Smoothed analysis of the low-rank approach for smooth semidefinite programs

Thomas Pumir*, **Samy Jelassi*** and **Nicolas Boumal†**

* ORFE Department, equal contribution

† Mathematics Department

Setting

Burer-Monteiro factorization

- We consider SDPs of the form:

$$\min_{X \in \mathbb{S}^{n\times n}} \langle C, X\rangle \text{ s.t. } \mathcal{A}(X) = b,\ X \succeq 0. \qquad \text{(SDP)}$$

- Solvable in poly time but enforcing PSD constraint can be expensive.
- Burer-Monteiro factorization: Set $X = YY^*$ and (SDP) becomes

$$\min_{Y \in \mathbb{K}^{n\times k}} \langle C, YY^*\rangle \text{ s.t. } \mathcal{A}(YY^*) = b. \qquad \text{(P)}$$

Advantages:

- PSD constraint naturally enforced.
- Moreover, if SDP is compact, it always has a solution of rank r with $\dim \mathbb{S}^{r\times r} \leq m$ (# constraints): can reduce dimension to $k \sim \sqrt{m}$.

Issues:

- The problem becomes non-convex.
- Besides, algorithms can only guarantee *approximate* second-order optimality conditions in a finite number of iterations: they return ASOSPs.

Smoothness assumption

- Search space of (P): $\mathcal{M}_k = \{Y \in \mathbb{K}^{n\times k} : \mathcal{A}(YY^*) = b\}$.
- The set $\mathcal{M}_n$ is a smooth manifold (implies $\mathcal{M}_k$ is smooth for $k \leq n$).

Overview

Main Results

Main Result: With high probability, *approximate* second-order stationary points (ASOSPs) for a randomly perturbed objective function are *approximate* global optima.

Approach: Smoothed analysis.

Motivating applications: Phase retrieval, angular synchronization, max-cut, synchronization of rotations (SLAM, 3D registration), trust-region subproblem

Related work

- Burer and Monteiro [4] showed that if Y is a *rank-deficient* local optimum, then $X = YY^*$ is a global optimum.
- Under the same setting as us, Boumal et al. [3] showed that for *almost all* cost matrices, all second-order stationary points (SOSPs) are optimal.
- For a broader class of SDPs, *without* enforcing exact constraint satisfaction, Bhojanapalli et al. [1] showed that, under random perturbations, ASOSPs are approximately optimal.

Benign non-convexity in Burer-Monteiro factorization

Assumptions

- The search space of (SDP) is compact.
- The search space of (P) is a manifold.

Main theorem

- If we randomly perturb the cost matrix C, with $k = \tilde{\Omega}(\sqrt{m})$
- Then, w.h.p. on the perturbation, if $Y \in \mathbb{K}^{n\times k}$ is an ASOSP for (P),
- $X = YY^*$ is an approximate global optimum.

Proof sketch

Probabilistic argument

Perturbing the cost matrix in (P) with a Gaussian Wigner matrix $\Rightarrow$ w.h.p., any approximate first-order stationary point Y of the perturbed (P) is almost column-rank deficient.

Deterministic argument

If Y ASOSP for (P) and almost column-rank deficient, then $X = YY^*$ is an approximate global optimum for (SDP).

Example: Phase retrieval

Problem setting

Goal: Retrieve a signal $z \in \mathbb{C}^n$ from $b = |Az| \in \mathbb{R}^m_+$.

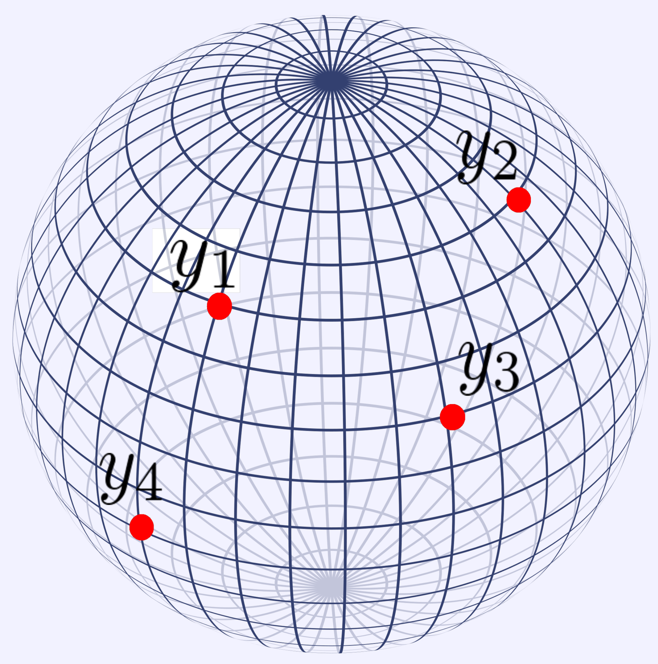
$$\min_{u \in \mathbb{C}^m} u^* C u \text{ s.t. } |u_i| = 1, \qquad \text{(PR)}$$

with $C = \mathrm{diag}(b)(I - AA^\dagger)\mathrm{diag}(b)$.

