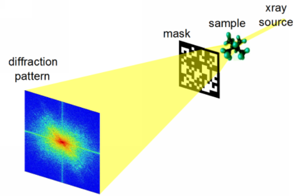
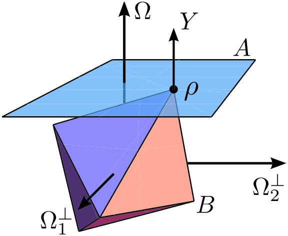
A Partial Derandomization of Phase Retrieval via PhaseLift





David Gross
Richard Küng
(Freiburg)

Felix Krahmer
(Göttingen)

TU Berlin
December 2013

[Building on, and evolving with, work by

Candès, Strohmer, Voroninski, Li, Soltanolkotabi]

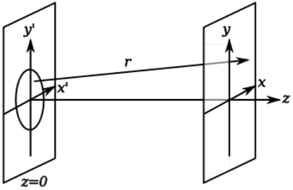
Outline

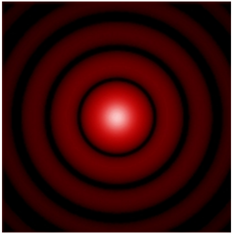
- Motivation
- Low-rank recovery
- Spherical Designs

- Results.

The Phase Retrieval Problem

Motivation: Far-Field Optics

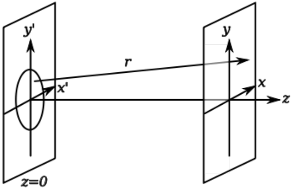


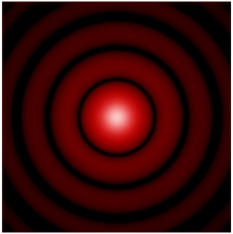


- Illuminate small object with coherent light...
- ...at screen:

$$\begin{aligned} &\text{observed intenstity} \\ \simeq\quad &\Big|\text{Fourier transform of object}\Big|^2. \end{aligned}$$

Motivation: Far-Field Optics





- Illuminate small object with coherent light...
- ...at screen:

$$\begin{aligned} & \text{observed intenstity} \\ \simeq \quad & \left|\text{Fourier transform of object}\right|^2. \end{aligned}$$

Q: How does one recover the object from screen intensities?

Motivation: Phase Retrieval

Discrete model:

- Unknown $x \in \mathbb{C}^n$.
- Measurement vectors $A_1, \ldots, A_k$,
  (represented as rows of sensing matrix A).

Motivation: Phase Retrieval

Discrete model:

- Unknown $x \in \mathbb{C}^n$.
- Measurement vectors $A_1, \ldots, A_k$,
  (represented as rows of sensing matrix A).

The general *phase retrieval problem*:

Recover x from quadratic measurements: $y_i = |A_i x|^2$.

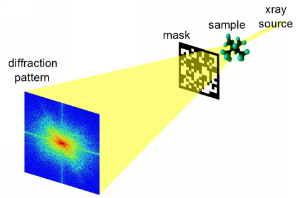
Motivation: Phase Retrieval

Discrete model:

- Unknown $x \in \mathbb{C}^n$.
- Measurement vectors $A_1, \ldots, A_k$, (represented as rows of sensing matrix A).

The general *phase retrieval problem*:

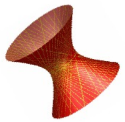
Recover x from quadratic measurements: $y_i = |A_i x|^2$.



- $\{A_i\}$ a basis clearly not enough
- Small number of bases seems to work
- Practice: 2-D FT after masks

Problem has long history

Applications, geometry, algorithms studied since '70s. . .



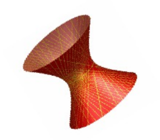
Solution set is *algebraic variety*:

$$V := \left\{ x \,\middle|\, (|A_i x|^2 - y_i) = 0 \,\forall\, i \right\}.$$

$\Rightarrow$ algebraic geometry seems right language.

Problem has long history

Applications, geometry, algorithms studied since '70s. . .



Solution set is *algebraic variety*:

$$V := \left\{x \,\middle|\, (|A_i x|^2 - y_i) = 0 \,\forall\, i\right\}.$$

$\Rightarrow$ algebraic geometry seems right language.

One central result in this language:

- $O(4n)$ "generic" measurements required and sufficient [Mixon, Voroninski, Wolf, . . . ; past few years]

From AG to convex optimization

Problems with AG:

- Ill-equipped to handle *approximations, noise, uncertainties*.
- Doesn't naturally yield *efficient algorithms*.
- Difficult to handle *specific*, non-generic cases.

