Fast and Robust Phase Retrieval

Aditya Viswanathan
aditya@math.msu.edu

MICHIGAN STATE UNIVERSITY

ICOSAHOM 2014, Salt Lake City, UT
26^{th} June 2014

Joint work with

Yang Wang

Mark Iwen

Research supported in part by National Science Foundation grant DMS 1043034.

Outline

The Phase Retrieval Problem

$$\text{find} \quad \mathbf{x} \in \mathbb{C}^d \quad \text{given} \quad |M\mathbf{x}| = \mathbf{b} \in \mathbb{R}^D,$$

where

- $\mathbf{b} \in \mathbb{R}^D$ are the magnitude or intensity measurements.
- $M \in \mathbb{C}^{D \times d}$ is a measurement matrix associated with these measurements.

Let $\mathcal{A} : \mathbb{R}^D \to \mathbb{C}^d$ denote the recovery method.

The phase retrieval problem involves designing measurement matrix and recovery method pairs.

Applications of Phase Retrieval

Important applications of Phase Retrieval

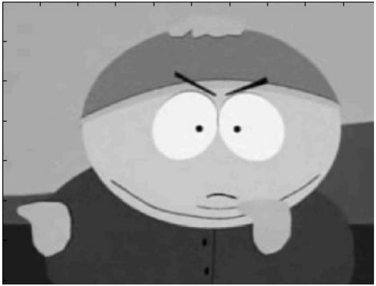
- X-ray crystallography
- Diffraction imaging
- Transmission Electron Microscopy (TEM)

In many molecular imaging applications, the detectors only capture intensity measurements.

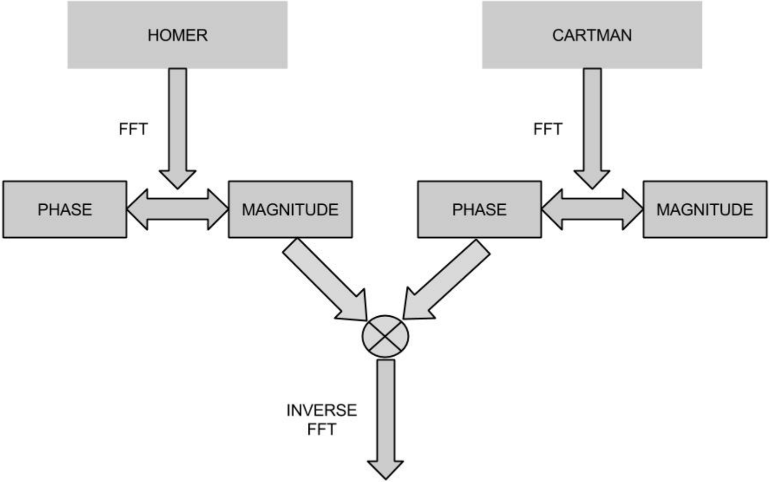
Indeed, the design of such detectors is often significantly simpler than those that capture phase information.

The Importance of Phase – An Illustration

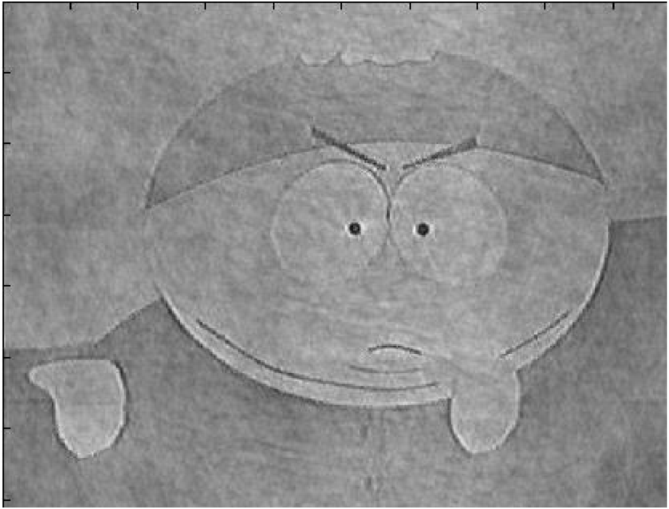
- The phase encapsulates vital information about a signal
- Key features of the signal are retained even if the magnitude is lost



