Phase Retrieval, MAXCUT and Complex Semidefinite Programming

Alexandre d'Aspremont, *CNRS & Ecole Polytechnique*,

with **Fajwel Fogel**, **Irène Waldspurger** and **Stéphane Mallat**, *ENS*.

Support from ERC (project SIPA).

Introduction

Focus on the **phase retrieval** problem, i.e.

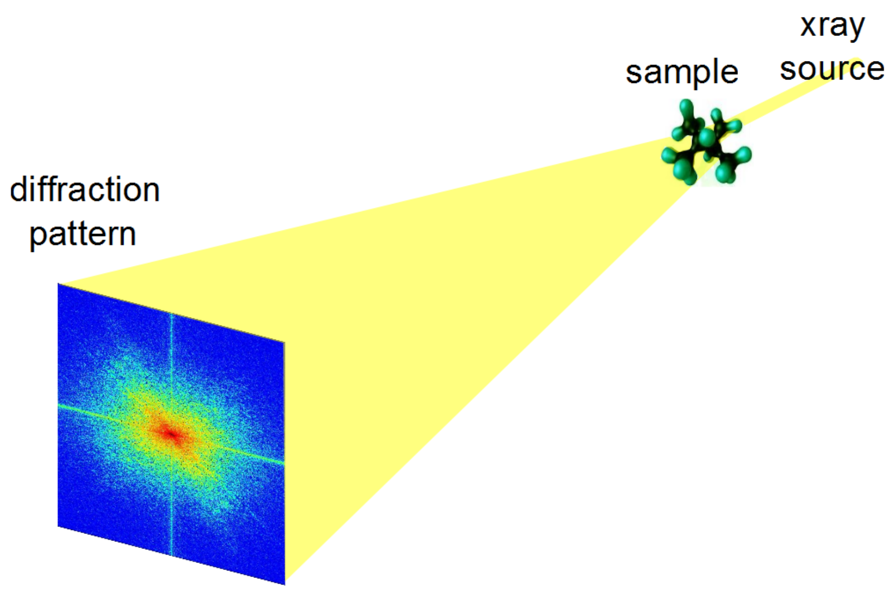
$$\begin{array}{ll} \text{find} & x \\ \text{such that} & |\langle a_i, x\rangle|^2 = b_i^2, \quad i = 1, \ldots, n \end{array}$$

in the variable $x \in \mathbb{C}^p$.

- Reconstruct a signal x from the **amplitude of n linear measurements**.
- We seek a **tractable** procedure, i.e. a polynomial time algorithm with explicit approximation and complexity bounds.

Introduction

Applications in e.g. molecular imaging



(from [Candes et al., 2011b])

- CCD sensors only record the **magnitude** of diffracted rays, and loose the **phase**
- **Fraunhofer diffraction:** phase is required to invert the 2D Fourier transform

Introduction

Problem is almost 100 years old, infinite list of references. . .

Algorithms

- Greedy algorithm [Gerchberg and Saxton, 1972]
- Classical survey of algorithms by [Fienup, 1982].
- NP-complete [Sahinoglou and Cabrera, 1991].
- Matrix completion formulation [Chai, Moscoso, and Papanicolaou, 2011] and [Candes, Strohmer, and Voroninski, 2011a]

Applications

- X-ray and crystallography imaging [Harrison, 1993], diffraction imaging [Bunk et al., 2007] or microscopy [Miao et al., 2008].
- Audio signal processing [Griffin and Lim, 1984].

Introduction

Classical **greedy algorithm** [Gerchberg and Saxton, 1972].

Input: An initial $y^1 \in \mathbb{C}^n$, i.e. such that $|y^1| = b$.

1: **for** $k = 1, \ldots, N-1$ **do**

2: Set

$$w = AA^\dagger y^k$$

3: Set

$$y_i^{k+1} = b_i \frac{w}{|w|}, \quad i = 1, \ldots, n.$$

4: **end for**

Output: $y_N \in \mathbb{C}^n$.

Very similar to **alternating projections**:

- Project on $\mathcal{R}(A)$.
- Adjust the magnitude to match b
- Repeat. . .

Introduction

- [Chai et al., 2011] and [Candes et al., 2011a] use a **lifting** procedure from [Shor, 1987, Lovász and Schrijver, 1991] to write

$$|\langle a_i, x \rangle|^2 = b_i^2 \quad \Longleftrightarrow \quad \mathbf{Tr}(a_i a_i^* x x^*) = b_i^2$$

  and formulate phase recovery as a **matrix completion** problem

$$\begin{array}{ll} \text{Minimize} & \mathbf{Rank}(X) \\ \text{such that} & \mathbf{Tr}(a_i a_i^* X) = b_i^2, \quad i = 1, \ldots, n \\ & X \succeq 0 \end{array}$$

  in the matrix $X \in \mathbf{H}_p$.

- [Recht et al., 2007, Candes and Recht, 2008, Candes and Tao, 2010] show that under certain conditions on A and x_0, it suffices to solve

$$\begin{array}{ll} \text{Minimize} & \mathbf{Tr}(X) \\ \text{such that} & \mathbf{Tr}(a_i a_i^* X) = b_i^2, \quad i = 1, \ldots, n \\ & X \succeq 0 \end{array}$$

  which is a (convex) **semidefinite program** in $X \in \mathbf{H}_p$.

Outline

- Introduction
- **MAXCUT formulation**
- Tightness
- Algorithms & Structure
- Numerical Results

MAXCUT formulation

We can **decouple** the phase and magnitude reconstruction problems.

- In the noiseless case, write $Ax = \mathbf{diag}(b)u$ where $u \in \mathbb{C}^n$ is a **phase vector** with $|u_i| = 1$.
- The phase recovery problem can be written

$$\min_{\substack{u \in \mathbb{C}^n,\, |u_i|=1, \\ x \in \mathbb{C}^p}} \quad \|Ax - \mathbf{diag}(b)u\|_2^2,$$

- The inner minimization problem in x is a standard least squares, with solution $x = A^\dagger \,\mathbf{diag}(b)u$, so phase recovery becomes

$$\begin{array}{ll} \text{minimize} & u^* M u \\ \text{subject to} & |u_i| = 1, \quad i = 1, \ldots n, \end{array}$$

  in $u \in \mathbb{C}^n$, where the Hermitian matrix $M = \mathbf{diag}(b)(\mathbf{I} - AA^\dagger)\,\mathbf{diag}(b)$ is positive semidefinite.

MAXCUT formulation

MAXCUT. Classical algorithm in combinatorial optimization.