Dropping the rank constraint, relaxes the problem to the following SDP:

$$\min_{X \in \mathbb{H}^{m\times m}} \langle C, X\rangle \text{ s.t. } \mathrm{diag}(X) = 1, \quad X \succeq 0. \qquad \text{(PhaseCut)}$$

Burer-Monteiro factorization



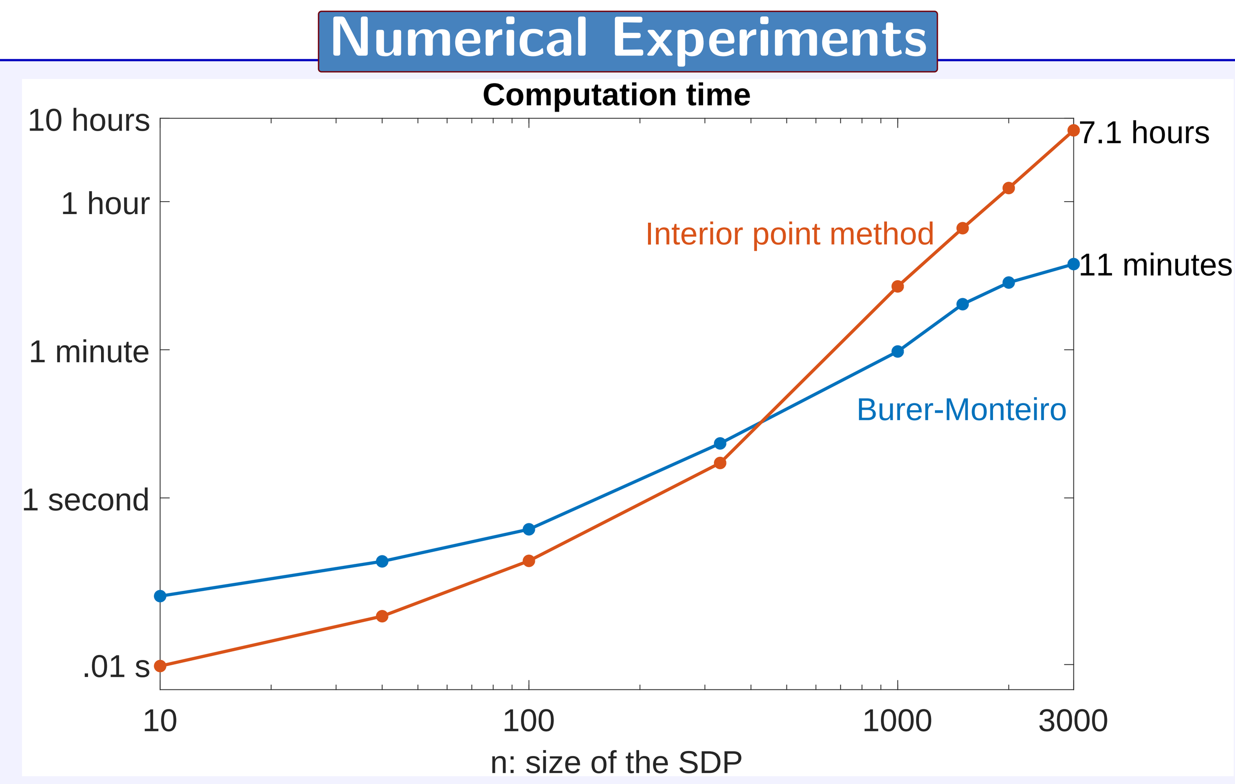
Each y_i is a point on the unit sphere in $\mathbb{C}^k$.

$$\min_{Y \in \mathbb{C}^{m\times k}} \langle CY, Y\rangle \text{ s.t. } y_i^* y_i = 1,\ \forall i,$$

with $Y = [y_1^*, \ldots, y_m^*]$ and $y_i \in \mathbb{C}^k$.

- The main theorem applies
- Related work: Mei et al. [5].
  - Holds for ASOSPs without perturbation.
  - More general result since it holds for any k.
  - But when k $= \tilde{\Omega}(\sqrt{m})$, does not capture optimality as we do here.

Numerical Experiments



Computation time of a dedicated interior-point method (IPM) and of the Burer-Monteiro approach (BM) using Manopt [2] to solve (PhaseCut). As the number of measurements increases, BM outperforms IPM.

References

[1] S. Bhojanapalli, N. Boumal, P. Jain, and P. Netrapalli. Smoothed analysis for low-rank solutions to semidefinite programs in quadratic penalty form. *arXiv preprint arXiv:1803.00186*, 2018.

[2] N. Boumal, B. Mishra, P.-A. Absil, and R. Sepulchre. Manopt, a Matlab toolbox for optimization on manifolds. *Journal of Machine Learning Research*, 15:1455–1459, 2014. URL `http://jmlr.org/papers/v15/boumal14a.html`.

[3] N. Boumal, V. Voroninski, and A. Bandeira. The non-convex Burer-Monteiro approach works on smooth semidefinite programs. In *Advances in Neural Information Processing Systems*, pages 2757–2765, 2016.

[4] S. Burer and R. D. Monteiro. A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization. *Mathematical Programming*, 95(2):329–357, 2003.

[5] S. Mei, T. Misiakiewicz, A. Montanari, and R. I. Oliveira. Solving sdps for synchronization and maxcut problems via the grothendieck inequality. In *Proceedings of the 30th Conference on Learning Theory, COLT 2017, Amsterdam, The Netherlands, 7-10 July 2017*, pages 1476–1515, 2017.