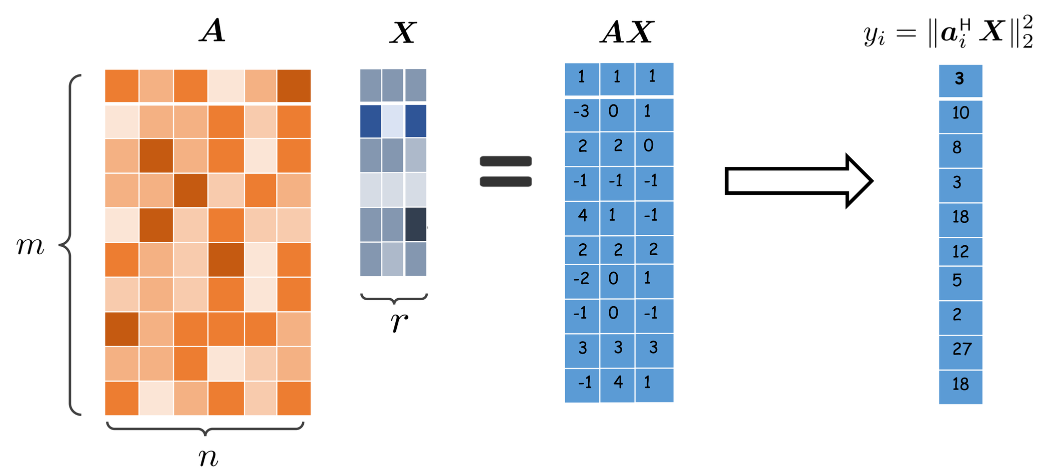
Solving Quadratic Equations via Amplitude-based Nonconvex Optimization

Vincent Monardo, Yuanxin Li and Yuejie Chi
Department of Electrical and Computer Engineering, Carnegie Mellon University
Email: {vmonardo,yuanxinl,yuejiec}@andrew.cmu.edu

Carnegie Mellon University

Quadratic Inverse Problem Setup

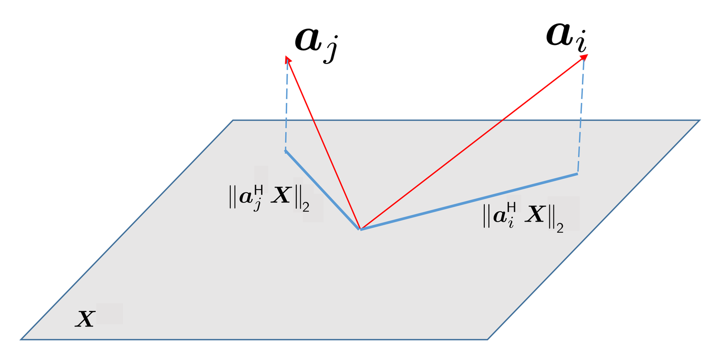


- **Goal:** recover $\boldsymbol{X} \in \mathbb{C}^{n\times r}$ from m quadratic measurements,

$$y_i = \|\boldsymbol{a}_i^H \boldsymbol{X}\|_2^2, \quad i = 1, ..., m.$$

- This problem has many applications in coherence retrieval in optical imaging, covariance sketching of high-dimensional streaming data for the general rank-r case, and phase retrieval for the rank-1 case.

Geometric Interpretation



- We lose a generalized notion of "phase" in $\mathbb{C}^r$: $\boldsymbol{a}_i^H \boldsymbol{X}/\|\boldsymbol{a}_i^H \boldsymbol{X}\|_2$

Amplitude Based Optimization

- **Typical approach**: minimizing an intensity-based loss function with measurements $y_i = \|\boldsymbol{a}_i^H \boldsymbol{X}\|_2^2, \quad i = 1, ..., m,$

