Randomstrasse101: Open Problems of 2024

Afonso S. Bandeira[*]   Anastasia Kireeva[§]   Antoine Maillard[♯]   Almut Rödder[¶]

May 29, 2025

Abstract

`Randomstrasse101` is a blog dedicated to Open Problems in Mathematics, with a focus on Probability Theory, Computation, Combinatorics, Statistics, and related topics. This manuscript serves as a stable record of the Open Problems posted in 2024, with the goal of easing academic referencing. The blog can currently be accessed at `randomstrasse101.math.ethz.ch`.

Introduction

In this manuscript we include the blog entries in `Randomstrasse101` of 2024, containing a total of fifteen Open Problems. `Randomstrasse101` is a blog created and maintained by our group at the Department of Mathematics at ETH Zürich[1].

The focus is on mathematical open problems in Probability Theory, Computation, Combinatorics, Statistics, and related topics. The goal is not to necessarily write about the most important open questions in these fields but simply to discuss open questions and conjectures[2] that each of us find particularly interesting or intriguing, somewhat in the same style as the first author's –now almost a decade old– list of forty-two open problems [Ban16].

The blog was created and is currently maintained by Afonso S. Bandeira, Daniil Dmitriev, Anastasia Kireeva, Antoine Maillard, Chiara Meroni, Petar Nizic-Nikolac, Kevin Lucca, and Almut Rödder at ETH Zürich. Each blog entry is generally written by one author and the author list of this manuscript is the union of the entry authors in it.

Given the nature of this material, the writing of this manuscript is more informal than a typical academic text[3]. Nevertheless, we hope it is useful and inspires thought on these questions. Happy solving!

Contents

---

[*]ASB: Department of Mathematics, ETH Zürich, Rämistrasse 101, 8092 Zurich, Switzerland. `bandeira@math.ethz.ch`

[§]AK: Department of Mathematics, ETH Zürich, Rämistrasse 101, 8092 Zurich, Switzerland. `anastasia.kireeva@math.ethz.ch`

[♯]AM: ARGO Team, Inria Paris, 48 Rue Barrault, 75013 Paris, France. `antoine.maillard@inria.fr`. Blog entry written while at the Department of Mathematics, ETH Zürich, Rämistrasse 101, 8092 Zurich, Switzerland.

[¶]AR: Department of Mathematics, ETH Zürich, Rämistrasse 101, 8092 Zurich, Switzerland. `almutmagdalena.roedder@ifor.math.ethz.ch`

[1]The inspiration for the name of the blog will be clear to the reader after looking up the department's address and recalling the Probability Theory focus of the blog.

[2]Conjectures here should be interpreted as mathematical statements that we do not know not to be true, and for which a proof or a refutal would be interesting progress. We do not necessarily have very high confidence that conjectures here are true.

[3]If you would like to refer to an open question or blog entry, we encourage you to refer to this manuscript instead, since it is a more stable reference.

Entry 1   Welcome & Matrix Discrepancy (ASB)

Let me start with one of my favourite Open Problems.

Conjecture 1 (Matrix Spencer). *There exists a positive universal constant C such that, for all positive integers n, and all choices of n self-adjoint $n \times n$ real matrices $A_1, \ldots, A_n$ satisfying, for all $i \in [n]$, $\|A_i\| \leq 1$ (where $\| \cdot \|$ denotes the spectral norm) the following holds*

$$\min_{\varepsilon \in \{\pm 1\}} \left\| \sum_{i=1}^{n} \varepsilon_i A_i \right\| \leq C\sqrt{n}. \tag{1}$$

I first learned about this question from Nikhil Srivastava at ICM 2014, who pointed me to this very nice blog post of Meka [Mek14]. To the best of my knowledge the question first appeared in a paper of Zouzias [Zou12]. I have also written about it in [Ban16] and more recently in [Ban24].

The non-commutative Khintchine inequality (or a matrix concentration inequality) gives the bound up to a $\sqrt{\log n}$ factor.

In the particular case of commutative matrices, one can simultaneously diagonalize the matrices and the question reduces to a vector problem (representing the diagonal of the respective matrices) where the spectral norm is replaced by the ℓ_∞ norm. This is precisely the setting of Joel Spencer's celebrated "Six Standard Deviations Suffice" theorem [Spe85b] which establishes the conjecture in that setting, for $C = 6$. It is noteworthy that in the commutative setting a random choice does not give (1), the argument is of entropic nature. In a sense, in the other extreme situation in which the matrices behave "very non-commutatively" (or, more precisely, "freely") we know that a random choice of signs works, from our (myself, Boedihardjo, and van Handel) recent work on matrix concentration [BBvH23]. The fact that the reason for the conjecture to be true in these two extreme situations appears to be so different (one based on entropy the other on concentration), together with the fact that we still do not know if it is true in general makes this (in my opinion) a fascinating question!

In recent work, Hopkins, Raghavendra, and Shetty [HRS22] has established connections between this problem and Quantum Information, and Dadush, Jiang, and Reis [DJR22] establishes a connection with covering estimates for a certain notion of relative entropy.

More recently, Bansal, Jiang, and Meka [BJM23] has proved Conjecture 1 for low-rank matrices (whose rank is n over a polylog factor) using the matrix concentration inequalities in [BBvH23] within a very nice decomposition argument.

Motivated by Conjecture 1 and the suspicion that "ammount of commutativity" may play an important role in this question, we (myself, Kunisky, Mixon, and Zeng) proposed a Group theoretic special case of this conjecture in [BKMZ22].

Conjecture 2 (Group Spencer). *Let G be a finite group of size n. Conjecture 1 holds in the particular case in which $A_1, \ldots, A_n$ are the $n \times n$ matrices corresponding to the regular representation of G.*

In [BKMZ22] we prove Conjecture 2 for simple groups, but the general case is still open. While the commutative case follows from Spencer's theorem, giving an explicit construction is not trivial (see this Mathoverflow post: https://mathoverflow.net/q/441860).

Entry 2   Global Synchronization (ASB)

In the 17th century, Christiaan Huygens (inventor of the pendulum clock) observed that pendulum clocks have a tendency to spontaneously synchronize when hung on the same board. This phenomenon of spontaneous synchronization of coupled oscillators has since become a central object of study in dynamical systems.

We will focus here on spontaneous synchronization of n coupled oscillators with pairwise connections given by a graph with adjacency matrix $A \in \mathbb{R}^{n \times n}$. The celebrated Kuramoto model [Kur75] models the oscillators as gradient flow on $\mathcal{E}(\theta) = \frac{1}{2} \sum_{i,j=1}^{n} A_{ij} \left(1 - \cos(\theta_i - \theta_j)\right)$, where the parameterization $\theta \in [0, 2\pi[^n$ represents each oscillators angle in the moving frame of its natural frequency (for the sake of the sequel the derivation of this potential is not crucial you can see, e.g. [ABKS23] for more details). This motivates the following definition which is central to what follows.

