Phase Retrieval,

New Results on an Old Problem.

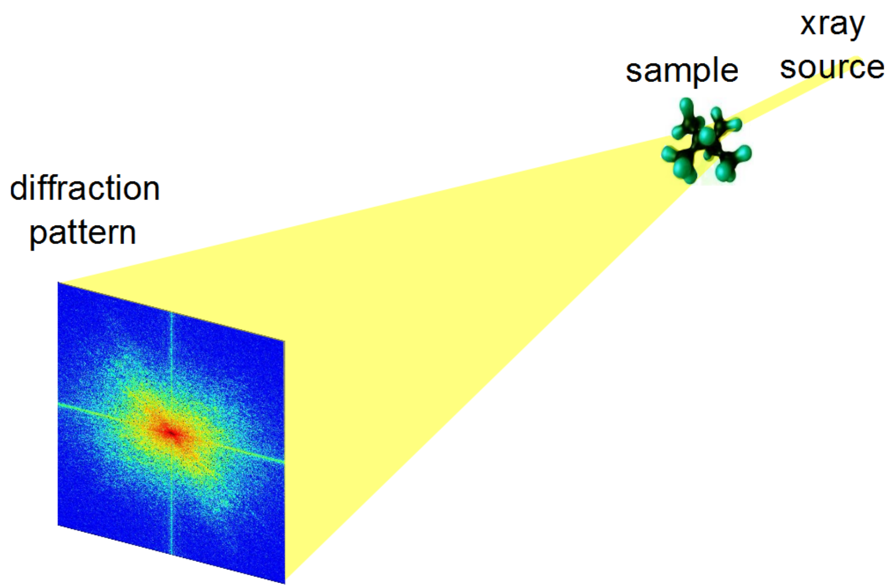
Alex d'Aspremont, *CNRS & École Normale Supérieure, Paris,*

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

Support from ERC (project SIPA).

Introduction: diffraction imaging

Diffraction imaging



[Candes et al., 2011]

- Sensors only record the **magnitude** of diffracted rays, and lose the **phase**.
- The phase is required to invert the 2D Fourier transform and reconstruct the sample density.

Introduction: phasing

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

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

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

- Reconstruct a signal x from the **amplitude of n linear measurements A**.
- Easy to write, very **hard to solve** in general.
- We seek a **tractable** procedure, i.e. a polynomial time algorithm with explicit approximation and complexity guarantees.

Introduction: efficiency & stability

We want **more than uniqueness** of the solution.

- A **tractable** algorithm to solve the phasing problem in polynomial-time.
- A solution that is **stable** and **robust** to noise.

For certain measurement matrices A, this is indeed possible. . .

Introduction

Greedy algorithm [Gerchberg and Saxton, 1972], find $y = Ax$ given $b = |Ax|$

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

$$w = AA^\dagger y^k, \quad \textbf{(project } y \textbf{ on } \mathcal{R}(A)\textbf{.)}$$

3: Set

$$y_i^{k+1} = b_i \frac{w}{|w|}, \quad \textbf{(match } |y| \textbf{ with } b\textbf{.)}$$

4: **end for**

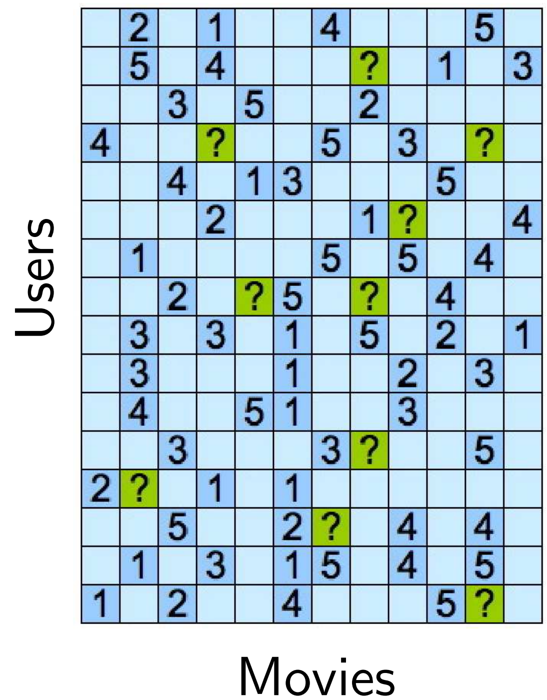
Output: $y_N \in \mathbb{C}^n$.

Similar to **alternating projections**. Sometimes it works, sometimes it doesn't. . .

Can we do better?

NETFLIX

Given user ratings



Make **personalized** recommendations for other movies. . .

Introduction: collaborative prediction

- A **linear prediction** model

$$\text{rating}_{ij} = u_i^T v_j$$

where u_i represents user characteristics and v_j movie features.

- Collaborative prediction is a **matrix factorization** problem

$$M = U^T V$$

$U \in \mathbb{R}^{n \times k}$ user types, $V \in \mathbb{R}^{k \times m}$ movie features, $M \in \mathbb{R}^{n \times m}$ ratings.

- Assume M is **low rank**.

Introduction: matrix completion

Matrix completion. [Recht et al., 2007, Candes and Recht, 2008, Candes and Tao, 2010].

- The **NETFLIX** problem can be written as

$$\begin{array}{ll} \text{Minimize} & \mathbf{Rank}(X) \\ \text{subject to} & \mathbf{Tr}(A_i X) = b_i, \quad i = 1, \ldots, n \\ & X \succeq 0 \end{array}$$

- For certain matrices A_i, it suffices to solve

$$\begin{array}{ll} \text{Minimize} & \mathbf{Tr}(X) \\ \text{subject to} & \mathbf{Tr}(A_i X) = b_i, \quad i = 1, \ldots, n \\ & X \succeq 0 \end{array}$$

which is a **convex problem** in $X \in \mathbf{S}_n$.

Introduction: phase retrieval as a SDP

- [Chai et al., 2011, Candes et al., 2013a], **lifting** technique from [Shor, 1987]

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

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

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

- [Candes, Strohmer, and Voroninski, 2013a] show that under certain conditions on A and x_0, it suffices to solve

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

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

Introduction

A very sparse (and incomplete) list of references. . .

Algorithms

- Greedy algorithm [Gerchberg and Saxton, 1972]
- Classical survey of early algorithms by [Fienup, 1982].
- NP-complete [Sahinoglou and Cabrera, 1991].
- Many algorithms. [Miao et al., 1998, Bauschke et al., 2002, Luke, 2005].
- Matrix completion formulation [Chai, Moscoso, and Papanicolaou, 2011] and [Candes, Strohmer, and Voroninski, 2013a]

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].

Outline

- Introduction
- **Algorithms**
- Exploiting structure
- Numerical results
- Experimental setup?

Introduction: semidefinite programming

A **linear program** (LP) is written

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array}$$

where $x \geq 0$ means that the coefficients of the vector x are nonnegative.

A **semidefinite program** (SDP) is written

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}(CX) \\ \text{subject to} & \mathbf{Tr}(A_i X) = b_i, \quad i = 1, \ldots, m \\ & X \succeq 0 \end{array}$$

where $X \succeq 0$ means that the matrix variable $X \in \mathbf{S}_n$ is **positive semidefinite**.

- Nesterov and Nemirovskii [1994] showed that the **interior point algorithms** used for linear programs could be extended to semidefinite programs.
- Efficient solvers, many (unexpected) applications.

Phase problem in phase

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

- $Ax = \mathbf{diag}(b)u$ 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 $M = \mathbf{diag}(b)(\mathbf{I} - AA^\dagger)\,\mathbf{diag}(b) \succeq 0$.

Tightness

Exact phase reconstruction in polynomial-time.

- [Candes et al., 2013a,b] 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 $n = O(p)$ observations a_i picked randomly (sphere or coded Fourier).

- [Waldspurger, d'Aspremont, and Mallat, 2012] Semidefinite relaxation for phase recovery, called **PhaseCut**.

