FITTING ELLIPSOIDS TO RANDOM POINTS

Antoine Maillard

ETH zürich

- arXiv:2310.05787 (with A. Bandeira)
- arXiv:2310.01169 (with D. Kunisky)
- arXiv:2307.01181 (with A. Bandeira, S. Mendelson, E. Paquette)

FITTING ELLIPSOIDS TO RANDOM POINTS

$$\boxed{\begin{array}{l} x_1, \cdots, x_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \mathrm{I}_d/d) \\ n, d \to \infty \end{array}} \longrightarrow \mathbb{E}\|x_i\|^2 = 1$$

Ellipsoid Fitting Property (EFP)

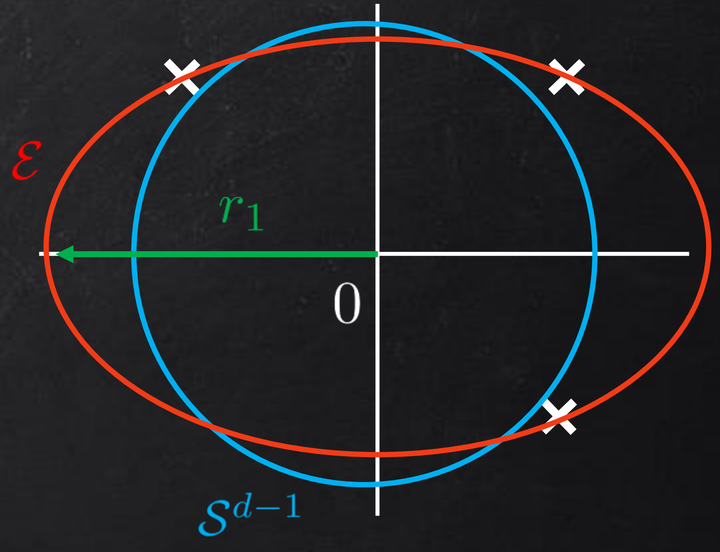
$\boxed{\text{Centered ellipsoid } \mathcal{E} \text{ with } x_1, \cdots, x_n \in \mathcal{E}}$?

⇕

$\boxed{\exists S \in \mathbb{R}^{d \times d} : S \succeq 0 \text{ and } x_i^\top S x_i = 1 \text{ for all } i \in [n]}$?

EFP is a semidefinite program

- Convex
- Efficient algorithms

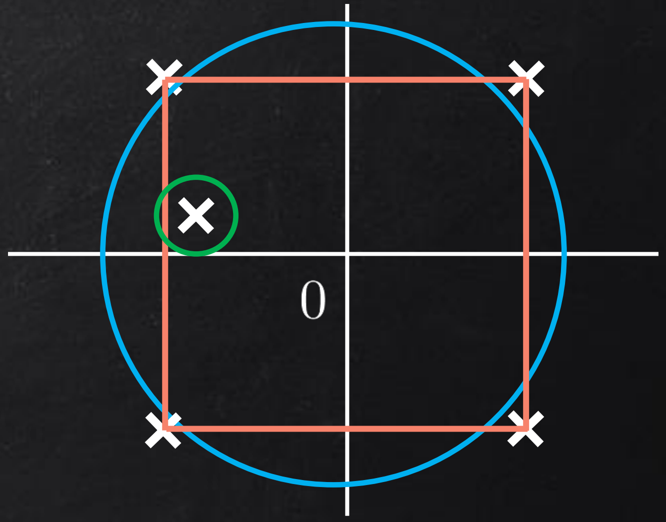


Principal axes of $\mathcal{E}$ ⟺ Eigenspaces of S

$$r_i = \lambda_i(S)^{-1/2}$$

$$p(n, d) \coloneqq \mathbb{P}[(x_1, \cdots, x_n) \text{ satisfies EFP}]$$

A FEW MOTIVATIONS



- **Geometry :** $\mathrm{EFP} \Rightarrow (\pm x_1, \cdots, \pm x_n) \in \mathrm{Bound}(\mathrm{Conv}(\pm x_1, \cdots, \pm x_n))$

