Exact Minimax Estimation for Phase Synchronization

Chao Gao
University of Chicago

Oct 2021


Anderson Zhang

Phase Synchronization

Model $$Y_{jk} = z_j \bar{z}_k + \sigma W_{jk} \in \mathbb{C} \qquad 1 \leq j < k \leq n$$

Phase Synchronization

Model $$Y_{jk} = z_j \bar{z}_k + \sigma W_{jk} \in \mathbb{C} \qquad 1 \leq j < k \leq n$$

Parameter space $$z_j \in \mathbb{C}_1 = \{z \in \mathbb{C} : |z| = 1\}$$

Phase Synchronization

Model $$Y_{jk} = z_j \bar{z}_k + \sigma W_{jk} \in \mathbb{C} \qquad 1 \leq j < k \leq n$$

$$\boxed{\mathrm{Im}(W_{jk}), \mathrm{Re}(W_{jk}) \sim N(0, 1/2)}$$

Parameter space $$z_j \in \mathbb{C}_1 = \{z \in \mathbb{C} : |z| = 1\}$$

Phase Synchronization

Model

$$Y_{jk} = z_j \bar{z}_k + \sigma W_{jk} \in \mathbb{C} \qquad 1 \leq j < k \leq n$$

$$\boxed{\mathrm{Im}(W_{jk}), \mathrm{Re}(W_{jk}) \sim N(0, 1/2)}$$

Parameter space

$$z_j \in \mathbb{C}_1 = \{z \in \mathbb{C} : |z| = 1\}$$

Loss function

$$\ell(\widehat{z}, z) = \min_{a \in \mathbb{C}_1} \frac{1}{n} \sum_{j=1}^{n} |\widehat{z}_j a - z_j|^2$$

Algorithms

Algorithms

MLE

$$\max_{z \in \mathbb{C}_1^n} z^{\mathrm{H}} Y z$$

Algorithms

MLE

$$\max_{z \in \mathbb{C}_1^n} z^{\mathrm{H}} Y z$$

GPM

$$z_j^{(t)} = \frac{\sum_{k \in [n] \setminus \{j\}} Y_{jk} z_k^{(t-1)}}{\left| \sum_{k \in [n] \setminus \{j\}} Y_{jk} z_k^{(t-1)} \right|}$$

Algorithms

MLE

$$\max_{z \in \mathbb{C}_1^n} z^{\mathrm{H}} Y z$$

GPM

$$z_j^{(t)} = \frac{\sum_{k \in [n] \setminus \{j\}} Y_{jk} z_k^{(t-1)}}{\left| \sum_{k \in [n] \setminus \{j\}} Y_{jk} z_k^{(t-1)} \right|}$$

SDP

Algorithms

MLE

$$\max_{z \in \mathbb{C}_1^n} z^{\mathrm{H}} Y z = \max_{z \in \mathbb{C}_1^n} \mathsf{Tr}(Y z z^{\mathrm{H}})$$

GPM

$$z_j^{(t)} = \frac{\sum_{k \in [n] \setminus \{j\}} Y_{jk} z_k^{(t-1)}}{\left| \sum_{k \in [n] \setminus \{j\}} Y_{jk} z_k^{(t-1)} \right|}$$

SDP

Algorithms

MLE

$$\max_{z \in \mathbb{C}_1^n} z^{\mathrm{H}} Y z = \max_{z \in \mathbb{C}_1^n} \mathsf{Tr}(Y z z^{\mathrm{H}})$$

GPM

$$z_j^{(t)} = \frac{\sum_{k \in [n] \setminus \{j\}} Y_{jk} z_k^{(t-1)}}{\left| \sum_{k \in [n] \setminus \{j\}} Y_{jk} z_k^{(t-1)} \right|}$$

SDP

$$\max_{Z = Z^{\mathrm{H}} \in \mathbb{R}^{n \times n}} \mathsf{Tr}(YZ) \quad \text{subject to} \;\; \mathrm{diag}(Z) = I_n \text{ and } Z \succeq 0$$

Literature

MLE

GPM

SDP

Literature

MLE [Bandeira, Boumal & Singer 17] $\ell(\widehat{z}_{\mathrm{MLE}}, z^*) \leq C\frac{\sigma^2}{n}$

GPM

SDP

Literature

MLE [Bandeira, Boumal & Singer 17] $\ell(\widehat{z}_{\mathrm{MLE}}, z^*) \leq C\frac{\sigma^2}{n}$

GPM

SDP [Bandeira, Boumal & Singer 17] $\widehat{Z}_{\mathrm{SDP}} = \widehat{z}_{\mathrm{MLE}}\widehat{z}_{\mathrm{MLE}}^{\mathrm{H}} \quad \text{when } \sigma^2 = O(n^{1/2})$

Literature

MLE [Bandeira, Boumal & Singer 17] $\ell(\widehat{z}_{\text{MLE}}, z^*) \leq C\frac{\sigma^2}{n}$

GPM [Boumal 16] $z^{(t)} \to \widehat{z}_{\text{MLE}}$ when $\sigma^2 = O(n^{1/3})$

SDP [Bandeira, Boumal & Singer 17] $\widehat{Z}_{\text{SDP}} = \widehat{z}_{\text{MLE}}\widehat{z}_{\text{MLE}}^{\text{H}}$ when $\sigma^2 = O(n^{1/2})$

Literature

MLE [Bandeira, Boumal & Singer 17] $\ell(\widehat{z}_{\mathrm{MLE}}, z^*) \leq C\frac{\sigma^2}{n}$

GPM

[Boumal 16] $z^{(t)} \to \widehat{z}_{\mathrm{MLE}}$ when $\sigma^2 = O(n^{1/3})$

[Zhong & Boumal 18] $z^{(t)} \to \widehat{z}_{\mathrm{MLE}}$ when $\sigma^2 = O\left(\frac{n}{\log n}\right)$

SDP [Bandeira, Boumal & Singer 17] $\widehat{Z}_{\mathrm{SDP}} = \widehat{z}_{\mathrm{MLE}}\widehat{z}_{\mathrm{MLE}}^{\mathrm{H}}$ when $\sigma^2 = O(n^{1/2})$

Literature

MLE [Bandeira, Boumal & Singer 17] $\ell(\widehat{z}_{\mathrm{MLE}}, z^*) \le C\frac{\sigma^2}{n}$

GPM

[Boumal 16] $z^{(t)} \to \widehat{z}_{\mathrm{MLE}}$ when $\sigma^2 = O(n^{1/3})$

[Zhong & Boumal 18] $z^{(t)} \to \widehat{z}_{\mathrm{MLE}}$ when $\sigma^2 = O\left(\frac{n}{\log n}\right)$

SDP

[Bandeira, Boumal & Singer 17] $\widehat{Z}_{\mathrm{SDP}} = \widehat{z}_{\mathrm{MLE}}\widehat{z}_{\mathrm{MLE}}^{\mathrm{H}}$ when $\sigma^2 = O(n^{1/2})$

[Zhong & Boumal 18] $\widehat{Z}_{\mathrm{SDP}} = \widehat{z}_{\mathrm{MLE}}\widehat{z}_{\mathrm{MLE}}^{\mathrm{H}}$ when $\sigma^2 = O\left(\frac{n}{\log n}\right)$

Main Results

Model $$Y_{jk} = z_j^* \bar{z}_k^* + \sigma W_{jk} \in \mathbb{C} \qquad 1 \leq j < k \leq n$$

Main Results

Model $$Y_{jk} = z_j^* \bar{z}_k^* + \sigma W_{jk} \in \mathbb{C} \qquad 1 \leq j < k \leq n$$

Missing data $$A_{jk} \sim \mathrm{Bernoulli}(p)$$

Main Results

Model

$$Y_{jk} = z_j^* \bar{z}_k^* + \sigma W_{jk} \in \mathbb{C} \qquad 1 \leq j < k \leq n$$

