# Nonconvex landscapes for phase retrieval

Andrew D. McRae

EPFL
Lausanne, Switzerland

ICCOPT
July 23, 2025

# The phase retrieval problem

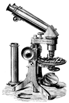
Generalized linear model: for unknown $x_* \in \mathbf{C}^d$, suppose we observe

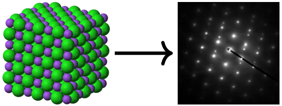
$$y_i \approx |\langle a_i, x_* \rangle|^2, \quad i = 1, \ldots, n,$$

where $a_1, \ldots, a_n \in \mathbf{C}^d$ are known measurement vectors.

**Recovery problem:** estimate $x_*$

## Motivation: optical imaging





- Electromagnetic field (complex amplitude) is often linear...
- However, measured light **intensity** is the (squared) magnitude

# Least-squares estimation

We observe

$$y_i \approx |\langle a_i, x_* \rangle|^2, \qquad a_1, \ldots, a_n \in \mathbf{C}^n \text{ known}, \qquad x_* \in \mathbf{C}^n \text{ unknown}$$

How do we efficiently **compute** as estimate of $x_*$?

- (∃ vast literature)

Least-squares estimator of $x_*$:

$$\min_{x \in \mathbf{C}^d} \sum_{i=1}^{n} (y_i - |\langle a_i, x \rangle|^2)^2$$

**Nonconvex:** could have bad local minima

- How can we overcome this?

# Low-rank matrix sensing approach

We observe

$$y_i \approx |\langle a_i, x_* \rangle|^2 \underbrace{= \langle a_i a_i^*, x_* x_*^* \rangle}_{\text{linear in } x_* x_*^*}, \qquad (\text{“lifting” trick})$$

We can then use the techniques of (linear) **low-rank matrix sensing**

- $x_* x_*^*$ is a rank-1 positive semidefinite matrix

“Lifted” matrix estimator ($A_i = a_i a_i^*$):

$$\min_{Z \succeq 0} \sum_{i=1}^{n} (y_i - \langle A_i, Z \rangle)^2 \text{ s.t. } \operatorname{rank}(Z) = 1$$

One approach: drop rank constraint to get **convex** semidefinite program (“PhaseLift”)

- This is computationally expensive ($\approx d^2$ variables)
- Can we use the **nonconvex** problem directly?

We are interested in the **landscape**: when is *any* local optimum a good solution?

# Challenge 1: no restricted isometry property

Ignoring noise and using Burer-Monteiro, we have ($A_i = a_i a_i^*, Z_* = x_* x_*^*$)

$$\min_{x \in \mathbf{C}^n} \sum_{i=1}^{n} \langle A_i, xx^* - Z_* \rangle^2$$

- Landscape of such problems well-studied...
- Most theory assumes **restricted isometry property**:

$$(1-\delta)\|H\|_F^2 \le \frac{1}{n}\sum_{i=1}^{n}\langle A_i, H\rangle^2 \not\le (1+\delta)\|H\|_F^2 \quad \text{for low-rank } H.$$

Upper restricted isometry **fails** for phase retrieval

- More specialized analysis needed

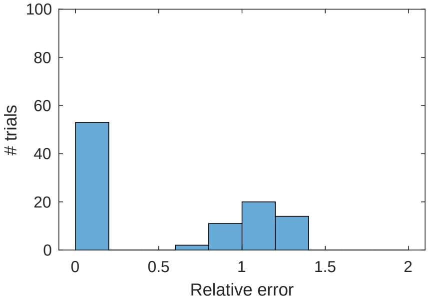
# Challenge 2: existing phase retrieval results make strong assumptions

Despite lack of RIP, $\exists$ theoretical results for phase retrieval

- For example, Sun et al. (2018), Cai et al. (2023)

**Limitations**

- Assume Gaussian measurements
- Require $n \gtrsim d \log d$ measurements **(statistically suboptimal)**
- For "harder" problem instances, nonconvex landscape is **not benign** in general!



# Relaxation

To try to improve the landscape, we relax the **rank constraint**

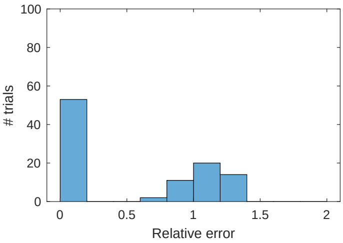
$$\min_{Z \succeq 0} \sum_{i=1}^{n} (y_i - \langle A_i, Z \rangle)^2 \text{ s.t. } \operatorname{rank}(Z) \leq r \iff \min_{X \in \mathbf{C}^{d \times r}} \sum_{i=1}^{n} (y_i - \langle A_i, XX^* \rangle)^2$$

- Motivated by work in matrix sensing and synchronization (Ling, 2023; Zhang, 2024)
- $r = n \longleftrightarrow$ SDP

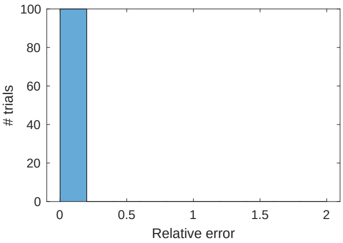
**Theoretically**, not obvious this helps!

