Clustering Mixture Models: Beyond Gaussians

Ainesh Bakshi
MIT

Central Open Question

Can we efficiently cluster mixtures of non-spherical log-concave distributions?

Plan :

1. Problem Setup
2. Bottlenecks
3. Recent Progress

Setup :

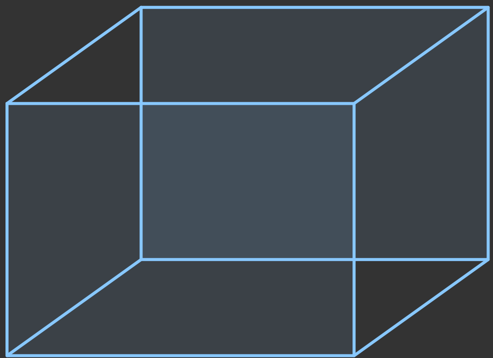
Let D be the uniform distribution over a convex body $K \in \mathbb{R}^d$ with mean μ and covariance Σ.

(log-concave)

Ex. ① Unit sphere

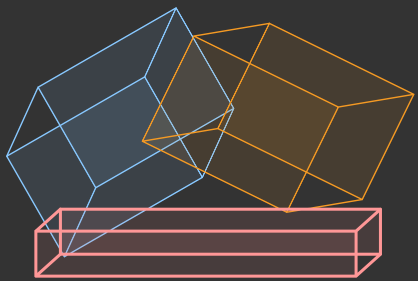
② Unit Cube

③ Simplex



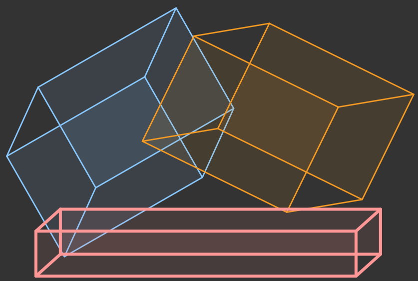
Setup :

Let $M = \sum_{i \in [k]} w_i \, D(\mu_i, \Sigma_i)$ be a mixture of log-concave distributions.



Setup :

Let $M = \sum_{i \in [k]} w_i \, D(\mu_i, \Sigma_i)$ be a mixture of log-concave distributions.

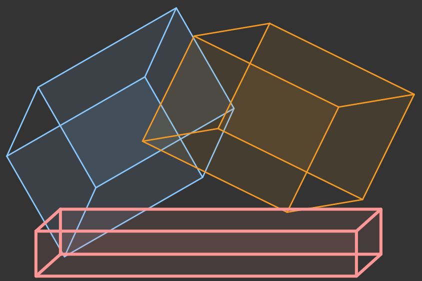


Definition: M is "clusterable" if $\forall\, i,j \in [k]$ $D(\mu_i, \Sigma_i)$ and $D(\mu_j, \Sigma_j)$ are either

1. Mean - seperated
2. Spectrally - seperated
3. Shell / Frobenius - seperated

Setup :

Definition: $\mathcal{M}$ is "clusterable" if $\forall\, i,j \in [k]$
$D(\mu_i, \Sigma_i)$ and $D(\mu_j, \Sigma_j)$ are either



1. Mean - seperated
2. Spectrally - seperated
3. Shell / Frobenius - seperated

- For Gaussians, this captures seperation in total-varition distance
  [B-Kothari'20, Diakonikolas - Hopkins - Kane - Karmalkar '20]

- Implies all known* notions of parameter seperation

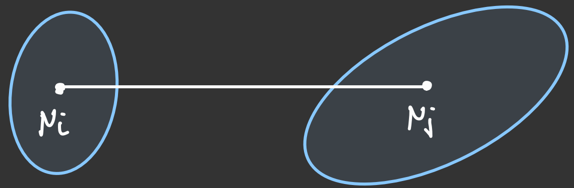
- Generalizes clustering spherical mixtures

Setup :

Definition: M is "clustrable" if $\forall\, i,j \in [k]$
$D(\mu_i, \Sigma_i)$ and $D(\mu_j, \Sigma_j)$ are either

1. Mean-seperated : $\exists v$ s.t.

$$\langle \mu_i - \mu_j, v \rangle^2 \gg v^T(\Sigma_i + \Sigma_j)v .$$



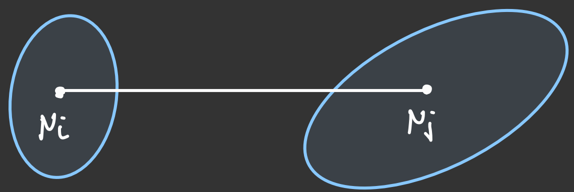
Setup :

