DETECTION WITH PHASELESS MEASUREMENTS

Amitai Koretz[⋆] *Ami Wiesel*[⋆] *Yonina C. Eldar*[†]

[⋆] Rachel and Selim Benin School of Computer Science and Engineering,
The Hebrew University of Jerusalem, Israel
[†]Department of Electrical Engineering, Technion, Israel Institute of Technology

ABSTRACT

We consider the problem of hypothesis testing for detection of a signal in Gaussian noise. We assume that the vector of measurements is unobserved, and that our observations consist of phaseless inner products with a set of known measurement vectors. This is typical of the phase retrieval problem, where the goal is to recover the vector of measurements. We provide a simple estimator for the test statistic that does not necessitate a phaseless recovery method to reconstruct the measurements. Our analysis shows that for random measurement vectors, we can reconstruct the test statistic for any signal from a sufficient number of observations, quadratic in the signal length, using a simple least-squares approach. The primary advantage of this method its simplicity and computational efficiency, which comes at the expense of requiring many more measurements. We show that for Fourier measurements vectors, our approach works only when the signal is also a Fourier vector.

Index Terms— detection, phase retrieval, least-squares approximation

1. INTRODUCTION

The problem of recovering a complex signal from phaseless measurements is of significant interest in fields including crystallography [1], optics [2], astronomy [3] and possible smart grid tasks [4] and has received increased attention in recent years. Typical scenarios include recovering a signal from Fourier magnitude measurements, or recovery from the absolute values of random i.i.d. linear measurements [5]. A unique signal consistent with the measurements does not always exist. In some situations, increasing the number of measurements (oversampling) will lead to the existence of a unique solution, up to inherent ambiguities. Other methods exploit sparsity of the unknown signal to facilitate recovery and reduce ambiguities [6–8].

In this paper, we focus on estimating a positive functional of the unknown signal, rather than the signal itself. Our motivation comes from a situation where we would like to perform hypothesis testing, which utilizes a functional of the unknown signal (a test statistic). A straightforward approach is to apply a phaseless recovery method to estimate the signal, and then use it to calculate the test statistic. Instead, we propose a simple method to estimate the test statistic using Least Squares estimation, which has very low computational complexity compared with typical phaseless recovery algorithms. We show that this method can be successfully used with N^2 random measurement vectors, (where N is the length of the unknown signal). This confirmed by numerical simulations of detector performance, which show that the probability of detection at a given false alarm rate improves with additional measurements. However, this approach is not suitable for Fourier measurements, due to the structure of the DFT basis vectors. Specifically, the test statistic can be successfully recovered only if the hypothesis signal (known up to a constant) has the same structure as the measurement vectors. In this case, $2N-1$ observations are sufficient for perfect recovery.

We mention here a partial list of existing methods for phaseless signal recovery. Early methods for phase retrieval featured non-convex projections [9,10] and often included assumptions on the signal such as knowledge of the bandwidth or support. More recently, we mention methods on semidefinite programming using convex relaxations [11–13]. For sparse signals, there are phaseless recovery algorithms based on convex relaxations [6] as well as nonlinear formulations [7]. Other methods are based on Wirtinger Flow descent [14, 15], alternating minimization [16] and generalized approximate message passing (GAMP) [17]. A method with low complexity which requires a quadratic number of measurements was described in [18], though a specific structure is assumed for the measurement vectors. Theoretical results on the existence of a unique solution include [19, 20] for Gaussian measurements, [21, 22] for Fourier measurements and [23] for the sparse signals. For a recent survey of the phase retrieval problem, see [5].

The outline of remainder of the paper is as follows. In Section 2 we outline the detection problem and present our detector using our proposed least-squares approximation. In Section 3 a theoretical analysis of the proposed approximation is presented for the case of random measurement vectors and for the case of Fourier measurement vectors. In Section 4 we display simulation results which complement our analysis for the case of random measurements. We then conclude the paper.

