Title
One-dimensional phase retrieval: regularization, box relaxation and uniqueness

Journal
Inverse Problems, 36(9)

ISSN
0266-5611

Authors
Wong, Hong
Lou, Yifei
Marchesini, Stefano
et al.

Publication Date
2020-09-01

DOI
10.1088/1361-6420/aba2bc

Peer reviewed

One-Dimensional Phase Retrieval: Regularization, Box Relaxation and Uniqueness

Wing Hong Wong[1], **Yifei Lou**[2], **Stefano Marchesini**[3], **Tieyong Zeng**[1]

[1]Department of Mathematics, the Chinese University of Hong Kong, Hong Kong
[2]Mathematical Sciences Department, University of Texas Dallas, Richardson, TX 75080, USA
[3]Computational Research Division, Lawrence Berkeley National Laboratory, USA

E-mail: `whwong@math.cuhk.edu.hk`, `yifei.lou@utdallas.edu`, `smarchesini@lbl.gov`, `zeng@math.cuhk.edu.hk`

Abstract. Recovering a signal from its Fourier magnitude is referred to as phase retrieval, which occurs in different fields of engineering and applied physics. This paper gives a new characterization of the phase retrieval problem. Particularly useful is the analysis revealing that the common gradient-based regularization does not restrict the set of solutions to a smaller set. Specifically focusing on binary signals, we show that a box relaxation is equivalent to the binary constraint for Fourier-types of phase retrieval. We further prove that binary signals can be recovered uniquely up to trivial ambiguities under certain conditions. Finally, we use the characterization theorem to develop an efficient denoising algorithm.

PACS numbers: 00.00, 20.00, 42.10

Keywords: Phase Retrieval, binary signals, box relaxation, ambiguities.

Submitted to: *Inverse Problems*

1. Introduction

In many fields of physics and engineering, one can only measure the magnitude of the Fourier Transform of a discrete signal $\boldsymbol{x} \in \mathbb{C}^N$. Denote the discrete Fourier Transform by $\mathcal{F}$. Recovering $\boldsymbol{x}$ from $|\mathcal{F}\boldsymbol{x}|$ is referred to as *phase retrieval* (PR), since the phase is completely lost in measurements. Phase retrieval originated from X-ray crystallography [1, 2], trying to determine atomic and molecular structures of a crystal. This approach was later used to reconstruct an image of a sample with resolution at a nano-meter scale from its X-ray diffraction pattern, known as coherent diffraction imaging (CDI) [3]. The PR techniques now occur in various applications such as astronomy [4] and laser optics [5]; please refer to [6] for a contemporary overview.

Phase retrieval is a very challenging problem largely due to its nonconvexity and solutions being non-unique [7]. Specifically for the nonuniqueness (a.k.a., *ambiguities*), there are trivial ambiguities and non-trivial ambiguities [6]. Trivial ambiguities of $|\mathcal{F}\boldsymbol{y}| = |\mathcal{F}\boldsymbol{x}|$ can be summarized as follows,

$$\begin{aligned} &\text{global phase shift: } y_k = x_k \cdot e^{i\phi_0} \\ &\text{conjugate inverse: } y_k = \overline{x_{-k}} \\ &\text{spatial shift: } y_k = x_{k+k_0}, \end{aligned} \tag{1.1}$$

where the indices are taken cyclically up to N, $\bar{\cdot}$ denotes the complex conjugate, and $\phi_0 \in [0, 2\pi)$, $k_0 \in \mathbb{Z}$ are the phase shift and spatial shift, respectively. Note that every combination of (1.1) is also a trivial ambiguity. Non-trivial ambiguities of one-dimensional signals can be classified by the roots of the Z-transform of the *autocorrelation* of the signal [8], while almost all multi-dimensional signals only have non-trivial ambiguities [9], since the Z-transform of their autocorrelation being reducible is of measure zero in the space of all polynomials [8, 10].

For unique recovery of a real signal of size N in up to trivial ambiguities, at least $2N - 1$ random measurements are needed, provided the sampling matrix has full spark [11]. This result was later extended to the complex case in [12, 13], requiring at least $4N - 4$ measurements. Other sufficient conditions for unique recovery include minimum phase signals [14], sparse signals with non-periodic support [15], and signals with collision-free [16]. For s-sparse signals in $\mathbb{R}^N$, the number of Fourier magnitude measurements is in the order of $O(s\log(N/s))$ [17, 18], while $\min\{2s, 2n-1\}$ for random measurements [19].

In addition to taking more measurements than the ambient dimension, one often relies on regularization to refine the solution space with an attempt to reduce ambiguities. Stemming from image processing, a common choice is a gradient-type formalism. For example, Chang et al. [20] considered the total variation, which is the ℓ_1 norm of the gradient for phase retrieval. Computationally, many optimization algorithms can be used to solve the (regularized) phase retrieval problems, including alternating projections [21], Wirtinger flow [22], alternating direction method of multipliers (ADMM) [20], and a preconditioned proximal algorithm [23].

This paper contributes to a new set of characterization theorems for phase retrieval, indicating that gradient-based regularization is redundant to the magnitude measurements. We also impose additional constraints on the underlying signal in order to resolve the ambiguities. In particular, we focus on binary signals [24] due to its simplicity and a wide variety of applications such as bar code [25, 26] and obstacle detection [27]. Specifically for phase retrieval, binary signals are considered in magnetism to describe the x-ray energies of some chemical compound films such as the $SmCO_5$ film [28], and in block copolymers to describe films [29]. It was observed empirically in [30] that incorporating a box constraint into the ADMM framework, referred to ADMMB, often gives an exact recovery of binary signal, which motivates us to give a theoretical explanation. In this paper, we prove that the phase retrieval problem with binary constraint is equivalent to phase retrieval with box relaxation. We describe a new type of *trivial ambiguities* for binary phase retrieval and show that unique recovery is possible under certain conditions. A related work [31] proved binary signals that cannot be uniquely recovered by Fourier magnitude is a zero-measure set. Finally, we take the noise into consideration and develop a denoising algorithm.

Our contributions are three-fold: (1) We give a characterization theorem (Theorem 3.5), revealing the fact that $\|\nabla^n \boldsymbol{x}\|_2$ is completely determined by $|\mathcal{F}\boldsymbol{x}|$ for an arbitrary integer order n. (2) We give thorough analysis of phase retrieval problem in a binary setting. We show that the box relaxation to binary constraint is equivalent to the original binary phase retrieval problem (Theorem 4.1). We then describe a new type of ambiguities and guarantee the uniqueness of binary phase under certain conditions. (3) We conduct a series of error analysis (Propositions 5.1–5.2 and Corollary 5.3) of phase retrieval, which motivates a new denoising scheme.

The rest of the paper is organized as follows. In Section 2, we set up notations and review some practical ways of taking magnitude measurements. In Section 3, we give a new characterization theorem and discuss its consequences. In Section 4, we prove that the phase retrieval of binary signals can be relaxed to the box constraint. Furthermore, we show it is possible to relax the set of vectors having the same norm to its convex hull. In Section 4.1, we describe a new type of ambiguities for binary signals and show that the unique recovery of binary signals is possible under some special circumstances. Several extensions from the Fourier case to other types of sampling schemes are presented in Section 4.2. In Section 5, we estimate recover accuracy with respect to noise and propose a denoising algorithm that empirically yields better performance compared to a naïve approach. Section 6 concludes the paper. Appendix provides all the proofs for the theorems presented.

2. Preliminaries

2.1. Notations

Let $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^N$ be arbitrary signals, we define some notations that are used throughout the paper,

- x_k denotes the k-th entry of $\boldsymbol{x}$, i.e. $\boldsymbol{x} = (x_0, x_1, x_2, \ldots, x_{N-1})^T$
- $\|\boldsymbol{x}\|_p$ denotes the ℓ_p-norm of $\boldsymbol{x}$, i.e. $\|\boldsymbol{x}\|_p = (\sum_{k=0}^{N-1} |x_k|^p)^{\frac{1}{p}}$, where $p > 0$. For $p = 0$, we define $\|\boldsymbol{x}\|_0$ to be the ℓ_0 "norm" by counting the number of its nonzero elements.
- $\boldsymbol{e}_k$'s denotes the standard basis in $\mathbb{C}^N$, i.e. the vector with a 1 in the k-th coordinate and 0's elsewhere, e.g., $\boldsymbol{e}_0 = (1, 0, 0, \ldots, 0)^T$ and $\boldsymbol{e}_1 = (0, 1, 0, \ldots, 0)^T$.
- $\mathcal{F}_{N \to M} : \mathbb{C}^N \to \mathbb{C}^M$ denotes the matrix representing discrete Fourier transform

(DFT), i.e.

$$\mathcal{F}_{N\to M} = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \omega & \omega^2 & \dots & \omega^{N-1} \\ 1 & \omega^2 & \omega^4 & \dots & \omega^{2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{M-1} & \omega^{2(M-1)} & \dots & \omega^{(M-1)(N-1)} \end{bmatrix}, \tag{2.1}$$

where $\omega = e^{\frac{-2\pi i}{M}}$. Note that $\frac{1}{\sqrt{N}}\mathcal{F}_{N\to N}$ is unitary. If $M > N$, we refer it as an *oversampling* Fourier matrix.

- We define

$$\boldsymbol{x} \odot \boldsymbol{y} = (x_0y_0, x_1y_1, \dots, x_{N-1}y_{N-1}),$$

where $\odot$ denotes the Hadamard product (i.e. entrywise multiplication).

- The *discrete (periodic) convolution* $\boldsymbol{x} * \boldsymbol{y}$ is defined by

$$(\boldsymbol{x} * \boldsymbol{y})_j = \sum_{k=0}^{N-1} x_k y_{(j-k)\mathrm{mod}N}, \tag{2.2}$$

for $j = 0, 1, \dots, N-1$.

- The *(regular) autocorrelation* is defined by

$$(\mathrm{Aut}(\boldsymbol{x}))_j = \sum_{k=0}^{N-1} x_{(k+j)}\overline{x_k}, \tag{2.3}$$

where $j = -N+1, -N+2, \dots, N-1$ and $x_k = 0, \forall k < 0$ and $k > N-1$.

- By replacing the zero boundary condition in the regular autocorrelation with periodic boundary condition, we consider *periodic autocorrelation* defined as

$$(\mathrm{Aut}_p(\boldsymbol{x}))_j = \sum_{k=0}^{N-1} x_{(k+j)\mathrm{mod}N}\overline{x_k}, \tag{2.4}$$

for $j = 0, 1, \cdots, N-1$. These definitions will be used in the proofs of some interesting results.

For the rest of the paper, we denote $\mathcal{F}_{N\to N}$ by $\mathcal{F}$, $\mathcal{F}_{N\to M}$ by $\mathcal{F}_M$, and omit ${}_{\mathrm{mod}N}$ if the context is clear.

2.2. Sampling Schemes

In practice, there are numerous ways [32, 33, 34, 35, 36, 37] to take magnitude measurements of a signal. This paper develops new theoretical characterizations in PR focusing on the following sampling schemes.

- **Classic Fourier Transform.** One aims to find an unknown signal $\boldsymbol{x} \in \mathbb{C}^N$ from the magnitude measurements $\boldsymbol{b} := |\mathcal{F}\boldsymbol{x}|$, i.e.

$$b_n = \Big|\sum_{k=0}^{N-1} x_k e^{\frac{-2\pi kni}{N}}\Big|, \quad \forall n = 0, 1, \ldots, N-1.$$

- **Oversampling Fourier Transform.** An M-point ($M > N$) oversampling discrete Fourier Transform (DFT) of a signal $\boldsymbol{x} \in \mathbb{C}^N$ is defined by

$$b_n = \Big|\sum_{k=0}^{N-1} x_k e^{\frac{-2\pi kni}{M}}\Big|, \quad \forall n = 0, 1, \ldots, M-1.$$

  One wants to recover an N-point signal $\boldsymbol{x}$ based on M measurements of $|\mathcal{F}_M\boldsymbol{x}|$. A typical choice of M is $M = 2N$ [38], which is experimentally adopted by Miao et al [39]. Also, a sufficient number of measurements is crucial in avoiding false solutions [40]. However, we show in Theorem 3.6 theoretically that more measurements (i.e. $M \geq N$) do not resolve ambiguities in the noiseless PR problem.

- **Short-Time Fourier Transform** (STFT) [34, 41]. Let $\boldsymbol{x} \in \mathbb{C}^N$ be a signal of length N and $\boldsymbol{w} \in \mathbb{C}^W$ be a window function of length W. The Short-Time Fourier Transform (STFT) of $\boldsymbol{x}$ with respect to $\boldsymbol{w}$ is defined as

$$\boldsymbol{z}_{n,m} = \sum_{k=0}^{N-1} x_k w_{mL-k} e^{\frac{-2\pi kni}{N}}, \tag{2.5}$$

  for $n = 0, 1, \ldots, N-1$ and $m = 0, 1, \ldots, R-1$, where L denotes the separation in time between adjacent short-times sections, R $= \lceil \frac{N+W-1}{L} \rceil$ denotes the number of short-time sections considered, and $w_k := 0$ for all $k < 0$ and $k > W-1$.

