A Problem Proposed by E. J. Candès et al. in Phase Retrieval

Song Li
Joint works with Huiping Li and Qun Mo

School of Mathematical Sciences
Zhejiang University
songli@zju.edu.cn

Sichuan University, 2018-Apr.11

Outline

Outline I

Background

The Phase retrieval problem is of considerable importance in many different areas of science, where capturing phase information is hard or even infeasible. Problems of this kind occur, for example, in X-ray crystallography, diffraction imaging, and astronomy.

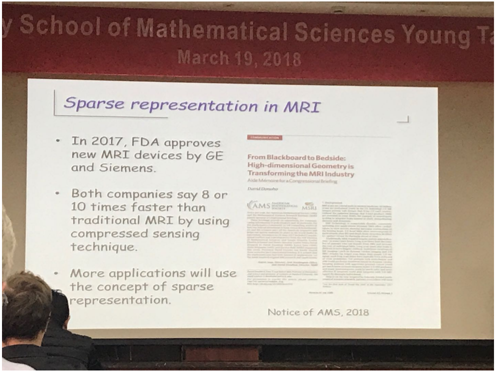

Figure: MRI

Problem Formulation

- **Classical Phase Retrieval Problem**
  - Reconstruct a signal $x_0 = (x_0[0], x_0[1], ..., x_0[n-1])^T \in \mathbb{C}^n$ whose $supp(x) \subseteq [0, n-1]$ from a collection of Fourier magnitude-square measurements $|\hat{x}_0[\omega]|^2$:

$$\hat{x}_0[\omega] = \frac{1}{\sqrt{n}} \sum_{t=0}^{n-1} x_0[t] e^{(\frac{-i2\pi\omega t}{n})}, \quad \omega \in \Omega,$$

    where Ω is a grid of sampled frequencies. A special case is $\Omega = \{0, 1, ..., n-1\}$ with the corresponding unitary discrete Fourier transform.
  - The problem is ill-posed since there are many different signals whose Fourier transforms have the same magnitude.

- **Trivial ambiguities**:
  - cx is also a solution for any scalar $c \in C$ obeying $|c| = 1$.
  - $y[t] = e^{i\phi}x[t]$ is also a solution if $x[t]$ is a solution.

General Models

- **General Problem**:
  - Reconstruct a signal $x_0 \in \mathbb{C}^n$ from its magnitude-square measurements:

$$b_i = |\langle a_i, x_0 \rangle|^2, \quad where\ i = 1, 2, \ldots, m.$$

    where a_i is general measurement vector.
- **Several Special Measurements**:
  - The measurement a_i is sampled independently from some random distribution $\mathbb{D}$, for instance, a_i is sampled independently from $N(0, I_n)$.
  - The measurement a_i is masked Fourier transform. For instance, let $D = diag(\{e^{i2\pi t}\}_{1 \le t \le n})$ be the modulation of a signal $x_0 \in \mathbb{C}^n$. Then we consider $3n$ measurements of the form

$$\{|F_n(x_0)|^2, |F_n((I_n + D^s)x_0)|^2, |F_n((I_n - D^s)x_0)|^2\},$$

    where F_n is the $n \times n$ unitary DFT.
  - x_0 can be recovered up to a global phase if and only if $(n, s) = 1$ provided that the DFT of x_0 does not vanish.

Main Questions

- Questions:
  - Is the signal or image uniquely determined by the noiseless magnitude measurements (up to trivial ambiguities such as global phase)?
  - Is there an efficient algorithm that can provably compute an optimal and stable estimate of the signal or image according to a suitable criterion?

Outline I

Methodology

- Straightforward Choice:

$$\textit{Find} \quad x, \quad \text{s.t.} \quad A(x) = b,$$

  where $b = A(x_0) = \{|\langle a_i, x_0\rangle|^2, i = 1, 2, \ldots, m\}$, $A \in \mathbb{C}^{m\times n}$ is the measurement matrix with rows a_i^* and $x_0 \in \mathbb{C}^n$ is the true vector.
- Obviously, the phase retrieval problem is a nonlinear and nonconvex problem.