- Given an undirected graph with weights w_{ij} on its edges (i,j), *MaxCut* seeks to partition the vertices in two sets S and $\bar{S}$ to **maximize the weight of the cut**

$$\max_{S \subset [1,n]} \sum_{\{i \in S,\ j \in \bar{S}\}} w_{ij}$$

- This can be written as a quadratic program

$$\begin{array}{ll} \text{maximize} & x^T L x \\ \text{subject to} & x_i^2 = 1, \quad i = 1, \ldots, n \end{array}$$

  where L is the graph Laplacian, $L = \mathbf{diag}(We) - W$.

- Other interpretations as computing the ground state of spin glass models [Mezard and Montanari, 2009], computing mixed matrix norms [Nemirovski, 2005], approximating the CUT-norm [Alon and Naor, 2004], etc...

MAXCUT formulation

MAXCUT. We know a lot about how to find an approximate solution

$$\begin{array}{ll} \text{maximize} & x^T L x \\ \text{subject to} & x_i^2 = 1, \quad i = 1, \ldots, n \end{array}$$

- [Goemans and Williamson, 1995] produce a polynomial algorithm with an approximation ratio of $0.878\ldots$, using a semidefinite relaxation

$$\begin{array}{ll} \text{maximize} & \mathbf{Tr}(XL) \\ \text{subject to} & \mathbf{diag}(X) = 1, \; X \succeq 0 \end{array}$$

  combined with a randomization argument.
- Approximating the solution with an approximation ratio better than $16/17$ is NP-Hard, etc.

MAXCUT formulation

The **phase recovery** problem was written (in phase) as

$$\begin{array}{ll} \text{minimize} & u^* M u \\ \text{subject to} & |u_i| = 1, \quad i = 1, \ldots n, \end{array}$$

- We can write a relaxation for phase recovery similar to the MAXCUT SDP, and recycle all the efficient algorithms designed for MAXCUT to solve it.
- Nesterov [1998] produces approximation bounds for generic nonconvex quadratic programs. [Goemans and Williamson, 2001, Zhang and Huang, 2006] extend these results to complex valued problems and show a $\pi/4$ approximation ratio for

  $$\begin{array}{ll} \text{maximize} & u^* M u \\ \text{subject to} & |u_i| = 1, \quad i = 1, \ldots n, \end{array}$$

  when $M \succeq 0$.
- Tightness results on very similar maximum-likelihood channel detection problems [Luo et al., 2003, Kisialiou and Luo, 2010, So, 2010].

Outline

- Introduction
- MAXCUT formulation
- **Tightness**
- Algorithms & Structure
- Numerical Results

Outline

Tightness. [Waldspurger, d'Aspremont, and Mallat, 2012] Write a semidefinite relaxation for phase recovery, similar to the MAXCUT SDP

$$\begin{array}{ll} \text{Minimize} & \mathbf{Tr}(MU) \\ \text{such that} & \mathbf{diag}(U) = 1, \; X \succeq 0 \end{array}$$

call it **PhaseCut**. When do we perfectly recover the signal x?

- [Candes et al., 2011a] show exact recovery w.h.p. for the **PhaseLift** relaxation

$$\begin{array}{ll} \text{Minimize} & \mathbf{Tr}(X) \\ \text{such that} & \mathbf{Tr}(a_i a_i^* X) = b_i^2, \quad i = 1, \ldots, n \\ & X \succeq 0 \end{array}$$

  when $n = O(p \log p)$ observations a_i are picked uniformly on the unit sphere.
- [Waldspurger et al., 2012] show

PhaseCut is tight whenever PhaseLift is.

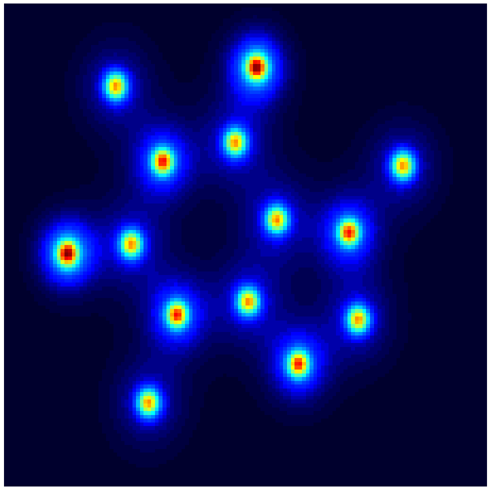
- Empirically, slightly more robust to noise.

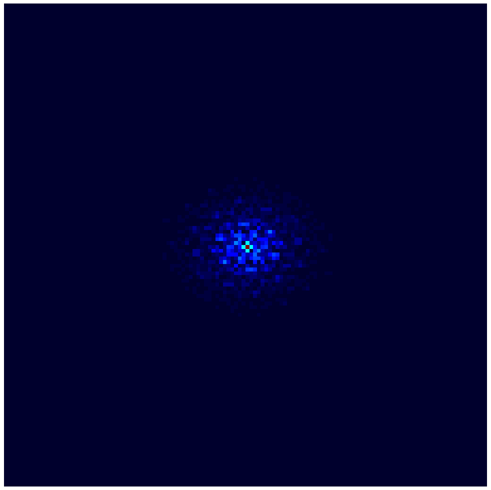
Outline

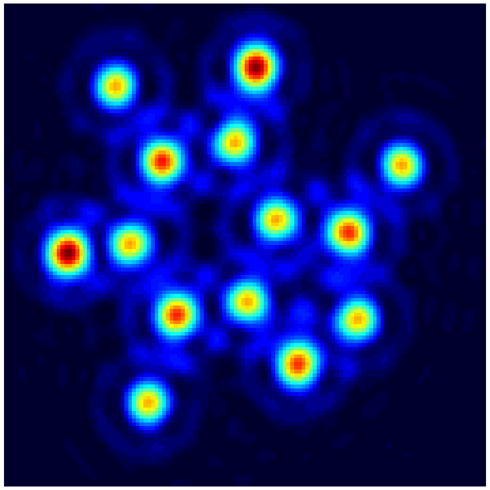
- Introduction
- MAXCUT formulation
- Tightness
- **Algorithms & Structure**
- Numerical Results

Sparsity: known support in 2D

- Molecular imaging: the samples are approximately sparse with **known support.**
- Most of the coefficient in b are close to zero.







Electronic density for the caffeine molecule (left), its 2D FFT transform (diffraction pattern, center), the density reconstructed using 3% of the coefficients at the core of the FFT (right).

Positivity

- We observe the magnitude of the Fourier transform of a discrete nonnegative signal $x \in \mathbb{R}^p$ so that
$$|\mathcal{F}x| = b$$
- We seek to reconstruct **positive** signals $x \geq 0$.
- This introduces additional **convex** restrictions on the phase vector u.

A function $f : \mathbb{R}^s \mapsto \mathbb{C}$ is *positive semidefinite* if and only if the matrix B with $B_{ij} = f(x_i - x_j)$ is Hermitian positive semidefinite for any sequence $x_i \in \mathbb{R}^s$.