Missing data

$$A_{jk} \sim \text{Bernoulli}(p)$$

Loss function

$$\ell(\widehat{z}, z) = \min_{a \in \mathbb{C}_1} \frac{1}{n} \sum_{j=1}^{n} |\widehat{z}_j a - z_j|^2$$

Main Results

Theorem [G & Zhang]. Assume $\sigma^2 = o(np)$ and $\frac{np}{\log n} \to \infty$. Then,

$$\inf_{\widehat{z} \in \mathbb{C}_1^n} \sup_{z \in \mathbb{C}_1^n} \mathbb{E}_z \ell(\widehat{z}, z) \geq (1 - o(1)) \frac{\sigma^2}{2np}.$$

Main Results

Theorem [G & Zhang]. Assume $\sigma^2 = o(np)$ and $\frac{np}{\log n} \to \infty$. Then,

$$\inf_{\widehat{z} \in \mathbb{C}_1^n} \sup_{z \in \mathbb{C}_1^n} \mathbb{E}_z \ell(\widehat{z}, z) \geq (1 - o(1)) \frac{\sigma^2}{2np}.$$

Moreover, the MLE, GPM initialized by the leading eigenvector of $A \circ Y$, and the leading eigenvector of SDP all achieve

$$\ell(\widehat{z}, z) \leq (1 + o(1)) \frac{\sigma^2}{2np}$$

with high probability.

Main Results

Theorem [G & Zhang]. Assume $\sigma^2 = o(np)$ and $\frac{np}{\log n} \to \infty$. Then,

$$\inf_{\widehat{Z}} \sup_{z \in \mathbb{C}_1^n} \mathbb{E}_z \frac{1}{n^2} \|\widehat{Z} - zz^{\mathrm{H}}\|_{\mathrm{F}}^2 \geq (1 - o(1)) \frac{\sigma^2}{np} .$$

Main Results

Theorem [G & Zhang]. Assume $\sigma^2 = o(np)$ and $\frac{np}{\log n} \to \infty$. Then,

$$\inf_{\widehat{Z}} \sup_{z \in \mathbb{C}_1^n} \mathbb{E}_z \frac{1}{n^2} \|\widehat{Z} - zz^{\mathrm{H}}\|_{\mathrm{F}}^2 \geq (1 - o(1)) \frac{\sigma^2}{np}.$$

Moreover, SDP achieves

$$\frac{1}{n^2} \|\widehat{Z} - zz^{\mathrm{H}}\|_{\mathrm{F}}^2 \leq (1 + o(1)) \frac{\sigma^2}{np}$$

with high probability.

How to prove the results?

Statistical and Computational Guarantees of Lloyd's Algorithm and Its Variants

Yu Lu[1] and Harrison H. Zhou[1]

[1] *Yale University*

December 8, 2016

Abstract

Clustering is a fundamental problem in statistics and machine learning. Lloyd's algorithm, proposed in 1957, is still possibly the most widely used clustering algorithm in practice due to its simplicity and empirical performance. However, there has been little theoretical investigation on the statistical and computational guarantees of Lloyd's algorithm. This paper is an attempt to bridge this gap between practice and theory. We investigate the performance of Lloyd's algorithm on clustering sub-Gaussian mixtures. Under an appropriate initialization for labels or centers, we show that Lloyd's algorithm converges to an exponentially small clustering error after an order of $\log n$ iterations, where n is the sample size. The error rate is shown to be minimax optimal. For the two-mixture case, we only require the initializer to be slightly better than random guess.

In addition, we extend the Lloyd's algorithm and its analysis to community detection and crowdsourcing, two problems that have received a lot of attention recently in statistics and machine learning. Two variants of Lloyd's algorithm are proposed respectively for community detection and crowdsourcing. On the theoretical side, we provide statistical and computational guarantees of the two algorithms, and the results improve upon some previous signal-to-noise ratio conditions in literature for both problems. Experimental results on simulated and real data sets demonstrate competitive performance of our algorithms to the state-of-the-art methods.

1 Introduction

Lloyd's algorithm, proposed in 1957 by Stuart Lloyd at Bell Labs [40], is still one of the most popular clustering algorithms used by practitioners, with a wide range of applications from computer vision [3], to astronomy [45] and to biology [26]. Although considerable innovations have been made on developing new provable and efficient clustering algorithms in the past six decades, Lloyd's algorithm has been consistently listed as one of the top ten data mining algorithms in several recent surveys [55].

Structured Linear Models

$$Y = \mathscr{X}_z(B) + w$$

[Gao, van der Vaart & Zhou 15]

Structured Linear Models

structure

$$Y = \mathscr{X}_z(B) + w$$

[Gao, van der Vaart & Zhou 15]

Structured Linear Models

structure

$$Y = \mathscr{X}_z(B) + w$$

linear operator

[Gao, van der Vaart & Zhou 15]

Structured Linear Models

structure

$$Y = \mathscr{X}_z(B) + w$$

linear operator

parameter

[Gao, van der Vaart & Zhou 15]

Structured Linear Models

structure

$$Y = \mathscr{X}_z(B) + w$$

linear operator

parameter

$$B \in \mathcal{B}_z$$

[Gao, van der Vaart & Zhou 15]

Structured Linear Models

Clustering $$Y_i \sim N(\theta_{z_i}, I_d)$$

Ranking $$Y_{ij} \sim N(\beta(z_i - z_j), 1)$$

Regression $$Y_i \sim N(X_i^T \beta, 1)$$

Structured Linear Models

Clustering

$$Y_i \sim N(\theta_{z_i}, I_d)$$

$$\mathscr{X}_z : [\theta_1, ..., \theta_k] \mapsto [\theta_{z_1}, ..., \theta_{z_n}]$$

Ranking

$$Y_{ij} \sim N(\beta(z_i - z_j), 1)$$

Regression

$$Y_i \sim N(X_i^T \beta, 1)$$

Structured Linear Models

Clustering

$$Y_i \sim N(\theta_{z_i}, I_d)$$

$$\mathscr{X}_z : [\theta_1, ..., \theta_k] \mapsto [\theta_{z_1}, ..., \theta_{z_n}]$$

Ranking

$$Y_{ij} \sim N(\beta(z_i - z_j), 1)$$

$$\mathscr{X}_z : \beta \mapsto (\beta(z_i - z_j))_{ij}$$

Regression

$$Y_i \sim N(X_i^T \beta, 1)$$

Structured Linear Models

Clustering

$$Y_i \sim N(\theta_{z_i}, I_d)$$

$$\mathscr{X}_z : [\theta_1, \ldots, \theta_k] \mapsto [\theta_{z_1}, \ldots, \theta_{z_n}]$$

Ranking

$$Y_{ij} \sim N(\beta(z_i - z_j), 1)$$

$$\mathscr{X}_z : \beta \mapsto (\beta(z_i - z_j))_{ij}$$

Regression

$$Y_i \sim N(X_i^T \beta, 1)$$

$$\mathscr{X}_z : \beta \mapsto X\beta$$

Iterative Algorithm

Iterative Algorithm

some local statistic T_j

Iterative Algorithm

T_j

Iterative Algorithm

$$\mathbb{E}T_j = \mu_j(B, z_j)$$

Iterative Algorithm

$$\mathbb{E}T_j = \mu_j(B, z_j)$$

$$\widehat{z}_j = \underset{a}{\operatorname{argmin}} \|T_j - \mu_j(B, a)\|^2$$

Iterative Algorithm

$$\mathbb{E}T_j = \mu_j(B, z_j)$$

$$\widehat{z}_j = \operatorname*{argmin}_a \|T_j - \mu_j(B, a)\|^2$$

unknown

Iterative Algorithm

$$\mathbb{E}T_j = \mu_j(B, z_j)$$

$$\widehat{z}_j = \operatorname*{argmin}_a \|T_j - \mu_j(B, a)\|^2$$

unknown, but $B \in \mathcal{B}_z$

Iterative Algorithm

$$\mathbb{E}T_j = \mu_j(B, z_j)$$

$$z_j^{(t)} = \underset{a}{\operatorname{argmin}} \|T_j - \mu_j(\widehat{B}(z^{(t-1)}), a)\|^2$$