2. PHASELESS DETECTOR

2.1. Model

Consider an unknown vector $\mathbf{x} \in \mathbb{C}^N$. We are given a set of M squared magnitude observations $b_1, \ldots, b_M$ where

$$b_i = |\langle \mathbf{a}_i, \mathbf{x} \rangle|^2 \tag{1}$$

with respect to a given set of complex measurement vectors $\mathbf{a}_1, \ldots, \mathbf{a}_M$. Given a complex vector $\mathbf{s} \in \mathbb{C}^N$, our goal is to approximate a test statistic objective

$$h(\mathbf{x}) = |\langle \mathbf{s}, \mathbf{x} \rangle|^2 . \tag{2}$$

The work of A. Koretz and A. Wiesel was partially supported by the Israel Smart Grid (ISG) consortium. The work of Y. Eldar was funded by the European Unions Horizon 2020 research and innovation programme under grant agreement ERC-BNYQ, by the Israel Science Foundation under Grant no. 335/14, and by ICore: the Israeli Excellence Center Circle of Light.

Our motivation comes from a scenario of deciding between a pair of hypotheses H_0, H_1 given a vector $\mathbf{x} \in \mathbb{C}^N$:

$$H_1 : \mathbf{x} = A\mathbf{s} + \mathbf{w}$$
$$H_0 : \mathbf{x} = \mathbf{w}$$

where $\mathbf{s} \in \mathbb{C}^N$ is a deterministic known vector, $\mathbf{w} \sim \mathcal{CN}\left(0, \sigma^2 \mathbf{I}_N\right)$ where σ^2 is the noise variance and $A \in \mathbb{R}$ is an unknown deterministic scalar. We assume our observations are $b_1, \ldots, b_M$ as defined in (1) for an associated set of measurement vectors $\mathbf{a}_1, \ldots, \mathbf{a}_M$. A standard approach is to perform a generalized likelihood ratio test (GLRT), which requires the computation of the joint density of $b_1, \ldots, b_M$. A closed-form solution is intractable in general due to the statistical dependence between the observations. An exception to this is when $\mathbf{a}_1, \ldots, \mathbf{a}_M$ are an orthogonal set, in which case we can write the joint density of the observations as a product of chi-squared densities (this limits us to N observations).

We focus on the general case and consider a generalized likelihood ratio test (GLRT) with respect to the observation $\mathbf{x}$ to decide between H_0 and H_1:

$$h(\mathbf{x}) = |\langle \mathbf{s}, \mathbf{x} \rangle|^2 \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{3}$$

for a threshold $\eta > 0$. We refer to (3) as the squared correlator detector. Our proposal is to replace $h(\mathbf{x})$ in (3) by an estimate, $\widehat{h(\mathbf{x})}$, from the measurements.

2.2. Least-Squares Approximation

Denote

$$\mathcal{S}_M \triangleq \operatorname{span}_{\mathbb{R}} \left\{ \mathbf{a}_1 \mathbf{a}_1^H, \ldots, \mathbf{a}_M \mathbf{a}_M^H \right\} = \left\{ \sum_{i=1}^{M} \beta_i \mathbf{a}_i \mathbf{a}_i^H : \beta_i \in \mathbb{R} \right\}. \tag{4}$$

We write the objective as $h(\mathbf{x}) = \mathbf{x}^H \mathbf{s}\mathbf{s}^H \mathbf{x}$ and define an approximate test statistic $\widehat{h(\mathbf{x})} \triangleq \mathbf{x}^H \mathbf{S} \mathbf{x}$ where $\mathbf{S} \in \mathcal{S}_M$. We would like to choose $\mathbf{S}$ which minimizes the absolute distance from the true test statistic

$$\left| h(\mathbf{x}) - \widehat{h(\mathbf{x})} \right| = \left| \mathbf{x}^H \left(\mathbf{s}\mathbf{s}^H - \mathbf{S} \right) \mathbf{x} \right|. \tag{5}$$

Since $\mathbf{x}$ is unknown, direct minimization of (5) is not possible. Instead, we minimize a surrogate term. From the Cauchy-Schwarz inequality

$$\left| \mathbf{x}^H \left(\mathbf{s}\mathbf{s}^H - \mathbf{S} \right) \mathbf{x} \right| \leq \left\| \mathbf{s}\mathbf{s}^H - \mathbf{S} \right\|_F \|\mathbf{x}\|_2^2 . \tag{6}$$

Therefore choose $\mathbf{S}$ to minimize the right hand side of (6), i.e.