- **Frequency-resolved optical gating trace** (FROG) [32, 33, 37]. Let

$$z_{n,m} = x_n x_{n+mL},$$

  where L is a fixed integer. The FROG trace is equivalent to the one-dimensional Fourier magnitude of $z_{n,m}$ for each fixed m, i.e.

$$|\hat{z}_{n,m}|^2 = |\sum_{k=0}^{N-1} x_k x_{k+mL} e^{\frac{-2\pi kni}{N}}|^2, \tag{2.6}$$

  for $n = 0, \ldots, N-1, m = 0, \ldots, \lceil \frac{N}{L} \rceil - 1$.

Both STFT and FROG make experimentally plausible means of additional phaseless measurements to improve the accuracy of phase retrieval. For example, the STFT measurements can be obtained by a set of shifted versions of a single mask, while FROG measures the product of the signal with a shifted version of itself. It was claimed in [41] that the STFT magnitude leads to better performance than an oversampled DFT with the same number of measurements.

3. Regularization and Constraint in Phase Retrieval

Mathematically, the Fourier-type of phase retrieval problems in one dimensional case is formulated as follows,

$$\text{Find } \boldsymbol{x} \in \mathbb{C}^N, \text{ s.t. } |\mathcal{F}_M \boldsymbol{x}| = \boldsymbol{b}.$$

It is desirable and often necessary to impose some regularization term in order to regularize the solution and avoid ambiguities in PR as much as possible. A classic choice is the use of $\|\boldsymbol{x}\|_2$ and $\|\nabla^n \boldsymbol{x}\|_2$ to enforce the smoothness of an underlying signal $\boldsymbol{x}$, where ∇^n is the n-th order discrete finite difference operator. For simple notations, we define $\nabla^0 \boldsymbol{x} := \boldsymbol{x}$. In other words, a regularized PR problem can be expressed as

$$\underset{\boldsymbol{x}}{\text{minimize}} \|\nabla^n \boldsymbol{x}\|_2 \quad \text{s.t.} \quad |\mathcal{F}\boldsymbol{x}| = \boldsymbol{b}.$$

Unfortunately, Theorem 3.1 shows that $\|\nabla^n \boldsymbol{x}\|_2$ is completely determined by $|\mathcal{F}\boldsymbol{x}|$, which implies that such gradient-based regularization cannot resolve any ambiguities. But on the other hand, adding gradient-based regularizations may help to escape from local optima due to the nonconvex nature of the phase retrieval problem.

Theorem 3.1. *Given $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^N$, if $|\mathcal{F}\boldsymbol{x}| = |\mathcal{F}\boldsymbol{y}|$, then $\|\nabla^n \boldsymbol{x}\|_2 = \|\nabla^n \boldsymbol{y}\|_2$, for all $n = 0, 1, 2, \ldots$.*

One may wonder whether it is helpful to take more measurements and then impose regularizations. Theorem 3.2 implies that the gradient-type regularization is insufficient for the PR problem with more than phaseless $2N-1$ measurements.

Theorem 3.2. *Let $M \geq 2N-1$, given $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^N$, if $|\mathcal{F}_M \boldsymbol{x}| = |\mathcal{F}_M \boldsymbol{y}|$, then $\|\nabla^n \boldsymbol{x}\|_2 = \|\nabla^n \boldsymbol{y}\|_2$, for all $n = 0, 1, 2, \ldots$.*

Remark 3.3. When $N < M < 2N-1$, gradient-based regularization may help. For example, let $\boldsymbol{x} = (0,0,0,0,1,0,1,0,0,1,1)$ and $\boldsymbol{y} = (0,0,0,1,0,0,0,1,0,1,1)$. Both of them are of length 11 and have the same $|\mathcal{F}_M \boldsymbol{x}| = |\mathcal{F}_M \boldsymbol{y}|$ for $M = 12$, but $\|\nabla^3 \boldsymbol{x}\|_2^2 = 7.5 \neq 7 = \|\nabla^3 \boldsymbol{y}\|_2^2$, where the third order finite scheme $\nabla^3 \boldsymbol{x}$ is defined by $(\nabla^3 \boldsymbol{x})_k := -\frac{1}{2}x_{k-2} + x_{k-1} - x_{k+1} + \frac{1}{2}x_{k+2}$.

To prove Theorems 3.1-3.2, we need to review a classical result that $\text{Aut}(\boldsymbol{x})$ is determined by $|\mathcal{F}_{2N-1}\boldsymbol{x}|$ and vice versa, as stated in Theorem 3.4.

Theorem 3.4 ([9, 31])**.** *Given $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^N$, the following statements are equivalent:*

(1) $|\mathcal{F}_{2N-1}\boldsymbol{x}| = |\mathcal{F}_{2N-1}\boldsymbol{y}|$;

(2) $Aut(\boldsymbol{x}) = Aut(\boldsymbol{y})$.

We extend this analysis to an arbitrary number of measurements (not just $2N-1$) as well as to period autocorrelation (from regular autocorrelation). Specifically in Theorem 3.5, we show that when $M = N$, $\text{Aut}_p(\boldsymbol{x})$ and $\|\boldsymbol{v} * \boldsymbol{x}\|_2$ for $\boldsymbol{v} \in \mathbb{C}^N$ are determined by $|\mathcal{F}\boldsymbol{x}|$, and vice versa. A similar result for $M \geq 2N-1$ is presented in Theorem 3.6.

Theorem 3.5. *Given $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^N$, the following statements are equivalent:*

(1) $|\mathcal{F}\boldsymbol{x}| = |\mathcal{F}\boldsymbol{y}|$*;*

(2) $Aut_p(\boldsymbol{x}) = Aut_p(\boldsymbol{y})$*;*

(3) $\|\boldsymbol{v} * \boldsymbol{x}\|_2 = \|\boldsymbol{v} * \boldsymbol{y}\|_2 \ \forall \boldsymbol{v} \in \mathbb{C}^N$*.*

Theorem 3.6. *Given* $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^N$*,* $M \geq 2N - 1$*, the following statements are equivalent:*

(1) $|\mathcal{F}_M \boldsymbol{x}| = |\mathcal{F}_M \boldsymbol{y}|$

(2) $Aut(\boldsymbol{x}) = Aut(\boldsymbol{y})$

Also, either (1) or (2) implies that $Aut_p(\boldsymbol{x}) = Aut_p(\boldsymbol{y})$ *and* $\|\boldsymbol{v}*\boldsymbol{x}\|_2 = \|\boldsymbol{v}*\boldsymbol{y}\|_2 \ \forall \boldsymbol{v} \in \mathbb{C}^N$*. The converse does not necessarily hold.*

Remark 3.7. For $M < 2N-1$ and $M \neq N$, we cannot determine the autocorrelation from M magnitude measurements of $|\mathcal{F}_M(\boldsymbol{x})|$, due to an insufficient number of measurements.

To the best of our knowledge, the equivalence of phaseless measurements to $\|\boldsymbol{v}*\boldsymbol{x}\|_2, \forall \boldsymbol{v}$ is novel in the literature, which leads to useful consequences as characterized in Theorems 3.1 and 3.2. In particular, Theorem 3.1 directly follows from Theorem 3.5 (1) $\Rightarrow$ (3) and the fact that $\nabla^n \boldsymbol{x} = \boldsymbol{v}_n * \boldsymbol{x}$ for some $\boldsymbol{v}_n \in \mathbb{C}^N$. Similarly, Theorem 3.2 follows from Theorem 3.6.

4. Box Relaxation to Binary Constraint

We now restrict our attention to binary signals $\boldsymbol{x} \in \{0,1\}^N$, as another way of imposing additional prior knowledge to facilitate phase retrieval. Mathematically, we formulate the binary phase retrieval problem as follows,

$$\text{Find } \boldsymbol{x} \in \{0,1\}^N, \text{ s.t. } |\mathcal{F}\boldsymbol{x}| = \boldsymbol{b}. \tag{P}$$

Since the binary constraint is nonconvex, we relax it to a box constraint in a similar way as a linear problem [42]:

$$\text{Find } \boldsymbol{x} \in [0,1]^N, \text{ s.t. } |\mathcal{F}\boldsymbol{x}| = \boldsymbol{b}. \tag{Q}$$

Clearly, if (P) has a solution, then (Q) also has a solution. The question is whether we can recover $\boldsymbol{x}$ from $\boldsymbol{b}$ through (Q). Computationally, the binary constraint in (P) can be posed as a minimization problem of $\boldsymbol{x}(1-\boldsymbol{x})$ subject to $\boldsymbol{x} \in [0,1]^N$, which can be solved via the difference of the convex algorithm (DCA) [43, 44]. Each DCA iteration requires to a subproblem similar to the (Q) problem and it takes a few iterations for DCA to converge. Therefore, solving (Q) is computationally more efficient compared to (P). Theoretically, we prove in Theorem 4.1 that all the solutions to (Q) are solutions to (P) and have the same number of 1's as the ground-truth signal.

Theorem 4.1. *Given* $0 \leq \alpha < \beta$*,* $\boldsymbol{x} \in \{\alpha, \beta\}^N$ *and* $\boldsymbol{y} \in [\alpha, \beta]^N$*, if* $|\mathcal{F}\boldsymbol{x}| = |\mathcal{F}\boldsymbol{y}|$*, then* $\boldsymbol{y} \in \{\alpha, \beta\}^N$ *and* $\boldsymbol{y}$ *has the same number of* α*'s and* β*'s as* $\boldsymbol{x}$*.*

If $\{0,1\}^N$ in problem (P) is replaced by a set such that every element has the same modulus, one can also relax the problem to its convex hull.

Theorem 4.2. *Suppose $\mathcal{E}$ is a set of complex number and there exists some constant $c > 0$ such that $|z| = c \geq$ for all $z \in \mathcal{E}$. Given $\boldsymbol{x} \in \mathcal{E}^N$ and $\boldsymbol{y} \in$ conv $\mathcal{E}^N$, if $|\mathcal{F}\boldsymbol{x}| = |\mathcal{F}\boldsymbol{y}|$, then $\boldsymbol{y} \in \mathcal{E}^N$, where conv $\mathcal{E}$ denotes the convex hull of $\mathcal{E}$.*

We have a similar version of Theorem 4.1 when $\boldsymbol{x} \in \{-1, 1\}^N$.

Corollary 4.3. *Given $\boldsymbol{x} \in \{-1, 1\}^N$ and $\boldsymbol{y} \in [-1, 1]^N$, if $|\mathcal{F}\boldsymbol{x}| = |\mathcal{F}\boldsymbol{y}|$, then $\boldsymbol{y} \in \{-1, 1\}^N$, and the number of 1's in $\boldsymbol{y}$ is the same as the number of 1's in $\boldsymbol{x}$ or the number of -1 in $\boldsymbol{x}$.*

We then characterize trivial ambiguities for binary phase retrieval in Section 4.1 and extend to other sampling schemes in Section 4.2.

4.1. Ambiguities and Uniqueness

In addition to trivial ambiguities (1.1) for general PR, there is another type of ambiguity in the binary setting. For example, one has

$$|\mathcal{F}(1, 1, 1, 1, 0, 0, 1, 0, 0, 0)^T| = |\mathcal{F}(0, 0, 0, 0, 1, 1, 0, 1, 1, 1)^T|,$$

in which the two signals are not related by (1.1), but rather by switching zeros and ones. We present this ambiguity for binary phase retrieval in Corollary 4.5. In fact, this result can be easily extended to the complex case:

Proposition 4.4. *Given $\boldsymbol{x} \in \mathbb{C}^N$, $|\mathcal{F}\boldsymbol{x}| = |\mathcal{F}(c\mathbb{1} - \boldsymbol{x})|$ if and only if $c = \frac{1+e^{-i\theta}}{N} \sum x_i$ for some $\theta \in [0, 2\pi)$, where $\mathbb{1}$ denotes the vector of all one's, i.e. $\mathbb{1} = (1, 1, \ldots, 1)^T$.*

Applying Proposition 4.4 with $\theta = 0$ and noting that $\sum x_i = \|\boldsymbol{x}\|_0$ for binary signal $\boldsymbol{x}$, one easily obtains:

Corollary 4.5. *Given $\boldsymbol{x} \in \{0, 1\}^N$ and N is even, if $\|\boldsymbol{x}\|_0 = N/2$, then $|\mathcal{F}\boldsymbol{x}| = |\mathcal{F}(\mathbb{1} - \boldsymbol{x})|$.*

As a by-product from the proof of Proposition 4.4, we reveal an interesting fact, stating that if $\boldsymbol{x}$ and $\boldsymbol{y}$ have the same Fourier magnitude, then so do $(\mathbb{1} - \boldsymbol{x})$ and $(\mathbb{1} - \boldsymbol{y})$:

Proposition 4.6. *Given $\boldsymbol{x}, \boldsymbol{y} \in \{0, 1\}^N$, $|\mathcal{F}\boldsymbol{x}| = |\mathcal{F}\boldsymbol{y}|$ if and only if $|\mathcal{F}(\mathbb{1} - \boldsymbol{x})| = |\mathcal{F}(\mathbb{1} - \boldsymbol{y})|$.*

We show in Proposition 4.7 that the exact recovery of $\boldsymbol{x}$ up to trivial ambiguities (1.1) is guaranteed when $\|\boldsymbol{x}\|_0 \leq 3$ and $\|\boldsymbol{x}\|_0 \geq N - 3$. The proof uses the fact that $(\text{Aut}_p(\boldsymbol{x}))_k$ is the number of pairs of 1's with distance‡ k for a binary signal $\boldsymbol{x} \in \{0, 1\}^N$. The combinatorial nature of $\text{Aut}_p(\boldsymbol{x})$ guarantees the uniqueness of $\boldsymbol{x}$ up to trivial ambiguities.