- **Statistical estimation**
  - Minimum Trace Factor Analysis [Saunderson &al '12]

$$X = \underset{\text{Diagonal}}{D^\star} + \underset{\succeq 0 \text{ + low-rank}}{L^\star} \in \mathbb{R}^{n\times n}$$

$$\mathrm{MTFA} := \min_{\substack{D,L\,:\,X=D+L \\ L\succeq 0}} \mathrm{Tr}(L)$$

$$\mathrm{col}(L^\star) \sim \mathrm{Unif}[r\text{-dim subspaces}] \Rightarrow \boxed{\mathbb{P}[\mathrm{MTFA\ recovers}\ (L^\star, D^\star)] = p(n, n-r)}$$

$$p(n,d) = \mathbb{P}[x_1, \cdots, x_n \in \mathbb{R}^d \text{ satisfy EFP}]$$

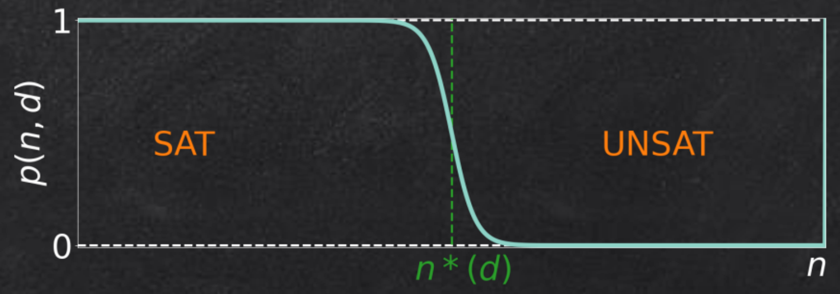
  - Independent Component Analysis [Podosinnikova&al '19]

- **Theoretical computer science**
  - Discrepancy of random matrices
  - Characterization of SDPs in average-case scenarios…

[Potechin &al '22]

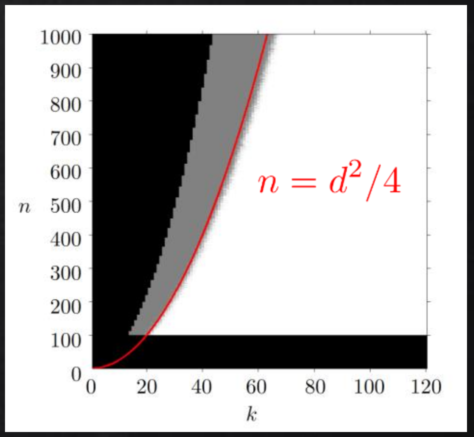
THE ELLIPSOID FITTING CONJECTURE

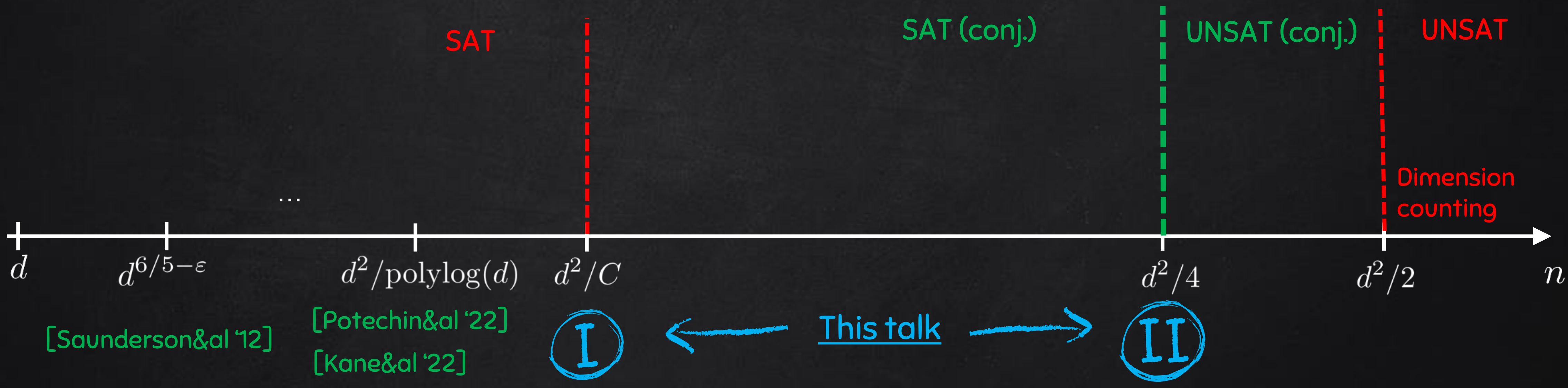
$$p(n,d) = \mathbb{P}[\exists S \succeq 0 \,:\, x_i^\top S x_i = 1 \ (\forall i \in [n])]$$



$\boxed{n^\star(d)}$?