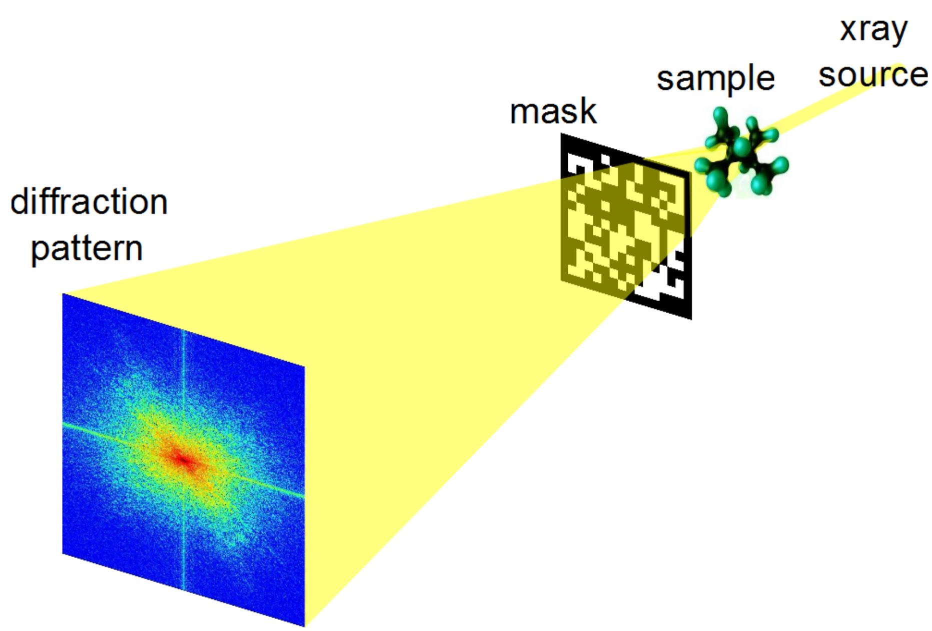
$$\begin{array}{ll} \text{Minimize} & \mathbf{Tr}(MU) \\ \text{such that} & \mathbf{diag}(U) = 1, \; U \succeq 0 \end{array}$$

  similar to MAXCUT relaxation.

- [Waldspurger et al., 2012] show PhaseCut is tight when PhaseLift is.

Which observations A?

[Candes et al., 2013b]: The observations A are constructed from **multiple** coded diffraction patterns



More on this later. . .

Block Coordinate Method. PhaseCut & MAXCUT

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

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

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

3: Compute

$$\mathbf{u} = \mathbf{U^k_{i^c,i^c} M_{i^c,i}} \quad \text{and} \quad \gamma = u^* M_{i^c,i}$$

4: If $\gamma > 0$, set

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

else

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

5: **end for**

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

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

Algorithms

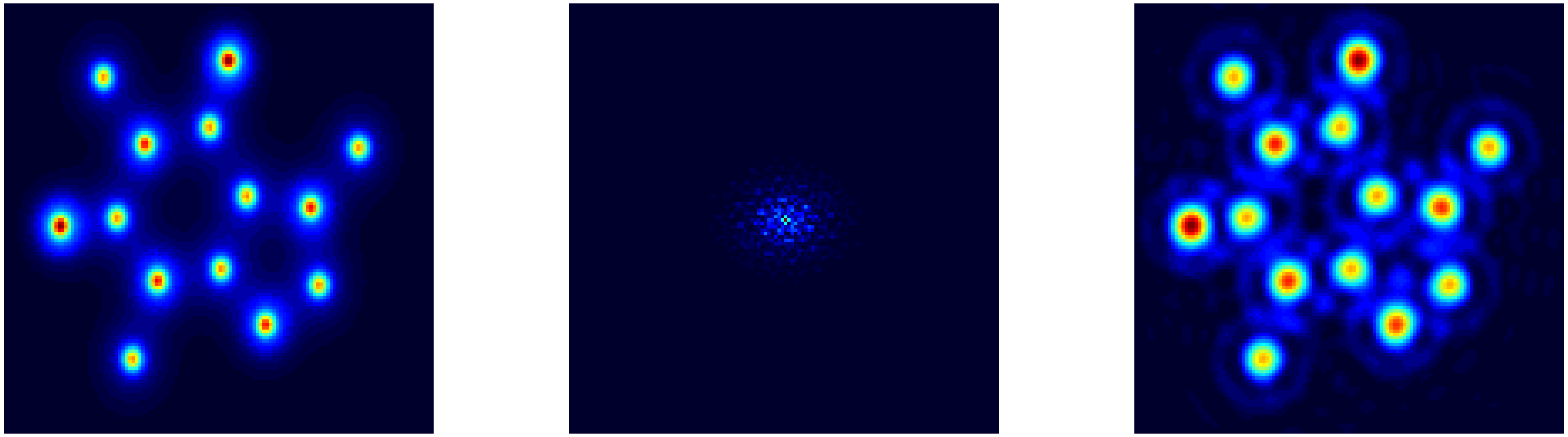
Complexity.

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

Outline

- Introduction
- Algorithms
- **Exploiting structure**
- Numerical results
- Experimental setup?

Sparsity: known support in 2D



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

- Molecular imaging: data is **sparse with known support.**
- Most coefficients in b close to zero, so **most coefficients in u can be set to zero** in

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

which means significant computational savings.

Positivity

- We observe the magnitude of the Fourier transform of a discrete signal $x \in \mathbb{R}^p$

$$|\mathcal{F}x| = b$$

- We seek to reconstruct **positive signals $x \geq 0$.**

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)

Fourier on positive signals. *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

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

- We define the Toeplitz matrix $B_{ij}(y) = y_{|i-j|+1},\ i,j = 1,\ldots,p$, so that

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

- **Bochner's theorem.**

$$x \geq 0 \quad \Longleftrightarrow \quad B(\mathbf{diag}(b)u) \succeq 0,$$

which is a (convex) **linear matrix inequality in u.**

Real signals

Real valued signal. Phase problem on real valued signal is

$$\begin{array}{ll} \text{minimize} & \left\| \mathcal{T}(A) \begin{pmatrix} x \\ 0 \end{pmatrix} - \mathbf{diag}\begin{pmatrix} b \\ b \end{pmatrix} \begin{pmatrix} \Re(u) \\ \Im(u) \end{pmatrix} \right\|_2^2 \\ \text{subject to} & u \in \mathbb{C}^n,\ |u_i| = 1 \\ & x \in \mathbb{R}^p. \end{array}$$

Here $x = A_2^\dagger B_2 v$, where

$$A_2 = \begin{pmatrix} \Re(A) \\ \Im(A) \end{pmatrix}, \quad B_2 = \mathbf{diag}\begin{pmatrix} b \\ b \end{pmatrix}, \quad \text{and} \quad v = \begin{pmatrix} \Re(u) \\ \Im(u) \end{pmatrix}$$

the phase problem is equivalent to

$$\begin{array}{ll} \text{minimize} & \|(A_2 A_2^\dagger B_2 - B_2)v\|_2^2 \\ \text{subject to} & v_i^2 + v_{n+i}^2 = 1, \quad i = 1, \ldots, n, \end{array}$$

in the variable $v \in \mathbb{R}^{2n}$.