Definition: M is "clusterable" if $\forall\, i,j \in [k]$

$D(\mu_i, \Sigma_i)$ and $D(\mu_j, \Sigma_j)$ are either

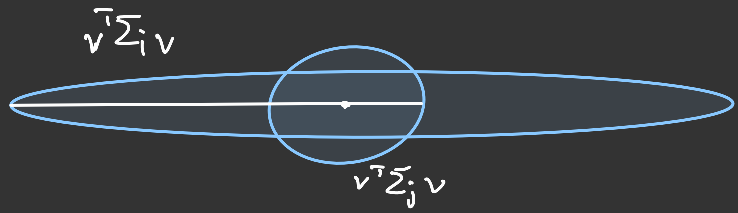
① Mean-seperated : $\exists v$ s.t.

$$\langle \mu_i - \mu_j, v \rangle^2 \gg v^T (\Sigma_i + \Sigma_j) v .$$



② Spectrally-seperated : $\exists v$ s.t.

$$v^T \Sigma_1 v \gg v^T \Sigma_2 v .$$

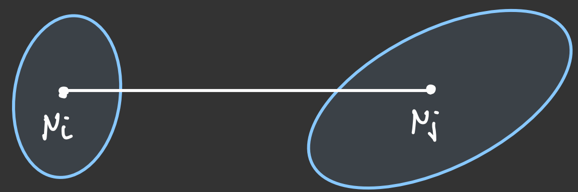


Setup :

Definition: M is "clustrable" if $\forall\, i,j \in [k]$
$D(\mu_i, \Sigma_i)$ and $D(\mu_j, \Sigma_j)$ are either

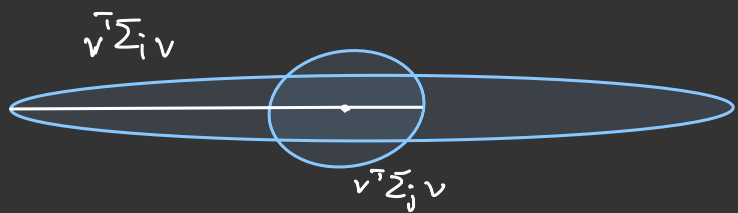
① Mean - seperated : $\exists v$ s.t.

$$\langle \mu_i - \mu_j, v \rangle^2 \gg v^T (\Sigma_i + \Sigma_j) v .$$



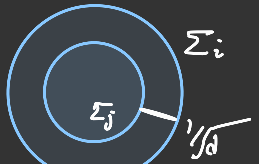
② Spectrally - seperated : $\exists v$ s.t.

$$v^T \Sigma_1 v \gg v^T \Sigma_2 v .$$



③ Shell / Frobenius - seperated :

$$\| \Sigma_i^{-1/2} \Sigma_j \Sigma_i^{-1/2} - I \|_F^2 \gg \| \Sigma_i^{-1/2} \cdot \Sigma_j^{1/2} \|_{op}^2$$



Recap : Gaussian Mixtures

① Hyper-contractivity of linear forms $\Rightarrow$ Mean-Seperated Clusting

[Kothari-Steinhardt'18, Hopkins-Li'18]

② Hyper-contractivity of deg-2 polynomials $\Rightarrow$ Shell / Frobenius-Sep Clustering

[B.-Kothari'20, Diakonikolas-Hopkins-Kane-Karmalkar'20]

③ Anti-Concentration $\Rightarrow$ Spectral-Sep Clustering

[B.-Kothari'20, Diakonikolas-Hopkins-Kane-Karmalkar'20]

Recap : Gaussian Mixtures

① Hyper-contractivity of linear forms* $\Rightarrow$ Mean-Seperated Clusting

[Kothari-Steinhardt'18, Hopkins-Li'18]

② Hyper-contractivity of deg-2 polynomials* $\Rightarrow$ Shell / Frobenius-Sep Clustering

[B.-Kothari'20, Diakonikolas-Hopkins-Kane-Karmalkar'20]

③ Anti-Concentration* $\Rightarrow$ Spectral-Sep Clustering

[B.-Kothari'20, Diakonikolas-Hopkins-Kane-Karmalkar'20]

* SoS certifiable versions of these conditions imply efficient algorithms

Beyond Gaussians:

Claim: Any log-concave distribution satisfies:

1. Hyper-contractivity of linear forms
2. Hyper-contractivity of deg-2 polynomials
3. Anti-Concentration

Beyond Gaussians:

Claim: Any log-concave distribution satisfies:

1. Hyper-contractivity of linear forms
2. Hyper-contractivity of deg-2 polynomials
3. Anti-Concentration

Proof Sketch:

$\forall\, v \in \mathbb{R}^d$, the PDF is point-wise bounded by $\exp(-\langle x, v\rangle / c)$.

How do we get efficient algorithms?