Proposition 4.7. *Given $\boldsymbol{x} \in \{0, 1\}^N$, if $\|\boldsymbol{x}\|_0 = 0, 1, 2, 3, N - 3, N - 2, N - 1$ or N, then we can uniquely recover $\boldsymbol{x}$ from $|\mathcal{F}\boldsymbol{x}|$ up to the trivial ambiguities (1.1).*

‡ Note that it is a wrap-around distance. For example, x_0 and x_{N-1} are considered of distance 1.

Remark 4.8. The above does not hold for $4 \le \|\boldsymbol{x}\|_0 \le N-4$ in general. For example, $(0,0,0,0,0,1,0,1,0,0,1,1)^T$ and $(0,0,0,0,1,0,0,0,1,0,1,1)^T$ have the same magnitude after Fourier Transform, but they are not related to each other by trivial ambiguities.

Next, we would like to discuss the uniqueness in oversampling case. Recall that the Z-transform of a signal $\boldsymbol{x} \in \mathbb{C}^N$ is defined by

$$P_{\boldsymbol{x}}(z) = \sum_{k=0}^{N-1} x_k z^k,$$

which is a complex polynomial. The reciprocal polynomial $\tilde{P}_{\boldsymbol{x}}(z)$ of $P_{\boldsymbol{x}}(z)$ is defined by $\tilde{P}_{\boldsymbol{x}}(z) = z^n P_{\boldsymbol{x}}(z^{-1})$, where n is the degree of the polynomial $P_{\boldsymbol{x}}(z)$. If the Z-transform of an unknown binary signal $P_{\boldsymbol{x}}$ is either reciprocal or irreducible, then $\boldsymbol{x}$ can be recover uniquely up to conjugate inverse. Using this fact, the exact recovery up to trivial ambiguities in the oversampling case is characterized in Propositions 4.9-4.10.

Proposition 4.9. *Given $M \ge 2N-1$ in the setting of the oversampling Fourier PR, $\boldsymbol{x} \in \{0,1\}^N$, if $x_n = x_{N-1-n}$ for all $n = 0,1,\dots,N-1$, then we can recover $\boldsymbol{x}$ uniquely.*

Proposition 4.10. *Given $M \ge 2N-1$ in the setting of the oversampling Fourier PR, we can recover a random unknown binary $\boldsymbol{x} \in \{0,1\}^N$ uniquely up to the equivalence relation defined by $y_n = x_{N-1-n}$ with probability at least $\frac{c}{logN}$ for a constant $c > 0$.*

Note that the factor $\frac{c}{logN}$ in Proposition 4.10 is a lower bound. In fact, there is a conjecture in [45] that most of all polynomial with $\{0,1\}$ coefficients are irreducible. If it holds, a much better lower bound can be expected.

4.2. Extensions to other sampling schemes

We extend the analysis of Theorem 4.1 to the oversampling case, STFT, and FROG in Theorems 4.11–4.14, respectively. Also, it can be extended to $\{0,\alpha\}^N$, $\{-\alpha,\alpha\}^N$ simply by scaling, which are omitted.

Theorem 4.11. *Let $M \ge N$, given $\boldsymbol{x} \in \{0,1\}^N, \boldsymbol{y} \in [0,1]^N$, if $|\mathcal{F}_{N\to M}\boldsymbol{x}| = |\mathcal{F}_{N\to M}\boldsymbol{y}|$, then $\boldsymbol{y} \in \{0,1\}^N$ and $\|\boldsymbol{y}\|_0 = \|\boldsymbol{x}\|_0$.*

Theorem 4.12. *Given $\boldsymbol{x} \in \{0,1\}^N$ and $\boldsymbol{y} \in [0,1]^N$, if $\boldsymbol{x}$ and $\boldsymbol{y}$ have the same STFT under non-zero constant window, with $W \ge L$, as defined in (2.5), then $\boldsymbol{y} \in \{0,1\}^N$.*

Theorem 4.13. *Given $\boldsymbol{x} \in \{0,1\}^N$ and $\boldsymbol{y} \in [0,1]^N$, if $\boldsymbol{x}$ and $\boldsymbol{y}$ have the same FROG trace (2.6), then $\boldsymbol{y} \in \{0,1\}^N$ and $\|\boldsymbol{y}\|_0 = \|\boldsymbol{x}\|_0$.*

Theorem 4.14. *Given $\boldsymbol{x} \in \{-1,1\}^N$ and $\boldsymbol{y} \in [-1,1]^N$, if $\boldsymbol{x}$ and $\boldsymbol{y}$ have the same FROG trace, then $\boldsymbol{y} \in \{-1,1\}^N$.*

Remark 4.15. Unlike Theorem 4.3, the number of 1's in $\boldsymbol{x}$ is not necessarily the same as the number of 1's nor -1's in $\boldsymbol{y}$. For example, if we take $\boldsymbol{x} = (1,1)^T$ and $\boldsymbol{y} = (1,-1)^T$, then $\boldsymbol{x}$ and $\boldsymbol{y}$ have the same FROG trace.

5. Denoising

The preceding sections focus on the noiseless case, where the measured data we obtain is $\boldsymbol{b} = |\mathcal{F}\boldsymbol{x}|$. However, noise is inevitable in practice and there is a need to develop denoising techniques for phase retrieval. For this purpose, we consider a corrupted measurement $\tilde{\boldsymbol{b}} = \boldsymbol{b} + \boldsymbol{\eta}$ with a noise term $\boldsymbol{\eta}$. In the proof of Theorem 3.5 (specifically Lemma A.1), we reveal that $\mathcal{F}^{-1}(\boldsymbol{b} \odot \boldsymbol{b}) = \text{Aut}_p(\boldsymbol{x})$. If the noise $\boldsymbol{\eta}$ is small enough, then $\mathcal{F}^{-1}(\tilde{\boldsymbol{b}} \odot \tilde{\boldsymbol{b}})$ can be approximated by $\mathcal{F}^{-1}(\boldsymbol{b} \odot \boldsymbol{b})$, which is equivalent to $\text{Aut}_p(\boldsymbol{x})$. Proposition 5.1 is about the approximation error.

Proposition 5.1. *Given* $\epsilon > 0$, $\boldsymbol{x} \in \mathbb{C}^N \setminus \{\boldsymbol{0}\}$, $\boldsymbol{b} = |\mathcal{F}\boldsymbol{x}|$, $\tilde{\boldsymbol{b}} = \boldsymbol{b} + \boldsymbol{\eta}$ *for some noise* $\boldsymbol{\eta} \in \mathbb{C}^N$, *if* $\|\boldsymbol{\eta}\|_\infty < min\{\frac{\epsilon}{4\|\boldsymbol{b}\|_\infty}, \frac{\epsilon}{2}, 1\}$, *then* $\|\mathcal{F}^{-1}(\tilde{\boldsymbol{b}} \odot \tilde{\boldsymbol{b}}) - Aut_p(\boldsymbol{x})\|_\infty < \epsilon$.

Ideally, it would be helpful to analyze the error to the ground-truth signal, which is unfortunately impossible due to trivial and non-trivial ambiguities.

In the following, we restrict the ground-truth signal $\boldsymbol{x} \in \{0,1\}^N$ and observe a denoising scheme based on Proposition 5.1 often gives good results. For binary signal $\boldsymbol{x}$, we know $\text{Aut}_p(\boldsymbol{x}) \in \mathbb{Z}^N$. If the noise $\boldsymbol{\eta}$ is small such that $\|\mathcal{F}^{-1}(\tilde{\boldsymbol{b}}\odot\tilde{\boldsymbol{b}}) - \text{Aut}_p(\boldsymbol{x})\|_\infty < 1/2$, we can round off each entry of $\mathcal{F}^{-1}(\tilde{\boldsymbol{b}} \odot \tilde{\boldsymbol{b}})$ to the nearest integer to perform denoising. Since $(\text{Aut}_p(\boldsymbol{x}))_k$ is the number of pair of 1's with distance k, $\|\mathcal{F}^{-1}(\tilde{\boldsymbol{b}}\odot\tilde{\boldsymbol{b}}) - \text{Aut}_p(\boldsymbol{x})\|_\infty \geq 1/2$ means the measurements cannot give us the true number of pairs of 1's with distance k. In this circumstance, one should not expect to have a successful recovery.

Proposition 5.2. *Given* $\boldsymbol{x} \in \{0,1\}^N \setminus \{\boldsymbol{0}\}$, $\boldsymbol{b} = |\mathcal{F}\boldsymbol{x}|$, $\tilde{\boldsymbol{b}} = \boldsymbol{b} + \boldsymbol{\eta}$ *for some noise* $\boldsymbol{\eta} \in \mathbb{C}^N$, *if* $\|\boldsymbol{\eta}\|_\infty < \frac{1}{8\|\boldsymbol{x}\|_0}$, *then* $\|\mathcal{F}^{-1}(\tilde{\boldsymbol{b}} \odot \tilde{\boldsymbol{b}}) - Aut_p(\boldsymbol{x})\|_\infty < \frac{1}{2}$.

Since $\|\boldsymbol{x}\|_0 \leq N$, it is straightforward to have Corollary 5.3. We can also express the error analysis in Proposition 5.2 in terms of signal-to-noise ratio (SNR).

Corollary 5.3. *Given* $\boldsymbol{x} \in \{0,1\}^N$, $\boldsymbol{b} = |\mathcal{F}\boldsymbol{x}|$, $\tilde{\boldsymbol{b}} = \boldsymbol{b} + \boldsymbol{\eta}$ *for some noise* $\boldsymbol{\eta} \in \mathbb{C}^N$, *if* $\|\boldsymbol{\eta}\|_\infty < \frac{1}{8N}$, *then* $\|\mathcal{F}^{-1}(\tilde{\boldsymbol{b}} \odot \tilde{\boldsymbol{b}}) - Aut_p(\boldsymbol{x})\|_\infty < \frac{1}{2}$.

Recall SNR is defined by

$$\text{SNR}_{\text{dB}} = 10 \log_{10} \frac{\|\boldsymbol{x}\|_2^2}{\|\boldsymbol{\eta}\|_2^2}.$$

Proposition 5.4 presents a condition to safely round off each entry to 0 and 1.

Corollary 5.4. *Given* $\boldsymbol{x} \in \{0,1\}^N \setminus \boldsymbol{0}$, *if*

$$SNR_{dB} > 10 \log_{10}(64) + 30 \log_{10} \|\boldsymbol{x}\|_0,$$

then $\|\mathcal{F}^{-1}(\tilde{\boldsymbol{b}} \odot \tilde{\boldsymbol{b}}) - Aut_p(\boldsymbol{x})\|_\infty < \frac{1}{2}$.

The proposed denoising scheme, referred to as *rounding scheme*, is described as follows: given a corrupted measurement $\tilde{\boldsymbol{b}} \in \mathbb{C}^N$,

Algorithm 1 Fourier phase retrieval subject to a box constraint (5.5) via ADMM.

Input: $\boldsymbol{b}$ and two positive parameters ρ_1, ρ_2

Initialize $k = 0, \boldsymbol{w}^0 = 0, \boldsymbol{d}^0 = 0, \boldsymbol{y}^0 = 0, \boldsymbol{z}^0 = \boldsymbol{b}e^{i\boldsymbol{\phi}}$ with a random vector $\boldsymbol{\phi}$

while stopping conditions are not satisfied **do**

$\quad \boldsymbol{x}^{k+1} = (\rho_1 + \rho_2)^{-1}(\rho_1 \mathcal{F}^* \boldsymbol{z}^k + \mathcal{F}^* \boldsymbol{d}^k + \rho_2 \boldsymbol{y}^k - \boldsymbol{w}^k)$

$\quad \boldsymbol{y}^{k+1} = \min(\max(\boldsymbol{x}^{k+1} + \boldsymbol{w}^k/\rho_2, 0), 1)$

$\quad \boldsymbol{z}^{k+1} = \mathbf{prox}_{\rho_1}(\mathcal{F}\boldsymbol{x}^{k+1} - \boldsymbol{d}^k/\rho_1)$

$\quad \boldsymbol{d}^{k+1} = \boldsymbol{d}^k + \rho_1(\boldsymbol{z}^{k+1} - \mathcal{F}\boldsymbol{x}^{k+1})$

$\quad \boldsymbol{w}^{k+1} = \boldsymbol{w}^k + \rho_2(\boldsymbol{x}^{k+1} - \boldsymbol{y}^{k+1})$

$\quad$ k = k+1

end while

Output the solution $\boldsymbol{x}^* = \boldsymbol{x}^k$

(i) Round off each entry $\mathcal{F}^{-1}(\tilde{\boldsymbol{b}} \odot \tilde{\boldsymbol{b}})$ to nearest integer to get the autocorrelation $\mathrm{Aut}_p(\boldsymbol{x})$.