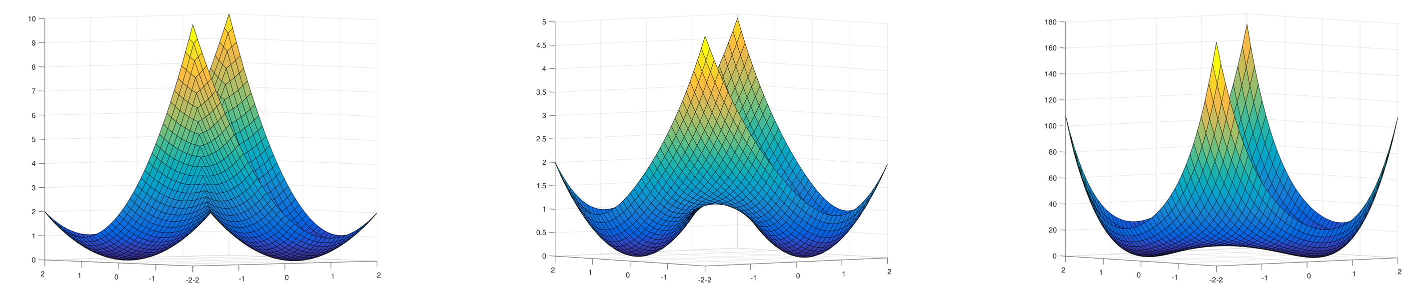
$$\ell_{in}(\boldsymbol{U}) = \frac{1}{m}\sum_{i=1}^{m}\left(y_i - \|\boldsymbol{a}_i^H \boldsymbol{U}\|_2^2\right)^2.$$

  - Provable statistical and computational guarantees, e.g. Li et. al., 2018.
- **Proposed approach**: minimizing an amplitude-based loss function with measurements $z_i = \sqrt{y_i}, \quad i = 1, ..., m,$

$$\ell(\boldsymbol{U}) = \frac{1}{m}\sum_{i=1}^{m}\left(z_i - \|\boldsymbol{a}_i^H \boldsymbol{U}\|_2\right)^2.$$

  - Only involves the second power of variable, *reducing complexity*!
  - Proposed for phase retrieval in Gerchberg-Saxton 1972, Zhang et.al. 2017, Wang et. al. 2017.

Why Amplitude-based Loss?



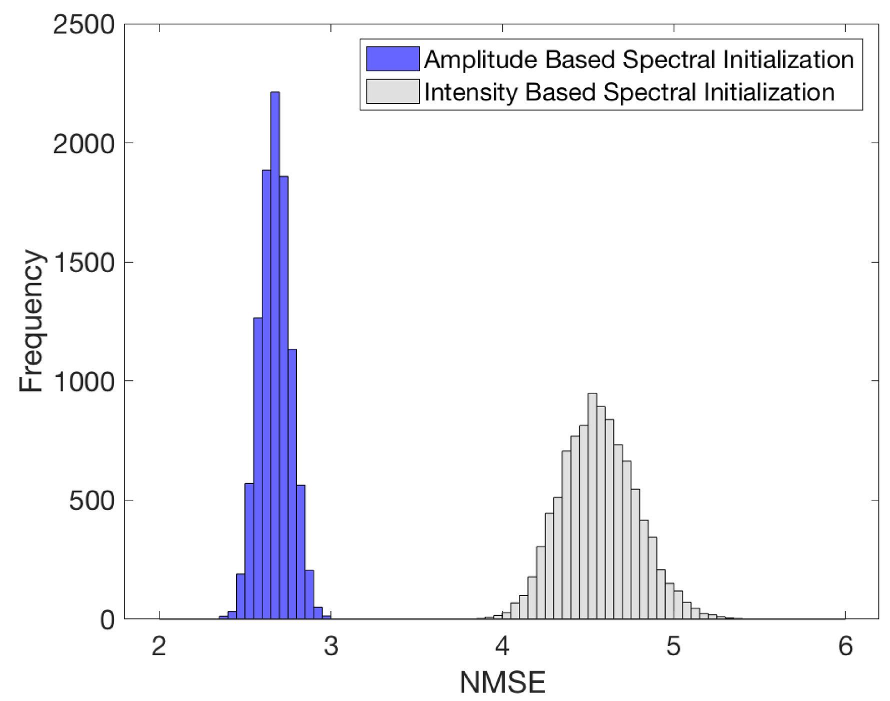
(a) Quadratic Surface about Minima (b) Amplitude-Based Loss Surface (c) Intensity-Based Loss Surface

- The curvature of the amplitude based loss surface is similar to the quadratic function around the global optima, see Zhang et. al., 2017.

