PhaseLin: Linear Phase Retrieval

Ramina Ghods[1], Andrew S. Lan[2], Tom Goldstein[3], and Christoph Studer[1]

[1]Cornell University, Ithaca, NY; rg548@cornell.edu, studer@cornell.edu
[2]Princeton University, Princeton, NJ; andrew.lan@princeton.edu
[3]University of Maryland, College Park, MD; tomg@cs.umd.edu

Abstract—**Phase retrieval deals with the recovery of complex- or real-valued signals from magnitude measurements. As shown recently, the method PhaseMax enables phase retrieval via convex optimization and without lifting the problem to a higher dimension. To succeed, PhaseMax requires an initial guess of the solution, which can be calculated via spectral initializers. In this paper, we show that with the availability of an initial guess, phase retrieval can be carried out with an ever simpler, linear procedure. Our algorithm, called PhaseLin, is the linear estimator that minimizes the mean squared error (MSE) when applied to the magnitude measurements. The linear nature of PhaseLin enables an exact and nonasymptotic MSE analysis for arbitrary measurement matrices. We furthermore demonstrate that by iteratively using PhaseLin, one arrives at an efficient phase retrieval algorithm that performs on par with existing convex and nonconvex methods on synthetic and real-world data.**

I. INTRODUCTION

Phase retrieval recovers the N-dimensional signal vector $\mathbf{x} \in \mathcal{H}^N$, with $\mathcal{H}$ being either the set of real ($\mathbb{R}$) or complex ($\mathbb{C}$) numbers, from the following nonlinear measurement process:

$$\mathbf{y} = |\mathbf{Ax} + \mathbf{n}^z|^2 + \mathbf{n}^y. \tag{1}$$

Here, the measurement vector $\mathbf{y} \in \mathbb{R}^M$ contains M real-valued and phase-less observations, the absolute square function $|\cdot|^2$ operates element-wise on vectors, $\mathbf{A} \in \mathcal{H}^{M \times N}$ is a known measurement matrix, and the vectors $\mathbf{n}^z \in \mathcal{H}^N$ and $\mathbf{n}^y \in \mathbb{R}^M$ model signal and measurement noise, respectively. In what follows, we assume a deterministic (and known) measurement matrix $\mathbf{A}$, but randomness in the signal $\mathbf{x}$ to be estimated as well as the two noise sources $\mathbf{n}^z$ and $\mathbf{n}^y$.

A. Relevant Prior Art

Phase retrieval has been studied extensively over the last few decades [1], [2] as it finds use in numerous applications, including X-ray crystallography [3]–[5], microscopy [6], [7], and imaging [8]. In its original form, the phase retrieval problem is nonconvex and was solved traditionally using alternating projection methods, such as the Gerchberg-Saxton [1] or Fienup [2]

methods. During the last few years, it has been shown that lifting the phase retrieval problem to a higher-dimensional space enables the use of semidefinite programming [9]–[11], which led to a revival of phase retrieval research. While lifting-based phase retrieval methods provide strong theoretical guarantees, their computational complexity and storage requirements quickly become prohibitive for high-dimensional problems (e.g., for the recovery of high-resolution images). To perform phase retrieval for high-dimensional problems with methods that provide theoretical performance guarantees, a number of nonconvex methods have been proposed and analyzed within the last years; see, e.g., [12]–[20]. All these methods rely on an accurate initial guess of the true signal to be recovered, which can be obtained by means of so-called spectral initializers [12], [15], [16], [19]–[22]. More recently, it has been shown in [23]–[27] that given such initial guesses, one can perform phase retrieval via the convex program PhaseMax, which avoids lifting and provides accurate performance guarantees.

B. Contributions

In this paper, we show that the availability of an initial guess enables the use of *linear* estimators to perform phase retrieval. Concretely, we propose a novel mean squared error (MSE)-optimal linear phase retrieval method we call *PhaseLin*. Our method provides exact and nonasymptotic expressions of the recovery MSE and is suitable for scenarios in which the measurement matrix is finite-dimensional, deterministic, and (possibly) structured—these aspects are in stark contrast to most existing theoretical results that are either of asymptotic nature and/or require randomness in the measurement matrix. We furthermore show that by iteratively using PhaseLin, one arrives at a phase retrieval algorithm that performs on par with existing methods on synthetic and real-world data.

C. Notation

Column vectors and matrices are denoted with lower- and upper-case bold letters, respectively. For a matrix $\mathbf{A}$, its transpose, Hermitian transpose, and (entry-wise) complex conjugate are denoted respectively by $\mathbf{A}^T$, $\mathbf{A}^H$, and $\mathbf{A}^*$. The entry on the mth row and nth column of $\mathbf{A}$ is denoted as $[\mathbf{A}]_{m,n}$, and $\mathbf{a}_m$ denotes the mth column vector. For a vector $\mathbf{a}$, its kth entry is denoted by a_k; the ℓ_2-norm norm is denoted by $\|\mathbf{a}\|_2$. The (entry-wise) Hadamard product and the trace operator are denoted by $\odot$ and $\text{tr}(\cdot)$, respectively. The notation

diag($\mathbf{a}$) stands for the square matrix with the vector $\mathbf{a}$ on its main diagonal; diag($\mathbf{A}$) denotes the column vector comprising the diagonal elements of the matrix $\mathbf{A}$. The operators $\Re(\cdot)$ and $\Im(\cdot)$ extract the real and imaginary parts of a complex-valued number, respectively; for a complex-valued vector $\mathbf{a}$, we use $\mathbf{a}_{\mathcal{R}}$ and $\mathbf{a}_{\mathcal{I}}$ to denote the real and imaginary parts. Expectation with respect to the random vector $\mathbf{x}$ is denoted by $\mathbb{E}_{\mathbf{x}}[\cdot]$.

II. MAIN RESULTS

We now present PhaseLin. We separately provide results for the recovery of real-valued and complex-valued signals. We then provide an exact expression for the recovery MSE of PhaseLin. We finally show how one can iteratively use PhaseLin to arrive at a powerful phase retrieval algorithm.

A. The Real Case: PhaseLin-$\mathbb{R}$

We first focus on the case where the signal to be recovered $\mathbf{x} \in \mathbb{R}^N$ and the measurement matrix $\mathbf{A} \in \mathbb{R}^{M \times N}$ are both real-valued. We need the following assumptions.

Assumptions 1. *We assume that the signal noise vector $\mathbf{n}^z$ is zero-mean Gaussian with covariance matrix $\mathbf{C}_{\mathbf{n}^z}$, i.e., $\mathbf{n}^z \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_{\mathbf{n}^z})$; the measurement noise vector $\mathbf{n}^y$ is Gaussian with mean $\bar{\mathbf{n}}^y$ and covariance matrix $\mathbf{C}_{\mathbf{n}^y}$, i.e., $\mathbf{n}^y \sim \mathcal{N}(\bar{\mathbf{n}}^y, \mathbf{C}_{\mathbf{n}^y})$. For the signal vector $\mathbf{x} \in \mathbb{R}^N$, we assume that we have an initial guess $\bar{\mathbf{x}} \in \mathbb{R}^N$ (e.g., obtained from a spectral initializer). We furthermore assume that the true signal can be written as $\mathbf{x} = \bar{\mathbf{x}} + \mathbf{e}$, where the error vector $\mathbf{e}$ denotes the difference between the initial guess $\bar{\mathbf{x}}$ and the true signal $\mathbf{x}$. We assume that the error vector follows a zero-mean Gaussian distribution with covariance $\mathbf{C}_{\mathbf{e}}$, i.e., $\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_{\mathbf{e}})$.*

With these assumptions, we can derive PhaseLin-$\mathbb{R}$; the proof of the following result is given in Appendix A.

