Hyperbolic Concentration, Anti-Concentration, and Discrepancy

Zhao Song
Adobe Research, Seattle, WA, USA

Ruizhe Zhang
The University of Texas at Austin, TX, USA

Abstract

Chernoff bound is a fundamental tool in theoretical computer science. It has been extensively used in randomized algorithm design and stochastic type analysis. Discrepancy theory, which deals with finding a bi-coloring of a set system such that the coloring of each set is balanced, has a huge number of applications in approximation algorithms design. Chernoff bound [Che52] implies that a random bi-coloring of any set system with n sets and n elements will have discrepancy $O(\sqrt{n \log n})$ with high probability, while the famous result by Spencer [Spe85] shows that there exists an $O(\sqrt{n})$ discrepancy solution.

The study of hyperbolic polynomials dates back to the early 20th century when used to solve PDEs by Gårding [Går59]. In recent years, more applications are found in control theory, optimization, real algebraic geometry, and so on. In particular, the breakthrough result by Marcus, Spielman, and Srivastava [MSS15] uses the theory of hyperbolic polynomials to prove the Kadison-Singer conjecture [KS59], which is closely related to discrepancy theory.

In this paper, we present a list of new results for hyperbolic polynomials:

- We show two nearly optimal hyperbolic Chernoff bounds: one for Rademacher sum of arbitrary vectors and another for random vectors in the hyperbolic cone.
- We show a hyperbolic anti-concentration bound.
- We generalize the hyperbolic Kadison-Singer theorem [Brä18] for vectors in sub-isotropic position, and prove a hyperbolic Spencer theorem for any constant hyperbolic rank vectors.

The classical matrix Chernoff and discrepancy results are based on determinant polynomial which is a special case of hyperbolic polynomials. To the best of our knowledge, this paper is the first work that shows either concentration or anti-concentration results for hyperbolic polynomials. We hope our findings provide more insights into hyperbolic and discrepancy theories.

2012 ACM Subject Classification Theory of computation → Randomness, geometry and discrete structures

Keywords and phrases Hyperbolic polynomial, Chernoff bound, Concentration, Discrepancy theory, Anti-concentration

Digital Object Identifier 10.4230/LIPIcs.APPROX/RANDOM.2022.10

Category RANDOM

Related Version *Full Version*: https://arxiv.org/abs/2008.09593

Funding *Ruizhe Zhang*: This work was supported by Dana Moshkovitz's NSF Grant CCF-1648712.

Acknowledgements We thank the anonymous reviewers for helpful comments. The authors would like to thank Petter Brändén and James Renegar for many useful discussions about the literature of hyperbolic polynomials. The authors would like to thank Yin Tat Lee and James Renegar, Scott Aaronson for encouraging us to work on this topic. The authors would like to thank Dana Moshkovitz for giving comments on the draft.

Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2022).
Editors: Amit Chakrabarti and Chaitanya Swamy; Article No. 10; pp. 10:1–10:19
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Introduction

The study of concentration of sums of independent random variables dates back to Central Limit Theorems, and hence to de Moivre and Laplace, while modern concentration bounds for sums of random variables were probably first established by Bernstein [15] in 1924. An extremely popular variant now known as Chernoff bounds was introduced by Rubin and published by Chernoff [20] in 1952.

Hyperbolic polynomials are real, multivariate homogeneous polynomials $p(x) \in \mathbb{R}[x_1, \dots, x_n]$, and we say that $p(x)$ is hyperbolic in direction $e \in \mathbb{R}^n$ if the univariate polynomial $p(te - x) = 0$ for any x has only real roots as a function of t (counting multiplicities). The study of hyperbolic polynomials was first proposed by Gårding in [27] and has been extensively studied in the mathematics community [28, 32, 14, 71]. Some examples of hyperbolic polynomials are as follows:

- Let $h(x) = x_1 x_2 \cdots x_n$. It is easy to see that $h(x)$ is hyperbolic with respect to any vector $e \in \mathbb{R}^n_+$.
- Let $X = (x_{i,j})_{i,j=1}^n$ be a symmetric matrix where $x_{i,j} = x_{j,i}$ for all $1 \le i, j \le n$. The determinant polynomial $h(x) = \det(X)$ is hyperbolic with respect to $\widetilde{I}$, the identity matrix I packed into a vector. Indeed, $h(t\widetilde{I} - x) = \det(tI - X)$, the characteristic polynomial of the symmetric matrix X, has only real roots by the spectral theorem.
- Let $h(x) = x_1^2 - x_2^2 - \cdots - x_n^2$. Then, $h(x)$ is hyperbolic with respect to $e = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^\top$.

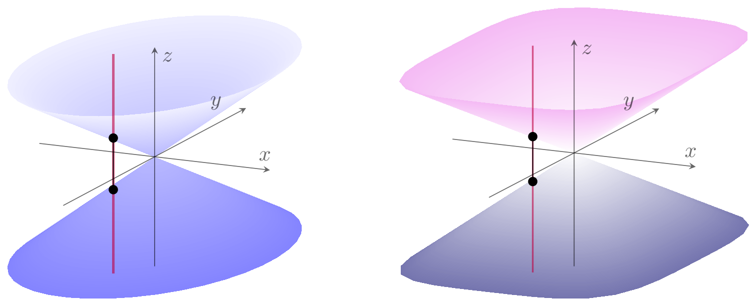

Figure 1 The function on the left is $h(x, y, z) = z^2 - x^2 - y^2$, which is hyperbolic with respect to $e = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^\top$, since any line in this direction always has two intersections, corresponding to the two real roots of $h(-x, -y, t - z) = 0$. The function on the right is $g(x, y, z) = z^4 - x^4 - y^4$, which is *not* hyperbolic with respect to e, since it only has 2 intersections but the degree is 4.

Inspired by the eigenvalues of matrix, we can define the hyperbolic eigenvalues of a vector x as the real roots of $t \mapsto h(te - x)$, that is, $\lambda_{h,e}(x) = (\lambda_1(x), \dots, \lambda_d(x))$ such that $h(te - x) = h(e) \prod_{i=1}^d (t - \lambda_i(x))$ (see Fact 12). In other words, the hyperbolic eigenvalues of x are the zero points of the hyperbolic polynomial restricted to a real line through x. In this paper, we assume that h and e are fixed and we just write $\lambda(x)$ omitting the subscript. Furthermore, similar to the spectral norm of matrix, the *hyperbolic spectral norm* of a vector x can be defined as

$$\|x\|_h = \max_{i \in [d]} |\lambda_i(x)|. \tag{1}$$

In this work, we study the concentration phenomenon of the roots of hyperbolic polynomials. More specifically, we consider the hyperbolic spectral norm of the sum of randomly signed vectors, i.e., $\| \sum_{i=1}^n r_i x_i \|_h$, where $r \in \{-1, 1\}^n$ are uniformly random signs and $\{x_1, x_2, \cdots, x_n\}$ are any fixed vectors in $\mathbb{R}^m$. This kind of summation has been studied in the following cases:

1. **Scalar case:** $x_i \in \{-1, 1\}$ and the norm is just the absolute value, i.e., $|\sum_{i=1}^n r_i x_i|$, the scalar version Chernoff bound [20] shows that

$$\Pr_{r \sim \{-1,1\}^n} \left[\left| \sum_{i=1}^n r_i x_i \right| > t \right] \leq 2 \exp\left(-t^2/(2n)\right),$$

corresponding to the case when $h(x) = x$ for $x \in \mathbb{R}$ and the hyperbolic direction $e = 1$.

2. **Matrix case:** x_i are d-by-d symmetric matrices and the norm is the spectral norm, i.e., $\|\sum_{i=1}^n r_i x_i\|$, the matrix Chernoff bound [86] shows that

$$\Pr_{r \sim \{-1,1\}^n} \left[\left\| \sum_{i=1}^n r_i x_i \right\| > t \right] \leq 2d \cdot \exp\left(-\frac{t^2}{2 \|\sum_{i=1}^n x_i^2\|} \right),$$

corresponding to $h(x) = \det(X)$ and $e = I$.

We try to generalize these results to the hyperbolic spectral norm for any hyperbolic polynomial h, which is recognized as an interesting problem in this field by James Renegar [74].

1.1 Our results

In this paper, we can prove the following "Chernoff-type" concentration for hyperbolic spectral norm. We show that, when adding uniformly random signs to n vectors, the hyperbolic spectral norm of their summation will concentrate with an exponential tail.