Real signals

Real valued signal. The last problem can be relaxed as

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}(VM_2) \\ \text{subject to} & V_{ii} + V_{n+i,n+i} = 1, \quad i = 1, \ldots, n, \\ & V \succeq 0, \end{array}$$

which is a semidefinite program in the variable $V \in \mathbf{S}_{2n}$, where

$$M_2 = (A_2 A_2^\dagger B_2 - B_2)^T (A_2 A_2^\dagger B_2 - B_2) = B_2^T (\mathbf{I} - A_2 A_2^\dagger) B_2.$$

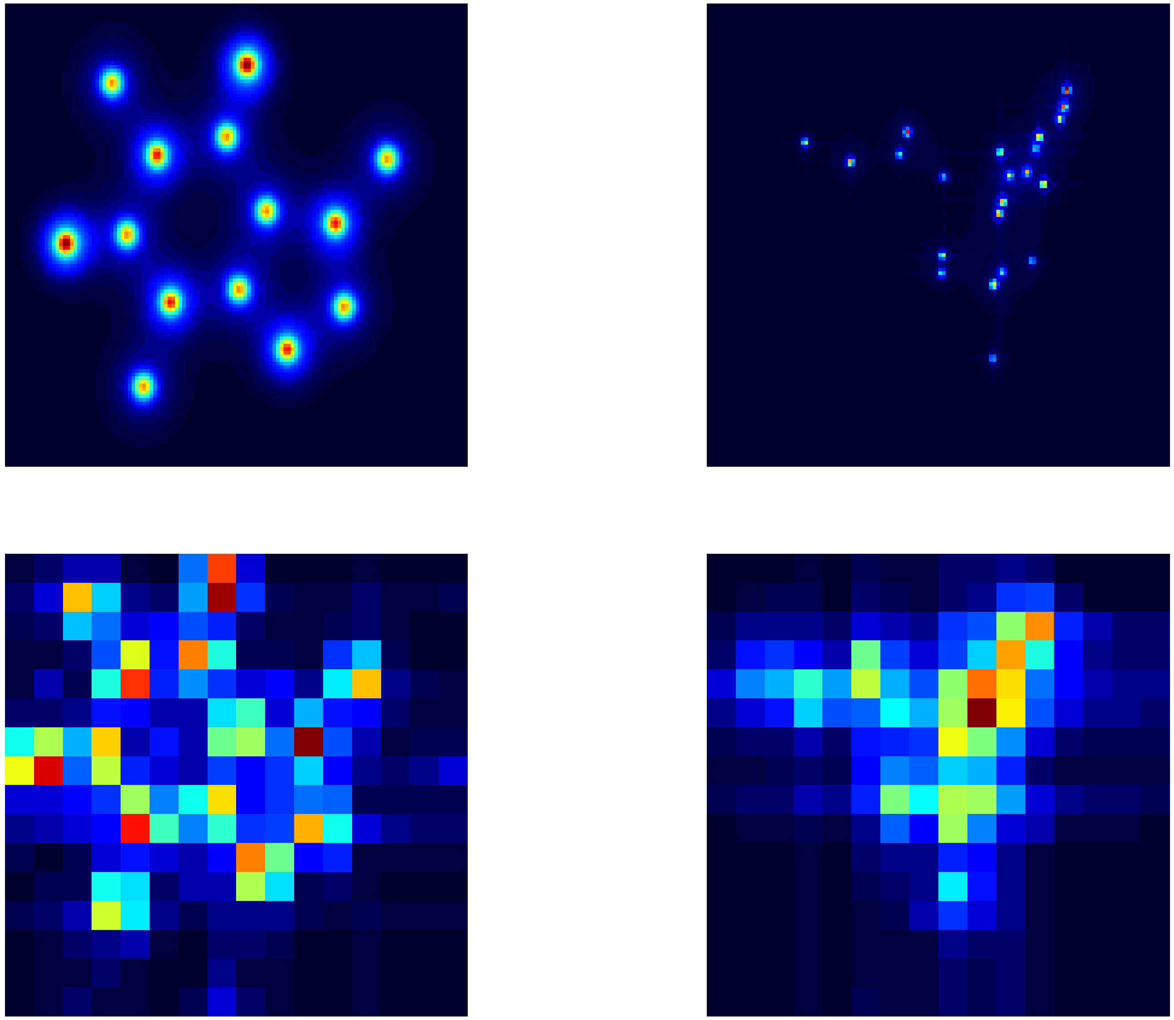
- Explicitly **constrains the solution x to be real valued.**
- Small increase in complexity.

Outline

- Introduction
- Algorithms
- Exploiting structure
- **Numerical results**
- Experimental setup?

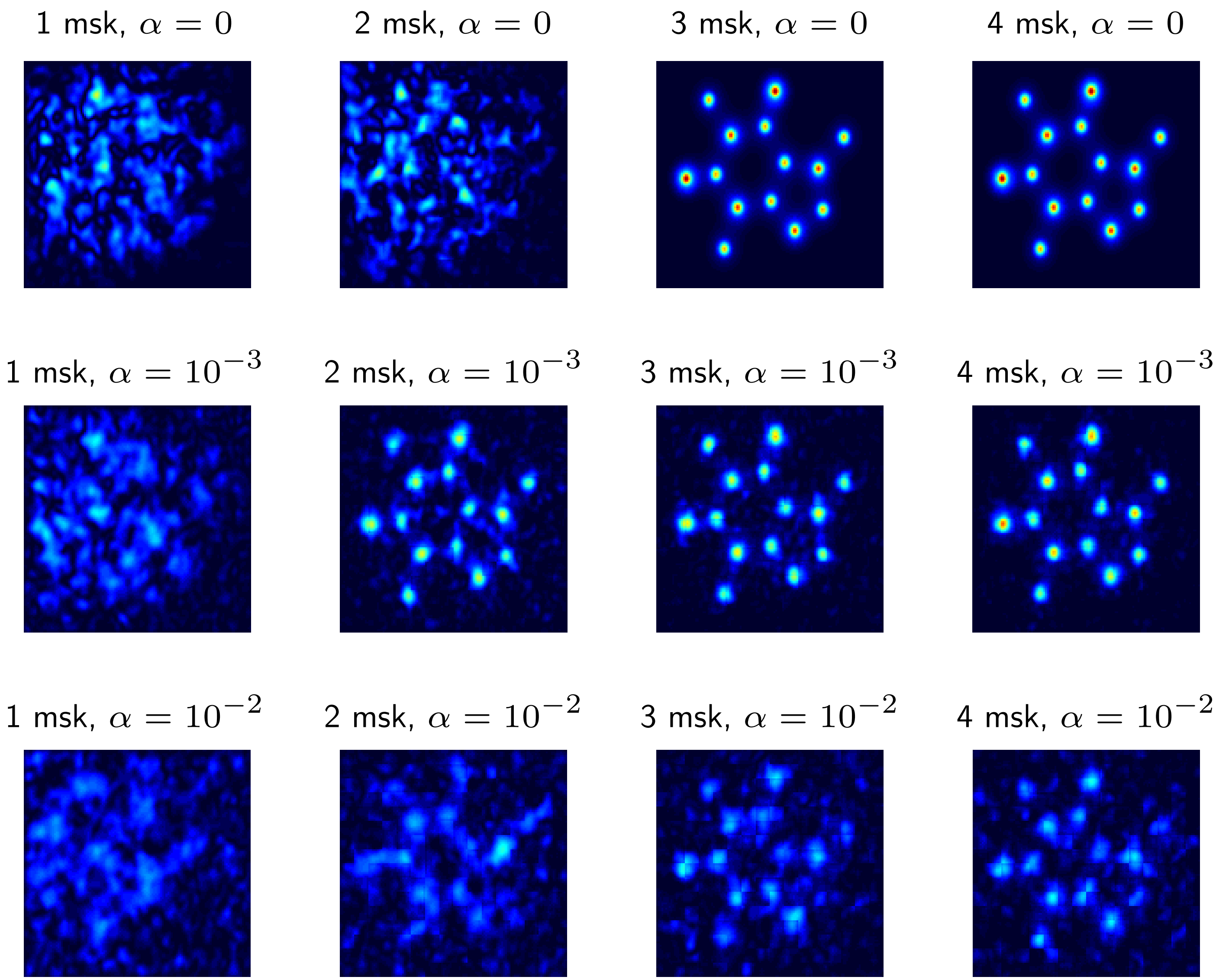
Numerical Experiments: PDB molecules

Two molecules, two resolutions: 16x16 and 128x128.