Theorem

Bochner. *A function $f : \mathbb{R}^s \mapsto \mathbb{C}$ is positive semidefinite if and only if it is the Fourier transform of a (finite) nonnegative Borel measure.*

Positivity

- Reconstruct a phase vector $u \in \mathbb{C}^n$ such that $|u| = 1$ and

$$\mathcal{F}x = \mathbf{diag}(b)u.$$

In 1D (for simplicity), if we define the Toeplitz matrix

$$B_{ij}(y) = y_{|i-j|+1}, \quad i,j = 1,\ldots,p,$$

so that

$$B(y) = \begin{pmatrix} y_1 & y_2^* & & \cdots & & y_n^* \\ y_2 & y_1 & y_2^* & & \cdots & \\ & y_2 & y_1 & y_2^* & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \\ & \cdots & & y_2 & y_1 & y_2^* \\ y_n & & \cdots & & y_2 & y_1 \end{pmatrix}$$

- When $\mathcal{F}x = \mathbf{diag}(b)u$, Bochner's theorem means $B(\mathbf{diag}(b)u) \succeq 0$ iff $x \geq 0$.
- The contraint $B(\mathbf{diag}(b)u) \succeq 0$ is a **linear matrix inequality** in u, hence is convex.

Algorithms

PhaseCut is a complex semidefinite program, written

$$\begin{array}{ll} \text{Minimize} & \mathbf{Tr}(MU) \\ \text{such that} & \mathbf{diag}(U) = 1, \; X \succeq 0 \end{array}$$

where $U \in \mathbf{H}_n$ with $n = Jp$, where p is the size of the signal.

- The complexity of solving this SDP using the algorith in Helmberg et al. [1996] is
$$O\left(J^{3.5}\, p^{3.5} \log \frac{1}{\epsilon}\right) \quad \text{and} \quad O\left(K\, J^2 p^{4.5} \log \frac{1}{\epsilon}\right)$$
for *PhaseCut* and *PhaseLift* respectively.
- Solving a generic linear system is $O(p^3)$, solving a LP is $O(p^{3.5})\ldots$
- Using first-order solvers such as TFOCS [Becker et al., 2012], based on [Nesterov, 1983], the dependence on the dimension can be further reduced, to become
$$O\left(\frac{J^3\, p^3}{\epsilon}\right) \quad \text{and} \quad O\left(\frac{K J\, p^3}{\epsilon}\right)$$
for solving *PhaseCut* and *PhaseLift* respectively, serious impact on precision.

Algorithms

Block Coordinate Method. [Wen et al., 2009]

Input: An initial $X^0 = \mathbf{I}_n$ and $\nu > 0$ (typically small). An integer $N > 1$.

for $k = 1, \ldots, N$ **do**

Pick $i \in [1, n]$.

Compute

$$x = X^k_{i^c,i^c} M_{i^c,i} \quad \text{and} \quad \gamma = x^* M_{i^c,i}$$

If $\gamma > 0$, set

$$X^{k+1}_{i^c,i} = X^{k+1*}_{i,i^c} = -\sqrt{\frac{1-v}{\gamma}}\, x$$

else

$$X^{k+1}_{i^c,i} = X^{k+1*}_{i,i^c} = 0.$$

end for

Output: A matrix $X \succeq 0$ with $\mathbf{diag}(X) = 1$.

Writing i^c the index set $\{1, \ldots, i-1, i+1, \ldots, n\}$.

Algorithms

Complexity.

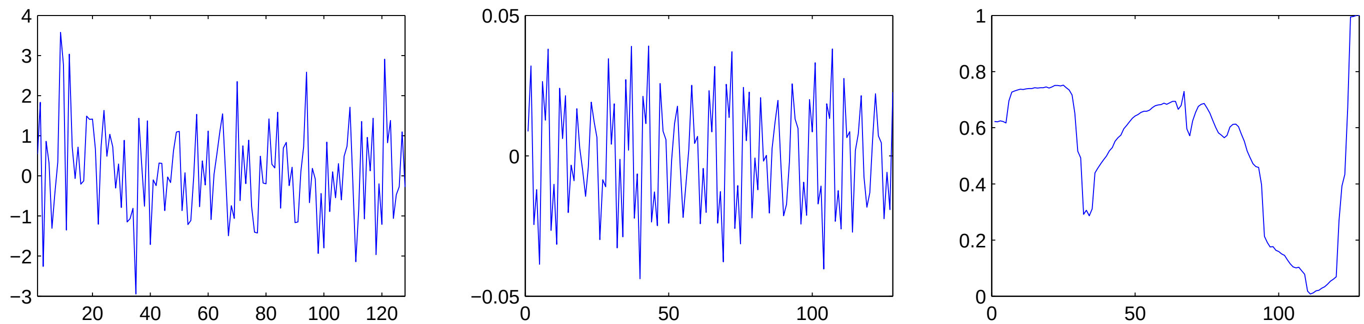
- Each iteration only requires matrix vector products $O(n^2)$.
- Cost per iteration very similar to the greedy algorithm by [Gerchberg and Saxton, 1972].
- In signal applications, the matrix vector product can be computed efficiently using the **FFT**, and the cost per iteration is reduced to $O(n \log n)$.

Outline

- Introduction
- MAXCUT formulation
- Tightness
- Algorithms & Structure
- **Numerical Results**

Numerical Experiments: 1D

- Three random signal classes: (a) Gaussian white noise. (b) Sum of 6 sinuoids of random frequency & random amplitudes. (c) Random scan-line of an image.



(a) (b) (c)

- The linear sampling operator A is an **oversampled Fourier transform**, multiple filterings with **random filters**, or a **wavelet transform**.
- We measure the error both in signal and in modulus

$$\epsilon(x,\tilde{x}) = \min_{c\in\mathbb{C},|c|=1} \frac{\|x - c\,\tilde{x}\|}{\|x\|} \qquad \text{and} \qquad \epsilon(|Ax|,|A\tilde{x}|) = \frac{\||Ax| - |A\tilde{x}|\|}{\|Ax\|}.$$

Numerical Experiments: 1D

	Fourier	Random Filters	Wavelets
Gerchberg-Saxton	5%	49%	0%
PhaseLift with reweighting	3%	100%	62%
PhaseCut	4%	100%	100%

Percentage of perfect reconstruction from $|Ax|$, over 300 test signals, for the three different operators A (columns) and the three algorithms (rows).