From AG to convex optimization

Problems with AG:

- Ill-equipped to handle *approximations, noise, uncertainties*.
- Doesn't naturally yield *efficient algorithms*.
- Difficult to handle *specific*, non-generic cases.

New approach in 2012: PhaseLift, based on convex optimization [Candès, Strohmer, Voroniski].

Lift it...

"Lift" quadratic equation to matrix version [Balan *et al.*]:

$$y_i = |A_i x|^2 = \operatorname{tr}(A_i A_i^*) x x^*.$$

Lift it. . .

"Lift" quadratic equation to matrix version [Balan *et al.*]:

$$y_i = |A_i x|^2 = \mathrm{tr}(A_i A_i^*) x x^*.$$

Then recovery equivalent to affinely constrained rank-minimization:

$$\begin{aligned} & \min \operatorname{rank} X \\ \text{s.t.} \quad & X = X^\dagger \\ & \mathrm{tr}(A_i A_i^*) X = y_i. \end{aligned}$$

($X = xx^*$ is solution).

Lift it...

"Lift" quadratic equation to matrix version [Balan *et al.*]:

$$y_i = |A_i x|^2 = \mathrm{tr}(A_i A_i^*) x x^*.$$

Then recovery equivalent to affinely constrained rank-minimization:

$$\begin{aligned} &\min \operatorname{rank} X \\ \text{s.t.} \quad & X = X^\dagger \\ & \mathrm{tr}(A_i A_i^*) X = y_i. \end{aligned}$$

($X = xx^*$ is solution).

- But these are NP-hard in general.
- Any hope?

Lift it. . .

"Lift" quadratic equation to matrix version [Balan *et al.*]:

$$y_i = |A_i x|^2 = \operatorname{tr}(A_i A_i^*) x x^*.$$

Then recovery equivalent to affinely constrained rank-minimization:

$$\begin{aligned} & \min \operatorname{rank} X \\ \text{s.t.} \quad & X = X^\dagger \\ & \operatorname{tr}(A_i A_i^*) X = y_i. \end{aligned}$$

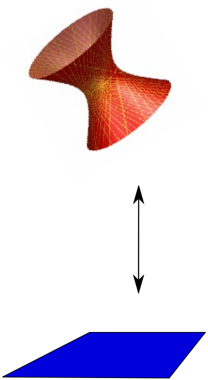
($X = xx^*$ is solution).

- But these are NP-hard in general.
- Any hope?

Well. . . rewind to 2009 & recall low-rank recovery results . . .

Low-rank matrix recovery via convex optimization

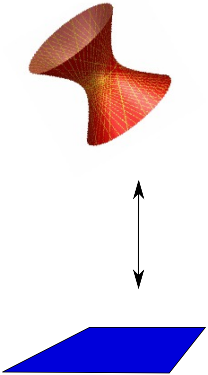
Matrix version of compressed sensing



- Consider hermitian $(n \times n)$-matrix X of rank $r \ll n$:

$$X = \sum_{i=1}^{r} v_i v_i^* .$$

Matrix version of compressed sensing

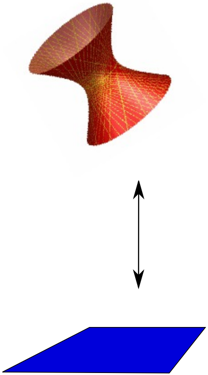


- Consider hermitian $(n \times n)$-matrix X of rank $r \ll n$:

$$X = \sum_{i=1}^{r} v_i v_i^* .$$

- Described by $\simeq (r \times n) \ll n^2$ parameters.

Matrix version of compressed sensing



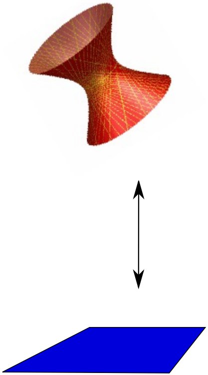
- Consider hermitian $(n \times n)$-matrix X of rank $r \ll n$:

$$X = \sum_{i=1}^{r} v_i v_i^*.$$

- Described by $\simeq (r \times n) \ll n^2$ parameters.
- Now choose (overcomplete) *matrix* basis $\{B_i\}_{i=1}^{n^2}$

$$X = \sum_{i=1}^{n^2} c_i B_i.$$

Matrix version of compressed sensing



- Consider hermitian $(n \times n)$-matrix X of rank $r \ll n$:

$$X = \sum_{i=1}^{r} v_i v_i^*.$$

- Described by $\simeq (r \times n) \ll n^2$ parameters.
- Now choose (overcomplete) *matrix* basis $\{B_i\}_{i=1}^{n^2}$

$$X = \sum_{i=1}^{n^2} c_i B_i.$$

Problem statement:

> *(Low-rank matrix recovery).* Can one recover a rank-r matrix X from only $O(r\,n)$ randomly chosen expansion coefficients c_i w.r.t. some fixed matrix basis?