$$\hat{\mathbf{S}} = \arg \min_{\mathbf{S} \in \mathcal{S}_M} \left\| \mathbf{s}\mathbf{s}^H - \mathbf{S} \right\|_F^2 = \sum_{i=1}^{M} \hat{\beta}_i \mathbf{a}_i \mathbf{a}_i^H \tag{7}$$

for appropriate coefficients $\hat{\beta}_1, \ldots, \hat{\beta}_M$. The approximate test statistic becomes

$$\begin{aligned} \widehat{h(\mathbf{x})} &= \mathbf{x}^H \hat{\mathbf{S}} \mathbf{x} = \mathbf{x}^H \left(\sum_{i=1}^{M} \hat{\beta}_i \mathbf{a}_i \mathbf{a}_i^H \right) \mathbf{x} \\ &= \sum_{i=1}^{M} \hat{\beta}_i \left| \mathbf{a}_i^H \mathbf{x} \right|^2 = \sum_{i=1}^{M} \hat{\beta}_i b_i. \end{aligned} \tag{8}$$

Expanding the norm in (7) and some algebra gives the following optimal coefficients:

$$\hat{\boldsymbol{\beta}} = \left(\left| \mathbf{A}^H \mathbf{A} \right|^2 \right)^{-1} \left| \mathbf{A}^H \mathbf{s} \right|^2, \tag{9}$$

where $\hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_1 & \ldots, & \hat{\beta}_M \end{bmatrix}^T$ and $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \ldots, & \mathbf{a}_M \end{bmatrix}$. The matrix operator $|\cdot|^2$ indicates entrywise squared absolute value in this context. The derivation is straightforward and is omitted due to space constraints.

The proposed least-squares approach is a fast way to approximate the objective. Additionally, if the measurement vectors are known ahead of time, the coefficients $\hat{\beta}_1, \ldots, \hat{\beta}_M$ may be computed offline and $\widehat{h(\mathbf{x})}$ is calculated at a cost of M multiplications. We will see that N^2 random measurements will recover $h(\mathbf{x})$ with probability one. This is in contrast to some methods such as PhaseLift, which require $\mathcal{O}(N)$ random measurements for accurate recovery of $\mathbf{x}$ (as shown in [12, 13]), albeit at a much higher computational cost. Therefore, our method is appropriate when phaseless measurements are cheap and plentiful and computational time is at a premium.

3. ANALYSIS

In this section, we analyze our approximation $\widehat{h(\mathbf{x})}$ for random measurements and for Fourier measurements. The key attribute for the viability of the method is the dimension of the span of outer products of the measurement vectors. We see that for random measurements, the dimension increases to the dimension of the entire ambient space. However, the particular structure of Fourier vectors prevents the dimension from increasing beyond a trivial level.

3.1. Random Measurements

3.1.1. Rank Sufficiency

We consider the linear span of the rank-one Hermitian matrices $\mathbf{a}_1 \mathbf{a}_1^H, \ldots, \mathbf{a}_M \mathbf{a}_M^H$ as a subspace of the linear space of Hermitian matrices of size $N \times N$. Note that the Hermitian matrices form a vector space over the field of real numbers (but not over the complex numbers) with dimension N^2.

Proposition 1. *With probability one,* $\dim \operatorname{span}\left\{ \mathbf{a}_1 \mathbf{a}_1^H, \ldots, \mathbf{a}_M \mathbf{a}_M^H \right\} = \min\left\{ M, N^2 \right\}$ *when the joint probability distribution of* $\mathbf{a}_1, \ldots, \mathbf{a}_M$ *has a continuous density.*

Proposition 1 states that each additional measurement vector increases the dimension of the subspace spanned by the outer products. As a result, N^2 measurement vectors span the entire space of $N \times N$ complex matrices.

Proof. Assume that $M = N^2$. Consider the matrix $\mathbf{Y} \in \mathbb{C}^{N^2 \times N^2}$:

$$\mathbf{Y}\left(\mathbf{a}_1, \ldots, \mathbf{a}_{N^2}\right) \triangleq \begin{bmatrix} \operatorname{vec}\left(\mathbf{a}_1 \mathbf{a}_1^H\right) & \ldots, & \operatorname{vec}\left(\mathbf{a}_{N^2} \mathbf{a}_{N^2}^H\right) \end{bmatrix}.$$