(ii) Calculate $\boldsymbol{b} = \sqrt{\mathcal{F}(\mathrm{Aut}_p(\boldsymbol{x}))}$, where the square root is taken entrywise.

(iii) Solve the minimization problem:

$$\boldsymbol{x}^{\boldsymbol{*}} = \underset{\boldsymbol{x}}{\mathrm{argmin}} \| |\mathcal{F}\boldsymbol{x}| - \boldsymbol{b}\|_2^2 \quad \text{s.t.} \quad \boldsymbol{x} \in [0, 1]^N. \tag{5.5}$$

(iv) Round off each entry of $\boldsymbol{x}^*$ to be either 0 or 1.

We compare the proposed scheme with a *naïve scheme* with the following steps: given a corrupted measurement $\boldsymbol{b} \in \mathbb{C}^N$,

(i) Solve the minimization problem (5.5).

(ii) Round off each entry of $\boldsymbol{x}^*$ to be either 0 or 1.

Both rounding and naïve schemes require to find a solution to (5.5), which can be solved via the alternating direction methods of multiplier (ADMM) [46]. We summarize in Algorithm 1 for Fourier phase retrieval subject to the $[0, 1]$-box constraint (5.5) via ADMM; for more details, please refer to [30]. Notice that ADMM requires two parameters: ρ_1 and ρ_2. We examine the effects of these two parameters on the naïve scheme and the rouding scheme in terms of success rates. We consider a binary vector of length 50 with 5 nonzero element as the ground-truth $\boldsymbol{x}_{\text{true}}$, which is contaminated by noise with SNR= 16 dB. We choose ρ_1, ρ_2 among a candidate set of $\{10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}\}$ and plot the success rates in Figure 1 based on 1000 random realizations; we declare a trial is successful if $\| |\mathcal{F}\boldsymbol{x}_{\text{recovered}}| - \boldsymbol{b}\| < 10^{-6}$. We observe no significant difference when $\rho_1 = \rho_2$ and hence we choose $\rho_1 = \rho_2 = 10^{-5}$ for both rounding and naïve schemes throughout the experiments. Figure 1 also shows that our rounding scheme outperforms the naïve scheme when $\rho_1 = \rho_2$.

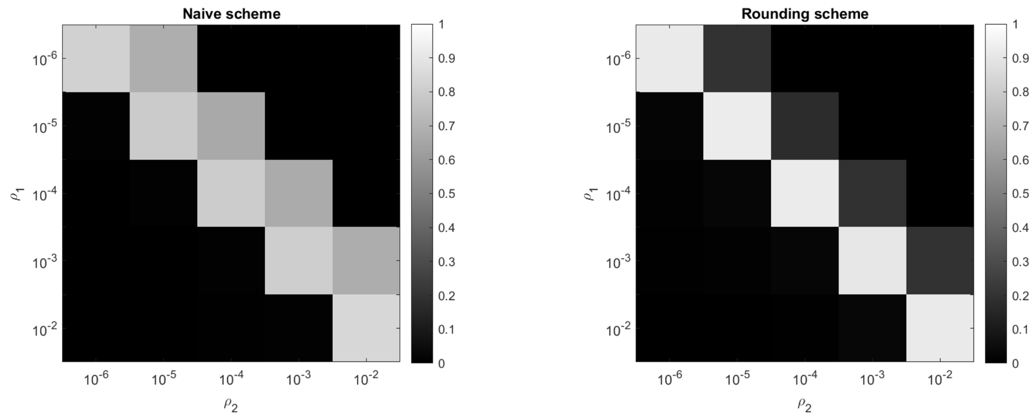

Figure 1. Influence of ρ_1, ρ_2 on the naïve scheme (left) and the rounding scheme (right) in terms of success rates when SNR = 16 dB and $\|\boldsymbol{x}_{\text{true}}\|_0 = 5$ for $\boldsymbol{x}_{\text{true}}$ of length 50.

5.1. Fourier phase retrieval

We then compare the performance of both schemes in terms of success rates. We consider the ground-truth signal $\boldsymbol{x}_{\text{true}}$ is a binary vector with different combinations of sparsity and noise levels in the Fourier measurements. In particular, we examine ten sparsity levels $(1, 2, \ldots, 10)$ and generate the noisy measurements $\tilde{\boldsymbol{b}}$ by adding Gaussian noise with SNR $= (36, 32, \ldots, 0)$ dB. We plot the success rates of recovering signals of length 50 and 100 based on 1000 random realizations in Figure 2 and 3, respectively. Compared to the naïve scheme, the rounding scheme works much better when the signal is sparse, which is expected by Proposition 5.2 that sparser signals allow for larger tolerance of the noise. According to Corollary 5.4, the exact recovery bound of SNR is calculated as $18 + 30\log_{10}\|\boldsymbol{x}\|_0$, which aligns well with Figures 2-3. Figure 4 gives some examples on false reconstructions, which implies that one scheme does not dominate the other, as there exist examples when the naïve scheme succeeds and the rounding one fails, and vice versa. The conclusion that the rounding scheme is better is based on the success rates.

5.2. Extension to oversampling Fourier Transform

One may extend our method to oversampling schemes to find the periodic autocorrelation or the regular autocorrelation. We conduct numerical simulations for this case, while leaving the theoretical analysis for the future investigation. When the number of measurements do not match the number of coefficients in autocorrelation, one can perform a polynomial regression and round off to the nearest integer to find the autocorrelation, following equation A.4. The extension for the rounding scheme is summarized as follows, similar for the naïve scheme.

(i) Use polynomial regression to estimate the degree $2N-1$ polynomial $A(z)$ by $e^{\frac{2\pi i k(N-1)}{M}} A(e^{-\frac{2\pi i k}{M}}) = \tilde{b}_k^2$.

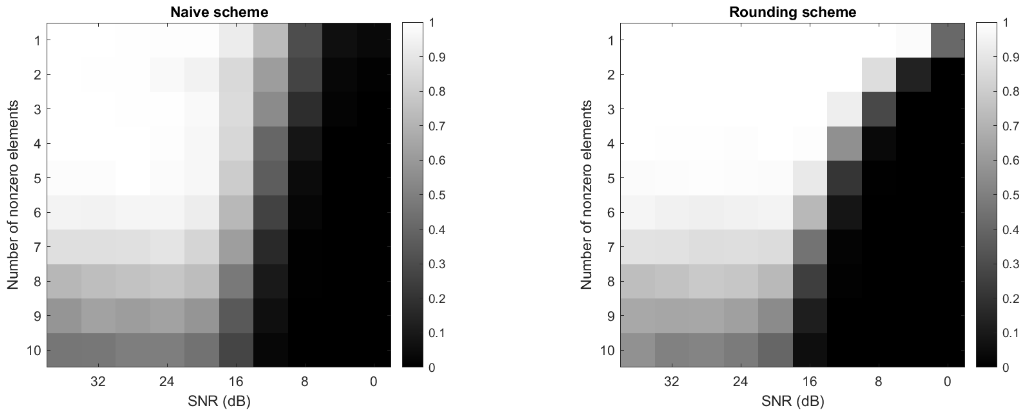

Figure 2. Comparison of the naïve scheme (left) and rounding scheme (right) in terms of success rates of Fourier phase retrieval for a signal of length 50. The value at each combination of sparsity and SNR is based on 1000 random realizations.

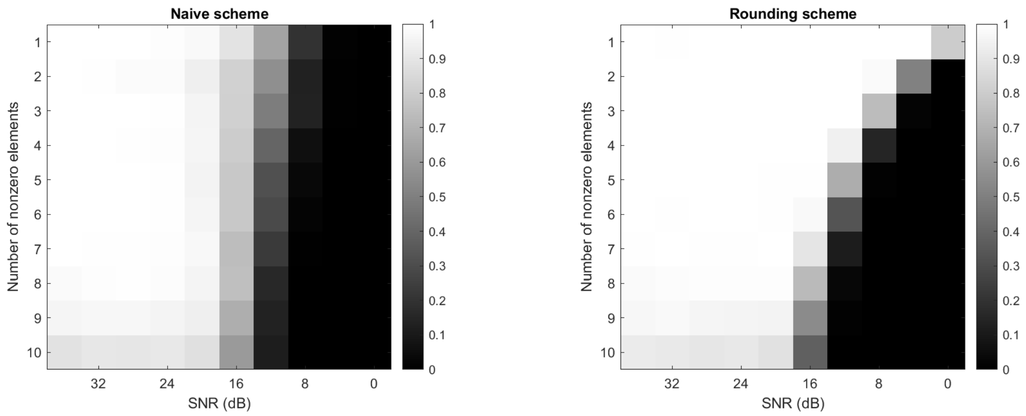

Figure 3. Comparison of the naïve scheme (left) and rounding scheme (right) in terms of success rates of Fourier phase retrieval for a signal of length 100.

(ii) Round off each coefficient of $A(z)$ to the nearest integer to get the polynomial $B(z)$ with $\mathrm{Aut}(\boldsymbol{x})$ as its coefficient.

(iii) Calculate $b_k = \sqrt{e^{\frac{2\pi i k(N-1)}{M}} B(e^{-\frac{2\pi i k}{M}})}$

(iv) Solve the minimization problem:

$$\boldsymbol{x}^* = \underset{\boldsymbol{x}}{\mathrm{argmin}} \| |\mathcal{F}_M \boldsymbol{x}| - \boldsymbol{b} \|_2^2 \quad \text{s.t.} \quad \boldsymbol{x} \in [0,1]^N. \tag{5.6}$$

(v) Round off each entry of $\boldsymbol{x}^*$ to be either 0 or 1

Again, we compare the performance of the naïve scheme and the rounding scheme in terms of success rates. We consider the ground-truth signal $\boldsymbol{x}_{\text{true}}$ is a binary vector of length 50 with different combinations of sparsity and noise levels. We take 99 oversampled Fourier magnitude measurements for each signal, i.e. $N = 50, M = 99$ in (2.1). We consider ten sparsity levels $(1, 2, \ldots, 10)$ and generate the noisy measurements

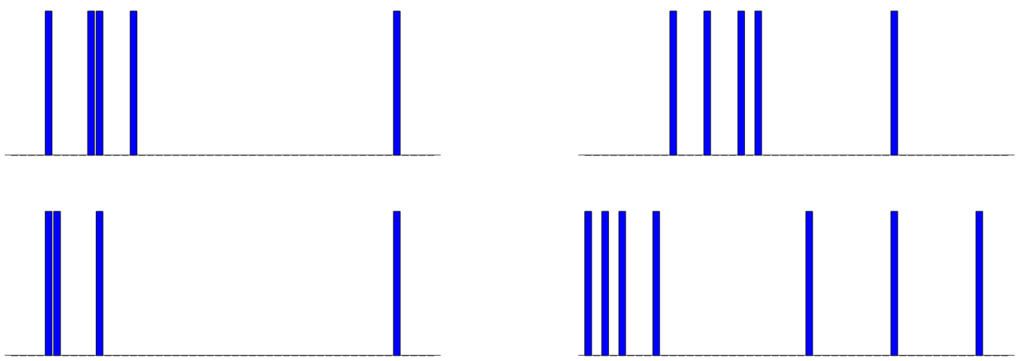

Figure 4. Failed reconstructions by the naïve scheme (left) and the rounding scheme (right), when the other scheme succeeds. The ground-truth signals are plotted on the top, while the reconstructed ones are on the bottom.

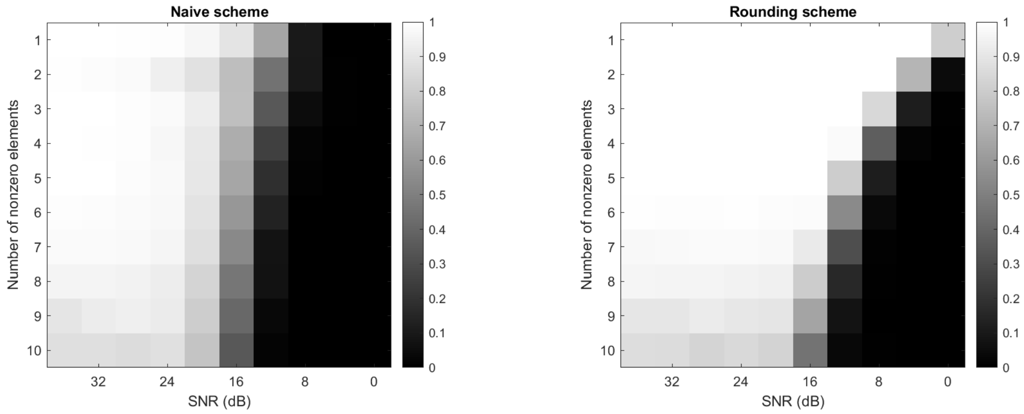

Figure 5. Comparison of the naïve scheme (left) and rounding scheme (right) in the case of oversampled Fourier phase retrieval.