Analytic Certificates

Def [Certifiable Hypercontractivity of linear forms]:

$$\forall v \in \mathbb{R}^d, \quad C_k \cdot \Big(\underbrace{\mathbb{E}\langle x, v\rangle^2}_{\text{variance}}\Big)^k - \mathbb{E}\underbrace{\langle x, v\rangle}_{\text{linear form}}^{2k} = \text{sos}(v).$$

Analytic Certificates

Def [Certifiable Hypercontractivity of linear forms]:

$$\forall v \in \mathbb{R}^d, \qquad c_k \cdot \Big(\underbrace{\mathbb{E}\langle x, v\rangle^2}_{\text{variance}}\Big)^k - \mathbb{E}\underbrace{\langle x, v\rangle}_{\text{linear form}}^{2k} = \operatorname{sos}(v).$$

↪ apriori, a statement with infinite constraints

↪ admits a polynomial size representation

Analytic Certificates imply Efficient Algorithms

Mean Seperation

Shell seperation

Spectral Seperation

Clustering

Analytic Certificates imply Efficient Algorithms

Certifiable Hypercontractivity of linear forms $\longrightarrow$ Mean Seperation

Certifiable Hypercontractivity of deg 2 polynomials $\longrightarrow$ Shell seperation

Certifiable Anti-concentration $\longrightarrow$ Spectral Seperation

$\Bigg\}$ Clustering

Analytic Certificates imply Efficient Algorithms

Poincaré Inequality $\xrightarrow{\text{Kothari-Steinhardt}}$ Certifiable Hypercontractivity of linear forms $\longrightarrow$ Mean Seperation

Poincaré Inequality $\xrightarrow{\text{no SoS proof}}$ Certifiable Hypercontractivity of deg 2 polynomials $\longrightarrow$ Shell seperation

Certifiable Anti-concentration $\longrightarrow$ Spectral Seperation

Clustering

Analytic Certificates imply Efficient Algorithms

KLS Conjecture

↓

Poincaré Inequality $\xrightarrow{\text{Kothari - Steinhardt}}$ Certifiable Hypercontractivity of linear forms ⟶ Mean Seperation

Poincaré Inequality $\xrightarrow{\text{no sos proof}}$ Certifiable Hypercontractivity of deg 2 polynomials ⟶ Shell seperation

Certifiable Anti - concentration ⟶ Spectral Seperation

Mean Seperation, Shell seperation, Spectral Seperation } Clustering

Analytic Certificates imply Efficient Algorithms

KLS Conjecture

↓

Poincaré Inequality $\xrightarrow{\text{Kothari - Steinhardt}}$ Certifiable Hypercontractivity of linear forms → Mean Seperation

Poincaré Inequality ⇢ (no SoS proof) Certifiable Hypercontractivity of deg 2 polynomials → Shell seperation