The determinant $\det \mathbf{Y} : \mathbb{R}^{2N^3} \longrightarrow \mathbb{C}$ is a polynomial in the real and imaginary parts of the elements of $\mathbf{a}_1, \ldots, \mathbf{a}_{N^2}$. Since a polynomial is nonzero almost everywhere (with respect to the Lebesgue measure on the domain), the set

$$\mathcal{A} \triangleq \left\{ \mathbf{a}_1, \ldots, \mathbf{a}_{N^2} : \det \mathbf{Y}\left(\mathbf{a}_1, \ldots, \mathbf{a}_{N^2}\right) = 0 \right\} \subseteq \mathbb{R}^{2N^3}$$

has zero Lebesgue measure. As a result, $\mathbb{P}\left(\mathbf{a}_1, \ldots, \mathbf{a}_{N^2} \in \mathcal{A}\right) = 0$. Since the determinant of $\mathbf{Y}$ is nonzero with probability one, the N^2

outer products span the entire ambient dimension with probability one. This implies that for $M < N^2$, $\mathbf{Y}(\mathbf{a}_1, \ldots, \mathbf{a}_M)$ is full rank since, if we assume the converse, then we are in contradiction with the result stated above. □

It is evident Proposition 1 holds for any joint distribution of $\mathbf{a}_1, \ldots, \mathbf{a}_M$ which possesses a continuous density. Our proof is similar to the one given in [24, Theorem 2.1]. Proposition 1 implies that given N^2 measurements, any Hermitian matrix can be expressed as a linear combination of $\mathbf{a}_1\mathbf{a}_1^H, \ldots, \mathbf{a}_N\mathbf{a}_N^H$ and the test statistic will be fully recovered (with probability one). In comparison, if we use a state-of-the-art phaseless recovery method to recover $\mathbf{x}$ and then calculate the statistic, the number of noiseless measurements required is much less. For example, PhaseLift requires $\mathcal{O}(N)$ measurements [13] for recovery with high probability, but involves solving a semidefinite program of size $N \times N$. The Truncated Wirtinger Flow method requires $\mathcal{O}(N \log N)$ measurements and yields a solution with ϵ relative accuracy in $MN^2 \log 1/\epsilon$ flops [15]. In contrast, our proposed method requires $\mathcal{O}(M^2N)$ flops for the least-squares solution (which can be performed offline if the measurement vectors are known ahead of time), and M multiplications plus summations to calculate the statistic for each set of measurements.

3.2. Fourier Measurements

3.2.1. Toeplitz Approximation

In the case where the measurement vectors are rows of a Discrete Fourier Transform (DFT) matrix (or a matrix representing an oversampled DFT), the collection of outer products will not approximate a given Hermitian matrix in general. Assume that $\{\omega_1, \ldots, \omega_M\} \subset [0,\, 2\pi)$ form a distinct set. The measurement vectors have the form (up to a normalization constant)

$$\mathbf{a}_i = \left[1, e^{-i2\pi\omega_i}, \ldots, e^{-i2\pi(N-1)\omega_i}\right]^T, \quad 1 \le i \le M. \tag{10}$$

Lemma 1. *For $M \ge 2N-1$, the span of $\mathbf{a}_1\mathbf{a}_1^H, \ldots, \mathbf{a}_M\mathbf{a}_M^H$ over the field of reals is equal to*
$\mathcal{L} = \left\{\mathbf{S} \in \mathbb{C}^{N\times N} : \mathbf{S} = \mathbf{S}^H \text{ and } \mathbf{S} \text{ is Toeplitz}\right\}$.

Proof. Let $\mathbf{a}_1, \ldots, \mathbf{a}_M$ represent an M-point uniform sampling of the Discrete Time Fourier Transform (DTFT) spectrum, i.e. $[\mathbf{a}_m]_n = \frac{1}{\sqrt{N}} e^{-i2\pi\frac{(m-1)(n-1)}{M}}$. It is evident that $\mathbf{a}_m\mathbf{a}_m^H$ has both Hermitian and Toeplitz structure, both of which form linear spaces over the field of reals. This shows the left inclusion. Arranging a subset of N vectors of $\mathbf{a}_1, \ldots, \mathbf{a}_N$ in a matrix, we have a Vandermonde structure which can be seen to have a nonzero determinant. As a result, every subset of N vectors is linearly independent. We next use the following proposition:

Proposition 2. [24, Proposition 4.1] *Suppose every subset of $\mathbf{a}_1, \ldots, \mathbf{a}_M$ containing no more than N vectors is linearly independent, and that $M \le 2N-1$. Then $\mathbf{a}_1\mathbf{a}_1^H, \ldots, \mathbf{a}_M\mathbf{a}_M^H$ is linearly independent with probability one.*

Assume that $M \ge 2N-1$. The span of $\mathbf{a}_1\mathbf{a}_1^H, \ldots, \mathbf{a}_M\mathbf{a}_M^H$ over the complex field has dimension $2N-1$ and is equal to the set of Toeplitz matrices (which have $2N-1$ degrees of freedom). Therefore, the Hermitian Toeplitz matrices are a subset of $\operatorname{span}\left\{\mathbf{a}_1\mathbf{a}_1^H, \ldots, \mathbf{a}_M\mathbf{a}_M^H\right\}$, and are obtained only by all real coefficients. This shows the reverse inclusion. □

As a result of Lemma 1, the detector performance will generally not reach the squared correlator performance. This is because the approximation error, which in general is nonzero, is not necessarily orthogonal to the unobserved vector $\mathbf{x}$, leading to an error in the estimation of the test statistic.

3.2.2. Amplitude Spectrum

We observe that in the case of Fourier measurements, $h(\mathbf{x})$ can be recovered only when the signal $\mathbf{s}$ has the same structure as the measurement vectors. Lemma 1 implies that a necessary and sufficient condition for recovering $h(\mathbf{x})$ is that $\mathbf{s}\mathbf{s}^H$ is Toeplitz. It is not hard to show that this holds only when $\mathbf{s}$, up to a normalization constant, is of the form defined by Eq. (10).

As immediate result, we can derive a general property of the Fourier amplitude spectrum of discrete signals, which agrees with a result stated in [21]. The squared amplitude spectrum of a discrete signal $\mathbf{x}$, denoted as $\Psi_{\mathbf{x}}(\omega)$, is defined as the pointwise squared absolute value of the DTFT of $\mathbf{x} = \left[\; x[0] \quad, \ldots, \quad x[N-1] \;\right]^T$, i.e.

$$\Psi_{\mathbf{x}}(\omega) = \left|\sum_{n=0}^{N-1} x[n] e^{-in\omega}\right|^2 = h(\mathbf{x}) \tag{11}$$

where i denotes the complex unit, $h(\mathbf{x})$ is defined in Eq. (2) for some $\mathbf{s}$ of the form in Eq. (10) for some appropriate $\omega \in [0,\, 2\pi)$. In general, $2N-1$ observations samples define the amplitude spectrum of $\mathbf{x}$, as the following proposition shows.

Proposition 3. *Given a discrete signal $\mathbf{x} \in \mathbb{C}^N$ and a fixed set of angular frequencies $\Omega = \{\omega_1, \ldots, \omega_M\}$ containing at least $M \ge 2N-1$ distinct values, the amplitude spectrum of $\mathbf{x}$ evaluated at any $\omega' \in [0,\, 2\pi)$ is given as*

$$\Psi_{\mathbf{x}}(\omega') = \sum_{i:\omega_i \in \Omega} \beta_i \Psi_{\mathbf{x}}(\omega_i) \tag{12}$$

for appropriate coefficients $\beta_1, \ldots, \beta_M$. Similarly, given a 2-dimensional discrete signal $\mathbf{X} \in \mathbb{C}^{N_a \times N_b}$ and two sets of angular frequencies Ω_a, Ω_b containing at least $M_a \ge 2N_a - 1, M_b \ge 2N_b - 1$ distinct values, respectively, the amplitude spectrum of $\mathbf{X}$ evaluated at any (ω_a', ω_b') is given as

$$\Psi_{\mathbf{X}}(\omega_a', \omega_b') = \sum_{i:\omega_i \in \Omega_a} \sum_{j:\omega_j \in \Omega_b} \beta_i \gamma_j \Psi_{\mathbf{X}}(\omega_i, \omega_j) \tag{13}$$

for appropriate coefficients $\beta_1, \ldots, \beta_{M_a}$, $\gamma_1, \ldots, \gamma_{M_b}$.

Proposition 3 implies that the amplitude spectrum of a discrete signal is completely defined by its samples, where the number of samples in each dimension is at least double the dimension minus one. This is known to be the case for any number of dimensions [21].