Theorem 1 (PhaseLin-$\mathbb{R}$). *Under Assumptions 1, the linear estimate $\hat{\mathbf{x}}$ that minimizes the recovery MSE defined as*

$$MSE = \mathbb{E}_{\mathbf{e}, \mathbf{n}^z, \mathbf{n}^y} \left[\|\hat{\mathbf{x}} - \mathbf{x}\|_2^2 \right] \qquad (2)$$

is given by

$$\hat{\mathbf{x}} = \mathbf{C}_{\mathbf{x}, \mathbf{y}} \mathbf{v} + \bar{\mathbf{x}} \quad with \quad \mathbf{C}_{\mathbf{y}} \mathbf{v} = \mathbf{y} - \bar{\mathbf{y}},$$

where $\bar{\mathbf{x}}$ is an initial guess and

$$\bar{\mathbf{y}} = \mathrm{diag}(\mathbf{C}_{\mathbf{z}}) + |\bar{\mathbf{z}}|^2 + \bar{\mathbf{n}}^y$$
$$\mathbf{C}_{\mathbf{x}, \mathbf{y}} = 2 \mathbf{C}_{\mathbf{e}} \mathbf{A}^{\mathrm{T}} \mathrm{diag}(\bar{\mathbf{z}})$$
$$\mathbf{C}_{\mathbf{y}} = (4 \bar{\mathbf{z}} \bar{\mathbf{z}}^{\mathrm{T}} + 2 \mathbf{C}_{\mathbf{z}}) \odot \mathbf{C}_{\mathbf{z}} + \mathbf{C}_{\mathbf{n}^y},$$

with $\bar{\mathbf{z}} = \mathbf{A} \bar{\mathbf{x}}$ and $\mathbf{C}_{\mathbf{z}} = \mathbf{A} \mathbf{C}_{\mathbf{e}} \mathbf{A}^{\mathrm{T}} + \mathbf{C}_{\mathbf{n}^z}$.

The above result describes the linear estimate $\hat{\mathbf{x}}$ of the signal $\mathbf{x}$ to be recovered that minimizes the recovery MSE, given the phase-less measurements in $\mathbf{y}$ and an initial guess $\bar{\mathbf{x}}$. As shown in Lemma 3, we are able to provide a closed form expression for the recovery MSE of PhaseLin-$\mathbb{R}$.

Remark 1. *If the initial guess $\bar{\mathbf{x}}$ is zero, then the quantity $\bar{\mathbf{z}}$ is zero. In this situation, the matrix $\mathbf{C}_{\mathbf{x}, \mathbf{y}}$ is zero and, hence, the obtained estimate $\hat{\mathbf{x}}$ is zero as well. Clearly, for such an*

initial guess, PhaseLin fails at estimating the signal $\mathbf{x}$. By using spectral initializers to set the mean $\bar{\mathbf{x}}$, such as the ones proposed in [12], [15], [16], [19]–[22], PhaseLin performs well in practice; see Section III for simulation results.

B. The Complex Case: PhaseLin-$\mathbb{C}$

We now focus on the case where the signal to be recovered $\mathbf{x} \in \mathbb{C}^N$ and the measurement matrix $\mathbf{A} \in \mathbb{C}^{M \times N}$ are both complex-valued. The measurements, however, remain real-valued. We need the following assumptions.

Assumptions 2. *We assume that the signal noise vector $\mathbf{n}^z$ is circularly symmetric complex Gaussian with covariance matrix $\mathbf{C}_{\mathbf{n}^z}$, i.e., $\mathbf{n}^z \sim \mathcal{CN}(\mathbf{0}, \mathbf{C}_{\mathbf{n}^z})$. We assume that the error vector is also circularly symmetric complex Gaussian with covariance matrix $\mathbf{C}_{\mathbf{e}}$, i.e., $\mathbf{e} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C}_{\mathbf{e}})$. The remaining assumptions are identical to those in Assumptions 1.*

We can now derive PhaseLin-$\mathbb{C}$; the proof of the following result is given in Appendix B.

Theorem 2 (PhaseLin-$\mathbb{C}$). *Under Assumptions 2, the linear estimate $\hat{\mathbf{x}}$ that minimizes the MSE in (2) is given by*

$$\hat{\mathbf{x}} = \mathbf{C}_{\mathbf{x}, \mathbf{y}} \mathbf{v} + \bar{\mathbf{x}} \quad with \quad \mathbf{C}_{\mathbf{y}} \mathbf{v} = \mathbf{y} - \bar{\mathbf{y}},$$

where $\bar{\mathbf{x}}$ is an initial guess and

$$\bar{\mathbf{y}} = \mathrm{diag}(\mathbf{C}_{\mathbf{z}}) + |\bar{\mathbf{z}}|^2 + \bar{\mathbf{n}}^y$$
$$\mathbf{C}_{\mathbf{x}, \mathbf{y}} = \mathbf{C}_{\mathbf{e}} \mathbf{A}^{\mathrm{H}} \mathrm{diag}(\bar{\mathbf{z}})$$
$$\mathbf{C}_{\mathbf{y}} = 2 \Re \left\{ (\bar{\mathbf{z}} \bar{\mathbf{z}}^{\mathrm{H}}) \odot \mathbf{C}_{\mathbf{z}}^* \right\} + \mathbf{C}_{\mathbf{z}} \odot \mathbf{C}_{\mathbf{z}}^* + \mathbf{C}_{\mathbf{n}^y},$$

with $\bar{\mathbf{z}} = \mathbf{A} \bar{\mathbf{x}}$ and $\mathbf{C}_{\mathbf{z}} = \mathbf{A} \mathbf{C}_{\mathbf{e}} \mathbf{A}^{\mathrm{H}} + \mathbf{C}_{\mathbf{n}^z}$.

As for PhaseLin-$\mathbb{C}$, the above result describes the linear estimate $\hat{\mathbf{x}}$ of the signal $\mathbf{x}$ that minimizes the recovery MSE given the phase-less measurements in $\mathbf{y}$ and an initial guess $\bar{\mathbf{x}}$.

C. Exact Expression for the Recovery MSE

For both PhaseLin-$\mathbb{R}$ and PhaseLin-$\mathbb{C}$, the following result provides an exact and nonasymptotic expression of the recovery MSE; the proof is given in Appendix C.