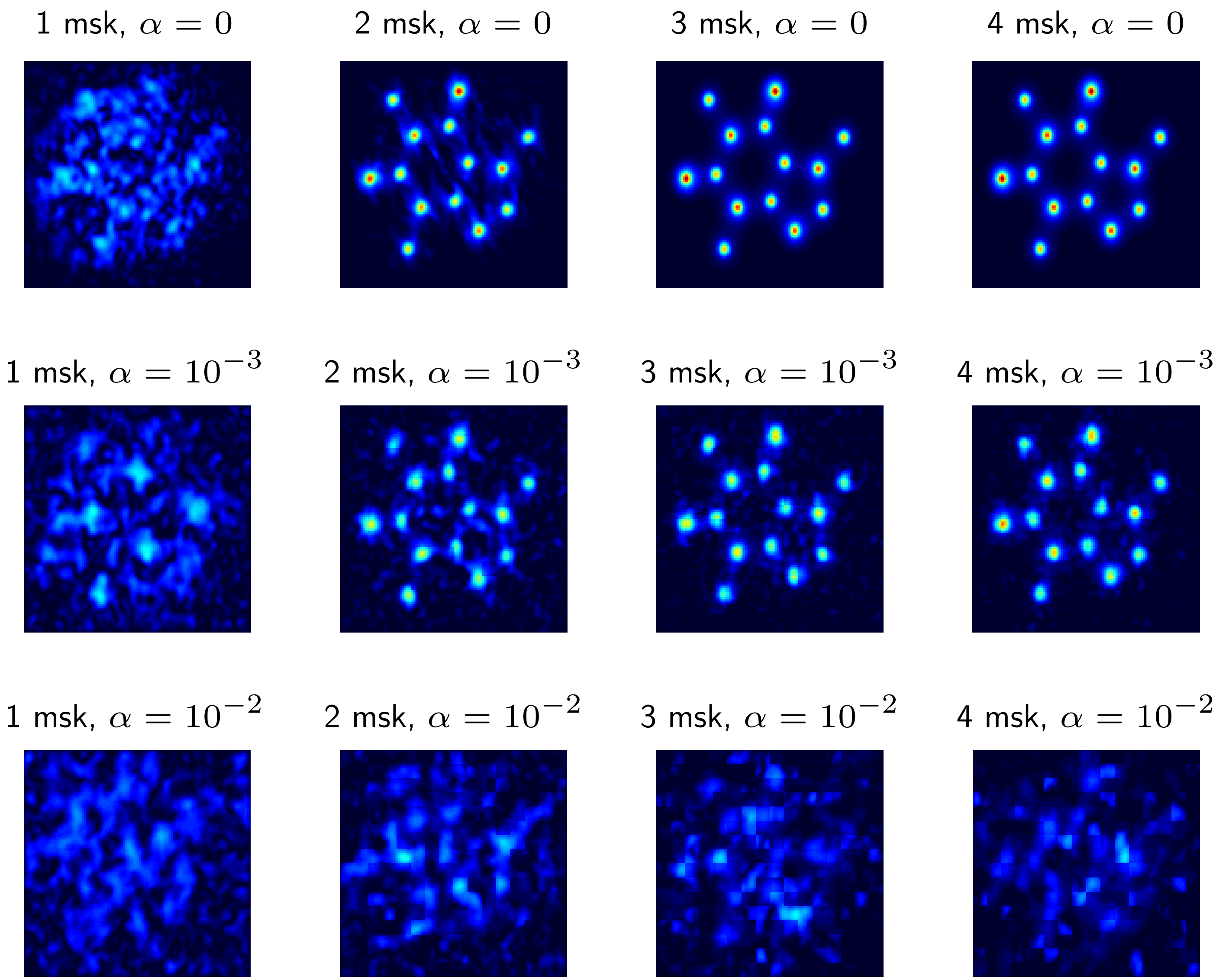
Caffeine **Cocaine**

Numerical Experiments: PDB molecules



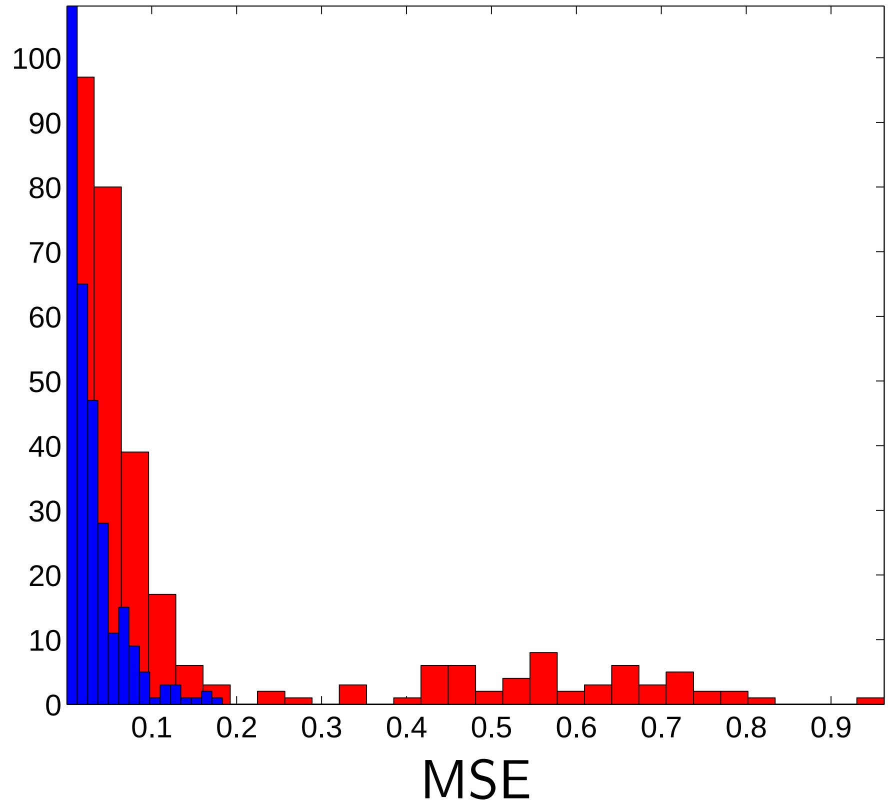
Solution of the **greedy algorithm** on caffeine molecule, for various values of the **number of masks** and **noise level** α.

Numerical Experiments: 2D



Solution of the **PhaseCut SDP** followed by greedy refinements, for various values of the **number of masks** and **noise level** α.

