Kadison–Singer conjecture for strongly Rayleigh measures

Benjamin Matschke

1. Introduction

Marcus, Spielman and Srivastava [5] proved the following theorem. It implies the long-standing Kadison–Singer conjecture [4, 1, 6], which asserts that every pure state on the abelian von Neumann algebra $D(\ell_2)$ of bounded diagonal operators on ℓ_2 has a unique extension to a pure state on $B(\ell_2)$.

Theorem 1 (MSS)**.** *Let $V_1, \dots, V_k$ be independent random vectors in $\mathbb{R}^d$, each of which take only finitely many values, and let $\varepsilon > 0$ be such that $\sum \mathbb{E}[V_i V_i^t] = \mathrm{id}_d$ and $\mathbb{E}\big[||V_i||^2\big] \leq \varepsilon$ for all $i = 1, \dots, k$. Then*

$$\mathbb{P}\big[||\sum V_i V_i^t|| \leq (1+\sqrt{\varepsilon})^2\big] > 0.$$

Anari and Oveis Gharan [2] proved a version of the MSS theorem (see Section 3) in which in some sense they managed to weaken the independence assumption for the random vectors $V_1, \dots, V_k$. This allowed them to apply this technique to the Asymmetric Traveling Salesman Problem. In particular they proved a new upper bound for the integrality gap of its natural LP-relaxation.

2. Strongly Rayleigh measures.

Borcea, Brändén and Liggett [3] recently introduced the notion of strongly Rayleigh measures. Let P_n denote the set of all probability measures on $2^{[n]}$. For such a measure $\mu \in P_n$, let $g_\mu := \sum_{S \subseteq [n]} \mu(S) x^S \in \mathbb{R}[x_1, \dots, x_n]$ denote its generating function. We say that $\mu \in P_n$ is homogeneous of degree d if and only if g_μ is. Homogeneous $\mu \in P_n$ of degree 1 are the same as probability measures on $[n]$. There is a product map $P_{n_1} \times P_{n_2} \to P_{n_1+n_2}$, the product of μ_1 and μ_2 being given via $g_{\mu_1 \times \mu_2}(x_1, \dots, x_{n_1+n_2}) := g_{\mu_1}(x_1, \dots, x_{n_1}) \cdot g_{\mu_2}(x_{n_1+1}, \dots, x_{n_1+n_2})$. We call $\mu \in P_n$ strongly Rayleigh if g_μ is a real stable polynomial, i.e. when g_μ has no complex roots $(x_1, \dots, x_n)$ with $\mathrm{im}(x_i) > 0$ for all i. A basic example of homogeneous strongly Rayleigh measures are homogeneous $\mu \in P_n$ of degree 1, and products of such measures.

3. MSS theorem for strongly Rayleigh measures.

Anari and Oveis Gharan [2] proved the following version of the MSS theorem.

Theorem 2 (AO)**.** *Let μ be a homogeneous strongly Rayleigh probability measure on $2^{[m]}$ that satisfies $\mathbb{P}_{S\sim\mu}[i \in S] \leq \varepsilon_1$ for all $i = 1, \dots, m$. Let $v_1, \dots, v_m \in \mathbb{R}^d$ such that $\sum v_i v_i^t = \mathrm{id}_d$ and $||v_i||^2 \leq \varepsilon_2$ for all i. Then*

$$\mathbb{P}_{S\sim\mu}\big[||\sum_{i\in S} v_i v_i^t|| \leq 4(\varepsilon_1+\varepsilon_2) + 2(\varepsilon_1+\varepsilon_2)^2\big] > 0.$$

Theorems 1 and 2 are related as follows. The random vectors $V_1, \ldots, V_k$ from Theorem 1 have finite supports and can thus be considered as homogeneous measures on $2^{[n_i]}$, respectively. Let $\mu \in P_m$ be their product measure according to the previous section, with $m = \sum n_i$. Thus μ is supported on (some of) the k-subsets of the multiset $\{v_1, \ldots, v_m\} := [n_1] \dot\cup \ldots \dot\cup [n_k]$. With this correspondance we see that Theorem 2 holds for more general probability measures, but in turn it needs a bound on each $||v_i||$ and not only on some expected norms, and the assertion is also not exactly the analog of the one in Theorem 1.

4. Motivation: Asymmetric Travelling Salesman Problem.

Let $G = (V, E)$ be a directed graph on n vertices with cost function $c : E \to \mathbb{R}_{\geq 0}$. The Asymmetric Travelling Salesman Problem (ATSP) askes for the shortest tour in G that visits each vertex *at least* once. (Equivalently one can write "exactly once" instead of "at least once" if one further requires the triangle inequality for c.) If c is symmetric, $c(u, v) = c(v, u)$, then this is called the Symmetric TSP, for which it is considerably easier to find approximate solutions. On the other hand, the associated decision problems for both ATSP and STSP are NP-complete.

The ATSP has a natural LP relaxation (by Held and Karp '70). The integrality gap is defined as the quotient between the costs of the optimal tours for the LP relaxation and for the original ATSP. It is known that this gap can be at least 2. It is unknown whether it is bounded from above by a constant. The prevous best upper bound was $O(\log(n)/\log\log(n))$, and Anari and Oveis Gharan were able to improve it to $O((\log\log(n))^a)$ for some a. Their approach was via so-called α-spectrally thin trees, which are defined as follows.