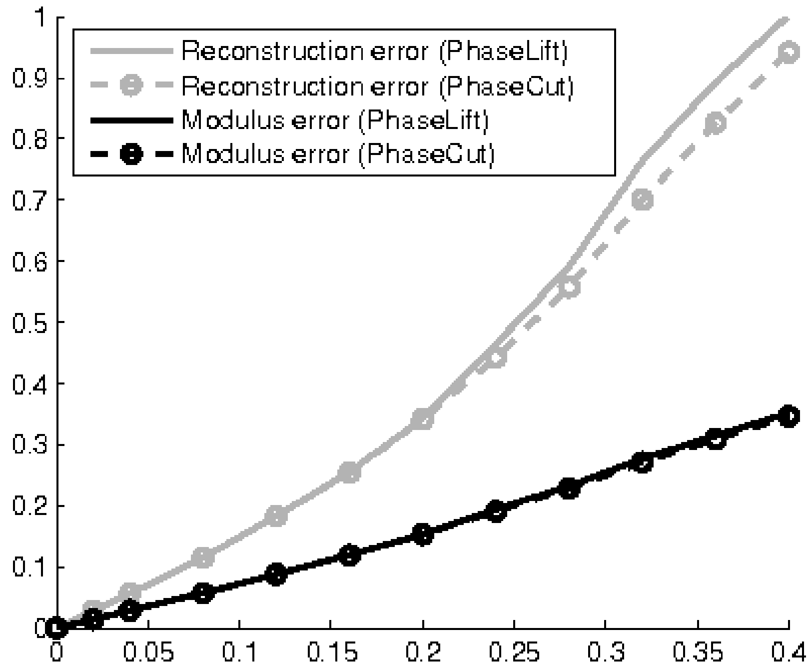
	Fourier	Random Filters	Wavelets
Gerchberg-Saxton	0.9	1.2	1.3
PhaseLift with reweighting	0.8	exact	0.5
PhaseCut	0.8	exact	exact

Average relative signal reconstruction error $\epsilon(\tilde{x}, x)$ over all test signals that are not perfectly reconstructed, for each operator A and each algorithm.

	Fourier	Random Filters	Wavelets
Gerchberg-Saxton	9.10^{-4}	0.2	0.3
PhaseLift with reweighting	5.10^{-4}	exact	8.10^{-2}
PhaseCut	6.10^{-4}	exact	exact

Average relative error $\epsilon(|A\tilde{x}|, |Ax|)$ of coefficient amplitudes, over all test signals that are not perfectly reconstructed, for each operator A and each algorithm.

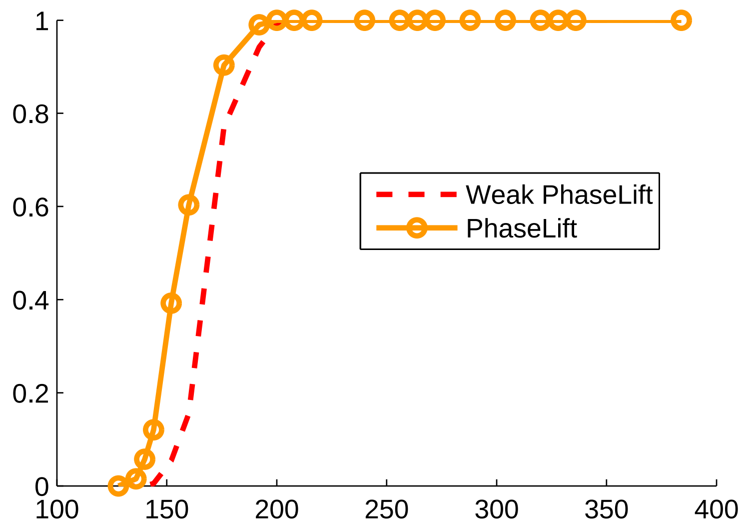
Numerical Experiments: 1D



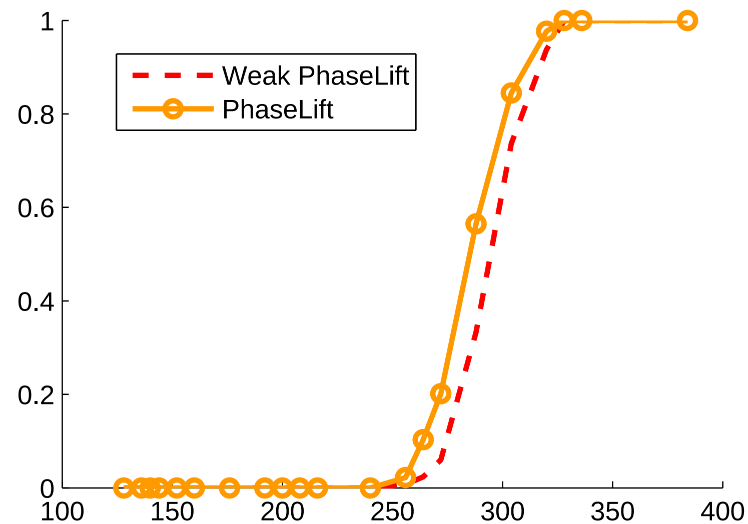
Mean performances of *PhaseLift* and *PhaseCut*, followed by some greedy iterations, for 4 gaussian random illumination filters. The x-axis represents the relative noise level, $||b_{\text{noise}}||_2/||Ax||_2$ and the y-axis the relative error on the result (signal and modulus).

Numerical Experiments: 1D

[Demanet and Hand, 2012] show that the solution to the relaxation is unique (trace minimization is unnecessary).



(a)

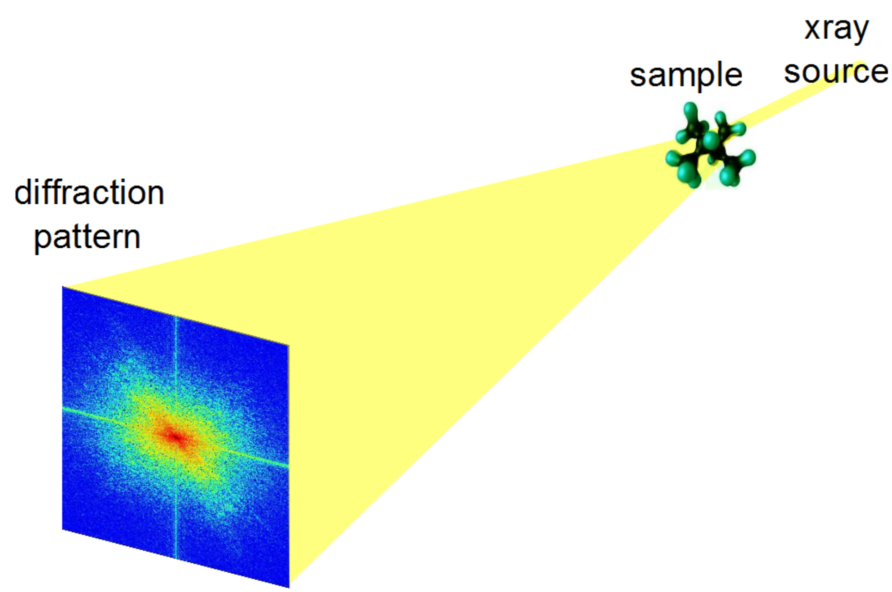


(b)

PhaseLift performance, for 64-sized signals, as a function of the number of measurements.
(a) Proportion of reconstructed signals, postprocessing using after GS iterations.
(b) Proportion of rank 1 (tight) solutions in the relaxation.