[Saunderson&al '12]





SAT

SAT (conj.)

UNSAT (conj.)

UNSAT

Dimension counting

d

$d^{6/5-\varepsilon}$

…

$d^2/\mathrm{polylog}(d)$

d^2/C

$d^2/4$

$d^2/2$

n

[Saunderson&al '12]

[Potechin&al '22]

[Kane&al '22]

I

This talk

II

- Upper bound $n \gtrsim d^2/2$ by dimension counting.
- EFP is not universal $x_i \overset{\text{i.i.d.}}{\sim} \mathrm{Unif}(\{\pm 1\}^d) \Rightarrow p(n,d) = 1$

$S = \mathrm{I}_d$

Fluctuations of the norm are critical ⚠

LOWER BOUNDS

[Bandeira, M., Mendelson & Paquette '23]

Goal: $p(n,d) \to 1$ for $n < n_c(d)$

Existing works on EFP rely on an explicit estimate:

- $$\hat{S}_{\mathrm{LS}} := \underset{\{x_i^\top S x_i = 1\}}{\arg\min} \|S\|_F$$

[Potechin&al '22]

Theorem: $\hat{S}_{\mathrm{LS}} \succeq 0$ w.h.p. if $n \lesssim d^2/\mathrm{polylog}(d)$

Non-rigorous analysis shows this holds for $n \leq d^2/10$ [M.&Kunisky '22]

- "Identity perturbation"

$$\hat{S}_{\mathrm{IP}} := \mathrm{I}_d + \sum_{i=1}^{n} q_i x_i x_i^\top$$

$\{x_i^\top \hat{S}_{\mathrm{IP}} x_i = 1\}_{i=1}^n$ n linear equations in $q \in \mathbb{R}^n$

Theorem: $\hat{S}_{\mathrm{IP}} \succeq 0$ w.h.p. if

- $n \lesssim d^2/\mathrm{polylog}(d)$ [Kane & Diakonikolas '22]
- $n \leq d^2/C$ [Bandeira, M., Mendelson & Paquette '23]

Numerically: $C \simeq 10$

Similar result obtained w. different estimates in [Hsieh&al'23 ; Tulsiani & Wu '23]

LOWER BOUNDS – SKETCH OF PROOF

[Bandeira, M., Mendelson & Paquette '23]

$\boxed{x_i = \sqrt{d_i}\omega_i}$ $\omega_i \overset{\text{i.i.d.}}{\sim} \mathrm{Unif}(\mathcal{S}^{d-1})$

Define $\begin{cases} D = \mathrm{Diag}(\{d_i\}) \\ \Theta_{ij} := \langle \omega_i, \omega_j \rangle^2 \end{cases}$

$$\hat{S}_{\mathrm{IP}} := \mathrm{I}_d + \underbrace{\sum_{i=1}^n q_i x_i x_i^\top}_{\text{We show } \|\cdot\|_{\mathrm{op}} \leq 1} + \{x_i^\top \hat{S}_{\mathrm{IP}} x_i = 1\}_{i=1}^n \Rightarrow \boxed{q = D^{-1}\Theta^{-1}(D^{-1}\mathbf{1}_n - \mathbf{1}_n)}$$

Goal: $$\left\| \sum_{i=1}^n [\Theta^{-1}\underbrace{(D^{-1}\mathbf{1}_n - \mathbf{1}_n)}_{\text{i.i.d., independent of } \omega_i}]_i \, \omega_i \omega_i^\top \right\|_{\mathrm{op}} \leq 1$$

- <u>Key difficulty</u>: controlling $\|\Theta^{-1}\|_{\mathrm{op}}$. ⚠
- Rest of the proof: classical ε–net argument.