Rotational Invariance → Certifiable Anti-concentration → Spectral Seperation

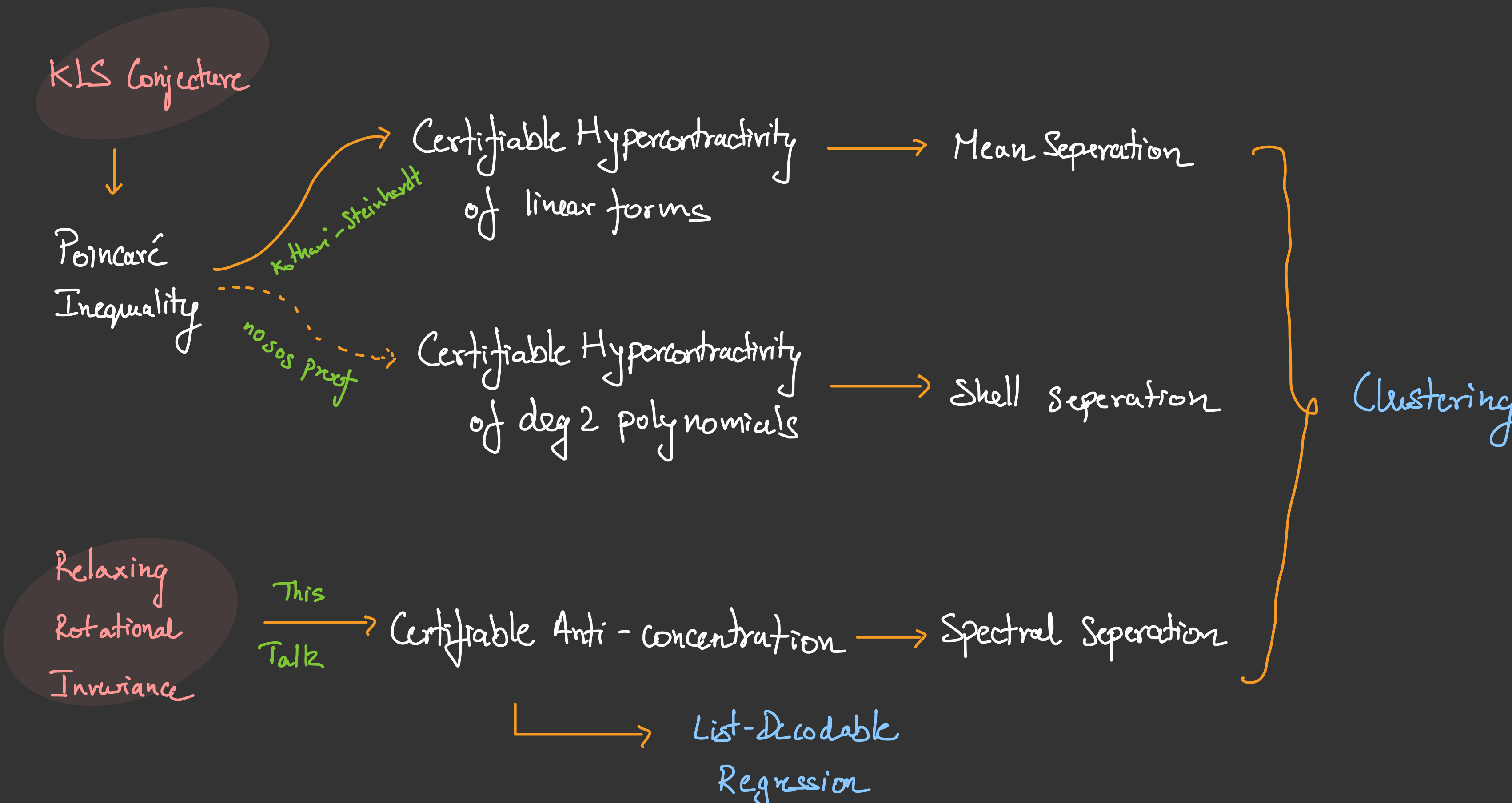
Raghavendra - Yau

Karmalkar - Klivans - Kothari

B. - Kothari

Clustering

Analytic Certificates imply Efficient Algorithms



Certifiable Anti-Concentration: Beyond Gaussians

joint w/ Pravesh Kothari, Goutham Rajendran, Madhur Tulsiani & Aravindan Vijayaraghavan

Certifiable Anti-Concentration: Beyond Gaussians

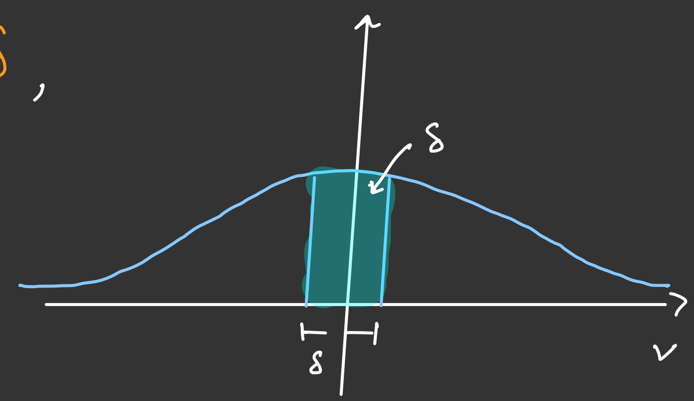
joint w/ Pravesh Kothari, Goutham Rajendran, Madhur Tulsiani & Aravindan Vijayaraghavan

Def [Anti-concentration]: Given a distribution D over $\mathbb{R}^d$, $\forall$ directions v

and all intervals I of length $\delta\sqrt{v^T \Sigma v}$, if

$$\Pr_{x \sim D}\left[\langle x, v\rangle \in I \right] \leq \delta,$$

then D is δ-anti-concentrated.

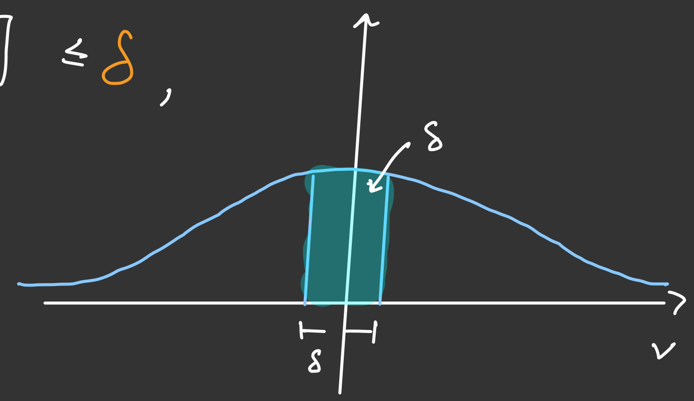


Certifiable Anti-Concentration: Beyond Gaussians

Def [Anti-concentration]: Given n iid samples $\{x_i\}_{i\in[n]}$ from a distribution D, for all directions v in $\mathbb{R}^d$, if

$$\Pr_{x_i \sim \{x_i\}_{i\in[n]}} \left[\langle x_i, v\rangle^2 \leq \delta \cdot v^T \Sigma v \right] \leq \delta ,$$

then D is δ-anti-concentrated.