Iterative Algorithm

$$\mathbb{E}T_j = \mu_j(B, z_j)$$

$$z_j^{(t)} = \operatorname*{argmin}_a \|T_j - \mu_j(\widehat{B}(z^{(t-1)}), a)\|^2$$

$$\widehat{B}(z) = \operatorname*{argmin}_{B \in \mathcal{B}_z} \|Y - \mathscr{X}_z(B)\|^2$$

Iterative Algorithm

$$\mathbb{E}T_j = \mu_j(B, z_j)$$

$$\begin{cases} z_j^{(t)} = \underset{a}{\operatorname{argmin}} \|T_j - \mu_j(\widehat{B}(z^{(t-1)}), a)\|^2 \\ \\ \widehat{B}(z) = \underset{B \in \mathcal{B}_z}{\operatorname{argmin}} \|Y - \mathscr{X}_z(B)\|^2 \end{cases}$$

Conditions

Conditions

loss function

$$\ell(z, z^*) \quad = \quad \sum_j \|\mu_j(B^*, z_j) - \mu_j(B^*, z_j^*)\|^2$$

Conditions

loss function

$$\begin{aligned}\ell(z,z^*) &= \sum_j \|\mu_j(B^*,z_j)-\mu_j(B^*,z_j^*)\|^2 \\ &\geq \Delta_{\min}^2 \sum_j \mathbf{1}_{\{z_j\neq z_j^*\}}\end{aligned}$$

Conditions

loss function

$$\ell(z, z^*) \quad = \quad \sum_j \|\mu_j(B^*, z_j) - \mu_j(B^*, z_j^*)\|^2$$

error conditions

$$\frac{\mathsf{diff}\left(\mu_j(B(z), a), \mu_j(B(z^*), a)\right)}{\ell(z, z^*)} = o_{\mathbb{P}}(1)$$

$$\mathsf{diff}\left(\mu_j(\widehat{B}(z^*), a), \mu_j(B^*, a)\right) = o_{\mathbb{P}}(1)$$

Conditions

loss function

$$\ell(z, z^*) \quad = \quad \sum_j \|\mu_j(B^*, z_j) - \mu_j(B^*, z_j^*)\|^2$$

error conditions

$$\max_{\{z:\ell(z,z^*)\leq\tau\}} \frac{\mathsf{diff}\left(\mu_j(B(z), a), \mu_j(B(z^*), a)\right)}{\ell(z, z^*)} = o_{\mathbb{P}}(1)$$

$$\mathsf{diff}\left(\mu_j(\widehat{B}(z^*), a), \mu_j(B^*, a)\right) = o_{\mathbb{P}}(1)$$

Convergence

Theorem [G & Zhang]. Assume

$$\ell(z^{(0)}, z^*) \leq \tau$$

and the conditions hold with the same τ.

Convergence

Theorem [G & Zhang]. Assume

$$\ell(z^{(0)}, z^*) \leq \tau$$

and the conditions hold with the same τ. Then, we have

$$\ell(z^{(t)}, z^*) \leq [\text{ideal error}] + \frac{1}{2}\ell(z^{(t-1)}, z^*)$$

for all $t \geq 1$.

Iterative Algorithm

$$\mathbb{E}T_j = \mu_j(B, z_j)$$

$$\begin{cases} z_j^{(t)} = \underset{a}{\operatorname{argmin}} \|T_j - \mu_j(\widehat{B}(z^{(t-1)}), a)\|^2 \\ \\ \widehat{B}(z) = \underset{B \in \mathcal{B}_z}{\operatorname{argmin}} \|Y - \mathscr{X}_z(B)\|^2 \end{cases}$$

Clustering

model $$Y_i \sim N(\theta_{z_i}, I_d)$$

Clustering

model $$Y_i \sim N(\theta_{z_i}, I_d)$$

specialization $$T_i = Y_i$$

Clustering

model $$Y_i \sim N(\theta_{z_i}, I_d)$$

specialization $$T_i = Y_i \qquad \mu_i(B, a) = \theta_a$$

Clustering

model $$Y_i \sim N(\theta_{z_i}, I_d)$$

specialization $$T_i = Y_i \qquad \mu_i(B, a) = \theta_a$$

algorithm
$$\begin{cases} z_i^{(t)} = \underset{a \in [k]}{\operatorname{argmin}} \, \|Y_i - \widehat{\theta}_a(z^{(t-1)})\|^2 \\ \\ \widehat{\theta}_a(z) = \dfrac{\sum_{i=1}^n \mathbf{1}_{\{z_i = a\}} Y_i}{\sum_{i=1}^n \mathbf{1}_{\{z_i = a\}}} \end{cases}$$

Clustering

model $$Y_i \sim N(\theta_{z_i}, I_d)$$

specialization $$T_i = Y_i \qquad \mu_i(B,a) = \theta_a$$

algorithm
$$\begin{cases} z_i^{(t)} = \underset{a \in [k]}{\operatorname{argmin}} \|Y_i - \widehat{\theta}_a(z^{(t-1)})\|^2 \\ \\ \widehat{\theta}_a(z) = \dfrac{\sum_{i=1}^n \mathbf{1}_{\{z_i=a\}} Y_i}{\sum_{i=1}^n \mathbf{1}_{\{z_i=a\}}} \end{cases}$$

[Lloyd 57]

Clustering

Theorem Assume $k = o(n)$, and

Clustering

Theorem Assume $k = o(n)$, and

$$\frac{\min_{a\neq b}\|\theta_a^* - \theta_b^*\|^2}{k^2(kd/n+1)} \to \infty.$$

Clustering

Theorem Assume $k = o(n)$, and

$$\frac{\min_{a \neq b} \|\theta_a^* - \theta_b^*\|^2}{k^2(kd/n+1)} \to \infty .$$

Initialize by spectral clustering, and then

$$\frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\left\{z_i^{(t)} \neq z_i^*\right\}} \leq \exp\left(-(1+o(1))\frac{\min_{a \neq b} \|\theta_a^* - \theta_b^*\|^2}{8}\right) + 2^{-t}$$

for all $t \geq 1$ with high probability.

Clustering

Theorem Assume $k = o(n)$, and

$$\frac{\min_{a \neq b} \|\theta_a^* - \theta_b^*\|^2}{k^2(kd/n + 1)} \to \infty .$$

Initialize by spectral clustering, and then

$$\frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\left\{z_i^{(t)} \neq z_i^*\right\}} \leq \exp\left(-(1+o(1))\frac{\min_{a \neq b} \|\theta_a^* - \theta_b^*\|^2}{8}\right) + 2^{-t}$$

for all $t \geq 1$ with high probability.

minimax rate

Clustering

Theorem [Lu & Zhou 16]. Assume $k = o(n)$, and

$$\frac{\min_{a \neq b} \|\theta_a^* - \theta_b^*\|^2}{k^2(kd/n + 1)} \to \infty \quad .$$

Initialize by spectral clustering, and then

$$\frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\left\{z_i^{(t)} \neq z_i^*\right\}} \leq \exp\left(-(1+o(1))\frac{\min_{a \neq b} \|\theta_a^* - \theta_b^*\|^2}{8}\right) + 2^{-t}$$

for all $t \geq 1$ with high probability.