$p = \binom{d+1}{2}$

$\Theta_{ij} = \langle \omega_i \omega_i^\top, \omega_j \omega_j^\top \rangle$ ⇒ **Lemma:** $\|\Theta - \mathbb{E}\Theta\|_{\mathrm{op}} \lesssim \sqrt{\frac{n}{d^2}}$ ⇒ $\|\Theta^{-1}\|_{\mathrm{op}} \leq 2$ for small enough $\frac{n}{d^2}$. ■

Gram matrix of sub-exp.
random vectors in $\mathbb{R}^p$

[Bartl & Mendelson '22]

ELLIPSOID FITTING CONJECTURE REVISITED

$p = \binom{d+1}{2}$

$h_i = x_i x_i^\top$

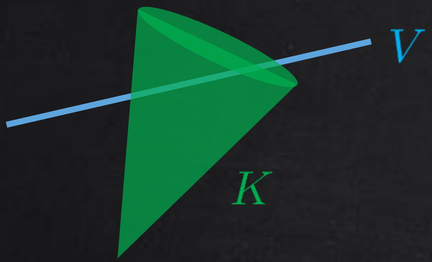
EFP: $\begin{cases} \mathrm{Tr}[S x_i x_i^\top] = 1 \longrightarrow S \in V = \{x \in \mathbb{R}^p \, : \, \langle x, h_i \rangle = 1 \; (\forall i \in [n])\} \\ S \succeq 0 \longrightarrow S \in K \end{cases}$

$S \in V$: random affine subspace in $\mathbb{R}^p$

$S \in K$: closed convex cone

General question

$\boxed{\mathbb{P}[V \cap K \neq \emptyset]}$?



Heuristic

What if the directions of V were uniformly randomly oriented ?

"Gaussian Fitting" (GF) $\begin{cases} x \in V = \{x \in \mathbb{R}^p \, : \, \langle x, g_i \rangle = 1 \; (\forall i \in [m])\} \\ x \in K \end{cases}$

$g_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \mathrm{I}_p)$

Theorem [Gordon '88, ...] GF is
- SAT (whp) if $n \leq (1-\varepsilon)\omega(K)^2$
- UNSAT (whp) if $n \geq (1+\varepsilon)\omega(K)^2$

$$\omega(K) := \mathbb{E} \max_{\substack{x \in K \\ \|x\|=1}} \langle g, x \rangle$$ Gaussian width

$g \sim \mathcal{N}(0, \mathrm{I}_p)$

For PSD matrices $\omega(\mathcal{S}_d^+) \sim \frac{d}{2} \Longrightarrow \boxed{\tilde{n}(d) \simeq \frac{d^2}{4}}$ Heuristic for the ellipsoid fitting conjecture

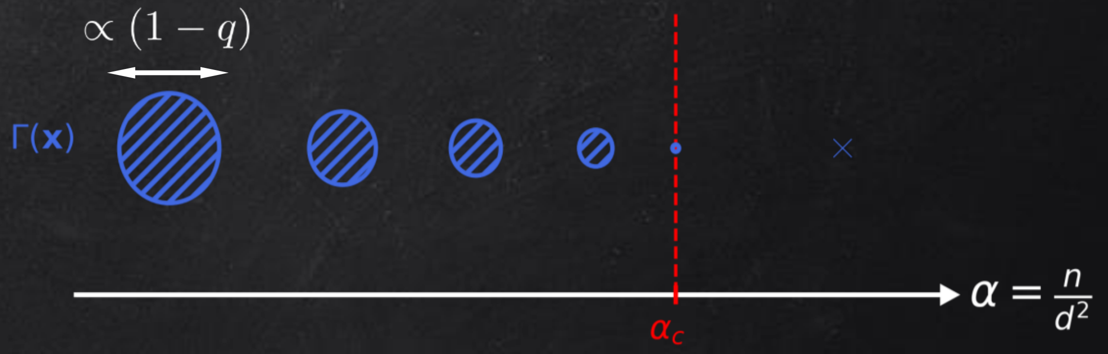
NON-RIGOROUS RESULTS (Statistical physics of disordered systems)