▶ **Theorem 1** (Nearly optimal hyperbolic Chernoff bound for Rademacher sum)**.** *Let h be an m-variate, degree-d hyperbolic polynomial with respect to a direction $e \in \mathbb{R}^m$. Let $1 \leq s \leq d$, $\sigma > 0$. Given $x_1, x_2, \cdots, x_n \in \mathbb{R}^m$ such that* $\mathrm{rank}(x_i) \leq s$ *for all $i \in [n]$ and $\sum_{i=1}^n \|x_i\|_h^2 \leq \sigma^2$, where* $\mathrm{rank}(x)$ *is the number of nonzero hyperbolic eigenvalues of x. Then, we have*

$$\mathop{\mathbb{E}}_{r \sim \{\pm 1\}^n} \left[\left\| \sum_{i=1}^n r_i x_i \right\|_h \right] \leq 2\sqrt{\log(s)} \cdot \sigma.$$

Furthermore, for every $t > 0$, and for some fixed constant $c > 0$,

$$\Pr_{r \sim \{\pm 1\}^n} \left[\left\| \sum_{i=1}^n r_i x_i \right\|_h > t \right] \leq 2 \exp\left(-\frac{ct^2}{\sigma^2 \log(s+1)} \right).$$

We discuss the optimality of Theorem 1 in different cases:

- **Degree-1 case:** When the hyperbolic polynomial's degree $d = s = 1$, the hyperbolic polynomial is $h(z) = z$. Then, we have $\|x\|_h = |x|$ and we get the the Hoeffding's inequality [37]:

$$\Pr_{r \sim \{\pm 1\}^n} \left[\left| \sum_{i=1}^n r_i x_i \right| > t \right] \leq \exp\Big(-\Omega\Big(t^2 / (\sum_{i=1}^n x_i^2)\Big)\Big).$$

  It implies that our result is optimal in this case.

- **A special degree-2 case:** $h(z) = z_1^2 - z_2^2 - \cdots - z_m^2$. Let $v_1, \ldots, v_n$ be any $(d-1)$-dimensional vectors. Then, we define $x_i := \begin{bmatrix} 0 & v_i \end{bmatrix} \in \mathbb{R}^d$ for $i \in [n]$. We know that $\|x_i\|_h = \|v_i\|_2$, and Theorem 1 gives the following result:

$$\Pr_{r \sim \{\pm 1\}^n} \left[\left\| \sum_{i=1}^n r_i v_i \right\|_2 > t \right] \leq \exp(-\Omega(t^2/\sigma^2)),$$

  where $\sigma^2 := \sum_{i=1}^n \|v_i\|^2$, which recovers the dimension-free vector-valued Bernstein inequality [63].

- **Constant degree case:** When $d > 1$ is a constant, consider h being the determinant polynomial of d-by-d matrix. Since $s \le d = O(1)$, we can show that $\sigma = (\sum_{i=1}^n \|x_i\|^2)^{1/2} = \Theta(\|\sum_{i=1}^n x_i^2\|^{1/2})$, and Theorem 1 exactly recovers the matrix Chernoff bound [86], which implies that our result is also optimal in this case.
- **Constant rank case:** When all the vectors have constant hyperbolic rank, we still take $h = \det(X)$, but $X_1, \dots, X_n$ are constant rank matrices with arbitrary dimension. In this case, we can obtain a dimension-free matrix concentration inequality:

$$\Pr_{r\sim\{\pm1\}^n}\left[\left\|\sum_{i=1}^n r_i X_i\right\| > t\right] \le 2\exp\left(-\Omega(t^2/\sigma^2)\right).$$

  It will beat the general matrix Chernoff bound [86] when σ is not essentially larger than $\|\sum_{i=1}^n X_i^2\|^{1/2}$. Thus, Theorem 1 is nearly optimal in this case. However, Theorem 1 is also sub-optimal in this case if we consider the high degree polynomial $h(z) = \prod_{i=1}^n z_i$, and $x_i = e_i \in \mathbb{R}^n$. Then, we have $\|x_i\|_h = 1$, and $\|\sum_{i=1}^n r_i x_i\|_h = 1$ for any $r \in \{\pm1\}^n$. Therefore, the probability density function of the hyperbolic spectral norm of the Rademacher sum is a delta function[i] in this case. But our concentration result cannot characterize such a sharp transition.

Theorem 1 works for arbitrary vectors in $\mathbb{R}^m$. We also consider the maximum and minimum hyperbolic eigenvalues of the sum of random vectors in the hyperbolic cone, which is a generalization of the positive semi-definite (PSD) cone for matrices. Recall that for independent random PSD matrices $\mathbf{X}_1, \dots, \mathbf{X}_n$ with spectral norm at most R, let $\mu_{\max} := \lambda_{\max}(\sum_i \mathbb{E}[\mathbf{X}_i])$. Then, matrix Chernoff bound for PSD matrices [86] shows that $\Pr[\lambda_{\max}(\sum_i \mathbf{X}_i) \ge (1+\delta)\mu_{\max}] \le de^{-\Omega(\delta\mu_{\max})}$ for any $\delta \ge 0$. The following theorem gives a hyperbolic version of this result:

▶ **Theorem 2** (Hyperbolic Chernoff bound for random vectors in hyperbolic cone)**.** *Let h be an m-variate, degree-d hyperbolic polynomial with hyperbolic direction $e \in \mathbb{R}^m$. Let Λ_+ denote the hyperbolic cone*[ii] *of h with respect to e. Suppose $\mathsf{x}_1, \dots, \mathsf{x}_n$ are n independent random vectors with supports in Λ_+ such that $\lambda_{\max}(\mathsf{x}_i) \le R$ for all $i \in [n]$. Define the mean of minimum and maximum eigenvalues as $\mu_{\min} := \sum_{i=1}^n \mathbb{E}[\lambda_{\min}(\mathsf{x}_i)]$ and $\mu_{\max} := \sum_{i=1}^n \mathbb{E}[\lambda_{\max}(\mathsf{x}_i)]$.*

Then, we have

$$\Pr\left[\lambda_{\max}\left(\sum_{i=1}^n \mathsf{x}_i\right) \ge (1+\delta)\mu_{\max}\right] \le d\cdot\left(\frac{(1+\delta)^{1+\delta}}{e^{\delta}}\right)^{-\mu_{\max}/R} \quad \forall\delta \ge 0,$$
$$\Pr\left[\lambda_{\min}\left(\sum_{i=1}^n \mathsf{x}_i\right) \le (1-\delta)\mu_{\min}\right] \le d\cdot\left(\frac{(1-\delta)^{1-\delta}}{e^{-\delta}}\right)^{-\mu_{\min}/R} \quad \forall\delta \in [0,1].$$

1.2 Hyperbolic anti-concentration

Anti-concentration is an interesting phenomenon in probability theory, which studies the opposite perspective of concentration inequalities. A simple example is the standard Gaussian random variable, which has probability at most $O(\Delta)$ for being in any interval of length Δ.

---

[i] The delta function is defined as $\delta(x) = \begin{cases} 1 & \text{if } x = 1, \\ 0 & \text{otherwise.} \end{cases}$

[ii] The hyperbolic cone is a set containing all vectors with non-negative hyperbolic eigenvalues. See Definition 13 for the formal definition.

For Rademacher random variables $x \sim \{\pm 1\}^d$, the celebrated Littlewood-Offord theorem [53] states that for any degree-1 polynomial $p(x) = \sum_{i=1}^d a_i x_i$ with $|a_i| \geq 1$, the probability of $p(x)$ in any length-1 interval is at most $O(\frac{\log d}{\sqrt{d}})$. Later, the theorem was improved to $O(\frac{1}{\sqrt{d}})$ by Erdös [26], and generalized to higher degree polynomials by [22, 79, 61]. From a geometric prospective, the Littlewood-Offord theorem says that the maximum fraction of hypercube points that lay in the boundary of a halfspace $\mathbf{1}_{\langle a,x\rangle \leq \theta}$ with $|a_i| \geq 1$ for $i \in [d]$ is at most $O(\frac{1}{\sqrt{d}})$. [67] extended this result from half-space to polytope and [11] further extended to positive spectrahedron.

Following this line of research, we prove the following hyperbolic anti-concentration theorem, which shows that the hyperbolic spectral norm of Rademacher sum of vectors in the hyperbolic cone cannot concentrate within a small interval.

▶ **Theorem 3** (Hyperbolic anti-concentration theorem, informal). *Let h be an m-variate degree-d hyperbolic polynomial with hyperbolic direction $e \in \mathbb{R}^m$. Let $\{x_i\}_{i\in[n]} \subset \Lambda_+$ be a sequence of vectors in the hyperbolic cone such that $\lambda_{\max}(x_i) \leq \tau$ for all $i \in [n]$ and $\sum_{i=1}^n \lambda_{\min}(x_i)^2 \geq 1$.*

Then, for any $y \in \mathbb{R}^m$ and any $\Delta \geq 20\tau \log d$, we have

$$\Pr_{\epsilon \sim \{-1,1\}^n}\left[\lambda_{\max}\left(\sum_{i=1}^n \epsilon_i x_i - y\right) \in [-\Delta, \Delta]\right] \leq O(\Delta).$$

From the geometric viewpoint, we can define a "positive hyperbolic-spectrahedron" as the space $\{\alpha \in \mathbb{R}^n : \lambda_{\max}(\alpha_1 x_1 + \cdots + \alpha_n x_n - y) \leq 0\}$, where $x_1, \ldots, x_n$ are in the hyperbolic cone. Then, Theorem 3 states that the hyperbolic spectral norm of a positive hyperbolic-spectrahedron cannot be concentrated in a small region.