Proof. We begin with the one-dimensional case. Choose some $\omega' \in [0\, 2\pi)$. Define

$$\mathbf{a}_i = \left[1, e^{-i2\pi\omega_i}, \ldots, e^{-i2\pi(N-1)\omega_i}\right]^T, i = 0, \ldots, M \tag{14}$$

$$\mathbf{a}' = \left[1, e^{-i2\pi\omega'}, \ldots, e^{-i2\pi(N-1)\omega'}\right]^T. \tag{15}$$

It holds that

$$\Psi_{\mathbf{x}}(\omega_i) = \mathbf{x}^H \mathbf{a}_i \mathbf{a}_i^H \mathbf{x}, \quad 0 \le i \le M. \tag{16}$$

$$\Psi_{\mathbf{x}}(\omega') = \mathbf{x}^H \mathbf{a}' \mathbf{a}'^H \mathbf{x}. \tag{17}$$

Clearly $\mathbf{a}_i\mathbf{a}_i^H$ have Hermitian and Toeplitz structure for $0 \leq i \leq M$. Lemma 1 implies that $\mathbf{a}'\mathbf{a}'^H \in \text{span}\left(\mathbf{a}_1\mathbf{a}_1^H, \ldots, \mathbf{a}_M\mathbf{a}_M^H\right)$ so we can write $\mathbf{a}'\mathbf{a}'^H = \sum_{i:\omega_i \in \Omega}^{M} \beta_i \mathbf{a}_i\mathbf{a}_i^H$ for some coefficients $\beta_1, \ldots, \beta_M$. Substituting this expression into Eq. (17) and applying Eq. (16) gives Eq. (12).

For the two dimensional case, note that we can write

$$\Psi_{\mathbf{X}}\left(\omega_a', \omega_b'\right) = \left|\mathbf{a}_i^H \mathbf{X} \mathbf{a}_j\right|^2 = \text{Tr}\left(\mathbf{a}_i\mathbf{a}_i^H \mathbf{X} \mathbf{a}_j\mathbf{a}_j^H \mathbf{X}^H\right) \qquad (18)$$

where $\omega_a' \in \Omega_a, \omega_b' \in \Omega_b$ and, with a slight abuse of notation, $\mathbf{a}_i, \mathbf{a}_j$ are defined according to Eq. (14) using angular frequency paramters ω_i, ω_j, respectively. Here $\text{Tr}(\cdot)$ denotes the trace operator. The remainder of the proof proceeds in identical fashion as the one-dimensional case, yielding Eq. (13) after substituting two expressions $\sum_{i:\omega_i \in \Omega_a}^{M} \beta_i \mathbf{a}_i\mathbf{a}_i^H$, $\sum_{j:\omega_j \in \Omega_b}^{M} \gamma_j \mathbf{a}_j\mathbf{a}_j^H$. □

4. SIMULATIONS

We demonstrate the performance of the least-squares approximation for i.i.d. standard Gaussian measurement vectors $\mathbf{a}_1, \ldots, \mathbf{a}_M$. The model parameters were chosen as $N = 10, A = 2, \sigma^2 = 1$. The signal vector $\mathbf{s}$ is defined in the individual subsections. The number of observations M is a simulation variable.

4.1. Squared Approximation Error

We define the squared approximation error as

$$\hat{e}(M) \triangleq \left\|\mathbf{s}\mathbf{s}^H - \hat{\mathbf{S}}\right\|_F^2 = \min_{\mathbf{S} \in \mathcal{S}_M} \left\|\mathbf{s}\mathbf{s}^H - \mathbf{S}\right\|_F^2.$$

The empirical average squared approximation error, calculated by averaging 10000 independent realizations of $\hat{e}(M)$, is displayed in Fig. 1. We chose $\mathbf{s} = \begin{bmatrix} 1 & 0 & \ldots, & 0 \end{bmatrix}^T$. The results indicate that the expected squared approximation error, defined as $e(M) \triangleq \mathbb{E}\left[\hat{e}(M)\right]$, is dominated by $1 - \frac{M}{N^2}$ (represented by the dashed line). This implies that the approximation a rank-one Hermitian matrix by outer products of Gaussian i.i.d. vectors is at least as good (actually slightly better) as the approximation using Gaussian i.i.d. matrices, in terms of the expected squared approximation error.