Definition 2.1. *We say an $n \times n$ matrix A is globally synchronizing if the only local minima of $\mathcal{E} : \mathbb{S}^{n-1} \to \mathbb{R}$, parameterized by $\theta \in [0, 2\pi[^n$ and given by*

$$\mathcal{E}(\theta) = \frac{1}{2} \sum_{i,j=1}^{n} A_{ij} \left(1 - \cos(\theta_i - \theta_j)\right),$$

are the global minima corresponding to $\theta_i = c, \forall_i$. A graph G is said to be globally synchronizing if its adjacency matrix is globally synchronizing.

Abdalla, Bandeira, Kassabov, Souza, Strogatz, and Townsend [ABKS23] recently showed that a random Erdős-Rényi graph is globally synchronizing above the connectivity threshold with high probability (you can also see the Quanta article covering this at https://www.quantamagazine.org/new-proof-shows-that-expander-graphs-synchronize-20230724/). The same paper also shows that a uniform d-regular graph is globally synchronizing with high probability for $d \geq 600$, but it leaves open the following question.

Conjecture 3 (Globally Synchronizing Regular Graphs). *A uniform random 3-regular graph is globally synchronizing with high probability (probability going to 1 as $n \to \infty$).*

The following question is motivated by trying to understand the Burer-Monteiro algorithmic approach to community detection in the stochastic block model [Kur16]. The most recent progress is in [MABB24] which answers a high rank version of this question.

Conjecture 4 (Global Synchrony with negative edges). *Given any $\varepsilon > 0$, the $n \times n$ random matrix A with zero diagonal and off-diagonal entries given by*

$$A_{ij} = \begin{cases} 1 & \text{with probability} & \frac{1}{2} + \delta \\ -1 & \text{with probability} & \frac{1}{2} - \delta, \end{cases}$$

with $\delta \geq (1 + \varepsilon)\sqrt{\frac{\log n}{2n}}$, is globally synchronizing with high probability.

Another elusive question about global synchrony is the "extremal combinatorics" version of the question.

Conjecture 5 (Density threshold for Global Synchrony). *For any $\varepsilon > 0$, there exists $n > 0$ and a graph G on n nodes such that the minimum degree of G is at least $\left(\frac{3}{4} - \varepsilon\right) n$ and G is not globally synchronizing.*

Kassabov, Strogatz, and Townsend [KST21] showed an upper bound on the $\frac{3}{4}$ threshold in Conjecture 5, and the same paper contains evidence that $\frac{3}{4}$ is the correct threshold (by arguing that an upper bound below $\frac{3}{4}$ would need to go beyond linear analysis of the dynamical system).

Entry 3   Fitting ellipsoids to random points (AM)

In this entry I will discuss a seemingly simple question of high-dimensional stochastic geometry, which originated (as far as I know) in the series of works [Sau11, SCPW12, SPW13]:

For $n, d \gg 1$, when can a sequence of n i.i.d. random points in $\mathbb{R}^d$, drawn from $\mathcal{N}(0, \mathrm{I}_d/d)$, be fitted by the boundary of a centered ellipsoid?
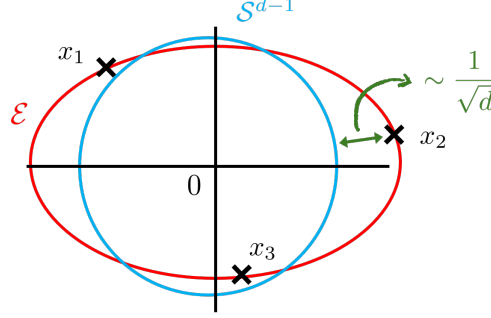
Figure 1:   Fitting Gaussian random points $x_i \sim \mathcal{N}(0, \mathrm{I}_d/d)$ to an ellipsoid. Notice that the unit sphere itself is close to being a fit by simple concentration of measure: a random $x \sim \mathcal{N}(0, \mathrm{I}_d/d)$ has (with high probability) distance $\mathcal{O}(1/\sqrt{d})$ to it.

The covariance being I_d/d is a convention which ensures $\mathbb{E}[\|x_i\|^2] = 1$: it is clear that assuming a generic positive-definite covariance $\Sigma \succ 0$ does not change this question, so we do not lose any generality with this assumption.

The motivation of [Sau11, SCPW12, SPW13] for this question came from statistics: they showed that the probability of existence of an ellipsoid fit was equal to the one of the success of a canonical convex relaxation (called Minimum-Trace Factor Analysis) in a problem of low-rank matrix decomposition. Several other motivations were uncovered later on, and I recommend reading the introduction of [PTVW23] to learn more about them.

Interestingly, it was soon conjectured that a sharp transition occurs in the regime $n/d^2 \to \alpha > 0$, exactly at $\alpha = 1/4$, see e.g. Conjecture 2 of [SPW13].

Conjecture 6 (Ellipsoid fitting). *Let $n, d \geq 1$ and $x_1, \cdots, x_n \sim \mathcal{N}(0, \mathrm{I}_d/d)$. For any $\varepsilon > 0$:*

$$\limsup_{d \to \infty} \frac{n}{d^2} \leq \frac{1 - \varepsilon}{4} \Rightarrow \lim_{d \to \infty} \mathbb{P}[\exists \mathcal{E} \text{ an ellipsoid fit to } (x_i)_{i=1}^n] = 1, \tag{2}$$

$$\liminf_{d \to \infty} \frac{n}{d^2} \geq \frac{1 + \varepsilon}{4} \Rightarrow \lim_{d \to \infty} \mathbb{P}[\exists \mathcal{E} \text{ an ellipsoid fit to } (x_i)_{i=1}^n] = 0. \tag{3}$$

Upper bounds – Ellipsoid fitting can be written as a semidefinite program (SDP), since any origin-centered ellipsoid satisfies $\mathcal{E} = \{x \in \mathbb{R}^d : x^T S x = 1\}$ for some positive semidefinite symmetric matrix S, so:

$$\mathbb{P}[\exists \mathcal{E} \text{ an ellipsoid fit to } (x_i)_{i=1}^n] = \mathbb{P}[\exists S \in \mathbb{R}^{d \times d} : S \succeq 0 \text{ and } x_i^T S x_i = 1 \text{ for all } i \in [n]]. \tag{4}$$

It's then not hard to convince oneself that $\{x_i^T S x_i = 1\}_{i=1}^n$ is an independent system of n linear equations on S: this already gives an upper bound of $\binom{d+1}{2} \simeq d^2/2$ to the number of points that can admit an ellipsoid fit (with high probability). As of now, this "silly" argument (which does not take into account the constraint that $S \succeq 0$) is the best upper bound established for Conjecture 6.