minimax rate

Phase Synchronization

$$\boxed{Y = zz^{\mathrm{H}} + \sigma W}$$

Phase Synchronization

$$\boxed{Y = zz^{\mathrm{H}} + \sigma W}$$

quadratic

Phase Synchronization

$$\boxed{Y = zz^{\mathrm{H}} + \sigma W}$$

quadratic

structure

$$Y = \mathscr{X}_z(B) + w$$

linear operator

parameter

$$B \in \mathcal{B}_z$$

Phase Synchronization

$$\boxed{Y = zz^{\mathrm{H}} + \sigma W}$$

quadratic

structure

$$Y = \mathscr{X}_z(B) + w$$

can deal with structure
parameters in two positions

linear operator

parameter

$$B \in \mathcal{B}_z$$

Phase Synchronization

$$Y = z\beta^{\mathrm{H}} + \sigma W$$

Phase Synchronization

$$Y = z\beta^{\mathrm{H}} + \sigma W \qquad \mathscr{X}_z : \beta \mapsto z\beta^{\mathrm{H}}$$

Phase Synchronization

$$Y = z\beta^{\mathrm{H}} + \sigma W \qquad \mathscr{X}_z : \beta \mapsto z\beta^{\mathrm{H}}$$

$$\beta \in \mathcal{B}_z = \{z\}$$

Phase Synchronization

$$Y = z\beta^{\mathrm{H}} + \sigma W \qquad \mathscr{X}_z : \beta \mapsto z\beta^{\mathrm{H}}$$

$$\beta \in \mathcal{B}_z = \{z\}$$

linear in z (instead of quadratic)

Phase Synchronization

$$Y = z\beta^{\mathrm{H}} + \sigma W \qquad \mathscr{X}_z : \beta \mapsto z\beta^{\mathrm{H}}$$

$$\beta \in \mathcal{B}_z = \{z\}$$

linear in z (instead of quadratic)

specialization $T_i = Y_i$

Phase Synchronization

$$Y = z\beta^{\mathrm{H}} + \sigma W \qquad \mathscr{X}_z : \beta \mapsto z\beta^{\mathrm{H}}$$

$$\beta \in \mathcal{B}_z = \{z\}$$

linear in z (instead of quadratic)

specialization $\qquad T_i = Y_i \qquad \mu(\beta, a) = a\beta$

Phase Synchronization

$$Y = z\beta^{\mathrm{H}} + \sigma W \qquad \mathscr{X}_z : \beta \mapsto z\beta^{\mathrm{H}}$$

$$\beta \in \mathcal{B}_z = \{z\}$$

linear in z (instead of quadratic)

specialization $\quad T_i = Y_i \qquad \mu(\beta, a) = a\beta$

$$\begin{cases} \beta^{(t)} = \underset{\beta \in \{z^{(t-1)}\}}{\operatorname{argmin}} \|Y - z^{(t-1)}\beta^{\mathrm{H}}\|_{\mathrm{F}}^2 \\ z_j^{(t)} = \underset{a \in \mathbb{C}_1}{\operatorname{argmin}} \|Y_j - a\beta^{(t)}\|^2 \end{cases}$$

Phase Synchronization

$$Y = z\beta^{\mathrm{H}} + \sigma W \qquad \mathscr{X}_z : \beta \mapsto z\beta^{\mathrm{H}}$$

$$\beta \in \mathcal{B}_z = \{z\}$$

linear in z (instead of quadratic)

specialization $\quad T_i = Y_i \qquad \mu(\beta, a) = a\beta$

$$\begin{cases} \beta^{(t)} = \underset{\beta \in \{z^{(t-1)}\}}{\operatorname{argmin}} \|Y - z^{(t-1)}\beta^{\mathrm{H}}\|_{\mathrm{F}}^2 = z^{(t-1)} \\ z_j^{(t)} = \underset{a \in \mathbb{C}_1}{\operatorname{argmin}} \|Y_j - a\beta^{(t)}\|^2 \end{cases}$$

Phase Synchronization

$$Y = z\beta^{\mathrm{H}} + \sigma W \qquad \mathscr{X}_z : \beta \mapsto z\beta^{\mathrm{H}}$$

$$\beta \in \mathcal{B}_z = \{z\}$$

linear in z (instead of quadratic)

specialization $\qquad T_i = Y_i \qquad \mu(\beta, a) = a\beta$

$$\begin{cases} \beta^{(t)} = \underset{\beta \in \{z^{(t-1)}\}}{\operatorname{argmin}} \|Y - z^{(t-1)}\beta^{\mathrm{H}}\|_{\mathrm{F}}^2 = z^{(t-1)} \\ z_j^{(t)} = \underset{a \in \mathbb{C}_1}{\operatorname{argmin}} \|Y_j - a\beta^{(t)}\|^2 = \underset{a \in \mathbb{C}_1}{\operatorname{argmin}} \|Y_j - az^{(t-1)}\|^2 \end{cases}$$

Phase Synchronization

$$Y = z\beta^{\mathrm{H}} + \sigma W \qquad \mathscr{X}_z : \beta \mapsto z\beta^{\mathrm{H}}$$

$$\beta \in \mathcal{B}_z = \{z\}$$

linear in z (instead of quadratic)

specialization $\quad T_i = Y_i \qquad \mu(\beta, a) = a\beta$

$$z_j^{(t)} = \frac{\sum_{k\in[n]\setminus\{j\}} Y_{jk} z_k^{(t-1)}}{\left|\sum_{k\in[n]\setminus\{j\}} Y_{jk} z_k^{(t-1)}\right|}$$

Phase Synchronization

$$Y = z\beta^{\mathrm{H}} + \sigma W \qquad \mathscr{X}_z : \beta \mapsto z\beta^{\mathrm{H}}$$

$$\beta \in \mathcal{B}_z = \{z\}$$

linear in z (instead of quadratic)

specialization $\quad T_i = Y_i \qquad \mu(\beta, a) = a\beta$

$$z_j^{(t)} = \frac{\sum_{k\in[n]\setminus\{j\}} A_{jk} Y_{jk} z_k^{(t-1)}}{\left|\sum_{k\in[n]\setminus\{j\}} A_{jk} Y_{jk} z_k^{(t-1)}\right|}$$

with missing data

Phase Synchronization

$$z_i^{(t)} = \frac{[(A \circ Y) z^{(t-1)}]_i}{|[(A \circ Y) z^{(t-1)}]_i|}$$

Phase Synchronization

$$z^{(t)} = f(z^{(t-1)}) \quad \Longleftrightarrow \quad z_i^{(t)} = \frac{[(A \circ Y) z^{(t-1)}]_i}{|[(A \circ Y) z^{(t-1)}]_i|}$$

Phase Synchronization

$$z^{(t)} = f(z^{(t-1)}) \quad \Longleftrightarrow \quad z_i^{(t)} = \frac{[(A \circ Y)z^{(t-1)}]_i}{|[(A \circ Y)z^{(t-1)}]_i|}$$

Lemma. Assume $\sigma^2 = o(np)$ and $p \gg \frac{\log n}{n}$. For any $\gamma = o(1)$, we have

$$\mathbb{P}\left(\ell(f(z), z^*) \leq \delta \ell(z, z^*) + (1+\delta)\frac{\sigma^2}{2np} \text{ for any } z \in \mathbb{C}_1^n \text{ s.t. } \ell(z, z^*) \leq \gamma\right) \geq 1 - \delta$$

for some $\delta = o(1)$.

Phase Synchronization

Corollary. Assume $\sigma^2 = o(np)$ and $p \gg \frac{\log n}{n}$.
Initialized by PCA, the power method satisfies

$$\ell(\widehat{z}, z) \leq (1 + o(1))\frac{\sigma^2}{2np}$$

with high probability after $\log\left(\frac{np}{\sigma^2}\right)$ iterations.

Phase Synchronization

MLE $$\max_{z \in \mathbb{C}_1^n} z^{\mathrm{H}} (A \circ Y) z$$

Phase Synchronization

MLE $\max_{z \in \mathbb{C}_1^n} z^{\mathrm{H}}(A \circ Y)z$

$$\widehat{z} = f(\widehat{z})$$

Phase Synchronization

MLE $\max_{z \in \mathbb{C}_1^n} z^{\mathrm{H}}(A \circ Y)z$ $\widehat{z} = f(\widehat{z})$

Corollary. Assume $\sigma^2 = o(np)$ and $p \gg \frac{\log n}{n}$.
The MLE satisfies

$$\ell(\widehat{z}, z) \leq (1 + o(1))\frac{\sigma^2}{2np}$$

with high probability.