$\tilde{\boldsymbol{b}}$ by adding Gaussian noise with SNR $= (36, 32, \ldots, 0)$ dB. In Figure 5, we plot the success rates based on 1000 random realizations, which shows the rounding scheme outperforms the naïve one.

6. Conclusions

In this paper, we improved upon an autocorrelation-based characterization of Fourier phase retrieval. We discuss several choices of regularization terms and measurements. Our analysis suggested that a gradient-based regularization, i.e. $\|\nabla^n \boldsymbol{x}\|_2$, is redundant to the magnitude measurements, thus not helpful to phase retrieval. Furthermore, we proved that binary signals can be recovered by imposing a box constraint. We also presented ambiguities and uniqueness for binary phase retrieval. Finally, we proposed a denoising scheme suggested by characterization theorems. Since the proposed denoising scheme involves rounding, it is interesting to extend to 2D images, in which the measured data are often integer-valued. This will be our future work. Another future direction involves theoretical analysis of oversampling schemes and noisy measurements for phase retrieval.

Acknowledgements

WHW and TZ are sponsored in part by the National Natural Science Foundation of China under Grant 11671002, CUHK start-up and CUHK DAG 4053296, 4053342. YL is partially supported by NSF CAREER 1846690. SM was funded by the Center for Advanced Mathematics for Energy Research Applications, US Department of Energy, contract number DOE-DE-AC03-76SF00098.

Appendix A. Proof of Theorems 3.5 and 3.6

To prove Theorem 3.5, we introduce Lemma A.1 and A.2. Note that Lemma A.1 is a periodic version of a similar result in [47, P. 215] and Lemma A.2 is Parseval's Theorem.

Lemma A.1. *$\mathcal{F}(Aut_p(\boldsymbol{x})) = |\mathcal{F}\boldsymbol{x}| \odot |\mathcal{F}\boldsymbol{x}|, \ \forall \boldsymbol{x} \in \mathbb{C}^N$.*

Proof. It is straightforward that for all $j = 0, 1, \ldots, N-1$, we have

$$\begin{aligned}(\mathcal{F}(\mathrm{Aut}_p(\boldsymbol{x})))_j &= \sum_{m=0}^{N-1}\sum_{n=0}^{N-1} x_{n+m}\overline{x_n}\omega^{mj} \\ =&\sum_{m=0}^{N-1}\sum_{n=0}^{N-1} x_m\overline{x_n}\omega^{(m-n)j} = (\sum_{m=0}^{N-1} x_m\omega^{mj})\overline{(\sum_{n=0}^{N-1} x_n\omega^{nj})} \\ =&(\mathcal{F}\boldsymbol{x})_j\overline{(\mathcal{F}\boldsymbol{x})}_j = |(\mathcal{F}\boldsymbol{x})_j|^2.\end{aligned}$$

□

Lemma A.2 (Application of Parseval's Theorem)**.** *Given $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^N$, if $|\mathcal{F}\boldsymbol{x}| = |\mathcal{F}\boldsymbol{y}|$, then $\|\boldsymbol{x}\|_2 = \|\boldsymbol{y}\|_2$.*

Proof. Since $\frac{1}{\sqrt{N}}\mathcal{F}$ is unitary, we have

$$\|\boldsymbol{y}\|_2 = \|\frac{1}{\sqrt{N}}\mathcal{F}\boldsymbol{y}\|_2 = \frac{1}{\sqrt{N}}\||\mathcal{F}\boldsymbol{y}|\|_2 = \frac{1}{\sqrt{N}}\||\mathcal{F}\boldsymbol{x}|\|_2 = \|\boldsymbol{x}\|_2.$$

□

Proof of Theorem 3.5. (1) $\Rightarrow$ (2). Suppose $|\mathcal{F}\boldsymbol{x}| = |\mathcal{F}\boldsymbol{y}|$, by Lemma A.1, $\mathcal{F}(\mathrm{Aut}_p(\boldsymbol{x})) = |\mathcal{F}\boldsymbol{x}| \odot |\mathcal{F}\boldsymbol{x}|$. Hence, $\mathrm{Aut}_p(\boldsymbol{x}) = \mathcal{F}^{-1}(|\mathcal{F}\boldsymbol{x}| \odot |\mathcal{F}\boldsymbol{x}|) = \mathcal{F}^{-1}(|\mathcal{F}\boldsymbol{y}| \odot |\mathcal{F}\boldsymbol{y}|) = \mathrm{Aut}_p(\boldsymbol{y})$.

(2) $\Rightarrow$ (1). Suppose $\mathrm{Aut}_p(\boldsymbol{x}) = \mathrm{Aut}_p(\boldsymbol{y})$. By Lemma A.1, we have $|\mathcal{F}\boldsymbol{x}| = \sqrt{\mathcal{F}(\mathrm{Aut}_p(\boldsymbol{x}))} = \sqrt{\mathcal{F}(\mathrm{Aut}_p(\boldsymbol{y}))} = |\mathcal{F}\boldsymbol{y}|$, where the square root is taken entrywisely.

(1) $\Rightarrow$ (3). By the Convolution Theorem, we have $\forall \boldsymbol{v} \in \mathbb{C}^N$ and $j = 0, 1, \ldots, N-1$,

$$(\mathcal{F}(\boldsymbol{v} * \boldsymbol{x}))_j = (\mathcal{F}\boldsymbol{v})_j \times (\mathcal{F}\boldsymbol{x})_j,$$

thus leading to,

$$|(\mathcal{F}(\boldsymbol{v} * \boldsymbol{x}))_j| = |(\mathcal{F}\boldsymbol{v})_j||(\mathcal{F}\boldsymbol{x})_j|.$$

Similar result holds for $\mathcal{F}(\boldsymbol{v} * \boldsymbol{y})$. Since $|\mathcal{F}\boldsymbol{x}| = |\mathcal{F}\boldsymbol{y}|$ (by assumption), we have $|\mathcal{F}(\boldsymbol{v} * \boldsymbol{x})| = |\mathcal{F}(\boldsymbol{v} * \boldsymbol{y})|$, which implies that $\|\boldsymbol{v} * \boldsymbol{x}\|_2 = \|\boldsymbol{v} * \boldsymbol{y}\|_2$ by Lemma A.2.

(3) $\Rightarrow$ (1). Suppose $\|\boldsymbol{v} * \boldsymbol{x}\|_2 = \|\boldsymbol{v} * \boldsymbol{y}\|_2$ for all $\boldsymbol{v} \in \mathbb{C}^N$. Since $\mathcal{F}$ is invertible, we can choose $\boldsymbol{v}_k = \mathcal{F}^{-1}\boldsymbol{e}_k \in \mathbb{C}^N$. Then we have

$$\begin{aligned}
&\|\boldsymbol{v}_k * \boldsymbol{x}\|_2^2 = \|\frac{1}{\sqrt{N}}\mathcal{F}(\boldsymbol{v}_k * \boldsymbol{x})\|_2^2 \\
=&\frac{1}{N}\sum_{j=0}^{N-1}|(\mathcal{F}(\boldsymbol{v}_k * \boldsymbol{x}))_j|^2 = \frac{1}{N}\sum_{j=0}^{N-1}|(\mathcal{F}\boldsymbol{v}_k)_j|^2|(\mathcal{F}\boldsymbol{x})_j|^2 \\
=&\frac{1}{N}\sum_{j=0}^{N-1}|(\boldsymbol{e}_k)_j|^2|(\mathcal{F}\boldsymbol{x})_j|^2 = \frac{1}{N}|(\mathcal{F}\boldsymbol{x})_k|^2.
\end{aligned}$$

Similarly, we have $\|\boldsymbol{v}_k * \boldsymbol{x}\|_2^2 = \|\boldsymbol{v}_k * \boldsymbol{y}\|_2^2 = \dfrac{1}{N}|(\mathcal{F}\boldsymbol{y})_k|^2$ and hence $|\mathcal{F}\boldsymbol{x}| = |\mathcal{F}\boldsymbol{y}|$. □

Proof of Theorem 3.6. (1) $\Rightarrow$ (2) Define

$$A_{\boldsymbol{x}}(z) = z^{N-1}\sum_{n=-(N-1)}^{N-1}(\mathrm{Aut}(\boldsymbol{x}))_n z^n, \tag{A.3}$$

and similarly for $A_{\boldsymbol{y}}(z)$. Note that

$$e^{\frac{2\pi ik(N-1)}{M}}A_{\boldsymbol{x}}(e^{-\frac{2\pi ik}{M}}) = |(\mathcal{F}_M\boldsymbol{x})_k|^2 = |(\mathcal{F}_M\boldsymbol{y})_k|^2 = e^{\frac{2\pi ik(N-1)}{M}}A_{\boldsymbol{y}}(e^{-\frac{2\pi ik}{M}}), \tag{A.4}$$

for all $k = 0, 1, \ldots, M-1$. Since $A_{\boldsymbol{x}}$ and $A_{\boldsymbol{y}}$ are polynomials of degree at most $2N-1$, their coefficients are determined by $|\mathcal{F}_M\boldsymbol{x}| = |\mathcal{F}_M\boldsymbol{y}|$, which is a system of M linear equations with $M \geq 2N-1$. Thus, $\mathrm{Aut}(\boldsymbol{x}) = \mathrm{Aut}(\boldsymbol{y})$.

(2) $\Rightarrow$ (1). Suppose $\mathrm{Aut}(\boldsymbol{x}) = \mathrm{Aut}(\boldsymbol{y})$. Then $A_{\boldsymbol{x}}(z) = A_{\boldsymbol{y}}(z)$. Since $M \geq 2N-1$, we have

$$|(\mathcal{F}_M\boldsymbol{x})_k|^2 = e^{\frac{2\pi ik(N-1)}{M}}A_{\boldsymbol{x}}(e^{-\frac{2\pi ik}{M}}) = e^{\frac{2\pi ik(N-1)}{M}}A_{\boldsymbol{y}}(e^{-\frac{2\pi ik}{M}}) = |(\mathcal{F}_M\boldsymbol{y})_k|^2, \tag{A.5}$$

for all $k = 0, 1, \ldots, M-1$.

It remains to prove that (1) implies $\mathrm{Aut}_p(\boldsymbol{x}) = \mathrm{Aut}_p(\boldsymbol{y})$ and $\|\boldsymbol{v} * \boldsymbol{x}\|_2 = \|\boldsymbol{v} * \boldsymbol{y}\|_2$ $\forall \boldsymbol{v} \in \mathbb{C}^N$. This directly follows from (1) $\Rightarrow$ (2) and (1) $\Rightarrow$ (3) in Theorem 3.5 by considering $M = N$ in equation (A.5). □

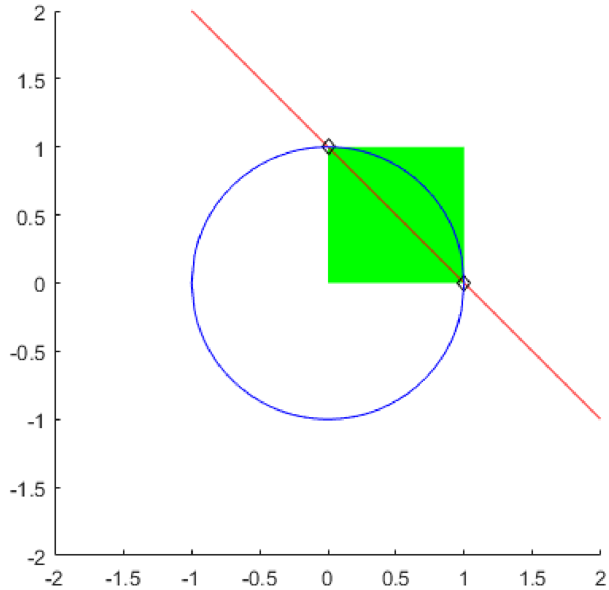

Figure B.1. Illustration of Theorem 4.1 with $\alpha = 0$ and $\beta = 1$ when $\boldsymbol{x} = (1, 0)$. The plane constraint is $y_1 + y_2 = 1$, indicated by the red line. The sphere constraint is $y_1^2 + y_2^2 = 1$, indicated by the blue circle. The box constraint is $0 \le y_1, y_2 \le 1$, indicated by the green square. The solutions to the problem (Q) lie on the intersection of these 3 sets, indicated by the black points. Both black points are ambiguous due to Proposition 4.5.

