zero. The sub-differential set of the function L (defined in (74)) can be formally characterized as follows,

$$\partial L(\mathbf{x}_0) = \{-\mathbf{x}_{\text{init}} + \frac{\lambda}{\sqrt{k}} \mathbf{v} : \mathbf{v} \in \mathbb{R}^n \text{ s.t. } \mathbf{v}[i] = \text{sign}(\mathbf{x}_0[i]) \text{ for } i \in \Delta, |\mathbf{v}[i]| \leq 1 \text{ for } i \in \Delta^c\}. \quad (82)$$

Here $\Delta^c := [n] - \Delta$ represents the zero entries of \mathbf{x}_0 . Without the loss of generality, we are going to assume that the first k entries of \mathbf{x}_0 are non-zero, while the rest are zero. Then, cone of the sub-differential can be written as

$$\text{cone}(\partial f(\mathbf{x})) = \{\beta \cdot (-\mathbf{x}_{\text{init}} + \frac{\lambda}{\sqrt{k}} \mathbf{v}) : \mathbf{v}[1:k] = \mathbf{1}, \|\mathbf{v}[k+1:n]\|_\infty \leq 1, \beta \geq 0\}. \quad (83)$$

The squared distance to the normal cone can be formulated as the following optimization:

$$\begin{aligned} \text{dist}^2(\mathbf{g}, N_L(\mathbf{x}_0)) &= \min_{t \geq 0} \left(\sum_{i \in \Delta} (\mathbf{g}[i] + t\mathbf{x}_{\text{init}}[i] - t\lambda \text{sign}(\mathbf{x}_0[i]))^2 \right. \\ &\quad \left. + \min_{j \in \Delta^c, |u_j| < t} (\mathbf{g}[j] + t\mathbf{x}_{\text{init}}[j] - \lambda u_j)^2 \right) \end{aligned} \quad (84)$$

Define $\mathbf{z} := \mathbf{z}(t) = \mathbf{g} + t\mathbf{x}_{\text{init}}$. We can rewrite the equation (84) as follows,

$$\begin{aligned} \text{dist}^2(\mathbf{g}, N_L(\mathbf{x}_0)) &= \min_{t \geq 0} \left(\sum_{i \in \Delta} (\mathbf{z}[i] - t\lambda \text{sign}(\mathbf{x}_0[i]))^2 \right. \\ &\quad \left. + \sum_{j \in \Delta^c} \text{shrink}(\mathbf{z}[j], t\lambda)^2 \right), \end{aligned} \quad (85)$$

where the function $\text{shrink}(\cdot, \cdot)$ is defined as, $\text{shrink}(x, T) = \begin{cases} x + T, & x < -T \\ x, & -T \leq x \leq T \\ x - T, & x > T \end{cases}$. This function

is known as ℓ_1 -shrinkage function and is used in sparse denoising. Taking the expectation with respect to \mathbf{g} will provide us with the quantity we would like to bound. We are bounding the expectation of the squared distance by bounding each of the two terms of the sum. For the first term, we have:

$$\mathbb{E} \left[\sum_{i \in \Delta} (\mathbf{z}[i] - t\lambda \text{sign}(\mathbf{x}_0[i]))^2 \right] = k + t^2(\lambda^2 k - \|\mathbf{x}_{\text{init}}^\Delta\|^2) - 2t\lambda \text{sign}(\mathbf{x}_0)^T \mathbf{x}_{\text{init}}^\Delta \quad (86)$$

Bounding the expectation of the second term in (85) is more involved. Using the same techniques as Appendix C in [11], we can show the following:

$$\mathbb{E} \left[\sum_{j \in \Delta^c} \text{shrink}^2(\mathbf{z}[j], t\lambda) \right] \leq \frac{2(n-k)\sigma^3}{\sqrt{2\pi}t\lambda} \exp\left(\frac{-t^2\lambda^2}{2\sigma^2}\right), \quad (87)$$

where $\sigma^2 := 1 + t^2 \frac{\|\mathbf{x}_{\text{init}}^\Delta\|^2}{n-k}$.

Using the result of equations (86) and (87), we can see that when $\lambda > \frac{c}{\sqrt{k}}$, then by choosing $t = \frac{\sqrt{2 \log \frac{n}{k}}}{\lambda}$, both of the terms in the sum are bounded by $Ck \log \frac{n}{k}$, where c and C are constants that are independent of the problem's parameters.

Therefore, when $\lambda > \frac{c}{\sqrt{k}}$ (where k is the number of non-zero entries), the statistical dimension of $T_{L_\lambda}(\mathbf{x}_0)$ is bounded by $Ck \log \frac{n}{k}$. Using the result of Theorem 1, we can conclude that for the sparse phase retrieval problem the required sample complexity of the regularized PhaseMax is $\mathcal{O}(k \log \frac{n}{k})$. This indicates that regularized PhaseMax is order-wise optimal.