4

Lower bounds – On the other hand, there has been an abundance of works to establish lower bounds [Sau11, SCPW12, SPW13, KD23, PTVW23, HKPX23, TW23, BMMP24]. Not diving into details, they essentially all rely on carefully picking a candidate solution $S^\star$ to the linear system $\{x_i^T S x_i = 1\}_{i=1}^n$ (e.g. the least Frobenius norm solution), and establishing that $S^\star \succeq 0$ for small enough n. The best results in this vein are due to [TW23, HKPX23, BMMP24], which independently proved the following.

Theorem 3.1 ([TW23, HKPX23, BMMP24]). *Let $n, d \geq 1$ and $x_1, \cdots, x_n \sim \mathcal{N}(0, \mathrm{I}_d/d)$. There is $\delta > 0$ such that:*

$$\limsup_{d \to \infty} \frac{n}{d^2} \leq \delta \Rightarrow \lim_{d \to \infty} \mathbb{P}[\exists \mathcal{E} \text{ an ellipsoid fit to } (x_i)_{i=1}^n] = 1.$$

While it is not known if these methods can be pushed all the way to $\delta = 1/4$, it is conjectured that picking $S^\star$ as the minimal nuclear norm solution to the linear system $\{x_i^T S x_i = 1\}_{i=1}^n$ gives this optimal value, and could be a path towards proving eq. (2) [MK24].

The transition at $d^2/4$ – Notice that the $d^2/4$ threshold can be uncovered both by numerical simulations (solving a SDP can be done efficiently), but also by a heuristic argument: if the linear equations $\{\mathrm{Tr}[S x_i x_i^T] = 1\}_{i=1}^n$ were replaced by $\{\mathrm{Tr}[S G_i] = 1\}_{i=1}^n$ with G_i Gaussian i.i.d. matrices, the existence of a sharp transition would be a direct consequence of Gordon's theorem [Gor88], and the value $d^2/4$ can even be recovered in this case as the squared Gaussian width of the cone of positive semidefinite matrices! While this heuristic was well-known, we formalized it in [MB23], essentially by showing that the volume of the space of solutions is universal in both problems. We established from it the following sharp transition result.

Theorem 3.2 ([MB23]). *For any $\varepsilon, M > 0$, we define*

$$\mathrm{EFP}_{\varepsilon, M} \ : \ \exists S \in \mathbb{R}^{d \times d} \ : \ \mathrm{Sp}(S) \subseteq [0, M] \text{ and } \frac{1}{n} \sum_{i=1}^n |x_i^T S x_i - 1| \leq \frac{\varepsilon}{\sqrt{d}}.$$

Notice that the original ellipsoid fitting problem of Conjecture 6 is $\mathrm{EFP}_{0, \infty}$. We prove:

$$\limsup_{d \to \infty} \frac{n}{d^2} = \alpha < \frac{1}{4} \ \Rightarrow \ \exists M_\alpha > 0 \ : \ \forall \varepsilon > 0, \ \lim_{d \to \infty} \mathbb{P}[\mathrm{EFP}_{\varepsilon, M_\alpha}] = 1, \tag{5}$$

$$\liminf_{d \to \infty} \frac{n}{d^2} = \alpha > \frac{1}{4} \ \Rightarrow \ \exists \varepsilon_\alpha > 0 \ : \ \forall M > 0, \ \lim_{d \to \infty} \mathbb{P}[\mathrm{EFP}_{\varepsilon_\alpha, M}] = 0. \tag{6}$$

A few remarks are in order to clarify the conclusion of Theorem 3.2.

- The setting $\varepsilon \ll 1$ (but not going to 0 with d) is precisely the regime where the problem is not trivially solved by the unit sphere (i.e. $S = \mathrm{I}_d$), since $(1/n) \sum_{i=1}^n |x_i^T x_i - 1| = \mathcal{O}(1/\sqrt{d})$ (cf. also Fig. 1).

- In the regime $n/d^2 < 1/4$, Theorem 3.2 shows that there exists ellipsoids which are: (i) well-behaved (the spectral norm of S is bounded), and (ii) they fit the points (x_i) up to an arbitrarily small error (but not going to 0 with d).

- In the regime $n/d^2 > 1/4$, Theorem 3.2 rules out any well-behaved ellipsoid (i.e. with bounded spectral norm) as a possible fit, even allowing a small fitting error.

Theorem 3.2 provides the first mathematical result identifying a satisfiability transition in the ellipsoid fitting problem at the conjectured $d^2/4$ threshold. While we were not yet able to close Conjecture 6 from it, some parts seem tantalizingly close: e.g. the non-existence of ellipsoid fits for $n/d^2 > 1/4$ would now follow from showing that there is no ill-behaved ellipsoid fit (or that if there is an ill-behaved ellipsoid fit, there must also be a well-behaved one).

Other works – To keep this post at a reasonable length, I have not mentioned some other recent works which are in some way related to this problem, and to which I contributed [MK24, MTM$^+$24, ETB$^+$24].

In [MK24] we use non-rigorous tools that originate in statistical physics to extend the conjecture to other classes of random points beyond Gaussian distributions (and the conjectured satisfiability threshold can then be very different from $d^2/4!$), but also to study the typical geometrical shape of ellipsoid fits, and even to predict analytically the performance of the methods used to prove the lower bounds discussed above (cf. the conjecture on the minimal nuclear norm solution mentioned below Theorem 3.1).

Finally, in [MK24] we realized that the ideas behind Theorem 3.2 can be adapted to problems in statistical learning: precisely, we characterize optimal learning from data in a model of so-called "extensive-width" neural network, a regime which is both particularly relevant in practical applications and which had so far largely resisted to theoretical analysis. In another recent preprint [ETB$^+$24], we build upon these ideas to study a toy model of a transformer architecture.

Entry 4   Detection in Multi-Frequency synchronization (AK)

In the study of average-case complexity, one is often interested in the critical threshold of the signal-to-noise ratio at which a problem becomes tractable. In this entry, we will consider a multi-frequency synchronization problem, where one obtains noisy measurements of the relative alignments of the signal elements through multiple frequency channels. We are interested whether it becomes possible to detect the signal using a time-efficient algorithm at a lower signal-to-noise ratio compared to the single-frequency model.

Let us define the synchronization problem more formally. In a general synchronization problem over a group G, one aims to recover a group-valued signal $g = (g_1, \ldots, g_n) \in G^n$ from its noisy pairwise information on $g_k g_j^{-1}$. One way to model such observations is through receiving a function of $g_k g_j^{-1}$ corrupted with additive Gaussian noise

$$Y_{kj} = f(g_k g_j^{-1}) + W_{kj}, \quad W_{kj} \sim \mathcal{N}(0, 1).$$

In this entry, we will focus on a setting where measurements are available for all pairs (j, k).

