#### Fast and Robust Phase Retrieval

Aditya Viswanathan aditya@math.msu.edu

# MICHIGAN STATE

CCAM Lunch Seminar – Purdue University April 18 2014

#### Joint work with



Yang Wang



Mark Iwen

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

# Outline

- 1 The Phase Retrieval Problem
- 2 Existing Approaches
- 3 Computational Framework
- 4 Mathematical Foundations Block Circulant Matrices Angular Synchronization
- 5 Numerical Results
- 6 Future Directions

#### The Phase Retrieval Problem

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

where

- $\mathbf{b} \in \mathbb{R}^D$  are the magnitude or intensity measurements.
- $M \in \mathbbm{C}^{D \times d}$  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.

*Note:* We are interested in recovering the signal modulo trivial ambiguities such as multiplication by a unimodular constant.

# 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, *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

1 The Phase Retrieval Problem

#### 2 Existing Approaches

- 3 Computational Framework
- 4 Mathematical Foundations Block Circulant Matrices Angular Synchronization
- **5** Numerical Results
- 6 Future Directions

# Alternating Projection Methods

- 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, ...

#### Gerchberg – Saxton Algorithm

- Oversampled Fourier magnitude measurements  $\{\mathbf{b}[\omega]\}_{\omega\in\Omega}$
- Known support T (supp $(\mathbf{x}) \subset T$ )
- 1 Choose initial guess  $\mathbf{x}_0$ . Set

$$\widehat{\mathbf{y}}_0[\omega] = \mathbf{b}[\omega] \cdot \frac{\widehat{\mathbf{x}}_0[\omega]}{|\widehat{\mathbf{x}}_0[\omega]|}$$

2 For 
$$k = 1, 2, ...$$

$$\mathbf{x}_{k}[t] = \begin{cases} (\mathcal{F}^{-1}\widehat{\mathbf{y}}_{k-1})[t] & t \in T \\ 0 & \text{else} \end{cases}$$

$$\widehat{\mathbf{y}}_k[\omega] = \mathbf{b}[\omega] \cdot \frac{\widehat{\mathbf{x}}_k[\omega]}{|\widehat{\mathbf{x}}_k[\omega]|}$$

## Gerchberg – Saxton Algorithm

- 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.

- Modify the problem to that of finding the rank-1 matrix  $X = \mathbf{x}\mathbf{x}^*$
- Let  $\mathbf{w}^m$  be a mask. Then the measurements may be written as

$$\begin{aligned} |\langle \mathbf{w}^m, \mathbf{x} \rangle|^2 &= \mathsf{Tr}(\mathbf{x}^* \mathbf{w}^m (\mathbf{w}^m)^* \mathbf{x}) = \mathsf{Tr}(\mathbf{w}^m (\mathbf{w}^m)^* \mathbf{x} \mathbf{x}^*) \\ &:= \mathsf{Tr}(A^m X). \end{aligned}$$

- Let A be the linear operator mapping positive semidefinite matrices into {Tr(A<sup>m</sup>X) : k = 0,...,L}.
- The phase retrieval problem then becomes

$$\begin{array}{ll} \mbox{find} & X \\ & \mathcal{A}(X) = b \\ \mbox{subject to} & X \succeq 0 \\ & \mbox{rank}(X) = 1 \end{array}$$

- Modify the problem to that of finding the rank-1 matrix  $X = \mathbf{x}\mathbf{x}^*$
- Let  $\mathbf{w}^m$  be a mask. Then the measurements may be written as

$$\begin{split} |\langle \mathbf{w}^m, \mathbf{x} \rangle|^2 &= \mathsf{Tr}(\mathbf{x}^* \mathbf{w}^m (\mathbf{w}^m)^* \mathbf{x}) = \mathsf{Tr}(\mathbf{w}^m (\mathbf{w}^m)^* \mathbf{x} \mathbf{x}^*) \\ &:= \mathsf{Tr}(A^m X). \end{split}$$

- Let A be the linear operator mapping positive semidefinite matrices into {Tr(A<sup>m</sup>X) : k = 0,...,L}.
- The phase retrieval problem then becomes

$$\begin{array}{ll} \mbox{minimize} & \mbox{rank}(X) \\ \mbox{subject to} & & \mathcal{A}(X) = b \\ & & X \succeq 0 \end{array}$$

