The Benefits of Side Information for Structured Phase Retrieval

M. Salman Asif
ECE Department, University of California
Riverside, CA, USA
sasif@ece.ucr.edu

Chinmay Hegde
Tandon School of Engineering, New York University
Brooklyn, NY, USA
chinmay.h@nyu.edu

***Abstract*—Phase retrieval, or signal recovery from magnitude-only measurements, is a challenging signal processing problem. Recent progress has revealed that measurement- and computational-complexity challenges can be alleviated if the underlying signal belongs to certain low-dimensional model families, including sparsity, low-rank, or neural generative models. However, the remaining bottleneck in most of these approaches is the requirement of a carefully chosen initial signal estimate. In this paper, we assume that a portion of the signal is already known a priori as "side information" (this assumption is natural in applications such as holographic coherent diffraction imaging). When such side information is available, we show that a much simpler initialization can provably succeed with considerably reduced costs. We supplement our theory with a range of simulation results.**

I. Introduction

A. Overview

The problem of *phase retrieval* refers to the challenge of recovering a real- or complex-valued signal from its magnitude-only measurements. This problem owes its roots to applications in diffraction imaging, X-ray crystallography, and ptychography [1]–[3]. Let us assume that an unknown signal, $\mathbf{x}^* \in \mathbb{R}^n$, is measured via (possibly noisy) measurements of the form:

$$y_i = |\langle \mathbf{a}_i, \mathbf{x}^* \rangle| + e_i, \tag{1}$$

The goal is to recover an estimate of $\mathbf{x}^*$ from the (phaseless) measurements $\mathbf{y} \in \mathbb{R}^m$. While classical, the phase retrieval problem has attracted renewed interest in the signal processing community due to several recent breakthroughs [4], [5]. In particular, a variety of phaseless signal recovery algorithms have emerged in the literature that enjoy provable guarantees.

The majority of these results have focused on the "over-complete" case where the number of measurements exceeds the signal dimension, sometimes by a significant amount. This may cause compute and storage problems in real-world applications. To alleviate measurement- and related computational-complexity challenges, subsequent works have considered the case where the unknown signal $\mathbf{x}^*$ is *structured* in the sense that it belongs to a low-dimensional sub-manifold of the signal space that is known *a priori*. Examples of signal structure include sparsity [6], [7], low-rank [8], [9], or neural generative models [10], [11]. Such structural assumptions can provably reduce the measurement complexity of the phase retrieval problem to $m = o(n)$ using specialized algorithms; for example, for s-sparse signals whose coefficients obey a power-law behavior, the measurement-complexity drops to $O(s \log n)$ [12].

However, there is a price to be paid in achieving provable signal recovery. Many of these newer algorithms involve solving a fundamentally *non-convex* problem using iterative gradient-style methods; therefore, they need to be carefully initialized. The initialization technique of choice is spectral initialization, first proposed in the context of phase retrieval in [5], and extended to the sparse signal case in [6], [7]. However, the running time of spectral initialization is *quadratic* in the signal dimension ($\Omega(n^2)$) and constitutes the main algorithmic bottleneck.

B. Our contributions

In this paper, we propose, analyze, and experimentally validate a signal recovery algorithm for a variant of the phase retrieval problem. In particular, we will assume that a portion of the signal (or measurements) is already known as "side information". This assumption is natural in optics applications such as holographic coherent diffraction imaging. Such situations can be mathematically modeled as follows:

$$y_i = |\langle \mathbf{a}_i, \mathbf{x}^* \rangle + b_i| + e_i, \tag{2}$$

where $\mathbf{b} = [b_1, \ldots, b_m]$ is assumed to be "side information" that is available to the reconstruction algorithm.

Our recovery algorithm proceeds in two stages. We first construct a coarse initial estimate of the signal; crucially, this is done via a simple, *linear* estimation that leverages the side information. Then, we iteratively refine this initial estimate using an alternating minimization method in which we alternately estimate: (i) the phases (signs) of the observations and (ii) the signal $\mathbf{x}^*$ using (conjugate/projected) gradient descent.