Motivated by the Fourier decomposition of a non-linear objective of the non-unique games problem [BCLS20], we consider the model as receiving measurements through several frequency channels. For example, for $G = U(1) = \{e^{i\varphi}, \varphi \in [0, 2\pi)\}$, we consider measurements of the following form:

$$\begin{cases} Y_1 = \dfrac{\lambda}{n} x x^* + \dfrac{1}{\sqrt{n}} W^{(1)}, \\[2mm] Y_2 = \dfrac{\lambda}{n} x^{(2)} \left(x^{(2)}\right)^* + \dfrac{1}{\sqrt{n}} W^{(2)}, \\[2mm] \quad \vdots \\[2mm] Y_L = \dfrac{\lambda}{n} x^{(L)} \left(x^{(L)}\right)^* + \dfrac{1}{\sqrt{n}} W^{(L)}, \end{cases}$$

where $W^{(1)}, \ldots, W^{(L)}$ are independent matrices, and $x^{(k)} = (x_1^k, \ldots, x_n^k)$ denotes entrywise power. This case corresponds to the angular synchronization, where the objective is to determine phases $\varphi_1, \ldots, \varphi_n \in [0, 2\pi]$ from their noisy relative observation $\varphi_k - \varphi_j \mod 2\pi$, and equivalent to the synchronization over $SO(2)$, as $SO(2) \cong U(1)$.

In the general case of compact groups, we consider the Peter-Weyl decomposition instead. Informally, in this case, the Fourier modes correspond to irreducible representations, and each pairwise measurement corresponding to an irreducible representation ρ is a block matrix with blocks given by

$$Y_{kj}^\rho = \frac{\lambda}{n} \rho(g_k) \rho\left(g_j^{-1}\right) + \frac{1}{\sqrt{n}} W_{kj}^{(\rho)} \text{ for each } \rho \in \Psi.$$

The single-frequency model (i.e., when $L = 1$) reduces to the Wigner spiked matrix model. In this model, the celebrated BBP transition [BAP05] postulates that above a critical value $\lambda \geq \lambda^* = 1$, detection is possible based on the top eigenvalue, while below this threshold, the top eigenvalue does not provide reliable

information on the presence of the signal as n grows to infinity. Moreover, for a variety of dense priors on signal x, this procedure is optimal in the sense that no algorithm can detect the signal reliably below the spectral threshold, $\lambda^* = 1$, including algorithms with no constraints on the runtime [PWBM16].

Does the situation change for the multi-frequency model? For simplicity, let us assume that the signal x is sampled uniformly from L-th roots of unity, $x \sim \text{Unif}(\{e^{2\pi i k/L}\}_{k=0}^{L-1})$. This corresponds to synchronization over the cyclic group $\mathbb{Z}/L\mathbb{Z}$. In this case, [PWBM16] showed that the statistical threshold is $\lambda_{\text{stat}} = \Theta(\log L/L)$, i.e., below this threshold it is impossible to detect the signal, while above the threshold, there exists an inefficient algorithm for this task. The authors give upper and lower bounds on λ_{stat} with exact constants, and the upper bound is lower than the spectral threshold for a sufficiently large number of frequencies, namely, $L \geq 11$.

From the computational point of view, in [KBK24], it was shown that assuming the low-degree conjecture, no polynomial-time algorithm can detect the signal below the spectral threshold, $\lambda^* = 1$, regardless of receiving additional information through additional channels. This result applies to the setting when the number of frequencies L is constant compared to the dimension n and when the signal is sampled uniformly from $SO(2)$ or any finite group of constant size. Combining this result with the optimality of PCA for the single-frequency model, this suggests that receiving only constant number of additional frequencies does not lower the computational threshold. This opens up an intriguing question: how much more frequencies one requires so that the detection becomes possible using an efficient algorithm below the spectral threshold? We can formulate the following conjecture.

Open Problem 7. *Consider synchronization model over $SO(2)$ or synchronization model over a finite group, where the signal x is sampled uniformly over group elements. Find a scaling $L = L_n$ of number of frequencies such that there exists a polynomial-time algorithm that can detect the signal reliably for all $\lambda > \lambda_{comp,L}$ for some $\lambda_{comp,L} < 1$ as $n \to \infty$.*

There is empirical evidence supporting the conjecture that, as L diverges, the computational threshold becomes lower than the spectral threshold: numerical simulations in [GZ19] suggest that the variant of AMP with a carefully performed initialization can surpass the threshold when the dimension n and the number of frequencies L are comparable.

In the computationally hard regime, we may also hope that there exists a sub-exponential algorithm whose runtime possibly depends on the signal strength. Such a tradeoff was observed in the sparse PCA problem [DKWB23]. For the synchronization model over finite groups, [PWBM16] proposed an algorithm for detecting the signal that works in the entire computationally hard regime. This algorithm involves maximizing a certain test function over all possible solutions $g \in G^n$ and thus has an exponential runtime.

Open Problem 8. *Consider synchronization model over $SO(2)$ or synchronization model over a finite group, where the signal x is sampled uniformly over group elements. Fix number of frequencies L that does not depend on n. In the computationally hard regime, i.e., when $\lambda \in (\lambda_{stat}, 1)$, does there exist a sub-exponential algorithm, i.e., an algorithm running in time $\exp n^\delta$ for some $\delta < 1$, that can detect the signal reliably as $n \to \infty$?*

Entry 5   Tensor PCA and the Kikuchi algorithm (ASB)

In 2014, Andrea Montanari and Emile Richard [MR14] proposed a statistical model for understanding a variant of Principal Component Analysis in Tensor data. This is currently referred as the Tensor PCA problem. We will consider symmetric version of the problem in which the signal of interest is a point in the hypercube. Given n, r and λ, the goal is to estimate (or detect) an unknown "signal" $x \in \{\pm 1\}^n$ (drawn uniformly from the hypercube), from "measurements" as follows: for $i_1 < i_2 < ... < i_r$,

$$Y_{i_1,i_2,...,i_r} = \lambda x^{\otimes r} + Z_{i_1,i_2,...,i_r}$$

where $Z_{i_1,i_2,...,i_r}$ are iid $\mathcal{N}(0,1)$ random variables (and independent from x).

We will consider r (and ℓ, to be introduced below) fixed and $n \to \infty$, all big-O notation will be in terms of n. Tensor PCA is believed to undergo a statistical-to-computaional gap: without regards for computational efficiency, it is possible to estimate x for $\lambda = \Omega\left(n^{-\frac{r}{2}+\frac{1}{2}}\right)$. Efficient algorithms, such as the Sum-of-Squares hierarchy, are able to solve the problem at $\lambda = \tilde{\Omega}\left(n^{-\frac{r}{4}}\right)$, where $\tilde{\Omega}$ hides logarithmic factors. Local methods, such as gradient descent and approximate message passing succeed at $\lambda = \tilde{\Omega}\left(n^{-\frac{1}{2}}\right)$. For $r = 2$, the problem simply involves matrices, and indeed all these thresholds coincide. We point the reader to [WAM19] and references therein for more on each of these thresholds.

In 2019, Alex Wein, Ahmed El Alaoui, and Cris Moore [WAM19] proposed an algorithm for this problem based on the so-called Kikuchi Free Energy, it roughly corresponds to a hierarchy of message passing algorithms. They showed that this approach achieves (up to logarithmic factors) the threshold for the sum-of-squares approach, shedding light on the gap between the message passing and sum-of-squares frameworks.