Appendix B. Proof of Theorem 4.1 to Corollary 4.3

We given a geometry interpretation to facilitate the proof of Theorem 4.1. For $\alpha = 0$ and $\beta = 1$, we have $\boldsymbol{y} \in [0, 1]^N$. Lemma A.2 implies that $\boldsymbol{y}$ must lie on a sphere, while $\sum x_i = \sum y_i$ implies that $\boldsymbol{y}$ must lie on a plane. Therefore, the solution $\boldsymbol{y}$ must be on the intersection of these three sets, as illustrated in Figure B.1.

Proof of Theorem 4.1. Rewrite $|(\mathcal{F}\boldsymbol{x})_0| = |(\mathcal{F}\boldsymbol{y})_0|$, we know that $\boldsymbol{y}$ lies on the plane $P : \sum x_i = \sum y_i$, which is convex. The box constraint $\boldsymbol{y} \in [\alpha, \beta]^N$ is also convex. Therefore, we define $\mathcal{C} := P \cap [\alpha, \beta]^N$, which is a convex compact set. By Krein-Milman Theorem [48, Theorem 3.23], C is the closure of the convex hull of its extreme points.

We claim that the set of extreme points $\mathcal{E} = \{\boldsymbol{z}_i\}_{i \in \mathcal{I}}$ is a subset of points in $\{\alpha, \beta\}^N$ with the same number of α's and β's as $\boldsymbol{x}$. Given $\boldsymbol{w}$ be an extreme point of $\mathcal{C}$, assume that $\boldsymbol{w}$ does not belong to $\{\alpha, \beta\}^N$. Since $\sum w_i = \sum x_i$ and $\boldsymbol{x} \in \{\alpha, \beta\}^N$, there exists some $i < j$ such that $w_i, w_j \neq \alpha$ and β (otherwise, we will have $\boldsymbol{w} \in \{\alpha, \beta\}^N$). Choose small $\epsilon > 0$ such that $w_i, w_j > \alpha + \epsilon$ and $w_i, w_j < \beta - \epsilon$. Let $\boldsymbol{w}_1 = (w_0, w_1, \ldots, w_i + \epsilon, \ldots, w_j - \epsilon, \ldots, w_{N-1})^T$ and $\boldsymbol{w}_2 = (w_0, w_1, \ldots, w_i - \epsilon, \ldots, w_j + \epsilon, \ldots, w_{N-1})^T$. Then $\boldsymbol{w}_1, \boldsymbol{w}_2 \in \mathcal{C}$ and $\boldsymbol{w} = \frac{1}{2}(\boldsymbol{w}_1 + \boldsymbol{w}_2)$, contradicting the fact that $\boldsymbol{w}$ is an extreme point of $\mathcal{C}$. Hence, we have $\boldsymbol{w} \in \{\alpha, \beta\}^N$. It follows from $\sum w_i = \sum x_i$ that $\boldsymbol{w}$ has the same number of α's and β's as $\boldsymbol{x}$. Since $\mathcal{E}$ is a finite set, the convex hull of $\mathcal{E}$ is compact and thus equal to $\mathcal{C}$.

Since $\boldsymbol{y} \in \mathcal{C}$, we write $\boldsymbol{y} = \sum \lambda_i \boldsymbol{z}_i$ for some $0 \le \lambda_i \le 1, \sum \lambda_i = 1$. Since $\boldsymbol{z}_i$ has the same number of α's and β's as $\boldsymbol{x}$, then $f(\boldsymbol{x}) = f(\boldsymbol{z}_i)$ for all $i \in \mathcal{I}$, where $f(\boldsymbol{w}) := \|\boldsymbol{w}\|_2^2$, which is a strictly convex function. By Lemma A.2, we have $f(\boldsymbol{y}) = f(\boldsymbol{x})$. If $\boldsymbol{y}$ does not belong to $\mathcal{E}$, then we have

$$f(\boldsymbol{y}) < \sum \lambda_i f(\boldsymbol{z}_i) = \sum \lambda_i f(\boldsymbol{x}) = f(\boldsymbol{x}) \sum \lambda_i = f(\boldsymbol{x}),$$

which is a contradiction. So $\boldsymbol{y} \in \mathcal{E}$, i.e. $\boldsymbol{y} \in \{\alpha, \beta\}^N$ and has the same number of α's and β's as $\boldsymbol{x}$. □

The proof of Theorem 4.2 is based on the convexity to show that every $\boldsymbol{y}$ lies in the convex hull with some entries y_i having smaller value than α.

Proof of Theorem 4.2. Since $\boldsymbol{x} \in \mathcal{E}^N$, we have $\|\boldsymbol{x}\|_2^2 = Nc^2$. By Lemma A.2, we have $\|\boldsymbol{y}\|_2^2 = \|\boldsymbol{x}\|_2^2 = Nc^2$.

Note that for all $i = 1, 2, \ldots, N, y_i = \sum \lambda_{ik} z_{ik}$ for some $\sum \lambda_{ik} = 1, 0 \le \lambda_{ik} \le 1, z_{ik} \in \mathcal{E}$ and

$$|y_i| = |\sum \lambda_{ik} z_{ik}| \le \sum \lambda_{ik} |z_{ik}| = \sum \lambda_{ik} c = c.$$

If there exists some y_i such that $y_i \in \text{conv } \mathcal{E} \setminus \mathcal{E}$,then

$$|y_i| = |\sum \lambda_{ik} z_{ik}| < \sum \lambda_{ik} |z_{ik}| = \sum \lambda_{ik} c = c.$$

Now,

$$\|\boldsymbol{y}\|_2^2 = \sum |y_i|^2 < \sum c2 = Nc^2,$$

which leads to a contradiction. Thus, we must have $y_i \in \mathcal{E}$ for all $i = 1, 2, \ldots, N$, i.e. $\boldsymbol{y} \in \mathcal{E}^N$ □

Proof of Corollary 4.3. The fact that $\boldsymbol{y} \in \{-1, 1\}^N$ follows from Theorem 4.2 directly. Now, $|(\mathcal{F}\boldsymbol{x})_0| = |(\mathcal{F}\boldsymbol{y})_0|$ implies that $\sum_{i=0}^{N-1} x_i = \pm \sum_{i=0}^{N-1} y_i$. Denote the number of 1's in $\boldsymbol{x}$ by $n_{\boldsymbol{x}}$, and define $n_{\boldsymbol{y}}$ similarly, then we have $n_{\boldsymbol{x}} - (N - n_{\boldsymbol{x}}) = \pm\big(n_{\boldsymbol{y}} - (N - n_{\boldsymbol{y}}\big)$. We either have $n_{\boldsymbol{x}} = n_{\boldsymbol{y}}$ or $n_{\boldsymbol{x}} = N$. The result now follows. □

Appendix C. Proof of Propositions 4.4–4.10

Proof of Proposition 4.4. Suppose $|\mathcal{F}\boldsymbol{x}| = |\mathcal{F}(c\mathbb{1} - \boldsymbol{x})|$, i.e. $(\mathcal{F}\boldsymbol{x})_0 = e^{i\theta}(\mathcal{F}(c\mathbb{1} - \boldsymbol{x}))_0$ for some $\theta \in [0, 2\pi)$. Thus, $\sum x_i = e^{i\theta}(Nc - \sum x_i)$, $c = \frac{1+e^{-i\theta}}{N} \sum x_i$.

On the other hand, suppose $c = \frac{1+e^{-i\theta}}{N} \sum x_i$ for some $\theta \in [0, 2\pi)$. Since $\mathcal{F}\boldsymbol{x} + \mathcal{F}(c\mathbb{1} - \boldsymbol{x}) = c\mathcal{F}\mathbb{1} = Nc\boldsymbol{e}_0$, one has $(\mathcal{F}\boldsymbol{x})_j + (\mathcal{F}(c\mathbb{1} - \boldsymbol{x}))_j = 0$ for $j = 1, 2, \ldots, N-1$. In particular, we obtain $|(\mathcal{F}\boldsymbol{x})_j| = |(\mathcal{F}(\mathbb{1} - \boldsymbol{x}))_j|$ and clearly $|(\mathcal{F}\boldsymbol{x})_0| = |(\mathcal{F}(c\mathbb{1} - \boldsymbol{x}))_0|$ due to the choice of c. □

Proof of Proposition 4.6. Similar to Proposition 4.4, we have

$$|(\mathcal{F}(\mathbb{1}-\boldsymbol{x}))_j| = |(\mathcal{F}\boldsymbol{x})_j| = |(\mathcal{F}\boldsymbol{y})_j| = |(\mathcal{F}(\mathbb{1}-\boldsymbol{y}))_j|$$

for $j = 1, 2, \ldots, N-1$. When $j = 0$, we get

$$(\mathcal{F}(\mathbb{1}-\boldsymbol{x}))_0 = N - (\mathcal{F}(\boldsymbol{x}))_0 = N - (\mathcal{F}(\boldsymbol{y}))_0 = (\mathcal{F}(\mathbb{1}-\boldsymbol{y}))_0.$$

Therefore, $|\mathcal{F}(\mathbb{1}-\boldsymbol{x})| = |\mathcal{F}(\mathbb{1}-\boldsymbol{y})|$. Similar analysis for the other direction. □

Proof of Proposition 4.7. For a binary signal $\boldsymbol{x} \in \{0,1\}^N$, $(\mathrm{Aut}_p(\boldsymbol{x}))_k$ is the number of pairs of 1's with distance k. As a result, when $\|\boldsymbol{x}\|_0$ is either too small or too large, the uniqueness can be guaranteed thanks to the combinatorial nature of $\mathrm{Aut}_p(\boldsymbol{x})$.

When $\|\boldsymbol{x}\|_0 = 0$, $\boldsymbol{x}$ is the zero vector and hence the recovery is unique.

When $\|\boldsymbol{x}\|_0 = 1$, we get $\boldsymbol{x} = \boldsymbol{e}_k$ for some k, which is related by spatial shifts to each other. Therefore, the recovery is unique up to trivial ambiguities.

When $\|\boldsymbol{x}\|_0 = 2$, we obtain the $\mathrm{Aut}_p(\boldsymbol{x})$ from $|\mathcal{F}\boldsymbol{x}|$ by Theorem 3.5. Without loss of generality, up to spatial shift, we assume $x_0 = 1$. Let k be the smallest positive number such that $(\mathrm{Aut}_p(\boldsymbol{x}))_k$ is nonzero. Since $(\mathrm{Aut}_p(\boldsymbol{x}))_k$ is equal to the number of pairs of 1's with distance k and there are only two 1's in $\boldsymbol{x}$, i.e. only one pair of 1's. This pair must contain x_0. Say the pair contains x_0 and x_j. We know that x_j and x_0 has distance k. Hence, $j = k$ or $N-k$, i.e. we either have $x_0 = x_k = 1$ or $x_0 = x_{N-k} = 1$, which are spatial shifts of each other.

When $\|\boldsymbol{x}\|_0 = 3$, given $|\mathcal{F}\boldsymbol{x}|$, we obtain $\mathrm{Aut}_p(\boldsymbol{x})$. Let k be the smallest positive number such that $(\mathrm{Aut}_p(\boldsymbol{x}))_k$ is nonzero. Since there are three 1's in $\boldsymbol{x}$, there are ${}_3C_2$, i.e. 3 pairs of 1's in $\boldsymbol{x}$. Thus, $(\mathrm{Aut}_p(\boldsymbol{x}))_k = 1, 2$ or 3. By spatial shift, we may assume one of the pairs contains x_0 and x_k.

If $(\mathrm{Aut}_p(\boldsymbol{x}))_k = 2$ or 3, then there is still at least one pair of 1's containing x_0 or x_k and the remaining 1. If it contains x_0, then the 1 should lie in x_{N-K} since x_k is already occupied. If the pair contains x_k, by similar reasoning, the 1 should lie in x_{2k}. In both cases, all three 1's are placed and these 2 cases are spatial shift of each other.

If $(\mathrm{Aut}_p(\boldsymbol{x}))_k = 1$, let l be the smallest positive number greater than k such that $(\mathrm{Aut}_p(\boldsymbol{x}))_l$ is nonzero. By considering the position of 1, we have 4 cases: $x_{N-l} = 1$, $x_{N-l+k} = 1$, $x_l = 1$ or $x_{l+k} = 1$. The cases that $x_{N-l+k} = 1$ and $x_l = 1$ are impossible, otherwise it will contradicts the minimality of l, k and the fact that $(\mathrm{Aut}_p(\boldsymbol{x}))_k = 1$, i.e. there is a pair of 1 with distance $(l-k) < l$ while this pair is not the pair corresponding to the pair of distance k. Hence, we either have $x_0 = x_k = x_{l+k} = 1$ or $x_0 = x_k = x_{N-l} = 1$. Note that these two cases are equivalent to each other through conjugate inverse and spatial shift.