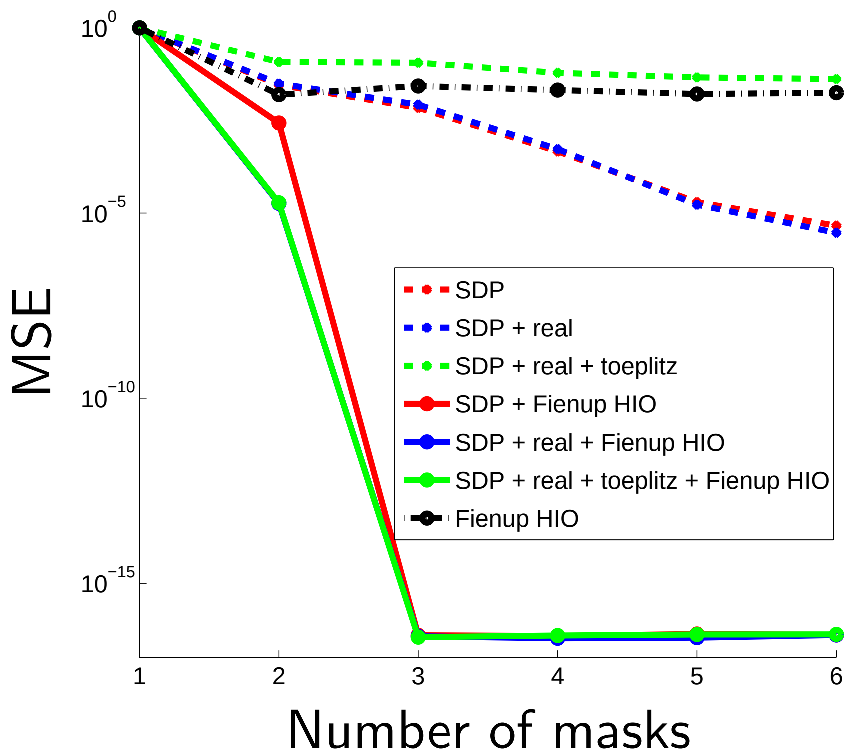
Numerical Experiments: 2D

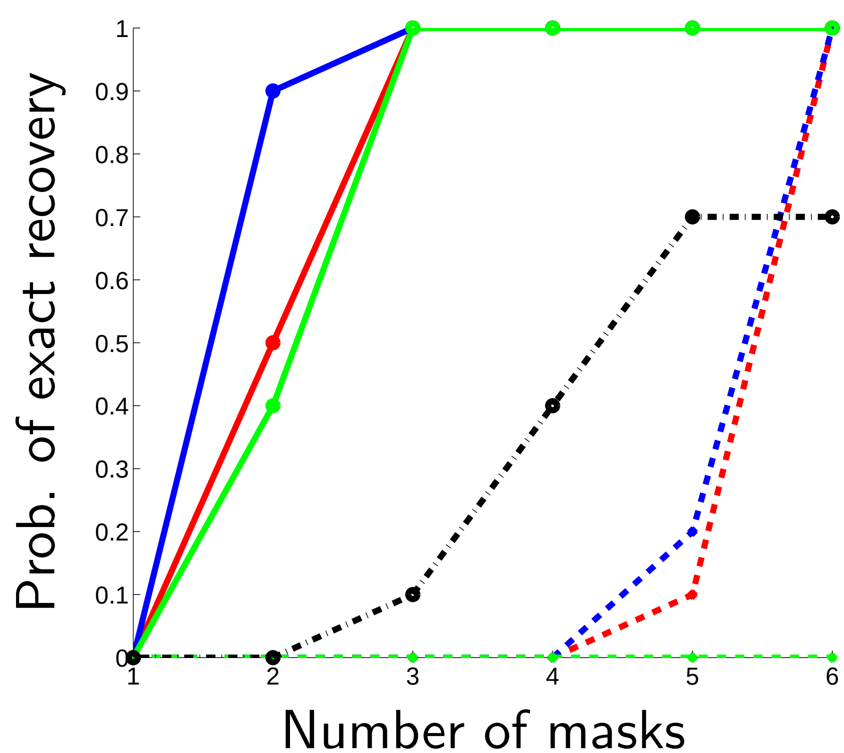


MSE between reconstructed image and true image for **2 random illuminations** without noise, using **SDP then Fienup (blue)**, and **Fienup only (red)**.

Numerical Experiments: comparing algorithms

16x16 caffeine image. No oversampling.

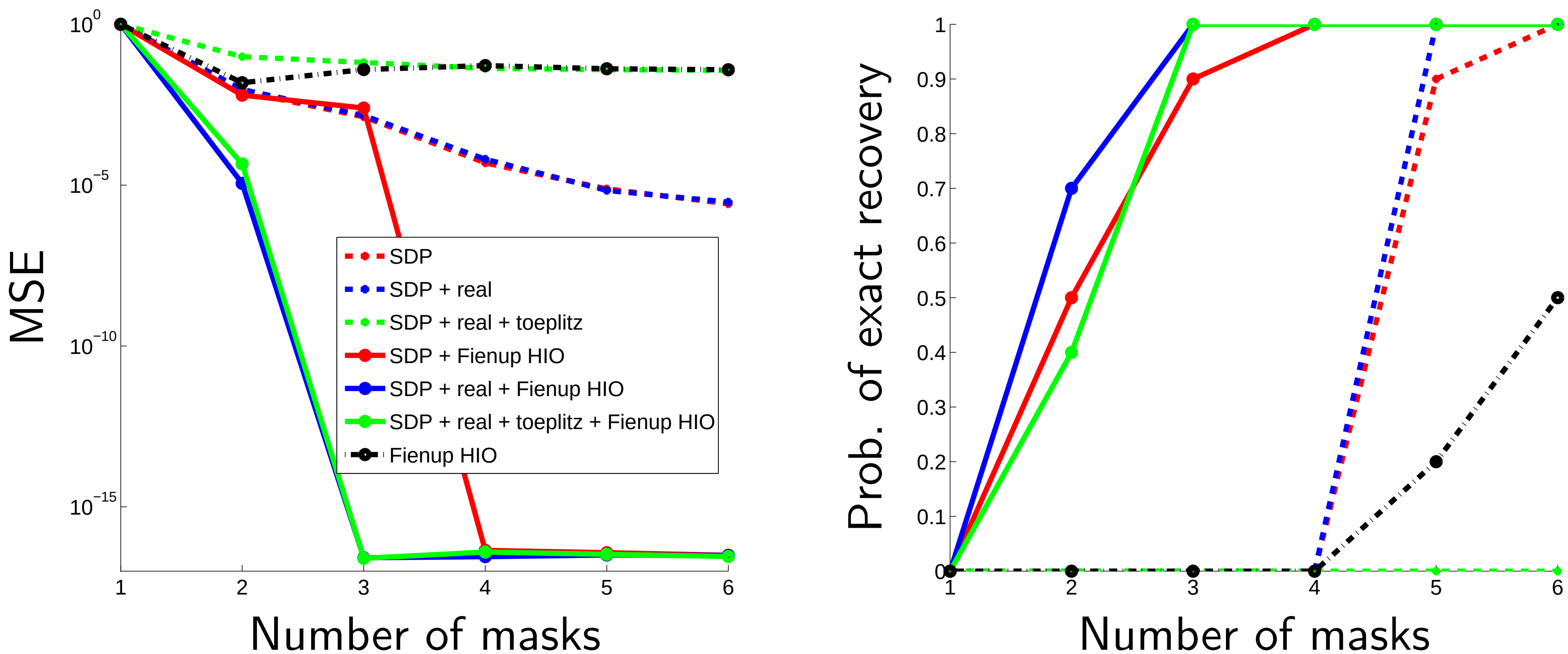




Left: MSE (relative to b) vs. number of random masks.
Right: Probability of recovering molecular density ($MSE < 10^{-4}$) vs. number of random masks.

Numerical Experiments: comparing algorithms

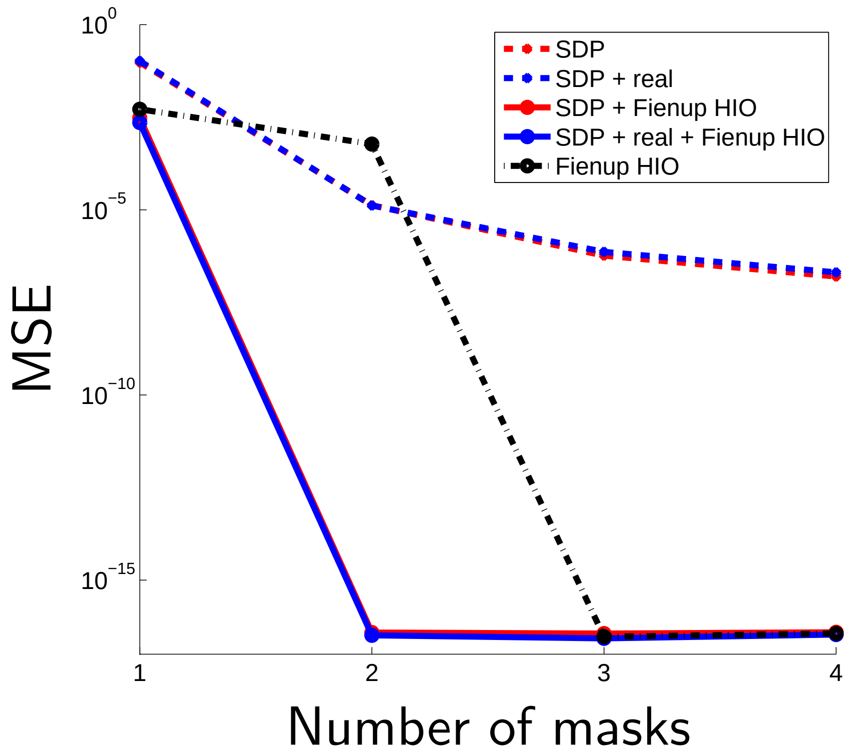
16x16 cocaine image. **No oversampling.**