The Importance of Phase – An Illustration



The Importance of Phase – An Illustration



Objectives

- Computational Efficiency – Can the recovery algorithm $\mathcal{A}$ be computed in $O(d)$-time?
- Computational Robustness: The recovery algorithm, $\mathcal{A}$, should be robust to additive measurement errors (i.e., noise).
- Minimal Measurements: The number of linear measurements, D, should be minimized to the greatest extent possible.

Outline

Alternating Projection Methods

[Gerchberg and Saxton, 1972] and [Fienup, 1978]

- These methods operate by alternately projecting the current iterate of the signal estimate over two sets of constraints.
- One of the constraints is the magnitude of the measurements.
- The other constraint depends on the application – positivity, support constraints, . . .

Alternating Projection Methods

[Gerchberg and Saxton, 1972] and [Fienup, 1978]

Issues

- Convergence is slow – the algorithm is likely to stagnate at stationary points
- Requires careful selection of and tuning of the parameters
- Mathematical aspects of the algorithm not well known. If there is proof of convergence, it is only for special cases.

Applications

- Can be used as a post-processing step to speed up more rigorous (but slow) computational approaches

PhaseLift [Candes et. al., 2012]

- Modify the problem to that of finding the rank-1 matrix $X = \mathbf{x}\mathbf{x}^*$
- Uses multiple random illuminations (or masks) as measurements.
- The resulting problem can be cast as a rank minimization optimization problem (`NP hard`)
- Instead, solve a convex relaxation – trace minimization problem (`SDP`)

PhaseLift [Candes et. al., 2012]

Advantages

- Recovery guarantees for random measurements
- Optimization problems of the above type are well-understood
- Mature software for solving the resulting optimization problem.

Disadvantages

- SDP solvers are still slow!
- General-purpose solvers have complexity $\mathcal{O}\left(d^3\right)$; FFT-based measurements may be solved in $\mathcal{O}\left(d^2\right)$ time

Outline

Overview of the Computational Framework

1. Use compactly supported masks and correlation measurements to obtain phase difference estimates.

$$|\mathtt{corr}(\mathbf{w}, \mathbf{x})| \longrightarrow x_j \overline{x}_{j+k}, \quad k = 0, \ldots, \delta$$

- $\mathbf{w}$ is a mask or window function with $\delta + 1$ non-zero entries.
- $x_j \overline{x}_{j+k}$ gives us the (scaled) difference in phase between entries x_j and x_{j+k}.

Constraints on $\mathbf{x}$: We require $\mathbf{x}$ to be non-sparse.
(The number of consecutive zeros in $\mathbf{x}$ should be less than δ)

Overview of the Computational Framework

1. Use compactly supported masks and correlation measurements to obtain phase difference estimates.

$$|\mathtt{corr}(\mathbf{w}, \mathbf{x})| \longrightarrow x_j \overline{x}_{j+k}, \quad k = 0, \ldots, \delta$$

   - $\mathbf{w}$ is a mask or window function with $\delta + 1$ non-zero entries.
   - $x_j \overline{x}_{j+k}$ gives us the (scaled) difference in phase between entries x_j and x_{j+k}.

2. Solve an angular synchronization problem on the phase differences to obtain the unknown signal.

$$x_j \overline{x}_{j+k} \longrightarrow x_j$$

Constraints on $\mathbf{x}$: We require $\mathbf{x}$ to be non-sparse.
(The number of consecutive zeros in $\mathbf{x}$ should be less than δ)

Correlations with Support-Limited Functions

- Let $\mathbf{x} = [x_0 \ x_1 \ \ldots \ x_{d-1}]^T \in \mathbb{C}^d$ be the unknown signal.
- Let $\mathbf{w} = [w_0 \ w_1 \ \ldots \ w_\delta \ 0 \ \ldots 0]^T$ denote a support-limited mask. It has $\delta + 1$ non-zero entries.
- We are given the (squared) correlation measurements

$$(b^m)^2 = |\mathtt{corr}(\mathbf{w}^m, \mathbf{x})|^2, \qquad m = 0, \ldots, L$$

corresponding to $L + 1$ distinct masks.