- Modify the problem to that of finding the rank-1 matrix  $X = \mathbf{x}\mathbf{x}^*$
- Let  $\mathbf{w}^m$  be a mask. Then the measurements may be written as

$$\begin{aligned} |\langle \mathbf{w}^m, \mathbf{x} \rangle|^2 &= \mathsf{Tr}(\mathbf{x}^* \mathbf{w}^m (\mathbf{w}^m)^* \mathbf{x}) = \mathsf{Tr}(\mathbf{w}^m (\mathbf{w}^m)^* \mathbf{x} \mathbf{x}^*) \\ &:= \mathsf{Tr}(A^m X). \end{aligned}$$

- Let A be the linear operator mapping positive semidefinite matrices into {Tr(A<sup>m</sup>X) : k = 0,...,L}.
- The phase retrieval problem then becomes

$$\begin{array}{ll} \text{minimize} & \mathsf{rank}(X) \\ \text{subject to} & & \mathcal{A}(X) = b \\ & & X \succeq 0 \end{array}$$

• Unfortunately, this problem is NP hard!

- Modify the problem to that of finding the rank-1 matrix  $X=\mathbf{x}\mathbf{x}^*$
- Let  $\mathbf{w}^m$  be a mask. Then the measurements may be written as

$$\begin{split} |\langle \mathbf{w}^m, \mathbf{x} \rangle|^2 &= \mathsf{Tr}(\mathbf{x}^* \mathbf{w}^m (\mathbf{w}^m)^* \mathbf{x}) = \mathsf{Tr}(\mathbf{w}^m (\mathbf{w}^m)^* \mathbf{x} \mathbf{x}^*) \\ &:= \mathsf{Tr}(A^m X). \end{split}$$

- Let A be the linear operator mapping positive semidefinite matrices into {Tr(A<sup>m</sup>X) : k = 0,...,L}.
- Use the convex relaxation

minimize trace(X)  
subject to 
$$\begin{array}{c} \mathcal{A}(X) = b\\ X \succeq 0 \end{array}$$

• Implemented using a semidefinite program (SDP).

# Outline

- 1 The Phase Retrieval Problem
- 2 Existing Approaches

#### 3 Computational Framework

- 4 Mathematical Foundations Block Circulant Matrices Angular Synchronization
- **5** Numerical Results
- 6 Future Directions

## Overview of the Computational Framework

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

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

- 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 x: We require x to be non-sparse. (The number of consecutive zeros in x should be less than  $\delta$ )

# Correlations with Support-Limited Functions

- Let  $\mathbf{x} = [x_0 \ x_1 \ \dots \ x_{d-1}]^T \in \mathbb{C}^d$  be the unknown signal.
- Let w = [w<sub>0</sub> w<sub>1</sub> ... w<sub>δ</sub> 0 ... 0]<sup>T</sup> denote a support-limited mask. It has δ + 1 non-zero entries.
- Define the shift operator

$$\tau(\mathbf{x}) = [x_{d-1} \ x_0 \ x_1 \ \dots \ x_{d-2}]^T$$

 $\mathbf{v}_k = \tau^k(\mathbf{w})$  denotes a (circular) k-shift of the mask.

• Then, the entries of  $\mathtt{corr}(\mathbf{w},\mathbf{x})$  are given by

$$\langle \mathbf{v}_k, \mathbf{x} \rangle, \quad k = 0, \dots, d-1.$$

#### Correlation Measurements

Now consider the correlation measurements

$$|\langle \mathbf{v}_k^m, \mathbf{x} \rangle| = b_k^m, \qquad k = 0, \dots, d-1, \quad m = 0, \dots, L.$$

Here, L + 1 distinct masks are used.

We have  

$$\begin{split} (b_k^m)^2 &= |\langle \mathbf{v}_k^m, \mathbf{x} \rangle|^2 = \left| \left\langle \tau^k(\mathbf{w}^m), \mathbf{x} \right\rangle \right|^2 = \left| \sum_{j=0}^{\delta} \overline{w}_j^m \cdot x_{k+j} \right|^2 \\ &= \sum_{i,j=0}^{\delta} w_i^m \overline{w}_j^m x_{k+j} \overline{x}_{k+i} =: \sum_{i,j=0}^{\delta} w_i \overline{w}_j Z_{k+j,i-j} \end{split}$$

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.