Lemma 3 (MSE of PhaseLin). *Let either Assumptions 1 or Assumptions 2 hold. Furthermore, assume that $\mathbf{C}_{\mathbf{y}}$ is full rank. Then, the recovery MSE is given by*

$$MSE = \mathrm{tr}(\mathbf{C}_{\mathbf{e}} - \mathbf{C}_{\mathbf{x}, \mathbf{y}} \mathbf{C}_{\mathbf{y}}^{-1} \mathbf{C}_{\mathbf{x}, \mathbf{y}}^{\mathrm{H}}).$$

We emphasize that most existing phase retrieval methods either provide theoretical results that are exact in the asymptotic regime or provide upper bounds on the recovery error. In addition, virtually all existing theoretical results require randomness in the measurement matrix $\mathbf{A}$. In contrast, Lemma 3 assumes randomness in the signal to be recovered and the noise sources, is nonasymptotic and exact, and holds for arbitrary and deterministic measurement matrices.

Remark 2. *Lemma 3 requires the matrix $\mathbf{C}_{\mathbf{y}}$ to be full rank. It can be shown that this full-rank requirement is satisfied for most scenarios with nondegenerate measurement matrices $\mathbf{A}$ or for situations with nonzero measurement noise.*
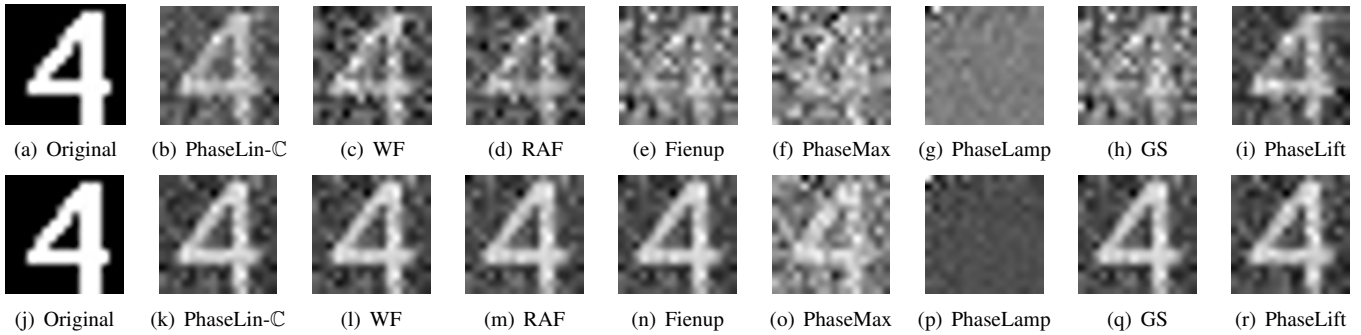
Fig. 1. (Top) Reconstructions of a 16×16-pixel image captured through a scattering medium (real-world dataset provided in [28], reconstructions using PhasePack [29]) with $M = 3N$ measurements using different phase retrieval methods. (Bottom) The same experiment but with $M = 9N$ measurements. See Table I for the N-MSE values and runtimes of each phase retrieval algorithm.

D. Iterative Variant of PhaseLin

The authors of [26] proposed an iterative scheme that can improve upon the performance of the PhaseMax formulation. The iterative method, called PhaseLamp, first applies PhaseMax with an estimate obtained from a spectral initializer. In each subsequent iteration, the result of the previous iteration is then used as a new initial guess for PhaseMax—this procedure is repeated for a predefined number of iterations or until convergence. Inspired by PhaseLamp, we propose to iteratively apply PhaseLin. The resulting method proceeds as follows.

We start at iteration $t = 0$ with a spectral initializer $\bar{\mathbf{x}}^{(0)}$. We run PhaseLin with this initial guess to obtain an estimate of the signal to be recovered $\hat{\mathbf{x}}^{(0)}$. We then take this estimate as a new initial guess, i.e., $\bar{\mathbf{x}}^{(1)} = \hat{\mathbf{x}}^{(0)}$ and re-use PhaseLin to obtain a (hopefully improved) estimate. Concretely, we perform

$$\bar{\mathbf{x}}^{(t+1)} = \hat{\mathbf{x}}^{(t)} = \text{PhaseLin}(\bar{\mathbf{x}}^{(t)}) \quad \text{for} \quad t = 0, \dots, t_{\max}$$

with the final estimate being $\hat{\mathbf{x}}^{(t_{\max})}$. In our experiments, we simply keep the same covariance matrix $\mathbf{C_e}$ during all iterations. We note that more sophisticated methods for selecting the error covariance matrix $\mathbf{C_e}$ on a per-iteration basis may yield improved performance; the design and analysis of such methods is left for future work.

III. NUMERICAL RESULTS

We now compare the performance of PhaseLin against existing phase retrieval methods for both real and synthetic data. Algorithm and initializer implementations and experimental setups were provided by PhasePack [29].

A. Image Recovery

We first test the performance of PhaseLin on an image reconstruction task. In this experiment, an image is captured through multiple scattering media, producing phase-less measurements [28]; our task is to reconstruct the original image. In Figure 1, we show the recovered images for a 16×16-pixel image taken from the dataset provided in [28] with $M = 3N$ and $M = 9N$ measurements, respectively. We compare PhaseLin to the Wirtinger flow (WF) [14], reweighed amplitude flow (RAF) [22], Fienup [2], PhaseMax [23]–[25], PhaseLamp [26], Gerchberg-Saxton (GS) [1], and PhaseLift [9]

TABLE I
RELATIVE RUNTIME (RR) AND NORMALIZED MSE (N-MSE) VERSUS OVERSAMPLING RATIO (OSR) FOR IMAGE RECOVERY AS IN FIGURE 1.

Metric OSR	RR $M = 3N$	N-MSE	RR $M = 9N$	N-MSE
PhaseLin-$\mathbb{C}$	1.00	0.3252	1.00	0.2783
WF [14]	0.61	0.4492	0.13	0.3069
RAF [22]	0.70	0.4769	0.16	0.2946
Fienup [2]	19.1	0.6070	0.45	0.2899
PhaseMax [25]	0.96	0.6254	0.42	0.4872
PhaseLamp [26]	13.3	0.6843	5.31	0.6848
GS [1]	17.9	0.6036	0.43	0.2899
PhaseLift [9]	170	0.3195	35.0	0.2786

methods. For each method, we use the asymptotically-optimal spectral initializer [20]. For PhaseLin, we use the complex version and perform $t_{\max} = 10$ iterations. Table I provides the associated normalized MSE (N-MSE), which is defined as [29]

$$\textit{N-MSE} = \min_{\alpha \in \mathbb{C}} \|\mathbf{x} - \alpha \hat{\mathbf{x}}\|^2 / \|\mathbf{x}\|^2$$

as well as the relative runtime (compared to that of PhaseLin).

We see that for $M = 3N$, PhaseLin achieves a lower N-MSE than all other methods except for PhaseLift, which is significantly more complex as it must solve a large semidefinite program. Furthermore, for this image dimension, the runtime of PhaseLin is comparable to that of the Wirtinger Flow, RAF, and PhaseMax methods. For $M = 9N$, PhaseLin achieves the lowest N-MSE. However, increasing the number of measurements also increases the complexity of PhaseLin.