We briefly describe the Kikuchi approach. We will focus on even r. There is a design parameter ℓ (with $n \gg \ell \geq \frac{r}{2}$) which we will consider fixed. The Kikuchi matrix M is the $\binom{n}{\ell} \times \binom{n}{\ell}$ matrix (with rows and columns indexed by subsets $I \subset [n]$ of size ℓ) given by

$$M(\lambda)_{I,J} = \begin{cases} Y_{I\Delta J} & \text{if} & |I\Delta J| = r, \\ 0 & \text{otherwise}, \end{cases}$$

where $I\Delta J = (I \cup J) \setminus (I \cap J)$ denotes the symmetric difference.

The goal is to understand when the top of the spectrum of M reveals the spike x. It is shown in [WAM19] that this happens for $\lambda = \tilde{\Omega}\left(n^{-\frac{r}{4}}\right)$. Since Z is rotation invariant, we can assume WLOG that $x = \mathbf{1}$, the all-ones vectors. The following conjecture also appears in [Ban24] and is a reformulation of a conjecture in [WAM19]:

Conjecture 9 (Kikuchi Spectral Threshold). *Given r, ℓ, n positive integers satisfying $n \gg \ell \geq \frac{r}{2}$ and r even (r and ℓ will be fixed and $n \to \infty$). For each $S \subset [n]$ with $|S| = r$ let $Z_S \sim \mathcal{N}(0,1)$, and all independent. Let $\lambda \geq 0$ and let M be the $\binom{n}{\ell} \times \binom{n}{\ell}$ matrix (with rows and columns indexed by subsets $I \subset [n]$ of size ℓ) given by*

$$M(\lambda)_{I,J} = \begin{cases} \lambda + Z_{I\Delta J} & \text{if} & |I\Delta J| = r, \\ 0 & \text{otherwise}, \end{cases}$$

where $I\Delta J = (I \cup J) \setminus (I \cap J)$ denotes the symmetric difference.

Let $\lambda_{r,\ell}^{\natural}$ denote the threshold at which eigenvalues "pop-out" of the spectrum of $M(\lambda)$: in other words $\lambda_{r,\ell}^{\natural}$ is the real number such that, for all $\lambda > \lambda_{r,\ell}^{\natural}$, there exists $\varepsilon > 0$ such that $\mathbb{E}\lambda_{\max}M(\lambda) > (1 + \varepsilon + o(1))\mathbb{E}\lambda_{\max}M(0)$, where $o(1)$ is a term that goes to zero as $n \to \infty$.

For fixed r, we have

$$n^{\frac{r}{4}}\lambda_{r,\ell}^{\natural} \to 0$$

as $\ell \to \infty$ (note that this is after one has taken $n \to \infty$).

This would establish the very interesting phenomenon that there is no sharp threshold for polynomial time algorithms in Tensor PCA in the following sense: This conjecture would imply that by increasing ℓ (while remaining $\Theta(1)$, corresponding to increasing the computational cost of the algorithm while keeping it polynomial time), one would be able to decrease the critical signal-to-noise ratio in the sense of $n^{\frac{r}{4}}\lambda_{r,\ell}^{\natural}$ getting arbitrarily close to zero. Since the bound in [WAM19] contains logarithmic factors on n (present in $\tilde{\Omega}(\cdot)$), it does not allow to see this phenomenon for $\ell = \Theta(1)$.

For even r and $\frac{1}{2}r \leq \ell < \frac{3}{4}r$ the threshold has been characterized in my work with Giorgio Cipolloni, Dominik Schröder, and Ramon van Handel [BCSvH24], which is a follow up to the matrix concentration inequalities (leveraging intrinsic freeness) in my work with March Boedihardjo, and Ramon van Handel [BBvH23].

Entry 6   Did just a couple of deviations suffice all along? (ASB)

In September 2024, at a conference on Mathematical Aspects of Learning Theory in Barcelona (`https://www.crm.cat/mathematical-aspects-of-learning-theory/`) Dan Spielman gave a beautiful talk on discrepancy theory and some of its applications in clinical trials. Even though this was not the precise focus of the talk, it sparked a discussion that day about lower bounds for Joel Spencer's Six Deviations Suffice Theorem. I will describe some of the unanswered questions from this discussion. Many thanks to Dan Spielman, Tselil Schramm, Amir Yehudayoff, Petar Nizic-Nikolac, and Anastasia Kireeva with whom discussing this problem contributed to making the workshop a memorable week!

Given n, a positive integer, and A, an $n \times n$ matrix, we define the discrepancy of A, $\mathrm{disc}(A)$, as

$$\mathrm{disc}(A) = \inf_{x \in \{\pm 1\}^n} \|Ax\|_\infty.$$

One of the motivations of this question is to understand how much smaller is $\inf_{x \in \{\pm 1\}^n} \|Ax\|_\infty$ compared to the value, e.g. in expectation, when x is drawn uniformly from the hypercube. We point to the references in this post for a more thorough discussion of the context, motivations, and state of the art.

In 1985, Joel Spencer [Spe85a] showed the cellebrated "Six Deviations Suffice" Theorem, which states that for any positive integer n and any $A \in \{\pm 1\}^{n \times n}$ $\mathrm{disc}(A) \le 6\sqrt{n}$. There have since been improvements on this constant, but the question we will pursue here concerns lower bounds.

When the condition of ± 1 entries is replaced by an ℓ_2 condition on the columns, the corresponding question is the famous Kómlos Conjecture (which coincidentally is the first open problem in my lecture notes from around a decade ago [Ban16]).

Conjecture 10 (Kómlos Conjecture). *There exists a universal constant K such that, for all square matrices A whose columns have unit ℓ_2 norm, $\mathrm{disc}(A) \le K$.*

It is easy to see that the statement of this conjecture implies Spencer's Theorem, since the columns of an $n \times n$, ± 1 matrix, have ℓ_2 norm $\sqrt{n}$. The best current lower bound for Kómlos is a recent construction by Tim Kunisky [Kun23] which shows a lower bound on K of $1 + \sqrt{2}$. Unfortunately, the resulting matrix is not a ± 1 matrix (even after scaling), and so it does not provide a lower bound for the Spencer setting.

Open Problem 11 (How many deviations are needed afterall?). *What is the value of the following quantity?*

$$\sup_{n \in \mathbb{N}} \sup_{A \in \{\pm 1\}^{n \times n}} \frac{1}{\sqrt{n}} \mathrm{disc}(A)$$

It is easy to see that the matrix $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ has discrepancy $2 = \sqrt{2}\sqrt{2}$, giving a lower bound of $\sqrt{2}$ to this quantity. Also, Dan Spielman mentioned a numerical construction (of a specific size) that achieved a slightly larger discrepancy value. To the best of my knowledge, there is no known lower bound above 2, motivating the question: "Did just a couple of deviations suffice all along?".