[M. & Kunisky '23]

Free energy / volume of solution set:

$$\Phi = \mathbb{E}\frac{1}{d^2}\log\int P_0(\mathrm{d}S)\prod_{i=1}^{n}\delta(x_i^\top S x_i - 1)$$

$\mathrm{supp}(P_0) \subseteq \mathcal{S}_d^+$



Replica method + convexity
("replica symmetry")

Asymptotic formula

$$\boxed{\Phi \to_{d\to\infty} \sup_{q\in[0,1]}\ \sup_{\mu\in\mathcal{M}_1^+(\mathbb{R})} \left[F(\alpha,q,\mu) + I_{\mathrm{HCIZ}}\left(\frac{1}{\sqrt{1-q}},\mu,\sigma_{\mathrm{s.c.}}\right)\right]}$$

"Overlap"

Typical spectrum of solutions

$$I_{\mathrm{HCIZ}}(\theta,A,B) = \lim_{d\to\infty}\frac{1}{d^2}\log\int_{\mathcal{O}(d)}\mathcal{D}O\,\exp\{\theta\mathrm{Tr}[OAO^\top B]\}$$

Hard asymptotic expressions via PDEs [Matytsin '94 ; Guionnet&al'02]

Dilute expansion ($\theta\to\infty$) of $I_{\mathrm{HCIZ}}(\theta,\mu,\nu)$ [Bun&al '16] ✚

$\alpha\to\alpha_c$

$$\boxed{\begin{array}{c} q\to 1 \\ \Phi\to-\infty \end{array}} \Longrightarrow \boxed{\alpha_c = \frac{1}{4}}$$

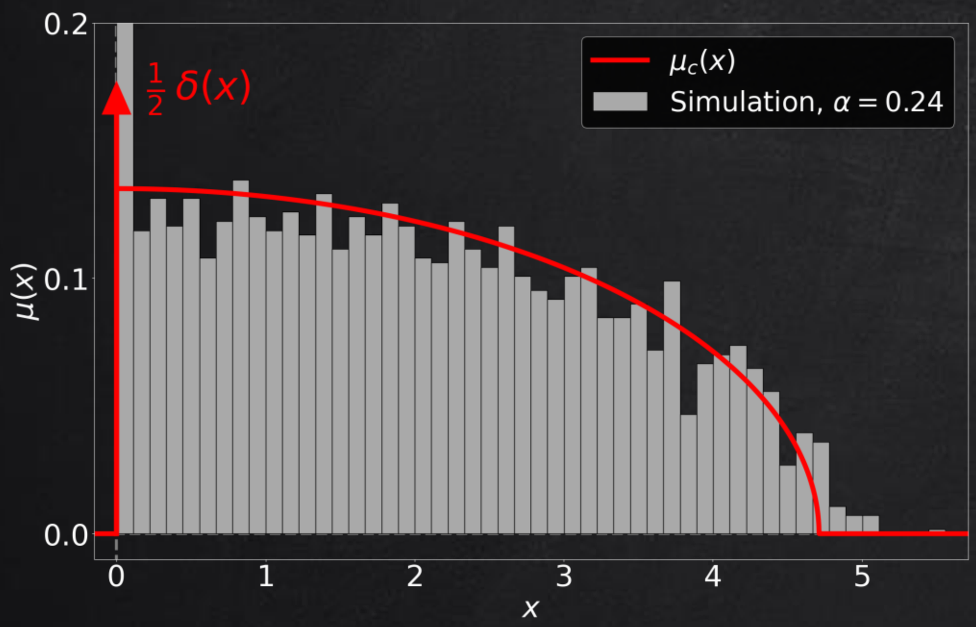
⚠ Replica method hints at universality of Φ with the "Gaussian fitting" problem.

