Sparse Phase Retrieval with Near Minimal Measurements: A Structured Sampling Based Approach

{ Heng Qiao and Piya Pal } Department of Electrical and Computer Engineering, University of Maryland, College Park

Introduction

Phaseless Measurement:

$$y_i = |\langle \mathbf{x}, \mathbf{f}_i \rangle|^2 + n_i \quad 0 \le i \le M-1$$

- $\mathbf{x} \in \mathbb{C}^N$ unknown data
- $\mathbf{f}_i \in \mathbb{C}^N$ measurement vector
- n_i denotes additive noise

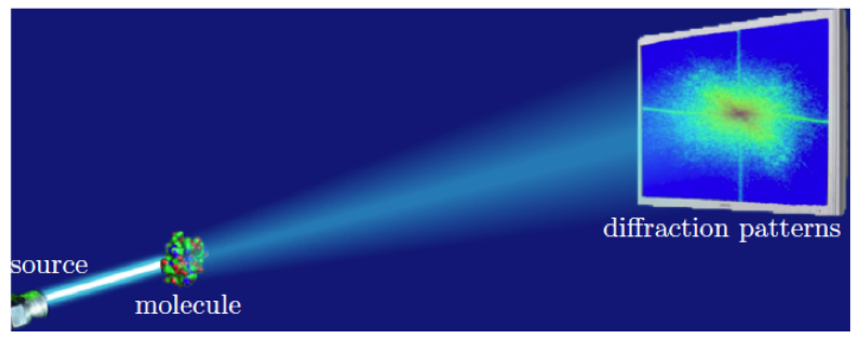


Figure 1: [Adopted from Candès, Li and Soltanolkotabi, 2014]

In noiseless case, estimation of $\mathbf{x}$ can be achieved up to $\mathbf{x}^\# = c\mathbf{x}$ with $|c| = 1$.

Applications:

- X-ray crystallography
- Astronomical Imaging
- Electron microscopy
- Computational optical imaging
- Blind deconvolution

Conjecture: $4N-4$ measurements are necessary for exact recovery on $\mathbb{C}^N$.

Partial Nested Fourier Sampler

Acquired Image: $y_i = |\sum_{k=1}^{N} e^{j2\pi ik/N} x_k|$ ⟷ **Equivalent Formulation**: $\text{ACF}_\mathbf{x}(k) = \sum_{n=1}^{N-k} x_n x_{n+k}$

- Recovering $\mathbf{x}$ from $ACF_\mathbf{x}(k)$: Non-unique due to spectral factorization.
- Equivalent representation

$$y_i = \left(\mathbf{f}_i^T \otimes \mathbf{f}_i^H\right) \text{Vec}\left(\mathbf{x}\mathbf{x}^H\right) + n_i$$

Difference sets arise naturally in magnitude-only-Fourier measurements.

$$y_i^2 = \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} x_m x_n^* e^{j2\pi i(m-n)/N}$$