(*matrix* or *basis-independent* or *non-commutative* version of c.s.)

Incoherences

- Some pairs of bases and matrices don't work:

$$\begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix}$$

Incoherences

- Some pairs of bases and matrices don't work:

$$\begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix}$$

- Introduce "well-posedness parameter" μ:

$$|\operatorname{tr} B_i X|^2 \simeq \mu \frac{r}{n}.$$

Two ways to guarantee well-posed problems:

Incoherences

- Some pairs of bases and matrices don't work:

$$\begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix}$$

- Introduce "well-posedness parameter" μ:

$$|\operatorname{tr} B_i X|^2 \simeq \mu \frac{r}{n}.$$

Two ways to guarantee well-posed problems:

- Restrict *basis*:
  - demand B_i to have "small operator norm"; akin to Fourier basis in c.s.
  - *Unitary operator bases* optimal [DG09]

Incoherences

- Some pairs of bases and matrices don't work:

$$\begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix}$$

- Introduce "well-posedness parameter" μ:

$$|\operatorname{tr} B_i X|^2 \simeq \mu \frac{r}{n}.$$

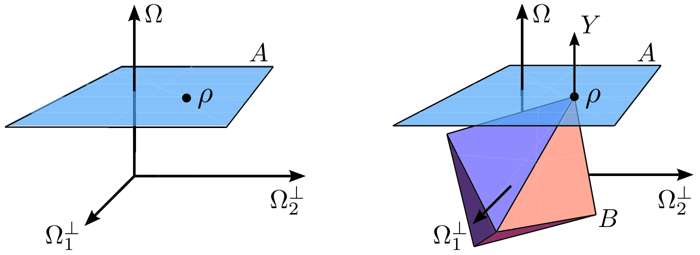
Two ways to guarantee well-posed problems:

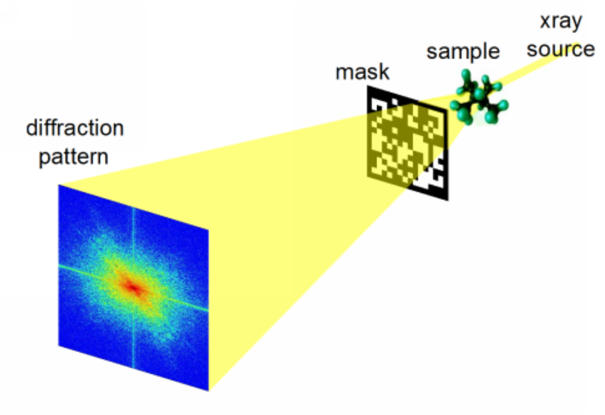
- Restrict *basis*:
  - demand B_i to have "small operator norm"; akin to Fourier basis in c.s.
  - *Unitary operator bases* optimal [DG09]
- Restrict *signals*:
  - demand singular vectors of X be "spread out" [Candès *et al.* 09].

General Low-Rank Recovery

Theorem [Candès, Recht, Tao. . . ; DG '09]

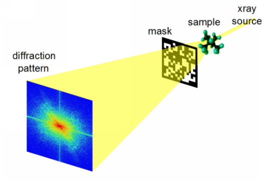
- Any rank-r matrix X can be *exactly* recovered (w.h.p.) from $rn\mu\,(\log^2 n)$ randomly chosen expansion coefficients
$$c_k = \operatorname{tr} B_k X.$$
- The unknown matrix X minimizes the nuclear norm in the affine space defined by the known coefficients.





...back to phase retrieval.

PhaseLift

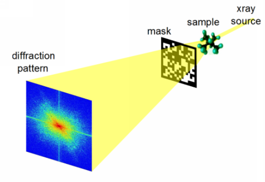


$$\begin{aligned} &\min \|X\|_* \\ \text{s.t.} \quad & X = X^\dagger \\ & \operatorname{tr}(A_i A_i^*)X = y_i. \end{aligned}$$

Neither condition for μ small works for phase retrieval:

- $B_i = A_i A_i^*$ are rank-1: worst case.
- $X = xx^*$ with x highly structured: worst case.