Can we formulate this as an integer program?

Certifiable Anti-Concentration: Integer Program

Given $\{x_i\}_{i\in[n]}$,

Def [Anti-concentration]: Given samples $\{x_i\}_{i\in[n]}$ from a distribution D, D is δ-anti-concentrated if for all directions v in $\mathbb{R}^d$,

$$\Pr\left[\,|\langle x_i, v\rangle| \le \delta \cdot v^T \Sigma v\,\right] \le \delta$$

$$\max_{v\in\mathbb{R}^d} \quad \frac{1}{n}\sum_{i\in[n]} \mathbb{1}\left[\,\langle x_i, v\rangle^2 \le \delta\, v^T \Sigma v\,\right]$$

Certifiable Anti-Concentration: Integer Program

Given $\{x_i\}_{i\in[n]}$,

Def [Anti-concentration]: Given samples $\{x_i\}_{i\in[n]}$ from a distibution D, D is δ-anti-concentrated if for all directions v in $\mathbb{R}^d$,

$$\Pr\left[|\langle x_i, v\rangle| \le \delta \cdot v^T \Sigma v \right] \le \delta$$

$$\max_{v\in\mathbb{R}^d} \frac{1}{n}\sum_{i\in[n]} \mathbb{1}\left[\langle x_i, v\rangle^2 \le \delta v^T \Sigma v \right]$$

Claim: If $OPT \le \delta$, then the uniform distribution over $\{x_i\}_{i\in[n]}$ is δ-anticoncentrated.

Certifiable Anti-Concentration: Integer Program

Def [Anti-concentration]: Given samples $\{x_i\}_{i \in [n]}$ from a distribution D, D is δ-anti-concentrated if for all directions v in $\mathbb{R}^d$,

$$\Pr\left[|\langle x_i, v \rangle| \leq \delta \cdot v^T \Sigma v \right] \leq \delta$$

Given $\{x_i\}_{i \in [n]}$,

$$\max_{v, w} \quad \frac{1}{n} \sum_{i \in [n]} w_i \quad \text{s.t.}$$

$$\forall i \in [n] \qquad w_i^2 = w_i \qquad \text{[Indicator variables]}$$

$$\forall i \in [n] \qquad w_i \langle x_i, v \rangle^2 \leq w_i \delta \cdot v^T \Sigma v \qquad \text{[Counting concentrated points]}$$

Certifiable Anti-Concentration: Integer Program

Def [Anti-concentration]: Given samples $\{x_i\}_{i\in[n]}$ from a distribution D, D is δ-anti-concentrated if for all directions v in $\mathbb{R}^d$,

$$\Pr\left[|\langle x_i, v\rangle| \leq \delta \cdot v^T \Sigma v \right] \leq \delta$$

Given $\{x_i\}_{i\in[n]}$,

$$\max_{v,w} \quad \frac{1}{n}\sum_{i\in[n]} w_i \qquad \text{s.t.}$$

$$\forall i\in[n] \qquad w_i^2 = w_i \qquad \text{[Indicator variables]}$$

$$\forall i\in[n] \qquad w_i \langle x_i, v\rangle^2 \leq w_i \delta \cdot v^T \Sigma v \qquad \text{[Counting concentrated points]}$$

How do we certify upper bounds on the objective value of this program?

Certifiable Anti-Concentration: Efficient Certificates

$$\mathcal{A} := \left\{ \begin{array}{ll} \forall i \in [n] & w_i^2 = w_i \\ \forall i \in [n] & w_i \langle x_i, v \rangle^2 \leq \delta \cdot v^\top \Sigma v \end{array} \right\}$$

Key Idea: Derive an upper bound on the objective in the sum-of-squares proof system

$\Rightarrow$ Efficiently representable certificate

Certifiable Anti-Concentration: Efficient Certificates

$$\mathcal{A} := \left\{ \begin{array}{ll} \forall i \in [n] & w_i^2 = w_i \\ \forall i \in [n] & w_i \langle x_i, v\rangle^2 \leq \delta \cdot v^\top \Sigma v \end{array} \right\}$$

Key Idea: Derive an upper bound on the objective in the sum-of-squares proof system

$\Rightarrow$ Efficiently representable certificate

Claim:

$$\mathcal{A} \vdash^{v,w} \left\{ \delta - \frac{1}{n} \sum_{i \in [n]} w_i \geq 0 \right\}$$