The cases when $\|\boldsymbol{x}\|_0 = N-3, N-2, N-1$ or N now follow from above. If $\|\boldsymbol{x}\|_0 = N-3, N-2, N-1$ or N, then $\|\mathbb{1}-\boldsymbol{x}\|_0 = 0, 1, 2$ or 3. Hence, we can recover $(\mathbb{1}-\boldsymbol{x})$ up to trivial ambiguities. Since $\boldsymbol{x} = \mathbb{1} - (\mathbb{1}-\boldsymbol{x})$, the recovery of $\boldsymbol{x}$ is unique up to trivial ambiguities. □

To prove Propositions 4.9 and 4.10, we need to introduce some results in algebra. Specifically, Theorem C.3 and C.4 are summarized from the proof of [31, Theorem 2.1].

Theorem C.1 (Theorem 1 in [49])**.** *Let $\boldsymbol{x} \in \{0,1\}^N$ with $x_0 = x_{N-1} = 1$, the Z-transform of $\boldsymbol{x}$ is irreducible with probability at least $c/\log N$ for some constant $c > 0$*

Theorem C.2. *[50] If a $f(x)$ is 0, 1 reciprocal polynomial and its constant term is* 1*, then $f(x)$ is not divisible by a non-reciprocal polynomial in $\mathbb{Z}[x]$*

Theorem C.3. *Given $\boldsymbol{x} \in \{0,1\}^N$ and $Aut(\boldsymbol{x})$, if $P_{\boldsymbol{x}}(z)$ is reciprocal, then there does not exist $\boldsymbol{y} \in \{0,1\}^N$ such that $\boldsymbol{y} \neq \boldsymbol{x}$ and $Aut(\boldsymbol{y}) = Aut(\boldsymbol{x})$.*

Proof. Define $A_{\boldsymbol{x}}$ by equation A.3. Then $A_{\boldsymbol{x}} = P_{\boldsymbol{x}}\tilde{P}_{\boldsymbol{x}}$. Write $P_{\boldsymbol{x}} = f_1 f_2 \ldots f_k$ be its factorization such that each f_j is irreducible. It follows from Theorem C.2 that each f_j is also reciprocal. If there exists some $\boldsymbol{y} \in \{0,1\}^N$ such that $\mathrm{Aut}(\boldsymbol{y}) = \mathrm{Aut}(\boldsymbol{x})$. Then $P_{\boldsymbol{y}}\tilde{P}_{\boldsymbol{y}} = A_{\boldsymbol{y}} = A_{\boldsymbol{x}} = P_{\boldsymbol{x}}\tilde{P}_{\boldsymbol{x}} = f_1^2 f_2^2 \ldots f_k^2$. If f_j divides $P_{\boldsymbol{y}}$, we also have $\tilde{f}_j = f_j$ divide $\tilde{P}_{\boldsymbol{y}}$, and vice versa. So $(P_{\boldsymbol{y}}/f_1)(\tilde{P}_{\boldsymbol{y}}/f_1) = f_2^2 f_3^2 \ldots f_k^2$. Inductively, we have $P_{\boldsymbol{y}} = \tilde{P}_{\boldsymbol{y}} = f_1 f_2 \ldots f_k = P_{\boldsymbol{x}}$, i.e. $\boldsymbol{x} = \boldsymbol{y}$. □

Theorem C.4. *Given $\boldsymbol{x} \in \{0,1\}^N$, $Aut(\boldsymbol{x})$, if $P_{\boldsymbol{x}}(z)$ is irreducible, then the only $\boldsymbol{y} \in \{0,1\}^N$ satisfying $Aut(\boldsymbol{y}) = Aut(\boldsymbol{x})$ is either $\boldsymbol{x}$ or $\boldsymbol{z}$, which is defined by $z_n = x_{N-1-n}$. for $n = 0, 1, \ldots, N-1$.*

Proof. Similar to the above, we have $P_{\boldsymbol{y}}\tilde{P}_{\boldsymbol{y}} = P_{\boldsymbol{x}}\tilde{P}_{\boldsymbol{x}}$. Since $P_{\boldsymbol{x}}$ is irreducible, we have either $P_{\boldsymbol{x}}$ divides $P_{\boldsymbol{y}}$ or $\tilde{P}_{\boldsymbol{y}}$. Suppose the first case. We have $\tilde{P}_{\boldsymbol{x}}$ divides $\tilde{P}_{\boldsymbol{y}}$. Together with the fact that $P_{\boldsymbol{x}}$ divides $P_{\boldsymbol{y}}$, this implies $P_{\boldsymbol{x}} = P_{\boldsymbol{y}}$ and $\tilde{P}_{\boldsymbol{x}} = \tilde{P}_{\boldsymbol{y}}$, i,e, $\boldsymbol{y} = \boldsymbol{x}$. Similarly, the second case implies that $P_{\boldsymbol{y}} = \tilde{P}_{\boldsymbol{x}}$, i,e, $\boldsymbol{y}$ is the conjugate inverse of $\boldsymbol{x}$.

□

Proof of Proposition 4.9. By Theorem 3.6, $\mathrm{Aut}(\boldsymbol{x})$ is uniquely determined when $M \geq 2N-1$. Note that $P_{\boldsymbol{x}}(z)$, the Z-transform of $\boldsymbol{x}$, is a reciprocal polynomial since $\boldsymbol{x}$ is equal to its conjugate inverse. According to Theorem C.3, there does not exist $\boldsymbol{y} \neq \boldsymbol{x}$ such that $\mathrm{Aut}(\boldsymbol{y}) = \mathrm{Aut}(\boldsymbol{x})$. Therefore, we can uniquely recover $\boldsymbol{x}$ from $\mathrm{Aut}(\boldsymbol{x})$ up to trivial ambiguities. □

Proof of Propsotion 4.10. Theorem C.1 shows that for a random binary $\boldsymbol{x}$ with $x_0 = x_{N-1} = 1$, the Z-transform of $\boldsymbol{x}$ ($P_{\boldsymbol{x}}(z)$) is irreducible with probability at least $\frac{c'}{logN}$ for a constant $c' > 0$. Note that we have 2^{N-2} binary signals under the constraint $x_0 = x_{N-1} = 1$ while we have 2^N binary signals in total. For a random binary $\boldsymbol{x}$, $P_{\boldsymbol{x}}(z)$ is irreducible with probability at least $\frac{1}{4}\frac{c'}{logN} = \frac{c}{logN}$, where $c = \frac{c'}{4} > 0$ is a fixed constant. The remaining now follows from Theorem C.4 directly. □

Appendix D. Proof of Theorems 4.11–4.14

Proof of Theorem 4.11. Write $|\mathcal{F}_{N\to M}\boldsymbol{x}| = |\mathcal{F}_{M\to M}\tilde{\boldsymbol{x}}|$ with

$$\tilde{\boldsymbol{x}} = (x_0, x_1, \ldots, x_{N-1}, 0, 0, \ldots, 0)^T \in \{0,1\}^M.$$

Similarly, we write $|\mathcal{F}_{N\to M}\boldsymbol{y}| = |\mathcal{F}_{M\to M}\tilde{\boldsymbol{y}}|$ and define $\tilde{\boldsymbol{y}}$. Note that $\tilde{\boldsymbol{x}} \in \{0,1\}^M$, $\tilde{\boldsymbol{y}} \in [0,1]^M$ and $|\mathcal{F}_{M\to M}\tilde{\boldsymbol{x}}| = |\mathcal{F}_{M\to M}\tilde{\boldsymbol{y}}|$. By Theorem 4.1 with $\alpha = 0$ and $\beta = 1$, one has $\tilde{\boldsymbol{y}} \in \{0,1\}^M$ and $\|\tilde{\boldsymbol{y}}\|_0 = \|\tilde{\boldsymbol{x}}\|_0$. Since $\tilde{\boldsymbol{x}}$ and $\tilde{\boldsymbol{y}}$ are obtained by appending zeros to $\boldsymbol{x}$ and $\boldsymbol{y}$, we have $\boldsymbol{y} \in \{0,1\}^N$ and $\|\boldsymbol{y}\|_0 = \|\tilde{\boldsymbol{y}}\|_0 = \|\tilde{\boldsymbol{x}}\|_0 = \|\boldsymbol{x}\|_0$. □

Proof of Theorem 4.12. Without loss of generality, we may assume the windows $\boldsymbol{w}$ is an all one vector $\mathbb{1}$ by scaling. Recall the STFT of $\boldsymbol{x}$ is defined by

$$\boldsymbol{z}_{n,m} = \sum_{k=0}^{N-1} x_k w_{mL-k} e^{\frac{-2\pi kni}{N}}.$$

Since $W \geq L$, for each $l = 0, 1, \ldots, N-1$, there is some m such that $w_{mL-l} = 1$. For such m, define $\tilde{x}_k = x_k w_{mL-k}$ for all $k = 0, 1, \ldots, N-1$ and define $\tilde{\boldsymbol{y}}$ in a similar way. Then, $\tilde{\boldsymbol{x}} \in \{0,1\}^N$ and $\tilde{\boldsymbol{y}} \in [0,1]^N$ by our assumption on $\boldsymbol{w}$.

Now, $|\mathcal{F}\tilde{\boldsymbol{x}}| = \boldsymbol{z}_{\cdot,m} = |\mathcal{F}\tilde{\boldsymbol{y}}|$. Applying Theorem 4.1 with $\alpha = 0$ and $\beta = 1$, we have $\tilde{\boldsymbol{y}} \in \{0,1\}^N$. In particular, $y_l = y_l w_{mL-l} = \tilde{y}_l \in \{0,1\}$. Since l is arbitrary, we have $\boldsymbol{y} \in \{0,1\}^N$. □

Proof of Theorem 4.13. Denote $|\hat{z}_{k,m}|^2$ and $|\hat{w}_{k,m}|^2$ be the FROG trace (2.6) of $\boldsymbol{x}$ and $\boldsymbol{y}$, respectively. We consider $m = 0$ and define $\boldsymbol{z}_0 = (z_{0,0}, z_{1,0}, \ldots, z_{N-1,0})^T$ and similarly for $\boldsymbol{w}_0$. As $x_n \in \{0,1\}$, we obtain $z_{n,0} = x_n^2 = x_n$ and $w_{n,0} = y_n^2 \in [0,1]$. Now, our assumption translates to $|\hat{z}_{k,0}| = |\hat{w}_{k,0}|$ for $k = 0, \ldots, N-1$, i.e. $|\mathcal{F}\boldsymbol{z}_0| = |\mathcal{F}\boldsymbol{w}_0|$. Since $\boldsymbol{z}_0 \in \{0,1\}^N$ and $\boldsymbol{w}_0 \in [0,1]^N$, we have $\boldsymbol{w}_0 \in \{0,1\}^N$ by Theorem 4.1 with $\alpha = 0$ and $\beta = 1$, i.e. $w_{n,0} = y_n^2 \in \{0,1\}$ for all $n = 0, \ldots, N-1$. Therefore, we obtain $\boldsymbol{y} \in \{0,1\}^N$, which implies that $\boldsymbol{y} = \boldsymbol{w}_0$. Since $\boldsymbol{x} = \boldsymbol{z}_0$, we have $|\mathcal{F}\boldsymbol{x}| = |\mathcal{F}\boldsymbol{z}_0| = |\mathcal{F}\boldsymbol{w}_0| = |\mathcal{F}\boldsymbol{y}|$ and $\|\boldsymbol{y}\|_0 = \|\boldsymbol{x}\|_0$ by Theorem 4.1. □

Proof of Theorem 4.14. The proof is similar to the proof of Theorem 4.13 by noting that $z_{n,0} = x_n^2 = 1 \in \{-1,1\}$ and using Theorem 4.3. □

Appendix E. Proof of Proposition 5.1 to Corollary 5.3

Proof of Proposition 5.1.

$$\begin{aligned}
&\|\mathcal{F}^{-1}(\tilde{\boldsymbol{b}} \odot \tilde{\boldsymbol{b}}) - \mathrm{Aut}_p(\boldsymbol{x})\|_\infty = \|\mathcal{F}^{-1}(\tilde{\boldsymbol{b}} \odot \tilde{\boldsymbol{b}}) - \mathcal{F}^{-1}(\boldsymbol{b} \odot \boldsymbol{b})\|_\infty \\
\leq &\|\mathcal{F}^{-1}(\tilde{\boldsymbol{b}} \odot \tilde{\boldsymbol{b}}) - \mathcal{F}^{-1}(\boldsymbol{b} \odot \boldsymbol{b})\|_2 = \frac{1}{\sqrt{N}}\|\tilde{\boldsymbol{b}} \odot \tilde{\boldsymbol{b}} - \boldsymbol{b} \odot \boldsymbol{b}\|_2 \\
\leq &\|\tilde{\boldsymbol{b}} \odot \tilde{\boldsymbol{b}} - \boldsymbol{b} \odot \boldsymbol{b}\|_\infty = \|2\boldsymbol{b} \odot \boldsymbol{\eta} + \boldsymbol{\eta} \odot \boldsymbol{\eta}\|_\infty \\
\leq &2\|\boldsymbol{b}\|_\infty \|\boldsymbol{\eta}\|_\infty + \|\boldsymbol{\eta}\|_\infty^2 < \frac{\epsilon}{2} + \frac{\epsilon}{2} \leq \epsilon,
\end{aligned}$$

where the first and second inequalities come from the fact that $\|\boldsymbol{x}\|_\infty \leq \|\boldsymbol{x}\|_2 \leq \sqrt{N}\|\boldsymbol{x}\|_\infty$ for all $\boldsymbol{x} \in \mathbb{C}^N$ and the third inequality comes from the fact that $\|\boldsymbol{x} \odot \boldsymbol{y}\|_\infty \leq \|\boldsymbol{x}\|_\infty \|\boldsymbol{y}\|_\infty$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^N$. □

Proof of Proposition 5.2. Note that

$$b_n = \left| \sum_{k=0}^{N-1} x_k e^{\frac{-2\pi kni}{N}} \right| \leq \sum_{k=0}^{N-1} |x_k| = \|\boldsymbol{x}\|_1 = \|\boldsymbol{x}\|_0.$$

So $\|\boldsymbol{b}\|_\infty \leq \|\boldsymbol{x}\|_0$.