We analyze the performance of our proposed algorithm. Our main result is given in Theorem IV.1. We also analyze the empirical performance of our algorithm using simulated measurements.

This work was partially supported by NSF grants CCF-2005804 and CCF-1815101, and the DARPA REVEAL program.

Algorithm 1 Linear initialization

Input: $\mathbf{A}, \mathbf{y},$ side information $\mathbf{b},$ signal sub-manifold S

1: Compute the linear estimate:

$$\bar{\mathbf{x}} \leftarrow \frac{1}{2m} \sum_{i=1}^{m} y_i^2 b_i \mathbf{a}_i \tag{4}$$

2: (Optional) Project onto feasible sub-manifold:

$$\hat{\mathbf{x}} \leftarrow \underset{\mathbf{x} \in S}{\operatorname{argmin}} \ \|\mathbf{x} - \bar{\mathbf{x}}\|_2^2.$$

Output: $\mathbf{x}^0 \leftarrow \hat{\mathbf{x}}$.

II. Background

The phase retrieval problem has been extensively studied over the last few decades [4], [13], [14] and it appears in several applications, including optical imaging [14], [15], microscopy [16], [17], and X-ray crystallography [18].

Phase retrieval is a non-convex problem and classical solution methods rely on alternating projection heuristics; examples include Gerchberg-Saxton [13] and Fienup algorithms [14]. In recent years, lifting-based methods were introduced that reformulate phase retrieval as a semidefinite program. [4]. Subsequently, non-convex methods have been proposed for solving phase retrieval problem with theoretical performance guarantees [5], [12], [19]–[23]. Most non-convex methods rely on estimating a good initial solution via the so-called *spectral initialization* method. In its simplest form, spectral initialization requires computation of the top singular vector of the following Hermitian matrix:

$$\frac{1}{m} \sum_{i=1}^{m} y_i^2 \mathbf{a}_i \mathbf{a}_i^T, \tag{3}$$

In our proposed algorithm, we circumvent this requirement of spectral computations, and show that when a portion of the signal is known, then a simpler, linear initialization gives us the necessary conditions required to establish convergence.

III. Phase retrieval with side information

Our goal is to reconstruct the underlying signal $\mathbf{x}^*$ from measurements of the form 2. We will assume that $\mathbf{x}^*$ lies on a signal submanifold $S \subseteq \mathbb{R}^n$; this represents the prior structural information about the signal available to the reconstruction algorithm[1]. If there is no such information available we can simply set $S = \mathbb{R}^n$.

Our overall reconstruction algorithm is described in pseudocode form in Algorithms 1 and 2. Our algorithm follows the same familiar two-stage approach that is common in the phase retrieval literature.

(a) *Initialization*: We first construct a coarse initial estimate of the signal. Specifically, we obtain a signal estimate $\mathbf{x}^0$ using a *linear* initialization process described in Algorithm 1. Such an estimate is expected to satisfy the following condition:

$$\operatorname{dist}\left(\mathbf{x}^0, \mathbf{x}^*\right) \leq \delta \left\|\mathbf{x}^*\right\|_2$$

for some small constant δ, where $\operatorname{dist}(\cdot,\cdot)$ is a suitably defined distance measure. In the following section we will derive conditions under which the above condition is provably met.

(b) *Iterative refinement*: We then iteratively refine this coarse initial estimate using alternating minimization following [5], [12], [23]. The refinement procedure is described in Algorithm 2. In each iteration, the (unknown) phases of the observations are hallucinated using the current signal estimate, and the corresponding linear problem is solved using (constrained) least squares. In the following section, we will derive conditions under which our refinement strategy demonstrates linear convergence to $\mathbf{x}^*$ when no noise is present[2]. Specifically, for $t = 0, 1, 2, \ldots$, our sequence of estimates satisfies:

$$\operatorname{dist}\left(\mathbf{x}^{t+1}, \mathbf{x}^*\right) \leq \rho \operatorname{dist}\left(\mathbf{x}^t, \mathbf{x}^*\right).$$

Algorithm 2 Iterative refinement

Input: $\mathbf{A}, \mathbf{y},$ side information $\mathbf{b},$ signal sub-manifold S