SDP: A Non-Convex View

SDP: A Non-Convex View

MLE

$$\boxed{\begin{array}{c} \max \mathsf{Tr}((A \circ Y) z z^{\mathrm{H}}) \\ \\ \text{s.t. } |z_j| = 1 \text{ for all } j \in [n] \end{array}}$$

SDP: A Non-Convex View

MLE

$$\boxed{\begin{array}{c} \max \mathsf{Tr}((A \circ Y) z z^{\mathrm{H}}) \\ \\ \text{s.t. } |z_j| = 1 \text{ for all } j \in [n] \end{array}}$$

GPM

$$\boxed{z_j^{(t)} = \frac{\sum_{k \in [n] \setminus \{j\}} A_{jk} Y_{jk} z_k^{(t-1)}}{\left| \sum_{k \in [n] \setminus \{j\}} A_{jk} Y_{jk} z_k^{(t-1)} \right|}}$$

SDP: A Non-Convex View

MLE

$$\max \mathsf{Tr}((A \circ Y) z z^{\mathrm{H}})$$

$$\text{s.t. } |z_j| = 1 \text{ for all } j \in [n]$$

fixed point

GPM

$$z_j^{(t)} = \frac{\sum_{k \in [n] \setminus \{j\}} A_{jk} Y_{jk} z_k^{(t-1)}}{\left| \sum_{k \in [n] \setminus \{j\}} A_{jk} Y_{jk} z_k^{(t-1)} \right|}$$

SDP: A Non-Convex View

MLE

$$\max \operatorname{Tr}((A \circ Y) z z^{\mathrm{H}})$$

$$\text{s.t. } |z_j| = 1 \text{ for all } j \in [n]$$

fixed point

GPM

$$z_j^{(t)} = \frac{\sum_{k \in [n] \setminus \{j\}} A_{jk} Y_{jk} z_k^{(t-1)}}{\left| \sum_{k \in [n] \setminus \{j\}} A_{jk} Y_{jk} z_k^{(t-1)} \right|}$$

SDP

$$\max \operatorname{Tr}((A \circ Y) Z)$$

$$\text{s.t. } \operatorname{diag}(Z) = I_n \text{ and } Z \succeq 0$$

SDP: A Non-Convex View

MLE

$$\max \mathsf{Tr}((A \circ Y) z z^{\mathrm{H}})$$

$$\text{s.t. } |z_j| = 1 \text{ for all } j \in [n]$$

fixed point

GPM

$$z_j^{(t)} = \frac{\sum_{k \in [n] \setminus \{j\}} A_{jk} Y_{jk} z_k^{(t-1)}}{\left| \sum_{k \in [n] \setminus \{j\}} A_{jk} Y_{jk} z_k^{(t-1)} \right|}$$

SDP

$$\max \mathsf{Tr}((A \circ Y) Z)$$

$$\text{s.t. } \operatorname{diag}(Z) = I_n \text{ and } Z \succeq 0$$

$$Z = V^{\mathrm{H}} V$$

SDP: A Non-Convex View

MLE

$$\max \mathsf{Tr}((A \circ Y) z z^{\mathrm{H}})$$

$$\text{s.t. } |z_j| = 1 \text{ for all } j \in [n]$$

fixed point

GPM

$$z_j^{(t)} = \frac{\sum_{k \in [n] \setminus \{j\}} A_{jk} Y_{jk} z_k^{(t-1)}}{\left| \sum_{k \in [n] \setminus \{j\}} A_{jk} Y_{jk} z_k^{(t-1)} \right|}$$

SDP

$$\max \mathsf{Tr}((A \circ Y) Z)$$

$$\text{s.t. } \operatorname{diag}(Z) = I_n \text{ and } Z \succeq 0$$

$$Z = V^{\mathrm{H}} V$$

$$\max_{V \in \mathbb{C}^{n \times n}} \mathsf{Tr}((A \circ Y) V^{\mathrm{H}} V)$$

$$\text{s.t. } \|V_j\|^2 = 1 \text{ for all } j \in [n]$$

SDP: A Non-Convex View

MLE

$$\max \mathsf{Tr}((A \circ Y) z z^{\mathrm{H}})$$

$$\text{s.t. } |z_j| = 1 \text{ for all } j \in [n]$$

fixed point

GPM

$$z_j^{(t)} = \frac{\sum_{k \in [n]\setminus\{j\}} A_{jk} Y_{jk} z_k^{(t-1)}}{\left|\sum_{k \in [n]\setminus\{j\}} A_{jk} Y_{jk} z_k^{(t-1)}\right|}$$

SDP

$$\max \mathsf{Tr}((A \circ Y) Z)$$

$$\text{s.t. } \operatorname{diag}(Z) = I_n \text{ and } Z \succeq 0$$

$$Z = V^{\mathrm{H}} V$$

$$\max_{V \in \mathbb{C}^{n \times n}} \mathsf{Tr}((A \circ Y) V^{\mathrm{H}} V)$$

$$\text{s.t. } \|V_j\|^2 = 1 \text{ for all } j \in [n]$$

[Burer & Monteiro 03]

SDP: A Non-Convex View

MLE

$$\max \mathsf{Tr}((A \circ Y) z z^{\mathrm{H}})$$
$$\text{s.t. } |z_j| = 1 \text{ for all } j \in [n]$$

fixed point

GPM

$$z_j^{(t)} = \frac{\sum_{k \in [n]\backslash\{j\}} A_{jk} Y_{jk} z_k^{(t-1)}}{\left|\sum_{k \in [n]\backslash\{j\}} A_{jk} Y_{jk} z_k^{(t-1)}\right|}$$

SDP

$$\max \mathsf{Tr}((A \circ Y) Z)$$
$$\text{s.t. } \operatorname{diag}(Z) = I_n \text{ and } Z \succeq 0$$

$$Z = V^{\mathrm{H}} V$$

$$\max_{V \in \mathbb{C}^{n \times n}} \mathsf{Tr}((A \circ Y) V^{\mathrm{H}} V)$$
$$\text{s.t. } \|V_j\|^2 = 1 \text{ for all } j \in [n]$$

$$V_j^{(t)} = \frac{\sum_{k \in [n]\backslash\{j\}} A_{jk} \bar{Y}_{jk} V_k^{(t-1)}}{\left\|\sum_{k \in [n]\backslash\{j\}} A_{jk} \bar{Y}_{jk} V_k^{(t-1)}\right\|}$$

[Burer & Monteiro 03]

SDP: A Non-Convex View

MLE

$$\max \mathsf{Tr}((A \circ Y) z z^{\mathrm{H}})$$

$$\text{s.t. } |z_j| = 1 \text{ for all } j \in [n]$$

fixed point

GPM

$$z_j^{(t)} = \frac{\sum_{k \in [n] \setminus \{j\}} A_{jk} Y_{jk} z_k^{(t-1)}}{\left| \sum_{k \in [n] \setminus \{j\}} A_{jk} Y_{jk} z_k^{(t-1)} \right|}$$

SDP

$$\max \mathsf{Tr}((A \circ Y) Z)$$

$$\text{s.t. } \operatorname{diag}(Z) = I_n \text{ and } Z \succeq 0$$

$Z = V^{\mathrm{H}} V$

$$\max_{V \in \mathbb{C}^{n \times n}} \mathsf{Tr}((A \circ Y) V^{\mathrm{H}} V)$$

$$\text{s.t. } \|V_j\|^2 = 1 \text{ for all } j \in [n]$$

fixed point

$$V_j^{(t)} = \frac{\sum_{k \in [n] \setminus \{j\}} A_{jk} \bar{Y}_{jk} V_k^{(t-1)}}{\left\| \sum_{k \in [n] \setminus \{j\}} A_{jk} \bar{Y}_{jk} V_k^{(t-1)} \right\|}$$