Numerical Experiments: 2D

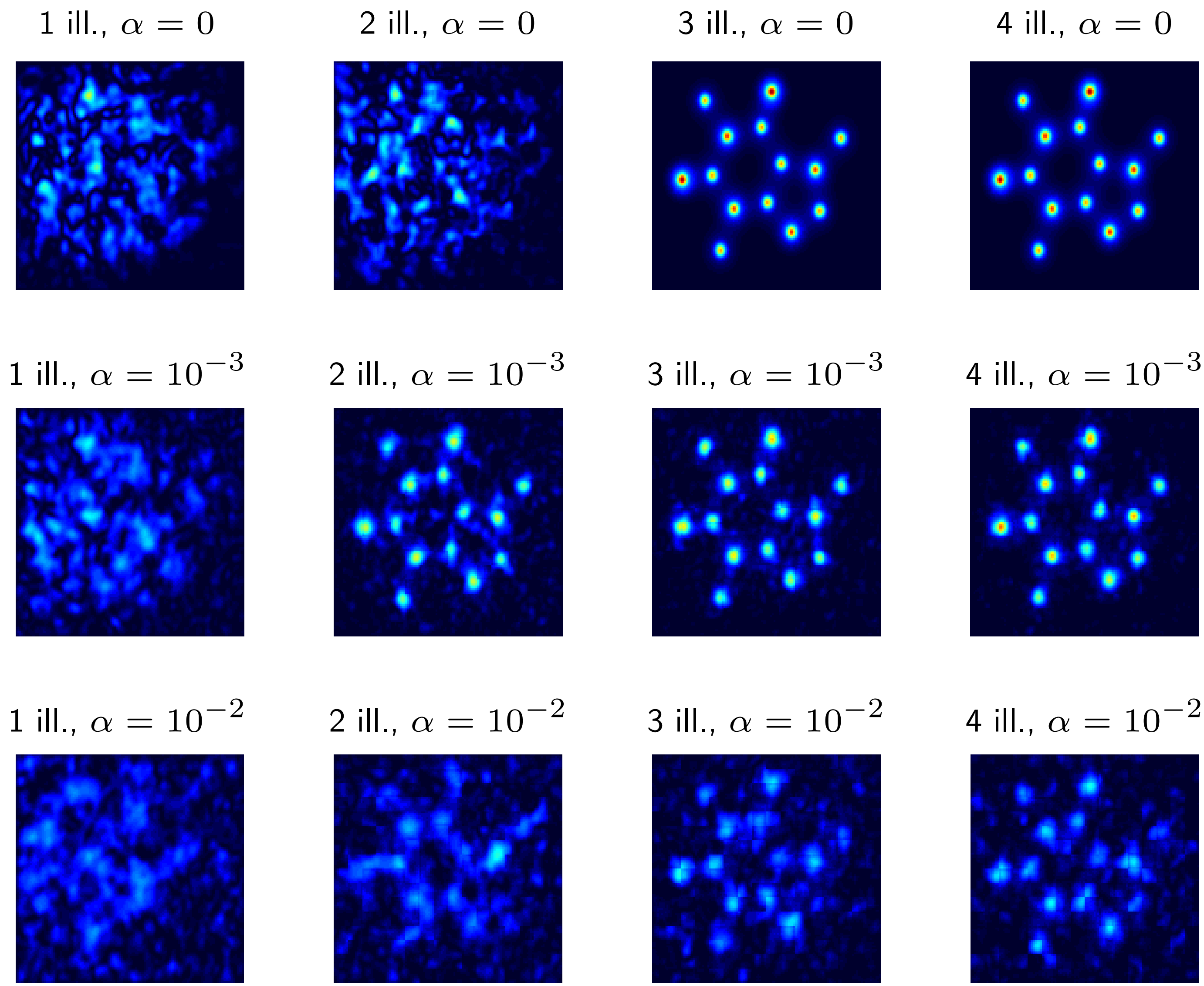
Applications in e.g. molecular imaging



(from [Candes et al., 2011b])

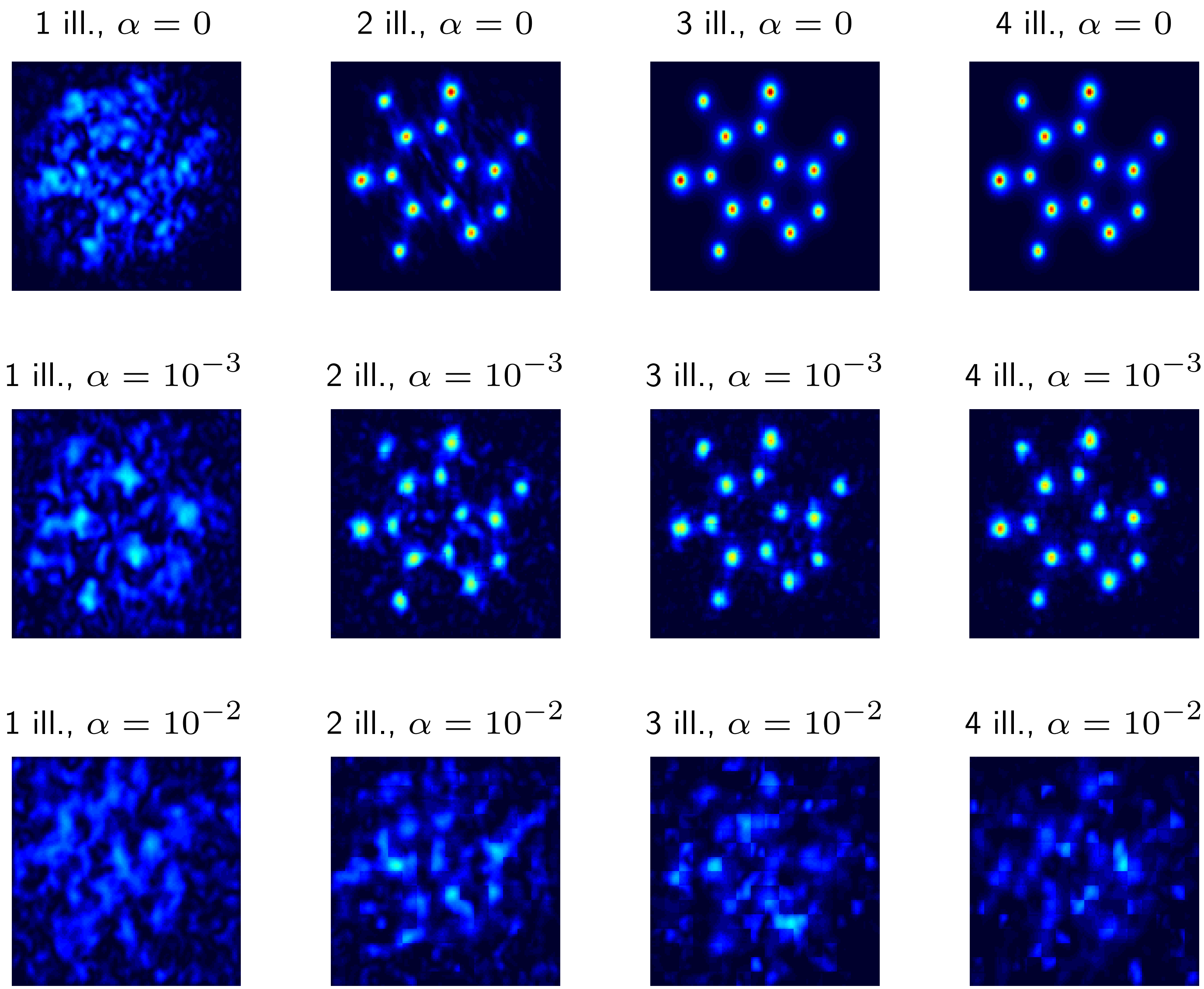
- CCD sensors only record the **magnitude** of diffracted rays, and loose the **phase**
- **Fraunhofer diffraction:** phase is required to invert the 2D Fourier transform
- Simulate diffraction using molecules from PDB and Poisson noise.