Also, to the best of my knowledge, there is no known construction (or proof of existence) of an infinite family achieving a lower bound strictly larger than one. More explicitly, we formulate the following conjecture.

Conjecture 12. *Given the definitions above, we have*

$$\limsup_{n \to \infty} \sup_{A \in \{\pm 1\}^{n \times n}} \frac{1}{\sqrt{n}} \mathrm{disc}(A) > 1.$$

Hadamard matrices
A natural place to search for a family proving a lower bound is to use the Sylvester construction of Hadarmard matrices. An Hadarmard matrix is a ± 1 matrix whose columns are all orthogonal (in other words, after scaling by $\frac{1}{\sqrt{n}}$ it is an orthogonal matrix). In 1867 James Joseph Sylvester proposed the

following construction for sizes n that are powers of 2. $H_0 = 1$ and H_k is the $2^k \times 2^k$ matrix given by $H_k = H_1 \otimes H_{k-1}$. In other words,

$$H_k = \begin{bmatrix} H_{k-1} & H_{k-1} \\ H_{k-1} & -H_{k-1} \end{bmatrix}.$$

Recall that $n = 2^k$. Since $\frac{1}{\sqrt{n}} H_k$ is an orthogonal matrix, it is easy to see that $\|H_k x\|_2 = n$ for all $x \in \{\pm 1\}^n$ and thus $\mathrm{disc}(H_k) \geq \sqrt{n}$. This is not tight for H_1 (as its discrepancy is 2), however it is tight for H_2: e.g. taking the eigenvector $y = (1, 1, 1, -1)$.

To give an upper bound on $\mathrm{disc}(H_k)$ we proceed with an explicit construction of a ± 1 vector, using the (optimal) eigenvector y for H_2. Let $x_{(0)}^\flat = 1$, $x_{(1)}^\flat = (1, 1)$, and for $k \geq 2$, $x_{(k)}^\flat \in \{\pm 1\}^{2^k}$ is given by

$$x_{(k)}^\flat = y \otimes x_{(k-2)}^\flat = \begin{bmatrix} x_{(k-1)}^\flat \\ x_{(k-1)}^\flat \\ x_{(k-1)}^\flat \\ -x_{(k-1)}^\flat \end{bmatrix}.$$

Since $H_k = H_1 \otimes (H_1 \otimes H_{k-2}) = H_2 \otimes H_{k-2}$, the mixed-product property gives

$$H_k x_{(k)}^\flat = (H_2 \otimes H_{k-2})(y \otimes x_{(k-2)}^\flat) = (H_2 y) \otimes (H_{k-2} x_{(k-2)}^\flat) = (2y) \otimes (H_{k-2} x_{(k-2)}^\flat).$$

Inducting this argument shows that

$$\left\| H_k x_{(k)}^\flat \right\|_\infty = \begin{cases} \sqrt{2}\sqrt{2^k} & \text{if } k \text{ is odd,} \\ \sqrt{2^k} & \text{if } k \text{ is even.} \end{cases}$$

While this settles the discrepancy of H_k for k even, the odd case is, to the best of my knowledge, open.

Conjecture 13. *Let H_k be the $2^k \times 2^k$ Hadamard matrix obtained by the Sylvester construction (see above). For odd k, we have*

$$\mathrm{disc}(H_k) = \sqrt{2}\sqrt{2^k}.$$

Note that this would settle Conjecture 12.

Remark 6.1 (Boolean Fourier Analysis)**.** *I am not going to discuss this at length here, but since H_k essentially corresponds to the boolean Fourier transform for boolean functions on k variables, the discrepancy of H_k has a natural interpretation in Boolean Analysis. We point the reader to [O'D14] for more on this subject.*

In a final note, I could not mention Hadamard matrices without stating a famous conjecture regarding their existence.

Conjecture 14 (Hadamard Conjecture)**.** *For any positive n a multiple of 4, there exists an $n \times n$ Hadamard matrix.*

Entry 7   Sampling from the Sherrington-Kirkpatrick Model (AR)

In recent years, significant progress has been made in understanding the mixing behavior of Glauber dynamics for Ising models, particularly in the Sherrington-Kirkpatrick (SK) model. Despite the classical Dobrushin-uniqueness condition being too restrictive for this model, recent breakthroughs employing techniques like stochastic localization and spectral analysis have improved the understanding of the fast-mixing regime. In this entry, we will discuss these improved mixing bounds for Glauber dynamics in the SK model and the open questions that remain.

Let us start with some preliminaries. Consider the following measure on the hypercube $\{\pm 1\}^n$:

$$\mu_{J,h}(x) \propto \exp\left(\frac{1}{2}x^T J x + h^T x\right),$$

where $\propto$ corresponds to equal up to a normalizing constant. Here, the interaction matrix $J \in \mathbb{R}^{n \times n}$ encodes pairwise interactions between spins, while h represents the influence of an external field. We will particularly focus on $J = \beta W$ being a random matrix with $W \sim GOE(n)$ and inverse temperature $\beta > 0$. This model was introduced by Sherrington and Kirkpatrick [SK75] and we refer to it as SK-model. The reader can focus on the case $h = 0$. All statements below are with high probability on the disorder W.

Glauber Dynamics is an MCMC algorithm which chooses a site $i \in [n]$ uniformly at random and updates the spin according to μ conditioned on the remaining spins. The dynamics can be represented by the following transition kernel:

$$P(x,y) = \begin{cases} \frac{1}{n}\mu(X_i = y_i \mid X_{-i} = x_{-i}), & \text{if } d_H(x,y) = 1 \\ 1 - \frac{1}{n}\sum_{j=1}^n \mu(X_j \neq x_j \mid X_{-i} = x_{-i}), & \text{if } d_H(x,y) = 0 \\ 0, & \text{otherwise}, \end{cases}$$

representing the probability of transitioning from x to y. Here, d_H denotes the Hamming distance in the hypercube. Mixing in time T_ϵ means that, regardless of the starting point, the measure of the Markov chain at time T_ϵ is ϵ-close to its stationary measure μ with respect to the total variation distance (although other notions of distances can also be used). The reader should think of ϵ as a small constant.

One celebrated and generally sufficient condition of Glauber Dynamics mixing in $O(n\log(n))$ time is the Dobrushin-uniqueness condition [Dob68] which requires $\max_i \sum_j |J_{ij}| < 1$. While this condition is tight for some models (e.g. the Curie-Weiss-model), it only proves fast mixing for very high temperatures ($\beta = O(\frac{1}{\sqrt{n}})$) in the SK-model. In fact, a long standing open problem asked to extend the results above for some $\beta = O(1)$.