Difference Set

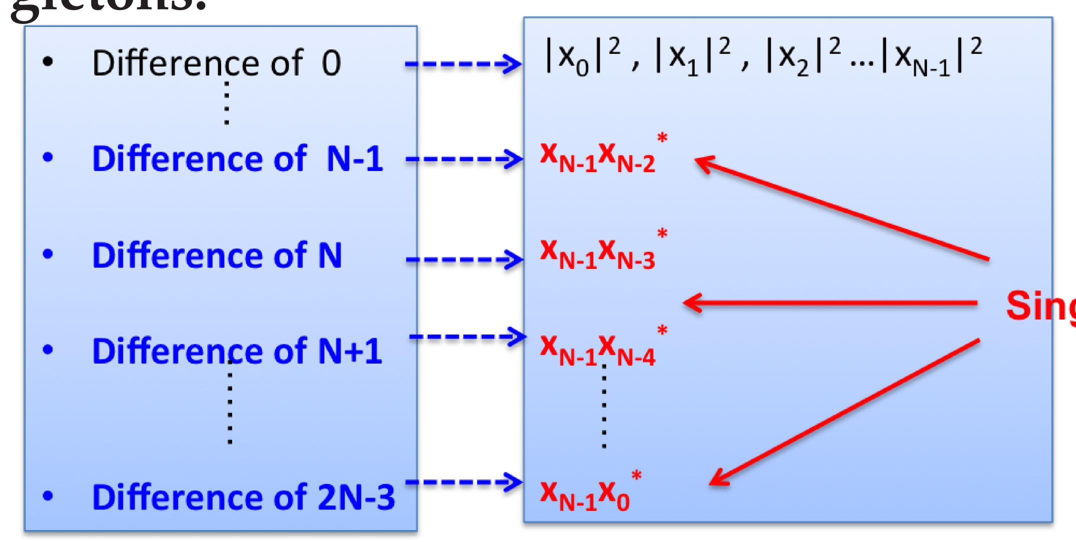
Definition 1 (Partial Nested Fourier Sampler:) *A Partial Nested Fourier Sampler (PNFS) of dimension N, consists of measurement vectors given by*

$$\mathbf{f}_i^{(N)} = \frac{1}{\sqrt[4]{4N-5}}[z_i^1, z_i^2, \cdots z_i^{N-1}, z_i^{2N-2}]^T, \quad (1)$$

where $z_i = e^{j2\pi m_i/4N-5}, m_i \in [0, 4N-6]$.

Index Set of Vectors	Difference Set
1, 2, 3, ..., N-1, 2N-2	0,1,..., N-2, N-1, N, N+1, N+2,...2N-3

PNFS reveals the support of x by obtaining singletons:



Recovery Guarantee for Non-Sparse x: If $n_i \equiv 0$, $M = 4N-5$ PNFS measurements are sufficient to exactly recover non-sparse $\mathbf{x}$.

When $\mathbf{x}$ is sparse and no prior knowledge is available, we use randomized version of PNFS

Definition 2 (Randomized PNFS) *A Randomized PNFS (R-PNFS) consists of measurement vectors*

$$\mathbf{f}_i^{(\text{R-PNFS})} = [\mathbf{I}_{N,N} \quad \mathbf{v}]\,\mathbf{f}_i^{(N+1)}$$

where $\mathbf{v} \in \mathbb{C}^N$ is a random vector with independent entries, and $\mathbf{f}_i^{(N+1)}$ is defined in (1) for dimension $N+1$.

A Cancellation Based Algorithm and Main Result

Algorithm:

1. Collect two sets of (noisy) phaseless measurements $\mathbf{y}^{(1)}, \mathbf{y}^{(2)} \in \mathbb{C}^{\tilde{M}}$ as
$$y_i^{(1)} = \left|\left(\mathbf{f}_i^{\text{R-PNFS}}\right)^H \mathbf{x}^\star\right|^2 + n_i^{(1)}$$
$$y_i^{(2)} = \left|\left(\tilde{\mathbf{f}}_i^{(N+1)}\right)^H \mathbf{x}^\star\right|^2 + n_i^{(2)} \quad (2)$$
Assuming $|n_i^{(k)}| \le \eta, k = 1, 2.$ $M = 2\tilde{M}$.
2. Compute the difference measurements $\mathbf{\Delta y} = \mathbf{y}^{(1)} - \mathbf{y}^{(2)}$. The *key step is to notice that*
$$\mathbf{\Delta y} = \mathbf{Z}\hat{\mathbf{x}} + \mathbf{\Delta n} \quad (3)$$
where $\hat{\mathbf{x}} \in \mathbb{C}^{4N-1}$ given by
$$[\hat{\mathbf{x}}]_m = \begin{cases} |x_{N+1}|^2 & m = 0 \\ 0 & m = 1, 2, \cdots, N-1 \\ x_{2N-m}\bar{x}_{N+1} & m = N, \cdots, 2N-1 \\ \overline{[\hat{\mathbf{x}}]}_{-m} & m < 0 \end{cases}$$
3. Obtain an estimate of $\hat{\mathbf{x}}$ as the solution to the following l_1-minimization problem:
$$\min_\theta \|\theta\|_1 \quad \text{subject to } \|\mathbf{\Delta y} - \mathbf{Z}\theta\|_2 \le \eta\sqrt{\tilde{M}} \quad (\mathbf{P1})$$
4. Given the solution $\hat{\mathbf{x}}^\#$ to (**P1**), the estimate for each entry of $\mathbf{x}^\star$ is given by $x_q^\# = [\hat{\mathbf{x}}^\#]_{2N-q}/x_{N+1}^\#$ for $1 \le q \le N$ and $x_{N+1}^\# = |\sqrt{[\hat{\mathbf{x}}^\#]_0}|$.