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



Unknowns:  $[Z_{0,0} \ Z_{0,1} \ Z_{1,0} \ Z_{1,1} \ Z_{2,0} \ Z_{2,1} \ Z_{3,0} \ Z_{3,1}]^T$ Measurements:  $\begin{bmatrix} b_0^0 \ b_1^0 \ b_2^0 \ b_3^0 \ b_1^1 \ b_1^1 \ b_2^1 \ b_3^1 \end{bmatrix}^T$ 

# Angular Synchronization

Re-ordering the phase difference variables  $\{Z_{n,l}\}$ , we obtain entries of the rank-1 matrix  $\mathbf{x}\mathbf{x}^T$  along a band.

$$\begin{bmatrix} |x_0|^2 & x_0\overline{x}_1 & \dots & x_0\overline{x}_{\delta} & 0 & 0 & 0 \\ 0 & |x_1|^2 & x_1\overline{x}_2 & \dots & x_1\overline{x}_{\delta+1} & 0 & 0 \\ & & \dots & \dots & & \\ & & & \dots & \dots & \\ x_{d-1}\overline{x}_0 & \dots & x_{d-1}\overline{x}_{\delta-1} & \dots & \dots & \dots & |x_{d-1}|^2 \end{bmatrix}$$

- The magnitudes of each component of the signal can be estimated from the diagonal entries.
- The phase of each entry can be computed as the phase of the leading eigenvector of this matrix.
- A few iterations of least-squares can also be performed to correct for magnitude errors on the diagonal entries.

# Angular Synchronization

Re-ordering the phase difference variables  $\{Z_{n,l}\}$ , we obtain entries of the rank-1 matrix  $\mathbf{x}\mathbf{x}^T$  along a band.

$$\begin{bmatrix} 0 & \angle x_0 \overline{x}_1 & \dots & \angle x_0 \overline{x}_{\delta} & 0 & 0 & 0 \\ 0 & 0 & \angle x_1 \overline{x}_2 & \dots & \angle x_1 \overline{x}_{\delta+1} & 0 & 0 \\ & & \dots & \dots & & \\ & & & \dots & \dots & \\ \angle x_{d-1} \overline{x}_0 & \dots & \angle x_{d-1} \overline{x}_{\delta-1} & \dots & \dots & 0 \end{bmatrix}$$

- The magnitudes of each component of the signal can be estimated from the diagonal entries.
- The phase of each entry can be computed as the phase of the leading eigenvector of this matrix.
- A few iterations of least-squares can also be performed to correct for magnitude errors on the diagonal entries.

# Outline

- 1 The Phase Retrieval Problem
- 2 Existing Approaches
- 3 Computational Framework
- 4 Mathematical Foundations Block Circulant Matrices Angular Synchronization
- **5** Numerical Results
- 6 Future Directions

# Block Circulant Matrices

Rewriting the system matrix from before,



This a block circulant matrix.

## Spectral Properties of Block Circulant Matrices

Consider the block circulant matrix

- Here,  $A_k \in M_{p imes q}(\mathbb{C})$  are the constituent blocks.
- We are interested in the singular values of M, and, in particular, the condition number of M.

## A Simple Reduction

Let  $\omega_d = e^{2\pi i/d}$  be the primitive *n*-th root of unity. Denote  $\tau_k := \omega_d^k$ ,  $k = 0, 1, \dots, n-1$ . Then

is unitary. Moreover, we can easily show that

$$U_{d,p}^* \ M \ U_{d,q} = \text{block-diag} \left[ J(1), J(\tau_1), \dots, J(\tau_{d-1}) \right],$$

where  $J(t) := A_0 + tA_1 + \dots + t^{n-1}A_{d-1}$ .