1.3 Hyperbolic discrepancy theory

Hyperbolic polynomial is an important tool in the discrepancy theory, which is an important subfield of combinatorics, with many applications in theoretical computer science. Following Meka's blog post [60], by combining scalar version Chernoff bound and union bound, we can easily prove that, for any n vectors $x_1, \ldots, x_n \in \{-1,1\}^n$, there exists $r \in \{-1,1\}^n$ such that $|\langle r, x_i\rangle| \leq O(\sqrt{n \log n})$ for every $i \in [n]$. In a celebrated result "Six Standard Deviations Suffice", Spencer showed that it can be improved to $|\langle r, x_i\rangle| \leq 6\sqrt{n}$ [83].

For the matrix case, by the matrix Chernoff bound, it follows that for any symmetric matrix $X_1, \ldots, X_n \in \mathbb{R}^{d\times d}$ with $\|X_i\| \leq 1$, for uniformly random signs $r \in \{-1,1\}^n$, with high probability, $\|\sum_{i=1}^n r_i X_i\| \leq O(\sqrt{\log(d) n})$.

An important open question is, can we shave the $\log(d)$ factor for *some* choice of the signs?

▶ **Conjecture 4** (Matrix Spencer Conjecture). *For any symmetric matrices $X_1, \ldots, X_n \in \mathbb{R}^{d\times d}$ with $\|X_i\| \leq 1$, there exist signs $r \in \{-1,1\}^n$ such that $\|\sum_{i=1}^n r_i X_i\| = O(\sqrt{n})$.*

The breakthrough paper by Marcus, Spielman and Srivastava [56] proved the famous Kadison-Singer conjecture [41], which was open for more than half of a century.

▶ **Theorem 5** (Kadison-Singer, [41, 56]). *Let $k \geq 2$ be an integer and ϵ a positive real number. Let $x_1, \ldots, x_n \in \mathbb{C}^m$ such that $\|x_i x_i^*\| \leq \epsilon \quad \forall i \in [n]$, and $\sum_{i=1}^n x_i x_i^* = I$. Then, there exists a partition $S_1 \cup S_2 \cup \cdots \cup S_k = [n]$ such that $\|\sum_{i\in S_j} x_i x_i^*\| \leq (\frac{1}{\sqrt{k}} + \sqrt{\epsilon})^2 \quad \forall j \in [k]$.*

The Kadison-Singer theorem implies that for rank-1 matrices $X_1, \dots, X_n$ with $\|X_i\| \le \epsilon$ in isotropic position[iii], there exists a choice of $r \in \{-1,1\}^n$ such that $\|\sum_{i=1}^n r_i X_i\| \le O(\sqrt{\epsilon})$.[iv]

Theorem 5 can be generalized for higher rank matrices by Cohen [21] and Brändén [18] independently. However, their results still need the isotropic condition. On the other hand, Kyng, Luh, and Song [44] proved a stronger version of rank-1 matrix Spencer theorem (Conjecture 4) by showing that when the spectral norm of the sum of the squared matrices (the variance of the random matrices) is bounded, the matrix discrepancy upper bound is at most four deviations. Formal theorem statements will be presented in the full version of this paper [82].

Similar to the scalar and matrix cases, the discrepancy theory can be further generalized to the hyperbolic spectral norm. Brändén [18] proved a hyperbolic Kadison-Singer theorem, which generalizes Theorem 5 to the hyperbolic spectral norm and vectors with arbitrary rank and in isotropic condition. Our first result relaxes the isotropic condition to sub-isotropic:

▶ **Theorem 6** (Hyperbolic Kadison-Singer with sub-isotropic condition, informal). *Let $k \ge 2$ be an integer and $\epsilon, \sigma > 0$. Suppose h is hyperbolic with respect to $e \in \mathbb{R}^m$, and let $x_1, \dots, x_n$ be n vectors in the hyperbolic cone such that*

$$\mathrm{tr}_h[x_i] \le \epsilon \quad \forall i \in [n], \text{ and } \quad \Big\| \sum_{i=1}^n x_i \Big\|_h \le \sigma. \tag{2}$$

where $\mathrm{tr}_h[x] := \sum_{i=1}^d \lambda_i(x)$. Then, there exists a partition $S_1 \cup S_2 \cup \dots \cup S_k = [n]$ such that for all $j \in [k]$,

$$\Big\| \sum_{i \in S_j} x_i \Big\|_h \le \left(\sqrt{\epsilon} + \sqrt{\sigma/k}\right)^2.$$

Theorem 6 implies the high rank case of [56] result (Theorem 5) without the isotropic condition. We note that there is a naive approach to relax the isotropic condition in [56, 18]'s results by adding several small dummy vectors to make the whole set in isotropic position. (See [30] for more details.) However, Theorem 6 is slightly better than this approach, since the naive approach will increase the number of vectors which results in a worse bound.

Theorem 6 also implies the following hyperbolic discrepancy result:

▶ **Corollary 7** (Hyperbolic discrepancy for sub-isotropic vectors). *Let $0 < \epsilon \le \frac{1}{2}$. Suppose $h \in \mathbb{R}[z_1, \dots, z_m]$ is hyperbolic with respect to $e \in \mathbb{R}^m$, and let $x_1, \dots, x_n \in \Lambda_+(h, e)$ that satisfy Eq. (2). Then, there exist signs $r \in \{-1,1\}^n$ such that*

$$\Big\| \sum_{i=1}^n r_i x_i \Big\|_h \le 2\sqrt{\epsilon(2\sigma - \epsilon)}.$$

We note that this result is incomparable with [44] due to the following reasons: 1) [44] only works for rank-1 matrices while our result holds for arbitrary rank vectors in the hyperbolic cone; 2) the upper bound of [44] depends on $\|\sum_{i=1}^n X_i^2\|^{1/2}$ while our result depends on the hyperbolic trace and spectral norm of the sum of vectors.

To obtain a hyperbolic discrepancy upper bound for arbitrary vectors (as in the case of Conjecture 4), we can apply hyperbolic Chernoff bound (Theorem 1) and get the following discrepancy result which holds with high probability:

[iii] Isotropic means $X_1 + \dots + X_n = I$.

[iv] For more details and consequences of the Kadison-Singer theorem, we refer the readers to [19, 58].

▶ **Corollary 8.** *Let h be a degree-d hyperbolic polynomial with respect to $e \in \mathbb{R}^m$. We are given vectors $x_1, x_2, \cdots, x_n \in \mathbb{R}^m$ such that $\|x_i\|_h \le 1$ and $\mathrm{rank}(x_i) \le s$ for all $i \in [n]$ and some $s \in \mathbb{N}_+$. Then for uniformly random signs $r \sim \{-1,1\}^n$,*

$$\left\| \sum_{i=1}^n r_i x_i \right\|_h \le O(\sqrt{n \log(s+1)})$$

holds with probability at least 0.99*.*

This result may not be tight when the ranks of the input vectors are large. It is thus interesting to study whether we can do better to improve the $\sqrt{\log d}$ factor in the non-constructive case. We thus conjecture the following hyperbolic discrepancy bound:

▶ **Conjecture 9** (Hyperbolic Spencer Conjecture)**.** *We are given vectors $x_1, x_2, \cdots, x_n \in \mathbb{R}^m$ and a degree d hyperbolic polynomial $h \in \mathbb{R}[z_1, \ldots, z_m]$ with respect to $e \in \mathbb{R}^m$, where $\|x_i\|_h \le 1$ for all $i \in [n]$. Then, there exist signs $r \in \{-1,1\}^n$, such that*

$$\left\| \sum_{i=1}^n r_i x_i \right\|_h \le O(\sqrt{n}).$$

Note that Conjecture 9 is more general than the matrix Spencer conjecture (Conjecture 4). And for constant degree d or constant maximum rank s, this conjecture is true by Corollary 8.

1.4 Related work

Chernoff-type bounds

There is a long line of work generalizing the classical scalar Chernoff-type bounds to the matrix Chernoff-type bound [77, 5, 78, 85, 55, 29, 45, 66, 10, 40]. [77, 78] showed a Chernoff-type concentration of spectral norm of matrices which are the outer product of two random vectors. [5] first used Laplace transform and Golden-Thompson inequality [31, 84] to prove a Chernoff bound for general random matrices. It was improved by [85] and [68] independently. [55] proved a series of matrix concentration results via Stein's method of exchangeable pairs. Our work further extends this line of research from matrix to hyperbolic polynomials and can fully recover the result of [5]. On the other hand, [29] showed an expander matrix Chernoff bound. [45] prove a new matrix Chernoff bound for Strongly Rayleigh distributions.