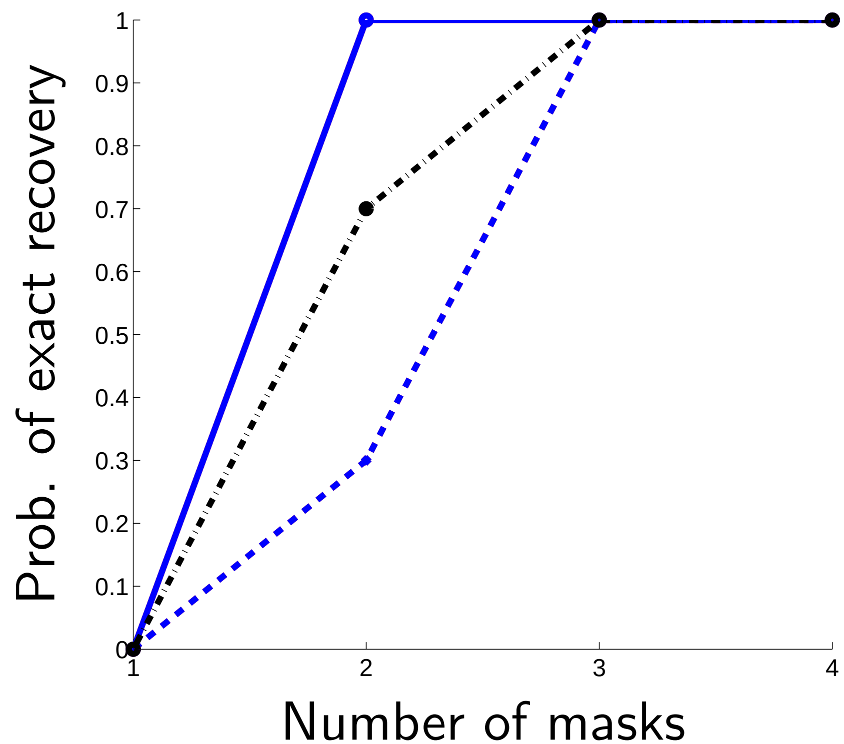


Left: MSE (relative to b) vs. number of random masks.
Right: Probability of recovering molecular density ($MSE < 10^{-4}$) vs. number of random masks.

Numerical Experiments: comparing algorithms

16x16 caffeine image. **2x oversampling.**

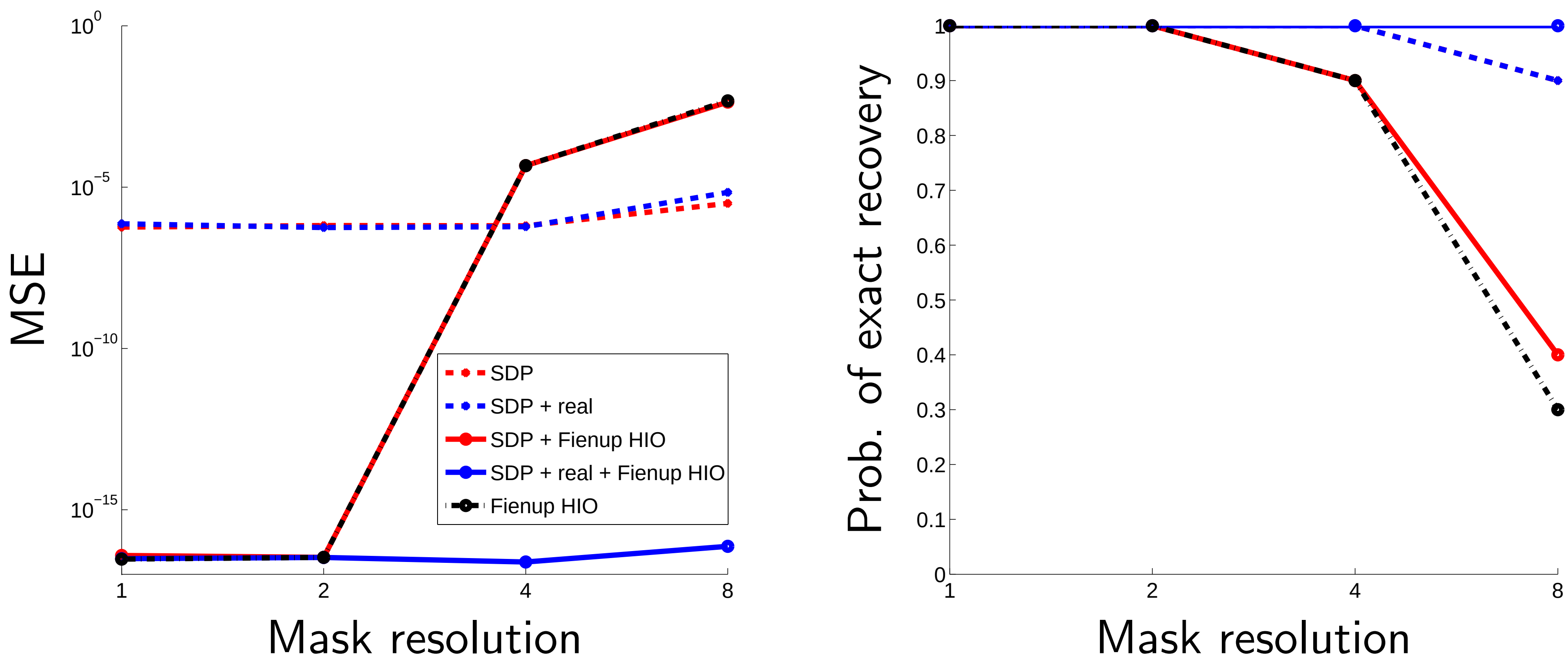




Left: MSE vs. number of random masks.
Right: Probability of recovering molecular density ($MSE < 10^{-4}$) vs. number of random masks.

Numerical Experiments: comparing algorithms

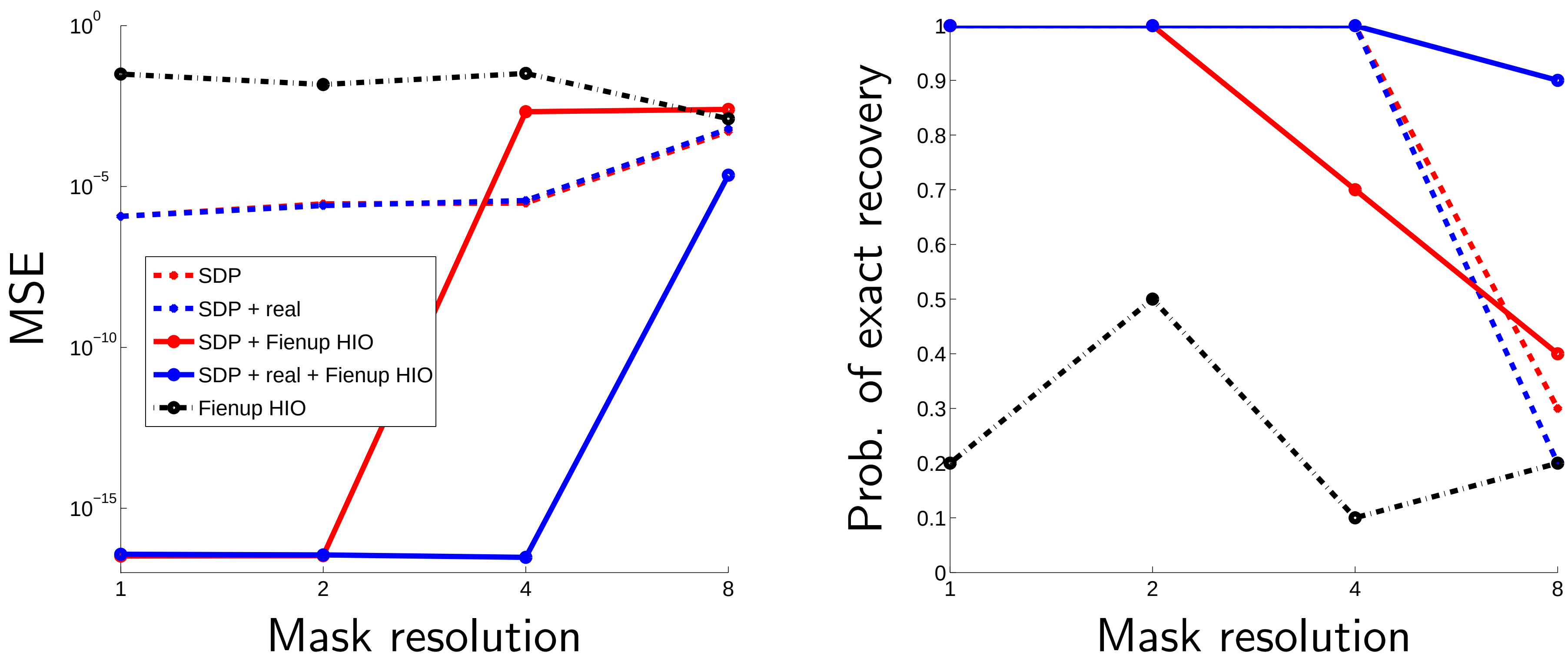
16x16 caffeine image. Mask resolution (1x1 to 8x8 pixels).