PhaseLift



$$\begin{aligned} & \min \|X\|_* \\ \text{s.t.} \quad & X = X^\dagger \\ & \operatorname{tr}(A_i A_i^*) X = y_i . \end{aligned}$$

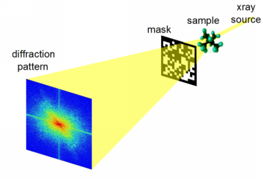
Neither condition for μ small works for phase retrieval:

- $B_i = A_i A_i^*$ are rank-1: worst case.
- $X = xx^*$ with x highly structured: worst case.

First results [Candès, Strohmer, Voroninski, Li '12]:

- Assume *Gaussian* measurements, $A_{i,j} \sim \mathcal{N}(0,1)$.
- $\Rightarrow O(n)$ samples guarantee recovery.

Great. But.

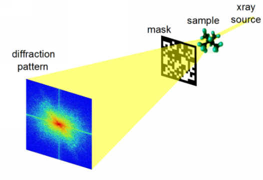


$$\begin{aligned} & \min \|X\|_* \\ \text{s.t.} \quad & X = X^\dagger \\ & \mathrm{tr}(A_i A_i^*) X = y_i. \end{aligned}$$

Achieved – Put phase retrieval into convex optimization realm:

- Rigorous recovery guarantees
- Formulation is robust against noise
- Efficient algorithms (at least in computer science sense)

Great. But.



$$\begin{aligned} &\min \|X\|_* \\ \text{s.t.} \quad & X = X^\dagger \\ & \mathrm{tr}(A_i A_i^*) X = y_i. \end{aligned}$$

Achieved – Put phase retrieval into convex optimization realm:

- Rigorous recovery guarantees
- Formulation is robust against noise
- Efficient algorithms (at least in computer science sense)

To be done – Gaussian measurements not completely satisfactory:

- Practically: doesn't cover common use cases
- Conceptually: unclear which properties "make it work"

$\Rightarrow$ derandomize these results!

Two simultaneous pre-prints arXiv:1310

[DG, Richard Küng, Felix Krahmer]:

- *Conceptual* approach: find right "well-posedness parameter"
- Quite general, but rather abstract, condition
- Based on "spherical designs"

[Emmanuel Candès, Xiadong Li, Mahdi Soltanolkotabi]:

- *Practical* point of view
- Solution for masked Fourier case

Two simultaneous pre-prints arXiv:1310

[DG, Richard Küng, Felix Krahmer]:

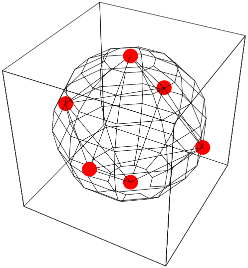
- *Conceptual* approach: find right "well-posedness parameter"
- Quite general, but rather abstract, condition
- Based on "spherical designs"

[Emmanuel Candès, Xiadong Li, Mahdi Soltanolkotabi]:

- *Practical* point of view
- Solution for masked Fourier case

(Technical parts of papers surprisingly similar).

Spherical designs



Painless motivation for spherical designs

Brainstorm about properties of measurement vectors A_i:

- $\{A_i\}$ ortho-normal basis clearly not enough

Painless motivation for spherical designs

Brainstorm about properties of measurement vectors A_i:

- $\{A_i\}$ ortho-normal basis clearly not enough
- It is clear from

  $$y_i = \mathrm{tr}(A_i A_i^*) x x^*$$

  that reconstruction possible if $\{A_i A_i^*\}$ form matrix basis
- (That observation is Balan's *painless* method)

Painless motivation for spherical designs

Brainstorm about properties of measurement vectors A_i:

- $\{A_i\}$ ortho-normal basis clearly not enough
- It is clear from

  $$y_i = \mathrm{tr}(A_i A_i^*) x x^*$$

  that reconstruction possible if $\{A_i A_i^*\}$ form matrix basis
- (That observation is Balan's *painless* method)

Hence:

What is condition on A_i such that outer products $A_i A_i^*$ form ortho-normal matrix basis?

(their traceless part a tight frame in space of traceless matrices).

Painless motivation for spherical designs

A set of vectors A_i is...

- ...tight frame iff

$$\mathbb{1} \propto \sum_i A_i A_i^* \propto \int_{S^{n-1}} A A^* \, \mathrm{d}A,$$

i.e. if it agrees with Haar measure to 2nd moments.