Correlations with Support-Limited Functions

Explicitly writing out the correlation, we have

$$(b_k^m)^2 = \left| \sum_{j=0}^{\delta} \overline{w}_j^m \cdot x_{k+j} \right|^2$$

$$= \sum_{i,j=0}^{\delta} w_i^m \overline{w}_j^m x_{k+j} \overline{x}_{k+i}$$

$$=: \sum_{i,j=0}^{\delta} w_i \overline{w}_j Z_{k+j,i-j},$$

where $Z_{n,l} := x_n \overline{x}_{n+l}, \quad -\delta \leq l \leq \delta.$

Solving for Phase Differences

Ordering $\{Z_{n,l}\}$ lexicographically, second index first, we obtain a linear system of equations for the phase differences.

Example: $\mathbf{x} \in \mathbb{R}^d, d = 4, \delta = 1$

$$
\begin{pmatrix}
(w_0^0)^2 & 2w_0^0w_1^0 & (w_1^0)^2 & 0 & 0 & 0 & 0 & 0 \\
(w_0^1)^2 & 2w_0^1w_1^1 & (w_1^1)^2 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & (w_0^0)^2 & 2w_0^0w_1^0 & (w_1^0)^2 & 0 & 0 & 0 \\
0 & 0 & (w_0^1)^2 & 2w_0^1w_1^1 & (w_1^1)^2 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & (w_0^0)^2 & 2w_0^0w_1^0 & (w_1^0)^2 & 0 \\
0 & 0 & 0 & 0 & (w_0^1)^2 & 2w_0^1w_1^1 & (w_1^1)^2 & 0 \\
(w_1^0)^2 & 0 & 0 & 0 & 0 & 0 & 2w_0^0w_1^0 & (w_1^0)^2 \\
(w_1^1)^2 & 0 & 0 & 0 & 0 & 0 & 2w_0^1w_1^1 & (w_1^1)^2
\end{pmatrix}
\begin{pmatrix}
Z_{0,0} \\ Z_{0,1} \\ Z_{1,0} \\ Z_{1,1} \\ Z_{2,0} \\ Z_{2,1} \\ Z_{3,0} \\ Z_{3,1}
\end{pmatrix}
=
\begin{pmatrix}
b_0^0 \\ b_0^1 \\ b_1^0 \\ b_1^1 \\ b_2^0 \\ b_2^1 \\ b_3^0 \\ b_3^1
\end{pmatrix}
$$

The system matrix is *block circulant*!

Consequences of the Block Circulant Structure

- There exists a unitary decomposition of the system matrix
- The condition number of the system matrix is a function of the individual blocks. In particular, we have

$$\kappa = \frac{\max\limits_{|t|=1} \ \sigma_1(t)}{\min\limits_{|t|=1} \ \sigma_\delta(t)},$$

  where $\sigma_1(t), \sigma_\delta(t)$ are the largest/smallest singular values of

$$J(t) = A_0 + tA_1 + \cdots + t^{\delta-1} A_\delta$$

- The linear system for the phase differences can be solved efficiently using FFTs

Entries of the Measurement Matrix

Two strategies

- Random entries (Gaussian, Uniform, Bernoulli, . . .)
- Structured measurements

Random Measurements

Representative Condition Numbers

	$\delta = 3$	$\delta = 6$	$\delta = 9$
Critical sampling	34.89	124.05	465.32
Oversampling factor 2	3.73	11.73	13.30
Oversampling factor 3	3.29	6.08	8.60

- Critical sampling: $D = (2\delta + 1)d$, where D denotes the number of measurements.
- Oversampling: $D = \gamma \cdot (2\delta + 1)d$, where γ is the oversampling factor
- Typically, we use $\gamma = 1.5$.