"Lifting" Technique

- As is well known, quadratic measurements can be lifted up and interpreted as linear measurements about the rank-one matrix $X_0 = x_0 x_0^*$:

$$|\langle a_i, x_0\rangle|^2 = Tr(a_i a_i^* x_0 x_0^*) = Tr(a_i a_i^* X_0), \quad i = 1, 2, \ldots, m.$$

- Let $\mathcal{A}$ be the linear operator

$$\begin{aligned} \mathcal{A} : \quad & \mathbb{H}^{n\times n} \to \mathbb{R}^m \\ & X \to \mathcal{A}(X) = \sum_{i=1}^{m} tr(a_i a_i^* X) e_i, \end{aligned} \tag{1}$$

then phase retrieval problem reduces to finding a rank-one positive semidefinite matrix X which satisfies these affine measurement constraints as the following reformulation:

$$\begin{aligned} \min_X \quad & rank(X) \\ \text{s. t.} \quad & b = \mathcal{A}(X) \\ & X \geq 0. \end{aligned} \tag{2}$$

Phaselift Method for Phase Retrieval

- However, $rank(X)$ is a nonconvex function of X, we can relax it by the convex surrogate $\text{Tr}(X)$(provided by E. J. Candès and B. Recht et.al 2009-2010) to get a convex SDP program (Phaselift method proposed by Candès, Eldar, Strohmer and Voroninshi, SIAM J. Imaging Sci, 2013):

$$\begin{aligned} \min_{X} \quad & Tr(X) \\ \text{s. t.} \quad & b = \mathcal{A}(X) \\ & X \geq 0. \end{aligned} \tag{3}$$

- Remark: If the solution X of the above problem has rank 1, we factorize it as xx^*. Thus we can recover the true vector up to a global phase.

Stable Models for Noisy Case

- In most application of interest, we have observations of the form

$$b_i = |\langle a_i, x\rangle|^2 + \varepsilon_i, \quad where\ i = 1, 2, \ldots, m,$$

where ε_i is a noise term with $\|\varepsilon\|_2 \leq \eta$. Then we can recover the true vector by solving

  - (Proposed by E. J. Candès and T. Stromher and V. Voroninshi, Comm. Pure Appl. Math, 2013):

$$\begin{aligned} \min_X \quad & Tr(X) \\ \text{s. t.} \quad & \|b - \mathcal{A}(X)\|_2 \leq \eta \\ & X \geq 0, \end{aligned} \tag{4}$$

  - (Proposed by E. J. Candès and X. Li, Found. Comput. Math., 2013):

$$\begin{aligned} \min_X \quad & \|b - \mathcal{A}(X)\|_1 \\ \text{s. t.} \quad & X \geq 0. \end{aligned} \tag{5}$$

General Phase Retrieval in Noiseless Case

Considering the general phase retrieval problem $b_i = |\langle a_i, x_0 \rangle|^2$, where $i = 1, \ldots, m$, the following is the existing guarantees of exact recovery for Gaussian or uniform distributed measurements based on PhaseLift method.

Theorem (E. J. Candès, T. Strohmer, and V. Voroninski, Comm. Pure Appl. Math., 2013)

Consider an arbitrary signal $x_0 \in \mathbb{C}^n$ or $\mathbb{R}^n$ and suppose that $a_i's$ are independently and uniformly distributed on the unit sphere of radius $\sqrt{n}$ and the number of measurements obeys $m \geq c_0 n \log n$, where c_0 is a sufficiently large constant. Then in both the real and complex cases, the solution of (3) is exact with high probability at least $1 - 3e^{-\frac{\gamma m}{n}}$ where γ is a positive absolute constant.

General Phase Retrieval in Noiseless Case

Further, the number of such measurements is improved as follows:

Theorem (E. J. Candès and X. Li, Found. Comput. Math., 2013)

Considering an arbitrary signal $x_0 \in \mathbb{C}^n$ *or* $\mathbb{R}^n$ *and suppose that* $a_i's$ *are independently and uniformly distributed on the unit sphere of radius* $\sqrt{n}$ *and the number of measurements obeys* $m \geq c_1 n$*, where* c_1 *is a sufficiently large absolute constant, then for all* $x_0 \in \mathbb{R}^n$ *or* $\mathbb{C}^n$*, the solution of* (3) *is exact with probability at least* $1 - O(e^{-\gamma m})$*. Thus, exact recovery holds simultaneously over all input signals.*

General Phase Retrieval in Noiseless Case

The result improves upon earlier bounds which requires the number of equations to be at least on the order of $n \log n$ and is optimal for recovering any vector $x_0 \in \mathbb{R}^n$ *or* $\mathbb{C}^n$ exactly from on the order of n quadratic equations.
The above result was also appeared at plenary speaker in ICM 2014 by E.Candes.