B. Synthetic Data

To further illustrate the efficacy of PhaseLin, we study its performance for different oversampling ratios M/N using synthetic data in Figure 2. We use the same (empirical) measurement matrix from the above experiment and generate synthetic signals of dimension 256 from a zero-mean Gaussian distribution as $\mathbf{x} \sim \mathcal{CN}(0, 2\mathbf{I}_{256 \times 256})$. We then apply each method to 10 randomly generated instances of $\mathbf{x}$ and plot the median N-MSE. PhaseLin-$\mathbb{C}$ with only 15 iterations outperforms existing methods at small oversampling ratios except PhaseLift (which has much higher computational complexity).
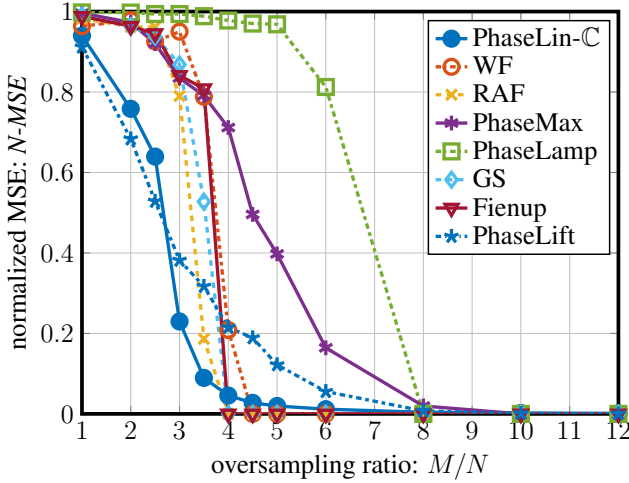
Fig. 2. Normalized MSE of different phase retrieval methods as a function of the oversampling ratio M/N for Gaussian data with $N = 256$ and the same measurement process used in Figure 1. PhaseLin with only 15 iterations outperforms most existing methods in the regime of small oversampling ratios.

We expect that performance could be improved at higher sampling ratios with more sophisticated strategies to select the error covariance $\mathbf{C_e}$ in each PhaseLin iteration.

IV. CONCLUSIONS

We have proposed PhaseLin, an MSE-optimal linear estimator that recovers signals from magnitude measurements. PhaseLin requires an initial guess of the true signal that can be obtained from spectral initializers and enables an exact and nonasymptotic analysis of the recovery MSE. Furthermore, we have demonstrated that PhaseLin performs on par with existing phase retrieval methods using real images and synthetic data when used in an iterative manner.

There are many avenues for ongoing work. First, analyzing the MSE of the iterative version of PhaseLin is a challenging open research problem. Second, tuning the error covariance in each PhaseLin iteration and further reducing our method's computational complexity is part of ongoing work.

APPENDIX A
PROOF OF THEOREM 1

Our goal is to derive a linear (or affine) estimate $\hat{\mathbf{x}}^{\text{L-MMSE}} = \mathbf{W}\mathbf{y} + \mathbf{b}$ for the unknown signal $\mathbf{x}$ that minimizes the MSE defined in (2). The necessary quantities of this linear minimum MSE (L-MMSE) estimator are given by

$$\mathbf{W} = \mathbf{C_{x,y}}\mathbf{C_y}^{-1} \quad \text{and} \quad \mathbf{b} = \bar{\mathbf{x}} - \mathbf{W}\bar{\mathbf{y}}, \qquad (3)$$

with $\bar{\mathbf{x}} = \mathbb{E}[\mathbf{x}]$, $\bar{\mathbf{y}} = \mathbb{E}[\mathbf{y}]$, and

$$\mathbf{C_{x,y}} = \text{Cov}[\mathbf{x}, \mathbf{y}] = \mathbb{E}\big[(\mathbf{x} - \bar{\mathbf{x}})(\mathbf{y} - \bar{\mathbf{y}})^{\text{H}}\big] = \mathbb{E}\big[\mathbf{x}\mathbf{y}^{\text{H}}\big] - \bar{\mathbf{x}}\bar{\mathbf{y}}^{\text{H}}$$
$$\mathbf{C_y} = \text{Var}[\mathbf{y}] = \mathbb{E}\big[(\mathbf{y} - \bar{\mathbf{y}})(\mathbf{y} - \bar{\mathbf{y}})^{\text{H}}\big] = \mathbb{E}\big[\mathbf{y}\mathbf{y}^{\text{H}}\big] - \bar{\mathbf{y}}\bar{\mathbf{y}}^{\text{H}},$$

where we assumed that $\mathbf{C_y}$ is full rank. Our task is to compute the remaining quantities $\bar{\mathbf{y}}$, $\mathbf{C_{x,y}}$, and $\mathbf{C_y}$. We will frequently use the following lemma with proof given in [30, Sec. 3.1].

Lemma 4. *Let* $(u_1, u_2) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ *be a pair of real-valued jointly Gaussian random variables with covariance matrix*

$$\boldsymbol{\Sigma} = \left[\begin{array}{cc} \sigma_1^2 & \sigma_{1,2}^2 \\ \sigma_{1,2}^2 & \sigma_2^2 \end{array} \right].$$

Then, for $i = 1, 2$, *the pair of random variables* (ν_1, ν_2) *with* $\nu_1 = u_1^2$ *and* $\nu_2 = u_2^2$ *follows the bivariate folded normal distribution with moments*

$$\bar{\nu}_i = \mathbb{E}\big[u_i^2\big] = \sigma_i^2 + \mu_i^2$$
$$[\mathbf{C_\nu}]_{1,2} = \mathbb{E}[(\nu_1 - \bar{\nu}_1)(\nu_2 - \bar{\nu}_2)] = 4\mu_1\mu_2\sigma_{1,2}^2 + 2\sigma_{1,2}^4$$
$$[\mathbf{C_\nu}]_{i,i} = \mathbb{E}\big[(\nu_i - \bar{\nu}_i)^2\big] = 2\sigma_i^4 + 4\mu_i^2\sigma_i^2.$$

Define the vector $\mathbf{z} = \mathbf{A}\mathbf{x} + \mathbf{n}^z$ to contain the phased measurements, $\bar{\mathbf{z}} = \mathbb{E}[\mathbf{z}]$ its mean vector, and $\mathbf{C_z}$ its covariance matrix. Clearly, under Assumptions 1, the phased observations in $\mathbf{z}$ are jointly Gaussian, with mean $\bar{\mathbf{z}} = \mathbf{A}\bar{\mathbf{x}}$ and covariance matrix $\mathbf{C_z} = \mathbf{A}\mathbf{C_e}\mathbf{A}^{\text{T}} + \mathbf{C_{n^z}}$. We are now ready to compute the missing quantities $\bar{\mathbf{y}}$, $\mathbf{C_{x,y}}$, and $\mathbf{C_y}$.