Numerical Experiments: 2D



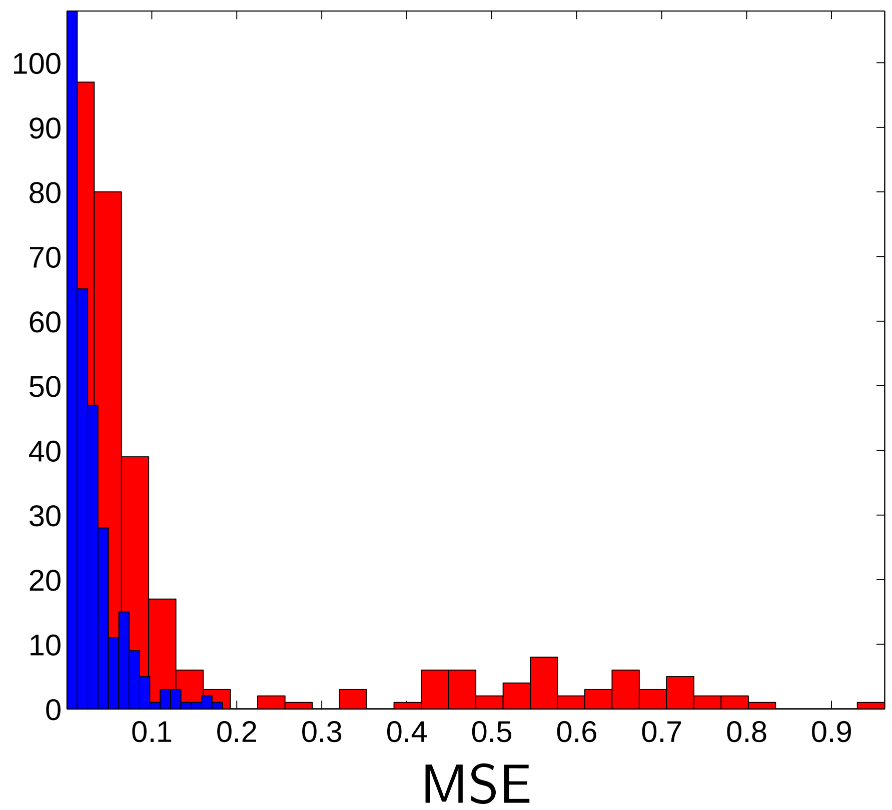
Solution of the greedy algorithm on caffeine molecule, for various values of the number of filters and noise level α.

Numerical Experiments: 2D



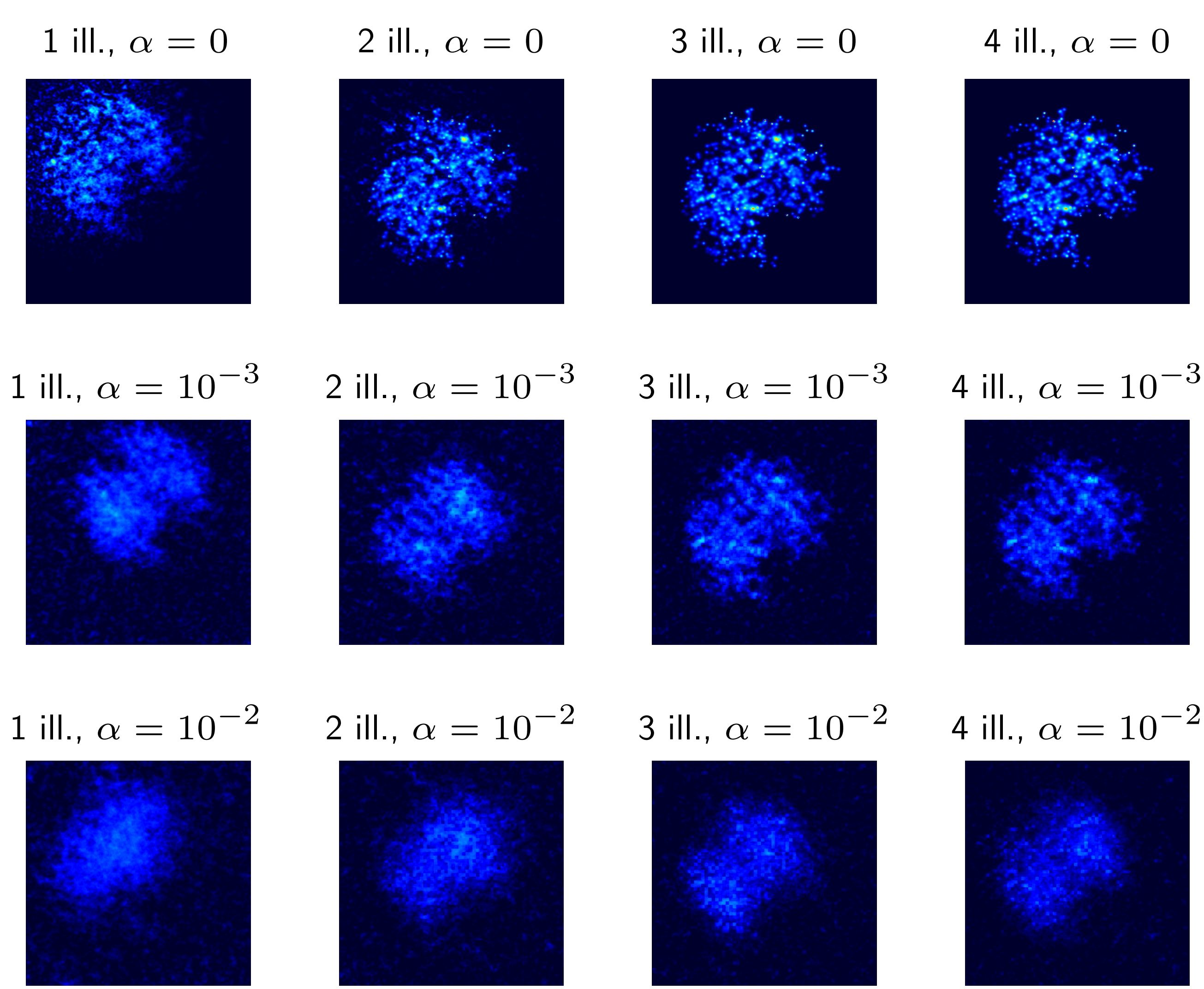
Solution of the semidefinite relaxation algorithm followed by greedy refinements, for various values of the number of filters and noise level α.