Left: MSE vs. mask resolution. **(2x oversampling, no noise, 3 masks).**
Right: Probability of recovering molecular density ($MSE < 10^{-4}$) vs. number of random masks.

Numerical Experiments: comparing algorithms

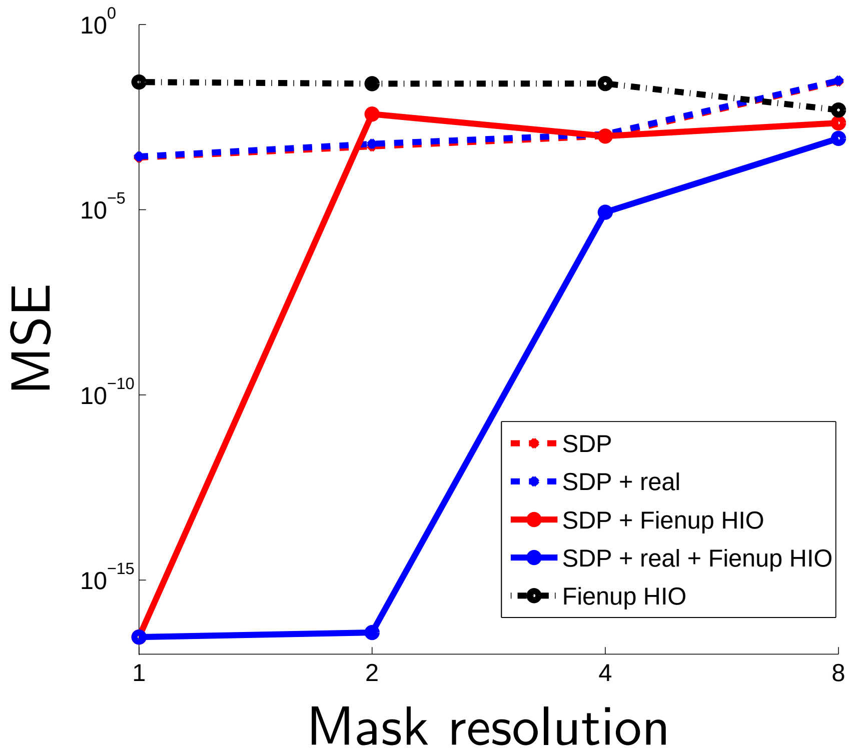
16x16 cocaine image. **Mask resolution (1x1 to 8x8 pixels).**

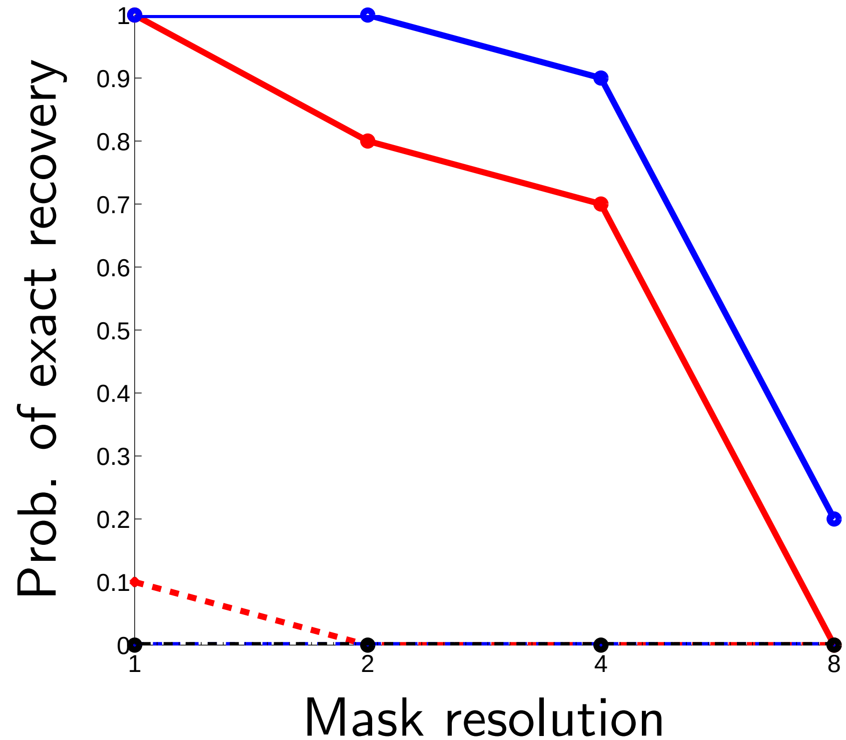


Left: MSE vs. mask resolution. **(2x oversampling, no noise, 3 masks).**
Right: Probability of recovering molecular density ($MSE < 10^{-4}$) vs. number of random masks.

Numerical Experiments: comparing algorithms

16x16 cocaine image. Mask resolution (1x1 to 8x8 pixels).

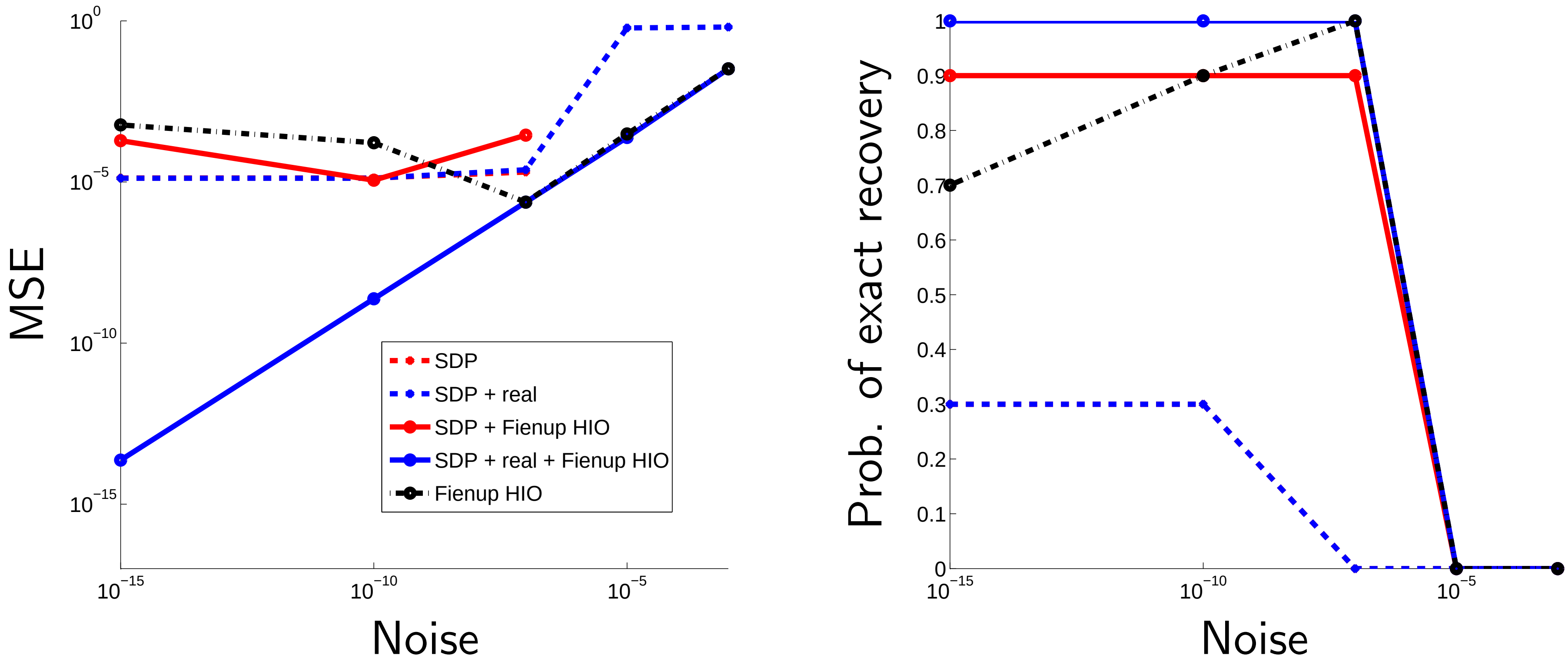




Left: MSE vs. mask resolution. **(2x oversampling, no noise, 2 masks).**
Right: Probability of recovering molecular density ($MSE < 10^{-4}$) vs. number of random masks.