1) Phased Measurement Mean $\bar{\mathbf{y}}$: Using Lemma 4, we have

$$\bar{y}_m = \mathbb{E}\big[|z_m|^2 + \mathbf{n}^y\big] = \sigma_m^2 + |\bar{z}_m|^2 + \bar{n}_m^y, \qquad (4)$$

where $\sigma_m^2 = [\mathbf{C_z}]_{m,m}$. Hence, the quantity $\bar{\mathbf{y}}$ reads

$$\bar{\mathbf{y}} = \text{diag}(\mathbf{C_z}) + |\bar{\mathbf{z}}|^2 + \bar{\mathbf{n}}^y.$$

2) Cross-Covariance Matrix $\mathbf{C_{x,y}}$: To compute the cross-covariance matrix $\mathbf{C_{x,y}}$, we will use a classical result due to Brillinger given in [31, Lem.1] that states

$$\begin{aligned} \text{Cov}[y_m, x_n] &= \text{Cov}\big[|z_m|^2, x_n\big] \\ &= \frac{\text{Cov}\big[|z_m|^2, z_m\big]}{\text{Var}[z_m]} \text{Cov}[z_m, x_n]. \end{aligned}$$

By using Stein's lemma [32], we have

$$\frac{\text{Cov}\big[|z_m|^2, z_m\big]}{\text{Var}[z_m]} = \mathbb{E}[2z_m] = 2\bar{z}_m.$$

We also have

$$\begin{aligned} \text{Cov}[z_m, x_n] &= \mathbb{E}[(z_m - \bar{z}_m)(x_n - \bar{x}_n)^*] \\ &= \mathbf{a}_m^{\text{T}} \mathbb{E}[(\mathbf{x} - \bar{\mathbf{x}})(x_n - \bar{x}_n)^*] = \mathbf{a}_m^{\text{T}}[\mathbf{C_e}]_{:,n}, \end{aligned}$$

where $[\mathbf{C_e}]_{:,n}$ corresponds to the nth column of $\mathbf{C_e}$ and $\mathbf{a}_m^{\text{T}}$ to the mth row of $\mathbf{A}$. Hence, we obtain

$$[\mathbf{C_{y,x}}]_{m,n} = 2\bar{z}_m \mathbf{a}_m^{\text{T}}[\mathbf{C_e}]_{:,n}, \qquad (5)$$

and $\mathbf{C_{x,y}} = \mathbf{C_{y,x}}^{\text{H}}$. In compact vector form, the cross-covariance matrix $\mathbf{C_{x,y}}$ reads

$$\mathbf{C_{x,y}} = 2\mathbf{C_e}\mathbf{A}^{\text{T}}\text{diag}(\bar{\mathbf{z}}).$$

3) Observation Covariance Matrix $\mathbf{C_y}$: The last quantity required is the matrix $\mathbf{C_y}$. We compute the following quantity:

$$\begin{aligned} [\mathbf{C_y}]_{m,m'} &= \mathbb{E}[(y_m - \bar{y}_m)(y_m - \bar{y}_m)^*] \\ &= \mathbb{E}[(|z_m|^2 + n_m^y - (\mathbb{E}\big[|z_m|^2\big] + \bar{n}_m^y)) \\ &\quad \times (|z_{m'}|^2 + n_{m'}^y - (\mathbb{E}\big[|z_{m'}|^2\big] + \bar{n}_{m'}^y))^*] \end{aligned}$$

$$= \mathbb{E}[(|z_m|^2 - \mathbb{E}\left[|z_m|^2\right])(|z_{m'}|^2 - \mathbb{E}\left[|z_{m'}|^2\right])^*]$$
$$+ [\mathbf{C_{n^y}}]_{m,m'}.$$

Using Lemma 4, we can compute the above expression as follows. For $m = m'$, we have

$$[\mathbf{C_y}]_{m,m} = 2\sigma_m^4 + 4\bar{z}_m^2\sigma_m^2,$$

where $\sigma_m^2 = [\mathbf{C_z}]_{m,m}$. Similarly, for $m \neq m'$, we have

$$[\mathbf{C_y}]_{m,m'} = 4\bar{z}_m\bar{z}_{m'}\sigma_{mm'}^2 + 2\sigma_{mm'}^4,$$

where $\sigma_{mm'}^2 = [\mathbf{C_z}]_{m,m'}$. In summary, the observation covariance matrix $\mathbf{C_y}$ is given by

$$\mathbf{C_y} = (4\bar{\mathbf{z}}\bar{\mathbf{z}}^\mathsf{T} + 2\mathbf{C_z}) \odot \mathbf{C_z} + \mathbf{C_{n^y}}.$$

APPENDIX B
PROOF OF THEOREM 2

For complex-valued signals, the so-called widely linear minimum MSE (WL-MMSE) estimator [33], [34], which is of the form $\mathbf{W}_1\mathbf{y} + \mathbf{W}_2\mathbf{y}^*$, often provides superior results compared to the standard L-MMSE estimator. In our setting, both the WL-MMSE and L-MMSE estimators yield the same results. Hence, for Assumptions 2, the linear estimator is simply given by $\hat{\mathbf{x}} = \mathbf{W}\mathbf{y} + \mathbf{b}$ with $\mathbf{W}$ and $\mathbf{b}$ given in (3). As a consequence, we only need to compute the following quantities for complex-valued signals: $\bar{\mathbf{y}}$, $\mathbf{C_{x,y}}$, and $\mathbf{C_y}$.

Let $\mathbf{z} = \mathbf{A}\mathbf{x} + \mathbf{n}^z$ contain the phased measurements, $\bar{\mathbf{z}} = \mathbb{E}[\mathbf{z}]$ denote its mean vector, and $\mathbf{C_z}$ denote its covariance matrix. Given Assumptions 2, the phased observations in $\mathbf{z}$ are jointly complex Gaussian, where we can easily compute the mean $\bar{\mathbf{z}} = \mathbf{A}\bar{\mathbf{x}}$ and covariance $\mathbf{C_z} = \mathbf{A}\mathbf{C_e}\mathbf{A}^\mathsf{H} + \mathbf{C_{n^z}}$. From the covariance $\mathbf{C_z}$, we can easily extract the covariance of the real and imaginary parts of $\mathbf{z}$ separately as follows:

$$\mathbb{E}[\mathbf{z}_\mathcal{R}\mathbf{z}_\mathcal{R}^*] \stackrel{(a)}{=} \mathbb{E}[\mathbf{z}_\mathcal{I}\mathbf{z}_\mathcal{I}^*] = 1/2\Re\{\mathbf{C_z}\} = 1/2\mathbf{C}_{\mathbf{z},\mathcal{R}} \qquad (6)$$

$$\mathbb{E}[\mathbf{z}_\mathcal{I}\mathbf{z}_\mathcal{R}^*] \stackrel{(a)}{=} -\mathbb{E}[\mathbf{z}_\mathcal{R}\mathbf{z}_\mathcal{I}^*] = 1/2\Im\{\mathbf{C_z}\} = 1/2\mathbf{C}_{\mathbf{z},\mathcal{I}}. \qquad (7)$$

Here, (a) follows from the circular symmetry of the complex-valued random variable $\mathbf{x}$. We will use these covariances to compute $\mathbf{C_y}$ later in the section. We are now ready to compute the missing quantities $\bar{\mathbf{y}}$, $\mathbf{C_{x,y}}$, and $\mathbf{C_y}$.