1: Initialize $\mathbf{x}^0$ according to Algorithm 1
2: **for** $t = 1, \cdots, T$ **do**
3: $\quad \mathbf{p}^t \leftarrow \operatorname{sign}\left(\mathbf{A}\mathbf{x}^{t-1} + \mathbf{b}\right),$
4: $\quad \mathbf{x}^t = \operatorname{argmin}_{\mathbf{x} \in S} \left\|\mathbf{p}^t \circ \mathbf{y} - \mathbf{b} - \mathbf{A}\mathbf{x}\right\|_2^2$
5: **end for**

Output: $\hat{\mathbf{x}} \leftarrow \mathbf{x}^T$.

We now estimate the computational complexity of the overall reconstruction algorithm. It can be seen that Line 1 in Alg. 1 and Line 3 in Alg. 2 involve a single matrix-vector multiply involving A, and hence can be achieved in time $O(mn)$. The main burden now becomes the (constrained) least squares step (Line 4 in Alg. 2), but this can be solved in principle using approximate techniques. For example, in the unconstrained case, one can use conjugate gradient for a fixed number of steps, while in the case of s-sparsity, one can use iterative algorithms such as CoSaMP [24] or iterative hard thresholding [25]. In either case, this step can be achieved in time $\tilde{O}(mn)$. Therefore, the overall running time of the reconstruction procedure is $O(mnT)$. When $m = o(n)$ and linear convergence is achieved, our proposed algorithm runs *sub-quadratic* in the signal dimension n.

IV. Analysis

We now provide sufficient conditions under which our proposed reconstruction algorithm provably converges to the

[1]For example, if the signal is known to be s-sparse then S represents the union of all s-dimensional canonical subspaces of $\mathbb{R}^n$.

[2]When noise is present, we achieve linear convergence up to a ball around $\mathbf{x}^*$ whose radius depends on the noise level

desired solution. We will use the Euclidean norm to define the distance measure for convergence:

$$\text{dist}\left(\mathbf{x}, \mathbf{x}'\right) = \left\|\mathbf{x} - \mathbf{x}'\right\|_2.$$

A. Assumptions

While our algorithm is applicable in general settings, for our theoretical analysis we make the following assumptions:

(a) The elements of $\mathbf{A}$ are chosen i.i.d. from a unit Gaussian.
(b) The elements of $\mathbf{b}$ are also chosen i.i.d. from any distribution of mean 0, variance 1 (for example, a unit Gaussian).
(c) (optional) There exists an ε-approximation algorithm for solving the manifold-constrained least squares step.
(d) The measurements are noiseless, i.e., $e_i = 0$.

The first assumption is standard in the (generalized) phase retrieval literature. The second assumption is new and particular to our "side-information" setting, and we expect that it can be relaxed even further. The third assumption is reasonable since there has been a lot of recent effort towards deriving efficient approximation algorithms for solving constrained least squares [26], [27]. The fourth assumption is made purely for simplicity, and analogous theoretical results can be derived even in the presence of noise.

B. Linear initialization

The first stage of the reconstruction algorithm constructs a coarse initial signal estimate. To do so, we define a *unbiased* estimator of the underlying signal using the measurements plus side information as follows:

$$\bar{x} = \frac{1}{2m}\sum_{i=1}^{m} y_i^2 b_i \mathbf{a}_i.$$

To provide the intuition behind the above expression, observe the following:

$$\begin{aligned} y_i^2 &= (\langle \mathbf{a}_i, \mathbf{x}^* \rangle + b_i)^2, \\ &= \mathbf{a}_i \mathbf{x}^* \mathbf{x}^{*T} \mathbf{a}_i^T + b_i^2 + 2 b_i \mathbf{a}_i^T \mathbf{x}^*, \quad \text{i.e.,} \\ \mathbb{E}[y_i^2 b_i \mathbf{a}_i] &= \mathbb{E}\left[b_i (\mathbf{a}_i \mathbf{x}^* \mathbf{x}^{*T} \mathbf{a}_i^T + b_i^2 + 2 b_i \mathbf{a}_i^T \mathbf{x}^*) \mathbf{a}_i \right], \\ &= \mathbb{E}[b_i]\mathbb{E}[\mathbf{a}_i \mathbf{x}^* \mathbf{x}^{*T} \mathbf{a}_i^T \mathbf{a}_i] \\ &\quad + 2\mathbb{E}[b_i^2]\mathbb{E}[\mathbf{a}_i^T \mathbf{x}^* \mathbf{a}_i] + \mathbb{E}[b_i^3]\mathbb{E}[\mathbf{a}_i], \\ &= 2\mathbf{x}^*, \end{aligned}$$