Hyperbolic polynomials

The concept of hyperbolic polynomials was originally studied in the field of partial differential equations [27, 39, 42]. Güler [32] first studied the hyperbolic optimization (hyperbolic programming), which is a generalization of LP and SDP. Later, a few algorithms [71, 64, 75, 72, 65, 73] were designed for hyperbolic programming. On the other hand, a lot of recent research focused on the equivalence between hyperbolic programming and SDP, which is closely related to the "Generalized Lax Conjecture" and its variants [36, 52, 17, 43, 80, 7, 69]. In addition to the hyperbolic programming, hyperbolic polynomial is a key component in resolving Kadison-Singer problem [56, 18] and constructing bipartite Ramanujan graphs [57]. Gurvits [34, 35] proved some Van der Waerden/Schrijver-Valiant like conjectures for hyperbolic polynomials, giving sharp bounds for the capacity of polynomials. [81] gave an approach to certify the non-negativity of polynomials via hyperbolic programming, generalizing the Sum-of-Squares method.

Discrepancy theory

For discrepancy theory, we give a few literature in Section 1.3 and we provide more related work here. For Kadison-Singer problem, after the breakthrough result [56], Anari and Oveis Gharan [8] generalized it for Strongly Rayleigh distributions. Alishahi and Barzegar [6] extended the "paving conjecture" to real stable polynomials[v]. Zhang and Zhang [89] further relaxed the determinant polynomial in [8] and [44] to homogeneous real-stable polynomials. More recently, [38, 23] proved some special cases of the matrix Spencer conjecture. For algorithmic results, Bansal [12] proposed the first constructive version of partial coloring for discrepancy minimization. Based on this work, more approaches [54, 76, 51, 25, 13, 24] were discovered in recent years. For applications of the discrepancy theory, [8, 9] used the Strongly Rayleigh version of Kadison-Singer theorem to improve the integrality gap of the Asymmetric Traveling Salesman Problem. [47] used the rank-1 matrix Spencer theorem in [44] to obtain a two-sided spectral rounding result. For more applications, we refer to the excellent book by Matousek [59].

1.5 Technique overview

In this section, we provide a proof overview of our results. We first show how prove hyperbolic Chernoff bounds by upper bounding each polynomial moment. After that, we show how to apply our new concentration inequality to prove hyperbolic anti-concentration. Finally, we show how to relax the isotropic condition in [18], and also how to get a more general discrepancy result via hyperbolic concentration.

1.5.1 Our technique for hyperbolic Chernoff bound for Rademacher sum

The main idea of our proof of hyperbolic Chernoff bound is to upper bound the polynomial moments.

By definition, the hyperbolic spectral norm of X is the ℓ_∞ norm of the eigenvalues $\lambda(X)$. Inspired by the proof of the matrix Chernoff bound by Tropp [87], we can consider the ℓ_{2q} norm of $\lambda(X)$, for $q \geq 1$. When the hyperbolic polynomial h is the determinant polynomial, this norm is just the Schatten-$2q$ norm of matrices. For general hyperbolic polynomials, we define hyperbolic-$2q$ norm as $\|x\|_{h,2q} := \|\lambda(x)\|_{2q}$. By the result of [14], hyperbolic-$2q$ norm is actually a norm in $\mathbb{R}^m$. And the following inequality shows the connection between a hyperbolic spectral norm and hyperbolic-$2q$ norm:

$$\mathop{\mathbb{E}}_{r\sim\{\pm 1\}^n}[\|X\|_h] \leq \Big(\mathop{\mathbb{E}}_{r\sim\{\pm 1\}^n} \big[\|X\|_{h,2q}^{2q}\big]\Big)^{1/(2q)}.$$

In order to compute $\|X\|_{h,2q}^{2q} = \sum_{i=1}^{\mathrm{rank}(X)} \lambda_i(X)^{2q}$, we use a deep result about hyperbolic polynomials: the Helton-Vinnikov Theorem [36], which proved a famous conjecture by Lax [48], to translate between hyperbolic polynomials and matrices. The theorem is stated as follows.

▶ **Theorem 10** ([36]). *Let $f \in \mathbb{R}[x, y, z]$ be hyperbolic with respect to $e = (e_1, e_2, e_3) \in \mathbb{R}^3$. Then there exist symmetric real matrices $A, B, C \in \mathbb{R}^{d\times d}$ such that $f = \det(xA + yB + zC)$ and $e_1 A + e_2 B + e_3 C \succ 0$.*

[v] A polynomial is real stable if it is hyperbolic with respect to every $e \in \mathbb{R}^n_{\gtrsim} 0$.

Gurvits [33] proved a corollary (Corollary 22) that for any m-variate hyperbolic polynomial h, and $x, y \in \mathbb{R}^m$, there exist two symmetric matrices $A, B \in \mathbb{R}^{d\times d}$ such that for any $a, b \in \mathbb{R}$, $\lambda(ax+by) = \lambda(aA+bB)$, where the left-hand side means the hyperbolic eigenvalues of the vector $ax+by$ and the right-hand side means the eigenvalues of the matrix $aA+bB$.

Therefore, we try to separate and consider one random variable r_i at a time. We first consider the expectation over r_1. By conditional expectation, let $X_2 := \sum_{i=2}^n r_i x_i$ and we have

$$\mathop{\mathbb{E}}_{r\sim\{\pm1\}^n}\left[\|X\|_{h,2q}^{2q}\right] = \mathop{\mathbb{E}}_{r_2,\ldots,r_n\sim\{\pm1\}}\left[\mathop{\mathbb{E}}_{r_1\sim\{\pm1\}}\left[\|r_1x_1+X_2\|_{h,2q}^{2q}\right]\right],$$

By Corollary 22, there exist two matrices A_1, B_1 such that $\lambda(r_1x_1+X_2) = \lambda(r_1A_1+B_1)$ holds for any r_1. And it follows that

$$\mathop{\mathbb{E}}_{r_1\sim\{\pm1\}}\left[\|r_1x_1+X_2\|_{h,2q}^{2q}\right] = \mathop{\mathbb{E}}_{r_1\sim\{\pm1\}}\left[\|r_1A_1+B_1\|_{2q}^{2q}\right].$$

It becomes much easier to compute the expected Schatten-$2q$ norm of matrices. We can prove that

$$\mathop{\mathbb{E}}_{r\sim\{\pm1\}^n}\left[\|X\|_{h,2q}^{2q}\right] \le \sum_{k_1=0}^{q}\binom{2q}{2k_1}\|x_1\|_h^{2k_1}\cdot\mathop{\mathbb{E}}_{r_2,\ldots,r_n}\left[\|X_2\|_{h,2q-2k_1}^{2q-2k_1}\right].$$

Now, we can iterate this process for the remaining expectation $\mathbb{E}_{r_2,\ldots,r_n}\left[\|X_2\|_{h,2q-2k_1}^{2q-2k_1}\right]$. After $n-1$ iterations, we get that

$$\left(\mathop{\mathbb{E}}_{r\sim\{\pm1\}^n}\left[\|X\|_{h,2q}^{2q}\right]\right)^{1/(2q)} \le \sqrt{2q-1}\cdot s^{1/(2q)}\cdot\sigma, \tag{3}$$

where $\sigma^2 = \sum_{i=1}^n \|x_i\|_h^2$ and s is the maximum rank of $x_1, \ldots, x_n$. Then, by taking $q := \log(s)$ and $\|X\|_h \le \|X\|_{h,2q}^{2q}$, we get the desired upper bound for the expectation $\mathbb{E}_{r\sim\{\pm1\}^n}[\|\sum_{i=1}^n r_ix_i\|_h]$ in Theorem 1.

To obtain the concentration probability inequality, We can apply the result of Ledoux and Talagrand [49] for the concentration of Rademacher sums in a normed linear space, which will imply:

$$\Pr_{r\sim\{\pm1\}^n}[\|X\|_h > t] \le 2\exp\Big(-t^2\Big/\Big(32\mathop{\mathbb{E}}_{r\sim\{\pm1\}^n}[\|X\|_h^2]\Big)\Big). \tag{4}$$

However, we need to verify that the hyperbolic spectral norm $\|\cdot\|_h$ is indeed a norm, which follows from the result of Gårding [28]. Since by Khinchin-Kahane inequality ([82, Theorem A.16]) the second moment of $\|X\|_h$ can be upper-bounded via the first moment. Hence, we can put our expectation upper bound into Eq. (4) and have

$$\Pr_{r\sim\{\pm1\}^n}[\|X\|_h > t] \le C_1\exp\left(-\frac{C_2t^2}{\sigma^2\log(s+1)}\right),$$

for constants $C_1, C_2 > 0$, and hence Theorem 1 is proved. We defer the formal proof in the full version [82, Section B].