1) Phased Measurement Mean $\bar{\mathbf{y}}$: To compute the entries of $\bar{\mathbf{y}}$, we have

$$\bar{y}_m = \mathbb{E}\left[|z_m|^2 + \mathbf{n}^y\right] = \mathbb{E}\left[|z_{m,\mathcal{R}}|^2 + |z_{m,\mathcal{I}}|^2 + n_m^y\right]$$
$$\stackrel{(b)}{=} |\bar{z}_{m,\mathcal{R}}|^2 + \frac{\sigma_{m,\mathcal{R}}^2}{2} + |\bar{z}_{m,\mathcal{I}}|^2 + \frac{\sigma_{m,\mathcal{R}}^2}{2} + \bar{n}_m^y$$
$$\stackrel{(c)}{=} |\bar{z}_m|^2 + \sigma_m^2 + \bar{n}_m^y, \qquad (8)$$

where $\sigma_{m,\mathcal{R}}^2 = [\mathbf{C}_{\mathbf{z},\mathcal{R}}]_{m,m}$ and $\sigma_m^2 = [\mathbf{C_z}]_{m,m}$. Here, (b) follows from Lemma 4, and (c) follows from the fact that we can conclude $[\mathbf{C}_{\mathbf{z},\mathcal{R}}]_{m,m} = [\mathbf{C_z}]_{m,m}$ from equations (6) and (7). Hence, $\bar{\mathbf{y}}$ in vector form reads

$$\bar{\mathbf{y}} = \text{diag}(\mathbf{C_z}) + |\bar{\mathbf{z}}|^2 + \bar{\mathbf{n}}^y.$$

2) Cross-Covariance Matrix $\mathbf{C_{x,y}}$: We next compute the individual entries of the cross-covariance matrix

$$[\mathbf{C_{x,y}}]_{n,m} = \mathbb{E}[(x_n - \bar{x}_n)(y_m - \bar{y}_m)^*] = \mathbb{E}[x_n y_m^*] - \bar{x}_n\bar{y}_m^*.$$

Let us first focus on the quantity $\mathbb{E}[x_n y_m^*]$. We have

$$\mathbb{E}[x_n y_m^*]$$
$$= \mathbb{E}\left[x_n\left((\sum_{i=1}^N A_{m,i}x_i + n_m^z)(\sum_{j=1}^N A_{m,j}x_j + n_m^z)^* + n_m^y\right)^*\right]$$
$$= \mathbb{E}\left[x_n \sum_{i=1}^N A_{m,i}x_i \sum_{j=1}^N A_{m,j}^* x_j^*\right] + \bar{x}_n[\mathbf{C_{n^z}}]_{m,m} + \bar{x}_n\bar{n}_m^y$$
$$= \bar{x}_n \sum_{i=1}^N \sum_{j=1}^N A_{m,i}A_{m,j}^* \bar{x}_i\bar{x}_j^* + \bar{x}_n \sum_{i=1}^N \sum_{j=1}^N A_{m,i}A_{m,j}^* [\mathbf{C_e}]_{i,j}$$
$$+ \sum_{i=1}^N A_{m,i}\bar{x}_i \sum_{j=1}^N A_{m,j}^* [\mathbf{C_e}]_{n,j} + \bar{x}_n[\mathbf{C_{n^z}}]_{m,m} + \bar{x}_n\bar{n}_m^y. \qquad (9)$$

Next, using (8), we obtain

$$\bar{x}_n\bar{y}_m^* = \bar{x}_n(\sigma_n^2 + |\bar{z}_m|^2 + \bar{n}_m^y)$$
$$= \bar{x}_n([\mathbf{A}\mathbf{C_e}\mathbf{A}^\mathsf{H} + \mathbf{C_{n^z}}]_{m,m} + |\mathbf{a}_m^\mathsf{H}\bar{\mathbf{x}}|^2 + \bar{n}_m^y)$$
$$= \bar{x}_n \sum_{i=1}^N \sum_{j=1}^N A_{m,i}A_{m,j}^* [\mathbf{C_e}]_{i,j} + \bar{x}_n[\mathbf{C_{n^z}}]_{m,m}$$
$$+ \bar{x}_n \sum_{i=1}^N \sum_{j=1}^N A_{m,i}A_{m,j}^* \bar{x}_i\bar{x}_j^* + \bar{x}_n\bar{n}_m^y. \qquad (10)$$

Hence, from (9) and (10), we finally get

$$[\mathbf{C_{x,y}}]_{n,m} = \sum_{i=1}^N A_{m,i}\bar{x}_i \sum_{j=1}^N A_{m,j}^* [\mathbf{C_e}]_{n,j},$$

which, in compact form, reads $\mathbf{C_{x,y}} = \mathbf{C_e}\mathbf{A}^\mathsf{H}\text{diag}(\bar{\mathbf{z}})$.

3) Observation Covariance Matrix $\mathbf{C_y}$: To obtain the entries of the matrix $\mathbf{C_y}$, we first calculate

$$[\mathbf{C_y}]_{m,m'} = \mathbb{E}[(y_m - \bar{y}_m)(y_m - \bar{y}_m)^*]$$
$$= \mathbb{E}[(|z_m|^2 + n_m^y - (\mathbb{E}\left[|z_m|^2\right] + \bar{n}_m^y))$$
$$\times (|z_{m'}|^2 + n_{m'}^y - (\mathbb{E}\left[|z_{m'}|^2\right] + \bar{n}_{m'}^y))^*]$$
$$= \mathbb{E}[(|z_{m,\mathcal{R}}|^2 - \mathbb{E}\left[|z_{m,\mathcal{R}}|^2\right] + |z_{m,\mathcal{I}}|^2 - \mathbb{E}\left[|z_{m,\mathcal{I}}|^2\right])$$
$$\times (|z_{m',\mathcal{R}}|^2 - \mathbb{E}\left[|z_{m',\mathcal{R}}|^2\right] + |z_{m',\mathcal{I}}|^2 - \mathbb{E}\left[|z_{m',\mathcal{I}}|^2\right])^*]$$
$$+ [\mathbf{C_{n^y}}]_{m,m'}.$$