which follows from Assumptions (a) and (b) above. If a constraint sub-manifold S has been given beforehand, then the signal estimate $\bar{x}$ is orthogonally projected onto S to produce the initial estimate:

$$\mathbf{x^0} = \underset{\mathbf{x} \in S}{\text{argmin}} \left\|\bar{\mathbf{x}} - \mathbf{x}\right\|_2^2.$$

By applying tail concentration bounds for sub-Gaussian random variables, we can infer with high probability that for m/n exceeding a certain constant, we get:

$$\text{dist}\left(\mathbf{x^0}, \mathbf{x}^*\right) \leq \delta_0 \left\|\mathbf{x}^*\right\|_2,$$

with high probability. We omit detailed proofs here due to space constraints, but refer to [12], [28] for details for both the unstructured and the structured cases.

C. Iterative refinement

The second stage of the reconstruction algorithm refines the initial signal estimate to progressively become closer and closer to the solution. The method involves alternately estimating the (unknown) phases of the (linearized) measurements $\langle \mathbf{a}_i, \mathbf{x}^* \rangle + b_i$, and using these linearized measurements to reconstruct the signal estimate using standard regression or sparse recovery techniques.

The intuition is as follows. In the absence of noise (Assumption (d) above), the observation model in (2) reduces to:

$$\text{sign}\left(\langle \mathbf{a}_i, \mathbf{x}^* \rangle + b_i\right) \circ y_i = \langle \mathbf{a}_i, \mathbf{x}^* \rangle + b_i,$$

for all $i = \{1, 2, \ldots, m\}$. To ease notation, denote the *phase vector* $\mathbf{p} \in \mathbb{R}^m$ as a vector that contains the unknown signs of the measurements, i.e., $p_i = \text{sign}\left(\langle \mathbf{a}_i, \mathbf{x} \rangle + b_i\right)$ for $i = \{1, 2, \ldots, m\}$. Let $\mathbf{p}^*$ denote the true phase vector and let $\mathcal{P}$ denote the set of all phase vectors, i.e. $\mathcal{P} = \{\mathbf{p} : p_i = \pm 1, \forall i\}$. Therefore, the recovery of $\mathbf{x}^*$ can be posed as a (non-convex) optimization problem:

$$\min_{\mathbf{x} \in S, \mathbf{p} \in \mathcal{P}} \left\|\mathbf{A}\mathbf{x} + \mathbf{b} - \mathbf{p} \circ \mathbf{y}\right\|_2 \tag{5}$$

To solve this problem, we alternate between estimating $\mathbf{p}$ and $\mathbf{x}$. We perform two estimation steps:

(a) if we fix the signal estimate $\mathbf{x}$, then the minimizer $\mathbf{p} \in \mathcal{P}$ is given in closed form as:

$$\mathbf{p}_i = \text{sign}\left(\langle \mathbf{a}_i, \mathbf{x} \rangle + b_i\right), \tag{6}$$

(b) and if we fix the phase vector $\mathbf{p}$, the signal vector $\mathbf{x} \in S$ can be obtained via (approximately) solving the constrained least squares problem:

$$\min_{\mathbf{x} \in S} \left\|\mathbf{p} \circ \mathbf{y} - \mathbf{A}\mathbf{x} - \mathbf{b}\right\|_2^2. \tag{7}$$

We now analyze our proposed descent scheme. We obtain:

Theorem IV.1. *Given an initialization $\mathbf{x^0} \in \mathcal{M}$ satisfying $\text{dist}\left(\mathbf{x}^0, \mathbf{x}^*\right) \leq \delta \left\|\mathbf{x}^*\right\|_2$, for $0 < \delta < 1$, if the number of (Gaussian) measurements,*

$$m > \tilde{\Omega}_{\delta,\rho}(\omega^2(S))$$

then with high probability, the iterates $\mathbf{x}^{t+1}$ of Alg. 2, satisfy:

$$\text{dist}\left(\mathbf{x}^{t+1}, \mathbf{x}^*\right) \leq \rho\, \text{dist}\left(\mathbf{x}^t, \mathbf{x}^*\right), \tag{8}$$

where $0 < \rho < 1$ and $\omega(\cdot)$ represents Gaussian width.

Proof: The high level idea behind the proof is that with a δ-ball around the true signal $\mathbf{x}^*$, the "phase noise" can be suitably bounded in terms of a constant times the signal estimation error. To be more precise, suppose that $\mathbf{z}^* = \mathbf{A}\mathbf{x}^* + b_i = \mathbf{p}^* \circ \mathbf{y}$. Then, at any iteration t, we have:

$$\begin{aligned} \mathbf{z}^t &= \mathbf{p}^t \circ \mathbf{y} \\ &= \mathbf{p}^* \circ \mathbf{y} + (\mathbf{p}^t - \mathbf{p}^*) \circ \mathbf{y} \\ &= \mathbf{z}^* + \mathbf{e}^t, \end{aligned}$$

where $\mathbf{e}^t$ can be viewed as the "phase noise". Now, examining Line 4 of Algorithm 2, we have:

$$\left\|\mathbf{A}\mathbf{x}^t + \mathbf{b} - \mathbf{p}^t \circ \mathbf{y}\right\|_2 \leq (1+\varepsilon)\left\|\mathbf{A}\mathbf{x}^* + \mathbf{b} - \mathbf{p}^t \circ \mathbf{y}\right\|_2.$$

This follows from two reasons: (i) in Assumption (c) above we assumed that $\mathbf{x}^t$ is an ε-approximate solution to the constrained least squares problem 7. (ii) $\mathbf{x}^*$ is a feasible point in S. Therefore, we have:

$$\begin{aligned}\left\|\mathbf{A}\mathbf{x}^t + \mathbf{b} - \mathbf{p}^t \circ \mathbf{y}\right\|_2 &\leq (1+\varepsilon)\left\|\mathbf{z}^* - \mathbf{z}^t\right\|_2 \\ &= (1+\varepsilon)\left\|\mathbf{e}^t\right\|_2.\end{aligned}$$

On the other hand, we have:

$$\begin{aligned}\left\|\mathbf{A}\mathbf{x}^t + \mathbf{b} - \mathbf{p}^t \circ \mathbf{y}\right\|_2 &= \left\|\mathbf{A}\mathbf{x}^t + \mathbf{b} - \mathbf{z}^t\right\|_2 \\ &\geq \left\|\mathbf{A}\mathbf{x}^t + \mathbf{b} - \mathbf{z}^*\right\|_2 - \left\|\mathbf{z}^* - \mathbf{z}^t\right\|_2 \\ &= \left\|\mathbf{A}(\mathbf{x}^t - \mathbf{x}^*)\right\|_2 - \left\|\mathbf{e}^t\right\|_2 \\ &\geq (1-\delta_0)\left\|\mathbf{x}^t - \mathbf{x}^*\right\|_2 - \left\|\mathbf{e}^t\right\|_2,\end{aligned}$$

where the second inequality follows from the triangle inequality and the last inequality follows from standard matrix concentration arguments [29]; in general, this holds provided:

$$M \geq \frac{C}{\delta_0^2}(\omega^2(S)),$$

where ω represents the Gaussian width[3]. Rearranging the inequalities gives us the following bound:

$$\left\|\mathbf{x}^t - \mathbf{x}^*\right\|_2 \leq \frac{2(1+\varepsilon)}{1-\delta_0}\left\|\mathbf{e}^t\right\|_2.$$