4.2. Detector Performance

The performance of the detector can be evaluated and compared using region of convergence (ROC) plots for different values of M. Fig. 2 displays estimated ROC curves for the our detector, denoted as "least-squares", as well as for the squared correlator detector defined in (3) (which uses the true value of $\mathbf{x}$) and for a reference method. Here we use a random signal $\mathbf{s}$ drawn from a standard complex normal distribution, normalized to have unit Euclidean norm (the resulting signal vector is held fixed for all subsequent simulation trials). The reference method, denoted as "PhaseLift", is based on producing an estimate of $\mathbf{x}$ using the PhaseLift algorithm and plugging it into (3). It is seen that the ROC curves for the Phaselift method increase at a much faster rate than for our method (at the expense of computation time). For both methods, for any given False Alarm rate, the probability of detection increases with the number of measurements. Fig. 2 also indicates that our detector attains the squared correlator detector performance when the number of observations is large enough ($M = N^2 = 100$).

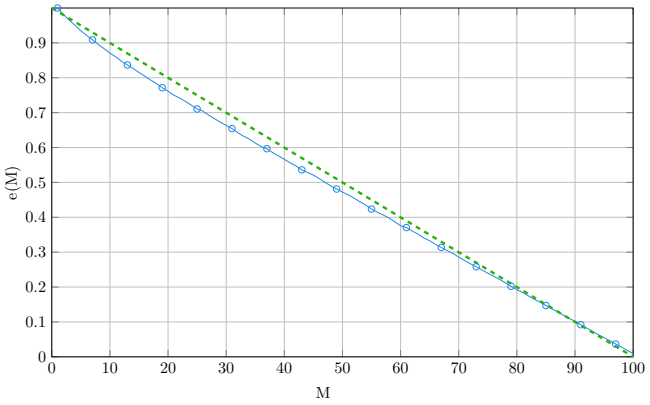

Fig. 1: Empirical squared approximation error ($N = 10$). The dashed line is the graph of $1 - \frac{M}{N^2}$ while the line with the circular markers represents the estimate of the squared approximation error. Each point is averaged over 10000 realizations.

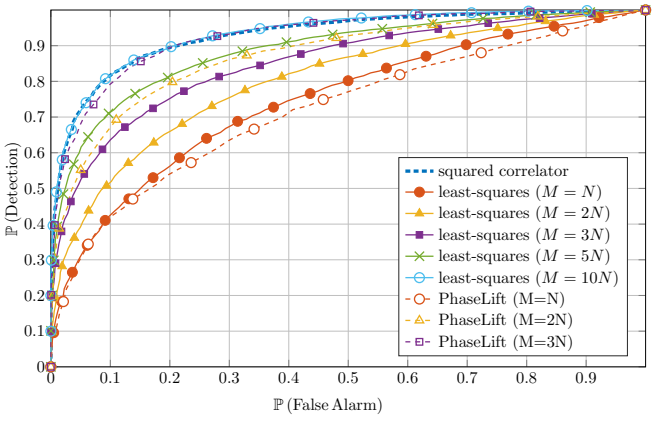

Fig. 2: Comparison of ROC curves for different values of M for our detector and for a reference method. Simulation parameters are $N = 10, A = 2, \sigma^2 = 1$. Simulations for our method are averaged over 10000 independent trials. PhaseLift simulations are averaged over 5000 independent trials.

5. CONCLUSION

We proposed a detector for the presence of a signal in noise given the squared amplitudes of linear measurements of an unobserved acquisition vector. For random measurements, the detector performance increases, on average, with the number of observations and attains the squared correlator performance when N^2 observations are used, where N is the signal length. In practice, sufficiently good detection performance may be obtained for fewer observations. Finally, we demonstrated that when using Fourier basis measurement vectors, our detector attain the squared correlator performance only when the signal vector has the same form as the measurement vectors.

6. REFERENCES

[1] R. P. Millane, "Phase retrieval in crystallography and optics," *JOSA A*, vol. 7, no. 3, pp. 394–411, 1990.

[2] A. Walther, "The question of phase retrieval in optics," *Journal of Modern Optics*, vol. 10, no. 1, pp. 41–49, 1963.

[3] C. Fienup and J. Dainty, "Phase retrieval and image reconstruction for astronomy," *Image Recovery: Theory and Application*, pp. 231–275, 1987.