[Burer & Monteiro 03]

SDP: A Non-Convex View

$$V_j^{(t)} = \frac{\sum_{k\in[n]\setminus\{j\}} A_{jk}\bar{Y}_{jk}V_k^{(t-1)}}{\left\|\sum_{k\in[n]\setminus\{j\}} A_{jk}\bar{Y}_{jk}V_k^{(t-1)}\right\|}$$

SDP: A Non-Convex View

$$V^{(t)} = f(V^{(t-1)}) \iff V_j^{(t)} = \frac{\sum_{k \in [n] \setminus \{j\}} A_{jk} \bar{Y}_{jk} V_k^{(t-1)}}{\left\| \sum_{k \in [n] \setminus \{j\}} A_{jk} \bar{Y}_{jk} V_k^{(t-1)} \right\|}$$

SDP: A Non-Convex View

$$V^{(t)} = f(V^{(t-1)}) \quad \Longleftrightarrow \quad V_j^{(t)} = \frac{\sum_{k\in[n]\backslash\{j\}} A_{jk}\bar{Y}_{jk}V_k^{(t-1)}}{\left\|\sum_{k\in[n]\backslash\{j\}} A_{jk}\bar{Y}_{jk}V_k^{(t-1)}\right\|}$$

Lemma. Assume $\sigma^2 = o(np)$ and $p \gg \frac{\log n}{n}$. For any $\gamma = o(1)$, we have

$$\mathbb{P}\left(\ell(f(V),z^*) \leq \delta\ell(V,z^*) + (1+\delta)\frac{\sigma^2}{2np} \text{ for any } z \in \mathbb{C}_1^n \text{ s.t. } \ell(V,z^*) \leq \gamma\right) \geq 1-\delta$$

for some $\delta = o(1)$.

SDP: A Non-Convex View

$$V^{(t)} = f(V^{(t-1)}) \quad \Longleftrightarrow \quad V_j^{(t)} = \frac{\sum_{k\in[n]\setminus\{j\}} A_{jk}\bar{Y}_{jk}V_k^{(t-1)}}{\left\|\sum_{k\in[n]\setminus\{j\}} A_{jk}\bar{Y}_{jk}V_k^{(t-1)}\right\|}$$

Lemma. Assume $\sigma^2 = o(np)$ and $p \gg \frac{\log n}{n}$. For any $\gamma = o(1)$, we have

$$\mathbb{P}\left(\ell(f(V), z^*) \leq \delta\ell(V, z^*) + (1+\delta)\frac{\sigma^2}{2np} \text{ for any } z \in \mathbb{C}_1^n \text{ s.t. } \ell(V, z^*) \leq \gamma\right) \geq 1 - \delta$$

for some $\delta = o(1)$.

$$\boxed{\ell(\widehat{V}, z^*) = \min_{a\in\mathbb{C}^n:\|a\|^2=1} \frac{1}{n}\sum_{j=1}^n \|\widehat{V}_j - \bar{z}_j^* a\|^2}$$

Phase Synchronization

SDP

$$\max \mathsf{Tr}((A \circ Y)Z)$$

$$\text{s.t.} \quad \operatorname{diag}(Z) = I_n \text{ and } Z \succeq 0$$

Phase Synchronization

SDP

$$\max \mathsf{Tr}((A \circ Y)Z)$$

$$\text{s.t.} \quad \text{diag}(Z) = I_n \text{ and } Z \succeq 0$$

$$\xrightarrow{\widehat{Z} = \widehat{V}^{\mathrm{H}}\widehat{V}}$$

$$\widehat{V} = f(\widehat{V})$$

Phase Synchronization

SDP

$$\max \mathsf{Tr}((A \circ Y)Z)$$

$$\text{s.t.} \quad \text{diag}(Z) = I_n \text{ and } Z \succeq 0$$

$$\xrightarrow{\widehat{Z} = \widehat{V}^{\mathrm{H}} \widehat{V}}$$

$$\widehat{V} = f(\widehat{V})$$

Corollary. Assume $\sigma^2 = o(np)$ and $p \gg \frac{\log n}{n}$.
The SDP satisfies

$$\begin{aligned} \ell(\widehat{V}, z) &\leq (1 + o(1)) \frac{\sigma^2}{2np}, \\ \frac{1}{n^2} \|\widehat{Z} - zz^{\mathrm{H}}\|_{\mathrm{F}}^2 &\leq (1 + o(1)) \frac{\sigma^2}{np}, \\ \ell(\widehat{z}, z) &\leq (1 + o(1)) \frac{\sigma^2}{2np} \end{aligned}$$

with high probability.

Phase Synchronization

SDP
$$\max \mathsf{Tr}((A \circ Y)Z)$$
$$\text{s.t.}\quad \mathrm{diag}(Z) = I_n \text{ and } Z \succeq 0$$

$$\xrightarrow{\widehat{Z} = \widehat{V}^{\mathrm{H}}\widehat{V}}$$

$$\widehat{V} = f(\widehat{V})$$

Corollary. Assume $\sigma^2 = o(np)$ and $p \gg \frac{\log n}{n}$.
The SDP satisfies

$$\begin{aligned} \ell(\widehat{V}, z) &\leq (1+o(1))\frac{\sigma^2}{2np}, \\ \frac{1}{n^2}\|\widehat{Z} - zz^{\mathrm{H}}\|_{\mathrm{F}}^2 &\leq (1+o(1))\frac{\sigma^2}{np}, \\ \ell(\widehat{z}, z) &\leq (1+o(1))\frac{\sigma^2}{2np} \end{aligned}$$

normalized
leading eigenvector

with high probability.

Lower Bound

Lower Bound

proof of lower bound

Lower Bound

proof of lower bound

$$\inf_{\widehat{z}\in\mathbb{C}_1^n} \sup_{z\in\mathbb{C}_1^n} \mathbb{E}_z \min_{\theta\in\mathbb{R}} \sum_{j=1}^{n} |\widehat{z}_j e^{i\theta} - z_j|^2$$

Lower Bound

proof of lower bound

$$\inf_{\widehat{z}\in\mathbb{C}_1^n}\sup_{z\in\mathbb{C}_1^n}\mathbb{E}_z\min_{\theta\in\mathbb{R}}\sum_{j=1}^{n}|\widehat{z}_je^{i\theta}-z_j|^2$$

not separable

Lower Bound

proof of lower bound

$$\inf_{\widehat{z}\in\mathbb{C}_1^n} \sup_{z\in\mathbb{C}_1^n} \mathbb{E}_z \min_{\theta\in\mathbb{R}} \sum_{j=1}^{n} |\widehat{z}_j e^{i\theta} - z_j|^2$$

not separable

$$\geq \quad \frac{1}{2n} \inf_{\widehat{z}\in\mathbb{C}_1^n} \sup_{z\in\mathbb{C}_1^n} \mathbb{E}_z \|\widehat{z}\widehat{z}^{\mathrm{H}} - zz^{\mathrm{H}}\|_{\mathrm{F}}^2$$