Structured Measurements

Theorem

Choose entries of the measurement matrix, S, as follows:

$$w_i^\ell = \begin{cases} \frac{\mathbb{e}^{-i/a}}{\sqrt[4]{2\delta+1}} \cdot \mathbb{e}^{\frac{2\pi \mathbb{i} \cdot i \cdot \ell}{2\delta+1}}, & i \leq \delta \\ 0, & i > \delta \end{cases} \qquad a \in \left[4, \frac{\delta-1}{2}\right), \quad 0 \leq \ell \leq L.$$

Then,

$$\kappa(S) < \max\left\{144\mathbb{e}^2, \frac{9\mathbb{e}^2}{4} \cdot (\delta-1)^2\right\}.$$

Note:

- w_i^ℓ are scaled entries of a DFT matrix.
- δ is typically chosen to be 6–12.
- No oversampling necessary!

Angular Synchronization

The Angular Synchronization Problem
Estimate n unknown angles $\theta_1, \theta_2, \ldots, \theta_n \in [0, 2\pi)$ from m noisy measurements of their differences $\theta_{ij} := \theta_i - \theta_j \bmod 2\pi$.

- Applications include time synchronization in distributed computer networks and computer vision.
- Problem formulation similar to that of partitioning a weighted graph.
- Problem can be cast as a semidefinite program (SDP) or an Eigenvector problem

A Greedy Algorithm for Angular Synchronization

The unknown phases (modulo a global phase offset) may be obtained by solving a simple greedy algorithm.

1. Set the largest magnitude component to have zero phase angle; i.e.,

$$\angle x_j = 0, \qquad j = \underset{i}{\operatorname{argmax}} Z_{i,i}.$$

   Note: (by definition) $|x_j|^2 = Z_{i,i}, \qquad i = 0, \ldots, d-1.$

2. Use this entry to set the phase angles of the next δ entries; i.e.,

$$\angle x_k = \angle x_j - \angle Z_{j,k}, \qquad k = 1, \ldots, \delta.$$

3. Use the largest magnitude component from these δ entries to repeat the process.

Outline

Numerical Results

- Test signals (Real and Complex) – iid Gaussian, uniform random, sinusoidal signals
- Noise model

$$\tilde{\mathbf{b}} = \mathbf{b} + \tilde{\mathbf{n}}, \quad \tilde{\mathbf{n}} \sim U[0, a].$$

  Value of a determines SNR

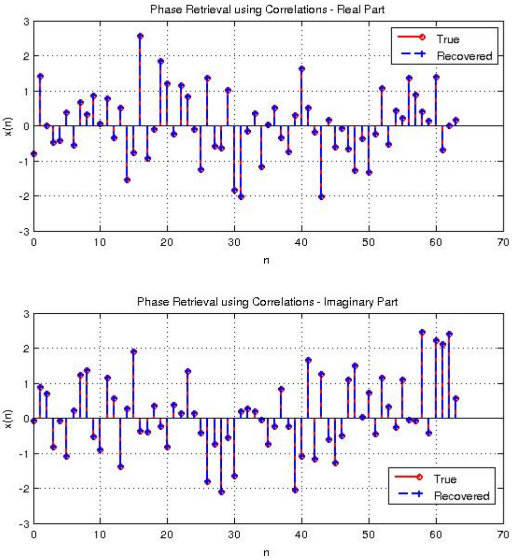
$$SNR = 10\log_{10}\left(\frac{\text{noise power}}{\text{signal power}}\right) = 10\log_{10}\left(\frac{a^2/3}{||b||^2/d}\right)$$

- Errors reported as SNR (dB)

$$\text{Error (dB)} = 10\log_{10}\left(\frac{\|\hat{\mathbf{x}} - \mathbf{x}\|^2}{\|\mathbf{x}\|^2}\right)$$