Certifiable Anti-Concentration: Efficient Certificates

$$\mathcal{A} := \left\{ \begin{array}{ll} \forall i \in [n] & w_i^2 = w_i \\ \forall i \in [n] & w_i \langle x_i, v \rangle^2 \leq \delta \cdot v^\top \Sigma v \end{array} \right\}$$

Key Idea: Derive an upper bound on the objective in the sum-of-squares proof system

$\Rightarrow$ Efficiently representable certificate

Claim:

$$\mathcal{A} \vdash^{v,w} \left\{ \delta - \frac{1}{n} \sum_{i \in [n]} w_i \geq 0 \right\}$$

OR

$$\delta - \frac{1}{n} \sum_{i \in [n]} w_i = \underbrace{\mathrm{sos}(v, w)}_{\text{non-negative}} + \underbrace{\sum_{i \in [n]} q_i^2 \, w_i \left(\delta v^\top \Sigma v - \langle x_i, v \rangle^2 \right)}_{\text{non-negative whenever constraints are satisfied}}$$

Certifiable Anti-Concentration: Efficient Certificates

$$\mathcal{A} := \left\{ \begin{array}{ll} \forall i \in [n] & w_i^2 = w_i \\ \forall i \in [n] & w_i \langle x_i, v\rangle^2 \leq \delta \cdot v^T \Sigma v \end{array} \right\}$$

Main Theorem: Given n samples from a 'reasonably anti-concentrated' distribution there is a degree-$O_\delta(\log d)$ certificate of anti-concentration, i.e.

$$\mathcal{A} \vdash_{O_\delta(\log d)} \left\{ \delta - \frac{1}{n} \sum_{i \in [n]} w_i \geq 0 \right\}.$$

Certifiable Anti-Concentration: Efficient Certificates

$$\mathcal{A} := \left\{ \begin{array}{ll} \forall i \in [n] & w_i^2 = w_i \\ \forall i \in [n] & w_i \langle x_i, v\rangle^2 \leq \delta \cdot v^\top \Sigma v \end{array} \right\}$$

Main Theorem: Given n samples from a "reasonably anti-concentrated" distribution there is a degree-$O_\delta(\log d)$ certificate of anti-concentration, i.e.

$$\mathcal{A} \vdash_{O_\delta(\log d)} \left\{ \delta - \frac{1}{n} \sum_{i \in [n]} w_i \geq 0 \right\}.$$

1. Running Time: $n^{O_\delta(\log d)}$ (Quasi-polynomial)
2. δ-Dependence: $\exp(1/\delta)^{\exp(1/\delta)}$ (Doubly-exponential)
3. No direct sum-of-squares proof

Certifiable Anti-Concentration: Efficient Certificates

$$\mathcal{A} := \left\{ \begin{array}{ll} \forall i \in [n] & w_i^2 = w_i \\ \forall i \in [n] & w_i \langle x_i, v\rangle^2 \leq \delta \cdot v^T \Sigma v \end{array} \right\}$$

Main Theorem: Given n samples from a "reasonably anti-concentrated" distribution there is a degree-$O_\delta(\log d)$ certificate of anti-concentration, i.e.

$$\mathcal{A} \vdash_{O_\delta(\log d)} \left\{ \delta - \frac{1}{n} \sum_{i \in [n]} w_i \geq 0 \right\}.$$

④ Certificate works for affine transformations of uniform distributions over L_p balls, anti-concentrated product distributions etc.

⑤ All prior certificates required rotational invariance.

Certifiable Anti-Concentration: Efficient Certificates

Main Theorem: Given n samples from a "reasonably anti-concentrated" distribution there is a degree-$O_\delta(\log d)$ certificate of anti-concentration, i.e.

Applications:

1. **Clustering:** A $n^{O_\delta(\log d)}$ time (robust) algorithm for clustering spectrally-seperated components.

   $\hookrightarrow$ A $n^{O_{k,t}(\log^2 d)}$ time (robust) algorithm for clustering mixtures.

Certifiable Anti-Concentration: Efficient Certificates

Main Theorem: Given n samples from a "reasonably anti-concentrated" distribution there is a degree-$O_\delta(\log d)$ certificate of anti-concentration, i.e.

Applications:

1. **Clustering:** A $n^{O_\delta(\log d)}$ time (robust) algorithm for clustering spectrally-seperated components.
   $\hookrightarrow$ A $n^{O_{k,t}(\log^2 d)}$ time (robust) algorithm for clustering mixtures.

2. **List-Decodable Regression:** A $n^{O_{\varepsilon,\alpha}(\log d)}$ time algorithm for outputting a list of size $O(1/\alpha)$ s.t. $\|\hat{\Theta} - \Theta\|_2 \leq \varepsilon$.

Certifiable Anti-Concentration: Overview

Starting Point: For uniform distributions over L_p balls, the marginals along random directions are Gaussian-like.