Numerical Experiments: 2D



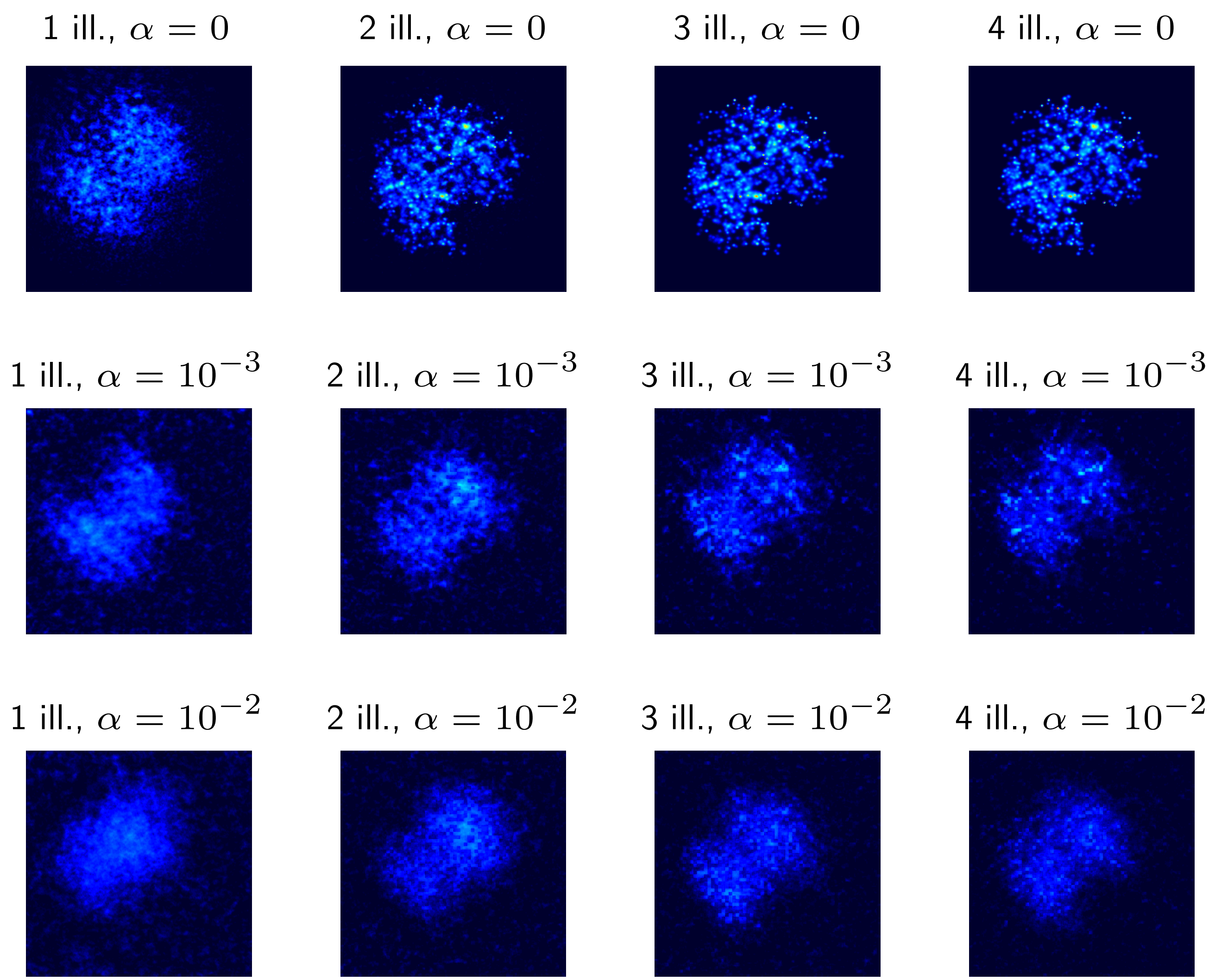
MSE between reconstructed image and true image for 2 illuminations without noise, using SDP then Fienup **(blue)**, and Fineup only **(red)**.

Numerical Experiments: 2D



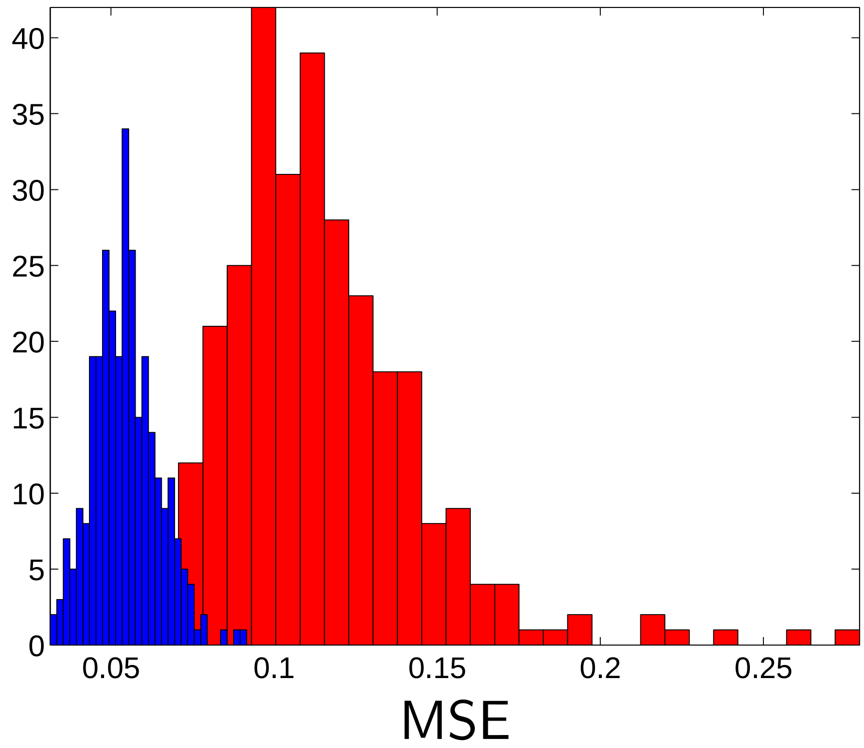
Solution of the greedy algorithm on 2LYZ (Lysozyne), for various values of the number of filters and noise level α.