Using the expressions in (6) and (7) together with Lemma 4, we can compute the above expressions. For $m = m'$, we have

$$[\mathbf{C_y}]_{m,m} = 2\frac{\sigma_{m,\mathcal{R}}^4}{4} + 4\bar{z}_{m,\mathcal{R}}^2\frac{\sigma_{m,\mathcal{R}}^2}{2} + 2\frac{\sigma_{m,\mathcal{R}}^4}{4} + 4\bar{z}_{m,\mathcal{I}}^2\frac{\sigma_{m,\mathcal{R}}^2}{2}$$
$$+ 4\bar{z}_{m,\mathcal{I}}\bar{z}_{m,\mathcal{R}}\frac{\sigma_{m,\mathcal{I}}^2}{2} + 2\frac{\sigma_{m,\mathcal{I}}^4}{4} + 4\bar{z}_{m,\mathcal{R}}\bar{z}_{m,\mathcal{I}}\frac{\sigma_{m,\mathcal{I}}^2}{2}$$
$$+ 2\frac{\sigma_{m,\mathcal{I}}^4}{4} + [\mathbf{C_{n^y}}]_{m,m},$$

where $\sigma_{m,\mathcal{R}}^2 = [\mathbf{C}_{\mathbf{z},\mathcal{R}}]_{m,m}$ and $\sigma_{m,\mathcal{I}}^2 = [\mathbf{C}_{\mathbf{z},\mathcal{I}}]_{m,m}$. From (7) we see that $\sigma_{m,\mathcal{I}}^2 = 0$. Hence, $[\mathbf{C}_{\mathbf{y}}]_{m,m}$ can be written as

$$[\mathbf{C}_{\mathbf{y}}]_{m,m} = \sigma_{m,\mathcal{R}}^4 + 2(\bar{z}_{m,\mathcal{R}}^2 + \bar{z}_{m,\mathcal{I}}^2)\sigma_{m,\mathcal{R}}^2 + [\mathbf{C}_{\mathbf{n}^y}]_{m,m}.$$

Similarly, for $m \neq m'$ we have

$$\begin{aligned}
[\mathbf{C}_{\mathbf{y}}]_{m,m'} = {}& 4\bar{z}_{m,\mathcal{R}}\bar{z}_{m',\mathcal{R}}\frac{\sigma_{mm',\mathcal{R}}^2}{2} + 2\frac{\sigma_{mm',\mathcal{R}}^4}{4} \\
& + 4\bar{z}_{m,\mathcal{I}}\bar{z}_{m',\mathcal{I}}\frac{\sigma_{mm',\mathcal{R}}^2}{2} + 2\frac{\sigma_{mm',\mathcal{R}}^4}{4} \\
& - 4\bar{z}_{m,\mathcal{R}}\bar{z}_{m',\mathcal{I}}\frac{\sigma_{mm',\mathcal{I}}^2}{2} + 2\frac{\sigma_{mm',\mathcal{I}}^4}{4} \\
& + 4\bar{z}_{m,\mathcal{I}}\bar{z}_{m',\mathcal{R}}\frac{\sigma_{mm',\mathcal{I}}^2}{2} + 2\frac{\sigma_{mm',\mathcal{I}}^4}{4} \\
& + [\mathbf{C}_{\mathbf{n}^y}]_{m,m'},
\end{aligned}$$

where $\sigma_{mm',\mathcal{R}}^2 = [\mathbf{C}_{\mathbf{z},\mathcal{R}}]_{m,m'}$ and $\sigma_{mm',\mathcal{I}}^2 = [\mathbf{C}_{\mathbf{z},\mathcal{I}}]_{m,m'}$. As a result, the observation covariance matrix $\mathbf{C}_{\mathbf{y}}$ is given by

$$\mathbf{C}_{\mathbf{y}} = 2\Re\left\{(\bar{\mathbf{z}}\bar{\mathbf{z}}^{\mathsf{H}}) \odot \mathbf{C}_{\mathbf{z}}^*\right\} + \mathbf{C}_{\mathbf{z}} \odot \mathbf{C}_{\mathbf{z}}^* + \mathbf{C}_{\mathbf{n}^y}.$$

APPENDIX C
PROOF OF LEMMA 3

The MSE of the PhaseLin estimator is given in (2). If $\mathbf{C}_{\mathbf{y}}$ is full rank, then $\hat{\mathbf{x}} = \mathbf{C}_{\mathbf{x},\mathbf{y}}\mathbf{C}_{\mathbf{y}}^{-1}(\mathbf{y} - \bar{\mathbf{y}}) + \bar{\mathbf{x}}$. Inserting this PhaseLin estimator in the MSE expression leads to

$$\begin{aligned}
MSE = {}& \mathbb{E}_{\mathbf{e},\mathbf{n}^z,\mathbf{n}^y}\left[\|\mathbf{C}_{\mathbf{x},\mathbf{y}}\mathbf{C}_{\mathbf{y}}^{-1}(\mathbf{y} - \bar{\mathbf{y}}) + \bar{\mathbf{x}} - \mathbf{x}\|_2^2\right] \\
= {}& \mathrm{tr}\left(\mathbb{E}_{\mathbf{e},\mathbf{n}^z,\mathbf{n}^y}\left[\mathbf{C}_{\mathbf{x},\mathbf{y}}\mathbf{C}_{\mathbf{y}}^{-1}(\mathbf{y} - \bar{\mathbf{y}})(\mathbf{y} - \bar{\mathbf{y}})^{\mathsf{H}}\mathbf{C}_{\mathbf{y}}^{-1}\mathbf{C}_{\mathbf{x},\mathbf{y}}^{\mathsf{H}}\right]\right) \\
& - \mathrm{tr}\left(\mathbb{E}_{\mathbf{e},\mathbf{n}^z,\mathbf{n}^y}\left[\mathbf{C}_{\mathbf{x},\mathbf{y}}\mathbf{C}_{\mathbf{y}}^{-1}(\mathbf{y} - \bar{\mathbf{y}})(\mathbf{x} - \bar{\mathbf{x}})^{\mathsf{H}}\right]\right) \\
& - \mathrm{tr}\left(\mathbb{E}_{\mathbf{e},\mathbf{n}^z,\mathbf{n}^y}\left[(\mathbf{x} - \bar{\mathbf{x}})(\mathbf{y} - \bar{\mathbf{y}})^{\mathsf{H}}\mathbf{C}_{\mathbf{y}}^{-1}\mathbf{C}_{\mathbf{x},\mathbf{y}}^{\mathsf{H}}\right]\right) \\
& + \mathrm{tr}\left(\mathbb{E}_{\mathbf{e}}\left[(\mathbf{x} - \bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}})^{\mathsf{H}}\right]\right) \\
= {}& \mathrm{tr}\left(\mathbf{C}_{\mathbf{x},\mathbf{y}}\mathbf{C}_{\mathbf{y}}^{-1}\mathbf{C}_{\mathbf{x},\mathbf{y}}^{\mathsf{H}}\right) - \mathrm{tr}\left(\mathbf{C}_{\mathbf{x},\mathbf{y}}\mathbf{C}_{\mathbf{y}}^{-1}\mathbf{C}_{\mathbf{y},\mathbf{x}}\right) \\
& - \mathrm{tr}\left(\mathbf{C}_{\mathbf{x},\mathbf{y}}\mathbf{C}_{\mathbf{y}}^{-1}\mathbf{C}_{\mathbf{x},\mathbf{y}}^{\mathsf{H}}\right) + \mathrm{tr}\left(\mathbf{C}_{\mathbf{e}}\right) \\
= {}& \mathrm{tr}\left(\mathbf{C}_{\mathbf{e}} - \mathbf{C}_{\mathbf{x},\mathbf{y}}\mathbf{C}_{\mathbf{y}}^{-1}\mathbf{C}_{\mathbf{y},\mathbf{x}}\right),
\end{aligned}$$

where we have used the fact that $\mathbf{C}_{\mathbf{y},\mathbf{x}} = \mathbf{C}_{\mathbf{x},\mathbf{y}}^{\mathsf{H}}$.