1.5.2 Our technique for hyperbolic Chernoff bound for positive vectors

We can use similar techniques in the previous section to prove Theorem 2.

For any random vectors $\mathsf{x}_1, \ldots, \mathsf{x}_n \in \Lambda_+$, we may assume $\|\mathsf{x}_i\|_h \leq 1$. Using the Taylor expansion of the mgf, we can show that:

$$\Pr\left[\lambda_{\max}\left(\sum_{i=1}^n \mathsf{x}_i\right) \geq t\right] \leq \inf_{\theta>0} \ e^{-\theta t} \cdot \sum_{q \geq 0} \frac{\theta^q}{q!} \, \mathbb{E}\left[\left\|\sum_{i=1}^n \mathsf{x}_i\right\|_{h,q}^q\right]. \tag{5}$$

Then, for the q-th moment, we separate x_1 and $\sum_{i=2}^n \mathsf{x}_i$ and have

$$\mathbb{E}_{\geq 1}\left[\left\|\sum_{i=1}^n x_i\right\|_{h,q}^q\right] = \mathbb{E}_{\geq 2}\mathbb{E}_1\left[\operatorname{tr}\left[(A_1 + B_1)^q\right]\right],$$

where A_1 and B_1 are two PSD matrices obtained via Gurvits's result (Corollary 22) such that A_1 depends on x_1 and B_1 depends on $\mathsf{x}_2, \ldots, \mathsf{x}_n$. The next step is different from the case of Rademacher sum, since we cannot drop half of the terms by the distribution of x_1. Instead, we can fully expand the matrix products in the trace and use Horn's inequality to upper bound the eigenvalue products. We have

$$\mathbb{E}_{\geq 2}\mathbb{E}_1\left[\operatorname{tr}\left[(A(x_1) + B)^q\right]\right] \leq \mathbb{E}_1\left[\sum_{k_1=0}^q \binom{q}{k_1} \lambda_{\max}(\mathsf{x}_1)^{k_1} \cdot \mathbb{E}_{\geq 2}\left[\left\|\sum_{i=2}^n \mathsf{x}_i\right\|_{h,q-k_1}^{q-k_1}\right]\right].$$

By repeating this process, we finally have

$$\mathbb{E}\left[\left\|\sum_{i=1}^n \mathsf{x}_i\right\|_{h,q}^q\right] \leq d \cdot \mathbb{E}\left[\left(\sum_{i=1}^n \|\mathsf{x}_i\|_h\right)^q\right].$$

Then, we put the above upper bound into Eq. (5), which gives:

$$\Pr\left[\lambda_{\max}\left(\sum_{i=1}^n \mathsf{x}_i\right) \geq t\right] \leq \inf_{\theta>0} \ e^{-\theta t} \cdot d \cdot \prod_{i=1}^n \mathbb{E}\left[e^{\theta\|\mathsf{x}_i\|_h}\right].$$

Now, we use some similar calculations in the matrix case [85] to prove that

$$\Pr\left[\lambda_{\max}\left(\sum_{i=1}^n \mathsf{x}_i\right) \geq t\right] \leq \inf_{\theta>0} \ d \cdot \exp\left(-\theta t + (e^\theta - 1)\mu_{\max}\right).$$

By taking $\theta := \log(t/\mu_{\max})$ and $t := (1+\delta)\mu_{\max}$, we get that

$$\Pr\left[\lambda_{\max}\left(\sum_{i=1}^n \mathsf{x}_i\right) \geq (1+\delta)\mu_{\max}\right] \leq d \cdot \left(\frac{(1+\delta)^{1+\delta}}{e^\delta}\right)^{-\mu_{\max}} \tag{6}$$

For the minimum eigenvalue case, we can define $\mathsf{x}_i' := e - \mathsf{x}_i$ for $i \in [n]$. Then, by the property of hyperbolic eigenvalues (Fact 19) and the assumption that $\|\mathsf{x}_i\|_h \leq 1$, we know that x_i' are also in the hyperbolic cone and $\lambda_{\max}(\mathsf{x}_i') = 1 - \lambda_{\min}(\mathsf{x}_i')$. Therefore, we can obtain the Chernoff bound for the minimum eigenvalue of x by applying Eq. (6) with x_i'. We defer the formal proof in the full version [82, Section C].

1.5.3 Our technique for hyperbolic anti-concentration

In this part, we will show how to prove the hyperbolic anti-concentration result (Theorem 3) via the hyperbolic Chernoff bound for vectors in the hyperbolic cone (Theorem 2).

In [67], they studied the unate functions on hypercube $\{-1,1\}^n$, which is defined as the function being increasing or decreasing with respect to any one of the coordinates. Then, they showed that the Rademacher measure of a unate function is determined by the expansion of its indicator set in hypercube. In particular, for the maximum hyperbolic eigenvalue, it is easy to see that the indicator function $\left[\lambda_{\max}\left(\sum_{i=1}^n \epsilon_i x_i^j - y_j\right) \in [-\Delta, \Delta]\right]$ is unate when $x_i \in \Lambda_+$. Hence, we can show the anti-concentration inequality by studying the expansion in the hypercube, which by [11], is equivalent to lower-bound the minimum eigenvalue of each vector. However, for the initial input x_i, we only assume that $\sum_{i=1}^n \lambda_{\min}(x_i)^2 \geq 1$, but we need a $\Omega(\frac{1}{\sqrt{\log d}})$ lower bound for each x_i to prove the theorem. To amplify the minimum eigenvalue, we follow the proof in [11] that uses a random hash function to randomly assign the input vectors into some buckets and considers the sum of the vectors in each bucket as the new input. They proved that the "bucketing" will not change the distribution. Then, we can use Theorem 2 to lower bound the minimum hyperbolic eigenvalue of each bucket, which is a sum of independent random vectors in the hyperbolic cone. Hence, we get that

$$\Pr\left[\lambda_{\min}\left(\sum_{i=1}^n z_{i,j} x_i\right) \leq \Omega(\frac{1}{\sqrt{\log d}})\right] \leq \frac{1}{10},$$

which $z_{i,j} \in \{0,1\}$ is a random variable indicating that x_i is hashed to the j-th bucket. Then, by the standard Chernoff bound for negatively associated random variables, we can prove that most of the buckets have large minimum eigenvalues, which concludes the proof of the hyperbolic anti-concentration theorem. We defer the formal proof in the full version [82, Section D].

1.5.4 Our technique for hyperbolic discrepancy

To relax the isotropic condition in [18], we basically follow their proof. The high-level idea is to construct a compatible family of polynomials[vi] such that the probability in the hyperbolic Kadison-Singer problem (Theorem 6) can be upper-bounded by the largest root of the expected polynomial of the family, which can be further upper-bounded by the largest root of the mixed hyperbolic polynomial $h[v_1,\dots,v_n] \in \mathbb{R}[x_1,\dots,x_m,y_1,\dots,y_n]$, defined as $h[v_1,\dots,v_n] := \prod_{i=1}^m (1 - y_i D_{v_i}) h(x)$, where D_{v_i} is the directional derivative with respect to v_i. In particular, we can consider the roots of the linear restriction $h[v_1,\dots,v_n](te+\mathbf{1}) \in \mathbb{R}[t]$. Then, using Gårding's result [28] on hyperbolic cone, we know that the largest root equals the minimum $\rho > 0$ such that the vector $\rho e + \mathbf{1}$ is in the hyperbolic cone Γ_+ of $h[v_1,\dots,v_n]$, which can be upper-bounded via similar techniques in [56, 44] to iteratively add each vector v_i while keeping the sum in the hyperbolic cone. Our key observation is that the proof in [18] essentially proved that

$$\frac{\epsilon\mu e + \left(1 - \frac{1}{n}\right)\delta \sum_{i=1}^n v_i}{1 + \frac{\mu-1}{n}} + \mathbf{1} \in \Gamma_+$$

holds for any vectors $v_i \in \Lambda_+$. Hence, once we assume that $\|\sum_{i=1}^n v_i\|_h \leq \sigma$, then by the convexity of the hyperbolic cone, we get that $\rho \leq \frac{(\epsilon\mu + (1-\frac{1}{n})\delta\sigma)}{1+\frac{\mu-1}{n}}$, which will imply the upper bound in Theorem 6. We defer the formal proof in the full version [82, Section E].

[vi] The compatible family of polynomials is closely related to the interlacing family in [56, 57]. See [82, Definition E.16].