Numerical Experiments: comparing algorithms

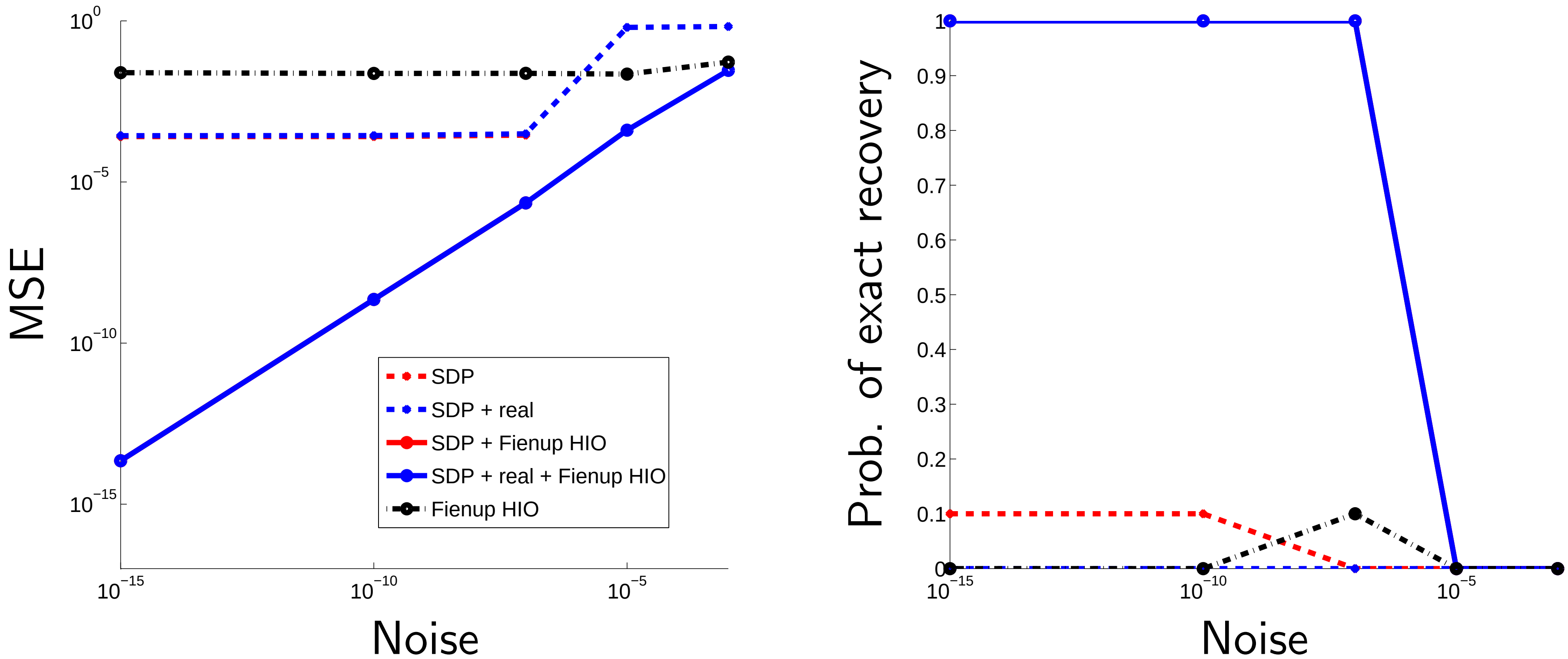
16x16 caffeine image. Noise.



Left: MSE vs. noise level **(2x oversampling, 2 masks).**
Right: Probability of recovering molecular density ($MSE < 10^{-4}$) vs. number of random masks.

Numerical Experiments: comparing algorithms

16x16 cocaine image. Noise.



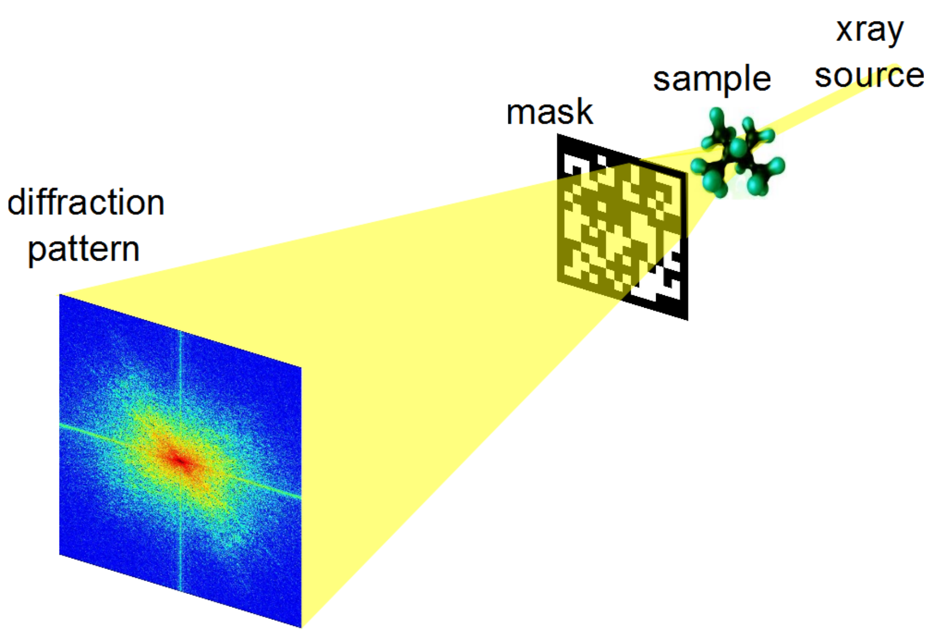
Left: MSE vs. noise level **(2x oversampling, 2 masks).**
Right: Probability of recovering molecular density ($MSE < 10^{-4}$) vs. number of random masks.

Outline

- Introduction
- Algorithms
- Exploiting structure
- Numerical results
- **Experimental setup?**

Observations A: implementation

Construct observations A from **multiple** coded diffraction patterns



- Split the beam?
- Mask before/after the sample?

Conclusion

- Tractable algorithms for phase recovery
- Exact recovery results
- Exploit structure

Open questions. . . .

- Is the SDP relaxation optimal?
- Experimental setup?

References

Heinz H Bauschke, Patrick L Combettes, and D Russell Luke. Phase retrieval, error reduction algorithm, and fienup variants: a view from convex optimization. *JOSA A*, 19(7):1334–1345, 2002.

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*, 66(8):1241–1274, 2013a.

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*, 2011.

Emmanuel J Candes, Xiaodong Li, and Mahdi Soltanolkotabi. Phase retrieval from coded diffraction patterns. *preprint*, 2013b.

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

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.

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.

D Russell Luke. Relaxed averaged alternating reflections for diffraction imaging. *Inverse Problems*, 21(1):37, 2005.

J Miao, D Sayre, and HN Chapman. Phase retrieval from the magnitude of the fourier transforms of nonperiodic objects. *JOSA A*, 15(6): 1662–1669, 1998.

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.

Y. Nesterov and A. Nemirovskii. *Interior-point polynomial algorithms in convex programming*. Society for Industrial and Applied Mathematics, Philadelphia, 1994.

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. Singer. Angular synchronization by eigenvectors and semidefinite programming. *Applied and computational harmonic analysis*, 30(1): 20–36, 2011.

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