Indeed, in recent years, new proof techniques based on stochastic localization schemes have emerged and, when applied to the SK-model, significantly improved the bounds on β for guaranteeing fast mixing. Eldan, Koehler and Zeitouni [EKZ22] were able to prove the following general (deterministic) spectral condition: As long as

$$\lambda_{\max}(J) - \lambda_{\min}(J) < 1 - \delta$$

for some $\delta > 0$, then Glauber Dynamics mixes in $O(n^2)$ time with respect to total variation distance (big-O notation may hide constants depending on ϵ). Applied to the SK-model this guarantees fast mixing for inverse temperature $\beta < \frac{1}{4}$.

The bound on the width of the spectrum of the Interaction Matrix proves to be tight again in the Curie-Weiss-Model. Moreover, based on the Low-degree-conjecture, Kunisky [Kun24] gave evidence that no polynomial-time sampling algorithm exists that can succeed in general for all J with $\lambda_{\max}(J) - \lambda_{\min}(J) \leq 1 + \delta$. Hence, for improving bounds in the SK-model using spectral analysis, it does not suffice to measure the distance of the largest and the smallest eigenvalue of the interaction matrix, but additional properties of the interaction matrix in the SK model have to be used. One nice example is the symmetry of the spectrum: In contrast to the all-ones-matrix which defines the Curie-Weiss-Interaction matrix, the spectrum of W is almost symmetric around 0. Anari, Koehler and Vuong [AKV24] used exactly this property to improve the bound on β: When

$$\frac{|\lambda_{\max}(J)|}{|\lambda_{\min}(J)|} = 1 + o(1),$$

Glauber Dynamics mixes with respect to Kullback-Leibler divergence in $O(n\log(n))$ time, when

$$\lambda_{\max}(J) - \lambda_{\min}(J) < 1.18$$

and $\lambda_{\min}(J) < 0$. The proof is based on a trickle down equation in a stochastic localization scheme and the authors introduce novel techniques to bound the operator norm of covariance matrices, depending on both

the ratio of $\lambda_{\max}(J)$ and $\lambda_{\min}(J)$ and $\lambda_{\max}(J) - \lambda_{\min}(J)$. This leads to the currently best-known upper bound $\beta < 0.295$ and it remains an open question whether this bound could be improved up to 1:

Open Problem 15. *Consider the SK-model with an arbitrary external field. Does Glauber Dynamics fast mix (meaning in polynomial time) up to the replica symmetry breaking threshold $\beta_* = 1$?*

El Alaoui, Montanari and Sellke [AMS22] proposed an algorithm for sampling from the SK-model without an external field that applies a stochastic localization scheme and makes use of the Approximate Message Passing framework. They proved fast mixing in Wasserstein Distance (which is a weaker metric than Total Variation Distance) up to $\beta < \frac{1}{2}$. Celentano [Cel24] extended the argument to all β up to the threshold $\beta_* = 1$. Moreover, El Alaoui, Montanari and Sellke, proved in the same paper [AMS22] that it is impossible for stable algorithms to sample from the SK model without external field for $\beta > 1$, which also rules out fast mixing for their algorithm.

Updates

While we generally do not expect to continuously update with progress on the questions discussed here, we try to include such updates if they happen (or we learn of them) before the posting of this manuscript. In this context, we are happy to report on two lines of progress: Conjecture 4 was answered by McRae in [McR25], following up on work of Rakoto Endor and Waldspurger [EW24]. Furthermore, Rares-Darius Buhai [Buh25] showed that Conjecture 13 is false by identifying a connection between discrepancy of H_k and a theory of non-linearity in Boolean Functions dating back, at least, to the work of Paterson and Wiedemann [PW83]; unfortunately, to the best of our knowledge, it does not rule out that the H_k family of matrices can be used to establish Conjecture 12.

References

[ABKS23] P. Abdalla, A. S. Bandeira, M. Kassabov, and V. Souza. Expander graphs are globally synchronizing. *Preprint, available on arXiv*, 2023.

[AKV24] Nima Anari, Frederic Koehler, and Thuy-Duong Vuong. Trickle-down in localization schemes and applications. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 1094–1105, 2024.

[AMS22] Ahmed El Alaoui, Andrea Montanari, and Mark Sellke. Sampling from the sherrington-kirkpatrick gibbs measure via algorithmic stochastic localization. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 323–334. IEEE, 2022.

[Ban16] Afonso S. Bandeira. Ten lectures and forty-two open problems in the mathematics of data science. *Available at $https://people.math.ethz.ch/\sim abandeira/TenLecturesFortyTwoProblems.pdf$*, 2016.

[Ban24] Afonso S. Bandeira. Ten open problems involving matrices, randomness, graphs, and more. In *Oberwolfach Report for Workshop 2417 (21.04.2024–26.04.2024). Available at $https://publications.mfo.de/handle/mfo/4149$*, 2024.

[BAP05] Jinho Baik, Gérard Ben Arous, and Sandrine Péché. Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *The Annals of Probability*, 33(5):1643–1697, 2005.

[BBvH23] A. S. Bandeira, M. Boedihardjo, and R. van Handel. Matrix concentration inequalities and free probability. *Inventiones Mathematicae*, (234):419–487, 2023.

[BCLS20]   Afonso S Bandeira, Yutong Chen, Roy R Lederman, and Amit Singer. Non-unique games over compact groups and orientation estimation in cryo-em. *Inverse Problems*, 36(6):064002, apr 2020.

[BCSvH24]   A. S. Bandeira, G. Cipolloni, D. Schroder, and R. van Handel. Matrix concentration inequalities and free probability ii. two-sided bounds and applications. *Preprint, available on arXiv*, 2024.

[BJM23]   N. Bansal, H. Jiang, and R. Meka. Resolving matrix spencer conjecture up to poly-logarithmic rank. *STOC*, 2023.

[BKMZ22]   A. S. Bandeira, D. Kunisky, D. Mixon, and X. Zeng. On the concentration of gaussian cayley matrices. *Preprint, available on arXiv*, 2022.

[BMMP24]   Afonso S Bandeira, Antoine Maillard, Shahar Mendelson, and Elliot Paquette. Fitting an ellipsoid to a quadratic number of random points. *To appear in Latin American Journal of Probability and Mathematical Statistics.*, 2024.

[Buh25]   Rares-Darius Buhai. *Private Communication*, 2025.

[Cel24]   Michael Celentano. Sudakov–fernique post-amp, and a new proof of the local convexity of the tap free energy. *The Annals of Probability*, 52(3):923–954, 2024.

[DJR22]   Daniel Dadush, Haotian Jiang, and Victor Reis. A new framework for matrix discrepancy: Partial coloring bounds via mirror descent. *STOC*, 2022.

[DKWB23]   Yunzi Ding, Dmitriy Kunisky, Alexander S Wein, and Afonso S Bandeira. Subexponential-time algorithms for sparse pca. *Foundations of Computational Mathematics*, pages 1–50, 2023.

[Dob68]   PL Dobruschin. The description of a random field by means of conditional probabilities and conditions of its regularity. *Theory of Probability & Its Applications*, 13(2):197–224, 1968.