Numerical Experiments: 2D



Solution of the semidefinite relaxation algorithm followed by greedy refinements on 2LYZ (Lysozyne), for various values of the number of filters and noise level α.

Numerical Experiments: 2D



MSE between reconstructed image and true image for 2 illuminations of 2LYZ without noise, using SDP then Fienup **(blue)**, and Fineup only **(red)**.

Conclusion

- Write the phase recovery problem as a MAXCUT like problem.
- Tightness properties equivalent to the matrix completion approach.
- Very fast/scalable algorithms.

Open questions. . . .

- Tightness results in the noisy case, or in the positive case?
- Is the SDP relaxation optimal?

*

References

N. Alon and A. Naor. Approximating the cut-norm via Grothendieck's inequality. In *Proceedings of the thirty-sixth annual ACM symposium on Theory of computing*, pages 72–80. ACM, 2004.

S. Becker, E.J. Candes, and M. Grant. Tfocs v1. 1 user guide. 2012.

O. Bunk, A. Diaz, F. Pfeiffer, C. David, B. Schmitt, D.K. Satapathy, and JF Veen. Diffractive imaging for periodic samples: retrieving one-dimensional concentration profiles across microfluidic channels. *Acta Crystallographica Section A: Foundations of Crystallography*, 63 (4):306–314, 2007.

E. J. Candes, T. Strohmer, and V. Voroninski. Phaselift : exact and stable signal recovery from magnitude measurements via convex programming. *To appear in Communications in Pure and Applied Mathematics*, 2011a.

E.J. Candes and B. Recht. Exact matrix completion via convex optimization. *preprint*, 2008.

E.J. Candes and T. Tao. The power of convex relaxation: Near-optimal matrix completion. *Information Theory, IEEE Transactions on*, 56(5): 2053–2080, 2010.

E.J. Candes, Y. Eldar, T. Strohmer, and V. Voroninski. Phase retrieval via matrix completion. *Arxiv preprint arXiv:1109.0573*, 2011b.

A. Chai, M. Moscoso, and G. Papanicolaou. Array imaging using intensity-only measurements. *Inverse Problems*, 27:015005, 2011.

L. Demanet and P. Hand. Stable optimizationless recovery from phaseless linear measurements. *Arxiv preprint arXiv:1208.1803*, 2012.

J.R. Fienup. Phase retrieval algorithms: a comparison. *Applied Optics*, 21(15):2758–2769, 1982.

R. Gerchberg and W. Saxton. A practical algorithm for the determination of phase from image and diffraction plane pictures. *Optik*, 35: 237–246, 1972.

M.X. Goemans and D. Williamson. Approximation algorithms for max-3-cut and other problems via complex semidefinite programming. In *Proceedings of the thirty-third annual ACM symposium on Theory of computing*, pages 443–452. ACM, 2001.

M.X. Goemans and D.P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. ACM*, 42:1115–1145, 1995.

D. Griffin and J. Lim. Signal estimation from modified short-time fourier transform. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 32(2):236–243, 1984.

R.W. Harrison. Phase problem in crystallography. *JOSA A*, 10(5):1046–1055, 1993.

C. Helmberg, F. Rendl, R. J. Vanderbei, and H. Wolkowicz. An interior–point method for semidefinite programming. *SIAM Journal on Optimization*, 6:342–361, 1996.

M. Kisialiou and Z.Q. Luo. Probabilistic analysis of semidefinite relaxation for binary quadratic minimization. *SIAM Journal on Optimization*, 20:1906, 2010.

L. Lovász and A. Schrijver. Cones of matrices and set-functions and 0-1 optimization. *SIAM Journal on Optimization*, 1(2):166–190, 1991.

Z.Q. Luo, X. Luo, and M. Kisialiou. An efficient quasi-maximum likelihood decoder for psk signals. In *Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03). 2003 IEEE International Conference on*, volume 6, pages VI–561. IEEE, 2003.

M. Mezard and A. Montanari. *Information, physics, and computation*. Oxford University Press, USA, 2009.

J. Miao, T. Ishikawa, Q. Shen, and T. Earnest. Extending x-ray crystallography to allow the imaging of noncrystalline materials, cells, and single protein complexes. *Annu. Rev. Phys. Chem.*, 59:387–410, 2008.

A.S. Nemirovski. *Computation of matrix norms with applications to Robust Optimization*. PhD thesis, Technion, 2005.

Y. Nesterov. A method of solving a convex programming problem with convergence rate $O(1/k^2)$. *Soviet Mathematics Doklady*, 27(2): 372–376, 1983.

Y. Nesterov. *Global quadratic optimization via conic relaxation*. Number 9860. CORE Discussion Paper, 1998.

B. Recht, M. Fazel, and P.A. Parrilo. Guaranteed Minimum-Rank Solutions of Linear Matrix Equations via Nuclear Norm Minimization. *Arxiv preprint arXiv:0706.4138*, 2007.

H. Sahinoglou and S.D. Cabrera. On phase retrieval of finite-length sequences using the initial time sample. *Circuits and Systems, IEEE Transactions on*, 38(8):954–958, 1991.

N.Z. Shor. Quadratic optimization problems. *Soviet Journal of Computer and Systems Sciences*, 25:1–11, 1987.

A.M.C. So. Non–asymptotic performance analysis of the semidefinite relaxation detector in digital communications. 2010.

I. Waldspurger, A. d'Aspremont, and S. Mallat. Phase recovery, maxcut and complex semidefinite programming. *ArXiv: 1206.0102*, 2012.

Z. Wen, D. Goldfarb, S. Ma, and K. Scheinberg. Row by row methods for semidefinite programming. Technical report, Technical report, Department of IEOR, Columbia University, 2009.

S. Zhang and Y. Huang. Complex quadratic optimization and semidefinite programming. *SIAM Journal on Optimization*, 16(3):871–890, 2006.