## The Compressive Phase Retrieval Problem

Let $x \in \mathbb{C}^n$ be a k-sparse signal, with $k \ll n$. Given squared magnitude measurements

$$y = |Mx|^2 + n,$$

where $M \in \mathbb{C}^{m \times n}$ denotes a measurement matrix and $n \in \mathbb{R}^n$ denotes measurement noise, the compressive phase retrieval problem seeks to recover the unknown signal $x$ (up to some global phase offset) using only $m \ll n$ phasor measurements, $y \in \mathbb{R}^m$.

We are interested in measurement constructions $\mathcal{M}$ and associated recovery algorithms $\mathcal{A}_M : \mathbb{R}^m \rightarrow \mathbb{C}^n$ which are efficient, use a minimal number of measurements, and are robust to measurement errors.

The phase retrieval problem occurs in several fields of science such as X-ray crystallography, optics, astronomy and quantum mechanics, where, either due to the underlying physics or instrumentation limitations, we are unable to acquire phase information.

## Ingredients: (I) Fast (Non-Sparse) Phase Retrieval

### Main Result

There exists a deterministic algorithm $\mathcal{A}_M : \mathbb{R}^m \rightarrow \mathbb{C}^n$ for which the following holds. Let $\epsilon \in (0,1]$, $x \in \mathbb{C}^n$ with sufficiently large $n$, and $k \in \{1,2, \ldots, n\} \subseteq \mathbb{N}$. Then, one can select a random measurement matrix $M \in \mathbb{C}^{m \times n}$ such that

$$\min_{x \in \mathbb{C}^n} | \mathbb{E} [y - \mathcal{A}_M(\mathbb{E}[Mx])^2] | \leq \frac{2\| x - x^* \|_{\infty}^2}{\| x - x^* \|_{\infty}^2} \mathbb{V}_C,$$

where $y$ is true with probability at least $1 - \exp\left(-\frac{k^2}{\epsilon^6} n^{-1} \log \left(\frac{n}{\epsilon^2}\right) \cdot \log (\log n) \cdot \log \log n \cdot \log n \right)$.

Furthermore, the algorithm will run in $O \left(\frac{\| x - x^* \|_{\infty}^2}{\| x - x^* \|_{\infty}^2} \mathbb{V}_C \right)$ time.

This is the first sub-linear time compressive phase retrieval algorithm.

Both the sampling and runtime complexities are sub-linear in the problem size and (poly)log-linear in the sparsity.

### Proposed Algorithm

Let $P \in \mathbb{C}^{m \times n}$ denote an admissible phase retrieval matrix associated with the phase retrieval method $\Delta_P$. Let $C \in \mathbb{C}^{m \times n}$ denote a compressive sensing matrix associated with the phase retrieval algorithm $\Delta_C$. Construct the measurement matrix $M$ for the compressive phase retrieval problem as $M = PC$.

Note: Consider the following simple two-stage formulation:

1. Apply a fast phase retrieval method $\Delta_P : \mathbb{R}^m \rightarrow \mathbb{C}^n$, to the phasor measurements $y$ and recover an intermediate compressed signal $z \in \mathbb{C}^n$, where $d = (k \log k \cdot \log n)$.
2. Next, use a sub-linear time compressive sensing algorithm, $\Delta_C : \mathbb{C}^n \rightarrow \mathbb{C}^n$, to recover the unknown signal $x$.

We can show that $\Delta_C \circ \Delta_P : \mathbb{R}^m \rightarrow \mathbb{C}^n$ recovers the unknown signal $x$ up to a global phase factor accurately and stably.

## Ingredients: (II) Sub-Linear Time Compressive Sensing

Choose the measurement matrix $C$ to be a random sparse binary matrix obtained by randomly sub-sampling rows of a well-chosen incoherent matrix (for example, the adjacency matrix of certain unbalanced expander graphs). In [2], it is shown that these matrices satisfy certain combinatorial properties which permit the use of fast compressed sensing recovery algorithms.

The recovery algorithm then proceeds in two phases:

1. Identify the $k$ largest magnitude entries of $x$ using standard bit-testing techniques.
2. Estimate these $k$ largest entries using median estimates and techniques from computer science streaming literature.

The sampling and runtime complexities of this method are both $O(k \cdot \log k \cdot \log n)$.

## Numerical Results

Left panel figure shows execution time as a function of sparsity for various problem dimensions. We observe that the overall execution time is sub-linear in the problem size $n$ and (poly)log-linear in the sparsity $k$.

The right panel figure illustrates robustness of the method to (i.i.d) Gaussian measurement noise. It plots the reconstruction error in dB as a function of the added noise level (in dB) for a length $n = 2^{20}$ signal using less than 10% of measurements.

In both cases, complex sparse test signals with i.i.d complex Gaussian non-zero entries were used, with non-zero index locations chosen by $k$-permutations.

### References and Acknowledgement


This work was supported in part by NSF grant DMS-1445725 and SSS grant H2823-13-1-0273.