[EKZ22]   Ronen Eldan, Frederic Koehler, and Ofer Zeitouni. A spectral condition for spectral gap: fast mixing in high-temperature ising models. *Probability theory and related fields*, 182(3):1035–1051, 2022.

[ETB⁺24]   Vittorio Erba, Emanuele Troiani, Luca Biggio, Antoine Maillard, and Lenka Zdeborová. Bilinear sequence regression: A model for learning from long sequences of high-dimensional tokens. *arXiv preprint arXiv:2410.18858*, 2024.

[EW24]   Faniriana Rakoto Endor and Irène Waldspurger. Benign landscape for burer–monteiro factorizations of maxcut-type semidefinite programs. *arXiv preprint arXiv:2411.03103*, 2024.

[Gor88]   Yehoram Gordon. On Milman's inequality and random subspaces which escape through a mesh in $\mathbb{R}^n$. In *Geometric Aspects of Functional Analysis: Israel Seminar (GAFA) 1986–87*, pages 84–106. Springer, 1988.

[GZ19]   Tingran Gao and Zhizhen Zhao. Multi-frequency phase synchronization. In *International Conference on Machine Learning*, 2019.

[HKPX23]   Jun-Ting Hsieh, Pravesh Kothari, Aaron Potechin, and Jeff Xu. Ellipsoid fitting up to a constant. In *International Colloquium on Automata, Languages and Programming, ICALP*, number 2023, 2023.

[HRS22]   Samuel B Hopkins, Prasad Raghavendra, and Abhishek Shetty. Matrix discrepancy from quantum communication. *STOC*, 2022.

[KBK24]   Anastasia Kireeva, Afonso S Bandeira, and Dmitriy Kunisky. Computational lower bounds for multi-frequency group synchronization. *arXiv preprint arXiv:2406.03424*, 2024.

[KD23] Daniel Kane and Ilias Diakonikolas. A nearly tight bound for fitting an ellipsoid to gaussian random points. In *The Thirty Sixth Annual Conference on Learning Theory*, pages 3014–3028. PMLR, 2023.

[KST21] M. Kassabov, S. H. Strogatz, and A. Townsend. Sufficiently dense kuramoto networks are globally synchronizing. *Chaos* **31**, 2021.

[Kun23] Dmitriy Kunisky. The discrepancy of unsatisfiable matrices and a lower bound for the komlós conjecture constant. *SIAM J. Discrete Mathematics (SIDMA)*, 37(2), 2023.

[Kun24] Dmitriy Kunisky. Optimality of glauber dynamics for general-purpose ising model sampling and free energy approximation. In *Proceedings of the 2024 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 5013–5028. SIAM, 2024.

[Kur75] Yoshiki Kuramoto. Self-entrainment of a population of coupled non-linear oscillators. *International Symposium on Mathematical Problems in Theoretical Physics*, pages 420–422, 1975.

[Kur16] Yoshiki Kuramoto. On the low-rank approach for semidefinite programs arising in synchronization and community detection. *COLT*, 2016.

[MABB24] A. McRae, P. Abdalla, A. S. Bandeira, and N. Boumal. Nonconvex landscapes for z2 synchronization and graph clustering are benign near exact recovery thresholds. *Preparing, available on arXiv*, 2024.

[MB23] Antoine Maillard and Afonso S Bandeira. Exact threshold for approximate ellipsoid fitting of random points. *arXiv preprint arXiv:2310.05787*, 2023.

[McR25] Andrew D. McRae. Benign landscapes for synchronization on spheres via normalized laplacian matrices. *arXiv preprint arXiv:2503.18801*, 2025.

[Mek14] Raghu Meka. Discrepancy and beating the union bound. *Windows On Theory Blog Post. Available at `https: // wind ows on the ory. or g/ 20 14 /0 2/ 07/ d is cr ep an cy -a nd -b ea ti ng -t he -u ni on -b ou nd`*, 2014.

[MK24] Antoine Maillard and Dmitriy Kunisky. Fitting an ellipsoid to random points: predictions using the replica method. *IEEE Transactions on Information Theory*, 2024.

[MR14] Andrea Montanari and Emile Richard. A statistical model for tensor pca. *Neural Information Processing Systems (NIPS 2014)*, 2014.

[MTM⁺24] Antoine Maillard, Emanuele Troiani, Simon Martin, Florent Krzakala, and Lenka Zdeborová. Bayes-optimal learning of an extensive-width neural network from quadratically many samples. *arXiv preprint arXiv:2408.03733*, 2024.

[O'D14] Ryan O'Donnell. *Analysis of Boolean Functions*. Cambridge Press, 2014.

[PTVW23] Aaron Potechin, Paxton M Turner, Prayaag Venkat, and Alexander S Wein. Near-optimal fitting of ellipsoids to random points. In *The Thirty Sixth Annual Conference on Learning Theory*, pages 4235–4295. PMLR, 2023.

[PW83] K. G. Paterson and D. H. Wiedemann. The maximum number of linearly independent vectors in a Bent function. *Proceedings of the National Academy of Sciences of the United States of America*, 80(22):6278–6279, 1983.

[PWBM16] Amelia Perry, Alexander S Wein, Afonso S Bandeira, and Ankur Moitra. Optimality and sub-optimality of pca for spiked random matrices and synchronization. *arXiv preprint arXiv:1609.05573*, 2016.

[Sau11] James James Francis Saunderson. *Subspace identification via convex optimization*. PhD thesis, Massachusetts Institute of Technology, 2011.

[SCPW12] James Saunderson, Venkat Chandrasekaran, Pablo A Parrilo, and Alan S Willsky. Diagonal and low-rank matrix decompositions, correlation matrices, and ellipsoid fitting. *SIAM Journal on Matrix Analysis and Applications*, 33(4):1395–1416, 2012.

[SK75] David Sherrington and Scott Kirkpatrick. Solvable model of a spin-glass. *Physical review letters*, 35(26):1792, 1975.

[Spe85a] Joel Spencer. Six standard deviations suffice. *Transactions of the American Mathematical Society*, 289(2):679–706, 1985.

[Spe85b] Joel H. Spencer. Six standard deviations suffice. *Transactions of the American Mathematical Society*, 289:679–706, 1985.

[SPW13] James Saunderson, Pablo A Parrilo, and Alan S Willsky. Diagonal and low-rank decompositions and fitting ellipsoids to random points. In *52nd IEEE Conference on Decision and Control*, pages 6031–6036. IEEE, 2013.

[TW23] Madhur Tulsiani and June Wu. Ellipsoid fitting up to constant via empirical covariance estimation. *arXiv preprint arXiv:2307.10941*, 2023.

[WAM19] A. Wein, A. El Alaoui, and C. Moore. The kikuchi hierarchy and tensor pca. *FOCS*, 2019.

[Zou12] Anastasios Zouzias. A matrix hyperbolic cosine algorithm and applications. In *International Colloquium on Automata, Languages, and Programming*, 2012.