Lower Bound

proof of lower bound

$$\inf_{\widehat{z}\in\mathbb{C}_1^n} \sup_{z\in\mathbb{C}_1^n} \mathbb{E}_z \min_{\theta\in\mathbb{R}} \sum_{j=1}^{n} |\widehat{z}_j e^{i\theta} - z_j|^2$$

not separable

$$\geq \quad \frac{1}{2n} \inf_{\widehat{z}\in\mathbb{C}_1^n} \sup_{z\in\mathbb{C}_1^n} \mathbb{E}_z \|\widehat{z}\widehat{z}^{\mathrm{H}} - zz^{\mathrm{H}}\|_{\mathrm{F}}^2$$

separable

Lower Bound

proof of lower bound

$$\inf_{\widehat{z}\in\mathbb{C}_1^n} \sup_{z\in\mathbb{C}_1^n} \mathbb{E}_z \min_{\theta\in\mathbb{R}} \sum_{j=1}^{n} |\widehat{z}_j e^{i\theta} - z_j|^2$$

not separable

$$\geq \quad \frac{1}{2n} \inf_{\widehat{z}\in\mathbb{C}_1^n} \sup_{z\in\mathbb{C}_1^n} \mathbb{E}_z \|\widehat{z}\widehat{z}^{\mathrm{H}} - zz^{\mathrm{H}}\|_{\mathrm{F}}^2$$

separable

$$\geq \quad \frac{1}{2n} \inf_{\widehat{z}} \sum_{1\leq j\neq k\leq n} \left(\int \prod_{l=1}^{n} \pi(z_l) \right) \mathbb{E}_z |\widehat{z}_j \bar{\widehat{z}}_k - z_j \bar{z}_k|^2 dz$$

Lower Bound

proof of lower bound

$$\inf_{\widehat{z}\in\mathbb{C}_1^n} \sup_{z\in\mathbb{C}_1^n} \mathbb{E}_z \min_{\theta\in\mathbb{R}} \sum_{j=1}^n |\widehat{z}_j e^{i\theta} - z_j|^2$$

not separable

$$\geq \frac{1}{2n} \inf_{\widehat{z}\in\mathbb{C}_1^n} \sup_{z\in\mathbb{C}_1^n} \mathbb{E}_z \|\widehat{z}\widehat{z}^{\mathrm{H}} - zz^{\mathrm{H}}\|_{\mathrm{F}}^2$$

separable

$$\geq \frac{1}{2n} \inf_{\widehat{z}} \sum_{1\leq j\neq k\leq n} \left(\int \prod_{l=1}^n \pi(z_l)\right) \mathbb{E}_z |\widehat{z}_j \bar{\widehat{z}}_k - z_j \bar{z}_k|^2 dz$$

$$\geq \frac{1}{2n} \sum_{1\leq j\neq k\leq n} \mathbb{E}_{z_{-(j,k)}\sim\pi} \inf_{\widehat{T}} \int \pi(z_j)\pi(z_k)\mathbb{E}_z|\widehat{T} - z_j\bar{z}_k|^2 dz_j dz_k$$

Lower Bound

proof of lower bound

$$\inf_{\widehat{z}\in\mathbb{C}_1^n} \sup_{z\in\mathbb{C}_1^n} \mathbb{E}_z \min_{\theta\in\mathbb{R}} \sum_{j=1}^{n} |\widehat{z}_j e^{i\theta} - z_j|^2$$

not separable

$$\geq \frac{1}{2n} \inf_{\widehat{z}\in\mathbb{C}_1^n} \sup_{z\in\mathbb{C}_1^n} \mathbb{E}_z \|\widehat{z}\widehat{z}^{\mathrm{H}} - zz^{\mathrm{H}}\|_{\mathrm{F}}^2$$

separable

$$\geq \frac{1}{2n} \inf_{\widehat{z}} \sum_{1\leq j\neq k\leq n} \left(\int \prod_{l=1}^{n} \pi(z_l) \right) \mathbb{E}_z |\widehat{z}_j \bar{\widehat{z}}_k - z_j \bar{z}_k|^2 dz$$

$$\geq \frac{1}{2n} \sum_{1\leq j\neq k\leq n} \mathbb{E}_{z_{-(j,k)}\sim\pi} \inf_{\widehat{T}} \int \pi(z_j)\pi(z_k)\mathbb{E}_z |\widehat{T} - z_j \bar{z}_k|^2 dz_j dz_k$$

$$\geq \frac{1}{2n} \sum_{1\leq j\neq k\leq n} (1-\delta)\frac{\sigma^2}{p}$$

van Tree's inequality

Lower Bound

proof of lower bound

$$
\begin{aligned}
& \inf_{\widehat{z}\in\mathbb{C}_1^n} \sup_{z\in\mathbb{C}_1^n} \mathbb{E}_z \min_{\theta\in\mathbb{R}} \sum_{j=1}^{n} |\widehat{z}_j e^{i\theta} - z_j|^2 \\
\geq \quad & \frac{1}{2n} \inf_{\widehat{z}\in\mathbb{C}_1^n} \sup_{z\in\mathbb{C}_1^n} \mathbb{E}_z \|\widehat{z}\widehat{z}^{\mathrm{H}} - zz^{\mathrm{H}}\|_{\mathrm{F}}^2 \\
\geq \quad & \frac{1}{2n} \inf_{\widehat{z}} \sum_{1\leq j\neq k\leq n} \left(\int \prod_{l=1}^{n} \pi(z_l) \right) \mathbb{E}_z |\widehat{z}_j \bar{\widehat{z}}_k - z_j \bar{z}_k|^2 dz \\
\geq \quad & \frac{1}{2n} \sum_{1\leq j\neq k\leq n} \mathbb{E}_{z_{-(j,k)}\sim\pi} \inf_{\widehat{T}} \int \pi(z_j)\pi(z_k)\mathbb{E}_z |\widehat{T} - z_j \bar{z}_k|^2 dz_j dz_k \\
\geq \quad & \frac{1}{2n} \sum_{1\leq j\neq k\leq n} (1-\delta)\frac{\sigma^2}{p} \\
= \quad & (1-o(1))\frac{\sigma^2}{2p}
\end{aligned}
$$

not separable

separable

van Tree's inequality

Z2 Synchronization

Model

$$Y_{jk} = z_j z_k + \sigma W_{jk}$$

Z2 Synchronization

Model	$Y_{jk} = z_j z_k + \sigma W_{jk}$
Parameter space	$z_j \in \{-1, 1\}$

Z2 Synchronization

Model

$$Y_{jk} = z_j z_k + \sigma W_{jk}$$

Parameter space

$$z_j \in \{-1, 1\}$$

Missing data

$$A_{jk} \sim \mathrm{Bernoulli}(p)$$

Z2 Synchronization

MLE $$\max_{z \in \{-1,1\}^n} z^{\mathrm{T}} (A \circ Y) z$$

Z2 Synchronization

MLE

$$\max_{z \in \{-1,1\}^n} z^{\mathrm{T}}(A \circ Y)z$$

GPM

$$z_j^{(t)} = \frac{\sum_{k \in [n] \setminus \{j\}} A_{jk} Y_{jk} z_k^{(t-1)}}{\left| \sum_{k \in [n] \setminus \{j\}} A_{jk} Y_{jk} z_k^{(t-1)} \right|}$$

Z2 Synchronization

MLE

$$\max_{z \in \{-1,1\}^n} z^{\mathrm{T}} (A \circ Y) z$$

GPM

$$z_j^{(t)} = \frac{\sum_{k \in [n] \setminus \{j\}} A_{jk} Y_{jk} z_k^{(t-1)}}{\left| \sum_{k \in [n] \setminus \{j\}} A_{jk} Y_{jk} z_k^{(t-1)} \right|}$$

SDP

$$\max_{Z = Z^{\mathrm{T}} \in \mathbb{R}^{n \times n}} \mathsf{Tr}((A \circ Y) Z) \quad \text{subject to} \quad \mathrm{diag}(Z) = I_n \text{ and } Z \succeq 0$$

Z2 Synchronization

Theorem [G & Zhang]. Assume $\sigma^2 = o(np)$ and $\frac{np}{\log n} \to \infty$. Then,

$$\inf_{\widehat{z}\in\{-1,1\}^n} \sup_{z\in\{-1,1\}^n} \mathbb{E}_z \ell(\widehat{z}, z) \geq \exp\left(-(1+o(1))\frac{np}{2\sigma^2}\right).$$

Moreover, the MLE, GPM initialized by the leading eigenvector of $A \circ Y$, and the leading eigenvector of SDP all achieve

$$\ell(\widehat{z}, z) \leq \exp\left(-(1-o(1))\frac{np}{2\sigma^2}\right)$$

with high probability.