Certifiable Anti-Concentration: Overview

Starting Point: For uniform distributions over L_p balls, the marginals along random directions are Gaussian-like.

Lemma: $\forall \; \|v\|_4^4 \leq \lambda \cdot \|v\|_2^4$ } analytic density

then $$\underbrace{\mathbb{E}_{x\sim D} \langle x, v\rangle^{2k}}_{\text{2k-th moment along } v} = \underbrace{\frac{(2k)!}{2^k k!} \|v\|_2^{2k}}_{\text{Gaussian 2k-th moment}} \pm \underbrace{\lambda \cdot k^k \|v\|_2^{2k}}_{\text{dim-independent deviation}}$$

Certifiable Anti-Concentration: Overview

Starting Point: For uniform distributions over L_p balls, the marginals along random directions are Gaussian-like.

Lemma: If $\|v\|_4^4 \le \lambda \cdot \|v\|_2^4$ } analytic density

then

$$\underbrace{\mathop{\mathbb{E}}_{x\sim D} \langle x, v\rangle^{2k}}_{\text{2k-th moment along } v} = \underbrace{\frac{(2k)!}{2^k k!} \|v\|_2^{2k}}_{\text{Gaussian 2k-th moment}} \pm \underbrace{\lambda \cdot k^k \|v\|_2^{2k}}_{\text{dim-independent deviation}}$$

↳ We provide an explicit sum-of-squares proof of this (in the indeterminate v).

Certifiable Anti-Concentration: Overview

Starting Point: For uniform distributions over L_p balls, the marginals along random directions are Gaussian-like.

Lemma: If $\|v\|_4^4 \leq \lambda \cdot \|v\|_2^4$

then $$\mathop{\mathbb{E}}_{x \sim D} \langle x, v \rangle^{2k} = \frac{(2k)!}{2^k k!} \|v\|_2^{2k} \pm \lambda \cdot k^k \|v\|_2^{2k}$$

How do we handle the remaining directions?

↳ Do not admit a small cover ex: $v = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2\sqrt{d}} \cdots \frac{1}{2\sqrt{d}} \right)$

Certifiable Anti-Concentration: Overview

Lemma: If $\|v\|_4^4 \leq \lambda \cdot \|v\|_2^4$

then $$\mathop{\mathbb{E}}_{x \sim D} \langle x, v \rangle^{2k} = \frac{(2k)!}{2^k k!} \|v\|_2^{2k} \pm \lambda \cdot k^k \|v\|_2^{2k}$$

How do we handle the remaining directions?

Lemma: Any direction v can be decomposed in $v_L + v_H$ s.t. v_L has $1/\lambda$ non-zero coordinates and v_H is analytically dense i.e.

$$\|v_H\|_4^4 \leq \lambda \|v_H\|_2^4 .$$

Does not seem to help if v is an indeterminate.

Certifiable Anti-Concentration: Overview

Lemma: Any direction v can be decomposed in $v_L + v_H$ s.t. v_L has $1/\lambda$ non-zero coordinates and v_H is analytically dense i.e.

$$\|v_H\|_4^4 \leq \lambda \|v_H\|_2^4 .$$

① Switch to dual:

Fact: If $\forall$ deg t pseudo-distributions μ $\tilde{\mathbb{E}}_\mu\, p(x) \geq 0$, $\exists$ a SoS proof of $p(x) \geq 0$.

Certifiable Anti-Concentration: Overview

Lemma: Any direction v can be decomposed in $v_L + v_H$ s.t. v_L has $1/\lambda$ non-zero coordinates and v_H is analytically dense i.e.

$$\|v_H\|_4^4 \leq \lambda \|v_H\|_2^4 .$$

① Switch to dual:

Fact: If $\forall$ deg t pseudo-distributions μ $\tilde{\mathbb{E}}_{\mu}\, p(x) \geq 0$, $\exists$ a sos proof of $p(x) \geq 0$.

② Assume for contradiction there is such a pseudo-distribution

③ Condition on "large" coordinates of v by "re-weighting" the pseudo-distribution

Certifiable Anti-Concentration: Overview

Lemma: Any direction v can be decomposed in $v_L + v_H$ s.t. v_L has $1/\lambda$ non-zero coordinates and v_H is analytically dense i.e.

$$\|v_H\|_4^4 \leq \lambda \|v_H\|_2^4 .$$

① Switch to dual:

Fact: If $\forall$ deg t pseudo-distributions μ $\tilde{\mathbb{E}}_\mu\, p(x) \geq 0$, $\exists$ a SoS proof of $p(x) \geq 0$.