Introducing a second sampling vector $\tilde{\mathbf{f}}_i^{(N+1)} \in \mathbb{C}^N$ as

$$\tilde{\mathbf{f}}_i^{(N+1)} = [\mathbf{I}_{N,N} \quad \mathbf{0}]\,\mathbf{f}_i^{(N+1)}$$

The algorithm is summarized on the left.

Theorem 1 *Given $\mathbf{x}^\star \in \mathbb{C}^N$ with sparsity s, and the measurement vector $\mathbf{v} \in \mathbb{C}^N$, using (2) where the indices m_i of $\mathbf{f}_i^{(N+1)}, i = 1, 2\cdots, M$ are chosen uniformly at random from $[0, 4N-2]$. If $\tilde{M} \ge c_0(2s+1)\log(4N-1)\log(\varepsilon^{-1})$ and $|x_{N+1}|^2 > c_1\sqrt{2s+1}\eta$, with probability at least $1-\varepsilon$, the estimates $x_q^\#$ of $x_q^\star, 1 \le q \le N$, satisfy*

$$\sum_{q=1}^{N} |x_q^\star - e^{j\phi_0}x_q^\#| \le \frac{c_1\sqrt{(2s+1)(4N-1)}}{\sqrt{|x_{N+1}|^2 - c_1\sqrt{2s+1}\eta}}\eta + \|\mathbf{x}^\star\|_1 \left(\frac{1}{\sqrt{1 - c_1\frac{\sqrt{2s+1}\eta}{|x_{N+1}|^2}}} - 1\right)$$

where $x_{N+1} = \mathbf{v}^H\mathbf{x}^\star$, $\phi_0 = arg_{\phi\in[0,2\pi)} x_{N+1}/|x_{N+1}|$, and c_0, c_1 are universal constants.

With R-PNFS, $M = O(s\log N)$ is sufficient for stable recovery by implementing cancellation based algorithm.