Painless motivation for spherical designs

A set of vectors A_i is...

- ...tight frame iff

$$\mathbb{1} \propto \sum_i A_i A_i^* \propto \int_{S^{n-1}} AA^* \, \mathrm{d}A,$$

  i.e. if it agrees with Haar measure to 2nd moments.

- ...such that $\{A_i A_i^*\}$ is tight frame in matrix space iff

$$P_{\mathrm{Sym}^2} \propto \sum_i A_i A_i^* \otimes A_i A_i^* \propto \int_{S^{n-1}} AA^* \otimes AA^* \mathrm{d}A,$$

  i.e. if it agrees with Haar measure to 4th moments.

Painless motivation for spherical designs

A set of vectors A_i is...

- ...tight frame iff

$$\mathbb{1} \propto \sum_i A_i A_i^* \propto \int_{S^{n-1}} AA^* \,\mathrm{d}A,$$

i.e. if it agrees with Haar measure to 2nd moments.

- ...such that $\{A_i A_i^*\}$ is tight frame in matrix space iff

$$P_{\mathrm{Sym}^2} \propto \sum_i A_i A_i^* \otimes A_i A_i^* \propto \int_{S^{n-1}} AA^* \otimes AA^* \mathrm{d}A,$$

i.e. if it agrees with Haar measure to 4th moments.

- ...a *k-design* iff

$$P_{\mathrm{Sym}^k} \propto \sum_i \left(A_i A_i^*\right)^{\otimes k} \propto \int_{S^{n-1}} \left(AA^*\right)^{\otimes k} \mathrm{d}A,$$

i.e. if it agrees with Haar to moments of order $2k$.

2-dimensional example

Map $\mathbb{C}^2$ to $\mathbb{R}^3$ by:

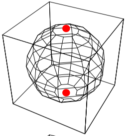
$$(\sin\theta, e^{i\phi}\cos\theta) \to (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta).$$

2-dimensional example

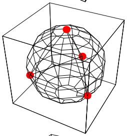
Map $\mathbb{C}^2$ to $\mathbb{R}^3$ by:

$$(\sin\theta, e^{i\phi}\cos\theta) \to (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta).$$

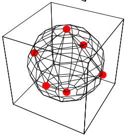
Then:



- ortho-normal basis is 1-design



- set of equi-angular lines is 2-design



- set of “mutually unbiased bases” is 3-design

Spherical Designs: Summary

Spherical k-designs:

- Provide extrapolation between “basis” and “Haar random”
- Contain information about moments up to order $2k$
- Appear naturally in Balan’s painless phase retrieval scheme

Spherical Designs: Summary

Spherical k-designs:

- Provide extrapolation between "basis" and "Haar random"
- Contain information about moments up to order $2k$
- Appear naturally in Balan's painless phase retrieval scheme

Some facts:

- Efficient, randomized constructions for any order
- Explicit construction for any dimension up to order 3
- Cardinality: $O(n^{2k})$

Results

Results

Basic philosophy:

Use "degree of agreement with Haar measure" as "well-posedness" parameter.

Results

Basic philosophy:

Use "degree of agreement with Haar measure" as "well-posedness" parameter.

Theorem 1 [DG, Küng, Krahmer]
PhaseLift works from random subset of k-design of size

$$O\left(d^{1+2/k}\log^2 d\right).$$

- Non-trivial for $k \geq 3$
- Tight if $k = 2\log n$
- Conjecture: $O(n\, polylog(n))$ possible for $k \geq 3$.

Results

Basic philosophy:

Use “degree of agreement with Haar measure” as “well-posedness” parameter.

Theorem 2 [DG, Küng, Krahmer] Conversely, for $k = 2$, phase retrieval from sub-quadratic number of measurements impossible.

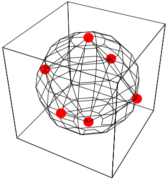
- I.e. subsampling from Balan-type ensemble impossible.

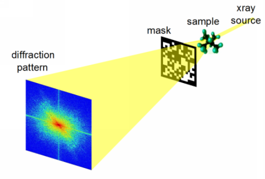
Summary

We have...

- ...constructed a partial derandomization of PhaseLift,
- ...pointed to spherical designs as general-purpose derandomization tools for structured signal recovery schemes. [See also M. Fickus talk ("shockingly bad" in his context)]

Thank you for your attention!





David Gross
University of Freiburg

TU Berlin
December 2013