② Assume for contradiction there is no such pseudo-distribution

③ Condition on "large" coordinates of v by "re-weighting" the pseudo-distribution

④ The resulting vector is analytically dense and we can invoke the explicit proof for analytically dense directions

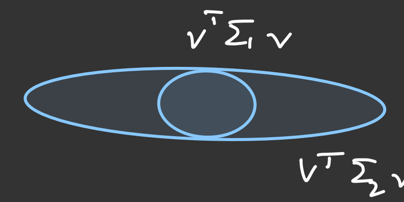
Certifiable Anti-Concentration: Toy Application

$\{x_i\}_{i\in[n]}$ samples from $M = \frac{1}{2} D(0, \Sigma_1) + \frac{1}{2} D(0, \Sigma_2)$ s.t.

1. $\forall v \quad v^T \Sigma_1 v \le v^T \Sigma_2 v$

2. $\exists v$ s.t. $v^T \Sigma_1 v < \delta^2 v^T \Sigma_2 v$

(w.l.o.g. assume M is isotropic)

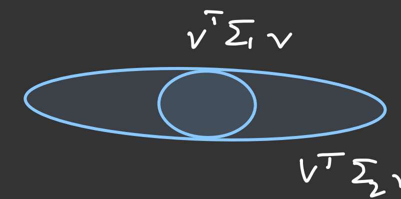


Certifiable Anti-Concentration: Toy Application

$\{x_i\}_{i\in[n]}$ samples from $M = \frac{1}{2}D(0,\Sigma_1) + \frac{1}{2}D(0,\Sigma_2)$ s.t.

1. $\forall v \quad v^T\Sigma_1 v \leq v^T \Sigma_2 v$

2. $\exists v \quad$ s.t. $\quad v^T\Sigma_1 v < \delta^2 v^T\Sigma_2 v$



(w.l.o.g. assume M is isotropic)

$$A := \begin{cases} \forall i\in[n] & w_i^2 = w_i \\ \forall i \in [n] & w_i\langle x_i, v\rangle^2 \leq w_i\,\delta^2\|v\|^2 \\ & \sum w_i = n/2 \end{cases}$$

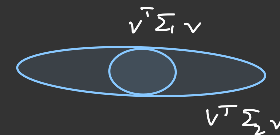
Certifiable Anti-Concentration: Toy Application

$\{x_i\}_{i\in[n]}$ samples from $M = \frac{1}{2}D(0,\Sigma_1) + \frac{1}{2}D(0,\Sigma_2)$ s.t.

1. $\forall v \quad v^T\Sigma_1 v \leq v^T\Sigma_2 v$

2. $\exists\, v$ s.t. $v^T\Sigma_1 v < \delta^2\, v^T\Sigma_2 v$

(w.l.o.g. assume M is isotropic)



$$\mathcal{A} := \begin{cases} \forall i \in [n] & w_i^2 = w_i \\ \forall i \in [n] & w_i \langle x_i, v\rangle^2 \leq w_i\, \delta\, \|v\|^2 \\ & \sum w_i = n/2 \end{cases}$$

Algorithm:

1. Compute a deg $O_\delta(\log d)$ pseudo-distribution
2. Sample $g \sim N(0, \tilde{\mathbb{E}}\, vv^T)$
3. Output $g/\|g\|$.

Certifiable Anti-Concentration: Toy Application

$\{x_i\}_{i\in[n]}$ samples from $M = \frac{1}{2}D(0,\Sigma_1) + \frac{1}{2}D(0,\Sigma_2)$ s.t.

1. $\forall v \quad v^T\Sigma_1 v \leq v^T\Sigma_2 v$
2. $\exists v \quad$ s.t. $\quad v^T\Sigma_1 v < \delta^2 v^T\Sigma_2 v$

$$A := \begin{cases} \forall i\in[n] & w_i^2 = w_i \\ \forall i\in[n] & w_i\langle x_i, v\rangle^2 \leq w_i\,\delta\,\|v\|^2 \\ & \sum w_i = n/2 \end{cases}$$

Key SoS Lemma:

$$A \vdash \left\{ \langle \Sigma_1, vv^T\rangle^2 \leq O(\delta)\|v\|^4 \right\}$$

Finishing the Proof:

$$\langle \Sigma_1, \tilde{\mathbb{E}}\, vv^T\rangle^2 \leq \tilde{\mathbb{E}}\langle \Sigma_1, vv^T\rangle^2 \leq O(\delta)$$

Observe $\mathbb{E}\, gg^T = \tilde{\mathbb{E}}\, vv^T$ and the claim follows from Markov's.

Take aways:

KLS Conjecture

Poincaré Inequality

Kothari - Steinhardt

Certifiable Hypercontractivity of linear forms → Mean Seperation

no sos proof

Certifiable Hypercontractivity of deg 2 polynomials → Shell seperation

Uniform over Lp balls

This Talk

Certifiable Anti - concentration → Spectral Seperation

List-Decodable Regression

Clustering