Algorithmic Approaches

- **Step 1: Spectral Initialization**

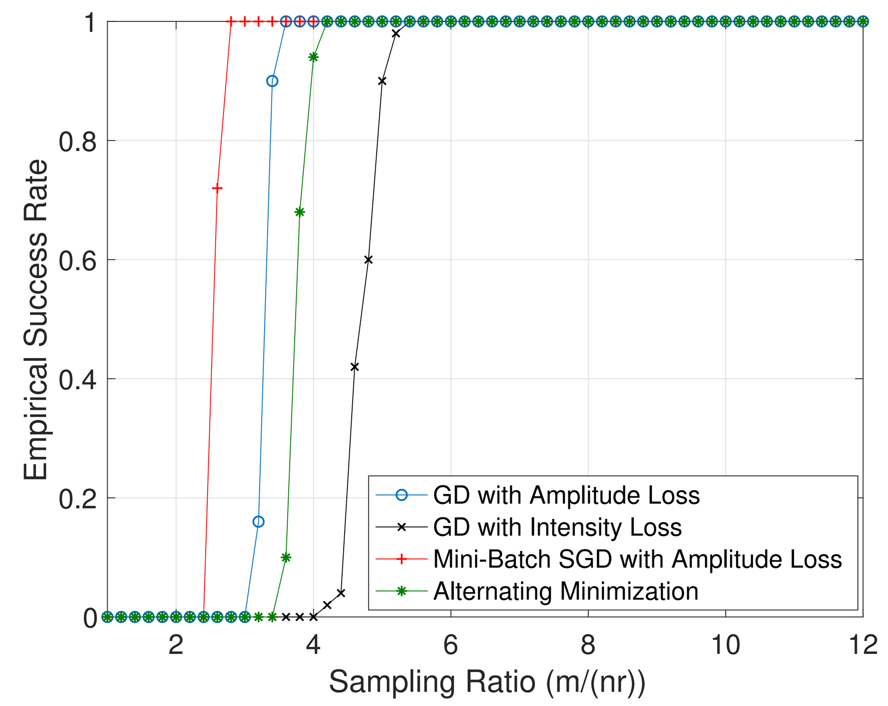
1. With $\{z_i\}_{i=1}^m$, and $\{\boldsymbol{a}_i\}_{i=1}^m$, create $\boldsymbol{D} = \frac{1}{2m}\sum_{i=1}^m z_i \boldsymbol{a}_i \boldsymbol{a}_i^H$.
2. Obtain the r normalized eigenvectors $\boldsymbol{Z}_0 \in \mathbb{C}^{n\times r}$ corresponding to the r largest eigenvalues of $\boldsymbol{D}$.
3. Obtain the diagonal matrix $\boldsymbol{\Lambda}_0 \in \mathbb{C}^{r\times r}$, with entries on the diagonal given by $[\boldsymbol{\Lambda}_0]_i = \lambda_i(\boldsymbol{D}) - \lambda,\ i = 1, ..., r$ where $\lambda = \frac{1}{m}\sum_{i=1}^m z_i$ and $\lambda_i(\boldsymbol{D})$ is the i^{th} largest eigenvalue of $\boldsymbol{D}$.
4. $\boldsymbol{U}_0 = \boldsymbol{Z}_0 \boldsymbol{\Lambda}_0^{1/2}$.