It remains to show that $\|\mathbf{e}^t\|_2$ can be bounded in terms of $\left\|\mathbf{x}^{t-1} - \mathbf{x}^*\right\|_2$. We do this by invoking Lemma C1 in [12] (whose proof involves chaining techniques and is lengthy, so we omit it here). Consequently, we get

$$\left\|\mathbf{x}^t - \mathbf{x}^*\right\|_2 \leq \frac{2(1+\varepsilon)\rho'}{1-\delta_0}\left\|\mathbf{x}^{t-1} - \mathbf{x}^*\right\|_2,$$

where ρ' is a small enough constant. Setting various constants, we achieve a per-step error reduction scheme of the form:

$$\left\|\mathbf{x}^t - \mathbf{x}^*\right\|_2 \leq \rho_0\left\|\mathbf{x}^{t-1} - \mathbf{x}^*\right\|_2,$$

if the initial estimate $\mathbf{x}^0$ satisfies $\left\|\mathbf{x}^0 - \mathbf{x}^*\right\|_2 \leq \delta_0\left\|\mathbf{x}^*\right\|_2$.

V. Experiments

In this section, we present several simulation results to demonstrate the performance of our proposed algorithm under different scenarios.

[3]See [29] for bounds that link Gaussian width to the complexity of the signal manifold S. For example, for s-sparse vectors, $\omega(S) = \sqrt{s\log n}$.

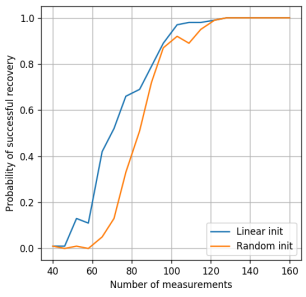
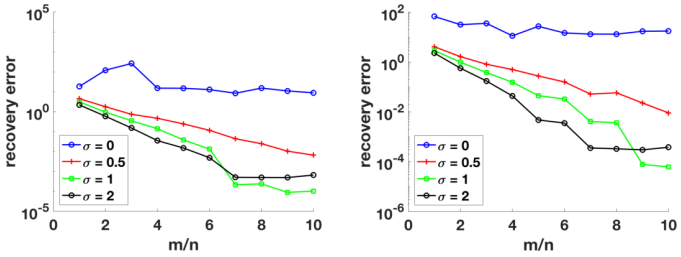


Fig. 1: Phase transition for reconstructing sparse signals using side information ($n = 1000, s = 5$).



(a) Fourier measurements (b) Gaussian measurements

Fig. 2: Reconstruction error for different values of oversampling factor m/n and reference signal strength σ with $n = 1024$.

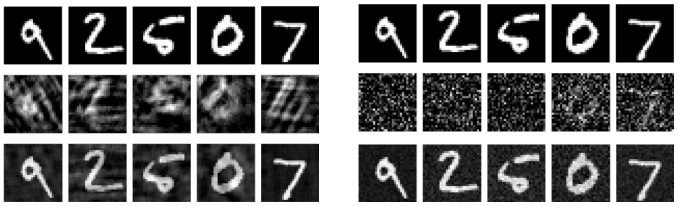
a) 1D signals: We first test our method for the recovery of 1D signals from their (phaseless) measurements with an additive known reference signal $\mathbf{b}$. Our simulation setup is as follows. We generate $\mathbf{x}^* \in \mathbb{R}^n$ with coefficients drawn from $\mathcal{N}(0, 1/m)$ and the known "side information" $\mathbf{b} \in \mathbb{R}^m$ with independent entries drawn from $\mathcal{N}(0, \sigma^2)$. We simulate phaseless measurements $\mathbf{y} = |\mathbf{A}\mathbf{x}^* + \mathbf{b}|$ and estimated $\hat{\mathbf{x}}$ using our proposed algorithm. We performed 100 independent trials. For every trial, we record the normalized recovery error $\frac{\|\mathbf{x}^* - \hat{\mathbf{x}}\|_2}{\|\mathbf{x}^*\|_2}$.

First, we consider recovering s-sparse signals $\mathbf{x}^*$ with $s = 5, n - 1000$ using our proposed algorithm. Figure 1 shows the phase transition curve (in blue) plotting successful recovery as a function of m (where success is defined as signal recovery with error less than 1e-4.) Interestingly, a *randomly* initialized method succeeds only a little worse; however, randomly initialized alternating minimization without the side information was unable to reconstruct the signal for the given range of m.