To obtain the discrepancy result for arbitrary vectors (Corollary 8), we can use the hyperbolic Chernoff bound for Rademacher sum (Theorem 1) to derive the discrepancy upper bound. For any vectors $x_1, \ldots, x_n$ with maximum rank s, by setting $t = O(\sigma\sqrt{\log s})$ in Theorem 1, we get that $\|\sum_{i=1}^n r_i x_i\|_h \leq O(\sigma\sqrt{\log s})$ holds with high probability for uniformly random signs $r \sim \{\pm 1\}^n$.

1.6 Discussion and Open problems

In this paper, we initiate the study of concentration with respect to the hyperbolic spectral norm, and we generalize several classical concentration and anti-concentration results to the hyperbolic polynomial setting. Our results are closely related to the discrepancy theory and pseudorandomness. We provide some open problems in below.

Tighter hyperbolic Chernoff bound?

Our current result has a worse dependence on the variance σ^2 than the matrix Chernoff bound [86]. Can we match the results when $h = \det(X)$? We note that there is a limitation for using the techniques like Golden-Thompson inequality and Lieb's theorem, which were used in [68, 85] to improve the original matrix Chernoff bound [5], to tighten our result. Because for any symmetric matrix X, we can define a mapping such that $\phi(X)$'s eigenvalues are the p-th power of X's eigenvalues for any $p > 0$, where the mapping is just X^p. However, we cannot find such a mapping for vectors with respect to the hyperbolic eigenvalues. Some new techniques may be required to get a hyperbolic Chernoff bound matching the matrix results.

Resolving the hyperbolic Spencer conjecture?

Inspired by the matrix Spencer conjecture (due to Meka [60]), we came up with a more general conjecture for hyperbolic discrepancy. Can we prove or disprove this conjecture? It is also interesting to study the connection between hyperbolic Spencer conjecture and the generalized Lax conjecture [36, 52, 17, 43, 80, 7, 69]. If we assume the matrix Spencer conjecture and the generalized Lax conjecture, can we prove the hyperbolic Spencer conjecture? On the other hand, in a very recent work by Reis and Rothvoss [70], they conjectured a weaker matrix Spencer by considering the Schatten-p norm of matrices. We can also define such an ℓ_p version of the hyperbolic Spencer conjecture by looking at the ℓ_p-norm of hyperbolic eigenvalues (the hyperbolic-p norm). Any progress towards the ℓ_p-hyperbolic Spencer conjecture will provide more insights in matrix and hyperbolic discrepancy theory.

Fooling hyperbolic cone?

One of the results in this paper is showing an anti-concentration inequality with respect to the hyperbolic spectral norm, which generalizes the results in [67, 11]. They actually combined the anti-concentration results with the Meka-Zuckerman [62] framework to construct PRGs fooling polytopes/positive spectrahedrons. Hence, an open question in complexity theory and pseudorandomness is: can we apply the hyperbolic anti-concentration inequality to construct a PRG fooling positive hyperbolic-spectrahedrons, or even hyperbolic cones?

Concentration of random tensors?

Tensor concentration is another natural generalization of matrix concentration. Although there have been a large number of works on this problem [46, 50, 4, 3, 88, 2], it is still unclear what is the optimal concentration bound for the Euclidean norm of random tensor $X \in \mathbb{R}^{n^d}$, even in the simple case when $X = x_1 \otimes \cdots \otimes x_d$ for random vectors $x_1, \ldots, x_d \in \mathbb{R}^n$. On the other hand, people also care about whether random tensors are well-conditioned, which is more related to TCS problems including tensor decompositions and learning Gaussian mixtures. The current results [88, 1, 16] have a large gap between the matrix case. For these tensor concentration problems, is it possible to study them via hyperbolic polynomials and obtain tighter bounds?

References

1. Emmanuel Abbe, Amir Shpilka, and Avi Wigderson. Reed–muller codes for random erasures and errors. *IEEE Transactions on Information Theory*, 61(10):5229–5252, 2015.
2. Radosław Adamczak, Rafał Latała, and Rafał Meller. Moments of gaussian chaoses in banach spaces. *Electronic Journal of Probability*, 26:1–36, 2021.
3. Radosław Adamczak and Paweł Wolff. Concentration inequalities for non-lipschitz functions with bounded derivatives of higher order. *Probability Theory and Related Fields*, 162(3):531–586, 2015.
4. Radosław Adamczak and Rafał Latała. Tail and moment estimates for chaoses generated by symmetric random variables with logarithmically concave tails. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 48(4):1103–1136, 2012. doi:10.1214/11-AIHP441.
5. Rudolf Ahlswede and Andreas Winter. Strong converse for identification via quantum channels. *IEEE Transactions on Information Theory*, 48(3):569–579, 2002.
6. Kasra Alishahi and Milad Barzegar. Paving property for real stable polynomials and strongly rayleigh processes. *arXiv preprint*, 2020. arXiv:2006.13923.
7. Nima Amini. Spectrahedrality of hyperbolicity cones of multivariate matching polynomials. *Journal of Algebraic Combinatorics*, 50(2):165–190, 2019.
8. Nima Anari and Shayan Oveis Gharan. The kadison-singer problem for strongly rayleigh measures and applications to asymmetric tsp. *arXiv preprint*, 2014. arXiv:1412.1143.
9. Nima Anari and Shayan Oveis Gharan. Effective-resistance-reducing flows, spectrally thin trees, and asymmetric tsp. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 20–39. IEEE, 2015.
10. Richard Aoun, Marwa Banna, and Pierre Youssef. Matrix Poincaré inequalities and concentration, 2020. In *Advances in Mathematics*, volume 371. arXiv:1910.13797.
11. Srinivasan Arunachalam and Penghui Yao. Positive spectrahedrons: Geometric properties, invariance principles and pseudorandom generators. *arXiv preprint*, 2021. arXiv:2101.08141.
12. Nikhil Bansal. Constructive algorithms for discrepancy minimization. In *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, pages 3–10. IEEE, 2010.
13. Nikhil Bansal, Daniel Dadush, Shashwat Garg, and Shachar Lovett. The gram-schmidt walk: a cure for the banaszczyk blues. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 587–597, 2018.
14. Heinz H Bauschke, Osman Güler, Adrian S Lewis, and Hristo S Sendov. Hyperbolic polynomials and convex analysis. *Canadian Journal of Mathematics*, 53(3):470–488, 2001.
15. Sergei Bernstein. On a modification of chebyshev's inequality and of the error formula of laplace. *Ann. Sci. Inst. Sav. Ukraine, Sect. Math*, 1(4):38–49, 1924.
16. Aditya Bhaskara, Moses Charikar, Ankur Moitra, and Aravindan Vijayaraghavan. Smoothed analysis of tensor decompositions. In *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*, pages 594–603, 2014.

17 Petter Brändén. Hyperbolicity cones of elementary symmetric polynomials are spectrahedral. *Optimization Letters*, 8(5):1773–1782, 2014.

18 Petter Brändén. Hyperbolic polynomials and the Kadison-Singer problem. *arXiv preprint*, 2018. arXiv:1809.03255.

19 Peter G Casazza and Janet C Tremain. Consequences of the Marcus/Spielman/Srivastava solution of the Kadison-Singer problem. In *New Trends in Applied Harmonic Analysis*, pages 191–213. Springer, 2016.

20 Herman Chernoff. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *The Annals of Mathematical Statistics*, pages 493–507, 1952.

21 Michael Cohen. Improved spectral sparsification and Kadison-Singer for sums of higher-rank matrices. In *Banff International Research Station for Mathematical Innovation and Discovery*. https://open.library.ubc.ca/cIRcle/collections/48630/items/1.0340957, 2016.

22 Kevin P Costello, Terence Tao, and Van Vu. Random symmetric matrices are almost surely nonsingular. *Duke Mathematical Journal*, 135(2):395–413, 2006.

23 Daniel Dadush, Haotian Jiang, and Victor Reis. A new framework for matrix discrepancy: Partial coloring bounds via mirror descent. *arXiv preprint*, 2021. arXiv:2111.03171.

24 Daniel Dadush, Aleksandar Nikolov, Kunal Talwar, and Nicole Tomczak-Jaegermann. Balancing vectors in any norm. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1–10. IEEE, 2018.

25 Ronen Eldan and Mohit Singh. Efficient algorithms for discrepancy minimization in convex sets. *Random Structures & Algorithms*, 53(2):289–307, 2018.

26 Paul Erdös. On a lemma of littlewood and offord. *Bulletin of the American Mathematical Society*, 51(12):898–902, 1945.

27 Lars Gårding. Linear hyperbolic partial differential equations with constant coefficients. *Acta Mathematica*, 85:1–62, 1951.

28 Lars Gårding. An inequality for hyperbolic polynomials. *Journal of Mathematics and Mechanics*, pages 957–965, 1959.

29 Ankit Garg, Yin-Tat Lee, Zhao Song, and Nikhil Srivastava. A matrix expander chernoff bound, 2018. In *STOC*. arXiv:1704.03864.