Let L_G denote the discrete Laplace operator on G, now regarded as an undirected graph. A matrix representation of L_G is $L_G = \sum_{e \in E} b_e b_e^t \in \mathbb{R}^{n \times n}$, where b_e is the vector $\mathbb{1}_u - \mathbb{1}_v \in \mathbb{R}^n$. Similarly for a spanning tree $T \subseteq G$, define $L_T = \sum_{e \in T} b_e b_e^t \in \mathbb{R}^{n \times n}$. Now, a spanning tree $T \subseteq G$ is called α-spectrally thin, $\alpha \in \mathbb{R}_{>0}$, if $L_T \preceq \alpha L_G$.

A sufficient condition for T being α-spectrally thin is $|| \sum_{e \in T} v_e v_e^t || \leq \alpha$, where $v_e := L_G^{\dagger/2} \cdot b_e$, $L_G^{\dagger/2}$ denoting the square root of the pseudo inverse of L_G. This is precisely a condition that can be obtained from Theorem 2. For this one needs further an adequate probability distribution on the set of spanning trees of G. In [3] it was proved that for any $\gamma : E \to \mathbb{R}$ the measure μ supported on the spanning trees of G and given via $P_\mu(T) \sim \prod_{e \in T} \exp(\gamma(e))$ is a homogeneous and strongly Rayleigh measure in $P_{|E|}$.

5. Mixed characteristic polynomials

Let $\mu \in P_m$ be a homogeneous probability distribution on $2^{[m]}$ of degree d_μ. For m given vectors $v_1, \ldots, v_m \in \mathbb{R}^d$, the mixed characteristic polynomial of μ at $v_1, \ldots, v_m$ is defined as

$$\mu[v_1, \ldots, v_m](x) = \mathbb{E}_{S \sim \mu} \chi \big[\sum_{i \in S} 2 v_i v_i^t \big](x^2) \in \mathbb{R}[x].$$

where $\chi[M]$ denotes the ordinary characteristic polynomial of a square matrix M.

Theorem 3 ([2]). *$\mu[v_1, \ldots, v_m](x)$ equals*

$$x^{d-d_\mu} \cdot \big(\prod(1-\partial_{z_i}^2)\big) \cdot \big(g_\mu(x \cdot \mathbb{1} + z) \cdot \det(x \cdot \mathrm{id}_d + \sum z_i v_i v_i^t)\big)\Big|_{z_1=\ldots=z_m=0} \in \mathbb{R}[x].$$

Here, $z_1, \ldots, z_m$ are m further variables. In the formula of the theorem, the differential operators $(1-\partial_{z_i}^2)$ are applied to $g_\mu(\ldots)\det(\ldots)$ *before* the variables z_i are put to zero. Note that both factors $g_\mu(\ldots)$ and $\det(\ldots)$ are are linear in each z_i, whence each operator $\partial_{z_i}^2$ gets "distributed", one ∂_{z_i} for each factor.

This representation of the mixed characteristic polynomial opens the way to apply the theory of stable polynomials. Using the lemmas from [5] the following corollary follows immediately.

Corollary. *If μ is strongly Rayleigh, then $\mu[v_1, \ldots, v_m]$ is real rooted.*

6. Interlacing families

Let $\mathcal{F} := \{S \subseteq [n] \mid \mu(S) \neq 0\}$. Let $\{q_S\}_{S\in\mathcal{F}}$ denote the family of polynomials given by $q_S(x) = \mu(S) \cdot \chi\big[\sum_{i\in S} 2v_i v_i^t\big](x^2)$. The characteristic polynomial at $v_1, \ldots, v_m$ is clearly the sum of the q_S. In fact $\{q_S\}_{\mathcal{F}}$ is a so-called interlacing family (in the sense of [2]; the proof uses the previous corollary), by which one obtains the following theorem.

Theorem 4. *There exists an $S \in \mathcal{F}$ such that the largest root of q_S is less or equal to the largest root of $\mu[v_1, \ldots, v_m](x)$.*

7. Proof scheme for Theorem 2.

By an extension of the so-called multivariate barrier argument of [5], Anari and Oveis Gharan proved that the largest root of $\mu[v_1, \ldots, v_m](x)$ is at most $4(2\varepsilon+\varepsilon^2)$, where $\varepsilon = \varepsilon_1+\varepsilon_2$. We omit this part, as this is given in large detail in the next talk by Romanos Malikiosis. Then one applies Theorem 4 and obtains the existence of some $S \in \mathcal{F}$ such that all roots of q_S are bounded from above. As q_S is essentially the characteristic polynomial of a matrix $\sum_{i\in S} v_i v_i^t$, this bounds the operator norm of that matrix. And this finishes the proof of Theorem 2.

References

[1] C. A. Akemann and J. Anderson, *Lyapunov theorems for operator algebras*, Mem. Amer. Math. Soc. 94, 1991.

[2] N. Anari and S. Oveis Gharan, *The Kadison–Singer Problem for Strongly Rayleigh Measures and Applications to Asymmetric TSP*, arXiv:1412.1143, 2014.

[3] J. Borcea, P. Brändén and T. M. Liggett, *Negative dependence and the geometry of polynomials*, J. Amer. Math. Soc. 22 (2009), 521–567.

[4] R. V. Kadison and I. M. Singer, *Extensions of pure states*, Amer. J. Math. 81(2) (1959), 383–400.

[5] A. Marcus, D. A. Spielman and N. Srivastava, *Interlacing Families II: Mixed Characteristic Polynomials and The Kadison–Singer Problem*, arXiv:1306.3969, 2014.

[6] N. Weaver, *The Kadison–Singer problem in discrepancy theory*, Discrete Math. 278(1-3) (2004), 227–239.