  ($\hat{\mathbf{x}}$ – recovered signal, $\mathbf{x}$ – true signal)

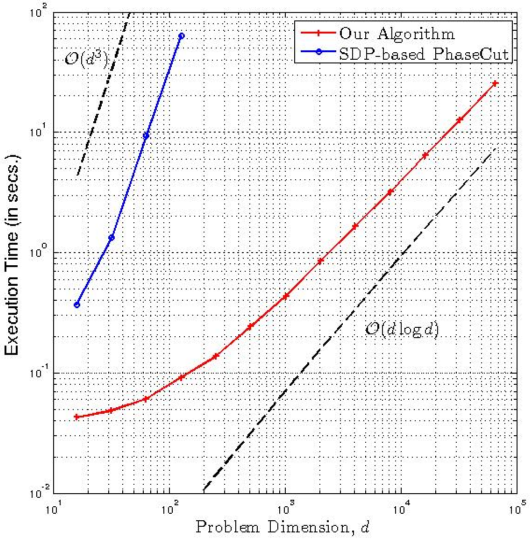
Noiseless Case



- iid Complex Gaussian signal
- $d = 64$
- $\delta = 1$ ($3d$ measurements)
- No noise
- Reconstruction Error

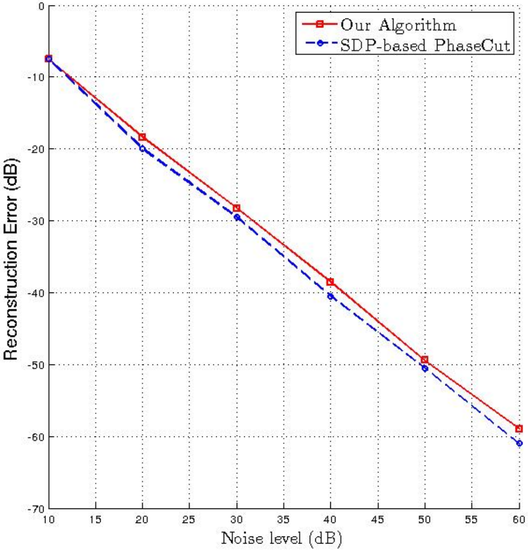
$$\frac{\|\hat{\mathbf{x}} - \mathbf{x}\|^2}{\|\mathbf{x}\|^2} = 6.436 \times 10^{-15}$$

Efficiency



- iid Complex Gaussian signal
- $\delta = 8$ ($17d$ measurements)
- SDP solver – $4d$ measurements
- 40 dB noise

Robustness



- iid complex Gaussian signal
- $d = 64$
- $\delta = 4$ (i.e., $9d$ measurements)
- SDP solver – $4d$ measurrements

Outline

1. The Phase Retrieval Problem
2. Existing Approaches
3. Computational Framework
4. Numerical Results
5. Future Directions

Discussion

(+)

- Well-conditioned measurement matrices with explicit condition number bounds
- Significantly faster (FFT time) than comparable SDP-based methods

(-)

- Requires $2\times$ to $4\times$ more measurements than equivalent SDP-based methods

Future Research Directions

- Writing out a formal recovery guarantee
- Efficient implementations for two dimensional signals.
- Extension to sparse signals – number of required measurements can be reduced.

References

1. R.W. Gerchberg and W.O. Saxton. *A practical algorithm for the determination of phase from image and diffraction plane pictures*. Optik, 35:237–246, 1972
2. J.R. Fienup. *Reconstruction of an object from the modulus of its Fourier transform*. Optics Letters, 3:27–29, 1978.
3. E. J. Candes, T. Strohmer, and V. Voroninski. *Phaselift: exact and stable signal recovery from magnitude measurements via convex programming*. Comm. Pure Appl. Math., 66, 1241–1274
4. I. Waldspurger, A. d'Aspremont, and S. Mallat. *Phase recovery, MaxCut and complex semidefinite programming*. Math. Prog., 2013.