REFERENCES

[1] R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik*, vol. 35, pp. 237–246, Aug. 1972.

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

[3] R. W. Harrison, "Phase problem in crystallography," *J. Opt. Soc. Am. A*, vol. 10, no. 5, pp. 1046–1055, May 1993.

[4] J. Miao, T. Ishikawa, Q. Shen, and T. Earnest, "Extending X-ray crystallography to allow the imaging of noncrystalline materials, cells, and single protein complexes," *Ann. Rev. Phys. Chem.*, vol. 59, pp. 387–410, Nov. 2008.

[5] F. Pfeiffer, T. Weitkamp, O. Bunk, and C. David, "Phase retrieval and differential phase-contrast imaging with low-brilliance X-ray sources," *Nat. Phys.*, vol. 2, no. 4, pp. 258–261, Apr. 2006.

[6] S. S. Kou, L. Waller, G. Barbastathis, and C. J. Sheppard, "Transport-of-intensity approach to differential interference contrast (TI-DIC) microscopy for quantitative phase imaging," *Opt. Lett.*, vol. 35, no. 3, pp. 447–449, Feb. 2010.

[7] H. Faulkner and J. Rodenburg, "Movable aperture lensless transmission microscopy: A novel phase retrieval algorithm," *Phys. Rev. Lett.*, vol. 93, no. 2, Jul. 2004.

[8] J. Holloway, M. S. Asif, M. K. Sharma, N. Matsuda, R. Horstmeyer, O. Cossairt, and A. Veeraraghavan, "Toward long-distance subdiffraction imaging using coherent camera arrays," *IEEE Trans. Comput. Imag.*, vol. 2, no. 3, pp. 251–265, Sept. 2016.

[9] E. J. Candès, T. Strohmer, and V. Voroninski, "PhaseLift: Exact and stable signal recovery from magnitude measurements via convex programming," *Commun. Pure Appl. Math.*, vol. 66, no. 8, pp. 1241–1274, 2013.

[10] E. J. Candès and X. Li, "Solving quadratic equations via phaselift when there are about as many equations as unknowns," *Found. Comput. Math.*, vol. 14, no. 5, pp. 1017–1026, Oct. 2014.

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

[12] P. Netrapalli, P. Jain, and S. Sanghavi, "Phase retrieval using alternating minimization," in *Adv. Neural Inf. Process. Syst.*, 2013, pp. 2796–2804.

[13] P. Schniter and S. Rangan, "Compressive phase retrieval via generalized approximate message passing," *IEEE Trans. Sig. Process.*, vol. 63, no. 4, pp. 1043–1055, Feb. 2015.

[14] E. J. Candès, X. Li, and M. Soltanolkotabi, "Phase retrieval via Wirtinger flow: Theory and algorithms," *IEEE Trans. Inf. Theory*, vol. 61, no. 4, pp. 1985–2007, Feb. 2015.

[15] Y. Chen and E. Candès, "Solving random quadratic systems of equations is nearly as easy as solving linear systems," in *Adv. Neural Inf. Process. Syst.*, 2015, pp. 739–747.

[16] G. Wang, G. B. Giannakis, and Y. C. Eldar, "Solving systems of random quadratic equations via truncated amplitude flow," *arXiv preprint: 1605.08285*, Aug. 2017.

[17] K. Wei, "Solving systems of phaseless equations via kaczmarz methods: A proof of concept study," *Inverse Probl.*, vol. 31, no. 12, p. 125008, 2015.

[18] W. Zeng and H. So, "Coordinate descent algorithms for phase retrieval," *arXiv preprint: 1706.03474*, June 2017.

[19] Y. M. Lu and G. Li, "Phase transitions of spectral initialization for high-dimensional nonconvex estimation," *arXiv preprint: 1702.06435*, Apr. 2017.

[20] M. Mondelli and A. Montanari, "Fundamental limits of weak recovery with applications to phase retrieval," *arXiv preprint: 1708.05932*, Sep. 2017.

[21] P. Chen, A. Fannjiang, and G. Liu, "Phase retrieval with one or two diffraction patterns by alternating projections of the null vector," *arXiv preprint: 1510.07379*, Apr. 2015.

[22] G. Wang, G. B. Giannakis, Y. Saad, and J. Chen, "Solving almost all systems of random quadratic equations," *arXiv preprint: 1705.10407*, May 2017.

[23] S. Bahmani and J. Romberg, "Phase retrieval meets statistical learning theory: A flexible convex relaxation," in *Proc. Intl. Conf. on Artificial Intelligence and Statistics (AISTATS)*, May 2017, pp. 252–260.

[24] T. Goldstein and C. Studer, "PhaseMax: Convex phase retrieval via basis pursuit," *arXiv preprint: 1610.07531*, Oct. 2016.

[25] ——, "Convex phase retrieval without lifting via PhaseMax," in *Proc. Intl. Conf. Mach. Learn.*, Aug. 2017, pp. 1273–1281.

[26] O. Dhifallah, C. Thrampoulidis, and Y. M. Lu, "Phase retrieval via linear programming: Fundamental limits and algorithmic improvements," *arXiv preprint: 1710.05234*, Oct. 2017.

[27] O. Dhifallah and Y. M. Lu, "Fundamental limits of PhaseMax for phase retrieval: A replica analysis," *arXiv preprint: 1708.03355*, Aug. 2017.

[28] C. A. Metzler, M. K. Sharma, S. Nagesh, R. G. Baraniuk, O. Cossairt, and A. Veeraraghavan, "Coherent inverse scattering via transmission matrices: Efficient phase retrieval algorithms and a public dataset," in *Proc. IEEE Intl. Conf. Comput. Photograph.*, May 2017, pp. 1–16.

[29] R. Chandra, Z. Zhong, J. Hontz, V. McCulloch, C. Studer, and T. Goldstein, "PhasePack: A phase retrieval library," *arXiv preprint: 1711.10175*, Nov. 2017.

[30] R. Kan and C. Robotti, "On moments of folded and truncated multivariate normal distributions," *J. Comput. Graph Stat.*, pp. 1–5, Apr. 2017.

[31] D. Brillinger, *A generalized linear model with "Gaussian" regressor variables*, ser. Wadsworth Statistiscs/Probability Series. Chapman & Hall/CRC, 1982, pp. 97–114.

[32] C. Stein, "Estimation of the mean of a multivariate normal distribution," *Ann. Stat.*, vol. 9, no. 6, pp. 1135–1151, Nov. 1981.

[33] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Trans. Sig. Process.*, vol. 43, no. 8, pp. 2030–2033, Aug. 1995.

[34] T. Adali, P. J. Schreier, and L. L. Scharf, "Complex-valued signal processing: The proper way to deal with impropriety," *IEEE Trans. Sig. Process.*, vol. 59, no. 11, pp. 5101–5125, Nov. 2011.