Z2 Synchronization

Theorem [G & Zhang]. Assume $\sigma^2 = o(np)$ and $\frac{np}{\log n} \to \infty$. Then,

$$\inf_{\widehat{z} \in \{-1,1\}^n} \sup_{z \in \{-1,1\}^n} \mathbb{E}_z \ell(\widehat{z}, z) \geq \exp\left(-(1+o(1))\frac{np}{2\sigma^2}\right).$$

Moreover, the MLE, GPM initialized by the leading eigenvector of $A \circ Y$, and the leading eigenvector of SDP all achieve

$$\ell(\widehat{z}, z) \leq \exp\left(-(1-o(1))\frac{np}{2\sigma^2}\right)$$

with high probability.

"p=1" by [Fei & Chen 20]

Comparison

Comparison

phase sync

$$\max \quad \mathsf{Tr}((A \circ Y)Z)$$

$$\text{s.t.} \quad \text{diag}(Z) = I_n \text{ and } Z \succeq 0$$

Comparison

phase sync

$$\max \quad \mathsf{Tr}((A \circ Y)Z)$$

$$\text{s.t.} \;\; \text{diag}(Z) = I_n \text{ and } Z \succeq 0$$

Z2 sync

$$\max \quad \mathsf{Tr}((A \circ Y)Z)$$

$$\text{s.t.} \;\; \text{diag}(Z) = I_n \text{ and } Z \succeq 0$$

Comparison

phase sync	$\max \quad \mathsf{Tr}((A \circ Y)Z)$ $\text{s.t.} \quad \text{diag}(Z) = I_n \text{ and } Z \succeq 0$	$(1 + o(1))\dfrac{\sigma^2}{2np}$
Z2 sync	$\max \quad \mathsf{Tr}((A \circ Y)Z)$ $\text{s.t.} \quad \text{diag}(Z) = I_n \text{ and } Z \succeq 0$	

Comparison

phase sync	$\max \quad \mathsf{Tr}((A \circ Y)Z)$ $\text{s.t.} \;\; \mathrm{diag}(Z) = I_n \text{ and } Z \succeq 0$	$(1 + o(1))\dfrac{\sigma^2}{2np}$
Z2 sync	$\max \quad \mathsf{Tr}((A \circ Y)Z)$ $\text{s.t.} \;\; \mathrm{diag}(Z) = I_n \text{ and } Z \succeq 0$	$\exp\left(-(1 - o(1))\dfrac{np}{2\sigma^2}\right)$

Comparison

phase sync	$\max_{\boxed{Z=Z^{\mathrm{H}}\in\mathbb{C}^{n\times n}}} \mathsf{Tr}((A\circ Y)Z)$ s.t. $\mathrm{diag}(Z)=I_n$ and $Z\succeq 0$	$(1+o(1))\dfrac{\sigma^2}{2np}$
Z2 sync	$\max_{\boxed{Z=Z^{\mathrm{T}}\in\mathbb{R}^{n\times n}}} \mathsf{Tr}((A\circ Y)Z)$ s.t. $\mathrm{diag}(Z)=I_n$ and $Z\succeq 0$	$\exp\left(-(1-o(1))\dfrac{np}{2\sigma^2}\right)$

Comparison

phase sync	$\max_{\boxed{Z=Z^{\mathrm{H}}\in\mathbb{C}^{n\times n}}} \mathsf{Tr}((A\circ Y)Z)$ s.t. $\mathrm{diag}(Z)=I_n$ and $Z\succeq 0$	$(1+o(1))\frac{\sigma^2}{2np}$
Z2 sync	$\max_{\boxed{Z=Z^{\mathrm{T}}\in\mathbb{R}^{n\times n}}} \mathsf{Tr}((A\circ Y)Z)$ s.t. $\mathrm{diag}(Z)=I_n$ and $Z\succeq 0$	$\exp\left(-(1-o(1))\frac{np}{2\sigma^2}\right)$

For any $x, y\in\mathbb{C}^n$ *such that* $\|y\|=1$ *and* $\mathrm{Re}(y^{\mathrm{H}}x)>0$, *we have*

$$\left\|\frac{x}{\|x\|}-y\right\|^2\le\frac{\|(I_n-yy^{\mathrm{H}})x\|^2+|\mathrm{Im}(y^{\mathrm{H}}x)|^2}{|\mathrm{Re}(y^{\mathrm{H}}x)|^2}$$

Comparison

phase sync	$\max_{\boxed{Z=Z^{\mathrm{H}}\in\mathbb{C}^{n\times n}}} \mathsf{Tr}((A\circ Y)Z)$ s.t. $\mathrm{diag}(Z)=I_n$ and $Z\succeq 0$	$(1+o(1))\dfrac{\sigma^2}{2np}$
Z2 sync	$\max_{\boxed{Z=Z^{\mathrm{T}}\in\mathbb{R}^{n\times n}}} \mathsf{Tr}((A\circ Y)Z)$ s.t. $\mathrm{diag}(Z)=I_n$ and $Z\succeq 0$	$\exp\left(-(1-o(1))\dfrac{np}{2\sigma^2}\right)$

For any $x,y\in\mathbb{C}^n$ *such that* $\|y\|=1$ *and* $\mathrm{Re}(y^{\mathrm{H}}x)>0$, *we have*

$$\left\|\frac{x}{\|x\|}-y\right\|^2\le\frac{\|(I_n-yy^{\mathrm{H}})x\|^2+\boxed{|\mathrm{Im}(y^{\mathrm{H}}x)|^2}}{|\mathrm{Re}(y^{\mathrm{H}}x)|^2}$$

extra term in complex space

Rotation Synchronization

Rotation Synchronization

Model $Y_{ij} = Z_i Z_j^{\mathrm{T}} + \sigma W_{ij} \in \mathbb{R}^{d \times d}$ $Z_i \in \mathsf{SO}(d)$

Rotation Synchronization

Model $Y_{ij} = Z_i Z_j^{\mathrm{T}} + \sigma W_{ij} \in \mathbb{R}^{d \times d}$ $Z_i \in \mathsf{SO}(d)$ $A_{jk} \sim \mathrm{Bernoulli}(p)$

Rotation Synchronization

Model $Y_{ij} = Z_i Z_j^{\mathrm{T}} + \sigma W_{ij} \in \mathbb{R}^{d \times d}$ $\quad Z_i \in \mathsf{SO}(d)$ $\quad A_{jk} \sim \mathrm{Bernoulli}(p)$

Loss

$$\ell(\widehat{Z}, Z) = \min_{U \in \mathsf{SO}(d)} \frac{1}{n} \sum_{j=1}^{n} \|\widehat{Z}_j U - Z_j\|_{\mathrm{F}}^2$$

Rotation Synchronization

Model $Y_{ij} = Z_i Z_j^{\mathrm{T}} + \sigma W_{ij} \in \mathbb{R}^{d\times d}$ $\quad Z_i \in \mathsf{SO}(d)$ $\quad A_{jk} \sim \mathrm{Bernoulli}(p)$

Loss $$\ell(\widehat{Z}, Z) = \min_{U\in\mathsf{SO}(d)} \frac{1}{n}\sum_{j=1}^{n} \|\widehat{Z}_j U - Z_j\|_{\mathrm{F}}^2$$

Theorem [G & Zhang]. Assume $\sigma^2 = o(np)$ and $\frac{np}{\log n} \to \infty$. Then,

$$\inf_{\widehat{Z}\in\mathsf{SO}(d)^n} \sup_{Z\in\mathsf{SO}(d)^n} \mathbb{E}_n \ell(\widehat{Z}, Z) \geq (1-o(1))\frac{(d-1)d\sigma^2}{2np}.$$

Moreover, both MLE and GPM achieve

$$\ell(\widehat{Z}, Z) \leq (1+o(1))\frac{(d-1)d\sigma^2}{2np}$$

with high probability.

Thank You