[4] P. Yang, Z. Tan, A. Wiesel, and A. Nehora, "Power system state estimation using PMUs with imperfect synchronization," *Power Systems, IEEE Transactions on*, vol. 28, no. 4, pp. 4162–4172, 2013.

[5] Y. Shechtman, Y. Eldar, O. Cohen, H. Chapman, J. Miao, and M. Segev, "Phase retrieval with application to optical imaging: A contemporary overview," *Signal Processing Magazine, IEEE*, vol. 32, no. 3, pp. 87–109, May 2015.

[6] H. Ohlsson, A. Yang, R. Dong, and S. Sastry, "CPRL–an extension of compressive sensing to the phase retrieval problem," in *Advances in Neural Information Processing Systems*, 2012, pp. 1367–1375.

[7] Y. Shechtman, A. Beck, and Y. Eldar, "GESPAR: Efficient phase retrieval of sparse signals," *Signal Processing, IEEE Transactions on*, vol. 62, no. 4, pp. 928–938, Feb 2014.

[8] K. Jaganathan, S. Oymak, and B. Hassibi, "Sparse phase retrieval: Convex algorithms and limitations," in *Information Theory Proceedings (ISIT), 2013 IEEE International Symposium on*, July 2013, pp. 1022–1026.

[9] J. R. Fienup, "Phase retrieval algorithms: a comparison," *Appl. Opt.*, vol. 21, no. 15, pp. 2758–2769, Aug 1982.

[10] R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik*, no. 35, pp. 237–246, 1972.

[11] I. Waldspurger, A. D'aspremont, and S. Mallat, "Phase recovery, MaxCut and complex semidefinite programming," *Math. Program.*, vol. 149, no. 1-2, pp. 47–81, Feb. 2015.

[12] E. J. Candes, Y. C. Eldar, T. Strohmer, and V. Voroninski, "Phase retrieval via matrix completion," *SIAM Review*, vol. 57, no. 2, pp. 225–251, 2015.

[13] E. J. Candès and X. Li, "Solving quadratic equations via phaseLift when there are about as many equations as unknowns," *CoRR*, vol. abs/1208.6247.

[14] E. Candès, X. Li, and M. Soltanolkotabi, "Phase retrieval via wirtinger flow: Theory and algorithms," *Information Theory, IEEE Transactions on*, vol. 61, no. 4, pp. 1985–2007, April 2015.

[15] Y. Chen and E. J. Candes, "Solving Random Quadratic Systems of Equations Is Nearly as Easy as Solving Linear Systems," *ArXiv e-prints*, 2015.

[16] P. Netrapalli, P. J. 0002, and S. Sanghavi, "Phase retrieval using alternating minimization." *CoRR*, 2013.

[17] P. Schniter and S. Rangan, "Compressive phase retrieval via generalized approximate message passing," in *Communication, Control, and Computing (Allerton), 2012 50th Annual Allerton Conference on*, Oct 2012, pp. 815–822.

[18] R. Balan, B. G. Bodmann, P. G. Casazza, and D. Edidin, "Painless reconstruction from magnitudes of frame coefficients," *Journal of Fourier Analysis and Applications*, vol. 15, no. 4, pp. 488–501, 2009.

[19] R. Balan, P. Casazza, and D. Edidin, "On signal reconstruction without phase," *Applied and Computational Harmonic Analysis*, vol. 20, no. 3, pp. 345 – 356, 2006.

[20] A. S. Bandeira, J. Cahill, D. G. Mixon, and A. A. Nelson, "Saving phase: Injectivity and stability for phase retrieval," *Applied and Computational Harmonic Analysis*, vol. 37, no. 1, pp. 106–125, 2014.

[21] M. Hayes, "The reconstruction of a multidimensional sequence from the phase or magnitude of its fourier transform," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 30, no. 2, pp. 140–154, Apr 1982.

[22] K. Jaganathan, Y. C. Eldar, and B. Hassibi, "STFT phase retrieval: Uniqueness guarantees and recovery algorithms," *arXiv preprint arXiv:1508.02820*, 2015.

[23] H. Ohlsson and Y. C. Eldar, "On conditions for uniqueness in sparse phase retrieval," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 1841–1845.

[24] J. Cahill and X. Chen, "A note on scalable frames," *ArXiv e-prints*, 2013.