- In matrix sensing, sometimes "overparametrization" can introduce spurious local optima!
- How do we ensure the relaxation is tight?

**Empirically**, seems promising:




$r = 1$




$r = 2$

## (Some) theoretical results

Relaxed nonconvex estimator ($y_i \approx \langle A_i, x_* x_*^* \rangle$, $A_i = a_i a_i^*$):

$$\min_{X \in \mathbf{C}^{d \times r}} \sum_{i=1}^{n} (y_i - \langle A_i, XX^* \rangle)^2 \tag{BM-$r$}$$

### Theorem (Representative)

*If $a_1, \dots, a_n$ are sub-Gaussian random vectors satisfying the assumptions of Krahmer and Stöger (2020), as long as*

$$n \gtrsim d \qquad \text{and} \qquad r \gtrsim \log d,$$

*every second-order critical point of* (BM-$r$) *satisfies (if there is no measurement error)* $XX^* = x_* x_*^*$.

Comments:

- In some cases, first statistically optimal result without SDP
- Requires significantly different analysis than those assuming RIP
- Can be generalized to other PSD measurement and ground truth matrices
- The particular PSD structure avoids possible dangers of overparametrization
- Deterministic result looks suspiciously like a **condition number**

# What's next?

Forthcoming

- **Different loss functions**
- Nonparametric/infinite-dimensional results

Future work

- Theory for more realistic (e.g., optical) measurements
- Additional structure (e.g., **sparsity**)

# Improvements with modified loss

Quartic "intensity" estimator:

$$\min_{X \in \mathbf{C}^{d \times r}} \sum_{i=1}^{n} (y_i - \langle A_i, XX^* \rangle)^2$$

- **Pros:** Smooth, fits into matrix sensing framework nicely
- **Con:** Landscape guarantees (that are statistically optimal) require $r \gtrsim \log d$

"Amplitude" estimator:

$$\min_{X \in \mathbf{C}^{d \times r}} \sum_{i=1}^{n} (y_i^{\mathbf{1/2}} - \langle A_i, XX^* \rangle^{\mathbf{1/2}})^2$$

- New result: good landscape with only $r = O(1)$
- Similar results for
  - Poisson MLE loss
  - Nonconvex PhaseCut (phase retrieval via synchronization; Waldspurger et al., 2015)
- Preprint coming "soon"

# Open problem—nonconvex estimator with sparsity

Old paper: Andrew D. McRae, Justin Romberg, and Mark A. Davenport (2023). "Optimal convex lifted sparse phase retrieval and PCA with an atomic matrix norm regularizer". In: *IEEE Trans. Inf. Theory* 69.3, pp. 1866–1882

- Promising empirical results with estimator of the form

$$\min_{X \in \mathbf{C}^{d \times r}} \sum_{i=1}^{n} (y_i - \langle A_i, XX^* \rangle)^2 + \theta(X) \quad \longleftarrow \quad \text{penalty based on } \ell_1 \text{ norm}$$

- Difficulty: every version of this I can think of with an $\ell_1$ norm has spurious local optima due to nonsmoothness
- **Questions:**
  - Why does it work so well empirically?
  - Is there a formulation more amenable to theory?

Preprint (quartic loss): Andrew D. McRae (2025). "Phase retrieval and matrix sensing via benign and overparametrized nonconvex optimization". In: arXiv: 2505.02636 [math.OC]

# Thanks!

# References I


- Cai, Jian-Feng, Meng Huang, Dong Li, and Yang Wang (2023). "Nearly optimal bounds for the global geometric landscape of phase retrieval". In: *Inverse Probl.* 39.7.
- Krahmer, Felix and Dominik Stöger (2020). "Complex Phase Retrieval from Subgaussian Measurements". In: *J. Fourier Anal. Appl.* 26.89.
- Ling, Shuyang (2023). "Solving Orthogonal Group Synchronization via Convex and Low-Rank Optimization: Tightness and Landscape Analysis". In: *Math. Program.* 200, pp. 589–628.
- McRae, Andrew D. (2025). "Phase retrieval and matrix sensing via benign and overparametrized nonconvex optimization". In: arXiv: 2505.02636 [math.OC].
- McRae, Andrew D., Justin Romberg, and Mark A. Davenport (2023). "Optimal convex lifted sparse phase retrieval and PCA with an atomic matrix norm regularizer". In: *IEEE Trans. Inf. Theory* 69.3, pp. 1866–1882.
- Sun, Ju, Qing Qu, and John Wright (2018). "A Geometric Analysis of Phase Retrieval". In: *Found. Comput. Math.* 18.5, pp. 1131–1198.
- Waldspurger, Irène, Alexandre d'Aspremont, and Stéphane Mallat (2015). "Phase recovery, MaxCut and complex semidefinite programming". In: *Math. Program.* 149, pp. 47–81.

# References II

Zhang, Richard Y. (2024). "Improved global guarantees for the nonconvex Burer-Monteiro factorization via rank overparameterization". In: *Math. Program.*