Figure 2 summarizes error curves for our simulations using our refinement algorithm with both Fourier and Gaussian measurements with $n = 1024$, $\sigma = 0, 0.5, 1, 2$, and a range of m/n. Every point on the error curves represents the recovery error for the given values of m/n and σ. We observe that for nonzero values of σ, the recovery error reduces as m/n increases. We also observe that the quality of reconstruction depends on the strength of the reference signal $\mathbf{b}$ that we

denote as σ. Note that $\sigma = 0$ refers to the standard phase retrieval problem when $\mathbf{b} = 0$; this problem is especially challenging for the Fourier measurements. We can observe in Fig. 2(a) that the phase retrieval method fails to recover signal from Fourier amplitude measurements when $\sigma = 0$ but as σ increases, the recovery performance improves. In the case of Gaussian measurements, the signal is recovered almost perfectly if $m/n \geq 6$ with or without reference. We also observe that the presence of a strong known reference in the measurements allows signal recovery even if $m/n \approx 3$.

b) 2D images: In the second set of experiments, we test our method for the recovery of sample MNIST images from their Gaussian and Fourier (phaseless) measurements with an additive known reference signal $\mathbf{b}$. We select five test images from MNIST dataset, each of size 28×28 (i.e., $n = 784$). We captured $m = 1764$ phaseless measurements as $|\mathbf{A}\mathbf{x}^* + \mathbf{b}|$, where $\mathbf{A}$ either denotes an $m \times n$ Gaussian matrix or a 2D Fourier transform operator with appropriate zero padding to achieve $4\times$ oversampling, $\mathbf{x}^*$ denotes unknown image, and $\mathbf{b}$ denotes the reference. We generated the reference signal $\mathbf{b} \in \mathbb{R}^m$ with independent entries drawn from uniform distribution in interval $[0, 1]$.



(a) Fourier measurements (b) Gaussian measurements

Fig. 3: Sample reconstruction of MNIST images from Fourier and Gaussian phaseless measurements.

The results are summarized in Fig. 3. The original 28×28 images are shown in the top row. The middle row shows reconstruction from the regular phase retrieval method (i.e., $\mathbf{b} = 0$), which fails to recover any image accurately. The bottom row shows the reconstruction using our proposed method when a known reference signal is added to the real- or complex-valued measurements before recording their amplitude. The results clearly illustrate the benefits of the side information available to the reconstruction algorithm.

REFERENCES

[1] R. Harrison, "Phase problem in crystallography," *JOSA a*, vol. 10, no. 5, pp. 1046–1055, 1993.

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

[3] Y. Shechtman, Y. Eldar, O. Cohen, H. Chapman, J. Miao, and M. Segev, "Phase retrieval with application to optical imaging: a contemporary overview," *IEEE Sig. Proc. Mag.*, vol. 32, no. 3, pp. 87–109, 2015.

[4] 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*, vol. 66, no. 8, pp. 1241–1274, 2013.

[6] G. Wang, L. Zhang, G. Giannakis, M. Akçakaya, and J. Chen, "Sparse phase retrieval via truncated amplitude flow," *IEEE Transactions on Signal Processing*, vol. 66, no. 2, pp. 479–491, 2018.

[5] P. Netrapalli, P. Jain, and S. Sanghavi, "Phase retrieval using alternating minimization," in *Adv. Neural Inf. Proc. Sys. (NIPS)*, 2013, pp. 2796–2804.

[7] Gauri Jagatap and Chinmay Hegde, "Fast, sample-efficient algorithms for structured phase retrieval," in *Advances in Neural Information Processing Systems*, 2017, pp. 4924–4934.

[8] G. Jagatap, Z. Chen, S. Nayer, C. Hegde, and N. Vaswani, "Sample efficient fourier ptychography for structured data," *IEEE Trans. Computational Imaging*, Oct. 2019.

[9] Z. Chen, G. Jagatap, S. Nayer, C. Hegde, and N. Vaswani, "Low rank fourier ptychography," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, Apr. 2018.