# FFT Implementation

- The unitary matrix  $U_{d,r}$  in the previous reduction can be efficiently computed using r FFTs.
- Therefore, the linear system for the phase differences  $\{Z_{n,l}\}$  can be solved using FFTs in conjunction with a block diagonal solve.
- Operation count is

$$\underbrace{(\delta+1)\mathcal{O}(d\log d)}_{\text{to evaluate }U_{d,q}} + \underbrace{(L+1)\mathcal{O}(d\log d)}_{\text{to evaluate }U_{d,p}^*} + \underbrace{\mathcal{O}\left((\delta+1)(L+1)d\right)}_{\text{block-diagonal solve}}$$

- +  $\delta+1$  is the number of non-zero entries in a mask
- L+1 is the number of masks used
- Typical values for  $\delta$  and L are 8 and 12 respectively.

# Singular Value Decomposition

Assume that for each  $t \in \mathbb{C}$  the SVD of J(t) is given by

$$Q(t) \operatorname{diag}_{p \times q} [\sigma_1(t), \dots, \sigma_r(t)] R(t), \qquad r = \min(p, q)$$

where Q(t) is  $p \times p$  and unitary, R(t) is  $q \times q$  and unitary, and

$$\operatorname{diag}\left[c_{1}, c_{2}, \dots, c_{r}\right] = \begin{cases} \left[\operatorname{diag}\left[c_{1}, c_{2}, \dots, c_{r}\right] \ 0\right], & q \geq p \\ \left[\operatorname{diag}\left[c_{1}, c_{2}, \dots, c_{r}\right] \\ 0 & \end{array}\right], & q < p. \end{cases}$$

Then, we have the following theorem

# Singular Value Decomposition

#### Theorem 1

 ${\cal M}$  has the following singular value decomposition

$$M = Q \ U_{d,p} \ D \ U_{d,q}^* \ R^*,$$

where

$$\begin{split} Q &= \textit{block-diag}\left[Q(1), Q(\tau_1), \dots, Q(\tau_{d-1})\right], \\ R &= \textit{block-diag}\left[R(1), R(\tau_1), \dots, R(\tau_{d-1})\right], \\ D &= \textit{block-diag}\left[\textit{diag}_{p \times q}\left(\sigma_1(1), \dots, \sigma_r(1)\right), \dots, \\ \dots &\textit{diag}_{p \times q}\left(\sigma_1(\tau_{n-1}), \dots, \sigma_r(\tau_{d-1})\right)\right]. \end{split}$$

*Remark*: D is in fact not true diagonal if  $p \neq q$ . The true SVD will require permutations to make D a diagonal matrix. But this is just academic.

# Singular Value Decomposition

#### Corollary 1

The singular values of M are

 $\Sigma_M = \{ \sigma_j(\tau_k) : 1 \le j \le \min(p, q), \quad 0 \le k \le d - 1 \}.$ 

The condition number of M when  $p\geq q$  is

$$C = \frac{\max\left(\Sigma_M\right)}{\min\left(\Sigma_M\right)}.$$

#### Corollary 2

Let  $\sigma_1(t) \geq \sigma_2(t) \geq \cdots \geq \sigma_q(t)$  be the singular values of J(t). Let

$$\sigma_1^* = \max_{t \in \mathbb{C}, |t|=1} \sigma_1(t), \qquad \sigma_q^* = \min_{t \in \mathbb{C}, |t|=1} \sigma_q(t).$$

Then the condition number C of M is

$$C \le \sigma_1^* / \sigma_q^*$$

### Block Circulant Matrices - Banded Case

• Of particular interest is the case with  $A_k = 0$  for k > L. Assuming that  $A_0, \ldots, A_L$  are given and fixed (the size d may vary), the condition number of M is (assuming  $p \ge q$ )

$$\frac{\max_{k} \sigma_1(\tau_k)}{\min_{k} \sigma_q(\tau_k)}$$

- Since  $\tau_k = \omega_d^k$ ,  $0 \le k \le d-1$ , as  $d \to \infty$ , the set  $\{\tau_k\}$  becomes increasingly dense on |z| = 1.
- Now,  $\sigma_1(t), \sigma_q(t)$  are that largest and smallest singular values of  $J(t) = A_0 + tA_1 + \cdots + t^{L-1}A_L$ , and they are continuous functions of t.