NON-RIGOROUS RESULTS: SOME CONSEQUENCES

$\alpha = n/d^2$

Spectrum of solutions/shape of ellipsoids

Near the transition: $\alpha \uparrow \frac{1}{4}$

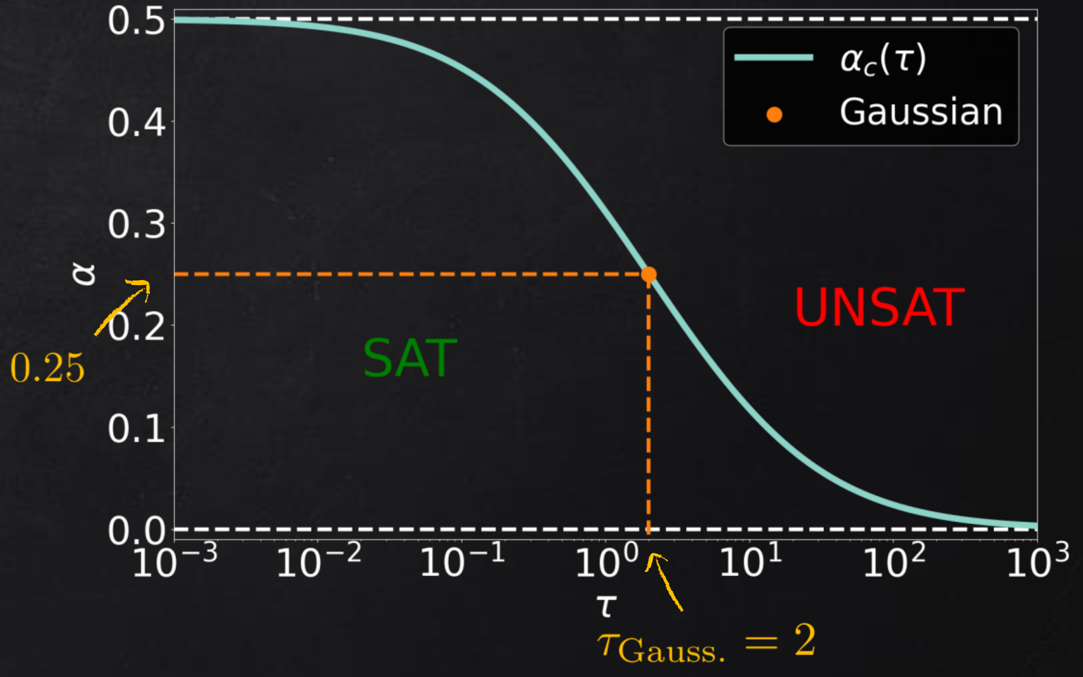


- As $\alpha \uparrow \frac{1}{4}$ ellipsoid fits are "cylinders" in half directions !
- Truncated semicircular distribution

Universality with "Gaussian fitting" problem.

Generalization to non-Gaussian random vectors

$$x_i = \sqrt{r_i}\,\omega_i \quad \begin{cases} \omega_i \sim \mathrm{Unif}(\mathcal{S}^{d-1}) \\ \mathbb{E}[r_i] = 1 \ ; \ \boxed{\mathrm{Var}(r_i) = \frac{\tau}{d}} \end{cases}$$



Larger norm fluctuations → Ellipsoid fits harder to find

A RIGOROUS APPROACH INSPIRED BY PHYSICS

[M. & Bandeira '23]

Regularizing $\delta(\cdot)$; $\beta \gg 1$

"Free energy"

$$\begin{cases} \Phi := \mathbb{E}\frac{1}{d^2} \log \int P_0(\mathrm{d}S) \exp\Big\{-\beta \sum_{i=1}^{n} \ell\Big[\sqrt{d}(x_i^\top S x_i - 1)\Big]\Big\} \\ \Phi_G := \mathbb{E}\frac{1}{d^2} \log \int P_0(\mathrm{d}S) \exp\Big\{-\beta \sum_{i=1}^{n} \ell\Big[\sqrt{d}(\mathrm{Tr}[S Y_i] - 1)\Big]\Big\} \end{cases}$$

⟺ "Gaussian fitting" problem

Gaussian, same order-2 moments as $x_i x_i^\top$

"Gaussian equivalence"

Lemma: $\Phi \simeq \Phi_G$ if "$\sup_S$"$|\mathbb{E}\,\varphi(x^\top S x) - \mathbb{E}\,\varphi(\mathrm{Tr}[SY])| \xrightarrow[d\to\infty]{} 0$

[Hu & Lu '20 ; Montanari & Saeed '22 ; Gerace & al '22, ...]