[10] R. Hyder, V. Shah, C. Hegde, and S. Asif, "Alternating phase projected gradient descent with generative priors for solving compressive phase retrieval," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, May 2019.

[11] G. Jagatap and C. Hegde, "Algorithmic guarantees for inverse imaging with untrained network priors," in *Adv. Neural Inf. Proc. Sys. (NIPS)*, Dec. 2019.

[12] G. Jagatap and C. Hegde, "Sample-efficient algorithms for recovering structured signals from magnitude-only measurements," *IEEE Trans. Inform. Theory*, vol. 65, no. 7, pp. 4435–4456, July 2019.

[13] R.W. Gerchberg and A Saxton W. O., "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik*, vol. 35, pp. 237–250, 11 1971.

[14] James R Fienup, "Phase retrieval algorithms: a comparison," *Applied optics*, vol. 21, no. 15, pp. 2758–2769, 1982.

[15] Jason Holloway, M. Salman Asif, Manoj Kumar Sharma, Nathan Matsuda, Roarke Horstmeyer, Oliver Cossairt, and Ashok Veeraraghavan, "Toward long-distance subdiffraction imaging using coherent camera arrays," *IEEE Transactions on Computational Imaging*, vol. 2, no. 3, pp. 251–265, 2016.

[16] Lei Tian, Xiao Li, Kannan Ramchandran, and Laura Waller, "Multiplexed coded illumination for fourier ptychography with an led array microscope," *Biomedical optics express*, vol. 5, no. 7, pp. 2376–2389, 2014.

[17] JM Rodenburg, "Ptychography and related diffractive imaging methods," *Advances in Imaging and Electron Physics*, vol. 150, pp. 87–184, 2008.

[18] Jianwei Miao, Pambos Charalambous, Janos Kirz, and David Sayre, "Extending the methodology of x-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens," *Nature*, vol. 400, no. 6742, pp. 342–344, 1999.

[19] S. Bahmani and J. Romberg, "Efficient compressive phase retrieval with constrained sensing vectors," in *Adv. Neural Inf. Proc. Sys. (NIPS)*, 2015, pp. 523–531.

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

[21] G. Wang, G. Giannakis, Y. Saad, and J. Chen, "Solving most systems of random quadratic equations," in *Adv. Neural Inf. Proc. Sys. (NIPS)*, 2017.

[22] Emmanuel J Candes, Xiaodong Li, and Mahdi Soltanolkotabi, "Phase retrieval from coded diffraction patterns," *Applied and Computational Harmonic Analysis*, vol. 39, no. 2, pp. 277–299, 2015.

[23] G. Jagatap and C. Hegde, "Fast, sample efficient algorithms for structured phase retrieval," in *Adv. Neural Inf. Proc. Sys. (NIPS)*, 2017, pp. 4924–4934.

[24] D. Needell and J. Tropp, "Cosamp: Iterative signal recovery from incomplete and inaccurate samples," *Ap. Comp. Har. An.*, vol. 26, no. 3, pp. 301–321, 2009.

[25] Thomas Blumensath and Mike E Davies, "Iterative hard thresholding for compressed sensing," *Applied and computational harmonic analysis*, vol. 27, no. 3, pp. 265–274, 2009.

[26] R. Baraniuk, V. Cevher, M. Duarte, and C. Hegde, "Model-based compressive sensing," *IEEE Trans. Inform. Theory*, vol. 56, no. 4, pp. 1982–2001, Apr. 2010.

[27] C. Hegde, P. Indyk, and L. Schmidt, "A nearly linear-time framework for graph-structured sparsity," in *Proc. Int. Conf. Machine Learning (ICML)*, July 2015.

[28] H. Zhang and Y. Liang, "Reshaped wirtinger flow for solving quadratic system of equations," in *Adv. Neural Inf. Proc. Sys. (NIPS)*, 2016, pp. 2622–2630.

[29] K. Jaganathan, S. Oymak, and B. Hassibi, "Sparse phase retrieval: Convex algorithms and limitations," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*. IEEE, 2013, pp. 1022–1026.