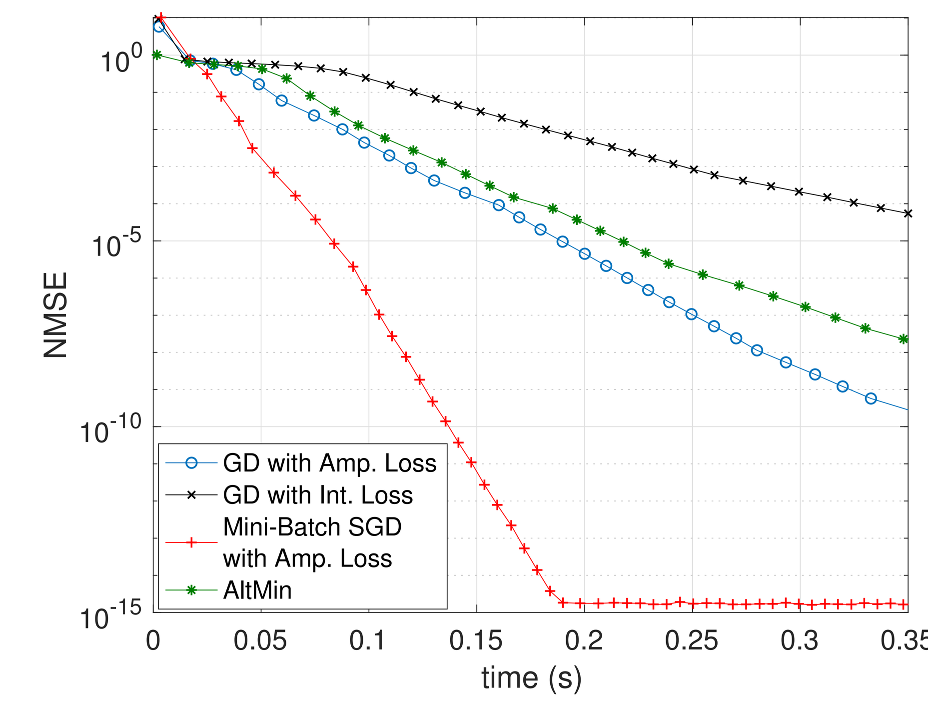
- **Step 2: Iterative Refinement:**
  - Gradient Descent (GD):
  $$\boldsymbol{U}_{k+1} = \boldsymbol{U}_k - \frac{2\mu_k}{m}\boldsymbol{A}^H \operatorname{diag}\left(\left[\frac{\|\boldsymbol{a}_1^H \boldsymbol{U}_k\|_2 - z_1}{\|\boldsymbol{a}_1^H \boldsymbol{U}_k\|_2}, ..., \frac{\|\boldsymbol{a}_m^H \boldsymbol{U}_k\|_2 - z_m}{\|\boldsymbol{a}_m^H \boldsymbol{U}_k\|_2}\right]\right)\boldsymbol{A}\boldsymbol{U}_k.$$
  - Mini-batch Stochastic Gradient Descent (SGD):
  $$\boldsymbol{U}_{k+1} = \boldsymbol{U}_k - \frac{2\mu_k}{m}\boldsymbol{A}_{\Gamma_k}^H \operatorname{diag}\left(\left\{\frac{\|\boldsymbol{a}_i^H \boldsymbol{U}_k\|_2 - z_i}{\|\boldsymbol{a}_i^H \boldsymbol{U}_k\|_2}\,\middle|\, i \in \Gamma_k\right\}\right)\boldsymbol{A}_{\Gamma_k}\boldsymbol{U}_k.$$
  - Alternating Minimization (AltMin):
  $$\boldsymbol{U}_{k+1} = \boldsymbol{A}^{\dagger}\text{diag}\left(\left[\frac{z_1}{\|\boldsymbol{a}_1^H \boldsymbol{U}_k\|_2}, ..., \frac{z_m}{\|\boldsymbol{a}_m^H \boldsymbol{U}_k\|_2}\right]\right)\boldsymbol{A}\boldsymbol{U}_k.$$

Numerical Experiments

Statistical Performance ($n = 50$ and $r = 4$)



Computational Performance ($n = 50$, $r = 4$ and $m = 800$)



References

[1] Y. Li, C. Ma, Y. Chen, and Y. Chi. Nonconvex matrix factorization from rank-one measurements. 2018.

[2] G. Wang, G. B. Giannakis, and Y. C. Eldar. Solving systems of random quadratic equations via truncated amplitude flow. 2018.

[3] H. Zhang, Y. Zhou, Y. Liang, and Y. Chi. A nonconvex approach for phase retrieval: Reshaped Wirtinger flow and incremental algorithms. 2017.

Acknowledgements

This work is supported in part by ONR under the grant N00014-18-1-2142, by ARO under the grant W911NF-18-1-0303, and by NSF under grants CCF-1806154 and ECCS-1818571.