1. We show this "uniform CLT of projections" using a Berry–Esseen–type CLT → Focus on Φ_G
2. We leverage Gordon's theorem to study Φ_G

Limitation: supremum over S with bounded spectrum

TRANSITION FOR APPROXIMATE EFP

[M. & Bandeira '23]

EFP_ε Find $S \succeq 0$ such that $\underbrace{\frac{1}{n}\sum_{i=1}^{n} \sqrt{d}|x_i^\top S x_i - 1|}_{= \Theta(1) \text{ for } S = \mathrm{I}_d} \leq \varepsilon$

"Relaxed" problem: $\mathrm{EFP} = \mathrm{EFP}_0$

Theorem

- $n/d^2 \to \alpha < 1/4$: $\forall \varepsilon > 0$, we can find $\hat{S}_\varepsilon$ solution to EFP_ε, and $\mathrm{Sp}(\hat{S}_\varepsilon) \subseteq [\lambda_-(\alpha), \lambda_+(\alpha)] \subseteq (0, \infty)$
- $n/d^2 \to \alpha > 1/4$: $\exists \varepsilon(\alpha) > 0$ s.t. $\forall \lambda_+ > 0$, there is no solution S to EFP_ε such that $\mathrm{Sp}(S) \subseteq [0, \lambda_+]$

⚠ Rigorous characterization of the SAT/UNSAT transition in (approximate) ellipsoid fitting at $n \simeq \frac{d^2}{4}$

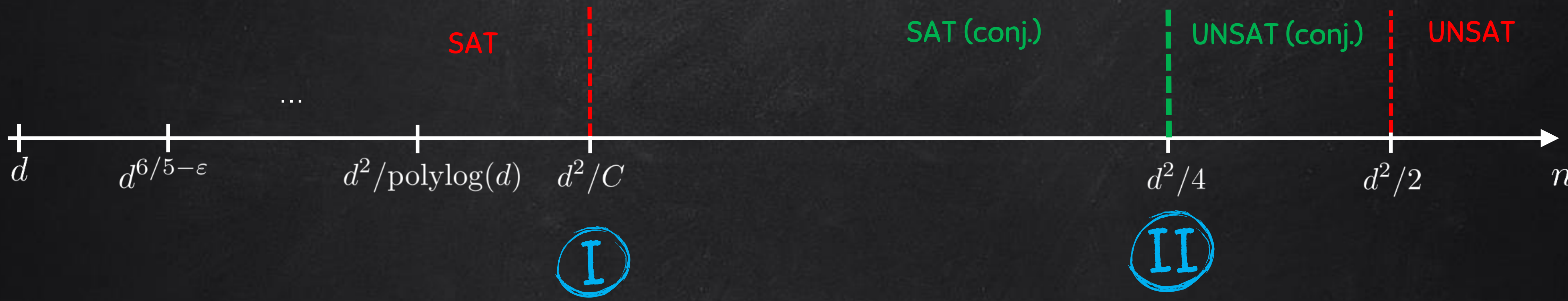
$\alpha < 1/4$

- Approximate solutions, up to arbitrary accuracy
- We control the spectrum of solutions in the SAT phase
  (shape of ellipsoid fits)

$\alpha > 1/4$

- Rule out solutions with bounded spectrum
  (ellipsoid with axes not too small)

SUMMARY & OUTLOOK



Best-known lower bound + Rigorous transition for approximate ellipsoid fitting

- Other proof approaches for lower bounds ?

$$\boxed{\hat{S}_{\mathrm{NN}} := \operatorname*{arg\,min}_{\{x_i^\top S x_i = 1\}} \|S\|_{\mathrm{NN}} \succeq 0 \text{ for any } \alpha < \frac{1}{4}}$$

?

✓ Numerics
✓ Replica prediction

- Universality proof for non-Gaussian random vectors ?

From approximate to exact ellipsoid fit ?

Rule out matrices with diverging spectral norm ?
(ellipsoids with very small axes)

THANK YOU !