Simulation Results

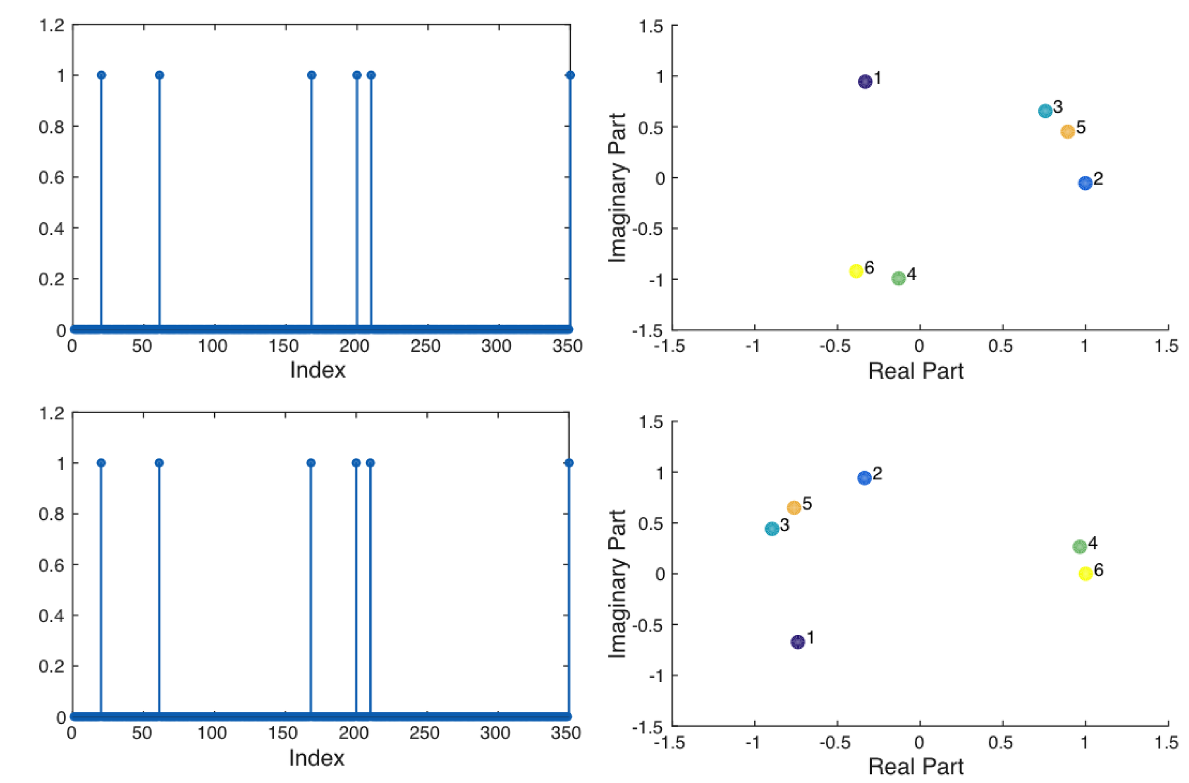


Figure 2: Amplitudes and complex plane representations of the nonzero part of the original and recovered data in noiseless case.

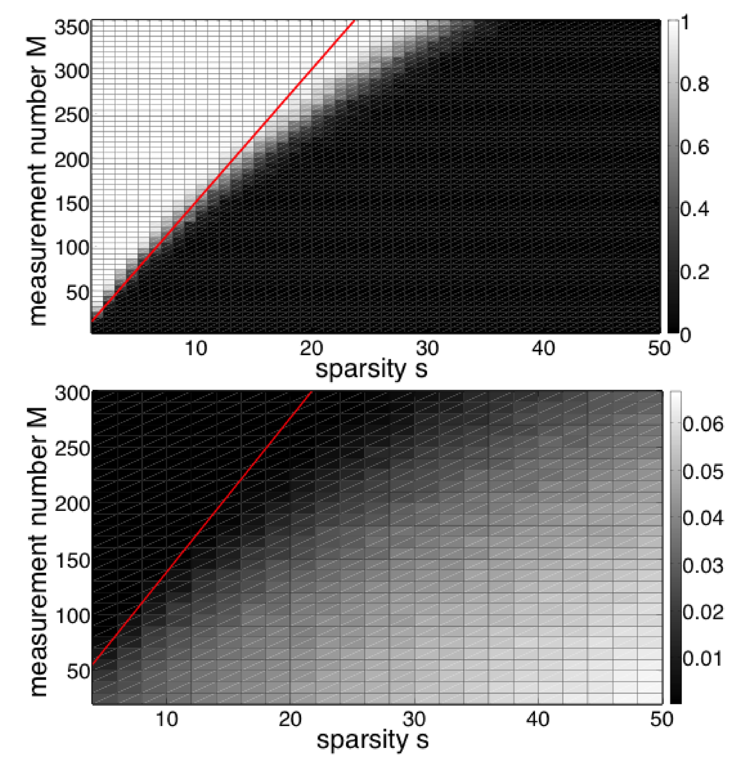


Figure 3: Phase transition plots. (Top) noiseless case. (Bottom) noisy case. The red line represents $M = 3s\log N$ for both.