30 Shayan Oveis Gharan. Proof of kadison-singer conjecture and the extensions, 2015.

31 Sidney Golden. Lower bounds for the helmholtz function. *Physical Review*, 137(4B):B1127, 1965.

32 Osman Güler. Hyperbolic polynomials and interior point methods for convex programming. *Mathematics of Operations Research*, 22(2):350–377, 1997.

33 Leonid Gurvits. Combinatorics hidden in hyperbolic polynomials and related topics. *arXiv preprint*, 2004. arXiv:math/0402088.

34 Leonid Gurvits. Hyperbolic polynomials approach to van der waerden/schrijver-valiant like conjectures: sharper bounds, simpler proofs and algorithmic applications. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 417–426, 2006.

35 Leonid Gurvits. Van der waerden/schrijver-valiant like conjectures and stable (aka hyperbolic) homogeneous polynomials: one theorem for all. *arXiv preprint*, 2007. arXiv:0711.3496.

36 J William Helton and Victor Vinnikov. Linear matrix inequality representation of sets. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 60(5):654–674, 2007.

37 Wassily Hoeffding. Probability inequalities for sums of bounded random variables. In *The collected works of Wassily Hoeffding*, pages 409–426. Springer, 1994.

38 Samuel B Hopkins, Prasad Raghavendra, and Abhishek Shetty. Matrix discrepancy from quantum communication. *arXiv preprint*, 2021. arXiv:2110.10099.

39 L Hormander. The analysis of linear partial differential operators ii. *Grundlehren*, 257, 1983.

40 He Jia, Aditi Laddha, Yin Tat Lee, and Santosh S Vempala. Reducing isotropy and volume to KLS: An $O^*(n^3\psi^2)$ volume algorithm. *arXiv preprint*, 2020. arXiv:2008.02146.

41 Richard V Kadison and Isadore M Singer. Extensions of pure states. *American journal of mathematics*, 81(2):383–400, 1959.

42 N.V. Krylov. On the general notion of fully nonlinear second-order elliptic equations. *Transactions of the American Mathematical Society*, 347(3):857–895, 1995.

43 Mario Kummer, Daniel Plaumann, and Cynthia Vinzant. Hyperbolic polynomials, interlacers, and sums of squares. *Mathematical Programming*, 153(1):223–245, 2015.

44 Rasmus Kyng, Kyle Luh, and Zhao Song. Four deviations suffice for rank 1 matrices, 2020. In *Advances in Mathematics*. `arXiv:1901.06731`.

45 Rasmus Kyng and Zhao Song. A matrix chernoff bound for strongly rayleigh distributions and spectral sparsifiers from a few random spanning trees, 2018. In *FOCS*. `arXiv:1810.08345`.

46 Rafał Latała. Estimates of moments and tails of gaussian chaoses. *The Annals of Probability*, 34(6):2315–2331, 2006.

47 Lap Chi Lau and Hong Zhou. A spectral approach to network design. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, pages 826–839, 2020.

48 Peter D Lax. Differential equations, difference equations and matrix theory. Technical report, New York Univ., New York. Atomic Energy Commission Computing and Applied, 1957.

49 Michel Ledoux and Michel Talagrand. *Probability in Banach Spaces: isoperimetry and processes*. Springer Science & Business Media, 2013.

50 Joseph Lehec. Moments of the gaussian chaos. In *Séminaire de Probabilités XLIII*, pages 327–340. Springer, 2011.

51 Avi Levy, Harishchandra Ramadas, and Thomas Rothvoss. Deterministic discrepancy minimization via the multiplicative weight update method. In *International Conference on Integer Programming and Combinatorial Optimization*, pages 380–391. Springer, 2017.

52 Adrian Lewis, Pablo Parrilo, and Motakuri Ramana. The lax conjecture is true. *Proceedings of the American Mathematical Society*, 133(9):2495–2499, 2005.

53 John Edensor Littlewood and Albert Cyril Offord. On the number of real roots of a random algebraic equation (iii). *Rec. Math. [Mat. Sbornik] N.S.*, 12(3):277–286, 1943.

54 Shachar Lovett and Raghu Meka. Constructive discrepancy minimization by walking on the edges. *SIAM Journal on Computing*, 44(5):1573–1582, 2015.

55 Lester Mackey, Michael I Jordan, Richard Y Chen, Brendan Farrell, and Joel A Tropp. Matrix concentration inequalities via the method of exchangeable pairs. *The Annals of Probability*, 42(3):906–945, 2014.

56 Adam W Marcus, Daniel A Spielman, and Nikhil Srivastava. Interlacing families II: Mixed characteristic polynomials and the Kadison-Singer problem, 2015. `arXiv:1306.3969`.

57 Adam W Marcus, Daniel A Spielman, and Nikhil Srivastava. Interlacing families IV: Bipartite ramanujan graphs of all sizes. *SIAM Journal on Computing*, 47(6):2488–2509, 2018.

58 Adam W Marcus and Nikhil Srivastava. The solution of the Kadison-Singer problem, 2016. In *Current Developments in Mathematics*. `arXiv:1712.08874`.

59 Jiri Matousek. *Geometric discrepancy: An illustrated guide*, volume 18. Springer Science & Business Media, 2009.

60 Raghu Meka. Discrepancy and beating the union bound. In *Windows on theory, a research blog*. `https://windowsontheory.org/2014/02/07/discrepancy-and-beating-the-union-bound/`, 2014.

61 Raghu Meka, Oanh Nguyen, and Van Vu. Anti-concentration for polynomials of independent random variables. In *Theory Of Computing*. arXiv preprint, 2017. `arXiv:1507.00829`.

62 Raghu Meka and David Zuckerman. Pseudorandom generators for polynomial threshold functions. *SIAM Journal on Computing*, 42(3):1275–1301, 2013.

63 Stanislav Minsker. On some extensions of bernstein's inequality for self-adjoint operators. *Statistics & Probability Letters*, 127:111–119, 2017.

64 Tor Myklebust and Levent Tunçel. Interior-point algorithms for convex optimization based on primal-dual metrics. *arXiv preprint*, 2014. `arXiv:1411.2129`.

65 Simone Naldi and Daniel Plaumann. Symbolic computation in hyperbolic programming. *Journal of Algebra and Its Applications*, 17(10):1850192, 2018.

66 Assaf Naor, Shravas Rao, and Oded Regev. Concentration of markov chains with bounded moments. In *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, volume 56, pages 2270–2280. Institut Henri Poincaré, 2020.

67 Ryan O'Donnell, Rocco A Servedio, and Li-Yang Tan. Fooling polytopes. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pages 614–625, 2019.

68 Roberto Imbuzeiro Oliveira. Concentration of the adjacency matrix and of the laplacian in random graphs with independent edges. *arXiv preprint*, 2009. `arXiv:0911.0600`.

69 Prasad Raghavendra, Nick Ryder, Nikhil Srivastava, and Benjamin Weitz. Exponential lower bounds on spectrahedral representations of hyperbolicity cones. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2322–2332. SIAM, 2019.

70 Victor Reis and Thomas Rothvoss. Vector balancing in lebesgue spaces. *arXiv preprint*, 2020. `arXiv:2007.05634`.

71 James Renegar. Hyperbolic programs, and their derivative relaxations. *Foundations of Computational Mathematics*, 6(1):59–79, 2006.

72 James Renegar. "Efficient" subgradient methods for general convex optimization. *SIAM Journal on Optimization*, 26(4):2649–2676, 2016.

73 James Renegar. Accelerated first-order methods for hyperbolic programming. *Mathematical Programming*, 173(1-2):1–35, 2019.

74 James Renegar. Personal communication, 2019.

75 James Renegar and Mutiara Sondjaja. A polynomial-time affine-scaling method for semidefinite and hyperbolic programming. *arXiv preprint*, 2014. `arXiv:1410.6734`.

76 Thomas Rothvoss. Constructive discrepancy minimization for convex sets. *SIAM Journal on Computing*, 46(1):224–234, 2017.

77 Mark Rudelson. Random vectors in the isotropic position. *Journal of Functional Analysis*, 164(1):60–72, 1999.

78 Mark Rudelson and Roman Vershynin. Sampling from large matrices: An approach through geometric functional analysis. *Journal of the ACM (JACM)*, 54(4), 2007.

79 Mark Rudelson and Roman Vershynin. Hanson-wright inequality and sub-gaussian concentration. *Electronic Communications in Probability*, 18, 2013.

80 James Saunderson. A spectrahedral representation of the first derivative relaxation of the positive semidefinite cone. *Optimization Letters*, 12(7):1475–1486, 2018.

81 James Saunderson. Certifying polynomial nonnegativity via hyperbolic optimization. *SIAM Journal on Applied Algebra and Geometry*, 3(4):661–690, 2019.