General Phase Retrieval in Noisy Case

The stable recovery based on PhaseLift method for Gaussian or uniform distributed measurements:

Theorem (E. J. Candès, T. Strohmer, and V. Voroninski, Comm. Pure Appl. Math, 2013)

Under the above assumption, suppose that the number of measurements obeys $m \geq c_0 n \log n$, where c_0 is a sufficiently large constant, the solution of (4) X obeys $\|X - x_0 x_0^*\|_2 \leq C_0 \eta$ for some positive numerical constant C_0. By finding the eigenvector x corresponding the largest eigenvalue of X, we also have $\|x - e^{i\phi} x_0\|_2 \leq C_0 \min\{\|x_0\|_2, \frac{\eta}{\|x_0\|_2}\}$ for some $\phi \in [0, 2\pi]$. Both these estimates hold with probability at least $1 - 3e^{-\frac{\gamma m}{n}}$ where γ is a positive absolute constant.

General Phase Retrieval in Noisy Case

Further, the error estimate is improved:

Theorem (E. J. Candès and X. Li, Found. Comput. Math., 2013)

Under the above assumption, suppose that the number of equations obeys $m \geq c_1 n$ where c_1 is a sufficiently large constant. Then for all $x_0 \in \mathbb{R}^n$ or $\mathbb{C}^n$, the solution X of (5) obeys $\|X - x_0 x_0^*\|_2 \leq C_1 \frac{\|\varepsilon\|_1}{\sqrt{m}}$ for some numerical constant C_1. By finding the eigenvector x corresponding to the largest eigenvalue of X, we also have $\|x - e^{i\phi} x_0\|_2 \leq C_1 \min\{\|x_0\|_2, \frac{\|\varepsilon\|_1}{\sqrt{m}\|x_0\|_2}\}$ for some $\phi \in [0, 2\pi]$. Both these estimates hold with probability at least $1 - O(e^{-\gamma m})$ where γ is a positive absolute constant.

General Phase Retrieval in Noisy Case

The above result improves upon earlier bounds to the order of n and is optimal for recovering any vector $x_0 \in \mathbb{R}^n$ *or* $\mathbb{C}^n$ exactly from on the order of n quadratic equations. The result was also mentioned at plenary speaker in ICM 2014 by E.Candes.

Theoretical Guarantees for Other Cases

- Candès, Eldar, Strohmer and Voroninshi, SIAM J. Imaging Sci(2013)
- X.Li and V.Voroninski, SIAM J. Math. Anal(2013)
- R. Kueng and H. Rauhut and U. Terstiege, Appl. Comput. Harmon. Anal.(2014)
- J.Topp, IEEE.Trans.Inform.Theory(2014)
- T.Cai and A.Zhang, Ann. Stat(2015)
- D.Gross, F.Krahmer and R. Kueng, J Four. Anal.Appl(2015)
- K.Jaganathan, Y.Eldar and B.Hassibi,IEEE-ISIT(2015)
- F.Krahmer and Y.Liu, IEEE Trans.Inform.Theory(2018)

Outline I

Structured Illumination

Recently, many researchers have focused on their attention to the phase retrieval problem occurred in diffraction imaging. E. J. Candès, X. Li and M. Soltanolkotabi firstly proposed that this problem can be solved by a combination of structured illuminations and Phaselift approach.

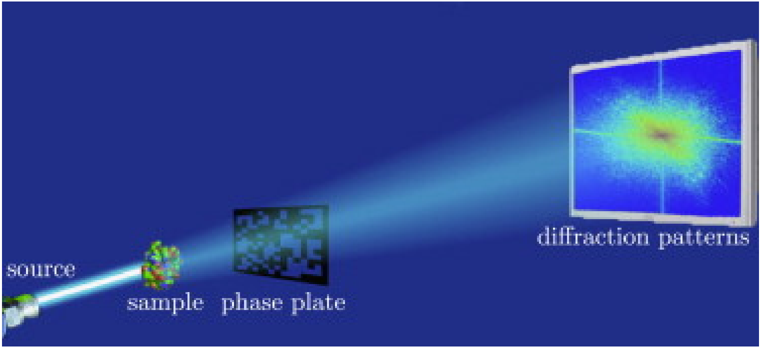

Figure: A typical setup for structured illuminations in diffraction imaging using a phase mask