Let $\epsilon = 1/2$, we have $\|\boldsymbol{\eta}\|_\infty < \frac{1}{8\|\boldsymbol{x}\|_0} \leq \frac{1}{8\|\boldsymbol{b}\|_\infty} = \frac{\epsilon}{4\|\boldsymbol{b}\|_\infty}$. Also, since $\boldsymbol{x} \neq \mathbf{0}$, $\|\boldsymbol{x}\| \geq 1$. $\|\boldsymbol{\eta}\|_\infty < \frac{1}{8\|\boldsymbol{x}\|_0} \leq \frac{1}{8} \leq \min\{\frac{\epsilon}{2}, 1\}$. The remaining follows from Proposition 5.1. □

Proof of Corollary 5.3. When $\boldsymbol{x} = \mathbf{0}$, then $\boldsymbol{b} = |\mathcal{F}\boldsymbol{x}| = \mathbf{0}$, $\mathrm{Aut}_p(\boldsymbol{x}) = \mathbf{0}$ and $\tilde{\boldsymbol{b}} = \boldsymbol{b} + \boldsymbol{\eta} = \boldsymbol{\eta}$.

$$\begin{aligned} &\|\mathcal{F}^{-1}(\tilde{\boldsymbol{b}} \odot \tilde{\boldsymbol{b}}) - \mathrm{Aut}_p(\boldsymbol{x})\|_\infty = \|\mathcal{F}^{-1}(\boldsymbol{\eta} \odot \boldsymbol{\eta})\|_\infty \leq \|\mathcal{F}^{-1}(\boldsymbol{\eta} \odot \boldsymbol{\eta})\|_2 \\ =&\frac{1}{\sqrt{N}}\|\boldsymbol{\eta} \odot \boldsymbol{\eta}\|_2 \leq \|\boldsymbol{\eta} \odot \boldsymbol{\eta}\|_\infty \leq \|\boldsymbol{\eta}\|_\infty^2 < \frac{1}{64N^2} < \frac{1}{2}. \end{aligned}$$

When $\boldsymbol{x} \neq \mathbf{0}$, note $\|\boldsymbol{x}\|_0 \leq N$ and $\|\boldsymbol{\eta}\|_\infty < \frac{1}{8N} \leq \frac{1}{8\|\boldsymbol{x}\|_0}$. The rest is straightforward from Proposition 5.2. □

Proof of Corollary 5.4. The inequality

$$\mathrm{SNR}_{\mathrm{dB}} > 10\log_{10}(64) + 30\log_{10}\|\boldsymbol{x}\|_0,$$

is equivalent to $\frac{\|\boldsymbol{x}\|_2^2}{\|\boldsymbol{\eta}\|_2^2} > 64\|\boldsymbol{x}\|_0^3$. Since $\boldsymbol{x} \in \{0,1\}^N$, $\|\boldsymbol{x}\|_2^2 = \|\boldsymbol{x}\|_0$. Thus, we have $\|\boldsymbol{\eta}\|_\infty^2 \leq \|\boldsymbol{\eta}\|_2^2 < \frac{1}{64\|\boldsymbol{x}\|_0^2}$. □

References

[1] Millane RP. Phase retrieval in crystallography and optics. J Opt Soc Am A. 1990;7(3):394–411.

[2] Kim W, Hayes MH. The phase retrieval problem in X-ray crystallography. In: International Conference on Acoustics, Speech, and Signal Processing; 1991. p. 1765–1768 vol.3.

[3] Miao J, Charalambous P, Kirz J, Sayre D. Extending the methodology of X-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens. Nature. 1999;400(6742):342.

[4] Dainty JC, Fienup JR. Phase Retrieval and Image Reconstruction for Astronomy. Image Recovery: Theory and Application, ed by H Stark, Academic Press. 1987;p. 231–275.

[5] Seifert B, Stolz H, Donatelli M, Langemann D, Tasche M. Multilevel Gauss–Newton methods for phase retrieval problems. J Phys A: Math Gen. 2006 Mar;39(16):4191–4206.

[6] Shechtman Y, Eldar YC, Cohen O, Chapman HN, Miao J, Segev M. Phase Retrieval with Application to Optical Imaging: A contemporary overview. IEEE Signal Process Mag. 2015 May;32(3):87–109.

[7] Hofstetter EM. Construction of time-limited functions with specified autocorrelation functions. IEEE Trans Inf Theory. 1964;10(2):119–126.

[8] Beinert R, Plonka G. Ambiguities in One-Dimensional Discrete Phase Retrieval from Fourier Magnitudes. J Fourier Anal Appl. 2015 Dec;21(6):1169–1198.

[9] Bendory T, Beinert R, Eldar YC. Fourier Phase Retrieval: Uniqueness and Algorithms. In: Boche H, Caire G, Calderbank R, März M, Kutyniok G, Mathar R, editors. Compressed Sensing and its Applications. Springer; 2017. p. 55–91.

[10] Hayes MH, McClellan JH. Reducible polynomials in more than one variable. Proceedings of the IEEE. 1982;70(2):197–198.

[11] Balan R, Casazza P, Edidin D. On signal reconstruction without phase. Appl Comput Harmon A. 2006;20(3):345–356.

[12] Conca A, Edidin D, Hering M, Vinzant C. An algebraic characterization of injectivity in phase retrieval. Appl Comput Harmon A. 2015;38(2):346–356.

[13] Fickus M, Mixon DG, Nelson AA, Wang Y. Phase retrieval from very few measurements. Linear Algebra Its Appl. 2014;449:475–499.

[14] Huang K, Eldar YC, Sidiropoulos ND. Phase Retrieval from 1D Fourier Measurements: Convexity, Uniqueness, and Algorithms. IEEE Trans Signal Process. 2016 Dec;64(23):6105–6117.

[15] Jaganathan K, Oymak S, Hassibi B. Sparse phase retrieval: Convex algorithms and limitations. In: IEEE International Symposium on Information Theory; 2013. p. 1022–1026.

[16] Ranieri J, Chebira A, Lu YM, Vetterli M. Phase Retrieval for Sparse Signals: Uniqueness Conditions. arXiv preprint arXiv:13083058. 2013;.

[17] Li X, Voroninski V. Sparse Signal Recovery from Quadratic Measurements via Convex Programming. SIAM J Numer Anal. 2013;45(5):3019–3033.

[18] Ohlsson H, Eldar YC. On conditions for uniqueness in sparse phase retrieval. In: IEEE Int Conf Acoust. Speech and Signal Process; 2014. p. 1841–1845.

[19] Wang Y, Xu Z. Phase retrieval for sparse signals. Appl Comput Harmon A. 2014;37(3):531–544.

[20] Chang H, Lou Y, Ng M, Zeng T. Phase Retrieval from Incomplete Magnitude Information via Total Variation Regularization. SIAM J Sci Comput. 2016;38:A3672–A3695.

[21] Gerchberg RW, Saxton WO. A practical algorithm for the determination of the phase from image and diffraction plane pictures. Optik. 1972;35(2):237–246.

[22] Candès EJ, Li X, Soltanolkotabi M. Phase retrieval via Wirtinger flow: Theory and algorithms. IEEE Trans Inf Theory. 2015;61(4):1985–2007.

[23] Chang H, Marchesini S, Lou Y, Zeng T. Variational Phase Retrieval with Globally Convergent Preconditioned Proximal Algorithm. SIAM J Imaging Sci. 2018;11(1):56–93.

[24] Keiper S, Kutyniok G, Lee DG, Pfander GE. Compressed sensing for finite-valued signals. Linear Algebra Its Appl. 2017;532:570 – 613.

[25] Esedoglu S. Blind deconvolution of bar code signals. Inverse Probl. 2003 Nov;20(1):121–135.

[26] Lou Y, Esser E, Zhao H, Xin J. Partially blind deblurring of barcode from out-of-focus blur. SIAM J Imaging Sci. 2014;7(2):740–760.

[27] Litman A, Lesselier D, Santosa F. Reconstruction of a two-dimensional binary obstacle by controlled evolution of a level-set. Inverse Probl. 1998;14(3):685–706.

[28] Shi X, Fischer P, Neu V, Elefant D, Lee J, Shapiro D, et al. Soft x-ray ptychography studies of nanoscale magnetic and structural correlations in thin SmCo5 films. Appl Phys Lett. 2016;108(9):094103.

[29] Stein A, Wright G, Yager KG, Doerk GS, Black CT. Selective directed self-assembly of coexisting morphologies using block copolymer blends. Nature communications. 2016;7:12366.

[30] Hart R, Chang H, Lou Y. Empirical studies on phase retrieval. In: IEEE Image, Video, and Multidimensional Signal Processing (IVMSP) Workshop; 2018. p. 1–5.

[31] Yuan Z, Wang H. Phase retrieval for sparse binary signal: uniqueness and algorithm. Inverse Probl Sci Eng. 2018;26(5):641–656.

[32] Trebino R, DeLong KW, Fittinghoff DN, Sweetser JN, Krumbgel MA, Richman BA, et al. Measuring ultrashort laser pulses in the time-frequency domain using frequency-resolved optical gating. Rev Sci Instrum. 1997;68(9):3277–3295.

[33] Trebino R. Frequency-resolved optical gating: the measurement of ultrashort laser pulses. Springer Science & Business Media; 2012.

[34] Jaganathan K, Eldar YC, Hassibi B. STFT Phase Retrieval: Uniqueness Guarantees and Recovery Algorithms. IEEE J Sel Topics Signal Process. 2016;10(4):770–781.

[35] Candès EJ, Li X, Soltanolkotabi M. Phase retrieval from coded diffraction patterns. Appl Comput Harmon A. 2015;39(2):277–299.

[36] Chang H, Enfedaque P, Lou Y, Marchesini S. Partially coherent ptychography by gradient decomposition of the probe. Acta Cryst A. 2018 May;74(3):157–169.

[37] Bendory T, Edidin D, Eldar YC. On signal reconstruction from FROG measurements. Appl Comput Harmon A. 2018;.

[38] Hayes MH. The reconstruction of a multidimensional sequence from the phase or magnitude of its Fourier transform. IEEE Trans Acoust Speech Signal Process. 1982;30:140–154.

[39] Miao J, Kirz J, Sayre D. The oversampling phasing method. Acta Cryst D. 2000;56(10):1312–1315.

[40] Soldovieri F, Liseno A, D'Elia G, Pierri R. Global convergence of phase retrieval by quadratic approach. IEEE transactions on antennas and propagation. 2005;53(10):3135–3141.

[41] Eldar YC, Sidorenko P, Mixon DG, Barel S, Cohen O. Sparse phase retrieval from short-time Fourier measurements. IEEE Signal Process Letters. 2014;22(5):638–642.

[42] Mao Y. Reconstruction of binary functions and shapes from incomplete frequency information. IEEE Trans Inf Theory. 2012;58(6):3642–3653.

[43] Pham-Dinh T, Le-Thi HA. Convex Analysis Approach to D.C. Programming: Theory, Algorithms and Applications. Acta Math Vietnam. 1997;22(1):289–355.

[44] Pham-Dinh T, Le-Thi HA. The DC (difference of convex functions) programming and DCA revisited with DC models of real world nonconvex optimization problems. Ann Oper Res. 2005;133(1-4):23–46.

[45] Odlyzko AM, Poonen B. Zeros of polynomials with 0, 1 coefficients. Enseign Math. 1993;39:317–348.

[46] Boyd S, Parikh N, Chu E, Peleato B, Eckstein J. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. Found Trends Mach Learn. 2011;3(1):1–122.

[47] Beinert R. Non-negativity constraints in the one-dimensional discrete-time phase retrieval problem. Inf Inference A J IMA. 2016 Dec;6(2):213–224.

[48] Rudin W. Functional analysis. McGraw-Hill; 1991.

[49] Konyagin SV. On the number of irreducible polynomials with 0,1 coefficients. Acta Arith. 1999;88(4):333–350.

[50] Filaseta M, Meade DB. Irreducibility testing of lacunary 0, 1-polynomials. Journal of Algorithms. 2005;55(1):21–28.