Nonconvex optimization landscapes in phase retrieval

Andrew D. McRae

Institute of Mathematics, EPFL

Statistical problem

We want to recover a low-rank positive semidefinite (PSD) $d \times d$ matrix $Z_* \succeq 0$ from measurements

$$y_i \approx \langle A_i, Z_* \rangle, i = 1, \dots, n,$$

where $A_1, \dots, A_n \succeq 0$ are $d \times d$ PSD sensing matrices.
Canonical example: **phase retrieval**, where

$$Z_* = x_* x_*^*, \qquad \text{and} \qquad A_i = a_i a_i^*$$

for vectors $x_* \in \mathbf{F}^d$ ($\mathbf{F}$ could be $\mathbf{R}$ or $\mathbf{C}$) and $a_1, \dots, a_n \in \mathbf{C}^d$ (in general).

Nonconvex least-squares estimator

We estimate Z_* by the Burer-Monteiro factored least-squared problem

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

where p is a search rank hyperparameter.
For phase retrieval with the obvious choice of $p = 1$, this becomes the classic quartic problem

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

These objective functions are smooth (quartic polynomials) but nonconvex, hence they potentially could have spurious local optima.
We want to understand the **nonconvex landscape** of (BM-LS). Can we show that all local minima are global or are at least good statistical estimators?

Prior work and its limitations

There is much existing work on landscapes of problems like (BM-LS) in low-rank matrix sensing, but it generally assumes a *restricted isometry property* (RIP): this requires that, for all low-rank Hermitian S,

$$\alpha \|S\|_F^2 \le \frac{1}{n} \sum_{i=1}^n \langle A_i, S \rangle^2 \le \beta \|S\|_F^2 \tag{RIP}$$

for $\alpha, \beta > 0$ with sufficiently small ratio β / α.
For phase retrieval, **RIP fails**: in general (e.g., with Gaussian measurements), we have

$$\frac{\beta}{\alpha} \gtrsim \frac{d^2}{n},$$

which is far too large without unreasonably large sample size n. Thus another approach is needed.
By more problem-specific methods, the phase retrieval problem (PR-LS) has been studied and shown to have a benign landscape when the a_i's are Gaussian and $n \gtrsim d \log d$ [Cai+23]. However, the **sample complexity** requirement is suboptimal, and the analysis depends on the fact that the measurements are Gaussian.

New analysis and overparametrization

- We develop a **novel landscape analysis** that is inspired by but greatly extends existing techniques for phase retrieval.
- As in other recent works on low-rank matrix sensing and synchronization, we benefit from **overparametrization**: setting search rank $p > \operatorname{rank}(Z_*)$.
- Our analysis exploits the **positive semidefinite structure** (that both Z_* and the measurements A_i are PSD).

Deterministic landscape result (rank-1, noiseless)

***Theorem* 1.** *Suppose $Z_* = x_* x_*^*$ has rank 1, and suppose we observe*

$$y_i = \langle A_i, Z_* \rangle \quad \text{for} \quad A_1, \dots, A_n \succeq 0.$$

Furthermore, suppose the measurements $\{A_i\}$ satisfy, for some $\alpha, L > 0$,

$$\frac{1}{n} \sum_{i=1}^n \langle A_i, XX^* - Z_* \rangle^2 \ge \alpha \|XX^* - Z_*\|_F^2 \quad \forall X \in \mathbf{F}^{d \times p}, \text{ and}$$

$$\left\| \frac{1}{n} \sum_{i=1}^n \langle A_i, Z_* \rangle A_i \right\|_{\mathrm{op}} \le L \|x_*\|^2.$$

Finally, suppose the rank parameter p satisfies

$$p + 2 > \frac{2L}{\alpha}.$$

Then every second-order critical point X of (BM-LS) *satisfies $XX^* = Z_*$.*

- This can be extended to Z_* of larger rank (see Theorem 3).
- Lower isometry requirement is similar to RIP; upper isometry only required on a restricted space.
- A refined version recovers the result of [Cai+23] for (PR-LS) with Gaussian measurements (see preprint).
- A larger search rank p allows looser assumptions on $\{A_i\}$. Or does it?

Challenge: universal lower isometry bound

To get the full benefits of overparametrization with Theorem 1, we want a lower isometry bound independent of the search rank p:

$$\frac{1}{n} \sum_{i=1}^n \langle A_i, Z - Z_* \rangle^2 \ge \alpha \|XX^* - Z_*\|_F^2 \quad \forall Z \succeq 0.$$

For general low-rank matrix sensing, this is impossible unless $n \gtrsim d^2$. However, the **PSD structure** lets us do more.
Intuitively, we can build on the observation in the PhaseLift literature (e.g., [CL13]) that, in many cases,

$$\{Z \succeq 0 : \langle A_i, Z \rangle = \langle A_i, Z_* \rangle, i = 1, \dots, n\} = \{Z_*\}.$$

Some applications

To apply Theorem 1 with a universal (p-independent) lower isometry bound, we build on two existing methods in the phase retrieval literature.
For phase retrieval, we can use the framework of Krahmer et al. for general sub-Gaussian measurements:

***Theorem* 2.** *With $Z_* = x_* x_*^*$, under the assumptions of [KS20] on x_* and sub-Gaussian measurements, if $n \gtrsim d$, with high probability, for search rank*

$$p \gtrsim 1 + \frac{d \log n}{n},$$

every second-order critical point X of (BM-LS) *satisfies $XX^* = Z_* = x_* x_*^*$.*

As another method, we adopt and extend the dual certificate approach of (among others) [CL13]. The following result is one application:

***Theorem* 3** (General rank, Gaussian measurements). *If $Z_* \succeq 0$ has rank r, and $A_i = a_i a_i^*$ for i.i.d. Gaussian $a_1, \dots, a_n$, then, if $n \gtrsim rd$, with high probability, for*

$$p \gtrsim \left(1 + \frac{d \log n}{n}\right) \frac{\operatorname{tr} Z_*}{\sigma_r(Z_*)},$$

every second-order critical point X of (BM-LS) *satisfies $XX^* = Z_*$.*

- This can be extended to handle noise (see preprint). The resulting error bound is nearly optimal even without explicit regularization.
- We hope to apply next the dual certificate approach to more realistic models like masked Fourier measurements.

The tasks of Theorems 2 and 3 had previously only been solved with semidefinite programming (PhaseLift) or other more elaborate schemes.

References

[Cai+23] J.-F. Cai et al. "Nearly optimal bounds for the global geometric landscape of phase retrieval". In: *Inverse Probl.* (2023).

[CL13] E. J. Candès and X. Li. "Solving Quadratic Equations via PhaseLift When There Are About as Many Equations as Unknowns". In: *Found. Comput. Math.* (2013).

[KS20] F. Krahmer and D. Stöger. "Complex Phase Retrieval from Sub-gaussian Measurements". In: *J. Fourier Anal. Appl.* (2020).

Preprint

https://go.epfl.ch/prscapes