Coded Diffraction Patterns

- The measurements of Coded Diffraction Patterns are the following form

$$b_{k,l} == \left| \sum_{j=0}^{n-1} x_j \bar{\epsilon}_{j,l} e^{-\frac{i2\pi kj}{n}} \right|^2, \quad 1 \leq k \leq n, 1 \leq l \leq L, \tag{6}$$

where $\epsilon_{l,j}$ are the entries of mask matirces and the corresponding measurement vector is $a_{k,l} = [\bar{\epsilon}_{1,l} e^{\frac{-i2\pi k}{n}}, \ldots, \bar{\epsilon}_{n-1,l} e^{\frac{-i2\pi(n-1)k}{n}}, \bar{\epsilon}_{n,l}]^*$. Two main masks are as follows:

  - (Proposed by E. Candès, X. Li and M. Soltanolkotabi, ACHA, 2015) $\epsilon_{j,l}$ $(1 \leq j \leq n,\ 1 \leq l \leq L)$ are independent copies of a complex random variable ϵ which obeys

$$\begin{aligned} &\mathbb{E}[\epsilon] = \mathbb{E}[\epsilon^2] = 0, \\ &|\epsilon| \leq b \quad \textit{almost surely for some } b > 0, \\ &\mathbb{E}[|\epsilon|^4] = 2\mathbb{E}[|\epsilon|^2]. \end{aligned} \tag{7}$$

- (Proposed by D. Gross and F. Krahmer and R. Kueng, ACHA, 2017) $\epsilon_{j,l}$ $(1 \leq j \leq n,\ 1 \leq l \leq L)$ are independent copies of a real-valued random variable ϵ which obeys

$$\begin{aligned} &\mathbb{E}[\epsilon] = \mathbb{E}[\epsilon^3] = 0, \\ &|\epsilon| \leq b \quad \textit{almost surely for some } b > 0, \\ &\mathbb{E}[\epsilon^4] = 2\mathbb{E}[\epsilon^2]^2, \ \textit{and} \ \ \nu := \mathbb{E}[\epsilon^2]. \end{aligned} \tag{8}$$

The Number of Coded Diffraction Patterns

Theorem (E. J. Candès, X. Li and M. Soltanolkotabi, Appl. Comput. Harmon. Anal, 2013)

Let $x_0 \in \mathbb{C}^n$ be an unknown signal and suppose that the number L of coded diffraction patterns with masks (7) obeys $L \geq c \log^4 n$ for some fixed numerical constant c. Then with probability at least $1 - \frac{1}{n}$, the Phaselift (3) program reduces to a unique point, namely, $x_0 x_0^*$, and thus recovers x_0 up to a global phase. For $\gamma \geq 1$, setting $L \geq c\gamma \log^4 n$ leads to a probability of success at least $1 - \frac{1}{n^\gamma}$.

Remark:

- The result was mentioned in ICM (2014) by the plenary speaker E. J. Candès.
- They expect that the number L of coded diffraction patterns can be reduced to be a figure independent of the dimensions of complex signals.

Further, the number of coded diffraction patterns is improved as follows:

Theorem (D. Gross, F. Krahmer and R. Kueng, Appl. Comput. Harmon. Anal, 2014)

Let $x_0 \in \mathbb{C}^n$ be an unknown signal with $\|x_0\|_2 = 1$ and let $n \geq 3$ be an odd number. Suppose that L coded diffraction patterns with masks (8) are performed. Then with probability at least $(1 - e^{-\omega})$, Phaselift (3) with the additional constraint $tr(X_0) = 1$, recovers x_0 up to a global phase, provided that $L \geq c\omega \log^2 n$. Here, $\omega \geq 1$ is an arbitrary parameter and C a dimension-independent constant that can be explicitly bounded.

Remark: The number of masked Fourier measurements has dropped to $O(n \log^2 n)$ when the signals are limited to odd dimensional complex signals.

Our Results

We partially improve their results. Namely, $O(\log^2 n)$ coded diffraction patterns are sufficient to recover real-valued signals of even dimensions.