## Condition Number

• In the banded case, the condition number of  ${\cal M}$  is asymptotically

$$\frac{\max_{\substack{|t|=1}} \sigma_1(t)}{\min_{|t|=1} \sigma_q(t)}$$

as d becomes large.

Representative Condition Numbers

	q = 3	q = 6	q = 9
p = q	34.89	124.05	465.32
p = 2q	3.73	11.73	13.30
p = 3q	3.29	6.08	8.60

# 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  $\theta_{ij} := \theta_i - \theta_j \mod 2\pi$ .

- 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).
- There also exists an eigen-problem formulation.

## The Eigenvector Method (Singer, 2011)

Let S be the  $\{i,j\}$  indices for which the phase angle differences are known. Start with the  $n\times n$  matrix H

$$H_{ij} = \begin{cases} e^{i\theta_{ij}} & \{i,j\} \in S \\ 0 & \{i,j\} \notin S \end{cases}$$

Since  $\theta_{ij}=-\theta_{ij},\,H$  is Hermitian. Now consider the maximization problem

$$\max_{\theta_1,\dots,\theta_n \in [0,2\pi)} \sum_{i,j=1}^n e^{-i\theta_i} H_{ij} e^{i\theta_j}.$$
 (1)

Each correctly determined phase difference contributes

$$e^{-i\theta_i}e^{i(\theta_i-\theta_j)}e^{i\theta_j} = 1$$

to the sum.

# The Eigenvector Method (Singer, 2011)

Unfortunately, (1) is non-convex. Instead, we solve the following relaxation

$$\max_{\substack{z_1,\dots,z_n \in \mathbb{C} \\ \sum_{i=1}^n |z_i|^2 = n}} \overline{z_i} \ H_{ij} \ z_j.$$
  
i.e., 
$$\max_{||z||^2 = n} z^* \ H \ z.$$

This is a quadratic maximization problem whose solution is the leading eigenvector of H.

The phase angles are then given by

$$e^{\hat{\theta}_i} = \frac{v_1(i)}{|v_1(i)|}, \qquad i = 1, \dots, n,$$

where  $v_1$  is the eigenvector corresponding to the largest eigenvalue of  ${\cal H}.$ 

# Outline

- 1 The Phase Retrieval Problem
- 2 Existing Approaches
- 3 Computational Framework
- 4 Mathematical Foundations Block Circulant Matrices Angular Synchronization
- 5 Numerical Results
- 6 Future Directions

#### Numerical Results

- Test signals 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  $\boldsymbol{a}$  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)

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

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

## Noiseless Case



- iid Gaussian signal
- d = 128
- $\delta = 1$ (2d measurements)
- No noise
- Reconstruction Error

$$\frac{\|\hat{\mathbf{x}} - \mathbf{x}\|^2}{\|\mathbf{x}\|^2} = 1.214 \times 10^{-16}$$

# Efficiency



#### Robustness



• iid Gaussian signal

• 
$$d = 256$$

• 
$$\delta = 8$$

• oversampling factor – 1.5

#### Robustness



- iid Gaussian signal
- d = 1024
- $\delta = 8$
- oversampling factor 1.5

#### Robustness



- iid Gaussian signal
- *d* = 64
- 4d measurements
- (oversampling factor of 1.5 for our method)

# Outline

- 1 The Phase Retrieval Problem
- 2 Existing Approaches
- 3 Computational Framework
- 4 Mathematical Foundations Block Circulant Matrices Angular Synchronization
- **5** Numerical Results
- 6 Future Directions

## Current and Future Research Directions

- Use deterministic masks. This will also allow us to write out the condition number explicitly.
- Efficient implementations for two dimensional signals.
- Extension to sparse signals number of required measurements can be reduced.
- Formulation and prescriptions for Fourier measurements convolutions, bandlimited window/mask functions.

## References

- 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
- I. Waldspurger, A. d'Aspremont, and S. Mallat. *Phase recovery, MaxCut and complex semidefinite programming*. Math. Prog., 2013.
- 5 A. Singer, Angular synchronization by eigenvectors and semidefinite programming. Appl. Comp. Harm. Anal., 30:20–36, 2011.