82 Zhao Song and Ruizhe Zhang. Hyperbolic concentration, anti-concentration, and discrepancy, 2020. `arXiv:2008.09593`.

83 Joel Spencer. Six standard deviations suffice. *Transactions of the American mathematical society*, 289(2):679–706, 1985.

84 Colin J Thompson. Inequality with applications in statistical mechanics. *Journal of Mathematical Physics*, 6(11):1812–1813, 1965.

85 Joel A Tropp. User-friendly tail bounds for sums of random matrices. *Foundations of computational mathematics*, 12(4):389–434, 2012.

86 Joel A Tropp. An introduction to matrix concentration inequalities. *Foundations and Trends in Machine Learning*, 8(1-2):1–230, 2015.

87 Joel A Tropp. Second-order matrix concentration inequalities. *Applied and Computational Harmonic Analysis*, 44(3):700–736, 2018.

88 Roman Vershynin. Concentration inequalities for random tensors. *Bernoulli*, 26(4):3139–3162, 2020.

89 Ruizhe Zhang and Xinzhi Zhang. A real stable generalization of Anari, Oveis Gharan and Kyng, Luh, Song. *manuscript*, 2021.

A Basics of Hyperbolic Polynomial

A.1 Basic definitions of hyperbolic polynomials

We provide the definition of hyperbolic polynomial.

▶ **Definition 11** (Hyperbolic polynomial). *A homogeneous polynomial $h : \mathbb{R} \to \mathbb{R}$ is hyperbolic with respect to a vector $e \in \mathbb{R}^m$ if $h(e) \neq 0$, and for all $x \in \mathbb{R}^m$, the univariate polynomial $t \mapsto h(te - x)$ has only real zeros.*

The following fact shows how to factorize a hyperbolic polynomial, which easily follows from the homogeneity of the polynomial:

▶ **Fact 12** (Hyperbolic polynomial factorization). *For a degree-d polynomial $h \in \mathbb{R}[z_1, \ldots, z_m]$ hyperbolic with respect to $e \in \mathbb{R}^m$, we have*

$$h(te - x) = h(e) \prod_{i=1}^{d} (t - \lambda_i(x))$$

where $\lambda_1(x) \geq \lambda_2(x) \geq \cdots \geq \lambda_d(x)$ are real roots of $h(te - x)$.

All the vectors with nonnegative hyperbolic eigenvalues form a cone, which is proved by Gårding [28]. It is a very important object related to the geometry of hyperbolic polynomials. The formal definition is as follows:

▶ **Definition 13** (Hyperbolic cone). *For a degree d hyperbolic polynomial h with respect to $e \in \mathbb{R}^m$, its hyperbolic cone is*

$$\Lambda_+(e) := \{x : \lambda_d(x) \geq 0\}.$$

The interior of Λ_+^m is

$$\Lambda_{++}(e) := \{x : \lambda_d(x) > 0\}.$$

Gårding [28] showed the following fundamental properties of the hyperbolic cone:

▶ **Theorem 14** ([28]). *Suppose $h \in \mathbb{R}[z_1, \ldots, z_m]$ is hyperbolic with respect to $e \in \mathbb{R}^n$. Then,*

1. *$\Lambda_+(e), \Lambda_{++}(e)$ are convex cones.*
2. *$\Lambda_+ + (e)$ is the connected component of $\{x \in \mathbb{R}^m : h(x) \neq 0\}$ which contains e.*
3. *$\lambda_{\min} : \mathbb{R}^m \to \mathbb{R}$ is a concave function, and $\lambda_{\max} : \mathbb{R}^m \to \mathbb{R}$ is convex.*
4. *If $e' \in \Lambda_{++}(e)$, then h is also hyperbolic with respect to e' and $\Lambda_{++}(e') = \Lambda_{++}(e)$.*

For simplicity, we may use Λ_+ and Λ_{++} to denote $\Lambda_+(e), \Lambda_{++}(e)$, when e is clear from context. In this paper, we always assume that e is any fixed vector in the hyperbolic cone of h.

We define the trace, rank and spectral norm respect to hyperbolic polynomial h.

▶ **Definition 15** (Hyperbolic trace, rank, spectral norm). *Let h be a degree d hyperbolic polynomial with respect to $e \in \mathbb{R}^m$. For any $x \in \mathbb{R}^m$,*

$$\mathrm{tr}_h[x] := \sum_{i=1}^{d} \lambda_i(x), \quad \mathrm{rank}(x) := \#\{i : \lambda_i(x) \neq 0\}, \quad \|x\|_h := \max_{i \in [d]} |\lambda_i(x)| = \max\{\lambda_1(x), -\lambda_d(x)\}.$$

We define the p norm with respect to hyperbolic polynomial h.

▶ **Definition 16** ($\|\cdot\|_{h,p}$ norm)**.** *For any $p \geq 1$, we define the hyperbolic p-norm $\|\cdot\|_{h,p}$ defined as:*

$$\|x\|_{h,p} := \|\lambda(x)\|_p = \Big(\sum_{i=1}^{d} |\lambda_i(x)|^p\Big)^{1/p} \quad \forall x \in \mathbb{R}^m.$$

It has been shown that $\|\cdot\|_h$ and $\|\cdot\|_{h,p}$ are indeed norms:

▶ **Theorem 17** ([28, 18, 73])**.** *$\|\cdot\|_h$ is a semi-norm.*
Furthermore, if Λ_+ is regular, i.e., $(\Lambda_+ \cap -\Lambda_+) = \{0\}$, then $\|\cdot\|_h$ is a norm on $\mathbb{R}^m$.

▶ **Theorem 18** ([14])**.** *For any $p \geq 1$, $\|\cdot\|_{h,p}$ is a semi-norm. Moreover, if the hyperbolic cone Λ_+ is regular, then $\|\cdot\|_{h,p}$ is a norm.*

A.2 Basic properties of hyperbolic polynomials

We state a fact for the eigenvalues $\lambda(\cdot)$ of degree-d hyperbolic polynomial h.

▶ **Fact 19** ([14])**.** *For all $i \in [d]$,*

$$\lambda_i(s \cdot x + t \cdot e) = \begin{cases} s \cdot \lambda_i(x) + t, & \text{if } s \geq 0; \\ s \cdot \lambda_{d-i}(x) + t, & \text{if } s < 0. \end{cases}$$

Then, we show that the elementary symmetric sum-products of eigenvalues can be computed from the directional derivatives of the polynomial.

▶ **Observation 20** ([14])**.** *For a degree-d hyperbolic polynomial h with respect to e, we have*

$$h(te + x) = p(e) \cdot \prod_{i=1}^{d} (t + \lambda_i(x)) = \sum_{i=0}^{d} s_i(\lambda(x)) \cdot t^{d-i},$$

where $\lambda(x) = (\lambda_1(x), \cdots, \lambda_d(x))$ are the hyperbolic eigenvalues of x and $s_i : \mathbb{R}^d \to \mathbb{R}$ is the i-th elementary symmetric polynomial:

$$s_i(y) := \begin{cases} \sum_{S \in \binom{[d]}{i}} \prod_{j \in S} y_j, & \forall i \in [d]; \\ 1 & \textit{if } i = 0. \end{cases}$$

Furthermore, for each $i \in \{0, 1, \cdots, d\}$,

$$h(e) \cdot s_i(\lambda(x)) = \frac{1}{(d-i)!} \cdot \nabla^{d-i} h(x) \underbrace{[e, e, \ldots, e]}_{(d-i) \text{ terms}}.$$

If $i \in [d]$, then $s_i \circ \lambda$ is hyperbolic with respect to e of degree i.

▶ **Corollary 21.** $\mathrm{tr}[x]$ *is a linear function.*

Proof. By Observation 20, we have

$$\mathrm{tr}_h[x] = s_1(\lambda(x)) = \frac{1}{h(e) \cdot (d-1)!} \cdot \nabla^{d-1} h(x)[e, e, \ldots, e].$$

Since h is of degree d, $\nabla^{d-1} h$ is a degree-1 polynomial. Hence, $\mathrm{tr}_h[x]$ is a linear function. ◀

A.3 Helton-Vinnikov Theorem

We state a corollary of Helton-Vinnikov Theorem (Theorem 10), proved by Gurvits [33]:

▶ **Corollary 22** (Proposition 1.2 in [33]). *Let $h(x)$ be a m-variable degree-d hyperbolic polynomial. Then, for $x, y \in \mathbb{R}^m$, there exists two symmetric real matrices $A, B \in \mathbb{R}^{d\times d}$ such that for any $a, b \in \mathbb{R}$, the ordered eigenvalues $\lambda(ax + by) = \lambda(aA + bB)$.*

This Corollary relates the hyperbolic eigenvalues of a vector $ax + by$ to the eigenvalues of matrix $aA + bB$, which allows us to study some properties of hyperbolic eigenvalues using results in matrix theory.