Theorem (Huiping Li, Song Li and Qun Mo, 2018)

Let $x_0 \in \mathbb{R}^n$ be an unknown signal with $\|x_0\|_{l_2} = 1$ and let $n > 3$ be an even number. Then with probability at least $(1 - e^{-\theta})$, Phaselift (3) with the additional constraint $tr(X_0) = 1$ recovers x_0 up to a global phase, provided that $L \geq C_0 \theta \log^2 n$. Here $\theta \geq 1$ is an arbitrary parameter and C_0 is a constant which is independent of the signal's dimension.

Technical Lemmas

Lemma (Huiping Li, Song Li and Qun Mo, 2018)

Let n be even and $\mathcal{R} : \mathbb{H}^{n\times n} \to \mathbb{H}^{n\times n}$ be defined by

$$\mathcal{R}(Z) = \frac{1}{\upsilon^2 nL}\sum_{l=1}^{L}\sum_{k=1}^{n}\Pi_{a_{k,l}a_{k,l}^*}(Z) = \frac{1}{\upsilon^2 nL}\sum_{k=1}^{n} tr(a_{k,l}a_{k,l}^* Z)a_{k,l}a_{k,l}^*. \quad (9)$$

Then we have

$$\mathbb{E}[\mathcal{R}] = \Pi_{I_n} + \mathcal{I} + \mathcal{N} \ \text{ or } \ \mathbb{E}[\mathcal{R}](Z) = tr(Z)I_n + Z + \mathcal{N}(Z) \ \forall Z \in \mathbb{H}^{n\times n}. \quad (10)$$

where $\mathcal{N}(Z) = \sum_{k=1}^{\frac{n}{2}} z_{k+\frac{n}{2},k} e_k e_{k+\frac{n}{2}}^* + \sum_{k=\frac{n}{2}+1}^{n} z_{k-\frac{n}{2},k} e_k e_{k-\frac{n}{2}}^*$ $, Z = (z_{i,j})_{1\le i,j\le n}$
and $\Pi_Y : \mathbb{H}^{n\times n} \to \mathbb{H}^{n\times n}$ *is defined by*

$$\Pi_Y(Z) = Y\langle Y, Z\rangle = tr(YZ)Y, \quad \forall Z \in \mathbb{H}^{n\times n}.$$

Robust Injectivity, Lower Bound

Lemma (Huiping Li, Song Li and Qun Mo, 2018)

For any $\delta \in (0,1)$, with high probability at least $1 - n^2 e^{(-\frac{\nu^4\delta^2 L}{C_1 b^8})}$, the inequality

$$\frac{1}{\upsilon^2 nL}\|\mathcal{A}(Z)\|_{l_2}^2 > (1-\delta)\|Z\|_2^2, \tag{11}$$

is valid for all nonzero matrices $Z \in T = \{x_0 z^T + z x_0^T : \forall z \in \mathbb{R}^n\} \subset \mathbb{S}^{n\times n}$ simultaneously. Here n is an even number and C_1 is an absolute constant.

Remark: (11) is limited to nonzero symmetric matrices in T. And it fails to be true for the corresponding $T^{'} := \{x_0 z^* + z x_0^* : \forall z \in \mathbb{C}^n\}$ if the dimension of the signal $x_0 \in \mathbb{C}^n$ is even.

Counter-Example

- Let n be even. For example, if we choose $x_0 = [1-i, \underbrace{0, \cdots, 0}_{\frac{n}{2}-1}, 3-3i, \underbrace{0, \cdots, 0}_{\frac{n}{2}-1}]^* \in \mathbb{C}^n$, then there exists $z = [-2-2i, \underbrace{0, \cdots, 0}_{\frac{n}{2}-1}, -4-4i, \underbrace{0, \cdots, 0}_{\frac{n}{2}-1}]^* \in \mathbb{C}^n$ such that

$$Z = x_0 z^* + z x_0^* = \begin{pmatrix} 0 & \cdots & 0 & 4i & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ -4i & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}$$

with $Z \in T^{'}$ and $\|Z\|_2 = 32$. Based on a simple calculation, we find that $\frac{1}{\upsilon^2 \cdot nL}\|\mathcal{A}(Z)\|_{l_2}^2 = 0$, which implies that (11) is failed to be true.

Thank you!

Song Li
School of Mathematical Sciences,
Zhejiang University
songli@zju.edu.cn